NYPD Shooting Incident Data Report

~From Japanese tourist's & data scientist's views~

2022.10.23

Daisaku Ikoma

Agenda

1. Objective of this analysis
2. Data description
3. Tidy data
4. Visualize data
    1. Shooting incidents counts by Borough
    2. Shooting incidents counts by Time Zone
5. Modeling
6. Potential biases and future analysis
7. Summary

1. Objective of this analysis

- New York is one of the most attractive cities for Japanese tourists (museums, Yankee Stadium, Statue of Liberty, musicals, etc.)
- For safe and fun, need to understand the security in NY
- Japanese are less sensitive to shooting incidents
- **Analyze NYPD Shooting incidents as a reference for traveling**
- Shooting incidents by **Borough** and **Time zone**
- For further analysis, the features of the data and specific analysis procedures also be described

Analyzing NY shooting incidents for Japanese travelers

2. Data description

- Analyzing "NYPD Shooting Incident Data (Historic)" https://catalog.data.gov/dataset/nypd-shooting-incident-data-historic
- Every shooting incident that occurred in NYC going back to 2006 through the end of the previous year
- Dataset includes, Rows: 25,596 Columns: 19
- Column specification:

data type	#	columns
character	10	OCCUR_DATE, BORO, LOCATION_DESC, PERP_AGE_GROUP, PERP_SEX, PERP_RACE, VIC_AGE_GROUP, VIC_SEX, VIC_RACE Lon_Lat
double	7	INCIDENT_KEY, PRECINCT, JURISDICTION_CODE, X_COORD_CD, Y_COORD_CD, Latitude, Longitude
logical	1	STATISTICAL_MURDER_FLAG
time	1	OCCUR_TIME

3. Tidy data

- First, check basic statistics (Median, Max/Min, # of NA)
- Feature selection (BORO, OCCUR_TIME, etc.) & drop NAs
- Group by BORO & OCCUR_TIME

Group by BORO

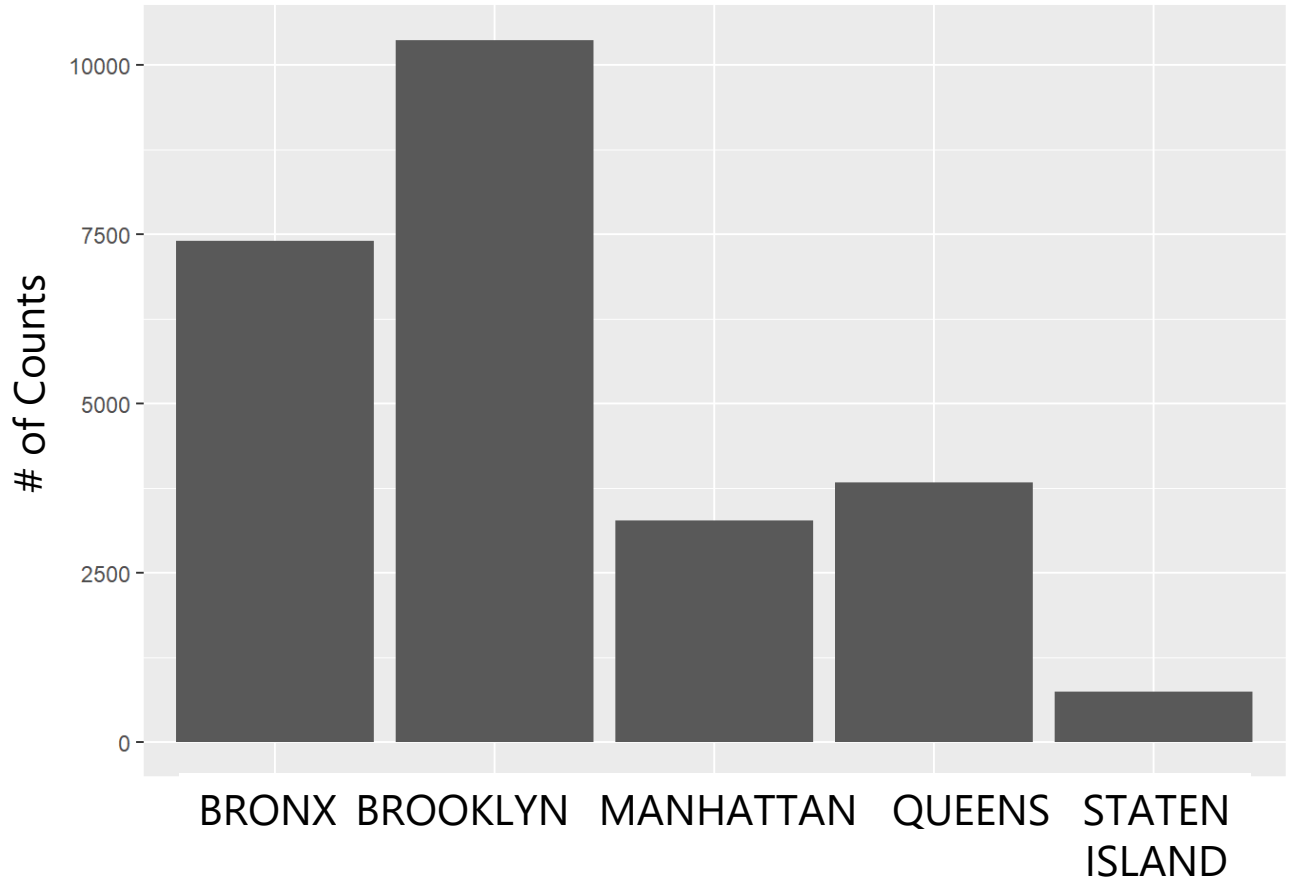
BORO	# of cases
BRONX	7402
BROOKLYN	10365
MANHATTAN	3264
QUEENS	3827
STATEN ISLAND	736

