

# 75.06/95.58 Organización de Datos

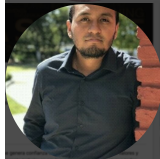
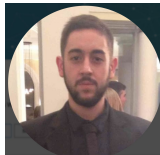
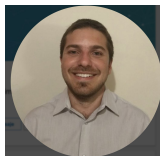
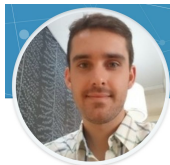
Segundo Cuatrimestre de 2018

Trabajo Práctico 1

Equipo de oyentes: “Rigoxu-Data”

## Integrantes:

- Joaquín Pascual  
joacox86@gmail.com
- Juan Augusto Sanchez  
juanaugusto\_sanchez@hotmail.com
- Matías Moris  
moris.matias@yahoo.com.ar
- Rigoberto Malca  
rigoxu@outlook.com



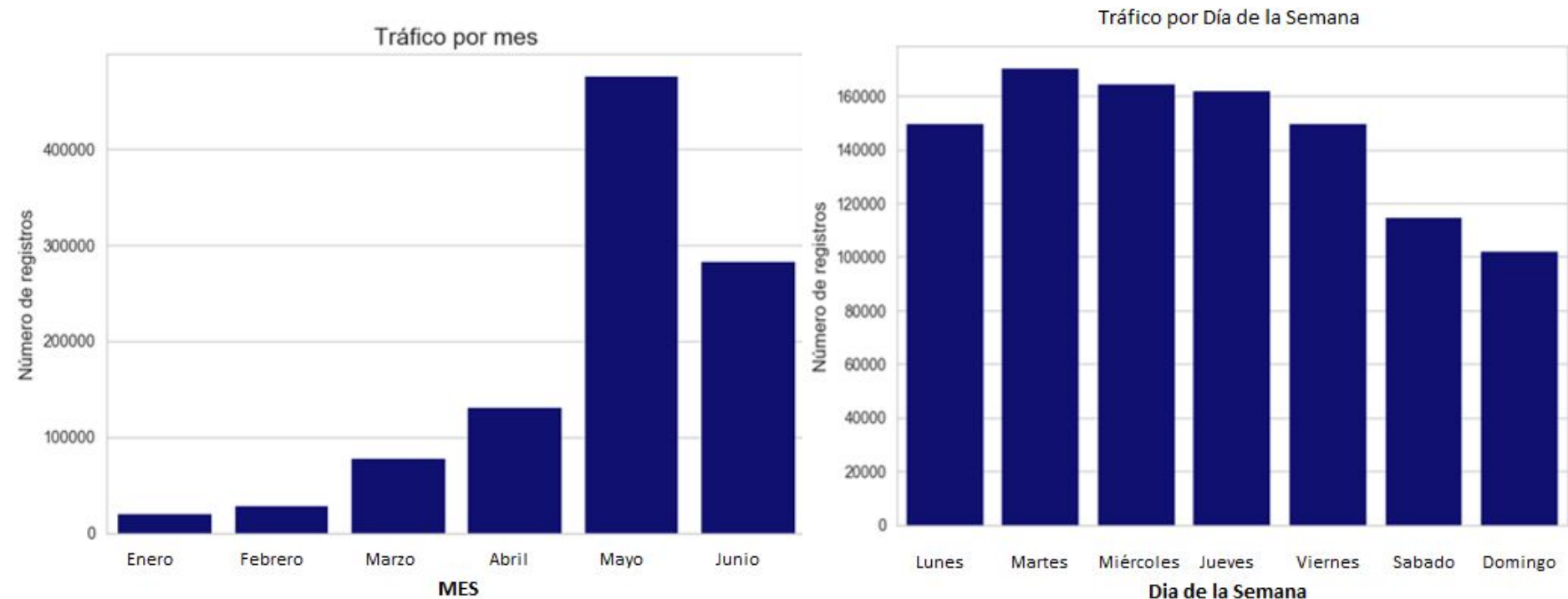
# Nivel de tráfico del sitio

El análisis temporalmente está situado dentro de los meses de Enero y Junio del 2018, puntualmente desde el 01/01/2018 al 15/06/2018.

Algunos datos interesantes que pudimos observar fueron:

- El mes de Mayo fue el de más tráfico, se observa una tendencia creciente
- La menor cantidad de visitas en la página fueron en los horarios matutinos de las 7.00 y 8.00 am, con picos máximos de visitas entre las 23:00 y las 0.00 de la noche.
- Sorpresivamente los fines de semana eran los días menos usados para navegar en el sitio, siendo los Martes y Miércoles los más utilizados.
- Hay 1.172 conversiones que representan el 0.12% del tráfico generado total, a su vez estas conversiones son generados por 716 personas, es decir cerca del 63% compró más de un producto.

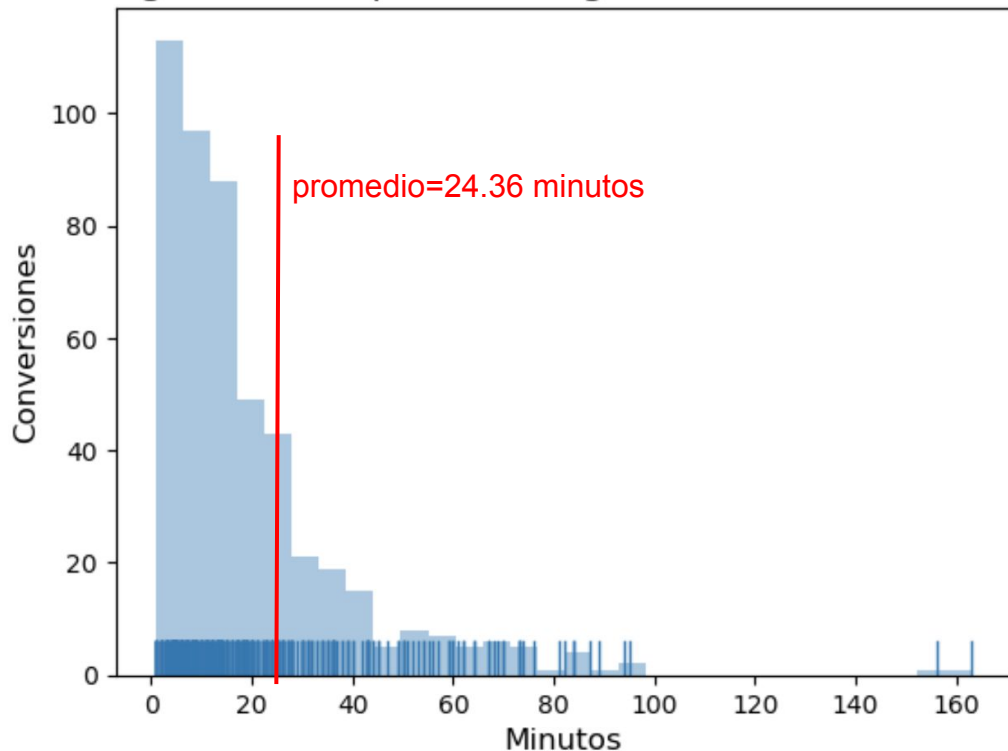
# Tráfico del sitio por mes y por día de la semana.



Para entender el comportamiento general del dataset es importante visualizar el pico de tráfico que se observa durante el mes de mayo y que el tráfico en los días del fin de semana es menor al tráfico de los días de semana.

