MINE SOCIAL MEDIA SENTIMENT

SENTIMENT ANALYSIS

- The key aspect of sentiment analysis is to analyze a body of text for understanding the opinion expressed by it. Typically, we quantify this sentiment with a positive or negative value, called polarity. The overall sentiment is often inferred as positive, neutral or negative from the sign of the polarity score.

- Sentiment analysis is widely used, especially as a part of social media analysis for any domain, be it a business, a recent movie, or a product launch, to understand its reception by the people and what they think of it based on their opinions or, you guessed it, sentiment!

Tweepy

An easy-to-use Python library for accessing the Twitter API.

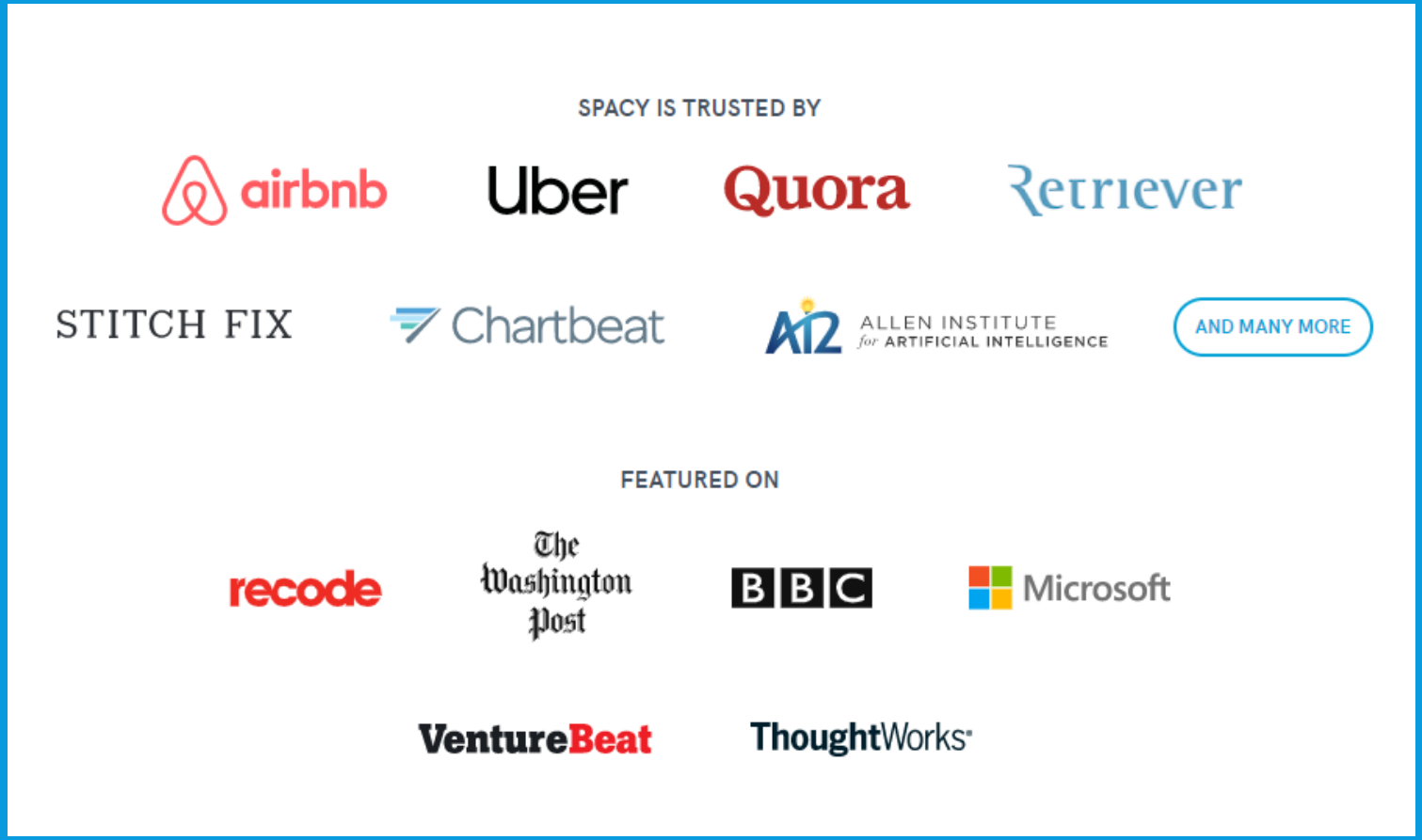
- Tweepy is an open source Python package that gives you a very convenient way to access the Twitter API with Python. Tweepy includes a set of classes and methods that represent Twitter's models and API endpoints, and it transparently handles various implementation details, such as:

- Data encoding and decoding
- HTTP requests
- Results pagination
- OAuth authentication
- Rate limits
- Streams

spaCy

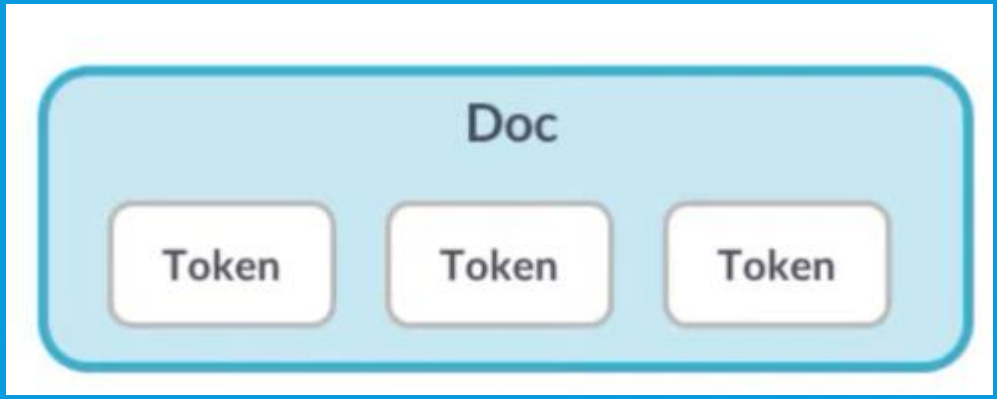
spaCy is a free, open-source library for advanced Natural Language Processing (NLP) in Python.

spaCy is designed specifically for production use and helps you build applications that process and “understand” large volumes of text. It can be used to build information extraction or natural language understanding systems, or to pre-process text for deep learning.



spaCy

- Spacy splits the input text into tokens which it can recognize
- Verbs
- Nouns



Tweepy

An easy-to-use Python library for accessing the Twitter API.

```python
import tweepy
from tweepy import OAuthHandler
from config import consumer_key, consumer_secret, access_token, access_secret
auth = OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_secret)

api = tweepy.API(auth)

for status in tweepy.Cursor(api.home_timeline).items(10):
    # Process a single status
    print(status.text)
```

```
También integra apple carplay y android auto, mecánicamente parece no haber sufrido modificaciones en cuanto al mot… https://t.co/qwhLUstub5
- Alerta de tráfico cruzado trasero
- Advertencia de punto ciego
- Advertencia de cambio de carril
- Asistente de… https://t.co/1Uwyxn5MXb
Nuevo #Nissan #Kicks 2021 🆕

El fabricante japonés presentó el rediseño del Kicks, el exterior e interior fueron re… https://t.co/1G4GQ6Pxne
De hecho no es buena relación valor precio https://t.co/mnAdp8tWff
Los integrantes acordaron seguir fortaleciendo la relación entre los #países y enfoncar sus esfuerzos hacia una… https://t.co/LL6t2kQfTG
Se revisaron los avances alcanzados por este bloque de países en materia de #integración económica y financiera, fi… https://t.co/3XeVjZFdyQ
Hoy, se llevó a cabo la XXIII Reunión de Ministros de Finanzas de la @A_delPacifico, en la que participó el secreta… https://t.co/FvdbHoMJ4K
📺Con una inversión de mil 100 MDD, la #TermoeléctricaHuexca funcionará con dos grandes insumos: aguas negras tratad… https://t.co/974nlYOukb
RT @cromero_aranda: Durante la conferencia del presidente @lopezobrador_, hoy por la mañana, firmamos un acuerdo entre sector #obrero, #emp…
RT @cromero_aranda: Todos coincidimos en que es necesario trabajar en conjunto para acabar con prácticas ilegales y #garantizar los #derech…
```

EXTRACTING TWEETS

- Here we use tweepy with Hashtag the following parameters:
- Hashtag = #Elections2020
- Number of tweets = 10,000
- Language = English
- Since = 2020 – 01 – 11
- Encoding = UTF8

```python
import tweepy
from tweepy import OAuthHandler
import csv
import pandas as pd
from config import consumer_key, consumer_secret, access_token, access_secret


auth = OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_secret)

api = tweepy.API(auth, wait_on_rate_limit=True)
user = api.me()

#Create a CSV
csvFile = open('data.csv', 'a')
csvWriter = csv.writer(csvFile)

for tweet in tweepy.Cursor(api.search,q="#Elections2020",count=10000,
                           lang="en",
                           since="2020-01-11").items():
    print (tweet.created_at, tweet.text)
    csvWriter.writerow([tweet.created_at, tweet.text.encode('utf-8')])
```

