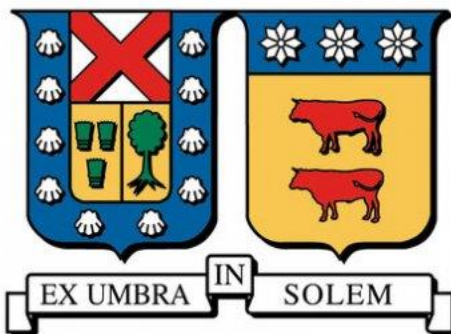




***UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INGENIERÍA COMERCIAL***

***ANÁLISIS DE LA RELACIÓN ENTRE MATERIAL AUDIOVISUAL SOBRE
CRIPTOMONEDAS EN REDES SOCIALES Y LA WEB FRENTE AL MERCADO
DE BITCOIN***

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO COMERCIAL

***AUTOR
SEBASTIÁN CATALÁN VALENCIA***

***PROFESOR GUÍA
ZÓCIMO CAMPOS J.***

***PROFESOR CORREFERENTE
JUAN TAPIA G.***

VALPARAÍSO, 02 DE NOVIEMBRE, 2022.

Índice de contenidos

1. RESUMEN EJECUTIVO	8
2. PROBLEMA DE INVESTIGACIÓN	10
3. OBJETIVOS	16
3.1. OBJETIVO GENERAL.....	16
3.2. OBJETIVOS ESPECÍFICOS	16
4. MARCO TEÓRICO	18
4.1. ANTECEDENTES	18
4.1.1. Criptomonedas	18
4.1.1.1. Industria	19
4.2. BITCOIN	20
4.2.1. Definición	20
4.2.2. Mercado.....	21
4.2.3. Inversionistas.....	22
4.2.4. Evolución del precio en el tiempo	23
4.2.5. Valor intrínseco de la moneda.....	24
4.2.6. Volatilidad	26
4.3. ESTUDIOS ENTRE BITCOIN Y OTRAS VARIABLES.....	28
4.3.1. Relación con variables económicas, financieras y políticas.....	29
4.3.2. Relación con otras monedas.....	31
4.3.3. Relación con otras criptomonedas	32
4.3.4. Relación con medios de comunicación.....	34
4.4. REDES SOCIALES Y LA WEB	35
4.4.1. Definiciones.....	35
4.4.2. Principales plataformas a considerar	36
4.5. ESTUDIOS SOBRE BITCOIN, REDES SOCIALES Y LA WEB	38
4.5.1. Twitter.....	38
4.5.2. Facebook	39
4.5.3. YouTube.....	41
4.5.4. Reddit.....	42
4.5.5. Google	43
4.5.6. Wikipedia.....	45
4.6. CONTENIDO ACERCA DE BITCOIN EN REDES SOCIALES	45
4.6.1. Panorama general	45
4.6.2. Material audiovisual sobre criptomonedas.....	48
4.6.2.1. Películas, series y documentales	48
4.6.2.2. Videos de YouTube	52
4.6.2.3. Transmisiones en vivo por Twitch.....	54
4.6.2.4. Potencial de este material como fuente de datos	57
4.7. DATOS DISPONIBLES PARA LA INVESTIGACIÓN.....	59
4.7.1. Parámetros de Bitcoin.....	60

4.7.1.1.	Parámetros de Bitcoin - Bases de datos existentes	60
4.7.1.2.	Parámetros de Bitcoin - Base de datos a utilizar: Coingecko	65
4.7.2.	<i>Índices de audiencia de material cinematográfico</i>	66
4.7.2.1.	Índices de audiencia – Bases de datos existentes.....	67
4.7.2.2.	Índices de audiencia - Base de datos a utilizar: Google Trends	70
4.7.3.	<i>Métricas en canales de YouTube sobre criptomonedas</i>	71
4.7.3.1.	Métricas en canales de YouTube - Posibles fuentes de datos.....	71
4.7.3.2.	Métricas en canales de YouTube - Fuente a utilizar: SocialBlade	74
4.7.4.	<i>Métricas de canales de Twitch sobre Bitcoin</i>	75
4.7.4.1.	Métricas de canales en Twitch - Posibles fuentes de datos	75
4.7.4.2.	Métricas de canales en Twitch - Fuente a utilizar: TwitchTracker.....	79
4.8.	BASES METODOLÓGICAS.....	81
4.8.1.	<i>Modelos presentes en la literatura</i>	81
4.8.1.1.	Modelos de regresión.....	81
4.8.1.2.	Modelos para analizar correlación	84
4.8.1.3.	Modelos autoregresivos	85
4.8.2.	<i>Modelos a utilizar</i>	86
4.8.3.	<i>Modelos de regresión</i>	88
4.8.3.1.	Regresión Lineal Simple.....	88
4.8.3.2.	Regresión Lineal Múltiple	90
4.8.3.3.	Significancia estadística de los coeficientes	92
4.8.3.4.	Significancia estadística del modelo regresivo	95
4.8.3.5.	Verificación de supuestos	96
4.8.3.5.1.	Linealidad de los parámetros	97
4.8.3.5.2.	Valor esperado de los residuos igual a cero.....	98
4.8.3.5.3.	Condición de homocedasticidad en los residuos	99
4.8.3.5.4.	No hay autocorrelación entre residuos.....	101
4.8.3.5.5.	Variables independientes no estocásticas	106
4.8.3.5.6.	Residuos con distribución normal.....	107
4.8.3.5.7.	No existe multicolinealidad	109
4.8.3.5.8.	Serie de tiempo estacionarias	112
4.8.4.	<i>Metodologías para analizar Correlación</i>	115
4.8.4.1.	Coeficiente de correlación de Pearson.....	116
4.8.4.2.	Coeficiente de correlación de Spearman.....	119
4.8.4.3.	Consideraciones para ambos coeficientes.....	120
4.8.4.4.	Significancia estadística de los coeficientes	122
5.	METODOLOGÍA	124
5.1.	DATOS	124
5.1.1.	<i>Bitcoin</i>	124
5.1.1.1.	Recolección de datos diarios.....	124
5.1.1.2.	Transformación de datos a formato semanal	127
5.1.2.	<i>Índices de audiencia</i>	129

5.1.2.1.	Recolección de datos	129
5.1.2.1.1.	Selección del material audiovisual sobre criptomonedas	129
5.1.2.1.2.	Recopilación de datos vía Google Trends	130
5.1.3.	<i>Métricas en canales de YouTube</i>	135
5.1.3.1.	Recolección de datos	135
5.1.3.1.1.	Selección de canales sobre Critpomonedas	135
5.1.3.1.2.	Elaboración de ranking	137
5.1.3.1.3.	Obtención de datos en SocialBlade	138
5.1.3.1.4.	Consolidación de una única serie de tiempo	141
5.1.4.	<i>Métricas en canales de Twitch</i>	142
5.1.4.1.	Recolección de datos	142
5.1.4.1.1.	Selección de categorías sobre criptomonedas	146
5.1.4.1.2.	Obtención de datos en TwitchTracker	146
5.1.4.1.3.	Consolidación de una única serie de tiempo	147
5.1.4.2.	Transformación de datos a formato semanal	148
5.2.	ANÁLISIS	149
5.2.1.	<i>Software a Utilizar</i>	149
5.2.2.	<i>Índices de audiencia</i>	149
5.2.3.	<i>YouTube y Twitch</i>	151
6.	RESULTADOS	156
6.1.	ÍNDICES DE AUDIENCIA - ANÁLISIS DE CORRELACIÓN	156
6.1.1.	<i>Verificación de supuestos</i>	156
6.1.2.	<i>Análisis de correlación</i>	159
6.1.2.1.	Niveles de Búsqueda de Google	160
6.1.2.2.	Niveles de búsqueda de YouTube	164
6.1.2.3.	Comparación de ambos casos	168
6.2.	REGRESIÓN LINEAL ENTRE BITCOIN Y YOUTUBE	168
6.2.1.	<i>Verificación de supuestos</i>	168
6.2.2.	<i>Modificaciones en los modelos</i>	179
6.2.3.	<i>Resultados de la regresión</i>	180
6.2.3.1.	Regresión de los retornos	180
6.2.3.2.	Regresión de las variaciones del volumen de intercambio	183
6.3.	REGRESIÓN LINEAL ENTRE BITCOIN Y TWITCH	185
6.3.1.	<i>Verificación de Supuestos</i>	185
6.3.2.	<i>Resultados de la regresión</i>	198
7.	CONCLUSIONES Y RECOMENDACIONES	200
8.	LIMITACIONES	203
8.1.	METODOLOGÍAS SELECCIONADAS	203
8.2.	RECOPILACIÓN DE DATOS	204
9.	REFERENCIAS	206

10. ANEXOS	240
10.1. APÉNDICE 1: LISTA DE PRODUCCIONES AUDIOVISUALES.....	240
10.2. APÉNDICE 2: MUESTRA DE CANALES DE YOUTUBE	241
10.3. APÉNDICE 3: MUESTRA DE CANALES DE TWITCH.....	245
10.3.1. <i>Procedimiento para la recopilación de datos para canales de Twitch:</i>	245
10.3.1.1. Selección de canales sobre criptomonedas	245
10.3.1.2. Elaboración de ranking	249
10.3.1.3. Obtención de datos en TwitchTracker	257
10.3.1.4. Consolidación de una única serie de tiempo	258

ÍNDICE DE FIGURAS

FIGURA 1: RESULTADOS DE GOOGLE TRENDS PARA "BITCOIN".	69
FIGURA 2: EJEMPLO DE MATRIZ DE CORRELACIÓN.....	111
FIGURA 3: GRÁFICOS DE SOCIALBLADE PARA EL CANAL "BITBOY CRYPTO"	139
FIGURA 4: GRÁFICO DE CATEGORÍA "CRYPTO"	147
FIGURA 5: GRÁFICOS DE LA FUNCIÓN DE CORRELACIÓN PARA BÚSQUEDAS EN GOOGLE FRENTA A LOS RETORNOS.....	162
FIGURA 6: GRÁFICOS DE LA FUNCIÓN DE CORRELACIÓN PARA BÚSQUEDAS EN GOOGLE FRENTA A LAS VARIACIONES EN EL VOLUMEN DE INTERCAMBIO.....	163
FIGURA 7: GRÁFICOS DE LA FUNCIÓN DE CORRELACIÓN PARA BÚSQUEDAS EN YOUTUBE FRENTA A LOS RETORNOS.....	166
FIGURA 8: GRÁFICOS DE LA FUNCIÓN DE CORRELACIÓN PARA BÚSQUEDAS EN YOUTUBE FRENTA A LAS VARIACIONES EN EL VOLUMEN DE INTERCAMBIO.....	167
FIGURA 9: GRÁFICOS DE DISPERSIÓN - MODELO CON VARIABLES DE YOUTUBE PARA ESTIMAR RETORNOS	171
FIGURA 10: GRÁFICO DE RESIDUOS EN FUNCIÓN DE VALORES ESTIMADOS - MODELO CON VARIABLES DE YOUTUBE PARA ESTIMAR RETORNOS.....	172
FIGURA 11: GRÁFICOS DE DISPERSIÓN - MODELO CON VARIABLES DE YOUTUBE PARA ESTIMAR DIFERENCIAS DEL VOLUMEN DE INTERCAMBIO	173
FIGURA 12: GRÁFICO DE RESIDUOS EN FUNCIÓN DE VALORES ESTIMADOS - MODELO CON VARIABLES DE YOUTUBE PARA ESTIMAR LAS DIFERENCIAS EN EL VOLUMEN DE INTERCAMBIO	174
FIGURA 13: RESUMEN DEL MODELO LINEAL CON VARIABLES DE YOUTUBE PARA ESTIMAR RETORNOS	180
FIGURA 14: VALORES P CORREGIDOS SEGÚN EL MÉTODO DE NEWEY-WEST – MODELO CON VARIABLES DE YOUTUBE PARA ESTIMAR RETORNOS.....	182
FIGURA 15: RESUMEN DEL MODELO LINEAL CON VARIABLES DE YOUTUBE PARA ESTIMAR DIFERENCIAS EN EL VOLUMEN DE INTERCAMBIO	183
FIGURA 16: GRÁFICOS DE DISPERSIÓN - MODELO CON VARIABLES DE TWITCH PARA ESTIMAR RETORNOS	188

FIGURA 17: GRÁFICO DE DISPERSIÓN - APLICACIÓN DE LOGARITMOS AL NÚMERO DE ESPECTADORES DE TWITCH AL ESTIMAR RETORNOS	189
FIGURA 18: GRÁFICO DE RESIDUOS EN FUNCIÓN DE VALORES ESTIMADOS - MODELO CON VARIABLES DE TWITCH PARA ESTIMAR RETORNOS	189
FIGURA 19: GRÁFICOS DE DISPERSIÓN - MODELO CON VARIABLES DE TWITCH PARA ESTIMAR DIFERENCIAS DEL VOLUMEN DE INTERCAMBIO	190
FIGURA 20: GRÁFICO DE DISPERSIÓN - APLICACIÓN DE LOGARITMOS AL NÚMERO DE ESPECTADORES DE TWITCH AL ESTIMAR DIFERENCIAS EN EL VOLUMEN DE INTERCAMBIO	191
FIGURA 21: GRÁFICO DE RESIDUOS EN FUNCIÓN DE VALORES ESTIMADOS - MODELO CON VARIABLES DE TWITCH PARA ESTIMAR LAS DIFERENCIAS EN EL VOLUMEN DE INTERCAMBIO	191

ÍNDICE DE TABLAS

TABLA 1: VARIABLES DE PRODUCCIONES AUDIOVISUALES	134
TABLA 2: PRUEBAS ADF Y JB PARA VARIABLES DE BITCOIN.	156
TABLA 3: PRUEBAS ADF Y JB PARA BÚSQUEDAS EN GOOGLE.....	157
TABLA 4: PRUEBAS ADF Y JB PARA BÚSQUEDAS EN YOUTUBE	158
TABLA 5: COEFICIENTES DE CORRELACIÓN PARA LOS NIVELES DE BÚSQUEDA EN GOOGLE .	160
TABLA 6: COEFICIENTES DE CORRELACIÓN PARA LOS NIVELES DE BÚSQUEDA EN YOUTUBE	164
TABLA 7: PRUEBAS ADF Y PP PARA EL MODELO CON VARIABLES DE YOUTUBE.....	169
TABLA 8: PRUEBAS ADF Y PP PARA EL MODELO CON VARIABLES DE YOUTUBE LUEGO DE APLICAR LAS TRANSFORMACIONES NECESARIAS	170
TABLA 9: CÁLCULO DEL VALOR ESPERADO DE LOS RESIDUOS PARA AMBOS MODELOS CON VARIABLES DE YOUTUBE	174
TABLA 10: RESULTADOS DE LA PRUEBA DE BREUSCH-PAGAN PARA AMBOS MODELOS	175
TABLA 11: RESULTADOS DE LA PRUEBA DE DURBIN-WATSON PARA AMBOS MODELOS CON VARIABLES DE YOUTUBE	175
TABLA 12: CÁLCULO DE LA COVARIANZA ENTRE LOS RESIDUOS DE CADA MODELO CON VARIABLES DE YOUTUBE Y LAS VARIABLES INDEPENDIENTES DE ESTOS.....	176
TABLA 13: RESULTADOS DE LA PRUEBA DE JARQUE-BERA PARA LOS RESIDUOS DE AMBOS MODELOS CON VARIABLES DE YOUTUBE	177
TABLA 14: RESULTADOS DE LA PRUEBA DE JARQUE-BERA PARA CADA VARIABLE INDEPENDIENTE DE LOS MODELOS SOBRE YOUTUBE	178
TABLA 15: MATRIZ DE CORRELACIÓN PARA LAS VARIABLES INDEPENDIENTES DE LOS MODELOS SOBRE YOUTUBE.....	178
TABLA 16: PRUEBAS ADF Y PP PARA EL MODELO CON VARIABLES DE TWITCH.....	186
TABLA 17: PRUEBAS ADF Y PP PARA EL MODELO CON VARIABLES DE TWITCH LUEGO DE APLICAR LAS TRANSFORMACIONES NECESARIAS	187
TABLA 18: CÁLCULO DEL VALOR ESPERADO DE LOS RESIDUOS PARA AMBOS MODELOS CON VARIABLES DE TWITCH	192

TABLA 19: RESULTADOS DE LA PRUEBA DE BREUSCH-PAGAN PARA AMBOS MODELOS SOBRE TWITCH	192
TABLA 20: RESULTADOS DE LA PRUEBA DE DURBIN-WATSON PARA AMBOS MODELOS CON VARIABLES DE TWITCH	193
TABLA 21: CÁLCULO DE LA COVARIANZA ENTRE LOS RESIDUOS DE CADA MODELO CON VARIABLES DE TWITCH Y LAS VARIABLES INDEPENDIENTES DE ESTOS	194
TABLA 22: RESULTADOS DE LA PRUEBA DE JARQUE-BERA PARA LOS RESIDUOS DE AMBOS MODELOS CON VARIABLES DE TWITCH.....	194
TABLA 23: RESULTADOS DE LA PRUEBA DE JARQUE-BERA PARA CADA VARIABLE INDEPENDIENTE DE LOS MODELOS SOBRE TWITCH	195
TABLA 24: MATRIZ DE CORRELACIÓN PARA LAS VARIABLES INDEPENDIENTES DE LOS MODELOS SOBRE TWITCH.....	196
TABLA 25: RESUMEN DEL MODELO LINEAL CON VARIABLES DE TWITCH PARA ESTIMAR RETORNOS	198
TABLA 26: RESUMEN DEL MODELO LINEAL CON VARIABLES DE TWITCH PARA ESTIMAR DIFERENCIAS EN EL VOLUMEN DE INTERCAMBIO	199

1. RESUMEN EJECUTIVO

Bitcoin, la criptomoneda de mayor antigüedad y aquella que posee la mayor capitalización de mercado entre sus pares, ha sido ampliamente estudiada en cuanto a los factores que podrían jugar un rol importante en su formación de precio. Uno de ellos corresponde a las redes sociales y la web, donde existe evidencia suficiente para decir que la actividad de las personas en estos medios sí genera efectos en el mercado de Bitcoin. Sin embargo, la literatura disponible suele enfocarse mayoritariamente en aquellas formas de comunicación ligadas a texto, excluyendo así las interacciones por medio de contenido audiovisual. Lo cual causa resulta intrigante, dado que hay estudios que demuestran la existencia de vínculos entre lo audiovisual y otros mercados financieros.

Es por esta razón que la presente memoria busca investigar si existe un vínculo entre el material audiovisual sobre criptomonedas y las variables presentes en el mercado de Bitcoin. Para ello primero se examinan y evalúan las fuentes de datos más adecuadas para el estudio, con un enfoque destinado a producciones cinematográficas (películas, series y documentales), videos de YouTube y transmisiones en vivo por Twitch. Luego se opta por revisar y discutir las metodologías disponibles para el análisis, estableciendo el uso de coeficientes de correlación y modelos de regresión lineal múltiple, ya que la literatura sugiere una hipótesis acerca de relaciones con carácter lineal.

Las conclusiones obtenidas señalan que la atención dirigida a las producciones cinematográficas (medida por sus niveles de búsqueda en Google y YouTube) posee una asociación monótona baja frente a Bitcoin, que se mantiene dentro de un mismo rango de intensidad a lo largo de tiempo. Mientras que las capacidades predictivas de los suscriptores de YouTube a canales sobre criptomonedas logran explicar una parte considerable de los

cambios en el mercado de la criptomoneda. Confirmando así la existencia de relaciones lineales o monótonas entre variables audiovisuales y Bitcoin.

En contraste a lo anterior, Twitch no logra alcanzar resultados significativos como variable explicativa. Lo cual conduce a cuestionar la capacidad de los datos para representar la actividad de potenciales inversionistas en la plataforma, o bien establecer que no hay relación alguna entre las estadísticas de este sitio web y Bitcoin.

Finalmente, se propone que futuras investigaciones podrían extender este tipo de estudios a otras plataformas de contenido en formato de video, tales como Instagram o TikTok. Para así expandir la literatura que toma en consideración los formatos de video acerca de criptomonedas.

2. PROBLEMA DE INVESTIGACIÓN

Las criptomonedas pueden definirse como “monedas alternativas que son exclusivamente electrónicas y, por ende, no tienen forma física” (Kristoufek, 2013). Se cree que el surgimiento de este tipo de activos es una consecuencia de la crisis financiera de 2007, donde la credibilidad de las instituciones financieras se vio duramente afectada, motivando así la búsqueda de activos que prescindan de intermediarios (Papathanasiou, 2019). A partir de ese punto, las criptomonedas se han convertido en un tópico muy relevante en las finanzas, al ser un nuevo tipo de activo emergente que se distingue entre los demás, llamando así la atención de diversos grupos de interés. Tal como gobiernos (Zhang y Wang, 2020; Yatsyk, 2020), medios de comunicación e inversionistas (Urquhart, 2018), economistas, consumidores, entre otros (Wolfson, 2015).

La criptodivisa con más antigüedad, Bitcoin, es un sistema de efectivo electrónico de usuario a usuario creado por Satoshi Nakamoto en el año 2009, que tiene como propósito llegar a ser una moneda que garantice la confianza en las transacciones sin necesidad de un intermediario y, por ende, que disminuyan los costos como las comisiones a bancos u otras entidades financieras (Nakamoto, 2008). Actualmente es la criptomoneda con mayor precio y capitalización en el mercado (Coinmarketcap, s.f.), siendo así un punto de referencia constante para la industria. Su funcionamiento se basa en el concepto de Blockchain, el cual es una especie de cadena de transacciones únicas registradas a modo de un libro contable muy extenso (Torras y Castellana, 2021). Dicha tecnología permite que las actividades realizadas con esta moneda cuenten con gran seguridad y confianza, dado que se protegen los datos de cada una de las partes (garantizando el anonimato de ellas). También existe total

transparencia de esos registros para todas las personas, y los usuarios pueden revisar cada una de las operaciones para así permitir que la cadena siga creciendo de manera correcta.

Esta moneda desde sus inicios tiene como propósito ser una divisa ligada a la deflación, dado que existe un tope de 21 millones de Bitcoins que pueden llegar a existir (Torras y Castellana, 2021). El proceso por el que se incorporan más de ellas al mercado se conoce como “Mining” y consiste en descifrar un código único con el que se puede extender la cadena de Blockchain y así generar ganancia para quien haya logrado hacerlo. No obstante, la cantidad monetaria que se obtiene en este proceso va disminuyendo con el paso del tiempo, por lo tanto, cada vez es menos atractivo hacer ese tipo de esfuerzos (Torras y Castellana, 2021). A la fecha de esta memoria hay aproximadamente 19,1 millones de Bitcoins en circulación (Coinmarketcap, s.f.).

Pese a que la moneda busca crear transacciones completamente confiables y su actual valor ha permitido que se establezca como un activo relevante para los inversionistas, se ha puesto en duda su capacidad para alcanzar y mantener esa posición debido a la alta volatilidad en su precio (Yermack, 2015). Frente a lo cual existen distintas posturas, sobre si Bitcoin posee un carácter especulativo (Yermack, 2015), de moneda (Hazlett y Luther, 2020), o ambos en distintas proporciones (Baur, Lee y Hong, 2015).

Es por ello que existen numerosos enfoques de estudio sobre los factores que participan en la formación del precio del Bitcoin, ya que su entendimiento repercute en las posibles utilidades que se le puede dar a la criptomoneda (Hayes, 2017). Más aún cuando diversos académicos resaltan la importancia de este tipo de investigación, al estar enfocadas en una moneda pionera dentro de los activos financieros (Karalevicius, Degrande y Weerdt, 2018), que además goza de gran popularidad y aceptación en el mercado (Hayes, 2017).

Por ejemplo, se ha estudiado los efectos en el precio de Bitcoin frente a variables como la incertidumbre política (Fang et al., 2019), índices bursátiles (Chan et al., 2019), monedas fiduciarias (Adnan Fendi et al., 2019), otras criptomonedas (Yi et al., 2018), medios de comunicación (Polasik, Piotrowska, Wisniewski, Kotkowski y Lightfoot, 2015), redes sociales (Pant et al., 2018), y la web (Kristoufek, 2015).

Entre los distintos factores que influyen en el mercado de Bitcoin cobra especial relevancia la actividad en las redes sociales y el internet. Un medio donde en la actualidad las personas interactúan con aquello que le genera interés y las comunidades presentes alrededor de esos tópicos. Por ende, cuando el tema común es Bitcoin, en esos canales se da origen a una potencial fuente de información capaz de describir las emociones de las personas, su atención y comportamiento mostrados frente a esta criptomoneda.

La relación entre las plataformas de red social y la web frente a Bitcoin ha sido estudiada desde frentes como Twitter (Pant et al., 2018; Jahjah y Rajab, 2020), Facebook (Jerdack, 2018), YouTube (Partida, 2019), Google (Abraham et al., 2018; Zhang y Wang, 2020), entre otras plataformas. Hallando así resultados que confirman la influencia que estos medios tienen sobre el mercado de la criptomoneda pionera. No obstante, la presente memoria ha detectado que la literatura existente sobre Bitcoin se limita únicamente a la información que posee formato de texto en las redes. Excluyendo así todo tipo de ideas expresadas dentro de videos o transmisiones en vivo, traduciéndose en que una enorme cantidad de datos no están siendo considerados al estudiar cómo las personas interesadas en Bitcoin interactúan entre sí.

El vínculo entre materiales audiovisuales y mercados financieros ha sido observado en otro tipo de sectores. Tal como es la investigación realizada por (Lepori, 2015), quien comparte que los finales de temporada en series de televisión son capaces de incidir en las

emociones de las personas, y por consiguiente en la actividad registrada dentro del mercado de valores de Estados Unidos. Al igual que (Zhang et al., 2017) confirma las repercusiones de la audiencia televisiva sobre el mercado de valores chino.

Frente a tales antecedentes en otros sectores financieros, cabe cuestionar si existen efectos similares en el mercado de Bitcoin. Tomando en cuenta la gran cantidad de videos producidos en redes sociales como YouTube o Twitch, que poseen creadores de contenido dedicados a comentar acontecimientos recientes relativos a las criptomonedas y compartir sus experiencias como inversionistas. Siendo este tipo de material capaz de crear comunidades con un interés común en las criptomonedas, fomentando interacciones que pueden traducirse en cambios dentro del mercado de Bitcoin.

Es en esos tipos de conversaciones donde se pueden transmitir ideas que intervengan en las decisiones de aquellos que consumen dichos videos e invierten en Bitcoin. Tomando como referencia la investigación realizada por (Rani y Prerana, 2021), cuyas encuestas realizadas demuestran que el contenido de YouTube es capaz de influenciar las decisiones de inversión de las personas.

A esto se suma la presencia de otro segmento audiovisual, compuesto por películas, series y documentales relacionados con Bitcoin. Las cuales adquieren relevancia al estudiar las redes sociales puesto que a través de ellas las personas suelen adquirir interés por este tipo de producciones cinematográficas (Bhattacharjee et al., 2017), y posteriormente comparten sus evaluaciones o comentarios en el mismo medio (Hsu et al., 2014). Como resultado, este tipo de material audiovisual es otro posible factor que podría relacionarse con el mercado del Bitcoin.

De esa manera, teniendo en cuenta el potencial de los formatos de video para comunicarse en redes sociales, cabe mencionar las consecuencias de no tomar en cuenta este

tipo de información al estudiar las variables que pueden ser relevantes en la formación de precio de Bitcoin.

Por un lado, los formatos de video están presentes en todas partes. Videos de YouTube pueden ser compartidos en plataformas como Twitter y Reddit (que ya han sido estudiados frente a Bitcoin), pero no han sido tomados en cuenta en las investigaciones. Generando así intriga acerca de cuál es el rol que cumplen en las conversaciones sobre Bitcoin y si haberlos considerado hubiese cambiado las conclusiones obtenidas.

Luego no considerar como variable al material audiovisual desaprovecha la oportunidad de incluir un nuevo enfoque al estudio de formación de precio de Bitcoin. Siendo esto equivalente a perder los posibles nuevos conocimientos que se puedan generar al entender cuál es el rol que juegan los formatos de video al ser consumidos por potenciales inversionistas. Sobre todo, cuando se considera que el estudio de las variables que tienen influencia en Bitcoin es un área a la cual le resta mucho por ser explorado (Uras et al., 2020), por lo tanto, diversificar los enfoques de estudio ayuda al desarrollo de la literatura. De esa manera, investigar los efectos de nuevas variables representa un aporte significativo para futuros estudios centrados en Bitcoin, y no hacerlo es perder oportunidades de aprendizaje que podrían influir en el uso que se da a las criptomonedas en la actualidad (como moneda, instrumento de inversión, entre otros.).

En conclusión, la literatura existente sobre redes sociales y Bitcoin posee muy poca presencia de variables basadas en el contenido audiovisual. Razón por la cual investigar esta brecha significa crear entendimiento que revele más información sobre cómo funciona un activo emergente como Bitcoin y el resto de criptomonedas.

Un estudio con ese tipo de objetivos buscaría alcanzar respuestas para preguntas como: ¿Existe correlación entre el consumo de contenido con formato de video y variables

como el precio del Bitcoin? ¿Las variables representativas de material audiovisual son capaces de explicar el comportamiento del mercado? ¿Qué tipos de videos son aquellos que poseen una relación más fuerte con Bitcoin?

Al darles respuesta se estaría generando información valiosa para futuros análisis fundamentales a Bitcoin, sean estos de carácter académico o para ser utilizados por un inversionista. Del mismo modo que se motivaría a indagar más sobre la influencia que ejercen los videos a lo largo de las redes sociales y la web.

3. OBJETIVOS

3.1. Objetivo General

Describir y analizar cómo es que el material audiovisual sobre criptomonedas afecta las principales variables del mercado de Bitcoin. Por medio del estudio de la correlación y capacidad predictiva de los indicadores de audiencia o interés del contenido en formato de video, frente a los datos históricos propios de la criptomoneda. Para así poder determinar el potencial del contenido audiovisual para esta área de estudio, junto a entregar información que apoye a inversionistas y académicos al momento de investigar la formación del precio de Bitcoin.

3.2. Objetivos Específicos

- Consolidar antecedentes acerca de las metodologías y supuestos utilizados previamente en la literatura sobre la relación de Bitcoin frente a otras variables, con el propósito de posicionar el presente estudio dentro de la literatura actual.
- Evaluar las fuentes de información disponibles sobre Bitcoin y los tres tipos de material audiovisual considerados, al estudiar sus distintas características, con el objetivo de seleccionar aquellas más adecuadas para su utilización.
- Seleccionar los modelos que sean aptos para resolver la problemática de investigación, describiendo sus fortalezas y limitaciones, para así definir la metodología a desarrollar.
- Elaborar una base de datos con las distintas variables a analizar, por medio de las fuentes de información seleccionadas, con el fin de contar con series de tiempo en las que aplicar los modelos escogidos.

- Analizar la correlación y capacidad predictiva de las variables provenientes del material audiovisual frente a Bitcoin, a través los modelos seleccionados, con tal de obtener cifras que describan dicha relación.
- Concluir acerca de los efectos del contenido en formato de video sobre Bitcoin, destacando el potencial observado en este tipo de datos dentro del área de estudio, con el fin de dar solución a la problemática de investigación.

4. MARCO TEÓRICO

4.1. Antecedentes

Para dar respuestas a las preguntas de investigación señaladas con anterioridad es prudente elaborar una revisión de literatura detallando qué tipo de criptodivisa es Bitcoin y las características que eso conlleva. Al mismo modo que se explica su posición con respecto a otros tipos de criptomonedas en la misma categoría.

4.1.1. Criptomonedas

Por definición, una criptodivisa es dinero virtual disponible únicamente de manera encriptada (Jumde y Cho, 2020), gracias a tecnología que busca garantizar seguridad en su circulación (blockchain).

El nombre de la palabra “Criptomoneda” hace alusión a la Criptografía, que es un proceso por el que se traduce información legible en un código que difícilmente puede ser transgredido (Milutinovic, 2018). Para luego ser registrado en una cadena de bloques o blockchain, que funciona a modo de un gran libro contable de las transacciones realizadas.

Las criptomonedas inicialmente se posicionaron como dinero electrónico centralizado, que dependía de entidades bancarias para procesar y validar transacciones (Han et al., 2019). Lo cual las hacía bastante similares al dinero fiat de cada país (Dólar estadounidense, Peso chileno, Yen japonés, etc.), que son las monedas aprobadas por cada país para ser oficialmente emitidas e intercambiadas (Jumde y Cho, 2020). No obstante, con el surgimiento de Bitcoin se comenzó a prescindir de estos intermediarios y se optó por ser un medio de pago descentralizado (Han et al., 2019). Es decir, dejó de ser necesario depositar la confianza en una entidad determinada ya que el sistema es confiable por sí mismo.

A la fecha de entrega de esta memoria existen aproximadamente 20.331 tipos de criptomonedas, siendo Bitcoin (BTC) aquella con un mayor precio y capitalización en el

mercado, seguida por Ethereum (ETC) y Tether (USDT). Teniéndose así una capitalización de todo el mercado equivalente a \$1,049 trillones de dólares estadounidenses (Coinmarketcap, s.f.).

Cabe mencionar que, dentro del universo de criptodivisas, la mayoría de las monedas distintas a Bitcoin son conocidas como “Altcoins”, ya que son idénticas a Bitcoin, pero con modificaciones en algunos de sus parámetros. Tales como la cantidad de monedas en circulación y el tiempo de bloqueo. Con la excepción de pocas criptomonedas que han innovado en su funcionamiento (Hileman y Rauchs, 2017).

A pesar de la buena evolución de las criptomonedas en los últimos años, actualmente existen numerosos casos de manipulación deliberada de los precios para beneficio personal y creación de expectativas demasiado altas en nuevas criptomonedas, prometiendo un crecimiento elevado para engañar a inversores (Gandal et al., 2021). Provocando que el riesgo percibido en este mercado sea elevado, afectando así la inversión y uso de las criptodivisas.

La actividad de las criptomonedas ha llevado a que se desarrollen agentes dentro del área de las finanzas u otros sectores de la economía para generar valor en torno a las criptodivisas, es decir, una industria.

4.1.1.1. Industria

Según Hileman y Rauchs (2017), la industria de las criptomonedas está compuesta por agentes que facilitan el uso de este tipo de dinero digital. Estos pueden clasificarse en cuatro sectores:

- Intercambio: Está relacionada con la compra, venta e intercambio de criptomonedas y divisas nacionales. Por lo tanto, se otorga liquidez a las criptomonedas.
- Billeteras: Permiten almacenar de manera segura las criptomonedas.
- Compañías de pago: Comprende empresas que entregan servicios para facilitar el pago con criptomonedas.
- Minería: Se encargan de confirmar y asegurar el registro de transacciones (cadena de bloques o blockchain) (p.9)

Con tales condiciones en el mercado e industria, resulta interesante estudiar la criptomoneda de mayor relevancia, el Bitcoin. Tomando en cuenta su posición número uno entre sus competidores, además de su antigüedad, que dota de una extensa lista de estudios alrededor de él.

4.2. Bitcoin

A continuación, se desarrollará una breve descripción del Bitcoin, al igual que del mercado donde es comercializado y los agentes involucrados en este. Para luego detallar cifras históricas de su evolución y recopilar estudios que aportan a la discusión para explicar el comportamiento de esta criptomoneda.

4.2.1. Definición

Existen numerosas definiciones de Bitcoin, las cuales comparten el hecho de haber sido creada por Satoshi Nakamoto en 2009, ser independiente de una entidad reguladora y estar basada en la tecnología blockchain. No obstante, se opta por destacar a Fosso Wamba et al. (2019), que definen Bitcoin de manera más holística, como “una moneda basada en sistemas computacionales que no posee una contraparte legal física, usada como un medio

de intercambio a través de sistemas computacionales y protocolos de comunicación online” (p.3).

En base a lo anterior, es prudente elaborar más acerca de cómo es que funciona este mercado basado en sistemas computacionales.

4.2.2. Mercado

Bitcoin ocupa un 43,08% del mercado de criptomonedas, posicionándose como aquella con mayor capitalización y precio en el mercado (Coinmarketcap, s.f.) a la fecha de publicación de esta memoria.

Las transacciones realizadas con esta criptodivisa poseen un carácter de igual a igual, sin intermediarios. Por lo tanto, la operación solo requiere el consenso de ambas partes para ser realizada (Fosso Wamba et al., 2019). Estas pueden realizarse en cualquier momento, las veinticuatro horas del día, los siete días de la semana. Y, a modo general, la mayor parte de ellas corresponden a intercambios entre inversionistas que desean especular en vez de compras de bienes o servicios (Dyhrberg et al., 2018).

En cuanto al precio al cual es comercializado el Bitcoin, este depende individualmente de cada transacción que se esté realizando. De hecho, diversas plataformas como Coinmarketcat, Coingecko o Binance se encargan de calcular un valor que represente el precio al que está siendo intercambiada cada criptomoneda en un determinado periodo de tiempo.

Por otro lado, acerca de las características relativas a la eficiencia del mercado, la investigación de Feng et al. (2018) establece que el mercado del Bitcoin es altamente ineficiente, debido a los comportamientos irracionales de los inversionistas y la falta de un

mecanismo que fije el precio de la criptomoneda. Por lo cual no recomiendan invertir en él a menos que se tengan solamente fines especulativos (p.3).

Sin embargo, también existe evidencia de que, si bien el mercado del Bitcoin es ineficiente en la actualidad, presenta mejoras que indican un aumento de la eficiencia con el paso del tiempo (Dyhrberg et al., 2018). Lo cual genera expectativas positivas en el mercado que contrastan las críticas al mercado.

Frente a tales condiciones cabe preguntar cuál es la manera de actuar de los inversionistas y qué los caracteriza.

4.2.3. Inversionistas

Como primera característica de los inversionistas en el mercado del Bitcoin se propone que los inversionistas no suelen ser de carácter institucional, debido a que el mercado no posee un organismo regulador (Dyhrberg et al., 2018). Además, según los mismos autores, la mayor parte de los inversionistas corresponden a especuladores. Es decir, personas que tienen expectativas de un futuro aumento de precio del activo, por lo tanto, invierten en el tiempo presente para así obtener rentabilidad posteriormente (p.1).

La actividad de los inversionistas se ve muy influenciada por la información disponible a cada uno de ellos, ya que al ser un mercado no regulado no existe una entidad que supervise y sancione los intercambios realizados por inversionistas que poseen información que alerte de eventos relevantes para el mercado. Por ende, se ha investigado que normalmente los inversionistas informados son capaces de acomodar su posición en el mercado para obtener mejores rentabilidades frente a fluctuaciones en el precio. Especialmente cuando existe la opción de hacer intercambios las veinticuatro horas del día (Feng et al., 2018).

Además, la actividad de los inversionistas cobra aún más importancia cuando se toma en consideración que, en el mercado de Bitcoin, las personas que son responsables de una gran cantidad de transacciones (compras y ventas) ejercen mayor influencia en el precio en comparación a los efectos provocados por otros mercados financieros (Mizerka et al., 2020). Por lo tanto, al estudiar la formación de precio de esta criptomoneda se debe tener especial atención al comportamiento de las personas que actúan en él.

De esta forma, considerando tanto las características del mercado como las de los inversionistas, se procede a continuación a describir los cambios del precio del Bitcoin en los últimos años.

4.2.4. Evolución del precio en el tiempo

Desde los comienzos de Bitcoin (2009), el precio de la criptomoneda fue en aumento lentamente. En el año 2011 su valor estaba bajo los \$100 USD (Baur y Hoang, 2021), manteniéndose bajo ese valor hasta que en 2013 ocurrió una rápida alza del precio. La cual consistió en pasar desde \$150 hasta más de \$1.000 dólares estadounidenses en tan solo dos meses (Gandal et al., 2021). No obstante, se mantuvo sobre el umbral de los \$1.000 dólares por poco tiempo, ya que en 2014 volvió a estar bajo él. Manteniéndose en esa posición hasta los inicios del año 2017, donde lo volvió a superar.

Es en dicho año, 2017, el precio de Bitcoin alcanzó su valor más alto hasta entonces, alrededor de los \$19.500 USD (Guo et al., 2021). Siendo este producido en un muy corto periodo de tiempo, ya que en agosto de 2017 su valor estaba alrededor de los \$2.700 dólares, y posteriormente se acercó a los \$19.500 en diciembre (Wu et al, 2018). Sin embargo, ese valor también resultó ser un alza efímera, ya que en 2018 su valor cayó por debajo de los

\$10.000 USD, permaneciendo en oscilación alrededor de ese valor hasta su siguiente gran aumento de precio.

Durante los últimos meses del año 2020 el precio de Bitcoin vuelve a subir, provocando que durante el año 2021 que ocurra un nuevo incremento brusco. Donde en los primeros meses del año ya se tenían posturas que indicaban la presencia de un patrón de burbuja especulativa (Shilov y Zubarev, 2021). Dado que el precio había pasado desde aproximadamente \$5.000 USD en marzo de 2020 hasta \$63.564 USD el 13 abril de 2021 (su valor más alto hasta entonces), que después en junio del mismo año disminuyó drásticamente a 31.634 el 21 de junio de 2021 (Shu et al., 2021). Por consiguiente, las expectativas que se tenían sobre la presencia de una burbuja financiera eran ciertas en lo que respecta a un pronto desplome del precio.

A partir de ese punto la criptomoneda logró recuperarse y alcanzar una nueva cifra más alta, llegando esta vez cerca de los \$68.000 USD en noviembre de 2021 (Coinmarketcap – Bitcoin, s.f.). Pero desde entonces su precio ha seguido una tendencia negativa, ubicándose cercanamente a los \$21.500 USD durante julio de 2022. Generando así mucha incertidumbre sobre cuál será su valor en los próximos meses (NextAdvisor, 2022; Streissguth, 2022; Browne, 2022).

4.2.5. Valor intrínseco de la moneda

En el presente subcapítulo se abordará parte de los argumentos que critican a Bitcoin de ser una moneda no válida, al igual que otros en donde se contradice esa idea.

Existen detractores de la idea que Bitcoin sea capaz de ser utilizado como una moneda. Algunos argumentan que su volatilidad hace que no se pueda desempeñar bien como una unidad contable y un contenedor de valor, luego de someter la divisa a una serie de

pruebas relativas al mercado de futuros. Por lo tanto, concluyen que el Bitcoin es únicamente un activo especulativo y no una moneda válida (Corbet et al., 2018).

También se critica que el mercado de esta criptodivisa no es lo suficientemente eficiente para poder desempeñarse como un sistema de intercambio monetario, y que requiere mayores regulaciones para alcanzar dicho estado (Apopo y Phiri, 2021).

Por otro lado, otros autores han propuesto maneras de demostrar el valor intrínseco del Bitcoin, tal como Chico-Frias (2021), que se enfoca en aislar los efectos de la especulación y asocia el valor intrínseco de la moneda a los costos de producción que tiene en su minería, tal como la electricidad utilizada y la remuneración de aquel profesional que asume el cargo de minero (p. 2).

Incluso es posible encontrar posturas que no contradicen ni aceptan el valor completamente el valor intrínseco del Bitcoin, dado que se ve a dichas cifras como irrelevantes, al ser algo que se ve tan opacado por la especulación intensa que existe alrededor de este activo (Hayes, 2015)

Aun cuando la literatura contiene estos distintos puntos de vista sobre el valor intrínseco de Bitcoin y su utilización como moneda, las personas que no se dedican a estas áreas de estudio no ha demostrado percibir la controversia. Tal como fue demostrado por (Papathanasiou, 2019), quienes obtuvieron que la población suele percibir a Bitcoin como un medio para realizar transacciones y pagos.

En síntesis, la volatilidad del Bitcoin es definitivamente un tema frente al que hay mucha discusión, razón por la cual es considerada una de las causas por las que esta investigación busca analizar, desde el punto de vista del material audiovisual, el efecto de las redes sociales y la web sobre esta criptomonedas.

4.2.6. Volatilidad

La volatilidad de Bitcoin es considerada una de las principales variables que representa el riesgo presente en su mercado (Hu, 2019; Gkillas et al., 2021). Siendo esta uno de los obstáculos que dificulta la utilización de esta criptomoneda como moneda de intercambio y la razón por la que se posiciona como instrumento de inversión de alto riesgo (Baur y Dimpfl, 2017; Baur y Dimpfl, 2021).

A lo largo de los últimos años la percepción que se tiene de Bitcoin en cuando a volatilidad ha cambiado en diversas ocasiones. Tal es caso de los años previos a las grandes alzas de precio ocurridas en 2017, donde se creía que la variabilidad del activo estaba en descenso, pero sin descartar la posibilidad de un futuro cambio en esta tendencia, tomando en consideración la inmadurez del mercado (Bouoiyour y Selmi, 2016). Además, se estimaba que, si la tendencia de ese entonces se extendía lo suficiente, Bitcoin alcanzaría la volatilidad del resto de monedas fiduciarias para el año 2019 o 2020 (Cermak, 2017).

Sin embargo, dichas proyecciones no se cumplieron, dado que en los años posteriores a 2017 se han observado en diferentes oportunidades rápidas alzas de precio y decrecimientos abruptos. Generando que la volatilidad de Bitcoin sea en ocasiones catalogada como una criptodivisas con variabilidad 30 veces mayor a aquella que poseen otras monedas fiduciarias como el dólar estadounidense o el yen japonés (Baur y Dimpfl, 2017).

De hecho, durante la evolución de Bitcoin se ha podido apreciar que la volatilidad puede sufrir cambios a medida que se incorporan nuevos elementos al mercado. Tal como ocurrió en el año 2017 con la adición de derivados financieros basados en Bitcoin por Chicago Board Options Exchange (CBOE), donde inicialmente la volatilidad aumentó significativamente, pero con el paso del tiempo logró estabilizarse (Kim et al., 2020).

Dichas proyecciones y diagnósticos sobre el mercado del Bitcoin han ido modificándose a medida que mejora el entendimiento de los distintos factores que influyen en la volatilidad. Es por ello que en la literatura se puede encontrar que los valores de volatilidad pasada son buenos predictores de sus valores futuros, al igual que el volumen de intercambio (Aalborg et al., 2019). Al igual que se ha observado a variables macroeconómicas o financieras teniendo muy poca influencia en la variabilidad de Bitcoin (Guizani y Nafti, 2019).

No obstante, publicaciones recientes proponen que la volatilidad de Bitcoin posee características diferentes dependiendo de cómo se esté estudiando esta variable. Por ejemplo, Bitcoin tiene una muy alta volatilidad en cuanto a su desviación estándar, pero si se está estudiando cuán volátil en el sentido de comportamiento irregular o impredecible, su volatilidad es menor a la de otros mercados financieros como el oro o el índice S&P 500 (Dehouche, 2021). Más aún cuando, según el mismo autor, “la alta volatilidad de Bitcoin es un fenómeno que ocurre dentro de la semana, ya que al estudiarla con periodos semanales se ve fuertemente atenuada”. Por lo tanto, para estudiar la volatilidad de Bitcoin se debe poner especial atención en el tipo de parámetro que la mide y la granularidad de los datos. De esa manera, se opta por entregar una interpretación más apropiada y se evita la simplificación de la volatilidad como algo que posee solo un significado.

Siguiendo la misma línea, hay quienes sostienen que la volatilidad de Bitcoin no es tan alta cuando se tiene en consideración el volumen de intercambio en el mercado (Sapuric et al., 2020). Añadiendo que la volatilidad se muestra con altos valores cuando el volumen de intercambio es bajo. Por consiguiente, otro criterio a tener presente para interpretar la volatilidad es observar qué variables están presentes en el análisis, para concluir sobre ellas en conjunto y no necesariamente de manera individual.

En conclusión, la variabilidad de Bitcoin es conocida por presentar valores muy altos, pero sus antecedentes históricos demuestran que esta no es una característica que haya permanecido constante durante la existencia de la criptomoneda. Más aún cuando hay presencia de un gran interés investigativo para revelar más aspectos sobre la variabilidad, mejorando el entendimiento y las proyecciones en esta materia. Lo cual se materializa en una serie de publicaciones que evalúan la influencia de variables externas sobre el mercado del Bitcoin, tal como ocurre con el resto de variables alrededor de esta criptomoneda.

De esa forma, en la siguiente sección se lleva a cabo una revisión de dichos estudios entre Bitcoin y factores ajenos a su mercado.

4.3. Estudios entre Bitcoin y otras variables

En el presente subcapítulo se revisará la literatura sobre vínculos entre Bitcoin y otros factores externos a su mercado. Con el propósito de poder posicionar dónde se ubican en la conversación científica las publicaciones sobre la relación entre Bitcoin, redes sociales y la web, que luego conducirá a investigar la influencia del material audiovisual acerca de criptomonedas sobre Bitcoin. Para ello se cubrirán los principales hallazgos encontrados, al igual que los focos de discusión existentes.

Las relaciones a considerar son:

- Relación con variables económicas, financieras y políticas
- Relación con otras monedas
- Relación con otras criptomonedas
- Relación con medios de comunicación

4.3.1. Relación con variables económicas, financieras y políticas

Las publicaciones que examinan los vínculos existentes entre este tipo de variables y Bitcoin pueden ser caracterizadas en tres grupos. Primero están aquellas que analizan a la criptomoneda frente a variables de incertidumbre, identificando los efectos existentes y poniendo principal atención a la capacidad de cobertura que presenta Bitcoin. Luego, como segundo lugar, están las interacciones frente a los índices financieros más relevantes. Las cuales se enfocan en ver cómo estas afectan a Bitcoin. Y como tercer punto se investiga cómo las relaciones mencionadas anteriormente se comportan durante determinados periodos con alta incertidumbre como la pandemia de Covid-19.

Comenzando por el primer grupo mencionado, se poseen antecedentes que respaldan a la volatilidad de Bitcoin viéndose afectada significativamente por la incertidumbre política (Fang et al., 2019; Chen et al., 2021). Donde se ha descubierto que Bitcoin puede servir de cobertura cuando tanto los retornos de Bitcoin como la incertidumbre (medida por el índice de volatilidad VIX) presentan valores extremos, sean estos altos o bajos (Bouri et al., 2017; Demir et al., 2018). Además de poder ser utilizado como “safe heaven” en momentos donde la incertidumbre es muy alta, tal como se suele utilizar al oro, ya que sus retornos y volatilidad aumentan durante esos periodos (Paule-Vianez et al., 2020). Este tipo de resultados también se repiten ante otras medidas de incertidumbre, como el “Chinese Economic Policy Uncertainty (CEPU)” (Chen et al., 2021).

De manera similar, esta criptomoneda ha sido analizada frente a otras variables de incertidumbre, como es el riesgo geopolítico. El cual se define como “el riesgo asociado con guerras, terrorismo y tensiones entre diferentes estados que puedan afectar tener relaciones internacionales pacíficas” (Lacoviello et al., 2018). Ante ello se ha propuesto que Bitcoin se ve influenciado positivamente con el aumento del riesgo geopolítico, además de ser un activo

que puede ser utilizado para evadir los efectos de este riesgo (Su et al., 2020). Mientras que el riesgo geopolítico por sí mismo posee una gran capacidad predictiva frente a los retornos y volatilidad de Bitcoin (Aysan et al., 2019).

Luego, como parte del segundo grupo, Bitcoin también ha sido estudiado frente a otros mercados financieros más tradicionales, que suelen tener mayor antigüedad y por ende literatura estudiando sus comportamientos. Por ejemplo, Bitcoin ha probado poder ser utilizado como instrumento de cobertura al largo plazo frente a índices como el Euro STOXX, Nikkei, Shanghai A-Share, S&P 500 y el TSX (Chan et al., 2019). A lo que se suma que dicha relación es mayoritariamente observable al utilizar frecuencias de datos mensuales (Chan et al., 2019).

Del mismo modo, otras investigaciones han indagado sobre cómo los volúmenes de negociación en dichos tipos de mercados influyen sobre Bitcoin, observando la existencia de una relación negativa con el volumen en el mercado de la criptomoneda (Jerdack, 2018). Y en lo que respecta a los retornos de Bitcoin, estos se ven afectados por los retornos del Índice S&P 500, el precio del petróleo y el precio del oro (Jareño et al., 2020).

No obstante, la relación entre Bitcoin y el oro es compleja, dado que hay posturas diferentes para describirla. Por una parte, se han dado origen a resultados implicando que el precio del oro no se relaciona significativamente con los precios de Bitcoin, en lo que respecta al uso de la criptodivisa como “safe haven” (Kristoufek, 2015). Mientras que, por la otra, publicaciones señalan que es muy común en la literatura posicionar a Bitcoin y el oro como dos activos muy similares en cuanto al uso que se les da como “safe haven”, su descentralización, oferta limitada y la minería alrededor de ambos recursos financieros (Baur y Hoang, 2021). Lo cual contrasta con resultados empíricos que indican describen a ambas

variables con un comportamiento muy distinto, por lo tanto, mencionar que son activos similares no es acertado (Klein et al., 2018).

La discusión sobre cómo se vincula a estos dos activos intenta incorporar distintos diseños de investigación para intentar descifrar de mejor manera una respuesta más clara. Siendo así como surgen otras publicaciones que proponen dinámicas diferentes entre el oro y Bitcoin, argumentando que se complementan el uno al otro como instrumentos de cobertura (Su et al., 2020).

Finalmente, en el tercer grupo, hay quienes estudian la relación entre Bitcoin y las variables financieras tomando en cuenta periodos específicos de incertidumbre, tal como la pandemia de Covid-19. Donde se ha descubierto que, al haber alta incertidumbre durante la contingencia, los retornos en el índice S&P 500 impactan los retornos de la semana siguiente de Bitcoin (Nguyen, 2022). De la misma manera que se ha señalado a Bitcoin como una buena respuesta ante la incertidumbre provocada por la pandemia, en un contexto de portafolio de inversión (Zhai, 2020).

Con ello se tiene un diagnóstico de la manera en que Bitcoin es estudiado frente a variables de este tipo. Observándose distintos enfoques y objetivos de investigación, que buscan revelar aspectos de la relación existente frente a la criptomoneda. No obstante, los estudios sobre Bitcoin se extienden aún más a lo largo de otros grupos de variables. Razón por la que en la siguiente sección se inspecciona la literatura sobre otras divisas o activos.

4.3.2. Relación con otras monedas

Como Bitcoin es una criptomoneda cuya utilización como moneda de cambio es muy polémica, en la literatura hay diversas premisas que la comparan con monedas fiduciarias.

Por ejemplo, frente a otras monedas como el dólar estadounidense (USD), Yen japonés (JPY) y el Euro (EUR), Bitcoin es considerada un activo con retornos significativamente más inestables (Adnan Fendi et al., 2019).

También se ha obtenido que el precio de Bitcoin se correlaciona de manera muy débil con las tasas de cambio del Euro y el Yuan chino (Giudici y Abu-Hashish, 2019). Teniéndose similares resultados con el dólar estadounidense, ya que no afecta de manera significativa a Bitcoin (Ozyesil, 2019). Aunque, a diferencia de las otras monedas, para el dólar sí se ha comprobado la existencia de una correlación con valor negativo y suficiente significancia estadística (Kwon, 2020).

Si bien los estudios entre Bitcoin y otras monedas fiduciarias coinciden en la débil relación que existe entre ellos, esto no se replica al momento de estudiar su relación con otras criptomonedas. Ya que Bitcoin presenta mayor correlación con otras criptodivisas en comparación a las modelos tradicionales (Giudici y Abu-Hashish, 2019). Por esa razón, en el siguiente apartado se cubrirán los dichos estudios, para poder entender los aportes en esa área de la literatura.

4.3.3. Relación con otras criptomonedas

Como Bitcoin es la criptomoneda más antigua y con mayor capitalización de mercado, es una variable muy presente al momento de analizar cómo se relacionan estos activos entre sí.

Aun así, pese a que Bitcoin es la criptodivisa con mayor capitalización, no se considera que domina el mercado de las criptomonedas (Yi et al., 2018). Es decir, el resto de criptodivisas también son capaces de generar efectos en Bitcoin y el mercado en general. Especialmente cuando se toma en consideración a otras de ellas que también tienen una capitalización alta, porque existe evidencia respaldando que los cambios bruscos en la

volatilidad de monedas con mayor capitalización bursátil suelen influir en las variabilidades de las otras criptodivisas con menor capitalización (Yi et al., 2018).

Aunque Bitcoin no tenga un dominio único en el mercado, su valor ha demostrado tener correlación con el precio de altcoins como Litecoin, Ethereum, Dash y Monero (Silhavy et al., 2020; Hossain e Ismail, 2021). Al igual que se ha concluido que los retornos de criptomonedas como Bitcoin, Dogecoin, Ethereum, Litecoin y Ripple (entre otras) suelen aumentar en conjunto cuando el valor de una de ellas incrementa rápidamente (Bouri et al., 2020).

Luego, en cuanto a relaciones entre solo dos criptomonedas, se tiene que Litecoin se ha mostrado con mayor volatilidad que Bitcoin (Miglietti et al., 2019). Al igual que Bitcoin y Ethereum presentan causalidad bidireccional entre sus precios, es decir, los valores de una influyen en las futuras cifras de la otra (Sifat et al., 2019).

No obstante, la manera en que las criptomonedas se relacionan no parece ser constante en el tiempo, puesto que después de las grandes alzas y caídas de precio registradas en el periodo de 2017 a 2018 la persistencia de shocks de precio ha aumentado, al igual que se aprecia mayor divergencia de precio entre las criptomonedas (Yaya et al., 2019). Por consiguiente, los acontecimientos históricos en la trayectoria de Bitcoin y estos activos cobran alta relevancia, dado que parecen ser factores de cambio frente a las dinámicas existentes en el mercado de criptoactivos.

