



**FACULTAD
DE INGENIERIA**

Universidad de Buenos Aires



75.06

Organización de
Datos

Trabajo Práctico 1
Análisis Exploratorio

Grupo 34

- *Fandos Nicolás 101018*
- *Ferreyra Javier 100680*
- *Scakosky Matias 99627*
- *Sinisi Fernando 99139*

Link del repositorio de GitHub:

https://github.com/javier2409/TP1_Datos_2_C2018/

Indice

1.	<i>Introducción.....</i>	<i>5</i>
2.	<i>Descripción de las interacciones por usuario.....</i>	<i>6</i>
2.1	<i>Lista de eventos posibles por usuario.....</i>	<i>6</i>
2.2.1	<i>Relación entre las personas y eventos.....</i>	<i>7</i>
2.2.2	<i>Cantidad de personas según sus primeras 4 acciones.....</i>	<i>8</i>
2.2.3	<i>Primera acción del usuario en Trocafone en 2018.....</i>	<i>8</i>
2.2.4	<i>Relación entre cantidad de eventos y cantidad de personas que lo realizan</i>	
3.	<i>Compras / Conversiones.....</i>	<i>11</i>
3.1	<i>Visitas por persona y compras por usuario.....</i>	<i>11</i>
3.2	<i>Distribución de los eventos de usuarios hasta 1ª compra.....</i>	<i>12</i>
3.3	<i>Conclusión del 100% de usuarios y compras.....</i>	<i>13</i>
4.	<i>Productos y modelos.....</i>	<i>15</i>
4.1	<i>Los modelos más populares.....</i>	<i>15</i>
4.1.1	<i>Top 7 de los modelos más buscados.....</i>	<i>16</i>
4.2	<i>Los productos más vistos.....</i>	<i>17</i>
4.3	<i>Los productos más comprados.....</i>	<i>18</i>
4.3.1	<i>Los productos comprados, elegidos por más usuarios.....</i>	<i>18</i>
4.3.2	<i>Los modelos más elegidos.....</i>	<i>19</i>
5.	<i>Análisis temporal de los datos.....</i>	<i>21</i>
5.1	<i>Horarios de mayor actividad en la plataforma.....</i>	<i>21</i>
5.2	<i>Tráfico semanal de la web.....</i>	<i>22</i>
5.3	<i>Tráfico semanal respecto de las horas del día.....</i>	<i>23</i>
5.3.1	<i>Compras en la semana, según su horario.....</i>	<i>24</i>
5.4	<i>Compras de celulares por mes.....</i>	<i>25</i>
5.5	<i>Visitas por mes al site.....</i>	<i>26</i>
5.5.1	<i>Evolución de las visitas.....</i>	<i>27</i>
5.6	<i>Distribución del tiempo de permanencia en la página de las personas que compraron.....</i>	<i>28</i>
6.	<i>Análisis Geográfico de los datos.....</i>	<i>30</i>
6.1	<i>Dispositivo más utilizado según la región.....</i>	<i>30</i>
6.2	<i>Visitas según la región.....</i>	<i>31</i>
6.3	<i>Otras visitas fuera de Brasil.....</i>	<i>32</i>

7. Análisis de la interacción con anuncios web.....	34
7.1 Clicks en ads según la hora del día.....	34
7.2 Cantidad de clicks según la campaña de publicidad.....	35
8. Sistemas operativos de los usuarios.....	37
8.1 Visitas a la página según Sistema Operativo.....	37
8.2 Cantidad de conversiones según sistemas utilizados para ingresar al sitio.....	38
8.3 Cantidad de conversiones según sistemas utilizados para ingresar al sitio.....	39
8.4 Comparación entre visitas y conversiones.....	40
9. Tipo de dispositivo.....	42
9.1 Cantidad de visitas según dispositivo.....	42
9.2 Compras según dispositivos utilizados.....	42
9.3 Canales según dispositivo.....	43
10. Análisis según resolución de la pantalla.....	45
10.1 Vistas según resolución del dispositivo.....	45
10.2 Cantidad de personas que visitaron/compraron vs. Resolución que usaron para ingresar.....	46
10.2.1 Cantidad de personas que visitaron/compraron vs. Resolución que usaron para ingresar(log).....	47
11. Análisis de los motores de búsqueda.....	48
11.1 Cantidad de visitas según el motor de búsqueda.....	48
11.1.1 Cantidad de visitas según el motor de búsqueda(log)....	49
12. Análisis del canal origen de eventos.....	50
12.1 Visitas según el canal que las originó en función del tiempo.....	50
12.2 Cantidad de personas que visitaron por primera vez el sitio según canal.....	51
12.3 Cantidad de personas que compraron según desde qué canal ingresaron al sitio.....	52
13. Conclusiones.....	54

1. Introducción

El objetivo de este trabajo práctico es analizar un conjunto de datos brindados por la empresa Trocafone. Los usuarios que visitaron la plataforma de ecommerce de Brasil, realizaron diferentes eventos a ser evaluados. Los mismos son tomados desde una web de analytics y procesados.

Trocafone es un sitio de compra-venta de dispositivos celulares, con un mercado muy fuerte en Brasil, y en amplio crecimiento en el mercado Argentino.

Para comenzar a realizar este trabajo, la cátedra nos brindó un archivo de formato .CSV con la información seleccionada específicamente para nosotros. Los mismos contienen un grupo muy amplio de datos que nos permiten hacer un análisis heterogéneo de las interacciones entre los usuarios que ingresaron a Trocafone desde Enero a Junio de 2018.

En las siguientes páginas se mostrarán los resultados obtenidos del análisis previamente mencionado.

2. Descripción de las interacciones de usuario

El usuario que ingrese a Trocafone interactúa con la plataforma de diferentes maneras, las cuales serán detalladas en esta sección, de manera tal que pueda entenderse que significa cada acción que la persona haga dentro de la página.

Dentro del set de datos, el dominio de estas interacciones son llamados eventos.

2.1 Lista de eventos posibles de un usuario

- Visitar la página de un producto
- Visitar un listado específico de una marca
- Ingresar a una determinada url
- Visitar la homepage
- Realizar una búsqueda de productos desde el site
- Ingresar al sitio mediante un motor de búsqueda web
- Visitar una página
- Ingresar al checkout de compra de un producto
- Comprar un producto
- Registrarse para recibir notificaciones de disponibilidad de productos y de stock.

2.2 Análisis de eventos de Trocafone

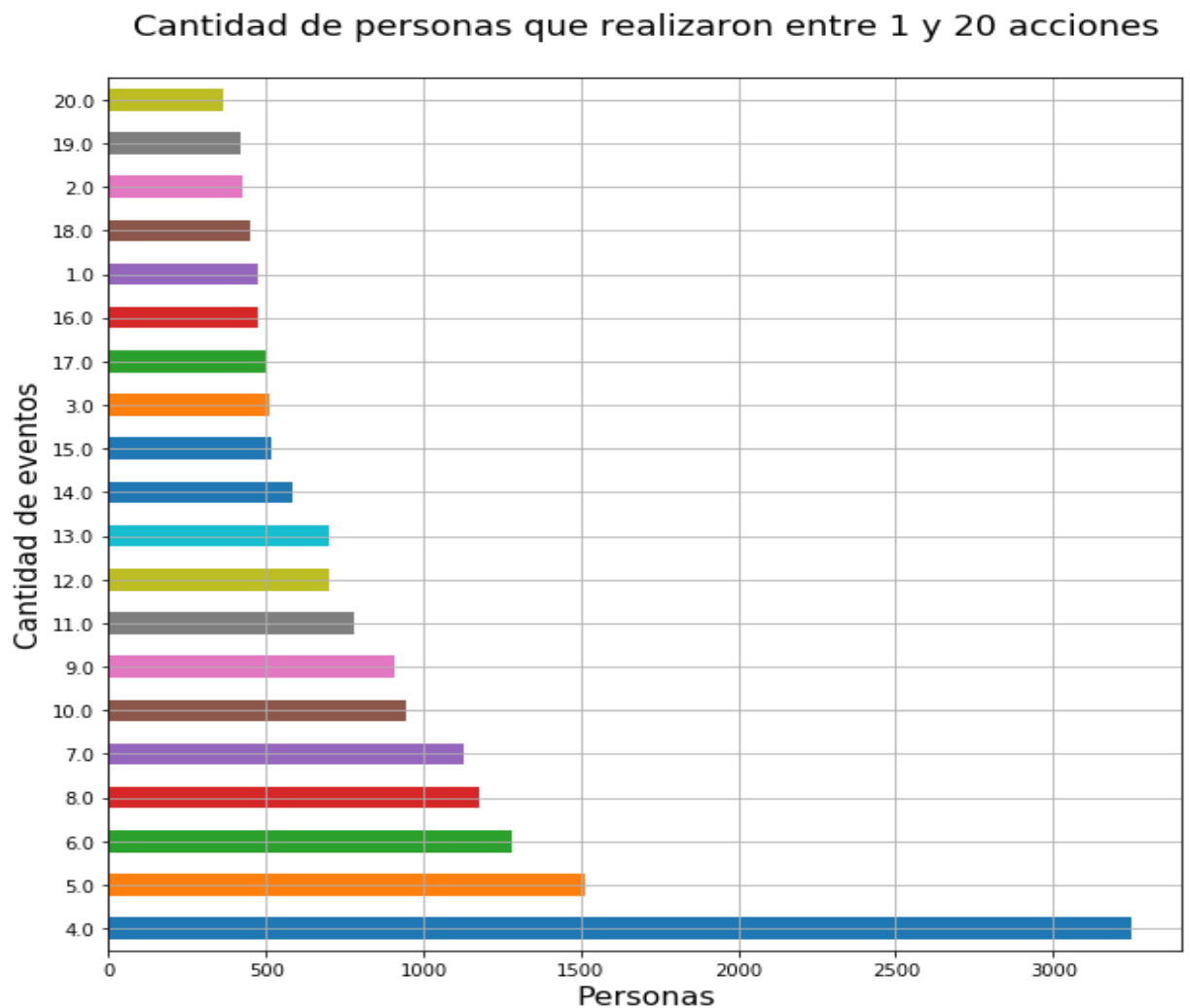
Los eventos mencionados en la sección 2.1 se suceden con diferentes frecuencias entre los usuarios de la plataforma web. Este es el Top 5 de los eventos ocurridos en la página.



Se puede apreciar en el gráfico que el evento predominante del set de datos es “viewed product”, correspondiente a la acción de visitar la página de un producto específico.

2.2.1 Relación entre las personas y eventos

Los usuarios que entran a la página no realizan una sola acción, teniendo en cuenta que a mayor cantidad de acciones hay más probabilidades que una de ellas sea una compra de producto, es interesante estudiar cuantas acciones suelen hacer las personas que entran a la plataforma



Se puede apreciar con claridad que la gran mayoría de usuarios que entran a la plataforma realizan cuatro acciones, con amplia diferencia respecto de la segunda cantidad de acciones por usuario. De todas maneras, los usuarios que ingresan al site, realizan 4, 5 o 6 en porcentaje mayoritario respecto de otras cantidades de acciones de los usuarios que ingresan.

Una vez obtenida esta información sería interesante saber cuales son esas 4 acciones que realizan los usuarios que ingresan a la página, para poder saber entre otras cosas cuál es su navegación en la misma o como llegó a Trocafone

2.2.2 Cantidad de personas según sus primeras 4 acciones

El estudio inmediatamente posterior a conocer que la mayoría de los usuarios realizan cuatro acciones, una buena pregunta para hacer es ¿Cuáles son los cuatro primeros eventos que un usuario realiza al entrar a Trocafone?

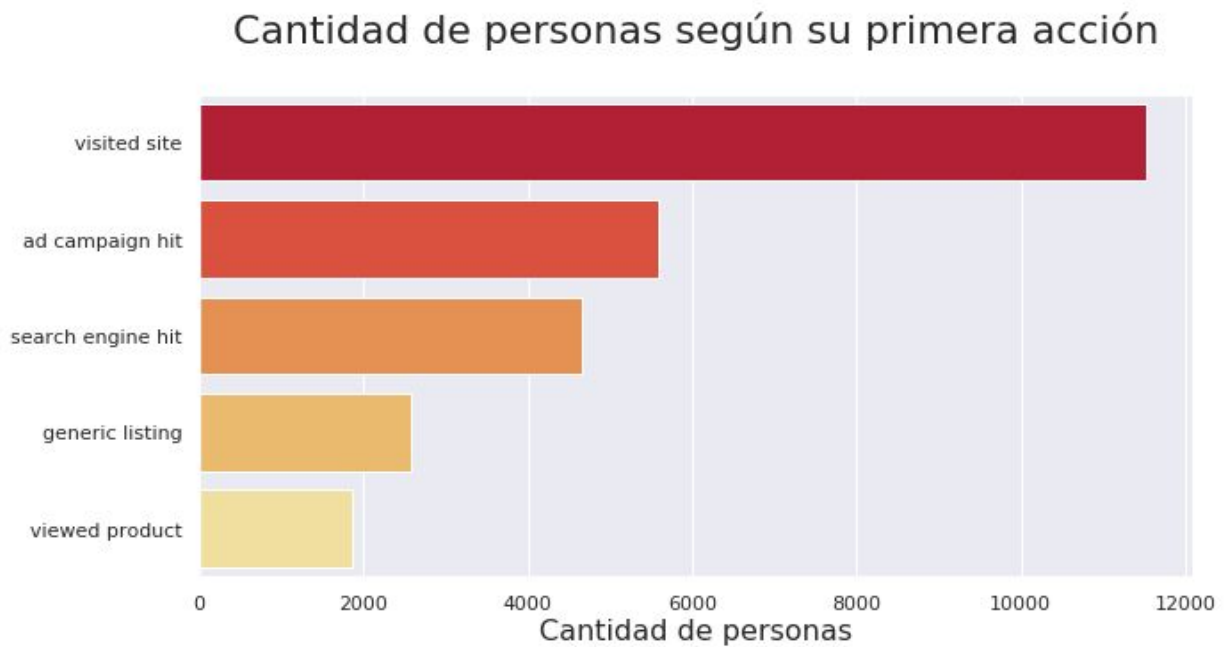
Cantidad de personas según sus primeras cuatro acciones



Los primeros cuatro eventos fueron tomados para cada usuario del dataset, desde la primera vez que se registró un evento del mismo hasta el cuarto, sin excepción.

2.2.3 Primera acción del usuario en Trocafone en 2018

Del estudio anterior, y de manera análoga se puede preguntar cuál es el primer evento que realiza un usuario al entrar al website.

