Personalisation- Assignment 1

Korina - Kyriaki Zaromytidou

The code for this assignment is taken from the in-class notebook Week 3.1

I run the code for all tasks. I have elaborated on the Task 2 - Song Similarity & Task 3 - Top N Recommendations including comments based on the suggested questions for each of those tasks. ## I used a personal playlist for this assigment. The playlist is called 'Make it Rain - Live'. It has 18 songs of varied artist in English.

Starting with installing libraries.

In [58]:
```
!pip install spotipy
```

```
Requirement already satisfied: spotipy in /opt/miniconda3/envs/pml/lib/pytho
n3.9/site-packages (2.23.0)
Requirement already satisfied: redis>=3.5.3 in /opt/miniconda3/envs/pml/lib/
python3.9/site-packages (from spotipy) (4.5.5)
Requirement already satisfied: requests>=2.25.0 in /opt/miniconda3/envs/pml/
lib/python3.9/site-packages (from spotipy) (2.31.0)
Requirement already satisfied: urllib3>=1.26.0 in /opt/miniconda3/envs/pml/l
ib/python3.9/site-packages (from spotipy) (2.0.2)
Requirement already satisfied: six>=1.15.0 in /opt/miniconda3/envs/pml/lib/p
ython3.9/site-packages (from spotipy) (1.16.0)
Requirement already satisfied: async-timeout>=4.0.2 in /opt/miniconda3/envs/
pml/lib/python3.9/site-packages (from redis>=3.5.3->spotipy) (4.0.2)
Requirement already satisfied: certifi>=2017.4.17 in /opt/miniconda3/envs/pm
l/lib/python3.9/site-packages (from requests>=2.25.0->spotipy) (2023.5.7)
Requirement already satisfied: idna<4,>=2.5 in /opt/miniconda3/envs/pml/lib/
python3.9/site-packages (from requests>=2.25.0->spotipy) (3.4)
Requirement already satisfied: charset-normalizer<4,>=2 in /opt/miniconda3/e
nvs/pml/lib/python3.9/site-packages (from requests>=2.25.0->spotipy) (3.1.0)
```

In [59]:
```
! pip install matplotlib
```

```
Requirement already satisfied: matplotlib in /opt/miniconda3/envs/pml/lib/py
thon3.9/site-packages (3.7.1)
Requirement already satisfied: cycler>=0.10 in /opt/miniconda3/envs/pml/lib/
python3.9/site-packages (from matplotlib) (0.11.0)
Requirement already satisfied: pillow>=6.2.0 in /opt/miniconda3/envs/pml/li
b/python3.9/site-packages (from matplotlib) (9.5.0)
Requirement already satisfied: kiwisolver>=1.0.1 in /opt/miniconda3/envs/pm
l/lib/python3.9/site-packages (from matplotlib) (1.4.4)
Requirement already satisfied: importlib-resources>=3.2.0 in /opt/miniconda
3/envs/pml/lib/python3.9/site-packages (from matplotlib) (5.12.0)
Requirement already satisfied: fonttools>=4.22.0 in /opt/miniconda3/envs/pm
l/lib/python3.9/site-packages (from matplotlib) (4.39.4)
Requirement already satisfied: packaging>=20.0 in /opt/miniconda3/envs/pml/l
ib/python3.9/site-packages (from matplotlib) (23.1)
Requirement already satisfied: contourpy>=1.0.1 in /opt/miniconda3/envs/pml/
lib/python3.9/site-packages (from matplotlib) (1.0.7)
Requirement already satisfied: numpy>=1.20 in /opt/miniconda3/envs/pml/lib/p
ython3.9/site-packages (from matplotlib) (1.23.5)
Requirement already satisfied: pyparsing>=2.3.1 in /opt/miniconda3/envs/pml/
lib/python3.9/site-packages (from matplotlib) (3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in /opt/miniconda3/envs/
pml/lib/python3.9/site-packages (from matplotlib) (2.8.2)
Requirement already satisfied: zipp>=3.1.0 in /opt/miniconda3/envs/pml/lib/p
ython3.9/site-packages (from importlib-resources>=3.2.0->matplotlib) (3.15.
0)
Requirement already satisfied: six>=1.5 in /opt/miniconda3/envs/pml/lib/pyth
on3.9/site-packages (from python-dateutil>=2.7->matplotlib) (1.16.0)
```

