Project_Stat184

Catherine Galant

Project Introduction

My goal with this project is to look at how crime in Montgomery County has changed over time.

To answer my guiding questions I plan to look at how crime changes over time for a location. My first guiding question is really to help narrow down my scope, by picking out 5-10 of the most prominent offenses in Montgomery County I can take a closer look at they change through the past 9 years.

I plan to facet out the data by crime and/or city and look at how it changes by month/year by plotting the total number of crimes.

Primary Dataset

My primary datasets is from catalog.Data.gov and is public crime data from Montgomery County, MD that catalogs every crime reported in the county from July 1 2016 to Aug 2 2025. It has 38 variables detailing the specifics of the crime like the exact time, location, charge, victims, location, etc.

This dataset is useful for my guiding questions because it gives specific location, crime, and time data. It will tell me exactly what crime was committed, divided into categories 3 times so to specify the exact incindent. It also includes multiple different location categories, like the exact address, city, and police district.

```
-- Attaching core tidyverse packages ---------------------- tidyverse 2.0.0 --
v dplyr     1.1.4     v readr     2.1.5
v forcats   1.0.0     v stringr   1.5.1
v ggplot2   3.5.2     v tibble    3.3.0
v lubridate 1.9.4     v tidyr     1.3.1
v purrr     1.1.0
-- Conflicts -------------------------------------------- tidyverse_conflicts() --
```

```
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to beco

  Incident.ID Offence.Code CR.Number    Dispatch.Date...Time
1   201194204         2305 180031343 06/25/2018 07:03:12 PM
2   201192754         3562 180029527 06/15/2018 09:28:52 AM
3   201197266         5404 180035053 07/15/2018 10:14:12 PM
4   201197538         2399 180035451 07/17/2018 10:11:04 PM
5   201194239         1399 180031372 06/25/2018 10:14:39 PM
         Start_Date_Time          End_Date_Time NIBRS.Code Victims
1 06/22/2018 06:00:00 PM 06/22/2018 08:00:00 PM        23F       1
2 06/15/2018 09:28:00 AM 06/15/2018 09:28:00 AM        35A       1
3 07/15/2018 10:14:00 PM 07/15/2018 11:45:00 PM        90D       1
4 07/17/2018 01:40:00 PM 07/17/2018 02:20:00 PM        23H       1
5 06/25/2018 10:14:00 PM                               13B       1
            Crime.Name1                 Crime.Name2
1 Crime Against Property    Theft From Motor Vehicle
2  Crime Against Society    Drug/Narcotic Violations
3  Crime Against Society Driving Under the Influence
4 Crime Against Property           All other Larceny
5   Crime Against Person              Simple Assault
                      Crime.Name3 Police.District.Name
1              LARCENY - FROM AUTO   MONTGOMERY VILLAGE
2        DRUGS - MARIJUANA - POSSESS             WHEATON
3 DRIVING UNDER THE INFLUENCE LIQUOR            WHEATON
4        LARCENY (DESCRIBE OFFENSE)          GERMANTOWN
5              ASSAULT - 2ND DEGREE          GERMANTOWN
                  Block.Address                City State Zip.Code Agency
1     19200 BLK  WATKINS MILL RD MONTGOMERY VILLAGE    MD    20886   MCPD
2 17000 BLK  BATCHELLORS FOREST RD              OLNEY    MD    20832   MCPD
3                                       SILVER SPRING    MD    20906   MCPD
4                                          GERMANTOWN    MD    20874   MCPD
5        12600 BLK  GREY EAGLE CT          GERMANTOWN    MD    20874   MCPD
                     Place Sector Beat PRA Address.Number Street.Prefix
1    Parking Lot - Commercial      R  6R2 546          19200
2 School/College - DO NOT USE      J  4J2 391          17000
3         Street - In vehicle      K  4K2 336             NA
4         Street - In vehicle      N  5N2 594             NA
5 Residence - Apartment/Condo      N  5N1 447          12600
        Street.Name Street.Suffix Street.Type Latitude Longitude
1     WATKINS MILL                         RD 39.17057  -77.2089
```

```
2 BATCHELLORS FOREST                      RD 39.13580  -77.0453
3            WELLER                        RD 39.06421  -77.0690
4           CLOPPER                        RD 39.14830  -77.2377
5         GREY EAGLE                       CT 39.16706  -77.2635
  Police.District.Number          Location
1                     6D (39.1706, -77.2089)
2                     4D (39.1358, -77.0453)
3                     4D  (39.0642, -77.069)
4                     5D (39.1483, -77.2377)
5                     5D (39.1671, -77.2635)
```

