Owen Transcript

Katie [00:00:01] OK. Yeah, OK, sorry, go ahead. You said you had a look at the code

Owen [00:00:10] Yes, so I had a quick look, but only a quick look because I wasn't entirely sure what I was supposed to be looking for, when I was looking at it so. Yes, so you you'll have to explain exactly what sort of, feedback on the code that you're after.

Katie [00:00:31] Yes, I will do. Now, er, do you have a bias towards like either R or Python, I'm kind of gathering that, you probably like R.

Owen [00:00:43] Yes! I'm an R person. I wrote a book on R, so...

Katie [00:00:51] Quite strong bias towards R then?

Owen [00:00:53] So having said that, I wrote my book some time ago. So I'm actually not enormously up to date on the whole tidy verse broohaha. I've actually been reading Hadley Wickham's book on the Tidey verse just recently, so. And the thing that's quite interesting is. The definite opinion on graphics in particular and how to present them and also write code for the graphics.

Katie [00:01:36] I haven't seen that might be interesting to look into, definitely.

Owen [00:01:40] Okay, so who I saw I mean, your code uses ggplot, I saw. So, yeah. So, I mean, that's very, very different to the way graphics was done in R when it first came out. There've been a couple of iterations actually, so the- ggplot has essentially its own programming language. Whereas standard R, it's got a bunch of printing primitives which are just that, you know, they let you they can print points, lines. Yeah, well, they can fill polygons and things-

Katie [00:02:25] I actually meant to try and use ggplot for the assignment I've just turned in and then realised, oh, I can't use that.

Owen [00:02:34] Ah, yes, for the for the next one I have, for the next one I have eased that restriction. I realised that because I think once you're in the MMORS programme, aren't you?

Katie [00:02:49] Yeah, yeah.

Owen [00:02:50] Because everyone who's in the MSc programme, the OR and applied stats do the stats module and first semester where they were introduced to ggplot, so I've realised, you know, most of the class is familiar with it, so probably I will let people start using it if they wish.

Katie [00:03:13] Yeah, that'll be cool.

Owen [00:03:13] But it's not actually necessary.

Katie [00:03:21] So, yeah, that's interesting.

Owen [00:03:25] Anyway, OK, so yes, so, so I'm biased.

Katie [00:03:29] I mean I'm also biased towards R which is- yeah, I've interviewed a lot of Python programmers so far.

Owen [00:03:37] And Vince hasn't converted you yet?

Katie [00:03:39] That's not yet no. I, um, I did complete a whole year of just doing R on placement, so I'm pretty much solid on that's my favourite language.

Owen [00:03:55] We'll have to get you onto Julia next, that's-.

Katie [00:03:59] Huh?

Owen [00:03:59] Or have you- you haven't come across Julia?

Katie [00:04:01] No, I haven't.

Owen [00:04:02] So that Julia's is designed for, er, numerical work. So it's fast, much faster than both R and Python, but still not too hard to use. So nothing like C, which is soul destroying.

Katie [00:04:23] Yeah, I have some engineer friends that have had to use C, and it looks...not great.

Owen [00:04:29] Yeah, it makes you worry about things we don't really need to worry about, or we don't want to maybe.

Katie [00:04:37] Yeah, I'll have a look into that as well.

Owen [00:04:41] In your spare time. Anyway, but meanwhile, I should probably let you get on with your interview.

Katie [00:04:49] So, I mean, the first question. Well, the next question is just do you have the initial comments on the code or like similarities and differences between them, sort of readability of them.

Owen [00:05:00] Oh, OK. Right.

Katie [00:05:04] Just like your subjective opinions on them.

Owen [00:05:05] So, you know. OK, well, in fact, I did have I actually had. One impression, strong impression I had with the code was to do with the size of the functions, and it's something that many people have noted, not- not just me, but small functions are much easier to digest. So so some of your code was written in, know, little sort of 10 line bites or chunks which are sort of then stuck together. And that, of course, is is much easier to to read than a great big, long things. So that's that's a sort of a general impression. It seems to me that your R code, those seemed to come in longer bits than the Python code. So I'm just just reminding myself, as I said it was only a quick look that I had yet. Yes, if I if I had to be honest, I say I probably found the Python code easier to read than the R code, in terms of what it was trying to do, but neither of them were particularly difficult. OK, and you've you clearly haven't documented it, so you're relying on the the names of things to convey the meaning.

Katie [00:07:53] I'd spoken to Vince about the documentation of things, and we kind of decided for this interview to just kind of leave the codes as they are just to sort of fully, sort of, compare the coding itself. And define variables further up and stuff.

Owen [00:08:11] Yeah, but I mean, with this code it wasn't- there wasn't a problem particularly so the meaning is generally fairly clear just from the sort of the names that you've chosen for the the functions of the various bits. So, yes. So that's a fairly vague response. Did you want something more quantitative or.

Katie [00:08:31] That's good. Yeah. And then the next follow on from that is what kind of do you feel either code could be changed in any way, in your opinion?

Owen [00:08:45] Well, I suppose unless I knew what was being used for, I wouldn't I wouldn't be inclined to to try, I suppose. I suppose the other general impression I got is this is a sort of code I would expect to be sitting behind the scenes with some other sort of interface for me to use. In which case, as long as it works, I don't really care too much, I suppose.

Katie [00:09:29] Oh, that's good. And then sort of, um, so in terms of how well suited for visualisation each language is. Just sort on, I don't know, based on these codes and your sort of knowledge of programming, how well suited you think each language is to visualising and visualisation.

