

# Trabajo Práctico N°1: Análisis Exploratorio de datos de Trocafone

Grupo I: Raúl Dip, Roberto Irahola, Emanuel Lemos & Lautaro Peña

Organización de Datos  
Universidad de Buenos Aires

25 de septiembre de 2018

## **Abstract**

En el presente trabajo realizamos un análisis exploratorio sobre un subconjunto representativo de datos de eventos de web analytics de usuarios que visitaron la plataforma de ecommerce de Brasil de [www.trocafone.com](http://www.trocafone.com). El objetivo es obtener nuevos insights que puedan resultar útiles para la toma de decisiones.

Se abordó el comportamiento del usuario antes de la compra del producto, la relación entre búsquedas de distintos modelos, la fidelidad de los clientes hacia las marcas y las características geográficas de los usuarios entre otras cosas.

Link a Github: [https://github.com/rauldip22/TP1\\_TROCA.git](https://github.com/rauldip22/TP1_TROCA.git)

## 1. Orientación del trabajo

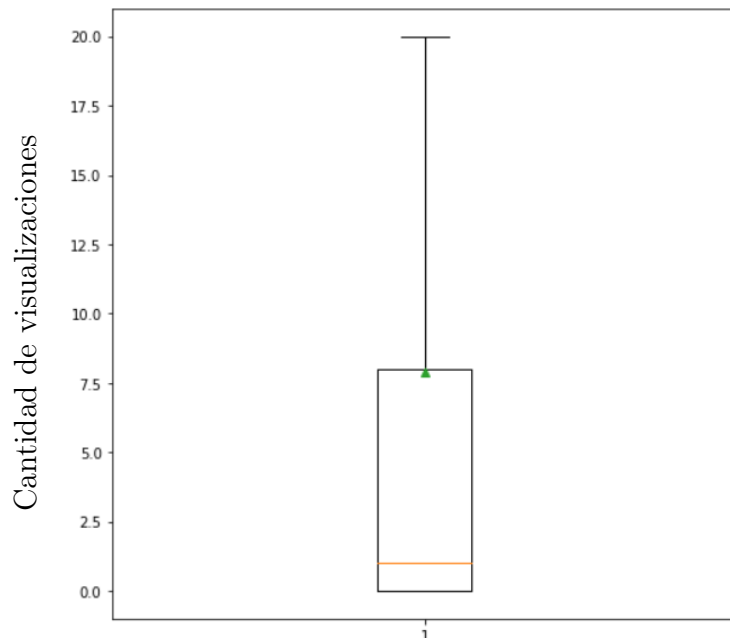
A pesar de ser sólo un subconjunto de datos, los mismos nos permiten realizar una gran gama de estadísticas descriptivas. En este análisis nosotros optamos por buscar cosas fuera de lo obvio; para tratar de generar nuevos insights que permitan tomar decisiones. El trabajo está dividido en 5 secciones. En cada una se aborda una dimensión diferente; yendo de lo más superfluo a lo diferenciador.

## 2. Comportamiento de los compradores

Esta sección tendrá como objetivo ver algunas características de los compradores.

En primer lugar, nos interesamos por la búsqueda que realizan los usuarios antes de comprar los productos. Quedándonos con los clientes que efectivamente compren un producto, pasamos a contar la cantidad de veces que visualiza un modelo de celular antes de comprarlo. El resultado se vuelca en la Figura 1.

