

**UNIVERSIDAD DE BUENOS AIRES FACULTAD DE INGENIERÍA**



75.06 Organización de Datos

Trabajo Práctico N°1

Análisis exploratorio de datos de Jampp

Primer Cuatrimestre de 2019

**Grupo 43**

**Repositorio:** www.github.com/cruzpa95/Datos2019

|  |  |  |
| --- | --- | --- |
| **Alumno** | **Padrón** | **Mail** |
| Cenizo, Facundo | 96657 | [facundocenizo@hotmail.com](mailto:facundocenizo@hotmail.com) |
| Cruz, Pablo | 97553 | [cruzpa95@gmail.com](mailto:cruzpa95@gmail.com) |
| Villalba, Braian | 97641 | [braian.e.a.villalba@gmail.com](mailto:braian.e.a.villalba@gmail.com) |

Introducción

1.1 Acerca de Jampp



Jampp es una plataforma de marketing para adquirir y captar clientes móviles. La compañía combina comportamiento de datos con tecnología predictiva para generar beneficios para los anunciantes al mostrar anuncios relevantes y personales que inspiran a los consumidores a comprar por primera vez o más a menudo.

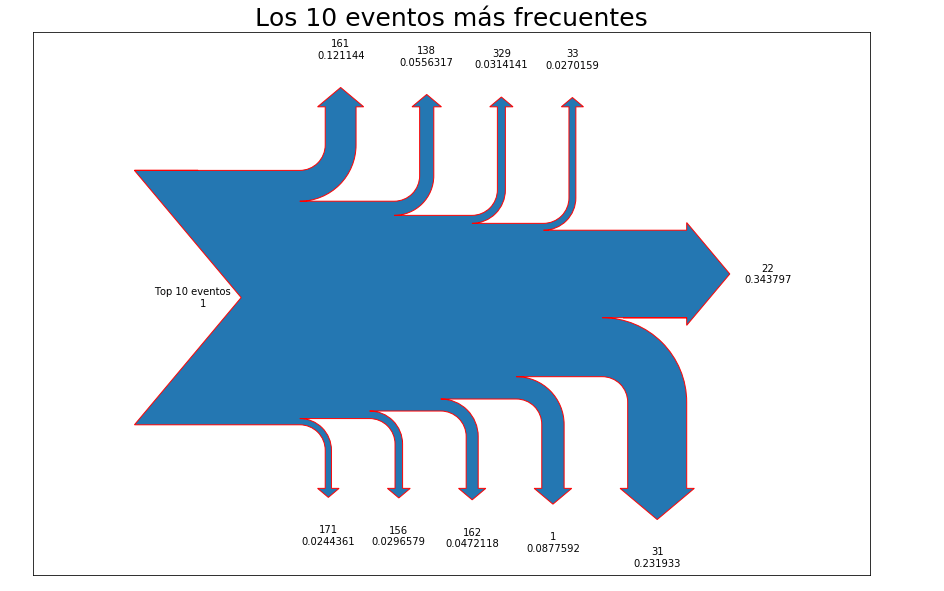
Fundado en 2013, Jampp atiende a una base de clientes global desde las oficinas de San Francisco, Londres, Berlin, São Paulo, Singapur, [Ciudad del Cabo](https://www.booking.com/city/za/cape-town.es.html) y Buenos Aires.

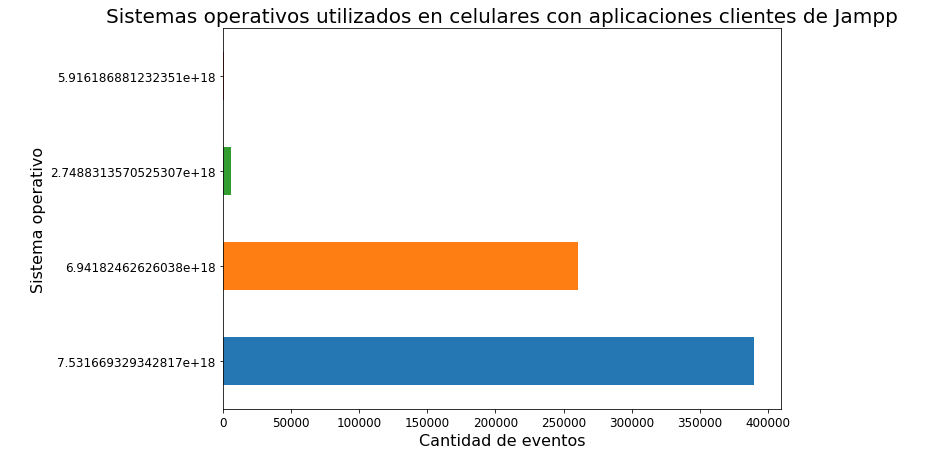
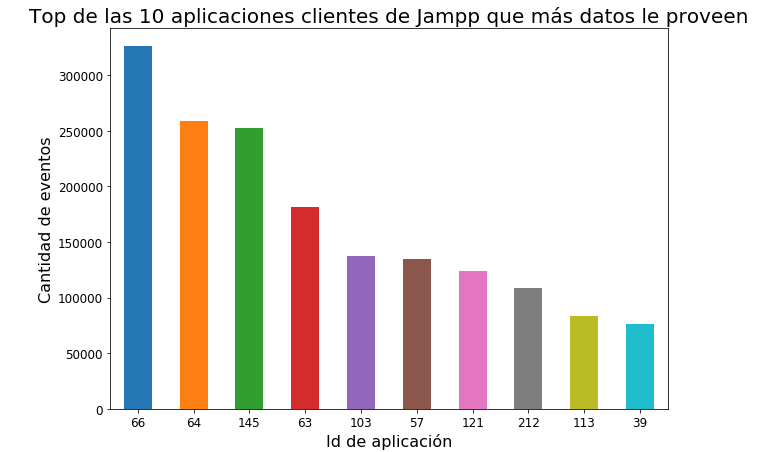
1.2 Objetivo

