Project 1 Writeup

Our project is to analyze the towed cars in city of Chicago. We found our first dataset from Kaggle – towed\_cars.csv, then we found our second dataset of abandoned\_vehicles.csv from city of Chicago site.

After our initial analysis of the datasets, we decided to analyze the data from 6 perspectives (6 questions).

Before we start analyzing the data, we cleaned our datasets.

For the first dataset – towed\_cars.csv,

1, we cleaned up ‘Tow Date’ field to make it more readable.

2, then, we removed ‘Model’ field and dropped empty data from ‘State’ field

3, we also combined similar value in ‘Make’ field, for example.

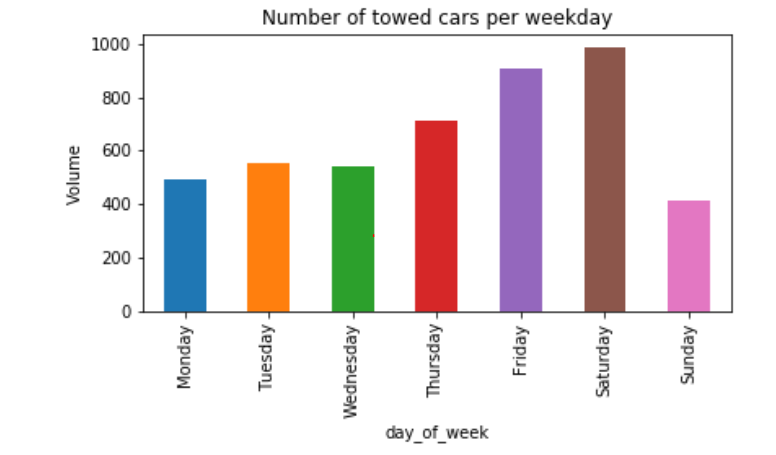
4, for our further analysis, we added a new column named ‘day\_of\_week’ in the dataset

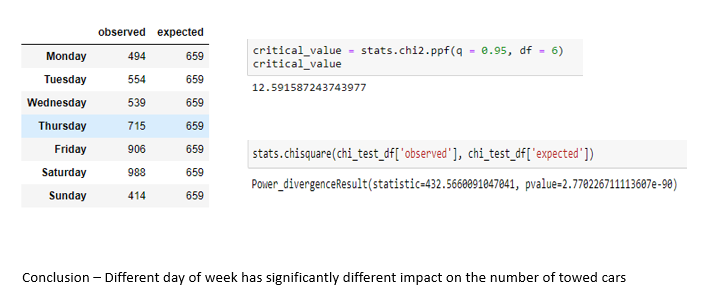
For second dataset

After we cleaned up our datasets, we analyzed the data and answered the 6 questions below

Question 1: Does the day of the week have an any impact?

According to the bar chart below, most of the cars were towed on Fridays and Saturdays, so we think the different day of week does have different impact on the volume of towed cars and we used Chi-Square test to do the validation. The conclusion shows positive result.

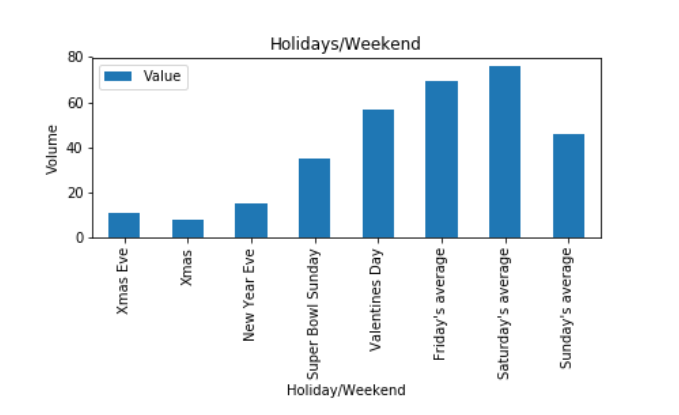


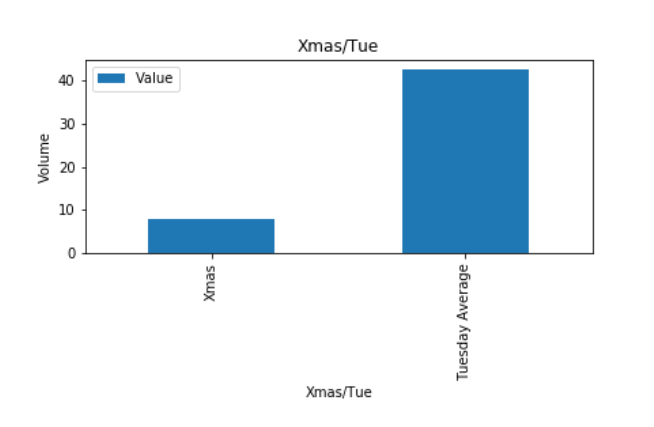


Question 2: What impact did holidays and cultural events have on towing?

According to the bar chart below, we think holidays, especially family holidays do have impact on towing. We can see the numbers of towed cars on Christmas, Christmas eve and New Year eve are significantly less than the other days. We also compared the number of towed cars with average number of towed cards on Tuesday, so we see this year’s Christmas (Last Tuesday of December) number of towed cars is much less than the average number of towed cards on Tuesday.

Our conclusion is holidays especially holidays do have impact on towing.





Question 5: We wanted to know to which facility of the four are the cars towed?

Most of the cars were towed to the 10300 S Doty location (50%), followed by the 701 N. Sacramento site (48%), then Lower Wacker and (2%) and finally ORD.

Question 6: Using the Plate variable in the dataset, we looked at to which lot each vehicle went. We see here in the states represented in each of the tow lots.

Looking at this chart, we see 37 different states represented in Chicago tow lots. Illinois is represented most frequently and in each of the four lots. Indiana, Wisconsin, and Michigan are well represented by percentage as well as in each three of the four lots.

The 701 N Sacramento Lot contains the most diverse collection of cars, Wacker has the least diverse collection, at least by license plate.