I = interviewer

UCSB = respondent, employee of the University of California, Santa Barbara (UCSB) Library

Interview 1

I: I want to start this interview by acknowledging that teaching and learning has been significantly disrupted in the past months due to the pandemic. So, for any questions I'll be asking you, please feel free to answer them with reference to your normal teaching practices, your teaching practices adapted for the crisis situation, or even considering both. I would like to start listening, and learning a little bit about your experience teaching undergrads with data and how does your teaching relate to your current or past research?

UCSB1: Okay, well, I teach data analysis and archaeology for undergrads. I also teach it for graduate students, and I cross list them often. And so grad students and undergrads are in there together. In general, well, I consider it one of the most important classes I teach because regardless of what career they go forward with, they need to understand good graphs and bad graphs, misleading graphs, misleading statistics, knowing how to sort of do basic things, and just interpret things they read on the news, where people show them ridiculous graphs, that may not mean anything, and you're like, what? So, there is that and I imagine any job they're in, they're going to have to have some statistical knowledge. Just to be conversant with others. So the other thing I would say is, it is a required course, for the archaeology track in our department. But there are other options to fulfill that requirement with it is not the one I teach just to have maximum flexibility. I try to teach it every year, I've actually had to cancel it this year and was going to teach it in winter and I did not feel there was a way I could do it with COVID. So, I had to switch out other classes because I was going to do both the grad and undergrad in winter. And now I'm doing something completely different.

I: Okay, so let me just understand, you have this stats class, and you would like to have a combination of grad students and undergrads for a new offer. But have you ever taught this class like having undergrads and grads together?

UCSB1: Yeah, I teach it that way, every other year, because the grad students have a much higher bar in terms of the assignments and what I'm expecting in terms of interpretation, and meaning after they do all the statistics. So there's a higher bar for them [graduate students] to meet. And it is required to the graduate curriculum, but we have fewer grad students. So, it's an every other year offering.

I: Okay. And in terms of getting the data that students are going to be working and engaging with, how would you say they usually obtain the data? Do they usually collect and generate the data themselves? Or do they search for and identify some pre-existing data sources?

UCSB1: I have all of the data that they get. So there are specific assignments for different types of modules that they have to use. And I provide the data set, because it's archaeological data.

And I don't want them working with non archaeological data. Because the whole point is, how to understand like, how we actually learn things from archaeology, that it's not about a find, it's about a large, complex data set that you analyze. And I don't have confidence that any of these students can find appropriate archaeological data to do that, and it's hard enough for them as it is, week by week. I mean, these are anthropology majors, you know. So I mean, of all the classes I teach, the lowest student evaluations I get is always in that class. And it's the same for any other data analysis classes in our department we have. So we have cultural anthropology, biological anthropology, and archaeology. And they all have their distinct data analysis classes. Then, everyone who teaches them gets lower evaluations than any of their other classes. It's just common and expected.

I: I see. So, you have mentioned that you provide the data sets and data sources, how do you usually share or provide access to these data sets? How do you usually share these datasets with your students?

UCSB1: I upload everything to our UCSB platform Gaucho space, and they download everything from there.

I: Okay, I see, and how do you find and select these datasets? Have you been working with the same datasets for a while?

UCSB1: I have been working with the same datasets for a while, because they produce the patterning I want the students to see. And any other data that I could pull in may not actually produce the patterning that they're supposed to see. So yeah, once I find the data set, you know, it works. And, and then over time, I mean, I know the nuance, I know the meaning and I like to see the students struggle with that meaning, which is pretty interesting.

I: Would you say that it's an easy task, or a challenge to find good datasets in your field, for instruction?

UCSB1: I think there are plenty of good data sets. And various reports, there's no doubt about that there is a lot of data available. And I've scoured those reports. And you know, every year, so I change up one week, as well, to find, you know, new datasets, but these students have never seen these before, it almost doesn't matter how old the data sets are. It's just the basic information that was collected that we're looking at. And so I have to step along with the data set in an assignment sheet. And it is at the top, it defines all of the variables that are listed in the data set, and what they mean. And then there's usually two parts, there's usually two data sets per assignment. And then each part has like step-by-step instructions that relates to the statistical program they are using. Because they're learning a new statistical program at the same time as they're learning to work with data. And this statistical program, I can no longer use, and I'm actually in the midst of having to learn in order to keep this class going. Because that [referring to the software R] is the future in archaeology and anthropology, in terms of what people are using. So, yeah, so that's gonna take some work for me.

I: That actually relates to my next question about how your students work with data. So how do they manipulate, analyze and interpret data in your courses? And you mentioned that you're transitioning to another program, right? What tools or software that your students are using?

And then if they need to have, like, any pre required knowledge to get to your classes, that you expected them to have? Or do you teach them explicitly how to use and manage the tool?

UCSB1: Yeah, I teach them how to use and manage the tool. And that's why there are step by step instructions and the assignments to sort of explain a lot of these things to them. I was using Stata, I used it for years, there was a free student version called My Stata. And our computer labs on campus also installed these. So it was perfect! And then this year, no, it was last year was the winter. So, it was in 2020 right before it shut down. I was teaching the course. And I get into the computer lab by the third assignment. It's clear that they updated the operating systems and about half of the modules no longer worked on that program. That was a really ridiculous semester having to figure out how to do complex things on Excel, but they had access to it. So, they learned to use Excel a little bit better than they had already. But it was ridiculous, it was just half [of the intended content]. And you didn't know if, if something was going to work until you got to it. So, after the first one [referring to the assignment] didn't work, I went through to try to run every assignment. And at least five out of the 10 weeks, those modules said error will not run. So yeah, that's why I decided it was time to switch to R. At first, I thought I would use data and bought my own version of Stata. But then in talking to all the grad students, they're all using R, they all took my class, but now they're using R and they figured it out. So I can figure this out. But yeah, there's gonna be a learning curve. There's not a lot of time in my life right now, for course redesign. I just finished redesigning one that I'm going to start teaching on Thursday.

I: Yeah, so your classes are highly dependable on the software?

UCSB1: Yes.

I: And...do you face any challenges relating to your students' abilities to work with data? Like, this process of getting to teach them how to use a specific tool, and then also go through the process of working with the data? Would you say it's a challenging process in any way, to get them working with the data?

UCSB1: Oh, it varies. First, it really does. Some of them come in with, you know, more of a math background, or have had statistics classes before some of them come in with an Archaeology background. And so the assignments just make more sense to them. So it really varies, but I work with students closely. This is the class that I have more, one on one officer contact with students. And I have computers in my lab too. And I had a student that couldn't get it because the whole mindset was supposed to be downloadable for students to work on their own computers. But that was also bugging out on them. And so I had students that I had, I have multiple versions, you know, over the years. And so I have a lab computer that has a version of Stata, the full software package. And so I had students that I would say that would come in, and I would have them use one of those two computers, and then they would get through everything and print their stuff. And so I make myself really accessible. And I try to have as many workarounds as possible.

I: Great. So you mentioned in the very beginning that you are concerned about teaching them how to identify misleading statistics, you know, for them to think more critically about the data.

What would say, like specific data skills that you teach, that can connect to these and other ways of thinking critically about the data?

UCSB1: Well, we spend a whole week just looking at tables and figures that I throw up on the screen. I first start by having like a 30 minute presentation to where I'm showing them where we talk about the elements of a graph. So you know, they are not just your axes, but your, your data box, and then you brought a rectangle, the tick marks everything, the placement of everything. And I'm losing my train of thought...so yeah, there's that and then I have all sorts of bad graphs that I have found. I am at the point where I have students that took this course with me 15 years ago, will be at a conference will take a picture of somebody's terrible graph and just send it to me, you know, and so I show good examples and bad examples, some that just have a few mistakes, some that you're like, “What in the hell is that even saying?”... and we go through them. And I asked the students, you know, “What does this mean? Do you know what this means? You know, what is the problem with this graph?” And so through that, they get to see the misleading nature of things. Like, you know, I can only think right now, of course, I'll never have access to this graph, but the one that Trump showed in that Axios interview is a bar chart that meant nothing was hilarious, you know. And so it's those types of things I want my students to leave, see something like that and be like, “What is that? What does that even mean?” That's ridiculous! [referring to the Axios interview episode]. So, those are like the critical thinking skills, I want them to have to be able to assess right away, if you know, some representation of data that someone is putting out there is problematic. And usually, you can tell it's problematic, just by the way, they have graphed it, there are ways to graph things to make the pattern look less clear and to make a pattern that's not there look like an actual pattern. And so we talked about where people place their axes and all of that.

I: Okay, great, yeah, many ways to make them look more closely to the data, and then kind of extract meaning and also to identify some misleading statistics. So I would like to hear your thoughts, about policies or like cultural changes at UCSB, or even like in the field as a whole, that have influenced your ways of teaching with data? Have you noticed, like some cultural changes, or that has impacted your teaching?

UCSB1: Absolutely. And that's why I've just finished redesigning a course that I'm teaching in the fall completely and utterly. In order to, to address that. In terms of the data analysis class, I have already thought about redesigning that, and I am going to work in graphs in my graph section and, you know, a lot of them aren't about archeological data. They're just examples, good graphs and bad graphs. And in order to understand if they're good or bad, we have to assess what they're actually trying to say. So I am going to work in data and statistics in the future relating to some of these issues that you're experiencing right now related to an equities, Black Lives Matter and that sort of thing, where, you know, I could show a graph of percentage of white people killed by police in this year span and, and you know, just go categorically through that, and maybe show how, if you use the data a different way, it might show no disparity. And so we can go through that. And I'm not engaging in the debate, right, I'm just going to be showing them different ways that these data can be represented to be misleading, and that sort of thing. So that's, that's how I'm going to work it into that in terms of the datasets, where they're all archaeological. And so that's more difficult. And it's not the type of class where we can get into the more complex interpretations archaeologists have about, you know, ethnic diversity in urban cities, that sort of thing. Because those require multiple lines of evidence of different kinds that are really too complex for me to introduce into this sort of introductory data analysis class. But I'm, I'm still thinking, trying to think of other ways as well to do that. But right now, I'm just seizing on the graph presentation section.

I: And would say that there are like ethical challenges in teaching with data, in your field? And if so, are they important and to which extent?

UCSB1: Let me think about that. So now, if we look outside the field and within academia in general, one thing that is, you're hearing more about. I'm sure it has happened throughout time. But people are actually being more just discovered for falsifying data and patterns to arrive at what findings they want that we've seen happen in the news to professors. And so to that end, I do talk about how serious it is that you represent the data as best you can. And I have a story of a paper I had submitted for publication, you know, I got it revised and resubmitted it, they [referring to reviewers] wanted me to look closer at some of these things. I put it aside for like, a year because I was busy. I went back to the data. And I had a realization about one of the sites and the data sets. And I realized I had to remove this one context, because it was completely and utterly different from everything else. And once I did that, the primary pattern for which the paper was based disappeared into so that's one of those situations where I'm glad I didn't revise and resubmit right away. But at that point, you know, anyone has a choice, right? And an ethical choice. You know, do I just move forward and pretend that doesn't exist? Or do I drop the paper because that pattern doesn't hold up anymore. And I dropped the paper. And there was one other pattern in that paper that was really interesting. And I was able to fold that into a different paper. But yeah, I had to drop that. And, so I bring that up to students. And I think I want to find some of those news pieces about the past five years. I think there's been, I don't know, two or three professors in the sciences who've been called out on this. And maybe I should share those as well.

I: So, you basically use that as an example for your classes and for demonstrating to your students how, you know, some things have to be reconsidered when you notice their mistakes?

UCSB1: Yes. And, you know, data analysis is never truly complete. You reach a point where like, okay, I'm ready to publish, but it's never complete. And so you always have to reevaluate things, if they agree.

I: So, we are now going to move on to some questions more related to training and support. First of all, I would like to hear a little bit about your interaction with other divisions or if you have collaboration or support at UCSB with other instructors, librarians, teaching assistants, or even people that do in class presentations for your class, for providing extra content related to teaching, sorry, not teaching, but working with data. Do you have this type of collaboration?

Does anyone other than you provide instruction or support for your students?

UCSB1: Sometimes, yes. Not always. But yes, sometimes this past year I had. I had an archaeologist from [name of the institution]. And he is one of the issues in archaeology right now is data curation. Archaeologists don't know how to design and use databases. And so it always would cost a lot to find someone to do that, like you have. And I'm talking about every site, every assemblage anyone's ever worked with it. All of it is in like, accessible via integrated data system, I don't know, like access or something. And the argument here is that we need to bring this conversation first to undergrads because they're going to be the future archaeologists. And they need to value the importance of actually maybe training and this sort of thing. And I mean, it's the difference between, you know, saying, “Oh, yes, I have this information on that site!”.

And then maybe you have to, maybe you've got folders and folders that you're trying to get through to find it or hard copies, you got to find a site report. And if you had an integrated database, every piece of information would be in there. And you could just click on the site and see what's there, and then click on field files, or this or that. And it was really illuminating. And it's something I cannot take on in my life, because I don't know how to code. But sort of a new generation of students these days, it's all about coding, coding, coding. I mean, there's all these things for our kids, to teach them how to code. They're learning it in elementary school even.

And so I think I may keep having him come, or, you know, however long COVID is, I would have to zoom in, because it is a really important topic. And we have to be looking towards the future, if, you know, if we want good data curation.

I: Great. So, to your knowledge, are there any ways in which your students are learning to work with data outside their formal coursework? For example, online tutorials, internships like we have the carpentry workshops on campus. Do you know if you're like looking for these external resources and like to learn how to deal with data?

UCSB1: Yes, some of them definitely are. So let me think. So I have other upper division classes I teach. And the students in my class that have already taken my data analysis class, like if they have a project, where they have to analyze the data set in that in, in the non data analysis class, I have a different set of expectations. You know, I told students, if you have not taken my class class, then here are the basic things you need to do. Those of you who have taken my class, I'm expecting this and this and this, in addition to those other things, because I know you guys know how to do that. I want you to push it. So there's that. Also, the guy I told you about that came to my class in the winter, he recruited two of my undergrads as interns to learn that, and to help him continue the database construction and curation at the [name of the institution previously mentioned] He's the archaeologist for that so it was really cool. Students do senior theses, and those that are doing them in archaeology, well, definitely, my students have already taken that class. And so they can hit the ground running with their data analysis, and I don't have to be with them every second, basically, which is what I want. And let's see, oh, yeah, and of course, the class in terms of graduate students, that takes them a step further, towards actually being able to put together good publications where they, where the arguments they make from their data are believable.

I: Nice. So, from what you have mentioned, I can tell you encourage this kind of extracurricular learning activities, correct?

UCSB1: Yes.

I: And how do you usually approach that? Like, do you mention these things in classes? How do you stimulate or encourage your students to look for other ways of learning about data and about tools to work with data, and so on?

UCSB1: Oh, let's see. Okay, so I'm in these other classes I have, you know, if I throw something up on a slide, and I'm like, What does this mean? And then nobody responds, I will say ‘Okay, come on, those of you who are in my class, you know, the answer to this, “what is it and how do you know it?”. And so I will call them out to share that information, which demonstrates other students the value of having taken a data analysis class more broadly in their education. So there's that, in terms of grad students, I have to tell you, I know you're interested in undergrads. But every archaeology grad student in our department has taken this course. And all of my students, it's clear they're using their skills. But I might get to like a dissertation of some other student, of a student of some other faculty who's taken this class, you read through the whole thing, and they haven't done any data analysis. They like, give you a table with numbers, and maybe some percentages, and they're making complicated arguments from these sort of basic things that are just stepping stones. And, you know, at that point, I'm hard pressed when the student comes to talk to me to not get really annoyed. I'm just like, you know, your PhD, you've taken it to the point where I'm like, “Do you remember taking my quantitative analysis class?”. You know, it's just, it's astounding. Oh, one other thing that I do too though, because for me, it's really about learning and understanding the connection between the pattern and the meaning, that the students do every week, there's an assignment, and it's due the next week. And it requires that you show all your statistical data that you present things and tables, that you write up your results and interpret them. And it's the last part that students often don't really engage with, which is the whole point.

I: Right.

UCSB1: I've taught this.this is the whole point, we want to understand the past. So you've got to push it, my students can rewrite their assignments until they get an A, if they want to do it four times, that's fine. As long as they're all turned in, by the last day of class, all the rewrites. And some of them, I go through three or four with them. And then, I get more contact facetime office hours with these students than any other students to take any other classes. So I know these students, by the time they leave, I know a lot about them. And so that's another thing I like about it is that I really get to know the students. I mean, I think it's that interpersonal interaction, that can create the excitement they see. I'm excited about it.

I: Right.

UCSB1: So, in that regard, the class also becomes less daunting, and they can perceive it as you know, something they can accomplish, that they can get through. It's not scary. Because of you know, you think math fears people to go into ANTH [referring to the class]. I didn't even when I was an anthropology major, I did not have any math or requirements.

I: Oh, okay.

UCSB1: So, these kids are afraid of math.

I: Yeah. And that connects actually to my next question which is...I would like to hear a little bit about your own training...so...if you have received any training for teaching with data, other than your graduate degree, let's say, if you participate in workshops, if you had some help from peers...it doesn't have to be formal training, like more informal too, how have you acquired your skills?

UCSB1: Even though my degree, my BA did not have a requirement for math, I nevertheless took a statistics course for Anthropology that was offered. And that was really key, but I didn't really know how to really apply that to archeological data. Oh, I also took just a regular like a sociology stats class too as an undergraduate, I did both of those things. And I did them because I thought I would need to for archeological data. But it wasn't until I finished my BA when I was actually working on a conference paper. One person who was also in grad school sort of looked at it and he's like, “So you really don't show a pattern and there's a way you do it. And here's sort of an index you might use in order to control for sample size, different sites”.

And so I was like, and when I did that, yes, the pattern emerged, and it was incredible. And that was right before our first year of grad school started, we'd already moved there. And then we had a graduate required course as well. But before I took this graduate class, I think it was in our second year, we had to take it. I had done a research project the first year. And I had used my statistical skills before the grad school class, to find a pattern. I was reanalyzing a data set that was already published. And I found a pattern that called into question this person's entire thesis, basically, what they thought was going on. And I was able to show a correlation between the number of food species people ate, and the amount of soil that was excavated, right, so more soil, more diversity, more possible diversity in terms of what you find. And so I brought that paper, you know, to one of my professors who was not my advisor, but he also taught the quantitative class and my advisor told me to go to him, he reads the paper, and he says, "You need to use this module, [interviewee’s name]! And you can actually demonstrate statistically the sample sizes, the reason for this!” And it was this old program, that archaeologists have created, well, in the 1980s or something, I still use this. And it's hard on these new operating systems. But I learned now, once I learned, that I can create a whole new module that does the same thing. So I'm really excited about that. But in any case, and it was that that sort of clinched for me. And it was really hard to use, because it was so old, it was hard to make sure you set up the data data file, right, you'd have to do it as a Word doc and save it as a txt file, and then change the extension, and all of this, and sometimes it works. And sometimes it didn't. So, I became really good at using that program. And it really changed. Sort of, it sort of brought me to the next level in that moment, and then, in the following year, taking the stats class, and that was all, again, archaeological data sets, problem sets. So we're learning statistics, while we're learning how to use them. And that is how I base my quantitative class on that class that I took there.

I: All right, so would you say that you learned it by yourself in most cases, because you were triggered by curiosity, and then you kind of explored these tools by yourself? Or would say you went through some kind of other training? What about some help from your peers?

UCSB1: Oh, yes. I mean, oh, when we took that class, all of us, like, first, I don't know, like two cohorts of archaeologists got together. We had a computer lab for the archaeologists, and we would all basically sit there in the same room while we were working on it. And when someone encountered a problem, like, "I can't get the program to show me this!" And we're supposed to see this, somebody else figures it out and says, "Okay, here's, here's the button, you press or here's the thing you go into!". So, it was very collaborative in terms of learning the program and how to do it. We all of course, you know, dealt with our interpretations and all that independently. So there was a lot of collaboration. Absolutely. And that collaboration continued throughout grad school, in terms of reading each other's papers, checking out each other's patterns, seeing what we think all of that. It was a very, from my understanding, rare experience in grad schools, from what I understand, but it was impressive. So I had formal training, I had a sort of collaborative interaction, which pushed us all further. And then I had sort of my own initiative based on specific things I was working on. So all of those three things sort of came together.

I: Okay. Yeah. Nice. And I see that you're kind of applying some of these concepts to your classes as well, like when you engage students with data, for them to think critically...

UCSB1: Okay, so another thing I also tell them, I say, “You guys can work together, as long as you don't work on your write ups together”. Okay, work on the analysis and figuring things out”. So yeah, it's not that they don't have to do it all by themselves. Because no one does this all by themselves, really, in the real world, you know, we work with groups or partners.

I: Yes. Nice. And you mentioned in the beginning that you usually provide the data sets [for your students], and that you have this collection of data sets that you believe are good, like examples for your classes, because you know them thoroughly and you're, and they're good for like, things you want to explain in some of the skills you want students to develop. But in terms of like assignment plans, or syllabi, or even other instructional resources that you might, you know, apply to your teaching plan. Do you usually find external materials from other schools or do you share with, you know, other faculty those materials on how to put together a class that involves instruction with data?

UCSB1: No, because everyone else is clueless, I was hired to do this. And evidently, when they interviewed people for this job, that was one of the key things, so I showcase all my quantitative skills in my job talk. You know, I made sure I had all sorts of things they saw I could do, I didn't spend a lot of time explaining how they worked. I told them, they could ask me that later. But I would, you know, refer to them. And usually everything was pretty clear. And nobody else had done this, no one else showcased any quantitative skills for a job where they were asking someone to teach these quantitative courses. That's why I got it [referring to the position], or one reason anyway. So at this point, some of my grad students have moved way beyond me. I mean, way beyond me. And that's, of course, why I want to move on to R. So if I'm collaborating with anybody, it is not other faculty members, it is graduate students.

I: Okay, and do you usually share these...your own resources, make available to them if they want to kind of, you know, if they're going to be teaching classes with data, and having the experience of taking your classes...

UCSB1: I, you know, I share all of my course materials with my students that they ask. Okay, they just have to ask, they ask I will share.

I: So, considering evolving trends in the field of archaeology, what types of training or assistance would you say that are most beneficial to instructors in teaching with data?

UCSB1: Training us on programs! Training us on new programs! It takes a long time to figure out a new program. And I would love it if there was some sort of, I don't know, three week workshop on how to use R, something like that. I, you know, I rely on my grad students acquire, I will be relying on my grad students to help me with that they've sent me all of the material, like all the books that they think are the best ones, they've sent me the YouTube videos that walk you through it. So I have gotten all these resources from two students who have previously taken my quantitative class and did not know are at that time. So basically, I've taught them, they have gone beyond me, and now they're teaching me and that is the way learning should work even with our children.

I: Nice. Okay. So we are almost at the end. But to wrap up, I would like to ask if there's anything else from your own experience, or your perspective as an instructor, or on the topic of teaching as a whole, teaching with data more broadly, that I should know...something you would like to consider about your experience as an instructor?

UCSB1: Well, I can say that I've deeply dedicated [myself to] undergrads instruction, and undergraduate experience. I've been the faculty undergrad advisor for my department for years, and vice chair on the undergraduate council at UCSB. I was a first generation college student myself. And I interact with students a lot, I have every quarter between five to 10 undergraduate interns working in my lab on research projects with me and my graduate students. They both apply to work with me, but I also invite students from my classes that are high achieving. This is also how I can ensure that I get some racial and ethnic diversity in my lab, I invite people. Most of the people that apply to work with me or are all white, and I think this is changing. I wasn't the faculty undergraduate advisor for a few years. And so now that I'm back to that...I am emailing the students a lot more about any and all opportunities and that sort of thing. So I'm hoping that makes everything seem more accessible to them. And I'm talking, you know, opportunities with me, opportunities with other faculty. Any sort of external internships or jobs come through, I mean, I, I send everything like that to people. So, I mean, I even have a rule that I don't take any interns in the summer. I mean, this is the only time I can get my own stuff done. And I broke my rule twice, one for a program that was meant to provide some sort of experience to understand the PhD track by students applying to this program. underrepresented students, students of color, in their master's program, and so I accepted a student to work with me. It was summer 2018. I don't know. So a few years back, and then this has not this past summer of 2019. Our classes end in June.

I: Right.

UCSB1: I got this email from this high school kid who is going to be a junior. It's clear from his name that he is Asian American. I just can't tell you how hard it is to find students that are not white in archaeology. And this is, you know, and it's, it's almost too late by the undergraduate period. And so I jumped at the chance. And he came into my lab and he worked every week until COVID. basically like a, like four hours a week. He got to know my undergrads, he got to know grad students, he got to know all of these possibilities for the future. And he was just such a great kid. Right? And I, he just recently asked me to write a letter for his college applications and sent me his materials. And it wasn't until I looked at his materials, I was stunned. He's actually going to major in anthropology.

I: Nice.

UCSB1: Yeah...and I was like, wow, you know, I just, that just is very exciting for me. So I have to say that yeah, I really, I really interact with, with young students, like with undergrad students, I may actually start recruiting some sort of student from high school every year for the summer. We'll see. But that was a really rewarding experience and I just feel like wow, you know, I made maybe one difference here towards you know, diversifying my field. So yeah...

I: It's very inspiring. Great. Oh, thank you so much. I will stop recording but I would like to say a few things before we end this call. Is that okay?

UCSB1: Yeah, that's fine.

I: So it will stop recording now...(the non-recorded portion of the conversation, for about 5min or so, was used to advertise existing campus resources)

Interview 2

I: First of all I want to start by acknowledging that teaching and learning has been disrupted in the past months due to the pandemic. For any of the questions I'm about to ask, please feel free to answer with your reference to your normal teaching practice or your teaching practice that is adapted for the crisis situation or even both, ok? So, I was just checking if the recording was okay... yeah. So, I'd like to start listening about your experience teaching undergraduate students and if you could briefly describe how your teaching relates to your current or past research and in which of the courses that you teach do students work with data?

UCSB2: So how does my research affect my teaching?

I: How it relates to your teaching...

UCSB2: So yeah, so like, well, so basically, I guess that I, so the primary for undergraduate data classes, so far that I've talked, I primarily have been teaching a GIS class. And a lot of my sort of research ultimately relies on GIS in one capacity or another, to do different kinds of spatial analyses to understand the data that I produce, and analyze. And so, you know, within the, through the teaching, you know, what I try to do is I try to think about things that I wish I knew as an undergraduate, that would have kind of, like, pushed me along a little bit faster, in my own research, and explain things in such a way that I think, you know, sort of my past self would have appreciated for, you know, sort of, like some kind of teaching situation. And so ultimately, you know, kind of like, when you're the way that I sort of go about teaching things, I think about like, what are the core things that a student really needs to know. And then in order to sort of be functional, and you know, to be productive using sort of the data tools, because a lot of it, you know, it being a GIS class, it's as much about the software and teaching the software as it is also about the data. And so trying to, you know, you have to kind of combine that. And so, you know, it's sort of like thinking about, you know, how you can kind of combine those two elements to, you know, sort of create some kind of result, and you have a slight challenge. Well, I'll just sort of leave it there for now, since we have a while ago.

I: Alright, so I’d like to hear a little bit about how your students usually obtain the data they engage with. For instance, they collect and generate data themselves, if they search and identify pre existing data, or if you provide them with the data sets that they are going to use for your classes.

UCSB2: It’s usually a combination of searching and identifying and I provide some data for them. Pretty rare that… only a graduate student would be able to provide their own data. An undergraduate student typically there’s no way for them to have usually their own data that’s truly their own.

I: Okay. So, but say that in general they identify existing data sources. What would be the primary method because you also supply them with some data sets, right? Would you say that the one that's more common to happen?

UCSB2: So usually what ends up happening, you know for these GIS classes is I’ll give them a list of potential repositories for data and they sometimes… depending on their… you know, the majority of the students will just go to one of the repositories that I list for them and they just try to find something that works for them there. Every once in a while you’ll get a student who is a little more gung ho and can actually - realizes how you look for data on their own and will find something a little bit different - figure out what other repositories are or some other sources… especially for GIS geospatial data set...But the vast majority of the students, you know like I said, I sort of give them a list of repositories from the word “go” in the class or even as a part of laboratory assignments. You know, part of the laboratory assignments is going and getting data to then use in the laboratory assignment. And so, I will… they will just use other data sets they might come across when they are looking for other data that they are supposed to use for the labs, things like that.

I: Would you say that you face any challenges relating to students’ ability to find appropriate data sets?

UCSB 2: Not really. There's two places that I see challenges. One of them isn't a challenge, the other one is…One place there is a challenge, the other place there isn’t. There isn’t a challenge in the sense that there are plenty of repositories out there for students to go and get data if they want to. It may not always be like the precise data set they want, you know? Not everything exists in the world, you know, in a ready to go computer/data format that they would use. But where the challenge is getting them to understand how to actually go about finding it a little bit on their own. So even with fairly comprehensive explanations, trying to sort of give them tutorials, and walkthroughs on how you actually go about finding data and things like that when you're sometimes when push comes to shove they do not necessarily always find the data on their own. They require a little bit of extra hand holding, or they just don't do it, or something along those lines.

I: For the data sets you provide, there are cases, how do you usually make these data sets available for your students? Let’s say GauchoSpace?

UCSB2: On GauchoSpace. Usually what ends up happening is … the reason why… my choice is usually to try to get them to go and get the data themselves off of a repository to get them used to that idea. But, if I provide data the reason why I provide it is because the data set requires a certain amount of cleaning. This is like an intro to GIS class that I’m teaching where they potentially come in with zero experience doing any kind of data analysis whatsoever. Every once in a while there’s… you know one of the labs that I like to have them do I like to have them work with New York City taxi data that’s available. But those data sets it’s a month’s worth of every taxi trip in the city and it’s like a gigabyte or a 2 gigabyte text file. It’s too much for them to work with. So I slice it down a bit and try to clean it up a little bit so it's a little bit easier for them to deal with.

I: In terms of working with data, how do students manipulate, analyze, or interpret data in your courses? Would you say, could you mention some tools or software that you use for that for example?

UCSB2: The entire course is taught using Q-GIS. All the data analysis happens in Q-GIS and then there’s one extra assignment that’s a kind of extra credit assignment where I try to get them to use Grass. It either can operate as an extension of Q-GIS or it can operate independently. Typically I have them use it as an extension of Q-GIS rather than sort of having them learn the software out of nothing. But those are the two major software packages that we use.

I: Is it a pre-requirement for them to know these tools? Or do you expect them to…

UCSB2: No. It’s all taught as part of the class.

I: You mentioned that these tools are highly important for your teaching because it connects to the data that you use. Would you say that there are any challenges relating to students’ abilities to work with these platforms for your classes?

