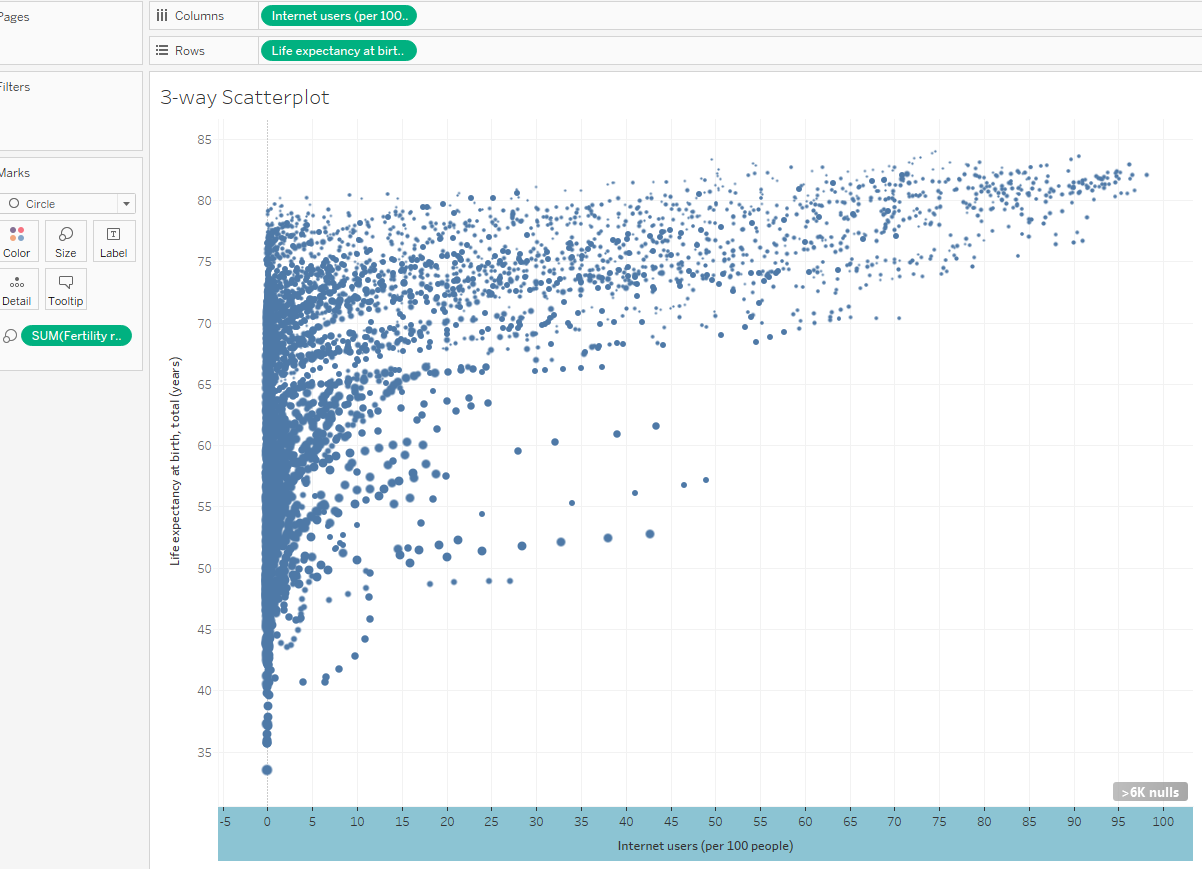
**3-way scatterplots**

We now turn to a common way of visualizing relationships between 3 quantitative variables, the 3-way scatterplot. You may be familiar with the bubble plots at [Gapminder.org](http://www.gapminder.org/tools/#_locale_id=en;&chart-type=bubbles), which are essentially 3-way scatterplots that we can re-create in Tableau. Along the way, we will encounter the following Tableau topics:

* Group creation
* Parameter creation
* Pages (aka animation)

**Task: visualize the relationship of life expectancy with internet usage and fertility rate across the world.**

Whenever you are interested in the relationship between 2 quantitative variables, you should immediately think “scatterplot.” 3-way scatterplot simply indicates the color or size of the points are encoding some third variable. 3-way scatterplots are easy to create in Tableau. Here’s the first pass; note that we must set life expectancy and internet usage as **Dimensions** so they do not aggregate:



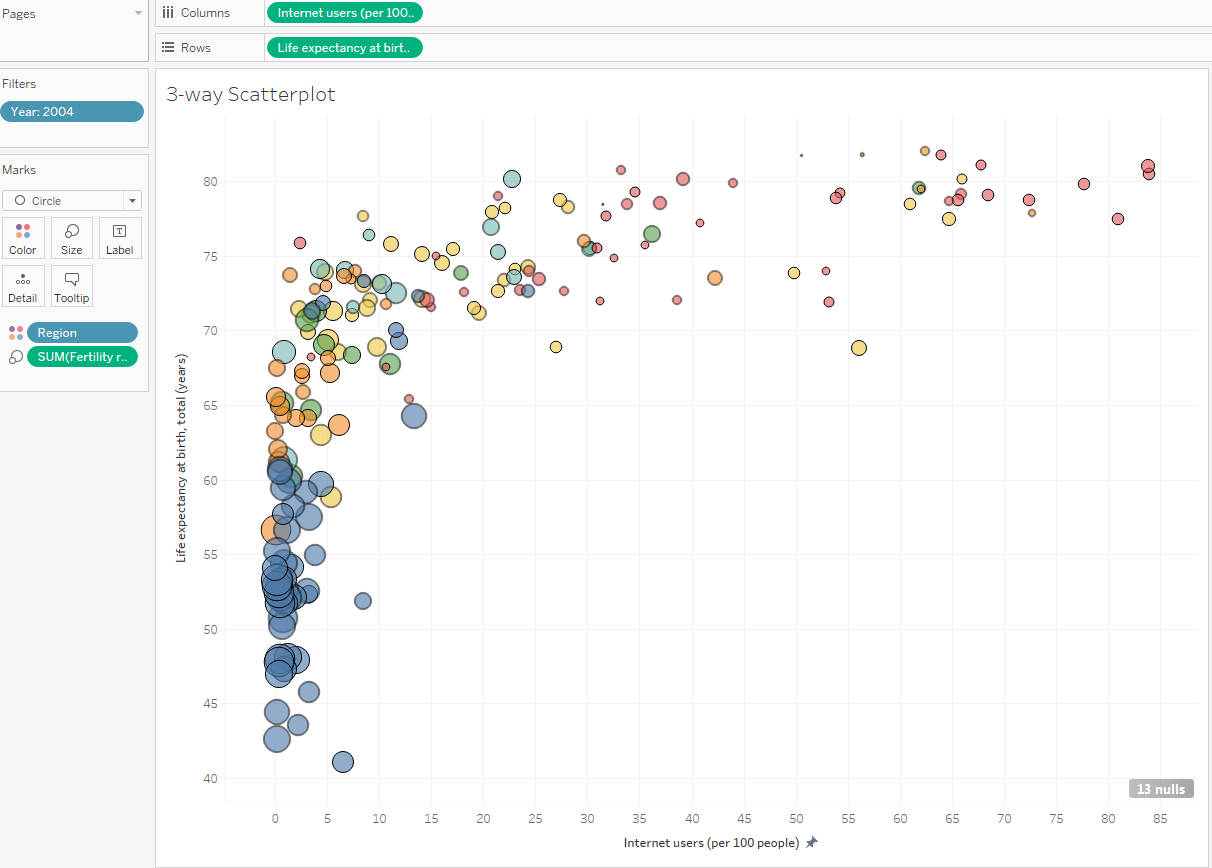
There are several issues we should be able to identify right away:

