

Spotify Song Attributes

Prepared in the partial fulfillment of the Summer Internship Program on Data Analysis

AT



Under the guidance of

Mrs. Meenakshi and Mrs. Narmada Mani , APSDC

Submitted by

P. CHARISHMA – AP21110011376

P. HARSHA – AP21110011450

K. BALAJI – AP21110011477

SRM University AP (August, 2023)

ACKNOWLEDGEMENT

I would like to express my heartfelt gratitude to all those who have contributed to the successful completion of my summer internship project at **Andhra Pradesh Skill Development Corporation (APSSDC)**. This opportunity has been an enriching and transformative experience for me, and I am truly thankful for the support, guidance, and encouragement I have received along the way.

First and foremost, I extend my sincere regards to Mrs Meenakshi and Mr Narmada Mani , my supervisors and mentors, for providing me with valuable insights, constant guidance, and unwavering support throughout the duration of the internship. Their expertise and encouragement have been instrumental in shaping the direction of this project.

I would like to thank the entire team at **Andhra Pradesh Skill Development Corporation (APSSDC)** for fostering a collaborative and innovative environment. The camaraderie, knowledge sharing, and feedback I received from my colleagues significantly contributed to the development and success of this project.

In conclusion, I am honoured to have been a part of this internship program, and I look forward to leveraging the skills and knowledge gained to contribute positively to future endeavours.

Thank you.

ABSTRACT

The purpose of “Spotify Song Attributes” is to make more efficient & meaningful of the existing system by the help of programming language and their useful libraries, their built-in-functions, and methods... for reaching their business requirements, so that their valuable data/information can be used for analysis with ease. Data Manipulation & Data Visualization of the data is done separately. The required programming language is easily available and easy to work with.

The aim is to analyze the whole dataset and on the basis of the dataset we can get some insights, trends & patterns. It can assist the users to concentrate on their main activities rather than anything else. Thus, it will help organization in better utilization of resources. Basically, the project describes how to analyze data and get insights, trends & patterns out of it.

TABLE OF CONTENTS

S.No	Content	Pg.No
1.	Introduction.....	5
2.	Methodology.....	6
3.	Implementation	8
4.	Results.....	11
5.	Conclusion	19

1. INTRODUCTION

The “Spotify Song Attributes” has been developed to analyze the dataset to improve better decision in business. This code is supported to eliminate and, in some cases, reduce the uncertainty faced by the existing system. Moreover, the code is designed for the particular need of the company to carry out operations in a smooth and effective manner.

The application is reduced as much as possible to avoid redundancy while analyzing the data. It also provides extra insights while analyzing invalid/null data. A little bit knowledge of Python is needed for the user to understand the Data Analysis.

Spotify song attributes can lead to more valuable insights, better decisions, increase sales and customer retention. It can suggest to the user to recommendation on their search activities. Every organization, has challenges to overcome and managing the information of Artist, Albums & their songs. This is coded to analyze in strategic decision, and will help you ensure that your organization is equipped with the right level of insights and details for future goals. This will ultimately allow you to better see the user’s preferences and efficiently manage resources accordingly.

2. METHODOLOGY

The development of the Spotify Song Attributes followed a systematic and iterative methodology to ensure the successful implementation of the project objectives. The methodology encompassed several phases, including requirements gathering, design, implementation, testing and deployment. The following sections outline each phase of the methodology in detail.

2.1. Requirements Gathering

The project commenced with an analysis of the requirements and expectations of the Spotify Song Attributes. This phase involved interactions with mentors. Feedback sessions were conducted to identify key features.

2.2. Design

Based on the collected requirements, a system design was formulated. The architecture of the system was planned, outlining the whole data set.

2.3. Implementation

The implementation phase involved translating the design specifications into a functional system. The project is totally built at open source end and thus anyone can access the program. It starts with analyzing the dataset of application's data like Artist, Streams, Albums, Tracks, Duration, Loudness, and many other parameters. It manages all the information about Scale, Mode, Danceability, Loudness, Duration...

2.4. Testing

The testing phase encompassed integration testing and user acceptance testing. Integration testing is the testing of multiple components in an application to ensure that they work as

expected when integrated all together. The point of integration testing, as the name suggests, is to test whether separated developed modules work together as expected. Integration testing is mainly done in grouping and sorting implementation methods. User acceptance testing involved real users interacting with the system to identify and rectify any usability issues or bugs.

2.5. Deployment

The deployment of the Spotify Song Attributes was executed in Jupyter Notebook for analyzing and visualizing data with excel sheet for the CSV data. The system's components were configured to ensure optimal performance, security, and scalability. Continuous monitoring and optimization were performed to maintain system stability.

