CHAPTER VII

# Visualizing Data

Now that we have learned how to source, scrap, explore, manage, store and narrate, we can use our data to create visualizations.

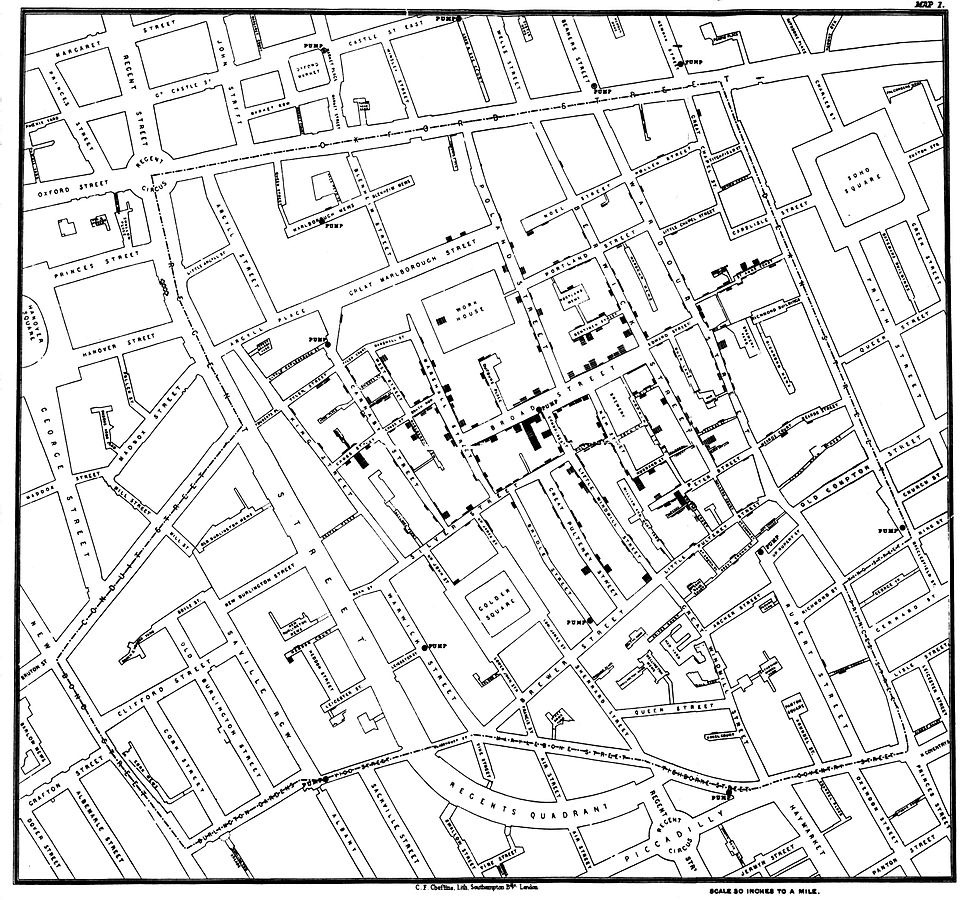
## What is data visualization?

Data visualization is a way of communicating information visually that shows patterns, trends, or correlations. It displays highly complex data and information in a way that is quick and easy to understand. It also communicates information that may have been overlooked or under-lighted in text. Often, it is generated to be able to be understood by the broader public. It is broadly used in digital and data journalism.

Interactive data visualizations, such as Tableau and CartoDB softwares, give the user the capability to filter data in a way that is of most interest to them. This is extremely useful because the audience has more freedom to include or exclude data from very large datasets made available to the general public. As seen in the following famous John Snow visualization, excluding less relevant information can declutter the drawing and make the patterns more clearly visible.

## Origins of Data Visualization

***John Snow is credited with creating the first data visualization.***

In 1854, John Snow created a visualization using a map designed by cartographer Charles Cheffins. Following a cholera outbreak in 1854, Snow and others talked to building residents and Snow mapped deaths in the Soho neighborhood of London. He drew bars inside on top of each building’s location; each bar represented one death. From this map and the information gathered to create this map, he showed Broad Street pump to be the culprit of the outbreak in a way that was easy for people to understand.

*Right: 1854 Broad Street pump, London*

Visualizations can be created two ways: 1) drawings that seek to show an answer to a question with the answer already clearly defined, and 2) drawings that seek to draw associations and test theories and assumptions with no answer yet clearly defined. In this visualization, Snow was using the former method by trying to prove the “answer” to the origins of the deaths that cholera was water-borne.

He included deaths and water pump locations, and excluded other unnecessary information such as bakeries, hospitals, parks or numbers of residents, and horses, etc.

