# CYSE 368 Final Term Paper

Dylan Anderson Kaitlyn McCoy Brooks Crossing Innovation Lab Spring 2024 4/15/2024

# **Table of Contents**

Introduction	3
The Work Environment	4
The Starting Point	4
Finding Program Alternatives	5
Further Educating Myself	5
Learning Data Analysis	6
Why Were These Needed?	6
Similarities with Cybersecurity	6
How did ODU Help?	7
Were My Outcomes Achieved?	8
The Aspects of My Internship	8
My Recommendations for the Data Analyst Position	9
Conclusion	10
References	11

#### **Introduction**

Throughout my time in college, I had never really decided on a proper internship position to pursue. I had always known in the back of my mind that I needed an internship but hadn't put much thought into what specific company I should choose. I ended up being recommended a few places but after overviewing their requirements, they either had no positions opened or none which fit the criteria I was expecting. This stayed about the same until the summer before the start of my final year of college. I had seen an advertisement about an internship at Brooks Crossing Innovation Lab. A friend originally recommended this place as they had previously interned here. I sent an application in and received a response a few days later. After a few back-and-forth emails, I eventually landed an opportunity for a new position as a data analyst for the following spring. As for why it was in the spring, I had chosen that to coincide with my required internship class.

A bit of background information on my internship. The Brooks Crossing Innovation Lab is an Old Dominion University community outreach facility centered around K12 and group learning. The facility also serves events and walk-ins during its daily hours of operation (ODU, 2024). Due to being a remote worker, the data analyst position fits the best with my needs. Most of my position would be organizing and analyzing data from field trips and entrances into the facility. While it wasn't exactly a cybersecurity internship, it was similar enough for me to believe I could gain something worthwhile from it.

As for my initial training, after a series of introductions and online meetings with my supervisor, Kaitlyn McCoy, I was given access to read and write data from multiple spreadsheets and Google forms. My task was to sort and organize it into a digestible form while also allowing enough data to be readily available for any summer reports by other workers inside the iLab. The task seemed simple enough. Due to this, I decided to set a few goals for myself to achieve during the semester. I would first familiarize myself with the main form of data used, Google Sheets. As I had a decent background in Excel and Google Sheets before, I believed further delving into detail on both of them would be beneficial for me and the iLab in the future. The second goal I had hoped to achieve would be to find out if data analysis was a potential alternative for me in the work field. My final learning goal was to understand how remote work affected my mentality on work and my general productivity.

#### The Work Environment

As I was a remote worker, I cannot say how the work environment was as an in-person intern. Though my friend who had recommended the internship told me the environment was very easygoing and helpful. From my emails with my supervisor, this seemed to be the sentiment I found. My supervisor was extremely supportive of the goals I had laid out and allowed me to work at my own pace over a direct work schedule. I believe this was a double-edged sword however, as due to the aspect of a nonexistent work schedule, lack of interaction with other employees, and the zero pay of the internship, I sometimes felt extremely unmotivated on my tasks. I do not believe that this is any fault of the iLab however, and just a result of the structure of the job itself. I believe that in the future, more mandatory meetings and specific interaction with other workers would be extremely beneficial in promoting work motivation. Adding on to this, I was only one of two remote interns both tasked with the data analyst role. This role was added this spring semester, and as such no expectations were created around it. I believe this could only be fixed through time as more is learned.

# **The Starting Point**

As stated previously, my main responsibility was the organization and implementation of data into a main data sheet. This was a broad responsibility, as many of the other duties were my own that I decided upon. Those were learning data analysis, finding data program alternatives, and educating myself on the current programs used by the iLab. Upon starting my internship, there were multiple forms of data either not implemented or scattered throughout random pages. Before tackling this behemoth with no experience, I decided first to educate myself on data sorting techniques and organization options. I believed that without first educating myself on what I was exactly doing, I'd end up messing things up further or quitting out of frustration. As the entirety of the main data sheet was on Google Sheets, I first started there. After a few hours of learning sorting methods and filter options, I believed I was ready to start my internship. After a staggering amount of time sorting through and organizing data, I managed to get through it.

#### **Finding Program Alternatives**

I decided the amount of time I had spent deciphering and sorting that data had taken far too long for my liking. My next self-imposed task would be to find data-sorting program alternatives to Excel and Google Sheets. After a bit of research, I found a few alternatives like Python and Tableau, an interchangeable, AI analytic software focusing directly on data visualization tools and methods (Dambrioso, 2023). Though none of these sat quite right with me. After being given another task by my supervisor of sorting through an email list and removing any potential duplicates, I was stumped. While talking about it to one of my roommates, who had a background in data analysis, he had recommended to me a software known as Google Refine. After a bit of searching, I found it was exactly what I was looking for, and decided to try it on the task I was given. Google Refine which was a previous version of what is now known as OpenRefine, was centered directly on cleaning, sorting, and transforming messy data into a presentable form. One of the main aspects of this was a tool for detecting duplicate entries throughout databases. This along with the sorting options inside OpenRefine made this email task a breeze compared to what I was working with before. I decided this program would be the best alternative I could find for data organization.

