

Interviewer: Let's turn this on. So basically today we're just going to do some reflecting on the workshops after about three months so you know what you thought about it, your kind of experience of the workshop. Then we'll take a look at your workflow to see if you're doing anything different. You know, changes you've made or changes you maybe plan on making. And then we'll just wrap it up with that same checklist that we did. Yeah. So I have an hour booked, but so far it hasn't taken that long. So let's just get started with thinking back to the workshops. So this is early March and you were in our R workshop, right?

Researcher: Yes.

Interviewer: So what first of all, were you able to attend both of the days?

Researcher: Ah, yes.

Interviewer: Okay. And what did you think of the workshop?

Researcher: It was really interesting. So I was a little hesitant because it started off with lot of Unix and Git and I wasn't actually interested in that, but then I realized that it is going to help me somewhere down the line, and especially GitHub because I usually use the packages I download it and just run it. Usually GUI based but didn't know that what is possible with GitHub. It was really interesting. And also the version control I liked that it was all the time, I cannot remember what version it is, you know, it doesn't lead I always get confused and I have to go back. That's something that I think I would use. But right now I'm not using it but I think that's a skill that should help me.

Interviewer: Yeah.

Researcher: As for the R, it was basic commands, which is useful, but I had already done, like there is a small library, so I had done a little bit of that. It was kind of like doing it again, but it was interesting.

Interviewer: Yeah. So what do you think was your biggest takeaway from the course?

Researcher: You know, it was like a formal introduction to coding which I had never had. So it was nice to know what is possible. And you know, I have always seen people do the command line on the computers and I'm like, ah, that's super geeky and now I know how to do it and I have like telling friends, "Yeah I have done Unix," and my sister's a programmer so she's like, "That's very basic."

Interviewer: No, that's very cool.

Researcher: I like it. I think I liked it a little bit more now and I understand there are a lot of possibilities and that's something they want to get trained on. So I have been actually doing a little bit of other workshops too.

Interviewer: Oh cool.

Researcher: What I can find like, and I'll have signed up for a single cell workshop, July.

Interviewer: The single cell RNA sequence?

Researcher: Yeah.

Interviewer: Great.

Researcher: The UC Davis one.

Interviewer: Oh, oh neat. Very cool.

Researcher: It's very expensive, but I think it would be worth it. They also had a single cell symposium that Gladstone arranged.

Interviewer: Oh cool.

Researcher: So one day I didn't get the hands on workshop, but I got the notes from it. So I have to play around and figure out what I can actually do. The training database and see.

Interviewer: And was there anything that you think we could have done differently to make your learning experience better at the workshop?

Researcher: I think really well organized and I was happy to see that some of the postdocs who just learned it just the way I did was there to coach us. I was a little bit upset especially because I didn't know there was going to be Unix and it almost like half of the two days. But if you had put it forth and said that, "Okay this is going to be R plus Unix," I would have taken it more nicely. I think it's good, it's just unexpected.

Interviewer: Yeah, yeah. We'll have to think about our messaging a little bit. All right, so thinking about the-

Researcher: And I also really liked, I can't remember the name of the person who came in who was a assistant professor in some other institute.

Interviewer: Was that [Name]?

Researcher: I think so.

Interviewer: Yeah he works at [University].

Researcher: Uh-huh.

Interviewer: Yeah, yeah.

Researcher: He's really good and he was so involved and it made sense to me actually and stuff.

Interviewer: That's great. Yeah, I like him a lot. We tried to get him to come over here and teach. Great. So thinking about this kind of workflow that we talked about, I think back in January actually. So you know, one of the things we were wondering about is this is kind of the workflow that you detailed where you're taking a lot of data, you're working with a data analyst who would actually do a lot of the analysis and the data merging. And then it sounded like you were using some other data packages with GUIs to do your own kind of processing and analysis. So first of all, is there anything that's different about your workflow now?

Researcher: Not really. So when I met you last time the data analysts was going away, she ended up being staying longer. It has slowed down things for me. Because I know I can depend on her for some more times. But it's good in a way that there. But I have been doing the same exact thing. We've worked on little bit more of the data using the merge data. So she has sent it to the Excel and I go over it and ask her to do kind of analysis. And you know, I have learned a little bit of more about regression and what kind of data modeling. It's more statistics than computation but yeah, it has been nice.

Interviewer: So you think because that your data analyst is-

Researcher: Staying.

Interviewer: Stayed, kind of it takes the pressure off a little bit or?

Researcher: Because last time when I met you, I was like, I really need this. Now I'm like, I have time, I have one more year. But I think some of the things that I'm interested in, like RNA sequence, sequence data which is more high dimensional than what I'm doing right now. She has no experience in that, so I have to do it and at least get the data structure and know what exactly I want as an output and then move on. So that's why I have been doing all these courses.

