**Course Name - EDA Project**

**Course Code- INT 353**

**Continuous Assessment-III**

**Topic – Spotify Song Attributes**

Submitted by **Aman Verma 12114325**

### RK21UTB70

Submitted to

### Shivangini Gupta

**School of Computer Science & Engineering Lovely Professional University, Phagwara**

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| Serial NO. | Topic | Page No. |
| 1. | Introduction | 3 |
| 2. | Domain Knowledge | 4 - 5 |
| 3. | Reasons for choosing this dataset | 6 – 7 |
| 4. | Libraries Used | 8 – 9 |
| 5. | Data Description | 10 - 11 |
| 6. | Data cleaning and Data exploration | 12 |
| 7. | Univariate Analysis | 13 – 16 |
| 8. | Bivariate Analysis | 17 – 18 |
| 9. | Multivariate Analysis | 19 |
| 10. | Distributions | 20 |
| 11. | Hypothesis Testing | 21 |
| 12. | Findings and insights | 22 - 23 |
| 13. | Recommendations | 24 |
| 14. | Conclusion | 25 |
| 15. | References | 26 |
| 16. | Acknowledgement | 27 |
| 17. | Project Code | 28 |

**1.INTRODUCTION**

**Exploratory Data Analysis refers to the critical process of performing initial**

**investigations on data so as to discover patterns, to spot anomalies, to test hypothesis and to check assumptions with the help of summary statistics and graphical representations.**

#### **Dataset Overview**

The Spotify Song Attributes Dataset contains detailed information about songs available on the Spotify platform. This dataset is compiled from a diverse collection of music genres and artists. Key attributes within the dataset include:

1. **Track ID**: A unique identifier for each song on Spotify.
2. **Track Name**: The title of the song.
3. **Artist Name**: The name of the artist(s) who performed the song.
4. **Album Name**: The title of the album to which the song belongs.
5. **Popularity**: A metric representing the song's popularity on Spotify.
6. **Audio Features**: These include quantitative measures that describe various aspects of the song, such as danceability, energy, instrumentalness, loudness, valence, tempo, and more.

**Objectives of EDA:** The objective of this project is to perform exploratory data analysis (EDA) on a Spotify songs dataset to understand the music that people listen to. We will explore the following aspects of the dataset:

* The most popular songs, artists, and genres
* Trends in listening habits over time
* The relationship between people's musical preferences and their demographics, social networks, and other factors.

**2.DOMAIN KNOWLEDGE**

#### **Spotify: A Music Streaming Platform**

Spotify is a leading music streaming platform that allows users to access a vast library of songs, albums, and playlists. It offers both free and premium subscription tiers, making music easily accessible to millions of users worldwide. Spotify provides a rich user

experience by offering personalized playlists, recommendations, and a wide range of features for music enthusiasts.

#### **Audio Features**

Understanding the significance of audio features is vital for interpreting the dataset accurately:

* **Danceability**: These metric measures how suitable a song is for dancing based on factors like rhythm, tempo, and beat stability. A high danceability score suggests a more danceable song.
* **Energy**: Energy quantifies the intensity and activity of a song. High-energy songs tend to be more dynamic and livelier.
* **Loudness**: This attribute indicates the overall volume of a song, measured in decibels (dB).
* **Valence**: Valence represents the positivity or happiness conveyed by a song. Higher valence values suggest a more positive emotional tone.
* **Tempo**: Tempo refers to the speed of a song, measured in beats per minute (BPM).
* **Instrumentalness**: This metric reflects the presence of vocals in a song. A high instrumentalness score indicates minimal or no vocals.
* **Acousticness**: Acousticness quantifies the acoustic nature of a song, with higher values indicating a more acoustic sound.

#### **Applications of the Dataset**

The Spotify Song Attributes Dataset has several practical applications:

1. **Music Recommendation**: Spotify leverages this dataset to recommend songs to users based on their listening history and preferences. Audio features like danceability and valence play a crucial role in suggesting suitable tracks.
2. **Playlist Generation**: Users can create playlists or explore curated playlists based on audio features. For example, a user can create a "Relaxing Sunday Morning" playlist by selecting songs with low energy and acoustic qualities.
3. **Genre Classification**: Machine learning models can classify songs into genres by analyzing audio features. This aids in organizing and categorizing the vast music library.
4. **Mood Analysis**: Valence and energy attributes help in determining the mood of a song, making it easier to create mood-based playlists like "Happy Vibes" or "Chill Out."
5. **Music Industry Insights**: Record labels and artists use this data to gain insights into the characteristics of successful songs, assisting in marketing and production

decisions.

