2020 Chicago Crime Rate Analysis

Michaela Dobbs
Ryann Hawkins
Laura Gonzalez
Mario Martinez

Agenda

Introduction & Objective

Trends: Districts, Types, and Monthly Rates

Data Collection & Analysis

API Chicago Crime Rates
(Cleaned Dataset: Year 2020)

Recommendations & Conclusion

Accuracy & Hesitation
Future: Weather Trends

Q&A

Investigators

Michaela Dobbs, Md

Which district has the highest percentage in crime rate?

Ryann Hawkins, BS Ed

What trends do we see in violations?

Laura Gonzalez, MS-IS

What trends do we see among districts?

Mario Martinez, BS

What is the correlation between location vs crime rate?

```
per_year_counts = crimes_df["Year"].value_counts()
per_year_counts
```

[7] ✓ 0.0s

```
Year
2020    8046
2021    1519
2023     610
2022      44
Name: count, dtype: int64
```

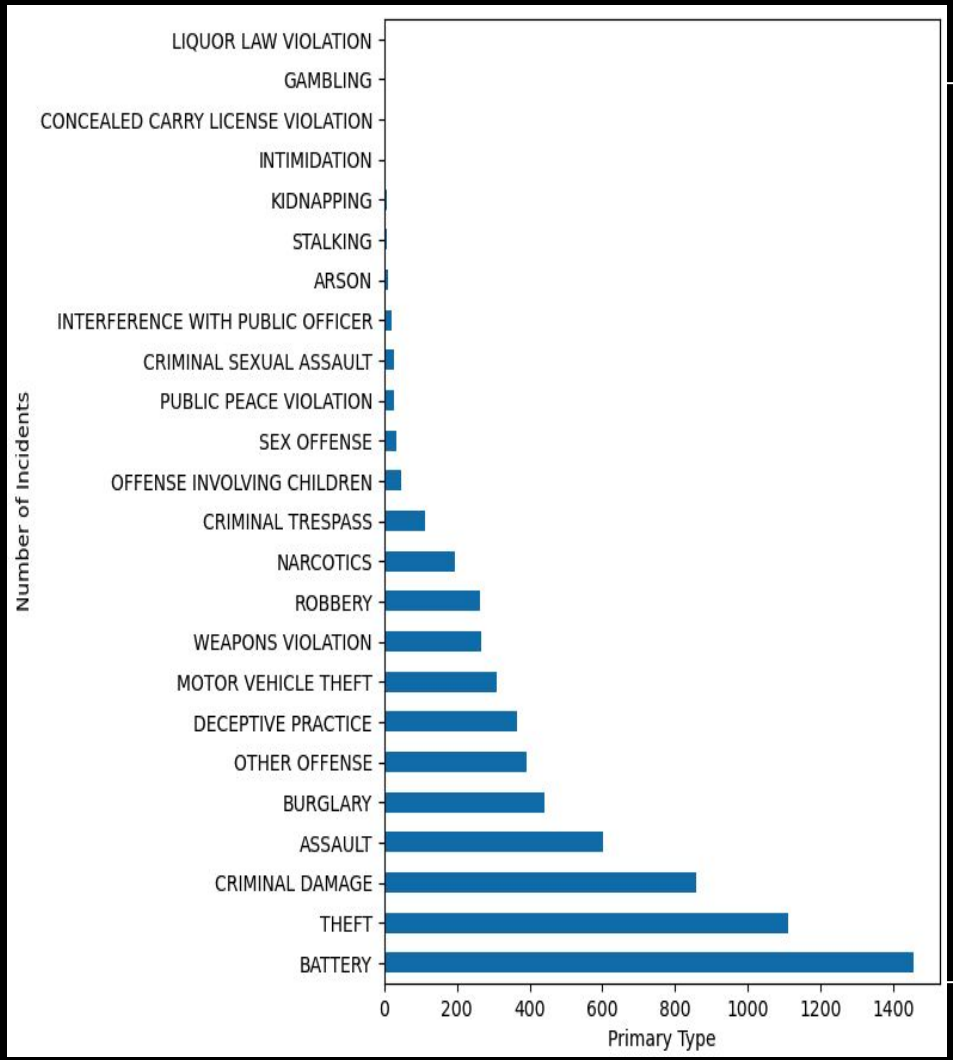
```
per_month_counts = Clean_Crimes20["Month_Name"].value_counts()
per_month_counts
```

