# Project Tableau Presentation

## Project/Goals

Review provided datasets in Tableau 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:

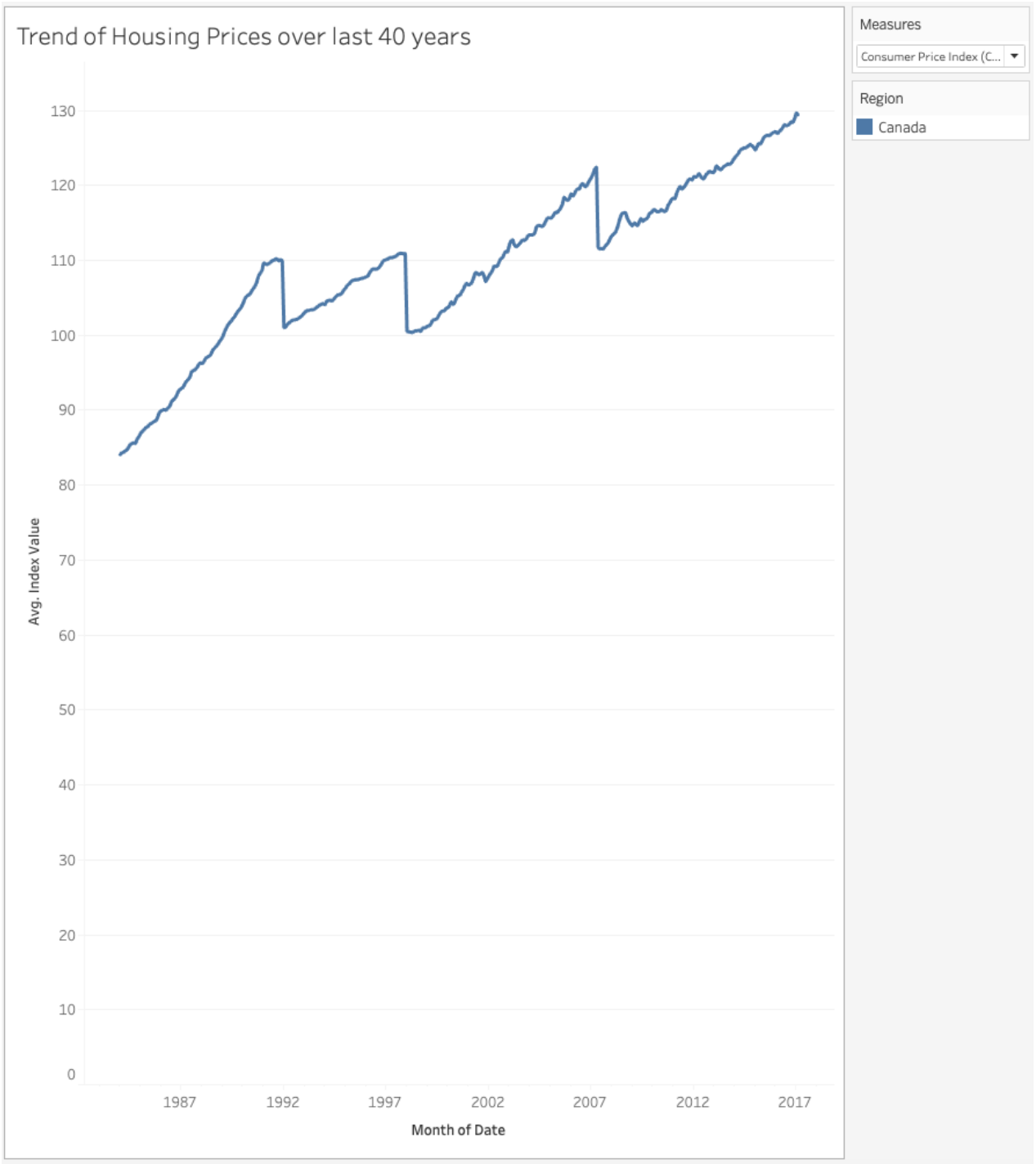
- Weekly earnings from 1.1.2001 to 15.4.2015 (weekly_earnings - CSV)
- Housing constructions from 1955 to 2019 (real_estate_numbers - CSV) with housing starts/completions and geographical regions
- House prices from 1.1.2005 to 1.9.2020 (real_estate_prices - EXCEL)
- Housing_price_index from November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

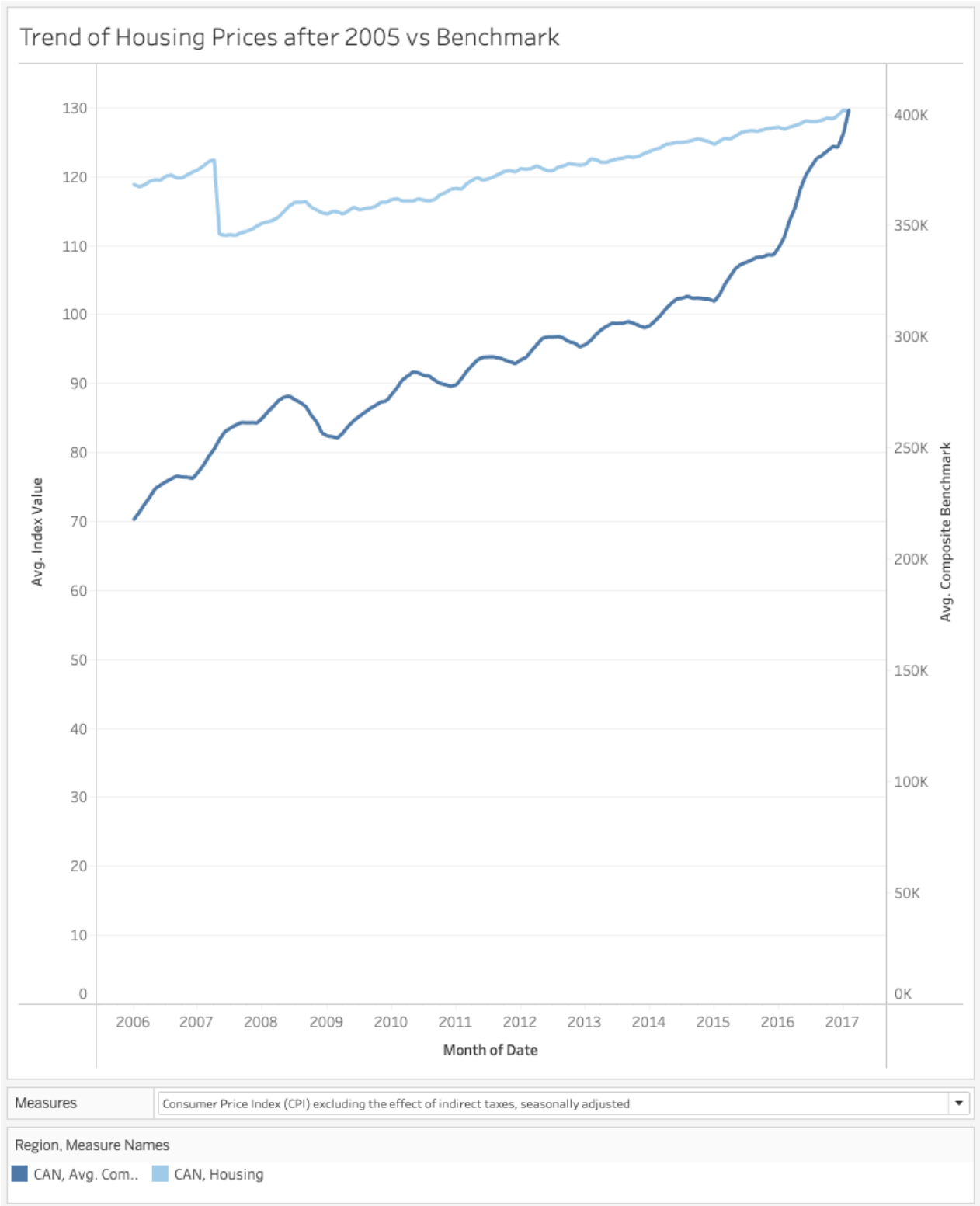
> Option 1 Selected for Standard Final Project

