

>> SPEAKER: Good afternoon, everybody. My name is Kathie Mason. I am the collections librarian at Eastern Michigan University, and today, I'm going to share my adventures with understanding what I don't know. I will, from the beginning, issue this caveat that I did not put my contact information on the slide, so I am actually going to quickly put it in the chat for anybody who wishes to contact me after the session. I also put our hashtag that we came up with here, which was growyourknow. We don't have much time together, so I'd like to get started sooner rather than later. Um, I have some lofty goals for this session, however, I do want to clarify that we're not going to actually identify deficiencies, define key steps, and make a plan. We're going to talk about how to accomplish these things. I have very simple ground rules when I present, which is ask me to speak up, slow down, clarify what I just said, repeat myself. I'm not, even though we don't have a lot of time today, I don't have a problem with, um, saying something twice, if it hasn't come across clearly. As for why I picked this topic and why I submitted this proposal, it's a combination of a selfless desire to share my experience and hopefully save somebody a little bit of pain and suffering if they, too, want to know more about how to maximize their ability to use statistics in an effective way, and I also have some selfish motivations in that I wanted to motivate myself to finally do what I said I wanted to do all along and also learn a new tool for presenting. I've never presented virtually before, so I thought that to be a really meaningful opportunity, so I'm very excited to really dig into this topic. I learned that, um, when I was really thinking about this, one of my biggest problems that I had was I never considered my understanding of statistics. I thought, hey, I took a class, I know something, I know what a mean is, I know what a mode, median, all those terms that you kind of hear as the general beginning statistical concepts, you know, I'm fine, I'll figure it out, if I have a question, I'll ask the question, I'll look at Google, you know, I was figuring out little things bit by bit, but I truly wasn't growing my understanding of statistical analysis, and I was often struggling.

I'd actually like to know from my audience, how about you? How do you characterize your current understanding of statistical concepts and methods? Are you an apprentice, do you consider yourself a journeyman, or are you a master? I'll admit, I'm a low-level journeyman. I do know a bit about statistics, I'm getting to know some of the lingo, but I'm still struggling, and I have plenty of room to grow. Well, I want to talk to these three masters, and I'm not just joking. I think, actually, what I'm seeing here is quite a comment, and I acknowledge this is a group effort, so if no one person wants to answer this, perhaps the library's understanding of statistical concepts and methods might be a way to answer this. It looks like we're doing pretty good here. I think this actually may be pretty representative of what I've noticed in library land, is we know a little bit, and some of us even know a little more, but I have encountered very few people who truly, um, are comfortable and have mastered data analysis and analytics, and I think that's okay. I'm trying to move on here. Just a second. I learned a few lessons while I was on this journey, and the first one, which is really such commonsense, it's almost sad, is the realization that I need to have a plan. I need a systematic strategy for improving my skills and understanding. Sure, it was good enough to ask that question when it came up or find that process when need be. Still, though, I found myself over and over coming into roadblocks, unnecessarily wasting time on something that shouldn't take so long. When I sat down and thought about this, I realized what I really wanted to know is I need to improve my vocabulary. When I speak

about statistics, I want to do so in an intelligent manner and not misrepresent what I'm trying to say. I also wanted to ease my pain. Like I said, um, before, I would figure it out, but I spent so much time, um, fumbling around, in Q & As and forums and heading down a path to realize it was the wrong path and starting over and over again.

Most importantly, though, I want my analyses to count. In the past, I would gather some information and throw together a couple of charts that looked maybe interesting to somebody, put together some bulleted lists of factoids about the information I had gathered, but it was lacking that oomph. You know, I couldn't say, hey, this is statistically significant, because to be honest, I wasn't quite sure how to calculate statistical significance, and I think that's okay, again. Another part of this having a plan lesson was I need to figure out where I was at. You know, I said, hey, as I said earlier, I know this, I know how to calculate a median, what else do I need to know? You know, Excel will give me the chart that I need. So, what I did is I just started seeing what available resources were out there, and there were a few online tutorials and short quizzes, and it didn't take me long to realize I could start one step past that mean/median level. Sure, there may be some refreshers and things I already knew, but for me, personally, I was best off starting almost at the beginning. I started exploring resources. I checked out books from the library, requested some materials, I started looking at informational websites, I started playing little statistical games. One thing I found that was really helpful, K-12 AP statistics tutorials were intelligent enough to speak to what I needed to know, but not so advanced I found myself being confused. I found those especially helpful. Another thing that I did in exploring my resources, um, my previous campus had a statistical consulting center, and I made several appointments to meet with them and talk through issues I was having about things that I was struggling to understand, and while that was somewhat helpful, I often left those meetings thinking, I have no idea what I'm doing.

So, I proceeded on and developed a plan, and my initial plan was I'm going to take advantage of whatever's available to me, and that led me to some poor outcomes. You know, it was random explorations. I would see a pre-conference for, I remember one in particular where I was learning about pivot tables, and I wasn't to that point yet, so I was able to manually regurgitate the instructions on the screen, but then when I got home, I didn't truly quite understand what I was trying to do. In another case, I was asked to find the P value to see, um, if what I was looking at had any significance, and I had no idea what that meant, which speaking of P value, I actually want to ask another question. Suppose your ILL office offers 24-hour turnaround document delivery. Your experience and other feedback suggest a longer wait for document scans than promised. You gather details about document delivery requests and calculate the P value of your data sample. The P value equals 0.004. What does this mean? I think there should be a quiz pop-up. I'm not seeing the quiz pop-up, so if you want to answer in the chat, that'll be fine as well. Oh. I'll be honest, um, when I first started working with this, I could not answer this question correctly. I'll leave it open for a few more seconds, and as I said before about the group situation, feel free to answer as a group. Power of many minds, right? Let's see here. Yes, what is exactly how I answered this, and I'm not yessing because I'm happy everybody feels the same way I did, but this is definitely how I felt the first time somebody asked me what the P value of the data that I was presenting was, because I had no idea. Well, I'm not going to go into a huge description of what P value is and how it's calculated.

Simply put, this is a statistically significant result, and if the 24-hour turnaround time is inaccurate, the lower the P value, the stronger the evidence for statistical significance. There are calculators to help you find this out, and there's even, like, a line that you can chart in Excel, you know, there's definitely, um, tools to help make this easier. Now, at least on this survey and as I mentioned, we're not alone in not really understanding what this means, and that's okay, because we can learn, and I brought this up because this is one of these situations where I was running before I could walk. When I started working with the P value, um, I was up higher in the, um, in the taxonomy here, and this image is, um, well, a variation of Bloom's taxonomy of educational objectives. It's interesting, I found this, I was aware of this, but I actually found this after I started trying to figure out how do I manage this crazy world of statistics and analytics, and on the left is my list of what I came up with how I shouldn't navigate this. I need to learn the vocabulary and the basic concepts before I move into applications of those concepts, and then I can learn some more advanced concepts before I move into more applications, and then I can worry about developing some software aptitude, whether Excel or other programs. You can do enough through tutorials and, you know, Google Help, but I don't need to spend, I was spending too much time with the software, if you will, too much time on trying to figure out the big things before I understood the little things. In essence, that was the cause of my woes, too much time without mastery of the basics. Where it became crystal clear to me, I was at an SPS workshop last year, and I was so excited to learn more about SPS, I had heard about it, I knew what it was, I also had never seen it before, I had not worked in an institution that had it available, and I thought, hey, I'm going to learn this tool, and I'm going to learn how to, um, do better research, if you will. I even had a research project in mind, and I showed up at the workshop, and there were nice, um, how-to documents provided, which were very helpful, and I discovered that I didn't really understand what people were talking about half the time, and, so, I learned the basics about SPS and, you know, kind of like muscle memory, I learned to click here and then click there to perform basic operations, but I didn't really understand what I was doing or why I was doing it, so this was a real a-ha moment for me, that perhaps my scattered, chaotic, you know, I'm going to take advantage of the available resources whenever they're available method of planning was really a poor idea, and that's okay, again. I'll say I figured it out.

I do have another question. I want to shift gears a little bit into output of research. This is a real life example. I was looking at book usage, print and e-book usage, and I had, um, the information split out by the type of format and call numbers. Which chart do you think would be most informative? A pie graph, a line with markers chart, a stacked column chart, or a three-D area chart. I was suspecting this was going to be a fairly easy one. I'm going to give it a few more seconds here. Okay, let's see. In this case, you actually could use different options depending on exactly how you want to, if you want to do one chart or multiple charts. Um, the answer that I went with was C, the stacked column chart. You could use a pie graph, if you wanted to show, um, the proportion of e-book versus print book usage on the whole level or within each individual column number region. You could use a line chart. Um, it would not be possible to show the proportion of both e-book and print book usage in that manner. I tried to use a three-D area chart, and it just had this really cool, awesome look, but the more I looked at it, the less I understood what I was trying to say with the chart. In the end, the stacked column chart

conveyed the information that I wanted to convey, it showed me, um, how much usage in each area as a total, it also showed me which ratio of that usage was print versus e-book usage. Another thing I could have done was normalize the data a little bit and look at it as a percentage of the whole. You know, what percentage of the D call number e-books circulate at least one time, but that's a whole other conversation for another time. I had one final lesson that I learned during this exploration that I'm still learning. Use it. Take advantage of the resources around you, whether it's the books, the experts in the building, locally. Your consultant center, even if you don't work in the university, they're often willing to help the community at no or little charge. For that matter, you could get an intern, a student, a graduate student from a local university who might be willing to work with you on planning and implementing a statistical analysis project of some sort. I've also learned that you have to use the skills. Practice really does make perfect. Like the P value, I learned something about it a couple years ago and never used it again, and when I realized I wanted to know more, I saw that I don't know enough. I bring up family projects here because I think, you know, you don't have to apply these skills only in a work environment for them to be, you know, be something useful to you. A little nerdy thing I've always wanted to do is do a gantt chart of my family's immigration to the United States. I think that would be a really cool visual thing to see.

With that said, I would like to ask one last question; what resources do you use when working with statistics? That's actually going to partway into the Q & A. We can also, um, move this into general chat as well, because I do have one last slide of information I'd like to share with everybody, and I agree with Excel, Excel is my primary tool that I use at this time. I plan on using SPS in the near future. I put two tools that you could use, if you're interested. The first one is I created a very simple, brief quiz to get you thinking about your level of understanding of statistics and what you know and don't know. I also create a resource list of websites and books that I use and still use to help me understand better statistics. I think we have about one minute for questions, if anybody has any, and I will also be in the forum as well.

>> SPEAKER: We are actually out of time. So, I'd like to, sorry about that, Kathie.

>> SPEAKER: I am too.

>> SPEAKER: But this is the reason why we have the online forum. So, I'm going to go ahead and encourage everyone to please engage with Kathie on your questions and her presentation in the online forum. Um, we're going to move into our first break now, and we just have a couple announcements as we head into that.