****

**Spotify Analysis**

**Submitted by: Submitted to:**

**Nalleswaran N-12213255 Dr. Punam Rattan**

**Kalepu poorna sai Nagesh-12219945**

**Kurugonda Venu-12212481.**

**Table of content:**

|  |  |  |
| --- | --- | --- |
| S.l | Name | Page.no |
|  | Introduction | 3 – 4 |
|  | Review of literature | 4 - 5 |
|  | Methodology | 5 – 27 |
|  | 1. Questionnaire For Data Collection | 5 |
|  | 1. Algorithms Applied | 5 – 6 |
|  | 1. Datasets | 6 – 7 |
|  | 1. Data Cleaning | 8 – 14 |
|  | 1. Visualization Techniques Applied | 14 – 27 |
|  | 1. Gantt Chart | 27 |
|  | Analysis Result | 27-28 |
|  | Coding and Visualization | 1. – 45 |

1. **INTRODUCTION:**

Spotify is a digital music, podcast, and video service that gives you access to millions of songs and other content from creators worldwide. Basic functions such as playing music are free, but you can also upgrade to Spotify Premium.

Whether you have Premium or not, you can:

1. Get recommendations based on your taste
2. Build collections of music and podcasts

Spotify is available across a range of devices, including computers, phones, tablets, speakers, TVs, and cars, and you can easily transition from one to another with Spotify Connect.

**Spotify Analysis:**

Spotify analysis typically refers to the examination of data collected by the Spotify music streaming service. Spotify gathers vast amounts of data on user listening habits, preferences, and behaviors. This data includes information such as which songs users listen to, how often they listen to them, which artists they follow, playlists they create, and more.

Analyzing this data can provide valuable insights into trends in music consumption, popular genres, emerging artists, and user preferences. It can also help music labels, artists, and Spotify itself make informed decisions regarding marketing strategies, content recommendations, and playlist curation.

Spotify analysis may involve techniques such as data mining, statistical analysis, machine learning, and data visualization to extract meaningful patterns and insights from the large datasets generated by the platform. It can be used for various purposes, including market research, content optimization, and personalization of the user experience.

**2)** Review of literature:

**Introduction:**

The advent of music streaming platforms has revolutionized the way we consume music, with Spotify emerging as a dominant player in the industry. Understanding user behavior, recommendation systems, and the impact on the music industry are pivotal in comprehending the dynamics of Spotify analysis.

**User Behavior Analysis:**

Research on user behavior within the Spotify ecosystem has shed light on various aspects of engagement and interaction. Studies have explored patterns such as frequency of listening, playlist creation, and response to personalized features like Discover Weekly. Factors influencing user retention and satisfaction have also been investigated, providing insights into enhancing the user experience.

**Recommendation Systems:**

Spotify's recommendation system lies at the core of its user experience, employing algorithms to deliver personalized music suggestions. Research has evaluated the efficacy of collaborative filtering, content-based filtering, and hybrid approaches in generating recommendations. Understanding the nuances of these algorithms is crucial for enhancing user engagement and satisfaction.

**Music Data Mining and Analysis:**

The vast amount of data generated by Spotify presents a rich opportunity for analysis. Researchers have leveraged data mining and machine learning techniques to classify music genres, model artist similarity, and predict song popularity. Challenges such as data bias and scalability have been addressed, paving the way for deeper insights into music consumption patterns.

**Impact on Music Industry:**

The rise of Spotify and other streaming platforms has had profound implications for the music industry. Traditional revenue streams such as album sales and radio airplay have been disrupted, prompting a reevaluation of distribution and promotion strategies. Understanding these shifts is essential for stakeholders to adapt and thrive in the digital age.

**3) METHODOLOGY :**

1. **QUESTIONNAIRE FOR DATA COLLECTION:**
2. Are you using Spotify
3. Full Name
4. Age
5. Gender
6. What is the Spotify usage period?
7. which device you are using?
8. What is the spotify\_subscription\_plan
9. Do you have Premium Account?
10. Which plan will you prefer.
11. which type of content you will listen?
12. which type of genres do you like?
13. Did you Satisfied or not?
14. **ALGORITHMS APPLIED**

**Kmeans Algorithm :**

K-Means Clustering is an Unsupervised Learning algorithm, which groups the unlabeled dataset into different clusters. Here K defines the number of pre-defined clusters that need to be created in the process, as if K=2, there will be two clusters, and for K=3, there will be three clusters, and so on.

It is a centroid-based algorithm, where each cluster is associated with a centroid. The main aim of this algorithm is to minimize the sum of distances between the data point and their corresponding clusters.

The working of the K-Means algorithm is explained in the below steps:

**Step-1:** Select the number K to decide the number of clusters.

**Step-2:** Select random K points or centroids. (It can be other from the input dataset).

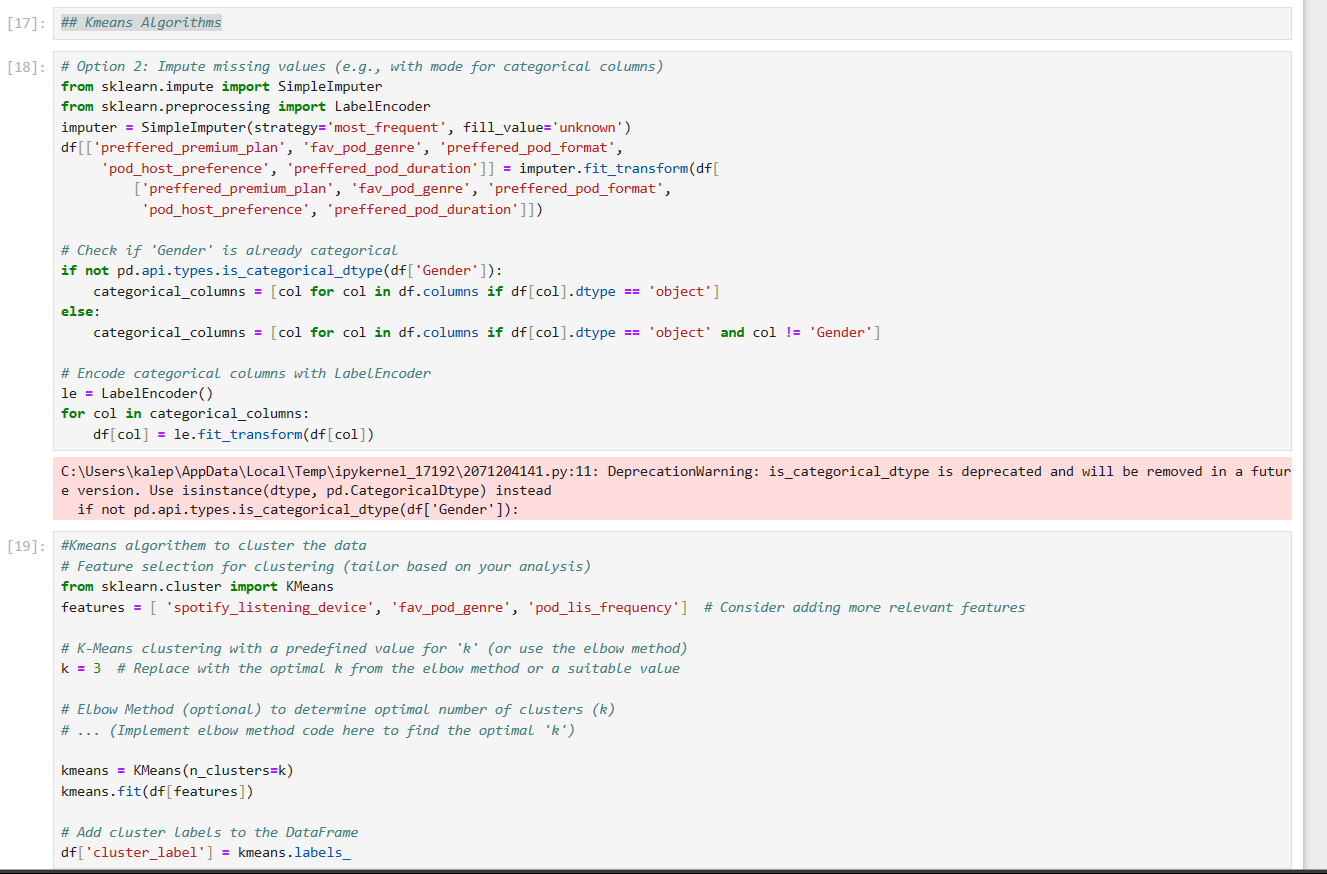
**Step-3:** Assign each data point to their closest centroid, which will form the predefined K clusters.

