Choosing Chart Types

We can't open up a magazine or watch the news without seeing some kind of chart or survey results. In the brief moment we glance at it or hear the commentary that goes along with it, we take away a message. This is good or this bad or this is going up or this is going down. The graphic may highlight a trending issue or present a hand-picked stat that will shock you, and so on. Most of us don't have the statistics background to know whether to believe what we are seeing in these charts. We don't think immediately about how many people were surveyed or how the data was manipulated to support the story.

I bring this up so that we can begin to appreciate the care that is taken to choose the right chart type or visual in order to communicate or persuade us to the message or story that is meant to be conveyed. In a few weeks you will have completed this course and have gained tools and understanding to improve the signal (the message you are trying to convey with the data) versus the noise (things of lesser import that distract from the signal). This noise may include background images or the use of fancy fonts or colors used for aesthetics rather than to convey meaning. Besides reducing clutter and noise from the story that you want to tell with the data, it is important to recognize the importance of the vehicle—the chart type—that you use to convey the information.

In summary, the data that you present can be categorized as the following: Comparison (best and worst, highs and low), Transition (trends over time), or Composition (part of a whole). The site, Fusion Chart, illustrates examples of these different categories in a useful infographic at http://www.fusioncharts.com/charting-best-practices/selecting-the-right-chart/ (also referenced below).

We are familiar with a handful that we see all the time such as bar charts or pie charts but there may also be some that you are not aware of that can be chosen to better illustrate the points you are making. In this week, I'd like for us to understand when to use what type of chart and then in later weeks we will dive into each type in more detail so that once chosen it is optimized for your audience and used to its fullest capacity.

In 2015, Jami Oetting wrote a great article called "A Simple Guide to Choosing the Right Chart for Your Data," which summarizes different types of charts and when to use them. I've included it in the readings this week so be sure to read it but I wanted to touch on parts of the article here.

There is more to just choosing a visual based on what you are used to seeing or have a template for or what would look the best. I'm not saying we all resort to this but often times when under time crunches we do just use what we have or what we have done before. The article really asks you to think about the data you are working with or visualizing and ask yourself what goals you have and what you are trying to learn or demonstrate from the data.

He came up with some key questions to answer that will help point you to the right chart to use. These questions really help direct you to the right types of charts to use. Each chart has its strengths and weaknesses that we will learn about but answering these simple questions will get you in the right direction and narrow down the list of charts you should be considering. You'll have to read the article to see what charts are suggested for each response to the question.

1. Do you want to compare values?
2. Do you want to show the composition of something?
3. Do you want to understand the distribution of your data?
4. Are you interested in analyzing trends in your data set?
5. Do you want to better understand the relationship between value sets?

Another question you should ask yourself that he does not mention is whether your data is static or does it change over time?

Finally he goes on to summarize these 13 different types of charts: Column, Bar, Line, Dual Axis, Area, Stacked Bar, Pie, Scatter Plot, Bubble, Waterfall, Funnel, Bullet, and Heat Map. You will learn what they are best for and what their weaknesses are at a high level. In future weeks we dig in deeper to learn how to optimize many of these types of charts for usability, storytelling, and fitting your audience. Do you want to

be informative or persuasive or both?

So besides answering those 5 questions that lead you to a chart type set, you must then also make sure you are considering your audience if it isn't you. What is important to your audience? What are their biases? What are their limitations? What do you know about who will be consuming the data and how will that influence your design choices and what to bring attention to. In my work experience, knowing what problems my audience is trying to solve and what decisions they need to make to perform their job are critical to know when designing, for example, a dashboard where multiple pieces of data are presented in webpages.

Color Use in Data Visualizations

Now having touched on how to choose the right chart type, let's briefly discuss a question that comes up all the time. That question is, should I use color and how should I use color in data visualizations.

First of all, it is important to understand the difference between hue and saturation. For example, you choose the hue Red, adjusting its saturation lower would make it pinkish. For scatterplots where you want to call out certain dots as all being of the same family or category, it would be good to use different hues to identify to which family it belongs. Too much color becomes distracting, so as a general rule it should be use to highlight key data in your visualization. Anything that is less important should fade to the background like axis or labels.

We can also leverage conventions and reserve certain colors. For example green, in general, means a positive change and red means a negative change as we commonly see when viewing stock market data. Therefore they should be reserved and not used arbitrarily for a variable’s line color. Saturation is used to show patterns while changes in hue show groups of attributes. When different colors are used to call out different chart elements or represent a change in the data, make sure to use a lighter or darker shade of that second color. Our eyes are better at distinguishing lightness and darkness than color. Using enough contrast between values will make your visualizations also more accessible to all users.

Color should enhance the comprehension of the data and not detract from it. When showing comparisons it can be helpful to show them in opposing colors (across the color circle from each other). For example you could show a blue bar chart next to an orange bar chart to emphasize the comparison. However, make sure to use different values for the colors in terms of lightness. Be consistent and ensure you represent the same variable with the same color throughout the chart. Don't use multiple colors to represent the same type of data. It is helpful to use shades of the same color or colors on the same side of the color wheel. This will help keep keep a bar chart uncluttered and noisy with color so your audience can focus on the underlying data.

