A blue and white logo

Description automatically generated1

**Data Programming**

**Group-2**

**(BDAT1004-23F-10981)**

## **Final Project**

Spotify Data visualization

December 6th, 2023

## **Submitted by**

## Joyson Gonsalves 200533709

## Sulav Thapa 200575683

## Oluwakorede Balogun 200572710

## Laxman Rokaya 200562874

## Darpan Aryal 200569576

**Submitted to**

## Prof. Roshan Shahu

We embarked on a comprehensive exploration and analysis of music data from Spotify; The following steps were engaged in the process:

* Data retrieval from Spotifys API
* Data Processing using inbuilt python functions
* Data Analysis using Pandas
* Data Visualization using Matplotlib and Seaborn
* Data Storage using MongoDB Atlas

Firstly, the script acquires an access token from the Spotify API using client credentials, enabling it to query a diverse set of music genres and gather intricate audio features for individual tracks within each genre. The dataset contains details such as Track ID, Name, Artist, Popularity Score, Release Year, and Genre. This dataset is then enriched through a merge operation with Spotify's audio feature data, creating a comprehensive and detailed representation of the music landscape.

The next stage of the analysis was marked by insightful exploratory data analysis and visualization.

The code does the following:

* Filters the dataset to spotlight popular tracks
* Calculates average popularity scores for each genre
* Constructs visualizations to view trends over release years.

These visualizations take the form of bar charts showcasing the average popularity score by genre and line plots delineating popularity trends across the years for the top 10 genres. Another aspect of the analysis is the correlation analysis between popularity and various audio features, a relationship that was visualized using a heatmap.

For future analyses and to ensure data accessibility, the processed dataset was stored in a CSV file named 'Final\_file.csv.' The script establishes a connection with a MongoDB database using a URI, creating a collection named 'Final\_project' and inserting the data.

The visuals used (Bar Charts, Line Plots, and Heat Maps) offer intuitive representations of music trends and correlations between features, catering to a diverse audience that includes both technical and non-technical stakeholders.

In conclusion, this Python code forms a foundation for a data-driven exploration/analysis of music trends, genre popularity, and the relationship between musical features and audience preference. It explores the intricacies of Spotify's music landscape and presents the findings in a comprehensible and visually appealing manner.