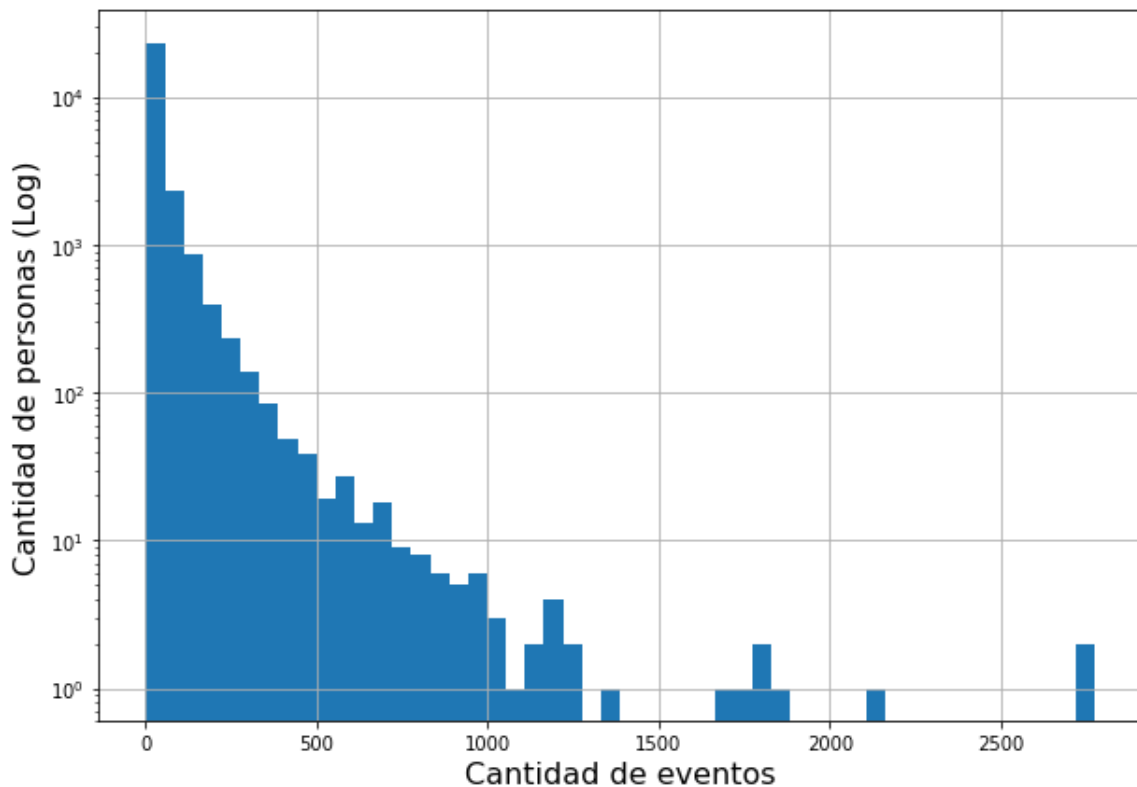


Se puede apreciar que la fuerte aparición del evento "visited site" se corresponde fuertemente con el análisis de los primeros 4 eventos del usuario presentado en el gráfico anterior, ya que en todas las combinaciones más frecuentes de los primeros cuatro eventos, el evento primordial de "visited site" aparece en todos. De la misma forma aparece "ad campaign hit"

2.2.4 Relación entre cantidad de eventos y cantidad de personas que lo realizan

A partir de la información presentada en la sección de relación entre personas y eventos (2.2.1) se puede apreciar que a mayor cantidad de acciones, menos personas son las que lo realizan.

Cantidad de personas que realizaron cierta cantidad de acciones (Log)



El gráfico está en escala logarítmica, para lograr una mejor visualización de los datos, ya que muchos datos están concentrados en pocos lugares y no se puede apreciar las diferencias (Se puede apreciar que en la zona de 2500, hay un usuario, eso es una excepción que aparece en este gráfico debido a estar representado en escala logarítmica ya que hay un par de usuarios que efectivamente realizaron alrededor de 2500 eventos pero no es un dato útil pues no es el comportamiento usual de un usuario típico de Trocafone).

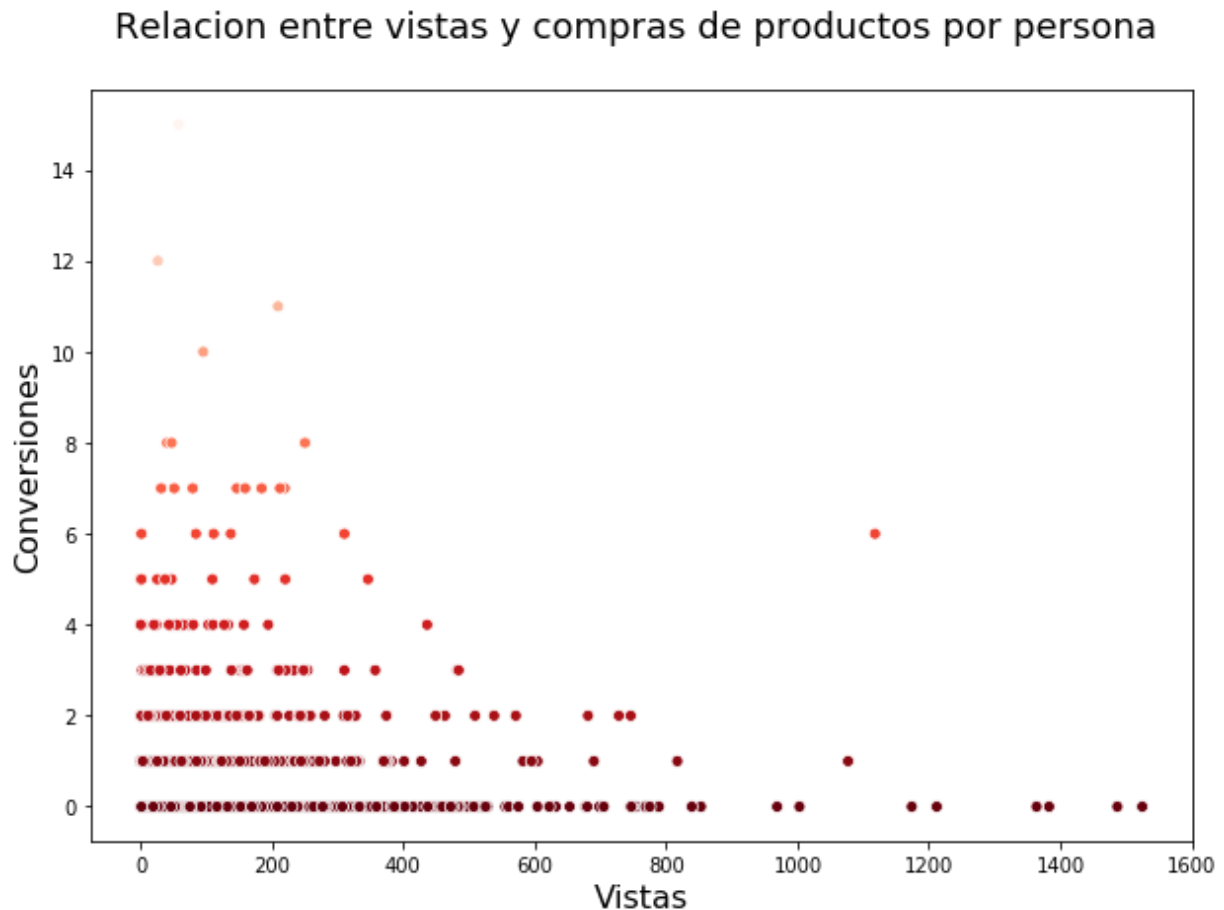
3. Compras / Conversiones

Está claro que dentro de todas las acciones que pueden realizar los usuarios de Trocafone, las compras, o también llamadas conversiones, se destacan por sobre el resto, debido a que el objetivo de la empresa es que las personas que visiten el site terminen realizando una conversión.

En esta sección se analizará la relación entre las compras y otras variables presentes en el dataset.

3.1 Cuántas visitas hace una persona y cuántas veces compra un usuario

Existe una relación entre las veces que un usuario visita la página de un producto y termina comprando, existen personas que vieron muchas veces un celular y capaz compraron o no, la idea de esta sección es analizar las compras de los usuarios según la cantidad de veces que visitaron la página de uno o varios teléfonos



Se puede apreciar que cada punto representa, para un usuario particular, la relación entre la cantidad de veces que visitó la página de un producto y la cantidad de conversiones que realizó.

En un primer vistazo de este gráfico se puede apreciar que muchas personas vieron productos y no compraron nada.

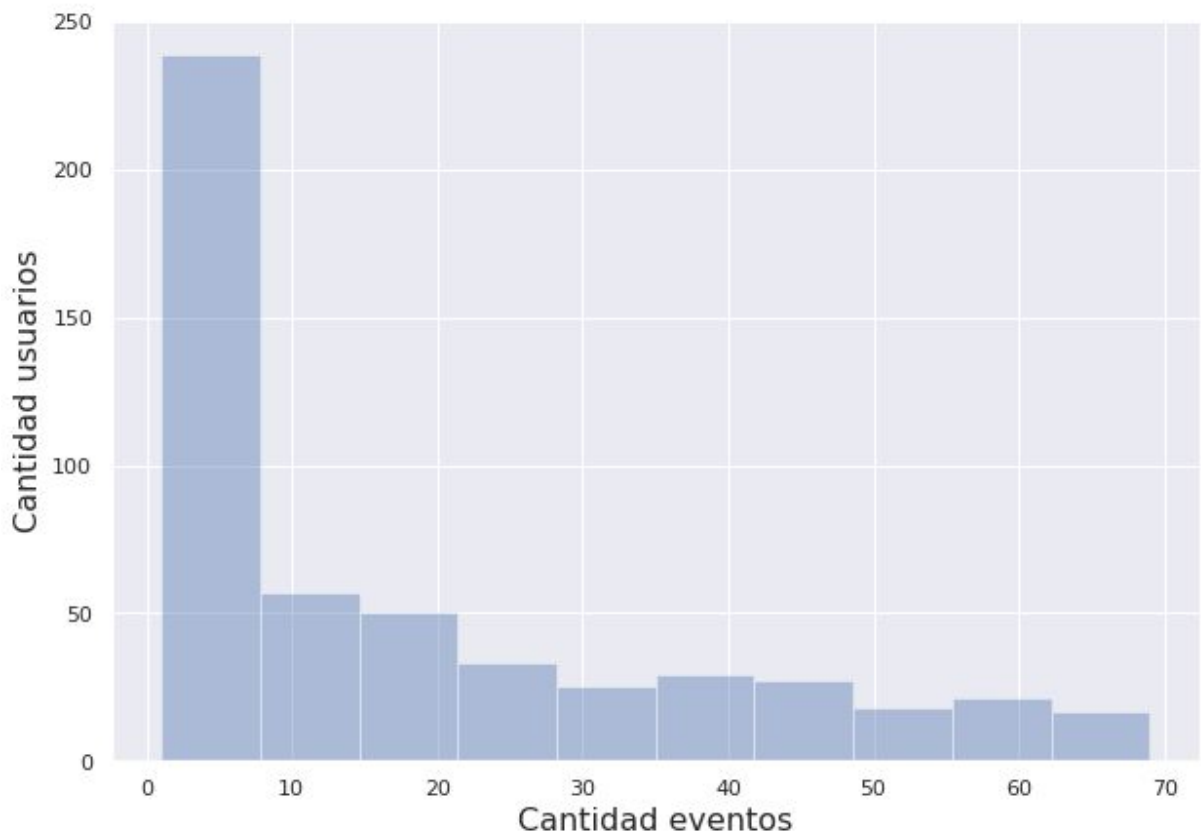
También se puede percibir que muchos de los usuarios que realizaron algunas conversiones no presentan una cantidad exagerada de visitas al sitio.

3.2 Cómo se distribuyen los eventos de los usuarios hasta su primera compra

La cantidad de eventos que hacen las personas hasta su primera compra podría pensarse como una variable aleatoria. Para la misma se puede generar su función

histograma y poder estudiar cuantos eventos diferentes hacen diferentes

Distribucion de la cantidad de eventos hasta la primera compra

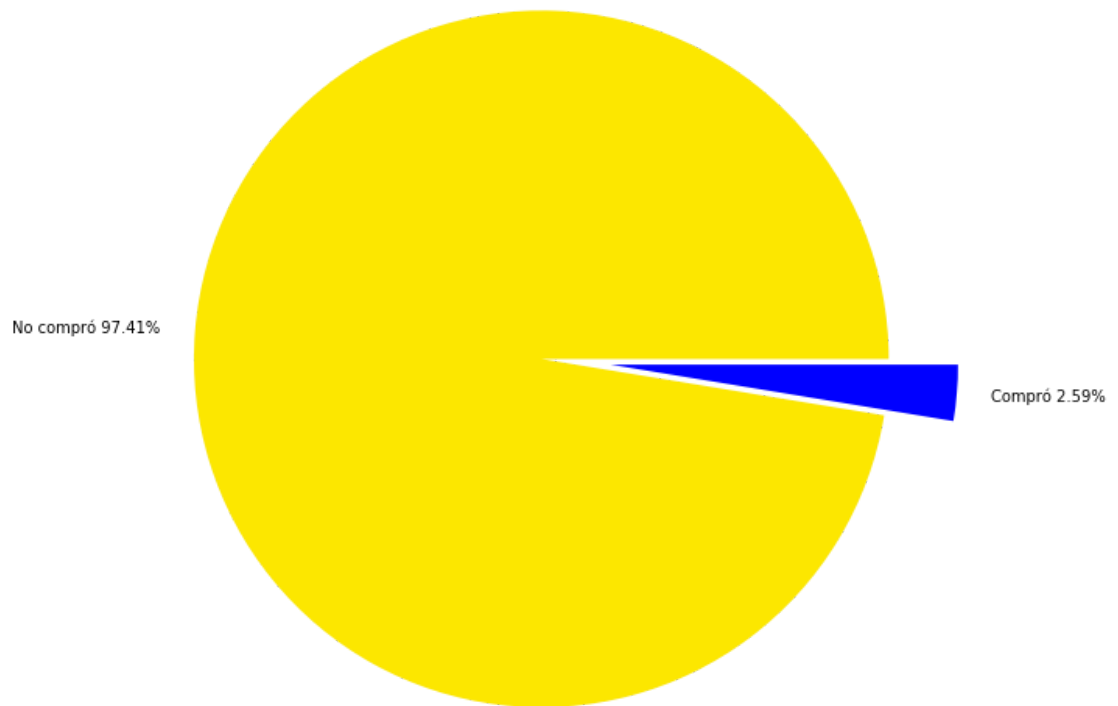


La mayoría de los usuarios hicieron alrededor de 10 eventos para hacer su primera conversión.

3.3 Conclusión del 100% de usuarios y compras

Por último, para analizar puramente al evento de conversión, se puede analizar a todos los usuarios que aparecen registrados en el set de datos y saber porcentualmente cuantos realizaron al menos una conversión.

Porcentaje de personas que compraron vs no compraron



Se puede ver que casi el 3% de las personas que entraron terminaron realizando una conversión y el 97% entra a trocafone y se va sin comprar nada.

4. Productos y modelos

Esta sección se avoca a analizar los modelos de celulares, sus características y las preferencias de los usuarios respecto de los mismos, también las conversiones siguen siendo un factor importante a la hora del análisis, pero en esta sección el enfoque estará puesto en función de los productos.

Los productos en Trocafone pueden elegirse con diferentes estados, colores y capacidades de almacenamiento. Cada combinación única de modelo, estado, color, capacidad es identificada mediante un código numérico único llamado Sku, el cual por sí mismo visto a simple vista no ofrece mucha información, pero su carácter de único nos permite realizar un análisis de las compras de cada producto específico

4.1 Los modelos más populares

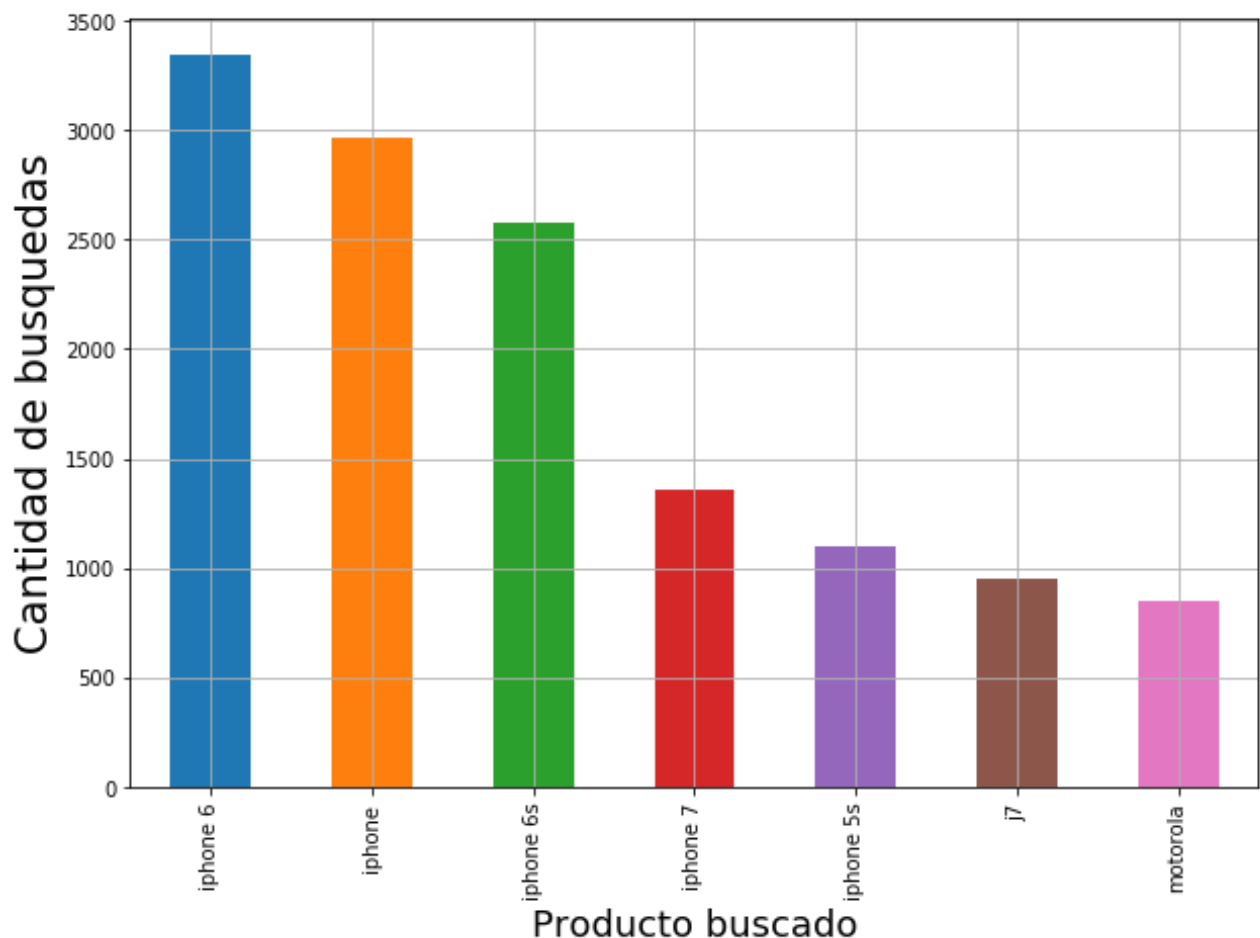


Los usuarios tienen preferencias según el celular que quieran comprar, las búsquedas más populares de equipos de telefonía móvil están representados según su jerarquía de aparición en el motor de búsqueda