[14] ✓ 0.0s

```
Month_Name
March        3199
August       2063
December     1076
July           55
November       53
February       48
January        23
April          11
June            9
October         9
May             8
September       1
Name: count, dtype: int64
```

About our Data

- Chicago data from data.gov
- Contained data from 2020 to present
- Focused on 2020-present, refined to only 2020
- Interested in months, but data was unreliable
- Started with 10,219 data points ended with 6,555



What trends do we see in violations?

```
crime_counts = Clean_Crimes20["Primary Type"].value_counts()
crime_counts
```

✓ 0.1s

Primary Type	
BATTERY	1455
THEFT	1110
CRIMINAL DAMAGE	858
ASSAULT	601
BURGLARY	441
OTHER OFFENSE	391
DECEPTIVE PRACTICE	367
MOTOR VEHICLE THEFT	308
WEAPONS VIOLATION	268
ROBBERY	265
NARCOTICS	196
CRIMINAL TRESPASS	112
OFFENSE INVOLVING CHILDREN	46
SEX OFFENSE	32
PUBLIC PEACE VIOLATION	27
CRIMINAL SEXUAL ASSAULT	26
INTERFERENCE WITH PUBLIC OFFICER	19
ARSON	9
STALKING	8
KIDNAPPING	6
INTIMIDATION	4
CONCEALED CARRY LICENSE VIOLATION	3
GAMBLING	2
LIQUOR LAW VIOLATION	1

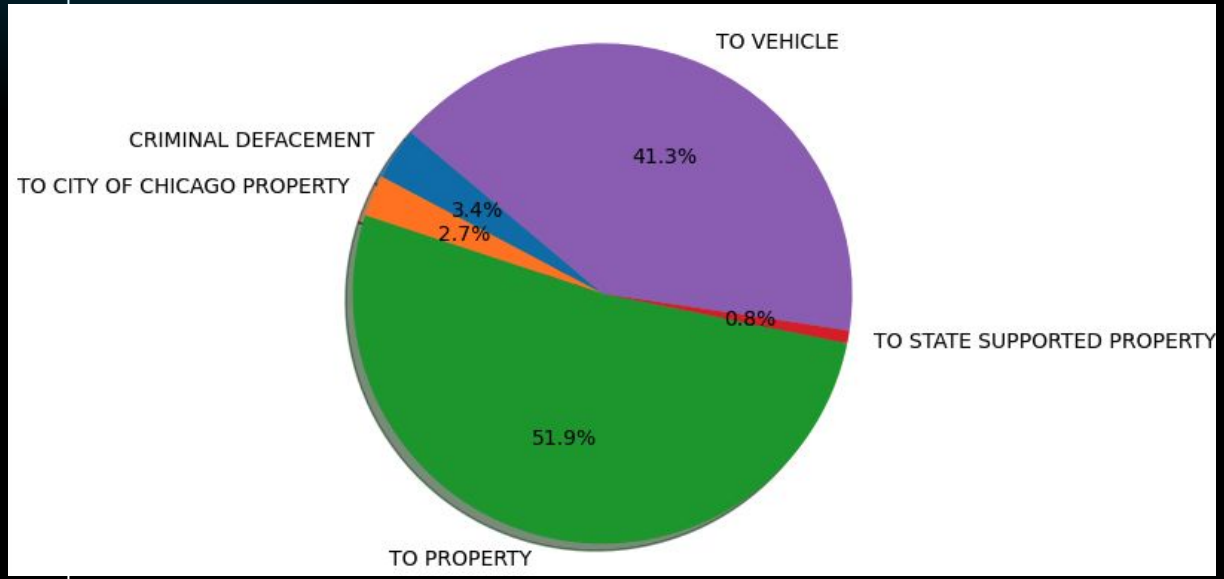
Name: count, dtype: int64

- The top three violations are Battery, Theft, and Criminal Damage.

