



Análisis exploratorio

Grupo dataPhone

Apellido y nombre

Padrón

Fonseca, Matias Manuel

98591

Zambelli Tello, Brian

98541

Zambianchi, Victoria

98570

TROCAFONE™

Índice.

Introducción - - - - -	3
Análisis de usuarios- - - - -	4
Análisis temporal - - - - -	10
Análisis de dispositivos - - - - -	14
Análisis regional - - - - -	22
Anexo técnico de los datos- - - - -	26
Conclusion - - - - -	27

Introducción

Trocafone es una plataforma de canje de celulares creada en 2014 con sede en Brasil. Por medio de su tienda online, ofrece un catálogo de dispositivos reacondicionados. Sus mercados más importantes se concentran en Argentina y Brasil, aunque recientemente se está expandiendo al resto de sudamérica.

Teniendo todo esto en cuenta, se llevó a cabo un análisis de datos, provistos por la misma, plasmado en este informe. Nuestro objetivo es mostrar características de los datos por medio de nuestros análisis y los gráficos que adjuntamos.

Luego de ver el set, decidimos seleccionar la información en distintas etapas que tienen como núcleo un enfoque particular. En cada una de ellas, visualizamos distintos aspectos del mercado y de la situación de la empresa, para culminar con una sección de conclusiones y propuestas para optimizar la plataforma.

Para finalizar, el análisis exploratorio del set se realizó utilizando el lenguaje de programación Python y las librerías pandas, numpy, matplotlib, seaborn y geopandas.

“

En la actual era del marketing y la web 2.0, la web de una empresa es la clave de todo su negocio

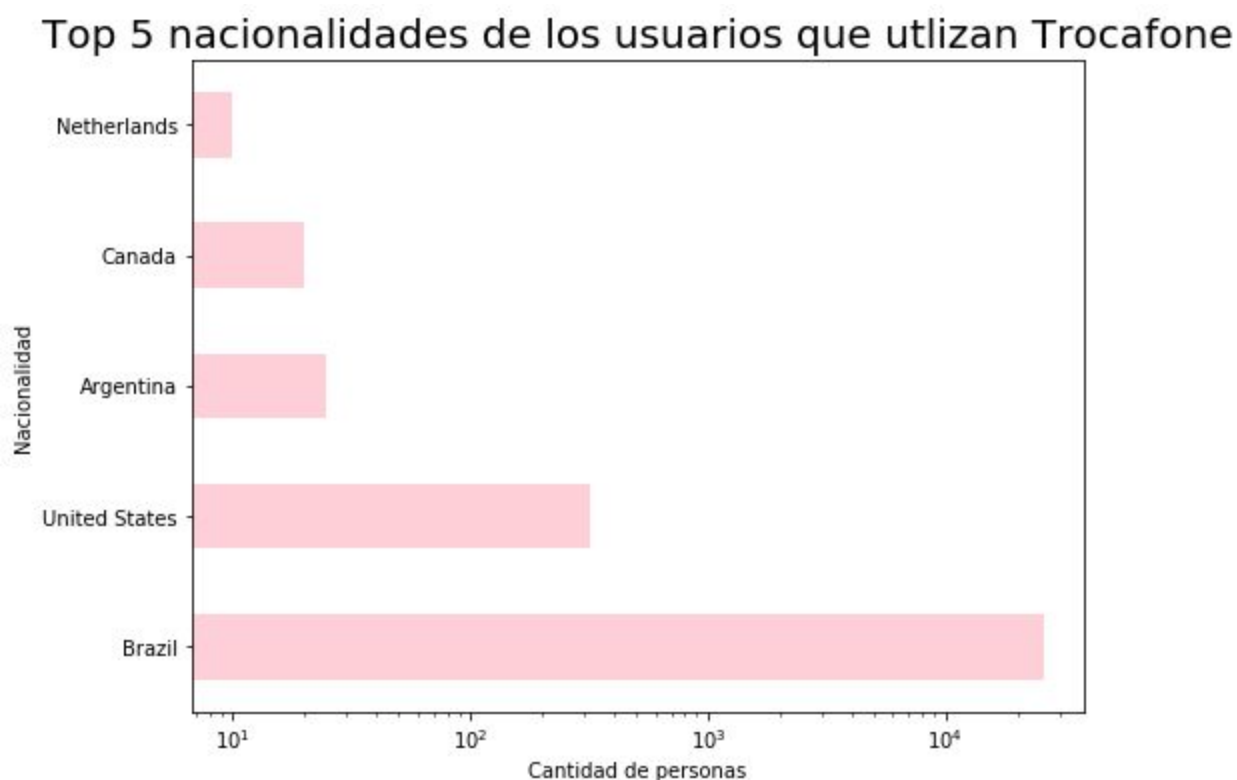
-Marcus Sheridan

”

Análisis de usuarios.

Lo principal al analizar una compañía comercial, es saber el target de clientes al que está enfocada. Es por ello que nuestro análisis comienza por el perfil del usuario de Trocafone.

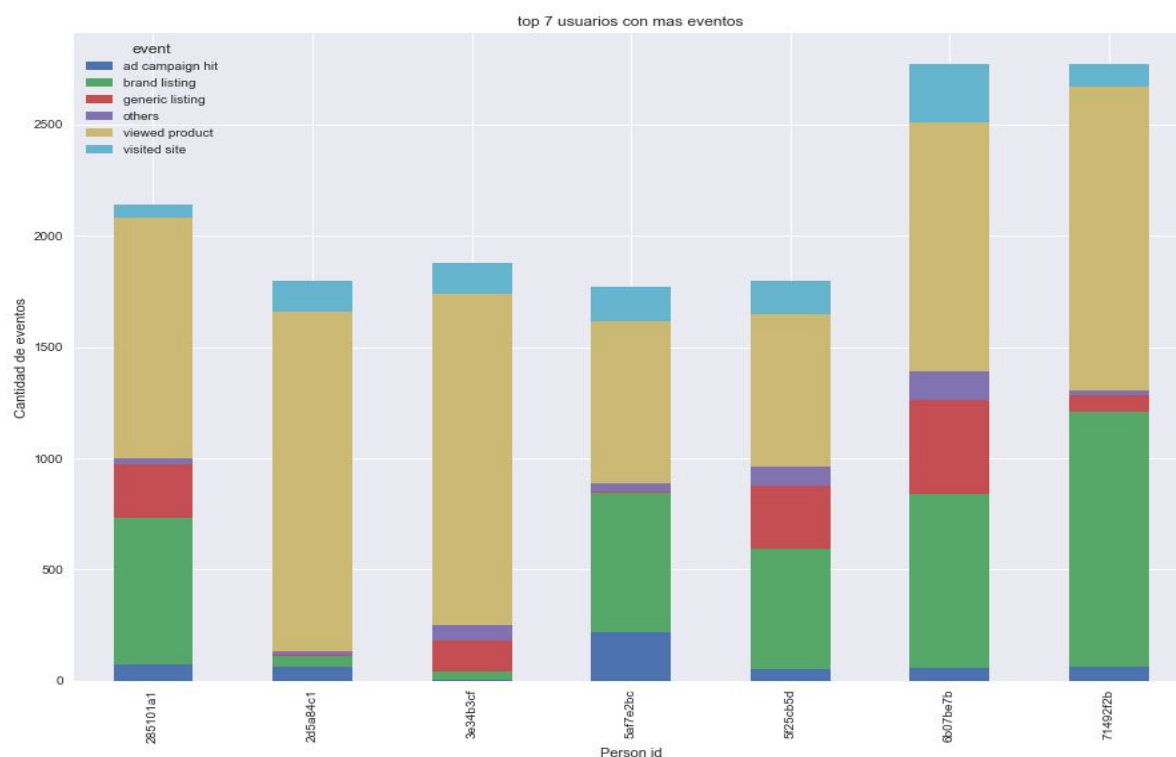
El paso inicial fue buscar las nacionalidades de cada uno de ellos y ver cuales predominan. Al ser una plataforma creada en Brasil, supusimos que los usuarios brasileños iban a liderar el top, y así fue. En el siguiente gráfico se puede observar cómo los usuarios de Brasil son los que más utilizan la plataforma, seguidos por Estados Unidos finalizando con Argentina, en el tercer puesto.



A pesar de que se podría esperar que la plataforma tuviera más presencia en los países vecinos, como Argentina, el gráfico nos muestra que tuvo más éxito en países que no pertenecen a Latinoamérica.

