Taste in Popular Music in the United States and around the Globe

An Exploratory Data Analysis using the Spotify Web API, by Jake Mullins

Goal:

- Explore Spotify's datasets to gain an understanding of the features that their app uses to classify audio tracks and tailor its music recommendations for its users.
- Use this data to compare popular music in select regions

Questions:

- How similar is popular music in the United States, United Kingdom, Mexico, and Canada?
- Which country’s musical taste is most similar to the global population’s musical taste?

Hypothesis:

- The USA is the country whose popular music is most similar to the Global

Method

Analyze the current "Top 50" playlists for each region.

- ❏ Similarity in Popular Tracks
- ❏ Similarity in Popular Genres
- ❏ Similarity in the ‘Features’ of tracks

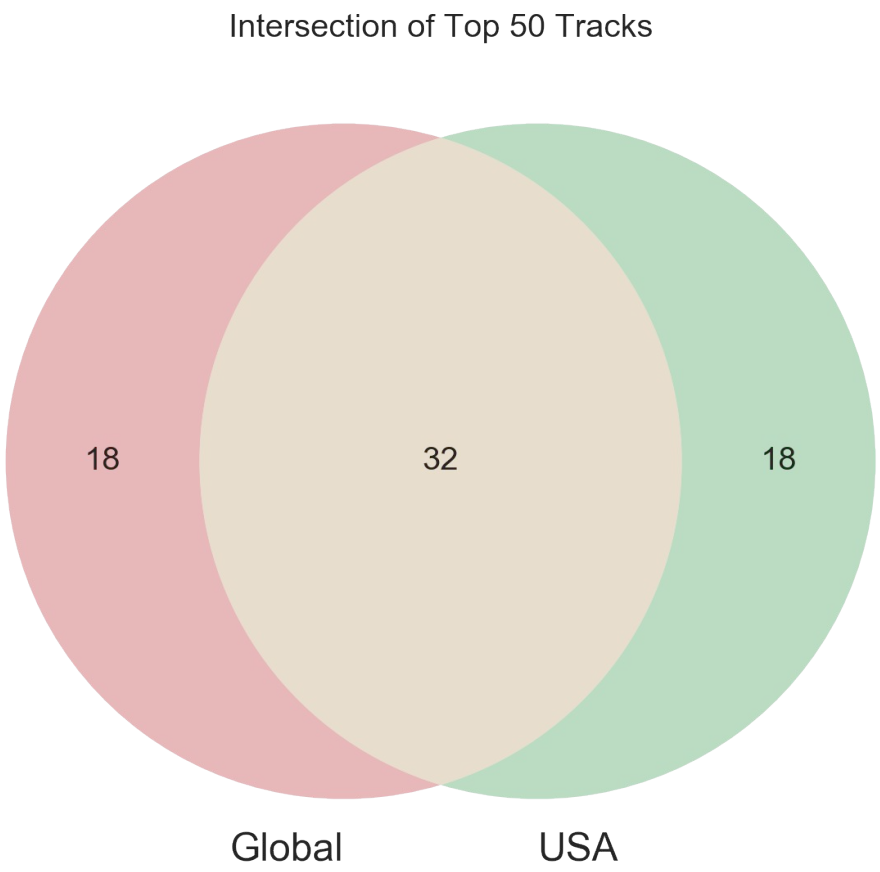
Create functions for calculating these metrics that are easily scalable to include much larger sample sizes.

