API Project

Oscar Lidheim

APIs

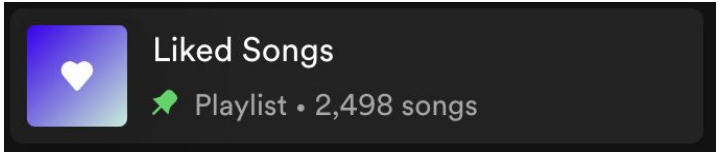
- **Retrieve info from my personal Spotify account including:**
  - Discover Weekly Playlist (30 tracks) and All Liked Songs (2500 tracks)
  - Valence (happiness/sadness index), tempo, instrumentalness, loudness, acousticness and others (10 in total)

- **Used to enrich data (genres) and summarize data**
- Need to use latest GPT-4 (real time and more advanced)
- Paid service and quite expensive (if you were to run on large datasets)

Two Main Analyses

1 **Analysing 10 years of my Spotify Listening**

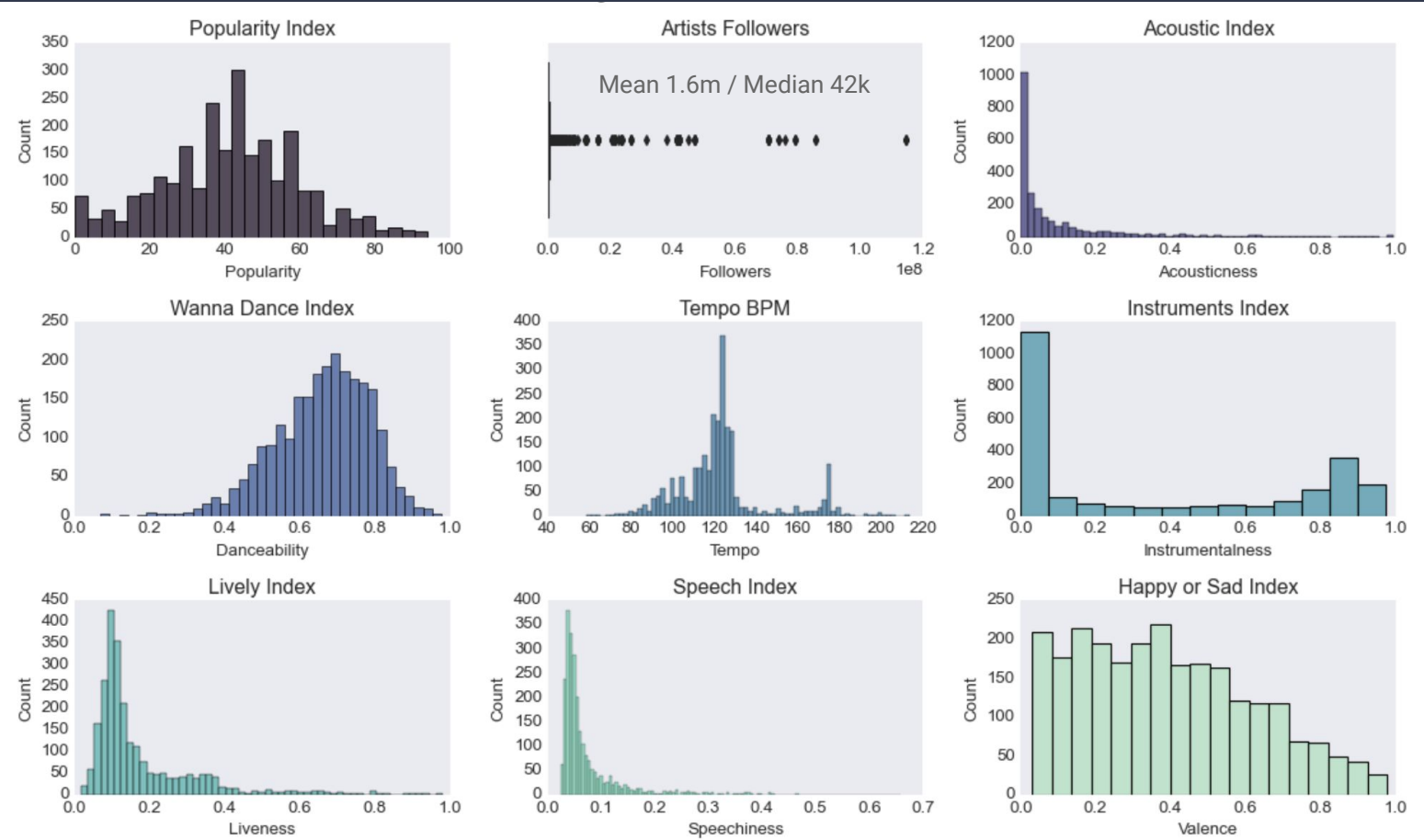
- Any clear trends in my listening data?



