NYC Parking Tickets: An Exploratory Analysis

**Examine the data**

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

>>> Total ticket count=10803029

1. Find out the number of unique states from where the cars that got parking tickets came. *(Hint: Use the column 'Registration State'.)*  
   *There is a numeric entry '99' in the column, which should be corrected. Replace it with the state having the maximum entries. Provide the number of unique states again.*

>>> Number of unique states are 67

>>>After replacing 99 unique states are 66

Here we replaced 99 by NY as it has maximum entries

**Aggregation tasks**

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

[(violation code count=1528588, Violation Code='21'),

(Violation code count =1400614, Violation Code='36'),

(Violation code count =1062304, Violation Code='38'),

(Violation code count =893498, Violation Code='14'),

(violation code count =618593, Violation Code='20')]

>>>> Top 5 Violation Code are 21, 36,38,14,20

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

[(violation\_summon\_count=3719802, Vehicle\_Body\_Type='SUBN'),

(violation\_summon\_count=3082020, Vehicle\_Body\_Type='4DSD'),

(violation\_summon\_count=1411970, Vehicle\_Body\_Type='VAN'),

(violation\_summon\_count=687330, Vehicle\_Body\_Type='DELV'),

(violation\_summon\_count=438191, Vehicle\_Body\_Type='SDN')]

Top 5 Violation Body type are SUBN / 4DSD / VAN / DELV and SDN

[(violation\_summon\_count=1280958, Vehicle\_Make='FORD'),

(violation\_summon\_count=1211451, Vehicle\_Make='TOYOT'),

(violation\_summon\_count=1079238, Vehicle\_Make='HONDA'),

(violation\_summon\_count=918590, Vehicle\_Make='NISSA'),

(violation\_summon\_count=714655, Vehicle\_Make='CHEVR')]

Top 5 Vehicle Body Type are FORD / TOYOTA / HONDA / NISSAN and Chevrolet

1. 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 each of the following:
2. 'Violation Precinct' (This is the precinct of the zone where the violation occurred). Using this, can you draw any insights for parking violations in any specific areas of the city?

>>>> Top five 'Violating Precinct' are 19, 14,1,18 and 114

1. 'Issuer Precinct' (This is the precinct that issued the ticket.)  
   *Here, you would have noticed that the dataframe has the'Violating Precinct' or 'Issuing Precinct' as '0'. These are erroneous entries. Hence, you need to provide the records for five correct precincts. (****Hint****: Print the top six entries after sorting.)*

>>>>> top five Issuing Precinct' are 19, 14,1,18 and 114

1. Find the violation code frequencies for three precincts that 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?   
   ***(Hint****: In the SQL view, use the 'where' attribute to filter among three precincts.)*

>>>> Violation code frequencies for three precincts that have issued the most number of tickets i.e. for 19, 14 and 1

>>>>>> Yes these precinct zones have an exceptionally high frequency of certain violation codes

>>>>>> Top three codes are not common across precincts 19, 14 and 1

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

* Find a way to deal with missing values, if any.  
  ***(Hint****: Check for the null values using 'isNull' under the SQL. Also, to remove the null values, check the 'dropna' command in the API documentation.)*

>>>>> as we getting 0 results we are go to go

7. The fines collected from all the instances of parking violation constitute a source of revenue for the NYC Police Department. Let’s take an example of estimating this for the three most commonly occurring codes:

* Find the total occurrences of the three most common violation codes.

>>> total occurrences of the three most common violation codes =3991506

* Then, visit the website:  
  <http://www1.nyc.gov/site/finance/vehicles/services-violation-codes.page>  
  It lists the fines associated with different violation codes. They’re divided into two categories: one for the highest-density locations in the city and the other for the rest of the city. For the sake of simplicity, take the average of the two.
* Using this information, find the total amount collected for the three violation codes with the maximum tickets. State the code that has the highest total collection.
* What can you intuitively infer from these findings?

>>>>> Amount collected for the violation code 21 =84072340.0

>>>>> Amount collected for the violation code 36 = 70030700.0

>>>>> Amount collected for the violation code 38 = 53115200.0

Violation Code 21 has highest collection

No parking (Violation Code 21) where parking is not allowed by sign, street marking or traffic control device is most ignored violation