In [60]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

In [61]:
```
SPOTIPY_CLIENT_ID='4547ac9120de4cbba75a19ee6dc3b908'
SPOTIPY_CLIENT_SECRET='75fa035f5a0b47b7868e69a93ef878b4'
```

In [62]:
```
import spotipy
from spotipy.oauth2 import SpotifyClientCredentials

auth_manager = SpotifyClientCredentials(SPOTIPY_CLIENT_ID,SPOTIPY_CLIENT_SEC
sp = spotipy.Spotify(auth_manager=auth_manager)
```

In [63]:
```
playlist_id = '2sU7avcVJBJ7rJrimHcoli'
```

In [64]:
```
limit = 50
trim = True
```

In [65]:
```
#Get playlsit
playlist = sp.playlist(playlist_id)
#Extract tracks
tracks = playlist["tracks"]["items"]
if trim:
    tracks = tracks[-limit:]
#Extract track IDs
ids = [track["track"]["id"] for track in tracks]
#Get audio features for tracks
features =  pd.DataFrame(sp.audio_features(ids))
#Label dataframe rows with track name and artist
labels = [track["track"]["artists"][0]["name"] + " - "+ track["track"]["name
features.index = labels
```

In [46]:
```
features
```

Out[46]:

	danceability	energy	key	loudness	mode	speechiness	acousticness	instrume
Foy Vance - Make It Rain - Live	0.367	0.309	9	-10.167	0	0.0832	0.1450	
Michael Kiwanuka - Love & Hate	0.308	0.513	0	-7.808	1	0.0292	0.1740	
Jack Johnson - Good People	0.721	0.588	11	-6.018	1	0.0331	0.2190	
Son Little - Mad About You	0.599	0.377	7	-9.829	1	0.1050	0.4410	
Foy Vance - She Burns	0.730	0.403	4	-10.146	1	0.0483	0.7180	
KALEO - Way down We Go - Recorded at Spotify Studios NYC	0.602	0.459	8	-7.167	0	0.0360	0.6300	
Leon Bridges - Coming Home	0.427	0.465	6	-6.595	1	0.0291	0.4610	
Michael Kiwanuka - Home Again	0.564	0.320	9	-11.473	0	0.0306	0.7060	
Michael Kiwanuka - One More Night	0.684	0.764	6	-5.342	0	0.0254	0.0129	
Robert Plant - Can't Let Go	0.713	0.824	10	-6.895	0	0.0502	0.1150	
Michael Kiwanuka - Cold Little Heart - Radio Edit	0.629	0.580	7	-6.208	0	0.0278	0.3570	
The Teskey Brothers - So Caught Up	0.744	0.692	4	-6.561	0	0.0260	0.3090	
David Gray - Sail Away	0.651	0.568	0	-7.882	0	0.0399	0.1190	
Irene Skylakaki	0.549	0.454	10	-7.785	1	0.0326	0.5270	

	danceability	energy	key	loudness	mode	speechiness	acousticness	instrume
- Tonight								
Ed Sheeran - Thinking out Loud	0.781	0.445	2	-6.061	1	0.0295	0.4740	
Sam Smith - I'm Not The Only One	0.677	0.485	5	-5.795	1	0.0361	0.5290	
Sia - Elastic Heart	0.421	0.791	9	-4.998	1	0.0496	0.0117	
Sia - Chandelier	0.399	0.787	1	-2.880	1	0.0499	0.0197	

Audio Features

In [47]:
```
features.columns
```

Out[47]:
```
Index(['danceability', 'energy', 'key', 'loudness', 'mode', 'speechiness',
       'acousticness', 'instrumentalness', 'liveness', 'valence', 'tempo',
       'type', 'id', 'uri', 'track_href', 'analysis_url', 'duration_ms',
       'time_signature'],
      dtype='object')
```

In [48]:
```
my_features = ['danceability', 'energy', 'key',
               'loudness', 'mode', 'speechiness',
               'acousticness', 'instrumentalness',
               'liveness', 'valence', 'tempo', 'duration_ms',
               'time_signature']
features = features[my_features]
features
```

