Naavica Adunur
Markelle Kelly
Jenna Landy

Spotify Top Songs: Shiny App

The data set we used is the top 100 Spotify songs of 2017, with variables including danceability, energy, loudness, genre, sex of the artist, tempo, and valence. Our app has four main tabs, three of which have interactive plots describing the data: quantitative variables by genre, artists by genre, and regression analysis. The final tab contains descriptive statistics of all variables of interest.

Within the first tab, you can create a density plot of any quantitative variable across any subset of genres by selecting the desired genres in the "Genre" section, and choosing a quantitative variable from "Variable to Plot." For example, users could plot valence (the happiness of a given song) for the genres Hip-Hop/Rap and Latino. The density plot shows that on average, Latino songs are far more happy than Hip-Hop/Rap. Main takeaways from our density plot include that R&B songs are most danceable, Latino songs have the most energy, and Alternative songs are the loudest and most acoustic.

The second tab allows you to select one or more genres and see the number of top songs by each artist, colored by genre. Here, you can see that the genre with the most top songs is Pop, and within that genre, Ed Sheeran has the most (4) top songs. You can also further divide the artists by gender. This will allow you to see which gender has the most top songs. In this case the males clearly have a lead over the females in dominating the charts. You can further visualize this data by identifying the genres of the top songs that pertain to each gender. Here you can see that Pop is the genre that most frequently makes the list of top songs for both females and males. Although Pop is the clear winner for females, for males genres such as Dance and Hip-Hop/Rap come close behind Pop in terms of popularity. If we were to predict the next successful artist based off our bar chart, an artist's chances for having a top song would be higher if the person was a male in the Pop genre.

In the third tab, you can choose x and y quantitative variables to form a scatter plot and run a regression. Furthermore, you can choose to separate the plots by genre, sex of the artist, or mode (whether the song is in a major or minor key). When the plot is not separated by a categorical variable, the regression equation is printed and significance is tested, represented by a green line for significant slope values, at the 0.05 significance level. When plotting danceability as x and valence as y separated by sex of the artist, you can see that the variables seem to have a positive association in top songs by male artists, but no association in top songs by female artists (though no significance has been tested). Furthermore, all top songs by female artists have high danceability, while top songs by male artists have a spread of danceability values.

Overall, our app provides graphical visualizations to analyze the characteristics of Spotify's top songs in 2017 using regression, density plots, and bar plots.