## Análisis del tiempo que toma la conversión desde el momento que ingresan los usuarios (evento “visited site”).

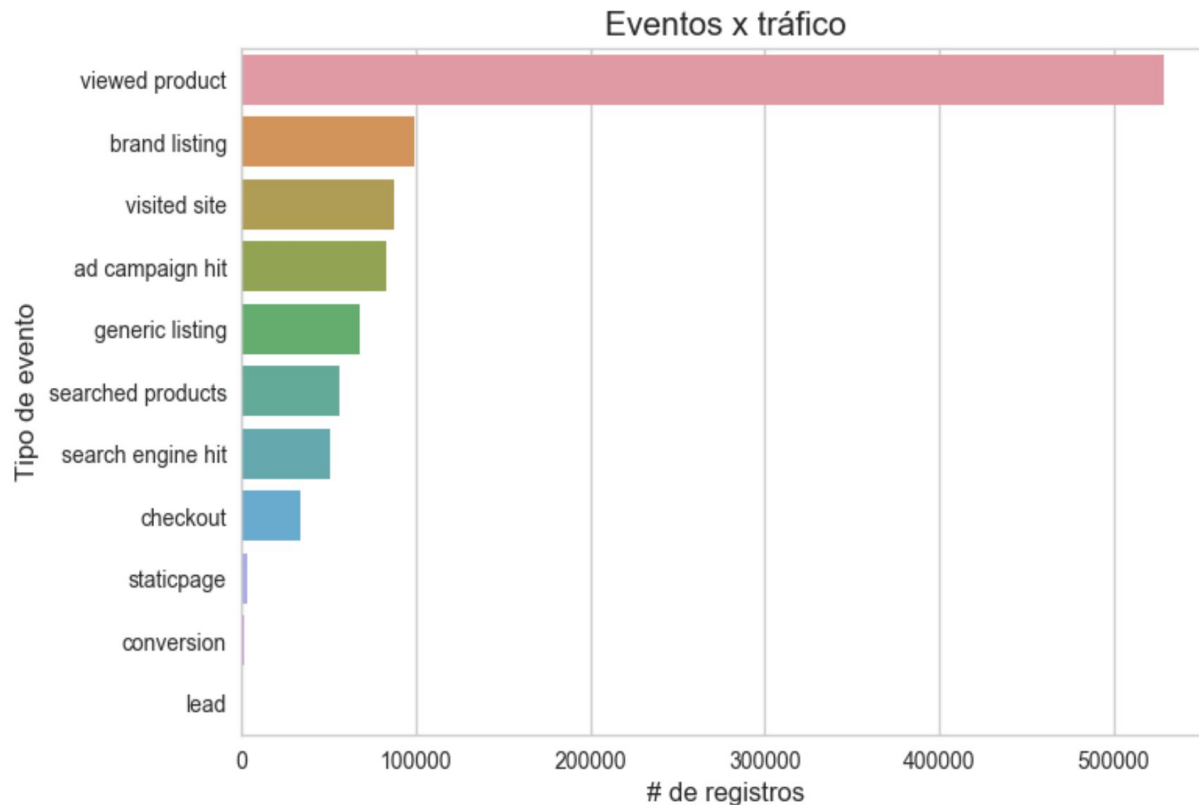
Histograma: tiempo de navegacion hasta la conversion



La mayoría de las conversiones se dan antes de los 20 minutos.

La distribución tiene una asimetría positiva, presentando una media de 24.36 minutos.

# ¿Cuáles son los eventos con mayor frecuencia?



- Puede visualizarse que el evento que más registros genera en el dataset es “viewed product” que representa un 52.30% del tráfico total.

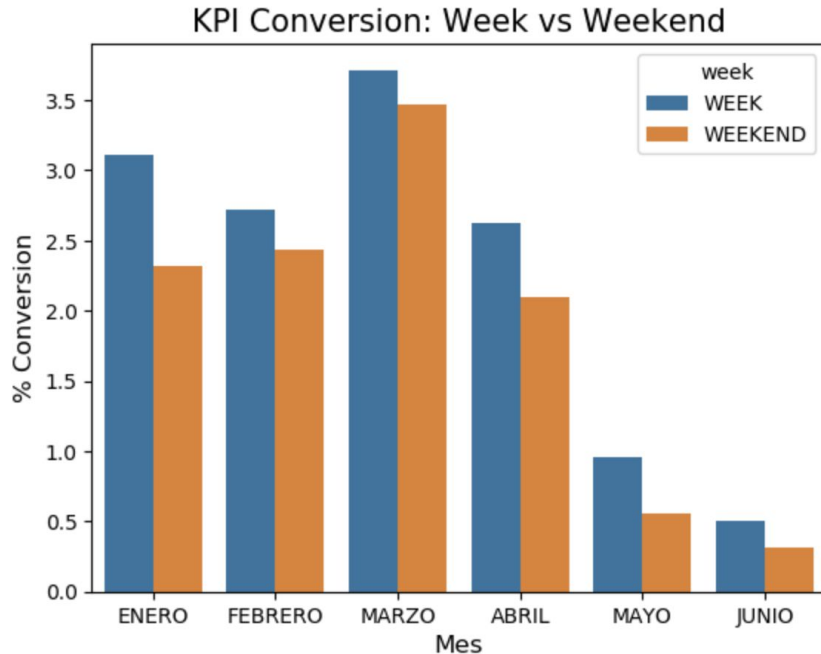
- Por otro lado es importante tener en cuenta que el evento conversión (compra) representa el 0.12 % de todo el tráfico.

**KPI** - Cantidad de eventos por cada conversión:

- Por cada 451.31 viewed products hay una conversión.
- Por cada 28.78 checkouts hay una conversión.

# CONVERSIÓN (Week vs. Weekend)

$$\text{KPI \% Conversión} = \frac{\text{\# Usuarios que convirtieron}}{\text{\# Visitas}}$$

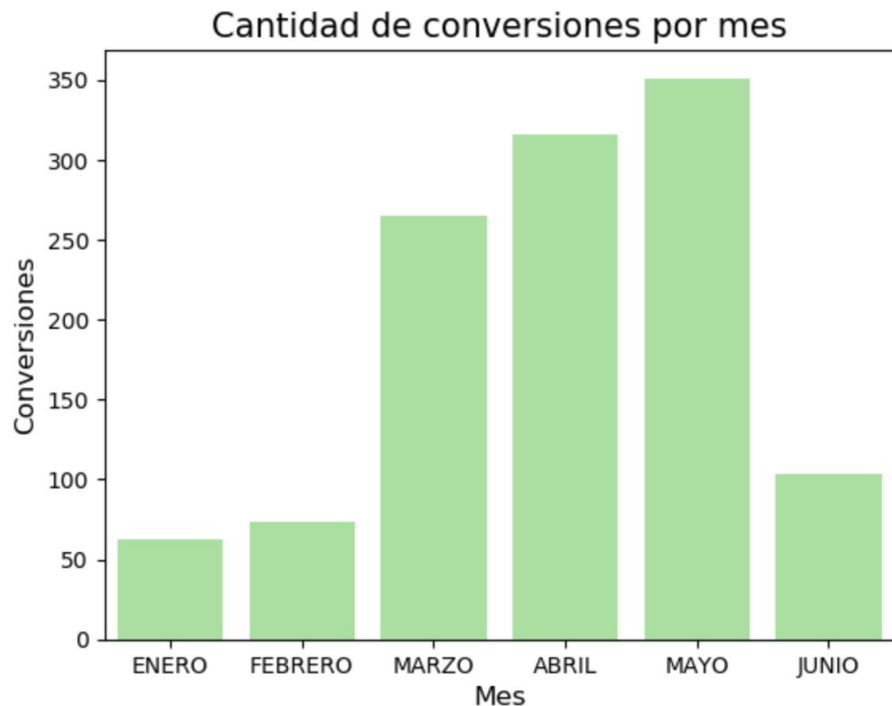


Estudiamos la conversión de los usuarios que navegaron en Trocafone.

Hicimos dos observaciones importantes:

- Los días de semana la conversión es superior a los fines de semana.
- Hay una clara caída de la conversión a partir del mes de marzo.

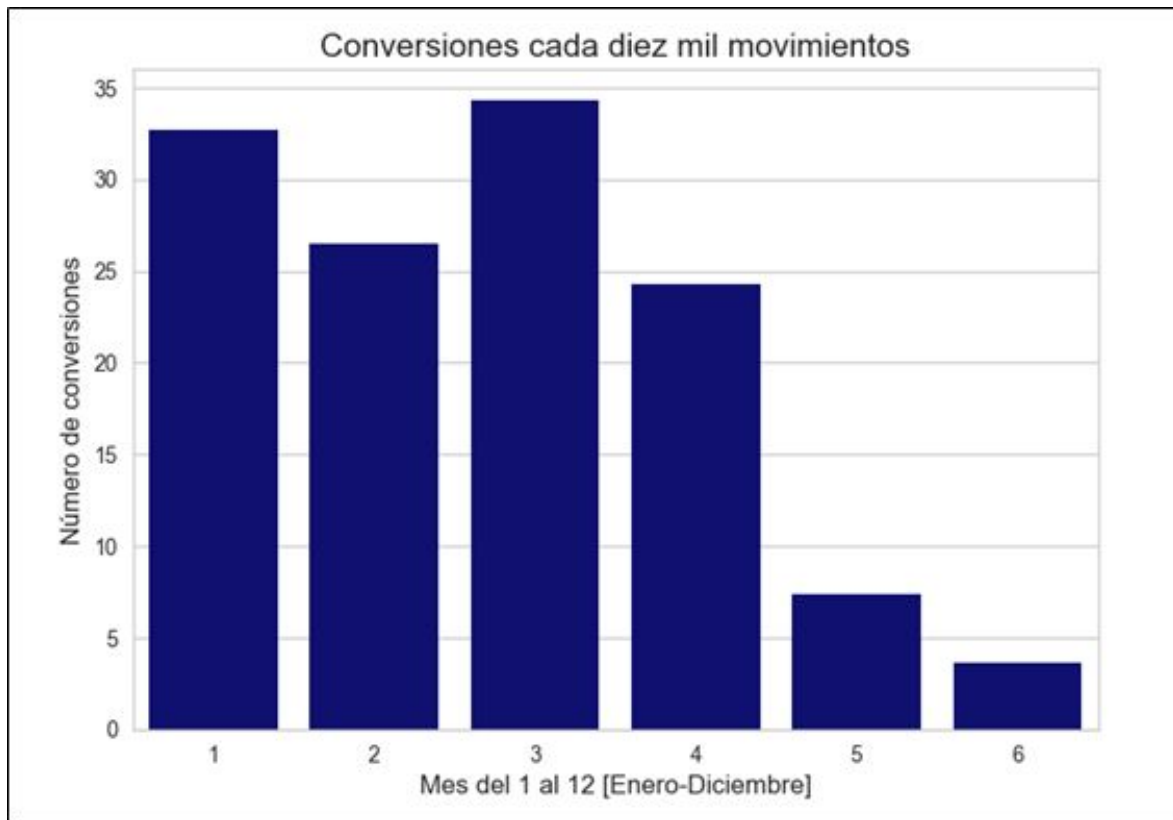
## Nivel de ventas por mes (conversiones).



