[Enrique Sotomayor](mailto:riquesoto2003@gmail.com)

Computer science principles

Prof.Cohen

5/3/2023

Data tools

I was unaware of how useful the Csv files actually are when it comes to data storage for various things. I was unaware that there were some negative aspects to these types of files which are that they have somewhat of a limited space to store information.The other negative aspect being that they do not have built in tools to manipulate the data on hand. I found the spreadsheet section to be rather informational and resourceful. As much as I dislike spreadsheets, their results cannot be argued. I found the reading rather informative over concepts that I have not given much thought to.

Big data

The Idea of increasing storage space was always a puzzling question that I have wondered about. Zetabytes, although impressive, seem to be a bit much more than is needed in terms of storage space. I was completely unaware that the satellites in orbit are still active and storing multiple terabytes worth of sensory information. I like the idea of internet archives and data libraries as their information is accessible to all who wish to learn. Overall the material was very educational and simple to comprehend.

Bias in machine learning

Machine learning is a very fascinating subject that has become more prominent in the recent months. The concept of a neural neural network is quite fascinating; it always appears more complex than it actually is on paper. facial recognition software is quite impressive seeing how much it has advanced in the last few years. It is unfortunate that the software also has a few problems as it can perceive two separate people as one person. Language translation software is also an astonishing feat in recent years. However, it still requires some improvement as certain words in a language have multiple meanings and the software will oftentimes give the wrong one. The concepts were quite informative and only added to the information that I already know.

Unit test

I found the test to be somewhat difficult as the wording on the questions was a bit confusing at times. The questions with one answer were often easier to understand and answer as the wording was less comprehensive. The questions that had multiple answers were rather difficult in my opinion as I had to read the question multiple times for it to make sense. I found the questions with a scatterplot graph to be simple and straightforward.For the hypothesis questions I found them to be quite easy as there was only one correct answer which was simple to deduce. The test had its ups and downs however it was a proper test of knowledge.

Hypothesis for data collection

The life expectancy has improved for males and females since the early nineteen hundreds.

Findings

Both males and females have improved from 1900 to 1950 the rate had a bit of a mixed result as it was inconsistent as when it increased it did so it was by 2.5%. When it decreased it was by 2.3%. From 1950 to 2018 the expectancy rate was more linear however there were a few years where it stayed the same. The rate increased by about 0.2% and did not really decrease.

For females the life expectancy rate is not linear in the years between 1900 to 1950.The life expectancy rate increases by 3 percent in certain years. When the rate for expectancy decreases it is by 1.8%. From 1950 to 2018 the rate for expectancy is much more linear and does not decrease all that much, however it does stay the same for a few years. When the rate increases it is by 0.4%. When the rate does decrease it is by 0.1%.

The male expectancy rate from 1900 to 1950 is not linear as it increases and decreases in certain years. When the rate increases for males it is about 2.5 to 3 percent. When the expectancy rate decreases for males it is by 3.4 percent. From 1950 to 2018 the life expectancy rate for males is more of a linear increase however it does decrease more often than females. The rate increases by 0.1 to 0.2 percent. When the rate decreases it is by 0.4 percent.