Snow was not the first one to create this map, but he is credited as being the first one to create this map that was *used to tell a story.*

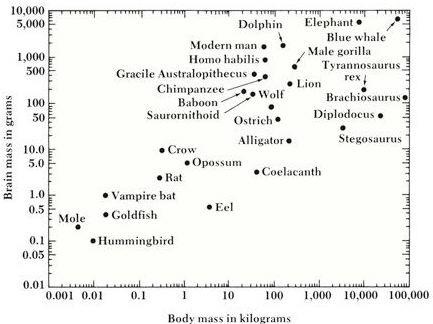
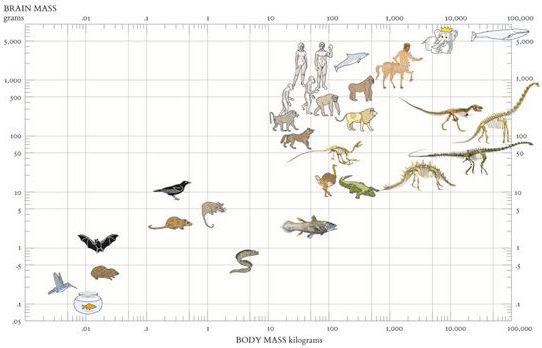
***Charles Minard created one of the most famous visualizations.***

Below is a design of **Napoléon Bonaparte’s famous march** to Moscow in 1812-1813 designed by Charles Minard in 1869. He started with 400,000 troops and returned with 10,000.[[1]](#footnote-1)



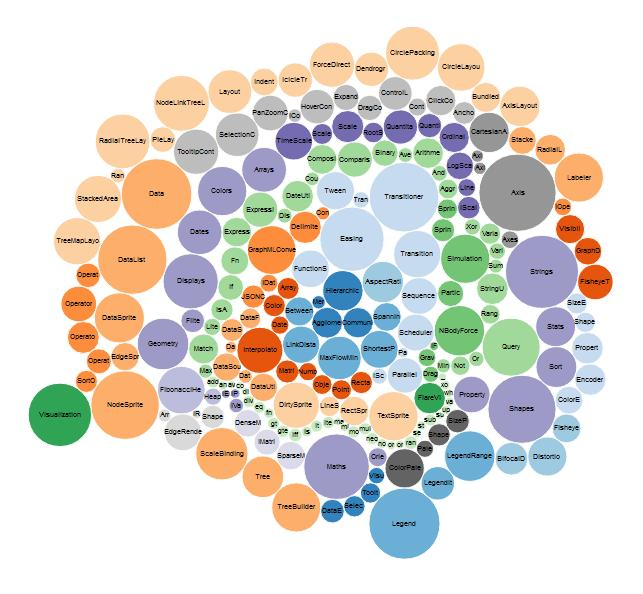
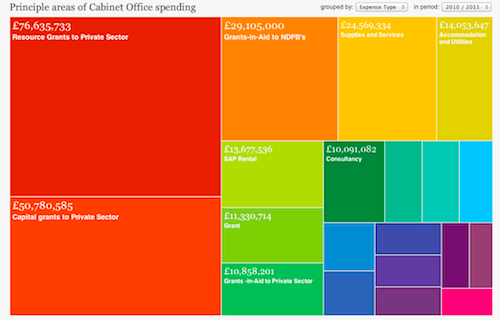
***Edward Tufte is known as a pioneer in data visualization.***

Edward Tufte is an American statistician born 1942. He is Professor Emeritus in Political Science, Statistics, and Computer Science at Yale University. He is known for his contributions to information and visual literacy. He first published *The Visual Display of Quantitative Information* in 1983, and has published four in a series of five books, the latest is *Beautiful Evidence.* He has been called the “da Vinci of Data” and the “Galileo of Graphics.”[[2]](#footnote-2) Below is a small *before* and *after* example of what smart visualization can do, the *before* is of Carl Sagan’s original diagram, and the *after* is the one designed by Tufte.[[3]](#footnote-3)

*Before After*

## Types of Data Visualizations

|  |  |
| --- | --- |
| **Choropleth** | **Dot distribution** |
| https://lh3.googleusercontent.com/oNsRBQZsr-ZL39pCPGnh5TimiS_SXZnNPstj5jxqaCHXaptHZ7yyo9JmOuC20P1DGdWuXvja0kNN4lgUD5vGF0MqM2ng3Z1t6q6qcjVw1fUI6zxDi7J9f2h-XWOl2j0g_E-GZlE | https://lh6.googleusercontent.com/6kBIhMqSxMipdKmYRKy4GiwVvngwlw-wxdDn9FCWeBj_6jgXBSvKNcenSXN-StOYLsxv95k_wDWeBDt0uiLsqRKxw1s6ZcbvDb1g5qRqxiODByyZvoknLsEcw1q_MKH6GmDMCbc |
| **Scatterplot** | **Bar graph** |
| https://lh4.googleusercontent.com/3xvJDH3Ta39NMVNk65PPH5LMdJGSb-yjxbIo4WP1McjdSA6uT0WjP1H5YnQ81lrWKkFPVNtbeOBpGBtLEOeMu3g4jBZ1mFrpHnx-PjZrYn2Zu7cZ_0nVQGcqnYfQ9vZRZToUowU | https://lh5.googleusercontent.com/JayZyTeAvFl0zhgyhPL3D5Pi7_WHvVJrsZ_LsIDOA-tjvY1djpb-lHm9UovQ-pLrG0-D0yXvEr1E-bsIxxJ4F7n5tAZ4Tn5E4u0s1VBS7XJQvy-YyE8PBU1Rk5f7hfbdBiDNuPw |
| **Time series** | **Pie Chart** |
| https://lh6.googleusercontent.com/8piS-7FtwdE9gVFLmwMyt7l6GJvVm_Rdjwwryidfsdvb3hCXQXFqvvcJ2gA_o1v1GKqlePlzIRCd2QDDfQTwOIoEZ7gJ_nkopjoIWJeO3DWDlPwjA5MRv8sWfRI6U4QAK0Gl_ik | https://lh3.googleusercontent.com/Ad6lPH4gDkmPCeLQf-oo8EipH7a2yYKJt75SrOT7MpCGShuREYlySlBtue4JALM6NRyCWiKiIHNS4IkOw3zQdfxwi99gsESK2JEy6SqFadBzwlvqm2k3CXKcXuuRk4IcnJ16wRA |
| **Histogram** | **Bubble chart** |
| https://lh3.googleusercontent.com/SIoSRpPT_gvpLk6M6UROA4i9Lb99mo48MSdhIrw8FrI0Sk6p9um90HcTpKyX9I6r--5zpyNzej0YstDIE84Qnt_7UgGhMClbVqXPNOo_ZQlTXnb1o9LaKzU-j2WD2QaWYunXX4w | https://lh5.googleusercontent.com/ts4faIbRpf9jsWOQKVaQ-uzzsc8IZLpreTVy-xUweq_HrOZM6VYcUTRsc1JtSMZ6jWDFiNbIjIvLXGCX_w7491E-CU7mgsi4y7Efq0Km4A2cipxisDDPIb4nmxnN9abLrzLLqiA |
| **Bubble cloud** | **Tree map** |
|  |  |
|  |  |



The following six pages are designed to help you explore data visualization techniques and create visualizations of your own by using available free or open-source software.

## Excel

**What is Excel[[4]](#footnote-4)**

Excel is a Microsoft spreadsheet application, one of the most powerful and widely-used data processing software. With Excel, you will be able to create tables, format tables, sort and filter table data, and use formulas with tables. This tool enables you to manipulate and visualize data in a simple, handy manner.

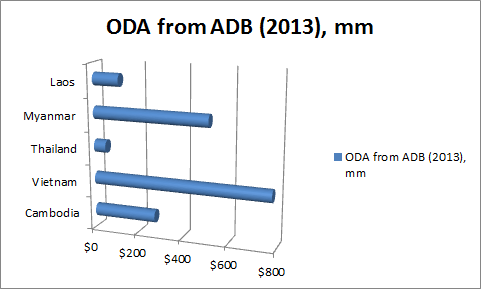
**Who Uses Excel**

Excel is used by students, scholars, businesses, citizens etc. to track the data trends, perform calculations and build data models. It is often installed by default as part of the Microsoft Office Suite on computers. Excel workbooks contain individual worksheets, which you can use to create lists and spreadsheets. More advanced applications of Excel include creating pivot tables (see *Exploring and Managing Data with Excel* section) and generating basic visuals.

**When to Use Excel**

Excel is widely used. People have used Excel for all sorts of purposes, from financial modeling in big companies to serious statistical analysis. In financial industry, Excel’s ability to create new spreadsheets where users can define and apply formulas to extend the model and create forecasts makes it highly valuable. Excel is also frequently used for common information organization and tracking such as a list of staff names, project status reports, sales records, and invoicing. Additionally, Excel is a useful tool for scientific and statistical analysis for larger datasets. Excel’s statistical formulas and graphing can help researchers perform variance analysis, correlation and regressions, chi-square testing, and more.[[5]](#footnote-5)

**Data visualization with Excel**

Excel allows users to visualize the data currently stored in the worksheet into different types of graphics. By combining Excel formulas and visualization, users can even create interactive data visualizations with this tool. While making charts in Excel is fairly easy, most of the time we use Excel visualizations to explore and discover trends in the data. For commonly used, and more complex data visualization tools, please refer to the open-source software that are mentioned below.

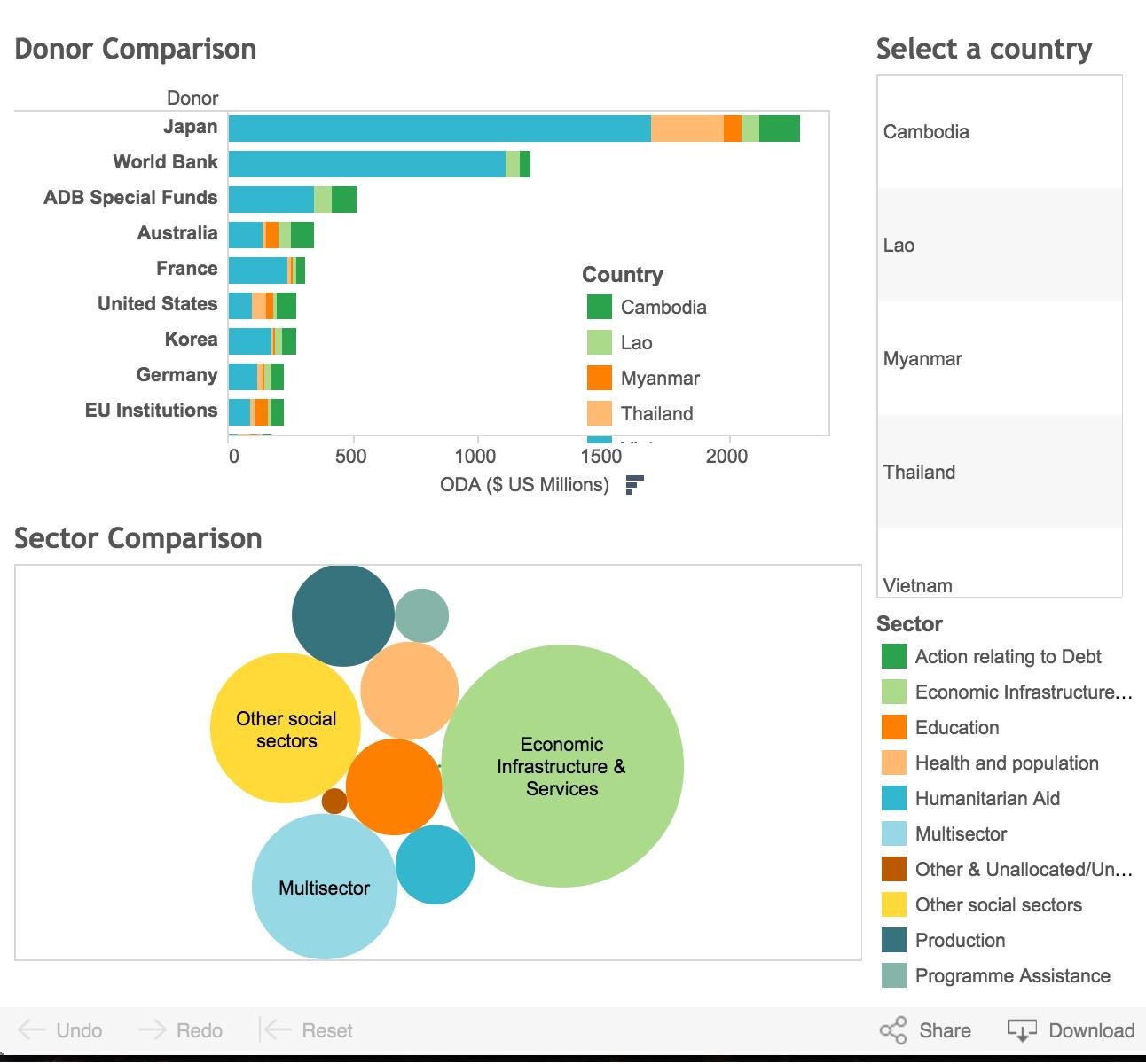
Excel enables you to create different data visualization models including bar charts, pie charts, line charts, scatter plots etc, all in 2-D or 3-D format. Here we provide an example of how to create a bar chart. The steps are as follows: 1) Select the data that you want to plot in the bar chart. 2)On the **Insert** tab, in the **Charts** group, click bar type you would like to create. 3) Click the chart area of the chart. This displays the **Chart Tools**, adding the **Design**, **Layout**, and **Format** tabs. 4)Right-clicking the chart area, you will be able to change the chart type, re-select data and move the chart. Below is an example of bar-chart visualization using Excel.

## Tableau Public

**What is Tableau Public?**

Tableau is an interactive data visualization software. A free version of their program is available through Tableau Public, which is available for organizations to use as an introductory service.[[6]](#footnote-6) With Tableau, you can connect to data, create visualizations such as bar charts, line charts, pie graphs, circle charts, histograms, scatterplots, maps, tree maps, and more, and then combine multiple visualizations into an interactive dashboard. Once your data is on the web through Tableau Public, anyone can download it and create their own visualizations from it.

**Who uses Tableau Public?**

Tableau is great because you do not need any programming knowledge to use it. It is used by writers, bloggers, students, professors, critics, citizens and more. Content created with Tableau Public can be embedded directly into websites, or shared via links.

**When to use Tableau Public**

Tableau is best used with quantitative data in order to visualize relationships in interesting ways. It is great for comparisons and allows users to filter results on certain characteristics. It is useful for when multiple variables can be compared simultaneously since you can combine two or more visualizations into a dashboard for greater interactivity. You can also create “Story Points” with Tableau to guide users through a narrative to help reveal a certain pattern or information you would like to highlight.

**Getting started with Tableau Public**

The very first step is to format your data correctly. Tableau needs raw data. This means the very first row of your Excel sheet has the column headers (these will be the variable names), each subsequent row is one observation, and the data begins in cell A1 (no titles, descriptions, etc). Also make sure to take out rows with totals – this is not a data point and Tableau can easily create totals, averages, and more once the data is uploaded.

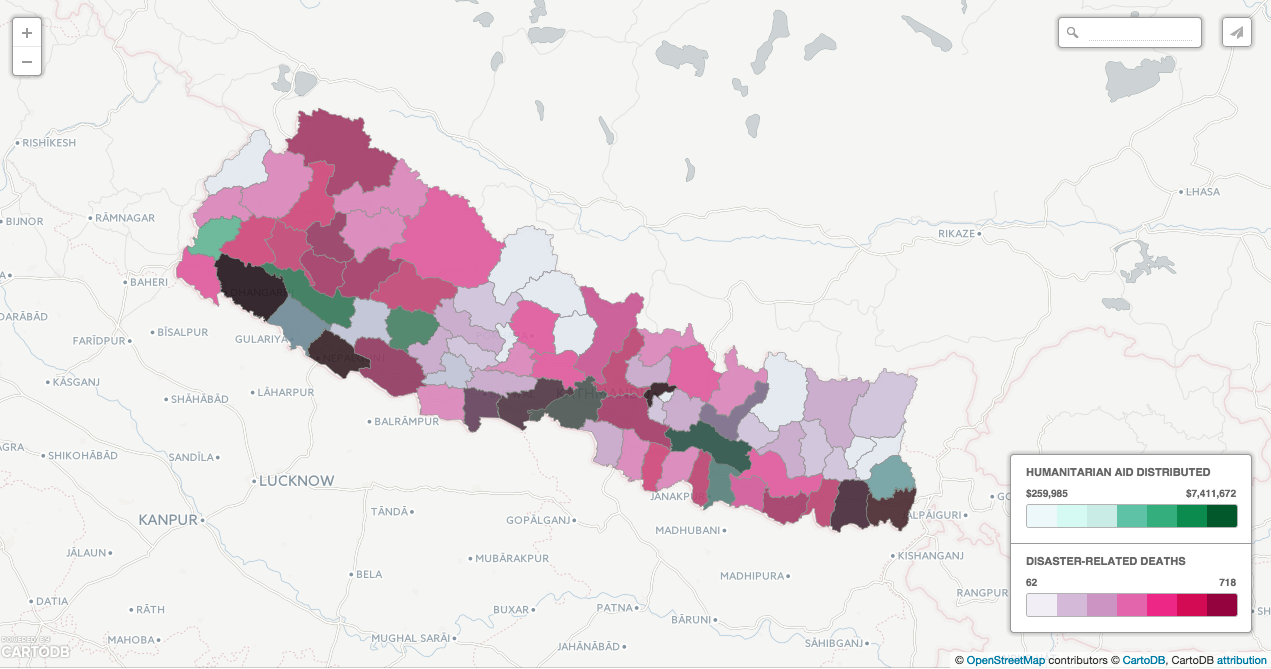
Once the data is formatted, you can upload your data to Tableau. You will notice that Tableau automatically classifies variables as either *dimensions* or *measures*. A *dimension* is a qualitative, categorical, or independent variable. Examples of dimensions are country, province, year, or product title. A *measure* is a quantitative or dependent variable, such as official development assistance received, sales, latitude and longitude, number of projects. Drag and drop dimensions and measures into the workspace to begin creating visualizations. The *Show Me* tab will display options for visualizations that are possible using the selected variables.

## CartoDB

**What is CartoDB?**

CartoDB is interactive mapping software. *Carto* from “cartography” – the study and practice of map-making, and *DB* from “database.” CartoDB allows users to upload data, and use others’ datasets to map trends geographically. With CartoDB, you can connect to data, create visualizations and then combine multiple tables and maps into an interactive dashboard that lets users filter across variables. Once your data is shared with the web, anyone can interact with the data, download the dashboard, and create their own visualizations from it.

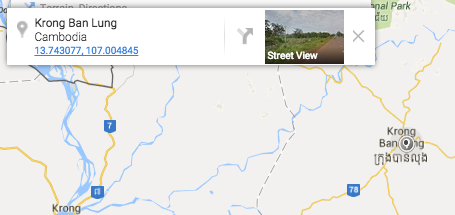
**Who uses CartoDB?**

It is used by journalists, academics, government officials, NGOs, and more. Content created with CartoDB can be embedded directly into websites, or shared via links.

To the right is a map of Disaster-Related Deaths and Humanitarian Aid Received, mapped by district in Nepal in February 2015. Each color represents a variable and the opacity represents the severity.

**How to use CartoDB**

CartoDB uses Excel to upload its data. If you do not have your latitude and longitude points already, you can use CartoDB’s data based on cities, towns, etc. or you can use GoogleMaps; the latter tends to give more precise locations. To do this, type in your point on GoogleMaps, and right click on the drop down menu to “What’s Here?” This will drop a pin into your map. The latitude (first) and longitude (second) will appear in blue under your search. Use those to create columns for your geo points. You can upload data by clicking on “new table.” After you upload your data, there are two views, the Table view and the Map view that allow users to alter their data in real time, and customize their own maps.



**Getting Started**

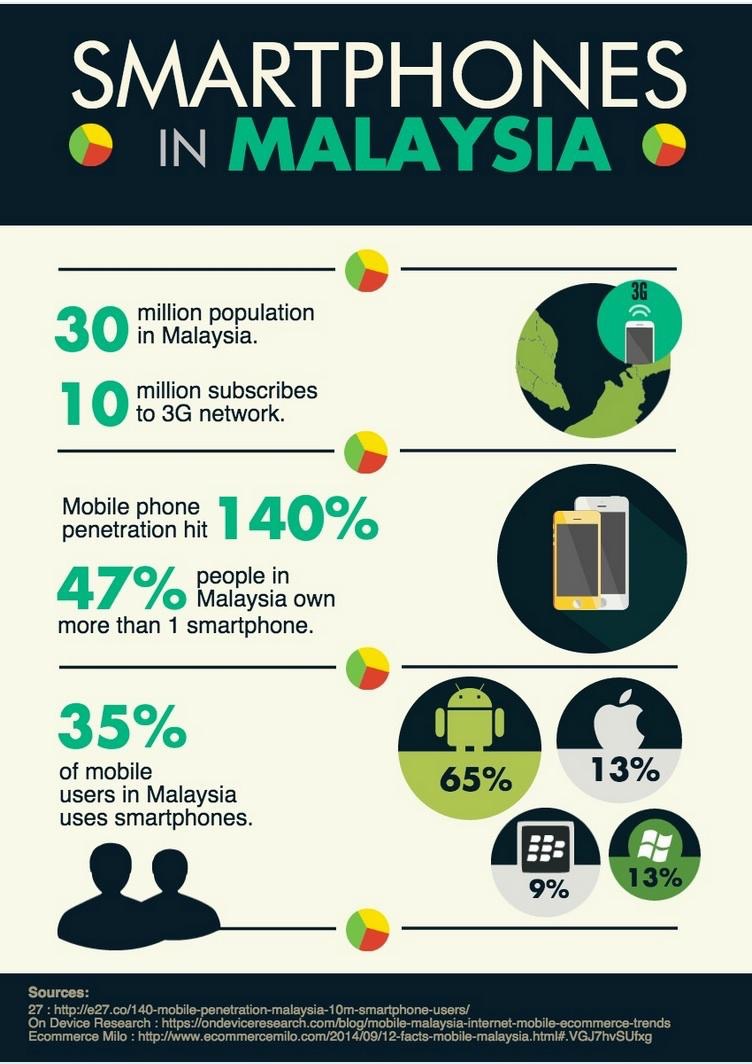
You can sign up for a free CartoDB account [here,](http://cartodb.com/) which allows you to upload 75.0 megabytes of your own data. When you sign up for an account, CartoDB also allows you to use public data to explore and design your own maps using chloropleths, categories, and more. You can publish your designs to the web.

