Determining ‘Descriptive Genres’ for Songs

By

Naina Jain

Pranav Deshpande

Vishv Brahmbhatt

RUTGERS
THE STATE UNIVERSITY
OF NEW JERSEY

Introduction

- Genre classification is too mainstream
- Determining descriptive genres through clustering
- Analysing the similarity between various tracks based on the audio features

Dataset

- Tracks.csv - The main dataset that contains all the information like date, duration, id, tags, name, composers, genres and so on for album, artist and tracks.
- Genres.csv – Contains details about genres of respective tracks along with the subgenres
- Echonest.csv – Contains information about the audio features like acousticness, danceability, energy, valence and so on along with other information like date, artist, title, etc. of any particular song

Data Pre-Processing

- tracks.csv
- Dropping extra columns, converting year to readable format, replacing missing values

album id	album informatic	album listens	album producer	album tags	album title	album tracks	album type	artist active_yea	artist active_yea	artist associated	artist bio
1		6073		[]	AWOL - A	7	Album	########			<p>A Way
1		6073		[]	AWOL - A	7	Album	########			<p>A Way
1		6073		[]	AWOL - A	7	Album	########			<p>A Way
6		47632		[]	Constant H	2	Album			Mexican S	<p><span
4	<p>Â "spiri	2710		[]	Niris	13	Album	########	########		<p>Songs
4	<p>Â "spiri	2710		[]	Niris	13	Album	########	########		<p>Songs
4	<p>Â "spiri	2710		[]	Niris	13	Album	########	########		<p>Songs
4	<p>Â "spiri	2710		[]	Niris	13	Album	########	########		<p>Songs
4	<p>Â "spiri	2710		[]	Niris	13	Album	########	########		<p>Songs
1		6073		[]	AWOL - A	7	Album	########			<p>A Way
58	<p>A	3331		[]	mp3	4	Single Tracks				
58	<p>A	3331		[]	mp3	4	Single Tracks				
59	<p>Here's	1681		['lafms']	Live at LAC	2	Live Perfor	########	########	Los Angele	<p>Airwa
59	<p>Here's	1681		['lafms']	Live at LAC	2	Live Perfor	########	########	Los Angele	<p>Airwa
60	<p>A full e	1304		[]	Every Man	2	Album	########			<p>The Ey
61	<p>Alec K.	1300	Alec K. Ref	[]	The Blind S	1	Album	########			<p>The Ey
60	<p>A full e	1304		[]	Every Man	2	Album	########			<p>The Ey
62	<p>Recor	845		[]	The Quiet	1	Album	########			<p>The Ey
64	<p><em>A	2014	Tom Buckl	[]	Amoebiasi	0	Album	########	########		<p>The
64	<p><em>A	2014	Tom Buckl	[]	Amoebiasi	0	Album	########	########		<p>The
65	<p>"Excell	1446		[]	Limbic Rag	0	Album	########	########		<p>The
65	<p>"Excell	1446		[]	Limbic Rag	0	Album	########	########		<p>The

Data Pre-Processing

- Tracks.csv

```
tracks.head()
```

track_id	year_released	album	album_id	artist_id	artist	genres_all	track	duration	genre_top
2	2009	AWOL - A Way Of Life	1	1	AWOL	[21]	Food	168	Hip-Hop
3	2009	AWOL - A Way Of Life	1	1	AWOL	[21]	Electric Ave	237	Hip-Hop
5	2009	AWOL - A Way Of Life	1	1	AWOL	[21]	This World	206	Hip-Hop
10	2008	Constant Hitmaker	6	6	Kurt Vile	[10]	Freeway	161	Pop
20	2009	Niris	4	4	Nicky Cook	[17, 10, 76, 103]	Spiritual Level	311	UNKNOWN

Data Pre-Processing

- echonest.csv
- Extracting only the audio features

	echonest	echonest	echonest	echonest	echonest	echonest	echonest	echonest	echonest	echonest	echonest	echonest	echonest
	audio_fea	audio_fea	audio_fea	audio_fea	audio_fea	audio_fea	audio_fea	audio_fea	metadata	metadata	metadata	metadata	metadata
	acousticne	danceabili	energy	instrument	liveness	speechine	tempo	valence	album_dat	album_nar	artist_latit	artist_loca	artist_long
track_id													
2	0.416675	0.675894	0.634476	0.010628	0.177647	0.15931	165.922	0.576661			32.6783	Georgia, U	-83.223
3	0.374408	0.528643	0.817461	0.001851	0.10588	0.461818	126.957	0.26924			32.6783	Georgia, U	-83.223
5	0.043567	0.745566	0.70147	0.000697	0.373143	0.124595	100.26	0.621661			32.6783	Georgia, U	-83.223
10	0.95167	0.658179	0.924525	0.965427	0.115474	0.032985	111.562	0.96359	########	Constant H	39.9523	Philadelph	-75.1624
134	0.452217	0.513238	0.56041	0.019443	0.096567	0.525519	114.29	0.894072			32.6783	Georgia, U	-83.223
139	0.10655	0.260911	0.607067	0.835087	0.223676	0.030569	196.961	0.160267			41.8239	Providence	-71.412
140	0.376312	0.734079	0.265685	0.669581	0.085995	0.039068	107.952	0.609991			41.8239	Providence	-71.412
141	0.963657	0.435933	0.075632	0.345493	0.105686	0.026658	33.477	0.16395			41.8239	Providence	-71.412
142	0.662881	0.379065	0.823856	0.910266	0.088705	0.07909	147.781	0.092868	2005	The Quiet	41.8239	Providence	-71.412
144	0.909011	0.443643	0.641997	0.924092	0.267669	0.089659	128.537	0.788251			41.8239	Providence	-71.412
145	0.235506	0.438672	0.487752	0.716122	0.070359	0.047298	120.79	0.650452			41.8239	Providence	-71.412
146	0.532019	0.417681	0.476422	0.4025	0.172105	0.035361	135.468	0.682397			41.8239	Providence	-71.412
147	0.77841	0.706681	0.866116	0.806703	0.10465	0.065083	120.218	0.917613			41.8239	Providence	-71.412
153	0.988306	0.255661	0.979774	0.973006	0.121342	0.05174	90.241	0.034018					
154	0.970135	0.352946	0.023852	0.957113	0.113261	0.032177	53.758	0.035632					
155	0.981657	0.142249	0.912122	0.967294	0.36351	0.087527	91.912	0.034325					
169	0.989141	0.225978	0.722835	0.263076	0.092371	0.053406	94.322	0.028347					
170	0.88666	0.298518	0.744333	0.92095	0.139587	0.088781	97.88	0.073548					
171	0.698278	0.285816	0.213494	0.955691	0.087036	0.064094	125.645	0.150599					
172	0.815549	0.144125	0.892721	0.90043	0.104703	0.102294	138.68	0.034916					
173	0.842113	0.285293	0.564689	0.951624	0.110481	0.040611	166.552	0.254299					