**3. REASONS FOR CHOOSING THIS DATASET**

1. **Relevance to the Research or Project Goals:** Researchers or data analysts may choose this dataset because their project or research objectives are related to music, music streaming platforms, or music analysis. Understanding song attributes can be valuable for various applications, such as music recommendation, genre

classification, or mood analysis.

1. **Interest in Music**: Individuals with a passion for music may naturally gravitate towards a dataset that allows them to explore and analyze music-related data. This dataset provides an opportunity to combine data analysis skills with a personal

interest in music.

1. **Applications in Data Science and Machine Learning:** The Spotify Song Attributes Dataset is rich in audio features, making it suitable for machine learning and data

science projects. Researchers and data scientists often choose datasets with a variety of attributes to develop predictive models, recommendation systems, or

classification algorithms.

1. **Practical Applications**: This dataset has practical applications in the music industry. Record labels, music artists, and streaming platforms like Spotify can use this data to gain insights into music trends, user preferences, and song characteristics. For example, artists may want to understand what attributes make a song popular.
2. **Educational Purposes:** This dataset can serve as an educational resource for students and aspiring data scientists. Analyzing real-world data like this can help learners practice data analysis techniques and gain valuable experience.
3. **Availability and Accessibility:** The Spotify Song Attributes Dataset may be readily available and accessible to the public, making it an attractive choice for research or analysis. Ease of access and availability of data is often a deciding factor.
4. **Innovation and Creativity**: Some individuals and organizations are interested in exploring creative and innovative ways to use music data. This dataset provides opportunities for creative data projects, such as generating playlists based on mood or building AI-powered music recommendation systems.
5. **Data Visualization**: The dataset's audio features offer opportunities for creative data visualization projects. Visualizations can help convey insights about music in an engaging and informative way.

**4. LIBRARIES USED**

The libraries that I have used in this project are:

* NumPy
* Pandas
* Matplotlib
* Seaborn
* SciPy

**NumPy**

Numpy is a python library to perform numerical calculations.

It provides high-level mathematical functions to work with arrays.

**Pandas**

Pandas is used to import a dataset and provide operations to work on that dataset.

Pandas library is used for data analysis and manipulation.

**Matplotlib**

It is a library used for data manipulation.

Various types of plots like line plot, bar chart , histogram and scatter plot can be

plotted using this library.

**Seaborn**

Seaborn is a also a data visualization library but built on top of matplotlib.

Seaborn provides more specialized visualizations, such as heatmap, correlation

matrices, and regression plot.

**SciPy**

SciPy is an open-source Python library which is used to solve scientific and

mathematical problems.

It is built on the NumPy extension and allows the user to manipulate and visualize

data with a wide range of high-level commands.

Approach to solve the problem:

1. First of all, import a dataset from Kaggle or Machine Learning Repository.
2. Then implement summary statistics see how the mean standard deviation is

behaving.

1. Through summary we can see that columns have negative value for some records.
2. After removing those invalid values univariate analysis can be implemented

through bar graph and histogram.

1. After getting insight from univariate, bivariate analysis and multivariate

analysis can also be implemented which will give us detail how more than one

variables are related to each other.

1. Then using histplot () visualize which distribution is the dataset following.
2. At last, develop some hypothesis and test whether the dataset is following

those hypotheses or not.

**5.DATA DESCRIPTION**

The Spotify Song Attributes Dataset is a comprehensive collection of information about songs available on the Spotify music streaming platform. Spotify Song Attributes Dataset, gaining a thorough understanding of the data is essential for meaningful analysis and

interpretation.

