Table of Contents

Business Understanding & Objective

The analysis is to compare phenomenon related to parking violation tickets in New York City over three different years - **2015, 2016, 2017**

Assumptions for Data Cleaning and Data Analysis

1. **Column Names**
   - There are spaces in column names, we are replacing those with underscore.
2. **Removal of Columns**
   Removing these fields during EDA from 2015 and 2016 dataset , as these fields have only null values . These fields are not present in 2017 dataset.
   - Latitude
   - Longitude
   - Community_Board
   - Community_Council_
   - Census_Tract
   - BIN
   - BBL
   - NTA
3. **New York City Precincts & Parking Violation Codes**
   - Precinct '0' is treated as Valid
   - For violation codes fine amounts, reference link -
     http://www1.nyc.gov/site/finance/vehicles/services-violation-codes.page
4. **New City Seasons**
   - Categorizes as 4 seasons Summer, Winter, Spring & Fall
   - Reference link used - https://www.nyc.com/visitor_guide/weather_facts.75835
5. **Fiscal Year Wise Tickets data**
   We have filtered the datasets and considered based on fiscal year from each dataset. We have decided fiscal year based on this link https://www.kaggle.com/new-york-city/nyc-parking-tickets
   - 2015 dataset has data from July 2014 to June 2015
   - 2016 dataset has data from July 2015 to June 2016
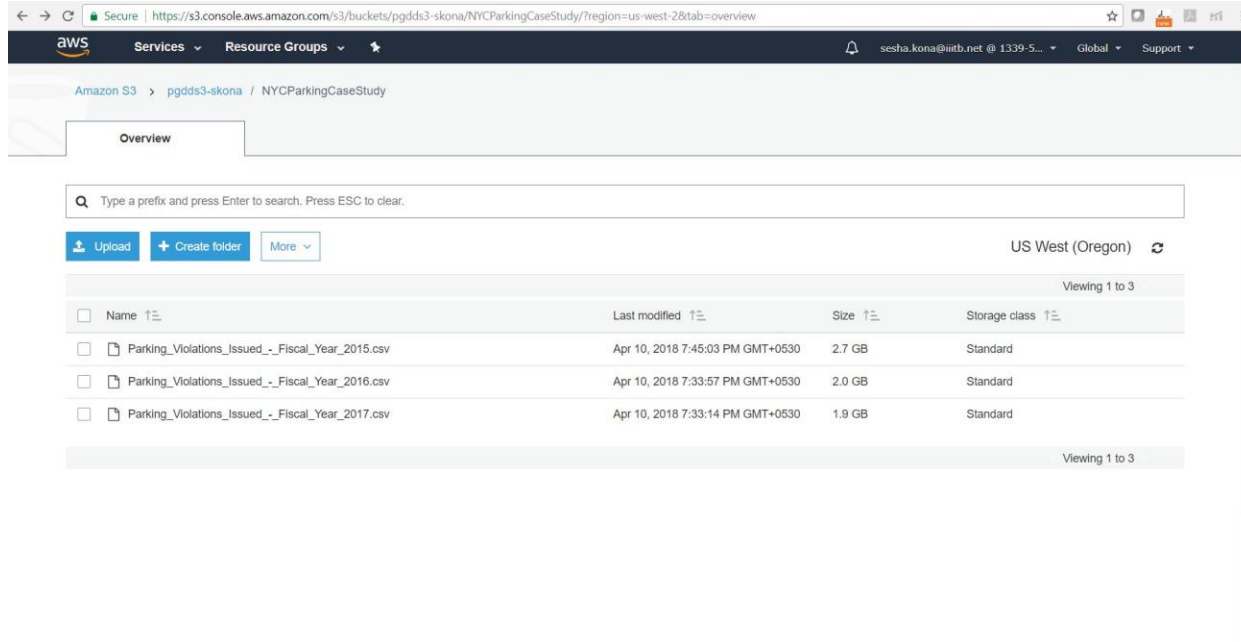   - 2017 dataset has data from July 2016 to June 2017
6. **Data Analysis - Vehicle Registration Codes**
   - We have not included the registration state with code '99' as it's being a invalid state
   - There are 50 states in USA, but in all the three datasets have few additional state codes, which are assumed as those trips are done by vehicles which has registration state outside of USA.

Screenshots of S3 buckets of Case Study Group Members

Sailendra Kona

Rituraj Achuthan

Swati Bansal

Pradosh Kumar Jena

Q1 Data Analysis - How often does each violation code occur? (frequency of violation codes - find the top 5)

Inference(s)