Data Pre-Processing

```
echonest_audio_features.head()
```

track_id	acousticness	danceability	energy	instrumentalness	liveness	speechiness	tempo	valence
2	0.416675	0.675894	0.634476	0.010628	0.177647	0.159310	165.922	0.576661
3	0.374408	0.528643	0.817461	0.001851	0.105880	0.461818	126.957	0.269240
5	0.043567	0.745566	0.701470	0.000697	0.373143	0.124595	100.260	0.621661
10	0.951670	0.658179	0.924525	0.965427	0.115474	0.032985	111.562	0.963590
134	0.452217	0.513238	0.560410	0.019443	0.096567	0.525519	114.290	0.894072

Data Pre-Processing

- Tracks_cleaned.csv

```
df.head()
```

track_id	acousticness	danceability	energy	instrumentalness	liveness	speechiness	tempo	valence	year_released	album	album_id	artist_id
2	0.416675	0.675894	0.634476	0.010628	0.177647	0.159310	165.922	0.576661	2009	AWOL - A Way Of Life	1	1
3	0.374408	0.528643	0.817461	0.001851	0.105880	0.461818	126.957	0.269240	2009	AWOL - A Way Of Life	1	1
5	0.043567	0.745566	0.701470	0.000697	0.373143	0.124595	100.260	0.621661	2009	AWOL - A Way Of Life	1	1
10	0.951670	0.658179	0.924525	0.965427	0.115474	0.032985	111.562	0.963590	2008	Constant Hitmaker	6	6
134	0.452217	0.513238	0.560410	0.019443	0.096567	0.525519	114.290	0.894072	2009	AWOL - A Way Of Life	1	1

Data Normalisation

- Normalising the tempo feature which is exceptionally high

```
min_tempo=np.min(df.tempo)
max_tempo=np.max(df.tempo)
```

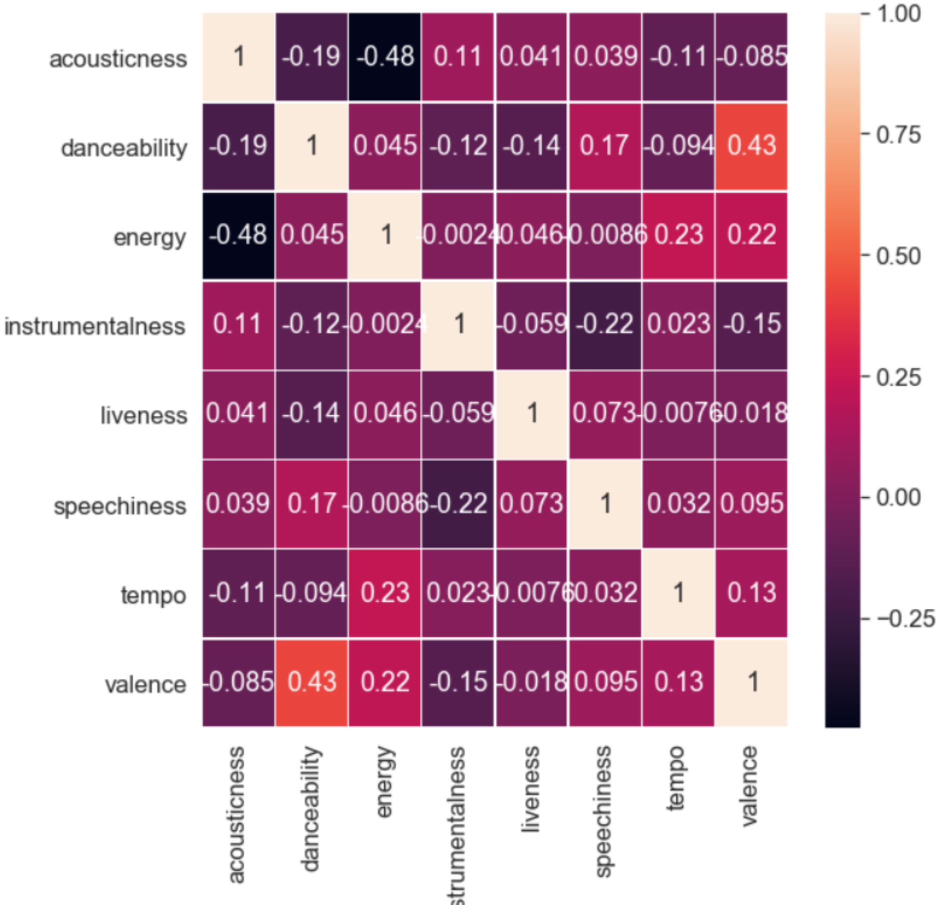
```
df['tempo']=(df.tempo-min_tempo)/(max_tempo-min_tempo)
```

```
audio_cols = ['acousticness', 'danceability', 'energy', 'instrumentalness', 'liveness', 'speechiness', 'tempo', 'valence']
```

```
audio_features = df[audio_cols]
audio_features.head()
```

	acousticness	danceability	energy	instrumentalness	liveness	speechiness	tempo	valence
track_id								
2	0.416675	0.675894	0.634476	0.010628	0.177647	0.159310	0.642706	0.576661
3	0.374408	0.528643	0.817461	0.001851	0.105880	0.461818	0.479206	0.269240
5	0.043567	0.745566	0.701470	0.000697	0.373143	0.124595	0.367184	0.621661
10	0.951670	0.658179	0.924525	0.965427	0.115474	0.032985	0.414608	0.963590
134	0.452217	0.513238	0.560410	0.019443	0.096567	0.525519	0.426055	0.894072

Correlation



Machine learning models

- K-Means Algorithm
- Hierarchical clustering
- Principle component analysis (for multi-dimensional data)
- kNN

Why Clustering?

