CSC343 Final Project Discussion

Shimon Nauenberg, Sagnik Adusumilli

Introduction

For this project we were interested in investigating cycling in Toronto. Our data was collected from open source data from open.toronto.ca. We focused on bike usage as a form of transportation and how it interacts with traffic, and the subway system. We used information about traffic, bike share stations, bike shops, bike parking spots, and street traffic, and street lights. This project involved developing a database system for all this information, and utilizing queries on the database to investigate questions. For certain questions we utilized graphical, and statistical tools in order to determine results.

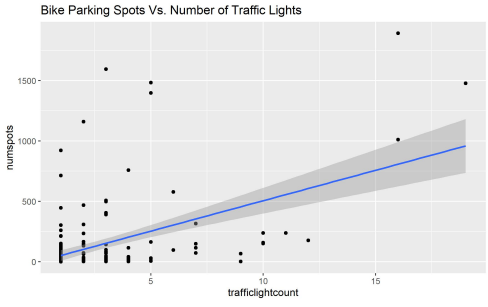
Questions

There were four primary questions which we were originally interested in investigating. The first of these questions was to investigate bike lane information in Toronto, but due to certain issues with the open source data on bike lanes we were unable to proceed with this investigation. Our Second question involved the relationships between number of traffic lights, bike parking spots, and bike share stations. Our third question involved investigating the streets which were best for cycling. We decided on a certain factors which we thought would be ideal for cyclists. Our fourth question involved the relationship between subway station usage and cycling. We wanted to know if certain cycling friendly features of subway stations were encouraging usage of those stations. Following these investigations we developed an interest in investigating the relationship between intersection traffic and bike share availability. Throughout the investigation we developed some further sub-investigations which were valuable and interesting.

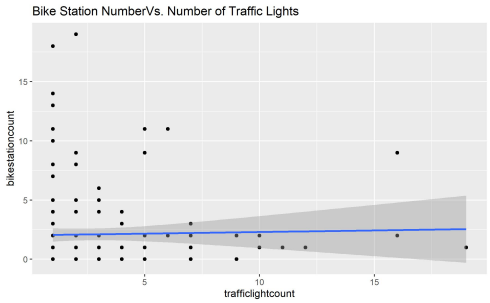
Question 1

Our first question was: is there a relationship between the number of bike parking spots and the number of traffic lights available? Further, does this relationship change once we consider the number of bike stations available? We Used the GGPLOT2 library in R, and the stats package to investigate linear models to identify relationships.

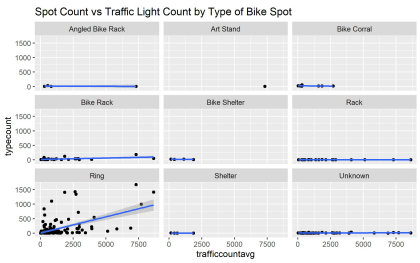
After doing the queries to prepare the data to import into R, we determined that there is a positive relationship between the number of traffic lights



and the number of parking spots, as is evident from the graphic to the right. Afterwards we investigated the relationship between the number of traffic lights and the number of bike share stations available. The graphic below shows this relationship. The graphic seems to show no relationship; however, we built a linear model, and determined from the statistical test that there is no relation between these two characteristics, (a p-value of about .7).



Lastly, we investigated if a model based on both the number of traffic lights and the number of bike stations, caused an increase in the number of bike parking spots available. For this we were unable to produce graphics, but created a linear model, and determined that when considering both number of traffic lights, and the number of bike stations, both these factors are deemed significant predictors in determining the number of bike parking spots available,(both with p-values below .05). After this investigation we become interested to see if the number of a certain type of bike parking spot had a relationship with the number of traffic lights. We created a graphic using this data, and graphically determined that there only seems to be a relationship with the "ring" type of spot, and no relationship with the other types of spots. However, there is many more "rings" than the other types of spots, as can be seen from the graphic, and so we would not make general conclusions about this information.



Question 2

Our second question we investigated the areas which streets which are best for cycling. We thought that the following metrics would be an indicator for which streets people like cyling on: the number of traffic lights, the number of bike parking spots, the number of bike shops, and the number of bike share stations available. We had thought that less traffic lights would be ideal for cyclists, as they would not have to start and stop as often. We also determined that the number of bike share stations, and parking spots would be ideal, as they made

it easier to access bikes and park the bikes. The number of bike stores was also used, as we figured that cyclists may choose these roads as they may need to access these shops in the event of a problem with their bikes. We ranked the streets by each of these metrics, and then determined the streets which had the greatest average score. After running the queries, we determined that the top five streets for cycling according to these metrics are 1) Davenport Road: score 46 2) Lake Shore Blvd W: score 62 3) Carlton st: score 76 4) King Street W: score 92 5) Elizabeth St: score 98. These results were very interesting. The high ranking of Lake Shore Blvd, and King Street W align nicely with our experience of cyclist usage of both of these streets.

Question 3

This question looked at investigating the relationship between subway usage, and the presence bike share stations and bike repair stands at the stations. We wanted to know if these cycling characteristics of the stations influenced people's decisions to use the stations. We first investigated, the average station usage in the presence of a bike share station and noticed that this did not cause an increase in usage(difference in average usage about 13000, in favor of not station). Similarly, we investigated the same metric regarding the presence of a repair stand, and recognized that this again does not cause an increase in usage(about 20000 difference in favor of no repair stand). We thought that perhaps, terminal stations lacked these attributes and had a high average usage which could skew the results. Further, it is possible that cyclists did prefer these stations, but not enough people cycle to make a meaningful difference.

Question 4

For our final question we where interested in looking at traffic at intersections in which neither street has bike share stations on it. We determined that there was a meaningful difference between the streets which have a bike share station(average traffic 316.35), and those which do not (526.8). This difference was very interesting. It could indicate that the use of bike share stations is causing a meaningful decrease in the average amount of traffic. This could be a strong support in policy development to implement more bike share stations at intersections which have a high amount of traffic.

Conclusion

We investigated four different questions about cycling in Toronto. Some of these questions led to surprising results. These were all very interesting questions and provided information that can be valuable in policy development .