Spotify Charts Report

By: Salma Moustafa

28/8/2024

DEBI/CLS

Executive Summary

This report provides an in-depth analysis of the Spotify streaming dataset, focusing on identifying music trends across various regions from 2017 to 2021. Key findings reveal the most popular artists and songs in different regions, as well as changes in music preferences over time. The analysis includes enhancements like feature engineering and the use of interactive visualizations for better insights. Recommendations for further exploration and the development of a city-specific Spotify Wrapped feature are also provided.

Table of Contents

1. Introduction

1.1 Purpose of the Report

The purpose of this report is to summarize the findings from the Spotify Charts data analysis project. This document aims to provide a comprehensive overview of the data analysis process, key findings, and insights that can inform Spotify's business strategies, particularly in developing a city-specific version of the Spotify Wrapped feature.

1.2 Scope

Dataset: The analysis was conducted on the Spotify Charts dataset, which includes data on song streams across various countries from 2017 to 2021.

1.3 Data Sources

Kaggle Dataset: The dataset used for this analysis was sourced from Kaggle: https://www.kaggle.com/datasets/dhruvildave/spotify-charts.

2. Methodology

2.1 Data Collection

The Spotify Charts dataset, containing over 26 million records from 2017 to 2021, was used for this analysis. The dataset includes information about song titles, artists, streams, and regions.

2.2 Data Cleaning

The data cleaning process involved removing null values in the title, artist, and streams columns, filling missing values in the streams column with zeros, dropping duplicate records, converting the date column to a datetime format, and converting the streams column to an integer format for better processing.

2.3 Data Transformation and Feature Engineering

New Features: year (extracted from the date column), month and month_name (derived from the date column to facilitate time-based analysis).

Enhancements: Aggregated data by artist and song to identify top performers, filtered data by region to focus on specific country-level trends.

2.4 Analytical Methods

Exploratory Data Analysis (EDA): Descriptive statistics and visualizations were employed to understand the distribution and trends within the data.

Trend Analysis: The dataset was analyzed to identify the most popular songs, artists, and regions over time.

Visualization Techniques: Various plots, including bar charts, line plots, and interactive visualizations using Plotly, were used to represent the data visually.

3. Data Analysis

3.1 Exploratory Data Analysis (EDA)

The data was explored using descriptive statistics and visualizations to understand the distribution of streams, identify popular artists and songs, and examine trends over time.

3.2 Key Findings

Top Artists (2017-2021): Ed Sheeran emerged as the most popular artist with over 23 billion streams, followed by Post Malone and Drake.

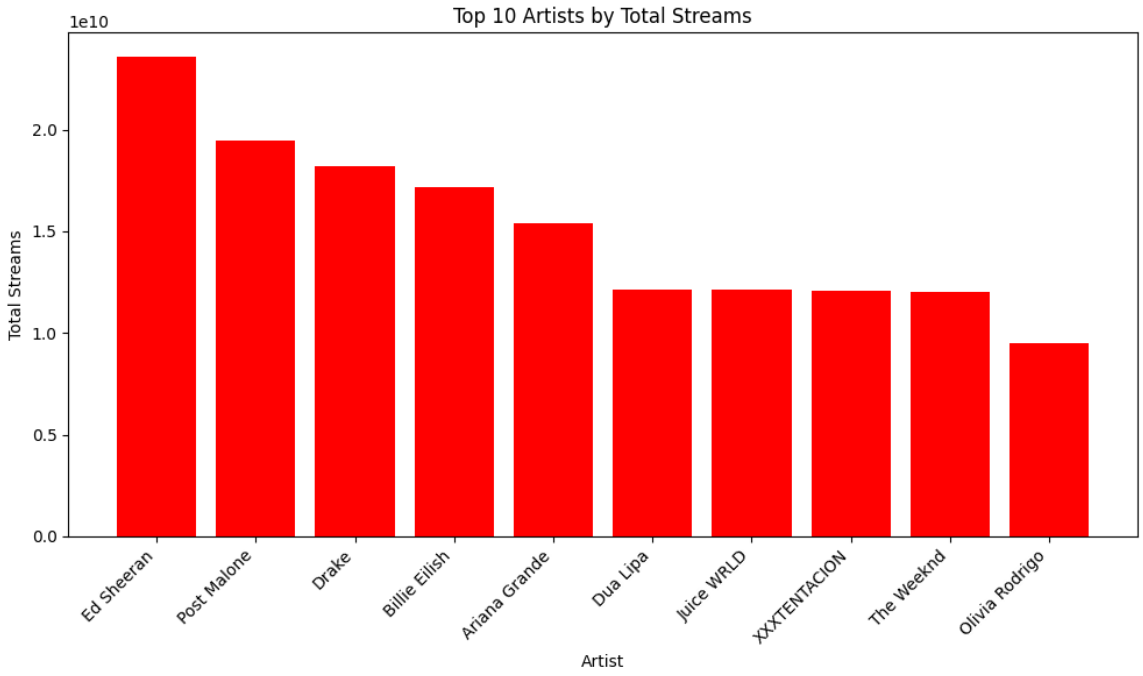


Figure 1: Top 10 Artists by total streams

The most popular songs are shape of you, blinding lights, and dance monkey.