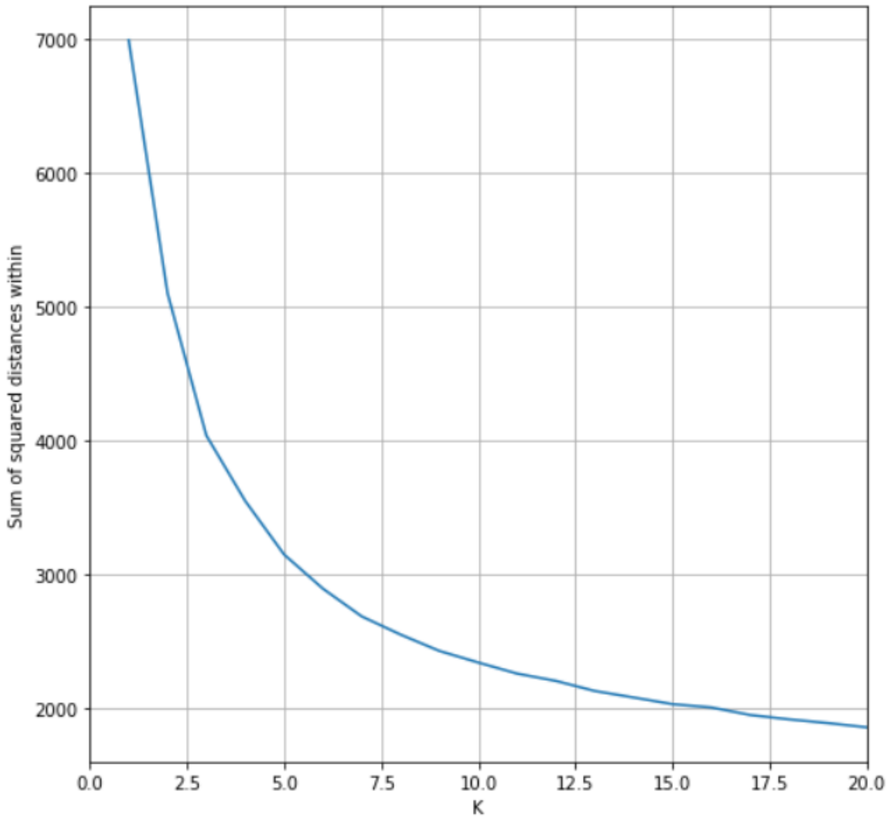
- Classification algorithms lead to genre classification
- Genre classification is subjective
- By clustering we can discover the specific similar properties among the tracks
- Based on those properties we can create or own labels for the clusters created

K-Means Algorithm

- Divides the data into pre-determined k clusters
- Points in the clusters are have maximum similarity
- All clusters vary from each other a lot