4.1.1 Top 7 de los modelos más buscados

Una forma de respaldar el gráfico de la figura es contabilizando los 7 modelos de celulares más buscados

Top 7 de productos mas buscados en Trocafone



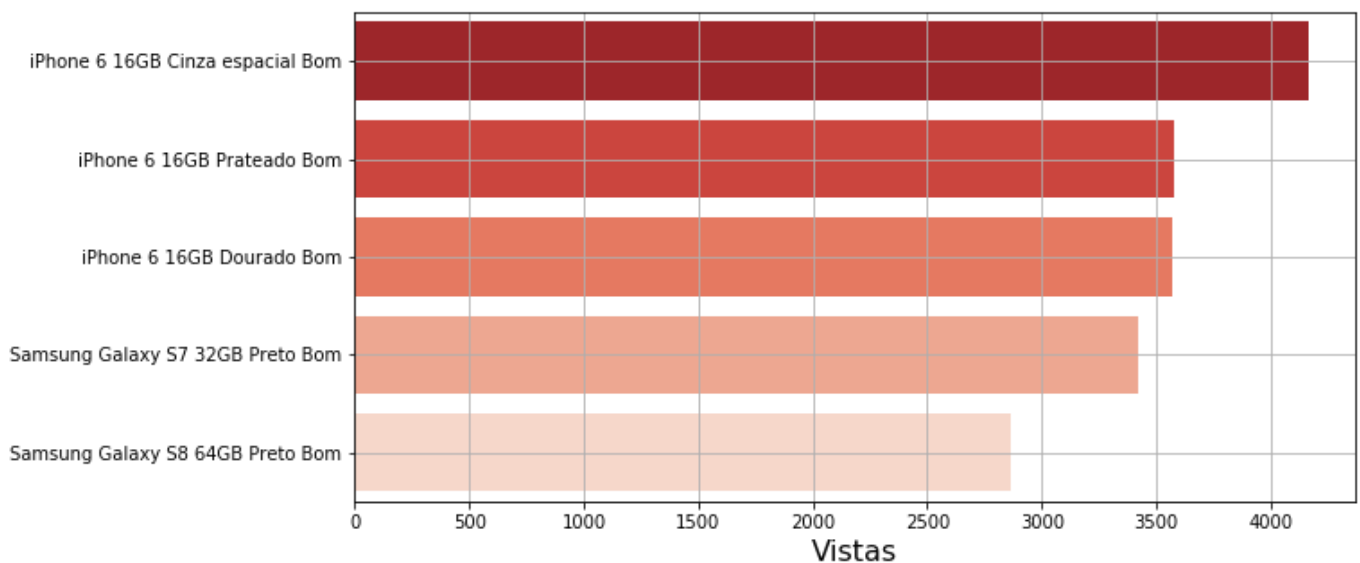
Se puede apreciar como el iPhone es un producto bastante popular entre las búsquedas de Trocafone, lo cual se corresponde con las jerarquías mencionadas en 4.1

4.2 Los productos más vistos

Los productos más vistos son aquellos que los usuarios estuvieron interesados, entrando en la página del mismo, dentro del set de datos esto es identificado como el evento "viewed product".

Teniendo en cuenta que este es el evento que predomina de manera mayoritaria en el set de datos, este análisis es interesante ya que es el modelo más popular, dentro del evento más influyente en el conjunto

Los 5 productos más vistos

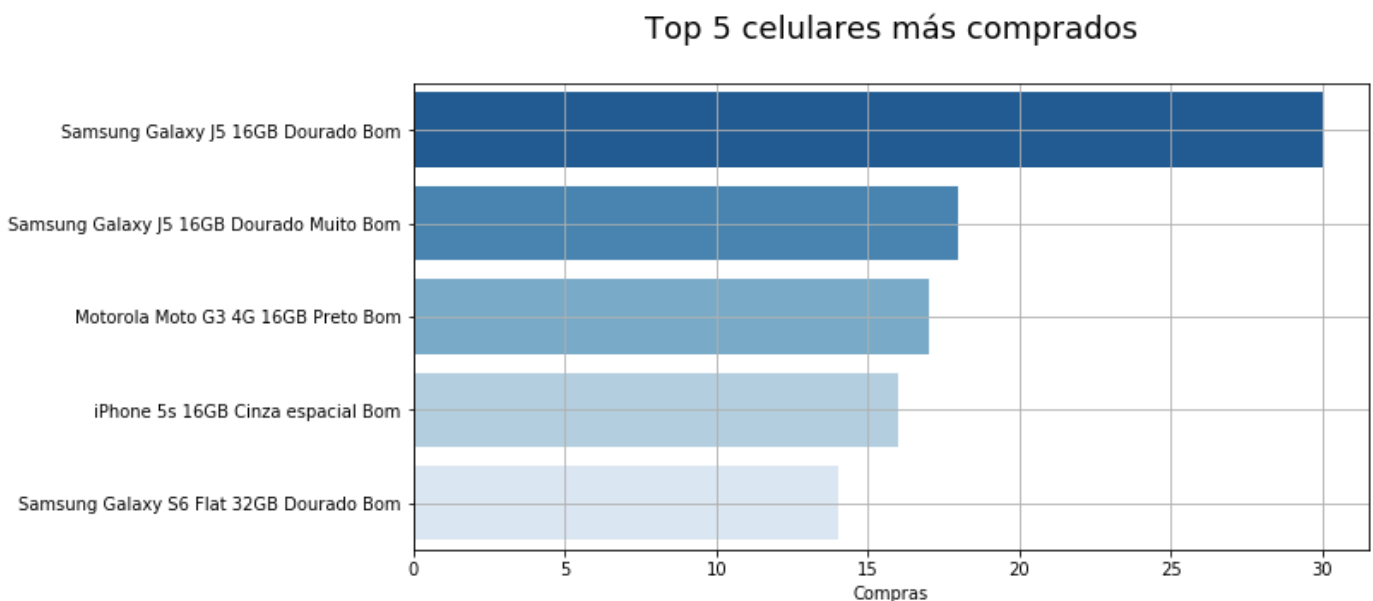


Teniendo en cuenta que producto ya incluye, además del modelo de celular en si, la capacidad de almacenamiento interna del teléfono y el estado en el que el celular se encuentra.

Ya que para la página, el mismo celular pero en diferentes estados es considerado como dos productos diferentes.

4.3 Los productos más comprados

Estos son los 5 productos más vendidos:

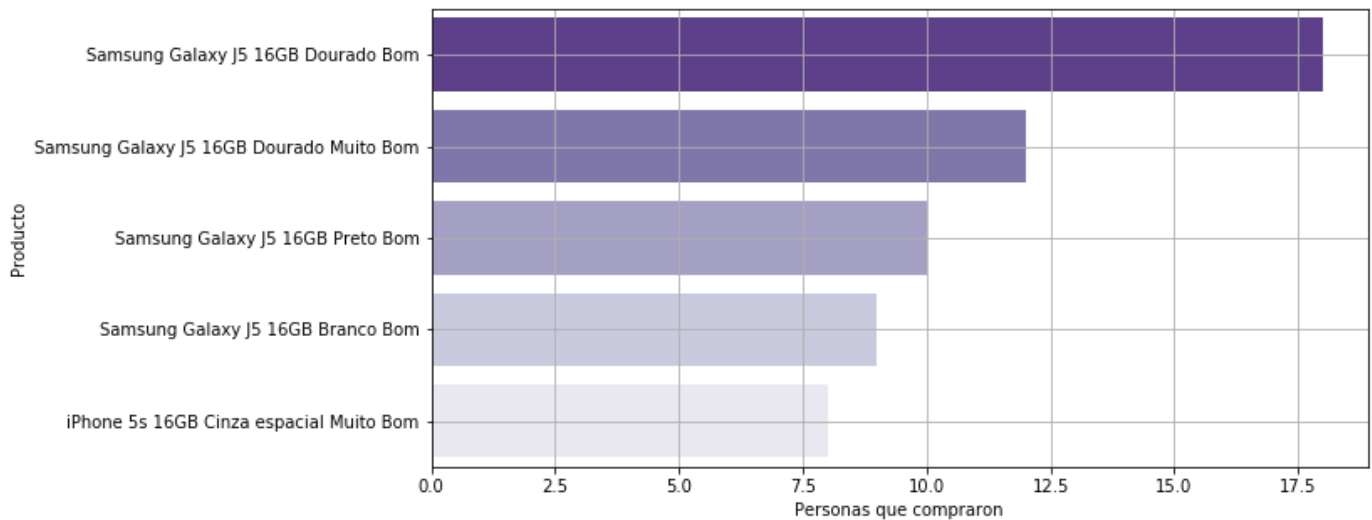


Se puede ver que los que encabezan la lista son celulares que son conocidos en el mercado de la telefonía como celulares de no muy alto costo pero de buen rendimiento.

4.3.1 Los productos comprados, elegidos por más usuarios.

A diferencia de lo analizado en 4.3 los productos más comprados, son aquellos que fueron comprados en diferentes ocasiones y varios ejemplares del mismo producto. Aquí se detalla los productos elegidos por más usuarios, que decidieron realizar una conversión.

Top 5 productos con más compras (1 por persona)

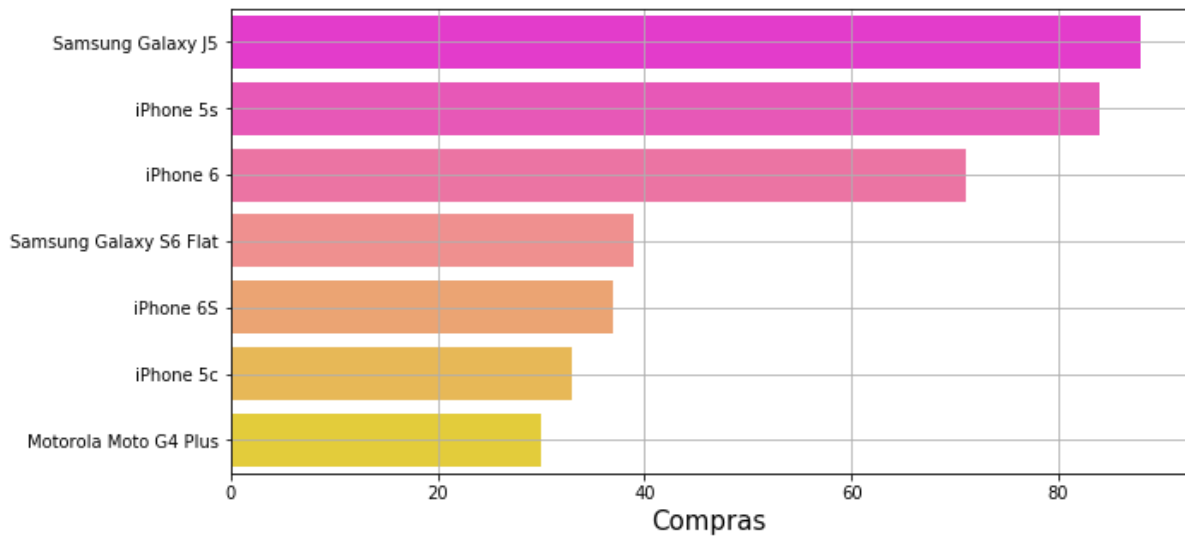


Siguiendo con la línea del gráfico anterior, se puede apreciar que algunos de los productos más comprados coincide con los más elegidos por los usuarios, aunque pareciera que la línea es muy delgada en cuanto a la diferencia, es significativa respecto del análisis.

4.3.2 Los modelos más elegidos

Como fue introducido al inicio de esta sección, se marca una diferencia entre producto y modelo, al poder hacer esta separación, podemos ofrecer una visualización la cual no repita el modelo cambiando features que hacen diferentes al producto y poder enfocarnos en los modelos más elegidos por los usuarios, independientemente de su estado o capacidad de almacenamiento.

Top 7 modelos mas vendidos



La ventaja que nos da este estudio del modelo más allá del producto, respecto del gráfico presentado en 4.3.1, donde parecía que el Samsung Galaxy J5 se repetía en sus diferentes variantes, aquí es coherente con la preferencia del modelo respecto de los usuarios, pero da lugar a un análisis más heterogéneo de modelos de celular.

5. Análisis temporal de los datos

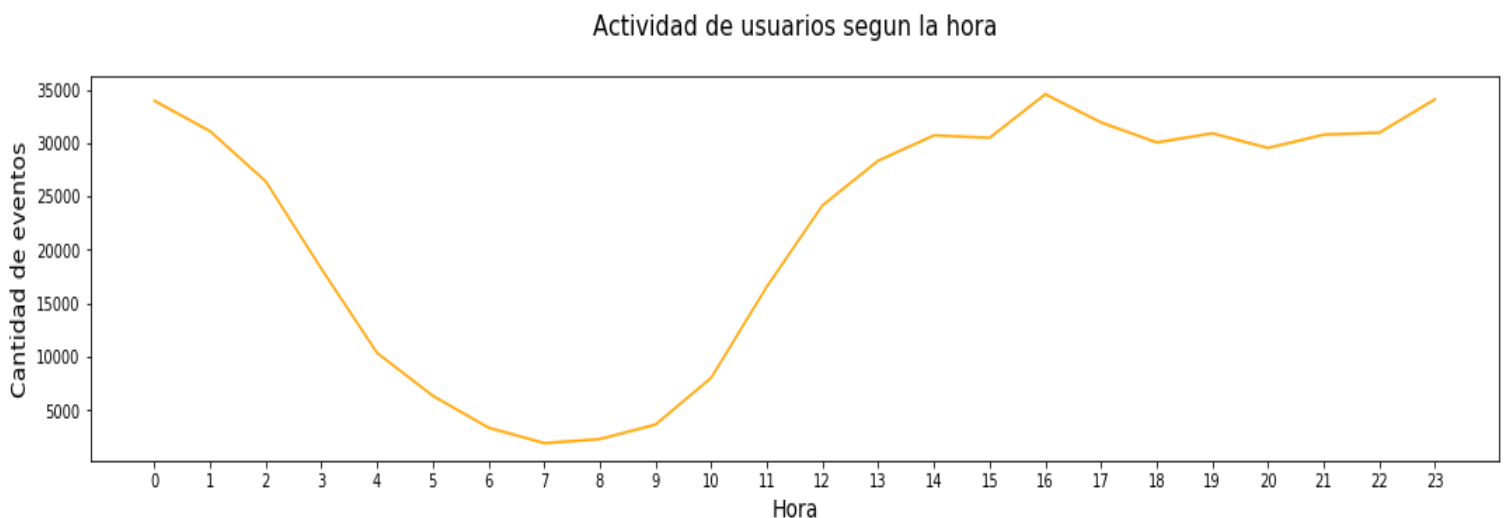
Dentro de la sección de análisis temporal de los datos, se estudiarán distintas evoluciones de los elementos del dataset respecto del tiempo.

Analizar esta sección es importante debido a que el tiempo es un factor que permite entender como función el flujo de compras de los usuarios de Trocafone, y otra vez, el evento de conversión dentro de esta sección volverá a ser importante para descubrir nuevas cosas acerca de los datos.

También entender cómo el factor tiempo afecta a los usuarios de la plataforma debe ser tomado como algo importante ya que puede ser una herramienta de utilidad conocer el tráfico de usuarios en pos de querer mejorar el servicio web.

5.1 Horarios de mayor actividad en la plataforma

Unas de las primeras cosas a analizar es la actividad de los usuarios al día, es decir cuando se produce más tráfico en la página.

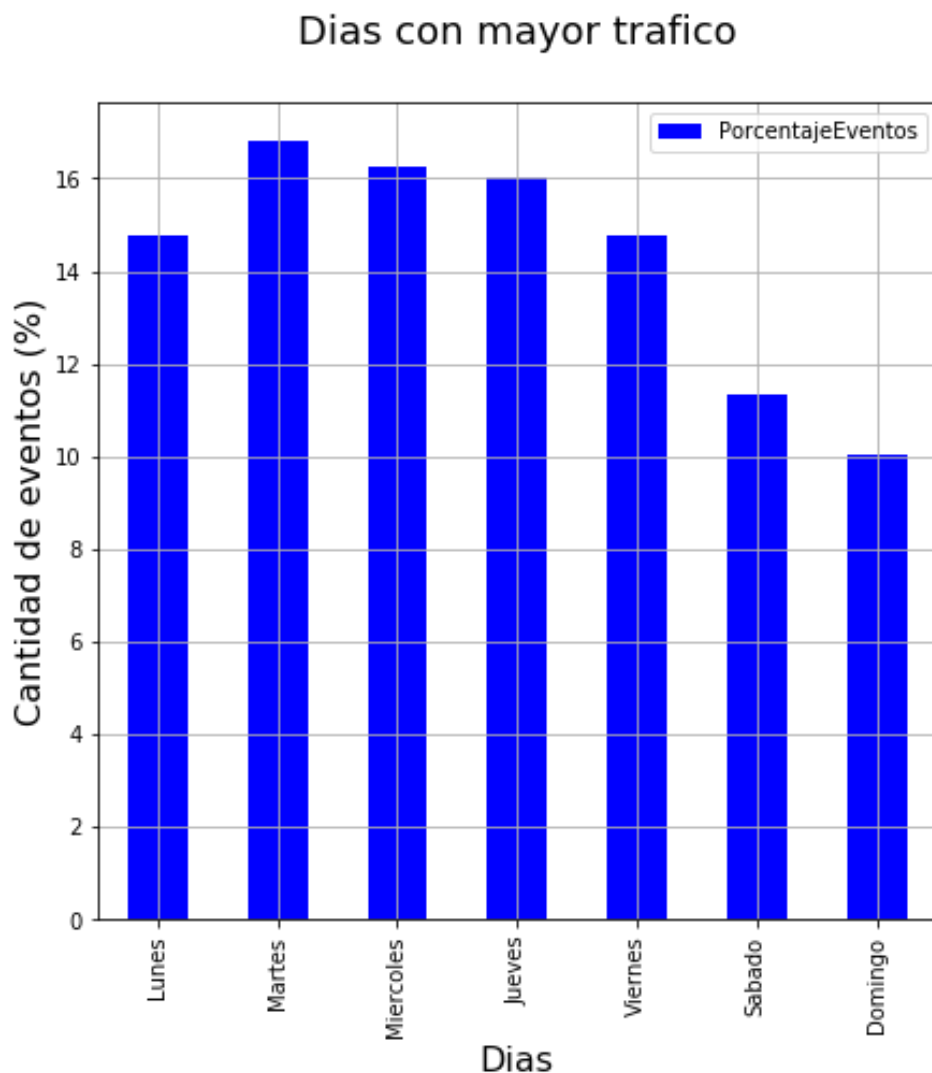


Como conclusión podemos rescatar que los horarios de mayor tráfico son en los vespertinos y nocturnos, pero con

una disminución a la madrugada y llegando al menor pico al amanecer.

5.2 Tráfico semanal de la web

Dentro de los días de la semana se puede intentar saber si algún día de la semana tiene alguna falta importante de tráfico de eventos, o también saber si algún día de la semana la plataforma tiene una actividad mayor a lo usual para, de alguna manera, poder aprovechar que los usuarios están dentro de la página.

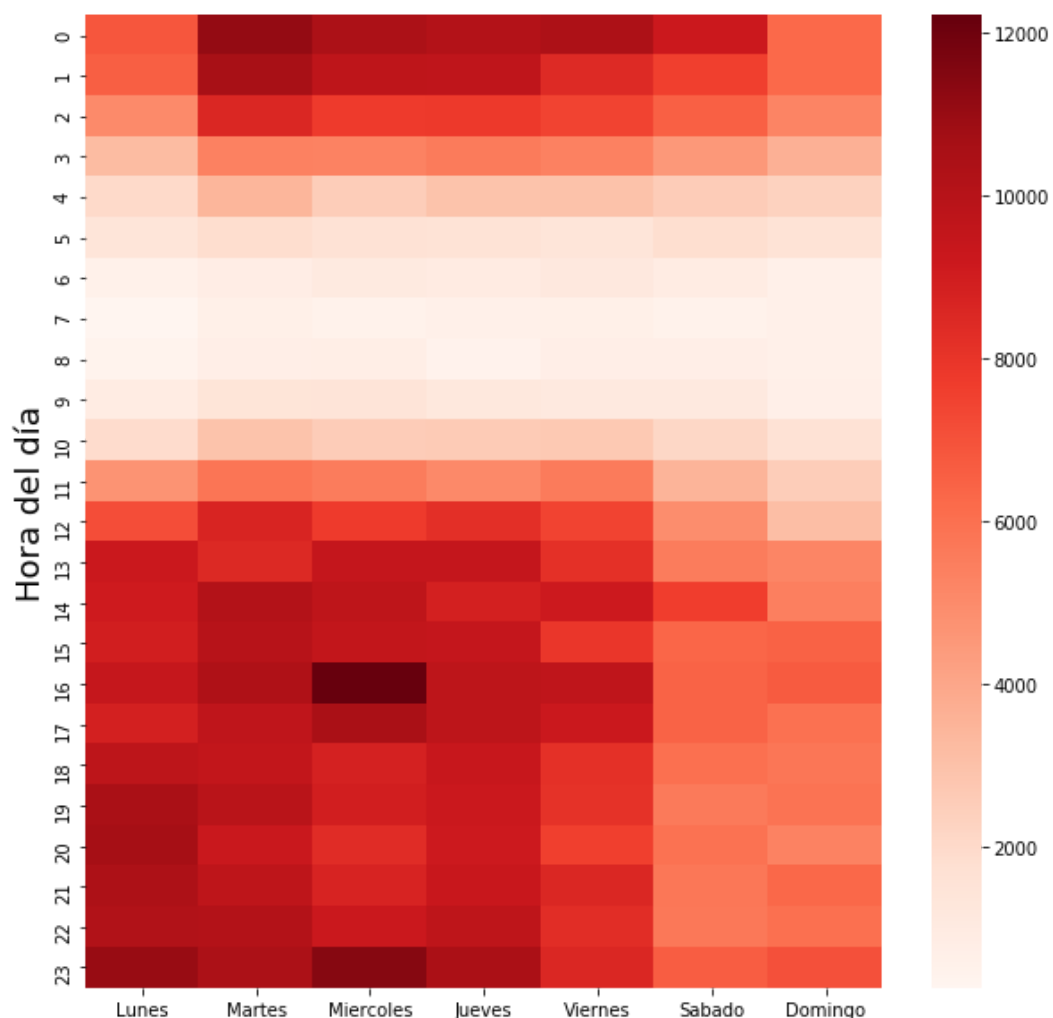


Como puede apreciarse dentro de esta visualización, el tráfico de eventos está representado en forma porcentual y los días de menor actividad son durante el fin de semana, ya que sábado y domingo la plataforma presenta una caída significativa en la interacción de los usuarios. Lo cual a priori es algo extraño debido a que en el tiempo libre del fin de semana la gente entra menos a la página.

5.3 Tráfico semanal respecto de las horas del día

Unión los datos del gráfico 5.1 y los datos del gráfico 5.2 podemos representar más justamente en que momento de la semana se produce mayor actividad en la página web.

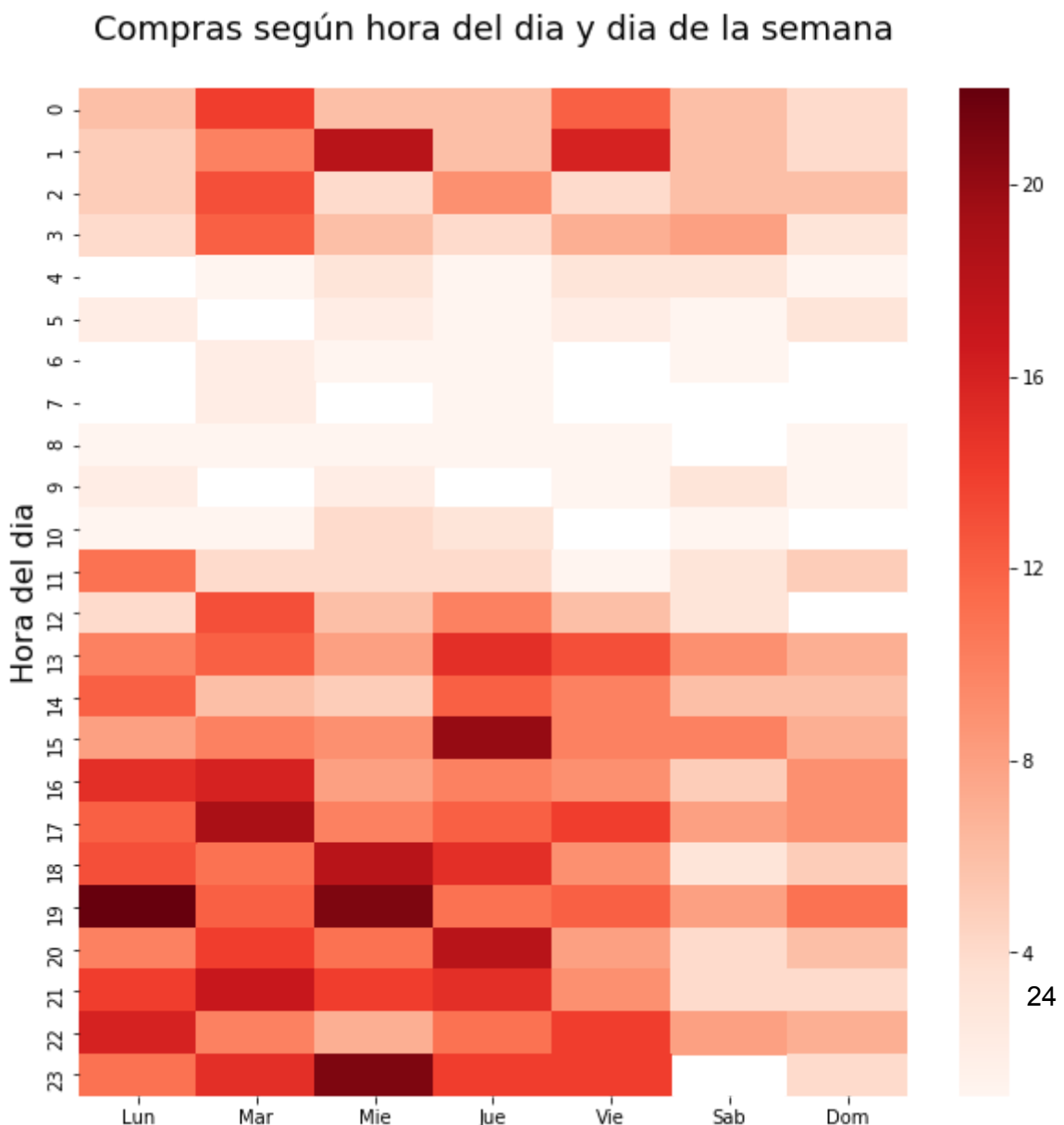
Eventos por día de la semana segun hora del día



Como conclusión se obtiene que los fines de semana tanto como a principio del día (horas de la mañana) son los horarios con menos movimiento en la página, por lo contrario se produce el mayor tráfico en las horas de la tarde/noche y más predominante en los días laborales (lunes a viernes)

5.3.1 Compras en la semana, según su horario

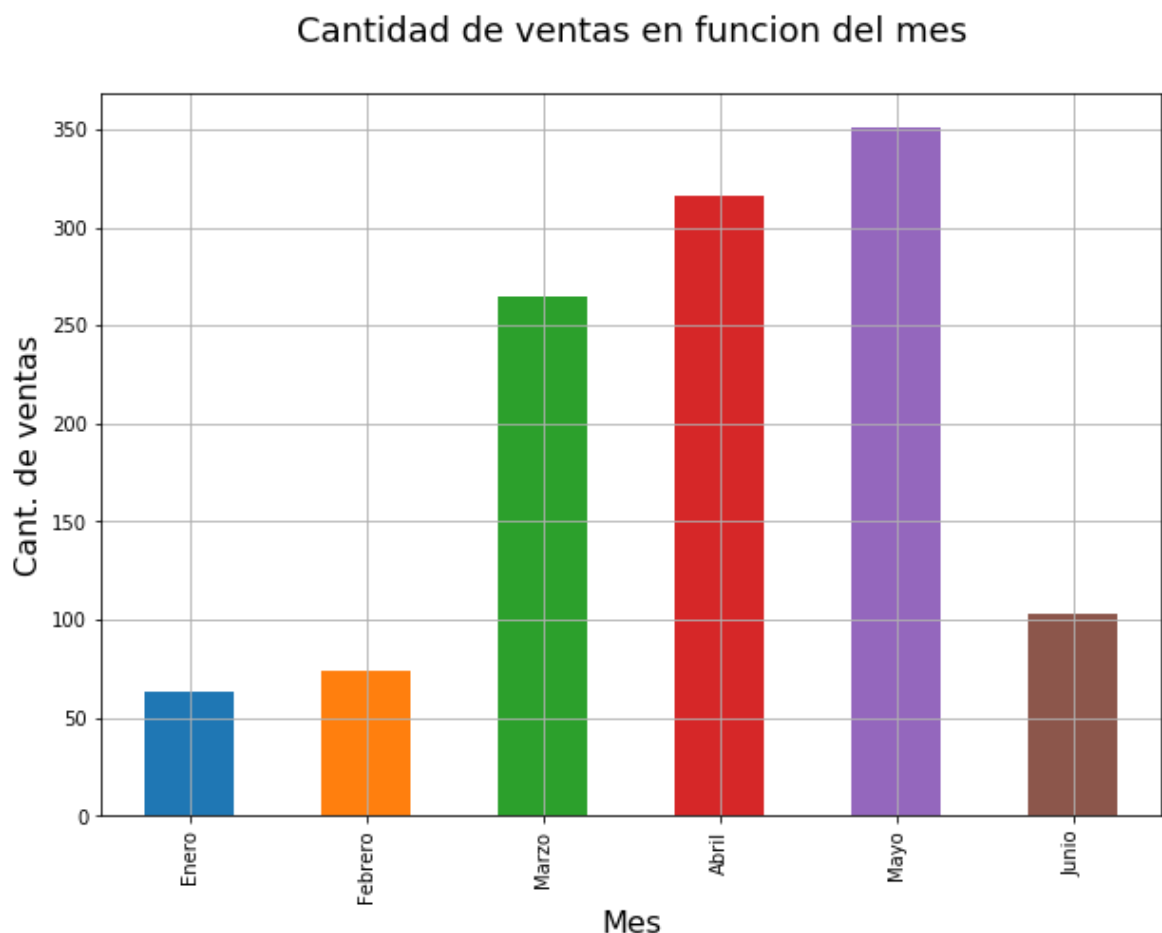
Algo interesante para tomar acerca del estudio inmediato anterior acerca de los eventos y sus ocurrencias en los días de la semana y sus horarios, es particularmente el caso de las conversiones, donde lo que buscamos ahora es saber en qué momento de la semana, con mayor exactitud posible, se realizan la mayor cantidad de conversiones.



