GREATER TORONTO AREA

REAL-ESTATE VALUE ESTIMATION

DONE BY: ARVIND SUBRAMANIAN

SCOPE

- Background
- Data Used
- Methodology
- Results & Discussion
- Conclusions

BACKGROUND

- Home purchases are key financial decisions
- Buyers usually succumb to sales-pitch by realtors
  - Surface-level knowledge of real-estate
  - Cumbersome to trawl through data
- Waste of $
- Higher physical and financial risk

CHALLENGE: Provide quantitative and qualitative analysis for buyers in the Greater Toronto Area (GTA) to estimate the value of a house quickly and accurately.

DATA USED

- Real-Estate Sales - Complete record of all houses sold in the GTA for 2019, from Zoocasa.
- GTA Neighbourhoods - Names and postal codes of GTA neighbourhoods, scrubbed from a Wikipedia page.
- Nearby Venues - Dataset produced by feeding above dataset to the Foursquare API using 'get venues/explore' function.
- Crime Rates - 2019 crimes data from Toronto Metropolitan Police database.
- Fire Incidents - 2019 fires data obtained from Toronto Fire Department database.

METHODOLOGY

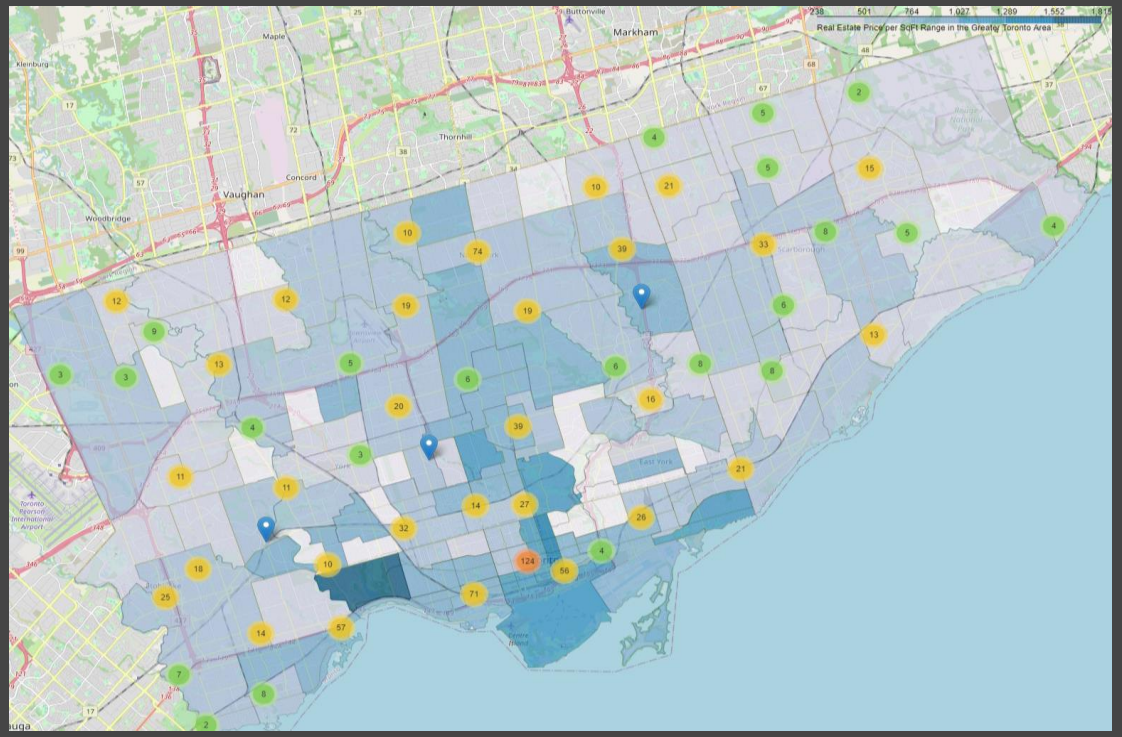
- Quantitative
  - 5 Variables
  - Analyse and visualize each variable using relevant plots
- Qualitative
  - 3 Variables
  - Analyse using Machine Learning Model (K-means Clustering)
  - Visualize with plots and heatmaps
- Price Calculator
  - Test and evaluate regression models

