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

Violation_Code	Frequency_of_Occurence
21	1500396
36	1345237
38	1050418
14	880152
20	609231

Code 21: *Street Cleaning: No parking where parking is not allowed by sign, street marking or traffic control device.*
Code 36: *Exceeding the posted speed limit in or near a designated school zone.*
Code 38: *Failing to show a receipt or tag in the windshield.*

The top three violations codes (21, 36, 38) in terms of frequency constituents approx. 37% of the total parking tickets.

Insights

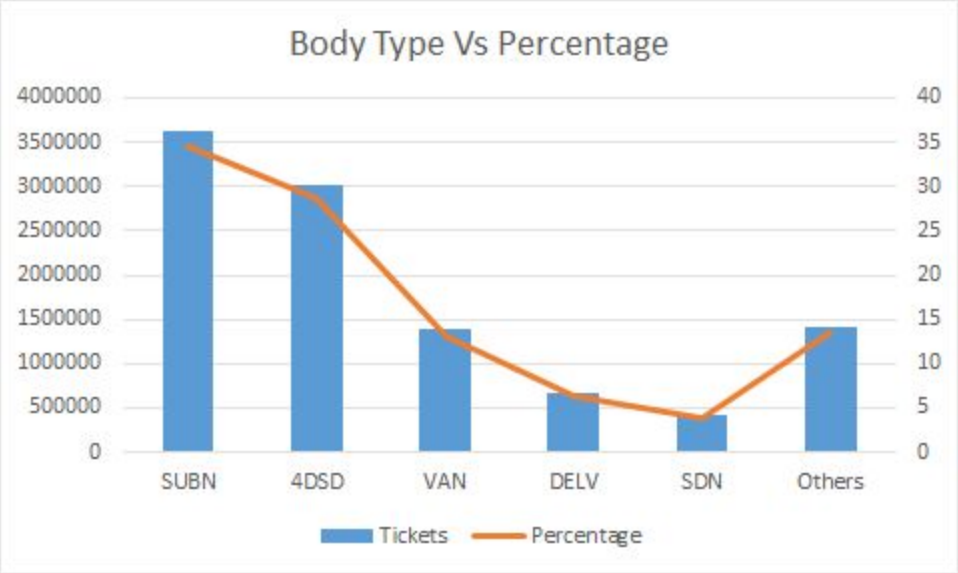
1. 14% of the parking tickets are issues for parking in a non-parking zone where parking is not allowed by sign, street marking or traffic control device.
2. 12.976% of the tickets are issued due to overspeeding near a designated school zone.
3. 09.96% of the tickets drivers failed to show a receipt or tag on the windshield.

Ques 2: How often does each 'vehicle body type' get a parking ticket? How about the 'vehicle make'?

Body Type	Tickets	Percentage
SUBN	3632003	34.4606603
4DSD	3017372	28.6290048
VAN	1384121	13.1326223
DELV	672123	6.37714296
SDN	414984	3.93739285
Others	1418960	13.4631768

Insights:

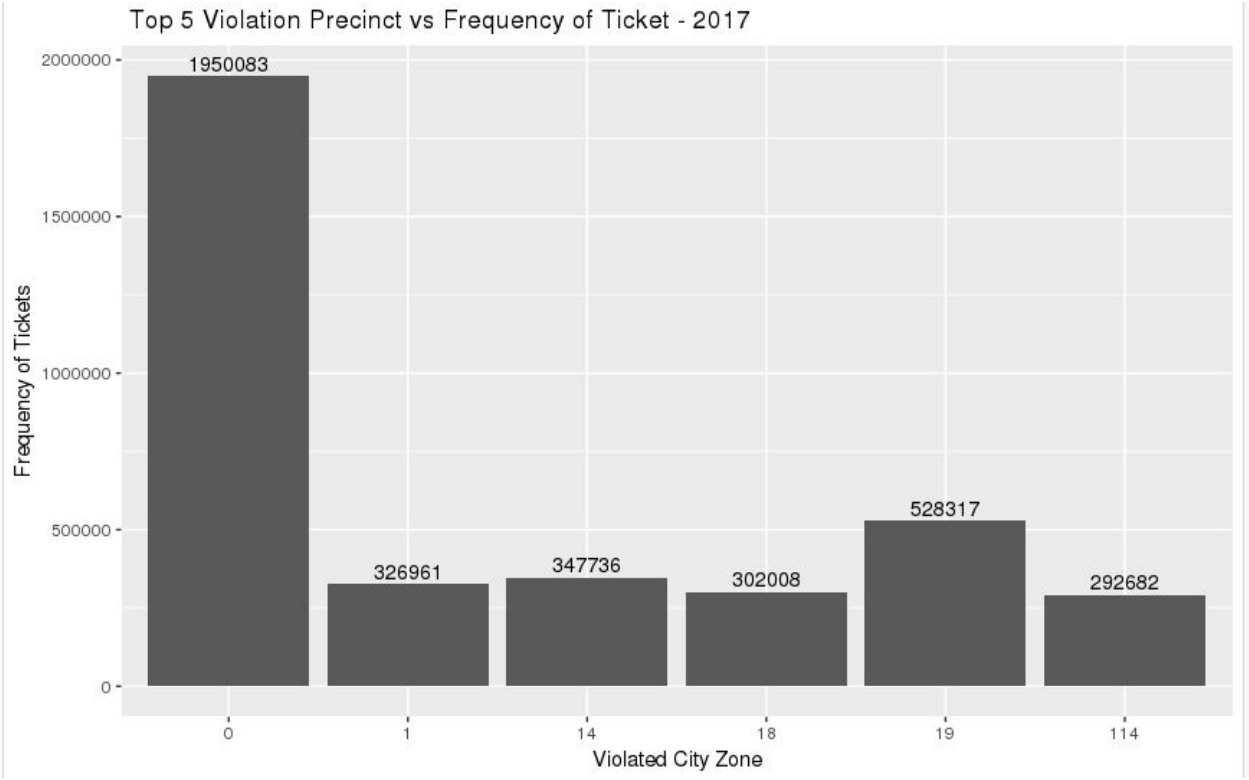
1. Sub urban and 4DSD body type vehicles are responsible for a huge percentage of parking tickets ie approx. **63%**

