

2018

ANÁLISIS INTELIGENTE DE WEBLOGS DE UN PORTAL WEB UNIVERSITARIO

SALGADO DÍAZ, JORGE

<https://hdl.handle.net/11673/47042>

Repositorio Digital USM, UNIVERSIDAD TECNICA FEDERICO SANTA MARIA

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INFORMÁTICA
VALPARAÍSO - CHILE



“ANÁLISIS INTELIGENTE DE WEBLOGS
DE UN PORTAL WEB UNIVERSITARIO”

JORGE SALGADO DÍAZ

MEMORIA DE TITULACIÓN PARA OPTAR AL TÍTULO DE INGENIERO CIVIL
EN INFORMÁTICA

PROFESOR GUÍA	: CECILIA REYES COVARRUBIAS
PROFESOR CORREFERENTE	: LUIS HEVIA RODRÍGUEZ

Agosto - 2018

*Dedicado a mis padres Jorge e Isabel,
sin ellos nada de esto sería posible.*

Agradecimientos

A mis padres Jorge e Isabel, por el enorme esfuerzo que realizaron durante toda mi etapa universitaria, por todo su apoyo en los diferentes momentos de mi vida, por todas las veces que me recibieron en casa con mucho afecto. Muchas gracias por todo.

A mi hermana Cony, por brindarme su apoyo incondicionalmente, porque sé que siempre puedo contar con ella para lo que sea y por ser mi ejemplo a seguir en los estudios. A Joaco, por ser un buen hermano, por todos los favores que me hiciste y sobre todo por ser una gran persona.

A todos mis amigos de la Universidad, con los que compartí los mejores momentos de mi vida, con los que muchas veces trasnoché estudiando o carreteando. Mención especial a Vito, Tari, Mato, Camilo, Oscar, Jose, Rodolfo. A mis amigos caídos, que también fueron parte importante de esta etapa, Pájaro, Raúl y Negro. A mis compañeros de tenis: Enzo, Nico, Portus y el profesor Enrique por acompañarme cada martes y jueves en los que iba a entrenar, por los gratos momentos vividos en cada torneo que participamos, muchas gracias.

A mis amigos de la vida: Tomacho, Pipe, Waso y Pedro porque gracias a ellos la rutina universitaria pasaba a otro plano. Por todos los asados, juntas, viajes, partidos y por todos los buenos momentos que pasé. Como no olvidar a PMT, que a pesar de tus dificultades fuiste capaz de obtener tu título universitario, muchas gracias por todos los buenos momentos y ser una inspiración de que todo es posible.

A Pato por toda la ayuda brindada, por todos los buenos momentos, por alojarme en su casa durante mi práctica y mi primer año de trabajo, por ser parte de esta memoria.

A mis compañeros de trabajo, muchas gracias por la buena disposición y acogida, en especial a Camilo, Marta, Fabián, Belén, Majo y Rocío. A todos ellos muchas gracias, son más de lo que esperaba...

Resumen

Uno de los principales problemas de las organizaciones es que muchas veces no utilizan su información disponible, lo que lleva a conclusiones erradas, mala gestión y el no seguimiento de objetivos.

Esta memoria se basa en el diseño, construcción e implementación de un sistema de Inteligencia de Negocios, con el fin de obtener el perfil de los visitantes web de un portal universitario, mediante el análisis de sus navegaciones, y orientar la estrategia de marketing digital de una carrera universitaria.

Para ello, se identifican los indicadores claves de rendimiento que entregan la información relevante para apoyar a la toma de decisiones y se utiliza una herramienta de visualización de datos, que permite mostrarlos a través de reportes y gráficos de fácil interpretación para el tomador de decisiones.

Abstract

One of the main problems of organizations is that they often do not use their available information which leads to wrong conclusions, bad management and non-pursuit of objectives.

This memoir is based on the design, construction and implementation of a Business Intelligence system in order to obtain the profile of web visitors through the analysis of information from a university web portal, and guide the digital marketing strategy of a university career.

Initially, key performance indicators are identified that will be used to generate relevant information, through a data visualization tool, that allows the creation of reports and graphs in order to support decision making.

Glosario

DI: Departamento de Informática.

DM: Data Mart.

DW: Data Warehouse.

ETL: Extract, Transform and Load (Extraer, Transformar y Cargar).

GA: Google Analytics.

KPI: Key Performance Indicator (Indicador Clave de Rendimiento).

OLAP: Online Analytical Processing (Procesamiento Analítico en Línea).

SEO: Search Engine Optimization (Trabajo de optimización y aumento de popularidad de un sitio web).

SEM: Search Engine Marketing (Promoción de un sitio web mediante el uso de anuncios de pagos).

URL: Uniform Resource Locator (Localizador Uniforme de Recursos).

UTFSM: Universidad Técnica Federico Santa María.

Índice general

Introducción	1
1. Definición del Problema	3
1.1. Contexto	3
1.2. Situación Actual	5
1.3. Objetivos de una Solución	6
1.3.1. Objetivo General	6
1.3.2. Objetivos Específicos	6
2. Estado del Arte	7
2.1. Inteligencia de Negocios	7
2.1.1. Componentes de BI	8
2.1.1.1. ETL	8
2.1.1.2. Data Warehouse	9
2.1.1.3. Data Mart	9
2.1.1.4. OLAP	10
2.1.1.5. Dashboards	10
2.1.2. Aplicaciones de BI	10
2.2. Análisis Web	11
2.2.1. Análisis web en empresas	12
2.2.2. Google Analytics	13

ÍNDICE GENERAL

2.2.2.1.	Analítica de Contenido	13
2.2.2.2.	Analítica de Móviles	14
2.2.2.3.	Analítica de Conversiones	14
2.2.2.4.	Analítica Social	14
2.2.2.5.	Analítica de Anuncios	15
2.2.3.	Archivos Log de Servidores Web	15
2.2.4.	<i>Clickstream</i>	16
2.2.5.	<i>Metodología para Análisis Web</i>	16
2.2.5.1.	<i>Métricas Básicas</i>	16
2.2.5.2.	<i>Conversiones, objetivos y KPI</i>	17
2.2.5.3.	<i>Análisis Básicos: Web y Contenido</i>	21
2.2.5.4.	Origen y salida de las visitas	22
2.2.5.5.	Perfiles de Navegación	26
2.2.5.6.	Captación, Fidelización y Asiduidad	27
2.2.5.7.	Elaboración de Informes	29
2.3.	Marketing Digital	30
2.3.1.	Tipos de Marketing Digital	31
2.3.2.	Métricas	31
3.	Diseño de la Solución	34
3.1.	Metodología utilizada	34
3.2.	Identificación de Objetivos y KPIs	35
3.2.1.	Objetivos	35
3.2.2.	KPIs o Métricas	36
3.3.	Fuente de Datos y Nivel de Granularidad	37
3.3.1.	Fuentes de Datos	37
3.3.2.	Nivel de Granularidad	40
3.4.	Diseño Modelo de Datos Lógico	41

ÍNDICE GENERAL

3.4.1. Data Mart Audiencia	41
3.4.2. Data Mart Origen	41
3.4.3. Data Mart Comportamiento	41
4. Construcción e Implementación de la Solución	44
4.1. Selección de Herramienta de BI	44
4.2. Conexión Microstrategy a Google Analytics	48
4.3. Modelo de datos Microstrategy	50
4.4. Visualización en Escritorio	51
4.4.1. Dashboard Audiencia	51
4.4.2. Dashboard Comportamiento	55
4.4.3. Dashboard Origen	58
5. Análisis de Resultados	59
5.1. Origen	59
5.1.0.1. General	59
5.1.0.2. Fuente/Medio	60
5.1.0.3. Referencia	61
5.2. Audiencia	62
5.2.1. General	62
5.2.2. Datos demográficos	63
5.2.2.1. Edad	63
5.2.2.2. Género	63
5.2.3. Intereses	64
5.2.4. Información Geográfica	65
5.2.5. Tipo de visitante, Frecuencia y Asiduidad	66
5.2.5.1. Tipo de visitante	66
5.2.5.2. Frecuencia	67
5.2.5.3. Asiduidad	67

ÍNDICE GENERAL

5.2.6. Tecnología y Dispositivos	68
5.2.6.1. Tecnología	68
5.2.6.2. Dispositivos	69
5.2.6.3. Dispositivos Móviles	70
5.3. Comportamiento	70
5.3.1. General	70
5.3.2. Páginas de entrada	72
5.3.3. Páginas de salida	73
5.3.4. Desglose de contenido	74
5.4. Síntesis del Análisis de Resultados	77
6. Conclusiones	80
6.1. Cumplimiento de Objetivos	80
6.1.1. Identificar qué tipo de información es importante o relevante para obtener el perfil de los visitantes	80
6.1.2. Establecer los indicadores claves de rendimiento (KPIs) para obtener el perfil de los visitantes de la web	81
6.1.3. Diseñar una solución de Inteligencia de Negocios en base a los indicadores anteriores	81
6.1.4. Construir e implementar la solución de Inteligencia de Negocios al portal del Departamento de Informática de la UTFSM	81
6.1.5. Realizar una validación de la propuesta mediante el análisis de los resultados obtenidos utilizando metodología de Analítica Web	82
6.2. Beneficios de la solución	82
6.3. Trabajo a futuro	83
Bibliografía	84

Índice de Figuras

1.1. Cómo influye la publicidad online en las compras.	4
1.2. Solución de Inteligencia de Negocios	4
2.1. Modelo BI	11
2.2. Google Analytics: Métrica fidelización	28
3.1. Diseño de la Solución	35
3.2. Api Google Analytics: Habilitación API	39
3.3. Api Google Analytics: Tipos de Credenciales	39
3.4. Api Google Analytics: URLs autorizadas	40
3.5. Data Mart Audiencia	42
3.6. Data Mart Origen	43
3.7. Data Mart Comportamiento	43
4.1. Cuadrante Mágico de Gartner 2018	45
4.2. Api Google Analytics: Log in Gmail Analytics	49
4.3. Api Google Analytics: Otorgar permisos a Microstrategy	49
4.4. Google Analytics: Importar datos	50
4.5. Microstrategy: Vista previa	51
4.6. Visualización en Escritorio: Dashboard Audiencia	52
4.7. Geolocalización.	52

ÍNDICE DE FIGURAS

4.8. Afinidad.	54
4.9. Tipo de Dispositivo.	54
4.10. Marca y S.O.	55
4.11. Visualizacion en Escritorio: Dashboard Comportamiento - General	56
4.12. Visualización: Dashboard Comportamiento - Páginas de destino	56
4.13. Visualización: Dashboard Comportamiento - Páginas de salida	57
4.14. Visualización: Dashboard Comportamiento - Desglose de contenido	57
4.15. Visualizacion en Escritorio: Dashboard Origen	58
5.1. Análisis y resultados: Origen - Fuentes de tráfico	60
5.2. Análisis y resultados: Origen - Fuente/Medio	60
5.3. Análisis y resultados: Fuentes por Referencia	61
5.4. Análisis y resultados: Audiencia	62
5.5. Análisis y resultados: Edad	63
5.6. Análisis y resultados: Edad	63
5.7. Análisis y resultados: Sexo	64
5.8. Análisis y resultados: Intereses de los Usuarios	64
5.9. Análisis y resultados: Información geográfica	65
5.10. Análisis y resultados: Ciudades	66
5.11. Análisis y resultados: Tipos de Usuario	67
5.12. Análisis y resultados: Tipos de Usuarios	67
5.13. Análisis y resultados: Frecuencia	68
5.14. Análisis y resultados: Ciudades	68
5.15. Análisis y resultados: Navegadores	69
5.16. Análisis y resultados: Tipos de dispositivo	69
5.17. Análisis y resultados: Tipos de dispositivo	69
5.18. Análisis y resultados: Marca dispositivo	70
5.19. Análisis y resultados: Comportamiento	71

ÍNDICE DE FIGURAS

5.20. Análisis y resultados: Páginas de destino	73
5.21. Análisis y resultados: Páginas de Salida	74
5.22. Análisis y resultados: Páginas de destino	75

Introducción

En los últimos años, la Inteligencia de Negocios o Business Intelligence (BI) ha ido tomando fuerza en las organizaciones que necesitan obtener información valiosa a partir de sus datos, con el fin de mejorar su toma de decisiones. BI consiste en la creación de estrategias, conocimiento y aspectos relevantes para gestionar el negocio, a través del análisis de datos existentes.

Dentro de BI, el análisis de las páginas web de una organización permite mejorar las estrategias de negocio. Analizar archivos log, tráfico web o los click realizados, generan datos de los visitantes que son de utilidad principalmente para que el Departamento de Informática de una organización pueda mejorar el diseño web del portal del negocio, como para que el Departamento de Marketing mejore su estrategia para llegar a los clientes.

Además con el crecimiento elevado del uso de Internet, se han ampliado los mercados a través del comercio electrónico. Sumado a que más gente utiliza estos medios, generando un mayor volumen de datos que debe analizar la empresa y transformarla en información valiosa.

Por otra parte, surge el marketing digital con una fuerte inversión en publicidad online para dar a conocer un producto o servicio en una web, con el fin de llegar a la mayor cantidad de público posible.

En este ámbito se propone una memoria que permita crear un sistema de BI para analizar un portal web en búsqueda de mejorar la estrategia de marketing digital del Departamento de Informática de la Universidad Técnica Federico Santa María para su carrera de Ingeniería Civil Informática.

El presente trabajo de memoria está estructurado de la siguiente forma:

- *Capítulo 1: Definición del problema*, se describe la situación actual junto con los principales inconvenientes detectados, finalizando con los objetivos de la solución que se propone para abordar los problemas identificados.
- *Capítulo 2: Estado del Arte*, base teórica en la que se sustenta esta memoria: Inteligencia de Negocios, Analítica Web y Marketing Digital.
- *Capítulo 3: Diseño de la Solución*, se detalla la solución diseñada, a través de las herramientas utilizadas, el modelo de datos construido y los informes generados.
- *Capítulo 4: Construcción e Implementación de la Solución*, detalles de la de la solución diseñada, herramientas a utilizar, etapas de implementación.
- *Capítulo 5: Análisis de Resultados*, se valida que la solución propuesta cumple con los objetivos iniciales planteados para ella.

En la parte final se presentan las conclusiones derivadas del trabajo desarrollado, y los anexos que complementan este mismo.

Capítulo 1

Definición del Problema

1.1. Contexto

La gran cantidad de información que se genera hoy en día en la web se está transformando en un problema para las empresas. Si las organizaciones no tienen o no utilizan su información no pueden mejorar sus estrategias de negocios. Algunos problemas o pérdida de beneficios que genera esto son los siguientes:

- Incapacidad de realizar análisis de información
- No es posible el seguimiento de objetivos
- Mala gestión de datos
- Existen especulaciones

Por otro lado, el número de personas en Internet ha crecido exponencialmente. Actualmente existen 19.4 millones de conexiones en Chile [4] y 4.000 millones a nivel mundial [16], generando un mayor tráfico en la web. Es por esto, que el marketing digital se ve afectado debido a los grandes volúmenes de datos que deben procesar para generar sus respectivas estrategias. Además, todavía existen organizaciones que no tienen información relevante o no generan información importante para la elaboración de buenas campañas de marketing.

Sumado a la importancia que tiene el marketing digital¹ en los últimos años a nivel mundial, es necesaria una mayor inversión en publicidad online. Lo anterior, se complementa con la necesidad de llegar a la mayor

¹Es el uso de Internet y las redes sociales con el objetivo de mejorar la comercialización de un producto o servicio.

CAPÍTULO 1. DEFINICIÓN DEL PROBLEMA

cantidad posible de público (internacionalización). Se puede observar en la figura 1.1 la importancia de la publicidad online a la hora de comprar un producto o servicio en un portal web. Un 36 % de los que compró provienen de un aviso publicitario online.

Sin embargo, a pesar de que una marca esté bien posicionada en Internet o tenga una buena campaña de marketing digital, es necesario comprobar que dicha campaña o inversión sea la correcta. Esto se puede lograr mediante herramientas de analítica web, las que toman información generada en la web y la procesan con el fin de obtener reportes y resultados con los que se pueden tomar decisiones o detectar errores de una estrategia de marketing digital.

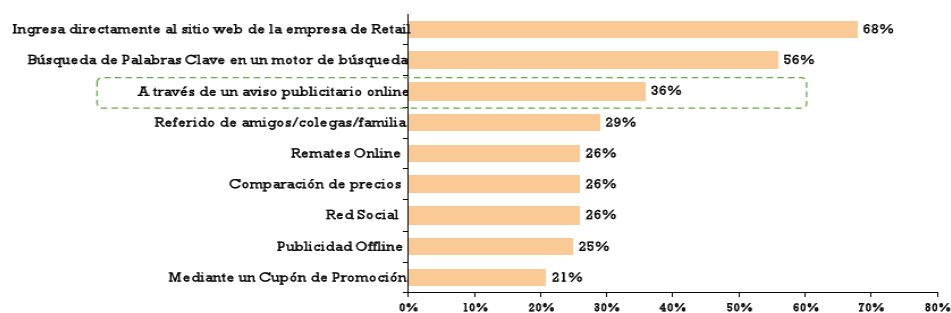


Figura 1.1: Cómo influye la publicidad online en las compras.

Fuente:[2]

Para abordar los problemas vistos anteriormente, hoy en día, existen soluciones de BI que centralizan los datos de las organizaciones y generan información relevante para la toma de decisiones. Además, es posible hacer seguimiento a objetivos, realizar análisis y conocer el comportamiento de los usuarios. Un ejemplo de esto lo podemos observar en la figura 1.2.

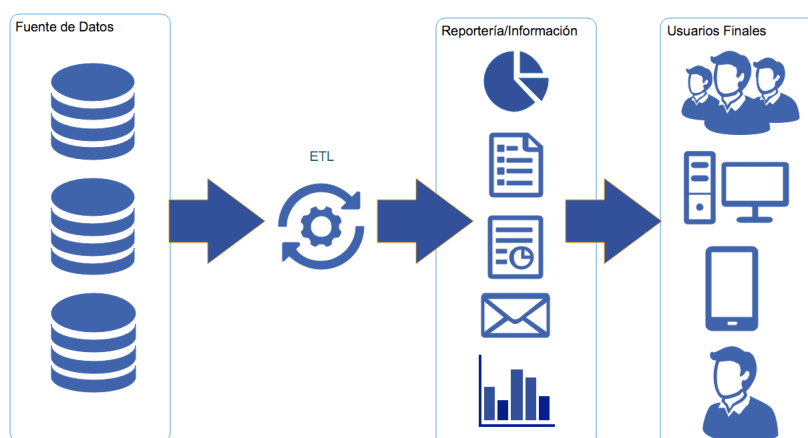


Figura 1.2: Solución de Inteligencia de Negocios

Fuente: Elaboración Propia

1.2. Situación Actual

El *Departamento de Informática (DI) de la UTFSM*², primeramente llamado Departamento de Ciencias en Computación en 1975, fue fundado en el año 1980 y en el año 2002 se extiende a la ciudad de Santiago.

El DI al año 2018, cuenta con una planta académica de 27 profesores, 16 en la Casa Central y 11 en Campus Santiago San Joaquín. Actualmente tiene alrededor de 900 estudiantes de pregrado y más de 100 estudiantes en el postgrado, sin contar a los estudiantes que participan en programas de continuidad de estudios.

En el año 2016 se crea el nuevo sitio web del DI para ajustarse a los lineamientos de la Universidad en el tema. Además, de unificar los sitios web que el DI tenía para ambos campus, junto con mejorar la navegación y experiencia de usuario, principalmente para potenciar la difusión de las carreras impartidas.

A pesar de todas estas mejoras, al día de hoy no se tiene conocimiento de las visitas que acceden al sitio web. La información está disponible, pero no se utiliza ni se realiza análisis de ella. Al no utilizar esta información, no es posible hacer un seguimiento de los objetivos (si es que se tienen), y por ende, no se pueden tomar decisiones para mejorar el sitio o atraer a mayor cantidad de visitas.

Otros problemas que presenta el DI actualmente son:

- No se tiene información de búsquedas en el sitio, por tanto no se sabe realmente qué buscan los usuarios en el sitio web.
- No existe el seguimiento de eventos, lo que permitiría saber las interacciones con elementos en Flash y AJAX o clics en enlaces de descarga.
- No puede identificar el comportamiento y preferencias de los usuarios. Por lo que no puede mejorar la experiencia de navegación.
- No conocer el origen que está aportando más usuarios al sitio, por lo tanto, no se conoce el tipo de sitio web en función de la fuente de tráfico.
- No conocer los contenidos más populares, las páginas de fuga, las preferencias de los usuarios por ubicación geográfica, las acciones dentro del sitio. Sin esto, no es posible segmentar a los usuarios.
- No se conocen las páginas de entrada con peor rendimiento o las páginas por las que más salen, por lo que no es posible optimizarlas.

²<https://www.inf.utfsm.cl/quienes-somos/presentacion-del-departamento>

CAPÍTULO 1. DEFINICIÓN DEL PROBLEMA

- No cuenta con informes o dashboard que presenten la información de forma clara y precisa. Estos, se adaptan a las necesidades de cada usuario, departamento o sección. Además, deben utilizar un lenguaje acorde al receptor para que pueda entender cada uno de los indicadores. Sin esto, no es posible tener una primera foto de cómo es el tráfico general del sitio web.

1.3. Objetivos de una Solución

1.3.1. Objetivo General

Diseñar e implementar un sistema de BI que permita analizar el tráfico de un portal web universitario, de tal manera de obtener el perfil de los visitantes para orientar la estrategia de marketing digital de una carrera universitaria.

1.3.2. Objetivos Específicos

- Identificar qué tipo de información es importante o relevante para obtener el perfil de los visitantes.
- Establecer los indicadores claves de rendimiento (KPIs) para obtener el perfil de los visitantes de la web.
- Diseñar una solución de BI en base a los indicadores anteriores.
- Construir e implementar la solución de BI para el portal del Departamento de Informática de la UTFSM.
- Realizar una validación de la propuesta, mediante el análisis de los resultados obtenidos utilizando metodología de Analítica Web.

Capítulo 2

Estado del Arte

2.1. Inteligencia de Negocios

Es un concepto que engloba todos los procesos, técnicas, metodologías, estándares y herramientas de alta tecnología, que se aplican con el objetivo de poner a disposición de los distintos niveles de una organización, información precisa y relevante con el fin de una mejor toma de decisiones.[1]

También es llamado *Business Intelligence* (BI) en inglés y algunos autores lo definen como el proceso de analizar los bienes o datos acumulados en la empresa y extraer una cierta inteligencia o conocimiento de ellos. Dentro de los bienes se incluyen bases de datos de clientes, información de la cadena de suministros, y cualquier actividad de marketing relevante para la empresa [14].

En [17] se describe a BI como un término que se relaciona con software analítico y soluciones que permiten acceder a información para tomar mejores decisiones de negocios. Además está incluido el proceso de extracción, transformación y carga de los datos (ETL por sus siglas en inglés), así como los términos almacén de datos (data warehouse), consultas de base de datos y reportes, procesamiento analítico en línea (OLAP por sus siglas en inglés) y *minería de datos* (data mining).

A partir de las diferentes definiciones de BI, se puede concluir que el objetivo es generar conocimiento desde diversas fuentes de datos. Muchas veces el volumen de datos es muy grande y al aplicarle BI, se puede generar una mejor toma de decisiones por parte de gerentes o altos mandos. Algunas características que posee BI son:

- **Accesibilidad a la información:** Los datos son la principal fuente de todo sistema de información, esta herramienta permite y garantiza que los datos estén accesibles para los usuarios, independiente de su procedencia.
- **Apoyo en la toma de decisiones:** Con la obtención de conocimientos a partir de estas herramientas, se puede apoyar de manera descriptiva o predictiva la toma de decisiones. No sólo se busca mostrar la información, sino que también generar un valor agregado.
- **Orientación al usuario final:** Provee facilidad de uso, y la independencia de las habilidades técnicas que tengan los usuarios para utilizar este tipo de herramientas, así personas de distintas áreas o especialización pueden usar la herramienta.

Una de las áreas de mayor aplicación de BI es el marketing, el que ha tenido un crecimiento destacado debido a las tendencias a nivel mundial. Más adelante se abordará el tema del marketing digital con mayor profundidad.

A continuación algunos conceptos con sus definiciones que permitirán un mayor entendimiento del proceso de diseño e implementación de un sistema de BI.

2.1.1. Componentes de BI

2.1.1.1. ETL

Componente BI que incluye los procesos de extracción (Extract), transformación (Transform) y carga (Load), los cuales permiten migrar los datos desde múltiples fuentes, reformatearlos y limpiarlos, y cargarlos en un data warehouse (DW) o data mart (DM) para apoyar un proceso de negocio.

Extracción: Es la primera parte del proceso de ETL y consiste en extraer los datos desde su origen. Las fuentes de origen son principalmente bases de datos relacionales, ficheros planos y en algunas ocasiones bases de datos no relacionales. Un requerimiento importante en este proceso es que la tarea de extracción cause el menor impacto posible en el sistema de origen. Debido a que generalmente las organizaciones

CAPÍTULO 2. ESTADO DEL ARTE

disponen de horarios en los cuales el sistema no puede ralentizarse.

