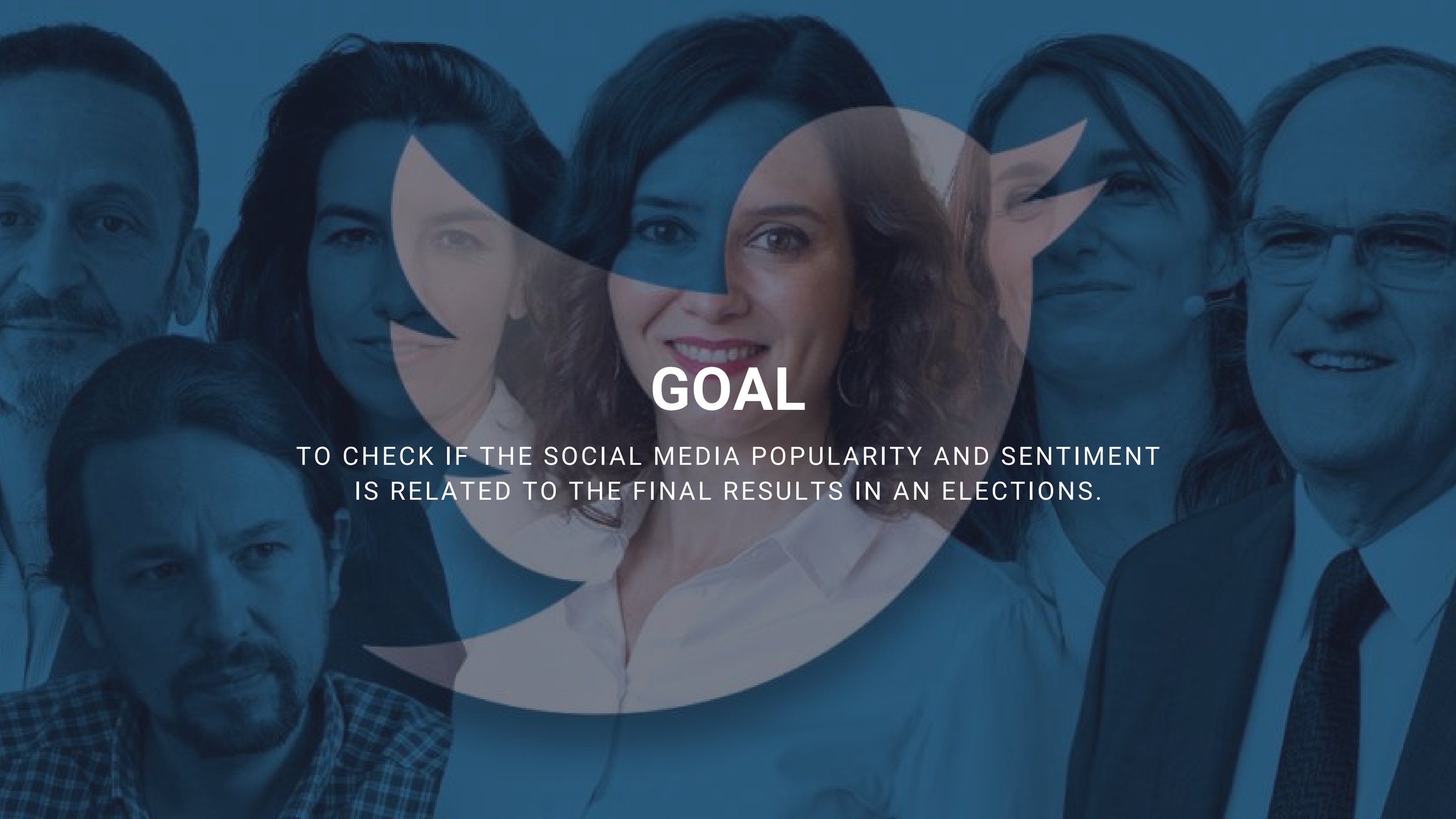
SENTIMIENTO POLÍTICO

DATA PROJECT USING TWEEPY API
AND NLP MACHINE LEARNING



GOAL

TO CHECK IF THE SOCIAL MEDIA POPULARITY AND SENTIMENT IS RELATED TO THE FINAL RESULTS IN AN ELECTIONS.


HAZLO MADRID


tve

PROCESS

CONNECT TO THE API

Ask for a Twitter developer account and get the keys.

EXTRACT TWEETS AND MENTIONS

Extract tweets of different spanish political parties and their representatives, and also, extract user mentions of each party.