Pudiendo hacer la comparación con el análisis anterior, los resultados son lo suficientemente acoplados como para poder sostener la conclusión mencionada en 5.3.1 para las conversiones también.

5.4 Compras de celulares por mes

También es interesante analizar la relación de la cantidad de ventas por mes que tuvo la página, aunque se debe tener en cuenta que solo contamos con los datos de la mitad del mes de Junio (los primeros 15 días) pero decidimos graficarlos igualmente.



Como se puede observar el mes con más ventas fue Mayo, donde las mismas fueron creciendo mes a mes desde enero y considerando que al mes de Junio solo tenemos la mitad de los datos, no sabemos si hubiera seguido el crecimiento mensual o si por la cantidad de ventas en proporción a los primeros días del mes se produjera un descenso de ventas.

5.5 Visitas por mes al site

Teniendo en cuenta que en Junio solo están tomados los primeros 15 días de los datos, las visitas al sitio web se ven de la siguiente manera, representadas por un gráfico continuo



El periodo de visitas de Enero a Mayo se ve como una función creciente, pero en mayo, donde alcanza el máximo y comienza a decrecer, es debido a que no se extrapolaron los datos del mes de Junio, ya que como fue mencionado anteriormente, los datos fueron tomados hasta los primeros 15 días del mes

5.5.1 Evolución de las visitas

Para solucionar el problema de tener los datos de Junio incompletos, el siguiente gráfico muestra cómo evolucionan las visitas de los usuarios en función del tiempo, evitando hacer un promedio ponderado de las visitas y llevando un control hasta mitad de Junio.



5.5.2 Evolución de los eventos

Si realizamos un análisis similar para los eventos realizados diariamente obtenemos:

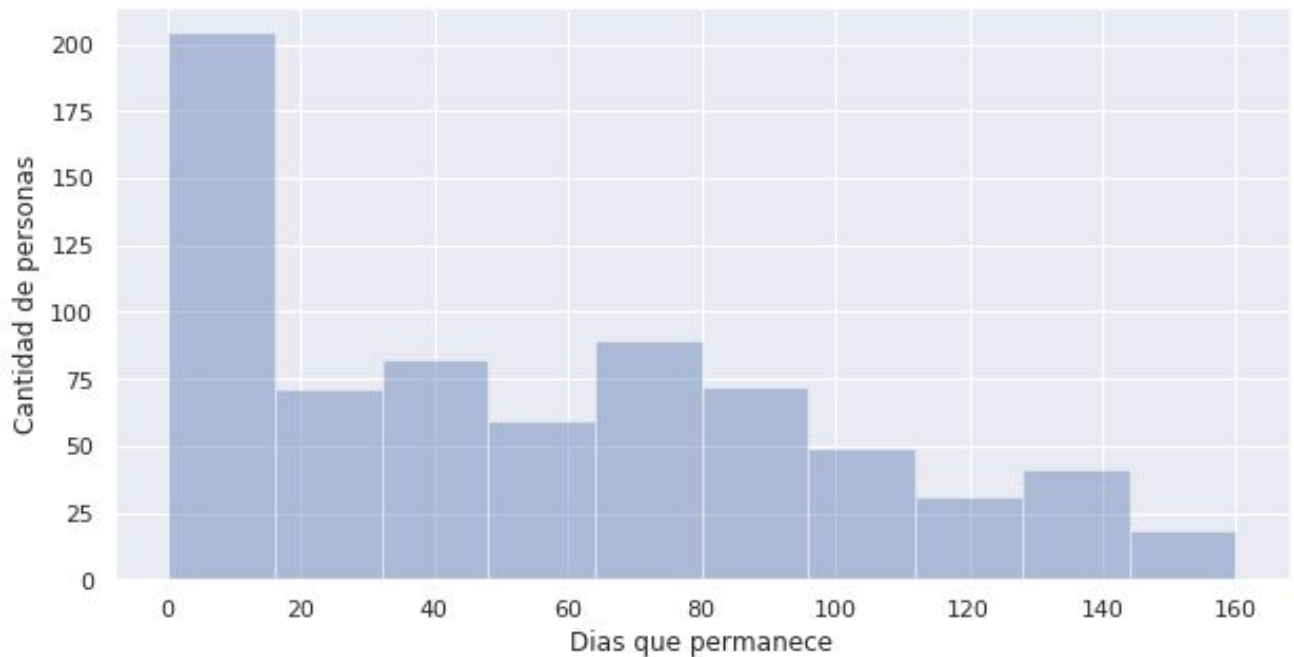


Aquí también se observa una evolución similar de los eventos totales realizados en el sitio, y evidencia que a mediados de mayo fue cuando el sitio comenzó a tener más actividad. Y que sólo hay datos disponibles hasta mediados de Junio.

5.6 Distribución del tiempo de permanencia en la página de las personas que compraron

Las personas que entran a Trocafone pueden tardar varios días en decidirse a comprar e incluso aquellos que se deciden pueden luego regresar a la página para seguir viendo otros productos y/o realizar más compras. Entonces es interesante ver cuántos días suele permanecer en la página las personas que realizaron al menos una compra en Trocafone.

Distribución de los días que permanecen en la página las personas que compraron



Se puede apreciar que la gran mayoría de las personas que realizan una venta no se queda más de 20 días, pero a su vez la distribución de la gente que compró que se quedó entre 20 y 100 días es casi constante, lo que muestra que el interés en la página de las personas que compraron no es algo que decaiga rápidamente.

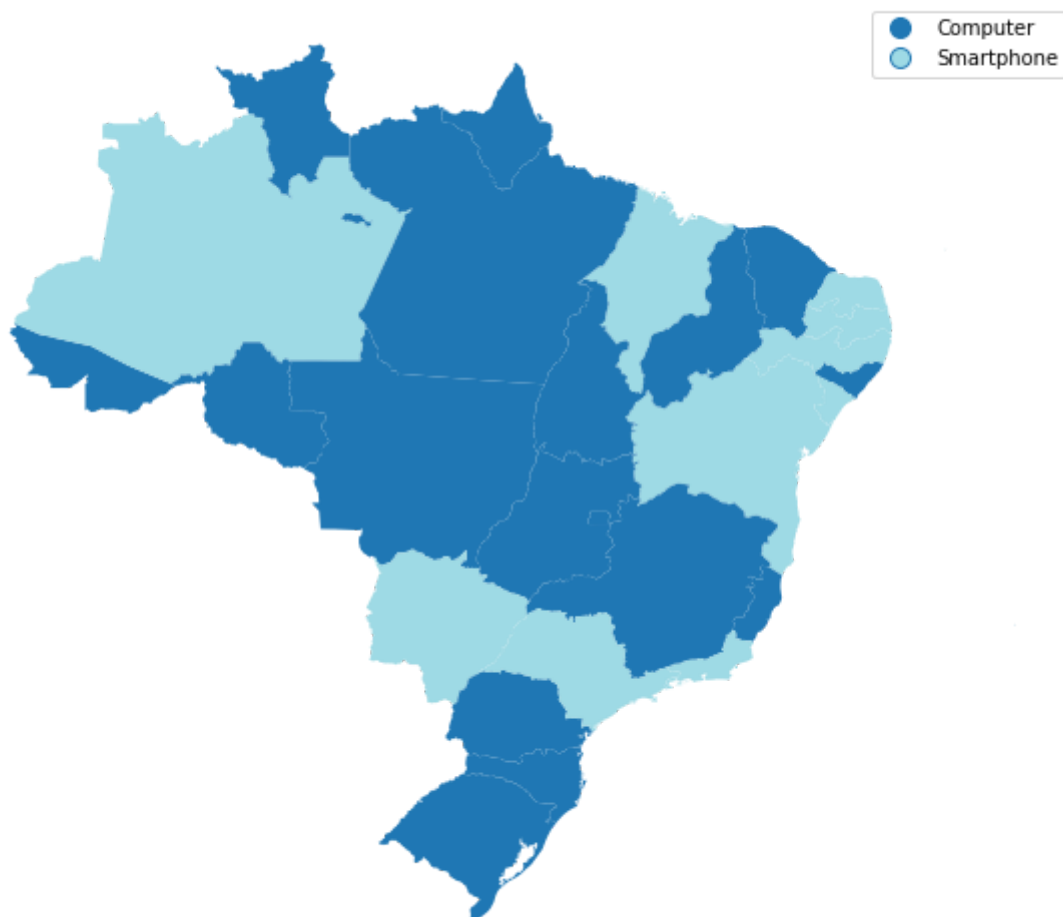
6. Análisis Geográfico de los datos

Dentro de esta sección, se podrán evaluar diferentes características del set de datos referidas al lugar donde ocurren los eventos. Como la gran mayoría de los datos son referidos a Brasil, los consideraremos como el 100% en nuestro caso.

6.1 Dispositivo más utilizado según la región

Conociendo los dispositivos en los que los usuarios entraron al website, se puede analizar desde qué parte de Brasil los usuarios de Trocafone utilizan más sus PC o sus smartphone

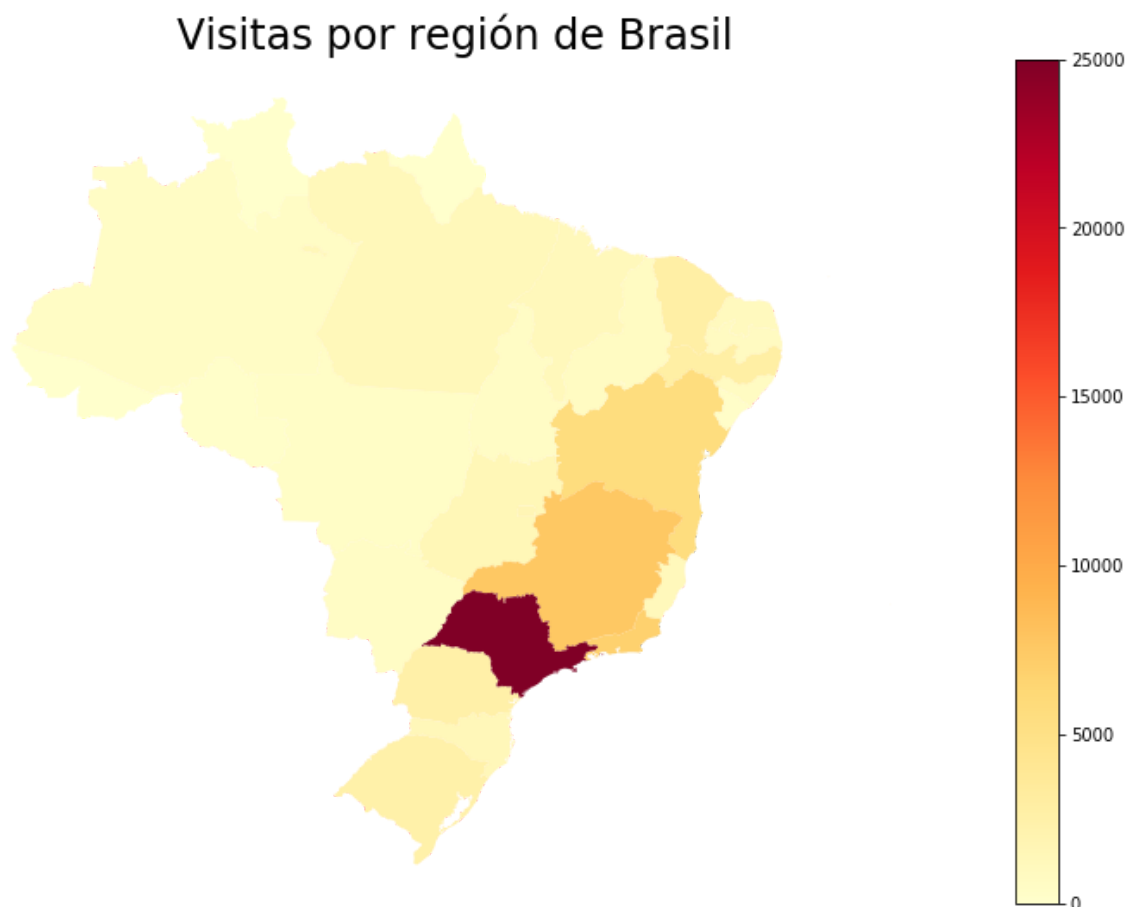
Dispositivo más utilizado según región



Como puede apreciarse en el la visualización del mapa, es pareja la preferencia que tienen los usuarios respecto de cuál dispositivo ingresar al sitio.

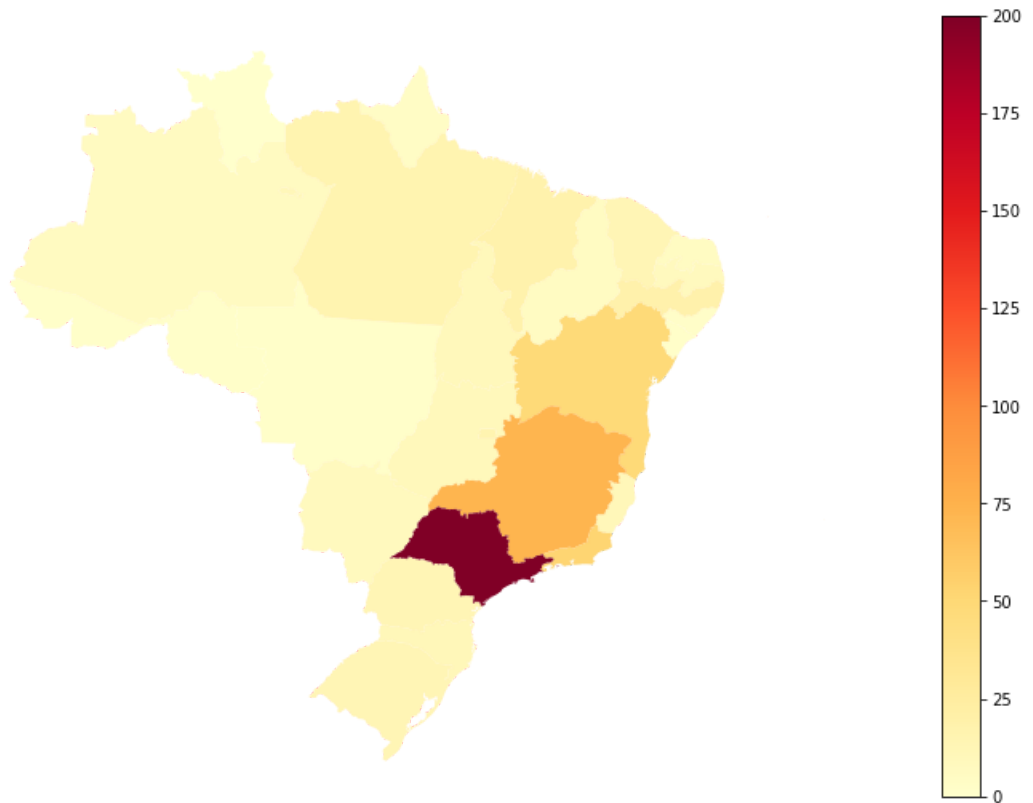
6.2 Visitas según la región

Conociendo el origen, de la ciudad o región, desde donde se originó la visita, se puede llegar a diversas conclusiones relacionadas a la densidad de población o los estratos sociales y segmentaciones económicas dentro de Brasil



Además se puede analizar la cantidad de compras según la región, la cual, como quedó demostrado en la sección 2 acerca de eventos no debería diferir mucho respecto de donde hay más visitas, ya que se demostró que a mayor cantidad de visitas son mayores las compras de productos.