Transformación: Es la segunda fase del ETL y en ella se aplica una serie de reglas de negocio sobre los datos extraídos para convertirlos en los datos que posteriormente serán cargados. En este proceso se suelen aplicar las siguientes transformaciones: No cargar columnas con valores nulos, transformar datos de la fuente de origen a otro tipo de datos para la fuente de destino, codificar datos, diferenciar entre distintas abreviaciones que significan lo mismo. Todo lo relacionado con detección de anomalías, transformación de datos o validaciones se puede observar en esta fase.

Carga: La última fase consiste en la carga de los datos ya procesados de la etapa anterior hacia su destino. Puede ser que se sobrescriban bases de datos anteriores con los nuevos datos, pero en un DW generalmente se guarda un historial con los cambios de los registros para así disponer los distintos estados a lo largo del tiempo de un dato.

2.1.1.2. Data Warehouse

Es un repositorio central o colección de datos en el cual se encuentra integrada la información de la organización utilizada como soporte para el proceso de toma de decisiones gerenciales.

Otra definición corresponde al lugar físico donde se almacena la información que genera la empresa. Esta información puede provenir de distintos sistemas utilizados en una organización para sus actividades diarias, es decir, los sistemas transaccionales de distintas áreas de la empresa son los encargados de proveer la información al DW. Algunas funcionalidades importantes de los DW son:

- **Integración:** De los datos provenientes de bases de datos de las diferentes unidades de la organización y que con frecuencia tienen diferentes formatos.
- **Separación:** De los datos usados en operaciones diarias de aquellos usados en el Data Warehouse para propósitos de divulgación, de ayuda en la toma de decisiones, análisis y operaciones de control. No deben estar en la misma base de datos ya que obedecen a distintos objetivos.

2.1.1.3. Data Mart

Es un subconjunto de un DW que tiene por objetivo apoyar a un área específica dentro de la organización para mejorar su toma de decisiones. Generalmente los DM son consultados mediante una herramienta

OLAP.

2.1.1.4. OLAP

Procesamiento analítico en línea o en inglés Online Analytical Processing. Es una solución utilizada en BI para agilizar la consulta de grandes cantidades de datos, a través de herramientas que generan cubos o estructuras multidimensionales, que permiten mostrar información del negocio de una forma intuitiva y ordenada, lo que facilita el monitoreo del desempeño del negocio en el tiempo.

Los sistemas que utilizan OLAP permiten un acceso rápido a la información. Además, esta tecnología permite responder a preguntas del tipo ¿qué? y ¿quién?. También responder preguntas como ¿por qué? y ¿qué pasa si?, o sea, especular con escenarios futuros usando datos históricos.

2.1.1.5. Dashboards

Son herramientas que tienen como objetivo mostrar la información resumida de la organización de manera gráfica y en tiempo real.

En la figura 2.1 se puede observar el modelo que integra BI con los procesos, etapas y aplicaciones involucradas. Se nota a simple vista que los datos parten de sistemas transaccionales, pasan por un proceso de ETL y posteriormente son llevados a un DW o DM para que sean utilizados por un cubo OLAP y finalmente se generen reportes o se presenten a través de un dashboard.

2.1.2. Aplicaciones de BI

En general, las soluciones basadas en BI están presentes en áreas como:

- **Ventas:** Análisis de ventas, análisis de productos, mercados, pronósticos y proyecciones.
- **Marketing:** Segmentación y análisis de clientes, seguimiento de nuevos productos. Entrega de mensajes más exactos a la audiencia. Como ejemplo se puede considerar el análisis de las encuestas para identificar comentarios más comunes, solicitudes más frecuentes y buscar perfiles en cada caso. Otro ejemplo es el estudio de foros, cartas, mensajes electrónicos.
- **Finanzas:** Análisis de gastos, rotación de cartera y tasas financieras.

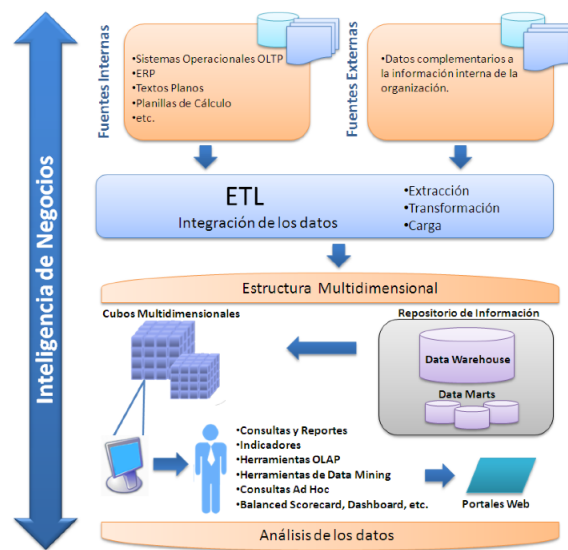


Figura 2.1: Modelo BI con sus proceso, etapas y aplicaciones. Fuente:[1, pag 21]

- **Manufactura:** Productividad en líneas, análisis de desperdicios, análisis de calidad, rotación de inventarios y productos críticos.
- **Banca:** Análisis de la cartera de clientes, identificación de perfiles de clientes para otorgamiento de créditos.
- **Salud y Medicina:** Asignación de tratamientos médicos en base a patrones encontrados. Prevención de futuras enfermedades. Mejoramiento en diagnósticos médicos.

2.2. Análisis Web

Consiste en el conjunto de técnicas relacionadas con el análisis de datos relativo al tráfico en un sitio web. Existen dos categorías de análisis web: *off-site* y *on-site*, la primera hace referencia a mediciones web y análisis independiente de si es el dueño del sitio web o es la persona encargada del mantenimiento. Aquí se incluye la medición de la audiencia potencial, participación de voz y el denominado murmullo o comentarios acerca de lo que está pasando en Internet como un todo.

El análisis on-line mide la ruta de un visitante una vez que ya ha ingresado al sitio web de propiedad de la persona que realiza el análisis. Estas mediciones comparan indicadores de desempeño clave y se usan para mejorar un sitio web o para ver la respuesta de la audiencia o visitante frente a una campaña de marketing.

CAPÍTULO 2. ESTADO DEL ARTE

Existen dos técnicas principales para recolectar datos:

Analizadores de ficheros log del servidor: Consiste en la lectura de los archivos log que un servidor guarda de las navegaciones que se realizan a través de un sitio web. Sus principales ventajas son:

- Los servidores siempre producen ficheros de logs, por lo cual la información siempre está disponible.
- Los servidores capturan la totalidad de los accesos al sitio web.
- La información reside en los propios servidores y tiene un formato estándar.
- Los ficheros de logs almacenan información sobre peticiones fallidas, en otras técnicas, ésta se pierde.

Etiquetado de páginas: Es un método más reciente que el anterior y se basa en la incorporación de un script a cada una de las páginas de un sitio. Cada vez que se visita una página, el script envía a una base de datos cierta información relevante del sitio e información del visitante obtenida de las cookies del sitio. Sus ventajas son:

- Puede capturarse información no disponible en los ficheros de log, incluso, modificarla sin más que cambiar el script.
- Puede realizarse a pesar de que los dueños del sitio no tengan acceso por estar alojados en servidores ajenos a los logs del servidor.

2.2.1. Análisis web en empresas

El análisis web comenzó a tomar un papel importante en organizaciones que tienen presencia en Internet. No sólo los departamentos de sistema y diseño web están interesados en obtener información acerca del tráfico web, sino que también, un departamento muy importante hoy en día: el departamento de marketing.

En los últimos años, el uso y crecimiento de Internet ha sido notorio, muchos usuarios domésticos pasaron a utilizar en mayor medida los sitios web de las empresas, ya sea para compras, obtener información o muchas veces presentar un reclamo. Debido a lo anterior nacen las llamadas herramientas de análisis web, que permiten a los altos mandos o gerentes tomar decisiones de marketing de mejor manera y además mejorar su estrategia en el posicionamiento de Internet, lo que lleva a un potencial incremento de clientes y la necesidad de optimizar el sitio web. Así, que es posible conocer de mejor forma a los visitantes o clientes

CAPÍTULO 2. ESTADO DEL ARTE

del sitio, lo que permite segmentar el público objetivo y mejorar las estrategias.

Además es posible determinar oportunidades para el negocio, comprender de mejor forma los llamados indicadores clave de rendimiento (KPI), averiguar los productos más rentables, o saber qué es necesario potenciar del sitio web. Así, es posible hacer un mejor seguimiento de clientes, averiguar la procedencia y destino de los visitantes, aumentar la fidelidad de ellos, en definitiva, gestionar todo lo relacionado con el negocio en Internet.

También es aplicable a organizaciones sin fines de lucro, mejorando el servicio entregado, el diseño web para hacerlo más atractivo y fácil de navegar, posicionamiento en buscadores y maximizando el retorno de la inversión en publicidad.

2.2.2. Google Analytics

Google Analytics [8] es una herramienta de analítica web de la empresa Google. Ofrece información agrupada del tráfico que llega a los sitios web según la audiencia, la adquisición, el comportamiento y las conversiones (Acciones que realiza un usuario para el cumplimiento de objetivos. Ej: descargar PDF, llenar un formulario, etc.) que se llevan a cabo en el sitio web.

Se pueden obtener informes como el seguimiento de usuarios exclusivos, el rendimiento del segmento de usuarios, los resultados de las diferentes campañas de marketing online, las sesiones por fuentes de tráfico, tasas de rebote, duración de las sesiones, contenidos visitados, conversiones, etc. Google Analytics se desarrolló basándose en la compra de Urchin (hasta entonces la mayor compañía de análisis estadístico de páginas web) por parte de Google.

El modo de funcionamiento de esta herramienta, se basa en incorporar un código JavaScript a cada una de las páginas que se desee analizar, este código recibe el nombre de Google Analytics Tracking Code o GATC, quien es el encargado de monitorizar todo el tráfico de las páginas donde se ha implementado y enviarlo a los respectivos servidores de Google para almacenarla en las cuentas de usuario que corresponda.

Existen diferentes funcionalidades de análisis web:

2.2.2.1. Analítica de Contenido

Descubrir con qué frecuencia visitan las personas cada una de las páginas del sitio web, durante cuánto tiempo y con qué frecuencia realizan conversiones.

CAPÍTULO 2. ESTADO DEL ARTE

- **Búsqueda en el sitio:** Permite descubrir realmente qué buscan los usuarios mediante el análisis de información sobre las palabras más buscadas en el sitio web.
- **Analítica de página:** Se puede analizar visualmente cómo se mueven los visitantes en el sitio.
- **Análisis de velocidad del sitio:** Permite ver cuánto tardan las páginas en cargarse, lo cual es información relevante para corregir aquellas páginas lentas y dar un mejor servicio al visitante.
- **Seguimiento de eventos:** Permite saber qué hacen los visitantes en el sitio. Por ejemplo, interacciones con elementos en Flash y AJAX o clics en enlaces de descarga de folletos; incluso configurar el seguimiento de eventos para saber más detalles del comportamiento de los visitantes.

2.2.2.2. Analítica de Móviles

Analizar anuncios, aplicaciones y sitios para móviles, y descubrir cómo se puede obtener ingresos de ellos, es cada día más importante. Google Analytics permite evaluar los sitios web y las aplicaciones para móviles, así como las visitas desde dispositivos móviles con conexión a Internet, incluidos los tablets y los teléfonos, tanto de gama alta como básicos.

- **Analytics para aplicaciones móviles:** Entrega la cantidad de usuarios que ejecutan la aplicación móvil, mostrando desde qué tipo de dispositivo y/o sistema operativo se realiza la ejecución.
- **Informes de interacción:** Permite saber cómo interactúan los usuarios con la aplicación. Además, obtener fidelización, frecuencia y el nivel de interacción.

2.2.2.3. Analítica de Conversiones

Se puede observar cómo funcionan todos sus canales de marketing digital en conjunto para atraer visitantes al sitio. Medir las ventas, las descargas, las reproducciones de vídeo y otras acciones para el cliente. Saber por qué algunos visitantes compran productos y otros no, para poder ajustar el sitio y programas de marketing a fin de alcanzar los objetivos de la empresa.

Para el caso del comercio electrónico permite identificar los productos más vendidos, promociones más valiosas, motivos de compra y tipos de compra de los clientes.

2.2.2.4. Analítica Social

Los informes sociales ayudan a medir el impacto de las redes sociales en los objetivos de la organización y en las conversiones relacionadas.

CAPÍTULO 2. ESTADO DEL ARTE

- **Conversiones:** Se obtiene el porcentaje de conversiones de los visitantes que provienen desde una red social.
- **Fuentes:** Permite saber desde qué red social provienen los visitantes hasta llegar al sitio web.
- **Compartir:** Observar qué publicaciones o artículos del sitio son más compartidos, los botones sociales más usados (+1 de Google o Me gusta de Facebook).

2.2.2.5. Analítica de Anuncios

Comprobar todos los canales digitales, incluidos los de búsqueda, los de display, los de redes sociales, los de afiliación y los de correo electrónico, y ver cómo afectan a los porcentajes de conversiones y de retorno.

- **Marketing en motores de búsqueda:** El marketing en motores de búsqueda atrae visitantes. Se puede enlazar el comportamiento de los visitantes con las campañas de anuncios pagados (SEM del inglés Search Engine Marketing) y tomar decisiones sobre ofertas y palabras claves.
- **Remarketing:** Es una forma eficaz de ampliar la cobertura y acercar a los clientes a la compra. Un cliente al agregar un producto al carrito de compras o visitantes que dedican un tiempo importante a una página concreta, son acciones precisas que permiten identificar un público objetivo y publicar anuncios adaptados a dicho público.
- **Medición de campañas:** Evaluar las campañas en todos sus canales, para así, ver qué actividades de marketing son rentables. Realizar informes segmentados por canal, datos en tiempo real a minutos de iniciada una campaña. Observar la interacción y conversión de los visitantes.

2.2.3. Archivos Log de Servidores Web

Es un archivo de log o en inglés *server log* que contiene un historial con las solicitudes de un sitio web. Existe un formato estándar para los archivos log de servidores web. La información que se puede obtener de las solicitudes que recibe un sitio web son:

- Dirección IP del visitante
- Día/tiempo de la solicitud
- Página solicitada
- Código HTTP

CAPÍTULO 2. ESTADO DEL ARTE

Usualmente estos archivos no son accesibles por los usuarios de Internet, sólo por los *webmaster* o administradores del sitio web. El departamento de marketing de cualquier organización que posee un sitio web debería tener un entendimiento de lo que podría hacerse con esta valiosa información.

2.2.4. *Clickstream*

Es el seguimiento de los click que realiza un usuario sobre las distintas partes de la pantalla de un computador mientras navega en un sitio web, a través de un navegador o mientras usa algún software de seguimiento. Mientras el usuario hace click en cualquier parte del sitio web o aplicación, la acción es archivada o guardada en un servidor web. Análisis de *Clickstream* es usado para actividades de análisis web, testeo de programas, investigación de mercado y para analizar la productividad del empleado.

2.2.5. *Metodología para Análisis Web*

Una de las metodologías más populares de la Analítica Web es la propuesta por [5]. En primer lugar se analizan las métricas básicas de los sitios web, posteriormente se deben definir los objetivos y los principales KPI's asociados a éstos.

En la segunda parte se comienza a realizar el análisis, el cual consiste en revisar la navegación de las personas en el sitio web, seguido de las conversiones, y de un análisis cualitativo sobre lo que pasó y por qué puede haber pasado.

Finalmente se termina con la elaboración de informes, los que permiten obtener conclusiones para la toma de decisiones o pasos a seguir.

2.2.5.1. *Métricas Básicas*

Las estadísticas de tráfico en la web son la primera y más importante fuente de información a la hora de medir el rendimiento de ésta. Estas métricas permiten interpretar cuánto tráfico se tiene en el sitio web, qué páginas son preferidas por los usuarios, cuánto tiempo se quedan en el sitio, si realmente es lo que están buscando y cuántos usuarios son fieles (vuelven). En resumen, ayudan a entender cómo navegan los usuarios para así poder ofrecerles lo que realmente están buscando.

- **Páginas vistas:** Cada vez que se descarga una página se cuenta como una página vista. Con esta métrica se puede determinar cuántas páginas distintas se han descargado en un período de tiempo, y por otro lado, saber cuántas descargas ha tenido una página determinada.
- **Visitas:** Sesiones que se abren del sitio web durante un período de tiempo.

CAPÍTULO 2. ESTADO DEL ARTE

- **Visitantes únicos:** Son los distintos dispositivos que entran a la página. Si una persona se conecta al sitio web desde 3 dispositivos distintos, se contabilizarán 3 usuarios únicos.
- **Tiempo de estancia en una página:** Diferencia de tiempo en que el usuario deja una página y el tiempo en que llegó a ésta. Si la persona se va del sitio web, no es posible calcular el tiempo de estancia de esa página.
- **Tiempo de estancia en el sitio:** Es la suma de los tiempos de estancia en cada página.
- **Tasa de rebote:** Es el porcentaje de visitas que, o bien estuvieron un par de segundos en la página, o solamente vieron una página.
- **Tasa de salida:** Es el número de visitas que han salido por una determinada página, independiente de la cantidad de páginas que vió.

2.2.5.2. Conversiones, objetivos y KPI

El primer paso para cumplir los objetivos es medir. Los objetivos se relacionan con las conversiones que se producen en el sitio web. Lo principal es que el usuario cumpla con los objetivos y no sólo ingrese a la página.

Tasa de Conversión:

El objetivo estará determinado por lo que se desea haga el usuario en el sitio web. Convertirse en cliente, contestar una acción, comprar un producto, contratar un servicio, llegar a un determinado punto del sitio web. La tasa de conversión es el mejor KPI para medir el rendimiento de una web y se calcula como:

$$Tasa\ de\ conversion = \frac{Objetivos\ conseguidos}{Numero\ total\ de\ visitas}$$

Se recomienda medir la conversión de cada proceso en particular:

$$Tasa\ de\ conversion\ proceso = \frac{Objetivos\ conseguidos\ proceso}{Numero\ total\ de\ visitas}$$

Con estos KPI se puede dar respuesta a preguntas como:

¿Qué proceso es el que mejor está funcionando? ¿De cuál se puede aprender? ¿Los procesos cortos funcionan mejor que los largos? ¿Qué tienen en común los que poseen tasa de conversión más baja? ¿Y los que tienen la más alta?

Objetivos del sitio web:

Los objetivos deben ser específicos, realistas y sobre todo se deben poder medir. Estos objetivos se deben definir junto con los responsables del negocio. Las preguntas más comunes en este proceso son:

CAPÍTULO 2. ESTADO DEL ARTE

- ¿Por qué existe la web?
- ¿Qué aporta al negocio?
- ¿Qué se pretende obtener y cómo se puede lograr?

Posteriormente, estos objetivos se deben convertir en métricas. La táctica que permite hallar la métrica que mejor define el objetivo es pensar en la acción que debe realizar el visitante. Para ayudar se realizan las siguientes preguntas enfocadas en el usuario:

- ¿Debe rellenar un formulario?
- ¿Llamar a algún teléfono?
- ¿Descargar un PDF?
- ¿Terminar de ver un video?

Los objetivos pueden focalizarse en el propio negocio o en las necesidades del cliente. Los que se refieren al negocio se clasifican en dos grupos: bajar los costos o incrementar las ganancias.

Algunos KPI's señalados en [3] para sitios web enfocados en el negocio son los siguientes:

Ecommerce

- Ratio de conversión de compradores nuevos y compradores repetidores
- Conversiones por campañas
- Porcentaje de pedidos respecto a los que iniciaron el proceso de compra
- Nivel de atracción de la página inicio o landing page
- Porcentaje de errores

Turismo

- Promedio de visitantes que realizan reservas
- Porcentaje de visitas que necesitaron asistencia
- Promedio de usuarios que se registran
- Porcentaje de ingresos provenientes de ofertas

CAPÍTULO 2. ESTADO DEL ARTE

- Número de búsquedas por visita

Banca y Seguros

- Número de visitas que contratan algún servicio
- Porcentaje de accesos por cliente
- Porcentaje de altas de usuarios
- Interés productos
- Efectividad de las campañas internas

Los objetivos que se centran en el cliente buscan conseguir un canal eficiente para los visitantes, además, proveer de un alto nivel de personalización y relevancia a la vez. Algunos ejemplos típicos son los siguientes:

Redes Sociales

- Menciones de la marca contra menciones de la competencia
- Distribución de contenido
- Participación frente a visualizaciones
- Índice de satisfacción
- Ratio de sentimiento
- Relación visitantes de Redes Sociales (SM) respecto a visitantes del sitio

Publicaciones

- Frecuencia de las visitas
- Tasa de rebote
- Efectividad del buscador
- Media de subscripciones RSS (Really Simple Syndication, formato en XML para distribuir contenido en la web)
- Ratio de visitantes nuevos frente a repetidores

CAPÍTULO 2. ESTADO DEL ARTE

Cuando los objetivos de la web están claros, y los de cada departamento, es necesaria la generación de KPI's, o sea, definir las métricas necesarias para saber qué tan cerca o lejos se está de alcanzar los objetivos, estableciendo un plazo de control sobre los mismos.

Selección de KPI:

Cada sitio web tiene su objetivo general, cada proceso tiene su objetivo particular. Es necesario determinar qué métricas son las que mejor se adaptan al negocio y cuáles pueden representar mejor los objetivos del sitio. Según el propósito del sitio web, éstas son:

Sitio web de tipo comercial: Aquí el objetivo es vender, gastando lo mínimo posible, por lo que las métricas serían:

- Tasa de conversión
- Media de venta (ingresos/conversiones)
- Valor de la visita (ingresos/visitas)
- Fidelidad (visitantes que repiten / nuevos)
- Retención (> 2 páginas / visitante)
- Visitas desde buscadores (visitas buscadores / visitas)

Sitio web de soporte o servicios: El objetivo será dar respuesta inmediata al cliente:

- Tasa de conversión (conversiones/visitas)
- Tasa de rebote (visitas menores a 30 segs.)
- Contenidos vistos (cantidad de contenidos vistos / visitas)
- Top palabras claves (¿qué buscan?)

Sitio web de captación (Lead Generation): El objetivo será incrementar la red de contactos:

- Tasa de conversión (leads/visitantes): Cuántos leads se registran o consiguen
- Costo por lead (leads/inversión en marketing): Costo de cada lead
- Porcentaje de nuevas visitas (nuevos visitantes / visitantes únicos): ¿Cuántos son potenciales?
- Visitantes que vieron 1 página (visitantes 1 página / total páginas): ¿Cuántos se van sin ir más allá?

CAPÍTULO 2. ESTADO DEL ARTE

Sitio web Corporativo: Por un lado incrementar la fidelidad del lector (web de contenido), por otro, conseguir que la marca genere interés:

- Tasa de conversión (visitas a la página de contacto o registros en newsletter/total de visitas): ¿Cuántos convierten?
- Profundidad de la visita (páginas vistas / visita): Generar interés
- Visitantes que repiten (nº visitantes que repiten / visitantes únicos): ¿Son fieles?
- Visitas cortas (visitas < 1 minuto / total visitas): ¿Se quedan con el concepto?
- Adquisición de tráfico (visitas de tráfico directo + visitas de buscadores con palabra clave de la marca / Total de visitas): Visitas que llegan conociendo la marca o que han oído hablar de ella

Marco de trabajo Objetivos + KPI:

La metodología sugerida en [5] para identificar las acciones asociadas a cada objetivo y su medición consisten en:

- Identificar los objetivos del negocio.
- Identificar los objetivos de la web, los cuales tratarán de satisfacer los objetivos del negocio.
- Seleccionar los KPI's que mejor representan los objetivos del negocio y del sitio.

2.2.5.3. Análisis Básicos: Web y Contenido

Consiste en un primer acercamiento, entender bien quién visita la web, desde dónde, qué está buscando, qué contenido consume, cuándo y por dónde se va.

Entendiendo cómo las visitas ven el sitio web se puede optimizar el contenido para lograr los objetivos planteados. Algunos elementos a analizar son los siguientes:

- **Origen y salida de las visitas:** Al conocer bien las fuentes de origen de las visitas y analizando las palabras claves (keywords) se pueden ir mejorando las estrategias de captación. Analizando por donde salen se pueden optimizar las páginas que tienen problemas.
- **Perfiles de navegación:** Conociendo a las visitas por el contenido que consumen y su comportamiento, a largo plazo, se podría aprender de la experiencia y anticiparse o adelantarse a las intenciones del visitante. Establecer un patrón sobre por dónde entran las visitas, qué contenido les interesa más, por dónde se mueven en el sitio y por dónde se van, permitirá mejorar el sitio donde haga falta.

CAPÍTULO 2. ESTADO DEL ARTE

- **Captación y fidelidad de las visitas:** Saber si se atrae a las visitas en el mismo grado en que se fidelizan, ayudará a saber si se están haciendo bien las cosas.

2.2.5.4. Origen y salida de las visitas

Es necesario realizar un análisis de las fuentes de entrada, para ello se definen 4 fuentes de origen de tráfico en un sitio web:

- **Buscadores:** Todo el tráfico proveniente de los distintos buscadores (orgánico o pagado)
- **Referrals:** Enlaces en otros sitios web que traen tráfico al sitio.
- **Campañas:** Tráfico de todas las campañas externas al sitio web y que han sido etiquetadas correctamente.
- **Tráfico Directo:** El resto, visitantes que escribieron la URL del sitio en el navegador o tienen el sitio web en favoritos.