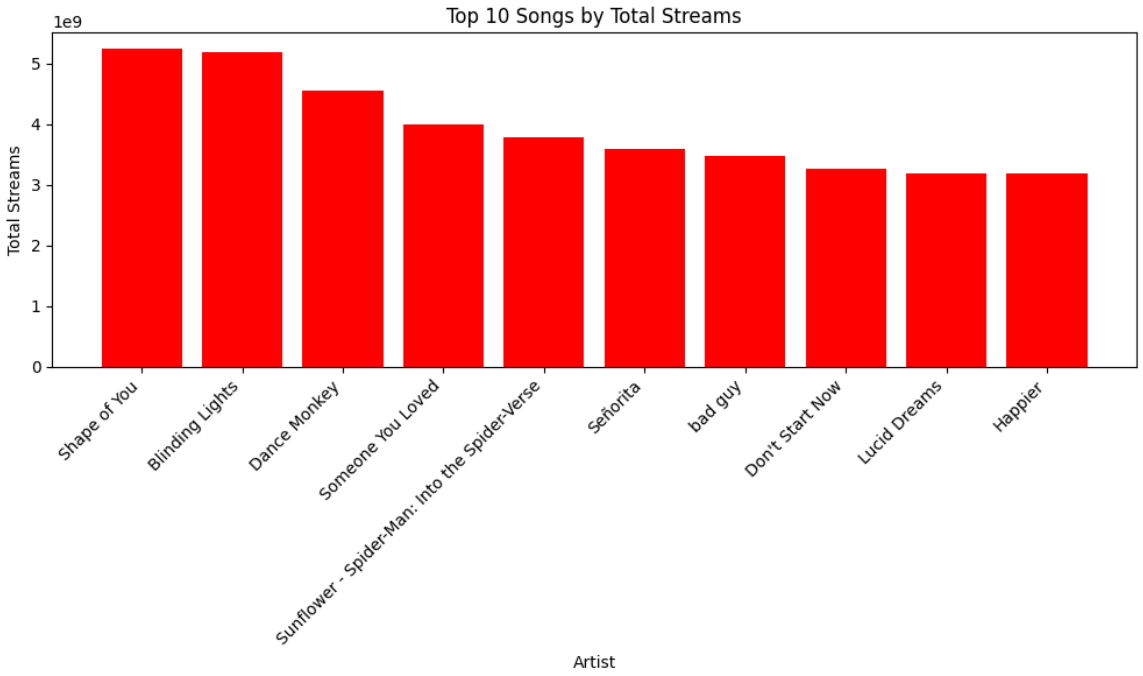


Figure 2: Top 10 songs by total streams

Most Popular Songs by Region: "Me Gusta" was the most popular song in Argentina, while "Happier" by Marshmello topped the charts in Australia.

region	most_popular_song
Andorra	Friday (feat. Mufasa & Hypeman) - Dopamine Re-...
Argentina	Me Gusta
Australia	Happier
Austria	Ohne mein Team
Belgium	Shape of You
Bolivia	Me Rehúso

region	most_popular_song
Brazil	Believer
Bulgaria	SICKO MODE
Canada	Say You Won't Let Go
Chile	Me Rehúso

Most popular singer in each country:

	region	most_popular_song
0	Andorra	Bad Bunny
1	Argentina	Maluma
2	Australia	Ed Sheeran
3	Austria	Ed Sheeran
4	Belgium	Ed Sheeran
5	Bolivia	Bad Bunny
6	Brazil	Marília Mendonça
7	Bulgaria	Pop Smoke
8	Canada	Drake
9	Chile	Bad Bunny

Yearly Trends: Each year from 2017 to 2021 was dominated by different artists, with Ed Sheeran leading in 2017 and Olivia Rodrigo in 2021.

	year	artist	streams
0	2017	Ed Sheeran	8802271569
1	2018	Drake	7310761327
2	2019	Billie Eilish	7954483427
3	2020	Juice WRLD	5320473705
4	2021	Olivia Rodrigo	9443262677

Trend Over Time: The analysis of total streams over time for the top 200 charts revealed significant seasonal patterns, with peak streaming periods typically in late December and early January.

Trend Distribution: The majority of songs experienced a "MOVE_DOWN" trend, indicating a common pattern of decline in popularity over time.

4. Insights and Discussion

4.1 Key Insights

Artist Dominance: Ed Sheeran was consistently the most popular artist across multiple regions, highlighting his global appeal.

Global vs. Regional Preferences: While some songs were globally popular, others had a strong regional appeal, indicating varied music tastes across different cultures.