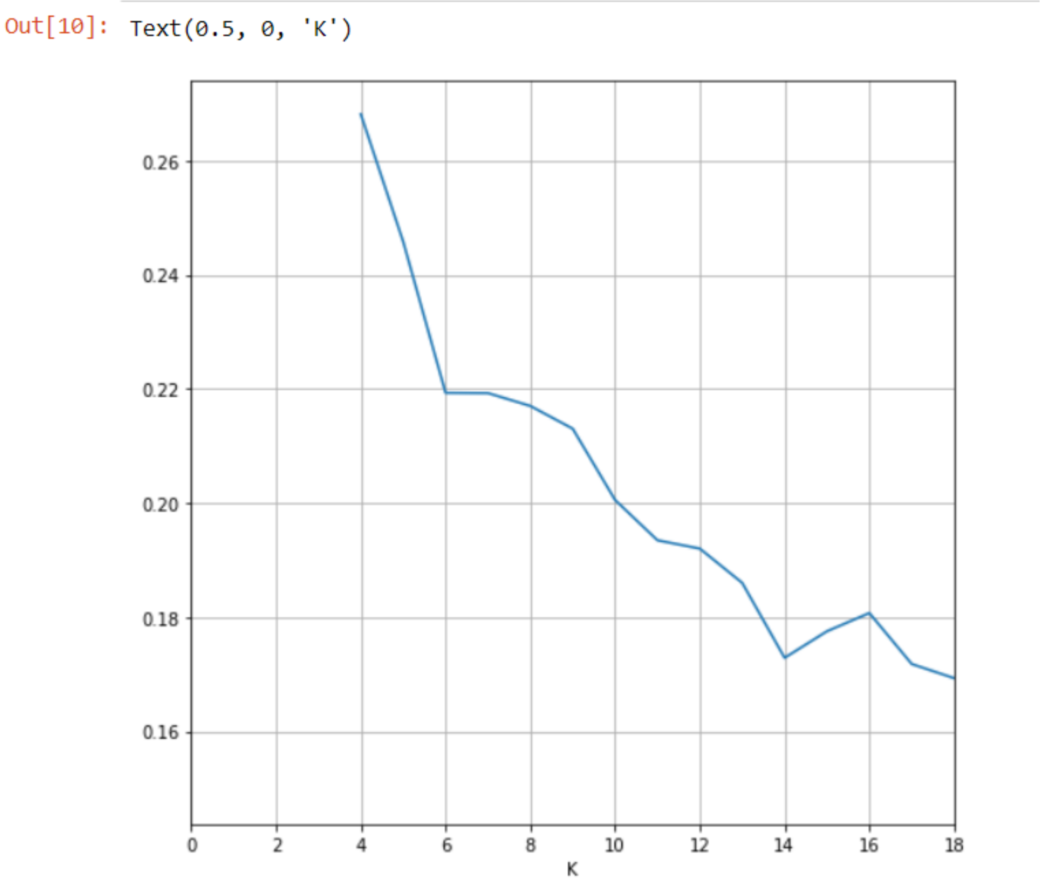
Deciding optimal K

- Sum of squared distance



Deciding optimal K

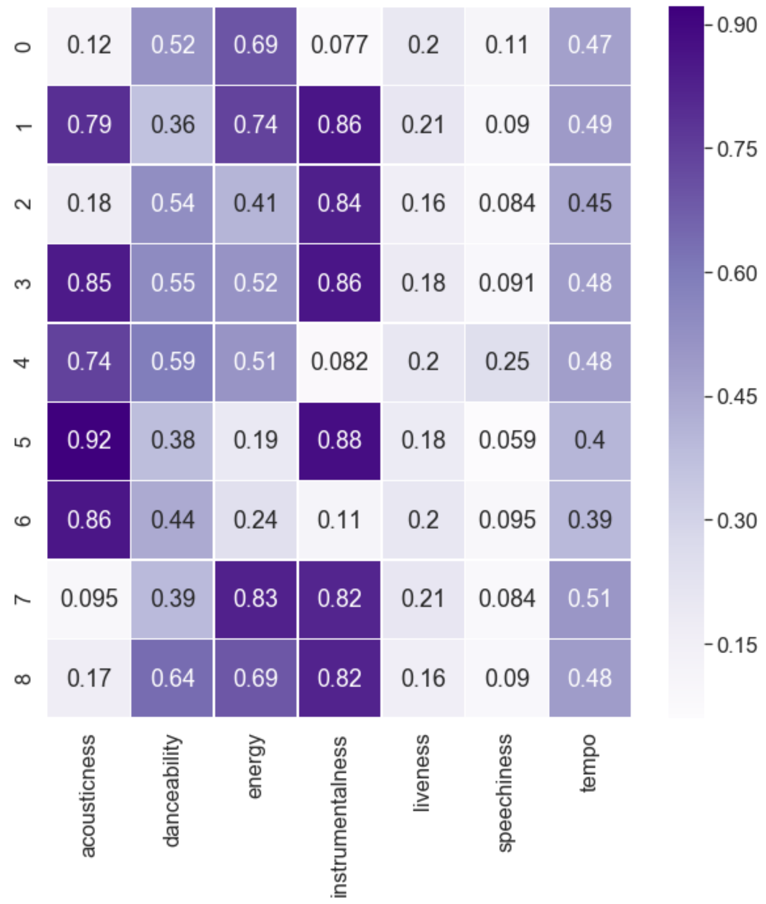
- Silhouette Score



Deciding optimal K

- The plot for within sum of squared distance can sometimes be ambiguous
- In such case we check for the silhouette score
- In our graph there is a sharp curve at 9
- Hence, k=9

K-Means Clustering



Creating Descriptive genres

KM0: Highly instrumental and danceable. "Upbeat danceable songs"

KM1: Highly acoustic, instrumental, and with low valence. "satanic songs"

KM2: Moderately danceable, energitic, and with moderate valence. "Uplifting calm songs"

KM3: Highly danceable, energitic, instrumental, and with high valence. "Lively empowering songs"

KM4: Highly acoustic with high valence and moderate speechiness. "Pleasant relaxing songs"

KM5: Highly energitic, instrumental, and moderate tempo. "Druggie songs"

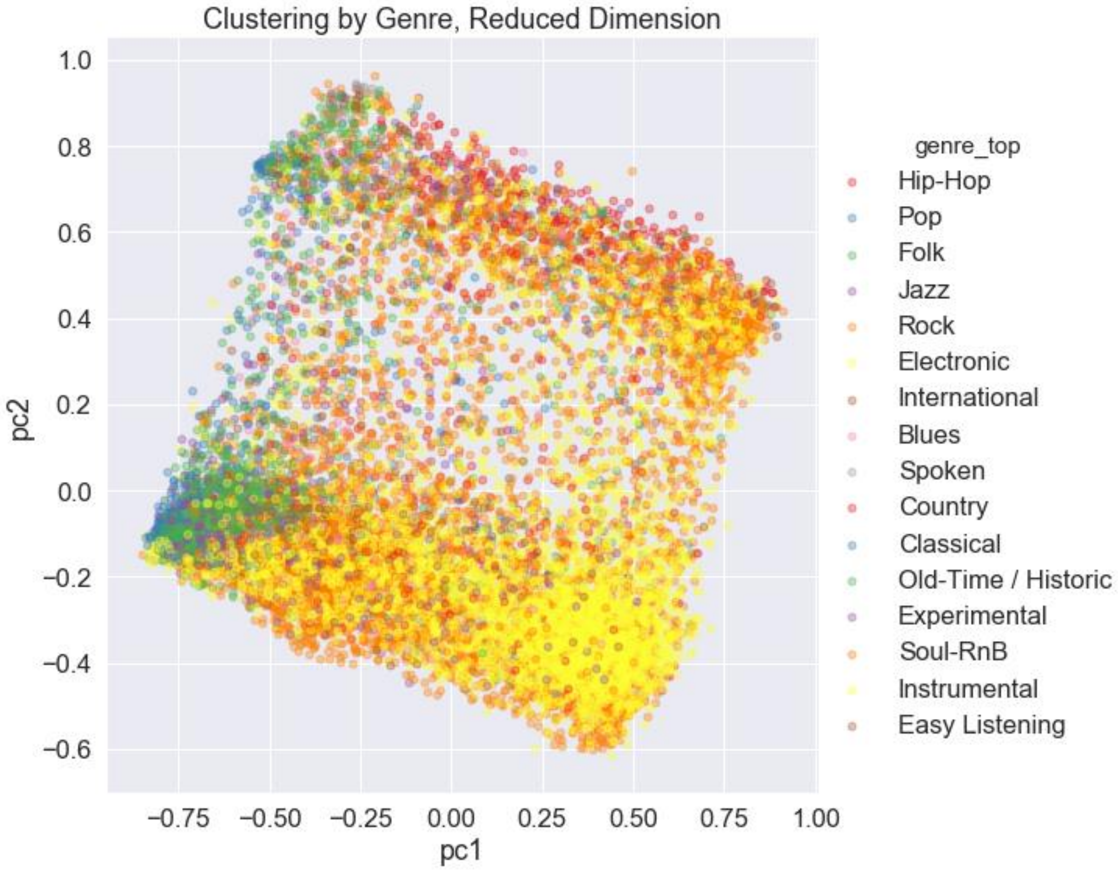
KM6: Highly acoustic, instrumental, and with high valence. "Enlighting happy songs"

KM7: Highly acoustic with low valence. "Dull sad songs"

KM8: Highly acoustic, energitic, instrumental with high tempo. "exciting fast songs"

