1. Hypothesis

Hypothesis: The total NYC Arrest rates increased during the initial period COVID-19 pandemic.

Measurement Definitions:
- **Arrest**: Each individual arrest in NYC recorded by the NYPD
- **Sampled timeframe for the initial period of the COVID-19 pandemic:** February 29th to July 2nd 2020.

2. Datasets of Interest

Arrest Data Source: [NYPD Citywide Crime Statistics](#)

A   Total New York City (NYC) Arrests
- Available date range: 2013 to 2019, January 1st 2020 to July 2nd, 2020.
- **Source**: This data is manually extracted every quarter and reviewed by the Office of Management Analysis and Planning.
- **Data Size**: Historical dataset 1.1 GB / Current year dataset 16 MB

B   **How data was used in the analysis**:
  - **Excel**: Total count of arrest by day and month was used based on the ARREST_DATE field.
  - **Tableau**: Latitude and Longitude data from each arrest was processed to create a heatmaps of arrests over a map of NYC.

C   **Data pre-processing before analysis**:
  - OData import of the data to Tableau failed due to file size.
  - Data was narrowed for date range February 29th to July 2nd in 2020, 2019 and 2018.
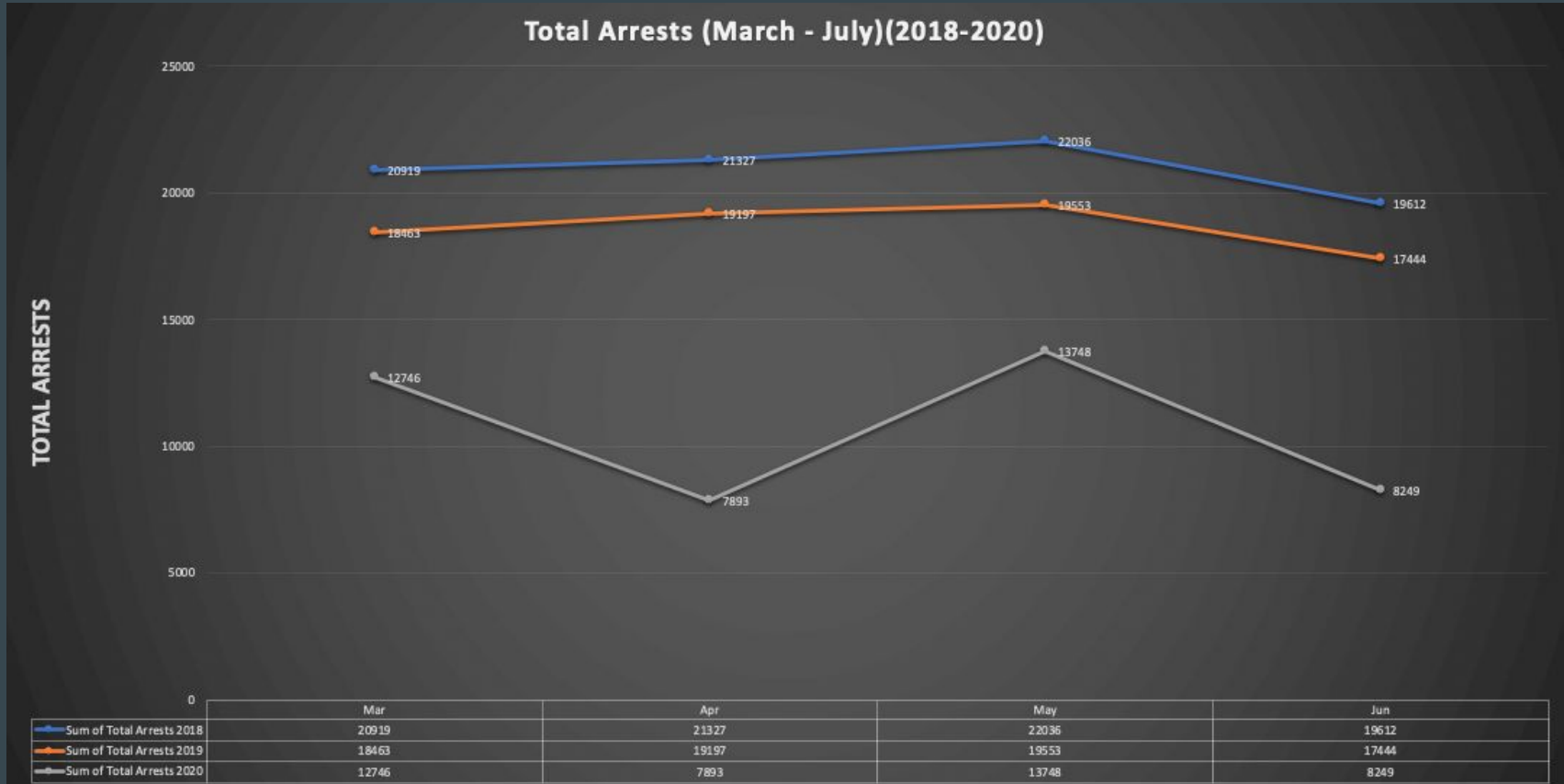 - Data was converted from CSV to xls to be processed in Tableau.

