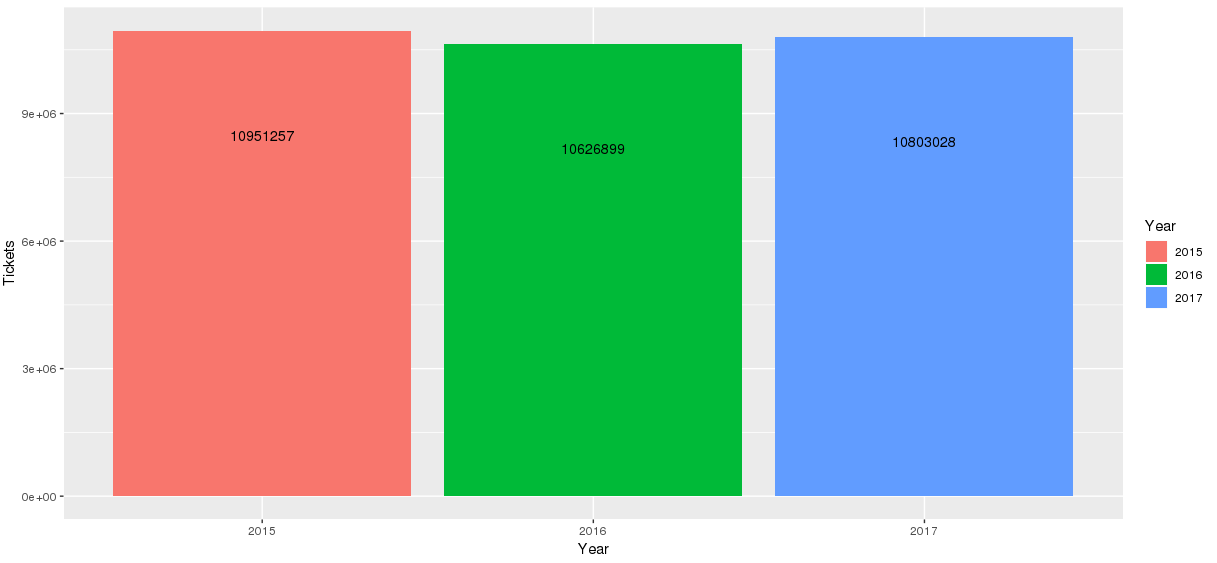
**NYC Parking Tickets - Analysis**

**Following details the approach and results of analyzing the NYC parking tickets for years 2015,2016 and 2017.**

**( Note : We have assumed Fiscal year for all our calculations.)**

1. Find the total number of tickets for each year.



**Approach:**

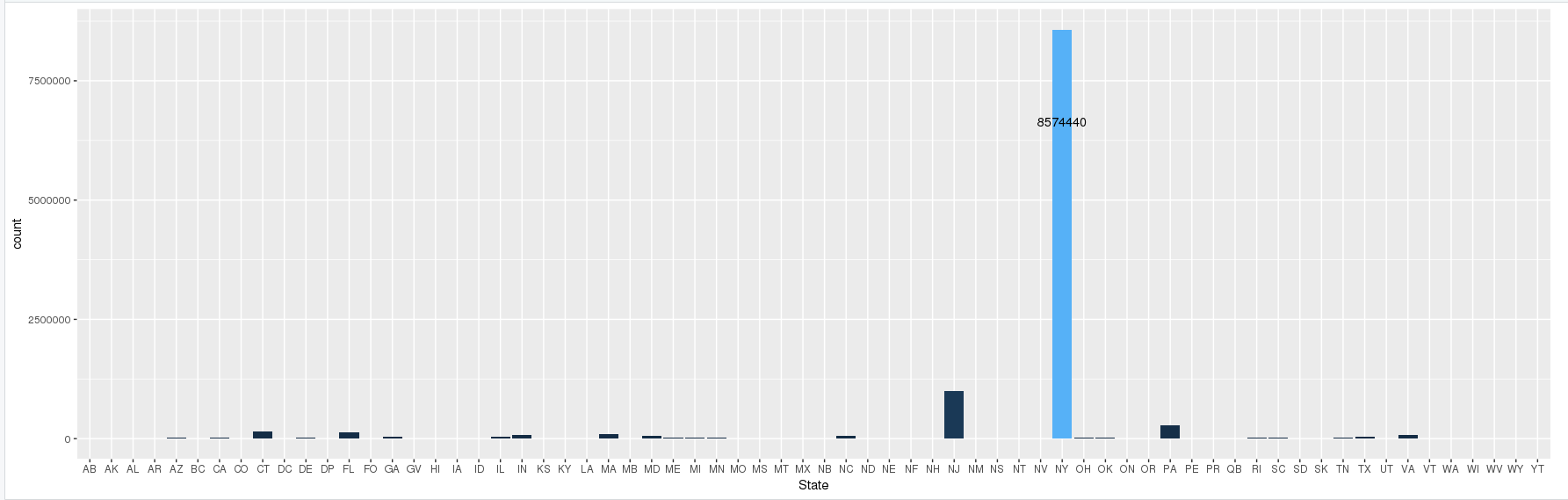
* Remove the duplicates from the data set
* Count the number of rows in the data set to get the number of tickets

**Analysis:**

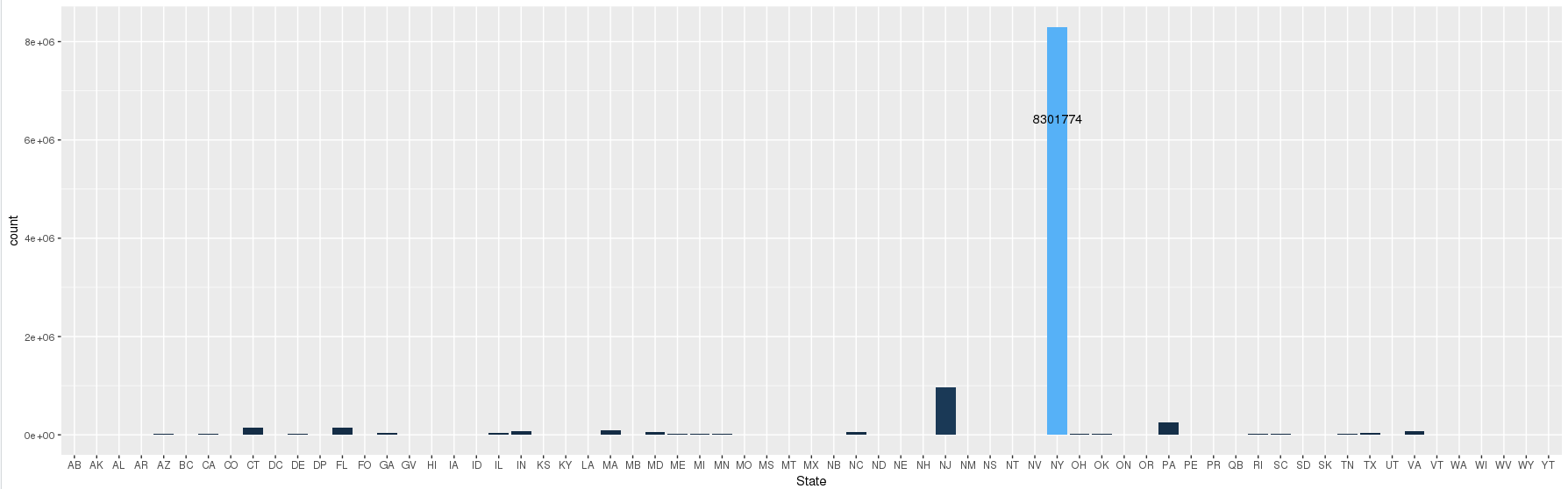
* 2015 has highest occurrences of parking tickets
* 2017 follows 2015 , while number of violations in 2016 is lowest.

