1. Data Collection

Source: Excel

- Title: Amazon Prime Movies and TV Shows
- Type : Movie and TV shows.
- Release Year**: The year the content was released.
- Runtime**: Duration of the movie or episodes.
- Rating**: IMDb rating or Amazon Prime's internal rating.
- Cast**: Lead actors or directors.
- Language**: Primary language of the content.
- Region**: Availability based on geographical location.

2. Data Cleaning and Preparation
- **Handling Missing Data**: Fill or remove rows with missing values in critical fields like ratings or genres.
- **Standardization**: Convert runtime to minutes, genres into categorical variables, and release years into decades for better grouping.
- **Deduplication**: Remove duplicate entries to ensure data integrity.

3. Exploratory Data Analysis (EDA)
- Analyze the distribution of movies vs. TV shows.
- Explore popular genres and their ratings.
- Examine trends in content release over the years.
- Identify top-rated actors or directors on the platform.
- Determine the average runtime and its correlation with viewer ratings.

4. Visualizations
Using tools Power BI.
- bar chart
- filled map
- donut chart
- area chart

5. Insights a
From the analysis, you can provide insights such as:
- Content Trends**: Genres gaining popularity over time.
- Audience Preferences**: Genres and languages that receive the highest ratings.
- Optimization Opportunities**: Regions with fewer titles could be targeted for more localized content production.
- Release Timing**: Optimal periods for releasing new content based on trends.

6. Advanced Features
- Machine Learning (Optional)**: Build a recommendation engine using collaborative filtering to suggest movies/shows to users based on their viewing history.
- Sentiment Analysis**: Analyze user reviews to understand viewer sentiment towards specific content.

Final Presentation
- Dashboard*: Use tools Power BI