UCSB2: Well. Yes and no. The whole point of the class is for them to learn how to use the software in addition to how to use the data that they use in the software. It’s a challenge in the sense that the class is a challenge and I have to learn to use the software... you know, figure out how to parse these things out, right? Some students are more computer literate than others. Part of it is more advanced software than they are accustomed to using. It’s more complicated, the details of how you use it is more complicated than just using Microsoft Word or something like that. Some of them get it right away. Generally speaking, all of the students do fine. So far I haven’t had a student who really broke down because they couldn’t figure out how to use the software. Part of that is because, like I said before, I mean the whole part of the design of the class is to do both data and software hand-in-hand.

I: Would you say there are challenges in terms of students abilities to actually interpret or kind of explore the data using these tools - not connected to literacy with the tools, but actually dealing with the data and understanding what they are getting from these data sets.

UCSB2: There are some challenges...In terms of really getting…Usually they do ok with getting something out of it. They are able to export and create maps, I mean so it is all mapping software, right? So they are able to, there haven’t been any significant challenges in terms of the core functionality of the software and the core goals of the software.

I: How do the ways in which you teach with data relate to the learning goals in your field. Let’s say, if you can talk a little bit about how you teach your students to think critically about the sources and uses of data they encounter in everyday life and how it connects to the field that’s covered in the classes.

UCSB2: Ultimately the class is sort of understanding spatial analysis as it relates to human behavior. In the class I end up having all of these different lab assignments, which is the core work they end up doing in those core assignments as they are building up their knowledge in software they are also building up… dealing with more complicated data sets that help… or more complicated data sets that include more and more synthetic analysis. That has to do with how people behave within their lived environment and then also with just things you can locate in space. So, how do you think about human behavior happening in a spatial way? And so, ultimately the way that the classes are designed, the kind of data sets they end up interacting with, like I say, kind of build up over time to begin to...force them to think about some of those kinds of questions about - how the landscape or space affects people’s behavior and sort of being able to interpret synthetically a lot of that information. The key thing is, all social sciences - being able to think synthetically with some of the data and take in data sources from a bunch of different places that may not necessarily have obvious connections or sometimes have very obvious connections but, just getting them to create some a new conclusion data set or explanatory data set that allows them to make some kind of conclusion.

I: Would you say that these are specific data skills you prepare students for their future careers?

UCSB2: Yeah. Ultimately the goal with it, within research, certainly within archeological research, but then just more broadly anthropological research, you end up using GIS all the time. So, if one of these students was going on a field project, part of the skills that they are actually developing as part of their everyday work allows them to function - begin using GIS or begin being able to use the software in a productive way, as part of a field project, if necessary.

I: Have you observed any policy or cultural changes at UCSB that would influence the way in which you teach with data?

UCSB2: Can you repeat the beginning of that?

I: Have you observed any policies or cultural changes at UCSB that would influence or affect the ways you teach with data?

UCSB2: No. I wouldn’t say that there is anything that is UCSB-specific.

I: Ok. Would you say that other than on our campus, but outside, have any of these changes influenced your teaching?

UCSB2: The one thing I’ll say… So, right now, one of the things that I’m working into my class is some subject matter that has to do with or is related to Black Lives Matter teaching and trying to highlight how we might use some of these data sources to understand social inequities and things like that in an anthropological way. In the past I typically, what I was doing was pretty plain or vanilla...data analysis for the sake of learning the mechanics of it. But, with everything that is going on in the political climate without necessarily directly teaching the message, I am including things that highlight a kind of interpretation that would be seen as… you know, like I say, some of the things that would be in line with the Black Lives Matter stuff that is going on and racial equity and stuff like that.

I: Would you say that instructors in your field face any ethical challenges in teaching with data? Do you see any ethical challenges that affect the way you teach with data?

UCSB2: No. Right now there isn’t...well there are some ethical challenges that could exist. But I avoid them. You know, where ethical challenges potentially come up really have to do with using let’s say indigenous information or data that would have to do with indigenous groups that may not have been procured in an ethical way. But, ultimately what ends up happening is I simply totally avoid teaching that kind of stuff as a way to deal with it.

I: In terms of… now we are going to move to some questions of training and support. I would like to hear if anyone other than you provides instruction or support for your students in obtaining and working with data. So, for example if you have co-instructors, if you work with librarians or teaching assistants, or if you have drop-in sessions or presentations?

UCSB2: So right now I do it all myself. The nature of the class is it’s only, well I think there’s about 25 students enrolled in it and within the way the university does things they don’t give me a TA or anything for that, so I do it all myself. To be honest, specifically for this GIS class it’s so fast paced that there isn’t a lot of time to bring in a lot of outside people. If it was on a semester system or something like that, I could see doing it, but in 10 weeks we barely make the objectives that I set out for the course and there isn’t a lot of time for outside resource use. The only kind of resources that do come up is I do tell the students to potentially get in contact with the Geospatial Librarian at the university, or take advantage of resources that exist at the university. But, it is not, none of it is structured.

I: That ties to my follow up question about the ways in which your students are learning to work with data outside the formal coursework. You mentioned that you recommend them to talk to the GIS Librarian and also to learn about data resources that are available on campus, so how do you encourage this kind of extracurricular learning? Do you talk to students throughout your course? Do you have this listed on your syllabus? How do you usually encourage them to look for this extra help?

UCSB2: It ends up being one of these things where there’s certainly a lot of just conversation throughout the course...when there are sort of things or places that there would be...it would be useful for them to consider talking to somebody outside of the class, I certainly suggest it and bring it up and make reference to the fact that there are other places to look for some of this information. I encourage it throughout the class. Part of it is, in their coursework, one of the things that I often suggest is in the syllabus I have all kinds of external resources that are both associated with the university and also just kind of wiki pages on the internet that are good for learning GIS and where people ask questions to learn how to do things. I just try to encourage them to go...or think a little outside of the box in terms of how you might develop some independent ability to find sources and information about whatever it is they are doing. Part of that is even discussing the applicability of what we are doing beyond the scope of the coursework. Within the coursework itself I try to expose them to a bunch of different kinds of situations that would be associated with different kinds of work that would...so they sort of think about how does GIS...not just anthropologists, but if you’re a urban planner, or if you are a geologist, or if you’re a...I don’t know, some other professions, so how different professions might deal with this kind of material. They can think about where on their own if they decide to pursue it where you would find somebody else who is using this software, thinking about the software and things like that and the data itself as well.

I: Okay, and would you say they're looking for these external resources? Do you get any feedback from them?

UCSB2: Yeah, some do. I think that it's a Oh, it's very hit or miss, it depends on how engaged the students are really, you know, some students are more engaged than others. And so if the students are more engaged, certainly are more likely to go out and look for some of that material.

I: Okay. Now we are going to move to some questions about your own training. I would like to hear if you have received any training in teaching with data, other than your graduate degree. If you have participated, like even in informal settings such as workshops, or if you get help from your peers, or even like more formal types of activities to learn about how to teach with data?

UCSB2: Yeah, that's a good question. So at academic conferences, I have dropped a lot, you know, attended different kinds of workshops that talk about how you might approach digital pedagogies, and things like that. And that's been the primary source of external experience, where I've kind of learned a little bit about how one might teach data or teach things that have to do with data.

I: What factors would you say that have influenced your decision to receive or not to receive training or assistance?

UCSB2: Sorry, can you repeat that? It broke up a little bit.

I: What were the factors that have influenced your decision to receive or not to receive training or assistance on teaching?

UCSB2: It's ease of availability is really and even...the sort of awareness. I think that, you know, in terms of a lot of graduate school, or sort of things that are more directly associated with graduate school I haven't seen that many. So like, for instance, like at UCSB, you know, sort of, I haven't seen a whole lot of things, at least the sort of emails that I read, you know, one of these challenges where you get a million emails, so maybe there have been a bunch of emails that I've just passed over. But, you know, I haven't seen a lot of opportunities necessarily outside of academic conferences to sort of have discussions about that kind of pedagogy, let's say.

I: Okay, and do you use any datasets, assigning plans, syllabi, or the instruction resources that you received from others, let's say from your peers, or from other schools?

UCSB2: I use other...so the basis with the class, you know, I mean, it's almost any class I teach, I always look to other syllabi, to get some ideas of how it's taught. And so I certainly use some other syllabi by of how a class like this has been taught by other people, but then, you know, everything else, so I sort of, you need to actually select the data sets are all, you know, it's stuff that I, you know, found myself, I have identified for the labs and things like that myself.

I: And do you make your own resources available to others, like openly available so that people can have access to your course materials? Or students?

UCSB2: Well, I don't have, I don't have a big public website where I post all the material, but if somebody asked for it, I certainly would make it available to others.

I: Okay, still concerning training. What would you say are evolving trends in your field, and what types of training or assistance would be most beneficial to you and other instructors, about teaching with data?

UCSB2: Well, I think that there's the trend in the field is sort of broadly one where people are interested in like anything digitally, I sort of attach the word digital to it, and people are all of a sudden, more attracted to it, you know, in one way or the other. And so there's, I mean, there's this kind of thing where people are interested in data or sort of, you know, but one of the things is that the vast majority of the sort of tenured faculty or people who are, you know, been around in one capacity or another, don't necessarily, it's certainly at UC Santa Barbara, they don't necessarily always sort of know what to do with data, they sort of like, they use data, I mean, it's sort of an interesting thing, because they, it's not that they're sort of unaware of, you know, what sort of happens with data, but the sort of this kind of more modern computational trend, where people are doing things on a larger scale with sort of bigger data, so to speak. And using, you know, sort of dataset, I guess, geospatial data sets somehow factor into this, they're sort of understood to be a little bit more complex than just, you know, sort of something that you would use a spreadsheet for something like that. Yeah, that so that, that sort of, I guess, sort of where things have been going and sort of this move away from just kind of really basic on your, on your own computer, spreadsheet analysis in Excel or something like that, you know, using, you know, some kind of older statistical package that sort of trying to use something that's a little bit more complex, in that, you know, might involve some kind of machine learning or, you know, even though that's not really what necessarily happens. Those kinds of things are what people get a little bit more, it's certainly, for instruction gets a little bit of attention, because I think they're interested in developing translational skills. Since there's this recognition that, you know, not everybody with an anthropology degree is going to end up being a university professor or anthropologists, and there aren't, you know, there's, there's not a lot of jobs out there that have to title anthropologists. And so it's kind of, you know, sort of justifying the degree or sort of some of the skills that can come along with it in the wider workspace and sort of how you could teach anthropology to students while at the same time developing skills that are useful outside of anthropology or something like that. That's a little bit of a rambling answer, but...

I: No, it was great. So in terms of types of training, could you list or name a few things you'd like to have available on campus for instructors for teaching with data...

UCSB2: Yeah, I think I mean, to me, so I'll tell you, like my, like, where I have the hardest time or….but and this is actually I mean, this is, you know, this is what is this taking also sort of admitted this extends from a very wide definition of what is data in the sense that sort of like anytime that somebody, a student is trying to is this like, as almost as much to do with just how a student makes an argument is anything else. But you know, how it deals with data sort of, has a slightly different tack than if you were teaching it for like a writing class or something like that. So the key thing that I'd like to see is sort of more discussions on how you sort of teach students to be synthetic in their analysis. I think one of the things that often they run into, over time, as you know, sort of like students are very good at following, you know, directions in a very, you know, sort of descriptive, like a normative way where you sort of like, say, do this, do A, B, and C, and they follow, they do A, B, and C and whatever else, but then once it comes time for them to sort of take A and B and combine it with C and D, and but you don't actually give them D, that's where, you know, kind of some of the sort of challenges sort of common, so it ends up being one of these things, sort of like teaching some of these things with the aim of, you know, stoking some kind of imagination or sort of synthetic understanding of whatever is going on so that they will feel like they're really like learning the the core elements of what it is that you're teaching and then are able to actually apply it in a concrete coherent way that's a little bit more complex.

I: Great. We are actually at our final question now. And I would like to hear if there's anything else from your experiences or perspectives as an instructor, or on the topic of teaching with data more broadly that he would like to share?

UCSB2: No, think it was a pretty wide ranging conversation. I think that. Having some kinds of, you know, resources that are specific to teaching data are, would be, it is sort of like an interesting one. I'll just say that, like, sort of one of the challenges for me, even in some ways with, sort of, maybe this came across a little bit in this early conversation is that, you know, the class that I specifically have been teaching that, you know, is sort of truly this kind of, like data related class is both teaching software and the data at the same time, which has its own kind of unique challenges. Because, you know, you're sort... there's, they require two different, you know, sort of mental states in a way to, you know, fully understand and, it's the software itself is entirely very complex and then the data is also really complex. And, you know, I, in some ways, I mean, I've talked a little bit to [inaudible] about how you might separate the class that I teach into two quarters, which would maybe, like, help out that a little bit. But at the same time, I'm not sure that there's really a desire to make it a two quarter class. And so what do I think, you know, on campus, sort of one of the things that I sort of be curious to sort of, like, see more of, or sort of understand more of is sort of like how, you know, that unique element, just sort of things where it is, I think this is probably true for if you're like thinking about teaching.I would say if you're teaching a statistics class, it's a similar kind of problem, because oftentimes, if you're really teaching a statistics class, you're also teaching something like Stata or SPSS, or R know, whatever statistical package you're using to use to process this statistical information, which requires teaching the students at the same time, that kind of thing. And I'd actually know that that's like something though, that I don't really, I'm curious, you know, sort of like how that kind of like….there's where the pedagogical elements or how these things are combined. And I think that the challenge for me is that, unfortunately, geospatial data is more complex than just a statistical dataset, just because it's, in addition to having statistical information, you also have all this geospatial stuff. And so you have to teach them about geospatial stuff on top of the statistics, and it's sort of a lot of work. And it's sort of hard to, you know, hammer in some kind of hard prerequisite that would….sort of in terms of like, what they really should know, to succeed in the class. But, you know, again, like, sort of this idea of teaching both data and software at the same time as its own kind of thing would be sort of like, also, like an interesting thing to learn about, or teach or sort of add some kind of exercises in or something.

I: Okay, so you mentioned that it was an idea to split into two quarters. And in this case, it would go about like, the explanation of the tool and how to use it, in a more general context, and the next quarter would be more engaging with data, that would be the plan, or how would you handle it?

UCSB2: In a more ideal setting, I think that I would spend more of the first quarter teaching the software and the second quarter would be more on sort of the data analysis, you know, so you're able to go from, because the challenge is, is that, you know, by the time that the 10 weeks are up, they're just really starting to feel comfortable, I think, in a serious way with software. And we're trying to end because it's a one core class and are also trying to deal with more complex datasets and more complex processes within the software, which they can usually handle on a basic function. But you if you really want them to remember what they're doing, or sort of have some kind of, you know, have this actually, have a more lasting effect, make spending more time, you know, specifically on the software, learning some of the theoretical elements of what is geospatial data...and what is like, what is how is that different from like other kinds of data sets? And you know, how is geospatial software, or software that can handle geospatial data inherently different from these other kinds of data sets? And then kind of moving on in a second quarter to like, some of the specific data elements would be certainly useful for the class that I've been teaching. And it's actually, you know, I should say that, you know, in the geography class, or geography department, they also teach GIS, but they teach it as a three core sequence. And I think they do it, you know, for that very reason for the reasons that I'm describing. But in anthropology, it's because it's more of...because GIS is more, is even more than just a tool, as opposed to the basis for...the theory, like map theory is not that important for an anthropologist what ends up being more important is that you're able to practically use the software. And so the class is very much a practice or nothing more than a theory oriented type of class.

I: Yeah. So, thank you so much. I'm going to stop recording now.

Interview 3

I: So, just to start here, a note for the research with covid 19 disruption. We just wanted to acknowledge that changes teaching and learning. So, for any of these questions that I'm going to ask you can either answer with reference to your normal teaching pre covid or your teaching practices now that you've adapted or both whatever you feel like.

All right. So, first we just want to start getting some background. So, could you briefly describe your experience teaching undergraduates, including how your teaching relates to your current or past research, and in which of the courses you teach do your students work with data?

UCSB3: So, I've taught two different undergraduate courses here. The first one is [intro to econometrics]. And I'm not teaching that this year but I have a lot in the past. And this is basically the application of statistics in economics, so it's entirely about learning methods to analyze data, which is what I do in large parts of my research is analyze data. And then this fall, I'm teaching a brand-new course remotely, asynchronously, with great difficulty, which is [data wrangling for economics]. And this, this one is specifically about handling data. More handling than actually analyzing though we do a little bit of analyzing. It's trying to teach students skills for how to take the kind of dirty data you get in the real world and massage it into something that's usable for analysis or for making decisions, that kind of thing.

I: Great, I guess we'll have to talk to you more on that because that's also the kind of workshops we try to put on the Collaboratory So yeah, that's good to hear.

UCSB3: Alright.

I: So, as far as those two courses go, the next question is how do your students usually obtain the data they engage with? So I'm going to ask questions on how they get data. You know, then how you work with it in the course and then what kind of training and support there is. But first we're wondering how they get that data. Do they collect and generate it themselves, do they search and identify pre-existing data, or do you provide the data sets for them?

UCSB3: We provide the data sets for them.

I: You provide them. Okay. Can you describe the general process that students go through then?

UCSB3: In the introduction to econometrics class, we teach statistical or econometric techniques. And we then give them some mild instruction on using an econometrics software package. We supply them data and then we ask them to do a particular kind of analysis, basically. So, the idea is that in lecture they've learned some theory they've seen some examples. And then they're being asked to apply what they learn. In the new course that we're doing the lectures are all asynchronous. And what we do is we supply them with data and ask them to do a specific analysis which requires them to write some computer code, and then requires them to write up the results that they get.

I: Okay. And then, have you found in either of those courses that they’re may be different challenges or the same challenges, that students have challenges with?

UCSB3: Yeah. Very, this is going to be, may sound a little strange, but students seem to be decreasingly computer savvy. And I think some part of this is the shift to everybody being on cell phones all the time. But I have had, I'll give you an example from an earlier course, I had a student come to my office and say okay I've got to turn in this write up. And I got this output on the screen and the econometrics package. How do I get this stuff into my Microsoft Word file? And I said, copy and paste. And he said, oh, how do you do that? And literally he was so thrilled when I could show him how to copy and paste. And I'm not the only one who's had this experience. So the course I'm teaching now where the delivery mechanism is R, and we are assuming that students have never coded anything before. There are some of the students who are in total, complete bewilderment as to how to think about writing a piece of computer code, and you know we give them lots of examples and we're trying to do hand holding and. But for some of them, it's just a complete mystery, and some part of this is in both courses but more, particularly in this new one. It's not about memorization. It's not. I mean, you know, we give the lectures, but it's not about regurgitating what's in the lecture, it's being actually asked to go out and do things. And I think in a lot of courses here at UCSB, partially because class sizes are so large, that's actually rather unusual. And for some of them, I think they're having some trouble learning how to deal with this.

I: Yeah, and I know it's it's definitely a different mind space that you have to be in like, you know, coding and programming requires a different kind of logic than you have to use, and other you know other circumstances under that was it took me a little bit to get my mind around that when I started.

UCSB3: Yeah, I mean, that's really true, but I think I'm saying there was actually something a little bit more to it. In a lot of courses, you're shown some material, and you're then asked to be able to feed that material back on an exam, maybe in you know a slightly new problem you haven't seen before. Whereas, what we're doing is we're saying look, here's the task you'd like to accomplish. Here's an example of that kind of task and how it could be done. Now here is a new task. You now have to figure out which of the tools you've learned to apply and so forth. And for some people that is just fine but for a lot of them. This idea that they have not been told what to do, is actually pretty hard.

I: I see, I see what you mean there. Yeah, and I know I TA’d for some courses in grad school and we did a lot of HTML coding. So, you know, I think more people now are familiar with that, than like R coding. But even then, I mean, it was unzipping files that people had difficulties with.

UCSB3: Well right so we're having students, we're very fortunate that Google donated computer time and so we have a server running in a Google Cloud, and students write code on that. And then they have to push a button marked export to download their code to their own computer so they can turn it in and things, a bunch of them have no idea how to do that. And having had it explained they still don't really know how to do it. Well, you know, if you've never used anything like this on a computer before, then it's, you know, I wouldn't want to criticize somebody for not easily knowing how to do that. But it just seems very strange. So I have a colleague in the English department who tells me she's had students write term papers on their cell phone.

I: Oh yeah, I've heard that too. I was reading a piece a couple weeks ago that was kind of talking about this problem with students that are coming in that are not as computer savvy, and they're talking about how it has a lot to do with how you know UI design has gotten so good that you don't really need to do much yourself anymore in these apps.

UCSB3: Yeah, I think that's right. And this is something which has sort of surprised me because I just assumed that over time, more and more people would be computer savvy and it does not seem to be going that way.

I: Yeah, that's interesting. I think I was caught right in the middle of that transition there. But getting back to, you know, the data sets that you're working with. How do you go about obtaining those data sets that you use for teaching and vetting them? And then do you find any challenges in finding data sets for teaching?

UCSB3: Yes, so we do find some challenges. So, I try very hard to use real data, rather than toy examples. And because it's their economics classes we typically use economics data or business data but mostly economics. The particular challenge that makes stuff hard to find is it would really be nice to get data from businesses. And they are almost never willing to do that because you know I don't know exactly what they're worried about but there's nothing in it for them, I guess.

I: Yeah that's interesting. I tuned into the very end there's like the, I think it's a new center but Center for Responsible machine learning, on campus and they did a little conference last week, and a little bit I was tuning in on there talking about how they were actually able to track Covid cases better because they had gotten data from insurance companies. I was just sitting there wondering how they managed to convince the insurance companies to give them that data.

UCSB3: Right. I should say that this has been a big change in economics research, more and more people have been doing things to get access, sometimes confidential from industry sometimes from governments. But when you're doing it, I mean so sometimes people will invest a couple of years trying to get data. You can't do that for teaching. Right. And so, lots of data is, you can get access to it but only on a confidential basis well that doesn't do any good for teaching. So one of the best datasets, I actually, I happen to have is, you know, [name redacted].

I: [clarifying question]?

UCSB3: No, you know, [name redacted]. He runs the, I don't remember the name of it but institutional research. Great guy is actually a [PhD graduate] originally I think I saw years ago. He basically gave me data on all the student grades for a year. And now, of course he very carefully took out the names which we care about, and put in fake identifying numbers and we, you know, did some other things to protect privacy, so it was perfectly safe. But this lets me, in teaching, use actual grades and talk about how different departments grade and things like that. Well, you know, that's really good for student motivation, something they can really identify with. And so that's an example of something that's really good but it's hard to get that sort of stuff.

I: Yeah, I know we've been in for library workshops aimed at the librarians we've been trying to use more of our own data too but we have to do that as well because if it's like the door counts or, who's swiping cards, you know, we all have to do de-identify all of that before using it. Yeah. Okay, well, I'm gonna move on to the next question, which is more on working with data. So generally, how do the students manipulate, analyze, or interpret the data in your course? You know what kind of tools or software do they use? Do you expect them to come in with prior knowledge, things like that?

UCSB3: Okay, so in the econometrics class that I'm not teaching at the moment, we use Eviews, which is an econometrics software package. We certainly do not expect them to have any background in using it whatsoever. In the class that I'm teaching now it's been done in R. And we not only don't expect them to have any background in R, we sort of assume they've never programmed anything at all. We, to some extent, kind of expect students may have used Excel at some point. We don't, I don't use that very heavily. And it turns out there are quite a few students who are not very familiar with Excel which I find a little shocking but there you have it.

I: Yeah. And with those courses are they typically freshmen?

UCSB3: No, no, no, no, they're their upper division courses, the econometrics class is a required upper division class. And this new class I'm teaching is an upper division elective.

I: Okay, so that's pretty surprising that some of them haven't used Excel much. And then, to what extent are each of those tools do you feel they're important for teaching? I mean, it sounds like the econometrics and R are kind of the core of the courses but yeah.

UCSB3: So, so, right. So, using Eviews, which is widely used all around the country for the thing. Some of the other programs which people might substitute for STATA. I use Eviews because I think it's better, I’m familiar with it. And then when for this course I'm doing now we chose R. Mostly because it's very widely used in data science, and some of our accounting firms that help support the department had expressed some interest in students, knowing more, knowing some R. And actually, the accounting firms donated some, one in particular donated, some money to help develop the course, which was very important. Yeah, when you teach a course which is very heavily tech.

Well, there's a lot of tech involved in getting the damn thing running, that's especially true doing it remotely. So, we had a crew of three graduate students, and two undergraduates working all summer, putting together material, figuring out the technology and all those sort of things like that, and it was invaluable, we would not have worked without that, in fact, I can tell you that we, so we have a nectir channel. Are you familiar with it?

I: Yeah, I've heard about it, I haven't tried it.

UCSB3: It's kind of like Slack, except that it's apparently FERPA compliant. The event, since the term began for this course, We've had 1800 messages from students there.

I: Oh wow.

UCSB3: Yeah, which in some ways is great. So, we have actually three undergraduates who are each spending I don't know something like 10 hours a week, hanging out to answer questions from the students, and that's invaluable. One of the things that's much harder right at the moment is we really really encourage students to work together. You know they have to all turn in their own work but they should certainly be talking to each other. And that's very hard to arrange under current circumstances because they can't physically get together.

I: Yeah, it's something that's hard with our workshops too. So those, the undergraduates that are, you know, supporting nectir is that for just the one course that you're teaching or all of the courses?

UCSB3: That's all for my one course and we’re at the end of the second week.

I: That's incredible.

UCSB3: Yeah, yeah, we, I mean, we couldn't do the course without what they're doing for us.

I: Yeah I have had no clue how you know the different departments have been adapting for teaching because we've, unfortunately, had to put so much time into setting up our own remote lab. It takes a lot of time, I believe it took you know, all of those students to do it the whole summer.

UCSB3: So, I've been fortunate because my department chair has just been very supportive of giving us the resources we need to make this work.

I: That's good to hear. I know as since I'm new to the university, It seems like it's a little disjointed between some of the departments. They’re run so differently.

UCSB3: Yep. That's a very nice mild way of putting it. Yeah.

I: Alright, so the next question I have here is how do the ways in which you teach with data relate to the learning goals in your field? Do you teach them to think critically about the sources and uses of data or other specific data skills that prepare them for future careers? I know you touched on R, so anything else?

UCSB3: One of the things in the econometrics course that is very central is to help students, let me give you the cliche version which is that correlation does not imply causation. And trying to get that through to people is unbelievably hard. And so that's an important learning goal. But for me, an important learning goal in both these courses is that students walk away with tools they can actually use themselves, not just idhired of the business and could go run a basic regression and interpret the regression. In this new course, they are learning R. It's kind of a, in principle, we're trying to talk to them about the concepts of how to mess with data. In practice, they're spending a huge amount of their time on R, but I explicitly said there's a final project. And the final project will be a little longer than the homework they've done and one of the goals, is that they should end up with some code and a written document that they could use as part of a job interview. So they can actually show something that they have done.

I: So, do you encourage them to put it up on GitHub, so that it's public or to have a place where they can share it from?

UCSB3: No, we don’t. We, internally, we're using GitHub, but we've stayed away from it as one too many tech tools to teach.

I: Yeah, I can, I can imagine that would be overloading at that point. Yes, and I, and I meant to ask earlier. Do you have push from, from businesses or anywhere else for using Python or is it mostly R?

UCSB3: Well, we've had very little push of any kind, the push we have had was sort of on the level of R. Now the computer science department teaches Python which I think clearly is the right decision. Because that's a very good way to just teach the principles of programming. When I think about the course we're doing, which it's labeled explicitly as we are not doing big data. Okay, I'd like to just learn how to do small data.

I: Which is a much better place to start when you're teaching programming for the first time.

UCSB3: Yeah, I want to say that the actual programming in this course is pretty minimal to somebody who knows how to program. So, it's like the next to last lecture or something that we're going to explain to them what a loop is. And one of the reasons for using R is, to a large extent, you can use it in the same way you would a statistical package, which is, you just have a series of commands that do one thing after another.

I: Yeah I know in our R workshops we're doing one right now, actually I was teaching that earlier today. We rarely, if at all, use loops at all because with R there are so many packages available. You really don't need to know how to do loops and that's one of the I think being big hang ups with programming

UCSB3: I agree, yeah, absolutely.

I: All right. And then kind of a follow up question but slightly different direction. Have you observed any policies or cultural changes, either here at UCSB or in general that influenced the ways that you teach with data?

UCSB3: No, only one, and I'm not quite sure this is responsive, but I do make an attempt to see to it that there are examples that are I guess I'll use the word inclusive. I'll give you an example of what we're doing in this new course. All the students are told that they have to take an implicit bias test which we provide online. And then we're going to take that data and have them analyze it. It's not that the analysis is messy, the analysis could have been done on some other data set. But it seemed like a good opportunity to bring in some other lessons in a very gentle way.

I: Yeah, I think that sounds like a great activity. Because you're, you know you're bringing in the implicit bias and doing that training but you're also analyzing the same data. And, like you said, like having a connection with the student grades, that helps with motivation too.

UCSB3: Yeah, yeah, that's right. Yeah.

I: And then, you know, slightly different question, but along the same lines, do instructors in your field, you know in econ specifically face any ethical challenges with teaching with data? I know we mentioned, you know, anonymization being one of them.

UCSB3: I would say, anonymization. I can't think of really other ethical things that I would call ethical challenges.

I: I guess, do you have, is it often data about individuals?

UCSB3: Yes, very often about individuals. So, in the econometrics class, a number of the examples I use data with from what's called the Current Population Survey. You probably know about that. And I often do things like gender wage gaps or things like that. So, I will always say much of the data is actually about individuals but never identified individuals.

I: Okay, never identified ones. So, do you end up teaching about anonymization in your courses?

UCSB3: Nope.

I: Alright. So then, I mean the next questions here are on training and support for those courses and I know you mentioned already, the students that you have helping out at least you know with these asynchronous courses. But do you have any other people that provide instruction or support, you know now or before?

UCSB3: We always have graduate teaching assistants, who play an incredibly important role in meeting with sections, you know, answering student questions, that sort of thing.

I: How big are the classes usually?

UCSB3: The econometrics class is typically 200. And this new class, I think we're down to about 165. Some of the students decided to drop out, which I can understand. Yeah, and so the new one, one of the things that's going to be a challenge going forward is you need to really, I think, need to change the homework each year. Because, as you probably know everything floats around, nowadays. And developing these homework is no longer just a matter of getting an idea. You have to get an idea, get some data, put together. And then, we’re using automated grading for the computer code. Which means somebody actually has to write the code that does the coding, the grading code. And so it's going to need on an ongoing basis, some more technical support than most of our courses.

I: Right. It's such a lifesaver to have those though but yeah it does take a lot of work still to keep them up. Do you encourage them to go to drop in sessions around campus or in the Econ department, I don't know if there are drop in sessions in the econ department that are available.

UCSB3: Well, in the econometrics class they typically are not, I mean the graduate students and me, will have office hours. And students are certainly, you know, welcome to go to the office hours and they generally don't. In the new class, the original plan before we all got Covid was, we have an undergraduate computer lab. And that was going to be staffed with some undergraduates as many hours as we could manage. So that exactly that would be a place for students to just drop in with questions, well we can't do that. So that's what this nectir channel we're using as a substitute for, and I gotta tell you the undergraduates we've hired to do this, they're fantastic.

