**IST652 SCRIPTING FOR DATA ANALYTICS**

**University Of Syracuse**

**Exeario Dario Boscan**

**11 December 2021**

Introduction

The purpose of this project is to conduct a data analysis with Python, an interpreted, object-oriented, high-level programming language with dynamic semantics. The analysis was done on the National Incident-Based Reported System (NIBRS). The NIBRS captures details on each single crime incident as well as on separate offenses within the same incident, to include information on victims, known offenders, relationships between victims and offenders, arrestees, and property involved in crime. Specifically, I researched crime data for North Carolina derived from the Federal Bureau of Investigations (FBI) Crime Data Explorer reports voluntarily submitted to the FBI.

Data and Data Sources

In 2020, there were 29, 908 violent-crime incidents, and 39,880 offenses reported in North Carolina by 377 law enforcement agencies that submitted National Incident-Based Reporting System (NIBRS) data and covers 90% of the total population. The data is divided in the following subjects;

1. All Violent Crime Offender vs. Victim Demographics (Offender Age, Victim Age).
2. All Violent Crime Victim Demographics (Location Type, Victim’s Relationship to the Offender).
3. All Violent Crime Offense Characteristics (Type of Weapon Involved by Offense, Offense linked to another offense).
4. Homicide Offender vs. Victim Demographics (Offender Age, Victim Age).
5. Homicide Victim Demographics (Location Type, Victim’s Relationship to the Offender).
6. Homicide Offense Characteristics (Type of Weapon Involved by Offense, Offense Linked to Another Offense).
7. <https://crime-data-explorer.app.cloud.gov/pages/home> Main Federal Bureau of Investigations Crime Data Explorer page.
8. <https://crime-data-explorer.fr.cloud.gov/pages/explorer/crime/crime-trend> Main page to gather the crime data for North Carolina derived from the both the National Incident-Based Reporting System (NIBRS) and Summary Reporting System (SRS) voluntarily submitted to the FBI.

Preprocessing

I extracted and downloaded .csv files for each of the data subjects stated above. Subsequently, I proceeded to upload those files on Jupyter Notebook for processing with Python code and Pandas.

Method of Analysis

|  |  |  |
| --- | --- | --- |
| ANALYSIS METHOD 1 | DATA INPUT | DATA OUTPUT |
| Imported Pandas as pd, created a datapath to equal the csv file. Created a function to read the file to extract the shape of the function along with the dataframe. | CSV file | Data Shape and Data Frame |

|  |
| --- |
| ANALYSYS METHOD 1 CODE (EXAMPLE) |
| import pandas as pd  import numpy as np  datapath = 'Rate-of-Violent-Crime-Offenses-by-Population12\_10\_2021.csv'  rateofviolence\_data = pd.read\_csv(datapath, sep=',')  rateofviolence\_data.shape |

|  |  |  |
| --- | --- | --- |
| ANALYSIS METHOD 2 | DATA INPUT | DATA OUTPUT |
| Utilized Python code to show visualization of the csv files on each analyzed category | Dataframe and X,Y data. | Bar chart |

|  |
| --- |
| ANALYSIS METHOD 2 CODE (EXAMPLE) |
| NorthCarolina = [362.7, 346.3, 353.5, 341.0, 329.1, 346.1, 371.8, 370.4, 356.2, 378.7, 419.3]  UnitedStates = [404.5, 387.1, 387.8, 369.1, 361.6, 373.7, 397.5, 394.9, 383.4, 380.8, 398.5]  index = ['2010', '2011', '2012', '2013', '2014', '2015', '2016', '2017', '2018', '2019', '2020']  df = pd.DataFrame({'NorthCarolina': NorthCarolina,  'UnitedStates': UnitedStates}, index=index)  ax = df.plot.bar(rot=0) |

Questions to be Answered

1. What is the average offenders’ age?
2. What is the average victims age?
3. What was the most predominant location for crimes committed during 2020?
4. What was the victim’s relationship to the offender in homicide crimes?
5. What was the most predominant crime committed in North Carolina during 2020?