Compras segun region de la primer visita al sitio de la persona

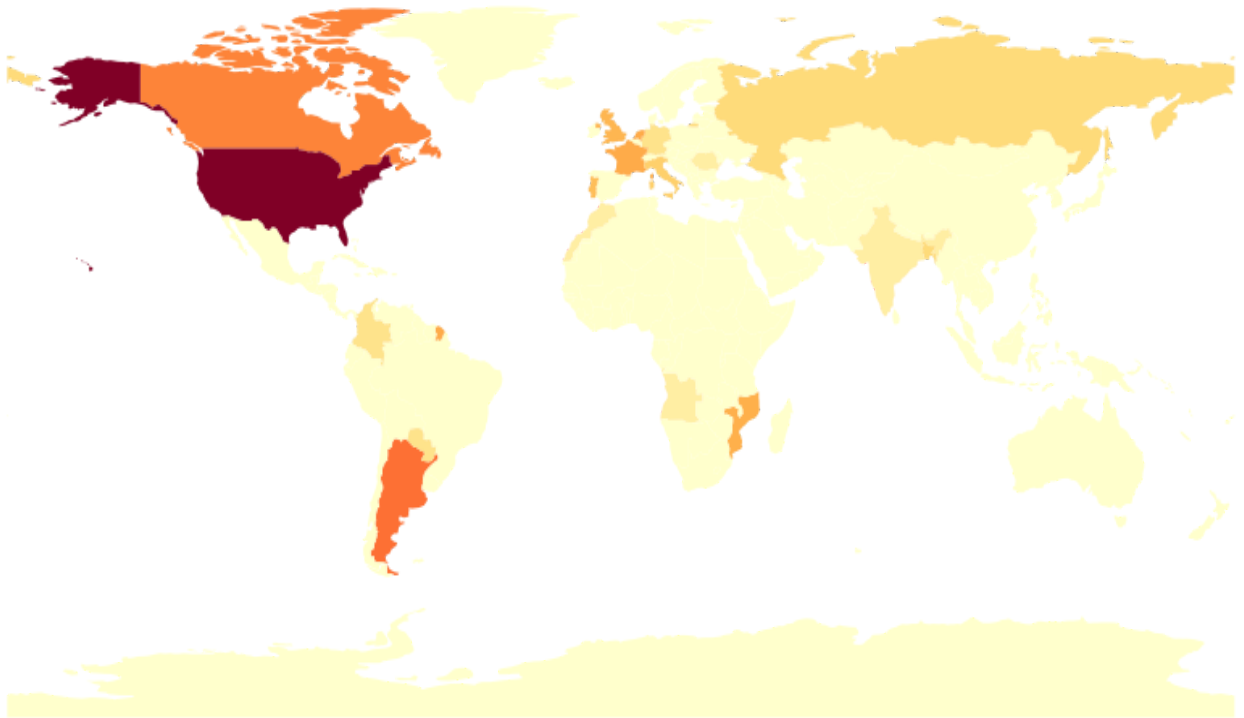


Efectivamente, la distribución de visitas y compras es parecida a posteriori según las personas que entraron por primera vez al sitio.

6.3 Otras visitas fuera de Brasil

Dentro del set de datos hay registradas otras vistas desde otras partes del mundo, las cuales son interesantes para hacer un posible análisis respecto de si hay personas interesadas en comprar en la plataforma

Visitas por país (log)



Puede verse que hay algunas visitas originadas desde Estados Unidos, Canadá y Argentina.

De nuevo, se utilizó la escala logarítmica para resaltar las diferencias entre los valores altos y bajos, para lograr una mejor visualización de los datos.

7. Análisis de la interacción con anuncios web

Una gran cantidad de gente conoce la página e ingresa por primera vez debido a los anuncios que aparecen en distintas páginas web, lo que genera un mayor tráfico de personas que visitan la página que son compradores potenciales.

7.1 Clicks en ads según la hora del día

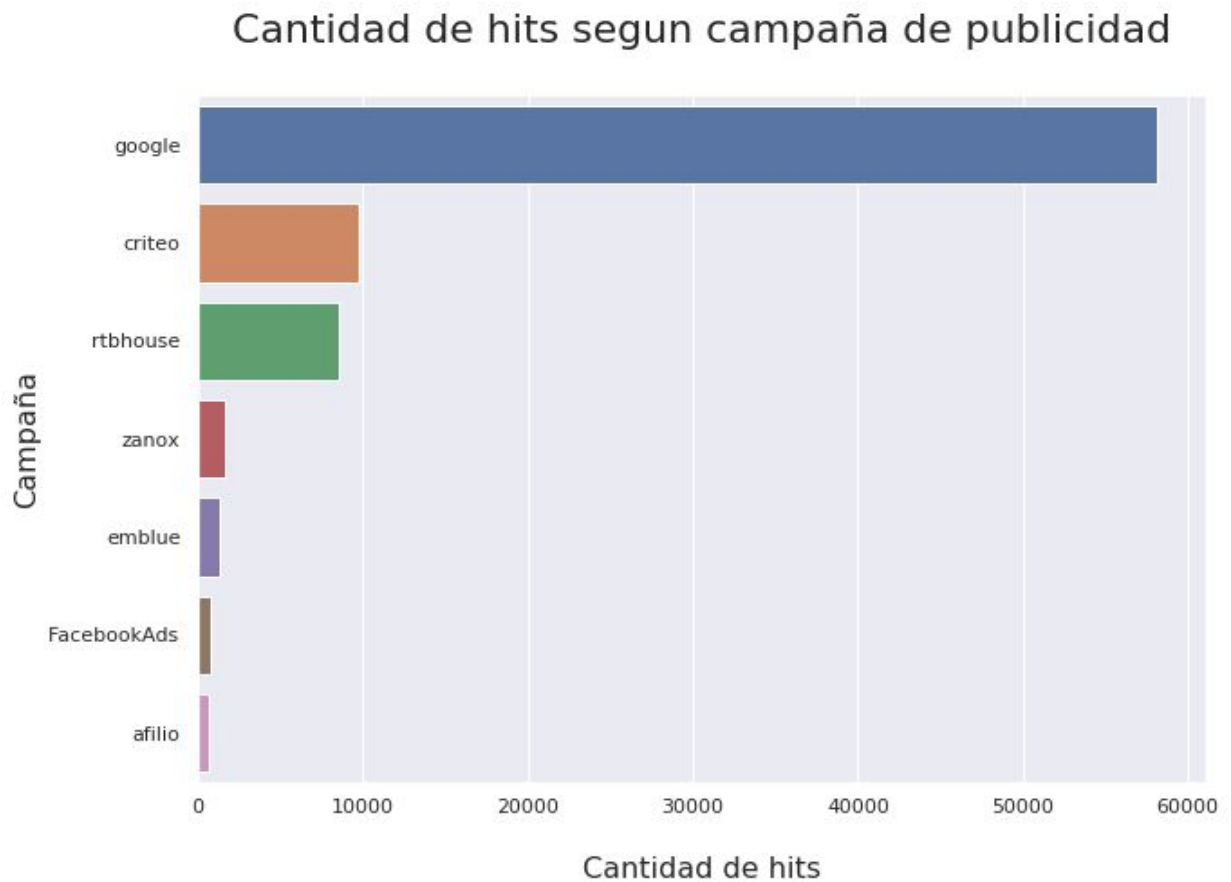
Considerando al tiempo como continuo queda así el gráfico que muestra la cantidad de gente que entró a la página debido a un click en un anuncio en función de la hora del día.



Esto muestra que el tráfico generado debido a interacciones con anuncios se da en mayor proporción entre las 14 y las 0 horas, horario en el que la gente en general empieza a estar más desocupada de sus tareas laborales.

7.2 Cantidad de clicks según la campaña de publicidad

Los clicks o hits sobre los ads que generan tráfico en trocáfone son mostrados a las personas que navegan la web según distintas campañas de publicidad que muestran justamente los anuncios a los usuarios de distintas páginas de la red. Estas son las que generaron mayor tráfico sobre la página.



Siendo la campaña de google la más exitosa en relación a la cantidad de clicks totales generados.

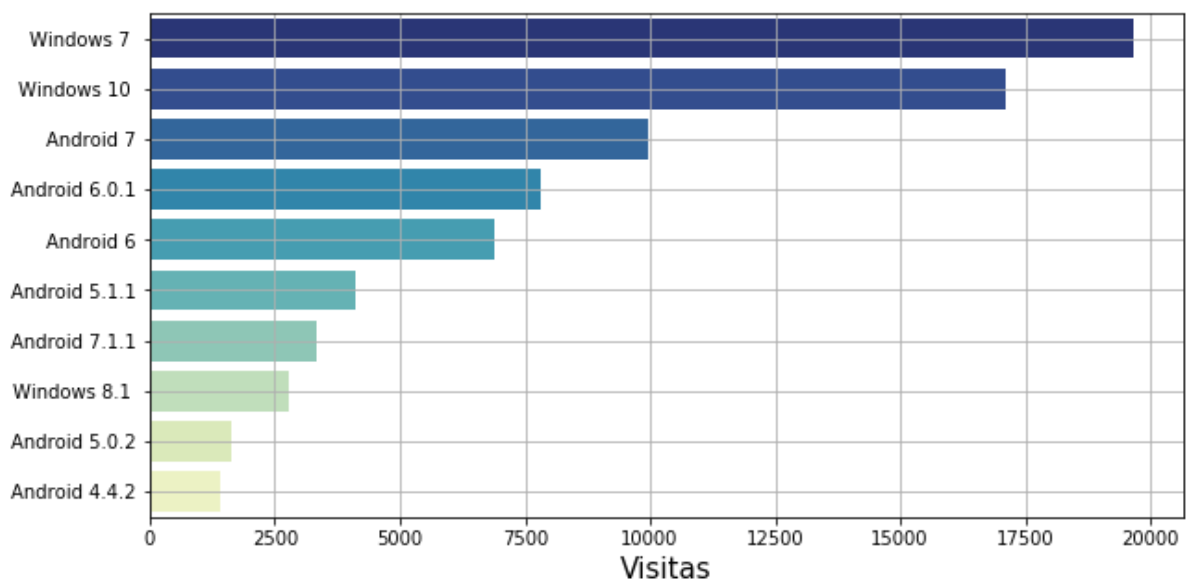
8. Sistemas operativos de los usuarios

En esta sección se evaluarán las variables relacionadas al sistema operativo y al comportamiento de los usuarios provenientes de ellos con respecto a las compras.

8.1 Visitas a la página según Sistema Operativo

Como primera instancia vamos a analizar desde qué sistemas hay más ingresos al sitio.

Visitas según sistema operativo de origen

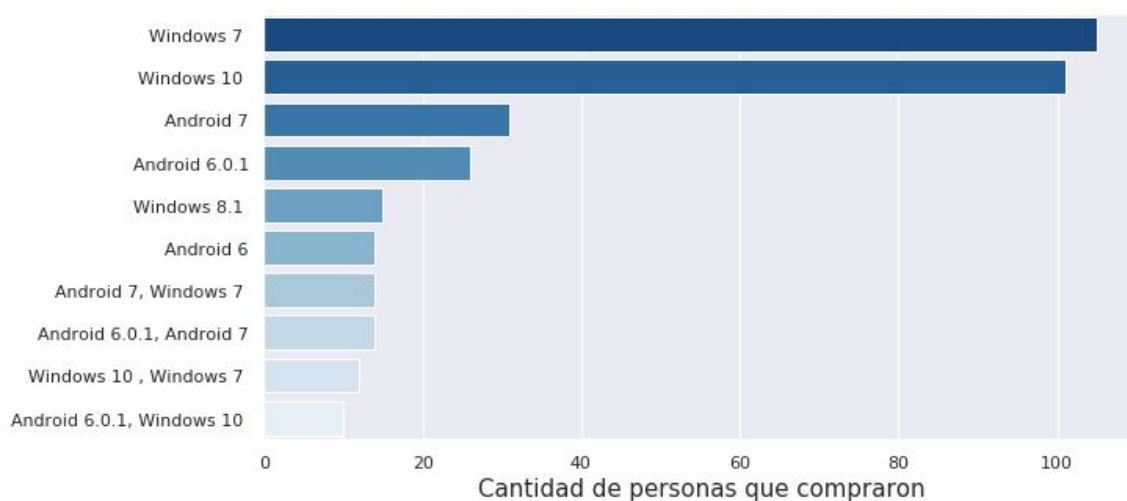


Podemos ver que los sistemas más utilizados entre los usuarios son Windows 10 y Windows 7, es decir, sistemas utilizados en computadora.

8.2 Cantidad de conversiones según sistemas utilizados para ingresar al sitio.

Para realizar un análisis de las conversiones debemos tener en cuenta que una persona puede haber ingresado al sitio desde distintos dispositivos, por lo tanto vamos a clasificar a las personas según desde cuáles de ellos ingresó.

Cantidad de personas que realizaron una conversión
vs Sistemas desde los que ingresó



Aquí podemos ver que muchas personas utilizaron únicamente una computadora con Windows 10 o Windows 7, y que lideran en conversiones realizadas. Además se observa que en la mayoría de las categorías del ranking aparecen estos sistemas. Por otro lado, también se nota la presencia de los sistemas Android 6 y 7 pero las conversiones que realizan las personas que ingresaron por estos dispositivos están muy por debajo de Windows.

8.3 Cantidad de conversiones según sistemas utilizados para ingresar al sitio.

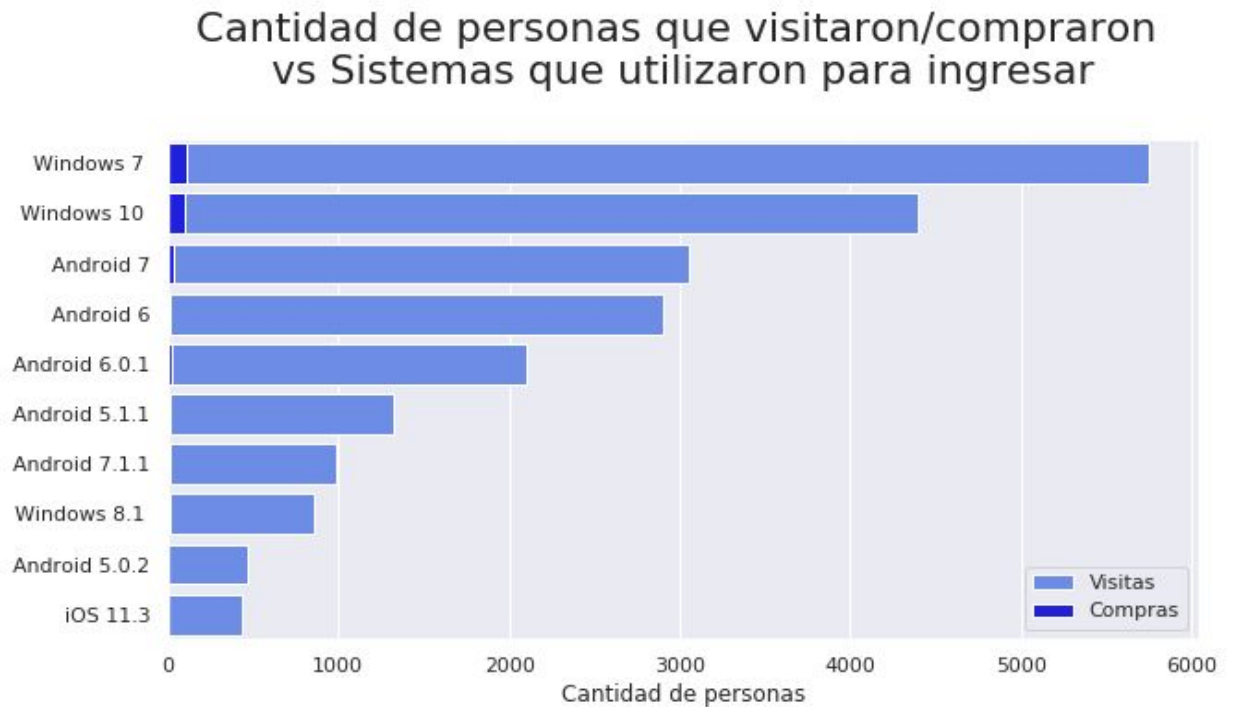
Si hacemos un análisis similar al anterior pero esta vez considerando las visitas al sitio según los dispositivos utilizados obtenemos el siguiente gráfico.



