Crime in Montgomery county

Dawit H. Hailu

dhailu@montgomerycollege.edu

Immersive Data Analytics Bootcamp April, 2020

Outline

The problem

- I want to have a look at crimes in Montgomery county
- This information is believed to help fight crime and can be used as input for site selection (residence, bussiness)

aim

Let us achieve our aim by making use of our newly acquired data science tools

Data Collection

- Data is collected from Montgomery County open data website:[1]
- Then wrangle, clean it, and read it into a *pandas* dataframe before making use of it
- Consult for the website for data description (meaning of column names)

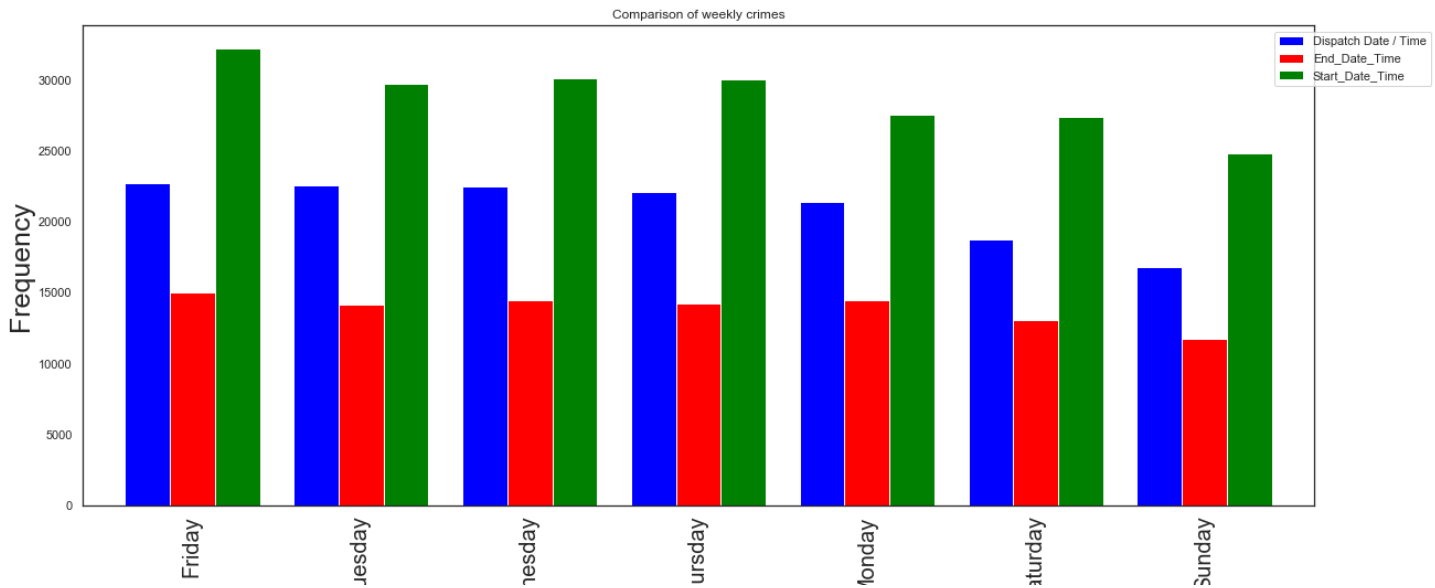
[1] (https://data.montgomerycountymd.gov/)

Snippet of the dataset

- Dataframe has 202094 rows and 30 columns
- redundant information about location is observed, so I will drop that
- I will focus on columns with information regarding crime time rows and crime district
- Information about time is encoded as dispatch time , start time and end times
- There are missing entries on the dataset

