# HOUSE PRICING COMPARISON

Sherana Syed, Avinash Dhawan, Jasmeet Aujla and Raj Periyasamy

# MOTIVATION & SUMMARY SLIDE

- As we are aware that the house pricing and cost of living has increased over the last 5 years in the cities surrounding the Greater Toronto Area; we decided to study a couple of the underlining factors such as; school scores and commute times in 4 cities.
- Our objective was to study the correlation between house pricing to school scores as well as commute times to Toronto. In order to obtain this data, we studied the school scores and the commute times from the Go Train Stations in Brampton, Mississauga, Oakville and Burlington.
- There are some strong and weak correlations between the two factors that we analyzed in comparison to housing prices.

# DATA QUESTIONS

Do school ratings affect house pricing in these cities?

Does the commute time from the Go Train stations to the Union Station in Toronto affect house pricing in these cities?

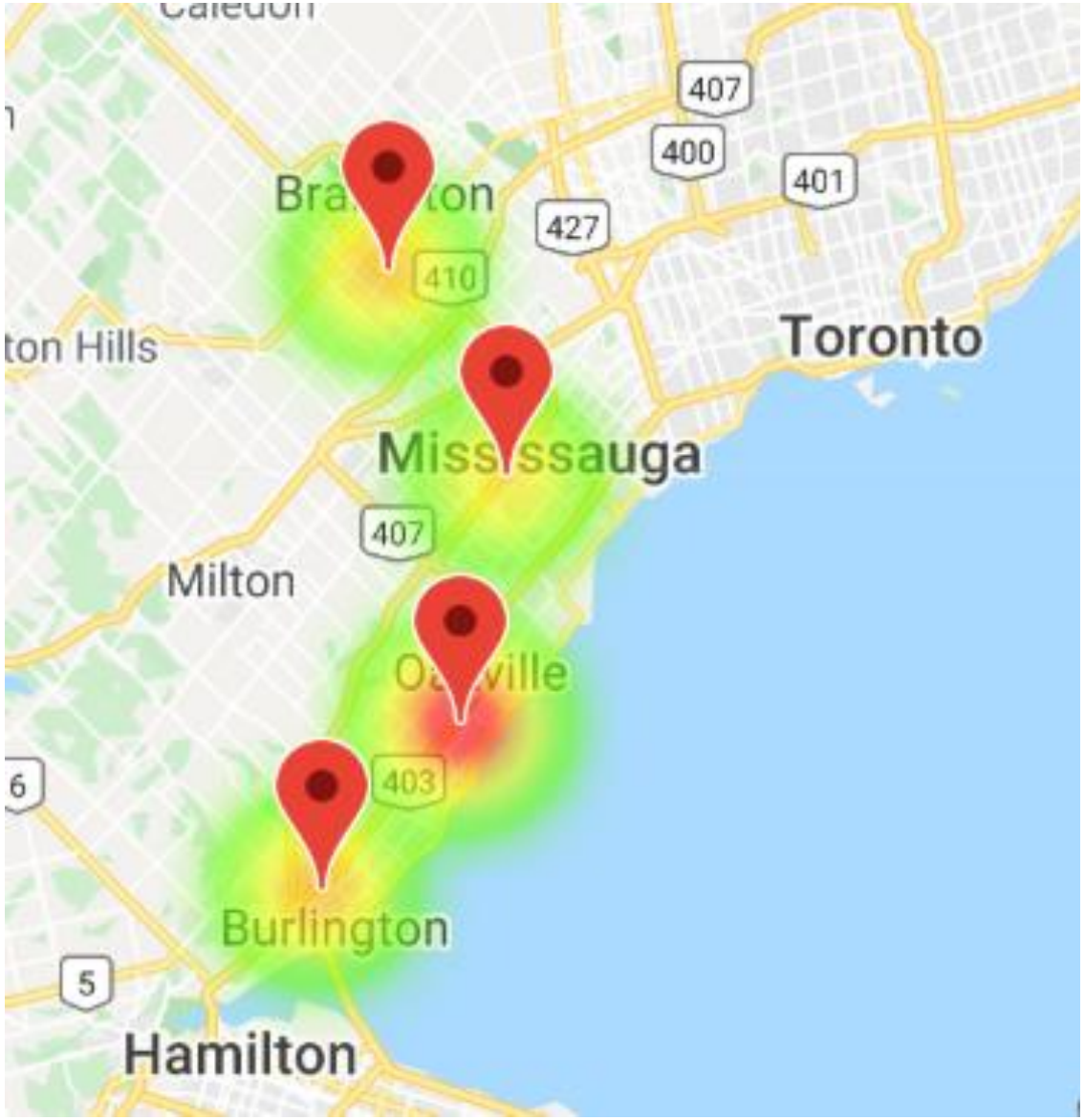
What is the correlation and what are some other factors we could have studied?