## Piktochart

**What is Piktochart?**

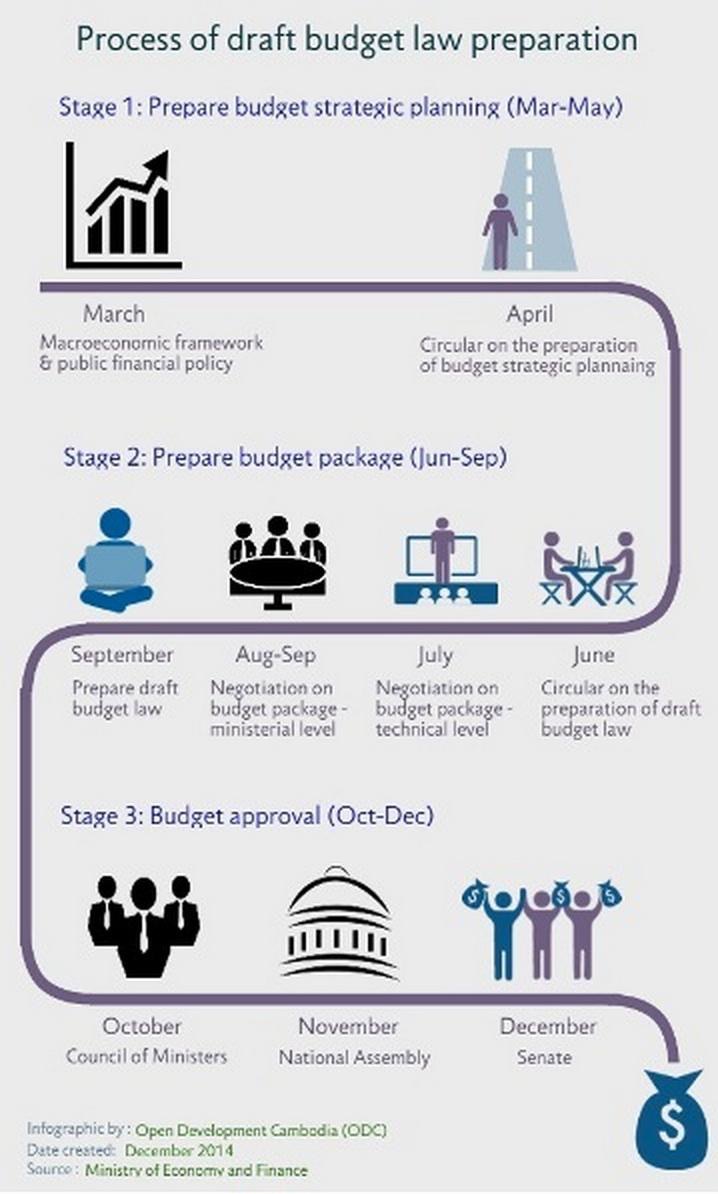
Piktochart is an online tool for creating graphics. With the website, you can easily build infographics, reports, posters, and simple data visualizations that look stylish and modern. The user can easily combine text, graphics, pictures, and charts within a single visual to display information. Users can customize a template or theme, or they can create their own visual from scratch. The website has an easy-to-use drag and drop interface for editing. The visual can be embedded with HTML (allowing for clickable elements), or saved in PDF, PNG, or JPEG formats.

**Who uses Piktochart?**

Piktochart can be used by anyone with any level of graphic design experience. Users can easily produce professional-grade infographics with only a few clicks. Beautiful infographics can be produced without a graphic designer. Forbes Magazine named Piktochart the infographic tool of choice for the “graphically challenged” or for those who need to produce a graphic quickly.

**When to use Piktochart**

Since text, images, graphics, and charts can be combined within a single visual, Piktochart is ideal for producing infographics. Infographics display information, data, or knowledge quickly and clearly. Infographics are great when you have multiple key facts about a single topic; by combining those facts through charts and text in a single visual, the reader is given a complete picture of a complex topic.

**Getting started with Piktochart**

Users can sign up for a free account with Piktochart online at www.magic.piktochart.com[[7]](#footnote-7) Once registered for an account, the easiest place to begin is to choose a pre-made infographic template. After opening the template, explore the features of customizing text, adding pictures, choosing illustrated graphics to incorporate and manipulating the size and color of elements. Once familiar with the interface, you can create beautiful infographics, reports, and presentations from scratch.

Piktochart can produce some simple data visualizations through bar, line, area, dot, pie, stacked venn, doughnut, progress bar, and bubble charts. Different types of charts can be used within a single visual to illustrate information in various ways. Try presenting data with at least one or two types of charts along with text to highlight key statistics or to briefly explain the information presented in the charts. Keep your visuals simple and clean - the goal is to present complex information in a simple way.

## HighCharts Cloud

**What is HighCharts?**

HighCharts cloud is an online visualization software that allows you to create data visualization using the templates available. The templates are not very flexible but they allow users to create, edit and share basic charts. High Software has a more complicated software that must be licensed.

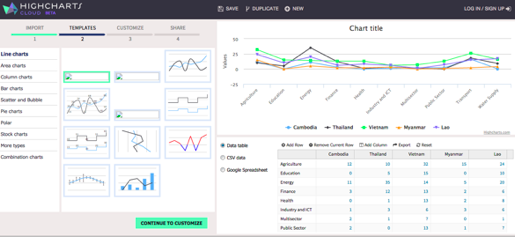
**Who uses HighCharts Cloud?**

HighCharts Cloud is specially designed for “non-techies” which means that you do not have to know how to write code and is easily shareable on blogs, websites and social media. They are served in Amazon.

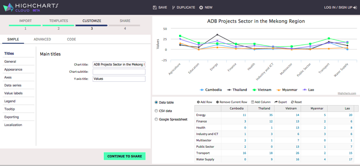
**Getting Started with HighCharts Cloud:**

Visit:<https://cloud.highcharts.com/> and click on “Get Started Now”. Then, drag your data in .csv or .xls format.

Click on “continue to template” to pick the appropriate chart for your data.



Then customize it by adding a title, axes titles, changing the appearance, editing the label, changing the tooltip, etc.:

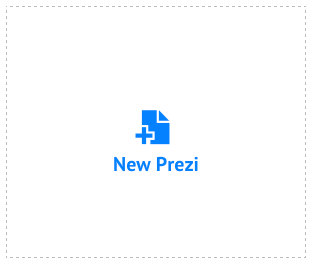


Click on share to download your chart in the format you want or share it on social media, or to get the code to embed it on a website:

## Prezi

**Why and When to use Prezi?**

Prezi is a cloud based presentation software, which means that you would be able to access your presentation from any computer connected to the internet. Prezi is very easy to use and helps you create engaging presentations.



**Creating a Prezi**

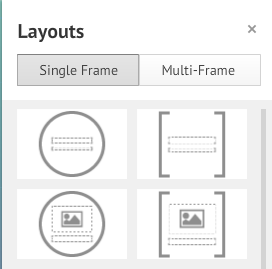
To start using Prezi, you need to create an account first. Then, click New Prezi. You can start from a blank presentation or use a one of the preloaded templates to create your presentation. Prezi works like a canvas, you can add content and navigate through it using the zoom buttons located in the right-hand side of your Prezi.

**Text and images**

To add text, you can click anywhere on your canvas and start typing, on top og your text you will see a formatting bar which will allow you to change the font, the size or the style of the text.

To add images, right click on the insert button in the top menu of prezi, and select what you would like to add to your presentation, it gives you the option of adding images, symbols and shapes, a youtube video or content uploaded from your own computer.

**Layouts**

To organize the design of the presentation, Prezi has the option of inserting a layout: click on the insert button in the top menu and pick from a single-frame or a multiframe and drag to the canvas.

**Borrowing Content**

Prezi allows you to take content from other presentations that are “reusable”. To do this, just click “save a copy” to save it in your own library, from there you can use it to update the content.



1. This visualization, “Napoleon’s March,” was taken from Edward Tufte’s website: <http://www.edwardtufte.com/tufte/minard> [↑](#footnote-ref-1)
2. “The da Vinci of Data,” New York Times (March 30, 1998). http://www.nytimes.com/1998/03/30/business/the-da-vinci-of-data.html and Adam Aston’s “Tufte’s Invisible Yet Ubiquitous Influence,” Bloomberg BusinessWeek (June 10, 2009). http://www.bloomberg.com/bw/stories/2009-06-10/tuftes-invisible-yet-ubiquitous-influencebusinessweek-business-news-stock-market-and-financial-advice [↑](#footnote-ref-2)
3. Bloomberg, “The Vision of Edward Tufte.” <http://www.bloomberg.com/ss/09/06/0608_tufte/3.htm> *.*  [↑](#footnote-ref-3)
4. For additional resources, please visit Microsoft Office Support: <https://support.office.com/> [↑](#footnote-ref-4)
5. “What is Excel Used For”, <http://www.opengatesw.net/ms-excel-tutorials/What-is-Excel-Used-For.htm> [↑](#footnote-ref-5)
6. Tableau Public gives organizations the opportunity to explore its capabilities. If an organization wants to put data online for the public, please contact Tableau Public at info-public@tableausoftware.com to discuss a commercial relationship. [↑](#footnote-ref-6)
7. Nonprofits and social enterprises can register for a PRO account for $39.99 USD per year (go to <http://piktochart.com/pricing/nonprofit/> to request consideration for nonprofit pricing.) [↑](#footnote-ref-7)