El trabajo práctico tiene como objetivo extraer la información que se considere más significativa para la empresa Jampp a partir de un conjunto de datos facilitados por un integrante de la compañia y realizar un análisis en profundidad que pueda resultarle útil a la misma.

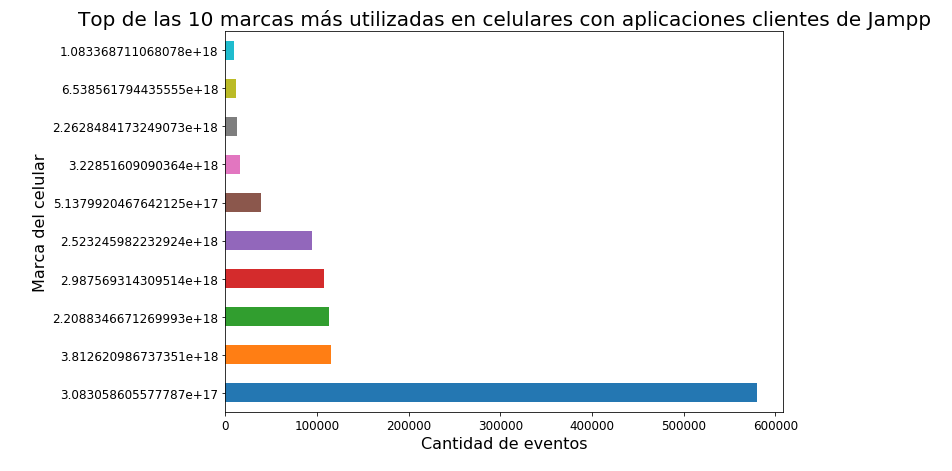
2. ANÁLISIS DEL SET DE DATOS (EVENTOS!!)

Comenzando por el set de datos de eventos, en el siguiente grafico se puede observar la distribución de los eventos mas frecuentes.

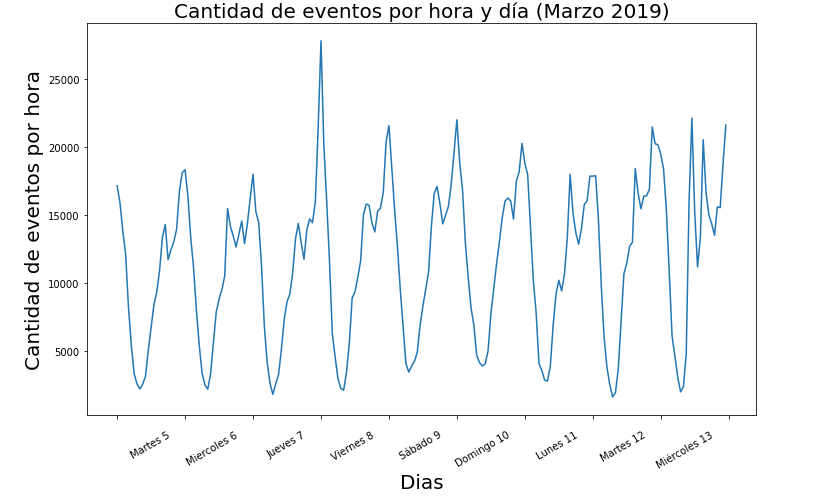




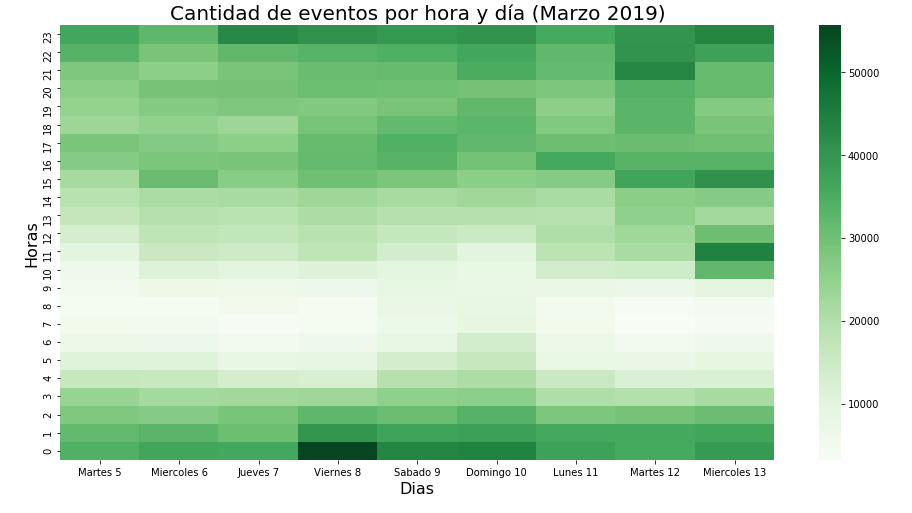
Se puede observar claramente que hay dos sistemas operativos que lideran los eventos. Suponemos que son Android y iOS.