# **Further Educating Myself**

After finding out about OpenRefine, I decided my next step would be to explore in detail every aspect of Excel, Google Sheets, and OpenRefine. This task proved to be the longest one throughout my entire internship lasting for over a month and a half. As this was a self-imposed task I worked on it while performing minimal data changes whenever new field trips and events were completed. I ended up finding and reviewing a variety of different tutorials both written and in video form. This research resulted in me growing far more experience than before in both Excel and Google Sheets, and a relative comfort in OpenRefine. However, this is where I encountered a slight issue. OpenRefine had relatively no tutorials or written documentation online, meaning most of my learning was out of date. Though the project was routinely updated, I had to mainly learn features from testing and tweaking generic data. I did this by creating data that was structured similarly to my internship data and then manipulating it over and over to see which results worked and which did not.

#### **Learning Data Analysis**

The final self-imposed task I put on myself was learning data analysis and which methods would best suit the iLab. Most of my information was found in a comprehensive online guide by Simplilearn. This guide went over the direct importance of data analysis for business, the process of data analysis, and specific data analysis methods. While I did a quick overview of each method listed, I found four methods that I believe suited the iLab the best. Those are text, descriptive, and predictive analysis. Text analysis or *natural language processing* involves extracting insights from textual data and is widely used for analyzing customer reviews, social media content, and news articles. I believe it would best work for the feedback sections of data in the iLab. Descriptive analysis involves summarizing and describing the main features of a dataset and is mainly used to provide neat overviews of data to find patterns or trends. Predictive analysis uses historical data to make predictions or forecasts about future outcomes (Simplilearn, 2024). All of these used in conjunction with each other could be used to analyze what works and doesn't work with the iLab.

# Why Were These Needed?

I believe that each self-imposed task I placed on myself was extremely necessary for both myself and the iLab. Without first understanding what exactly I was doing, I could not accurately complete my internship to the best of my ability. By finding program alternatives to the current Google programs used, I found potential avenues for quicker work and stronger organization methods. OpenRefine in particular was extremely easy to use and made my initial tasks a breeze. Learning data analysis methods and data analysis itself allowed me to create a report about my findings through my internship, and is directly necessary for businesses to succeed. Without also researching data analysis, I could not gain much insight from the data I was organizing and analyzing.

# **Similarities with Cybersecurity**

While the job was not directly related to cybersecurity, I believe some of the skills I had gained from my college curriculum of cybersecurity were able to be applied. The initial skills I had formed from my four years of college were problem-solving and adaptability. I directly used my problem-solving skills to find solutions to sorting and organizing data in a way that was both quick and efficient. However, the main similarity I noticed after reviewing my time at my internship was the ability to visualize data. While it was through different methods, data analysis and visualization is an extremely necessary skill to have as both a cybersecurity worker and a data analyst. Data analysis allows you to understand the data is telling you. Say an event log lists random suspicious data logs, or a server has seemingly random actions that should not be happening. Data analysis allows us to put the pieces together that a security threat or intrusion has taken place. Data analysis also allows cybersecurity workers to protect against threats before they even happen by understanding what the data is telling you. My direct experience with learning data analysis at my internship has allowed me better to connect data analysis with cybersecurity as a whole.

# How did ODU Help?

My cybersecurity curriculum had noticeable effects on my data analyst internship. While the direct material was not the same, I was able to use the concepts and skills I had learned from my classes to better adapt to a data analyst environment. My classes allowed me to develop problem-solving skills which I had not had before my time at college. Along with this, I learned a specific skill of adaptation. Adapting to problems on the fly has become something I've gotten quite good at over time at ODU. I believe the strongest strength my classes reinforced in me was the drive for further learning and knowledge in my career. These three skills worked in tandem and each was heavily reinforced during my time at my internship. My first task required all three to even start from a reasonable method. My ability to problem solve and my further drive for knowledge led me to find out about OpenRefine in the first place. I believe the strongest skill I gained from my internship that directly ties in with my career was the ability to strongly analyze and visualize data and its meaning. Most of the time in my classes, I was usually given a blanket explanation or an in-depth explanation of what I was looking at. Be it an intrusion method or a direct network protection, I was always given at least a brief description of it. In the case of my internship, I had none of this. I believe it was an eye-opening experience. In the field, most of the time the intrusions you face will be documented, but not to the extent as portrayed during college curriculums. Learning how to accurately analyze and visualize data allowed me to draw connections and conclusions on what the data was telling me.