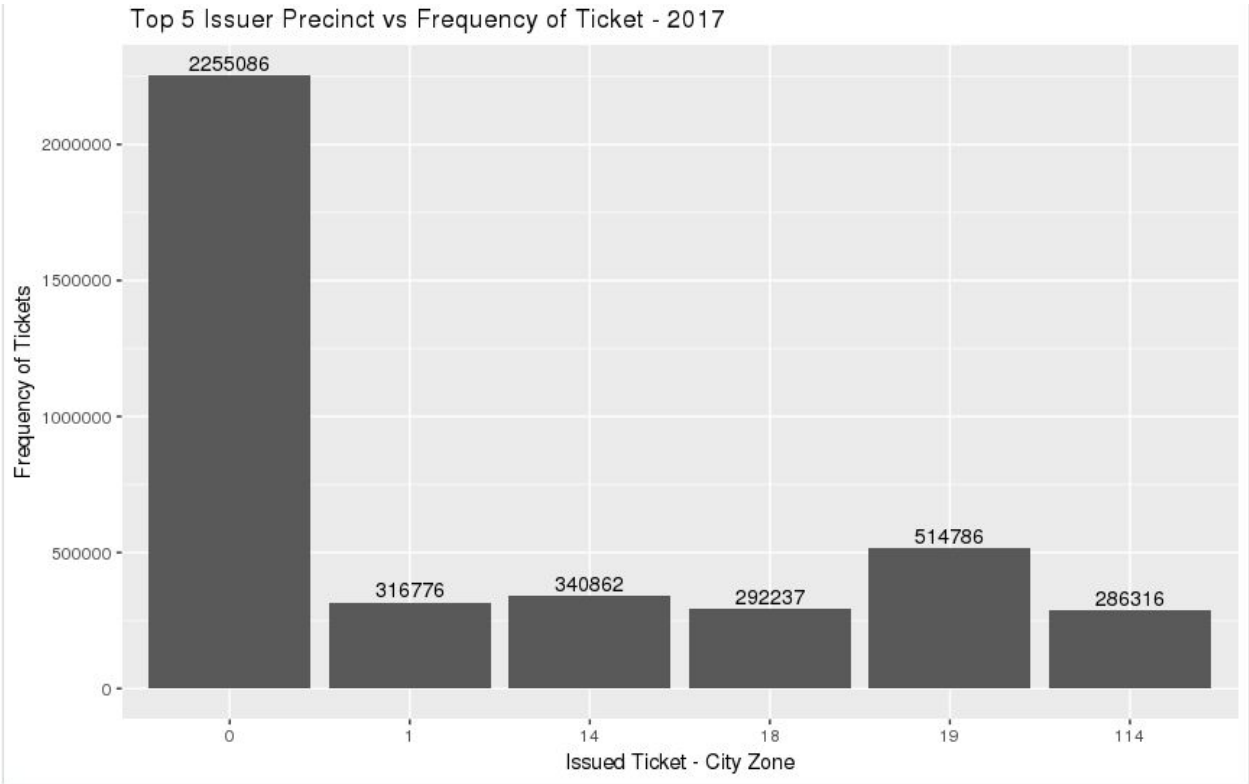


Ques 3: Find the (5 highest) frequency of tickets for each of the following:
3.1: Violation Precinct - (this is the precinct of the zone where the violation occurred)
3.2: 5 highest frequency of tickets for 'Violation Precinct'

Precinct	Frequency	Percentage
0	1950083	18.50250338
19	528317	5.012703088
14	347736	3.29933983
1	326961	3.102225396
18	302008	2.865469849
114	292682	2.776984207

Assumption: The Precinct “0” has a very high frequency. This is very unlikely looking at the data and the hints given in the question. Assuming that this data is corrupted and has not been entered properly, considering top 6 precinct.





19 Zone holds the highest frequency of tickets for both Violation & Issuer Precinct

Ques 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?

Note: Excluding the Precinct '0' entries

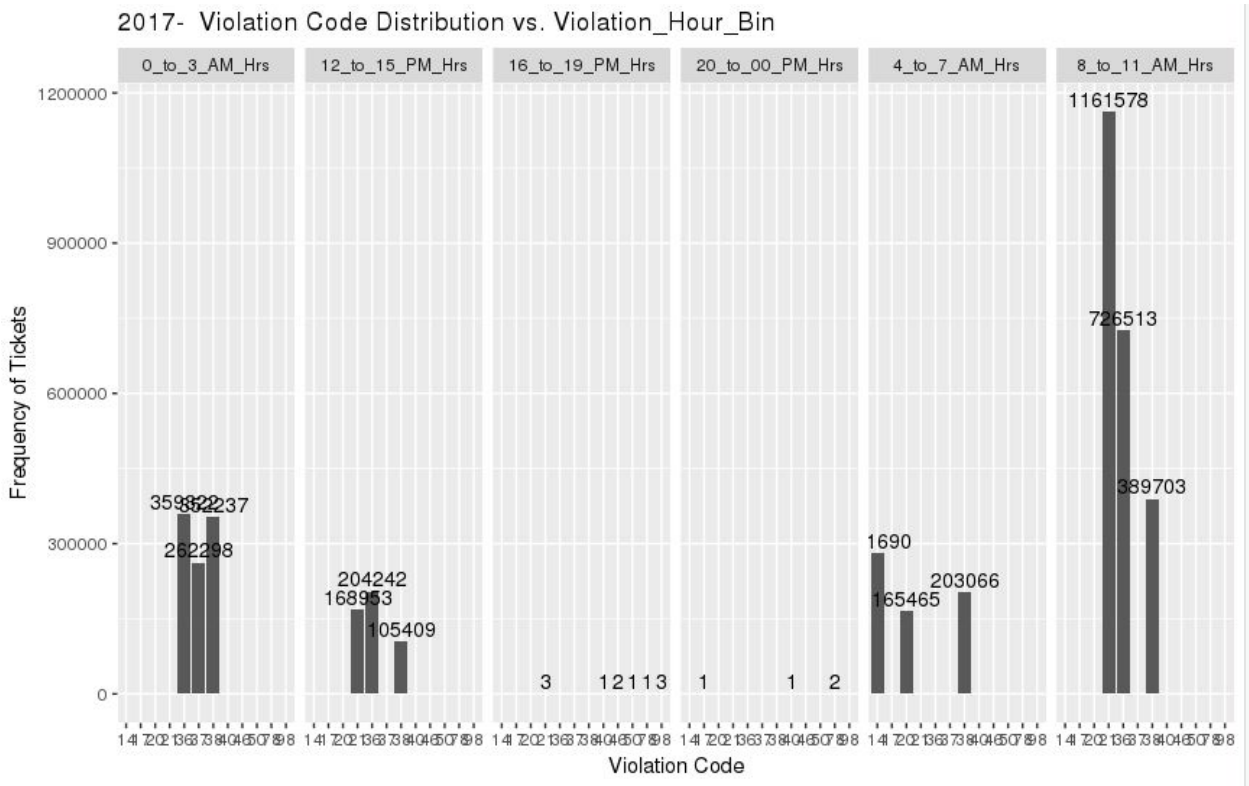
Precinct	Code	Frequency
19	46	84789
14	14	73007
1	14	72520
19	38	71631
19	37	71592
14	69	57316

Code 14: General No Standing: Standing or parking where standing is not allowed by sign, street marking or; traffic control device.