Interviewer: Yeah. So let's talk about kind of future plans. So is there anything that after taking this course you, or think you might do differently with future kind of workflows or...

Researcher: So I feel more comfortable with R now. I'm not as anxious so I think that's also why I have been comfortable signing up for courses are just like more data intensive. So that's definitely useful. And I want to attend more of those. So you had a three day workshop soon after the March one? Yeah, I thought maybe about attending that kind of a workshop once more so that I have a refresher once in a while.

Interviewer: Yeah. And is there anything for here that you would use R for, you could see doing different in the future?

Researcher: I think mostly we use R for almost everything. Except for this data, my data that I don't have the software access, so.

Interviewer: Do you want to draw in anywhere that you would use R that maybe isn't here now? That's right now.

Researcher: I think it was kind of like almost everything I do is here. I do use a new package, which is actually, this is the same kit, but even within that there is one thing called R Phenograph. I'm doing it a lot now and it's also R based and the data generated actually gets into a CSV file and then I analyze it on R, and so that's what basically heat maps. The heat maps and TSNE analysis. Just the clustering thing.

Interviewer: So those are things that you're doing now with this R?

Researcher: Yes.

Interviewer: Okay. That you didn't know how to do before.

Researcher: Yeah.

Interviewer: Very cool.

Researcher: Actually R is really useful. You just click through it and say like, Yeah.

Interviewer: And then, so we kind of touched on these, the red ones, which are the pain points, it sounds like maybe those are still not really pain points or they've been pushed off.

Researcher: Yes.

Interviewer: And then what about we have LaTeX here was circled as a pain point.

Researcher: Yes. I did a couple of graphs on it, but my advisor feels it's too complicated for him to edit it. So I have moved back to Word, but I think for submission, which I'm getting ready to in a week, I'll probably do LaTeX. I'm comfortable the platform I have a document which I can copy paste and you know, get going and I like the pdf files in it. I don't have to play around with the figures and the tables look awesome.

Interviewer: Nice. So that's something where...

Researcher: I'm currently using, but I didn't use it the past two months because my boss who does the...

Interviewer: So they don't really like the way it works or interacting with that?

Researcher: I think he's scared to actually download the LaTeX or it's something new and it's difficult to change.

Interviewer: Yeah. That definitely happens.

Researcher: I like they have the bibliography generation too. So that's something I'm looking forward to, not manually having to edit things that are in Endnote.

Interviewer: Yeah, yeah. That can save a lot of time. Let me just see what else we talked about back in the day. So I don't know if we talked about this, you mentioned back in January maybe being interested in doing some kind of machine learning stuff further down the line. Is that something that you are still thinking about?

Researcher: Yes, just didn't get going on it, but I haven't actually moved forward identifying courses or anything.

Interviewer: But that's something you think is an R project down the road, maybe?

Researcher: Yes.

Interviewer: Sounds good.

Researcher: Especially one of my aims of a grant that I have that's to actually identify clinical phenotypes, clinical endotypes based on the one of the phenotypes I have, that could require machine learning. So I'm interested.

Interviewer: Yeah, that sounds exciting.

Researcher: I feel like I'm moving to more towards computation from a molecular biologist.

Interviewer: I think a lot of people are finding themselves in that same position, so.

Researcher: And it's not boring, so.

Interviewer: Yeah, that's great. So we kind of talked about R and then you mentioned kind of messing around with Unix. So has Unix come up anywhere in your workflow?

Researcher: So far it hasn't.

Interviewer: Okay. And this ended-

Researcher: Git more

Interviewer: And where's the Git?

Researcher: Oh, all my packages are already pre-made packages, so I go on the GitHub website and look at versions and now I know if I ever need the problems, how to look up codes and like that.

Interviewer: So kind of just interacting with the packages that you're downloading?

Researcher: Yes.

Interviewer: Nice. And is there any other areas kind of in this workflow thinking towards the future that you think you might make changes in now?

Researcher: Okay. I think my plan is...

Interviewer: Pretty good? So we're interested to know kind of, it sounds like you were signing up for some other courses. Was there anything that really kind of helped you integrate some of these new R packages or really pushed you in the direction of being able to do more programming?

Researcher: I think yes, some of those things that the version control of the R packages. What I have been doing is like, before essentially downloading, the latest one and a lot of my packages don't work after I upgrade. So now I know version control, so I have linked it to the projects and things like that I didn't know before. It has been a huge help.

Interviewer: Oh that's great. And anything in terms of kind of like external factors, the supportive PI or any anybody or thing that kind of pushed you to be successful in programming or that helped you be successful in programming?

Researcher: I think I have my PI support when I am talking about this is given me money and time so I feel well supported and he understands that this is the way it has to go forward. And I'm hoping that he has some epidemiology projects that the data listed he doing together. So I'm hoping that I can join on some things and learn by joining.