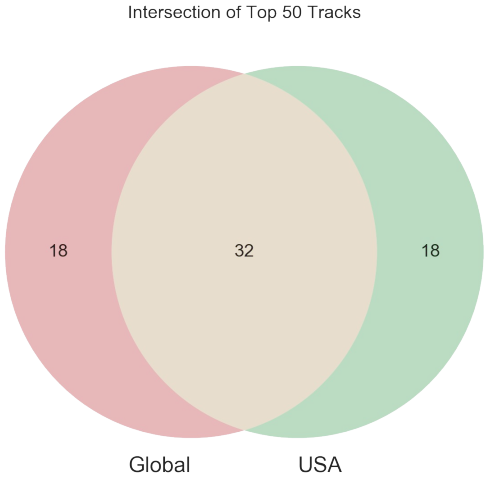
Example Data:

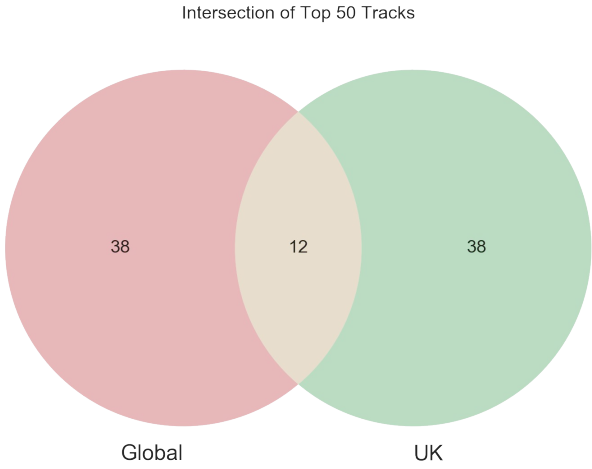
	track_id	name	album	artist	artist_id	genres	acousticness	danceability	energy	instrumentalness
0	0VjIjW4GlUZAMYd2vXMi3b	Blinding Lights	After Hours	The Weeknd	1Xyo4u8uXC1ZmMpatF05PJ	[canadian contemporary r&b, canadian pop, pop]	0.00146	0.514	0.730	0.000095
1	127QTOFJsJQp5LbJbu3A1y	Toosie Slide	Toosie Slide	Drake	3TVXtAsR1Inumwj472S9r4	[canadian hip hop, canadian pop, hip hop, pop ...	0.32100	0.834	0.454	0.000006
2	24Yi9hE78yPEbZ4kxyoXAI	Roses - Imanbek Remix	Roses (Imanbek Remix)	SAINt JHN	0H39MdGGX6dbnnQPt6NQkZ	[melodic rap, pop rap, rap, trap]	0.01670	0.770	0.724	0.010500
3	3PfIrDoz19wz7qK7tYeu62	Don't Start Now	Future Nostalgia	Dua Lipa	6M2wZ9GZgrQXHCFfjv46we	[dance pop, pop, uk pop]	0.01250	0.794	0.793	0.000000
4	0nbXyq5TXYPCO7pr3N8S4I	The Box	Please Excuse Me For Being Antisocial	Roddy Ricch	757aE44tKEUQEqRuT6GnEB	[melodic rap]	0.10400	0.896	0.586	0.000000

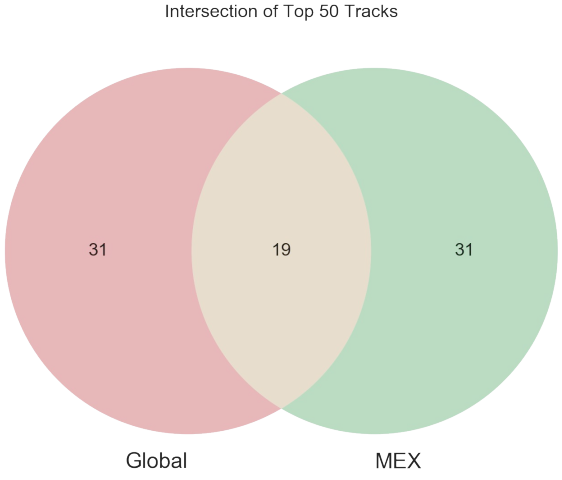
Track Set Similarity

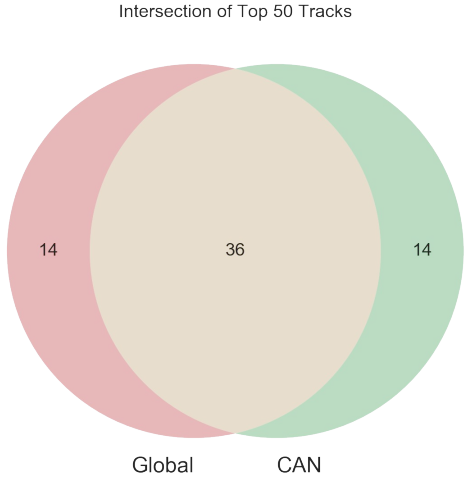
How many tracks does each playlist have in common?