Figura 1: Boxplot de Visualizaciones

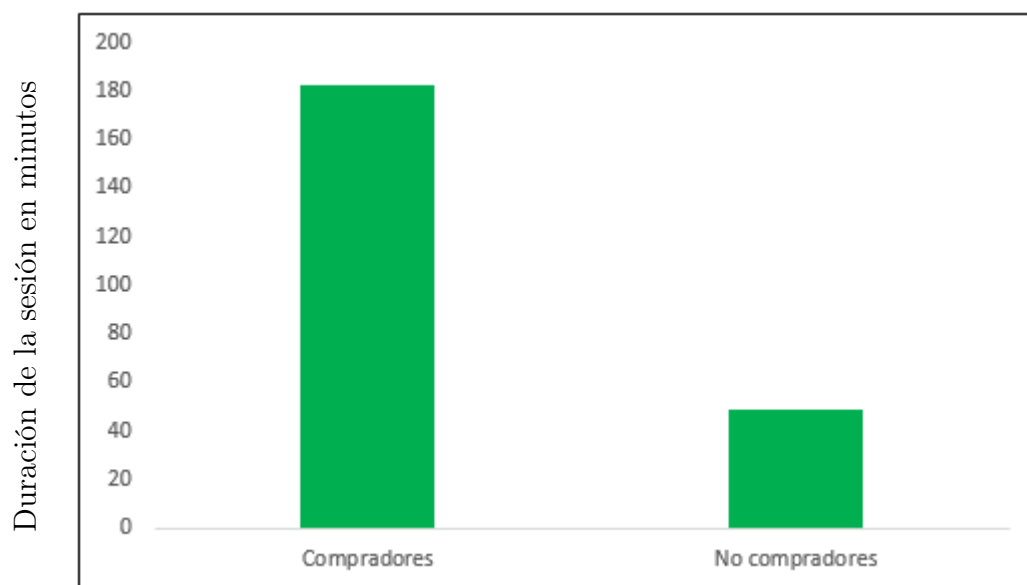


Como se ve en el Gráfico (en el que retiramos los Outliers), en promedio los compradores ven un modelo de celular aproximadamente 8 veces antes de comprarlo. Vemos además que los usuarios pertenecientes al primer cuartil no ven previamente un celular antes de comprarlo. Como así

también concluimos que el 50 % de los compradores vieron solamente 1 vez el modelo de celular que posteriormente compraron.

Probablemente, estos números escondan cierto conocimiento previo de los productos, adquirido por uso reiterativo del sitio. Para tratar de de evitarnos este sesgo, tomamos el primer día de navegación de los clientes nuevos; y comparamos el tiempo destinado por los compradores y por los no compradores a navegar por el sitio (obtenido como diferencia entre el primer y el último evento). Los resultados se vuelcan en el Gráfico 2.

Figura 2: Duración Promedio de la Sesión en minutos



Como es de esperar, los compradores destinan en promedio un poco más de 3 hs en el sitio; mientras que los no compradores rondan los 50 minutos.

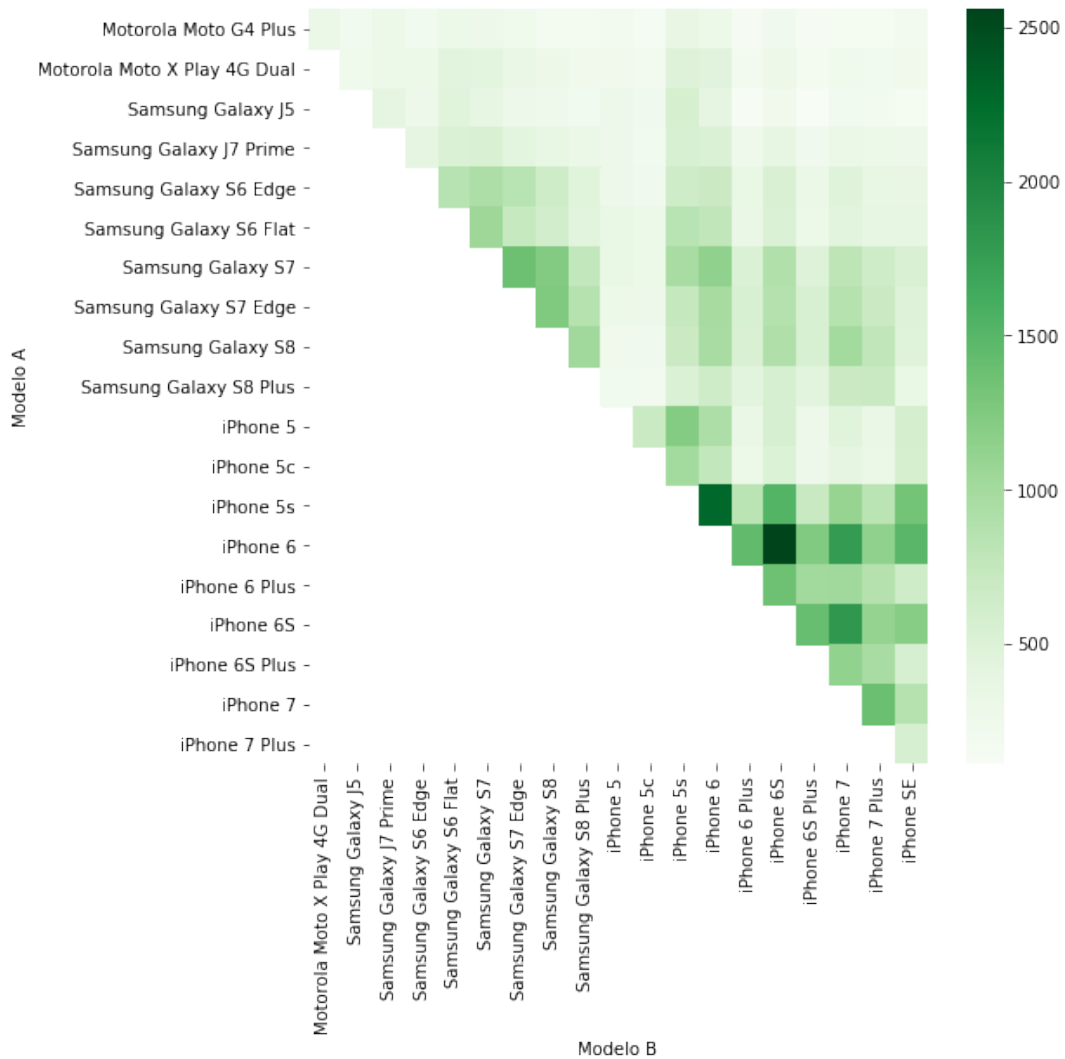
Por otra parte, **en promedio una persona compra por primera vez un producto 7 días 8 hs y 9 mins después de verlo por primera vez.**