Out[48]:

	danceability	energy	key	loudness	mode	speechiness	acousticness	instrume
Foy Vance - Make It Rain - Live	0.367	0.309	9	-10.167	0	0.0832	0.1450	
Michael Kiwanuka - Love & Hate	0.308	0.513	0	-7.808	1	0.0292	0.1740	
Jack Johnson - Good People	0.721	0.588	11	-6.018	1	0.0331	0.2190	
Son Little - Mad About You	0.599	0.377	7	-9.829	1	0.1050	0.4410	
Foy Vance - She Burns	0.730	0.403	4	-10.146	1	0.0483	0.7180	
KALEO - Way down We Go - Recorded at Spotify Studios NYC	0.602	0.459	8	-7.167	0	0.0360	0.6300	
Leon Bridges - Coming Home	0.427	0.465	6	-6.595	1	0.0291	0.4610	
Michael Kiwanuka - Home Again	0.564	0.320	9	-11.473	0	0.0306	0.7060	
Michael Kiwanuka - One More Night	0.684	0.764	6	-5.342	0	0.0254	0.0129	
Robert Plant - Can't Let Go	0.713	0.824	10	-6.895	0	0.0502	0.1150	
Michael Kiwanuka - Cold Little Heart - Radio Edit	0.629	0.580	7	-6.208	0	0.0278	0.3570	
The Teskey Brothers - So Caught Up	0.744	0.692	4	-6.561	0	0.0260	0.3090	
David Gray - Sail Away	0.651	0.568	0	-7.882	0	0.0399	0.1190	
Irene Skylakaki	0.549	0.454	10	-7.785	1	0.0326	0.5270	

	danceability	energy	key	loudness	mode	speechiness	acousticness	instrume
- Tonight								
Ed Sheeran - Thinking out Loud	0.781	0.445	2	-6.061	1	0.0295	0.4740	
Sam Smith - I'm Not The Only One	0.677	0.485	5	-5.795	1	0.0361	0.5290	
Sia - Elastic Heart	0.421	0.791	9	-4.998	1	0.0496	0.0117	
Sia - Chandelier	0.399	0.787	1	-2.880	1	0.0499	0.0197	

Task 1 - Sorting by Audio Features

In [68]:
```
sort_by = 'liveness'
features.sort_values(by = sort_by, ascending = False)[:16]
```

Out[68]:

	danceability	energy	key	loudness	mode	speechiness	acousticness	instrumer
Michael Kiwanuka - Cold Little Heart - Radio Edit	0.629	0.580	7	-6.208	0	0.0278	0.3570	0
Foy Vance - Make It Rain - Live	0.367	0.309	9	-10.167	0	0.0832	0.1450	0
David Gray - Sail Away	0.651	0.568	0	-7.882	0	0.0399	0.1190	0
The Teskey Brothers - So Caught Up	0.744	0.692	4	-6.561	0	0.0260	0.3090	0
Michael Kiwanuka - Love & Hate	0.308	0.513	0	-7.808	1	0.0292	0.1740	0
Ed Sheeran - Thinking out Loud	0.781	0.445	2	-6.061	1	0.0295	0.4740	0
Leon Bridges - Coming Home	0.427	0.465	6	-6.595	1	0.0291	0.4610	0
Sia - Elastic Heart	0.421	0.791	9	-4.998	1	0.0496	0.0117	0
KALEO - Way down We Go - Recorded at Spotify Studios NYC	0.602	0.459	8	-7.167	0	0.0360	0.6300	0
Michael Kiwanuka - One More Night	0.684	0.764	6	-5.342	0	0.0254	0.0129	0
Michael Kiwanuka - Home Again	0.564	0.320	9	-11.473	0	0.0306	0.7060	0
Robert Plant -	0.713	0.824	10	-6.895	0	0.0502	0.1150	0

	danceability	energy	key	loudness	mode	speechiness	acousticness	instrumer
Can't Let Go								
Son Little - Mad About You	0.599	0.377	7	-9.829	1	0.1050	0.4410	0
Foy Vance - She Burns	0.730	0.403	4	-10.146	1	0.0483	0.7180	0
Jack Johnson - Good People	0.721	0.588	11	-6.018	1	0.0331	0.2190	0
Irene Skylakaki - Tonight	0.549	0.454	10	-7.785	1	0.0326	0.5270	0.

