**Analyse Spotify trends using tableau**

1. **Introduction**
2. **Dataset Overview**
3. **Data Preparation Steps**
4. **Importing the Data**
5. **Data Cleaning in Power Query Editor**

**Step 1: Remove Duplicate Rows**

**Step 2: Handling Missing Values**

**Step 3: Standardize Date Format**

**Step 4: Correct Data Types**

**Step 5: Filter Out Irrelevant Data**

1. **Final Dataset Overview**
2. **Saving the Cleaned Data:**
3. **Visualization:**

**Dashboard1:**

**1. KPI Donut Charts:**

**2. Position and Streams by Country (Line Graph)**

**3. World Map**

**4. Streams by Artist (Scatter Plot)**

**5. Streams and Position by Date (Bar Chart with Line Trend)**

**6. Filters**

* **Key Insights**

**Dashboard2:**

1. **Top Streams and position**
2. **Key Metrics**
3. **Bar chart**
4. **4.Line chart**
5. **Global Distribution**
6. **Conclusion**

**Dashboard 3**

**7. Conclusion**

**8. Recommendation:**

**1. Introduction**

Spotify is one of the leading digital music streaming services that gives users access to millions of songs, podcasts, and other audio content from artists all over the world. This analysis provides insights into the streaming data of Spotify, focusing on various metrics such as total streams, average streams per track, and streams by country, using the provided dataset.

**2. Dataset Overview**

The dataset used for this analysis includes the following columns:

* Artist: The name of the artist.
* Country: The country where the stream occurred.
* Date: The date of the stream.
* Track Name: The name of the track.
* Track URL: The URL of the track on Spotify.
* Streams: The number of streams for the track.

This dataset helps in understanding the streaming patterns, popular tracks, and the performance of different artists across various countries.

This document details the data preparation process for the Spotify dataset using Excel Power Query Editor. The goal was to clean and transform the data to ensure it is ready for analysis in Tableau. The final cleaned dataset contains 5,705,133 rows and 7 columns.

**3. Data Preparation Steps**

1. **Importing the Data:**

The first thing I was download the dataset, which were available in CSV format. I imported the file into Excel. It was 3-year data set [2017,2018,2019]

* 1. Open Excel and navigate to Data > Get Data > From File > From Workbook.
  2. Select the Spotify dataset file and load it into Power Query Editor.

1. **Data Cleaning in Power Query Editor:**

**Step 1: Remove Duplicate Rows**

**Step 2: Handling Missing Values**

* 1. Check for any missing values in the dataset.
  2. For critical columns like Artist, Country, Date, Track Name, and Streams, remove rows where these values are missing.

**Step 3: Standardize Date Format**

* 1. Ensure the Date column is in a consistent format (e.g., YYYY-MM-DD).
  2. Select the Date column, go to the Transform tab, and use Data Type > Date to format the column correctly.

**Step 4: Correct Data Types**

* 1. Verify and correct the data types for each column:
     1. Artist: Text
     2. Country: Text
     3. Date: Date
     4. Track Name: Text
     5. Track URL: Text
     6. Streams: Whole Number
  2. Select each column and set the appropriate data type from the Transform tab.

**Step 5: Filter Out Irrelevant Data**

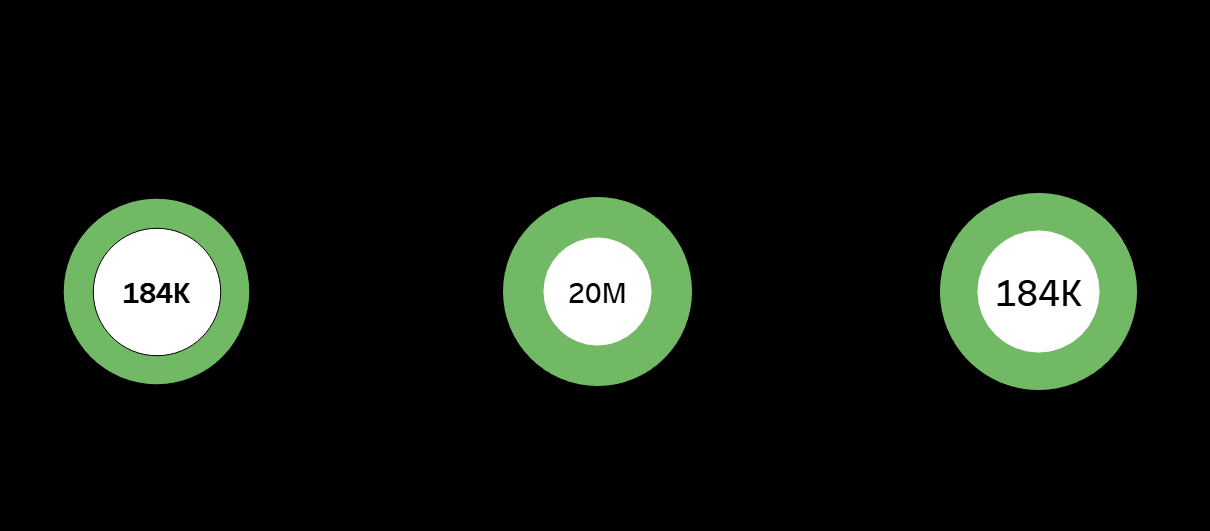
* 1. Filter out any irrelevant data that does not contribute to the analysis. For example, if the dataset includes test data or outliers, remove them.
  2. Use the filter options on each column to exclude unnecessary data.