#### Were My Outcomes Achieved?

I believe that each of my outcomes was achieved at least partially. My first learning goal of learning more about Excel and Google Sheets was most definitely reached as I spent over fifty hours learning in detail about every aspect I could find of them. I even ended up finding a third avenue which I hadn't expected to find in OpenRefine. I believe this internship gave me the drive and opportunity to finally have a reason to further my education in these programs.

My second learning goal was to see if I could potentially find a separate avenue of online work that did not relate exactly to cybersecurity. I believe that this goal was slightly fulfilled. While I have learned a lot about data analysis and visualization, I believe it is a career that I don't exactly see myself pursuing. However, I believe that the skills I've gained from my internship were extremely valuable and worth my time.

My final learning goal was much less a goal and more of a mental test for myself. I specifically wanted to find a remote-based internship to see how it directly affected my mental health and my overall work productivity. What I found was far different than what I expected. I realized how much motivation I had lost without a group environment. While I believe that remote work and learning can be beneficial, I believe there must be an aspect of group work and collaboration to promote work productivity. A few other aspects I believe I must mention were that this internship had a make-your-own-hours-type schedule and was unpaid. I believe these two aspects directly led to me having increased motivation and work issues. Though again this was not the fault of the iLab. Overall, I believe that this internship gave me great insight into every learning goal I set for myself, and I felt fulfilled on every one of them.

# **The Aspects of My Internship**

The most motivating and memorable part of this internship was the learning opportunities I gained. Without this internship, I would not have found OpenRefine or learned about as much detail as I had on Excel or Google Sheets. I also learned about a new skill that I can use in tandem with the skills I learned during my cybersecurity curriculum.

I believe the most discouraging aspect of this internship was the structure of the data analyst position. The overall structure I believe is a little too lax regarding remote

workers as I had gone for almost a month without notifying my supervisor of what I was researching. I believe this in tandem with how the structure of the internship is laid out leads remote employees to feel unmotivated and lack a desire to work. This along with the unpaid aspect and overall unorganized nature of this position made it feel sometimes like I was not a beneficial aspect of the iLab.

This then led to what I believe to have been the most challenging aspect of the job. That is the overall layout of what is expected by data analysts. While the position is labeled as a data analyst position, it is currently used as a way to sort and organize data for supervisors to accurately create summer reports for board heads. I believe that only having one main direct task of sorting through a main sheet of data led me to struggle to bring myself to work at some times. While I directly asked if I could use some of my time to research data analysis methods and find potential alternatives to current methods, this was not a stated or required goal set by the position. This leads me to what I believe the iLab can do better in the future for the data analyst position rather than what interns themselves can do.

#### My Recommendations for the Data Analyst Position

I believe the first recommendation that would promote work growth would be having more frequent and mandatory meetings with data analysts and the supervisor. This would make it easier for collaboration between data analysts. Remote work in its current state should also specifically not be allowed for this position due to the facts stated previously. A newer version of remote work would be necessary. A method that focuses on promoting work productivity and motivation would be enforcing deadlines, adding more tasks, or creating a general work schedule for interns to follow. I believe this along with a small report I created that overviews my findings would be another potential alternative to try for future interns. The report includes a general analysis of the current data of iLab along with a few potential program alternatives and methods to try in the future. After all, the position is labeled as a data analyst, not a data organizer.

# **Conclusion**

The main takeaways I gathered from this internship were an introduction to actual remote work and a few skills that I had not found or strengthened before my time started. This internship also allowed me to find out a very important aspect of my career that was previously not known. I believe this internship was extremely beneficial for myself and my career in the future, and I have gained extremely valuable knowledge and experience that I would never have gotten through my own time. This internship influenced the other two classes I was attending during this semester as well, as I tried to use my newfound knowledge on exams and quizzes. I believe this allowed me to further understand exactly what I was looking at. Without this internship, I believe that I would have potentially lost or taken a lot longer on potential problems I faced in the workforce than I can now. Overall, this internship influenced how I perceive data in my current career and I will use the concepts learned from this internship in my future professional career.

# **References**

Brooks Crossing Innovation Lab.

https://www.odu.edu/brooks-crossing-innovation-lab. (2024, March 27). https://www.odu.edu/brooks-crossing-innovation-lab

Dambrosio, L. (2023, May 24). *What is Tableau & What is it used for? an analyst's guide*. Mode. https://mode.com/blog/what-is-tableau

Simplilearn. (2024, February 16). What is Data Analysis: A comprehensive guide. Simplilearn.com.

https://www.simplilearn.com/data-analysis-methods-process-types-article