2. Find out the number of unique states from where the cars that got parking tickets came from.

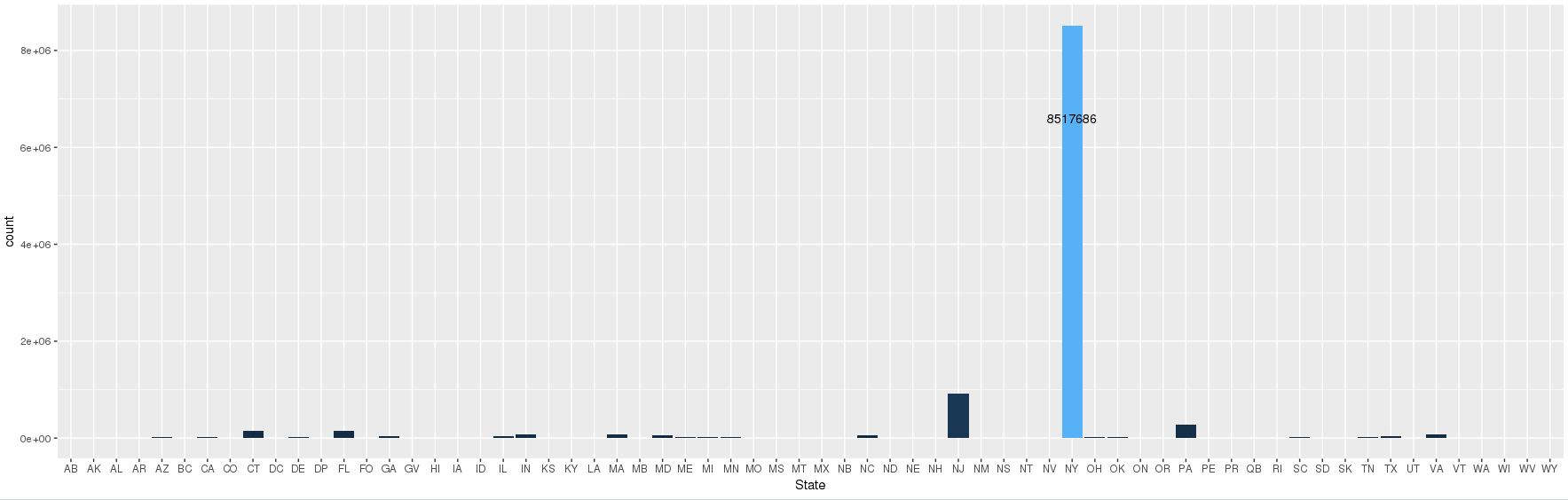
**2015**



**2016**



**2017**



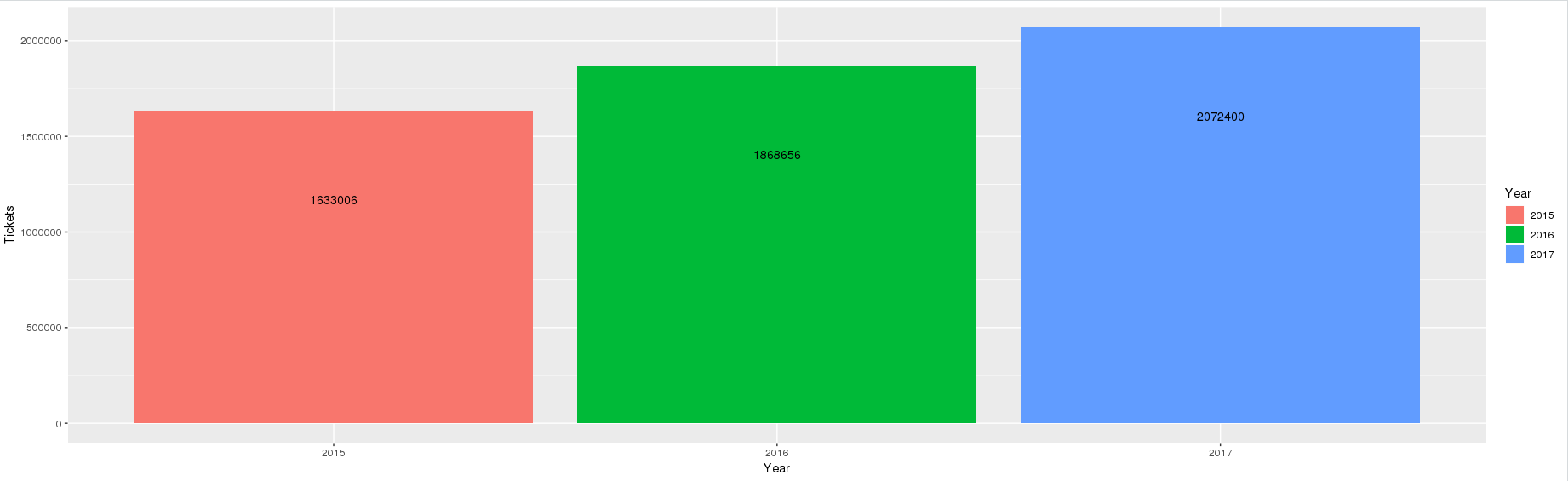
**Approach:**

* Group the data by **Registration State** and count the number of tickets in each grouping
* Arranging this data in descending order gives the results

**Analysis:**

* Number of tickets in each year belong to almost 66+ states .
* Each year mostly follws similar pattern in terms of occurrences of tickets in each state.
* Most number of tickets are from “**NY “**state.

3. Some parking tickets don't have the address for violation location on them, which is a cause for concern. Write a query to check the number of such tickets.



**Approach:**

* Check if Violation Location is Null and only select those which aren’t .
* Find the number of rows.

**Analysis:**

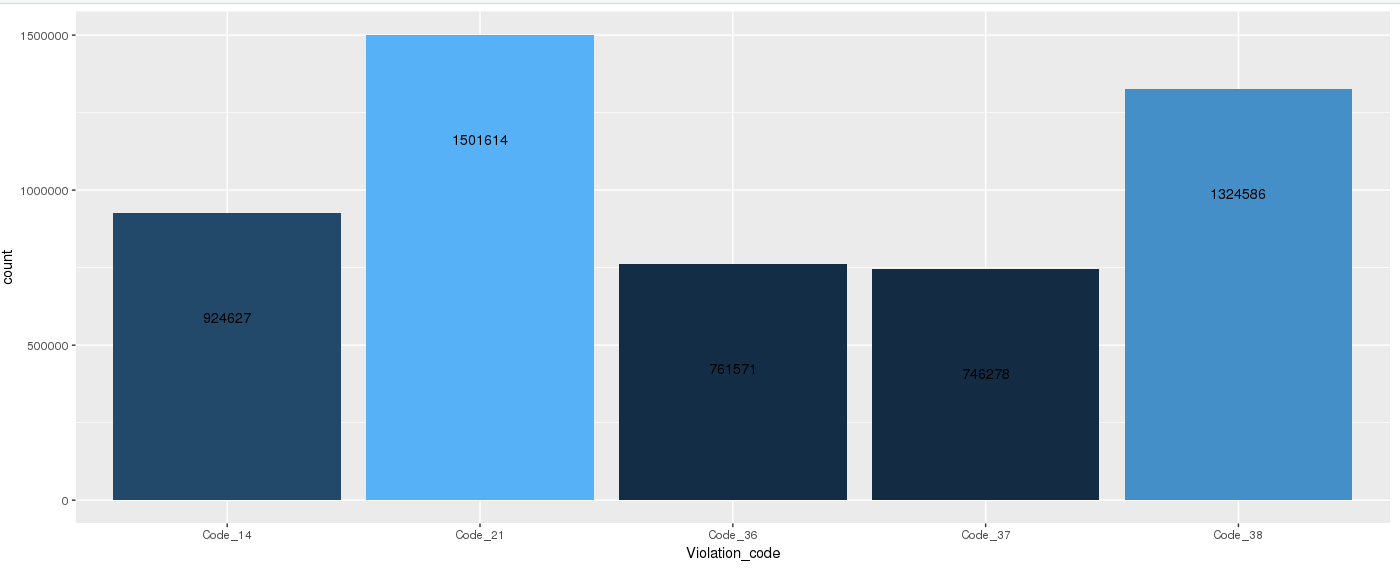
* From a statistical point of view, the most number of such cases where the parking tickets do not have violation location associated are from the year 2017
* 2016 ranks next , followed by year 2015.

