**BUSINESS OBJECTIVES:**

To investigate Top 100 Billboard music dataset from 1950-2015 to determine market trends. Machine learning will predict genre of song using other datapoints such as syllables, word counts, and Flesh Index.

**RESEARCH QUESTIONS & HYPOTHESES:**

Do songs from certain genres have more complex lyrics?

How does popularity of genres trend over time?

Are songs lyrics becoming more repetitive over time?

Which artitist has spent the most time on the top 100 Billboard charts?

Which artist have the most songs that have appeared on the top 100 Billboard?

Are the more successful artist cross over multiple genres?

**DATA CHALLENGES/LIMITATIONS:**

When reviewing lyrics we must take into consideration that certain genres will have more lyrics than others and might skew if we fail to average the data in our visuals.

Multiple genres for one song for instance Michael Jackson crosses multiple genres.

USA data only

For 1950-1955, there are only the top 30 songs on that site

Limited by the pre-established bins that Billboard has put together. For instance, R&B and Hip Hop are bunched together.

**DATA SOURCES:**

<http://www.bobborst.com/popculture/top-100-songs-of-the-year/?year=2015>

<https://developer.spotify.com/documentation/web-api/>

<https://en.wikipedia.org/wiki/Billboard_charts>

**WEBPAGE:**

Select artist and it will give you title and average position for the year.

Tabs will have interactive Tableau graphs

We will use bootstrap to build the HTML

We will utilize machine learning to see if it is possible to predict the genre of a song in the top 100 Billboards by looking at the number of syllables, number of words, year, title, and lyric readability.