YouTube And Spotify Data Analysis

Overview:

The “YouTube and Spotify Data Analysis” project is an exploration of user behaviour and trends on two of the most popular streaming platforms: YouTube and Spotify. By leveraging data from these platforms, this analysis aims to uncover valuable insights into user preferences, content trends, and engagement patterns.

Technologies Used:

* Python: Python was used for data preprocessing, analysis, and visualization. Libraries such as Pandas, NumPy, and Matplotlib were utilized for these tasks.
* SQL: SQL queries were employed to extract and manipulate data from the MySQL database, which serves as the primary data source.
* MySQL: MySQL was used to store and manage the structured data, making it accessible for analysis and reporting.
* Tableau: Tableau was utilized to create interactive and insightful data visualizations, allowing for a better understanding of the analysed data.
* Power BI: Power BI was used to generate interactive reports and dashboards, enabling stakeholders to explore the project’s findings.

Project Highlights:

* Data Preparation: Describe any significant data preprocessing steps, such as data cleaning, transformation, or feature engineering.
* SQL Queries: Highlight key SQL queries or operations performed on the database.
* Visualization: Mention the main visualizations created using Tableau and Power BI, and how they contribute to the project’s objectives.
* Insights: Summarize the key insights or findings derived from the analysis.

Folder Structure:

* DataBase : This folder contain the YouTube and Spotify Data.
* Power Bi : This folder contains the ‘POWER BI’ works for visualizations.
* Python Work : This folder contains the ‘PYTHON WORK’ for cleaning the DataBase.
* SQL Work : This folder contains the ‘SQL WORK’ .
* Tableau DashBoard : This folder contains the ‘TABLEAU WORK’ for

visualizations.

* Cleaned DataBase : This folder contains the cleaned data which we use for SQL Work.

SQL WORKS (Some of SQL Queries) –

1. list of unique Artist list

Query : select distinct Artist from Spotify\_Youtube;

1. Count total number of Unique Artist

Query : select count(distinct Artist) as Artist from Spotify\_Youtube;

1. Cout the number of tracks for each album

Query : select Album\_type ,( count(Track) ) as 'Total Tracks' from Spotify\_Youtube group by Album\_type

1. Total Album

Query : select sum(totaltracks) as totaltracks from (select Album\_type , count(\*) as totaltracks from Spotify\_Youtube group by Album\_type ) as subquery;

1. Album Contribution

Query : select Album\_type, count(\*) as 'Total Tracks' , concat(cast(((count(\*)\*1.0 / ( select sum(totaltracks) as totaltracks from (select count(\*) as totaltracks from Spotify\_Youtube ) as subquery ) \*1.0)\*100)as decimal(10,2)) , '%') as Precentage from Spotify\_Youtube group by Album\_type

1. Top 10 Artist basis on Views On YouTube

Query : select top 10 Artist , sum(Views) as 'Total number of views' from Spotify\_Youtube group by Artist order by sum(Views) desc

1. Top 10 Artist basis on Stream On Spotify

Query : select top 10 Artist , sum(Duration\_minute) as 'Total number of views' from Spotify\_Youtube group by Artist order by sum(Duration\_minute) desc

1. Top 10 commented songs On YouTube

Query : select top 10 Track , Artist, Views , Album,Uri , Channel, Comments as 'Total number of Comments' from Spotify\_Youtube order by Comments desc

1. Top 10 Liked songs On YouTube

Query : select top 10 Track , Artist, Album,Uri , Channel, Likes as 'Total number of Likes' from Spotify\_Youtube order by Likes desc

1. Top 10 channel Basis on views

Query : select top 10 Channel,sum(Views) as 'Total number of views' from Spotify\_Youtube group by Channel order by sum(Views) desc

1. Top 10 highest duration songs

Query : select top 10 Track ,concat(cast(Duration\_minute as decimal (10,2)) ,'Minute')as Total\_Minute from Spotify\_Youtube group by Track,Duration\_minute order by Duration\_minute desc

1. Total Licenced songs

Query : select count(Licensed) as 'Total Licensed Songs' from Spotify\_Youtube where Licensed=1