El análisis de las visitas se debe realizar en función a los porcentajes que cada fuente tiene sobre el total:

- **Alto porcentaje de tráfico directo:** Empresa o marca con una alta trayectoria, muy asentada en Internet, con un público muy fidelizado, marca muy reconocida.
- **Fuerte inversión publicitaria:** Si el porcentaje de tráfico es alto, quiere decir o bien que se ha gastado mucho dinero en poner campañas en todos lados o bien se ha dirigido a un público bien segmentado, o se trata de una campaña de emails a usuarios muy fieles que han respondido en masa.
- **Empresa/marca referentes en su sector:** Si se tienen muchas visitas desde otros sitios web, es o bien porque hay muchos sitios que enlazan al sitio o porque el sitio es enlazado desde webs del mismo sector. También está la opción que la página llame la atención por su calidad.
- **Tráfico orgánico bajo:** Este tráfico puede ser bajo por alguno de los siguientes motivos: el contenido de la web no está bien indexado, no se conoce demasiado la marca o es muy nueva, el producto o servicio tiene mucha competencia y no está bien posicionado, la web está muy abandonada y sus contenidos no están actualizados.
- **Tráfico orgánico alto:** Este tráfico puede ser alto por alguno de los siguientes motivos: se tiene un experto en SEO posicionando la web la URL es difícil de escribir o recordar y los usuarios utilizan el buscador para llegar al sitio, se tienen contenidos actualizados con mucha frecuencia.

CAPÍTULO 2. ESTADO DEL ARTE

Los porcentajes indicarán el tipo de sitio web que se tiene, pero lo importante es saber dónde se puede mejorar. No necesariamente hay que fijarse en el volumen de las visitas. Hay que analizar distintas métricas como: tasa de rebote, tiempo de estancia y las páginas vistas por visita.

Otra parte del estudio de una web por el cual se puede conocer a los usuarios, es el análisis de páginas de entrada en el sitio web. Esto consiste en estudiar las páginas más populares por las que ellos entran, así se podrán hacer cambios a estas páginas con el fin de mejorar el rendimiento total de la web.

La página por la que entra la visita desde la fuente de origen se denomina página de aterrizaje o landing page. Primero se debe realizar un análisis de la conducta de los visitantes en función al origen de la visita. Se debe identificar cada página de entrada, comenzando por la Home y las relativas a SEO, SEM, email Marketing, etc.

Identificar las páginas por las que más entran los usuarios, incluyendo además del número total de visitas, la tasa de rebote y el total de páginas vistas por visita, podría generar una idea sobre la aceptación de cada punto o fuente de entrada, el tráfico que atrae, si se está mostrando lo que el usuario quiere o espera ver (alta tasa de rebote), y si les gusta el contenido del sitio web (profundidad de la visita o páginas vistas por visita).

Estas páginas son la primera impresión, por lo tanto son importantes de analizar. Es importante conocer de dónde viene el tráfico para definir una buena estrategia, tanto para SEO como para SEM. Para una web en general, es mucho mejor tener una página para cada contenido, en vez de una página que los aglutine todos (excepto la Home). Para evaluar el rendimiento de las páginas de entrada, se debe medir su capacidad de persuadir a los visitantes a continuar hacia otras páginas en su camino a la conversión. En [5] se recomienda crear una tabla que contenga por cada página de entrada lo siguiente:

- Visitantes que entran por la página X / total visitantes del sitio web
- Páginas vistas por visitante que han entrado por la página X
- Tiempo medio pasado en la página X

El combinar métricas ayuda a entender el rendimiento y a identificar acciones de mejora de éste. Por otra parte, la segmentación de los orígenes del tráfico de cada página de aterrizaje ayuda a saber cuáles están funcionando bien y cuáles no, compararlas entre sí, y ver dónde se puede mejorar.

Se puede modificar el contenido de la tabla anterior por lo siguiente:

- Visitantes que entran desde el segmento X / total visitantes del sitio web
- Tiempo medio pasado en todas las páginas del segmento X

CAPÍTULO 2. ESTADO DEL ARTE

- Media de páginas vistas por visitantes que han entrado desde el segmento X

Este análisis permite saber el antes y durante de los visitantes por cada página de entrada. Para medir el después, las métricas que se pueden ocupar son las siguientes:

- Visitantes totales de cada una de las fuentes del segmento X
- Páginas vistas de visitantes que han entrado desde el segmento X
- Tasa de rebote por cada página de entrada

Con este análisis es más fácil saber qué porcentaje de visitantes provienen de una fuente determinada de tráfico, cómo funciona la página de entrada, de dónde provienen las visitas que terminan convirtiendo (cumpliendo objetivos), y por tanto, podemos tomar medidas para mejorar.

También se puede tomar el top 10 de URL de origen y combinarse con la tasa de conversión (cociente entre el número total de cumplimiento de objetivos y el número total de visitas durante un determinado periodo de tiempo). Se podría saber si la cantidad de visitas se corresponde con visitas rentable, y así, determinar nuevas estrategias basadas en la fuente que está mandando más visitas o cuál tiene la mejor tasa de conversión. Además, saber dónde invertir en un momento dado y dónde puede estar una oportunidad de negocio.

Las keywords son una parte importante del análisis de entrada de las visitas al sitio. Es la mayor pista de lo que se está buscando. Con esto se puede optimizar una posible inversión en SEM o SEO. Se podría combinar la palabra clave con métricas como el total de visitas del sitio web, páginas vistas (¿se han interesado?), promedio de tiempo en el sitio (¿les gusta el contenido?), porcentaje de visitas nuevas y porcentaje de rebote (¿no les gusta?).

Para **optimizar las páginas de entrada**, es conveniente medir su rendimiento teniendo en cuenta tres etapas:

1. Analizando la tasa de rebote de las páginas
2. Usuarios que ven la página de entrada y siguen navegando
3. Usuarios que ven la página de entrada, navegan y llegan a convertir

Con esto se puede tener una visión general del éxito o no de la landing page. No sirve de mucho atraer tráfico al sitio web si no se genera interés en el visitante. La métrica que mejor identifica si el usuario se interesa o no es la tasa de rebote.

CAPÍTULO 2. ESTADO DEL ARTE

Como primer paso se debería actuar en la página de entrada que tiene muchas visitas pero tiene una tasa de rebote alta. También, en la palabra clave que atrae muchas visitas, pero tiene una tasa de rebote muy alta. O, en la campaña que atrae muchas visitas, pero tiene una tasa de rebote muy elevada. No importaría lo que se está analizando, si por un lado trae una alta cantidad de visitas, pero por otro no tiene calidad ya que no avanzan en el sitio.

Otro estudio importante de realizar, es el **análisis de la tasa de rebote de las páginas de entrada**. Se puede realizar en base a las páginas de entrada o palabras claves de entrada al sitio. Además, esto combinarlo con métricas como: salidas, número de visitas y tasa de rebote.

Con estas variables se puede ordenar cada keyword según su importancia a la hora de atraer tráfico y a la hora de interesar y tomar decisiones del tipo: Esta palabra clave es lo suficientemente importante porque atrae una cantidad de tráfico por encima de la media, pero no se logra interesar a los usuarios una vez que llegan al sitio web. ¿Hay que modificar o mejorar la página de entrada?.

Posiblemente aquí es donde se detectarán las páginas de entrada genéricas, aquellas que no son personalizadas para el término de la búsqueda. Para maximizar la tasa de conversión se debe encontrar el equilibrio entre la respuesta inicial que se produce cuando el usuario llega al sitio web de entrada y la conversión final (objetivo).

Una vez detectado si el primer vistazo ha interesado o no a los usuarios, se debe centrar en las páginas de entrada con problemas para aplicarles un test A/B (experimento aleatorio en dónde las variables son A y B, siendo una la variable de control y la otra la variante) o incluso una reestructuración completa.

Un análisis más detallado sería combinar las variables de entrada, tasas de rebote, páginas vistas, ingresos obtenidos, objetivos conseguidos, tasa de conversión para cada una de las páginas de entrada. Así centrarse en las landing pages que no están funcionando bien.

Finalmente, cabe mencionar que un análisis de salida de las visitas, que implica investigar las páginas por dónde salen los usuarios, es tan importante como el conocer por dónde entra. Sabiendo la última página vista por la visita se puede recopilar información sobre las páginas que gatillan la salida del usuario y así trabajar en la retención de los visitantes o la tasa de conversión.

2.2.5.5. Perfiles de Navegación

Las páginas web proveen de un variado contenido, por lo que es necesario hacerse la siguiente pregunta: ¿Qué contenido interesa a nuestros visitantes?

Un análisis de las páginas más vistas permite conocer las páginas que tienen mayor número de visitas en un periodo de tiempo determinado. Al conocer los contenidos más visitados, por dónde entran al sitio web, los puntos de origen de las visitas y las palabras claves por las que han buscado el sitio, se tendrá una primera idea de lo que está pasando en el sitio web, se sabrá como el cliente percibe la página, pudiéndose así comparar el objetivo planteado inicialmente v/s lo que realmente está pasando.

Una vez analizado el origen y salida de las visitas, es necesario entender bien el ciclo de vida del contenido de la web, ya que esto es vital para optimizar el sitio.

El primer paso sería dividir el sitio web en secciones, establecer objetivos para cada uno y hacerse preguntas del tipo:

- ¿Cómo se comportan los visitantes de cada una de las secciones?
- ¿Por dónde vienen?
- ¿Desde qué fuente?
- ¿Hasta dónde llegan?
- ¿Qué relación hay entre las secciones?
- ¿Qué es lo que aporta esta sección a los objetivos globales del sitio web? ¿Funcionan las landing page de cada sección?

Dentro de cada sección se debe analizar la información de cada página para un periodo determinado (día, semana, mes, trimestre, etc.), así se conocerá la popularidad de cada contenido.

Las métricas más usadas o que mejor se adaptan a este tipo de análisis son:

- Cuántos (visitas/usuarios únicos)
- Cuántas (páginas vistas)

La clave es optimizar el contenido en base al análisis web. Se puede segmentar a las visitas en base a qué les gusta a los que repiten la visita, qué contenido atrae a los nuevos o qué diferencias de consumo de contenido hay entre las distintas fuentes de origen.

2.2.5.6. Captación, Fidelización y Asiduidad

Lo primero es entender bien el concepto de nuevos visitantes. Este consiste en aquellos visitantes que no tienen la cookie identificadora del sitio web, o sea, una persona que llega por primera vez al sitio.

Algunos KPI que miden el tipo de relación que tienen los visitantes con el sitio web son:

$$\%Nuevos\ Visitantes = \frac{Nuevos\ Visitantes \times 100}{Total\ Visitantes}$$

$$\%Visitantes\ que\ repiten = \frac{Visitantes\ que\ repiten \times 100}{Total\ visitantes}$$

Para una estrategia de marketing que se centra en la captación masiva de clientes, el porcentaje de nuevos visitantes deberá ser más alto que el de los que repiten. Por el contrario, si la estrategia se centra en la retención de usuarios (sitio web tipo blog), el porcentaje de visitantes que repiten deberá ser mayor al porcentaje de nuevos visitantes.

En [5] ocupan una fórmula o ratio que representa a los nuevos visitantes v/s los que repiten:

$$Ratio = \frac{Nuevos\ visitantes}{Visitantes\ que\ repiten}$$

Independiente de la estrategia a seguir, el ratio descrito anteriormente debe mantenerse más o menos estable, si no es así, se debe investigar el por qué de la baja de uno o la subida del otro. Además, identificar el KPI que ha provocado el cambio en el porcentaje.

Según Peterson [12] un ratio medio por tipo de sitio web sería:

- Sitio web basado en contenido: cercano a 1
- Sitio web basado en soporte: entre 1 y 2
- Sitio web basado en ventas: mayor que 2

El valor de este ratio depende mucho de cada sitio web, por lo que idealmente es necesario obtener el ratio actual y establecer objetivos para cada sitio. Si el ratio varía en exceso, lo importante que sea por alguna acción que se ha realizado en el sitio y no porque se está disgustando a los clientes habituales o descuidando la captación.

Cabe destacar que la métrica de fidelización está basada en sesiones. Y que por cada visita durante el periodo a analizar, esta métrica cuenta la historia de todo el tiempo disponible (lo que dure o haya durado la cookie). Puede pasar que se tenga una captación potente y que los esfuerzos en SEO, SEM, emailing, etc. se vean recompensados atrayendo a un número importante de visitas. Captar usuarios es importante, pero también lo es fidelizarlos. Si el usuario no vuelve, todos los esfuerzos y dinero invertidos en publicidad habrán sido en vano.

CAPÍTULO 2. ESTADO DEL ARTE

Una métrica que mide la fidelización en un sitio web es el número de sesiones. Un ejemplo de esto se puede observar en la figura 2.2, en la cual, el 70.43 % de las sesiones ha visitado el sitio 1 sola vez, el 19 % ha vuelto entre 1 y 8 veces y el 10.57 % ha visitado el sitio más de 9 veces.

Número de sesiones ?		Sesiones ?
		869 % del total: 100,00 % (869)
<input type="checkbox"/>	1. 1	612 (70,43 %)
<input type="checkbox"/>	2. 2	81 (9,32 %)
<input type="checkbox"/>	3. 3	35 (4,03 %)
<input type="checkbox"/>	4. 4	22 (2,53 %)
<input type="checkbox"/>	5. 5	13 (1,50 %)
<input type="checkbox"/>	6. 6	7 (0,81 %)
<input type="checkbox"/>	7. 7	4 (0,46 %)
<input type="checkbox"/>	8. 8	3 (0,35 %)
<input type="checkbox"/>	9. 10	3 (0,35 %)
<input type="checkbox"/>	10. 11	3 (0,35 %)

Mostrar filas: 10 Ir a: 1 1 - 10 de 72 Actualizar informe

Este informe se creó el 13/6/18 a las 21:15:11. - Actualizar informe

Figura 2.2: Métrica de fidelización en GA.
Fuente: Elaboración Propia

En un blog o sitio web de contenidos, el porcentaje de una sesión debe tender a la baja, dependiendo de cuántas veces se actualice el contenido en el periodo de tiempo analizado.

En un negocio de ventas, puede que un porcentaje alto en sesiones de una sola vez, en un mes, sea lo normal, por lo que quizás habría que ampliar el periodo de análisis.

El primer paso para analizar la fidelización es tomar datos del último año, mes a mes o en grupos de 3 meses, y ver la evolución de las sesiones que repiten en el sitio web. Así se puede tener una idea de los porcentajes de fidelidad del sitio.

Otro tipo de análisis es complementar la información anterior con la métrica de asiduidad (recencia), la cual, indica cuánto tiempo ha pasado desde la última sesión. Esta métrica permite evaluar el porcentaje de nuevos usuarios en el sitio, así como el nivel de familiaridad de todos los usuarios que regresan. Eso sí, es necesario segmentar por usuarios recurrentes, no tiene sentido analizar visitas que solamente han visitado el sitio una vez.

2.2.5.7. Elaboración de Informes

Una de las tareas importantes que a menudo se deja de lado, por enfocarse más en el análisis, es la presentación de la información.

Lo importante es resaltar lo que haya cambiado en el periodo de tiempo analizado, tanto para bien como para mal. Así se pueden identificar mejor las acciones a realizar y los segmentos afectados.

Para ello, existen los **dashboards** (o tableros de control), cuyas características principales son:

- **KPI:** Un dashboard debe tener muchos KPI, lo ideal es entre 5 y 10 por tablero.
- **Contexto:** Es de vital importancia dar un contexto histórico detallado, no solamente mostrar el dato. Los gráficos evolutivos son una buena alternativa.
- **Personalizable:** Tener un dashboard que se adapte a las necesidades de cada usuario, puede ser uno por cada departamento, área, sección, etc.
- **Recomendaciones y explicación de KPI:** Los dashboard siempre tienen que incluir recomendaciones o llamadas a la acción (botón ver más, opción último día, últimos 30 días en un gráfico, etc.) y una breve explicación de qué significa cada KPI, cómo se ha calculado y cómo podría mejorar.

Algunos ejemplos de KPI por áreas claves pueden ser:

- **KPI de compras:** Conversión, importe, costo por visita, costo por compra.
- **KPI de retención:** Número de visitas hasta la compra, número de visitas al mes.
- **KPI de fidelización:** Cada cuánto tiempo compran, cada cuánto tiempo vuelven.
- **KPI actitudinal:** Satisfacción del cliente, recomendación del sitio, uso de otros canales.

Según [5] es posible identificar los siguientes tipos de dashboard:

- **Dashboards ejecutivos:** Estos contienen KPIs de rendimiento del sitio web como: conversiones, beneficio, valor medio de la venta, costo medio de la visita, satisfacción del cliente; con la posibilidad de ser segmentado por tipo de cliente.
- **Dashboard de contenidos:** Poseen KPIs de compromiso con el cliente: tasa de rebote, uso del buscador interno, contenidos a los que accede, ratio de visitantes nuevos vs. visitantes que repiten, tiempo pasado por visita; con la posibilidad de ser segmentado por tipo de cliente.

CAPÍTULO 2. ESTADO DEL ARTE

- **Dashboard de campañas:** Tienen KPIs de rendimiento por campaña: beneficio, tasa de conversión, conversión por keyword, ratio de abandono de compra, costo por adquisición; con la posibilidad de ser segmentado por tipo de fuente de entrada.
- **Dashboard de usabilidad del sitio y mantenimiento:** Poseen KPIs técnicos como: velocidades, consistencia, carga de páginas, uso de navegadores.

Un factor importante para que un dashboard tenga éxito es la forma en que se muestran los datos. Si el lenguaje utilizado está completamente adaptado al receptor, existe una alta probabilidad de que entienda el indicador. La idea es captar la atención del receptor, su interés y poder sacar el máximo de partido al análisis web.

Otro punto importante es la gráfica. No siempre es fácil decidir cómo se visualiza una tendencia, un porcentaje.

Para el análisis web es esencial, un **dashboard informativo** que entregue una primera foto de cómo es el tráfico en general en la web, principalmente para hacerse una idea de qué está pasando, más que para la toma de decisiones. Las métricas que mejor definen el tráfico del sitio web son las siguientes:

- ¿Cuántas visitas tenemos? = n° visitas totales
- ¿Cuántos usuarios únicos? = % usuarios únicos
- ¿Son clientes o no clientes? = % visitas logradas vs. % visitas no logradas
- ¿De dónde vienen? = % visitas desde buscadores vs. % visitas directas
- ¿Son fieles? = % visitas que repiten vs. % visitas que sólo vienen 1 vez
- ¿Les gustamos? = % visitas que están menos de 30 segundos vs. % visitas que se quedan más de 30 segundos. De éstas, % de visitas que solamente ven 1 página vs. % de visitas que ven 2 o 3 páginas vs. % visitas que ven más de 3 páginas.
- ¿Son rentables? = % visitas que convierten y de éstas, % conversiones hechas por clientes vs. % conversiones hechas por no clientes.

2.3. Marketing Digital

Marketing es reconocido como la estrategia comercial, de promoción o propaganda de un producto o servicio [15].

Marketing digital se basa en la utilización de recursos tecnológicos y de medios digitales para desarrollar

CAPÍTULO 2. ESTADO DEL ARTE

comunicaciones directas, personales y que provoquen una reacción en el destinatario. Se utiliza a través de dispositivos electrónicos como: computadores personales, smartphones, tablets, consolas de videojuegos, para lograr que un cliente o potencial cliente compre cierto producto o servicio.

El marketing digital se aplica en tecnologías como sitios web, correos electrónicos, aplicaciones móviles y redes sociales. El término comenzó a ser usado durante los años 90 y su principal crecimiento fue entre los años 2000 y 2010. Hoy en día, las estadísticas muestran que el marketing digital sigue en crecimiento.

Algunas ventajas del marketing digital son:

- **Personalizado:** Permite hacer marketing a medida (personalizado). Analizar información, feedback recibido, puede ayudar a conseguir una base de datos segmentada con lo que es posible lanzar una campaña de marketing adecuada.
- **Masivo:** Con poco dinero es posible llegar a mucha gente a través de distintas tecnologías.
- **Rapidez:** Prácticamente es inmediato, las mediciones se pueden hacer en línea y las posibles soluciones pueden aplicarse casi al instante, no así el marketing no digital.

2.3.1. Tipos de Marketing Digital

Existen dos tipos de marketing digital [13]:

- **Pull digital marketing:** El visitante o consumidor no busca activamente el contenido de marketing, generalmente lo encuentra navegando en la web, en buscadores, al abrir un correo electrónico, al recibir un mensaje de texto, o a través de algún medio de transmisión de audio o video. El visitante no anda en busca de contenido con marketing, pero en el proceso de navegación, lo encuentra.
- **Push digital marketing:** Aquí el vendedor o dueño del sitio web envía un mensaje sin el consentimiento de los destinatarios, impone la publicidad en sitios web o blogs de noticias, en correos electrónico, mensajería de texto y vínculos web.

2.3.2. Métricas

En cualquier tipo de negocio o estrategia, para poder tomar decisiones se necesita información clave. Sin embargo, es necesario medir los resultados obtenidos a través de métricas específicas. Es importante para los directores de marketing medir su actividad y los resultados, además, tener alguna forma de validar el desempeño, por ello se necesitan de métricas que permitan una mejor toma de decisiones.

CAPÍTULO 2. ESTADO DEL ARTE

Las métricas que se deben considerar para el marketing digital dependen de lo que se quiere buscar, puede ser que el enfoque del sitio web sea ventas, de información o de captura de clientes. Es por esto que no hay métricas generales, sino que dependen de lo que se busca. Tomando en cuenta esto, algunas métricas relacionadas a un sitio web de información podrían ser las que permitan contactar a la empresa. Por ejemplo:

- **Porcentaje de clientes originados por el marketing digital:** Esta proporción muestra qué porcentaje del negocio está siendo generado por actividades de marketing digital. Para calcularlo se necesita el número total de nuevos clientes y de éstos, revisar la cantidad que se contactó a través del sitio web (potencial cliente interesado).
- **Porcentaje de clientes influenciados por el marketing digital:** Esta medida es parecida a la anterior, con la diferencia que la influencia de un cliente está dada por el manejo de potenciales clientes. Se analiza si un cliente que tuvo contacto con el sitio web, tuvo contacto con un ejecutivo de ventas o si después volvió a visitar el sitio. También se ve la influencia de las redes sociales sobre el cliente.
- **Cuánta gente contactó al sitio:** Número de personas que visitaron un sitio web y que en el transcurso de la visita se contactó con el sitio.
- **Cuántas personas llenaron un formulario:** Cantidad de personas que llenaron el formulario de contacto de la página web.
- **Número de personas que leyó un tema específico:** Cantidad de personas que visitó un tema del sitio web, así se puede conocer los intereses del visitante.

Actualmente una de las tendencias que está tomando fuerza es la de redes sociales (social media). El crecimiento se debe en gran medida a las redes sociales, lo que ha cambiado un poco el concepto de marketing. Las métricas importantes a seguir dentro de *social media* son:

- **Actividad:** Esfuerzo de comunicación de la empresa en sus diversos canales sociales y blog. Ejemplos: Frecuencia de tweets (Twitter), Respuestas a comentarios (Blogs), Frecuencia de publicaciones (Facebook), Frecuencia de publicación de videos (Youtube) y Contactos directos (Linkedin).
- **Tamaño de la comunidad:** La audiencia estimada de los medios sociales (número de seguidores, fans, suscriptores). Ejemplos: Número de suscriptores al RSS (Blogs), Número de seguidores y su evolución (twitter), Cantidad de Fans y evolución (Facebook) y Menciones (Linkedin).
- **Visibilidad:** Total de menciones de la empresa u organización, productos, marcas o todo lo relacionado con ellas en los medios sociales.

CAPÍTULO 2. ESTADO DEL ARTE

- **Interacción:** Reacción que generan las publicaciones en las redes sociales.
- **Difusión:** Tamaño real de impactos en las redes sociales mediante la viralización de contenidos.

Capítulo 3

Diseño de la Solución

3.1. Metodología utilizada

Para realizar un análisis basado en BI de un portal web universitario se propone utilizar el conjunto de pasos que se entregan en Figura 3.1.

El primer paso, Identificación de Objetivos y KPIs, consiste en definir objetivos e indicadores claves de rendimiento que permitan obtener el perfil de los visitantes del sitio web.

En el paso 2, Selección Fuentes de Datos y Nivel de Granularidad, se identifica la fuente de datos desde la cuál se obtiene la información a ser analizada y se conoce el nivel de detalle de los datos.

El paso 3, se diseña el Modelo de Datos Lógico que buscará satisfacer los objetivos de la solución y el nivel de granularidad explicados en el paso anterior.

Estos 3 pasos constituyen el diseño de la solución que será entregado en el presente capítulo.

El paso 4, permite la construcción e implementación de la solución mediante la selección de la herramienta de BI que se utilizará, las diferentes configuraciones que se realizarán para conectar Google Analytics y esta herramienta, la implementación del modelo lógico en la herramienta de BI y la creación de dashboard con información relevante que servirá para cumplir los objetivos planteados inicialmente.

Finalmente, en el paso 5, se realiza el análisis de los resultados obtenidos con el fin de obtener el perfil de visitantes del portal del Departamento de Informática de la UTFSM.



Figura 3.1: Pasos de Metodología Utilizada para el Diseño de la Solución.
Fuente: Elaboración Propia.

3.2. Identificación de Objetivos y KPIs

3.2.1. Objetivos

Uno de los objetivos de la solución es analizar el tráfico de un portal web universitario, con el fin de obtener el perfil de los visitantes y así orientar la estrategia de marketing digital de una carrera universitaria.