RESULTS & DISCUSSION

Chloropleth Map of GTA neighbourhoods

- Birds-eye view of the price-range across GTA
- Allows for a quick budget check to scope further search

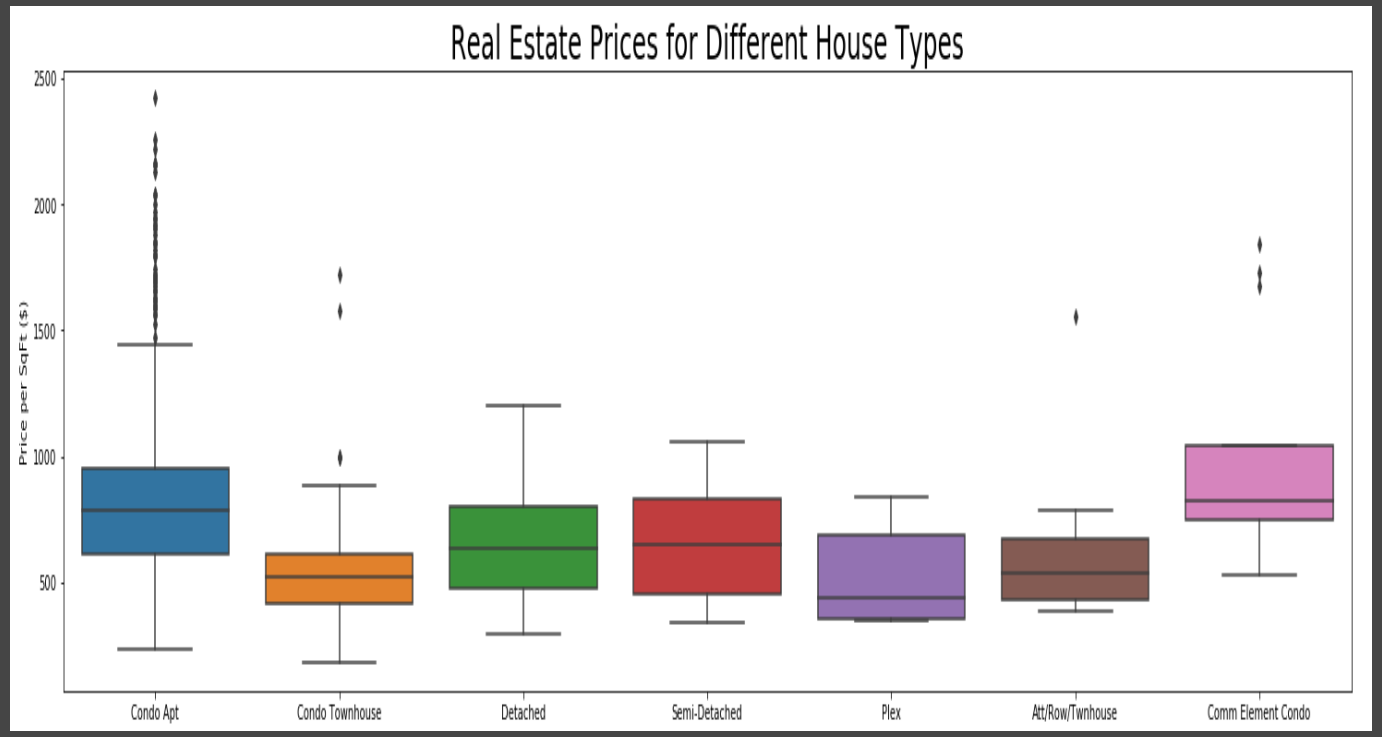
Plotted each listing for further details (zoom-in)



RESULTS & DISCUSSION

Box Plot of price per square foot (p/sqft) range for varying home types.

