TRENDING HOUSING PRICE IN TORONTO AREA AT 2018

By: Sripranavan Sritharan
No.304378

One of the largest investments that most individuals will make in their lifetime is the buying of a house. In these times of "hot" real-estate markets, the more meaningful information that individuals have on the predictors of price, the better.

1. Buyers want to get the most value for their money as well as getting all the points covered off on their "wishlist' for features they want to have in a home.
2. Sellers would like to have top dollar for their home without having their home sit on the market for an inordinate amount of time.
3. Real-estate agents want to help their clients in both groups reach their goals in the real-estate market.
4. Real-estate boards and municipalities are very interested in their area's real-estate market trends.

Our data consists of home attributes taken from Toronto, Canada an area that includes Seattle. It includes homes sold in 2018.

The Goal of this analysis is to answer the following question:

"Can the sale price of a house be predicted accurately based on physical characteristics such as the Population, House-hold, House-hold Income, Education Status, Age of living area (0 to 4 'Child' and 25 to 64)?"

Our data was acquired from own collection data :

Attached two CSV file and such as,Toronto2018 and Location_Toronto

On assessment the following was found:

- The Toronto2018 data has 96 rows (observations) and 18 columns and Location_Toronto data has 104 rows (observations) and 4 columns
- All features in the data are numeric with the exception of date, which is an object.
- There are no missing values in this data and no unvalued numbers .

Main Heading and Factors from Toronto2018 data sets:

Name

Forward Sortation Area ID - IN

Number of Basics | Total Population, 2018 - (1)

Number of Basics | Total Households, 2018 – (2)

Households by Income (Current Year) | Median Household Income (Current Year $), 2018 – (3)

% of Household Population 25 to 64 Years by Educational Attainment | Household Population 25 To 64 Years | No Certificate, Diploma Or Degree, 2018 – (4)

% of Household Population 25 to 64 Years by Educational Attainment | Household Population 25 To 64 Years | University Certificate, Diploma Or Degree At Bachelor Level Or Above, 2018 – (5)

Number of Household Population by 5-Year Mobility | Household Population For 5 Year Mobility Status, 2018 (6)

Number of Population 15 Years or Over by Marital Status | Total Population 15 Years Or Over, 2018 – (7)

Number of Male Population by Age | Males, 2018 – (8)

Number of Household Population by Total Immigrants and Place of Birth | Total Household Population, 2018

Number of Total Population by Age | Total Childern Population | Total 0 To 4, 2018 – (9)

Detached Housing Prices| Median ($),2018 - OUT

Semi-Detached Housing Price | Median ($),2018 - OUT

Condo Housing Prices| Median ($),2018 - OUT

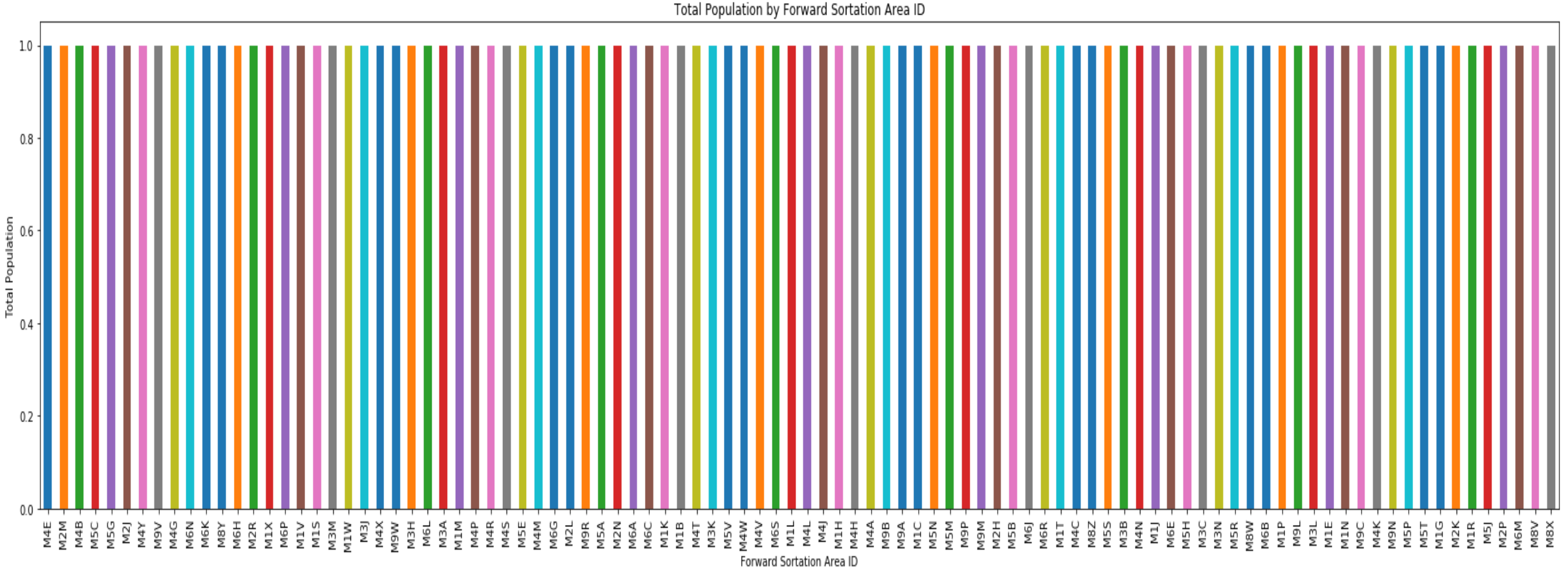
Main Heading and Factors from Location_Toronto data sets:

FSA

Latitude

Longitude

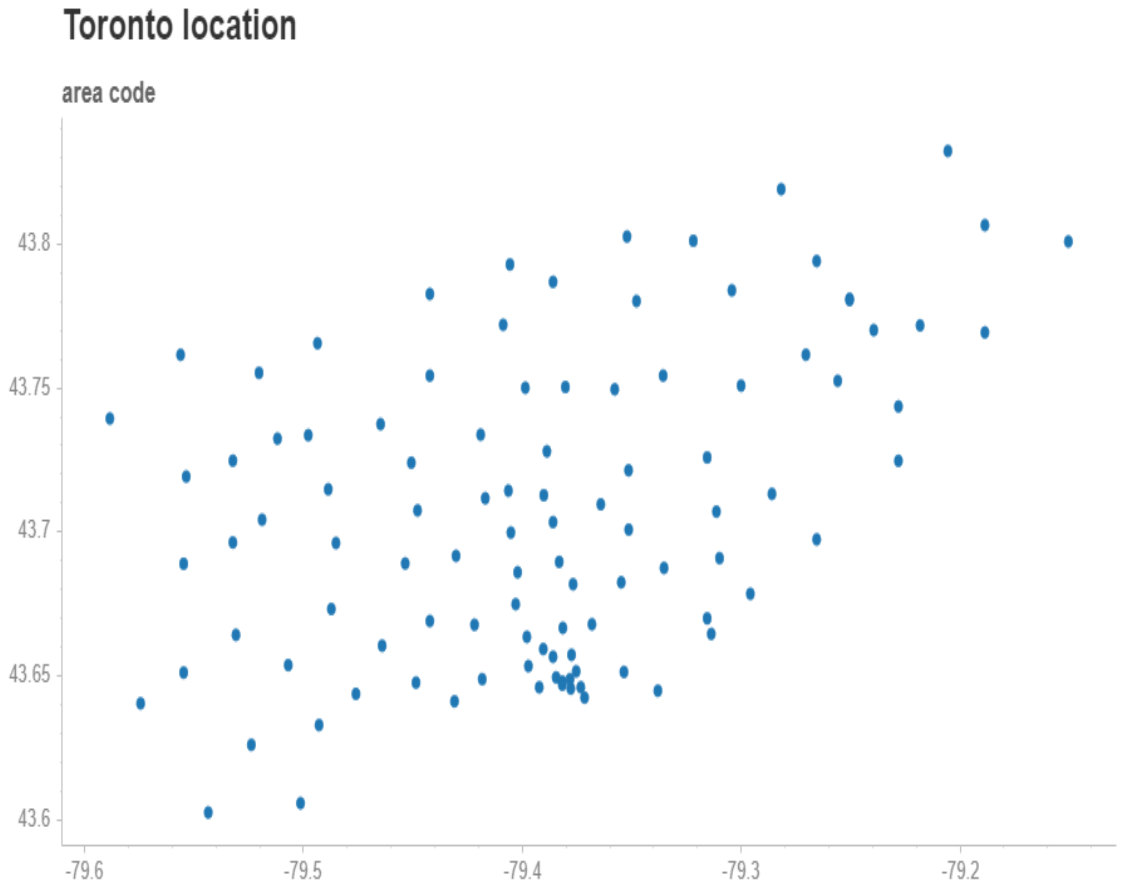
Place Name

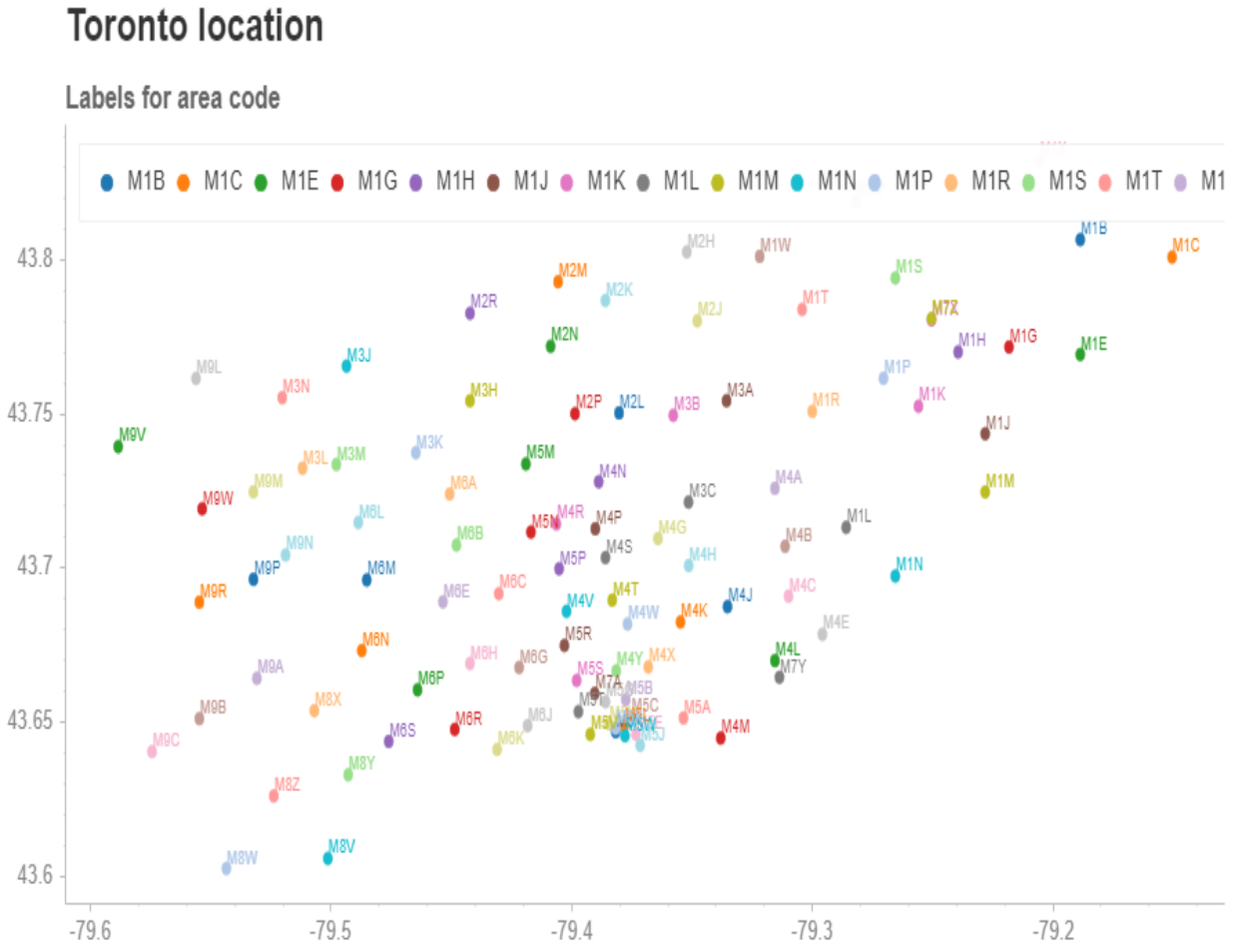


Business Problem | Data Preparation | **Model Development** | Model Evaluation | Solution Deployment

Modelling Build 1:

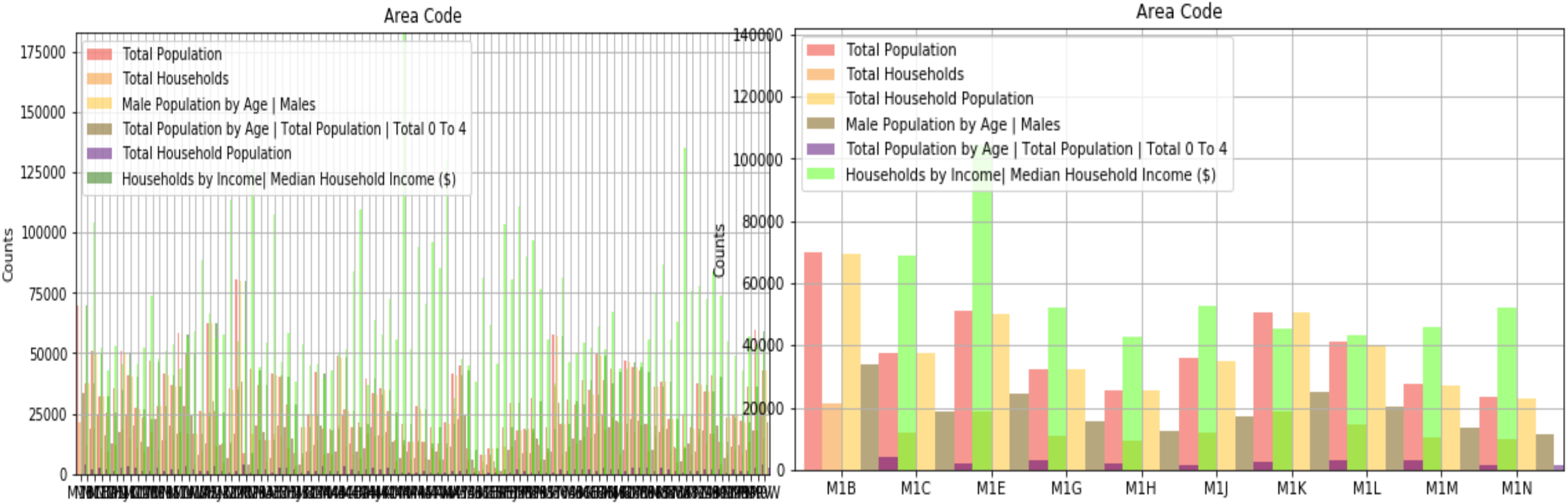
Location Mapping ### Report





Modelling Build 2:

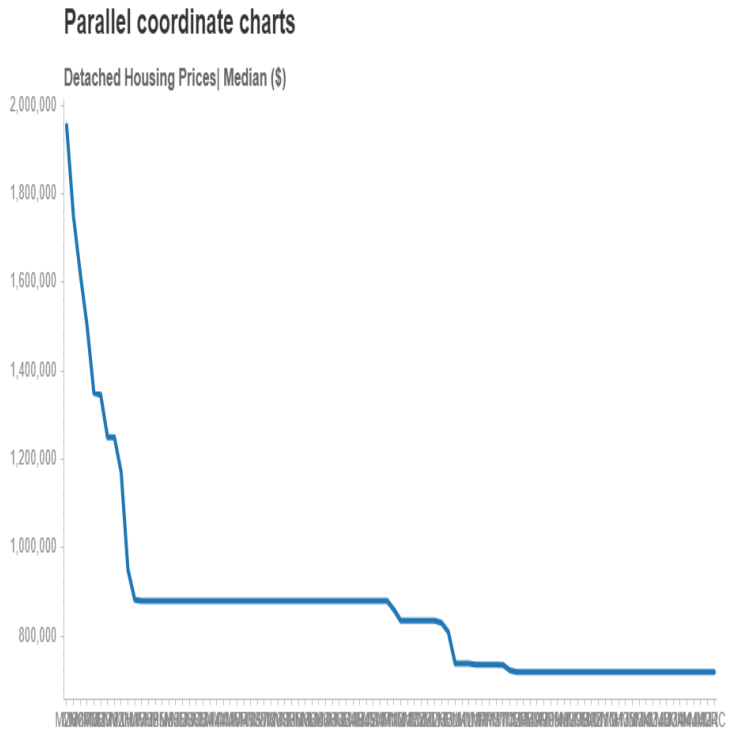
Plot the chart for all factors with Toronto Area Name ### Report