Para conocer el tráfico del portal web y obtener el perfil de los visitantes se definieron los siguientes objetivos:

1. **Identificar audiencia:** Permite identificar cómo es y qué cosas tienen en común los visitantes de la web. Se puede saber si la visita está fidelizada, si son nuevos visitantes, la frecuencia de éstos, su información geográfica, datos demográficos, intereses y tecnología ocupada.
2. **Conocer el origen de las visitas:** Esto con el fin de entender las fuentes de origen de la visita. Además, identificar canales y URLs por dónde ingresa el usuario.

CAPÍTULO 3. DISEÑO DE LA SOLUCIÓN

3. **Identificar el comportamiento del usuario:** Este objetivo permite obtener información sobre lo que hacen las visitas cuando llegan a la web, también, qué contenido es el que más les gusta. Se pueden identificar las páginas de entrada, por dónde se mueven y por dónde salen con el fin de optimizar el sitio web donde realmente hace falta.

3.2.2. KPIs o Métricas

A continuación se detallarán los KPI asociados a los objetivos mencionados en la sección anterior, los cuales permiten medirlos y posteriormente tomar decisiones en base a su valor.

1. **Usuarios:** Usuarios que han iniciado al menos una sesión en un periodo determinado.
2. **Usuarios nuevos:** Número de usuarios nuevos en un periodo específico.
3. **Usuarios recurrentes:** Número de usuarios que vuelven a la web en un periodo determinado.
4. **Fidelización:** Porcentaje de usuarios que vuelven vs usuarios que repiten:

$$Fidelizacion = \frac{Recurrentes}{Unicos}$$

5. **Sesiones:** Número total de sesiones en un periodo. Una sesión es el periodo en el cual el usuario interactúa con el sitio web.
6. **Porcentaje de rebote:** Porcentaje de sesiones en las que no se ha interactuado con la página:

$$Porcentaje\ de\ Rebote = \frac{Visitas\ que\ vieron\ 1\ pagina\ y\ se\ fueron * 100}{Total\ de\ Visitas}$$

7. **Profundidad:** Promedio de páginas vistas por sesión:

$$Profundidad = \frac{Paginas\ Vistas}{Sesion}$$

8. **Duración media de la sesión:** Promedio en tiempo de una sesión:

$$DuracionMediaSesion = \frac{Tiempo\ en\ Paginas}{Sesion}$$

9. **Número de visitas a páginas:** Número total de páginas vistas.
10. **Número de páginas vistas únicas:** Cantidad de sesiones durante las cuales se ha visto la página especificada al menos una vez.
11. **Promedio de tiempo en la página:** Tiempo medio durante el cual los usuarios han visitado una determinada página:

$$PromedioenPagina = \frac{Tiempo\ en\ pagina}{Pagina}$$

CAPÍTULO 3. DISEÑO DE LA SOLUCIÓN

12. **Entradas:** Número de veces que los visitantes han accedido al sitio desde una página o conjunto de páginas determinadas.

13. **Porcentaje de nuevas sesiones:** Porcentaje estimado de visitas realizadas por primera vez:

$$\text{PorcentajeNuevasSesiones} = \frac{\text{Nuevas Sesiones}}{\text{Total Sesiones}}$$

14. **Salidas:** Número de veces que los visitantes han abandonado el sitio desde una página o conjunto de páginas determinadas.

15. **Porcentaje de Salidas:** Porcentaje de salidas del sitio desde una página o conjunto de páginas determinadas:

$$\text{PorcentajeSalida} = \frac{\text{Visitas que salieron por pagina}}{\text{Visitas a pagina}}$$

16. **Tiempo medio de carga de la página:** Promedio de tiempo (en segs.) que demora la página en cargar.

Las métricas Usuarios, Usuarios Nuevos, Usuarios Recurrentes, Fidelización, Sesiones, Porcentaje de Rebote, Profundidad, Duración Media de la Sesión, Número de Visitas a páginas, ayudan a identificar el primer y tercer objetivo: **Identificar Audiencia** y **Conocer el Origen de las Visitas**.

Finalmente las métricas Número de Visitas a Páginas, Número de Páginas Vistas Únicas, Promedio de Tiempo en la Página, Porcentaje de Rebote, Entradas, Salidas, Porcentaje de Salida, Tiempo Medio de Carga de la Página, se relacionan con el tercer objetivo: **Identificar el Comportamiento del Usuario**.

3.3. Fuente de Datos y Nivel de Granularidad

3.3.1. Fuentes de Datos

La fuente de datos utilizada es Google Analytics (GA), a través de la cual se obtiene información sobre el tráfico del sitio web del Departamento de Informática de la UTFSM para su carrera de pregrado de Ingeniería Civil Informática.

La información se puede obtener mediante consultas preestablecidas o consultas manuales en las que se especifica lo que se quiere. Para ello se utiliza la herramienta de BI (MicroStrategy Desktop); se debe partir configurando la cuenta de GA de la siguiente forma:

1. En la cuenta de desarrollador de GA, seleccionar el proyecto y navegar hasta API Manager, Overview para habilitar la Api de GA.

CAPÍTULO 3. DISEÑO DE LA SOLUCIÓN

2. Navegar a API Manager, Credenciales y crear una Nueva credencial del tipo OAuth client ID.
3. Seleccionar el tipo de Aplicación (Web, Android, Chrome, iOS, PS4, Otro). En este caso Aplicación Web. Además, rellenar los campos Authorized JavaScript origins y Authorized redirect URLs con la información proporcionada por la herramienta de BI.
4. Finalmente el Cliente ID queda creado para ser usado posteriormente en la obtención de la información.

CAPÍTULO 3. DISEÑO DE LA SOLUCIÓN

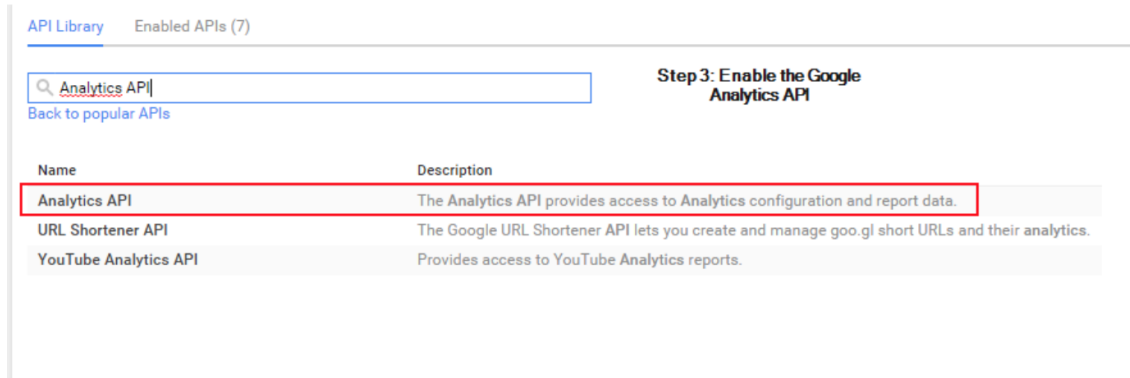


Figura 3.2: Habilitación API de Google Analytics.
Fuente: Elaboración Propia.

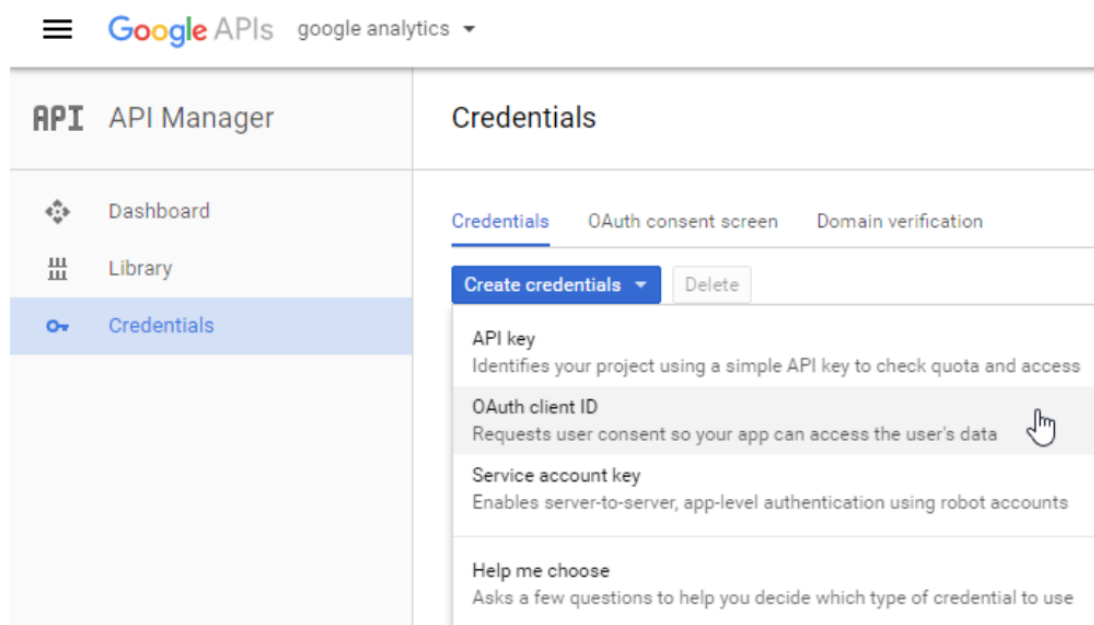


Figura 3.3: Tipos de Credenciales.
Fuente: Elaboración Propia.

CAPÍTULO 3. DISEÑO DE LA SOLUCIÓN

The screenshot shows the Google API Manager interface. On the left is a sidebar with 'API Manager' and a menu containing 'Dashboard', 'Library', and 'Credentials' (which is selected). The main area is titled 'Credentials' and contains a 'Create client ID' form. The 'Application type' section has 'Web application' selected. The 'Name' field contains 'Web client 2'. The 'Restrictions' section has two text input fields: 'Authorized JavaScript origins' with the value 'http://www.example.com' and 'Authorized redirect URIs' with the value 'http://www.example.com/oauth2callback'. At the bottom are 'Create' and 'Cancel' buttons.

Figura 3.4: Rellenar URL's autorizadas.
Fuente: Elaboración Propia.

3.3.2. Nivel de Granularidad

El nivel de granularidad consiste en el nivel de detalle de los datos a obtener. En este caso para cada objetivo se desea obtener lo siguiente:

- **Identificar audiencia:** El nivel de detalle corresponde al registro que se hace por Día, Usuario (Edad, Sexo, Tipo de Usuario), Intereses (Afinidad, Segmentos del mercado, Otras categorías), Información Geográfica (Idioma, País, Ciudad), Sistema (Navegador, Sistema Operativo, Proveedor de Servicios), Tipo de Dispositivo (Dispositivo, Sistema Operativo, Proveedor de Servicios, Resolución de Pantalla) y Frecuencia (Número de Sesiones).
- **Conocer del origen de las visitas:** El nivel de detalle corresponde a los registros por Día, Canales, Medio y Fuente.
- **Identificar el comportamiento de los usuarios:** El nivel de detalle corresponde a los registros por Página (Nombre Página, Nivel de Ruta, Página de Destino, Página de Salida), Navegador y País.

Una vez obtenido el nivel de detalle, se pueden indentificar las características o dimensiones que permiten dar respuesta a los objetivos. Éstas se pueden observar en la tabla 3.1.

Tabla 3.1: Dimensiones de la solución.

Fuente: Elaboración Propia.

Dimensiones
Tiempo
Usuario
Zona
Intereses
Dispositivo
Sesiones
Origen
Página

3.4. Diseño Modelo de Datos Lógico

El modelo lógico de la solución estará constituido por 3 Data Marts (DM): BT_Audiencia, BT_Origen y BT_Comportamiento, los cuales buscan satisfacer los objetivos de la solución y están creados de acuerdo a las dimensiones y métricas explicadas en la sección 3.2 y 3.4 respectivamente. Además, satisface el nivel de granularidad que se explicó en la sección anterior.

3.4.1. Data Mart Audiencia

Con este DM se dará respuesta al objetivo relacionado con la Audiencia. Permitirá identificar cómo son los visitantes del sitio web. Saber si son usuarios fieles, nuevos, su frecuencia. Se puede observar este DM en la figura 3.5

3.4.2. Data Mart Origen

Con este DM se pueden obtener las fuentes de origen de los visitantes. Ya sea una fuente orgánica, tráfico directo, fuente pagada. Además desde qué buscadores vienen. La figura 3.6 muestra el modelo del DM.

3.4.3. Data Mart Comportamiento

Con el siguiente DM se puede obtener información sobre el comportamiento de los visitantes en la web. Cuál es el contenido que consumen, el que más les gusta, páginas de entrada, de salida. El modelo de DM

CAPÍTULO 3. DISEÑO DE LA SOLUCIÓN

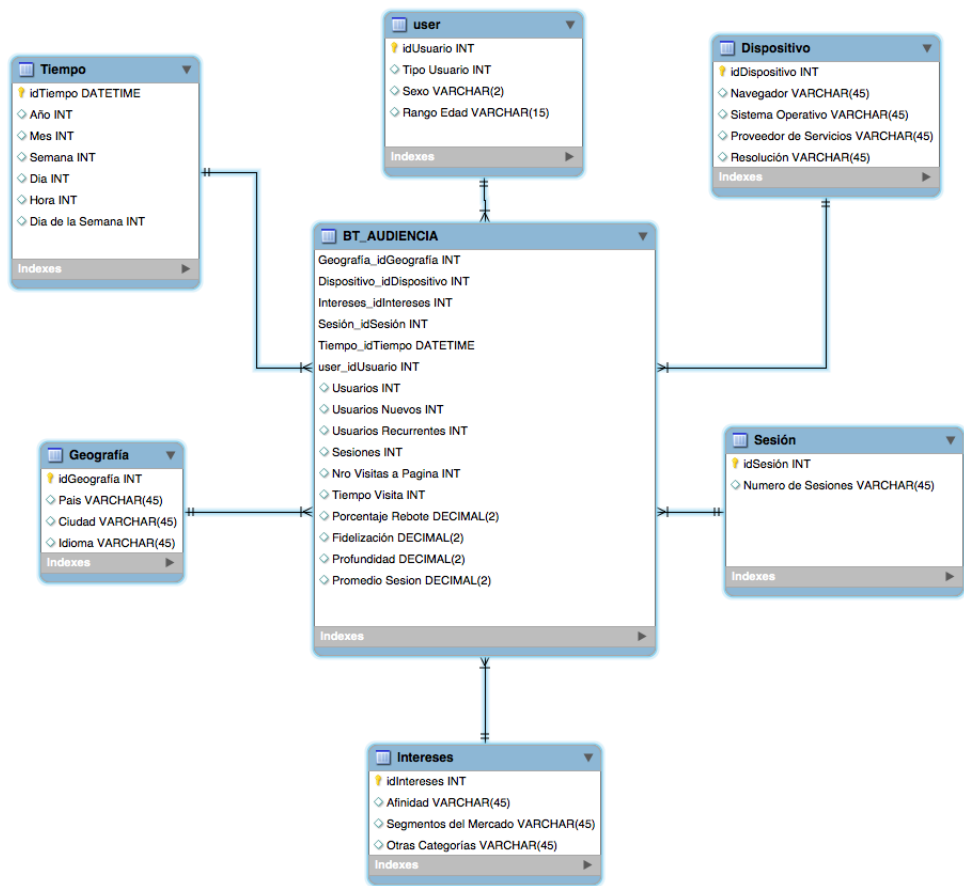


Figura 3.5: Data Mart Audiencia.
Fuente: Elaboración Propia.

se puede observar en la figura 3.7

CAPÍTULO 3. DISEÑO DE LA SOLUCIÓN

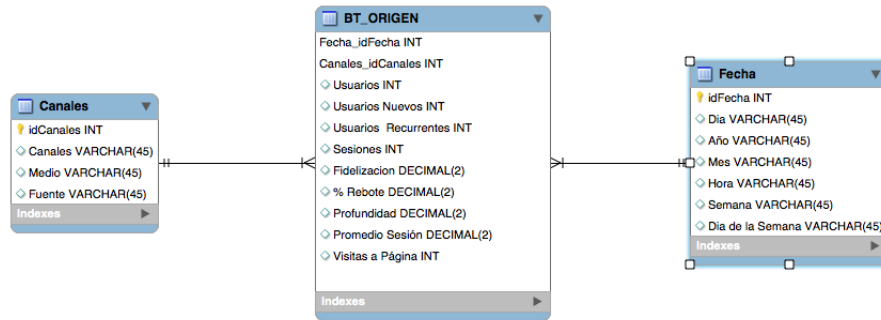


Figura 3.6: Data Mart Origen.
Fuente: Elaboración Propia.

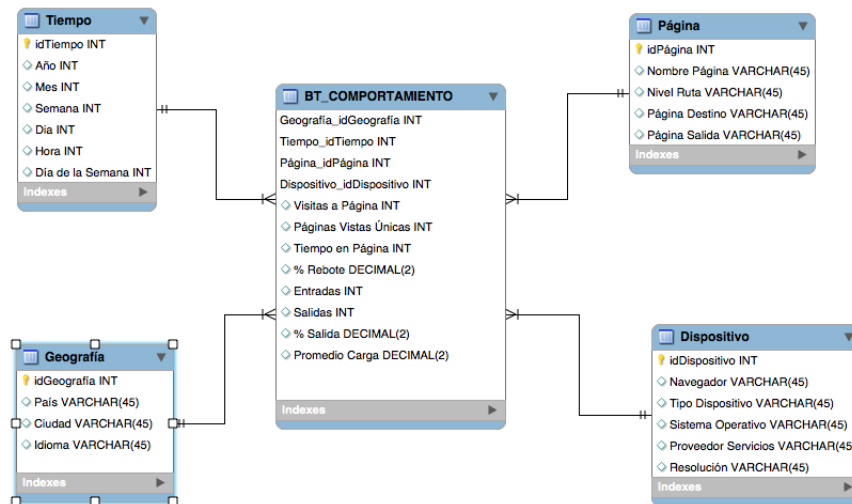


Figura 3.7: Data Mart Comportamiento.
Fuente: Elaboración Propia.

Capítulo 4

Construcción e Implementación de la Solución

4.1. Selección de Herramienta de BI

Para seleccionar la herramienta de BI a utilizar, se consideró lo propuesto en el Cuadrante Mágico de Gartner para la Inteligencia de Negocios [2018], donde se clasifican a los distintos proveedores de herramientas en cuadrantes según su grado de experiencia en el mercado y funcionalidades ofrecidas. Ver Figura 4.1.

A continuación se destacan las 4 herramientas más importantes de los últimos 4 años según Gartner:

- Tableau: Es una herramienta de BI que permite analizar grandes volúmenes de información en forma rápida, flexible y con altos estándares de usabilidad. Tableau está orientada para que usuarios comunes y corrientes sean capaces de interpretar y comprender los datos. Según Gartner es la herramienta líder en Inteligencia de Negocios del 2018, gracias a su capacidad de ejecución y visión integral. Las características principales son:
 - Fomenta el trabajo colaborativo: Se pueden compartir dashboards de manera segura, controlar datos con alertas y suscripciones.
 - Permite análisis avanzados: Potencia la obtención de información a partir de cálculos y estadísticas de nivel avanzado.
 - Provee gobernanza para datos seguros y confiables: Permite administrar y optimizar el uso de

CAPÍTULO 4. CONSTRUCCIÓN E IMPLEMENTACIÓN DE LA SOLUCIÓN

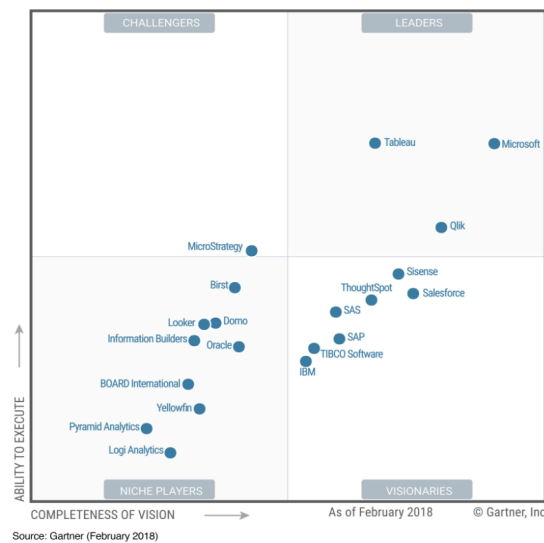


Figura 4.1: Cuadrante Mágico de Gartner 2018
Fuente: Gartner-2018.

datos con fuentes de datos centralizadas, certificación de fuentes de datos, permisos de nivel de fila y análisis de uso.

- Ofrece altos niveles de usabilidad: Tiene una interfaz visual e interactiva con la cual se puede preparar el dato. Se pueden crear uniones o combinaciones. Permite modelamiento automático para ver la información con más claridad.
 - Diferentes formas de acceso a los datos: Se puede acceder a los datos desde la nube o instalaciones físicas, además, combinarlos sin escribir código. También se puede acceder en tiempo real o para extracciones dependiendo de la necesidad.
- Qlik: Empresas de software que ofrece soluciones intuitivas en una plataforma para visualización de datos de autoservicio, aplicaciones de analítica, analítica integrada e informes. Es la creadora de los softwares QlikView (Inteligencia de Negocios), Qlik Sense (Plataforma de Visualización de Datos) y Qlik Sense Cloud (Plataforma para crear apps en la nube). Algunas características de estas herramientas son:
- Conexión con múltiples orígenes de datos: Permite utilizar cualquier fuente válida de datos, desde archivos Excel hasta todo tipo de base de datos.
 - Indexación de datos: Utiliza la tecnología QIX (Qlik Indexing Engine), la cual hace que todos los datos se relacionen o asocien de manera correcta, independiente si vienen de diferentes fuentes de datos.

CAPÍTULO 4. CONSTRUCCIÓN E IMPLEMENTACIÓN DE LA SOLUCIÓN

- Creación de paneles personalizados: Permite la creación de informes o paneles personalizados para todos los usuarios.
 - Generación de análisis de diversos tipos: Permite descubrir información para la correcta toma de decisiones.
- Microsoft: Empresa tecnológica de nivel mundial que desarrolla, manufactura, licencia y provee soporte de softwares para computadores personales, servidores, dispositivos electrónicos y servicios. En el área de la Inteligencia de Negocios posee la herramienta Power BI, la cual consiste en un conjunto de herramientas de análisis empresarial que pone el conocimiento al alcance de toda organización. Posee características como:
- Conexión a cientos de orígenes de datos
 - Preparación de datos simplificada
 - Generación de informes en la web y dispositivos móviles
 - Creación de paneles personalizados
 - Generación de análisis de diversos tipos
- Microstrategy: Empresa que provee software de Inteligencia de Negocios para realizar análisis de datos a partir de una base de datos relacional o de otros tipos de fuentes. Sus características más importantes son:
- Conectividad de datos: Cuenta con docenas de controladores y puertas de enlace hacia distintas fuentes de datos.
 - Informes para empresas: Creación de informes personalizados y cuadros de mandos para las distintas áreas de la organización.
 - Descubrimiento de datos: Permite que toda la organización mejore la productividad mediante el acceso a los datos de múltiples equipos de trabajo. Ofrece libertad para explorar datos por cuenta propia, sin sacrificar la gobernabilidad, escalabilidad y seguridad de ellos.
 - Distribución: Automatiza la entrega de informes, documentos y cuadros de mando a cada usuario. Asegura que los encargados de tomar las decisiones tengan siempre acceso a la información cuándo y dónde la necesiten.
 - Movilidad: Posibilidad de crear aplicaciones de productividad para posteriormente visualizar la información en móviles o tablets.

CAPÍTULO 4. CONSTRUCCIÓN E IMPLEMENTACIÓN DE LA SOLUCIÓN

- **Credenciales digitales:** Credenciales para cada sistema físico y lógico de la empresa. La idea es reemplazar las tarjetas de identificación por tarjetas de identificación móvil que los empleados pueden utilizar para abrir puertas, desbloquear computadores y acceder a los sistemas.

En general las herramientas presentadas tienen características similares, que se pueden sintetizar en:

- Generación de reportes y dashboards.
- Generación de diferentes tipos de análisis a partir de los datos para la toma de decisiones.
- Realizar predicciones a partir de los datos existentes.
- Disponibilizar los datos a las distintas áreas.
- Recopilar información de diferentes orígenes de datos.

Para seleccionar cuál de estas herramientas se utilizará en la presente memoria, se parte de la base que se ocupará la versión gratuita de cada una de ellas, y se definen los siguientes criterios de selección:

- **Tipo de Sistema Operativo:** El Sistema Operativo juega un rol fundamental a la hora de elegir una herramienta. Si la herramienta está disponible para la mayoría de los S.O, se facilita la implementación de esta.
- **Conexión a GA:** Las herramientas de BI tienen diversas opciones de conexión a fuentes de datos. No todas tienen conexión directa a Google Analytics, por lo que se privilegiará alguna que sí tenga conexión a esta porque es la fuente de datos de la solución.
- **Conocimiento Previo:** Es un aspecto importante a la hora de elegir la plataforma de visualización, debido a que al tener un conocimiento un poco más avanzado, se minimizan las horas de aprendizaje que provoca una nueva herramienta.

Otro factor importante es si posee conexión directa a GA ya que es el origen de datos de la solución. Además, el conocimiento previo de la herramienta es relevante. Para finalizar, otro factor importante es la existencia de versiones para los diferentes sistemas operativos.