Crime day

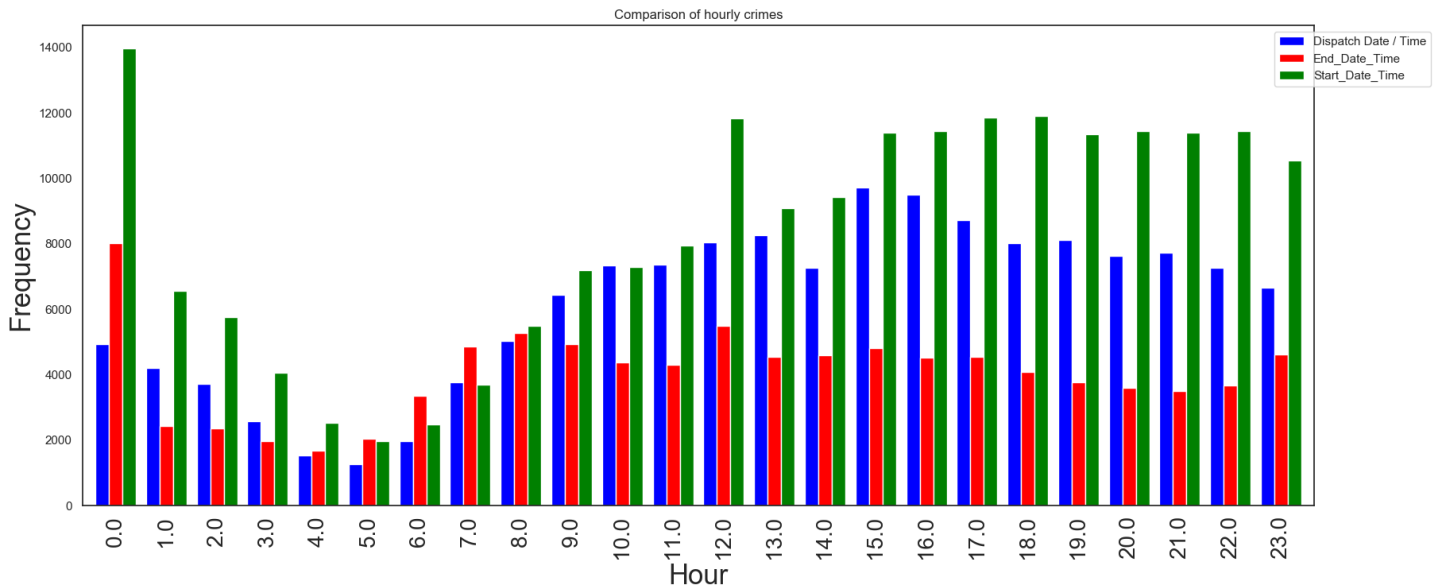
Based on Dispatch, End, and Start time data



Highest crime day **Friday**, decreases on *Saturday-Sunday*

Crime hour

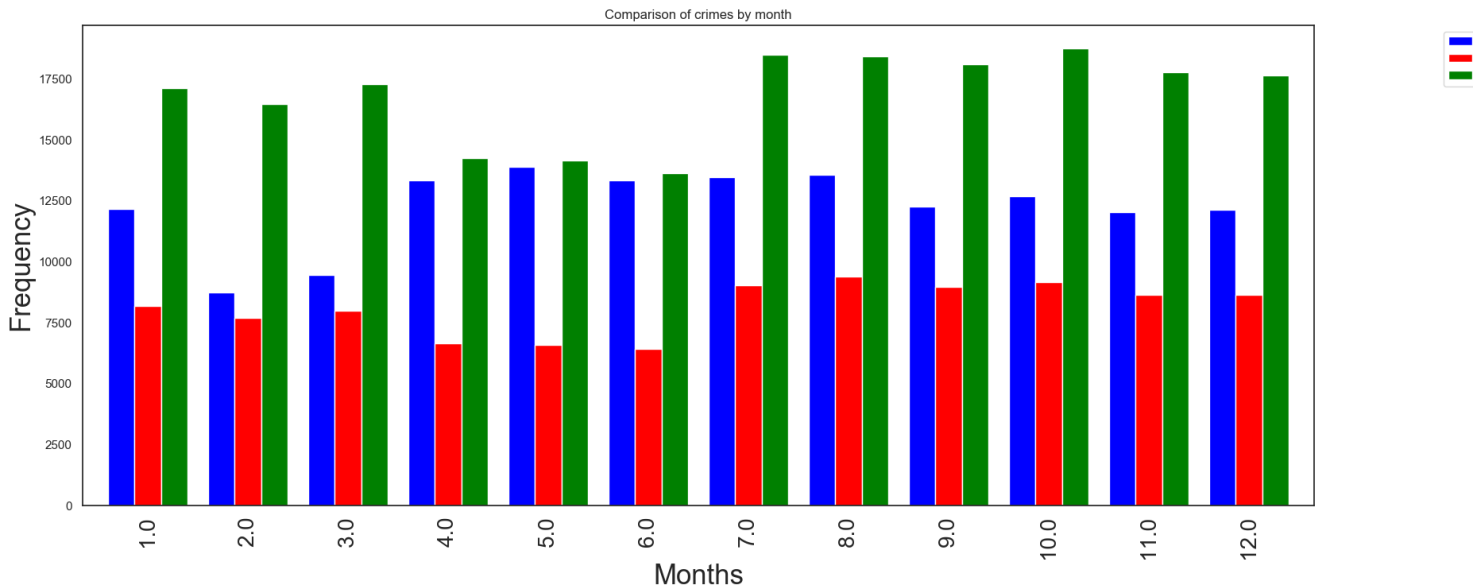
- Which hours do high crime incidence register?
- Similar to previous slide we use data from columns Dispatch, End. and Start time data



