


 [cboyda](#) / [LighthouseLabs](#) Public[Code](#) [Issues](#) [Pull requests](#) [Actions](#) [Projects](#) [Wiki](#) [Security](#) [Insights](#) [Settings](#) [main](#) [LighthouseLabs / Project-Tableau / README.md](#) Go to file

t

...



cboyda Update README.md

now

...



93 lines (77 loc) · 9 KB

Project Tableau Presentation

Project/Goals

Review provided datasets in Tableau [.twbx](#) or [.twb](#) to demonstrate proficiency and insights from this data.

Process

Key

The key to this dataset was making sure the data was filtered to only include there relevant categories. Datasets used included:

- Weekly earnings from 1.1.2001 to 15.4.2015 (weekly_earnings - CSV)
- Housing constructions from 1955 to 2019 (real_estate_numbers - CSV) with housing starts/completions and geographical regions
- House prices from 1.1.2005 to 1.9.2020 (real_estate_prices - EXCEL)
- Housing_price_index from November 1979 to September 2020
- Office_realestate_index from November 1979 to September 2020
- Consumer index from November 1979 to September 2020

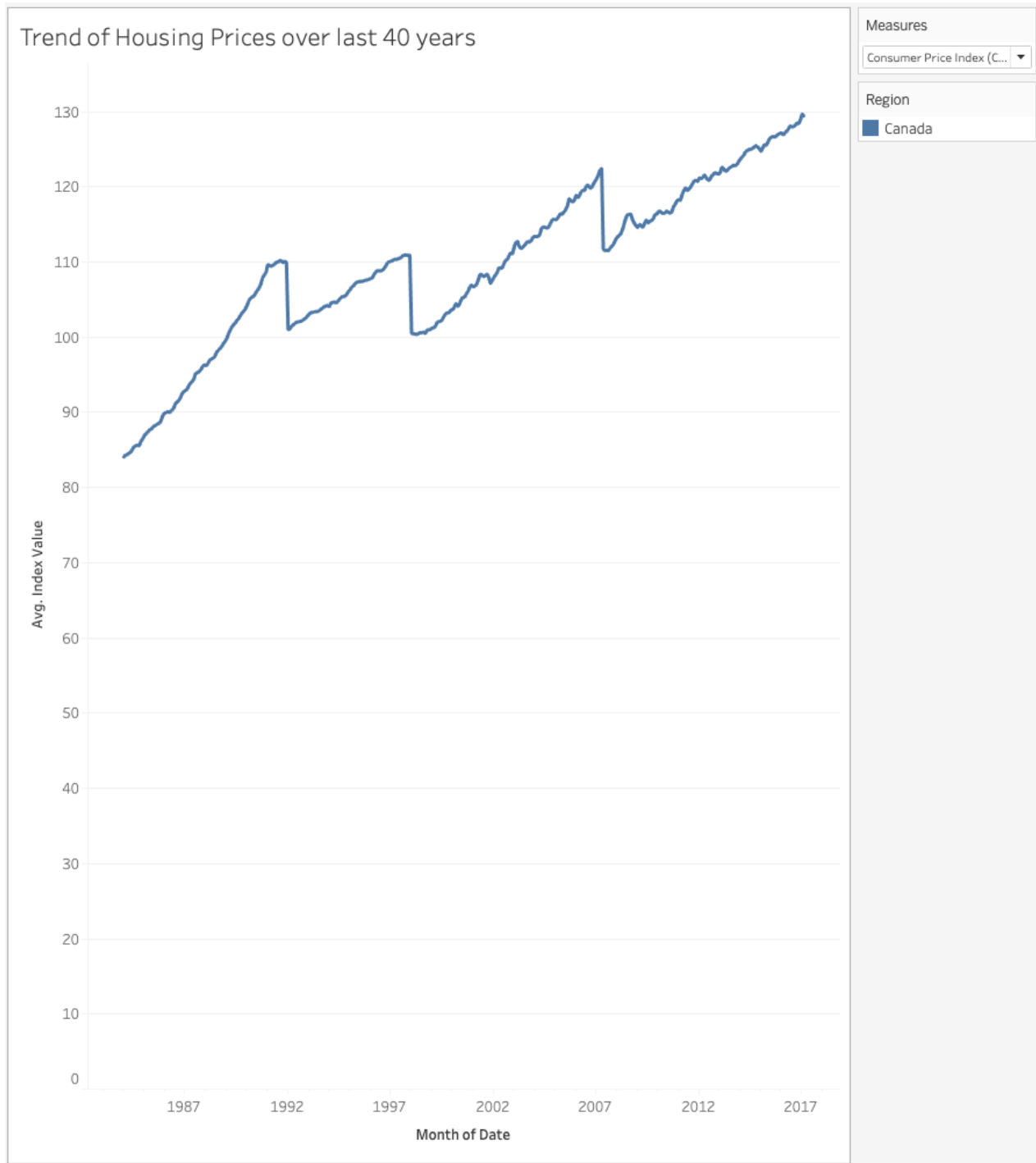
Utilized [Python Google Colab](#) to parse the JSON file and create a cleaner csv for Tableau to visualize.

Results

Option 1 Selected for Standard Final Project

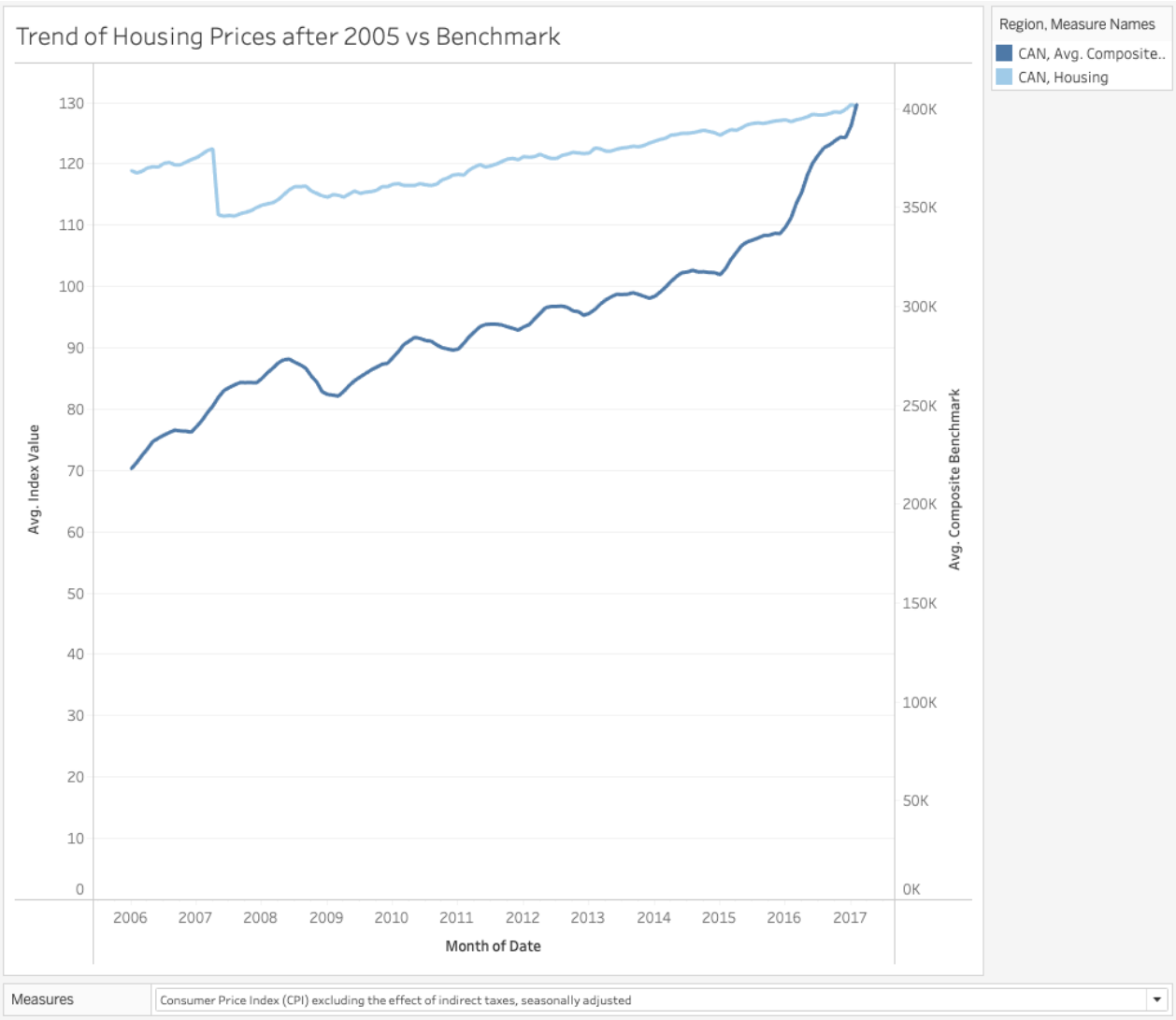
Entire results are available in a [PDF](#), or as a [Power Point Presentation](#), or [Presentation Summary in PDF format](#).

