Yet Another Analysis on NYPD Shooting Incidents with visualization

Aaron Li

26/3/2022

This report leverages **five sheets methodology** to create a preliminary visualization design.

Brainstorming

The data that we are working on

We take the historic data set that includes a list of every shooting incident occurred in NYC going back to 2006 through the end of the previous calendar year. You may find the same CSV file used in this report as below :

https://data.cityofnewyork.us/api/views/833y-fsy8/rows.csv

And it contains the following data points:

```
##  [1] "INCIDENT_KEY"          "OCCUR_DATE"
##  [3] "OCCUR_TIME"            "BORO"
##  [5] "PRECINCT"              "JURISDICTION_CODE"
##  [7] "LOCATION_DESC"         "STATISTICAL_MURDER_FLAG"
##  [9] "PERP_AGE_GROUP"        "PERP_SEX"
## [11] "PERP_RACE"             "VIC_AGE_GROUP"
## [13] "VIC_SEX"               "VIC_RACE"
## [15] "X_COORD_CD"            "Y_COORD_CD"
## [17] "Latitude"              "Longitude"
## [19] "Lon_Lat"
```

Goals

The goal for working on this data set is to analyze the number of incidents and deaths per precinct and borough recorded by NYPD since 2006 to previous calendar year. These are the questions to answer:

- How many shooting incidents per precinct or borough in New York city?
- How many deaths caused by shooting per precinct or borough in New York city?
- How many shooting incidents in a borough? How are the incidents distributed?
- Do they likely happen from few precincts with big number or most of precincts with small number each?

Tasks

- The purpose (high-level) - User can use the the visualization to have good understand about shooting incidents happened in New York city with details and the distribution.

- Navigation (middle-level)
  - User can view circles plotted on the visualization and look at the values on X-axis and Y-axis
  - User can also hover over a particular circle to find the details
  - User can select a borough by clicking on one circle belongs to it and view the details on the sub view plotted next to the main view

- Data attributes(low-level)
  - Marks are denoted by circles in the main view
  - Each circle represents a few data attributes attached to it:
    1. Name of precinct
    2. Name of borough
    3. Number of shooting incidents in the precinct
    4. Number of deaths caused by shooting in the precinct

Interations

3 iterations of low-fidelity prototypes are doodled down a few pieces of paper
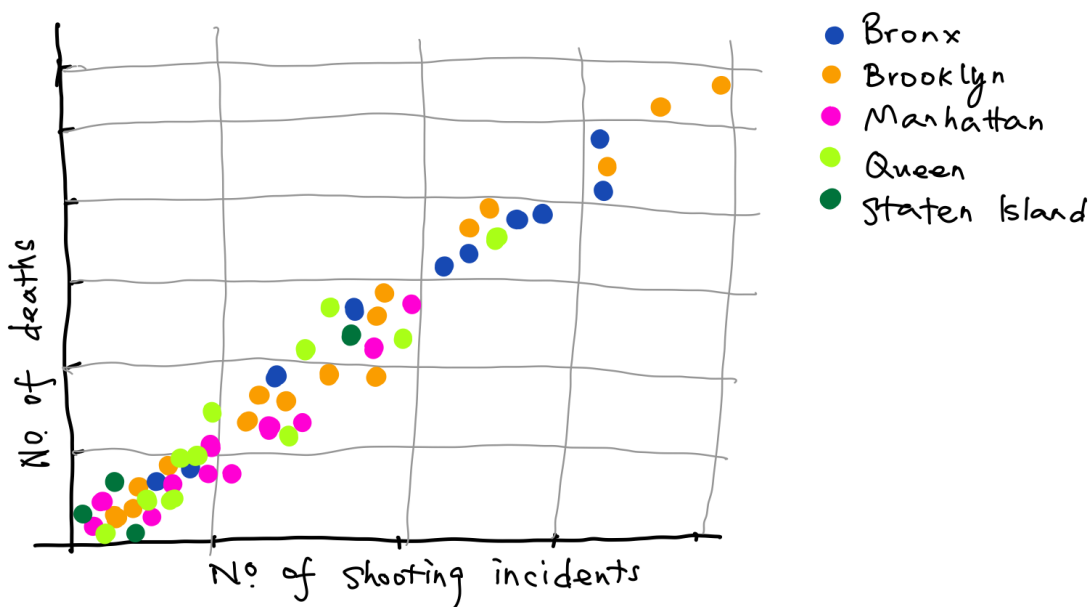
Low-fidelity Prototypes

Figure 1: Low-fdelity Interation 1 .