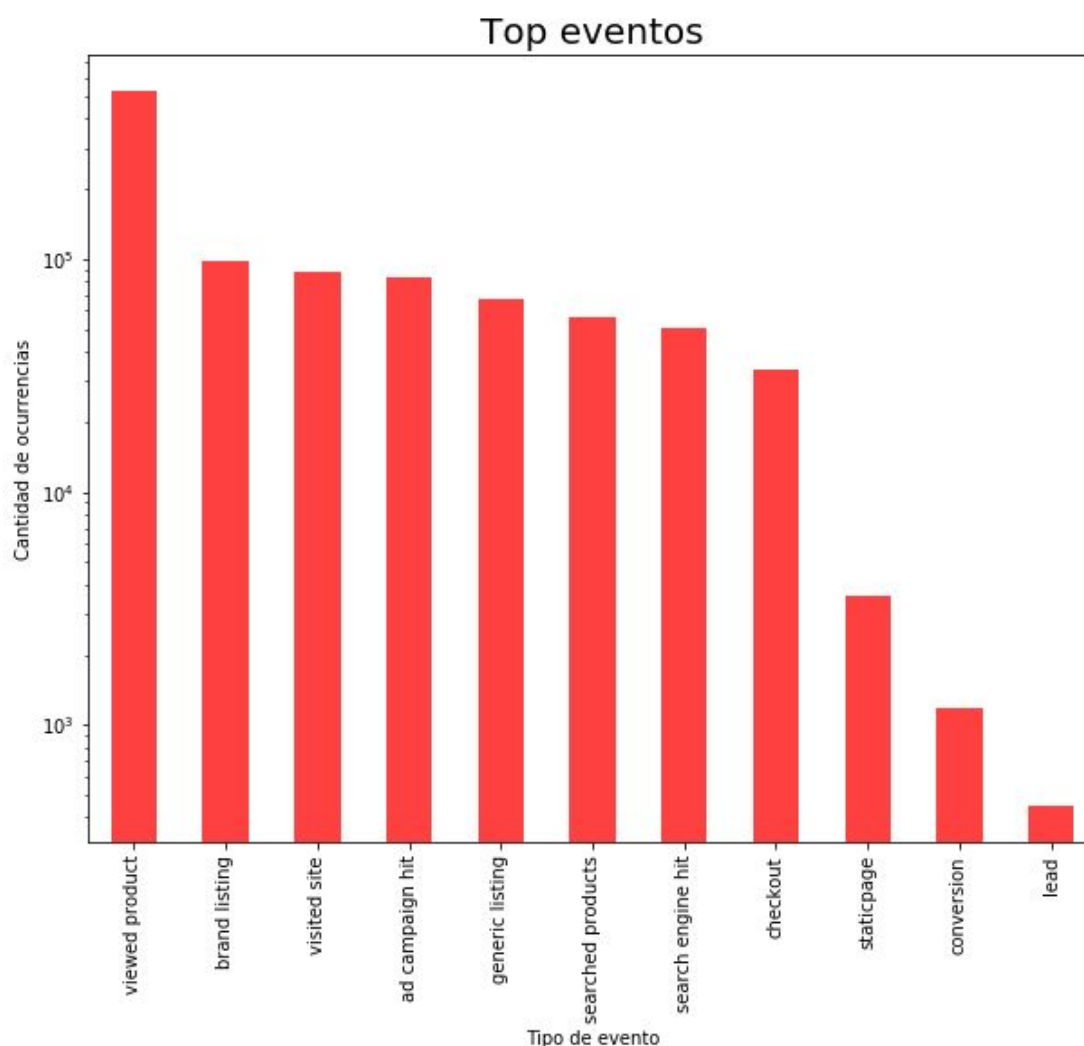
Pensamos que el bajo éxito en la región, puede ser por la falta de confianza de los argentinos a la hora de comprar móviles usados. En los Estados Unidos es muy común la venta de celulares “refurbished” (reacondicionados), los cuales se pueden adquirir en cualquier local de telefonía móvil. Esta práctica no está tan arraigada en Argentina, lo cual podría explicar la diferencia entre estos dos.

Luego de tener una primera visión sobre la distribución geográfica de los usuarios que utilizan la plataforma, nos centramos en analizar a los 7 usuarios que más utilizaron la plataforma y estudiar la proporción entre los distintos tipos de eventos que realizaron con la intención de hallar patrones.



Como era de esperarse, notamos un patrón en el comportamiento de los usuarios. El evento que más realizaron, fue el de mirar un producto, lo cual nos parece muy lógico. Estos usuarios generaron, también, una gran cantidad de entradas en el evento “brand listing”. Este hecho nos hace pensar que el sistema de recomendación de modelos funciona de manera correcta, generando el interés necesario en el usuario que entra a ver un producto específico y sigue navegando en busca de otros modelos recomendados.

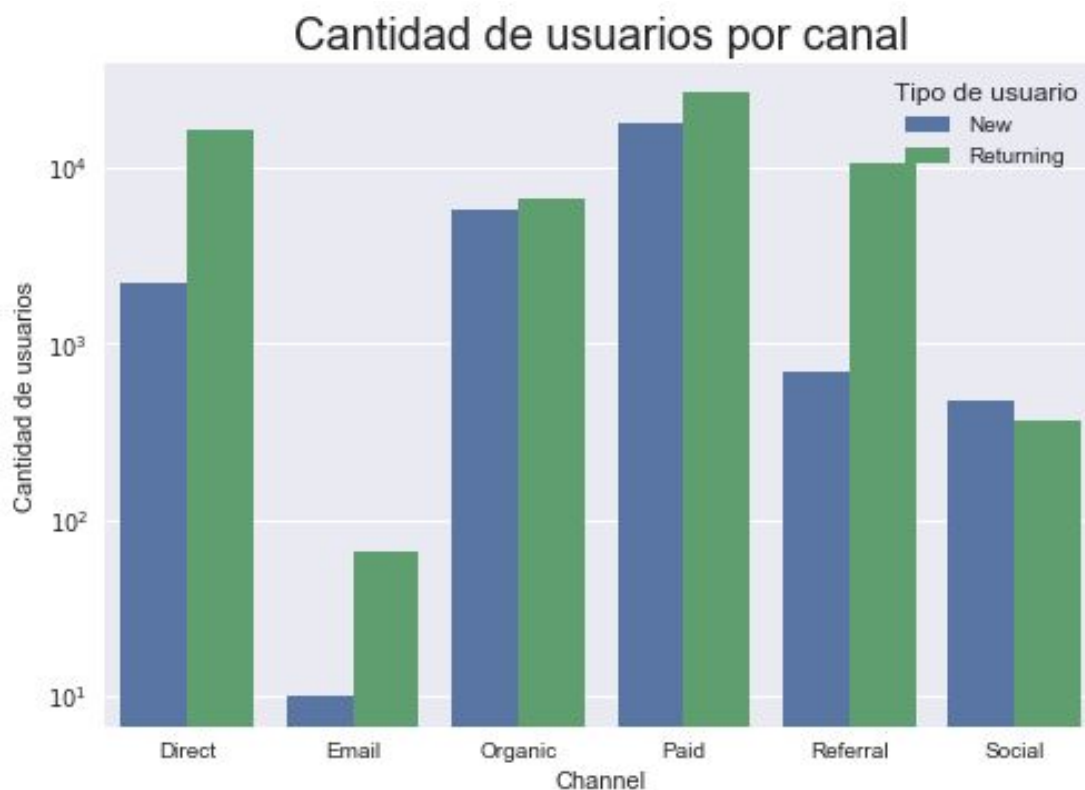
Al abordar los eventos realizados por los usuarios en general, tuvimos como resultado el siguiente gráfico.



Podemos observar que la mayor parte de los registros, corresponde con la visita de un producto en específico. A partir de aquí, una cantidad menor llega a la fase de checkout. Sin embargo, una cantidad mínima termina realizando la compra, lo que se ve reflejado en la categoría de conversión en el gráfico anterior.

Esto genera incertidumbre al respecto, ya que podría llegar a ser un problema para la plataforma. Es de esperar que una mayor proporción de los usuarios que llegan hasta la instancia de checkout, finalicen con éxito la compra, generando una ganancia para la plataforma. Esto puede deberse a la familiaridad que tengan con la página. Un usuario que ya conoce Trocafone, supone un cliente con mayor probabilidad de compra por el historial que tiene con la plataforma.

Es por eso que decidimos analizar más exhaustivamente la proporción de usuarios teniendo como prioridad su status de nuevo o reincidente y cuál había sido el canal por el cual habían llegado a Trocafone.