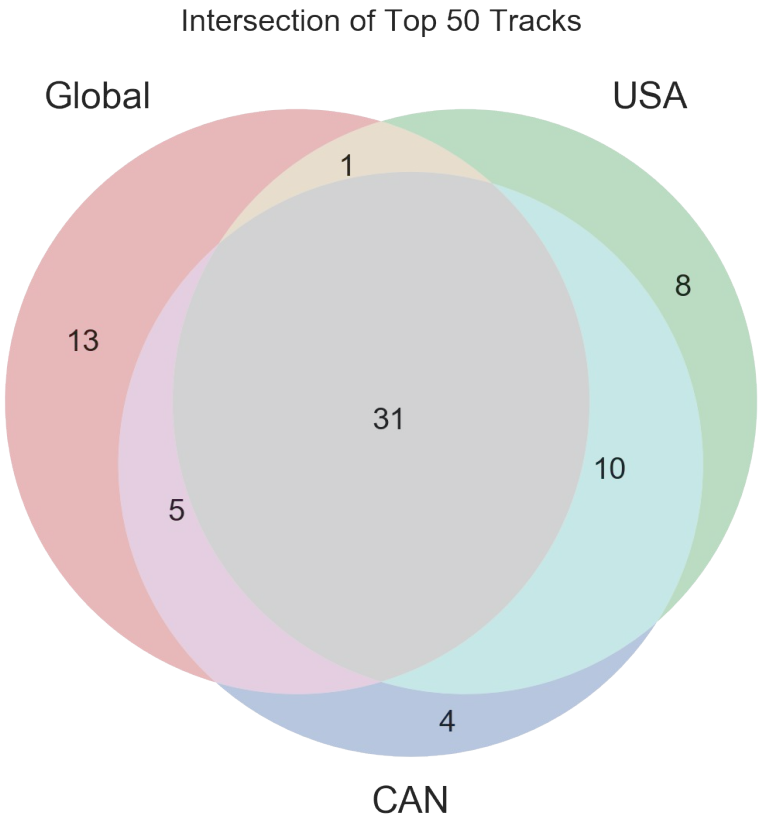

Intersection of Top 50 Tracks
18
32
18
Global
USA


Intersection of Top 50 Tracks
18
32
18
Global
USA


Intersection of Top 50 Tracks
38
12
38
Global
UK