What trends do we see in violations?

Each violation can be categorized into a description.

The pie chart helps to visualize the descriptions of Criminal Damage.




```
CD_Pie = CD_df.groupby(["Description"]).count()
CD_Pie
```

Description	ID	Case Number	Date
CRIMINAL DEFACEMENT	29	29	29
TO CITY OF CHICAGO PROPERTY	23	23	23
TO PROPERTY	445	445	445
TO STATE SUPPORTED PROPERTY	7	7	7
TO VEHICLE	354	354	354

```
crimes_df_clean["Domestic"].value_counts()
```

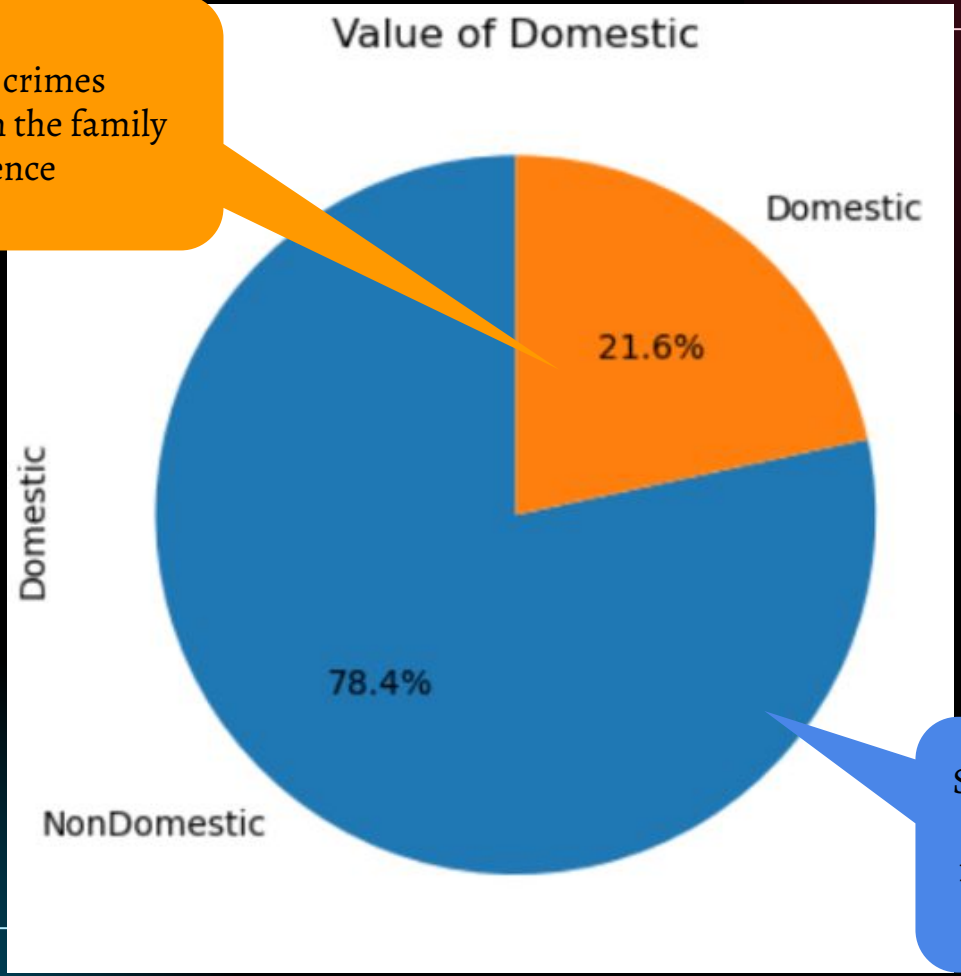
```
False    8009
True     2210
Name: Domestic, dtype: int64
```

```
Domestic = crimes_df_clean["Domestic"].value_counts()

Domestic.plot(kind='pie', labels=["NonDomestic", "Domestic"], autopct='%1.1f%%', startangle=90)
plt.title('Value of Domestic')

plt.show()
```