Task 2 - Song Similarity

In [50]:
```
!pip install -U scikit-learn
```

```
Requirement already satisfied: scikit-learn in /opt/miniconda3/envs/pml/lib/
python3.9/site-packages (1.2.2)
Requirement already satisfied: joblib>=1.1.1 in /opt/miniconda3/envs/pml/li
b/python3.9/site-packages (from scikit-learn) (1.2.0)
Requirement already satisfied: numpy>=1.17.3 in /opt/miniconda3/envs/pml/li
b/python3.9/site-packages (from scikit-learn) (1.23.5)
Requirement already satisfied: scipy>=1.3.2 in /opt/miniconda3/envs/pml/lib/
python3.9/site-packages (from scikit-learn) (1.10.1)
Requirement already satisfied: threadpoolctl>=2.0.0 in /opt/miniconda3/envs/
pml/lib/python3.9/site-packages (from scikit-learn) (3.1.0)
```

In [51]:
```
from sklearn.decomposition import PCA
import matplotlib.pyplot as plt
from sklearn.metrics.pairwise import cosine_similarity as cosine
from sklearn.preprocessing import StandardScaler
```

In [52]:
```
my_subset_features = ['danceability', 'energy', 'key',
            'loudness', 'mode', 'speechiness',
            'acousticness', 'instrumentalness',
            'liveness', 'valence', 'tempo', 'duration_ms',
            'time_signature']
```

In [90]:
```
#Get subset
subset_features = features[my_subset_features]
#Standardise
scaled_features = StandardScaler().fit_transform(subset_features)
#Get cosine distances
similarities = cosine(scaled_features)
#Visualise
similarities = pd.DataFrame(similarities, columns = features.index, index =
similarities.style.background_gradient(cmap='Greens')
```

Out[90]:

	Foy Vance - Make It Rain - Live	Michael Kiwanuka - Love & Hate	Jack Johnson - Good People	Son Little - Mad About You	Foy Vance - She Burns	KALEO - Way down We Go - Recorded at Spotify Studios NYC	Leon Bridges - Coming Home	
Foy Vance - Make It Rain - Live	1.000000	0.353336	-0.447927	0.211017	-0.196242	-0.037723	0.165624	
Michael Kiwanuka - Love & Hate	0.353336	1.000000	-0.480976	-0.135053	-0.200916	-0.372784	0.154782	-
Jack Johnson - Good People	-0.447927	-0.480976	1.000000	-0.263814	0.131830	-0.331941	0.028010	-
Son Little - Mad About You	0.211017	-0.135053	-0.263814	1.000000	0.401505	0.020201	-0.133272	
Foy Vance - She Burns	-0.196242	-0.200916	0.131830	0.401505	1.000000	0.182974	0.081430	
KALEO - Way down We Go - Recorded at Spotify Studios NYC	-0.037723	-0.372784	-0.331941	0.020201	0.182974	1.000000	0.068731	
Leon Bridges - Coming Home	0.165624	0.154782	0.028010	-0.133272	0.081430	0.068731	1.000000	
Michael Kiwanuka - Home Again	0.196324	-0.142596	-0.337017	0.202878	0.459199	0.788646	0.099042	
Michael Kiwanuka - One More Night	-0.419857	-0.218802	0.407509	-0.536893	-0.546651	-0.275331	-0.317333	-
Robert Plant - Can't Let Go	-0.216641	-0.382891	0.123922	0.240372	-0.236773	-0.173143	-0.388234	-
Michael Kiwanuka - Cold Little Heart - Radio Edit	0.032405	-0.266967	0.026885	-0.484082	-0.363601	0.138773	-0.142949	-
The Teskey Brothers -	-0.308971	-0.249589	0.110258	-0.485010	-0.284014	-0.077428	-0.361815	-