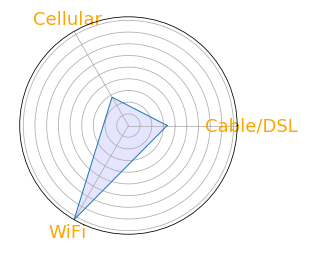
Se puede observar que hay una marca que supera a las demás en un gran porcentaje. Suponemos que es Samsung ya que líder mundial en venta de celulares.



El 8 de marzo se celebró el Black Friday

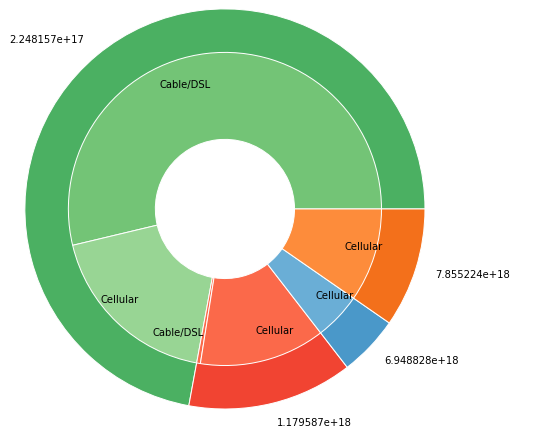


Se puede observar que, como es de esperar, la cantidad de eventos es casi nula en la franja horaria de las 4 a las 9 de la mañana, ya que es el momento donde la mayoría de la gente duerme o recién comienza a despertarse. También se puede notar que el día viernes 8 a las 00:00 horas se da el mayor número de eventos. Esto puede deberse a que ese día se celebró el Black Friday.



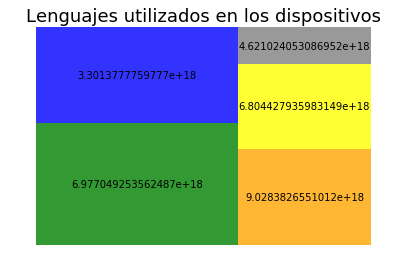
Para este análisis se consideró interesante presentar un gráfico de radar. En este tipo de esquemas se define según escala un valor a cada radio del radar y se divide la circunferencia en tantas partes como variables se quieran mostrar. Luego se le asigna el valor que alcanza cada variable y se genera el gráfico, en el cual se podrá ver a simple cuáles son las conexiones más frecuentes y su relación proporcional entre ellas. Se puede observar, como era de esperar, que en la mayoría de las aplicaciones clientes de Jampp se navega con una conexión de WiFi. En segundo lugar, no muy distante del tercero, se ubica la conexión mediante Cable/DSL, lo que se puede suponer es que, es muy probable que éstos sean juegos ya que en ellos se necesita una conexión más directa para que rinda al máximo. También se podría suponer que son computadoras que no tienen tecnología WiFi, pero esto es muy poco frecuente a esta altura de la tecnología. En tercer lugar quedó la conexión mediante red Cellular, que debe ser 3G, 4G, etc. Lo cual es lógico ya que generalmente se trata de ahorrar los datos para navegar que uno tiene en el celular debido a que son costosos.

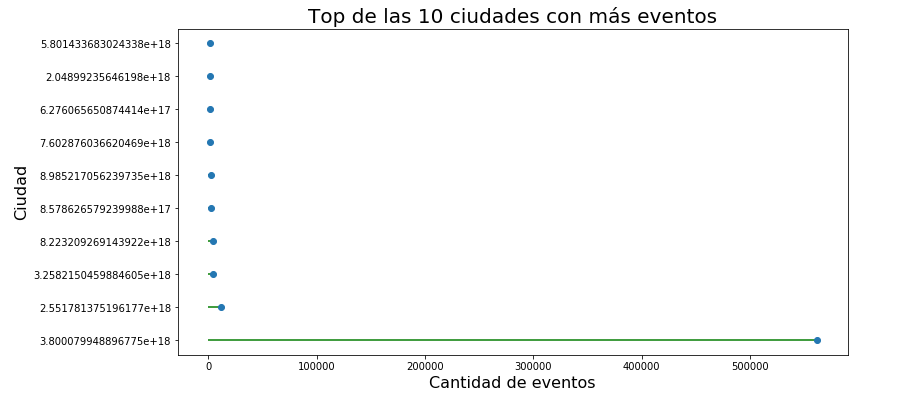
Se puede observar que hay cuatro compañías que se diferencian con las restantes en cuanto a usuarios, por lo que se graficarán sólo éstas.