Trends Over Time: The analysis showed a clear pattern of declining popularity for certain songs, which could inform future playlist curation strategies.

4.2 Implications

These insights suggest that Spotify could benefit from customizing its "Wrapped" feature to reflect regional music preferences more accurately. Additionally, understanding declining trends can help in the strategic placement of new releases.

4.3 Limitations

Data Size: The large dataset posed challenges in terms of processing time and memory usage, which may have impacted the comprehensiveness of the analysis.

Missing Data: Some regions had missing data, which could affect the accuracy of the trend analysis.

5. Recommendations

City-Specific Wrapped Feature: Develop a city-specific "Wrapped" feature that highlights the most popular songs and artists in different regions.

Data Optimization: Implement more efficient data processing techniques to handle large datasets.

Trend Monitoring: Regularly monitor song trends to identify potential hits early and promote them accordingly.

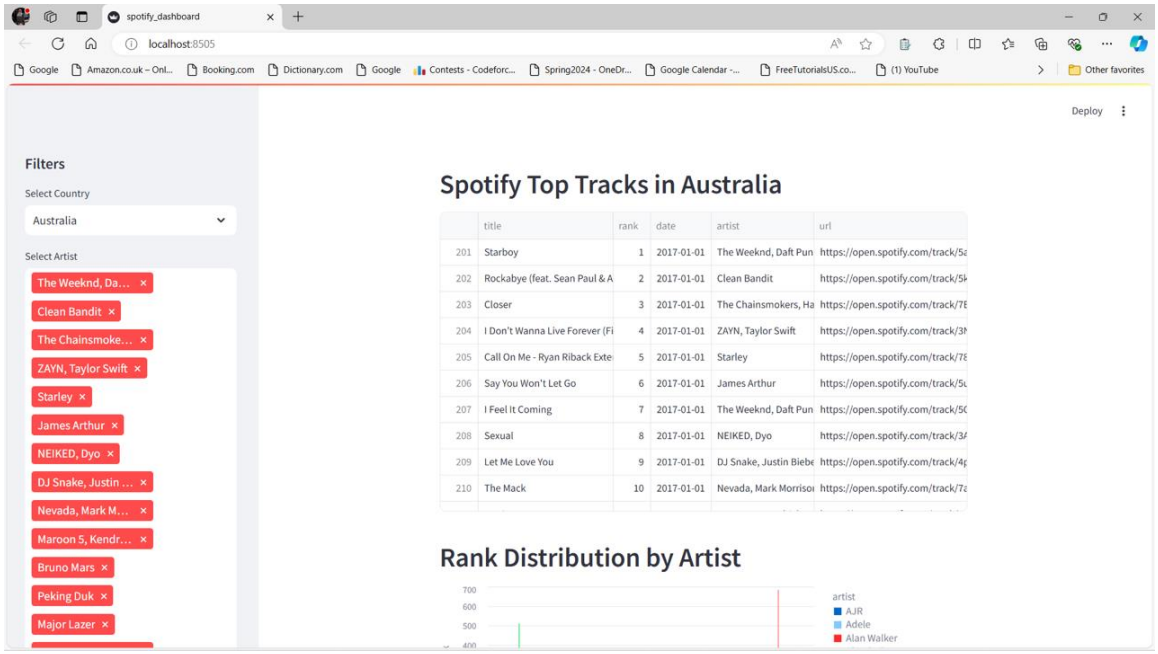
6. Conclusion

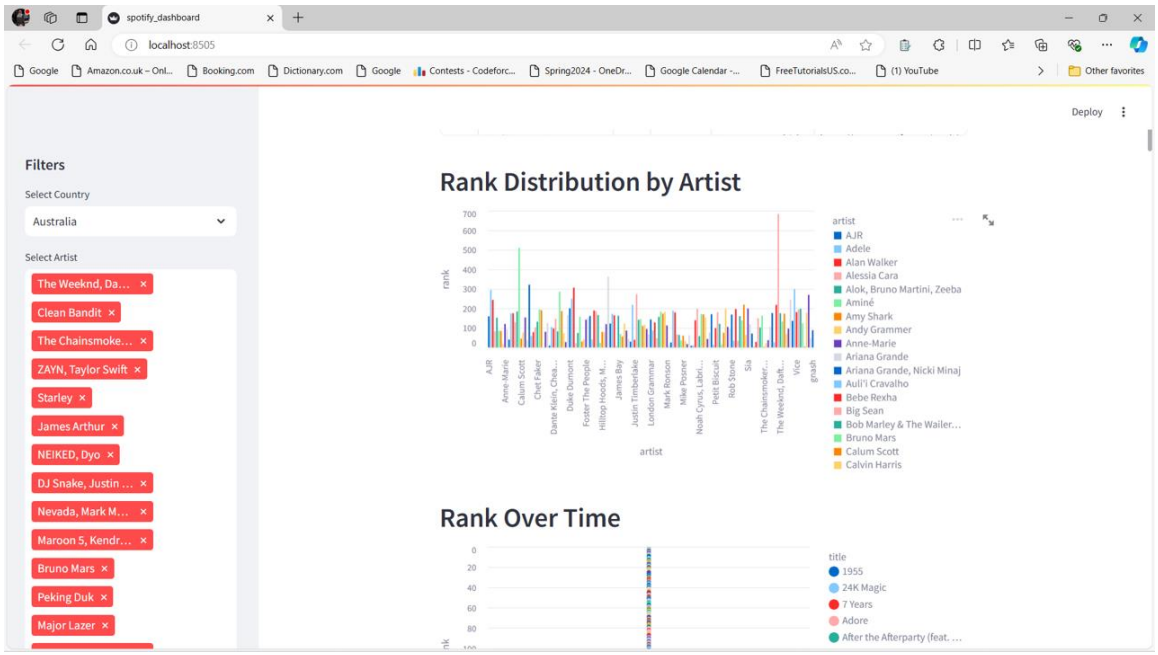
This report provides valuable insights into music streaming trends across various regions, highlighting the dominance of certain artists and songs. The findings support the development of a city-specific Spotify Wrapped feature and emphasize the importance of monitoring trends to stay ahead in the competitive music streaming market.

