

# Drake Lyrics Projects

What features could influence Genius.com view on Drake albums

Project Module 2

Orphee

<https://github.com/MaxCrmr/Ironhack-project-m1#title-of-my-project>

# Data Source n°1

## Drake lyrics



Format : This is a dataset consisting of Drake lyrics and other information gathered from [Genius.com](https://www.genius.com).

<https://www.kaggle.com/juicobowley/drake-lyrics>

< **drake\_data.csv** (781.85 KB)



Detail Compact Column

5 of 5 columns ▾

**About this file**

CSV contains lyrics, song title, album title, url, view count (at this time)

album	lyrics_title	lyrics_url	lyrics
Album Title	Song Title	Link to lyrics page on genius.com	Lyrics of song
Unreleased Songs 11% Scorpion 9% Other (234) 81%	278 unique values	278 unique values	277 unique values
Certified Lover Boy	Certified Lover Boy* Lyrics	<a href="https://genius.com/Drake-certified-lover-boy-lyrics">https://genius.com/Drake-certified-lover-boy-lyrics</a>	[Verse] Put my feelings on ice Always been a gem Certified lover b somehow still heartless Heart .

# Challenges

## Challenge 1

**Original dataset columns had weird string**

1st : Clean album and lyrics title columns

## Challenge 2

**Dataset very limited : featured engineering**

1st : Added the following columns : Add demo, featurings, number of verses, number of chorus, number of words, average length sentences, lexical diversity, positive / Negative / Neutral sentiments

## Challenge 3

**Gathering into albums**

1st : Groupby Albums

# Challenge details

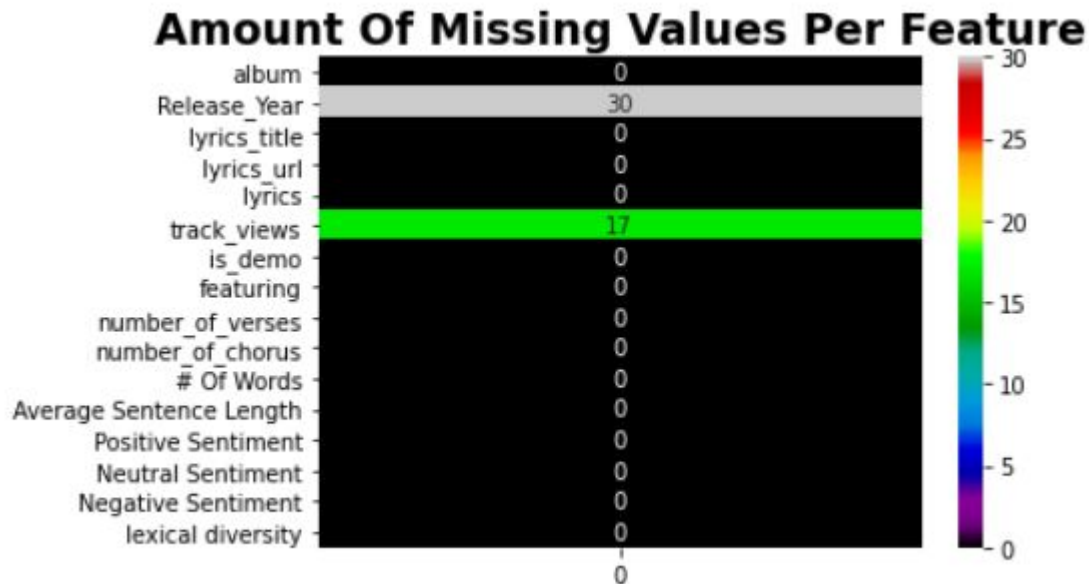
## Challenge 1 : Lyrics field is not complete

```
Entrée [5]: drake_data.loc[0,'lyrics']
```

```
Out[5]: "[Verse]\nPut my feelings on ice\nAlways been a gem\nCertified lover boy, somehow still heartless\nHeart is only gettin' colder"
```

# Challenge details

Challenge 2 : Missing values because songs have not been released



# Findings

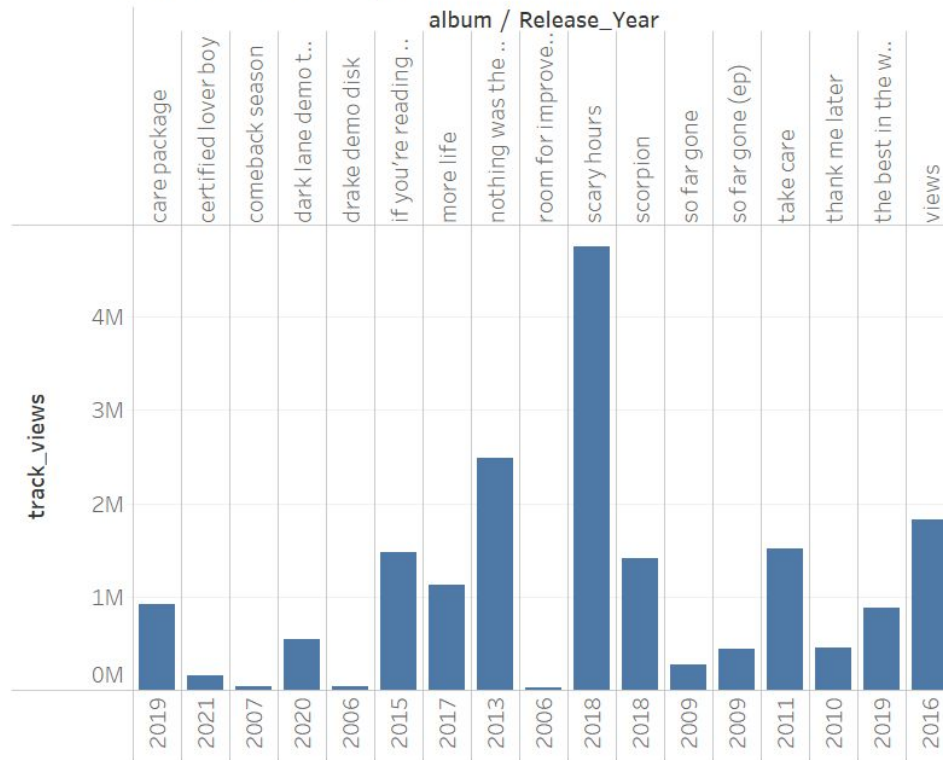
Features influencing Drake  
views on Genius.com



# Findings n°1

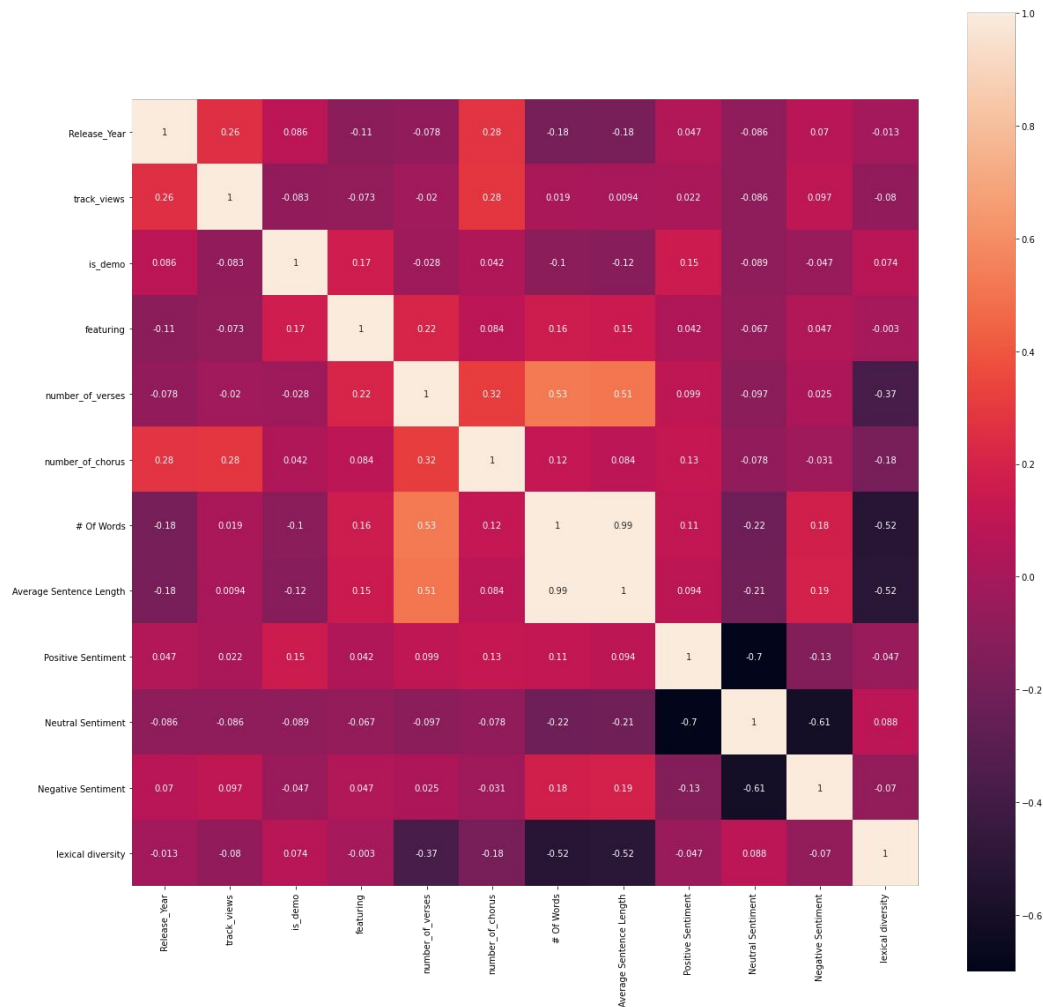
In 2018 “Scary Hours”  
was the most popular  
albums