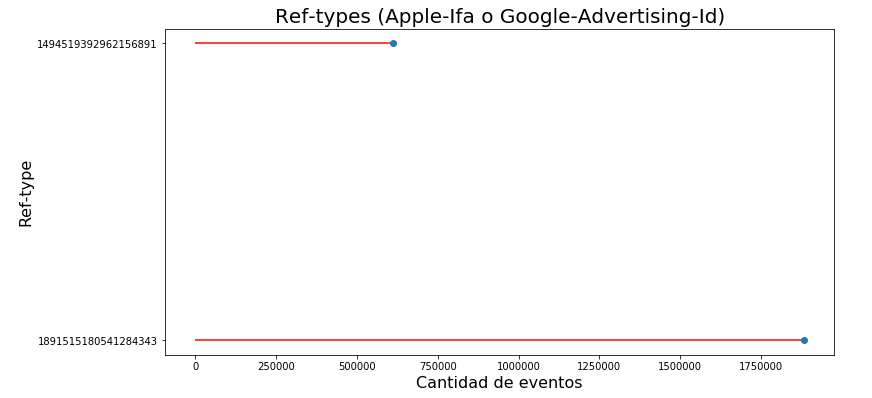
# Compañías proveedoras de internet y telefonía móvil

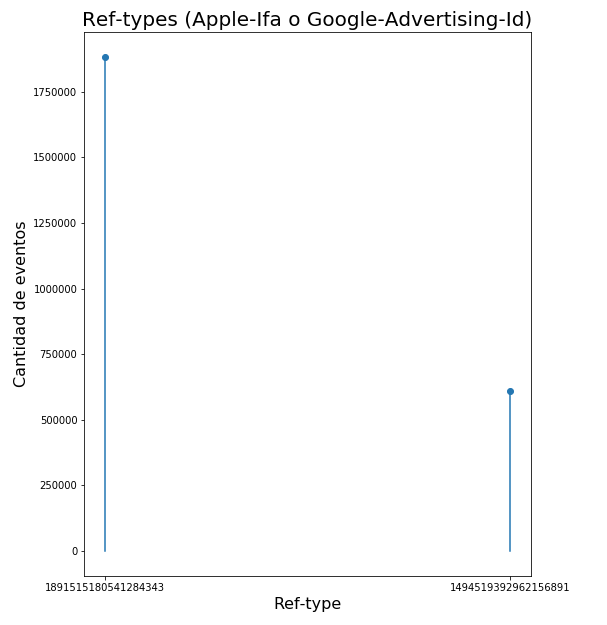
De las cuatro compañias analizadas, se puede observar que hay una que contiene a la gran mayoría de los usuarios, a los cuales les provee tanto el servicio de telefonía móvil así como el de internet. La segunda empresa con más usuarios (representada por el color rojo) también brinda los dos servicios, pero mayormente el de celular. Las dos restantes sólo tienen servicio para dispositivos móviles.

   
En el gráfico se muestran los cinco lenguajes más usados en los dispositivos. El que está representado por el cuadrado con más área (el verde) se podría afirmar que es el "español", seguido del "inglés" representado por el cuadrado de color azul, si seguimos la lógica de los idiomas más hablados en Occidente. Vale aclarar que todos los eventos fueron originados en un solo país. Lo curioso es que los lenguajes que se utilizan en los dispositivos son muchos, es por ello que en un primer momento se había pensado que el país de los datos era Estados Unidos, ya que allí hay mucho turismo y también mucha gente viviendo que habla diferentes idiomas, pero se supo que el país era Uruguay. Resulta muy raro entonces, que en un país tan pequeño haya celulares con tantos idiomas distintos.

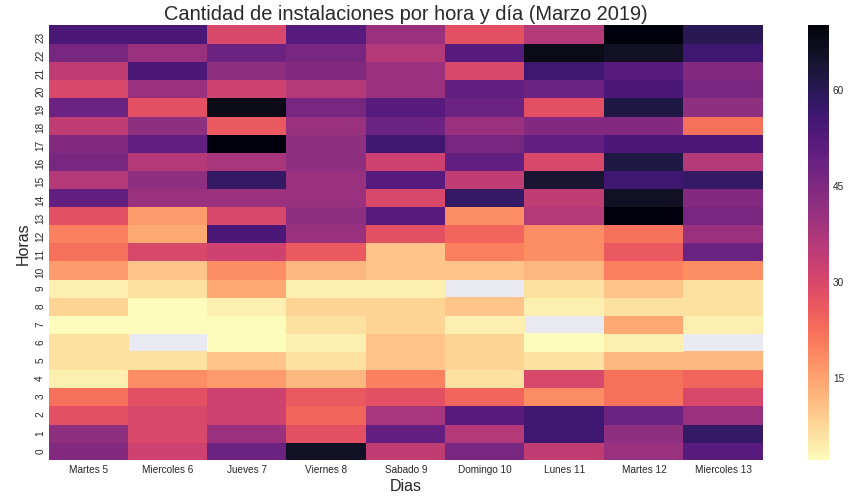


Se puede analizar que hay una ciudad en la que se originan casi todos los eventos. Es muy probable que esta sea la capital del país de los datos. Sabiendo que dicho país es Uruguay, se podría decir que la ciudad es Montevideo y sería muy lógico ya que la población charrúa se concentra principalmente allí.



   
Se puede observar que de los dos ref-type que hay, uno genera mucísimos más eventos que el otro. Sabiendo que estos dos son Apple-Ifa (celulares iPhone) y Google-Advertising-Id (celulares con Android), se supone que el mayor corresponde a los dispositivos con Android ya que en el mundo hay más celulares con ese sistema operativo.

