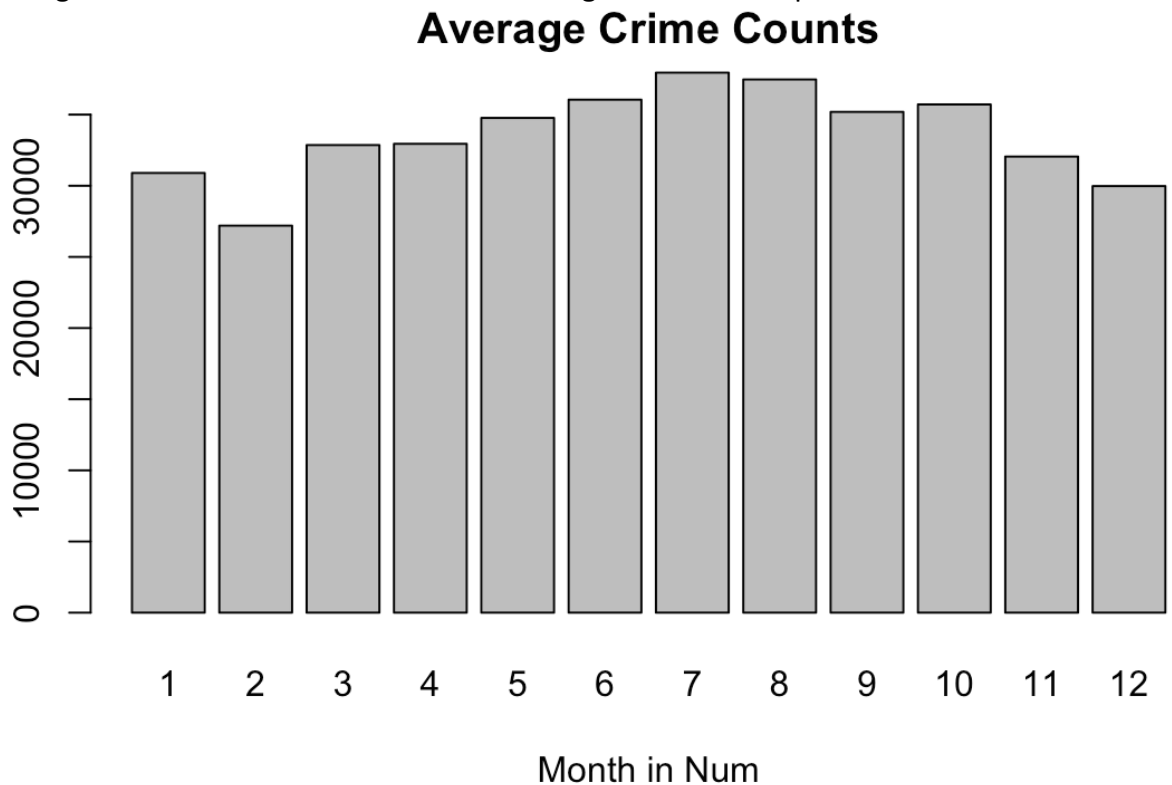


**Problem 1:**

The histogram below has the counts for the average crime events per month.



We can see that the average crime is highest during the summer months and lowest during the winter months. This is probably due to people being more outside during the summer and staying in during the winter.

**Problem 2:**

Part 1:

The three (all 10 can be seen in output) streets with the largest amount of crime raw are:

N State, W Terminal, N Michigan Avenue

All these streets are located in The Loop and it makes sense that the true number of crime is highest there due to there being more people

Part 2:

The beats with the highest correlation are:

2132	2311	0.999999
2113	2312	0.999999

## Adit Venkataraman Report

2133	2331	0.999999
2123	2324	0.999999

### Part 3:

I used a paired sample T-test to compute to see if there is a statistically significant difference in crime on an averaged per-year basis. The T-critical value is 25.5 which is very significant and there is enough evidence to reject the null hypothesis such that there is a difference between Daly and Emmanuel. This makes sense as crime is steadily decreasing in Chicago

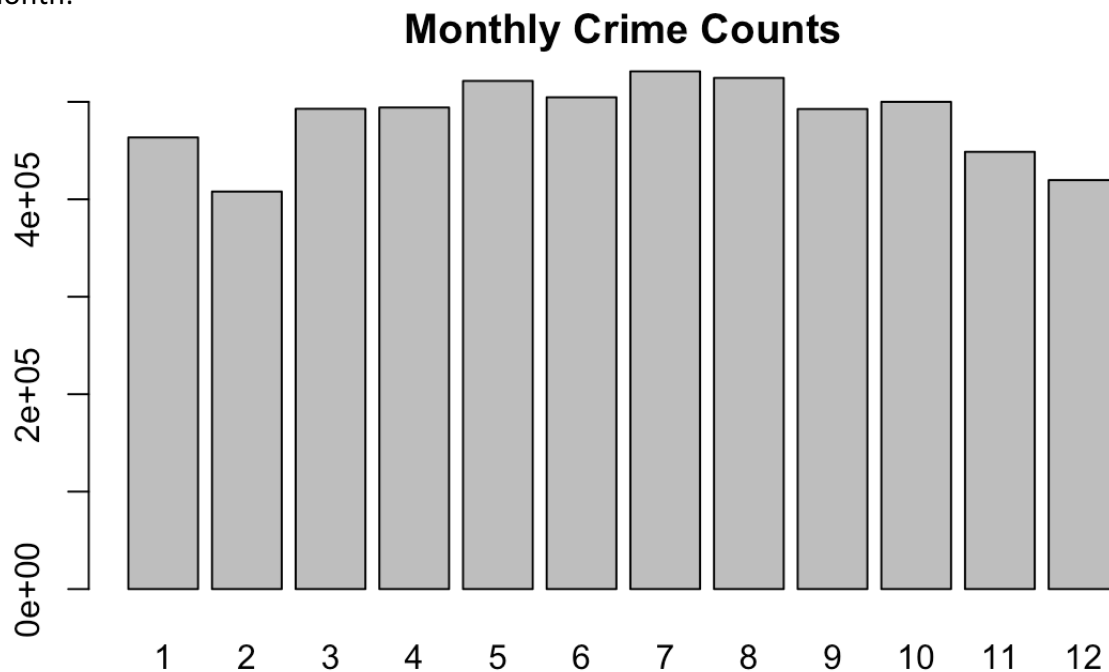
### Problem 3:

For this problem, I predicted the crime in the next week using crime from the previous week for the past three years. The MSE was 40.78 and the  $R^2$  value is 0.6433. I had trouble trying to use the model to predict future events, but from the predictions that were run on the test set, the model predicted the crime to be similar to the previous week.

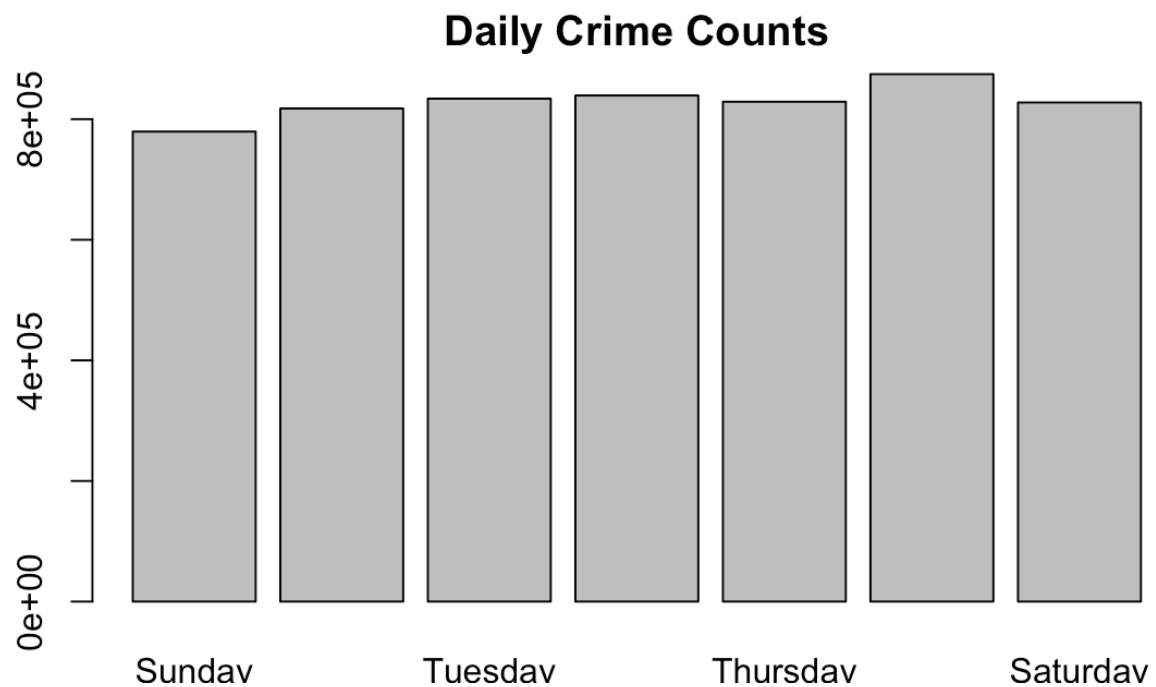
### Problem 4:

We will now look at histograms between three different granularities to see if there is a difference between time periods.

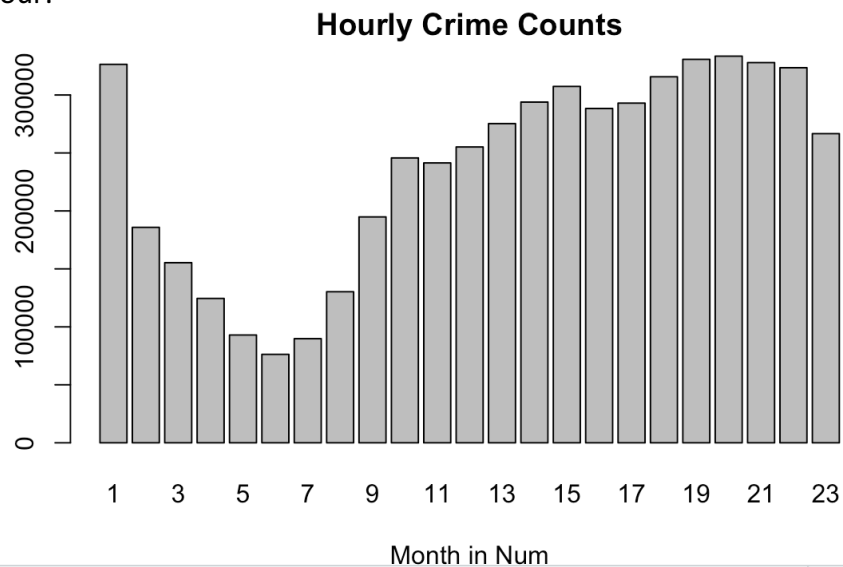
Month:



Day of Week:



Hour:



For the monthly counts, it is evident that there are more crimes in the summer than in the winter as people are hibernating for the winter or more likely spending time with family. For the day of week, there is a bit of a spike on Friday and Saturday as people are more likely to go out. For the hourly crime counts, there is a higher count of crime in the night hours but

## Adit Venkataraman Report

significantly dwindles down around 4-7 am as people are too tired to be committing crimes so early in the morning.