**Project Title: Amazon Prime Video Titles**

**Objective:**

To analyze the Amazon Prime Video catalog to uncover trends in content types, country contributions, genres, ratings, and popular creators. The project aims to deliver insights into how the platform has evolved over time and what themes dominate the streaming service.

**Project Phases:**

**1. Data Collection**

* Load the dataset (typically CSV) containing Amazon Prime Video titles.
* Check for columns such as title, type, country, date\_added, release\_year, rating, listed\_in, director, cast, and description.

**2. Data Cleaning & Preparation**

* Convert date\_added to datetime format and extract month and year.
* Handle missing values (e.g., missing country, cast, or director fields).
* Normalize entries (e.g., split multiple countries or genres).
* Remove or correct inconsistent data (e.g., duplicates, malformed ratings).

**3. Exploratory Data Analysis (EDA)**

* **Content Type Distribution**: Movies vs. TV Shows
* **Time Trends**:
  + Monthly and yearly additions
  + Release year distribution
* **Geographic Trends**: Top contributing countries
* **Genres & Ratings**: Most common genres and age ratings
* **Creator Analysis**:
  + Cross-analysis of directors/actors by genre or rating

**4. Text Analysis**

* Perform basic Natural Language Processing (NLP) on the description column:
  + Clean and tokenize text
  + Generate WordClouds

**5. Visualization & Interpretation**

* Create clear and informative plots using:
  + Matplotlib and Seaborn for bar plots, line charts, and histograms
  + WordCloud for visualizing frequent keywords in descriptions
* Annotate visualizations to highlight key insights

**Tools & Technologies:**

* **Programming Language**: Python
* **Libraries**: Pandas, WordCloud, datetime, dash, base64, BytesIO
* **IDE**: Jupyter Notebook