* 1. **trackName:** This attribute represents the title of the track, which is a fundamental identifier for a song.
  2. **artistName:** It refers to the name of the artist or band responsible for creating the track. In the music industry, artists have unique styles and fan bases.
  3. **msPlayed:** This attribute indicates the duration in milliseconds that a track was played. Understanding this helps in analyzing user listening habits.
  4. **genre**: Genre classification is a fundamental aspect of music categorization, reflecting the stylistic characteristics of a song or artist.
  5. **danceability:** A measure of how suitable a track is for dancing. Higher values indicate tracks with rhythmic elements that encourage dancing.
  6. **energy:** This attribute quantifies the intensity and liveliness of a track. High-energy songs are typically more dynamic and engaging.
  7. **key**: The key of a track (e.g., C, D, E) provides information about the musical scale and tonality of the song.
  8. **loudness**: Loudness represents the overall volume of the track in decibels (dB). It plays a crucial role in audio mixing and perception**.**
  9. **mode:** Mode indicates whether a track is in a major (1) or minor (0) key. It influences the emotional character of the music.
  10. **speechiness**: Speechiness reflects the presence of spoken words or non-musical elements in the track. It can be useful for distinguishing spoken-word content from music.
  11. **acousticness:** This attribute quantifies the acoustic nature of the track. High values

indicate a more natural, acoustic sound, while low values suggest electronic or synthesized elements.

* 1. **instrumentalness**: It represents the probability of the track being instrumental (i.e., without vocals). This is useful for identifying instrumental music.
  2. **liveness:** Liveness indicates the presence of a live audience in the track. It helps distinguish live recordings from studio recordings.
  3. **valence:** Valence measures the musical positiveness or happiness conveyed by the track. High valence values indicate a more positive or uplifting mood.
  4. **tempo**: Tempo refers to the speed of the track in beats per minute (BPM). It influences the rhythm and pace of the music.
  5. **type:** This attribute specifies the type of Spotify track, which can include "track," "episode," or other categories.
  6. **id:** It serves as a unique identifier for the track, facilitating data management and retrieval.
  7. **uri:** The Spotify URI is a link to the track on the Spotify platform, allowing for easy access to the song.
  8. **track\_href**: This attribute is a link to the Spotify Web API endpoint for the track, enabling further data retrieval and analysis.
  9. **analysis\_url:** It provides a link to the audio analysis of the track, which includes detailed information about the song's acoustic properties.
  10. **duration\_ms:** This attribute specifies the duration of the track in milliseconds, allowing for precise measurement of track length.
  11. **time\_signature:** Time signature indicates the rhythmic structure of the track, such as 4/4 or 3/4 time. It affects the feel and groove of the music

**6. DATA CLEANING AND DATA EXPLORATION**

Before starting the analysis, first we need to clean and filter the data.

So first it need to checked that how many null values are present in each column.

Use d.is.null().sum() to check total null values in each column.

No null values were found .

Then d.drop\_duplicates() was used to drop duplicates if any duplicate value is

present.

After that invalid values were removed from 'msPlayed', 'danceability', 'energy', 'loudness', 'speechiness','acousticness', 'instrumentalness', 'liveness', 'valence', 'tempo', 'duration\_ms'

Then summary statistics was perform using d.describe() to find the count, mean,

standard deviation, min, max, 25 percentile, 50 percentile and 75 percentile of the

dataset.

After removing invalid values the size of data points were reduced to (4290,22) from

(10082,22).

**7. Univariate Analysis**

**Q1.** How diverse are the genres represented in the dataset (Top 50), and which genres are most prevalent?

Ans- Most prevalent genre being the-alt z, pop and filmi.

* alt z 328
* pop 301
* filmi 206
* dance pop 86
* singer-songwriter pop 82
* alternative metal 75
* anime lo-fi 68
* art pop 63
* drift phonk 62
* brostep 58
* modern alternative rock 56
* lo-fi study 55
* edm 50
* anime 49
* chill pop 47
* classical 46
* j-pop 46
* cloud rap 44
* la pop 42
* modern rock 40
* lo-fi sleep 37
* boy band 36
* lo-fi chill 35
* german soundtrack 34
* baroque pop 34
* modern indie pop 34
* icelandic indie 33
* desi hip hop 33
* desi pop 32
* bedroom pop 32
* hip hop 32
* alternative dance 31
* gen z singer-songwriter 31
* anime score 29
* classic bollywood 29
* dark r&b 26
* sad lo-fi 26
* canadian pop 25
* pov: indie 24
* electropop 24
* folk-pop 24
* album rock 23
* indie pop 23
* gym phonk 23
* alternative rock 22
* lo-fi beats 22
* canadian contemporary r&b 21
* glitchcore 20
* detroit hip hop 20
* australian pop 20

**Q2.** What are the most common keys for the tracks in the dataset?

Ans- The key 0.0, 1.0, 7.0, 2.0 and 9.0 are the major keys in the dataset.