- **Violation Code 21** has the highest frequency of tickets across all the 3 years
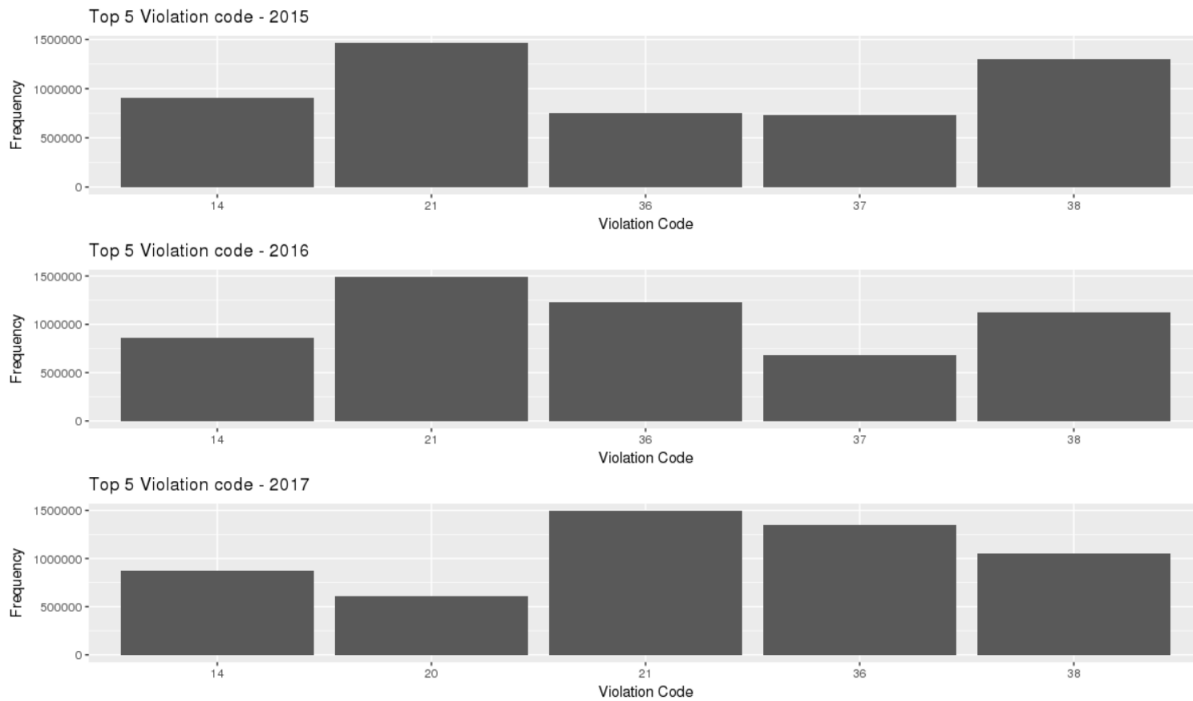
Q2.1 Data Analysis - How often does each vehicle body type get a parking ticket?

Inference(s)

- **Vehicle Body Type SUBN** receives the most parking tickets across all 3 years

Top 5 Vehicle Body Type - 2015
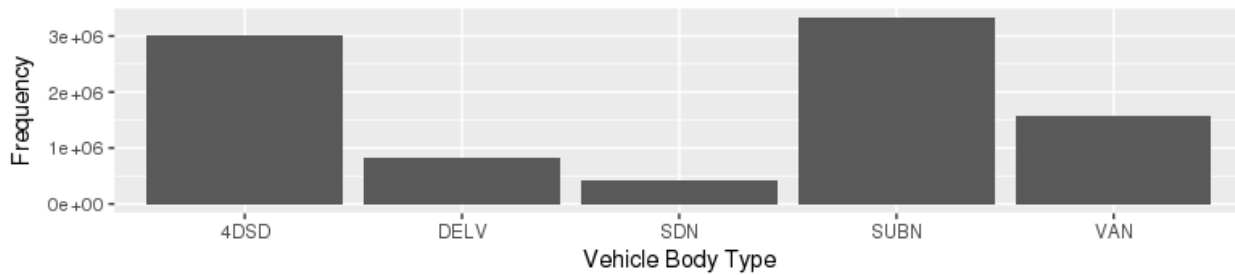
Top 5 Vehicle Body Type - 2016
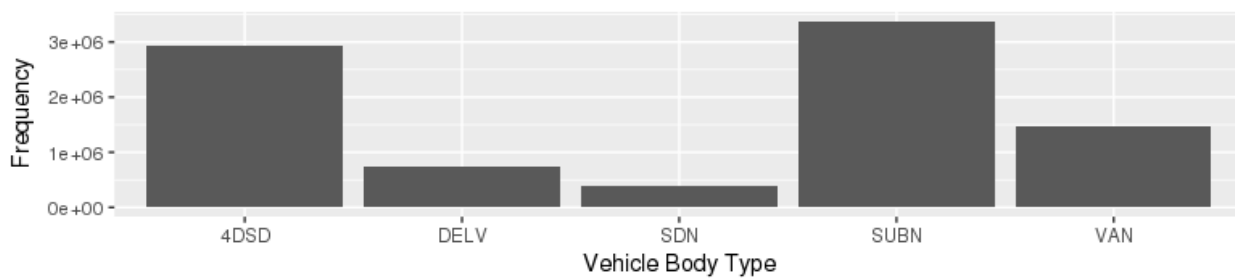
Top 5 Vehicle Body Type - 2017
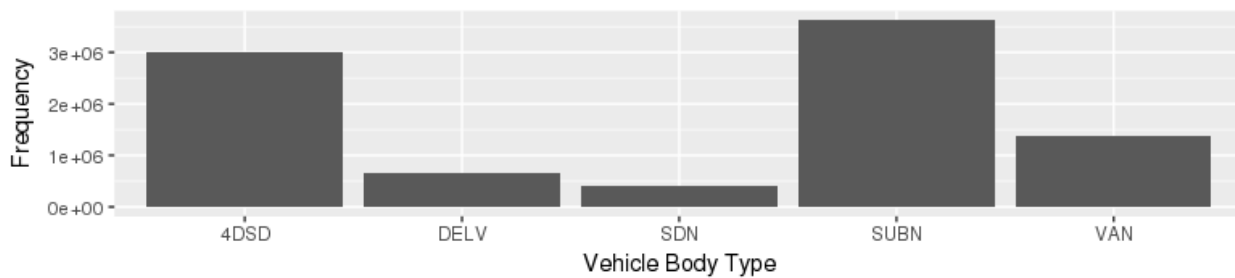
Q2.2 Data Analysis - How about the vehicle make? (find the top 5 for both)