En Tabla 4.1, se puede visualizar la comparación de las 4 herramientas, teniendo en cuenta que la puntuación otorgada varía entre 0 y 2, donde 0 indica que no posee la característica, 1 la posee a medias y 2 la posee completamente.

Finalmente, se opta por utilizar Microstrategy Desktop ya que cumple con la conexión a GA, está disponible para todos los sistemas operativos y el conocimiento de la herramienta es alto. Además, hay que

CAPÍTULO 4. CONSTRUCCIÓN E IMPLEMENTACIÓN DE LA SOLUCIÓN

Tabla 4.1: Tabla comparativa herramientas BI.

Fuente: Elaboración Propia.

Herramienta	Características		
	S.O	Conexión GA	Conocimiento Previo
Tableau	2	1	1
Qlik Sense	1	1	1
Power BI	1	2	1
Microstrategy	2	2	2

señalar que posee una herramienta de preparación de datos la que permite tener una vista previa del dato y se pueden especificar las opciones antes de la importación o realizar cambios después que se hayan importado.

4.2. Conexión Microstrategy a Google Analytics

En las versiones 10.x de Microstrategy los usuarios pueden importar los datos estadísticos desde GA de manera rápida. Existen dos opciones para importar estos datos:

1. Importar los datos desde GA usando consultas preestablecidas.
2. Crear consultas especificando los datos a importar.

Existen una serie de pasos para conectar Microstrategy con GA y poder importar la información:

1. En Microstrategy Desktop se deben agregar nuevos datos. Posterior a esto, se debe acceder con la cuenta de Google asociada a GA y otorgar permisos a Microstrategy, tal como se puede observar en las figuras 4.2 y 4.3
2. Finalmente aparece una interfaz en la cual se pueden importar los datos con las consultas preestablecidas o manuales. Ver figura 4.4

CAPÍTULO 4. CONSTRUCCIÓN E IMPLEMENTACIÓN DE LA SOLUCIÓN

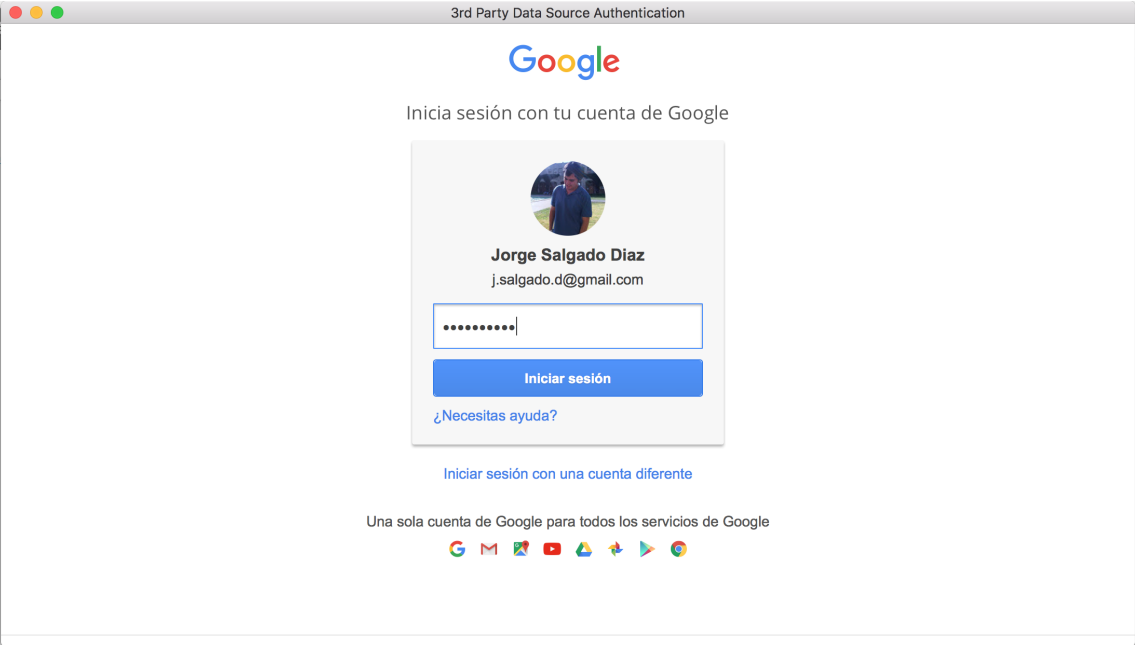


Figura 4.2: Log in Gmail Analytics.
Fuente: Elaboración Propia.

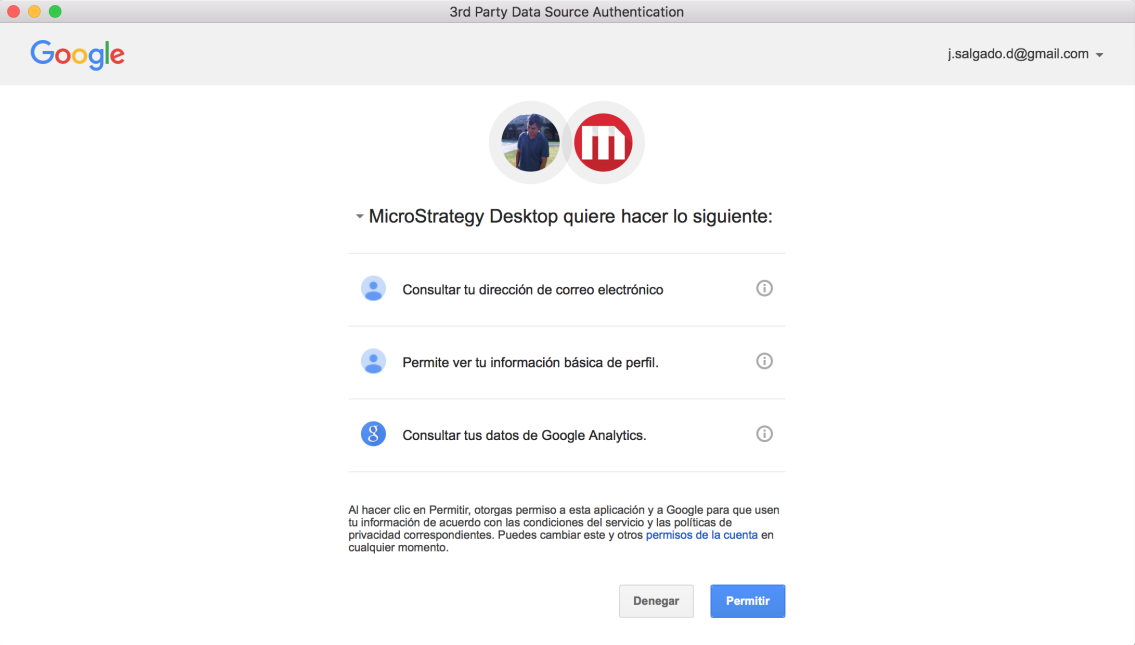


Figura 4.3: Otorgar permisos a Microstrategy.
Fuente: Elaboración Propia.

CAPÍTULO 4. CONSTRUCCIÓN E IMPLEMENTACIÓN DE LA SOLUCIÓN

Google Analytics: Importar datos de Web Analytics desde su cuenta

Seleccione el archivo de Google Analytics

Jorge Salgado Díaz (j.salgado.d@gmail.com) [Cerrar sesión](#)

Seleccione su cuenta, propiedad web y perfil [Ocultar](#)

Cuenta: Sitio Informatica USM

Propiedad web: web DI

Perfil: Todos los datos de sitios web

Seleccione el intervalo de fechas: [Ocultar](#)

Intervalo de fechas: Mes anterior

Desde: MM/DD/YYYY

Hasta: MM/DD/YYYY

☒ Elegir un informe ☐ Escriba su propia consulta

Archivos	Descripción
Detailed Audience Device Data	
Detailed Audience Geo-Location Data	
Detailed Ecommerce Data	
Detailed Event Tracking Data	

[Atrás](#) [Finalizar](#) [Preparación de datos](#) [Cancelar](#)

Figura 4.4: Importar datos de Web Analytics desde su cuenta

Fuente: Elaboración Propia.

4.3. Modelo de datos Microstrategy

Para cargar datos y poder visualizarlos en Microstrategy, se deben ocupar consultas preestablecidas o manuales. La herramienta permite relacionar las tablas en base al atributo que tengan en común. El proceso para cargar las tablas es el siguiente:

1. Se copia el script obtenido de la web <https://ga-dev-tools.appspot.com/query-explorer> o se utiliza una consulta preestablecida.
2. Se edita el script para obtener las métricas y dimensiones necesarias para la solución. Existe un límite de 10 métricas y 7 dimensiones por script.
3. Microstrategy permite la definición explícita entre las tablas y no es necesario que las dimensiones o atributos tengan el mismo nombre.
4. Es posible modificar los datos antes que se importen o realizar cambios en los datos después que se hayan importado. Además, es posible definir manualmente las columnas de datos como atributos o indicadores antes de guardar el conjunto de datos. También, existe la posibilidad de transformar los datos. Se puede mejorar la calidad de estos antes de importarlo: quitar espacios en blanco, filtrar datos, eliminar filas o celdas duplicadas, reemplazar datos.

En la figura 4.5 se puede observar la vista previa de las tablas antes de ser importadas. Las dimensiones tienen un ícono de color verde y las métricas un ícono de indicador. Estas tablas están en memoria, por

CAPÍTULO 4. CONSTRUCCIÓN E IMPLEMENTACIÓN DE LA SOLUCIÓN

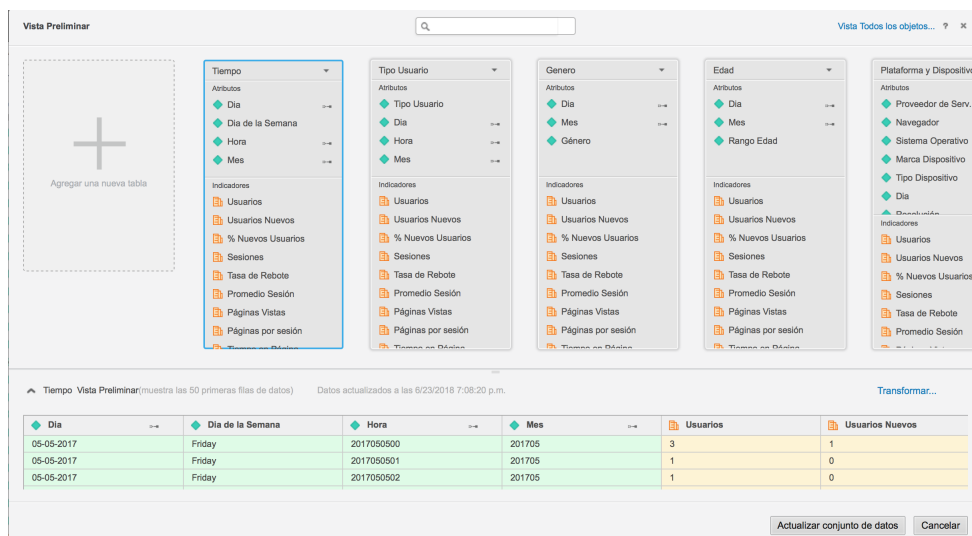


Figura 4.5: Vista previa de datos en Microstrategy.

Fuente: Elaboración Propia.

lo que el acceso es mucho más rápido que la lectura de los mismos en un disco duro. La mejora en el rendimiento de las consultas, permite que su modificación sobre la marcha y la navegación en general, sean mucho más fluidas, lo que mejora la experiencia de usuario.

4.4. Visualización en Escritorio

Para la visualización en la herramienta de escritorio se crearon dashboards con información general y pestañas con información más específica. En todos es posible filtrar la data por rango de fechas.

4.4.1. Dashboard Audiencia

Este panel contiene las pestañas general, geolocalización, afinidad, categorías del mercado, otras categorías, dispositivos y navegador y sistema operativo.

La primera tiene información general de los usuarios que visitan la web (figura 4.6). Existen métricas de cantidad de usuarios, tasa de páginas vistas por sesión, usuarios nuevos, sesiones, total páginas vistas y sesiones por usuario. Además, un gráfico con la cantidad de usuarios que acceden por hora. Finalmente, cantidad de usuarios por país, idioma y por ciudad.

La pestaña geolocalización (figura 4.7) posee información de sesiones por continente y un mapa con el detalle de sesiones por ciudad.

CAPÍTULO 4. CONSTRUCCIÓN E IMPLEMENTACIÓN DE LA SOLUCIÓN

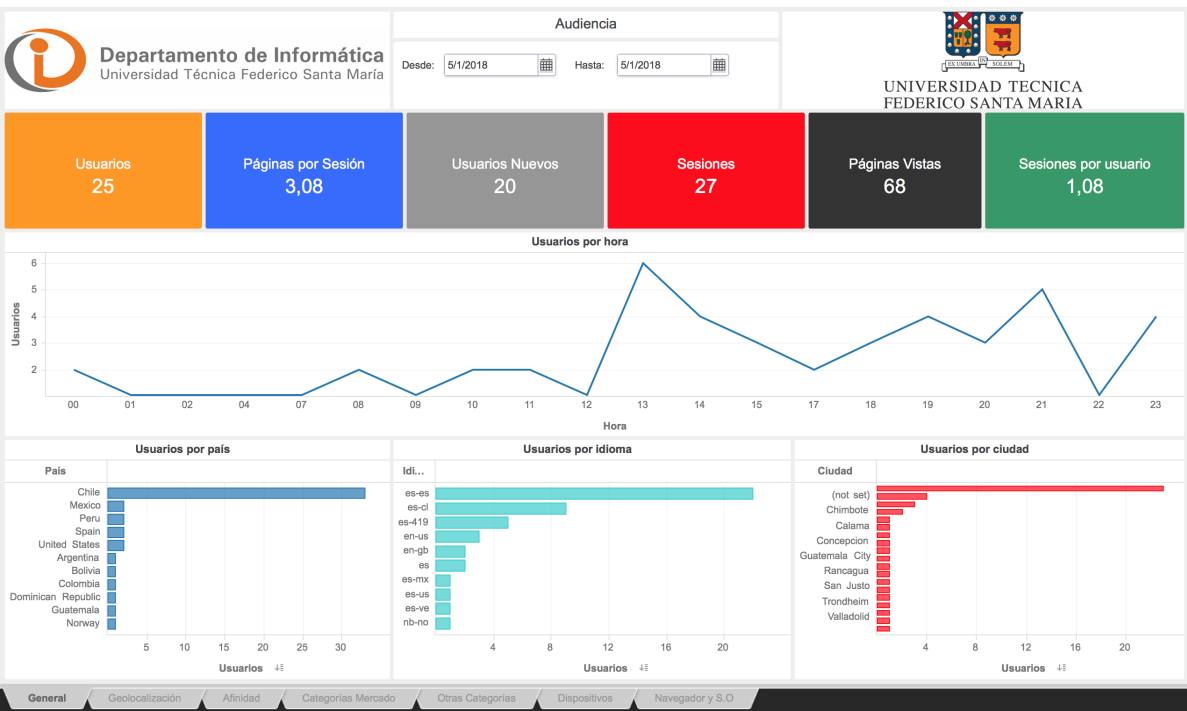


Figura 4.6: Dashboard Audiencia.
Fuente: Elaboración Propia.

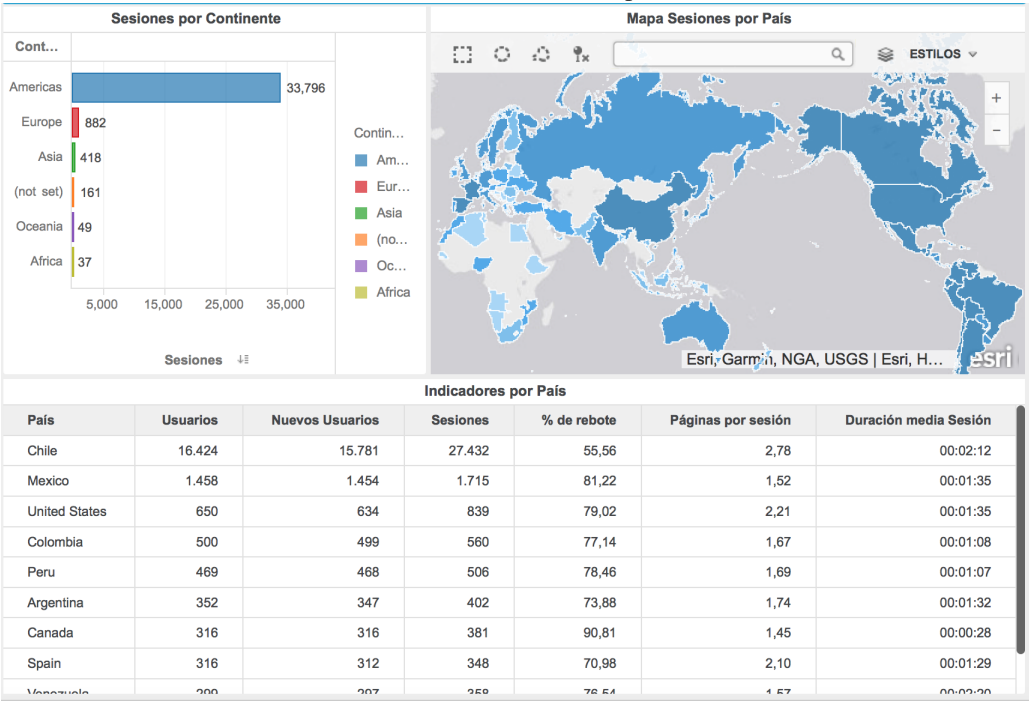


Figura 4.7: Geolocalización.
Fuente: Elaboración Propia.

CAPÍTULO 4. CONSTRUCCIÓN E IMPLEMENTACIÓN DE LA SOLUCIÓN

En cuanto a afinidad (figura 4.8), se puede generar un gráfico con las distintas categorías de afinidad de los usuarios. También, una cuadrícula segmentada por las categorías de afinidad e indicadores como usuarios, usuarios nuevos, sesiones, porcentaje de rebote, páginas por sesión y promedio de la sesión.

En la figura 4.9, se puede observar información sobre los tipos de dispositivos (computadores, móviles y tablets). Además, información de los navegadores de estos dispositivos.

Finalmente en la figura 4.10, en lo que respecta a los dispositivos móviles por los que accede el usuario al portal web, se visualiza la cantidad de usuarios por marca y por sistema operativo.

CAPÍTULO 4. CONSTRUCCIÓN E IMPLEMENTACIÓN DE LA SOLUCIÓN

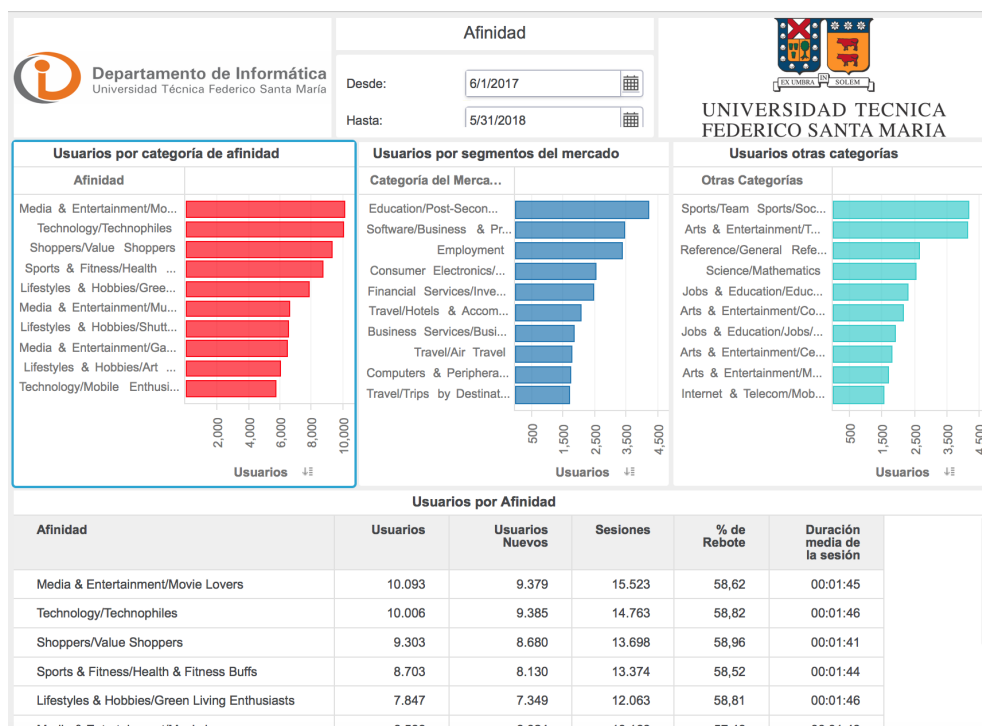


Figura 4.8: Afinidad.

Fuente: Elaboración Propia.

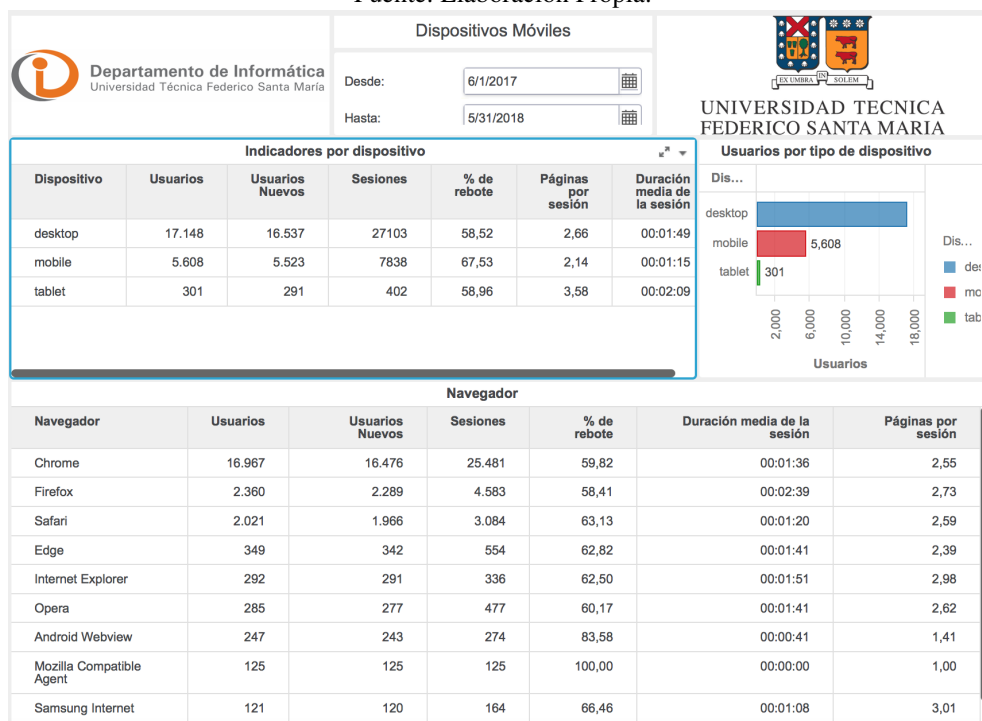


Figura 4.9: Tipo de Dispositivo.

Fuente: Elaboración Propia.

CAPÍTULO 4. CONSTRUCCIÓN E IMPLEMENTACIÓN DE LA SOLUCIÓN

Marca Dispositivo						
Marca Dispositivo	Usuarios	Usuarios Nuevos	Sesiones	% de rebote	Páginas por sesión	Duración media de la sesión
Samsung	1.753	1.726	2.569	62,79	2,46	00:01:33
Apple	1.581	1.562	2.103	70,14	2,13	00:01:04
Huawei	606	596	849	67,49	2,01	00:01:20
Motorola	558	535	807	71,13	1,89	00:00:52
LG	336	334	478	66,95	2,44	00:01:29
(not set)	251	247	317	70,98	1,87	00:01:02

S.O						
Sistema Operativo	Usuarios	Usuarios Nuevos	Sesiones	% de Rebote	Páginas por sesión	Duración media de la sesión
Android	4.293	4.218	6.079	66,21	2,23	00:01:22
iOS	1.581	1.562	2.103	70,14	2,13	00:01:04
(not set)	91	91	91	98,90	1,01	00:00:00
Windows Phone	10	10	17	58,82	2,12	00:00:34
Firefox OS	2	2	10	40,00	4,20	00:00:46
Tizen	2	2	2	100,00	1,00	00:00:00
BlackBerry	1	1	1	100,00	1,00	00:00:00

Figura 4.10: Marca y S.O.
Fuente: Elaboración Propia.

4.4.2. Dashboard Comportamiento

Este dashboard está constituido por un resumen general del comportamiento del usuario en las distintas páginas de la web, información de las páginas de destino, páginas de salida y desglose de contenido.

En la pestaña general (figura 4.11), se encuentra información de páginas vistas, páginas únicas vistas, porcentaje de salidas, promedio de tiempo en página y porcentaje de rebote. También, un gráfico con el número de páginas vistas por hora. Finalmente, dos cuadrículas con información de los títulos de página y rutas de las páginas.

En la pestaña páginas de destino (figura 4.12), se puede encontrar información relacionada a las distintas páginas de la web. En la cuadrícula hay información de las sesiones, % de nuevas sesiones, nuevos usuarios, % de rebote, páginas por sesión y duración media de la sesión.

