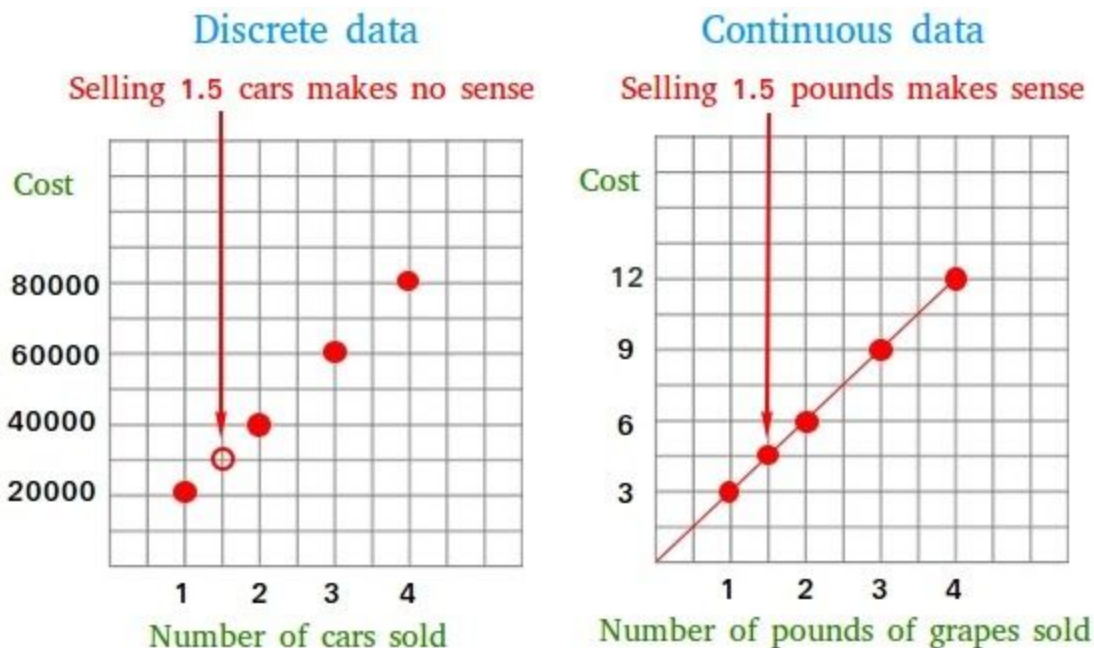


Data Visualization

Data visualization refers to various techniques for making graphical representations of data. The visualization includes graphs, plots, charts, and Venn diagrams. There are two significant distinctions used for data visualization: quantitative visualizations, used to illustrate quantitative (information about quantities/numbers) data, and categorical, used to represent characteristics. Quantitative data comes in two categories: continuous or discrete. Discrete data has fixed points of measurement that are possible.

An example of this is the number of people in a room since you can have 1, 2, 3, etc., people in a room, but there can't be 1.05 people in a room. Continuous data can be anything within a range of data. For instance, data on heights are persistent because the measured size can be 100 cm, or 101, or 101.0001 cm without being fixed at specific values. Categorical data can be ordinal or nominal. Nominal data includes categories that aren't ranked or ordered, such as country of origin. Ordinal data is requested, such as attributes on scales.

Continuous vs Discrete data example



((2008). Difference between discrete and continuous data [Image]. Basic Mathematics.com. Taken from <https://www.basic-mathematics.com/discrete-and-continuous-data.html>)

Ordinal

Center	Status	Response Treatment	Poor	Moderate	Excellent
1	1	Active	3	20	5
1	1	Placebo	11	14	8
1	2	Active	3	14	12
1	2	Placebo	6	13	5
2	1	Active	12	12	0
2	1	Placebo	11	10	0
2	2	Active	3	9	4
2	2	Placebo	6	9	3

Pennsylvania State University. (2018). 1.2 - Discrete Data Types and Examples. Retrieved December 02, 2020, from <https://online.stat.psu.edu/stat504/node/1/>

Nominal

What is your gender? <input checked="" type="radio"/> M - Male <input type="radio"/> F - Female	What is your hair color? <input checked="" type="radio"/> 1 - Brown <input type="radio"/> 2 - Black <input type="radio"/> 3 - Blonde <input type="radio"/> 4 - Gray <input type="radio"/> 5 - Other	Where do you live? <input checked="" type="radio"/> A - North of the equator <input type="radio"/> B - South of the equator <input type="radio"/> C - Neither: In the international space station
--	---	---

((2020). Types of Data & Measurement Scales: Nominal, Ordinal, Interval and Ratio [Image]. My Market Research Methods. Taken from <https://www.mymarketresearchmethods.com/types-of-data-nominal-ordinal-interval-ratio/>)

Importance

Data visualization is useful for organizing complex information in ways that can be understood by a broad audience. Suppose a scientist needs to explain their work's main findings to a group of policymakers to convince them to change an individual policy. In that case, a graphic can help illustrate concepts that would otherwise be harder for non-scientists to understand. Visualizing data can also let people see patterns that would be harder to notice when just given the data itself without the visual to help them contextualize their information. Thus, data visualizations must be accurate and understandable. Otherwise, they will be ineffective or actively misleading. Useful data visualizations organize large sets of data into digestible information segments, are organized logically, allow for comparison between separate areas of data, reveal multiple levels of detail, and integrate with other descriptions of the data, have