Inference(s)

- **Vehicle Make FORD** has the most parking tickets across all 3 years
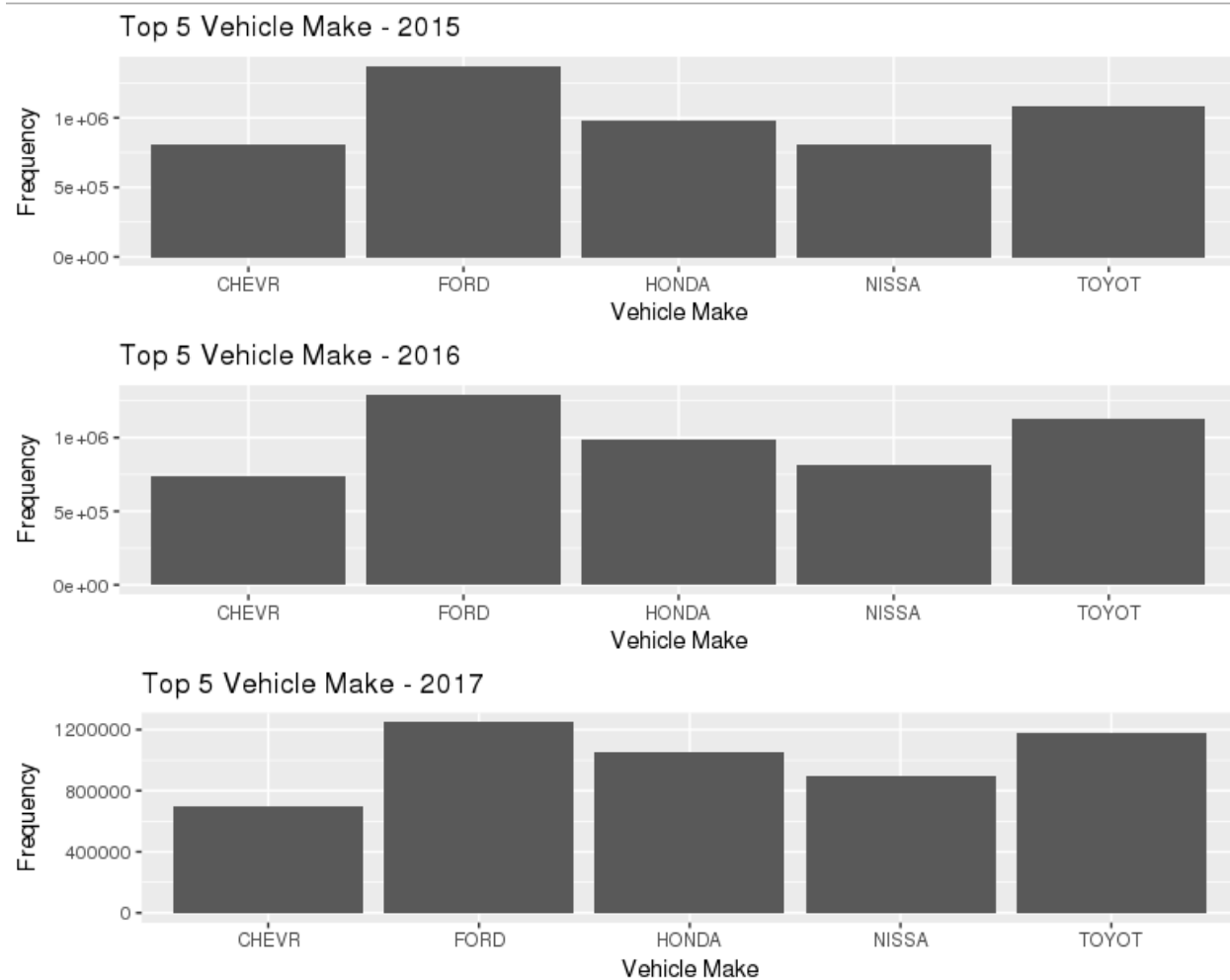
Q3.1 Data Analysis - Precincts (violation occurred) wise Violations Frequency

Inference(s)

