Home Price Index with Highlighting and Custom Index Date
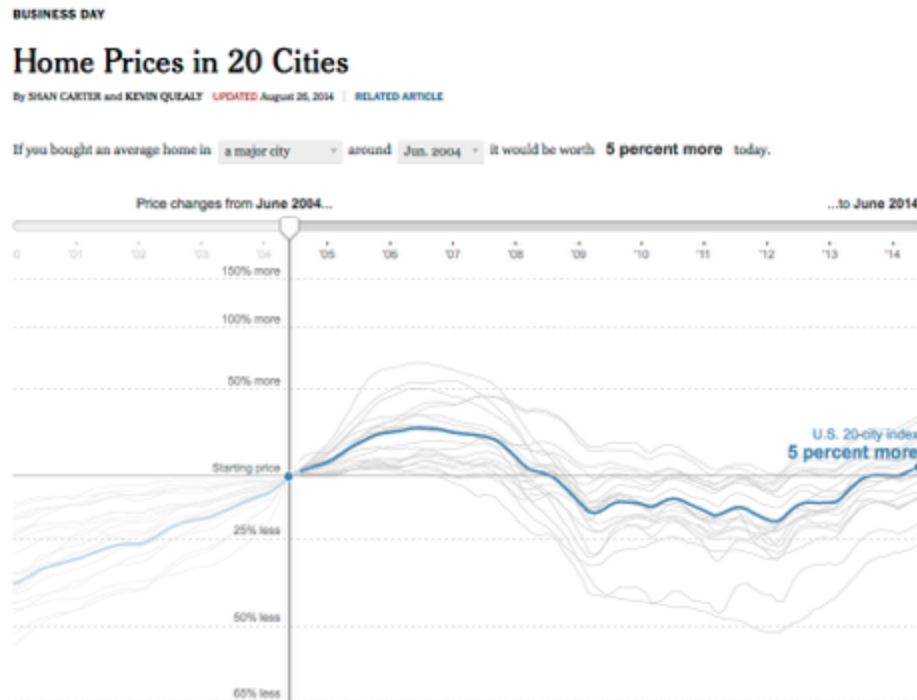
Based on work by Robert Kosara, Tableau Research
Updated by Caitlin Havlak (September 17 2016)

We will recreate this piece from the NY Times - http://nyti.ms/NUUD2e - but with Canadian housing data.

If you would like to recreate the original NYT graphic, the US Case-Shiller Index data lives here - http://us.spindices.com/index-family/real-estate/sp-case-shiller. To download it, click on the *Additional Info* dropdown and select *Home Price Index Levels* or *Seasonally Adjusted Home Price Index Levels*.

The National Bank House Price Index measures the changes in home price in major Canadian urban centres between March 1999 and July 2016.

(1) Import and reshape the data
- First connect the house_price_index.xlsx data to Tableau
- Highlight all of the columns with a city name, including the aggregate *Composite 6* and *Composite 11* "cities". Click the arrow in the last column and perform a pivot.
- Rename the columns to "Date", "City" and "Index".
- Make sure that all variables are assigned the correct variable type. City is okay as a string.

(2) Create the initial view

- Drag *Date* onto the *Columns* shelf
- Drag *Index* onto *Rows*
- Drag *City* onto *Colour* (in the Marks card)
- Click on *Date* in the Columns shelf to switch it to continuous *Month* (the *Month* item in the second list in the menu)

(3) Let's only look at data from that date forward.
- Drag *Date* onto the *Filters* card
- In the dialog that pops up, select *Range of Dates,* click Next
- On the next page, click on the *Starting Date* field at the top
- Click on the date to manually enter the date as January 2000 or select from the calendar popup. Select the first of the month.
- Click *OK*

(4) Our visualization is a bit too colourful. Also, when you click one line, you will notice that the others fade but don't change to grey. We want to have a clearer selection. To do this, we need a little workaround. We will create a parameter that contains the selected city, and then use a calculated field to compare the city name of each data point to that. If it matches, it will turn blue, if not it will be grey.
- Right-click in the data pane on the left and select *Create Parameter…* from the menu
- Give it a name like *City to Highlight*
- For the data type, choose *String*
- At the bottom for *Allowable values*, pick *List*
- In the bottom of the window, use the *Add from Field* button to select the *City* field (this populates the possible values from all the different city names in the *City* column of our data)
- Select OK to close the window
- The new parameter shows up at the bottom of the data pane on the left
- Click it to show the menu and select *Show Parameter Control*

(5) We now have a little dropdown box in the right of the view, where we can select cities. That's a start, but it doesn't change the view yet when we select one.
- Right-click in the data pane and select *Create Calculated Field…*
- Name it something like Highlight
- Type the equation [City] == [City to Highlight]
  - Words between square brackets denote variables like fields, parameters, etc.
  - The field on the left is the name of the column in the data
  - The field on the right is the parameter we created above
  - The == operator is the comparison
- Click OK to close the formula editor
- Drag the newly created Highlight field from the data pane onto Color

- If you only see two lines, drag *City* onto *Detail*. This tells Tableau how to break down the data. Since we only have two possible values from the Highlight field (*true* and *false*), that groups all the cities that aren't selected together, which we don't want.
- Click on the little triangle in the top right part of the colour legend and select *Edit Colours…* from the menu
- Select a dark blue for the value *True* and a light grey for *False*
- In the colour legend window, drag the blue colour to the top if it isn't already. That draws blue lines on top of grey lines, which is what we want.
- Now drag *Highlight* onto *Size* as well
- Click on the little triangle in the top right of the size legend and select *Edit Sizes…*
- Check the *Reversed* box to make the selected value heavier, rather than the unselected ones
- Adjust the size range with the slider and click Apply until you like the way it looks, then click OK to close the window.