Intersection of Top 50 Tracks
31
19
31
Global
MEX

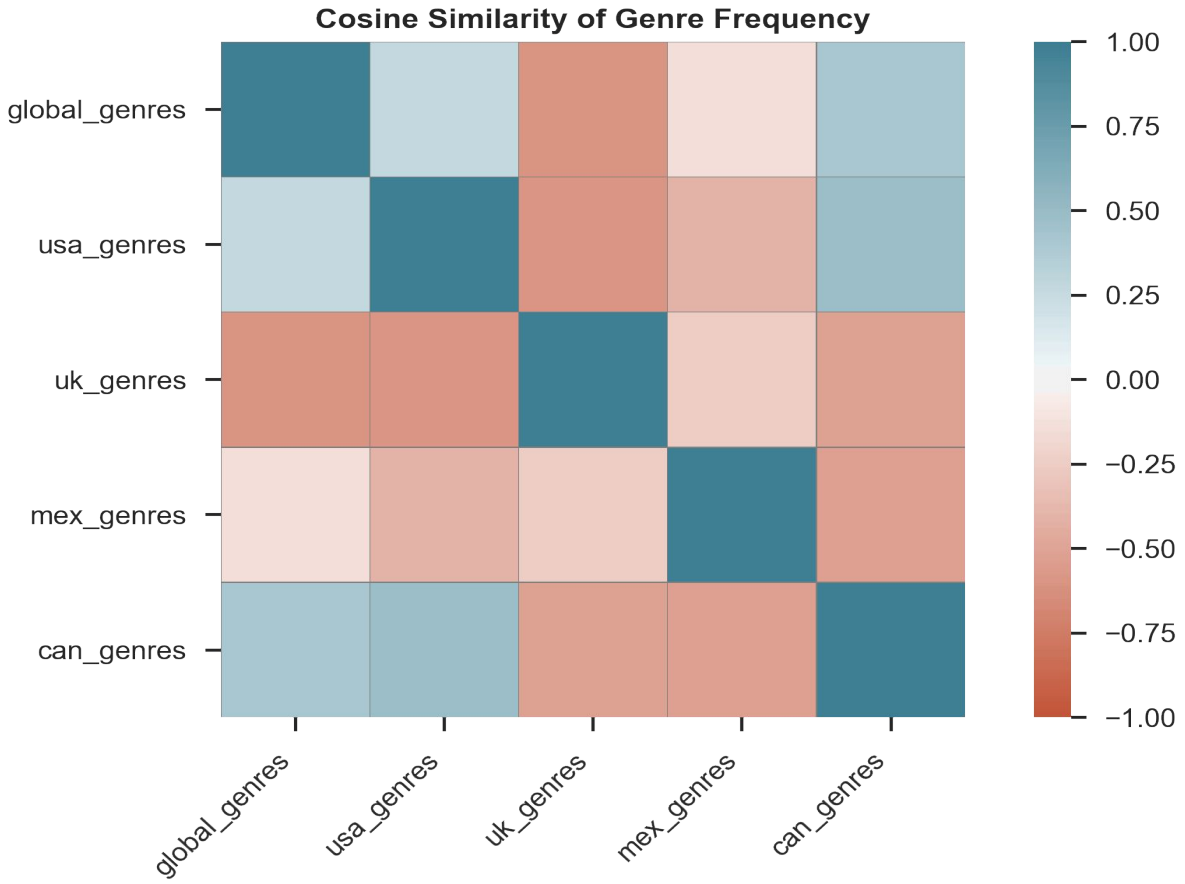

Intersection of Top 50 Tracks
14
36
14
Global
CAN



Result: Canada has the highest number of tracks in common with the Global set.

