

CRIME ANALYSIS IN PHILADELPHIA

DSCI 521: DATA ANALYSIS AND INTERPRETATION

Term Project Phase 1: Scoping an analytics project

OVERVIEW

A. Background Report on Team Members and Contributions

- ✓ 1. *Anish Kania (ajk462@drexel.edu):*
- 2. *Aryan Jain (aj3246@drexel.edu):*
- 3. *Darshit Rai (dr3264@drexel.edu):*
- 4. *Manav Bhagat (mb4474@drexel.edu):*

Together, we aim to conduct a thorough and insightful data analysis, utilizing our collective strengths in programming, data visualization, and statistical analysis. Collaboration among team members will be key in leveraging our individual proficiencies to achieve a comprehensive and successful outcome for this proposed data analysis project..

B. Project scope

- ✓ The proposed project aims to conduct a comprehensive crime analysis in Philadelphia, utilizing a rich dataset obtained from data.phila.gov. The primary objective is to uncover patterns, trends, and correlations within the crime data, focusing on various factors such as the timing of crimes, weather conditions, unemployment rates, and geographic distribution. This analysis will be powered by an array of sophisticated data analysis and machine learning tools, including but not limited to pandas, NumPy, Seaborn, Folium, GeoPandas, scikit-learn, NLTK, and spaCy.

C. Project Motivations

- ✓ *The motivations driving this project stem from a collective desire to:*
 - *Understand the underlying factors contributing to crime in Philadelphia..*
 - *Identify potential hotspots and temporal patterns that could assist in crime prevention.*
 - *Explore the impact of socio-economic variables, such as unemployment rates, on crime frequencies.*
 - *Enhance public safety through data-driven insights and recommendations.*

D. Data Utilization

- ✓ *The project will leverage pre-processed data available from data.phila.gov, which is expected to include detailed crime records along with associated variables such as date, time, location, type of crime, weather conditions, and possibly socio-economic indicators. The preprocessing stage will ensure the data is clean, structured, and ready for in-depth analysis.*

The project will leverage pre-processed data available from data.phila.gov, which is expected to include detailed crime records along with associated variables such as date, time, location, type of crime, weather conditions, and possibly socio-economic indicators. The preprocessing stage will ensure the data is clean, structured, and ready for in-depth analysis.

E. Exploratory Analysis

- ✓ *An initial exploratory analysis will be conducted to:*
 - *Understand the distribution and frequency of various types of crimes.*
 - *Analyze crime trends over time and correlate them with weather patterns and unemployment rates.*
 - *Identify geographical areas with higher crime rates using spatial analysis tools like Folium and GeoPandas.*

This phase will involve extensive use of visualization libraries like Seaborn to create intuitive and informative plots that reveal underlying patterns and insights within the data.

F. Analytical Tools and Frameworks

- ✓ *The project will employ a wide range of analytical tools and frameworks, each chosen for its specific capabilities:*
 - **Pandas and NumPy:** For data manipulation and numerical analysis.
 - **Seaborn and Folium:** For advanced data visualization, including geographical mapping.
 - **GeoPandas:** For spatial data analysis, enhancing the geographical aspect of the crime data.
 - **Scikit-learn:** For applying machine learning models to predict crime rates and identify significant predictors.
 - **NLTK and spaCy:** For natural language processing, potentially to analyze textual data from crime reports or social media for additional insights.

G. Objectives

- ✓
 1. **Pattern Identification:** Uncover patterns related to the timing and location of crimes, and investigate the influence of weather and unemployment rates..
 2. **Hotspot Analysis:** Use geographical data analysis to identify and visualize crime hotspots.
 3. **Predictive Modeling:** Develop predictive models to forecast crime rates and understand the impact of various factors on crime
 4. **Recommendations:** Provide data-driven recommendations to local authorities and stakeholders for targeted interventions and policy-making.

H. Documentation and Adaptation

- ✓ *Throughout the project, meticulous documentation will be maintained to track objectives, methodologies, findings, and changes in strategy. This living document will serve as a roadmap and a historical record, accommodating shifts in focus or approach as new data, tools, or insights become available.*

I. Conclusion

- ✓ *This crime analysis project in Philadelphia is not just an academic exercise but a step towards making tangible impacts on community safety and well-being. By combining a comprehensive dataset with a suite of powerful analytical tools, the project aims to provide actionable insights that can guide policy and preventive measures in the city.*

```
data.columns
```

```
Index(['crime_type', 'point_x', 'point_y', 'Year', 'Month', 'Day', 'Hour',  
      'Minute', 'Weekday', 'IsWeekend', 'street_name', 'block_number',  
      'distance_to_nearest_police_station', 'temperature_2m_max (°F)',  
      'temperature_2m_min (°F)', 'temperature_2m_mean (°F)',  
      'precipitation_sum (mm)', 'precipitation_hours (h)',  
      'Unemployment Rate of a Population'],  
      dtype='object')
```

Fig : Data Columns





