Project Title	**Netflix Data Analysis Using POWERBI**
Skills take away From This Project	• **Data Cleaning and Preprocessing** • **Exploratory Data Analysis (EDA)** • **Data Visualization using Power BI** • **Business Insights and Decision Making** • **Customer Behavior Analysis** • **Content Performance Analytics** • **Subscription and Engagement Analysis** • **Dashboard Creation in Power BI**
Domain	**Entertainment, Media, and OTT Analytics**

Problem Statement:

Analyze Netflix's content and customer viewing behavior to derive meaningful business insights. The goal is to clean and visualize the data to understand customer engagement, content popularity, and subscription trends.

Business Use Cases:

1. Identify the most popular movies and TV shows on Netflix based on watch time and ratings.
2. Analyze customer demographics to understand target audience preferences.
3. Understand the impact of subscription types on watch time and customer engagement.
4. Explore content ratings and their effect on viewer behavior.
5. Provide insights on the best time for content release to maximize engagement.
6. Assist Netflix in refining its content strategy based on genre popularity and user ratings.
7. Compare customer retention and engagement patterns across different countries.

Approach:

Data Cleaning & Transformation (Excel + Power BI)

1. Data Consistency Check:
    - Identify and fix missing values in director, cast, and country columns in the Netflix dataset.
    - Standardize date formats in the date_added column.
    - Ensure rating values are properly categorized.

2. Customer Demographics Analysis:
    - Group customers by age range (e.g., 13-18, 19-30, 31-45, 46-60, 60+) and analyze their distribution.
    - Identify which age group watches Netflix the most.
3. Subscription Type vs. Watch Time:
    - Analyze the relationship between subscription type (Basic, Standard, Premium) and total watch time.
    - Find which subscription tier has the highest engagement.

Descriptive & Comparative Analysis

4. Top 10 Most Watched Shows/Movies:
    - Identify the top 10 most-watched Netflix shows or movies based on total watch time.
    - Compare watch time for Movies vs. TV Shows.
5. Genre Popularity by Country:
    - Find the most popular genre per country based on the total number of viewers.
    - Visualize the trend in a world map chart.
6. Customer Ratings Distribution:
    - Analyze the distribution of user ratings (1-5 stars) for Netflix content.
    - Find the average rating per genre and per country.
7. Trending Content Analysis:
    - Identify shows or movies with the most recent additions (last 2 years).
    - Compare the average ratings of newly added vs. older content.

Market Insights & Revenue Analysis

8. Subscription Market Share Analysis:
    - Show the market share of Netflix subscriptions (Basic, Standard, Premium).
    - Identify which subscription plan dominates in different countries.
9. Customer Retention Insights:
    - Find out how many customers watch multiple shows vs. just one show.
    - Identify whether watch time influences customer retention (do heavy watchers give higher ratings?).

10. High-Rated vs. Low-Rated Content:
    - Identify which genres get the highest and lowest ratings.
    - Compare the ratings of TV Shows vs. Movies.

Business Decision-Making Scenarios

11. Director & Cast Influence on Ratings:
    - Analyze if movies from popular directors or casts receive higher ratings.
    - Find the top 5 highest-rated directors based on average user ratings.
12. Content Localization Strategy:
    - Identify countries where Netflix is underperforming in terms of engagement and ratings.
    - Suggest content strategies to improve engagement in those regions.
13. Best Time to Release New Content:
    - Analyze which month Netflix adds the most content.
    - Compare user ratings of content added in different months to determine best release periods.
14. Customer Engagement Dashboard:
    - Create a dashboard showing customer engagement trends (watch time, ratings, most popular genres, and subscription types).
    - Use slicers to filter by country, age group, and subscription type.
15. Competitor Benchmarking Simulation:
    - Compare Netflix's content availability and ratings by country.
    - If Netflix had a competitor with different content strategies, what improvements should Netflix make based on user preferences?