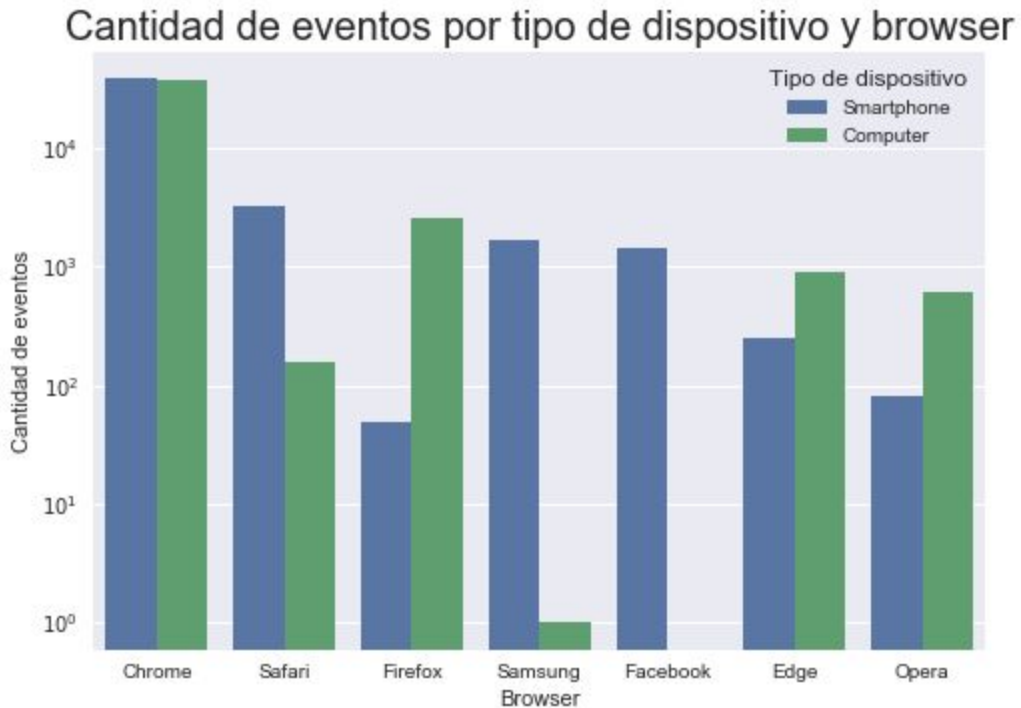
Podemos observar varias cosas. Por un lado, sea cual sea el tipo de canal, se puede apreciar la gran tasa de permanencia que tiene el usuario con la plataforma. Esto es especialmente positivo, ya que un usuario que vuelve, significa una recomendación del servicio.

Esto ayuda a entender mejor el gráfico anterior. Se puede concluir que un usuario medio probablemente visite más de una vez una página de un producto antes de comprarlo. Esto explica la gran cantidad de “viewed products” que se observan. Este tipo de comportamiento es más común en un usuario familiarizado con la plataforma.

Observando cada uno de los canales que se utilizan, se puede concluir que el servicio de propaganda que representan los mails, es el menos eficiente con diferencia. Esto sucede porque la tendencia de un usuario nuevo hacia una marca con la que no está familiarizada, es marcar el mail en la casilla de spam, mermando enormemente su efectividad.

Desconociendo el impacto económico que significa para la compañía el canal ‘paid’, no podemos comentarlo. Sin embargo, sí queremos destacar el papel que tiene el canal directo, como principal captador de clientes. Finalmente, el último canal interesante son las redes sociales. Su papel es destacable por ser el único canal que genera más usuarios nuevos que reincidentes.

Otra característica relevante a analizar es con qué tipo de dispositivo y browser ingresan los usuarios. Como Trocafone funciona sobre una plataforma web, estas dos características deberían ser muy importantes. En el siguiente gráfico, podemos ver los resultados que obtuvimos de la cantidad de eventos según el tipo de dispositivo y el browser usado.



Como se observa Chrome lidera como el browser más elegido por usuarios, tanto de dispositivos smartphones como de computadoras. En smartphones, es seguido por safari. En computadoras, en cambio, por Firefox. Esto es de importancia a la hora de desarrollar los entornos donde interactúan con el usuario. También, dependiendo de el browser, se podría asociar al usuario con una marca. Por ejemplo, alguien que ingresa con safari puede que tenga preferencia por productos de la marca Apple.

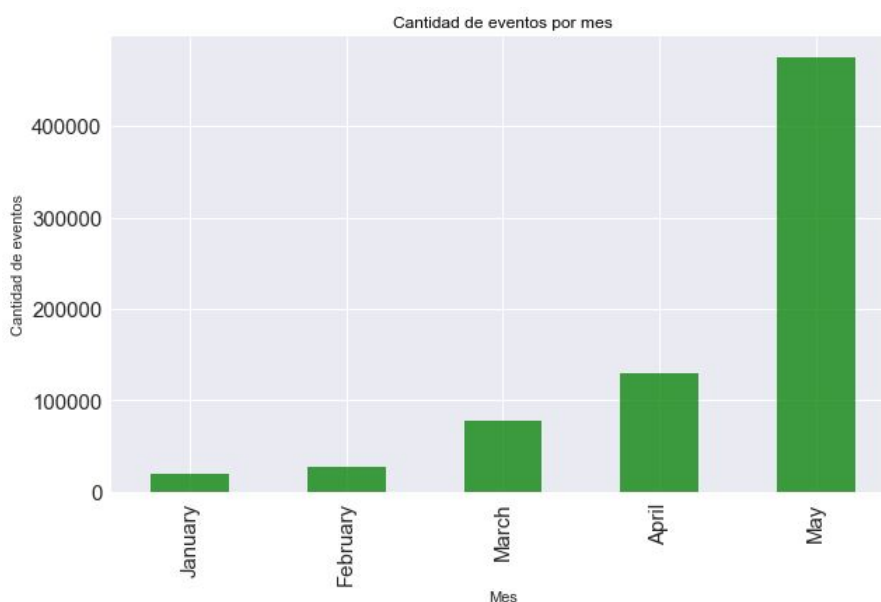
Con los datos que nos fueron provistos, este es el maximo análisis que se puede realizar acerca de los usuarios.

Análisis Temporal

Luego de haber realizado un análisis sobre los usuarios, decidimos analizar la temporalidad de los eventos que estos realizaron.

Para ello, utilizamos el feature timestamp, el cual indica la fecha y hora del evento. A partir de este pudimos ver que el set de datos era de los primeros 6 meses de 2018, aunque el mes de junio no está completo sino que tiene datos sobre los primeros 15 días.

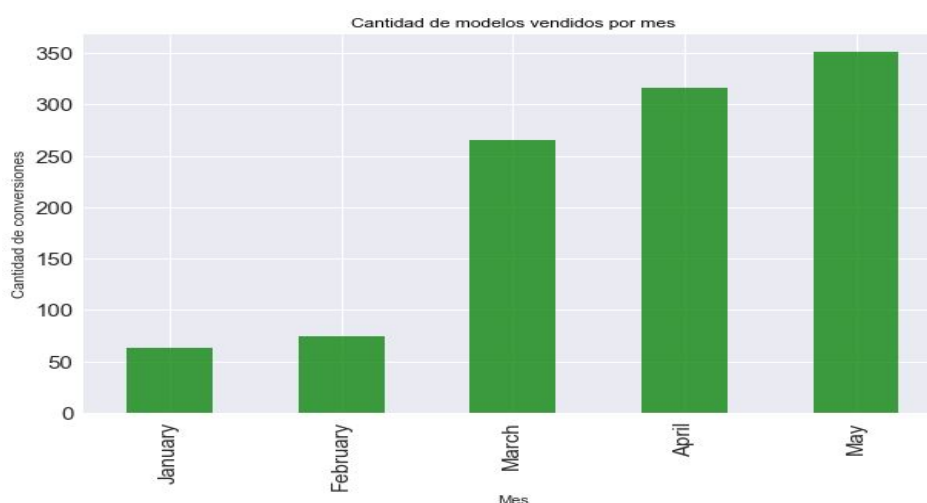
Notamos algo particular. En los primeros 5 meses de 2018, las cantidades de eventos eran muy distintas por mes, las cual mostraremos con el siguiente gráfico:



No se tiene en cuenta al mes de junio porque no está completo en el set

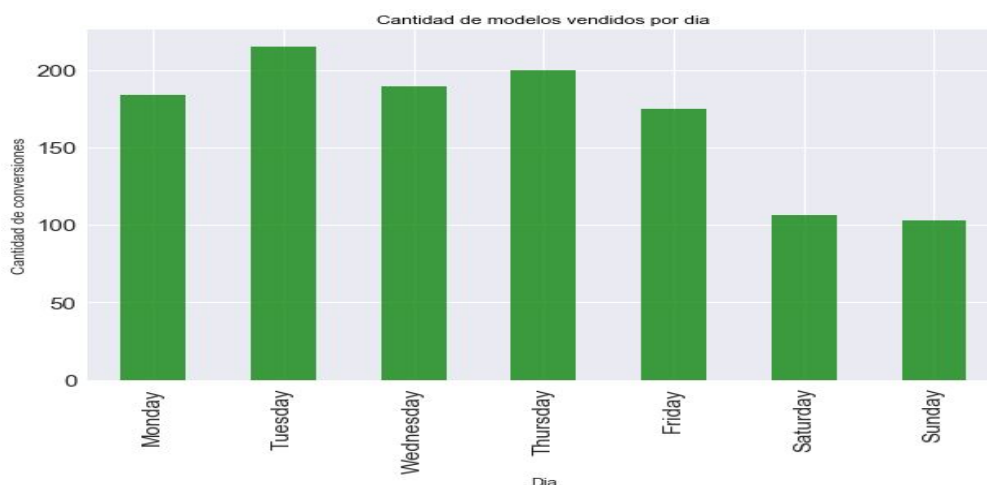
Como podemos ver, a medida que pasaron los meses, la cantidad de eventos aumentó de manera muy abrupta. Las razones pueden ser varias. No lo sabemos. Pero podemos suponer que el set de datos está formado por datos particulares como, por ejemplo, un conjunto de usuarios y su actividad durante estos meses. Por lo tanto, un análisis temporal demasiado profundo puede que no sea la mejor opción en base a este set, ya que por las proporciones de las cantidad de eventos por meses son muy desproporcionadas. Por lo tanto no son comparables entre ellos.

De todas maneras analizaremos algunos aspectos temporales que de todas formas puede llegar a resultar de interés. Uno de ellos, es la proporción entre ventas y eventos generales por mes.



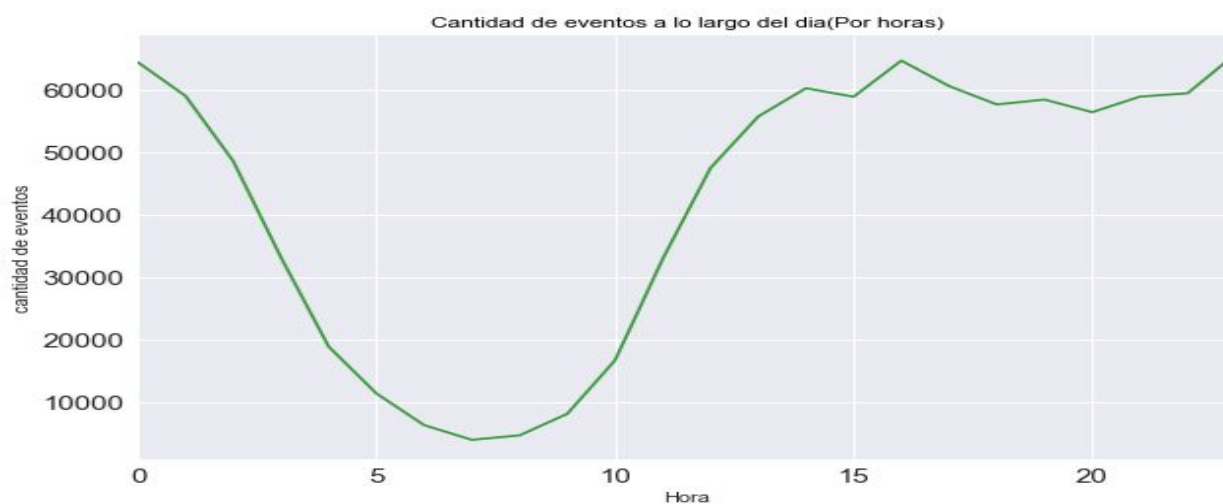
Podemos ver que la cantidad de ventas por mes también aumentan a medida que pasa el mes pero a comparación de los eventos en los tres últimos meses no es tan desproporcionada, se esperaría que Mayo tenga muchas más ventas dado que tiene muchos más eventos que los de mas meses pero no es así.

Algo que es comparable es la cantidad de compras por día de la semana. De esta manera podemos ver que días hay más actividad y, a partir de esta información, se podrían tomar decisiones de marketing que puedan ser beneficiosas para la empresa. En el siguiente gráfico veremos cómo varía la cantidad de compras por día de la semana.



Se observa que los fines de semana son donde menos ventas hay y, por lo tanto, durante la semana se concentra la mayoría de las ventas.

Otro aspecto temporal a analizar, serían los horarios donde se concentra más actividad y donde se realizan más compras. En los siguiente dos gráficos mostraremos esto:



En este gráfico, podemos ver que la actividad de los usuarios, es decir, la cantidad de eventos, decae durante la madrugada/mañana: entre los horarios aprox 4 a 10 am. Esto es intuitivo ya que es el horario donde la mayoría de las personas duermen.

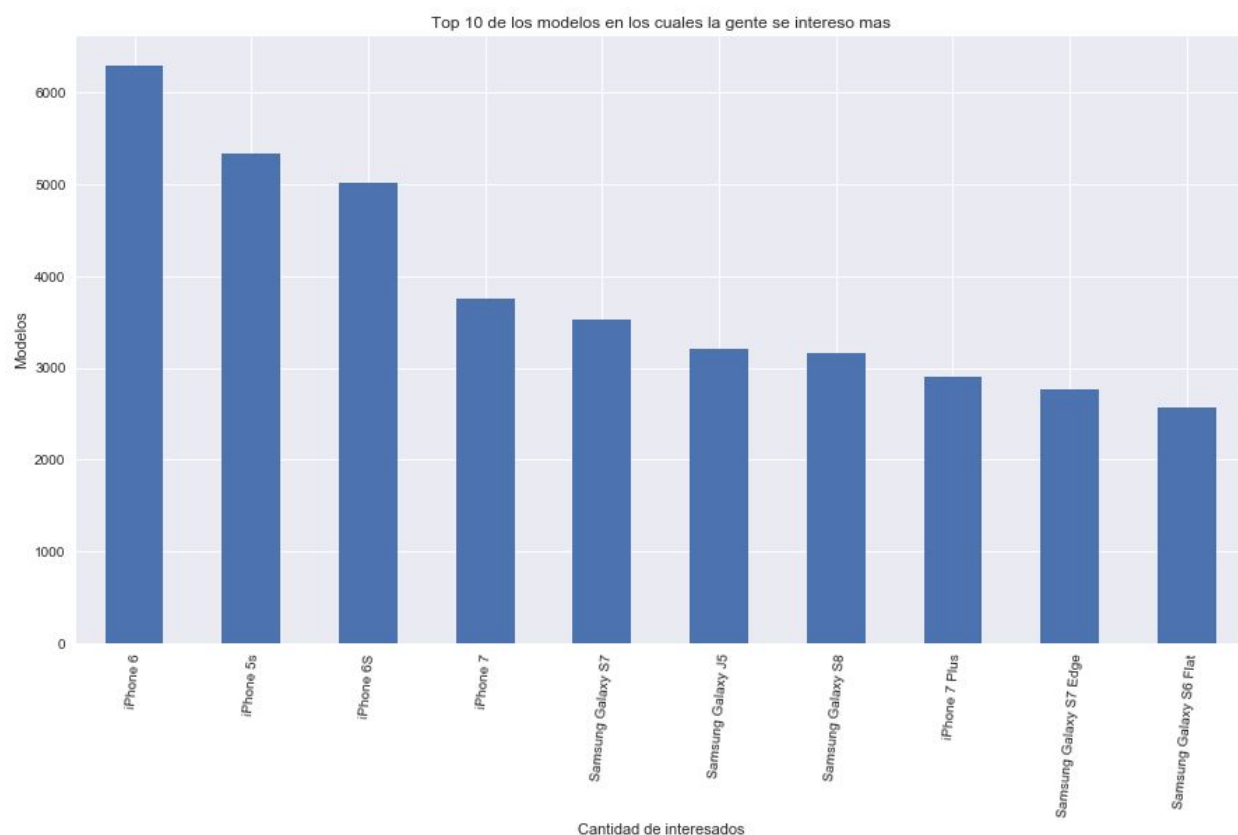


En este último gráfico también se puede ver la actividad de los usuarios, pero ahora ,solamente nos basamos en las compras. Se podría decir que la mayoría de las compras ocurren de las 10 de la mañana en adelante. Como en el anterior gráfico, se nota que en la madrugada/mañana es donde hay menos cantidad de compras en este caso.

