Exploring the Raleigh Police Department's NIBRS data (2014-2021)

Siddhartha Vanam
NC State Datathon
03.21.21

Agenda

1. Background
2. Executive Summary
3. Analysis
    a. Exploratory Data Analysis (EDA)
    b. Key Results
4. Recommendations for Improvement
5. Foreword: Machine Learning Approach

1. Background (½)

FBI Implementation: Suggested a shift in Reporting Systems.

- **Past**: <u>Summary Reporting System</u> (**SRS**) = An *aggregate monthly tally* **of crimes.**

- **Current**: <u>National Incident-Based Reporting System</u> (**NIBRS**) = *Detailed* context to incidents such as location and time of day.

● FBI's *Uniform Crime Reporting* (**UCR**) has officially **retired the SRS (01-01-2021)** and is **encouraging** and assisting agencies **to make the switch to NIBRS** .

1. Background (2/2)

Cost of Violent Crime is received by Health Care Systems for :

- Security
- Safety
- Prevention
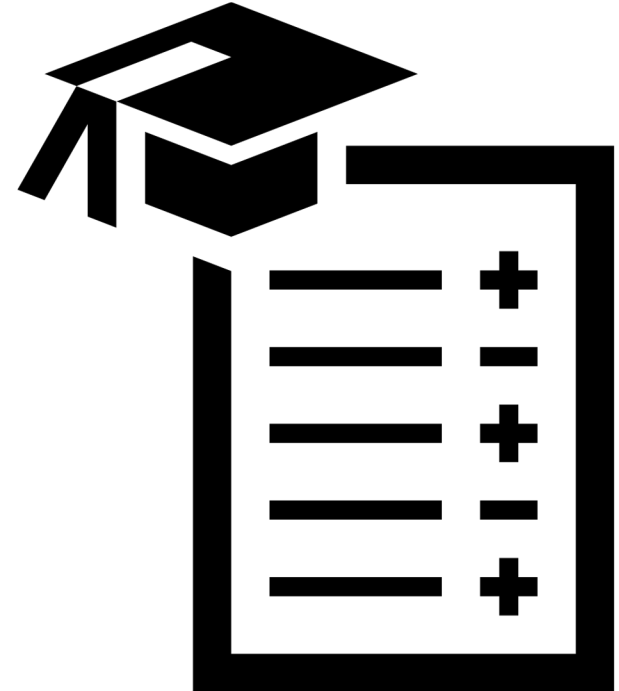
$9M
Per homicide

