

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>



Crime Incidents from February 2017 to Present

Home Data Catalog Transforming Neighborhoods Initiative Open Performance My Prince George's County Resources and Feedback ▾



Sign In

Crime Incidents February 2017 to Present

The Prince George's County Police Department (PGPD) provides the displayed data as a ▸



More Views Filter Visualize Export Discuss Embed About

Incident_case_...	Date	Clearance_code_inc...	PGPD Reporting ...	PGPD Sector	PGPD B...	Street_Numbe ↑	Street_Address	Latitude	Longitude
PP19093000000542	09/30/2019	ACCIDENT	201	B	B7	00 BLOCK	0 BLK SB CAP BELT HWY ...	39.0172362327576	-76.9250212
PP21042600000281	04/26/2021	THEFT FROM AUTO	827	D	D2	10000 BLOCK	10007 GREENBELT RD	38.9909895658493	-76.832174
PP21050100000775	05/01/2021	THEFT FROM AUTO	970	B	B7	10000 BLOCK	10000 BALTIMORE AVE SB	39.0205170661211	-76.926662
PP21102200000466	10/22/2021	ACCIDENT WITH IMPOUND	828	D	D5	10000 BLOCK	10003 LINDLEY CT	38.9640867263079	-76.832830
PP21052000001217	05/20/2021	ACCIDENT	632	F	F4	10000 BLOCK	10000 THRIFT RD	38.7518168091774	-76.895822
PP21051600001231	05/16/2021	AUTO, STOLEN	815	D	D2	10000 BLOCK	10001 AEROSPACE RD	38.9932861477137	-76.832174
PP21103100001859	10/31/2021	ACCIDENT	828	D	D5	10000 BLOCK	10000 ELLARD DR	38.965399056673	-76.831846
PP21052100000428	05/21/2021	THEFT FROM AUTO	828	D	D5	10000 BLOCK	10005 MARTIN AVE	38.9781842112612	-76.816754

< Previous Next >

Showing rows 1 to 100 out of 115,601

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