1. Show the trend of house prices across Canada in the last 40 years (table housing_price_index).

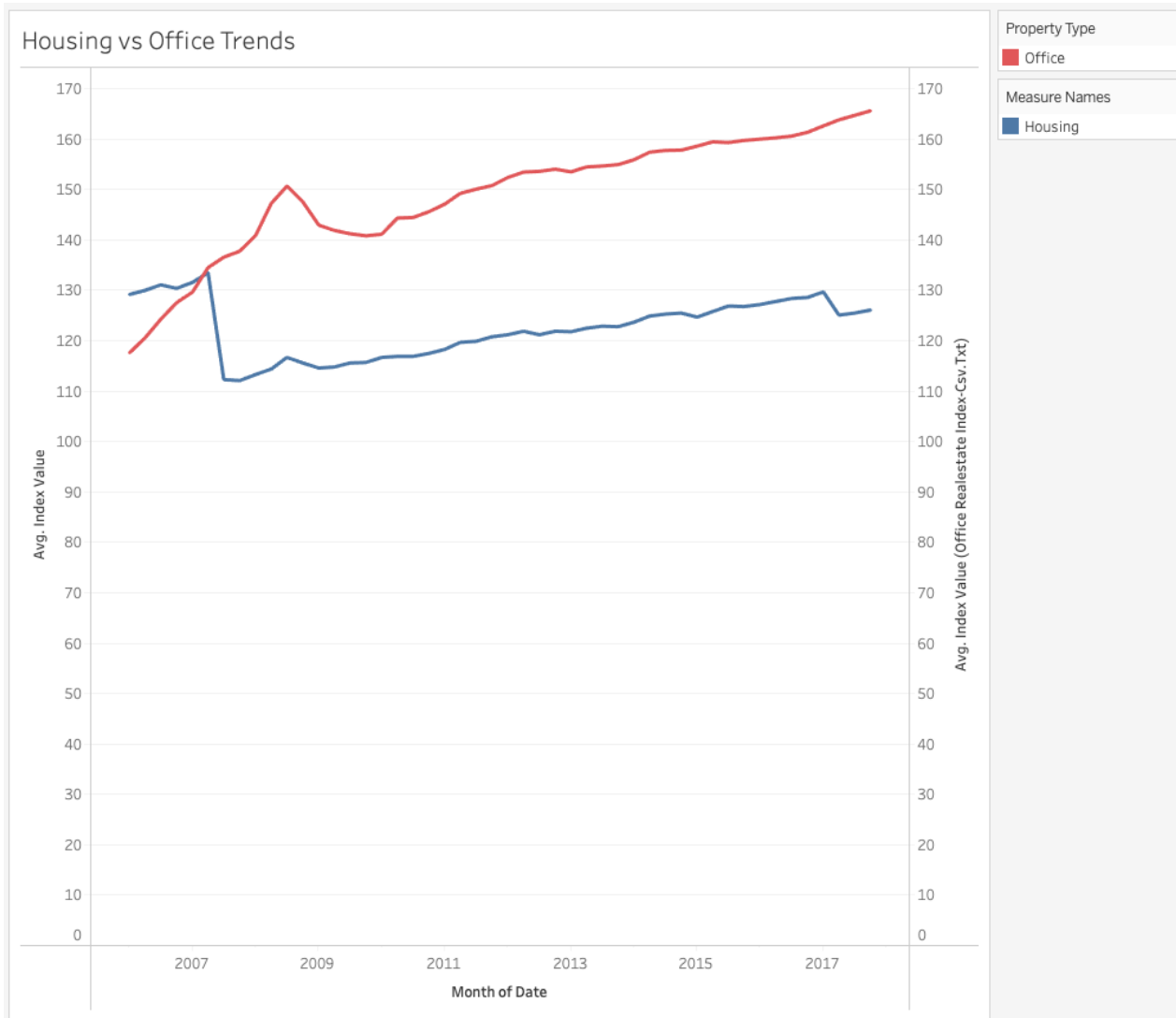


Note: This is using the measure: Consumer Price Index (CPI) which excludes indirect taxes, seasonally adjusted.

2. Compare the trend after 2005 with actual benchmark prices in table real_estate_prices to see if there are any differences.



3. Compare this trend with the trend of office prices.



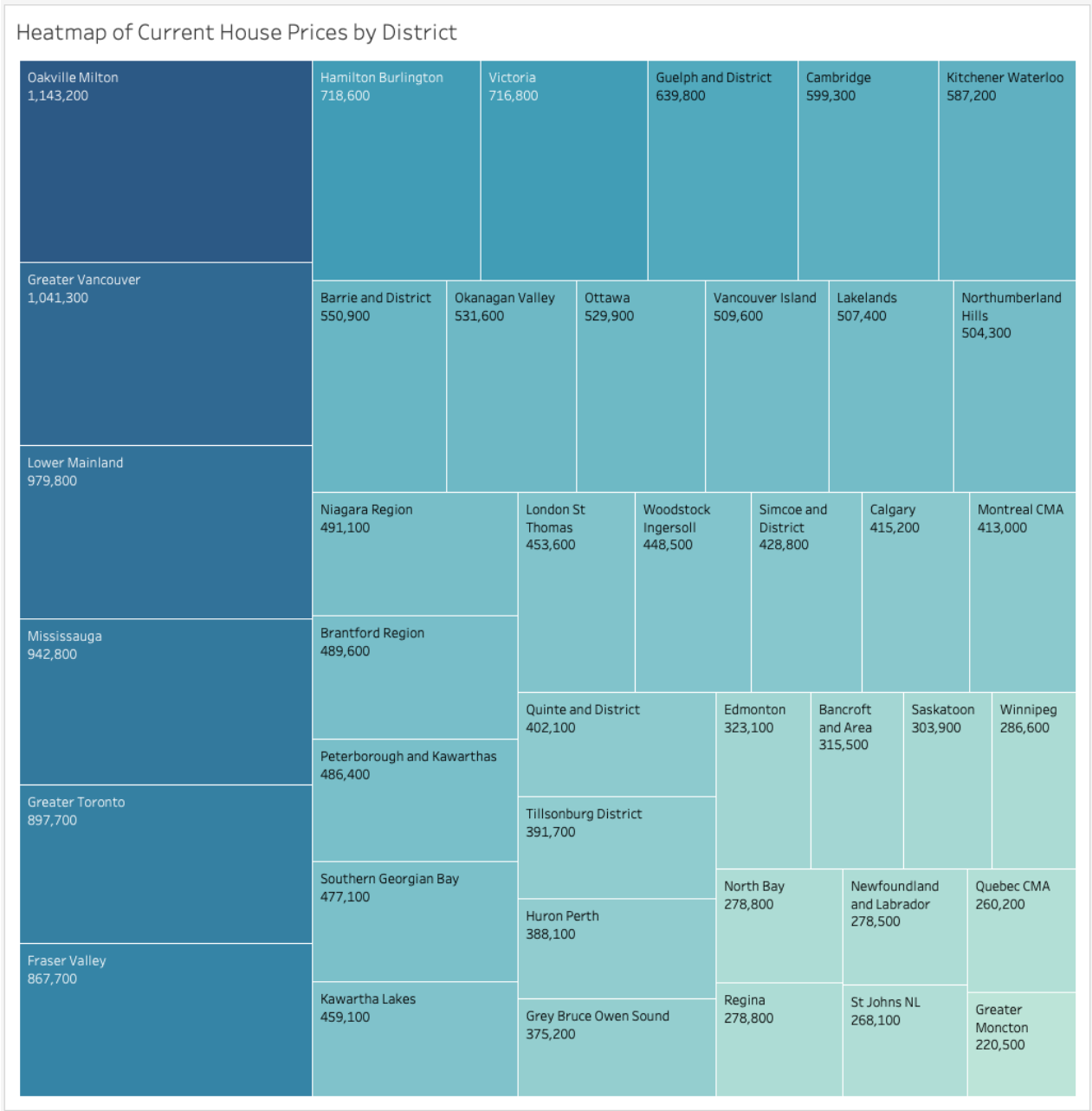
Notice that in the first chart our housing prices go to 2021 but once we compare with office prices their data only goes to 2017 so much of the trend is cropped off.

Which one is getting more expensive, faster? OFFICE SPACE is observed with a larger slope/increase vs housing prices.

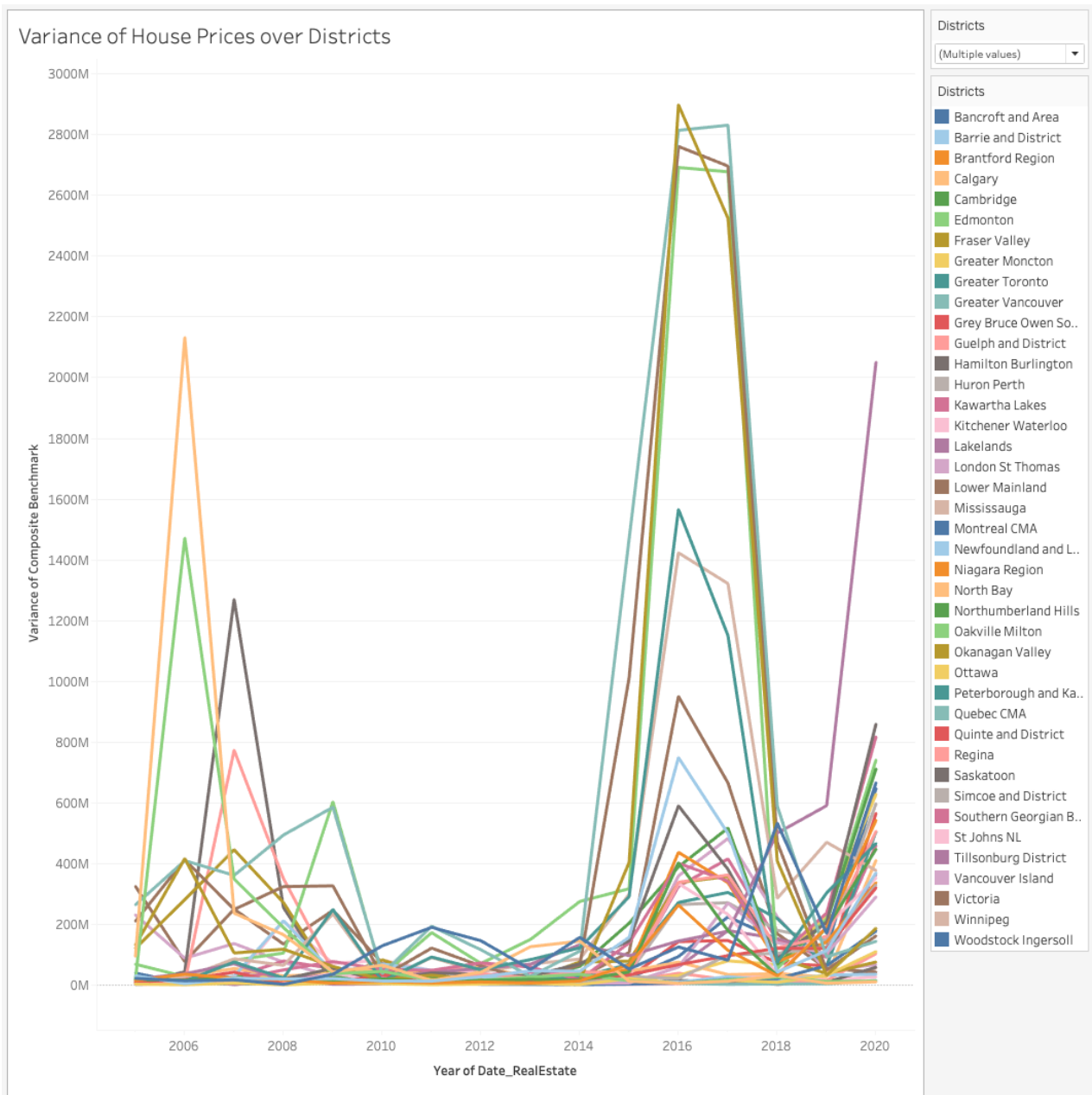
4. Create a heatmap of Canada with current house prices for each available district.

Latest time available was Sept 2020, with the size of the composite benchmark relating the size of

each district.