**Step-4:** Calculate the variance and place a new centroid of each cluster.

**Step-5:** Repeat the third steps, which means reassign each datapoint to the new closest centroid of each cluster.

**Step-6:** If any reassignment occurs, then go to step-4 else go to FINISH.

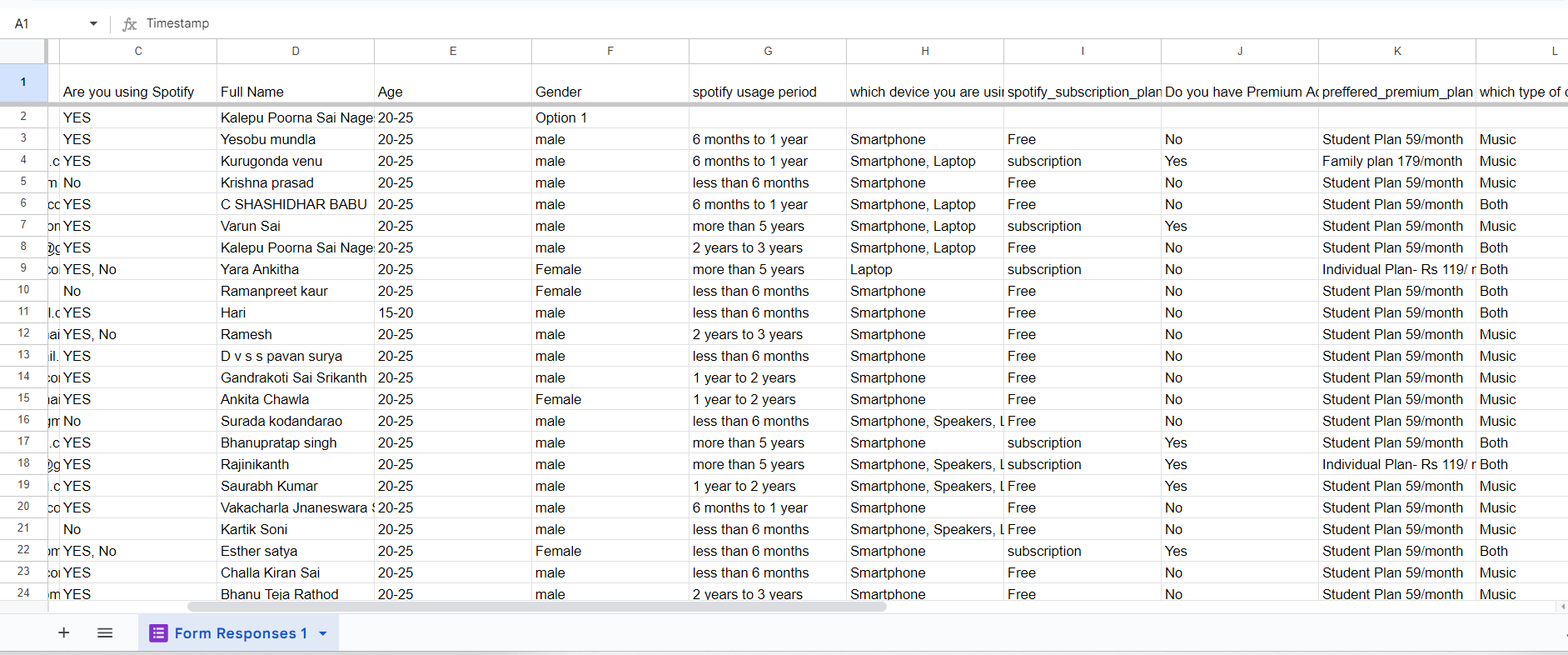
**Step-7**: The model is ready.



1. **DATA SET**

****1.First data set will be collected from the kaggele, the data set look like this.

1. The second data is collected from the user by using Google Forms. the data set looks like this.



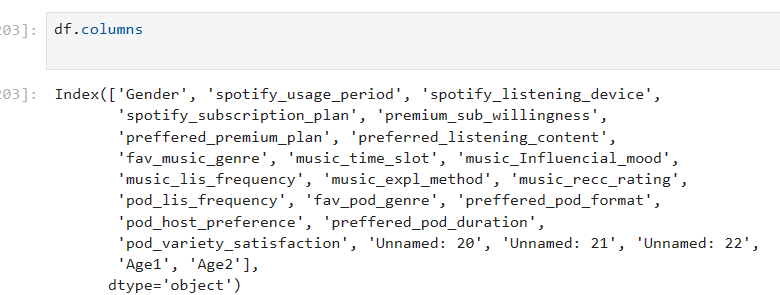
1. **DATA CLEANING:**
2. **Finding first 5 Rows 1 by using .head()**