1. Total Unlicenced songs

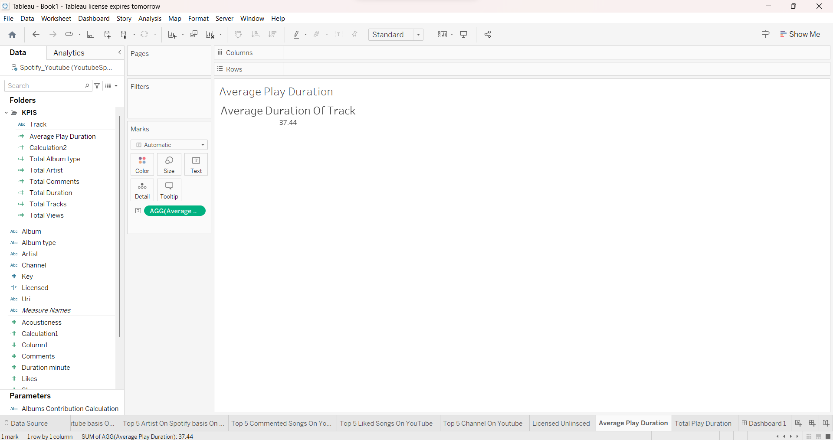
Query : select count(Licensed) as 'Total UnLicensed Songs' from Spotify\_Youtube where Licensed=0

1. Average time duration for Tracks

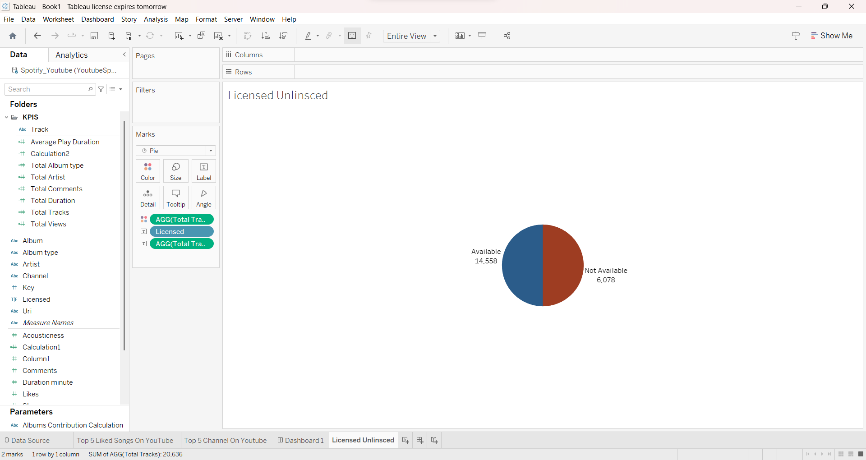
Query : select concat(cast((sum(Duration\_minute) / count(\*)) as decimal(10,2)) ,' Minute') as 'Average Time Duration' from Spotify\_Youtube

Creating Visualizations:

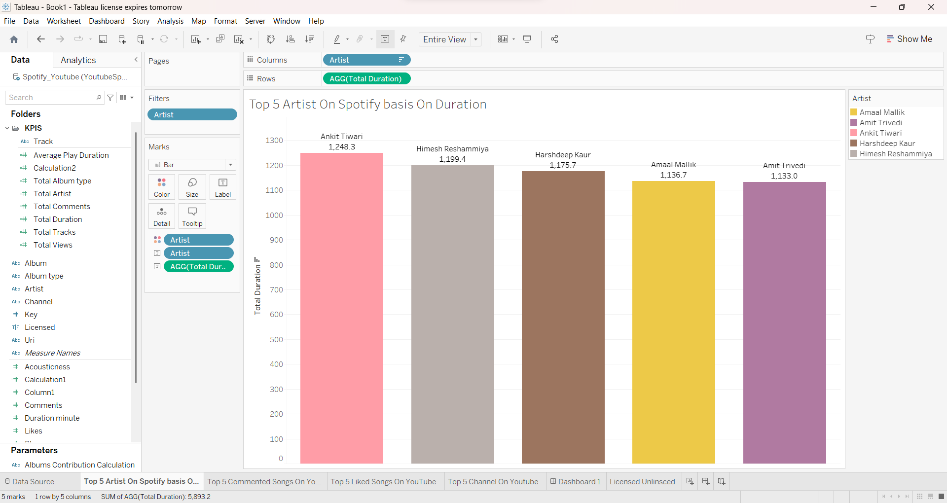
1. Tableau Visualizations
2. Average Play duration