- **Precinct 0** has highest violations across all 3 years
- **Precincts 19 and 18** have next highest violations
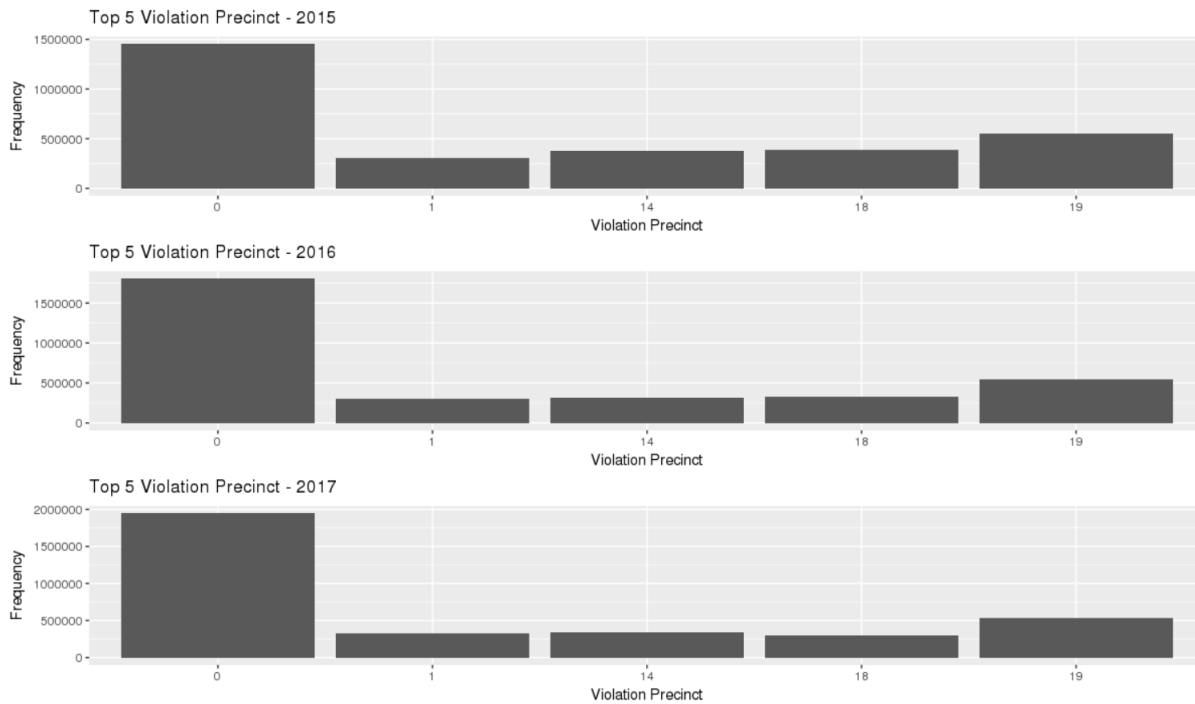
Q3.2 Data Analysis - Issuing Precincts (this is the precinct that issued the ticket)

Inference(s)

- **Precinct 0** has highest issues tickets for parking violations
- **Precincts 19 and 18** have next highest violations
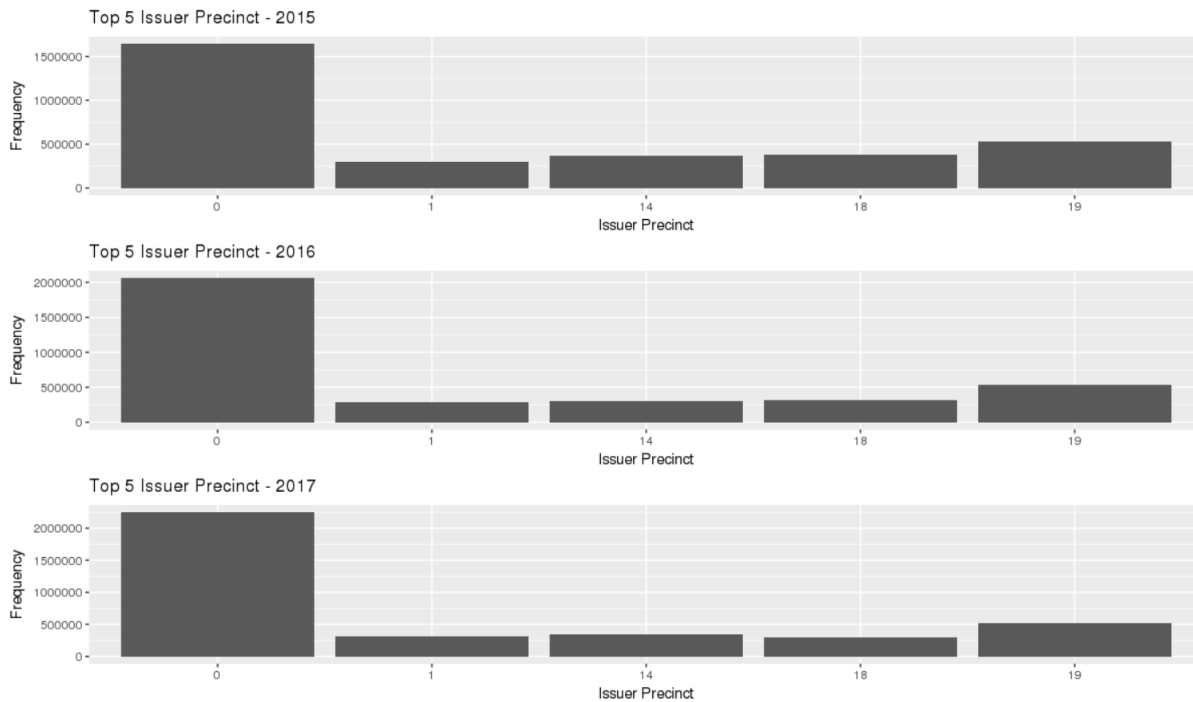
Q4 Data Analysis - Violation Code frequency across 3 precincts (issued the most number of tickets) - do these precinct zones have an exceptionally high frequency of certain violation codes? Are these codes common across precincts?