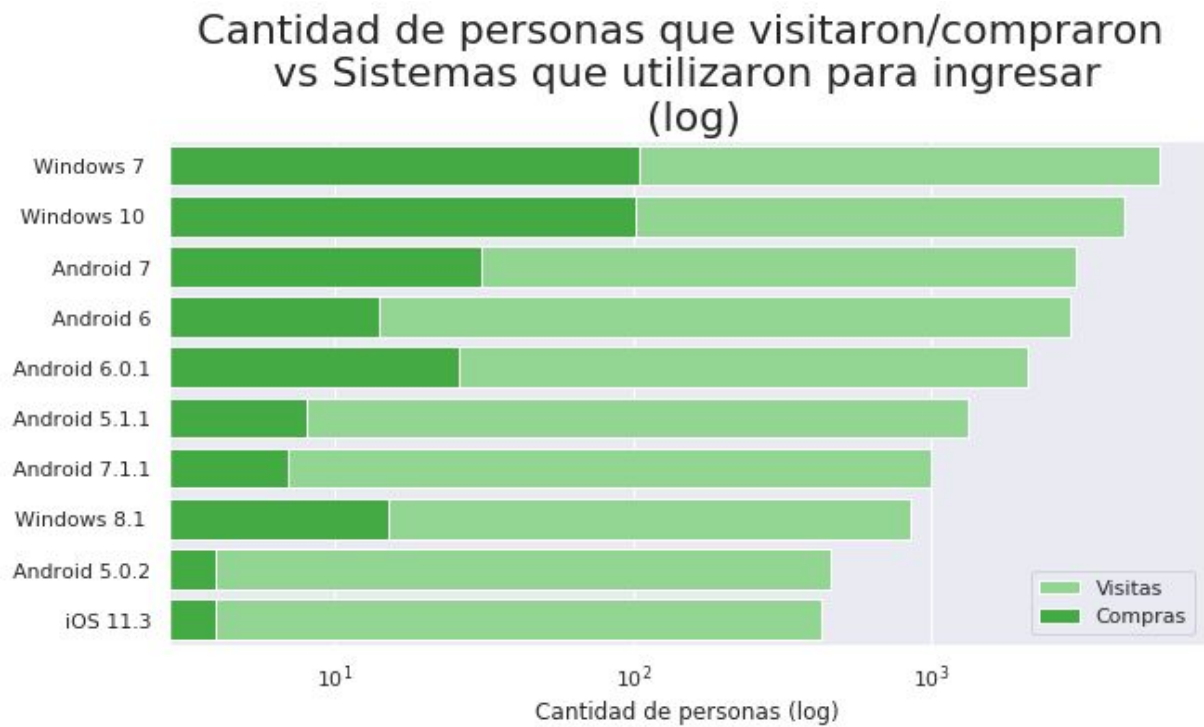
En él podemos observar que se mantiene el liderazgo de los sistemas Windows pero esta vez la diferencia no es tan pronunciada. Incluso se ve que aparecen en el ranking personas que ingresaron únicamente desde dispositivos Android e iOS, las cuales no se encontraban en el ranking anterior. Podemos decir que realizaron más compras las personas que decidieron ingresar al menos una vez desde Windows. En cambio los que ingresaron únicamente desde Android o iOS realizaron menos compras.

8.4 Comparación entre visitas y conversiones

En el siguiente gráfico analizamos la relación entre visitas y conversiones dentro de la misma categoría de sistemas utilizados



Al ser tan bajo el número de conversiones en relación a las visitas podemos realizar el gráfico en escala logarítmica.



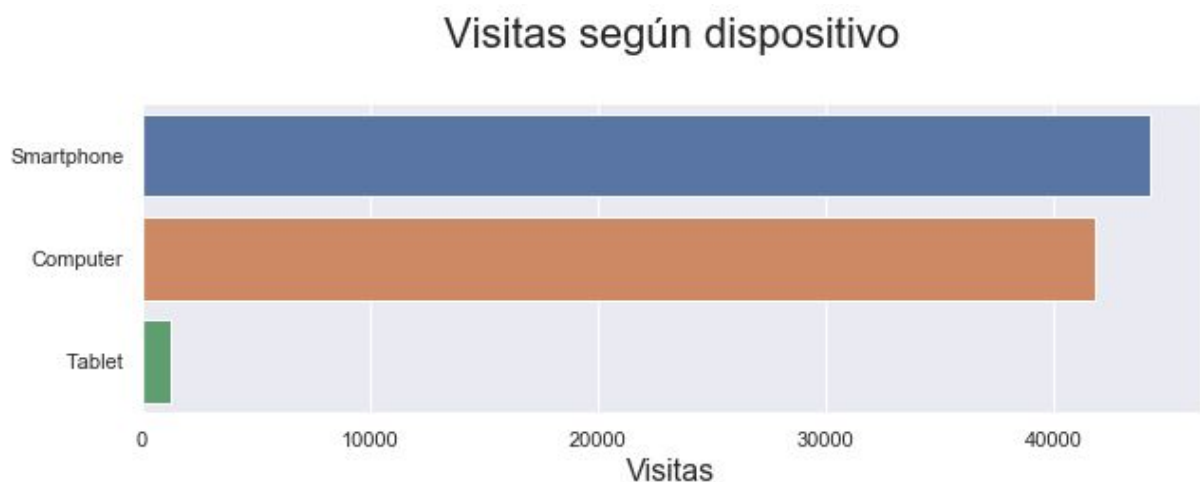
Estos últimos dos gráficos coinciden con los resultados vistos anteriormente, por ejemplo se puede ver que Windows 8.1 se encuentra muy abajo en visitas pero en compras supera a otros sistemas Android.

9. Tipo de dispositivo

El análisis del tipo de dispositivo utilizado está relacionado con el sistema operativo utilizado y por lo tanto es de esperarse que se obtengan resultados similares.

9.1 Cantidad de visitas según dispositivo

Si obtenemos el total de visitas según dispositivo utilizado podemos obtener lo siguiente.

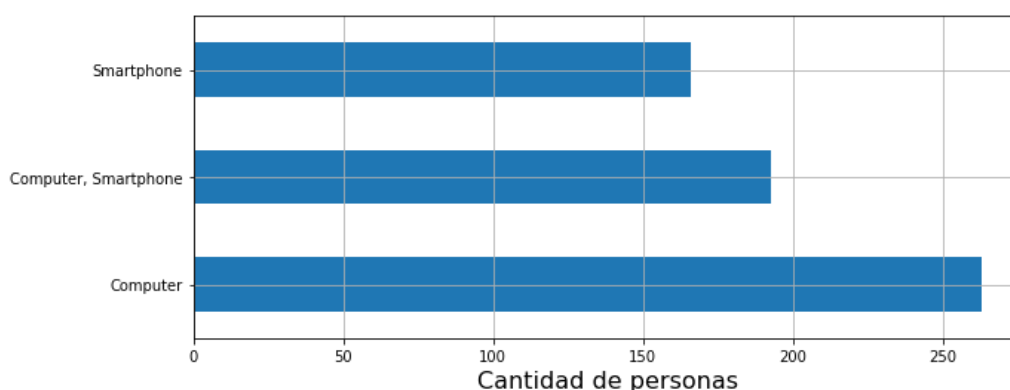


Aquí observamos que las visitas realizadas desde una computadora y desde un celular son similares aunque es ligeramente superior en el último. Las visitas desde tablet son insignificantes por lo que vamos a excluirlas del análisis.

9.2 Compras según dispositivos utilizados

Aquí nuevamente debemos agrupar a las personas según los distintos dispositivos que utilizó la persona ya que este puede ser más de uno.

Cantidad de personas que compraron segun dispositivos utilizados para ingresar

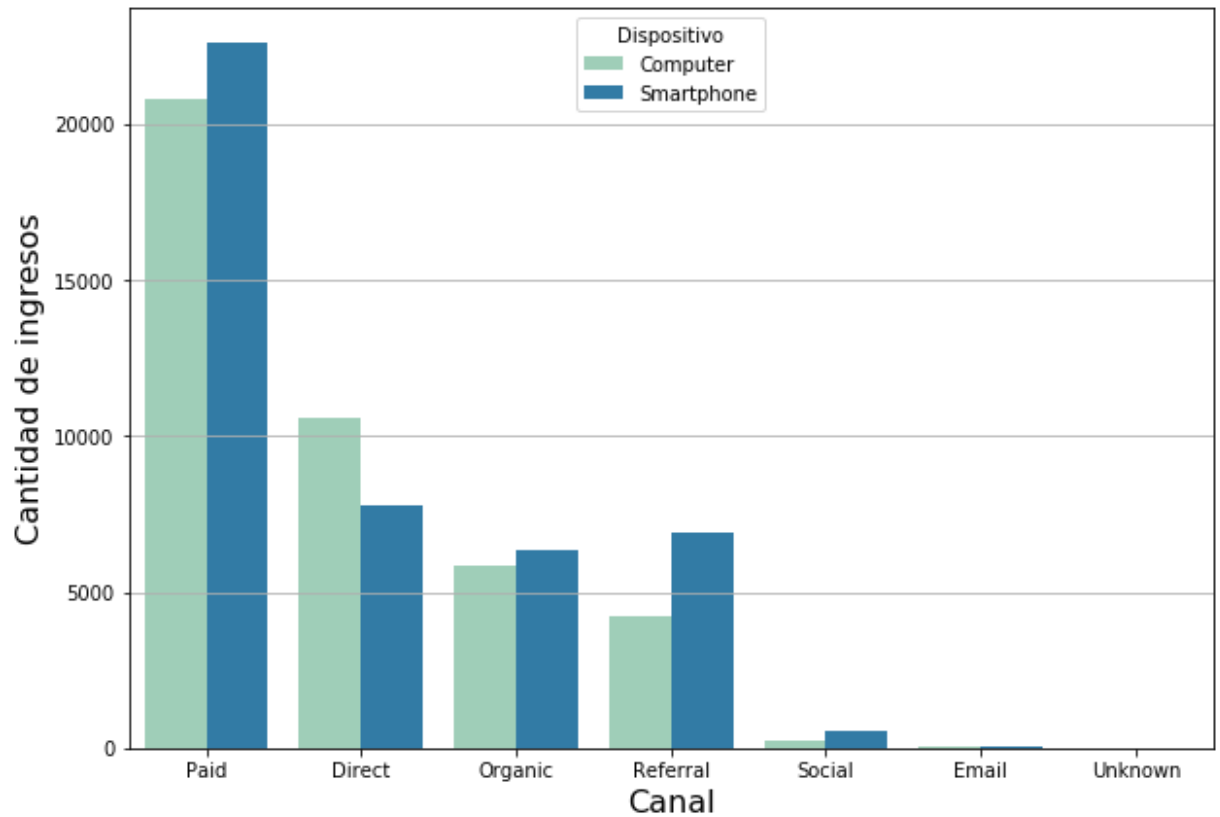


En concordancia con lo observado en el análisis de Sistema Operativo podemos observar que las compras que realizaron las personas que utilizaron una computadora fueron mayores aunque haya ligeramente menos visitas desde ellas. Aquí se puede evidenciar una diferencia en el comportamiento de las personas que ingresan por distintos tipos de dispositivo, ya que en la cantidad de compras es mayor en computadora que en smartphone aunque en estos últimos hay más visitas al sitio.

9.3 Canales según dispositivo

Otra manera de diferenciar el comportamiento de los usuarios según el dispositivo es el canal desde el que ingresan.

Ingresos según canal según dispositivo



En este gráfico se puede ver que en general las visitas desde smartphone son mayores en desde cualquier canal de tráfico, aunque en el canal Direct se ve una diferencia ya que hay más ingresos al sitio por ese canal en computadoras.

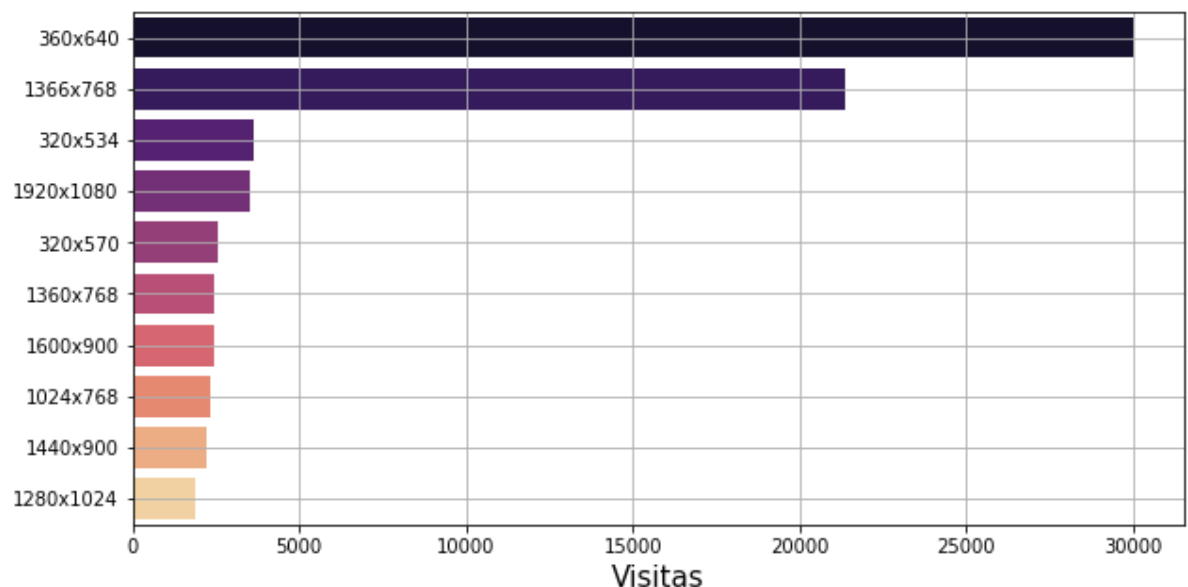
10. Análisis según resolución de la pantalla

La resolución de la pantalla del dispositivo utilizado para ingresar a una página de venta de celulares puede dar una idea del tipo de tecnología que maneja la persona que desea realizar una compra o informarse sobre precios y características de los dispositivos a la venta en Trocafone.

10.1 Vistas según resolución del dispositivo

Lo primero a analizar será cuántas vistas se realizan a los distintos productos de la página dependiendo de la resolución del dispositivo de las personas que la visitan.

Visitas según resolución del dispositivo



Siendo las predominantes 360x640 (resolución típica en dispositivos móviles) y 1366x768 (resolución típica en computadoras de escritorio y notebooks discretas).

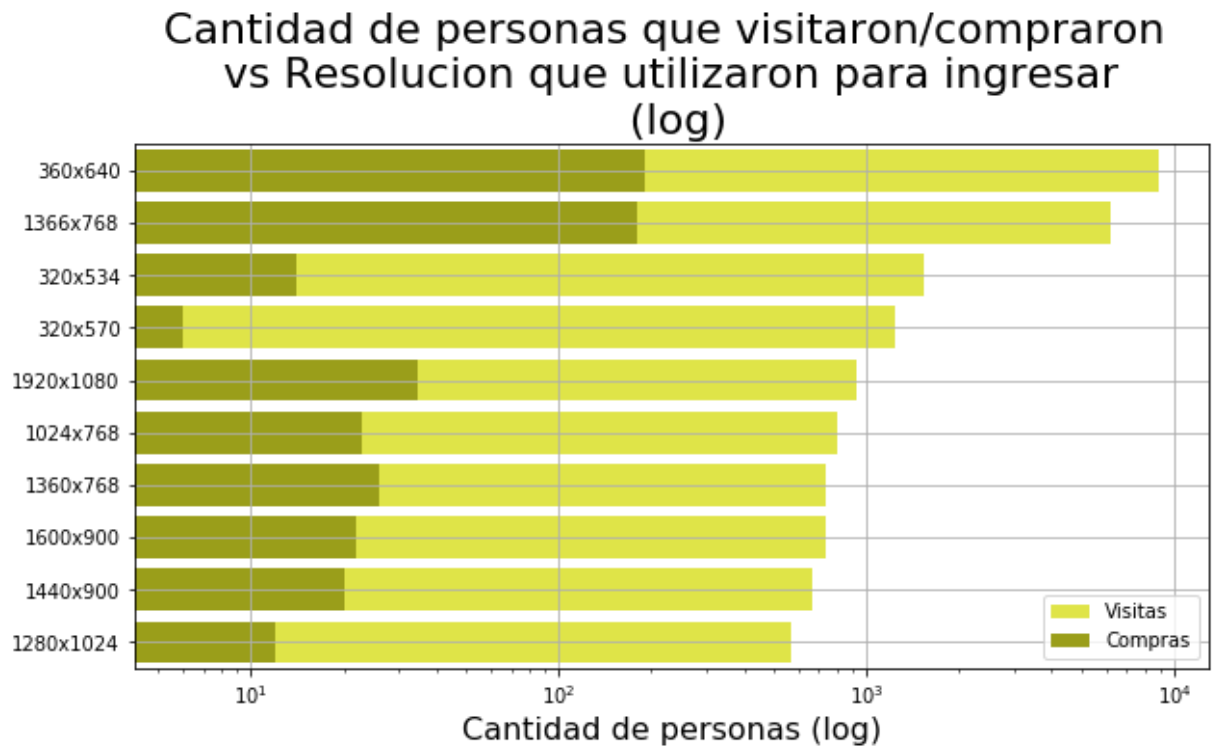
10.2 Cantidad de personas que visitaron/compraron vs. Resolución que usaron para ingresar

Luego, la cantidad de vistas de los productos puede ser comparada con la cantidad de ventas que realizaron las personas que ingresaron a la página con una resolución específica.



