# Question:  
# 1: Which light rail stop has the most pedestrian traffic?

Calculate total riders by stop looking for all quarters and then years. We could have two results here to show on a quarterly basis this stop has the most traffic and then etc. for year.

A bar graph may be a good visual representation but that will only work if we use something like the top 5 stops so we do loose readability if there are too many values on display.

We could even break this down by the quarters per year in their own graphs to show a few graphs and possibly identify any seasonal trends.

# 2: What are the busiest light rail locations annually?

Use total riders by year calculation here again.

Could reflect in a pie chart to show total of the overall rider population in our study.

\*\*Something to note we may be trying to answer the same question twice in 1 & 2. Curios what you guys think on Sat if I am being too general here.

# 1: How many bike stations are w/i a 1 mile radius of a transit station?

Convert lat/long into a list for both bike data and light rail data to determine stops locations.

Then compare lists using loops to determine what stops/transit centers have overlapping locations within the set radius (1 mile). Maybe we extend this to look at 5 miles as an additional step, even 10? 10 may be irrelevant but curious on teams’ feedback.

This would be a perfect use of a heat mat to show the relationships of bike stations w/i proximity to one another.

# 2: During a high-volume period of light rail usage do we also see a correlation in bike share usage?

From the above variables created for light rail rider peaks, break down by quarter and see if seasonal changes exist in relation to the bike usage. Since we do not have biker share rider numbers, we can use number of locations within a 5 mile radius of the highest volume light rail stations to see if there is a high or low number of bike share stations.

If we find that the number of bike share stations are high in that 5-mile radius we could assume light rail use is comparable to bike share use. If the number of stations is low in that radius, we could conclude that bike share use is not correlated to high light rail usage times.

We could use the creation time of the bike station to compare to the quarterly rider numbers to see the changes over time.

This could be a good use of a scatter plot I think. Map the different locations and then maybe implement a heat signal(not exactly sure what those are called) map and show the radius on

# 3: In high profit quarters for light rail are there any reductions in traffic?

For this we could again pull in the variables created for light rail riders and compare quarterly or annually how that increases or decreases based on the number of traffic slowness or traffic collisions. I forget exactly which one we were using but either one could make a good hypothesis.

I believe we will need to convert the traffic data to lists and/or dictionaries so we can then see if the periods with high slowness or collisions are also times when light rail riders number increase.

I believe a bar graph would be best here but curios on group thoughts.

# 4: On days where we see higher traffic do we see an increase in light rail use?

This is an extension of the above questions and we could probably easily answer this one with the data we put together on the above question.

Graph here could be a line graph with 2 different data findings overlain on one another.