Inferences
- **Top 3 Precincts** – **0, 18 and 19**
- High Frequency Violation Codes – **36, 7 and 21**
- Common Violation Codes across precincts
  - **14 and 38** common to **Precincts 18 and 19**
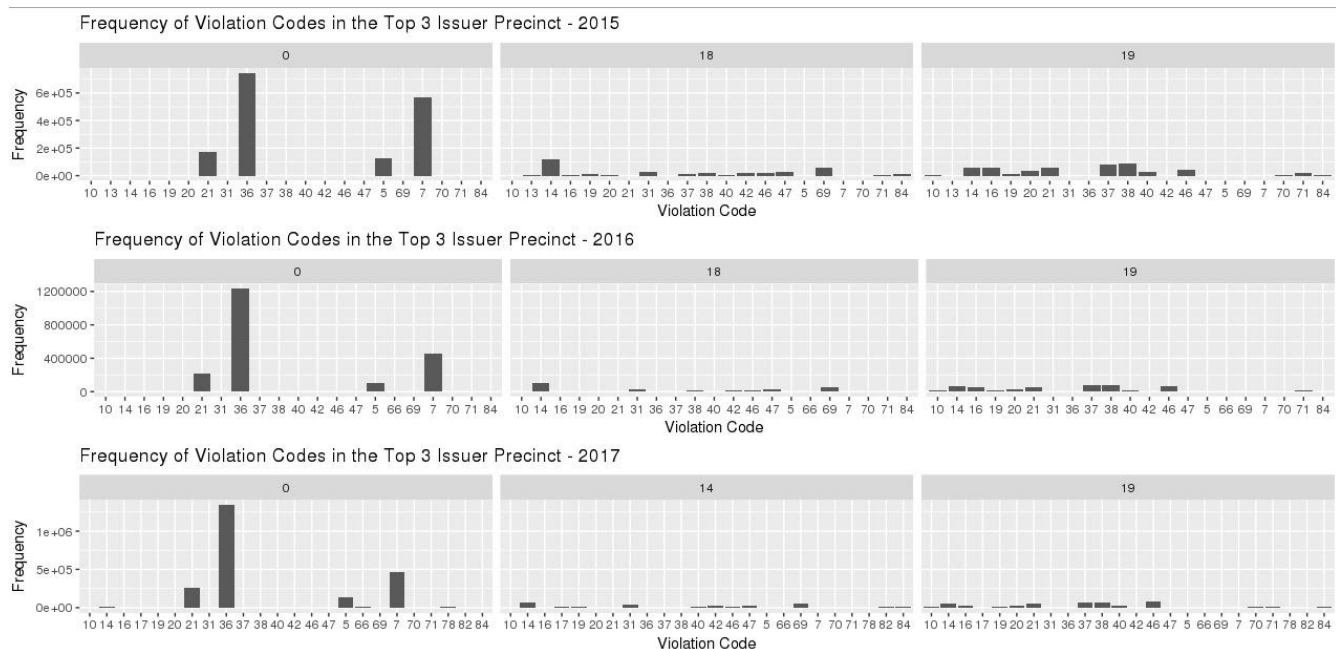
Q5.1 Data Analysis - Top3 most commonly occurring violations Parking Violations across different times of the day (24 hours groups)

Inferences

- Violation trend has been similar for all three years for *Early-Morning, Morning, Afternoon, Evening, Night* and *Mignight-4AM*
- *Evenings* which has minimum Violations
- *Nights* have moderate violations

Violation Code Frequency for Top3 most commonly occurring violations - Time Slots of The Day
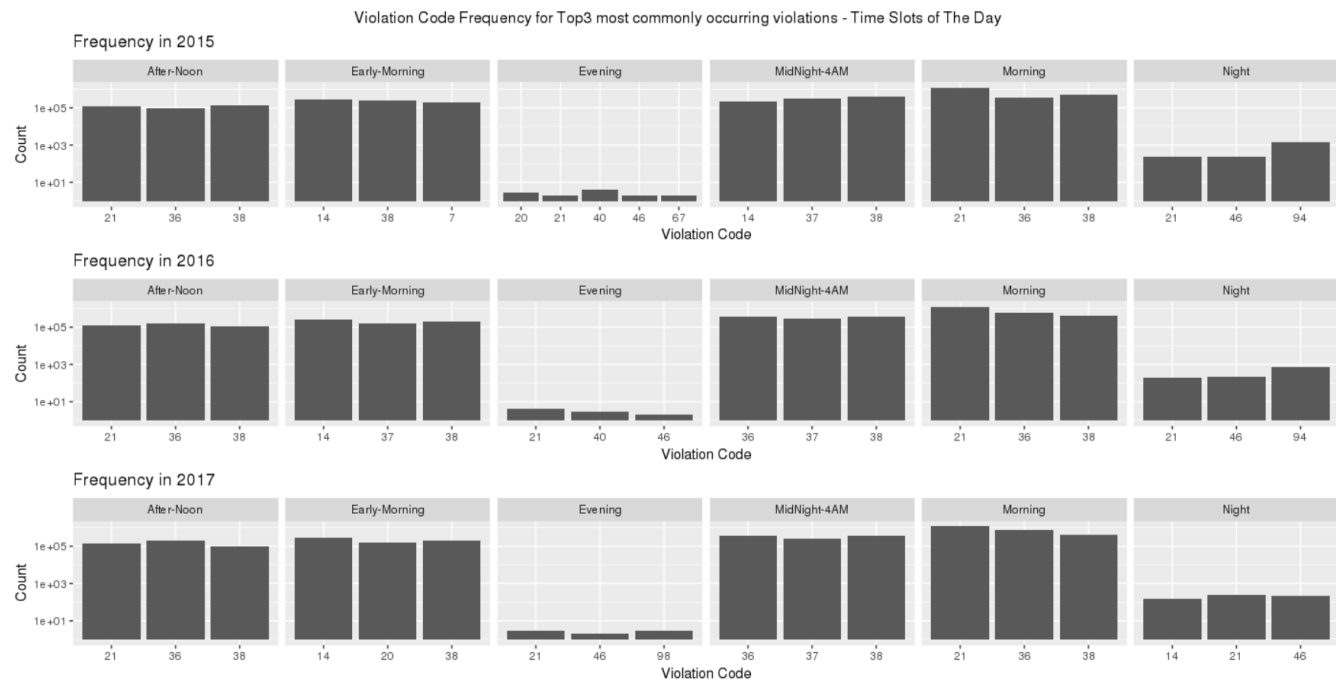
Q5.2 Data Analysis - Parking Violations Pattern across different times of the day - 3 most commonly occurring violation codes and most common times of day