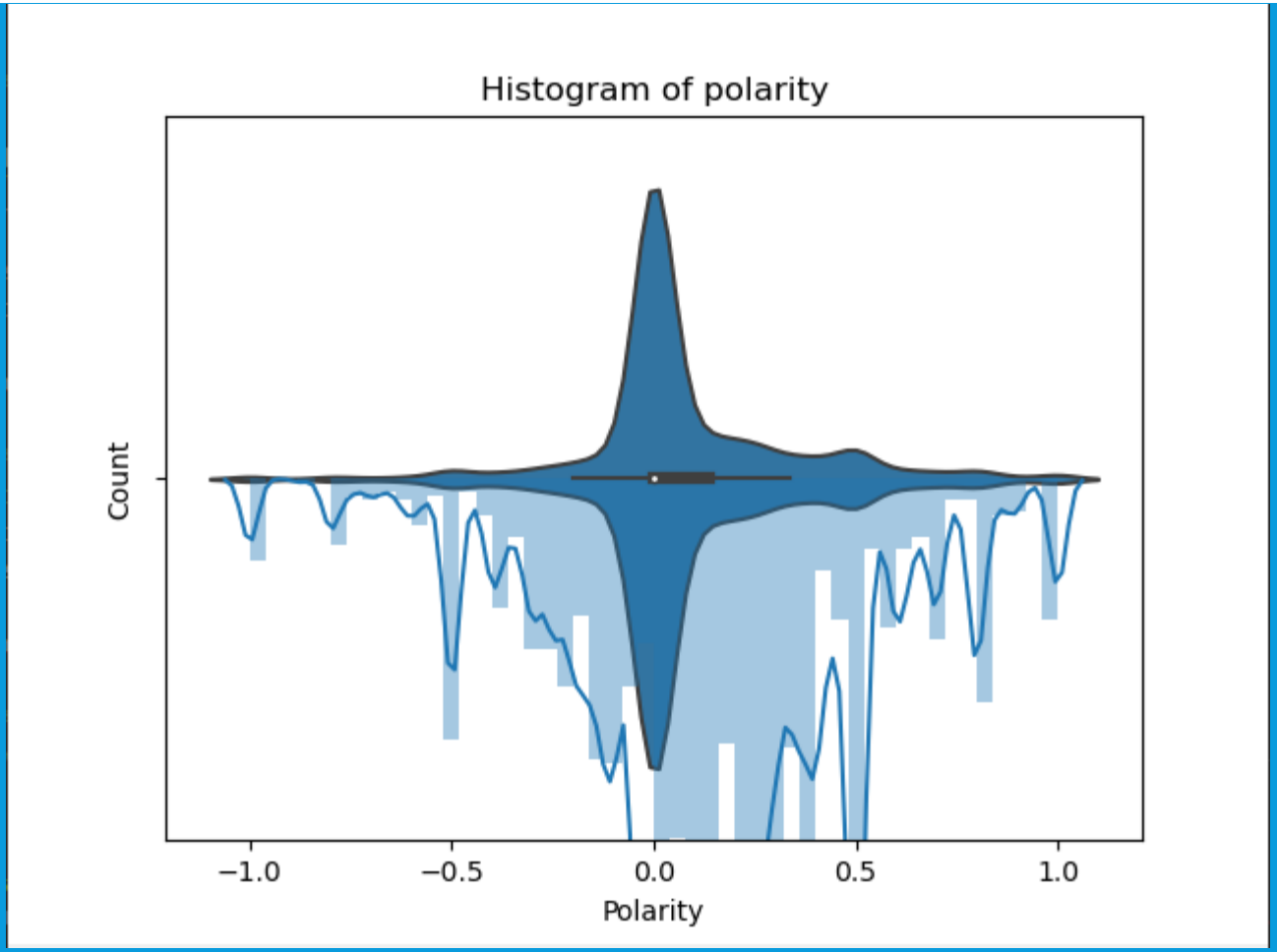
DATA OBTAINED

- Then the data obtained is processed resulting in:
- Total Tweets : 10,000
- Unique Tweets: 5,648
- Unique Non RT: 4,377

POLARITY

- Typically, we quantify this sentiment with a positive or negative value, called polarity. The overall sentiment is often inferred as positive, neutral or negative from the sign of the polarity score.