Los vínculos ente criptomonedas incluso han sido analizados desde enfoques más implícitos que el simple comportamiento del mercado. En vista de que algunas de las criptodivisas que poseen desarrolladores en común han manifestado una sincronización de sus retornos luego de la intervención de dichos programadores, poniendo así en riesgo la transparencia que caracteriza a este tipo de activos (Lucchini et al., 2020). Generando así que

nuevos enfoques y cuestionamientos se sumen a la literatura, para comprender de mejor manera los efectos que condicionan la formación de precio de estos activos.

De esta manera se comprende los distintos enfoques ante los que se estudian las criptomonedas entre sí. Pero esta área de investigación sobre Bitcoin también va más allá de los tópicos financieros, extendiéndose a áreas como las comunicaciones y atención de los inversionistas. Por esa razón en la siguiente sección se hará revisión de publicaciones que analicen a los medios de comunicación frente a Bitcoin.

4.3.4. Relación con medios de comunicación

La atención dirigida a Bitcoin comenzó a ser muy alta a partir de sus rápidas alzas y caídas de precio a contar del año 2017 (Zhang y Wang, 2020). Provocando que medios de comunicación de carácter noticioso elaboren contenido para informar de los acontecimientos más recientes sobre esta criptomoneda.

Sin embargo, estas noticias pueden cargar con más que solo información objetiva, traduciéndose en que también se transmiten emociones. Las que serán luego recibidas por el público, y entre ellos, los inversionistas.

El análisis de dicho aspecto emocional que poseen los medios ha llevado a descubrir que, por medio de herramientas de Aprendizaje Automático (“machine learning”), el texto al interior de las noticias sobre Bitcoin pueden servir para predecir el precio de esta criptomoneda (Yao et al., 2019). De la misma forma que los retornos de Bitcoin también se ven influenciados por los sentimientos en reportes noticiosos en la web, que se obtienen a través de técnicas de lingüística computacional (Polasik, Piotrowska, Wisniewski, Kotkowski y Lightfoot, 2015).

Luego, a un nivel más específico, se ha obtenido que las noticias relacionadas con expertos en Bitcoin generan en el corto plazo un movimiento en el precio que sigue la

dirección del sentimiento transmitido (Karalevicius, Degrande y Weerdt, 2018). Es decir, si el sentimiento es positivo el precio sube, mientras que si es negativo esta baja. Lo cual revela parte de las mecánicas que explican cómo los medios de comunicación logran influir en el mercado del Bitcoin.

Aun así, la literatura sobre la relación entre Bitcoin y los medios de comunicación es bastante escasa cuando se no se toma en consideración la actividad en redes sociales. En vista de que estas plataformas son un potente medio por el que viajan las reacciones a dichas noticias. Por esa razón, en el próximo subcapítulo se cubre a mayor profundidad el concepto de redes sociales y las publicaciones en torno a cada plataforma individual.

4.4. Redes sociales y la web

En este subcapítulo se abordará los conceptos de redes sociales y la web, además de presentar los sitios web considerados dentro de ellos para esta investigación.

4.4.1. Definiciones

Dentro de las definiciones de Redes sociales, aquella que más se ajusta al sentido que se le da al concepto en esta investigación es: “Sitios web y aplicaciones tecnológicas que permiten a los usuarios compartir contenido y participar de actividades de carácter social” (Aichner et al., 2021). Abarcando así correctamente los sitios web como Twitter, Facebook, YouTube o Reddit. Que son los que se quieren analizar en la presente memoria a modo de red social.

Por otro lado, la Web (también conocida como World Wide Web) se puede definir como la parte más prominente de internet, que corresponde a un sistema tecno-social¹ para la interacción de las personas, empleando para ello redes de tecnología (Aghaei et al., 2012).

¹ Se entiende por sistema tecno-social a “un sistema que mejora la cognición humana, la comunicación y la cooperación” (Aghaei et al., 2012).

No obstante, al ser un concepto tan amplio y con componentes menos reconocibles a los de las redes sociales (como son las distintas plataformas), es necesario delimitar aquel sector de la web que será considerado.

La presente investigación solo hace referencia al sector de internet al que es accesible por buscadores de fácil acceso como Google (Google, s. f.), Bing (Microsoft Bing, s. f.), Yahoo! (Yahoo Search, s. f.), entre otros. Es decir, se toma en consideración solo el 4% de las fuentes disponibles en internet, en vista de que el 96% restante corresponde a sectores como la internet profunda (“Deep web”) y la internet oscura (“Dark web”), a los que no se puede acceder por buscadores comunes (Odey y Okoro, 2021).

4.4.2. Principales plataformas a considerar

A continuación, se detallan breves descripciones de los sitios web a estudiar:

- Twitter: Es un servicio de red social en línea que ofrece comunicación a través de mensajes breves y frecuentes, llamados “Tweets”, que pueden contener fotos, videos, enlaces y texto (New user FAQ - Tweeting, s. f.). En él las personas pueden optar por ser seguidores de otras, con el propósito de ver sus mensajes de manera habitual. Dichos mensajes pueden ser visualizados por todos los usuarios de la plataforma utilizando el buscador, a menos que se trate de una cuenta privada, que restringen quiénes pueden ver sus Tweets a solo sus seguidores autorizados.
- Facebook: plataforma de red social utilizada para conectar con personas, enviar mensajes, compartir contenido, formar grupos, comercializar, entre otras funciones. Uno de sus propósitos es ayudar a sus usuarios a conectar con amigos, familia y gente con intereses similares (Facebook, s. f.).

- YouTube: sitio web que permite compartir videos, ya sean de elaboración propia o creados por otros usuarios. Las cuentas registradas tienen la opción de suscribirse a los canales de otras personas, pasando a ser un “seguidor” de ellos, lo que permite que sean notificados cuando un nuevo video está disponible. También es posible dejar comentarios en los videos y entregar retroalimentación, ya sea con un “me gusta” o un “no me gusta”. La plataforma cuenta con una amplia variedad de categorías en sus videos, además de herramientas que facilitan la exploración del sitio web para los usuarios, tales como: buscadores, recomendaciones y noticias (Ever wonder how YouTube works?, s. f.).
- Reddit: Red social que corresponde a una colección de comunidades construidas alrededor de diferentes tópicos, llamados “Subreddits”, a los cuales las personas pueden suscribirse, redactar publicaciones y comentar (Phillips y Gorse, 2018a). También los usuarios pueden valorar aquellos mensajes que son útiles para las conversaciones (How does voting work on Reddit?, 2022) y reaccionar a cada uno de ellos (Reddit Help, 2021).
- Google: Fundado en 1998, es uno de los principales buscadores disponibles en la web (Pandey y Pradhan, 2018), que cuenta con la mayor cuota de mercado entre sus competidores (Statista, 2022). Permite realizar búsquedas para encontrar sitios web de interés para el usuario, resguardando que las respuestas sean lo más precisas y seguras posibles (Busca de forma segura, s. f.).

- Wikipedia: Página web considerada como una de las “wikis”² más reconocidas, que funciona como una enciclopedia gratis disponible para ser editada por sus usuarios alrededor del mundo (Obregón-Sierra y González-Fernández, 2020). Su contenido suele estar acompañado por múltiples enlaces de referencia, razón por la cual este sitio cumple la función de intermediario entre los buscadores y páginas web que contienen información de interés para el usuario (Piccardi et al., 2021).

4.5. Estudios sobre Bitcoin, Redes sociales y la Web

A continuación, se revisan publicaciones que relacionan a Bitcoin con cada red social y sitio web observados en la literatura, a modo de recopilar antecedentes sobre cómo es que se han estudiado estos vínculos en los últimos años. Para así poder enfocarse en aquellas que se relacionan de manera más estrecha con el material audiovisual.

4.5.1. Twitter

Esta red social ha motivado académicos para utilizar la información que hay en ella como representación del pensamiento, emociones o comportamientos de los inversionistas. En vista de que Twitter alberga un gran flujo de mensajes breves creados por los usuarios, que en conjunto dan pistas sobre las dinámicas seguidas por los distintos grupos de personas que hablan sobre un mismo tópico.

Es así como en la literatura sobre Bitcoin se encuentran múltiples artículos que analizan el texto contenido en Tweets, identificando los sentimientos contenidos en ellos. Lo cual luego es utilizado para predecir el comportamiento del precio de la criptomoneda

² Una “Wiki” es un tipo de sitio web o conjunto de estos que pueden ser editados por toda persona a la que se le permite el acceso (Aghaei et al., 2012).

(Colianni et al., s. f.; Pant et al., 2018; (Sattarov et al., 2020). De manera similar, las emociones provenientes de Twitter han demostrado estar relacionadas con variables de Bitcoin como el precio, volumen de intercambio, retornos y volatilidad (Kaminski, 2014; (Gao et al., 2021).

También hay quienes no optan por utilizar un análisis de emociones, sino que estiman la atención de los inversionistas sobre Bitcoin con el número de Tweets que contienen el término “Bitcoin” dentro del texto. Descubriéndose que dicha cifra es capaz de impulsar los valores del día siguiente propios de la volatilidad y el volumen de intercambio de esta criptodivisa (Shen et al., 2019).

Incluso hay publicaciones que dirigen su atención a únicamente los Tweets emitidos por cuentas individuales que a menudo pertenecen a celebridades. Tal es el caso de Elon Musk, cuyos mensajes en la plataforma han sido estudiados en relación al comportamiento de Bitcoin, pero sin obtener resultados significativos (Ante, 2021). Por el contrario, otros investigadores han hallado que los sentimientos en los Tweets emitidos por el expresidente de los Estados Unidos, Donald Trump, sí están correlacionados y pueden predecir variables del mercado de Bitcoin (Huynh, 2021).

De esa forma se cubren los principales antecedentes encontrados sobre Twitter y Bitcoin, por lo que se continúa con la plataforma Facebook.

4.5.2. Facebook

Al momento de identificar la literatura que relaciona a Bitcoin con Facebook se observan distintas maneras de utilizar la información que entregan los servicios de este sitio web. Ya sea enfocándose en sus características como red social, empresa, desarrolladora de herramientas o impulsora de nuevas propuestas sobre Blockchain.

Por ejemplo, el volumen de intercambio de Bitcoin ha sido analizado frente a las emociones contenidas en noticias financieras presentes en un grupo de gran popularidad en Facebook, obteniéndose que no existe una relación significativa entre ambas variables (Jerdack, 2018). Dicha publicación opta por enfocar el estudio de Facebook al grupo sobre Bitcoin con más seguidores a la fecha, llamado “Bitcoin Chart”. Lo cual es un punto importante para los diseños investigativos vinculados a redes sociales, ya que destaca la necesidad de hallar fuentes de datos que puedan representar a la mayor cantidad de interacciones posibles en la red. De hecho, una de las críticas que los autores hacen de su investigación es que hubiese sido valioso recolectar más noticias desde otros grupos de gran popularidad en Facebook (Jerdack, 2018). Por lo tanto, se confirma que al seleccionar datos para esta memoria una buena consideración sería abarcar la mayor cantidad de actividad posible dentro de cada red social.

Luego, en cuanto a los estudios que analizan a Facebook desde un punto de vista empresarial, se ha descubierto que precios de los instrumentos financieros de Facebook han demostrado influir en el precio de Bitcoin (Bartos, 2015). Como también se ha propuesto que el crecimiento de esta red social y Bitcoin siguen el comportamiento de modelos matemáticos como la función de Gompertz (Peterson, 2019). Lo cual, según la misma fuente, impulsa la idea de que el precio de Bitcoin no es necesariamente el resultado de las emociones al invertir, sino que sigue principios característicos de las economías de red (Peterson, 2019).

Por otro lado, hay publicaciones que han optado por estudiar a Bitcoin a través del algoritmo de código abierto desarrollado por Facebook: “Prophet”. Con él se han modelado las proyecciones del consumo de energía y la cantidad de desechos tecnológicos generadas por la minería de Bitcoin (Jana et al., 2022). Al igual que la relación entre esta criptomoneda, el mercado del oro y los cambios monetarios de estados unidos (Akdag y Emsen, 2021).

Por último, la última rama identificada en la literatura sobre Bitcoin y Facebook comprende la investigación acerca de la criptomoneda creada por Facebook: “Libra”. Con la cual se estudia su impacto frente a la volatilidad de Bitcoin (Senarathne, 2019), además de comparársele con otras criptomonedas como Ethereum y Bitcoin (Li & He, 2020).

Con esta descripción se logra visualizar la literatura ligada a Facebook, por lo que se continúa con la siguiente red social: YouTube.

4.5.3. YouTube

Otra red social de gran importancia en la actualidad es YouTube, que se basa principalmente en información con formato de video, distinguiéndose así de las otras redes sociales consideradas en esta sección (cuyo elemento principal es el texto). No obstante, los estudios que relacionan a esta plataforma con Bitcoin son bastante escasos, ya que solo se logró localizar una única publicación que relaciona ambos términos. Por lo tanto, como este trabajo de memoria tiene el objetivo enfocarse en el estudio de contenidos audiovisuales, hacer revisión del único antecedente encontrado es muy importante.

La investigación realizada por (Partida, 2019) analiza las dinámicas del precio de Bitcoin frente a las visualizaciones de videos en YouTube y búsquedas realizadas en Google, describiendo una relación de causalidad desde el mercado de la criptomoneda hacia el resto de variables (p.3). Por lo tanto, su conclusión respalda la idea sobre la existencia de vínculos entre Esta red social y la criptomoneda, siendo así un argumento más a favor de la importancia del material audiovisual sobre Bitcoin para el mercado.

En cuanto a la metodología para obtener datos, esta publicación hace uso de “Web Crawling”, una herramienta de programación que logra identificar los videos de YouTube sobre Bitcoin y registrar cifras como las visualizaciones de cada video (Partida, 2019). Tal procedimiento sirve como referencia para esta memoria, ya que muestra una forma de extraer

datos de YouTube. La cual posee características como considerar un gran número de videos y seleccionar delicadamente cuáles de ellos son incluidos en la investigación.

De esta manera se logra cubrir los aspectos más importantes de la única investigación entre YouTube y Bitcoin, razón por la cual se continúa con la plataforma Reddit.

4.5.4. Reddit

Como esta plataforma está compuesta por distintos tópicos de discusión, la literatura ha demostrado tener interés sobre qué es lo que se discute en esta red social cuando el tema principal es Bitcoin. Ante lo cual se ha encontrado que las publicaciones de Reddit con carácter subjetivo sobre Bitcoin y otras criptomonedas suelen provocar discusiones de mayor tamaño, duración y probabilidad de ser virales en internet (Glenski et al., 2019).

Las conversaciones encontradas en subreddits como “/r/Bitcoin” (Reddit, s.f.) han permitido descubrir que las cifras del número de publicaciones creadas, autores novicios y nuevos suscriptores son capaces de predecir el precio de Bitcoin al largo plazo (Phillips y Gorse, 2018a). Al igual que, en conjunto con el subreddit “/r/ BitcoinMarkets”, preceden el comportamiento de la volatilidad del precio y los retornos de la criptomoneda (Phillips y Gorse, 2018b).

Además, en sintonía con la literatura de otras redes sociales, las interacciones dentro de Reddit han sido analizadas en cuanto a las emociones que contienen. Encontrando así resultados que afirman la facultad de estas variables para predecir el precio de Bitcoin (Prajapati, 2020). Al mismo tiempo que otras publicaciones agregan al estudio de las emociones en Reddit la consideración de variables como el idioma utilizado, disminuyendo así los errores al predecir el precio de Bitcoin (Glenski, 2019).

También los datos de Reddit han sido estudiados en cuanto a: la precisión con la que pueden predecir la dirección de los movimientos del precio de Bitcoin (Wooley et al., 2019),

las palabras que más se correlacionan con las dinámicas del precio de esta criptomoneda (Burnie y Yilmaz, 2019), y la capacidad para pronosticar burbujas financieras en los precios de distintas criptomonedas, siendo Bitcoin una de ellas (Phillips y Gorse, 2017).

4.5.5. Google

Esta plataforma se caracteriza por ser uno de los más grande buscadores disponibles en internet, que es considerada uno de los principales lugares donde los inversionistas de criptomonedas acceden a la información sobre el mercado, por sobre otros buscadores como Yahoo, Bing o Ask (Aslanidis et al., 2022).

Los datos correspondientes a este sitio web suelen provenir de Google Trends, una herramienta que permite observar una aproximación relativa de los niveles de búsquedas³ en función del tiempo, según los términos que sean consultados. A partir de ellos suele medirse la atención que las personas dirigen hacia los términos buscados, que en este subcapítulo serán las criptomonedas o Bitcoin.

Existen estudios que han logrado predecir la dirección de los cambios en el precio de Bitcoin a partir de información proveniente de Twitter y Google, utilizando entre sus variables a los niveles de búsqueda relativos de palabras como “Bitcoin” o “BTC” (Abraham et al., 2018). Lo cual coincide con lo obtenido por (Zhang y Wang, 2020), quienes midieron la atención de los inversionistas con datos de Google Trends e indican que posee un significativo poder predecir el comportamiento las principales 20 criptomonedas a la fecha del estudio (entre ellas Bitcoin).

³ Se le considera una aproximación relativa puesto que la plataforma no entrega el número de búsquedas realizadas, sino que dispone de un valor normalizado a cifras que van de 0 a 100, donde 0 es muy baja popularidad y 100 es la popularidad más alta dentro del periodo consultado.

Adicionalmente, los niveles relativos de búsqueda para la palabra “Bitcoin” presentan altos niveles de correlación con el precio de dicha criptomoneda (Kristoufek, 2013; Mittal et al., 2013), al igual que con los retornos y el volumen (Zhang et al., 2018; Matta et al., 2015). Es más, hay evidencia de que la atención de los inversionistas medida por Google Trends también intensifica los cambios bruscos en el precio de Bitcoin, impulsando las alzas y empeorando aún más las caídas (Kristoufek, 2015).

Por otro lado, se tiene registro de investigaciones que señalan a variables como los retornos de las criptomonedas (considerando a Bitcoin), la volatilidad o el volumen como aquellas que provocan cambios en la atención de los inversionistas (Lin, 2021; Urquhart, 2018). Es decir, a diferencia de publicaciones anteriores, postulan que las criptodivisas son aquellas que influyen en la atención de los inversionistas, y no lo inverso. Generando así discusión sobre cuál es la dirección causal dentro de este vínculo.

De manera alternativa, hay publicaciones que comparan los niveles relativos de búsquedas correspondientes a diferentes términos de búsqueda frente a Bitcoin. Tal como es el caso de (Yelowitz y Wilson, 2015), quienes analizan la correlación de Bitcoin con términos propios de la programación, los inversionistas especulativos, el libertarismo y las actividades criminales. Obteniendo que la criptomoneda se correlaciona positivamente con términos criminales y de programación (Yelowitz y Wilson, 2015).

Con esta revisión es posible comprender las maneras en las que se utiliza Google Trends frente a Bitcoin, al igual que los principales hallazgos descubiertos. Ante esto, se continúa con el siguiente sitio web: Wikipedia.

4.5.6. Wikipedia

Al igual que Google, Wikipedia es frecuentemente utilizado en la investigación sobre criptomonedas debido a su buscador. A partir del cual se obtienen datos que apuntan representar aquellos tópicos que captan el interés de las personas, cuando ellas consultan términos dentro de esta enciclopedia.

Las búsquedas en este sitio web también han demostrado estar correlacionadas con el precio de Bitcoin (Kristoufek, 2013). Aunque existen resultados que limitan dicha correlación a periodos de largo plazo (Kristoufek, 2015). E incluso se hay ocasiones donde Wikipedia no describe tener relación alguna con Bitcoin (Yu et al., 2019).

A partir de lo anterior se concluye que la literatura sobre Wikipedia y Bitcoin es relativamente escasa, además de estar polarizada en cuanto a la relación que existe entre ambos conceptos. Aun así, su revisión aporta conocimiento sobre las capacidades de este sitio web para la investigación, para tenerlo en consideración a la hora de hallar una metodología que logre dar respuesta al problema de investigación.

Finalmente, Wikipedia es el último sitio web a considerar, por lo que se da comienzo a la revisión de literatura propia del contenido acerca de Bitcoin en redes sociales.

4.6. Contenido acerca de Bitcoin en redes sociales

4.6.1. Panorama general

La literatura sobre redes sociales, la web y Bitcoin cubierta en las secciones anteriores posee como principal eje de interés a textos en línea que mencionan a esta criptomoneda, siendo estos emitidos por personas partícipes de las conversaciones o búsquedas que se dan a lugar en las distintas plataformas cubiertas. Sin embargo, no existe abundante literatura que identifique los contenidos que puedan dar origen a ese tipo de conversaciones sobre Bitcoin. Razón por la cual resulta de interés investigar más sobre estos.

Es de esperar que quienes comparten sus opiniones y conocimientos en las redes sociales tengan algo en común con aquellos que leen esos mensajes para luego responder. De otra forma, no habría manera de que las personas pudieran coincidir entre ellas y charlar. Por ejemplo, puede que una publicación o video de carácter noticioso proponga hablar sobre Bitcoin, atrayendo personas que quieran hablar sobre dicho tema. Tomando en consideración que, normalmente, las cuentas de redes sociales como YouTube promueven la participación de la audiencia por medio de indicadores como los comentarios o “me gusta”. Ya que esto mejora el posicionamiento del contenido (rankings), da la imagen de una audiencia activa, y aumenta el potencial de estos videos como herramientas de marketing (Khan, 2017).

Algunas de las materias de discusión que pueden existir son los estudiados en relación al Bitcoin, son la seguridad y usabilidad de la criptomoneda en su rol de método de pago digital (Alshamsi y Andras, 2019), puesto que las personas poseen opiniones diferentes frente a usar Bitcoin comercialmente. Del mismo modo, también se discute sobre cuál es el mejor uso que se le puede dar al Bitcoin como instrumento de cobertura, de diversificación o de refugio seguro (Kliber et al., 2019). Debido a que las respuestas varían según el escenario económico global, la situación de cada país y las preferencias del inversionista.

De esa forma es posible identificar en redes sociales focos que generan conversación, siendo los mensajes de dichas interacciones aquellas que han sido analizadas indirectamente por la literatura existente. Tal como ocurre al estudiar los efectos de una muestra de tweets frente a Bitcoin, y otros tipos de muestreo similares en plataformas como Facebook o Reddit, que fueron cubiertos en los subcapítulos anteriores.

En el presente estudio se busca analizar uno de aquellos focos de conversación, específicamente el material audiovisual que tiene como tema principal a las criptomonedas. Ya que la literatura en torno a las plataformas enfocadas en archivos de video, como

YouTube, es escasa a la hora de analizarlas frente al mercado de las criptodivisas. Siendo esto algo inusual al considerar la creciente importancia que tienen los elementos visuales en los medios digitales para la comunicación, que es transversal entre las plataformas existentes y requiere que se expanda su entendimiento académico usando nuevas metodologías innovadoras (Highfield y Leaver, 2016).

Por lo tanto, este análisis propuesto constituye un aporte a la literatura que busca esclarecer parte de los múltiples factores que pueden explicar el comportamiento del precio de Bitcoin, desde un enfoque relacionado con la parte audiovisual de las redes sociales. Distinguiéndose así de otras publicaciones centradas en la información con formato de texto disponible en la web (tweets, publicaciones y foros).

Dentro de la categoría de material audiovisual se consideran, para propósitos de este estudio, a los videos disponibles en YouTube, las transmisiones en vivo de la plataforma Twitch y las producciones cinematográficas (películas, series y documentales). Todos ellos teniendo en común, dentro de la muestra a estudiar, que el tema principal son las criptomonedas.

La elección de estas tres categorías de material audiovisual se fundamenta en que ellas suelen construir comunidades de personas que comparten ideas periódicamente en la red, respaldado por la información a ser cubierta en la siguiente sección (“Material audiovisual sobre Bitcoin”). Produciéndose así que este comportamiento de las personas interesadas en las criptomonedas pueda ser estimado, y luego analizado frente al comportamiento en el mercado de Bitcoin.

En cuanto a la decisión del tema principal, cabe mencionar que se define a este como las criptomonedas y no a Bitcoin únicamente. Dado que se considera como innecesario aislar el contenido que habla de forma exclusiva de esta criptodivisa, excluyendo aquel material

que cubre a las criptomonedas en general. Especialmente cuando existen tantas similitudes entre ellas (aludiendo a las altcoins), pueden ser incluidas en un mismo portafolio, y suelen ser constantemente comparadas entre sí como elementos dentro de un mismo sector de las finanzas, tal como fue visto en la sección 3.3.3.- “Relación con otras criptomonedas”. Adicionalmente, se opina que tomar en cuenta solo las producciones que hayan mencionado a Bitcoin desembocaría posteriormente en diseñar una recolección de datos ineficiente, dada la complejidad que conlleva tener que separar cada video donde se haya mencionado solo a Bitcoin o únicamente al resto de criptomonedas.

Por consiguiente, ubicar a las criptomonedas como el eje que conecta todo el material audiovisual a estudiar logra abarcar de mejor manera todos los focos de conversación que podrían tener efectos sobre el mercado de Bitcoin.

Con objetivo de comprender de mejor forma la muestra de contenidos a estudiar, en la siguiente sección se les contextualizará dentro de la literatura científica existente.

4.6.2. Material audiovisual sobre criptomonedas

A continuación, se destacan las principales características de los tres tipos de material audiovisual a examinar. Estos son: Películas, series y documentales; luego Videos de YouTube; y finalmente, Transmisiones en vivo por Twitch.tv.

4.6.2.1. Películas, series y documentales

Las producciones del área cinematográfica son indispensables a la hora de analizar los contenidos en formato de video que están presentes en redes sociales. En vista de que es dentro de las grandes plataformas como YouTube o Twitter donde las personas comparten sus puntos de vista acerca de películas (Ahmed et al., 2015), dejando así registros que permiten analizar este tipo de comentarios. Por esa razón, en esta investigación se considera

que las películas⁴, series⁵ y documentales⁶ acerca de Bitcoin son capaces de describir datos útiles para analizar si los emisores de estos puntos de vista tienen correlación con la actividad en el mercado del Bitcoin.

Los materiales cinematográficos han sido estudiados como agentes capaces de influir en las ideas de las personas. Tal es el caso de la investigación realizada por (Bieniek-Tobasco et al, 2019), donde se destaca que los documentales sobre el cambio climático son de gran importancia para esa área de estudio, en vista de que son capaces de alcanzar un gran rango de personas y generar reacciones en la relación que ellas tienen con las problemáticas medio ambientales. Es decir, se generan respuestas en el comportamiento de quienes consumen el contenido. Lo cual se complementa con las ideas de (Borum et al., 2016), quien menciona que este tipo de contenido audiovisual exponen historias que conectan emocionalmente con la audiencia.

Por otro lado, uno de los medios por donde es transmitido este contenido, como es la televisión, también ha sido estudiado dentro de otras disciplinas en cuanto a su relación con los espectadores. Existen estudios pertenecientes a las ciencias políticas donde se estudia el vínculo de los programas de televisión satíros sobre temas políticos con el conocimiento, compromiso y confianza de la audiencia latinoamericana en Estados Unidos hacia el sistema gubernamental (Martinez y Atouba, 2021).

Del mismo modo, las películas han demostrado servir como fuente de estudio al analizar sus efectos en la población a la hora de cubrir temas relacionados a los movimientos

⁴ Se considera “película” a las obras cinematográficas (Película, s. f.).

⁵ Una serie (o “serial”) se define como “una obra que se difunde en emisiones sucesivas” (Serie, s. f.).

⁶ Según la real academia española, un documental se define como: “Dicho de una película cinematográfica o de un programa televisivo: Que representa, con carácter informativo o didáctico, hechos, escenas, experimentos, etc., tomados de la realidad.” (Documental, s. f.).

de anti-vacunación (Bradshaw et al., 2020). Logrando comprender cuáles son los recursos utilizados en la cinta para convencer a los miembros de ese movimiento, que creen fervientemente en la veracidad de determinados filmes contra la vacunación. Para luego discutir la veracidad de esas ideas.

Con estos antecedentes se observa que este tipo de cintas de video efectivamente cuentan con el potencial para motivar una investigación donde la hipótesis busca comprobar una posible correlación entre la atención a este material audiovisual y el comportamiento de sus espectadores. Que, para este estudio, corresponden a las producciones de video sobre criptomonedas y los posibles usuarios de Bitcoin, respectivamente.

Por ese motivo también es necesario revisar otras publicaciones que han utilizado material audiovisual dentro del ámbito de las finanzas. Para así verificar la viabilidad de aplicar dichos vínculos entre producción y espectador a la temática de criptomonedas.

Primeramente, el material audiovisual posee la cualidad de aportar conocimientos a los espectadores a través de historias, al exponer elementos de la realidad con una narrativa llamativa que mantiene la atención de las personas. De hecho, un estudio realizado por (Asaad, 2016) evaluó la utilidad de estos filmes a la hora de aportar conocimiento a estudiantes acerca de temas financieros. Logrando reunir evidencia respaldando que consumir este tipo de producciones ayuda a comprender conceptos complejos de las finanzas, gracias a la manera en que estos son presentados.

Luego la literatura también aporta estudios que enlazan las películas a redes sociales, a través de publicaciones que analizan cómo las productoras cinematográficas llevan a cabo proyectos para posicionar sus películas en las principales redes y en la televisión (Bhattacharjee et al., 2017). Que luego, según la misma publicación, repercuten en el desempeño de las ventas para la película. Es decir, la actividad en redes sociales entorno a

estas producciones son capaces de influir en las acciones del público interés, traducándose en determinados resultados financieros.

Adicionalmente, hay investigaciones que han trabajado con variables provenientes de la televisión, específicamente series. Este es el caso de (Lepori, 2015), donde se evalúa la relación existente entre los finales de temporada de series televisivas y el mercado de acciones estadounidense. Para lo cual se utiliza literatura respaldando el hecho de que los últimos episodios de las series provocan un impacto emocional negativo en los televidentes. Sus conclusiones afirman la existencia de una correlación, habiendo una baja en los retornos de la bolsa alrededor de los días donde una serie llegaba a su fin. Dejando también como acotación que este efecto fue observado con mayor fuerza en mercados con baja capitalización y alta volatilidad, al ser más sensibles al sentimiento de los inversionistas.

Similares resultados fueron obtenidos por (Zhang et al., 2017) en su publicación acerca de la relación entre el sentimiento de las provincias de China y el mercado bursátil del mismo país. Donde dicha variable sentimental está basada en la audiencia televisiva de cada territorio. Obteniéndose así que existe una correlación positiva entre el mercado bursátil y el sentimiento de los inversionistas.

Estas dos últimas investigaciones refuerzan los antecedentes que respaldan investigaciones entre un mercado financiero y la atención hacia material audiovisual en un medio de comunicación. Al igual que se presentan cimientos que ayudarían a justificar la hipótesis de una posible correlación positiva entre el material audiovisual sobre Bitcoin y el mercado de esta criptomoneda.

Adicionalmente, con el propósito de ejemplificar el entorno cinematográfico sobre criptomonedas, es posible mencionar algunas de las producciones audiovisuales relacionadas con las criptomonedas. Estas comprenden a películas como “Dope” (2015) y “Unfriended:

Dark Web” (2018), series tales como “How to Sell Drugs Online (Fast)” (2019), y documentales como “Cryptopia: Bitcoin, Blockchains and the Future of the Internet” (2020). A las que en un futuro se seguirán sumando nuevas cintas, como por ejemplo “The Infinite Machine”, basada en la vida del cofundador de Ethereum Vitalik Buterin (Frater, 2022). Al igual que una trilogía de películas planificadas por un estudio cinematográfico japonés, las cuales buscarán acercar a la audiencia hacia las criptomonedas a través de la ficción (Bhardwaj, 2022). Por lo tanto, se puede decir que las criptodivisas son un tópico que tiene presencia en los registros del cine y la televisión actual, siendo una trama que seguirá estando activa dentro de los próximos años.

De esa manera se concluye que las películas, series y documentales poseen antecedentes investigativos que sostienen el propósito de esta memoria, junto a una presencia en el mundo cinematográfico que las hacen un atractivo objeto de estudio en la actualidad. Por esas razones en los siguientes subcapítulos se discutirá la mejor manera de representar este material audiovisual frente a Bitcoin.

4.6.2.2. Videos de YouTube

El segundo tipo de material audiovisual a considerar en este estudio son los videos disponibles en la plataforma de YouTube. Este sitio web ha crecido desde ser un repositorio de videos aficionados hasta lograr posicionarse como una de las más grandes plataformas de videos en el mundo, con cerca de dos billones de usuarios a nivel global y 694.000 horas de contenido reproducidas cada minuto (Ceci, 2022).

Su desarrollo, en conjunto con el de otras plataformas de transmisión como Netflix o Amazon Prime, han conducido cambios muy rápidos dentro del mercado audiovisual en línea según un estudio realizado por (Budzinski et al., 2021). Quienes además señalan que la audiencia antes se centraba únicamente en la televisión, pero en la actualidad se observa a

las personas más jóvenes prefiriendo las nuevas plataformas al momento de consumir contenido audiovisual.

Es por esta razón que un medio como YouTube resulta muy relevante en la actualidad para analizar las visualizaciones que recibe el material audiovisual en línea. Aún más cuando se tiene en consideración que este sitio web no solo se enfoca en entregar videos para que sean consumidos pasivamente, como sucede a menudo en la televisión. Sino que también se busca construir comunidades y promover la interacción social entre los usuarios de YouTube (About YouTube, s.f.).

Es así como esta página web ha sido estudiada alrededor de cómo se vincula con su público. Existen publicaciones que analizan los anuncios presentes en los videos (Tan, 2018), identificando cuál es su consumidor objetivo y cómo esta exposición puede afectar los comportamientos alimenticios de las personas. Mientras que en otros artículos se opta por explorar el rol educativo que pueden cumplir los videos de YouTube, al indagar si los videos explicativos para clases de contabilidad logran aportar al aprendizaje de los alumnos (D'Aquila et al., 2019). Encontrando que son de gran utilidad para ayudar a entender los contenidos del curso. Razón por la cual se etiqueta a YouTube como parte de los sistemas de aprendizaje online.

De manera similar, se ha comprobado que la cantidad de anuncios contratados en YouTube se correlaciona positivamente con las ventas de una empresa (Zhou et al., 2021). Es decir, a medida que se disponía de más anuncios se observaba un mayor nivel de ventas.

A partir de esta literatura, los videos de YouTube poseen antecedentes que respaldan su influencia en el comportamiento de las personas. Ya sea motivando el consumo de un producto o entregando conocimientos sobre una temática en particular. Siendo esto relevante al estudiar si estos videos también son capaces de correlacionarse con otros mercados o temas

de interés, tal como el Bitcoin para esta memoria. Por ende, esta categoría de contenido audiovisual califica para poder ser considerada como fuente de videos acerca de Bitcoin, con el objetivo de comprobar su posible influencia en el mercado de dicha criptodivisa.

Dentro de los canales de YouTube con mayores visualizaciones se puede destacar a “BitBoy Crypto”, que se especializa en noticias sobre criptomonedas, recomendaciones de inversión, entre otras temáticas (Bitboy Crypto, s.f.). De la misma forma es posible reconocer a otros creadores de contenido dedicados a las criptodivisas y de similar número de visualizaciones, como son: “Andrei Jikh”, “Coin Bureau”, “Altcoin Daily” y “TechLead”.

4.6.2.3. Transmisiones en vivo por Twitch

La tercera y última clasificación de contenido audiovisual a cubrir son las transmisiones en directo en la plataforma de Twitch.tv. Este sitio web está posicionado como la plataforma más popular en cuanto a transmisiones en vivo, superando en horas de sintonía a competidores como YouTube Gaming Live y Facebook Gaming en conjunto (Twitch statistics & facts, s.f.). Su funcionamiento se basa en que los creadores de contenidos, también conocidos como “streamers”, transmiten contenido en vivo por medio de sus canales (cuentas), atrayendo así público en línea que desee observar e interactuar con lo que presentan. Lo cual logra conectar con la audiencia gracias a que el sitio web posee una ventana de texto llamado “Chat” donde los espectadores pueden conversar con el anfitrión, quien lee constantemente esta herramienta y responde a ella (Johnson y Woodcock, 2019). De hecho, algunos “streamers” son capaces de recibir remuneraciones por conceptos de patrocinio, publicidad, suscripciones, donaciones, entre otros tipos de transacciones (Woodcock y Johnson, 2019; Johnson y Woodcock, 2019). Generando así que esta actividad sea más que solo una actividad de recreación, pasando a ser un trabajo.

Adicionalmente, en 2022 el total de sus transmisiones en vivo hasta la fecha describe un promedio de 2,72 millones de espectadores constantes en línea (Twitch Statistics & Charts, s. f.). Lo cual ayuda a estimar que se tiene un gran flujo constante de personas consumiendo el contenido transmitido en Twitch, siendo esto un potencial indicador del alto nivel de interacciones sociales que se dan a lugar durante estos eventos (Diwanji et al., 2020). Es decir, existe un gran número de personas intercambiando información acerca de las temáticas de las transmisiones por medio del chat. Siendo esto de mucho interés para el presente estudio cuando el contenido es sobre Bitcoin o alguna otra criptomoneda.

La manera en que las personas se comunican dentro de Twitch es reconocida como similar a la dinámica presente en Twitter (Oh et al., 2020), en lo que respecta a la emisión de emoticonos y mensajes breves que expresan sentimientos. Con la diferencia que Twitch los presenta durante las transmisiones en vivo y no como una lista de orden cronológico propia de Twitter. Teniéndose así un atributo a favor de la investigación en Twitch, ya que Twitter es un medio con mucha presencia en los estudios sobre redes sociales y el Bitcoin (cuyas publicaciones fueron cubiertas en secciones anteriores de este marco teórico).

Incluso hay investigaciones que resaltan a Twitch por ser un sitio web en donde los académicos pueden estudiar las emociones de la población frente a eventos que afectan a la sociedad, del mismo modo que ocurre con Twitter o Facebook (Chae y Lee, 2022). Lo cual corresponde a una razón más para considerar a Twitch como una fuente de información apta para investigar variables emocionales o del comportamiento.

Adicionalmente las dinámicas de comunicación en Twitch han sido bastante exploradas, teniéndose conocimiento de cómo los espectadores interactúan entre sí y a su vez con el locutor de la transmisión en vivo. Este es el caso de (Kim et al., 2022), quienes analizan los significados de los emoticonos disponibles en Twitch. Según los autores, estos recursos

enriquecen la expresión de quienes lo usan y cumplen un rol no verbal en la comunicación. Pudiendo ser estudiados para entender a los espectadores y la cultura en donde participan.

De la misma manera, (Diwanji et al., 2020) identifican distintos métodos con que los locutores interactúan con su audiencia y cómo es el posterior comportamiento de los espectadores. En donde es posible destacar la construcción de una comunidad alrededor de cada canal de Twitch⁷, en vista de los intereses en común que posee el público y el locutor.

Por otro lado, hay quienes postulan que las nuevas maneras que Twitch utiliza para vincular a un locutor y su audiencia desafían, al mismo tiempo que revitalizan, conceptos provenientes de los estudios acerca de las audiencias televisivas (Spilker et al., 2018). Vinculando así ambos medios audiovisuales dentro de la literatura.

A su vez, la misma publicación recalca que las interacciones en la plataforma no están limitadas a la creencia de Twitch siendo un sitio web destinado a los videojuegos, ya que transmisiones en vivo que no pertenecen a esa categoría han adquirido significativa atención con el crecimiento de este medio audiovisual (Spilker et al., 2018). Por lo tanto, Twitch puede ser considerada una plataforma que atrae audiencias con intereses que pueden ser distintos a los videojuegos. Pudiéndose comprobar al encontrar categorías como “Crypto”, donde diversos canales crean contenido sobre criptomonedas.

Por último, en la actualidad no existe literatura que vincule variables provenientes de Twitch con tópicos de carácter económico. No obstante, la discusión científica acerca de las comunidades en este sitio web (expuesta en los párrafos anteriores) describen características en común con las producciones cinematográficas y YouTube. En el sentido de que son

⁷ Un canal de Twitch es el nombre que recibe el lugar donde cada locutor presenta y almacena cada una de sus transmisiones en vivo, al igual que un conjunto de información complementaria dirigida a describir cuáles son los temas a tratar y la identidad de quien transmite.

medios con una gran influencia sobre las emociones que construyen comunidades a su alrededor. Esto conduce a pensar que Twitch también es una fuente de información con gran potencial para ser estudiada en correlación a otras variables. Por lo que, a través de esta memoria, se propone incorporar a esta plataforma en la literatura al investigar si existe relación entre sus visualizaciones y el actuar de los espectadores. Específicamente cuando las transmisiones en vivo son sobre criptomonedas y se esperan movimientos correlativos en el mercado de Bitcoin.

Entre los canales de Twitch que se dedican a producir contenido sobre criptodivisas se puede destacar a MarketTradersTV, que en su descripción se identifica como una comunidad orientada a enseñar sobre las inversiones en mercado de valores y criptodivisas (MarketTradersTV, s.f.). Otros creadores de contenido de la misma índole son: Tradersamwise, Doctor_degen, TR4D3R10, entre otros.

4.6.2.4. Potencial de este material como fuente de datos

Ante la descripción de estos tres tipos de material audiovisual, es necesario argumentar sobre la manera en que se puede obtener datos de estos medios.

Primeramente, para el conjunto compuesto por películas, series y documentales sobre criptomonedas se requiere una forma de cuantificar la atención que reciben por parte de la audiencia, para luego estudiarlo a lo largo del tiempo frente a variables del mercado del Bitcoin, como el precio. Tomando en cuenta que este tipo de diseño investigativo está en línea con estudios similares comentados anteriormente, involucrando audiencia y factores financieros, realizados con la televisión o producciones cinematográficas (Bhattacharjee et al., 2017; Lepori, 2015; Zhang et al., 2017).

Por esa razón, a la hora de evaluar distintas bases de datos para este tipo de contenido, se considerará aquella información que represente la atención de los espectadores. Por

ejemplo, índices de audiencia (rating), número de visualizaciones en plataformas de transmisión (streaming) o registros de búsqueda en los sitios web que permiten a las personas acceder al material audiovisual.

En segundo lugar, sobre los datos a seleccionar como representativos de la atención que reciben los videos de YouTube sobre criptomonedas, se buscarán cifras relacionadas con los principales indicadores que posee la plataforma. Es decir, la cantidad de visualizaciones para cada video, el número de suscriptores⁸ de cada canal o el nivel de búsquedas relativos a videos sobre criptomonedas.

Al igual que ocurre con el material cinematográfico, este acercamiento está basado en la literatura cubierta alrededor de YouTube y su relación con otras variables (Zhou et al., 2021; D'Aquila et al., 2019)). Procurando que también sea similar al análisis del material cinematográfico y televisivo, produciendo sintonía entre las tres categorías audiovisuales a estudiar.

Finalmente, como la plataforma Twitch no tiene precedentes académicos que la vinculen con variables económicas, se opta por definir los datos a seleccionar en base a lo definido para las otras dos categorías anteriores. En otras palabras, también se considerarán los indicadores que ofrece Twitch, como la cantidad de visualizaciones de las transmisiones en directo, la cantidad de seguidores⁹ y los datos disponibles en las herramientas de búsqueda. Con ello se lograría tener pistas de la atención que recibe este medio audiovisual, cuando el contenido es sobre criptomonedas.

⁸ Un suscriptor de un canal de YouTube es aquel que configura su cuenta para estar siempre al tanto de los últimos videos en dicho canal. Por lo tanto, es un potencial indicador de cuántas personas ven su contenido con regularidad.

⁹ Los seguidores de un canal de Twitch son aquellos usuarios que configuran su cuenta para saber con mayor facilidad cuándo una transmisión de dicho canal está en directo. Para así poder asistir a esta y ver el contenido.

Con esta caracterización de los datos a identificar en las siguientes secciones, la presente investigación establece los cimientos para recopilar dicha información y aplicar una metodología cuyos resultados aporten a la resolución de la problemática presentada en esta memoria. Por lo tanto, en el siguiente subcapítulo se detallan las fuentes de datos disponibles que puedan satisfacer las necesidades de información. Al igual que se establece cuáles de ellas son las más apropiadas entre todas las alternativas, decidiendo en cuanto a la literatura existente.

4.7. Datos disponibles para la investigación

Dentro de los requerimientos de información de este estudio se pueden reconocer cuatro grupos de variables a analizar. Estos son:

- Parámetros del mercado de Bitcoin (Por ejemplo: Precio y Volumen de intercambio).
- Índices de audiencia de material cinematográfico (Atención dirigida a películas, series o documentales sobre criptomonedas)
- Métricas de YouTube (Visualizaciones, suscriptores o nivel de búsquedas de contenido sobre criptomonedas)
- Métricas de Twitch (Visualizaciones, seguidores o nivel de búsqueda de contenido sobre criptomonedas)

Cada uno de ellos será cubierto en su propia sección, primero discutiendo todas las fuentes de datos disponibles sobre sus fortalezas y debilidades. Para luego argumentar las razones por las que se elige una de ellas como la más apta para ser utilizada.

4.7.1. Parámetros de Bitcoin

Para este estudio, las variables acerca de Bitcoin deben ser capaces de representar aspectos relevantes del mercado del Bitcoin, pudiendo ser analizados a lo largo del tiempo frente a datos del contenido audiovisual. También se requiere seleccionar información que posea alta confiabilidad y a su vez sea accesible a la hora de recolectar los datos.

4.7.1.1. Parámetros de Bitcoin - Bases de datos existentes

A lo largo de internet es posible encontrar múltiples plataformas que registran los precios de las criptomonedas, cada una con su propio enfoque y metodologías para generar sus reportes. Algunas de ellas son Coinmarketcap (Coinmarketcap, s.f), Binance (Binance, s.f.), Crypto (Crypto, s.f.), Coingecko (Coingecko, s.f.), entre otras.

Aun cuando hay estudios que atienden distintos problemas de investigación con información proveniente de plataformas conocidas como Coinmarketcap o Binance, las investigaciones que mencionan las diferencias que existe entre estos sitios es muy reducida, sea ese el tema de investigación principal o no. Razón por la cual resulta difícil identificar cuál de ellas presenta mayor precisión y confiabilidad.

Una de las publicaciones que ha aportado a revertir esta escasez en la literatura es (Alexander y Dakos, 2019), donde los investigadores alzan evidencia acerca de la obtención de resultados incongruentes al aplicar un mismo modelo a diferentes bases de datos sobre criptomonedas. Las razones de esta disparidad recaen en las distintas maneras que tiene cada plataforma para determinar los precios de las criptodivisas y una serie de errores de almacenamiento de los datos, como el registro de los precios a diferentes horas del día. Por ende, se debe verificar que la base de datos a escoger no presente grandes errores que puedan perjudicar la veracidad de los resultados a obtener. En caso de existir errores menores, se

buscarán maneras de corregirlas, en vista de que esta última investigación menciona procedimientos recomendados para ello.

Posteriormente, haciendo referencia a la problemática expuesta por el estudio anterior, Vidal-Tomás (2022) analiza la obtención de resultados congruentes al emplear una misma metodología para calcular la eficiencia de mercado de Bitcoin con cada una de las principales bases de datos sobre criptomonedas (Coinmarketcap, Coingecko, BraveNewCoin y Cryptocompare) y las Plataformas de intercambio más utilizadas (Coinbase, Bitstam, Bittrex, Cexio y Exmo) (p.1). Sus conclusiones mencionan que tanto las bases de datos como las plataformas de intercambio estudiadas poseen procesos implícitos idénticos a la hora de calcular los precios de las criptomonedas, razón por la cual se obtienen resultados concordantes entre las distintas fuentes, presentando diferencias mínimas. Por lo tanto, los datos de todas ellas son aptos para ser utilizados con fines investigativos. De hecho, los autores mencionan que bases de datos como Coinmarketcap, Coingecko y BraveNewCoin son las mejores opciones, dado que incluyen la mayoría de las transacciones en criptomonedas. Describiendo así adecuadamente las características del mercado de las criptodivisas (Vidal-Tomás, 2022).

Frente a esta última referencia, se decide profundizar y analizar dichas tres plataformas recomendadas que figuran como las mejores en cuanto a representatividad de los datos ante el mercado. Con el objetivo de obtener información adicional para decidir entre una de ellas.

Coinmarketcap (CoinMarketCap, s. f.) es la plataforma de seguimiento de precios para criptomonedas más citada a nivel global, con apariciones en citas de empresas como Bloomberg, CNBC e incluso el gobierno de Estados Unidos (CoinMarketCap About, s. f.). Fue fundada en 2013 y a partir de 2020 es propiedad de la empresa Binance Capital Mgmt,

quien también posee el control de Binance.com, una de las más grandes plataformas dedicadas al intercambio de activos digitales. Aunque, según el sitio web de la empresa, Coinmarketcap y Binance poseen una estricta independencia entre sí (CoinMarketCap About, s. f.).

Al interior de su sitio web se puede acceder a datos históricos de Bitcoin, donde se cuenta con información diaria de los precios en la apertura y cierre del día (según la zona horaria UTC), precio más alto y más bajo durante el día, volumen, y capitalización de mercado (CoinMarketCap, s. f.). Lo cual puede ser extraído a través de una Interfaz de Programación de Aplicaciones (API por sus siglas en inglés)¹⁰ o manualmente, no existiendo la opción de exportarla directamente como un archivo. Teniendo presente en todo momento que Coinmarketcap permite usar sus datos o gráficos con fines académicos siempre y cuando se cite al sitio web como la fuente de información (CoinMarketCap FAQ, s. f.).

Cabe mencionar que para acceder a datos históricos mediante una Interfaz de Programación de Aplicaciones (API) se debe acreditar un pago de acuerdo a los planes ofrecidos por la plataforma. Donde el plan gratuito no ofrece acceso a datos históricos, pero aquellos que poseen montos más altos ofrecen datos históricos con longitud de un año o más (CoinMarketCap API, s. f.). Por lo tanto, se debe tener en consideración dichos valores a la hora de decidir cómo se accederá a la información.

Como referencia, los datos provenientes de Coinmarketcap han sido utilizados en diversos estudios relativos a Bitcoin, tal como al pronosticar los precios de cierre de Bitcoin y otras criptomonedas (Uras et al., 2020), analizar la capacidad de cobertura de los

¹⁰ Las Interfaces de Programación de Aplicaciones (Application Programming Interface – API) son herramientas para poder acceder a datos de los sitios web actuales vía diversos lenguajes de programación. Requieren identificarse con los proveedores de la plataforma, quienes pueden autorizar o no la utilización de su API (Education, 2022).

instrumentos financieros en base a Bitcoin (Sebastião y Godinho, 2020), y examinar la manipulación de precios en el mercado de esta criptodivisa (Gandal et al., 2018).

Por otro lado, Coingecko (Coingecko, s.f.) es una empresa fundada en 2014 que busca democratizar el acceso a las criptomonedas y dotar a los usuarios de información procesable (Información sobre - Coingecko, s.f.). Corresponde a la mayor fuente independiente de datos sobre criptomonedas del mundo, con más de 13.000 criptoactivos diferentes en más de 600 intercambios de todo el mundo (Coingecko, s.f.). De hecho, es citada frecuentemente en publicaciones de Forbes, The Wall Street Journal, CNBC, Coindesk, International Business Times, Bitcoin Magazine, entre otras firmas. Además de haber servido como base de datos para estudios sobre: el vínculo de Bitcoin con el mercado de valores durante la pandemia de Covid-19 (Nguyen, 2022), la eficiencia del mercado de Bitcoin y otras criptodivisas (Apopo y Phiri, 2021), y las propiedades de esta criptomoneda como activo de inversión (Ram, 2019).

Su plataforma, además de presentar una amplia cobertura sobre el estado actual de Bitcoin en el mercado, ofrece disponibilidad de datos históricos que trazan la trayectoria de su precio, capitalización de mercado y volumen. Teniendo observaciones desde el año 2013 hasta la actualidad, que pueden ser exportadas rápidamente en formato de Microsoft Excel o CSV. Lo cual se traduce en ventajas como, no requerir hacer una compra para acceder a la información al necesitar esta cantidad de información (solo sobre Bitcoin) ni tampoco utilizar una Interfaz de Programación de Aplicaciones (API), que es más conveniente cuando se requiere usar una gran cantidad de datos.

También esta plataforma permite utilizar su API, y por consiguiente los datos, para fines comerciales. Siempre que se incluya una referencia que indique a Coingecko como la fuente de estos. Por lo tanto, se asume que también es válido utilizarlo para fines académicos, tal como esta investigación, al mencionar los enlaces del sitio utilizados en las referencias.

No obstante, estas dos fuentes de datos (Coinmarketcap y Coingecko) fueron criticadas por (Alexander y Dakos, 2019), a causa del procedimiento para calcular los precios entre criptomonedas y un supuesto desfase de un día en las cifras de Coingecko. Pero luego Vidal-Tomás (2022), siendo un estudio más reciente que considera información más actualizada, analizaron nuevamente estas dos bases de datos (junto a otras mencionadas anteriormente) y determinaron que ambas no presentaban irregularidades al ser utilizadas (p.1). Frente a tal controversia en la literatura, el presente estudio decide optar por seguir las conclusiones más actuales sobre esta materia. Por lo tanto, se considera a Coinmarketcap y Coingecko como bases de datos válidas para esta investigación.

Finalmente, BraveNewCoin (Brave New Coin, s. f.-a) se define a sí misma como “una empresa de datos e investigación centrada en la industria de los activos criptográficos y cadenas de bloques (blockchain)” (Brave New Coin, s. f.-b). Fue fundada en 2014, se enfoca en brindar soluciones para datos y ha sido mencionada en Bloomberg, CNBC, Forbes, Washington Business Journal, Yahoo! Finance, Market insider, EconoTimes, entre otras grandes compañías (Brave New Coin, s. f.-b).

Junto a los servicios que ofrece BraveNewCoin también se cuenta con datos históricos, que para el caso del precio se disponen a modo de gráfico en su sitio web, sin existir la opción de exportar los datos directamente. Aunque existe la posibilidad de solicitar una Interfaz de Programación de Aplicaciones (API) para obtener aquellos datos. Lo cual significa tener que contactar a la empresa y, luego de ser aprobada la solicitud, recibir una llave para utilizar la interfaz. Sin considerar que puede ser necesario realizar un pago por este servicio.

Esta plataforma ha sido referenciada en investigaciones relativas a las propiedades de cobertura de Bitcoin (Smales, 2019), la presencia de actividades ilegales alrededor del

mercado de Bitcoin (Hu, 2019), y las consecuencias de que el mercado del Bitcoin admita transacciones en todo momento para las investigaciones académicas (Vidal-Tomás, 2021).

De esta manera se consolida una descripción de las tres posibles fuentes de datos a utilizar para obtener datos de Bitcoin, por lo que en la siguiente sección se argumentará cuál de ellas es la mejor opción para este estudio.

4.7.1.2. Parámetros de Bitcoin - Base de datos a utilizar: Coingecko

Las tres bases de datos analizadas poseen una buena reputación en los medios de comunicación, al ser citadas por firmas de gran reconocimiento en el ámbito financiero. También cuentan con respaldo acerca de la veracidad de los resultados a obtener, porque al usar una misma metodología con ellas se obtienen resultados muy similares (Vidal-Tomás, 2022), generando que no haya razones para discernir entre ellas en cuanto a la calidad de la información. Sin embargo, sus diferencias se hacen notar al analizar el tipo de datos a obtener y su accesibilidad.

Al analizar el tipo de datos a obtener, se tiene que Coinmarketcap y Coingecko son capaces de proveer información histórica sobre el precio de Bitcoin, su volumen y la capitalización de mercado. Mientras que BraveNewCoin, a juzgar por su sitio web, solo presenta registros históricos para el precio en formato de gráfico. Por consiguiente, BraveNewCoin es percibida como una fuente con poca claridad, ya que se desconoce la verdadera composición de sus datos antes de acceder a su Interfaz de Programación de Aplicaciones (API).

Por otra parte, acerca de la accesibilidad, tanto Coinmarketcap como Coingecko poseen maneras para extraer los datos sin necesidad de ocupar una Interfaz de Programación de Aplicaciones (API), lo cual es ventajoso ya que se evita el deber de gestionar una solicitud para obtener acceso a dicha API, que podría extender el periodo de recolección de datos

significativamente. Por ende, estas dos plataformas se posicionan por sobre BraveNewCoin nuevamente, en vista de que esta última plataforma solo admite obtener datos vía una solicitud de acceso a su API.

Posteriormente el criterio que permite discriminar entre Coinmarketcap y Coingecko es la eficiencia para recopilar los datos. Por un lado, excluyendo el uso de las API, Coinmarketcap solo permite acceder a la información manualmente. Es decir, al recolectarlos sería necesario copiar las tablas del sitio web a un archivo de Microsoft Excel o similar, generando la posibilidad de obtener errores en los registros. O bien se extendería innecesariamente el tiempo utilizado en la recolección, ya que se debería revisar datos históricos diarios de los últimos años para eliminar los posibles errores, siendo un proceso sería lento. Es por esa razón que Coingecko es la mejor plataforma para este estudio, debido a que permite extraer la información directamente desde la plataforma a un archivo de Microsoft Excel. Provocando que el proceso de recolección de datos sea mucho más eficiente y con menos probabilidad de error.

Tomando esta decisión en cuenta, se continúa con la selección de bases de datos para el resto de variables audiovisuales que participan de este estudio.