Owen [00:09:54] Oh, I think it's a strength of R absolutely and the ggplot in particular, so I would say that's extremely well suited to visualisation. Well, ggplot is and again, talking about Python, I think, again, you have to specify what pack- or your library you're using for graphics. I must admit my understanding of Python graphics is that most people use something based on Matlab graphics, don't they met matplotlib or something?

Katie [00:10:39] Oh yeah, yeah

Owen [00:10:40] Yeah, so so that's a sort of a python import of a matlab. Graphics. Language, isn't it, Google graphics and so. Yeah, I mean, Matlab certainly has good graphics as well. I don't have enough experience, I suppose, with Python graphics, to have a really informed opinion. I think like most people, once you once you've learnt how to do graphics you are happy with, you tend to stick to the system you've learnt because learning a new system takes so long. So, for example, I mentioned Julia as a language, so I've learnt that fairly recently. And doing graphics in Julia, well, they've, they've imported matplotlib, but it's also fairly easy to use R to do plots for Julia output. And so that's what I do. And so instead of learning a new system.

Katie [00:11:56] I mean, that's definitely interesting.

Owen [00:11:58] Cling to the one I know.

Katie [00:12:00] So, yeah, it is interesting because I can sort of. So I might look into that a bit for the project now, and sort of discuss that a little bit. So thank you for, bringing that up.

Owen [00:12:15] Absolutely. Oh, yeah. No, I've tenaciously cling to the graphics, so I know it works.

Katie [00:12:28] Yeah, I, generally use ggplot for everything.

Owen [00:12:34] Yes, well, I am yes, I am learning, in fact, I've just been so I've just been writing a proposal to update the we've got a 10 credit module on our coding for the MSc students in the second semester. So. So I've just been updating the syllabus for that to specifically include the tidy verse and shiny, in fact. Which is why I've been actually reading up about these things just recently.

Katie [00:13:07] Is that the the statistical packages module? I'm taking out one.

Owen [00:13:14] Well, you yes, shiny won't appear until next year, I'm afraid. But I do think Andrei touches on ggplot a little bit. But what's there to think? Yes. So there'll be more of that hopefully next year.

Katie [00:13:33] OK. Yeah.

Owen [00:13:36] And so this was meant to think this. Because there are some new programming paradigms, too, that sort of have been introduced with the whole tidy verse, the idea of pipes in particular is, new to me as well. I think I've drifted off topic there a little bit.

[00:14:08] Well I kind of drifted along with you from so again just comparing the two well, comparing the two libraries this time, so ggplot and matplotlib. So for a beginner with sort of equal experience in both languages, which library do you feel will be easier to learn?

Owen [00:14:40] Ooh, I think actually probably ggplot. Was the. And again, as long as you're not trying to do anything too different from this sort of standard set of plots that it caters to. Which is fairly rich, I think, and then it's got, I think, a nice, logical system for for doing things. As soon as you want to do something a bit out of the ordinary, then life becomes much more difficult and my suspicion, not one that I've really tested, is that making ggplot do something a bit different is going to be harder than making matplotlib do something a bit different. So that's your a bridge to cross when you come to it. So for a beginner, that shouldn't be an issue. And I suppose I mean, another thing for a beginner, something that's very important, is getting a feel for what good plots look like and what's possible as well. And I think ggplot makes that quite easy. Well, relatively easy. Yeah.

Katie [00:16:12] So now I'm going to show you some of the plots that were made, and I'd just sort of like your opinion on which language you feel, the output is more sort of publication ready. So these are just so you can see from the codes I've tried to keep to sort of default scales, default scale numbers as much as possible and things. So I'll just share those with you. So this is the first- so these are just the control. So I haven't done any- haven't made any alterations to the scaling, or anything that's obviously on the left is ggplot, and on the right is the python. Or if you have no preference for either.

Owen [00:17:05] I have no preference. I feel like I'm at the optometrist, so I think these are both equally good.

Katie [00:17:14] OK. Then, yeah, just got the same thing, but for the so it's the logarithmic scale and again these are the numberings that are just to- come as the default if you change to a log scale.

Owen [00:17:34] Oh, interesting, okay, well, the matplotlib on the right. Yes, so that certainly makes it clearer that it's using a log scale, which I think is a good thing. That would be my preference out of those two.

Katie [00:17:57] And sort of just as a final thing, how much freedom do you feel each sort of language or package allows for sort of customisation of plots?

Owen [00:18:09] Well, I think I've probably already answered that I think matplotlib allows more freedom. Yeah, well, for the the less experienced user, I suppose that's the caveat.

Katie [00:18:24] And yeah, just any other comments that you've thought of?

Owen [00:18:30] Oh. Yes, I suppose one- Something that- an advantage that Matlab has with its graphics that Python and R don't have is interactive graphics. Yes, so matlab, you can you can you can actively tweak it and see the changes in front of you as opposed to you know going back and rewriting your code. And that that certainly has some advantages that can speed up production of graphs. But it's- It's not a deal breaker, I suppose, but.

Katie [00:19:31] Yeah, no that is interesting.

Owen [00:19:33] Yeah, I suppose that's the advantage of a commercial package over a free package.

Katie [00:19:42] Yes.

Owen [00:19:42] One of.

Katie [00:19:47] That's all my questions.

Owen [00:19:52] OK, excellent.

Katie [00:19:53] Well, thank you for your time.

Owen [00:19:56] You're very welcome. So if you have any follow up questions, let me know.