Group by OCCUR_TIME

hours	# of cases	hours	# of cases	hours	# of cases
0	2053	8	206	16	968
1	1981	9	199	17	1010
2	1726	10	273	18	1139
3	1544	11	345	19	1364
4	1374	12	462	20	1573
5	667	13	498	21	1859
6	332	14	733	22	2022
7	222	15	854	23	2190

4. Visualize data (1/2)

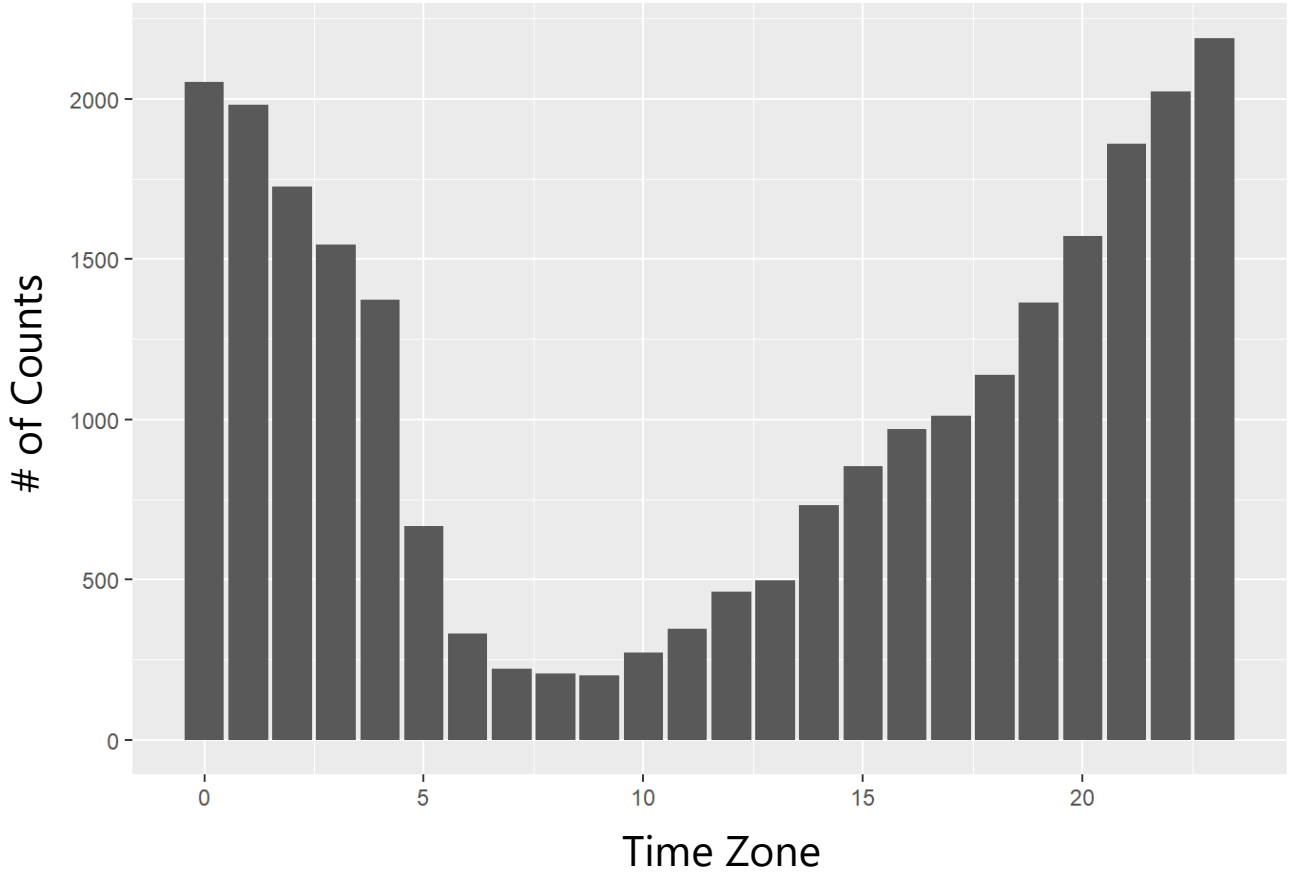
Shooting incidents counts by **Borough**



- **Brooklyn** and **Bronx** have many Shooting incidents
- **Staten Island** is safe with few shooting incidents