Análisis de dispositivos.

Marca lider

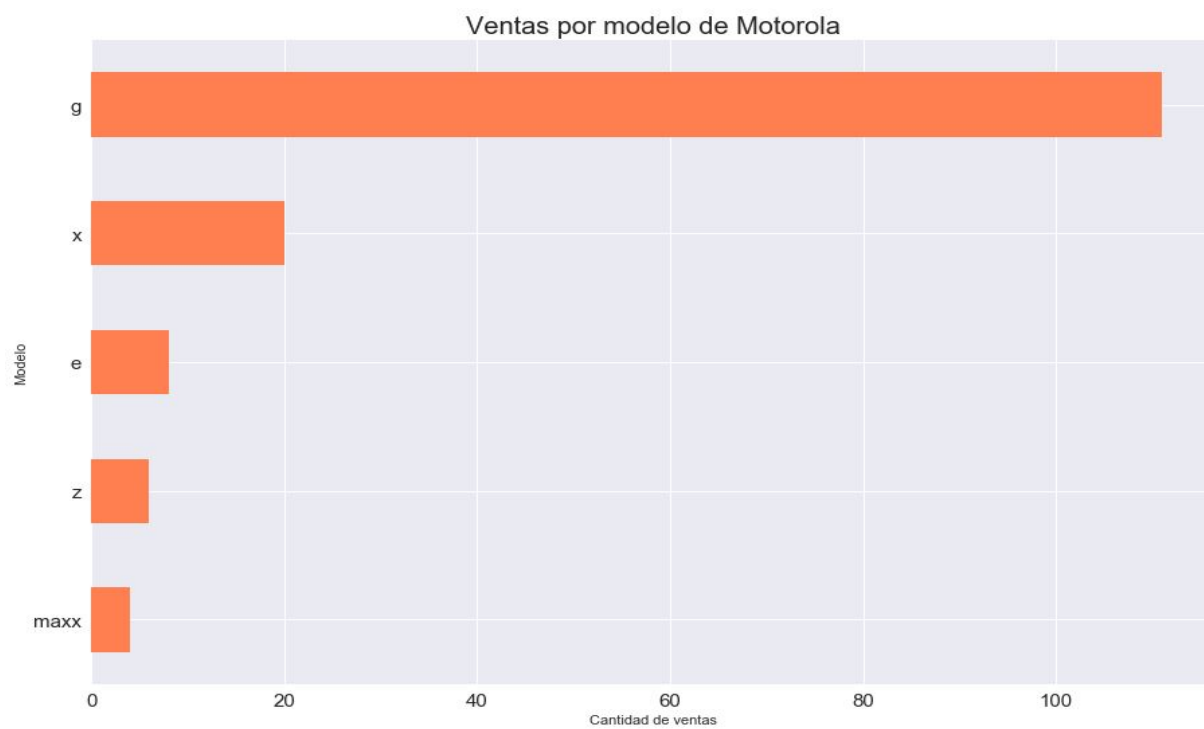
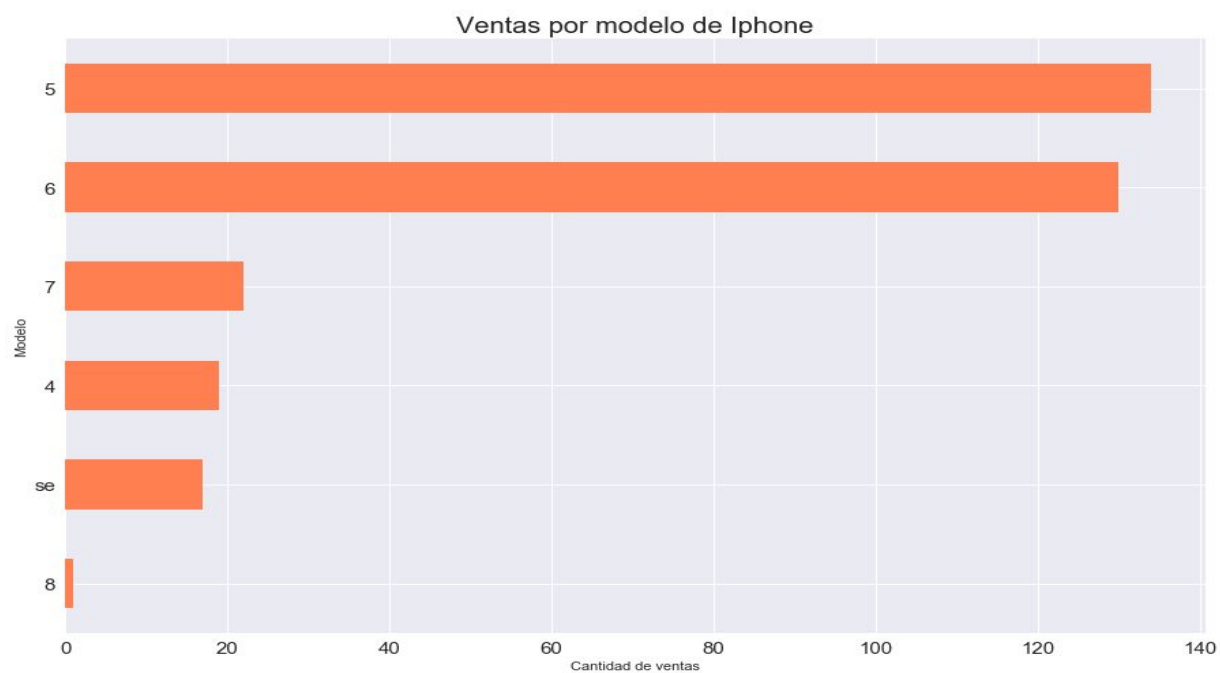
Habiendo analizado el perfil del usuario de Trocafone, el siguiente paso lógico, es indagar sobre los productos. Empezando por lo básico, a continuación se muestra un gráfico que esquematiza los modelos más populares.

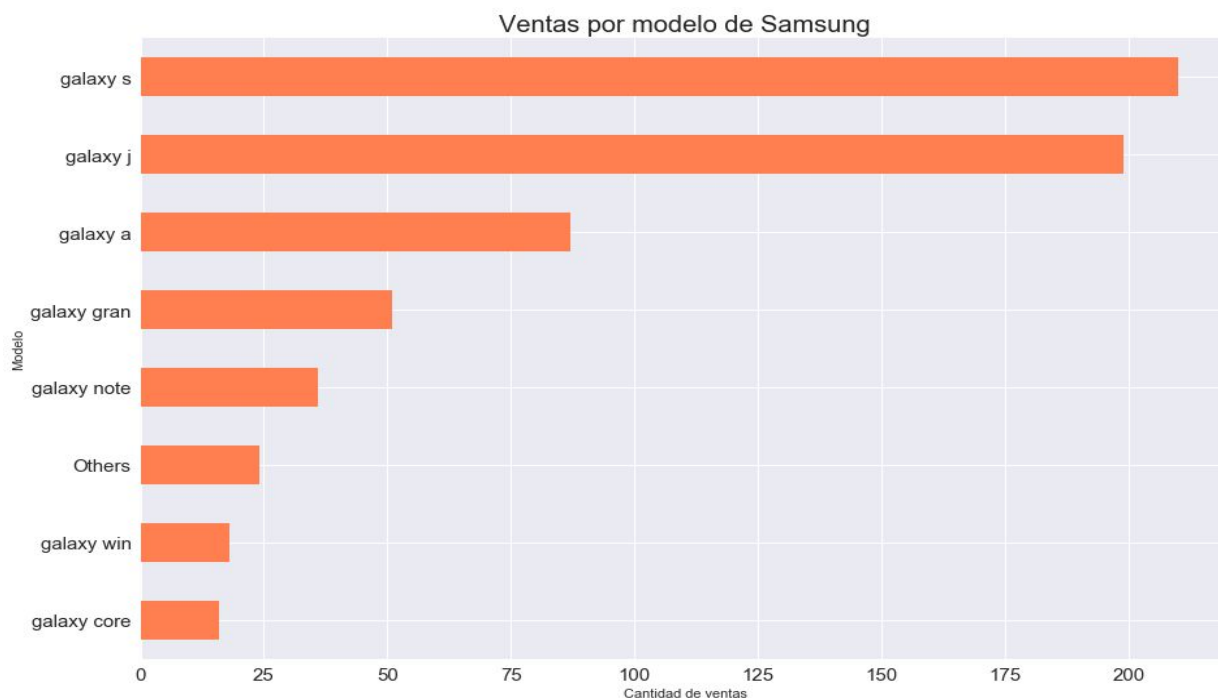


La predilección por los celulares de la marca Apple se nota claramente al ocupar estos los primeros cuatro puestos. Con esta observación, se propone potenciar el reacondicionamiento de los dispositivos de esta marca ya que, queda demostrado, es un buen negocio. Aun así, esto, se contradice con lo siguiente:



Este gráfico nos muestra una tendencia. Los usuarios buscan en su mayoría dispositivos de la marca Apple, pero la relación entre búsquedas y compras efectivas es mucho menor que la de los celulares de la marca Samsung. Habiendo visto esto, se procedió a un análisis más exhaustivo de las ventas para estas tres marcas:

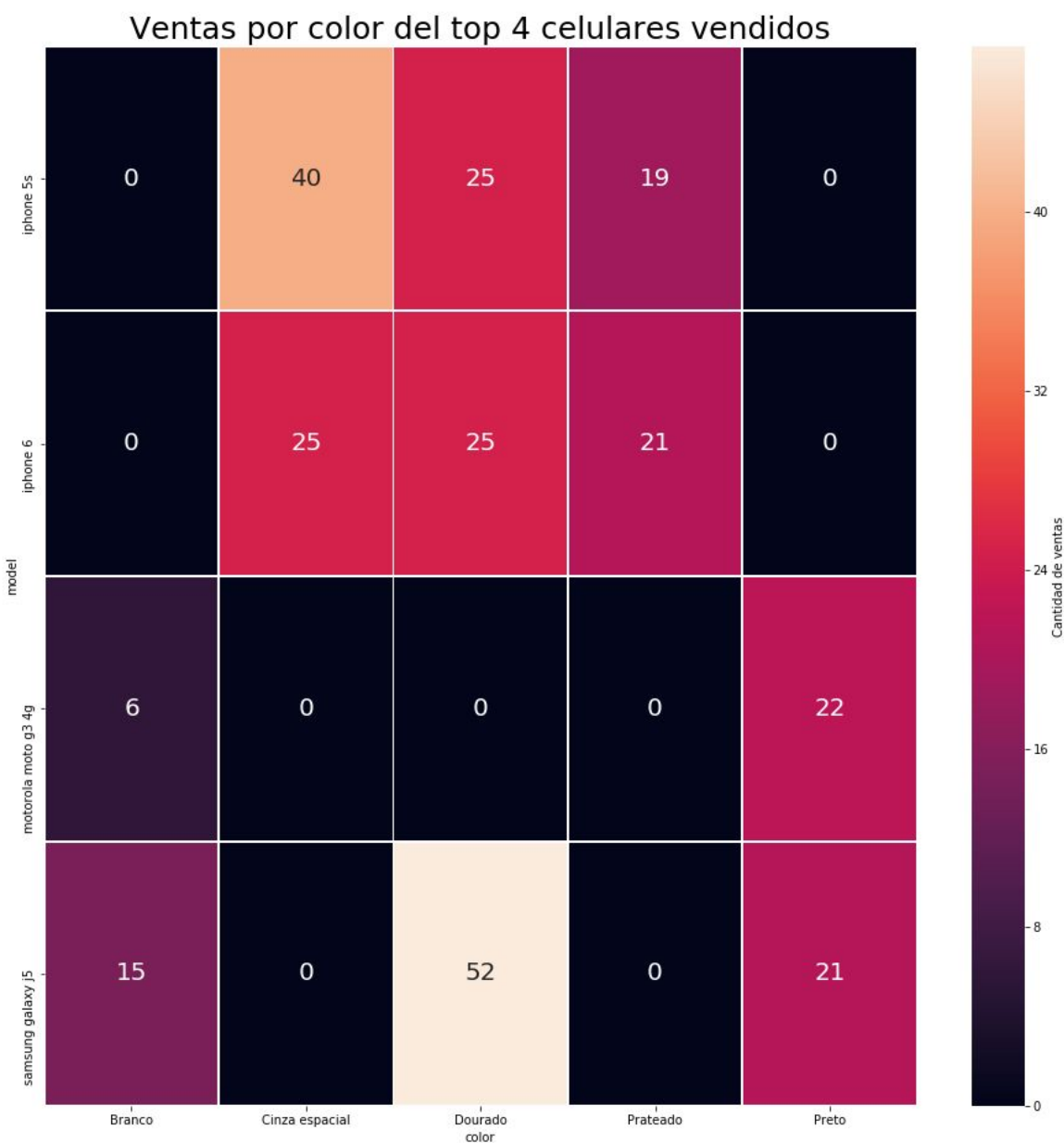




Con estos tres gráficos queda más clara la situación. Las ventas de la marca Samsung para sus dos modelos estrella, son considerablemente mayores que las de Apple y Motorola. La razones detrás de esto son variadas. Por ejemplo, el reacondicionamiento de un móvil de Apple, es más caro debido al precio de los repuestos. Esto, influye de manera inevitable en el precio que percibe el usuario final. Por otra parte. Para realizar un examen más exhaustivo de esta situación, se necesitan datos de precios que no han sido provistos. Para llegar a una conclusión válida, se buscó otro enfoque.

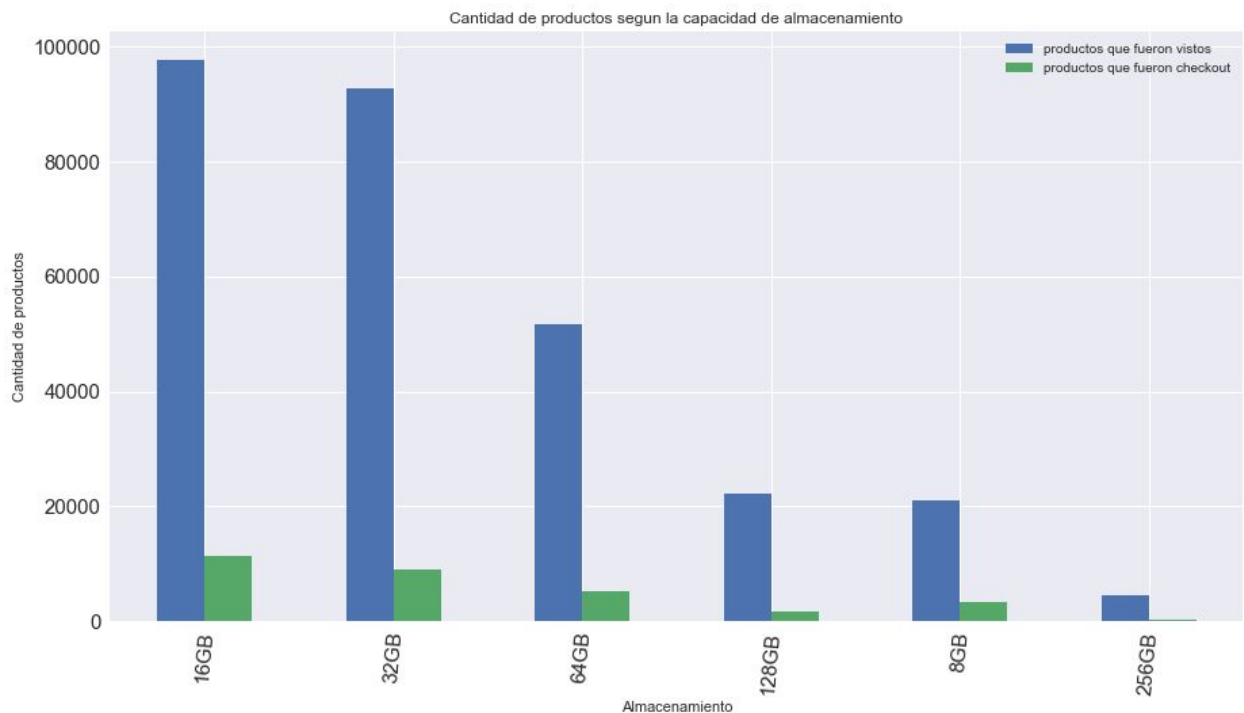
Características comunes

Como no se puede hacer un examen competitivo entre marcas, basaremos los siguientes análisis en características que compartan las terminales de ambas compañías, para buscar otro tipo de patrones. Empezaremos con uno estético, el color. Para ello, analizaremos las combinaciones de colores de celulares que son comprados.