### 3. Productos

En esta sección continuaremos desarrollando el comportamiento de los usuarios; pero esta vez orientado a los productos. Particularmente veremos si existe alguna relación entre los productos vistos por los clientes; y si existe un componente de "fidelidad" de estos mismos a la hora de comprar.

Para ver la relación entre las visualizaciones, extrajimos por cada usuario todos los modelos que vió y se armaron todas las combinaciones (de dos elementos) posibles de modelos vistos. Luego, contamos la cantidad de combinaciones iguales de todos los usuarios. Obteniendo el heatmap mostrado en la Figura 3.

Figura 3: Visualización conjunta de modelos



El Gráfico mostrado en la Figura 3 muestra con la cantidad de personas que vieron el Modelo A y también vieron el Modelo B. Por ejemplo, el número más alto de combinaciones es 2559 (iPhone 6, iPhone 6S), lo

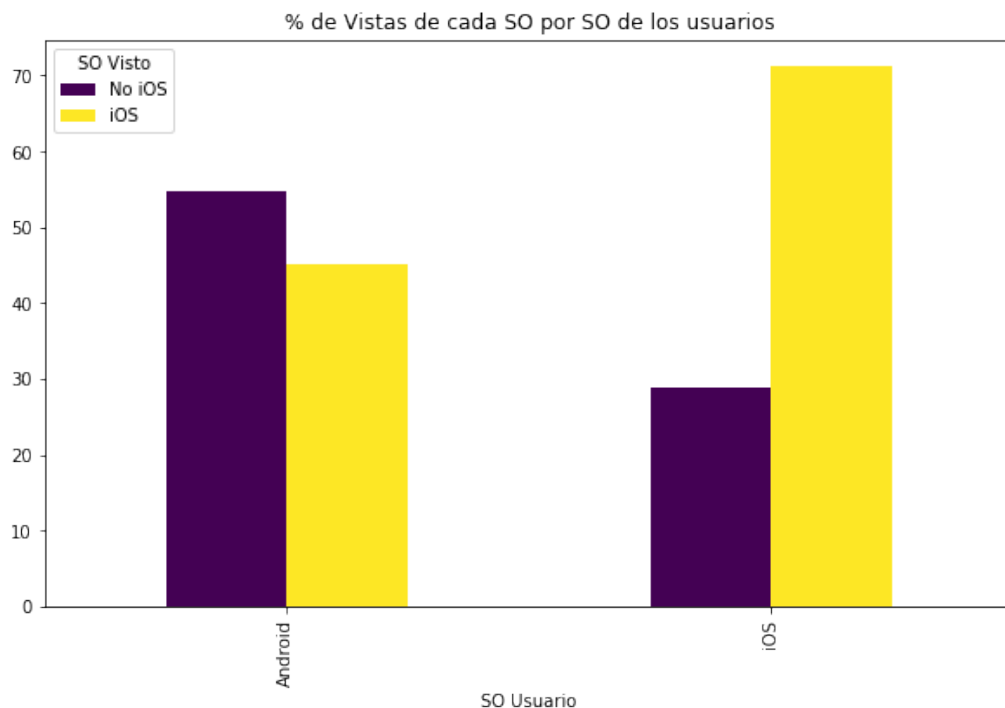
que quiere decir que hay 2559 personas que vieron el modelo iPhone 6z además vieron el modelo iPhone 6S"