# CLEANUP AND EXPLORATION

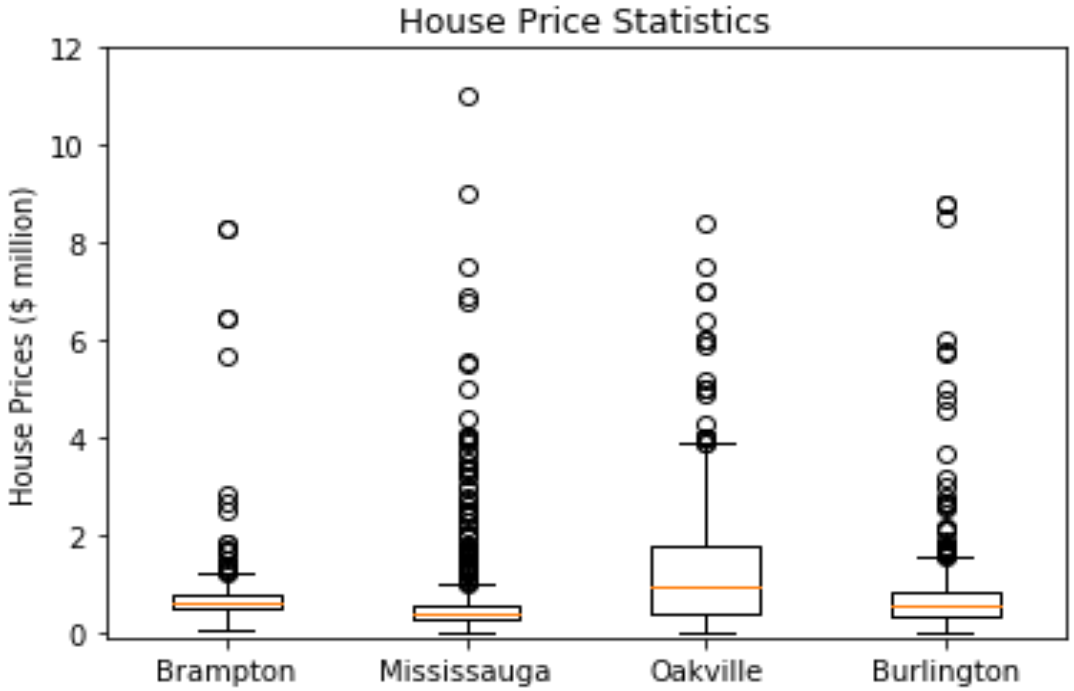
- Kaggle was used to obtain the data set for Ontario house sales in July 2016. The data was cleaned up and contained the home address, price, neighborhood, latitude and longitude. We constructed four different data frames by slicing the data from the original data frame using .loc and concentrating on Area Name. After getting data for each city, we used the describe() function to get the statistics for prices for each city and merged that into one data frame so that we can plot the data points from this final data frame.
- We also obtained the elementary and high school scores data of each city from the Fraser Institute. As this was not available as a data set, we had to use HTML scraping to create the data frames of each city.
- For the commute times, we took all the GO train stations from the Go Website in Brampton, Mississauga, Oakville and Burlington into consideration. We looked at the morning and evening commute times to and from the Union Station for each of the stations in these cities. The average commute time from each station to the Union Station was determined. These averages were further combined to generate an average commute time from each city.
- After we merged all the data, we used scatter plots and box plots to compare both the commute times and the school scores with the house prices.
- We also used Google API to create a heatmap to compare all four cities.