- Condo apartments and common element condos most expensive
- Condo townhouses most consistently priced
- Plex is cheapest
- Condo apartments have several high outliers



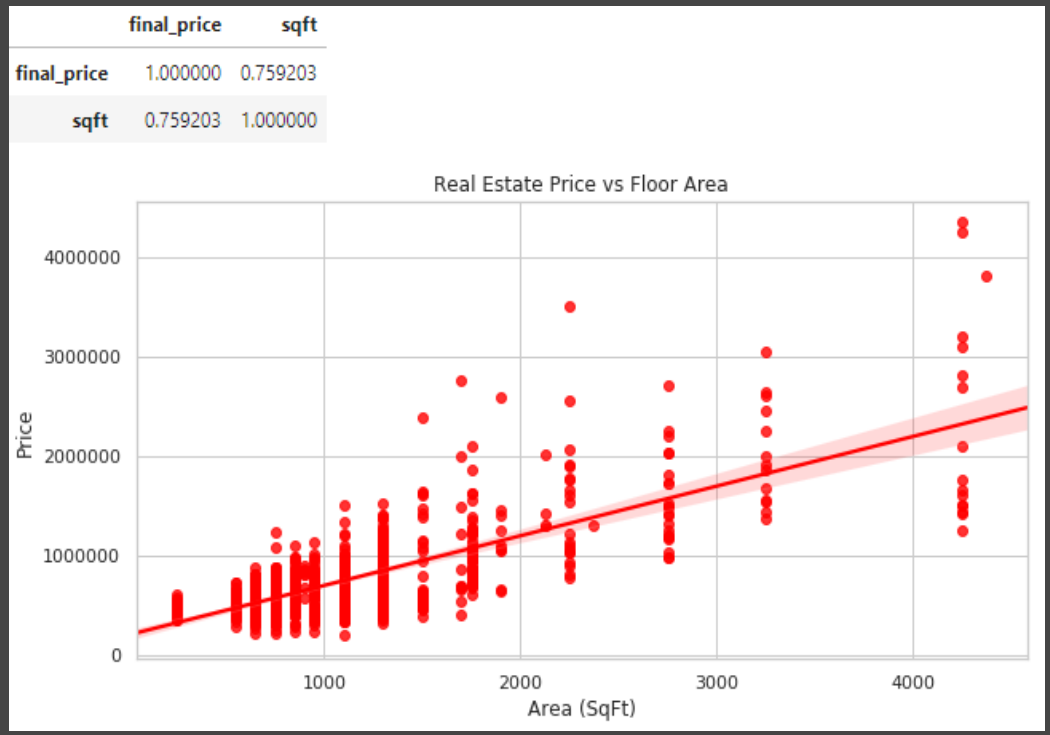
RESULTS & DISCUSSION

Good correlation between price and area of the houses

- 0.76

Size likely to have the biggest impact on price across all neighbourhoods.

	final_price	sqft
final_price	1.000000	0.759203
sqft	0.759203	1.000000

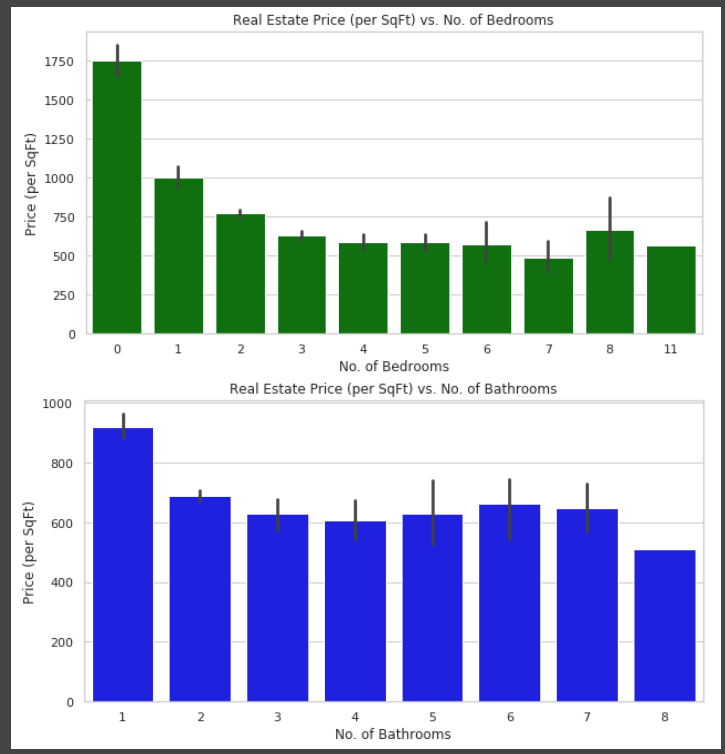


RESULTS & DISCUSSION

P/sqft generally decreases with increasing bedrooms

P/sqft does not change much with varying bathrooms

- 4 Bed, 4 Bath home has the best p/sqft or value for money.