INSTALLS

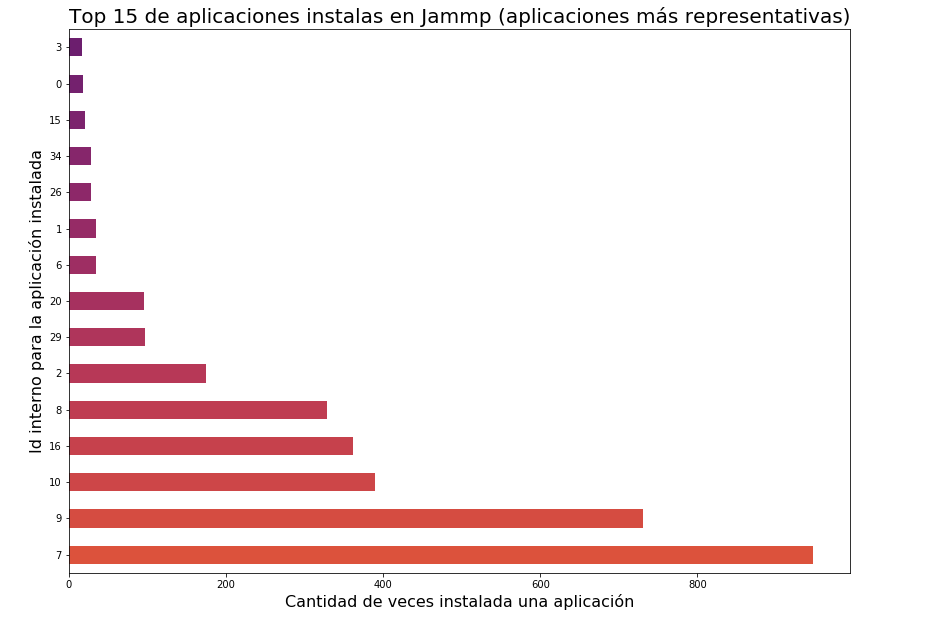


\*mejorar grafico de true false\*

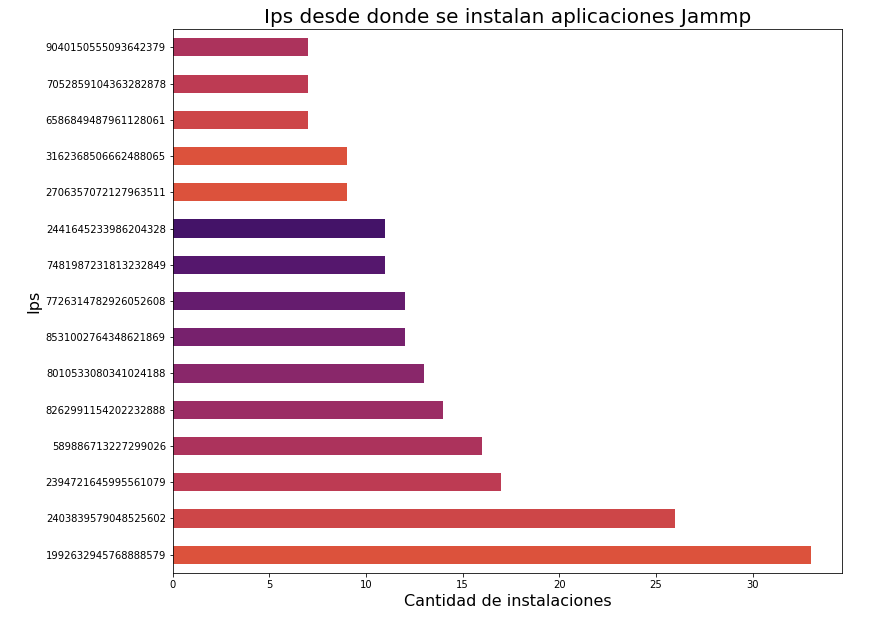
REPUESTA DE JAVIER SOBRE INSTALLS """En installs se almacenan instalaciones atribuidas o implícitas de las apps que son clientes de Jampp. En las subastas RTB puede venir un espacio de publicidad desde cualquier app, no solo las que son clientes de Jampp (a estas apps se las llama publishers). Estas apps en general ofrecen su inventario (espacios de publicidad) a un ad exchange (lo que nosotros llamamos sources) y desde ahí se hacen las subastas en las que participa Jampp. Cuando se instala desde un ad mostrado por Jampp el exchange es notificado y ellos nos avisan. Si luego vemos un nuevo device en la data orgánica, podemos inferir la instalación por otro canal. Una simplificación bastante correcta de la respuesta a esa pregunta sería que un DSP (jampp) muestre publicidad y por como es el mecanismo de atribucion no se le atribuye si hay instalación. Entonces, la instalación no es "atribuida" pero siendo que tenemos el callback de la instalación, sabemos que pasó, como no lo vimos desde eventos posteriores, no es "implícita"."""



