A Model-based Approach to Music Genre Assignment

A case study using technical attributes of music from Spotify

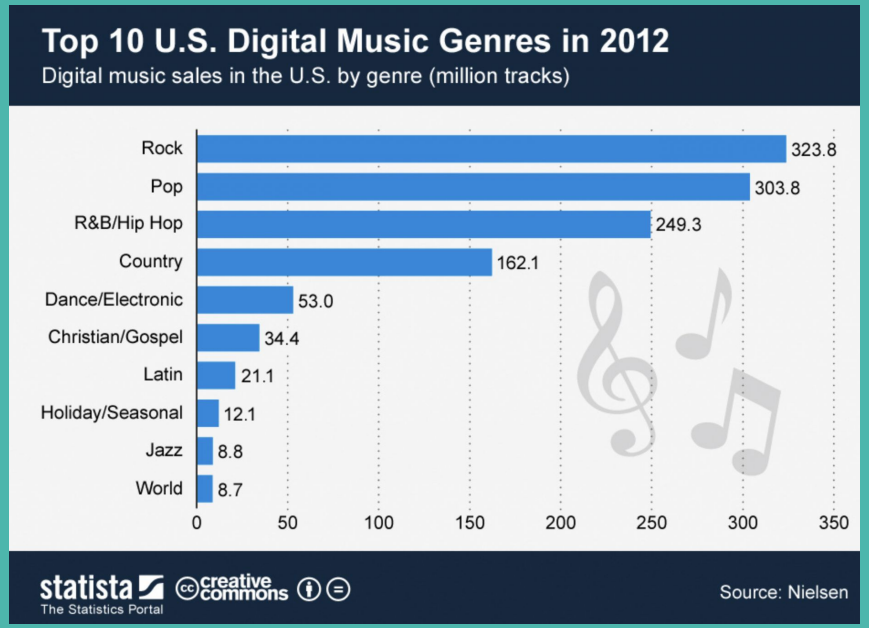
Preet Khowaja, John Owusu Duah, Rashaad Ratliff-Brown, Clarissa Ache

What is pop music?

What is Latin music?

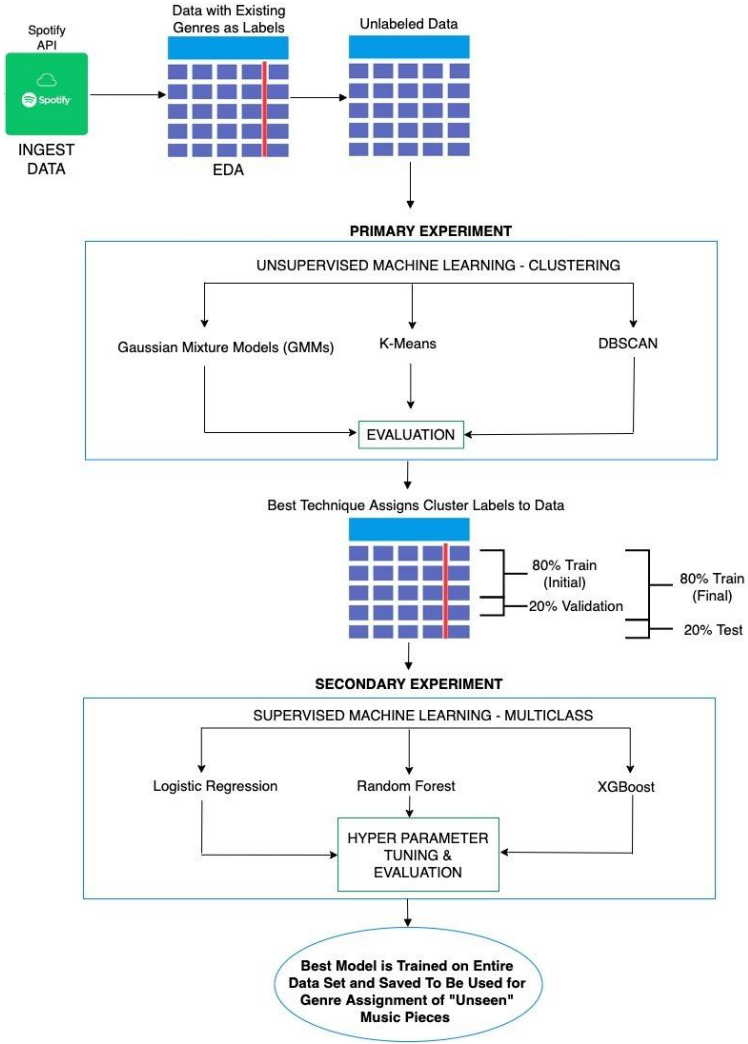
What is Bollywood music?