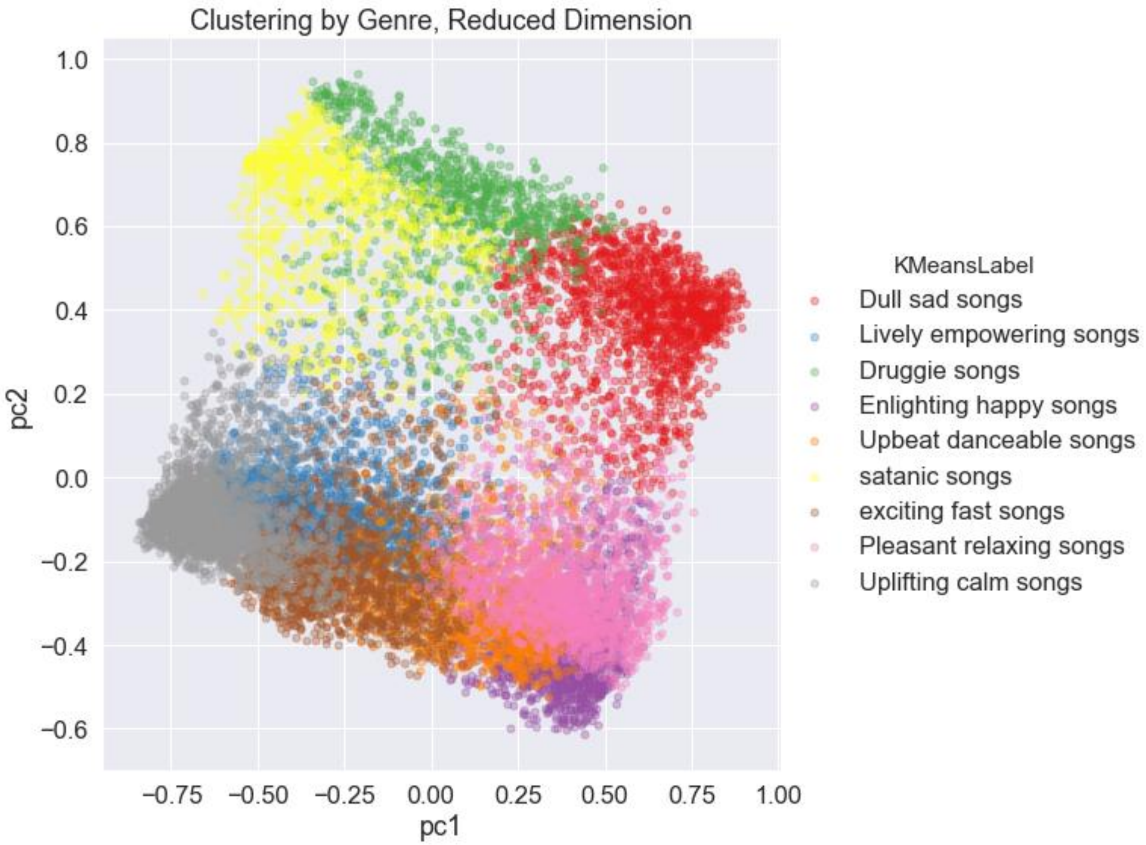
Principle component analysis

- Before Clustering



Principle component analysis

- After Clustering

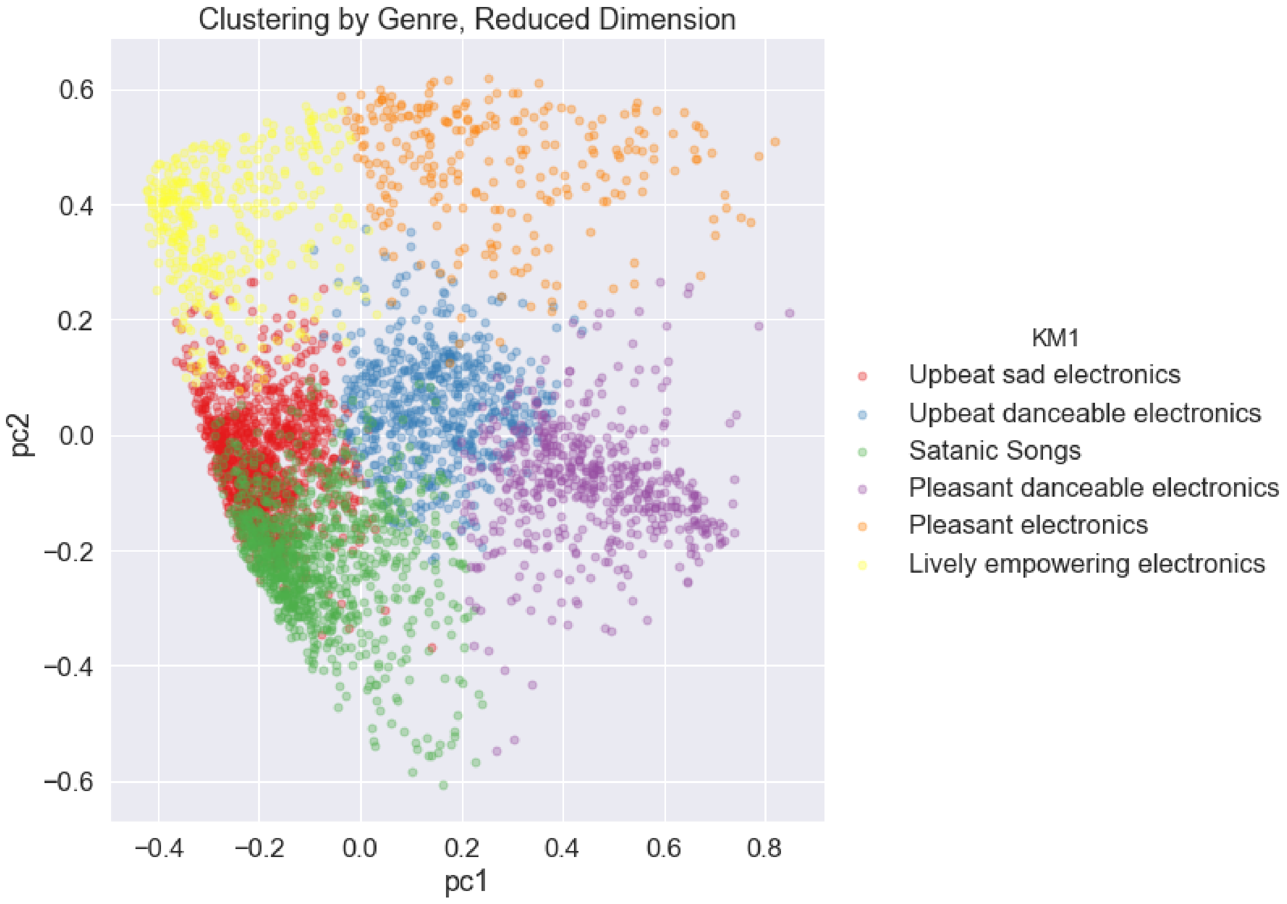


K-Means Clustering

- For “Electronic” songs
- By plotting the sum of squared distances and silhouette score, we get K=6