Acá evaluamos el nivel de ventas por mes de la compañía y vemos con hay un ascenso claro en los meses de Marzo a Mayo.

El descenso en Junio no es re relevante ya que, como mencionamos al principio, sólo contamos con 15 días de información

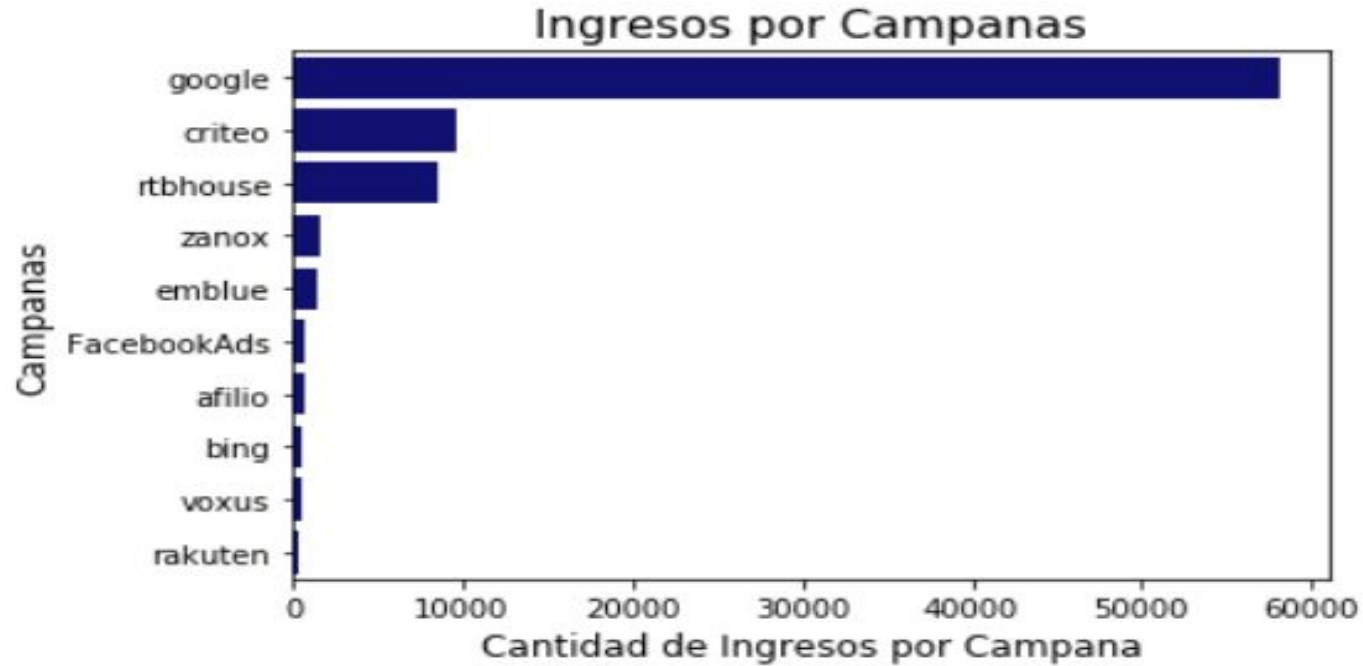
## Nivel de ventas por mes respecto al total del tráfico.



- Por otro lado notamos que en los meses en los que hay menor tráfico (entre enero y abril) es más alta la relación entre cantidad de conversiones y tráfico del mes. De esta manera podríamos decir que en los meses de menor cantidad de movimientos hay mayor efectividad de conversiones.

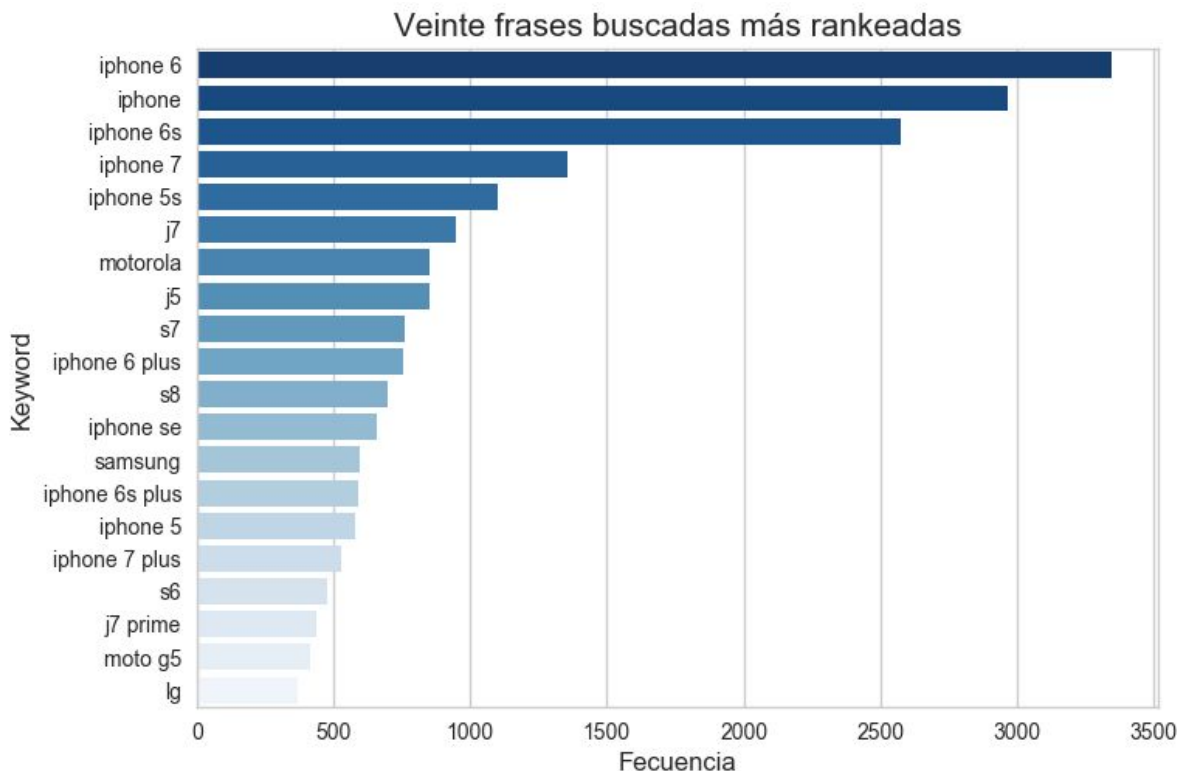


¿Cuántos ingresos generaron cada una de las campañas realizadas? ¿Cuál fue su efectividad?



Puede notarse que las campañas con mayor alcance fueron las de Google, Criteo y rtbhouse.

