Visualizing Time-series Data in Venngage

Part of a Series of Tutorials on using Google Sheets to work with data for making charts in Venngage

TERC

CU University of Colorado Boulder

Sometimes, you will have data that shows how some quantity changes over time, on a yearly, monthly or possibly even daily basis. A good way to visualize such data is in a time series graph. This document will demonstrate how you can get time-series data from the Web and put it into Venngage.

We'll start with data from a website that specializes in time-series data called Gapminder. Gapminder has a repository for time-series data presented in a very useable format at https://www.gapminder.org/data/. The data are taken from reputable international data sources such as the World Health Organization, the World Bank, and the International Labour Organization. These sources collect and aggregate massive datasets from around the world, on topics varying from demographics and population to health, education, and trade. In all of these data files, each case is a single country.

Let's pick something not too morbid like Cell phones (per 100 people). This is more useful than Cell phones (total) since it gives us a sense of the proportion of people who own cell phones rather than the total number, which would naturally be larger in country with a larger population.

Indicator name	Data provider	Category	Subcategory	Download View Visualize
Cell phones (per 100 people)	World Bank	Infrastructure	Communication	
Cell phones (total)	World Bank	Infrastructure	Communication	

Download the Excel file for this dataset by clicking on the Excel icon under "Download." The file will be saved to your default download location.

In Google Drive in your web browser, create a new Google Sheet file, going to New -> Google Sheets.
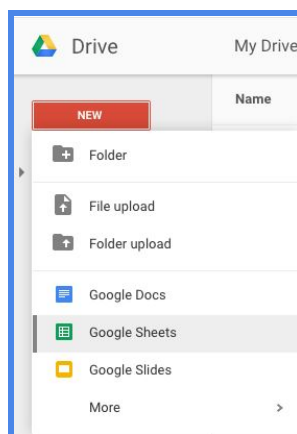
Once there, go to File ->Import... and then upload the file.
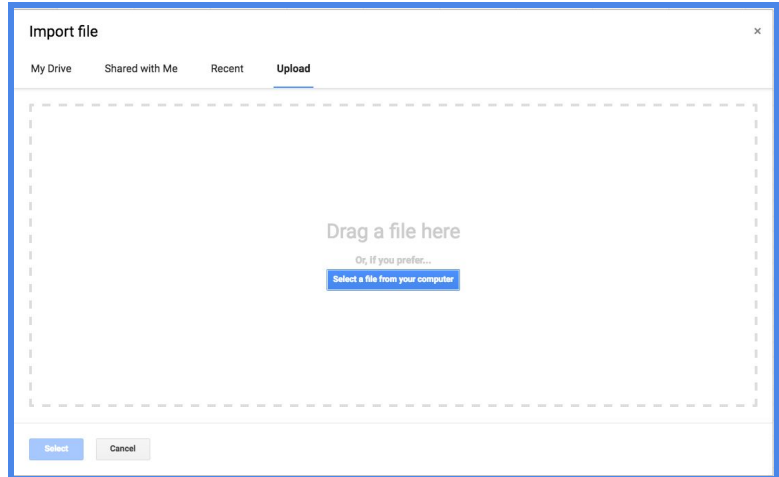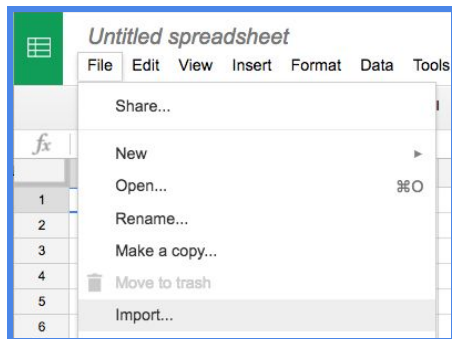
If Google Spreadsheets ask whether you'd like to replace the spreadsheet or create a new one, select the option to replace the spreadsheet. You should be starting with a blank spreadsheet, otherwise, choose to create a new one. Your file will look like this:

Notice that for these data, each row (or case) is a country and each column is a year. The data start at the year 1965 and go to 2011 in single year increments. Since these numbers represent cell phones per 100 people, a number less than 1.0 means that fewer than one person out of 100 owns a cell phone and a number greater than 100 means people own more than one cell phone per person, on average. Since this is an overwhelming amount of information, let's focus on just 2 countries for now.

As we scroll through the data, we see that most countries have 0 or a blank (probably the data weren't collected that year) for the years between 1965 and 1980. Finland is the

country that has the earliest non-zero value, starting in 1980, which suggests that people started owning cell phones in Finland before many other countries. Let's make a graph comparing the growth of cell phone ownership over time in Finland to the United States.

In order to visualize the data from Finland and the United States in Venngage, it will be easier if we have each case be a YEAR, rather than a COUNTRY. This means we want each row to be an individual year, and there will be two columns, one each for Finland and the United States. Create a new Google sheet with this structure. The first row will be for the column labels, so enter "Year" in A1, "Finland" in B1, and "United States" in C1.