**################# 4. Aggregation Tasks #############################################**

# 1. How often does each violation code occur? Display the frequency of the top five violation codes.

The violation codes indicates the type of violations indicated in each parking ticket. Among the different years 2015, 2016 and 2017, it can be different types of violations that can be prominent.

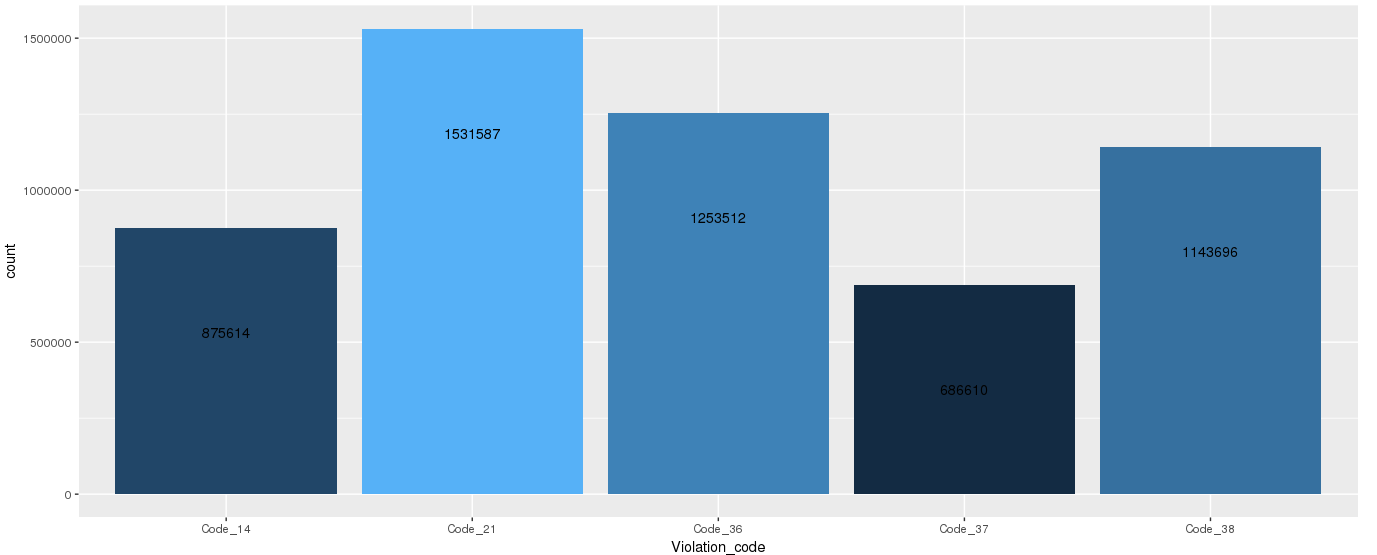
**2015**

****

**Analysis:**

* It is observed that the violations indicated by violation code 21 is highest in terms of count.
* This is followed up by 38, 14, 36 and 37 in the decreasing order of total count.

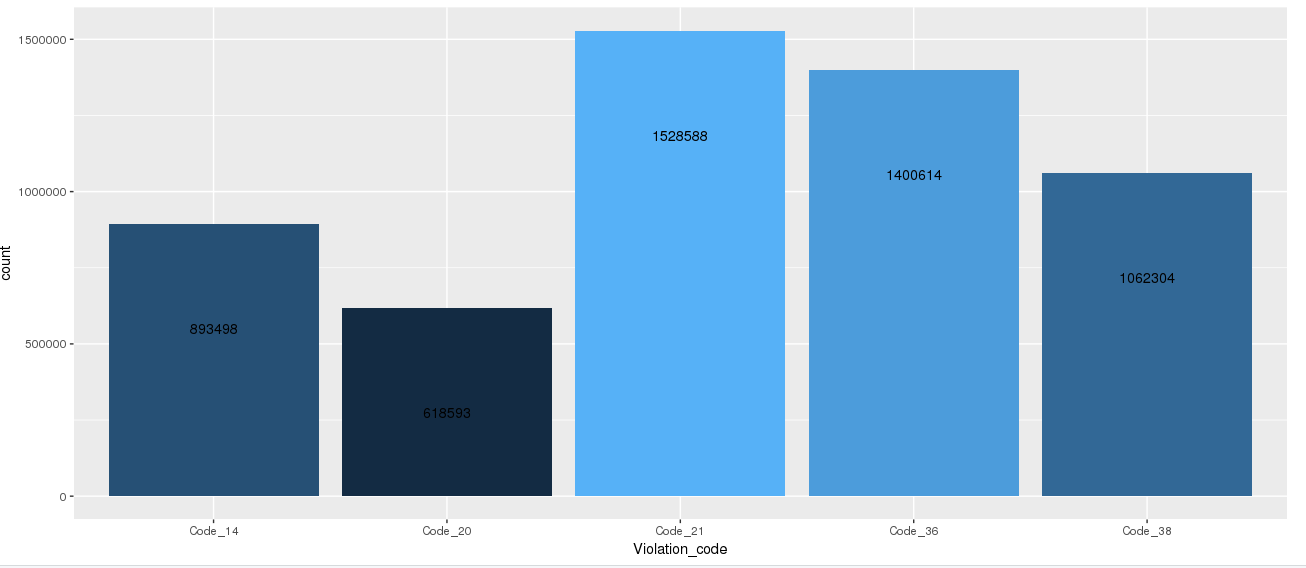
**2016**



**Analysis:**

* 2016 follows similar pattern.Violations indicated by violation code 21 is the highest in terms of count.
* This is followed up by 36, 38, 14 and 37 in the decreasing order of total count.

**2017**

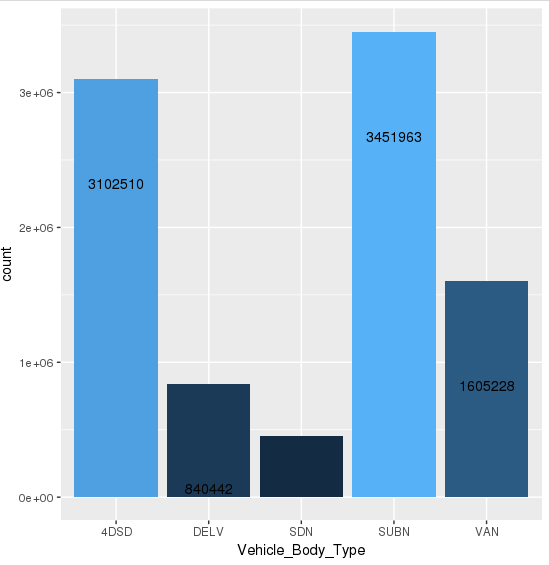
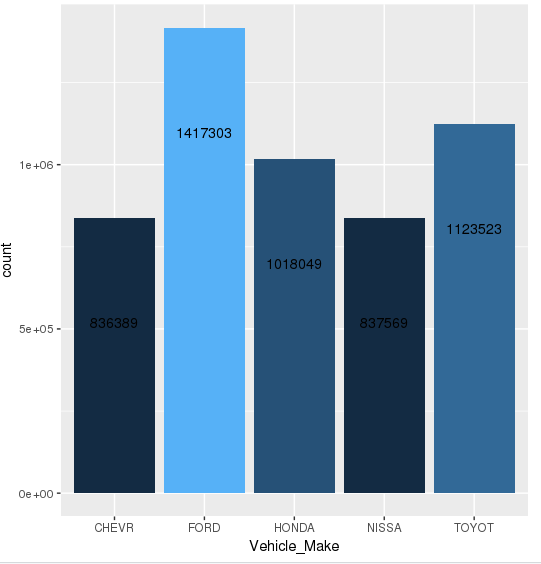