Secondary Dataset

My secondary dataset is from opendata.Maryland.gov, it includes a tally of violent crimes committed every year from 1975 to 2022 broken up by counties in Maryland. It also gives the percent change of each crime by year, the population of the county, and the crime rate per 100,000 people for each violent crime.

This dataset is important for my guiding questions because it will help compare how Montgomery County compares to it's surrounding counties in terms of crime rates.

```
        JURISDICTION YEAR POPULATION MURDER RAPE ROBBERY AGG..ASSAULT B...E
1 Allegany County 1975      79655       3    5      20          114   669
2 Allegany County 1976      83923       2    2      24           59   581
3 Allegany County 1977      82102       3    7      32           85   592
4 Allegany County 1978      79966       1    2      18           81   539
5 Allegany County 1979      79721       1    7      18           84   502
  LARCENY.THEFT M.V.THEFT GRAND.TOTAL PERCENT.CHANGE VIOLENT.CRIME.TOTAL
1          1425        93        2329             NA                 142
2          1384        73        2125           -8.8                  87
3          1390       102        2211            4.0                 127
4          1390       100        2131           -3.6                 102
5          1611        99        2322            9.0                 110
  VIOLENT.CRIME.PERCENT VIOLENT.CRIME.PERCENT.CHANGE PROPERTY.CRIME.TOTALS
1                  6.1                           NA                  2187
2                  4.1                        -38.7                  2038
3                  5.7                         46.0                  2084
4                  4.8                        -19.7                  2029
5                  4.7                          7.8                  2212
  PROPERTY.CRIME.PERCENT PROPERTY.CRIME.PERCENT.CHANGE
1                   93.9                           NA
2                   95.9                          -6.8
```

```
3                     94.3                       2.3
4                     95.2                      -2.6
5                     95.3                       9.0
  OVERALL.CRIME.RATE.PER.100.000.PEOPLE
1                             2923.9
2                             2532.1
3                             2693.0
4                             2664.9
5                             2912.7
  OVERALL.PERCENT.CHANGE.PER.100.000.PEOPLE
1                                     NA
2                                  -13.4
3                                    6.4
4                                   -1.0
5                                    9.3
  VIOLENT.CRIME.RATE.PER.100.000.PEOPLE
1                             178.3
2                             103.7
3                             154.7
4                             127.6
5                             138.0
  VIOLENT.CRIME.RATE.PERCENT.CHANGE.PER.100.000.PEOPLE
1                                             NA
2                                          -41.8
3                                           49.2
4                                          -17.5
5                                            8.2
  PROPERTY.CRIME.RATE.PER.100.000.PEOPLE
1                             2745.6
2                             2428.4
3                             2538.3
4                             2537.3
5                             2774.7
  PROPERTY.CRIME.RATE.PERCENT.CHANGE.PER.100.000.PEOPLE
1                                              NA
2                                           -11.6
3                                             4.5
4                                             0.0
5                                             9.4
  MURDER.PER.100.000.PEOPLE RAPE.PER.100.000.PEOPLE ROBBERY.PER.100.000.PEOPLE
1                       3.8                     6.3                       25.1
2                       2.4                     2.4                       28.6
3                       3.7                     8.5                       39.0
```

```
4                          1.3                   2.5                        22.5
5                          1.3                   8.8                        22.6
  AGG..ASSAULT.PER.100.000.PEOPLE B...E.PER.100.000.PEOPLE
1                        143.1                 839.9
2                         70.3                 692.3
3                        103.5                 721.1
4                        101.3                 674.0
5                        105.4                 629.7
  LARCENY.THEFT.PER.100.000.PEOPLE M.V.THEFT.PER.100.000.PEOPLE
1                         1789.0                          116.8
2                         1649.1                           87.0
3                         1693.0                          124.2
4                         1738.2                          125.1
5                         2020.8                          124.2
  MURDER..RATE.PERCENT.CHANGE.PER.100.000.PEOPLE
1                                             NA
2                                          -36.7
3                                           53.3
4                                          -65.8
5                                            0.3
  RAPE.RATE.PERCENT.CHANGE.PER.100.000.PEOPLE
1                                         NA
2                                      -62.0
3                                      257.8
4                                      -70.7
5                                      251.1
  ROBBERY.RATE.PERCENT.CHANGE.PER.100.000.PEOPLE
1                                             NA
2                                           13.9
3                                           36.3
4                                          -42.2
5                                            0.3
  AGG..ASSAULT..RATE.PERCENT.CHANGE.PER.100.000.PEOPLE
1                                                   NA
2                                                -50.9
3                                                 47.3
4                                                 -2.2
5                                                  4.0
  B...E.RATE.PERCENT.CHANGE.PER.100.000.PEOPLE
1                                          NA
2                                       -17.6
3                                         4.2
4                                        -6.5
```

```
5                                                       -6.6
  LARCENY.THEFT..RATE.PERCENT.CHANGE.PER.100.000.PEOPLE
1                                                         NA
2                                                       -7.8
3                                                        2.7
4                                                        2.7
5                                                       16.3
  M.V.THEFT..RATE.PERCENT.CHANGE.PER.100.000.PEOPLE
1                                                         NA
2                                                      -25.5
3                                                       42.8
4                                                        0.7
5                                                       -0.7
```

