**Aplicaciones móviles**

**Segmentación de usuarius en función del comportamiento**

**Objetivo**

Tomar las acciones realizadas en la aplicación para definir grupos que difieren en términos de métricas del producto:

* tasa de retención,
* tiempo dedicado a la aplicación,
* frecuencia de ocurrencia de eventos,
* conversión en el evento de destino, contacts\_show

**Descripción de los datos**

Los datasets contienen datos sobre los eventos que ocurrieron en la aplicación móvil Trash to Treasure, en la que los usuarios y las usuarias publican anuncios para vender cosas que ya no necesitan.

Los datos corresponden a quienes realizaron sus primeras acciones en la aplicación del 7 de octubre de 2019 al 03 de noviembre de 2019, o a favoritos

**mobile\_dataset\_us.csv** contiene las siguientes columnas:

* event.time: cuando tuvo lugar el evento,
* event.name
* user.id

Detalles del evento:

* advert\_open: abrir una publicación publicitaria,
* photos\_show: ver fotos en el anuncio,
* tips\_show: se mostraron anuncios recomendados a usuari@,
* tips\_click: Usuari@ hizo clic en un anuncio recomendado,
* contacts\_show y show\_contacts: Usuari@ hizo clic en el botón "mostrar número de teléfono" en el anuncio,
* contacts\_call: Usuari@ marcó el número del anuncio
* map: Usuari@ abrió el mapa de los anuncios publicados
* search\_1 - search\_7: varios eventos relacionados con la búsqueda en el sitio web
* favorites\_add: el usuario o la usuaria agregó el anuncio a favoritos

**mobile\_sources\_us.csv** contiene las siguientes columnas:

* userId: ID de usuario/a
* source: la fuente desde la cual el usuario o la usuaria descargó la aplicación

**Plan de Trabajo**

**Preprocesamiento**

* + Cambiar nombre de columnas,
  + Complementar/combinar "dataset" con la información de "sources"
  + juntar (en otra columna) los eventos de "search"

**Análisis exploratorio a mobile\_dataset\_us**

* + replantear objetivo: tasa de retención o tasa de éxito → embudo → conversión en el evento de destino, contacts\_show o contact\_calls.
  + tiempo dedicado a la aplicación por usuario, en promedio considerando totales (eventos y usuarios)
  + frecuencia de ocurrencia de eventos
  + cantidad de eventos realizados por cada usuario
  + cantidad de eventos realizados en general

**Análisis exploratorio a mobile\_sources\_us**

* + Cantidad de usuarios únicos por fuente de adquisición,
  + Graficar cantidad de usuarios únicos por fuente de adquisición.

**Análisis exploratorio a dataset combinado**

* + Cantidad de usuarios únicos a nivel general y desglosados por fuente de adquisición,
  + totales por evento,
  + Agrupación de eventos por fuente de adquisición.
  + Histograma de eventos promedio por usuario (desglosado por fuente de adquisición)
  + Cantidad de eventos a través del tiempo (agrupando diariamente/semanalmente, desglosado por fuente de adquisición)
  + Para los eventos más populares, cantidad de eventos a través del tiempo
  + adquisión de usuarios
  + eventos más populares (gráfico de barras con conteo de eventos, desglosado por fuente de adquisición)

**Segmentar a los usuarios**

* Construir métricas de interacción (engagement) (e.g., % de eventos en fin de semana, % de eventos en la mañana/tarde/noche, cantidad de eventos, número de días distintos que un usuario realizó un evento, número de eventos por tipo de evento)
* graficar un scatter matrix para identificar visualmente si existe una separación por grupos
* usar un algoritmo de clusterización (K means, KNN, t-SNE, UMAP, etc.) para crear segmentos

**Pruebas de hipótesis**

* prueba de hipótesis de la tarea
* conversión a otros eventos
* conversión entre segmentos creados en la sección anterior
* curvas de retención semanales con diferentes eventos (o con todos)

**Entrega del proyecto**

<https://www.notion.so/Entrega-del-proyecto-a4bdeb206bdd42c485d67ee88d68d70a?pvs=21>

Tu proyecto constará de tres componentes:

* Un notebook con tu código (.ipynb)
* Una presentación (.pdf)
* Un link a tu dashboard en Tableau Public (dashboard.txt) (opcional)
* Asegúrate de que la solución esté completa y comprueba que el código funciona en Jupyter antes de enviarlo para su revisión. No te sorprendas si te piden que hagas algunas mejoras.
* Carga el archivo en Google Drive y copia el enlace para compartir al principio de tu notebook.
* Si quieres crear un dashboard más tarde, informa al líder de tu equipo cuando entregues tu proyecto por primera vez.
* El proyecto será aceptado solo después de que envíes tu dashboard.

Haz una lista de 5 a 10 fuentes (documentación, artículos, etc) que usaste mientras trabajabas en el proyecto. Brinda explicaciones breves sobre las preguntas que cada fuente te ayudó a responder. Prepara una presentación con tus conclusiones. Puedes usar cualquier herramienta que quieras para hacer la presentación, pero debes enviarla en PDF. Coloca el enlace en el cuerpo del proyecto principal.

**Dashboards**

Opción 1

1. Traza un diagrama que muestre la distribución de eventos por tipo de evento

2. Agrega un indicador para el número de usuarios y usuarias

3. Agrega un filtro de fecha de evento

Opción 2

1. Traza un diagrama que muestre el número de eventos por día

2. Traza un histograma que muestre el número de usuarios o y usuarias que provienen de cada fuente

3. Agrega un filtro de tipo de evento