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

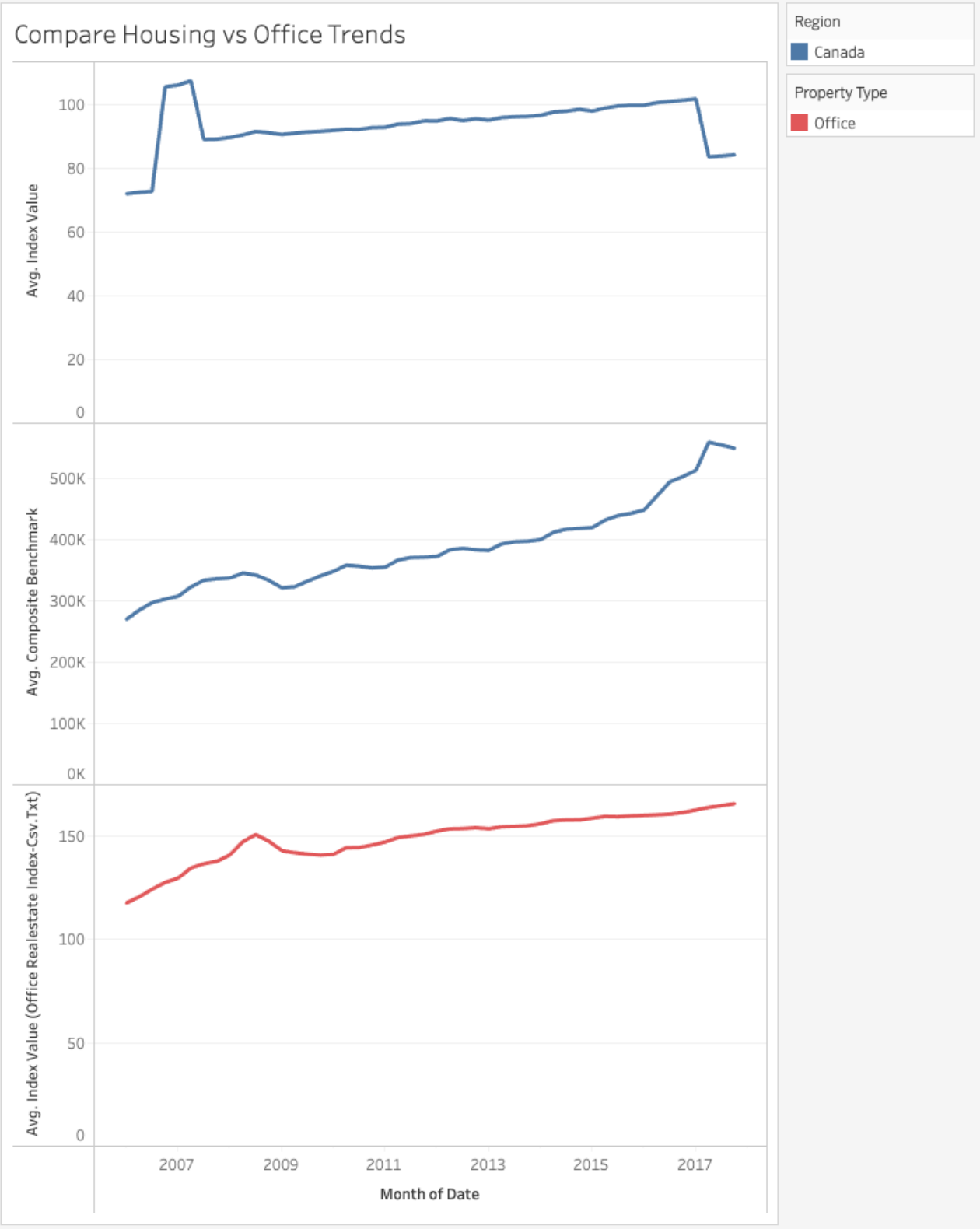



Note: This is using the measure: Consumer Price Index (CPI) which excludes indirect taxes, seasonally adjusted.

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

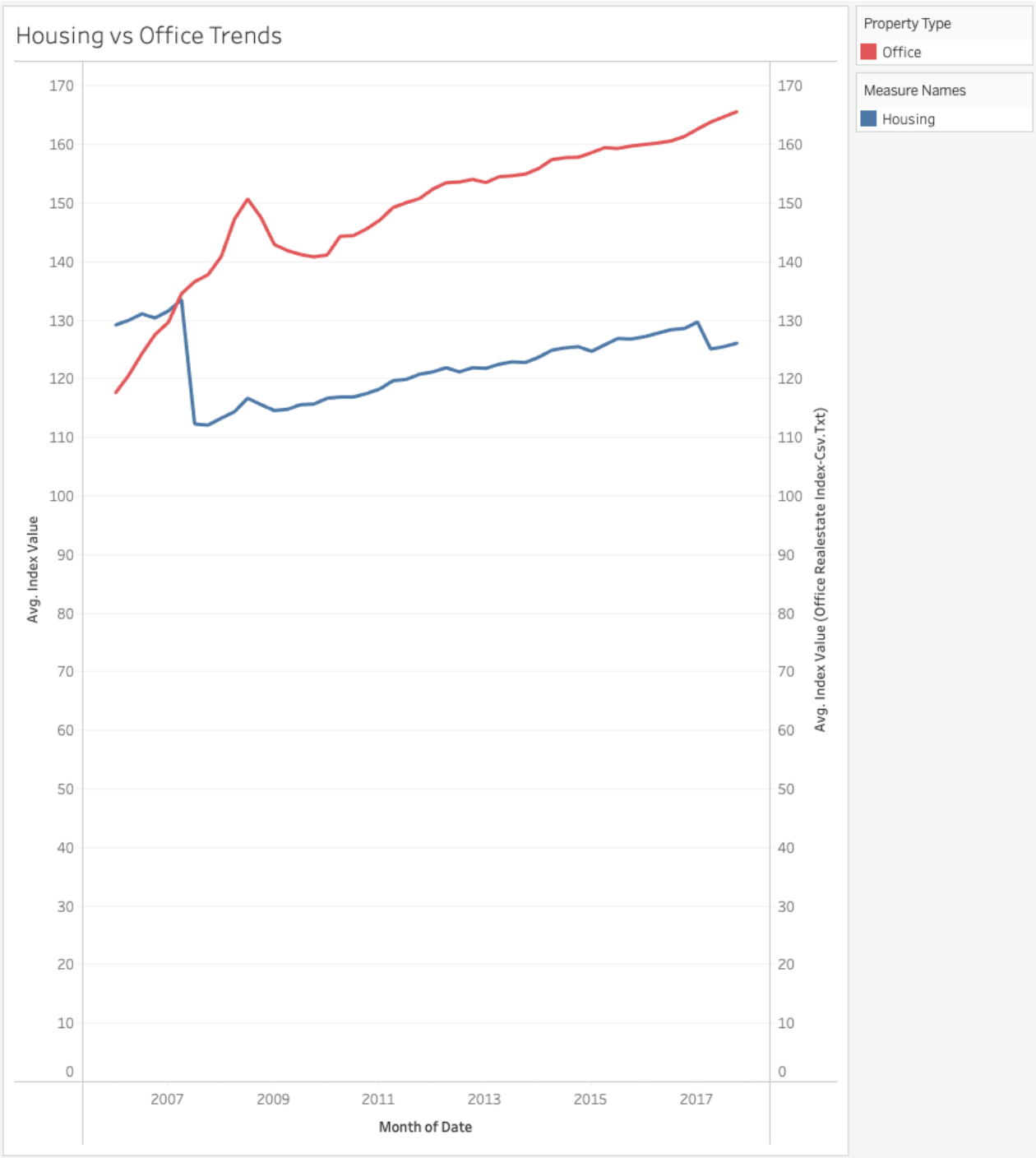

Trend of Housing Prices after 2005 vs Benchmark
Avg. Index Value
Avg. Composite Benchmark
Month of Date
130
120
110
100
90
80
70
60
50
40
30
20
10
0
400K
350K
300K
250K
200K
150K
100K
50K
0K
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
Measures
Consumer Price Index (CPI) excluding the effect of indirect taxes, seasonally adjusted
Region, Measure Names
CAN, Avg. Com..
CAN, Housing

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




This is too complicated to compare easily.

We have 3 graphs with different scales, so by using dual axies and synchronizing the scales we get:

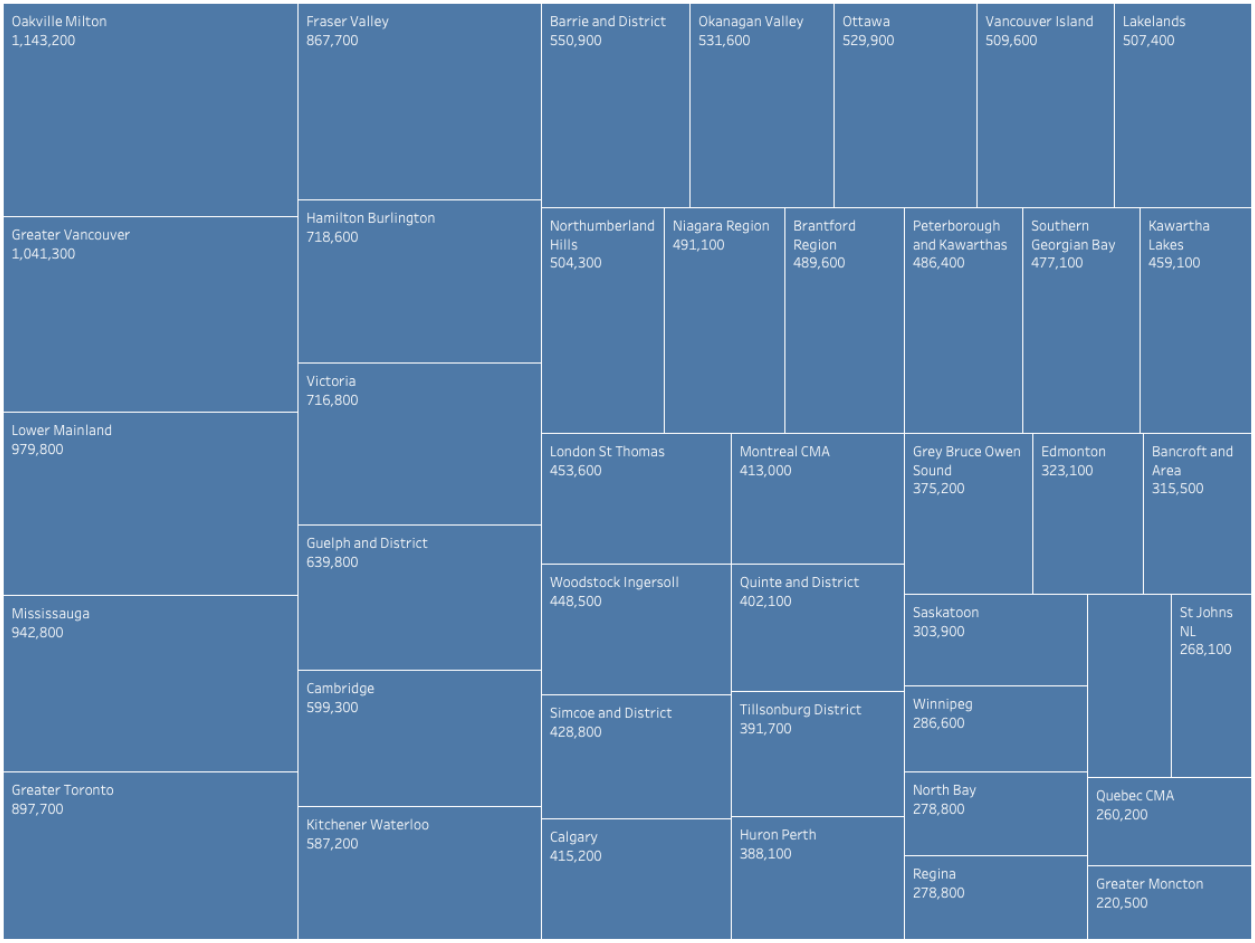



Notice that in the first chart our housing prices go to 2021 but once we compare with office prices their data only goes to 2017 so much of the trend is cropped off. Which one is getting more expensive, faster? OFFICE SPACE is observed with a larger slope/increase vs housing prices.

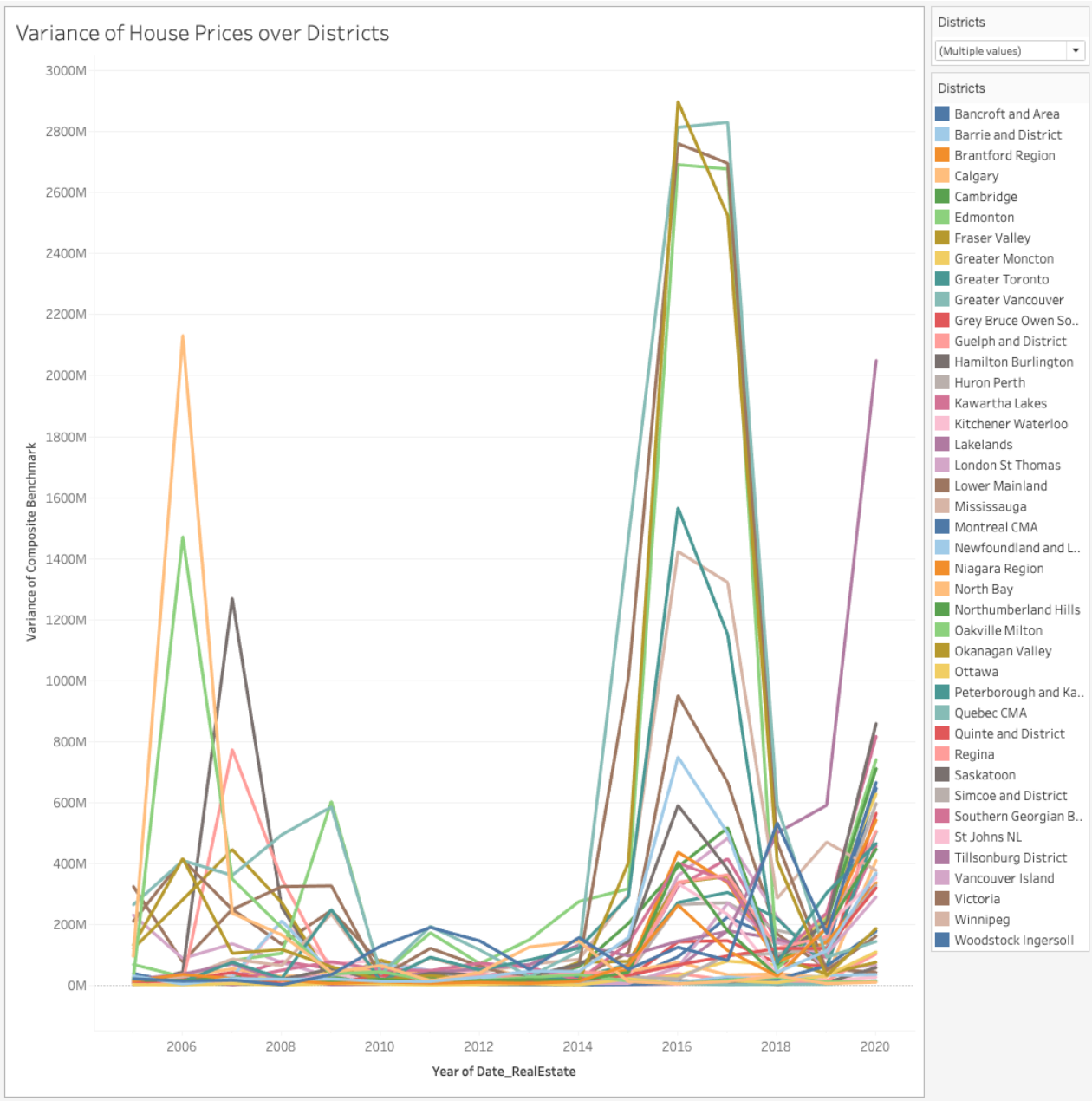
4. Create a heatmap of Canada with current house prices for each available district.
   Latest time available was Sept 2020, with the size of the composite benchmark relating the size of

each district.

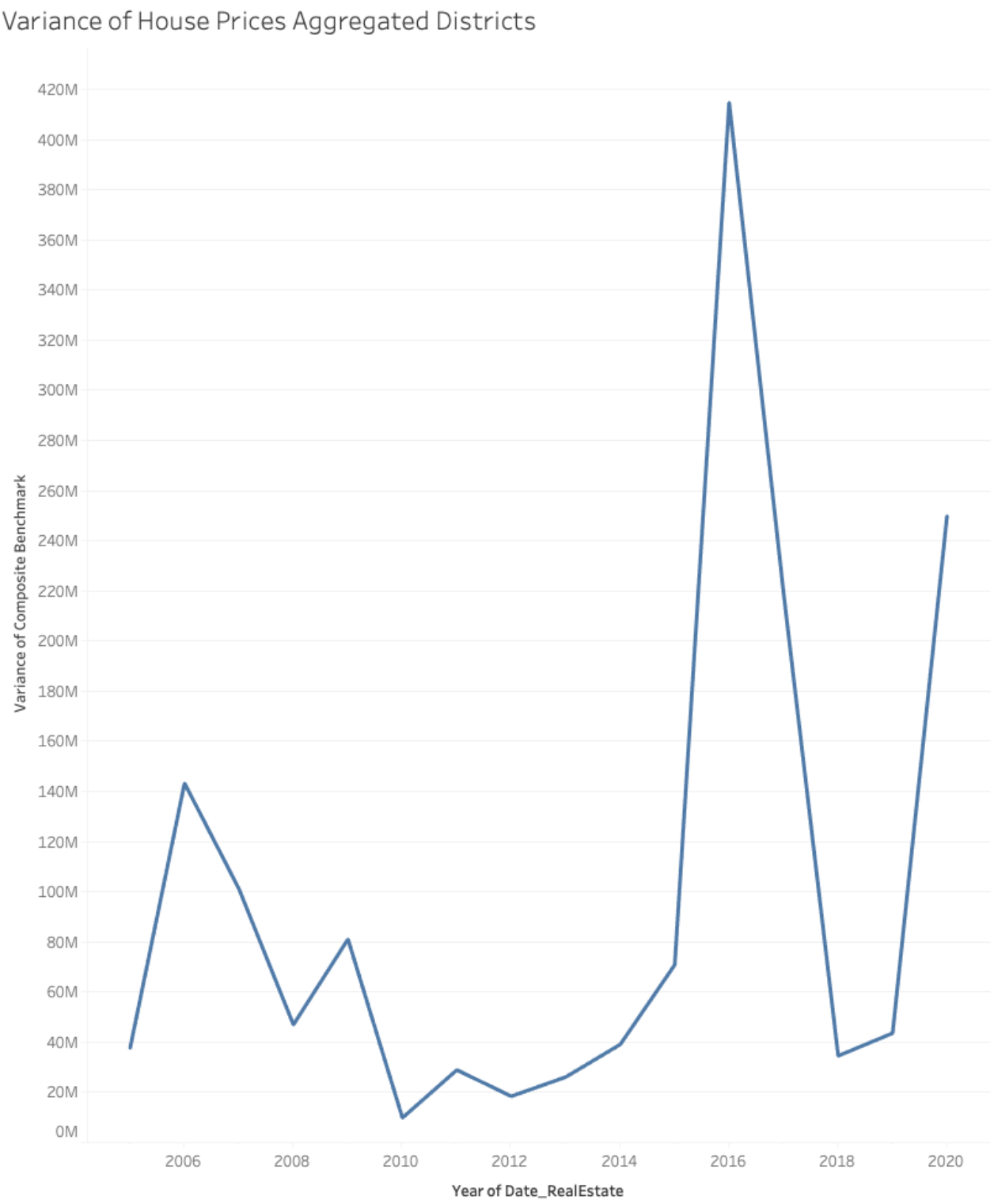
Heatmap of Current House Prices by District



