

Crime Hotspot Prediction and Visualization in Atlanta

Aravind Rajeev Nair, Karan Nahar, Vamsi V Kalidindi, Krishna Raj, and Naveen Sethuraman
Team 40

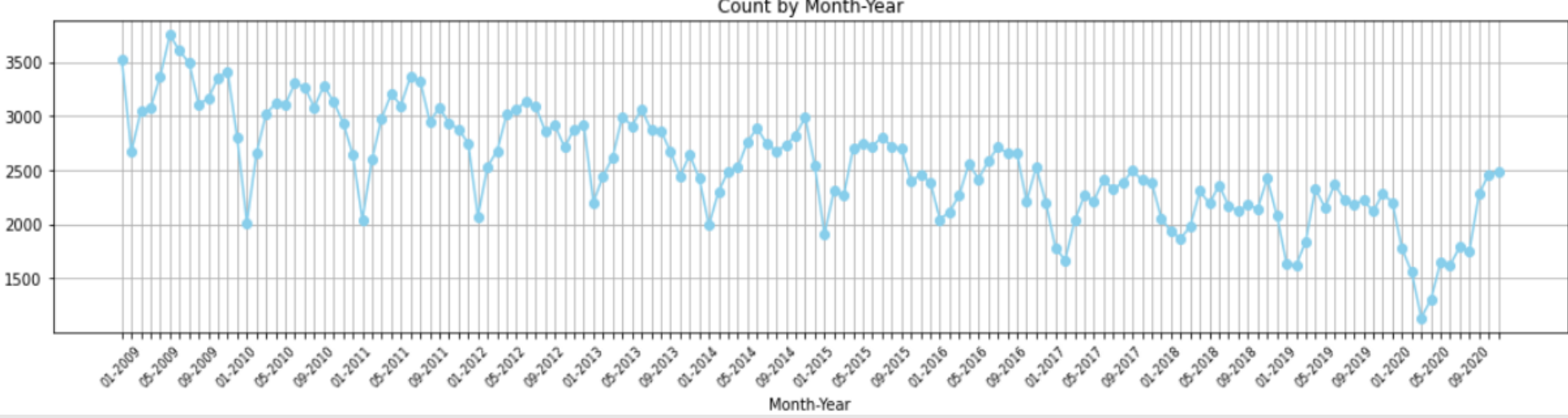
Introduction and Overview of Methodology

Our project aims to predict crime hot spots in Atlanta, assign risk scores to each neighborhood in Atlanta, and Visualize the results using an interactive map. The main components of our project are:

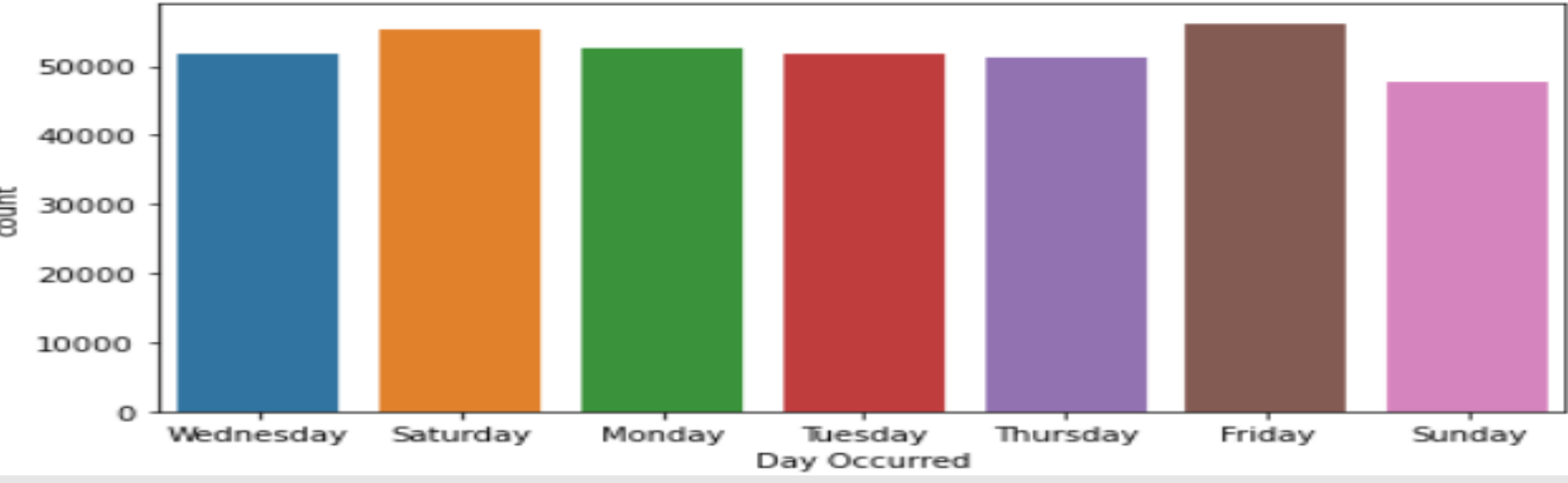
- Basic exploration of the data (EDA)
- Clustering of crimes using the GMM algorithm based on the latitude and longitude of the crime
- Geographical clusters assigned a risk score reflecting crime severity, providing a monthly and yearly risk profile for each area
- Time series model, Exponential Smoothing, used to forecast geographical crime risk scores from 2019 to 2020, utilizing data from 2009 to 2018
- Visualize this result on an Interactive map of Atlanta