**Q3.** Top 50 most popular artists in the dataset.

Ans- why mona 4.209253e+07

* George Michael 3.831582e+07
* Yoh kamiyama 3.769777e+07
* Austin Farwell 3.577789e+07
* ØZI 3.501439e+07
* The Honeysticks 3.421363e+07
* Blake Rose 2.863643e+07
* Michael Sembello 2.771202e+07
* Johann Pachelbel 2.733168e+07
* Ralph Alvern 2.442189e+07
* the engy 2.232900e+07
* Kygo 2.124032e+07
* Petit Biscuit 2.019390e+07
* Air Review 1.879505e+07
* Before You Exit 1.750644e+07
* Stephen Sanchez 1.624433e+07
* Chris Norman 1.590254e+07
* Kate Bush 1.584831e+07
* Farhan Saeed 1.537492e+07
* Christian Leave 1.537115e+07
* Absofacto 1.433911e+07
* Mumm-ra 1.410444e+07
* John-Robert 1.325516e+07
* Wolfmother 1.027873e+07
* a-ha 1.023117e+07
* Cigarettes After Sex 9.618610e+06
* Apparat 8.987626e+06
* Sarah Barrios 8.893614e+06
* Paper Idol 8.813955e+06
* Jax 8.794378e+06
* Jim Croce 8.674595e+06
* RADWIMPS 7.855428e+06
* Golden Features 7.832768e+06
* The Clash 7.808602e+06
* Eurythmics 7.747612e+06
* Max Coveri 7.651124e+06
* Tommee Profitt 7.555178e+06
* Loote 7.544652e+06
* Polaris 7.390426e+06
* Devon Cole 7.359670e+06
* luvbyrd 7.227135e+06
* Riria. 6.844262e+06
* Young Ejecta 6.693531e+06
* Wild Rivers 6.614394e+06
* Johann Sebastian Bach 6.560980e+06
* Babygirl 6.408913e+06
* Yunomi 6.394658e+06
* CHUNG HA 6.303704e+06
* TOTO 6.292962e+06
* Chelsea Cutler

**Q4.** Which artist has the most tracks in the dataset (Top 50)?

Ans- Blackbear- 62

**Q5.** Are major key tracks more common than minor key tracks?

Ans- minor-1.0 and major-0.0

1.0 2644

0.0 1646

**Q6.** How speechy are the tracks on average?

Ans- This value indicates the overall level of spoken words or lyrics in the music, with higher values suggesting a greater presence of speech.

Average Speechiness: 0.07832058275058275

**Q7.** How does the time signature vary across tracks?

Ans- By examining this bar chart, you can visually assess how the time signature varies across tracks in your dataset. The chart will show the distribution of different time signatures, allowing you to identify which time signatures are most common or prevalent among the tracks. In this dataset time signature -4.0 has the most values.

**Q8.** Which track has the longest duration and which has shortest duration?

Ans- Longest – Acardia Max

Shortest- [Stef]

**Q9.** What is the distribution of loudness across tracks?

Ans- you can visualize the distribution of loudness values across tracks. It will show you how loudness levels are distributed in your dataset, allowing you to identify common loudness ranges and any potential outliers.

**8. Bivariate Analysis**

**Q10.** Is there a correlation between 'energy' and 'loudness' attributes?

Ans- If the points tend to follow a clear upward or downward trend, it suggests a positive or negative correlation, respectively. If the points are scattered randomly, there may be little to no correlation between the two attributes.

In the dataset shows upward trend meaning it’s a positive trend.

**Q11.** How does the 'tempo' attribute vary across different genres?

Ans-The box plot displays the median, quartiles, and potential outliers in tempo for each genre, providing insights into the tempo characteristics of each genre in dataset.

Like pop, bedroom pop has 4 outliers each and each boxplot shows the mean, max and min values of the genre.

**Q12.** Trends in 'acousticness' across different genres.

Ans- Genres with higher bars have a higher mean 'acousticness,' indicating a greater presence of acoustic elements in the music of those genres. Conversely, genres with lower bars have a lower mean 'acousticness' and are likely to have a more electronic or non-acoustic sound. In dataset classical, lofi-sleep, anime-lofi and lofi beats have greater acoustic elements whereas genres like brostep, glitchcore are some with more electronic or non-acoustic sound.