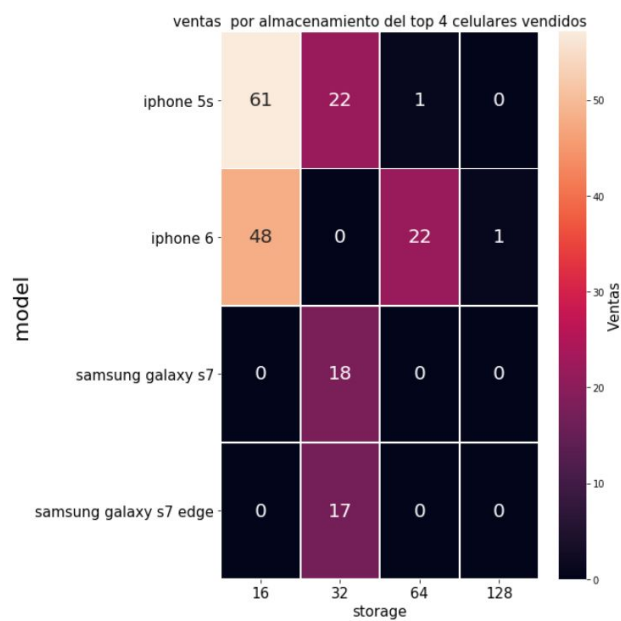
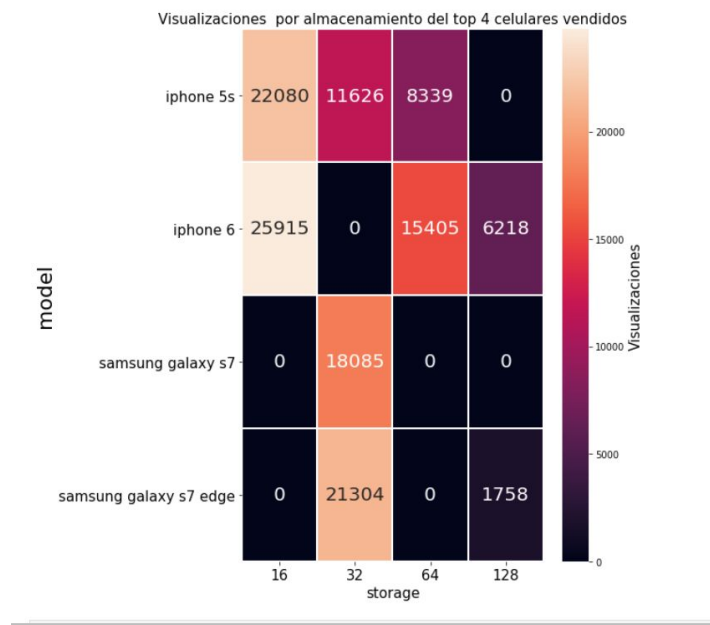


Como conclusión obvia, se puede llegar a que los colores metálicos están siendo tendencia entre los nuevos celulares. La razón detrás de esto, es que el color metálico, especialmente el color dorado, le confiere a un dispositivo el estatus de especial, deluxe, etc, por lo que atrae mas al publico.

Siguiendo con las características en común, tenemos el almacenamiento. Siguiendo las pautas hasta ahora, tendria sentido que los dispositivos con mayor capacidad sean los que más vende. Esto es cierto en parte.



Como se ve en el gráfico, los dispositivos con menor capacidad de almacenamiento (16 ~ 32GB) son los más vistos. Sin embargo, también tienen la peor tasa de conversión a compras. Antes de llegar a una conclusión, veamos un gráfico un par de gráficos más en los cuales tomaremos como referencia 4 modelos con gran variedad de almacenamientos disponibles.



Con esta nueva información, se puede ampliar el panorama hasta ahora descrito, percibiendo una nueva tendencia de los consumidores. Siguiendo las pautas anteriormente detalladas en el apartado del

color, se esperaría que, si el usuario decide un color en base a un status de calidad, el mismo razonamiento se utiliza en cuestión de almacenamiento. Sin embargo a diferencia de las distintas gamas cromáticas, las distintas opciones en el ámbito de memoria del dispositivo, si modifican el precio del mismo. Para entender mejor esto, tomemos como ejemplo dos casos: el Samsung Galaxy S7 Edge y el iPhone 5s.

Como vemos en el primer gráfico, se tiene un interés inicial enorme para con ambos dispositivos en una amplia variedad de almacenamientos. No solo eso, las dos opciones con mayor capacidad tienen una distribución equiparable a pesar de que cuentan con precios completamente distintos y estar destinados a dos públicos potenciales distintos. Esta diferencia final, se denota en el segundo gráfico y es justamente el comportamiento que queremos destacar.

Aquellas personas que buscan un iPhone 5s, el modelo más económico, tienden a elegir un terminal con menor memoria, descartando el almacenamiento extra. Por otro lado, con el Samsung Galaxy S7 Edge sucede lo contrario. Existe una porción considerable de los compradores que deciden mantenerse con la elección más cara. Teniendo esto como referencia, se recomienda llevar a cabo optimizaciones del sistema de recomendaciones de la página.

Análisis regional.

Como se explicó y demostró al principio, Trocafone tiene como principales clientes a los residentes de Brasil. Es por ello que la mayoría de los datos proveídos, representan a usuarios del país y es, por lo tanto, de vital importancia analizar las características intrínsecas de este mercado.

Se puede prever que las ciudades más grandes representan un mayor número de clientes. A modo de pantallazo inicial y para constatar esto, se creó la visualización que se muestra a continuación.

Regiones con mayor cantidad de visitas al sitio



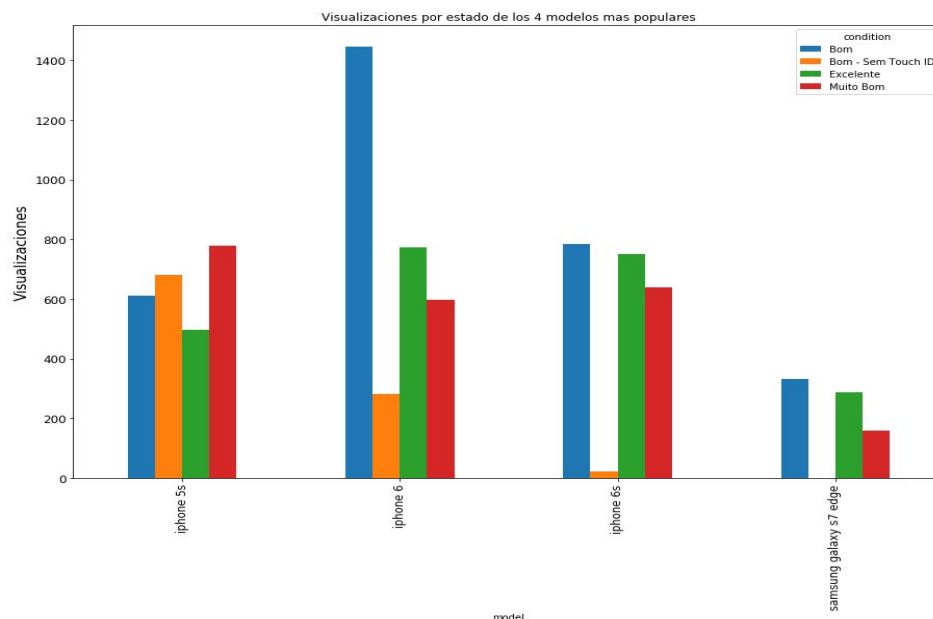
Como se mencionó, las regiones con mayor población son las que más han visitado la página, destacando Minas Gerais como la primera. Sin embargo asombra que, haya regiones del calibre de Sao Paulo que apenas tengan clientes. Dado el volumen de datos, no podemos llegar a la conclusión de que sean regiones en donde la propuesta de Trocafone no funcione.

A partir de acá, se vuelve más interesante realizar un examen más exhaustivo del mercado.

Entre las actividades que se realizan en Trocafone, reacondicionar los dispositivos que posteriormente se venderán, está entre las más importantes económicamente. Es necesario saber qué dispositivos es rentable reparar, es decir, saber cuáles son aquellos modelos y/o marcas de dispositivos que más se adquieren por medio de la plataforma.

Sin embargo, también hay otro factor a tener en cuenta. El estado final del dispositivo. Cuanto mejor sea el estado, mayor será el precio, lo cual puede perjudicar la cantidad de usuarios que pueden adquirirlos. Por otra parte, si el estado no es cercano al óptimo, el producto generará desinterés.