1. Exploratory Data Analysis (EDA)

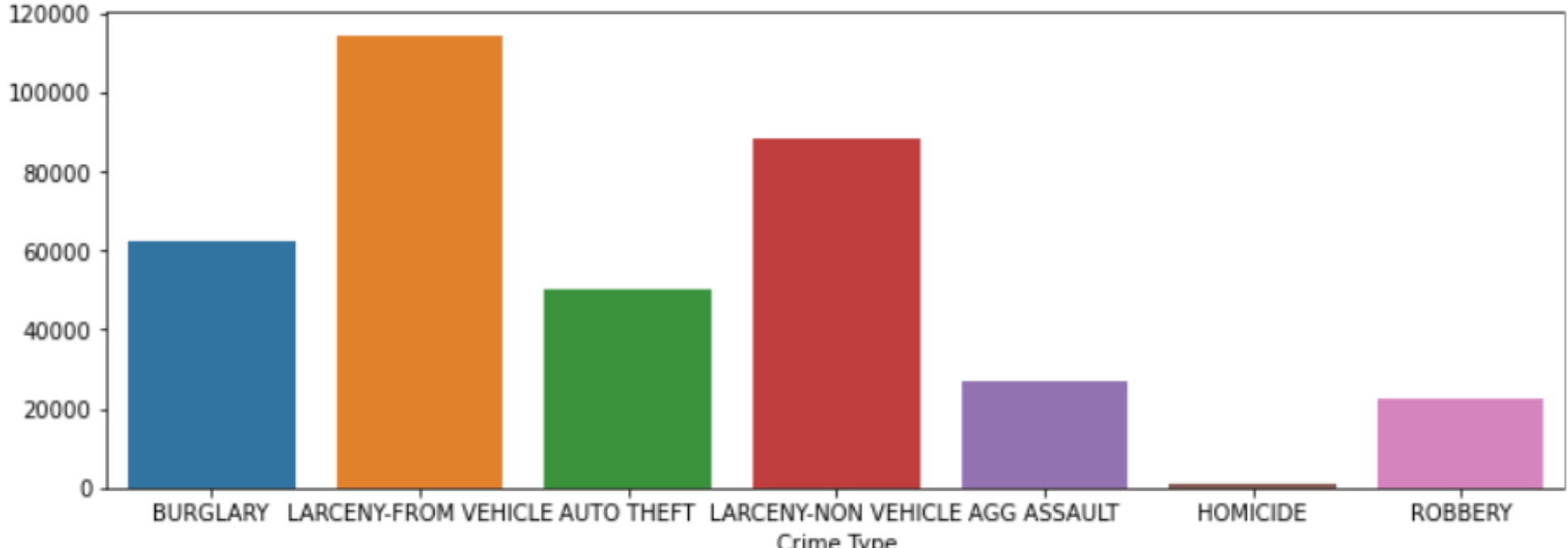
➤ Decreasing trend of the crimes with an added seasonality factor in the occurrence of the crimes