2.b Descriptive Statistics of NYC Arrest Datasets

Total Arrests 2020	
Mean	353.6147541
Standard Error	13.53099102
Median	315
Mode	328
Standard Deviation	149.4546808
Sample Variance	22336.7016
Kurtosis	0.270488722
Skewness	0.97189548
Range	632
Minimum	151
Maximum	783
Sum	43141
Count	122

Total Arrests 2019	
Mean	622.483607
Standard Error	11.1473499
Median	633.5
Mode	715
Standard Deviation	123.126504
Sample Variance	15160.1361
Kurtosis	-1.1030455
Skewness	0.01697759
Range	526
Minimum	362
Maximum	888
Sum	75943
Count	122

Total Arrests 2018	
Mean	695.836066
Standard Error	12.5120257
Median	689.5
Mode	560
Standard Deviation	138.199841
Sample Variance	19099.196
Kurtosis	-0.8578555
Skewness	-0.0308321
Range	579
Minimum	421
Maximum	1000
Sum	84892
Count	122

- The total arrests are highest in 2018 and the total arrests are lowest in 2020
- The maximum arrests were in 2018 and the minimum arrests were in 2020
- The statistics that we're going to use is the sum of each 3 years for comparison
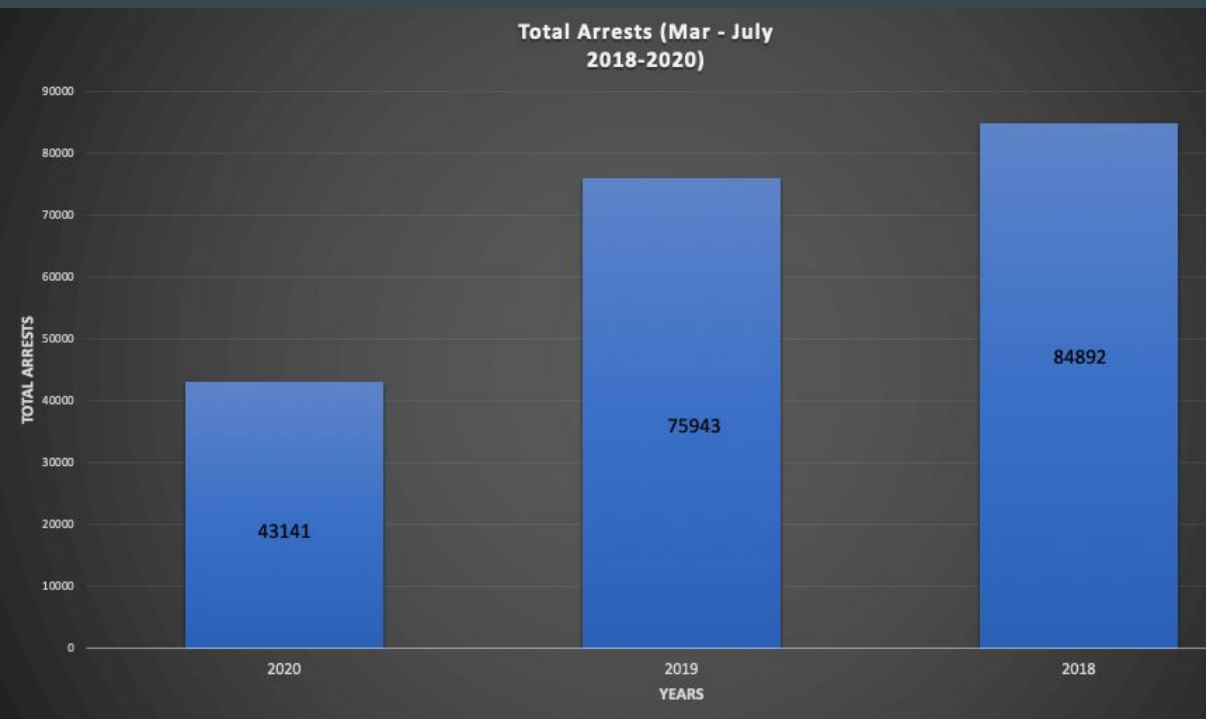
3. Excel Visualizations: Line Chart