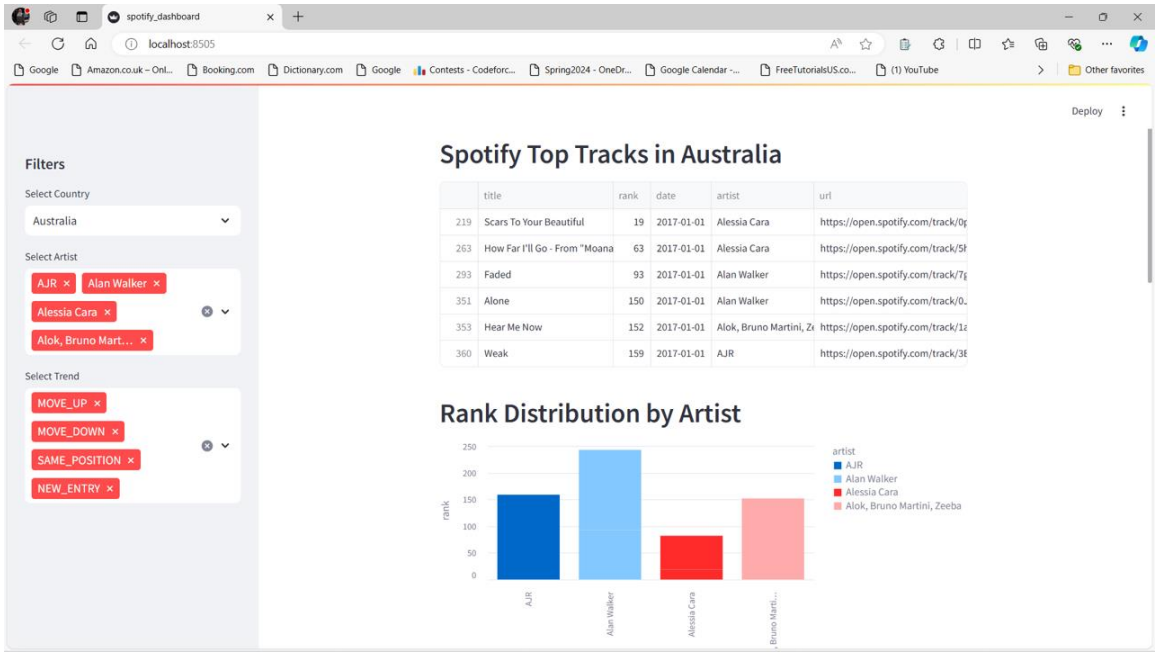
7. Appendices

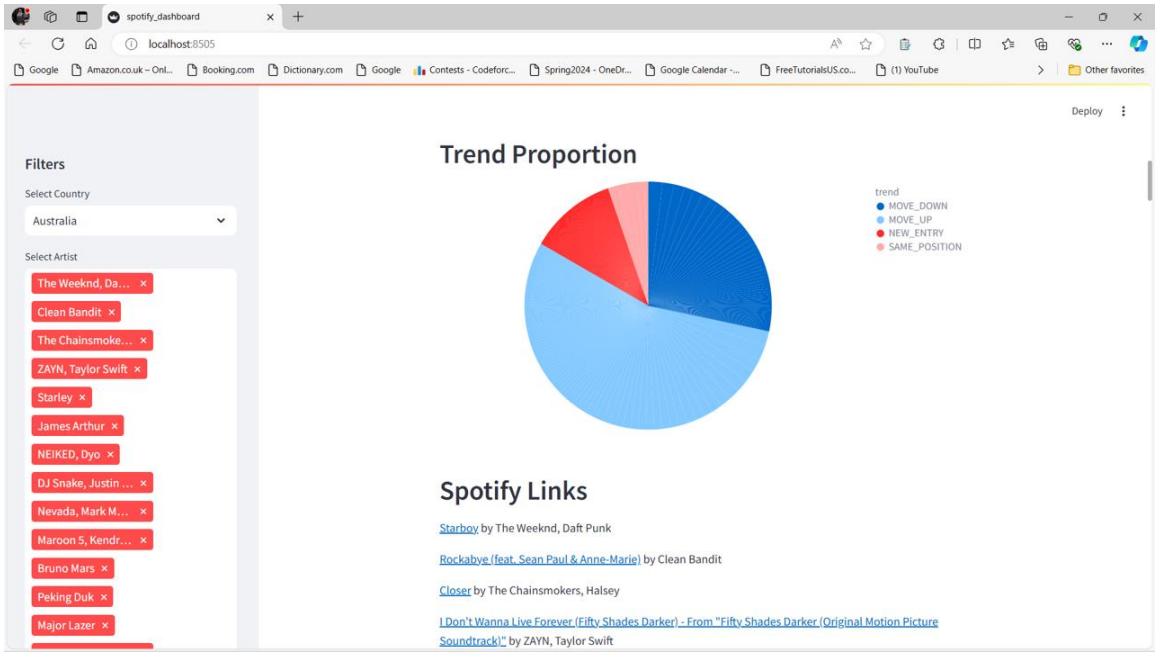
Data Preprocessing Steps: Detailed description of the data cleaning