**Q13.** Patterns in the 'time\_signature' attribute across different genres.

Ans.- The chart will show how common each time signature is within each genre, providing insights into the rhythmic characteristics of music in genres like alt z, album rock, etc.

**Q14.** The relationship between danceability and valence (positivity).

Ans- you can visually assess the relationship between 'danceability' and 'valence.' If the points tend to follow a clear pattern (e.g., moving from the bottom-left to the top-right), it suggests a positive correlation between the two attributes, indicating that more danceable tracks tend to be more positive or happy in mood. Conversely, if the points are scattered randomly, there may be little to no correlation between the two attributes.

**Q15.** How do song attributes like energy, danceability, and valence differ across genres (Top 50).

Ans- Energy has higher mean than the others, danceability has smaller box plot as compared to other 2 and valence has low mean. All three have outliers in them.

**Q16.** Are certain genres associated with higher tempos?

Ans- chillhop, uk alternative hip hop, singer songwriter pop, pop etc has highest tempos.

**Q17.** Are instrumental tracks more common than vocal tracks?

Ans- Yes,

Instrumental - 3958

Vocal - 332

**Q18.** Are live performances more prevalent in certain genres?

Ans- you can assess whether certain genres tend to have a higher prevalence of live performances compared to others. The chart will show the counts of live and non-live performances within each genre, helping you identify genres where live recordings are more common. Genres like alternative metal, indie poptisim, etc.

**Q19.** Correlation between 'energy' and 'valence'.

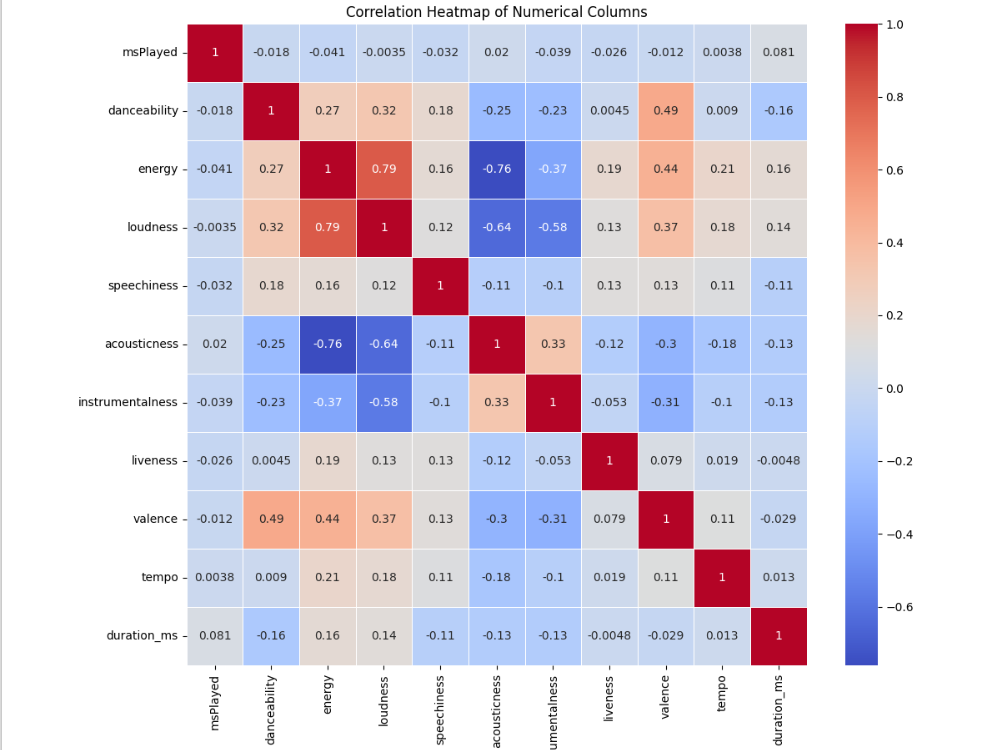
Ans- If the points tend to follow a clear upward or downward trend, it suggests a positive or negative correlation, respectively. If the points are scattered randomly, there may be little to no correlation between the two attributes.

In this dataset points are scattered randomly, so little to no relation between the 2 attributes.