Inferences

- For all three years, the most common violations are **14, 21 and 38.**
- As we saw in the previous graph, ***Evening*** has the least violations.
- The **Violation Code 21** occurs mostly during the ***Mornings***
- **Violation Code 38** which is increases during ***Midnight-4AM*** and ***Morning***
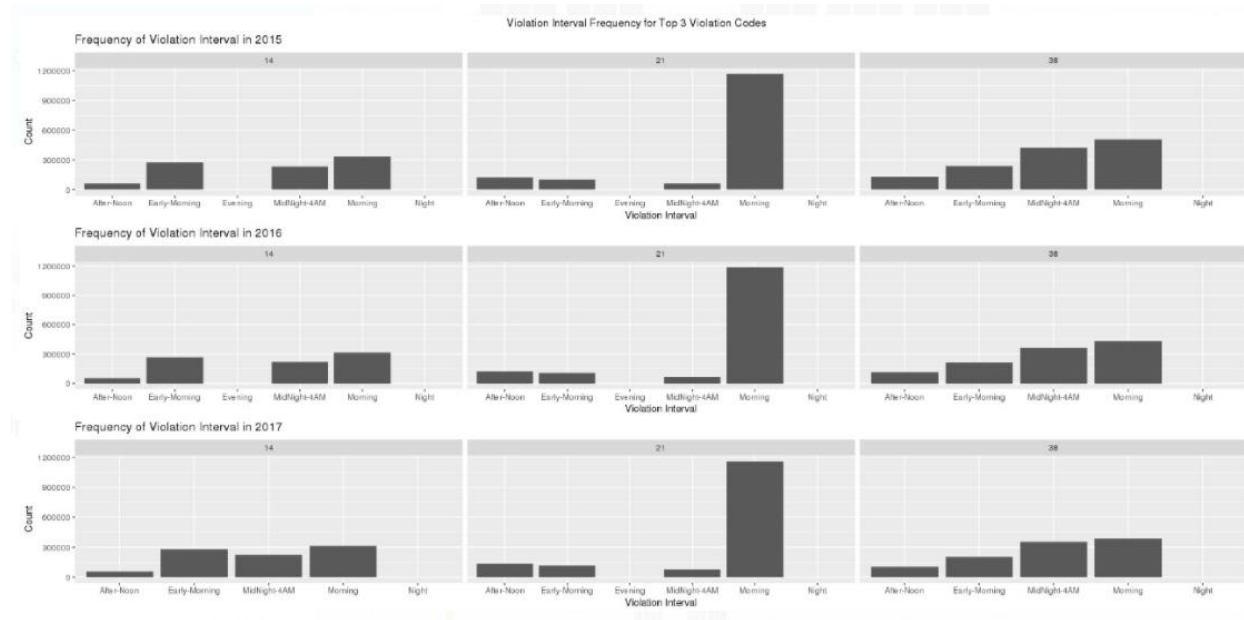
Violation Interval Frequency for Top 3 Violation Codes

Q6 Data Analysis - Frequencies of tickets for each season and Top 3 most common violations for each of these seasons

Inferences for Year 2015 – Season wise Top 3 Violation Codes

Violation codes 21, 38 and 14 high across all seasons

- Fall Season - Violation Codes 21, 38, 14
- Spring Season – 21, 38, 14
- Summer Season – 21, 38, 14
- Winter Season – 38, 21, 14

