? Real-World Population Growth Analysis Using UN Dataset & Logistic Simulation

---

□ Objective

To analyze historical population growth of **Pakistan** using actual UN-provided population data, and **compare it** with a **theoretical simulation** based on the **logistic growth model**, using `NumPy`, `Pandas`, and `Matplotlib`.

□ Dataset Details

- **Filename:** `population.csv`
- **Source:** UN World Population Data
- **Columns in CSV:**
  - `Country Name`
  - `Country Code`
  - `Year`
  - `Value` *(Population in that year)*
- **Time Range:** 1960 to most recent available year
- **Target Country:** Pakistan

□ Tools & Libraries Used

- **Python Libraries:**
  - `pandas`: Data loading, filtering, and processing
  - `numpy`: Simulation of logistic growth (mathematical modeling)
  - `matplotlib.pyplot`: Data visualization
  - `seaborn`: Improved styling of plots
- **Environment:** Jupyter Notebook
- **Platform:** Local (Anaconda, Windows 10)

□ Project Steps (Jupyter Notebook Flow)

Step 1: Load and Inspect Dataset

- Loaded `population.csv` with `pandas.read_csv()`
- Inspected columns and values using `.head()`, `.info()`, `.describe()`

Step 2: Filter Data for Pakistan

- Selected rows where `Country Name == 'Pakistan'`
- Created a DataFrame `pak_data` with two columns:
  - `Year`
  - `Population`

Step 3: Visualize Real Population Growth

- Used `matplotlib` to plot real growth over time
- Title: "□ Pakistan Population Growth (1960 - Present)"

Step 4: Simulate Logistic Growth

- Used the **logistic growth model**:

  P(t)=K1+(K−P0P0)e−rtP(t) = \frac{K}{1 + \left(\frac{K - P_0}{P_0} \right)e^{-rt}}P(t)=1+(P0K−P0)e−rtK

  Where:

  - P0P_0P0: Initial population (1960)
  - rrr: Growth rate (chosen value, e.g., 0.03)
  - KKK: Carrying capacity (e.g., 400 million)
  - ttt: Time (years)
- Used `numpy` to apply this formula and generate simulated values

Step 5: Compare Real vs Simulated Growth

- Plotted **both curves** on the same graph:
  - Real data (blue)
  - Simulated logistic growth (orange)
- Added legends, titles, and axis labels

□ Key Outputs

- A **line plot** showing real growth trend of Pakistan's population (from 1960 onward)
- A **simulated curve** that predicts future growth, slowing as it nears carrying capacity
- A **comparison chart** that shows how real-world data aligns or differs from the model

---

□ Insights Gained

- Pakistan's actual population growth is **exponential** up to recent years
- Logistic model helps **predict a saturation point** where growth slows
- Combining real data and simulation gives more **realistic forecasting tools**

□ File Structure

- `population.csv` → The source dataset
- `real_population_analysis.ipynb` → Your full working notebook
- `Population_Growth_Project_Summary.pdf` → Summary export (optional)