and feature engineering processes.

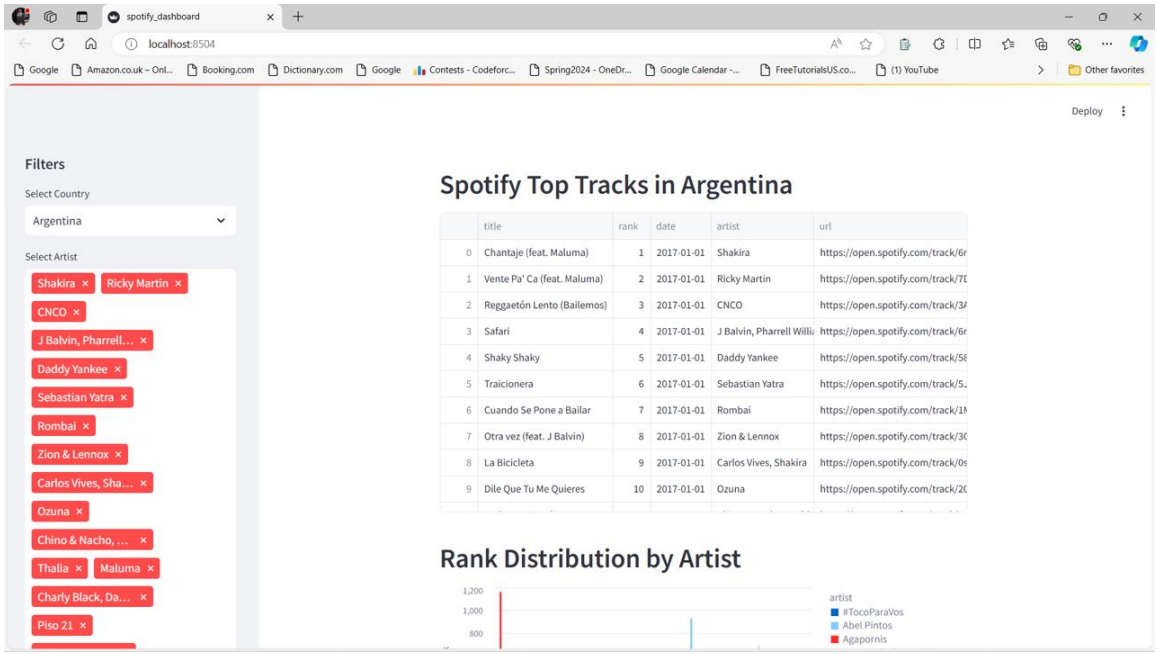
Visualization Samples: Examples of the visualizations created during the analysis.

