Examine the data

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

2015 : 11809233

2016 : 10626899

2017 : 10803028

2. Find out the number of unique states from where the cars that got parking tickets came from. (Replaced 99 with *NY* since had max tickets)

2015 : 68

2016 : 67

2017 : 66

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.

2015 : 1799170 (~15%)

2016: 1868656 (~17.6%)

2017: 2072400 (~19%)

Looks like the number of tickets with null violation location is on the rise by almost 2% per year.

Aggregation tasks

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

2015

	Violation Code	count
1	21	1630912
2	38	1418627
3	14	988469
4	36	839197
5	37	795918

2016

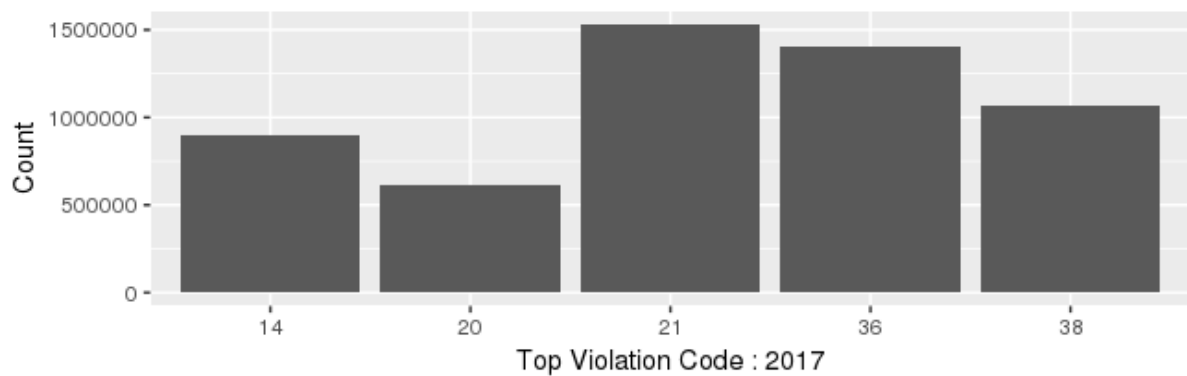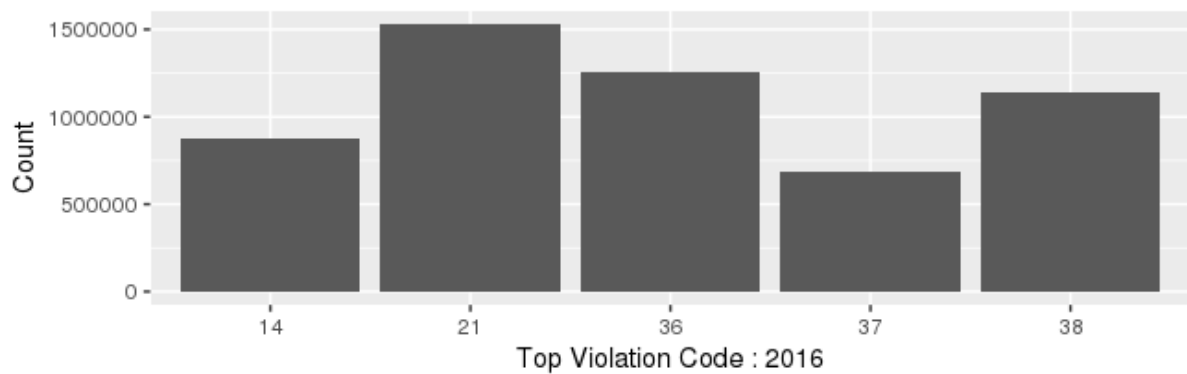|	Violation Code	count
1	21	1531587
2	36	1253512
3	38	1143696
4	14	875614

```
5             37          686610

2017
        Violation Code    count
1            21          1528588
2            36          1400614
3            38          1062304
4            14          893498
5            20          618593
```