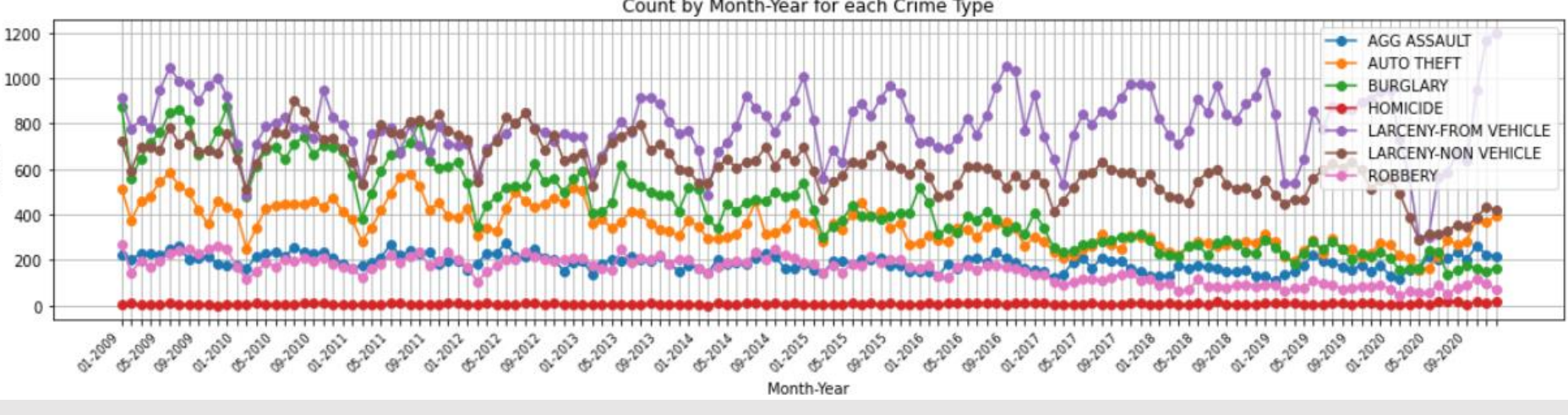
➤ Distribution of crimes for different days of the week looks even



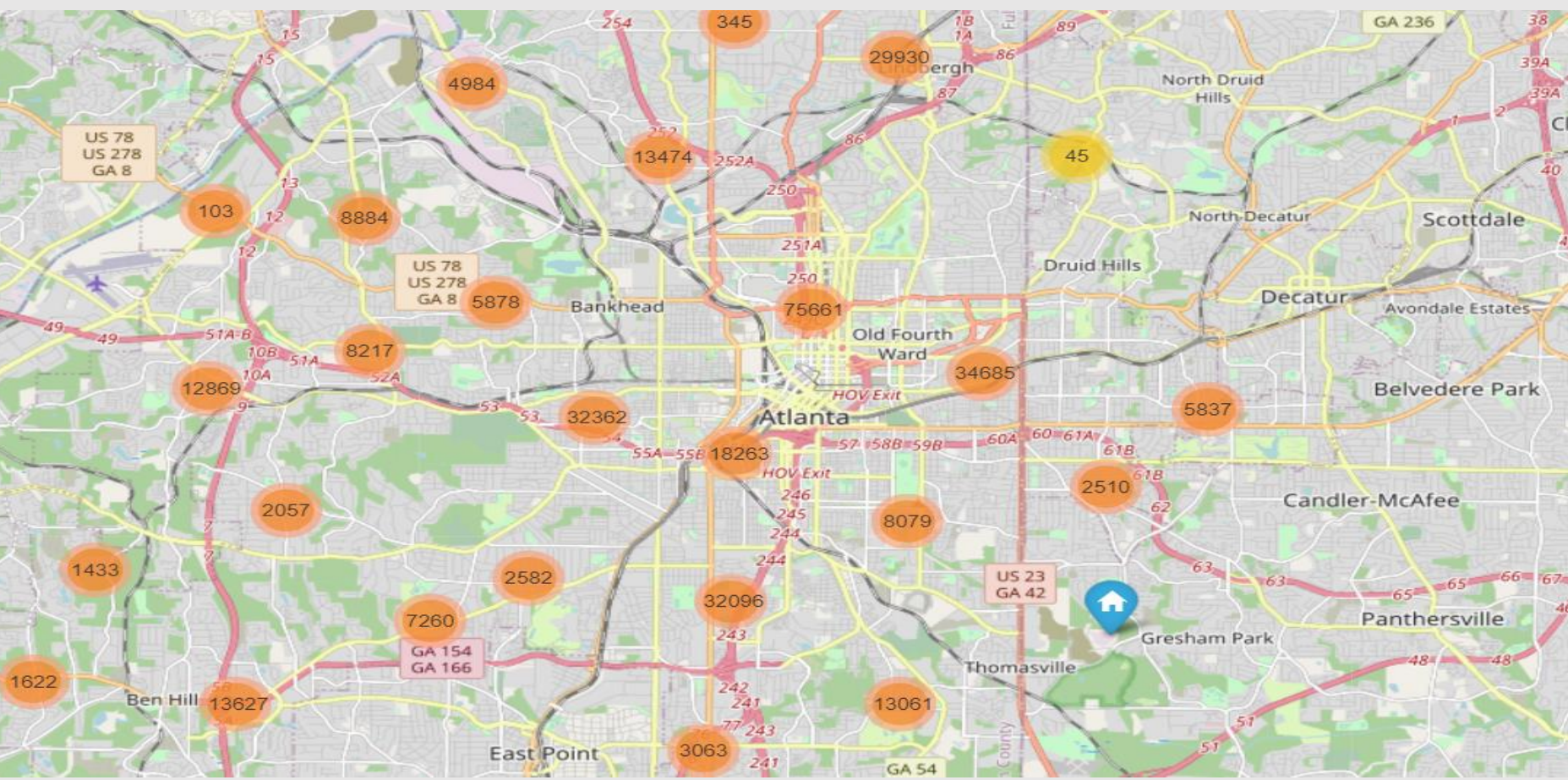
➤ Larceny (vehicular and non-vehicular), burglary, and auto theft are the three highest types of crime