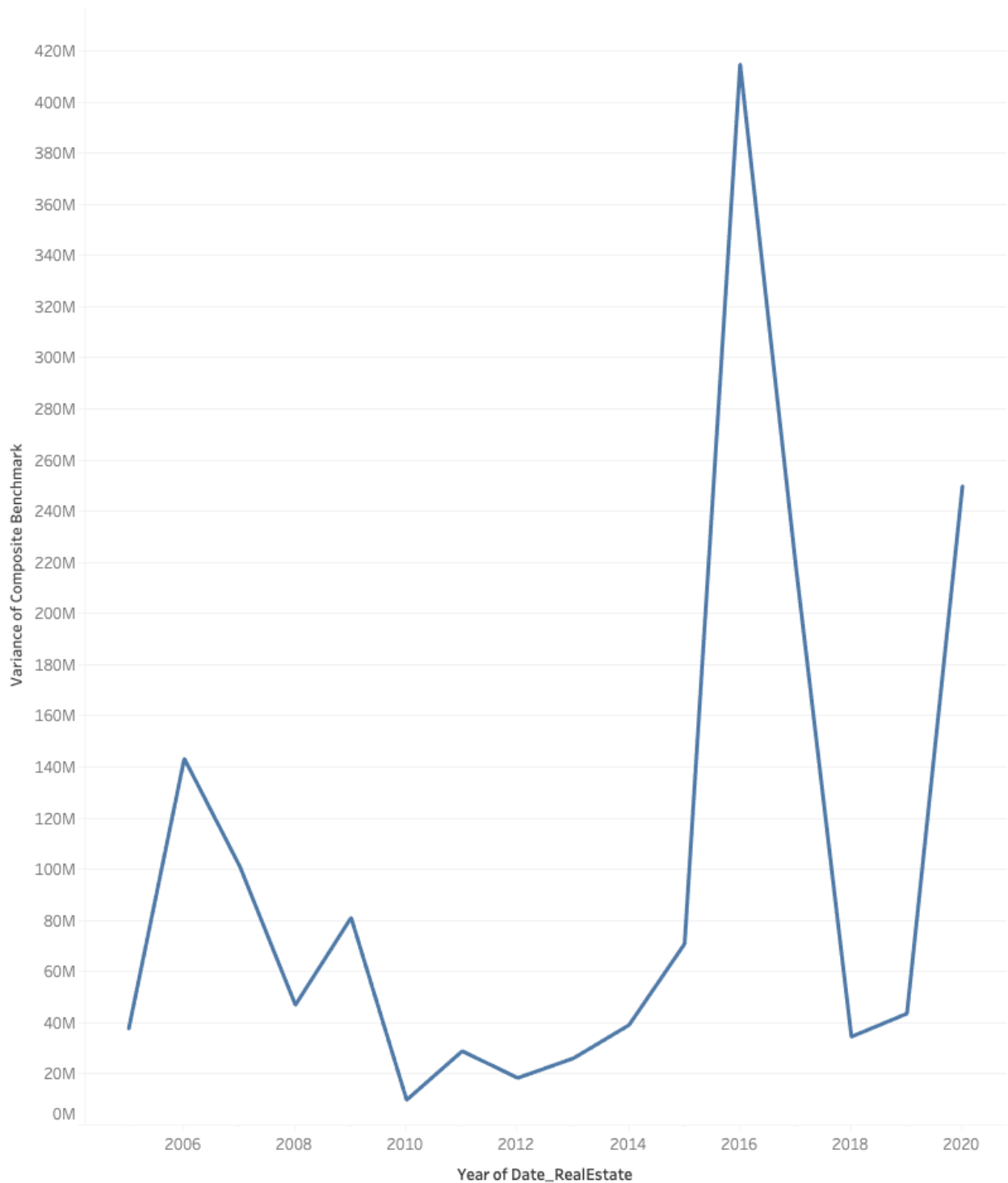


5. Are the price differences between different districts increasing? Variance: is another measure of the dispersion of prices over the districts. It is the squared value of the standard deviation. Variance provides an understanding of the overall variability of prices and can be useful for comparing the spread between districts.



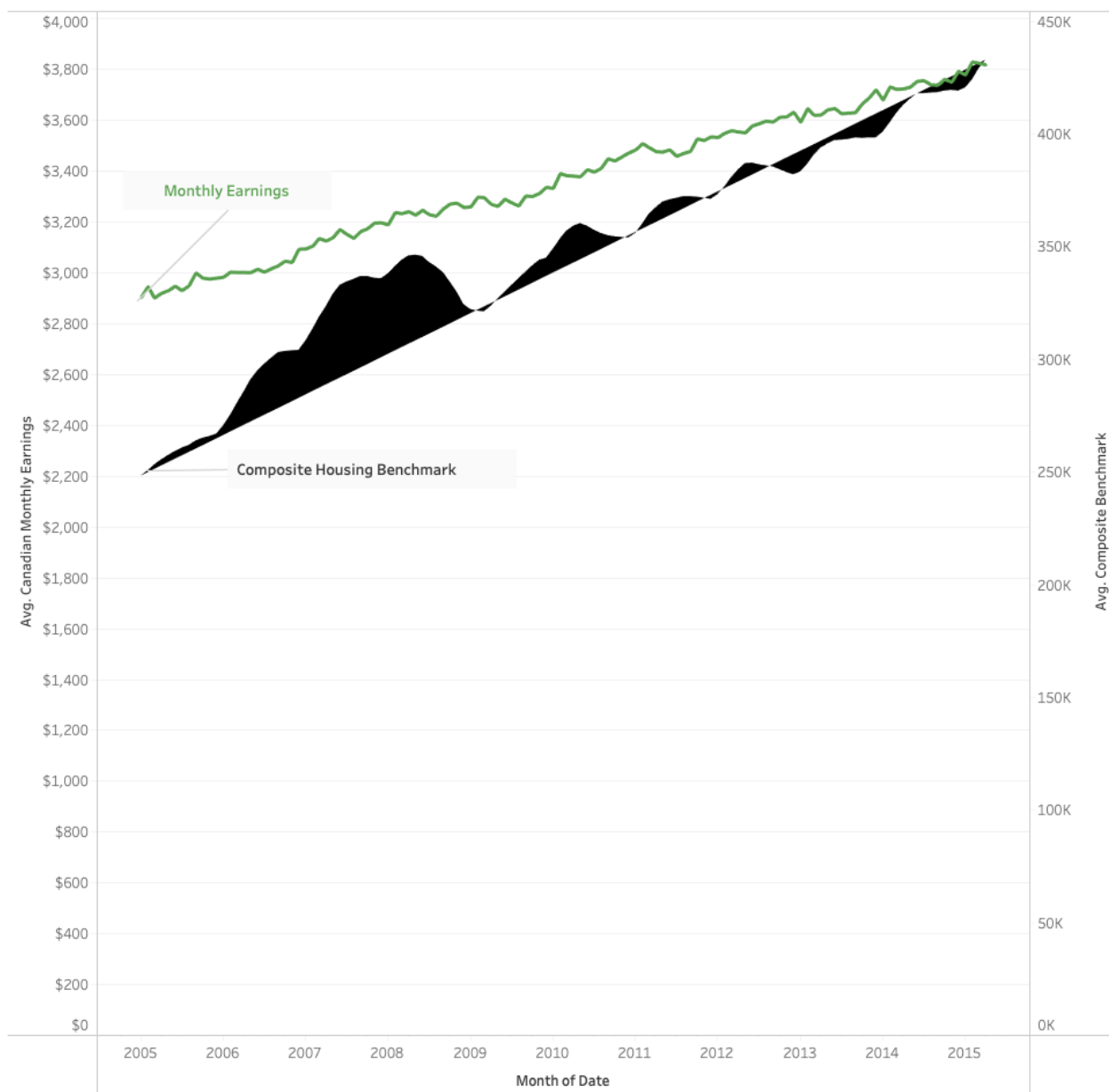
In this chart we aggregate results to reduce noise.

Variance of House Prices Aggregated Districts



6. Compare the trend of house prices with earnings. *In case you want to plot monthly salary, be aware that the earnings value is per week.

