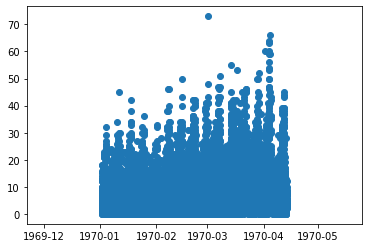
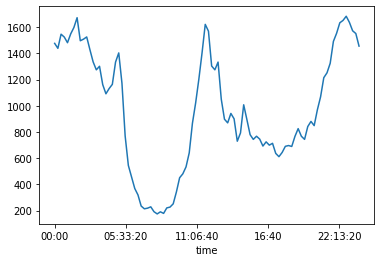
Part 1 Report:

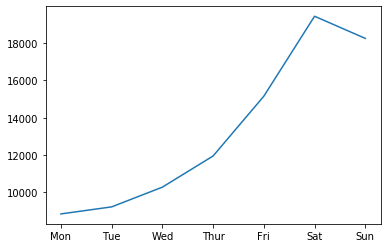
First, I put all data into 15 mins time intervals. I used the scatter plot to represent the date. The plot does not make too much sense to me since all the data points are clustered together, not able to extract meaningful information from the plot. I decide to do more analysis without considering date information.



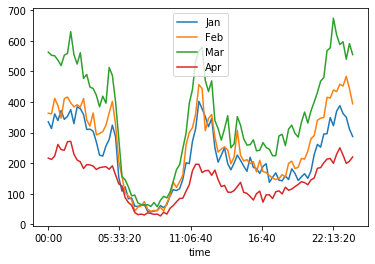
Second, I add in another column into dataset, time column, regardless of the date information. I utilized simple line plot to demonstrate the total counts of each 15 minutes in a day. The plot shows two peaks around 11am to 12pm and also 10pm to 12am which means the website has maximum clicks during these two periods of time.



Third, I add in another column, weekday column into data frame, trying to analyze the total click trends between weekdays. A simple line plot is also utilized here to show the trend. The maximum clicks happen at Saturday and start to decline in Sunday. Monday and Tuesday have the least number of clicks.



At last, I did the analysis by month and also by the 15 minutes time intervals within each month. The plot agrees with the second plot showing two peaks in a day and also showing the declining trend from Jan to Apr.



Part 2 –Experiment and metrics design

1. I would compare the number of times that the driver goes through the bridge before and after reimbursing all the toll costs as the key measure of success. If the number of times that the driver goes through the bridge has dramatically increase, that means the policy would be highly likely successful.

2.a I will do below three steps:

* I would first try to get historical data from both cities administration to see that how many times that all cars go through the bridge on each weekday. The historical data would be around three month long or six month long.
* And then, after the policy is implemented, I will collect data for the same length of period time as step one to see if the total number of times increases for the each weekday.
* At the same time, I would send out surveys to drivers to get feedbacks if they are more willing to serve the other city if the toll costs will be reimbursed. Also, what can the cities do to encourage them to serve the other city? Let the drivers say what they care about most.

2.b I will use hypothesis test to verify the effective of the policy. The null hypothesis is no effect in encouraging drivers to serve both cities and the alternative hypothesis would be the opposite way that the policy does encouraging drivers to serve both cities. The significant level is 5%.

2.c I will use chi-square test to calculate chi-square value and then to get the P-value from the chi-square value. If the p-value is equal or less than 5%, we can reject the null hypothesis and accept the alternative hypothesis. If the p-value is greater than 5%, we can accept the null hypothesis.