Rhodes Ferris

Data Analysis - Application and Presentation

Khan Academy Report

Data Tools

In this module I learned different ways to store data as well as different ways to interpret data. The main lesson I took away from this module is that data is contextual. Just because there is a correlation doesn't mean that there is causation. Applying the necessary context to data is key in order to ensure it is being used properly. Data can then be used to make inferences or predictions based on the data when the data is properly used in its context.

Big Data

In this module I realized how much data there actually is in the world. Big data can be pretty tricky to sort however with the use of parallel processing data can be sorted much faster. Storing data is another issue that is facing the tech industry as more and more data continues to be collected. Server rooms are gigantic and require serious amounts of space and hardware in order to function. Based on the sheer amount of data there is and the current limitations we face with sorting and storing data, something needs to change soon in order to make big data more manageable if we are going to continue to collect more and more data.

Bias in Machine Learning

This module made it clear to me that bias from humans can be very easily put onto a machine learning algorithm even if the humans creating it didn't intend for that to be the case. Training a machine learning algorithm is incredibly difficult and impossible to truly do bias free. Since bias exists in the world and machine learning algorithms need a reference in order to be trained, then they will inherently have bias. This problem seems to be growing over time given our society's increased dependency on machine learning algorithms. Said dependency ironically exists as a means to combat bias, however the use of machine learning algorithms seems to only be perpetuating the problem, since the biases we are trying to combat are the same biases that are used when training a machine learning algorithm.

Unit Test

Overall, the unit test wasn't very difficult and I got a perfect score. Most likely this was because the test contained recycled questions from exercises within the modules I had done previously. The questions were good though and helped put the lessons from the modules into more real life context.