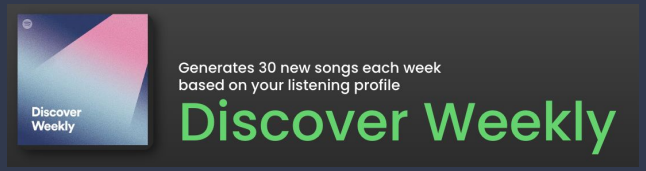
2 **Comparing this week's Discover Weekly 30 with my 10 years**

- Why always so much electronic music?

Audio Features of 10 Years of my Music



Discover Weekly Tracks



2

	Song	Artist(s)	Genre(s)	Popularity	Followers	Acousticness	Danceability	Tempo	Instrumentalness	Liveness	Loudness
0	Celestial	Nomyn	[]	37	20017	0.001200	0.481000	82.451000	0.923000	0.107000	-6.078000
1	Seaside	Attom	[]	44	41870	0.076700	0.655000	105.107000	0.904000	0.141000	-6.781000
2	Meet You There	LE YORA, SOMMA, Jewels, yuma., MAGNUS	[]	27	1175	0.012600	0.689000	122.011000	0.820000	0.153000	-9.441000
3	falling easy	sad face.	[]	26	2258	0.464000	0.728000	107.026000	0.021000	0.112000	-7.749000
4	Way Back	Boeuv	[]	20	573	0.884000	0.668000	126.028000	0.873000	0.101000	-12.513000
5	Damaged Friday	eeph, Jani, Feverkin	[]	11	667	0.170000	0.508000	159.898000	0.489000	0.054600	-10.127000
6	Forever Young - Morgin Madison Remix	JES, Morgin Madison	[trance, vocal trance]	41	28140	0.011600	0.556000	125.045000	0.000132	0.186000	-3.822000
7	Manzanita	Surf Mesa	[pop edm]	60	162133	0.003510	0.612000	126.986000	0.858000	0.144000	-5.081000
8	Goosebumps	Whammyboy	[]	30	2368	0.048000	0.742000	130.037000	0.000568	0.075500	-5.780000
9	Tell Me	Hermei	[wave]	32	3887	0.142000	0.562000	110.997000	0.853000	0.152000	-11.483000
10	Unravel	Zorro, Ooyy	[]	37	2295	0.134000	0.578000	106.031000	0.001280	0.600000	-9.042000

Using OpenAI API to Enrich Data

Query

Model: gpt-4 (latest)

Assistant Role:

“You are specialist in music and categorizing genres of songs”

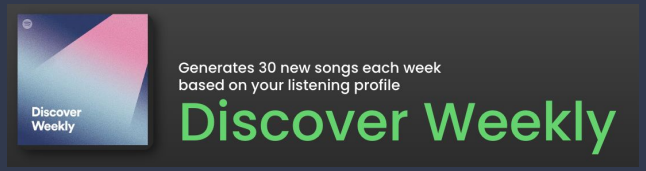
Message:

“Can you add the genre of each of these songs and return the updated list. Please only change the genre for those with no genre. Answer with only the updated list in json format. {dw_data_gpt}”

Key Learnings

- Did not work with earlier models such as gpt-3 or gpt-3.5
- High Costs of adding genre to 30 rows c. $0.10
- Testing could be done to correctly prompt / adjust temperature (randomness of response)