K-Means Clustering

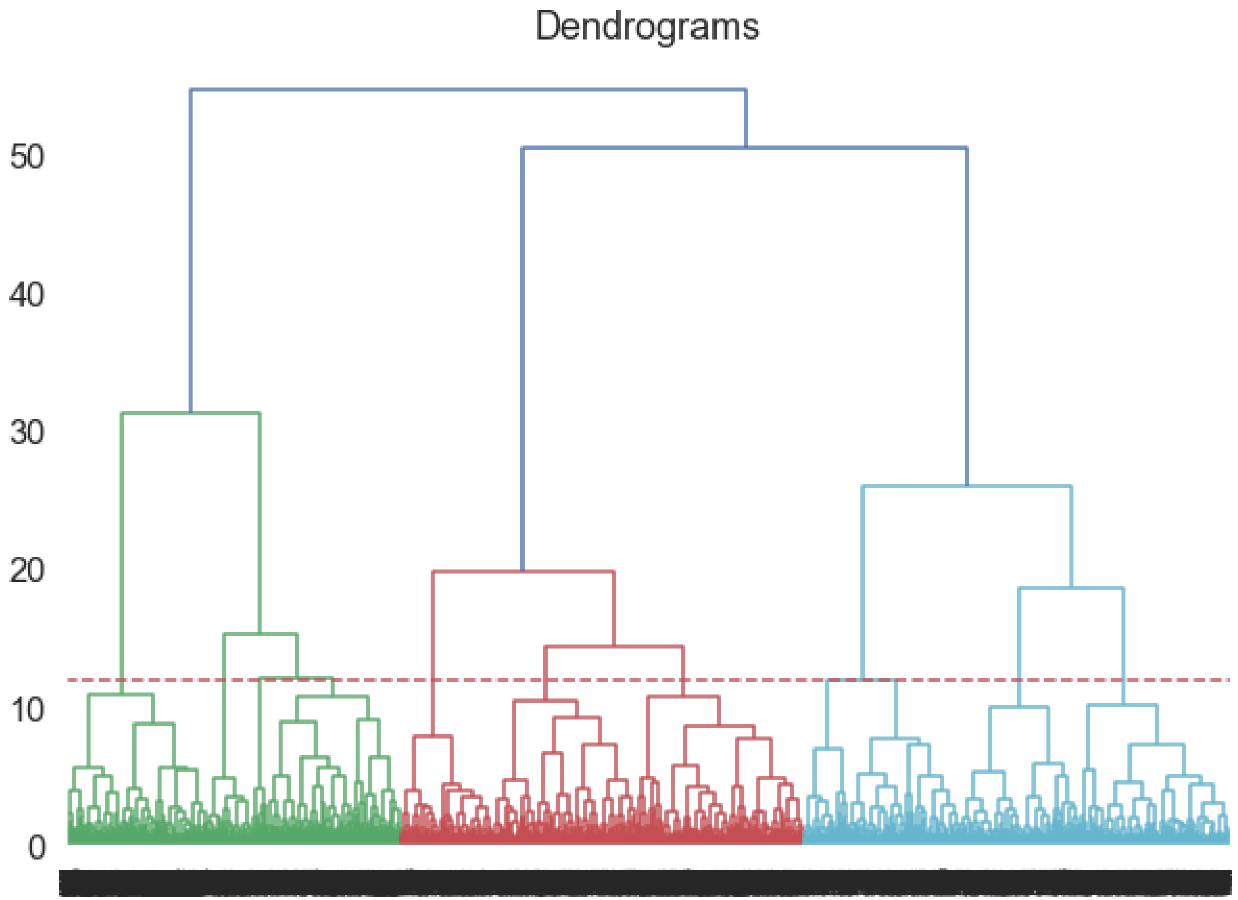
Principle Component Analysis



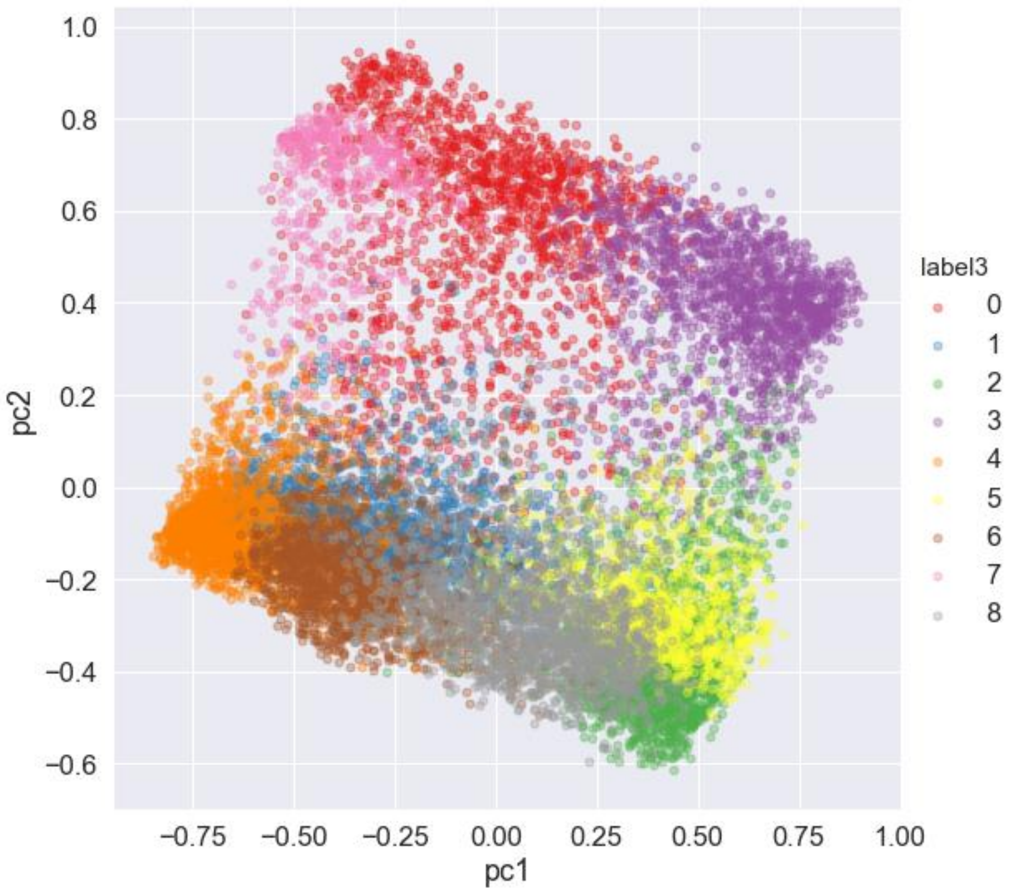
Hierarchical Clustering

- Alike K-Means, K is ‘not’ pre-determined
- Each point is a cluster initially
- Based on similarity, based on the distance, clusters are created and dendrograms are produced
- The number of clusters ‘k’ is decided based on the largest distance (longest vertical line) of the dendrogram