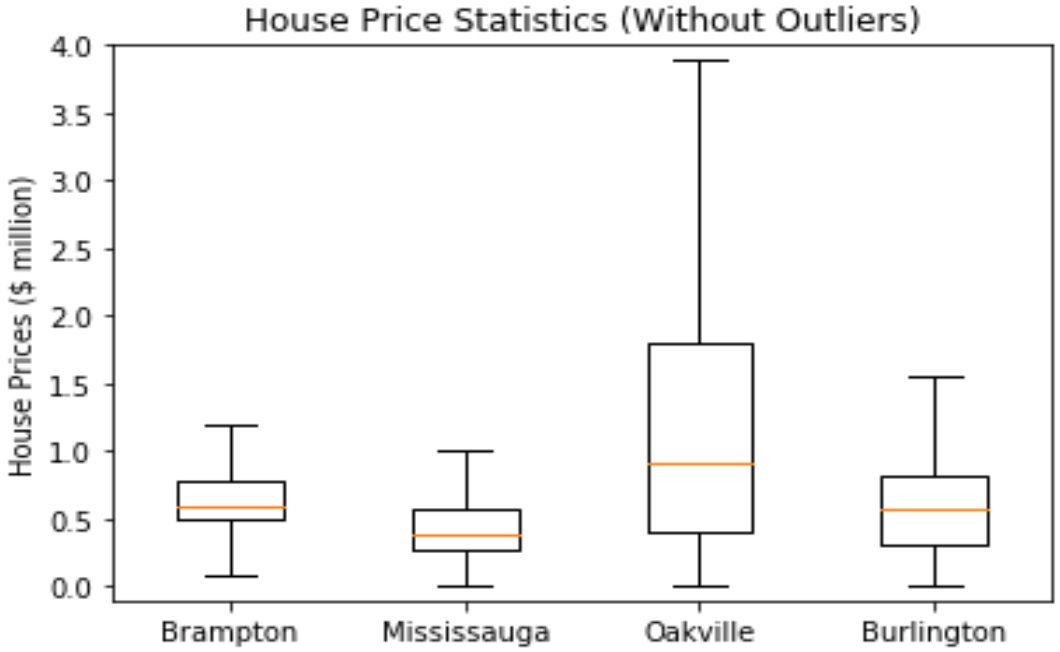
# ANALYSIS

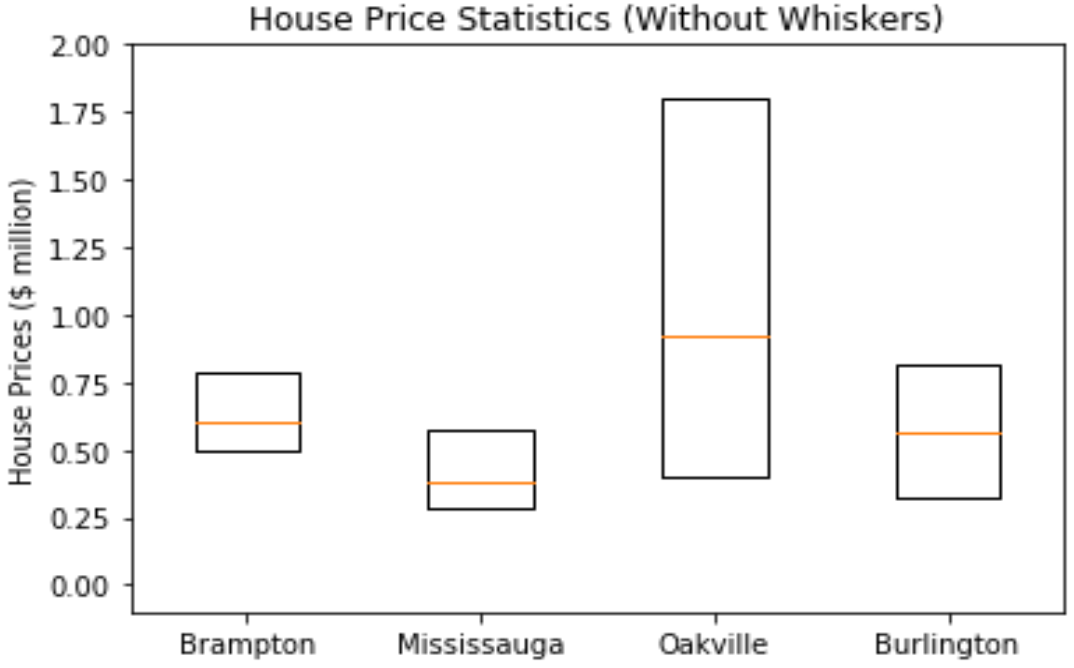
- The Anova test value (2.0e-24), has confirmed that the variance in the house prices is statistically significant.
- The average house prices in Oakville are higher than the other 3 cities. House pricing is proportional to the average school scores in Oakville.
- The average commute times are statistically insignificant to the average house prices in all 4 cities.