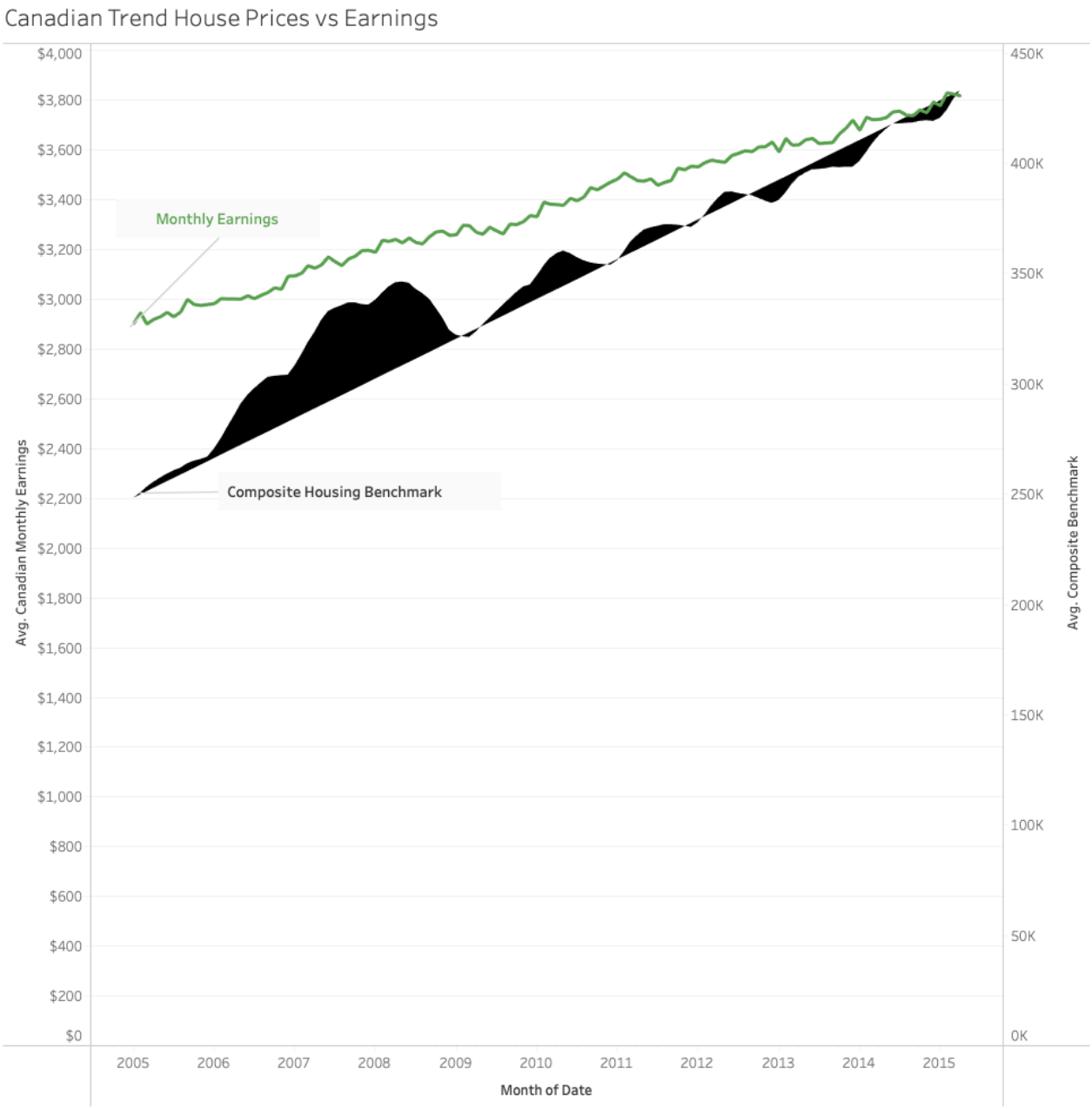

5. Are the price differences between different districts increasing? Variance: is another measure of the dispersion of prices over the districts. It is the squared value of the standard deviation. Variance provides an understanding of the overall variability of prices and can be useful for comparing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




In this chart we aggregate results to reduce noise.

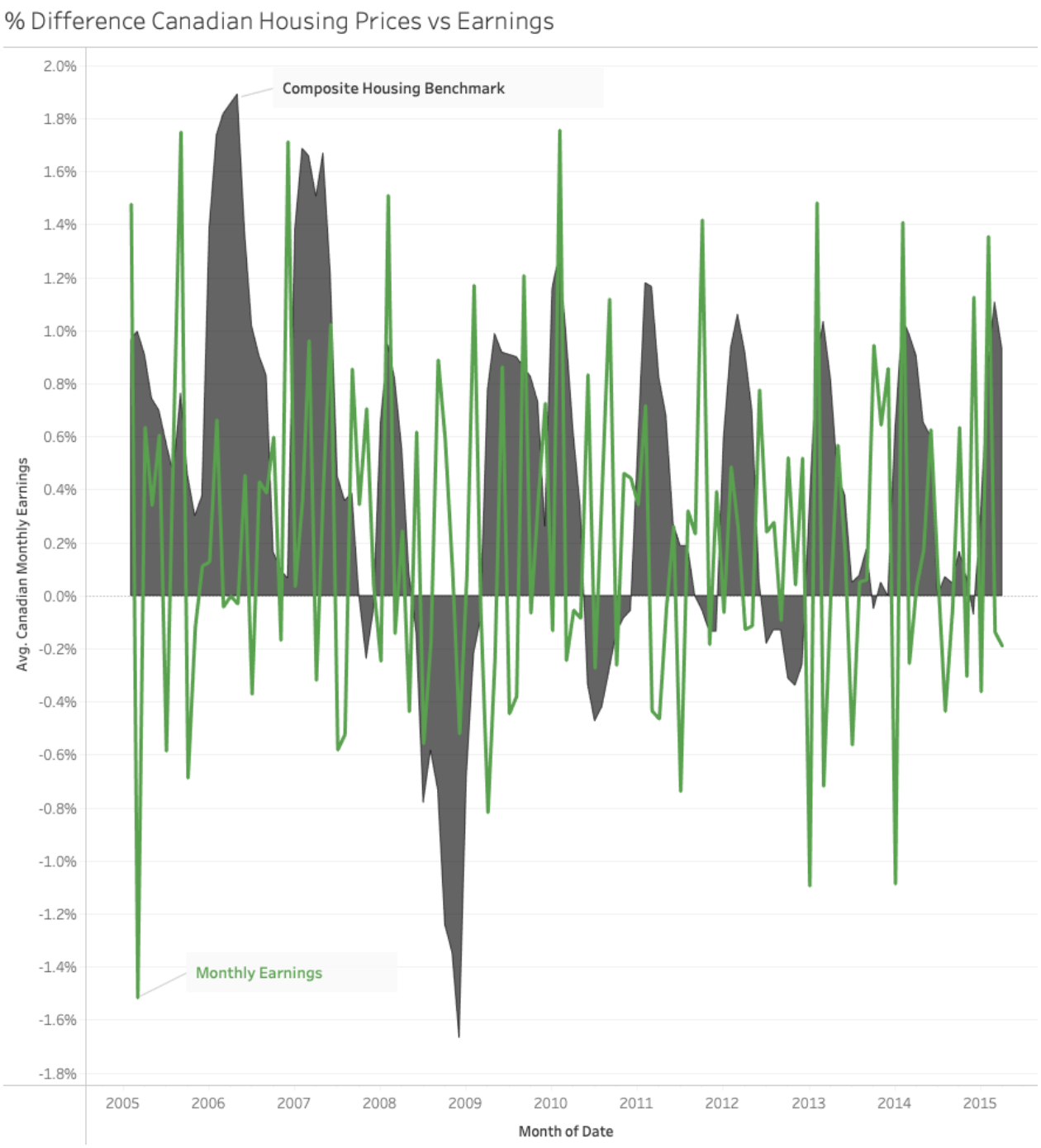



6. Compare the trend of house prices with earnings. *In case you want to plot monthly salary, be aware that the earnings value is per week.



