John Shumway

SNHU : CS 499

August 13th, 2021

Self-Assessment

The coursework from the SNHU computer science course has really improved my skills in the computer science field. When I originally entered this field, I was planning on learning various languages and aiming to be an all-around programmer. However, as I studied further into the curriculum I found that data related courses and data mining like courses was something I found enjoyable. After some additional research, the data science and engineering field is something that interests me and would like to pursue further. Even after the curriculum, I plan to continue researching and studying further into data science.

One of the most notable features and programming tools I’ve learned in this course is the inclusion of GIT. The team collaboration tool is invaluable for any developer working in a team environment and knowing how to use it is a crucial skill for new and veteran programmers. I originally practiced GIT in JAVA, though I’ve also learned how to use it with the GIT bash shell, Python, and Visual Studio. Even now, when learning a new language GIT is one of the main features I also research into and is a core concept to practice the new language with.

In the computer science field, knowing which data structure to use is part of being a good engineer. I’ve learned and experienced more about Hash tables and Binary trees, on top of already being familiar with lists and dictionaries. Hash tables is a feature used commonly in computer science and something I would like to practice with more often outside of school projects.

The various Database work throughout the curriculum has vastly expanded my knowledge with various features and the different kinds of databases. In the past I have used the Microsoft SQL database for projects, but the courses here have allowed me to work with mySQL, SQLite, and MongoDB. Additionally, I’ve also learned how to work with databases in both Windows and Linux OS. As data science is a field I would like to explore more into, I’m happy to have learned and practiced how to effectively use both Linux and Windows for database tools. As I go forward with database practice, database security is another feature I’m always looking out for. Stored procedures, CRUD framework, encryption, all of these and more are items I’m always looking out to improve upon to make the data secure.

While not covered in the curriculum as much, cloud computing is another emerging technology that is quickly becoming useful in the computer science and data engineering fields. I’ve dabbled with Amazon web services before, and have been looking into exploring Azure technology going forward. I have also noticed Microsoft offers Azure certificates which is something I would like to explore after this course.

Overall, the courses have really improved upon how I approach and engineer solutions. One of the bigger concepts I’ve learned is the importance of unit testing. Even in my current programming job I’ve started to use unit testing more often. The process of creating a suite that checks if the program I’m working on has not been modified too much and still retains the core functions has been very useful. Even minor changes that wouldn’t affect the code too much can still break another core feature and then the programmer is wasting time debugging the error. Test cases have helped me solve issues before releasing code and create stronger programs.

The goal of the capstone project and my portfolio is to provide a showcase introduction into data science. The SNHU capstone project contains 4 major artifacts: The CRUD framework, the test cases, the database, and the algorithms to build the data charts. The language used for the project is Python as it’s one of the languages suited for data science. Python is a well supported language that has stood out in performance and ease-of-use in the data world. This project will exhibit my ability to design a data server, populate it with random aggregated data, then gather and display various information into readable charts. The CRUD framework and test case artifacts were designed to show what I learned in building well written code to connect and retrieve data from a database.

While the CRUD framework was done in Python, many of the other components used different languages. The test cases originally done in JAVA were reviewed and converted into Python. The test cases were a good pick as I wanted to showcase the use of the concept and that it isn’t unique to any specific language. Along with the CRUD framework, It was also a great way to start the foundation of the Python project.

Next we have the algorithms and data structures used. Looking over the many different kinds of structures, I choose something simple to use – lists and dictionaries. Hash tables and Binary Trees have more advanced features and are more efficient with searching or sorting data, however many of those features would be covered in the CRUD framework. Additionally, Lists uses very little memory compared to the other structures, and was easy to convert to dictionaries which were used by the algorithms to create graphs. The algorithms use a python import called “matplotlib” which is a popular tool in both Python and the data science world where it can easily create well designed flexible charts. That being said, I designed the modules so that each module can be independent of each other. While the sample graphs I created used lists, more advanced modules could require more detailed results and use Hash tables or Binary Trees, or any other structure as needed.

The final artifact is the database component. Due to the time constraint, I wanted to keep the design simple to show the ability to create and use a database using mySQL. Since I’ve used Microsoft SQL, SQLite and Mongo in the past, I wanted to try my hand at another major database repository. If given more time and larger scope I would have liked to build something closer to a data warehouse or data lake. I have worked with CRUD in MongoDB before, and that’s where I got the general idea of using the framework. Usually I work with existing databases so designing the database from the ground up was a fun experience, and something I would like to do more of.