I’m fascinated by the analytics, data science, and visualizations that appear in programs like “David Attenborough: A Life on our Planet” and “Seven World’s, One Planet.” Using satellite images, aerial data, and drone data to do things like calculate forest density, determine resource availability, track animal migration, and visualize dispersion of dust from one continent to another (such as the Africa to South America discovery) is my dream job.

As such, I was intrigued when I found the Astraea website. While I enjoy the technical challenges of my current job, I’ve started wondering how I can get involved in the kinds of projects that affect the world we live in. I’ve considered returning to school for a PhD, but know that I’m more interested in the practical applications of data science than the theory behind it.

In my current position I get to apply machine learning to the cat-and-mouse game of fraud. I started as an analyst for about six months before moving to the machine learning team. My current position is Risk Modeling Scientist.

One of my first major projects on the machine learning team was to tackle the free text provided by customers to describe their property damage or loss. Nearly three years later, twenty-four of the features that I engineered are still major contributors to the models we currently have in production.

My most recent project is related to network theory. I’ve been working on a way to associate claims as a network to help adjusters see the kind of related fraud that’s perpetuated by organized fraud rings. In the course of this work, I was afforded the opportunity to be a featured speaker at the 2020 NODES conference hosted by Neo4j, one of the largest graph database companies. You can view the speech on Neo4j’s website at <https://neo4j.com/videos/24-not-all-visualizations-are-created-equal/>.

I’ve also been documenting the general principles that I’ve been working on in my Data Science Diaries blog at <https://julielinx.github.io/>. I’ve learned a lot in the course of writing the entries and while the early code and problem statements are messy, I still find myself referencing the early posts for the concept explanations and visualizations.

In a recent post on Graph Counts (<https://julielinx.github.io/blog/g06_global_counts/>) I referenced one of the first entries (Entry 5: <https://julielinx.github.io/blog/05_EDA/>) to illustrate the concept represented by Anscombe’s Quartet. Additionally, all of the posts and code can be reviewed on my GitHub page (<https://github.com/julielinx/datascience_diaries>) since I used that platform to host the website.

I would be interested in learning more about Astraea and the open Data Science position.