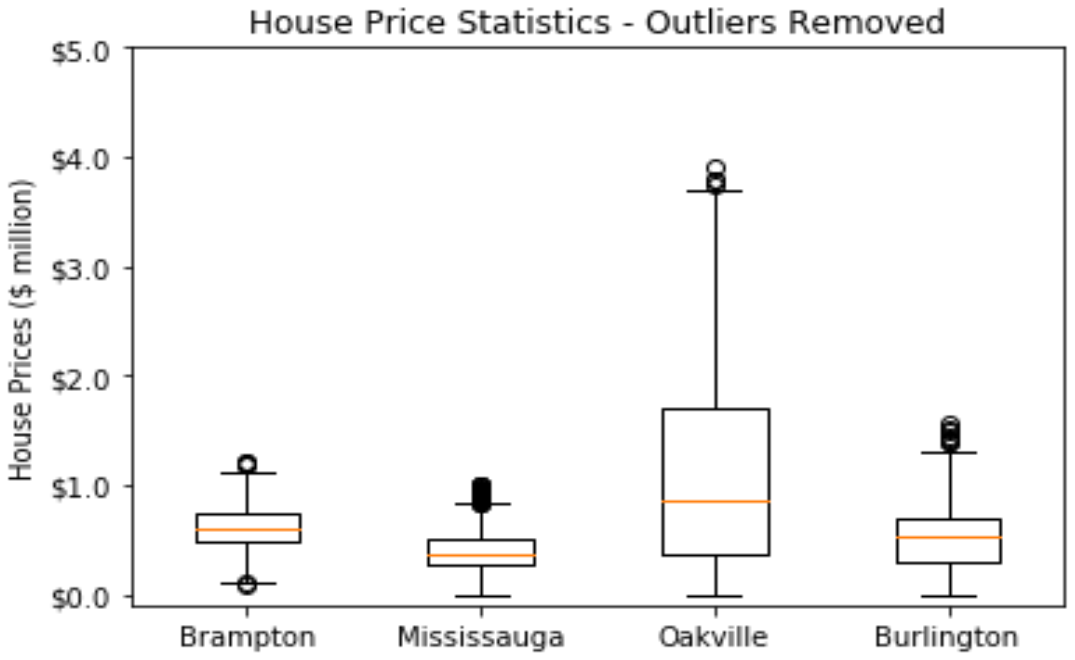



# HEAT MAP OF ALL CITIES

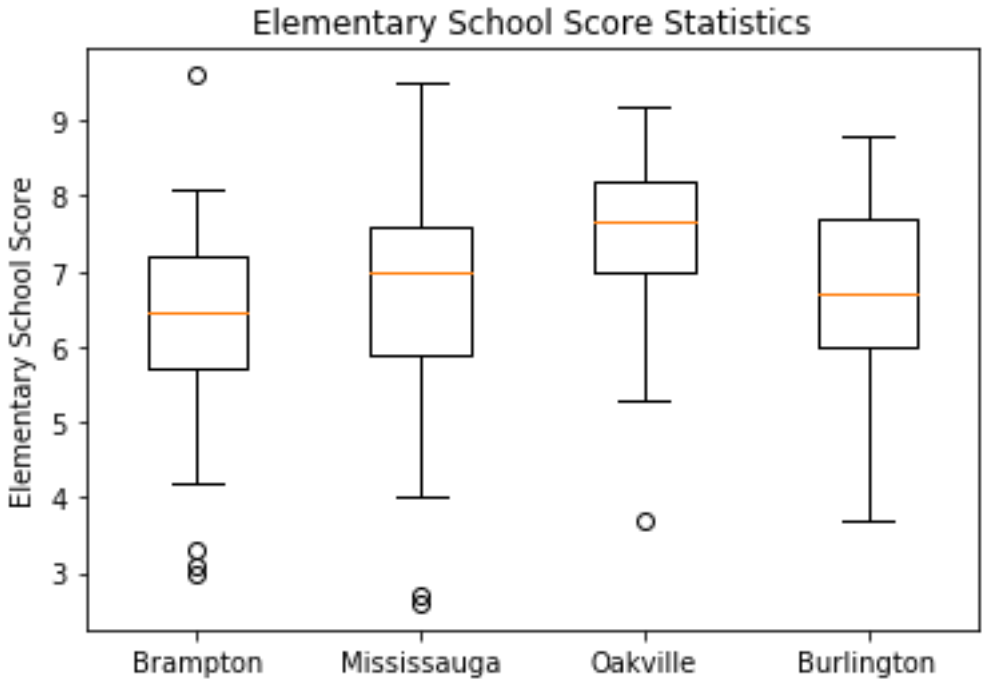

House Price Statistics
House Prices ($ million)
0
2
4
6
8
10
12
Brampton
Mississauga
Oakville
Burlington

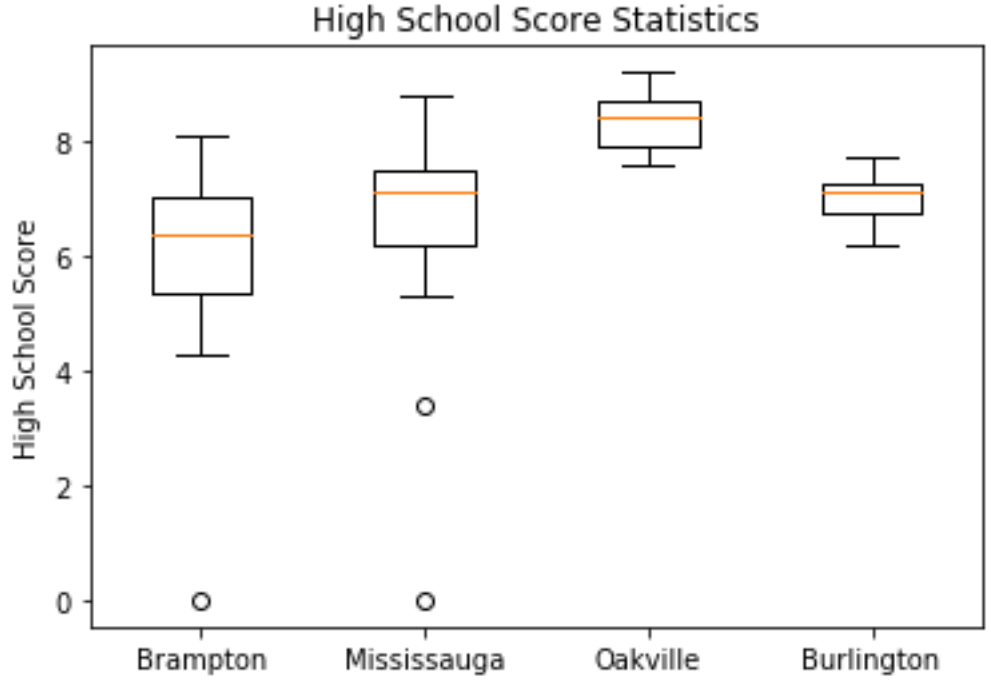

House Price Statistics (Without Outliers)
House Prices ($ million)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Brampton
Mississauga
Oakville
Burlington