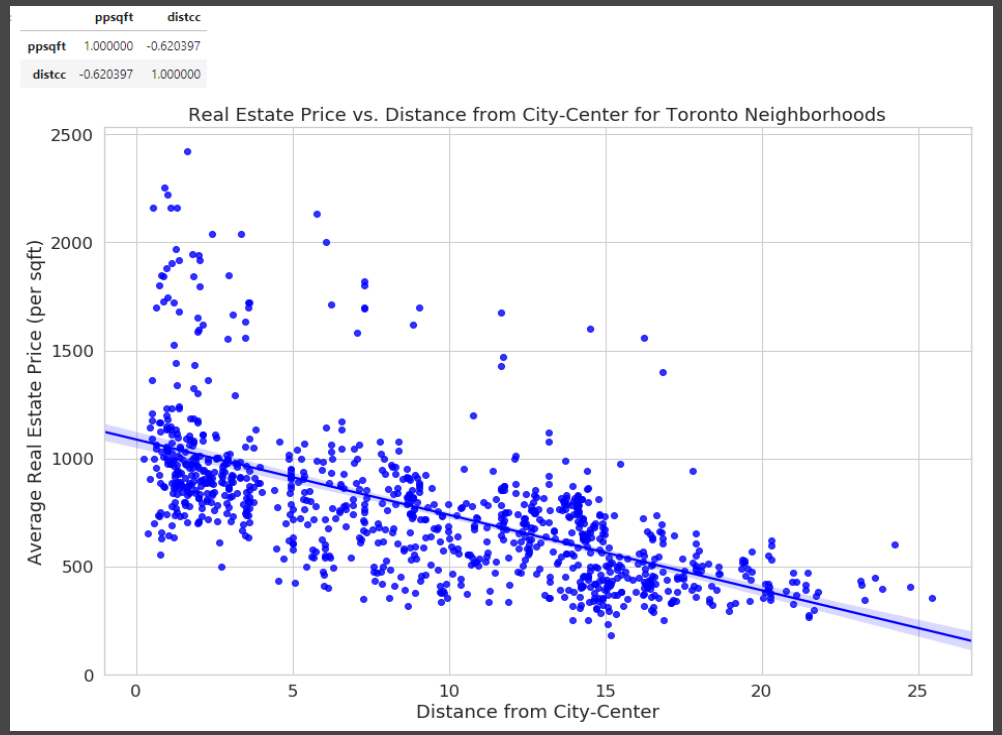
RESULTS & DISCUSSION

Decent correlation between proximity to the city-centre and real-estate prices

- -0.62

Moderate premium for houses near city-centre

	ppsqft	distcc
ppsqft	1.000000	-0.620397
distcc	-0.620397	1.000000



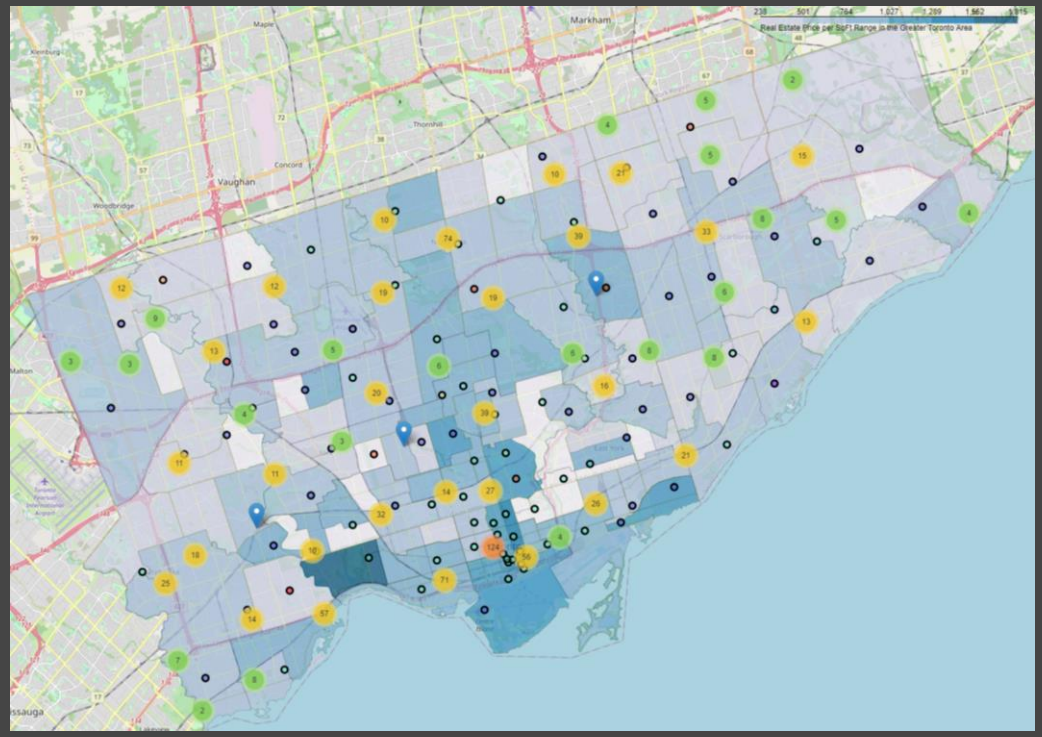
RESULTS & DISCUSSION

Clusters close to the city-centre consist of cafes, coffee shops etc.