➤ Trend across time for the above types of crime shows that crimes like larceny (non-vehicle) have seen a dip in recent years, particularly after COVID-19



➤ We also experimented with the folium library in Python to render an interactive map that has the count of crimes in each geographic area of Atlanta



Conclusion and Interpretation of Results

In conclusion, our analysis reveals a consistent pattern in crime distribution across Atlanta.

- Notably, neighborhoods near Five Points and the Hotel District consistently exhibit the highest risk scores or crime ratings, while areas such as Carey Park, Kirkwood, and East Lake Highlands, further from the city center, show the least.
- The central zones, including Home Park/Atlantic Station and Midtown, are identified as common crime hotspots, predominantly for vehicular larceny and auto theft. This suggests a targeted need for increased policing of these specific crime types within these regions. Additionally, an encouraging trend observed is the general reduction in crime ratings over the years, indicating the potential effectiveness of current crime prevention strategies.

What's New??

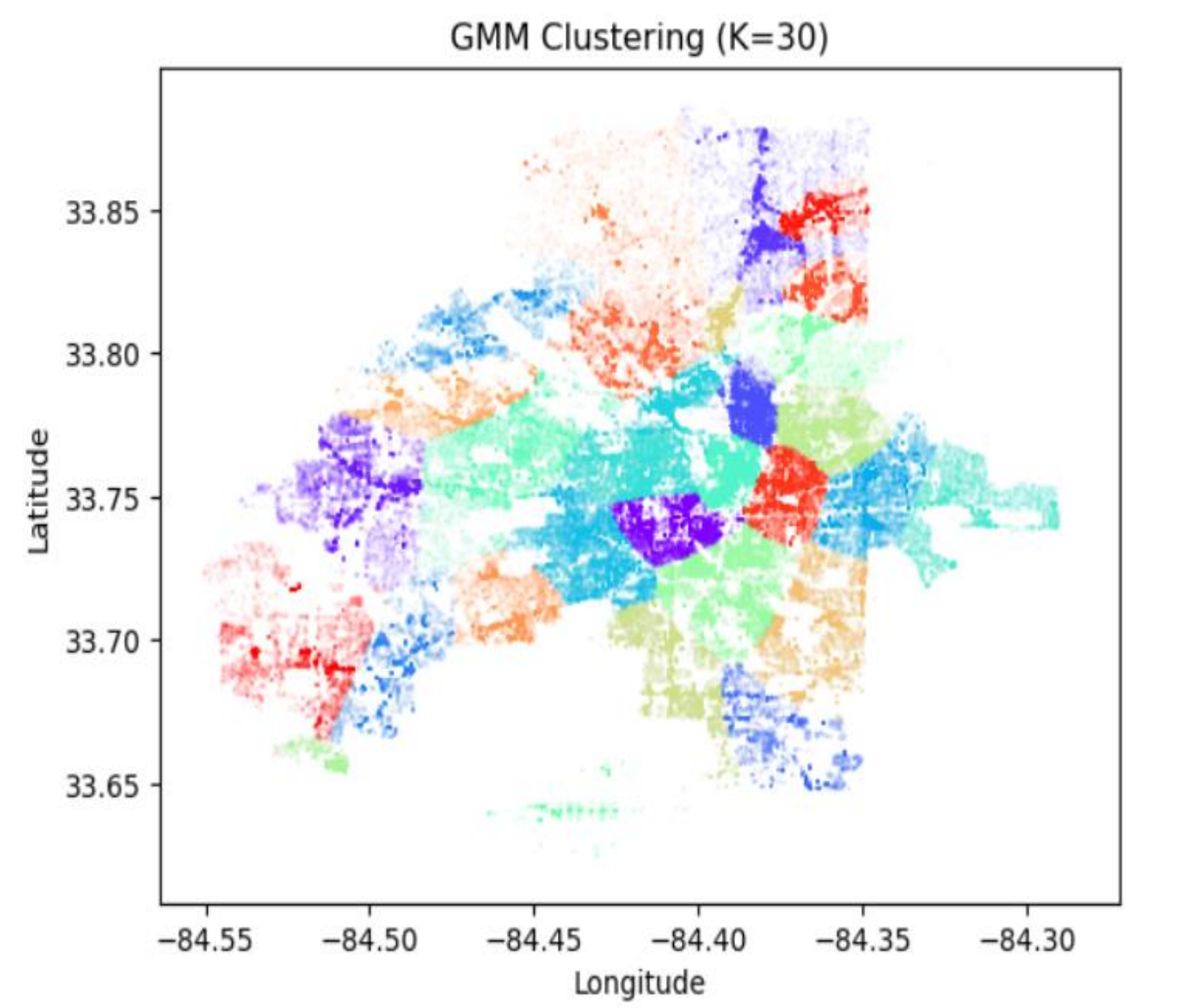
This project introduces a novel approach to crime prediction and visualization in urban areas, utilizing a combination of Gaussian Mixture Models (GMM) for clustering and time series modeling for risk score prediction. Some unique aspects are :

- The specific application of GMM to crime data, allowing for the accurate identification of varied crime hotspots, especially in complex urban landscapes like Atlanta
- A detailed focus on neighborhood-level crime variations, enabling a more granular understanding and response to urban crime dynamics.

2. Clustering Crimes by Location

➤ The clustering model, a Gaussian Mixture with a K value of 30, successfully distinguished the airport as a unique cluster due to its distinct crime distribution.

➤ For our crime prediction project, GMM proved superior to K-means due to its ability to handle clusters of varied shapes and densities, crucial for accurately mapping the complex and irregular distribution of crime data.



Our clustering approach effectively singled out the airport as its own entity due to its unique crime pattern, a decision substantiated by hierarchical analysis showing the airport's early divergence in a cosine similarity comparison with other hotspots.

4. Interactive Visualisation on Atlanta Map

The tool utilizes Dash and Plotly to create an interactive dashboard for crime data analysis with the following features

- A dropdown menu for selecting specific months and years to dynamically update the displayed data.
- An interactive map that visualizes crime clusters geographically, with points color-coded based on crime ratings and equipped with hover functionality for detailed cluster information.
- When a user hovers over a cluster on the map, a horizontal bar chart updates to show various crime types and their counts for the selected cluster and time period.
- The map employs a clear and readable carto-positron Mapbox style and offers zoom functionality for detailed exploration.
- Hovering over map clusters provides additional information on crime rates and predictions, enhancing the user's understanding of crime patterns.