**9. Multivariate Analysis**

**Q20**. Correlation between different numerical columns.

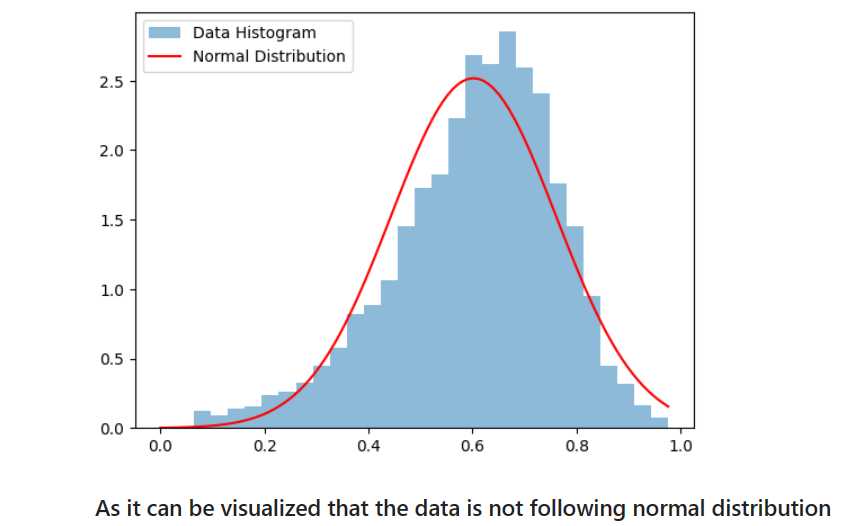
Ans- The heatmap reveals the strength and direction of correlations between numerical columns; for example, a strong positive correlation between 'energy' and 'loudness' suggests that energetic tracks tend to be louder.

****

**10. Distributions**

Distribution for age of casualty column was visualized using sns.histplot(d[‘danceability’], kde=True)

A normal distribution plot was plotted on the danceability column using

****plt.plot(x, stats.norm.pdf(x, np.mean(data), np.std(data)), ‘r-‘, label=’Normal Distribution’)

As it can be seen that the data is not following a normal distribution.

Nor is it following any other standard distributions.

Then skewedness was checked for the distribution using

if skewness > 0: The data is right-skewed (positively skewed).

Elif skewness < 0: The data is left-skewed (negatively skewed).

The distribution was found to be left skewed.

Therefore, the distribution of danceability column is a left skewed distribution

**11. Hypothesis Testing**

Following hypothesis were developed and then tested on the data.

**1.Normality test using Shapiro-Wilk Test** : **tests If data is normally distributed.**

Normality test was conducted using Shapiro-Wilk method and the distribution was found to be not normal distribution.

**2.T Test : Comparing mean age of the casualty of male and female**

Lets assume mean danceability and energy have no major difference

if p\_value < 0.05:

print(“Reject the null hypothesis. The means of danceability and energy are significantly different.”)

else:

print(“Fail to reject the null hypothesis. No significant difference in danceability and energy.”)

the null hypothesis was rejected

**12. Finding and Insights**

1. **Genre Diversity:**
   1. The dataset encompasses a diverse range of music genres, with varying prevalences.
   2. Some genres exhibit a higher prevalence of live performances, indicating a potential association between musical styles and live recordings.
2. **Instrumental Tracks:**
   1. Instrumental tracks are common, pointing to a significant presence of purely instrumental music in the dataset.
3. **Key and Modality:**
   1. Major key tracks are more prevalent than minor key tracks, suggesting a general inclination towards a more upbeat tonality.
   2. The prevalence of instrumental tracks indicates a significant presence of purely instrumental music in the dataset.
4. **Correlations:**
   1. A correlation between 'energy' and 'loudness' suggests that tracks with higher energy levels tend to be louder, highlighting a natural relationship between these attributes.
5. **Genre-Specific Characteristics:**
   1. Exploration of song attributes like 'energy,' 'danceability,' and 'valence' across genres reveals distinct musical characteristics associated with different genres.
6. **Tempo Preferences:**
   1. Tempo preferences vary across genres, with some favoring faster tempos while others gravitate towards slower tempos.

**7. Acousticness:**

* 1. Analysis of 'acousticness' across genres reveals that some genres, such as classical and acoustic, predominantly feature acoustic instruments.

**8. Speechiness:**

* 1. The 'speechiness' of tracks varies across genres, indicating differences in the presence of vocals.

**9. Time Signatures:**

* 1. Examination of time signatures across genres provides insights into the typical musical structures associated with different genres.

**10. Emotional Content:**

* 1. The correlation between 'danceability' and 'valence' suggests that more danceable tracks tend to convey more positive emotions.

**11. Top Artists and Track Durations:**

* 1. Identification of the top 50 most popular artists showcases those with the most significant presence in the dataset.
  2. Exploration of track durations highlights extreme outliers in terms of track length.

**13. RECOMMENDATIONS**

Data Collection and Preparation

1. Acquire a comprehensive dataset of Spotify song attributes. This dataset should include a variety of attributes, such as acousticness, danceability, duration, energy, instrumentalness, key, liveness, loudness, mode, tempo, and valence.
2. Clean and prepare the data for analysis. This may involve handling missing values, removing outliers, and transforming variables as needed.
3. Explore the distribution of each attribute to understand the range and variability of the data.

Exploratory Data Analysis

1. Examine the relationships between song attributes using correlation analysis and scatter plots. Identify any patterns or trends that emerge.
2. Investigate the relationship between song attributes and popularity metrics, such as track popularity and danceability rating.
3. Analyze how song attributes vary across different genres and eras of music.
4. Utilize grouping and filtering techniques to uncover hidden patterns and insights within the data.

Visualization and Storytelling

1. Create compelling visualizations to effectively communicate the findings of the EDA. Utilize charts, graphs, and maps to present the data in a clear and engaging manner.
2. Employ storytelling techniques to weave the findings of the EDA into a coherent narrative. Highlight key insights and draw meaningful conclusions from the data.
3. Share the EDA findings with a wider audience through presentations, reports, or interactive dashboards.

**14. CONCLUSION**

In conclusion, the Exploratory Data Analysis (EDA) conducted on Spotify song attributes has unveiled a rich tapestry of musical diversity within the dataset. The prevalence of varied genres, ranging from those inclined towards live performances to genres predominantly featuring instrumental tracks, reflects the heterogeneous nature of musical expression.

Key findings, such as the dominance of major key tracks, a correlation between 'energy' and 'loudness,' and the exploration of attributes like 'energy,' 'danceability,' and 'valence' across genres, provide valuable insights into the musical landscape. These observations not only deepen our understanding of musical preferences but also hint at inherent connections between different musical elements.

Tempo preferences across genres, distinctions in acousticness, variations in speechiness, and insights into time signatures contribute to a nuanced understanding of the stylistic nuances within various musical genres. The correlation between 'danceability' and 'valence' further illuminates the emotional content of danceable tracks, suggesting a positive emotional association.

The identification of top artists, exploration of track durations, and recognition of outliers in terms of track length contribute to a comprehensive snapshot of the dataset. Altogether, this EDA serves as a foundation for further analyses and underscores the richness and complexity of the musical landscape captured in the Spotify song attributes dataset.

**15.References**

* + Kaggle
  + Google
  + Stackoverflow
  + Libraries Used

1. NumPy

2. pandas

3. seaborn

4. matplotlib

5. SciPy

**16.Acknowledgments**

I would like to express my sincere gratitude and appreciation to everyone who contributed to the successful completion of this project. This endeavor has been a journey of growth, learning, and collaboration, and I am thankful for the support and guidance provided by various individuals and resources.

First and foremost, I extend my heartfelt thanks to Shivangini Gupta my project supervisor, for her invaluable mentorship, encouragement, and constructive feedback throughout the duration of this project. Their expertise and insights have been instrumental in shaping the project’s direction and improving its quality.

I would like to acknowledge the assistance of my friend who provided support, shared ideas, and engaged in meaningful discussions that contributed to the project’s development. Their collaboration has been a source of inspiration and motivation.

Furthermore, I extend my appreciation of the Anaconda Application on which I executed the code to get insight into the project.

This project has been a fulfilling and enriching experience, and I am thankful for the collective effort that has gone into its realization.

**17.Project Code**

• IPYNB file

<https://drive.google.com/file/d/1Hf2XPgNNLX5X4V6F5yiimVBLsdwuvXDI/view?usp=drive_link>

• Dataset

<https://drive.google.com/file/d/1ye0O9mqssYciViVGgmXeWKjDSIYgZOlx/view?usp=drive_link>

• Presentation

<https://docs.google.com/presentation/d/1UlIq_UXfz_ahNLFIREh51ZZF7yLcqa3D/edit?usp=drive_link&ouid=102021664191856274230&rtpof=true&sd=true>