# Twitter Avro Data

**An avro dataset is provided with 3 columns**

**Username, tweet and the timestamp.**

Solve the following on the Avro dataset

1. Identify the top 5 users with the highest number of tweets
2. Most commonly used words in all the tweets, excluding stopwords
3. Filter the tweets that are in Spanish
4. Which country has the highest number of tweets?
5. Separate the tweets by a specific username ex: All tweets by Mark