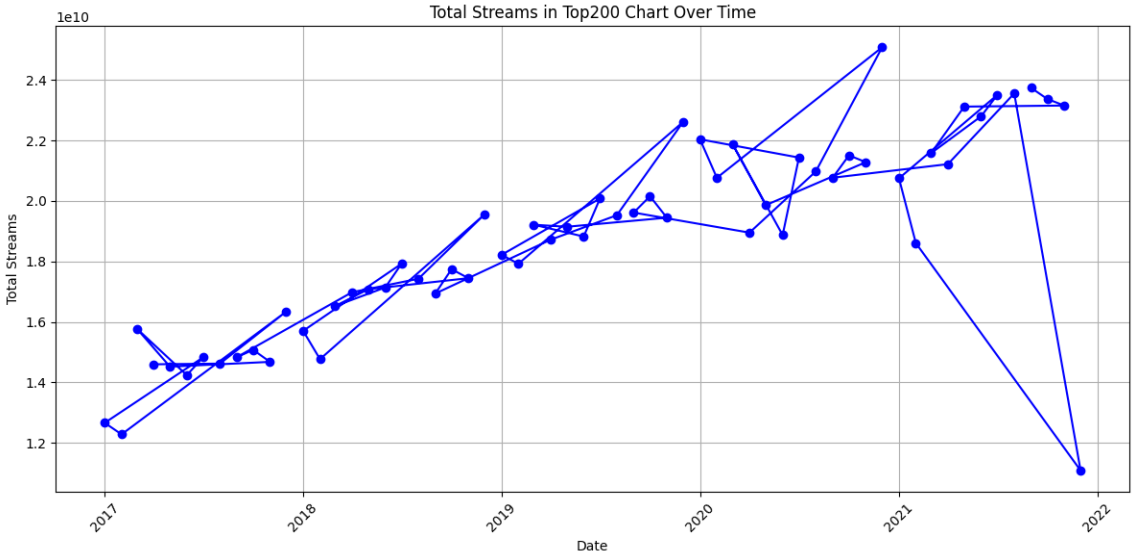


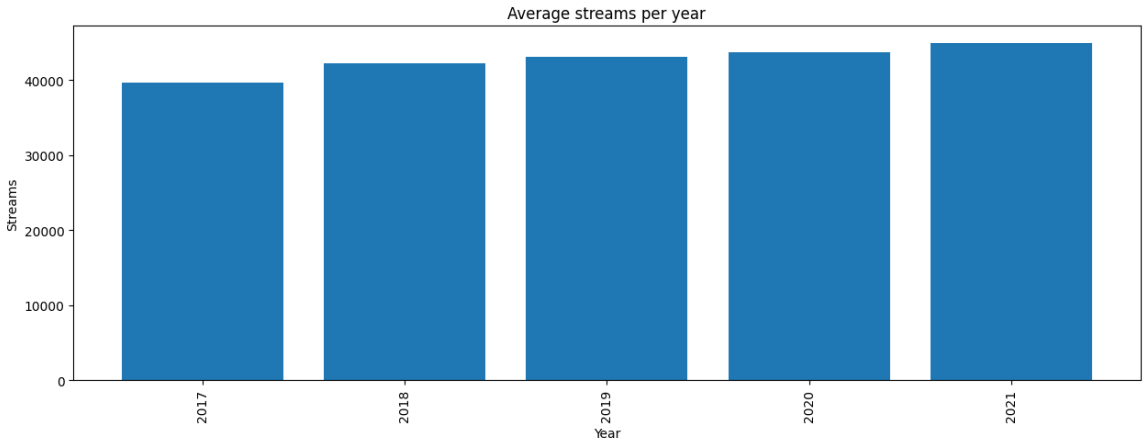


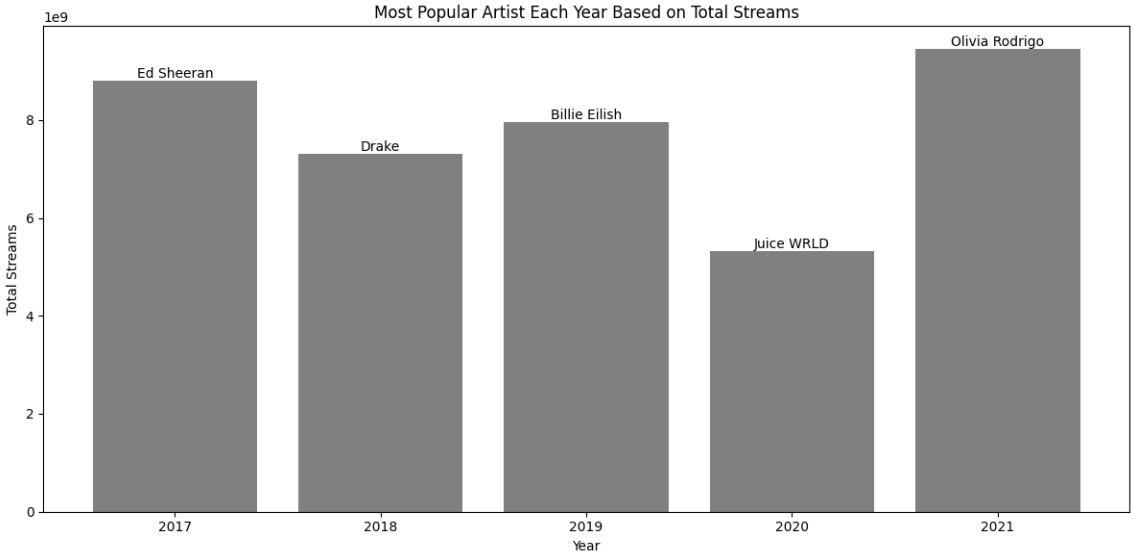

spotify_dashboard
localhost:8505
Deploy
Filters
Select Country
Australia
Select Artist
AJR
Alan Walker
Alessia Cara
Alok, Bruno Mart...
Select Trend
MOVE_UP
MOVE_DOWN
SAME_POSITION
NEW_ENTRY
Spotify Top Tracks in Australia
title rank date artist url
219 Scars To Your Beautiful 19 2017-01-01 Alessia Cara
263 How Far I'll Go - From "Moana 63 2017-01-01 Alessia Cara
293 Faded 93 2017-01-01 Alan Walker
351 Alone 150 2017-01-01 Alan Walker
353 Hear Me Now 152 2017-01-01 Alok, Bruno Martini, Z
360 Weak 159 2017-01-01 AJR
Rank Distribution by Artist
rank
artist
AJR
Alan Walker