I: we have a couple undergraduate student workers too and they're wonderful.

UCSB3: But you know there's no substitute for personal contact. Yeah, it's no great substitute, and, you know, being able to sit down in front of a computer screen and say, here's what I did, I don't know what's going on, what I did wrong. That doesn't work so well online. And we're actually just in fact I think, with some of the undergraduates’ advice, we're going to add in some more dropping office hours, where they'll do zoom sessions instead of just this nectir channel. So we're trying hard to keep the students as connected as possible. Once a week, I have an hour, which I have labeled relaxing beverage hour, one student at least has actually taken me up on, where it's like office hours but actually the rule is, we'll talk about anything except for coding and R and the homework. And so it's a way to get to talk to the students about whatever they want to talk about. And so students come in with all sorts of just different questions which is nice.

I: So, is that for your undergraduate student workers or the whole class?

UCSB3: No no no no that's for the class.

I: Oh, that's wonderful.

UCSB3: Not that many of them show up, but more show up than used to show up for office hours.

I: Yeah, I think everyone's kind of missing the lack of connection and like conversations from the beginning and end of the class are really like are part of the learning process.

UCSB3: They very much are.

I: Yeah, we've had to do a lot of workarounds for these carpentry workshops too, you know, a lot of times such simple questions that people have or like just a spelling error in their code, and unless you can see their screen, you would have no idea how to, how to help them.

UCSB3: Yeah. So, one of the things which nearly is I can tell from what my students are explaining to me on how they've arranged the technology. So, I don't know if you know UCSB uses a product called Gradescope for grading. And okay, I realize you guys are not grading.

I: No we're not. And I'm pretty new here too.

UCSB3: Yeah. And, but part of what this is, this is a place where students can actually upload their programs. And the idea is the machine will grade it but it also turns out it means that the teaching assistance, and my staff can go in and they can actually while they're trying to answer a student question they can see what the students submitted so they at least have some shot at seeing the code sometimes.

I: Right. And then, do you know if there are other ways that they're learning outside of the formal coursework here? Like for the students that do come in with some knowledge too, do you know of the ways that they're learning or is it self-taught?

UCSB3: I don't know, I mean I will say, one of the other things we did as it was one of the, my students who suggested this something called I think Groupme, where we set up something with the students can actually get together in groups, which we cannot participate in which is good, so they can talk without worrying about us, you know, what we're thinking about them. We are putting up a few links to some other videos online since there are lots on R. It's also clear some of the students have come in with significant backgrounds, some of them know the R coding just fine. But I think most of them don't.

I: Okay. And do you often encourage them to watch those videos? I mean I imagine you do since you put them up

UCSB3: Well, we've been encouraging them, and this is sort of new because again, this is the first time we've done this course and we're sort of learning. One of the things we're learning is that getting started doing this is even harder than we thought it was going to be for the students. So we put a lot of effort into doing hand holding, especially at the beginning. And it turns out, it hasn't been enough. So as we're going we keep trying to add some more hand holding things.

I: Some more materials?

UCSB3: Yes.

I: Yeah, and I know getting started is also a challenge for these workshops we put on and you need to be very explicit and, you know, especially online if you skip a step then then people are lost and you don't even know that they're lost.

UCSB3: Yes.

I: All right, and then as far as your own training goes. Have you received any training in teaching with data or, you know, any more data analysis techniques or software since your graduate degree?

UCSB3: I'm not quite sure how to answer that. I've helped develop some software. I mean, you know, developing techniques for statistical analysis is a half or two thirds of what I do for a living.

I: Okay. So, this question is probably not as related to your work since that's, yeah an integral part of it. Yeah. I think I think yeah they meant for this question, you know, other conferences or workshops or, you know, institutions that put on training.

UCSB3: No, I mean I you know, I never programmed in R more than about three lines before teaching this course so you know I went out and bought some books and learned and taught myself.

I: So, you mostly use books for that?

UCSB3: Yeah, I, you know I watched some online videos that I thought were useful. You know the folks who do R studio have lots of good resources.

I: Yeah, Rstudio has great resources. And do you end up, do you access them through the UCSB library or get them for yourself?

UCSB3: I just get them for myself. Yeah.

I mean we do these workshops, but I need to go through and look more at what we have in the library here. I just haven't quite had time to do that since I've started working here.

I: That's good to hear. Okay, the second to last question, considering the trends in your field, it's less on you know, the training that's available but what should be available. What kinds of training or assistance do you think would be beneficial to you, or other instructors in teaching with data.

UCSB3: It's very clear, it's the tech support sort of not not generic but focused for people helping create the course and create all the technical stuff we need for the course. And then the other issue is access to computing facilities for the students. So, you know, we are managing somewhat because Google made this nice donation to us. Otherwise, I don't know what we would do.

But this of course relies on the students all having good internet access and having a laptop or something anyhow. And so, that second part could be a bit of an issue but it's really the people costs that I'm concerned about. Yeah.

I: Yeah, I know. You know, we've set up this remote lab because we also got a gift from Microsoft Azure, very timely for all of this for remote computing but I'm surprised at how little there was for remote computing, when I got the campus here.

UCSB3: That's very true. And they may, I mean how far they've come in six months is amazing.

I: Right. It's really prioritized. Yeah, set the priorities. But the manpower is a hard part of it for sure, and you know it doesn't help if like you said, the students don't have access to the internet or a computer that can connect.

UCSB3: Right. And you know, let's assume that next year, this will all be over and we'll be back in person which will be great, but then we're back at the issue of computer labs and things like that, where students can get together and where they can actually get help. So for example the open access labs on campus are generally only open Monday through Friday. Right. And you know I understand this is a cost reason for that, that's not criticizing, and they’re also open mostly just during the day. Well, it would be a big help if there was just a lot more accessibility to computers and stuff like that for the undergraduates.

I: Yeah, and I know on my part, it seems like the these open computer labs that are available on campus are, you know, so all over the place and decentralized that it took me forever to get a grasp of where there were open computing labs and you know I'm still finding out about different labs that are you know were available on campus before we went remote. And I don't know if the students have an easier way of finding that out then I do as staff but…

UCSB3: I mean, the letters and sciences actually runs a whole bunch of the labs and have a, well not a very good website telling you where they are when they're open and what they have. And then of course different departments have their own labs that are just for internal department use.

I: So actually your undergrad lab and Econ, how many computers, do you have there?

UCSB3: I'm not sure I think they're 35.

I: Okay.

UCSB3: And again, a lot of this was we have some donors who have given us some money to do some of it so you know we're fortunate in that way.

I: That always helps. Yeah. Well, those I mean those are all the questions I have and the final thing is just, is there anything else from your experiences and perspectives that you can think of that might be relevant to this or something that you've thought of from a question earlier that you'd like to share?

UCSB3: No, no, I, I think you've covered most of the stuff. I guess I should say, by way of background. That economists typically do not do huge amounts of original data collection.

I: Okay.

UCSB3: We tend to use data that somebody else's collected. There are exceptions to that. My daughter is an economist and runs a survey crew in Nigeria. But that's why some of my answers might be different from some of the other social science groups and anonymization just doesn't come up all that much compared to places where you're well, compared to like what you're doing right now.

I: That's really interesting because yeah I'm not familiar with the ins and outs of, you know, economics data or, you know, analysis versus other departments, it's just it's you know it's it's not what I studied and, and we have a lot of the econ related software in the Collaboratory so it's hard because you know we have the software but I don't have the knowledge to know how it might be used.

Interview 4

I: All right. And then I just wanted to start off here with the note regarding you know covid 19 to acknowledge that obviously teaching and learning has been significantly disrupted in the past months. So, for any of the questions that we're asking you can answer with reference to your normal teaching practices before this started, or your teaching practices that have been adapted for this situation.

UCSB4: Okay

I: Either are fine. And then just want to start out the interview here with some background. So, if you want to talk about just generally your experience with teaching undergraduate students, especially, especially regarding how your teaching relates to your current and past research, and then which of the courses that you teach, do you actually have students work with data?

UCSB4: Okay. So, there's a variety of courses that I've taught here at UCSB but I've only ever taught here at UCSB. So, I've tried a couple of different types of courses, so I have a larger, upper division undergrad course, which is supposed to be taken sort of between the very large lower division courses and then the smaller upper division courses. And then I've taught a couple of smaller upper division undergrad courses. Couple of grad courses. And then one thing that I particularly had in mind, with regard to this interview and the data use in the classroom is an interdisciplinary honor seminar which is not through my department but it's the [course code].

I: Okay.

UCSB4: And that's very small, it's undergrads but only those who are in the university wide Honors Program whatever that distinction is there with the department level Honors Program. So, in the large upper division undergrad course, it's not so much that I'm using my personal data from my research, should I explain my research?

I: Yeah, you can explain your research, you know, not necessarily what you use in the courses themselves, but just generally your current and past research and how that relates to your teaching or not. As much as you want to get into it.

UCSB4: Yeah, well I figured just for context that could be helpful. So the data that I collect is on [primates]. However, when it comes to publishing, I will then sort of, whenever possible, get additional sources of data. For instance, from other researchers other publications, and then combine that with my own data. So, the conclusions I'm trying to reach in my publications are not, I don't want them to just be applicable to my study groups and study species. I want them to be general principles for primates, at least, and ideally for animal behavior writ large. So, the data that I use for this large upper division undergraduate course, it's not the data that I collect from the monkeys explicitly it's more sort of the general patterns that come out from doing the combination of the literature and my own data. And so it's just broader results at that level. And the reason I do that is because primate behavioral ecology research is driven by a set of hypotheses that are, I think, fundamentally flawed. And this is part of my, what do we call it, I'm trying to give undergrads a different experience than what I had when I was an undergrad, where I had the impression that everything was known, and there was nothing that I could contribute to the field. And so I try to show them the older hypotheses, the gaps in the hypothesis, and you know how those gaps are illustrated by certain examples from the collated data. And, you know, try to get them to think about what's missing from the original hypothesis, why is it not working, why is, why are there all these species that are not conforming to the predictions? And so that is to get them to think about, you know, this is still a growing and developing field, perhaps they have some contributions to make and kind of trying to get them excited about it.

I: That's really interesting.

UCSB4: So, Yeah, I mean there's this push and pull of getting them interested in the observations, the hypotheses. Thinking constructively about the patterns that are being observed. And at the same time I feel kind of like a hypocrite because doing private field work is extremely challenging and limited. Very few people have the opportunity to actually go out and do that. So on the one hand I'm saying you have something to contribute. On the other hand, it's very much a pay-to-play system and very few people have the resources to get involved in that. So, yeah, So there's that. Then in the smaller upper division undergrad courses, I guess I'm not really using data in those courses so much, but the place where I most intensively integrate data into the courses the [small interdisciplinary honor seminar] and the students go out and collect data on animals in the local campus environment, things like that. And then I teach them how to start looking for patterns in that data, graphing it, making conclusions.

I: Okay, that actually feeds into the next question here, which is more along the lines of how you get data for your courses. So I don't know if you want to expand on that honors course a little more and how they collect data. But the question here is, you know, what's the primary method for how you provide data? Do you have students collect or generate it themselves, search and identify pre- existing data, or do you provide them datasets? And maybe you could expand a little more on that honors course but then also talk about that upper division one as well.

UCSB4: Yeah, okay. So, for the other seminar. I teach them the methods for collecting the data and then we go outside and they have pieces of paper. Basically, tables that they fill out while they're observing animals so they are literally collecting their own data. And then I have them enter that data into a spreadsheet, like on Google sheets or Excel, and we go from there so it's very raw, in that sense. Actually, there's another small upper division undergrad course that I'd forgotten about which is similar to the honor seminar and that the students collected their own samples that then generate data so this is a lab methods course, where I teach endocrinology, and assays, and they collect their own urine and they specify sort of sampling regime, so certain times a day certain number of days. And then, you know, we start, we've labeled all the urban samples, store them in the lab on campus, and then I teach them how to do radio amino acids and assays, and to look for cortisol urinary c- peptide and a few other biomarkers. So again, that stated that they generate and then I teach them how to make sense of that data. First is troubleshooting did they do the assay correctly and then if they did do the assay correctly, what did the results mean for their cortisol levels, you know, cortisol is a biomarker of stress both metabolic and social so expect to make some inferences from those patterns.

I: Oh, so you teach quite the wide breadth of courses, actually. That's really, that's really interesting too. So, for those two courses that you have students collect data, do you face any challenges with them being able to find or to create those data sets or, it sounds like the process is actually pretty simple for like creating a table.

UCSB4: Oh god no, it sounds simple and for some reason it's just absolutely not. So, I should also say that for the honor seminar, I take a little bit of glee in that it's a bait and switch. So, the course is called observing behavior and the, you know, the one paragraph description that's on the list of courses for the university is all about observing animals understanding behavior. It says nothing about spreadsheets and statistics. So, I use this as a way to get students doing things that they would normally run away from.

I: Okay.

UCSB4: And I have a deep love of data, and my perspective is that it changes how you think. And if they just go through a series of steps, hey haven't, you know, made a wrong turn, they haven't understood why it's wrong to do it one way and better to do it a different way. And so I want them to struggle with that process and figure out how to do it correctly. And in the end, they seem to really love it and appreciate that, the frustration was worth it. So, it is my favorite class.

I: It's really interesting to hear because I think often I hear more about people trying to limit those issues that students run into when working with data.

UCSB4: Yeah, Yeah and you know we do have another course. So, again, the honor seminars outside of my department but in my department, specifically in my unit of the department, there's a lower division undergrad course which is meant to prepare them to think about data in this way, and to become used to thinking about how we organize it and set it up so that we can test ideas. And it's very much in the formula as courses go, of a series of steps that the students do where, sure they extensively learn how to calculate a mean from raw data. But for some reason, there seems to be far lower rates of the students really grasping why they're doing what they're doing.

I: Okay.

UCSB4: I actually have not taught that course, so I don't, I personally don't really understand why the honor seminar, I think, seems to be a little bit more effective in that regard. I mean maybe it is a struggle that they go through.

I: And then, so what about the, you have so many courses, it's hard to keep them straight. That upper division course that you do use the monkey data. It sounds like that's data that you provide for them, then

UCSB4: Well so when I say data I guess it's more figures, and in all, it's not the raw data, in that sense, they're a bit more removed from the process in that respect, so they're seeing more the, the final product. Or I'll say for instance, you know, we expect females to participate in monkey wars under a certain set of conditions but you know here are 10 species that are supposed to show female participation and they don't, and then here seven species that are not supposed to show a female participation and they do. And, you know, these are some other factors that are involved so it's, yeah I've, I've already collected and processed the data so I know what it says, but I'm trying to lead them to that, that analysis through the figures.

I: So, that course is, it’s more about analyzing the data and critically engaging with it. Then, like working with the raw data itself.

UCSB4: Yeah, yeah.

I: Okay. All right. Well, I think I'm gonna move on to the next question, which is more, and I think you've described the process somewhat, but what is the process of students manipulating, analyzing, or interpreting data in your courses including what kind of tools or software they use? And what kind of knowledge, you expect them to come in with if anything?

UCSB4: Well, I have learned to basically have no expectations.

I: Okay.

UCSB4: Yeah, and it's not that I'm being excessively cynical about UCSB students. I remember as an undergrad I did take a statistics course and absolutely none of that made sense to me, even during the course.

I: I had the same experience.

UCSB4: And I did an honors thesis when I was an undergrad, and I had to relearn everything. Basically, there was a grad student who had to use entirely different words and concepts to teach me. And it was frustrating that of course that was supposed to have taught that or helped me learn that. It was so useless but anyway. So the courses here, I have the students using whatever spreadsheet program they have access to and usually that's Google Sheets. And I think it's really important for them not only to collect the data whatever possible, but also to enter the data because that structures a lot of how we use it right like people who've been using data for a decade or three. You know, we know basic things like you can't have a blank row in Excel right or a blank column, you have to have at least one cell with an entry. And it's, it's kind of delightful. What I show them is just little basic things that turned out to make things so much easier like pivot tables in Excel.

Because I try to tell them over and over and over again that even if they're not going to be an academic and go off to grad school and, you know, work with data for the rest of their lives, it's really useful to understand programs like Google sheets or Excel because that's a skill that you can take to many many many other jobs. So it's a translatable skill. And, yeah, so in the process of entering their data they always make a bunch of mistakes and then we go back and we fix them. And then we start sort of creating reports through pivot tables and inevitably there's a couple of students who don't understand what I'm saying, when I'm trying to, when I put up on the screen, you know how we go from a raw table to a pivot table and sort of create these summaries, and they end up doing it by hand. And then when they finally do start using the pivot table, that sense that they have of, oh my god this is so much easier, you know, why aren’t we learning this in elementary school.

I: Yeah.

UCSB4: Yeah. So, they do all have that. And I think that that part is almost as important as learning how to chunk the data, calculate averages and either standard deviations or standard errors of the mean and then graphing it. And then I teach them, you know, whether it's a bar graph or a line plot with, you know, confidence intervals. I teach them how to sort of do an eyeball scan of that figure to interpret. Is there likely to be a significant pattern here, and I'm using the old sense of statistical significance.

I: Okay. And do you find, other than the ways in which you've already highlighted, do you find that there are other ways that students are challenged in working with the data?

UCSB4: They have the syndrome where they kind of can't control their hands. I don't know if you've ever seen this but they, when they're not quite sure what to do or they're not, they may be a little distracted, their hands are sort of flicking around on the mouse pad. And so the spreadsheet is flying around, and it's like this negative feedback loop where once that spreadsheet is flying around, and they can't see like if they forget to freeze that top row so they don't, they don't even know what columns they're looking at. They then start to panic, that they don't know what to do and they're behind, and they're fucking everything up and etc etc. So, and this is why I haven't even tried to teach this course in the online version like you know I'm waiting until we can do in person classes again because that's when I say, take your hands away from the keyboard.

I: Yeah.

UCSB4: And, you know, start over. Tell me the steps and why you're doing each step. But there's this physical feedback with the hands and flipping around that I can't even wrap my head around. I mean, I don't know where that comes from.

I: I suppose and then they click something by accident and aren't seeing the same thing you are.

UCSB4: And you know it's one thing to sort of click around between programs but even within like if they're using Excel, they seem to not initially grasp the difference between a click, click and drag, a double click. All kinds of crazy things happening because they've just dragged a column over somewhere else and replaced the column, you know, and had no idea what they were doing.

I: Okay, I see. Interesting. Yeah like the scrolling or the accidental manipulation, that, yeah, that you don't know how to undo or they panic and don't realize they can undo it.

UCSB4: Yeah, exactly.

I: I see, and then moving on here. How do you see the ways that, I think you've touched on this already when talking about your research and your motivation for your courses but, how do you see the ways in which you teach data relating to the learning goals in your field?

UCSB4: It's a subset of the goals, and I think this is why in some of my classes I don't have them working with data at all, because those are a different set of goals, and the classes where I have integrated data, especially the extensive forms in the lab class and the honor seminar, that is the hands-on learning experience. So, the classes where I'm not using data. You know, these are classes where it's more of a big picture overview because, so for instance like the primary behavior, large upper division undergrad course, it's typically about 100 students, and especially this year I have more students who are not anthropology majors they're coming from ecology and psychology and MCBD and all these different departments.

I: And do you think that's just because the courses that they might have taken in their own departments are not available remotely?

UCSB4: So, enrollment was flagging within anthropology, for this particular course, and so I made a flyer, and I distributed that all over campus so that's why I have more. Departments represented in the courses, this quarter but yeah so the basic prerequisite for the courses that the students have taken some kind of intro to evolution and should, in theory, understand the mechanics of that, you know what is natural selection, how does it work, how do we know that natural selection is occurring in a population, things like that. And the thing about, of course like primate behavior is that you know there's depending on how you count, maybe 400 or 500 primate species in the world. And the reason that we study primates to begin with is sort of twofold. One is that they’re our closest relatives and so we learned about, at least in this course the ecological conditions that may have shaped human evolution, and we can sort of understand the possible routes that humans would have responded to those ecological conditions by looking at how other primates respond to those ecological conditions. And then the other sort of reason for studying primates is that they are hugely diverse in those hundreds of species. And so it's a great sort of testing ground for all kinds of theories. And so for that large upper division undergrad course, the data that they see is already sort of processed and set up in a way to get them thinking because, you know I am trying to keep a big picture. Many many primate species, many different patterns. With the courses where they are actually collecting data, learning how to enter it and summarize it and interpret it, that’s a very narrow window. So we're not talking about hundreds of primate species we're talking about one species at a time. And how do we start from the bottom up, start from the observations and how do we structure our observations in a way that we can actually test the research question. So, yeah, top down versus bottom up perspectives.

I: And do you generally find that your courses that are focusing on the raw, the raw data skills are smaller than the bigger picture ones?

UCSB4: Absolutely. The way that I teach them is very hands on for them and hands on for me because I have to do so much troubleshooting with them, you know, dealing with the frustration dealing with the inadvertent effects of twitchy hands. Yeah, there's you know scrolling frantically through pages and not thinking about what they're looking at. And I mean it yeah I wish that all my classes could be small, but I also wish that every student had the experience of being in something like that honors seminar.

I: And, and I know you touched on this with the Excel skills but do you also see that you are teaching them specific data skills that will directly prepare them for future careers either academic or otherwise?

UCSB4: So, the skills that I'm hoping that they learn that they can then apply to future jobs or studies is more about how do they figure out how to deal with a problem. Right, so maybe they're never again going to need to calculate and the standard error or the mean. But just sort of understanding that, A, they need a way to represent the data, and, you know, B, how do they go about figuring out the best way to do that without a statistics textbook at hand necessarily. So it’s the problem solving I want them to learn.

I: Yeah. And that part is so important, or just like thinking about things and in context of data. I know I've talked to grad students who you know are in PhD programs that don't have a grasp of that even. It's such an important thing. All right. And have you observed any policies or cultural changes at your institute, here at UCSB, or in your department specifically that influenced the ways that you teach with data? I don't know how, how many years you've been teaching these courses and if you've adjusted them based on policy changes or cultural changes.

UCSB4: I don't think that I've really seen policy changes. It's more for me, it's less of a temporal change and more of a difference between organizational units so my department versus the honors program right like there's a different approach there.

I: So, the honors program on campus here has their own approach for teaching their courses or for teaching with data specifically?

UCSB4: Just in the sense of course size. So I would really have to justify teaching such a small class within anthropology, and it's not that we have a ton of majors, but we don't have enough classes for the majors that we have. Anyway, the other sort of motivation for me personally for teaching the honor seminar, of course, is that there's like a little research stipend that comes with it for the faculty, and I can't use it for salary but I can use that to fund my research, and my research is constant and expensive, relatively speaking. It's a nice additional source of income.

I: Yeah. You know I'm not really aware of the Honors Program at all here yet. Even though I’ve been here a year I’m still learning a lot about UCSB. And then, do you find just generally in your fields that instructors or researchers in general face ethical challenges when teaching with data? And that's similar to the last question but yeah.

UCSB4: I don't know of any instances personally but I wouldn't be surprised if there were some just because. So, you know, again I'm an animal behavior researcher. I don't really study humans. But my colleagues do study humans, and you know they're applying for IRBs all the time and dealing with that. So I'm aware that there are those types of issues. I will say, in terms of ethical issues, whether it's data, or even papers that students are producing. You know, and each new quarter, students often want examples of what's expected of them, and they want to see, for instance, essays from the previous year. What's a good essay or, you know, what does the finished product look like when you're dealing with the data, and I don't want to show them something from a previous class because I haven't gotten permission from those students in the previous quarter to show it and future quarters. So there's that aspect of it.

I: And is that a concern with you for sharing data from previous quarters too?

UCSB4: Oh yeah, no I don't share it at all. So, yeah.

--- 30 minutes in ---

I: So, and I imagine that specifically with like, these challenges aren't as important to your teaching because as you said you're not. Well, other than the students, you're not working with human data. That's with animal data.

UCSB4: Yeah. Now the lab course. I've only taught that once. It was, it's something that we're supposed to now be teaching every year but we're trying to hire an instructor for that but the course is quite a lot of work. And in the course of the students collecting their own samples and labeling them and assaying them and collecting your, you know interpreting the data. I did sort of have at the back of my mind. Oh, what happens if we see that so what happens if we see that might be embarrassing for a student. And thankfully in that first quarter there, there weren't any such examples but I guess there's a possibility for that in the future so…

I: That thought popped into my mind too, so are all of the samples available to that, like, does the whole class see the results of everyone’s sample or is that person's sample private to them?

UCSB4: Well, that would be one option but I for various reasons I have them working in small groups within a small group they do see each other's data and typically in class what was happening is that, you know, a student would run into a problem and therefore this class time that was dedicated to ironing out these issues and so often they would share their screen in class and say you know this is what I saw how did I do this wrong and so you know we'd walk through the troubleshooting together so that, those are instances where they were showing their, their data to the class and it was fine in that quarter because nothing unusual was happening but again in future that might potentially be problematic.

I: Yeah, yeah that's such an interesting, like it's such an interesting set of data to be working with, I think I was talking with someone previously and they were talking about the importance of data that you have a personal connection to. But I could see a lot of questions coming up from that as well, or issues.

UCSB4: Yeah. Yeah. And you know, on the one hand, even if there's nothing unusual in the data, you know students might be hesitant, or think twice about showing the results of the class. On the other hand, they are so deeply invested in the process because it's their own samples and it's a level of interest that I haven't seen from students in other contexts. So, I'm reluctant to give up that level of student involvement.

I: There's a sense of ownership there since it is their own data.

UCSB4: Yeah, yeah. And even when they don't understand something they want to understand it. Because theirs, you know.

I: Yeah, that's so interesting I could see the same thing with DNA, as well. I mean, I think it's just something that everyone's intrinsically interested in. But anyway, I can move on.

UCSB4: But I thought I'd say one last thing there, so because we're starting off with urine samples. You know, they're starting with a pool of students who are not grossed out by this stuff right. They are happy to talk for hours about how like, Oh, I thought it would be easy but then it turns out like peeing into a tiny little cup is not so simple, because I peed all over my hand so how am I going to deal with that for the future. Or like, You know my pee was super dark and then you know this other person said well my pee was kind of a weird color should I be concerned and, you know, there's just all kinds of aspects to it that I had not anticipated because again this is the first time I had taught that course but again really had them invested in understanding what was going on with their own bodies.

I: Yeah, I wouldn't have expected that but I suppose, maybe the shared experience too, but they're all doing this makes people more open. Yeah, about their experiences.

UCSB4: That's true.

I: All right. Alright. One of the final portions here is about training and support in the courses. So do you have anyone other than you providing instruction or support, TA’s, other instructors, librarians, or drop in sessions anywhere on campus?

UCSB4: So when I'm doing the bulk of it, because the data intensive courses are small, I don't have a TA. However, during that honors seminar where the students observe animals on campus and then collect data and analyze it. I do every quarter have a one class meeting where we go to the library. In the past, Gosh, I don't even know how many times I've taught the course but there’ve been one or three times maybe where they've met with a librarian. But, I guess it was the last time I taught the course we actually had sort of a, what do you call that one, the library agrees to do a training session like a workshop?

I: We usually call them one shots, instruction sessions.

UCSB4: Yeah. So, we did a more formal one the last time.

I: And it was out with a subject librarian or someone in the Collaboratory.?

UCSB4: Yeah, I think it was a subject like librarian but, and so the last time I taught that course was, Oh gosh. So this year's 2020-2021. It would have either been early 2019 or late 2018 because last year I was on sabbatical. And that's why my memories are a little bit fuzzy here but so we had this sort of formal workshop with the subject librarian, and it was useful but the students were a little bit tuned out. And I think, to me it was interesting because it was more like a traditional class, and they had, most of the time with the class we were reading outside even if we are analyzing data. And so I think they kind of resented having to be put into a classroom.

I: I see, yeah I know that's something that we struggle with is, you know, setting the right tone or making sure we're talking about things that actually, you know, relate directly to the things that students are doing in their courses and it's hard if you haven't been in the course other than that one day.

UCSB4: Now every time I teach the course we do at least go to the library for one of the class meetings, maybe two, and then my reason for that as each year I'm horrified, I ask the students, How many of you have ever gone to the library and of course everybody raises their hand because they've studied in the library, say okay so how many of you use the library and of course they'll raise their hands again because to them that's using the library for studying, right, I say, Okay, how many of you have ever checked a book out. And, on average, it's not more than one student per quarter and usually it's maybe every other quarter.

I: No more than one student per quarter? Like one student out of your cohort, your course? Wow.

UCSB4: And then I say okay so you've checked a book out. Did you check the book out from this library, and they say no, they check the book out from their public library back home.

I: interesting.

UCSB4: I say okay so for all of you guys regardless of whether you've checked a book out. How do you figure out what book to check out, and they're like well we don't look for books, because, you know, unless it's online we don't care about. So, the class meeting where we're in the library and I started off with a series of questions. I say okay well I'm going to force you guys to, you know, use, I can't even remember what it's called the library website where you look up materials online. Yeah, you have to find the call number and then you have to find the book on the shelf and once you find the book on the shelf, you have to stay in that area for like half an hour and look at the other books around that book, because maybe that book is not actually going to be useful to you but you don't know until you pull it off the shelf and start looking at it. And then at the end of class you're actually going to check out the book and take it home. And, yeah, so that's the first time they've ever done that and I feel like a grumpy old lady, but I don't care.

I: That's so interesting with, like, it feels like just how quickly things have moved online, too. I mean, even between my undergrad program and my graduate program, I used the library all the time. And then I used to almost, I mean, even as a library student I barely used it because there were so many materials online.

UCSB4: Yeah. Yeah. And part of the reason that I have them go through this whole process is not just that, you know I want them to have the experience of checking out a book and holding a book in their hands and looking for information in a physical book. But time and time again, they realize that the material that they have access to online is partial, they're not getting the full picture and the stuff that they're looking up in these books about the Natural History of the animal that they have chosen to study for the course. The species. And in the course of looking up the, you know, finding the book and determining whether it's useful and then you know looking at the other books around, they find out how much information they are missing because it's not online.

I: Yeah, that's really important. Especially just the free resources online and there's so many things behind paywalls then too and then yeah the books, the physical books in the library.

UCSB4: Yeah, well, and this is a field where people were observing animals for decades, right before stuff went online and a lot of that stuff is in books, it's not even in journal articles. And those books have not been digitized so that's, I want them to get a sense of what's known about the Natural History of the species before they even start collecting data. And I want them. I want the book, and what's known to inform how they collect the data,

I: I see. Yeah, that's really interesting. I'm really shocked I mean I'm, I'm surprised but I'm also not surprised by the one per quarter number.

UCSB4: Yeah, yeah, no I mean I, I've never had more, I mean that's a small class right but this is the class I've taught most often, actually.

I: So, Do you find that students are learning to work with data outside their formal coursework either through things like, you know, like I'm thinking like data or not data camp, what is like coding, like different coding resources or data resources online, or from other internships or things like that before they come into the course or do they typically only have course experience with data.

