



ASSIGNMENT 1: LOGS - WhatsApp

Alejandro Polo

Lucía Poblador

El objetivo de la práctica es construir un notebook en Python capaz de analizar chats de Whatsapp.

En este caso, el primero de los obstáculos fue la lectura del texto que estaba en formato txt. Con el texto leído como un string, el siguiente paso es estructurarlo en un dataframe que facilite su análisis posterior. Este dataframe conserva para cada uno de los mensajes: autor, fecha y body (o cuerpo del mensaje).

A nivel particular, nos propusimos como objetivo tratar de construir el 100% del código desde cero. Esto nos ha permitido añadir funcionalidades extras y practicar en el uso de librerías como pandas. Algunos de nuestros propósitos fueron:

- Analizar cómo han evolucionado la mención de las asignaturas por meses
- Visualizar las urls y dominios web más comunes
- Los principales emojis que han sido utilizados o los días de la semana en los que más actividad ha sido registrada
- Extraer el top 10 de usuarios que más han interactuado en el chat (y también en qué meses están más activos)
- Analizar las palabras más mencionadas incluyendo un stop words en español.

Para todo ello nos hemos valido de la estructura temporal de los datos que han sido representados como una serie temporal para estudiar los picos de actividad y, a partir de ahí, se han realizado agregados con distintos propósitos. Por otro lado, para conseguir deshacernos de las palabras más comunes de la lengua española – y que carecen de relevancia - se ha construido un diccionario con todas esas palabras.