$241K
Per rape or sexual assualt

$107K
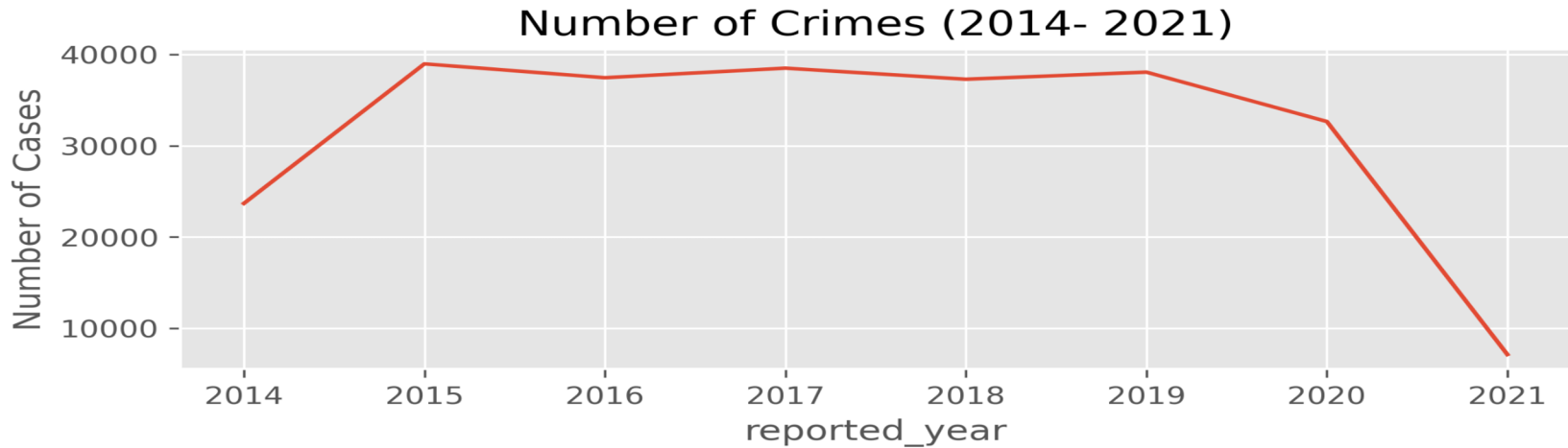*Per aggravated assault*

$42K
Per robbery

2. Executive Summary

- **+ Opportunity for Analysis**
  - + Seasonal Trends
- **- Inconsistent Reporting**
  - - Datetime transformations (Days)
- **- Organization/ Clarity**
  - - Cases can violate >1 Crime Code
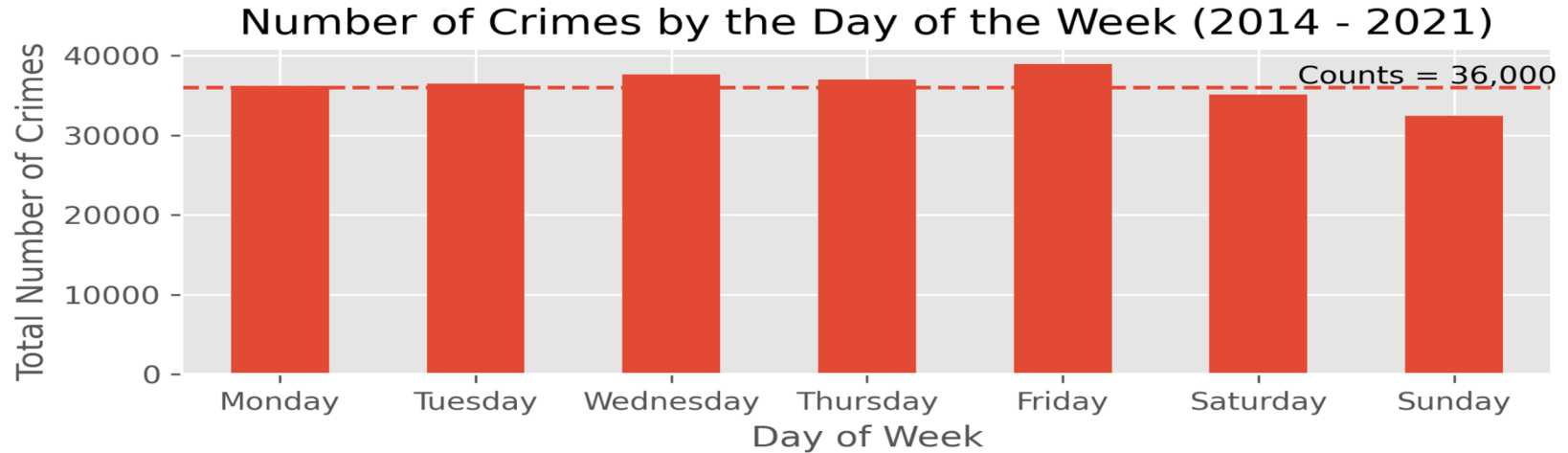 - - District based reporting
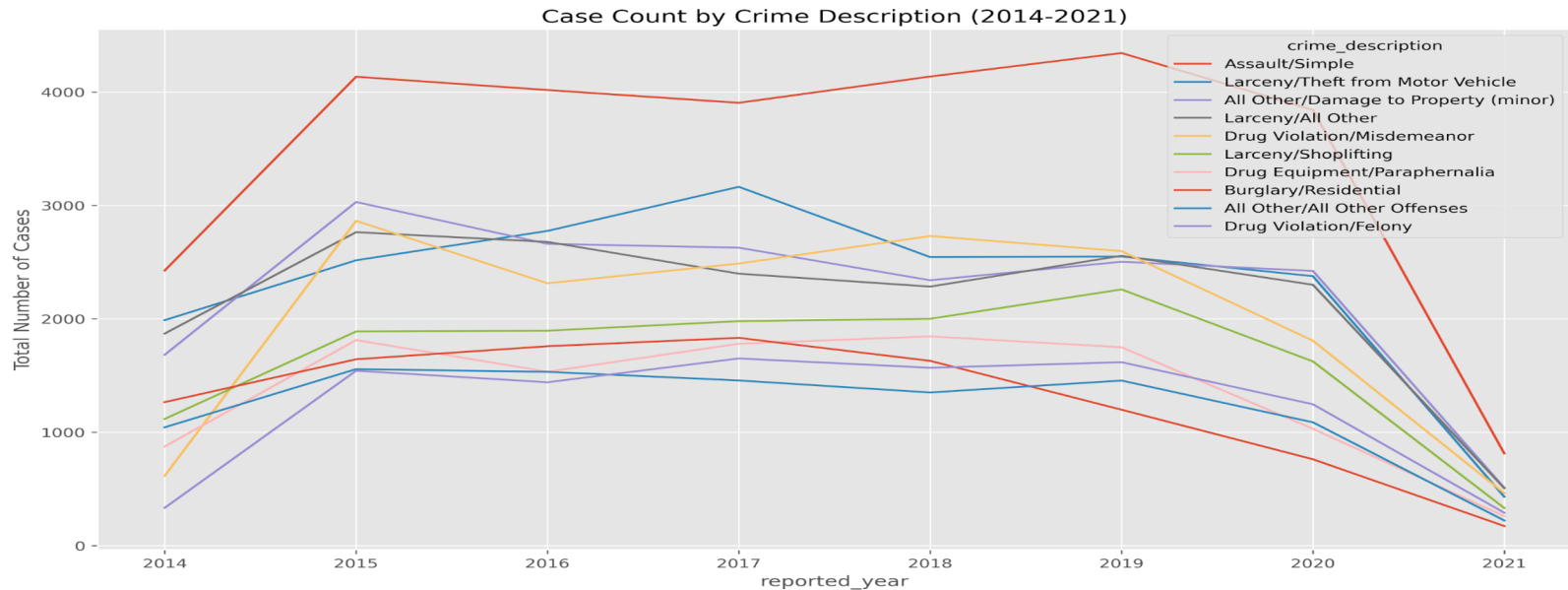
EDA – Exploratory Data Analysis

