

Cpts_575: Data Science

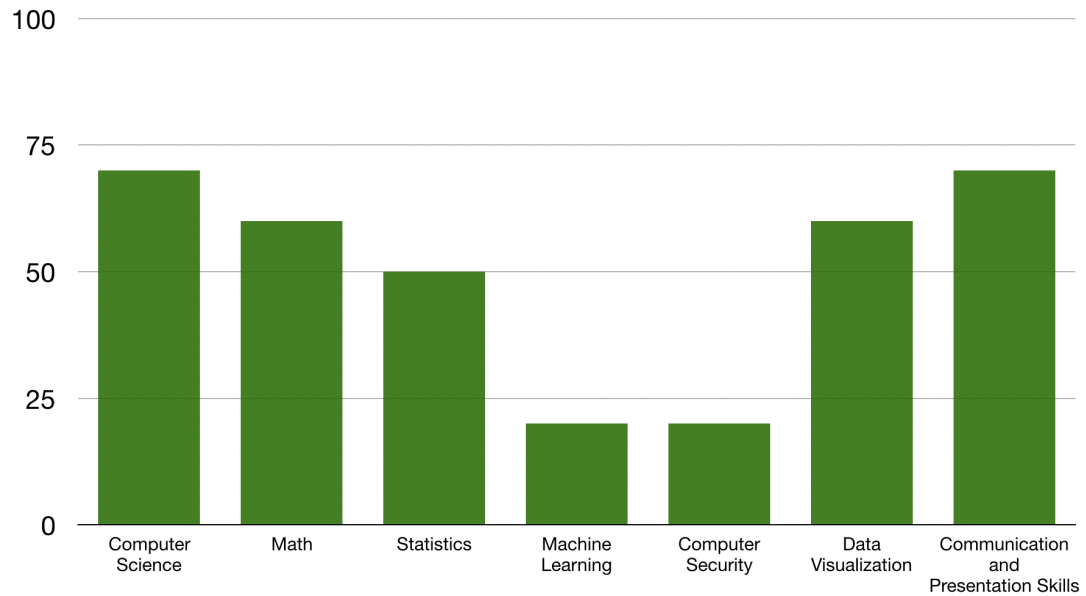
Assignment I

08/28/2019

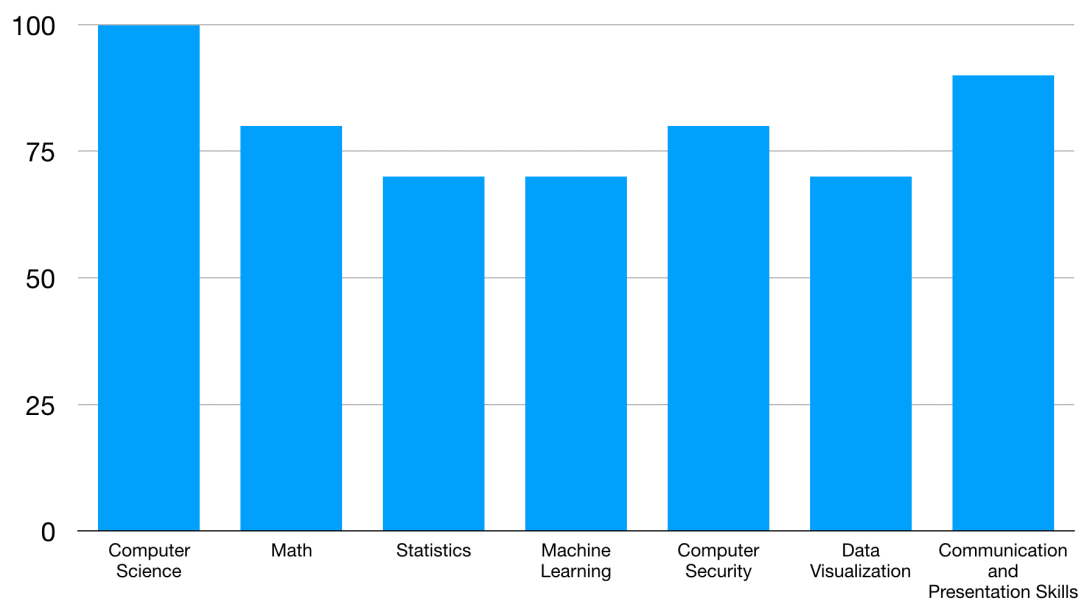
Instructor: Assefaw Gebremedhin

Student: Yu-Chieh Wang 11641327

1. Data science profile



Picture 1: Current Profile



Picture 2: Future Profile

- A. In these 7 areas, I believe that the skill of computer science is the most important one because of two reasons. First of all, as a software engineer, knowing how to operate computers well and solve program bugs are the basic requirements. Second, during learning classes in computer science, people are trained to think logically. It is a cornerstone for helping learning math and statistics.
- B. I believe that computer security should be added because when doing data science, people need to make sure the accuracy of data sources and analysis results. Also, I think that math should be removed because computer science and statistics already cover everything we need in data science.

2. The article "Data Science and Prediction"

- A. The author mentioned three differences between data science and statistics. First of all, the data become more unstructured because the relationships between data are more complicated now. Second, analyzing data requires more than two areas of cooperation, such as computer science, linguistics, econometrics, sociology, and other discipline. Therefore, it is impossible to complete the analysis by statistics alone. Finally, using the classification label attached to the data, computers can automatically interpret the data and make decisions which is the statistic cannot reach.
- B. The author points out that the reason why big data can make both theories of physical and social science solid is because the big data is able to reduce the errors of prediction. Therefore, the prediction results can be used to solid those theories. According to the article, there are three common prediction errors: misspecification of a model, fewer samples used for analysis, and randomness data in samples. However, the big data can solve two of these errors. It not only can make accurate prediction models, but also can provide a large number of parameter samples.

C. How data science affects humans future?

Nowadays, data science becomes a new tendency in our world. The reason why it is so popular is because people can make predictions through the analysis of big data. To compare with the traditional statistics and data science, data science is more effective because computers can make accurate predictions automatically. On the other hand, the traditional statistics depends on analyzing data by human which has more human mistakes. Indeed, computers can find some questions and problems that human is not able to find. Therefore, based on the ability of data science, human beings can predict many things in social, medical, and commercial fields that will help humans control future development.