In Chapter 3 of Visualize This, you read about different visualization tools that are available – based on what you read and what you know to be available, what are some of the pros and cons of the different options? Which tools do you think you will start to explore as you start getting deeper in this field?

For the past two weeks, I have noticed a lot of pros and cons with the different tools we have been using. It’s important to get a grasp on a lot of the tools. This is due to the difference in the tools. I found one of the software tools called Many Eyes interesting. I like the example on page 27 where it shows the US Constitution. I think there are a lot of neat graphs that can be built with this software. There is a down fall. The tools are java based. Java can be a big learning curve for most people. I think it’s important to learn multiple tools. I don’t believe there is one software that can do it all.

Programming language mentioned in Visualize this.

So far in this program we have worked with two different programming languages: Python and R. However, there are a lot of different programming languages that can be useful for data visualization. Imagine if we use a language pair like html, css, and JavaScript. We could literally run our graphs on the web browsers. I have never heard of action script, but it’s worth looking into. Most of the job postings that I have seen always mention that they want the applicant to know Java. I used an old reporting system called Birt where I used it to build graphs. Java wouldn’t be a bad language to also look at in your free time.

What is the difference between discrete and continuous? Give examples of each and describe which visualization is best for each scenario.

Discrete is numerical data type that is fixed and whole numbers. Think of it as a set number like how many pushups a person does in a day. It’s always a whole number like 15. Continuous is usually not fixed and includes measures that are not whole numbers like time interval or a baby’s weight 12.345 lbs. This is important because it tells us how to plot our graphs. For example, if you have a figure of 3.4446. It will plot different then 3. The visualization for discrete is like a single dot on a line graph. A continuous is more of a connected line. A great example of this can be found in the article I have provided.

https://www.g2.com/articles/discrete-vs-continuous-data

Why is it important to always start your value axis at 0 in a bar chart? What are the issues or implications that could happen if you don’t?

At first sight I noticed a problem with perception. It was almost like it was cutting out data by showing the graph smaller than it was. If our job is to give a clear indication of the data, this could be a misrepresentation of the whole picture. According to Manjob, the author of the article I shared, “when people look at the bar charts, they don’t see the legends right away. They see the relative differences between the height of the bars”. That’s exactly my first impression. Right away I noticed the size difference. It was like seeing two different graphs. This is why it’s important to start with zero.

**System directory in an operating system.**

I recommend everyone taking a moment to understand how the directories in a operating system work. I can’t tell you how many times this has saved me. For example, when you download a package, where does it go? If there is an error with it, how do you know where to look? The problem is that sometimes issues can’t be resolved because we don’t know where to look. When I updated my R and R studios it made changes all the way to the root directory. From there I had to follow the trail. This isn’t new though. In software engineering, it’s crucial to know where everything is. Now, everyone’s operating system might be different, but the core idea remains the same. Look at this article. There might be some good information.

<https://www.geeksforgeeks.org/structures-of-directory-in-operating-system/>

Git bash

Git bash is one of my favorite tools. One of the things that make it so special is that it’s easily transferable to different technology. Instead of having to learn multiple IDE, I can easily make the changes in my local and push them into github. This has saved me a lot of time in different technologies. I have tried uploading all my projects in school to github. I have even started uploading the data visualization. Some graphs won’t populate but others will. I think this will help us when we are looking for a job. How many of you are keeping records for future job prospects, and are you using github?

<https://www.geeksforgeeks.org/working-on-git-bash/>

Data Visualization And Color Blindness

I am starting to be realize that I might be a little color blind. For example, I have an issue with greens and blues. My wife says our house is green, but I see it as a light blue. I wonder how many people in the audience can’t relate with the graphs because of the color choice. I am starting to understand that color is important when dealing with graphs. This article goes into detail on how some people might perceive color. I think I am going to follow the color-blind palette that is referenced in it. I would take the time to look over this. This is a pretty good topic.