Visualization EDA 1

Source: https://github.com/siddley1001/NCState_Datathon21/blob/master/state_datathon_police.ipynb

EDA – Exploratory Data Analysis

Visualization EDA 1

Number of Crimes by the Day of the Week (2014 - 2021)

EDA – Exploratory Data Analysis

Visualization EDA 1

Case Count by Crime Description (2014-2021)

EDA – Exploratory Data Analysis

Visualization EDA 1

Top 10 Crimes by Description (2004 - 2021)
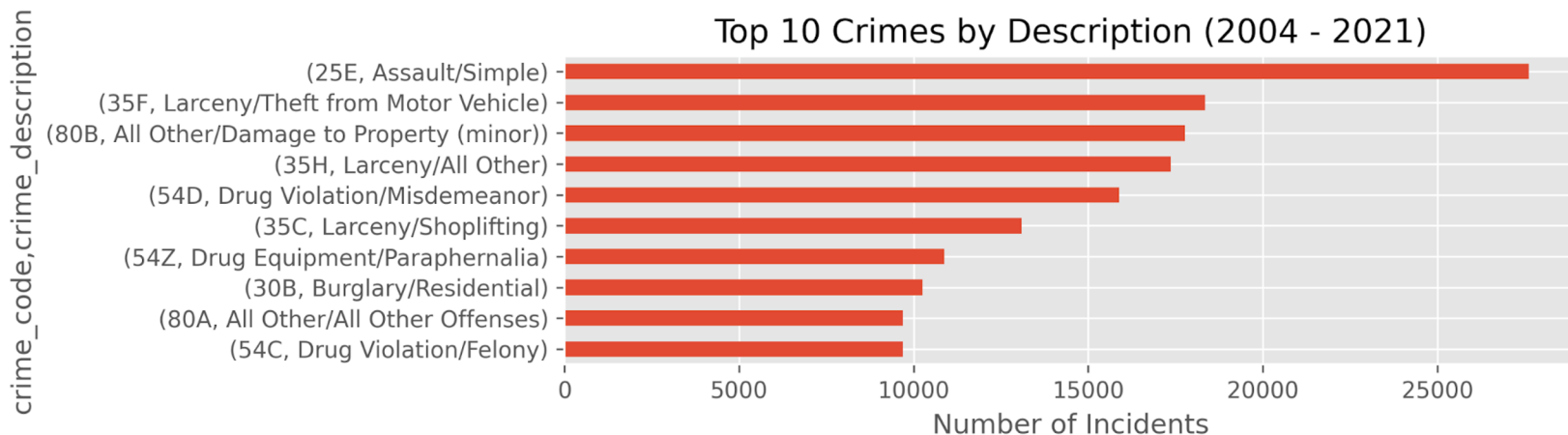
EDA – Exploratory Data Analysis

Visualization EDA 1

Distribution of Crime according to Region (2014- 2021)

North 20.6
Southeast 22.2
UNK 0.1
Northwest 10.6
Downtown 14.6
Southwest 14.7
Northeast 17.2

Visualization EDA 1

Visualization EDA

EDA – Exploratory Data Analysis

Visualization EDA 1

Top 10 Reported Block Cases for Crimes (2014- 2021)

EDA – Exploratory Data Analysis

Visualization EDA 1

Block Cases By Region

NE	3
Downtown	3
SE	2
NW	1
North	1

Visualization EDA 1

The Raleigh Police Department has six districts located throughout the city. Each district serves a specific geographic area: Northwest, North, Northeast, Southeast, Downtown, and Southwest.

