For this part of LA parking citations analysis, I was curious to find answers for the following questions.

* What time of the day more citations have been issued?
* What day of a week more citations have been issued?
* Are more citations issued on Holidays or weekdays?

So I imported necessary modules for the analysis.

**import** **numpy** **as** **np**

**import** **seaborn** **as** **sns**

**import** **pandas** **as** **pd**

**import** **csv**

**import** **matplotlib.pyplot** **as** **plt**

**import** **time**

**import** **calendar**

**import** **datetime**

**import** **holidays**

**import** **os**

**Data extraction:**

1.I got a sample **(**Date\_time\_df=dataframe.sample(frac=0.005, replace=True,axis=0)**)**

of 0.5% from 9 million rows which gave approx 47000 rows of data.

2. Removed Unwanted columns such as VIN, Route, Agency, Meter Id.

3. Dropped nan values which were 0.01% of my sample dataset.

4. Created a data frame with issued Time, Issued date and Ticket number columns.

5. Parsed Issued Date to standard YYYY-MM-DD format.

6. The issued time column was hard to manipulate. The time needed to convert in HH-MM format. For that, I added 0,00,00 depending on string length.

And then converted each time to HH-MM format using DateTime module.

7. Extracted Date, Month and year from Date and added into new columns of data frames.

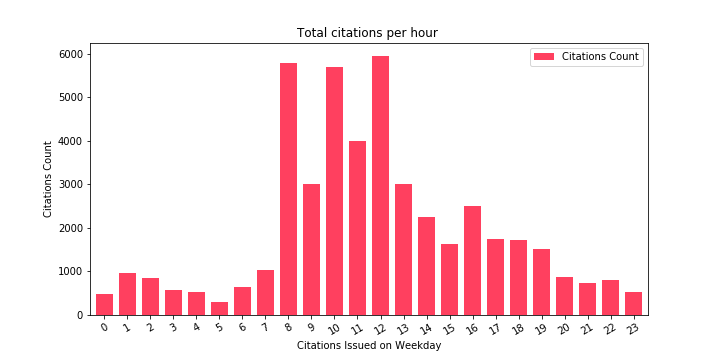
Data analysis:

* **What time of the day more citations have been issued?**

To answer this question I grouped the Time column to get all citations for each hour.

hourlygrp=Date\_time\_df.groupby("Issue time")

hourlycount=hourlygrp.count()

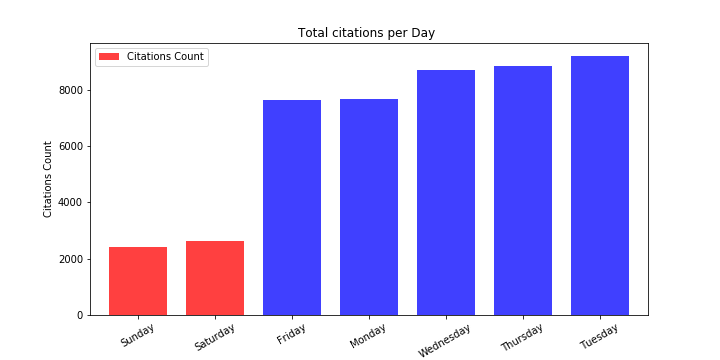
Stored the results into dataframe and plotted bar graph for better visualization.

* **What day of a week more citations have been issued?**

To answer this question I first found the days of a week each date represents.

Then I grouped them by Days\_Of\_week to get the citation count by Days of a week.

Stored the results into a data frame and plotted a bar graph out of it.



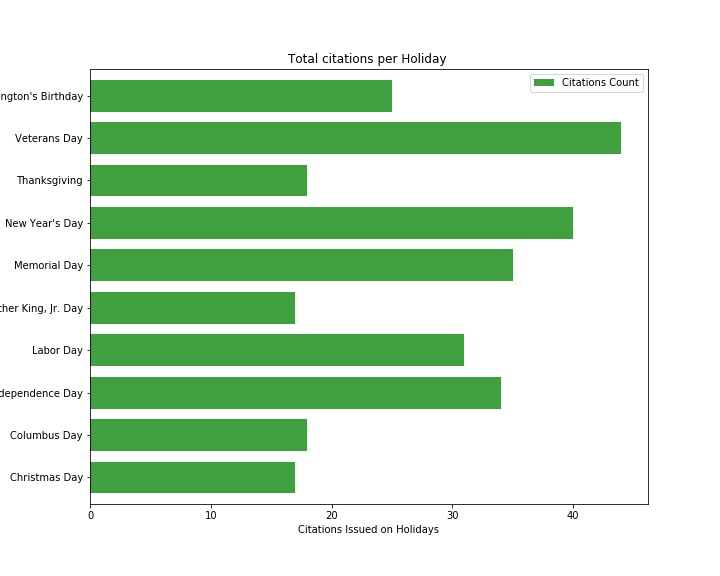
* **Are more citations issued on Holidays or weekdays?**

Imported holiday Module to do Holiday Citations count analysis.

Got the holidays from Date column using Date\_time\_df["Holiday"]=Date\_time\_df['Issue Date'].map(us\_holidays.get)

This gave similar holidays with different names. So used to **replace** function to make them unique.

Grouped them by Holidays and got the sum of Citations to count issued on that holiday.



**Created a heatmap for better visualization of total no of citations by the hour and day of the week using sns.heatmap function.**

daysofWeek=Date\_time\_df['day\_of\_week'].unique()

timeOfdayDF=Date\_time\_df.groupby(['Issue time','day\_of\_week'])["Ticket number"].count().reset\_index()

timeOfdayDF=timeOfdayDF.pivot('Issue time','day\_of\_week',"Ticket number")

fig, ax = plt.subplots(figsize=(10,10))

sns.heatmap(timeOfdayDF[daysofWeek] , cmap="YlOrRd")