Los datos nos permiten ir corroborando una hipótesis que muchos tenemos en la cabeza: los que quieren un iPhone no suelen buscar el modelo equivalente de Samsung. Por ejemplo, si bien parece haber un número de usuarios que busca el Samsung Galaxy S8 Plus y el Iphone 7; parece estar muy lejos de tener la relación que tienen el Iphone 6s y el 6.

A partir de este análisis, se podría sugerir nuevos productos a adquirir para los clientes en base a sus visualizaciones anteriores. Focalizando así el esfuerzo y, en el mejor de los casos, aumentando las ventas. La fuerte asociación recién descrita en los productos de iPhone no se ve entre los modelos de Samsung ni de otra marca, lo que podría estar evidenciando cierta fidelidad que pueda existir hacia los productos de Apple. Con el siguiente análisis se buscará arrojar más luz sobre este aspecto.

Dividimos a los usuarios según el sistema operativo que usan para navegar en el sitio; y dividimos los modelos buscados también según su sistema operativo. Una vez hecho esto, pasamos a ver las visualizaciones y compras hechas de cada producto según el sistema operativo a través del cual se realizó. Los resultados se ven en el Gráfico 4 y 5.

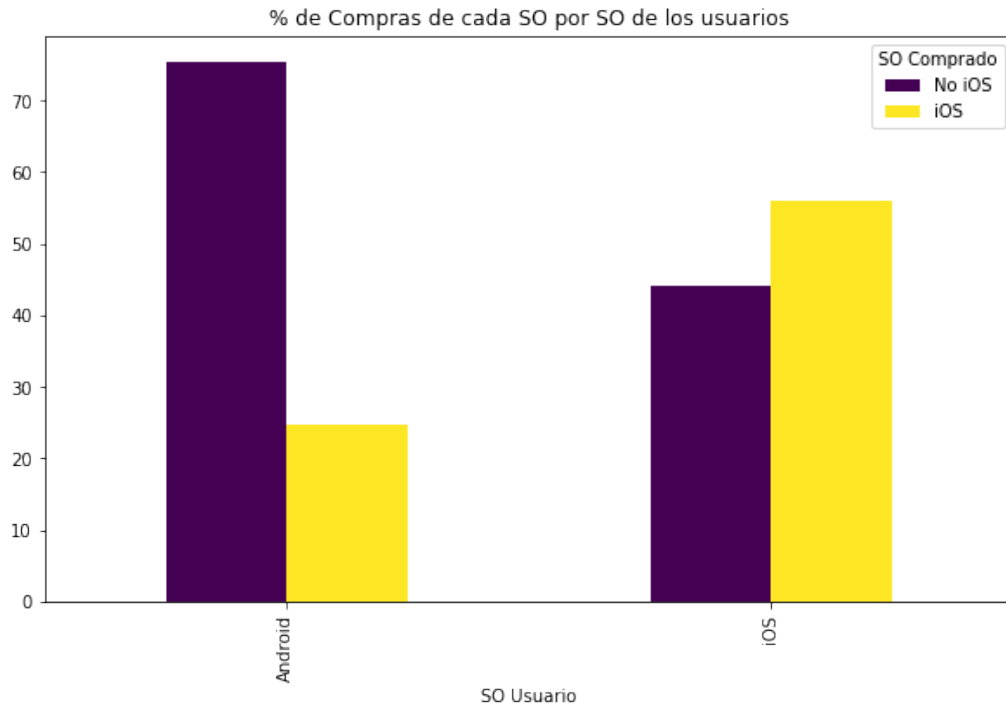
Figura 4: Visualización según sistema operativo