4.7.2. Índices de audiencia de material cinematográfico

Medir la magnitud con que un público dirige su interés hacia determinado tipo de contenido es una labor compleja en la actualidad, dado que los lugares en donde este material audiovisual es presentado son muy variados. Después de todo, un mismo filme puede ser visualizado en una plataforma de transmisión (streaming), en una sala de cine presencial, en la televisión, etc. Por lo tanto, tratar de identificar una forma de medir esta atención requiere tomar en cuenta la mayor cantidad posible de dichos espacios y traducir el movimiento de personas dentro de ellos hacia variables numéricas que puedan ser analizadas en este estudio.

En consecuencia, en la siguiente sección se describen las principales fuentes de información que podrían permitir obtener datos para esta variable.

4.7.2.1. Índices de audiencia – Bases de datos existentes

Los estudios que analizan cómo el nivel de audiencia influye frente a otras variables utilizan una serie de bases de datos con información de series, películas y documentales. Las cuales poseen sus propias variables registradas, criterios y reglas sobre la accesibilidad a la información.

Una de las plataformas más citadas al estudiar datos sobre producciones cinematográficas es el sitio web “Internet Movie Database (IMDb)” (IMDb, s.f.), que fue lanzado en 1990 y es la fuente autorizada de películas, televisión y celebridades más popular del mundo (IMDb Help, s.f.). Posee una gran galería de registros sobre material audiovisual, presentando datos sobre la valoración de cada película por parte de los espectadores, el desempeño financiero de los filmes y la información acerca de los profesionales que participaron en la producción de cada cinta (IMDb Datasets). Pero no posee datos que puedan describir la forma en que cada producción recibe atención del público a lo largo del tiempo.

La plataforma permite la utilización de sus datos de manera moderada, procurando que sea utilizada con fines no comerciales o personales, y solicitando la correspondiente mención a su sitio web (IMDb Help, s.f.). También cuenta con su propia Interfaz de Programación de Aplicaciones (API), que permite acceder a la información de los filmes de manera más rápida y específica. Siendo esto un servicio mayormente pagado y que requiere establecer contacto directo con la empresa (IMDb Developer, s.f.).

Dentro de la literatura es posible encontrar publicaciones que han hecho uso de esta fuente de datos con distintas metodologías. Sus valoraciones por parte de los espectadores y caracterizaciones de los filmes han sido utilizadas para analizar su correlación con distintas

variables, como la actividad de Twitter (Baeza-Yates et al., 2012), el desempeño financiero de las producciones (Bhattacharjee et al., 2017; (Cizmeci y Oguducu, 2018), indicadores de YouTube (Ahmed et al., 2015) y el mercado accionario de los Estados Unidos (Lepori, 2015).

Por otro lado, IMDb puede prestar su utilidad a través de su buscador y rankings, ya que permite identificar grupos de películas con cualidades en común de manera muy rápida y sencilla. Por ejemplo, (Apala et al., 2013; Shruti et al., 2014) uso dicha cualidad para predecir el desempeño financiero en conjunto con información de redes sociales.

Adicionalmente a IMDb existen otros sitios web que registran valoraciones elaboradas por quienes consumen películas o series, tal como RottenTomatoes (Rottentomatoes, s.f.). Que corresponde a una de las fuentes de confianza a la hora de compartir reseñas de producciones cinematográficas (Rottentomatoes: About, s.f.). Aunque, al igual que IMBd, no presenta información sobre la atención recibida por dicho contenido.

Con el objetivo de hallar una fuente de datos que efectivamente permita representar la atención de las personas, se recurre a la literatura acerca del análisis de datos relativos a una audiencia. Llegando así a la investigación realizada por Kalatzis et al. (2018), quienes emplean la herramienta Google Trends y datos de Twitter para predecir los niveles de audiencia en programas de televisión (p.39). Obteniendo como una de sus conclusiones que Google Trends corresponde a un recurso válido para estudiar la atención de los espectadores. Por lo tanto, se incorpora esta plataforma a las fuentes de datos consideradas para ser de utilidad en esta investigación.

Google Trends (Google Trends, s.f.) es uno de los servicios que ofrece Google como complemento a su habitual sitio web de búsquedas, que se caracteriza por proporcionar información sobre el nivel de búsquedas realizadas en las plataformas de Google a lo largo de una zona geográfica, pudiendo observar un estimado de la cantidad de búsquedas que los

usuarios han realizado alrededor de determinados términos. Por ejemplo, al buscar “Bitcoin” en Google Trends, seleccionando como país a Estados Unidos y como periodo a los últimos cinco años, se logra observar un gráfico como el siguiente:



Figura 1: Resultados de Google Trends para "Bitcoin".

Fuente: Google Trends (s.f.)

Por ende, al recurrir a este tipo de gráfico resulta posible identificar cómo varía el interés de las personas sobre un determinado término. Lo cual, al ser aplicado a los títulos de Películas, Series o Documentales, entregaría un estimado de la cantidad de veces que las personas buscan dicho material audiovisual en los servicios de Google. Especialmente cuando se considera que Google Trends también cuenta con una opción para entregar datos sobre búsquedas realizadas en YouTube. Siendo está una cualidad con un gran potencial, ya que YouTube es una plataforma donde están disponibles adelantos (en inglés “Trailers”) u otros tipos de videos sobre las producciones cinematográficas. Por ende, resulta de mucho interés poder adicionar el interés que la audiencia puede tener frente a ambas plataformas, YouTube y Google, aumentando el alcance del presente estudio.

De esa forma, Google Trends se consolida como un posible método para cuantificar la atención dirigida hacia el contenido cinematográfico sobre las criptomonedas. Aún más cuando la plataforma entrega la posibilidad de exportar los datos de este tipo de gráficos a

un archivo CSV (que puede ser utilizado con Microsoft Excel), facilitando así el proceso de recolección de información.

Con ello se considera tener una descripción completa de las posibles fuentes de información a ocupar para esta sección, razón por la que se procede a argumentar la elección que formará parte de la metodología de este estudio.

4.7.2.2. Índices de audiencia - Base de datos a utilizar: Google Trends

A partir de la información recopilada previamente, se decide optar por utilizar las plataformas “Internet Movie Database (IMDb)”, Rotten Tomatoes y Google Trends de manera complementaria.

Inicialmente, tanto IMDb como Rotten Tomatoes se posicionan como fuentes de información de gran potencial para seleccionar qué Películas, Series y Documentales se relacionan con temas relativos a criptomonedas. Gracias a su gran cantidad de registros, diversas categorías y rankings del material audiovisual. Pero, en contraste a lo anterior, ambas no podrían ser consideradas a la hora de obtener datos acerca de la atención dirigidas a dichos filmes, ya que no poseen este tipo de información en sus plataformas.

Para complementar dicha falta de información en las principales bases de datos cinematográficas, se obtiene que la herramienta Google Trends crea una nueva forma de acceder a datos que logren describir la atención dirigida a producciones relacionadas a las criptomonedas. La cual, al ser parte de Google, cumple con ser de gran confiabilidad y aporta datos en un formato conveniente para su posterior análisis (Microsoft Excel).

Además, desde el punto de vista del diseño de investigación, ocupar a Google Trends como fuente de información es una buena alternativa puesto que permite reconocer el interés de las personas sin necesidad de estar limitado a una plataforma de transmisión específica (como Netflix, Amazon Prime u otra). Permitiendo así que se observe la atención del público

por los filmes acerca de criptomonedas, sin importar dónde es que las personas van a interactuar con ese contenido. Por ejemplo, puede que una búsqueda de un determinado título conlleve a visualizar su tráiler en YouTube, leer noticias relacionadas, identificar la cuenta oficial de la película en una red social, o algún otro motivo. Todas esas posibilidades son consideradas al ocupar Google Trends, por lo que se considera una muy buena manera para definir esta variable a estudiar.

De esta manera se concreta la identificación de las fuentes de información para los Índices de audiencia de material cinematográfico. Por lo cual se da inicio a un análisis semejante para la siguiente variable a considerar.

4.7.3. Métricas en canales de YouTube sobre criptomonedas

Analizar la influencia que tienen los videos de YouTube sobre el comportamiento de las personas requiere poder identificar cómo se cuantifica cada visita que el contenido recibe y las acciones vinculadas a ello. Por esa razón es necesario recorrer las principales fuentes de información dedicadas a esta plataforma, sean estas primarias o secundarias con respecto a YouTube.

4.7.3.1. Métricas en canales de YouTube - Posibles fuentes de datos

La información propia de los videos almacenados en YouTube puede ser obtenida directamente de la plataforma, a través de la herramienta “YouTube Analytics”, como también por fuentes secundarias que se dedican a monitorear la actividad en dicha red social, elaborando sus propios resultados e informes estadístico, como es el caso de SocialBlade.

Primeramente, YouTube Analytics (YouTube Analytics, s. f.) es aquella sección de la plataforma que permite visualizar todos los indicadores correspondientes a un canal que produce contenido. Presenta información relativa a las visualizaciones, suscriptores,

anuncios, comentarios, entre otros parámetros que explican el funcionamiento y desempeño de un canal de YouTube.

Al ser una herramienta que forma parte de esta red social, es la fuente de datos con mayor exactitud y confiabilidad a la que se podría acceder. Sin embargo, posee el gran inconveniente de ser considerada como información privada de cada usuario. Por lo tanto, no se puede acceder a los datos de manera sencilla en el sitio web, sino que habría que contactar a los dueños de la información individualmente. Esto se traduce en considerar que una de las líneas de acción a seguir sería la colaboración con distintos canales de YouTube que produzcan contenidos sobre criptomonedas, donde se solicitaría la información luego de negociar con cada uno de estos creadores de contenido.

De esa forma, YouTube Analytics es una fuente de datos que ha sido utilizada en estudios vinculados con el rol de los videos de YouTube para la educación (D'Aquila et al., 2019). Donde los investigadores eran aquellos que tenían acceso al canal que producía los videos, por ende, fue posible acceder a la información usando YouTube Analytics.

Por otro lado, SocialBlade (Social Blade, s.f.) es una plataforma que procesa la información disponible públicamente en YouTube para elaborar estadísticas, gráficos y reportes que describen los indicadores de los distintos canales que producen videos en línea (Social Blade terms, s.f.). En otras palabras, entrega datos complementarios a los que podrían ser visualizados privadamente en YouTube Analytics, por medio del procesamiento de la información pública del sitio web. Por ejemplo, dispone datos semanales sobre los nuevos suscriptores y las visualizaciones obtenidas, junto a información sobre los resultados financieros estimados de cada canal.

El sitio web de SocialBlade fue fundado en el año 2008 y dedica a procesar datos de redes sociales como YouTube, Twitter, Twitch e Instagram para crear información

estadística que permita observar el crecimiento progresivo en cuentas populares de dichos medios (Social Blade info, s.f.). Contando con una base de datos que registra los cambios de más de veintitrés millones de canales de YouTube (Social Blade info, s.f.).

En adición, este sitio web también permite identificar a los creadores de contenido que posee YouTube, ya que cuenta con buscadores capaces de elaborar rankings de popularidad en torno a variadas categorías. De esa forma, SocialBlade aportaría al proceso de reconocer qué canales de YouTube producen videos acerca de criptomonedas, para así luego crear una base de datos con variables ligados a la audiencia, como son las visualizaciones o los nuevos suscriptores.

En cuanto a las características de los datos que pueden ser obtenidos de SocialBlade, se considera que estos poseen una alta confiabilidad. Puesto que en la página principal de su sitio web cuenta con un ícono indicando que poseen una certificación de YouTube (Social Blade, s.f.).

Además, esta veracidad de la información se ve reafirmada por la amplia participación que posee SocialBlade en la literatura de investigación académica. Lográndose identificar su presencia en estudios sobre temáticas como la visibilidad de videos en YouTube (Lopezosa et al., 2020), el carisma y liderazgo de los creadores de contenidos en dicha plataforma (Conde et al., 2020), los anuncios publicitarios de comida en videos para público infantil (Tan, 2018), y el análisis económico de las trayectorias correspondientes a celebridades de YouTube (Budzinski y Gaenssle, 2018). Al igual que se le ha hecho referencia en medios como “The New York Times”, “The Guardian”, “Forbes”, “The Wall Street Journal”, entre otros (Social Blade kit, s.f.).

Sin embargo, uno de los inconvenientes de la plataforma en la actualidad es que sus datos históricos solo pueden abarcar la información de los últimos tres años anteriores. Esto

debido a que YouTube les ha solicitado mantener dicho criterio a contar de noviembre del año 2017, para ajustarse a la legislación sobre privacidad establecidas por la unión europea (Social Blade FAQ, s.f.). Por ende, toda metodología que considere utilizar datos de esta fuente solo será capaz de estudiar los últimos tres años de información.

A partir de estos antecedentes sobre las posibles fuentes de información, en el siguiente subcapítulo se comparan ambas opciones y se establece cuál es la más apropiada para este estudio.

4.7.3.2. Métricas en canales de YouTube - Fuente a utilizar: SocialBlade

Inicialmente YouTube Analytics se posiciona como la mejor plataforma de la que es posible obtener información, al ser una fuente directa desde el sitio web donde los videos son almacenados. No obstante, las políticas de privacidad existentes hacen que solo se pueda acceder a estos datos por medio de colaboraciones directas con los creadores de contenido audiovisual. Lo cual posee la gran desventaja de ser muy extenso, riesgoso (al no poder garantizar la colaboración de todos los creadores) y costoso monetariamente (debido a una posible negociación con los canales). Por esa razón, se considera a esta fuente como no viable, al requerir esfuerzos demasiado grandes en cuanto a tiempo y dinero para esta investigación.

Por consiguiente, se decide que la mejor plataforma disponible para obtener datos de YouTube es SocialBlade. Dado que es posible acceder a la información de diversos canales, contando con la ayuda de rankings que destacan aquellas cuentas con mayor popularidad alrededor de las categorías disponibles (Ciencia y Tecnología, Educación, etc.). Además, sus registros poseen una alta confiabilidad, respaldada por una certificación oficial de YouTube, su presencia en otros estudios de investigación y el hecho de haber sido citada por importantes medios de comunicación.

No obstante, el presente estudio deberá adecuar su metodología para poder trabajar únicamente con información de YouTube recopilada en los últimos tres años, a causa de las políticas de privacidad comentadas en la sección anterior.

A partir de estos argumentos SocialBlade se consolida como la fuente de información a utilizar para estudiar las métricas en torno a canales de YouTube, cuyo tema principal sean las criptomonedas. Por ende, en el siguiente subcapítulo se da inicio al análisis de las fuentes de información para la última variable a participar de esta investigación.

4.7.4. Métricas de canales de Twitch sobre Bitcoin

Las transmisiones en vivo se caracterizan por ser eventos donde coincide un locutor con su audiencia en un mismo espacio, donde constantemente nuevos espectadores entran y otros salen. Esto genera que elaborar registros de este flujo de audiencia sea complejo, dado lo efímero de las transmisiones. Por esta razón se recurre a diversas plataformas especializadas en facilitar el acceso a este tipo de información, que serán descritas en el siguiente apartado, para luego decidir entre una de ellas como la fuente de datos más apta, considerándola así dentro de la metodología de este estudio.

4.7.4.1. Métricas de canales en Twitch - Posibles fuentes de datos

El conjunto de bases de datos identificadas que recopilan y analizan información de Twitch se compone por cuatro plataformas: TwitchTracker, TwitchMetrics, SullyGnome y SocialBlade. Cada una de ellas entrega datos de forma distinta, por lo que se hará revisión de las características que posee cada sitio web.

La primera fuente a describir es TwitchTracker (Twitchtracker, s.f.), una plataforma independiente de Twitch que se especializa en crear rankings de los canales que transmiten contenido en vivo. En sus reportes posee, primeramente, información diaria en función del tiempo para cada canal sobre el número de seguidores, número de espectadores promedio,

horas transmitidas en vivo y horas visualizadas por la audiencia. Mientras que, en segundo lugar, cuenta con un apartado para cada categoría (por ejemplo: “Crypto” o “Bitcoin”) que entrega datos sobre la cantidad promedio de espectadores diarios visualizando dicha programación y el número de canales que han transmitido con dichas etiquetas. En ambos casos se rastrea la actividad con hasta seis años de antigüedad con respecto a la fecha de este estudio, dependiendo de la fecha de creación de cada canal a consultar.

En cuanto al posicionamiento de TwitchTracker en la literatura, se aprecian múltiples investigaciones que la utilicen como fuente de información. Siendo sus cifras utilizadas en estudios sobre las emociones expresadas en Twitch durante la pandemia del Covid-19 (Chae y Lee, 2022), las razones de éxito en los creadores de contenido de la plataforma (Gutiérrez y Cuartero, 2020), las motivaciones de los espectadores para ver transmisiones en vivo (Speed et al., 2022), y las estrategias para construir marcas personales dentro de Twitch (Uva y Rosenthal, 2018).

Luego la segunda fuente a mencionar es TwitchMetrics (Twitchmetrics, s.f.), una empresa de tecnología dedicada a crear presentaciones con datos personales (recibiendo previa autorización), no personales y públicos provenientes de Twitch (Privacy Policy Twitchmetrics, s.f.). Su sitio web contiene rankings de los canales de la plataforma, al igual que el seguimiento diario a lo largo del tiempo de variables como la cantidad de visualizaciones, el número de seguidores y las horas de duración para cada transmisión en vivo. Disponiendo información con una antigüedad máxima de tres meses a contar de la fecha donde la página web es consultada. Por ejemplo, si se recolectan datos en junio de 2022, solo será posible ver los registros históricos hasta marzo del mismo año.

Adicionalmente, los rankings TwitchMetrics han sido consultados por estudios sobre las diferencias culturales presentes en la comunicación dentro de Twitch (Oh et al., 2020), el

entendimiento de los mensajes negativos compartidos durante las transmisiones en vivo (Kim et al., 2022), y la percepción de co-presencia manifestada por los espectadores (Diwanji et al., 2020).

Cabe destacar que sus datos han sido recomendados en investigaciones que analizan la capacidad que tiene Twitch como fuente de datos para distintos tipos de variables a estudiar, entregando una propuesta de estructura para trabajar con ellos (Harpstead et al., 2019).

Posteriormente, una tercera fuente de datos a estudiar es SullyGnome (Sullygnome, s.f.), un sitio web gestionado por una sola persona que dispone estadísticas creadas a partir de consultas periódicas cada quince minutos a la Interfaz de Programación de Aplicaciones (API) de Twitch (About Sullygnome, s.f.). Sus datos contienen información histórica que abarca hasta el año 2015, describiendo así para cada canal la trayectoria mensual de variables como la cantidad de espectadores promedio, el número de horas visualizadas por la audiencia, la cantidad de horas transmitidas en vivo, el número de seguidores ganados, y las visualizaciones obtenidas.

De manera adicional, SullyGnome posee un amplio historial dentro de la literatura, dentro de lo que se pueden destacar investigaciones usando sus cifras y rankings para analizar los mensajes en transmisiones de Twitch durante la pandemia del Covid-19 (Leith y Gheen, 2022) y la existencia de relaciones para-sociales en esta red social (Leith, 2021).

Mientras que sus datos históricos se han empleado para estudiar el potencial de los datos de la plataforma para investigaciones sobre videojuegos (Frommel et al., 2019), la influencia de las transmisiones en vivo de Twitch sobre el consumo de juegos de video en la plataforma de Steam (Oh y Kim, 2022), la relación entre permanecer más tiempo en el hogar durante la pandemia y el emprendimiento profesional en Twitch (Wollborn et al., 2021), y la

descripción de la actividad en torno a transmisiones en vivo de carácter educativo (Steinbeck et al., 2021).

Finalmente, Socialblade (Social Blade, s.f.) es la cuarta fuente de datos a describir, que también fue mencionada anteriormente al buscar recursos donde obtener información de los indicadores relativos a YouTube. Pero para esta sección se evalúa sus características aplicadas únicamente para Twitch.

Esta plataforma elabora reportes estadísticos y los presenta en su sitio web, con información de redes sociales como YouTube, Twitter, Twitch e Instagram (Social Blade, s.f.). Según sus cifras, poseen registros sobre más de cinco millones de canales en Twitch, priorizando hacer seguimiento de aquellos con mayor popularidad (Social Blade, s.f.).

Acerca de la composición de sus datos, se aprecia información histórica semanal y mensual, tanto para el número de seguidores como para la cantidad de visualizaciones del canal. Los cuales tienen una antigüedad extensa, puesto se cree que la restricción de datos que limitaba los indicadores de YouTube a tres años de antigüedad (Social Blade FAQ, s.f.) solo es válida para esa plataforma, al ser una solicitud directa desde esa red social en específico hacia SocialBlade.

En adición, el uso de las cifras y rankings Socialblade en la literatura se extiende a lo largo de investigaciones sobre los derechos de autor entre la industria de los videojuegos y Twitch (Jungar, 2016), la estructura de los sistemas de transmisiones en vivo (Tomassilli, 2019) y el consumo de contenido en Twitch sobre deportes electrónicos (Lux et al., 2019).

A partir de esta información sobre las cuatro bases de datos disponibles se procede a evaluar cuál de ellas es la más apropiada en la siguiente sección.

4.7.4.2. Métricas de canales en Twitch - Fuente a utilizar: TwitchTracker

Considerando las descripciones anteriores, las cuatro bases de datos coinciden en tener un buen posicionamiento dentro de la literatura, al existir numerosas publicaciones que hacen uso de ellas y destacan su potencial para generar conocimiento. Además, todas ellas cuentan con sistemas de búsquedas y rankings que permitirán localizar los principales canales con contenido sobre criptomonedas.

No obstante, al enfocarse en las cifras históricas a obtener, el objetivo de utilizar datos provenientes de Twitch en esta memoria busca identificar cómo estos se relacionan con el mercado de Bitcoin. Por ende, es necesario considerar que la fuente de información a seleccionar debe ser capaz de proveer datos que posean similitud en cuanto a extensión y frecuencia de muestreo con aquellos correspondientes a Bitcoin, que cuentan con observaciones de carácter diario y abarcan desde el año 2015 hasta la actualidad aproximadamente.

Bajo ese criterio, se decide que TwitchMetrics no es capaz de cumplir el requisito de disponer información con longitud semejante a los datos de Bitcoin, ya que solo abarca las métricas de los últimos tres meses desde la fecha en que se consulta. Generando la limitante de contar con muy pocas observaciones para el estudio. Más aún cuando se toma en cuenta que el resto de fuentes de información cuentan con datos de a lo menos dos años de antigüedad, lo cual se considera un periodo aceptable para el análisis, dado el número de observaciones que se tendrían con datos diarios o semanales es mayor.

Del mismo modo, se descarta a SullyGnome ya que su información histórica de larga extensión dispone únicamente en formato mensual. Lo cual no corresponde a la mejor opción para ser analizado frente a Bitcoin, puesto que se limita el análisis con el precio de la

criptomoneda solo a una escala mensual, sin tener la opción de evaluar si un estudio diario o semanal tendría un mejor desempeño para esta memoria.

Luego las dos fuentes de información restantes, SocialBlade y TwitchTracker, comparten cualidades como una longitud suficiente de sus datos históricos (a lo menos dos años) y la presentación de estos a nivel diario o semanal. Lo cual se traduce en que ambas permiten ubicar las variables provenientes de Twitch a un nivel donde pueden ser analizadas en similares condiciones con Bitcoin, en cuanto a la granularidad de los datos y el posible periodo de tiempo a estudiar.

También se aprecia que estas dos bases de datos cuentan con variables que representan apropiadamente la atención de la audiencia. En el caso de SocialBlade se tiene el número de nuevos seguidores en cada semana y la cantidad de visualizaciones semanales del canal. Mientras que TwitchTracker ofrece para cada canal el número de seguidores, horas transcurridas en vivo, cantidad de espectadores promedio, las horas de contenido visualizadas por la audiencia y la cuantificación de visualizaciones obtenidas. Al igual que para las categorías dispone del número de espectadores promedio y la cantidad de canales que transmiten las programaciones sobre criptomonedas.

Sin embargo, una diferencia entre ellas es que TwitchTracker posee datos diarios para sus seis métricas, en cambio, SocialBlade solo es capaz de otorgar cifras semanales para sus dos variables. Por lo tanto, TwitchTracker es la fuente de información más apta para ser utilizada en esta investigación a la hora de obtener datos históricos de Twitch sobre los canales con contenido sobre criptomonedas, ya que sus cifras diarias ofrecen una mayor granularidad de datos, asemejándose a la que poseen los datos de Bitcoin (que poseen formato diario). Permitiendo así decidir posteriormente en qué manera se trabajarán los datos, diariamente o semanalmente, tomando en cuenta el resto de elementos que componen esta

investigación. Al igual que se ofrece un mayor número de variables capaces de describir las interacciones de las personas con este tipo de contenido audiovisual, siendo apreciable que las variables presentes en SocialBlade también son obtenibles en TwitchTracker.

Con ello concluye la evaluación y selección de bases de datos para cada uno de los tipos de variables presentes en este estudio, por lo que se da comienzo a la revisión de las posibles metodologías a utilizar.

4.8. BASES METODOLÓGICAS

Al interior de este subcapítulo se recorrerán brevemente los distintos modelos utilizados en estudios sobre la relación de Bitcoin con otras variables (cubiertos en secciones anteriores), con el objetivo de comprender a grandes rasgos la funcionalidad de cada uno de ellos, junto a sus fortalezas y debilidades. Para así luego poder discutir cuáles de ellos resultan más apropiados para estudiar el vínculo entre el material audiovisual sobre Bitcoin y los datos históricos de la criptomoneda, que serán cubiertos a mayor profundidad en cuanto a su utilización.

4.8.1. Modelos presentes en la literatura

Las publicaciones revisadas sobre el estudio de la relación entre Bitcoin y factores externos a su mercado hacen uso de metodologías que pueden ser agrupadas en tres grupos: Modelos de regresión, Metodologías para analizar correlación y Modelos autoregresivos.

4.8.1.1. Modelos de regresión

Primeramente, los modelos de regresión son aquellos que se basan en la predicción de una variable objetivo, conocida como dependiente, frente a una variable o grupo de ellas (variables independientes). La manera en que las variables independientes son operadas para obtener los valores de la variable dependiente depende del tipo de regresión que se esté

utilizando, ya que estas pueden basarse en diversas funciones (lineal, cuadrática, logarítmica, etc.).

Dentro de estos modelos es posible identificar el uso de la Regresión Lineal Simple, que es capaz de probar si existe una relación lineal entre dos variables al buscar una recta que se ajuste de mejor manera a sus valores. Entre los estudios que han obtenido resultados significativos con ella se destaca el análisis de la relación de Bitcoin con variables como la incertidumbre global (Bouri et al., 2017) y la incertidumbre política (Paule-Vianez et al., 2020; Chen et al., 2021).

Del mismo modo que se ha empleado en publicaciones ligadas a lo audiovisual, como la influencia de la publicidad en videos de YouTube en los resultados financieros de la empresa que los produce (Zhou et al., 2021).

No obstante, este modelo se encuentra limitado a sólo poder detectar relaciones lineales, es decir, su desempeño no es óptimo cuando las variables se vinculan a través de otro tipo de función (Vetter & Schober, 2018).

Luego, a modo de extensión a dicho modelo, se encuentra la Regresión Lineal Múltiple. Esta posee características idénticas a la Regresión Lineal Simple, con la diferencia de que se incluye un mayor número de variables, mejorando de esa manera la estimación. Dado que se disminuye el sesgo generado por variables que tienen influencia en la variable dependiente y no habían sido incluidas (Lepori, 2015). Su uso ha sido de utilidad para analizar a Bitcoin frente a datos de Twitter y Google Trends (Abraham et al., 2018; Mittal et al., 2013), variables técnicas de las criptomonedas (Hayes, 2017), Promedio Industrial Dow Jones (Jerdack, 2018), el efecto de los días de cada semana (Aharon y Qadan, 2019), y precios de otras criptomonedas (Yatsyk, 2020).

En cuanto a las investigaciones sobre material audiovisual relacionadas, se tiene que el modelo de regresión lineal múltiple ha sido utilizado para hallar la influencia de los finales de temporada de series televisivas en el mercado bursátil estadounidense (Lepori, 2015), el vínculo entre la audiencia de la televisión en el mercado bursátil chino (Zhang et al., 2017), la interconexión de las visualizaciones en Twitch con el mercado de videojuegos (Oh y Kim, 2022), y el impacto de redes sociales con el desempeño financiero de producciones cinematográficas (Bhattacharjee et al., 2017).

También se puede apreciar en la literatura la presencia de la Regresión Cuantil, que fue propuesta por Koenker y Bassett en 1978 (Koenker y Bassett, 1978) y se caracteriza por poder entregar resultados de una misma regresión para los distintos cuantiles presentes en los datos (Bouoiyour y Selmi, 2017). De esa forma, se puede analizar la relación entre las variables a estudiar dependiendo de los valores que estas adquieren (de gran magnitud, de baja magnitud, cercanos a la mediana, etc.). Es considerada como una extensión del modelo de Regresión Lineal Múltiple, que resulta adecuada cuando no se cumplen los supuestos de normalidad y homocedasticidad que son exigidos en este último modelo (Paule-Vianez et al., 2020).

La Regresión Cuantil se ha utilizado para estudiar a Bitcoin frente a la incertidumbre global (Bouri et al., 2017), la atención medida a través de Google Trends (Zhang y Wang, 2020), la incertidumbre política (Paule-Vianez et al., 2020), y mercados bursátiles (Nguyen, 2022).

En contraste a lo anterior, no se poseen registros de investigaciones que utilicen esta metodología en estudios ligados al material audiovisual.

4.8.1.2. Modelos para analizar correlación

Posteriormente, al considerar las metodologías correspondientes al segundo grupo, están los coeficientes de correlación. En ellos se destaca el Coeficiente de Correlación de Pearson y el Coeficiente de Correlación de Spearman, cuya utilidad y valores son muy similares, pero se diferencian en sus requerimientos e interpretación. Por una parte, el coeficiente de Pearson permite cuantificar la correlación lineal entre dos variables (Xu & Deng, 2018), teniendo la necesidad de supuestos como la distribución normal de las mediciones a analizar (Schober et al., 2018). Mientras que el coeficiente de Spearman también estima la intensidad de la asociación entre dos variables, pero detecta aquellos vínculos que se apegan a una función monótona en vez de una con carácter lineal (Hauke & Kossowski, 2011). Además de ser considerado una alternativa que requiere menor rigurosidad de supuestos, pudiendo ser utilizado en presencia de valores extremos (Schober et al., 2018) y sin requerir una determinada distribución para las variables (Hauke & Kossowski, 2011).

Acerca de la presencia de estos indicadores en la literatura, el coeficiente de Pearson es ampliamente utilizado en la literatura sobre Bitcoin (Kaminski, 2014; Yatsyk, 2020; Abraham et al., 2018) y aquella relacionada con material en formato de video (Zhou et al., 2021; Kalatzis et al., 2018). Ya que suele ser incluido en análisis iniciales para identificar la asociación lineal entre el comportamiento de dos variables.

Lo anterior contrasta con la utilización del coeficiente de Spearman, ya que su uso se dentro de la literatura sobre Bitcoin es reducido en comparación al índice de Pearson, solo estando presente en el estudio de la relación entre Bitcoin y el interés captado por la criptomoneda según Google Trends (Dokic et al., 2015).

4.8.1.3. Modelos autoregresivos

Finalmente, en tercer lugar, se encuentran los modelos autoregresivos. Son considerados un tipo de modelo regresivo que posee entre sus variables explicativas valores retardados de la variable explicada (Das, 2019), es decir, sus cifras pasadas o futuras con respecto al tiempo donde se lleva a cabo la estimación. Una característica que los distingue de otros tipos de regresión anteriormente mencionados es que los modelos autorregresivos sí son capaces de ser utilizados en presencia de autocorrelación (Das, 2019).

Al interior de las distintas versiones existentes de modelos regresivos es posible destacar dos de ellas que tienen una amplia presencia en la literatura sobre Bitcoin, estos son: El modelo autorregresivo generalizado condicional heterocedástico (conocido como GARCH por sus siglas en inglés) y el modelo vectorial autorregresivo (conocido también como VAR por sus siglas en inglés).

El modelo GARCH corresponde a la generalización del modelo autorregresivo condicional heterocedástico (ARCH), un modelo econométrico que es utilizado para la estimación de la varianza presentada por series de tiempo, en situaciones donde esta es dependiente de la variable temporal (Das, 2019). La diferencia entre ambos es que el modelo ARCH toma en consideración que la varianza puede estar relacionada con el error de un periodo anterior en una serie de tiempo, mientras que el GARCH logra ampliar dicho análisis con un grupo de observaciones pasadas (Gujarati & Porter, 2008). De esa forma el modelo GARCH suele predecir los valores futuros de la varianza al utilizar el valor de los errores para observaciones pasadas (Naimy y Hayek, 2018).

En cuanto a su presencia en la literatura, este procedimiento es ampliamente utilizado para modelar la volatilidad de Bitcoin, su precio, y su relación con otras variables (Fang et al., 2019; Katsiampa, 2017; Shilov y Zubarev, 2021).

Por otro lado, el modelo vectorial autorregresivo es una metodología que no define inicialmente una variable dependiente y un grupo de variables independientes, sino que las trata a todas de manera simétrica, explicando sus valores futuros en base a sus propias cifras pasadas (Das, 2019). Su funcionamiento se basa en la utilización de vectores con múltiples variables y la presencia de valores retardados de las series de tiempo para analizar autocorrelación (Gujarati & Porter, 2008).

Una de sus mayores ventajas es que a menudo este modelo es complementado con las pruebas de causalidad de Granger, que permiten conocer la dirección de la capacidad predictiva entre dos variables al determinar cuán valiosa es la información que aporta una variable para predecir otra (Lin, 2021; Mai et al., 2018; Behzat, 2016).

Esta metodología ha participado de investigaciones sobre Bitcoin frente a la atención de los inversionistas (Lin, 2021; Nasir et al., 2019), la actividad en foros sobre Bitcoin (Mai et al., 2018), y el comportamiento de las personas en Twitter (Shen et al., 2019).

De esa forma concluye la revisión de los principales modelos observados en la literatura sobre la relación entre Bitcoin y factores externos a su mercado. Por lo que se da paso a seleccionar cuáles de ellos son los más apropiados para resolver la problemática planteada por la presente memoria.

4.8.2. Modelos a utilizar

Se selecciona la regresión lineal (simple o múltiple) como el modelo indicado para analizar la capacidad predictiva de las variables a considerar, porque con ello se logra estar en línea con las metodologías presentes en las publicaciones que motivan a estudiar si Bitcoin puede verse afectado por el material audiovisual (Lepori, 2015; Zhang et al., 2017; Oh y Kim, 2022; Bhattacharjee et al., 2017). De esa forma se apunta a obtener resultados que

permitan reconocer si dentro del mercado del Bitcoin ocurren efectos similares a los observados en otros mercados financieros.

Esta decisión también se fundamenta sobre la significancia de este estudio, que consiste en ser un primer antecedente sobre el análisis de la relación entre distintos tipos de material audiovisual y Bitcoin. Por lo tanto, se considera que lo más apropiado sería estudiar inicialmente si a partir de la evidencia disponible en la literatura es posible identificar un vínculo similar a los otros existentes, porque eso logra incluir a Bitcoin dentro de la discusión detectada en la literatura en torno al material audiovisual y las finanzas. Más aún cuando la problemática de investigación está centrada en que no existen antecedentes que incluyan a esta criptomoneda en tales conversaciones académicas.

Con posterioridad, frente a los resultados a obtener en este estudio (sean estos afirmativos o no), se contaría con un punto de conexión a la literatura para futuras investigaciones que busquen revelar más aspectos del vínculo analizado. Siendo ese un posible punto de partida para poder investigar a mayor profundidad si esta relación entre formatos de video y Bitcoin se ajusta de mejor manera a algún otro tipo de modelo, o si se obtendrían resultados más significativos un diseño de investigación alternativo.

Se descarta ocupar un modelo autorregresivo (GARCH o VAR) para esta investigación debido a que su uso significaría ir más allá de antecedentes que motivan a esta memoria, al incluir la autocorrelación para estudiar el vínculo entre material audiovisual y un mercado financiero. Una metodología con tales características entregaría resultados lejanos del conocimiento existente en la actualidad sobre esta materia, estando esto fuera de los objetivos de esta investigación, ya que se desea reconocer si los hallazgos encontrados en otros mercados financieros son similares en el caso de Bitcoin. Por consiguiente, se elige estudiar primero la posible relación lineal, sin tomar en cuenta la autoregresión. Dejando así

bases para que futuras investigaciones analicen otros enfoques de este posible vínculo, a través de diferentes metodologías.

Por último, también se hará uso de ambos índices de correlación cubiertos. Al ser prácticas que pueden ser utilizadas de manera complementaria a las regresiones lineales, que además aportan más información sobre la influencia que ejerce el material en formato de video frente a Bitcoin.

4.8.3. Modelos de regresión

Los modelos de regresión son aquellos que se basan en la predicción de una variable objetivo, conocida como dependiente, frente a una variable o grupo de ellas (variables independientes). La manera en que las variables independientes son operadas para obtener los valores de la variable dependiente depende del tipo de regresión que se esté utilizando. Por consiguiente, se hará revisión de aquellos que hayan sido observados en la literatura sobre Bitcoin frente a diferentes factores.

Se considera que la principal función de estos modelos es entregar información sobre la capacidad predictiva de las variables independientes frente a aquella que se quiere estimar.

4.8.3.1. Regresión Lineal Simple

Este tipo de regresión consiste en hallar una recta que represente los puntos de un gráfico de dispersión entre dos variables, para así determinar si existe capacidad predictiva enfocada a una de ellas y realizar predicciones (Schober et al., 2018). Para ello se plantea una ecuación lineal donde haya solo dos variables, una dependiente y otra independiente. De la forma:

$$y_1 = \beta_0 + \beta_1 \cdot x_1 + \varepsilon_1$$

Donde “y” es la variable dependiente o explicada, “x” es la variable independiente o explicativa, “ β ” son los coeficientes de regresión y “ ε ” es el error o residuo (la diferencia entre la estimación de “y” con su valor en cada observación).

Los coeficientes de regresión “ β ” indican cuál es el efecto que tiene la variable independiente sobre la variable dependiente, pudiendo ser positivos, negativos o cero. Mientras que el error es la representación de los factores que no son incluidos y afectan a la variable dependiente, en paralelo al efecto que tiene la única variable independiente de este modelo (Das, 2019).

En cuanto al procedimiento ocupado para su cálculo, la Regresión Lineal Simple suele basarse en procesos como el Método de Momentos, los Mínimos Cuadrados Ordinarios (MCO) o la Estimación por Máxima Verosimilitud (EMV) para obtener los coeficientes de regresión (Das, 2019).

Adicionalmente suele emplearse pruebas de hipótesis para verificar la significancia estadística de los coeficientes (β), en conjunto con pruebas de distribución t-student y la interpretación de valores P. Al igual que se aplica una prueba F de Fisher para determinar cuán representativa es la regresión con respecto a los valores observados de las variables, lo cual es indicado por el coeficiente de determinación (r^2).

No obstante, este modelo está limitado a representar únicamente relaciones simétricas (Jareño et al., 2020; Jareño et al., 2019) y lineales (Vetter & Schober, 2018). Implicando que no sería posible tomar en consideración ningún tipo de asimetría o función que represente la relación entre las variables.

También la regresión lineal simple no es capaz de identificar relaciones de causalidad entre las variables (Vetter & Schober, 2018).

Una de sus limitaciones es que no es capaz de indicar cuán fuerte es la relación entre las dos variables, sino que solo analiza la capacidad predictiva lineal y permite hacer proyecciones (Schober et al., 2018). De hecho, según el mismo autor, tal tipo de información es posible de obtener a partir de los coeficientes de correlación, que serán cubiertos en las siguientes secciones de esta memoria.

4.8.3.2. Regresión Lineal Múltiple

Teniendo un funcionamiento idéntico al modelo anterior, la Regresión Lineal Múltiple aporta una mayor capacidad explicativa para analizar la relación entre variables, ya que permite incluir un mayor número de ellas en la ecuación (Vetter & Schober, 2018). Esta posee la siguiente estructura:

$$y = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \dots + \beta_n \cdot x_n + \varepsilon$$

Siendo “y” la variable dependiente, las variables “x” las independientes (diferenciándose con subíndices) y “ε” el residuo.

De manera distinta a como ocurre en la Regresión Lineal Simple, donde los coeficientes β son llamados Coeficientes de Regresión, en este modelo los coeficientes β son llamados “Coeficientes de Regresión Parcial”. Esto se debe a que estos últimos representan el efecto que posee cada una de las variables explicativas sobre la variable explicada, cuando el resto de factores considerados en el modelo se mantienen constantes (Das, 2019). De hecho, la multiplicación entre la variación unitaria en el valor de una variable independiente y su coeficiente de regresión parcial entrega el cambio provocado por tal alteración en el promedio de la variable dependiente, considerando que el resto de variables se mantienen constantes (Gujarati & Porter, 2008).

Con motivo de determinar cuán bien un modelo regresivo se ajusta a los datos, se calcula la bondad de ajuste, que indica si la variable dependiente es explicada en mayor parte por las variables independientes o por el error (Das, 2019, p.76). Para ello se suele ocupar el Coeficiente de Determinación (R^2), que indica cuál es la proporción de la variación en “y” que es explicada por la regresión estimada (Das, 2019, p.78). Según el mismo autor citado, este coeficiente puede tomar valores entre cero y uno, siendo los valores más altos una señal de que la variable dependiente puede ser predicha con suficiente precisión con el modelo regresivo considerado. Mientras que cuando su magnitud se acerca a cero es posible interpretar que los cambios en las variables independientes no influyen en su contraparte dependiente, por lo tanto, la regresión no logra ser útil para realizar predicciones (p.78).

A pesar de la utilidad que el coeficiente de determinación (R^2) entrega para representar la bondad de ajuste, una de sus críticas es que corresponde a un valor que aumenta constantemente al agregar variables en un mismo modelo (Gujarati & Porter, 2008). Por esa razón se suele recomendar el uso de un coeficiente de determinación (R^2) ajustado, que toma en consideración dicho efecto y genera que R^2 solo aumente cuando las variables aumentan el poder de predicción del modelo (Shrestha & Bhatta, 2018).

Poder incluir un conjunto de variables explicativas puede ser utilizado para prevenir que los resultados obtenidos en una Regresión Lineal Simple sean en realidad debido a variables que no fueron incluidas en el modelo, tal como menciona (Lepori, 2015). Por ese motivo se suele recomendar incluir otras cuya relación con el factor dependiente ya haya sido confirmada, disminuyendo así el sesgo por variables omitidas (Stern, 2011). Para que de esa forma sea posible comparar si los coeficientes de regresión en los que se tiene interés permanecen constantes al adicionar elementos verificados a la ecuación.

Una desventaja de este modelo es la alta sensibilidad presentada frente a variables extremos dentro de las observaciones, cuando estos afectan que los datos puedan ser descritos por una distribución normal (Kurihara y Fukushima, 2017). Pero para hacer frente a esas eventualidades, tal como es utilizado en la última referencia mencionada, es posible aplicar la regresión utilizando Mínimos Cuadrados Robustos en vez de Mínimos Cuadrados Ordinarios.

Del mismo modo, esta metodología solo es capaz de describir relaciones de carácter simétrico (Jareño et al., 2020) y lineal (Uras et al., 2020), es decir, no detecta si las variables se relacionan de manera asimétrica o por medio de algún otro tipo de función que no sea una línea recta.

Otra problemática es la presencia de colinealidad entre las variables incluidas en el modelo (tanto dependientes como independientes), cuya aparición restaría veracidad a los resultados obtenidos con este modelo. Por esa razón es recomendado analizar la manera en que se correlacionan entre sí todas las variables, verificando que ellas no presenten valores por sobre 0,7 (Baird y Bieber, 2016; Jerdack, 2018).

También es necesario destacar que un modelo de Regresión Lineal Múltiple se limita a permitir establecer conclusiones sobre la predictibilidad de variables explicativas, y no así de causalidad entre ellas (Stern, 2011; Vetter & Schober, 2018)

4.8.3.3. Significancia estadística de los coeficientes

Tomando como referencia la explicación sobre regresiones lineales realizada por (Das, 2019) y (Gujarati & Porter, 2008), se procederá a revisar cómo se validan los coeficientes obtenidos al elaborar un modelo de regresión lineal.

Cuando el supuesto de la distribución normal de la variable error se cumple, es posible aplicar pruebas de hipótesis en torno a la significancia estadística que poseen los coeficientes β y el coeficiente de determinación. La significancia estadística puede ser interpretada como la cantidad de duda razonable que utiliza una prueba de hipótesis para tomar una decisión, ya que corresponde a la máxima probabilidad requerida para que sea posible rechazar una premisa, siendo este un límite definido por convención dentro de las investigaciones académicas. Los valores más usados son 0.05 (5%), 0.01 (1%) y 0.001 (0,1%), donde los porcentajes menores implican una mayor exigencia para que la hipótesis nula de la prueba sea rechazada.

Para llevar a cabo una prueba de hipótesis primero se define la hipótesis nula, que es un supuesto acerca del valor que posee un coeficiente, por ejemplo: " $H_0: \beta_1 = 0$ ". Luego se plantea una segunda hipótesis, que ocurre cuando la primera es negada, por ejemplo: " $H_1: \beta_1 \neq 0$ ". Si la segunda hipótesis involucra una desigualdad ($\neq$) la prueba de hipótesis considerará ambas colas de la distribución estadística, mientras que si posee símbolos de desigualdad como " $<$ " o " $>$ ", se utilizará solo una cola de la distribución.

Al analizar en detalle el ejemplo de hipótesis mencionado, la hipótesis nula: " $H_0: \beta_1 = 0$ " quiere decir que la variable independiente acompañando a dicho coeficiente no afecta a la variable dependiente. Por lo tanto, si esta premisa no es rechazada en la prueba de hipótesis se concluye que no es posible negar que tal coeficiente puede ser cero ni afirmar que es distinto de cero. Mientras que, si esta logra ser rechazada, se concluye que existe la suficiente evidencia para afirmar que la hipótesis " $H_1: \beta_1 \neq 0$ " es cierta, por lo tanto el coeficiente sí conlleva un efecto significativo entre una variable independiente y otra dependiente.

De manera clásica, una vez definidas ambas hipótesis, se realiza el cálculo de un estadístico de prueba, que indica cuán lejos está el valor estimado de un parámetro con respecto a su valor para la población. Su cálculo se basa en la distribución normal cuando la varianza de la población es conocida y se representa como “Estadístico z”. Mientras que cuando se desconoce dicho parámetro se utiliza la distribución normal en conjunto con una distribución de chi-cuadrado y se le reconoce como “Estadístico t”.

A partir de estos estadísticos se determina la probabilidad de que adquieran valores dentro de la hipótesis nula definida, que después es comparada con niveles de significancia estadística definidos por convención (0.05, 0.01 y 0.001). De esa forma, si la probabilidad obtenida con el estadístico es menor a un nivel de significancia escogido, se rechaza la hipótesis nula para ese nivel.

Sin embargo, otra alternativa para determinar si se rechaza la hipótesis nula es considerar los “Valores p”, que tienen como finalidad anular la subjetividad del método clásico sobre rechazar o no una hipótesis dependiendo de la significancia estadística escogida por cada investigador.

Los Valores p representan el menor nivel de significancia al que la hipótesis nula puede ser rechazada, en base a los valores de los estadísticos de prueba (z o t) y los grados de libertad (Gujarati & Porter, 2008). Su interpretación depende del investigador, dependiendo de si acepta o no la probabilidad asociada al valor p para rechazar la hipótesis. Pero aun así se puede comparar los valores p a los niveles de significancia, teniendo que si un valor p es menor a un nivel exigido significa que la hipótesis nula si puede ser rechazada para esa significancia estadística. Mientras que, si el valor p es mayor a un nivel de significancia, no se puede rechazar la hipótesis para ese nivel.

En softwares como Stata o R los valores p de cada coeficiente son entregados junto a los resultados de cada regresión, donde se asume que la hipótesis nula es del tipo “ $H_0: \beta_k = 0$ ”. Por consiguiente, estas cifras deben ser interpretadas como un análisis de dos colas de la distribución. En caso de querer realizar una prueba de hipótesis de una cola (involucrando símbolos como “<” o “>”) es posible dividir el valor p por dos e interpretar ese valor (Das, 2019).

Con tales antecedentes se poseen las herramientas necesarias para poder determinar la validez de los coeficientes obtenidos en una regresión lineal.

4.8.3.4. Significancia estadística del modelo regresivo

Adicionalmente se considerará realizar una evaluación de la significancia que posee la recta estimada en cada modelo regresivo, siguiendo como referencia lo expuesto por (Das, 2019) y (Gujarati & Porter, 2008). Con esos propósitos se recurrirá a la “Prueba F”, que consiste en ejecutar una prueba de hipótesis donde todos los coeficientes propios de las variables independientes son iguales a cero, es decir:

$$“H_0: \beta_1 = \beta_2 = \beta_3 = \dots = \beta_k = 0”.$$

Tal afirmación implica que no existe relación entre el conjunto de variables independientes incluidas en la regresión y la variable dependiente (Das, 2019). Lo cual se diferencia con lo evaluado en el análisis individual de cada coeficiente, ya que la Prueba F juzga la significancia de todos los coeficientes a modo de grupo.

De manera similar a las pruebas realizadas a coeficientes individuales en la sección anterior, la metodología tradicional para implementar la Prueba F se basa en construir un “Estadístico F” cuyo valor ayuda a decidir si la hipótesis nula que incluye a todos los coeficientes puede ser rechazada o no. Para lo cual se compara su magnitud con un punto de

referencia que depende del nivel de significancia considerado para la prueba. Si este supera el valor límite se rechaza la hipótesis nula, de lo contrario no es posible rechazar dicha afirmación.

Sin embargo, para la Prueba F también es posible utilizar los Valores p, que conservan su característica de ser el menor nivel de significancia para el cual es posible rechazar la hipótesis nula. Por consiguiente, la hipótesis nula puede ser rechazada dependiendo del Valor P obtenido junto al propio criterio de decisión del investigador o un nivel significancia escogido. Si un Valor P es menor a un nivel de significancia se puede rechazar la hipótesis nula bajo esos estándares.

Tales cifras son entregadas por Stata junto a los resultados de cada regresión. Se asume una hipótesis nula del tipo: “ $H_0: \beta_1 = \beta_2 = \beta_3 = \dots = \beta_k = 0$ ”, es decir, que todos los coeficientes son iguales a cero, por lo que el conjunto de variables explicativas no logra tener efecto sobre la variable explicada. Lo cual implica una hipótesis alternativa de la forma: “ $H_1 =$ no todos los coeficientes β son iguales a cero” (Gujarati & Porter, 2008). Por lo tanto, la interpretación es idéntica a la explicada en el párrafo anterior sobre los Valores p.

Finalmente, se concluye que se han revisado todas las herramientas necesarias para evaluar la significancia estadística de la recta entregada por un modelo regresivo. Dando así paso al siguiente modelo a utilizar.

4.8.3.5. Verificación de supuestos

Al realizar la estimación de un modelo de regresión lineal múltiple es necesario verificar si los supuestos que permiten obtener resultados adecuados con esta metodología se cumplen o no. De esa forma es posible identificar si las estimaciones involucradas poseen las

propiedades adecuadas para el modelo, como también reconocer si el fallo de algún supuesto requiere tener ciertas consideraciones en cuanto a los resultados a obtener (limitaciones).

Por esa razón en los siguientes apartados se resumen los ocho principales supuestos requeridos para construir un modelo de regresión lineal múltiple con datos en formato de series de tiempo. Estos han sido sintetizados a partir de las obras de (Gujarati & Porter, 2008).

4.8.3.5.1. Linealidad de los parámetros

Se define un modelo regresivo de la forma:

$$y_t = \beta_0 + \beta_1 \cdot x_{1t} + \dots + \beta_k \cdot x_{kt} + u_t$$

Donde “ y_t ” es la variable explicada, “ t ” el número de observaciones o periodos de tiempo considerados, “ x ” las variables explicativas, “ k ” la cantidad de variables explicativas y “ u_t ” son los errores o residuos de la regresión. Tal expresión es considerada un modelo lineal en cuanto a sus coeficientes β y es aquella que se desea estimar con la Regresión Lineal Múltiple.

Para poder determinar si es adecuada la formulación de este tipo de modelo para las variables a considerar es necesario estudiar la relación entre cada variable independiente “ x ” frente a la variable dependiente “ y ”. Esto es realizado a partir de un gráfico de dispersión entre las distintas combinaciones de dos variables, donde se busca obtener que cada dupla describa un gráfico donde los puntos sigan una tendencia cercana a una línea recta. Con ello es posible señalar que las variables posiblemente se relacionen de manera lineal, por lo tanto, son adecuadas para ser incluidos en este tipo de regresión.

Adicionalmente se requiere tener antecedentes que justifiquen la razón por la que se quiere estudiar la relación lineal entre la variable, para así tener una base empírica que conduzca a pensar que las variables pueden estar vinculadas de esa manera, tal como es recomendado por (Gujarati & Porter, 2008, p.59)

De no cumplirse que la variable dependiente se relaciona linealmente con una variable independiente no es recomendable incluirlas a ambas en un mismo modelo de este tipo. No obstante, es posible aplicar transformaciones como el cálculo de las primeras diferencias o la utilización de logaritmos naturales para poder lograr que las variables consideradas se relacionen linealmente y que así puedan ser incluidas en el modelo. Aunque esto puede cambiar la interpretación de los resultados, por lo que habría que tener en mente los cambios involucrados. Por ejemplo, en los modelos con logaritmos naturales la variación en las variables independientes puede tener una interpretación porcentual en la variable dependiente (Gujarati & Porter, 2008, p.162)

4.8.3.5.2. Valor esperado de los residuos igual a cero

Este supuesto sugiere que los errores u_t correspondientes a cada observación de las series de tiempo deben tener una esperanza igual a cero. Lo cual es necesario para poder afirmar que los factores no incluidos en el modelo como variables explicativas, y que por ende son representados por los residuos, no afectan sistemáticamente a la variable dependiente (Gujarati & Porter, 2008, p.64). Lo anterior se ve reflejado matemáticamente por la siguiente expresión:

$$E(u_t) = 0 \quad , t = 1, 2, \dots, n.$$

De no cumplirse este enunciado la estimación del coeficiente β_0 sería sesgada (Gujarati & Porter, 2008, p.317), es decir, el valor esperado de este coeficiente no correspondería al valor real en la población (Gujarati & Porter, 2008, p.72), o bien del proceso estocástico implícito descrito por las variables, debido a que tal es el nombre que semejante al modelo poblacional cuando se utilizan datos en serie de tiempo (Gujarati & Porter, 2008, p.740).

Sin embargo, esa consecuencia carece de relevancia para este estudio, debido a que el coeficiente β_0 no posee una interpretación práctica al analizar la relación entre las variables, a diferencia del resto de coeficientes β (Gujarati & Porter, 2008, p.317). Por lo tanto, el cumplimiento de este supuesto busca mejorar la capacidad predictiva del modelo, pero esto no es un impedimento para poder obtener conclusiones significativas del resto de estimadores.

4.8.3.5.3. Condición de homocedasticidad en los residuos

Adicionalmente se requiere que los residuos posean una varianza de valor finito y constante a lo largo de todas las observaciones, lo cual recibe el nombre de homocedasticidad (Gujarati & Porter, 2008, p.65) y, según (Brooks, 2014, p.91), se le representa por la siguiente expresión:

$$Var(u_t) = \sigma^2 < \infty, t = 1, 2, \dots, n.$$

En caso de no cumplirse tal igualdad se está en presencia de Heterocedasticidad, es decir, la varianza puede tomar valores diferentes entre las distintas observaciones.

La Heterocedasticidad puede ser detectada a través de las siguientes tres pruebas, que serán sintetizadas a continuación:

- Prueba de White

Consiste en estimar el modelo lineal al que se quiere aplicar la prueba, calcular sus residuos y crear una regresión auxiliar con el cuadrado de sus valores. Con ello se analiza si ellos están relacionados con alguna de las variables independientes del modelo a través de pruebas F o multiplicadores de Lagrange (Brooks, 2014, p.184). Su hipótesis nula señala que existe homoscedasticidad, es decir, que dicha relación estudiada es inexistente o muy débil. Esta puede ser rechazada según los valores p o los valores del estadístico formado para la prueba.

Este último es construido al multiplicar el coeficiente de determinación del modelo auxiliar con la cantidad de variables independientes consideradas, que se distribuye asintóticamente según la distribución chi-cuadrado (Gujarati & Porter, 2008, p.387).

Esta prueba no requiere que los residuos asemejen una distribución normal (Gujarati & Porter, 2008, p.386).

- Prueba de Breusch-Pagan-Godfrey:

Su procedimiento toma en consideración elaborar una regresión auxiliar de la varianza de los residuos, frente a las variables independientes tomadas en cuenta en el modelo lineal original. Al igual que en la prueba anterior, se establece una hipótesis nula que indica la inexistencia de relación entre las variables del modelo auxiliar, que equivale a la presencia de homocedasticidad (Gujarati & Porter, 2008, p.385). En base a esto también es posible construir un estadístico que sigue la distribución Chi-cuadrado, con lo cual es posible realizar las respectivas pruebas de hipótesis.

Sin embargo, una limitante de este recurso es que los residuos deben tener una distribución normal (Gujarati & Porter, 2008, p.385).

- Prueba de Koenker-Bassett

Por último, la prueba de Koenker-Bassett confecciona un modelo regresivo para estimar los residuos al cuadrado, teniendo como variables explicativas al cuadrado de cada una de las variables independientes del modelo original (Gujarati & Porter, 2008, p.388). Luego se establece una hipótesis nula que asume homocedasticidad, que puede ser puesta a prueba por medio de estadísticos t o F .