Refers to crimes committed in the family residence



Simple assault, aggravated assault, and other crimes not committed in a family residence

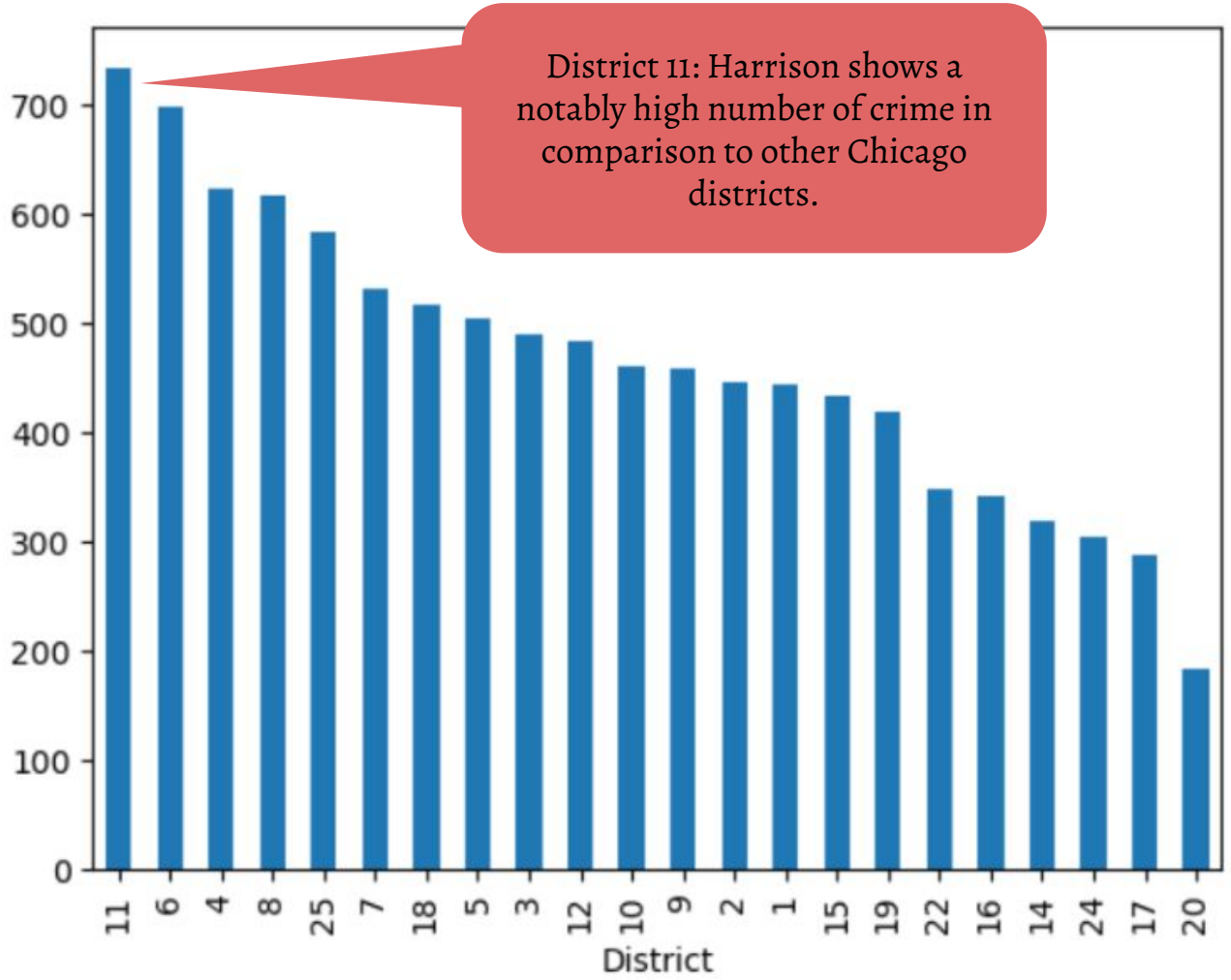

```
crimes_df_clean["District"].value_counts()
```

```
11    733
6     697
4     622
8     616
25    584
7     530
18    516
5     504
3     490
12    484
10    460
9     459
2     446
1     444
15    433
19    419
22    347
16    341
14    318
24    305
17    288
20    183
Name: District, dtype: int64
```

```
Crime_byDistrict = crimes_df_clean.groupby(["District"]).count()['Arrest']
Crime_byDistrict.sort_values(ascending=False).plot(kind='bar')
plt.show()
```