# ¿Cómo buscan los clientes? - Veinte frases más buscadas

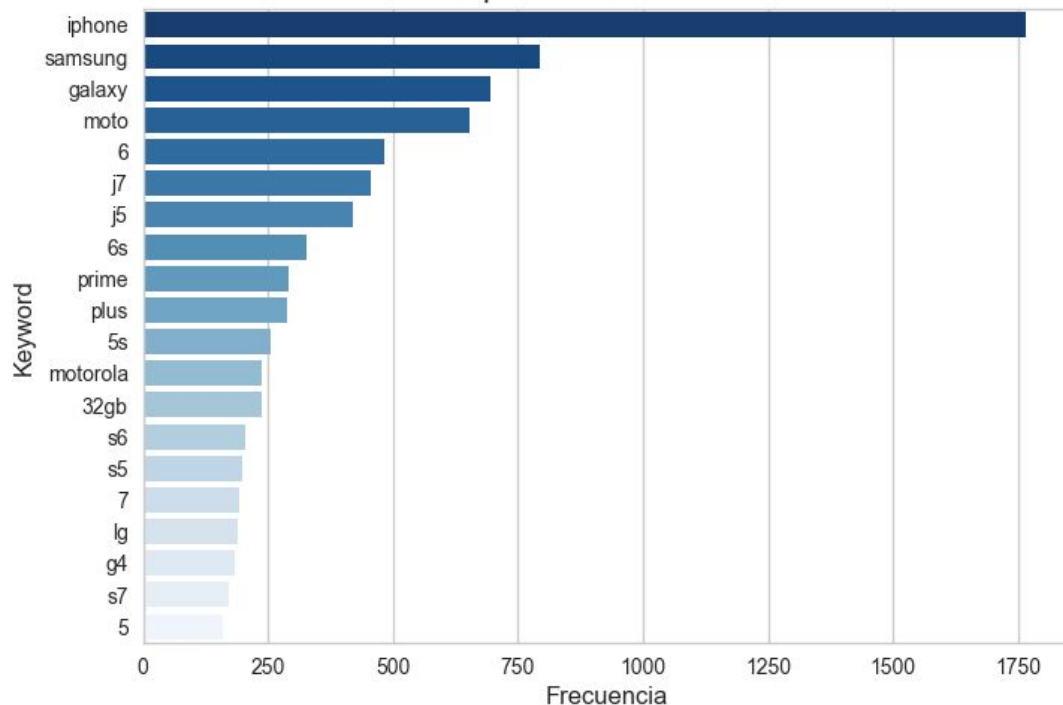


Un primer golpe de vista de este análisis es que los clientes buscan Iphone por la marca, en cambio cuando buscan a su competidor Samsung lo hacen por el modelo.

Realizando el mismo análisis para el dataset de sólo los usuarios de compra, se observa que no hay mayor variación al respecto, salvo por la aparición de la palabra 32 Gb.

# ¿Cómo buscan los clientes que compran? - Veinte palabras

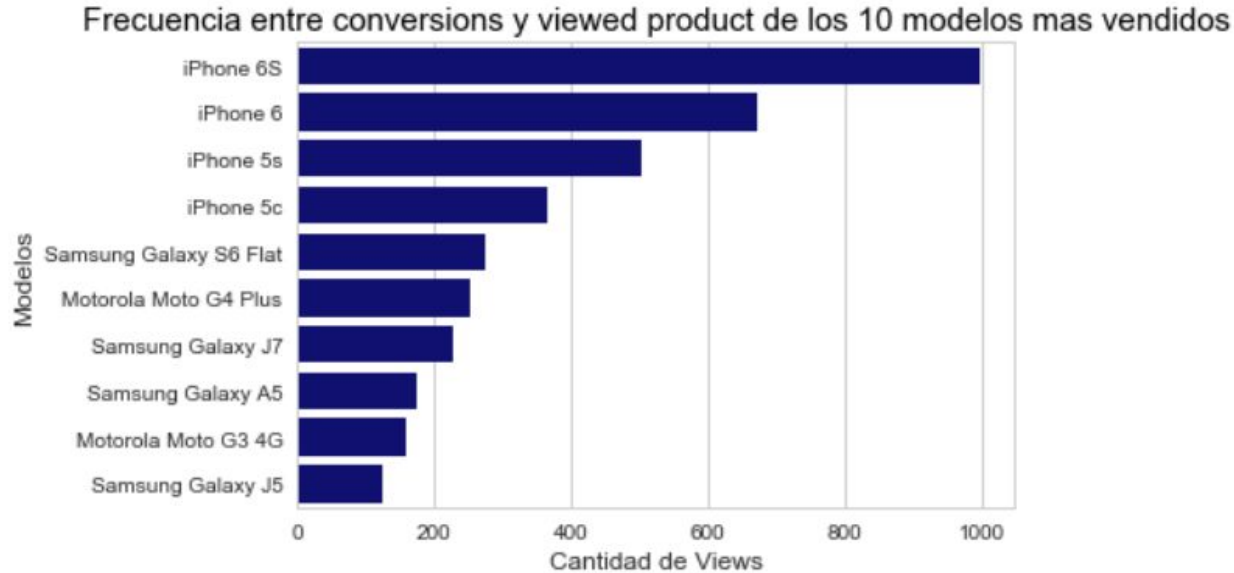
Veinte palabras más rankeadas



Un análisis interesante de eso, es que cuando se ingresó a la página de Trocafone y se realizó una búsqueda, claramente el buscador no tiene un indexación adecuada, ya que cuando se realizaba por ejemplo una búsqueda por ejemplo “J7” aparecía efectivamente los modelos de “Samsung J7”, sin embargo si se realiza la misma búsqueda con “J 7” o “7J” aparece el resultado por defecto. Se hace mención de esto ya que, las frases y keywords no pueden ayudar a realizar una mejor indexación y mejorar así nuestros resultados de búsquedas.



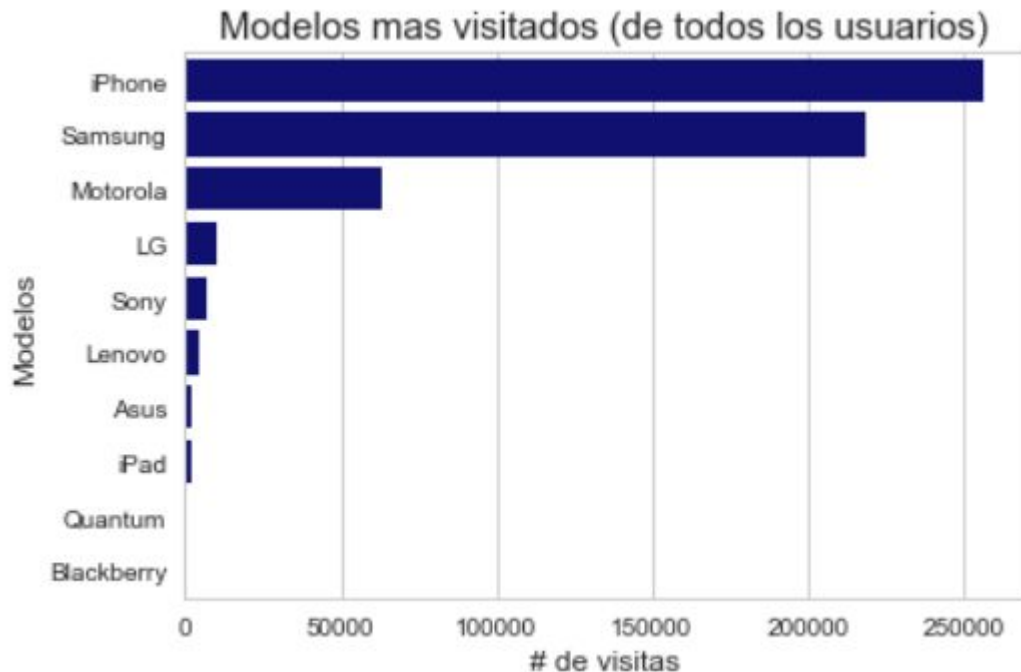
# ¿Cuántas visitas a un producto necesita un usuario para comprarlo?



Tomamos los 10 productos más vendidos en Trocafone y analizamos la frecuencia entre las visitas a un producto y las compras de estos:

- Notamos que hay 4 modelos de Iphone que los usuarios necesitan navegar mucho más para llevarse a cabo la venta: el Iphone 6S necesita de casi 1000 views para tener una venta.
- En cambio, el modelo más vendido de todos, el Samsung Galaxy J5, solo necesita de 125 views para darse una Venta.

# ¿Cuáles son los modelos más visitados por los usuarios?

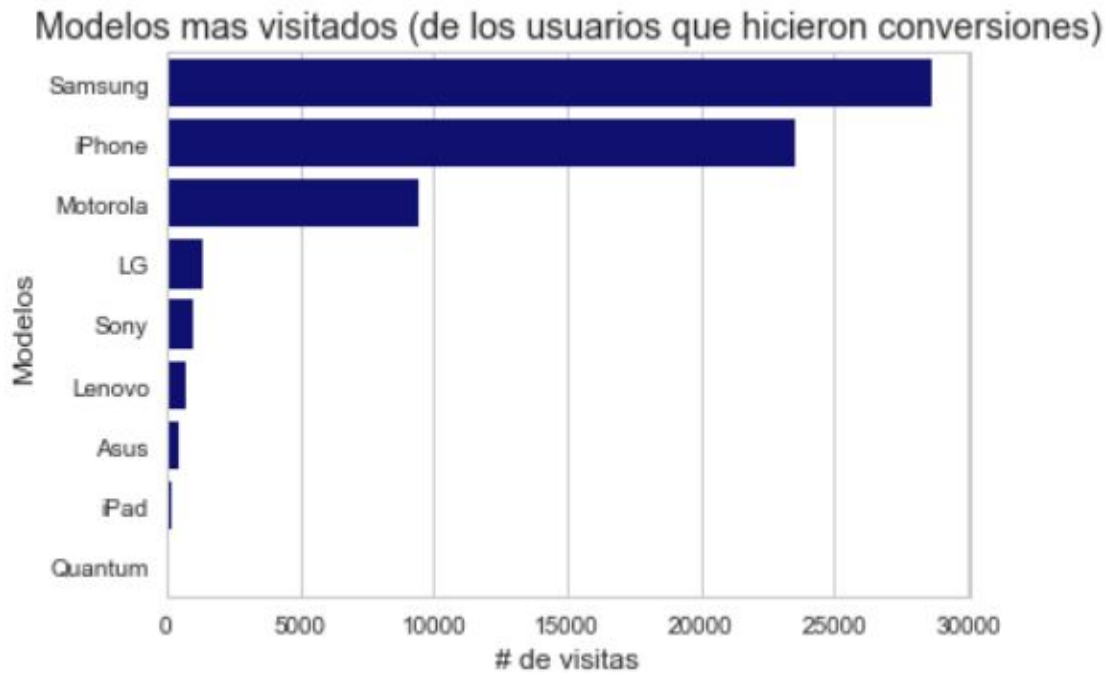


Investigamos las marcas que los usuarios de Trocafone visitaron independientemente del modelo buscado.

En base a esto se pudo observar:

- Entre los Iphones y Samsung abarcan el 84% de las visitas.
- Hay una clara caída de la conversión a partir del mes de marzo.

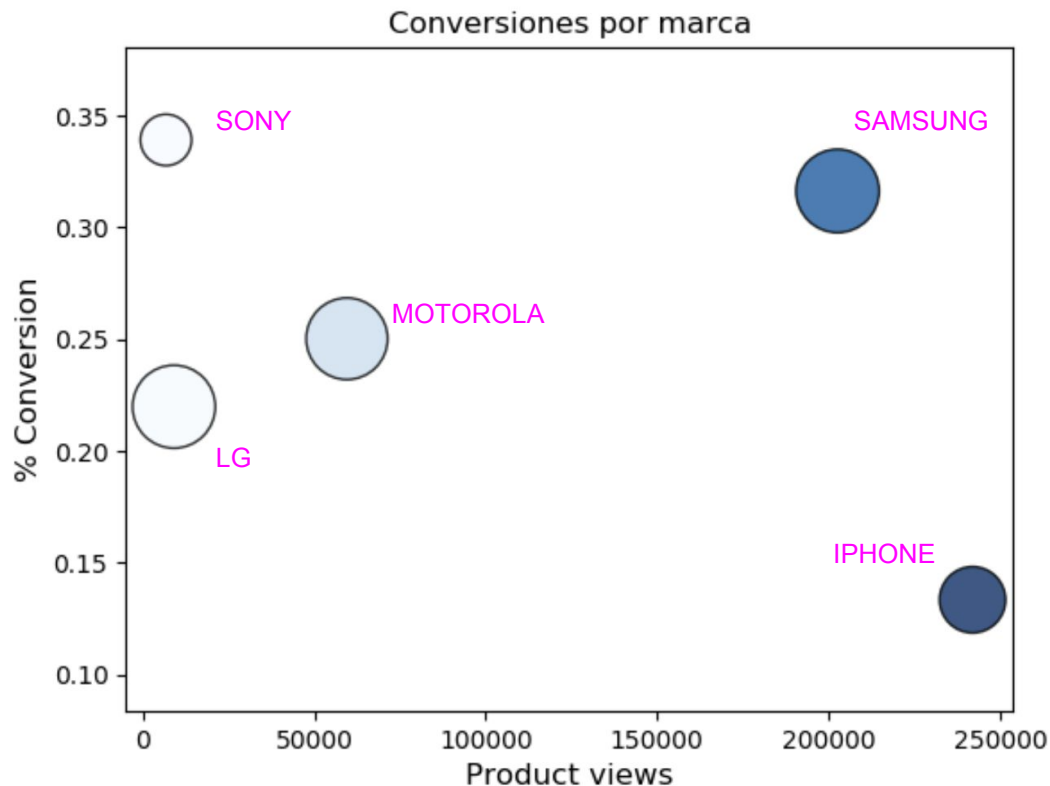
## Entonces... ¿Cuáles son los modelos más visitados por los usuarios que realizaron al menos una conversión?



Siguiendo el análisis de los modelos más visitados, llegamos a 2 conclusiones:

- Nuevamente entre los Iphones y Samsung abarcan más del 80% de las visitas.
- Vemos que, al sesgar los datos a solo usuarios que convirtieron, Samsung lidera las visitas superando a Iphone en un 8%

## ¿Cómo es la conversión por marca? ¿Cómo es la relación con el tiempo que transcurre entre que ingresan los usuarios y compran?



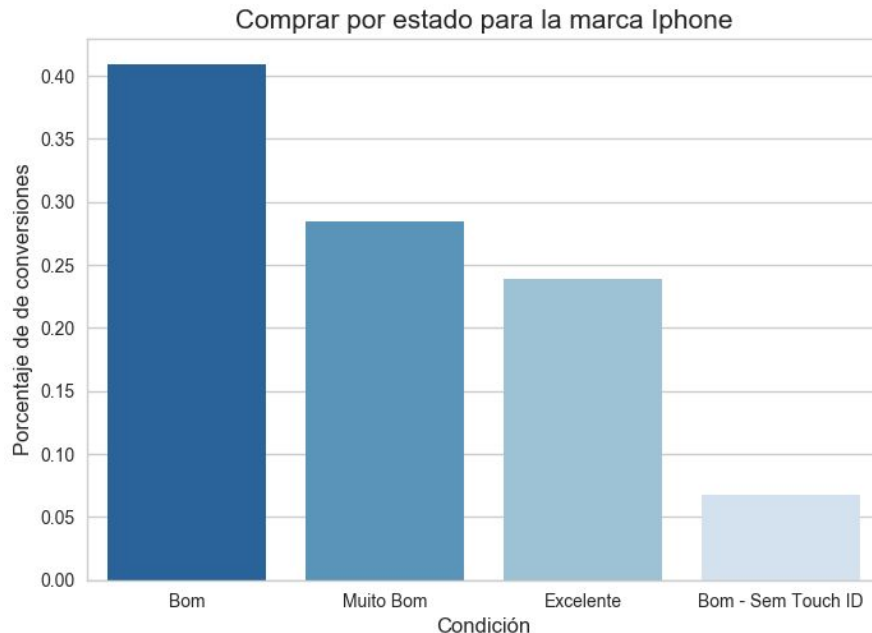
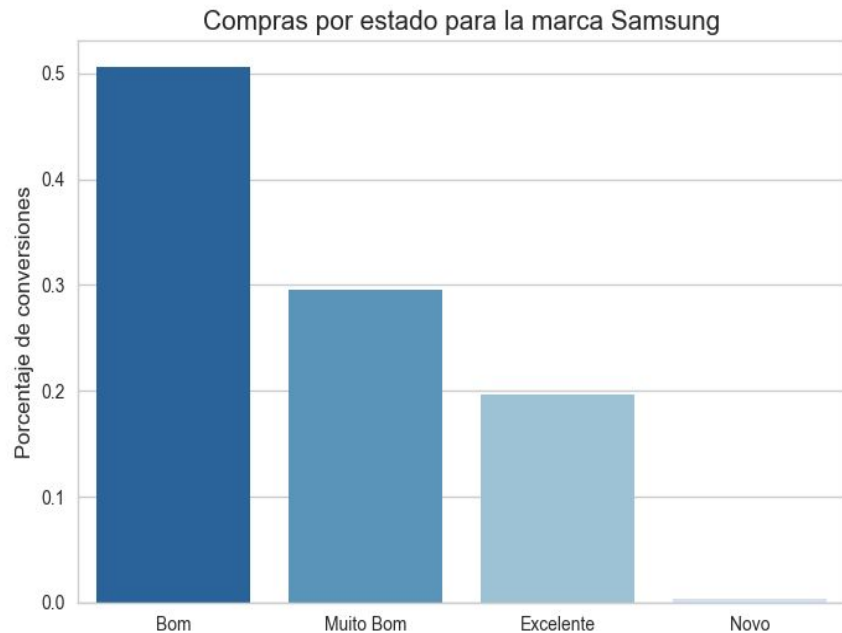
Acá el principal insight que vemos es que los celulares de marca Sony convierten muy bien y en un menor tiempo.