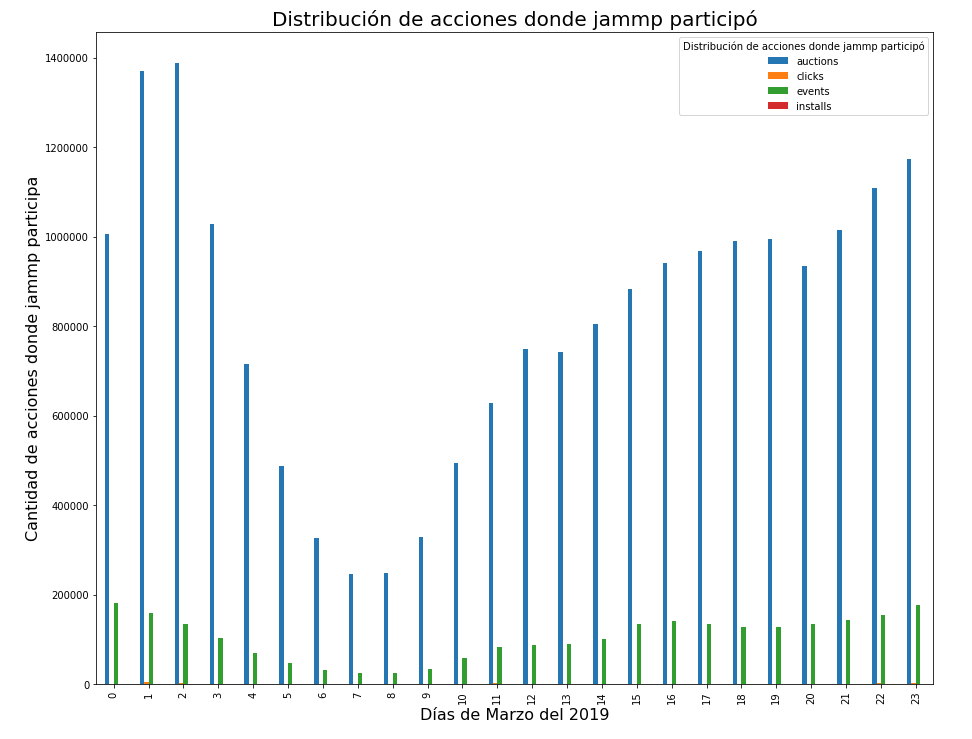
***Como se puede observar, las instalaciones realizadas en Jammps se centran en muchas ocurrencias que estan vinculadas con OPEN ( ya sea desde eventos para abrir la app) y LOGUEOS en la página.***



***Notamos que las instalaciones llevadas a mediante Jammp se centran en 6 aplicaciones principalmente, teniendo como pico la instalación de la aplicación "7" con más de 900 instalaciones.***