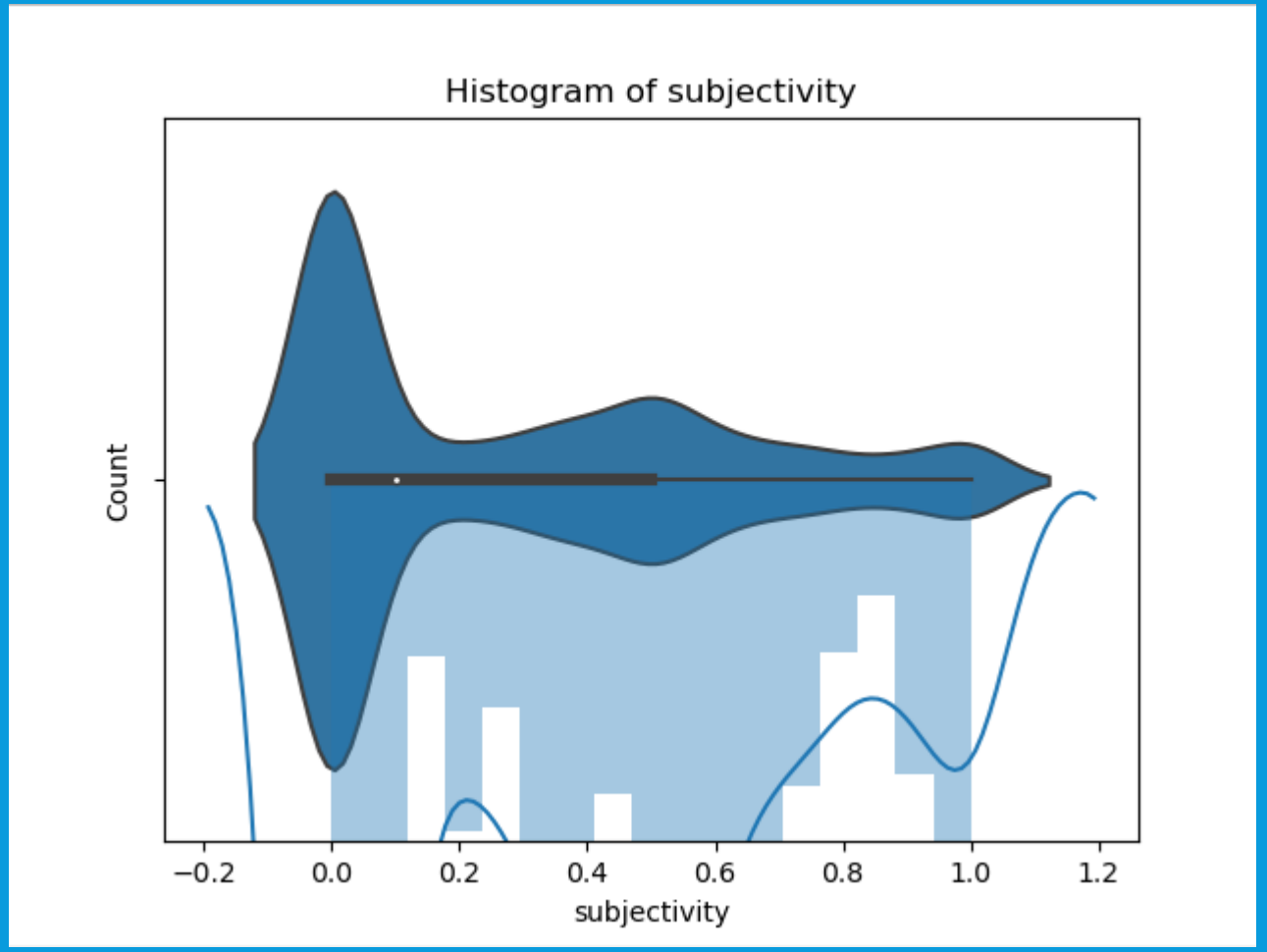
- Mean Polarity :
- `0.07`



SUBJECTIVITY

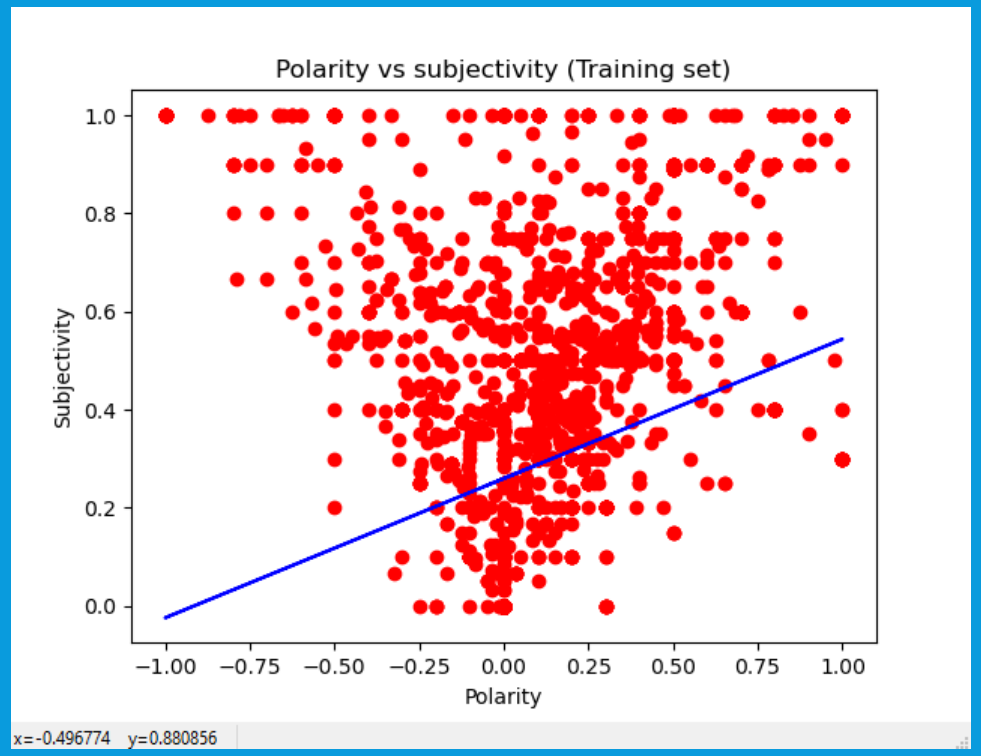
- Subjective text contains text that is usually expressed by a human