Davina F. Ward, Commander

District 11, Harrison



What district has the highest percentage in crime rate?

Made a for loop pulling all the Districts and the percentages From dataframe.

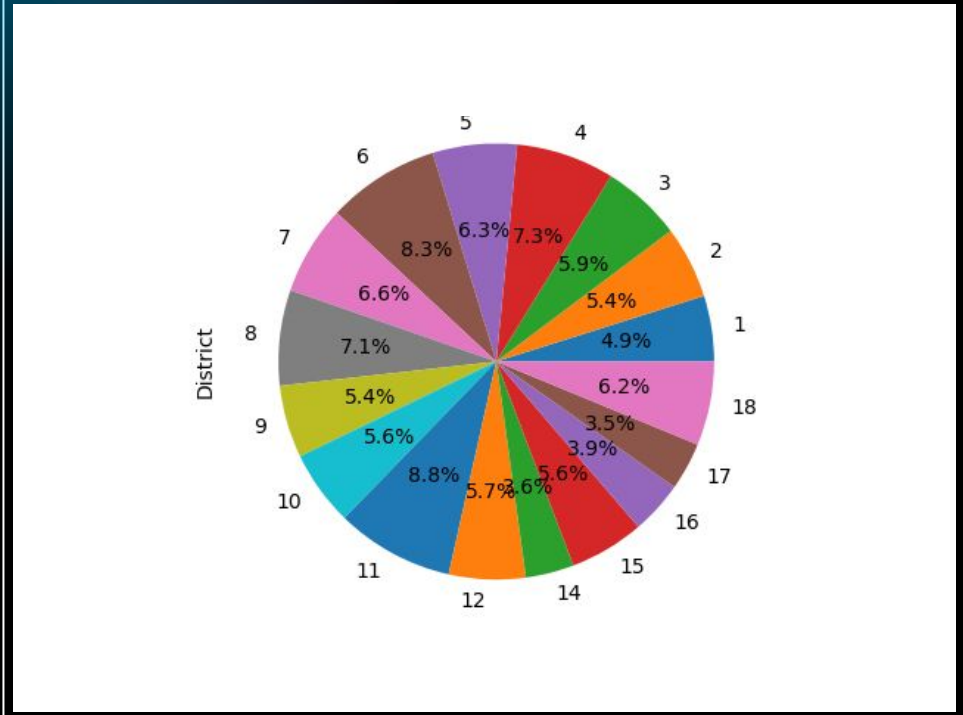
```
    crime_percent= []
    for number in district_number_list:
        percent= len(final_df.loc[final_df["District"] == number]["District"])/(len(final_df))
        crime_percent.append(percent)

    crime_percent
```

[21] ✓ 0.1s

```
[0.0486651411136537,
 0.05354691075514874,
 0.059496567505720827,
 0.07292143401983218,
 0.06254767353165523,
 0.08344774980930587,
 0.06636155606407322,
 0.07078565980167811,
 0.05385202135774218,
 0.05568268497330282,
 0.08756674294431732,
 0.05690312738367658,
 0.03585049580472922,
 0.0555301296720061,
 0.03920671243325705,
 0.03539282990083906,
 0.06224256292906179]
```