Top Violation Code : 2015

Top Violation Code : 2016

Top Violation Code : 2017
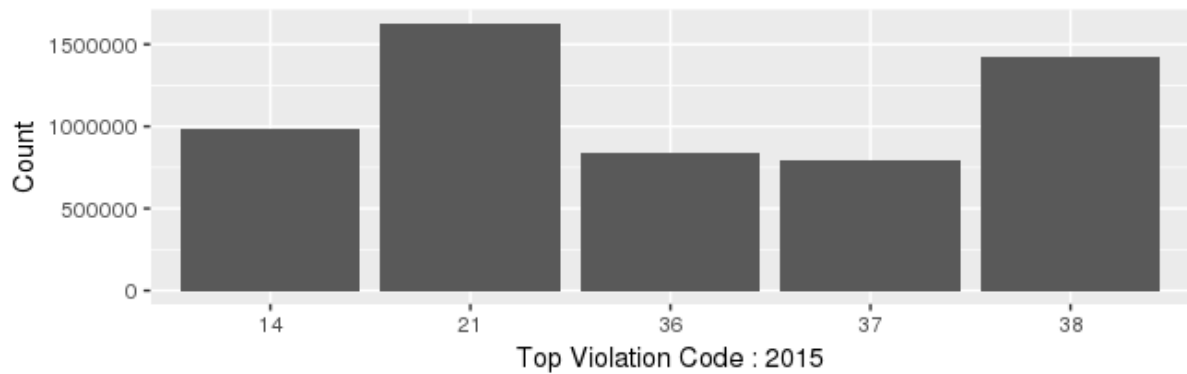
2. How often does each 'vehicle body type' get a parking ticket? How about the 'vehicle make'?

2015:

```
  Vehicle Body Type    count
1            SUBN 3729346
2            4DSD 3340014
```

```
3                VAN 1709091
4               DELV  892781
5                SDN  524596
```

2016:

```
Vehicle Body Type     count
1               SUBN 3466037
2               4DSD 2992107
3                VAN 1518303
4               DELV  755282
5                SDN  424043
```

2017:

```
  Vehicle Body Type     count
1               SUBN 3719802
2               4DSD 3082020
3                VAN 1411970
4               DELV  687330
5                SDN  438191
```

The vehicle body type frequency to get tickets is almost same throughout all years. Maybe the number of vehicles on road corelate to this and not the body type itself.

2015:

```
  Vehicle Make    count
1          FORD 1521874
2         TOYOT 1217087
3         HONDA 1102614
4         NISSA  908783
5         CHEVR  897845
```

2016:

```
  Vehicle Make    count
1          FORD 1324774
2         TOYOT 1154790
3         HONDA 1014074
4         NISSA  834833
5         CHEVR  759663
```

2017:

```
  Vehicle Make    count
1          FORD 1280958
2         TOYOT 1211451
3         HONDA 1079238
4         NISSA  918590
5         CHEVR  714655
```

Vehicle make ticket frequency are also similar throughout the years.

   3. A precinct is a police station that has a certain zone of the city unde
      r its command. Find the (5 highest) frequency of tickets for each of th
      e following:

'Violation Precinct'

2015:

	Violation Precinct	count
1	~~0~~	~~1799170~~
2	19	598351
3	18	427510
4	14	409064
5	1	329009
6	114	320963

2016:

	Violation Precinct	count
1	~~0~~	~~1868655~~
2	19	554465
3	18	331704
4	14	324467
5	1	303850
6	114	291336

2017:

	Violation Precinct	count
1	~~0~~	~~2072400~~
2	19	535671
3	14	352450
4	1	331810
5	18	306920
6	114	296514