The following data fields were used during my analysis;

|  |  |  |
| --- | --- | --- |
| Fields | Description | Example |
| Crime Offender Age | Age of individual | 21, 30, 40 |
| Victim Age | Age of individual | 21, 30, 40 |
| Location Type | Residence, Gas Station, etc. | Circle K gas station |
| Victim’s Relationship to the Offender | Stranger, Boyfriend/Girlfriend | Stranger, Boyfriend/Girlfriend |
| Type of Weapon Involved by Offense | Handgun, Knife, Firearm, etc. | N/A |
| Offense linked to another offense | Kidnapping, Burglary, etc. | N/A |

Output

- Rate of Violence Population Barchart. Rate per 100,000 people by year.

Chart, bar chart

Description automatically generated

- Offender Age (All Violent Crime) / 11 Variables / 31480 individuals

Chart

Description automatically generated

- Victim Age (All Violent Crime) / 11 Variables / 40,939 individuals

Chart

Description automatically generated

- Location Type (All Violent Crime) / 46 Variables / 45,516 locations

Histogram

Description automatically generated with low confidence

- Victim Relationship to Offender (All Violent Crime) / 26 Variables / 41,777 Relationships

Chart, histogram

Description automatically generated

- Offense Linked to Another Offense (All Violent Crime) / 50 Variables / 9,478 Links

Chart, histogram

Description automatically generated

- Type of Weapon (All Violent Crime) / 28 Variables / 29,796 Weapons

Chart, histogram

Description automatically generated

- Rate of Homicide by Population / Rate of 100,000 people, by year / 683 Homicides

Chart, bar chart

Description automatically generated

- Offender Age (Homicide) / 11 Variables / 832 Offenders

Chart

Description automatically generated

- Victim Age (Homicide) / 11 Variables / 742 Victims

Chart

Description automatically generated

- Location Type (Homicide) / 46 Variables / 742 Locations

A picture containing shape

Description automatically generated

- Victim Relationship to Offender (Homicide) / 26 Variables / 884 Relationships

Chart, histogram

Description automatically generated

- Type of Weapon (Homicide) / 28 Variables / 686 Weapons

Histogram

Description automatically generated

- Offense Linked to Another Offense (Homicide) / 50 Variables / 251 Links

Chart, histogram

Description automatically generated

Description of the Program

Graphical user interface, text, application

Description automatically generated

PLATFORM

Map

Description automatically generated with medium confidence

PLATFORM

Diagram

Description automatically generated

EXTRACTIONN

Table, whiteboard

Description automatically generated with medium confidence

ANALYSIS

Logo, company name

Description automatically generated

Conclusions

According to the National Incident-Based Reporting System (NIBRS) in North Carolina, in 2020, there were 29,908 violent crimes incidents, and 39,880 offenses reported by 377 law enforcement agencies and which covered 90% of the population.

I analyzed All Violent Crimes and Homicides during 2020 in North Carolina in order to answer the questions I stated in my final project proposal.

1. What is the average offenders’ age?

Due to the nature of the data, I could not establish the average offenders’ age, but the ages from 20 to 29 is the predominant range with 9,470 offenders of all violent crimes. With regards to Homicides, the ages from 20 to 29, is the predominant range with 344 offenders.

1. What is the average victims age?

Due to the nature of the analyzed data, I could not establish the average victims’ age, but the ages from 20 to 29 is the predominant range with 11,727 victims in all violent crimes. With regards to Homicides, the ages from 20 to 29, is the predominant range with 263 victims.

1. What was the most predominant location for crimes committed during 2020?

The most predominant location for all violent crimes was Residence Home with 23,207, and for homicides was Residence Homes with 387 instances.

1. What was the victim’s relationship to the offender in homicide crimes?

The most predominant relationship between victim and offender in all violent crimes is unknown with 14,599. For homicides, the most predominant relationship was also unknown with 437.

1. What was the most predominant crime committed in North Carolina during 2020?

The most predominant crime committed in North Carolinas was aggravated assault with 20,154.