Business Case: Analyzing Song Trends on Spotify to Improve the Popularity of Singers

Executive Summary: This business case aims to analyze the music industry, which is growing rapidly and becoming more competitive each year due to the emergence of new platforms and mediums. To stand out and reshape the way audiences connect with artists, companies are providing unique, artist-focused content or developing different pricing models. Spotify, founded in 2006, has become the most popular streaming platform worldwide. To help singers improve their popularity and competitiveness in the music industry, we will use the **"Spotify Charts"** dataset from Kaggle, which contains the top 200 streamed tracks on Spotify every day from Jan 1, 2017, to Dec 31, 2021, collected using the Spotify API. The dataset has **26,173,514 observations and 9 columns**, and it covers countries around the world. Through exploratory data analysis, we will identify song trends and suggest solutions to improve the popularity of singers.

Background: Our music streaming industry has seen rapid growth in recent years, driven by the increasing popularity of online music and the rise of smartphones and other mobile devices. The background of this project is the rapid growth of the music streaming industry and the challenges faced by music streaming companies in delivering high-quality content and personalized user experiences to their customers.

Purpose: The project aims to use the "Spotify Charts" dataset to **analyze song trends of singers**, identify solutions to improve their popularity, and help them become more competitive in the music industry. Through exploratory data analysis, the project will provide insights into user preferences and trends, which can help music streaming companies stay up to date with the latest music trends and adapt their offerings to remain competitive.

Methodology: We will use exploratory data analysis (EDA) techniques to analyze the dataset and identify trends in music streaming. We will use tools **PySpark** to clean and preprocess the data. The analysis will focus on the top songs, artists, and regions, and we will use visualizations to identify patterns and trends.

To perform the analyses, we will utilize **Google Compute Engine and Google Colab**, which provide cloud-based computing resources and enable collaborative coding and analysis. **We will store the cleaned and preprocessed dataset on Amazon S3**, which provides scalable and cost-effective storage. Additionally, we will **directly store the visualizations** generated from the analyses to the **S3 bucket** to facilitate easy access and sharing with stakeholders who may want to **create websites or dashboards in the future.**

Target Audience: The target audience for this business case includes **music streaming companies, record labels, and individual singers** who want to improve their popularity and competitiveness in the music industry. Music streaming companies can use the insights gained from this analysis to improve their platform offerings and attract more users. Record labels can use the analysis to make data-driven decisions about which artists to sign and how to promote their

music. Individual singers can use the analysis to understand the factors that contribute to a song's success and create music that resonates with their audience. The insights gained from this analysis can benefit the entire music industry by providing a deeper understanding of user preferences and trends.

Conclusion: The proposed analysis of the music industry using the Spotify Charts dataset will provide valuable insights into the trends and factors that contribute to the success of songs and artists. The analysis will help singers to improve their popularity and competitiveness and provide valuable information for music streaming platforms and record labels to make data-driven decisions.

Overall, this project aims to provide valuable insights into the music industry using EDA, clustering, and regression techniques. The results can be used by music streaming companies, record labels, and artists to improve their offerings and stay ahead of the competition in this rapidly evolving industry.

We are confident that the insights gained from this analysis will be beneficial to the music industry as a whole and lead to better content and user experiences.