Alessia Cara
Alok, Bruno Martini, Zeeba

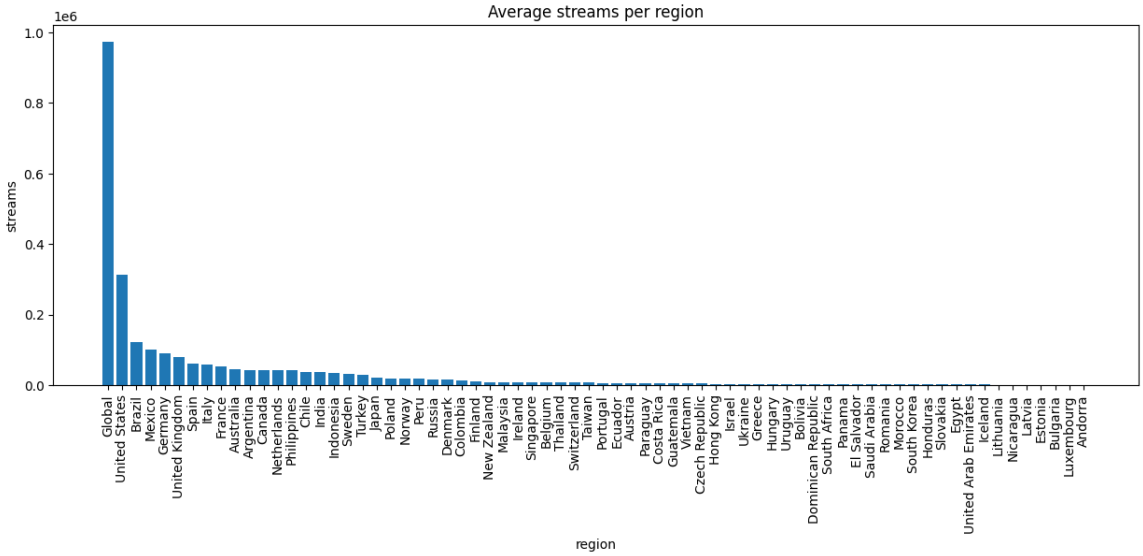

spotify_dashboard
localhost:8505
Deploy
Filters
Select Country
Australia
Select Artist
The Weeknd, Da...
Clean Bandit
The Chainsmoke...
ZAYN, Taylor Swift
Starley
James Arthur
NEIKED, Dyo
DJ Snake, Justin ...
Nevada, Mark M...
Maroon 5, Kendr...
Bruno Mars
Peking Duk
Major Lazer
Trend Proportion
trend
MOVE_DOWN
MOVE_UP
NEW_ENTRY
SAME_POSITION
Spotify Links
Starboy by The Weeknd, Daft Punk
Rockabye (feat. Sean Paul & Anne-Marie) by Clean Bandit
Closer by The Chainsmokers, Halsey
I Don't Wanna Live Forever (Fifty Shades Darker) - From "Fifty Shades Darker (Original Motion Picture Soundtrack)" by ZAYN, Taylor Swift

