

PRUEBA TÉCNICA

DataCompany

...

PABLO LOPEZOSA DEL PINO

Obtención de tweets

- Mediante webscraping .
- Librería snsrape.
- Obtención de 139 tweets desde el 01-01-22.
- Columnas : `Tweet_ID`, `Tweet_Content`, `Date`, `User_ID`, `User`, `Likes`, `Retweets`, `Replies`, `quoteCount`

Tablas SQL

- Disponibles en el repositorio de GitHub https://github.com/PabloLopezosa/PRUEBA_TECNICA_TB.git

EDA

- Tweet con mayor repercusión: `Tweet_ID: 1513848281801056258`
- Nueva columna de repercusión: $\text{likes} + \text{Retweets} * 2 + \text{coments} * 3 + \text{quote}$
- El usuario que más menciona @TheBridge_Tech es la propia escuela con 57 tw.
- Después @MujeresaSeguir con 7
- El mes que más concentración tiene es Abril con 49

MODELO

- Nos dice si el tweet contiene un comentario positivo o negativo.
- Es un modelo muy sencillo. En este caso mejoraría haciendo una mejor limpieza de los tweets y mejorar los scores del modelo.
- También se podría buscar otro modelo diferente y comparar resultados.
- Otras oportunidades:
 - Análisis de sentimiento para temas electorales.
 - Intenciones de voto.
 - Ver si una campaña publicitaria en concreto está teniendo impacto positivo o negativo.

