

Project Proposal – DS8007

Rafik Matta

Problem: Are Toronto Real Estate Prices Justified?

Introduction

Given the current state of the Toronto real estate market, using open economic data provided by the city, government and other open sources I am looking to discover and present an argument about whether or not there is enough economic justification for current residential real estate prices. I have yet to form an opinion one way or the other and will form a story around the data as it presents itself.

Dataset Descriptions

There are several good data open data catalogues with relevant historical and recent economic and price information as well as price index information. I have begun exploring them but will provide the sources here as I have yet to determine exactly which datasets from these sources will be the most appropriate.

- 1) City of Toronto:

<https://www.toronto.ca/city-government/data-research-maps/city-stats-in-detail/>

This is the City of Toronto's economic data that is aggregated through things such as property taxes, land registry, school enrollment and census data.

- 2) StatCan:

<https://www.statcan.gc.ca/dai-quo/economic-indicators-indicateurs-economiques-eng.htm>
<http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/dsbbscan-eng.htm>

This is StatCan's economic indicator data which provides things such as historical GDP growth rates, employment numbers etc.

- 3) Open Data Canada:

<https://open.canada.ca/en>

I'm still exploring the other possibly related datasets available via the Open Data Catalogue provided by the Government of Canada

- 4) World Bank Data:

<https://data.worldbank.org/country/Canada>

The World Bank Economic Data is economic data aggregated as well as collected by the World Bank through several sources and they provide this data to the public. This data is an alternative perspective and it can also allow comparisons between countries which can be very useful for identifying trends and patterns that are highly correlated with known economic phenomena.

Analysis

Fundamentally, I'm bearish on the Toronto real estate market. I've read plenty of analysis to support that position but I'd like to produce my own to see if my intuition is correct. I intend on dissecting the analysis into 4 parts:

- 1) Fundamental Economic Analysis of Toronto (GDP, CPI, employment rates etc.)
- 2) Comparative Economic Analysis with cities similar to Toronto
- 3) Historical Trends that justify certain economic events (e.g. increase in interest rates leading to recession leading to price decreases)
- 4) Geographical Trends (Immigration, land use etc.)

Each one of the parts will have at least one visualization but likely more. Ideally all of the visualizations tie in together well.

Technologies

Although I have a background in JS and HTML, I find the ease of running analysis and producing visualizations with R and Python to be a more efficient course of action for this project. I will stick to using R and Python with a mixture of libraries and packages such as:

- R Standard Library
- ggplot2
- Pandas and Numpy
- Plotly for both R and Python
- Seaborn for Python
- Matplotlib

The reason I'm choosing to use different visualization libraries and both R and Python is because certain visualizations and tasks tend to be easier to produce in different packages /languages.