Martin Jimenez

CS 1030-001

12-10-2023

Khan Academy Report

**Data Tools:**

The data tools module was very helpful because it allowed me to explore not only datasheets, but gave me exposure to CSV files. I didn’t focus on one in particular because I wanted to familiarize myself with both. Although I did copy that dataset into my own Google spreadsheet, I struggled to perform certain functions that Khan Academy was teaching me. It was also nice being able to see the trends in the data sets represented in a visual form with graphs. I got a 75% on the exercise for this module, so it said I was familiar with the information.

**Big Data:**

I am not sure why I struggled so much with the big data exercise at the end of this module. I was pretty familiar and understanding of the information that was being presented throughout the articles, but apparently I wasn’t when it came to answering the questions. This is something that I want to be familiar with because we do live in a world where there is a vast amount of data that is being collected every single day. We are going to be using all of this data to hopefully improve various aspects of our lives. I found it notable that in this module, one of the articles said, “Ultimately, we need to understand that a prediction is just that, a prediction. More data and better techniques helps us to predict the future better, but nothing can guarantee a perfectly accurate prediction.” There were some concepts here that were similar to the ones that were explored in *Hello World*.

**Bias in Machine Learning:**

I really enjoyed this module because it exposes the biases that could be imposed based on using machine learning systems.Depending on the information these systems are using, it could pose a bias in the results that it gives us. It is important to take these into consideration because we do not want the biases to have us go backwards as a society. If we can find alternatives and improvements to make in machine learning systems and eliminate the biases, then they will be really useful to everyone. This was my best exercise since I got a 100% on it.

**Unit Test Activity:**

Overall the unit test activity wasn’t too difficult. I received an 80 percent, and again it was mainly the questions similar to the ones in the “big data” module that I struggled with. The questions presented in the unit test were very similar to the ones that were given to me in the rest of the exercises, therefore it shouldn’t have been hard for anyone that was actually reading through the modules. The three modules didn’t take me very long to get through and it was a beneficial experience for me since I was able to take away some information and will be implementing it into my project.