* This visualization includes many years’ worth of data
* There are a lot of overlapping points (transparency might help)
* The X-axis includes -5. Axes should never include data values that are impossible or meaningless.
* A more subjective point: I don’t like the drop line at 0. Why do we need this?

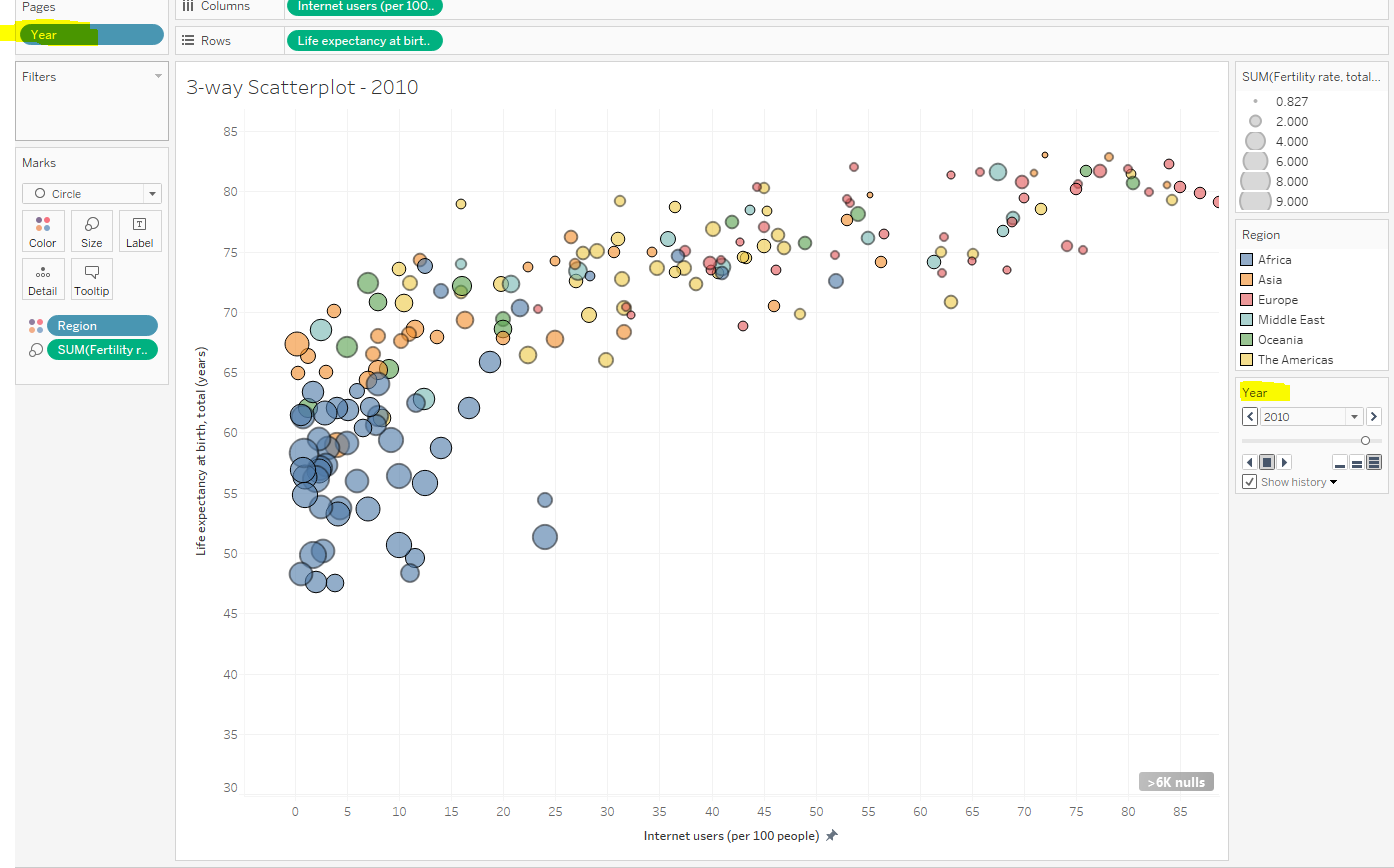
Let’s start editing:

* Add a filter on Year
* Use the Color menu on the Marks card to add borders to and make the points transparent
* Edit the axis. Try making it start at 0. If this doesn’t look good, try starting it at -4.9 so that you get the aesthetic look you want without the -5 tick mark label.
* Get rid of the 0 drop line (maybe grid lines too?) using the Format menu.

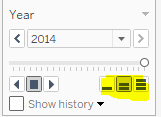
Edited visualization looks much better (here I have also color-coded by region; we are up to a 4-way scatterplot):



It might be interesting to see how this relationship changes over time. If we drag Year to Pages instead of Filter, we get a filtered-by-year scatterplot with a “control panel”:

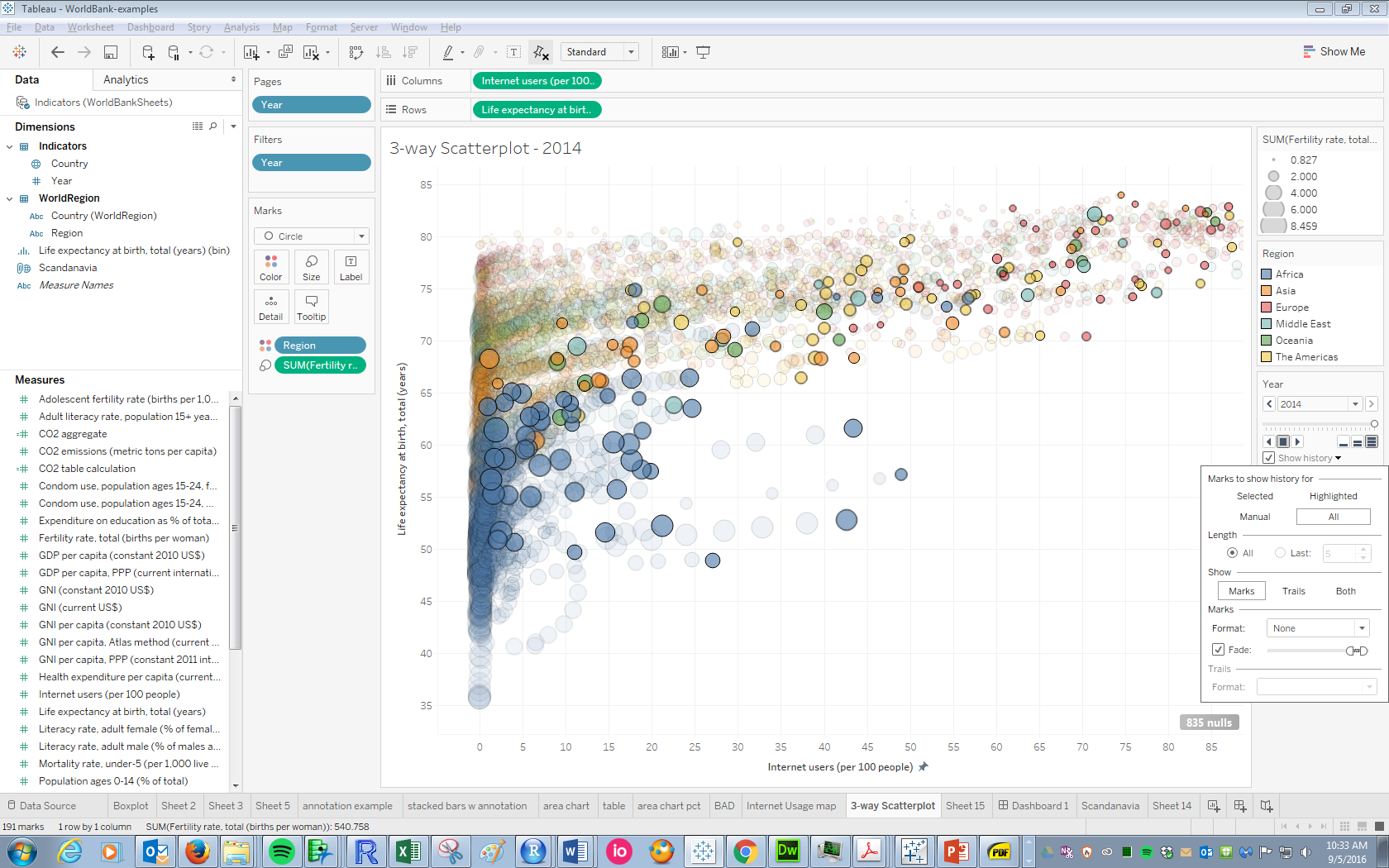


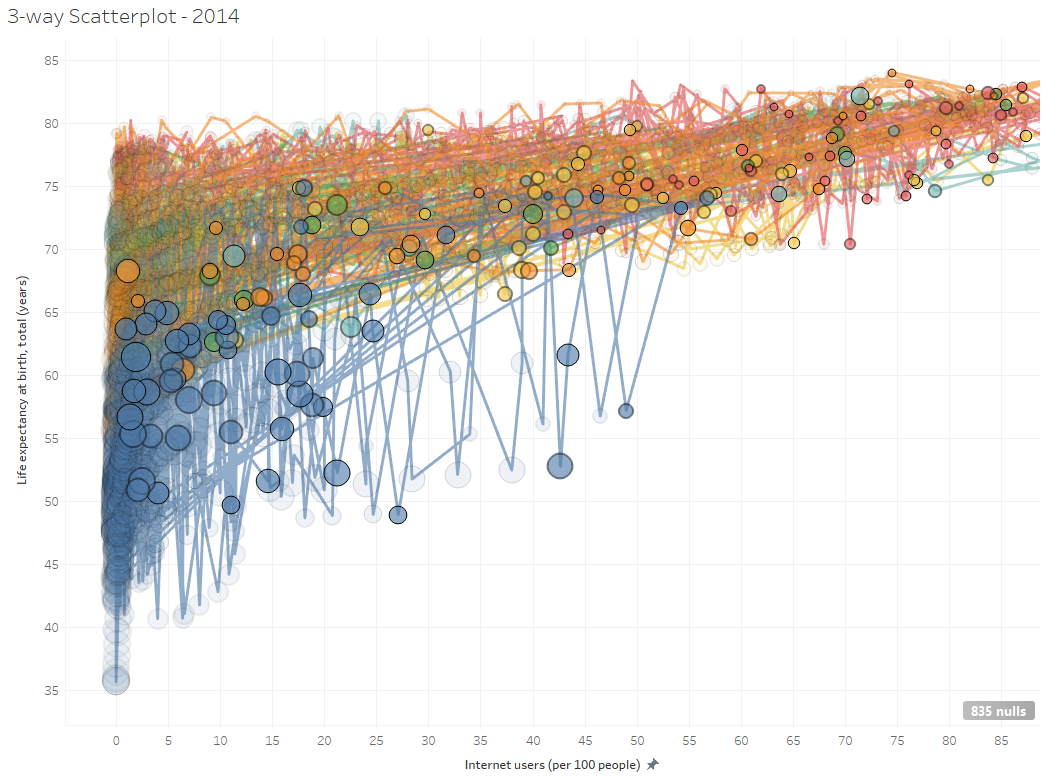
Let’s take a look at that control panel:



Clicking “Play” will now show the scatterplot for each year. The three highlighted buttons on the right control the speed of the animation. Note that Internet Usage is only measured for most countries starting in 1991 and does not have 2015 data available yet; hence we probably want to add **another** instance of Year to Filter, and include only the years 1991-2014. You can also check “Show history,” and have the points “tracked” over time.

Adding all the bells and whistles:



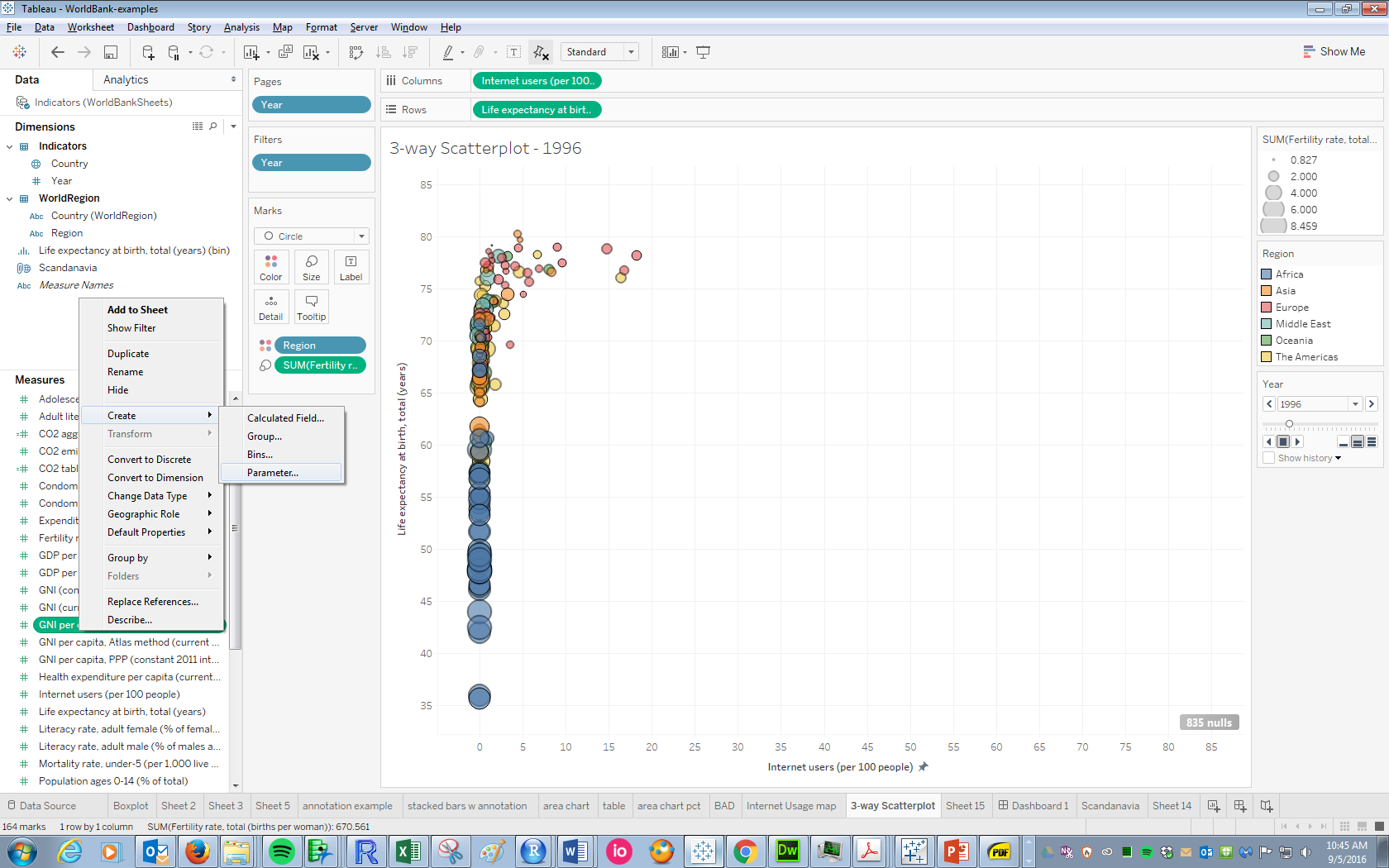
You can also have the trails of the pages tracked; but this is not recommended! (Unless, that is, you want your visualization to look like a two-year-old finger-painted all over it):

For most animations, tracking history makes the most sense if the X-axis is time. For this scatterplot, there is not much gained by tracking the history and it clutters the image. Hence, history probably shouldn’t be tracked here.

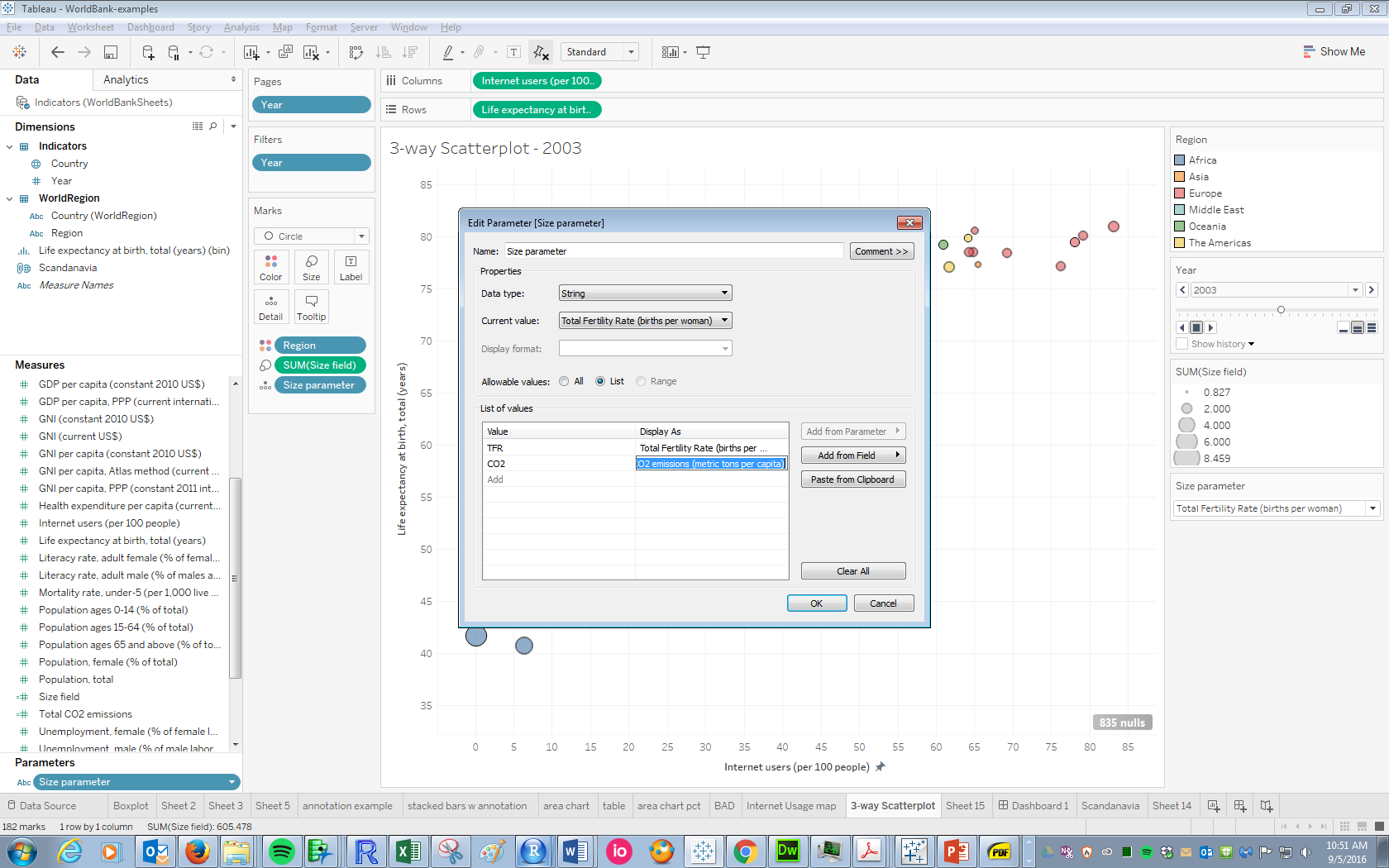
**Changing the “Size” variable using a parameter**

Suppose we want the viewer to have the option of sizing the points by Fertility Rate or CO2 emissions, similar to the Gapminder tool. We can allow this flexibility using a ***parameter***. ***Parameters*** allow for entire fields mapped to certain EPTs on the visualization to be selected by the viewer.