Al igual que la prueba de White, este procedimiento no requiere la distribución normal de los errores (Gujarati & Porter, 2008, p.389).

Tener residuos que presentan Heterocedasticidad provoca que en una regresión lineal los coeficientes continúen siendo consistentes e insesgados, pero estos ya no tendrán la mínima varianza entre el resto de los estimadores sin sesgo, dejando de ser los mejores estimadores lineales insesgados (Brooks, 2014, p.183) y pasando a ser considerados como ineficientes (Gujarati & Porter, 2008, p.371-374). Esto también repercute en los errores estándar, que tendrán valores no apropiados y generarán que las pruebas de hipótesis a realizar tampoco lo sean (Brooks, 2014, p.185).

Dentro de las posibles soluciones a la heterocedasticidad se puede mencionar la aplicación de logaritmos naturales (Brooks, 2014, p.186) o calcular la varianza y errores estándar robustos consistentes con la heterocedasticidad de White (Brooks, 2014, p.186). Aunque esta última opción suele disponer errores estándar más altos en comparación a los obtenidos usando mínimos cuadrados ordinarios, por lo que es más difícil rechazar una hipótesis nula al realizar ese tipo de pruebas (Gujarati & Porter, 2008, p.391).

4.8.3.5.4. No hay autocorrelación entre residuos

El presente supuesto estipula que la covarianza entre residuos de dos periodos de tiempo diferentes debe ser cero, es decir, no están correlacionados (Brooks, 2014, p.188). De lo contrario, según la misma referencia, se dice que los residuos están autocorrelacionados o correlacionados en serie.

Verificar esta suposición es muy relevante al trabajar con datos de series de tiempo, ya que se considera común que haya interdependencia entre sus observaciones (Gujarati & Porter, 2008, p.414). Más aún cuando estas series de tiempo no son estacionarias (Gujarati & Porter, 2008, p.418).

Al no corregir los efectos de la autocorrelación los coeficientes de regresión seguirán siendo insesgados y consistentes, pero ya no sean eficientes, es decir, no tendrán la mínima

varianza entre estimadores lineales (Brooks, 2014, p.199; Gujarati & Porter, 2008, p.423). También pueden verse modificados los errores estándar, generando que las pruebas de hipótesis no tengan una precisión aceptable. Junto a la posibilidad de que se produzca una sobrestimación del coeficiente de determinación para el modelo (Gujarati & Porter, 2008, p.423).

La autocorrelación de los residuos puede ser detectada a través de las siguientes pruebas:

- Prueba de Durbin-Watson:

Es capaz de percibir la autocorrelación de primer orden, en otras palabras, solo analiza si hay correlación entre los residuos de un periodo de tiempo con la observación anterior a ella (Brooks, 2014, p.194). Su proceso es crear una regresión estimando los residuos con periodo t , posicionando a los errores del periodo anterior como única variable explicativa, para luego calcular el estadístico de Durbin Watson (Brooks, 2014, p.194), cuya fórmula es:

$$d = \frac{\sum \hat{u}_t^2 + \sum \hat{u}_{t-1}^2 - 2 \sum \hat{u}_t \hat{u}_{t-1}}{\sum \hat{u}_t^2}$$

También es posible estimar su valor en base al coeficiente β obtenido, que representa la correlación entre ambos residuos y suele ser representado por el símbolo “ ρ ” (Gujarati & Porter, 2008, p.436):

$$d \approx 2(1 - \rho)$$

El estadístico de Durbin Watson adquiere cifras entre 0 y 4, donde un valor igual a cero representa que hay autocorrelación positiva perfecta, mientras que al ser igual a dos indica que no hay autocorrelación, y al ser cuatro existe autocorrelación negativa perfecta (Brooks, 2014, p.196). La interpretación de los valores intermedios solía ser realizada por medio de una tabla y criterios adicionales que indican cuándo rechazar la hipótesis nula

(Gujarati & Porter, 2008, p.435), sin embargo, en la actualidad los softwares estadísticos son capaces de calcular los valores p del estadístico. Esta prueba asigna como hipótesis nula la no existencia de autocorrelación, por lo tanto, si esta es rechazada se concluye que hay evidencia suficiente para decir que sí hay autocorrelación (Brooks, 2014, p.196).

La prueba de Durbin Watson, según (Gujarati & Porter, 2008, p.434) requiere los siguientes supuestos:

- La regresión debe tener un intercepto.
- Las variables X no deben ser estocásticas.
- Los residuos deben ser generados por el esquema autorregresivo de primer orden, es decir, con una regresión auxiliar que estime los residuos de un periodo " t " en función de aquellos correspondientes al instante " $t-1$ ". O sea, no es posible utilizar esta prueba para detectar autocorrelación de segundo orden.
- Los residuos siguen una distribución normal.
- El modelo regresivo no incluye valores anteriores de la variable " Y ".
- No debe haber observaciones con valores vacíos en los datos

Tomando en cuenta tales supuestos cabe mencionar que la prueba de Durbin Watson suele ser difícil de aplicar en datos de tipo serie de tiempo, específicamente por aquel que menciona la característica no estocástica de las variables X (Gujarati & Porter, 2008, p.438).

Por esa razón también se toma en cuenta una prueba adicional, llamada prueba de Breusch-Godfrey.

- Prueba de Breusch-Godfrey

A diferencia del proceso descrito anteriormente, la actual prueba es capaz de analizar la relación entre los residuos de manera general, pudiendo tomar en cuenta residuos con más de un periodo de distancia con respecto a un instante “t” (Brooks, 2014, p.197).

Su funcionamiento se basa en crear una regresión de los residuos de un periodo “t”, siendo las variables explicativas los residuos del resto de periodos anteriores a considerar (Brooks, 2014, p.198). De manera posterior, según la misma fuente, se utilizan los coeficientes estimados para formar el estadístico de la prueba, que se asemeja a la distribución chi-cuadrado y es utilizado para realizar las pruebas de hipótesis.

La hipótesis nula de este proceso supone que no hay autocorrelación con ninguno de los residuos considerados, por lo tanto, si esta es rechazada se concluye que sí existe autocorrelación (Brooks, 2014, p.198).

La principal ventaja de esta prueba es que no requiere de los supuestos exigidos por la prueba de Durbin Watson (Brooks, 2014, p.198), generando que sea una vía de acción alternativa en caso de no alcanzar dichas suposiciones.

En caso de hallar que sí existe autocorrelación en el modelo regresivo existen las siguientes alternativas para poder estimadores con las mejores propiedades posibles:

- Uso de Mínimos Cuadrados Generalizados

Considerando un modelo regresivo como el siguiente:

$$Y_t = \beta_1 + \beta_2 X_t + u_t$$

Al igual que la misma ecuación para el periodo anterior:

$$Y_{t-1} = \beta_1 + \beta_2 X_{t-1} + u_{t-1}$$

Si se multiplica el segundo modelo por el coeficiente de autocorrelación ρ y este se sustraído al primer modelo se obtiene:

$$(Y_t - \rho Y_{t-1}) = \beta_1 (1 - \rho) + \beta_2 (X_t - \rho X_{t-1}) + (u_t - \rho u_{t-1})$$

Lo cual puede ser considerado como un nuevo modelo de regresión lineal que puede ser elaborado por mínimos cuadrados ordinarios:

$$Y^\circ_t = \beta^\circ_1 + \beta^\circ_2 X^\circ_t + \varepsilon_t$$

Con él es posible hallar el coeficiente β°_2 , que posee las propiedades idóneas de la utilización de mínimos cuadrados ordinarios. Además de representar la relación entre las variables incluidas en el modelo.

No obstante, para poder seguir tal procedimiento es necesario conocer previamente el valor del coeficiente de autocorrelación ρ . Este puede ser estimado a partir del estadístico de Durbin Watson (d) cuando se trabaja con muestras grandes de datos (Gujarati & Porter, 2008, p.445), por medio de la siguiente expresión:

$$\rho \approx 1 - \frac{d}{2}$$

O bien es posible estimar una regresión de los residuos para un periodo “t”, frente a aquellos pertenecientes a un instante “t-1”:

$$\hat{u}_t = \rho * \hat{u}_{t-1} + v_t$$

Donde el coeficiente de regresión equivale al valor de ρ .

- Método de Newey West para corregir errores estándar

Corresponde a una extensión del método de White para tratar los efectos de la Heterocedasticidad y se caracteriza por generar nuevos errores estándar que reciben el nombre “Errores estándar HAC (Heteroscedasticity and Autocorrelation Consistent)” o “Errores estándar de Newey-West” (Gujarati & Porter, 2008, p.447).

Su uso está destinado únicamente a muestras de datos grandes y puede ser ejecutado a través de los principales softwares estadísticos (Gujarati & Porter, 2008, p.448). En donde cabe mencionar que se considera a una muestra grande a aquellas bases de datos con más de 50 observaciones (Gujarati & Porter, 2008, p.448).

A partir de sus errores estándar corregidos es posible afirmar que se controlan los efectos de la heterocedasticidad y autocorrelación, por lo que se puede interpretar los coeficientes de regresión y realizar las pruebas de hipótesis de manera normal.

Usualmente se recomienda utilizar esta solución en escenarios donde hay presencia de heterocedasticidad y autocorrelación simultáneamente, puesto que esta metodología funciona para ambos factores. Especialmente cuando corregir primero el primer concepto y luego el segundo (siendo ese el orden recomendado) suele ser de alta complejidad (Gujarati & Porter, 2008, p.450).

4.8.3.5.5. Variables independientes no estocásticas

Al crear un modelo de regresión lineal es necesario tener en cuenta si las variables independientes serán tratadas como valores fijos o como un proceso estocástico. Cuando estas son consideradas valores fijos es posible aplicar el resto de supuestos para este tipo de modelo y ejecutar la metodología. Sin embargo, cuando las variables son parte de un proceso estocástico es necesario establecer el supuesto que propone la necesidad de los residuos de ser independientes de cada una de las variables explicativas, es decir, su covarianza debe ser igual a cero:

$$Cov(u_i, x_i) = 0$$

El cumplimiento de esa igualdad permite poder trabajar con los datos de manera idéntica a cuando las variables independientes son fijas (Gujarati & Porter, 2008, p.317). Mientras que, por el contrario, cuando esto no se cumple los cambios en la variable

dependiente pueden ser producidos por los residuos en vez de las variables independientes, causando que los estimadores pierdan su consistencia (Brooks, 2014, p.209).

En ocasiones la obtención de datos a partir de una fuente secundaria puede facultar a que los datos puedan ser asumidos directamente como fijos, debido a que el investigador no se encuentra midiendo directamente el proceso estocástico y sus datos se encuentran condicionados a la medición realizada por la fuente (Gujarati & Porter, 2008, p.316). Teniendo eso en consideración las propiedades de los estimadores obtenidos con mínimos cuadrados ordinarios en cuanto a eficiencia, ausencia de sesgo y consistencia se mantienen (Gujarati & Porter, 2008, p.511).

4.8.3.5.6. Residuos con distribución normal

El actual supuesto se define como la necesidad de que los residuos de la regresión posean una distribución normal:

$$u_t \sim N(0, \sigma^2)$$

Su cumplimiento resulta fundamental para poder realizar inferencias estadísticas, específicamente pruebas de hipótesis. Debido a que muchas veces la distribución normal de los residuos es requisito para que estadísticos de prueba se comporten de acuerdo a una determinada distribución estadística.

Para poder diagnosticar si los residuos presentan este tipo de comportamiento se toman en cuenta las siguientes posibilidades:

- Prueba de Jarque-Bera:

Este test verifica si la asimetría y curtosis de los residuos se asemejan a los valores que caracterizan a las distribuciones normales (Brooks, 2014, p.209). Lo hace por medio del estadístico de Jarque-Bera, cuya fórmula es la siguiente:

$$W = T \left[\frac{b_1^2}{6} + \frac{(b_2 - 3)^2}{24} \right]$$

Siendo b_1 el coeficiente de asimetría, b_2 el coeficiente de curtosis y T el número de observaciones.

Dicho estadístico sigue asintóticamente (en muestras grandes) la distribución chi-cuadrado. Su hipótesis nula señala que los residuos poseen una distribución simétrica y mesocúrtica (ambas características de una distribución normal) (Brooks, 2014, p.210), la cual puede ser rechazada o no dependiendo de los valores p de la prueba.

En caso de que las pruebas anteriores indiquen que los residuos no se distribuyen de manera normal, se dispone de las siguientes soluciones:

- Recurrir al teorema del límite central

En tamaños de muestra grande se estima que la violación del supuesto de normalidad de los residuos no produce consecuencias adversas para los modelos de regresión lineal por mínimos cuadrados ordinarios (Brooks, 2014, p.210; Gujarati & Porter, 2008, p.509), conduciendo a que sus estimadores sigan siendo estimadores lineales insesgados y óptimos (Gujarati & Porter, 2008, p.509). Esto se debe a que se convoca al Teorema de Límite Central para señalar que “los estadísticos de las pruebas de hipótesis seguirán asintóticamente la distribución probabilística apropiada, incluso si los errores no se distribuyen de forma

normal” (Brooks, 2014, p.210). Por consiguiente, en esos casos es posible continuar con la estimación del modelo, aun cuando las pruebas de diagnóstico no confirmen el supuesto.

- Utilizar variables ficticias frente a valores extremos

En ocasiones no es posible obtener una distribución normal debido a la presencia de valores extremos en los residuos, principalmente porque se ubican en las colas de la curva y aumentan la curtosis (Brooks, 2014, p.211).

El procedimiento consiste en crear una variable ficticia con valor igual a uno solo en aquella observación donde hay presencia de un valor extremo, mientras que en el resto de periodos se le asigna valor cero (Brooks, 2014, p.212). Esto formará que el residuo de dicho punto deje de ser un valor extremo, mejorando así la distribución de esta variable.

En cuanto a la forma de identificar cuáles son dichos valores extremos, ellos pueden ser observados al crear un gráfico de dispersión entre la variable dependiente y cada variable independiente cuando se trata de una regresión simple. Mientras que al haber múltiples variables explicativas es mejor optar por graficar los residuos a lo largo del tiempo (Brooks, 2014, p.213).

4.8.3.5.7. No existe multicolinealidad

La no existencia de colinealidad se refiere a que las variables explicativas no deben estar correlacionadas entre sí, es decir, tienen que ser ortogonales (Brooks, 2014, p.217). De esa forma se logra que la incorporación de nuevas variables al modelo o la sustracción de una de ellas no modifique el efecto que tiene cada una frente a la variable explicada.

No obstante, en situaciones prácticas es muy poco usual que las variables presenten una correlación igual a cero con exactitud (Brooks, 2014, p.217), por ende, no es posible seguir la teoría en su totalidad al buscar variables sin correlación alguna. Es por esa razón que suele ser aceptado tener variables que presenten una ligera correlación entre ellas, ya que

esto no genera grandes pérdidas de precisión en el modelo (Brooks, 2014, p.217). Pero sí se hace la distinción con aquellas situaciones donde el nivel de asociación de las variables independientes es alto, ya que esto sí genera inconvenientes al estimar un modelo de regresión lineal. Tales tipos de relación se conocen como colinealidad (Brooks, 2014, p.217), debido a que la relación entre ambas variables se asemeja mucho a la de una línea recta.

Cuando dos variables independientes poseen entre sí una correlación igual a uno se le considera “Colinealidad perfecta”, mientras que en aquellas situaciones donde el coeficiente es alto pero distinto de uno se le llama “Casi-colinealidad” (Brooks, 2014, p.217). Cabe decir que un coeficiente de correlación es considerado alto cuando es mayor o igual a 0,8 (Gujarati & Porter, 2008, p.338).

Al estimar un modelo de regresión lineal simple en presencia de colinealidad se puede generar que el coeficiente de determinación R^2 sea muy alto, sugiriendo un buen desempeño del modelo, pero los errores estándar de los coeficientes serán elevados, generando que los resultados individuales de las variables no sean significativos (Brooks, 2014, p.218). Tal tipo de resultados indicarían que la colinealidad es un problema muy relevante en la regresión, por lo que habría que tomar medidas para corregirlo. Aunque no siempre la colinealidad es capaz de generar esas consecuencias, por lo que se suele recomendar evaluar la coherencia de los resultados (en cuanto a R^2 y errores estándar) antes de afirmar que la colinealidad influye gravemente en el modelo (Gujarati & Porter, 2008, p.347). En otras palabras, hay ocasiones donde la regresión entrega resultados apropiados a pesar de la presencia de colinealidad entre variables.

A pesar de esos inconvenientes, los coeficientes de una regresión lineal por mínimos cuadrados ordinarios conservan sus propiedades, por lo que continúan siendo estimadores lineales insesgados y óptimos (Gujarati & Porter, 2008, p.323).

La colinealidad puede ser detectada a través del siguiente proceso:

- Matriz de correlación

Esta herramienta es una representación gráfica en forma de matriz, que contiene los cálculos de coeficientes de correlación entre todas las posibles combinaciones de parejas que se pueden formar con las variables independientes. A modo de ejemplo, se presenta la siguiente matriz de correlación:

corr	x_2	x_3	x_4
x_2	—	0.2	<u>0.8</u>
x_3	0.2	—	0.3
x_4	<u>0.8</u>	0.3	—

Figura 2: Ejemplo de matriz de correlación

Fuente: (Brooks, 2014, p.218)

Este ordenamiento de los valores permite reconocer qué variables poseen niveles altos de correlación, que en el caso del ejemplo es entre las variables x_2 y x_4 .

Sin embargo, una de sus limitaciones es que no es capaz de detectar correlaciones que comprendan la combinación lineal de dos variables para formar una tercera, por ejemplo:

“ $x_4 = x_2 + x_3$ ” (Brooks, 2014, p.218).

Posteriormente, en cuanto a la manera de poder solucionar los inconvenientes generados por la colinealidad, se aprecian las siguientes técnicas:

- Remover una de las variables con casi-colinealidad

En caso de ser coherente con la teoría que respalda el modelo regresivo, una opción es sustraer a una de las variables involucradas de este, para así evitar los efectos de la casi-colinealidad. No obstante, esta alternativa corre el riesgo de aumentar el error por variable omitida (Brooks, 2014, p.219).

- No realizar correcciones al modelo y concluir

La presencia de colinealidad no corresponde a un problema propio del modelo de regresión lineal, sino que es una relación entre los datos provenientes de las fuentes de información (Gujarati & Porter, 2008, p.342). Por ese motivo, cuando la colinealidad genera resultados con un R^2 y errores estándar incoherentes se puede recomendar concluir que los datos son inapropiados para poder entregar estimadores precisos (Gujarati & Porter, 2008, p.342).

Debido a que una colinealidad fuerte se considera un problema de los datos, una posible solución es aumentar la cantidad de observaciones a considerar o tomar en cuenta nuevas variables en reemplazo de aquellas que no funcionaron (Gujarati & Porter, 2008, p.345).

- Transformar las variables

En ocasiones puede que optar por el cálculo de las primeras diferencias para las variables logre evitar los altos niveles de colinealidad. (Gujarati & Porter, 2008, p.344).

4.8.3.5.8. Series de tiempo estacionarias

Al trabajar con datos del tipo serie de tiempo es necesario incorporar un último supuesto al conjunto de premisas necesarias para la construcción de una regresión lineal múltiple, acerca de la estacionaridad de cada serie de tiempo (correspondientes a las variables a considerar dentro de la regresión).

Una serie de tiempo es estacionaria cuando su promedio y varianza no varían sistemáticamente a lo largo del tiempo (Gujarati & Porter, 2008, p.740). Esta característica cobra relevancia debido a que estimar una regresión a partir de series de tiempo no estacionarias podría llevar a presenciar el fenómeno de “Regresión Espuria” o “Regresión sin sentido” (Gujarati & Porter, 2008, p.748). La cual puede definirse como una regresión

con coeficiente de determinación (R^2) y coeficientes de estimación significativos, pero que se forma a partir de variables que no están correlacionadas realmente (lo cual puede ser verificado al llevarlas a una forma estacionaria). Por ende, las conclusiones que sugiere este tipo de regresión carece de veracidad o coherencia, razón por la que debe ser evadida o analizada con precaución. Especialmente cuando se toma en consideración que en estos casos los estadísticos t no siguen la distribución t -student, afectando así la validez de las pruebas de hipótesis (Gujarati & Porter, 2008, p.748).

También existe la posibilidad de estimar si una regresión presenta resultados espurios. Esto se lleva a cabo al comparar el coeficiente de determinación (R^2) con el estadístico de Durbin-Watson (d). Si el valor del primero (R^2) es mayor al segundo (d), entonces se cuenta con un buen indicio para sospechar que la regresión estimada es espuria (Gujarati & Porter, 2008, p.748).

Es por esto que, antes de formular una regresión lineal múltiple con series de tiempo, se requiere verificar que todas ellas sean estacionarias. De esa forma se evita obtener una regresión espuria, pudiendo así llegar a conclusiones que representen correctamente aquella información contenida en los datos.

Dentro de las distintas maneras por las que se puede detectar si las series de tiempo son estacionarias se destacan las siguientes pruebas:

- Prueba de Dickey-Fuller Aumentada

Consiste en hacer una regresión de la variable dependiente Y_t , en función de sus valores anteriores, de la forma (Gujarati & Porter, 2008, p.757):

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^m \alpha_i \Delta Y_{t-1} + \varepsilon_t$$

A partir de tal expresión se pone a prueba la hipótesis nula que supone al coeficiente δ como igual a cero, o sea, que la serie de tiempo no es estacionaria (Gujarati & Porter, 2008, p.757). Por consiguiente, si esta hipótesis es rechazada existe suficiente evidencia para afirmar que la serie de tiempo es estacionaria. Tal prueba de hipótesis puede ser llevada a cabo en base a los valores de un estadístico Tau (τ) (Gujarati & Porter, 2008, p.755) o bien por los valores p asociados.

- Prueba de Phillips-Perron (PP)

Es un proceso no paramétrico que permite identificar si una serie de tiempo es estacionaria o no, siendo capaz de entregar resultados adecuados cuando se está en presencia de heterocedasticidad o autocorrelación (Gujarati & Porter, 2008, p.758; Shrestha & Bhatta, 2018). Al igual que se le considera una alternativa más apropiada cuando se trabaja con muestras grandes (Shrestha & Bhatta, 2018).

Su funcionamiento se basa en la siguiente expresión:

$$\Delta Y_t = \pi y_{t-1} + \beta_i D_{t-i} + e_i$$

Donde D_{t-i} es un componente determinístico de tendencia (Shrestha & Bhatta, 2018).

La hipótesis nula de la prueba de Phillips-Perron (PP) es que el coeficiente π es igual a cero (Shrestha & Bhatta, 2018), siendo esto equivalente a decir que la serie de tiempo no es estacionaria. Si esta logra ser rechazada entonces se concluye que la serie de tiempo si presenta estacionaridad.

En caso de que las pruebas anteriores señalen que alguna de las series de tiempo contempladas no es estacionarias entonces se puede intentar remediar esta situación a través del cálculo de las primeras diferencias.

Tal como ha sido mencionado anteriormente, restar cada observación con respecto a su valor anterior puede crear una nueva variable estacionaria a partir de una serie de tiempo no estacionaria (Gujarati & Porter, 2008, p.760). Para poder verificar aquello se debe obtener la nueva variable y realizar las pruebas pertinentes que puedan determinar si es estacionaria o no.

Adicionalmente esta solución puede ser complementada con el uso de logaritmos naturales (Gujarati & Porter, 2008, p.762).

Con todas las referencias descritas a lo largo de estos ocho supuestos se logra dimensionar la importancia que tiene verificar cada uno de ellos, al igual que se reconocen las posibles líneas de acción para incorporar estos procedimientos a la metodología del presente estudio.

De esa manera concluye la revisión del modelo de regresión lineal, por lo que se continúa con las metodologías relativas a la correlación.

4.8.4. Metodologías para analizar Correlación

La correlación explora el grado de relación asociativa existente entre dos variables, que es medida a través de los coeficientes de correlación (Senthilnathan, 2019). De esa forma, “cuando dos variables están correlacionadas, la variación en la magnitud de una variable se asocia a un cambio en los valores de la segunda, ya sea en la misma dirección (correlación positiva) o en la opuesta (correlación negativa)” (Schober et al., 2018).

Entre las alternativas disponibles para ello se analizarán dos principales opciones: el Índice de Correlación de Pearson y el Índice de Correlación de Spearman.

4.8.4.1. Coeficiente de correlación de Pearson

El índice de Correlación de Pearson fue descrito por primera vez por Karl Pearson en 1896 (Hauke & Kossowski, 2011). Su cálculo permite reflejar la correlación lineal entre dos variables (Xu & Deng, 2018). Se le representa con una letra “R” y la expresión matemática para su cálculo es la siguiente:

$$R = \frac{n * (\sum x * y) - (\sum x) * (\sum y)}{\sqrt{n * (\sum x^2) - (\sum x)^2} * \sqrt{n * (\sum y^2) - (\sum y)^2}}$$

Donde “n” es el número de observaciones de la muestra, “x” son los valores de la primera variable e “y” las medidas de la segunda variable (Senthilnathan, 2019).

Es ampliamente utilizado en la literatura sobre Bitcoin (Kaminski, 2014; Yatsyk, 2020; Abraham et al., 2018) y aquella relacionada con material en formato de video (Zhou et al., 2021; Kalatzis et al., 2018). Ya que suele ser incluido en análisis iniciales para identificar la asociación lineal entre el comportamiento de dos variables.

En cuanto a sus valores y respectivas interpretaciones, su magnitud se ubica entre cero y uno, pudiendo tener un signo positivo o negativo. Si es positivo, significa que ambas variables tienden a aumentar o disminuir en los mismos tramos. Por el contrario, si es negativo, al momento que una aumenta la otra disminuye en su valor. Mientras que al ser cero significa que no existe asociación lineal entre las variables.

Adicionalmente se puede analizar la intensidad de esta relación, dependiendo del valor numérico indicado por el coeficiente. Para esto se decide seguir la escala propuesta por (Senthilnathan, 2019), que está construida específicamente para estudios dedicados a las ciencias sociales (p.4), siendo esta el área donde se ubica esta memoria. Su contenido es el siguiente:

- Valor absoluto 0,00: No existe correlación
- Valor absoluto entre 0,00 y 0,20: Correlación muy débil o despreciable
- Valor absoluto entre 0,20 y 0,35: Correlación débil, pero considerable
- Valor absoluto entre 0,35 y 0,50: Correlación moderada o razonable
- Valor absoluto entre 0,50 y 0,70: Correlación considerablemente alta
- Valor absoluto entre 0,70 y 1,00: Correlación muy fuerte y considerable
- Valor absoluto 1,00: Correlación perfecta

No obstante, la interpretación de estos valores será analizada con cautela en caso de obtener cifras que estén en los límites de cada rango. Ya que se toma en consideración las críticas realizadas por Schober et al. (2018) a este tipo de escalas, quienes señalan que resulta inapropiado discriminar entre una categoría o la otra en base a diferencias mínimas como al centésima de decimal (p.3)

Luego, en lo que respecta a los supuestos necesarios, este índice no es capaz obtener resultados representativos cuando se trabaja con muestras pequeñas de datos, que cuentan con valores extremos, o que presentan distribuciones distintas a la normal. En otras palabras, acorde a Schober et al. (2018), para poder aplicar el coeficiente de correlación de Pearson es necesario que se cumplan los siguientes supuestos:

- Los datos deben ser obtenidos por un muestreo aleatorio o ser representativos de la población de interés.
- Ambas variables deben ser continuas, aleatorias y con distribución bivariada normal.

- No debe haber presencia de valores extremos¹¹.
- Las variables son obtenidas de manera independiente (p.2).

Cabe mencionar que para asumir una distribución bivariada normal es necesario que ambas variables se distribuyan de forma normal y el gráfico de dispersión entre ellas debe describir una línea recta (Schober et al., 2018).

De cumplirse estos supuestos se podría señalar que los datos presentan las condiciones óptimas para generar resultados significativos con el Coeficiente de Correlación de Pearson.

Posteriormente, otra función del Coeficiente de correlación es su uso para verificar la presencia de multicolinealidad, que se refiere a la situación donde dos variables explicativas dentro de un mismo modelo de regresión poseen una alta correlación entre sí. Generando que ambas aporten el mismo tipo de información de manera duplicada, perjudicando así la predicción (Senthilnathan, 2019). Si bien este coeficiente no es la única metodología para examinar la existencia de multicolinealidad, se estima abarcarlo mencionarlo ya que una observación así es de utilidad al momento de elegir un modelo de regresión, que será abordado en próximos subcapítulos.

El último hecho a considerar sobre el Coeficiente de Correlación de Pearson es el indicador que se forma al elevar al cuadrado su valor, conocido como Coeficiente de Determinación (R^2). Esta nueva medida es utilizada para verificar la validez de una regresión, indicando la precisión con la que la variable independiente logra explicar a la variable dependiente (Senthilnathan, 2019). Por lo tanto, al calcular el coeficiente de correlación entre una variable explicada y una explicativa es posible conocer cuán precisa

¹¹ En caso de existir se debe hacer pruebas con o sin esas observaciones, u optar por implementar un coeficiente robusto como es el coeficiente de Correlación de Spearman.

sería una predicción si ambas fuesen incluidas en un modelo regresivo lineal. Y en aquellos casos donde se obtiene un valor alto del coeficiente de correlación es posible predecir una variable dependiente con el menor error posible (Senthilnathan, 2019).

4.8.4.2. Coeficiente de correlación de Spearman

Por otro lado, el Índice de Spearman corresponde a una medida no paramétrica para estimar la intensidad de asociación entre dos variables (Hauke & Kossowski, 2011). En donde “no paramétrico”, según la misma publicación, se refiere a la cualidad de no necesitar que los datos posean una determinada distribución probabilística, diferenciándose así del Coeficiente de Correlación de Pearson, que requiere datos distribuidos de manera normal.

Además, otra de sus características es que indica cuán bien una función monótona puede describir la relación entre dos variables (Hauke & Kossowski, 2011). Lo cual no es necesariamente igual a una relación lineal, ya que una relación monótona es aquella donde ocurren dos posibles escenarios: Mientras una variable aumenta, la otra también. O si una disminuye, la otra también (Schober et al., 2018). Es decir, una relación monótona es un concepto que incluye dentro de sí las relaciones lineales, ya que estas últimas solo se diferencian por tener variables que aumentan o disminuyen monótonamente a una razón de cambio constante. Siendo así consideradas como un caso especial dentro de las relaciones monótonas (Schober et al., 2018). Mientras que las relaciones monótonas abarcan todos los escenarios donde las variables se comportan de ese modo, pudiendo describir también otros tipos de funciones que se consideran monótonas, como los logaritmos naturales o la función exponencial.

Adicionalmente este indicador puede ser empleado con datos que posean valores extremos relevantes (Schober et al., 2018), siendo así un indicador robusto para analizar la correlación entre dos variables.

Debido a estas características robustas el Coeficiente de Spearman es capaz de medir la correlación entre dos variables incluso cuando no se cumplen la mayoría de los supuestos necesarios para utilizar el Coeficiente de Pearson. Razón por la cual se le considera como una herramienta que permitirá analizar la correlación independientemente de obtener o no estadísticas descriptivas que sean aptas para aplicar el Coeficiente de Pearson.

En cuanto a los valores que se pueden obtener con este coeficiente, estos también se encuentran dentro de un rango desde -1 a +1, donde un valor cero indica que no hay asociación monótona y aquellas cifras con magnitud cercana a 1 señalan que hay una fuerte relación monótona (Schober et al., 2018).

Su uso se dentro de la literatura sobre Bitcoin es reducido en comparación al índice de Pearson, estando presente en el estudio de la relación entre Bitcoin y el interés captado por la criptomoneda según Google Trends (Dokic et al., 2015).

4.8.4.3. Consideraciones para ambos coeficientes

Para ambos coeficientes existe una limitación común, que se basa en la diferencia entre el concepto de correlación y causalidad. El primero es un indicador de cuán fuerte es el grado de asociación entre dos variables, pero no entrega información sobre cuál es la relación entre ellas (por ejemplo: causalidad) (Senthilnathan, 2019; Schober et al., 2018). Mientras que el concepto de causalidad es posible de obtener mediante otras metodologías, como las “Pruebas de Causalidad de Granger”. Por consiguiente, se debe tener precaución de no

señalar que una variable es causante de la otra, ya que los coeficientes de correlación no son capaces de justificar tal tipo de conclusión.

También con estos dos coeficientes es posible utilizar pruebas de hipótesis e intervalos de confianza para estudiar la significancia de los resultados (Schober et al., 2018), que serán cubierto brevemente en el subcapítulo siguiente.

Otro punto a considerar son la obtención de valores iguales a cero, que para ninguno de los dos coeficientes implica que la relación entre las variables sea inexistente. Tanto el coeficiente de correlación de Pearson como Spearman sólo analizan relaciones lineales y monotónicas, respectivamente. Por lo tanto, en caso de obtener valores iguales a cero, solo es posible negar que existan esos dos tipos de vínculos, pero puede que ambas variables sí estén relacionadas con alguna otra forma de relación medida por alguna metodología distinta (Schober et al., 2018).

Por último, a modo de acercamiento con el tipo de datos que se ocupa en esta memoria, se debe señalar que la aplicación de la correlación entre dos series de tiempo pasa a ser conocida como “Correlación Cruzada” (Guevara, 2014). Siendo esto un caso particular dentro del estudio de la correlación, ya que la Correlación Cruzada también es capaz de analizar la relación existente tomando en consideración un retardo, es decir, la situación donde un punto de la serie de tiempo¹² se correlaciona con otro de la segunda serie, pero en un instante temporal desplazado, que puede ser anticipado o tardío con respecto al punto inicial. De esa forma, solo la correlación cruzada con retardo igual a cero es semejante a los análisis que pueden llevar a cabo los coeficientes de Pearson y Spearman, ya que se estudia

¹² Una serie de tiempo se define como una secuencia de datos que son recolectados en intervalos regulares a lo largo de un periodo de tiempo (Qiang et al., 2016).

la asociación cuando las observaciones corresponden a un mismo instante en el tiempo (Guevara, 2014).

Sin embargo, con estos cambios generados al estudiar la correlación entre dos series de tiempo también se da origen a nuevos supuestos necesarios para garantizar resultados confiables. Uno de ellos es el requerimiento de que ambas series de tiempo sean estacionarias, es decir, que los valores de sus medias y varianzas no dependan del tiempo (Guevara, 2014).

En caso de no obtener series de tiempo estacionarias, una alternativa para acercarlas a valores estacionarios es aplicar las primeras diferencias, al igual que si poseen heteroscedasticidad se recomienda aplicar logaritmos naturales (Guevara, 2014).

La no consideración de estos supuestos arriesga obtener resultados que indiquen una correlación, pero que en realidad se trate de una “Regresión Espuria” (Guevara, 2014), que se define como una relación estadística entre variables independientes que carece de sentido y que se obtiene al trabajar con series de tiempo no estacionarias (Gómez-Zaldivar et al., 2009). Por lo tanto, dichos casos no conducen a conclusiones válidas para la investigación, provocando que el cumplimiento de estos supuestos sea de gran importancia, para así apuntar a obtener los mejores resultados posibles.

4.8.4.4. Significancia estadística de los coeficientes

Tanto para el Coeficiente de Correlación de Pearson como el de Spearman Stata es capaz de proveer el Valor p del cálculo, permitiendo así poder juzgar la significancia estadística de los resultados. La hipótesis nula considerada para este tipo de prueba de hipótesis es que la correlación es cero, de la forma “ $H_0: R = 0$ ”, por lo que la segunda hipótesis sería “ $H_0: R \neq 0$ ”. De ese modo, al rechazar la hipótesis nula se concluye que se

reúnen los antecedentes necesarios para decir que coeficiente de correlación es distinto de cero, generando que sí haya asociación entre las variables. Mientras que, por el contrario, de no rechazarse la hipótesis nula es posible mencionar que con el cálculo realizado no es posible rechazar la posibilidad del coeficiente de correlación siendo igual a cero.

5. METODOLOGÍA

La presente memoria corresponde a una investigación empírica con un propósito exploratorio, que pretende comprobar cuantitativamente la existencia de una relación entre el material audiovisual sobre criptomonedas y el mercado de Bitcoin. Para ello, primero se extrajo información histórica de Bitcoin desde el banco de datos de Coingecko. Luego se crearon series de tiempo con los principales indicadores de audiencia provenientes de cada fuente de material audiovisual (Producciones cinematográficas, YouTube y Twitch). Para posteriormente elaborar las correspondientes estadísticas descriptivas, realizar las transformaciones de datos necesarias y aplicar los modelos seleccionados para resolver el problema de investigación, obteniéndose los resultados sintetizados en el capítulo de conclusiones.

5.1. Datos

En el presente subcapítulo se detallan los procesos para la recolección y transformación de datos, con el objetivo de que estos adquieran un formato adecuado para la aplicación de las metodologías. Además de mencionar cada una de las cifras consideradas como variables, junto a sus definiciones y características principales como series de tiempo.

5.1.1. Bitcoin

5.1.1.1. Recolección de datos diarios

La información histórica utilizada que describe al mercado del Bitcoin se obtiene desde el sitio web de Coingecko (Coingecko FAQ), específicamente el apartado de datos históricos de dicha criptomoneda (Datos históricos de Bitcoin USD, s. f.). En él se utiliza la función de exportar todos los registros en un archivo de formato Excel, obteniendo así 3365 observaciones en total a la fecha de recopilación (17 de julio de 2022).

Inicialmente la información se ubica solo en una columna del archivo Excel, por lo que se le separa con la herramienta “Texto en columnas” para obtener las siguientes variables en pilares independientes:

- Fecha (Date): Corresponde al día en el que es anotado el valor de las otras variables. Sus valores abarcan desde el 8 de Julio de 2019 hasta el 17 de julio de 2022, es decir, tres años aproximadamente.

- Hora (Time): Es la hora en la que se registran las variables de cada observación. Posee un valor constante de “00:00:00”, es decir, todos los elementos fueron tomados a la misma hora todos los días.

Al ser una variable constante y sin utilidad para este estudio, no se le vuelve a considerar nuevamente.

- Zona horaria (TimeZone): posee el valor “UTC” para todas las observaciones, indicando que la variable “hora” se vincula a la franja horaria “Coordinated Universal Time (UTC)”.

Tal como ocurre con la variable “Hora”, la zona horaria no es relevante para esta investigación, razón por la cual se le omite a contar de este punto.

- Precio (BTC_Price): Es el promedio ponderado (según el volumen de cada transacción) del precio del Bitcoin en todas las operaciones realizadas durante un mismo día (Coingecko FAQ, s.f.). Su valor es expresado en dólares estadounidenses.

- Capitalización de mercado (BTC_MktCap): Se define como el valor monetario total del conjunto de monedas en circulación de una criptomoneda para una fecha determinada. Se obtiene al multiplicar el precio por el

suministro circulante, siendo una medida de la cuota de mercado de un criptoactivo (Coingecko FAQ, s.f.). Al ser una variable derivada del precio y multiplicada por una escalar, su unidad también es en dólares estadounidenses (USD).

- Volumen (BTC_TrdVol): El volumen de intercambio de Bitcoin corresponde al valor monetario total de todas las cantidades compradas y vendidas en un día, a través de operaciones entre los pares comerciales¹³ existentes para esta criptomoneda (Methodology, s. f.). Su magnitud es expresada en dólares estadounidenses (USD).

Cabe mencionar que se utiliza el dólar estadounidense como la unidad monetaria para expresar el precio del Bitcoin, debido a que los índices en esa moneda poseen mayor relevancia en la literatura sobre criptomonedas y, al ser extraídos de bases de datos confiables como Coingecko, prometen generar resultados consistentes al trabajar con monedas de significativa liquidez (Vidal-Tomás, 2022). Tomando en cuenta que la liquidez de Bitcoin ha ido en aumento en los últimos años y es considerado una buena alternativa como diversificador de riesgo en cuanto a liquidez (Scharnowski, 2021)).

Con ello los datos acerca de Bitcoin se encuentran en un formato apropiado para poder ser analizados en cuanto a sus características como series de tiempo al verificar los supuestos de los modelos, seguido por la aplicación de transformaciones de datos necesarias y, finalmente, el análisis de los modelos elegidos para dar respuesta a la problemática de investigación.

¹³ Los pares comerciales son las combinaciones por las que puede ser intercambiada una criptomoneda, pudiendo ser con otra criptodivisa o con una moneda fiduciaria (USD, CLP, EUR, etc.) (Pair, 2021).

5.1.1.2. Transformación de datos a formato semanal

A partir de la información diaria de Bitcoin proporcionada por Coingecko (Coingecko, s.f.) se confecciona una segunda base de datos con información de la criptomoneda, que registra valores semanales de las dos principales variables anteriormente mencionadas. Su propósito es permitir analizar la actividad del mercado del Bitcoin frente a otras series de tiempo que poseen únicamente frecuencia semanal, como la información de SocialBlade sobre YouTube, siendo este un formato predefinido por la fuente de información de la que se extraen los datos.

La metodología aplicada para llevar los datos desde un formato diario a uno semanal se basa en la metodología ocupada por (Kristoufek, 2013) en su investigación sobre los vínculos entre Bitcoin, Wikipedia y Google Trends. En ella se utiliza el promedio de las observaciones correspondientes a cada semana, considerando los lunes como el primer día de cada periodo y a los domingos como el último.

Dicho calculo fue ejecutado a través de Microsoft Excel, al tener los datos diarios en un archivo con formato compatible a ese programa (.xlsx).

El estudio de series de tiempo semanales sobre Bitcoin ha sido cubierto al analizar la correlación de esta criptomoneda con la bolsa de valores estadounidense durante la pandemia del COVID-19 y otros periodos de alta incertidumbre (Nguyen, 2022). E incluso hay metodologías que consideran datos mensuales para indagar sobre la sensibilidad de Bitcoin ante el precio del oro, el precio del petróleo y otras variables relevantes para la economía global (Jareño et al., 2020).

Cabe destacar que al trabajar con información semanal se tiene evidencia respaldando la siguiente ventaja observada: “en el caso de Bitcoin, los datos con frecuencias de muestreo menores (como los datos mensuales frente a semanales, o semanales contra diarios) reducen

drásticamente la curtosis observada en el logaritmo de los retornos” (Dehouche, 2021). Siendo esto algo que, según el mismo autor, es opuesto a lo obtenido en su estudio con otras variables financieras como el precio del oro o el índice S&P 500, dado que al bajar la frecuencia de muestreo con esos datos las colas de la distribución se mostraron más pesadas, con una desviación estándar mayor y menor convergencia. Por esta razón también se menciona que “la volatilidad del Bitcoin es un efecto al interior de cada día y cada semana, que se muestra muy atenuado en una escala de tiempo semanal” (Dehouche, 2021).

En consecuencia, se considera que al transformar las series de tiempo diarias de Bitcoin a un formato semanal se obtienen ventajas en cuanto a las pruebas y criterios previos necesarios para aplicar los modelos a utilizar, como es el cálculo de curtosis. Además de adquirir un enfoque en sintonía con la literatura existente (al haber otros estudios usando frecuencias semanales), para obtener resultados que sean utilidad para la conversación académica sobre Bitcoin y su relación con variables externas a su mercado.

Cabe mencionar que con esta transformación los valores de la variable “Fecha (Date)” abarcan desde el día 8 de julio de 2019 hasta la semana del 11 de julio de 2022, considerando que una semana comienza un lunes y termina en domingo. No obstante, también se crea una segunda serie de tiempo que comienza el día 7 de julio de 2019 hasta el 10 de julio de 2022, a causa de que se necesita adoptar un formato de semana que comience en domingo y termine en sábado para trabajar con los datos de Google Trends. La justificación de esta decisión se explica en las próximas páginas, en la sección destinada a la obtención de datos de dicha plataforma (Índices de audiencia – Películas, series y documentales).

Finalmente, la serie de tiempo pasa a contar con 158 observaciones, para ambas versiones creadas.

5.1.2. Índices de audiencia

5.1.2.1. Recolección de datos

Para medir el interés de las personas sobre las películas, series o documentales con temática vinculada a criptomonedas primero se reunió una muestra de títulos que califiquen dentro de esta categoría. Luego se consultó en Google Trends cada uno de los nombres recopilados, obteniendo así series de tiempo semanales que describan el nivel de búsqueda que han presentado estos títulos desde 2017. Con ello culmina el proceso de recolección de datos y se procede a verificar los supuestos de las metodologías seleccionadas.

En las siguientes secciones se abordará cada etapa del proceso descrito, justificando cada una de las decisiones y presentando las variables a considerar.

5.1.2.1.1. Selección del material audiovisual sobre criptomonedas

Para identificar todas aquellas series, películas y documentales que se relacionen con las criptomonedas se recurrirá a las bases de datos IMDb y RottenTomatoes de manera complementaria. En ellas se buscarán la palabra como “Cryptocurrency”, “Crypto”, “Blockchain”, “Bitcoin” y “Bitcoin-mining”, obteniendo así una serie de rankings de popularidad con producciones que están ligadas a dichos términos. Siguiendo un orden de popularidad descendente, se verifica la aparición de las criptomonedas como temática, consultando la información dispuesta en las bases de datos utilizadas e información obtenida en Google. Si un título cumple con dicho criterio, es incorporado a la lista que será considerada para en este estudio.

La búsqueda del mayor número de obras audiovisuales disponibles generó una nómina con 17 producciones cinematográficas encontradas que cumplen con el criterio de selección. La cual se encuentra disponible en el Apéndice 1 de esta memoria. En ella se

registra el nombre de cada título, el año en el que fue estrenada, tipo de contenido (Película, Serie o Documental) y el enlace a su presentación en IMDb (IMDb, s.f.).

A partir de esta muestra es posible proceder con la compilación de datos sobre los niveles de búsqueda en Google.

5.1.2.1.2. Recopilación de datos vía Google Trends

Con la lista de filmes seleccionados se pasa a buscar cada uno de sus títulos en Google Trends, obteniendo valores que correspondan al nivel de búsqueda mundial a partir del año 2017. Tanto para los buscadores de Google como los de YouTube, dado que Google Trends dispone de ambas opciones.

Se elige estudiar las cifras mundiales dado que la actividad en el mercado de Bitcoin no se limita a ningún país en específico, generando que no sea relevante seleccionar solo aquellas búsquedas de google que provengan de un país determinado. Al igual que se decide recopilar datos desde el año 2017 porque es la fecha más antigua para la que Google Trends entrega cifras semanales. En caso de solicitar información anterior a dicho año, la plataforma entrega únicamente datos mensuales. Lo cual no resulta viable para el presente estudio, al estar estudiando todas las variables consideradas con un formato semanal.

El criterio para admitir un título en la muestra es, primero, que las sugerencias de búsqueda de Google Trends reconozcan dicho nombre como aquel asociado a la producción cinematográfica que se desea identificar, por ejemplo: “The Badger – Película de 2020”. Verificando que el nombre, tipo de material audiovisual y año de estreno coincidan con los datos registrados para este título. De esa forma se asegura que la plataforma esté rastreando efectivamente aquellas búsquedas dirigidas hacia el material audiovisual objetivo, y no aquellas que usan las mismas palabras, pero se refieren a algún otro tópico no relacionado con la película, serie o documental que se desea investigar.

Luego, en segundo lugar, se requiere que haya un nivel de búsqueda suficiente para que Google Trends pueda elaborar gráficos. De lo contrario, la plataforma no entregará datos sobre el nivel de búsqueda a lo largo del tiempo.

Y finalmente, tanto para datos de los buscadores de Google y YouTube, se hace revisión de los resultados de búsqueda obtenidos. Buscando que en el primer sitio existan enlaces que hagan efectivamente alusión al título buscado. Mientras que para YouTube se exige que existan videos en los resultados de búsqueda, ya sean tráileres u otro tipo de contenido.

Al igual que los nombres obtenidos en la sección anterior, la lista de las 19 producciones cinematográficas que logran disponer de datos a través de Google Trends se encuentran en el Apéndice 1.

De esa forma, la base de datos relativa a los índices de audiencia para cada título de material audiovisual sobre criptomonedas posee las siguientes variables:

- Fecha (Date): Indica la semana correspondiente a cada observación, abarcando valores desde el 7 de Julio de 2019 hasta la semana del 10 de julio de 2022, tomando en cuenta que una semana comienza en domingo y termina en sábado¹⁴. Lo cual forma un periodo de aproximadamente tres años.
- Nivel de búsqueda en Google (MSD_GTG): Representa la cantidad relativa semanal de búsquedas dentro de Google en torno a cada película, serie o documental relacionado a las criptomonedas. Su valor se define como relativo

¹⁴ A diferencia de las variables temporales de plataformas como YouTube (vía SocialBlade), que definen a una semana desde el día lunes hasta el día domingo, la plataforma de Google Trends asume que una semana se extiende a contar del día domingo hasta el día sábado. Por esa razón, de forma particular para los datos de Google Trends, su análisis se llevará a cabo con datos de Bitcoin que hayan sido transformados a formato semanal tomando a los días domingo como el comienzo de cada semana. De esa forma se evita trabajar con series de tiempo que tengan una convención distinta sobre qué días se consideran dentro de una semana.

puesto que esta variable toma valores normalizados de 0 a 100. Donde 100 es la magnitud de la observación con más búsquedas en el periodo seleccionado, siendo este el valor de referencia para normalizar todas las demás cifras junto a el nivel de búsquedas total para el periodo y zona geográfica seleccionados. Por ejemplo, un valor de 50 para una semana representaría que en esa fecha se observa la mitad e búsquedas que posee el día con más solicitudes registradas (de valor 100). Mientras que el valor 0 significa que en esa fecha se registran búsquedas a un nivel muy inferior al máximo (Google Trends, s.f.). Debido a tal normalización los valores de esta variable solo pueden ser comparados con aquellos que se obtienen en una misma búsqueda, ya que cada consulta de términos hecha a Google Trends tiene su propia escala. Por lo tanto, estas cifras tampoco pueden ser sumadas entre títulos de producciones cinematográficas (consultas) diferentes.

Cabe mencionar que Google Trends en ocasiones entrega celdas con valores “<1” (menor a uno), indicando que el nivel de búsqueda es muy bajo para asignarle valor 1, pero no inexistente (valor 0). Con motivos de poder trabajar con esta información, se aproximan todos los valores menores a uno como “0,5”, tomando como guía el procedimiento seguido por (Abraham et al., 2018) para los mismos efectos. Debido a que, como no se puede conocer a qué valor menor a 1 corresponde, se asigna 0,5 como un sustituto representativo.

- Nivel de búsqueda en YouTube (MSD_GTYT): Equivale al nivel relativo semanal de búsquedas en YouTube alrededor de cada producción

cinematográfica vinculada a las criptodivisas. Sus valores e interpretaciones son idénticas a las de la variable “Nivel de búsqueda en Google”.

A modo de observación, la plataforma de Google Trends es capaz de entregar datos diarios, semanales, mensuales y anuales. El formato en el cual es dispuesta la información en el sitio web depende del periodo para el cual se solicitan las cifras. Acorde a (Kalatzis et al., 2018), se entregan datos diarios con periodos de hasta 270 días de extensión, que fue verificado a la fecha de publicación de esta memoria y es correcto.

Con motivos de poder llevar a cabo un análisis de similares características a otras categorías de material audiovisual que poseen formato semanal (YouTube), también se decide adoptar el formato semanal para los datos de Google Trends sobre obras cinematográficas. De esa forma la investigación y resultados entre categorías son cohesivos, permitiendo llegar a conclusiones para todo el material audiovisual sobre criptomonedas basado en un mismo formato de datos. De lo contrario, este estudio solo sería capaz de concluir para cada fuente de información por separado, lo cual se aleja del propósito de analizar la relación entre Bitcoin y el material audiovisual acerca de criptomonedas en redes sociales y la web como un todo. Sin contar que las conclusiones de cada tipo de material audiovisual podrían variar únicamente por el formato en el que estaban los datos, y no necesariamente por el efecto que ellos tienen sobre Bitcoin con respecto a sus pares. Lo cual efectivamente sería visible si todas las bases de datos estuviesen en el mismo formato. En otras palabras, se piensa trabajar cada categoría en igualdad de condiciones para después compararlas entre sí y concluir frente a la problemática representando a todo el material audiovisual considerado en este estudio.

Además, optar por datos semanales provoca que la recolección de datos sea más eficiente, dado que solo se extraen datos de periodos para los cuales Google Trends coincide

entregando información semanal. Evitando así tener que recurrir a metodologías para estimar datos diarios a partir de cifras semanales, como aquella utilizada por (Abraham et al., 2018; Mittal et al., 2013; Phillips y Gorse, 2018a) en publicaciones científicas, que fue propuesta originalmente por Erik Johansson en su blog (Johansson, s. f.). La cual no habría sido eficiente para esta memoria, ya que se tendría que aplicar individualmente a cada título, convirtiéndolo en una tarea muy extensa.

Por último, para reconocer la serie de tiempo de cada producción audiovisual se decide asignar un nombre de variable “P” junto a un número identificador, según la siguiente tabla:

Variable	Título	Año	Tipo
P1	How to Sell Drugs Online (Fast)	2019	Serie de TV
P2	Unfriended: Dark Web	2018	Película
P3	Dope	2015	Película
P4	Crypto	2019	Película
P5	Hacker (Anonymous)	2016	Película
P6	Deep Web	2015	Documental
P7	Cryptopia: Bitcoin, Blockchains and the Future of the Internet	2020	Documental
P8	Bitcoin: The End of Money as We Know It	2015	Documental
P9	Hidden Secrets of Money	2013	Serie de TV
P10	Keiser Report	2009	Serie de TV
P11	The Badger	2020	Película
P12	Viraali	2017	Película
P13	Bitcoin Heist	2016	Película
P14	The Rise and Rise of Bitcoin	2014	Documental
P15	Banking on Bitcoin	2016	Documental
P16	StartUp	2016	Serie de TV
P17	Forhøret (Face to face / Cara a cara)	2019	Serie de TV

Tabla 1: Variables de producciones audiovisuales

Fuente: Elaboración propia

A partir de lo anterior se justifican todas las decisiones tomadas al crear la base de datos con índices de audiencia, que finalmente está compuesta por series de tiempo

semanales e individuales para cada Película, Serie o Documental. Siendo estas capaces de describir con cuanta intensidad relativa han sido buscados dichos filmes individualmente a lo largo del tiempo.

5.1.3. Métricas en canales de YouTube

5.1.3.1. Recolección de datos

Los canales de YouTube a considerados en esta muestra fueron seleccionados a través del propio sitio web de la plataforma (YouTube, s.f.). Por medio del buscador se identifican aquellos videos sobre criptomonedas más relevantes, al mismo tiempo que se examina si los canales en los cuales han sido publicados efectivamente se dedican a tópicos relacionados a las criptomonedas. A partir de esa muestra, se elabora un ranking de los creadores de contenido según su número total de reproducciones, para luego seleccionar un número determinado de ellos a ser considerados como aquellos que serán incluidos en este estudio. Con ello se acude a la plataforma Socialblade (Social Blade, s.f.) para obtener las principales estadísticas de cada canal como series de tiempo, que luego son sumadas para crear una única variable representativa de todo el contenido sobre critpomonedas en YouTube. La cual es aquella que se analiza frente a Bitcoin para estudiar su posible relación.

5.1.3.1.1. Selección de canales sobre Critpomonedas

Tal como se mencionó anteriormente, el sitio web donde se hallan los canales de interés es la plataforma oficial de YouTube (YouTube, s. f.). En ella se hizo uso del buscador para identificar aquellos videos sobre criptomonedas con más relevancia y reproducciones según los filtros disponibles¹⁵, inspeccionando los canales en los cuales han sido subidos para verificar que traten temas relacionados a las criptomonedas regularmente. Esta inspección

¹⁵ Se consideraron términos como “Cryptocurrency”, “Crypto” y “Bitcoin”. Mientras que los filtros son “Ordenar por: Relevancia” y “Ordenar por: Recuento de visitas”.

considera revisar los títulos de sus videos, la descripción del canal y los enlaces a las redes sociales de los creadores de contenido (Instagram, Twitter, entre otros). Si esas tres características cumplen con afirmar que se trata de un canal dedicado a las criptodivisas, se agrega el canal a la muestra.

Cabe mencionar que no se estima necesario distinguir entre los distintos lenguajes que se presentan los videos. A causa de que las criptomonedas son un concepto muy dependiente de la web, donde no existen límites impidiendo que personas accedan al contenido sobre estos activos debido a su lenguaje. En caso de visibilizar un canal con idioma distinto al español o inglés, se utilizará un traductor para leer los títulos de los videos y examinar el canal. En conclusión, solo se seleccionan los videos en base a su popularidad y relevancia según YouTube, no el lenguaje.

Además, tomando inspiración por la extracción de datos desde YouTube ejecutada por (Tan, 2018), se decide tomar la precaución de realizar todas las búsquedas de canales y videos a través de Google Chrome en modo incógnito. De esa forma se evita que la plataforma muestre resultados influidos por cookies¹⁶, dado que ello podría causar que se entreguen resultados sesgados para coincidir de mejor manera con el historial y preferencias del ordenador donde está siendo consultada la información. Por ende, emplear un navegador en modo incógnito permite obtener datos imparciales que tracen de mejor manera la verdadera disposición de los canales y videos en cuanto a popularidad.