4. Visualize data (2/2)

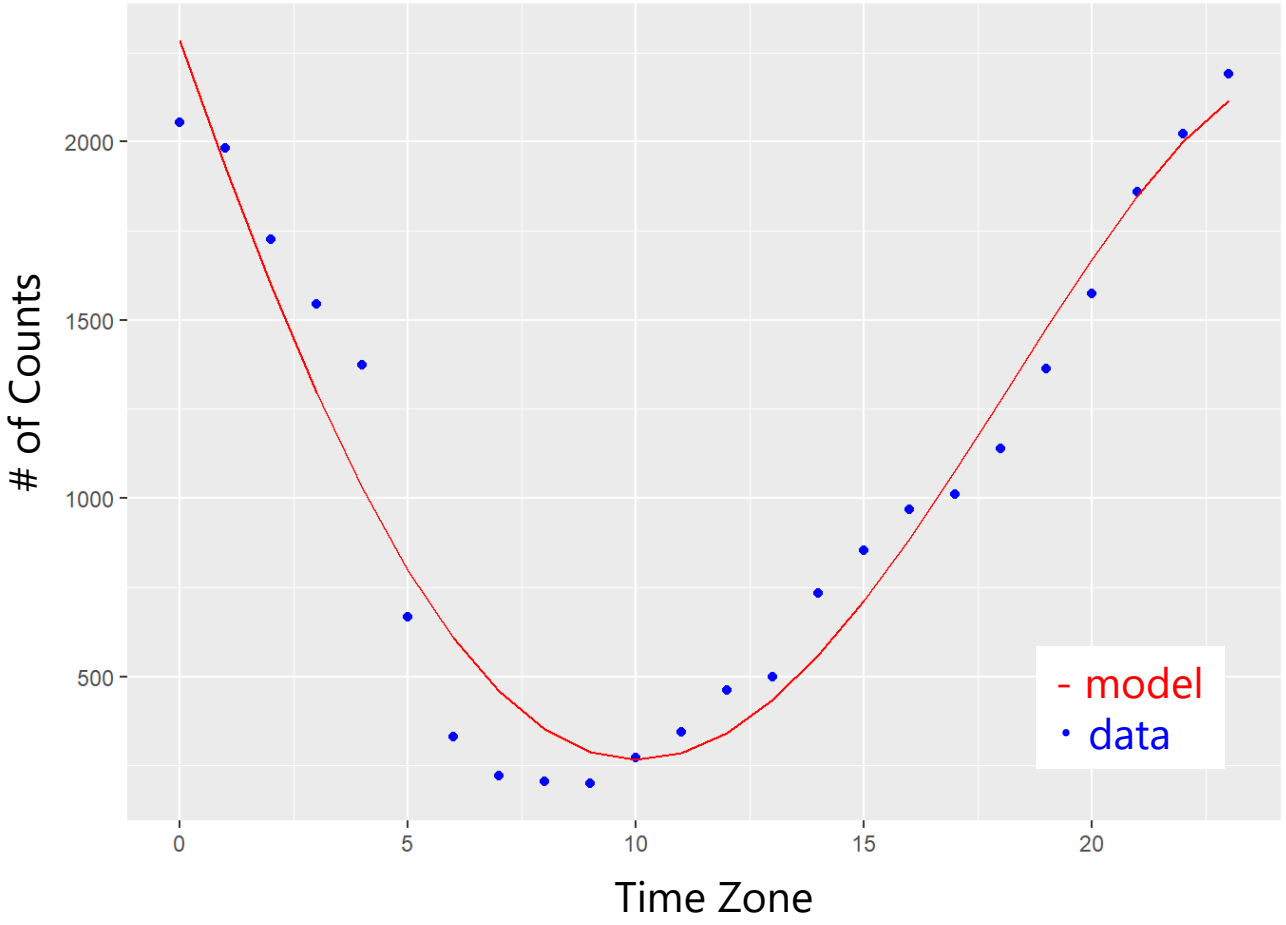
Shooting incidents counts by [Time Zone]



- **In the morning**, 9 o'clock is the minimum, and few shooting incidents
- Increases from **evening to night, peaking at 23:00**

5. Modeling

Shooting incidents counts by [Time Zome]



- 4-order polynomial model fit was performed
- Trends can be represented by models

6. Potential biases and future analysis

- Regarding bias, necessary to investigate the effects of **races**, **residents income**, **educational level**, etc. with objective data
- Brooklyn has the highest number of gun crimes and has the most population, so next need to look at **the number of crimes per capita**.
- **Cross analysis of time zone and borough** is required
- Trends in Serious Crime by **Correlation Analysis with Murder Cases**

Further in-depth analysis is required

7. Summary

- Brooklyn has the highest number of gun crimes in New York City by borough. Bronx is the second worst.

-> **Travelers should be careful in these areas**

- Staten Island is safe with the fewest shooting incidents
- The number of crimes is low in the morning, with a minimum at 9:00 a.m., increasing in the evening and peaking at 11:00 p.m.

-> **Travelers should avoid unnecessary outings late at night**

The end of my presentation