Este gráfico muestra en una escala real la relación entre las compras y las vistas de productos en función de la resolución del dispositivo usado, pero dado que la diferencia es tan grande conviene usar una escala logarítmica para poder apreciar mejor visualmente las relaciones entre vistas/compras.

10.2.1 Cantidad de personas que visitaron/compraron vs. Resolución que usaron para ingresar (En escala logarítmica)



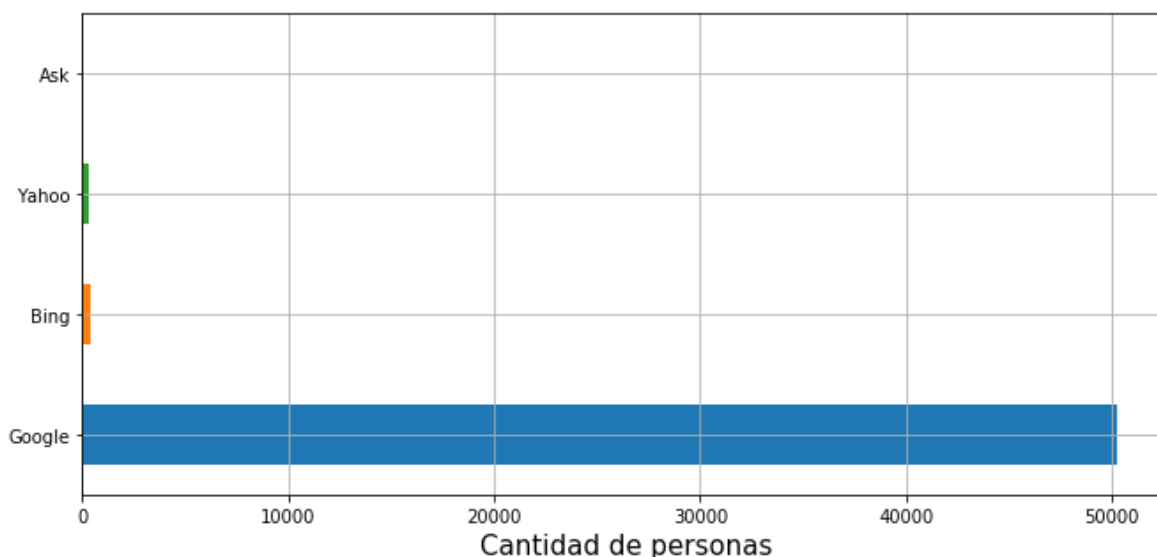
11. Análisis de los motores de búsqueda

Un resultado interesante del set de datos es saber como un usuario llegó a Trocafone, no siendo parte de un aviso publicitario, sino al intentar por sus propios medios buscar a través de los motores de búsqueda conocidos.

11.1 Cantidad de visitas según el motor de búsqueda

Los motores de búsqueda presentan las siguientes cantidades de entradas al sitio.

Cantidad de visitas según motor de busqueda

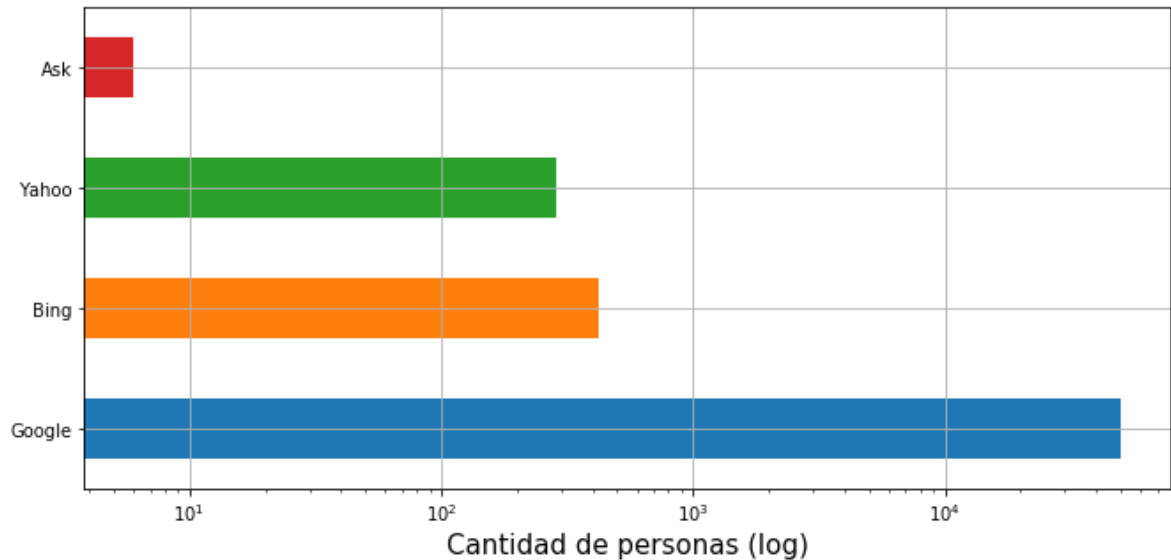


Como es de esperarse, debido a la popularidad de los buscadores fueron tomados los datos, Google presenta una superioridad numérica importante respecto de los otros buscadores, lo cual dificulta la visualización de una forma que permita comparar el resto de los motores de búsqueda.

Nuevamente puede ayudar ver el gráfico usando una escala logarítmica.

11.1.1 Cantidad de visitas según el motor de búsqueda (log)

Cantidad de visitas según motor de búsqueda (log)



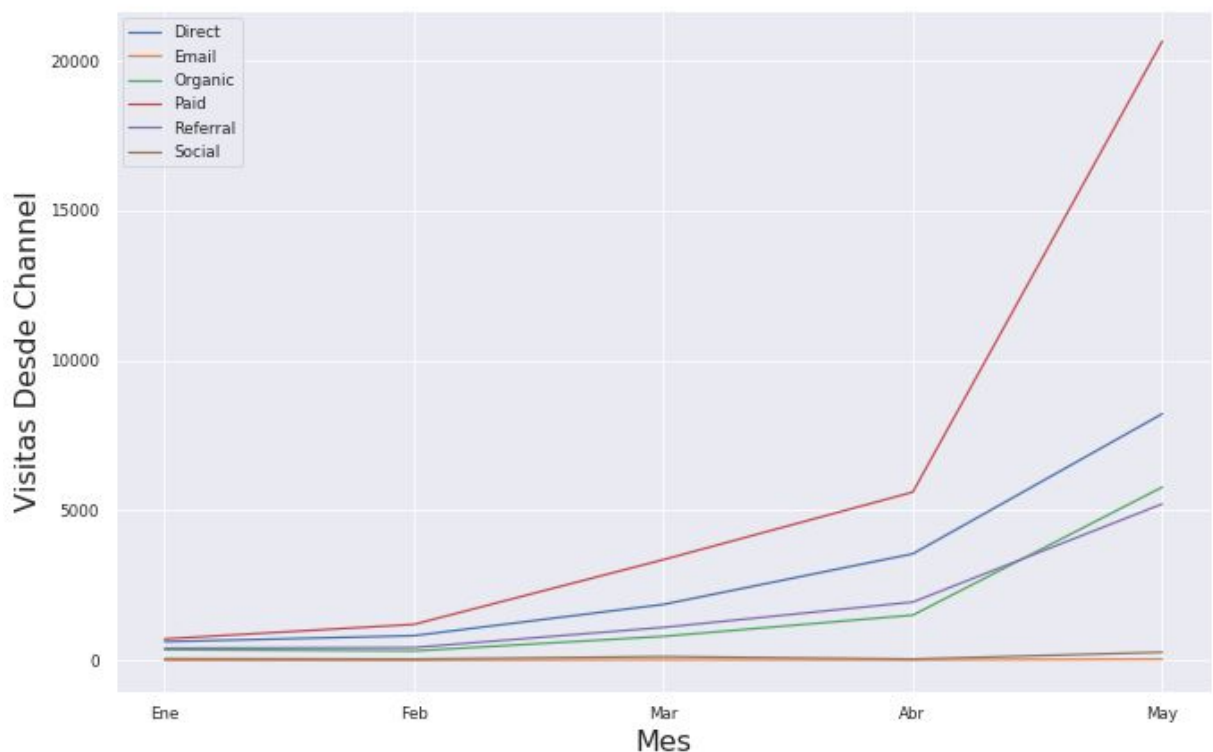
12. Análisis del canal origen de eventos

La última pieza de este análisis la compone el canal desde donde se origina un evento.

12.1 Visitas según el canal que las originó en función del tiempo

El canal desde donde se originó una visita a la página da una idea de las fuentes del tráfico en la página, el cual puede ser analizado como se desarrolla en el tiempo a través de los meses.

Visitas segun canal que las origino, evolucion en el tiempo



Puede verse que conforme pasa el tiempo, así como aumentó el tráfico en la página (puede verse en la sección temporal como con el paso de los meses aumenta considerablemente el número de eventos en Trocafone),

también aumentó la cantidad de gente que entró por los distintos canales, especialmente el canal Paid (que se refiere al tráfico que resulta de los resultados que aparecen como anuncios al buscar en distintas plataformas como Google) y el canal Direct (que es cualquier tipo de tráfico referido por una fuente desconocida).

12.2 Cantidad de personas que visitaron por primera vez el sitio según canal

Además de realizar un análisis temporal sobre cada canal, también resulta útil medir la cantidad total de gente que entra por primera vez a la página desde un determinado canal.



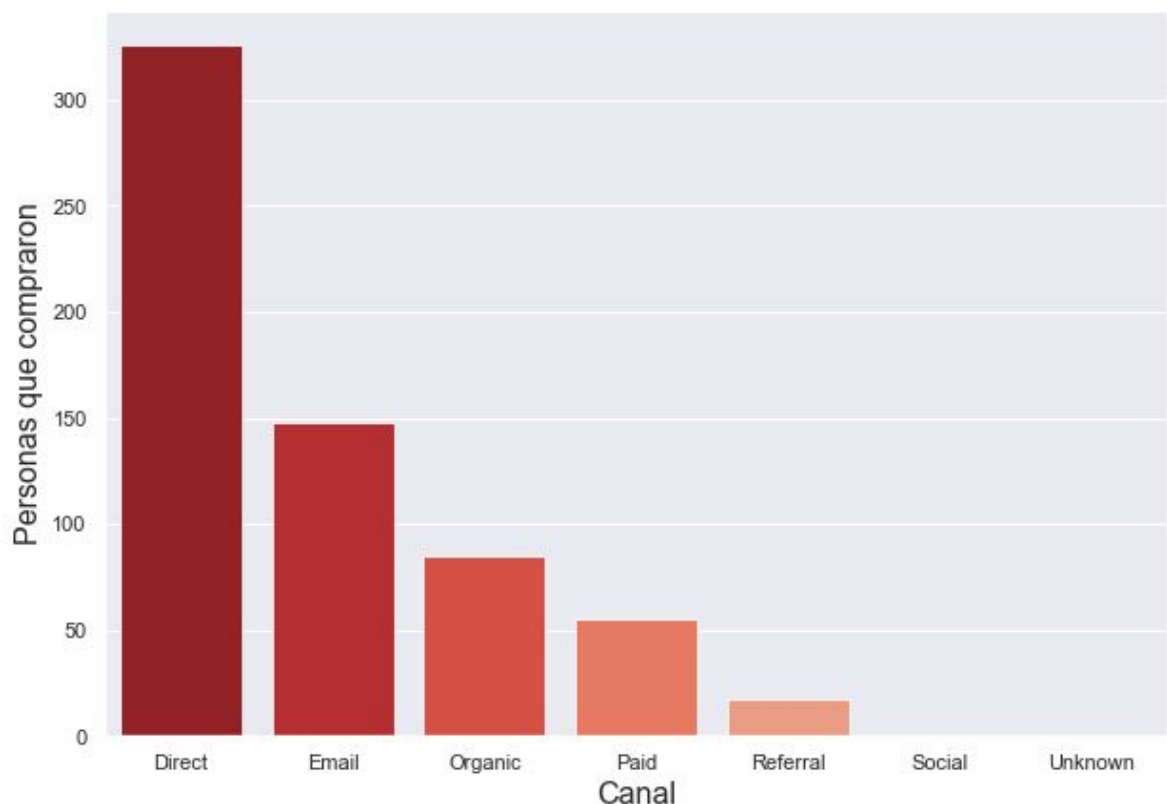
Este gráfico en general se corresponde con el mostrado anteriormente ya que el canal Paid que es el que mostraba el mayor crecimiento en el tiempo es el que suma la mayor

cantidad total, aunque puede apreciarse que la correspondencia no es exacta, pues el canal organic (Tráfico generado por resultados en motores de búsquedas que no fueron pagados) es en total mayor que el de direct y esto se debe a que solo se está considerando la primera visita de cada persona en la página.

12.3 Cantidad de personas que compraron según desde qué canal ingresaron al sitio

Finalmente de las personas que visitaron la página desde algún determinado canal es interesante quedarse con las que realizaron compras en el sitio.

Cantidad de personas que compraron segun desde donde ingresaron al sitio



Este gráfico muestra que de la gente que compró que ingresó al sitio desde un determinado canal, fue en su gran mayoría gente que entró desde el canal Direct, el cual se

refiere a que se desconoce la fuente del origen del enlace referido a la página desde el que entró una cantidad del tráfico. También puede apreciarse que aunque el canal paid sea el que más visitas genera, es uno de los que menos compras tiene.

13. Conclusiones finales

- El evento que tiene mayor cantidad de ocurrencias es visited site
- El flujo de compras resulta muy bajo, muy pocos de los que entran terminan comprando
- Si bien hay más ingresos desde dispositivos móviles, las compras suelen ser realizadas por personas que ingresan desde computadoras.
- La gran mayoría de los datos se corresponden a Brasil.
- Todos los usuarios realizaron al menos un checkout
- Los modelos más vendidos son Samsung Galaxy J5 y iPhone 5, pero los más vistos son iPhone 6, Samsung Galaxy S7 y S8.
- Los modelos más buscados son los de marcas Samsung y Apple(iPhone).
- Durante la tarde la gente suele entrar mas a Trocafone por Ads.
- La actividad en el sitio dio un salto importante en el mes de mayo, previamente la actividad en la página era mucho menor.
- Las horas de mayor actividad en el sitio suelen ser durante la tarde y la noche.
- La resolución más popular de los dispositivos de la gente que ingresa a la página es 360*640, seguida de cerca por la resolución 1366*768.
- El segundo país desde el que más gente realiza actividad en la página es Estados Unidos. Le siguen en menor medida Canadá y Argentina.
- La gran mayoría de gente que utiliza un motor de búsqueda para encontrar la página de Trocafone utiliza Google.
- El canal desde el que más se visitó la página es Paid, sin embargo de la gente que compró la mayor parte ingresó a la página por un canal Directo.
- Cerca de San pablo, la gente suele entrar mas a Trocafone.

- Los fines de semana disminuye la actividad de la página de manera considerable.
- Los usuarios realizan mayormente hasta 10 eventos antes de realizar una conversión.
- El número de mayor ocurrencias de eventos de los usuarios es cuatro