At the first iteration, a quick prototype is done with x-axis denoting *number of shooting incidents* and y-axis denoting *number of deaths from the*. Each circle represents a data mark of summarized data attributes for a precinct. 5 colors stand for 5 boroughs in New York city respectively.
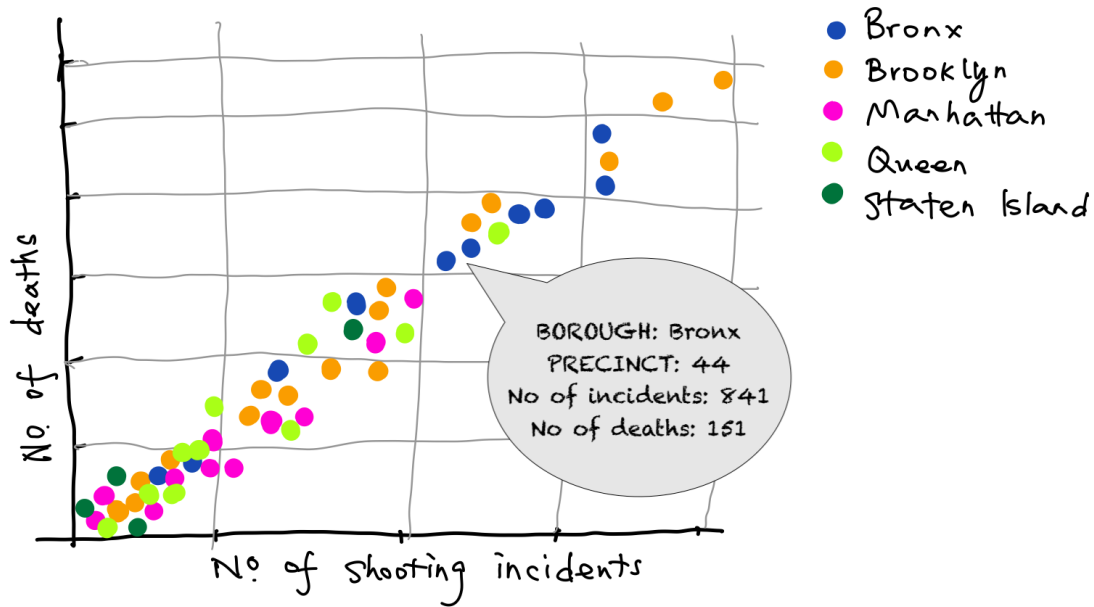
Figure 2: Low-fdelity Interation 2 .

At the second iteration, an overlay pops up when mouse hovers over a circle. It will detail the data attributes for the particular precinct.
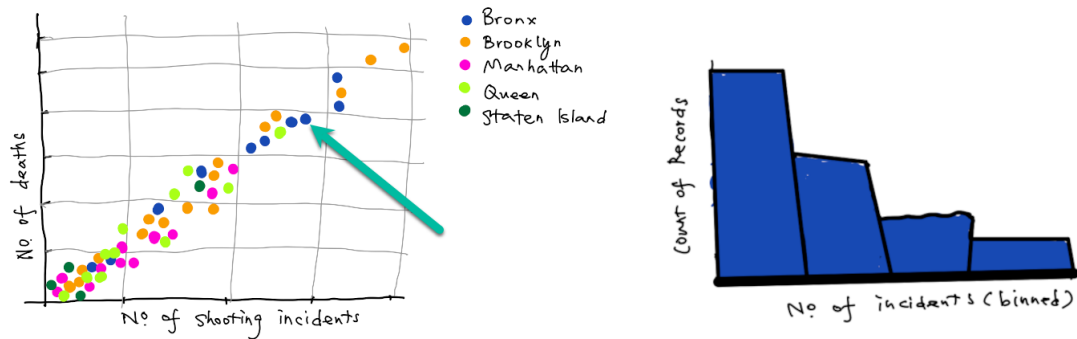

Figure 3: Low-fdelity Interation 3 .

The last iteration adds a sub view to link with the main view. When a borough is selected from main view, the histogram on sub view changes accordingly to give an idea how the incidents are distributed.

Realization

Altair implmentation

Above low-fidelity designs are finally realized into the final interactive high-fidelity design using Altair. The Jyupyter notebook has been uploaded to Google Colab and be viewable at https://colab.research.google.com/drive/1JQU-nRvBwN5sIEwgSKYSErkXcmQE9hvM?usp=sharing
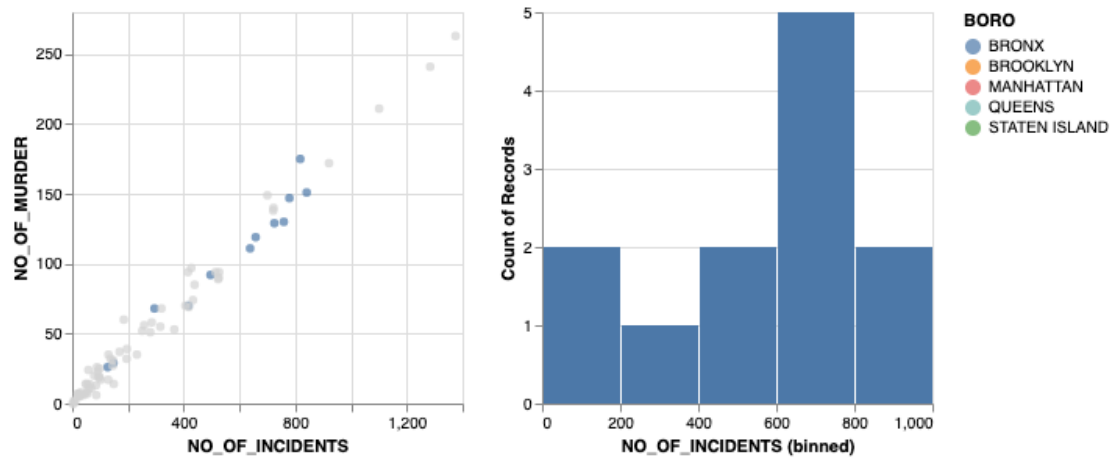
Key elements of design

Figure 4: High-fdelity design .

Here are the key elements put into the design:

1. Selection Once user selects interesting aspects of the data, a borough in this case, the main view make the data not selected less salient.
2. Semantic zoom When a borough is selected, the right view drill down into data representing selected borough only.
3. Overview + Detail Implemented *overview* and *detail* two views: zoom into specific subsets of the data in the *detail (right view)* to enable the analyst to explore their data; Show where that detail is with respect to the larger dataset through an *overview (left view)* indicating where the detail is focused.
4. Two view coordination This technique allows users to interact with data in *main view (left)* and to see that data in *another view (right)*.