Genre Similarity

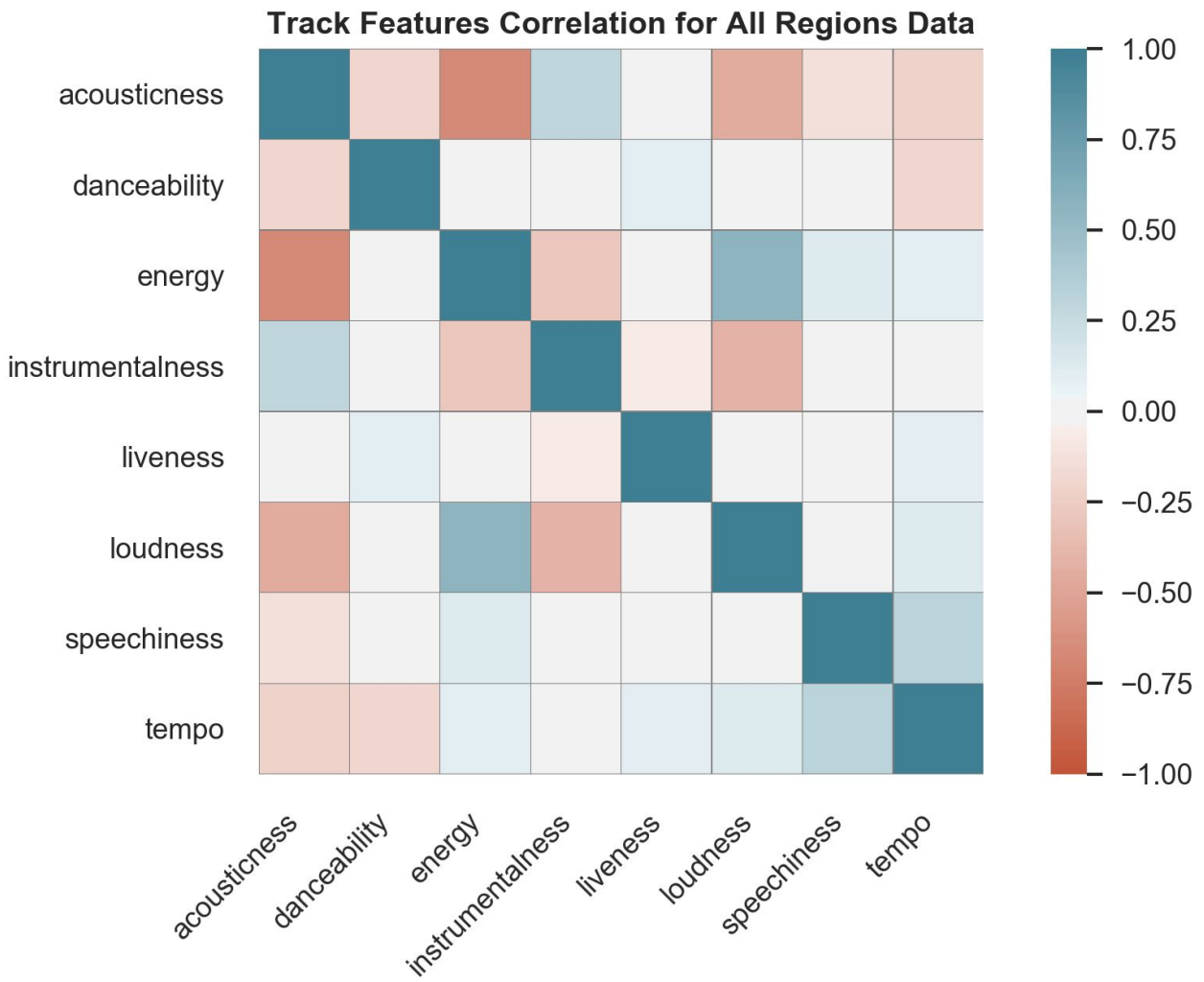
Count the frequency of each genre in a playlist, then calculate the cosine similarity for every combination of playlists.



Result: Canada's genres are most similar to the Global genres.

