**Part One: Histories**

**1. Peter Naur is famously quoted as saying data science *"deals with the data, while the actual relation of data to what they represent should occur in other fields."* Why do you think he'd choose to frame data science this way? What might be problematic in this statement?**

- I think Naur might have framed data science in this way because at the time, being a relatively new field of study, data scientists may have identified more as statisticians or mathematicians, but not as scientists who could dig into their findings to impact many fields of study.
- Thinking of data science as merely statistics or mathematics can be problematic in many ways. Data science is much more than math and numbers, it is being able to identify a bigger picture within any given data set, and more importantly, it is being able to tell a story, or stories, about what the data means. If one cannot help others understand the story of a data set they have analyzed, then the data is really not helpful at all. This is why it is crucial to not just think of data science as a singular job, but as a full circle of getting data, analyzing the data, and relaying any findings in understandable ways.

**2. In 2002, data science began to gain momentum as its own dedicated subfield. Compare and contrast the definitions of data science at that time, exemplified by the National Science Foundation, Data Science Journal, and Journal of Data Science, to those from Tukey & Naur in the 1970s.**

- In the early 2000's people began to think of data science as a science that crossed over into practically any field. It no longer had the static definition given by Turkey and Naur as being purely a field of mathematics and statistics, but was thought of rather as more of an "umbrella" like discipline.

**3. Data continues to grow at an exponential rate today. List at least three technological factors that contribute to this growth and what role they play. List three major sources of data that contribute to this growth and at least one way they're being used.**

- Pretty much everything we use in our lives today is being turned into data. Some technological factors that add to the exponential growth of data in our world today include faster compute times, cheaper storage, and the internet. Additionally to technological advancements, a few major sources that add to the exponential growth of data include big companies like Amazon and Google, social media platforms like Instagram, Snapchat, Twitter, and Facebook, and the ability to do pretty much anything we want digitally.

**Part Two: Terminal Crash Course:**

**4. How would you move into each of the following directories from the shell? Why do you think it is important to have shortcuts for each of these directories for navigating the file structure?**

A. Root

- c://

B. Home

- cd~

C. Parent

- cd..

It's important to have shortcuts for a few reasons.  Obviously it is much faster to use shortcuts and everyone likes saving time!  It is also a nice way to stay on top of where you are working in your computer.  It is easy to get lost, but if you know a few shortcuts, you can get back on track.


**5. Briefly describe what the following set of commands would achieve. What process would happen and what would be printed to the command line?**


cd ~

- Change directory to home directory

mkdir ./problem_set_1

- Make a new directory named problem_set_1

cd ..

- Change the current directory to the parent directory

pwd

- Tells you what directory you are currently working in

If you ran through this, you would be working in the directory named problem_set_1.