- Cater to working crowd/busy singles
- Food readily available

Clusters further out contain more grocery stores, family restaurants and parks

- For families that prefer to cook
- Spend time outdoors

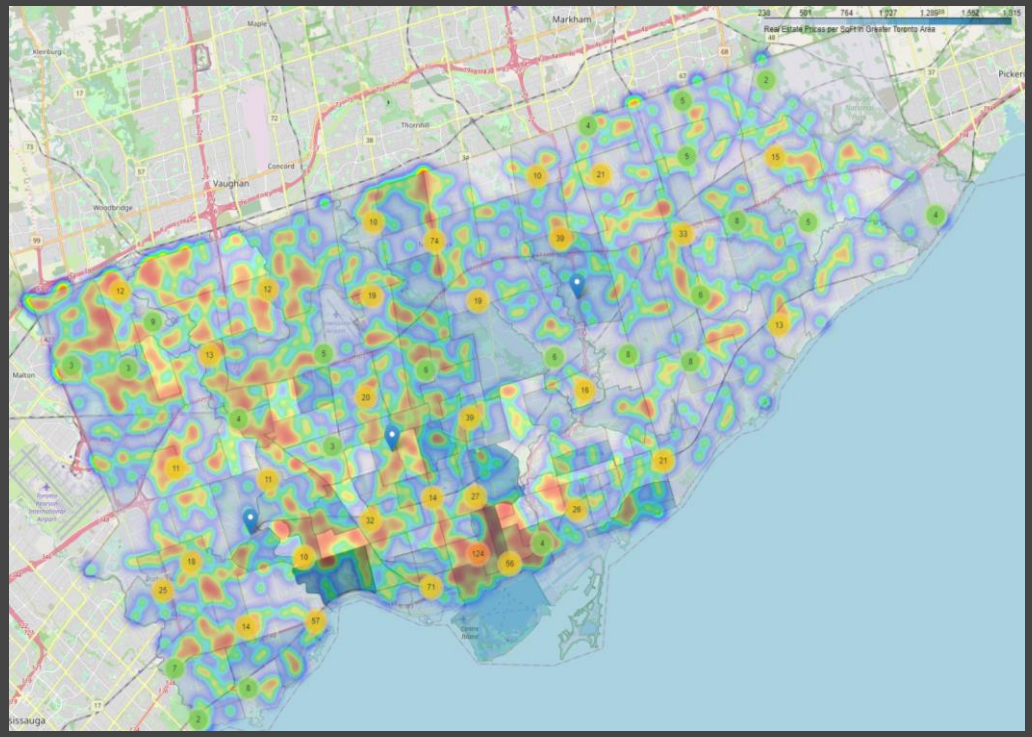


RESULTS & DISCUSSION

Heatmap of crime incidents in 2019

Red Zones = regions of high crime concentration