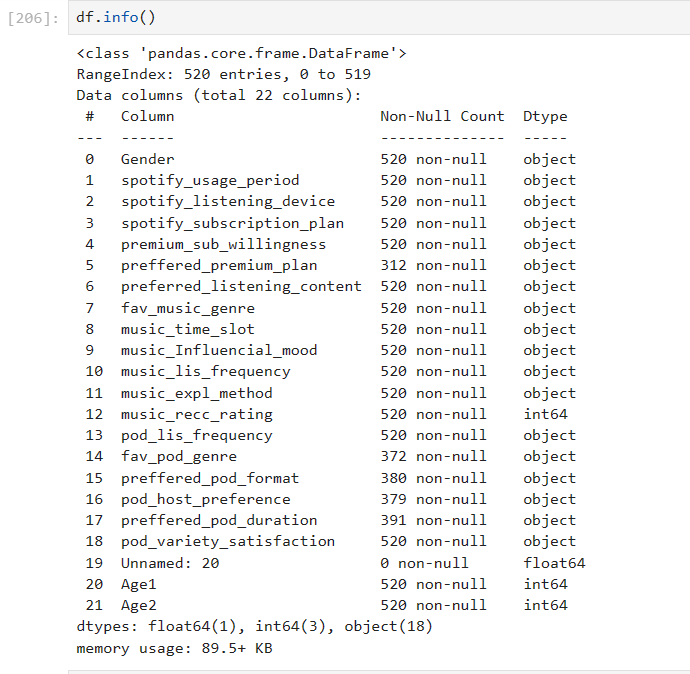
1. **Dropping Age column :**
2. **Shape of the dataset :**

****

1. **Columns in the Dataset :**

****

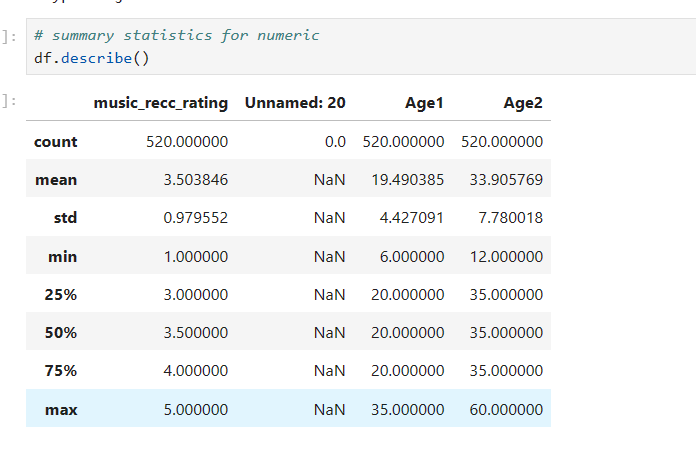
1. **Information about the Dataset :**

****

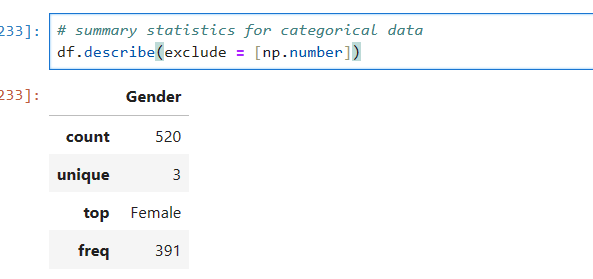
1. **DataTypes of Every column :**

****

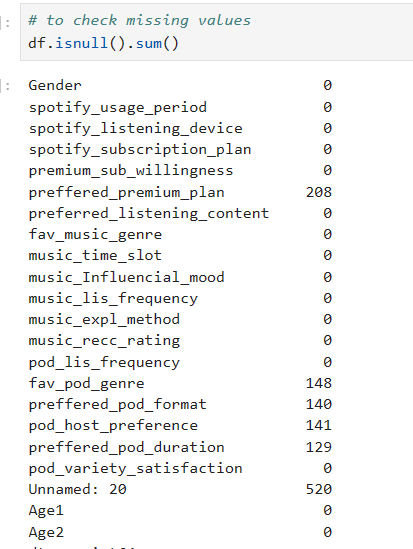
1. **Summary of Numeric Columns :**

****

1. **Summary of Categorical Values :**

****

1. **Finding the Null values :**

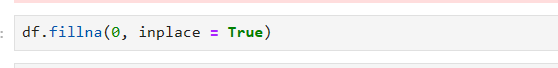
****

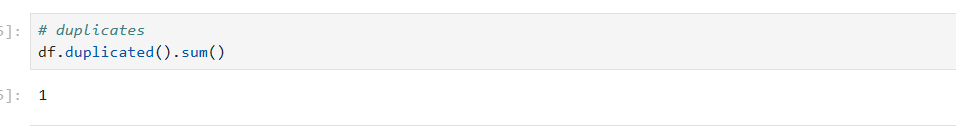
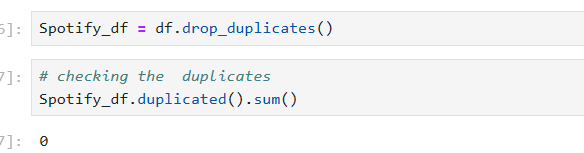
1. **Filling Null Values :**
2. **Filling null Values as NA.**

