NAAN MUDHALVAN – MODEL PRACTICAL

1. **EDA ON GLOBAL SUPERSTORE SALES DATASET**

   Dataset: Global Superstore Dataset (Excel)

   Objective: Analyze the sales performance across regions and customer segments.

   Instructions:

   - Load the dataset using Pandas.
   - Clean missing data and remove duplicates.
   - Calculate summary statistics (mean, median, standard deviation) for Sales and Profit.
   - ***Analyze:***
     - Total sales per region.
     - Top 5 most profitable product categories.
     - Year-wise sales trend.
   - ***Visualizations:***
     - Bar chart: Sales by region.
     - Line chart: Year-wise sales trend.
   - Write brief insights for each chart.

2. **EDA ON COVID-19 GLOBAL DATASET**

   Dataset: COVID-19 World Data (Johns Hopkins)
   Objective: Explore the spread and impact of COVID-19 across different states in India.
   Instructions:
   - Load the dataset and inspect the first few rows.
   - Handle missing data and convert date columns to datetime format.
   - ***Calculate:***
     - Total confirmed, recovered, and death cases for each state.
     - State with the highest number of cases.
     - Daily trend of new cases.
   - ***Visualizations:***
     - Pie chart: Top 5 states by confirmed cases.
     - Line graph: Trend of daily confirmed cases.
   - Write observations based on data and graphs.

3. **EDA ON YOUTUBE TRENDING VIDEOS DATASET**

   Dataset: YouTube Trending Videos
   Objective: Explore video trends, engagement metrics, and content categories.

Instructions:

- Load and clean the dataset using Pandas.
- Handle missing or null values and remove duplicates.
- ***Calculate:***
    - Most common video categories.
    - Top 5 channels with the highest number of trending videos.
    - Average likes, views, and comments.
- ***Visualizations:***
    - Bar chart: Video count by category.
    - Scatter plot: Likes vs. Views.
- Write insights describing trends in video popularity and engagement.