Percent of Crimes by District



Each district can be seen

With a percent of crimes. As

You can see by this visualization

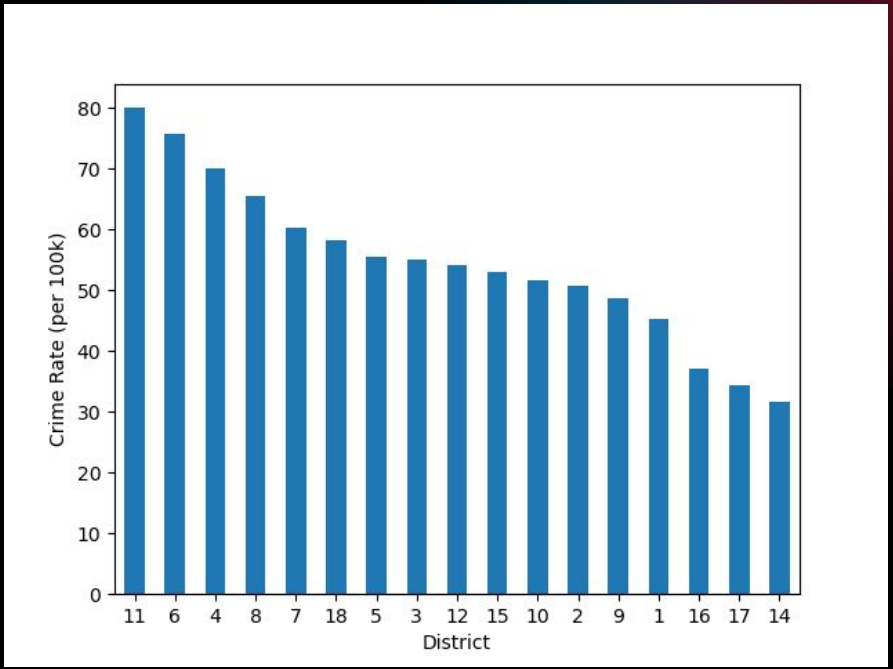
District 11 has the highest at

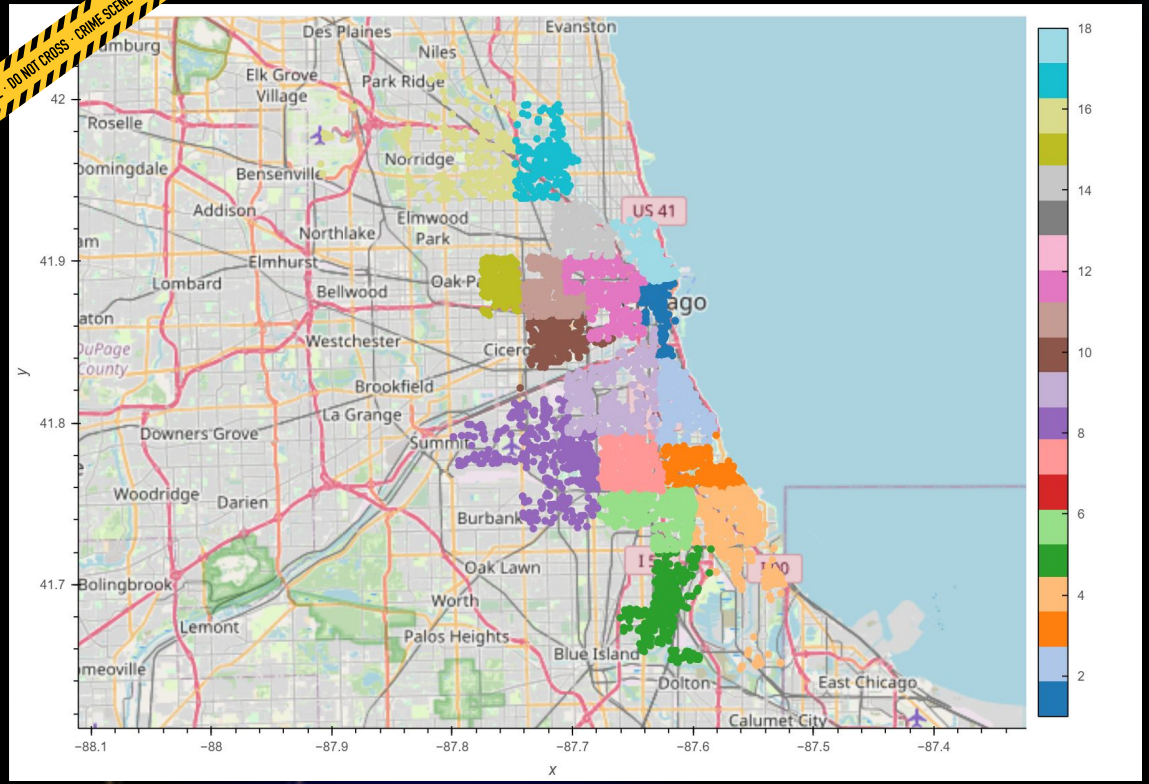
8.8%.

