NETFLIX DATA ANALYTICS PROJECT REPORT

1. Introduction

The rapid growth of digital streaming platforms has transformed the way audiences consume entertainment content. Netflix, being one of the world's leading streaming services, continuously expands its content library to cater to diverse audiences across different regions.

This project aims to analyze Netflix's movies and TV shows dataset to understand its content distribution, growth trends, genre preferences, audience targeting, and geographical focus. The analysis uses data analytics tools such as Python, SQL, and Power BI concepts to derive meaningful business insights.

---

2. Objective of the Study

The main objectives of this project are:

- To analyze the distribution of Movies and TV Shows on Netflix
- To understand Netflix's content growth trend over the years
- To identify popular genres and audience ratings
- To examine country-wise content production
- To derive business insights that can support content strategy decisions

---

3. Dataset Description

The dataset used in this project is **Netflix Movies and TV Shows Dataset**, sourced from Kaggle. It contains information about Netflix's content catalog, including:

- Title
- Type (Movie / TV Show)
- Country
- Date Added
- Release Year
- Rating

- Duration
- Genre
- Description

After cleaning, the dataset was used for SQL analysis and dashboard design.

---

4. Tools and Technologies Used

- **Python**: Used for data cleaning and preparation
- **SQL**: Used for data analysis and querying
- **Power BI (Conceptual Design)**: Used for dashboard planning and visualization logic
- **GitHub**: Used for project version control and submission

---

5. Data Cleaning and Preparation (Python)

Python was used to clean and preprocess the raw Netflix dataset. The following steps were performed:

- Removal of duplicate records
- Handling missing values for columns such as country, director, cast, and rating
- Conversion of date fields into proper date formats
- Extraction of year from the date added column
- Splitting the duration column into numeric and categorical fields
- Creation of a primary genre column for simplified analysis

The cleaned dataset ensured accuracy and consistency for further analysis.

---

6. SQL Data Analysis

SQL was used to analyze the cleaned dataset and answer key business questions. The analysis focused on:

- Total number of titles available on Netflix
- Distribution of Movies vs TV Shows

- Year-wise content addition trends
- Top content-producing countries
- Most popular genres
- Rating-wise content distribution
- Average duration of movies
- TV shows based on number of seasons

These queries helped uncover meaningful patterns in Netflix's content strategy.

---

7. Power BI Dashboard Design (Conceptual)

A Power BI dashboard was designed conceptually to visualize key insights. The dashboard includes:

- KPI cards showing total titles, movies, TV shows, and average movie duration
- Bar chart comparing Movies and TV Shows
- Line chart showing content growth over the years
- Country-wise content distribution chart
- Genre-wise content popularity
- Rating distribution chart

The dashboard structure supports quick decision-making and strategic analysis.

---

8. Key Business Insights

- Netflix has a higher proportion of movies compared to TV shows, indicating a strong movie-focused catalog while steadily expanding episodic content.
- There has been a significant increase in content addition after 2016, highlighting Netflix's aggressive global expansion strategy.
- The United States and India are among the top content-producing countries, showing Netflix's focus on both global and regional markets.

- Drama and Comedy are the most dominant genres, making them reliable and high-demand content categories.
- A large share of content is rated TV-MA, indicating Netflix's primary focus on mature audiences.

---

9. Conclusion

This project demonstrates how data analytics can be used to derive valuable business insights from streaming platform data. By combining Python for data cleaning, SQL for analysis, and Power BI visualization concepts, the study provides a clear understanding of Netflix's content strategy and audience focus.

The insights obtained from this analysis can help streaming platforms optimize content investments, target the right audience segments, and plan future growth strategies.

---

10. Future Scope

- Sentiment analysis using content descriptions
- Viewer engagement analysis using external datasets
- Predictive analysis for content demand trends