UCSB4: So, no, no. I've never met a student who had attended something that was not basically a formal class right for course credit, and of the students in the class, you know, over the many quarters, only one or two had ever taken something like a statistics course with data. And, you know, they had the same experience that I did where they were like yeah I know I was supposed to learn stuff and I did stuff but I don't. I didn't absorb it. And the thing about the honors seminar is that, you know, these are students not coming from just one major one department, they're all over the university. And part of the reason that I don't say in the course description that they will learn how to enter and graph and analyze and interpret data is that that's the sort of thing that a lot of students from social sciences and from the humanities will run away from.

I: Oh, yeah, I can attest to that. I didn't take a single statistics course until I was in grad school. I actively avoided it.

UCSB4: Yeah, exactly. And so it's easier for them to sort of get involved with it and then for me to say, guess what guys you're actually doing statistics and they go, No. And, you know, I have them doing this because I want them to lose the fear, you know, rather than thinking, Oh, this is something that I can't do. It's just part of course and not a big deal I guess yeah I mean it is a big deal but it's not.

I: And so do you encourage them to do any kind of other extracurricular learning outside of the course or after your course.

UCSB4: So, in the honor seminar, not really because you know that it's not actually a full credit course right it's only two units instead of four. So even there, I don't really give them much in the way of homework or readings. Because it's supposed to be less of a demand on their time. It's just enough that they make sure to show up every week because they're interested and you know they're there. To give you an example. The more convenient species to study on campus are the turtles in that little pond between the [student center] and and the tower, and almost none of the students come into the course thinking that turtles are interesting and by the end of the quarter, they've given every turtle a name, and they are highly invested in these as individuals. And, you know, they, they talk about it like a soap opera right like so and so did this did the other one and then you know this is how they responded and you know he's such a bully and blah blah blah blah blah right so so they do get invested in it. But, yeah, so if a student has an interest for instance in doing field research in ecology or animal behavior or thinking about grad school, I give them lots of information about that. But in terms of pursuing other opportunities nobody's really asked like how do i do more data analysis or how do I do more statistics.

I: Okay. And then, as far as your own training goes, have you received training in data, other than your graduate degree like have you done other workshops or training, since that formal training?

UCSB4: Yeah, so at UCSB just the last, I guess couple years so, so this year I did one of those data carpentry workshops for Python right, done a workshop for R, and I'm frustrated with myself because it didn't stick, but the reason for that is that I've been using STATA since mid grad school, and it makes sense to me and obviously Excel. It's not that I want to keep using Excel. But I have found over and over and over….Sorry.

I: Yeah. No, I was gonna say it takes a lot to learn a new language, I mean it's a different language, in the end.

UCSB4: Yeah, no I mean that's that's absolutely part of it I'm, I'm really short on time right so it's not like I have the luxury of doing a 12 hour dive into any one new software, or language. But maybe this is just a quirk of this, the grad students that I have experienced with and exposure to. But the students who adopt R or Python or something like that early on and do their utmost to avoid Excel, they make crazy mistakes. Right, so. Granted, I'm not super familiar with those software programs so basically they'll join data sets or manipulate them and then do analysis on them and come up with these amazing results in their analyses and when we backtrack and sort of check the flow of the data handling. I always discover some really, really serious fundamental problems. And my interpretation of this is that when they're using R, especially it's kind of like a black box and so they learn the code for doing a step and then they don't check it to make sure that the output makes sense.

I: Or they don't understand what's going on behind the code.

UCSB4: Exactly. And I think that when they learn how to do the analysis using Excel where they're actually looking at the numbers, it's a lot easier to catch those mistakes. So, I'm torn because I feel like a Neanderthal for not knowing how to use R properly at all, really. And it's what everybody uses everybody in the social sciences and at least the younger generations. And I want to be able to reduce, you know, the, the nice looking figures and to use the same language that everybody's using so that, you know, I can post my code with every manuscript and people can look at it and say yeah that's that's robust, I know that she did this right but at the same time, I am not letting go of the old methods because time and time again, I've seen how important it is to be physically keyed into the data at every step and to check it and think about what's happening. And then, I'm trying to teach that to my current grad students but it's a struggle.

I: Yeah, no that makes sense. Especially considering, you know, when you use R and Python that's usually at the step after doing the data collection. So after you've used Excel or whatever other collection method and it's hard I know with our workshops too. We try to give a grasp of everything, you know the collection portion and. And then, like how a spreadsheet should look, but it's hard to cover everything and of course the code part is like the flashy part right yeah. So, that's really interesting.

UCSB4: One of my goals during my sabbatical was to spend some time with R and that did not happen. That's my own fault.

I: That's okay and it's been, I can't imagine you have any spare time right now. I feel like I haven't had any and I can't imagine what it's like for people teaching courses right now. Let's see. Last thing I want to ask you, do you use any, like, data sets, or syllabi or other instructional resources that you received from other people or do you make them all yourself and do share those with others?

UCSB4: So, let's see. For my big upper division undergrad course I inherited the lectures from somebody else. And there are problems with them, and I keep meaning to make it my own and fix all those problems but that lack of time, lack of full interest. The smaller upper division undergrad courses and anthropology and the honors seminar, those are things that I have designed myself and I've shared the materials, both for the course that I inherited, and for at least one, maybe two of the other courses with people, and other universities like new faculty who have been hired and stuff like that.

I: Okay. And then I have one more question for you, and we'll wrap up here. Sorry if I'm keeping you over time. What is, considering the trends in your field, what types of training or assistance do you think would be most beneficial to instructors like yourself with teaching data?

UCSB4: So the thing that I'm frustrated with every year, and that it would be lovely if there were a better workaround for this. When students are doing the deep dive into data so they're collecting the data, entering, analyzing it, graphing etc. Most of the students don't have Excel on their personal computers and the only way that they could, that I know of, that they can have access to Excel is by going into one of the computer labs. And, you know, I don't want to lose the outdoor aspect of the class right because that's, it's the class that they look forward to because they know that they're going to be sitting in the sun and, you know, talking with members of their small group etc. So that means that a lot of the students are forced to use Google sheets which is super limited in its functions. And it's clunky. Yeah.

I: This might be an ignorant question since I am new here but don't, the students aren't able to get Microsoft Office with Excel for free, or is it just that they have computers that can't support it?

UCSB4: I don't think that they get it for free, I think they have to pay.

I: Oh, okay. I'll have to take a look at that. I've never had that experience at any other institution I've been at so

UCSB4: Yeah welcome to UCSB. So yeah, there's probably a workaround for that. But I think that's the one thing that sort of immediately comes to mind. I do keep the course kind of basic and lo fi, just because I don't want to get, I don't want to spend too much time on fancy tech and have the students be a step removed from the really hands on process of, for instance, entering their data and stuff.

I: Yeah, and obviously that brings in other issues and I mean there are always installation errors galore with carpentry workshops, yeah so that's a whole another set of issues to have to deal with, and troubleshoot. Okay. Well, is there anything else that you want to share what you thought of since the beginning here.

UCSB4: So, you know if there's ever a push to make that sort of sort of like a universal curriculum aspect of having students working with data from the beginning of their time in college, I'd be 1,000% behind that.

I: Yeah, I know once in a while we throw around the idea of an INT course for data management, but I think that would be even, even a larger push than that.

UCSB4: And I know I said this at the beginning of the interview but it's worth repeating, working with data changes your brain. And the students appreciate it once they've been sort of pushed into that experience or, you know, had a little bit of hand holding. And, you know, if you want to make sort of an argument that it's justified because it's useful to the students right like there's so many different careers where any little bit of handling data will improve their skill set improve their understanding of of the patterns and things like that that they're seeing with that data that they’re working with, big data is all the rage right, those are the jobs, in data science.

I: Thank you.

Interview 5

I: We just have a note regarding the Covid-19 disruption. Obvious teaching has changed since we've all gone remote. And it's been disrupted. So, if you want to answer for any of the questions. In reference to your pre coated your normal teaching or to how teaching has changed. Either is fine. You can talk to one or both.

UCSB5: Understood.

I: Alright. So can we start with some background just on one, how your teaching relates to your current and past research. And also, much of the courses that you teach the students work with data.

UCSB5: Yeah, so, my current and past research focuses on information science is specifically around the area of human information behavior. And basically how people look for information how they seek it would they do with it, how they how they use the technology around it. I have a prior background to that in engineering, I was a an electrical engineer computer engineer for many years before I changed careers and became an academic and, and got my PhD and information science. So, with that background that I have and now I'm really more interested in, issues of, you know, answering questions of social science, using computational methods that are pretty common in the computing sciences and engineering, which is where my background initially was. So that really informs a lot of what drives me to teach the classes that I teach here UCSB does not have an Information Science Department per se. So, I am shared between the [redacted] department and the Communication Department, which I love, I love having a foot in both those worlds. I have noticed that it's not a common thing on this campus. I am a lecturer so I'm not focused on research, currently, as much as I am on teaching.

So, the classes that I teach in the [redacted] department are [redacted], but in the Communication Department I teach a mix of both classic undergrad courses like research methods and also I teach courses that really are very much about data driven and computationally data driven research so I teach social networks to undergrads and I teach. Also, a special topics class that I created for the department for [data science]. The both of those classes, especially the latter though, utilize or introduce the students to the notions of big data. And in social networks. I don't take it too far with the data, because it is about introducing them to theory and it actually uses a lot of sociology stuff as well as you know application to come theory but the [data science course], it's all about data, it's all about the use of data so about a third of that class is on a mixture of like an intro of what it all means the technology required, and the ethical and societal impacts that big data has. I'd say it's more than a third, I'd say it's almost half of the class is on that. I also do a module on data visualization. And I do a small module on coding and programming or are the very basic fundamentals of programming, given that the audience in that class is not are not computer science majors. But, you know, I want to introduce them to the notions of what it takes to program something with data. And some other some other minor things that I talk about but that's, that's a good summary of that.

I: I think we’ll need to focus on the communication courses since it’s social sciences, the focus for this study. So, is the [data science course], do you use data sets in that course for doing the data, I imagine for the data visualization and the coding?

UCSB5: I do, and I introduced… I obtained a couple of data sets. We talked about also how one might obtain data sets as well so I talked them about some of the technicalities around. Finding data online using web scraper technology or using API technology to download data sets. We don't take it too far because again the audience is not really primed for to go into in too much technical depth. But I then provide them with like a small, or an example of what a big data set might look like and walk them through have them do exercises with analysis, and some basic statistical analysis of it. Some visualizations.This has turned out to be more of a survey course like, here's a bit of this here's a bit of that without delving too deep. I would love to come up with another class that say, is all about data visualization and actually have them do a database of their projects with that, but that's for the future.

I: Yeah, yeah. the next question is about how the students obtain the data that they're working with, whether they collect are generated themselves, whether they're searching are identifying, pre-existing data, or whether you provide the data set and it sounds like you give them an overview of all of those things but then you end up giving them that that data set to work actually work with and in the class.

UCSB5: That is correct. And I would have liked to experiment with having them, collect the data themselves or like a [unintelligible] exercise. However, when I came up with this class, I was slated to teach it in the spring last year. And as you know, we went, because of COVID, we went online with it. And so, I lost the ability to. I was planning on the TA and myself doing, you know, a fair amount of hand holding with these non-technical students to kind of show them on a computer. Look, do this, do that and this is how you get it right. And that's really hard to do online. So, I had to change my plans and curtail that, which is why I decided to tell them about, you know, hey, theoretically, this is how you go about and get it but look, I have one for you.

Here you go. Open up excel.

I: Yeah, we've been experiencing that with the carpentry workshops too. Yeah, I imagine for finding data sets, you know, there's all sorts of other scenarios that students might find themselves in and can't dig themselves out of

UCSB5: Yeah yeah exactly and the variability I found with these students who are social science students, mostly communication majors, the variability is very big, in terms of how easily they use technology. Some of they'll jump right in and they'll try this and that and others are just like what button do I press, ya know, they're a little more afraid of it.

I: Yeah, yeah, afraid of doing something irreversible.

UCSB5: Yeah, yeah that sense of that sense of experimentation isn't always there with undergrad students.

I: Yeah. Do you find it difficult to find those data sets that you use for that course, or did you have to do any kind of grooming of them to make them appropriate for the class?

UCSB5: Yeah, so, no it's not difficult to find data sets online so many sources. And, and, yes, eventually I did have to groom the data set, as you said, mostly just by taking out categories of the data like taking out fields, though, for example, though, the one assignment that I gave them was to analyze this data set of of movies, there's this data set on movies on like their financial performance and, you know, how much profit they made and all of that. So it's, you know, there's a lot of categorical data and numerical data it's a nice mix of different types of data. The big data set that I acquired was, I mean most datasets that you can find out there are either CSV or, or JSON and CSV is so easy to open up with something like Excel it's pretty common place now. Yeah, and like I said I just had to like clean up some things like remove some uses a very big data set I removed some categories that I didn't think they needed to work with. but that was minimal.

I: Okay. And I should ask how large is this course, the [data science course]?

UCSB5: Yeah, both this one and the social networks course in COMM, are both upper division, and we get at minimum, 80 students, maybe up to 100, 110. So, yeah, average about 100 students.

I: So, you have quite a few students that you're working with and do they then have like the actual lecture and then a lab, as well, with TAs or is it just lecture based?

UCSB5: it's just lecture based. There is no lab with this, it would have been fantastic. but no.

I: All right, so I think you've touched on this already a little bit but the next question is, you know, how do those students in that course, I'm thinking of the [data science course] specifically, how do they manipulate analyze and then interpret data in that course.

UCSB5: Um, so the exercise that I give them is...let me think, because I didn't know how many of them had access to Microsoft Excel. It's nice to assume that everyone can get that because actually I think they can get it for free on campus, if I'm not mistaken, but I ended up putting the data set on Google Sheets, sharing that with them. And it was a large data set, but it fit in Google Sheets so you know it's not like really big data. And what I had them do was do things like you know for categorical data, I taught them how to do frequency counts like you know how many times does this category appear versus that category and then like create a graph to represent it very basic data analysis type of tasks. Other numerical tasks, or you know, I'm sorry, categories that I mean, fields data types that were not categorical but were interval types or racial types, you know, that they could actually perform stats on. I taught them how to do, or I had them do you know find the mean, the median, the standard deviation, do a regression analysis between two variables, do a correlation between every pair of variables, things that they had learned in their...because they all have to take a beginners class in stats. You know how much of that sticks in their head either? Not that much. At least they all remember what an average was that was something.

I: I can say from my intro stats I don't know how much I retained. So, it's not surprising.

UCSB5: But you know the communication department is actually very quantitative oriented. And a lot of the professor's there, you know their research is very quantitative, and they really push their students to know this stuff. But despite that predilection, a lot of the students needed a lot of refreshing. And that's fine, I spend half the lecture just reminding them what this stuff is. But then I give them the task to analyze the data that way and create graphs from it. In my short experience right, I've only taught this class once in the spring, and once in the summer and then in summer it was a lot shorter time and just wasn't quite the same thing. So I don't have a big end to work with here but they did fairly well with it.

I: Okay. And so, were you able to do all of that in Google Sheets with them and, when you're demonstrating, do you show both in Excel and in Google sheets, are they close enough that people can figure it out?

UCSB5: No. A: it's close enough. B: what I was trying to avoid was giving them too much information and getting them confused because I know from past experience when I show them different tools, when my goal isn't to really teach them differences in the tools. My goal is to teach them the technique. And I showed them different tools. I get flooded with questions of “why does it work here and it doesn't work there?”, and you know it's because they're different, you can't do the same thing exactly. So no, I stuck with one tool, thank you. With Google Sheets and the fact that Google Sheets is free and accessible and shareable and they could share with me what they were doing so I could look into what they were doing. It has a lot of advantages over Excel in that sense.

I: And did you expect the students to have prior knowledge of Google Sheets and whatever other tools they ended up using in that course? Or did you teach it all from scratch?

UCSB5: No. Well, I did not question their background on using Google Sheets, I believe, there's no study for this but anecdotally, and from personal experience, most of these undergrads by the time they're seniors, they've used the Google tools in some class or another whether it's Docs or Sheets. And so I wasn't too worried about dealing with the technology. After all, I was giving them an already filled out sheet with the data. All they had to do was learn how to select columns of numbers and apply them in formulas which I had demonstrated in lecture earlier. I don't think it was a stretch for them to deal with the technology. At least my intention was to not make it about the tech, but to make it about the technique and the methods and the ideas and the concepts.

I: Right, I think that relates to one of the follow up questions here which was to what extent are those tools and software important for your teaching? I think you iterated that it's not about the tools, it's about learning the actual techniques.

UCSB5: For this class, but for example I've taught workshops for grad students on data acquisition from social media. And I've shown them how to do it, using an API to get tweets. And I showed them how to do that using R and I showed them how to do this using Python. Now, these were grad students in the social sciences, but they were pretty sophisticated when it came to knowledge of technology, even if they weren't programmers, they could deal with what I was telling them. This is a different audience, these are undergrads, and my purpose was different. The purpose of the grad student workshop was to introduce them to different tools and techniques. With the undergrads It was really more about the methods than the techniques.

I: Makes sense. Did you face any other challenges with their ability to work with data, rather than the tools?

UCSB5: Yeah, I would really like to do more with data visualization beyond just simple generation of bar charts and pie charts and scatter plots, which is really what a tool like Excel or Google Sheets that's really all they can do. And so, for the future classes I want to look into incorporating other tools for visualization like tableau, or something like it. Because I think data visualization can be a super interesting medium to teach them about big data actually because it speaks to communicating concepts, as well as dealing with tech, and it has this element of artistry to and having an eye for design and having an understanding how that works with human cognition, I mean it's a it's a lovely field that that just incorporate so much. But in order to do that effectively, I really do need a better tool, a more sophisticated tool than Google Sheets. So, I'll be looking into something like Tableau or something else like that to help me. I'd love to hear what other people might have in terms of suggestions I don't know if you do.

I: I was working with, in my grad degree, working as a TA for a design lab and a lot of what we did was teaching students how to do data visualization. And even though it was a year ago, it feels like forever ago, but maybe we can talk about that another time.

UCSB5: Yeah, yeah, and I'd love to compare notes with you and if you have that class syllabus somewhere. I'd love to take a look at it.

I: Well, that wasn't a class it was more like a consultation lab for students to come in and get help on their homework and yeah we did go and do instruction sessions and classes, too.

And I'm not sure if I still have some of those slides but I'm sure they'd be happy to share them as well. Yeah, they've done some research on, especially the human cognition side but the design side too.

UCSB5: So, yeah, I found this really great textbook that I photocopy the chapter and have them read it it's by Andy Kirk on data visualization, and it really just encapsulates all these different interdisciplinary areas that play into dataviz and I know there are departments on campus that do more with that. And I've just emailed back and forth with a couple of professors, but I'd love to find out more what's happening there too. I forget the department’s name. Their name, if it’s MAT, it's a grad program in design and art.

I: And that's here? I guess I don't have a full familiarity with all the departments here yet.

UCSB5: Yeah I've only been here four years myself. And there's a surprising amount of stuff. Yeah, it's Media Arts and Technology, MAT. So it's mat.ucsb.edu that's their website check it out.

I: All right, I'll have to take a look at that. I've heard that here and there but I really haven't looked into that department yet. Kind of changing directions I guess I know you said that, like, communications is really quantitatively focused here.

UCSB5: The research.

I: The research okay. the next question is how do the ways in which you teach with data relate to those learning goals in the communication field.

UCSB5: So the learning goal, the stated learning goal for, say[data science course] is to understand the different issues surrounding the use of big data in society today and more specifically income research, which is why we talk about things like ethics, like societal impact of data use and big data use, and gosh there's an endless amount of things we can talk about there in class. And among other things, the goal there is to have them have an idea of and an appreciation of an understanding of the technology that's involved and what it requires, and which is why a lot of what I talked about is big data and they've heard the term big data but what does it really mean? Like did you know that Twitter produces five petabytes of data a year, well what does that mean? How many high definition movies is that right? I try to put that into an everyday perspective for them so they have a sense of the scale and really the enormity of this thing. And then realize that that data is all about you. That's what that is, all of that mountain, those mountains of data, anyway. So the stated goal here, the goal of the class is to give them an awareness and an understanding of the high level topics around there. Be it ethnic societal impact, or technology and/or the computational skills required to work with this and navigate this and make sense of it.

I: Okay. Yeah, that makes sense. And this is moving on to a different topic here. But do you find that instructors in communication face ethical challenges in teaching with data? I guess more specifically with teaching with data, rather than using data in research but I think either is relevant.

UCSB5: Um, do I think the face, ethical?

I: Yeah, ethical challenges in teaching with data, like the data set that you're using or I guess even questions that come up with big data in that course.

UCSB5: Yeah. So, in this course, in [data science], we certainly spend a lot of time discussing the ethics of data collection, data use. And, you know, more to the point data misuse. So that's really very explicitly talked about in this course. In other courses, I don't know that there are many if any Comm courses that utilize big data sets with the undergrads. So I'm not aware, actually, of any Comm class that utilizes big data sets. But I do know that issues of the ethics of data use and ethics in research are very strongly underlined in the comm department. And in fact, in that research methods class I that I teach, it's a pre major course. Comm has a pre major, which they have to take these four courses and have a B average before they’re admitted into the major. Anywaythat is one such course that they have to do well in. We certainly talk about ethics in research and ethics and data use. So I can attest to the fact that this is an area that is talked about to the students in more than one class. But actually having students go and collect data, like big data sets, I don't know that any class does that.

I: Okay. Yeah. And have you seen any, you know policies or cultural changes here at UCSB with that or you know even wider communication field wise, that influence those ways in what you teach with data? Like there hasn't been any change over the years in that?

UCSB5: Well, just in the last year or so, we've started the Data Science Initiative on campus. [redacted]. And there's a push to turn that into a major eventually. So there's certainly been a cultural shift.When I joined the faculty four years ago, right after my PhD, I had a hard time frankly talking to people about what data science is and my PHD in information science, a lot of people were like “what is that?”, they had no idea. Which was surprising coming from other places, you know, other universities have had information schools for, you know, for a while.

Berkeley, Madison, Rutgers. Gosh, I mean I could, I could rattle off a dozen names. Yeah, so UCSB has come a long way just in the last two years at least with the awareness of the importance of data science in the world and not just in like stats and not just in computer science, but in the social sciences and the humanities. And so, you know, the faculty involved with this from all these different parts, they're very savvy about all the nuances of the use of data science in their own fields. So maybe that knowledge was always there but I think that the Data Science Initiative has kind of brought people together and brought it to the forefront. And computational social science has become more of a buzzword on campus now, at least in the past two years. This is just my feeling again. People are talking about that more in the social sciences. I've heard about more and more hirings in that area than before. But again, this is just my personal impression.

I: Yeah, and I think you know the majority of students we have coming to the R and Python carpentry workshops are from the social sciences here too.

UCSB5: Yeah, there's a real, you know, what's the word I'm looking for, demand for those courses from the social sciences, especially the grad students.

I: I mean, these online workshops we've been doing, especially registration will fill up within an hour of when we email them out. So I know there's a lot of demand there. You know we always have students asking for more and of course there's only so many of us to teach right.

UCSB5: Well, if you guys ever need more instructors please reach out. I’d be happy to help out.

I: Okay. Sounds good.

UCSB5: Yeah, I mean one of the courses I teach [redacted], which is open to everyone on campus so anyone can take that and it's highly impacted. And that just goes to speak to what you just said, you know, everyone wants to learn, at least take an introduction to programming, even if they don't quite know what it means. And mind you, many of them once they learn they go “oh, that's not what I thought it was, I don't want to do that. ”.

I: Oh they don’t want to do that? I had the opposite happen in grad school.

UCSB5: They’re like “What, it involves math? Logic? No, I don’t want to do that”.

I: That's funny. All right, I guess moving back to the courses that you teach. Do other people help or provide instruction with those like TAs or co-instructors, you know librarians or do you send them to any kind of drop in sessions on campus?

UCSB5: I usually have a TA in any class that’s over 60 students. If they have the budget for it, they will give us TAs. So, I did have a TA help me out in this class and you know their duties are really mostly grading and holding office hours. Again, if I was able to teach this class, you know, non-Covid, you know in person, I would love to have the TAs use their office hours as kind of like walk in labs, really they would actually demonstrate a lot of this stuff if they could. These grad students are our Comm grad students who have an interest, a research interest in computational social science so they usually know their way around a lot of the software and have been super helpful.

I: Yeah, nice. Sixty students until you get a TA that sounds like a huge workload.

UCSB5: I know, that's another topic of discussion, that's a totally different time. UCSB has huge classes, and they're only getting bigger.

I: And I guess the next question here is, do you know of any other ways that students are learning to work with data or, you know, learning to program or code outside of their formal coursework in your class? Does it seem like people come in and, you know, either before or during utilize other sources to get a grasp on things.

UCSB5: Yeah. Again, I don't have a big sample of students to, you know, talk about here, given this is a new course. But I have, and I'm sorry I maybe I misunderstood your question, but I did not see students go outside of class to learn more about especially the technical aspects.

I: Yeah, I'm thinking like Datacamp or coding camp, you know, those online resources.

UCSB5: Right. I've had a couple of students express interest but literally just two students interested in taking courses like that after. I've recommended, you know, paths that they could take, you know what courses to take. I don't believe they acted on those. Both of these students are graduating seniors and I am pretty sure they're busy with their own budding careers right now.

I: Do you often encourage people to do that kind of, you know, outside work outside of the courses? Especially if they're interested in coding or programming I guess that's where I'd see that naturally happen.

UCSB5: Absolutely, and then, I've had a couple of students take like directed research courses with me. Where you know we can set our own agenda. It's kind of like a, what's the word, yeah, directed research. Where we sit down and we talk about what it is that you want to do. Should we like, do a research project, do you want to learn more about this or that. And a couple of them wanted to learn more about programming. And I tailor the, you know, one on one intro to Python. And they did very well and they were very motivated and they understood it. So yeah, but that's a minority that's literally one or two students out of hundreds.

I: Yeah, okay. Yeah, I imagine they're really just too busy to think about other things like that and add them to their plates.

UCSB5: I mean it's a combination of they're too busy and they don't really know where to go. Yeah, but you know whenever they ask, I am happy to help. In fact I actually put it in my syllabus that if they ever want to learn the skills beyond what this class teaches, I'd be happy to refer them. And, you know, there are a lot of online resources for that. Some of them are free or very inexpensive. A lot of the Khan Academy courses are, you know, good intro courses and there's a bunch of stuff on YouTube, and I so I do point them out to these things.

I: And then as far as your own training, I know it sounds like you haven't been here all that long and you had just finished your PhD before that. But have you done other training yourself since you did that PhD?

UCSB5: Training on what?

I: Training on you know, working with data or more training on programming or coding. I think I imagine not since you just recently completed your PhD, I think this question is more oriented towards people that are further along in their careers.

UCSB5: Yeah, and also because I, the nature of the classes that I teach, have kept my coding skills, pretty fresh. So, if this is working with data and coding with data is something that I do on a very very regular basis, both for my own research, which is whenever I have the time I engage with, and in the courses that I teach so that actually works out well for me.

I: Yeah, so your, your learning is built into the teaching essentially.

UCSB5: Yes, yes. Having said that, I always keep an eye open to new things that are coming out. I keep up with the literature. I always query my, especially my students who go out and find, you know who go out and work in industry. What companies like Google and Facebook and Microsoft are doing with data, what tools they're using. I try to stay, you know I try to self-educate myself on what's going on there, on the trends and all of that.

I: Do you, it sounds like you do talk with other professors at UCSB about the courses that they're working on. But do you do a lot of sharing of syllabi and resources for instruction too?

UCSB5: Yes. I found the community here super helpful and willing to share ideas and syllabi and even slides and materials. So, yeah. And I do the same when someone asks, I share my stuff willingly yeah you know that it's part of the culture here.

I: Yeah, yeah. I'm seeing that, the only problem is how busy people are, right. It limits the time.

UCSB5: No everyone's crazy busy here and especially people who are immersed in research. And you know we're a premier research institute, I mean these people are busy and scarily productive, you know like, generating dozens of papers every year right that's unheard of at other universities.

I: Yeah, I can't imagine. Then the final question here is about, you know, considering those evolving trends. What types of training or assistance do you think would be beneficial to other instructors, when they're teaching with data?

UCSB5: I think teaching with data, especially teaching data analysis or data techniques, the training really is about doing the work yourself. Coding is a skill, like playing the piano, you can read books about it but you're not really going to learn how to play the piano unless you sit down and play it. Same with coding and working with big data and working with certain technologies, it’s just stuff you have to do, whether it's in your own research, or if you just take a course or a couple of workshops. So, that's what I would recommend an instructor to have as background before they teach a course like this.

I: Yeah, I think the same could be said for stats. I feel like that's why it doesn't stick because you're often just being lectured about it but you don't actually try using the methods on things and that's really where you end up learning it.

UCSB5: So true, a lot of applied stem is really about building skill, whether it's programming or using applied math or applied physics. There are theoretical aspects to all of those but the applied stuff is very much skill based. Just putting in the hours.

I: Yeah, that's interesting. Well, that's all I have for questions, is there anything that came up from the questions that you think you want to add in or anything from other experiences or perspectives that you might want to share?

UCSB5: No, I mean, I'm very glad I'm very excited that this activity is on the rise at UCSB. And I think the demand for classes like the workshops that you guys do in the library and courses like this data science course in the social sciences, that demand is only going to grow. And so, you know, this study that you and your colleagues are conducting is super timely. I'd love to know more about what your results are, what you find.

I: Yeah, we'll definitely send you the report when it's done both the local one. And the final one. It will take some time. I think the timeline is to have the local report done sometime I think late next spring.

Interview 6

I: I wanted to start by acknowledging that teaching and learning has been significantly disrupted in the past months due to our pandemic. For any of the questions I’m about to ask, please feel free to answer with reference to your normal teaching practices rather than the teaching practices you’ve had to adapt to during our crisis situation.

UCSB6: Ok

I: So, I first wanted to start by hearing about your experience teaching undergraduate, so can you briefly describe your experience and then how does your teaching relate to your current or past research? And what courses do you currently teach that work with data? And I can repeat any of that if you need.

UCSB6: Ok so you want to hear about all of my teaching, not just the class that has to do with data?

I: So we can start by talking if you wanted to start with that class, that's totally fine if that's the most relevant application that you have for this interview.

UCSB6: Okay well I guess just more generally speaking, I have been teaching at UCSB since [year given] and I taught a couple of other places before that as I was finishing my PhD and just afterwards. I am in the Department of Communication and I teach regularly these days. I teach a lower-division [course name], so that's directly about data and data analysis. At the end of graduate level, the other classes that I teach have to do with [course topic]. And at the graduate level, I teach again [course name] and also [course name], so those are the classes that I teach that are in my regular rotation here at UCSB. What else did you want to know?

I: So also just for the point of clarification, we’re specifically in this case referencing those courses you teach for undergraduates.