Los iPhones, por otro lado, tienen la peor conversión por la gran cantidad de visitas que tienen. No obstante el tiempo que se toman los usuarios en comprarlos (tamaño de la burbuja) es pequeño en comparación a las demás marcas.

brand_viewed	viewprod	conversion	kpi_conversion	tiempo_conv_mins
IPHONE	242014.0	323.0	0.133463	17.745098
LG	9098.0	20.0	0.219829	27.714286
MOTOROLA	59547.0	149.0	0.250223	26.880000
SAMSUNG	202648.0	641.0	0.316312	28.084337
SONY	6782.0	23.0	0.339133	10.600000

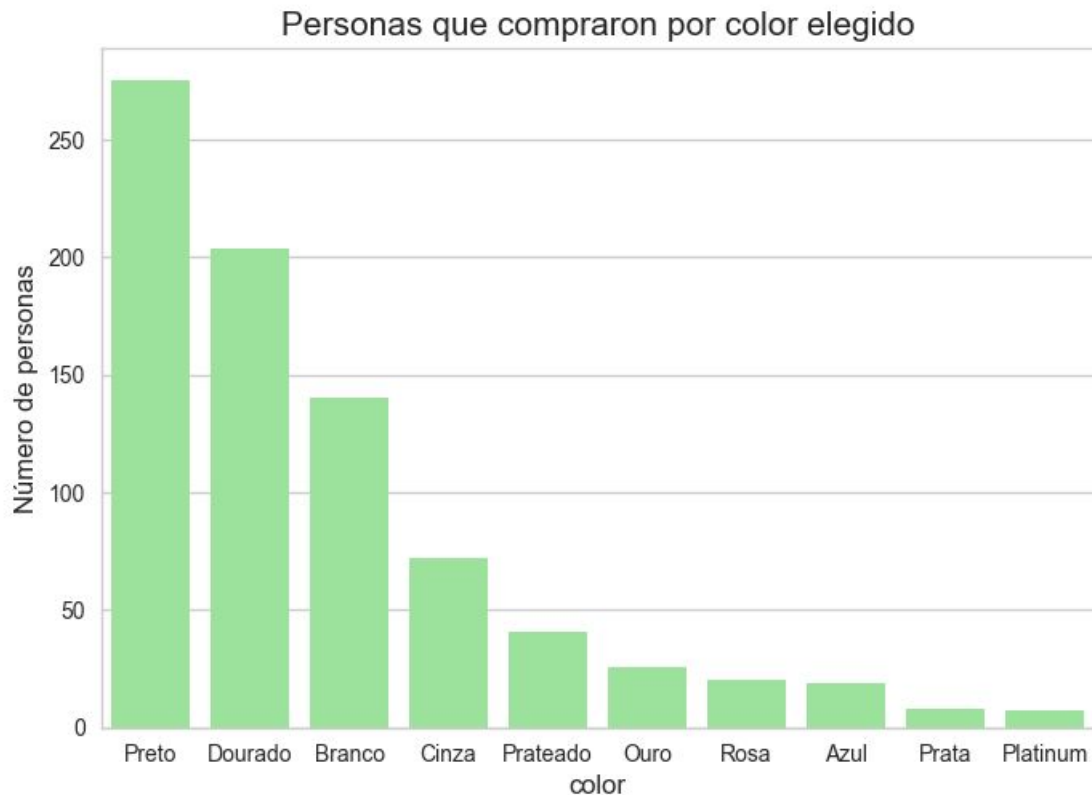


# Estado de productos que compraron los clientes: Samsung vs Iphone



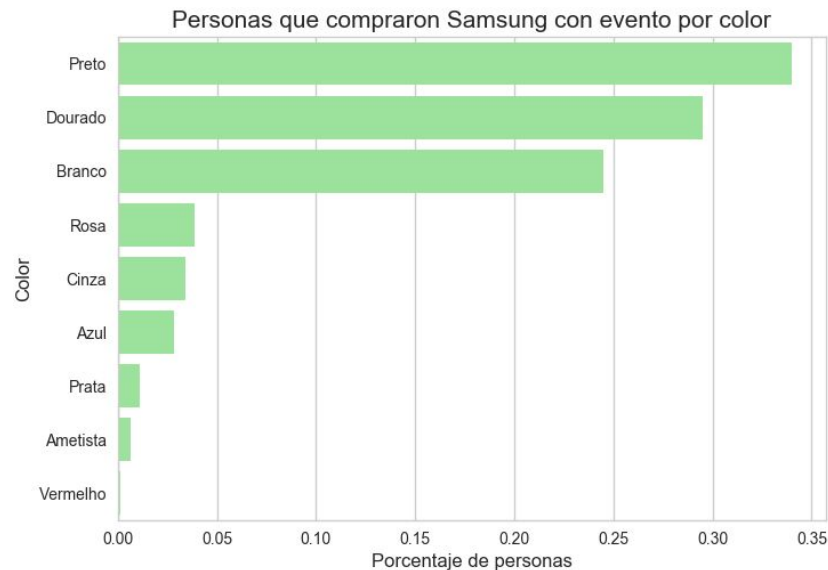
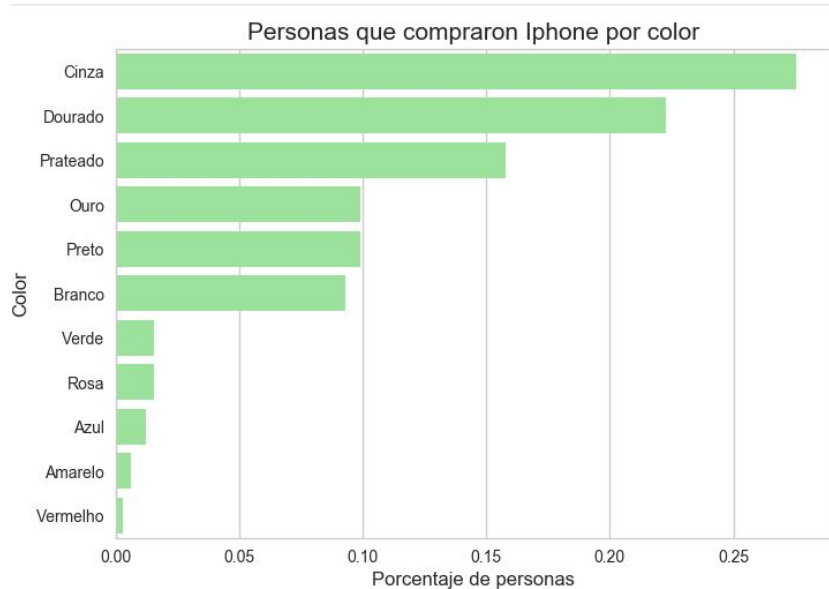
Se observa que tienen el mismo comportamiento de distribución de compra por condición, salvo un pequeño porcentaje de “equipos nuevos” para Samsung, y para iphone aparece “Bom- sem Touch Id”.