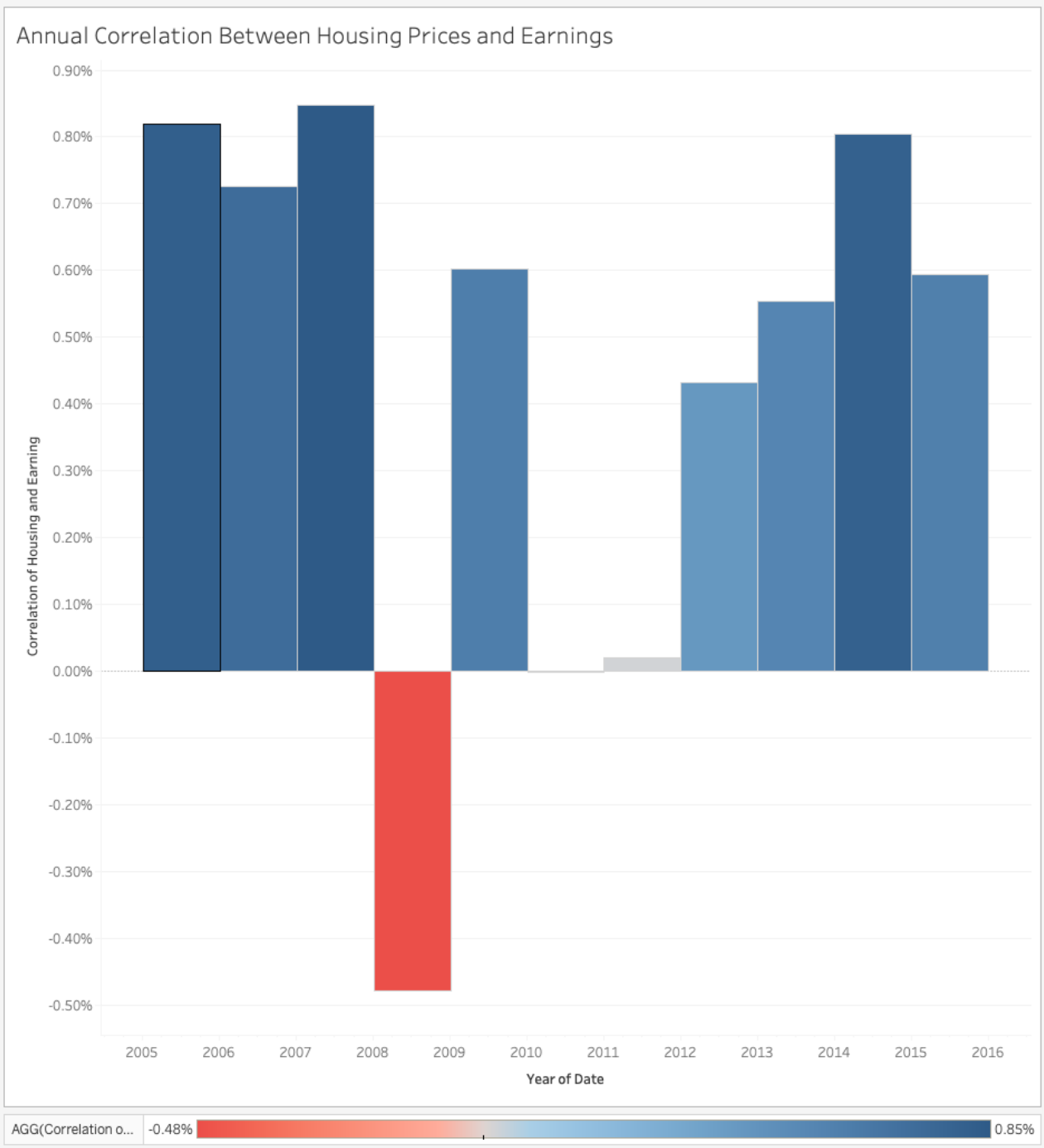

While this visualization shows the general slope over time there is a mismatch in scales that causes tremendous white space below. Given more time we could work to standardize these values to better "fill" our chart space.

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.




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

Let's check the correlation between earnings and housing prices:



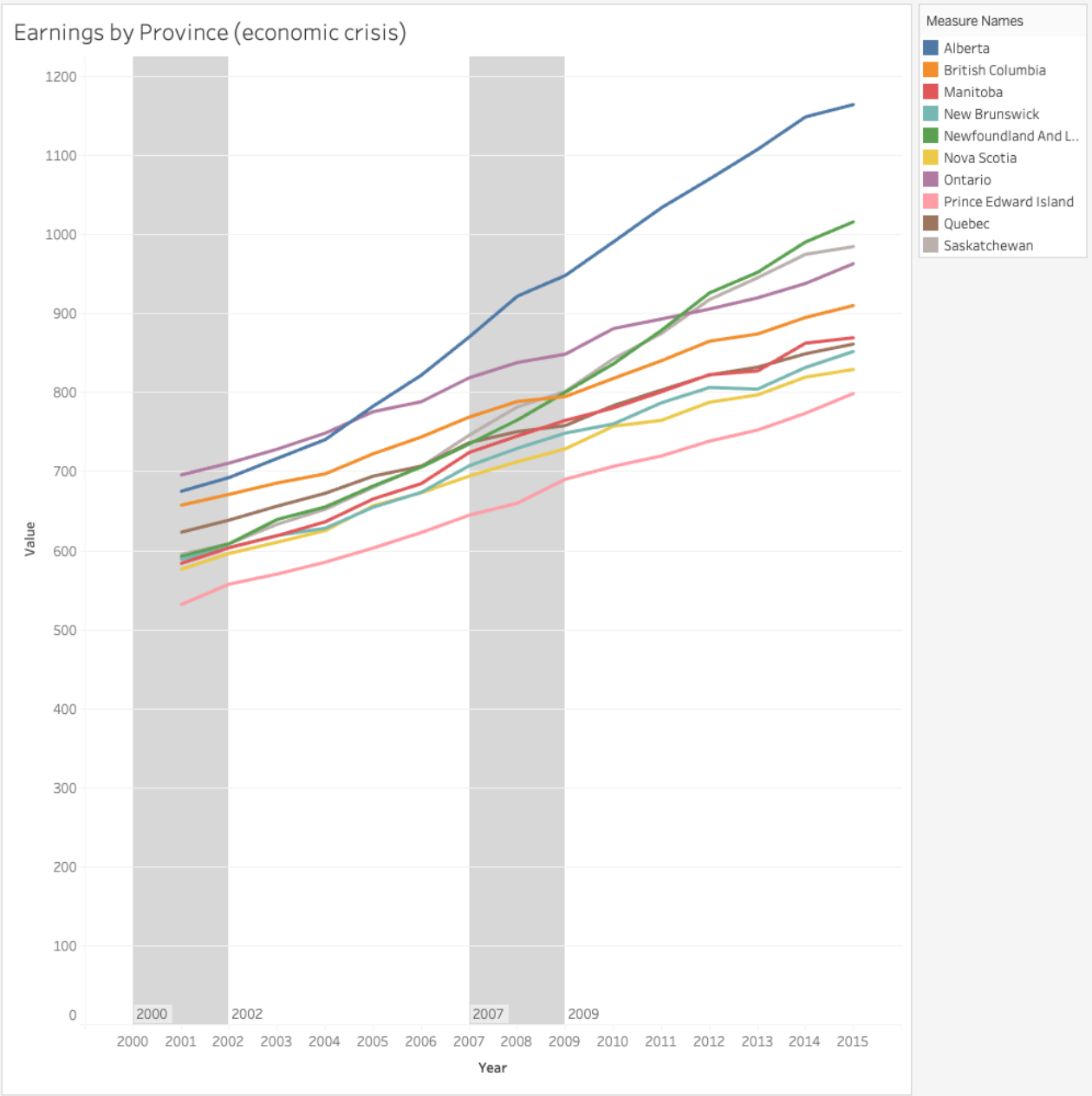

7. Did people spend more of their earnings in 2014 than they did in 2001? No expenses were provided, but we can compare earnings for 2014 vs 2001.


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 -

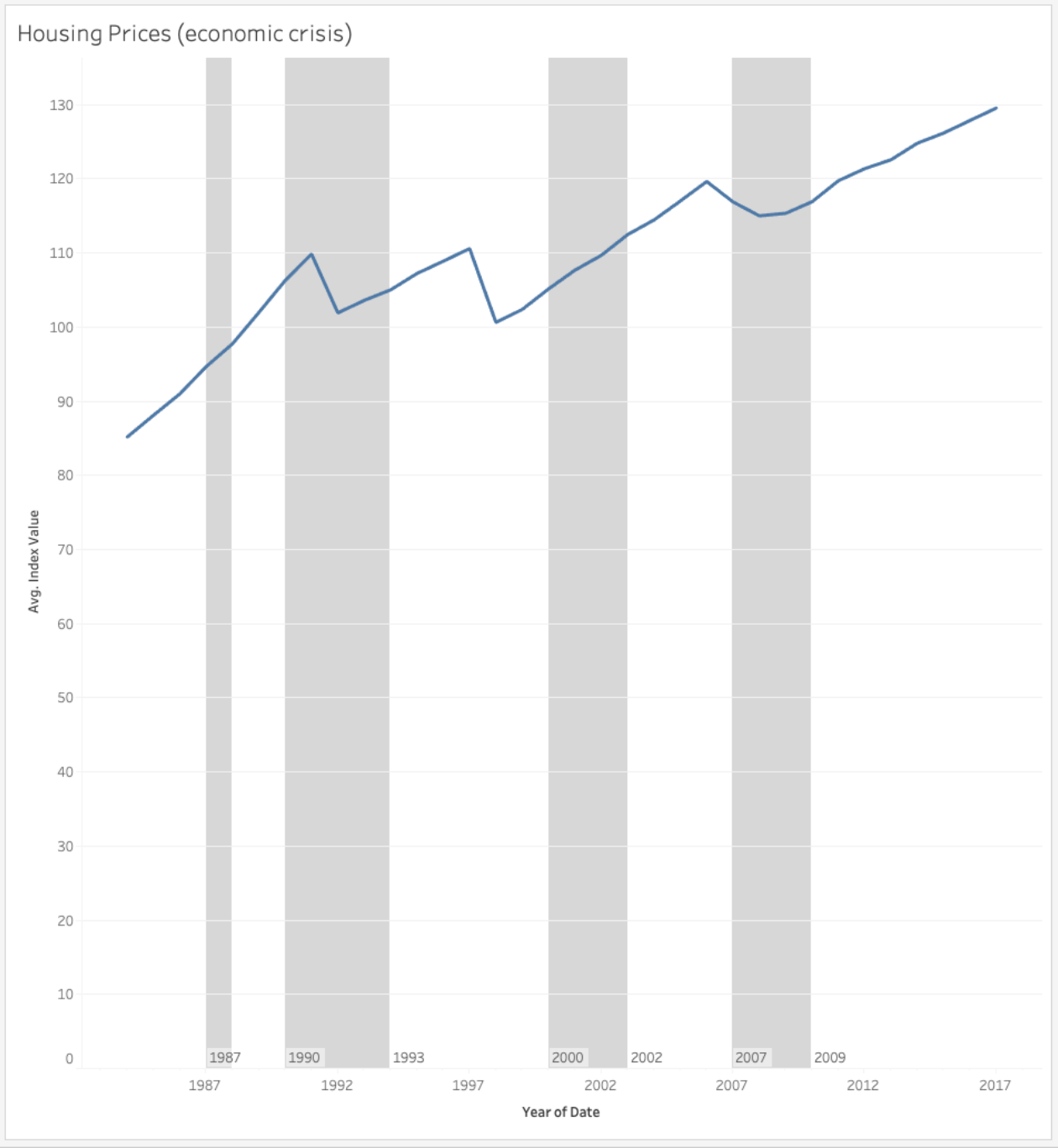
2009).

