Problem Identification

- Problem Statement
  - As the audio streaming platform Spotify, I want to give users a filtering option to select songs in their created playlists based on 26 genres, so that I can increase revenue of the service by converting the free users into premium users.
- Context
  - Spotify has free users and premium users for which under the section ‘liked songs’ allows users to filter songs based on genre. If I can create a good filtering option based on the songs in users playlists to mimic the current mobile filter option for users liked songs, I could potentially convert those free users into paying premium users as they might find the spotify service as desirable.
- Criteria for success
  - If the model allows users to filter songs in their created playlists based on 26 genres.
- Scope of solution space
  - I will use the kaggle spotify data set of 232,725 tracks in 26 genres to create a model. The model will be trained on this data and use it for testing.
- Constraints
  - A potential constraint is the lack of a larger number of genres that could be useful in creating a more accurate filter system of songs in users created playlists, but without being too large to be non user friendly.
- Stakeholders
  - Spotify executives and investors.
- Data sources
  - I will use this kaggle spotify data set of 232,725 tracks in 26 genres. It includes 18 columns of data: genre, acousticness, artist name, danceability, duration_ms, energy, mode, track id, instrumentalness, key, liveliness, loudness, song name, popularity, speechiness, tempo, valence, and time signature.
  - https://www.kaggle.com/zaheenhamidani/ultimate-spotify-tracks-db