UCSB6: Ok

I: So just for the sake of clarity. And then aside from those classes you’ve just mentioned, how does this relate to your current or past research that you do as well?

UCSB6: Well oddly, you know, a lot of my teaching doesn’t line up directly with my research and that's kind of an artifact of a couple of things that's a little bit unusual in our department typically. It’s more typical that faculty will teach, say a speciality course at least here or there that directly has to do with their research, but in my case, that's not really true because I do this [course name], which is just generally about, you know, research in general. And then the [course class and topic] that I teach undergrads, that's a [course level] class and that's more of just a passion that I have. I’ve done a tiny bit of research having to do with [course topic], but it’s not my area of focus. So, if I wanted to create some new classes, I could definitely link, you know, my teaching more directly with my research. I haven’t been, you know, eager to do a lot of new course preps, so I haven’t done that. I also do interact with undergrads about my research because we do a lot of undergraduate...we have a lot of research assistantships for undergraduates in our department, it’s a very common thing for our undergrads to do, so that’s working in very small teams on a particular project, and those projects do have to do directly with my research. So, I guess in that sort of class, it’s called the [course name] - teaching and interns happen there that’s directly related to my research.

I: Ok. Could you expand a little bit on what your research is?

UCSB6: Yeah sure. My primary area of research has to do with [research topic]. I’m looking at the social impacts and implications that they have. More particularly within that, I have established a thoroughly long research agenda looking at the [research topic], so the [research topic] particularly as it’s complicated in digital environments. And people used to hear me say that and go like “oh, you know, that's nice” you know with this kind of attitude like “hmm” you know, people you know they can't figure out you know what to believe in, they’re just stupid right? But ever since [year], now when I say what I study, people are like “wow that’s really important” so

I: Yes, I would imagine so.

UCSB6: I also study issues of [research topic].

I: Great, thank you. So we’re going to move into this whole getting data section. So, a little preamble. So, in your courses, how do your students usually obtain the data they engage with? Are they either generating it themselves, doing a search to identify pre-existing data, or do you provide them with data sets?

UCSB6: Ok. So now I’m going to switch to talking about my [course name] class which is, this is a statistics - it’s [course name] for the social sciences. It's in the department of communication, it serves as a [course level]. And that’s the only class in which I have students get, you know, like numbery data. In other classes, I might have them do literature searches, but I don’t think that’s what we’re talkin about here.

So, in the [course name] class that I teach, I do provide students with data. We have a data set that we use. It’s introduced in the first week of class. It's actually kind of loosely modeled off of the student evaluation survey just because that’s something that most students are fairly familiar with, but it's kind of, it was created and tailored for this specific class because it has some things about….it’s a hypothetical study in which students have taken a class either live or actually ironically remote through some kind of online learning type of thing, and then they either have a male or female instructure and they either have a certain, there’s these different type of structures, and so we have it’s a pre-generated data set that was created for this class over 20 years ago, I didn't create it myself, but it’s still fairly easy for students to grasp because they know the concept of a survey and they know the concept of course evaluations and so, in chapter 2 of our text book - this is a custom textbook that was made for again this class over 20 years ago, the data set and the hypothetical study are introduced, and then the students work with that data set every week so you know when we might be talking about t-test, another week we might be doing bivariate regression, and we use the same set of variables in that data set to demonstrate and illustrate those various statistical techniques that we are talking about.

So, they have, the sections for coming to section are two hours long. The first hour, the students meet with their TA in a classroom and it's been, you know, answering questions, going over their homework, those kinds of things. And then the second hour, they actually moved to a computer lab on campus - obviously before COVID. And the TA leads them through a set of computer exercises that are data analysis exercises that comes straight out of this custom textbook for our class. And so each chapter focuses on a different thing, we start out with just descriptive statistics and then we show them how to, you know, calculate frequencies and percentages and means and standard deviations and draw histograms and look for outliers and stuff like that. And then as the class progresses, we move into hypothesis testing and we do have them run T-tests and ANOVAs and do a little bit of regression. We keep it pretty simple - don't get into you know you can - I don't go into covariates much like it's just like the basic “this is an intro class” but that's how the data analysis part of it works. It’s done in the computer lab and then the students are assessed, not based on whether they themselves, you know, can figure out the right buttons to push in a data analysis software package, but rather we give them outputs that are similar to what they’ve seen in the lab and ask them like, you know we give them a scenario that a researcher wanted to know if “blah blah blah.” Below you’ll find - we use SPSS - so below you’ll find the SPSS output for this study and then basically the question is they need to interpret the meaning of the output, they need to understand what the numbers tell us about what the original research question or questions were.

I: I see. So, I think I’m going to pull back out just a little bit to ask about what’s the process that students actually go through to get access to the data set you have for these evaluations. Is it something that you just supply in the learning management system or is there a different means they get to it?

UCSB6: It used to be in the old days, we had the computer labs that we were using preloaded and so there was just a folder on the lab computers that the students would open up each week. Because of GauchoSpace, I just make the data set available on our GauchoSpace course page now and so every week when they go into the lab they get into the routine they just open it up.

And again, we use SPSS. I'm thinking about maybe changing that to R but I have to learn R to do that, and I haven’t invested the time to do that yet. So far we’ve used SPSS in the class, and I kind of sell it to the students as you know, a lot of jobs out there might find it useful that you have some familiarity with a data analysis package and so you know we're going to show you how to use this, we’re going to show you the basics of it, and you should put it on your resume that you have rudimentary knowledge of SPSS which is one of the most commonly used software packages in the industry. So I think that's fairly effective, like they’re like “oh wow that’s cool”.

I: Yeah, thank you. That helped a lot. So it sounds like you've only had this instance in this particular class where you supply the data set, there’s no other methodologies where they have to go and find their own at any point correct?

UCSB6: No, not in this class. In my department, we do have [class level] thesis in which they do an independent research, well they're mentored by faculty, but they come up with their own research question, they collect their own data….So it is possible for undergraduates in our department to do that. It's a year-long program.

I: Ok, so in that case, I mean is there any kind of standardization of, like, how you direct students to collect? I mean obviously, there’s methodology within the research itself or your discipline, but I’m thinking is there any structured guidance that’s given to them about how to go about doing that?

UCSB6: You know, in this case, and because we’re different from other fields that might have kind of common data sets that people, you know, might be able to access. I mean, certainly that’s a possibility for our students, but more typically, they're asking some research question where they need to generate their own data. Again, this is in the [course level] program. And so, they work for a quarter, it’s one quarter long where they develop - they’re in a class that’s led by a faculty member where they're developing hypotheses and ideas. You know, each person for their own project and also starting to think about the methodology, so sometimes the research question is going to lead to a more qualitative study maybe, it’s just interviews, something like that. Other times it’s going to be to a methodology with quantitative data. Typically for us, it’s more like surveys and experiments. I suppose someone can do big data, but our students typically don’t come in with the skill sets to do that to, you know, to crawl in and pull down data. Although some could probably do it, but so it’s wide open. So the possibilities are kind of endless there and so the guidance that they get is one-on-one. This is a small program, so the guidance they get is pretty much one-on-one for developing their ideas both in terms of honing in on their research questions as well as the methodologies to address those questions. Then in the second quarter, they're supposed to be collecting data, and if they're able to collect it in the second quarter, start to analyze it. And if they can't collect their data in the winter, if they don’t finish until the end of winter, then the data analysis will start in the spring. And then at that point, there is another faculty mentor who is available to help them do the data analysis. And I think that the reality of that and you could talk to that person like I can give you his contact information. He gets course credit for it. I think that the reality is he pretty much does the data analysis for the students because they just don’t - they’ve taken the introductory class but the introductory class is often too rudimentary to really serve the needs of these more complex research questions that they are asking, so he really kind of leads them by the hand, sits down with them, literally sits down with them in front of the computer. He has, I think, he tells them how to clean their data sets, but once they're clean, they meet together and do the analysis together. And he points out things as they go along, but there's no reason that we can expect the students would know how to run a….I was just looking at a, you know, moderated meditation analysis using you know process macro. There’s no way an undergraduate would know that.

I: Hmm okay. So, then this might be touching a little bit on what you’ve already mentioned before in the intro course, so feel free to repeat a little bit if you need to. So, we’re looking here for the how they work with that data, so how does students manipulate, analyze or interpret data in your courses? So, you already mentioned a tool, SPSS. Are there other tools that students use as well?

UCSB6: Not in my class.

I: Okay. Anything for visualisations?

UCSB6: We do in the early part of that class, SPSS, we have them produce histograms, line graphs, you know, you can do box plots, you know, it has within SPSS, it has its own kind of data visualization. Nothing fancy. And the class, like I said, we keep the class very very simple and this is not an advanced class. This is a class for students who have an absolutely terrible fear of math and a fear of quantitative data and so we keep it very simple.

I: So then, on the heels of that, I think I might know your answer. But what prior knowledge of these tools or software do you expect students to have? And this can even be like most basic of all understanding of math or…

UCSB6: Zero. Well, not zero. I tell them the first day that if, you know, most of them are worried about the class because, you know these are not engineering majors, these are not even economics majors, these are social science majors who like never dreamed they would have to take a statistics class. So I tell them, you know, “I know you’re nervous. However - and if I'm in person, I do this by raising a hand. I say, “Okay, by show of hands how many people in here can add?” And then I have them…..”How many people here can subtract?”, “How many people here can do multiplication?”, “How many people in here can do division? Not even long division and you'll get to use a calculator.” And so, everybody, you know, was raising their hands and so I’m like “Okay, if you raise your hand, you for all, for those four things and remember you can use a calculator, you are well prepared to succeed in this class.

I: Okay, great. And I guess on the other side of that, this question is pretty broad. This is “what do you teach them explicitly?” And I think this is in reference to the tools that you've already kind of gone through a little bit about you know like what kinds of features you’re trying to get them to engage with with the tool?

UCSB6: Okay yeah, that’s just, you know the tool just follows whatever, our use of tool just follows whatever concept we’re studying. So for example, the first week is we introduce them to this hypothetical data set and we have them do some data entry so we're just we're just kind of opening up SPSS, we’re explaining the environment like over here is where you can input data and over here is where you can look at a list of your variables and over here is you know, so we we just kind of orient them to the interface of SPSS the tool. And then we have them do a little bit of data entry just so they can get the hang of that. We have them label variables, we give them like the variables already have names that then we have them you'd have like variable labels that it’s you know a longer more descriptive name for the variable so that they don't have to remember that like a variable that’s abbreviated actually like “exp” actually refers to the expertness of the instructor, something like that. Then we have them do some labeling of values. So for example, if there are categorical variables you know gender we teach them that okay in our data set one it means female and two means male. We have them enter that in so they get familiar with data entry and setting up basically the codebook to understand the data. That’s week one. Week two, they start with doing just some basic descriptive statistics. You know, we have them use the tool to do the you know the mean, the min, the max, the mean, the standard deviation, you know really really basic stuff and just picking out a few variables here and there that “the instructor wanted to know the average... and how students perceived the instructor on average...or you know, the average rating of the course overall,” and so they’ll look at the average and we’ll talk about the standard deviation overall. The third week is where we do some of those graphical representations - histograms, line graphs, pie charts, again through the tool. And then after that, we start getting into hypothesis testing, but - oh Z-scores, having them calculate z-scores, show them how to do that - but a lot of my class is actually less about the computation and it's more about the understanding. Like they do, like for a T-Test, I tell them….well for variance for example, I say you know there’s a formula for calculating variance and most statistics classes are going to make you learn that formula and memorize that formula. And that's great if you want to look the formula and memorize the formula, do it. I am much much more interested in you being able to tell me what is the meaning of variance? What information does this give you? Why would you want to calculate a variance? You know, what kind of question can it help you to answer? So I'm much more conceptual than I think most stats classes are. So anyways, so then back to your question, though, about the tool, we just use whatever topic we’re teaching, we use the tool to illustrate - you know, okay now we want to know if if there's a difference between students’ rating of the course, if they had the course online versus live? And so we have them do a t test and then again the focus is not how do you produce that t-test, the focus becomes okay here are the results, here’s a bunch of numbers on the page you know how do you make sense of these results from the SPSS? So we take them through the output, or what we call the output, and we show them how to look at the numbers you know what to look at, what they mean, and how to then translate what that statistical output means in terms of okay how can we answer the question of is there a difference between how students rated the course if they took it live versus online and then they have to kind of like translate the statistics into words. And so, we just continue doing that with each statistical technique that we cover in the class. We go up to, we do, because the conceptual nature of the class, we go up past what most [course name] classes do, I actually have them do some multiple regression, again just some basic stuff, but looking at multiple predictors of one variable and asking questions about you know do all of the hypotheses predictors help to predict the dependent variable? Do only some of them do? Which ones predict better or worse? And have them answer questions like that.

I: I see. So, from what you’ve explained so far, it sounds like this particular tool SPSS is crucial for the teaching of this course. Is that accurate or...?

UCSB6: It is, although there are many other programs besides SPSS.

I: Right, you mentioned R...

UCSB6: Yeah R, SASS, I think there are many many others out there. I mean you could probably do most of what we do in Excel. I don’t know how to do it personally, but I can probably figure it out.

I: Mm hmm okay. So then, other than you already mentioned like there sounds like students kind of come in with a lot of reluctancy or like fears about their capabilities. Are there any other challenges that you face when you’re working with students to teach them about working with data?

UCSB6: I think the fear is the biggest one and just getting students to actually do their work. That's not specific to a data class, it’s any class, but the students that I find who struggle the most are the ones where it turns out that they really haven't put much effort in, they haven't bothered to do the readings, they haven't bothered to turn in their homework assignments, you know that sort of thing. And then, they come like “I don't know why I did so poorly on the test” and it’s like really? It’s pretty obvious to me. But that’s not specific to data in particular. I think the thing that’s most relevant to classes involved in data analysis is just students’ fear of math. And that’s going to vary across the divisions, right? I mean, I’m sure engineering students have a lot less fear of math than humanities and social sciences students do.

I: Okay, so how do the ways in which you teach with data relate to the learning goals of your field more broadly?

UCSB6: Ah okay. Well in my field, which is similar to a lot of social sciences, there are kind of two sides or two approaches to research - more qualitative approaches and more quantitative approaches. Historically, my department has taken a quantitative approach to communication research. It’s actually kind of in the process of changing right now. But this class was set up back when - or the purpose, really the reason that we set up the class that I teach in the lower division level is because a number of faculty in our department teach research - quantitative research - so they’re presenting findings of studies and they actually have students read the quantitative, you know, the journal articles that are kind of presenting quantitative research and communication. And so, the idea was we want students to have a basic familiarity with quantitative research and data so that when they are assigned these journal articles that present the results of experiments and surveys and things like this, they don't just kind of shut down when they get to the analysis section. You know, if they’re seeing a bunch of Greek letters and they’re seeing P’s less than .05 and they have no frame of reference for this, they often will just give up. And so, that's why we developed this class, which I actually the class that I teach, I actually kind of bill it more as more of a statistical literacy class than it is an actual math class. It involves some math, but like I said, basic simple stuff and what we’re really teaching them is kind of like “why do people do data analysis?”, “what kinds of questions that’s useful to answer?”, and “how do we interpret the results?”.

I: Okay. So you mentioned that this is specifically relating to their needs as students as they engage an academic text or literature, but do you ask them to think critically about those sources that they encounter in their everyday life as well?

UCSB6: Yeah actually, very much so. The first day of class, I kind of go through like you know a lot of you out there are probably wondering why do you have to take [course name]... “II want to major in communication or another special science, so what good is knowing about [course name]?” I do this whole song and dance about you know both what I said earlier about you know when you get into the upper division, when you get into a major or you may have to read these kinds of things, but besides that we’re surrounded by statistics all the time. We see you know reports, media reports and then if you go into your job, you are going to have to be able to interpret data if not you know collect data and analyze it yourself. You know, a lot of our students go into intellectual marketing and if you want to understand your markets, you want to understand consumer perceptions, you’re going to be, you know, if not doing the research yourself, you’re going to be trying to interpret research that someone else has done. And so I kind of give them a number of reasons including that it’s important to be a good information consumer. You don't want to just hear something and then take it in without being critical, you need to be a critical consumer to understand if you should believe what you’re being told.

I: So when you reference their potential future career aspects, are there...do you ever call attention to certain - I guess - exercises where you like “This is something that you'll be commonly asked to do and your professional contexts of…..”

UCSB6: Yeah, I mean not so specifically. I say you know a lot, if you want to be….like I try to work within the majors that, or sorry, the careers that Comm majors tend to gravitate towards, so that tends to be PR, journalism, marketing, I mean it can be anything, but a lot of those kinds of careers, and so I will say like….I’ve done this before like, “How many of you think you might want to be a journalist? You know raise your hand”, and then I say, “Well in journalism, you are going to be, you know, receiving information that you have to make sense of and then translate for your audience, and a lot of that information is going to have numbers, right? It's going to be statistical information. So I go through that. Or like I do the whole thing with you know if you're going into marketing and you want to do market research, you’re going to be collecting data from consumers, you know, or looking at sales figures. Or even if you’re Human Resources, you're going to be looking at statistics that are compiled about you know people within your organization or how your organization functions inside. I kind of go through some general examples like that, kind of hitting some of the careers that I know that our majors tend to gravitate towards.

I: Gotcha. Great, thank you. Have you observed any policies or cultural changes at your institution that influence the ways you teach with data? And you just mentioned the shift from quant to qual, but any other shifts that you’ve noticed as well?

UCSB6: Yeah, one thing that I have noticed and this isn’t in my own teaching, but I thought it was interesting, so the class that I teach - there isn't - is not the only class that's a [course name] that also fulfills [course level] requirement at the university. There are two other classes that are similar, and students therefore have a choice of which one they want to take. And I just found out recently that the one that's being taught in [department name] - I think it’s [course name] or something like that - has added sort of a big data component, which I thought was really intriguing and I wanted to find out more about how they're doing that because that's something that I think could be pretty pretty exciting to to work into my class. So I haven't done anything, COVID, you know, has just made it such that all I do is try to keep my head above water with regard to teaching, but you know, once this settles out, I think I'm going to reach out to them and ask to see, you know, maybe talk to the instructor to see what they're doing with that because there's a fine line. Again, these are students who are not strong enough, resent having to take anything related to math or data analysis and I don't want to overwhelm them, so I want to see how they, how they reached out to those kinds of students with big data. So it’s not an institutional thing, it’s just you know a [department name] made that choice to incorporate some of these new more computational kinds of approaches into their data analysis class.

Yeah...But institutionally, I can't really think of anything that's changed.

I: So this is kind of related, but do instructors in your field face any ethical challenges in teaching with data?

UCSB 6: ...not sure what that question is getting to. All of our data pretty much involves human subjects, so we get human subjects approval. This is for our research though. Teaching...the only thing that comes to mind with quantitative data it’s anonymized right like we're not really associating identities a person identifying information with it in a teaching context, and we’re very careful to avoid doing that in a research context also. The only thing I can think of is if you’re talking about maybe interview data where you haven’t secured permission of the interviewees to show that in class, like again, this is more on the research side. I've been involved in research projects where we have collected data in which participants were videotaped and we are very careful to ask and we get permission ahead of time for them to say yes, or no, uncomfortable, with your showing clips of my video in a classroom. So that's how we typically… I don’t know. Can you give me an example of ethical issues?

I: I guess one of the things that came in my head was actually not in reference to this intro course but into the honors course, especially when students are being asked to collect or generate their own data set essentially.

UCSB6: Oh. Oh. Ok.

I: We know with your other course, you supply a dataset that is 20 years old and has been anonymized, so for me I don’t see a connection....

UCSB6 and these are not real peoples’ responses. So, for the [course level] honors project they go through human subjects training and they have to get approval. We follow all ethical guidelines for conducting research and data collection.

I: ...and then for the sake of exploring it, you mentioned there was an individual who helps them with the analysis portion of that and that’s the only other area I could see… maybe less-so of an ethical issue.. But how they have another individual stepping into the process and whether or not the IRB process that we’ve just been through for this research project as well, and how that shifts things.

UCSB6: Those are just datasets that are anonymized. There are no names associated. The faculty member who also has human subjects training in general for his own research is never seeing the names of anyone. Literally like a spreadsheet of numbers.

I: We are now into training and support. Time check, it looks like we are doing pretty well. So, in your courses does anyone other than you provide instructional support to your students in obtaining or working with data, of course you already mentioned this other individual for the honors class, but are there others outside of that individual?

UCSB6: TAs. Whoever is teaching, a TA in the class would be involved.

I: How does their instruction or support relate to the rest of the course. I mean, it sounds like, if they're doing... are they mostly leading individual sections? I know TAs kind of function in different roles depending on the course. So, could you explain that a little bit further?

UCSB6: In our courses they lead sections and so in my data, in my classes that do data analysis they are… they go over the homework problems with students and help them understand because students do data analysis by hand there and they lead the lab sessions. So, they are the ones teaching SPSS and showing them how to obtain the outputs to then interpret.

I: So then how are you communicating with them about this instruction or the support that they are having to provide to the students in this process?

UCSB6: So, we have… so whenever I get new TAs I meet with them and I ask them about their own you know experience, comfort level, familiarity with SPSS and all of that. I give them access if they don't already have access to SPSS for example. Typically, the way that our TAs are assigned is it's a match process, so I’m typically not faced with a situation where I have a TA who has zero knowledge. So, it's, they typically come in with some kind of background and a comfort level with [course name]. So that isn’t too much of an issue. I have had… it varies a lot. You know, a huge range, and so I have a weekly meeting with my TAs and in that weekly meeting we... my primary goal is to make sure that they are ready to go into those sections/lab sessions with the students. So they they do the assignment ahead of time themselves. So, they will go through the computer problems, they will work them out. I also provide answer keys, sample output for the undergraduate students as well. They have instructors manual and then I ask them every week, I tell him you know, “Go through the problems ahead of time before our meeting so that if you have issues you can bring them up in the meeting and we can go through it together.” Because the class stays at such a simplistic level, I haven't really had to do very much. The TAs have been fairly comfortable. I haven't had a lot of issues with the TAs. Once or twice, I will have a TA come and say, “Ugh, can you just double check” or “I don't understand this…” and I meet with them maybe either privately or just in the context of our regular weekly meeting to make sure that they feel confident walking into the room.

I: I’m curious, how many sections are you working with a quarter?

UCSB6: How many TAs, or?

I: Yes… that’s a good question. I was making an assumption that there was only one TA per section, but maybe that’s not the case.

UCSb6: So, we have sections weekly and it depends on our enrollment. For every 75 students we get we get one TA. So right now I have a class of 75, I have one TA. But that one TA teaches three sections. So, in other words, the sections are 25 students each and the workload for the TAs is to teach three sections per week. They always occur at the end of the week, like right now. There's a section running right now that [section times]. So they are two-hour sections, so it’s six hours that each TA spends with 75 students.

I: Ok. How many times is the course offered in an academic year?

UCSB6: It’s offered either [range] times a year. So this year it is offered [number] times. I think that my department has kind of … it used to be offered [number] and now we're offering it all [number] quarters.

I: Okay. So, to your knowledge are their ways in which your students are learning to work with data outside of this formal coursework?

UCSB6: I think that would just be on an individual interest basis and only if the students chose to share that with me would I have any insight into that. So, maybe.

I: But you haven’t heard?

UCSB6: Nah. You know, it’s not the sort of thing a lot of students would share, and by the way, our classes are huge. So, we don’t really get to know our students. Sadly its one of my biggest frustrations with the university. You know, a small upper division class is 80 students. We’re not having idle chit chat a whole lot personally unless they happen to come into office hours and ask. I don’t get a lot of students mentioning to me off the cuff that they’re training… you know, teaching themselves data analysis.

I: Do you ever encourage, for the sake of following up on it, do you encourage them to other resources that they could reference outside of the coursework you are providing if they wanted to get further training?

UCSB6: I mean if a student did come to me, and if they were to say… actually, a lot of my grad students do that. They say, “This intro course has really got me excited about stats, how can I go learn more about this” and if an undergrad did the same thing I would tell them the same which is - there are a lot of other classes here at the university that you should look into. There is a great new data science program that the university has, I’m forgetting the actual name of it, but they have an intro course to data science and you know they can take P-stat courses, they can take classes in psych, they can take classes wherever they can find them if they want to pursue that interest. I would refer them to … the Comm major really doesn’t offer... at the undergraduate level doesn’t offer anything else. All of our stat classes besides this intro-level course, the course that I teach, are at the grad-level. I just tell them to explore their options.

They can also nowadays do like Coursera or other online options as well if they want to pursue that. That is not something that our students as communication majors really ask a lot about.

I: Sure. So this is pivoting more to your own training now. Have you received training in teaching with data other than your graduate degree? Examples would include a workshop, a technical support program, or help from peers.

UCSB6: Yeah, here and there I have. I think that it's actually sort of tragic that we don't have more opportunities to do this ‘cause statistics have changed since I was in grad school in the [decade]. There are a lot of common statistics that are used today in my field that I was never trained on because they did not exist. So, I really wish that there were more opportunities. I say that... there are opportunities and I am aware of them. There are R workshops, there are... I actually have participated in... there is a summer workshop that was done through the education school on specific topics, and they're really helpful. It's just very very hard to fit it into our busy lives. I mean we're not just teachers we're, in fact we’re first and foremost researchers and that’s a full time and half job by itself. Then we’re teachers. That’s another full-time job. Then we’re asked to you know administrators and we sit on all kinds of committees from the department to the campus to our field-level. So, fitting that in is very very difficult. So, typically...usually for faculty in my field if you need to use a statistical technique or a data analysis technique that you don't know, you either teach it to yourself. You can reach out to colleagues that know, or you can look online, or you can try to participate in... the I've been trying to participate in an R workshop for 3 years now and every single time they are offered during the time that I teach. Every time. Or I’ve started to do some computational work or work that requires computational analyses with big data and I don’t have the skill set to do that. I partner with computer scientists. So, I collaborate. I find people who have the data analysis skills that I don't. So that's how I do it.

I: This is a slight deviation, I think. Do you use any data sets, assignment plans, syllabi, or other instructional resources that you've received from others. Or do you make your own resources available to others related to teaching with data to undergrads.

UCSB6: I would say I do both. Like the class... I inherited these materials when I started teaching his class from others and because I felt they worked well I kept them. If I didn't feel they worked well I would have looked for something else or created my own. And as other people have begun teaching his class if they have asked me to share my resources and datasets and all that I'm happy to do so. So, I can imagine like if I were teaching some kind of specialty course and I needed to create resources, I would do that myself. So, I think it runs the range, but the easier thing is to take stuff that has already been out there and use it. But if you need to create something yourself, then we have the ability to do that.

I: This is a big picture question. Considering the evolving trends in your field that you have just referenced, what types of training or assistance would be most useful for instructors who are teaching with data?

UCSB6: I think that the workshops are good and I think that they should continue. I've just been frustrated that the timing, literally I don't know what it is, but the timing has never worked out. I do think it would be nice if there was a bigger effort on campus to... toward what I would call maybe data professional development because there are changes that happen in the data analysis world and legs we all kind of scurry around teaching ourselves. But it would be nice if there was kind of a one-stop-shop or a place that we knew that we could go get help, or that ran regular workshops that might vary the time here there - that would just provide more support for those of us who need to learn new techniques. Typically though, the need to learn new techniques is more in the are of research than it is teaching. But there is a lot of overlap between those two for people, so it could be both.

I: Final question. Is there anything else from your experience or your perspective as an instructor on the topic of teaching with data more broadly that you’d like to share?

UCSB6: For me I think that a number of students, or the type of students I teach, that is students who are not gravitating towards the physical sciences or math majors or engineering, so in other words, I’m talking more about the humanities and social science students… I think an emphasis on the conceptual understanding of data analysis and not just the procedures and not just the math… I have a graduate student right now, or there is a graduate student right now in our department who...she’s Chinese and comes incredibly, amazing data analysis skills. It’s incredible. Can pick up the most complex, or do the most complex analyses. But whenever I say, “Ok. Tell me what that means.” It’s like hitting a brick wall. We have a good relationship.

We tease each other, but I’m always like, “It’s great that you have these skills…” and she’s always like “I want to learn more skills, more skills, more skills!” I’m like, “That’s great! Do it. But now you need to translate. Now you need to be able to take those math skills that you have”... you know looking at matrixes and complex visualizations “and now you have to be able to tell somebody...tell somebody that has not had mathematical training what the heck these pretty data pictures mean. Or what the heck these numbers mean.” And she’s like, “Oh well, like you can see that this number is higher than that number” and she stops there. But, what does that mean for what you are researching? So, I really think that teaching students data analysis has to not just be here are the procedures or here’s the math of it and not just… you know conceptually you have to somehow marry those two things together and emphasize both equally.

I: Great! Thank you again for your time. I’m going to go ahead and stop the recording.

Interview 7

I: Okay, so I have a little note before we dive into the questions about COVID 19. I just wanted to start by acknowledging that teaching and learning during the pandemic has been a significant disruption in the past months. And for any of the questions I'm about to ask, please feel free to reference, with your normal teaching practices rather than what you're having to do right now to adapt to the crisis. Does that make sense?

UCSB7: Okay, so yes so do you want me to tell you whether. So, what are the regular, what is the regular practice that I do without covid during the pandemic and now.

I: You can do either but if you're able to draw upon what you've done previously more easily because you're not having to face the challenges of trying to teach online that's perfectly fine too. So whatever touches most relevant to your work.

UCSB7: Okay. That’s fine.

I: Alright, so I'd like to start by hearing about your experience teaching undergraduate students, briefly describe your experience with teaching them, including your current or past research.

UCSB7: So teaching, the teaching that I've done is mostly on research methods in sociology, the students do not necessarily have a background in statistics. I touch upon like basic basic concepts of survey research and sampling and things like that. I tried to teach them more like the general concept and not, and I try not to go through the formula a lot, because there is a lot of uneven students come to the classroom with very different knowledge on the statistics. And so, that is a like, based on that I tried first to level up the ground on using basic concepts and then I go on and teach them a little bit about formulas that are behind but very little, very few.

I: And what are these classes that you're doing the research methods? Can you give me titles of the courses?

UCSB7: So the title of the course is [course title] and this is an [course level] class that students need to take, this is an obligatory course of a series. They have to choose different methods classes and this is one of them. And in this case, I go over different methods of survey methods. And given that my specialty is in survey research. Half of the course, I teach research and sampling and mostly like history, publishing studies.

I: Sounds like we're having some trouble with that connection again. Yeah, it's breaking a little bit. Okay.

UCSB7: Do you hear me now?

I: I can hear you a little bit better now so we'll try. The last thing I heard was about the last class is your teaching work for [course level] undergraduates, they have to choose between two different courses, is that two different courses, about statistics or about research methods you were saying?

UCSB7: Yes, so there are like about like five or six classes in some methods that teach the students have to choose to take. And one of those methods classes is the one that I teach, which is a survey of method so we do quantitative and qualitative methods in sociology, right here.

I: Alright so we're going to move on to a section about how your students obtain access to data, so in your courses, how do your students usually obtain the data they engage with? For example, do they collect or generate the data themselves, or do they search to identify pre-existing data sets, or do you simply provide them with the data sets?