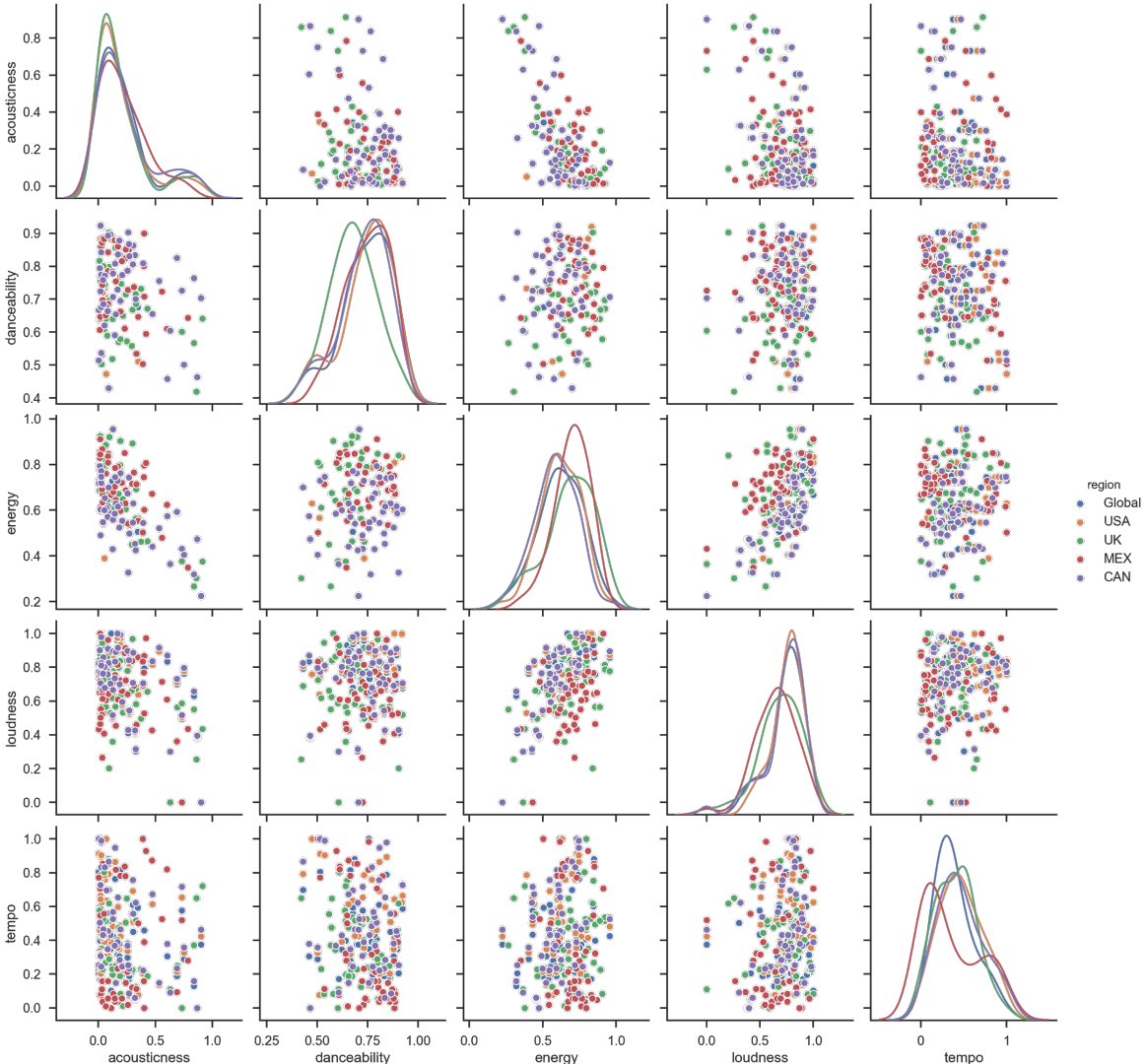
Feature Similarity

- Acousticness
  - 1.0 represents high confidence that the track is acoustic
- Danceability
  - 1.0 is most danceable
- Energy
  - 1.0 = death metal
- Instrumentalness
  - 1.0 represents high confidence that there are no vocals
- Liveness
  - 1.0 represents high confidence that a live audience is audible in the track
- Loudness :
  - Decibels (normalized)
- Speechiness
  - Rap is typically .33-.66
- Tempo
  - BPM (normalized)


Track Features Correlation for All Regions Data
acousticness
danceability
energy
instrumentalness
liveness
loudness
speechiness
tempo
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

Result: Canada's track features are the most similar to the Global track features.

