DATA ENGINEERING INDIVIDUAL COURSEWORK

SPOTIFY PLAYLIST DATABASE:

A SENTIMENT ANALYSIS

April 2022

Word Count:

TABLE OF CONTENTS

1.0 INTRODUCTION

"Music can heal the wounds which medicine cannot touch", says Debasish Mridha. From many decades ago, music has already been recognised as an opportunity to address mental health challenges (Schriewer and Bulaj, 2016). Nowadays, audience can enjoy the music more conveniently via music streaming services, instead of downloading the original audio file of a song. As one of the biggest music streaming platforms, Spotify had over 365 million users by February 2022 (Caddy, 2022).

One important feature of Spotify is the editorial playlists. Spotify's in-house teams curate these playlists by selecting collections of songs that somehow have some similarities so that meaningful playlists are created. Many playlists have emotion-related tags such as "happy", "moody" and "chilling", and each of them is designed to match the audience's emotions.

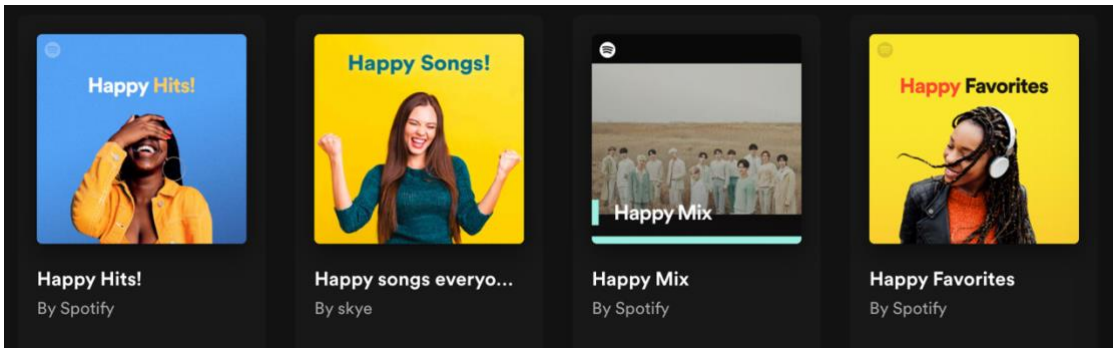

Figure 1: Spotify Curated Playlists Related to "Happy"

While the Spotify playlist is an important feature of the platform in helping the artists reach more of the target audiences, it is worthy of investigation that the exact criterion of how a song is featured on a certain playlist. This project aims to create a database of a specific Spotify playlist that contains a number of different attributes of the songs within that playlist and conduct further sentiment analyses of these songs based on the attributes.

In achieving the objective of the project, the following steps are executed and will be explained in detail throughout this report: First, real-time streaming data will be extracted through API scraping from Spotify and several other relevant platforms. The data will then be processed and stored in a suitable way so that further analyses are allowed. This project will then conduct sentiment analyses on these processed data so that we can evaluate whether the “emotion” of the songs in the playlist match with the playlist title. It is believed that successful execution of the project with provide precious guidelines for sentiment related research on a bigger database beyond the scope of this project.

Figure 2: Project Workflow

2.0
DATA COLLECTION

This project mainly used the Application Programming Interface (API) techniques in the Python language to extract data from four different platforms including Spotify, MusixMatch, YouTube and Twitter. The API is an important tool to obtain data from dynamic websites and allows some level of customisations (Medium,2022).

2.1 SPOTIFY DATA

As the main subject matter of this project, data from Spotify was scraped first. Spotify provides API services on its developer website and a number of functions can be achieved under the free user plan. On the developer website, a Python library called Spotipy (https://github.com/plamere/spotipy) is recommended for executing the API calls. This package was utilised by this project throughout the data extraction from Spotify.

First of all, the project wanted to retrieve a list of featured playlists (Editor's Picks) on Spotify to decide which playlist to extract and evaluate. By using the *featured_playlists()* method, a list of playlists which were the Editor's picks on the day when the data was scraped was returned:

```
Editor's picks
0 New Music Friday
1 Feel Good Friday
2 RapCaviar
3 Main Stage
4 I Love My '90s Hip-Hop
5 Mood Booster
6 Dance Hits
7 Today's Top Hits
8 just hits
9 Dance Party
10 Happy 80s
11 young & free
```

Figure 3: Featured Playlists on 8th April

The sixth result was a playlist named ***"Mood Booster"***, which was highly related to the research objective - to analyse the emotion of songs with sentimental analysis techniques and evaluate whether the emotion of the song match with the overall playlist genre.