Top Violation precinct : 2015

Top Violation precinct : 2016

Top Violation precinct : 2017
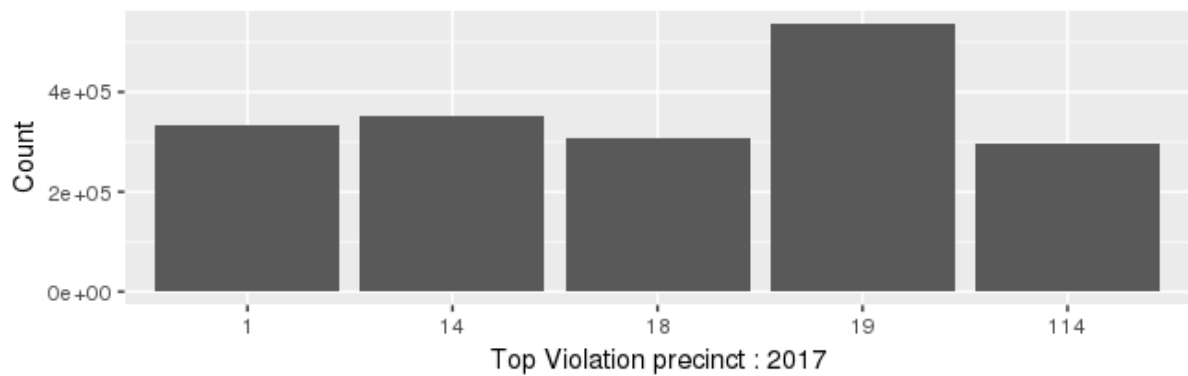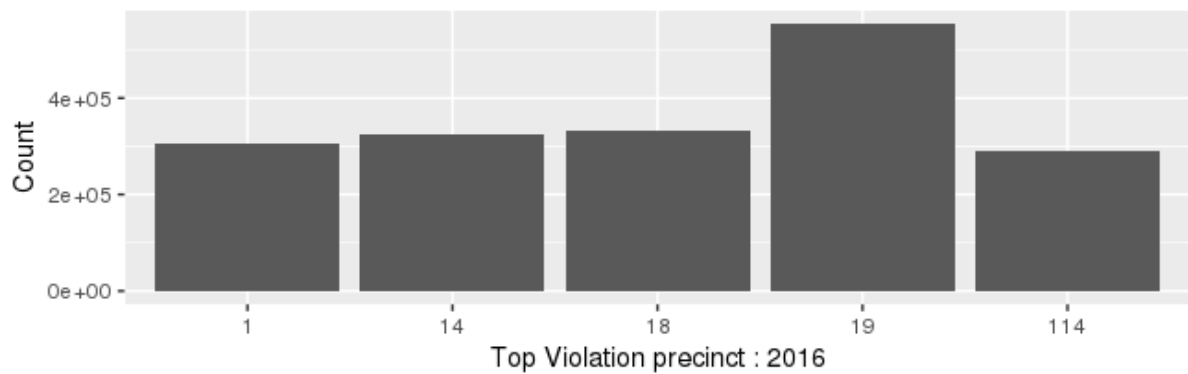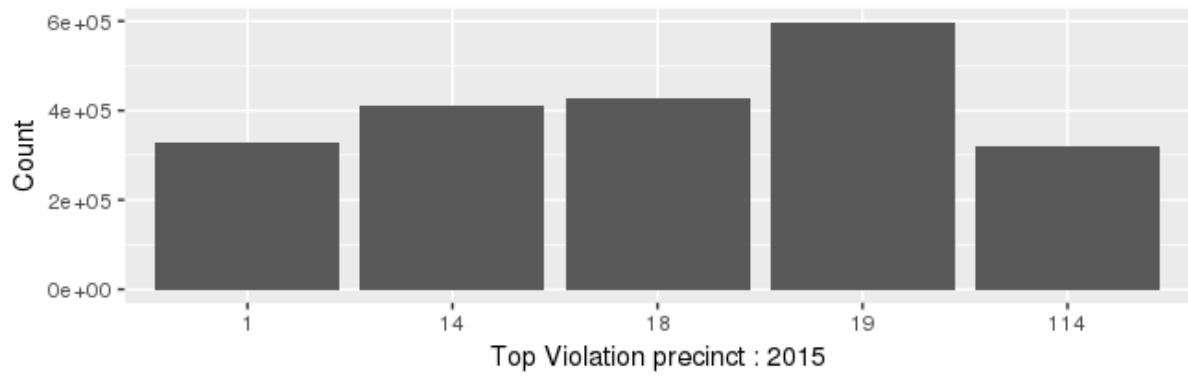
Even though the erroneous records (0) are on the rise, overall the violation precincts have the same frequency throughout the years.