* Right-click in the data window 🡪 Create 🡪 Parameter:



* Complete the parameter dialogue box as shown below:



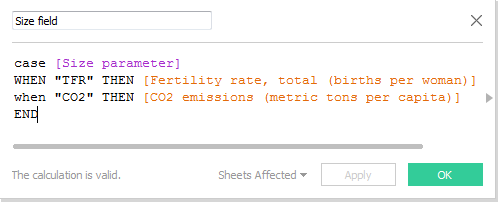
The important elements are the “Data Type”, “Value”, and “Display as.” We want a customized list of strings, and we want the display to show the full names of the variables. Under “Value,” we list some shorthand; we will use this to create a calculated field that depends on the selected parameter value.

Click “OK,” then right-click on the newly created parameter 🡪 “**Show Parameter Control**.” Note the control that appears on the worksheet:

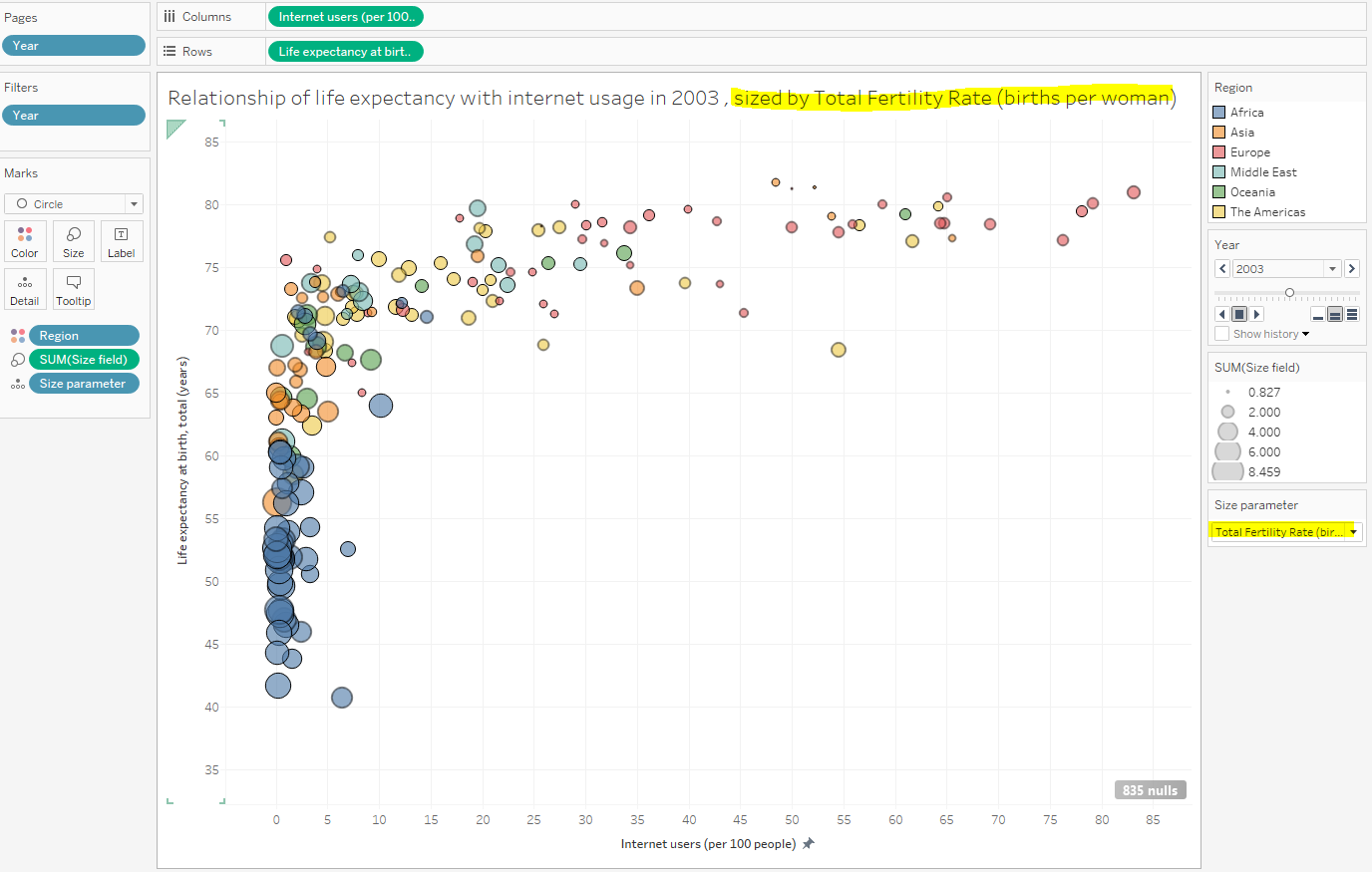


Right now, this doesn’t control anything. We need to create a calculated field that will respond to the selected values on this control.

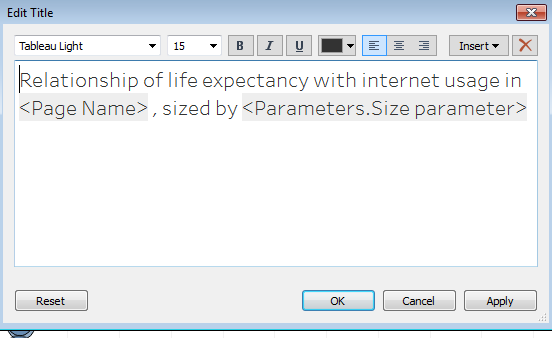
* Right-click in the data pane 🡪 Create 🡪 Calculated field. Fill out the dialogue box as shown. Note the use of “TFR” and “CO2,” which come from the parameter values we specified:



* Now drag this new “Size field” to size. Sizing will now be determined by whatever parameter value is selected:



* We can also get the title to reflect the sizing in the “Edit Title” dialogue box. Double-click on the title, then click “Insert” to have the title reflect the currently-selected parameter value (note that “Size Parameter” must be on Detail to have the option of adding it via the Insert menu):

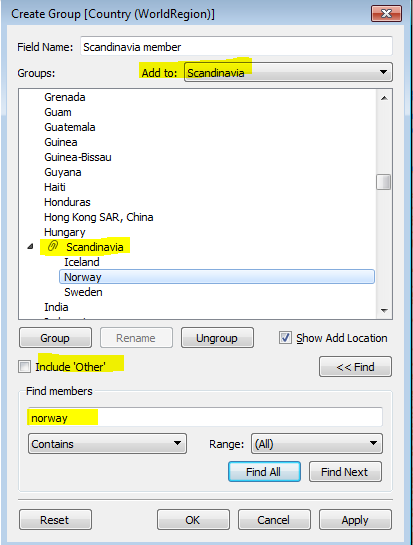


Note that a parameter could also very similarly be used to control which field is visualized on the X- and/or Y-axis.