	Foy Vance - Make It Rain - Live	Michael Kiwanuka - Love & Hate	Jack Johnson - Good People	Son Little - Mad About You	Foy Vance - She Burns	KALEO - Way down We Go - Recorded at Spotify Studios NYC	Leon Bridges - Coming Home
So Caught Up							
David Gray - Sail Away	0.276958	0.226006	-0.210348	-0.344963	-0.042966	-0.120364	-0.162261
Irene Skylakaki - Tonight	-0.270268	-0.171548	0.099026	0.222327	0.287947	0.440898	0.294361
Ed Sheeran - Thinking out Loud	-0.281750	0.151132	-0.190130	-0.072838	0.312450	0.043978	-0.032877
Sam Smith - I'm Not The Only One	-0.427900	-0.001168	-0.100596	0.097482	0.363954	0.331654	0.169252
Sia - Elastic Heart	-0.079916	0.088039	0.429484	-0.235865	-0.506180	-0.477075	0.109441
Sia - Chandelier	-0.477857	0.286834	0.138891	-0.174028	-0.318613	-0.357831	-0.289608

Based on the DataFrame that was created, the songs that appear to be mostly similar (based on all features) are "KALEO - Way Down We Go - Recorded at Spotify Studios NYC" (https://www.youtube.com/watch?v=0-7IHOXkiV8) and "Home Again" by Michael Kiwanuka (https://www.youtube.com/watch?v=kJ4s3G7hgR4). These songs share several similarities, such as a slow and gentle tempo, and both are performed by male artists. However, the Kaleo song does have some parts that exhibit higher intensity. With a similarity score of 0.788646, the similarity between these songs is quite high. Looking also at the sorting table above, both songs seem to have same mode and very similar speechiness.

I then re-run the code, changing the features that I thought that were similar in the 2 songs (tempo, danceability,liveness and valence) and was proven correct with the same songs have a very high similarity of 0.963927. Some other songs also appeared to be quite similar (similarity of 0.899831) in those features such as Sia's 'Elastic Heart' and Leon Bridges with 'Coming Home', which was quite interesting as these songs I wouldnt think that are similar. In the previous, all features DataFrame, their similarity was very low - 0.109.

In [94]:
```
my_subset_features2 = ['danceability',
            'liveness', 'valence', 'tempo',
          ]
```

In [95]:
```
subset_features = features[my_subset_features2]
scaled_features = StandardScaler().fit_transform(subset_features)
similarities = cosine(scaled_features)
similarities = pd.DataFrame(similarities, columns = features.index, index =
similarities.style.background_gradient(cmap='Greens')
```

Out[95]:

	Foy Vance - Make It Rain - Live	Michael Kiwanuka - Love & Hate	Jack Johnson - Good People	Son Little - Mad About You	Foy Vance - She Burns	KALEO - Way down We Go - Recorded at Spotify Studios NYC	Leon Bridges - Coming Home
Foy Vance - Make It Rain - Live	1.000000	0.664952	-0.713358	-0.115373	-0.704466	0.318696	0.452181
Michael Kiwanuka - Love & Hate	0.664952	1.000000	-0.637676	0.251300	-0.812866	0.235280	0.746316
Jack Johnson - Good People	-0.713358	-0.637676	1.000000	-0.606311	0.667204	-0.743041	-0.141082
Son Little - Mad About You	-0.115373	0.251300	-0.606311	1.000000	-0.184698	0.680230	-0.187816
Foy Vance - She Burns	-0.704466	-0.812866	0.667204	-0.184698	1.000000	-0.015001	-0.299766
KALEO - Way down We Go - Recorded at Spotify Studios NYC	0.318696	0.235280	-0.743041	0.680230	-0.015001	1.000000	0.067232
Leon Bridges - Coming Home	0.452181	0.746316	-0.141082	-0.187816	-0.299766	0.067232	1.000000
Michael Kiwanuka - Home Again	0.373298	0.366974	-0.690092	0.575735	-0.019160	0.963927	0.319545
Michael Kiwanuka - One More Night	-0.779402	-0.511378	0.621543	-0.021811	0.226334	-0.671707	-0.589329
Robert Plant - Can't Let Go	-0.758063	-0.454535	0.315600	0.374861	0.172389	-0.314767	-0.717125
Michael Kiwanuka - Cold Little Heart - Radio Edit	0.461298	-0.130059	-0.013001	-0.586509	-0.349067	-0.425615	-0.291411
The Teskey Brothers - So Caught Up	-0.227453	-0.503584	0.261911	-0.211345	-0.075003	-0.560083	-0.786347