2.6. Iterative Refinement

Throughout the development lifecycle, an iterative approach was adopted to incorporate feedback, address challenges, and refine the system. Regular interactions with mentors provided valuable insights that guided refinements and enhancements.

The methodology employed in the development of the Spotify Song Attributes, facilitated a comprehensive and user-centric approach, ensuring the creation of a functional, secure, and intuitive platform for users on recommendation of the required songs.

3. IMPLEMENTATION

The implementation phase of the Spotify Song Attributes involved translating the whole data set into a fully functional and user-friendly application. This section provides an overview of the key components and technologies used during the implementation process.

3.1. Analyzing Data Set

In this phase of implementation, there is a data set provided with a lots of columns and rows. We focus on understanding the whole data set and start implementing it by reading the data set. Removing all the null values present so not to collide with the other implementation process. Accessing only the required columns and renaming them as there are many columns present, we are supposed to remove all the unwanted columns and display only the required columns for better understanding and easy implementation.

3.2. Read Dataset

There is a data set provided with a lots of columns and rows. We focus on understanding the whole data set and start implementing it by reading the data set, for better understanding of what columns and their types are present.

3.3. Apply Basic Functions

After Reading the dataset, we start applying some functions on the data set like

- `df.shape` --- to know number of rows and number of columns present
- `df.columns` --- this method displays all columns present in the data set.

3.4. Data Cleaning

This process involves cleaning the data set. We can depict that Data Cleaning is a technique which helps to convert improper data into meaningful data. In the given data set, we are

provided with null columns by default. We start removing those columns so that we can utilize the data in an efficient manner. Therefore, It is important to process data before use.

3.5. Data Filtering

Data Filtering is one of the most frequent data manipulation operation. In terms of speed, python has an efficient way to perform filtering and aggregation.. Pandas filter() function in python is used for filters in the Data Frame rows and columns. The returned Data Frame contains only rows and columns that are specified with the function. It doesn't update the existing Data Frame instead it always returns a new one.

3.6. Grouping

After working on analyzing the data set, we start grouping the columns, on the user requirements, for the better comparison of the data set. For example: grouping the 'Artist' and 'Genre' column for getting relation between artist column and genre column. This grouping method is mainly helpful for getting comparisons between columns.

3.7. Sorting

Sorting is the process of arranging data into meaningful order so that you can analyze it more effectively. For example, you might want to get the artist name with respect to the type of genre so that you can produce a graph of Popular artists by genre type. The sorting can be both in ascending and descending of data elements. Python also has some built-in-functions, which help you sort the elements in the list.

3.8. Aggregation

Dataframe.aggregate() function is used to apply some aggregation across one or more column. Aggregate using callable, string, dict, or list of string/callables. Most frequently used aggregations are:

- **sum** --- Returns the sum of the values for the requested axis.
- **min** --- Returns the minimum of the values for the requested axis
- **max** --- Returns the maximum of the values for the requested axis

3.9. Data Visualization

Data Visualization with Matplotlib is the presentation of data in a graphical format. It helps people understand the significance of data by summarizing and presenting a huge amount of data in a simple and easy-to-understand format and helps communicate information clearly and effectively.

The successful implementation of the Spotify Song Attributes involved the seamless integration of steps like importing required modules, reading dataset, applying basic functions, data cleaning, data filtering, grouping, sorting, aggregations (min,max,sum...), data visualization. This implementation process ensured that the project's design was translated into a functional and user-friendly application.

4. RESULTS

The implementation of the Spotify Song Attributes yielded significant outcomes and accomplishments. This section highlights the key results achieved through the development and deployment of the system.

