1 Randell Boatright

After speaking to Randy I learned a lot about his role at New Ways Technology and the trajectory over his career. One of the key take-aways that Randy, and other professionals I've interviewed is to really learn SQL because you will spent a lot of time pulling data for analysis. Randy began his career studying computer science and moved into a role where ETL tasks take up the majority of his time. He explained to me the process of how data gets pulled from an oracle database, and the requirements that are specified prior to pulling the data. One of the things he mentioned specifically was to always be agile while learning new tools and tech stacks. He mentioned how over the past couple years they have been working to move the data from oracle to a cloud based service; AWS and databricks to store their data. Like others have mentioned, Randy spends a lot of his time aggregating and preparing data, than modeling. He also said that it is important to alwasy think about your projects in the context of the business purpose and be able to justify how your work supports the business functions.

2 Roylan Hernandez

Roylan began his career working in the data center at Blackrock. His role at that time was to maintain front end applications that related to key metrics within the data center. From there, Roy became friends with people in the info sec dept and when an open spot opened up they invited him to join the team. It was on this team that Roy became more familiar with data technologies, becoming familiar with Hadoop and Hive At first, his primary responsibilities were to bring in data from different sources and aggregate the data in dept dashboards. His team is currently working on automating a business process that reports anomalies, but on a monthly basis. He explained the issue with this is that by the time the right people see the anomaly, its already too late to make changes. This new process will run the report on a daily basis to better update management on key metrics. The advice Roy had to give was similar to what others have mentioned; make sure you know SQL really well. When asked about how much time includes modeling he said little, as a lot of the time is spent satisfying request for upper management. Another driving theme I have seen throughout my interviews is that companies, including Blackrock are moving away from on-site data storage to the cloud.

3 Mariam Aisset

Mariam began her career as an intern after completing a math degree at UPenn. She was offered a analytics position in the customer service team tracking power usage across the Lehigh Valley for PPL. One of the things that stood out for her when she began working was how a picture from someones life could be painted so easily with just power consumption data. She said that basically you could tell so much about a person based on their power usage. Somewhat recently, rf trackers for power consumption were installed which give dat points every 15 minutes and that this granularity really exposed people's behavioral patters. Marian then moved onto the data science team where most recently she was the owner of a project that used ensemble models to predict power consumption 7 days out. She explained the timeline it took to build the model, where the data was sourced from, and how the model was put into production. One of the biggest difficulties she faces on a daily basis is getting accurate data from third party vendors. One of the main vendors they get data from is a weather service, which are notoriously bad as predicting the weather. These inaccurate predictions from the the weather service influence her models in a negative way as she uses that data in her model. A lot of times she needs to defend her models to the business side of the company when they ask why the predictions are varying from the true value. When asked for advice to a person just starting in the field the first thing she mentioned was not to chase a title. A lot of times positions overlap and roles of data analyst and data scientist are often interchangeable.

4 Final Thoughts

The main take-away from my informational interviews was to be very well versed in SQL. In addition, a key theme across the interviews was that companies are transitioning from having data on premise to being stored in the cloud. Furthermore, all the interviewers mentioned to be prepared to not be building models all the time. A lot of time is spent determining what data is needed and how to obtain, clean, and organize it. And lastly, the work they do is in conjunction with business stakeholders, so communication across teams is important so that expectations are managed, and everyone involved agrees on a final outcome.