**635 - Project Update Report**

**Team Members:** Suprithson Dubba.

Teja Savatapalli.

Navyasree Putluri.

**Dataset:** Crime Forecasting.

**Introduction:**

We focused on analyzing and visualizing incident report data to identify patterns and aid in decision-making. We are utilizing the dataset from the “Crime Forecasting”, - “NJ12017\_MAR01\_MAYR31.xlsx”, which contains detailed incident logs over a three-month period.

**Progress Summary:**

We have successfully imported our dataset into Python using the pandas library and have begun exploratory data analysis. By setting sheet\_name=None, we loaded the data across multiple sheets into a dictionary of data frames for comprehensive access.

**Data Cleaning and Preparation:**

Our data cleaning efforts included stripping unnecessary whitespace from headers and cells and ensuring that all date and time columns are in the correct format. We've handled NaN values judiciously and transformed certain text fields into more analytically suitable formats.

**Analysis and Findings:**

Key analytical steps taken include:

* Counting the number of incidents per day, which revealed fluctuations in incident reports over time.
* Categorizing incidents and visualizing the distribution, which highlighted "Street Crimes" as the most frequent occurrence.
* Conducting a time series analysis to understand incident frequency by hour, providing insights into peak times for incident reports. Visualizations using matplotlib have proven invaluable in revealing the daily and categorical patterns of incidents, thereby offering a clear visual representation of the data.

**Challenges and Resolutions:**

We encountered challenges with data types and missing values. These were resolved by converting data types and imputing or excluding NaN values based on the context. There were also challenges with plotting the data effectively, which were overcome through iterative improvements to our visualization scripts.

**Next Steps:**

As we progress, our action plan includes the following key initiatives:

* Grid Overlays on Portland Map (Fishnet): We will create grid overlays on the Portland map to analyze spatial patterns and pinpoint hotspots of criminal activity. This "fishnet" analysis will help us visualize and quantify crime distribution across the city.
* Spatio-Temporal Data Analysis: Delving into the spatio-temporal aspects of our data will allow us to understand not just where, but also when crimes are more likely to occur. This will involve creating time series models that account for spatial dependencies.
* Feature Correlation with Crime for Burglary in the 1st Week of 2013: We will investigate which features, such as time of day, location, or socio-economic indicators, are most strongly correlated with the incidence of burglary during the first week of 2013.
* Feature Correlation with Crime for Burglary Over 2 Weeks in 2013: Expanding our analysis, we will study the correlation of the same or additional features with burglary incidents over a two-week period in 2013 to validate our findings from the initial one-week analysis.
* Predictive Modeling: Building on the correlations found, we will develop predictive models to forecast incident trends and identify the likelihood of burglary based on the correlated features.
* Final Report Preparation: We will synthesize our findings, models, and insights into a comprehensive final report that will detail our analytical journey, from data cleaning to predictive modeling, culminating in actionable recommendations based on our analysis.

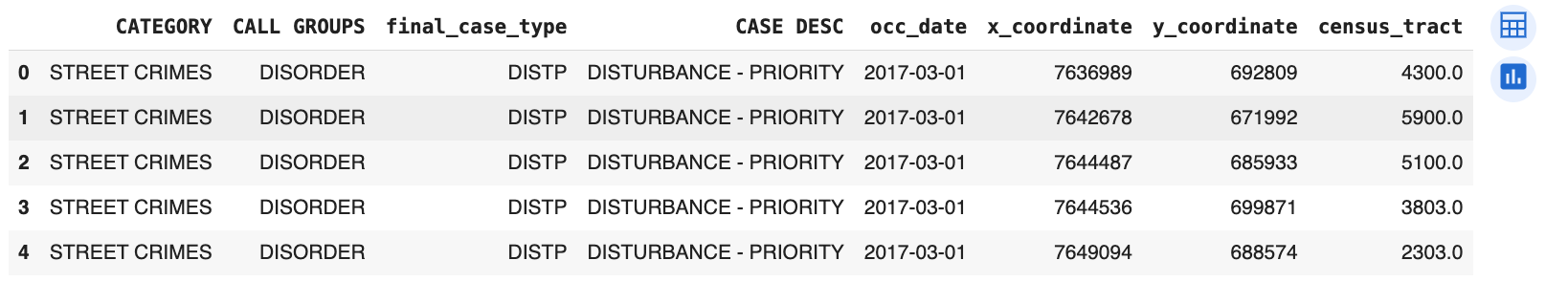
These steps are designed to enhance our understanding of the dynamics of crime within the area and will contribute significantly to the development of data-driven strategies for crime prevention and resource allocation.

**Conclusion:**

Our project is on track, having made substantial progress in data preparation, cleaning, and preliminary analysis. The insights gained thus far are promising and align well with our initial project goals of understanding incident dynamics and aiding predictive policing efforts.