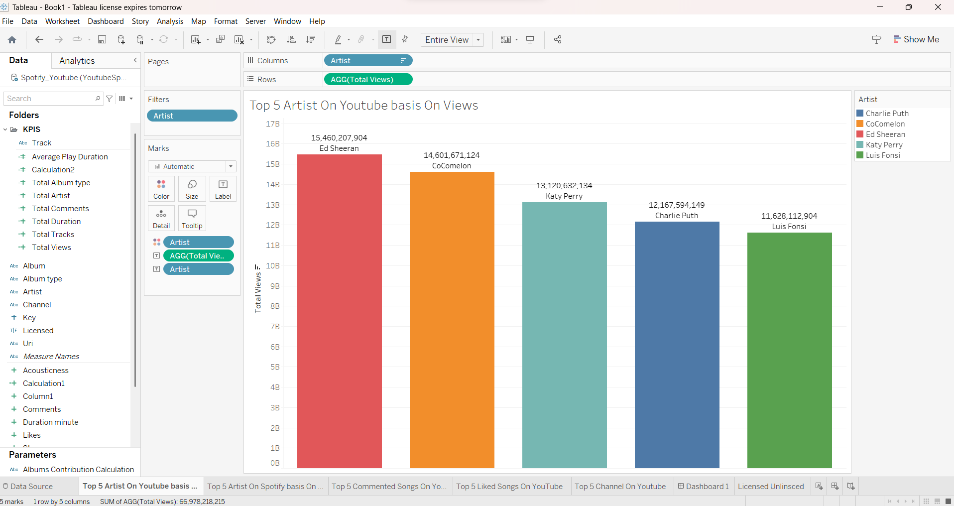
1. licensed And Unlicensed track



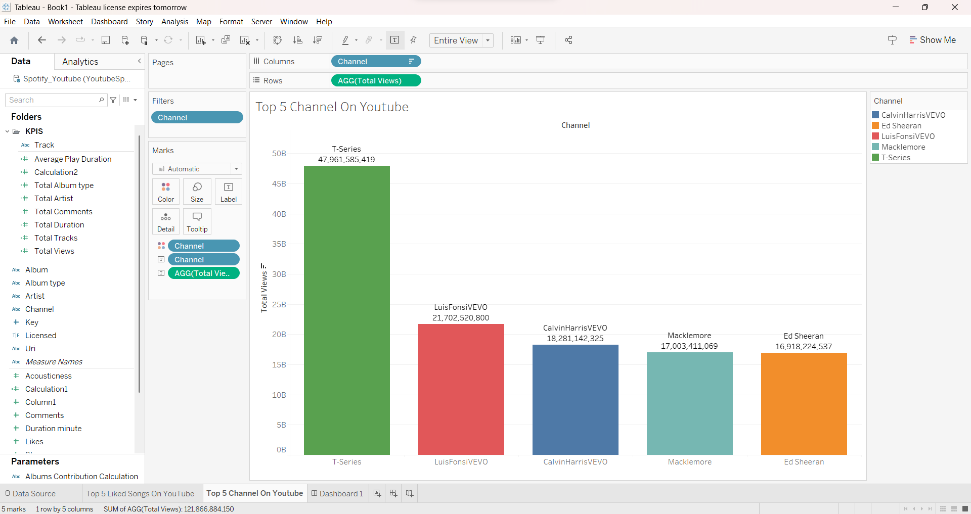
1. Top 5 artist on Spotify



