[00:00:00] It's time for Meaningful insides every researchers delighted stay tonight.

Kasha Ely: Hello, and welcome to Data Night with the Odum Institute. I'm your host, Kasha Ely, and joining me today is Dr. Eric Monson, data visualization specialist with the Duke University Libraries' data and visualizations services. Hi Eric!

Eric Monson: HI Kasha, it's good to be here. Thanks.

Kasha Ely: Thanks, glad to have you. Thanks for joining us. Later this year, you will be returning to teach a course at the annual Data Matters data science short course series, which is a week-long event sponsored by the National Consortium for Data Science, RENCI or the Renaissance Computing Institute at UNC Chapel Hill and Odum. This year, you'll be teaching class that you have taught before, Introduction to Effective Information Visualization. Can you just give us a quick overview of what you'll be covering in your course?

Eric Monson: Sure, I'd be happy to. Each day, I kind of start [00:01:00] with some basic principles. So, like on the first day, I start with some principles for effective data visualization. The second day starts off with some things about effective graphic design and something that goes, maybe beyond data, but is about visual communication and then both days are filled with hands-on exercises. We do a little bit with a program called OpenRefine, which helps you clean and transform data, kind of get it ready for visualization, and then a decent amount of time, the two days talking about and practicing with a program called Tableau, which was originally invented as a business intelligence tool, but now people from lots of different disciplines use it because it's really great for exploring data and visualizing it and then even making interactive dashboards with them. So, we cover a little bit of that, and then I also get into a bit of some web-based tools that were either designed for kind [00:02:00] of data journalism or do some somewhat unusual visualization types that you wouldn't find in something that's more standard like Tableau, and then sprinkled throughout, I give people practice doing the process of critiquing existing visualizations, so we might take something that's been published in a newspaper, or I might have gotten an email or something and go through it together because I think that practice of looking, trying to use the vocabulary we've built up through some of my talks to see what's good about it - try to figure out things that we could use in our own work and then what might not be as effective as it could be, what some ideas would be for improvement, because that's what we need to go through with our own work and with each other to help each other get better at visualizing. Really, that sort of feedback is the way you get better.

Kasha Ely: Absolutely. Thank you. Why do you think the data visualization courses at Data Matters [00:03:00] are among the most well-attended?

Eric Monson: I think there's a few different reasons. One, I think we just react to visuals; you know, everybody would like to make something that looks good, and it attracts attention, you know, for our work. I think visualization is something that people from almost every field, if they're dealing with data or even if they're just dealing with ideas, you know, I don't cover this as much in this course, but we're all trying to communicate, you know, different ideas, different patterns in data, no matter what our subject area. And so, visualization really cuts across disciplines. And then finally, I think visualization can help us in so many phases of dealing with data from the very beginning where it can help us spot errors or gaps in our data. It might help us figure out what types of analysis might be appropriate. And then at the end what, you know, we usually think about [00:04:00] is visualization is that communication phase, where we know what patterns are there, and we want to tell others about them. Visualization is a great tool for helping communicate those patterns. Just like almost all types of visuals, we kind of need to perform well to do our jobs. We need to be able to communicate results to our colleagues, our boss, audiences out there, to tell them about our work, but we're not really trained often in visualizing or visual communication, and I, so, I think people are really starving to get better at these things because it matters for getting our messages out or getting promoted or getting that grant, you know. So, I think it ends up being kind of important to people.

Kasha Ely: Definitely, and those are some great points. What are some current trends that you see in data visualization?

Eric Monson: I think, you know, I don't keep up with every new thing that's happening, partly, because I think the very basics could get you so far, [00:05:00] but I've noticed that there's a lot more tools that make it a little bit easier to visualize to hook up to data, and, you know, and look at the different patterns, and the defaults are better. This is one thing that I love is like if you even if you just look at Excel, over time, the default colors and the layouts and everything's have gotten so much better, and I think it's programs like Tableau frankly that are pushing this because it was built with a lot of good visualization principles and defaults built in, and so now, I think that's challenging other companies and other products to not only make it easier for a wider audience to visualize, but to do it better right off the bat, and not have to change so many defaults to have something that looks decent and communicates well.

Kasha Ely: Makes sense. When you see data visualization and popular media, what do you notice first?

Eric Monson: I think, yeah, I kind of noticed whether it grabs my attention in a good way, you know? [00:06:00] If, like, I think if the story is interesting, if it pulls me in and I think this is a really important part about visualization that it's really also about storytelling. It's not just about showing the data. It's about that visualization having a purpose and the purpose really pulling us into the story that's being told. And so, when I see visualizations in popular media, I will notice if it draws me in. I will notice if it's garish or ugly or off-putting, or if it's pleasant and well-designed. And, and really helps tell me a story and those things affect me a lot to the point where if it's not good looking in a kind of effective, visualization way, I'll just flip on by and I really won't engage with it.