UCSB7: So in the other undergrad course that I teach, I show them different databases that are available to the social science research. And I asked them to engage with one in particular that is called the General Social Survey. And I asked them to play a little bit with the variables in that survey, and to be able to like to do contingency tables, basically like two-by-two tables or at most, maybe using three variables. But I provide the data sources, I showed them the resources that are available there. And some of them know them, but most of them, they don't know they don't know all, that all the data is available, we have to show them. And I also ask them to collect some data, they do a little bit of data collection in the survey, when I teach survey methods I ask them to collect some data. But, that is an exercise that I asked them to do in the class.

I: Okay, so you're mostly providing them with like a pre-existing data set, and maybe having them do a small amount of generating their own data set is that accurate?

UCSB7: Yes. Okay.

I: Can you describe a little bit about what, you've kind of dove into it a little bit with the two table, or the table where there are three variable and there are two tables they are viewing simultaneously. Can you describe that general process a little bit more like what you're asking them to go through when you're having them look at a data set such as that?

UCSB7: Yes, so the General Social Survey is actually very friendly and then they already provide, their website already provides ways to construct tables. And so I, what I do is I have TAs and they teach the students how to construct this table so for example if you want to see. You might either want to see educational level and race and ethnicity. This would be a two-by-two table. We'd put race in one, in columns and rows educational level for example. So, I asked them to do these kinds of tables to see whether, for example, Hispanics are more likely to have lower levels of education and things like that to see like the relationships among variables. So, I ask them to construct these contingency tables. And, and I usually ask them to choose a topic they're interested in, to construct these tables based on those interests if they want to see the relationships between two or three variables.

I: Okay. So do you have them, in this case, this is an example of where they're, they're creating their own data set. Is that accurate or is this something where you have given them that data set before they're manipulating it to create the table?

UCSB7: No, no, this is an exercise that they do using a survey that was conducted by, that is out there, called the General Social Survey.

I: Okay.

UCSB7: And the idea is that they, they learn, first of all, how to, how to get access, they will have to have access to these. So, to handle these types of datasets that have a lot of variables. These General Social Surveys have a lot of different variables and they can cross tab these variables and look at the relationships among them. So, they usually choose a topic that they are interested in. And then they choose from the variables that are available, so I teach them how to navigate in the website of the General Social Survey. How to choose… how to choose the variables. They have to go and see which questions were asked, how the variables were named, and how to construct the tables, cross-tabulations.

I: Okay, that's helpful. Thank you for explaining that. In this scenario where the students are actually creating their own data set, could you describe what that process is like for them.

UCSB7: Yes, what I asked him to do this is, there is a section on survey research. And what I ask them to do first of all is to test, to create a questionnaire the idea is to design and test the questionnaire by conducting about like 10 interviews and to correct the questionnaire because people understand things differently. If the questions are not, are not very clear, people may not understand them so I ask them first to pilot, to make a pilot of the questionnaire, and once they correct the questionnaire, then they will get, they will interview maybe around like 20 people, not a lot of people this is not like, this is only an exercise of how to design a questionnaire and how to gather data. They might collect, at most, 20 interviews.

I: And then, you mentioned that they're piloting this survey. Is there, I guess a process by which you're having those questions vetted at all? Like, their pilot is that really their way of testing whether or not their questions get to the data that they're seeking, or is there some other like scaffolding you provide them before that pilot.

UCSB7: No Well, the idea is for, for them to realize that sometimes we think that our question is very clear and everyone is understanding the question in the same way. And when you go out to the field and you start to ask that question then you realize that your question is too long or the question is understood in very different ways that you are capturing other stuff that is not the one thing that you want to capture the concept that you want to capture it's more like, how do you operate operationalize that concept in a way that is reliable. That is the exercise.

I: Sure. Okay, great. In these scenarios both where you're providing them with the data sets and those when they're constructing their own, what have you found to be the greatest challenges and, and the one side finding the data sets for your teaching and then the other, you know, having them generate it. What's the biggest challenges with both of those?

UCSB7: I guess, these, these exercises are complimentary. So the first one, the first one requires skills of how to handle databases. So, the first one requires, the first one is designed for them to learn how a questionnaire looks, how those questions are coded in different ways, like how the coding is done. And then how you can recode those variables in order to be able to create a table or answer the question that you want to answer. So, it's more like on the side of using data that is available, and learn a little bit about how from a question you translate that into codes, and do the coding, then you translate that into a statistical analysis. So that is the first approach right when we look at secondary data. And in the case of survey research, where they are collecting data, my focus is more on how to do a questionnaire, how to write a question that is reliable, so the exercise is more on the survey design and how we can get questions that are basically valid. And it's different, the exercises are complimentary, so one one has one objective and another has another objective. And so the challenges are different so the challenges in the first one, are more like technical challenges.

Because they need to get into the website, learn how to use the website, learn how to read like how the variables are constructed, learn, they learn a bunch of skills about the data, how to manage data. And that is challenging, in the sense that sometimes they don't know how to navigate the websites. It applies technical skills, and then statistical skills, because they have to create a table and be able to read it. So, that is more like, the challenge is to like teach them how to do, to learn statistical tools but also data management tools, skills, and the other, the other exercise the challenges. The challenge, right now, during covid is that they will have to do this with only the people they know so that they can be safe and all that. But before they had to go outside and I think that it's less challenging because it requires listening skills. And they realize that people respond by understanding the questions they have to reframe the question. It is more like the process, learning process in the field that is more natural and it requires less technical skills, I guess.

I: Okay. So, we had a little bit of a glitch in the last response. I just wanna make sure I captured your response to that. The other side when they're out there in the field their greatest challenge right now you said with covid is that they can really only realistically ask people they know, for the sake of being safe. But before then, could you kind of repeat that a little bit just because we had your, your audio cut out just a tad.

UCSB7: Now before they, they just go like usually the final tests that they do to conduct the same in the same university or in the inner neighborhood. So the pilots they conduct with people they know and then when they collect the data they try to find people that are the appropriate sample that they want to target, because they are interested in that specific problem. So, they will need to go out and find the right people, that is, the people that are interested in surfing, before they could do that. Now it’s more difficult.

I: Right, right. I would imagine so. The next set of questions is more specifically to do with when students are actively working with the data. And in this question we're looking for what kinds of tools or software you're having students use, for example Excel or online platforms, or any analysis visualization tools.

UCSB7: Yes, for undergrads because they, it's very uneven, the levels we have when they come to my class. What I teach them in class is to use a platform that is already available online. Which is basically, I, I use the General Social Survey as the main platform in the class, I don't use any statistical package or, I don't teach them how to code because that would be very difficult for a class like this, because it's a survey of methods. I go over different kinds of methods that are quantitative and qualitative in one quarter, so I cannot devote a lot of time like in programming and doing a statistical analysis. So I basically devote like maybe, I don't know maybe like five weeks to survey research and measurement and all that but to analyze data I devote maybe like three classes, and that then data analysis. And so the, the, the way that I do it is I provide like this, in class I show them how to read cross tabulation. Basically, what is that dependent and independent variable and my TAs in their sections go over how to use specifically this platform that is for the data that I use the General Social Survey, they have this online platform where you can create tables, like once you choose the variables that you're going to use, you create the tables and you can also create graphs, so you, they have like these two things that you can create and this is the thing that they have to provide. In one of the exercises that they do for my class, they have to answer one research question that they have and they have to tell me, which are the variables that are involved, and how, and they have to provide that graph. And sometimes I ask them, well, I ask them to try to replicate the graph that these websites produce. I ask them to replicate it in Excel. That is very sometimes very challenging for them. Many of them don't know how to create a graph in Excel.

I: So just to clarify, earlier with the survey you were having them do, is the idea with the compilation of data with that survey that they have some quantitative data, as well as qualitative or is it dependent upon their topic?

UCSB7: So, when they do the survey, when they're collecting the data, it's more, the exercise is more focused on survey design and questionnaire design, it's less about analyzing the data. It's more like piloting a questionnaire and designing questions that are reliable.

I: Sure. And then just to follow up on the platform you're describing, did you provide... What is the name of it, just called the general social science?

UCSB7: Yeah, the name, the name of the service General Social Survey.

I: General Social Survey. Okay, great.

UCSB7: And the test, it is a representative sample of the US and they have information for every year like since a long time ago. So actually they can see changes like because it changes for a period of time, it's, it's a very interesting subject because it has many topics, included.

I: Okay, great. So, you kind of have addressed this a little bit already in the very early part of our conversation but what prior knowledge of the tools or the software do you expect students to have and then which do you teach to them explicitly?

UCSB7: Most of them, they don't know how to use the General Social Survey. Not all the students have, most of the students do not have experience with data sources like these. So, I start from scratch, I show them what is the question, how you download the questionnaire. Then I teach them how to do that. They need to know this is a representative sample, what is the population frame, then I go over the website and show them how to find variables based on the question or the questions, they are interested in how to read the tables, I go over all those things from scratch.

I: And then in terms of, you mentioned that they, in one exercise you've asked them to kind of recreate a graph that's in this platform using Excel so do you kind of expect that they have some familiarity with Excel?

UCSB7: Yes. Yeah, but I know now that many of them don't know how to use it. And sometimes I ask that they have to do that because the graph that the platform provides sometimes has categories that shouldn't be there, based on the research question that my students have. So, they would have to redo the graph. Basically, because they will have to get rid of one category or something, based on the research question so they will have to redo the graph. And also because presenting graphs in Excel sometimes they look better. So, the presentation is improved, if you know how to do it, like in Excel you have the flexibility to do it. Perhaps they are better formatted, you can have more flexibility. But they don't know, like many of them, they don't know how to use Excel so I have to provide examples of how to use it, and I basically do that in the sections like the TAs are the ones that show them how to use Excel, how to improve the graphs and the tables.

I: Okay. Are there other tools that you're having them engage with that you find yourself having to teach too in those sections or mostly is it Excel and then this General Social Survey?

UCSB7: Well, I have changed over time like this is the last iteration of it. Yeah, the last iteration. At the beginning I thought I could show them a little bit of programming and that didn't work. Obviously because you need much more time to do that so I changed it. And now I found this platform and I think this is ideal because you can at least show them the basic concepts of a survey, how it works, how it is coded and how you can use actual data in order to be able to read a table.

I: So I think I might know the answer to this next question but I'm going to ask it anyway. To what extent are the tools or software students use to work with data important for your teaching.

UCSB7: Well for this class in particular because it's a survey of methods. You don't have to go deeper into using data. It would be great if the students could already know how to program, how to use statistical packages. But I think that for the purposes and the goals of this particular class because it's a survey of methods that will not be necessary. But for other classes, for other methods classes that require more, like for example if there are just a lot of classes that are only survey research. It would be very useful for us as teachers if students already knew a little bit of coding, even if they knew a little bit of Excel. But for example with Excel, it seems to me that that would be very very useful if they knew of any of those things.

I: And then, do you face any challenges related to the students abilities working with data in your classes?

UCSB7: I think they are not very experienced in using data, I think that, I don't know maybe it's because it's sociology and people think that in sociology, you will not be required to do these kinds of things like data analysis. Most of the people, of the students that go to sociology at UCSB, they are more, they want to do more qualitative than quantitative work. And so I think that it's also like the profile of the students that they don't, they're not looking to do quantitative work. When they choose sociology, they don't know that actually sociology includes a bunch of like, half of sociologists do quantitative work.

I: Sure

UCSB7: But probably they don't know that and one of the problems that I think that I've come to realize is that this shouldn't be an upper division class, it should be a lower division class, so that people can choose how come they then the students are more aware of all the possibilities that sociology has like it's not only qualitative work, we do a lot of quantitative work. And if they have that since the beginning they would have known that you can do this and that -

I: It looks like we've had a little glitch in our audio again. Erica, are you still there?

UCSB7: Hey there. Yes.

I: Hi. I lost your last bit there but you know if it's, rather than trying to have you repeat it because you were kind of leading into what related to what my next question is do you mind if we move on to that one because that way you can elaborate more on what you were just saying?

UCSB7: Yes

I: So this question is about how do the ways in which you teach with data relate to the larger learning goals of your field and the last piece I had heard was students that come into sociology at UCSB tend to think you primarily work with qualitative data but that there's this whole other branch of sociology that specifically deals with quantitative so maybe if you could elaborate a little bit more on how the learning goals in your field kind of relate to the, what you're teaching in this class.

UCSB7: Yes. So, what I was telling you is that this class that I'm teaching, I think should be a lower division class instead of an upper division class. Because when you have freshmen and then you show them like the different methods and how these different methods answer different questions they might be exposed early enough to start to be interested in looking at taking more statistical classes, or other quantitative methods classes. But the problem is that they, they - Sorry, I’m in my house and I have to move, I am out of breath right now, haha.

I: Totally Okay

UCSB7: So, so I think that they need to be exposed early on in their careers to all the options that they have, because if they do it the way that it is right now, they just, they don't take the courses that need to be taken in order to be able to understand better statistical methods in here.

I: And I just want to make sure that also to go back to this question about the ways that you're teaching data represents the learning goals in your field and that's part of what you're trying to express I think here right is that by exposing them to this earlier, kind of, it's a better representation of what people in sociology are studying and how they go about conducting their work, is that right?

UCSB7: Yes, because right now it's, everything has a certain degree of bias towards qualitative methods and qualitative methods are geared to answer some questions but not all, and sociology includes a bunch of other things that are related to the population studies and quantitative methods and other questions that you also want to know about social structure and how social structure has an influence in individual behaviors. And so quantitative methods are better suited to answer those questions but students, because they don't take, I think, the methods classes early on, they don't realize that they can take other classes and statistical classes to be able to do these other kinds of sociology that is out there.

I: Okay. So this is a little bit of a change of course, how do you teach your students to think critically about the sources and uses of data they encounter in everyday life?

UCSB7: Right now I'm personally trying to… I've been trying to redesign this course so much so that they are interested, and they can see the usefulness of learning these types of methods in their own lives. When they read for example, when they get their news from the newspaper when they see the news on TV like I try to, right now I'm trying to incorporate more examples of everyday life, like everyday news. What are their interests, what are their opinions and I try to compare those opinions, what are the sources of their opinions. And then, how would they create a study where you can actually test those ideas and then I come across with studies that are already done conducted using quantitative methods for example, to answer those questions so that they can see that their opinion, maybe like bias towards their experiences or what is in the news and at the end, the way we are scientists, answer questions. Following these methods, in particular quantitative methods, may change our opinion of how something works because they are able to compare their own opinion at the beginning. So, I try to ask them what is your opinion about this topic at the beginning of the class. And then we go through the whole lesson and then I show them a paper about that question. And then they compare what was their opinion, and what does a study well conducted say about that phenomena. And they can compare. So, that is one way that I'm doing it.

I: Nice. And then how do you teach, or do you teach students specific data skills that will prepare them for future careers?

UCSB7: Statistical skills, so this class in particular is not designed for that. Like, at most, what I can do is to tell them okay if you are exposed to, if you read in the newspaper or if you are exposed to... Well, if you see a table to answer one specific question. What are the things that you need to ask yourself in order to know their study is well conducted? You know if you listen, if you read the newspaper, how can you know that the study where this is based from or this opinion or whatever, how do you know what this data represents? So that's the goal of the class. This is not a class where they learn how to manipulate data.

I: Back to that notion of being a survey class more so than, you know, targeted towards a specific research method design principle.

UCSB7: It's more like if you hear a news story, if you see a study like, what does that data represent? Or if you hear or post a survey, how are they conducted, what are they saying, what is the scope of the study? Like those things are the things that I asked them to do, because they hear that this is a survey and they believe everything that the survey is saying but at the same time they say well what is what is the population frame. What does this data represent? Is representative of what and things like that. There's a lot of attrition in the data like is this representative or not, there's a lot of mortality. What does this represent? That is what all the time I'm asking them to do when they are looking at data.

I: Right. Yeah. Have you observed any policy or cultural changes at our institution that influenced the way you teach with data?

UCSB7: At our institution, I don't think so. No, no, no, no. Right now I will try to push, just for example like to try to do this class instead of upper division, lower division because I think that is important. I will push for that. But changes, no. I've been trying to, in our department for example, most of the department is qualitative. And we have very, very few quantitative researchers as members of the faculty. Faculty members are very, very few quantitative and that has been also a problem because we need more. In order to be able to be stronger in quantitative methods in sociology here, and to be able to get more students that are quantitative that want to do quantitative work you need more faculty that does quantitative work. Because if not, you are signaling that this is a qualitative department. And so people that want to do qualitative work come here but we never have students, for example, that want to come to UCSB because we are very, very, like, around [number range] members that do quantitative research. It's very, very small. And it would be great if there was, like, maybe like with the previous Dean, he wanted to hire more specifically people that do quantitative research. And now that is gone. So, it would be great to have more people doing quantitative research in this department and I think that would help.

I: Okay, so then do the instructors in your field face any ethical challenges in teaching with data?

UCSB7: Ethical challenges, teaching with data.

I: Specifically, like you were mentioning, I mean this isn't, not an ethical challenge but your other challenge of not having enough representation in the department about quantitative sociology research but any other ethical orientations or challenges with teaching with the data?

UCSB7: With teaching, well, all the data that I use in class is data that is publicly available. When they collect the data they do it anonymously and I ask them to do the IRB training. To all my students I asked them to do that when they go to the field to pilot the questions. Basically, those surveys are anonymous. That is not a problem. So I think in that sense there is nothing, not that could be problematic. They basically do those pilots at the university and when they go out to see more specific populations that they are interested in, they are interested in, they have different interests. But I ask them to collect information in an anonymous way, and to get informed consent. So all the protocols that we need in order to get data. But it's more like an exercise, they don't gather a lot of data like that.

I: Thank you. So, our next section is on training and support and I'm just cognizant of the time it's 12:44 so I think it'll be about 15 more minutes if that works for you.

UCSB7: Yes.

I: So in your course does anyone other than you provide instruction or support which it sounds like in this case you have TAs. And for those people, obtaining or working with data, for example, the co- instructors, teaching assistants, or even librarians. How does their instructional support relate to the rest of the course, so maybe just describing a little bit of how the structure of the course runs with these TAs.

UCSB7: Yeah, so basically the TAs. So, the first part of the course, I ask them to teach the students how to learn basically the platform of the General Social Survey. So, first of all the TAs help me to define the research questions. I divide my class in different groups that work together, of maybe pairs or three, or groups of three. And the TAs are in charge of, first of all, helping them define our research question. And based on the questionnaires that are available in the General Social Survey so they helped me first teaching the students how to navigate the website of the General Social Survey. In order to be able to see the questionnaire first, so the topics that are available there. And then they kind of like put students into groups of students based on their interests. And then the students will develop a research question using the help of the TAs. And then they will, the TAs will help them navigate the website in order to be able to find the variables and to be able to produce tables, so that will be like the first thing. The TAs help me do that, like all the training using the website and defining the research question. Then for the survey research design, they also help the students. But that is more on the part of the students so only if the students have questions, they will ask the TAs but they will only go over the main concepts of how to design a survey, know the specific ways that you need to frame questions in order to avoid answers that are not understood. To avoid different meanings in one question. So there are like specific rules that you have to follow. And so they will go over that. And then, the students will do the pilot and they will change the survey, depending on the results. And the TAs will also help me so in the qualitative part of the course, but this is not related with the quantitative methods, it's more like in depth interviews and things like that.

I: I see, OK, and then how do you go about communicating with them for the instruction they provide.

UCSB7: Yeah, this is something that has changed over time.

I: Right. Feel free to refer to both obviously because I know what you're doing now is likely very different from how it was before March.

UCSB7: I mean, I've been changing that a lot. I have been teaching this class now for three years and the class looks very different than it did at the beginning. So, my communication with the TAs, I basically teach them because they don't know how to use the General Social Survey, so I show them how to use this General Social Survey. I ask them specifically what I want from students to do. So, all the things they teach to the students, I will teach to the TAs.

I: Okay. And then throughout the quarter, how does that communication change after you've done the initial, kind of preparing them?

UCSB7: So, it depends also on the TA, some kids are more independent but I usually like to talk to them and ask them. So, what are the issues with the students, if they are understanding the material. And over the years, what I've changed is that I'm like, I have to be more clear and give more clear expectations of the exercises. So I'm like, okay, if you produce a table, you need to include... I have to be super specific. You need a paragraph showing what is the specific relation between these two variables. I have to develop rubrics, like very clear rubrics, so that the TAs know specifically what I'm looking for and they can clearly convey that to the students so that there are not misunderstandings. And I think that that has helped a lot. So once they know exactly what I want and how to grade, and all that, then, that works better for the students.

I: Great. And then to your knowledge, is there any, I guess formal training or informal training that these students are doing outside of the, you know, help that you've given them to learn how to work with data for this course.

UCSB7: Um, I don't think that for this course they require that but some students struggle more than others for sure. Some students, and that is a problem with this course like some some students really, really, really know like the basic concepts in statistics so average median, mode and all those things they think they know. And, and sometimes I have to go over those concepts, those are very basic because they don't know. So, the field is not very well. It's not plain. It's like you have people at very different levels. And it's a little bit hard to be done so that everyone is happy like people like I need to go to the very very basics. And some of them, some of them require more help outside and they actually go with the TAs and ask for help.

I: Okay. Do you expect or encourage any kind of extra curricular learning in that case when there's folks that maybe need a little bit more education in order to teach the concepts in your course.

UCSB7: Actually, what I've been developing right now. Now with the coronavirus and I think that it's been very useful. I've developed for example, many videos on how to use Excel, and then I go, this is the way Excel works. And this is the way you produce a graph and I'm telling you, I am doing like mini videos so that they can see how I'm creating those graphs, for example. I think those are very useful because they can watch them whenever they want. And they can just basically copy the exercise that I'm doing. And then they realize okay so this is the way you do the graph, for example. I was like like mini videos like that that are like, Okay, how can you do this specific graph or how can you format a table in Excel, those can like, help them, you know, to be able to understand better how to use Excel how to learn how to use Excel how to bring us like better graphs or something like that. For my course because it's a survey of methods, I wouldn't say that they need like a bunch of stuff, but for other courses like survey research or courses where the idea is to use data, they would need a lot more resources.

I: Okay, great. Those videos sound like they'll be very helpful going forward for you.

UCSB7: Yeah, right now I'm actually thinking of changing. Well, it will change because now it's going to be an online hybrid course. I think that for the following term when there's no coronavirus, I will still rely on that course because I think that the use of my time in person might be better if I do more exercise and applied work. And maybe they can see the lecture online. I don’t know, I'm thinking of changing it.

I: Oh sure, like the flipped approach like where they kind of prepare for the session, the session ahead of time, and then they come in that they're able to actively work through some of these exercises with you there.

UCSB7: Yeah, I'm thinking of doing that.

I: Interesting. So this is the last section related to your own training. So, have you received training and teaching with data, other than your graduate degree. Examples would include workshops or technical support or even help from your peers.

UCSB7: Teaching with data. Well, I just took, not specifically with data, I haven't taken a lot of courses here at UCSB in instructional development of how to teach, pedagogical strategies. So I took in the summer through the [teaching professional development conference] offered to develop courses online, but it wasn't much more than that. It was also like the strategies to like to be more focused on our pedagogy, to be more focused on the concepts that we really want to convey and to be more clear when producing rubrics and that kind of stuff and so I've been taking those, not specifically for data. But I've also been teaching the other statistics course for grad students. And so over the course I've been developing different strategies and teaching them how to program and these are grad students that really don't know, they also lack a lot of knowledge of basic statistics. They don't know how to program. They require a lot, they need a lot of resources in order to be able to take these classes and these are the basic statistics classes for grad students. So, No, I haven't. I haven't had any training, specifically on data, like, how to teach data courses or quantitative courses. I've been doing that and I've been changing and changing and changing.

I: Right.

UCSB7: Based on the level of the students. With the pandemic it has been, it has been a little bit challenging.

I: Sure. And then, for any of the data sets that you use or assign, how do you share any of these instructional resources that you have or have you received from others before.

UCSB7: How do I share them with students?

I: No, this is more along lines of sharing your instructional resources from others who teach, I believe.

UCSB7: How do I use other resources?

I: One second. Let me look at this question again so it says do you use any data sets, assignments, plans, syllabi, or other instructional resources that you've received from others, and then do you make your own resources available to others? I think the idea behind this question is like fellow instructors who may teach similar courses. Is there any kind of cross collaboration, like here's the materials that I used to help students learn how to develop a survey and then pilot that survey.

UCSB7: Yeah, absolutely. A lot of sharing.

I: Could you describe it? How do you go about doing that, I suppose, is it just within UCSB or is it outside UCSB?

UCSB7: Yeah so many of my colleagues, I studied grad school at [institution name] so we share like, we do a lot of methods classes. Actually like most of my classmates at [institution name] when I was a grad student, now they are professors at different universities and they teach the methods class, so we share the way that we are teaching and and then we kind of like, change it and adapt it to our university and our knowledge and also to the topics that we really that we know. It's not the same class but we share the strategies that we use. And that is something that always happens. We are asking for how are you teaching this or what do you recommend? And we actually for example in Facebook, sometimes we ask like I’m looking for something interesting for my class for my students to learn this concept, do you have something, and then someone will post yes I have these I have used these. So sometimes we are, we're sharing all the time.

I: Sounds like a great community of practice.

UCSB7: Yes, and also from my previous instructors like the classes that I took at [institution name]. So, they had incredible instructors there, and teachers and so I use their syllabi and then I adapt.

Also, I take from different, so I use different syllabi from them and they create my own.

I: Nice. Okay, we're getting close. Considering the evolving trends in your field, what types of training or assistance would be most beneficial to instructors who are teaching with data like yourself?

UCSB7: What will be more beneficial for me?

I: Considering the evolving trends in the field, are there specific types of training out there that you think would be beneficial to teachers like you?

UCSB7: Oh yes yes yes yes. A lot! Well, yeah, for example, so there are different methodologies, statistical methodologies that I would like to learn how to implement and there are new ways of implementing those. Using new programs like, for example, R or STATA and I would like to learn those but those are super expensive and I can't afford to take those courses. Those courses would, like it would be great if we could get money to get into those courses easily. Or easier, get someone here to teach us or whatever, like, that would be great. Right now I am part of the [name of organization] the population center at UCSB, and so we ask some students to teach some like basic courses on R. And we as faculty, not all the faculty knows R and R is like something that is free, it is a software that is available for free and that might be very useful for students because they don't have to pay any license. But for example I don't, I don't use it I would like to learn it. So trainings like that would be incredible for us to continue our training and also to be able to train our students in these more affordable and software that are available and also more powerful.

I: Okay, great. And last but not least, is there anything else from your experience or perspective as an instructor on the topic of teaching with data that you'd like to share.

UCSB7: I don't know, I think that I already talked a lot.

I: It was a very very rich conversation so I really appreciate your time sharing what you have, happy to hear more if you had anything else that popped in.

UCSB7: No, I think one of the challenges and this, this has been something very useful right now, I took the [teaching professional development conference] and they gave fellowships to the people that took the course in order to develop our courses and we got money. I've been able to hire a grad student and an undergrad and actually I'm redesigning my methods classes so that it is more interesting. I want to make the topics more interesting for the students. So before, because my specialty is in [research topic]. My class was very focused on [research topic] and I think a lot of students were not interested. So right now I'm looking for new topics that are more exciting for them, like for example, marriage markets and dating, and maybe gender privilege and white privilege and things like that. That are about the US, I'm trying to focus on the US and how you can answer questions using data based on these topics that are of interest for, I think, for them. And once I defined the topics that I want, I've been asking that my grad student and my undergrad are helping me to find some material, and that has been really useful. And now that you are asking, this is like taking over the previous question, I asked all my colleagues in the UC, in the sociology department, the colleagues that are doing like mobile work in different topics like nationalism and violence and things like that, racial violence and so I asked them, would you suggest an article, would you suggest this is a good article or a bad article so that I can include it in my class. And then asking my colleagues and having the help of the students to redesign the course has been great. And this is only because I got the funding from the [teaching professional development conference].

I: Very nice. I was very lucky to be able to participate a bit in the RISE Institute as well. It was really, I thought, really well put together week-long exercise for folks to do so I was glad to see it myself.

UCSB7: Yeah I really liked it.

I: I’m going to go ahead and stop the recording since that is the final question.

Interview 8

I: Alright so I think now it’s recording. I want to start by saying that teaching and learning have been affected in the last few months because of the pandemic so in any of the questions I am going to ask you, please feel free to reference your normal teaching practice or even considering how you adapted your practice for the situation, or both. So, if you want to give examples and explain how you did things prior to Covid 19 that’s fine or if you want to reference your practices before March, that’s fine as well. Okay.

So, the first question is could you briefly describe your experience teaching undergrads like how does your teaching relate to your past research or current research, could you describe the courses where you teach students to work with data.

UCSB8: Sure, so I teach a statistics course for the psychology department in the pre-major. So most of the students are probably in their sophomore year, transfer students probably in their junior year. And it's actually the last class that students have to take before they're allowed into the major. So, I call it a high stakes class so they have to make a certain grade in the class, usually in order to enter this like the psych and brain sciences major. So it is a statistics class. So, in that sense it's grounded in a lot of data. And we cover statistical concepts both conceptually as well as we do hand calculations of statistical tests, as well as the course has a lab section, and in the lab section the students are shown how to conduct statistical analyses in R statistical software. So, it varies from quarter to quarter. I have anywhere from 125 to 250 students in the class so this year is a little bit smaller. And I usually have 6 to 10 TAs helping with the lab sections. And the lab sections are set up for students. We give them a data set and they work their way through statistical analyses and then they report those analyses in APA style. I wouldn't say that I linked the class directly to my own research per se, so I've set it up a couple of different ways. So previously I had given them fictitious databases to work with. So just kind of data that I made up and simulated to work with but this quarter, what I actually did is I sourced data from faculty and Psych and Brain Sciences. So, they actually gave me data from published papers that they had. And the students are running statistical analyses with their data, with faculty members data essentially. So, it's nice to ground the data analyses back into some real world, Psych and Brain Sciences data collections that have happened within the department.

I: Great. So that kind of ties to my second question which relates to how students usually get data. So, you mentioned that you usually provide them these data sets, right, based on publication.

UCSB8: For this class, yes.

I: Okay, are all their sources of data sets, like did you have any assignments for instance that the students were encouraged to find a given data set, find like primary resources to conduct some of the assignments they have or will just say, this is usually your primary method.

UCSB8: Previously I've only given them data sets. The reason being is because they have a weekly assignment that's do every week and it's really intensive grading for the teaching assistants and so it's much, I think, easier for the teaching assistants, as well as probably fair for the students for all students to be conducting the same analyses with the same data so that the TAs know what to look for when they're grading those assignments and it kind of puts everybody on an equal playing field with regards to the process of analyses. I can certainly see a situation where it would be useful to have students go out and source their own data. But for this class because it's a pre psychology major for example we don't have them collecting any of their own data, they do that in upper division classes after they’re in the psych major. But what we're trying to do in this class is really just to get them comfortable with statistics and, you know, using R for the first time, those sorts of things. So it's more of a, I mean they have a stats class before they take my class. I believe that class I think it's PSTAT5 or something like that focuses a lot on hand calculations and probability. And so what we do in this class is try, what I do in this class is try to make it a bit more conceptual and grounded in psychological data, and the ways that we use statistics in our fields so I really try to expose them to the specific analyses that we tend to use a lot. And the way that we would write up and interpret those results in APA style.