	Foy Vance - Make It Rain - Live	Michael Kiwanuka - Love & Hate	Jack Johnson - Good People	Son Little - Mad About You	Foy Vance - She Burns	KALEO - Way down We Go - Recorded at Spotify Studios NYC	Leon Bridges - Coming Home
David Gray - Sail Away	0.455306	-0.234483	0.050521	-0.642983	0.260010	0.101832	0.122413
Irene Skylakaki - Tonight	-0.248867	0.401691	-0.387631	0.897594	-0.208066	0.428148	0.081147
Ed Sheeran - Thinking out Loud	-0.287560	-0.544006	-0.213833	0.512279	0.196071	0.335824	-0.911147
Sam Smith - I'm Not The Only One	-0.400121	-0.187028	-0.307248	0.866783	0.316416	0.703880	-0.372519
Sia - Elastic Heart	0.176147	0.433033	0.296860	-0.504957	-0.012502	-0.276673	0.899831
Sia - Chandelier	-0.078983	0.545482	0.183408	-0.027368	-0.066694	-0.132514	0.834323

Task 3 - Top N Recommendations

I ran the code for the song "Good People" by Jack Johnson.

Output: Index(['Sia - Elastic Heart', 'Michael Kiwanuka - One More Night', 'Sia - Chandelier', 'Foy Vance - She Burns', 'Robert Plant - Can't Let Go', 'The Teskey Brothers - So Caught Up', 'Irene Skylakaki - Tonight', 'Leon Bridges - Coming Home', 'Michael Kiwanuka - Cold Little Heart - Radio Edit', 'Sam Smith - I'm Not The Only One'], dtype='object')

The results were quite surprising, especially for the first song "Elastic Heart" by Sia. The second and fourth songs, "One More Night" by Michael Kiwanuka and "She Burns" by Foy Vance, appeared to be more similar in terms of tone, feeling, and tempo. Around the fifth song, the songs started to deviate significantly. Hence, there was no point in re-running the code with n < 10, as the songs had already deviated. However, it should be noted that this is a small playlist.

I then re-ran the code for a smaller subset of features, specifically ['danceability', 'liveness', 'valence', 'tempo']. In this case, the results (see below) were generally closer in terms of similarity. The first recommendation is "She Burns" by Foy Vance, which likely has a similar tempo and feeling as both songs are performed by male artists. Similarly, "One More Night" by Michael Kiwanuka is also suggested, which may be slightly more upbeat. However, the third suggested song appears to deviate more compared to the previous two, and overall, the differences between the songs seem to be greater than their similarities.

In [99]:
```
track = "Jack Johnson - Good People"
```

In [100…
```
#How many tracks to recommend
n = 10
```

In [101…
```
similarities.sort_values(by = track, ascending=False)[track].index[1:n+1]
```

Out[101]:
```
Index(['Foy Vance - She Burns', 'Michael Kiwanuka - One More Night',
       'Robert Plant - Can't Let Go', 'Sia - Elastic Heart',
       'The Teskey Brothers - So Caught Up', 'Sia - Chandelier',
       'David Gray - Sail Away',
       'Michael Kiwanuka - Cold Little Heart - Radio Edit',
       'Leon Bridges - Coming Home', 'Ed Sheeran - Thinking out Loud'],
      dtype='object')
```

