1. Can you give a brief education and career history and discuss how this history led you to your current position at IRS and in LB&I? Why/how did you end up in LB&I?

I started my software engineering career coding for DECT wireless handsets in C/C++ and Assembly. I did that for 3+ years before pursuing an MS degree.

I got introduced to statistics and software for data analysis during my MS thesis work in Bioengineering. After my MS I did a 2.5-year fellowship with alloy data at NETL DOE labs in Pittsburgh; this was my first non-academic Data Science (DS) project. I worked with another researcher & remote team on this project. On completing the fellowship, I took up consulting work with Cognizant Technology, which brought me to Florida with a project with Florida Power & Light. Smart meters, Greenplum (Postgres), R/Python time series features for classification were used in this work to flag suspicious meters tampered with for investigation in the field. Subsequently, I did a data engineering project for Frontier Communications (FTR) which involved creating tables and views using Teradata for customers that call into FTR’s call centers. I also did excel report automation using Visual Studio after Teradata project completion.

I had applied to the IRS during my work at Frontier and got selected in the Product Management (PM) team. Throughout the interview process and final call with the reporting manager and senior manager, I got a sense of what the team was tasked with. I am glad I joined and can share my thoughts as a new hire. I do feel there is fair participation of both employees and management, and that processes are in place for both personal development while achieving organizational objectives.

1. What are your major duties as a data scientist? (3). What is a typical day (or week or month if that makes more sense) in your job like?

Data pull using queries, cleanup and transform (additional compute) and creating reports are essential elements of the role. Most of my work at the PM team involves pulling data from different sources. A lot of time is spent looking at raw data, correcting for any errors, aggregating, and creating summary stats and creating charts/visualizations. I have also looked at applying clustering techniques to find similarities in the data.

My typical workday/week activities include: 1) researching and trying out new libraries/tools for analysis and visualization. Currently I am learning Power BI and Tableau. 2) Looking at data closely, applying methods to discern statistical patterns, and figuring out which variables are important for instance for forecasting and inferential statistics. 3) Making sure the data meets the assumptions dictated by the method used to create the models etc. is important.

1. What is something that doesn’t come up often in your work but is interesting when you encounter it?

I am fascinated with programming languages and how different languages have evolved to solve different issues with a predecessor. I try to keep up with the tools of the trade, be it Haskell, Go, Rust, etc. Different teams and developer communities are actively adopting these because they believe they have the capabilities to take products to production and facilitate rapid application development.

1. What kind of traits would make a person a good fit for a job like yours in the Service? Do you recommend this line of work? When it’s rewarding what makes it so? (7) What’s something you want employees to know about data scientist in LB&I, maybe ones they underutilize?

Every day in a DS or Analyst’s role is a different day. Being agnostic of technology/platform/vendor and being open to all of them are key factors in embracing change in technology and/or business requirements.

1. Do you have an activity you do in your life away from work to relieve stress or use another part of your brain?

I love to listen to new music artists, podcasts and to go for walks and bike rides. Music and exercise are known to engage parts of brain that are non-quantitative and are activities that relieve stress, but I am not a neurologist, so take this as-is.

1. What are you working toward for the future of data scientist in LB&I (goals, milestones)? What do you want employees to know about data science as it stands right now? How can they get involved if they’re currently on the sidelines?

My goals within the IRS are to get involved with as many data projects as possible. Learning from veterans in the service who know what to look for and what works. I share what I have learnt from the past and in my current work. I provide support for different tools within PM and learn why they are using different features of the underlying data. Some sources of the data have been deemed more reliable than others from prior experience. I participate actively in code sharing, peer reviews and presenting findings to executives and subject matter experts, to ensure quality and correctness of the work.

I learn from job aides, ITM, LinkedIn learning, blogs, Twitter etc. These are available to most within the service and are great resources to expand your skills and to become a DS/Analyst. I also learn by maintaining a sandbox with the latest environment (R, Haskell, Go, Rust in my personal laptop). Some of the latest versions of software are restricted for valid cybersecurity concerns. Updates to the personal sandbox enables me to track what’s going on each of them and learn despite not having access to cutting edge versions of software or tools at work.