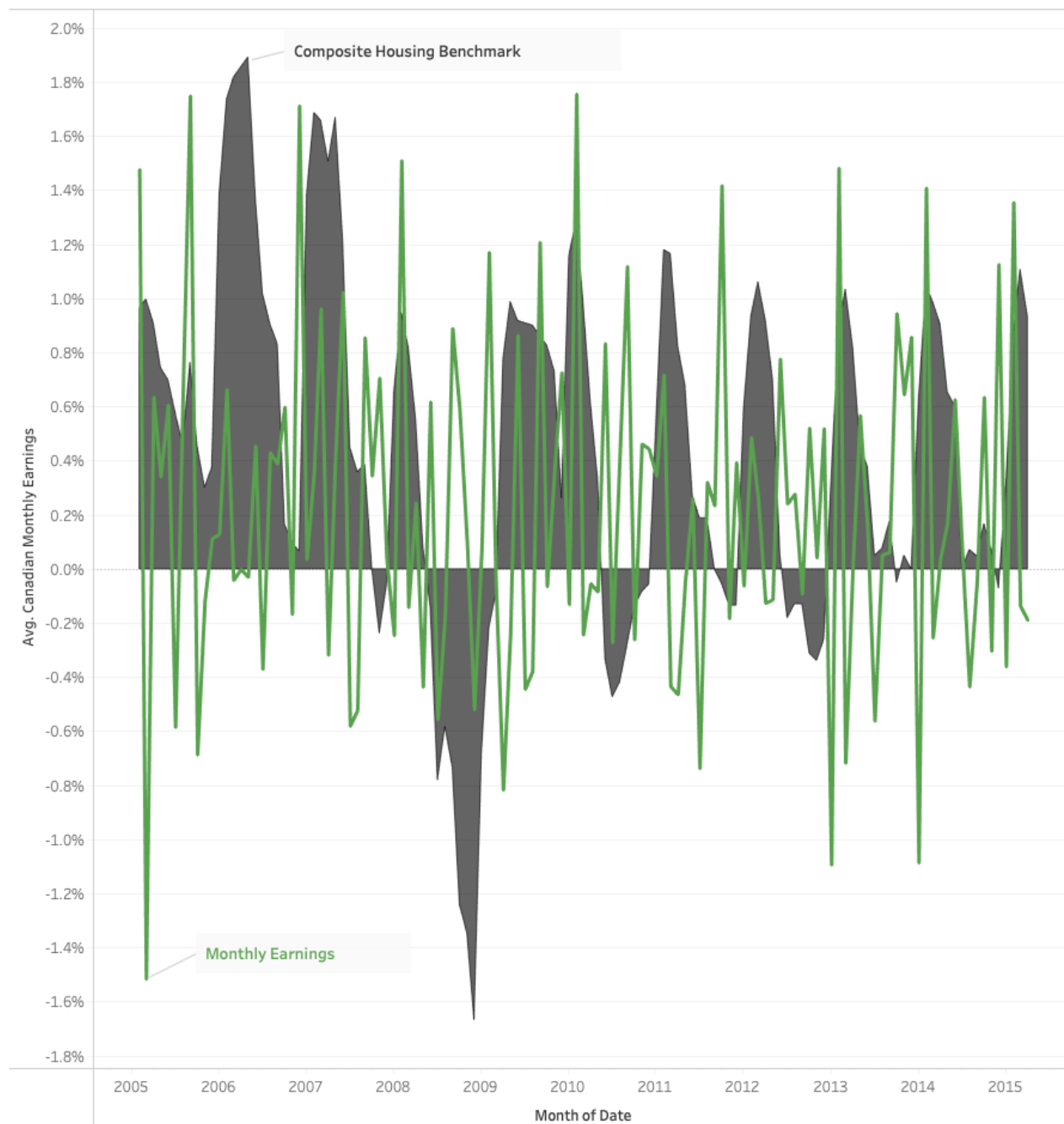
Canadian Trend House Prices vs Earnings



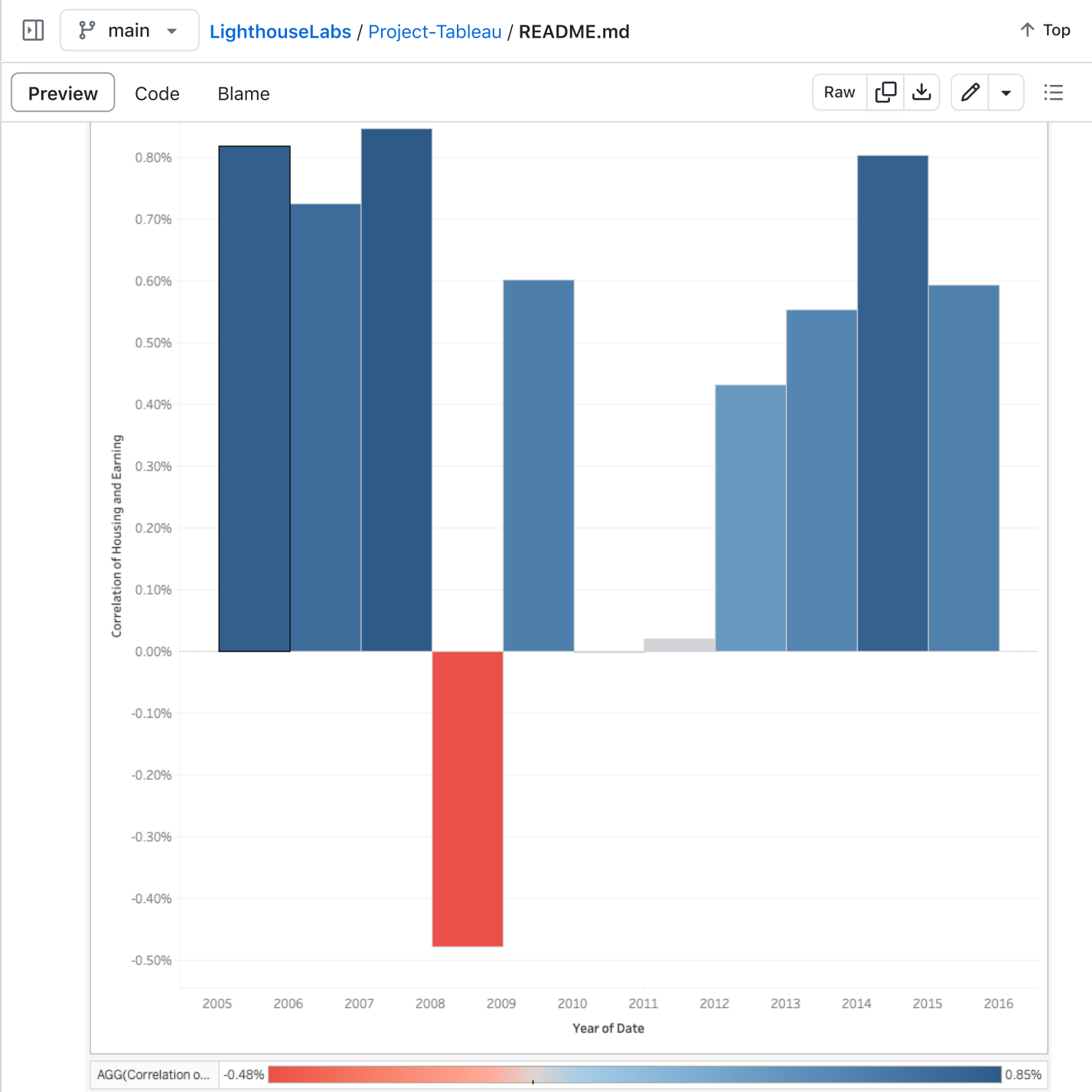
While this visualization shows the general slope over time there is a mismatch in scales that causes tremendous white space below. Given more time we could work to standardize these values to better "fill" our chart space.

An interesting thing to note is that the housing price has been filled above/below the trend line which really mimics the increases/decreases of the wages over time.