Frequency of Violation Codes in different Seasons of the Year 2015
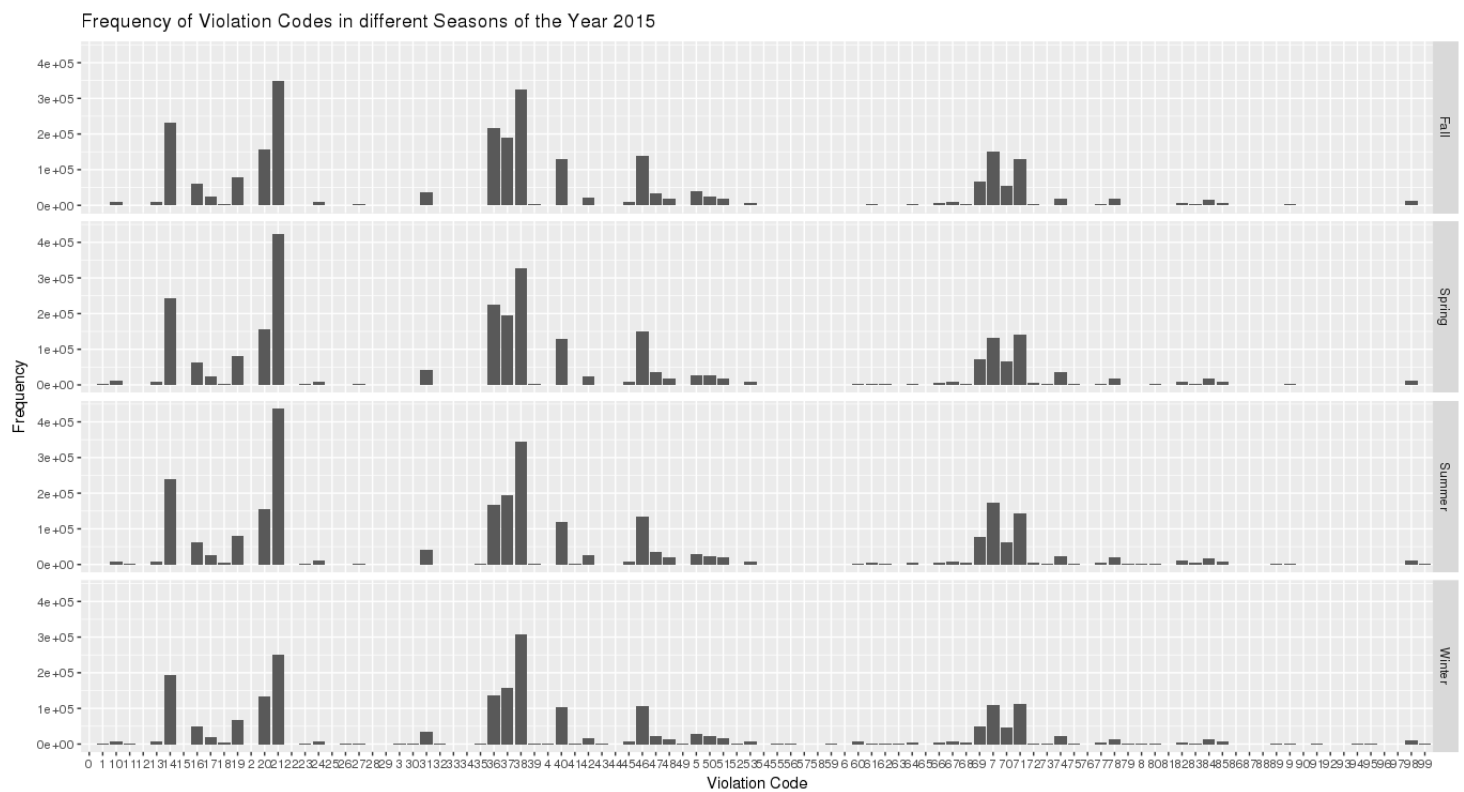
Inferences for Year 2016 – Season wise Top 3 Violation Codes

Violation codes 21 and 38 high across all seasons

- Fall Season - Violation Codes 21, 38, 14
- Spring Season – 21, 36, 38
- Summer Season – 21, 38, 14
- Winter Season – 21, 36, 14

Frequency of Violation Codes in different Seasons of the Year 2016

Inferences for Year 2017 – Season wise Top 3 Violation Codes

Violation codes 21 and 38 high across all seasons

- Fall Season - 36, 21, 38
- Spring Season – 21, 36, 38
- Summer Season – 21, 38, 14
- Winter Season – 21, 36, 38