Correlation between Location vs Crime Rate?

Crime Rate by District

- High: 79.99
- Low: 31.69
- Average: 54.51



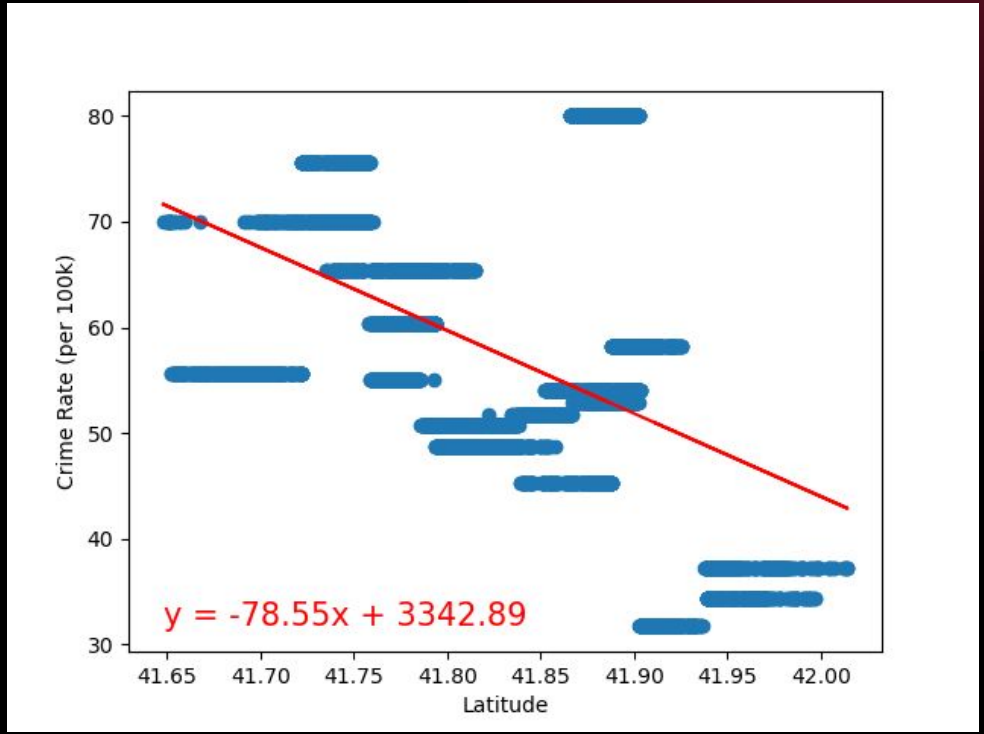


Crime Map by location

- 6555 individual data points
- Each color represents different district

Scatter plot Latitude vs Crime Rate

- Negative slope as Latitude increases
- R-value = 0.21



What can we conclude?

- There is no significant correlation in location and crime rate
- Vast majority of incidents are nondomestic
- Data is NOT reliable :(