APPLY NLP FOR SENTIMENT

To know the sentiment, we will apply NLP machine learning for each tweet.

Connect to the API

1

KEYS

Obtaining the developer account and the keys.

2

REST TIME

The API needs to rest some minutes (15 aprox.) for each 1.000 mentions.

3

WAIT

If the wait time is indicated as False, Twitter could ban out account. Solution = wait.

Extract tweets

1

HOW

Using API.user_timeline function of Tweepy.

2

LAST TWEET

Reading the current data base and obtaining the max id = last id.

3

FULL TWEET

Use tweet_mode='extended to obtain the full tweet.

Extract mentions

1

HOW

Using API.search function of Tweepy and giving it a query (examples).

2

AMOUNT OF DATA

The data frame breaks into 50 parts and apply machine learning for each part and add everything to database.

3

DATES

The amount of tweets and dates can be chosen inside the function (max 7 days).

Examples of queries

PP

'@populares OR "partido popular" OR "el pp" OR "los del pp" OR "Ayuso" OR "Isabel Díaz Ayuso" OR "Pablo Casado" -filter:retweets'

PSOE

'@psoe OR "partido socialista" OR "Partido Socialista Obrero Español" ' 'OR "el psoe" OR "los del psoe" OR "Pedro Sánchez" OR "Ángel Gabilondo" -filter:retweets'

PODEMOS

'@PODEMOS OR "podemos" OR "los de podemos" OR "podemitas" OR "el de podemos" 'OR "el coletas" OR "Pablo Iglesias" OR "el iglesias -filter:retweets'

VOX

'@vox_es OR "vox" OR "los de vox" OR "el de VOX" OR "la de VOX" ' 'OR "Santiago Abascal" OR "Rocío Monasterio" -filter:retweets'

Apply NLP

1

NLP MODEL

bert-base-multilingual-uncased-sentiment allows to have a classification of 1-5 starts of a text.

2

TWEETS

The tweets are extracted and classify in one process.

3

MENTIONS

First extract the mentions, export them to a invidual database and then apply NLP to append to main database.

Tableau Dashboard

1

FILTER BY PARTY

The dasboard shows the results by party, so a column is added with the name of each party.

2

CAMPAING

If the tweet contains "Madrid", it will belong to Madrid campaing, if not it will be General campaign.

