THE TOP MOST POPULAR SONGS ON SPOTIFY (2020-2021)

Maddie Niebling, Richa Shah, Emma Walsh DS 2001, Section 19



PROBLEM STATEMENT

WHAT MAKES A SONG **POPULAR?** *



- Source: Kaggle
 - spotify_dataset.csv: the top most popular songs (2020-2021)

	Index	Highest Charting Position	Number of Times Charted	Week of Highest Charting	Song Name	Streams	Artist	Artist Followers	Song ID	Genre	 Danceability	Energy	Loudness	Speechines
0	1	1	8	2021-07- 232021- 07-30	Beggin'	48,633,449	Måneskin	3377762	3Wrjm47oTz2sjlgck11l5e	['indie rock italiano', 'italian pop']	 0.714	0.8	-4.808	0.050
1	2	2	3	2021-07- 232021- 07-30	STAY (with Justin Bieber)	47,248,719	The Kid LAROI	2230022	5HCyWIXZPP0y6Gqq8TgA20	['australian hip hop']	 0.591	0.764	-5.484	0.048
2	3	1	11	2021-06- 252021- 07-02	good 4 u	40,162,559	Olivia Rodrigo	6266514	4ZtFanR9U6ndgddUvNcjcG	['pop']	 0.563	0.664	-5.044	0.15
3	4	3	5	2021-07- 022021- 07-09	Bad Habits	37,799,456	Ed Sheeran	83293380	6PQ88X9TkUIAUIZJHW2upE	['pop', 'uk pop']	 0.808	0.897	-3.712	0.034
4	5	5	1	2021-07- 232021- 07-30	INDUSTRY BABY (feat. Jack Harlow)	33,948,454	Lil Nas X	5473565	27NovPIUIRrOZoCHxABJwK	['lgbtq+ hip hop', 'pop rap']	 0.736	0.704	-7.409	0.061





M









Based on the dataset, here are the top 10 songs for 2020-2021 on Spotify, based on number of streams.

Figure 1 1e7 Song Streams for the Top 10 Songs in 2020-2021 Number of Streams (in 10 Millions) 4.5 4.0 3.5 3.0 2.5 og INDUSTRY BABY (feat. Jack Harlow) -bu w monTERO (Call Me By Your Name) -**Konaguni** WANNA BE YOUR SLAVE good 4 u Bad Habits Todo De Ti Beggin' STAY (with Justin Bieber) Kiss Me More (feat. SZA)

6

DESCRIPTION OF THE DATA



- Avid music listeners
- New songs come out daily
 - What contributes to a popular song?
- Overplayed songs on social media (i.e., TikTok during COVID)





- \sim 25 columns, qualitative and quantitative data
- We chose: Song Name, Index, Streams, Artist Followers, Genre, and Artist

METHODS

OUR METHODS

CREATING A DATA FRAME

User-centric data frame

Allowing user to input a specific genre and what parameters they want within that genre

NumPy & Pandas

CLEANING THE DATA

Original data set included over 1,500 songs, cleaned down to top 50

Dataset contained 23 columns per song, chose to look at only 6

DATA VISUALIZATION & LINEAR REGRESSION

Created different plots to show correlation between different columns

Very user centric, indicates correlations of the user's choice

Linear regression: correlation between Artist's Followers and their position on the chart

SAMPLE CODE ON THE NEXT SLIDE



```
def getGenre(this_genre, column_name):
```

```
'''Function: get song genre
   Parameters: genre_list, column_name
   Return: df['Song Name']
'''
```

```
df = data[data.apply(lambda row: this_genre in row['Genre'], axis = 1)]
return df[column_name]
```

```
def plot_categories(x, column1, column2):
    '''Function: plot_categories
    Parameters: x, column1, column2
    Return: plot showing a specified amount of time, comparing 2 columns
    '''
    plt.plot(data[column1][:x], data[column2][:x], color = "tomato")
    plt.xlabel(str(column1))
    plt.ylabel(str(column2))
    plt.title(str(column2)+" for the top "+str(x)+" "+str(column1)+" in 2020-2021")
    plt.xticks(rotation = 90)
    return plt.show
```

RESULTS, CONCLUSION, AND FUTURE WORK

Hypothesis:

- If the song ranks higher on the chart, it will have a higher number of streams.
- Results:
 - Måneskin was the top artist
 (based on number of streams)
 - Interesting because their songwas extremely popular on TikTok

Figure 2





index	Song Name	Streams
0	Beggin'	48633449.0
2	good 4 u	40162559.0
3	Bad Habits	37799456.0
4	INDUSTRY BABY (feat. Jack Harlow)	33948454.0
5	MONTERO (Call Me By Your Name)	30071134.0
6	Kiss Me More (feat. SZA)	29356736.0
7	Todo De Ti	26951613.0
9	I WANNA BE YOUR SLAVE	24551591.0
10	Levitating (feat. DaBaby)	23518010.0
12	Permission to Dance	22062812.0
13	Peaches (feat. Daniel Caesar & Giveon)	20294457.0
14	Butter	19985713.0
15	traitor	19480679.0
16	deja vu	18571755.0
17	Save Your Tears (with Ariana Grande) (Remix)	18053141.0
19	Need To Know	16908917.0
21	Ain't Shit	16126897.0
22	drivers license	15684978.0
23	Heat Waves	15089184.0
24	Blinding Lights	15011809.0
25	happier	14849078.0
28	Friday (feat. Mufasa & Hypeman) - Dopamine Re-Edit	14273637.0
30	Leave The Door Open	13963568.0
31	Don't Go Yet	13205250.0
32 33	favorite crime Pareja Del Año	12818935.0 12572307.0
34	Watermelon Sugar	11996689.0
37	You Right	11629418.0
38	911 - Remix	11553846.0
40	Miénteme	11477024.0
41	jealousy, jealousy	11287301.0
43	Dynamite	11196819.0
44	The Business	10739770.0
45	Heartbreak Anthem (with David Guetta & Little Mix)	10315460.0
50	Nicky Jam: Bzrp Music Sessions, Vol. 41	9799701.0

70% of the top 50 songs are considered within the 'pop' genre

Figure 3



- If the artist has a larger number of followers, their song will be more popular.
- Results:
 - This was not the case
 - ▷ Ed Sheeran → greatest follower count on Spotify
 - Does not correlate with song ranking in terms of streams





Hypothesis:

- If the artist has a larger number of followers, their song will be more popular.
- Results:
 - Similarly, the highest point (outlier) corresponds to Ed Sheeran, who is one of the most experienced artists on the list
 - Second outlier is Bad Bunny, also an experienced artist
 - $\triangleright \quad Lower points \rightarrow newer artists$



CONCLUSION

- Hoping to identify popularity using ranking, streams, and followers
- Figure 2: correlation between song ranking and number of streams
- ► Figures 3 & 4: correlation between artist followers and song popularity
 - $\blacktriangleright \quad \text{Higher # of followers} \rightarrow \text{more popular song}$
 - Other factors: TikTok, IG Reels, viral hits, "one-hit-wonders"
 - \triangleright Ed Sheeran and Bad Bunny \rightarrow two most experienced artists

FUTURE WORK

- Logistic regression to predict the next popular song
 - Elements included in the next popular song
 - ▶ i.e. danceability, bpm, acousticness, energy
 - How artists should create songs to ensure it tops the charts
- Emphasize how social media followers affect song popularity
 - Rise of TikTok
 - $\qquad \qquad \mathsf{Music going viral on social media platforms} \rightarrow \mathsf{bigger audience}$

THANK YOU ANY QUESTIONS?