Kasha Ely: Totally get that. I think a lot of people can relate to that.

Eric Monson: Yeah, and it's, you know, it's hard to... [00:07:00] none of us want that to happen with our own work. And so that's also why I think it's good to be practicing this stuff and learning some principles of how you might make your visualizations better is because, you know, whether it's at a poster session and you want people to walk up and engage with you, whether they see your article or you're, you know, out there and you want them to stop and notice it, visualizations can both attract that attention and then help hold people and teach them something and tell that story. So, none of us want to be that one that gets flipped by and people don't engage with it.

Kasha Ely: Who do you think should take this course, in your opinion?

Eric Monson: I think it's probably best for people that are very beginners or just have a little bit of experience with visualization and want to quickly get a lot further. So, I think some more advanced people could get things out of the, the talks about the principles and even the things I cover about Tableau, [00:08:00] you know, even if you've used Tableau before, I'm going to go through some pieces in a systematic way. So, it may fill in some gaps for you, but I think the people maybe who should not take it, are people that use Tableau every day, are super advanced, this is probably not going to be designed for you. It's designed more for the beginner who maybe does some data analysis or deals with data or would like to sort of move in this direction and be able to do this type of thing as part of their job. It'll give them a quick start into both some principles and some tools that I think, will, I think, empower them. You know, I've had good reactions in the past of people that said, it, it got them kick started really quickly, and now, they felt like they could take on a little project of their own or start moving in this direction and use some of these tools for their day-to-day work when maybe they were a little hesitant before. So, I've really tried to gear it towards the beginners but even some not true beginners will get something out of it, I think. [00:09:01]

Kasha Ely: Nice, and you mentioned Tableau and what was the other software, main software?

Eric Monson: OpenRefine was another one that I'll cover just for a little bit, but it's actually not a visualization tool; it's just a data cleaning and transformation tool, but people - I always hesitate yeah that's the thing, like, I hesitate to put it in with the nice start going through it. I'm like, I love this. I love it so much; I can't take it out and people - even though it's like the second thing of the two days. So, I cover some visualization principles and then I go into this, and everybody loves it because it's just super helpful to be able to look at the spreadsheet and see if there are mistakes or help fix those. Or if you know some visualization tools need the data in a certain form to work, and so it's just so helpful in dealing with the data and really, that's a huge part of visualization is trying to get the data ready, and so realistically I think it's a, [00:10:01] both a fun and a super useful tool, but then the rest of the time I'll spend on Tableau or there's something called RAWGraphs which does some things like alluvial diagrams, which can show relationships between different categories and numbers. I go into Datawrapper a little bit which is more of a journalism focused tool, but it in that it helps you kind of quickly get to a visualization that has good titles and labels and data sources cited and kind of gets you that point of telling a story really quickly because that's what journalists need to do.

Kasha Ely: That's really cool, and I'm definitely going to be a bookmarking that. Are all of these tools open to the public?

Eric Monson: They're open in some sense, at least. So, OpenRefine is free and open. It's an open-source program. RAWGraphs is open source and free. Datawrapper is free to use for the basic usage. I think they have some other tiers if you are a company using [00:11:01] it or different, want to save different amounts. Tableau is in a, is in a funny spot because it is free for students. It's free for some people that are doing nonprofit academic research. There is a free version called Tableau Public that people can use for the whole course and can use for quite a bit of work, but it has some certain restrictions such that it's free. You can only save your data up to Tableau's public server. It's up to kind of cloud thing. You can't just save files on your desktop, you know. So, there are some limitations to it, but nobody would have to spend any money on software to take the course or to do the things that we're going to talk about. It's just that Tableau is a bit of an edge because the real software, if you really have to pay for it, is somewhat expensive, and so, I hesitate teaching people that but to me, it's the most [00:12:01] flexible and easy tool for really exploring what form of the data, I don't know, gets, like, shows some interesting patterns. If I don't know if a bar chart or a map is really going to reveal the patterns that I want to see, or I want to do a scatter plot or, you know, there's just a lot of different things, built into Tableau, and it lets me really quickly and easily switch between those different forms of visualization. And so, I think it gets across certain principles and practices. So, even if you don't use it, you're going to learn something from that, but I do hesitate since the real product is not free. It's a very capable program, but I think it's worth learning.

Kasha Ely: Nice, thank you. What first got you into data viz? I was looking at your Data Matters bio, and it mentioned that you had a PhD in Applied Physics.