▼ Show the effect of these crises on:

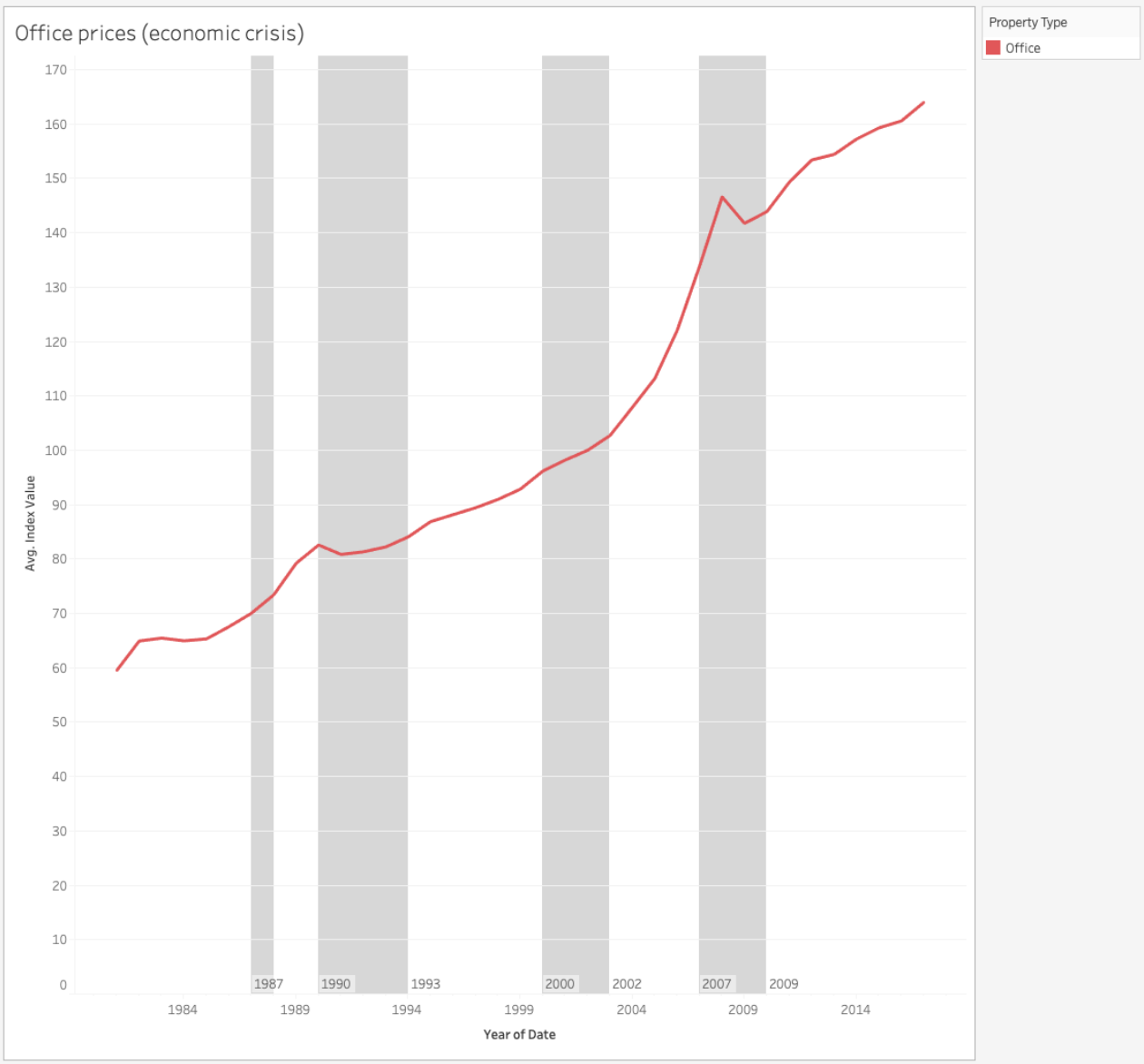
- a) Earnings



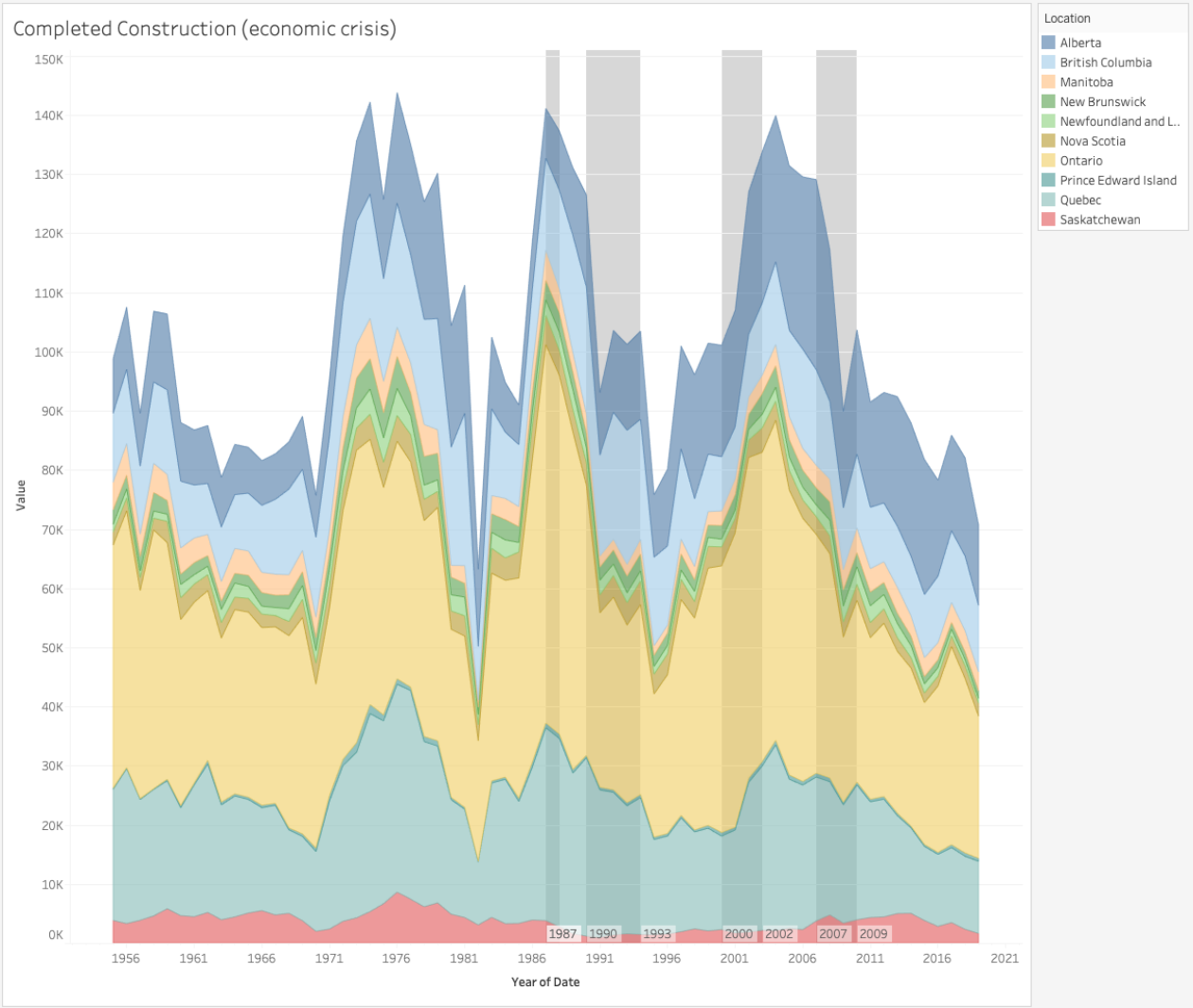
- b) House prices



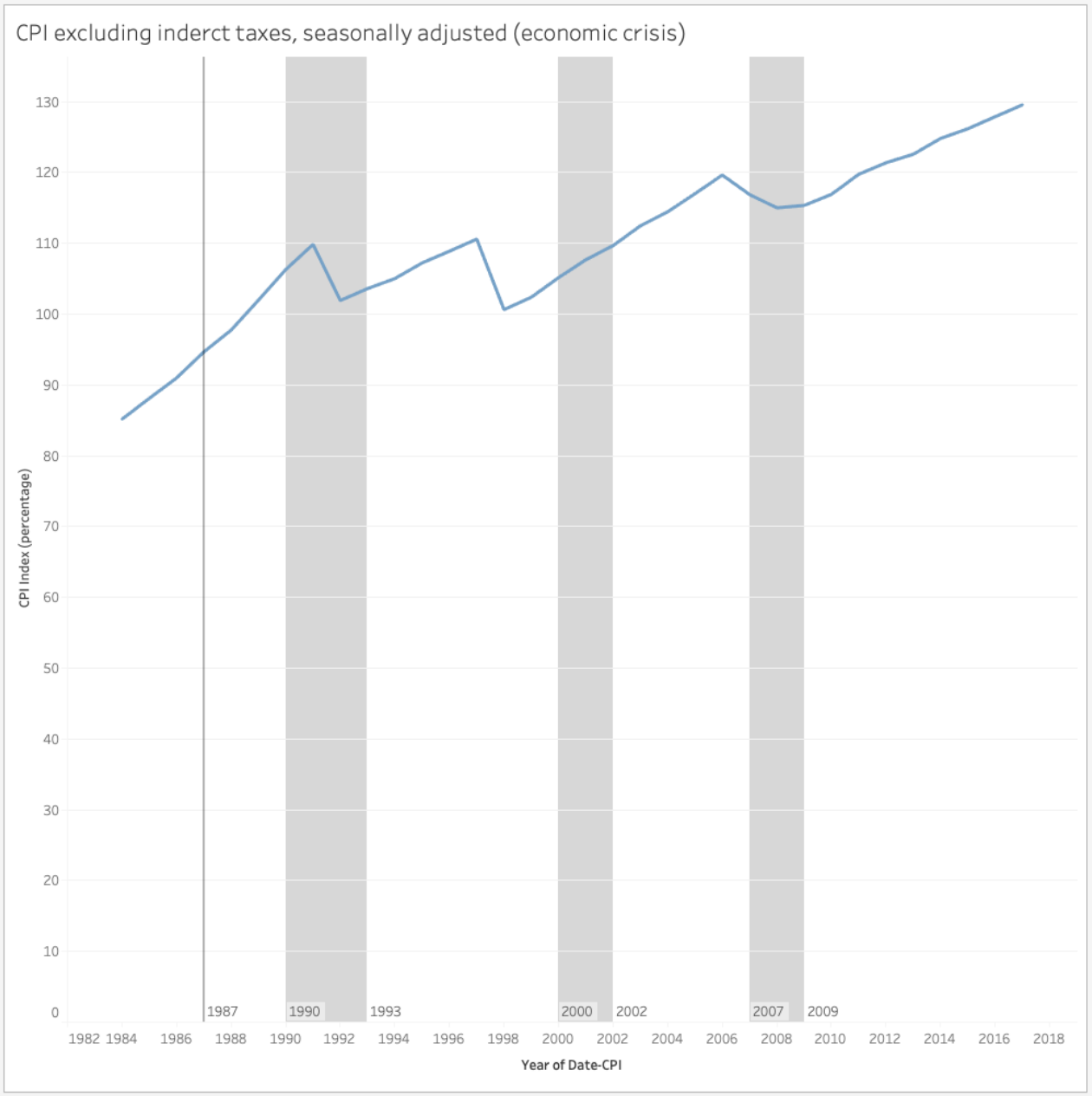
- c) Office prices



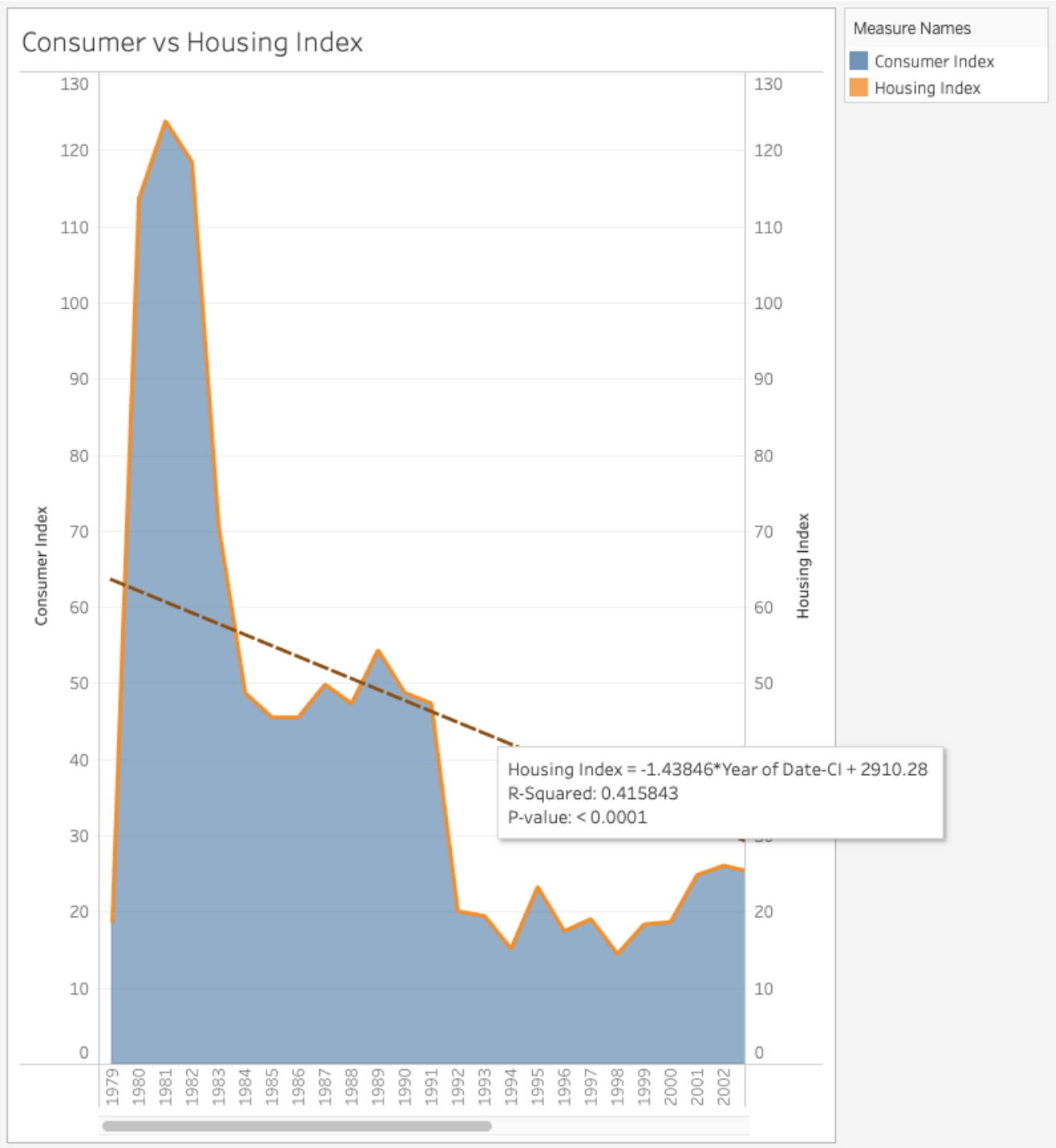
- d) House constructions



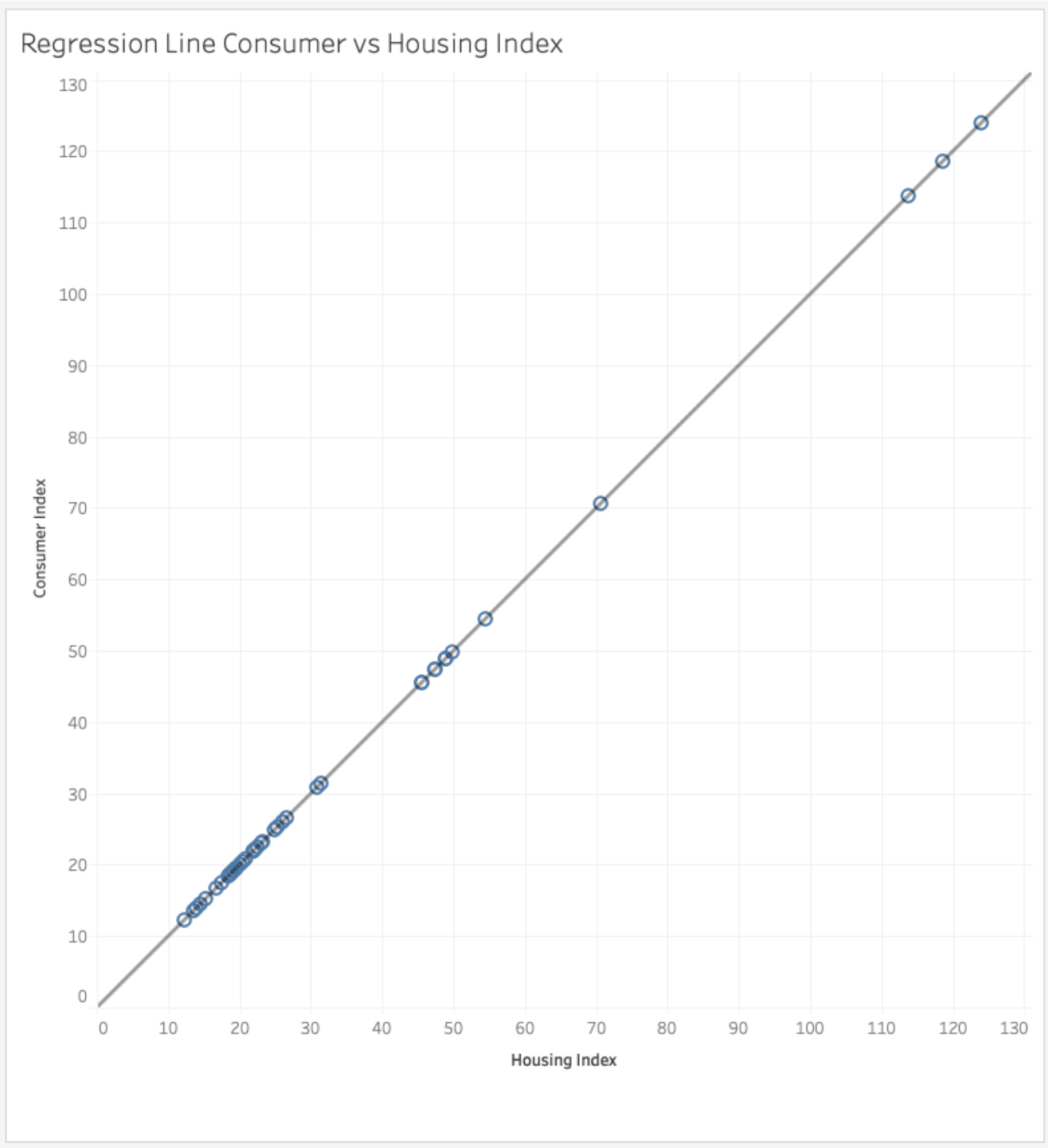
- e) Consumer index



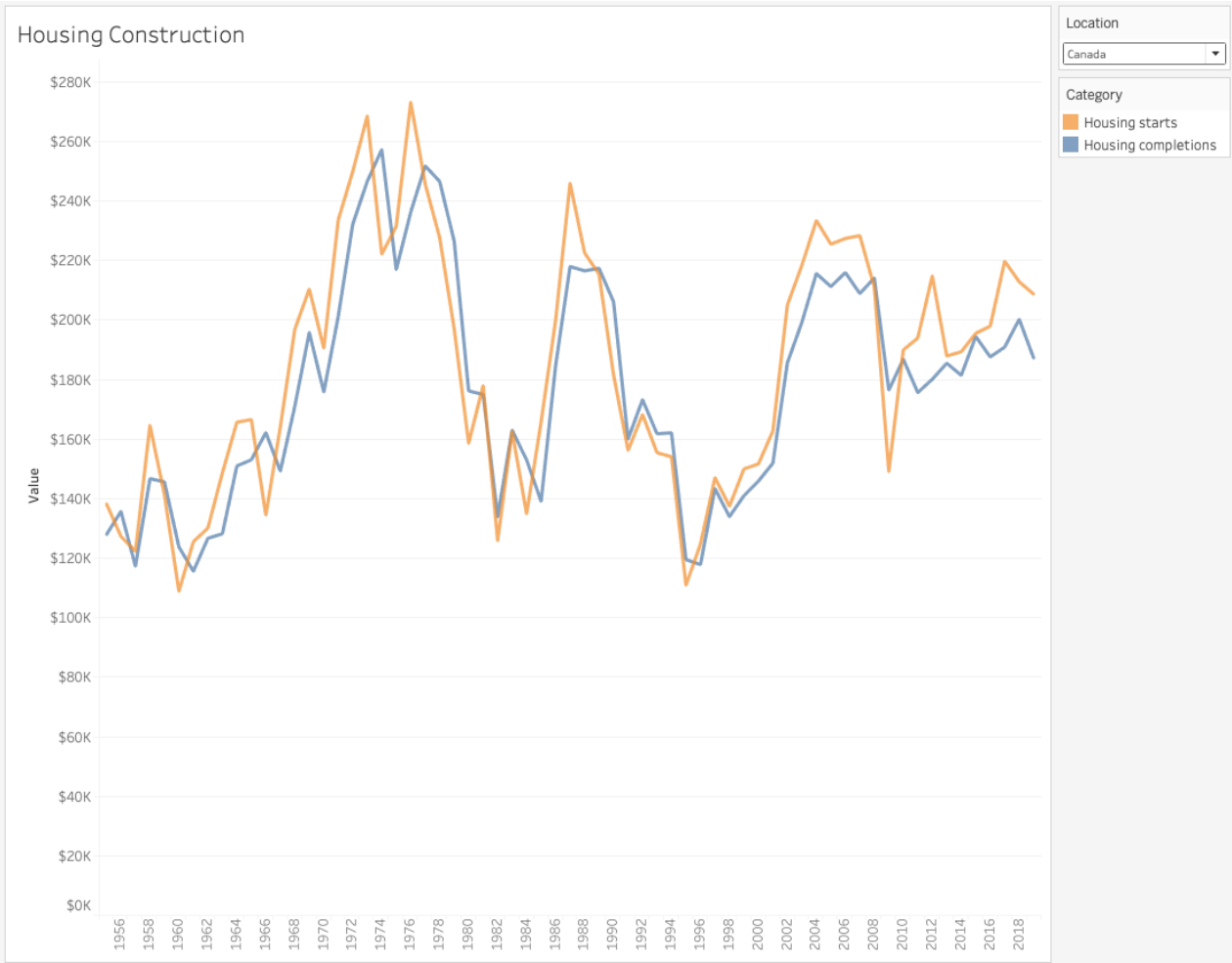

9. Plot consumer_index together with housing_price_index and fit the regression line between them. Can we predict consumer_index