Total Arrests (March - July)(2018-2020)

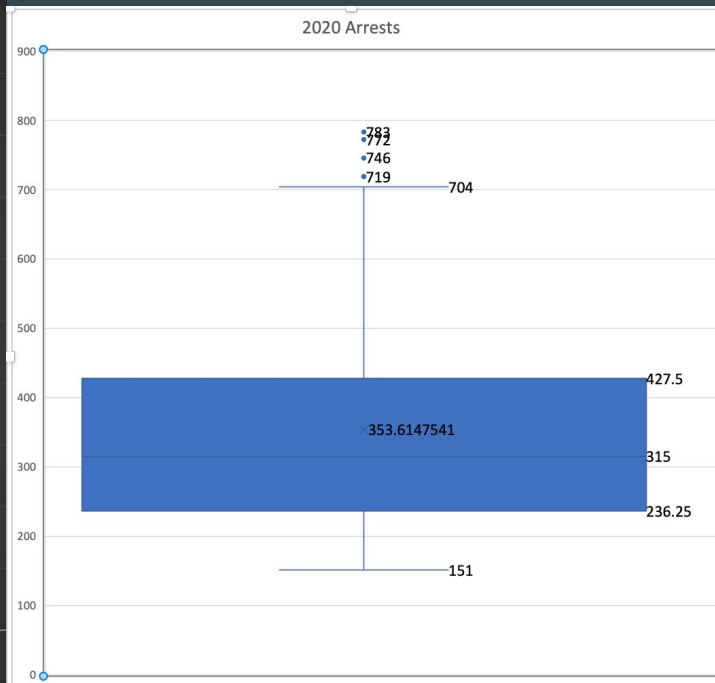|	Mar	Apr	May	Jun
Sum of Total Arrests 2018	20919	21327	22036	19612
Sum of Total Arrests 2019	18463	19197	19553	17444
Sum of Total Arrests 2020	12746	7893	13748	8249