Se puede observar que los usuarios que navegan desde iOS ven en su gran mayoría productos iOS (71 % de las vistas), mientras que esa "fidelidad de producto" no es tan marcada en los usuarios que navegan desde Android ya que un 45 % de los mismos buscan productos que tienen

el sistema operativo de Apple.

Figura 5: Compras según sistema operativo

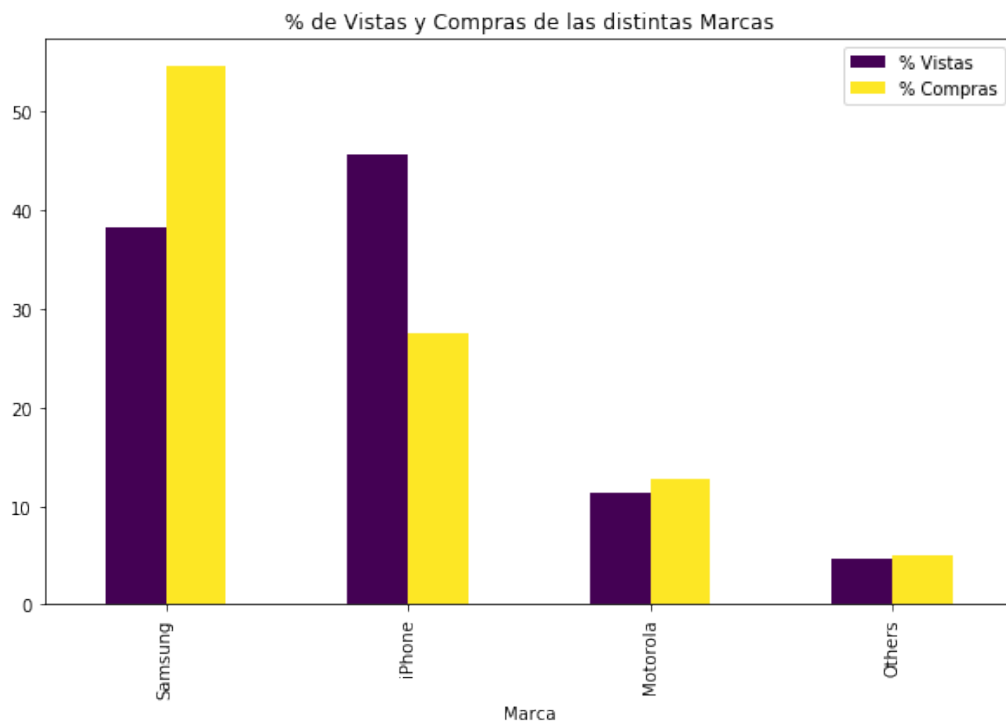


Cuando vemos las compras, vemos que ese "deseo de compra" de productos Apple por parte de usuarios Android no se lleva acabo ya que solamente un 25% de los usuarios Android que compran terminan optando por productos con sistema operativo iOS. En cambio, los usuarios de iOS sí terminan llevando a cabo la compra de productos Apple, aunque en menor proporción de fidelidad respecto a la gran diferencia que se veía en las búsquedas.

Aparte de la conclusión directa que sacamos del análisis anterior: "los clientes de Apple parecen ser más fieles"; hay un segundo detalle que no debemos dejar pasar. La cantidad de visualizaciones de productos de Apple hecha por productos con sistema operativo iOS es muy alta, pero las compras hechas desde los mismos celulares no lo es tanto. Esto nos llevó a poner el foco en las visualizaciones y las ventas de cada producto sobre los totales de vistas y venntas respectivamente. Los resultados se describen en la figura 6.