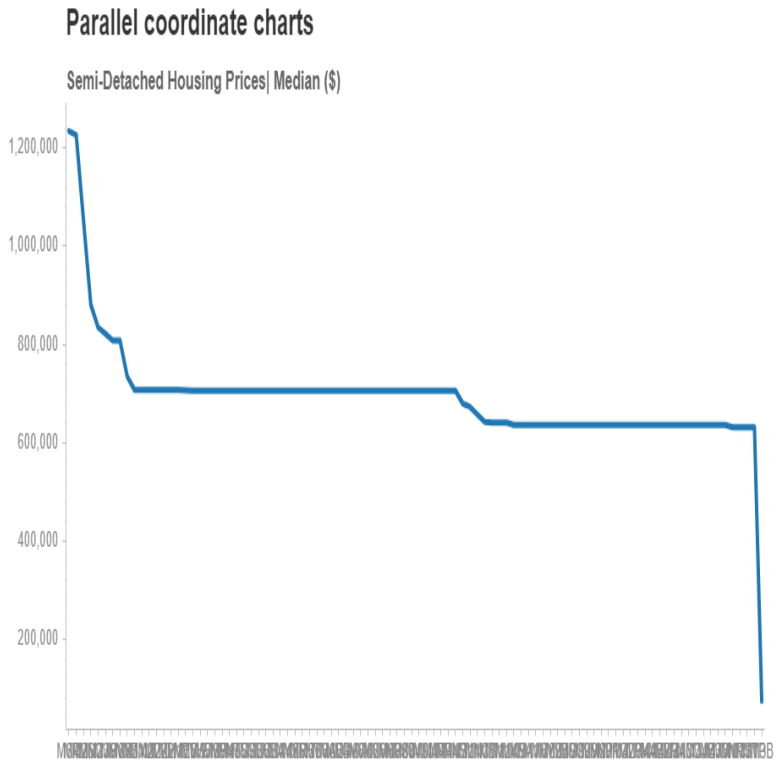


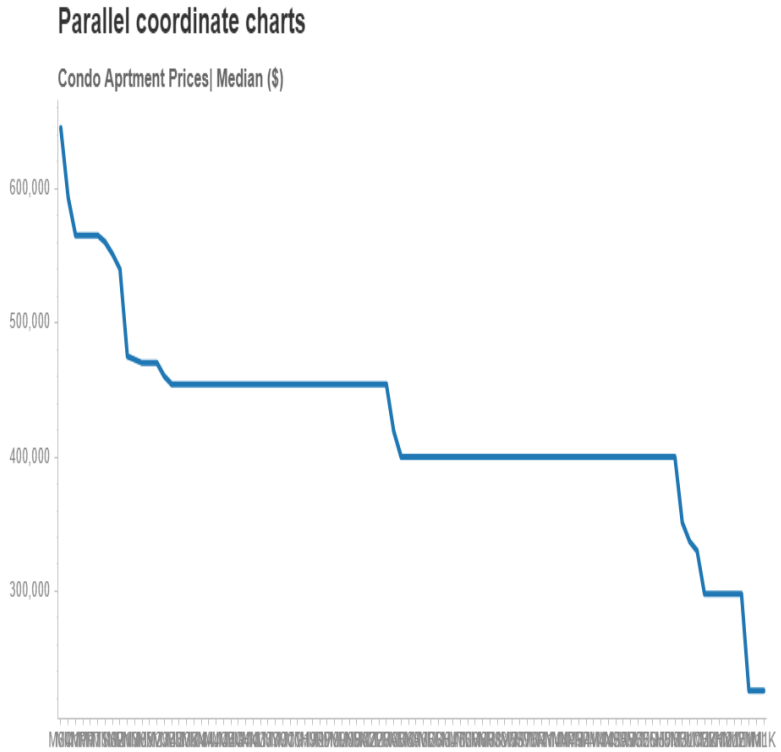
Modelling Build 3:

Plot the Housing price in Toronto Area### Report

Detached Housing # Semi-Detached Housing # Condo Apartment







Analysis the all factors using scatter plot model and Calculate some R-squared value using R programing

Out-Put	In-Put	Plot model	R-squared Value
Housing Price	Total Population	Bar Chart	>regmodel = lm(area name~Household Income,Toronto2018) > summary(regmodel)
Detached House	Total Households	Bar Chart	
Semi-Detached	Households by Income	Bar Chart	
Condo Apartment	Male Population by Age	Bar Chart	
	Childern Population	Bar Chart	
	25 to 64 Population	Bar Chart	
	25 to 64 Years by Educational Attainment	Bar Chart	

- Compare the results and analysis; the best model output

Model Evaluation and Analysis Summary

- Highest impact on predicted price is Total Population, Household and Household income and total population and household are highest M2N, M1B and M2J and Highest household income M1C. Household income has reverse and unpredictable effect
- Highest housing price area M2N.
- Avg housing price increase is with Total Population and increase with Household.
- Linear regression with 2 features gives the most accurate price prediction (75%) and 1 feature is negative regression.

The groups that would be interested in this report are:

1. The general public, house buyers and house sellers alike.

2. Real-estate agents.

3. The county and state real estate boards.

4. The local municipal government.

This analysis is based on the factual physical property features only. Emotional, social and financial position of the parties involved are also factors to be considered. Towards this end,additional demographic and income data would be useful for this prediction analysis.

This model should be updated on an annual basis.