# Las personas que compraron ¿Qué color eligieron?



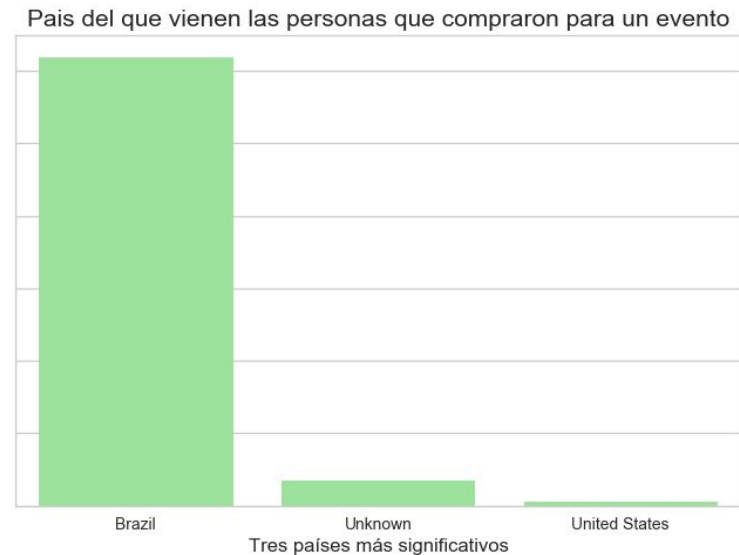
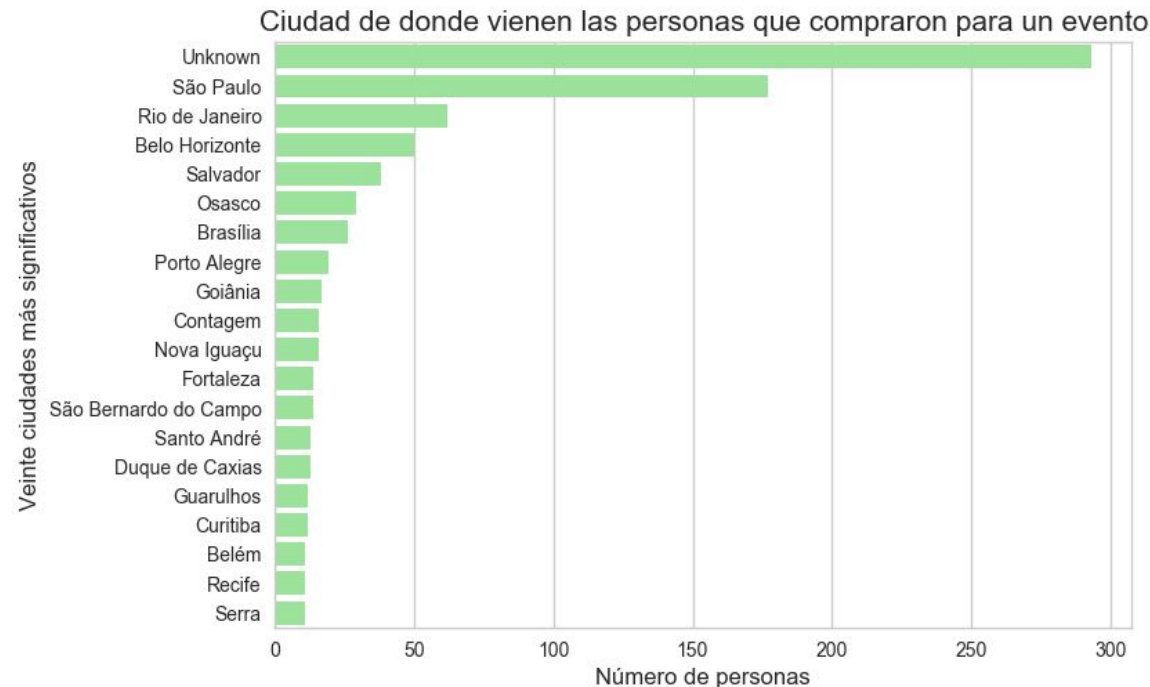
Los colores favoritos de los compradores son “Preto”, “Dourado” y “Branco”, esto nos puede dar un valor, al momento de establecer un peso de acuerdo al celular al momento de realizar la compra al proveedor.

# Color: Iphone vs Samsung



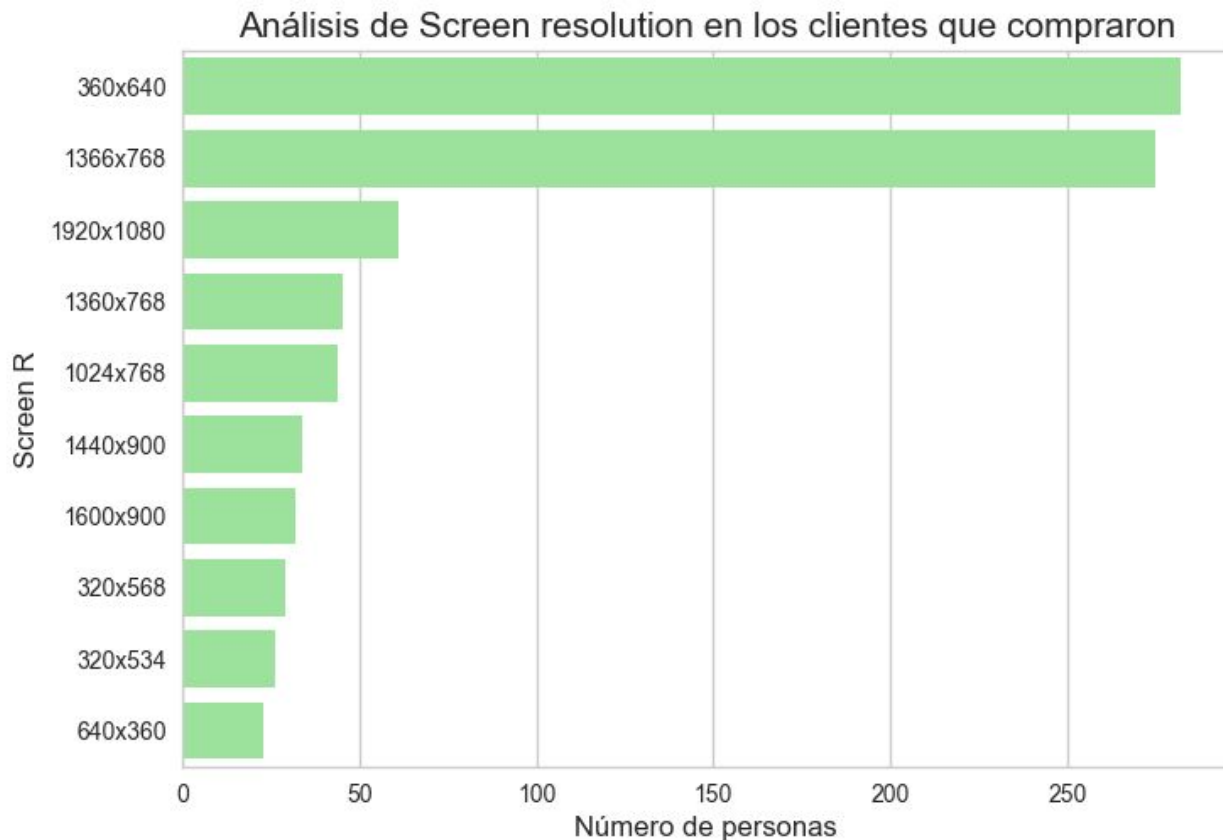
Siguiendo el análisis anterior, también se puede hacer enfoque por modelo, aquí un versus entre los modelos más comprados, claramente tienen distinta gama de colores requerido por los clientes.

# Ciudad y País de las personas que realizaron una compra



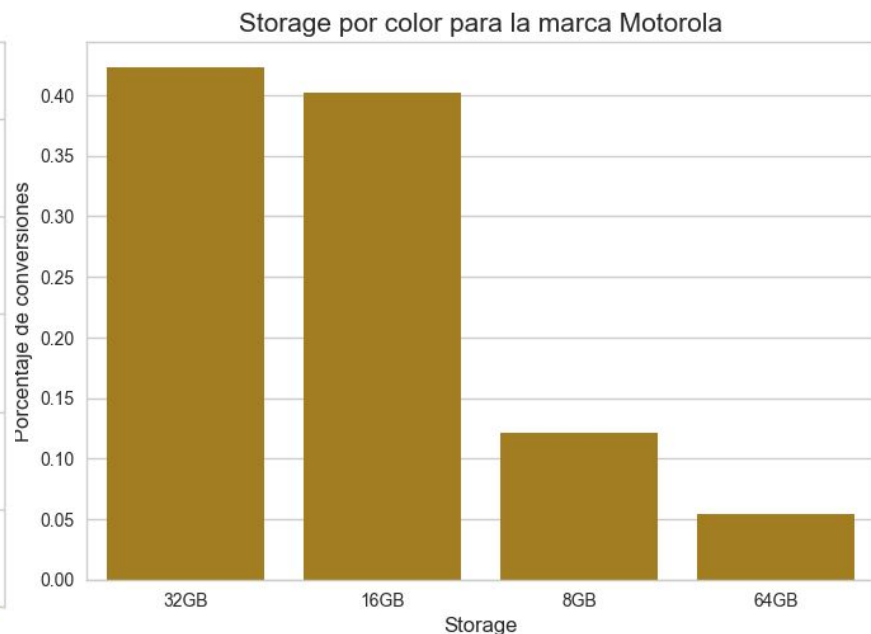
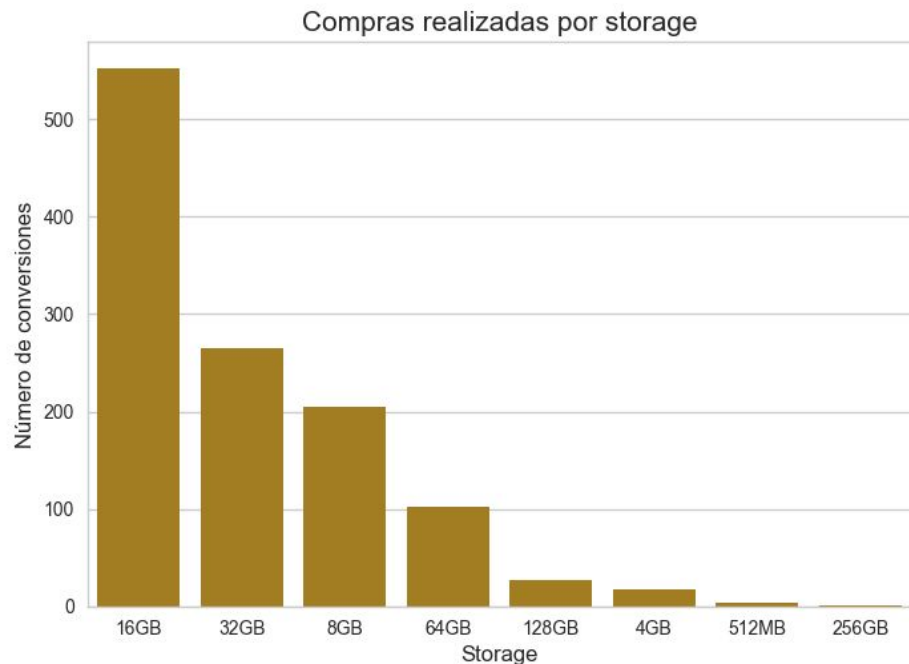
Se observa que el principal país de donde vienen nuestros clientes es Brasil, y por la distribución de las ciudades Sao Paulo, Rio de Janeiro y Belo Horizonte, como las tres ciudades más importantes. Considerar sin embargo que puede existir una long tail en las ciudades.

