**Report: Analysis of Spotify's Most Streamed Songs**

This report presents the analysis of the most streamed songs on Spotify, including the Python code used for:

- data handling

- cleaning

- exploration

- visualization along with explanations and interpretations of the results.

**Data Handling**

Code Snippet:

spotify\_data = pd.read\_csv('Spotify\_Most\_Streamed\_Songs.csv')

cleaned\_data = cleaned\_data.drop\_duplicates()

The dataset containing Spotify's most streamed songs is loaded into a pandas DataFrame. The dataset is now ready for further cleaning and exploration

**Data Exploration**

Code Snippet:

spotify\_data['streams'] = pd.to\_numeric(spotify\_data['streams'], errors='coerce')

cleaned\_data = spotify\_data.dropna(subset=['streams'])

I ensured that the streams column is in numeric format, converting any invalid entries into NaN values using pd.to\_numeric. The rows with missing stream data were dropped using dropna, as incomplete data would affect the accuracy of the analysis.

All rows with missing or invalid stream data have been removed, leaving clean and valid data for further analysis.

**Data Manipulation and Analysis:**

cleaned\_data = cleaned\_data.drop\_duplicates()

top\_10\_songs = cleaned\_data.sort\_values(by='streams', ascending=False).head(10)

print("Top 10 Most Streamed Songs on Spotify:")

print(top\_10\_songs[['track\_name', 'streams']])

**Explanation:**

We performed two key steps here:

**Duplicate Removal:** By removing any duplicate rows, we ensure that each song is only represented once in the analysis.

**Sorting and Selecting:** The data was sorted by the number of streams in descending order, and the top 10 most streamed songs were selected using head(10).

**Results:**  
  **track\_name streams**

**55 Blinding Lights 3,703,895,074**

**179 Shape of You 3,562,543,890**

**86 Someone You Loved 2,887,241,814**

**620 Dance Monkey 2,864,791,672**

**41 Sunflower - Spider-Man: Into the Spider-Verse 2,808,096,550**

**162 One Dance 2,713,922,350**

**84 STAY (with Justin Bieber) 2,665,343,922**

**140 Believer 2,594,040,133**

**725 Closer 2,591,224,264**

**48 Starboy 2,565,529,693**

**Data Visualization**

plt.figure(figsize=(12, 6))

sns.barplot(y=top\_10\_songs['track\_name'], x=top\_10\_songs['streams'], palette='Blues\_d', hue=top\_10\_songs['track\_name'], dodge=False)

plt.title('Top 10 Most Streamed Songs on Spotify')

plt.xlabel('Streams')

plt.ylabel('Track Name')

plt.legend([], [], frameon=False)

plt.gca().xaxis.set\_major\_formatter(plt.FuncFormatter(lambda x, loc: "{:,}".format(int(x))))

plt.tight\_layout()

plt.show()

**Explanation**:

A bar plot was created using Seaborn to visually represent the top 10 most streamed songs. The x-axis represents the number of streams (in full format), and the y-axis lists the track names. The legend was removed for a cleaner visualization, as it is redundant in this case.

**Result**:

The visualization clearly displays the most streamed songs, with the most popular tracks having close to 3.7 billion streams.