****

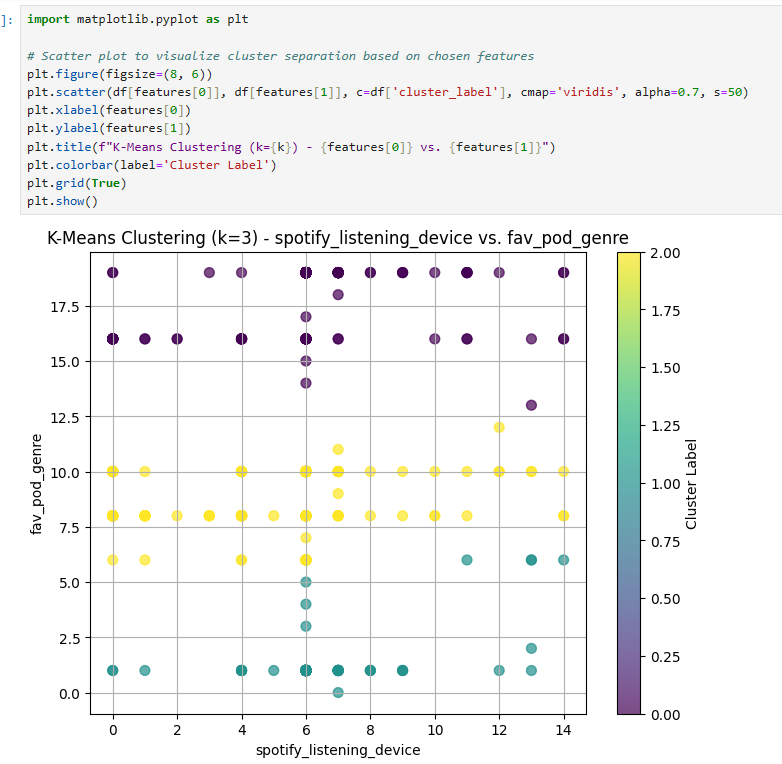
****

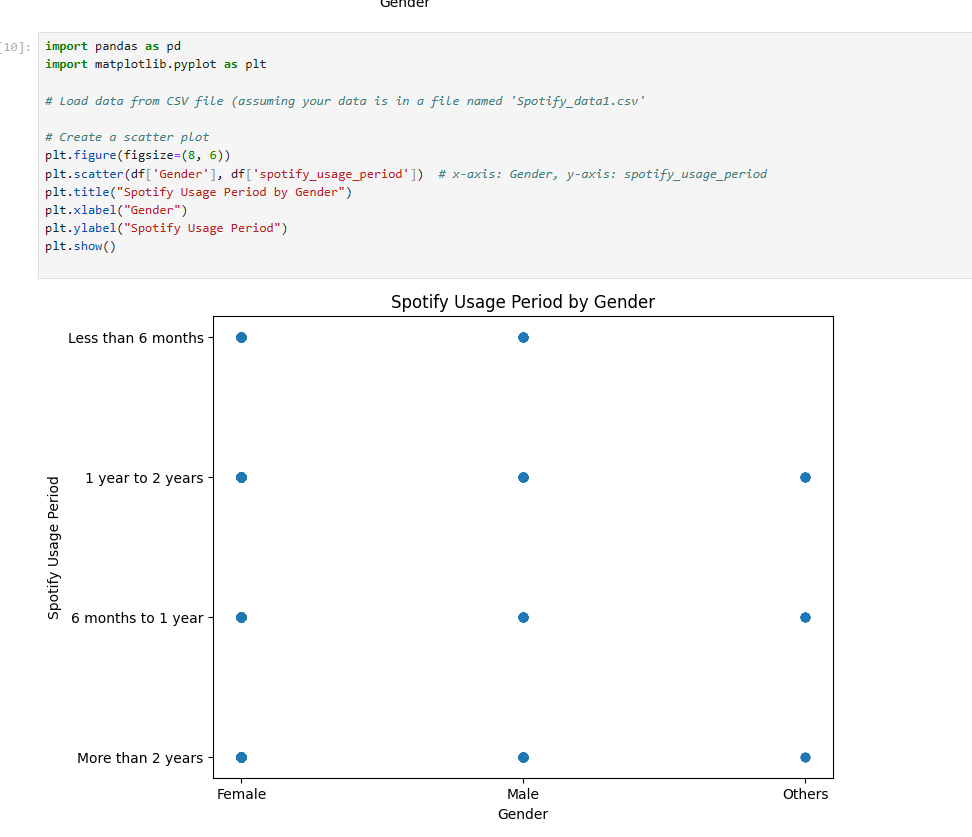
1. **Filling the Null Values with “Zero”**



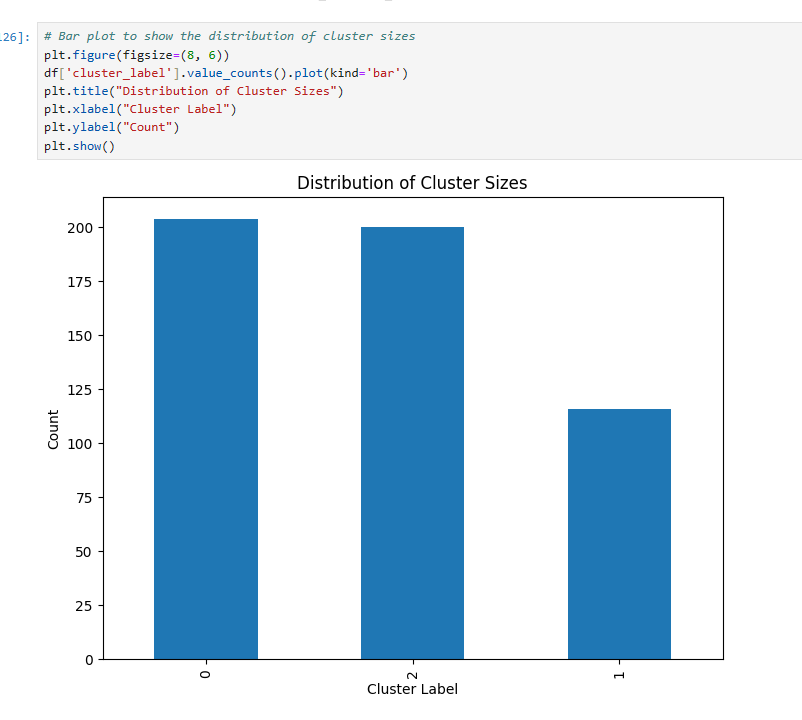
1. **Finding the Duplicate values :**
2. **Dropping the Duplicate value:**
3. **Visualization Techniques Applied :**

