CAL POLY
SAN LUIS OBISPO

COURSE SUMMARY

DEPARTMENT: STAT

INSTRUCTOR: Bodwin, Kelly

TERM:Summer 2020 10W

COURSE:STAT-331 70 Statistical Computing with R

RESPONSES: 8 OUT OF 26 - 30.77%

QUESTIONAIRE:STAT168ALL

QUESTION GROUP STATISTICS

INDICATOR	AVERAGE	MEDIAN	S.DEV
Instructor Evaluation Average (2.1 - 2.10)	4.39	5	1
Summary Evaluation Average (3.1 - 3.2)	4.38	5	0.81

STUDENT INFORMATION

	1 - Freshman	2 - Sophomore	3 - Junior	4 - Senior	5 - Graduate	Total
1.1) What is your class standing?	0.0	1.0	1.0	5.0	0.0	7.0

	1 - Elective	2 - Required	Total
1.2) What is your reason for taking this class?	0.0	7.0	7.0

	1 - CSC	2 - STAT	3 - MATH	4 - ENGR	5 - Other	Total
1.3) My major is:	1.0	1.0	0.0	1.0	5.0	8.0

	1 - 4+	2 - 3	3 - 2	4 - 1	5 - 0	Total
1.4) Number of previous courses I have taken that were taught by this instructor:	0.0	0.0	0.0	2.0	6.0	8.0

	1 - A	2 - B	3 - C	4 - D	5 - F	Total
1.5) Expected grade in this course:	3.0	3.0	1.0	1.0	0.0	8.0

STUDENT INFORMATION

						RESPONSES				
NUM	QUESTION	AVERAGE	MEDIAN	S.DEV	N =	(5)	(4)	(3)	(2)	(1)
1.6)	Effectiveness and quality of the text:	3.38	3.5	1.06	8	1	3	2	2	0

INSTRUCTOR EVALUATION

						RESPONSES				
NUM	QUESTION	AVERAGE	MEDIAN	S.DEV	N =	(5)	(4)	(3)	(2)	(1)
2.1)	The instructor's knowledge of the subject was:	4.88	5	0.35	8	7	1	0	0	0
2.2)	The instructor's course organization was:	4.38	5	1.19	8	6	0	1	1	0
2.3)	The instructor's ability to create interest in subject was:	4.25	5	1.16	8	5	1	1	1	0
2.4)	The instructor's interest in students was:	4.5	5	1.07	8	6	1	0	1	0
2.5)	The instructor's teaching techniques used in the class were:	3.75	4.5	1.49	8	4	1	0	3	0
2.6)	The appropriateness of the instructor's testing/assessment methods was:	4.38	5	0.92	8	5	1	2	0	0
2.7)	The instructor's grading procedures were:	4.13	4.5	1.13	8	4	2	1	1	0
2.8)	The instructor's ability to explain the subject matter was:	4.75	5	0.71	8	7	0	1	0	0
2.9)	The instructor's ability to answers student questions	4.75	5	0.46	8	6	2	0	0	0

Process Date: Aug 28, 2020

CC: PAF After 5 Working Days

CAL POLY
SAN LUIS OBISPO

COURSE SUMMARY

DEPARTMENT: STAT | **INSTRUCTOR: Bodwin, Kelly**

TERM:Summer 2020 10W | **COURSE:STAT-331 70 Statistical Computing with R**

RESPONSES: 8 OUT OF 26 - 30.77% | **QUESTIONAIRE:STAT168ALL**

NUM	QUESTION	AVERAGE	MEDIAN	S.DEV	N =	(5)	(4)	(3)	(2)	(1)
	was:									
2.10)	The instructor's overall teaching effectiveness was:	4.13	4.5	0.99	8	4	1	3	0	0

SUMMARY EVALUATION

					RESPONSES					
NUM	QUESTION	AVERAGE	MEDIAN	S.DEV	N =	(5)	(4)	(3)	(2)	(1)
3.1)	Overall, this instructor was educationally effective.	4.25	4.5	0.89	8	4	2	2	0	0
3.2)	Overall, this course was educationally effective.	4.5	5	0.76	8	5	2	1	0	0

COMMENTS

2.11) Please provide any comments about the instructor below.