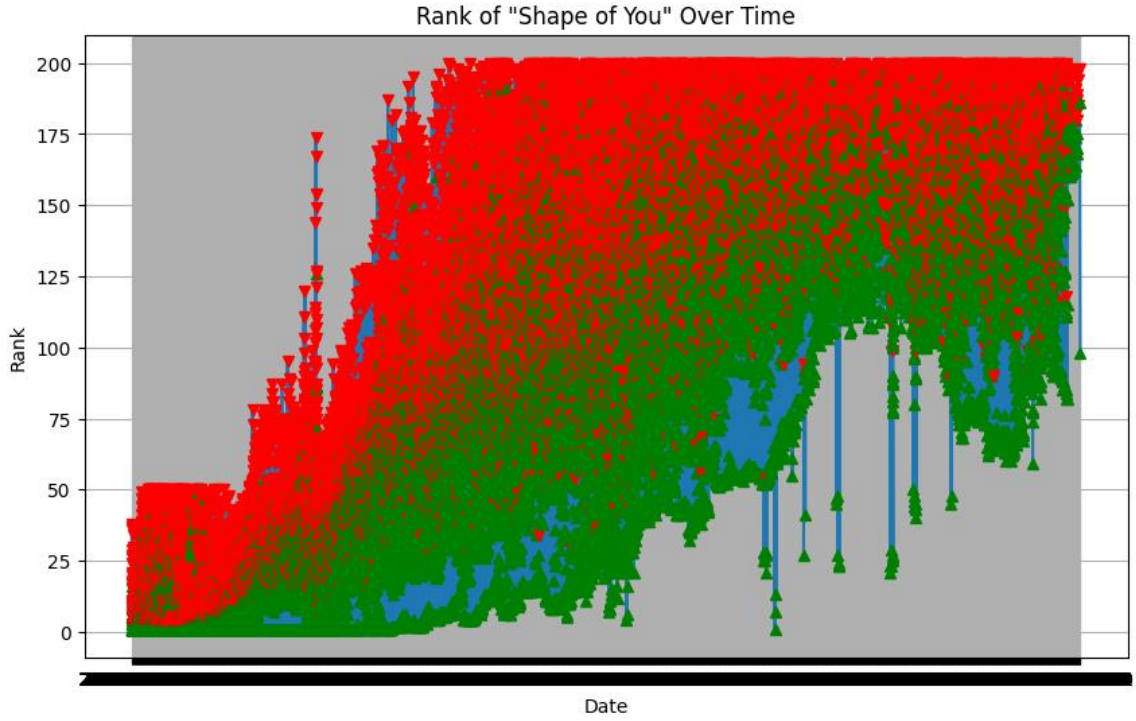

spotify_dashboard
localhost:8504
Deploy
Filters
Select Country
Argentina
Select Artist
Shakira
Ricky Martin
CNCO
J Balvin, Pharrell...
Daddy Yankee
Sebastian Yatra
Rombai
Zion & Lennox
Carlos Vives, Sha...
Ozuna
Chino & Nacho, ...
Thalia
Maluma
Charly Black, Da...
Piso 21
Spotify Top Tracks in Argentina
title rank date artist url
0 Chantaje (feat. Maluma) 1 2017-01-01 Shakira
1 Vente Pa' Ca (feat. Maluma) 2 2017-01-01 Ricky Martin
2 Reggaetón Lento (Bailemos) 3 2017-01-01 CNCO
3 Safari 4 2017-01-01 J Balvin, Pharrell Willi
4 Shaky Shaky 5 2017-01-01 Daddy Yankee
5 Traicionera 6 2017-01-01 Sebastian Yatra
6 Cuando Se Pone a Bailar 7 2017-01-01 Rombai
7 Otra vez (feat. J Balvin) 8 2017-01-01 Zion & Lennox
8 La Bicicleta 9 2017-01-01 Carlos Vives, Shakira
9 Dile Que Tu Me Quieres 10 2017-01-01 Ozuna
Rank Distribution by Artist
artist
#TocoParaVos
Abel Pintos
Agapornis


Total Streams in Top200 Chart Over Time
1e10
Total Streams
Date
2017
2018
2019
2020
2021
2022


Average streams per year
Streams
Year
2017
2018
2019
2020
2021


Most Popular Artist Each Year Based on Total Streams
1e9
Total Streams
Year
Ed Sheeran
Drake
Billie Eilish
Juice WRLD
Olivia Rodrigo
0
2
4
6
8
2017
2018
2019
2020
2021


Average streams per region
1e6
1.0
0.8
0.6
0.4
0.2
0.0
streams
Global
United States
Brazil
Mexico
Germany
United Kingdom
Spain
Italy
France
Australia
Argentina
Canada
Netherlands
Philippines
Chile
India
Indonesia
Sweden
Turkey
Japan
Poland
Norway
Peru
Russia
Denmark
Colombia
Finland
New Zealand
Malaysia
Ireland
Singapore
Belgium
Thailand
Switzerland
Taiwan
Portugal
Ecuador
Austria
Paraguay
Costa Rica
Guatemala
Vietnam
Czech Republic
Hong Kong
Israel
Ukraine
Greece
Hungary
Uruguay
Bolivia
Dominican Republic
South Africa
Panama
El Salvador
Saudi Arabia
Romania
Morocco
South Korea
Honduras
Slovakia
Egypt
United Arab Emirates
Iceland
Lithuania
Nicaragua
Latvia
Estonia
Bulgaria
Luxembourg
Andorra
region


Rank of "Shape of You" Over Time
200
175
150
125
100
75
50
25
0
Rank
Date

8. References

Spotify Charts Dataset. (2024). Retrieved from https://www.kaggle.com/datasets/dhruvildave/spotify-charts.