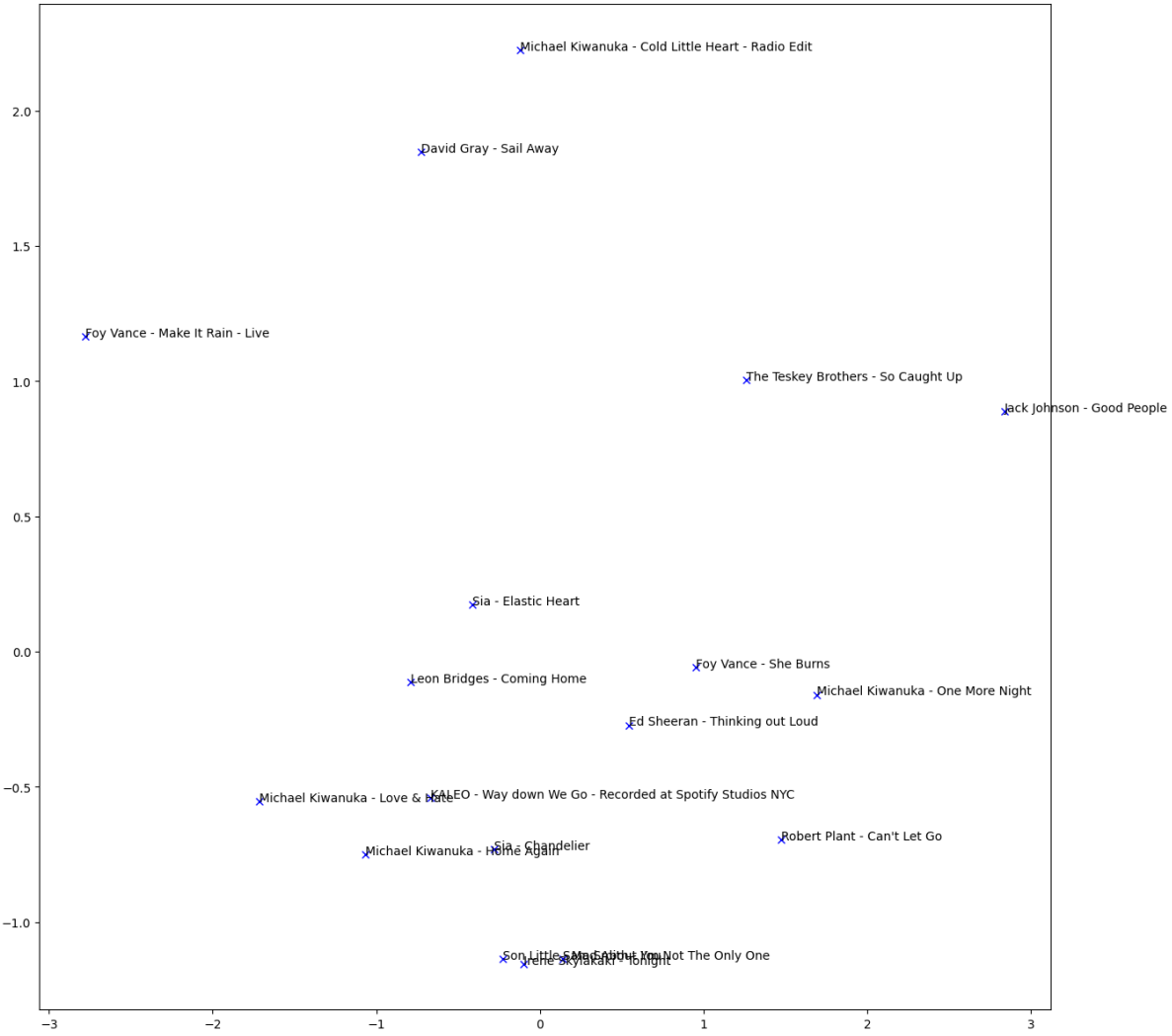
Task 4 - PCA Visualisation

In [104…
```
num_dimensions = 2

#Reduce dimensions
pca = PCA(n_components=num_dimensions)
reduced_features = pca.fit_transform(scaled_features)
```

In [103…
```
#Plot
plt.figure(figsize=(8, 8))
fig,ax = plt.subplots(figsize=(15,15))
x = reduced_features[:,0]
y = reduced_features[:,1]
ax.plot(x,y,"bx")
annotations = subset_features.index
for i, label in enumerate(annotations):
    ax.annotate(label, (x[i], y[i]),alpha=1)
```

```
<Figure size 800x800 with 0 Axes>
```



Thank you for reading through - hope it was not too confusing. :)

In [ ]: