**Project Title: Netflix Titles – Streaming Content Analysis and Visualization**

**Objective:**

To analyze and visualize trends in the Netflix catalog to understand how the platform’s content has evolved over time in terms of type, country of origin, genres, ratings, and creative contributors. Additionally, perform text analysis to extract common themes from content descriptions.

**Project Phases:**

**1. Data Collection**

* Load the Netflix Titles dataset (CSV format) which includes fields like:
  + title, type, 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 from it.
* Handle missing values (e.g., missing director, cast, or country).
* Normalize categorical fields (e.g., split multiple countries or genres).
* Clean and prepare text data for NLP analysis.

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

* **Content Type Breakdown**:
  + Number of Movies vs. TV Shows
* **Temporal Trends**:
  + Monthly and yearly additions to the platform
  + Distribution by release\_year
* **Geographic Distribution**:
  + Top countries contributing content
* **Genre and Rating Analysis**:
  + Most common genres and age ratings
  + Compare distributions for Movies vs. TV Shows
* **Creator Analysis**:
  + Most featured directors and actors

**4. Text Analysis**

* Preprocess description field:
  + Remove punctuation, stopwords; convert to lowercase
* Tokenize and generate WordClouds for:
  + All content
  + Separate Movies and TV Shows
* Highlight common themes and patterns

**5. Visualization & Dashboard**

* **Using Python (Matplotlib, Seaborn):**
  + Bar plots: content type, ratings, top genres
  + Line plots: trends in content additions over time
  + Histograms: release year distribution
* **Using Tableau:**
  + Interactive dashboard to:
    - Filter by type, genre, country, rating
    - Visualize time trends, regional breakdowns
    - Compare Movies vs. TV Shows with KPIs and charts

**Tools & Technologies:**

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