

UNIVERSIDAD DE BUENOS AIRES

FACULTAD DE INGENIERÍA

**75.06 Organización de Datos Trabajo Práctico 1**

**Primer Cuatrimestre de 2019**

Grupo 18

|  |  |
| --- | --- |
| Camussi, Alan |  |
| Catolino, Lucas | 95560 |
| Zugna, Federico | 95758 |

**Link de GitHub:** <https://github.com/LucasCatolino/Orgnizacion-de-Datos> **(Privado)**

# Índice

# Introducción

En este trabajo se nos pidió analizar un set de datos provisto por la empresa Jammp, plataforma enfocada en la promoción y remarketing de aplicaciones móviles. La empresa busca promover apps, lograr que quien instaló la app efectúe una primera acción, y quien la usa que la siga usando.

Desde la empresa pusieron a nuestra disposición cuatro archivos .csv comprimidos, y un archivo .json explicando los datos. Estos eran:

* auctions.csv.gzip
* clicks.csv.gzip
* events.csv.gzip
* installs.csv.gzip
* desc.json

La finalidad del trabajo es realizar un análisis exploratorio de los datos, familiarizarnos con las herramientas vistas en clase, investigar sobre la materia para así poder visualizar resultados y analizar conclusiones que podamos encontrar.

# Análisis de clicks

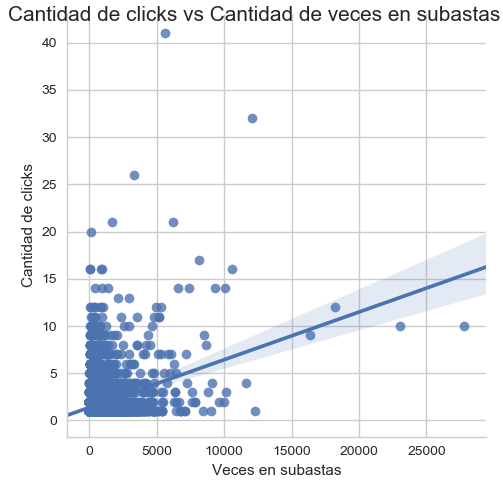
## Análisis de horarios

Depurando un poco los datos, puede verse que los datos correspondientes a los primeros días (es decir, el 5 y 6 de marzo), presentan muy pocos valores. Esto puede hacer que las conclusiones no sean representativas, por lo cual no se los tomó en cuenta. Por lo tanto, graficando la cantidad de clicks por día y hora, podemos ver los siguientes gráficos:

|  |
| --- |
|  |
|  |
|  |
|  |
|  |
|  |
|  |

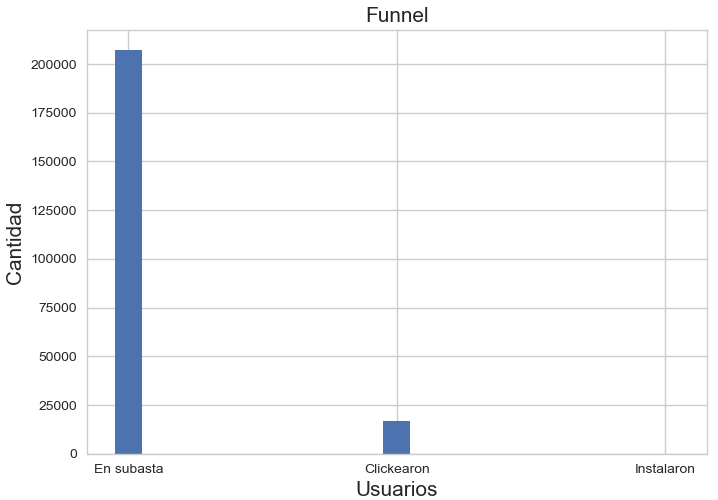
## Cantidad de clicks vs cantidad de veces en subasta

De los datos provenientes de los sets de subastas y clicks, se agrupó por id de usuario y luego se hizo un inner join para obtener la relación entre la cantidad de veces que un usuario aparece en una subasta, y la cantidad de veces que ese mismo usuario hizo click en la publicidad, obteniendo el siguiente gráfico:



## Funnel de conversión

Tomando como objetivo la instalación de una app, se tomaron los siguientes pasos para la conversión de un usuario: el usuario aparece en subasta, el usuario clickea una publicidad, el usuario instala la aplicación. Del total de los usuarios poco más de 200.000 usuarios que entraron en subasta en los días analizados, se encontró la siguiente relación:



## Conclusión

De los gráficos puede verse que en general los picos de actividad suelen darse al final y al inicio del día. Es decir, estamos hablando de usuarios en con fuerte actividad noctámbula. También puede verse, como era de esperar, muy baja actividad en los horarios de madrugada.

# Análisis de clicks