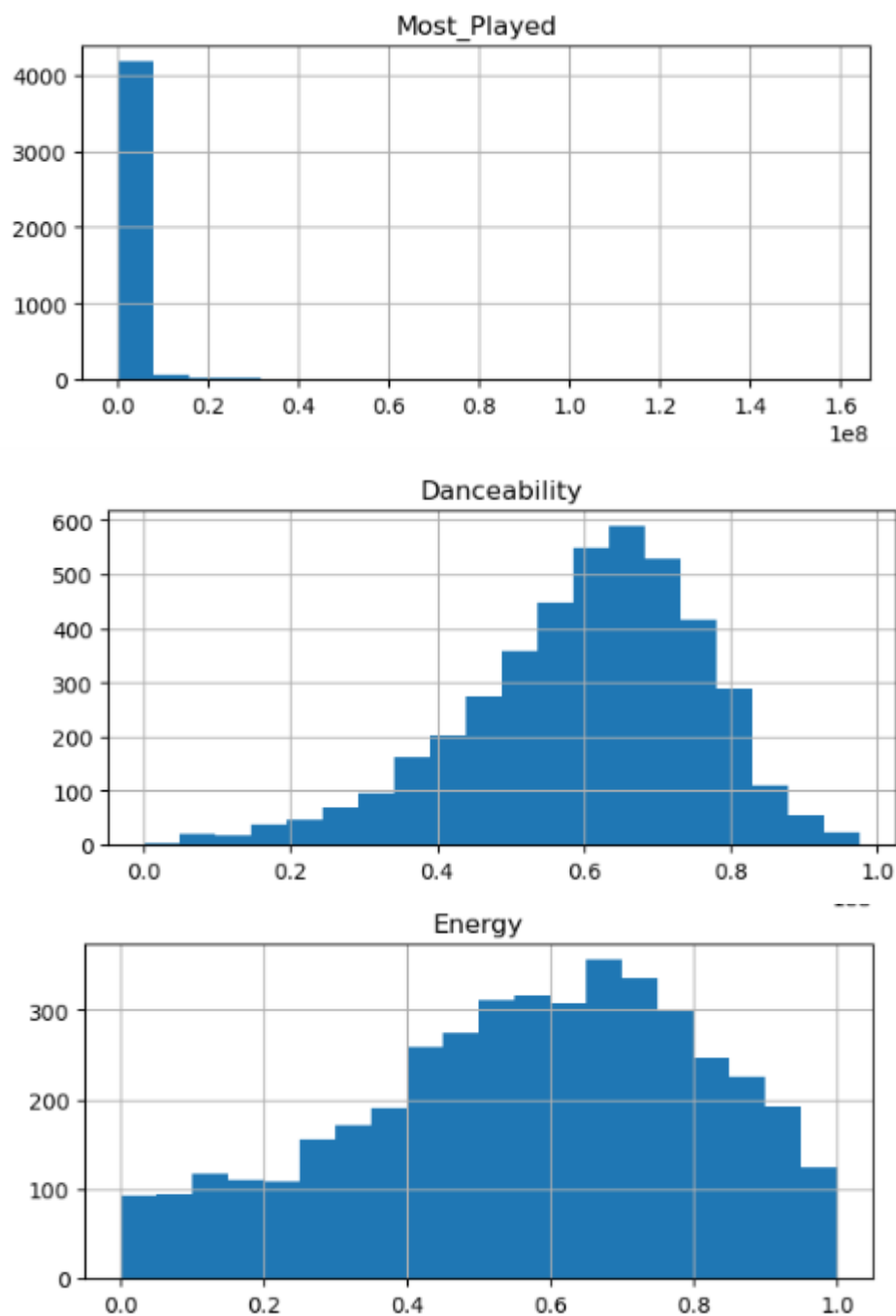
The main objective of the Project on 'Spotify Song Attributes' is to analyze the dataset of application's data like Artist, Streams, Albums, Tracks, Duration, Loudness and many other parameters. It manages all the information about Scale, Mode, Danceability, Loudness, Duration... The project is totally built at open source end and thus anyone can access the program. The purpose of the project is to build a program to analyze Spotify's data to get some insights, trends & patterns out of it. For analysis, we are using Python programming language and dataset is taken from Kaggle.

It may help drawing conclusions out of dataset. In a very short time, the insights will be obvious, simple and sensible. It will help in understanding the trends & patterns of users perfectly and vividly. It also helps in comparative analysis between Artist, Genres and Years. It will also help in recommendations feature. Our project aims at Business goals optimization, i.e., we have tried to analyze various parameters of Spotify's tracks.

- It satisfies the user requirement.
- Be easy to understand by the user and operator.
- Be easy to operate.
- Have a good user interface.
- Be expandable.
- Delivered on schedule within the budget.

