

Letitia Fickling

December 8, 2022

CS 1030

Data Analysis Project

1. Data Tools

Honestly this was mostly a review for me. I have an excel certification, I've taken the statistical methods course offered by the math department here at MSU, and I've taken the CS department's databases course. I feel very comfortable with data analysis and statistics. Because of that I went through this section quickly and got full marks on the quizzes. That said, I haven't used SQL in a while, and review is always useful. I definitely remembered some things I had forgotten.

2. Big Data

This was a shorter section. Again, more information about big data. This lesson was more surface level than the data tools lesson so I felt like I got less of a review from it. I'm always blown away by discussions of just how much data there is in the world. I'm glad they dipped into the ethical considerations but I would have liked them to address some of the climate ethics as well since data takes a lot of resources to maintain. There also should maybe have been discussion of companies like Palantir and the way big data is used to track people.

3. Bias in Machine Learning

I'm currently taking NLP with Deep Learning and my new job is working on a MADS (measurement and data science) team, so this entire unit is very relevant to me. More

importantly, I'm currently taking the computer ethics course required for the CS major and machine learning bias has been one of our main focuses. We also have covered a lot of the things brought up in this unit during this class through our reading of 'Hello World'. Between that and 'weapons of math destruction' I've heard a lot of these stories before. But it is very important to keep repeating them so they stay front of mind and it's great that khan academy includes them in this lesson aimed at beginners. So many people look at machine learning as magic, but it has major flaws and blind spots. One issue I had with the section was that it was very US centric. Khan academy is used by people all over the world but when talking about facial recognition use by police they spoke about the relatively small use of it by US law enforcement and not at all about countries like the UK and China where its use is much more widespread.

4. Unit Test Activity Experience

I just took the unit test. It seemed to follow very nicely from the section quizzes in the rest of the unit. Some of the questions even re-used set-up scenarios which is always nice because it means it takes less time to read the question and get the gist of what they are asking. It seemed completely doable if you'd read the articles, but not so easy that you didn't have to actually read the questions. I got a 100%. I've included a screenshot of my results below.



Unit test

Unit test: 100%
Data analysis



Computing > AP®/College Computer
Science Principles > Data analysis > Unit test

Unit test

Google Classroom



Leveled up: 4 skills

Leveled down: 0 skills

No change: 0 skills

9/9 correct · 395 energy pts

Skill level changes

Computing basic statistics



Finding patterns in data sets



Bias in machine learning



Big data