# Análisis del Screen R, en los eventos de las personas que compran



Esto requiere un análisis más profundo, ya que al intentar analizar Iphone y Samsung que son las marcas más compradas, no se dispone de ningún registro en los eventos del cliente que consigne screen resolution.

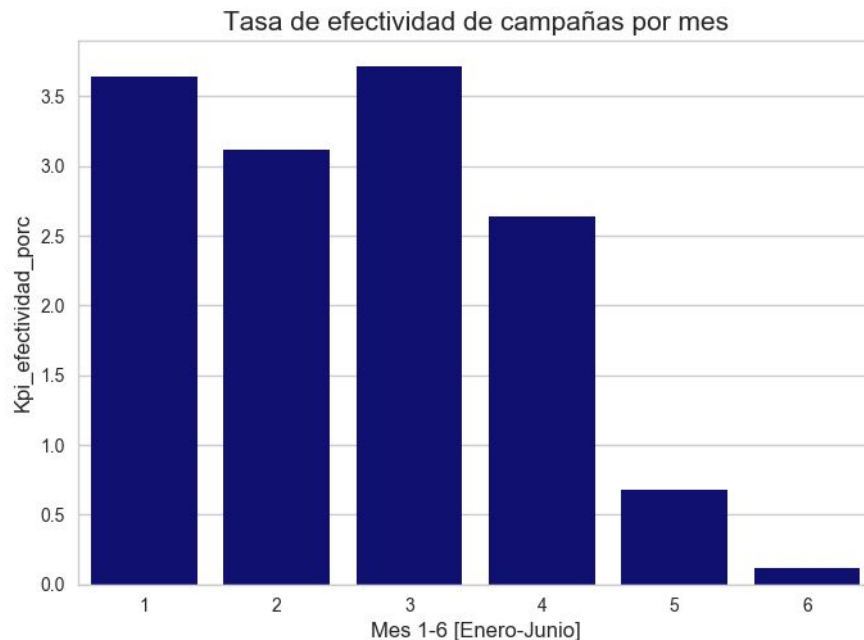
# Análisis de la variable storage en las compras



- En cuanto a las características de los equipos vendidos para el atributo storage los más requeridos son de 16GB siendo el aproximadamente el doble del segundo más requerido 32 GB.
- También analizando por marca para Iphone y Samsung se observó que tienen similar distribución al gráfico de 'Compras realizadas por storage', sin embargo para la marca Motorola nos damos cuenta que esta distribución cambia, siendo el storage más comprado de 32 GB

# ¿Y cómo nos fue con las campañas? Analisis mensual

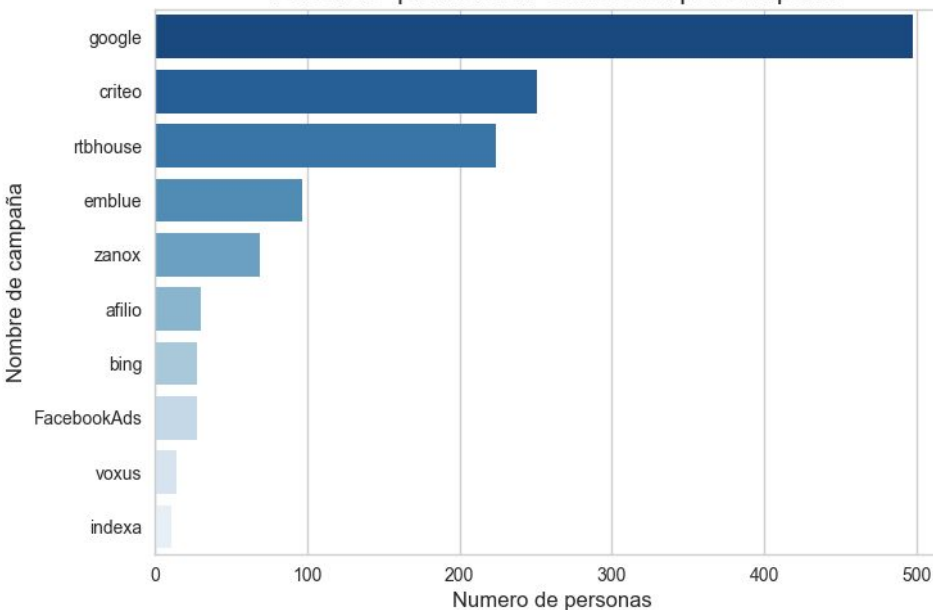
	person	checkout	conversion	ad campaign hit
month				
1	495	317.0	62.0	1702.0
2	693	413.0	73.0	2340.0
3	1529	1310.0	255.0	6872.0
4	2537	1833.0	287.0	10888.0
5	12707	12368.0	259.0	38108.0
6	10747	10747.0	26.0	22917.0



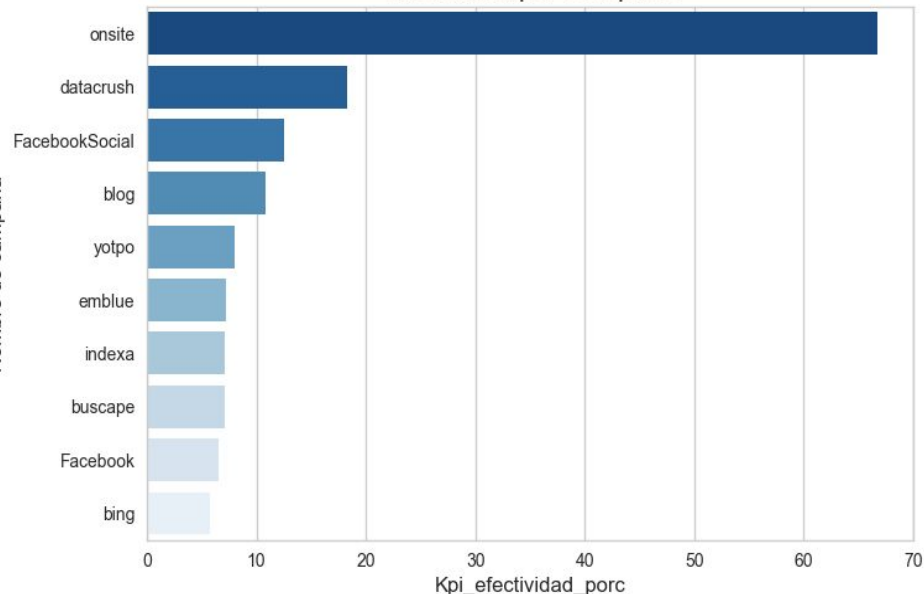
- De esto se desprende que los meses de enero a marzo hay menos conversiones pero también hay más efectividad, no así en los meses a partir del abril, quizás esto se deba a las campañas, ya que se nota un salto de los que vinieron por las mismas a partir de abril
- Del gráfico de la izquierda la efectividad **Kpi\_efectividad\_por**=  $\frac{[\text{Conversiones}]}{[\text{Trafico ad campaign hit}]}$

# ¿Y cómo nos fue con las campañas?

Personas que hicieron conversión por campaña



Efectividad por Campaña



- Y no, no siempre las campañas que nos dan más conversiones son las más efectivas, o sí. Tener en cuenta que el **kpi\_efectividad\_por** = **([Número personas conversión] / [Cantidad de tráfico generado])**, todo con respecto a la misma campaña. [por cuestiones de espacio sólo se muestran los diez primeros)



# FIN

