Deliverable 2: Project Proposal

Introduction

This project explores popular music in the 1960s which details the popularity, song duration, track name, and track ID. This dataset can help identify artists or songs that were especially popular. Identifying the artists and songs can provide feedback as to what makes songs or artists popular. This project aims to identify patterns and trends in popular music in the 60s. This is important because it can offer insight into cultural and historical phenomena.

Objectives

This project aims to identify correlations between different factors and the popularity of 1960s music in terms of track duration and artist influence. This project will also focus on the factors that may have led to possible outliers. Identifying these factors can allow an understanding of what made popular music in the 60s and can provide insight into what may make popular music currently.

Methodology

This project will start with cleaning the data and ensuring it is organized and readable. Next, generative statistics and visualizations will be generated to identify correlations. One example might be a contingency table to identify artists with bigger influences. Another example might be a scatterplot between rankings and song duration to determine if there is a correlation. Finally, there will be a summary of the key findings.

This project will use Rstudio for data analysis and visualization and packages such as dplyr, lubridate, ggplot2, and tidyverse for data tidying, data manipulation and visualizations.

The timeline is outlined in figure 1.

A black background with pink lines and white text

Description automatically generated

Figure 1: Project Timeline

Expected Outcomes

The final outputs of this project will be a compilation of all the visualizations and generative statistics. This will be in a report in an rmd or html format. The final project will also be put into a zip file and uploaded to Github. Finally, there will be a final presentation highlighting key findings with visualizations. A successful project is a complete understanding of the data and meta-data. It will also successfully communicate the findings and what they mean. In this case, that would mean understanding music in the 1960s and being able to communicate in the report and presentation the understanding.

Challenges and Limitations

Some challenges may include inconsistencies in track names or misspelled artist names. Additionally, since the data only includes data on popularity and track duration, factors that affect popularity can be overlooked which affects the data’s scope. Additionally, there are potential learning curves in terms of analyzing data and using the analysis tools.

Conclusions

This project aims to identify correlations between different factors and popularity in 1960s music. Analyzing this data through visualizations and different statistical techniques will allow further insight into what made music popular during that time. Delivering a thorough analysis, applying core concepts to create meaningful findings about 1960s music through Rstudios techniques.

References

<https://www.kaggle.com/datasets/kanchana1990/spotify-oldies-dataset>