1. **Final Dataset Overview:**
   1. After cleaning, the final dataset contains:
      1. **Rows:** 5,705,133
      2. **Columns:** 7 (Artist, Country, Date, Track Name, Track URL, Streams, and any other relevant column)
2. **Saving the Cleaned Data:**
   1. Once the data cleaning is complete, click on Close & Load to load the cleaned data back into Excel.
   2. Save the cleaned dataset as a new Excel file for further analysis in Tableau.
3. **Visualization:**

This visualization is a comprehensive dashboard designed to display various trends and metrics related to Spotify data. Below is a detailed description of each component in the dashboard:

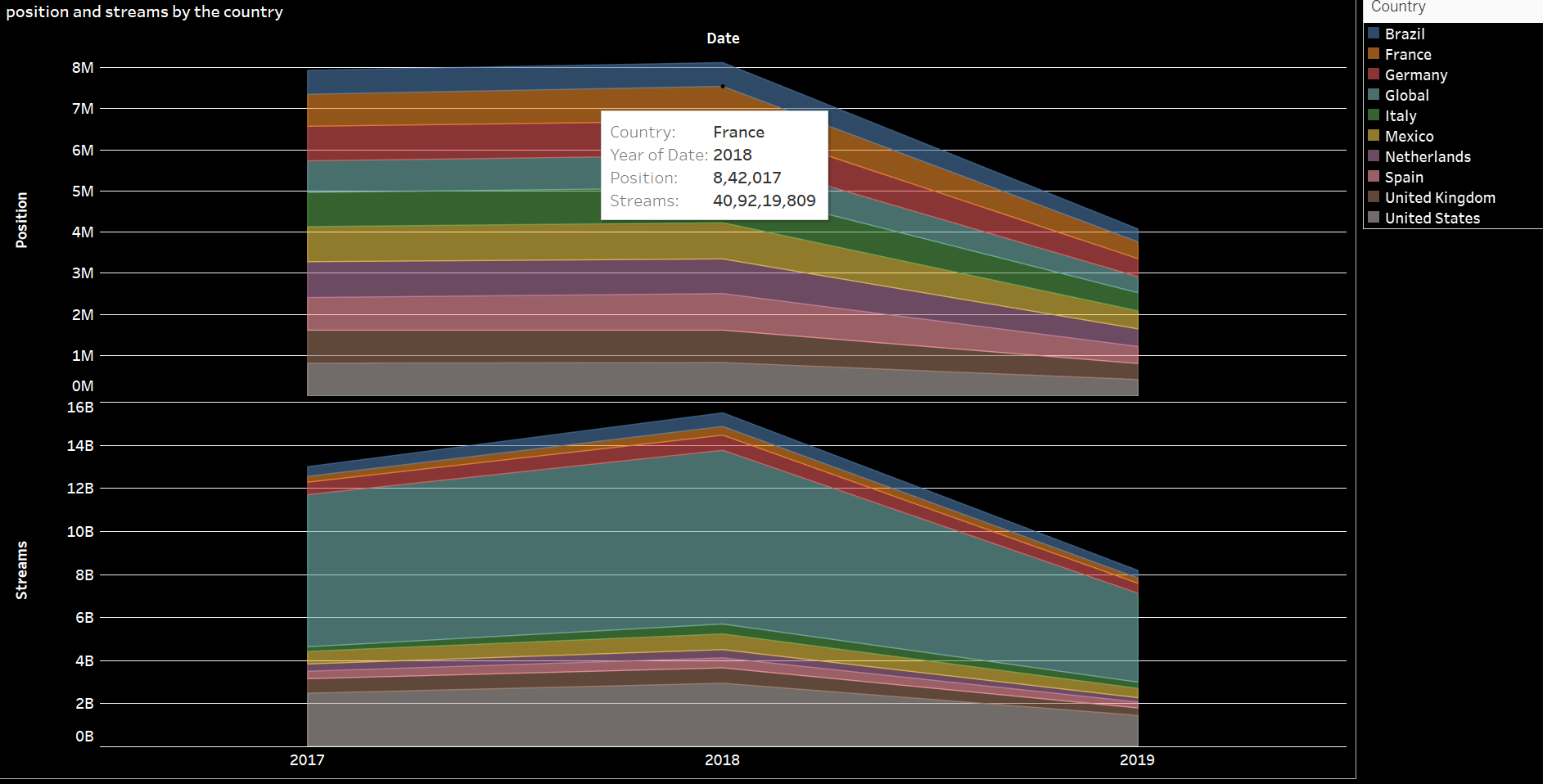