Screenshots of the outputs

Figure 1: Distribution of Numerical Features Using Histograms



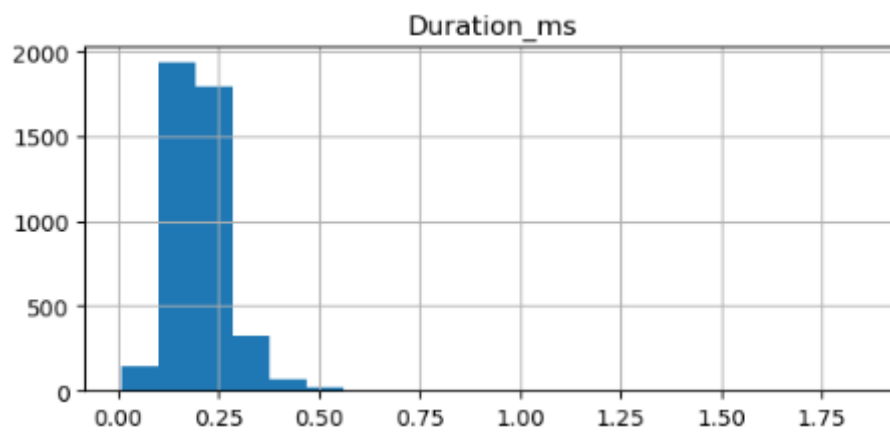
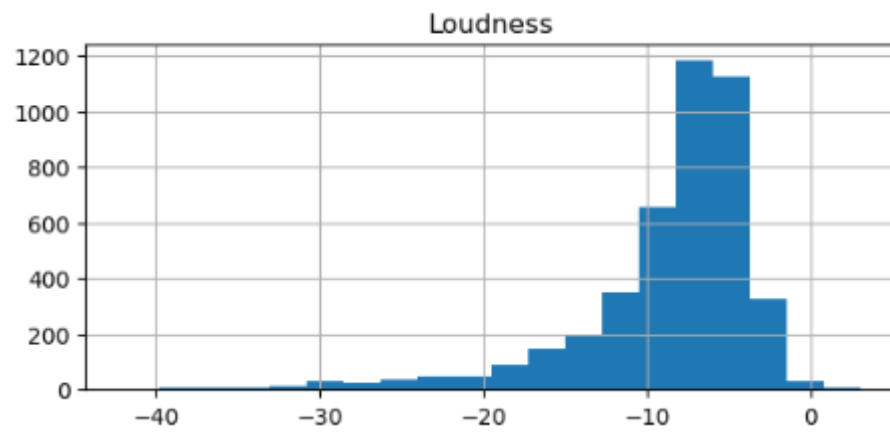