District Head Quarters   Top 10 Block Cases
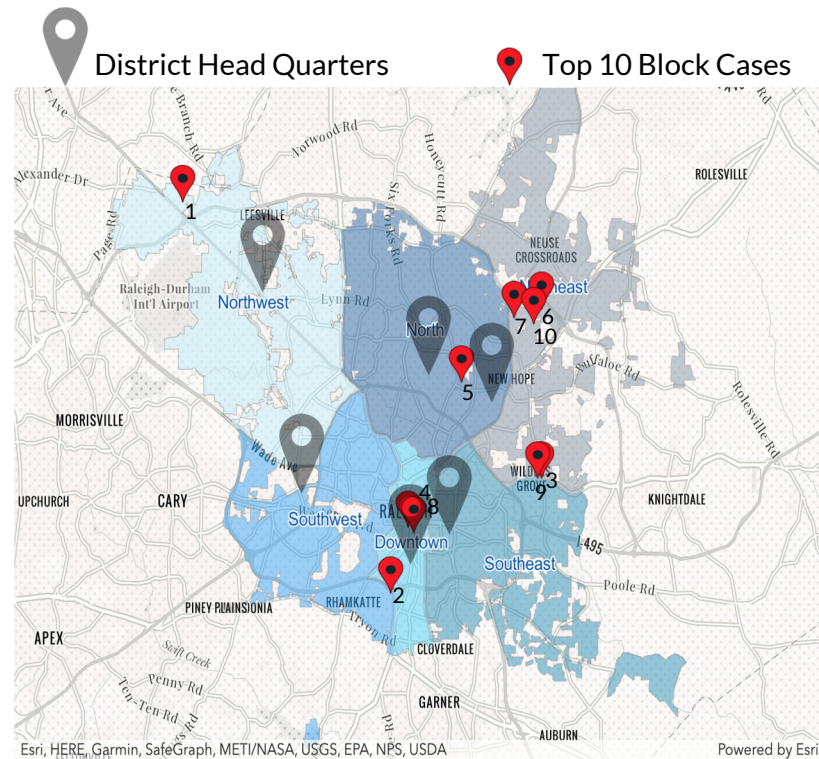
4. Recommendations

Comprehensive Prevention

- Using the Data to **predict future crimes**.
- Coordinate between Regional branches to **reach the site ASAP.**

Data Integrity

- Ensure **fields are consistent**
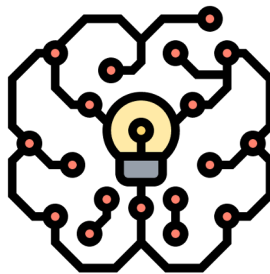- Keep a **data dictionary** to understand different features.

Public Transparency

- Keep the **public informed** with the status of cases.
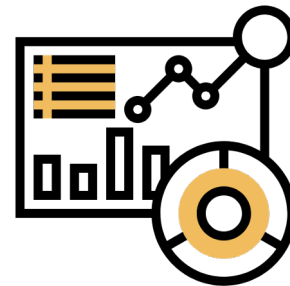- Market the available Data to keep the **public engaged** with **prevention plans** and **hotspots**.

Foreword: Machine Learning Approach

DateTime Components

Longitude and Latitude Interactions

Travel Time
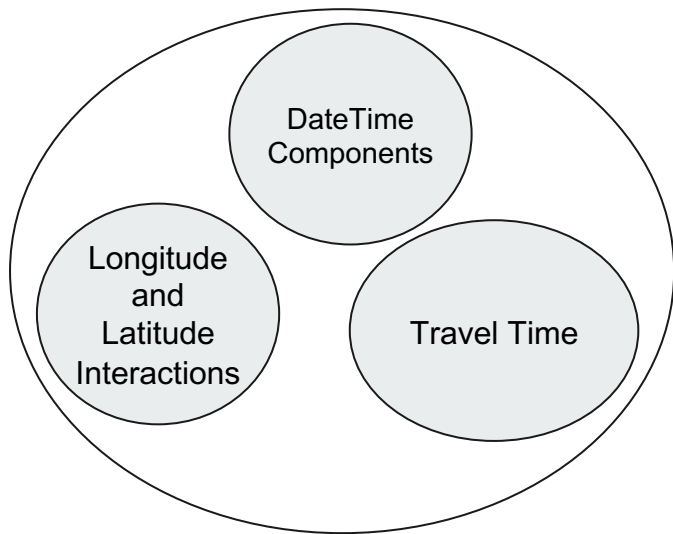
100k+ Total Features

Classification Model
(XGBoost)

GOAL: Homicide Prevention with high Accuracy and Recall.