- Wow.. a lot to say here. This came out longer than I expected, but I hope it's useful. Professor Bodwin has some real positives: she knows her stuff REALLY well; she is passionate about her class; she cares about her students and is online and available at all kinds of times throughout the week to answer questions. And in office hours, she is stellar. So none of the following criticism is, I believe, a result of a lack of trying, attentiveness, care, or experience on her part; I think she is brilliant, means really well, and has given a lot of thought to how she teaches.... but it really didn't work for me, and I don't suspect I'm alone in my experiences. What follows is not the most glowing feedback, but I hope I am clear in my intentions that this is *not* an attempt to hurt feelings. I am trying to give honest feedback that is constructive and can be used for improvement only; I edited this multiple times to ensure that what I say is true to my thoughts / experience, and to ensure that it is, above all, useful.

As background, I have taken tons of programming language classes. I taught myself Java, taken several classes in C and C++, Javascript, self-taught Python, and loved all of it thoroughly. I even considered changing my major to CS because of how awesome programming is to me. I've certainly struggled, but I've always enjoyed the struggle.

This is the first time I have ever resented a programming class, and a programming language. (Although I'm pretty certain the latter is a consequence of the former.) Professor Bodwin's approach seems to have the following errors, in my view: her reliance on the Tidyverse package, her philosophy that students should "figure it out," and an unnecessarily steep learning curve in her assignments and curriculum.

(Aside: Regarding "figure it out" -- I certainly believe in being able to figure things out, especially in programming where this will be the case throughout a career, but a more nuanced explanation is required and follows later.)

The class relies a lot on Tidyverse to get things done. This makes sense, in a way -- it's a highly ambitious project to attempt to learn as much data wrangling as we cover in 10 weeks, so this is likely a product of necessity in cramming so much content into a short window. However, what results is what I can best describe as "cookie-cutter" use of code, and "script-kid" levels of understanding. When relying on the functions in a library, the ability to manipulate data is really limited -- you can sort, filter, mutate, do all kinds of things.... but if you change the goal slightly and a function doesn't exist to do what you want, you are suddenly unable to achieve this goal because further manipulation, beyond the available functions in Tidyverse, is not within your grasp. (Either that, or the function might exist, but this information isn't conveyed to you. More on this later.) This is an even bigger problem when, inevitably, you need to Google your issue and wind up on StackOverflow, where you are suddenly "outside" of the familiar boundaries of the course, among R syntax that is entirely unfamiliar and difficult to comprehend. Many of the solutions on SO incorporate Base R, which means that when you find a solution to your problem, there is a huge impediment to your ability to understand what you are doing. As a result, I frequently found myself copy-pasting code and tinkering for 5-10 minutes to adapt it to my use case, when I'm quite certain that, had I really understood the base language, I could have parsed the logic and modified it quickly. Worse, this lack of understanding means you'll occasionally copy something from SO and tinker with it, only to realize that it *doesn't* actually do what you needed it to do... which you *could* have known if you understood the base code in the first place. Tidyverse is powerful, but it isn't all there is, and despite having covered a lot in this course, I don't feel proficient in R in the slightest. I don't even feel very proficient in Tidyverse. Mostly, I don't feel confident, and that's a problem because this is one of the best predictors of how likely a student is to reach for R to solve a problem beyond this class. Why do I always pull out Matlab or Python, even for Statistics courses? Because I'm very familiar in it... and I don't even know how to do that many things in it! (I barely know any libraries or any more advanced features.) However, I understand the base language enough that, when I inevitably run into a problem, I know how to troubleshoot it. I understand the meaning of the code I type, and when I do basic operations, I have a reasonable understanding of how they will work. That is, I have confidence in my ability with it, which isn't something I can say for R.