- Downtown/NW GTA – Higher population density etc.
- Try to avoid listings located in or near red zones

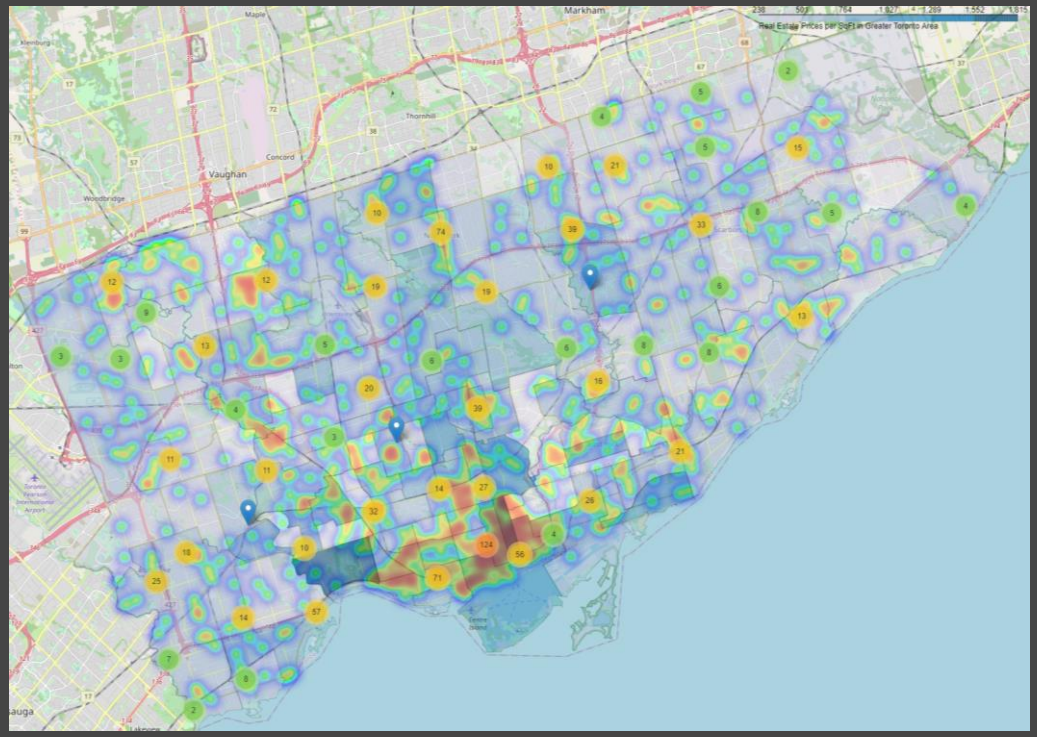


RESULTS & DISCUSSION

Heatmap of fire outbreaks in 2019

Red Zones = regions of high fire incidents

- Downtown – Higher population density etc.
- Try to avoid listings located in or near red zones



