Crime Data Analysis and Visualization

Final Project Proposal

Addis Yesserie(addisy1@umbc.edu)
MPS Data Science, UMBC
Data603 Platforms for Big Data Processing
Spring 2022

Introduction

"The rapid growth of cloud computing and data acquisition and storage technologies, from business and research institutions to governments and various organizations, have led to a huge number of unprecedented scopes/complexities from data that has been collected and made publicly available"

M. Huda, et. Al iJET: Insights into innovative environment for online learning resources," Int.J. Emerg. Technol. Learn., vol. 13, no. 1, pp. 2336, Jan. 2018. https://doi.org/10.3991/ijet.v13i01.6990

Introduction

- Big Data Analytics (BDA) can effectively address the challenges of data that are too vast, too unstructured, and too fast moving to be managed by traditional methods
- DBA can aid organizations to utilize their data and facilitate new opportunities

Introduction

- BDA has become an emerging approach for:
  - Analyzing data
  - Extracting information
  - Relations in a wide range of application areas
- BDA is a systematic approach for
  - analyzing and identifying
    - Patterns
    - Relations
    - Trends within a large volume of data

Objectives

- To analyze and visualize crime patterns of Prince George's County, Maryland

- The use of Big Data Analytics on this crime incident and pattern analysis will enable for hotspot detection and predictive policing

Motivation

- Pattern analysis uncovers the underlying interactive process between crime events by discovering
  - *Where*…location
  - *When*…Time
  - What and *Why* crimes are likely to occur
- Help tool development in the law enforcement industry and safety of the community

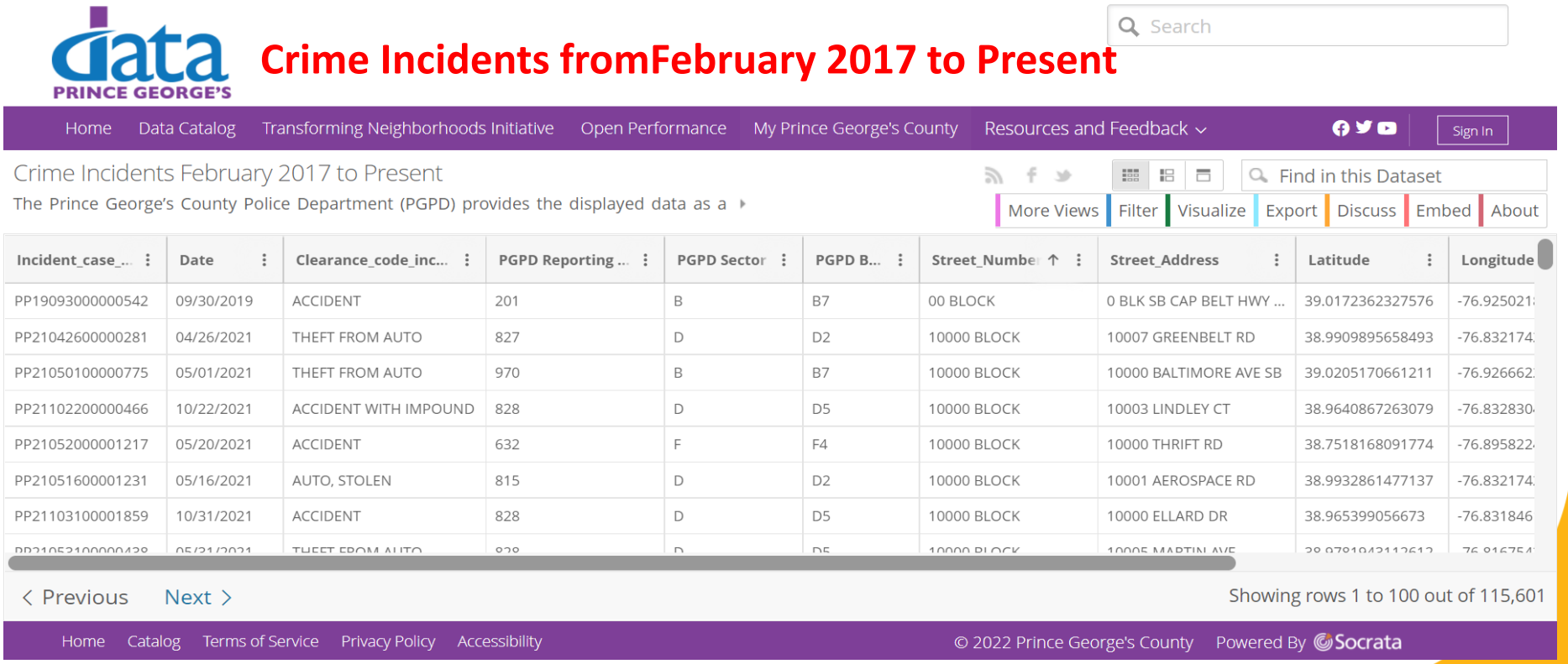
Motivation

- Using visualization techniques:
  - gain insights into an information space
  - provide qualitative overview
  - search for patterns, trends, structure, irregularities, relationships
  - help to find interesting regions and suitable parameters for further analysis

Datasets

- Publicly available datasets that consist of crime activities in Prince George's County, MD

https://data.princegeorgescountymd.gov/Public-Safety/Crime-Incidents-February-2017-to-Present/wb4e-w4nf/data



Methods

- After acquiring the data, big data processing python package including :
  - PySpark will be used to perform data reading, transforming, and querying and analysis
- Existing python libraries including:
  - Matplotlib will be used to visualize crime data