having typical moods, emotions, and feelings.

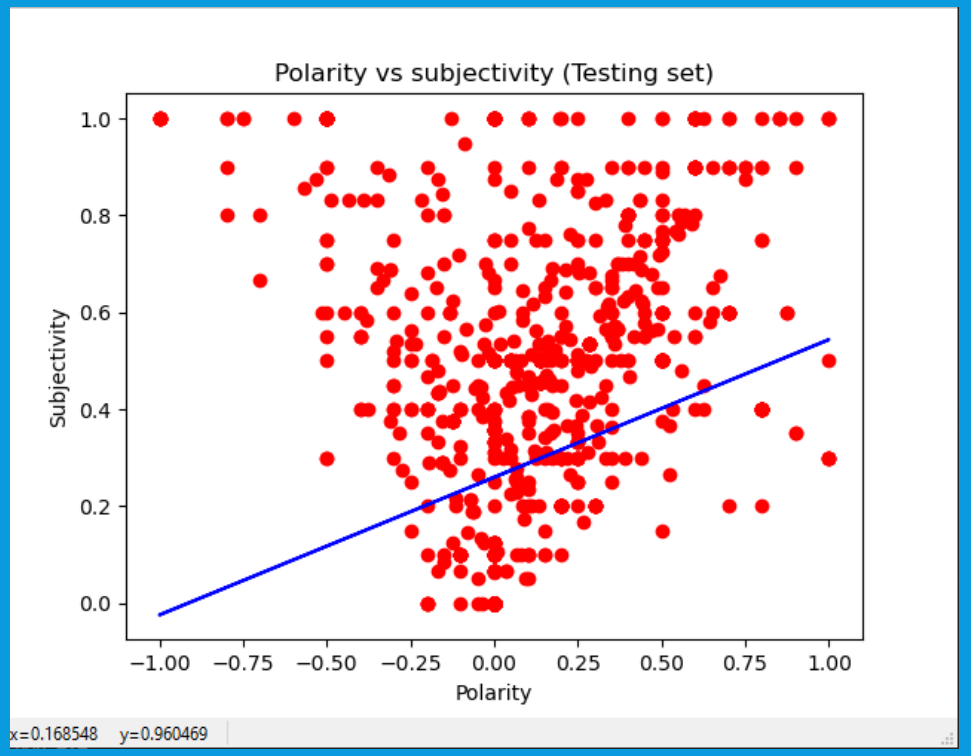
- Mean Subjectivity : 0.27



LINEAR REGRESSION

- Intercept: 0.25 , Coefficient: 0.28





Thank You!