from the housing_price_index? Using the CPIW index value, and limiting only to percentages we can easily compare consumer vs housing price indexes.




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



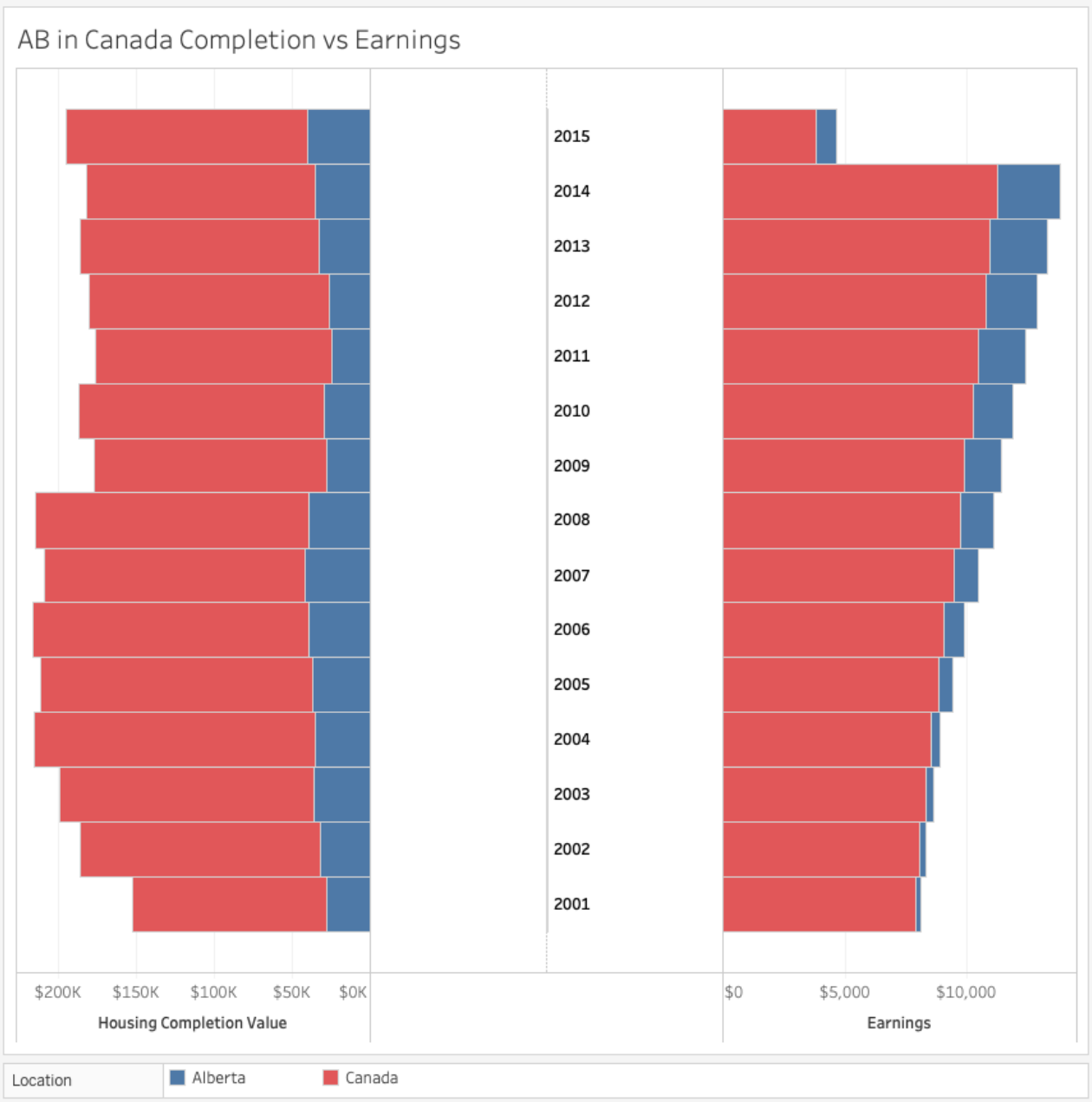

10. Try to find an interesting pattern, trend, outlier, etc. from the data used in the above questions. HINT : Double check all units in the table before any comparison.




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

- recreated earnings data to compare annually to housing, require Python modifications to original data
  Utilizing that reformated data we can now lookg at how much Alberta contributes to Canada in both construction completed and earnings. Notice there is not a direct correlation with earnings and construction completed, nor the % of earnings greater in Alberta to construction. I found this

counter-intuitive.




## 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 dashboard, but ensure collected data shares similar columns to leverage shared filters