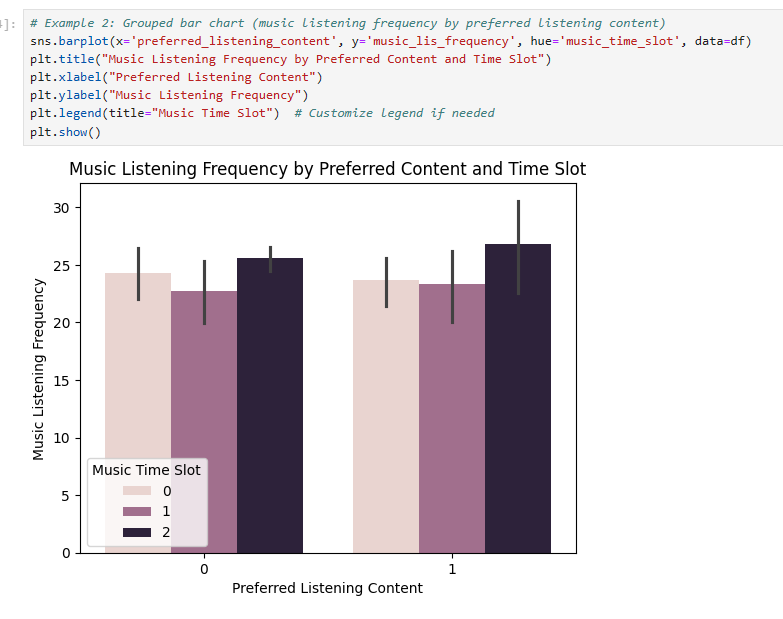
**Scatter Plot :**

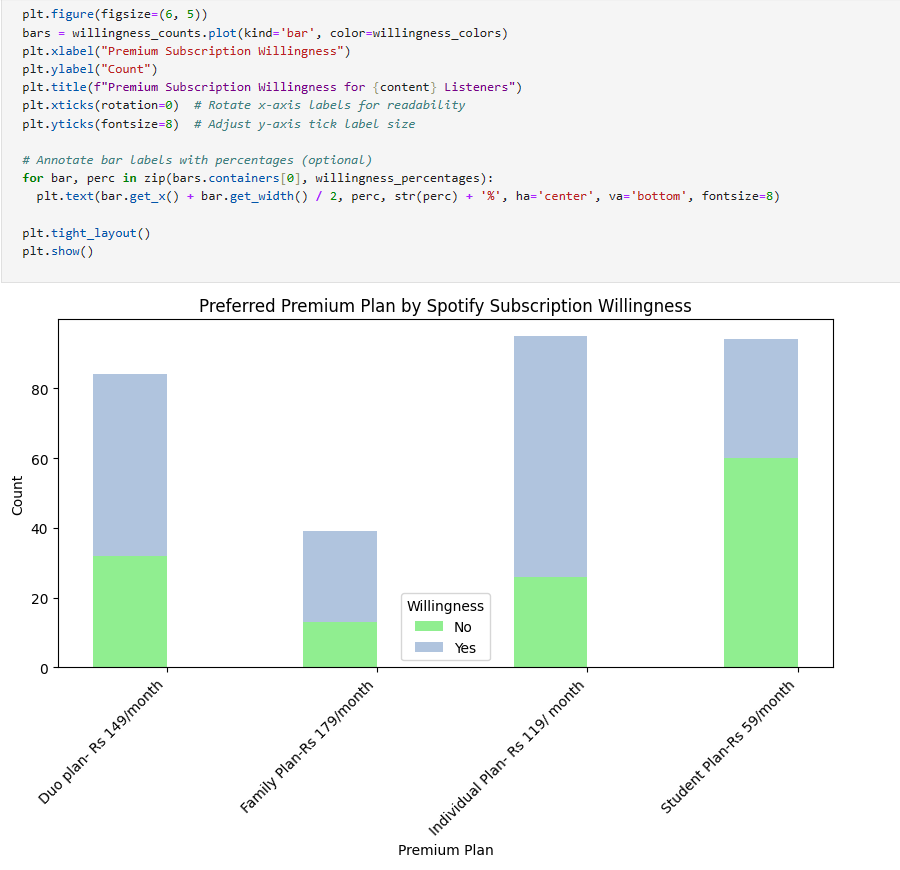
****



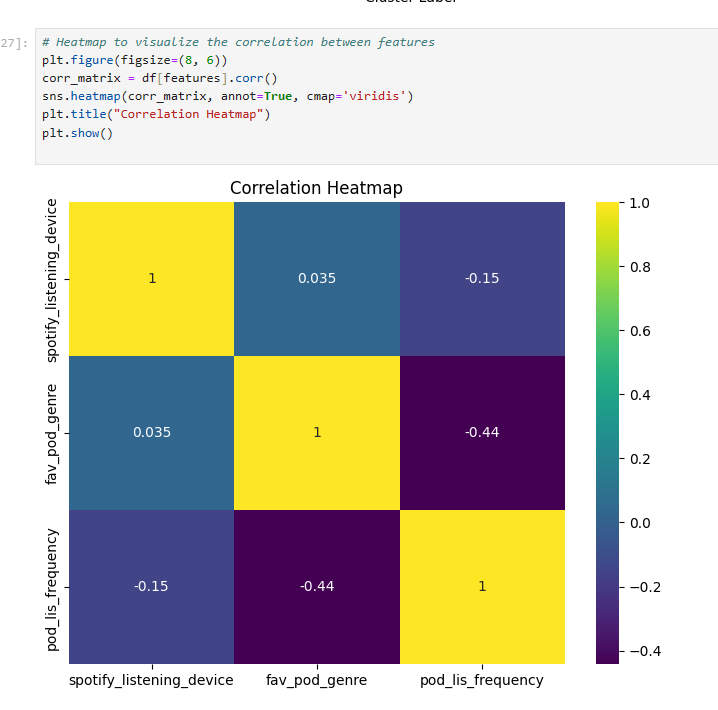
**Bar Plot :**



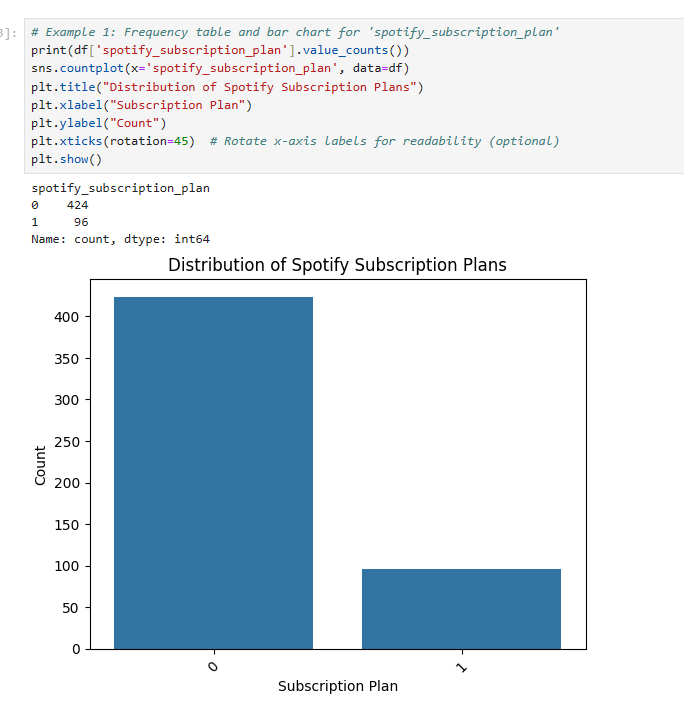




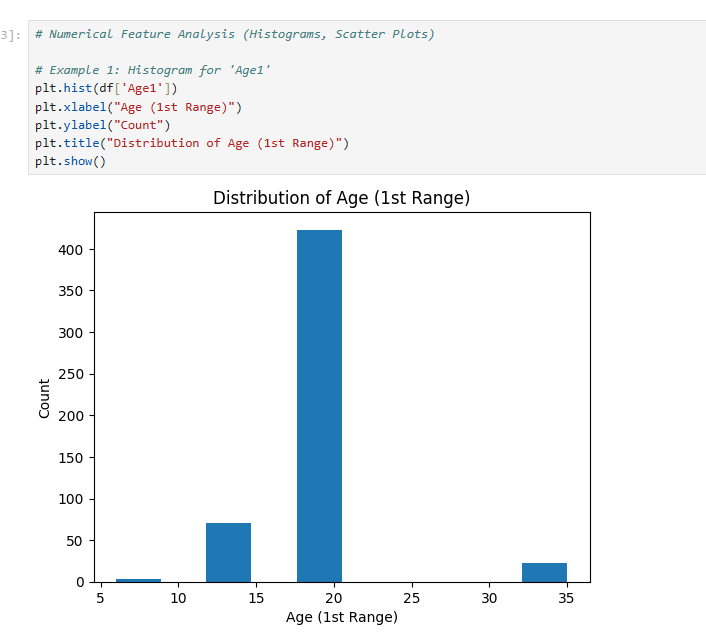
**Heat Map :**

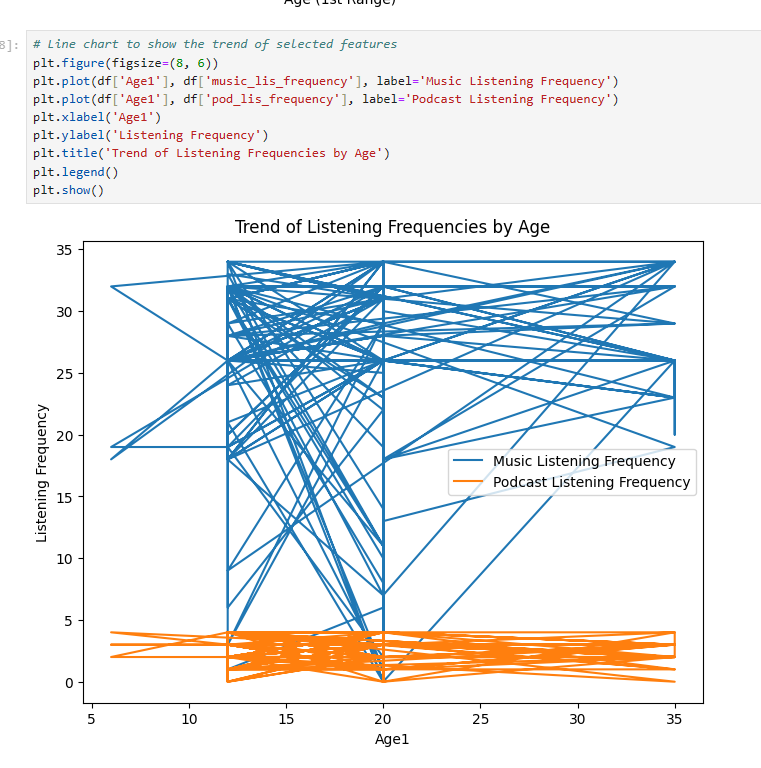


**Count Plot :**

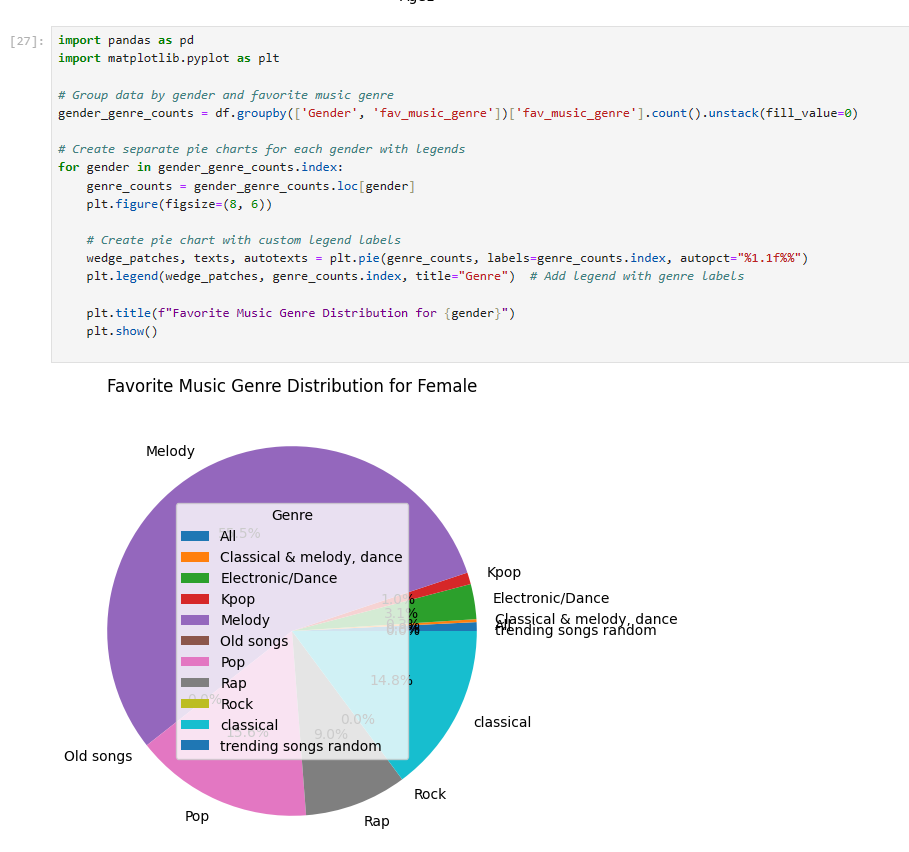


