**Project Outline Draft 1**

✓ Selected topic

The selected topic is analyzing and visualizing Spotify data to determine what factors contribute to song popularity. Additionally utilizing those factors to predict future song popularity.

✓ Reason why they selected their topic

We enjoy music, and it is an interesting choice to visualize.

✓ Description of their source of data

The data source is a Kaggle Spotify dataset in a csv file. There are approximately 40K sets of data within the csv.

✓ Questions they hope to answer with the data

* What components within the dataset determine popularity of a song?
* How can we use these components to predict future song popularity?
* What factors have no contribution?

✓ Description of the data exploration phase of the project

* Mock ML to clean and visualize the data
* EDA

✓ Description of the analysis phase of the project

* Description of data preprocessing
* Description of feature engineering and the feature selection, including the team's decision-making process
* Description of how data was split into training and testing sets
* Explanation of model choice, including limitations and benefits
* Explanation of changes in model choice (if changes occurred between the Segment 2 and Segment 3 deliverables)
* Description of how model was trained (or retrained, if they are using an existing model)
* Description and explanation of the model's confusion matrix, including final accuracy score Additionally, the model obviously addresses the question or problem the team is solving.

✓ Technologies, languages, tools, and algorithms used throughout the project

* Database stores static data for use during the project
* Database interfaces with the project in some format (e.g., scraping updates the database, or database connects to the model)
* Includes at least two tables (or collections, if using MongoDB)
* Includes at least one join using the database language (not including any joins in Pandas)
* Includes at least one connection string (using SQLAlchemy or PyMongo) Note: If you use a SQL database, you must provide your ERD with relationships.

✓ Result of analysis

* The dashboard presents a data story that is logical and easy to follow for someone unfamiliar with the topic. It includes all of the following:
* Images from the initial analysis
* Data (images or report) from the machine learning task
* At least one interactive element Either the dashboard is published or the submission includes a screen capture video of it in action.

✓ Recommendation for future analysis

✓ Anything the team would have done differently