In [125]:
```
import numpy as np
```

In [126]:
```
taxi = np.genfromtxt('nyc_taxis.csv', delimiter = ',', skip_header = True)
```

In [127]:
```
#what is the mean speed of all the cab rides
speed = taxi[:, 7]/(taxi[:, 8]/3600)
#now we have trip length in hrs and then speed is in miles/hr
```

In [128]:
```
mean_speed  = np.mean(speed)
print(mean_speed)
```

```
32.24258580925573
```

#so the above is the asnwer of mean speed

In [129]:
```
#Number of rides taken in february
rides_feb = taxi[taxi[:, 1] == 2]
print(rides_feb.shape[0])
```

```
13333
```

In [130]:
```
#Number of rides where tip more than $50
rides_tip = (taxi[taxi[:, -3] >50, -3].shape[0])
print(rides_tip)
```

```
16
```

In [131]:
```
#calculate the number of rides where drop was JFK airport
#drop at JFK airport will be checked by drop location = 2
no_of_rides_JFK = taxi[taxi[:, 6] == 2, 6].shape[0]
print(no_of_rides_JFK)
```

```
11832
```

In [132]:
```
#max tolls amount
maximum_tolls =  np.max(taxi[:, 11])
print(maximum_tolls)
```

```
805.54
```

In [133]:
```
#min tolls amount
minimum_tolls =  np.min(taxi[:, 11])
print(minimum_tolls)
```

```
-5.54
```

In [134]:
```
#Finding the unique pickup locations code
unique_pincode = np.unique(taxi[:, 6])
print(unique_pincode)
```

```
[0. 1. 2. 3. 4. 5. 6. 7.]
```

In [135]:
```
#finding variance of speed
var_speed = np.var(speed)
print(var_speed)
```

```
226054.47189294896
```

In [136]:
```
#sorting fare amount in ascending order
sort_fare_amount = np.sort(taxi[:, 9])
print(sort_fare_amount)
```

```
[-52.  -52.  -52.  ... 180.5 220.  400. ]
```

In [137]:
```
#knowing the shape of sort_fare_amount
print(np.shape(sort_fare_amount))
```

```
(89560,)
```

In [138]:
```
#searching 21 in trip_distance
search_distance = np.where(taxi[:, 7] == 21)
print(search_distance)
```

```
(array([    0,   682,   747,   860,  1144,  1192,  1564,  4269,  4300,
         4863,  6496,  6988,  7277,  8019,  8227, 12085, 12916, 13231,
       13709, 15595, 16038, 17194, 18529, 18822, 20340, 21323, 22290,
       22342, 24560, 24681, 25523, 25780, 26679, 28140, 29056, 30467,
       33631, 34592, 35310, 36835, 36930, 37386, 37492, 37850, 38201,
       38378, 38671, 39170, 39204, 39749, 40565, 42872, 43927, 44832,
       46265, 46975, 47077, 47130, 50026, 51151, 53363, 55714, 56136,
       56185, 56265, 56789, 58251, 58781, 59343, 60024, 60154, 61179,
       61355, 61374, 61890, 62040, 62283, 63188, 63430, 63563, 64297,
       64730, 64736, 64762, 65136, 65396, 65570, 67586, 68237, 71426,
       72324, 72328, 73331, 74045, 74602, 75649, 77469, 80370, 81312,
       82074, 82303, 83771, 84251, 87247], dtype=int64),)
```

#the above search function "where" gives the output the indexes where the value which we need is situated.