NYC Parking Tickets: An Exploratory Analysis

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

Field "Summons Number" has the ticket id information. So a count of distinct number of "summons number" gives the total number of tickets for the year – which comes to 5431918.

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

Field "Registration State" has the State information. Count of Distinct number of Registration state comes to 65. There is one numeric entry 99 for this field which is invalid. So eliminating it gives 64.

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

Grouping by field "Violation code" and aggregating on "Summon Number" field and arranging the resultant dataframe in descending order – gives the top violation codes.

+	·+	·+
•	on Code ticko	
+	+	+
21	768087	
36	662765	İ
38	542079	
14	476664	
20	319646	
+	+	+

only showing top 5 rows

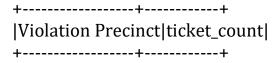
4.How often does each 'vehicle body type' get a parking ticket? How a bout the 'vehicle make'?

Grouping by field "Vehicle Body Type" and aggregating on "Summon Number" field and arranging the resultant dataframe in descending order – gives the top vechile body types that were issues tickets.

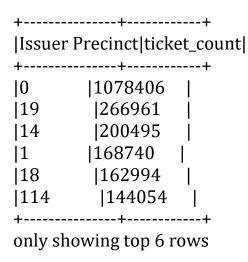
+	+	+		
Vehicle Body Type ticket_count				
+	+	+		
SUBN	1883954			
4DSD	1547312			
 VAN	724029			
DELV	358984			
SDN	194197			
+	+	+		

5.A precinct is a police station that has a certain zone of the city under its command. Find the (5 highest) frequencies of tickets for "Violation Precinct" and "Issuer Precinct".

Grouping by field "Violation Precinct" and aggregating on "Summon Number" field and arranging the resultant dataframe in descending order – gives the top Violation Precinct.



Grouping by field "Issuer Precinct" and aggregating on "Summon Number" field and arranging the resultant dataframe in descending order – gives the top Issuer Precinct.



6. Find the violation code frequencies for three precincts that have iss ued the most number of tickets. Do these precinct zones have an ex ceptionally high frequency of certain violation codes? Are these co des common across precincts?

Top 3 violation Precinct are $\mathbf{0}$, $\mathbf{19}$ and $\mathbf{14}$

+	+
Violatio	on Precinct
+	+
0	
19	1
14	

+----+

Violation code of 36 is highest for violation precinct 0 Violation Code of 46 is highest for violation precinct 19 Violation Code of 14 is highest for violation precinct 15

.....

+	+-	+-	+
0	36	662765	1
0	7	210174	
0	5	48076	
19	46	50785	
19	38	37483	
19	37	36468	
14	14	45885	
14	69	30465	
14	31	22649	

7. Find out the properties of parking violations across different times of the day:

Violation Time is assumed to be in format hhmmA (for AM) or hhmmP(f or PM) .Fields which doesn't fall in this format are dropped.

16), (12-15), (8-11), (4-7), (8-11), (12-15) -19),(20-23).

The Top 3 violation code for each hour group is shown as below

+-----+ |violation_hour_groups|Violation Code|ticket_count|violation_rank| +-----+ 10-3 |21 |26444 1 |40 |22420 12 10-3 |78 |13737 |3 |38 |240721 |2 10-3 |12-15

```
|12-15
            |37
                    |167025
                             |3
                             |1
|12-15
            |36
                    |286284
            |38
                             |1
|16-19
                    |102855
|16-19
            |14
                    |75902
                             |2
            |37
116-19
                    |70345
                            13
|20-23
            |7
                   26293
                            11
                            |2
                   122336
|20-23
            |40
|20-23
           |14
                    |21045
                            |3
|4-7
           |14
                  |74113
                           11
           |40
                  |60652
                           12
14-7
14-7
           |21
                  |57894
                           13
           |21
                   |598062
|8-11
                            |1
18-11
           |36
                   |348165
                             12
           |38
                   1176570
|8-11
+-----+
```

8. Find the three most commonly occurring violation codes, find the most common time of the day

```
+-----+
|Violation Code|violation_hour_groups|ticket_count|
+-----+
|38 |12-15 |240721 |
|21 |8-11 |598062 |
|36 |8-11 |348165 |
+------+
```

9. Seasonality Data:

We are dividing the year into 4 seasons – Season 0 – months (1-3), Season 1 – months(4-6), Season 2 – months(7-9), Season 3 – months(10-12). Then we are grouping on Seasons and aggregating the ticket counts .We get the following result:

|season|Violation Code|ticket_count|violation_rank| +-----+ |0 |21 |373874 |1 | |0 |36 |348240 |2 |

```
|0
    |38
             |287000
                         |3
             |314525
|1
    |36
                         |2
    |38
                         13
|1
             |255067
                        |1
    |21
             |393957
|1
             |149
                      |3
12
    |40
12
    |46
             |288
                      |1
12
    |21
             |212
                     |2
13
    |46
             |77
                      |1
    |21
                     |2
13
             |44
             |44
                      12
13
    |40
13
    |20
             |33
                      |3
```

9. Three most common violation codes

Taken the average of fines in localities of highest density and other area.

21	Street Cleaning: No parking where parking is not allowed by sign, street marking or traffic control device.	\$55
36	Exceeding the posted speed limit in or near a designated school zone.	\$50
38	Failing to show a receipt or tag in the windshield. Drivers get a 5-minute grace period past the expired time on parking meter receipts.	\$50

Violation 21 has the highest total collection: