Part I.

After conducting further research, I am compelled to dive into the field of big data. Information surrounds us and as time passes more information in being than can be analyzed. I can look out to the world and naturally see the data that can be captured. Big data connects to IoT and Cybersecurity. IoT can collect and produce a dataset that can be analyzed. How we interact to the world and our tools can now be visualized and interpreted. Big data is also connected to cybersecurity because this information is private and should be secured. In cyberspace, data and be manipulated by opportunist and used for personal gain. Big data is significant because it will provide us with more actionable information that will help us make better decisions and prevent mistakes.

I would describe big data to someone who is unfamiliar as “an extremely large set of information, collected to be interpreted to help make people make better decisions”. An example I would give is collecting weather information to predict when and if the next natural disaster will occur, so companies and disaster relief teams can gather supplies ahead of time and save lives. For example, Walmart has a weather center so they can track any issues with shipments and calculate possible delays. They used their system to deliver water to those who needed help after Hurricane Katrina. Using data, they were able to locate the closest Walmart locations and suppliers to deliver a certain amount of water to specific locations.

Future development for big data is data privacy and governance. Since data can be used to manipulate others, future development will be concentrated on creating standards to protect people from being taken advantage of. Development of best practices with other people’s information will adopt practices used in healthcare, HIPPA. In the previous paragraph I discussed an application of big data and how it can be used in society. Another application is studying health patterns. Is there an increase in eye problems as technology is being used more frequently and for longer durations? Using patterns in history we can determine is this is true. Another application is tracking traffic patterns on busy highways. Your GPS app already does this by telling you when there will be heavy traffic based on the data they collect and attempts to re-route you to the best path.

Some potential career opportunities in big data is titles like data analyst, data scientist, statistician, actuary, mathematician, and database administrator. Someone in big data would need to have business acumen. The ability to understand business operations and how data flows through the information system. They data professional would need to know what questions to ask and what information will be needed to get to the answer. Further, the data professional would need to know how to interpret data and communicate that data in a digestible way to managers and chief officers to help them make the correct decision.

Part II.

This course has led me to think about not only computer science but major trends to come in the future in several fields. I have been looking into which field would bring me the most satisfaction. I stumbled upon data science as “one of the sexiest jobs” and didn’t quite understand how data science can be sexy but I knew that many businesses thrive on collecting, trading, analyzing, and acting on information to make the best decisions. I have learned that information is an extremely valuable resource for businesses to get ahead of their competition and it pays to know how to extract and create value with this information. This course helped in identifying my choice by bringing it to my attention that big data is an emerging trend and not a fad. Out of any of the emerging trends I have the most experience with databases from courses in SQL, data science involved cleaning data, ETL into a database, and run SQL queries to pull the right data.

In the future I plan on developing a sense of how data products work and how databases are used to collect and store data. Further I want to learn how to scrape data from social media and derive insights to help people make more informed decisions. Lastly, the skills I want to develop is to learn how quants trade the stock market which I find interesting.

Data structures, I believe, are the foundation for data science and big data. I believe that the data structure is the “buckets” for which the data is assigned. The data is assigned from an algorithm that determines the condition for which “bucket” the data gets scraped. For example, if I were parsing strings from Twitter and wanted to parse the tweets for a specific group of keywords like “show” and check if the tweet contains this keyword, I can place that tweet in the “bucket”. At least this is what I think happens. Data structures and algorithms are important for data visualization as well. I have used pyplot and matplotlib to create visualizations in python. The data structures make it efficient to organize and store data for easy access and modification. Big data handles unstructured data which is where an algorithm can be used to analyze pictures, sound, and large text files.