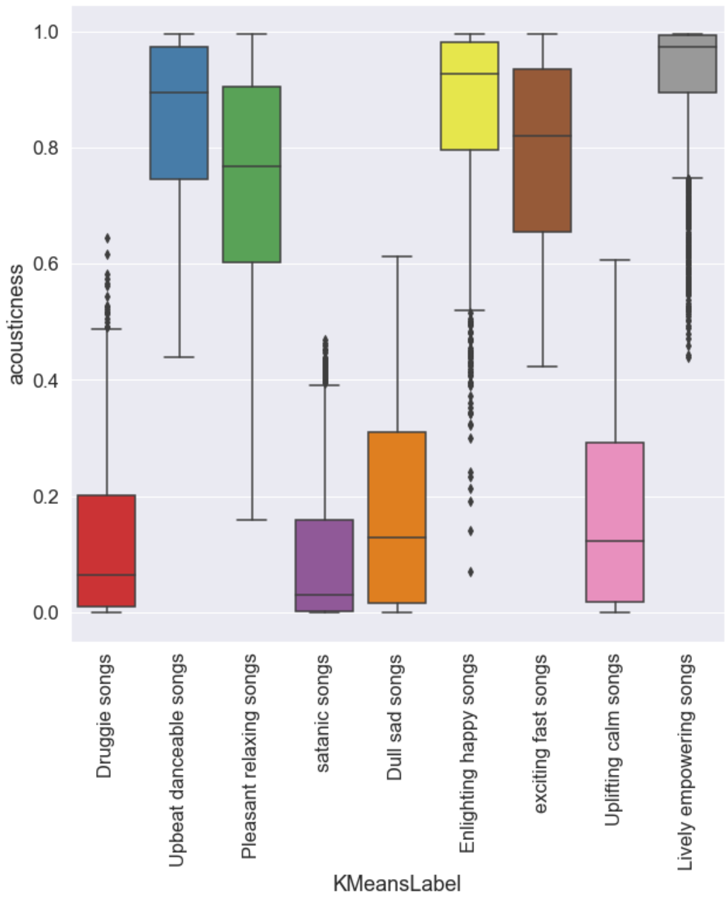
Dendrogram- Hierarchical Clustering

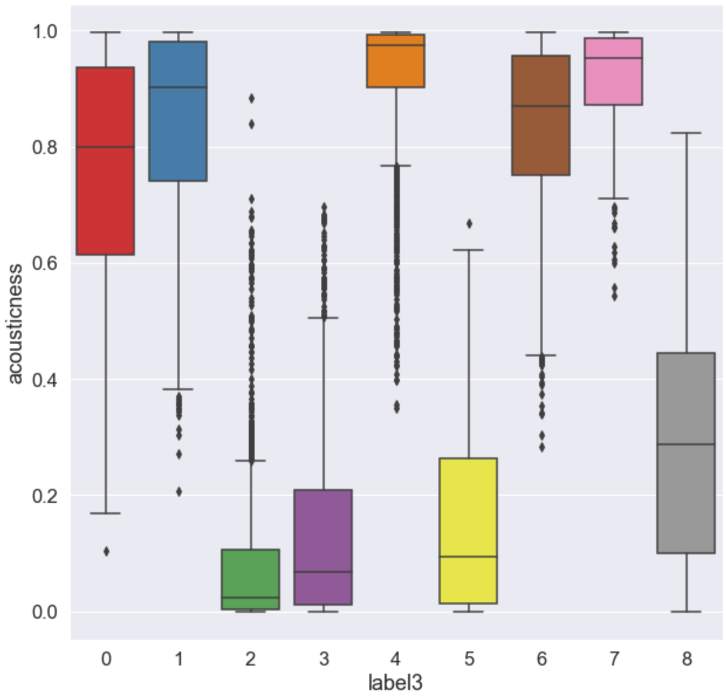


Visualisation-Hierarchical Clustering

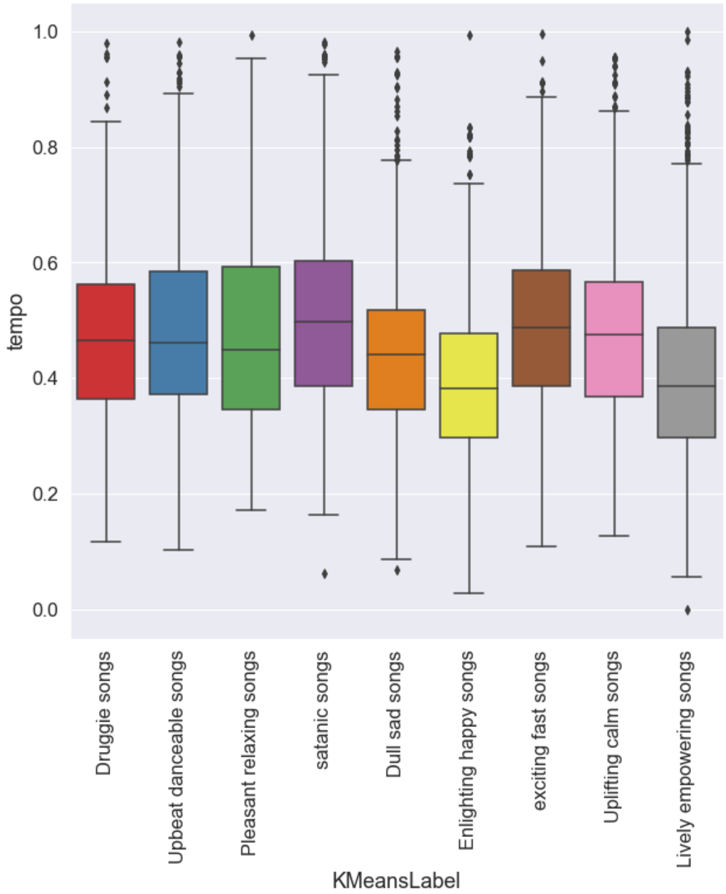


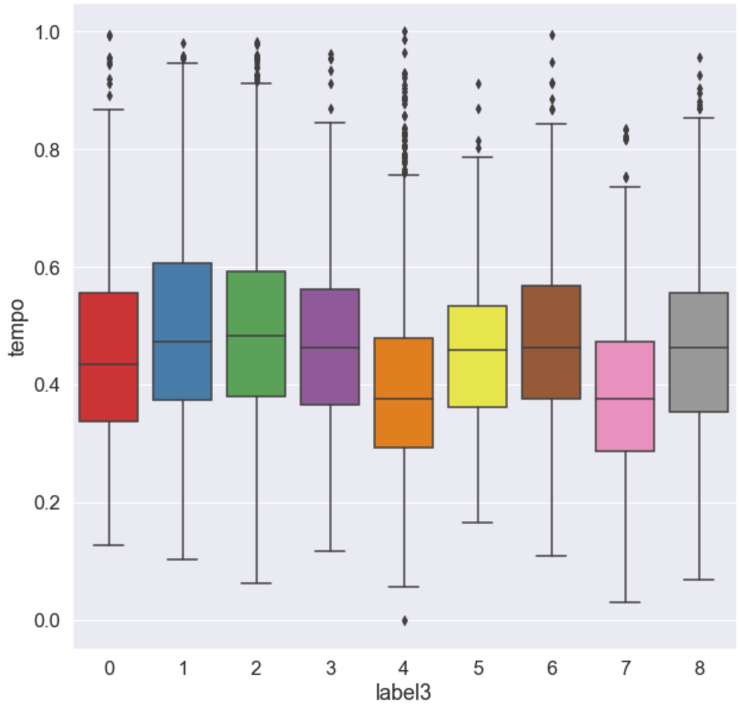
Model Comparison



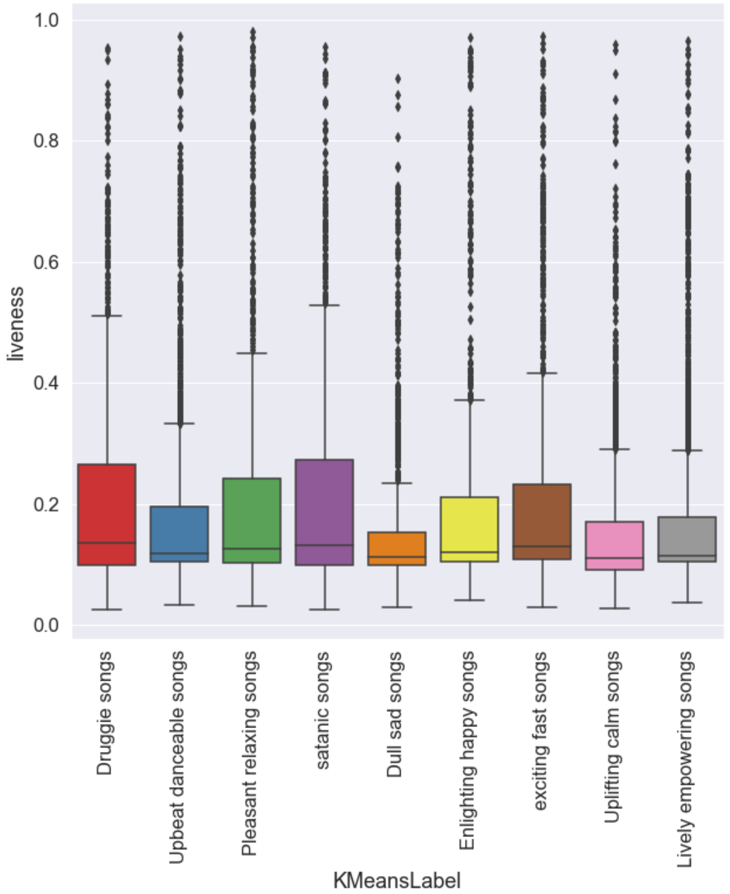


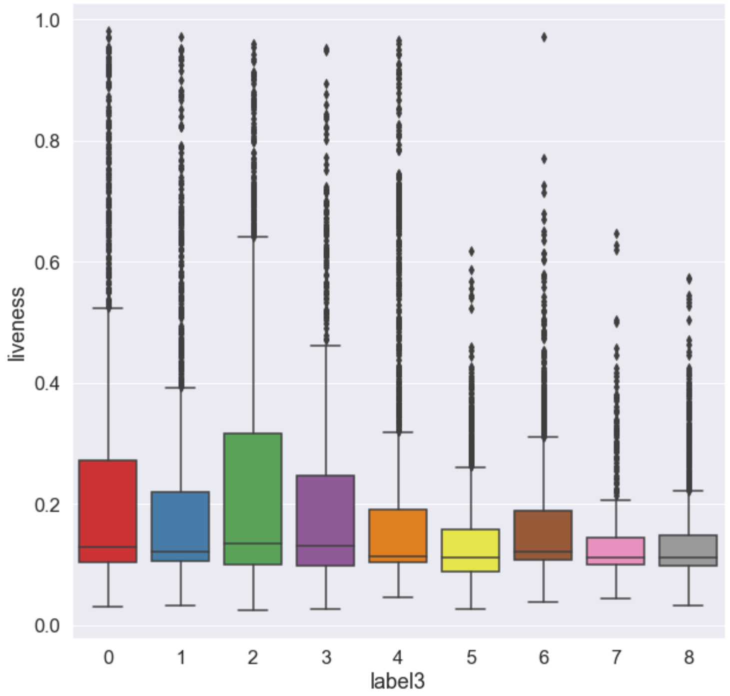
Model Comparison





Model Comparison





kNN Classification

- Classifies any new song into one of the above determined clusters

	precision	recall	f1-score	support
Druggie songs	0.95	0.94	0.94	307
Dull sad songs	0.91	0.92	0.92	268
Enlighting happy songs	0.93	0.96	0.94	339
Lively empowering songs	0.96	0.96	0.96	354
Pleasant relaxing songs	0.96	0.98	0.97	383
Upbeat danceable songs	0.92	0.91	0.92	273
Uplifting calm songs	0.98	0.91	0.94	193
exciting fast songs	0.92	0.96	0.94	207
satanic songs	0.96	0.92	0.94	302
micro avg	0.94	0.94	0.94	2626
macro avg	0.94	0.94	0.94	2626
weighted avg	0.94	0.94	0.94	2626

Conclusion

- Describing music in a new way
- Can be helpful for better music experience
- New songs successfully classified

Future scope

- Clustering on basis of sub genres
- A music recommendation system based on descriptive genres

Thank You