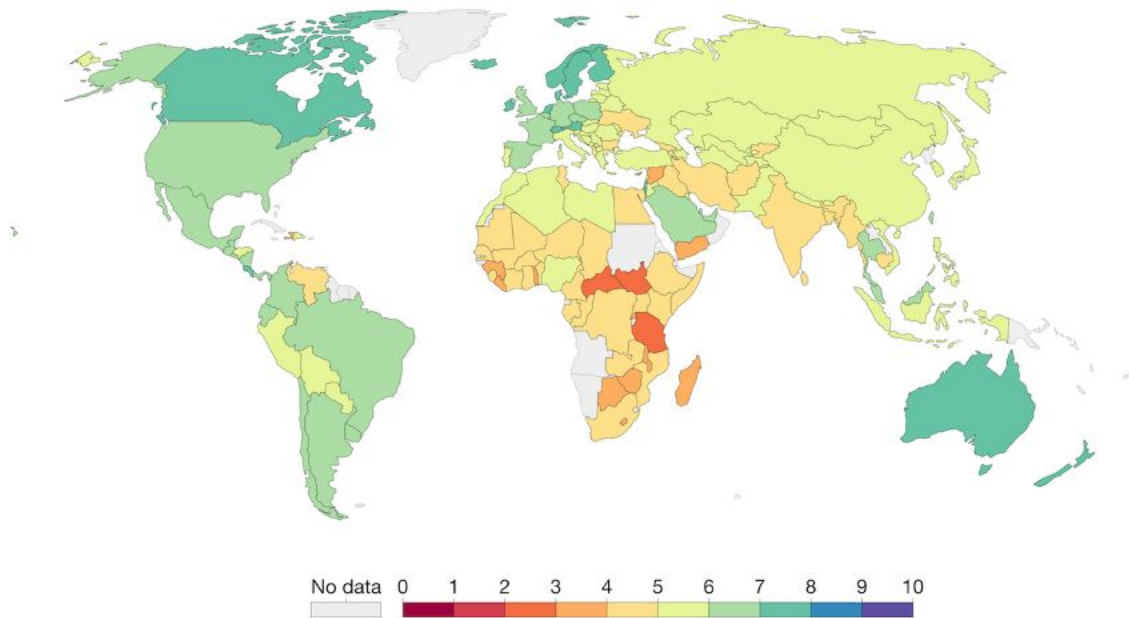
legends and scales. Bad data visualizations are cluttered, lack or have unclear codes or scales, follow conventions on graph usage, and distort the data they are supposed to represent.

Good example

Self-reported Life Satisfaction, 2016



Life satisfaction is self-reported as the answer to the following question: "Please imagine a ladder, with steps numbered from 0 at the bottom to 10 at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. On which step of the ladder would you say you personally feel you stand at this time?"

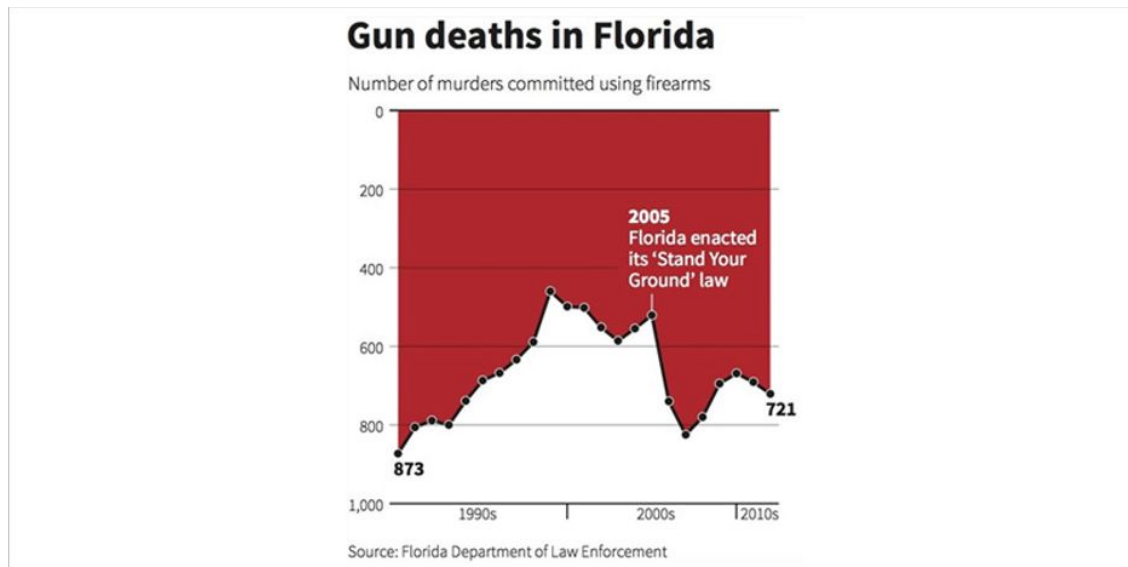


Source: World Happiness Report (2017)

OurWorldInData.org/happiness-and-life-satisfaction/ • CC BY-SA

(Clear scale, gives context with description, logical organization of representing data from countries using a map, allows comparisons between different countries)

Bad example



(Pay attention to the scales on this graph and where they start. The red in this graph represents murders. The chart begins at the top and goes further down when the number of murders increases, an inverse to this type of graph's conventions. This could mislead people into thinking that when the black dotted line goes up, it means more deaths when the opposite is true.)

Tools for students

There are numerous free tools to used to visualize data. These include applications like Microsoft Excel, Numbers, and Spreadsheets that you likely have access to because they come in bundles with things you already need for class. R is free software with some data visualization features. we can find various resources and tutorials through YouTube, online courses, or websites and databases with walkthroughs and shortcuts.