Results - Pretty good accuracy



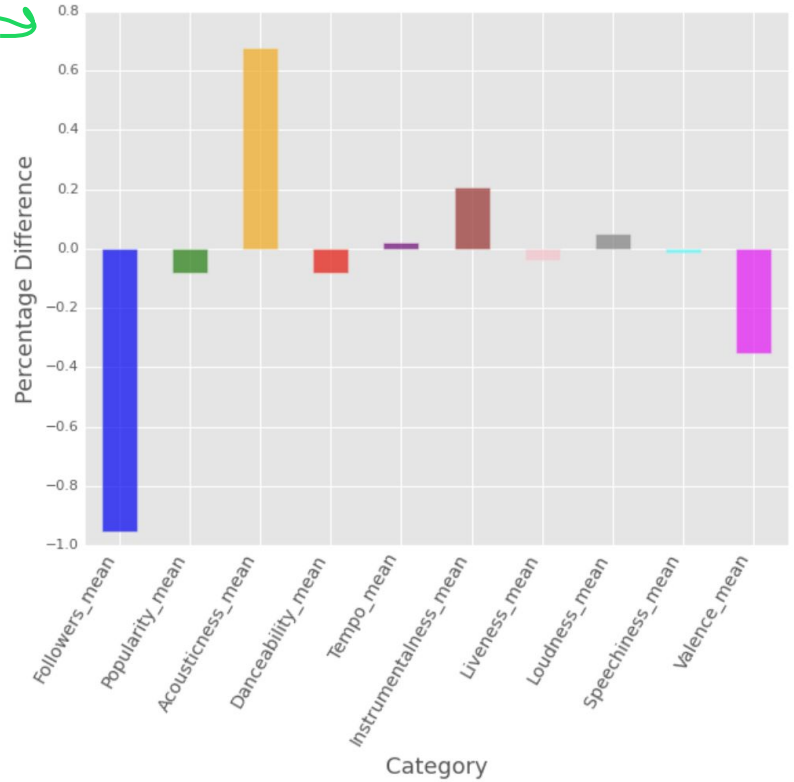
2

	Song	Artist(s)	Genre(s)	Genre(s)+GPT	Popularity	Followers	Acousticness	Danceability	Tempo	Instrumentalness	Liveness	Loudness
0	Celestial	Nomyn	[]	[ambient]	37	20017	0.001200	0.481000	82.451000	0.923000	0.107000	-6.078000
1	Seaside	Attom	[]	[chillwave]	44	41870	0.076700	0.655000	105.107000	0.904000	0.141000	-6.781000
2	Meet You There	LE YORA, SOMMA, Jewels, yuma., MAGNUS	[]	[pop]	27	1175	0.012600	0.689000	122.011000	0.820000	0.153000	-9.441000
3	falling easy	sad face.	[]	[lo-fi beats]	26	2258	0.464000	0.728000	107.026000	0.021000	0.112000	-7.749000
4	Way Back	Boeuv	[]	[deep house]	20	573	0.884000	0.668000	126.028000	0.873000	0.101000	-12.513000
5	Damaged Friday	eeph, Jani, Feverkin	[]	[electronica]	11	667	0.170000	0.508000	159.898000	0.489000	0.054600	-10.127000
6	Forever Young - Morgin Madison Remix	JES, Morgin Madison	[trance, vocal trance]	[trance, vocal trance]	41	28140	0.011600	0.556000	125.045000	0.000132	0.186000	-3.822000
7	Manzanita	Surf Mesa	[pop edm]	[pop edm]	60	162133	0.003510	0.612000	126.986000	0.858000	0.144000	-5.081000
8	Goosebumps	Whammyboy	[]	[trap]	30	2368	0.048000	0.742000	130.037000	0.000568	0.075500	-5.780000
9	Tell Me	Hermei	[wave]	[wave]	32	3887	0.142000	0.562000	110.997000	0.853000	0.152000	-11.483000
10	Unravel	Zorro, Ooyy	[]	[downtempo]	37	2295	0.134000	0.578000	106.031000	0.001280	0.600000	-9.042000

Analysis - Comparing DW to All Liked Tracks

2

	All Liked 2500 Songs	Discover Weekly 30	Difference	Difference Pct
Followers_mean	1629196.842507	76832.400000	-1552364.442507	-0.952840
Popularity_mean	41.413419	38.066667	-3.346752	-0.080813
Acousticness_mean	0.115428	0.193542	0.078114	0.676736
Danceability_mean	0.657942	0.605200	-0.052742	-0.080162
Tempo_mean	123.119257	125.483167	2.363909	0.019200
Instrumentalness_mean	0.353309	0.426461	0.073152	0.207048
Liveness_mean	0.177929	0.171503	-0.006426	-0.036113
Loudness_mean	-7.124739	-7.483833	-0.359095	0.050401
Speechiness_mean	0.076074	0.075127	-0.000947	-0.012447
Valence_mean	0.390446	0.252530	-0.137916	-0.353227



- Low follower count (even considering median)
- More acoustic instruments this week (wow!)
- Even with my right sadness skew this is even more sad!

Main Findings - why so much electronic?

Genres of All Liked Songs:

[uncategorized]	566 (23%)
[nightrun, popwave]	56
[progressive trance house]	51
[progressive house]	34
[complextro]	31

Asking GPT-4 to summarize genres of this week's Discover Weekly:

"20 songs fall under electronic music or a subgenre of electronic. EDM and Pop are the most common genres, with a few songs also falling under ambient, lo-fi, house, and trance."

- Is Spotify taking into account my 10 years of history and not accounting for **change in music taste?**
- **Subgenres are diversified** but when looking at broader genre is mostly electronics
- Does Spotify know my music taste better than myself (own self-image)?

Challenges & Learnings

- **Quick to get good data from popular API**
  - Both Spotify and OpenAI have python libraries and good documentation (chat GPT very helpful)
- **Quite a lot of data cleaning and thinking how it flows from one api to another**
  - Lists to DataFrame to JSON to DataFrame
- **Visualising took more time than expected**
  - Sizing, colors, titles
- **OpenAI will clearly be helpful for tasks in the future and cost will probably continue to go down**
  - Expensive to use gpt-4 for large data sets

Thank you

I will share code if you want to see your Spotify profile