La segunda pestaña, páginas de salida, contiene información de las páginas por las que sale el usuario. En la cuadrícula se pueden observar indicadores con la cantidad de salidas, páginas vistas y % de salidas de la ruta de la página (figura 4.13).

Finalmente, en la pestaña desglose de contenido (figura 4.14) se pueden apreciar indicadores de páginas vistas, páginas únicas vistas, promedio de tiempo en página, % de salidas y % de rebote por página por Ruta de Página y Página. Todo ordenado de mayor a menor por la cantidad de páginas vistas.

CAPÍTULO 4. CONSTRUCCIÓN E IMPLEMENTACIÓN DE LA SOLUCIÓN

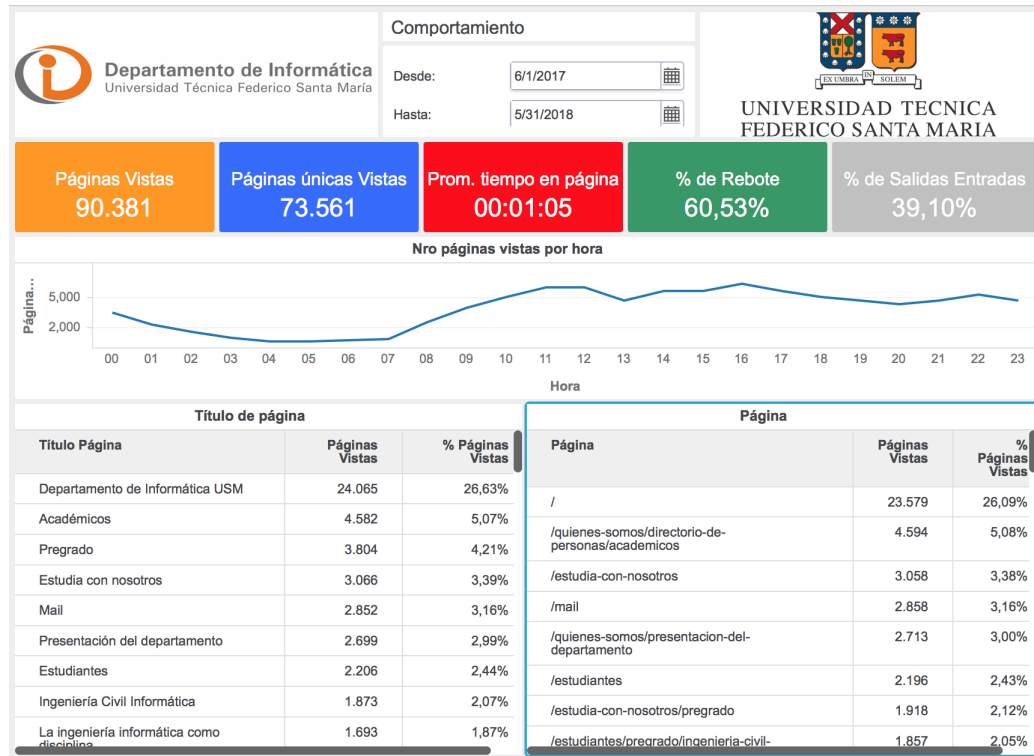


Figura 4.11: Dashboard Comportamiento.
Fuente: Elaboración Propia.

Páginas de destino						
Página de destino	Sesiones	% Nuevas Sesiones	Usuarios Nuevos	% de Rebote	Páginas/sesión	duración media sesión
/	18.362	58,22%	10690	47,11%	3,20	00:02:10
/quienes-somos/la-ingenieria-informatica-como-disciplina	1.300	91,08%	1184	87,08%	1,22	00:00:48
/vinculacion/programa-de-continuidad-de-estudios	1.187	91,07%	1081	74,73%	1,56	00:00:48
/estudiantes/pregrado/ingenieria-civil-informatica	1.063	63,12%	671	77,23%	1,70	00:01:05
/mail	957	35,53%	340	80,56%	1,37	00:01:01
/applications	787	57,56%	453	81,07%	1,65	00:00:59
/quienes-somos/directorio-de-personas/academicos	299	52,51%	157	44,15%	4,20	00:02:53
/quienes-somos/organizacion/estructura-organizacional	265	88,30%	234	84,53%	1,41	00:00:42
/quienes-somos/directorio-de-personas/academicos/s-dr-horst-von-brand	252	76,98%	194	75,40%	2,12	00:00:56
/noticias/194-fallecimiento-ex-alumna-constanza-matamoros-avila	214	73,83%	158	78,97%	1,68	00:01:26
/estudia-con-nosotros/pregrado/descripcion-de-la-carrera	211	61,14%	129	74,88%	2,06	00:01:38

Figura 4.12: Páginas de destino.
Fuente: Elaboración Propia.

CAPÍTULO 4. CONSTRUCCIÓN E IMPLEMENTACIÓN DE LA SOLUCIÓN

Ruta de Página			
Página	Salidas	Páginas Vistas	Porcentaje de Salidas
/	10.592	23.579	44,92%
/mail	2.293	2.858	80,23%
/quienes-somos/la-ingeniería-informatica-como-disciplina	1.307	1.709	76,48%
/estudiantes/pregrado/ingeniería-civil-informatica	1.223	1.857	65,86%
/vinculacion/programa-de-continuidad-de-estudios	1.072	1.480	72,43%
/applications	860	1.341	64,13%
/quienes-somos/directorio-de-personas/academicos	649	4.594	14,13%
/estudia-con-nosotros/pregrado	630	1.918	32,85%
/estudia-con-nosotros/pregrado/descripcion-de-la-carrera	564	1.128	50,00%
/estudia-con-nosotros/postgrado	416	931	44,68%
/estudia-con-nosotros	412	3.058	13,47%
/estudia-con-nosotros/postgrado/mti	316	488	64,75%

Figura 4.13: Páginas de salida.
Fuente: Elaboración Propia.

Nivel de ruta de página 1					
Ruta de Página 1(Agrupar)	Páginas Vistas	Páginas únicas Vistas	Promedio tiempo en la página	% de Rebote	% de Salidas
/	23.579	19.027	00:01:21	47,11%	44,92%
/quienes-somos/	23.417	18.145	00:00:53	72,46%	30,17%
/estudia-con-nosotros	8.973	7.559	00:00:52	71,21%	31,15%
/estudiantes	8.133	7.036	00:00:57	73,92%	35,67%
/investigacion/	3.949	2.983	00:00:37	71,97%	19,47%
/noticias/	3.418	2.997	00:01:41	83,03%	54,89%
/infraestructura-y-tecnologia/	3.344	2.833	00:00:43	80,16%	26,41%
/mail	2.858	2.505	00:04:07	80,56%	80,23%

Página					
Página	Páginas Vistas	Páginas únicas Vistas	Promedio tiempo en la página	% de Rebote	% de Salidas
/	23.579	19.027	00:01:21	47,11%	44,92%
/quienes-somos/directorio-de-personas/academicos	4.594	2.297	00:00:42	44,15%	14,13%
/estudia-con-nosotros	3.058	2.640	00:00:23	50,00%	13,47%
/mail	2.858	2.505	00:04:07	80,56%	80,23%
/quienes-somos/presentacion-del-departamento	2.713	2.184	00:00:30	31,34%	9,29%
/estudiantes	2.196	1.908	00:00:20	26,44%	11,57%
/estudia-con-nosotros/pregrado	1.918	1.648	00:00:45	76,03%	32,85%
/estudiantes/pregrado/ingeniería-civil-informatica	1.857	1.570	00:02:14	77,23%	65,86%
/quienes-somos/la-ingeniería-informatica-como-disciplina	1.709	1.516	00:02:45	87,08%	76,48%

Figura 4.14: Desglose de contenido.
Fuente: Elaboración Propia.

CAPÍTULO 4. CONSTRUCCIÓN E IMPLEMENTACIÓN DE LA SOLUCIÓN

4.4.3. Dashboard Origen

El dashboard Origen contiene información relacionada al origen por el cual llegan las visitas a la web. Se observan dos gráficos, el primero con la cantidad de usuarios que vienen por los distintos canales, el segundo, con la cantidad de usuarios por día.

Por último, se presenta información más detallada sobre los canales de ingreso a la web como: cantidad de usuarios, usuarios nuevos, sesiones, páginas vistas, páginas por sesión, % de rebote y promedio de la sesión. Todo esto ordenado de manera descendente por cantidad de usuarios (figura 4.15).

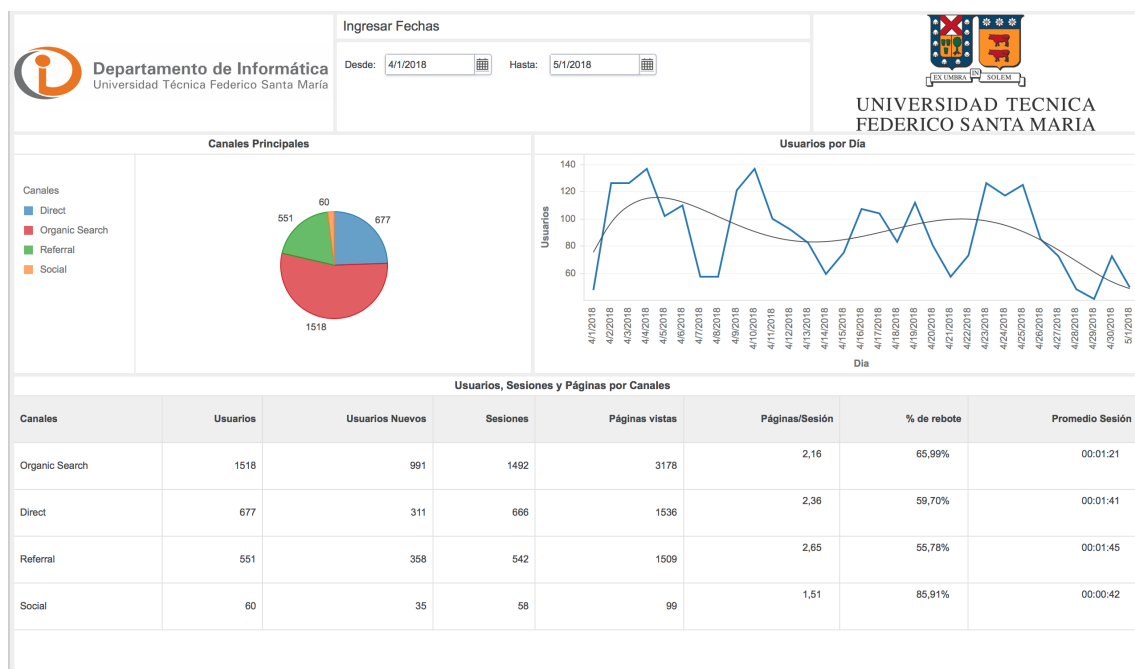


Figura 4.15: Dashboard Origen.
Fuente: Elaboración Propia

Capítulo 5

Análisis de Resultados

Para validar que la solución propuesta en esta memoria permite obtener el perfil de los visitantes al sitio web del Departamento de Informática de la UTFSM, se tomó en cuenta los datos de los visitantes a este sitio desde el 1 de junio del 2017 al 31 de mayo del 2018. De tal manera de comprobar si con datos de este año móvil, es posible generar un análisis que permita desarrollar estrategias de marketing digital y mejoras al sitio mismo.

5.1. Origen

5.1.0.1. General

Se puede apreciar en la figura 5.1 que la mayor cantidad de usuarios se concentra por el canal búsquedas orgánicas (organic search) con un 53.26 %, seguido del canal directo, referencia, social y otros. Esto quiere decir que la mayoría de los usuarios entra al sitio a través del buscador. Sin embargo, si nos fijamos en el indicador % de rebote, se nota que un 62.79 % llega y se va del sitio, lo que se traduce en tráfico de mala calidad (no encuentran lo que buscan o cuentan con muy poco tiempo).

El tráfico por referencia posee un 19.82 % del tráfico, además, de los usuarios que llegan un 51.95 % se va de inmediato, el promedio de tiempo por sesión es de 2 minutos aproximadamente y las páginas vistas por sesión son 2.96. Todos estos indicadores se encuentran en el primer lugar comparado con las otras fuentes de tráfico.

CAPÍTULO 5. ANÁLISIS DE RESULTADOS

Usuarios, Sesiones y Páginas por Canales									
Canales	Usuarios	% Peso Usuarios	Usuarios Nuevos	Sesiones	Páginas vistas	Páginas/Sesión	% de rebote	Promedio Sesión	Páginas únicas
Total	24078	100,00%	22351	35343	90381	2,56	59,00%	00:01:42	73561
Organic Search	13125	54,51%	12146	18819	45782	2,42	62,79%	00:01:33	37229
Referral	5497	22,83%	5004	7018	22186	2,96	51,95%	00:02:01	18212
Direct	4739	19,68%	4626	7927	19527	2,52	58,78%	00:01:50	15680
Social	715	2,97%	573	1577	2884	1,87	67,28%	00:01:20	2438
(Other)	2	0,01%	2	2	2	1,00	100,00%	00:00:00	2

Figura 5.1: Indicadores por fuente de tráfico.
Fuente: Elaboración Propia.

5.1.0.2. Fuente/Medio

La información obtenida desde esta dimensión, permite conocer las combinaciones de fuente y medio que traen usuarios al sitio web. En la figura 5.2 se puede observar el top 10 de combinaciones ordenadas por cantidad de usuarios. En primer lugar está la fuente Google como medio orgánico, sin embargo, posee un alto % de rebote. Hay que destacar que la combinación usm.cl/referencia que si bien no atrae la mayor cantidad de usuarios, posee el % de rebote más bajo de todos con un 47.26 % y tiene el mejor promedio de duración media de la sesión (02:12).

Usuarios, Sesiones y Páginas por Fuente/Medio								
Fuente/Medio	Usuarios	% Peso Usuarios	Usuarios Nuevos	Sesiones	Páginas vistas	Páginas/Sesión	% de rebote	Promedio Sesión
google / organic	12.984	54,59%	12.012	18.625	45.315	2,43	63,00%	00:01:33
(direct) / (none)	4.739	19,93%	4.626	7.927	19.527	2,52	58,78%	00:01:50
usm.cl / referral	4.427	18,61%	4.132	5.474	18.792	3,31	46,00%	00:02:12
progra.usm.cl / referral	248	1,04%	238	261	392	1,48	78,54%	00:00:47
facebook.com / referral	220	0,93%	135	594	1.316	2,24	47,98%	00:01:31

Figura 5.2: Indicadores por Fuente/Medio.
Fuente: Elaboración Propia.

CAPÍTULO 5. ANÁLISIS DE RESULTADOS

5.1.0.3. Referencia

Las fuentes de tráfico por referencia se destacan en el sitio web porque tienen la mejor tasa de rebote. En la figura 5.3 se presentan las 10 más importantes según cantidad de usuarios.

La que más se destaca entre todas es la fuente *usm.cl*, ya que atrae la mayor cantidad de tráfico (44.227 usuarios). Además, poseen el mejor % de rebote (46.44 %), la mayor cantidad de páginas por sesión (3.30) y el mejor promedio de tiempo por sesión (2:15) por lo que es la fuente por referencia más importante en la actualidad. Otra fuente destacada es *facebook.com* que si bien no atrae muchos usuarios, tiene indicadores destacados como páginas por sesión (2.34) y promedio de tiempo por sesión (01:47).

A partir de los resultados obtenidos, se puede concluir que se debe mantener y, ojalá, potenciar la fuente *usm.cl* debido a que es la fuente con mejor calidad de usuarios. También, aumentar la frecuencia de noticias y publicaciones en la fuente *facebook.com* con el fin de incrementar la cantidad de visitas que vienen desde ella. En cuanto a la fuente *admission.usm.cl*, debe mejorar la calidad de las visitas ya que, si bien, aporta una cantidad similar a la red social Facebook, tiene peores indicadores de páginas por sesión o % de rebote. Esto se puede lograr, mediante noticias más llamativas en el Home para retener al usuario en el sitio web.

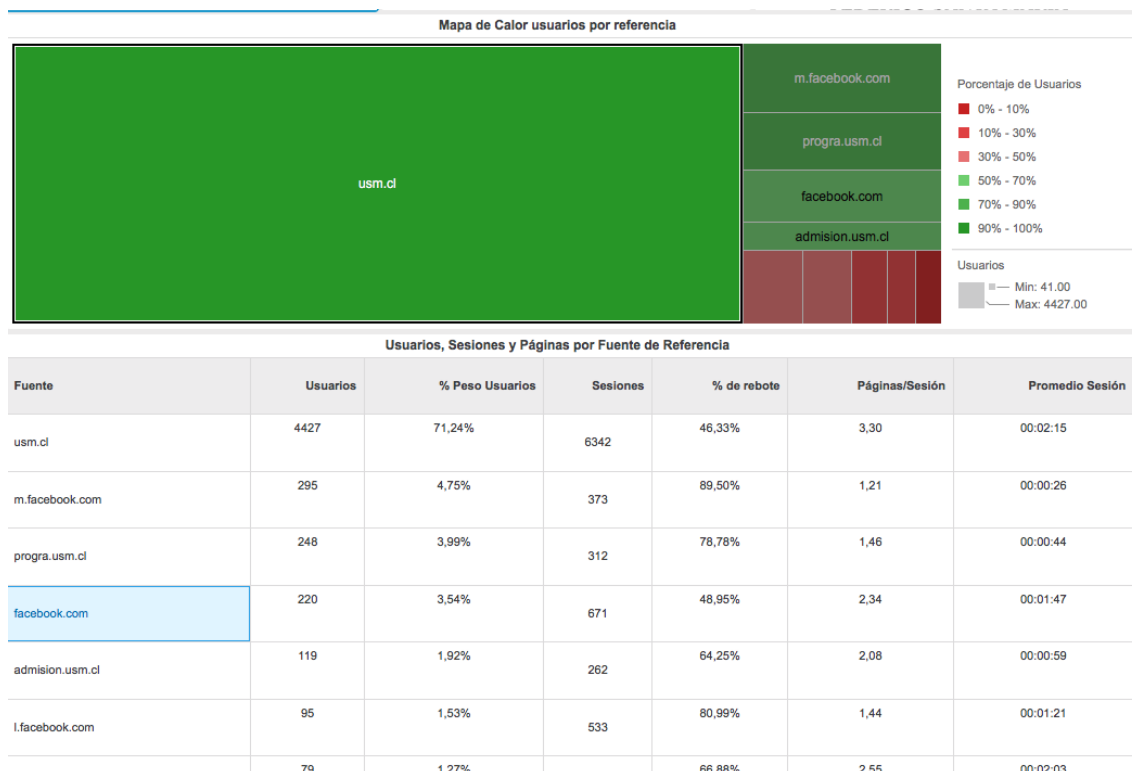


Figura 5.3: Indicadores por Fuentes de referencia.
Fuente: Elaboración Propia.

5.2. Audiencia

5.2.1. General

En esta sección se analizarán los antecedentes más generales sobre las características de la audiencia, éstos pueden ser observados en la figura 5.4.

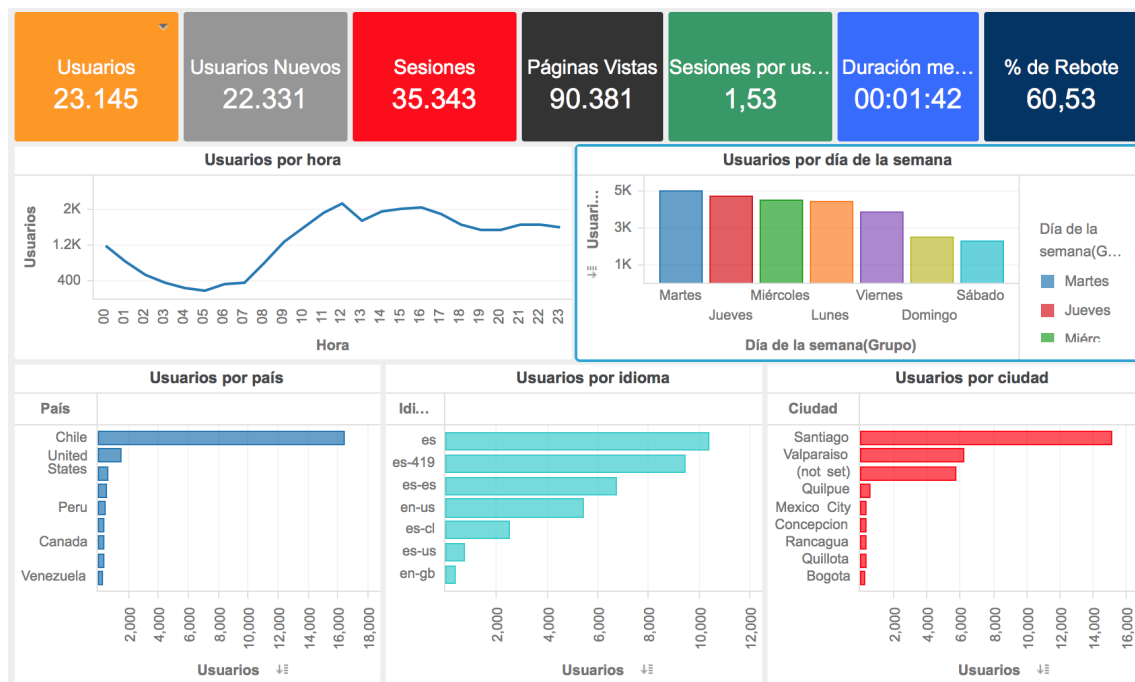


Figura 5.4: Indicadores generales Audiencia.

Fuente: Elaboración Propia.

La cantidad de usuarios del período es de 23.145, los nuevos usuarios 22.331, sesiones 35.343, páginas vistas 90.381, sesiones por usuario 1.53, duración media de la sesión 01:42 y el % de rebote un 60.53.

En cuanto a los usuarios que ingresan por hora, se puede destacar que el rango de horario por el que más acceden al sitio es entre las 11:00 y 17:00 hrs, siendo las 12:00 hrs el horario preferido por los usuarios con 2.095 visitas.

Si se analiza la cantidad de usuarios por día de la semana, se nota que entre los días lunes y jueves se concentra el mayor porcentaje de visitas con un 68.41 %. El día con más visitas es el martes con 5.000 visitas aproximadamente.

Finalmente, los 3 países que traen más tráfico son Chile, México y Estados Unidos. La ciudad con más visitas es Santiago, Valparaíso, Quilpué y Ciudad de México en orden descendente. El idioma de la mayoría

CAPÍTULO 5. ANÁLISIS DE RESULTADOS

de las visitas es el español, seguido del inglés y el portugués.

5.2.2. Datos demográficos

5.2.2.1. Edad

Según se observa en figuras 5.5 y 5.6, entre los 18 y 24 años representan un 31.44 % de las visitas y los usuarios entre los 25 y 34 años un 36.79 %. Entre los dos alcanzan un 68.23 % de las visitas. Además, estos dos rangos etarios se comportan de manera similar como queda demostrado al analizar el % de rebote (60.77 % vs 60.71 %), páginas por sesión (2.62 vs 2.50) y duración media de la sesión (01:35 vs 01:36). El resto posee una menor cantidad de visitas, sin embargo, son usuarios que tienen mejor % de rebote, páginas por sesión y duración media de la sesión.

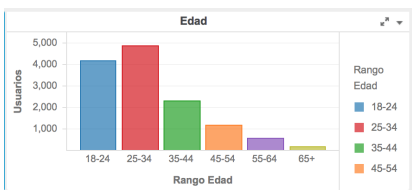


Figura 5.5: Gráfico edad.
Fuente: Elaboración Propia.

Indicadores por rango de edad						
Rango Edad	Usuarios	Usuarios Nuevos	Sesiones	% de Rebote	Páginas por sesión	Duración media de la sesión
25-34	4.885	4.487	6.895	60,77	2,62	00:01:35
18-24	4.174	3.840	5.655	60,71	2,50	00:01:36
35-44	2.322	2.125	3.705	57,88	2,90	00:01:58
45-54	1.159	1.058	1.817	56,75	2,83	00:01:55
55-64	544	490	1.046	52,87	2,57	00:01:41
65+	193	171	369	40,65	3,08	00:02:15

Figura 5.6: Indicadores por Edad.
Fuente: Elaboración Propia.

5.2.2.2. Género

Con la información obtenida por género (figura 5.7), se puede decir que los usuarios de sexo masculino corresponden a un 69.61 % y el sexo femenino a un 30.39 %, por lo que los primeros son los que traen mayor cantidad de tráfico al sitio.

Además, si se analizan los indicadores se obtiene que la tasa de rebote masculina corresponde a un 57.63 %, menor en un 6.21 % que la femenina; las páginas por sesión a 2.74 vs 2.40 de las mujeres; y la duración media de la sesión muestra 8 segundos más en las visitas masculinas.

Por lo tanto, el grupo de hombres entre 25 y 34 años representan a los usuarios de mejor calidad, o sea, los que ven más páginas por sesión, tienen mejor promedio de tiempo por sesión y menor tasa de rebote.

CAPÍTULO 5. ANÁLISIS DE RESULTADOS

Indicadores por género						
Género	Usuarios	Usuarios Nuevos	Sesiones	% de Rebote	Páginas por sesión	Duración media de la sesión
male	9.561	8.918	14.307	57,63	2,74	00:01:44
female	4.174	4.005	6.065	63,84	2,40	00:01:36

Figura 5.7: Indicadores por Sexo.
Fuente: Elaboración Propia.

