CLEANED_CRIME_DATA_SAMPLE

Project Documentation

NIKHIL FRANCIS
DAF 16 BATCH

Project Documentation

Cleaned_Crime_Data_Sample

1. **Project Overview**

   This project focuses on analyzing crime data to identify patterns, trends, and insights that can support law enforcement and policy-making. The dataset (Cleaned_Crime_Data_Sample.xlsx) contains 10,000 records of reported crimes, and a corresponding Power BI dashboard (.pbix) has been developed to visualize the findings interactively.

   Examine crime categories
   Analyze temporal trends
   Study victim demographics
   Support decision-making

2. **Tools Used**

   Microsoft Excel
   Used for initial data cleaning and formatting of the crime dataset.

   Power BI
   Main tool for data visualization and dashboard creation

3. **Dataset**

   The dataset used for this project is titled “Cleaned_Crime_Data_Sample.xlsx”, containing 10,000 crime records with 28 attributes. It provides structured information about crime incidents, victims, locations, and outcomes.

4. **Steps Followed**

   Step 1: Data Collection & Import
   Obtained the raw dataset of crime records.
   Imported the dataset into Excel cleaning and validation.

   Step 2: Data Cleaning & Preparation
   Checked for missing values, duplicates, and incorrect entries.
   Standardized date and time formats (DATE OCC, Date Rptd, TIME OCC).
   Verified categorical fields (e.g., Vict Sex, Vict Descent, AREA NAME).
   Validated numeric values like Vict Age and location coordinates (LAT, LON).

   Step 3: Data Exploration
   Identified major crime categories (Crm Cd Desc).
   Reviewed temporal trends (yearly, monthly, and daily distribution).
   Analyzed victim demographics (age, gender, descent).
   Checked crime hotspots using area and location data.

Step 4: Data Visualization (Power BI)
Imported the cleaned dataset into Power BI.
Built interactive dashboards and KPIs for:

- Top crime categories.
- Crimes over time (line charts, bar charts).
- Geographic distribution (maps).
- Victim demographics (pie charts, histograms).
- Case outcomes (arrests, investigations).

Step 5: Insight Generation
Summarized key patterns and trends from the visualizations.
Highlighted problem areas (e.g., high-crime zones, vulnerable groups).
Derived actionable insights for law enforcement and policy planning.

Step 6: Reporting
Compiled results into a structured written report.

5. Key Insights

Most Common Crimes

The top 5 crime categories are:
Vehicle Theft – 1,126 cases
Battery – Simple Assault – 840 cases
Burglary from Vehicle – 701 cases
Felony Vandalism – 686 cases
Aggravated Assault with Deadly Weapon – 590 cases

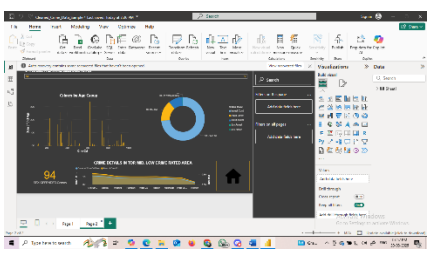
Crime Hotspots (Top Areas by Incidents)

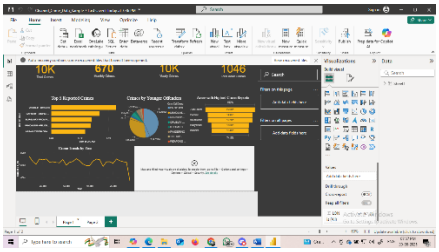
77th Street – 1,091 cases
Central – 992 cases
Southwest – 985 cases
Hollywood – 827 cases
Wilshire – 815 cases

Overall Observations

Vehicle theft and assault dominate the dataset.
Certain divisions (77th Street, Central, Southwest) are persistent hotspots.
Majority of victims are young males.
A large portion of cases remain under investigation, with fewer arrests.

6. Screenshots





7. Files Included

- Dataset File

  Name: Cleaned_Crime_Data_Sample.xlsx

  Format: Excel (.xlsx)

  Description: Contains 10,000 cleaned records of crime incidents with 28 attributes, including crime details, victim demographics, location data, and case outcomes.

- Dashboard File

  Name: Cleaned_Crime_Data_Sample.pbix

  Format: Power BI (.pbix)

  Description: Interactive dashboard built using the cleaned dataset.

- Report Document *(this report)*

  Format: PDF/Word (as required)

  Description: A structured document summarizing project overview, methodology, key insights, and visuals from the dataset.

  Purpose: Acts as a reference report for stakeholders who need insights without using Power BI.

8. How to Use

1. Excel Dataset (Cleaned_Crime_Data_Sample.xlsx)

   Open the file in Microsoft Excel (or any spreadsheet software).

   Use filters and pivot tables to explore crime categories, victim profiles, and area-wise distribution.

2. Power BI Dashboard (Cleaned_Crime_Data_Sample.pbix)

   Open the file in Microsoft Power BI Desktop.

   Navigate through different dashboard pages to explore:

   - Crime Frequency → Top crime types and trends.
   - Geographic Hotspots → Area-wise distribution and mapping.

9. Conclusion

The analysis of the crime dataset highlights several important patterns:

- Vehicle-related crimes and assaults are the most common incidents.
- 77th Street, Central, and Southwest divisions consistently emerge as high-crime hotspots.
- Young adults (20–40 years), especially males, form the majority of victims.
- A significant portion of cases remain under “Investigation Continuing”, with relatively low arrest rates.