Midnight outlier for End, and Start time data ?
Morning seems to have lesser crimes

Crime month

- Again we use data from columns Dispatch, End, and Start time data



April-June for Start & End time approach, **February-March** for Dispatch time approach,

Crimes by district

What we have

We can use information about location, i.e Police District Name, Block Address, Zip Code, Sector, Beat, Latitude, Longitude, Police District Number, Location, Address Number

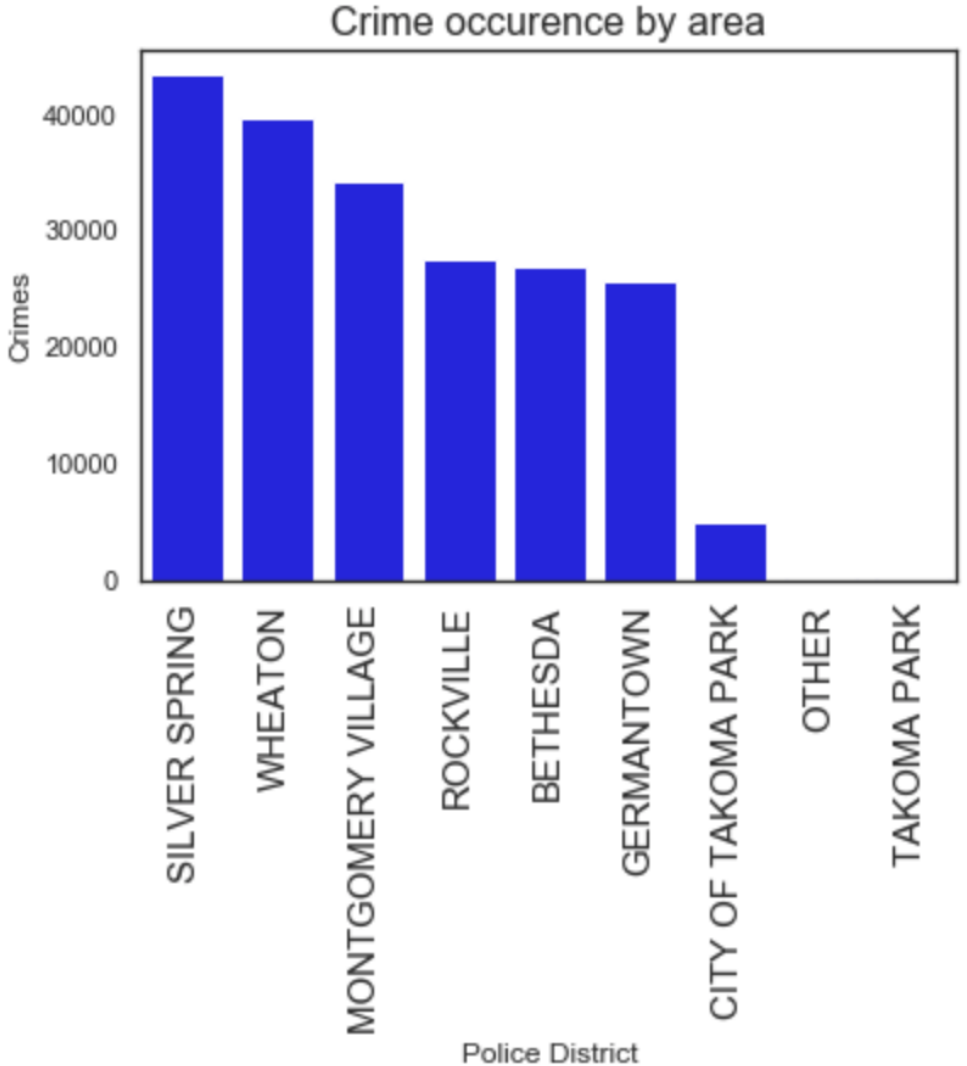
for our purpose

We chose to use Police District Name column as it has full information (i.e no missing values)