Teniendo esto en cuenta y, nuevamente analizando este mercado en particular, se buscó cuáles eran los celulares con mayor salida y cuáles eran sus estados al momento de la venta. Primero, se analiza qué dispositivos generan interés. Como medida de intención de compra, se analizan cuales son los dispositivos que llegan a la fase de checkout y la condición del mismo.

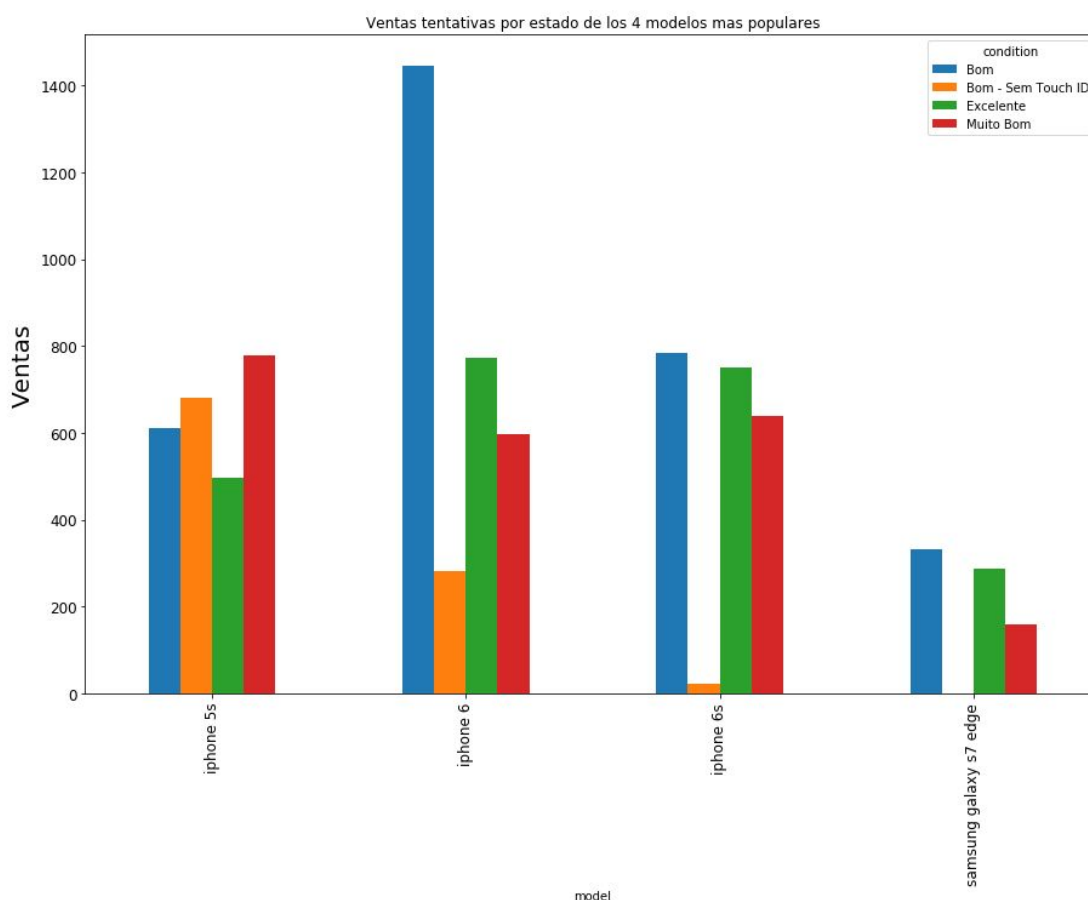


Se dejó la clasificación original de los dispositivos por motivos que se aclaran más adelante.

En primera instancia, observamos una gran diferencia en cuanto a lo que marca se refiere. Los dispositivos de Apple lideran con clara ventaja la preferencia de los usuarios. Si bien hubo importantes avances de la marca en Brasil en los últimos años, no deja de ser destacable la diferencia.

Lo que sí se podía intuir era que, aquellos modelos que se encuentran entre el mejor y peor estado, son vistos como las peores posibilidades de compra para el usuario brasileño. Esto se debe en gran medida a cómo está diseñada la página web de Trocafone.

Sin embargo, nuestro análisis no finaliza aquí. Si analizamos los dispositivos que pasan la fase de checkout veremos lo siguiente.



Una técnica de marketing muy utilizada es la siguiente. Si quiero vender el producto **A**, pongo entre mis opciones de venta un producto **B**

mucho más caro y otro, **C**, mucho más barato y de peor calidad. Está demostrado que de esta forma, la percepción a nivel neuronal del comprador sobre el producto **A** mejora muchísimo. Sin embargo, también se puede realizar una variante de este método que es lo realizado aquí.

Siguiendo con el ejemplo anterior, en este caso lo que queremos es vender son **B** y **C**. El producto **B** nos representa mayor ganancia a mayor inversión. El producto **C**, nos abre la posibilidad de obtener ganancias con poca inversión. En cambio, el **A** es una apuesta mayor con un margen de ganancia estándar. En casos así, lo que se hace es difuminar las líneas que separan productos. Para el comprador, la diferencia entre Excelente y Muito bom (muy bien) es difusa, por lo que ante la duda, buscará cuidar su capital no malgastándolo. De ahí, que la proporción visualización/venta del estado Muito Bom sea la peor. Pero no es lo único que se puede inferir.

Como se observa, el estado (condición) de dispositivo que genera mayor cantidad de compras, es el bueno, nuestro producto **C**. La razón detrás de esto, es lo explicado anteriormente. Al no estar seguro entre las diferencias entre dos categorías, la otra decisión razonable por parte del usuario es comprar el más barato.

Anexo técnico de los datos

Para poder realizar los análisis anteriores, trabajamos sobre el set que nos daban para poder entender los datos y, de esta manera, comprender mejor la información que nos brindaban.

En primer lugar, notamos que había una gran cantidad de NaN en varios features. Por lo tanto, decidimos empezar a analizar qué relaciones tenían los features entre sí para así poder rellenar esos datos faltantes.

Encontramos que en los eventos de visited site, los campos de la ubicaciones del usuario(país, región y ciudad) estaban completos. Por ello, decidimos tomar esos datos y asignarle a cada usuario en todos los eventos su determinado país, región y ciudad. De esta manera pudimos completar los features country, region y city.

Notamos que al feature “model” también le faltaban muchos datos, por lo que utilizamos el feature “url” para poder completarlo. Este último estaba asociado al evento “ad campaign hit”. Es por ello que, sacando la información sobre el modelo de este url, pudimos rellenar los campos de model asociados al evento antes mencionado.

En algunos casos creamos features nuevos. A la hora de realizar el análisis temporal, utilizamos la librería datetime, aplicando sus métodos sobre el feature timestamp pudimos obtener años,mes,día,hora.

Para el análisis de los modelos de celular, diferenciamos la marca por un lado y el modelo por otro. Esto nos sirvió para poder ver cuales eran las marcas con más ventas y dentro de esas marcas cuáles eran sus modelos estrellas.

Conclusión.

Luego de haber analizado el set de datos en lo que consideramos, su totalidad, llegamos a las siguientes conclusiones. Consideramos de vital importancia la implementación de una versión para móviles de la plataforma. Ya sea con una app y/o una versión para el navegador. Como se mostró en el apartado, Análisis de Usuario, un gran porcentaje de los usuarios deciden ingresar a Trocafone desde sus smartphone.

Siguiendo con la plataforma en sí, se ha notado una tendencia en el comportamiento de los usuarios. Un gran porcentaje de los mismos, no completa la transacción, es decir, solo llegan a la fase de checkout. Nuestra recomendación es la de simplificar este último paso de la transacción, promoviendo que el usuario compre más ágilmente.

Cambiando de enfoque, es necesaria una campaña publicitaria más fuerte en países como Argentina donde el servicio es totalmente nuevo. Para que Trocafone funcione en una región, es necesaria una gran cantidad de usuarios concentrada, como es el caso de Brasil.

Finalmente, en el aspecto de productos, con este informe se han mostrado las tendencias de los usuarios de Trocafone. Tanto Samsung como Apple, destacan del resto como las marcas favoritas. Ante esto, hay dos accionares distintos. Se puede potenciar esta segregación entre el público, o disminuirla. Dado el modelo de negocios de Trocafone, nuestra recomendación es que se potencie. Si se compran mas unidades de esta marca, su venta es más probable. Para ello, se pueden aplicar distintas medidas. Entre las recomendadas, se encuentra la de aprovechar el dato del navegador utilizado. Un usuario que ingrese desde Safari, tiene mayor probabilidad de buscar un iPhone, por ejemplo. Con esto, damos por concluido nuestro análisis de la plataforma Trocafone.
