Geospatial Analysis of Crime Rates
Group 2: Team Members: Anton LeBeque, Ben Anderson, Emerson Zahab, Michael Pond, Mikel Miller, Richard Kemonou, Shauna Allen

Analysis: The scope of this project is to see if there is or is not a correlation between crime rates in 2017 and 2020. It also was looking to see if there was any specific crime that was increasing and what happened to differently populated areas. Each group member selected one city and the list of the cities selected are: Seattle WA, Dallas TX, Indianapolis IN, Cleveland OH, San Francisco CA, Virginia MN, Missoula MT. The cities selected are a variety of differing locations and total populations. This was to get a better view on how different sized cities have changed over the last three years with regards to crime rate.

The two datasets that were used were taken from the USA Census and FBI data frames, utilizing api keys to pull in the day for analysis. The analysis was conducted on VS Code and utilizing Jupyter notebooks and collaborative efforts were through GitHub.

The first thing that we graphed was the poverty rates in the different cities for the years 2017 and 2020 to see the change in poverty rates during the three year span. As seen in the graph, all of the cities except Virginia, MN saw a reduction in poverty rates during this three year span. One main difference between

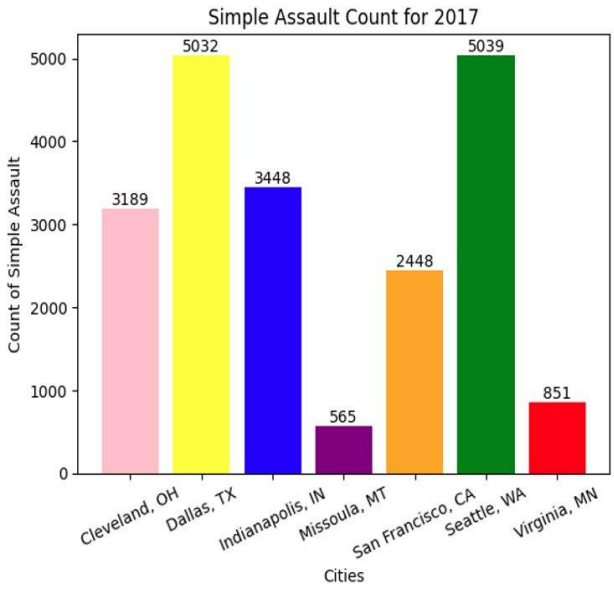
Virginia, MN and the other cities is that its population is drastically smaller than the other cities and that there are less job opportunities in this area. Meaning that during the pandemic and lockdowns, there were less resources available for the smaller community, leading to an increase in the poverty rate.

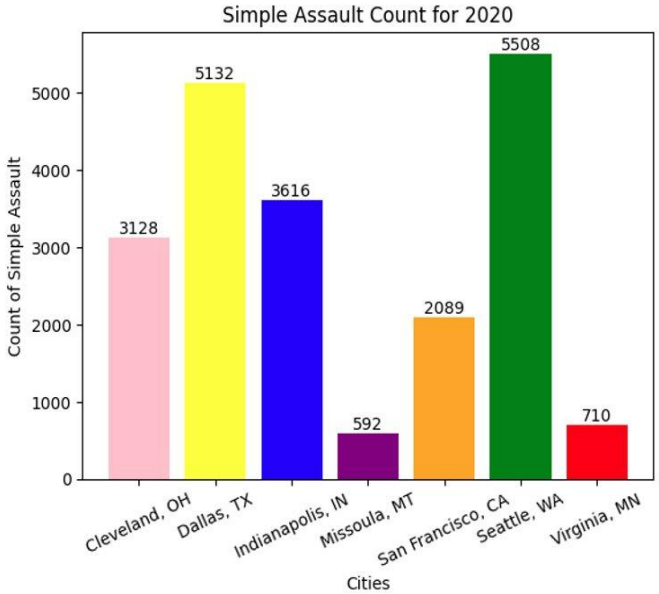
The following course of action was getting the total crime data in these cities for the years 2017 and 2020. The following data tables show the total simple assault, aggravated assault, robbery, murder, and manslaughter in the selected cities.

City Name	Year	Simple Assault	Aggravated Assault	Robbery	Murder	Manslaughter
Seattle, WA	2017	5039	1464	668	28	1
Dallas, TX	2017	5032	925	266	32	2
Indianapolis, IN	2017	3448	2225	781	71	4
Cleveland, OH	2017	3189	672	240	34	0
San Francisco, CA	2017	2448	2677	603	44	3
Virginia, MN	2017	851	214	16	8	1
Missoula, MT	2017	565	93	8	2	0

City Name	Year	Simple Assault	Aggravated Assault	Robbery	Murder	Manslaughter
Seattle, WA	2020	5508	1632	607	37	4
Dallas, TX	2020	5132	1069	217	26	5
Indianapolis, IN	2020	3616	1162	435	62	7
Cleveland, OH	2020	3128	703	220	31	0
San Francisco, CA	2020	2089	2172	463	42	3
Virginia, MN	2020	710	206	29	0	4
Missoula, MT	2020	592	170	20	4	0

Upon getting these data points we followed that up with a deeper dive into simple assault since it was the highest occuring offense in all the cities. This deep dive into simple assault showed that Cleveland, San Francisco, and Virginia had a decrease in total simple assault. In Dallas, Indianapolis, Missoula, and Seattle all had an increase during the same time. One thing that sticks out is that two of the three cities, Cleveland and Virginia, were the two highest in poverty rates in both years. Showing that although these cities have been struggling with poverty for years, they were able to reduce the total number of simple assaults.





Upon getting both the census data for the poverty rates and the crime data we used the T-test and the p-value to see if our findings were statistically significant and reject the null hypothesis that there is not a correlation between crime rates in the years 2017 and 2020 or accept the alternative hypothesis that there is a correlation between crime rates in the years 2017 and 2020.

```
T-test resuls for simple assault in 2017 and 2020 include TtestResult(statistic=0.3673383541517043, pvalue=0.7138709618251482, df=154)
T-test resuls for aggravated assault in 2017 and 2020 include TtestResult(statistic=0.6236669670510698, pvalue=0.5337689882162937, df=154)
T-test resuls for robbery in 2017 and 2020 include TtestResult(statistic=0.8271378969434836, pvalue=0.4094387170996685, df=154)
T-test resuls for murder in 2017 and 2020 include TtestResult(statistic=0.43758429604037385, pvalue=0.6623007303287021, df=154)
T-test resuls for manslaughter in 2017 and 2020 include TtestResult(statistic=-1.9946155618253274, pvalue=0.047849517019395035, df=154)
```

Conclusion:

The null hypothesis was that there was no correlation between poverty rates and crime rates. The alternative hypothesis was that there was a correlation between the two. The t-test for the data points in simple assault, aggravated assault, robbery, and murder has all increased during the same time. For the data points with the p value over 0.05, we accept the null hypothesis. For manslaughter, we reject the null hypothesis and accept the alternative hypothesis since the p value is less than 0.05. When looking at the final analysis we can see manslaughter was the only one to have statistical significance. Since the t-statistic was negative, that shows that it has decreased over the years. According to the data that we have gathered, as cities lowered their overall poverty rates, the crime rates increased.