Three things I want to improve on as a data scientist,

After completing my first regression analysis.

December 13th of 2021, I began studying at Flatiron School in the data science program. Since then, I have completed two of five phases and submitted and passed two major project each involving data handling and interpretation. Flatiron broadly falls into a category of vocational training many would refer to as a “coding bootcamp”; I do learn a lot of code at Flatiron and they do indeed market themselves as a bootcamp, but I don’t think that accurately describes what I am learning at Flatiron. The code itself is really just a substrate for an ecosystem of other tools to interact. This was really made apparent to me as soon as I completed all the lessons for phase two and was “ready” to begin the milestone project, which was to a) develop a predictive model using multiple linear regression on a provided dataset, and b) communicate the finding of that model to a non-technical audience. I essentially had to review every lesson in the phase all over again to even begin, because although I knew how to code a regression model, or a Q-Q plot, or any individual component in the entire process I would need to go through, I didn’t really understand what each component contributes to the process and therefore could not organize them (write the code) in an at least effective manner, never mind efficient or elegant. So, after taking the time to really understand each tool and component I had available and how they all interact with each other, I had slowly become quite confident with the project and eventually would pass my project review on the first attempt. Now I am beginning phase three and I don’t suspect it will get easier. So, here are three things I want to improve on as a data scientist moving forward…

I mentioned earlier that I am learning a lot more than just code at this “coding bootcamp”. One of the other things I am learning a lot about is mathematics, statistics, calculus, and algebra especially. I knew there would be a lot of math involved when I committed to this program but somehow its still so much more than I expected. My perspective on mathematics broadly has even evolved throughout this experience. Furthermore, I can intuit how one could be incredibly creative in their approach to problem solving, especially with data, with a strong understanding of the applications and interactions of certain mathematical tools. One day I would like to have a strong enough grasp of the mathematical aspect of what I do that I can be expressive and creative at a high level.

To continue with the theme that Flatiron is more than just a coding bootcamp, I am learning a lot more than just the technical aspects of data science but also the soft skills needed to perform in the professional world such as communication and public speaking. I don’t usually do well in customer facing roles which is certainly a major contributing factor to why I chose to pursue a career in tech, but as a data scientist it will be my job not only to discover actionable insights using data but also to communicate them successfully to non-technical individuals. I’ve never had quite the fear of public speaking that many people do, but its certainly feels like a chore to me at times. However, anytime that I have delivered a presentation or speech in a successful or even stylish way it always feels good, especially when the information you’re delivering can create change and help others. So moving forward I want to make a stronger effort to develop my communication skills especially those involving communicating about technical information in a non-technical way or to a non-technical audience. I especially want to expand my skills related to data visualization specifically. There is something to be said for having the right tool for the right job, and sometimes natural language isn’t the right tool to communicate certain information. Great data visualization can communicate information in ways that not only state what actions to take but actually invoke a feeling to take action, and I suspect that will be a massively important communication skill to have as I push my career forward.

I also want to expand what I can do with Python, or in other words I want to be able to do more real-world things using python (e.g. harvest real data, UI, etc.). For example, I discovered recently that there is an entire Python library called pptx which allows you to develop a PowerPoint presentation like its TKinter. Meanwhile I’m copy and pasting matplotlib charts from my notebook to my slideshow like a caveman! My point really is that I can now do a lot of things within Python that can have value in the real world, but there aren’t a lot of real-world things (i.e., outside of the direct Python ecosystem) that I can affect with Python. I want to tinker with technology the way my dad would with the mower. As a data scientist I will need to be really good with a relatively small set of skills but I think it would only broaden my scope to by fairly good with a lot of different aspects of tech. For instance, If I were doing analysis for a house flipping company it would only enhance my insight if I made a hobby of my own home renovation (e.g. DIY smarthome tools).

I am really excited about not just the future my career and the field I’m entering but technology broadly. The future is increasingly electric, increasingly digital. Hopefully refining my mathematics, technical communication, and inter-disciplinary prowess will give me the best show at exploring all the cool new things ahead.