I: And in terms of, like, providing access to these data sets. What are the most common approaches you follow like to provide your students with the data sets? Do you share through Gauchospace or email? How do you usually make these data sets available for your students, and the TAs that are helping you.

UCSB8: So usually I post the data sets up on Gauchospace. So, for example, I posted one up on the first week of this quarter and they've had a reading associated with it so the students read the empirical paper, I posted the data from the faculty member within the Psych and Brain Sciences Department, and the students have worked with that data, the same data set for a couple of weeks so running some descriptive statistics and T tests some one way ANOVAs.

And now we'll post up a second data set that the students will use for a few more lectures with an associated paper with it to run some correlations and regressions. So usually I just give it to them through Gauchospace.

I: And do face any challenges in finding or obtaining data sets for teaching, or would you say this is relatively easy in your field?

UCSB8: I think it's pretty easy. I mean like I said, in the past I've made up some data and sometimes I've used some of my own data prior to teaching here at UCSB. But this quarter I approached faculty members and said, this is the idea that I have, I'd like my students to work with real psychological data that faculty in our department have actually used and published with, can I have access to your data? And everybody was more than willing to share it, which was nice. I thought about using like national data sets that are available, but I think the number of variables and the number of cases would be really overwhelming for first time R users to open up a data set and see you know 1000 variables and 10,000 participants would be a little bit much I think.

I: Okay. So, we talked a little bit about getting access to data or providing access to data sets. Now I would like to shift my questions a little bit more about the working component so working with data. So how would you say your students manipulate, analyze, and interpret data in your courses and classes? You have provided descriptions that they are using R mostly right for this class to run some statistical tests. Could you describe a little bit more about this process, how they work, or give an example based on one of the assignments you have for the classes you teach.

UCSB8: Sure. So we actually use data in two ways. So, in the lectures I actually walked through how to hand calculate statistics and more of the conceptual background for a particular statistic like a T test for example. And I also, when was that, not this past summer but the summer before, created a series of learning glass videos. I’m not sure if you’ve seen learning glass before. It's like a one way or two way. Right so basically I created a series of, I call them stats class videos to show students how to hand calculate statistics. So, for example, in one week we might work with some conceptual ideas of what’s a T test, and then I'll show them how to hand calculate a T test and these stats class videos and then then the lab section that week, they would be given a data set like we just talked about, and they would learn how to use R. The TAs would walk them through the code that's used in R in order to run a T test with the dataset that they have. And in both cases we focus a lot on interpreting the outputs of R or the outputs of our hand calculations to determine whether or not results are statistically significant using null hypothesis statistical testing so we focus a lot on that in the class. And then there's a big emphasis like I said on just writing it up correctly, so I tend to err on the side of I really want students to be able to not just calculate statistics and values but to understand what those values mean on the other side, and how to report those values to someone else because I think that helps them with actually reading empirical papers. So when they practice writing a results section I think they are better equipped to actually read the results section of the paper.

I: Right. And so you mentioned R, are there other tools they use during your classes or would you say it’s mostly R?

UCSB8: We use Excel sometimes to help with graphs and developing charts and graphs and line graphs and so forth, but mostly we use R. Previous to coming to UCSB, at my previous university I taught SPSS to students but since being here, the focus has been on using R.

I: Okay. Yeah. And so you mentioned they have a previous statistics class, so that they learn how to hand calculate and also go through some of the tests like T tests and ANOVA in R. For this case, would you say they have any prior knowledge. Is there any prior knowledge you'd say, of tools or software do you expect our students to have before taking your courses or. You mentioned TAs are teaching and walking them through R, but then or, like, do they learn all the skills they need for this course throughout the course?

UCSB8: I think when it comes to R they learn everything in the lab section. So, I think they learn everything in this course, my understanding is that they haven't had that exposure in a previous stats class. In fact, I’ve been quite surprised, for example even working with Excel some students have a difficult time and don't seem to have much experience in working with Excel or uploading something to R for example. I mean it requires a little bit of coding knowledge but the process is kind of the similar, you know, similar to uploading a photo to to an email, but I do find that when it comes to statistics for whatever reason students don't think of the process as the same sometimes and so yeah I do feel like everything we teach them is pretty much brand new when it comes to actually using R to analyze their data, I mean the hand calculations and the conceptual ideas, some of that they've had before. I've never seen the like syllabus from the previous stats class that the students have to take before they get into my class. I've tried to access it a couple times. I think I saw an abbreviated version of one once. It's my understanding in the first class that they focus on some statistics in depth that we don't use as often in psychology so they might focus a lot on Chi squares and we don't use that a ton and so I try to focus on the analyses that we use a lot. So, T tests, ANOVAs and regression and correlation for example, we might dabble in chi square at the end of the quarter if we have time but we don't spend a lot of time on it because in psychology, we use a lot of the other analyses. But I do actually feel like even with the conceptual and hand calculations they're learning or relearning a lot of it in my class.

I: Okay. And do you face any challenges relating to students' abilities to work with data in your classes because you mentioned that you want them to go beyond just calculating but like be more critical about statistics. What would you say could be challenges for your students to work with data.

UCSB8: I think, well one of them I think is for students to move beyond thinking that data and statistics is just math. So, I think that's a hurdle that I try to overcome quite often to get a sense of it's more than math and that's out there in everything we see. So, there is a statistic on everything, perhaps, especially right now. So, there's lots of statistics out there and then it's up to us to understand and interpret them correctly, not necessarily up to the person presenting them sometimes. So, I do think that that's a big struggle is to move beyond statistics equals math to statistics is more than that. That data is more than that, that data can help us better understand things around us in terms of the things that we see in our everyday life. So, you know if they know how to read a table better from taking the stats class then we'll be able to use that in all sorts of situations. So, I think that's definitely one challenge. I do think that there's a general anxiety associated with data among students. So, like I mentioned before, I feel like I was like, upload a, you know, a photo onto an email, they know what I mean when I say upload. But when I say it in the context of “upload your data set into R” it feels very novel and foreign and there's a sense of anxiety associated with it sometimes. So, so yeah I think that's part of it as well as overcoming that kind of anxiety that students might have about interacting with data, and that when they push that little Run button, the whole thing's not going to melt in front of them.

I: Right. So in terms of your field, psychology, how do the ways in which you teach with data relate to the learning goals in your field? What will be your opinion on how it connects to the approaches that are taken in psychology?

UCSB8: Can you repeat the question?

I: Yeah, so how do the ways in which you teach with data relate to the learning goals in your field?

UCSB8: Well I think, I mean thinking critically as a big one in psychology and I think psychology is, I mean it's science and so I think any science field has data associated with it. That might be a hurdle that we have to overcome in psychology where it's thought, Psych and Brain Sciences particularly, that it's thought that perhaps psychology is just interacting with patients. You know, so a lot of students think it's about being a clinician or being a counselor, and overcoming that idea and helping students recognize psychology as a science, dealing with data, multiple types of data different types of data from humans and animal subjects and so forth, is important, and aligned with the goals of Psych and Brain Sciences.

I: Right. So, you mentioned about teaching students to think critically about the sources and the uses of data they may encounter like in papers. Could you describe, like, a situation or if you feel you have an anecdote or like a situation, actually, you had an assignment where students had to go and think critically about the data source they were using. Could you talk a little bit more about this. Like how do your students think critically about the data source they're using?

UCSB8: I think the critical thinking part in my class comes with the interpretation so everybody, if everybody's running the same analysis, for example, everybody's running a T test for the same data set everybody gets the same output. But that doesn't necessarily mean that everybody's going to write it up in the same way. So I think in this course in particular critical thinking comes with considering what the values that they get on their output mean in the context of the example study that they have. And everybody writes differently and I think that's the nice thing about having this writing component in the stats class is that you get more than just calculations and values you have students interpreting the data. Yeah, I think that comes with thinking critically. I think just in general, in order to conceptually understand statistics there has to be a component of critical thinking and I try to walk students through that process in my lectures in particular, so the lab sections focus a lot on R, and in my lectures I focus a lot on the conceptual understanding of concepts. And so sometimes I'll have, I don't know, three or four lectures where we don’t even calculate anything. We just talked about some of the concepts of statistics.

I: Right. So, if you could list out a few specific data skills that you believe you prepare your students, in terms of future careers. What would be those?

UCSB8: Like specifically for future careers?

I: Yeah.

UCSB8: So yeah, so, there's a couple skills, one is being able to see data represented visually so graphs, tables, charts, and being able to interpret that data is important. I also try to teach students that the onus is on them if they're the researcher to present the data accurately. So we talk a lot about that in the quarter how you can, you know, change the y axis on a graph and make an effect look bigger than it actually is for example and that the onus is on the researcher, the data collector to present that accurately. And in doing so I hope that means that they question the data that they see out there in the real world. Yeah. So I think those are some skills that I think would go forward. I mean, coding is a big thing now for I think a lot of professions. But I'm not trying to get my students to leave my class experts in R, they get it for nine weeks, and, you know, they'll probably look up whatever code they need to use again in the future. I don't make them memorize the coding for the R for example in this class. But I do think at least they get some exposure to coding which is important for their future.

I: Alright, the next questions about your perception in terms of cultural changes at UCSB or even outside UCSB in terms of the ways in which instructors that are teaching with data. So have you observed any policies or cultural changes at UCSB or outside that are influencing the way in which you teach with data?

UCSB8: I don't know if they're changing so much the way, I mean previously like I mentioned I was using made up data sets about kind of silly psych related examples, perhaps. But this quarter moving to using real psychological data sets, I did intentionally for example choose a data set that looked at cultural differences in the concept that will start working with the week after next in the class so we will have a data set that specifically focuses on cultural differences and to explore that in a data analysis. But, yeah, so I think that's part of it.

I: Yeah. Okay. Thank you. The next question is about your perception in terms of ethical challenges and teaching with data. So do you see that instructors in your field face any ethical challenges in teaching with data? And, if so, to what extent are these challenges important to your teaching?

UCSB8: I mean, speaking for myself, because I don't have students go out and collect data, it's very different so they're not interacting with participants, they're not asking each other questions. They're working with previously collected data that has all been de identified where there's no identifying information about who the participants were. Everything was confidential. So in the statistics class that I teach we don't have to deal with that too much. However, sometimes I teach [course number], once a year usually, which is a research methods class.

And in that class we talk about ethical principles in psychology and how to follow those principles when conducting research and collecting data, but not necessarily in the stats class that I'm teaching this quarter.

I: OK, so now the questions are going to move to training and support. So you mentioned that you have help from some TAs, correct, in this class, that are particularly involved in teaching R coding to your students. But in your course, would you say that you have any other type of support, in terms of helping to provide instruction and support for your students to, you know, obtaining or even working with data, rather than the students you already mentioned. So for instance if you have help from librarians or you have drop in sessions, like from different divisions or like different departments, coming in and kind of doing demos or getting involved in your instruction.

UCSB8: Not so much in this class because it really is an introduction to statistics. So, it is kind of a baseline level class and because I have so many TAs. There's quite a bit of knowledge among the graduate students from Psych and Brain Sciences about using R. I mean I have relied on tech support every once in a while, particularly you know during the pandemic where not all students are able to download R onto their computers or I think it doesn't work on, doesn't work on a particular type of iPad, I can't remember what, but we have access to R cloud. So the students have been able to use R cloud in the class, I think there's about five students out of 125 this quarter that are using it. But yeah, I don't have any drop in lectures. And that's probably it. They use Gauchospace a lot to communicate with students in terms of providing the data, providing lab slides for writing lecture slides, send you rely on that quite a bit, providing the, I mean the stats class videos I have. I recorded those over in the ID center so I had a lot of support with getting those done.

I: Great. So, to your knowledge, are there any ways in which your students are learning to work with data outside their formal coursework. So, for example online tutorials, internships, peers. Are you knowledgeable about these alternative ways they're learning how to work with data?

UCSB8: I think there's well in the Psych and Brain Sciences Department students can be research assistants, undergrad students can be research assistants and a faculty members lab. And that seems to me to be the next place where they get some additional experience outside of a formal class setting, I mean they are registered for it through a credit, but not in a formal way where they're receiving lectures and so forth. But yeah, I think most students probably have some experience and exposure, who are research assistants to data in that way, whether it's through data collection or through analyzing data and working with coding data.

I: And for your classes, do you have any type of, or do you expect or even encourage this kind of extra curricular learning. So, do you have any, like, additional resources that encourage your students to go forward and like to learn more about the, not only the statistical approaches or analysis but also like the tools that are available.

UCSB8: No, haven't used that much in my class. I mean, I feel like the classes so packed with work for the students that I'm not sure even if I recommended something like that, particularly now I think the students are totally overwhelmed with taking online classes so I mean I feel like even if I did recommend it, I'm not sure the extent to which a lot of the students would use it.

There'd probably be a handful that might, but I wouldn't want to put it out there and make students feel required to use it at this point so, anyway. I try to make things a little easier, not easier for them but less, less intense for them at the moment.

I: Yeah. Makes sense. Yeah. So I'd like to hear a little bit about your own training. So have you received any training teaching with data other than your graduate degree or you mentioned you had a past job in another institution. But let's say like, workshops, technical support or help from peers, how did you get your experience in teaching with data, how did you learn about like what would be good approaches to teach with data?

UCSB8: I think it probably started in graduate school at some really excellent statistics teachers in graduate school and when I finished my PhD. I can't remember what it was, I had taken so many stats classes that I had some sort of certificate in social science methodology and statistics. And then, I usually do when I attend conferences, if there's some novel statistical concept that I'm interested in but not familiar with, I often attend a workshop, I think, pretty much every conference I go to I try to do that, to attend a workshop at a conference about some novel statistics. I've attended some talks at the Q, the quantitative QMMS series. So, I've given a talk there and then I've also attended a few talks so on didactic analyses which I wasn't so familiar with. And then prior to coming at UCSB I was teaching statistics and methods at undergrad and graduate level for 10 years in the Caribbean, so my previous university, University of the West Indies, and Trinidad and Tobago, and there that's all I taught for 10 years was statistics and methods at both undergrad and graduate level.

I: So you mentioned that you upload these data sets, mostly to Gauchospace, that's like your environment actually students get access to these data sets. Do you use any data sets, some data sets you mentioned already, but like assignment plans, syllabi, or other instructional resources that you have received from others or from your peers? And how do you make available, or do you make them available, if so how do you make available your own resources to others. And when I mention others it’s like outside your classroom so not only your students, but also your peers.

UCSB8: Yeah, so maybe, maybe not so much, I’m trying to think. So when I first got to UCSB, when I first came to UCSB I did borrow resources from faculty in my department who had previously taught the course just to get a sense of the structure of the course kind of what students might be expecting what the department was sort of expecting with how the course was formatted and run. So, I did work with those previous lectures and assignment ideas and tests and so forth as a starting block. And then I've also shared my resources with graduate students who have taught the class in our departments, they usually teach in the summer. And then, outside of the department I'm trying to think if I've shared materials from this class. I did, I can’t remember if it was this class or different class but I think it was this class, yeah it was this class. So, the APA American Psychological Association has a syllabus resource center for faculty teaching methods and statistics and so I did share my syllabus there as part of the, I guess syllabus bank that they're developing. Yeah, that's kind of the extent of it yeah.

I: Yeah great APA.

UCSB8: Oh, and my stats class videos I offered those all to my colleagues to use.

I: Right, so there's like a lot of sharing in your community.

UCSB8: Yeah, in our department there seems to be.

I: In terms of, like, considering the evolving trends in your field, what types of training or assistance will be most beneficial to instructors in teaching with data? So, in your opinion what would be like beneficial for instructors to learn, to get prepared in teaching data.

UCSB8: Um, well I think, I don't know how other fields are but in Psych Brain Sciences, I know that R is the fastest growing software currently, and I wasn't trained in R, so I was trained a long time ago and I was trained in SAS and then I was trained in SPSS. And so actually I had never used R until I came to UCSB to teach it, which was pretty interesting. So I was kind of learning R just before I was teaching it to students and setting up lab sections to teach it. So I think that would have been a bit useful for me looking back to have a, you know, as a new faculty member to not feel kind of self conscious about not knowing the newest statistical software out there and have some resources to learn that would have been helpful. I mean, I feel like I have the hang of it now but it would have been nice early on to get that sort of resource. I think that’s the only thing.

I: Yeah. Do you see any other skills that will be important for instructors to learn or focus on before teaching with data.

UCSB8: I mean, I think maybe some, I mean I've been teaching it so long that I feel like it comes kind of second nature. But I remember initially when I first started teaching statistics I didn't realize how tentative students were about working with data and learning statistical concepts, that kind of, I think that sort of anxiety that students had and the fear that they had of learning a new statistical software program and learning about data and what data can tell you, I was kind of shocked by that sort of anxiety that they had. And so it might have been nice to kind of talk about that with colleagues ahead of time and \ help manage that with students.

I: Because it feels like it's all at once right they’re new

UCSB8: And they're all so anxious. Yeah.

I: Alright, so my last question to wrap up, I would like to ask if there's anything else from your experiences or perspectives as an instructor or on the topic of teaching with data more broadly that he would like to share.

UCSB8: I don't think so. I mean, I think, in our field, Psych and Brain Sciences, I think data is used in almost every class that's taught. So, I do think that our students get a lot of exposure to it, which I think is great. And like I say, for me, not all psych majors are going to go on to do something in psychology so it's more important for me, that they can appreciate other people's presentations of data and kind of weed out when it seems legitimate versus maybe when it doesn't. So, for me, that seems like a really important skill for students to leave with having that sense of feeling comfortable with seeing data, seeing graphs, hearing somebody say a percentage of people did x y and z and knowing what that means. It feels important.

I: Great, so I’m going to stop recording.

Interview 9

I: Okay. All right, so before my first question. I would like to acknowledge that due to covid there's a disruption in teaching. So, for any of the questions I'm about to ask, please feel free to answer with reference with your normal teaching practice or your teaching practices as adapted for the crisis situation, or even considering both. So, my first question is a background question that I would like to understand a little bit about your experience teaching undergrad students. So, could you briefly describe your experience teaching undergrad students including how does your teaching relate to your current or even your past research?

UCSB9: Okay. So, my background is in development economics. My PhD is in agricultural and applied Economics. And after my PhD I went and did three years of work at the Asian Development Bank. So, my interests are in development training as a quantitative social science and economics more generally, and the interests are already in public policy. You know, economic policy in general. Now, I teach in the Global Studies department so it's a little bit unusual because that isn't quite the mix that is covered in global studies. That is to say, there are people who do think about public policy, there are people who do think about development. But there's actually no other economists in the department and there's no other quantitative social scientists in the department at all. So as a result, when it comes to the undergraduates, you know it's a, you know, it's something that I can do. I can expose them to quantitative stuff for one quarter but never really any more than that. So, I teach this one class every year, it's called [name of the course]. It's taught to juniors and seniors, I have typically about 130 students in the class at a time. And I make it a point in that class to make them at least responsible consumers of quantitative data. That's that's the goal, right. To learn, to actually teach them anything about how to do quantitative analysis in any serious way is very difficult, because they're not seeing any of it in any other classes, many of them are taking classes where they're actually, you know, actively discouraged from dealing with quantitative social science. And you know they just don't have any training, many of them are math phobic. They have chosen global studies because they are either afraid of mathematics or somehow, you know, have some issue with it. So, that's kind of what I'm dealing with and so you know if you, the approach I take in class is to start out with, you know, big questions that they would want to know the answers to. And then try to start out by showing them some data showing them some correlations showing them, show them, you know, moving pictures of correlations over time.

The business website gap minder.com. That's extremely useful for this; it animates scatter plots over time. And so you can do a lot with that. Show them, you know, countries' infant mortality rates, GDP, etc . Watch these things moving. And then from there, launch into a critical discussion of, you know, what are the, accurate statements you can make based on these data, then which statements misinterpreting correlation as causation. Why, how do you fix it. Try to get them to read one or two papers that try to do it. And they can't actually read the whole thing, they can't make any sense of the guts of most of this work but they can, as you guide them through the arguments and say you know this is what this approach is trying to do you're comparing, you know, the rate of decrease in infant mortality in the same country over time during periods of high growth and low growth so how does that influence how you know, what does that relationship tell you is that close enough to causal for you to conclude that economic growth reduces infant mortality things like this, right. So, it's really taking data, showing them what the data are, animating it, trying to make statements, showing the logical pitfalls that you can make when interpreting the data, and then showing them that there are, in principle, ways to get around all this. But that's as far as I can pick any student in my department.

I: Okay, so I'm just to reiterate please correct me if I misunderstood, you teach [class name] and it's only focused on undergrads or you do also accept graduate students in this class?

UCSB9: We also have graduate students. So we have, we have had a master's program for a long time. And by the way, I should say I've been doing this since 2007. So we've had a master's program since 2006, and a PhD program since I think 2015. But it doesn't really make much of a difference at the, in terms of the curriculum they have in our department, our department is indifferent to quantitative research, and some faculty are actually hostile to it. So the students get no training in it. And so every once in a while we get a master student or a PhD student who has had some quantitative training as an undergraduate, maybe they majored in political science or economics. And then they are here. And then I sometimes end up having to work very closely with them to help to supervise a project to basically train them from the ground up, but it ends up being a solo act and something I'm increasingly keen not to do, it just takes too much time. So, you know, that's why that is but yes PhD students you know that's the place where you can find one or two that come in with some training.

I: Okay. Okay, thank you. In terms of getting access to data, for, for this particular class you teach, how do your students usually obtain the data they engage with? Let's say you provide the data, do you actually ask them to find data sets connected to the topic? How do you usually provide the data?

UCSB9: Oh, yeah, so what I do, honestly is I get the data myself from whatever source, whether it's the world development indicators or some research project that I'm working on, you know, because I'm always working on some data so I get the data myself. And then I organize it into an Excel file and I have to write very long, sort of instructional point by point instructions for them on how to then take it from there. If they had to download the data and organize it into Excel, nothing would get done. Because I, you know, bear in mind I also often have TAs who have very little quantitative training. And the other point I’ll just say quickly is you know when we do these exercises we have to give them stuff that is you know, the food has to be half chewed before it is given to them. Right. Then they'll finish it. And to finish it, they can only do it if we actually go to computer labs, book computer labs for discussion sections. I have to write detailed scripts for my TAs. And they then run it in sections, and then they also have office hours and computer labs and that's how that's how the students get through it.

I: So, you have mentioned some data sources that you usually explore to find data and that you do some work later on. How do you usually find data sets that are connected to your teaching? What are the main data sources you go for?

UCSB9: So there's the world development indicators. That's the World Bank right. UNDP has data on a lot of things but it tends not to be particularly well organized. So for teaching I don't take the time to use UNDP very often. And every once in a while I'll take a data set distilled from one that I'm working on. It's obviously got to be something that's in the public domain so if I've got survey data on something. Or, if there is, you know, a project I've been working on where I have accumulated data from many other sites. And then processed it into a form that an undergraduate would be able to work with without any prior training then I use that. So, one other site is the Groningen Growth and development center. Trying to think of what else I have used. I think those are the main ones, there's only so much you can actually do with untrained students, right. So, it's and, you know, the big questions in [class code], we have data from, from the world development indicators, or there's one other thing I do, which is in a different class and, which is a class on macro economics. I have them compile, I start with data sets produced by the Asian Development Bank which are just country sheets that sort of give you a country's balance of payments, the macro economic indicators, the macro economic structure, what the currencies have been doing etc. So I give them a bunch of that information. Just links to excel sheets that have this stuff put together. And then I have them produce PowerPoint presentations where they develop analyses of these countries, and support it with graphs that they make an Excel from these sheets. And for and for that I had to, you know, there was a parent of one of my former students who gave, you know, made a $2,000 donation to the department so I've been using that every year to hire lab assistants for a few hours to help the students with this.

I: Oh, nice, that was a generous act from the former student. So, you mentioned you usually use Excel sheets to distribute or to make these datasets. So, do you share like a Google Drive with them, with the class and then they have access to the data set.

UCSB9: I don't use Google.

I: Okay.

UCSB9: I try to avoid doing anything related to students on Google. I have very serious misgivings about the university using Google for everything and particularly for student records. Because I have every reason to suspect that Google is one day going to become an employment agency, and going to be selling information about students. I've gone through their privacy policies; they don't protect our students at all, so they might be using someone's performance in our classes to eventually tell employers whether they should hire this person or not. And so I won't use Google. So, I just use Excel on Gauchospace, I make Excel files and I put them on Gauchospace and that's it.

I: Okay. In terms of like, you mentioned different data sources, number of agencies that you find data, and you use data for your teaching. Would you say or have you experienced any challenges in obtaining data sets, even from the agencies you listed or I recall you mentioned that most of the datasets you use are in the public domain so I believe that not a recurrent thing you experience very often, but have you ever experienced any issues to find data?

UCSB9: So, so, yes. So, there are very serious data issues that arise when actually trying to do research. So, access issues, right, for teaching I don't need the particularly specialized data for teaching because what I'm trying to teach the students are some very basic principles about some very broad topics. But when teaching PhD students, that's basically research. And yes we do run into problems. So, for example, I have a student who's trying to do work on women farmers in India, and the Indian government websites are absolutely horribly run. The Indian government collects a massive amount of census information. And then, doesn't update it, puts it out on websites where it's incomplete. They never have working email addresses if you want to contact them to try and access the micro files, that kind of stuff is a real problem. And there are some universities that have access to these data. And so currently we're trying to get some through Berkeley. But I do know that you know if you're at Harvard and places like this, you know, and you want the census files you can get them.

I: Alright, so some of my next questions are going to be related to working with data. So, I would like to understand how do your students manipulate, analyze, or even do the interpretation of the data for your classes, for this particular class. So, what tools or software do your students use? You mentioned Excel, but are there other tools that are used for the statistical analysis or visualization, or any other tools manipulating or working with the data?

UCSB9: No, they use only Excel. You know my philosophy of teaching students to work with data is that, you know, the actual data manipulation is the easy part. And you know that you can sort of pick up in all kinds of ways with online help things and whatnot. What's really difficult to do, is to interpret what you get after you've manipulated it. So I teach them to manipulate data very minimally in very simple ways and think through what they can learn from it. And all that you can just do in Excel, by using tables and simple functions like minimum, maximum, drawing scatter plots, putting lines of best fit through them. And then you know just just thinking through that. Of course, you know if my students were doing anything resembling real data analysis, then, then I would use other things I would use STATA, for example, I have run into a problem with masters students and graduate students where they get they sometimes, I've had maybe two or three of them in the last 13 years who got to the point where they want to analyze the data set themselves, and they don't have access to STATA. And so, you know, I've had to buy licenses and things like that for them. So that's one thing that the university could really think about is, is licenses for software that students use, I know that increasingly people are switching to R which is, you know, free. I just haven't made the investment in learning to use R so maybe that's one way I can get around this. But certainly you know one thing I will emphasize, I've had people say you know why do we worry with Microsoft Excel etc. We should just have everybody using Google. And I don't think that is a good idea, both for these privacy reasons but also because I've never found Google Sheets to be completely satisfactory.

I: Yeah.

UCSB9: Maybe I need to learn, learn more about it but I find the operations for analysis a little bit rudimentary.

I: Yeah. So do you expect your students to have any prior knowledge of tools or software for these classes in particular that you teach for undergrads. Like do you anticipate them to be at least familiar with Excel prior to starting to class.

UCSB9: No, no I can't. They just don't have that familiarity. In my department we have people who approach the world through literature, through languages so there just isn’t that experience. The only thing I will not teach them is arithmetic. So, you know, I will say to them at the start of the class, you know, you have to be able to find the areas of rectangles and triangles and tell me about the equation for a straight line. And do you know your basic fractions and percentages, and actually get a lot of students who say that's not fair. Why should I have to do this, etc. You know that's what we're working with. So the problem, you know we don't really get into problems with, you know, software anything, I gotta teach them from the beginning.

I: So previously you mentioned that. It's not that challenging to get students manipulating the data but more like to make sense of the data, after and to do some analysis. Could you describe any challenges you have experienced relating to students’ ability to work with data for this class?

UCSB9: Well, yeah so I wouldn't say that it's very easy to get them to manipulate the data. I'm just saying relative to... You see what I was reacting to, you know students will often come to me and say, “I want to learn to use STATA.I want to learn to use R”. And my answer, my response to them is, “What do you want to do with it?”. That's like coming to me and saying “I want to learn to use an X-ray machine. But why do you want to learn? What is the question you want to answer? What will you do with these tools?”. And, you know, if they start out with oh I just want to learn to program this, that or the other, they will never produce anything of value with it. If they start out with, Here's a question, you know, here are some comparisons I want to make here are some sort of logical conclusions I want to be able to draw from a data analysis, we can design the entire analysis on a blackboard, without ever actually touching any software. And then once they have that, then we can throw it on the software and, you know, make the analysis happen. But otherwise, they're dealing with two challenges at the same time, the challenge of, what am I actually trying to accomplish here with the software and then how do I use the software. And that, so that's where the problem is. So what I really try to focus on first is showing them how to interpret data. So, the first four weeks of term we're looking at things on the board, where I have analyzed the data with the software, I show them the charts and we talk about what can you learn from this, what are the alternative interpretations etc. Then week 5, 6, 7 I'll get to your question now, in week 5, 6, 7 we go down to the computer lab, and we'll start actually manipulating the data in Excel. And the way to do that, I find because I have untrained TAs often, and I have totally untrained students is that I have to write step by step manuals on, you know, first you do this, then you can right click on that, then you highlight this column. Right. Then you copy and paste and you drag and drop it, we have to show them everything from the, from the start to the finish.

I: Describing the entire rationale behind it, you know, the statistical analysis first and then moving to work with the tool and the software, right?

UCSB9: Exactly. Exactly. Exactly. And to be honest, I have not tried anything other than spoon feeding them on how to use the software right. I know that there's things on lynda.com etc. where I could say you go learn Excel and I want you to produce this. And that could work, but it could also backfire and I'm often dealing with students who are quite upset about having to touch numbers and having to work with numbers in the first place, that if I tried something that pedagogically they're not capable of doing then I would really face a backlash. So I can't ask them to learn these tools on their own. I haven't yet had the guts to do that. Maybe I should. So okay, as far as difficulties with software I'll tell you what one of the big difficulties has been in the past. It's that the computer labs are often difficult to access; there are not enough computer labs and during section time when I try to get a reservation sometimes I've not been able to get one. So when I've not had a reservation we have actually tried and said to students, “Look come to office hours, bring your laptops, etc. You know, and we'll work it out”. Then you run into the problem that with, particularly those who use Macintosh is they're all using slightly different versions of Excel that have different bugs. And, you know, trying to get a TA who's not comfortable working with data to assist students who are working with five different versions of the software is a real nightmare. So, the truly important thing I would say for, you know, if the campus wants to help us teach students to work with data, we need more of these labs, and we need a reservation system that is a little bit friendlier that we can sort of access far ahead of time, book far ahead of time so that those of us who are actually running classes get first dibs. You know what happens is you'll often find that the labs have been booked up for open hours, but you know open hours should be scheduled after after sections.

