**Project One - Analysis of song popularity and pattern analysis in the audio files of popular songs.**

1. Team Name: (F) Data Pirates
2. Members: Alex Stanley, Jeni Arulraj, Ramyaa Vivek
3. Data Sources: Billboard Hot 100 and Spotify API
4. Some of the Spotify Audio features that we used for our analysis,
5. Danceability - It represents how a track is danceable based on tempo, rhythm stability, beat strength, and overall regularity. 0.0 is least danceable and 1.0 is most danceable.
6. Valence - 0.0 is a measure of low valence and 1.0 is a measure of high valence. Valence describes the musical positiveness of a track. Tracks with high valence are more positive (e.g. happy, cheerful, euphoric), while tracks with low valence are more negative (e.g. sad, depressed, angry).
7. Popularity - Represents how a track is popular based on various musical elements including danceability, valence, energy, duration, mode, loudness, tempo, speechiness, liveness, key etc.
8. Energy - 0.0 is a measure of low energy and 1.0 is a measure of high energy. Energetic tracks sound fast, loud, and noisy. For example, death metal has high energy, while a Bach prelude scores low on the scale. Perceptual features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate, and general entropy.
9. Duration - The duration of the track in milliseconds.
10. Mode - Mode indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived. Major is represented by 1 and minor is 0.
11. Loudness - The overall loudness of a track in decibels (dB). Loudness values are useful for comparing relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typical range between -60 and 0 db.
12. Tempo - The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration.
13. Questions we asked?
14. Is Country music more popular than Rap? (T-test for popularity)

Null Hypothesis (H0): No popularity difference between Country & Rap music

Alternate Hypothesis (H1 or Ha): There is difference in popularity b/w Country & Rap music

T-test Result: P-value = 6.14946668928311e-17

Decision: Since P-value < 0.05, we reject H0. Hence, there is difference in popularity b/w Country and Rap music.

1. Valence: Country music vs Hot-100 vs Rap music (Anova for valence).

Null Hypothesis (H0): No difference in Valence b/w Country, Rap music & Hot 100

Alternate Hypothesis (H1 or Ha): There is difference in valence b/w Country, Rap music & Hot 100

T-test Result: P-value = 0.0004471337243690721

Decision: Since P-value < 0.05, we reject H0. Hence, there is difference in valence b/w Country, Rap music and Hot-100

1. Danceability: Country music vs Hot-100 vs Rap music (Anova for Danceability).

Null Hypothesis (H0): No difference in Danceability b/w Country, Rap music & Hot 100

Alternate Hypothesis (H1 or Ha): There is difference in Danceability b/w Country, Rap music & Hot 100

T-test Result: P-value = 1.6941337881835844e-13

Decision: Since P-value < 0.05, we reject H0. Hence, there is difference in Danceability b/w Country, Rap music and Hot-100

1. **Key observations based on Scatter plot charts:**
2. Popularity vs Valence – Hot-100 and Rap music is more popular than Country music. However, Country music seems to have more positive vibe than Hot-100 and Rap music.
3. Popularity vs Danceability – Danceability is directly related to popularity of songs. Rap music is most popular and also danceable followed by Hot-100 and then Country music.
4. Popularity vs Energy – Although Country music is less popular than Hot-100 and Rap music, it appears to be most energetic compared to the other two genres.
5. **Key observations based on Bar charts:**
6. Danceability vs Genre - Among songs of top artists, Rap music is the most danceable genre followed by Hot-100 and then Country music.
7. Valence vs Genre - Among songs of top artists, Country music is the most positive genre followed by Rap music and then Hot-100.
8. **Key observations based on Regression Analysis:**
   1. Danceability is the most positive correlated feature to popularity.
   2. With the given top 100 songs, only 49% of the dataset participated in the linear regression.
   3. Prediction had 0.4 success rate.
9. **Final take:** If Country music is more danceable, it could be more popular too!!!