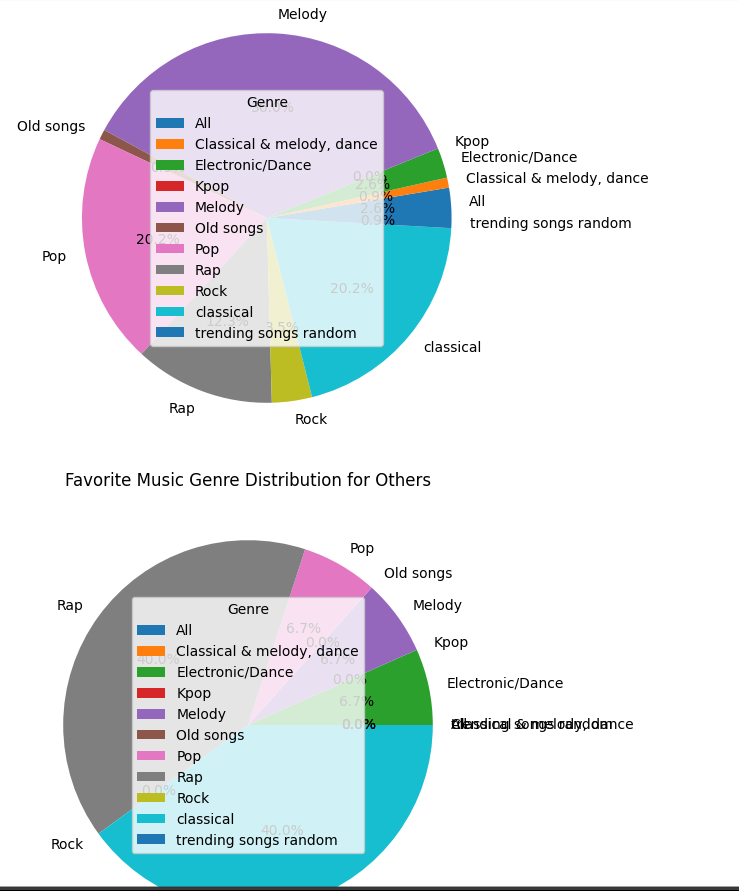
**Histrogram :**

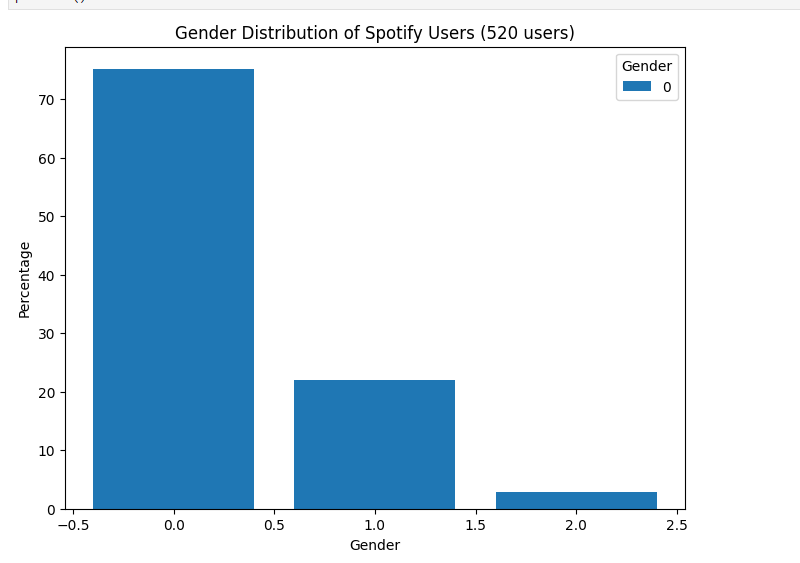


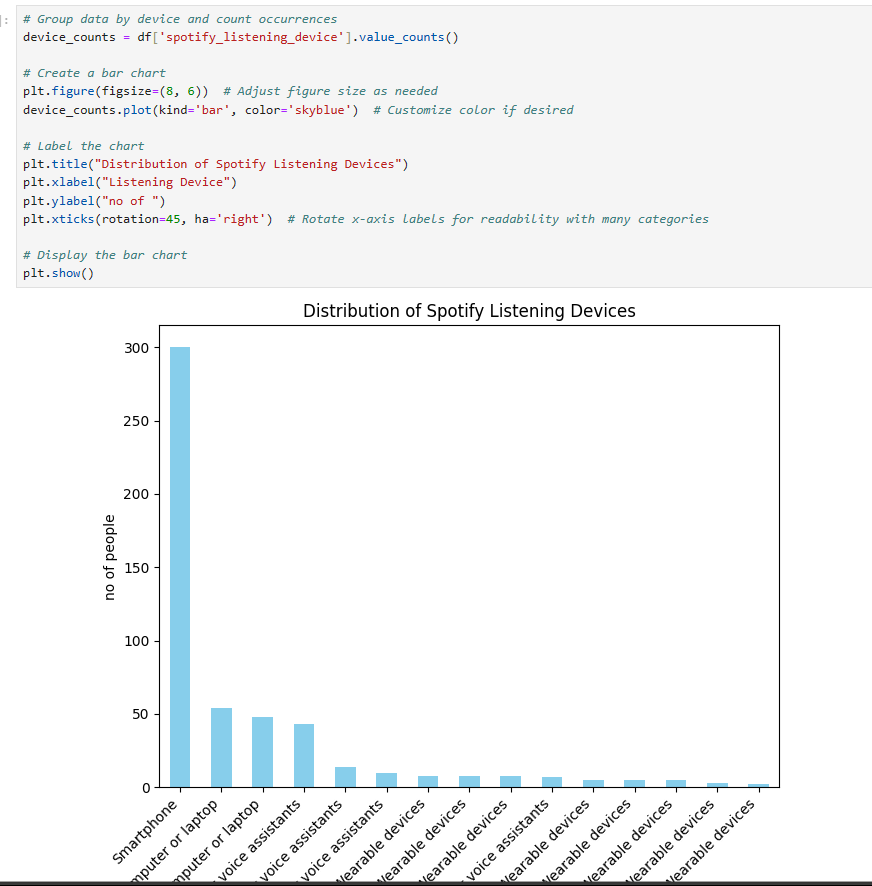


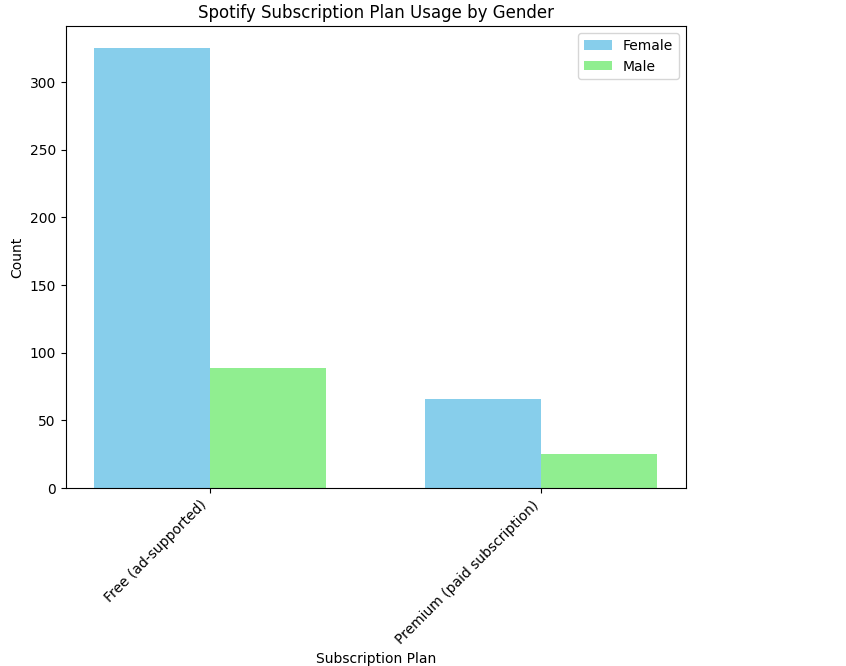
**Pie chart :**

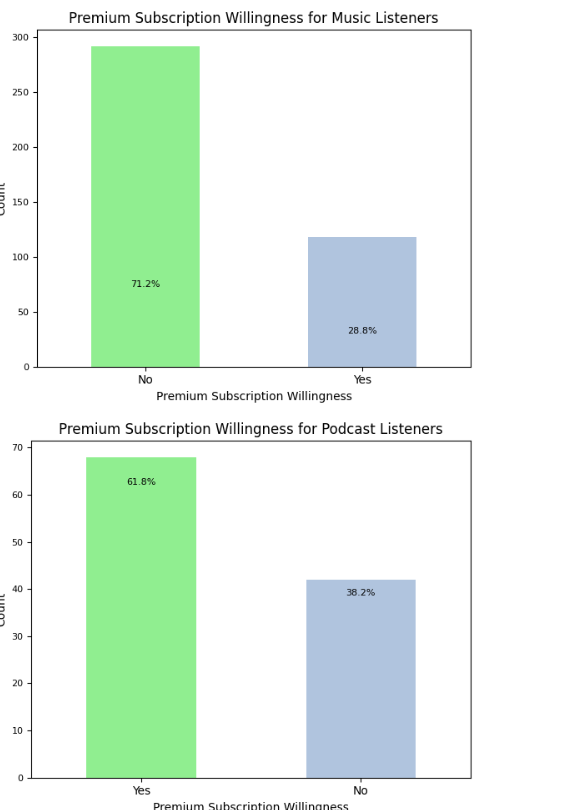




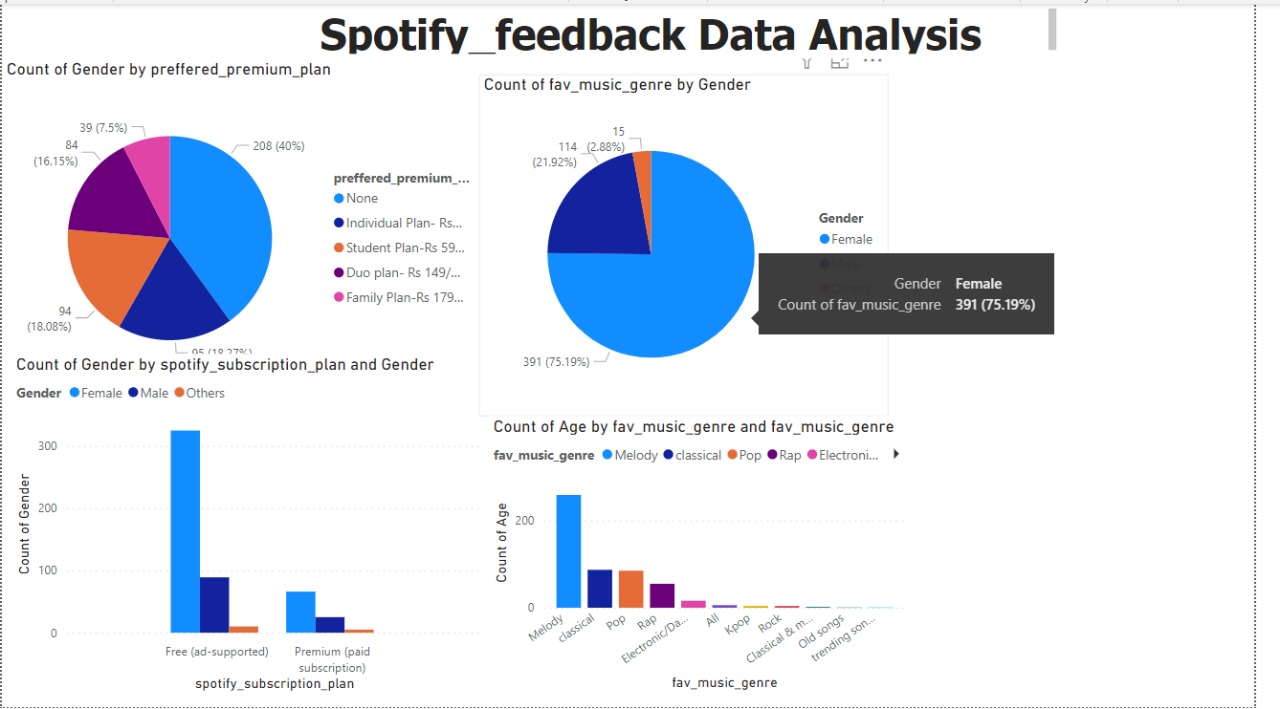




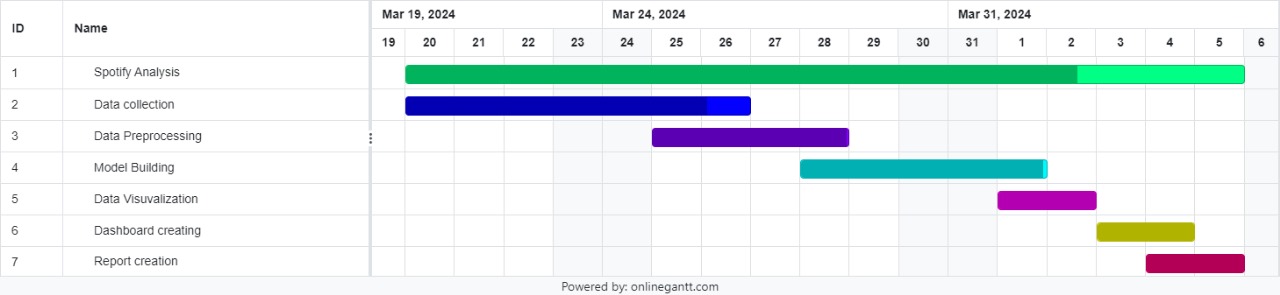




**Power Bi Dashboard :**



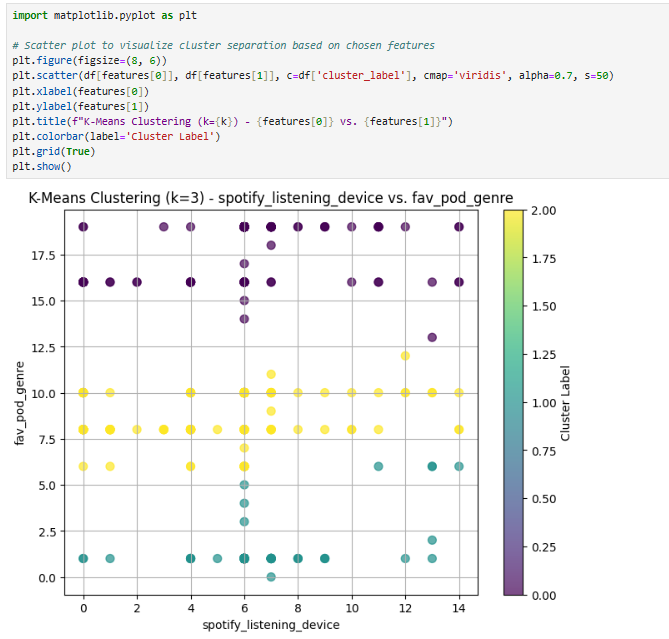