House Price Statistics (Without Whiskers)
House Prices ($ million)
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
Brampton
Mississauga
Oakville
Burlington

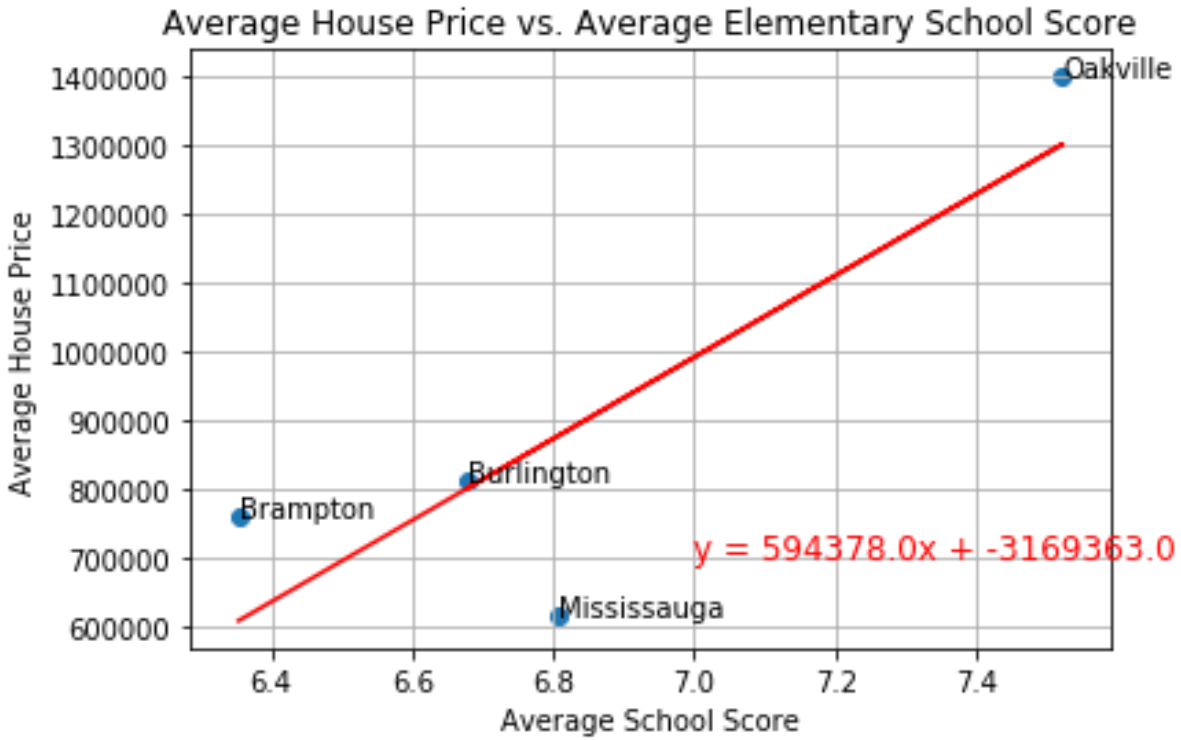

House Price Statistics - Outliers Removed
House Prices ($ million)
$0.0
$1.0
$2.0
$3.0
$4.0
$5.0
Brampton
Mississauga
Oakville
Burlington