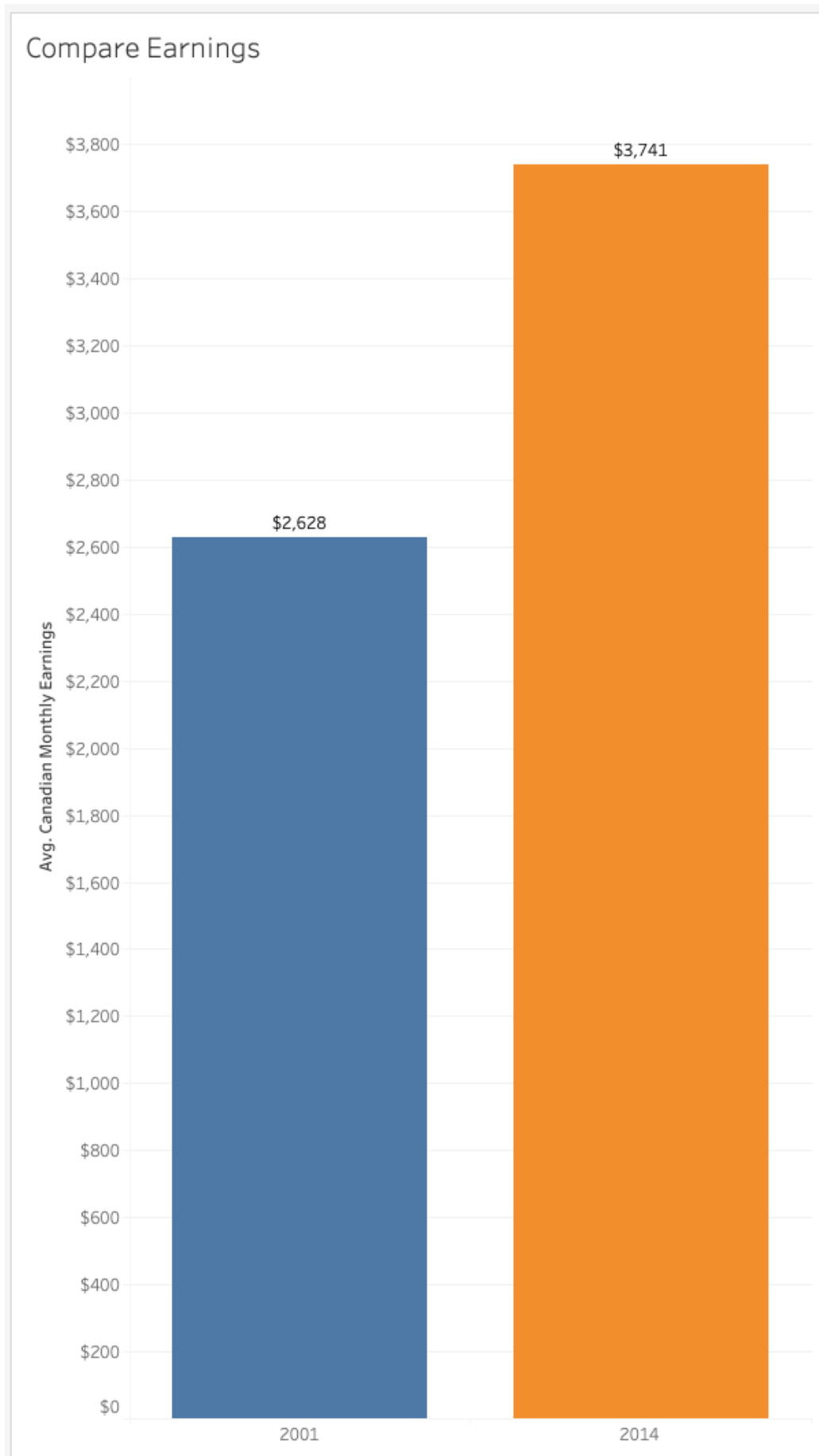
% Difference Canadian Housing Prices vs Earnings



By calculating the percentage difference we are able to place these trends over top of each other (on the same axis) and see there is a often a correlation between these values.



7. Did people spend more of their earnings in 2014 than they did in 2001? No expenses were provided, but we can compare earnings for 2014 vs 2001.

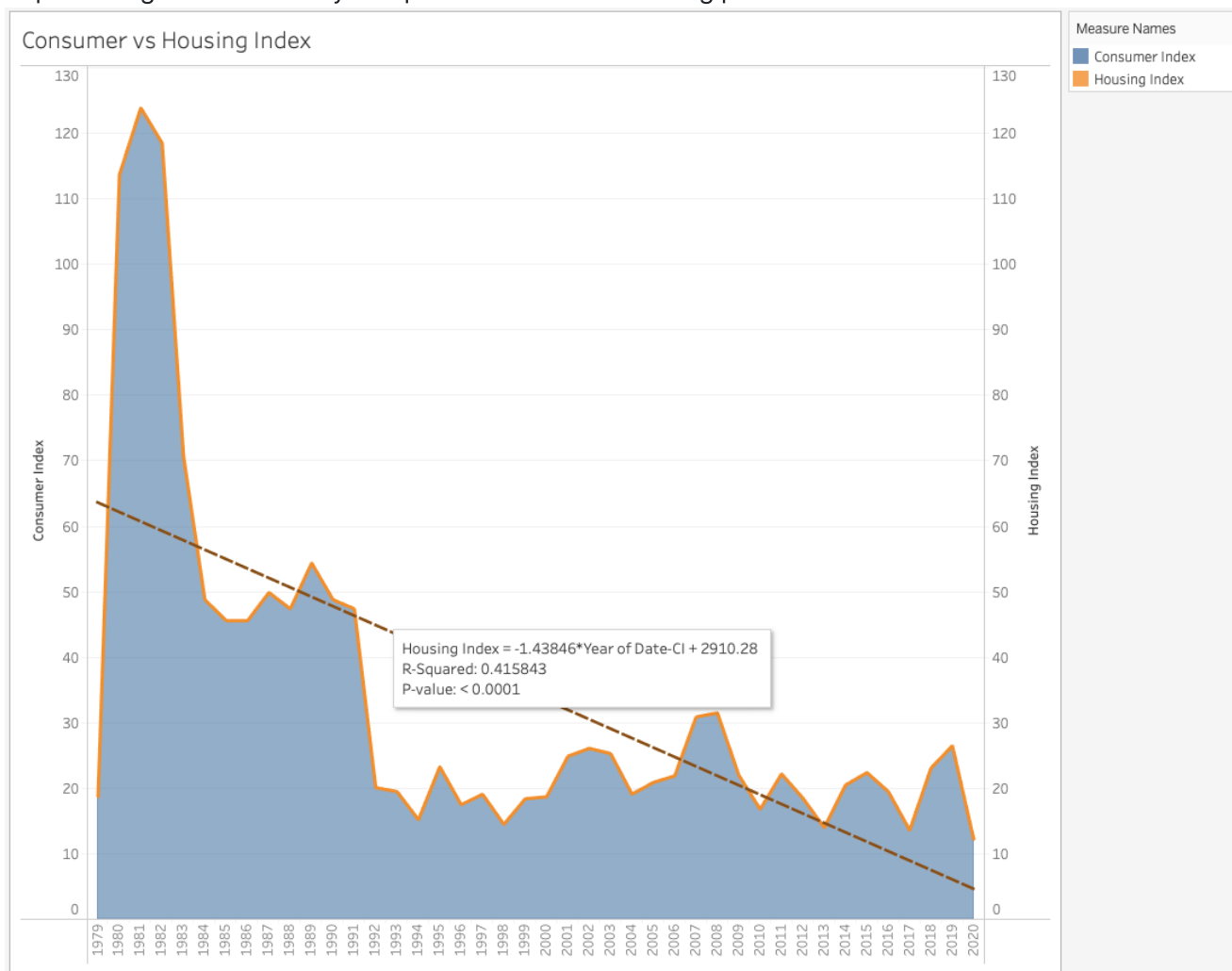


8. There were several economic crises in the world in the last 40 years, including these four: Black Monday (1987), Recession (early 1990s), dot com bubble (2000 - 2002), Financial crisis (2007 -

2009).