Tableau Dashboard

3

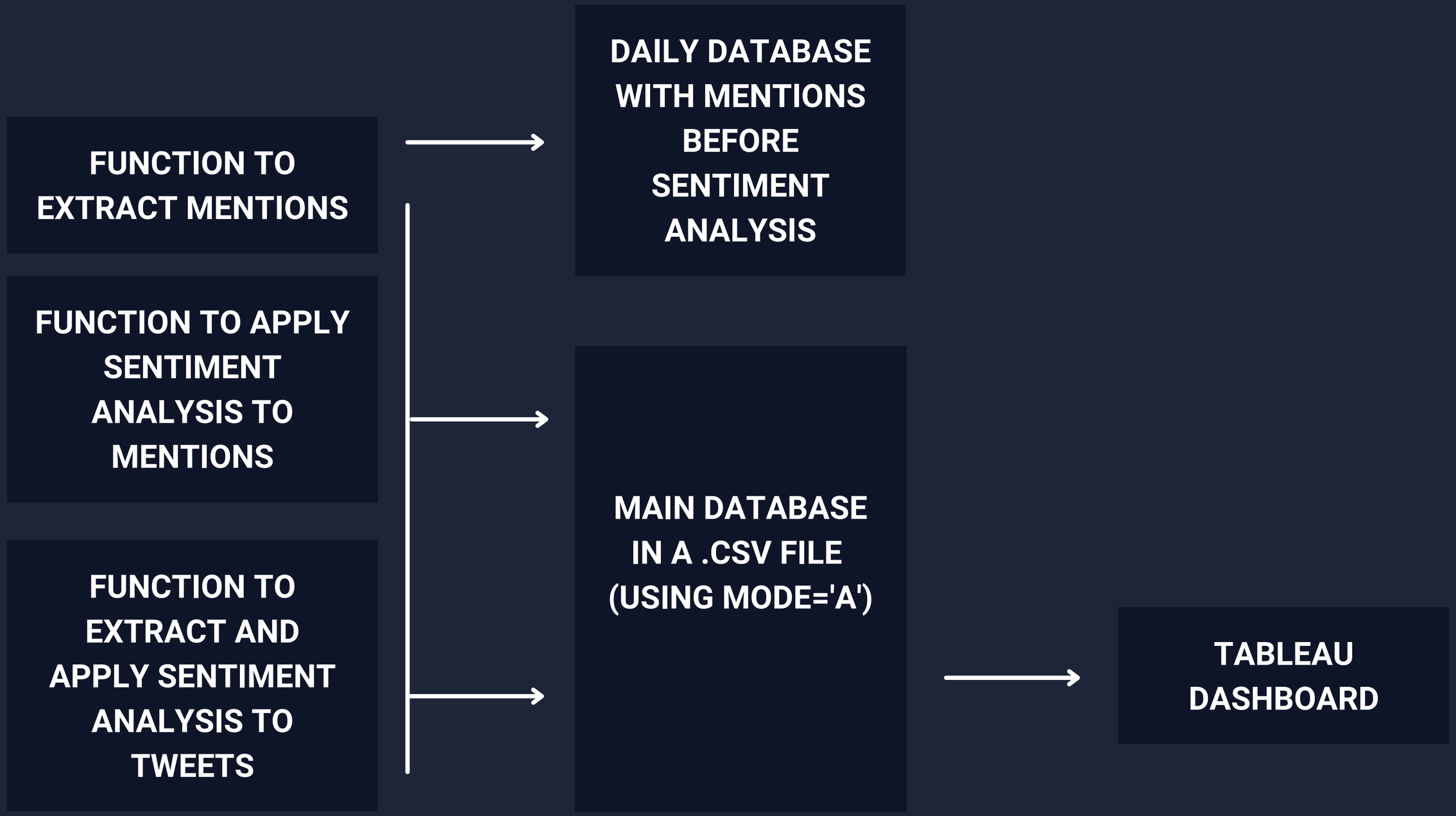
TYPE OF POST

Column that indicates if is a tweet (publicación) or a mention (mención).

4

NOT REWRITE

Surviving in the time is an important point of this project so I needed a way to update the data base every day. The solution has been export the final .csv with the last tweets and mentions in mode='a' and without header.


FUNCTION TO EXTRACT MENTIONS
FUNCTION TO APPLY SENTIMENT ANALYSIS TO MENTIONS
FUNCTION TO EXTRACT AND APPLY SENTIMENT ANALYSIS TO TWEETS
DAILY DATABASE WITH MENTIONS BEFORE SENTIMENT ANALYSIS
MAIN DATABASE IN A .CSV FILE (USING MODE='A')
TABLEAU DASHBOARD

FUNCTION TO EXTRACT MENTIONS

FUNCTION TO APPLY SENTIMENT ANALYSIS TO MENTIONS

FUNCTION TO EXTRACT AND APPLY SENTIMENT ANALYSIS TO TWEETS

DAILY DATABASE WITH MENTIONS BEFORE SENTIMENT ANALYSIS

MAIN DATABASE IN A .CSV FILE (USING MODE='A')

TABLEAU DASHBOARD

Next Steps

- Automate the process.
- Web using Streamlit.
- Add future political campaigns.

SENTIMIENTO POLÍTICO

¿Qué partido político español es más popular y mejor valorado en Twitter?

Número de tweets analizados: 223.982

tipo de post: (All)

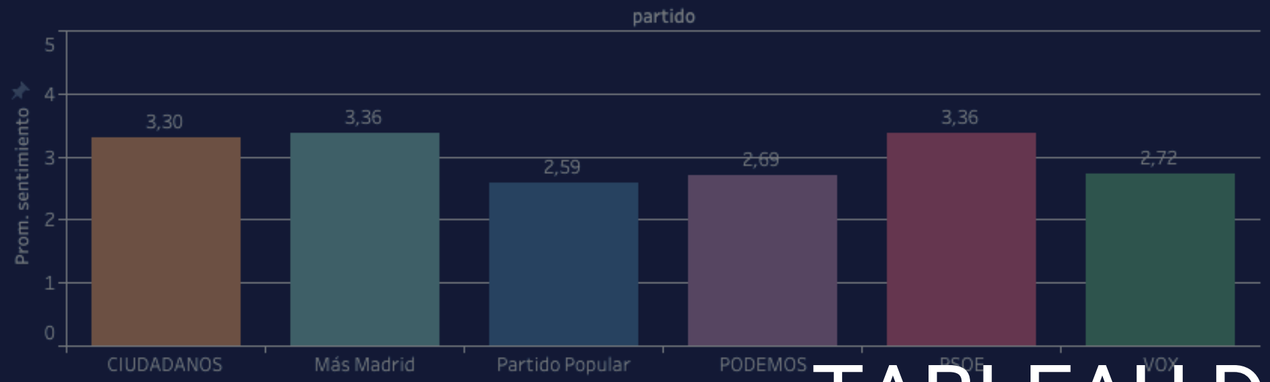
partido: (Multiple values)

campaña: (All)