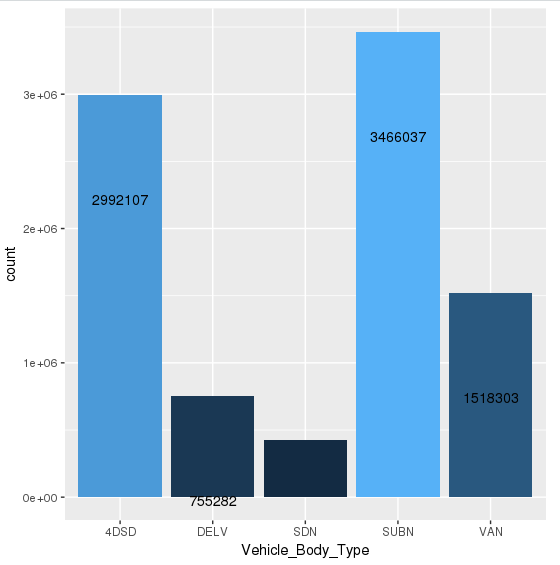
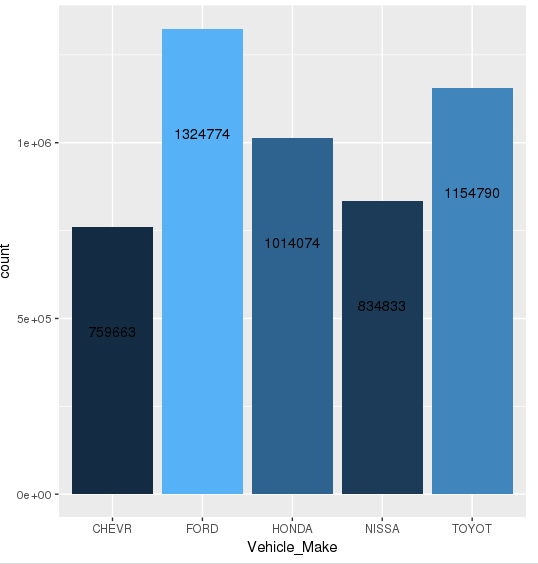
**Analysis:**

* In 2017 too violations indicated by violation code 21 is the highest in terms of count.
* his is followed up by 36, 38, 14 and 20 in the decreasing order of total count.

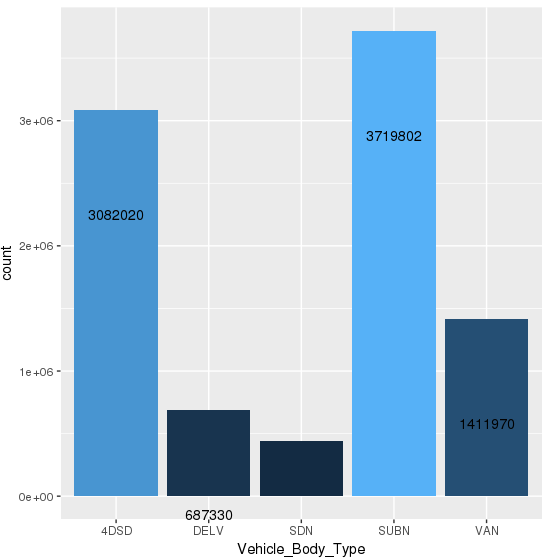
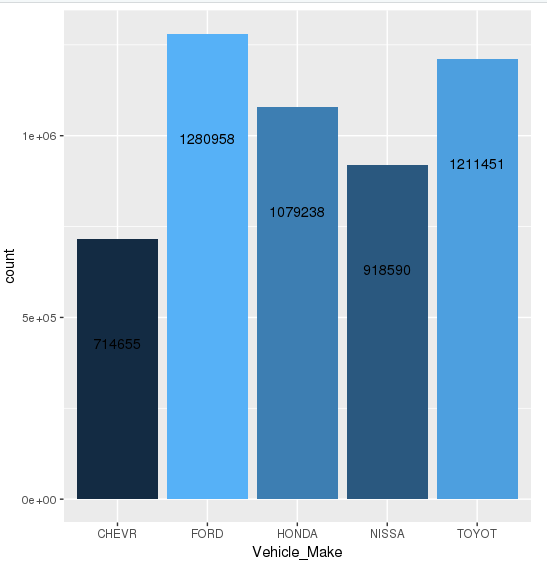
In general, across all the years, it is observed that the number of violations from the violation code 21 is the highest. And there is more frequency of violations from violation code 36, 38 and 14.

#2.How often does each 'vehicle body type' get a parking ticket? How about the 'vehicle make'? (Hint: find the top 5 for both)

**2015**

**2016**

**2017**



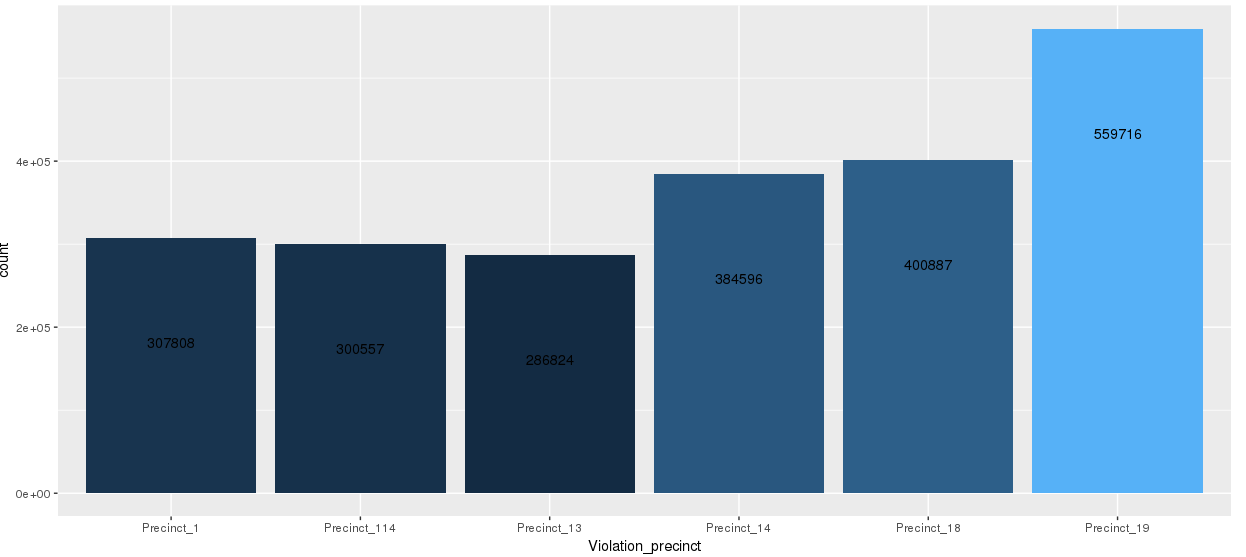
**Analysis:**

* We see similar patterns for various years match .
* Based on the vehicle body type analysis violations are the most from the vehicle body type ‘SUBN’
* 4DSD ranks next. The numbers for vehicle body type VAN, DELV and SDN follow.
* For the same year 2015, based on the vehicle make, FORD is involved in maximum number of parking violation tickets
* It’s followed by Toyota, HONDA, Chevrolet and Nissan in the decreasing order of counts.

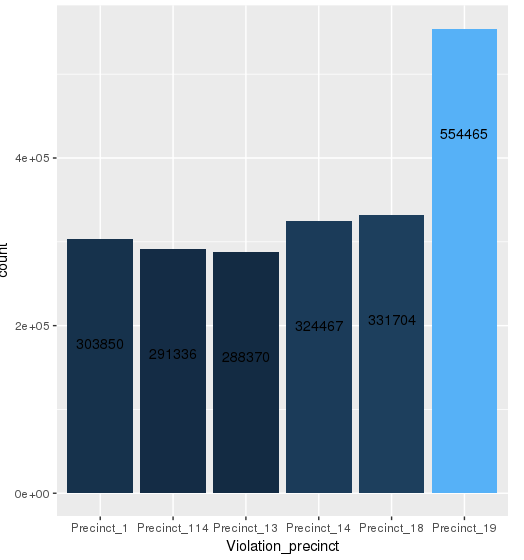
Find the (5 highest) frequency of tickets for each of the following:

# a. Violation Precinct

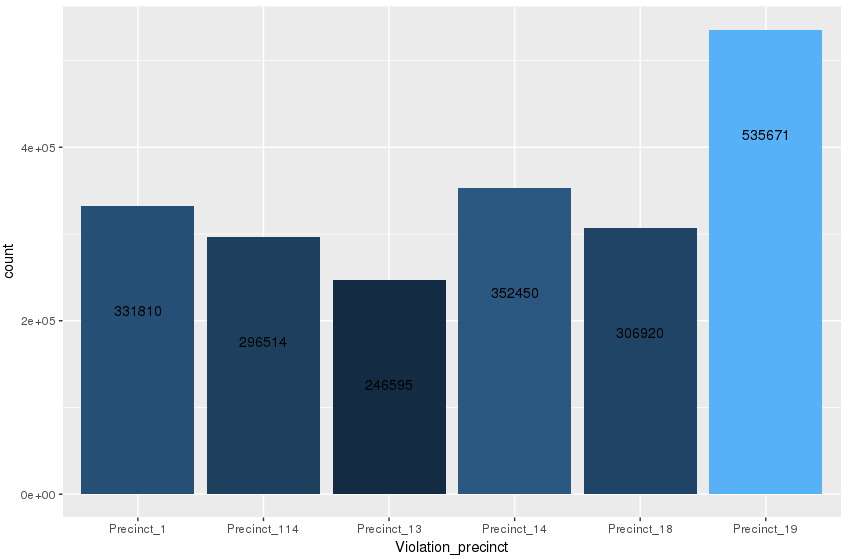
**2015**



**2016**



**2017**

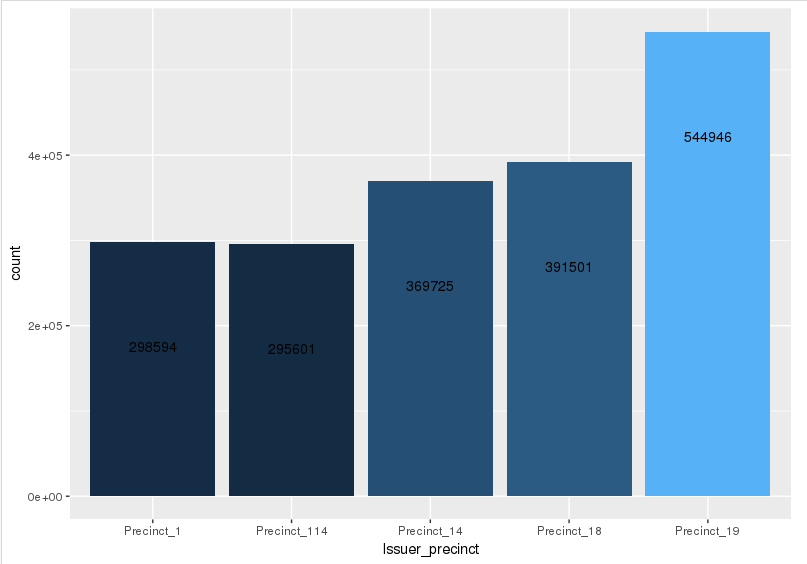


**Analysis :**

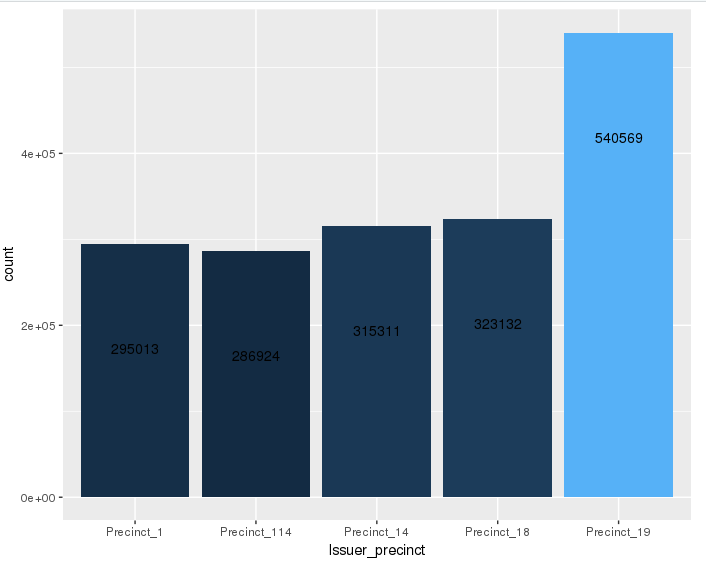
* Patterns for 2015 and 2016 are similar.
* Top 3 precicts for maximum number of violations are recorded from the violation precinct 19 , 18 and 14 in the decreasing order of violations counts.
* They are followed by Precinct 1, 114 and 13 in the decreasing order of parking violation counts.
* Overall violation count based on the violation precinct is the same in year 2015 and 2016.
* For year 2017 however, precinct 19 , 14 and 1 are the top3 precincts in the decreasing order of violation counts.
* They are followed by Precinct 18, 114 and 13 in the decreasing order of parking violation counts.
* However, the highest violation count are always obtained from the Precinct 19 across 3 years.