Back in the original sheet, start at Finland 1980 and copy the values all the way up to Finland 2011. Return to your new sheet, right-click (or control-click on a Mac) on cell B2 and select Paste special -> Paste transpose. This function in Google Sheets will make what used to be a row of values into a column of values, which is just what we want. To fill in the row labels for the years, enter "1980" in A2, and "1981" in A3. Then select A2 and A3, and click and drag the bottom right blue corner square of the selection to the last row that has a value in column B. It will automatically populate the cells with the years 1980 to 2011. Go back to the original data and copy and paste-transpose the data from the United States just as you did with the data from Finland, starting at 1980 even though it's a 0 for that year. There will be blank entries for the United States 1981 to 1983, but that's fine. The spreadsheet you just created should look like this and is ready to be transferred into Venngage:

|	A	B	C
1	Year	Finland	United States
2	1980	0.4913078556	0
3	1981	0.5890111245	
4	1982	0.7020069567	
5	1983	0.8700613469	
6	1984	1.065945789	0.03835937875
7	1985	1.379766579	0.1410971093
8	1986	1.733305806	0.280015091
9	1987	2.144160676	0.5005083055
10	1988	2.790161329	0.8331700865
11	1989	3.825464864	1.398778386
12	1990	5.171463976	2.085369002
13	1991	6.371068203	2.954234206
14	1992	7.667109057	4.271692476
15	1993	9.664807758	6.138520395
16	1994	13.28187422	9.160251427
17	1995	20.34389743	12.68593777
18	1996	29.30636347	16.34893582
19	1997	42.07950682	20.28741762
20	1998	55.25085886	25.07710643
21	1999	63.41410637	30.80808942
22	2000	72.07342603	38.75379151
23	2001	80.52925839	45.00169847
24	2002	86.90604073	49.15635015
25	2003	91.10525765	55.14660515
26	2004	95.44462404	62.85011191
27	2005	100.4892511	68.62738254
28	2006	107.6731658	76.64460341
29	2007	114.924474	82.47195844
30	2008	128.4719884	85.6752031
31	2009	144.1530224	89.14911634
32	2010	156.3972049	89.85645095
33	2011	166.0238042	105.9136009
34			