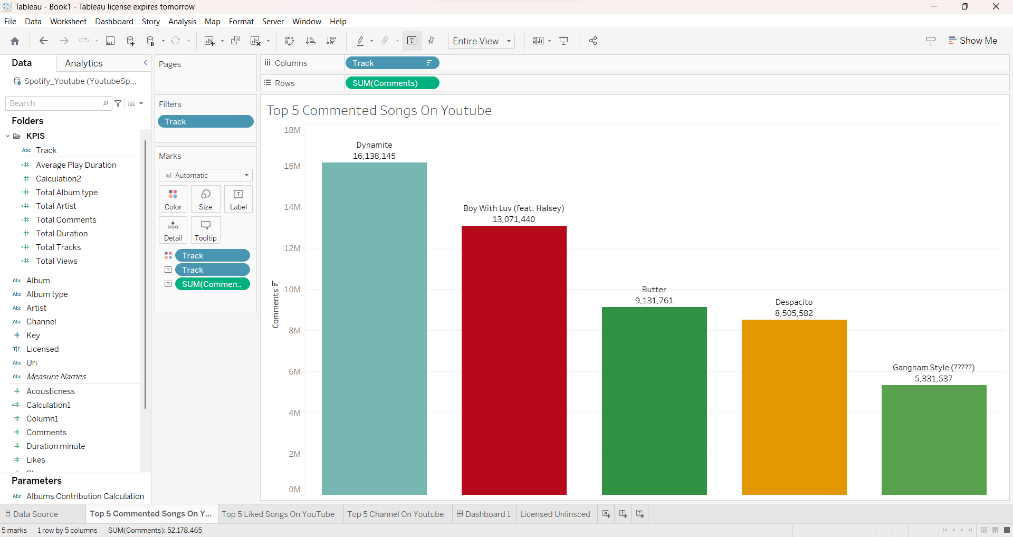
1. Top 5 artist on YouTube



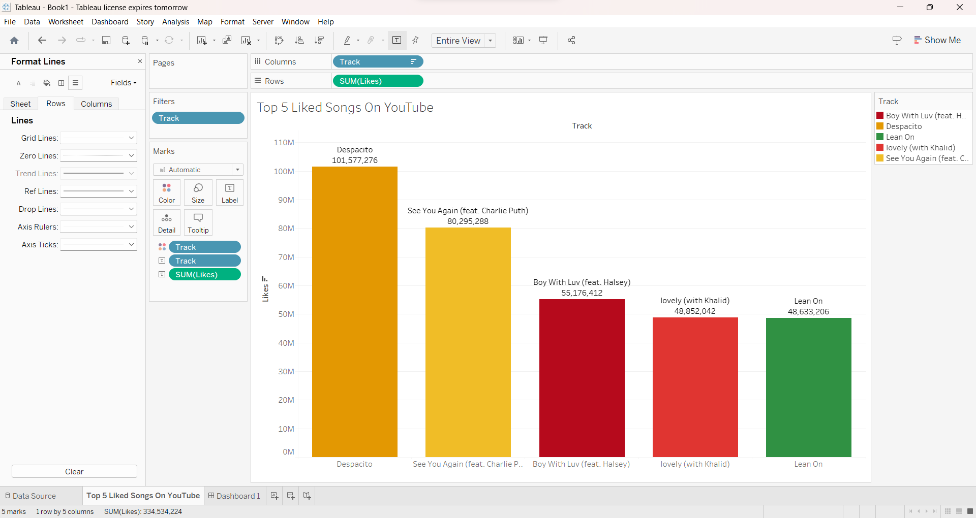
1. Top 5 channel on YouTube



