

Problem Identification

- Problem statement formation

How can a successful record label (RCA Records) and emerging artists leverage the elements of songs to predict track popularity and increase their following by 2x in their respective genre for 2024?

- Context

Songs sourced from Spotify are analyzed based on a variety of characteristics. These songs have rankings in well known charts, which is included in the dataset, to give an idea of general performance amongst the public.

- Criteria for success

Understanding the attributes of songs and artists that have thrived in recent years. Gather a fundamental view into what attributes matter and which have little impact towards stream volume.

- Scope of solution space

A field that is highlighted is track popularity. Spotify makes recommendations of songs to users based on popularity, which can be identified using streams (a continuous data point). Other metadata such as danceability, artisticness and beats per minute of the songs are also included for added value.

- Constraints

Incomplete history, sample size of dataset is less than 1,000 songs. Release year for the song dates back as far as 1930, so certain artists may be irrelevant for the problem statement question.

- Stakeholders

President, Creative

President, Strategy

VP, Director of Data Strategy

Artists with less than 2 million monthly streams

- Data sources

Spotify API CSV

A lookup CSV to supply Artist Genres