A graph of time-series data is just like a line chart, so to start, drag a line chart to the Venngage canvas. Double click to edit it. Delete the filler data that is already there, and then copy-paste your extracted data into the associated data table. Since Venngage interprets the first column as the X-axis variable and other columns as values for the Y-axis, you'll get this:
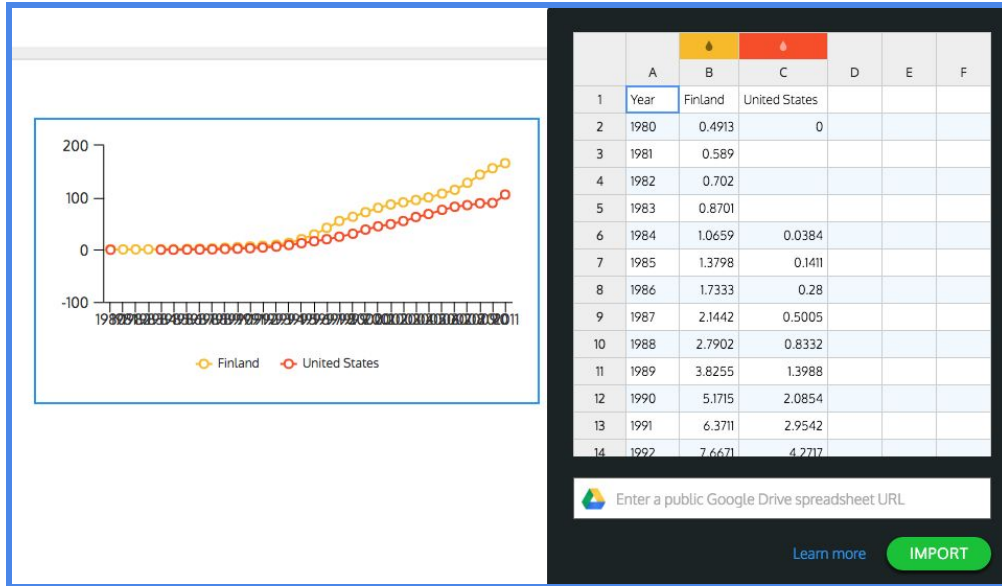
This graph is what we want; there are two time-series on the graph, one for Finland and one for the United States, but the x-axis is too crowded and it's impossible to read the years. One solution is to only use data for every fifth year, since it looks like the trend for cell phone ownership mostly goes up in both countries and doesn't have any sudden spikes that we would miss by deleting some of the interim data. In Venngage, select and delete multiple rows and leave only the rows for 1980, 1985, 1990, 1995, and so forth. You can select an entire row by clicking on the row number. Drag over other row numbers to select multiple rows. Delete the entire row by right-clicking and selecting "Remove row" from the pop-up menu.
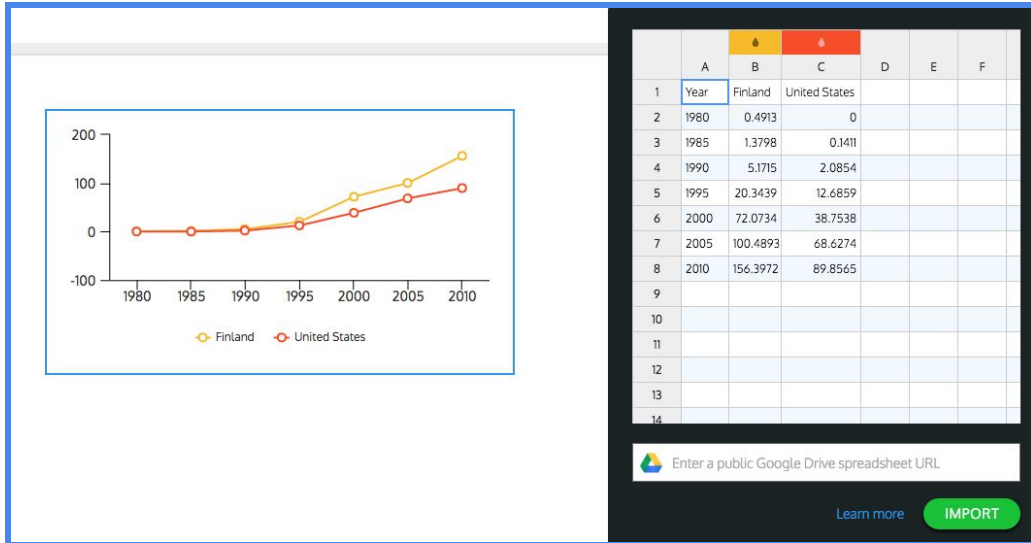
That looks much better! In the chart settings add a y-axis label based on the spreadsheet information, "Cellular subscriptions per 100 people".
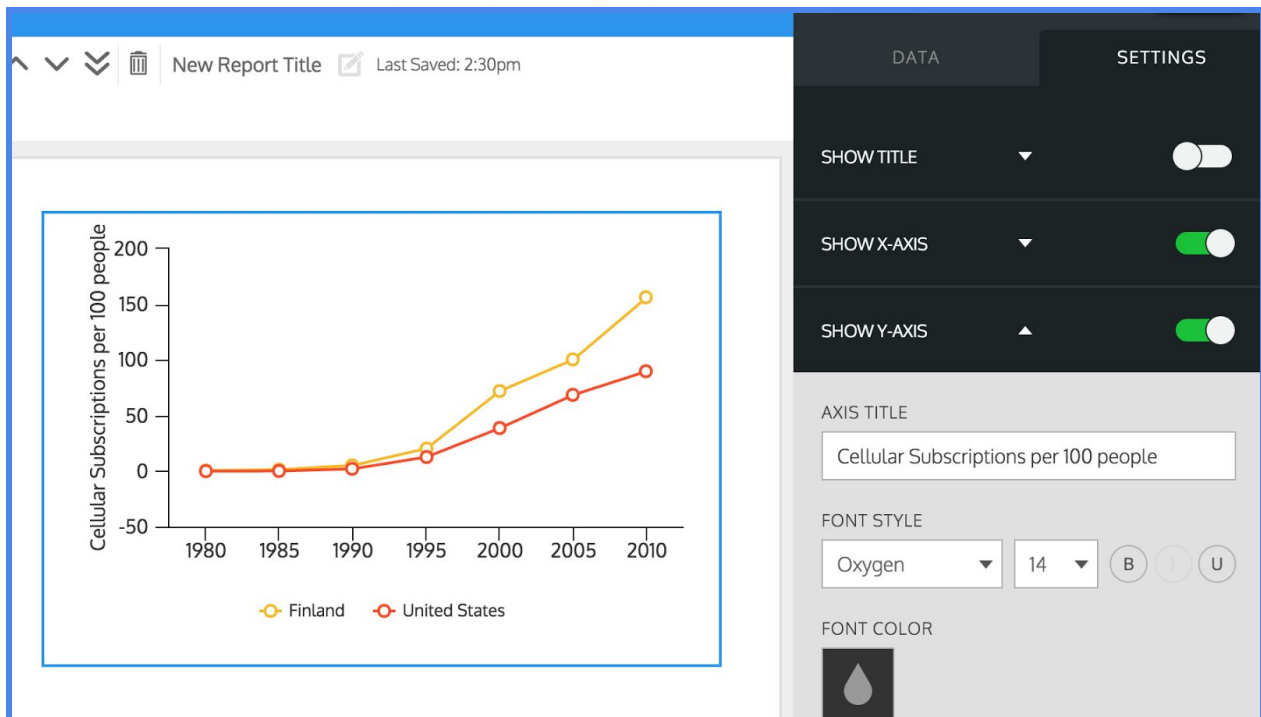
Now the numbers on the y-axis make more sense. You could add "Year" for the x-axis label for completeness, but the x-axis numbers look like years, and don't need much more explanation.

Now you have a story to consider, to tell and to ponder. Many Americans probably assume that more people in the United States have cell phones than in other countries. What might explain the fairly major difference between cellular subscriptions in Finland and the United States?