

Reflection Essay

Honglin Zheng(hz2682)

Time passed quickly. The first class was still vivid in my memory although the semester has already ended. It is my great pleasure to meet everyone in the class who I learned a lot from. In this reflection essay, I am going to share my gains from the course and how this wonderful journey shaped my mind.

Before the class, I imagined I would probably learn and grow to become a data scientist or data analyst. I expected I would become expertised at coding in Python. I felt very nervous about whether or not I could succeed in my studies. However, in the first class, we were actually informed that this course was supposed to be an introductory course about understanding how educational researchers might leverage data mining to fulfill educational goals. Educational researchers don't need to be excellent at analyzing data, which is a job of other experts majoring in statistics or math. The most important issue is about education per se. Not playing with data. This was how I started to understand the definition of educational data mining.

Actually I have already had experience in data analysis before this semester with SPSS, including linear regression, logistic regression and ANOVA. However, as long as I learnt how to use Python to run the regression analysis, the notebook interface in Anaconda impressed me with its elegance and clearance. I prefer to use Python rather than SPSS in some cases.

Not only learning regression, of course, we also learnt cluster analysis, social network analysis and principal component analysis during the semester. I appreciate that all the well-prepared ICE documents include example codes and representative educational related questions. For example, when I filled the questionnaire about all the courses that I had attended, I did not expect the data would be put into ICE to present how social network analysis works. It actually surprised me a little when I read the map and found my name on it. I forgot I had taken several courses with Karl and Pooja! This interesting moment impressed me and it was absolutely credited to your great effort in preparing all of the lectures.

Moreover, I also appreciate the ACA assignments, especially ACA1 and ACA3 that are about selecting universities. In fact, that is the largest data set I have ever worked on. In the beginning, I applied for a screening condition which only accepted universities with an education department. However, the information was not accurate, some outstanding universities with an education department were not excluded, such as Columbia University, UCLA and others. After a discussion, my group agreed on the opinion that missing information is better than wrong information. Therefore, we decided to ignore the requirement of having an education department in data selection. To some extent, students in any major can engage in an education career after graduation. This critical thinking process impressed me and I really enjoy it. In practice, it is impossible to have perfect data, we need to compromise with what we have and make the best use of it.

There are many gains in the creative EDM project as well. I learnt from my group members, and during the presentation, I also learnt from our audience. Kan asked that why we only considered

the friendship that last for a whole semester. This is a good question! Although those friendship that last less than a semester are excluded, we still need to control the factor of friendship duration, in case the friendship last for years might cause a difference than those that are just a little bit longer than a semester. After discussion, my group decided to change the study design to only assess friendship in the last week of the semester, while instructing students to report friendship duration and friendship quality, which are considered simultaneously in calculating the weight in the friendship map. If we didn't have such a presentation in class or if the class didn't listen at all, we wouldn't have corrected this problem in our research plan!

Although I feel reluctant to say goodbye to this course, it is time to carry on all of the gains and set off to the next journey. I am determined to continue learning coding skills and apply it into application.