What problem are we trying to solve?



https://www.statista.com/chart/1783/album-sales-in-the-us-by-genre/

Process Flow



Spotify Data

- **12 features**: Direct measurements + Confidence Scores

Tempo	Loudness	Key	Mode	Speechiness	Liveness	Instrumental	Danceability	Energy	Acousticness	Valence	Duration

- Obtained from Spotify API by querying the top ~5000 songs from the top 6 most popular genres in 2015: **EDM, POP, ROCK, RAP, LATIN, and R&B**

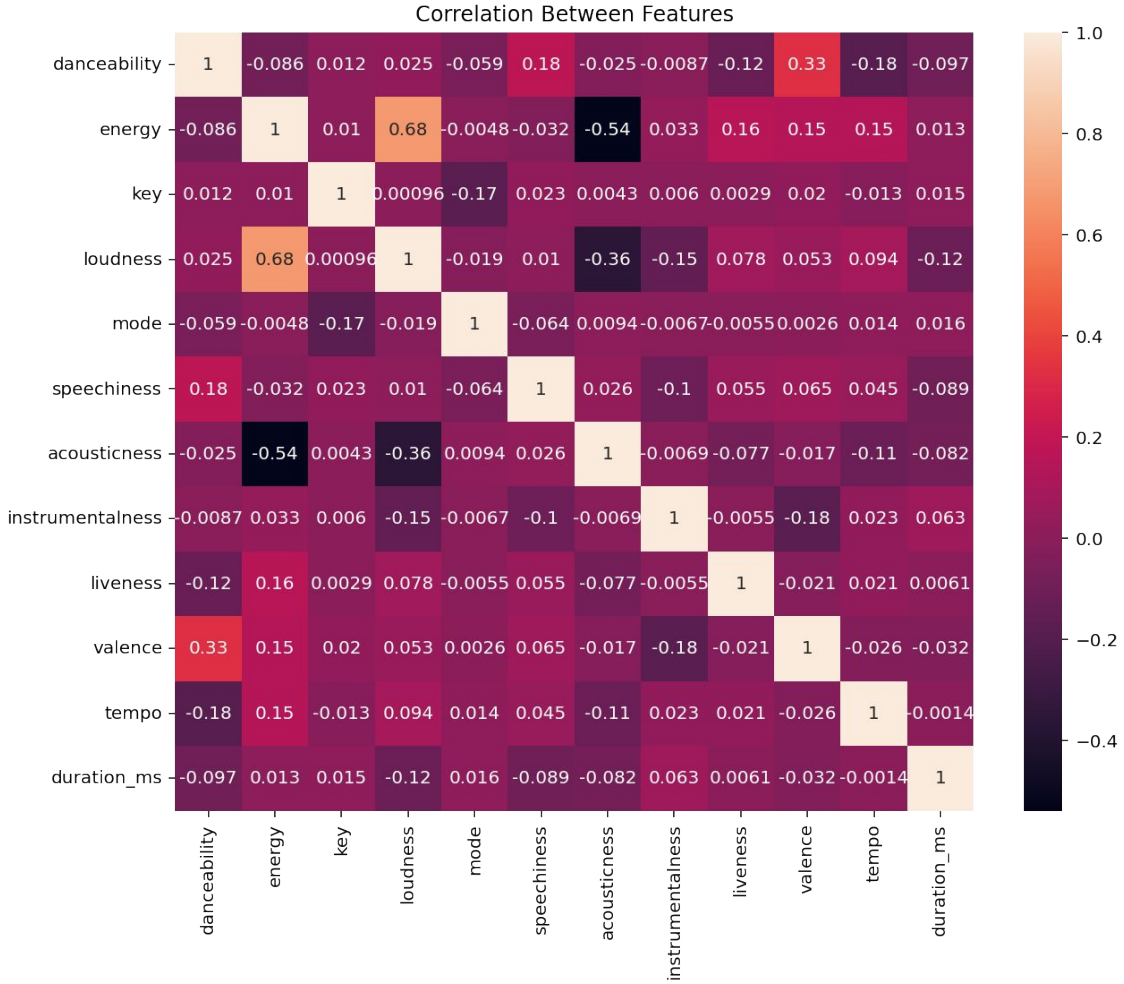
Example table:

track_name	track_artist	danceability	energy	key	loudness	mode	speechiness	acousticness	instrumentalness	liveness	valence	tempo	duration_ms
Lavish	WYLD	0.662	0.582	7	-7.937	1	0.1130	0.40400	0.000014	0.1920	0.599	99.904	204600
Work	Rihanna	0.725	0.534	11	-6.238	1	0.0946	0.07520	0.000000	0.0919	0.558	91.974	219320
Sleep Rider	K.D.S	0.819	0.825	3	-6.657	0	0.0654	0.14400	0.673000	0.0951	0.539	114.992	286500
Black Leather	Guns N' Roses	0.466	0.991	8	-3.272	1	0.1420	0.00551	0.174000	0.3540	0.205	137.595	248493

