**Project Title: Disney+ Titles – Streaming Content Analysis and Insights**

**Objective:**

To analyze the Disney+ streaming catalog and uncover trends in content distribution, country of origin, genres, ratings, and creator activity. The project also includes text analysis on content descriptions to reveal dominant themes and language patterns.

**Project Phases:**

**1. Data Collection**

* Load the dataset (CSV) that includes fields such as:
  + title, type (Movie/TV Show), date\_added, release\_year, country, rating, listed\_in (genres), director, cast, and description.

**2. Data Cleaning & Preparation**

* Convert date\_added to datetime; extract month and year for trend analysis.
* Handle missing or null values (e.g., in country, cast, or director).
* Normalize text columns:
  + Split genres and countries into lists
  + Strip whitespace and fix inconsistent formats
* Create additional features if useful (e.g., content age = current year - release year)

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

* **Content Type Distribution**:
  + Number of Movies vs. TV Shows
* **Temporal Trends**:
  + Monthly and yearly additions to the platform
  + Release year frequency
* **Geographic Insights**:
  + Top contributing countries
* **Genres & Ratings**:
  + Most frequent genres
  + Age rating distributions
* **Creator Insights**:
  + Most prolific directors and actors
  + Compare their outputs by content type or genre

**4. Text Analysis**

* Analyze description text field:
  + Clean text (lowercase, remove punctuation/stopwords)
  + Tokenize and generate WordClouds
    - Overall
    - Separate for Movies vs. TV Shows
    - Optional: per genre or time period
* Identify common keywords and themes

**5. Visualizations**

* **Matplotlib & Seaborn**:
  + Bar plots, pie charts, and histograms for:
    - Type, year, rating, genre, and country
  + Line plots for monthly/yearly trends
* **WordCloud**:
  + Visualize frequent words in content descriptions
  + Comparison between types/genres

**🧰 Tools & Technologies:**

* **Programming Language**: Python
* **Libraries**: Pandas, Matplotlib, Seaborn, WordCloud, datetime
* **IDE**: Jupyter Notebook