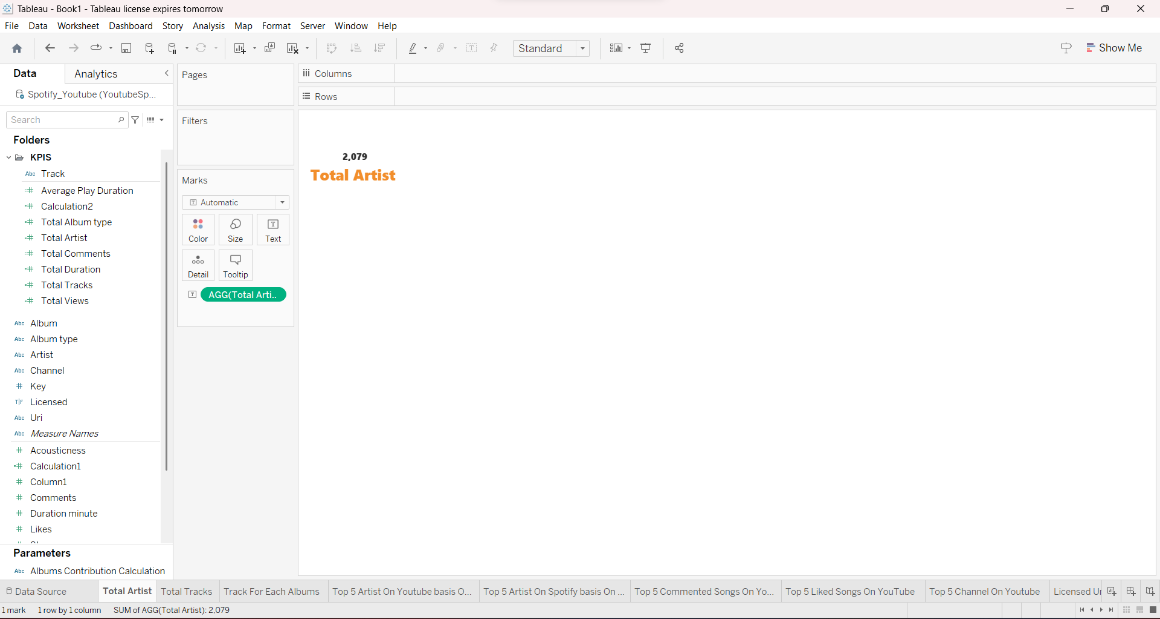
1. Top 5 commented songs



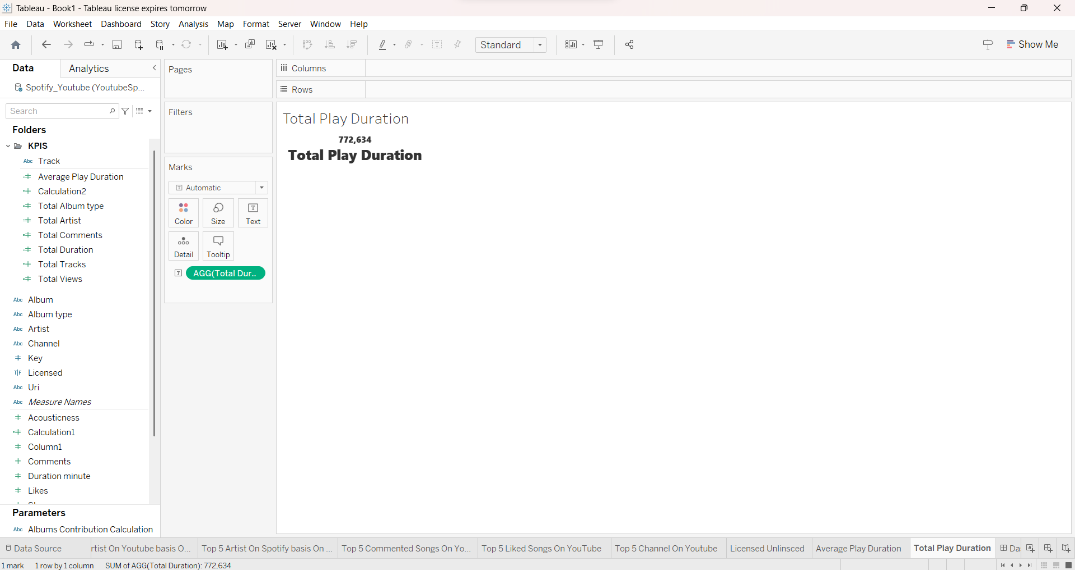
1. Top 5 liked Songs



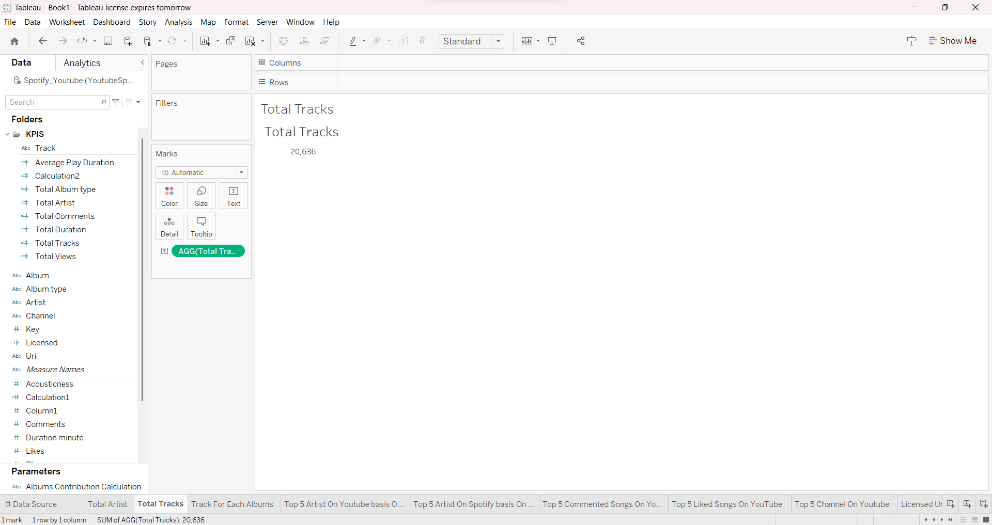