Results:

- Clean and structured datasets ready for analysis.
- A Power BI dashboard with key insights on content performance and customer behavior.
- Actionable business recommendations with PPT presentation for Netflix's content and subscription strategy

Project Evaluation metrics:

1. **Data Cleaning Efficiency:**
    - Proper handling of missing and inconsistent data.
2. **Depth of Business Insights:**

- Relevance and usefulness of insights derived.

3. **Visualization Effectiveness:**
   - Clarity, accuracy, and interactivity of Power BI dashboards.
4. **Analytical Approach:**
   - Logical approach in answering business questions.
5. **Presentation and Documentation:**
   - Well-structured report and presentation of findings.

Technical Tags:

- Power BI
- Excel
- Data Cleaning
- Business Intelligence
- Customer Analysis
- Content Analytics
- Data Visualization
- EDA

Data Set:

NETFLIX-DATASETS-2

1. **Netflix Shows Dataset**
   - **Format:** CSV
   - **Key Variables:** show_id, type, title, director, cast, country, date_added, release_year, rating, duration, genre.
2. **Netflix Customer Ratings Dataset**
   - **Format:** CSV
   - **Key Variables:** customer_id, show_id, customer_name, age, gender, country, subscription_type, watch_time_minutes, rating.

Data Set Explanation:

1. **show_id (String):** Unique identifier for each movie or TV show.
2. **type (String):** Specifies whether the content is a "Movie" or "TV Show."

3. **title (String):** Name of the movie or TV show.
4. **director (String):** Name of the director of the content (if available).
5. **cast (String):** List of main actors in the content.
6. **country (String):** The country where the content was produced.
7. **date_added (Date):** Date when the content was added to Netflix.
8. **release_year (Integer):** The year when the content was originally released.
9. **rating (String):** Content rating such as "PG", "TV-MA", etc.

G (General Audience) – Suitable for all ages.
PG (Parental Guidance Suggested) – Some material may not be suitable for children.
PG-13 (Parents Strongly Cautioned) – Some content may be inappropriate for children under 13.
R (Restricted) – Viewers under 17 require accompanying parent or adult guardian.
TV-Y (All Children) – Suitable for all children, designed for young kids.
TV-Y7 (Directed to Older Children) – Suitable for children aged 7 and older.
TV-G (General Audience for TV) – Suitable for all ages, no content unsuitable for children.
TV-PG (Parental Guidance Suggested for TV) – Some material may not be suitable for younger children.
TV-14 (Parents Strongly Cautioned for TV) – May be unsuitable for children under 14.
TV-MA (Mature Audience Only) – Intended for adults, may contain explicit content.

10. **duration (String):** Specifies duration in minutes for movies or seasons for TV shows.
11. **genre (String):** The genre or category of the content.

Customer Ratings Dataset Explanation:

1. **customer_id (String):** Unique identifier for each customer.
2. **show_id (String):** The unique identifier linking to the Netflix dataset.
3. **customer_name (String):** Name of the customer.
4. **age (Integer):** Age of the customer.
5. **gender (String):** Gender of the customer (Male, Female, Other).
6. **country (String):** The country where the customer resides.
7. **subscription_type (String):** Type of Netflix subscription (Basic, Standard, Premium).
8. **watch_time_minutes (Integer):** The total minutes the customer has watched content.
9. **rating (Integer):** Customer's rating for the show/movie (Scale of 1-5).

Project Deliverables:

- **Power BI Dashboard** with interactive visuals.
- **Excel File** with cleaned datasets (if required).
- **Presentation** with insights and recommendations with **Business Report** summarizing key findings..

Project Guidelines:

1. **Follow a structured approach** – Clean data first, then analyze and visualize.
2. **Ensure consistency in data formats** – Handle missing and incorrect data properly.
3. **Use relevant visualizations** – Choose the best charts (bar, line, pie, heatmaps) to convey insights.
4. **Focus on business impact** – Recommendations should be actionable and data-driven.
5. **Apply Power BI best practices** – Use slicers, filters, and interactive visuals to enhance usability.

Timeline:

The project must be completed and submitted **within 7 days from the assigned date**.

Created By:	Verified By:	Approved By:
Gaurav Kumar	Pooranam Annamalai	Pooranam Annamalai