

Billboard Top 10 Prediction

Problem Statement:

What drives the “success” of a song? What factors (such as: title, artist, time, lyrics, emotional score, and prior placement) can be used by artists, producers, or others in the music industry to determine whether or not a song will make it to the top 10 of the Billboard Hot 100?

Context:

With the increase and combination of music streaming and social media platforms, now more than ever, it is easier for artists to connect their music with their fan bases. This increased accessibility drives an increased competition between musicians. For over 80 years, Billboard has documented this competition by publishing a weekly chart that ranks song popularity. How can an artist, producer, or other individual in the music industry predict whether or not a new song will be “popular?” Using historical Billboard Hot 100 data we will look at what factors drive a song into the top 10 vs all others.

Criteria for success:

A model or list of factors that can be used to explain the success of a song. This will be used for upcoming music releases to determine whether or not the song has a higher chance of landing in the top 10 of the billboard 100.

Scope of solution space:

The initial focus of this project will focus on general information and emotional scoring of songs on the Billboard Top 100.

Constraints:

This project is constructed only with data that has been on the Billboard Top 100 at least once, all other songs are not included. This data is historical and only covers up to 2020. Potential confounding factors, such as social or economic status of society at the time are not included in this analysis.

Stakeholders (potential):

Music Data Manager, Music Producer, Music Artist, Director of Music Streaming Service

Data Sources:

Our Top 100 data is imported from this website:
<https://data.world/kcmillersean/billboard-hot-100-1958-2017>

Our emotion data is imported from this website:
<https://data.world/tazwar2700/billboard-hot-100-with-lyrics-and-emotion-mined-scores>

This data is gathered and cleaned according to this blog post:
<https://medium.com/codex/emotion-mining-lyrics-of-all-billboard-hot-100-songs-1958-2020-3007d6963115>