Project Description:

This project involves analyzing the **YouTube 2025 dataset** (available in CSV format on Kaggle) using Python and MySQL. The primary objective is to extract meaningful insights from the data by performing data cleaning, migration, and statistical analysis using popular Python libraries.

Key Steps:

1. **Dataset Acquisition:**
    - Download the **YouTube 2025 CSV dataset** from Kaggle.
    - Store it locally for further processing using Python.
2. **Python File Handling:**
    - Use Python's built-in file handling features to read and process the CSV file.
    - Ensure the file is opened, read, and parsed efficiently, with error handling in place for robustness.
3. **Data Cleaning & Sanitization:**
    - Utilize **Pandas** to load the dataset into a DataFrame.
    - Perform data cleaning tasks such as:
        - Removing duplicates
        - Handling missing or null values
        - Correcting data types and inconsistent formats
        - Renaming columns for clarity
        - Filtering irrelevant or noisy data
4. **MySQL Integration:**
    - Design and create an appropriate table structure in **MySQL** to store the cleaned dataset.
    - Use **MySQL Connector (or SQLAlchemy)** in Python to establish a connection with the MySQL database.
    - Migrate the cleaned dataset from the Pandas DataFrame into the MySQL table.
5. **Analytical Computation:**
    - Perform key **statistical and analytical calculations** such as:
        - Total number of videos

        - Most viewed categories
        - Average likes, dislikes, and comments
        - Distribution of views by channel or category
        - Trends based on upload date or region

6. **Console Output:**
    - Use **Python libraries** like pandas, numpy, and matplotlib/seaborn (optional for console summaries,) to display:
        - Summary statistics
        - Top trending videos
        - Category-wise comparisons
        - Any other meaningful patterns identified in the dataset

n