Insight

1. Violation Code = 14 has a high frequency in both Issuer Precincts 1 and 14
2. A lot of people are standing at parking spots where standing is not allowed.

Ques 5: You'd want to find out the properties of parking violations across different times of the day:
5.1 Find a way to deal with missing values, if any.
5.2 Violation Time field is specified in a strange format. Find a way to make this into a time attribute that you can use to divide into groups. Divide 24 hours into six equal discrete bins of time. The intervals you choose are at your discretion. For each of these groups, find the three most commonly occurring violations.



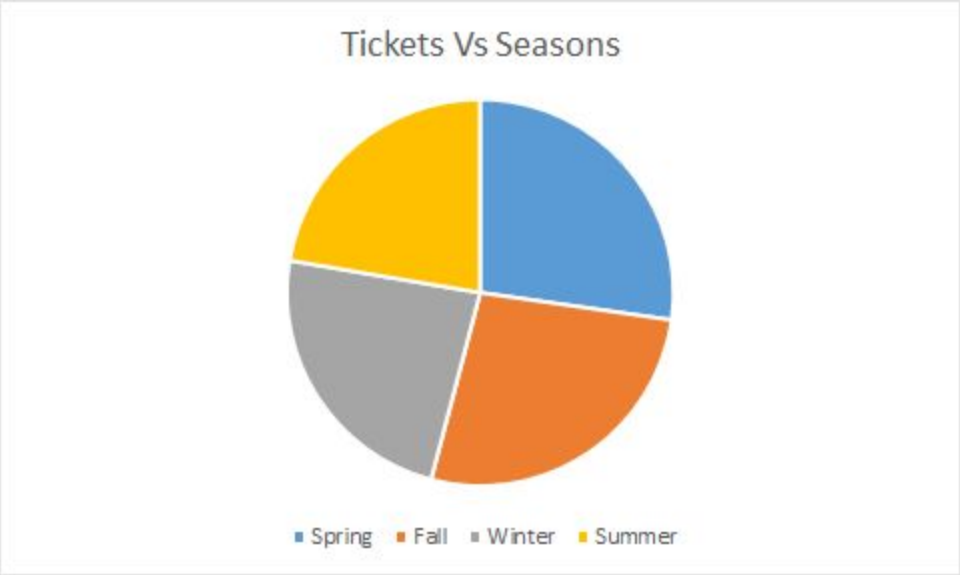
Violation Hour	Violation Code	Violation Bin
4	20	4_to_7_AM_Hrs
11	70	8_to_11_AM_Hrs
7	17	4_to_7_AM_Hrs
11	74	8_to_11_AM_Hrs
4	21	4_to_7_AM_Hrs
6	20	4_to_7_AM_Hrs

Insight:

1. Most violations happening from 8AM to 11AM slot
2. Top time slots for violations are 0-03AM, 12-15PM, 4-7AM and 8-11AM

Ques 6: Let's try and find some seasonality in this data
#Dividing the whole year into rainy,winter and summer seasons - 4 months each season
#6.1 first divide into seasons
#6.2 and then calculate the frequencies of tickets for each season

