Alison Clarke and Erika Heesen talk data
October, 2020
Transcript

Alison Clarke (A.C.)
Hey Erika, thanks for joining me today to talk about data. You and I get to talk about data a lot because we are both on the OPLA Research & Evaluation Committee, but today I thought we could talk about some of the fun and frustrating things about data that we have discovered over the years. I am the Coordinator of Performance Measures and Data at Brampton Library, so I am always looking at our data, and thinking about how to collect it, analyze it and share it. All data, all the time! Maybe you could talk a little bit about your role and how data comes into it.

Erika Heesen (E.H.)
Well, currently I'm on parental leave, as you know, but my regular job is that I'm the CEO and Chief Librarian of the Perth-Union Library in between Ottawa and Kingston. And generally speaking I do a little bit of everything. I wear many hats. We're a smaller library so that tends to be how we roll. But when it comes to data, generally I collect all the statistics, I present it to our library board, I also interpret that for the library board. So, "I'm showing you these numbers, what do these numbers mean". And often, that's also, "I see that number's down, why is it down?". So that's usually the question I need to answer, but often it's a variety of things from looking into a new program, looking into a new service, or seeing how we're doing overall in terms of our annual reports and that kind of thing.

A.C.
Very cool, thank you. So, I thought that we could start by talking about the fun side of data, because I know when I first tell people what I'm doing there's not necessarily an appreciation of how fun data can be. And some people are probably thinking "how can data be fun"? For example, I always tell people that pivot tables are fun, and they don't believe me, and then when I teach them how to use pivot tables, they completely agree that pivot tables are fun.

But what I really enjoy about data is finding the patterns in the data and the stories that the patterns represent. So for example I was doing some gate count analysis on our branches in Brampton, and instead of looking at the numbers I represented the data as a line graph. And it was very clear that the visual patterns of the gate count actually told the story of whether or not there were families and children in the catchment area. The branches with a lot of young families and children had a very well defined spike between 3 and 7 every weekday. They also had a big spike in July and a spike in March. And the branches where the demographic was older had much less of a spike or almost no spike at all. So, I didn't even really need to see the numbers, I just needed to see the pattern to understand the story of the demographic in that area.

E.H.
Well, one, I'm so excited about pivot tables, now that I know how to use them. And you

introduced me to pivot tables. I was really excited about it. My husband actually told me, he's like "I get that you're excited about pivot tables and that's cool and I understand why they're exciting." But maybe there's a little bit of nerdy happiness there. But finding the story behind the data, you described it as patterns, but certainly looking at it and going, "ok this is what the numbers are saying, but what are they actually saying, what does that mean?" So for example, I was looking at our data for payment of fines, and who's paying fines and who's not. And generally speaking we found that a lot of the fines don't get paid at all, and they're just outstanding, and eventually we delete them and so on. But the story that tells also is that when people have a certain amount of fines they often just don't come back. They don't come back to the library at all, so they're not paying them, but they're also not using our services any more. So that becomes a story we tell if we're looking at going fine free. I know Brampton has done that with children's fines. So, that it also informs policy changes that we make. So why are we doing what we're doing? We have the story, and then we have it affect what we do.

A.C.
Absolutely, evidence based decision making for sure. That is very cool. One of the other things I think people should know, and maybe they already know this, but mistakes can totally happen with data. Sometimes It's hard to get clean data. People will interpret questions in a completely different way than you expect. Even if you've tested it first, sometimes you'll get a response and think "hmm, I never thought this question could be interpreted in that way." And sometimes you just miscalculate. So, I was once convinced there was a baby boom in Brampton for a few weeks, because I missed some data in analysis of our summer learning numbers. And it wasn't until a discussion with a colleague who thought it was pretty suspect, that I went back and realized I'd completely missed some rows of data. So sadly, no baby boom in Brampton.

E.H.
I always go back to the whole GIGO principle - Garbage In, Garbage Out - that's where I see issues where data can be really frustrating, or it's not helping you the way you would expect. In a previous role I'd put together a survey that we were doing about patrons and one of the questions was around hours and the hours that we were open. We were trying to find out the hours they prefer, when people wanted to come in and that kind of thing. The way I phrased the question, and the way I had people answer the question, and it was a paper survey, that then a number of different branch staff were actually inputting into a spreadsheet. The way I phrased it and the way it got input meant the data was pretty much unusable. I was really sad - we had a few hundred responses which was great to have, but then it required a huge amount of time to just manually clean the data, and make it something that we could actually use and analyze. And if I'd thought about that end result when I was creating the survey, then that wouldn't have been the case. You have to have that; "this is what we want to know"; "this is what we're trying to get at"; "how do I make sure I"m getting the information I need to actually analyze it properly". And I find that with the program statistics, that's part as well. The annual survey for public libraries requires you to say "this is how many

programs", and "this is how many attendees" that are within set categories. And we hadn't been recording the categories of the programs, so again that meant a huge amount of time for me going through and manually coding all of the programs to say this is what it was, to be able to add it all up and figure it out. Whereas if we'd designed it from the beginning to account for that then it would have been much more efficient.

A.C.
I completely understand. That was one of the things I redid at Brampton was to add in some of that element, the ministry data, and things like that, into the program data collection from that we use, so that there didn't have to be the manual at the end of the year. Because people would forget and you might change your mind, so that was one of the things I did and that was super helpful because it saved a lot of time later.

E.H.
And then you can just use a pivot table to add it all up!

A.C.
I know right! And then we're right back to the fun! It's awesome, love the pivot table! Hopefully we've encouraged some people to learn pivot tables if they have not already done so. Because they are fun. Thanks so much for joining me today Erika. I hope we've been able to make data a little bit more accessible for anyone who's interested. And if you already love data then I'm sure you can relate to at least one of our stories. The OPLA Research & Evaluation Committee is always happy to talk about data, so please feel free to reach out with any questions if you want to talk about data. Thanks Erika!