

Group 6 ETL Project Proposal

Max Izotov, Julissa Guzman, MaryLouise Pabello

November 2, 2020

Our team obtained data from Kaggle to Analyze the Top Spotify Tracks of 2017 and 2018.

Original Data Source:

The audio features for each song were extracted using the Spotify Web API and the Spotify Python library. The Spotify data specifically contains the following information:

- Name of the song
- Artist(s) of the song
- Year
- Audio features for the song (such as danceability, duration, energy, etc.)

The dataset can be used to analyze the song patterns in 2017 compared to 2018 and analyze what audio features make them popular.

We used PostgreSQL pgAdmin to store and load all datasets.