¹⁶ De acuerdo al Instituto de Información Legal de la Universidad de Cornell (Legal Information Institute, s.f.), una cookie es “información creada por un servidor de internet mientras se visita un sitio web, la cual es enviada a un navegador (Google Chrome, Mozilla Firefox, etc.). Cuya función, según la misma fuente, es ayudar al servidor del sitio web a identificar y hacer seguimiento al navegador visitante, conocer sus preferencias o recordar información que haya sido compartida anteriormente (Contraseñas, direcciones, entre otro tipo de información).

A partir de esa selección, se obtiene una lista que finalmente cuenta con 100 creadores de contenido que cumplen con los criterios para ser considerados como aquellos que producen videos de YouTube sobre criptomonedas.

5.1.3.1.2. Elaboración de ranking

Con motivos de refinar el listado anterior hacia una la muestra más representativa de todos los videos y tomar en cuenta aquellos con mayor alcance de personas¹⁷, se elabora un ranking de los canales según su número total de reproducciones¹⁸. Para luego apartar los primeros 50 puestos y considerarlos como la muestra definitiva a estudiar.

Inicialmente se toma como referencia el trabajo de (Conde et al., 2020) y (Tan, 2018) sobre el uso de rankings para seleccionar muestras de cuentas en YouTube. No obstante, los videos sobre criptomonedas no poseen un ranking elaborado previamente por SocialBlade, como si los tienen las categorías de canales más populares en España o el contenido dirigido a niños (siendo estos los datos trabajados en ambas publicaciones, respectivamente). Por ende, en esta memoria nace la necesidad de levantar un criterio para crear un ranking propio, siendo este el número total de visualizaciones por canal.

Se escoge ordenar el ranking por número de reproducciones ya que se opina que es una mejor representación de las interacciones con la audiencia, por sobre el número de suscriptores o el número de videos publicados. En vista de que las visualizaciones traducen cuántas veces ha sido reproducido un determinado video, implicando que efectivamente hubo contacto con el público. Mientras que las suscripciones solo indican cuántas personas reciben

¹⁷ Siendo esta un enfoque tomado en referencia a (Lepori, 2015), donde para analizar la relación entre la bolsa de valores y los finales de series televisivas se selecciona a aquellas producciones que atraían más público.

¹⁸ El listado completo de las cuentas de YouTube seleccionadas está disponible en el Apéndice 2 (Sección 10.2), junto a una breve presentación de sus cifras y la correspondiente posición que ocupan en el ranking creado según número total de visualizaciones.

notificaciones o recordatorios al publicarse un video nuevo, ante lo cual el video no necesariamente llega a ser visto. Y de manera similar, el número de videos no es capaz de cuantificar a cuántas personas alcanza cada uno de ellos, sino que únicamente expresa la cantidad de contenido que ha sido creado.

A partir del ranking creado se toman los primeros 50 puestos para ser considerados en la consolidación de datos sobre YouTube. Sobre la base de esta nómina se consideran identificados los canales que componen una muestra significativa del contenido más popular sobre criptomonedas en YouTube. Siendo las estadísticas de estos las que serán consultadas a continuación en la plataforma SocialBlade (Social Blade, s.f.), para obtener datos en formado de series de tiempo individuales sobre cada canal.

5.1.3.1.3. Obtención de datos en SocialBlade

La recopilación de datos en esta plataforma es realizada manualmente al ingresar cada uno de los nombres pertenecientes a los canales de YouTube seleccionados en la sección anterior. Para ello se utiliza el buscador dispuesto en el sitio web, indicando a un costado de este que se desea consultar información de YouTube. Dado que SocialBlade también entrega estadísticas de otras redes sociales (Twitch, Instagram, TikTok, entre otras).

Al acceder a la presentación de estadísticas que SocialBlade dispone para cada canal buscado se ingresa al área “Detailed Statistics”, en la que se encuentran los gráficos de suscriptores ganados semanalmente y visualizaciones obtenidas cada semana. Un ejemplo de visualización de estos resultados puede ser apreciada en la Figura 3.

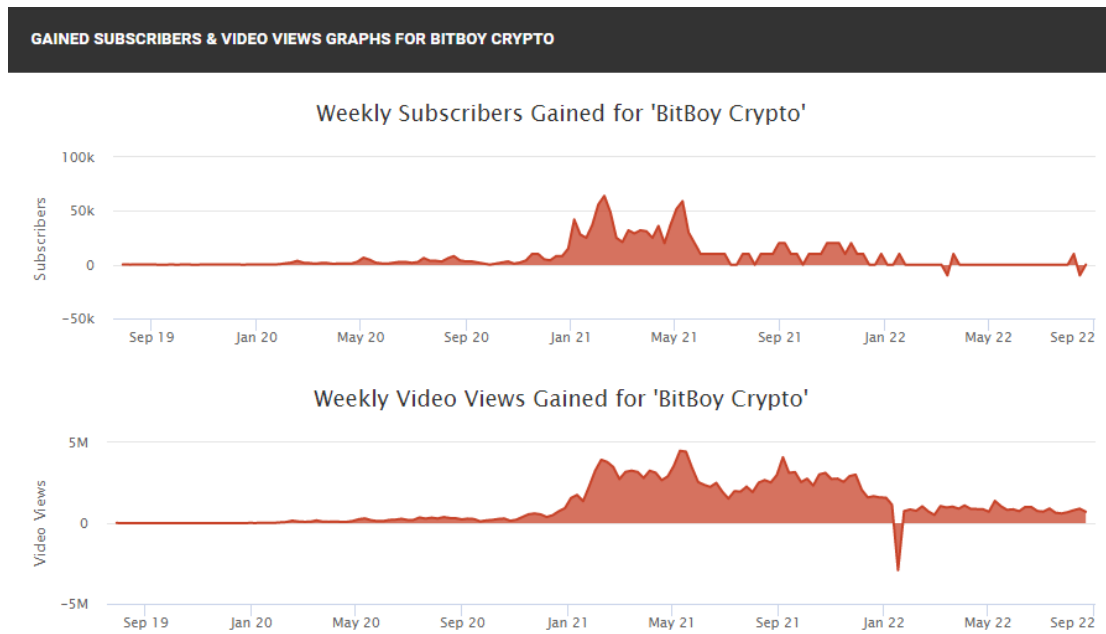


Figura 3: Gráficos de SocialBlade para el canal "BitBoy Crypto"

Fuente: Socialblade (s.f.)

A partir de ellos se transfiere manualmente la información a la base de datos de Excel, visualizando las cifras de cada observación al posicionar el cursor sobre cada punto del gráfico. Este proceso es repetido para cada uno de los canales en la muestra.

Igualmente, el sitio web contiene gráficos para la cantidad total de suscriptores y visualizaciones. Pero se decide usar los datos pertenecientes al aumento semanal de dichas variables puesto que los valores totales suelen ser cifras que aumentan constantemente. Lo cual coincide con la descripción de (Banerjee, Tutor in Economics and Barnett Fellow Anindya et al., 1993) sobre series de tiempo que no tienden a ser estacionarias. Trayendo consigo que probablemente se deban aplicar transformaciones a los datos para poder alcanzar series de tiempo con mayor estacionaridad, según los requisitos de cada modelo a aplicar en esta investigación. Siendo una de estas posibles transformaciones la aplicación de las primeras diferencias, por lo tanto, se estaría apuntando a llegar a tener una distribución de datos idéntica a la de los gráficos de suscriptores y visualizaciones ganados semanalmente.

Generando que todo el proceso de transformación y procesamiento de datos sea redundante, porque Socialblade ya dispone de un gráfico con dichas primeras diferencias. De esa forma se considera que la elección de esos gráficos para la extracción de datos es superior a los otros gráficos disponibles.

Por otro lado, en caso de encontrar que uno de los canales seleccionados no posee estadísticas elaboradas por SocialBlade, se procederá a tomar aquel creador de contenido siguiente dentro del ranking creado anteriormente. Con el objetivo de garantizar que se esté trabajando con el mismo número escogido de canales en este estudio.

Cabe mencionar que en SocialBlade pueden encontrarse valores negativos para los suscriptores ganados semanalmente y visualizaciones obtenidas cada semana. Esto no corresponde a un error de la plataforma, sino que representa aquellas oportunidades donde un determinado canal de YouTube retira de su cuenta un video o lo convierte a privado¹⁹, evitando que sea visto por el público (Social Blade FAQ, s.f.). Provocando que SocialBlade reconozca las estadísticas de ese video como perdidos, por ende, se registra un valor menor a cero.

No obstante, para efectos de esta investigación los valores negativos correspondientes a visualizaciones no son válidos, ya que eso indicaría que las visualizaciones pasadas de ese video dejan de existir. Siendo esto una interpretación poco realista, dado que el video aun así fue consumido por personas durante un periodo de tiempo, incluso cuando ya no se le pueda ver más a contar de una determinada fecha. Entonces, se estima que lo más adecuado es no registrar dichas cifras, reemplazándolas por cero. De esa forma los videos eliminados o quitados de la visión del público no son despojados de sus visualizaciones, sino que solo se

¹⁹ Convertir un video a privado es limitar el acceso a un video de YouTube, para que solo la persona que lo subió a la plataforma pueda visualizarlo.

tomará en cuenta su incapacidad para generar visualizaciones a partir de un punto en el tiempo.

Por el contrario, los valores negativos de suscripciones sí resultan valiosos, puesto que las personas pueden decidir no estar dispuestas a ser notificadas al haber un nuevo video de un creador de contenido, bajando así el contador de suscriptores. Por consiguiente, tener valores positivos y negativos para las suscripciones resulta una buena medida para entender cómo varía la cantidad de personas que efectivamente continúan atentos a un determinado canal.

A partir de este proceso de recopilación, se obtiene una base de datos con estadísticas de todos los canales de YouTube a considerar, que en la siguiente sección serán transformadas a una única serie de tiempo que será analizada en las estadísticas descriptivas y frente a las variables de Bitcoin.

5.1.3.1.4. Consolidación de una única serie de tiempo

Esta última etapa de la recolección de datos busca unir todas las series de tiempo recopiladas, para obtener solo una que simbolice las interacciones de las personas frente a todo tipo de material audiovisual sobre criptomonedas.

Para ello, se suman los valores semanales obtenidos para los canales seleccionados, resultando en una base de datos con 158 observaciones para las variables descritas a continuación:

- **Fecha (Date):** Indica la semana para la que corresponden los valores de las otras variables. Sus valores se extienden desde el 8 de julio de 2019 hasta la semana del 11 de julio de 2022, siendo así un periodo de tres años aproximadamente. Levantando como supuesto que una semana comienza en lunes y termina en domingo.

- Número de nuevos suscriptores (YT_NwSbs): Corresponde a la cantidad de personas que decidieron suscribirse a un canal que trata tópicos sobre criptomonedas durante una semana específica.
- Número de visualizaciones (YT_Vws): Señala las veces que un video subido por un canal sobre criptomonedas fue visibilizado por una persona durante una determinada semana.

Teniendo así datos en condiciones óptimas para llevar a cabo un análisis de estadísticas descriptivas, previo a la aplicación de los modelos escogidos para hallar la relación entre Bitcoin y el material audiovisual sobre criptomonedas en YouTube.

5.1.4. Métricas en canales de Twitch

5.1.4.1. Recolección de datos

Para recolectar los datos de Twitch primero es necesario determinar si se utilizarán variables correspondientes a cada canal individual²⁰, categoría de programación²¹ o una mezcla de ambos grupos. Ya que las variables disponibles y la forma en que son extraídas es diferente entre sí, al encontrarse en distintos lugares y formatos dentro de TwitchTracker. Es por ello que se decide describir brevemente los procedimientos necesarios en la extracción de ambas agrupaciones de variables, al igual que los resultados que se esperaban obtener, para así poder evaluar cuál es la mejor opción para obtener los datos a ser analizados en esta investigación.

El primer grupo, las variables propias de cada canal, requieren seleccionar una muestra de aquellas cuentas que suelen realizar transmisiones en vivo sobre criptomonedas

²⁰ Las cuales son: Número de seguidores, número de espectadores promedio, horas transmitidas en vivo y horas visualizadas por la audiencia.

²¹ Siendo estas: Cantidad promedio de espectadores diarios visualizando un tipo de programación y el número de canales que han transmitido en dichas categorías.

y crear una nómina con todos sus nombres. Para ello se utiliza complementariamente los buscadores proporcionados por TwitchTracker, Sullygnome y TwitchMetrics. En vista de que TwitchTracker, la plataforma con las series de tiempo más apropiadas para esta investigación, no cuenta con un ranking elaborado alrededor de cada categoría, mientras que los otros dos sitios web sí. Después se elabora un ranking de dichos creadores de contenidos, ordenándolos según los valores de la multiplicación entre su número promedio de espectadores y la cantidad de horas transmitidas, correspondientes solo a sus transmisiones sobre criptomonedas. Debido a que se considera a esa medida como una aproximación a la cantidad de personas alcanzadas durante las transmisiones emitidas por cada canal. A partir de ese ranking se toman los primeros 50 lugares, siendo estos aquellos nombres que son considerados para crear las series de tiempo a estudiar. Se toma registro de los valores de las cinco variables disponibles para cada canal, que luego serán sumadas, obteniendo así una única distribución de datos para cada variable.

Las ventajas de esta alternativa es la realización de una exploración de los canales sobre criptomonedas (al crear un ranking de ellos), obtener cinco variables en formato de series de tiempo, y poder tener claridad de los canales que serán considerados para el estudio. Mientras que su gran desventaja es el tiempo necesario para poder llevar a cabo este procedimiento.

Por otro lado, el segundo grupo de variables (propias de cada categoría) precisan identificar aquellas programaciones que identifican al contenido sobre criptomonedas en Twitch. Estas corresponden a “Crypto” y “Bitcoin”, que son aquellas categorías más utilizadas en los canales sobre criptomonedas, según lo obtenido en la exploración ligada a la creación del ranking de creadores de contenido dedicados a esta temática. Luego de conocerlas se ingresan ambos nombres en el buscador de la plataforma de TwitchTracker,

donde se presenta un apartado con las series de tiempo que se quieren tomar en consideración. De esa forma se toma registro de cada una de las observaciones, resultando dos series de tiempo: promedio de espectadores y cantidad de canales transmitiendo en dichas categorías. Posteriormente, estas son sumadas para obtener una única serie de datos para cada variable, que representan las interacciones de la audiencia con la plataforma Twitch.

Una de sus ventajas es la consideración de datos relativos a las categorías, lo cual implica que se abarca la totalidad del contenido en Twitch que posee etiquetas como “Crypto” y “Bitcoin”. Mientras que la primera solo toma en cuenta los primeros 50 canales más populares según la multiplicación entre sus espectadores promedios y horas de video transmitidas. Al igual que posee la cualidad positiva de ser un proceso que requiere menos tiempo en comparación a la primera alternativa, ya que solo se requiere registrar manualmente los datos diarios de dos categorías (en vez de cincuenta). Sin embargo, con esta opción solo se obtienen dos variables, teniendo así menos información para analizar frente a Bitcoin.

Estas dos alternativas de líneas de acción expuestas son probadas en este estudio. Inicialmente se siguen los procesos de los datos provenientes de canales, puesto que la creación de un ranking permite explorar en más detalle el comportamiento de las cuentas que crean contenido sobre criptomonedas (conociendo sus programaciones, redes sociales, comunidad, etc.). Además de apuntarse a obtener cinco variables para analizar, en vez de las dos que son posibles de obtener en la segunda opción. No obstante, al momento de extraer los datos diarios de cada canal se aprecia que este proceso de extracción supera las expectativas de tiempo destinadas a esta etapa de la metodología. Lo cual difiere de lo observado con YouTube (que comparte un proceso de obtención de datos similar y fue completado satisfactoriamente), porque aquellos datos poseen carácter semanal, y eso

implica un menor número de observaciones a ser anotadas con respecto a los valores diarios que posee Twitch en un mismo intervalo de tiempo. Por ejemplo, para YouTube se requiere una hora para cubrir las series de tiempo de una cuenta, resultando en cincuenta horas aproximadamente para toda la muestra a considerar. Mientras que para Twitch es necesario tres a cuatro horas por cuenta, resultando en doscientas horas en total. Por tal razón se decide que la recopilación de datos sobre cada canal de Twitch no es factible para esta memoria, dado que su ejecución manual es demasiado extensa, arriesgando así el cumplimiento de los periodos asignados para la realización de esta investigación.

En consecuencia, se opta por trabajar únicamente con los datos provenientes de las categorías “Crypto” y “Bitcoin”, ya que permite adquirir la mejor calidad de información alcanzable dentro los plazos de esta memoria con las herramientas que se cuenta para ello. Estos datos serán analizados frente a Bitcoin por medio de los modelos escogidos con estos fines en la sección de Bases Metodológicas. Aun así, la descripción de los métodos utilizados para elaborar el ranking y sus resultados se disponen en el Apéndice 3 “Muestra de canales de Twitch”, con el objetivo de detallar la metodología seguida y así comprender las decisiones tomadas en esa prueba.

A continuación, se expone en mayor detalle el procedimiento de recolección de datos ejecutado para las variables según categoría, describiendo los resultados obtenidos y justificando cada una de las decisiones tomadas.

5.1.4.1.1. Selección de categorías sobre criptomonedas

Para identificar aquellas categorías o “Juegos”²² en donde se emite contenido sobre criptomonedas se recurre al aprendizaje obtenido en la creación de un ranking para los canales destinados a esta temática, cuyo procedimiento se encuentra expuesto en el Apéndice 3 “Muestra de canales de Twitch”. Con ello se identifica que las dos principales categorías presentes son “Crypto” y “Bitcoin”, que están presentes en la mayor parte de los canales considerados. También se pudo apreciar la presencia de otras categorías ligadas a videojuegos que involucran criptomonedas, sin embargo, estas no son consideradas para la extracción de datos. Puesto que están presentes solo en un 5% de los canales de la muestra (considerado un porcentaje poco relevante para la investigación), al igual que no se estima necesario incluirlas en el estudio debido a los bajos valores que poseen sus series de tiempo con respecto a “Crypto” y “Bitcoin”.

5.1.4.1.2. Obtención de datos en TwitchTracker

Luego de identificar estas categorías sus nombres son ingresados en el buscador de TwitchTracker, para poder acceder a los apartados que contienen los datos de este tipo de programación. Estos están disponibles en formato de gráfico, donde es posible apreciar las cifras específicas al posicionar el cursor sobre cada una de las observaciones. En él se seleccionan los datos diarios (“Daily”), con su valor promedio (“Average”). Al igual que se habilita la aparición de ambas variables: Espectadores promedio (“Viewers”) y Canales promedio transmitiendo en esta categoría (“Channels”).

²² “Juegos” es la forma en la que se hace referencia a la temática central de cada transmisión en vivo. Aun cuando existen categorías que no corresponden legítimamente a videojuegos, como los programas de conversaciones, música, arte y Criptomonedas (Que se reconoce en Twitch primariamente como “Crypto”)

De esa forma, el gráfico del que se extraen los datos tiene la siguiente apariencia:

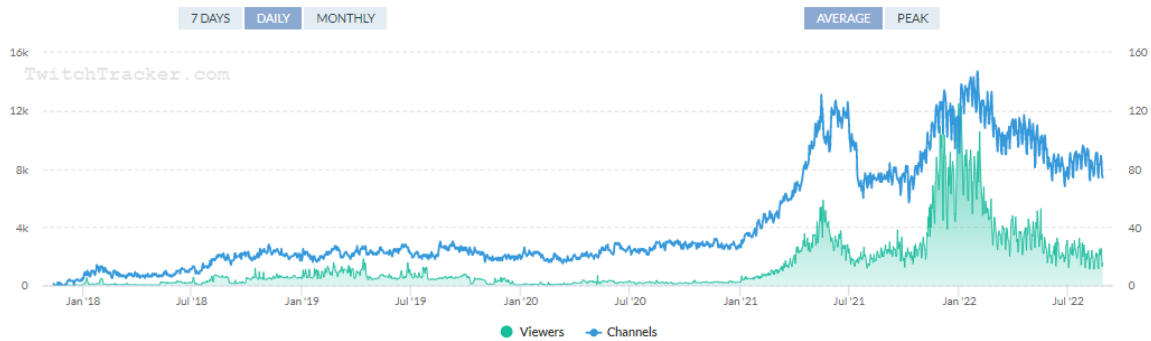


Figura 4: Gráfico de Categoría "Crypto"

Fuente: Twitchtracker (s.f.)

Cada uno de estos valores es registrado en una hoja de cálculo de Microsoft Excel, a modo de base de datos, obteniendo así series de tiempo para las dos variables de ambas categorías.

Una vez que todas las observaciones del periodo a estudiar (8 de julio de 2019 a 17 de julio de 2022) son anotadas se cuenta con una base de datos almacenando series de tiempo individuales para cada categoría, que comparten las mismas variables entre sí. Con ello se da paso a la siguiente sección, donde se crea un único conjunto de series de tiempo representativas de Twitch a utilizar frente a Bitcoin.

5.1.4.1.3. Consolidación de una única serie de tiempo

Dado que se deseaba estudiar la relación entre el material audiovisual sobre criptomonedas en Twitch frente a Bitcoin, fue necesario unir las series de tiempo individuales de cada canal para obtener solo una que estime a todo el material. Para lograrlo se sumó entre los canales todas las cifras correspondientes a una misma fecha, en cada una de las variables. Obteniendo así una base de datos definitiva compuesta con los siguientes valores:

- Fecha (TW_Date): Indica el día al cual corresponden las cifras de las demás variables. Se interpreta como la variable temporal de las series de tiempo.

Posee valores desde el día 8 de julio de 2019 al 17 de julio de 2022, describiendo aproximadamente un periodo de tres años.

- Cantidad promedio de espectadores (TW_AvgVwrs): Es la media calculada a partir del número de personas viendo las transmisiones en vivo, siendo este originalmente un valor que varía de manera constante mientras cada canal está en línea.
- Número de canales transmitiendo en dichas categorías (TW_Channls): Considera cuántas cuentas de Twitch han producido videos en vivo sobre criptomonedas en un día determinado.

Esta base de datos es aquella que participa en los modelos frente a las variables del mercado de Bitcoin, no sin antes pasar por las transformaciones de datos necesarias y la verificación de supuestos correspondientes.

5.1.4.2. Transformación de datos a formato semanal

Con motivos de poseer datos con formatos similares entre los tres tipos de material audiovisual (Cinematográfico, YouTube y Twitch), se determina aplicar una transformación a los datos de Twitch para que sus series de tiempo tengan valores semanales.

Al leer y entender de esta memoria, no existe literatura respaldando este tipo de transformación con datos de Twitch o similares. Por lo tanto, se decide tomar el mismo procedimiento ocupado para los datos de Bitcoin, respaldado por (Kristoufek, 2013). Es decir, se calcula la media aritmética de la información diaria dentro de una misma semana.

Luego de aplicar esta transformación las observaciones de la variable “Fecha (Tw_Date)” adquieren valores desde la semana del 8 de julio de 2019 hasta aquella que comienza el día 11 de julio de 2022, tomando como supuesto que una semana comienza un

día lunes y finaliza en domingo. Siendo así 158 observaciones dentro del intervalo de tiempo mencionado.

5.2. Análisis

5.2.1. Software a Utilizar

Se define que el Software estadístico a utilizar será “R” (R: The R Project for Statistical Computing, s. f.), dado que cuenta con una gran variedad de herramientas para llevar a cabo regresiones, realizar pruebas, calcular coeficientes y crear gráficos. Además, se le elige porque es un software gratuito, de fácil acceso, que posee bastos recursos disponibles en internet para poder aprender su lenguaje de programación y la forma de utilizar el programa en general.

5.2.2. Índices de audiencia

Para analizar este tipo de material audiovisual frente a Bitcoin se debe tomar en cuenta que se dispone de series de tiempo individuales para cada título de película, serie o documental seleccionado. Que, a diferencia de lo que ocurre con YouTube y Twitch, no son aptos para ser sumados y obtener una única serie temporal. Dado que la información proveniente de Google Trends se encuentra normalizada para cada título o consulta realizada en el sitio web, generando que los valores de cada serie de tiempo tengan una interpretación diferente según el material audiovisual al que corresponde.

Por esa razón se descarta la opción de crear regresiones para cada producción, optando así solo por calcular los Índices de correlación propuestos por Pearson o Spearman para cada producción cinematográfica, frente al precio y volumen de intercambio de Bitcoin. La elección de emplear un coeficiente o el otro depende de si las características que presenten los datos son apropiadas para los supuestos que exige cada uno de ellos. Este procedimiento

es realizado para ambos tipos de datos obtenidos en Google Trends, es decir, aquellos que describen las búsquedas de Google y los correspondientes a las búsquedas en YouTube.

En cuanto a los pasos a seguir, primero se ejecuta la prueba de Dickey-Fuller Aumentada²³ para identificar si las series de tiempo son estacionarias, ya que este es uno de los supuestos necesarios para poder obtener resultados fiables con los coeficientes de correlación. Se elige este test por sobre el procedimiento de Phillips-Perron (también cubierto en la revisión de literatura) debido a que Dickey-Fuller es considerada como más confiable (Shrestha & Bhatta, 2018).

Ante la eventualidad de una serie de tiempo siendo no estacionaria se consideró calcular la primera diferencia de dicha variable, para así obtener datos estacionarios que puedan ser utilizados para computar los coeficientes de correlación. Tal proceder para obtener series de tiempo estacionarias se ve justificado según (Gujarati & Porter, 2008, p.435).

En caso contrario, si las series de tiempo eran estacionarias, no era necesario aplicar una transformación.

Una vez que todas las variables son estacionarias se requiere definir si es posible aplicar el coeficiente de correlación de Pearson o aquel de Spearman. Con tal propósito se realiza la prueba de Jarque-Bera²⁴ sobre la distribución normal de las variables, ya que Pearson demanda como supuesto la distribución bivariada normal entre las variables a tomar en cuenta. Mientras que Spearman no requiere una distribución determinada para las variables, por lo que es tomado como un camino alternativo en un escenario donde los datos no se distribuyan de forma normal.

²³ Esta prueba es realizada a través de la función "adf.test ()" del paquete "tseries" en el software R.

²⁴ Esta prueba es realizada por medio de la función "jarque.bera.test ()" del paquete "tseries" en R.

Como resultado, si todas las variables poseían una distribución normal se utilizaba el coeficiente de Pearson, mientras que si eso no ocurría se optaba por emplear el coeficiente de Spearman.

Habiendo verificado ambos supuestos se continúa con la obtención del coeficiente de correlación entre las variables de Bitcoin y cada una de las series de tiempo correspondientes al material audiovisual. Lo cual es ejecutado a través de las funciones “cor ()” y “rcorr” en R.

Luego, como el análisis de correlación anterior solo se limita a estudiar la relación existente entre observaciones que comparten un mismo periodo de tiempo, se decide elaborar gráficos que estimen la función de correlación a través de la función “ccf ()” de R. Ellos permiten observar si las series de tiempo están correlacionadas de forma desplazada. Por ejemplo, se considera que hay una relación desplazada cuando una serie de tiempo se asocia con los valores pasados o futuros de otra serie temporal.

A partir de tales figuras se busca analizar si la correlación entre dos series de tiempo aumenta o disminuye al tomarse en cuenta desplazamientos, al mismo tiempo que se intenta identificar si existe un desfase común donde la correlación sea más alta con respecto a los resultados obtenidos cuando los datos se encuentran en fase.

De esa forma se abarcan todas las herramientas y procesos utilizados para analizar los índices de audiencia para películas, series y documentales. Por consiguiente, se da paso al análisis destinado a YouTube.

5.2.3. YouTube y Twitch

El procedimiento seguido por YouTube Twitch es idéntico en cuanto a las herramientas utilizadas para verificar cada supuesto e interpretar los resultados. Por tal razón,

en esta sección se cubren primero los modelos destinados a cada plataforma, para luego detallar los procedimientos comunes utilizados en ellos.

En primer lugar, el análisis de la capacidad predictiva que poseen las variables de YouTube sobre Bitcoin tomó en cuenta en un inicio los siguientes modelos de regresión lineal²⁵:

$$Precio_BTC = \beta_0 + \beta_1 \cdot (YT_NwSbs) + \beta_2 \cdot (YT_Vws) + \beta_3 \cdot Cntrl_1 + \beta_4 \cdot Cntrl_2 + \varepsilon$$

$$Vol_Interc_BTC = \beta_0 + \beta_1 \cdot (YT_NwSbs) + \beta_2 \cdot (YT_Vws) + \beta_3 \cdot Cntrl_1 + \beta_4 \cdot Cntrl_2 + \varepsilon$$

Se trata de dos estimaciones, una para el precio de Bitcoin (Precio_BTC) y otra para su volumen de intercambio (Vol_Interc_BTC), siendo explicadas por variables como el número de suscripciones a canales sobre criptomonedas (YT_NwSbs), las visualizaciones de ese contenido (YT_Vws) y las variables de control ($Cntrl_1$; $Cntrl_2$).

Posteriormente, en segundo lugar, los modelos iniciales a estimar para Twitch son:

$$Precio_BTC = \beta_0 + \beta_1 \cdot (TW_AvgVwrs) + \beta_2 \cdot (TW_Channls) + \beta_3 \cdot Cntrl_1 + \beta_4 \cdot Cntrl_2 + \varepsilon$$

$$Vol_Interc_BTC = \beta_0 + \beta_1 \cdot (TW_AvgVwrs) + \beta_2 \cdot (TW_Channls) + \beta_3 \cdot Cntrl_1 + \beta_4 \cdot Cntrl_2 + \varepsilon$$

En ellos las variables dependientes también son el precio y el volumen de intercambio de Bitcoin. Pero dentro de sus variables independientes se incluyó a las variables del número promedio de espectadores en canales de Twitch sobre criptomonedas (TW_AvgVws) y el número de canales que transmiten este tipo de contenido por semana (TW_Channls). Mientras que las variables de control son idénticas a las utilizadas en YouTube.

En ambos casos se optó por incorporar variables adicionales, llamadas variables de control, con el propósito de disminuir el sesgo por variables omitidas en la regresión. Por tal motivo se seleccionaron datos cuya influencia lineal sobre Bitcoin haya sido verificada previamente en otros estudios, siendo estas el índice CBOE VIX (Chicago Board Options

²⁵ Estos son ingresados al Software R por medio del comando "lm()"

Exchange Market Volatility Index) (Bouri et al., 2017) y el Promedio Industrial Dow Jones (Jerdack, 2018). La primera puede ser interpretada como una aproximación a la incertidumbre global, mientras que la segunda es semejante al volumen de intercambio en los mercados financieros estadounidenses. Para ambos casos los datos históricos semanales fueron obtenidos de la plataforma Yahoo! Finance (Dow Jones Industrial Average, s. f.; CBOE Volatility Index, s. f.).

Tales modelos son considerados como el estado inicial debido a que estos pueden sufrir modificaciones al estar revisando el cumplimiento de cada uno de los supuestos necesarios para el buen funcionamiento de las estimaciones.

Dicha verificación de supuestos comprende ocho puntos que es necesario evaluar, que fueron cubiertos en detalle en la sección del modelo de regresión lineal múltiple. A continuación, se mencionará cada uno de ellos en el orden que fueron probados y con una breve descripción de los comandos de R involucrados.

- Supuesto 8: Series de tiempo estacionarias
 - Las pruebas de Dickey-Fuller Aumentada (ADF) y Phillips Perron fueron realizadas con las funciones “`adf.test()`” y “`pp.test()`” respectivamente, siendo ambas del paquete “`tseries`”.
- Supuesto 1: Linealidad de los parámetros
 - Los gráficos de dispersión entre la variable dependiente y cada variable independiente fueron elaborados con la función “`plot()`”, al igual que los gráficos de residuos frente a valores estimados.

- Supuesto 2: Valor esperado de los residuos igual a cero
 - Se utilizó la función “mean()” para calcular el valor esperado.
- Supuesto 3: Condición de homocedasticidad en los residuos
 - Involucró a la función “bptest()” del paquete “lmtest” para realizar la prueba de Bresuch-Pagan.
- Supuesto 4: No hay autocorrelación entre residuos
 - Se ocupó la función “dwtest()” para ejecutar la prueba de Durbin-Watson, propia del paquete “lmtest”.
 - En los casos donde se detectó autocorrelación o heterocedasticidad se empleó la corrección de errores estándar del método Newey-West, por medio de la función “coeftest()”, que también forma parte del paquete “lmtest()”.
- Supuesto 5: Variables independientes no estocásticas
 - Se calculó la covarianza entre los residuos y cada variable independiente a través de la función “cov()”.
- Supuesto 6: Residuos con distribución normal
 - La prueba de Jarque-Bera se realizó con la función “jarque.bera.test()” del paquete “tseries”.
- Supuesto 7: No existe multicolinealidad
 - Nuevamente se hizo uso de la prueba de Jarque-Bera, para determinar si es posible usar el coeficiente de correlación de Pearson o aquel de Spearman. Se consideró la función “jarque.bera.test()” del paquete “tseries”.
 - La matriz de correlación se creó con el comando “cor()”, al ingresar una lista con las variables independientes como argumento.

- Los valores p de los coeficientes se obtuvieron con el comando “`rcorr()`” del paquete “Hmisc”.

Una vez que se verificaron todos los supuestos se hace síntesis de las principales modificaciones que adquiere el modelo, tal como el uso de primeras diferencias o la incorporación de logaritmos naturales.

Luego, al interpretar los resultados, se verificó que los coeficientes de determinación (R^2) sean menores al estadístico de Durbin-Watson, ya que de no ser cierto se estaría en presencia de una regresión espuria.

Por último, se analizan los resultados de cada modelo de regresión lineal con el comando “`summary()`”. Aunque, cuando había presencia de autocorrelación, para las pruebas de hipótesis se utilizaron los valores de la corrección de Newey-West (función “`coeftest()`”).

Se decidió tomar en cuenta los valores de los coeficientes de determinación ajustados, debido a que se trata de una regresión lineal múltiple. Estos también son proporcionados en el resumen de “`summary()`”.

6. RESULTADOS

6.1. Índices de Audiencia - Análisis de correlación

6.1.1. Verificación de supuestos

Los resultados de las pruebas de Dickey-Fuller Aumentada (ADF) y Jarque-Bera (JB) fueron resumidos en formato de tablas, indicando el valor p de cada prueba y la conclusión asociada a su respectiva prueba de hipótesis.

Con motivo de preservar la brevedad en los resultados se elige omitir a las variables que inicialmente no presentaban estacionaridad, reemplazándolas directamente a sus primeras diferencias, las cuales pueden ser reconocidas por tener la extensión “[Diff]” junto al nombre de la variable.

Variable	ADF		JB	
	Valor-p	Resultado	Valor-p	Resultado
Precio BTC [Diff]	0,011	Estacionaria	0,000	No normal
Vol. de Interc. BTC [Diff]	0,010	Estacionaria	0,000	No normal

Tabla 2: Pruebas ADF y JB para variables de Bitcoin.

Fuente: Elaboración propia.

Las primeras diferencias de ambas variables vinculadas a Bitcoin poseen valores p menores a 0,05 para la prueba de Dickey-Fuller Aumentada, indicando que se puede afirmar con un 5% de significancia estadística que son series de tiempo estacionarias. Luego para la prueba de Jarque-Bera los valores p son muy cercanos a cero, generando que se rechace la hipótesis nula con un 1% de significancia y se establezca que estas variables no poseen una distribución normal.

Variable	ADF		JB	
	Valor-p	Resultado	Valor-p	Resultado
P1G	0,010	Estacionaria	0,000	No normal
P2G[Diff]	0,010	Estacionaria	0,000	No normal
P3G[Diff]	0,010	Estacionaria	0,000	No normal
P4G	0,017	Estacionaria	0,000	No normal
P5G [Diff]	0,010	Estacionaria	0,000	No normal
P6G [Diff]	0,010	Estacionaria	0,427	Normal
P7G [Diff]	0,010	Estacionaria	0,003	No normal
P8G	0,010	Estacionaria	0,000	No normal
P9G	0,010	Estacionaria	0,000	No normal
P10G	0,010	Estacionaria	0,043	No normal
P11G	0,049	Estacionaria	0,000	No normal
P12G	0,010	Estacionaria	0,000	No normal
P13G	0,020	Estacionaria	0,000	No normal
P14G	0,020	Estacionaria	0,000	No normal
P15G [Diff]	0,010	Estacionaria	0,000	No normal
P16G [Diff]	0,010	Estacionaria	0,000	No normal
P17G	0,010	Estacionaria	0,000	No normal

Tabla 3: Pruebas ADF y JB para búsquedas en Google

Fuente: Elaboración propia.

Posteriormente las series de tiempo sobre los niveles de búsqueda en Google acerca del material audiovisual (que poseen una “G” en sus variables) poseen diferentes desempeños en la prueba de Dickey Fuller Aumentada (ADF). Las producciones 2, 3, 5, 6, 7, 15 y 16 no presentan estacionariedad inicialmente, por lo que se calculan sus primeras diferencias para lograr dicha característica. Mientras que el resto de variables sí logran ser estacionarias sin necesidad de aplicar dicha transformación. Los valores p para las series de tiempo ubicadas en la tabla son en su mayoría cercanos a 0,01, indicando que se rechaza la hipótesis nula con 1% de significancia en ellos. Con la excepción de los números 4, 11, 13 y 14, donde la hipótesis nula logra rechazarse con un 5% de significancia.

En cuanto a la prueba de Jarque-Bera (JB) solo la producción número 6 no logra rechazar la hipótesis nula, sugiriendo que posee una distribución normal. El resto de variables rechaza dicha hipótesis mayoritariamente con un 1% de significancia, exceptuando el número 10, que lo hace con un 5%.

Variable	ADF		JB	
	Valor-p	Resultado	Valor-p	Resultado
P1Y	0,014	Estacionaria	0,000	No normal
P2Y	0,019	Estacionaria	0,007	No normal
P3Y [Diff]	0,010	Estacionaria	0,000	No normal
P4Y [Diff]	0,010	Estacionaria	0,000	No normal
P5Y [Diff]	0,010	Estacionaria	0,000	No normal
P6Y [Diff]	0,010	Estacionaria	0,427	Normal
P7Y	0,010	Estacionaria	0,000	No normal
P8Y [Diff]	0,010	Estacionaria	0,000	No normal
P9Y	0,036	Estacionaria	0,000	No normal
P10Y [Diff]	0,010	Estacionaria	0,000	No normal
P11Y	0,010	Estacionaria	0,000	No normal
P12Y	0,010	Estacionaria	0,000	No normal
P13Y [Diff]	0,010	Estacionaria	0,000	No normal
P14Y [Diff]	0,010	Estacionaria	0,000	No normal
P15Y [Diff]	0,010	Estacionaria	0,806	Normal
P16Y [Diff]	0,010	Estacionaria	0,000	No normal
P17Y	0,010	Estacionaria	0,000	No normal

Tabla 4: Pruebas ADF y JB para búsquedas en YouTube

Fuente: Elaboración propia.

De manera similar los datos relativos a los niveles de búsqueda en YouTube (que poseen una “Y” en sus variables) poseen series de tiempo que no logran ser estacionarias al aplicar la prueba ADF, por lo que se calculan sus primeras diferencias. Después de aplicar las transformaciones necesarias se obtiene que todas las variables rechazan la hipótesis nula. Las series 1, 2 y 9 lo hacen con un 5% de significancia, mientras que el resto con un 1%.

Paralelamente la prueba de Jarque-Bera muestra que solo las producciones 6 y 15 se distribuyen de forma normal. Lo cual contrasta con las otras variables, pues en ellas se rechaza la hipótesis nula, indicando que no se distribuyen de esa forma.

A partir de los resultados generados para los tres grupos de variables es posible señalar que todos ellos contienen únicamente series de tiempo que cumplen con ser estacionarias, tomando en consideración que para algunas de ellas fue necesario calcular las primeras diferencias. Por ende, se cumple con el requerimiento de contar con series de tiempo apropiadas para el cálculo de los coeficientes de correlación, evitando así el riesgo de obtener resultados propios de relaciones espurias (producto de trabajar con series de tiempo no estacionarias).

Además, la prueba de Jarque-Bera permite apreciar que muy pocas variables de los tres grupos logran tener una distribución cercana a la normal. En consecuencia, los datos no cumplen con los supuestos necesarios para aplicar el coeficiente de correlación de Pearson, por lo que se considerará trabajar con el coeficiente de correlación de Spearman.

6.1.2. Análisis de correlación

Los valores del coeficiente de correlación de Spearman para cada producción audiovisual son almacenados en tablas que ordenan de mayor a menor las variables según la magnitud del coeficiente de correlación.

6.1.2.1. Niveles de Búsqueda de Google

Variable	Precio BTC		Variable	Vol. Interc. BTC	
	Corr. Spearman	Valor p		Corr. Spearman	Valor p
P1G	0,177	0,027*	P16G [Diff]	0,159	0,046*
P3G[Diff]	0,121	0,133	P10G	0,057	0,480
P4G	0,101	0,208	P11G	0,026	0,742
P9G	0,099	0,217	P15G [Diff]	0,022	0,789
P10G	0,08	0,317	P1G	0,021	0,791
P15G [Diff]	0,076	0,342	P7G [Diff]	0,004	0,964
P7G [Diff]	0,058	0,470	P6G [Diff]	-0,003	0,969
P16G [Diff]	0,045	0,579	P14G	-0,005	0,952
P11G	0,043	0,590	P12G	-0,015	0,847
P2G[Diff]	0,029	0,716	P9G	-0,018	0,819
P17G	0,011	0,891	P4G	-0,024	0,767
P12G	-0,005	0,952	P2G[Diff]	-0,039	0,624
P5G [Diff]	-0,037	0,643	P5G [Diff]	-0,073	0,366
P8G	-0,038	0,633	P3G[Diff]	-0,084	0,295
P6G [Diff]	-0,077	0,338	P8G	-0,086	0,286
P13G	-0,082	0,306	P13G	-0,091	0,259
P14G	-0,113	0,157	P17G	-0,138	0,084.
Significancia estadística: *** = 0,1% ; ** = 1% ; * = 5% ; . = 10%					

Tabla 5: Coeficientes de correlación para los niveles de búsqueda en Google

Fuente: Elaboración propia.

Las columnas que describen la correlación entre los niveles de búsqueda de Google y el precio de Bitcoin (BTC) muestran que la producción audiovisual con mayor asociación positiva es la variable “P1G”, con 0,177, que corresponde a la serie de televisión “How to sell Drugs Online (Fast) (2019)”. Mientras que aquella con la mayor asociación negativa es “P14G”, con -0,113, representando al documental “The Rise and Rise of Bitcoin (2014)”.

Sin embargo, todos los títulos analizados frente al precio de Bitcoin poseen una magnitud menor a 0,20, lo cual es considerada una correlación muy débil o despreciable, según la escala de interpretación propuesta por (Senthilnathan, 2019).

Además, la significancia estadística de estos cálculos indica que únicamente la variable “P1G2” es capaz de rechazar la hipótesis nula que supone al coeficiente de correlación como igual a cero, con un 5% de significancia. En otras palabras, se cuenta con evidencia para afirmar que la serie de televisión sobre criptomonedas “How to Sell Drugs Online (Fast)” posee una correlación monótona significativa de magnitud muy débil con el precio de Bitcoin.

Similarmente, al observar las columnas sobre el volumen de intercambio se aprecia que la variable con mayor correlación positiva es “P16G Diff”, con 0,159, que se refiere a la serie de televisión “Startup (2016)”. Luego, entre aquellas que poseen correlación negativa, la mayor magnitud recae en la variable “P17G”, de la serie de televisión del 2019 “Forhøret (Face to face / Cara a cara)”.

En términos generales, para el volumen de intercambio también se cumple que los coeficientes de correlación solo poseen magnitudes menores a 0,20, por ende, son consideradas relaciones muy débiles.

En adición, la variable “P16G [Diff]”, que cuenta con el mayor coeficiente de valor positivo, es la única producción audiovisual del grupo con significancia suficiente para establecer que se rechaza la hipótesis nula. En consecuencia, solo es posible aseverar que los cambios semanales en los niveles búsquedas sobre serie de televisión “Forhøret (Face to face / Cara a cara)” poseen una correlación monótona significativa de magnitud muy débil con el volumen de intercambio de Bitcoin.

También se analiza los gráficos de la función de correlación de cada variable para un periodo de 26 semanas, es decir, cercano a seis meses. Estos son:

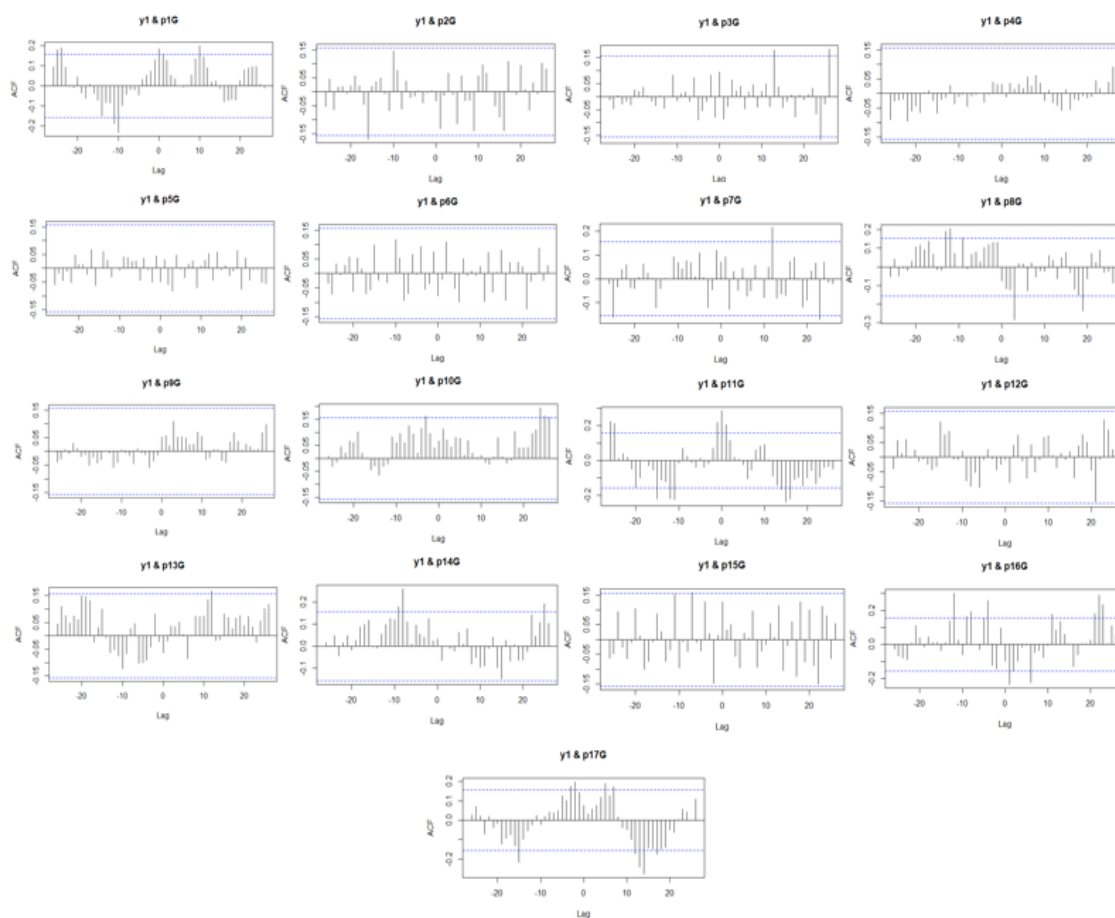


Figura 5: Gráficos de la función de correlación para búsquedas en Google frente a los retornos

Fuente: Elaboración propia.

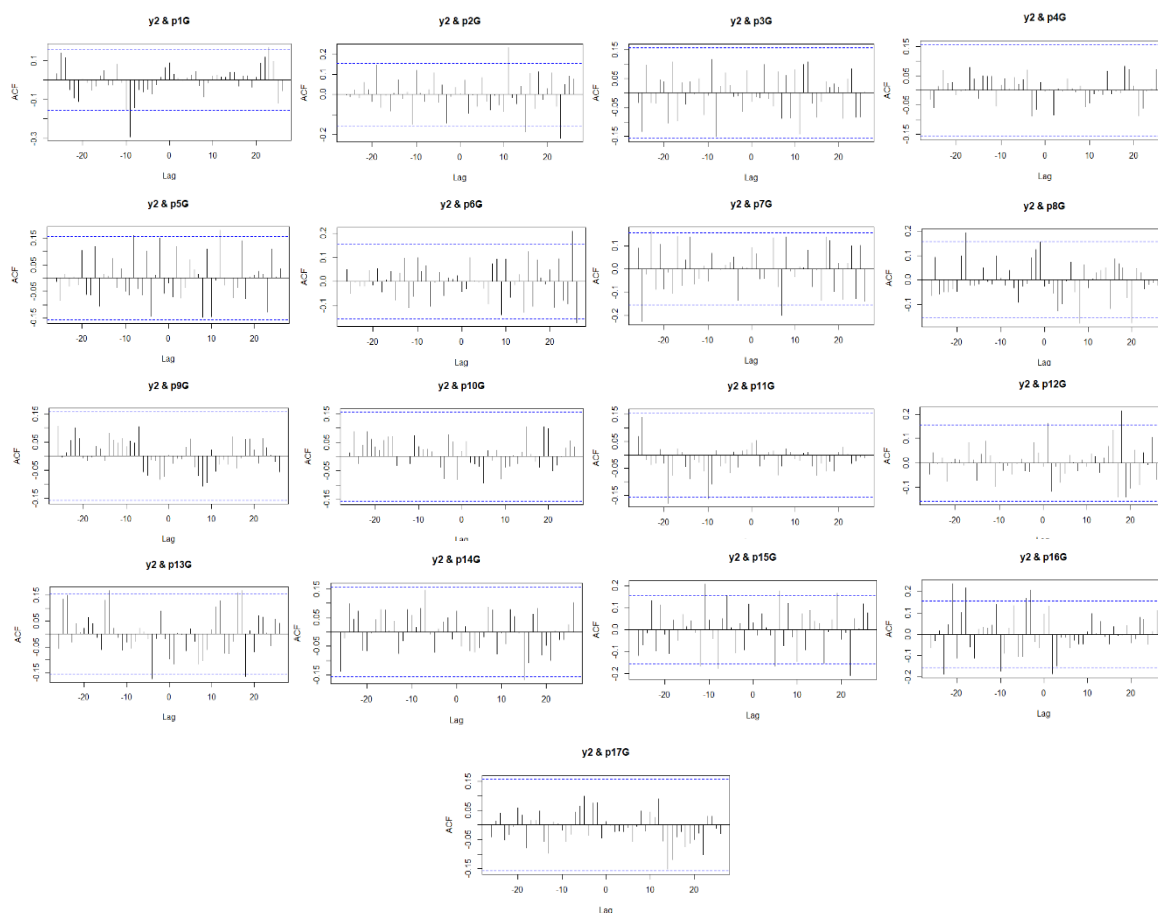


Figura 6: Gráficos de la función de correlación para búsquedas en Google frente a las variaciones en el volumen de intercambio

Fuente: Elaboración propia.

A partir de ellos es posible reconocer que cada combinación de dos series de tiempo mantiene niveles similares en magnitud, permaneciendo bajo el valor de 0,3. Por lo tanto alrededor del intervalo de tiempo seleccionado no se logra encontrar un desplazamiento para el cual las series de tiempo presenten una mayor correlación a la descrita en el formato de tabla.

6.1.2.2. Niveles de búsqueda de YouTube

Variable	Precio BTC		Variable	Vol. Interc. BTC	
	Corr. Spearman	Valor p		Corr. Spearman	Valor p
P14Y [Diff]	0,264	0,001***	P10Y [Diff]	0,161	0,044*
P17Y	0,252	0,002**	P4Y [Diff]	0,131	0,102
P1Y	0,18	0,024*	P13Y [Diff]	0,121	0,131
P9Y	0,162	0,043*	P5Y [Diff]	0,118	0,142
P15Y [Diff]	0,151	0,058.	P15Y [Diff]	0,113	0,159
P4Y [Diff]	0,12	0,134	P14Y [Diff]	0,086	0,286
P10Y [Diff]	0,09	0,263	P1Y	0,033	0,679
P2Y	0,044	0,583	P9Y	0,013	0,867
P8Y [Diff]	0,031	0,696	P7Y	0,003	0,971
P5Y [Diff]	0,027	0,738	P6Y [Diff]	-0,003	0,969
P3Y [Diff]	0,02	0,806	P12Y	-0,015	0,856
P16Y [Diff]	-0,042	0,605	P2Y	-0,027	0,741
P12Y	-0,043	0,593	P8Y [Diff]	-0,038	0,639
P11Y	-0,051	0,525	P17Y	-0,038	0,629
P6Y [Diff]	-0,077	0,338	P11Y	-0,058	0,474
P13Y [Diff]	-0,093	0,249	P3Y [Diff]	-0,061	0,451
P7Y	-0,115	0,150	P16Y [Diff]	-0,066	0,411
Significancia estadística: *** = 0,1% ; ** = 1% ; * = 5% ; . = 10%					

Tabla 6: Coeficientes de correlación para los niveles de búsqueda en YouTube

Fuente: Elaboración propia.

A modo de segunda parte del análisis de correlación se encuentran las variables con información derivada de los niveles de búsqueda en YouTube.

Al interior de las columnas dirigidas a la correlación frente al precio de Bitcoin se puede destacar que el mayor valor positivo lo posee la línea “P14Y [Diff]” (con 0,264), que alude al documental “The Rise and Rise of Bitcoin (2014)”. Mientras que la mayor magnitud de las correlaciones negativas es proporcionada en “P7Y” (con -0,155), del documental “Cryptopia: Bitcoin, Blockchains and the Future of the Internet (2020)”.

Abarcando la totalidad de las variables audiovisuales se aprecia que la mayoría de ellas posee una correlación menor a 0,20, lo cual equivale a decir que la asociación es muy débil. Tan solo dos variables se distinguen de dicha tendencia, siendo estas “P14Y [Diff]” y “P17Y”, que simbolizan al documental “The Rise and Rise of Bitcoin” y la serie de televisión “Forhøret (Face to face / Cara a cara)” respectivamente. Ambas poseen valores entre 0,20 y 0,35, que indican una correlación débil, pero considerable (Senthilnathan, 2019).

En cuanto a la significancia de los resultados, solo las variables 14, 17, 1, y 9 logran rechazar la hipótesis nula con una significancia de 5% o menos. Por esa razón se establece que las búsquedas de los títulos “The Rise and Rise of Bitcoin” (P14Y [Diff]) y “Forhøret (Face to face / Cara a cara)” (P17Y) poseen una correlación monótona significativa débil (aunque considerable), al igual que las producciones “How to Sell Drugs Online (Fast)” (P1Y) y “Hidden Secrets of Money” (P9Y) poseen una relación significativa, pero estas últimas presentan una asociación menos fuerte.

Posteriormente, los resultados que se encuentran en las columnas destinadas al volumen de intercambio describen que la correlación positiva más alta corresponde a la serie de televisión “Keiser Report (2019)” (P10Y [Diff]), con un valor de 0,161. Al igual que la serie “Startup (2016)” (P16Y [Diff]) presenta la relación negativa de mayor valor absoluto, con -0,066.

Tal como en los otros casos detallados, la correlación de las variables consideradas no logra superar el umbral de la magnitud 0,20. Por tal razón, las relaciones estudiadas poseen una intensidad muy débil.

Finalmente, en esta columna solo se logra obtener un coeficiente que cumple con tener una significancia del 5%, rechazando así la hipótesis nula. Por consiguiente, su variable

(“P10Y [Diff]”) permite determinar que la serie de televisión “Keiser Report” posee una correlación monótona significativa con el volumen de intercambio de Bitcoin.

Al igual que con la información de Google, a continuación, se muestran los gráficos de la función de correlación para un máximo de 26 semanas de desplazamiento entre ambas series de tiempo.