Figure 2: Top 10 Popular Artists

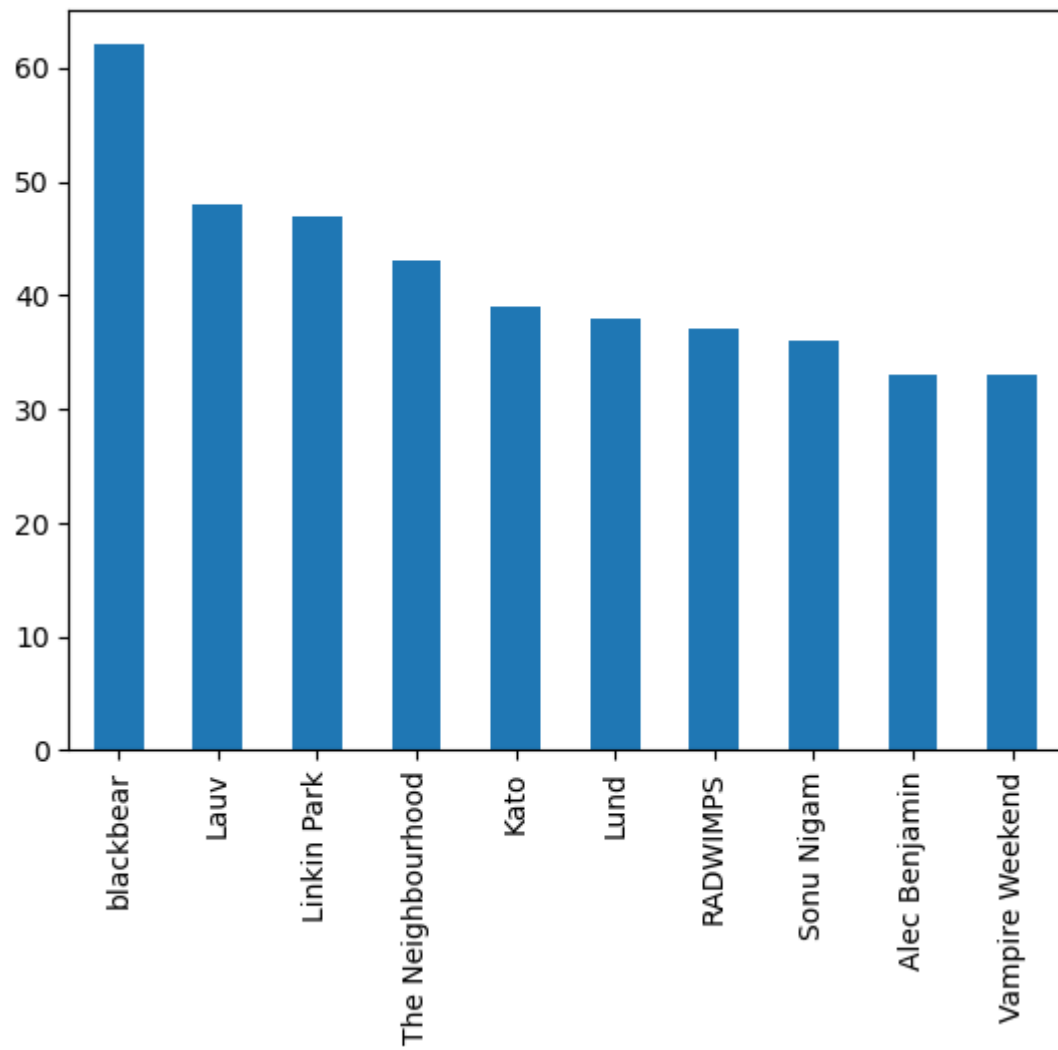


Figure 3: Top 10 Genre

