Ben Phan's End-of-Semester Independent Study Report

I. Background

Before joining the research group, I have some knowledge about Python coding. I have taken CSC 110 (Intro to Python), and two Global Tech Experience classes about Data Wrangling with SQL and Python. Through those courses, I learned some basics about Python coding and could create simple projects like a 1D chess game or Benford's Law checking for a data set. I know how to work on cleaning, analyzing, and visualizing data using SQL, Python Pandas Library, and Tableau. I also know a little bit about how to do photometry and spectroscopy through my Observational Astronomy course.

II. Progress

After joining the research group, I have learned much more about Python, especially the NumPy Library. NumPy array is what I studied for my first assignment, and I understood how to use NumPy array for image processing, as well as basic array operations such as slicing, changing shapes, and splitting arrays. My second assignment is to study Pandas Dataframe. Having learned about the Pandas Library a year prior to my Independent Study, I have quickly reviewed the topic. A month has passed since; however, I did not play with the code but studied things theoretically on paper instead, Matthew suggested I play around more so I did create a few Pandas Dataframes via coding. Around this time, Nikhil suggested that I create a GitHub account, install VS Code on my computer, and experiment with the Linux operating system. I was not familiar with Linux before so this was a very eye-opening experience for me. Through this assignment, I have learned so much about how different operating systems work and how Linux is like an inside-out Window. I was also taught how to run a virtual machine around this time. Running Linux through Ubuntu using a virtual machine is also a surreal experience for me. All my coding experience before was on Jupyter Notebook, so coding through Linux is new and exciting. My first assignment on Linux is to install VS Code using only the terminal. This assignment is also very interesting to me because I have always downloaded things by going on Google, searching for the download content, and then clicking on the "download" button. However, using the terminal, I have to figure out the code to successfully download the content, which to me is very intriguing to learn. Matthew gave

me the next assignment on Linux: installing an Apache web server, and I learned how to do it over the spring break. It was also a fun process. I remembered sending Nikhil a text, saying that I felt like a real programmer now that I have successfully created a simple website saying "Linear Algebra Sucks" on Linux (I had a Linear Algebra test that day). After getting familiar with Linux, Matthew gave me my first project (which is also my final task for this semester): I need to create a website whose function is to receive URL links, and then input the URL into a data frame shown as a table on the website. Nikhil suggested I get familiar with the programming language Streamlit, as they are very easy to learn and use to create clean websites. This is by far my favorite task of the semester, as it makes me work on both the front and back ends of a website, incorporating what I have learned while also using my problem-solving skills. I have run into some problems during this task, mainly how to save the input data via file since the input URL would immediately go away once the user entered the second URL. Another problem that for me, is even harder to solve is installing Python on my computer. At first, I could not run my localhost website despite having all the code run perfectly. It took me some days to even detect the problem as I was not used to running Python on my terminal. After a few days of learning how to install Python and running into so many other small troubles, I successfully launched the website. Despite this being a relatively small and simple project, I am proud of what I have done and happy with the results. However, I wish I could learn how to turn the URL into actual clickable URLs on my data frame, as that would be a bit more practical.

III. Conclusion

First of all, I would like to thank Professor Impey, Matthew, and Nikhil for even letting me join the research group in the first place. I do not have much coding experience, and the project has been around for 5 whole years by the time I joined. Despite knowing that I was not able to contribute to the project, you still accepted me with open arms, as well as spent time helping me learn a lot in the past semester. Even though I was not able to do much for the group, I have broadened my knowledge of computer science, which is a big cornerstone for me moving forward. I have learned more about NumPy, Pandas, Streamlit, and Linux during my Independent Study, which will become valuable assets

for me in the future. I wish all the best for the group as the project is coming to an end. It has been a pleasure for me this semester, and thank you very much once again.