

Program Assessment

Ajuni Sohota

Part One Program Assessment

1. Prepare datasets for solving problems.

I think that I am pretty confident by now knowing how to explore data to understand what is going on and how to prepare. We had a whole course on it data preparation and cleaning, which helped us understand the various ways we can deal with preparing data and then cleaning it for downstream analysis.

2. Recommend appropriate data modeling techniques to test hypotheses.

Through our projects is where I got comfortable dealing with this. Unfortunately just reading about things isn't enough to gain this kind of confidence in determining what to use for what kind of project. Actually doing the project is what helps us get to the bottom of the nitty gritty as to what kinds of things we need to consider when making the decisions of which way to go.

3. Communicate data science results into answers for domain challenges.

Through our presentations and final paper write ups, I got comfortable knowing whether or not I was doing a good job in communicating the details. Through our discussions I felt that we all collaborated and stressed the importance in being able to explain complex data science techniques into laymen's terms for an audience. This is important for funding as well as for understanding between cross-teams within a company.

4. Identify ethical considerations in dataset preparation and modeling.

We discussed many times throughout the course the ethical considerations we as data scientists need to take in order to maintain integrity. We also during our discussion posts wrote up about real-life examples of what kinds of things are occurring in data science that bring up the ethics of how we deal with data. Things like racism and breaching trust were big topics in this.

Part Two Program Assessment

- Where could the program have been stronger?

One thing I wish I learned more about was how to use SQL. Almost every job posting is asking that we have experience with this but we only had one assignment using sqlite and I didn't even feel like I understood how to do that assignment. I think it would benefit people in the program if after completing this program, we were extremely proficient in SQL as well. I also feel like R was mostly used in the statistics course, but I wished there was a little bit more hands on experience with R. The other thing is I think it would be important to step-by-step guide people how to use Anaconda/Jupyter notebook/Google Colab so that we understand how to use it and how to share to Github. I was still pretty confused by this for a long while. Lastly, I think it would be useful for us to know how to use AWS and other things like that because that is also on a lot of job posting.

- Are there any gaps or content you wish had been covered and wasn't?
(see above)
- Are there any suggestions for improving the program?
(see above)
- What are your next plans after completing the program?

Currently applying for jobs but am doing a course in SQL so that I can have a strong background with this.

- Do you feel the program prepared you for being a data scientist?
I think the program prepared me to be a data analyst, I still don't feel I am able to get a data scientist job yet. Working towards it though!