Track views over the year



# Findings n°2

Positive correlation  
between the number  
of views and the  
number of chorus

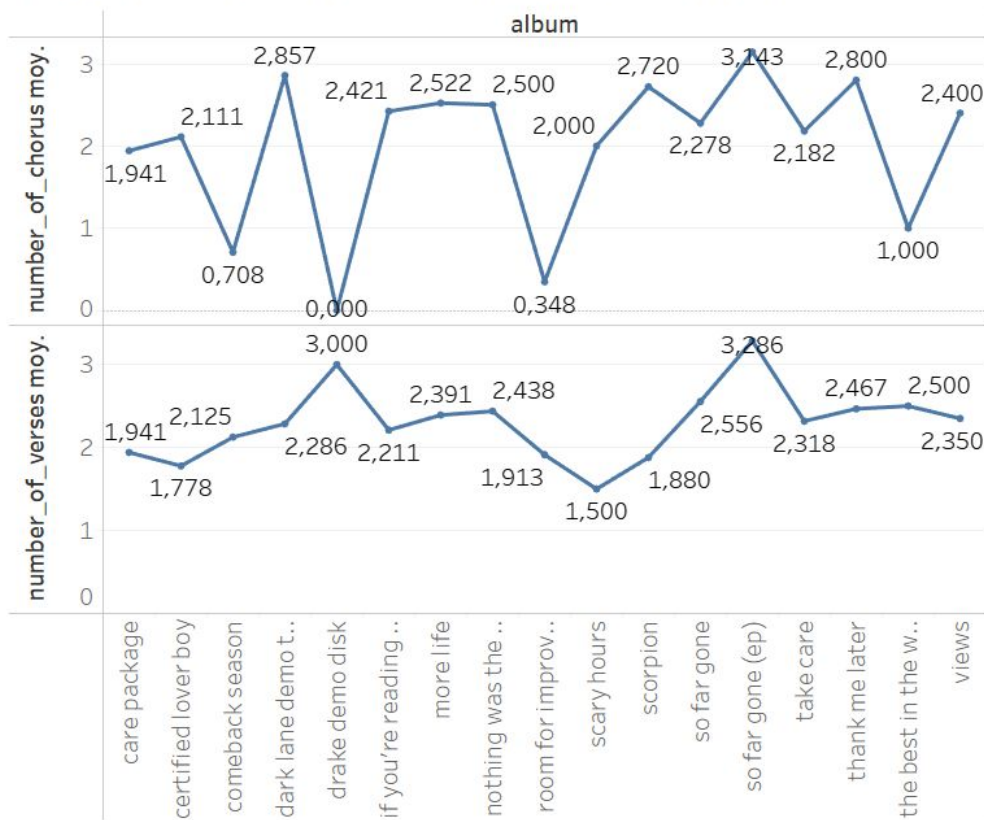




# Findings n°3

Scary Hours is not the album with the highest Chorus

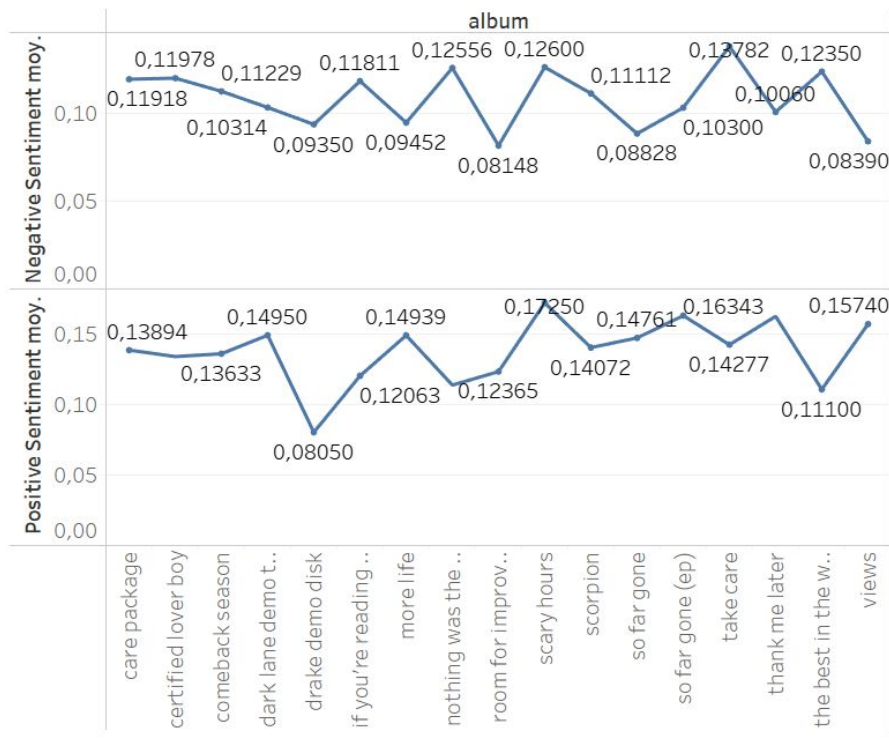
Number of verses and chorus over the years



# Findings n°4

Scary Hours is however the album with the highest positive and negative sentiment.

Sentiment analysis over the years



# Further analysis

-

Find more features related to the track features (SpotifyAPI)

Get the Full Lyrics to analyse the lyrics in more details

---

# Thanks

Orphee