Season	Tickets
Spring	2873383
Fall	2829224
Winter	2483036
Summer	2353920



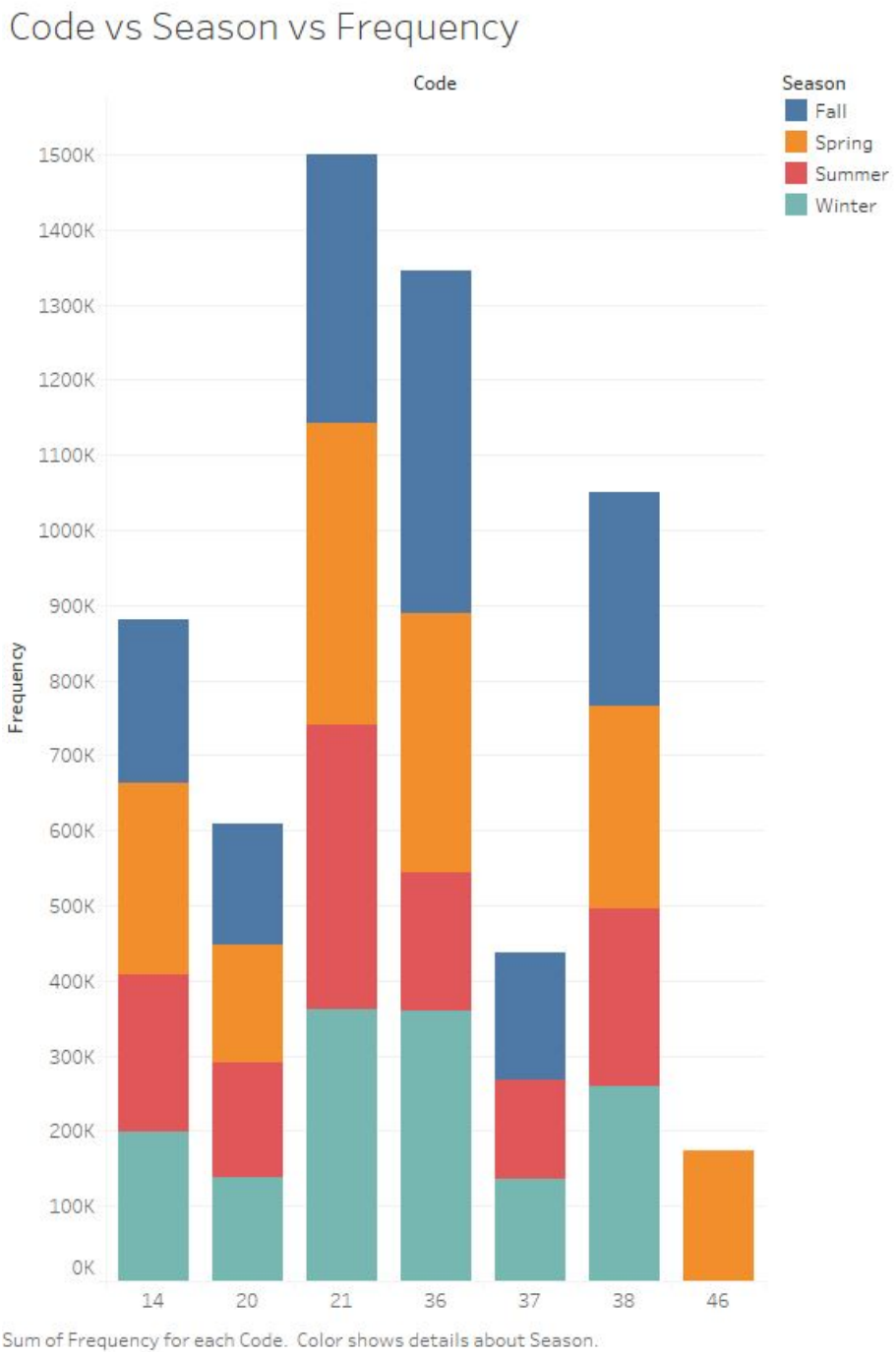
6.3 find the three most common violations for each of these seasons

Code vs Season vs Frequency

	Code						
Season	14	20	21	36	37	38	46
Fall	216,721	160,593	357,257	456,046	169,286	283,816	
Spring	256,397	157,122	402,424	344,834		271,167	173,440
Summer	207,495	154,465	378,699	185,019	132,283	235,725	
Winter	199,539	137,051	362,016	359,338	135,904	259,710	

Frequency
132,283 456,046

Sum of Frequency broken down by Code vs. Season. Color shows sum of Frequency. The marks are labeled by sum of Frequency.



Insight

Fall - Code 36 has very high value. Which means that drivers are over-speeding more near school areas in fall.

Ques 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.

Violation_Code	Frequency_of_Occurence
21	1500396
36	1345237
38	1050418

Violation code	Occurrences	Fine rate	Total Amount ($)
21	1500396	55	82521780
36	1345237	50	67261850
38	1050418	50	52520900

Highest amount of fines are collected under Violation Code = 21 with Average Fine Rate for city and other zones being $55 per ticket