

Visualizing Time-series Data

Part of a Series of Tutorials on using Google Sheets

Last updated: October 12, 2017









These materials are based upon work supported by the National Science Foundation under Grant Nos. IIS-1441561, IIS-1441471, & IIS-1441481. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

Getting a Sample Time-series Dataset

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 organize it to create a graph.

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 of people who own cell phones, which would naturally be larger in country with a larger population.

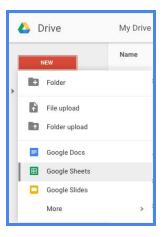


Download the Excel file for this dataset by clicking on the Excel icon under "Download." The file will be saved to your default download location.

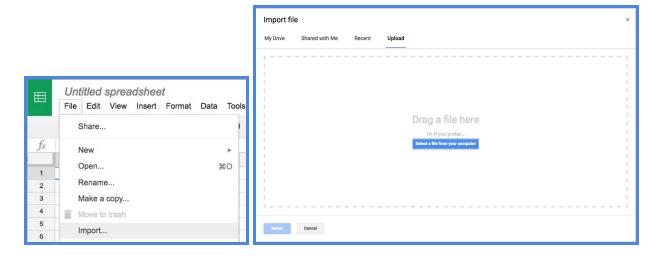


Importing into Google Sheets

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.

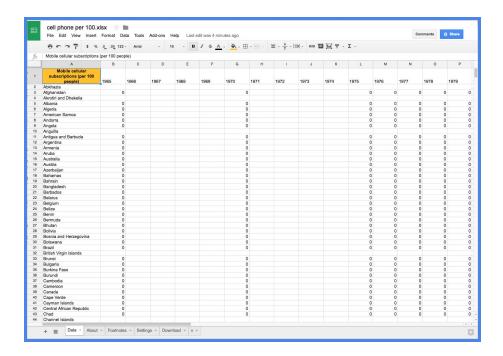


Google Sheets may ask whether you'd like to replace the spreadsheet or create a new one. If it does ask and you are starting with a blank spreadsheet, choose the option to replace the spreadsheet. If your spreadsheet already has data on it, choose to create a new one.



Reading the Data

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.

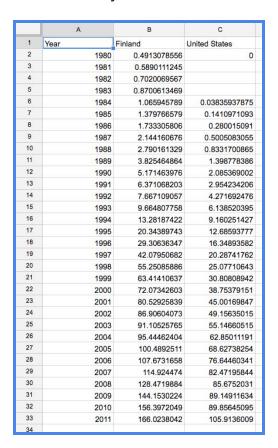
Reorganizing the Data to make a Chart

In order to visualize the data from Finland and the United States, it will be easier if we have each case (or row of data) be a YEAR, rather than a COUNTRY. That way, when creating a line graph, we can have the years on the horizontal x-axis and the cell phones per 100 people on the y-axis. As such, each row will have the information for an individual year, and there will be two columns, one for Finland and one for the United States. Create a new worksheet on the



bottom with this structure. The first row will be for the column labels, so enter "Year" in cell A1, "Finland" in B1, and "United States" in C1.

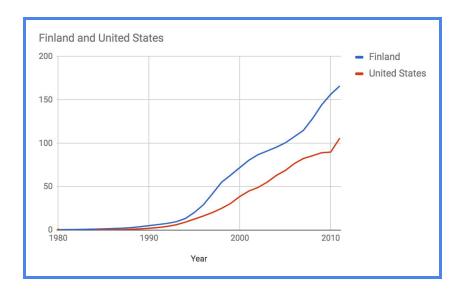
Back in the original sheet, start at Finland 1980 and copy the values all the way up to Finland 2011. Return to your new sheet, right-click (or control-click on a Mac) on cell B2 and select Paste special -> Paste transpose. This function in Google Sheets will make what used to be a row of values into 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 your data visualization tool of choice:



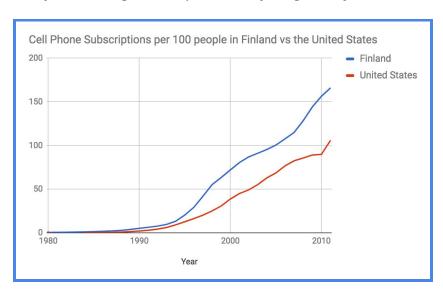
Making the Chart

First, let's test it out in Google Sheets. Select all the data and go to Insert->Chart in the menu options. It will automatically detect the decimal values and start with a line graph. A graph of time-series data is just like a line chart, so this is appropriate.





Edit the title to aptly match the data shown so viewers would know exactly what is shown in the graph. Let's go with "Cell Phone Subscriptions per 100 people in Finland vs the United States." This title may sound long but it captures everything exactly based on the original data.



Making the Chart Using other Tools

You can copy and paste the graph made in Google Sheets into your infographic, or re-make it in the infographic creation tool so that it can better match your design aesthetics. In your tool of choice, find the line graph or line chart option. Go into the "edit data" mode for the line graph. Different tools may have different names for this mode. Delete any filler data that may already be there, and then copy-paste your extracted data into the associated data table. Make sure you understand how the data table works. Some tools may make the first row the "header" row, meaning it uses that first row as the labels for the chart. When that's the case, paste in the numerical data starting in the second row.



Making it look good

If the tool shows every single year in the x-axis and makes it too crowded to see anything, see if you can adjust it in the graph settings. If not, one solution is to only use data for every fifth year or so. 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 not showing some of the interim data, let's create a graph with data for every fifth year.

In your tool, select and delete multiple rows and leave only the rows for 1980, 1985, 1990, 1995, and so forth. Here's a screenshot from Venngage:



Reflecting on the Data

Now that you've completed this walkthrough, here are some key points to take away:

- Starting in 1995, there is an increasing difference between cellular subscriptions in Finland and the United States. What might have caused this?
- Since the value shows number per 100 people, a value over 100 means that the average person has more than one cell phone subscription. This could mean some people have a personal phone and a work phone.
- When were cell phone invented? Are there countries that may have a higher rate of cell phone subscriptions? What could that say about people's lifestyles or the country's economy?