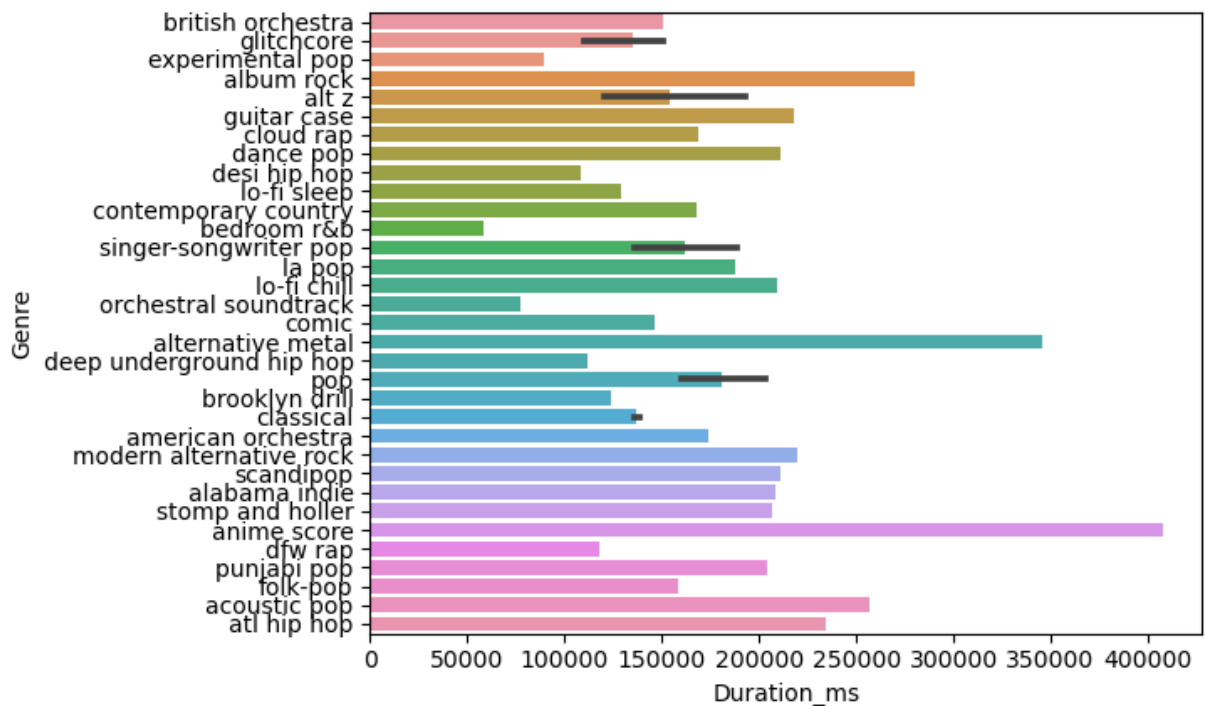
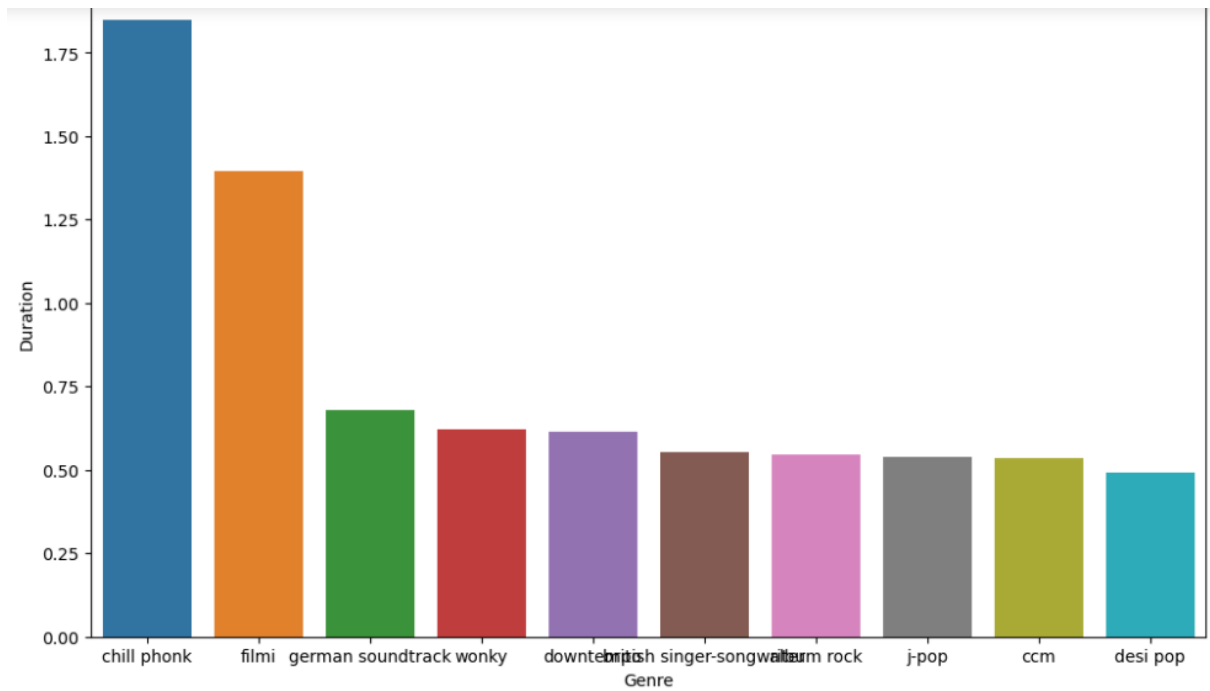