**1. KPI Donut Charts:**

* Track URL (184K): Indicates the total number of unique track URLs present in the dataset.
* Total Position (20M): Shows the aggregate total position value for all tracks.
* Total Streams (184K): Represents the total number of streams across all tracks.

KPI Donut chart

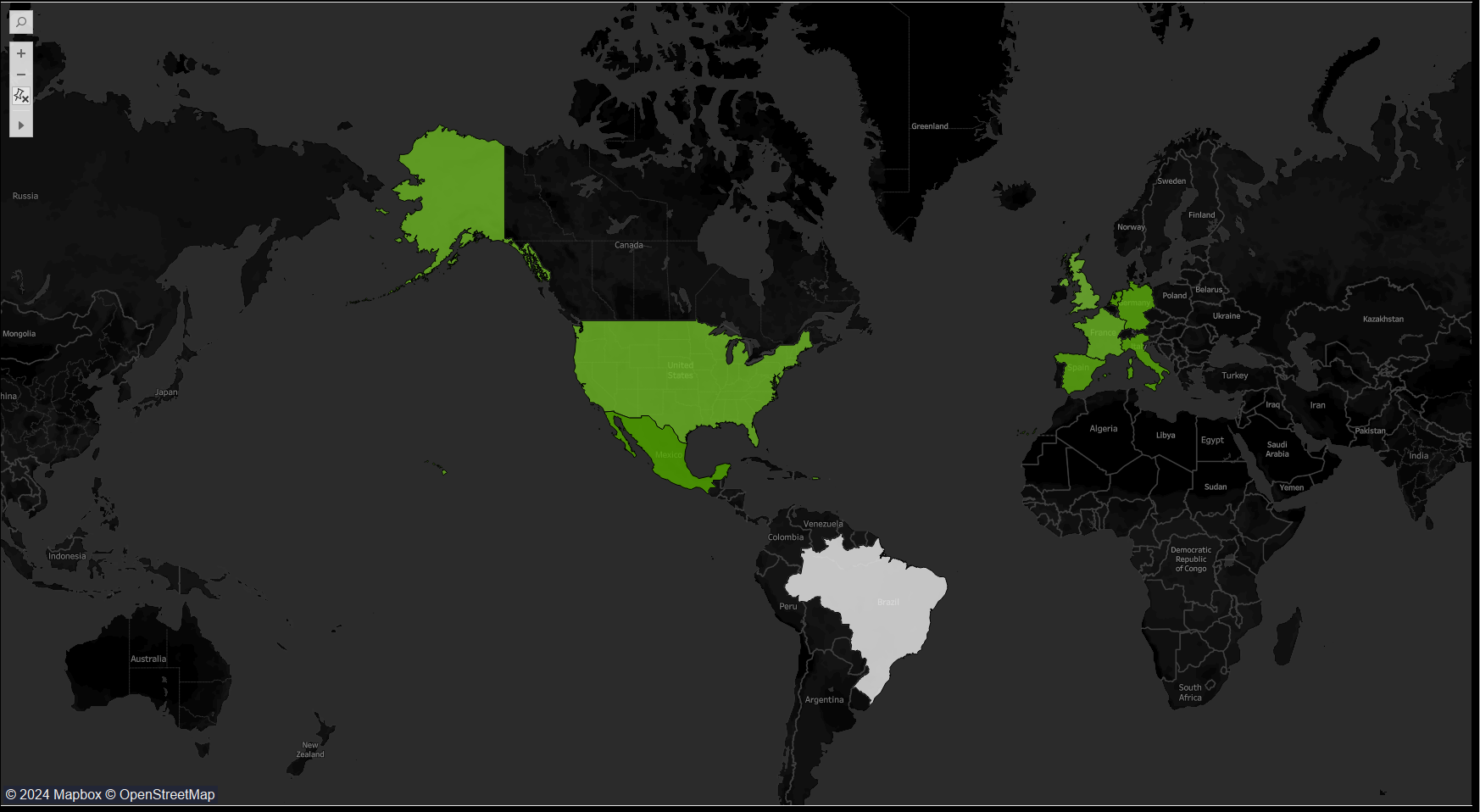
**2. Position and Streams by Country (Line Graph):**

* Y-Axis (Streams Position): Displays the position of streams, showing the popularity or ranking of tracks.
* X-Axis (Date): The timeline from 2017 to 2019, showing how the position and streams have changed over time.
* The colourful lines represent different countries, illustrating how each country's streams have varied over the given period.



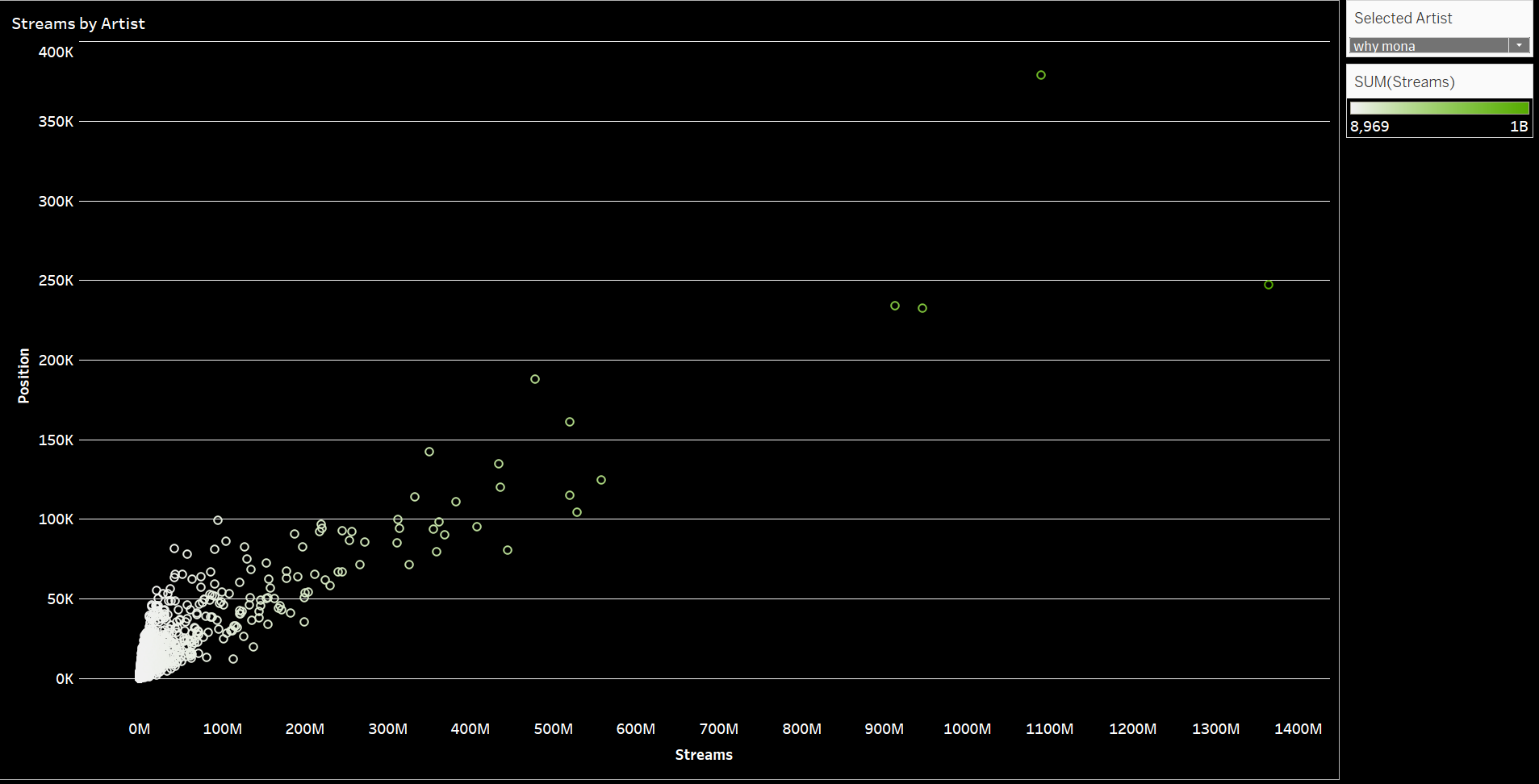
**3. World Map:**

* Country Highlight: A geographical representation of the data, where countries with significant stream counts are highlighted in green. This provides a visual insight into the global distribution of streams.

