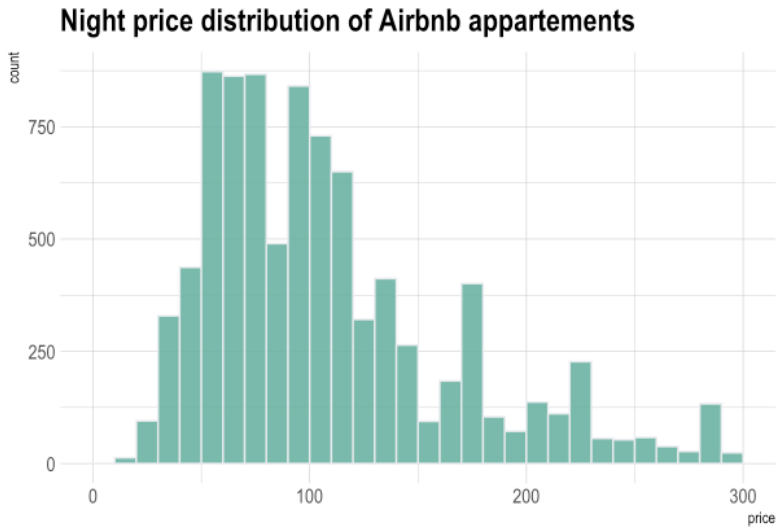
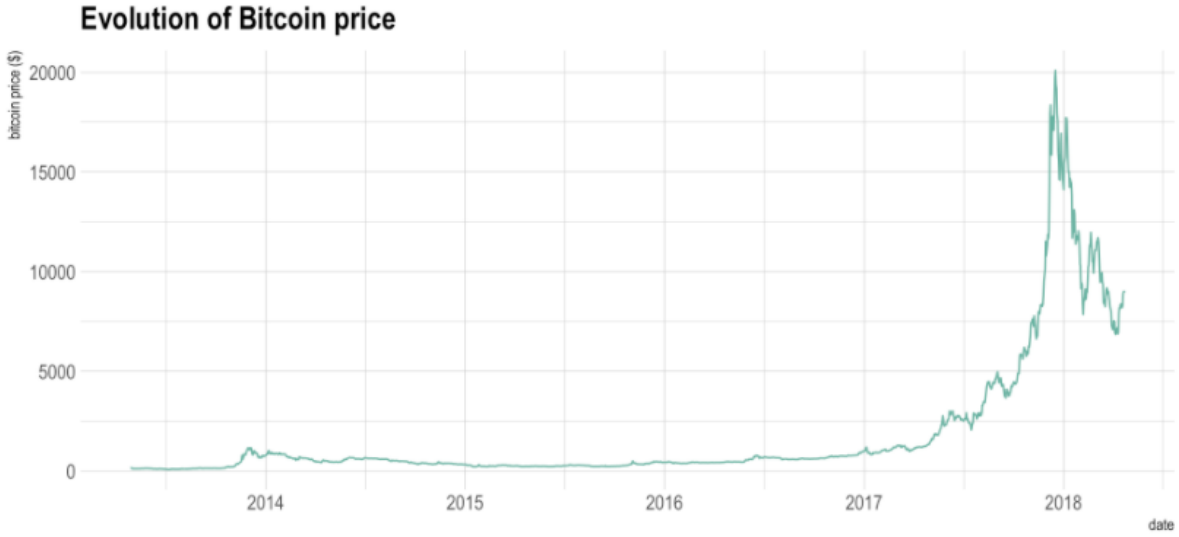
Histogram



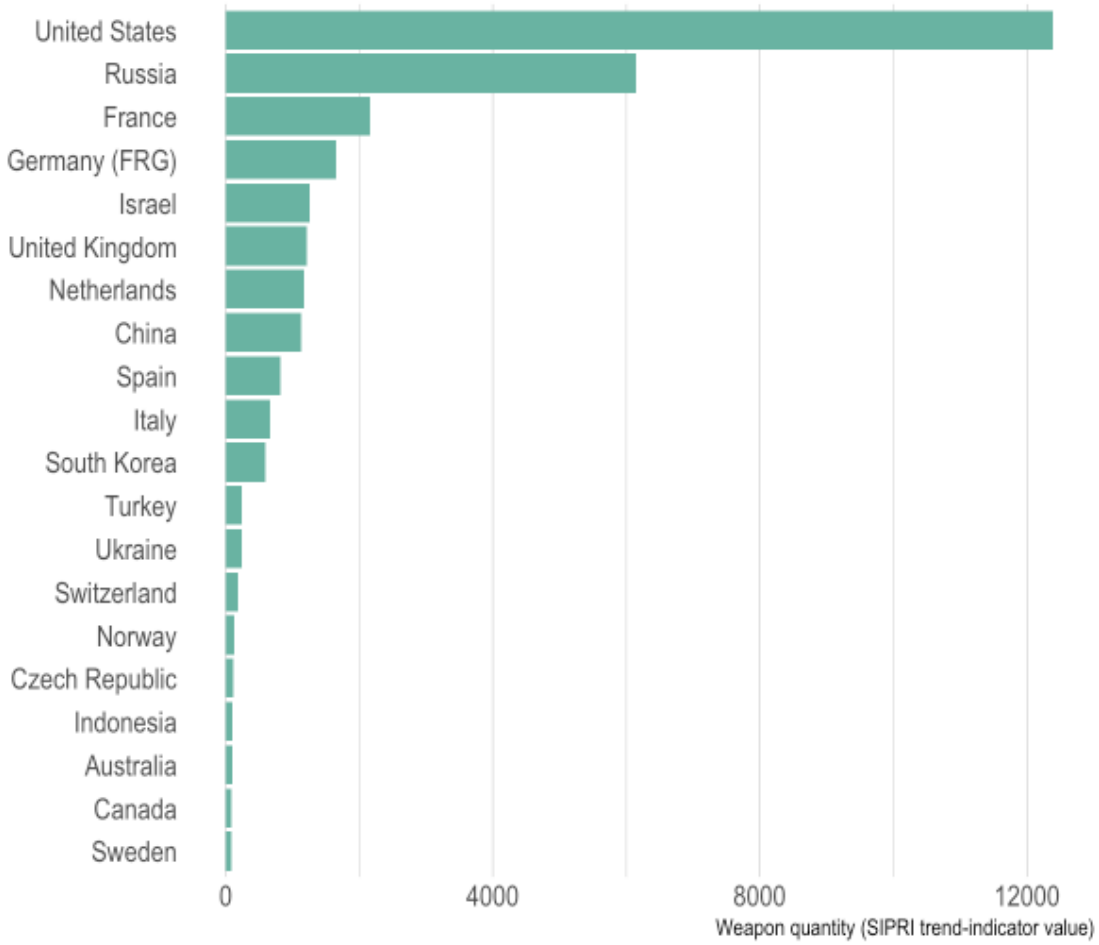
1. A histogram is a graphical representation of the distribution of a numeric variable. The variable is cut into several bins, and the number of observations per bin is represented by the height of the bar. Visually, histograms are similar to bar charts, but they differ in that bar charts compare variables while histograms display the distributions of variables.
2. The central message of a histogram is how a variable is distributed. For example, the included histogram displays how the variable of Airbnb apartment prices is distributed for some sample.
3. This chart is effective in that it is simple. The count variable is on the y-axis, allowing the mode, min, and max price values to be easily determined from the graph. One can even estimate the mean price and label the type of distribution from the graph, which indicates that it is a good visualization. The labels on the x and y axes could be a little larger, but that is also a function of the image size in this document.

Line



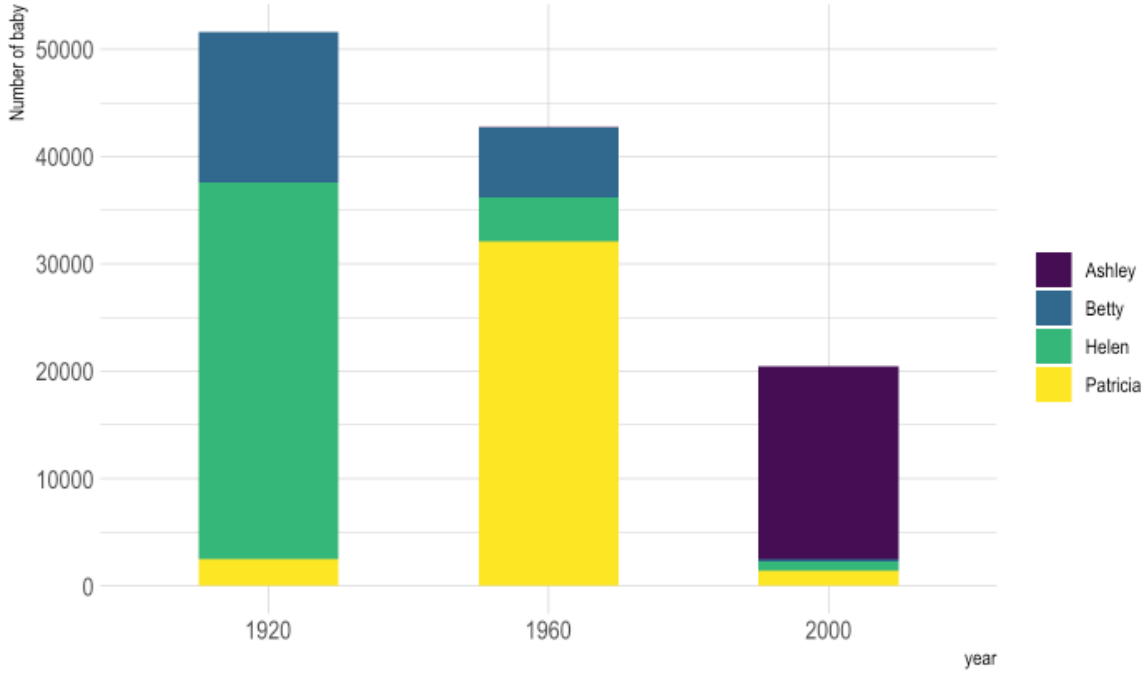
1. A line chart displays the evolution of one or several numeric variables, where data points are connected by straight line segments. It is similar to a scatter plot in that points are plotted, but it differs in that the points are ordered. Often, line graphs are used to visualize trends in data over time, with time on the x-axis.
2. The central message of a line chart is how a numeric variable evolves. For example, the included line chart displays how the price of Bitcoin has evolved from mid-2013 to mid-2018.
3. This chart is also effective in its simplicity. One can easily determine the max price from the graph, as well as guess the minimum price. The trend in the line graph is also very clear: the price increased dramatically until late 2017, when it started to decrease dramatically. The scale is also important in that it reveals just how dramatic the fluctuations have been. A log scale would not be quite as effective in this respect.

Bar



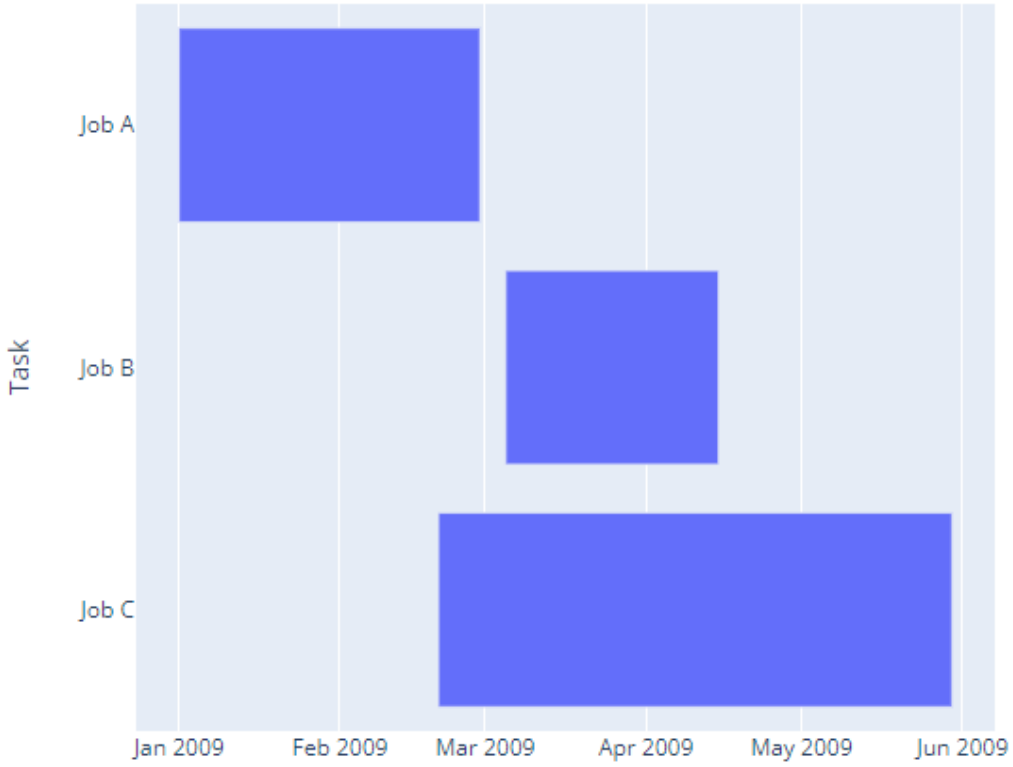
1. A bar plot displays the relationship between a categorical variable and a numeric one, where each entity of the categorical variable is a bar and the size of the bar represents its numeric value. Again, barplots and histograms look similar, but they differ in that histograms display the distribution of a numeric variable and bar plots compare numeric values of categorical variables.
2. The central message of barplots is how categorical values compare to one another. For instance, the included barplot displays how many weapons were exported from various countries. The number of weapons exported by a given country is being compared to that of other countries.
3. Once again, this plot is simple, but it's easy to understand. The bars are organized with the greatest numeric value at the top and the least at the bottom so the wide range of the data can be seen. The bars are all labeled clearly so as to indicate the entities of the categorical data (country). While the precise weapon quantity numbers are not known and would be difficult to decipher from this graph, it is easy to compare the numbers to one another.

Stacked Bar



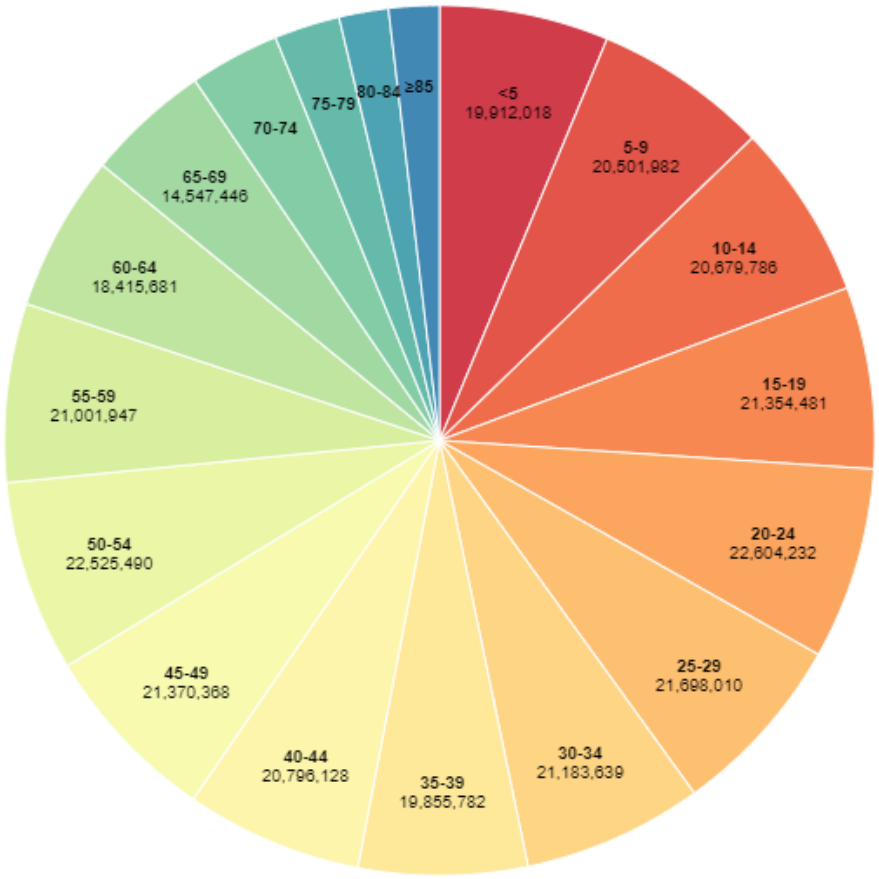
1. A stacked bar plot is a bar plot that displays several levels of grouping, but with the ability to break down and compare parts of a whole. Each bar in the chart represents a whole, and segments in the bar represent different parts or categories of the whole. It is similar to making several pie charts where one bar is like one pie chart.
2. Like barplots, the central message of a stacked barplot is to compare categories of data, but their use is slightly more restricted in that the bars represent a whole. For example, in the stacked barplot above, each bar represents the number of given names per year that were either Ashley, Betty, Helen, or Patricia. Each bar can then be analyzed as having a certain percentage of each name.
3. This plot is effective in that we can infer how name popularity has changed over time. Not only do we see that the bars are decreasing in size, indicating that these names were not as popular over time, but we also see what names grew or declined in popularity. The high-contrast color scheme facilitates this analysis, and the legend helps to indicate which name is represented by each color.

Gantt



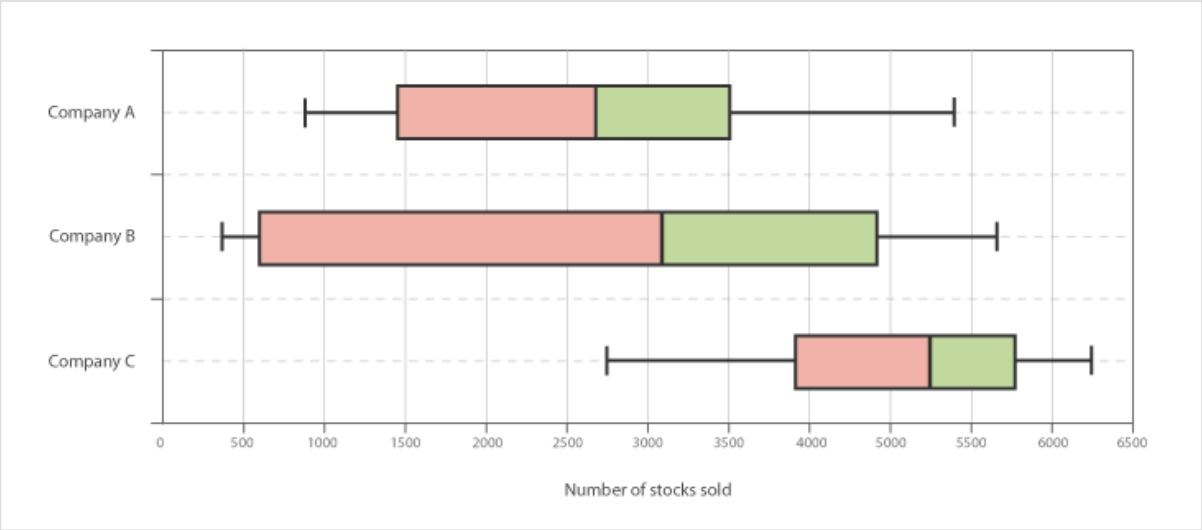
1. A Gantt chart is a bar chart in which a categorical variable is displayed against time. It is often used to map out tasks that need to be completed and can display both parallel tasks and sequential ones.
2. The central message of a Gantt chart is to display activities or events over time as in a project schedule. It does so by showing the start and end times of tasks. For instance, in the included Gantt chart, the user has one task, Job A, that starts in Jan. 2009 and ends in March 2009, then another that starts in March 2009 and ends in April 2009, and so forth.
3. The included chart is effective in that clearly displays the start and end time of labelled tasks. This could become more precise with better task descriptions and a finer scale on the x-axis, and the tasks may be easier to accomplish had the start dates been in order from top to bottom, but as is, this chart is legible and conveys the time-sensitive tasks.

Pie



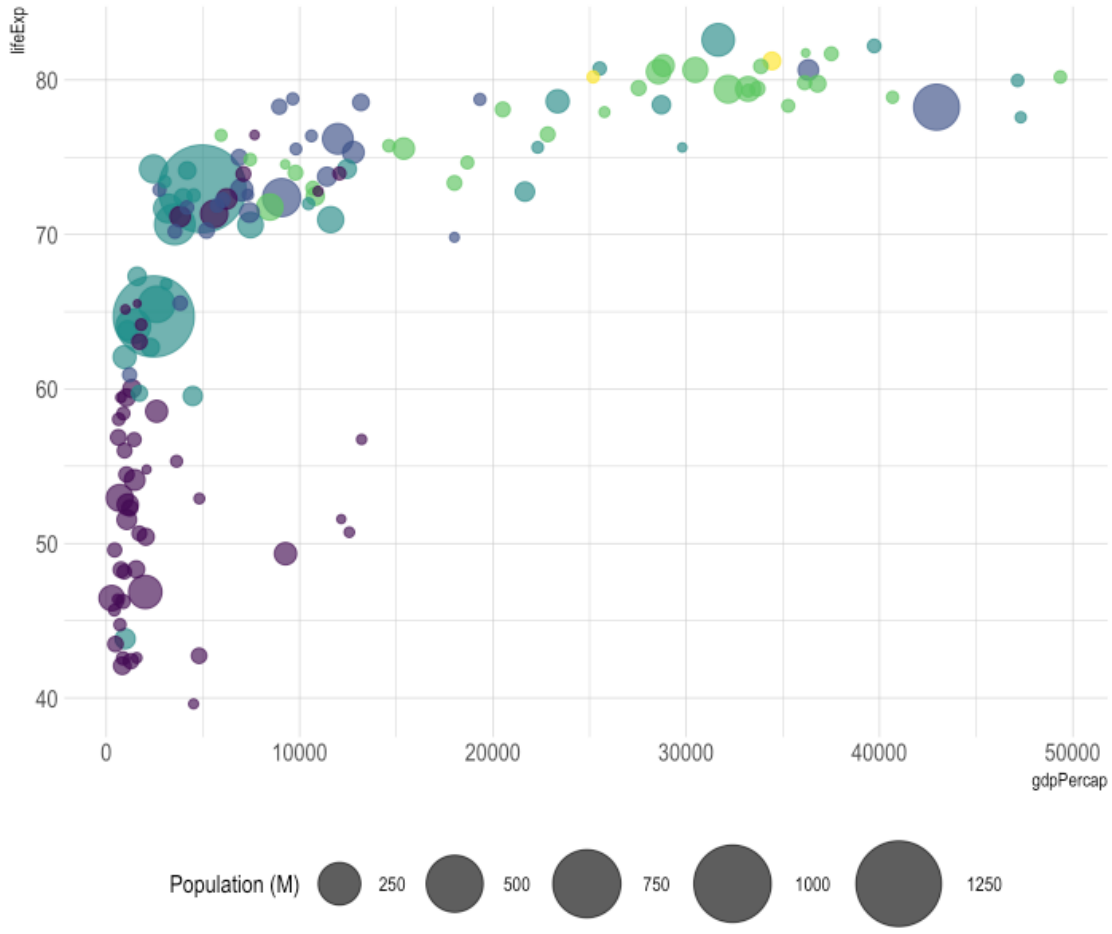
1. A pie chart is a circular graph that is divided into sectors, with each sector representing a portion or percentage of the whole. The sum of the sectors is 100%, similar to one bar in a stacked bar chart.
2. The central message of a pie chart is to compare data as parts of a whole. For instance, the included pie graph represents the estimated population by age in the United States as of 2015. Each age group is represented by a wedge in the pie, and all the age groups together make up the whole. The size of each age group's population is represented by the angle of the wedge.
3. What makes the included pie graph effective is the labels that are included: both the age group and the age group population are included for every wedge. The age groups are in order from youngest to oldest (in the clockwise direction), which is helpful for interpretation. Having a gradient for the pie graph is also useful in that it helps the graph reader to distinguish each section; however, looking at the graph as a whole, it is not easy to determine which age groups are the smallest/largest and what trends might be present. It would be helpful to explode or bold the min/max categories or use a different type of visualization. The graph also lacks a title which is imperative in data visualization.

Box Plot



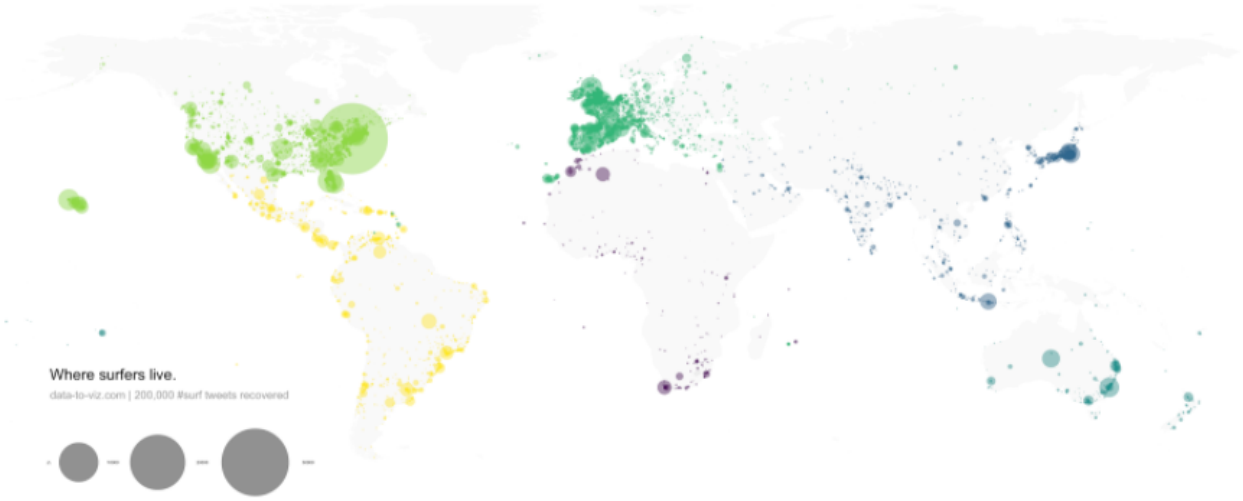
1. A box plot is a graph that gives a summary of one or more numeric variables. It consists of a box with a strikethrough, where the strikethrough delineates the median and the outer edges of the box designate the first and third quartiles. The box also has a line extending from it, where the endpoints represent the maximum and minimum values. Any outliers are depicted using points beyond the line. A box plot is similar to a violin plot, but it differs in that a violin plot displays the distribution of the data in addition to the descriptive statistics found in a box plot.
2. The central message of a box plot is conveying the distribution of a numeric variable by displaying some primary descriptive statistics. For example, in the included box plot, we can see that the median number of stocks sold at Company A is less than that of Company B, and the median number of stocks sold at Company B is less than that of Company C. Similarly, we can compare the maximum and minimum numbers sold by each company.
3. This chart is effective in its scale, for it is fine enough to be able to accurately estimate each statistic demonstrated in the box plots. The color scheme also assists in the readability of the plot - numbers that are below the median are blocked in red, and numbers above the median are blocked in green. However, one downfall to this plot is that we are unable to determine how many data points are represented by each box plot in the graph. Company A may have a drastically different amount of data points than Company B, and that would make for a somewhat deceptive comparison.

Bubble Chart



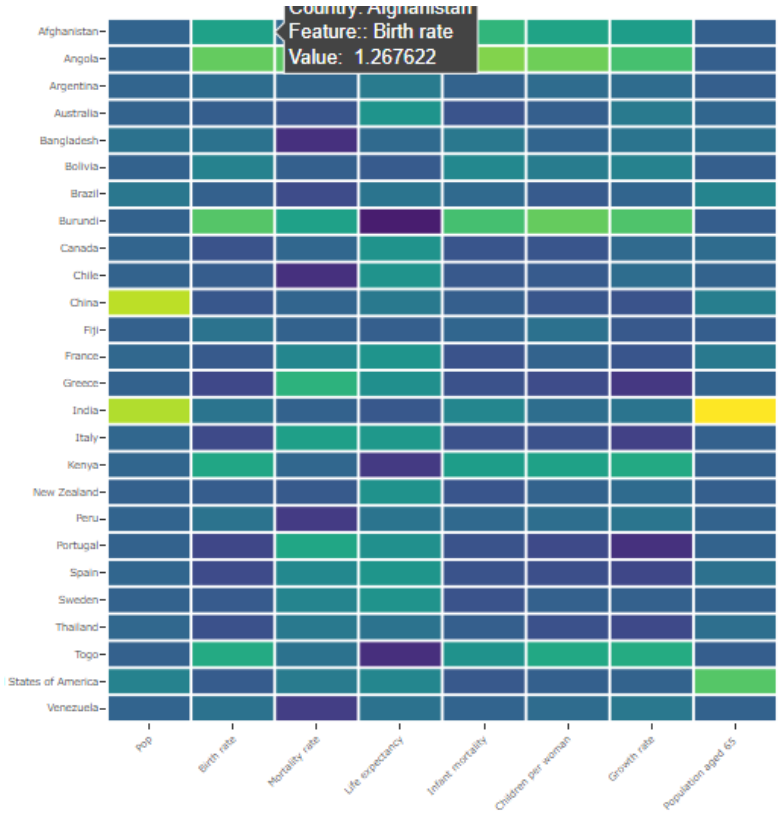
1. A bubble chart is a scatter plot in which a third dimension is added: the value of an additional numeric variable that is represented by the size of the dots. Thus, three numeric variables are needed: one for the x-axis, one for the y-axis, and one for the dot size.
2. The central message of a bubble chart is to communicate the relationship between three numeric variables. For instance, in the included bubble chart, the x-axis represents the gdp per capita, the y-axis represents life expectancy, and the bubble size represents population.
3. This chart is effective in showing a clear relationship between gdp and life expectancy. A good graph won't necessarily have such a relationship, but if the data exhibits any type of relationship, a graph should demonstrate it. This graph also makes use of a color gradient, but it does not provide a legend to indicate what the colors correspond to; this would be a good improvement. However, the included legend does indicate what population size the bubble size corresponds to, which is helpful for interpretation.

Bubble Map



1. A bubble map uses circles of different sizes to represent a numeric value on a territory. It displays one bubble per geographic coordinate or region. Two datasets are required for bubble maps: one containing a list of geographic coordinates (or a list of regions) and another with attributed numeric values.
2. The central message of a bubble map is to compare some numeric variable across a geographic region. For example, in the included bubble map, the size of the bubble is controlled by how the number of tweets with #surf, #kitesurf, or #windsurf. The idea behind this graph is that we can estimate where many surfers reside based on the size of the bubble or how many tweets contain the surf hashtags.
3. The bubble map above is effective in that it displays a range of bubble sizes and includes a legend to indicate what each bubble size signifies. The chart also has a range of colors that appears to depict the different continents; some clarification here would be good. The chart could also use a title to summarize the graph. However, on the whole, the graph reader can easily see that the east coast of the US and some other coasts have a greater amount of something, which turns out to be the number of tweets containing the surf hashtags. The graph reader may then be able to infer that more surfers reside in these regions, which is precisely what this bubble chart is trying to convey.

Heatmap



1. A heatmap is a graphical representation of data where the individual values contained in a matrix are represented as colors, similar to looking at a data table from above. Often, a dataset will need to be normalized before being represented graphically in a heatmap.
2. The central message of a heatmap is to show a general view of numerical data, not to extract specific data points. For example, in the included heatmap, which demonstrates eight general features (population, birth rate, etc.) from 30 countries in 2015, the specific values cannot be determined without interactivity, but certain values stand out in each column. We can use the values that stand out to interpret the graph. India and China, for instance, have a very high population compared to the other countries in the graph.
3. What makes this heatmap extremely effective is its use of interactivity. When the graph reader places the cursor over any sector of the heat map, a text box with the country, the feature, and the value appears. This allows the user to interpret the heatmap with more precision. The color scheme for heatmaps is very important, and this graph has a good one. The extreme values can easily be seen with the dark purple and the bright yellow boxes, which makes the map easy to interpret in the general sense.