

## Ejercicio 2

Tienes un archivo CSV llamado tweets.csv con varios miles de publicaciones en redes sociales.

El archivo contiene al menos las siguientes columnas:

- Id: identificador del tweet
- User: nombre del usuario
- Text: contenido del tweet
- Date: fecha de publicación

Qué debes hacer:

1. Cargar el archivo tweets.csv en un DataFrame de Spark.
2. Crear una vista temporal llamada tweets.
3. Usar Spark SQL para responder:
  - a) ¿Cuántos tweets hay en total?
  - b) ¿Cuál es el promedio de palabras por tweet?
  - c) ¿Cuáles son los 5 usuarios con mayor cantidad de tweets publicados?
4. Generar un gráfico de barras en Python/Matplotlib que muestre los 5 usuarios más activos con su número de tweets.