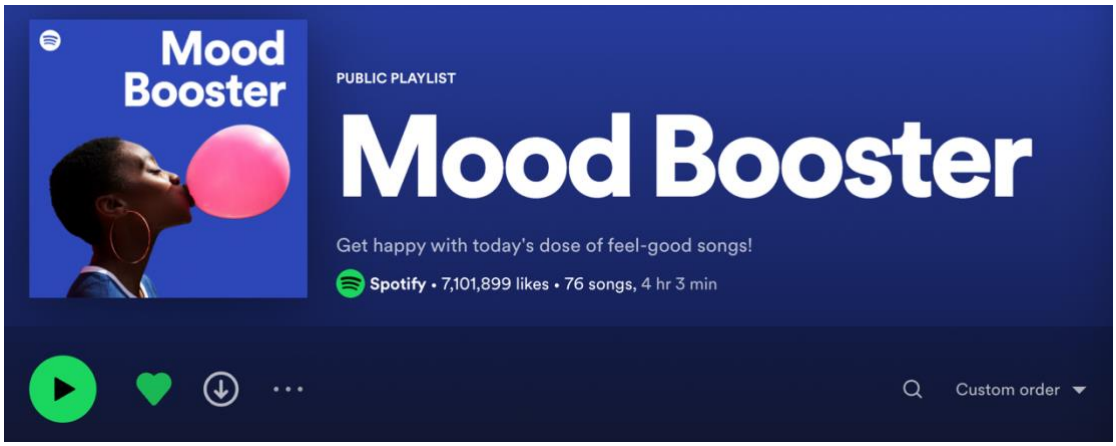

Figure 4: The *Mood Booster* Playlist

According to Spotify's description of this playlist, the songs in the playlist are supposed to make the audience "feel good" and "get happy". This project is interested in the if the emotions of the songs in this playlist actually match that purpose. To dig deeper into the tracks in this playlist, the playlist's ID on Spotify needs to be known. Unfortunately, at the moment, the only feasible way to get a playlist's ID is through getting a user's current playlist. Therefore, this playlist was manually followed on Spotify and added to the profile using a personal Spotify account.

With the playlist id being known, the project was able to retrieve more information about the tracks in this the playlist. A list 76 song names were first retrieved via the *playlist_items()* method, and the 76 songs' corresponding Spotify IDs were retrieved via similar ways. The songs' Spotify IDs allowed the project to retrieve a number of attributes of the track by using the Spotify API's *track()* method, including the artist information. By calling the method and looping into the nested dictionary returned by the method, this project was able to get the *artist name, album, Spotify popularity, release date, duration* information. The method could also tell whether a track contains explicit content or not: most music streaming platforms distinguish and differentiate between tracks that is suitable for mainstream consumption, and those songs that may contain a parental advisory or may be considered explicit content (Soundplate, 2022). On Spotify, a track with explicit content will have a "E" or "Explicit" symbol next to its title. With the API, boolean values of True or False was returned regarding the "explicit" attribute.

3.0
DATA PROCESSING

HEADING 2

View and edit this document in Word on your computer, tablet or phone. You can edit text, easily insert content such as pictures, shapes and tables, and seamlessly save the document to the cloud from Word on your Windows, Mac, Android or iOS device.

4.0 RELATIONAL DATABASE

HEADING 2

View and edit this document in Word on your computer, tablet or phone. You can edit text, easily insert content such as pictures, shapes and tables, and seamlessly save the document to the cloud from Word on your Windows, Mac, Android or iOS device.

5.0
DATA STORAGE

HEADING 2

View and edit this document in Word on your computer, tablet or phone. You can edit text, easily insert content such as pictures, shapes and tables, and seamlessly save the document to the cloud from Word on your Windows, Mac, Android or iOS device.

6.0
EXPLORATORY DATA ANALYSIS

HEADING 2

View and edit this document in Word on your computer, tablet or phone. You can edit text, easily insert content such as pictures, shapes and tables, and seamlessly save the document to the cloud from Word on your Windows, Mac, Android or iOS device.

7.0
FURTHER ANALYSES

HEADING 2

View and edit this document in Word on your computer, tablet or phone. You can edit text, easily insert content such as pictures, shapes and tables, and seamlessly save the document to the cloud from Word on your Windows, Mac, Android or iOS device.

8.0
DATA PIPELINE

HEADING 2

View and edit this document in Word on your computer, tablet or phone. You can edit text, easily insert content such as pictures, shapes and tables, and seamlessly save the document to the cloud from Word on your Windows, Mac, Android or iOS device.

9.0
FUTURE OPPORTUNITIES

HEADING 2

View and edit this document in Word on your computer, tablet or phone. You can edit text, easily insert content such as pictures, shapes and tables, and seamlessly save the document to the cloud from Word on your Windows, Mac, Android or iOS device.

10.0
CONCLUSION

HEADING 2

View and edit this document in Word on your computer, tablet or phone. You can edit text, easily insert content such as pictures, shapes and tables, and seamlessly save the document to the cloud from Word on your Windows, Mac, Android or iOS device.

11.0 BIBLIOGRAPHIES

REPORT REFERENCES

Schriewer, K. and Bulaj, G., 2016. Music Streaming Services as Adjunct Therapies for Depression, Anxiety, and Bipolar Symptoms: Convergence of Digital Technologies, Mobile Apps, Emotions, and Global Mental Health. *Frontiers in Public Health*, 4.

Caddy, B., 2022. *The best music streaming services 2022: Spotify, Apple Music, Tidal and more*. [online] TechRadar. Available at: <https://www.techradar.com/best/the-best-music-streaming-services-2021> [Accessed 18 April 2022].

Medium. 2022. *Using an API for Web Scraping: A List of the Best Advantages*. [online] Available at: <https://medium.com/api-world/using-an-api-for-web-scraping-a-list-of-the-best-advantages-586e9fec2660> [Accessed 19 April 2022].

Soundplate. 2022. *What Does 'Explicit Content' Mean on Spotify, Apple Music & Other Music Streaming Platforms*. [online] Available at: <https://soundplate.com/what-does-explicit-content-mean-on-spotify-apple-music-other-music-streaming-platforms/> [Accessed 19 April 2022].

12.0
APPENDIX

HEADING 2

View and edit this document in Word on your computer, tablet or phone. You can edit text, easily insert content such as pictures, shapes and tables, and seamlessly save the document to the cloud from Word on your Windows, Mac, Android or iOS device.