With regards to saturation, more saturation is usually used for higher values and lower saturation for smaller values.

What about backgrounds? Backgrounds should not detract from the data, so a white background is standard. Also avoid gradients in your background. Explore whether gridlines can be removed to let the data speak and not have the background command attention. If color is used for background, then use cooler colors for the background and warmer colors for the data. We must be sure to make our visualizations accessible for colorblind individuals. I will not get into that here but there are many sources for this available on the web in terms of safe colors.

With regards to fonts, black provides the highest contrast, therefore using black type on a white or light background is ideal. It is harder in general for all users to read a font that is colored. Also when you use a darker shade to highlight a particular piece of data you can also label that area directly with the use of a bold font. It is useful for accessibility reasons to provide redundant ways to communicate the call out so that anything that was distinguished with color that distinction is also clear without the color. Keep this in mind when designing legends. Also when using color in a scale it should always go from lightest to darkest. It should never alternate within the scale between dark and light strips. An example of this is used in heat maps showing how many people click where on a website.

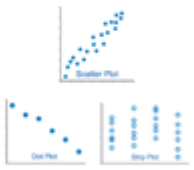
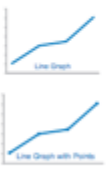
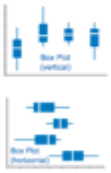
Information Dashboards

Finally, I want to briefly discuss how to display multiple data sets in close proximity or in a dashboard. One of the most common tasks a UX designer may face in their career is to design an information dashboard to help a certain audience make decisions. Here I want to provide some consideration and tips. We understand from eye tracking studies and reading gravity that users scan from the top left and move across and down. You may have seen the famous F pattern (https://www.nngroup.com/articles/f-shaped-pattern-reading-web-content/). We should design the dashboard so that the most important visualization is in the top left corner. We should use the layout to present information hierarchically, placing higher level visualization to the left and more detailed visualization to the right and bottom. Be sure to leverage the Gestalt principles you learned previously around proximity to ensure that elements that relate to a particular graph or chart are clearly associated with that correct visualization.

Guides and Recommendations for Choosing the Right Chart

Here is a great Chart Selection Tool below that is from Stephen Few's *Show Me The Numbers*, 2nd Edition https://www.perceptualedge.com/images/Effective_Chart_Design.pdf

Graph Selection Matrix

Featured Relationships	Value-Encoding Objects: Points	Lines	Bars	Boxes
				
Time Series Values display how something changed through time (yearly, monthly, etc.)	**Yes** (as a *dot plot*, when you don't have a value for every interval of time)	**Yes** (to feature overall trends and patterns and to support their comparisons)	**Yes** (vertical bars only, to feature individual values and to support their comparisons)	**Yes** (vertical boxes only, to display how a distribution changes through time)
Ranking Values are ordered by size (descending or ascending)	**Yes** (as a *dot plot*, especially when the quantitative scale does not begin at zero)	No	**Yes**	**Yes** (to display a ranked set of distributions)
Part-to-Whole Values represent parts (proportions) of a whole (for example, regional portions of total sales)	No	No	**Yes**	No
Deviation The difference between two sets of values (for example, the variance between actual and budgeted expenses)	**Yes** (as a *dot plot*, especially when the quantitative scale does not begin at zero)	**Yes** (when also featuring a time series)	**Yes**	No
Distribution Counts of values per interval from lowest to highest (for example, counts of people by age intervals of 10 years each)	**Yes** (as a *strip plot*, to feature individual values)	**Yes** (as a *frequency polygon*, to feature the overall shape of the distribution)	**Yes**	**Yes** (when comparing multiple distributions)
Correlation Comparison of two paired sets of values (for example, the heights and weights of several people) to determine if there is a relationship between them	**Yes** (as a *scatter plot*)	No	**Yes** (as a *table lens*, especially when your audience is not familiar with *scatter plots*)	No
Geospatial Values are displayed on a map to show their location	**Yes** (as bubbles of various sizes on a map)	**Yes** (to display routes on a map)	No	No
Nominal Comparison A simple comparison of values for a set of unordered items (for example, products, or regions)	**Yes** (as a *dot plot*, especially when the quantitative scale does not begin at zero)	No	**Yes**	No

Here are few handy tools and tips to help you pick the right type of chart for your data:

Juice Labs Interactive ChartChooser

"A Simple Guide to Choosing the Right Chart for Your Data" by Jami Oetting:
http://blog.hubspot.com/marketing/data-visualization-choosing-chart#sm.000yyabqg12tsex8rk51by0ly3cbd

"The Wall Street Journal Guide to Information Graphics: The do's and don'ts of presenting data, facts, and figures" by Dona M. Wong:
http://www.fusioncharts.com/charting-best-practices/selecting-the-right-chart/