Interviewer: So there's kind of that support from your...

Researcher: So basically the electronic health record data and help to get the information especially for my clinical phenotyping or endotyping project, I need that kind of information. And right now I don't know anything.

Interviewer: And what about barriers? So it sounded like maybe a little bit of a barrier was the fact that your data analyst actually stayed so you didn't feel that pressure to learn. Was there any other kind of barrier in terms of implementing your new programming skills or things to practice in the course?

Researcher: I think sometimes I struggle a lot and I wish it was a more formal course, like maybe a semester-wide course where I get to do things I still haven't signed up

for any of that. So it's also the time involved and being in a separate campus is a problem.

Interviewer: So coming over from the [campus]?

Researcher: I'm hoping that maybe for the spring semester I'll sign up for something, which then I have a little more time after my grants are done.

Interviewer: So is there like a schedule or a cycle of the year where it's maybe easier for you to spend time on this or is it just kind of whenever your project

Researcher: I just don't know about my spring schedule yet and fall. I know I'm really busy with stuff, so spring sounds like I'm thinking of signing up for a grant writing unrelated course, so I'm going to be here a lot more. So makes sense to do it. I think it's essential for my career to have those skills. So I move forward and I don't have to depend on a data analyst for everything.

Interviewer: Yeah, I was going to ask, what do you think kind of the impact, how would this impact your career?

Researcher: Yeah, because when I was doing my PHD work, all the data analysis was on me. I had actually learned SAS programming and do it and I had taken a course and that was really helpful. So I'm thinking maybe I'll do something similar here and maybe go for a formal course. So I know this is like, I'm troubleshooting now Google stuff and figure things out, but if I have a formal course, and maybe a little bit more involved in the process of learning, so.

Interviewer: And so do you think that will open up kind of new career paths?

Researcher: Yes, I'm really hopeful that, you know, especially the high dimensional data is everywhere. If I'm doing this or anything, more epidemiology intense, which is what I'm hoping I would, I need that.

Interviewer: Great, glad to hear that. Oh man, we're just blowing through this. So the last piece of this is our checklist, which we did last time. So we're just going to go through that again. And so the first question is, do you use any programming languages in your workflow right now, like R or Python?

Researcher: Yes, I do. R.

Interviewer: R. Do you have any step by step workflows that you've changed into scripts or functions?

Researcher: No

Interviewer: Do you use version control to manage your code?

Researcher: Yes. Well R not for git.

Interviewer: Yeah. Do you have any open source software?

Researcher: Yes.

Interviewer: Do you share any of your code publicly?

Researcher: Because I don't write the code, no.

Interviewer: Yeah. And do you-

Researcher: But I think when we publish our data, it's going to go in my lab script from the data analyst. The code is going to be going....

Interviewer: And will it be published as the supporting information file?

Researcher: I think so. Yeah.

Interviewer: And do you share your computational workflow or protocols publicly?

Researcher: No. Also because it's all open access and it's downloaded from someone else.

Interviewer: So you just point back to that?

Researcher: Yes.

Interviewer: That makes sense. And then we're also wondering, so one of the things that we're trying to do with these courses is, you know, we teach you the basics of programming and one of our kind of bigger picture goals is it'll help you to make your work more reproducible. And we're looking at a very kind of particular segment of reproducibility of just computational reproducibility. So you have your data, your code, your results, and somebody else could take the same data and code and get the same results. So I'm kind of wondering if that is something that you think about and if you feel like this workshop has helped you move in that direction?

Researcher: I think it has helped me. I think it's really important that it has to be reproducible. And at least, if I share my same data files, somebody, otherwise it's like manipulation of the data, and molecular biologist, we encounter that issue a lot, that somebody publishes a protocol and you try to do the same exact thing and it never works to same way. So at least in the computation world it's possible, to share the data. I think that's a way to go.

Interviewer: So you think it's helped you move...

Researcher: Yes.

Interviewer: And how is that?

Researcher: Because I have been actually using the training data sets. So it's basically standardizing the protocols that people use. So if you're going through the same questions and seeing controls, then you can be sure that your data is actually valid. And also reproducible.

Interviewer: Yeah. Sounds great. All right, well that was all the kind of formal questions I have. Is there anything else that you feel like is relevant about your experience in the workshop or things you're doing differently or thinking about since you in the past three months?

Researcher: No, I wish I had more to share because I didn't do as much as I hoped I could cause the pressure was off and so I don't have to actually to watch. But I'm interested in other courses if there are any that you could suggest to send them my way.

Researcher: Yeah, definitely.

[Note: researcher started asking about other Library courses, this has been cut from the transcript.]

Interviewer: Might be different group. Thank you. So, yeah. So yeah, thank you for chatting and good luck with your future, our stuff. And that's for you.

Researcher: Thank you.