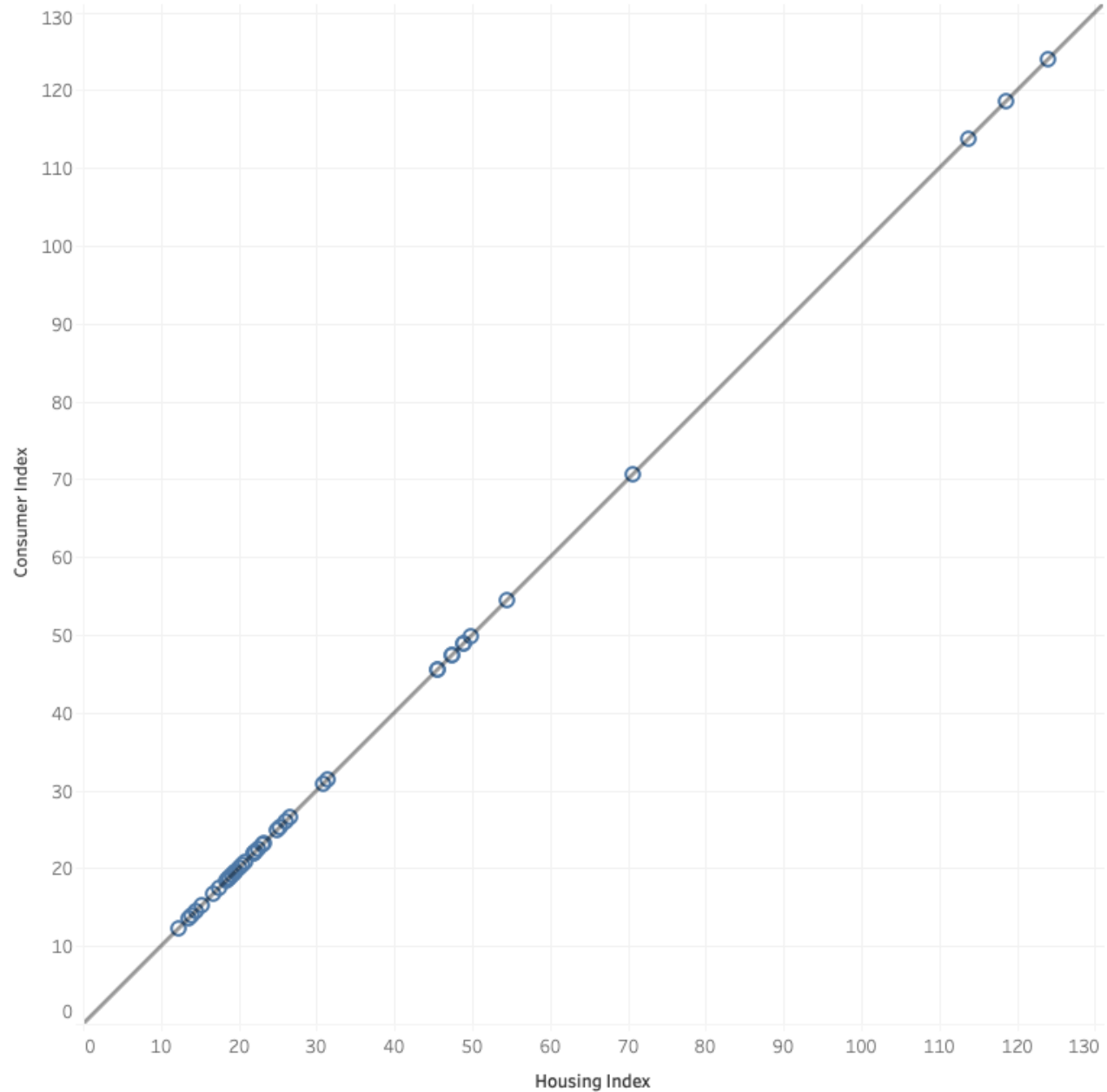
► Show the effect of these crises on:

9. Plot consumer_index together with housing_price_index and fit the regression line between them. Can we predict consumer_index from the housing_price_index? Using the CPIW index value, and limiting only to percentages we can easily compare consumer vs housing price indexes.

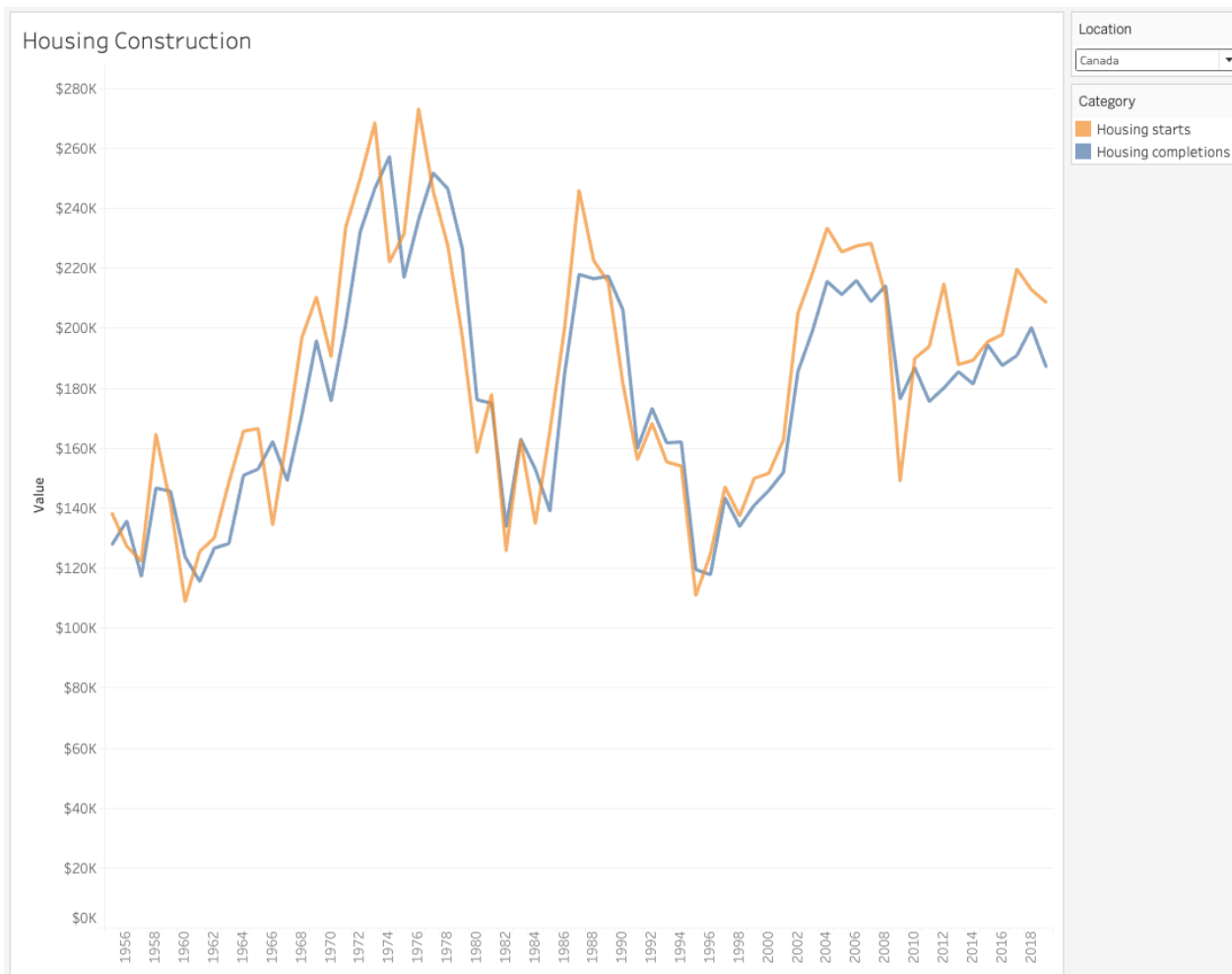


These follow each other so closely, even the linear regression line can be seen to accurately predict housing price from the consumer index.