'Issuer Precinct'

2015:

```
  Issuer Precinct    count
1               0 2037745
2              19  579998
3              18  417329
4              14  392922
5               1  318778
6             114  314437
```

2016:

```
  Issuer Precinct    count
1              0 2140274
2             19  540569
3             18  323132
4             14  315311
5              1  295013
6            114  286924
```

2017:

```
  Issuer Precinct    count
1              0 2388479
2             19  521513
3             14  344977
4              1  321170
5             18  296553
6            114  289950
```

Issuer precincts have the same qualities to that of violation precincts. Same frequency, Looks like the police are catching the offenders immediately.

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?

The top 3 precincts are 19,14,1 from previous questions.

2015:

```
  Violation Code  count
1             14 233299
2             38 123981
3             16 108355
4             69 104429
5             37 102141
6             46  84177
```

2016:

```
  Violation Code  count
1             14 199806
2             46 110983
3             38 101144
4             16  92933
5             37  92416
6             69  85438
```

2017:

```
  Violation Code   count
```

```
1                14 210865
2                46 129758
3                38  97816
4                37  89584
5                69  73875
6                16  72237
```

Violation code 14 has unusually high numbers in all 3 years. Lets check if its common across precincts.

2015:

```
  Issuer Precinct Violation Code count
1               19               38 97154
2               19               37 85007
3               19               14 64133
4               14               69 84895
5               14               14 81896
6               14               31 43928
7                1               14 79164
8                1               16 43915
9                1               20 25550
```

Violation Code 14 is common throughout precincts in 2015.

2016:

```
  Issuer Precinct Violation Code count
1               19               38 77183
2               19               37 75641
3               19               46 73016
4               14               69 67932
5               14               14 62426
6               14               31 35711
7                1               14 69200
8                1               16 37981
9                1               20 24725
```

In 2016, 14 was not common in precinct 19.

2017:

```
  Issuer Precinct Violation Code count
1               19               46 86390
2               19               37 72437
3               19               38 72344
4               14               14 73837
5               14               69 58026
6               14               31 39857
7                1               14 73522
8                1               16 38937
9                1               20 27841
```

Also in 2017: 14 was not common in precinct 19.

Looks like after 2015 , Issuer precinct 19 is doing a great job containing violation code 14, however 46,37,38 is common occurrence in precinct 19 throughout the years.

5. You'd want to find out the properties of parking violations across different times of the day

Most commonly occurring violations for each group:

2015:

```
   Violation period Violation Code count(1)
1              0-4               38    464332
2              0-4               37    347481
3              0-4               14    250345
4              4-8               14    303696
5              4-8               38    261888
6              4-8                7    242589
7             8-12               21   1292226
8             8-12               38    546381
9             8-12               36    396871
10           12-16               38    146026
11           12-16               21    145374
12           12-16               36    113334
13           16-20               40         4
14           16-20               20         3
15           16-20               46         2
16           20-24               46         1
17           20-24               40         1
18           20-24               51         1
```

2016:

```
   Violation period Violation Code count(1)
1              0-4               36    378436
2              0-4               38    367902
3              0-4               37    297686
4              4-8               14    275087
5              4-8               38    213478
6              4-8                7    165627
7             8-12               21   1209672
8             8-12               36    586791
9             8-12               38    441273
10           12-16               36    167282
11           12-16               21    134956
12           12-16               38    120784
13           16-20               21         4
14           16-20               40         3
15           16-20               46         2
16           20-24               19         2
17           20-24               20         2
18           20-24               98         1
```