5.2.3. Intereses

En esta sección se puede encontrar información acerca de las categorías de afinidad, segmentos del mercado y otras categorías de los usuarios (ver figura 5.8).

En relación a categorías de afinidad, se observa que la categoría Videos/Entretención/Amantes de cine es la más destacada, seguida de Tecnología y Compradores. Esta información es muy relevante a la hora de publicar contenido de interés para los usuarios, ya que existe una mayor probabilidad de que se sientan interesados por artículos relacionados con las categorías de mayor afinidad.

En segmentos del mercado, se observa que Educación/Educación Universitaria posee el mayor % de usuarios con un 4.06 %. Le sigue Software/Negocios y producción de Software con un 3.32 %. En tercer lugar está empleabilidad con un 3.26 %. También es una información de suma importancia, no sólo para conocer a la audiencia, sino para poder segmentar de la mejor manera el mensaje que se quiere transmitir.

En lo que respecta a información relacionada con otras categorías, se puede encontrar un mayor detalle que con respecto a categoría de afinidad. En primer lugar se encuentra Arte y Entretenimiento/TV y Video/-Videos en Línea con un 3.28 %, en segundo lugar Deportes/Equipos Deportivos/Fútbol con un 3.26 % y en tercer lugar Referencia/Referencias Generales/Diccionarios y Enciclopedias con un 2.09 %.

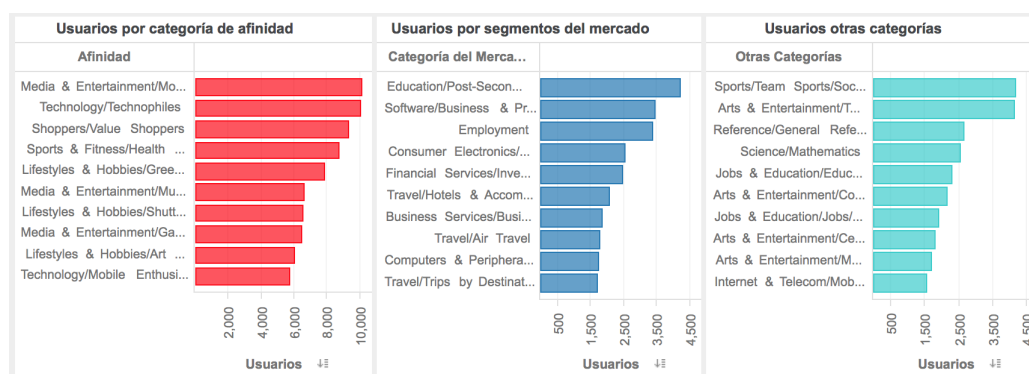


Figura 5.8: Gráfico por intereses del usuario.
Fuente: Elaboración Propia.

5.2.4. Información Geográfica

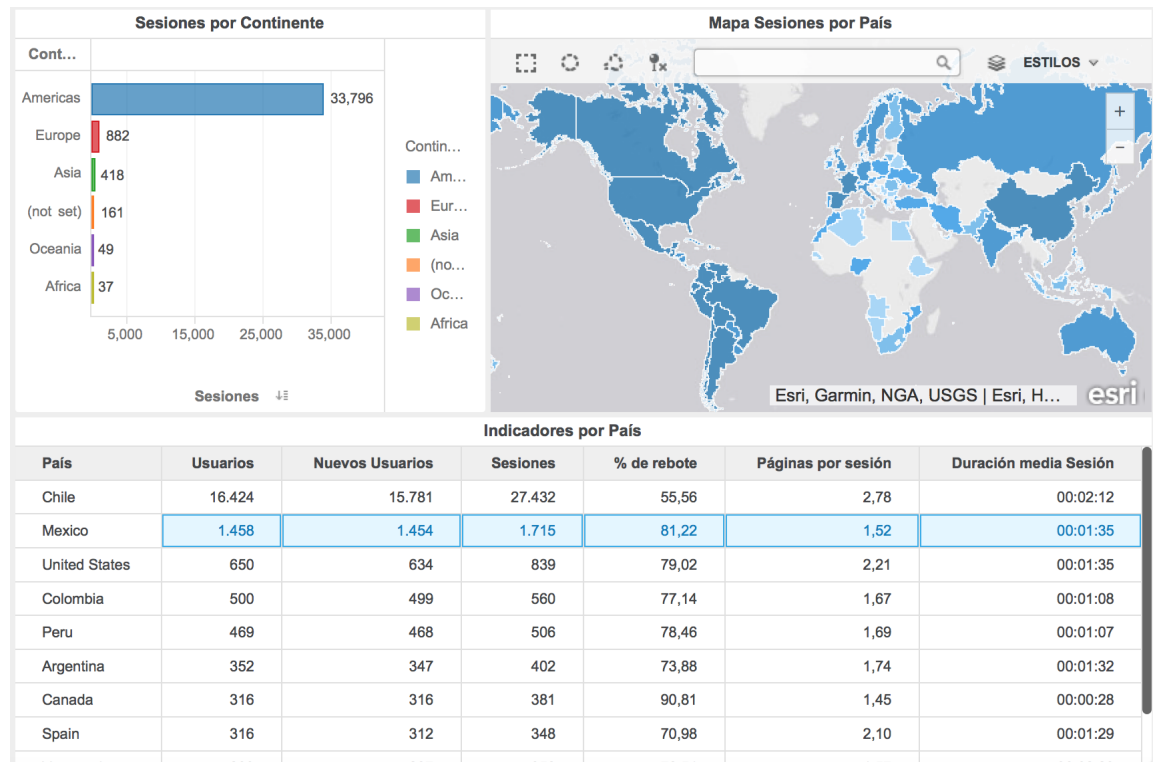


Figura 5.9: Indicadores geográficos.

Fuente: Elaboración Propia.

Según se observa en figura 5.9, y como era de esperarse, la mayor cantidad de sesiones viene del continente Americano (33.796). Le sigue Europa con 882 sesiones, lo cual es una diferencia muy grande, por lo que nos centraremos en América.

Los países que tienen mayor cantidad de usuarios son Chile (16.424), México (1.458) y Estados Unidos (650). De los tres países antes mencionados, Chile es el que posee el menor % de rebote con 55.56 %, la mayor cantidad de páginas por sesión con 2.78 y la mayor duración media de sesión con 2:12 es Chile, por lo cual se analizará con mayor detalle sus ciudades.

Las regiones que más aportan visitas al sitio web son Región Metropolitana (9.826), Región de Valparaíso (5.035), y en menor medida, la Región del Bío Bío (482). En relación a las ciudades (figura 5.10), las que más aportan son Santiago (9.823), Valparaíso (2.621) y Quilpué (338). Si se analiza el % de rebote, Villa Alemana aparece en primer lugar, en segundo lugar Valparaíso y en tercer lugar Quilpué. En cuanto a las páginas por sesión, Villa Alemana nuevamente es la mejor ubicada con 3.43, en segundo lugar aparece Temuco con 3.32 y en tercer lugar Quilpué con 3.06. Si se analiza la duración media por sesión, Villa Ale-

CAPÍTULO 5. ANÁLISIS DE RESULTADOS

mana nuevamente aparece en primer lugar con 02:30, en segundo lugar Temuco y en tercer lugar Quilpué y Valparaíso.

A pesar de que Santiago trae la mayor cantidad de tráfico, hay ciudades como Villa Alemana y Quilpué que aportan con mayor calidad de tráfico.

Indicadores por Ciudad						
Ciudad	Usuarios	Nuevos Usuarios	Sesiones	% de rebote	Páginas por sesión	duración media sesión
Santiago	9.823	9.341	14.094	56,50	2,80	00:01:46
Valparaíso	2.621	2.229	6.379	53,24	2,67	00:02:00
Quilpué	338	293	515	53,98	3,06	00:02:01
Concepción	250	247	289	63,32	2,66	00:01:45
Rancagua	243	223	293	59,73	2,74	00:01:50
Quillota	208	192	285	59,30	2,39	00:01:44
Temuco	143	136	188	54,79	3,32	00:02:27
Villa Alemana	133	106	188	48,94	3,43	00:02:30
La Serena	132	127	169	56,80	2,91	00:01:49
Talca	103	97	120	58,33	2,09	00:01:18

Figura 5.10: Indicadores por Ciudad.
Fuente: Elaboración Propia.

5.2.5. Tipo de visitante, Frecuencia y Asiduidad

5.2.5.1. Tipo de visitante

Los tipos de visitantes posibles son recurrentes o nuevos. Como se puede apreciar en la figura 5.12, el 84.02 % corresponde a usuarios que vienen por primera vez al sitio web, el 15.08 % son usuarios recurrentes. Analizando los indicadores por tipo de usuario (figura 5.12), se puede detectar que los usuarios más fidelizados (recurrentes) tienen mejores valores. El % de rebote de los recurrentes es de un 53.23 % vs un 64.77 % de los nuevos, las páginas por sesión son 2.84 para los recurrentes y 2.39 para nuevos. Finalmente, la duración media por sesión es de 02:16 para recurrentes y 01:22 para nuevos. Claramente, se puede notar que la calidad de los usuarios recurrentes es mucho mejor que la de los usuarios nuevos.

Basado en la información de la sección 2.2.6.6, se puede señalar que el % de usuarios nuevos es mucho mayor al de usuarios recurrentes, lo que permite concluir que la estrategia de marketing seguida hasta el momento ha sido más de captación masiva de clientes que de retención.

Además, el ratio entre usuarios nuevos vs usuarios recurrentes es de 5.63, lo que indica según la literatura [5], que es un sitio web basado en ventas (ratio mayor a 2). Un ratio menor a 2 indica un sitio web que

CAPÍTULO 5. ANÁLISIS DE RESULTADOS



Figura 5.11: Gráfico Tipo de Usuario.
Fuente: Elaboración Propia.

Tipo de usuario						
Tipo Usuario	Usuarios	Usuarios Nuevos	Sesiones	% de Rebote	Páginas por sesión	Duración media de la sesión
New Visitor	22.603	22.351	22.351	64,77	2,39	00:01:22
Returning Visitor	4.014	0	12.992	53,23	2,84	00:02:16

Figura 5.12: Indicadores por Tipo de Usuario.
Fuente: Elaboración Propia.

fidelize mucho al visitante (sitio web de contenido o soporte) o que no sabe atraer visitantes con asiduidad.

5.2.5.2. Frecuencia

Para medir la frecuencia (figura 5.13) se utiliza la dimensión número de sesiones, ésta indica la cantidad de sesiones que ha realizado el usuario en la web. Con esto se puede evaluar el porcentaje de nuevos usuarios en el sitio y el nivel de familiaridad de todos los usuarios que regresan.

La mayoría de los usuarios han ingresado hasta 3 veces al sitio web durante un período específico. Los usuarios con una sola sesión son usuarios nuevos, el resto son usuarios que vuelven al sitio web. Con esta información se concluye que la mayoría de las visitas al sitio son de nuevos usuarios, por lo que el nivel de fidelidad no es muy alto.

Sin embargo, para mejorar esta medición lo ideal sería segmentarla con algún acontecimiento importante (conversión de algún objetivo). Esto permitiría saber el número de visitas que un usuario tiene que realizar antes de la conversión. Si en promedio la conversión toma 5 visitas, se pueden tomar medidas para que un usuario con 3 o 4 visitas vuelva al sitio web. Como por ejemplo, el envío de un correo electrónico para que vuelva.

Es importante tener objetivos en el sitio web para así definir acciones o conversiones y poder mejorar la experiencia de los usuarios.

5.2.5.3. Asiduidad

Para medir asiduidad o recencia, se utilizan los días desde la última sesión. Con esto se puede saber la cantidad de días que han transcurrido desde la última visita de cada usuario. Este indicador permite ver la cantidad media de tiempo transcurrido entre las visitas para la base de usuarios del sitio.

De la figura 5.14, se puede decir que la mayoría de los usuarios vuelve en el rango de 1 y 2 días después

CAPÍTULO 5. ANÁLISIS DE RESULTADOS

de visitar el sitio web. Existe una tendencia a la baja hasta el día 5 y un incremento de sesiones al día 6. Después de esto, sigue la tendencia a la baja.

Se puede decir que a los 6 días desde la primera sesión, los usuarios están volviendo al sitio web. Un posible análisis sería combinar la información de frecuencia con recencia. La mayoría de los usuarios visita el sitio web hasta 3 veces, por lo que se podría crear un segmento para ese grupo y ver si esas visitas se producen dentro del período de 0-1 días o en el transcurso de los días siguientes.

Frecuencia		
Número de Sesiones(Grupo)	Sesiones	Páginas Vistas
1	22.351	53.511
2	3.702	11.048
3	1.546	5.118
4	892	2.652
5	622	2.144
6	451	1.498
7	347	1.004
8	283	852
Rango 09-14	1.144	3.200
Rango 15-25	952	2.684
Rango 26 - 100+	3.053	6.670

Figura 5.13: Indicadores de Frecuencia.
Fuente: Elaboración Propia.

Asiduidad		
Días desde la ultima sesión(Grupo)	Sesiones	Páginas Vistas
0	28.133	68.205
1	1.166	2.970
2	675	1.742
3	428	1.173
4	331	915
5	281	946
6	299	1.093
7	210	682
Rango 08-14	838	2.737
Rango 15-30	1.046	3.527
Rango 31-60	880	2.844
Todos los demás elementos	1.056	3.547

Figura 5.14: Indicadores de Asiduidad.
Fuente: Elaboración Propia.

5.2.6. Tecnología y Dispositivos

5.2.6.1. Tecnología

Como se puede observar en la figura 5.15, el navegador Chrome posee un 73.80 % de usuarios que ingresan por el sitio, Firefox un 10.27 % y Safari un 8.79 %, entre los tres abarcan el 92.86 % del total de visitas. Con esto se puede concluir que el sitio debe funcionar perfectamente y estar correctamente diseñado para ser usado con Chrome, Firefox y Safari.

Esto es muy importante a la hora del diseñar un sitio web, ya que la visualización puede variar mucho de un navegador a otro. Gracias a esta información se puede saber en cuántos navegadores como mínimo, debe funcionar perfectamente el diseño del sitio web.

Además, hay que destacar que el navegador que tiene mejor rendimiento en cuanto al comportamiento de los usuarios es Firefox. Posee el menor % de rebote (58,41 %), mayor cantidad de páginas por sesión (2.73) y mayor duración media de la sesión (02:31).

CAPÍTULO 5. ANÁLISIS DE RESULTADOS

Navegador						
Navegador	Usuarios	Usuarios Nuevos	Sesiones	% de rebote	Duración media de la sesión	Páginas por sesión
Chrome	16.967	16.476	25.481	59,82	00:01:36	2,55
Firefox	2.360	2.289	4.583	58,41	00:02:39	2,73
Safari	2.021	1.966	3.084	63,13	00:01:20	2,59
Edge	349	342	554	62,82	00:01:41	2,39
Internet Explorer	292	291	336	62,50	00:01:51	2,98
Opera	285	277	477	60,17	00:01:41	2,62
Android Webview	247	243	274	83,58	00:00:41	1,41
Mozilla Compatible Agent	125	125	125	100,00	00:00:00	1,00
Samsung Internet	121	120	164	66,46	00:01:08	3,01
Safari (in-app)	102	102	113	91,15	00:00:10	1,19

Figura 5.15: Indicadores de Navegadores.
Fuente: Elaboración Propia.

5.2.6.2. Dispositivos

En las figuras 5.16 y 5.17, se puede detectar que la cantidad de usuarios que ingresa por computador es de 17.148 (74.37 %), por celular 5.608 (24.32 %) y por tablet 301 (1.31 %). El dispositivo más destacado es el computador, ya que además de atraer la mayor cantidad de usuarios, posee un 58.52 % de rebote, 2.66 páginas por sesión y 01:49 de duración media de la sesión. El tablet posee mejores resultados en páginas por sesión y duración media de la sesión, pero la cantidad de usuarios que trae es muy baja comparada al computador.

El móvil tiene un % de rebote de 67.53 %, mucho más alto al promedio del sitio (60.53 %) y también una duración media de sesión más baja comparada al promedio (01:15 vs 01:42). Es necesario revisar por qué sucede esto, ya que claramente está bajando los promedios generales del sitio web.

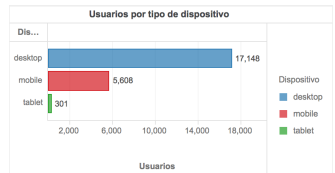


Figura 5.16: Tipo Dispositivo.
Fuente: Elaboración Propia.

Dispositivo	Usuarios	Usuarios Nuevos	Sesiones	% de rebote	Páginas por sesión	Duración media de la sesión
desktop	17.148	16.537	27103	58,52	2,66	00:01:49
mobile	5.608	5.523	7838	67,53	2,14	00:01:15
tablet	301	291	402	58,96	3,58	00:02:09

Figura 5.17: Indicadores por tipo de dispositivo.
Fuente: Elaboración Propia.

CAPÍTULO 5. ANÁLISIS DE RESULTADOS

5.2.6.3. Dispositivos Móviles

Al analizar las marcas de los dispositivos móviles por los que ingresan las visitas (5.18), se obtiene que entre la marca Samsung y Apple aportan un 56.39 % de los usuarios. Le siguen Huawei con un 10.25 % y Motorola con un 9.44 %. Con esta información se puede dar mantenimiento al sitio con la intención que se visualice mejor en el modelo y marca del dispositivo y así otorgar una mejor experiencia para el usuario.

Marca Dispositivo						
Marca Dispositivo	Usuarios	Usuarios Nuevos	Sesiones	% de rebote	(Páginas Vistas/Sesiones)	Duración media de la sesión
Samsung	1.753	1.726	2.569	62,79	2,46	00:01:33
Apple	1.581	1.562	2.103	70,14	2,13	00:01:04
Huawei	606	596	849	67,49	2,01	00:01:20
Motorola	558	535	807	71,13	1,89	00:00:52
LG	336	334	478	66,95	2,44	00:01:29
(not set)	251	247	317	70,98	1,87	00:01:02

Figura 5.18: Indicadores Marca dispositivo.
Fuente: Elaboración Propia.

Al analizar la información sobre los modelos de dispositivos móviles, se obtiene que el dispositivo móvil que tiene más usuarios es Apple Iphone con un 22.99 %, le sigue Motorola Moto G4 con 2.60 %, Apple Ipad con 2.50 % y Samsung J7 Prime con 2.40 %.

En relación al sistema operativo de los dispositivos, hay una clara tendencia para Android con un 76.66 % de las visitas y en segundo lugar se posiciona iOS con 26.76 %.

5.3. Comportamiento

5.3.1. General

A nivel general la información obtenida para el período es la siguiente:

- Número de páginas vistas: 90.381
- Número de visitas únicas a páginas: 73.561
- Promedio de tiempo en página 01:05
- Porcentaje de rebote: 60.53 %
- Porcentaje de salidas: 39.10 %

CAPÍTULO 5. ANÁLISIS DE RESULTADOS

La cantidad de páginas vistas por hora, mantiene la misma proporción que la de usuarios por hora vista en la sección 5.2. Entre las 11:00 y 17:00 hrs, se concentra la mayor cantidad de páginas vistas. Si se analiza la cantidad de páginas vistas por día de la semana, se ve que la mayoría accede al sitio de lunes a jueves.

En la figura 5.19 se puede observar la cantidad de visitas por página y por título de la página o secciones del sitio web.

Título de página			Página		
Título Página	Páginas Vistas	% Páginas Vistas	Página	Páginas Vistas	% Páginas Vistas
Departamento de Informática USM	24.065	26,63%	/	23.579	26,09%
Académicos	4.582	5,07%	/quienes-somos/directorio-de-personas/academicos	4.594	5,08%
Pregrado	3.804	4,21%	/estudia-con-nosotros	3.058	3,38%
Estudia con nosotros	3.066	3,39%	/mail	2.858	3,16%
Mail	2.852	3,16%	/quienes-somos/presentacion-del-departamento	2.713	3,00%
Presentación del departamento	2.699	2,99%	/estudiantes	2.196	2,43%
Estudiantes	2.206	2,44%	/estudia-con-nosotros/pregrado	1.918	2,12%
Ingeniería Civil Informática	1.873	2,07%	/estudiantes/pregrado/ingenieria-civil-	1.857	2,05%
La ingeniería informática como disciplina	1.693	1,87%			

Figura 5.19: Indicadores por Página y Título de la página.

Fuente: Elaboración Propia.

Para el primer caso, las páginas más vistas son:

1. Home con 26.09 %
2. /quienes-somos/directorio-de-personas/academicos con 5.08 %
3. /estudia-con-nosotros con 3.38 %
4. /mail con 3.16 %
5. /quienes-somos/presentacion-del-departamento con 3.00 %

La información por secciones, indica que las más destacadas en relación a la cantidad de visitas son:

1. Departamento de Informática USM
2. Académicos
3. Pregrado
4. Estudia con Nosotros
5. Mail

5.3.2. Páginas de entrada

El análisis de páginas de aterrizaje o de entrada sirve para conocer cómo se comportan las visitas al momento de llegar al sitio web y en cómo afectan estas páginas de entrada a la navegación de los usuarios.

En cuanto a número de sesiones, claramente la página *home* es la que más sesiones tiene con un 50.95 % del total, en segundo y tercer lugar le siguen */quienes-somos/la-ingenieria-informatica-como-disciplina* y */vinculacion/programa-de-continuidad-de-estudios* respectivamente.

Las que más aportan nuevos visitantes en comparación al total son: */quienes-somos/la-ingenieria-informatica-como-disciplina*, */vinculacion/programa-de-continuidad-de-estudios* respectivamente y */quienes-somos/organizacion/estructura-organizacional*. En teoría, si las páginas tienen un número elevado de nuevos visitantes en comparación a visitantes recurrentes, deben ser optimizadas para aumentar las visitas de usuarios que ya han ingresado al sitio y disminuir las visitas que nunca han ingresado. Sin embargo, estas tres, tienen resultados dentro de lo esperado, ya que contienen información enfocada hacia futuros estudiantes o postulantes, y en menor medida, a estudiantes que ya pertenecen a la carrera, por lo que se espera una mayor cantidad de nuevos visitantes que visitas recurrentes.

Los menores porcentajes de nuevas sesiones, los tienen la página */correo*, */academicos*, */applications* y la *home*. Para el correo es de un 35.53 %, lo que tiene sentido, ya que es una página destinada a estudiantes y, generalmente, no es la primera vez que ingresan a ella. Las restantes superan el 50 %, por lo que la mayoría de visitantes son nuevos.

Revisar la duración media de la sesión y páginas por sesión, sirve para ver si la página de entrada funcionó lo suficientemente bien para que el usuario interactúe con la web y navegue en ella. Los mejores resultados lo tienen la *home*, */academicos* y */dr-horst-von-brand*, ya que superan las 2 páginas por sesión. En cuanto a la duración media de la sesión se destacan */academicos* (02:53) y */home* (02:10).

Si se revisan los porcentajes de rebote de estas páginas, es posible destacar a dos: */academicos* con 44.15 % y */home* con un 47.11 %. Estas páginas tienen resultados bajo el promedio general del sitio y permiten que los usuarios sigan navegando y no se vayan al instante. Las que tienen los peores porcentajes de rebote son: */la-ingenieria-informatica-como-disciplina*, */estructura-organizacional*, */applications* y */correo*. Todas tienen sobre un 80 % de rebote, por lo que deben optimizarse, a excepción de */correo* debido a que posee link's a los sitios de correo de profesores y alumnos, para mejorar la experiencia del usuario.

CAPÍTULO 5. ANÁLISIS DE RESULTADOS

Páginas de destino						
Página de destino	Sesiones	% Nuevas Sesiones	Usuarios Nuevos	% de Rebote	Páginas/sesión	duración media sesión
/	18.362	58,22%	10690	47,11%	3,20	00:02:10
/quienes-somos/la-ingenieria-informatica-como-disciplina	1.300	91,08%	1184	87,08%	1,22	00:00:48
/vinculacion/programa-de-continuidad-de-estudios	1.187	91,07%	1081	74,73%	1,56	00:00:48
/estudiantes/pregrado/ingenieria-civil-informatica	1.063	63,12%	671	77,23%	1,70	00:01:05
/mail	957	35,53%	340	80,56%	1,37	00:01:01
/applications	787	57,56%	453	81,07%	1,65	00:00:59
/quienes-somos/directorio-de-personas/academicos	299	52,51%	157	44,15%	4,20	00:02:53
/quienes-somos/organizacion/estructura-organizacional	265	88,30%	234	84,53%	1,41	00:00:42
/quienes-somos/directorio-de-personas/academicos/5-dr-horst-von-brand	252	76,98%	194	75,40%	2,12	00:00:56
/noticias/194-fallecimiento-ex-alumna-constanza-matamoros-avila	214	73,83%	158	78,97%	1,68	00:01:26
/estudia-con-nosotros/pregrado/descripcion-de-la-carrera	211	61,14%	129	74,88%	2,06	00:01:38

Figura 5.20: Indicadores por página de entrada.
Fuente: Elaboración Propia.

5.3.3. Páginas de salida

En la figura 5.21 se puede observar el top 10 de páginas ordenadas por las veces en que son salida en el sitio web. Es normal que la página /home sea muchas veces la salida de la página, ya que es por donde ingresan las visitas. Le sigue /mail que también está dentro de la normalidad, debido a que contiene enlaces externos a páginas de correos (profesores y alumnos).