****

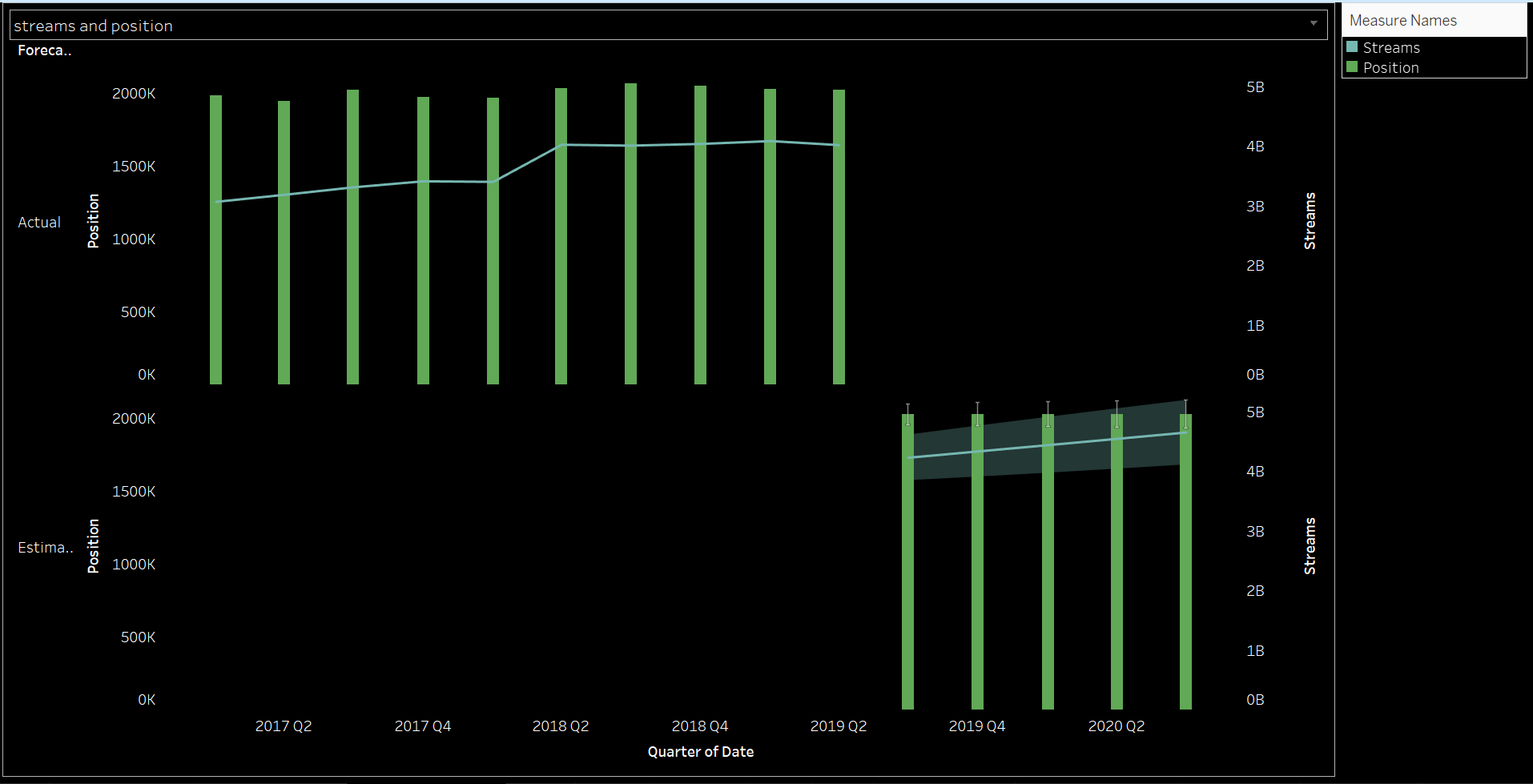
**4. Streams by Artist (Scatter Plot):**

* X-Axis (Streams): Number of streams ranging from 0 to 1.4 billion.
* Y-Axis (Position): Position ranging from 0 to 400K.
* Each dot represents an artist, showing their position relative to their total streams. This helps identify the most popular artists and their ranks.



**5. Streams and Position by Date (Bar Chart with Line Trend):**

* X-Axis (Quarter of Date): Represents quarters from Q4 2017 to Q4 2019.
* Left Y-Axis (Position): The vertical axis for position.
* Right Y-Axis (Streams): The vertical axis for streams.
* Green Bars: Show the actual streams for each quarter.
* Blue Line: Represents the trend of positions over time.
* This dual-axis chart provides a comparative view of streams and positions across different quarters, highlighting trends and patterns over time.



**6. Filters:**

* Country Filter: Allows users to select specific countries to view their data.
* Artist Filter: Enables the selection of specific artists to analyse their data in detail.