Como vemos en el gráfico de la Figura 6, iPhone es la única marca que tiene un porcentaje de visualizaciones sobre el total mayor a la cantidad de ventas sobre el total de ventas. La explicación parece ser clara: estos modelos son más atractivos a los ojos de los compradores, pero seguramente por cuestiones de precio optan por otras alternativas.

Figura 6: Vistas vs Compras por Marca



## 4. Análisis Geográfico

En esta sección nos propusimos caracterizar a los clientes espacialmente. Sabiendo dónde están los clientes que más utilizan la página y donde están los que más compran se podrán realizar acciones personalizadas según el lugar y lo que se necesite fomentar.

En un primer paso calculamos la cantidad de vistas y compras de productos en cada región, lo que plasmamos en el gráfico de la Figura 7 y 8.

Como vemos, claramente la cantidad de usuarios que hay en cada región influye en la cantidad total de visualizaciones y compras. En ambos gráficos vemos que San Pablo y Rio de Janeiro eclipsan la muestra. Por este motivo, calculamos la cantidad de usuarios en cada region y dividimos tanto la cantidad de vistas como de compras en la cantidad de personas que tiene la muestra en cada región. Así llegamos a los gráficos mostrados en las Figuras 9 y 10.

Del gráfico de la Figura 9 vemos que la región de Tocantins tiene la mayor cantidad de vistas por usuario y del gráfico de la Figura 10 vemos que esa misma región también está segunda en cantidad de compras por usuario, solamente superada por la región de Roraima (la cual tiene un número bastante alto ya que tiene solamente una población de 11 usuarios).

Otro datos que sacamos es la cantidad de vistas dividido por la cantidad de compras, para así determinar cada cuantas vistas hay una compra