1. Total Artist



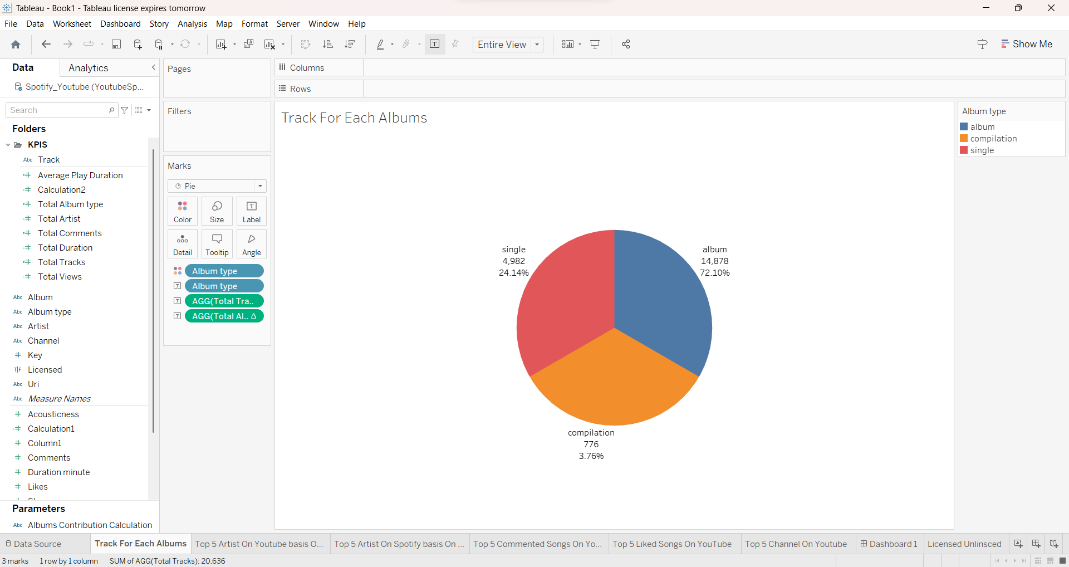
1. Total play duration



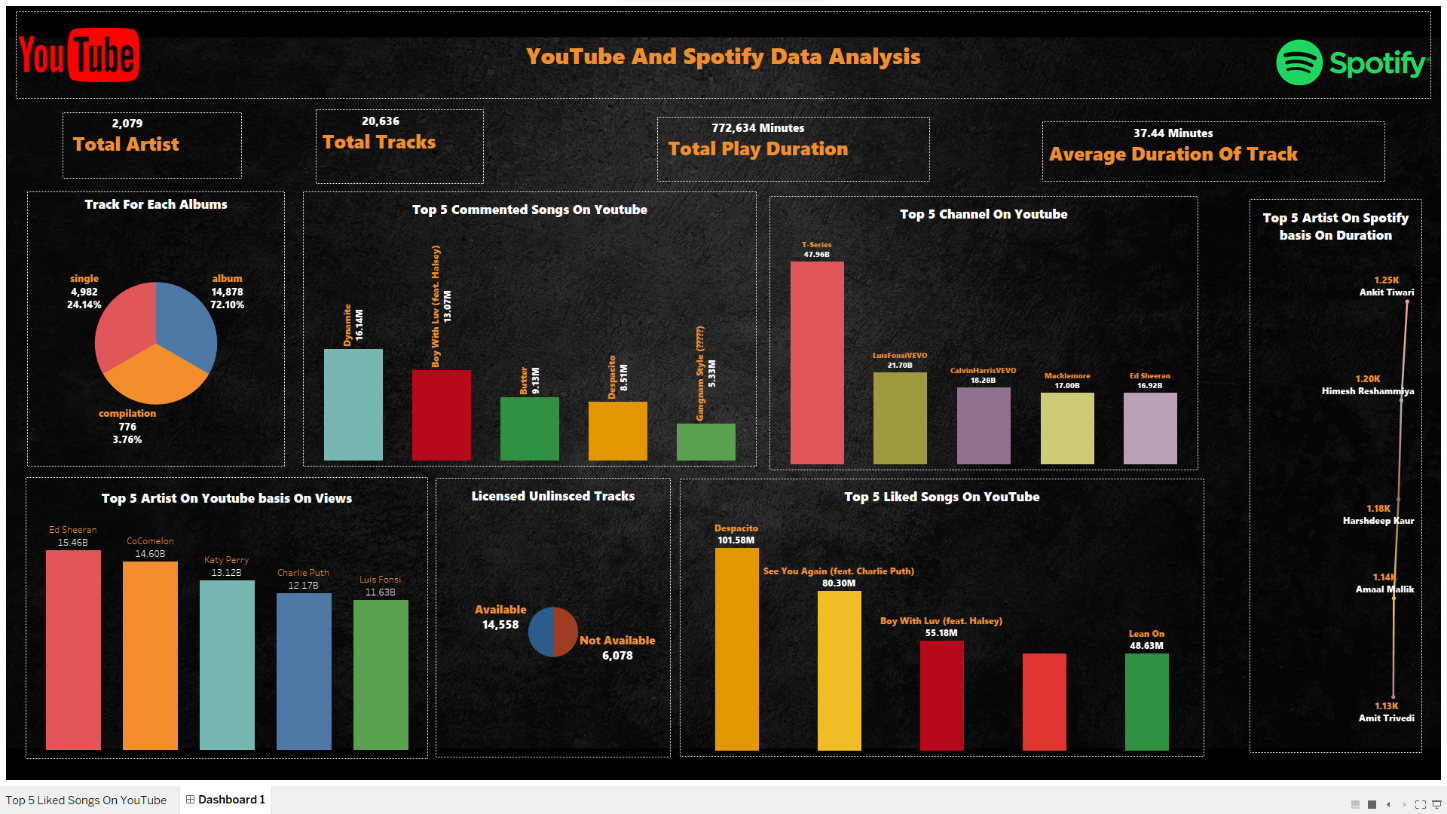
1. Total Tracks