Primero para la correlación entre los retornos de Bitcoin y con los niveles de búsquedas en Google:

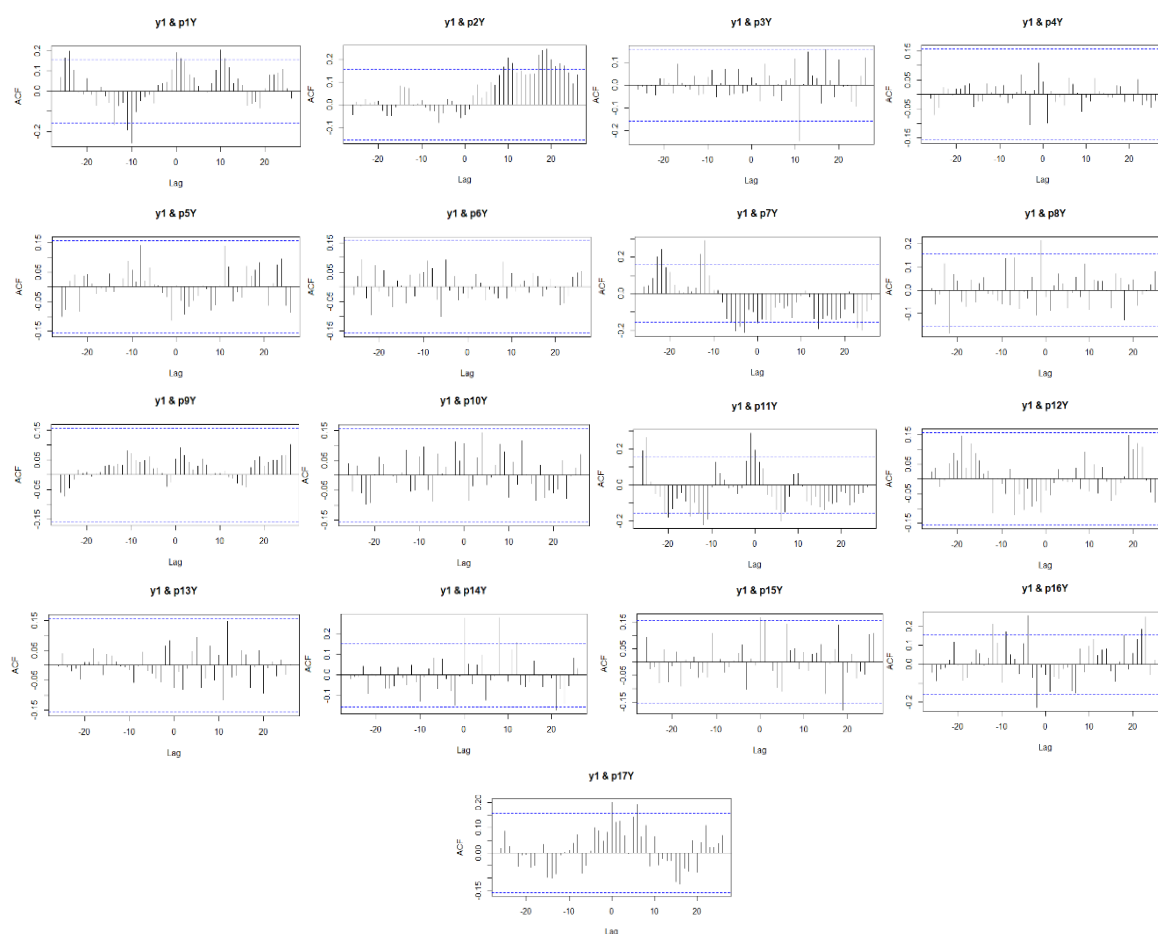


Figura 7: Gráficos de la función de correlación para búsquedas en YouTube frente a los retornos

Fuente: Elaboración propia.

Luego entre los cambios en el volumen de intercambio y el nivel de búsqueda en Google:

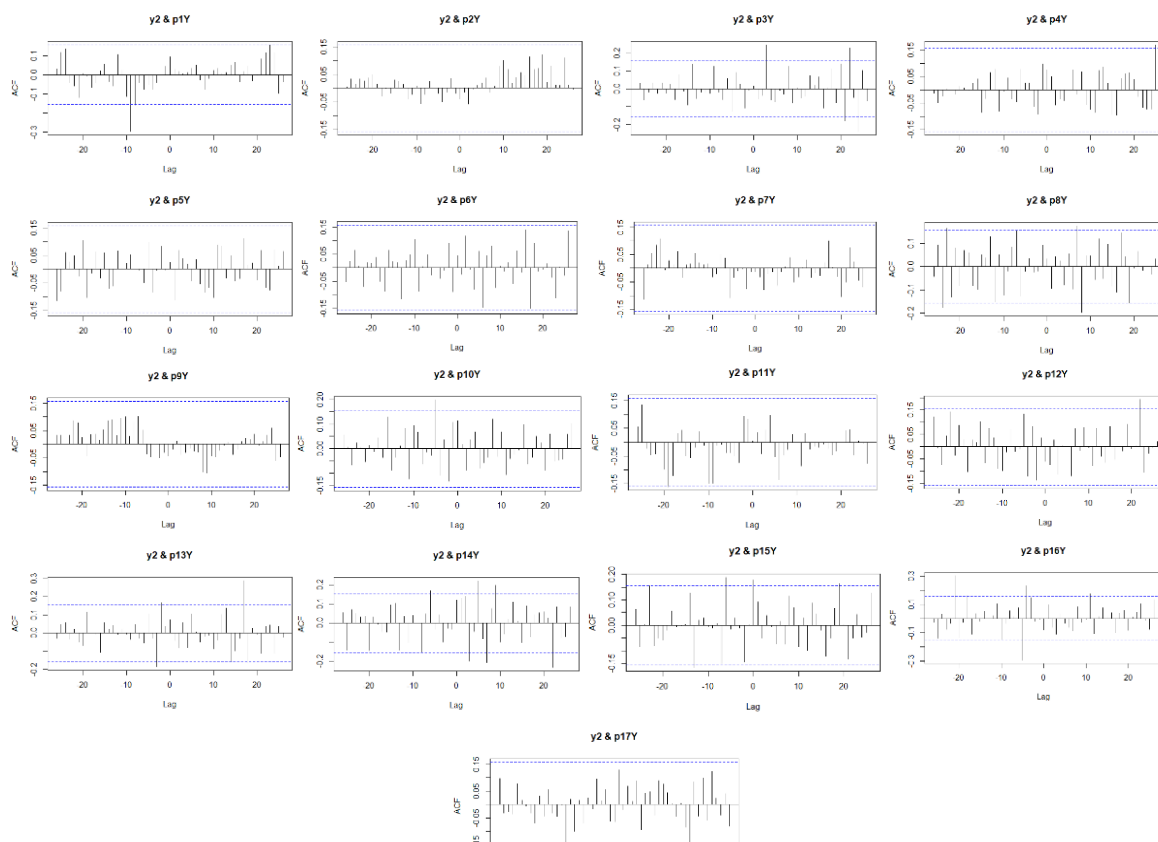


Figura 8: Gráficos de la función de correlación para búsquedas en YouTube frente a las variaciones en el volumen de intercambio
Fuente: Elaboración propia.

De esa forma se aprecia que para los datos de YouTube la correlación también presenta valores máximos similares a los encontrados sin desplazamiento, por lo tanto, se diagnostica que no existe un desplazamiento o desfase para el cual haya una mayor correlación en términos generales en comparación a la obtenida en el cálculo inicial.

6.1.2.3. Comparación de ambos casos

En síntesis, el análisis de correlación que utiliza datos de Google solo posee una relación significativa con dos títulos, mientras que YouTube presenta cinco vínculos de ese tipo. En consecuencia, las búsquedas en YouTube demuestran estar más asociadas con las producciones audiovisuales sobre criptomonedas. Más aún cuando se toma en cuenta la magnitud de su correlación, ya que alcanza valores iguales o más altos que aquellos presentes con los datos de Google.

6.2. Regresión lineal entre Bitcoin y YouTube

6.2.1. Verificación de supuestos

Los modelos iniciales para estimar el precio de Bitcoin y su volumen de intercambio son los siguientes:

$$\begin{aligned}
 \text{Precio_BTC} &= \beta_0 + \beta_1 \cdot (YT_Subs) + \beta_2 \cdot (YT_Views) + \beta_3 \cdot Cntrl_1 + \beta_4 \cdot Cntrl_2 + \varepsilon \\
 \text{Vol_Interc_BTC} &= \beta_0 + \beta_1 \cdot (YT_NwSubs) + \beta_2 \cdot (YT_Views) + \beta_3 \cdot Cntrl_1 + \beta_4 \cdot Cntrl_2 \\
 &+ \varepsilon
 \end{aligned}$$

Sin embargo, antes de poder analizar sus resultados es necesario verificar los ocho supuestos que indican su correcto funcionamiento.

- Supuesto 8: Series de tiempo estacionarias

Los resultados de las pruebas de Dickey-Fuller Aumentada y Phillips-Perron son sintetizadas en la siguiente tabla:

Variable	ADF		PP	
	Valor-p	Resultado	Valor-p	Resultado
Precio BTC	0,750	No estacionaria	0,919	No estacionaria
Vol. Interc. BTC	0,424	No estacionaria	0,045	Estacionaria
Subs YouTube	0,431	No estacionaria	0,388	No estacionaria
Views YouTube	0,835	No estacionaria	0,010	Estacionaria
CBOE VIX	0,224	No estacionaria	0,032	Estacionaria
DJIA	0,850	No estacionaria	0,635	No estacionaria

Tabla 7: Pruebas ADF y PP para el modelo con variables de YouTube

Fuente: Elaboración propia.

Se observa en los valores p que en la prueba ADF ninguna de las series de tiempo logra rechazar la hipótesis nula y por ende no son estacionarias. Mientras que en la prueba de PP el volumen de intercambio, las visualizaciones de YouTube y CBOE VIX sí son estacionarias con 5% de significancia. Como se busca que haya un consenso entre ambas pruebas se decide calcular las primeras diferencias de todas las variables y aplicar las pruebas nuevamente.

Variable	ADF		PP	
	Valor-p	Resultado	Valor-p	Resultado
Retornos BTC	0,010	Estacionaria	0,010	Estacionaria
Δ Vol. Interc. BTC	0,010	Estacionaria	0,010	Estacionaria
Δ Subs YouTube	0,010	Estacionaria	0,010	Estacionaria
Δ Views YouTube	0,010	Estacionaria	0,010	Estacionaria
Δ CBOE VIX	0,010	Estacionaria	0,010	Estacionaria
Δ DJIA	0,010	Estacionaria	0,010	Estacionaria

Tabla 8: Pruebas ADF y PP para el modelo con variables de YouTube luego de aplicar las transformaciones necesarias

Fuente: Elaboración propia.

Con esta transformación los valores p de ambas pruebas demuestran que todas las series de tiempo son estacionarias con un 1% de significancia estadística. Por lo tanto, es posible continuar con la verificación de supuestos.

Notar que la variable del precio de Bitcoin ahora es considerada los retornos de la criptomoneda.

- Supuesto 1: Linealidad de los parámetros

A continuación, se presentan los gráficos de dispersión entre la variable del precio de Bitcoin frente a cada variable independiente:

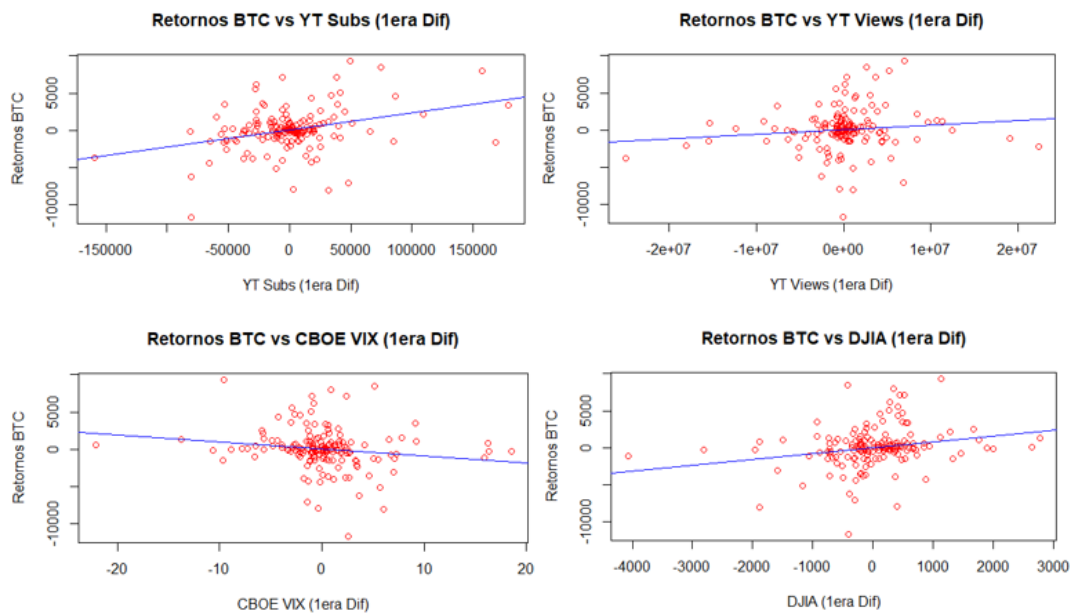


Figura 9: Gráficos de dispersión - Modelo con variables de YouTube para estimar retornos

Fuente: Elaboración propia.

En ellos se aprecia que la relación de ellas con la variable dependiente es lo suficientemente lineal, en vista de que los valores siguen la línea de tendencia, alrededor de la que hay una cantidad similar de puntos por sobre y debajo de ella. Por lo tanto, se cumple el supuesto de linealidad de los parámetros.

A modo de segunda verificación se adjunta el gráfico de los residuos en función de los valores estimados:

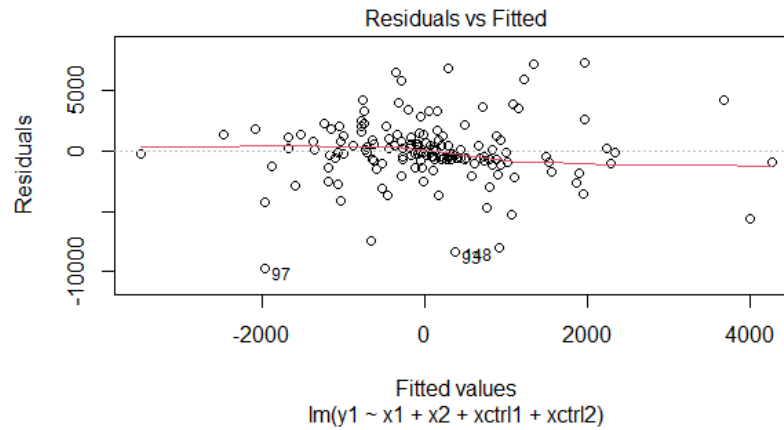


Figura 10: Gráfico de residuos en función de valores estimados - Modelo con variables de YouTube para estimar retornos

Fuente: Elaboración propia.

Tal representación logra mostrar que la totalidad del modelo sí logra describir una relación lineal entre el conjunto de variables, ya que los puntos se distribuyen alrededor de la horizontal y la curva roja se asemeja lo suficiente a una línea recta.

Por otro lado, la variable dependiente del volumen de intercambio también es analizada frente a las demás variables, obteniéndose los siguientes gráficos:

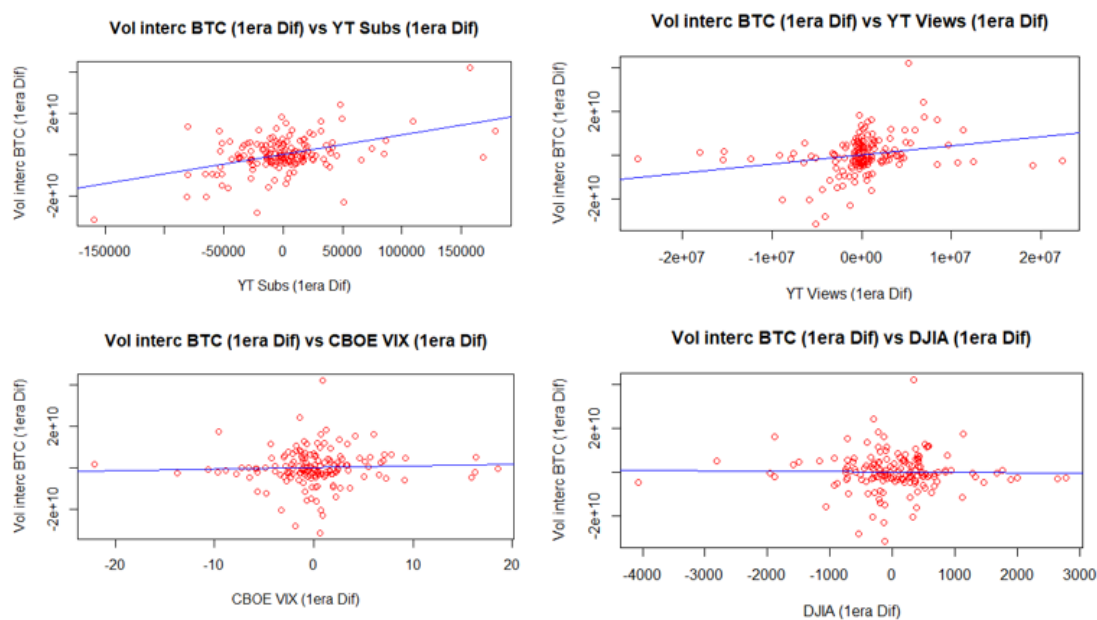


Figura 11: Gráficos de dispersión - Modelo con variables de YouTube para estimar diferencias del volumen de intercambio

Fuente: Elaboración propia.

Esto indica que las variables explicativas sí poseen una relación lineal con el volumen de intercambio de Bitcoin, por lo tanto, el supuesto se cumple.

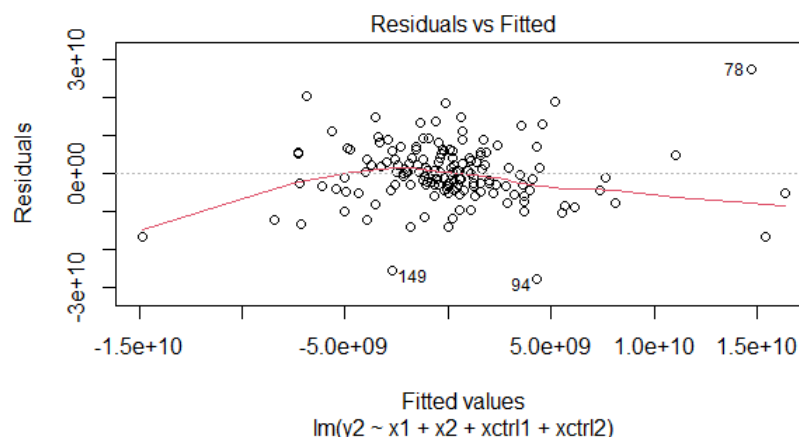


Figura 12: Gráfico de residuos en función de valores estimados - Modelo con variables de YouTube para estimar las diferencias en el volumen de intercambio

Fuente: Elaboración propia.

Luego, como segunda verificación se grafican los residuos frente a los valores estimados, obteniéndose que la relación es en gran parte lineal, con la excepción de la presencia de un valor extremo. Se decide conservar dicha observación puesto que se considera que aporta información útil para la evaluación del desempeño del modelo.

- Supuesto 2: Valor esperado de los residuos igual a cero

El cálculo de la media de los residuos para ambos modelos se muestra en la siguiente tabla:

Variable	Valor esperado de los residuos
Regresión del Precio BTC	$-6,129 \times 10^{-15}$
Regresión del Vol. de Interc.	$-1,726 \times 10^{-8}$

Tabla 9: Cálculo del valor esperado de los residuos para ambos modelos con variables de YouTube

Fuente: Elaboración propia.

Se aprecia que ambos valores son muy cercanos a cero, por lo tanto, se cumple el supuesto para ambos modelos

- Supuesto 3: Condición de homocedasticidad en los residuos

La aplicación de la prueba de Breusch-Pagan entrega los siguientes resultados:

Variable dependiente	Breusch-Pagan	
	Valor p	Conclusión
Precio BTC	0,126	No se rechaza H0
Vol. de Interc.	0,075	No se rechaza H0

Tabla 10: Resultados de la prueba de Breusch-Pagan para ambos modelos

Fuente: Elaboración propia.

En ambos casos no es posible rechazar la hipótesis nula, por lo tanto, ambas regresiones presentan homocedasticidad y se cumple el supuesto.

- Supuesto 4: No hay autocorrelación entre residuos

Los resultados obtenidos para la prueba de Durbin-Watson se presenta a continuación:

Variable Dependiente	Durbin-Watson		
	Estadístico	Valor-p	Resultado
Precio BTC	1,663	0,019	Rechaza H0
Vol. de Interc.	2,352	0,988	No rechaza H0

Tabla 11: Resultados de la prueba de Durbin-Watson para ambos modelos con variables de YouTube

Fuente: Elaboración propia.

La prueba de Durbin-Watson arroja que la regresión del precio de Bitcoin sí presenta auto correlación, mientras que para el caso del volumen de intercambio no hay auto correlación de primer orden.

De esa forma, el supuesto sobre la ausencia de autocorrelación no se cumple para la estimación del precio, por lo que al momento de interpretar los coeficientes de regresión se

corregirán sus errores estándar con el método de Newey-West. Mientras que para el modelo del volumen de intercambio no es necesario hacer correcciones.

Al aplicar tal procedimiento al primer modelo se mitigan los inconvenientes generados por la autocorrelación, pudiendo analizar las pruebas de hipótesis correctamente.

- Supuesto 5: Variables independientes no estocásticas

Para comprobar este supuesto se calcula la covarianza entre los residuos y cada variable independiente. Los valores resultantes se presentan a continuación:

Variable	Precio de BTC	Volumen de Intercambio BTC
	Covarianza	Covarianza
$\Delta\text{Subs YouTube}$	$-5,350 \times 10^{-10}$	-0,038
$\Delta\text{Views YouTube}$	$-5,954 \times 10^{-8}$	0,147
$\Delta\text{CBOE VIX}$	$-1,167 \times 10^{-13}$	$1,306 \times 10^{-7}$
ΔDJIA	$1,616 \times 10^{-13}$	0,000

Tabla 12: Cálculo de la covarianza entre los residuos de cada modelo con variables de YouTube y las variables independientes de estos

Fuente: Elaboración propia.

En el caso de la regresión del precio de Bitcoin la covarianza entre residuos y variables independientes es aproximadamente cero para todas las variables, por ende, el supuesto sí se cumple.

No obstante, para el volumen de intercambio los valores también son muy cercanos a cero, pero las variables sobre suscriptores y visualizaciones de YouTube poseen una covarianza mayor en comparación a sus pares. Por esa razón se concluye que para la regresión del volumen de intercambio este supuesto se cumple parcialmente, generando que los resultados tengan la limitación de estar ligeramente condicionados a los datos utilizados en este estudio.

- Supuesto 6: Residuos con distribución normal

Con el objetivo de comprobar si los residuos poseen una distribución normal se utiliza la prueba de Jarque-Bera sobre ellos. Estos son sus resultados:

Modelo	Jarque-Bera	
	Valor p	Conclusión
Precio BTC	$6,373 \times 10^{-12}$	Se rechaza H0
Vol. de Interc.	$1,997 \times 10^{-5}$	Se rechaza H0

Tabla 13: Resultados de la prueba de Jarque-Bera para los residuos de ambos modelos con variables de YouTube

Fuente: Elaboración propia.

Tomando en cuenta la tabla anterior, ambas regresiones muestran tener valores p capaces de rechazar la hipótesis nula de la prueba con más de un 0,1% de significancia. Esto quiere decir que sus residuos no poseen una distribución normal.

Ante esto se recurre al teorema del límite central, el cual señala que para muestras grandes los estadísticos t y F se distribuyen asintóticamente a sus distribuciones, aun cuando los residuos no se distribuyen de forma normal. Como la muestra considerada para estas regresiones contiene 158 observaciones aproximadamente (tomando en cuenta que con las primeras diferencias se pierde la primera) se decide continuar con los modelos bajo el supuesto de muestras grandes. Con ello no es necesario cumplir con el supuesto de residuos con distribución normal y las pruebas de hipótesis para los coeficientes de regresión pueden ser realizadas.

- Supuesto 7: No existe multicolinealidad

Con propósito de examinar si las variables independientes presentan multicolinealidad entre sí se elabora una matriz de correlación. El coeficiente de correlación que se utilizará para este procedimiento dependerá de si se cumplen los supuestos necesarios para aplicar el coeficiente de correlación de Pearson, de lo contrario se optará por una

alternativa robusta, el coeficiente de correlación de Spearman. Para ello se utiliza la prueba de Jarque-Bera sobre las cuatro variables independientes.

Modelo	Jarque-Bera	
	Valor p	Conclusión
Δ Subs YouTube	$2,2 \times 10^{-16}$	Se rechaza H0
Δ Views YouTube	$2,2 \times 10^{-16}$	Se rechaza H0
Δ CBOE VIX	$2,2 \times 10^{-16}$	Se rechaza H0
Δ DJIA	$2,2 \times 10^{-16}$	Se rechaza H0

Tabla 14: Resultados de la prueba de Jarque-Bera para cada variable independiente de los modelos sobre YouTube

Fuente: Elaboración propia.

Es posible apreciar que los valores p son prácticamente cero, por lo que en los cuatro casos se rechaza la hipótesis nula. A partir de tal resultado se concluye que ninguna de estas variables posee una distribución normal, provocando que el coeficiente de correlación de Spearman sea el más apropiado para crear la matriz de correlación.

Variable	Δ Subs YouTube	Δ Views YouTube	Δ CBOE VIX	Δ DJIA
Δ Subs YouTube	1,000	0,480***	-0,068	0,122
Δ Views YouTube	0,480***	1,000	-0,057	0,127
Δ CBOE VIX	-0,068	-0,057	1,000	-0,679***
Δ DJIA	0,122	0,127	-0,679***	1,000
Significancia estadística: *** = 0,1% ; ** = 1% ; * = 5% ; . = 10%				

Tabla 15: Matriz de correlación para las variables independientes de los modelos sobre YouTube

Fuente: Elaboración propia.

La matriz de correlación revela que solo hay dos pares de variables que poseen valores considerables de correlación, al mismo tiempo que poseen suficiente significancia estadística. El primer caso son las variables de suscriptores en YouTube y las visualizaciones de la misma plataforma, que tienen un coeficiente igual a 0,480, pudiendo ser etiquetada como una asociación moderada o considerable según la escala de (Senthilnathan, 2019). Sin

embargo, cabe mencionar que la colinealidad se considera perjudicial cuando la correlación posee magnitud demasiado alta, es decir, por sobre la cifra de 0,80.

Además, es de esperarse que ambas variables tengan cierta correlación entre sí, debido a que provienen de la misma fuente y están ligadas a un mismo efecto: el consumo de contenido audiovisual en YouTube. Por esa razón se considera que tener estos resultados es aceptable, sin representar un inconveniente relativo a una colinealidad dentro de los modelos regresivos.

El segundo par de variables que presentan correlación medianamente alta son las variables de control: el índice de volatilidad CBOE VIX (Chicago Board Options Exchange Market Volatility Index) y el promedio industrial de Dow Jones (DJIA). Entre ambas poseen un coeficiente de correlación de Spearman igual a -0,679, con alta significancia estadística. Tal cifra puede ser categorizada como considerablemente alta según (Senthilnathan, 2019). Aunque, tal como ocurre con el otro par de variables explicativas, no es un valor que sea suficiente para afectar de forma severa los resultados de la regresión. En vista de que el valor del coeficiente es menor a 0,80.

En conclusión, ambos casos donde existe correlación significativa no presentan un grado de asociación que impida el cumplimiento de este supuesto. En otras palabras, no se está en presencia de colinealidad.

6.2.2. Modificaciones en los modelos

A partir de la revisión de estos ocho supuestos se determina que los modelos a estimar son ahora los siguientes:

$$Retornos_BTC = \beta_0 + \beta_1 \Delta YT_Subs + \beta_2 \Delta YT_Views + \beta_3 \Delta Cntrl_1 + \beta_4 \Delta Cntrl_2 + \varepsilon$$

$$\Delta BTC_Vol_Interc = \beta_0 + \beta_1 \Delta YT_NwSubs + \beta_2 \cdot \Delta YT_Views + \beta_3 \Delta Cntrl_1 + \beta_4 \Delta Cntrl_2 + \varepsilon$$

Donde el símbolo “ Δ ” denota que ahora se trabaja con las primeras diferencias de las variables, que es el principal cambio realizado en comparación al estado del modelo previo a la revisión de supuestos.

Sin embargo, este análisis también revela que el modelo para predecir los retornos posee autocorrelación, por lo tanto, sus estimadores no logran cumplir con la propiedad de ser los estimadores lineales de menor varianza (eficientes).

6.2.3. Resultados de la regresión

Luego de haber comprobado cuáles supuestos logran cumplirse se es capaz de tener una vasta comprensión de cómo está formulado el modelo y cuáles propiedades se manifiestan, razón por la cual resulta apropiado analizar las estimaciones obtenidas.

6.2.3.1. Regresión de los retornos

```
Call:
lm(formula = y1 ~ x1 + x2 + xctrl1 + xctrl2)

Residuals:
    Min       1Q   Median       3Q      Max
-9701.3  -797.7  -118.5   1103.4   7337.2

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.002e+01  2.088e+02   0.144  0.885864
x1           2.121e-02  5.521e-03   3.843  0.000178 ***
x2           1.181e-05  4.335e-05   0.273  0.785607
xctrl1       1.825e+00  6.397e+01   0.029  0.977281
xctrl2       7.290e-01  3.795e-01   1.921  0.056606 .
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2609 on 152 degrees of freedom
Multiple R-squared:  0.1542,    Adjusted R-squared:  0.1319
F-statistic: 6.925 on 4 and 152 DF,  p-value: 3.789e-05
```

Figura 13: Resumen del modelo lineal con variables de YouTube para estimar retornos
Fuente: Elaboración propia.

La imagen muestra el modelo lineal referente al precio de Bitcoin (y_1), donde “ x_1 ” es la variable de suscriptores de YouTube, “ x_2 ” simboliza las visualizaciones del mismo sitio web y “ x_{ctrl1} ” junto a “ x_{ctrl2} ” son las variables de control.

De dichos resultados se desprende que el coeficiente de regresión que acompaña a la variable de suscriptores es igual a 0,021 dólares por suscripción, lo cual se traduce en que con un aumento de 1.000 suscriptores de una semana a otra (primera diferencia) se produce un cambio de 21 dólares estadounidenses dentro de los retornos semanales estimados de Bitcoin.

Posteriormente, al prestar atención en la variable de visualizaciones de YouTube, se observa que el coeficiente estimado es de 0,00001 dólares por visualización. Esto quiere decir que la relación es positiva y por cada aumento de 100.000 visualizaciones obtenidas semanalmente (primera diferencia) en canales sobre criptomonedas los retornos estimados de Bitcoin se incrementan en un dólar.

Además, el coeficiente de determinación ajustado (R^2) es igual a 0,1319, señalando así que las variables incluidas en el modelo logran describir un 13,19% de la variación presente en los retornos de Bitcoin. Si bien es un porcentaje bajo en comparación al total de las variaciones, se considera que es una magnitud considerable al tener en cuenta que este modelo solo incluye cuatro variables, razón por la cual es posible que otras variables relevantes no hayan sido incluidas. Entonces si este modelo logra aportar tal poder para explicar las variaciones implica que el aporte de las variables de YouTube es importante.

A modo de verificación sobre si la regresión es espuria se comprueba que el coeficiente de correlación de este modelo es menor al valor que posee el estadístico de Durbin-Watson que se calculó dentro del supuesto de autocorrelación (con magnitud 1,663).

También se aprecia que el estadístico F posee un valor p es igual a $3,789 * 10^{-5}$, por lo tanto se rechaza la hipótesis nula con todos los coeficientes iguales a cero, generando que el modelo sea válido.

En cuanto a la significancia estadística de estos resultados, no es apropiado guiarse por los valores p de la imagen anterior, ya que no toma en consideración los efectos de la autocorrelación encontrada en la verificación de supuestos.

De esa manera se recurre a la corrección de errores estándar según el método de Newey-West. Los valores p obtenidos se disponen a continuación:

```
t test of coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.0017e+01 2.3906e+02  0.1256 0.900247
x1           2.1214e-02 6.5716e-03  3.2281 0.001527 **
x2           1.1814e-05 2.9955e-05  0.3944 0.693843
xctrl1       1.8248e+00 3.9336e+01  0.0464 0.963061
xctrl2       7.2901e-01 3.2489e-01  2.2439 0.026286 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figura 14: Valores p corregidos según el método de Newey-West – Modelo con variables de YouTube para estimar retornos

Fuente: Elaboración propia.

Con ello es posible señalar que la relación descrita entre el cambio semanal de la cantidad de suscriptores y los retornos de Bitcoin logra rechazar la hipótesis nula con un 1% de significancia. Sin embargo, no es posible decir lo mismo de los cambios en el número de visualizaciones, ya que su valor p no logra rechazar la hipótesis nula. Por ende, en este segundo caso afirmar que el coeficiente es distinto de cero conllevaría tener un error demasiado grande para ser considerado un resultado confiable.

Por último, cabe mencionar que las variables de control enseñan un comportamiento inusual con respecto a los resultados esperados, que se basan en lo obtenido dentro de otros

estudios. Por ejemplo, el índice CBOE VIX (xctrl1) presenta un coeficiente positivo, siendo que se tenían expectativas de confirmar su efecto negativo sobre los retornos (Bouri et al., 2017). Aunque este no presenta una significancia suficiente para negar la hipótesis nula.

6.2.3.2. Regresión de las variaciones del volumen de intercambio

De la misma forma se analizan los coeficientes de regresión obtenidos al utilizar a la primera diferencia del Volumen de Intercambio de Bitcoin como variable dependiente. En este caso el software R provee el siguiente resumen:

```
call:
lm(formula = y2 ~ x1 + x2 + xctrl1 + xctrl2)

Residuals:
    Min       1Q   Median       3Q      Max
-2.762e+10 -4.104e+09 -3.042e+08  4.253e+09  2.735e+10

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -4.175e+07  6.306e+08  -0.066    0.947
x1           8.800e+04  1.668e+04   5.277 4.45e-07 ***
x2           1.624e+02  1.310e+02   1.240    0.217
xctrl1       7.744e+07  1.932e+08   0.401    0.689
xctrl2      -9.314e+04  1.146e+06  -0.081    0.935
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.88e+09 on 152 degrees of freedom
Multiple R-squared:  0.2078,    Adjusted R-squared:  0.1869
F-statistic: 9.966 on 4 and 152 DF,  p-value: 3.451e-07
```

Figura 15: Resumen del modelo lineal con variables de YouTube para estimar diferencias en el volumen de intercambio

Fuente: Elaboración propia.

Donde “y2” representa el volumen de intercambio de Bitcoin, “x1” es la variable de suscriptores de YouTube, “x2” simboliza las visualizaciones del mismo sitio web y “xctrl1” junto a “xctrl2” son las variables de control.

Para primera variable de interés, las visualizaciones de YouTube (x1), se obtiene un coeficiente positivo de 88.000 dólares por visualización. Es decir, cada cambio unitario en la

cantidad de videos vistos por semana genera un efecto de dicho valor sobre la cantidad estimada de los montos en compras o ventas de Bitcoin (Volumen de intercambio).

Por otro lado, la cantidad de nuevos suscriptores presenta un coeficiente positivo de 162.400 dólares por suscripción. Implicando que por cada persona que se suscribe a un canal sobre criptomonedas 162.400 dólares son movidos en transacciones dentro del mercado de Bitcoin.

Seguidamente, al observar el valor del coeficiente de determinación ajustado (R^2) se evidencia un valor de 0,1869, que es semejante a decir que las variables incluidas en este modelo logran explicar un 18,69% de la variación del volumen de intercambio que presenta Bitcoin. Esto, al igual que con el modelo para estimar los retornos, figura como un resultado bajo pero considerable.

Además, el estadístico de Durbin Watson para esta estimación posee un valor de 2,3525 (según lo expuesto en el supuesto de autocorrelación), lo cual es mayor al R^2 comentado el párrafo anterior. Esto conduce a decir que no se observan síntomas de que esta regresión sea espuria.

También se presta atención al valor p del estadístico F, el cual es igual a $3,451 \times 10^{-7}$, siendo esto muy cercano a cero, por consiguiente se rechaza la hipótesis nula y el modelo es válido.

A diferencia de como ocurre con el modelo sobre los retornos de Bitcoin, en el modelo del volumen de intercambio no se requiere usar la corrección de errores estándar según el método de Newey-West. Sino que basta con interpretar aquellos valores p entregados en el resumen de la regresión.

En él se indica que el efecto de los cambios en el crecimiento de suscriptores sobre el volumen de intercambio de Bitcoin posee un valor p de $4,45 \cdot 10^{-7}$, que se traduce en el rechazo de la hipótesis nula con un 0,1% de significancia. Mientras que la cantidad de suscriptores adquiere un valor p igual a 0,217, sin ser suficiente para rechazar la hipótesis nula. De esa forma, no existe evidencia suficiente para afirmar con un error bajo que esta última relación sea verdadera.

Por último, en el caso del modelo para predecir las diferencias del volumen de intercambio se comprueba que la variable de control del promedio industrial de Dow Jones presenta una relación negativa con el volumen de intercambio de Bitcoin. Lo cual coincide con lo obtenido en otras publicaciones (Jerdack et al., 2018). Aunque el coeficiente alcanzado no logra rechazar la hipótesis nula.

6.3. Regresión lineal entre Bitcoin y Twitch

6.3.1. Verificación de Supuestos

Los modelos iniciales para estimar el precio de Bitcoin y su volumen de intercambio son los siguientes:

$$Precio_BTC = \beta_0 + \beta_1 \cdot (TW_AvgVwrs) + \beta_2 \cdot (TW_Channls) + \beta_3 \cdot Cntrl_1 + \beta_4 \cdot Cntrl_2 + \varepsilon$$

$$Vol_Interc_BTC = \beta_0 + \beta_1 \cdot (TW_AvgVwrs) + \beta_2 \cdot (TW_Channls) + \beta_3 \cdot Cntrl_1 + \beta_4 \cdot Cntrl_2 + \varepsilon$$

Sin embargo, antes de poder analizar sus resultados es necesario verificar los ocho supuestos que indican su correcto funcionamiento.

- Supuesto 8: Series de tiempo estacionarias

Los resultados de las pruebas de Dickey-Fuller Aumentada y Phillips-Perron son sintetizadas en la siguiente tabla:

Variable	ADF		PP	
	Valor-p	Resultado	Valor-p	Resultado
Precio BTC	0,750	No estacionaria	0,918	No estacionaria
Vol. Interc. BTC	0,424	No estacionaria	0,045	Estacionaria
Prom. Espect. TW	0,010	Estacionaria	0,471	No estacionaria
Canales Prom. TW	0,344	No estacionaria	0,636	No estacionaria
CBOE VIX	0,224	No estacionaria	0,032	Estacionaria
DJIA	0,850	No estacionaria	0,635	No estacionaria

Tabla 16: Pruebas ADF y PP para el modelo con variables de Twitch

Fuente: Elaboración propia.

Los valores p correspondientes a la prueba ADF indica que solo la serie de tiempo de los espectadores promedio en Twitch logra ser estacionaria con 5% de significancia. Sin embargo, para la misma variable la prueba de Phillips Perron establece que esta no es estacionaria. Luego, al fijarse en la prueba PP, surgen el mismo tipo de contradicciones. Las variables del volumen de intercambio y CBOE VIX asemejan ser estacionarias con 5% de significancia, pero la prueba ADF rechaza tal afirmación. Debido a esta falta de consenso entre ambas pruebas se decide transformar las variables a sus primeras diferencias y volver a realizar las pruebas.

Variable	ADF		PP	
	Valor-p	Resultado	Valor-p	Resultado
Retornos BTC	0,010	Estacionaria	0,010	Estacionaria
Δ Vol. Interc. BTC	0,010	Estacionaria	0,010	Estacionaria
Δ Prom. Espect. TW	0,010	Estacionaria	0,010	Estacionaria
Δ Canales Prom. TW	0,010	Estacionaria	0,010	Estacionaria
Δ CBOE VIX	0,010	Estacionaria	0,010	Estacionaria
Δ DJIA	0,010	Estacionaria	0,010	Estacionaria

Tabla 17: Pruebas ADF y PP para el modelo con variables de Twitch luego de aplicar las transformaciones necesarias

Fuente: Elaboración propia.

Con esta modificación los valores p de ambas pruebas demuestran que todas las series de tiempo son estacionarias con un 1% de significancia estadística. Por lo tanto, es posible continuar con la verificación de supuestos.

Notar que la variable del precio de Bitcoin ahora es considerada los retornos de la criptomoneda.

- Supuesto 1: Linealidad de los parámetros

A continuación, se presentan los gráficos de dispersión entre la variable del precio de Bitcoin frente a cada variable independiente:

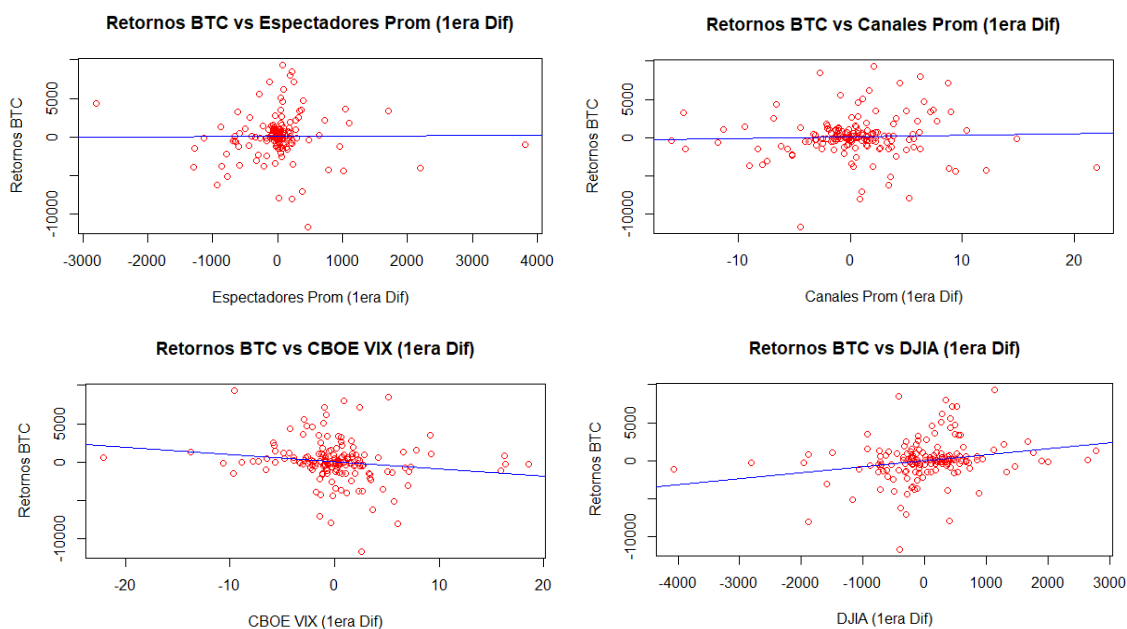


Figura 16: Gráficos de dispersión - Modelo con variables de Twitch para estimar retornos

Fuente: Elaboración propia.

En ellos se aprecia que en su mayoría la relación de ellas con la variable dependiente se asemeja a una línea recta, alrededor de la que hay una cantidad similar de puntos por sobre y debajo de ella. Por lo tanto, se cumple el supuesto de linealidad de los parámetros para aquellas que tienen esas características. El único gráfico que no cumple es el que toma en cuenta los retornos de BTC y los espectadores promedio. En él se muestra una alta concentración de las observaciones, que no permite que se pueda reconocer una relación de

línea recta. Por tal razón, se decide aplicar logaritmos naturales, para verificar si de esa forma la relación toma una apariencia lineal:

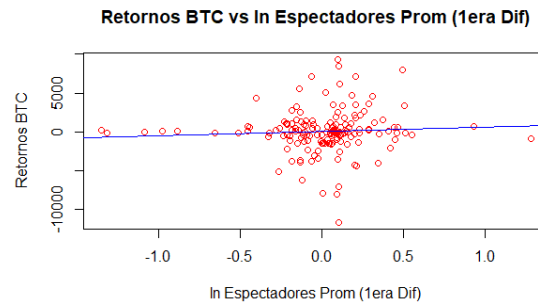


Figura 17: Gráfico de dispersión - Aplicación de logaritmos al número de espectadores de Twitch al estimar retornos

Fuente: Elaboración propia.

La imagen resultante demuestra que al aplicar tal función se logra que la variable de espectadores promedio en Twitch se relacione de forma lineal con los retornos de Bitcoin.

Para confirmar el resultado anterior se adjunta el gráfico de los residuos en función de los valores estimados:

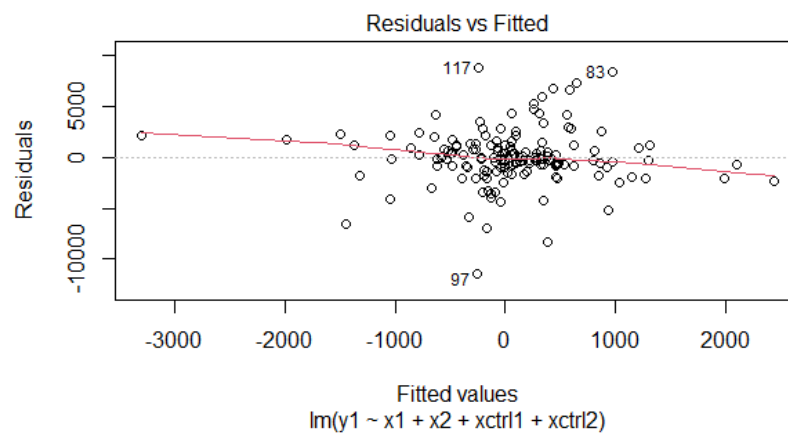


Figura 18: Gráfico de residuos en función de valores estimados - Modelo con variables de Twitch para estimar retornos

Fuente: Elaboración propia.

Tal representación logra mostrar que la totalidad del modelo sí logra describir una relación lineal entre las variables incluidas, ya que los puntos se distribuyen alrededor de la horizontal y la curva roja se asemeja lo suficiente a una línea recta.

Por otro lado, la variable dependiente del volumen de intercambio también es analizada frente a las demás variables, obteniéndose los siguientes gráficos:

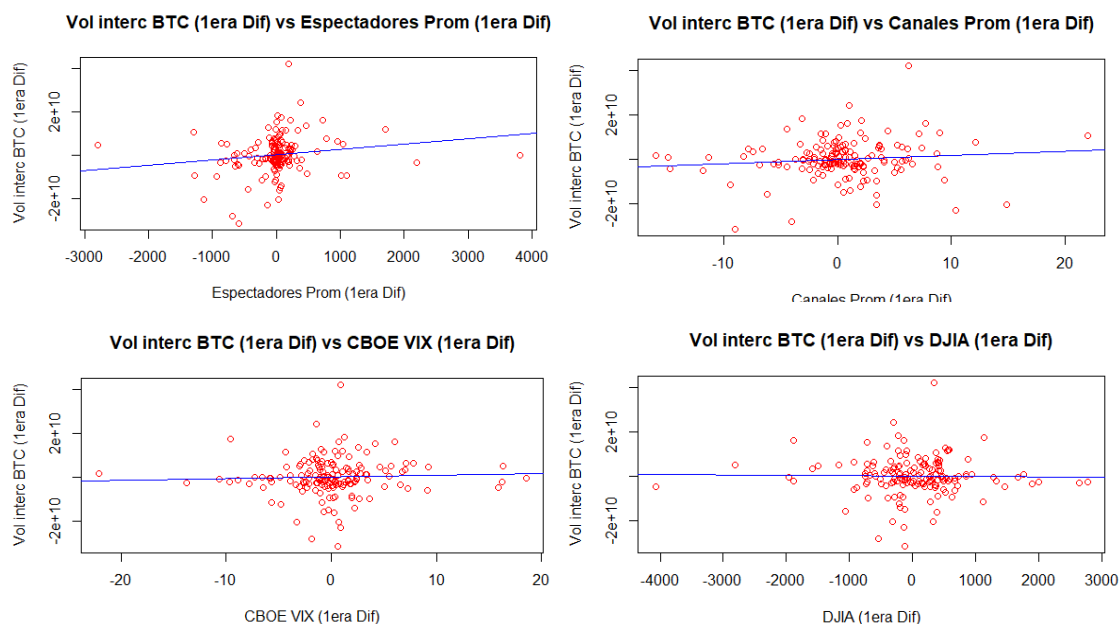


Figura 19: Gráficos de dispersión - Modelo con variables de Twitch para estimar diferencias del volumen de intercambio
Fuente: Elaboración propia.

Nuevamente se observa que la mayoría de las variables poseen una relación lineal con el volumen de intercambio de Bitcoin, con la excepción del número de espectadores

promedio. Esto causa que se opte por también utilizar la función de logaritmo natural sobre esta variable:

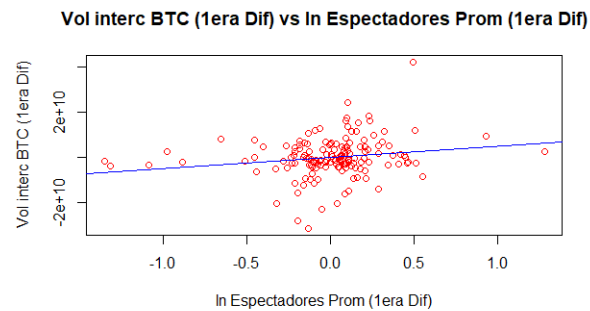


Figura 20: Gráfico de dispersión - Aplicación de logaritmos al número de espectadores de Twitch al estimar diferencias en el volumen de intercambio

Fuente: Elaboración propia.

Con ello la relación se confirma como lineal, por ende, en esas condiciones sí se cumple la totalidad del supuesto.

Luego, como segunda verificación se grafican los residuos frente a los valores estimados:

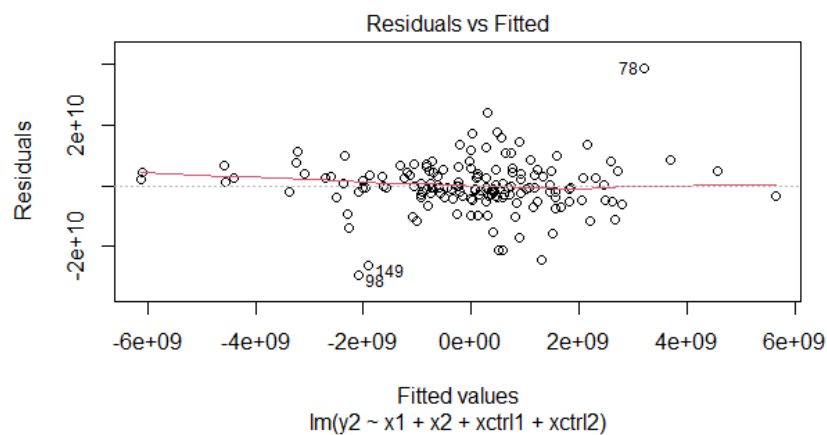


Figura 21: Gráfico de residuos en función de valores estimados - Modelo con variables de Twitch para estimar las diferencias en el volumen de intercambio

Fuente: Elaboración propia.

Este gráfico enseña que el conjunto de variables dentro del modelo logra describir una relación lineal, dado que la recta es cercana se asemeja a la horizontal y no posee curvas.

- Supuesto 2: Valor esperado de los residuos igual a cero

El cálculo de la media de los residuos para ambos modelos se muestra en la siguiente tabla:

Variable	Valor esperado de los residuos
Regresión del Retornos BTC	$-3,270 \times 10^{-14}$
Regresión del Vol. de Interc.	$1,166 \times 10^{-7}$

Tabla 18: Cálculo del valor esperado de los residuos para ambos modelos con variables de Twitch

Fuente: Elaboración propia.

Se aprecia que ambos valores son muy cercanos a cero, por lo tanto, se cumple el supuesto para ambos modelos

- Supuesto 3: Condición de homocedasticidad en los residuos

La aplicación de la prueba de Breusch-Pagan entrega los siguientes resultados:

Variable dependiente	Breusch-Pagan	
	Valor p	Conclusión
Retornos BTC	0,072	No se rechaza H0
Vol. de Interc.	0,058	No se rechaza H0

Tabla 19: Resultados de la prueba de Breusch-Pagan para ambos modelos sobre Twitch

Fuente: Elaboración propia.

En ambos casos no es posible rechazar la hipótesis nula, por lo tanto, ambas regresiones presentan homocedasticidad y se cumple el supuesto.

- Supuesto 4: No hay autocorrelación entre residuos

Los resultados obtenidos para la prueba de Durbin-Watson se presenta a continuación:

Variable Dependiente	Durbin-Watson		
	Estadístico	Valor-p	Resultado
Retornos BTC	1,545	0,002	Rechaza H0
Vol. de Interc.	2,317	0,977	No rechaza H0

Tabla 20: Resultados de la prueba de Durbin-Watson para ambos modelos con variables de Twitch

Fuente: Elaboración propia.

La prueba de Durbin-Watson arroja que la regresión de los retornos de Bitcoin sí presenta auto correlación, al mismo tiempo que para el caso del volumen de intercambio no hay auto correlación de primer orden.

De esa forma, el supuesto sobre la ausencia de autocorrelación no se cumple para el primer modelo, pero sí para el segundo. Por consiguiente, al momento de interpretar los coeficientes de regresión al estimar los retornos de Bitcoin se corregirán los errores estándar con el método de Newey-West. Mientras que para estimar el volumen de intercambio no es necesario aplicarlo.

- Supuesto 5: Variables independientes no estocásticas

Para comprobar este supuesto se calcula la covarianza entre los residuos y cada variable independiente. Los valores resultantes se presentan a continuación:

Variable	Precio de BTC	Volumen de Intercambio BTC
	Covarianza	Covarianza
Δ Prom. Espect. TW	$1,030 \times 10^{-14}$	$-7,127 \times 10^{-7}$
Δ Canales Prom. TW	$4,480 \times 10^{-13}$	$-7,850 \times 10^{-7}$
Δ CBOE VIX	$1,920 \times 10^{-13}$	$2,599 \times 10^{-7}$
Δ DJIA	$8,240 \times 10^{-11}$	$0,001 \times 10^{-1}$

Tabla 21: Cálculo de la covarianza entre los residuos de cada modelo con variables de Twitch y las variables independientes de estos

Fuente: Elaboración propia.

En ambas regresiones la covarianza entre residuos y variables independientes es aproximadamente cero para todas las variables, por ende, el supuesto sí se cumple.

- Supuesto 6: Residuos con distribución normal

Con el objetivo de comprobar si los residuos poseen una distribución normal se utiliza la prueba de Jarque-Bera sobre ellos. Estos son sus resultados:

Modelo	Jarque-Bera	
	Valor p	Conclusión
Retornos BTC	$3,220 \times 10^{-15}$	Se rechaza H0
Vol. de Interc.	$2,200 \times 10^{-16}$	Se rechaza H0

Tabla 22: Resultados de la prueba de Jarque-Bera para los residuos de ambos modelos con variables de Twitch

Fuente: Elaboración propia.

Tomando en cuenta la tabla anterior, ambas regresiones muestran tener valores p capaces de rechazar la hipótesis nula de la prueba con más de un 0,1% de significancia. Esto quiere decir que sus residuos no poseen una distribución normal.

Ante esto se recurre al teorema del límite central, el cual señala que para muestras grandes los estadísticos t y F se distribuyen asintóticamente a sus distribuciones, aun cuando

los residuos no se distribuyen de forma normal. Como la muestra considerada para estas regresiones contiene 158 observaciones aproximadamente (tomando en cuenta que con las primeras diferencias se pierde la primera) se decide continuar con los modelos bajo el supuesto de muestras grandes. Con ello no es necesario cumplir con el supuesto de residuos con distribución normal y las pruebas de hipótesis para los coeficientes de regresión pueden ser realizadas.

- Supuesto 7: No existe multicolinealidad

Con propósito de examinar si las variables independientes presentan multicolinealidad entre sí se elabora una matriz de correlación. El coeficiente de correlación que se utilizará para este procedimiento dependerá de si se cumplen los supuestos necesarios para aplicar el coeficiente de correlación de Pearson, de lo contrario se optará por una alternativa robusta, el coeficiente de correlación de Spearman. Para ello se utiliza la prueba de Jarque-Bera sobre las cuatro variables independientes.

Modelo	Jarque-Bera	
	Valor p	Conclusión
Δ Prom. Espect. TW	$2,2 \times 10^{-16}$	Se rechaza H0
Δ Canales Prom. TW	$2,2 \times 10^{-16}$	Se rechaza H0
Δ CBOE VIX	$2,2 \times 10^{-16}$	Se rechaza H0
Δ DJIA	$2,2 \times 10^{-16}$	Se rechaza H0

Tabla 23: Resultados de la prueba de Jarque-Bera para cada variable independiente de los modelos sobre Twitch

Fuente: Elaboración propia.

Es posible apreciar que los valores p son prácticamente cero, por lo que en los cuatro casos se rechaza la hipótesis nula. A partir de tal resultado se concluye que ninguna de estas variables posee una distribución normal, provocando que el coeficiente de correlación de Spearman sea el más apropiado para crear la matriz de correlación.

Variable	Δ Prom. Espect. TW	Δ Canales Prom. TW	Δ CBOE VIX	Δ DJIA
Δ Prom. Espect. TW	1,000	0,229**	0,113	-0,048
Δ Canales Prom. TW	0,229**	1,000	-0,065	0,143
Δ CBOE VIX	0,113	-0,065	1,000	-0,679***
Δ DJIA	-0,048	0,143.	-0,679***	1,000
Significancia estadística: *** = 0,1% ; ** = 1% ; * = 5% ; . = 10%				

Tabla 24: Matriz de correlación para las variables independientes de los modelos sobre Twitch

Fuente: Elaboración propia.

La matriz de correlación revela que solo hay dos pares de variables que poseen valores considerables de correlación, al mismo tiempo que poseen suficiente significancia estadística. El primer caso son las variables de espectadores promedio de Twitch y los canales promedio de la misma plataforma, que tienen un coeficiente igual a 0,229, pudiendo ser etiquetada como una asociación débil pero considerable según la escala de (Senthilnathan, 2019). Esto no representa un problema, ya que la colinealidad se considera perjudicial cuando la correlación posee magnitud demasiado alta, es decir, por sobre la cifra de 0,80.

Además, es de esperarse que ambas variables tengan cierta correlación entre sí, debido a que provienen de la misma fuente y están ligadas a un mismo efecto: el consumo de contenido audiovisual en Twitch. Por esa razón se considera que tener estos resultados es aceptable, sin representar un inconveniente relativo a una colinealidad dentro de los modelos regresivos.

El segundo par de variables que presentan correlación medianamente alta son las variables de control: el índice de volatilidad CBOE VIX (Chicago Board Options Exchange Market Volatility Index) y el promedio industrial de Dow Jones (DJIA). Entre ambas poseen un coeficiente de correlación de Spearman igual a -0,679, con alta significancia estadística. Tal cifra puede ser categorizada como considerablemente alta según Senthilnathan (2019). Aunque, tal como ocurre con el otro par de variables explicativas, no es un valor que sea

suficiente para afectar de forma severa los resultados de la regresión. En vista de que el valor del coeficiente es menor a 0,80.

En conclusión, ambos casos donde existe correlación significativa no presentan un grado de asociación que impida el cumplimiento de este supuesto. En otras palabras, no se está en presencia de colinealidad.

A partir de la revisión de estos ocho supuestos se determina que los modelos a estimar son ahora los siguientes:

$$\begin{aligned} \text{Retornos_BTC} &= \beta_0 + \beta_1 \Delta(\ln(TW_AvgVwrs)) + \beta_2 \Delta(TW_Channls) + \beta_3 \Delta Cntrl_1 + \beta_4 \Delta Cntrl_2 + \varepsilon \\ \Delta Vol_Int_BTC &= \beta_0 + \beta_1 \Delta(\ln(TW_AvgVwrs)) + \beta_2 \Delta(TW_Channls) + \beta_3 \Delta Cntrl_1 + \beta_4 \Delta Cntrl_2 + \varepsilon \end{aligned}$$

Donde el símbolo “ Δ ” denota que ahora se trabaja con las primeras diferencias de las variables, al igual que “ $\ln$ ” señala que se ocupa la función logaritmo natural en la variable de espectadores promedio. Ambos términos corresponden a los dos principales cambios realizados en comparación al estado del modelo previo a la revisión de supuestos.

Una observación adicional frente a la comprobación de estos supuestos es que el modelo para estimar los retornos posee autocorrelación, implicando que sus estimadores no son eficientes.

6.3.2. Resultados de la regresión

Luego de haber comprobado cuáles supuestos se procede a analizar las estimaciones obtenidas:

```
Call:
lm(formula = y1 ~ x1 + x2 + xctrl1 + xctrl2)

Residuals:
    Min       1Q   Median       3Q      Max
-11408.7  -942.6   -157.3   1128.0   8766.7

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    23.0352    220.9831   0.104   0.917
x1             612.5403    696.0602   0.880   0.380
x2              4.9111     45.0804   0.109   0.913
xctrl1         7.5997     67.4934   0.113   0.910
xctrl2         0.8357      0.3971   2.104   0.037 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2752 on 152 degrees of freedom
Multiple R-squared:  0.05898,    Adjusted R-squared:  0.03422
F-statistic: 2.382 on 4 and 152 DF,  p-value: 0.054
```

Tabla 25: Resumen del modelo lineal con variables de Twitch para estimar retornos

Fuente: Elaboración propia.

La imagen muestra el modelo lineal referente al precio de Bitcoin (y1), donde “x1” es la variable sobre espectadores promedio de Twitch, “x2” simboliza el número de canales transmitiendo contenido sobre criptomonedas en dicho sitio web y “xctrl1” junto a “xctrl2” son las variables de control.

Al observar el valor p del estadístico F se revela que este es igual a 0,054, traduciéndose en que no es posible rechazar la hipótesis nula al 5% de significancia. Esto implica que el modelo no es válido, por lo tanto, los datos de Twitch no son capaces de

sustentar que el modelo funcione correctamente bajo el diseño asignado para esta investigación.

De la misma forma se analizan los coeficientes de regresión obtenidos al utilizar a la primera diferencia del Volumen de Intercambio de Bitcoin como variable dependiente. En este caso el software R provee el siguiente resumen:

```
Call:
lm(formula = y2 ~ x1 + x2 + xctrl1 + xctrl2)

Residuals:
    Min       1Q   Median       3Q      Max
-2.941e+10 -3.671e+09 -4.509e+08  4.352e+09  3.882e+10