Regression Line Consumer vs Housing Index



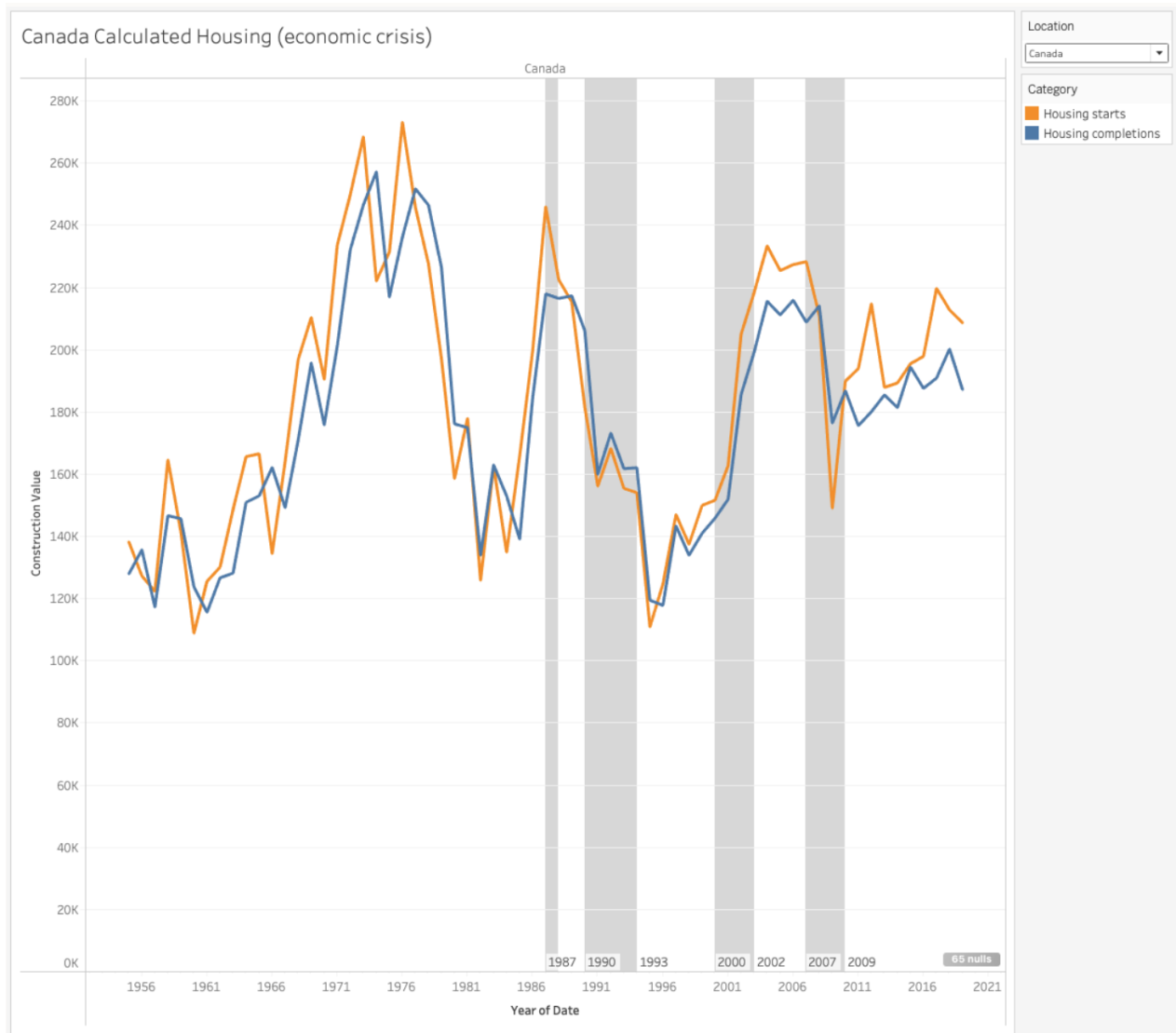
10. Try to find an interesting pattern, trend, outlier, etc. from the data used in the above questions. HINT : Double check all units in the table before any comparison.



Found it interesting that starts and completion differed inconsistently over time.

And certainly affected by recessions (1990) and mortgage collapses in our economy (2009). Notice

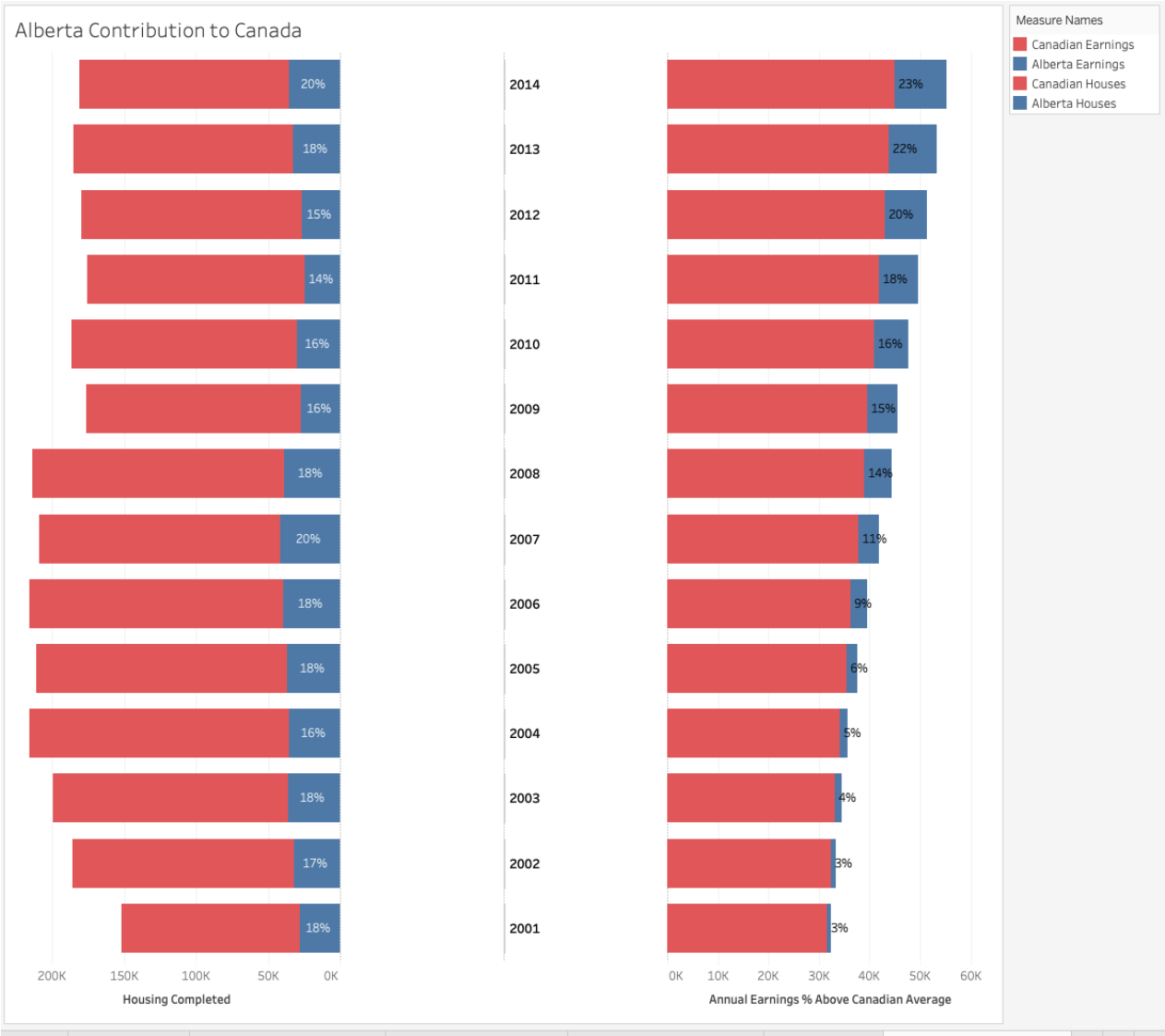
how the construction completion lags thereafter.



- recreated earnings data to compare annually to housing, require [Python modifications to original data](#)

Utilizing that reformed data we can now look at how much Alberta contributes to Canada in both

construction completed and earnings.



Challenges

Discuss challenges you faced in the project.

X greatest challenge was trying to understand the overall goal of the question, without clarifying questions or a sample output this is just a guess. The role of the subject matter expert is key to understanding the data and matching their knowledge with the clients chart visual goals.

X Many of the datasources tried to merge data. For example there are multiple CPI indexes, then the units are percentage AND index values. Without filtering these to match exactly the graphs/charts quickly can become erroneous.

Future Goals

(what would you do if you had more time?)

- create a much more interactive dashboards/animation/timelines, but ensure collected data shares similar columns to leverage shared filters

- with the data having such limited overlap it made shared universal dashboard controls very unrealistic (loses the perspective of change over time), would be important to choose better data in the future what shared elements/scales