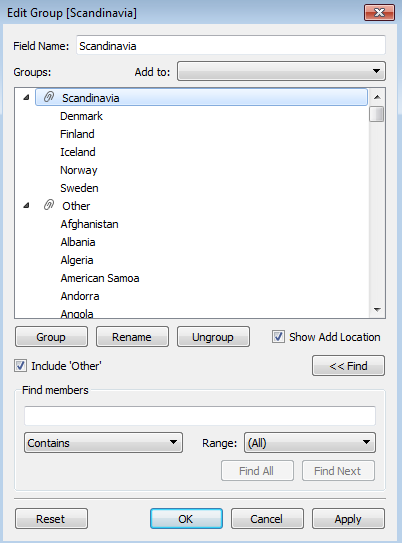
**Grouping**

One thing we notice from the scatterplot is the prominence of internet usage in the Scandinavian countries. We can emphasize this on its own worksheet by creating a new field which essentially groups all countries into one of two groups: “Scandinavia” (Denmark, Finland, Iceland, Norway, and Sweden) and everyone else.

* Right-click on the Country pill 🡪 Create 🡪 Group
* Call the field something intuitive. Then start searching for the Scandinavian countries using the “Find members” window.
* Repeat for each of the Scandinavian countries, adding them to the “Scandinavia” group:



* Check the “Include Other” box. Your final group should look like this:

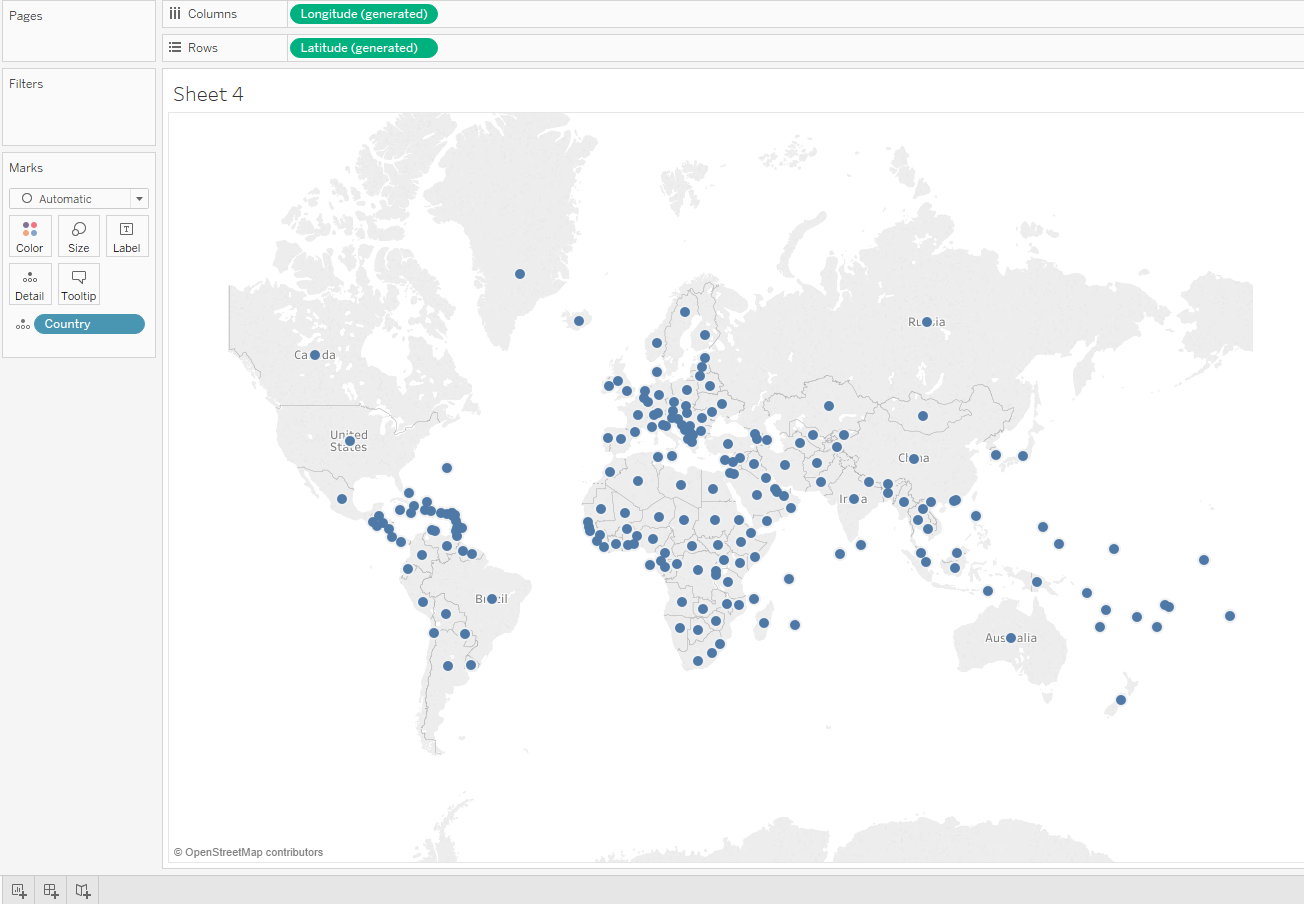


You can now use this field just as you would any other categorical (“discrete”) field.