Data Wrangling

To get my data into a usable format, I will select only the variables that I need from and filter it so that only cases relevant to my data are observed. For my primary dataset 'Crime.csv' from catalog.Data.gov that has every crime reported from July 1, 2016 to August 2, 2025.

I first separate the data into 2 different sets that I would like to look at, the first is how all crime rates change over the time period, the second is how the most common crimes change over time. To identify the most common crimes I group by the common crime name in Crime.Name2 and summarise the count, arrange in descending order and get the top 9 crimes.

These next steps I do to both dataframes.

I only want the start date and time, the specific crime, and the Police District it happened in. I have to convert the start date and time into a POSIXct date time so that I can use it later and I mutate the column to do that. Then I create new columns that store only the year and the month and get rid of the original Start_Date_Time column. Next I filter out all the minor offenses underneath "All Other Offenses" and all the unfinished Police District information, usually written as "OTHER" or left blank.

For the first dataframe that looks at all crimes over the months and years, named mont_crime_year_district. To get the summary data I need for graphing, I group it by Year, Month, and Police.District, and then summarize it by count. This gets me my tidied dataframe needed for graphing. For the second dataframe, named mont_year_bycrime, that looks at 9 of the top crimes by year, I first filter the data to only have those top 9 crimes, and then I group by year and crime, and summarize again by count. Now this dataframe is also ready to be plotted. Below are the first 5 lines of each dataframe.

```
`summarise()` has grouped output by 'Year', 'Month'. You can override using the
`.groups` argument.
`summarise()` has grouped output by 'Year'. You can override using the
`.groups` argument.


# A tibble: 5 x 4
# Groups:   Year, Month [5]
  Year  Month Police.District.Name   num
  <chr> <chr> <chr>                <int>
1 2017  10    SILVER SPRING          909
2 2017  09    SILVER SPRING          889
3 2019  05    SILVER SPRING          885
4 2017  08    SILVER SPRING          871
5 2019  10    SILVER SPRING          868


# A tibble: 5 x 3
# Groups:   Year [4]
  Year  Crime.Name2                 num
  <chr> <chr>                     <int>
1 2018  Drug/Narcotic Violations   6045
2 2017  Drug/Narcotic Violations   5889
3 2024  Shoplifting                5455
4 2020  Theft From Motor Vehicle   4993
5 2017  Theft From Motor Vehicle   4956
```

For my secondary dataset, it has the total number of the crimes it looks at as individual
columns, so I need to pivot the dataframe to be longer and make a new column that has what
crime was committed. After exploratory data analysis I discovered that

Data Visualization
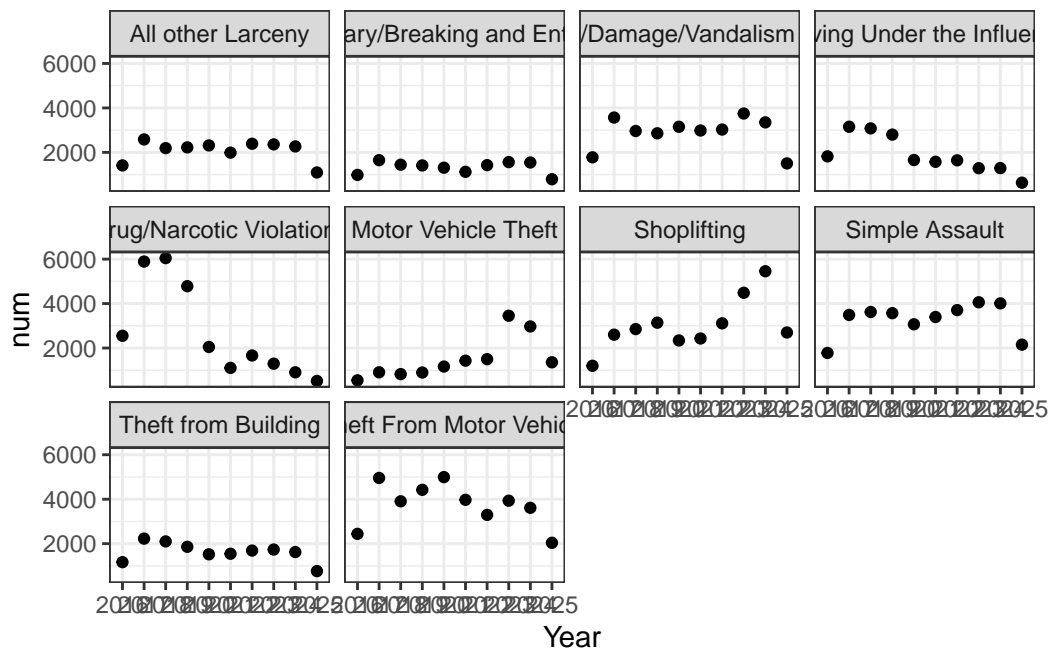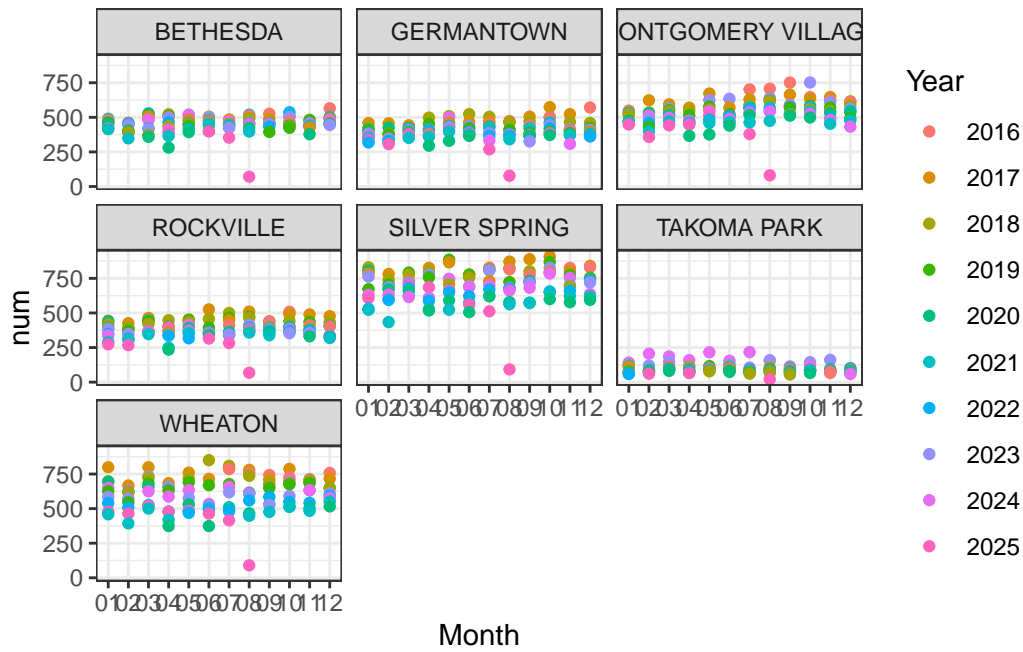
Number of other Crime Incidents in Prominent Maryland Cou
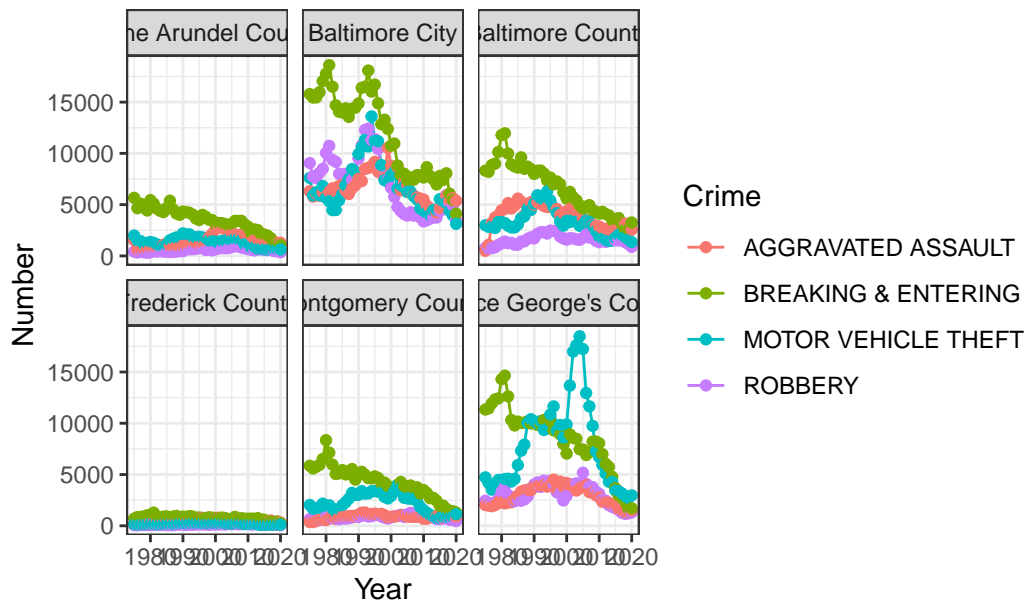
Number of Murder and Rape Incidents in Prominent Maryland
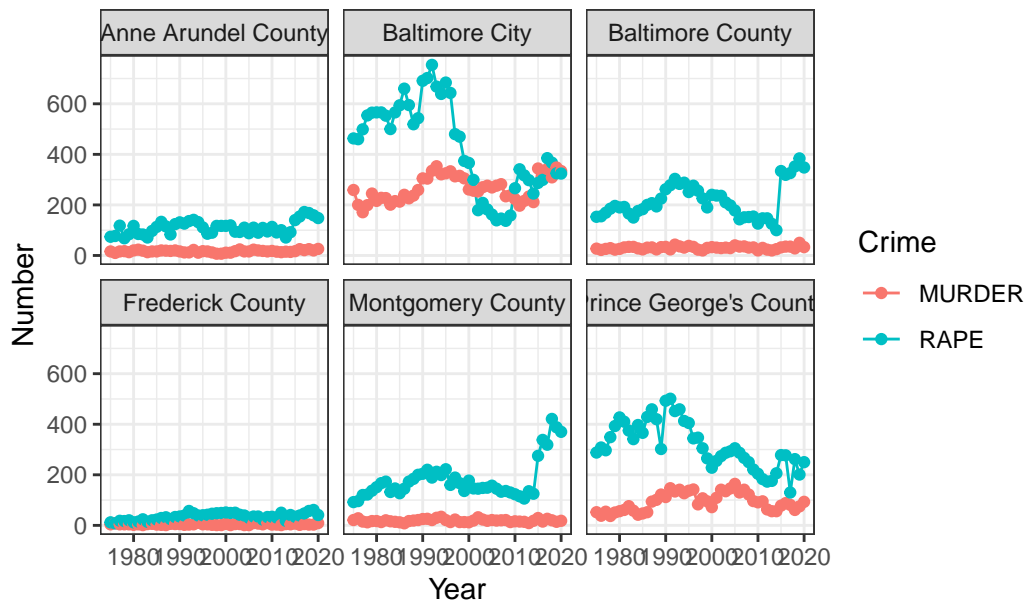
Code Appendix

```r
#My chosen style guide is the tidyverse style guide

library(tidyverse)
library(ggplot2)

mont_crime<-read.csv("C:\\Users\\catie\\Downloads\\Crime.csv")
head(mont_crime,5)
md_crime<-read.csv("C:\\Users\\catie\\Downloads\\Violent_Crime___Property_Crime_by_County__19

head(md_crime,5)
mont_crime<-mont_crime%>%select(Start_Date_Time,Crime.Name2,Police.District.Name)


mont_crime_year_district<-mont_crime%>%mutate(Start_Date_Time=as.POSIXct(Start_Date_Time,form
  mutate(Year=format(Start_Date_Time,format='%Y'))%>%mutate(Month=format(Start_Date_Time,form
  select(-Start_Date_Time)%>%filter(Crime.Name2!='All Other Offenses')%>%
  filter(Police.District.Name!='OTHER')%>%filter(Police.District.Name!='')%>%
  group_by(Year,Month,Police.District.Name)%>%summarise(num=n())%>%arrange(desc(num))
mont_crime_year_district<-na.omit(mont_crime_year_district)


top9_crimes<-(mont_crime%>%group_by(Crime.Name2)%>%summarise(count=n())%>%filter(Crime.Name2
mont_year_bycrime<-mont_crime%>%mutate(Start_Date_Time=as.POSIXct(Start_Date_Time,format = "%
  mutate(Year=format(Start_Date_Time,format='%Y'))%>%mutate(Month=format(Start_Date_Time,form
  select(-Start_Date_Time)%>%filter(Crime.Name2%in%top9_crimes)%>%
  filter(Police.District.Name!='OTHER')%>%filter(Police.District.Name!='')%>%
  group_by(Year,Crime.Name2)%>%summarise(num=n())%>%arrange(desc(num))
mont_year_bycrime<-na.omit(mont_year_bycrime)

head(mont_crime_year_district,5)
head(mont_year_bycrime,5)


md_crime<-md_crime%>%filter(JURISDICTION%in%c('Montgomery County','Baltimore City','Anne Arur

md_crime_high<-md_crime%>%select(JURISDICTION,YEAR,ROBBERY,AGG..ASSAULT,B...E,M.V.THEFT)
md_crime_low<-md_crime%>%select(JURISDICTION,YEAR,RAPE,MURDER)