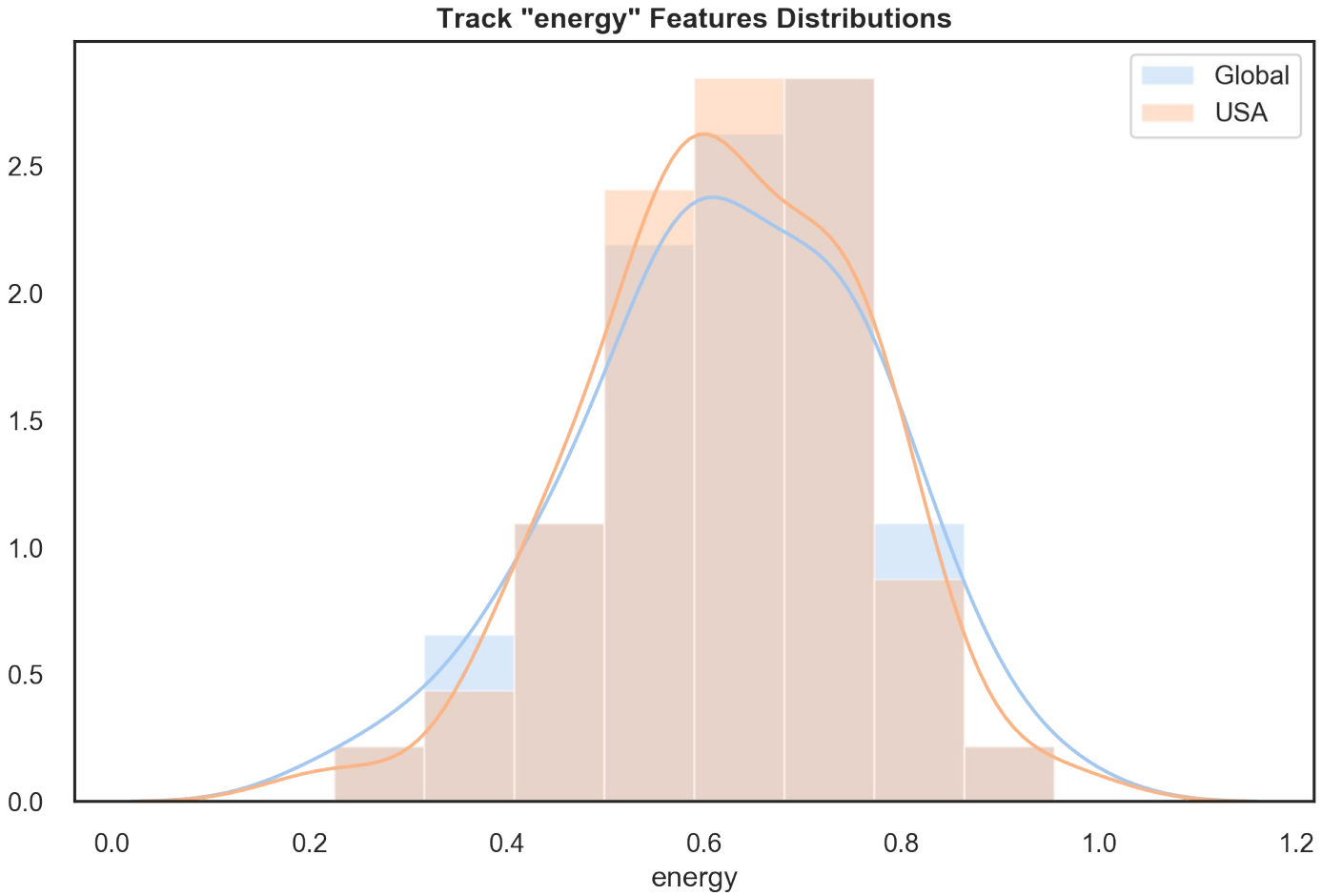
Feature Distributions


acousticness
danceability
energy
loudness
tempo
region
Global
USA
UK
MEX
CAN

Hypothesis Test

- Null Hypothesis: There is no difference in the mean “energy” of the Global tracks and the USA tracks.
- Use a two sided, independent t-test with alpha = .05 as our rejection threshold.

pval = 0.9966979993191145



Conclusions

- In this instance, we fail to reject the null hypothesis. We cannot say that the mean energy of the US tracks is different from the mean energy of the Global tracks.

- The data used in this project tells us a lot about a brief snapshot in time. Repeating these tests across a range of times would yield very interesting results about changes in musical taste over time.

Thank You.

GitHub
github.com/jakemull13

LinkedIn
linkedin.com/jacob.c.mullins

Email
jakemull13@gmail.com