(6) We have finished the visual design part, and now move on to adjusting where we calculate the index from.

- Create a new parameter called *Index Date* and set its data type to *Date*
- For the allowable values, select *Range* and set 2000/01/01 as the minimum and 2015/07/01 as the maximum
- Click OK, then click on the parameter in the Parameters area within the data pane and select *Show Parameter Control*
- Click in the data pane to create another calculated field and name it *Index Relative*
- Enter the following expression: SUM([Index]) / LOOKUP(SUM([Index]), FIRST() + DATEDIFF('month', #1/1/2000#, [Index Date])) * 100
  - This demo shows you the power of Tableau and calculated fields, but you're not expected to be able to come up with these formulas on your own! The steps below explain what's in this formula if you're curious.
  - This expression consists of two parts: SUM([Index]) / the Index at the location we want to do things relative to, and then multiply by 100
  - The second part figures out where the index for a particular data, specified by our parameter *Index Date*, can be found
  - The innermost part, DATEDIFF('month', #1/1/2000#, [Index Date]), calculates the difference between the date specified by the *Index Date* parameter and January 1, 2000 (the #1/1/2000# syntax is used to specify date constants).
  - We're doing this for each city separately, so we need to keep track of where in the table the first value for that city is, which is what the FIRST() function tells us. In the table on the right, the first row in each section is pointed to by that function.

City	Year of Date	Month of Date	Index	Relative ..
Atlanta	2015	January	118.8	100.0
		February	118.9	100.1
		March	119.9	100.9
		April	121.5	102.3
		May	123.2	103.7
		June	124.8	105.1
		July	125.9	105.9
Boston	2015	January	175.7	147.9
		February	175.4	147.6
		March	176.2	148.3
		April	176.6	148.6
		May	179.3	150.9
		June	181.9	153.1
		July	184.0	154.8
Charlotte	2015	January	129.0	108.6
		February	130.2	109.6
		March	131.2	110.5
		April	132.5	111.5
		May	133.5	112.3
		June	134.3	113.0
		July	134.5	113.2
Chicago	2015	January	126.6	106.5
		February	126.5	106.5
		March	127.8	107.5
		April	129.3	108.8
		May	130.9	110.2
		June	132.2	111.3
		July	133.4	112.2

- ○ The LOOKUP function then pulls a particular value of Index from the location we specified so we can divide the value we're interested in by it.
  - ○ Click OK to close the formula editor
- ● Drag the newly created field into the *Rows* shelf and drop it on the *Index* field, so that it gets replaced.
- ● Now when you change the date selected in the Index Date control, the visualization will update and show the recalculated data.

(7) Finally, we want to see the index date on the visualization and also show a line at 100 as a reference.
- ● At the top of the data pane on the left, click on the *Analytics* tab
- ● Drag a *Reference Line* from the *Custom* area into the view and drop it in the main view.
- ● In the window that appears, click on the drop-down next to *Value* and pick the *Index Date* parameter
- ● Next to the *Label*, pick *Value* (or *Custom* and enter text like *Index Date*)
- ● Click OK. This shows a vertical line at the index date (try changing it to see what happens)
- ● Drag a *Constant Line* from the *Summarize* section and drop it on the *Index Relative* field that appears.
- ● In the text window that pops up, enter 100 to set the line to the 100 reference value
- ● This now shows a horizontal line at 100
- ● There is also an average horizontal line that appears. Click on the horizontal line and select Remove

(8) To keep the view stable as you make changes, double-click on the vertical axis and select a *Range* of *Fixed*.
- ● Set the start to 25 and the end to 375.
- ● Delete the title as well for a cleaner look.
- ● Click OK.

Now as you change the index date, the values are always comparable because the visualization does not try to rescale to the data.

You visualization should look something like this:

Canadian housing index
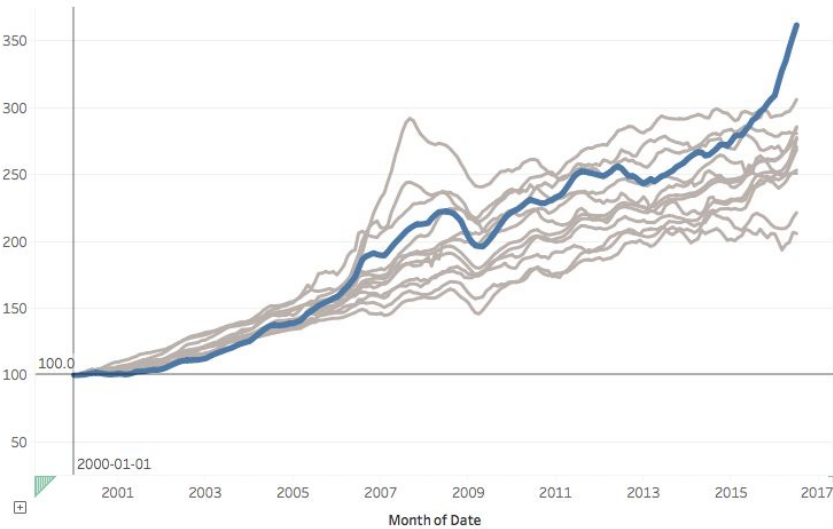
Highlight
- True
- False

City to Highlight

BC Vancouver

Highlight
- True
- False

Index Date

2000-01-01