**Crime Data Analysis in India (2001–2012)**

**🎯 Project Objective**

The objective of this project is to **analyze and visualize trends in crimes against Scheduled Castes (SCs) in India** between 2001 and 2012. The project aims to:

* Examine **yearly trends** in crime rates.
* Identify **state-wise** and **crime category-wise** patterns.
* Deliver **actionable insights** using interactive data visualizations.

**🛠️ Tools & Technologies**

* **Power BI** – For interactive dashboards and dynamic visual insights.
* **Python (Jupyter Notebook)** – For data cleaning, preprocessing, and exploratory analysis.
* **Libraries Used**: pandas, matplotlib, seaborn

**📊 Power BI Dashboard Design**

**📌 Page 1: National Overview**

* **KPI Cards**: Total Crimes, Max Crime Count, Case Count
* **Line Chart**: Crime trend (2001–2012)
* **Bar Chart**: Crimes by State/UT
* **Pie Chart**: Crime distribution by Category (Crime Head)
* **Gauge Chart**: Crime Count vs. Target
* **Table**: Year-wise crime data by state
* **Slicers**: Year, State/UT, Crime Head for interactive filtering

**🔎 Page 2: Detailed Drilldown**

* **Slicers**: State/UT, Crime Head, Year
* **Line Chart**: Crime Trend for Selected State
* **Bar Chart**: Crime Distribution by Crime Head
* **Matrix**: Year x Crime Head x Crime Count
* **Text Box**: Auto-generated insights
* **Navigation**: Back button for easy page transition

**📈 Key Insights**

* **Uttar Pradesh** had the **highest crime count against SCs in 2008**.
* **Murder cases** against SCs in **Bihar declined steadily** after 2010.
* **Rape cases** showed a **national increase from 2005 to 2012**.

**🐍 Python-Based Data Analysis**

**Tasks Performed:**

* Loaded and cleaned the dataset using pandas
* Performed grouping, aggregation, and plotting with seaborn and matplotlib
* Visualized trends by year and by state

**✅ Sample Python Code:**

python

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

# Load and clean the dataset

df = pd.read\_csv("crime\_data.csv")

df.dropna(inplace=True)

df.columns = df.columns.str.strip().str.upper()

# Aggregate crime data by year

yearly = df.groupby('YEAR')['VALUE'].sum().reset\_index()

# Plot crime trend over the years

sns.lineplot(data=yearly, x='YEAR', y='VALUE')

plt.title("Total Crimes Over Time")

plt.xlabel("Year")

plt.ylabel("Total Crime Count")

plt.xticks(rotation=45)

plt.tight\_layout()

plt.show()

**📁 Deliverables**

* 📊 **Power BI Dashboard** (.pbix)
* 📓 **Jupyter Notebook** (.ipynb)
* 📄 **README.md** – Project summary for GitHub
* 🖼️ **Visual Summary & Insights** – Charts and narrative interpretations