1. **Gantt Chart :**

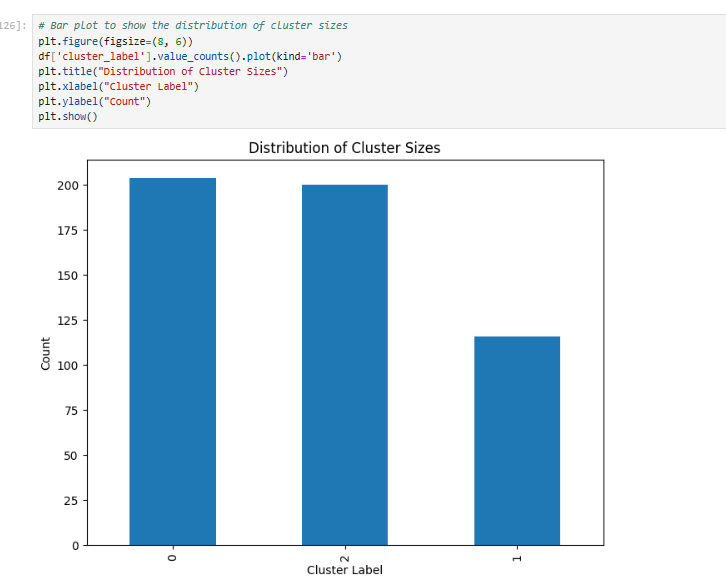
****

**4) Result Analysis :**

Result analysis of Spotify analysis involves examining the insights, trends, and patterns uncovered through data exploration, visualization, and statistical methods applied to Spotify's music streaming data. This analysis can reveal valuable information such as user preferences, popular genres, emerging artists, geographical trends, and the impact of various factors on music consumption behavior. By interpreting these results, Spotify can make informed decisions regarding content curation, personalized recommendations, marketing strategies, and platform enhancements to better serve its users and stay competitive in the music streaming industry

**5) Coding +Visualization**

****

****