# a. Issuer Precinct

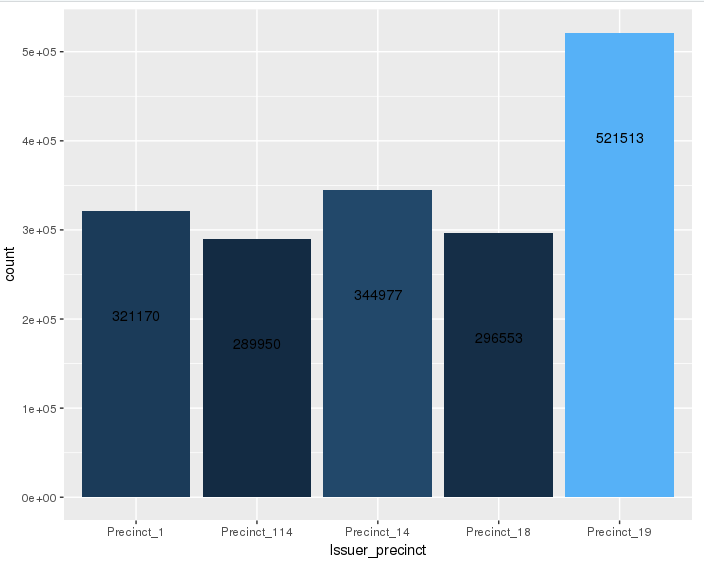
**2015**



**2016**



**2017**



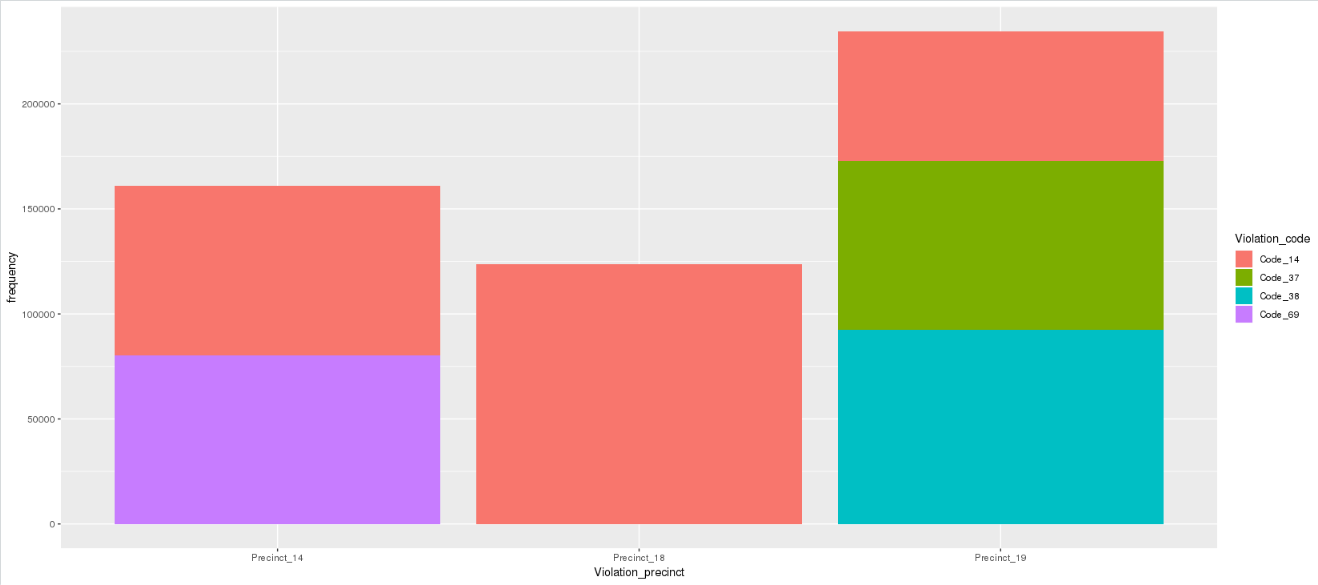
**Analysis:**

* 2015 and 2016 show similar patterns with , maximum number of violations are recorded from the issuer precinct 19 and closely followed by Precinct 18 and 14.
* These numbers are then followed by Precinct 1 and 114 in the decreasing order of parking violation counts.
* For year 2017, maximum number of violations are recorded from the issuer precinct 19 and closely followed by Precinct 14 and 1. These numbers are then followed by Precinct 18 and 114 in the decreasing order of parking violation counts.
* The top 3 precincts with maximum violations are same whether by issuer precinct or violation precinct.

#4. Find the violation code frequency across three precincts which have 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?

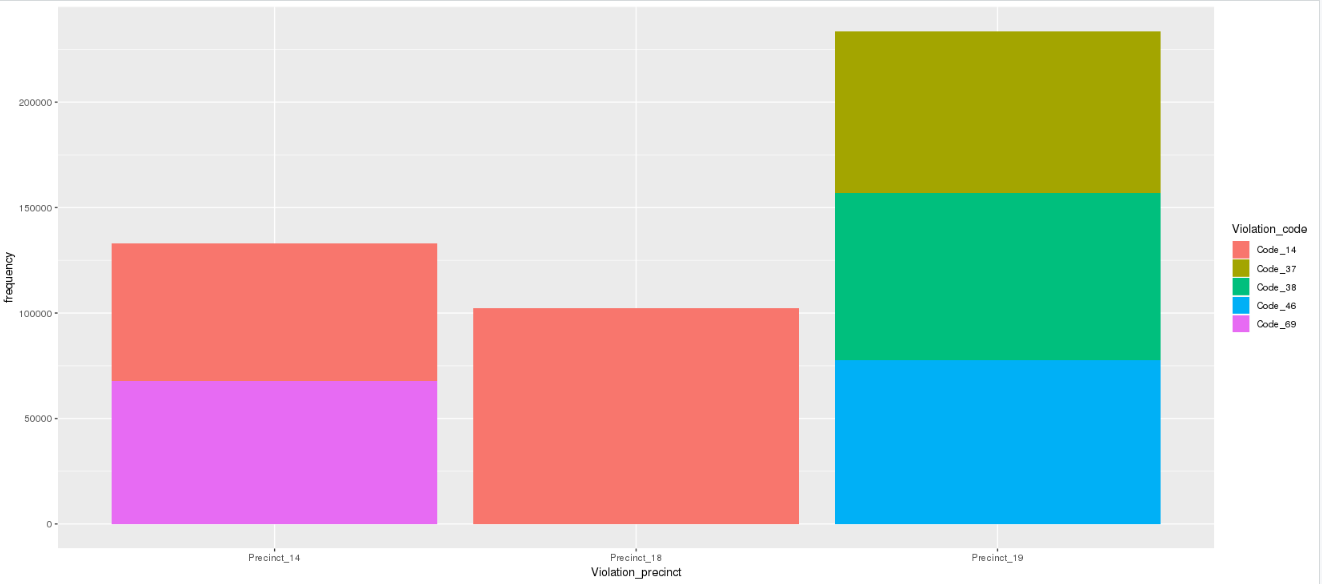
**2015**



**Analysis:**

* For the year 2015, the maximum number of violations have occurred in the precinct zones 19, 14 and 18 in the decreasing order of violations. Within these zones, the following is distribution of the type of violation based on the violation codes:
* Precinct 19 🡪 Violation code 38 closely followed by code 37 and then by code 14.
* Precinct 14 🡪 Violation code 69 followed by code 14.
* Precinct 18 🡪 Violation code 14 is the only type of violation.

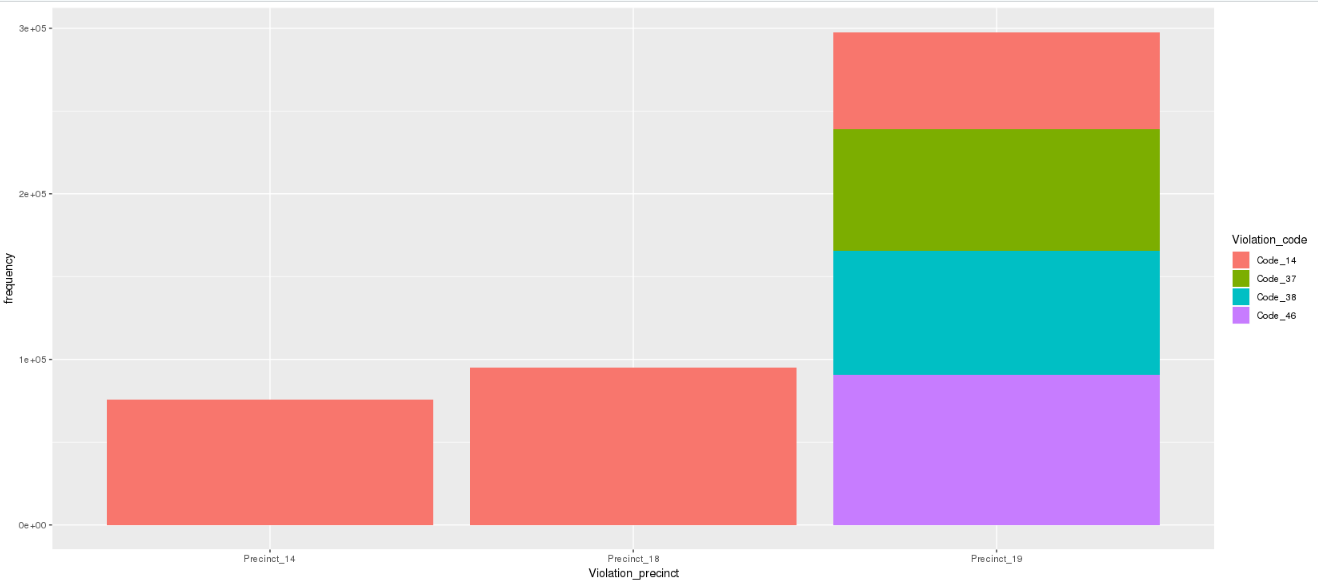
**2016**



**Analysis:**

* For the year 2016, the maximum number of violations have occurred in the precinct zones 19, 14 and 18 in the decreasing order of violations. Within these zones, the following is distribution of the type of violation based on the violation codes:
* Precinct 19 🡪 Violation code 46 closely followed by code 38 and then by code 37.
* Precinct 14 🡪 Violation code 69 followed by code 14.
* Precinct 18 🡪 Violation code 14 is the only type of violation.