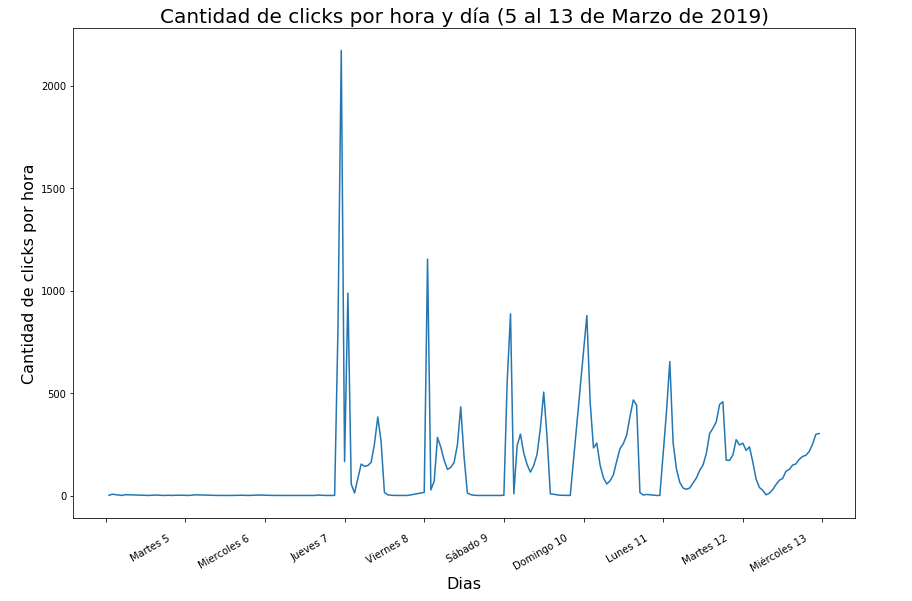


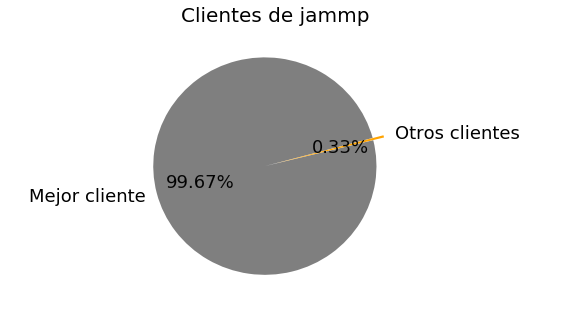
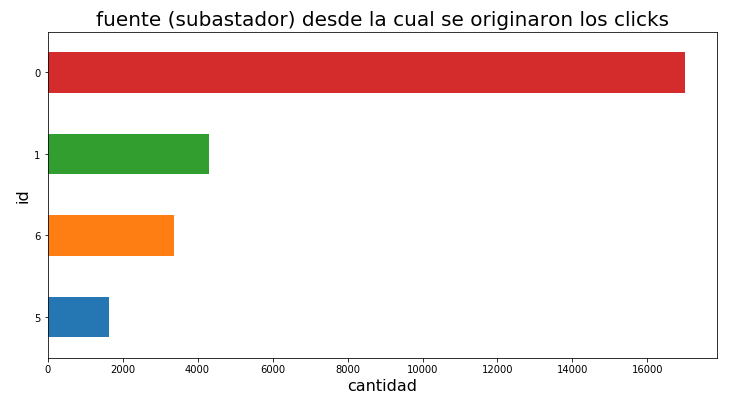
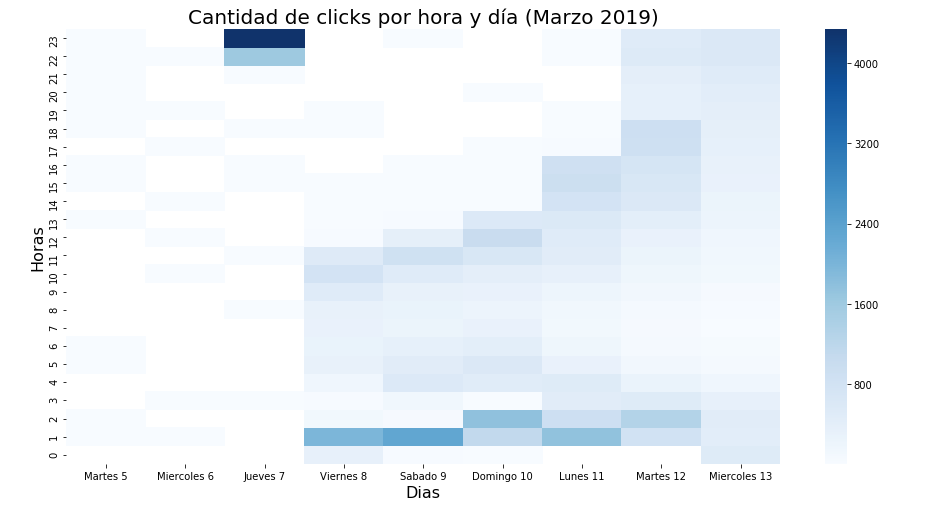
##### *Notamos que hay algunas ips que realizan muchas instalaciones de Jammp, suponemos que estas ips son de dominio público, por eso se realizan demasiadas conversiones desde una misma "localización"*

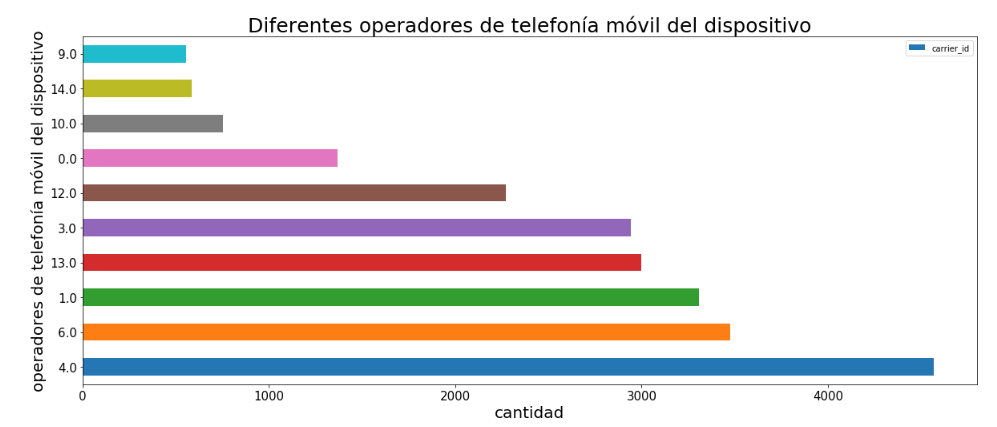


##### *Esté gráfico muestra la gran cantidad de subastas en las cuales puede participar jammps, contrastando mucho con la cantidad de eventos que finalmente se realizan sobre jammps.*