7. Legend:

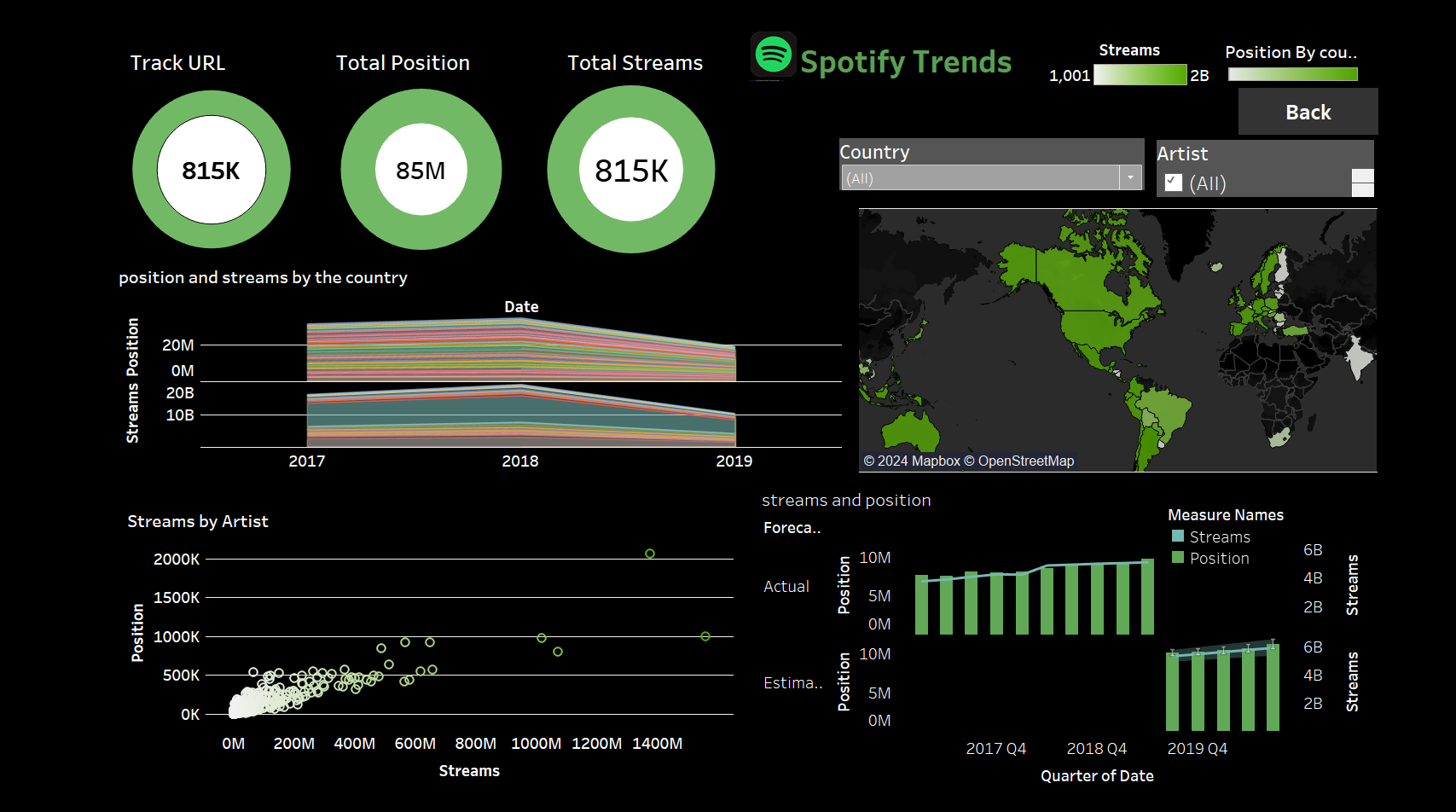
* Measure Names (Streams and Position): Clarifies the colour coding used in the bar chart and line trend chart.

**Key Insights:**

* The donut charts provide quick KPIs to understand the overall data metrics.
* The line graph and scatter plot offer detailed insights into trends over time and the popularity of tracks and artists.
* The world map helps visualize geographical distribution, highlighting key regions with significant streaming activity.
* The bar chart with the trend line allows for the analysis of how streams and positions have evolved quarterly.

This dashboard is a powerful tool for understanding Spotify trends, helping stakeholders make informed decisions based on historical data and trends.

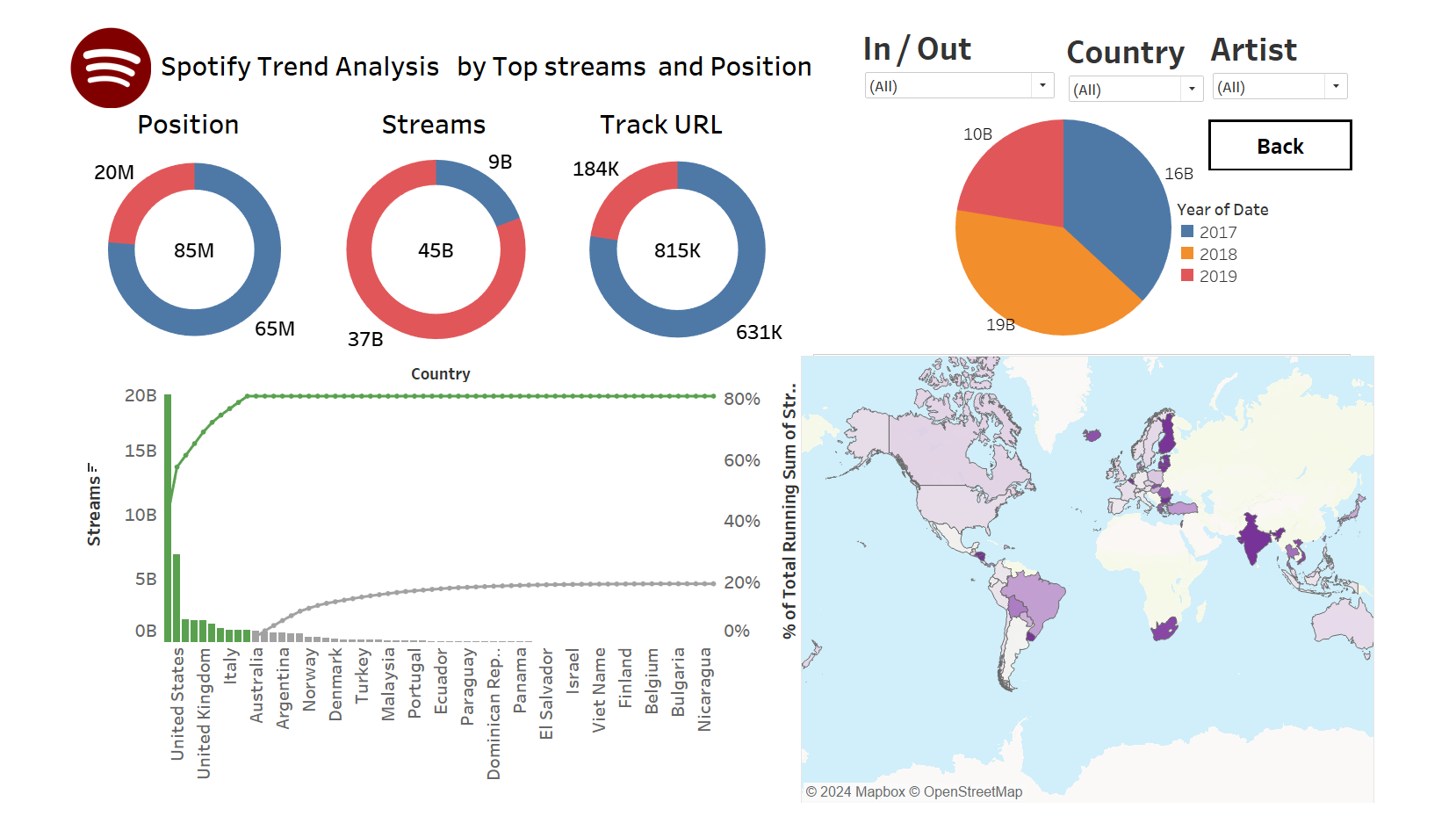
**Dash Board 1:**

****

**Dash board 2:**

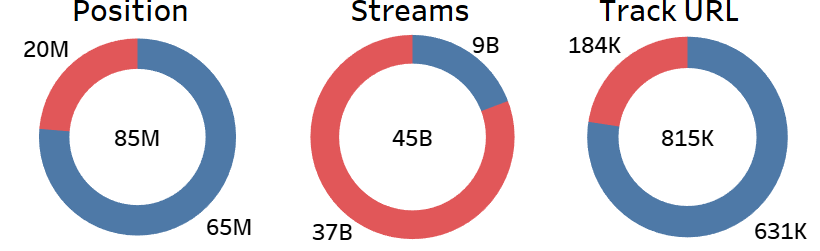
* 1. **Top Streams and position**

This document provides an analysis of Spotify trends based on the top streams. The data visualized includes information about the position, streams, track URL, and country-wise distribution of streams.