**2017**



**Analysis:**

* For the year 2015, the maximum number of violations have occurred in the precinct zones 19, 18 and 14 in the decreasing order of violations. Within these zones, the following is distribution of the type of violation based on the violation codes.
* Precinct 19 🡪 Violation code 46 closely followed by code 38 and 37 and then by code 14.
* Precinct 18 🡪 Violation code 14 is the only type of violation.
* Precinct 14 🡪 Violation code 14 is the only type of violation.

Overall, among the precincts 19, 18 and 14, it is observed from data that there are commonly occurring violation codes. Example of such violation codes are 37, 38 and 14. These appear across the different years and across the different precincts.

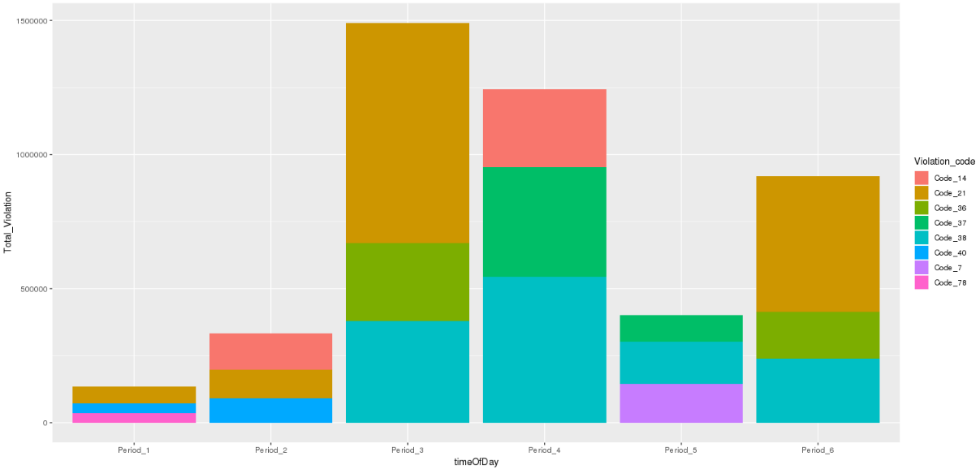
#5.find out the properties of parking violations across different times of the day

#Find the top 3 violation codes for each time of the day

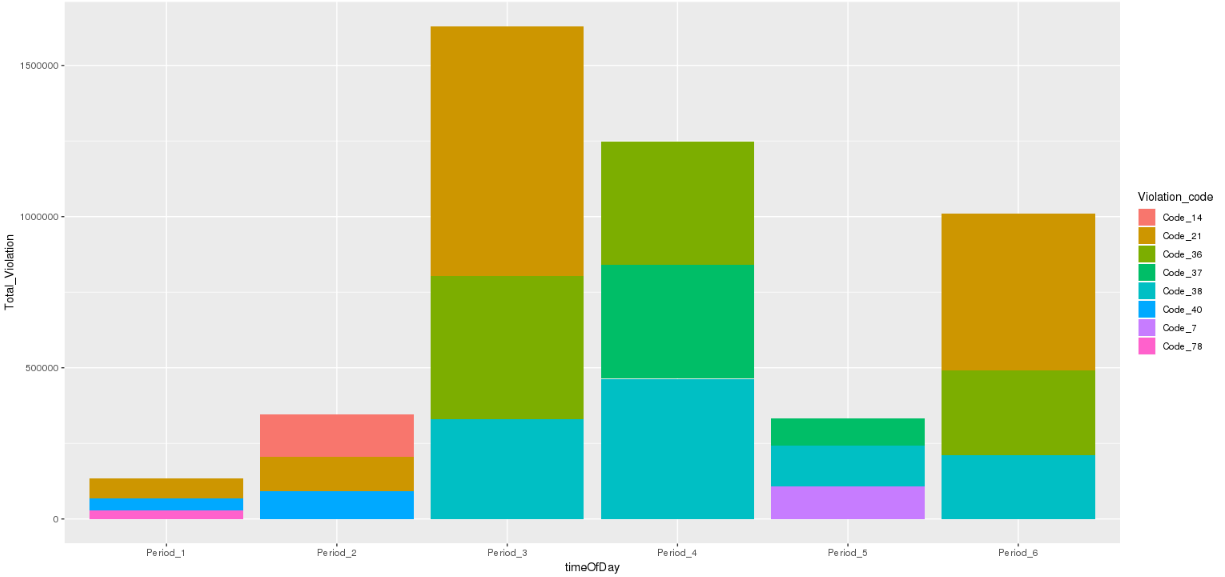
The entire day is split across 6 time zones as displayed below:

* Time zone 1 🡪 00 hrs and 03 hrs
* Time zone 2 🡪 03 hrs and 07 hrs
* Time zone 3 🡪 08 hrs and 12 hrs
* Time zone 4 🡪 12 hrs and 16 hrs
* Time zone 5 🡪 16 hrs and 20 hrs
* Time zone 6 🡪 20hrs and 00 hrs

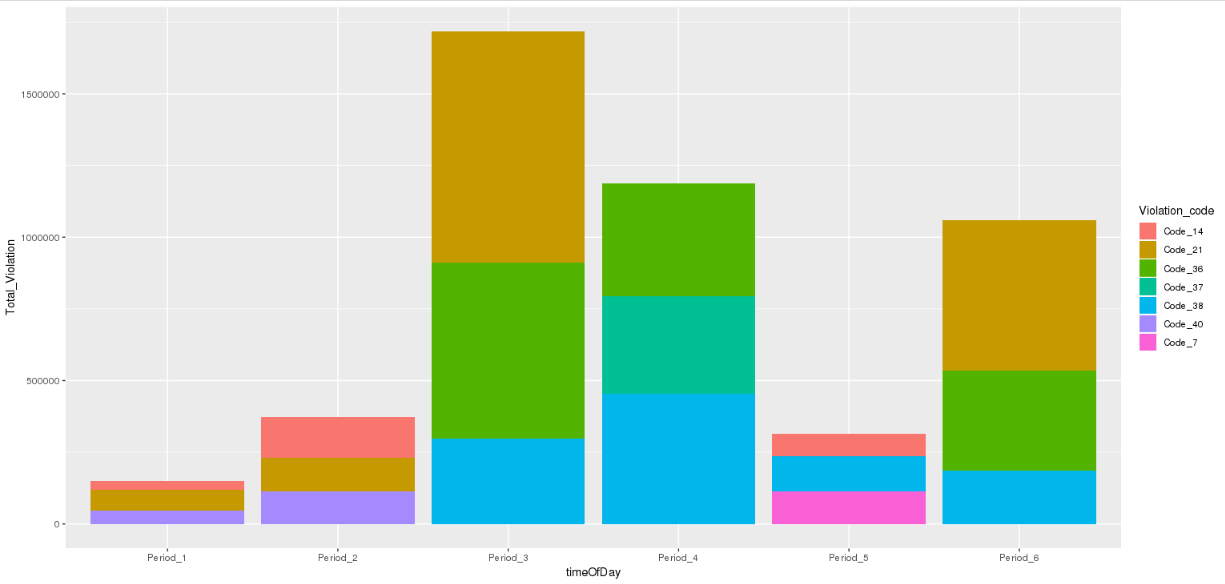
**2015**



**2016**



**2017**



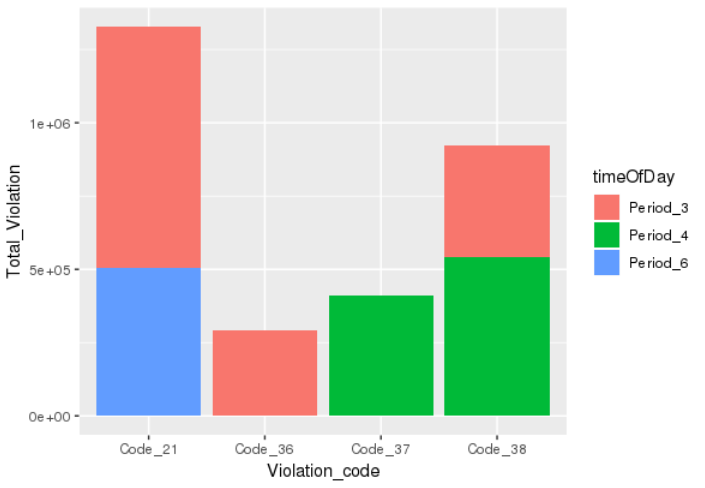
**Analysis:**

* In year 2015, it is observed that the maximum parking violations are in the time zone 3, i.e. between 08 hrs and 12 hrs and for the violation code 21.
* In year 2016, it is observed that the maximum parking violations are in the time zone 3, i.e. between 08 hrs and 12 hrs and for the violation code 21.
* In year 2017, it is observed that the maximum parking violations are in the time zone 3, i.e. between 08 hrs and 12 hrs and for the violation code 21.