Figura 7: Cantidad de Vistas en cada región

Vistas de productos en cada región

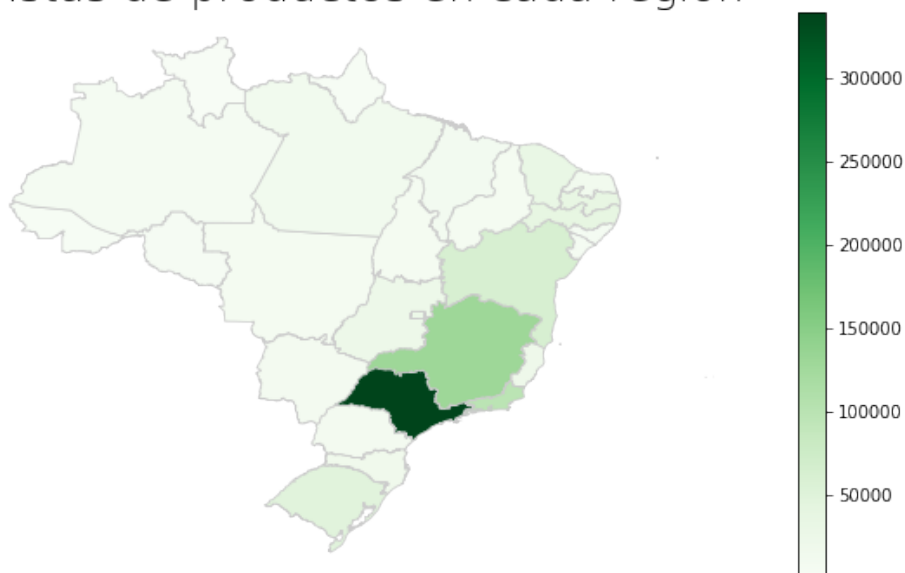


Figura 8: Cantidad de Compras en cada región

Compras de productos en cada región

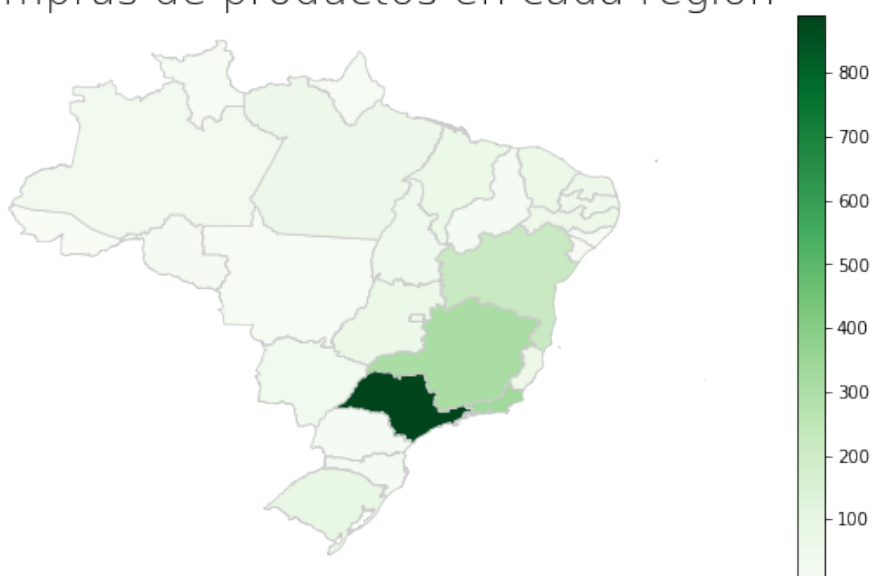




Figura 9: Cantidad de Vistas per capita en cada región

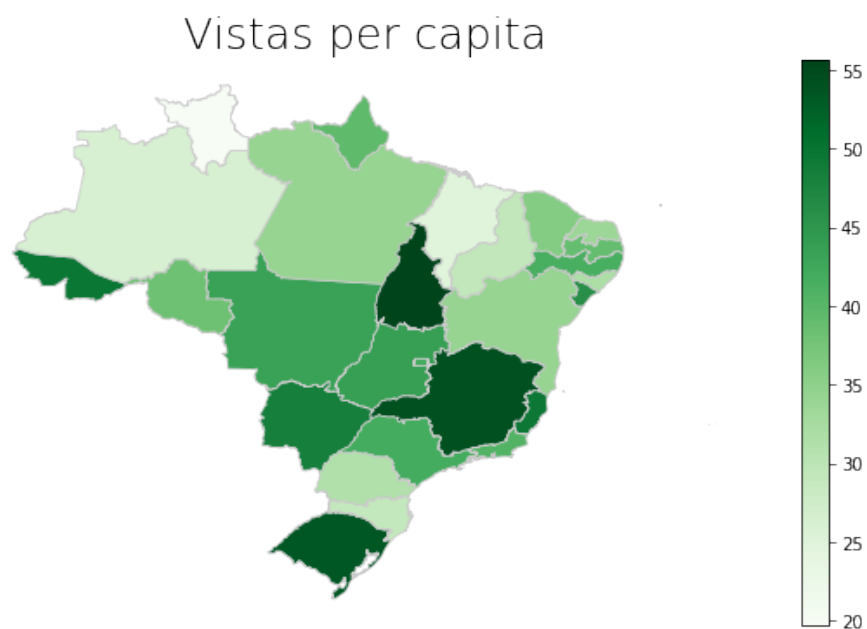
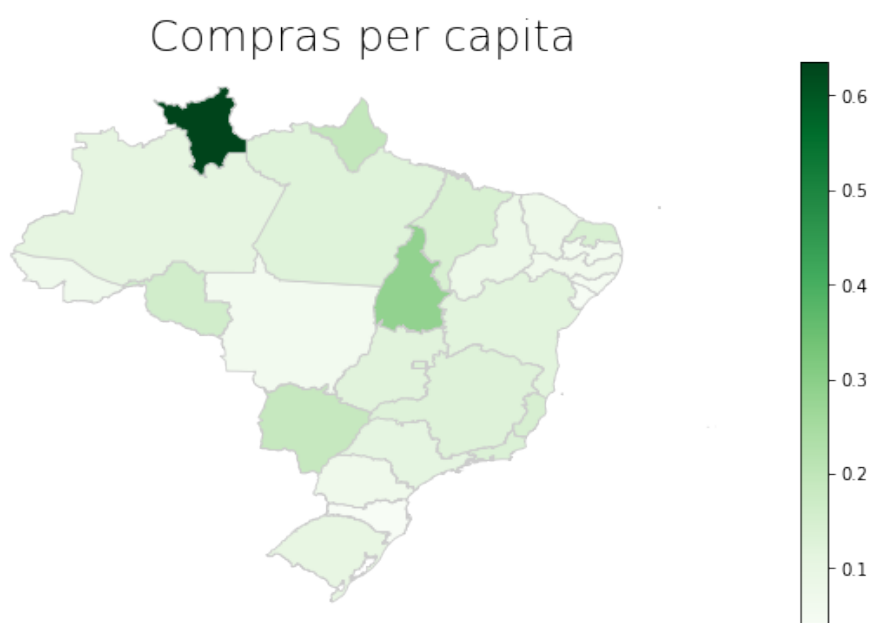