Frequency of Violation Codes in different Seasons of the Year 2017

Q7.1 Data Analysis - Total occurrences of the 3 most common violation codes

Inferences

- Violation Code 21 has been consistently Top most violation across all years
- Violation Code 38 has been consistently in Top 3 most violation across all years
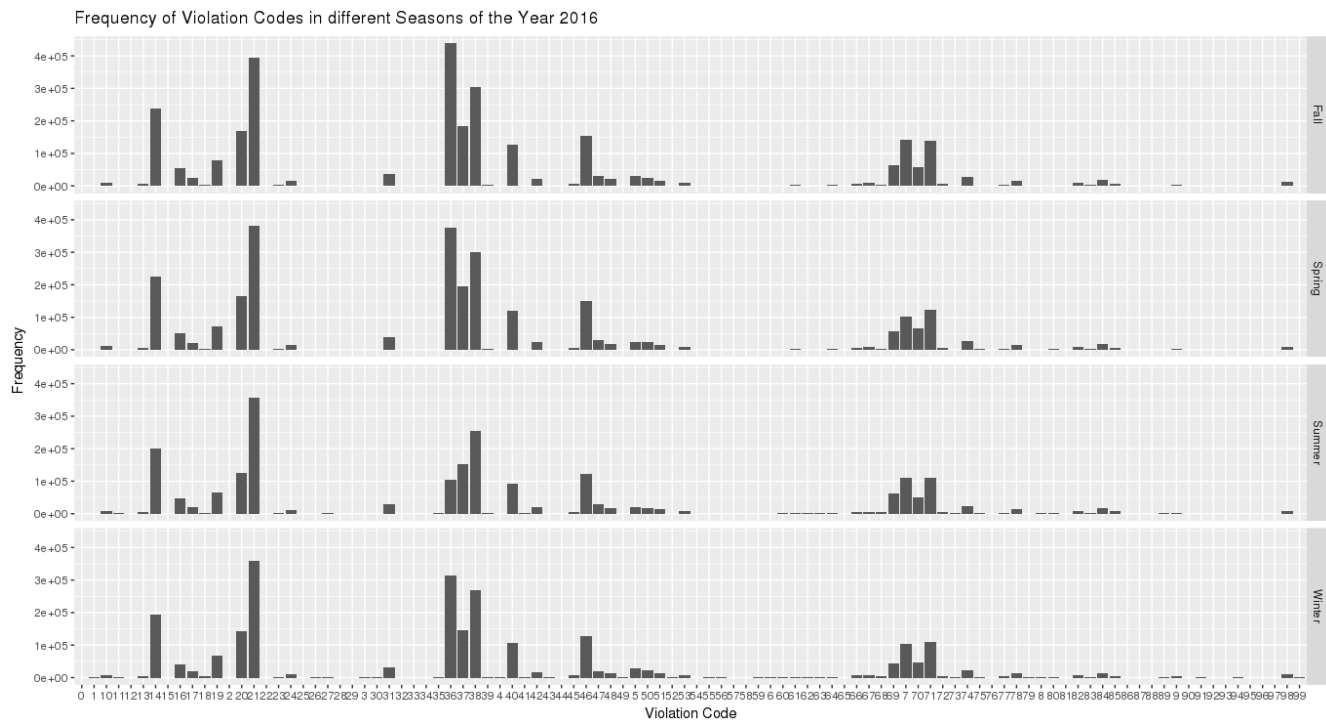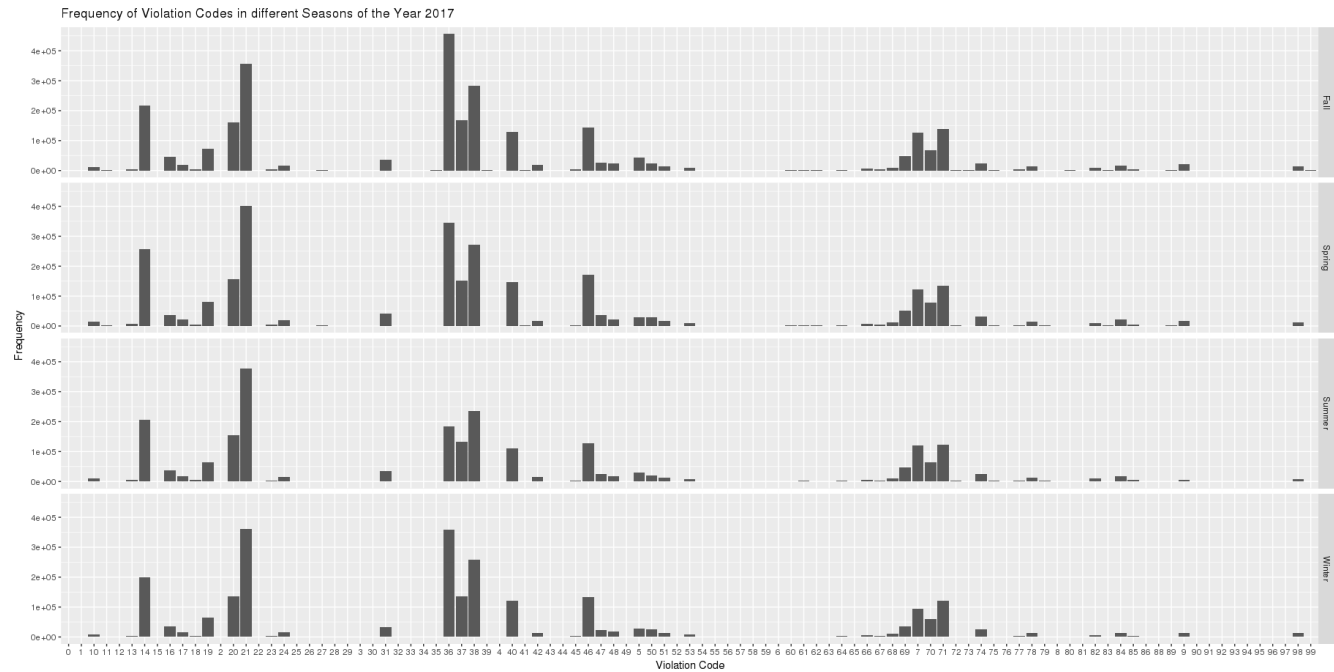- Violation Code 36 has been Top 2 in 2016 and 2017

Year – 2015

Violation_Code	count
21	1464127
38	1304009
14	905715

Year – 2016

Violation_Code	count
21	1490775
36	1232910
38	1125950

Year - 2017

Violation_Code	count
21	1494775
36	1345192
38	1049457

Q7.2 Data Analysis - Total Fine Amount Collected as Revenue for NYC police department

Inferences – Total Revenue of $ 651, 765, 360 (651.76 Million) collected over 3 Years

Year wise breakup

- Total Fine collected for **Year 2015** = USD 249,884,660
- Total Fine collected for **Year 2016** = USD 199,935,625
- Total Fine collected for **Year 2017** = USD 201,945,075

Q7.3 Data Analysis - What can you intuitively infer from these findings. ?

Inferences – Year wise Number of Tickets and % of Revenue

- **Year 2015**
  - **24.65%** of Tickets from **Violation Code 21**
  - **52.27%** Revenue generated for NYC police department
- **Year 2016**
  - **38.73%** of Tickets from **Violation Code 21**
  - **35.48%** Revenue generated for NYC police department
- **Year 2017**
  - **38.43%** of Tickets from **Violation Code 21**
  - **35.48%** Revenue generated for NYC police department