1. Track of each Album

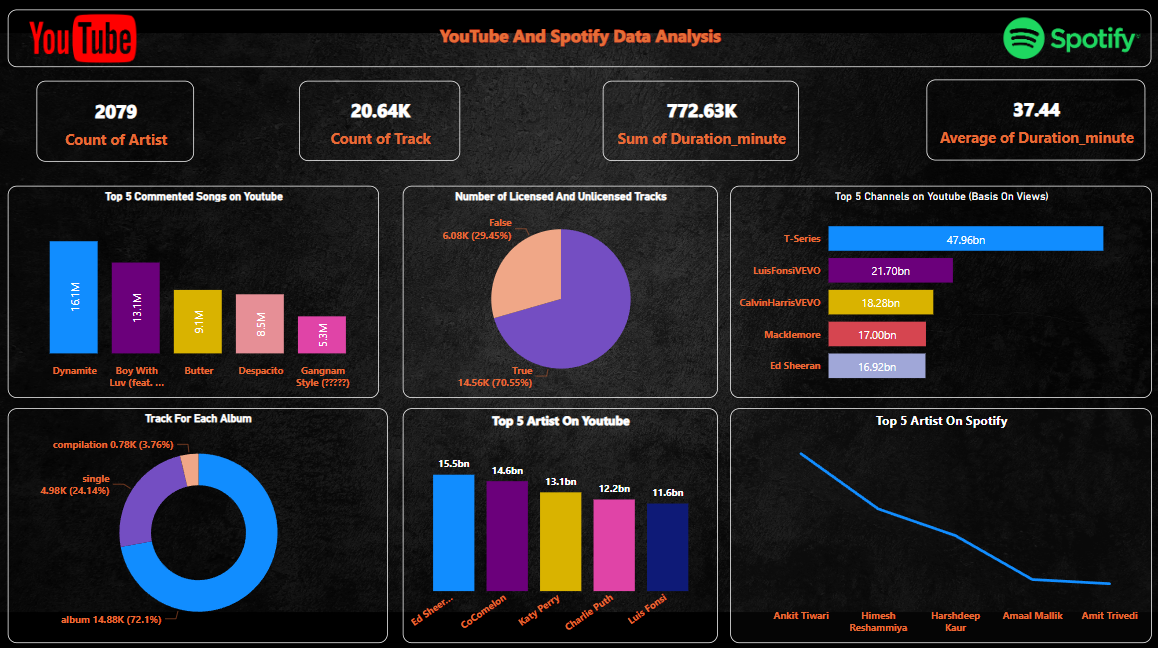


L. Final Dashboard



It’s Dynamic Dashboard

1. Power BI Dashboard



It’s Dynamic Dashboard