##### *Algo a destacar es la poca conversion (installs) que hay sobre Jammps, como a su vez el valle d e*

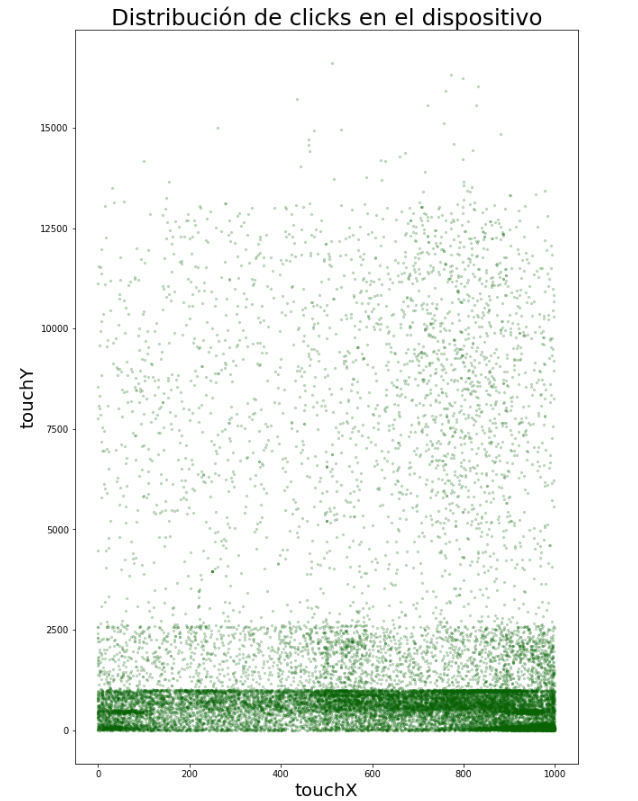
CLICKS



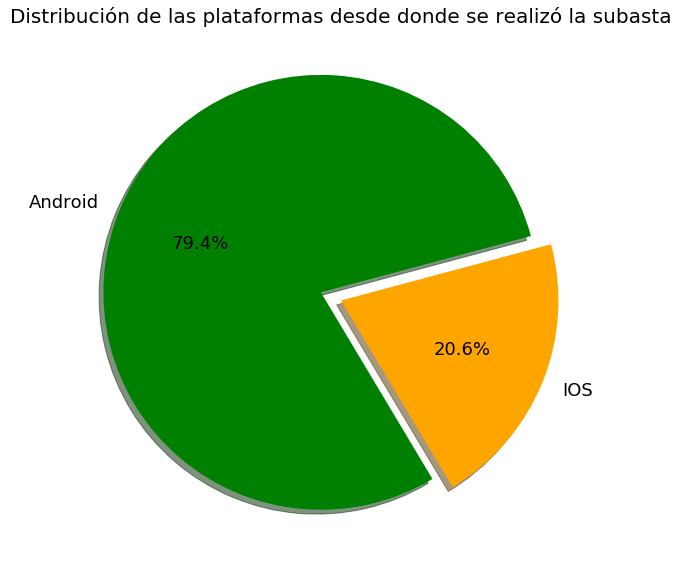


*#Podemos ver q todos los clicks registrados no fueron con conexion wifi*

*#datos registrados de un solo pais*



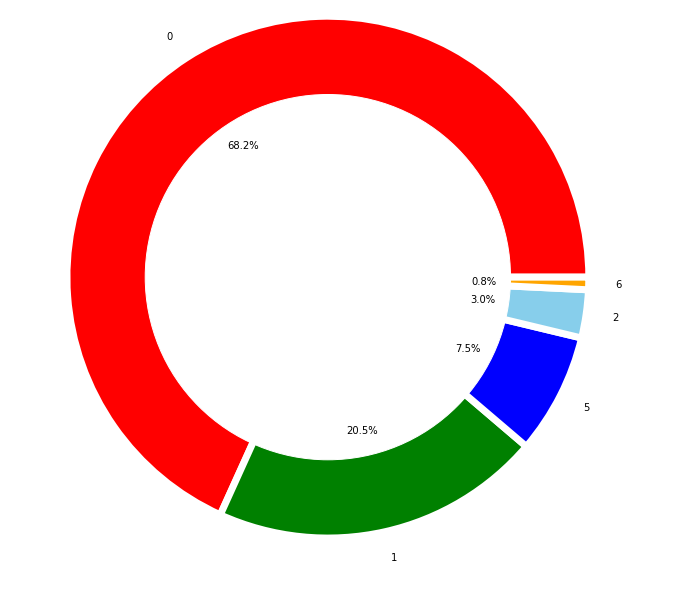
AUCTIONS



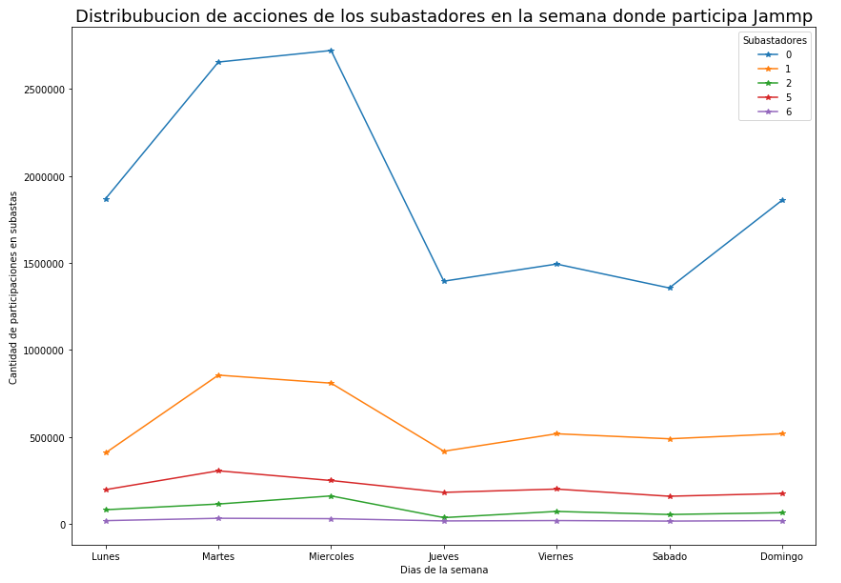
##### *Se puede observar que los uruguayos tienen en forma mayoritaria dispositivos android, desde donde salen las subastas donde participa Jampp*

\*titulo del grafico\*

Distribucion de fuentes de intercambio de las subastas



**En este gráfico podemos notar que con los datos provistos, hay una cierta preferencia (ocupando un 68.2% de los intercambios) por un motor de subastas, donde seguramente Jammp hace su subasta para poder ganarse un espacio.**



***Notamos que los subastadores registran una alza en sus acciones durante los primeros días de la semana, por lo cual, Jammp debe participar de manera muy activa para ganarse ese espacio en disputa, no así durante los siguiente días dónde el flujo de acciones de estos subastadores merma con respecto a los demás días.***