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -132376033   696388356  -0.190   0.8495
x1           4502414585  2193507871   2.053   0.0418 *
x2           153862054   142062622   1.083   0.2805
xctrl1       89798167    212693124   0.422   0.6735
xctrl2       213701     1251502    0.171   0.8646
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.671e+09 on 152 degrees of freedom
Multiple R-squared:  0.04071,    Adjusted R-squared:  0.01546
F-statistic: 1.613 on 4 and 152 DF,  p-value: 0.1739
```

Tabla 26: Resumen del modelo lineal con variables de Twitch para estimar diferencias en el volumen de intercambio

Fuente: Elaboración propia.

Donde “y2” representa el volumen de intercambio de Bitcoin, “x1” es la variable sobre espectadores promedio de Twitch, “x2” simboliza a los canales transmitiendo contenido sobre criptomonedas en dicho sitio web y “xctrl1” junto a “xctrl2” son las variables de control.

De manera similar a la regresión anterior, el estadístico F posee un valor de 0,1739, por lo que el modelo no logra rechazar la hipótesis nula. Esto conlleva a que no sea un modelo válido para su interpretación.

7. CONCLUSIONES Y RECOMENDACIONES

La actual literatura sobre Bitcoin comprende una gran variedad de estudios que intentan reconocer los factores que influyen en su comportamiento. Uno de los enfoques que ha sido analizado frecuentemente es la relación que existe entre redes sociales y esta criptomoneda, donde se ha descubierto que el comportamiento de las personas en la web puede ser de utilidad para predecir variables como el precio y su variabilidad. No obstante, las investigaciones presentes suelen enfocarse principalmente en la información contenida como texto acerca de Bitcoin y los grupos de interés a su alrededor. Dejando así de lado la actividad en línea que posee formato de video.

Es por ello que esta memoria decidió investigar la relación existente entre el material audiovisual sobre criptomonedas y el mercado de Bitcoin, basándose en antecedentes que respaldan la existencia de relaciones lineales entre estas variables dentro otros mercados financieros.

De esa forma, se examinó a través de un análisis de correlación si los niveles de búsquedas vinculados a películas, series o documentales sobre criptomonedas se relacionan con los retornos de Bitcoin y su volumen de intercambio. Al igual que se estiman regresiones lineales múltiples para estudiar la capacidad predictiva de estadísticas provenientes de plataformas como YouTube y Twitch.

En Primer lugar, los niveles de búsqueda sugieren que la popularidad con que las personas buscan los nombres de producciones sobre criptomonedas en Google y YouTube poseen en general una baja correlación monótona con los datos históricos de Bitcoin. Este vínculo suele mantenerse dentro de un rango de similar magnitud a lo largo del tiempo, sin lograrse identificar un periodo de desfase para el que las búsquedas de Google o YouTube se asocien con mayor intensidad a los cambios de las variables de Bitcoin en semanas diferentes.

Provocando así que la asociación entre los niveles de búsqueda y las variables de Bitcoin sea baja en la mayoría de los horizontes que es posible analizar.

Además, las relaciones encontradas de mayor magnitud resultan ser de carácter positivo, implicando que los periodos de tiempo donde YouTube y Google reciben más personas buscando sobre películas suele coincidir débilmente con aquellos donde los retornos de Bitcoin y el volumen de intercambio sufren cambios orientados a un alza.

Como segundo punto, al realizar la regresión lineal de estadísticas provenientes de YouTube se desprende que el número de nuevos suscriptores en los canales sobre criptomonedas tiene un efecto positivo con los retornos de Bitcoin y su volumen de intercambio. Tal comportamiento sugiere que la decisión de las personas por querer suscribirse a un canal de YouTube sobre criptomonedas podría estar ligada a el comportamiento de los inversionistas. Aunque también es posible que, en un sentido contrario, los acontecimientos del mercado sean aquellos que provoquen las suscripciones en la plataforma de videos. Por este motivo este hallazgo puede convertirse en una motivación para futuras investigaciones, las cuales podrían explorar con pruebas de causalidad cuál la dirección predictiva más probable en la que estas variables se vinculan.

Por otro lado, los desempeños de los modelos creados en base a la información de YouTube presentan una capacidad considerable para explicar las variaciones en los retornos de Bitcoin y su volumen de intercambio. Si bien esta capacidad implica que otras variables que no fueron incluidas en el modelo juegan un rol mayoritario en la estimación, los resultados obtenidos con YouTube podrían conllevar a que su capacidad predictiva pueda ser de aporte en modelos que consideren un mayor número de variables explicativas.

Después, como tercer punto, los datos provenientes de la plataforma Twitch no demostraron ser capaces de solventar un modelo regresivo para ambas variables de Bitcoin.

Generando que los efectos de las estadísticas de audiencia en sus transmisiones a tiempo real no sean lo suficientemente significativos. Una posible causa de ello puede que las variables recopiladas en TwitchTracker no sean lo suficientemente representativas de la actividad de las personas como para lograr detectar una posible relación con Bitcoin. O bien, efectivamente no existe vínculo alguno entre esta plataforma y el mercado de la criptomoneda, razón por la cual los resultados no son significativos.

De esa manera, al comparar los resultados obtenidos sobre capacidad predictiva entre YouTube y Twitch se determina que las variables de la primera plataforma mencionada son de mayor utilidad a la hora de estimar los retornos o los volúmenes de intercambio de Bitcoin.

A partir de las resoluciones anteriores se puede decir que esta investigación logra brindar un primer antecedente de la relación entre material audiovisual y Bitcoin, bajo el supuesto de que esta posee características lineales o monótonas. Al igual que se presentan nuevas formas para obtener datos acerca de material audiovisual en redes sociales y la web, gracias a la descripción y comparación de las distintas fuentes de información disponibles.

Por consiguiente, el presente estudio comprueba que el tipo de relaciones lineales observadas entre otros mercados financieros y medios audiovisuales también ocurre en el entorno de las criptomonedas. Aunque solo YouTube cumple con tales propiedades, en específico la cantidad de suscriptores que se unen cada semana a los canales sobre criptomonedas.

Futuras investigaciones podrían extender este tipo de estudios a otras plataformas de contenido en formato de video, tales como Instagram o TikTok (cuyos datos también están disponibles en los sitios web utilizados para esta memoria). Al igual que se pueden incorporar otros tipos de buscadores diferentes a Google y YouTube (Yahoo!, Bing, etc.), dependiendo de la disponibilidad de los datos.

8. LIMITACIONES

La presente investigación incluye múltiples limitaciones, que pueden ser agrupadas en dos grupos: las metodologías seleccionadas y la recopilación de datos

8.1. Metodologías seleccionadas

- El coeficiente de correlación de Spearman, que fue utilizado para la comprobación de supuestos y el análisis de los niveles de búsqueda, solo es capaz de detectar relaciones monótonas. Por lo tanto, cuando dos variables no presentan este tipo de correlación no significa que no hay relación entre ellas, ya que pueden estar asociadas de una forma que no es detectada por este coeficiente. Entonces, las conclusiones elaboradas acerca de la no existencia de colinealidad o la baja asociación entre dos series de tiempo podrían no ser ciertas en casos donde la relación no es detectada por este coeficiente.
- El uso de la metodología de regresión lineal múltiple y los coeficientes de correlación no son sinónimos de causalidad (Vetter & Schober, 2018), es decir, si una variable tiene efecto sobre otra en base a este tipo de cálculos eso no implica que una de ellas sea la causa de la otra. Por lo tanto, al estudiar si la información sobre contenido audiovisual acerca de criptomonedas tiene efectos sobre Bitcoin no se está asumiendo que dichos videos son la causa del movimiento del precio, sino que solo se evalúa su capacidad para estimar sus valores numéricos dentro de la regresión. Otras metodologías, como las pruebas de causalidad de Granger, son capaces de llegar más cerca de tales tipos de conclusiones.

- Para poder obtener series de tiempo estacionarias se aplicó el cálculo de las primeras diferencias. No obstante, tal proceder posee la limitación de perder parte de la relación a largo plazo entre las variables (Shrestha & Bhatta, 2018). De esa forma, esta investigación puede estar omitiendo dicha información de los datos debido a esta transformación.

8.2. Recopilación de datos

- Al obtener datos sobre Twitch se opta por utilizar aquellos que están asignados a una categoría o temática de las transmisiones en vivo. Existe la posibilidad de que haya creadores de contenidos que asignan erróneamente “Crypto” a su transmisión, causando que los datos puedan tener información de audiencias que no estén relacionadas con criptomonedas.
- El modelo lineal no siempre fue eficiente en esta aplicación, debido a la presencia de autocorrelación. Por ende, el uso de metodologías como los modelos autorregresivos podría relevar otros detalles sobre la relación entre el material audiovisual y Bitcoin.
- Como los datos recopilados provienen de fuentes secundarias se asume que sus valores son fijos, es decir, no corresponden a un proceso estocástico desde el punto de vista de esta investigación. Esta limitación se basa en lo expuesto por (Gujarati & Porter, 2008, p.316), en relación al supuesto de variables independientes no estocásticas para una regresión lineal múltiple.
- El intervalo de tiempo considerado para el estudio está limitado por la disponibilidad de información de SocialBlade, en vista de que es la fuente de datos considerada que posee el menor número de observaciones disponibles.

- Las variables de Bitcoin tomadas en cuenta para este estudio pueden estar relacionadas a variables que no fueron incluidas en los modelos de regresión, sin embargo, para esta memoria se utilizaron la mayor cantidad posible según disponibilidad de referencias que las respalden y de su fácil acceso en la web.

9. REFERENCIAS

Aalborg, H. A., Molnár, P., & de Vries, J. E. (2019). What can explain the price, volatility and trading volume of Bitcoin? *Finance Research Letters*, 29, 255–265.

<https://doi.org/10.1016/j.frl.2018.08.010>

About. (s. f.). CoinMarketCap. Recuperado 3 de agosto de 2022, de <https://coinmarketcap.com/about/>

About SullyGnome.com - SullyGnome. (s. f.). SullyGnome. Recuperado 3 de agosto de 2022, de <https://sullygnome.com/about>

About YouTube - YouTube. (s. f.). YouTube. Recuperado 2 de agosto de 2022, de <https://about.youtube/>

Abraham, J., Higdon, D., Nelson, J., & Ibarra, J. (2018). Cryptocurrency Price Prediction Using Tweet Volumes and Sentiment Analysis. *SMU Data Science Review*, 1. <https://scholar.smu.edu/datasciencereview/vol1/iss3/1/>

Adnan Fendi, U., Tahtamouni, A., Jalghoum, Y., & Jamal Mohammad, S. (2019). The study of bubbles in bitcoin behavior. *Banks and Bank Systems*, 14(4), 133–142. [https://doi.org/10.21511/bbs.14\(4\).2019.13](https://doi.org/10.21511/bbs.14(4).2019.13)

Aghaei, S., Nematbakhsh, M. A., & Farsani, H. K. (2012). Evolution of the World Wide Web: from Web 1.0 to Web 4.0. *International Journal of Web & Semantic Technology (IJWesT)*, 3(1). <https://doi.org/10.5121/ijwest.2012.3101>

Aharon, D. Y., & Qadan, M. (2019). Bitcoin and the day-of-the-week effect. *Finance Research Letters*, 31. <https://doi.org/10.1016/j.frl.2018.12.004>

Ahmed, M., Jahangir, M., Afzal, H., Majeed, A., & Siddiqi, I. (2015, diciembre). Using Crowd-Source Based Features from Social Media and Conventional Features to

Predict the Movies Popularity. *2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity)*, 273–278.

<https://doi.org/10.1109/smartcity.2015.83>

Akdag, M., & Emsen, H. S. (2021). EFFECT OF MONETARY GROWTH IN USA ON SELECTED ASSETS: S&P500, GOLD & BITCOIN ESTIMATES WITH FACEBOOK PROPHET. En *Administrative and Economic Science Research Theory* (pp. 219–234). Livre de Lyon.

https://www.researchgate.net/publication/357340269_EFFECT_OF_MONETARY_GROWTH_IN_USA_ON_SELECTED_ASSETS_SP500_GOLD_BITCOIN_ESTIMATES_WITH_FACEBOOK_PROPHET

Alexander, C., & Dakos, M. (2019). A critical investigation of cryptocurrency data and analysis. *Quantitative Finance*, 20(2), 173–188.

<https://doi.org/10.1080/14697688.2019.1641347>

ALL ABOUT SOCIAL BLADE. (s. f.). Social Blade. Recuperado 3 de agosto de 2022, de <https://socialblade.com/info>

Alshamsi, A., & Andras, P. P. (2019). User perception of Bitcoin usability and security across novice users. *International Journal of Human-Computer Studies*, 126, 94–110. <https://doi.org/10.1016/j.ijhcs.2019.02.004>

Analytics Made Easy. (s. f.). Social Blade. Recuperado 3 de agosto de 2022, de <https://socialblade.com/>

Ante, L. (2021). How Elon Musk’s Twitter Activity Moves Cryptocurrency Markets. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3778844>

Apala, K. R., Jose, M., Motnam, S., Chan, C. C., Liska, K. J., & de Gregorio, F. (2013, agosto). Prediction of movies box office performance using social media.

Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 1209–1214. <https://doi.org/10.1145/2492517.2500232>

Apopo, N., & Phiri, A. (2021). On the (in)efficiency of cryptocurrencies: have they taken daily or weekly random walks? *Heliyon*, 7(4).
<https://doi.org/10.1016/j.heliyon.2021.e06685>

Asaad, C. T. (2016). *Educational Entertainment: Integrating Business Ethics with Out-of-the-Classroom Financial Documentaries*. JSTOR. Recuperado 2 de agosto de 2022, de <http://www.jstor.org/stable/90000836>

Aslanidis, N., Bariviera, A. F., & López, S. G. (2022). The link between cryptocurrencies and Google Trends attention. *Finance Research Letters*, 47.
<https://doi.org/10.1016/j.frl.2021.102654>

Aysan, A. F., Demir, E., Gozgor, G., & Lau, C. K. M. (2019). Effects of the geopolitical risks on Bitcoin returns and volatility. *Research in International Business and Finance*, 47, 511–518. <https://doi.org/10.1016/j.ribaf.2018.09.011>

Baeza-Yates, R., de Vries, A. P., Zaragoza, H., Cambazoglu, B. B., Murdock, V., Lempel, R., Silvestri, F., & de Vries, A. P. (2012). *Advances in Information Retrieval* (Vol. 7224). Springer Publishing. https://doi.org/10.1007/978-3-642-28997-2_51

Baird, G. L., & Bieber, S. L. (2016). The Goldilocks Dilemma: Impacts of Multicollinearity -- A Comparison of Simple Linear Regression, Multiple Regression, and Ordered Variable Regression Models. *Journal of Modern Applied Statistical Methods*, 15(1), 332–357. <https://doi.org/10.22237/jmasm/1462076220>

Bamnerjee, A., Dolado, J. J., Galbraith, J. W., & Hendry, D. (1993). *Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data* (1.a ed.). Clarendon Press. <https://doi.org/10.1093/0198288107.001.0001>

Banerjee, Tutor in Economics and Barnett Fellow Anindya, Dolado, Chief Economist Research Department Juan, Galbraith, Associate Professor Department of Economics J W & Hendry, Professor of Economics David. (1993). *Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data* (1.^a ed.). Clarendon Press.

Bartos, J. (2015). Does Bitcoin follow the hypothesis of efficient market? *International Journal of Economic Sciences*, 4(2), 10–23.
<https://doi.org/10.52950/ES.2015.4.2.002>

Baur, D. G., & Dimpfl, T. (2017). Excess Volatility as an Impediment for a Digital Currency. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2949754>

Baur, D. G., & Dimpfl, T. (2021). The volatility of Bitcoin and its role as a medium of exchange and a store of value. *Empirical Economics*, 61(5), 2663–2683.
<https://doi.org/10.1007/s00181-020-01990-5>

Baur, D. G., & Hoang, L. (2021). The Bitcoin gold correlation puzzle. *Journal of Behavioral and Experimental Finance*, 32. <https://doi.org/10.1016/j.jbef.2021.100561>

Baur, D. G., Lee, A. D., & Hong, K. (2015). Bitcoin: Currency or Investment? *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2561183>

Behzat, M. (2016). External Borrowing and Inflation in Turkey Between 2003 and 2015: A Simple Linear Regression Analysis. *International Journal of Economics and Financial Issues*, 6(1), 45–54. <https://dergipark.org.tr/en/download/article-file/364683>

Bhardwaj, S. (2022). Japanese film studio to produce a movie series based on Crypto. Forbes India. Recuperado 29 de agosto de 2022, de
<https://www.forbesindia.com/article/crypto-made-easy/japanese-film-studio-to-produce-a-movie-series-based-on-crypto/77503/1>

Bhattacharjee, B., Sridhar, A., & Dutta, A. (2017). Identifying the causal relationship between social media content of a Bollywood movie and its box-office success - a text mining approach. *International Journal of Business Information Systems*, 24(3), 344–368. <https://doi.org/10.1504/ijbis.2017.082039>

Bieniek-Tobasco, A., McCormick, S., Rimal, R. N., Harrington, C. B., Shafer, M., & Shaikh, H. (2019). Communicating climate change through documentary film: imagery, emotion, and efficacy. *Climatic Change*, 154(1–2), 1–18. <https://doi.org/10.1007/s10584-019-02408-7>

Bitcoin - The Currency of the Internet. (s. f.). Reddit. Recuperado 3 de septiembre de 2022, de <https://www.reddit.com/r/Bitcoin/>

Borum, C., Das, A., Aufderheide, P., Clark, J., & Schneider, E. (2016, 27 abril). *Assessing the Social Impact of Issues-Focused Documentaries: Research Methods & Future Considerations*. Center for Media and Social Impact. Recuperado 2 de agosto de 2022, de <https://cmsimpact.org/resource/assessing-the-social-impact-of-issues-focused-documentaries-research-methods-future-considerations/>

Bouoiyour, J., & Selmi, R. (2016). Bitcoin: A beginning of a new phase? *Economics Bulletin*, 36(3), 1430–1440. https://www.researchgate.net/publication/305810738_Bitcoin_A_beginning_of_a_new_phase

Bouoiyour, J., & Selmi, R. (2017). The Bitcoin price formation: Beyond the fundamental sources. *Computational Finance*. <https://arxiv.org/abs/1707.01284>

BOUOIYOUR, J., SELMI, R., & TIWARI, A. K. (2015). IS BITCOIN BUSINESS INCOME OR SPECULATIVE FOOLERY? NEW IDEAS THROUGH AN IMPROVED

FREQUENCY DOMAIN ANALYSIS. *Annals of Financial Economics*, 10(01).

<https://doi.org/10.1142/s2010495215500025>

Bouri, E., Gupta, R., Tiwari, A. K., & Roubaud, D. (2017). Does Bitcoin hedge global uncertainty? Evidence from wavelet-based quantile-in-quantile regressions. *Finance Research Letters*, 23, 87–95. <https://doi.org/10.1016/j.frl.2017.02.009>

Bouri, E., Roubaud, D., & Shahzad, S. J. H. (2020). Do Bitcoin and other cryptocurrencies jump together? *The Quarterly Review of Economics and Finance*, 76, 396–409. <https://doi.org/10.1016/j.qref.2019.09.003>

Bradshaw, A. S., Treise, D., Shelton, S. S., Cretul, M., Raisa, A., Bajalia, A., & Peek, D. (2020). Propagandizing anti-vaccination: Analysis of Vaccines Revealed documentary series. *Vaccine*, 38(8), 2058–2069.

<https://doi.org/10.1016/j.vaccine.2019.12.027>

Brave New Coin. (s. f.-a). BraveNewCoin. Recuperado 3 de agosto de 2022, de <https://bravenewcoin.com/>

Brave New Coin. (s. f.-b). BraveNewCoin. Recuperado 3 de agosto de 2022, de <https://bravenewcoin.com/about>

Browne, R. (2022, 22 junio). *Bitcoin could plunge even further to a low of \$13,000, one strategist warns*. CNBC. Recuperado 29 de julio de 2022, de <https://www.cnbc.com/2022/06/22/bitcoin-btc-price-may-tank-as-low-as-13000-strategist-warns.html>

Budzinski, O., & Gaenssle, S. (2018). The economics of social media (super-)stars: an empirical investigation of stardom and success on YouTube. *Journal of Media Economics*, 31(3–4), 75–95. <https://doi.org/10.1080/08997764.2020.1849228>

Budzinski, O., Gaenssle, S., & Lindstädt-Dreusicke, N. (2021). The battle of YouTube, TV and Netflix: an empirical analysis of competition in audiovisual media markets. *SN Business & Economics*, 1(9). <https://doi.org/10.1007/s43546-021-00122-0>

Burnie, A., & Yilmaz, E. (2019). Social media and bitcoin metrics: which words matter. *Royal Society Open Science*, 6(10). <https://doi.org/10.1098/rsos.191068>

Busca. *De forma segura*. (s. f.). Google Safety Center. Recuperado 3 de septiembre de 2022, de <https://safety.google/products/#search>

Carvalho, A. (2018). Understanding What Drives Bitcoin Trading Activities. *Proceedings of the 2018 Annual Meeting of the Decision Sciences Institute*, 1864–1872. <https://doi.org/10.48550/arXiv.1901.00720>

CBOE Volatility Index. (s. f.). Yahoo! Finance. Recuperado 1 de noviembre de 2022, de <https://finance.yahoo.com/quote/^VIX?p=^VIX>

Ceci, L. (2022, 4 abril). *YouTube - Statistics & Facts*. Statista. Recuperado 2 de agosto de 2022, de https://www.statista.com/topics/2019/youtube/#topicHeader__wrapper

Cermak, V. (2017). Can Bitcoin Become a Viable Alternative to Fiat Currencies? An Empirical Analysis of Bitcoin's Volatility Based on a GARCH Model. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2961405>

Chae, S. W., & Lee, S. H. (2022). Sharing emotion while spectating video game play: Exploring Twitch users' emotional change after the outbreak of the COVID-19 pandemic. *Computers in Human Behavior*, 131, 107211. <https://doi.org/10.1016/j.chb.2022.107211>

Chan, W. H., Le, M., & Wu, Y. W. (2019). Holding Bitcoin longer: The dynamic hedging abilities of Bitcoin. *The Quarterly Review of Economics and Finance*, 71, 107–113. <https://doi.org/10.1016/j.qref.2018.07.004>

Chen, T., Lau, C. K. M., Cheema, S., & Koo, C. K. (2021). Economic Policy Uncertainty in China and Bitcoin Returns: Evidence From the COVID-19 Period. *Frontiers in Public Health*, 9. <https://doi.org/10.3389/fpubh.2021.651051>

Chico-Frias, A. (2021). Bitcoin, Intrinsic Value, and Herd Effect. *DLSU Business & Economics Review*, 31(1), 55–71. https://www.dlsu.edu.ph/wp-content/uploads/2021/08/DLSUBER.2021.July_.5frias0729.pdf

Cizmeci, B., & Oguducu, S. G. (2018, septiembre). Predicting IMDb Ratings of Pre-release Movies with Factorization Machines Using Social Media. *2018 3rd International Conference on Computer Science and Engineering (UBMK)*, 173–178. <https://doi.org/10.1109/ubmk.2018.8566661>

Clement, J. (2021, 29 julio). *Twitch - statistics & facts*. Statista. Recuperado 2 de agosto de 2022, de <https://www.statista.com/topics/7946/twitch/#dossierKeyfigures>

CoinGecko. (s. f.). *Cryptocurrency Prices, Charts, and Crypto Market Cap*. Recuperado 3 de agosto de 2022, de <https://www.coingecko.com/>

CoinMarketCap. (s. f.). *Cryptocurrency Prices, Charts And Market Capitalizations*. Recuperado 29 de julio de 2022, de <https://coinmarketcap.com/>

Colianni, S., Rosales, S., & Signorotti, M. (s. f.). *Algorithmic Trading of Cryptocurrency Based on Twitter Sentiment Analysis*. stanford.edu. Recuperado 2 de agosto de 2022, de http://cs229.stanford.edu/proj2015/029_report.pdf

Compra, haz trading y holdea más de 600 criptomonedas en Binance. (s. f.). Binance. Recuperado 3 de agosto de 2022, de <https://www.binance.com/es-LA>

Conde, M. A., Forteza-Martínez, A., & Andrade-Martínez, C. M. (2020). Análisis de la capacidad de liderazgo y el carisma de los principales youtubers españoles. *3C TIC*:

Cuadernos de desarrollo aplicados a las TIC, 9(3), 17–41.

<https://doi.org/10.17993/3ctic.2020.93.17-41>

cookie. (s. f.). LII / Legal Information Institute. Recuperado 3 de septiembre de 2022, de <https://www.law.cornell.edu/wex/cookie>

Corbet, S., Lucey, B., Peat, M., & Vigne, S. (2018). Bitcoin Futures—What use are they? *Economics Letters*, 172, 23–27. <https://doi.org/10.1016/j.econlet.2018.07.031>

Crypto.com | The Best Place to Buy, Sell, and Pay with Cryptocurrency. (s. f.). Crypto. Recuperado 3 de agosto de 2022, de <https://crypto.com/>

D'Aquila, J. M., Wang, D., & Mattia, A. (2019). Are instructor generated YouTube videos effective in accounting classes? A study of student performance, engagement, motivation, and perception. *Journal of Accounting Education*, 47, 63–74.

<https://doi.org/10.1016/j.jaccedu.2019.02.002>

Das, P. (2019). *Econometrics in Theory and Practice: Analysis of Cross Section, Time Series and Panel Data with Stata 15.1* (English Edition) (1st ed. 2019). Springer.

Datos históricos de Bitcoin USD. (s. f.). CoinGecko. Recuperado 4 de agosto de 2022, de https://www.coingecko.com/es/monedas/bitcoin/historical_data#panel

Dehouche, N. (2021). Scale matters: The daily, weekly and monthly volatility and predictability of Bitcoin, Gold, and the S&P 500. *Statistical Finance*.

<https://doi.org/10.48550/arXiv.2103.00395>

Demir, E., Gozgor, G., Lau, C. K. M., & Vigne, S. A. (2018). Does economic policy uncertainty predict the Bitcoin returns? An empirical investigation. *Finance Research Letters*, 26, 145–149. <https://doi.org/10.1016/j.frl.2018.01.005>

Diwanji, V., Reed, A., Ferchaud, A., Seibert, J., Weinbrecht, V., & Sellers, N. (2020). Don't just watch, join in: Exploring information behavior and copresence on Twitch. *Computers in Human Behavior*, 105. <https://doi.org/10.1016/j.chb.2019.106221>

Documental. (s. f.). Diccionario de la lengua española. Recuperado 29 de agosto de 2022, de <https://dle.rae.es/documental>

Dokic, K., Radman Funarić, M., & Potnik Galić, K. (2015). The Relationship between the Cryptocurrency Value (Bitcoin) and Interest for It in the Region. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3281915>

Dominguez, A. M. D. V., & Sierra, J. (2020). *Aproximación periódica y educativa al fen*. <https://dialnet.unirioja.es/servlet/articulo?codigo=7277176>

Dow Jones Industrial Average. (s. f.). Yahoo! Finance. Recuperado 1 de noviembre de 2022, de <https://finance.yahoo.com/quote/^DJI/history?p=^DJI>

Dyhrberg, A. H., Foley, S., & Svec, J. (2018). How investible is Bitcoin? Analyzing the liquidity and transaction costs of Bitcoin markets. *Economics Letters*, 171, 140–143. <https://doi.org/10.1016/j.econlet.2018.07.032>

Education, I. C. (2022, 28 septiembre). *Application Programming Interface (API)*. <https://www.ibm.com/cloud/learn/api>

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4), 987. <https://doi.org/10.2307/1912773>

Ever wonder how YouTube works? (s. f.). YouTube. Recuperado 3 de septiembre de 2022, de <https://www.youtube.com/howyoutubeworks/>

Facebook - Meld je aan of registreer je. (s. f.). Facebook. Recuperado 1 de noviembre de 2022, de <https://www.facebook.com/unsupportedbrowser>

Fang, L., Bouri, E., Gupta, R., & Roubaud, D. (2019). Does global economic uncertainty matter for the volatility and hedging effectiveness of Bitcoin? *International Review of Financial Analysis*, 61, 29–36. <https://doi.org/10.1016/j.irfa.2018.12.010>

FAQ. (s. f.-a). CoinMarketCap. Recuperado 3 de agosto de 2022, de <https://coinmarketcap.com/faq/>

FAQ. (s. f.-b). CoinGecko. Recuperado 4 de agosto de 2022, de <https://www.coingecko.com/es/faq>

Feng, W., Wang, Y., & Zhang, Z. (2018). Informed trading in the Bitcoin market. *Finance Research Letters*, 26, 63–70. <https://doi.org/10.1016/j.frl.2017.11.009>

Fosso Wamba, S., Kala Kamdjoug, J. R., Epie Bawack, R., & Keogh, J. G. (2019). Bitcoin, Blockchain and Fintech: a systematic review and case studies in the supply chain. *Production Planning & Control*, 31(2–3), 115–142. <https://doi.org/10.1080/09537287.2019.1631460>

Frater, P. (2022, 28 mayo). Crypto Film ‘The Infinite Machine’ Adds Metaverse Component in Pact With Decentraland, Lumiere. Variety. Recuperado 29 de agosto de 2022, de <https://variety.com/2022/digital/asia/the-infinite-machine-film-metaverse-decentraland-1235280769/>

FREQUENTLY ASKED QUESTIONS (FAQ). (s. f.). Social Blade. Recuperado 3 de agosto de 2022, de <https://socialblade.com/youtube/help>

Frommel, J., Gugenheimer, J., & Rogers, K. (2019). *Opportunities and Challenges of Using Game Video Stream Data for Games Research*. https://www.uni-ulm.de/fileadmin/website_uni_ulm/iui.inst.100/institut/verz-ma-

ehedem/frommel/Publications/CHI19_StreamingWorkshop_PositionPaper_Frommel_Opportunities_cameraReady.pdf

Gandal, N., Hamrick, J., Moore, T., & Oberman, T. (2018). Price manipulation in the Bitcoin ecosystem. *Journal of Monetary Economics*, 95, 86–96.
<https://doi.org/10.1016/j.jmoneco.2017.12.004>

Gandal, N., Hamrick, J. T., Moore, T., & Vasek, M. (2021). The rise and fall of cryptocurrency coins and tokens. *Decisions in Economics and Finance*, 44(2), 981–1014.
<https://doi.org/10.1007/s10203-021-00329-8>

Gao, X., Huang, W., & Wang, H. (2021). Financial Twitter Sentiment on Bitcoin Return and High-Frequency Volatility. *Virtual Economics*, 4(1), 7–18.
[https://doi.org/10.34021/ve.2021.04.01\(1\)](https://doi.org/10.34021/ve.2021.04.01(1))

Giudici, P., & Abu-Hashish, I. (2019). What determines bitcoin exchange prices? A network VAR approach. *Finance Research Letters*, 28, 309–318.
<https://doi.org/10.1016/j.frl.2018.05.013>

Gkillas, K., Tantoula, M., & Tzarakis, M. (2021). Transaction activity and bitcoin realized volatility. *Operations Research Letters*, 49(5), 715–719.
<https://doi.org/10.1016/j.orl.2021.06.016>

Glenski, M. (2019, 1 julio). *Improved Forecasting of Cryptocurrency Price using Social Signals*. arXiv.Org. Recuperado 3 de septiembre de 2022, de
<https://arxiv.org/abs/1907.00558>

Glenski, M., Saldanha, E., & Volkova, S. (2019). Characterizing Speed and Scale of Cryptocurrency Discussion Spread on Reddit. *The World Wide Web Conference*.
<https://doi.org/10.1145/3308558.3313702>

Gómez-Zaldivar, M., Manjarrez, O. & Ventosa-Santaularia, D. (2009). Regresión espuria en especificaciones dinámicas. *Ensayos Revista de Economía*, 18(1), 1-20.
https://www.researchgate.net/publication/227430758_Regresion_espuria_en_especificaciones_dinamicas

Google. (s. f.). Google Search. Recuperado 3 de septiembre de 2022, de
<http://www.google.com>

Google Traductor. (s. f.). Google Translate. Recuperado 14 de agosto de 2022, de
<https://translate.google.com/?hl=es>

Guevara, J. (2014). Uso correcto de la correlación cruzada en Climatología: el caso de la presión atmosférica entre Taití y Darwin. *Terra*, 30(47), 79-102.
http://ve.scielo.org/scielo.php?pid=S1012-70892014000100005&script=sci_abstract

Guizani, S., & Nafti, I. K. (2019). The Determinants of Bitcoin Price Volatility: An Investigation With ARDL Model. *Procedia Computer Science*, 164, 233–238.
<https://doi.org/10.1016/j.procs.2019.12.177>

Gujarati, D. N. & Porter, D. C. (2008). Basic Econometrics (5th Revised ed.). McGraw-Hill Education.

Guo, H., Zhang, D., Liu, S., Wang, L., & Ding, Y. (2021). Bitcoin price forecasting: A perspective of underlying blockchain transactions. *Decision Support Systems*, 151.
<https://doi.org/10.1016/j.dss.2021.113650>

Gutiérrez Lozano, J. F., & Cuartero, A. (2020). El auge de Twitch: nuevas ofertas audiovisuales y cambios del consumo televisivo entre la audiencia juvenil. *Ámbitos. Revista Internacional de Comunicación*, 50, 159–175.
<https://doi.org/10.12795/ambitos.2020.i50.11>

Han, X., Yuan, Y., & Wang, F. Y. (2019). A Blockchain-based Framework for Central Bank Digital Currency. *2019 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI)*. <https://doi.org/10.1109/soli48380.2019.8955032>

Harpstead, E., Rios, J. S., Seering, J., & Hammer, J. (2019). Toward a Twitch Research Toolkit. *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, 111–119. <https://doi.org/10.1145/3311350.3347149>

Hauke, J. & Kossowski, T. (2011). Comparison of Values of Pearson's and Spearman's Correlation Coefficients on the Same Sets of Data. *QUAGEO*, 30(2), 87-93. <https://doi.org/10.2478/v10117-011-0021-1>

Hayes, A. (2015). A Cost of Production Model for Bitcoin. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2580904>

Hayes, A. S. (2017). Cryptocurrency value formation: An empirical study leading to a cost of production model for valuing bitcoin. *Telematics and Informatics*, 34(7), 1308–1321. <https://doi.org/10.1016/j.tele.2016.05.005>

Hazlett, P. K., & Luther, W. J. (2020). Is bitcoin money? And what that means. *The Quarterly Review of Economics and Finance*, 77, 144–149. <https://doi.org/10.1016/j.qref.2019.10.003>

Highfield, T., & Leaver, T. (2016). Instagrammatics and digital methods: studying visual social media, from selfies and GIFs to memes and emoji. *Communication Research and Practice*, 2(1), 47–62. <https://doi.org/10.1080/22041451.2016.1155332>

Hileman, G., & Rauchs, M. (2017). 2017 Global Cryptocurrency Benchmarking Study. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2965436>

Hossain, M. J., & Ismail, M. T. (2021). Is There Any Influence of Other Cryptocurrencies on Bitcoin? *Asian Academy of Management Journal of Accounting and Finance*, 17(1), 125–152. <https://doi.org/10.21315/aamjaf2021.17.1.5>

How does voting work on Reddit? (2022, 30 junio). Reddit. Recuperado 3 de septiembre de 2022, de <https://reddit.zendesk.com/hc/en-us/articles/7419626610708-How-does-voting-work-on-Reddit->

Hsu, P. Y., Shen, Y. H. & Xie, X. A. (2014). Predicting Movies User Ratings with Imdb Attributes. *Rough Sets and Knowledge Technology*, 444-453. https://doi.org/10.1007/978-3-319-11740-9_41

Hu, J. (2019). Risk of Bitcoin Market: Volatility, Jumps, and Forecasts. *Risk Management*. <https://doi.org/10.48550/arXiv.1912.05228>

Hu, Y. (2019, 27 diciembre). *Characterizing and Detecting Money Laundering Activities on the. . .* arXiv.Org. Recuperado 3 de septiembre de 2022, de <https://arxiv.org/abs/1912.12060>

Huynh, T. L. D. (2021). Does Bitcoin React to Trump’s Tweets? *Journal of Behavioral and Experimental Finance*, 31. <https://doi.org/10.1016/j.jbef.2021.100546>

IMDb | Help. (s. f.-a). IMDb. Recuperado 3 de agosto de 2022, de https://help.imdb.com/article/imdb/general-information/what-is-imdb/G836CY29Z4SGNMK5?ref_=helpsect_cons_1_1#

IMDb | Help. (s. f.-b). IMDb. Recuperado 3 de agosto de 2022, de <https://help.imdb.com/article/imdb/general-information/how-do-i-link-to-a-specific-page-on-your-site/GTRTF6K8UA9JP9QQ#>

IMDb Datasets. (s. f.). IMDb. Recuperado 3 de agosto de 2022, de <https://www.imdb.com/interfaces/>

IMDb Developer. (s. f.). IMDb. Recuperado 3 de agosto de 2022, de <https://developer.imdb.com/>

IMDb: Ratings, Reviews, and Where to Watch the Best Movies & TV Shows. (s. f.). IMDb. Recuperado 3 de agosto de 2022, de <https://www.imdb.com/>

Información de BitBoy Crypto. (s. f.). YouTube. Recuperado 29 de agosto de 2022, de https://www.youtube.com/channel/UCjemQfjaXAzA-95RKoy9n_g/about

Información sobre. (s. f.). CoinGecko. Recuperado 3 de agosto de 2022, de <https://www.coingecko.com/es/informacion>

Brooks, C. (2014). *Introductory Econometrics for Finance* (3rd Revised ed.). Cambridge University Press. <https://www.cambridge.org/core/books/introductory-econometrics-for-finance/0E0DEC2BA49EFA566BE377806FBC37FE>

Jahjah, F. H., & Rajab, M. (2020). Impact of Twitter Sentiment Related to Bitcoin on Stock Price Returns. *Journal of Engineering*, 26(6), 60–71. <https://doi.org/10.31026/j.eng.2020.06.05>

Jana, R. K., Ghosh, I., & Wallin, M. W. (2022). Taming energy and electronic waste generation in bitcoin mining: Insights from Facebook prophet and deep neural network. *Technological Forecasting and Social Change*, 178. <https://doi.org/10.1016/j.techfore.2022.121584>

Jareño, F., González, M. D. L. O., Tolentino, M., & Sierra, K. (2020). Bitcoin and gold price returns: A quantile regression and NARDL analysis. *Resources Policy*, 67. <https://doi.org/10.1016/j.resourpol.2020.101666>

Jareño, F., Tolentino, M., González, M. D. L. O., & Oliver, A. (2019). Impact of changes in the level, slope and curvature of interest rates on U.S. sector returns: an

asymmetric nonlinear cointegration approach. *Economic Research-Ekonomska Istraživanja*, 32(1), 1275–1297. <https://doi.org/10.1080/1331677x.2019.1632726>

Jerdack, N., Dauletbek, A., Divine, M., Hult, M., & Cevalho, A. (2018, diciembre). *Understanding What Drives Bitcoin Trading Activities*. 1864–1872. pet

Johansson, E. (s. f.). *Creating daily search volume data from weekly and daily data*. Erik Johansson Blogspot. Recuperado 4 de agosto de 2022, de <http://erikjohansson.blogspot.com/2014/12/creating-daily-search-volume-data-from.html>

Johnson, M. R., & Woodcock, J. (2019). “And Today’s Top Donator is”: How Live Streamers on Twitch.tv Monetize and Gamify Their Broadcasts. *Social Media + Society*, 5(4). <https://doi.org/10.1177/2056305119881694>

Jumde, A., & Cho, B. Y. (2020). Can cryptocurrencies overtake the fiat money. *International Journal of Business Performance Management*, 21(1/2), 6–20. <https://doi.org/10.1504/ijbpm.2020.106107>

Jungar, E. E. E. (2016). Streaming Video Games: Copyright Infringement or Protected Speech? *Press Start*, 3(2). <https://press-start.gla.ac.uk/index.php/press-start/article/view/63>

Kalatzis, N., Roussaki, I., Matsoukas, C., Paraskevopoulos, M., Papavassiliou, S., & Tonoli, S. (2018, noviembre). *Social Media and Google Trends in Support of Audience Analytics: Methodology and Architecture*. 50–56. <https://biblio.ugent.be/publication/8603224/file/8603228.pdf#page=50>

Kaminski, J. (2014). Nowcasting the Bitcoin Market with Twitter Signals. *Social and Information Networks*. <https://doi.org/10.48550/arXiv.1406.7577>

Karalevicius, V., Degrande, N., & de Weerd, J. (2018). Using sentiment analysis to predict interday Bitcoin price movements. *The Journal of Risk Finance*, 19(1), 56–75.

<https://doi.org/10.1108/jrf-06-2017-0092>

Katsiampa, P. (2017). Volatility estimation for Bitcoin: A comparison of GARCH models. *Economics Letters*, 158, 3–6. <https://doi.org/10.1016/j.econlet.2017.06.023>

Khan, M. L. (2017). Social media engagement: What motivates user participation and consumption on YouTube? *Computers in Human Behavior*, 66, 236–247.

<https://doi.org/10.1016/j.chb.2016.09.024>

Kim, J., Wohn, D. Y., & Cha, M. (2022). Understanding and identifying the use of emoticons in toxic chat on Twitch. *Online Social Networks and Media*, 27.

<https://doi.org/10.1016/j.osnem.2021.100180>

Kim, W., Lee, J., & Kang, K. (2020). The effects of the introduction of Bitcoin futures on the volatility of Bitcoin returns. *Finance Research Letters*, 33, 101204.

<https://doi.org/10.1016/j.frl.2019.06.002>

Klein, T., Pham Thu, H., & Walther, T. (2018). Bitcoin is not the New Gold – A comparison of volatility, correlation, and portfolio performance. *International Review of Financial Analysis*, 59, 105–116. <https://doi.org/10.1016/j.irfa.2018.07.010>

Kliber, A., Marszałek, P., Musiałkowska, I., & Świerczyńska, K. (2019). Bitcoin: Safe haven, hedge or diversifier? Perception of bitcoin in the context of a country's economic situation — A stochastic volatility approach. *Physica A: Statistical Mechanics and its Applications*, 524, 246–257. <https://doi.org/10.1016/j.physa.2019.04.145>

Koenker, R., & Bassett, G. (1978). Regression Quantiles. *Econometrica*, 46(1), 33.

<https://doi.org/10.2307/1913643>

Kristoufek, L. (2013). BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. *Scientific Reports*, 3(1).

<https://doi.org/10.1038/srep03415>

Kristoufek, L. (2015). What Are the Main Drivers of the Bitcoin Price? Evidence from Wavelet Coherence Analysis. *PLOS ONE*, 10(4).

<https://doi.org/10.1371/journal.pone.0123923>

Kurihara, Y., & Fukushima, A. (2017). The Market Efficiency of Bitcoin: A Weekly Anomaly Perspective. *Journal of Applied Finance & Banking*, 7.

https://econpapers.repec.org/article/sptapfiba/v_3a7_3ay_3a2017_3ai_3a3_3af_3a7_5f3_5f4.htm

Kwon, J. H. (2020). Tail behavior of Bitcoin, the dollar, gold and the stock market index. *Journal of International Financial Markets, Institutions and Money*, 67.

<https://doi.org/10.1016/j.intfin.2020.101202>

Lacoviello, C., Lacoviello, D., & Lacoviello, M. (2018). Measuring Geopolitical Risk. *International Finance Discussion Paper*, 2018(1222), 1–66.

<https://doi.org/10.17016/ifdp.2018.1222>

Leith, A. P. (2021). Parasocial cues: The ubiquity of parasocial relationships on Twitch. *Communication Monographs*, 88(1), 111–129.

<https://doi.org/10.1080/03637751.2020.1868544>

Leith, A. P., & Gheen, E. (2022). Twitch in the time of quarantine: The role of engagement in needs fulfillment. *Psychology of Popular Media*, 11(3), 275–280.

<https://doi.org/10.1037/ppm0000372>

Lepori, G. M. (2015). Investor mood and demand for stocks: Evidence from popular TV series finales. *Journal of Economic Psychology*, 48, 33–47.

<https://doi.org/10.1016/j.joep.2015.02.003>

Let's find more that brings us together. (s. f.). Facebook. Recuperado 3 de septiembre de 2022, de <https://www.facebook.com/unsupportedbrowser>

Li, W. & He, M. (2020). Comparative Analysis of Bitcoin, Ethereum, and Libra. *2020 IEEE 11th International Conference on Software Engineering and Service Science (ICSESS)*. <https://doi.org/10.1109/icsess49938.2020.9237710>

Lin, Z. Y. (2021). Investor attention and cryptocurrency performance. *Finance Research Letters*, 40. <https://doi.org/10.1016/j.frl.2020.101702>

Lopezosa, C., Fernández, A., Apablaza, A., Codina, L., Dominguez, A. M. D. V., & Sierra, J. (2020). Video SEO y cibermedios estrategias de visibilidad y caracterización de los canales en Youtube de los principales medios españoles. En McGraw-Hill Interamericana De España (Ed.), *Aproximación periodística y educ comunicativa al fenómeno de las redes sociales* (pp. 195–212).

<https://dialnet.unirioja.es/servlet/articulo?codigo=7277176#>

Lucchini, L., Alessandretti, L., Lepri, B., Gallo, A., & Baronchelli, A. (2020). From code to market: Network of developers and correlated returns of cryptocurrencies. *Science Advances*, 6(51). <https://doi.org/10.1126/sciadv.abd2204>

Lux, M., Riegler, M., Halvorsen, P., Dang-Nguyen, D., & Potthast, M. (2019). Challenges for Multimedia Research in E-Sports Using Counter-Strike. En *Savegame* (pp. 197–206). https://doi.org/10.1007/978-3-658-27395-8_13

M., P., Nguyen, T. N., Hamdi, M., & Cengiz, K. (2021). Global cryptocurrency trend prediction using social media. *Information Processing & Management*, 58(6).
<https://doi.org/10.1016/j.ipm.2021.102708>

Mai, F., Shan, Z., Bai, Q., Wang, X. S., & Chiang, R. H. (2018). How Does Social Media Impact Bitcoin Value? A Test of the Silent Majority Hypothesis. *Journal of Management Information Systems*, 35(1), 19–52.
<https://doi.org/10.1080/07421222.2018.1440774>

MarketTradersTV. (s. f.). Twitch. Recuperado 3 de septiembre de 2022, de
<https://www.twitch.tv/marketraderstv>

Martinez, A., & Atouba, Y. (2021). Political Satire TV Shows in the Trump’s Era: Examining Their Impact on Latinx Viewers’ Political Knowledge, Political Engagement, and Trust in Institutions. *Southern Communication Journal*, 86(5), 460–471.
<https://doi.org/10.1080/1041794x.2021.1958913>

Matta, M., Lunesu, I., & Marchesi, M. (2015). *The predictor impact of Web search media on Bitcoin trading volumes*. 620–626.
<https://ieeexplore.ieee.org/abstract/document/7526987>

Methodology. (s. f.). CoinGecko. Recuperado 4 de agosto de 2022, de
<https://www.coingecko.com/es/metodolog%C3%Ada>

Miao, D., Pedrycz, W., Ślęzak, D., Peters, G., Hu, Q., & Wang, R. (2014). *Rough Sets and Knowledge Technology* (Vol. 8818). Springer Publishing.
https://doi.org/10.1007/978-3-319-11740-9_41

Microsoft Bing. (s. f.). Bing. Recuperado 3 de septiembre de 2022, de
<http://www.bing.com>

Miglietti, C., Kubosova, Z., & Skulanova, N. (2019). Bitcoin, Litecoin, and the Euro: an annualized volatility analysis. *Studies in Economics and Finance*, 37(2), 229–242. <https://doi.org/10.1108/sef-02-2019-0050>

Milutinovic, M. (2018). Cryptocurrency. *Ekonomika, Journal for Economic Theory and Practice and Social Issues*, 64(1). <https://doi.org/10.22004/ag.econ.290219>

Mittal, A., Dhiman, V., Singh, A., & Prakash, C. (2019, agosto). Short-Term Bitcoin Price Fluctuation Prediction Using Social Media and Web Search Data. *2019 Twelfth International Conference on Contemporary Computing (IC3)*, 1–6. <https://doi.org/10.1109/ic3.2019.8844899>

Mizerka, J., Stróżyńska-Szajek, A., & Mizerka, P. (2020). The role of Bitcoin on developed and emerging markets – on the basis of a Bitcoin users graph analysis. *Finance Research Letters*, 35. <https://doi.org/10.1016/j.frl.2020.101489>

Most Trusted Cryptocurrency Market Data API | CoinMarketCap. (s. f.). Coinmarketcap.Com. Recuperado 3 de agosto de 2022, de <https://coinmarketcap.com/api/>

Naimy, V. Y., & Hayek, M. R. (2018). Modelling and predicting the Bitcoin volatility using GARCH models. *International Journal of Mathematical Modelling and Numerical Optimisation*, 8(3), 197. <https://doi.org/10.1504/ijmmno.2018.088994>

Nakamoto, S. (2008, 31 octubre). *Bitcoin: A Peer-to-Peer Electronic Cash System*. Bitcoin.org. Recuperado 29 de julio de 2022, de <https://bitcoin.org/bitcoin.pdf>

Nasir, M. A., Huynh, T. L. D., Nguyen, S. P., & Duong, D. (2019). Forecasting cryptocurrency returns and volume using search engines. *Financial Innovation*, 5(1). <https://doi.org/10.1186/s40854-018-0119-8>

New user FAQ - Tweeting. (s. f.). Twitter. Recuperado 3 de septiembre de 2022, de <https://help.twitter.com/en/resources/new-user-faq>

- NextAdvisor. (2022, 28 julio). *Experts Say Bitcoin Could Hit \$100,000 In 2022. Here's What Investors Should Know*. NextAdvisor with TIME. Recuperado 29 de julio de 2022, de <https://time.com/nextadvisor/investing/cryptocurrency/bitcoin-price-predictions/>
- Nguyen, K. Q. (2022). The correlation between the stock market and Bitcoin during COVID-19 and other uncertainty periods. *Finance Research Letters*, 46. <https://doi.org/10.1016/j.frl.2021.102284>
- Obregón-Sierra, N., & González-Fernández, N. (2020). Wikipedia en las facultades de educación españolas. La visión de los estudiantes universitarios. *Alteridad*, 15(2), 218–228. <https://doi.org/10.17163/alt.v15n2.2020.06>
- Odey, J. A., & Okoro, A. T. (2021). ACCESSING THE INVISIBLE WEB: ISSUES AND CONCERNS. *FUW TRENDS IN SCIENCE & TECHNOLOGY JOURNAL*, 6(2), 607–611. <http://www.ftstjournal.com/Digital%20Library/62%20Article%2045.php>
- Oh, S., Kim, J., Ji, H., Park, E., Han, J., Ko, M., & Lee, M. (2020). Cross-cultural comparison of interactive streaming services: Evidence from Twitch. *Telematics and Informatics*, 55, 101434. <https://doi.org/10.1016/j.tele.2020.101434>
- Oh, T. T., & Kim, J. (2022). Influence of Livestreaming Viewership on the Life Cycle of Entertainment Products: Evidence from Twitch and Steam. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4011595>
- Ozyesil, M. (2019). A research on interaction between bitcoin and foreign exchange rates. *Pressacademia*, 6(1), 55–62. <https://doi.org/10.17261/pressacademia.2019.1028>
- Pair. (2021, 21 agosto). CoinMarketCap Alexandria. Recuperado 3 de septiembre de 2022, de <https://coinmarketcap.com/alexandria/glossary/pair>
- Pandey, K. K., & Pradhan, N. (2018, febrero). *Internet Search Engine: Performance Evaluating the Google, Yahoo and Bing Web Search Engine based on their Searching*

Capabilities. DMLS, Bilaspur, India.

https://www.researchgate.net/publication/324482784_Internet_Search_Engine_Performance_Evaluating_the_Google_Yahoo_and_Bing_Web_Search_Engine_based_on_their_Searching_Capabilities

Pant, D. R., Neupane, P., Poudel, A., Pokhrel, A. K., & Lama, B. K. (2018).

Recurrent Neural Network Based Bitcoin Price Prediction by Twitter Sentiment Analysis. *2018 IEEE 3rd International Conference on Computing, Communication and Security (ICCCS)*, 128–132. <https://doi.org/10.1109/cccs.2018.8586824>

Papathanasiou, S., Papamatthaiou, N., & Balios, D. P. (2019). Bitcoin as an alternative digital currency: exploring the publics' perception vs. experts. *International Journal of Financial Engineering and Risk Management*, 3(2), 146–171.

<https://doi.org/10.1504/ijferm.2019.101296>

Partida, P. (2019, 1 marzo). *Bitcoin Bubbles: Epidemic-Diffusion Analyses and Models*. Eidgenössische Technische Hochschule Zurich - Swiss Federal Institute of Technology Zurich - Entrepreneurial Risk. Recuperado 29 de julio de 2022, de https://ethz.ch/content/dam/ethz/special-interest/mtec/chair-of-entrepreneurial-risks-dam/documents/dissertation/master%20thesis/Master-project_Partida_March2019.pdf

Paule-Vianez, J., Prado-Román, C., & Gómez-Martínez, R. (2020). Economic policy uncertainty and Bitcoin. Is Bitcoin a safe-haven asset? *European Journal of Management and Business Economics*, 29(3), 347–363. <https://doi.org/10.1108/ejmbe-07-2019-0116>

Película. (s. f.). Diccionario de la lengua española. Recuperado 29 de agosto de 2022, de <https://dle.rae.es/pel%C3%ADcula>

- Peterson, T. (2019). Bitcoin Spreads Like a Virus. *SSRN Electronic Journal*.
<https://doi.org/10.2139/ssrn.3356098>
- Phillips, R. C., & Gorse, D. (2017). Predicting cryptocurrency price bubbles using social media data and epidemic modelling. *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*. <https://doi.org/10.1109/ssci.2017.8280809>
- Phillips, R. C., & Gorse, D. (2018a). Cryptocurrency price drivers: Wavelet coherence analysis revisited. *PLOS ONE*, 13(4).
<https://doi.org/10.1371/journal.pone.0195200>
- Phillips, R. C., & Gorse, D. (2018b). Mutual-Excitation of Cryptocurrency Market Returns and Social Media Topics. *Proceedings of the 4th International Conference on Frontiers of Educational Technologies - ICFET '18*, 80–86.
<https://doi.org/10.1145/3233347.3233370>
- Piccardi, T., Redi, M., Colavizza, G., & West, R. (2021). On the Value of Wikipedia as a Gateway to the Web. *Proceedings of the Web Conference 2021*.
<https://doi.org/10.1145/3442381.3450136>
- Polasik, M., Piotrowska, A. I., Wisniewski, T. P., Kotkowski, R., & Lightfoot, G. (2015). Price Fluctuations and the Use of Bitcoin: An Empirical Inquiry. *International Journal of Electronic Commerce*, 20(1), 9–49.
<https://doi.org/10.1080/10864415.2016.1061413>
- Prajapati, P. (2020, 16 enero). *Predictive analysis of Bitcoin price considering social sentiments*. arXiv.Org. Recuperado 3 de septiembre de 2022, de
<https://arxiv.org/abs/2001.10343>

Precio, gráficos, capitalización de mercado de Bitcoin (BTC). (s. f.).

CoinMarketCap. Recuperado 29 de julio de 2022, de

<https://coinmarketcap.com/es/currencies/bitcoin/>

PRESS KIT - SOCIAL BLADE. (s. f.). Social Blade. Recuperado 3 de agosto de 2022, de <https://socialblade.com/info/kit>

Privacy Policy. (s. f.). TwitchMetrics. Recuperado 3 de agosto de 2022, de <https://www.twitchmetrics.net/privacy>

Qiang, W., Zheng, X. & Hsu, C. (2016). *Cloud Computing and Big Data: Second International Conference, CloudCom-Asia 2015, Huangshan, China, June 17-19, 2015, Revised Selected Papers (Lecture . . . Science Book 9106) (English Edition)* (1st ed. 2015). Springer.

R: The R Project for Statistical Computing. (s. f.). Recuperado 1 de noviembre de 2022, de <https://www.r-project.org/>

Ram, A. J. (2019). Bitcoin as a new asset class. *Meditari Accountancy Research*, 27(1), 147–168. <https://doi.org/10.1108/medar-11-2017-0241>

Rani S, Y., & Prerana.M. (2021). Social-Media Influence on the Investment Decisions Among the Young Adults in India. *Asia-Pacific Journal of Management and Technology*, 02(01), 17–26. <https://doi.org/10.46977/apjmt.2021v02i01.003>

Rotten Tomatoes: About. (s. f.). Rotten Tomatoes. Recuperado 3 de agosto de 2022, de <https://www.rottentomatoes.com/about#:~:text=Rotten%20Tomatoes%20and%20the%20Tomatometer,in%20theaters%20and%20at%20home>

Rotten Tomatoes: Movies | TV Shows | Movie Trailers | Reviews. (s. f.). Rotten Tomatoes. Recuperado 3 de agosto de 2022, de <https://www.rottentomatoes.com/>

Sapuric, S., Kokkinaki, A., & Georgiou, I. (2020). The relationship between Bitcoin returns, volatility and volume: asymmetric GARCH modeling. *Journal of Enterprise Information Management, ahead-of*(ahead-of-print). <https://doi.org/10.1108/jeim-10-2018-0228>

Sattarov, O., Jeon, H. S., Oh, R., & Lee, J. D. (2020). Forecasting Bitcoin Price Fluctuation by Twitter Sentiment Analysis. *2020 International Conference on Information Science and Communications Technologies (ICISCT)*.
<https://doi.org/10.1109/icisct50599.2020.9351527>

Scharnowski, S. (2021). Understanding Bitcoin liquidity. *Finance Research Letters*, 38. <https://doi.org/10.1016/j.frl.2020.101477>

Schober, P., Boer, C. & Schwarte, L. A. (2018). Correlation Coefficients. *Anesthesia & Analgesia*, 126(5), 1763-1768.
<https://doi.org/10.1213/ane.0000000000002864>

Sebastião, H., & Godinho, P. (2020). Bitcoin futures: An effective tool for hedging cryptocurrencies. *Finance Research Letters*, 33. <https://doi.org/10.1016/j.frl.2019.07.003>

Senarathne, C. W. (2019). Possible Impact of Facebook's Libra on Volatility of Bitcoin: Evidence from Initial Coin Offer Funding Data. *Management of Organizations: Systematic Research*, 81(1), 87–100. <https://doi.org/10.1515/mosr-2019-0006>

Senthilnathan, S. (2019). Usefulness of Correlation Analysis. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3416918>

Sephton, C. (s. f.). *Bitcoin price today, BTC to USD live, marketcap and chart*. CoinMarketCap. Recuperado 3 de agosto de 2022, de
<https://coinmarketcap.com/currencies/bitcoin/historical-data/>

Serie. (s. f.). Diccionario de la lengua española. Recuperado 29 de agosto de 2022, de <https://dle.rae.es/serie>

Shen, D., Urquhart, A., & Wang, P. (2019). Does twitter predict Bitcoin? *Economics Letters*, 174, 118–122. <https://doi.org/10.1016/j.econlet.2018.11.007>

Shilov, K. D., & Zubarev, A. V. (2021). Evolution of bitcoin as a Financial Asset. *Finance: Theory and Practice*, 25(5), 150–171. <https://doi.org/10.26794/2587-5671-2021-25-5-150-171>

Shrestha, M. B. & Bhatta, G. R. (2018). Selecting appropriate methodological framework for time series data analysis. *The Journal of Finance and Data Science*, 4(2), 71-89. <https://doi.org/10.1016/j.jfds.2017.11.001>

Shruti, Roy, S. D., & Zeng, W. (2014, julio). Influence of social media on performance of movies. *2014 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, 1–6. <https://doi.org/10.1109/icmew.2014.6890664>

Shu, M., Song, R., & Zhu, W. (2021). The 2021 Bitcoin Bubbles and Crashes—Detection and Classification. *Stats*, 4(4), 950–970. <https://doi.org/10.3390/stats4040056>

Sifat, I. M., Mohamad, A., & Mohamed Shariff, M. S. B. (2019). Lead-Lag relationship between Bitcoin and Ethereum: Evidence from hourly and daily data. *Research in International Business and Finance*, 50, 306–321. <https://doi.org/10.1016/j.ribaf.2019.06.012>

Silhavy, P., Silhavy, P., & Prokopova, Z. (2020). *Software Engineering Perspectives in Intelligent Systems*. Springer Publishing. https://doi.org/10.1007/978-3-030-63322-6_55

Smales, L. (2019). Bitcoin as a safe haven: Is it even worth considering? *Finance Research Letters*, 30, 385–393. <https://doi.org/10.1016/j.frl.2018.11.002>

Speed, A., Burnett, A., Owens, C., & Parsons, M. (2022). Beyond the Game: Understanding Why People Enjoy Viewing Twitch. *SSRN Electronic Journal*.
<https://doi.org/10.2139/ssrn.4077503>

Spilker, H. S., Ask, K., & Hansen, M. (2018). The new practices and infrastructures of participation: how the popularity of Twitch.tv challenges old and new ideas about television viewing. *Information, Communication & Society*, 23(4), 605–620.
<https://doi.org/10.1080/1369118x.2018.1529193>

Statista. (2022, 27 julio). *Global market share of search engines 2010–2022*. Recuperado 3 de septiembre de 2022, de
<https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/>

Statistical software for data science | Stata. (s. f.). Stata. Recuperado 4 de agosto de 2022, de <https://www.stata.com/>

Steinbeck, H., Teusner, R., & Meinel, C. (2021). Teaching the Masses on Twitch: An Initial Exploration of Educational Live-Streaming. *Proceedings of the Eighth ACM Conference on Learning @ Scale*, 275–278. <https://doi.org/10.1145/3430895.3460157>

Stern, D. I. (2011). From Correlation to Granger Causality. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1959624>

Streissguth, T. (2022, 22 julio). *Bitcoin Price Prediction 2022*. Nasdaq. Recuperado 29 de julio de 2022, de <https://www.nasdaq.com/articles/bitcoin-price-prediction-2022>

Su, C. W., Qin, M., Tao, R., Shao, X. F., Albu, L. L., & Umar, M. (2020). Can Bitcoin hedge the risks of geopolitical events? *Technological Forecasting and Social Change*, 159. <https://doi.org/10.1016/j.techfore.2020.120182>

Su, C. W., Qin, M., Tao, R., & Zhang, X. (2020). Is the status of gold threatened by Bitcoin? *Economic Research-Ekonomska Istraživanja*, 33(1), 420–437.

<https://doi.org/10.1080/1331677x.2020.1718524>

Tan, L., Ng, S. H., Omar, A., & Karupaiah, T. (2018). What's on YouTube? A Case Study on Food and Beverage Advertising in Videos Targeted at Children on Social Media. *Childhood Obesity*, 14(5), 280–290. <https://doi.org/10.1089/chi.2018.0037>

TERMS OF SERVICE - SOCIAL BLADE. (s. f.). Social Blade. Recuperado 3 de agosto de 2022, de <https://socialblade.com/info/terms>

Tomassilli, A., Huin, N., & Giroire, F. (2019). The Structured Way of Dealing with Heterogeneous Live Streaming Systems. *Advances on P2P, Parallel, Grid, Cloud and Internet Computing*, 417–426. https://doi.org/10.1007/978-3-030-33509-0_38

Torras, L., & Castellana, A. (2021, 3 septiembre). *Perspectivas y retos de Bitcoin y la revolución Blockchain - IEF*. IEF Institut d'Estudis Financers. Recuperado 29 de julio de 2022, de <https://www.iefweb.org/es/publicacion-odf/perspectivas-y-retos-de-bitcoin-y-la-revolucion-blockchain/>

Trends. (s. f.). Google Trends. Recuperado 3 de agosto de 2022, de <https://trends.google.es/trends/?geo=CL>

Twitch Channels, Games and Global Statistics. (s. f.). TwitchTracker. Recuperado 3 de agosto de 2022, de <https://twitchtracker.com/>

Twitch Statistics & Charts. (s. f.). TwitchTracker. Recuperado 2 de agosto de 2022, de <https://twitchtracker.com/statistics>

Twitch Stats and Analytics - Games & Channels - SullyGnome. (s. f.). SullyGnome. Recuperado 3 de agosto de 2022, de <https://sullygnome.com/>

Twitchmetrics - Streamers and Games. (s. f.). Twitch Metrics. Recuperado 3 de agosto de 2022, de <https://www.twitchmetrics.net/>

Uras, N., Marchesi, L., Marchesi, M., & Tonelli, R. (2020). Forecasting Bitcoin closing price series using linear regression and neural networks models. *PeerJ Computer Science*, 6. <https://doi.org/10.7717/peerj-cs.279>

Urquhart, A. (2018). What causes the attention of Bitcoin? *Economics Letters*, 166, 40–44. <https://doi.org/10.1016/j.econlet.2018.02.017>

Uva, D., & Rosenthal, B. (2018, noviembre). *Competing online: a netnographic study on twitch influencers*. FGV Sistema de Bibliotecas - FGV Repositorio Digital. <https://bibliotecadigital.fgv.br/dspace/handle/10438/25851>

Vetter, T. R. & Schober, P. (2018). Regression. *Anesthesia & Analgesia*, 127(1), 277-283. <https://doi.org/10.1213/ane.00000000000003424>

Vidal-Tomás, D. (2021). An investigation of cryptocurrency data: the market that never sleeps. *Quantitative Finance*, 21(12), 2007–2024. <https://doi.org/10.1080/14697688.2021.1930124>

Vidal-Tomás, D. (2022). Which cryptocurrency data sources should scholars use? *International Review of Financial Analysis*, 81. <https://doi.org/10.1016/j.irfa.2022.102061>

What are awards and how do I give them? (2021, 9 octubre). Reddit Help. Recuperado 3 de septiembre de 2022, de <https://reddit.zendesk.com/hc/en-us/articles/360043034132-What-are-awards-and-how-do-I-give-them->

Wolfson, S. N. (2015). Bitcoin: The Early Market. *Journal of Business & Economics Research (JBER)*, 13(4), 201–214. <https://doi.org/10.19030/jber.v13i4.9452>

Wollborn, P., Holder, U., & Dornekott, D. (2021). Entrepreneurial Efforts and Opportunity Costs: Evidence from Twitch Streamers. *SSRN Electronic Journal*.

<https://doi.org/10.2139/ssrn.3801017>

Woodcock, J., & Johnson, M. R. (2019). The Affective Labor and Performance of Live Streaming on Twitch.tv. *Television & New Media*, 20(8), 813–823.

<https://doi.org/10.1177/1527476419851077>

Wooley, S., Edmonds, A., Bagavathi, A., & Krishnan, S. (2019). Extracting Cryptocurrency Price Movements from the Reddit Network Sentiment. *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*.

<https://doi.org/10.1109/icmla.2019.00093>

Wu, C. H., Lu, C. C., Ma, Y. F., & Lu, R. S. (2018). A New Forecasting Framework for Bitcoin Price with LSTM. *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, 168–175. <https://doi.org/10.1109/icdmw.2018.00032>

Xu, H. & Deng, Y. (2018). Dependent Evidence Combination Based on Shearman Coefficient and Pearson Coefficient. *IEEE Access*, 6, 11634-11640.

<https://doi.org/10.1109/access.2017.2783320>

Yahoo Search. (s. f.). Yahoo! Recuperado 3 de septiembre de 2022, de <https://search.yahoo.com/>

Yao, W., Xu, K., & Li, Q. (2019, junio). Exploring the Influence of News Articles on Bitcoin Price with Machine Learning. *2019 IEEE Symposium on Computers and Communications (ISCC)*, 1–6. <https://doi.org/10.1109/iscc47284.2019.8969596>

Yatsyk, T. (2020). A MULTIPLE LINEAR REGRESSION MODEL FOR CRYPTOCURRENCY PRICE IN THE FINANCIAL ANALYSIS AND ACCOUNTING.

European Journal of Economics and Management, 6(5), 100–107.

<https://doi.org/10.46340/eujem.2020.6.5.13>

Yaya, O. S., Ogbonna, A. E., & Olubusoye, O. E. (2019). How persistent and dynamic inter-dependent are pricing of Bitcoin to other cryptocurrencies before and after 2017/18 crash? *Physica A: Statistical Mechanics and its Applications*, 531.

<https://doi.org/10.1016/j.physa.2019.121732>

Yelowitz, A., & Wilson, M. (2015). Characteristics of Bitcoin users: an analysis of Google search data. *Applied Economics Letters*, 22(13), 1030–1036.

<https://doi.org/10.1080/13504851.2014.995359>

Yermack, D. (2015). Is Bitcoin a Real Currency? An Economic Appraisal. *Handbook of Digital Currency*, 31–43. <https://doi.org/10.1016/b978-0-12-802117-0.00002-3>

Yi, S., Xu, Z., & Wang, G. J. (2018). Volatility connectedness in the cryptocurrency market: Is Bitcoin a dominant cryptocurrency? *International Review of Financial Analysis*, 60, 98–114. <https://doi.org/10.1016/j.irfa.2018.08.012>

YouTube. (s. f.). YouTube. Recuperado 4 de agosto de 2022, de <https://www.youtube.com/>

YouTube Analytics. (s. f.). YouTube Analytics. Recuperado 3 de agosto de 2022, de <https://analytics.youtube.com/>

Yu, J. H., Kang, J., & Park, S. (2019). Information availability and return volatility in the bitcoin Market: Analyzing differences of user opinion and interest. *Information Processing & Management*, 56(3), 721–732.

<https://doi.org/10.1016/j.ipm.2018.12.002>

Zhai, Y. (2020). Research on Adding Biotech and Bitcoin to Portfolio. *2020 2nd International Conference on Economic Management and Model Engineering (ICEMME)*.
<https://doi.org/10.1109/icemme51517.2020.00042>

Zhang, W., & Wang, P. (2020). Investor attention and the pricing of cryptocurrency market. *Evolutionary and Institutional Economics Review*, 17(2), 445–468.
<https://doi.org/10.1007/s40844-020-00182-1>

Zhang, W., Wang, P., Li, X., & Shen, D. (2018). Quantifying the cross-correlations between online searches and Bitcoin market. *Physica A: Statistical Mechanics and its Applications*, 509, 657–672. <https://doi.org/10.1016/j.physa.2018.06.073>

Zhang, Y., Zhang, Y., Shen, D., & Zhang, W. (2017). Investor sentiment and stock returns: Evidence from provincial TV audience rating in China. *Physica A: Statistical Mechanics and its Applications*, 466, 288–294. <https://doi.org/10.1016/j.physa.2016.09.043>

Zhou, Y., Ahmad, Z., Alsuhabi, H., Yusuf, M., Alkhairy, I., & Sharawy, A. M. (2021). Impact of YouTube Advertising on Sales with Regression Analysis and Statistical Modeling: Usefulness of Online Media in Business. *Computational Intelligence and Neuroscience*, 2021, 1–10. <https://doi.org/10.1155/2021/9863155>

10. ANEXOS

10.1. Apéndice 1: Lista de producciones audiovisuales

N°	Título	Año	Tipo	Enlace IMDb
1	How to Sell Drugs Online (Fast)	2019	Serie de TV	https://www.imdb.com/title/tt9184994/?ref=kw_li_tt
2	Unfriended: Dark Web	2018	Película	https://www.imdb.com/title/tt4761916/?ref=kw_li_tt
3	Dope	2015	Película	https://www.imdb.com/title/tt3850214/?ref=kw_li_tt
4	Crypto	2019	Película	https://www.imdb.com/title/tt8563452/?ref=kw_li_tt
5	Hacker (Anonymous)	2016	Película	https://www.imdb.com/title/tt3173594/?ref=kw_li_tt
6	Deep Web	2015	Documental	https://www.imdb.com/title/tt3312868/?ref=kw_li_tt
7	Cryptopia: Bitcoin, Blockchains and the Future of the Internet	2020	Documental	https://www.imdb.com/title/tt9203586/?ref=kw_li_tt
8	Bitcoin: The End of Money as We Know It	2015	Documental	https://www.imdb.com/title/tt4654844/?ref=kw_li_tt
9	Hidden Secrets of Money	2013	Serie de TV	https://www.imdb.com/title/tt10375812/?ref=kw_li_tt
10	Keiser Report	2009	Serie de TV	https://www.imdb.com/title/tt2135453/?ref=kw_li_tt
11	The Badger	2020	Película	https://www.imdb.com/title/tt11222962/?ref=kw_li_tt
12	Viraali	2017	Película	https://www.imdb.com/title/tt4554360/?ref=kw_li_tt
13	Bitcoin Heist	2016	Película	https://www.imdb.com/title/tt4911780/?ref=fn_tt_tt_5
14	The Rise and Rise of Bitcoin (Bitcoin en Alza)	2014	Documental	https://www.imdb.com/title/tt2821314/?ref=fn_tt_tt_7
15	Banking on Bitcoin	2016	Documental	https://www.imdb.com/title/tt5033790/?ref=fn_tt_tt_2_1
16	StartUp	2016	Serie de TV	https://www.imdb.com/title/tt5028002/?ref=kw_li_tt
17	Forhøret (Face to face / Cara a cara)	2019	Serie de TV	https://www.imdb.com/title/tt9695118/?ref=kw_li_tt

10.2. Apéndice 2: Muestra de canales de YouTube

En el presente apéndice se dispone de la nómina de canales que fueron considerados para elaborar las series de tiempo sobre YouTube.

Nº	Nombre	Enlace	Visualizaciones totales
1	BitBoy Crypto	https://www.youtube.com/channel/UCjemQfjaXAzA-95RKoy9n_g/featured	209.050.875
2	Andrei Jikh	https://www.youtube.com/c/AndreiJikh/featured	199.646.169
3	Coin Bureau	https://www.youtube.com/c/CoinBureau/featured	148.885.670
4	Altcoin Daily	https://www.youtube.com/c/AltcoinDaily/featured	128.120.714
5	TechLead	https://www.youtube.com/c/TechLead/featured	116.044.421
6	CryptosRUs	https://www.youtube.com/c/CryptosRUs/featured	101.275.698
7	Waqar Zaka	https://www.youtube.com/c/waqarzaka/featured	98.100.395
8	The Moon	https://www.youtube.com/c/TheMoonCarl/featured	87.158.678
9	MMCrypto	https://www.youtube.com/c/MMCryptoTube/featured	82.415.786
10	Crypto Banter	https://www.youtube.com/c/CryptoBanterGroup/featured	64.893.584
11	FTX Official	https://www.youtube.com/c/FTXOfficial/featured	60.671.497
12	Lark Davis	https://www.youtube.com/c/TheCryptoLark/featured	56.041.600
13	Brian Jung	https://www.youtube.com/c/BrianJungy/featured	55.293.732
14	Michael Wrubel	https://www.youtube.com/c/MichaelWrubel/featured	54.523.780
15	Alex Becker's Channel	https://www.youtube.com/c/AlexBeckersChannel/featured	50.464.594
16	Kenan Grace	https://www.youtube.com/c/KenanGrace/featured	47.651.608
17	The Modern Investor	https://www.youtube.com/c/TheModernInvestor/featured	45.484.865
18	EllioTrades Crypto	https://www.youtube.com/c/FUDTV/featured	43.204.882

19	InvestAnswers	https://www.youtube.com/c/InvestAnswers/featured	41.854.661
20	Chico Crypto	https://www.youtube.com/c/ChicoCrypto/featured	39.136.984
21	JRNY Crypto	https://www.youtube.com/c/JRNYCrypto/featured	38.162.446
22	RUN GUYS	https://www.youtube.com/c/THERUNGUYS/featured	37.728.764
23	Crypto Mason 2	https://www.youtube.com/c/CryptoMason2/featured	37.586.105
24	Austin Hilton	https://www.youtube.com/c/AustinHilton/featured	35.555.825
25	99Bitcoins	https://www.youtube.com/c/Bitcoinwithpaypal/featured	34.657.949
26	MoneyZG	https://www.youtube.com/c/MoneyZG/featured	33.704.248
27	Crypto Zombie	https://www.youtube.com/c/CryptoZombie/featured	31.807.524
28	Rich Investing	https://www.youtube.com/c/RichInvesting/featured	31.069.751
29	Tyler S	https://www.youtube.com/c/TylerSCrypto/featured	29.721.679
30	Satoshi Stacker	https://www.youtube.com/channel/UCGDjpWZV-bU-sLSnhInCfKQ/featured	29.430.335
31	Paul Barron Network	https://www.youtube.com/c/PaulBarronNetwork/featured	28.436.573
32	Crypto Capital Venture	https://www.youtube.com/c/CryptoCapitalVenture/featured	27.131.589
33	Finance News	https://www.youtube.com/c/FinanceNewsMedia/featured	24.501.724
34	Jason Pizzino	https://www.youtube.com/c/JasonPizzinoOfficial/featured	24.279.635
35	Dapp University	https://www.youtube.com/c/DappUniversity	24.168.759
36	Crypto Jebb	https://www.youtube.com/c/CryptoJebb/featured	23.666.010
37	Jamie Tree	https://www.youtube.com/c/JamieTree/featured	23.076.843
38	Eric Krown Crypto	https://www.youtube.com/c/KrownsCryptoCave/featured	22.067.396
39	Crypto Love	https://www.youtube.com/c/CryptoLove/featured	21.480.837
40	Raishiz Crypto	https://www.youtube.com/c/RaishizCrypto/featured	21.274.843
41	Punto Cripto	https://www.youtube.com/c/PuntoCripto/featured	20.755.239

42	Crypto Crew University	https://www.youtube.com/c/CryptoCrewUniversity/featured	20.574.163
43	Matthew Perry	https://www.youtube.com/user/nikeocho11/featured	19.979.775
44	Alex Cobb	https://www.youtube.com/c/AlexCobbXRP/featured	15.351.118
45	Cheeky Crypto	https://www.youtube.com/c/CheekyCrypto/featured	15.155.508
46	ilme aalim	https://www.youtube.com/c/ilmeaalim/featured	15.011.222
47	Crypto Aman	https://www.youtube.com/c/CryptoAman/featured	14.199.140
48	Sami Loyal - Forflies	https://www.youtube.com/c/Forflies/featured	14.062.873
49	Joe Parys	https://www.youtube.com/c/JoeParysCrypto/featured	13.890.596
50	Mining Chamber	https://www.youtube.com/c/MiningChamber/featured	13.247.987
51	Tech Ranjan	https://www.youtube.com/c/TechRanjan/featured	13.014.081
52	Eddie Moon	https://www.youtube.com/c/eddiemoon/featured	12.903.881
53	Crypto Casey	https://www.youtube.com/c/CryptoCasey/featured	12.238.204
54	Late Night Grind	https://www.youtube.com/c/LateNightGrind/featured	11.095.664
55	7 INGRESOS	https://www.youtube.com/c/7INGRESOS/featured	11.033.771
56	Only The SAVVY	https://www.youtube.com/c/OnlyTheSAVVY/featured	10.861.857
57	CryptoJack	https://www.youtube.com/c/CryptoJackk/about	10.654.340
58	Crypto World	https://www.youtube.com/c/CryptoWorldJosh/featured	10.627.785
59	GNF Learning	https://www.youtube.com/channel/UCDxlvM8pMYjt1MYzUJFLB4A/featured	10.046.516
60	Crypto Tips	https://www.youtube.com/c/CryptoTips/featured	9.833.823
61	Token Metrics	https://www.youtube.com/c/TokenMetrics/featured	9.631.899
62	More Crypto Online	https://www.youtube.com/channel/UCnglhBkikUe6e7tZTjpKK7Q/featured	9.545.465
63	Crypto Anup	https://www.youtube.com/channel/UCY7zYMsilAA_sP1mY--l23Q/featured	9.509.497
64	Crypto Le Trone	https://www.youtube.com/c/LeTroneCrypto/featured	9.293.410

65	Savvy Finance	https://www.youtube.com/c/SavvyFinanceOfficial/featured	9.101.047
66	Crypto Jargon	https://www.youtube.com/c/CryptoJargon/featured	9.013.694
67	Crypto Mason	https://www.youtube.com/c/CryptoMason/featured	8.232.142
68	Crypto Rover	https://www.youtube.com/c/CryptoRover/featured	7.688.752
69	Michael Wyann	https://www.youtube.com/c/MichaelWyann/featured	7.173.235
70	Crypto Kirby Trading	https://www.youtube.com/channel/UCOaew10hdmtfa0MinTjOBqg/featured	6.975.429
71	Mayank Kharayat	https://www.youtube.com/c/LookAround2020/featured	6.943.049
72	Online Raketero	https://www.youtube.com/c/OnlineRaketero/featured	6.881.405
73	Crypto Notes	https://www.youtube.com/c/CryptoNotes/featured	6.375.987
74	Mauro Caimi - Trading On	https://www.youtube.com/c/TradingOn/featured	5.896.959
75	AllinCrypto	https://www.youtube.com/channel/UCIFNxD5mGJa-XIZlcX2ko2A	5.809.000
76	Criptoniano	https://www.youtube.com/c/Criptoniano/featured	5.718.838
77	AuzSOME Austin	https://www.youtube.com/c/AuzSOMEAustin/featured	5.671.270
78	TODAY CRYPTO	https://www.youtube.com/c/TODAYCRYPTO/featured	5.114.222
79	Roger Garcia	https://www.youtube.com/c/RogerGarciaa/featured	3.972.196
80	Blockgeeks	https://www.youtube.com/c/BlockGeeks/featured	3.789.208
81	Crypto Marg	https://www.youtube.com/c/cryptomarg/featured	3.502.630
82	Enter The Crypto Matrix	https://www.youtube.com/c/EnterTheCryptoMatrix/featured	3.378.624
83	The Crypto Factor	https://www.youtube.com/c/TheCryptoFactor/featured	3.274.069
84	Crypto SaQi	https://www.youtube.com/c/CryptoSaQi/featured	3.122.010
85	Classy Crypto	https://www.youtube.com/c/ClassyCryptoYT/featured	3.100.718
86	Crypto Sauce	https://www.youtube.com/channel/UC7VygmdW2oCaUvQf9uECetA/featured	2.813.674
87	Coinbase	https://www.youtube.com/c/coinbase/featured	2.606.497

88	To The Moon	https://www.youtube.com/c/Gacha4Life/featured	2.554.379
89	Crypto Gems	https://www.youtube.com/c/CryptoGemsOfficial/featured	2.505.451
90	James Crypto Guru	https://www.youtube.com/c/JamesCryptoGuru/featured	2.043.538
91	Amila Bro	https://www.youtube.com/c/AmilaBro/featured	1.814.623
92	Informe Crypto	https://www.youtube.com/c/InformeCrypto/featured	1.600.519
93	Crypto Giant	https://www.youtube.com/channel/UC6lyqeFttjBRFWuWvL6vJOQ/featured	1.526.461
94	SlavikCrypto	https://www.youtube.com/c/SlavikCrypto/featured	1.289.995
95	DailyDoseCrypto	https://www.youtube.com/c/DailyDoseCrypto/featured	1.014.050
96	HORA CRIPTO	https://www.youtube.com/channel/UCuCJEQsqun9ZYobkpROirsQ/about	687.503
97	Library Of Wealth	https://www.youtube.com/channel/UC6OOQMjYX09OyrSRPABNxA/featured	486.145
98	Michael Pizzino	https://www.youtube.com/c/MichaelPizzino	407.821
99	Crypto Professor	https://www.youtube.com/c/CryptoProfessor/featured	388.180
100	Latest Simplified	https://www.youtube.com/channel/UCkHHR09XwYHDv7qnrXUcMIA/featured	31.500

10.3. Apéndice 3: Muestra de canales de Twitch

En el este apéndice se dispone de la lista de canales que fueron escogidos para crear las series de tiempo sobre Twitch.

10.3.1. Procedimiento para la recopilación de datos para canales de Twitch:

10.3.1.1. Selección de canales sobre criptomonedas

Inicialmente, tomando en cuenta la elección de base de datos detallada en el marco teórico, se intenta usar TwitchTracker para reconocer a los canales que transmiten contenido acerca de las criptomonedas. Sin embargo, al buscar términos como “Cryptocurrency” o “Crypto” los buscadores del sitio web no entregan rankings alrededor de dichas palabras, sino que informan canales que poseen similitudes con los nombres de algunos canales o

“Juegos”²⁶. Ante tal dificultad para identificar los canales buscados únicamente por TwitchTracker, se decide complementar con los buscadores de algunas de las otras fuentes revisadas con anterioridad, como son TwitchMetrics y Sullygnome. En ellas los buscadores sí logran entregar rankings²⁷ construidos alrededor las criptomonedas, donde la categoría “Crypto” es aquella que es normalmente es utilizada para transmitir contenido sobre criptodivisas (en base a la exploración realizada en esta recolección de datos). La revisión de los rankings se realiza desde las posiciones más altas hacia las más bajas.

Frente a lo anterior se identifica a los canales por medio de TwitchMetrics y Sullygnome, para luego investigar esos creadores de contenido a mayor profundidad en TwitchTracker. Al ser esta la plataforma donde es posible encontrar los datos más apropiados para este estudio, acorde a la elección de bases de datos realizada.

A modo de criterio de selección se verifica que cada canal revisado efectivamente cree videos en vivo sobre criptomonedas, recurriendo a la pestaña “Games” que dispone TwitchTracker para cada cuenta examinada. En ella el sitio web entrega la distribución porcentual de las categorías o “Juegos” dentro de los que el canal suele transmitir.

El criterio para considerar que un canal tiene entre sus principales programaciones a las criptomonedas se compone de los siguientes puntos:

- Categorías como “Crypto”, “Bitcoin” u otras que contengan alguno estos dos términos anteriores (Por ejemplo: “CryptoMines”, “Bitcoin Billionaire”, “Bitcoin highway”, etc.) están presentes en los primeros diez puestos de la

²⁶ “Juegos” es la forma en la que se hace referencia a la temática central de cada transmisión en vivo. Aún cuando existen categorías que no corresponden legítimamente a videojuegos, como los programas de conversaciones, música, arte y Criptomonedas (Que se reconoce en Twitch primariamente como “Crypto”)

²⁷ Estos rankings son ordenados según el número de espectadores promedio en TwitchMetrics. Mientras que en Sullygnome es posible elegir el indicador por el que se desea ordenar los canales de mayor a menor.

lista presentada por TwitchTracker en la pestaña “Games” de cada canal. La cual se ordena con respecto a la cantidad de horas transmitidas en cada “Juego”.

- El promedio de espectadores correspondiente a “Crypto” y categorías afín no debe ser significativamente menor²⁸ a los valores obtenidos en otros “Juegos” más arriba dentro de la lista “Games”. Ya que eso indicaría que el creador de contenido atrae un mayor público al transmitir por más tiempo otras temáticas distintas a las criptomonedas. Por lo tanto, en esos casos “Crypto” no se consideraría una de las principales causas por las que la audiencia sintoniza un determinado canal.

Con esto se garantiza la selección de canales que poseen popularidad debido a las criptomonedas y no a otros tipos de contenido.

- Adicionalmente, se examina que la descripción (pestaña “Acerca de” o “About”) del canal de Twitch o las redes sociales vinculadas a la cuenta (Twitter, Instagram, u otra) rectifiquen que se produce contenido relacionado a las criptomonedas. Siendo esta una forma de verificar que el creador de contenido tenga a las criptomonedas como una de sus temáticas principales y recurrentes, excluyendo así aquellos casos donde se transmite “Crypto” de manera esporádica o casual.

A partir de estos tres enunciados se construye un procedimiento capaz de admitir solo aquellos creadores de contenido que logran alcanzar a un gran número de personas

²⁸ Para este estudio se considera a la cantidad de espectadores promedio de Crypto “significativamente menor” cuando su valor es menos de la mitad de lo obtenido en otras categorías.

interesadas en las criptomonedas, al mismo tiempo que muestran estar relacionados estrechamente con esta temática.

Cabe mencionar que al momento de seleccionar los canales y aplicar los criterios correspondientes no se discrimina en torno al lenguaje utilizado en cada uno de ellos. Dado que, tal como fue señalado en la selección de datos de YouTube, se considera a las criptomonedas como un activo muy ligado al internet, y por consiguiente no existen barreras que impidan a personas con diferentes idiomas participar de la industria. Por esa razón se hizo uso del traductor de Google (Google Traductor, s. f.) para inspeccionar aquellas cuentas con idiomas distintos al español o inglés, al ser estas las dos lenguas en las que se tiene las competencias necesarias para leer dentro de la presente investigación.

Luego, en cuanto a los datos registrados sobre cada canal, se anota desde el perfil de cada canal en TwitchTracker el nombre de usuario, enlace de Twitch, cantidad de espectadores promedio, número de seguidores, horas visualizadas por la audiencia, redes sociales, y enlace de TwitchTracker²⁹. Donde los valores numéricos corresponden a aquellos que son representativos de toda la información histórica disponible en la plataforma (sección “All Time”).

Adicionalmente se registran otras dos cifras que hacen alusión a solo las transmisiones sobre “Crypto” y demás categorías afines; el número de espectadores promedio y las horas transmitidas. Estos dos valores se obtienen desde la pestaña “Games” de cada perfil en TwitchTracker y tienen como propósito ser multiplicados entre sí para formar la medida por la que será ordenado el ranking a elaborar en la sección siguiente.

²⁹ Aquella lista se encuentra disponible en el Apéndice 3: Muestra de canales de Twitch.

Puesto que se considera a ese producto como una aproximación a la cantidad de personas que atrae cada canal de Twitch seleccionado a lo largo de todas las transmisiones realizadas.

Luego de obtener una muestra con 100 canales que cumplen los criterios de selección, termina el proceso de selección de canales sobre criptomonedas y se procede a la creación de un ranking a partir de la lista creada.

10.3.1.2. Elaboración de ranking

Con motivo de organizar y priorizar los canales según la cantidad de personas que alcanzan en sus transmisiones, se les ordena en base al producto entre el número de espectadores promedio y las horas transmitidas de aquellas transmisiones que son etiquetadas con categorías de criptomonedas. Lo cual se ajusta, tal como fue hecho en la sección de YouTube, a la línea de pensamiento propuesta por (Lepori, 2015). Quien al investigar los vínculos entre la bolsa de valores estadounidense y programas de televisión señaló que: “Un prerrequisito para que los finales en series de televisión generen una influencia en el precio de los activos financieros es que estos alcancen a un gran número de personas”. En consecuencia, la presente memoria toma como parámetro dicha idea, extendiéndola a lo que se desea estudiar tanto en Twitch como YouTube frente a Bitcoin.

Una vez se alcanza el estado final de ese ranking, se toma como la muestra definitiva a considerar en el estudio a los primeros 50 lugares³⁰. Considerados como aquellos que representan una gran porción de todo el contenido sobre criptomonedas en Twitch, y por lo tanto son capaces de ser analizados con respecto a las variables de Bitcoin.

³⁰ Tanto la elección de seleccionar los 100 principales canales sobre criptomonedas como la decisión de apartar los 50 primeros lugares se ve inspirada por la metodología seguida por (Steinbeck et al., 2021), quienes también utilizan datos de Twitch y consideran que los canales ubicados en esas posiciones dentro de un ranking demuestran poder mantener el interés y la atención de una audiencia (p.2).

Es importante mencionar que muestras de similar tamaño (50 canales) han sido utilizadas al estudiar en Twitch las diferencias interculturales dentro de esta plataforma (Oh et al., 2020) y el ecosistema de canales dedicados a enseñar contenidos en vivo (Steinbeck et al., 2021). Siendo esta una de las razones para seleccionar ese número de muestra, en conjunto con obtener una extensión apropiada para la recopilación de datos de acuerdo a la duración definida para realizar esta investigación (tomando en cuenta el tiempo necesario para extraer cada una de las observaciones en las series de tiempo a considerar, cuyo proceso implicado será cubierto detalladamente en la sección siguiente).

En la siguiente tabla se exponen las cien cuentas de Twitch consideradas como las más relevantes según el producto entre sus espectadores promedio y horas transmitidas bajo las programaciones “Crypto” y “Bitcoin”.

Nº Ranking	Nº muestra	Nombre	Enlace Twitch	% dedicac. a Crypto.	Espect. prom.	Número de seguidores	Hrs Visualiz.	Enlace Twitch Tracker	Espect. Prom. * Hrs Transm. (Crypto / BTC)
1	4	Sonecarox	https://www.twitch.tv/sonecarox	6,30%	738	194.000	14.000.000	https://twitchtracker.com/sonecarox	3.741.157,0
2	53	MarketTraders TV	https://www.twitch.tv/markettraderstv	35,70%	227	48.300	8.340.000	https://twitchtracker.com/markettraderstv	3.575.440,0
3	40	Doctor_Degen	https://www.twitch.tv/doctor_degen	80,60%	213	38.900	2.610.000	https://twitchtracker.com/doctor_degen	2.271.940,0
4	42	TraderSamwise	https://www.twitch.tv/tradersamwise	99,00%	213	22.100	1.820.000	https://twitchtracker.com/tradersamwise	1.823.630,0
5	41	TR4D3R10	https://www.twitch.tv/tr4d3r10	100,00%	111	16.100	1.800.000	https://twitchtracker.com/tr4d3r10	1.806.525,0
6	28	Alienigena404	https://www.twitch.tv/alienigena404	17,80%	244	69.900	2.220.000	https://twitchtracker.com/alienigena404	1.326.005,0

7	20	Lombrajr	https://www.twitch.tv/lombrajr	31,30%	183	25.500	1.210.000	https://twitter.com/lombrajr	833.748,0
8	29	brookz_	https://www.twitch.tv/brookz_	16,80%	259	51.400	1.350.000	https://twitter.com/brookz_	784.971,0
9	43	AlphaExchange	https://www.twitch.tv/alphaexchange	93,50%	103	8.600	590.000	https://twitter.com/alphaexchange	563.535,0
10	44	GooberArcade	https://www.twitch.tv/gooberarcade	97,90%	80	3.700	538.000	https://twitter.com/gooberarcade	544.231,0
11	8	The_Dent	https://www.twitch.tv/the_dent	28,10%	172	61.900	845.000	https://twitter.com/the_dent	502.512,0
12	46	Cryptokking	https://www.twitch.tv/cryptokking	96,70%	53	12.000	371.000	https://twitter.com/cryptokking	459.748,0
13	81	LuxAlgo	https://www.twitch.tv/luxalgo	94,90%	19	5.300	331.000	https://twitter.com/luxalgo	453.460,0
14	60	BitcoinTradingInformant	https://www.twitch.tv/bitcointradinginformant?lang=sv	87,90%	18	7.100	360.000	https://twitter.com/bitcointradinginformant	399.110,0
15	45	cryptogamer97	https://www.twitch.tv/cryptogamer97	86,30%	486	15.000	393.000	https://twitter.com/cryptogamer97	379.168,0
16	11	ZCoinTV	https://www.twitch.tv/zcointv	93,30%	323	13.900	353.000	https://twitter.com/zcointv	363.650,0
17	10	Scottfdx	https://www.twitch.tv/scottfdx	98,30%	281	14.200	339.000	https://twitter.com/scottfdx	350.400,0
18	23	Saser0101	https://www.twitch.tv/saser0101	85,40%	225	14.000	355.000	https://twitter.com/saser0101	335.587,0
19	47	MalakaTV	https://www.twitch.tv/malakatv	9,00%	167	41.400	942.000	https://twitter.com/malakatv	317.107,0
20	24	bellalopez	https://www.twitch.tv/bellalopez	49,60%	150	25.000	623.000	https://twitter.com/bellalopez	314.336,0
21	48	BomberOneBR	https://www.twitch.tv/bomberonebr	100,00%	160	21.100	291.000	https://twitter.com/bomberonebr	300.795,0

								om/bombe ronebr	
22	15	K1m6a	https://www.twitch.tv/k1m6a	100,00%	184	5.200	262.000	https://twitchtracker.com/k1m6a	284.675,0
23	70	Scalper_Chirurgical	https://www.twitch.tv/scalper_chirurgical	100,00%	52	3.300	85.100	https://twitchtracker.com/scalper_chirurgical	283.446,0
24	50	raedtoken	https://www.twitch.tv/raedtoken	82,60%	291	17.700	291.000	https://twitchtracker.com/raedtoken	282.975,0
25	86	whoistherabbit	https://www.twitch.tv/whoistherabbit	74,20%	154	21.600	288.000	https://twitchtracker.com/whoistherabbit	272.244,0
26	12	Khalistas	https://www.twitch.tv/khalistas	98,40%	328	19.800	238.000	https://twitchtracker.com/khalistas	241.670,0
27	51	kojisannft	https://www.twitch.tv/kojisannft	16,80%	453	82.200	855.000	https://twitchtracker.com/kojisannft	234.788,0
28	52	AromazTV	https://www.twitch.tv/aromaztv	36,80%	639	165.000	567.000	https://twitchtracker.com/aromaztv	234.459,0
29	61	리얼블랙카우 (real_black_cow)	https://www.twitch.tv/real_black_cow	71,70%	22	1.100	233.000	https://twitchtracker.com/real_black_cow	227.100,0
30	17	KManu88	https://www.twitch.tv/kmanu88	6,90%	680	213.000	1.260.000	https://twitchtracker.com/kmanu88	220.726,0
31	79	B1nomioCryptoSensei	https://www.twitch.tv/b1nomiocryptosensei	91,80%	70	7.400	221.000	https://twitchtracker.com/b1nomiocryptosensei	216.675,0
32	1	UpOnlyTV	https://www.twitch.tv/uponlytv	64,70%	2.174	32.800	304.000	https://twitchtracker.com/uponlytv	210.036,0
33	3	ColdBloodedShiller	https://www.twitch.tv/coldbloodedshiller	90,70%	518	20.700	199.000	https://twitchtracker.com/coldbloodedshiller	196.968,0
34	54	DrEmrys	https://www.twitch.tv/dremrys	34,50%	246	31.600	321.000	https://twitchtracker.com/dremrys	183.641,0
35	9	muleskinner1	https://www.twitch.tv/muleskinner1	40,50%	298	7.900	318.000	https://twitchtracker.com/muleskinner1	177.888,0

36	55	monnaiesloco	https://www.twitch.tv/monnaiesloco	54,70%	105	9.700	164.000	https://twitter.com/monnaiesloco	166.258,0
37	6	enterthecryptomatrix	https://www.twitch.tv/enterthecryptomatrix	52,20%	1.359	40.800	309.000	https://twitter.com/enterthecryptomatrix	161.376,0
38	13	Kel_Heset	https://www.twitch.tv/kel_heset	99,90%	218	7.500	155.000	https://twitter.com/kel_heset	159.120,0
39	56	ocarteironft	https://www.twitch.tv/ocarteironft	100,00%	223	8.000	147.000	https://twitter.com/ocarteironft	146.288,0
40	22	inversionesenelmundo	https://www.twitch.tv/inversionesenelmundo	33,60%	122	20.300	291.000	https://twitter.com/inversionesenelmundo	143.460,0
41	57	킹나도티비 (nadotv86)	https://www.twitch.tv/nadotv86	56,20%	94	21.400	594.000	https://twitter.com/nadotv86	142.120,0
42	33	BareNakedCrypto_	https://www.twitch.tv/barenakedcrypto_	91,40%	104	7.400	139.000	https://twitter.com/barenakedcrypto_	139.308,0
43	21	Capetlevrai	https://www.twitch.tv/capetlevrai	51,80%	292	27.100	183.000	https://twitter.com/capetlevrai	138.672,0
44	84	SheldonAssad	https://www.twitch.tv/sheldonassad	97,30%	84	5.800	138.000	https://twitter.com/sheldonassad	138.202,0
45	2	Hasheur	https://www.twitch.tv/hasheur	71,40%	1.727	62.700	180.000	https://twitter.com/hasheur	129.116,0
46	58	ShiniJuega	https://www.twitch.tv/shinijuega/about	46,70%	109	8.500	200.000	https://twitter.com/shinijuega	126.688,0
47	31	Jordan_FDX	https://www.twitch.tv/jordan_fdx	99,30%	158	7.100	125.000	https://twitter.com/jordan_fdx	124.504,0
48	34	CryptoLogikFR	https://www.twitch.tv/cryptologikfr	99,60%	165	11.300	124.000	https://twitter.com/cryptologikfr	124.415,0
49	18	0xpicsou	https://www.twitch.tv/0xpicsou	98,20%	467	17.600	126.000	https://twitter.com/0xpicsou	123.888,0
50	88	GiaccoAR	https://www.twitch.tv/giaccoar	51,70%	252	20.200	282.000	https://twitter.com/giaccoar	123.605,4

51	37	MinhArte	https://www.twitch.tv/minharte	26,40%	76	13.600	368.000	https://twitter.com/minharte	120.038,0
52	59	Chey365	https://www.twitch.tv/chey365	83,90%	116	5.600	123.000	https://twitter.com/chey365	114.165,0
53	62	FuturosTraders	https://www.twitch.tv/futurostraders	84,40%	402	16.400	161.000	https://twitter.com/futurostraders	111.878,0
54	82	NanoGiveaway	https://www.twitch.tv/nanogiveaway	100,00%	16	1.200	95.900	https://twitter.com/nanogiveaway	110.618,0
55	80	beecryptolive	https://www.twitch.tv/beecryptolive	100,00%	17	1.600	97.200	https://twitter.com/beecryptolive	107.749,0
56	19	WaveTrading	https://www.twitch.tv/wavetrading	100,00%	177	7.100	98.100	https://twitter.com/wavetrading	98.845,0
57	92	journalducoin	https://www.twitch.tv/journalducoin	54,80%	670	18.000	147.000	https://twitter.com/journalducoin	98.160,0
58	35	CryptoFutur	https://www.twitch.tv/cryptofutur	94,30%	420	26.600	99.900	https://twitter.com/cryptofutur	95.872,0
59	64	FunOnTheRide	https://www.twitch.tv/funontheride	96,00%	208	12.100	93.500	https://twitter.com/funontheride	92.016,0
60	30	cryptojack	https://www.twitch.tv/cryptojack	74,10%	126	8.500	93.000	https://twitter.com/cryptojack	90.848,0
61	89	mrcryptosniper	https://www.twitch.tv/mrcryptosniper	100,00%	30	4.600	75.000	https://twitter.com/mrcryptosniper	90.432,0
62	83	eunetoleao	https://www.twitch.tv/eunetoleao	63,90%	430	12.500	118.000	https://twitter.com/eunetoleao	85.400,0
63	67	desdeceroinvertion	https://www.twitch.tv/desdeceroinvertion	97,50%	113	7.200	85.400	https://twitter.com/desdeceroinvertion	84.870,0
64	63	DappCentre	https://www.twitch.tv/dappcentre	95,00%	329	9.200	81.600	https://twitter.com/dappcentre	84.365,0
65	7	NinjascaIp	https://www.twitch.tv/ninjascaip	100,00%	417	11.100	82.600	https://twitter.com/ninjascaip	82.764,0
66	85	Xiny_	https://www.twitch.tv/xiny_	39,90%	100	10.100	182.000	https://twitter.com/xiny_	82.038,0

67	87	onc394	https://www.twitch.tv/onc394	27,10%	41	8.200	122.000	https://twitter.com/onc394	76.234,0
68	38	CevapSichuan	https://www.twitch.tv/cevapsichuan	97,80%	125	7.200	75.000	https://twitter.com/cevapsichuan	74.549,0
69	74	TradingLatino	https://www.twitch.tv/tradinglatino	99,70%	157	24.900	72.800	https://twitter.com/tradinglatino	73.877,3
70	100	DojiSan	https://www.twitch.tv/dojisan	47,70%	14	2700	73600	https://twitter.com/dojisan	72.355,0
71	94	CryTraGa	https://www.twitch.tv/crytraga	91,10%	160	4.500	73.600	https://twitter.com/crytraga	71.649,0
72	93	ZeroKaal	https://www.twitch.tv/zerokaal	67,80%	111	9.200	88.000	https://twitter.com/zerokaal	64.080,0
73	91	ColomboNFT	https://www.twitch.tv/colombonft	18,40%	48	28.200	177.000	https://twitter.com/colombonft	63.020,0
74	90	TutanKoin	https://www.twitch.tv/tutankoin	42,10%	108	7.100	73.100	https://twitter.com/tutankoin	62.700,0
75	71	luvack36	https://www.twitch.tv/luvack36	70,00%	156	6.000	86.600	https://twitter.com/luvack36	58.824,0
76	95	sergechampagne netradinglive	https://www.twitch.tv/sergechampagnetradinglive	66,80%	45	1.300	113.000	https://twitter.com/sergechampagnetradinglive	58.380,0
77	99	CriptoJimmy	https://www.twitch.tv/cryptojimmy	47,30%	208	7.400	74.800	https://twitter.com/cryptojimmy	55.272,8
78	97	Kaneticia	https://www.twitch.tv/kaneticia	22,10%	64	10.100	169.000	https://twitter.com/kaneticia	54.126,0
79	96	Strompegasus	https://www.twitch.tv/strompegasus	71,60%	99	6.400	61.700	https://twitter.com/strompegasus	53.966,0
80	25	moro8leo	https://www.twitch.tv/moro8leo	91,20%	328	11.800	56.300	https://twitter.com/moro8leo	53.820,0
81	98	erickecuayt	https://www.twitch.tv/erickecuayt	98,00%	142	4.500	51.400	https://twitter.com/erickecuayt	50.976,0

82	32	c0tt0nc4ndyta	https://www.twitch.tv/c0tt0nc4ndyta	66,70%	89	5.100	82.100	https://twitchtracker.com/c0tt0nc4ndyta	48.585,0
83	26	ledgerstatus	https://www.twitch.tv/ledgerstatus	99,50%	251	6.700	37.600	https://twitchtracker.com/ledgerstatus	37.548,0
84	39	fecoxa	https://www.twitch.tv/fecoxa	30,40%	129	10.400	137.000	https://twitchtracker.com/fecoxa	35.098,0
85	5	jimtalbott	https://www.twitch.tv/jimtalbott	98,30%	561	9.200	30.400	https://twitchtracker.com/jimtalbott	31.280,0
86	69	CryptoMillionYT	https://www.twitch.tv/cryptomillionyt	61,70%	104	6.500	39.600	https://twitchtracker.com/cryptomillionyt	29.358,0
87	65	maruushae	https://www.twitch.tv/maruushae	99,00%	66	3.300	24.600	https://twitchtracker.com/maruushae	24.420,0
88	73	chiisansi	https://www.twitch.tv/chiisansi	66,00%	32	3.500	37.200	https://twitchtracker.com/chiisansi	22.920,0
89	14	bugrasukas	https://www.twitch.tv/bugrasukas	4,80%	93	13.600	46.800	https://twitchtracker.com/bugrasukas	16.426,8
90	66	Trader_Ex	https://www.twitch.tv/trader_ex	98,90%	140	2.200	15.400	https://twitchtracker.com/trader_ex	15.478,0
91	68	CryptoSoprano	https://www.twitch.tv/cryptosoprano	70,60%	100	3.600	20.300	https://twitchtracker.com/cryptosoprano	14.200,0
92	16	magustrade	https://www.twitch.tv/magustrade	81,20%	101	2.100	8.100	https://twitchtracker.com/magustrade	8.044,2
93	36	swisscryptojay	https://www.twitch.tv/swisscryptojay	56,10%	179	1.700	12.400	https://twitchtracker.com/swisscryptojay	6.690,8
94	72	EkoKripto	https://www.twitch.tv/ekokripto	100,00%	52	529	5.700	https://twitchtracker.com/ekokripto	5.724,0
95	76	MeraAlberto	https://www.twitch.tv/meraalberto	99,70%	28	1.400	4.800	https://twitchtracker.com/meraalberto	4.814,0
96	49	0xVKTR	https://www.twitch.tv/0xvktr	87,90%	96	4.800	4.500	https://twitchtracker.com/0xvktr	4.492,8
97	27	thepokencompany	https://www.twitch.tv/thepokencompany	60,50%	238	3.600	7.300	https://twitchtracker.com/thepokencompany	3.617,6
98	77	CitadelDispatch	https://www.twitch.tv/citadeldispatch	100,00%	18	634	2.700	https://twitchtracker.com/citadeldispatch	3.040,0

								om/citadel dispatch	
99	75	c2m_xbt	https://www.twitch.tv/c2m_xbt	100,00%	66	642	1.200	https://twitchtracker.com/c2m_xbt	1.199,3
100	78	locosxtrading	https://www.twitch.tv/locosxtrading	100,00%	9	239	929	https://twitchtracker.com/locosxtrading	1.071,1

10.3.1.3. Obtención de datos en TwitchTracker

Al tener el listado de los canales más populares sobre criptomonedas se comienza a construir una base de datos con todas las estadísticas objetivo de cada uno de ellos. Estas son: Espectadores promedio, Horas visualizadas por la audiencia, Horas transmitidas en vivo y la Cantidad de seguidores.

Para acceder a ellas es necesario ingresar en el buscador de TwitchTracker el nombre del canal que se desea consultar. Lo cual conducirá a un reporte con información general sobre la cuenta, donde se debe acceder a la pestaña “Statistics”. En su interior se presentan las cuatro variables que se desean recolectar en formato de gráfico, disponiendo de valores diarios según la frecuencia en que las transmisiones en vivo ocurran. Estos serán traspasados manualmente a una base de datos, repitiendo el proceso para todos los canales seleccionados con anterioridad.

Adicionalmente cada uno de dichos registros de transmisión poseen imágenes que señalan cuáles son las categorías cubiertas en ese día. Por lo tanto, se decide tomar en consideración solo aquellas fechas donde las categorías sobre criptomonedas están presentes, para así minimizar la inclusión de datos sobre otros tipos de programación (como los videojuegos).

Al completar el número total de revisiones se cuenta con una base de datos almacenando series de tiempo individuales para cada canal de Twitch, que comparten las

mismas variables entre sí. Con ello se da paso a la siguiente sección, donde se creará un único conjunto de series de tiempo a utilizar frente a Bitcoin.

10.3.1.4. Consolidación de una única serie de tiempo

Dado que se desea estudiar la relación entre el material audiovisual sobre criptomonedas en Twitch frente a Bitcoin, es necesario unir las series de tiempo individuales de cada canal para obtener solo una que estime a todo el material. Para lograrlo se sumará entre los canales todas las cifras correspondientes a una misma fecha, en cada una de las variables. Obteniendo así una base de datos definitiva compuesta con los siguientes valores:

- Fecha (TW_Date): Indica el día al cual corresponden las cifras de las demás variables. Se interpreta como la variable temporal de las series de tiempo.
- Espectadores promedio (TW_AvgVwrs): Es la media calculada a partir del número de personas viendo las transmisiones en vivo, siendo este originalmente un valor que varía de manera constante mientras cada canal está en línea.
- Horas visualizadas por la audiencia (TW_HrsWtd): Considera el tiempo en que el total de espectadores se mantiene viendo una transmisión en vivo.
- Horas transmitidas en vivo (TW_HrsStrmd): Cuantifica el tiempo que los canales de Twitch se mantienen emitiendo una transmisión en vivo.
- Seguidores ganados en cada periodo (TW_GndFwrs): Describe el número de personas que siguen a un canal.

Esta base de datos será aquella que participará en los modelos frente a las variables del mercado de Bitcoin, no sin antes pasar por las transformaciones de datos necesarias y estadísticas descriptivas correspondientes.