

SYRACUSE

IST-652 FINAL PROJECT

JUSTIN PATE

SUID: 445414259

jd pate@syr.edu

PROJECT SUBJECT: CRIMES IN MAJOR CITY ANALYSIS

DATASET:

Introduction and Overview

The purpose of this analysis is to explore the open crime data source of a select group of major cities based on open crime data for 2018.

1. Describe the data and its source(s), including any preprocessing

- The open crime database came from <https://osf.io/zyaqn/>
- This is a publicly posted open database that includes crime statistics. Wiki does a good job at describing the dataset in the following Link <https://osf.io/zyaqn/wiki/home/>
- Pre-Processing?
 - The datasource is relatively clean with a few fields that are someone duplicates that you would not want to consider in your analysis. Offense_code is directly related to offense type. For charting, it is necessary to sample as the datasource is extremely large. The data source by default is in a *.gz file so will need to be opened using compression = "gzip".

city_name	offense_code	offense_type	offense_group	offense_against	date_single	longitude	latitude	location_type	location_category	census_block	date_
Austin	22U	other burglary/breaking & entering	burglary/breaking & entering	property	2018-01-01 00:00	-97.710191	30.349433	other	other	484530018043000	
Austin	520	weapon law violations	weapon law violations	society	2018-01-01 00:00	-97.741558	30.411489	residence	residence	484530017541002	
Austin	23H	all other larceny	larceny/theft offenses	property	2018-01-01 00:00	-97.741119	30.305911	vehicle parking	open space	48453002031007	
Austin	290	destruction/damage/vandalism of property (exce...	destruction/damage/vandalism of property (exce...	property	2018-01-01 00:00	-97.672452	30.363035	residence	residence	484530018333009	
Austin	290	destruction/damage/vandalism of property (exce...	destruction/damage/vandalism of property (exce...	property	2018-01-01 00:00	-97.699980	30.258932	other	other	484530009021000	
...
Virginia Beach	35A	drug/narcotic violations	drug/narcotic offenses	society	2018-12-31 23:10	-76.065879	36.759361	NaN	NaN	518100454221000	
Virginia Beach	520	weapon law violations	weapon law violations	society	2018-12-31 23:10	-76.065879	36.759361	NaN	NaN	518100454221000	
Virginia Beach	35A	drug/narcotic violations	drug/narcotic offenses	society	2018-12-31 23:23	-76.159443	36.830235	NaN	NaN	518100460051035	
Virginia Beach	12U	other robbery	robbery	property	2018-12-31 23:30	-75.971098	36.833776	NaN	NaN	518100440031000	
Virginia Beach	23F	theft from motor vehicle (except theft of moto...	larceny/theft offenses	property	2018-12-31 23:59	-76.089190	36.800369	NaN	NaN	518100454053001	



2. Describe your methods of analysis, including the questions that will be answered, in what fields the data will be used, and what the resulting output will be

- i) Using python functions in google Collaboratory the following questions will be answered.
 - (1) Can the type of crime and density be identified by city so that law enforcement can alter a plan for prevention and detection?
 - (2) Can the latitude and longitude values be used as a tool for home buying and law enforcement concentration as well?

3. Include an overall description of the program

4. If your project is a group project, describe the tasks and roles of each member of the group (2 points)

Using variations of Visual Data representation and aggregates, conclusions can be made regarding city and crime type and density. The first primary tool used is a mapping tool where a function is created in the program that controls the initial map zoom and the re-centering of the map based on the data to be displayed. The second primary tool used is plotly.express which is a phenomenal interactive visualization tool that can be imported into python.

VISUALIZATIONS FOR ANALYSIS:

FIGURE 1: SHOWS RELATIONSHIPS THROUGH CITY, AND CRIME TYPE VALUES OVER DATA

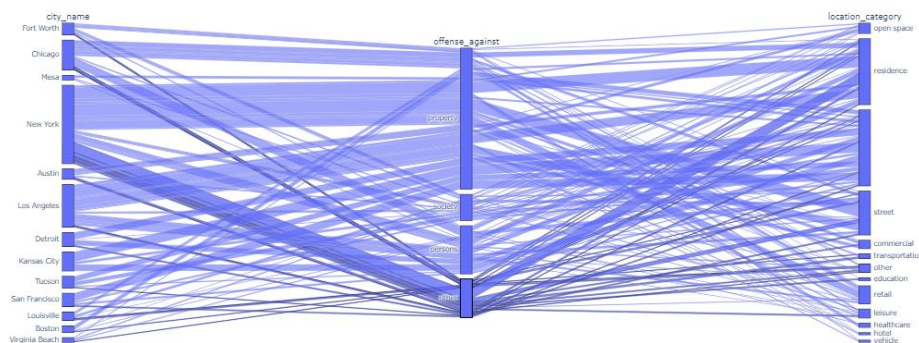


Figure 4:

Barchart by crime count by city

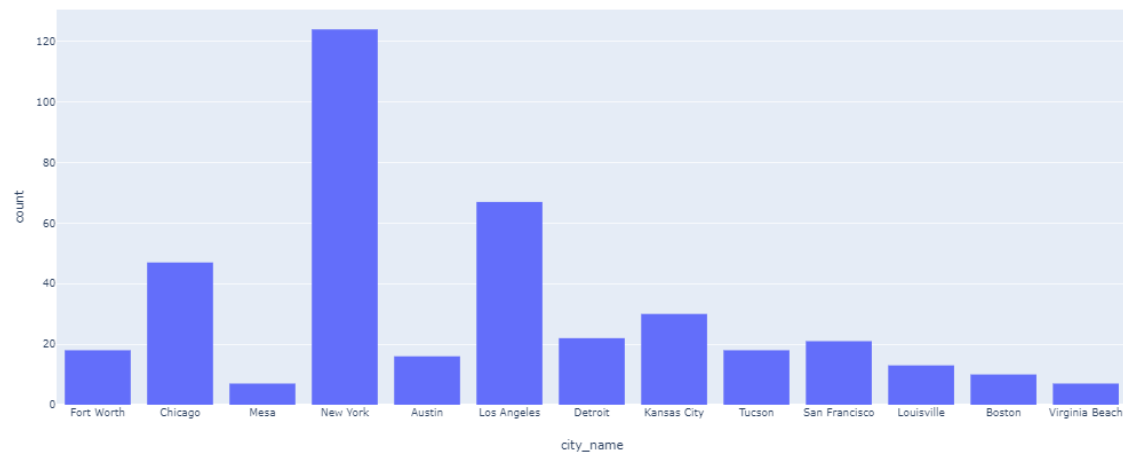
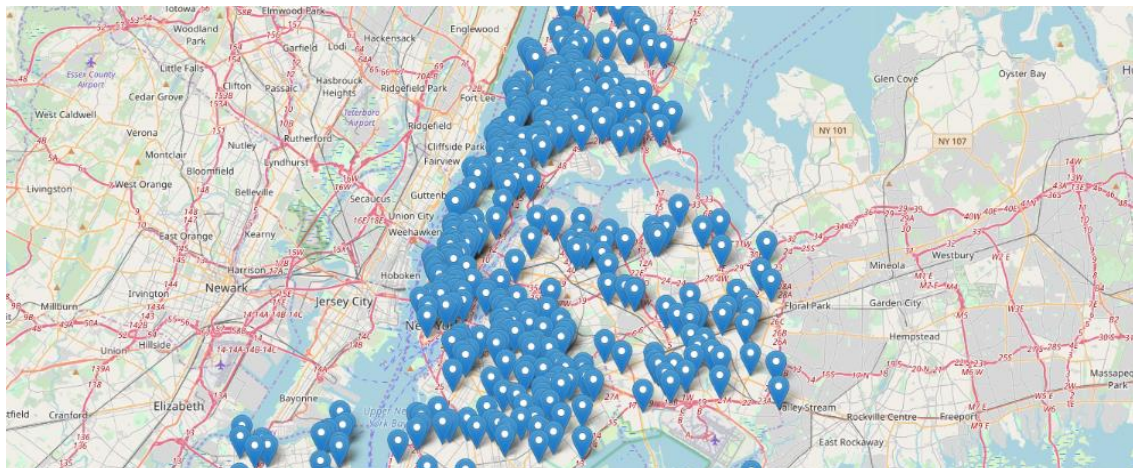
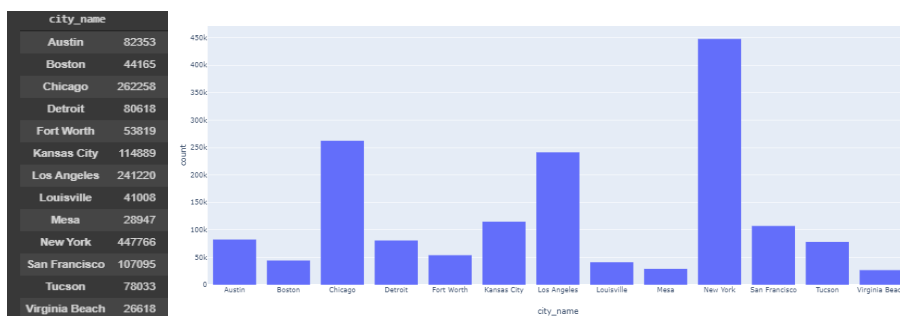


Figure 5: Sample crime map of city

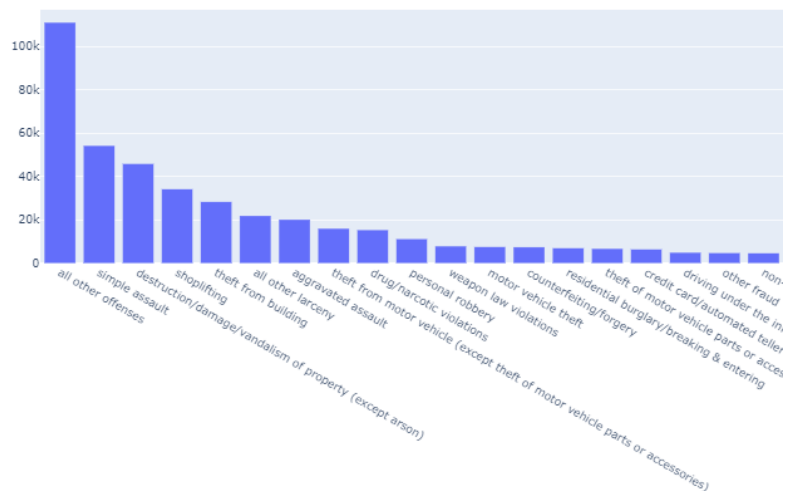


5. Draw conclusions from your results about your data

Some simple aggregation and visual analysis below allow some very confident and quick conclusions. New York has the highest number of reported crimes.



Break down the crime in only New York.



The largest number of identified crimes are simple assault at 54,219 reported.

If I still wanted to live in New York, I could use the mapping tool to find a location with less crime density.

Law Enforcement could use the same tool in planning officer daily routes.