Eric Monson: Yeah, so it's, I think it's a combination of things. [00:13:01] I think I've always been a very visually oriented person. I've done a fair amount of photography, a little bit of art. And so, I had that kind of bent already. Then, a lot of my research involved either imaging - So, I did lots of fluorescence, worked with fluorescence microscopes and kind of actually took pictures of things, or I did different types of experiments or studies that you needed to visualize the data to communicate it. And so, kind of with my design background and my data interest and research, you know, research experiences, I just kept wanting to get better at it. So, it was something that I liked doing. When I first came to Duke, I was doing a postdoc in physics, doing some biology related stuff and after a few years, my advisor, she left to get her, she left physics to get her PhD in Clinical Psychology, and so suddenly I needed a job and luckily, there [00:14:01] was a woman on campus that spanned engineering, computer science, and art history, and she needed someone that could work on grant-funded projects doing data visualization, and she had one that was in computational immunology, which she was like oh, it's an ex-physicist like you are. I think you'll get along, and so even though I wasn't completely qualified, she hired me, she knew I could learn on the job, so I spent seven or eight years doing visualization work on grant-funded projects and then about - almost 6 years ago, five and a half years ago, this position opened up in the library, and I have always loved doing research support more than doing my own research, and this just was a great opportunity to be involved with visualization, help other people, whether it is from art history or the med side or engineering or, you know, the school of the environment, all over campus. People come [00:15:01] to me and they just need a little bit of help getting further with their data, their visualization, their communication, and so I got involved with it through my experiences and just the love of it, but I think I still enjoy the fact that I get to cross across a bunch of disciplines, and it kind of reflects the fact that visualization really is a fairly universal need and language in a lot of lot of topics that we cover.

Kasha Ely: No, that's a great point, and like you said, if you're doing a poster or even just throwing stuff into a paper, there are so many applications that, until you stop to think about it, you really don't realize how prevalent it is.

Eric Monson: Yeah, I really, I mean we are asked to communicate visually a lot. Yeah, whether it's a paper, it's a presentation, it’s a poster, it's a newspaper article and even when we're just [00:16:01] explaining things to people, we might sketch a little idea or you know, we'll talk even visually. And so, I think this is that aspect where it's a. It's a very effective way to communicate. If we can do it, we are asked to do it a lot, and it really matters how well we do, both in terms of our own careers and in terms of how well people understand our stories, and so, and the only way you can get better at it is practice and to learn some principles and get some feedback on that. And so, for people that want to be better at that, for people who enjoy that process of trying to communicate visually, I think data visualization is a great place to sit because it can come in really handy.

Kasha Ely: Definitely. What would you tell people who might be on the fence about attending Data Matters?

Eric Monson: What - one of the things that I think about sometimes is that I think some people in it up attending [00:17:01] something like this and really engaging with it because it's somehow easier to take a day or two days or a week off of work and tell your people, I am taking this course; I need to be away from my stuff. If you're taking a one-hour webinar, at some point, you usually end up saying at the last minute, oh, I'm busy. I'm going to, you know, I'll just, I'll hope they send the recording. With this, it blocks off a certain amount of time. And it, I think a lot of the courses can really help give you boost in these areas, so very quickly, within a couple days - a day, depending on the course, it'll help you jump your skills. It will help you jump your conceptual levels, and it's time that you can really dedicate to this and it’s a live experience, okay? You can watch recordings, you can watch things on YouTube, but here there's not going to be recordings, the instructors are right there, you can ask them questions, you can get personal help and so these days, I [00:18:01] think that's quite valuable, and for people who learn well in that situation, you know, I think Data Matters is a great opportunity. And, you know, from past experience of participants giving feedback, people really seemed to enjoy it a lot.

Kasha Ely: They do. We've been looking over some of the responses from last year's attendees, and it's pretty awesome how many people get so much out of it?

Eric Monson: Yeah, you know, I think it's very self-selected. The people that show up there are, you know, they have something invested in it and I don't know. Yeah, people really seemed to enjoy it and people from a lot of different backgrounds seem to get something out of it, and that's both impressive and fun. Like, I think, I think the, the short course series as a whole is structured really well. Everybody behind it has a lot of experience at this point trying to make this event really effective and, you know, they have [00:19:01] gathered instructors that are good at doing this, and they only bring us back when we're good, you know, and people get dropped if they're not good and then they bring in other people. So, I think, you know, you're coming to an event that this isn't the first one, that it's been going on for quite a number of years, and they've really honed it and tuned it, such that I think it works quite well and is a lot of fun.

Kasha Ely: Agreed. Well, thank you so much for joining us on this episode, this has been great.

Eric Monson: Yeah, you're welcome. Thanks a lot for the invitation. It's been fun to talk to you about this stuff.

Kasha Ely: And to everyone listening, thank you for joining and keep an eye out for more information about this year's Data Matters. Until next time, stay safe and well.