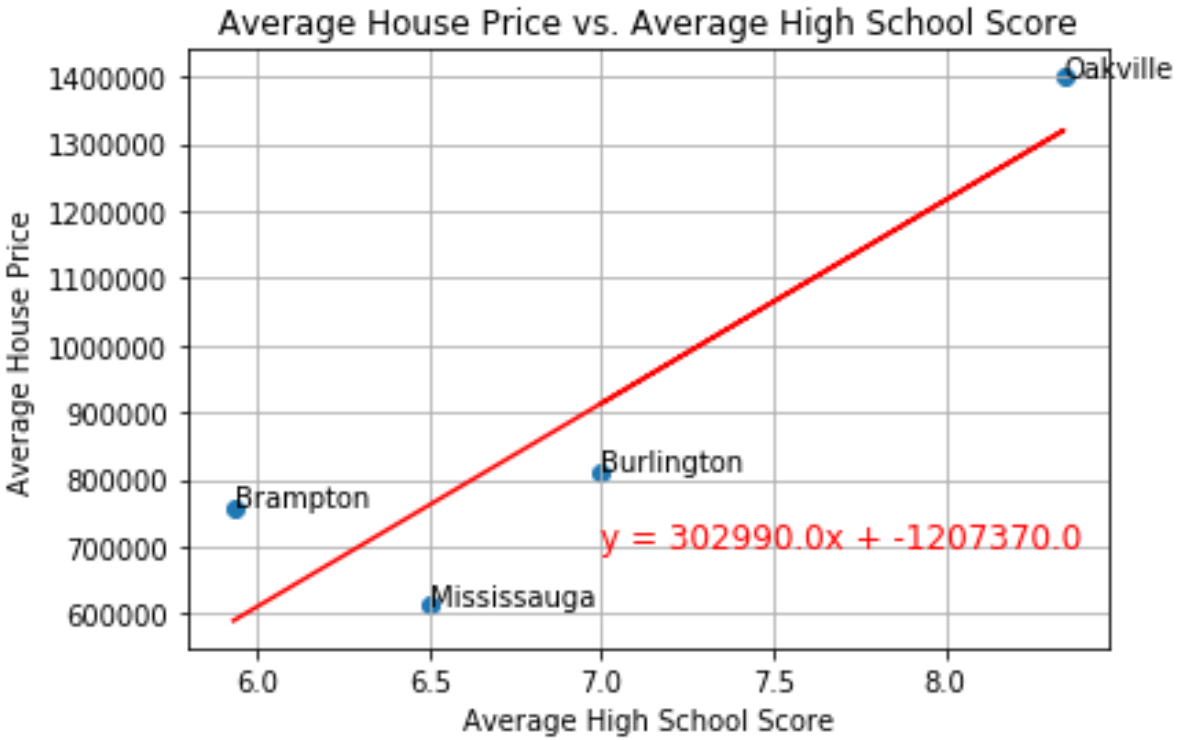
# SCHOOL STATISTICS

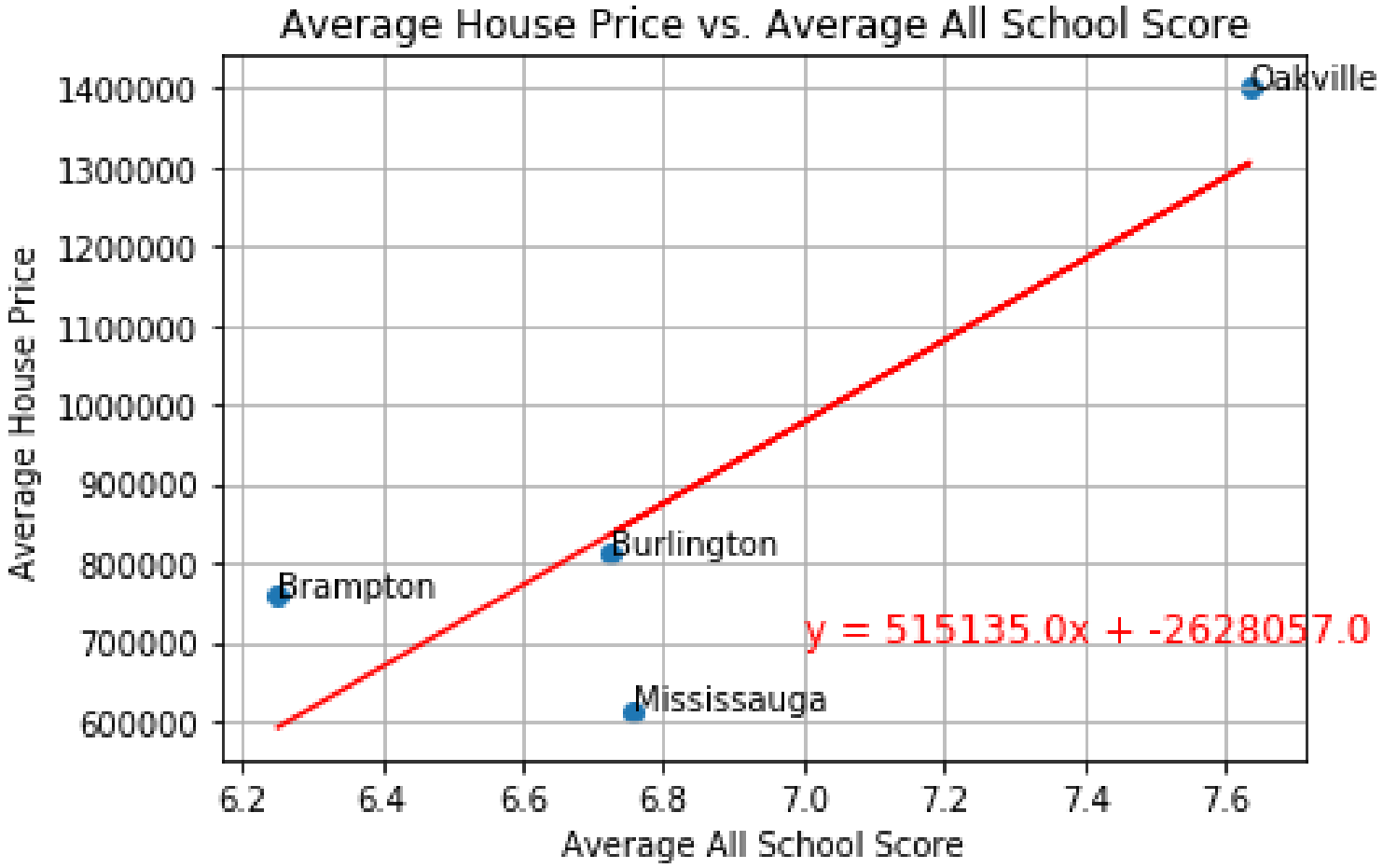


# SCHOOL SCORE VS. HOUSE PRICE

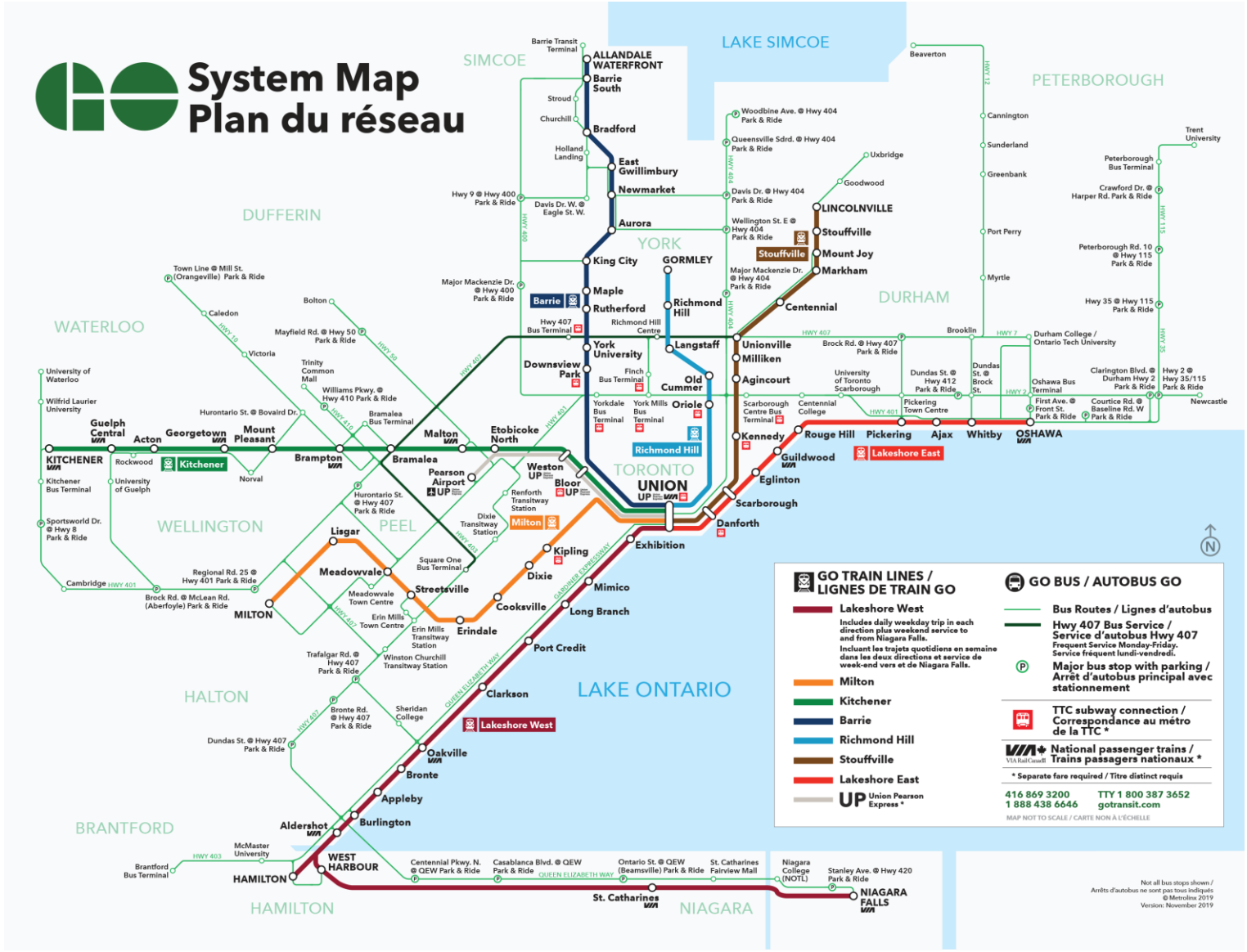


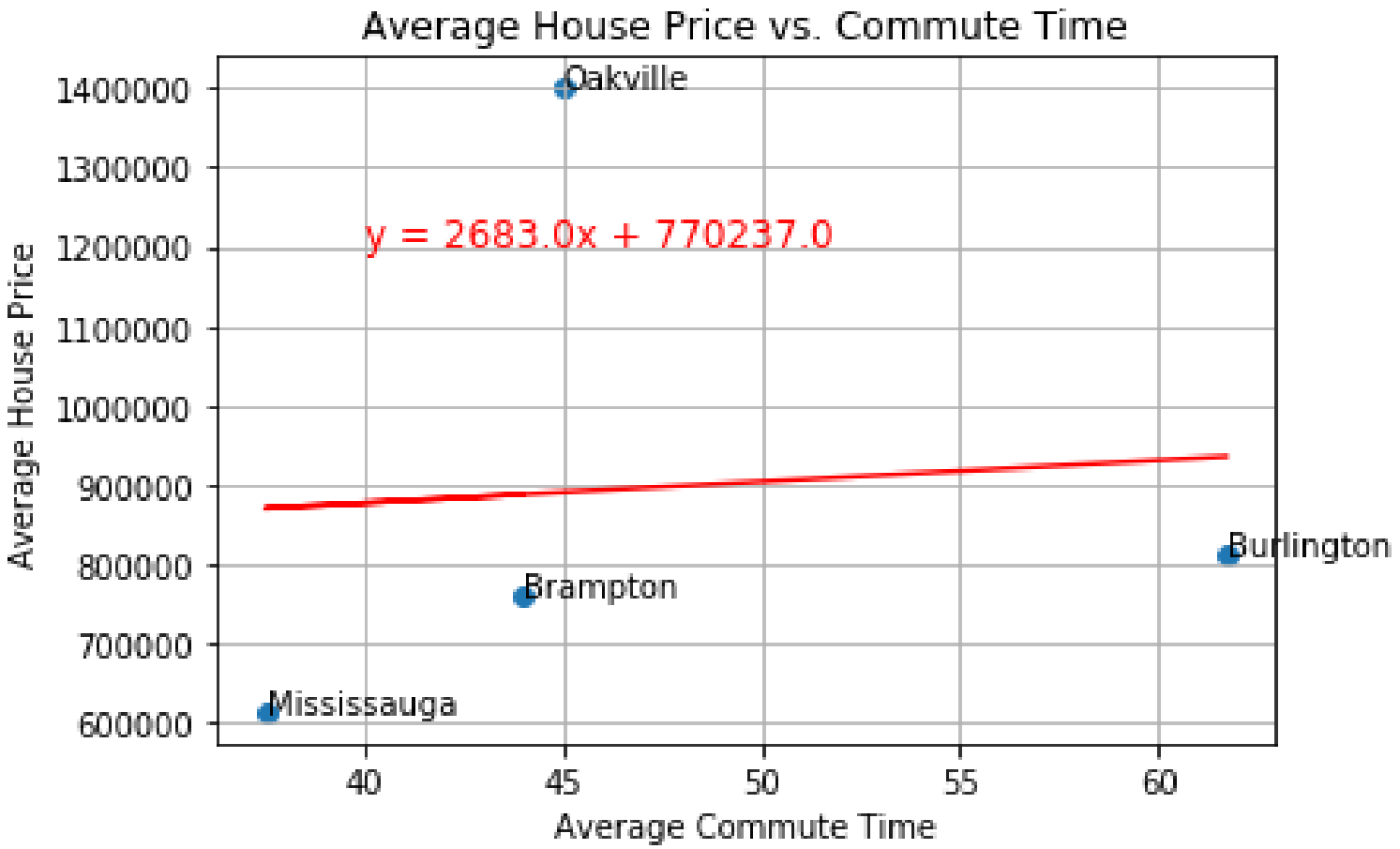






# AVERAGE HOUSE PRICE VS. AVERAGE ALL SCHOOL SCORES

# GO TRAIN SYSTEM MAP






# AVERAGE HOUSE PRICE VS. COMMUTE TIMES

# DISCUSSION

- The reason we thought of comparing this data was to see if there is a correlation between school scores vs. house prices and we found that there is a strong correlation between the average house prices and the average school scores in all cities except Mississauga.
- We suspect that Brampton, Oakville and Burlington have more families and school scores may be more important in these cities vs. Mississauga. The factors that may affect prices in Mississauga could be that there are more young professionals who want to be closer to downtown as well as families with younger adults that are out of the school system.
- In terms of the commute times, we determined the following:
    - The commute times are the highest in Burlington
    - The lowest in Mississauga
    - Oakville and Brampton have similar commute times

# POST MORTEM

- We obtained data sets from a few different resources such as; the Go Train website, the Fraser Institute website, Kaggle for the house sales data set and Google Maps.
- The house pricing could have also been impacted by the house sizes in each city and neighborhood. This is further research and analysis we would obtain to see the trends if we did not have time constraints or lack of access to more data compilation.
- The other factors that we could study are demographics, walk scores, grocery store access, mall access, parks and recreation, entertainment, restaurants, crime rates, what transit options residents are using and how many residents work in Downtown, Toronto. All these factors could also impact house prices in a city. This would be an interesting project to further study to observe these trends and therefore, we cannot conclude that the commute times and school scores are the only factors involved in a house purchase decision.

THANK YOU FOR LISTENING TO OUR PRESENTATION.
QUESTIONS ARE WELCOME.