****

1. **Key Metrics**

KPI Donuts:



1. **Country wise Stream Analysis Pareto chart:**

The Pareto chart depicts the total streams by country:

**Its divided two parts**

1. Top ten country by streams
2. Other country

**3.Bar chart:**

Bar chart tells total stream of the country

**4.Line chart:**

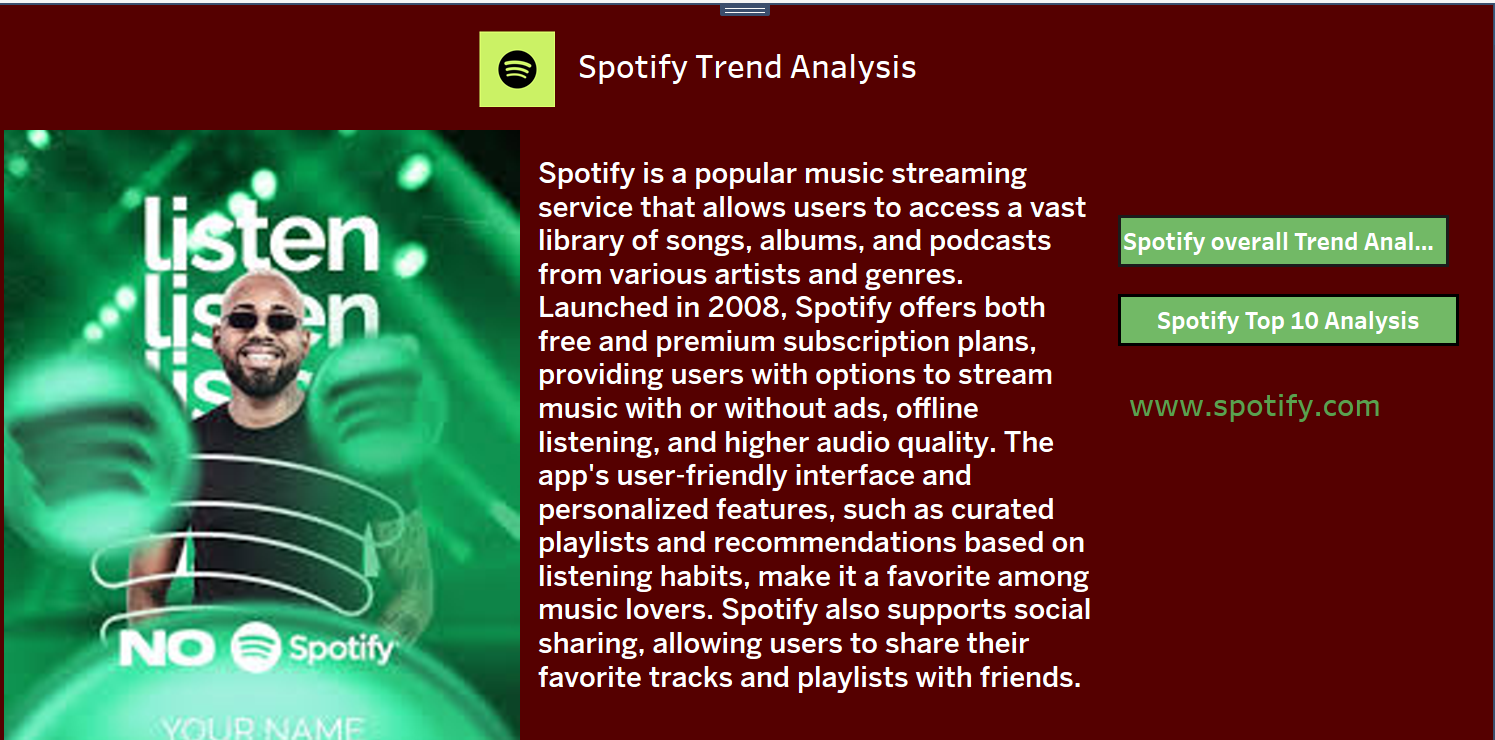
Line chart tell running total of the country

**5.Global Distribution**

The world map shows the geographical distribution of position, highlighting countries with the highest position in purple.

**Dash Board 3:**

This is the index page of the project



**Conclusion**

This analysis provides a comprehensive overview of Spotify's streaming data, offering valuable insights into artist popularity, track performance, and stream performance. The findings can be utilized by artists, record labels, and marketers to make data-driven decisions and tailor their strategies to better meet the needs and preferences of Spotify users.

**Recommendation:**

**Incorporate User Demographics and Behavioural Data:**

* Recommendation: Integrate user demographics (age, gender, location) and behavioural data (listening habits, playlist creation) into the analysis.
* Rationale: Identifying peak listening times and seasonal preferences can help in planning releases and promotional campaigns. It can also reveal insights into how user behaviour changes over time, which can inform content scheduling and platform optimization.
* Recommendation: Break down the genre data into more granular subgenres and analyse their performance.

**Conclusion**

The data preparation process involved importing the raw Spotify dataset into Excel Power Query Editor and performing a series of cleaning and transformation steps. By removing duplicates, handling missing values, standardizing date formats, correcting data types, and filtering out irrelevant data, we ensured that the dataset is clean and ready for analysis. The final cleaned dataset, containing 5,705,133 rows and 7 columns, provides a solid foundation for generating insights and visualizations in Tableau.