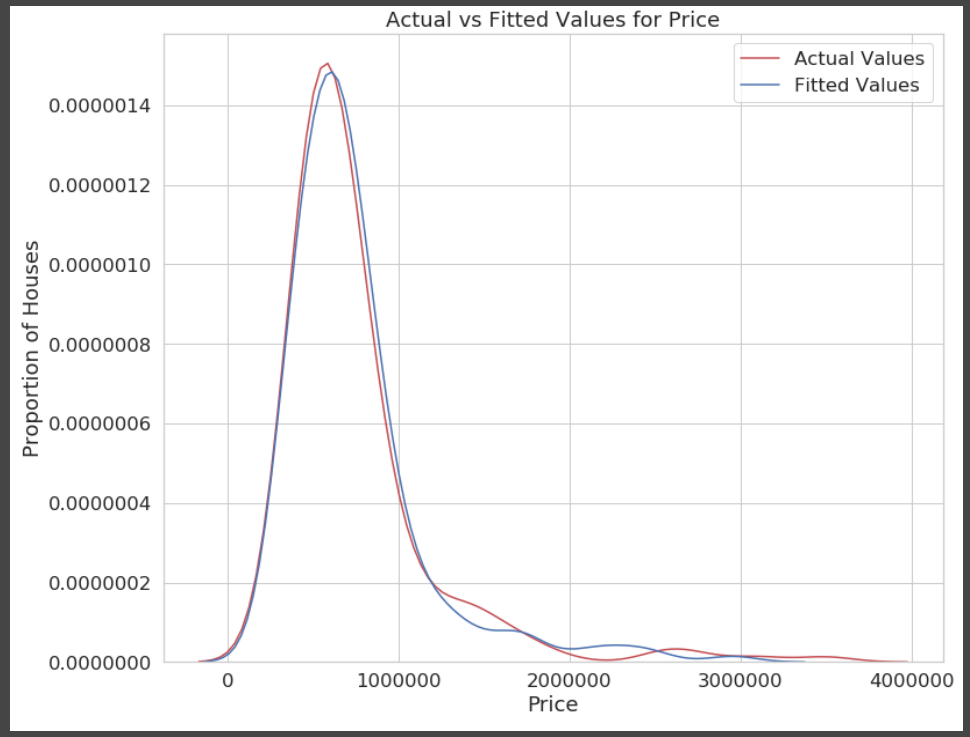
RESULTS & DISCUSSION

- Building and testing of various models

Model	R-Score	MSE (x 10^{10})
Multiple Linear Regression	0.687	6.57
Pipe Method	0.763	4.99
Ridge Regression	0.687	6.57
Polynomial with Ridge Regression	0.763	4.99

- Real-World Data testing of chosen model

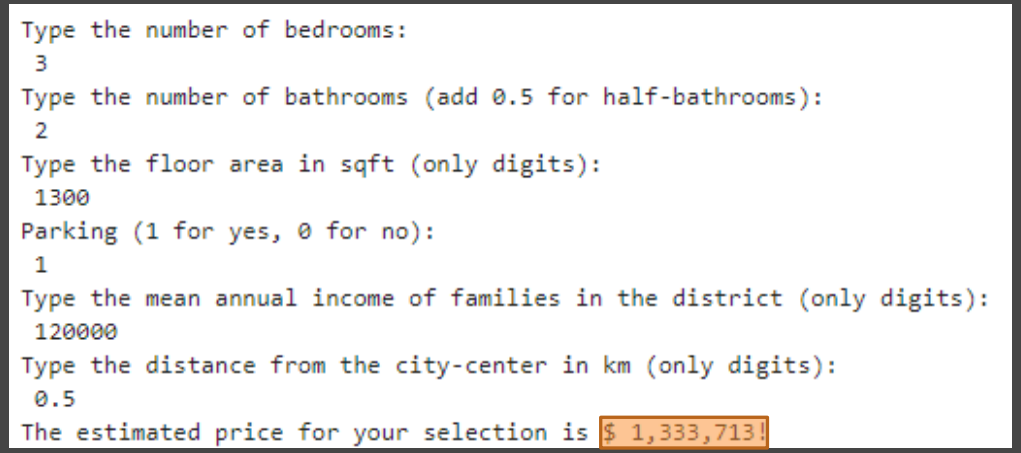
Model	R-Score	MSE (x 10^{10})
Polynomial with Ridge Regression	0.829	4.27



RESULTS & DISCUSSION

Chosen model used to build calculator:

- Translates user input of quantitative variables to price estimate
- Augment with qualitative analysis to tweak estimate as required



CONCLUSION

- Sound methodology (combine both quantitative and qualitative factors)
- Accurate estimation achieved
- Could use more variables so as to capture outliers:
  - Proximity to Train Stations
  - Renovations
  - Views
  - Floor/Level