Spotify Data

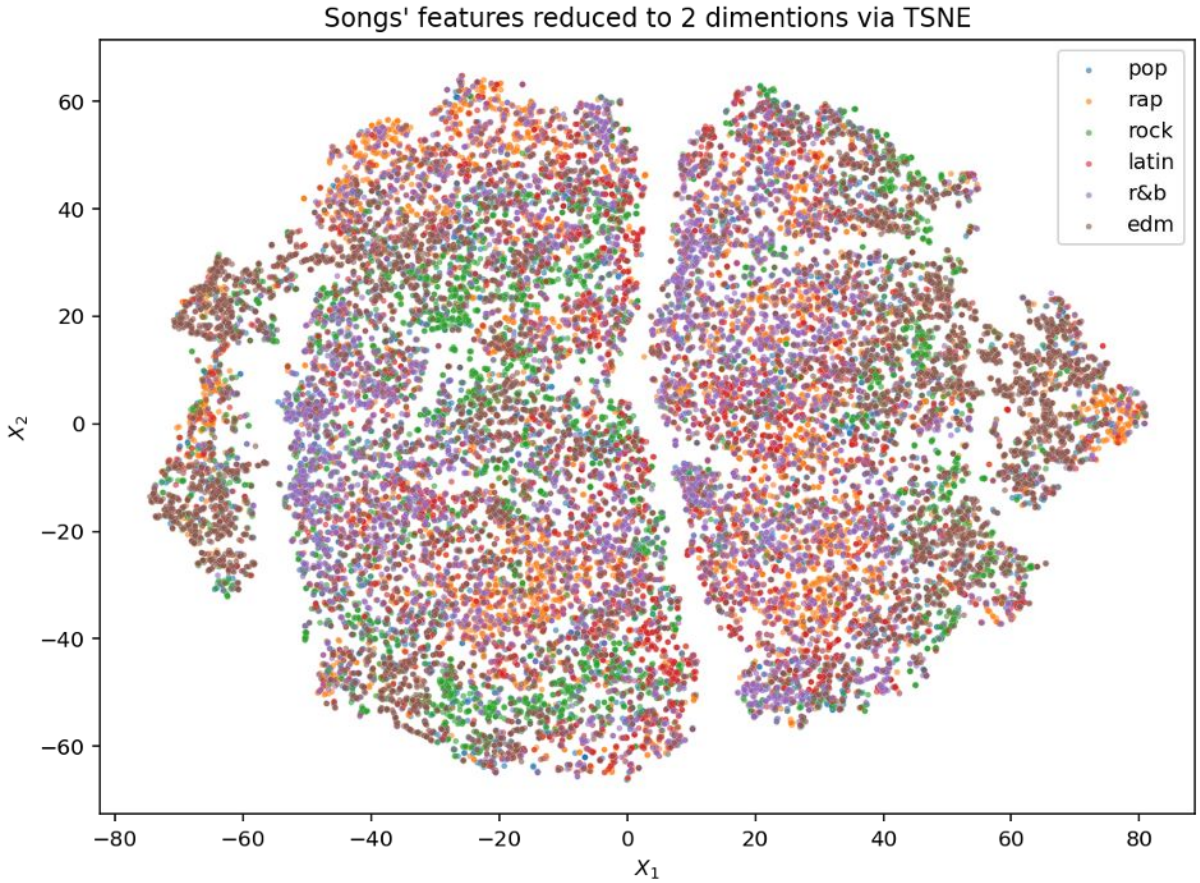
- Features are not highly correlated (except for loudness and energy)

- We have a good set of features that measure the songs in different ways!



Spotify Data

- Song genres seem inseparable when songs features are observed in 2 dimensions



Clustering methods

K-Means

- Assumes clusters are globular or spherical
- Tuned to provide optimal k = 5

DBSCAN

- Unclassified songs
- Too many clusters

Gaussian Mixture

- Assumes clusters have a Gaussian distribution
- Tuned to provide optimal k = 8

Clustering methods

DBSCAN

- Unclassified songs
- Too many clusters

K-Means

- Assumes clusters are globular or spherical
- Tuned to provide optimal k = 5

Gaussian Mixture

- Assumes clusters have a Gaussian distribution
- Tuned to provide optimal k = 8

Clustering methods

DBSCAN

- Unclassified songs
- Too many clusters

Gaussian Mixture

- Assumes clusters have a Gaussian distribution
- Tuned to provide optimal k = 8

K-Means

- Assumes clusters are globular or spherical
- Tuned to provide optimal k = 5

Evaluation Metrics for Clustering

Davies-Bouldin Score

Evaluates intra-cluster similarity and inter-cluster differences

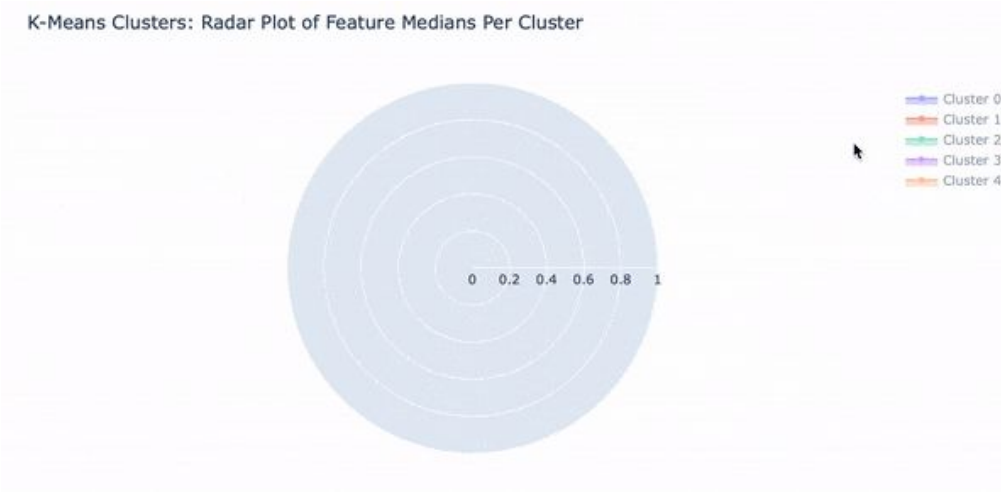
Silhouette Score

Measures the distance between each data point, the centroid of the cluster it was assigned to and the closest centroid belonging to another cluster

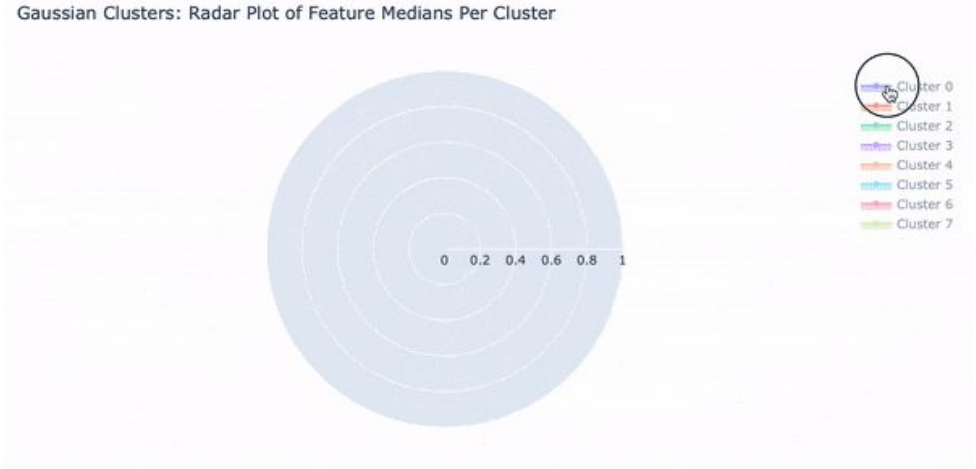
Evaluation Scores

	K-Means	Gaussian Mixture Model
Davies Bouldin Score	2.166	1.999
Silhouette Score	0.105	0.09

New genres obtained with K-means



New genres obtained with Gaussian Mixture Models



Supervised ML Results

A logistic regression model was used for prediction of new genres for unseen music

Classification Summary

	Precision	Recall	F-Score	Support
Cluster 0	0.97	0.99	0.98	497
Cluster 1	1.00	0.99	1.00	903
Cluster 2	0.99	0.99	0.99	995
Cluster 3	0.99	0.99	0.99	1841
Cluster 4	0.99	0.99	0.99	2331
Accuracy			0.99	6567
Macro Average	0.99	0.99	0.99	6567
Weighted Average	0.99	0.99	0.99	6567