Las páginas que tienen el mayor porcentaje de salidas, dejando de lado las nombradas anteriormente y que están en el top 10 de salidas son:

1. /quienes-somos/la-ingenieria-informatica-como-disciplina con 76.48 %
2. /vinculacion/programa-de-continuidad-de-estudios con 72.43 %
3. /estudiantes/pregrado/ingenieria-civil-informatica con 65.86 %
4. /applications con 64.13 %
5. /estudia-con-nosotros/pregrado/descripcion-de-la-carrera con 50 %

Hay que revisar constantemente las páginas con mayor número de salidas y ver si están dentro de los números esperados. Si hay casos en que no debiese ser así, revisar la causa: arquitectura, diseño, contenido, fechas y horas de salida hasta encontrar cuál es el problema.

CAPÍTULO 5. ANÁLISIS DE RESULTADOS

Ruta de Página			
Página	Salidas	Páginas Vistas	Porcentaje de Salidas
/	10.592	23.579	44,92%
/mail	2.293	2.858	80,23%
/quienes-somos/la-ingenieria-informatica-como-disciplina	1.307	1.709	76,48%
/estudiantes/pregrado/ingenieria-civil-informatica	1.223	1.857	65,86%
/vinculacion/programa-de-continuidad-de-estudios	1.072	1.480	72,43%
/applications	860	1.341	64,13%
/quienes-somos/directorio-de-personas/academicos	649	4.594	14,13%
/estudia-con-nosotros/pregrado	630	1.918	32,85%
/estudia-con-nosotros/pregrado/descripcion-de-la-carrera	564	1.128	50,00%
/estudia-con-nosotros/postgrado	416	931	44,68%
/estudia-con-nosotros	412	3.058	13,47%
/estudia-con-nosotros/postgrado/mti	316	488	64,75%

Figura 5.21: Indicadores por Página de Salida.
Fuente: Elaboración Propia.

5.3.4. Desglose de contenido

La información obtenida sobre el contenido que el usuario visita cuando entra al sitio web se puede observar en la figura 5.22.

La vista por nivel de ruta de página 1, muestra los indicadores por secciones del sitio web. Los más destacados son /home, /quienes-somos/, /estudia-con-nosotros y /estudiantes. Entre ellas concentran el 70.92 % de las páginas vistas en el sitio web.

La sección que tiene el mejor promedio de tiempo en la página es /mail, esto se puede explicar debido a que en esta página existen links hacia otros sitios de correos (profesores y alumnos), por lo que al abrirse la nueva pestaña sigue pasando el tiempo y es probable que los usuarios no cierren la página actual, aumentando la estadía de tiempo en el sitio. En segundo lugar, se puede ver a /applications/ con 02:07. Esto es de esperarse al tratarse de una página informativa para un puesto de trabajo en la Universidad y los usuarios deben estar cierto tiempo para leer la oferta de trabajo. Le sigue /noticias/ con 01:41, el /home con 01:21, /vinculacion/ con 01:39 y /estudiantes/ con 00:57.

En cuanto al porcentaje de rebote, la mayoría tiene una alta tasa de rebote a excepción de la /home. Los restantes están por sobre el 70 % y las más altas son las siguientes:

1. /noticias/ con 83.56 %
2. /applications con 81.07 %
3. /mail con 80.56 %

CAPÍTULO 5. ANÁLISIS DE RESULTADOS

Nivel de ruta de página 1					
Nivel de ruta de página 1	Páginas Vistas	Páginas únicas Vistas	Promedio tiempo en la página	% de Rebote	% de Salidas
/	23.579	19.027	00:01:21	47,11%	44,92%
/quienes-somos/	23.417	18.145	00:00:53	72,46%	30,17%
/estudiantes/	5.937	5.128	00:01:19	76,12%	44,58%
/estudia-con-nosotros/	5.915	4.919	00:01:14	73,23%	40,29%
/investigacion/	3.949	2.983	00:00:37	71,97%	19,47%
/noticias/	3.418	2.997	00:01:41	83,03%	54,89%
/infraestructura-y-tecnologia/	3.344	2.833	00:00:43	80,16%	26,41%
/estudia-con-nosotros	3.058	2.640	00:00:23	50,00%	13,47%
Página					
Página	Páginas Vistas	Páginas únicas Vistas	Promedio tiempo en la página	% de Rebote	% de Salidas
/	23.579	19.027	00:01:21	47,11%	44,92%
/quienes-somos/directorio-de-personas/academicos	4.594	2.297	00:00:42	44,15%	14,13%
/estudia-con-nosotros	3.058	2.640	00:00:23	50,00%	13,47%
/mail	2.858	2.505	00:04:07	80,56%	80,23%
/quienes-somos/presentacion-del-departamento	2.713	2.184	00:00:30	31,34%	9,29%
/estudiantes	2.196	1.908	00:00:20	26,44%	11,57%
/estudia-con-nosotros/pregrado	1.918	1.648	00:00:45	76,03%	32,85%
/estudiantes/pregrado/ingenieria-civil-informatica	1.857	1.570	00:02:14	77,23%	65,86%
/quienes-somos/la-ingenieria-informatica-como-disciplina	1.709	1.516	00:02:45	87,08%	76,48%
/vinculacion/programa-de-continuidad-de-estudios	1.490	1.208	00:01:54	74,73%	72,42%

Figura 5.22: Top 10 nivel 1 de ruta de página y Página.

Fuente: Elaboración Propia.

4. */infraestructura-y-tecnologia* con 80.16 %

5. */vinculacion* con 75 %

Los porcentajes de salidas más bajos son para */investigacion/* (19.47 %), */infraestructura-y-tecnologia/* (26.41 %), */quienes-somos/* (30.17 %), */estudia-con-nosotros/* (31.15 %) y */estudiantes/* (35.67 %). Esto quiere decir que generalmente éstas no son páginas por las que sale el usuario una vez que empieza a navegar por el sitio web.

Por otro lado, los peores porcentajes de salida son para */mail/* (80.23 %), */vinculacion/* (68.26 %) y */applications/* (64.13 %). El primer caso como se explicó anteriormente es normal. Pero para los otros dos es necesario revisar si corresponde que normalmente sean páginas por las que debe salir la visita.

Para la dimensión página, se puede analizar con mayor detalle las URL del sitio web. En la figura 5.22, se pueden ver las URL con mayor cantidad de páginas vistas. En la vista anterior (nivel de ruta de página 1) se señaló que una de las secciones con mayor visitas era */quienes-somos/*. Acá podemos observar que los contenidos que más aportan a esta sección son */academicos* con 4.594 páginas vistas, */presentacion-*

CAPÍTULO 5. ANÁLISIS DE RESULTADOS

del-departamento con 2.713 páginas vistas y /la-ingenieria-informatica-como-disciplina con 1.709 páginas vistas.

El top 3 de páginas con mayor **promedio de tiempo en la página**, ignorando /mail y /home, son las páginas:

1. /quienes-somos/la-ingenieria-informatica-como-disciplina con 02:45
2. /estudiantes/pregrado/ingenieria-civil-informatica con 02:14
3. /vinculacion/programa-de-continuidad-de-estudios con 01:54

Todas ellas tienen un promedio por sobre el general del sitio web (01:05)

Por otro lado, el top 3 de páginas con menor **promedio de tiempo en la página** y que deben ser evaluadas en caso que no estén dentro de lo esperado es:

1. /estudiantes con 00:20
2. /estudia-con-nosotros con 00:23
3. /quienes-somos/presentacion-del-departamento con 00:30

Las páginas que tienen el **mayor porcentaje de rebote** son /quienes-somos/la-ingenieria-informatica-como-disciplina con 87.08 %, /estudiantes/pregrado/ingenieria-civil-informatica con 77.23 % y /estudia-con-nosotros/pregrado con 76.03 %. Además, las dos primeras tienen un alto **porcentaje de salidas** con 76.48 % y 65.86 % respectivamente. La tercera, posee una tasa de 32.85 % por lo que concluimos que si bien la mayoría de las veces se van del sitio, cuando se quedan ésta no es principalmente de salida y probablemente siguen navegando en el sitio. En cambio, las dos primeras tienen un alto porcentaje de rebote y salida por lo que frecuentemente son páginas por las que la visita sale. Si se suma a esto el alto promedio de tiempo en la página, se obtiene que a pesar de ser páginas de salida, el usuario de queda una cantidad de tiempo considerable en el sitio antes de salir.

Las páginas con **bajo porcentaje de rebote y salidas** son /estudiantes y /quienes-somos/presentacion del departamento. Esto quiere decir que usualmente las visitas no salen por éstas páginas, pero siguen navegando debido a su bajo % de salidas.

Se puede añadir que es necesaria la revisión de estas páginas debido a su **bajo promedio de tiempo**, ya que probablemente el usuario no está encontrando la información que busca. En consecuencia, se dirigen hacia otras páginas del sitio.

5.4. Síntesis del Análisis de Resultados

Los resultados más relevantes del análisis, considerando los datos de visitantes en el período 1 de junio del 2017 al 31 de mayo del 2018, son los siguientes:

- Cantidad de usuarios es 23.145, usuarios nuevos 22.331, sesiones 35.343, páginas vistas 90.381, sesiones por usuario 1.53, % de rebote 60.53 %, páginas únicas vistas 73.561 y promedio de tiempo en página 01:05.
- El día en que más se reciben visitas es el día Martes, además, la mayoría de las visitas se concentra entre Lunes y Jueves. El rango de horario con mayor visitas es entre las 12:00 y 17:00 hrs, siendo las 12:00 hrs el horario preferido.
- Los países con mayor tráfico son Chile, México y Estados Unidos. Las ciudades con más visitas son Santiago, Valparaíso y Quilpué.
- La mayoría de las visitas está en el rango de los 18 y 34 años. El género masculino corresponde a un 70 % de las visitas y el femenino el 30 % restante.
- En relación a los tipos de visitantes se obtuvo que el 84 % son usuarios que vienen por primera vez al sitio, el 16 % son usuarios que ya lo habían visitado. Los recurrentes tienen una calidad mucho mejor en comparación a los nuevos, esto se debe a que son usuarios fidelizados y poseen indicadores de % de rebote, páginas por sesión y duración media de la sesión, mejores.
- El navegador más utilizado por las visitas es Chrome con 74 % aproximadamente. Le sigue Firefox y Safari con 10.27 % y 8.79 % respectivamente.
- Las visitas por tipo de dispositivo, mostraron que la mayoría de ellos accede por computador (74 %) y celular (24 %). En menor medida por tablets. La marca de dispositivo móvil más utilizada es Samsung y Apple. Entre las dos aportan un 56 % de los usuarios.
- La mayoría de los usuarios viene por el canal búsquedas orgánicas (56 %), seguido del canal directo y por referencia. La calidad del tráfico por referencia es mejor a la de búsquedas orgánicas en relación a % de rebote, promedio de tiempo por sesión y páginas vistas por sesión. La fuente orgánica que más trae visitas es Google y la fuente por referencia que más aporta es usm.cl. Además, hay que destacar la fuente por referencia facebook.com que, si bien, no trae mucha cantidad de visitas, posee buenos indicadores de páginas por sesión, % de rebote y promedio de tiempo por sesión.

CAPÍTULO 5. ANÁLISIS DE RESULTADOS

- Las páginas más vistas del sitio son: Home, /quienes-somos/directorio-de-personas/academicos, /estudia-con-nosotros, /mail y /quienes-somos/presentacion-del-departamento.
- Las secciones con mayor cantidad de páginas vistas son: Departamento de Informática USM, Académicos, Pregrado, Estudia con Nosotros y Mail.
- Las mejores landing pages (menor % de rebote) son /academicos y Home, las cuales mantienen a los usuarios navegando en el sitio.
- Las páginas que más veces son salida en el sitio web son: /quienes-somos/la-ingenieria-informatica-como-disciplina y vinculacion/programa-de-continuidad-de-estudios.
- Las páginas con mayor promedio de tiempo son: /quienes-somos/la-ingenieria-informatica-como-disciplina con 02:45 y /estudiantes/pregrado/ingenieria-civil-informatica con 02:14.
- Por otro lado, las páginas con peor promedio de tiempo son: /estudiantes con 00:20 y /estudia-con-nosotros con 00:23.
- Finalmente, las páginas con el mayor porcentaje de rebote son: /quienes-somos/la-ingenieria-informatica-como-disciplina con 87.08 %, /estudiantes/pregrado/ingenieria-civil-informatica con 77.23 % y /estudia-con-nosotros/pregrado con 76.03 %.

En base a los resultados mencionados anteriormente, se proponen las siguientes recomendaciones:

Generar metas a partir de los resultados obtenidos, de tal manera de poder establecer un seguimiento del sitio web en el tiempo, partiendo con los valores obtenidos en esta memoria como la situación base a partir de la cual comparar.

Se recomienda publicar noticias en el Home entre los días lunes y jueves, principalmente el día martes. Además, si se combina con el horario de mayor visita, las 12:00, se podría llegar a una mayor cantidad de usuarios en comparación a la situación actual.

El rango de edad de las visitas es entre 18 y 34 años y del género masculino. Por lo que, publicar noticias que llamen la atención de este público objetivo sería lo recomendado. Sin embargo, no se debe dejar de lado al género femenino, pensando en temas de inclusión y en el aumento de ellas a carreras universitarias que se está dando en el último tiempo, por lo que se recomienda la publicación de noticias que incentiven la visita de este género al sitio web.

Existe un 16 % de visitas que son recurrentes y éstas son las que aportan la mejor calidad de tráfico. Una campaña de correos, inversión en publicidad y publicar noticias con mayor frecuencia generan una mayor fidelidad de las visitas.

CAPÍTULO 5. ANÁLISIS DE RESULTADOS

El tráfico aportado por la fuente de referencia Facebook es de alta calidad, por lo que es necesario mejorar la frecuencia de publicaciones en este sitio. Con esto, se incrementaría el volumen de visitas desde esta fuente de referencia.

Hay problemas con las páginas que más veces son salida, por lo que es necesario revisar que está ocurriendo con ellas y así mejorar la retención de los visitantes. Problemas como el diseño o velocidad de carga son comunes en este tipo de páginas. También, es necesario revisar las páginas que tienen el mayor % de rebote, ya que el usuario no está encontrando la información que busca. Esto, puede deberse a que el contenido no es el adecuado en la página o, como se mencionó antes, existen problemas con la navegación (no sabe por dónde seguir) o problemas técnicos como enlaces erróneos.

Capítulo 6

Conclusiones

Para poder concluir sobre este trabajo es importante tener en cuenta que se planteó como objetivo de la memoria, el diseño e implementación de un sistema de BI que permita analizar el tráfico de un portal web universitario, de tal manera de obtener el perfil de los visitantes para orientar la estrategia de marketing digital de una carrera universitaria. Esto aplicado en el contexto del portal web del Departamento de Informática de la UTFSM para su carrera de Ingeniería Civil Informática. A continuación, se analizará el nivel de logro de cada uno de los objetivos por cada uno de los objetivos específicos definidos.

6.1. Cumplimiento de Objetivos

En la sección 1.3 de este trabajo se definieron 5 objetivos específicos, por lo que es importante analizar si se lograron satisfacer de forma correcta:

6.1.1. Identificar qué tipo de información es importante o relevante para obtener el perfil de los visitantes

Este objetivo se cumplió al realizar un estudio previo al diseño, donde se revisó qué información está disponible y es importante para conocer al visitante. De este estudio se obtuvo que conociendo cómo son las visitas en relación al sexo, rango de edad, el día u hora preferido de acceso al sitio, sus intereses o preferencias, la ubicación, el tipo de visitante, la fidelidad de la visita, el dispositivo por el que acceden, desde dónde vienen (fuente de origen) y su comportamiento en el sitio web, es posible generar un perfil del visitante. Sin embargo, existe información que no pudo ser obtenida como las palabras claves por las que

CAPÍTULO 6. CONCLUSIONES

los usuarios buscan el sitio web y las palabras claves utilizadas en el buscador interno del sitio. Para obtener esta información, el webmaster o encargado del sitio web, debe habilitar el seguimiento del buscador interno en Google Analytics. En cuanto a la información de palabras claves por las que los usuarios buscan el sitio, es necesario tener una cuenta en Search Console (gratuita) y habilitarla en Google Analytics.

6.1.2. Establecer los indicadores claves de rendimiento (KPIs) para obtener el perfil de los visitantes de la web

Los indicadores se definieron en base a la información disponible en Google Analytics, con el fin de poder medir la navegación que realizan los visitantes del sitio web. Los indicadores utilizados son: Usuarios, Nuevos Usuarios, Usuarios Recurrentes, Sesiones, Porcentaje de Rebote, Páginas por Sesión, Duración de la Sesión, Cantidad de Páginas Vistas, Tiempo Promedio en la Página, Salidas de una Página, Porcentaje de Salidas de una Página. Estas métricas en conjunto, más las dimensiones explicadas en la sección 3.2, permiten identificar el perfil de los visitantes del sitio web.

6.1.3. Diseñar una solución de Inteligencia de Negocios en base a los indicadores anteriores

El cumplimiento de este punto, se puede observar en el capítulo 3, donde primeramente se identifican los objetivos que darán respuesta a la solución. Posterior a esto se identificó y analizó la fuente de datos para determinar la información necesaria para dar respuesta a los objetivos planteados. Antes de comenzar el diseño de la solución fue muy importante tener conocimiento de los datos a utilizar, su estructura, su configuración, el tipo de datos y formatos. Esto permitió diseñar estructuras que logran integrar todos los datos necesarios en un mismo Data Mart. También se definió el nivel de granularidad de los datos, lo que permitió conocer las dimensiones que fueron utilizadas en la solución. En el diseño del Data Mart fue de gran importancia conocer el nivel de detalle requerido en los datos, para determinar el tamaño aproximado del almacenamiento y evitar una excesiva descripción.

6.1.4. Construir e implementar la solución de Inteligencia de Negocios al portal del Departamento de Informática de la UTFSM

La implementación de la solución tuvo una diferencia en comparación al diseño. La herramienta elegida, MicroStrategy, permite una conexión directa con la base de datos de Google Analytics y mediante consultas personalizadas fue posible implementar el diseño planteado en el capítulo 3, sin la necesidad de crear

CAPÍTULO 6. CONCLUSIONES

los Data Mart diseñados inicialmente. Además, la herramienta incorpora un módulo para previsualizar y modificar los datos a importar, lo que reemplazó el proceso de ETL utilizado normalmente en un sistema de BI. La elección de la herramienta se hizo en base a ciertos criterios como el conocimiento previo de la herramienta, gratuidad y, la más importante, conexión directa a Google Analytics.

6.1.5. Realizar una validación de la propuesta mediante el análisis de los resultados obtenidos utilizando metodología de Analítica Web

El análisis realizado se basó en parte a la metodología de analítica web explicada en la sección 2.2. En ella se plantea un análisis que se divide en 3 partes: origen de la visita, características de la visita y su comportamiento.

Para el origen de las visitas se obtiene que en su mayoría (53 %) acceden a través del buscador de Google.

En cuanto a las características de las visitas, se obtuvo que el promedio por sesión es de 1 minuto y 42 segundos, la tasa de rebote es de 60 % (6 de cada 10 visitas entran al sitio y se van sin realizar alguna acción), las páginas vistas por sesión en promedio son de 1.52, el 70 % corresponden a personas del género masculino, el rango de edad que más visita la web es entre los 25 y 34 años. También, el 85 % de usuarios que acceden al sitio son usuarios nuevos, el día de preferencia de acceso es entre los lunes y jueves, el horario preferido es entre las 11 y 17 horas, la frecuencia de visita es de 3 sesiones por año, el dispositivo preferido por el que acceden es el computador (74 %) y el celular (24 %).

En relación a las páginas del sitio, se obtiene que el promedio de tiempo en página es de 1 minuto y 5 segundos, el porcentaje de salidas es de un 39 %. Las secciones más visitadas, sin tomar en cuenta el home, son: quienes-somos, estudiantes y estudia-con-nosotros. Las páginas más vistas son: Académicos, Pregrado y Estudia con nosotros. La página que actúa más veces como página de destino y, coincidentemente, de salida es la-ingenieria-informatica-como-disciplina (a excepción de la página de inicio).

6.2. Beneficios de la solución

La solución de Inteligencia de Negocios planteada en esta memoria, automatiza directamente todo el procesamiento de datos proveniente de Google Analytics desde la extracción a través de consultas, modificación de los datos en la herramienta y su carga en memoria. Esto permite una reducción en los tiempos de desarrollo y en los gastos asociados a por ejemplo, la contratación de una consultora para la creación de reportes que generen la información requerida para la toma de decisiones por parte del cliente, en este caso

CAPÍTULO 6. CONCLUSIONES

el Departamento de Informática de la UTFSM.

La información obtenida permite al cliente estudiar en detalle el origen, comportamiento y características de los visitantes del sitio web, lo que genera un perfil más preciso de éstos y crear estrategias de marketing digital que permitan mejorar su navegación, la experiencia de usuario con el sitio, atraer nuevos visitantes o identificar las páginas más problemáticas.

El acceso a la información y toma de decisiones se hace más fácil con esta solución debido a que el Departamento de Informática, podrá acceder rápidamente a reportes o dashboards que cuentan con información general y específica. Finalmente, están diseñados de acuerdo a las necesidades del cliente, lo que facilita el entendimiento de éstos.

6.3. Trabajo a futuro

La solución de este trabajo fue construida específicamente para dar soporte a la problemática del Departamento de Informática de la UTFSM de identificar el perfil de los visitantes que acceden al portal web. Sin embargo, puede ser aplicada a otros departamentos de la UTFSM o incluso a la misma Universidad. Además, esta solución puede seguir creciendo al agregarle otras fuentes de información de redes sociales como: Facebook, Twitter y Youtube, con el fin de complementar y mejorar la información hacia el usuario final. Otra información que se puede agregar es la relacionada con los gastos en publicidad en general (si es que se tienen) o gastos de publicidad en redes sociales. Si a esto se le suman los ingresos percibidos por el mismo concepto, es posible obtener el margen en publicidad.

También, para complementar el análisis es necesario definir objetivos para medir actividades dentro de la página web (conversiones). Algunos ejemplos de objetivos son: realizar una compra (ecommerce), descargar un archivo o enviar un formulario de información de contacto. Google Analytics permite esto y para comenzar a medir es necesario que el webmaster o encargado del sitio web configure algunos parámetros y objetivos basados en eventos en el sitio mismo de Google Analytics. Finalmente debe añadir código dentro del sitio web para comenzar a medir estos eventos. Otra información que faltó agregar a la solución, por no estar disponible, y que es importante a la hora de realizar el análisis de los visitantes es: información de campañas de publicidad (anuncios), palabras claves por las que accede el usuario al sitio y palabras claves dentro del buscador.

Por otro lado, existe la posibilidad de implementar la solución a otros tipos de sitios web como: Ecommerce, sitios web de soporte o servicios, sitios web de captación y sitios web corporativos. La diferencia son los objetivos de cada sitio, ya que deben satisfacer las necesidades del negocio y el propósito de cada uno de ellos.

Bibliografía

- [1] ALDUNATE, E. M. P. Solución de inteligencia de negocios para estudiar el perfil de los postulantes a las carreras de ingeniería en Chile.
- [2] BUREAU, I. A. Estudios iab. <http://www.iab.cl/estudios-iab/>. Consultado: 27-10-2015.
- [3] CONSULTORIA, M. Analítica web. <http://www.analiticaweb.es/lista-de-kpis-por-industria-2/>. Consultado: 17-05-2018.
- [4] DE CHILE, G. Ministerio de transportes y telecomunicaciones - subsecretaría de telecomunicaciones. http://www.subtel.gob.cl/wp-content/uploads/2015/04/PPT_Series_DICIEMBRE_2017_VPUBL.pdf. Consultado: 10-06-2018.
- [5] ELOSEGUI, G. M. T. El arte de medir - manual de analítica web, 2011.
- [6] ELOSEGUI, G. M. T. Marketing analytics - como definir y medir una estrategia online, 2015.
- [7] GANGADHARAN, G. R., AND SWAMI, S. Business intelligence systems: design and implementation strategies. In *Information Technology Interfaces, 2004. 26th International Conference on* (2004), pp. 139–144 Vol.1.
- [8] GOOGLE. Google analytics features. http://www.google.com/intl/es_ALL/analytics/features/index.html. Consultado: 27-10-2015.
- [9] INC., F. Qué es marketing digital, 2013.
- [10] LECAROS, S. Q. Business intelligence en las empresas chilenas.
- [11] PERISIC, N. C. Modelo de inteligencia de negocios aplicado al control de proyectos en una empresa minera.

BIBLIOGRAFÍA

- [12] PETERSON, E., AND ANALYTICS DEMYSTIFIED, W. *The Big Book of Key Performance Indicators*. Web analytics demystified series. Web analytics demystified, 2006.
- [13] ROSALES, P. *Estrategia Digital*. MARKETING Y VENTAS. Deusto, 2010.
- [14] SÁNCHEZ, R. M. To bi or not to bi.
- [15] VÉRTICE, E. *Marketing digital*. Editorial Vértice, 2010.
- [16] Y HOOTSUITE, W. A. S. Digital in 2018 global overview. <https://www.slideshare.net/wearesocial/digital-in-2018-global-overview-86860338>. Consultado: 10-06-2018.
- [17] YAN, K.-Q., WANG, S.-C., WANG, S.-S., AND LIN, Y.-P. Application of data mining for enterprise digital marketing strategy making. In *Communications and Mobile Computing (CMC), 2011 Third International Conference on* (2011), pp. 509–512.