

Prompt: What are you learning about this kind of work? Does it appeal to you? Why/why not? Do you notice anything you have read for classes in this program playing out in the real world of your practicum?

Before I participated in this work, all I learn about institutional research (IR) were from the course I took. I understand the definition of IR, general tasks, and essential skills. However, the majority of them are book knowledge, which still has gaps with real work.

For this work, my main responsibility is to build an inferential model to explore factors that contribute to student retention at Ohio University. At first, I thought the most important part is to conduct analysis (i.e., choose variables, build models, run analysis, generate results, etc.). Nevertheless, when I opened the database, I was stuck with various variables. Even though I tried to read through those variables' descriptions, it is difficult to distinguish them. Some variables I am not familiar with because of cultural reasons, such as expected family contributions; some are because of their similar names and definitions, such as the relative distributed amount and distributed amount. With my supervisor's help and clarification, I still took plenty of time to figure out their meanings.

After I figured out the variables, choosing between them is another challenge. Theoretically, literature provides tons of factors that could impact students' retention. Practically, I attempted to find variables in the database I have to measure factors I have read about in the scholarly research. However, in practice, it is impossible and inefficient to put all variables that might impact students' retention in one model. I need to make the tradeoff between variables and choose the most powerful ones,

which need evidence and reasons. That evidence is not only from theory but also from practice.

Other than variables, model choice also needs a tradeoff. At first, we tried to use a multivariate probit model to analyze, but it spends more than two days to run a model. Worse, after two days, it had not finished running. Then, we switched to the multinomial probit model after reading several articles and comparing the pros and cons of those two models. Unfortunately, this model has many issues, such as convergence problems. Moreover, the package we choose has not been used for many years, which means it is hard to find solutions for our problems. Thus, we ended up with a multinomial logit model, which is user-friendly and easy to interpret.

During this process, I learned the importance of tracking my work, making notes of my progress and challenges, communicating with my supervisor, and seeking help from external experts. Unlike writing an academic paper, which requires me to finish alone, this work encourages me to communicate with my supervisor frequently. Discussing with him does not only offer suggestions for me to solve problems but also help to keep my work on track.

Previous works are the interesting parts of this practicum. I enjoy exploring the database, reading articles to find clues, brainstorming, and discussing with my supervisor. Nevertheless, this work also has some boring parts. Specifically, when we used the multivariate probit model and multinomial probit model to analyze, it had usually taken one or even two more days to run all models. Sitting in front of the computer, waiting for the software to run one model, and then clicking the next button to run another model seems like an endless process. Sometimes, after that process, I cannot generate results, or even meet another problem, which was pretty frustrating.