md_crime_high<-md_crime_high%>%pivot_longer(cols=c(ROBBERY,AGG..ASSAULT,B...E,M.V.THEFT),name
md_crime_high$CRIME[md_crime_high$CRIME=='AGG..ASSAULT']='AGGRAVATED ASSAULT'
md_crime_high$CRIME[md_crime_high$CRIME=='B...E']='BREAKING & ENTERING'
md_crime_high$CRIME[md_crime_high$CRIME=='M.V.THEFT']='MOTOR VEHICLE THEFT'
```

```r
md_crime_low<-md_crime_low%>%pivot_longer(cols=c(MURDER,RAPE),names_to = 'CRIME')
p_year_district<-ggplot(data=mont_crime_year_district,mapping = aes(x=Month,y=num,colour = Ye
  theme_bw()
p_year_district

p_year_crime<-ggplot(data=mont_year_bycrime,mapping = aes(x=Year,y=num))+geom_point(mapping =
p_year_crime

high<-md_crime_high%>%ggplot(mapping = aes(x=YEAR,y=value,colour = CRIME))+facet_wrap(vars(JU
  theme_bw()+ylab('Number')+xlab("Year")+labs(title = "Number of other Crime Incidents in Pro
low<-md_crime_low%>%ggplot(mapping = aes(x=YEAR,y=value,colour = CRIME))+facet_wrap(vars(JURI
  theme_bw()+ylab('Number')+xlab("Year")+labs(title = "Number of Murder and Rape Incidents in

high
low
```