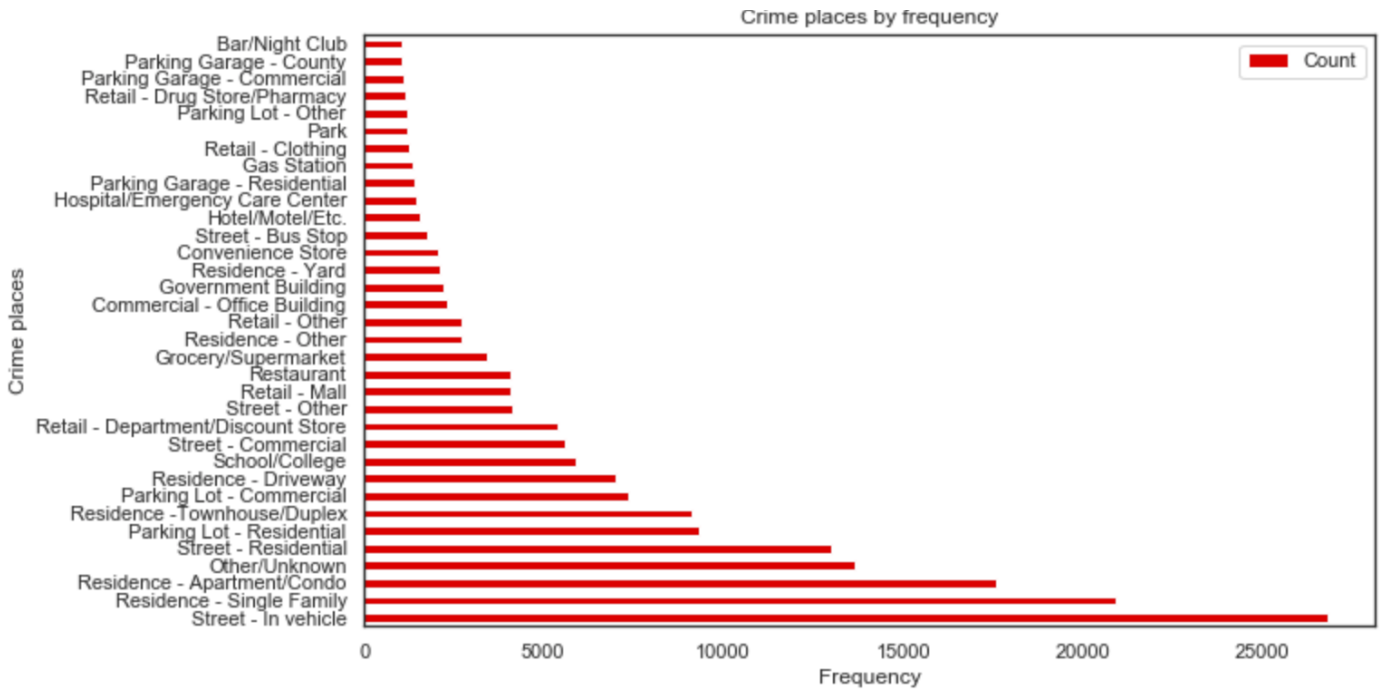
crime map?

For future purpose one can use the other information to create a map displaying rates of crimes

Crime by district



Crime by area



Low crime: Bar/Night Club, residential areas **high** crime place

Common crime types

How to get it

We pivoted our data frame with index = Crime Name3, values = Incident ID , then sort them by count

Top five crime types	
Crime Name	Frequency
Larceny-from auto	16811
Drugs-Marijuana-possess	13983
Police Information	11194
Driving under the influence of liquor	10878
Assault- 2nd degree	10858

lowest crimes

Bottom five crime types	
Crime Name	Frequency
Homicide- negligent manslaughter	1
Compounding crime	1
Condit release violation	1
Conservation- animals (describe offense)	1
Damage property- business-with explosive	1

Conclusions

- The beginning of the week was the period most crimes (Monday, Tuesday and Wednesday).
- It is also seen that February and March, have lesser crime incidents.
- Most of the crimes committed are in and around Silver Spring district.
- We also observed that the most crimes occur around residential places, and least crimes are committed in places like Bar/Club.
- The most common crime is Larceny followed by drugs/marijuana possession.
- The least crimes are crimes such as: homicide, damage property, etc.

Limitation of the study

- Is it possible to determine the longest and shortest police response time? Does the response time vary from district to district?
- It will be insightful and helpful to classify the types of crimes by whether they are violent or not.
- Furthermore, a relevant information could be to know in which places there are more occurrences of crimes and in which shift these crimes are most common.
- Further investigation is required to answer why 2016 registered low crimes.
- Will we arrive at different conclusion if we make analysis based on population density instead of district?

Thank you for your attention