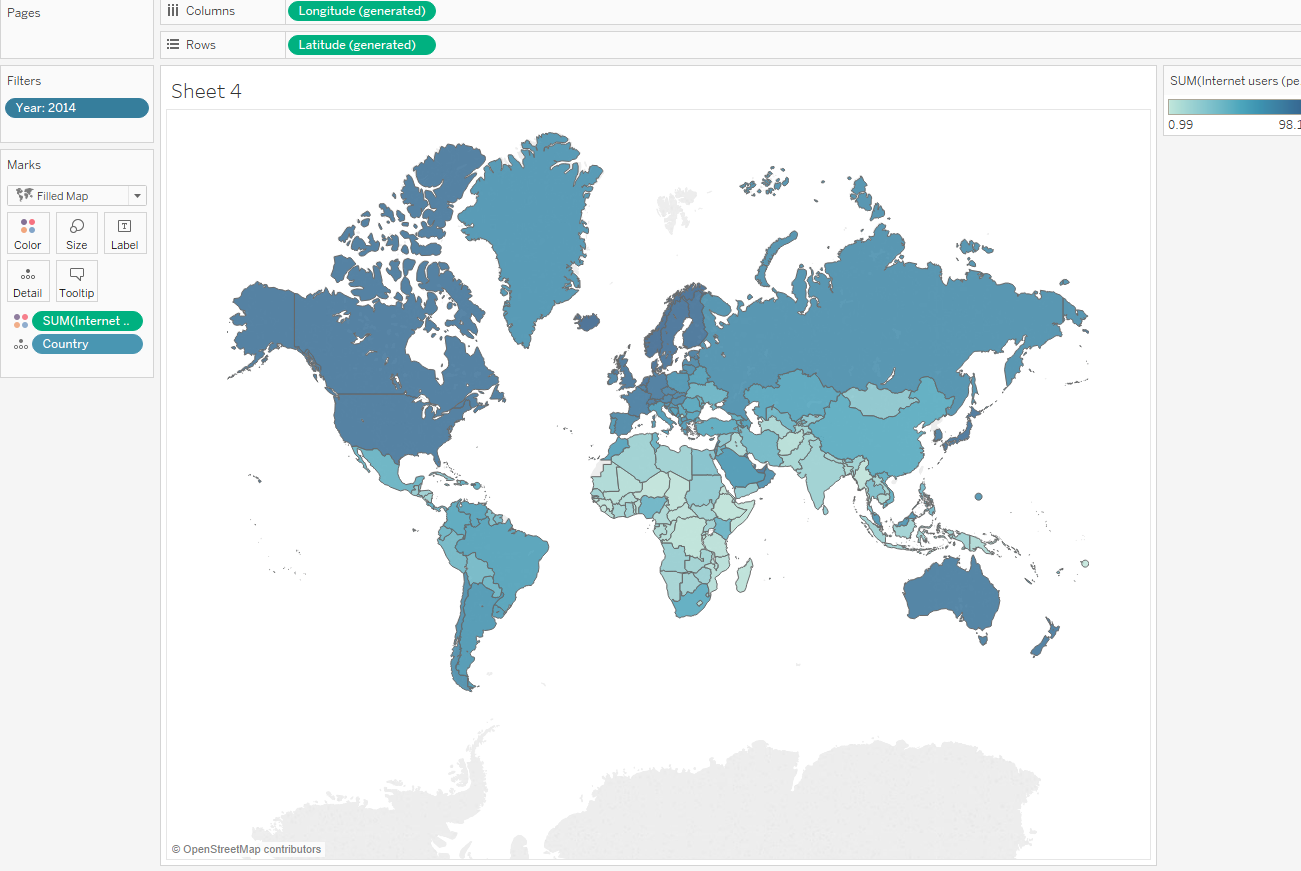
**Geographic fields & maps**

Tableau has extremely powerful geographic capabilities. Fields like countries, states, zip codes, counties, etc. can often be recognized as geographic fields. Tableau then creates customized latitude and longitude coordinates with shape files for each row of the geographic field.

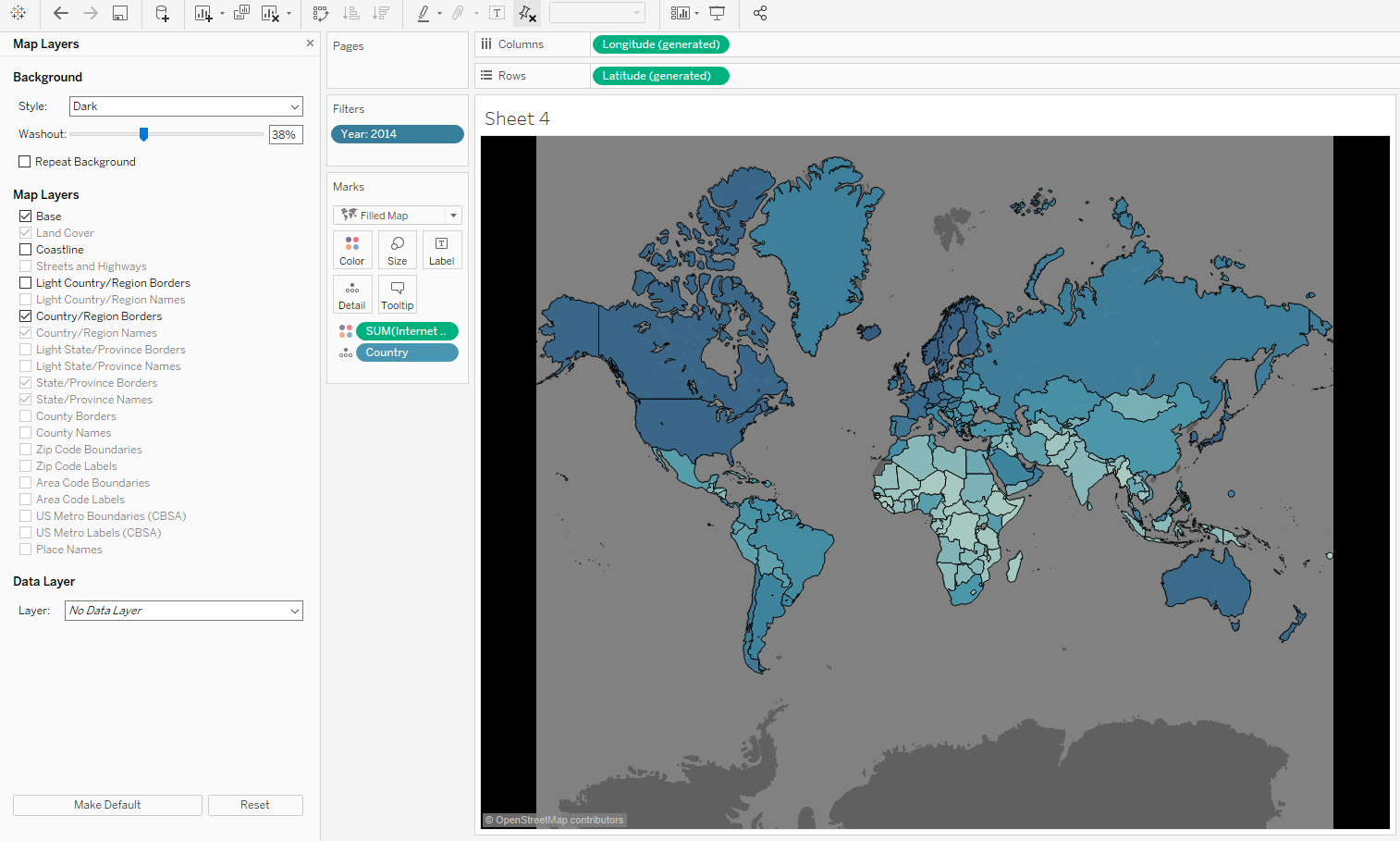
To create a map, simply drag a geographic field to Detail on the Marks card. Note that Latitude and Longitude appear automatically in Rows and Columns, and a map underlies the data points:



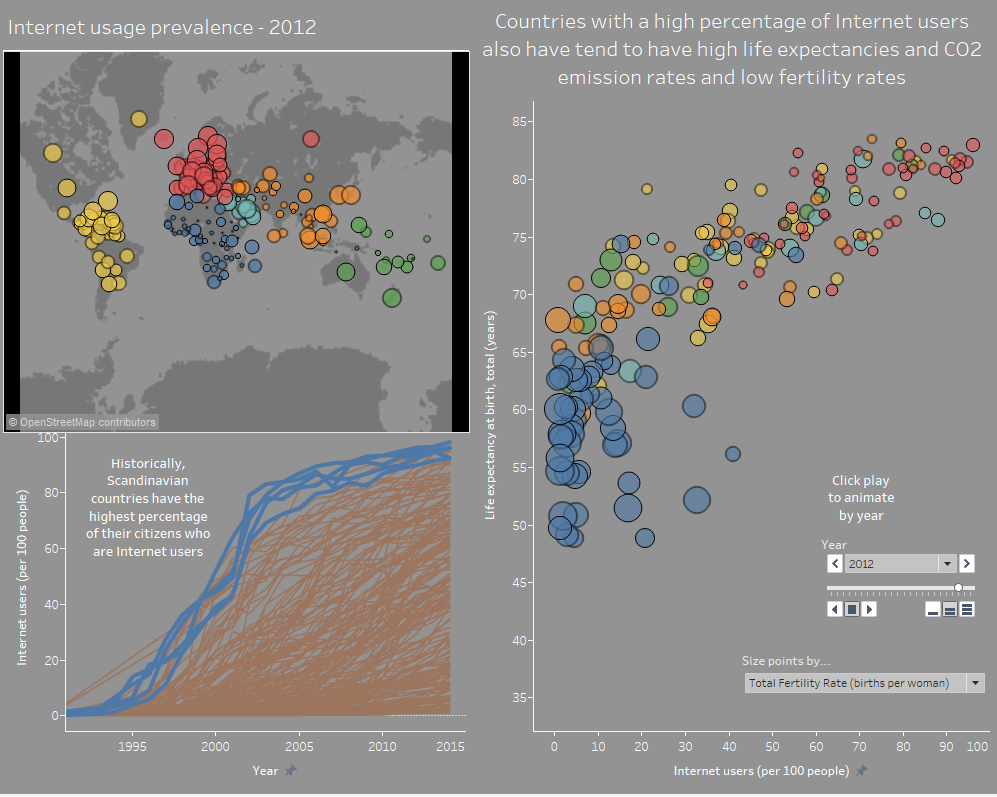
Changing Marks to “Filled Map” and color-coding by 2014 internet yields the following:



Clicking **Map 🡪 Map Layers** brings up the following, which allows for several useful options that are more applicable when dealing with smaller geographic regions (most are grayed out for countries):



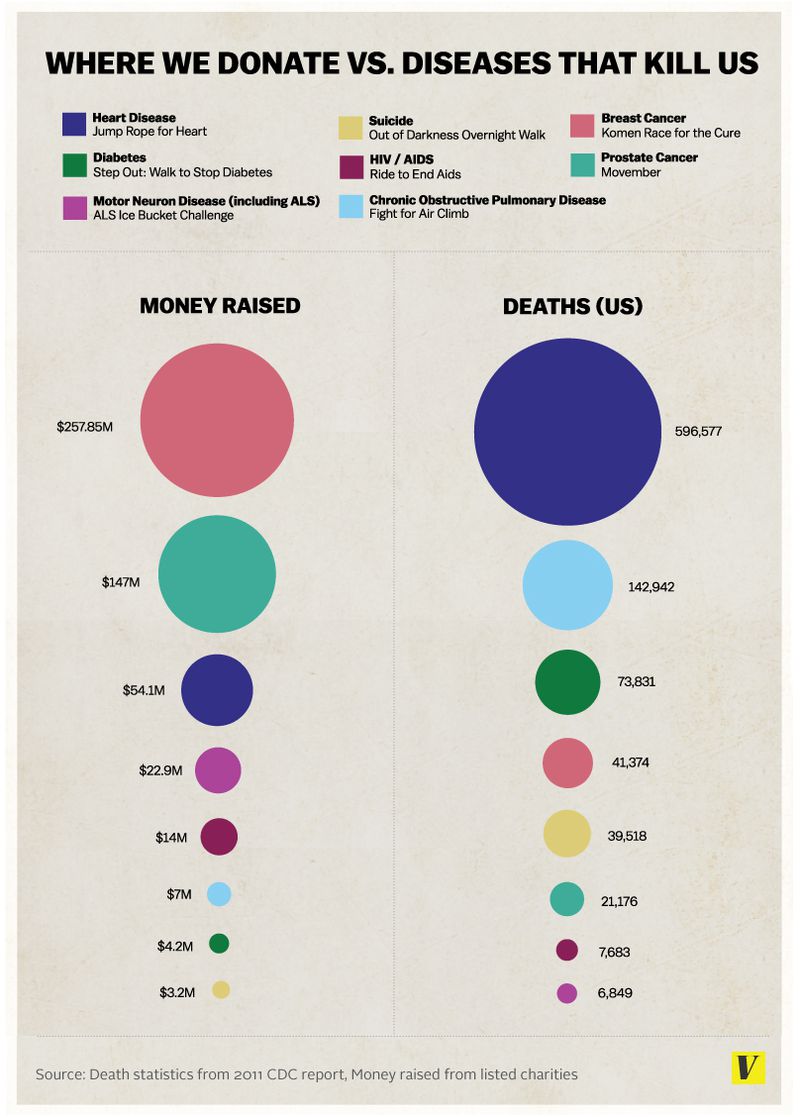
**Task**: re-create this dashboard that pulls all the information we’ve visualized together:



Unfortunately, Tableau Public does not allow “click play to animate,” yet, so one might want to reconsider the animation feature if publishing and sharing.

**Visualization critique/Task #1**

Consider the following visualization from [Vox](https://www.vox.com/2014/8/20/6040435/als-ice-bucket-challenge-and-why-we-give-to-charity-donate), visualizing data on money raised towards various causes and CDC data on death from those causes:



* What are the variables? What EPTs are used to visually map the variables?
* What are some issues with this visualization?
* The data are in Vox.csv. Create one or two alternative visualizations that improve upon the original.

**Task #2**

Using the IPUMS-I data, visualize inequality in educational attainment by sex. Do this for the United States in 2010, *restricted to those 18 or older*. You will need to work with the following variables:

SEX

AGE

EDATTAIN

GEOLEV1

Specifically, your task is to visualize state-by-state inequality across sex in the percent of people 18 or older who have completed at least a secondary education.

CREATIVE CHALLENGE: Visualize this question in 2 ways:

1. With a map, where the color or mark size for each state represents that state’s education inequality;
2. Without a map, using a single worksheet to encode all of the following quantities:
3. Each state’s female educational achievement (% of 18 or older females who have completed at least a secondary education);
4. Each state’s male educational achievement (% of 18 or older males who have completed at least a secondary education);
5. The gap between male and female educational achievement in each state.