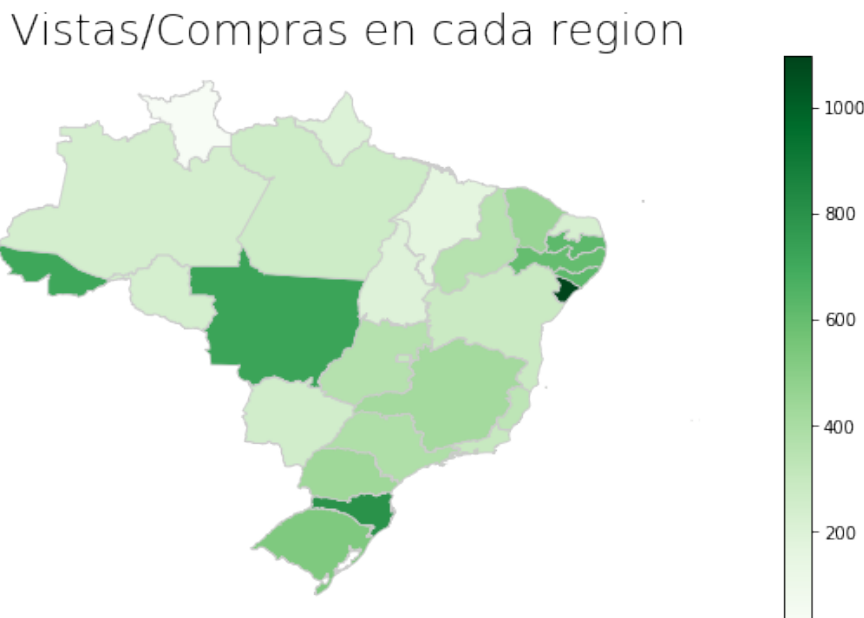
Figura 10: Cantidad de Compras per capita en cada región



en cada región.

El resultado se muestra en el gráfico de la Figura 11.

Figura 11: Vistas por cada Compra por región



En la Figura 11 vemos que Sergipe es la región que mayor cantidad de vistas tiene por cada compra, así también vemos que en Maranhao es donde menos vistas se registran por cantidad de compras, esto sin tener en cuenta a Roraima que muestra una cantidad poco significativa de usuarios.

## 5. Conclusiones

Del análisis anterior se desprenden varios puntos interesantes.

En primer lugar, descubrimos que hay una considerable proporción de clientes que no ven el producto antes de realizar la compra. Es vital poder diferenciarlos de los que sí navegan y ven distintas alternativas, ya que las acciones a tomar difieren en uno u otro caso. Por ejemplo, las publicidades o sugerencias de productos pueden ser mucho más efectivas en este último grupo.

Por otro lado, obtuvimos algunas particularidades de los modelos de iPhone. Estos tienen un mayor número de vistas, lo que parece evidenciar una atracción generalizada por estos productos, pero estas no se traducen en mayores compras; seguramente por cuestiones de precio. La decisión de proponer alternativas de financiamiento para estos productos podría hacer aumentar las ventas de estos modelos; y con esto, las ganancias de la empresa.

Además, la alta fidelidad de los usuarios de iPhone sugiere que el esfuerzo

para estos debiera ser destinado a sugerirles otros productos de la misma marca; ya que son los que más les interesan.

Por último, descubrimos que hay regiones con gran potencial para la empresa y regiones con bajos niveles de compras. En las primeras, destinar una inversión mayor podría ser rentable para la empresa. Para las segundas, en cambio, se podría considerar el uso de campañas de marketing para fortalecer la presencia de Trocafone.