

STATS 418: Final Project Proposal

Spotify Song Quiz

By Derek Wen

Data Collection & Dataset

- **Data Source:** Spotify API
- **Dataset:** Top 100 songs of 2024
- **Data attributes:**
 - Song Name (categorical)
 - Artist Name (categorical)
 - Song Attributes (numerical)
 - Danceability, Energy, Tempo, Loudness, etc
- **Collection Method:** Python script using Spotify API credentials

Exploratory Data Analysis & Modeling Plan

- **EDA Insights**

- **Top Genres:** Analysis of most frequent genres
- **Audio Features:** Distribution of key attributes in a song (tempo, energy, etc)
- **Artist Popularity:** Frequency of artists appearing in the top 100 list

- **Modeling Plan:**

- **Recommender System:** Implement a simple recommender system using Euclidean distance on selected audio features of a song to suggest similar songs from the top 100 list

Project & Application Proposal

- **Interactive Quiz App:**
 - Plays a random song from the list of the top 100 songs of 2024
 - User guess the song name and artist
 - Shows the answers and whether user guessed correctly or not
- **Song Recommendation Model**
 - After each quiz round, the application will recommend similar songs based on audio feature similarity
- **Deployment Plan**
 - Frontend: Shiny app (hosted on shinyapps.io)
 - Backend: Flask API hosted on Amazon EC2
 - Docker for easy deployment and scalability