Another issue with the course is Professor Bodwin's seeming philosophy that, because students will have to figure things out in the real world, they should learn to figure things out for themselves in the class. I find I largely agree with the sentiment -- most people should -- but the manifesting implementation of her course structure and curriculum, in particular, is something I strongly disagree with. I suspect Professor Bodwin cares greatly, and puts a lot of effort into this class for her students, but if not for this conviction, I could imagine that her "figure it out" philosophy is simply the best cop-out to write a minimal amount of teaching materials. (Eg. "I didn't teach you nearly all of the functions you'll need. Further, there are some really useful ways to use the functions that I did tell you about, but we didn't cover more than one or two ways to use them. However, **you'll have to deal with this in the real world when you use R, so it's good practice to learn about these on your own."**) The result is frustration. Sometimes, it feels like

pure luck to stumble upon a function that does what you need. (Eg. Googling "sum data frame rows by group in r" (Note: Not necessarily what you should search, but there are a ton of ways to frame that question and a new programmer isn't likely to know the right combination of terms.) .... and then you get lucky enough to find aggregate(). Meanwhile, another student might not be so lucky, instead resorting to a really clunky, roundabout set of steps to accomplish the same result. (And remember, this student doesn't know Base R, so they're going to try doing it with several cookie-cutter functions that hopefully put the data in the form they need.) It really sucks the joy out of the language. I have always completed my programming assignments before any other schoolwork, typically immediately after they're assigned (which is impressive and unusual for me), but I found myself constantly avoiding R and looking to the upcoming week with dread. Anyway, I'm not even sure that I wholeheartedly agree with the claim that Professor Bodwin makes about "learning in the real world" -- because yes, I will have to learn additional things in R once I move beyond the course.... but I'll be coming in with some basic understanding of the language. This is a huge difference. For comparison, in my major also, there are plenty of things that I'll need to figure out that I haven't learned here in school. Heck, I've already done two internships where this is the case. But just because my teachers can't teach me everything I'm going to need in my career, this doesn't mean they don't do their best to teach me what they can in the curriculum. It isn't license for any of my professors to say, "Instead of doing my best to convey what I know, I think it would be better if you opened the textbook and taught it to yourself." Instead, professors acknowledge that there is a lot to learn beyond the course, and say, "Let me give you the best foundation you can possibly have so that you're ready for any new material you'll encounter past this class." I'm sure I **could** learn a lot of my coursework from the textbook and other materials, but then why have a professor? My answer: the professor smooths the learning process by reducing the otherwise steep learning curve. In my view, if I'm going to spend the same amount of time teaching myself R in a class as I would teaching myself R on my own, it's difficult to see the point of a professor other than as a glorified troubleshooting resource, or a supplier of curriculum. (Both of which I could get online.) Anyway... like I said, I don't think Professor Bodwin intends for this outcome; in fact, I suspect it's a result of R being such a big language, and the vast array of options means it's impossible to show everything that's possible in it. However, she needs to do a better job of covering the content that she teaches, by introducing all of the functions that will be needed and giving thorough explanations of the most common ways we might use them, rather than simply covering one or two applications. The amount of browser history I have per assignment indicates that this standard is currently not being met.

Finally, assignments in this class have a steep learning curve and a great deal of confusion. Some of this is inevitable, as the R programming language and paradigm is a different way of thinking; however, some of it seems unnecessary. First, many of the lab instructions are not entirely clear in their aims. More precise wording would be beneficial in many cases. Moreover, though, it would be wonderful to have a clear example of what we should be aiming for. Obviously, in a real statistics project you don't know what the end result will look like, so you need to use your intuitions to guide you; however, when we are still learning the basic tools, it's hard to argue that these intuitions exist. Specific "check-in" points along the assignment would be nice to verify that we are on the right track. This exists in other programming courses, and I don't see why it couldn't exist here, too. (Eg. "You should end up with a table with 24 observations at this step.") Other times, the error is in my reading of the instructions, which perhaps are obvious in their aims if you know the language; however, I have spent considerable time trying to achieve a particular result, only to later find out that this isn't what she wanted us to do. Either way, less time spent "in the dark" wondering if the output you're getting is what you should have, would be great, especially in the beginning of each week when first familiarizing yourself with a tool.