Representantes y partidos: (All)

CIUDADANOS | Más Madrid | Partido Popular | PODEMOS | PSOE | VOX

Sentimiento medio por partido



Sentimiento medio diario en Twitter

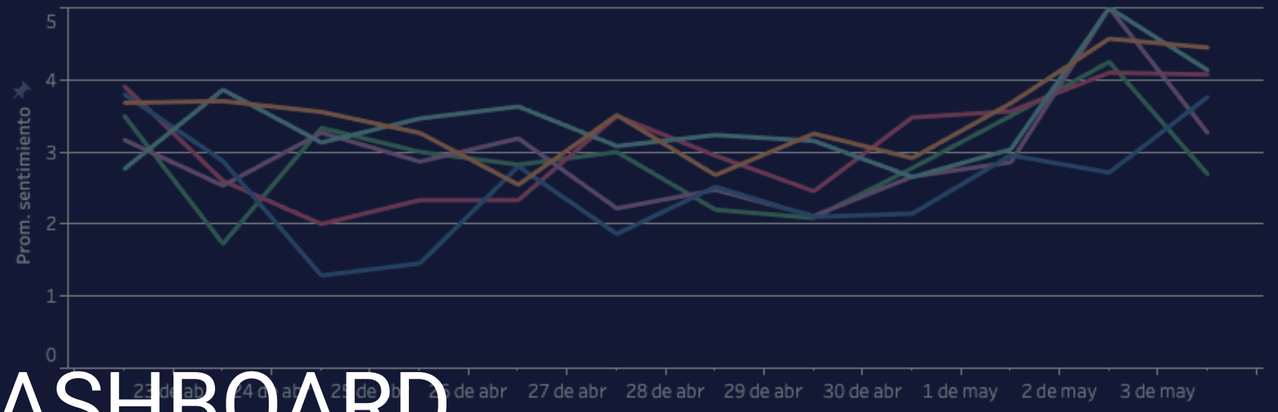
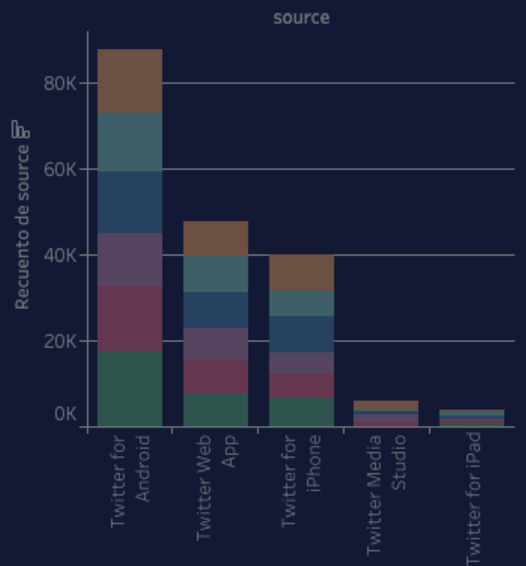


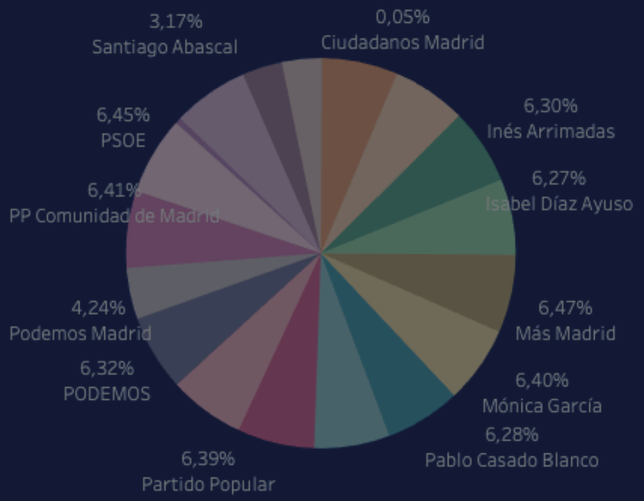
TABLEAU DASHBOARD

Fuente de tweets



Porcentaje menciones/publicaciones por partido

Porcentaje de menciones por cuenta



Número total de interacciones