**Team #6**

Angela Banko

Ashley Drewry

Dave Moorman

Todd Schanzlin

**Problem Statement a.k.a Decompose The Ask**

How has Covid impacted crime statistics in select major metropolitan cities?

**Identify Data Sources**

We will be using open source files and exporting data from local city police departments historical data. Additionally, we will be using data NY Times GitHub location.

Crime Data

* Seattle
  + <https://data.seattle.gov/Public-Safety/SPD-Crime-Data-2008-Present/tazs-3rd5>
* Denver
  + <https://www.denvergov.org/opendata/dataset/city-and-county-of-denver-crime>
* Kansas City
  + <https://data.kcmo.org/Crime/KCPD-Crime-Data-2020/vsgj-uufz/data>
* San Francisco
  + <https://data.sfgov.org/Public-Safety/Police-Department-Incident-Reports-2018-to-Present/wg3w-h783>
* Los Angeles
  + <https://data.lacity.org/A-Safe-City/Crime-Data-from-2020-to-Present/2nrs-mtv8>
* Philadelphia
  + <https://www.opendataphilly.org/dataset/crime-incidents/resource/a2373b85-408b-4144-84e9-cb24843e5f2b>
* Washington DC
  + <https://opendata.dc.gov/datasets/crime-incidents-in-2020>

Corona Virus Case Data

* NY Times GitHub for coronavirus data
  + <https://github.com/nytimes/covid-19-data/blob/master/us-counties.csv>

**Define Strategy and Metrics**

* Is there a correlation in crimes and open Covid cases by major cities?
* Change in crime over the last 18 months by type and by select city?
* Are there more or less specific crimes committed? Any noticeable changes in the types of crimes committed? Violent vs non-violent?

**Description of Data Analysis Tools You Plan to Use**

*Describe the tools you plan to use throughout the project. As you might expect, there*

*will be several stages in the project. Make sure to use at minimum the required tools*

*(see Project1 power point).*

* Initially will use Excel and Pivot tables to help clean and sort the data files
* Will create Jupyter Notebook and Pandas combine data from the two separate csv files:

1. Covid Case data
2. Master Crime Data with selected cities combined

* Panda data frames for initial data analysis and csv file merge
* MatPlotLib for graphs and further statistical analysis
* PowerPoint to summarize and present findings

**Describe the Data Products Your Project Will Produce**

*Data products include results of statistical tests, visualizations of the data, websites/apps,*

*or model parameters. What will be the tangible evidence of your discovered insights?*

* May use bar or trend graphs for each select city to detect any noticeable patterns year-over-year.
* Will see if there is a statistical correlation between open Covid case and change in # of crimes committed during that time period.