Instead, Professor Bodwin seems to get to the "challenge" and "testing students" far too early in each week's content. Perhaps she's just so good at R that she's out-of-touch with how large the knowledge gap is; I think she might underestimate how bewildering some simple things in R might be to students who are new to the language. I agree with Professor Bodwin's efforts to "test" and "challenge" students, but in my view, it comes too soon. The stages in which you are first learning how something works aren't the place for challenging tests; they should be straightforward and give you a strong feeling of confidence and competence with the content, so that you feel ready to face a challenge when it finally arrives. Thinking of CodeAcademy or any other "learn code online" curriculum, I realize that they "err" on the side of being boring. This is a gripe, for sure, but it's safer than drowning students in complexities when they're still getting a grip on the basic content. To this end, I think Professor Bodwin should (perhaps counterintuitively) make PAs and Check-Ins longer, so that they can be less "steep" and more slow-to-build; simpler and more straightforward, so that students have *more* practice using the tools they're given. (Maybe have the graded Check-In / PA questions only encompass those same final problems, though, so that students who are more proficient can skip straight to the challenging content instead of slogging through everything? After all, students are at different levels, and while the current approach is too steep, I would also argue against forcing more advanced students to wade through a "flatter" curriculum.) As one example of this, one PA asked students to combine some tables by column, row, and then a cross-reference kind of combination. I found it incredibly frustrating, because I didn't feel very comfortable using the tools to do this. We had only just learned joins(), and already I was being asked to use them at a high level to achieve a somewhat-complex cross-reference. (Which was tough to think about how to achieve even conceptually! Side note: This actually was one of the few times we *were* given the end-result upfront, so I was able to tell if what I was doing was on the right track. Hugely helpful. However, it was a really difficult task, and I would have appreciated an intermediate stage of the data transformation being given. This wasn't a lab, after all -- it's okay to show students a little bit of the steps, if it ultimately demonstrates a brand-new way for them to think about something.) Instead of having only a few problems where I do maybe 3 left_joins(), 3 right_joins(), and then need to do some complex operation, Professor Bodwin's assignments could have multiple problems that slowly build up the complexity, and even some that ask the student to do the same thing. (Repetition is okay here!) Instead, I'm at the end of the course, and constantly have to refer back to previous assignments to recall how to use (or even what the library is called -- was it Date? Wait, it was lubridate) the functions in lubridate, str_[anything], mapping (map() always feels like a leap of faith) and a bunch of other things because I've literally used them 5 times, max. (Another side note: I feel really confident doing filters(), mutates(), selects(), and all those operations, because **we use them on every single assignment**.) Basically, I *want* to be familiar with the tools, and right now, I'm not. We did some really impressive things with data in this course, but for all the time these tasks took, they were so "at my limit" of ability that it feels closer to hitting the gym and destroying my body one day a week, rather than working out multiple times per week. Yeah, I did some Herculean data manipulation, but I didn't get practice with the more basic and less impressive -- but more fundamental -- data manipulation regularly.

I want to conclude by stating that Professor Bodwin is a great professor. Despite all 2000+ words that might indicate otherwise, I am glad that I took this class, and I am glad to have taken it with her. I have strong feelings about how the class could be improved and where I think it missed the mark, but Professor Bodwin cares a lot and knows her stuff. Those two qualities are the most important, and so while this was certainly a challenging experience learning R in her class, I can see how it could have been a nightmare if she either wasn't experienced, didn't care about her students, or

both. Unfortunately, like a Yelp review, the good can be summed up in a few sentences, while negatives are always able to be written out at length. In short, I had some feedback which I think is important to convey, but I hope it's clear that Professor Bodwin is a great asset to the Statistics department at Cal Poly, and what I've written above should not overshadow that fact.

Best,
Alex

- Probably the most organized online class I've had.

- It was hard not having a consistent time when things would come available to students. Living at home and making our own schedules teachers need to take that into consideration. Especially when the work takes some students much longer than others.