PUTTING PREDICTIVE DATA SCIENCE TO WORK AT PPL

3Qs with Tom Tyler

Tom Tyler is an Operations Analyst in the Data Analytics Group at PPL Electric and is certified as a Lean Six Sigma Black Belt. He has more than 20 years of experience working for large corporations in the electric utility and industrial gases sectors. He has led corrective process improvement initiatives in relation to supply chain management, mechanical and operational failures, inventory reduction, and electric power reliability among others.

1

Tom, tell us about your data-science journey at PPL.

I joined PPL about four years ago after 19 years at Air Products. At the time, I joined a team of two charged with putting data into action to drive process and performance improvement. It was a broad mandate, but a good one – we were encouraged to get our hands dirty. No preconceived notions or demands of what we had to accomplish, just that we needed to get going and show results.

As a Six Sigma Black Belt in Data Analytics, I was up for the challenge. My boss at the time brought in TROVE as a partner, and we just clicked. It was all about teamwork, collaboration, and seeing what we could do together – no egos, just a lot of smart people sensing a great opportunity and not afraid to start with a blank piece of paper.

The first thing we did together was tackle a relatively small-scale project, but it yielded some big learnings. We had a Transmission Maintenance Program (TMP) defects database in the Transmission Asset Structures group, and we had a lot of holes in it. TROVE jumped right in and did a lot of data sleuthing and wrangling to help us bring this database up to date so we could use it more effectively.

A key takeaway for me from this project was the need to have a good data set from which to build. Now, a lot of companies let the thought of incomplete data paralyze them. I'm here to say that by working with the right partner you can improve your data quickly, and then you have a much more effective base from which to do predictive analytics.

So, after our early TMP defect database success, we spread our wings a little – both literally and figuratively – working with TROVE on a project to score our transmission structures based on their risk of avian contacts and lightning strikes. Again, the biggest lessons learned were about the primacy of data in predictive data science. TROVE did an awesome job of tackling this project with us in a really creative way. They didn't settle for simply using the data attributes PPL already possesses regarding physical structures, such as pole height, cross-arm width, and material composition. Instead, their data scientists took the extra step of adding a dozen or so new geo-spatial attributes that really honed in on "where" our assets were located and "what" typified those environments, so that lines that were on a bare hillside scored differently for risk than those in a low-land environment surrounded by trees.

Together, we built out the data and then ran it through a machine-learning model TROVE built for the job, which enabled us to risk score each structure. I'll talk about the benefit of that project and approach in a minute, but let me just say we started knocking out project after project together within the Asset Maintenance Group, and people across our company started to take notice, including our President. Now we are being asked by many different business units to do data-science work for them. We are in demand! →

PUTTING PREDICTIVE DATA SCIENCE TO WORK AT PPL (CONT.)

You mentioned results you have gotten to date. What have you seen so far?

I can tell you that we have driven positive results for over a dozen use cases so far at PPL, spanning maintenance Compatible Units (CU) prediction – where data has helped us understand where competitive bids should fall and shaved percentages off our total spend – to inventory-stocking for emergency restorations, where we've leveraged data to adopt a leaner, cost-saving inventory strategy. We've also had success applying predictive data science to mortality curves and replacement timelines for our Transmission Line Structures, "danger tree" prediction, and the avian and lightning mitigation I mentioned earlier.

But let's double-click on avian and lightning mitigation for a second, because I think it's important to understand what "results" often mean for utilities when it comes to predictive data science. I come from a manufacturing background, where it is possible to adjust a machine setting and dramatically increase product output. But electric utilities aren't making widgets, we are delivering a service and looking for ways to do that in the safest, most reliable, and most cost-effective way possible.

So, it is clear to PPL that a data-driven approach to avian and lightning mitigation is the right way to go, but it's harder to put a dollar number on the benefit. As a result of our data science work with TROVE and applying advanced predictive techniques, we were able to spend $400,000 to mitigate **12 different lines** instead of spending that same amount to mitigate **three**.

We let the data tell us that we could do partial protection of those 12 lines, based on the risk scoring we had performed, instead of full protection of the three, because we learned that some of the areas across the three lines only had a 2% chance of encountering debilitating bird contact or lightning strikes. Only time will tell if we are right, but the data is telling us the new approach is the right approach, and we think it behooves us and our customers to listen.

Any parting tips for peer utilities and other industries as they look to build predictive data science into their businesses?

At PPL we really have embraced the idea that data is an asset to our business, and we've seen that play out across many use cases. What I've learned through firsthand experience is that:

1. **The quality of your data really matters.** You don't need to put your business on hold and take years to get it right, but you do need to make it better. Before you try to use it for prediction, find the holes in your data and fill them. If you can't do it yourself, work with a company like TROVE that can help you. It is so doable, but it absolutely needs to get done.

2. **Know what you want to prove.** We have found it very useful to start with a question – tied to a use case important to our business – and then use the data to answer or prove it. It's better to start with something you really want to know, for example, "What is the optimum height for my structures, so trees won't knock down the lines?," and use the data to determine that, than it is to look at the data and try tease out where to aim it. Data works best when aimed well.

3. **Listen to what the data says, not your gut.** This is a really tough one, especially for seasoned professionals. Predictive data science is all about following where the data leads you. I'm not against intuition, I just want to validate it with data!