3. Excel Visualizations:

Histogram

Box Plot

3. Tableau Visualizations:

Heat Maps of Arrests by year from Feb 29 to July 2nd

2020   2019   2018

NYC Arrests 2020

NYC Arrests 2019

NYC Arrests 2018
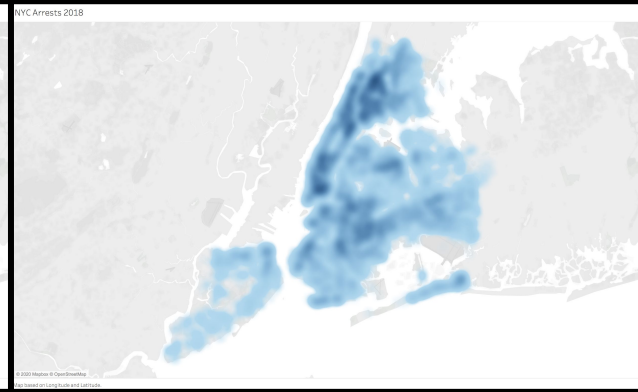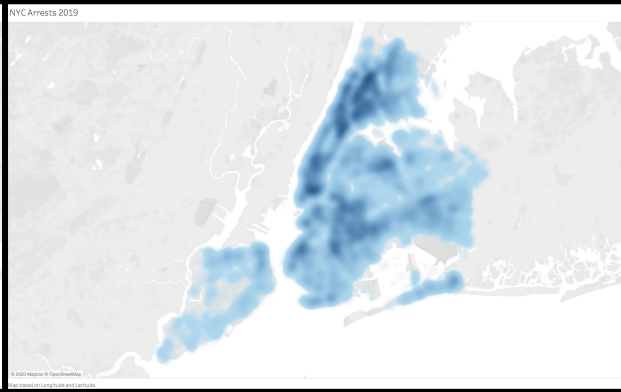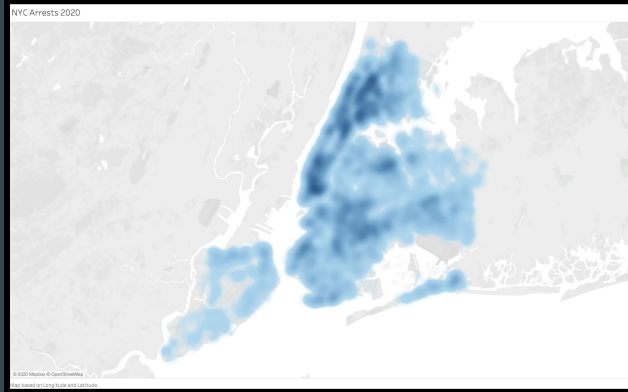
How to read a head map:
- Areas that are not blue had 0 arrests
- Arrests are marked on the NYC map in blue
- The more arrests in an area is indicated by darker shades of blue
- Lighter shades of blue indicate fewer arrests in an area

3. Charts' Insights Regarding our Hypothesis

Insights from Excel Charts

Pivot Table showed us that the arrests actually decreased during the early months of the pandemic. Arrests were actually higher in 2018 than 2019 and 2020. Arrests was higher in 2019 than 20.

The line chart and the histogram both tell us similar information as the pivot table. Arrests has been decreasing since 2018.

Insights from Tableau heat maps

- Heat maps are often useful for exploring data and are best when we have dramatic differences in data.
- Crime appears to be concentrated in similar areas each year.
- Data was too similar to visually see differences of Arrest data in NYC from 2018 to 2020.
- We'd expect dramatic differences in COVID-19 hotspots if it had an impact on arrests.

Alternative hypotheses that could explain results:

- COVID-19 could have reduced the number of arrests by reducing the total number of people outside.
  - Fewer targets of opportunity
- Police were less willing to expose themselves to COVID-19 to make arrests for less serious crimes.

4. Utility and Limitations

Utility:

- Police Departments
  - Allocation of police officers
  - Allocating funding and assets to prevent crime
- Hospitals
  - Staffing efficiency for possible overtime in heavily affected sectors
  - Budgeting for overtime

Limitations Breakdown

Data limitations	Issues	Generalizability of findings
- Borough distinction - Types of Arrests	- Data too similar to make heat map useful - General, no insight to felonies or misdemeanors	Arrests

5. General usefulness and limitations of descriptive analytics for decision making.

Purpose of Descriptive analytics
- Use of historical data to identify patterns or meaning

Usefulness of Descriptive analytics
- Descriptive analytics of similar events can help decision makers implement policies to prevent harm and strengthen communities.
- Even when the same measurement label at times can have similar but different definitions, as indicators they can still help estimate likely outcomes.

Limitations of Descriptive analytics
- Some situations could be considered overloaded with variables. This can make it difficult to identify the most impactful variables on a situation.
- Unique situations may not be accounted for with available historical data.

Links to data files used in this presentation
- [NYPD Arrest Data (Year to Date)](#)
- [NYPD Arrests Data (Historic)](#)