Figure 4: Top 10 Played Songs

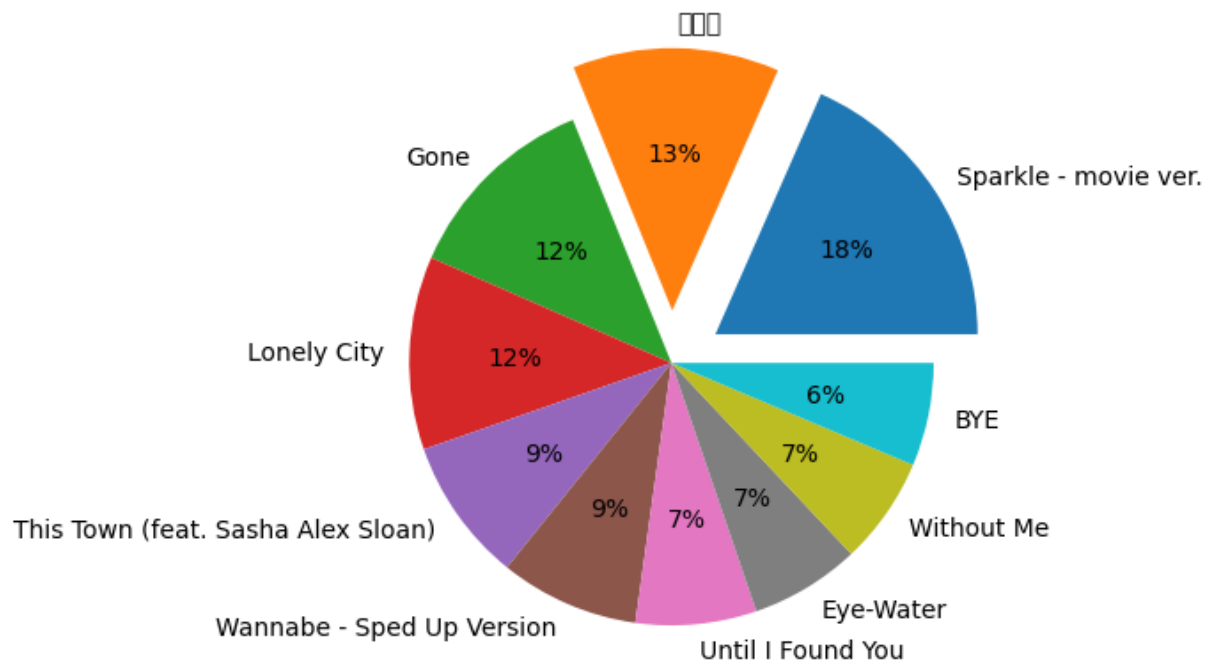


Figure 5: Top 10 Duration Songs

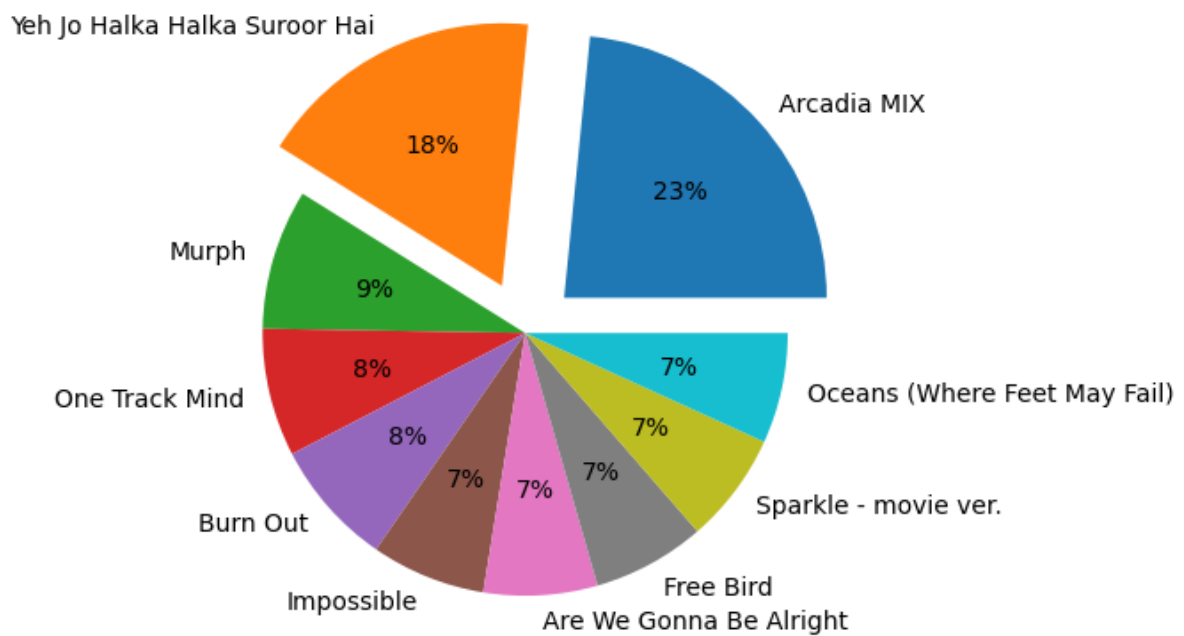


Figure 6: Relation Between Track Name and Danceability

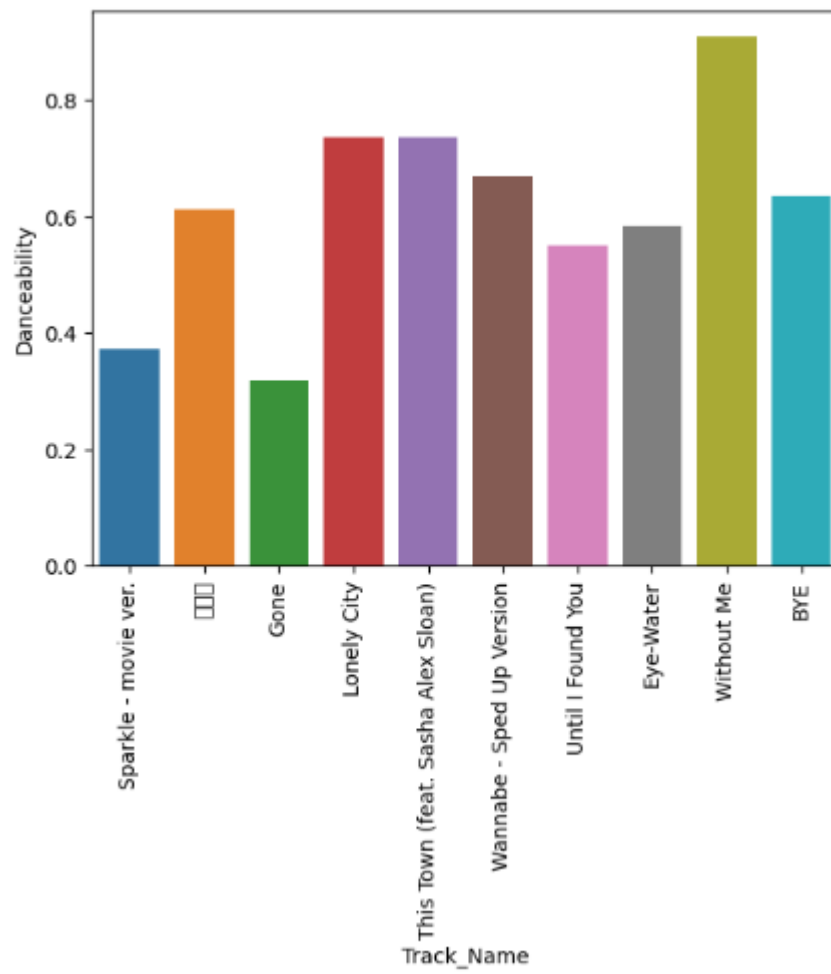
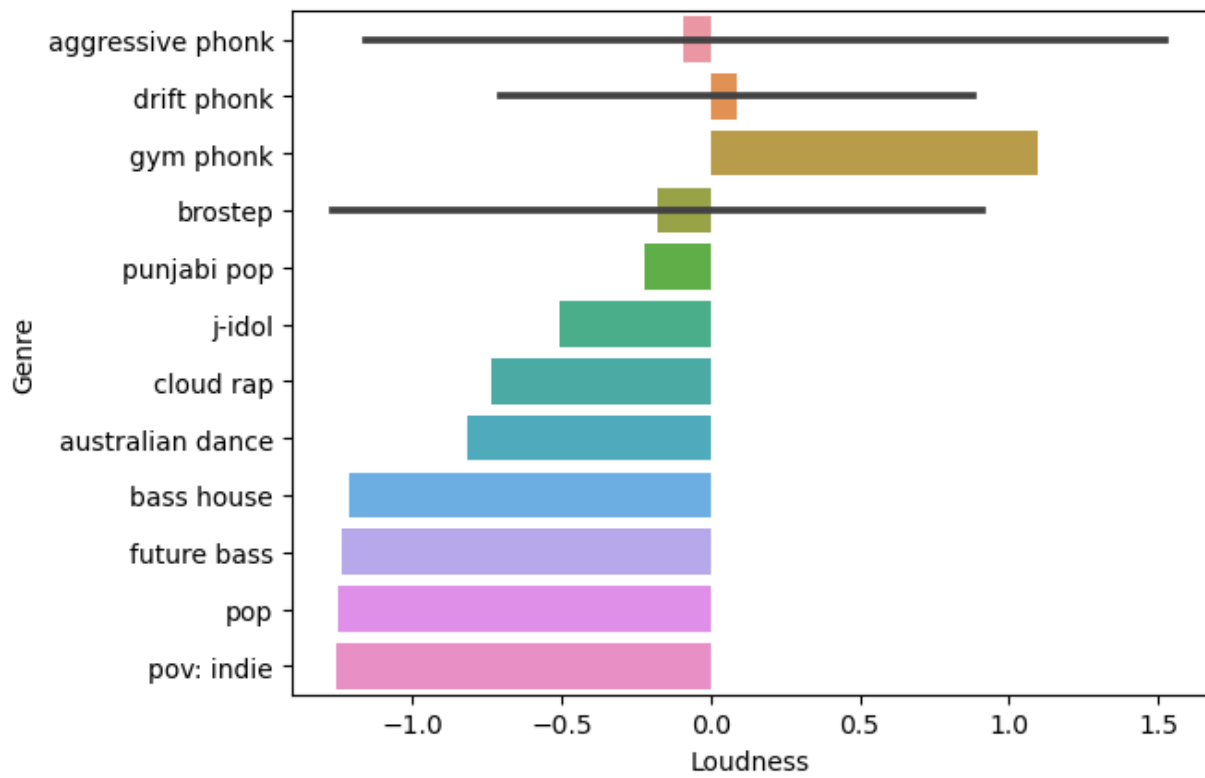


Figure 7: Relation Between Genre and Loudness



5. CONCLUSION

The development and deployment of the Spotify Song Attributes have culminated in the creation of a robust and user-centric platform that empowers users to come across various songs. This project sought to address the need for a modern Spotify Song Application, and the outcomes achieved through its implementation underscore its success. Our project is only a humble venture to satisfy the needs in an enterprise to manage their project work. Several user-friendly coding has also adopted. This package shall prove to be a powerful package in satisfying all the requirements of the enterprise in analyzing the data. The objective of program is to provide a framework that enables the organization to make reasonable business-oriented decisions within a limited time frame.

This project's approach is for the development and utilization of technologies that have resulted in a versatile tool that can adopt to the evolving Spotify Song Application. As the project continue to evolve, future enhancements and extensions hold the potential to elevate the system's functionality and impact even further.

In conclusion, the Spotify Song Attributes exemplifies the possibilities that arise when innovative solution are harnessed to address real-world challenges, enhancing the way users recommend and play their songs.