



Analyse de réactions sur Twitter

Introduction



python™

API Twitter

5,2Go de tweets récoltés entre le 24 et le 28 avril sur #Endgame

Ceci correspond à 806208 tweets dont 671290 en Anglais, Français ou Espagnol.

Premiers résultats

On appelle les 10 mots les plus fréquents :

```
[('RT', 425140), ('#Endgame', 322002), ('the', 269224), ('a', 144375), ('to', 132929), ('of', 126296), ('', 108431), ('I', 107497), ('and', 94594), ('#EndGame', 88400)]
```

On remarque directement la présence de mots inutiles

Stop words et Bigrams

Sans les stop words : Des résultats plus intéressants mais qui manquent de contexte :

```
[(' ', 108431), ('I', 107497), ('Happy', 49987), ('Day!!', 48241), ('https://t.co/CFipXu5uUj', 48164), ("It's", 39216), ('day', 39147), ('someone', 35660), ('waiting', 34711), ('Did', 33702)]
```

Bigrams : Des résultats un peu plus clairs :

```
[((' ', ' '), 70575), (('Happy', 'Day!!'), 48225), (('Day!!', 'https://t.co/CFipXu5uUj'), 48164), (('Did', 'someone'), 33110),  
('someone', 'say...'), 33089), (('say...', 'Tomorrow?'), 33075), (('Tomorrow?', 'https://t.co/RQrNKj5Hyv'), 33069), ((("we've",  
'waiting'), 32833), (('ride', 'lifetime.'), 32828), (("Today's", 'day'), 32827)]
```

Co-Occurrences

Des résultats illisibles alors on fait des requêtes supplémentaires

```
In [19]: 1 sorted(com['cool'].items(), key=operator.itemgetter(1), reverse=True)[:5]
```

```
Out[19]: [('see', 94), ('really', 88), ('movie', 83), ('people', 71), ('super', 69)]
```

```
In [21]: 1 sorted(com['boring'].items(), key=operator.itemgetter(1), reverse=True)[:5]
```

```
Out[21]: [('one', 11), ('movie', 10), ('marvel', 7), ('time', 7), ('movie.', 7)]
```

```
In [22]: 1 sorted(com['spoilers'].items(), key=operator.itemgetter(1), reverse=True)[:5]
```

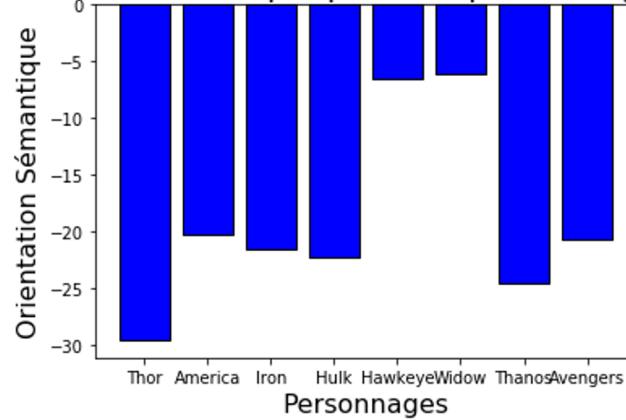
```
Out[22]: [('without', 10984),
```

Analyse sémantique

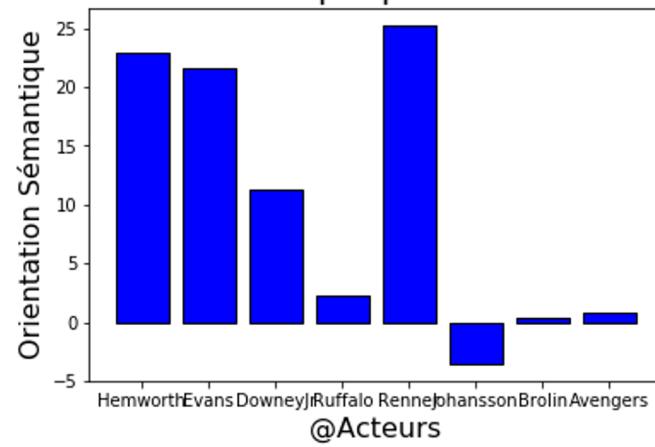
- Vérification et comptage du nombre de tweets
- Calcul de l'association des termes
- Calcul d'un « indicateur » d'orientation sémantique :
 - Associations avec des termes connotés positivement
 - Avec des termes négatifs
- Pourquoi « Happy » est si mal connoté ?

Conclusion

Orientation sémantique pour les personnages du film



Orientation sémantique pour les acteurs du film



- Limites :
 - Le problème d'Iron Man et de la « bonne tristesse »
 - Problème des « not-words »