

## DATS 6103 Final Project Proposal

*By Aditya Nayak, Alex Khater, Pooja Chandrashekara & Vaishnavi Nagraj*

**Project Proposal:** Using a dataset of nearly 11,000 songs from Spotify, we would like to observe the relationship between traditional music theory metrics (Key, Tempo, Time Signature, Duration, Tempo, Energy, Explicit, Mode) and a song's streaming popularity on Spotify and find out which factors contribute more towards a song be more popular.

**S:** This question is tied to 7 specific columns in the dataset and a clear response variable, it is quite a specific question.

**M:** Our response variable is clearly numeric and shows how many times a song is streamed. It is easily measurable.

**A:** This question is quite attainable with relatively basic modeling and regression methods as all variables are standard integer or categorical data.

**R:** This question will attempt to distill a popular song down to its core components from a music theory standpoint and give insight into creating one and will help in characterizing which songs can be trending or most played.

**T:** This project will be doable with a team of 4 as the dataset is quite clean from the start and the team is big enough to finish the project quickly.

We will use Multiple Linear Regression, Correlation Analysis or any other models that reveal anything interesting.

**Dataset Link:** <https://www.kaggle.com/datasets/maharshipandya/-spotify-tracks-dataset>

**Github Link:** <https://github.com/awinkhater/DATS6103T3>