Hence, the pattern of maximum tickets is same across all the different years.

#Find time of the day for top three violation codes

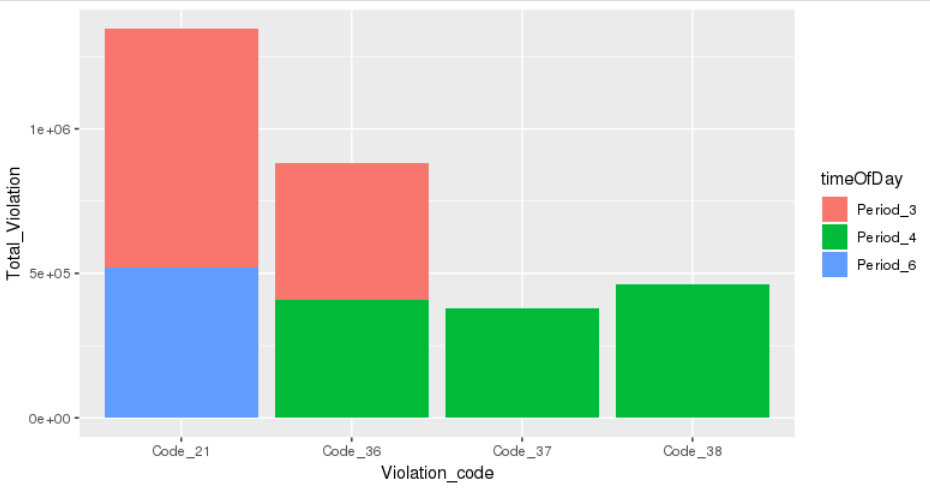
**2015**



**Analysis:**

* For year 2015, the top 3 violation codes in terms of the revenue are 21, 38 and 37. The time zones in which these appear are as shown in the plot above.:
* Code 21 🡪 Appear in period 3 and period 6.
* Code 38 🡪 Appear in period 4 and period 3.
* Code 37 🡪 Appear only in period 4.

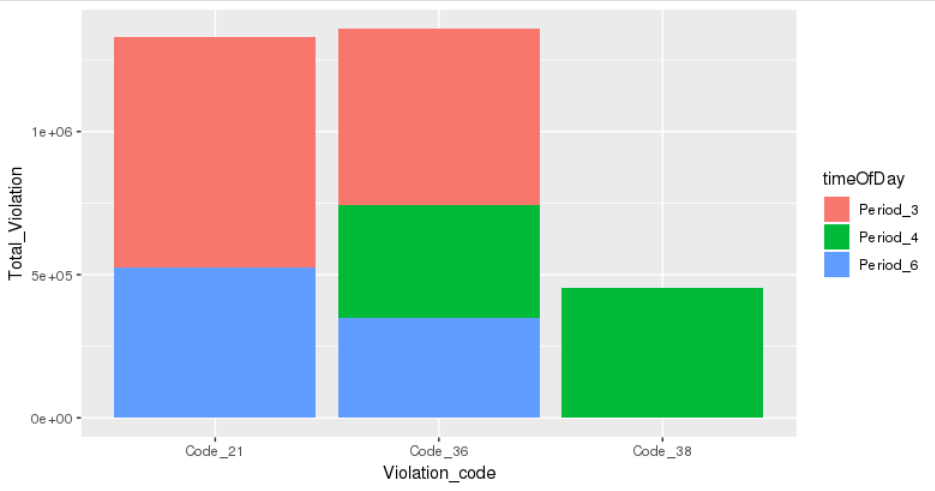
**2016**



**Analysis:**

* For year 2016, the top 3 violation codes in terms of the revenue are 21, 36 and 38. The time zones in which these appear are as shown in the plot above.:
* Code 21 🡪 Appear in period 6 and period 3.
* Code 36 🡪 Appear in period 4 and period 3.
* Code 38 🡪 Appear only in period 4.

**2017**



**Analysis:**

* For year 2017, the top 3 violation codes in terms of the revenue are 36, 21 and 38. The time zones in which these appear are as shown in the plot above.:
* Code 36 🡪 Appear in period 6, period 4 and period 3.
* Code 21 🡪 Appear in period 6 and period 3.
* Code 38 🡪 Appear only in period 4.

# 6. Let’s try and find some seasonality in this data

# 1. First, divide the year into some number of seasons, and find frequencies of tickets for each season. (Hint: Use Issue Date to segregate into seasons)

#2. Then, find the three most common violations for each of these seasons.(Hint: A similar approach can be used as mention in the hint for question 4.)

**Approach:**

* We have extracted month from Issue date for each year created a column, Issue Month.
* We have marked 4 seasons to these months respectively.
* April – June -> Summer
* July - September -> Rainy
* Oct - December -> Autumn
* Jan - March -> Winter

**Analysis:**

* Common violation codes across seasons are 21, 14 , 38.
* For Summer season order of Violation codes are 21,38 and 14 in decreasing order of violation counts.
* For Autumn season order of Violation codes are 21,38 and 14 in decreasing order of violation counts.
* For Winter season order of Violation codes are 38, 21 and 14 in decreasing order of violation counts.
* For Rainy season order of Violation codes are 21,38 and 14 in decreasing order of violation counts.

#7. The fines collected from all the parking violation constitute a revenue source for the NYC police department. Let’s take an example of estimating that for the three most commonly occurring codes.

1. Find total occurrences of the three most common violation codes.
2. Find the total amount collected for the three violation codes with maximum tickets. State the code which has the highest total collection.

What can you intuitively infer from these findings?

**Analysis:**

**2015:**

* Following are the figures for the total occurrences of most violation codes.

Violation.Code Total.Violation

1. 21 1501614
2. 38 1324586
3. 14 924627

* Total amount collected for the three violation codes with maximum tickets.

Violation.Code Total.Violation Fee Total\_Fine

1 21 1501614 55 82588770

2 38 1324586 50 66229300

3 14 924627 115 106332105

* Violation code 14 has the highest total collection.

**2016**

* Following are the figures for the total occurrences of most violation codes.

Violation.Code Total.Violation

1 21 1531587

2 36 1253512

3 38 1143696

* Total amount collected for the three violation codes with maximum tickets.

Violation.Code Total.Violation Fee Total\_Fine

1 21 1531587 55 84237285

2 36 1253512 50 62675600

3 38 1143696 50 57184800

* Violation code 21 has the highest total collection.

**2017**

* Following are the figures for the total occurrences of most violation codes.

Violation.Code Total.Violation

1 21 1528588

2 36 1400614

3 38 1062304

* Total amount collected for the three violation codes with maximum tickets.

Violation.Code Total.Violation Fee Total\_Fine

1 21 1528588 55 84072340

2 36 1400614 50 70030700

3 38 1062304 50 53115200

* Violation code 21 has the highest total collection.