2017:

```
   Violation period Violation Code count(1)
```

```
1               0-4          36    376961
2               0-4          38    356616
3               0-4          37    265924
4               4-8          14    286025
5               4-8          38    205532
6               4-8           7    175828
7              8-12          21   1183054
8              8-12          36    751422
9              8-12          38    393547
10            12-16          36    211434
11            12-16          21    147979
12            12-16          38    106606
13            16-20          21         4
14            16-20          98         3
15            16-20          46         3
16            20-24          78         2
17            20-24          14         1
18            20-24          40         1
```

Most commonly occurring violations:

2015: 14,38,21

2016: 21,36,38

2017: 21,36,38

Most Common time of day:

2015:

```
  Violation Code Violation period count(1)
1            14              8-12    366515
2            38              8-12    546381
3            21              8-12   1292226
```

2016:

```
  Violation Code Violation period count(1)
1            21              8-12   1209672
2            36              8-12    586791
3            38              8-12    441273
```

2017:

```
  Violation Code Violation period count(1)
1            21              8-12   1183054
2            36              8-12    751422
3            38              8-12    393547
```

It seems most violations happen during 8-12 which would be rush hour.

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

Season count:

2015:

```
   season   count
1 Spring 2956623
2 Summer 3157641
3   Fall 2794515
4 Winter 2898739
```

2016:

```
   season   count
1 Spring 2790866
2 Summer 2434349
3   Fall 2973044
4 Winter 2424360
```

2017:

```
   season   count
1 Spring 2880675
2 Summer 2606196
3   Fall 2830770
4 Winter 2485324
```

No clear insights from just the count itself. Perhaps seasons have no correlation to the violations.

Although winter has less violations somewhat.

Most common violations:

2015:

```
   season Violation Code   count
1  Spring             21 425354
2  Spring             38 327057
3  Spring             14 243771
4  Summer             21 476855
5  Summer             38 364578
6  Summer             14 257327
7    Fall             21 360548
8    Fall             38 329955
9    Fall             36 244788
10 Winter             38 397037
11 Winter             21 368155
12 Winter             14 250827
```

2016:

```
   season Violation Code   count
1  Spring             21 383752
2  Spring             36 374362
3  Spring             38 299459
```

```
4  Summer            21 391411
5  Summer            38 272161
6  Summer            14 215390
7    Fall            36 438320
8    Fall            21 395356
9    Fall            38 303397
10 Winter            21 360268
11 Winter            36 314764
12 Winter            38 268421
```

2017:

```
   season Violation Code  count
1  Spring             21 402807
2  Spring             36 344834
3  Spring             38 271192
4  Summer             21 405961
5  Summer             38 247561
6  Summer             36 240396
7    Fall             36 456046
8    Fall             21 357479
9    Fall             38 283828
10 Winter             21 362341
11 Winter             36 359338
12 Winter             38 259723
```

21,36,38 and somewhat 14 has the most amount of violations throughout the years.

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.

Frequency of violation codes and fines

2015:

```
  Violation Code    count
1             21 1630912
2             38 1418627
3             14  988463
```

```
  Violation Code    count       Fine
1             21 1630912  89700160
2             38 1418627  70931350
3             14  988463 113673245
```

14 has the most collection but 21 and 38 are not far behind.

2016:

```
  Violation Code    count
1             21 1530787
2             36 1253511
3             38 1143438
```

```
   Violation Code    count       Fine
1             21 1530787 84193285
2             36 1253511 62675550
3             38 1143438 57171900
```

21 has the most fines.

2017:

```
  Violation Code    count
1             21 1528588
2             36 1400614
3             38 1062304
```

```
   Violation Code    count       Fine
1             21 1528588 84072340
2             36 1400614 70030700
3             38 1062304 53115200
```

Again 21 has the most fines.

21,36,38 and 14 (due to the high fine amount) are bringing the most revenue.