Data Tools

This part wasn't too surprising to me since I am in an information foundations class right now, and I had to use Exel a ton. Datasheets tend to be something a lot of companies use since they are very easy to navigate for users. CVS's are harder to read but are the basic format for many tools. The flaw is the fact that it holds only one file, and often times once the file gets larger, the processing time increases. One thing that helps a lot is the ability to upload and change CVS sheets into spreadsheets and database formats and save them as a CVS later, that way it can be easily read and edited by most people.

Big Data

I knew there was a lot of data within the world of the interweb, but I never knew the exact number. 175 zettabytes is a much larger number than most humans can comprehend. My theory about what file sizes were bigger was correct too, from smallest to largest, text, image, gif, audio, video, then audio+video files. Meaning, keeping things as image and text files only, at least as much as possible, is the best way to save on data space. It was also fun seeing that these "data wear houses," are just large computer rooms with shelves upon shelves of hard-disc drives.

One last little side note facebooks data output is very unsurprising to me. From all I have read and seen, that app is one of the largest data mines to date.

Bias in Machine Learning

Reading about this in hello world, and seeing it in my criminal justice class, nothing was new about this. AI and machine learning programs learn from human input, and depending on the sources, can turn up very undesirable results. Supervised or unsupervised, most machine learning will learn human bias. Blacks being flagged more than whites being the prime example. Face recognition seeing people who look similar and causing innocent people to get arrested. It's just a big mess.

Unit Test

I don't really like tests of any form, but this one wasn't too stressful. If you got an answer wrong, it explained it to you, and it explained why the answer was right, which I did enjoy. It didn't really feel like a test, just a review. I wish there were more activities like this instead of high-stress one-chance tests.