**Appendices:**

Attached are various plots and charts generated during our analysis, including bar graphs for incident categories and line graphs for time-based trends.



A graph with blue lines

Description automatically generated

A white background with black text

Description automatically generated

A graph with blue squares

Description automatically generated

A graph with a green dot

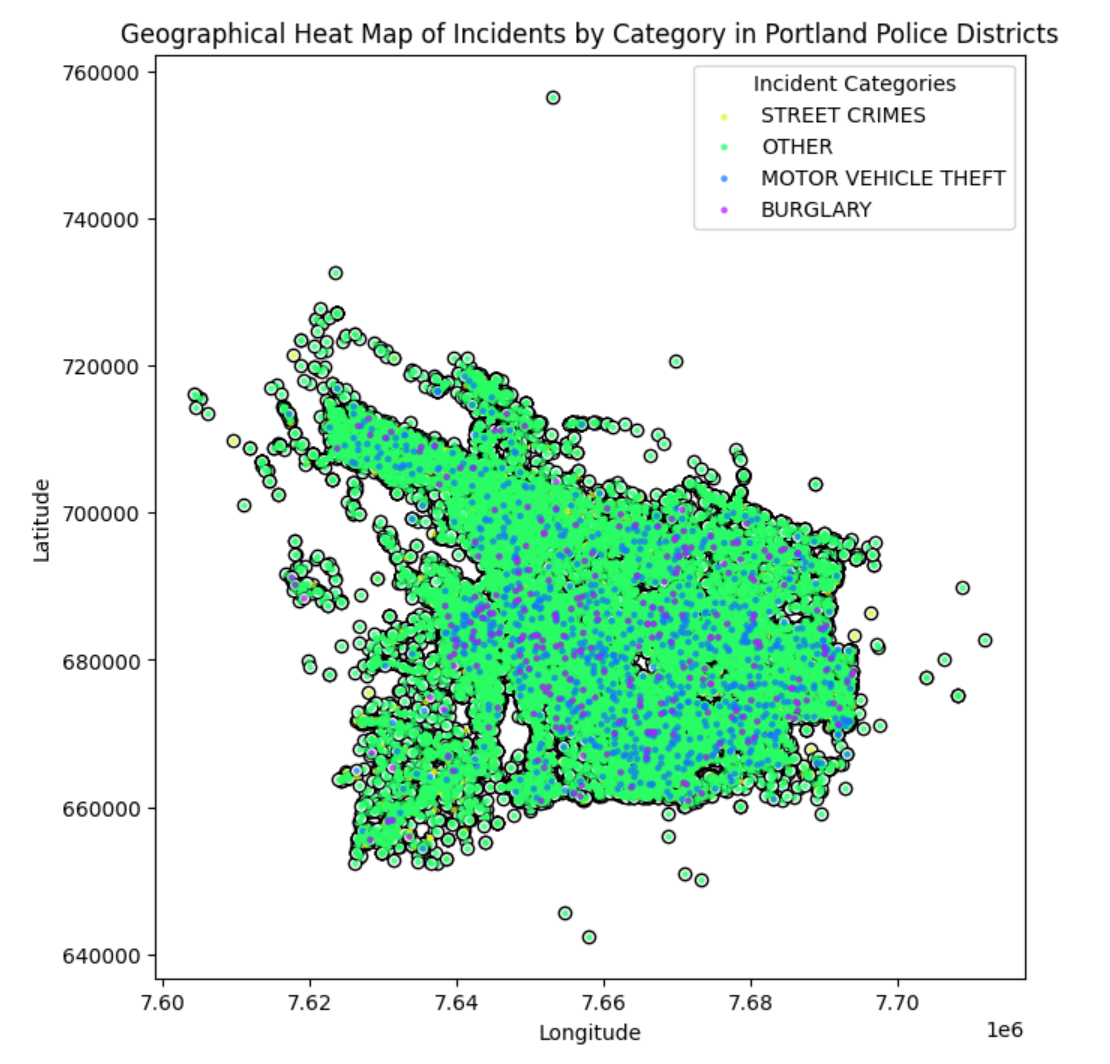
Description automatically generated

A screenshot of a computer code

Description automatically generated

A close-up of a list of text

Description automatically generated



**References:**

* Website: [Archived | Real-Time Crime Forecasting Challenge Posting | National Institute of Justice (ojp.gov)](https://nij.ojp.gov/funding/real-time-crime-forecasting-challenge-posting#data)
* Chainey, S., Thompson, L., & Uhligh, S. (2008). The Utility of Hotspot Mapping for Predicting Spatial Patterns of Crime. Security(21), 4-28.
* Hunt, J. (2016). Do Crime Hot Spots Move? Exploring the Effects of the Modifiable Areal Unit Problem and Modifiable Temporal Unit Problem on Crime Hot Spot Stability. Archived with ProQuest Dissertations & Theses.