I: Right.

UCSB9: You can have open hours at any time, you can't have sections at any time.

I: It's a good point. And indeed, like whenever we teach people with different systems like PCs and Macs, it's a nightmare. So, I can relate to that. Yeah, different versions of the same software as well. So, it makes it a little confusing. Yeah. So, anything else you'd like to add in terms of difficulties before we move to the next question?

UCSB9: So yeah. I think that's something that, you know, we really should address as a campus in some ways, which is this, this two cultures problem.

I: Right.

UCSB9: So, I don't know if you're familiar with C.P. Snow and his two cultures essay, it’s from the 1960s, but it's basically this idea that you know you've got the, the quantitative technical social scientists who imagine themselves a scientist and you’ve got the humanists. And we have a pretty difficult time communicating with each other. There is what I've said sometimes it's actually open hostility across this divide. And there is, in the old days that wasn't a problem because students could double major, they could go to classes outside their major etc. And so students were not as badly affected by this as much as faculty and PhD students but undergrads could get a well-rounded education at UCSB, play on both sides of the cultural divide and they'd be fine. Today, that's no longer true because all these majors are impacted and students can’t get into classes outside their major. And as a result, we're churning out students who, if they come from, sort of the more cultural studies side of things or from the humanities, are never exposed to quantitative social science at all. Even if they're coming out of the qualitative social sciences, they get no exposure at all to quantitative social science. And you know on the flipside too, right. You know the economics majors and the engineers don't know anything about literature, etc. And, you know, that actually has a lot to do with how you should interpret quantitative data too, you need a strong background and philosophy before you should ever touch a quantitative data set. And so I think that's, that's a major problem. And if the university is going to get serious about liberal arts education which means bridging the two cultures, there actually need to be, you know, core classes for freshmen where they’re actually required to do a bit of everything. And actually address this stuff head on, what I have seen frankly, you know, I know they've got these quantitative requirements to check boxes. But I have seen, I've encountered so many students who have checked the quantitative boxes and can't tell me what 4% of 200 is, much less, you know, tell me anything intelligent about, you know how to interpret a scatter plot, and how not to interpret a scatter plot, you know. So, so I think, I think there is this pedagogical work to be done, that would actually require faculty interest, faculty to come together to think about, sort of, you know, theory of knowledge classes, which include exposure the quantitative social science, hands on work early on, and that in turn would mean investments in computer labs. But these are the real problems we need to fix as a campus, I don't think that there's problems with accessing this data set or that data set, that's not the issue. The real constraint is contact with faculty with the time and the incentives to actually set up you know this kind of education.

I: Okay. Oh thanks for that. So, moving to the next question, I'd like to understand a little bit about the process of, like, teaching students with quantitative data and how it connects to the learning goals in your field. Can you describe how the ways in which you teach with data relate to the learning goals in your discipline at large, or even economics, considering that’s your background?

UCSB9: Sure. So I'll talk about economics, because frankly my department has more or less rejected quantitative social science. There just isn't any offered there if you read the journals, it's not there. If you look at the faculty, not just at UCSB but in general, anywhere in anywhere in the United States or, outside of Eastern Europe, all the Global Studies faculties that I've looked at I've seen are entirely non-quantitative. So, I'll talk about that sort of teaching. I would say Development Studies. Let's call the field development studies sort of halfway between economics and global studies, right. So it's interdisciplinary. And it is both quantitative and qualitative and that's what I think I'm teaching my students. So, how does quantitative work relate to the teaching goals. So, there are several things I want students to understand. One of the first things you have to do in teaching any kind of social science is you have to convince students that they're wrong about everything and they’re probably wrong about everything. So, they have to be confronted with a variety of kinds of evidence. And they have to be confronted with evidence that refutes their preconceived notions and quantitative evidence can be very good for doing this. Right, because it really forces them to sort of sit down and say, Okay, if this is my view of the world. First of all, you know, in how many places this is relevant? You know, forget about causation, do we even see correlations to suggest that my worldview is in fact the correct one, or that, or that these particular social forces operate the way that I see they do etc. And so, the one of the nicest things about it is it, you know using quantitative data in classes forces students to stand back and think about evidence in a way that does not presuppose the answer. I don't want to, you know, sort of suggest that quantitative evidence is better than qualitative evidence, that that isn't the argument. But systematic approaches to research are more likely to dispel confirmation biases than non-systematic ones. And the nice thing about quantitative data is that, you know, the only way to analyze it is systematically, I mean, there are people who will try to manipulate it etc., but if you give up on those dirty habits, then it's, you know, it's a systematic approach to data. So, you agree first on how am I going to analyze the data, how will I interpret what comes out, and only then do you get to see what comes out. And the effect of that is that often what comes out to something that makes you think about the world differently. And that's one of the main goals in teaching this kind of stuff but particularly at the undergraduate level.

I: My next question relates to how do you usually approach, like, what's your approach to make your students to think critically about the data sources they might encounter like in their everyday life. So can you expand a little bit about the approach that you usually follow to teach them when they are confronted with evidence and, you know, to assess and interpret the data?

UCSB9: Yeah. So, so there's, you know there's two things here there's the data, and there's the results of the data analysis. So, when I say the data, I mean, you know, the starting numbers that go into the analysis where do those come from. So, in, in my classes, whenever I introduce a variable, first of all, they have to understand, they’re obliged to understand every aspect of the definition of the units in which it's measured and the real world process by which somebody arrived at that number. So for example, you know, if we're going to look at a GDP number right. The question is, what units is it measured in, what does it capture, who made the number, who invented the number, what are the components of the calculation that were imputed, who imputed them. And to bear in mind, and then I always tell them a few stories about people who, who may have made up numbers or used dubious assumptions for some reason having to do with their political objectives or prior beliefs or whatever else.

I: Right.

UCSB9: So, start out by making sure they understand that all numbers are constructed and emphasize to them that that doesn't mean numbers are useless which is unfortunately what they get from some people, but rather that one must be aware of how they were constructed. Take a sensible decision about, you know, whether you accept the quality of these data or not. And then think about, then we get to the analysis, we think about how you could interpret, you know some relationship seen in the data if you believe the data absolutely fine. And then we also say we're looking at what if they weren’t. What kinds of bias measurement is in the data that would actually lead you to incorrect conclusions when you interpret the analysis from that data. We also take a lot of time to cover examples of misinterpretations.

I: Right.

UCSB: So, we go through lots of examples of reverse causality, lots of examples of omitted variables, biases, lots of examples of just spurious correlations. You know, I find that humorous examples help a lot. So, you know, things that people did something so thunderosly stupid it really sticks in the mind. So that's the approach.

I: Okay, thank you. That was very comprehensive. In terms of going back to institutional support, because you mentioned the issue of not having labs accessible or being able to schedule for lab timing. In the past year, have you observed any policies or cultural changes at UCSB that would influence the ways that you're teaching with data?

UCSB9: There was one set of changes with the scheduling of the computer labs. So, I don't know exactly how this happened, and this is a story from about two years ago because I was on sabbatical last year. And then this year I haven't been, this year I've taught one online class that didn't involve the labs at all. I don't know what I'm going to do for data teaching in the spring.

But I'm forgetting the names of the scheduling software but we used to have scheduling software, where we would put in a request, the request would immediately get elevated to somebody, I think at collaborate, and then, you know, they would sort it out and they write back and say okay you have labs at these times, etc. We went from there to scheduling the computer labs using the same system as being used to schedule the classrooms. It was difficult to get somebody on the other side. And I think also because there has been increasing demand for quantitative subjects and classes among students. We were running into constraints that the labs simply weren't available, and then there was nobody to talk to about it. And you know the response to it was largely well you know go figure it out. You know, what compromised positions can you come to, are there ways you can teach the stuff in office hours, etc. and that was a real problem. I've also, you know, I do sit on a campus technology committee that oversees Collaborate, etc. And I did hear some fairly disturbing things where people kept saying, maybe we don't need the labs everyone can just use Google stuff. And, you know, they can sort this out and they can learn online and they've all got laptops so we don't need to support all this.

And I think what that is really missing is the centrality of a teaching space of physical teaching space where students can come in and learn to just kind of jump through the hoops, monkey with the software and get the basic tasks done. You know, for, for, it may not be. Yeah, I mean I just, I just think this is going to be one of these things that if the campus does go in that direction of not supporting centralized learning computer labs, we're not going to be able to teach quantitative pedagogy, to students, particularly those who are quantitatively challenged. And one other related issue that comes up with this you know one of the things that I've actually been doing research on is sort of racial and economic stratification across majors at the UC’s including UCSB. And so we now, we're increasingly seeing since, over the past 15 years, a segmentation where the stem STEM majors are increasingly getting filled up with students who went to better high schools, they are you know less likely to be students of color less likely to be from working class backgrounds. And so this argument that you can, you know, just have students work on this stuff on their own, and we don't need the centralized learning spaces for them is going to make that kind of thing worse.

I: Right.

UCSB9: There's been, ever since the financial crisis, there was a migration of students from the qualitative social sciences and humanities into the quantitative Social Sciences and the STEM fields. And that has put tremendous pressure, I think, on the lab spaces. And, you know, the STEM departments are using these pre majors to decide who gets to take the majors and they are increasingly squeezing out students of color and lower income students. So there needs to be proper infrastructure to support this, to support students, and you know I've actually heard departments, saying oh we need our own dedicated lab so we can, these are departments that are pre majors who say, we need a dedicated lab computer lab for ourselves, so that we can sort of coach students to help them over these barriers. And I think, you know, whether that's a genuine attempt to help or not, I don't know that's a different question. But the fact that they're asking for this tells you that instructional spaces where people can get that, where students can be trained to get their hands dirty with data, or in very short supply on campus. So that's that's the constraint. I think that we need, we need to hit, we need to hit it hard.

I: Right, thank you. In terms of ethical challenges, do you think that instructors in your field have any ethical challenges in teaching with data? So, let's say, you know, sometimes the nature of the data that you're working with, or the topic? Have you ever experienced any challenges, ethical challenges in teaching?

UCSB9: No, because you know, when I teach a data I'm usually teaching about, you know, we're making comparisons between countries, and we're thinking critically about the data, and there's no, so at the country level there's no questions of privacy invasions or anything like that, these are national statistics. You know, yeah I mean I think if there's an ethical problem about teaching with data, it's that we don't do it enough. And the departments that teach our working-class students do it less than others And, you know, there's a big premium on this, right, you graduate from a four year college, if you have learned to work with quantitative data, it actually improves your future employment prospects. And we are systematically depriving, you know, the students that need upward mobility the most, access to that, that's our ethical problem. So we have to think systematically about, about the rationing of students into majors and think about what, and also about the representation of quantitative social science, particularly the interdisciplines on campus. Because unless we get that sorted we're depriving this, that's where these students are and we're depriving them of methodological training of this type.

I: Moving on to the training and support section of our interview. Earlier, you mentioned that there are a few PhD students that sometimes would work with you or assist your students in the class. Do you get any other support, let's say, other instructors, teaching assistants or drop-in sessions that involve other people to help with your students' questions? Because I remember you mentioned that being a solo task, but if you have some help, could you please describe a little bit more about the help you get?

UCSB9: So honestly. You know I don't I don't actually... It's sort of like this right, if you have a student who has done some undergraduate economics and has, you know, has operated some software and run some regressions, they now need to learn to do this better. This isn’t something you can learn through course instruction, you actually have to learn by dropping into a lab and so on. You know, there's one exception I have a PhD student who has been getting a lot of very nice help from somebody who's in your GIS lab in the library and that has been incredibly helpful, that kind of drop in thing. Because you know neither my student or I use GIS so for those startup costs, that's really helpful. So again, there's sort of two issues, there's training students in how to design research with data. And there's how to actually operate the software. The vast majority of the training time goes into how do you teach them to think about asking questions of data right that research design part? My difficulty is that, you know, to teach quantitative research design, you actually need like two years in the classroom and in seminars watching it done. But because no one else in my department will do it, you know, my graduate students don't have time in their schedules to get that kind of training. So that kind of training is what I'm having to do hands on on my own. And so that's just really hard, I mean that's that's the bit that, like, I don't, I don't see how the campus can fill in. You know, that's just a result of this two cultures problems that, that, you know, if you're in a department for your colleagues don't believe in quantitative social science or don't believe in the value of, you know, hiring a couple of other people to help teach it, and build it into the curriculum. You either do it on your own or you don't advise PhD students. That's, those are the choices.

I: And what about like external sources or learning tools or resources, to your knowledge, are there any ways in which your students are learning to work with data outside the formal coursework. I know it's only one class, and one new, one new course, but have you heard of other resources that they are trying to explore while they go through your course?

UCSB9: So, I do run into undergrads who are actually trying these sorts of things.

I: Right.

UCSB9: The best resource is actually a double major in a quantitative social science or in a STEM field. Because, look really this is the sort of like learning to work with data, it's something that you have. It's a craft. And like any craft you have to learn through apprenticeship, and you have to learn through sort of graduated sequences of classes. There is no substitute for that.

You can have as many little drop-in help centers as you want. But, you know, and that but that will help you to pick up the pieces, later on. So for example, you know now that I know how to do, you know I know how to do a whole lot of different kinds of quantitative social science, but I don't know GIS, I could drop into GIS and I know what to ask. And we can gradually figure things out. Same thing with graduate students, if they've had, if they've been through a sequence, if there's no sequence you can’t solve it. Undergraduates, you know, there’s really good courses in the statistics department where you actually get to, you know, turn the data and use it to ask questions so you're doing both design and programming in R etc. So those students that can access that are getting tremendous benefit from it. But those classes are severely overbooked. I have had some students say oh I've gone on lynda.com and learned to use Excel and it was kind of useful. But again, the trick is, you've got to have a project that has a set of tasks that you have to complete. Right, so that you're actually learning to apply the software to solve an actual problem. You can't learn to use statistical software to solve problems without solving problems. And therefore, you can't learn to use it at all because what's the point of using it if you're not solving a problem so there has to be, this is why I'm saying pedagogically you want to fix this, you got to have courses that students take, sequences of courses that they take in which they encounter problems and then use software to solve those problems. The other stuff, it exists, they use it some, but it's not real education. Right, it's it's plugging in a few weeks here and there but it's not not getting the job done. YouTube is a great source if you've got, if you've got no problem with, you know, I need to get Excel to do this then you can find a video that will be a YouTube video that will help you to do that. But that isn’t quantitative analysis, that is just making some manipulation in a piece of software. And it's not what's going to get our students, you know, elevated into, into positions where you know they can make a living doing this, or where it is even part of their profile. You know on the job. So I do think we've got to get serious about it in a systematic way. Sorry, I know you're looking for particular kinds of answers and I'm sitting over here answering my own question, and it’s irritating when people who are interviewed do it.

I: I think it's great because you're actually suggesting a few things and, yeah, so it's going even beyond, so it's really great because we really want to hear and know not only the current stage or things that instructors are struggling with, but also to understand and like, you know, see some possible solutions. So, it's really nice to hear your insights about these topics. The next and final section of the interview, and I want to check if we can go a little bit over time, if that's fine with you...

UCSB9: It’s fine.

I: So this question is about your own training. So, have you received training in teaching with data other than your graduate degree, for example workshops, technical support, or even like close help from peers? Can you explain a little bit about your own training with statistical analysis, and how you went about learning and advancing your knowledge in this field?

UCSB9: So, so I learned to do statistical analysis in graduate school. And I learned, you know, so I learned a lot by, so I went through the sequences, went through a lot of seminars etc. that was always very useful. The two greatest learning experiences, one was writing, working on working on a couple of papers and writing them with my advisor. Because again it's about sort of seeing how you design an analysis to solve a problem, right, to answer a question. So, that was extremely useful and then of course writing my own dissertation where I then had to apply this in my own work, and come up with my own analyses, etc. That was extremely useful. And then you know you, then you, you know once you're beyond your PhD then you start to practice, and you get feedback on your work and those feedback mechanisms are extremely important.

At UCSB, my main support mechanism are the micro economists in the economics department here. There's some really lovely faculty members who are very very sharp analysts, very very good at causal inference. So I go over and get a lot of help there. Well, I wouldn't exactly call it help as much, so on a paper if I hit a roadblock, you know where I know what the right thing to do is, I just can't figure out if it's good enough or how to frame something. I often get help with that. I take my work and present it there and get critique and that of course is extremely useful, because the, you know what, what people in my department think of as good work and what they think of as good work is very different. Right. Again, it's a two cultures question. So that's, that's how I learned to do statistical analysis, so from doing that, from presenting. Yeah, so it's all learning by doing and you know feedback from people who are, you know, either peers, or people who are even senior so earlier on when I was an assistant professor. You know I was getting schooled by them but you know now I just present work and we talk as peers I get feedback, so that's been good. The learning to teach with data, you know that that's something I've entirely have to figure out on my own. And, you know, I get a lot of practice with that, because nobody else in my department teaches quantitative anything. So I've had to do a lot of it, and I just sort of learned through trial and error there are, you know, what I recommend for that is actually books written by people who are good at it, not about how to do it but actually demonstrating it. So, you know, there are a lot of books written by, you know, good economists, sociologists, who can take technical material out of journals and express it in a way that is clear and easy to follow. And just reading a lot of those books, helps you to understand what it takes to make what seems like technical material accessible to lay people.

I: Okay.

UCSB9: And so that that's actually the best, those are some of the best examples for me.

I: In terms of exchanging assignment plans or syllabi or other instructional resources, have you ever done that with those departments you collaborate more closely with?

UCSB9: No, I can't. I can't because the people I collaborate closely with are in economics, and they can expect their students to know some things. So, their assignments simply wouldn't work for my students. And as I said in my department there's nobody who would actually do anything quantitative in class.

I: And what about outside UCSB, with other universities?

UCSB9: Honestly I've never really needed to do that. I mean I've done this enough that I've sort of got a knack for it. You know there are, every once in a while, I run across something on YouTube where there’s somebody who's actually you know, managed to teach people about data and a very creative way. So, again, you know I mentioned gap minder. You know, the founder of gap minder, he's dead now but he was this Norwegian public health guy who does extremely, extremely useful videos and you know shows you how you can use animated data to explain very simple things to students, I've been trying to push it a bit harder to use animated scatter plot to explain more complicated things. And it works but I mainly do this through trial and error. I don't get it from other people. Again, because there are very few places, and very, you know, because we live in this bifurcated world where either you do quantitative work or you to qualitative work, there are actually very few places where you have someone trying to teach

students to do quantitative work where those students will never before, and never again see quantitative work. Usually you can assume that there is a sequence of there's something before or there'll be something after. And because that never happens here I have to make it all up.

I: Okay. Considering evolving trends in your field, what types of training, or assistance would most be beneficial to instructors in teaching with data? What would you like to see, you know, being offered on campus for you to teach with data? You mentioned a few things but if you can explore it a little bit more that will be helpful.

UCSB9: When you say offered, are you talking about courses, are you talking about labs, or are you talking about infrastructure or support structures. What do you mean?

I: Like the types of training or assistance, it can be broadly understood as infrastructure, or even people to help. You know, what are types of training, or things that could be offered?

UCSB9: Okay, so I've heard, I've heard periodically so there's. Okay, so, you know, my position is you don't teach people to work with data, unless you teach them how to use data to solve problems or to answer questions right so you have to do that. That takes far more than just sort of dropping into a lab or a few class sessions that takes a sequence. So that's what needs to be there. Now there's two related things here, one, one issue that we have at UCSB is that if somebody wants to learn to program a computer, and they're not a computer science major there is no class they can take. So, all the programming classes are restricted entry. This has to change, this is a very serious problem. And the related, so, and I'll tell you where I'm going with this. You know, if we need to have sequences where people learn to solve problems and answer questions using quantitative data. And if we need to expand our access to training in, not just social scientific data analysis of the sort of classical variety that I'm doing but also the kinds of Baysian analysis, machine learning and things like that that students increasingly want to have because that's the current niche in the labor market. The way we solve this is by actually having, you know, either data sciences or a quantitative social science minor of some sort. Right, where we actually provide FTE [referring to full-time equivalent] for it. Right, so the campus, the campus could reach out and find faculty in other departments, right, including people like me. Right, who would absolutely love to spend you know one one quarter of the four that we teach, right, one of our teaching courses would be actually teaching students to apply data to solve social science problems. Right. So if we could do that, it would be incredibly popular, the demand is absolutely there, we would have to, you know, work at the campus level with the various deans, who would have to find the faculty FTE somewhere to make this happen. We would have to back it up with computer labs. But I think if we got a bunch of us like minded people around the table and we're actually told there will be the resources to make this happen, we could do great things. But that's, it's going to take sequencing, you can’t ad hoc this, this is a pedagogical issue and it has to be handled as a pedagogical issue, it has to be handled with sequences, problems, and resources.

I: Thank you. Yeah, wrapping up. So is there anything else from your experiences or perspectives as an instructor on the topic of teaching with data more broadly that you would like to share with our group.

UCSB9: Yes, I think in addition to this, so that I've talked about two different pedagogical models here the one I was just talking about is you can have a minor, where students who want to get serious with this stuff self select and then they kind of proceed and they go through it. I think that's really useful, it’s something to have. The other thing to have, is for freshmen to have, like I said, some kind of a core, a core knowledge sequence that students coming in so not, you know, change the way that we do, GE [referring to General Education requirements], I think, the way that we do GE is a bit scattershot. The general education credits. If you actually have, you know, some sequences the freshman year, taught by, you know, really well regarded instructors, find the ones who've been getting the teaching awards etc across campus, put them in the classroom for freshmen, make them work across the two cultures. So, you have some classes that are a little quantitative, some that are a little bit qualitative, some that are both and ensure that students coming in, who say things like I'm not a numbers person, etc. are forced to actually reconsider that position. So that early on, they can make decisions and actually start to get into the kinds of majors that they’re actually well suited to. And you know that actually interests them. I'll tell you. Almost every time I teach my [class code], I have a few students in the class who come to me and say I'm struggling with this, I'm not a numbers person, I shouldn't have to do this and please excuse me I won't learn this material, and almost every single one of the people who says this is female. Because that's what we do as a culture right, we've historically toward women they can't do math. And it's rubbish, all the empirical evidence says that's rubbish. And you know the other group of students I have who come into my classes are students of colors, working class students who say I wanted to major in economics, and I couldn't do it because I couldn't hack the math etc. But I know that it'd be actually really good at it. If they actually had the support systems early on. So, starting from freshman year. We've got to start building the support systems to teach applied mathematics to students of color, and to women in particular and I’m not saying target them but like make it something everyone's got to do so that everyone starts from the same basis. And maybe we'll start to get some of this stuff sorted. So, a sort of general education core knowledge approach that incorporates quantitative material in a challenging way, you know students shouldn't feel like they have to study hard or they might fail it and it'll be a problem.

I: I see. Well, thank you so much. Anything else. Sorry if I cut you off.

UCSB9: No, not at all. I just feel like we are depriving a lot of our students from things that their more privileged peers are getting, and then we try to put a bandaid on it and it’s not working.

I: Thank you.

Interview 10

I: I'm gonna start recording then. Okay. And the first note that I want to make here is obviously with Covid-19 we need to acknowledge that teaching and learning has been significantly disrupted in the past month, due to the pandemic and I'm not sure if that's affected your teaching specifically, but for any of the questions that I'm about to ask you can feel free to answer with reference to either your normal teaching or your teaching practices that you've had to adapt for remote teaching. All right, and we're going to start out here, I would like to hear a little more about your background teaching undergrad students, and especially where that background intersects with your current and past research. And then what specific courses do you teach with undergraduates that work with data.

UCSB10: Okay. So, my research mostly uses either survey data or cross national administrative data. So, it's big data sets. And I use almost all quantitative methods. And so that's what I teach, and I’m actually one of the few people in my department that is comfortable teaching quantitative methods so I teach a lot of required classes, and particularly I teach graduate students at the graduate level every year I teach a course, at least one of the two required courses and quantitative research methods, and I teach at the lower division, a required course at the undergraduate level. It's called methods of sociological research. And it's sort of one from a menu of methods classes that students can choose from they can choose, you know, different either quantitative or qualitative methods. And this is just one of them. And this one is sort of really designed to give you hands-on experience analyzing sociological quantitative data, mostly survey data. And I think that a lot of the skills that I teach are applicable to qualitative data analysis as well because I train students in how to pose sociological questions about relationships between variables, how to interpret the results and communicate the results. So, I have them write a short research paper at the end. So, I'm trying to teach them the logic of sociological inquiry. At the same time that they're learning about social statistics so I try very hard not to make it a statistics class but rather to make it a data analysis class. So, for the grad course I let students choose their own data. For the undergrad course, I ask them to use the General Social Survey of the US population, as you may know, this is a represent representative survey of US adults that has been conducted about every two years since the 1970s, and it has a lot of, it's huge and it has a lot of demographic questions about social background like your family of origin your current family, your educational attainment, your occupation. All the standard kinds of demographic categories, it also has a lot of questions about your attitudes toward different social issues, and political issues. So, students, sorry go ahead.

I: Oh, you can continue but you're actually answering my next question already, which is great. I'm wondering what went into the process of choosing that data set specifically? Is it just that it has such breadth or are there not other appropriate data sets.

UCSB10: Well, there is another one that I use a lot for teaching which is a world value survey that is where they can draw data from about 70 different countries. It has less demographic information and more attitudinal, so if you were interested for example in, the relationship between what your mother did and what you do or something like that you wouldn't be able to do it with the world value survey, but it has that cross national breath. Yeah, the reason that I use the General Social Survey is because it's very easy to use, it's well documented. And I find most undergrads are interested in sort of demographic and social issues in the United States. And I'm familiar with it so it makes it easier for me to teach them how to use it. There are some sort of longitudinal surveys where people are interviewed every couple of years like the same person over and over again, but I think for undergrads the data structure is too complex and they spend so much time trying to figure out how to use the data set and it's just not feasible in 10 weeks.

I: Yeah, that makes sense. And how many students do you have in each of those courses, when you teach them?

UCSB10: So, in our department the methods classes are limited to 60. Most other classes are like 100, upper division. But they have sections and the sections are limited to 15 so there's two TAs. And each TA has two sections of 15, so that really helps because it's an opportunity for them to get one on one instruction.

I: Right, right. So, the next questions I have are about how the students then work with the data in that course. So how do they manipulate and analyze, or interpret data, what tools or software to do they use? And do you expect any prior knowledge, versus what do you teach explicitly?

UCSB10: Yeah. Unfortunately, I cannot expect any prior knowledge so I teach them step by step how to use SPSS. SPSS is what we're using, statistical package for the social sciences. And the reason that we use that, I mean that's actually the oldest statistical package for the social scientists, it's been developed and further developed over time since I don't know, probably the 70s. It started out as kind of a mainframe you know, it was used with punch codes and stuff, punch cards. So, I make them actually, and it's oriented around the point and click with the mouse, but I make the students actually write their own code rather than use the mouse to do the data analysis. I want them to be able to see the logic of what they're doing and I find that writing code is a better way of keeping track of what they're doing to the data. So, I have them write a program. I give them an example of my own program and, you know, I go through that line by line and I ask them to come up with their own research question. And to find variables that they can use to address that research question from the General Social Survey. And the reason that I use SPSS, because I actually would prefer to use STATA, is that UCSB has a site license for SPSS, which means that all the students are able to download the software onto their own laptop. And the with STATA, which is actually the program that I think it's sort of easier to use and is more widely used lately. There is no site license and the only place that students can use it is in the lab, in the computer labs so we meet in the social science, the SSMS lab.

Okay, but then they have to go you know after hours there's little access to the labs and so a lot of them really prefer to be able to work on their laptops at home. And so for that reason, I use SPSS. The same thing holds for the grad students although for the grad students we really need to also teach them STATA so we've been, you know, sort of stuck with having to find time, you know, document carefully the opening hours of the lab and ask them not to go there at night alone or anything. Um, so yeah that's kind of a bit of a problem that I have. I would love, love it if the campus could get a site license for STATA.

I: Well, I can say that is something that's changed with remote work. They did negotiate a site license for STATA because it ended up being cheaper than getting the licenses they needed in order to, you know, host it remotely. But I can tell you, I'm not sure what's going on with that, like if they're making it widely available to everyone or not.

UCSB10: Really? Because I would love to know.

I: Yeah, I can tell you more about it after the interview.

UCSB10: Okay. Yeah, so what else did you want to know for me about the software.

I: That's a great explanation for us. Let's see. And I think you've already, one of the questions was, to what extent are those tools or software important for your teaching? I think you touched on that already, especially I appreciated the comparison between STATA and SPSS. Because I don't use those, I don't know why someone would choose one or the other. But are there any general challenges related to students abilities to work with data in those software's that you’ve run into other than you know timing for the labs?

UCSB10: Yeah, the labs have been, I mean I don't know if you want to know about this but the labs, it's been very difficult to work in those labs. In that particular lab, I've just, and it's not the fault of the student employees who are, you know, trying very hard to provide support but I think it's just there's a lack of resources or there's not enough resources and the classrooms are not proper, are not often set up like sometimes they have the partition in and I have to wait 20 minutes till they take it out or sometimes it just, you know, the stuff doesn't work properly and I have to sort of wing it and I can't I can't show my PowerPoints, in both classrooms simultaneously. So, I've taught generally like eight in the morning in those labs and it's just been an ongoing problem, and I don't think it's any particular person's fault. I just think that there's something structurally wrong.

I: And with SPSS itself. Are there any, like, obstacles for students that you find that they commonly run into?

UCSB10: SPSS, it's just, you know, learning to use a new program is hard and some students you know have a mental block and they think they can't do it and so. But, yeah, for the most part, I'd say 90% of students by the end of the quarter are somewhat competent in using it. And I mean what they end up doing is, is really replicating what a social science professional does is they use the same kind of data the same software package and they do an analysis it's very simple and bare bones but, but I think it gives a sense of the research process.

I: Yeah, I know from my experience being on the learning side it's hard to realize that you are learning exactly the same skills that people used to apply for data science.

UCSB10: Yeah, right.

I: Or to feel like that's enough. And there's always more to learn.

UCSB10: The thing is, I mean I can say one thing about I'm, I'm teaching this quarter remote and actually with grad students so it might not be the same, but I feel like the remote, I mean the zoom sharing of the desktop actually works pretty well. It replicates pretty well what I do in the lab, which is a lot of times I just go around from Terminal to Terminal and help the students troubleshoot their program and, you know, why doesn't this work and why do I get these error messages. And so being able to do it. You know, have a student ask the question and say I can't get this to work, and then I say can you share it, and then they show their problem and I'm sort of able to help them fix it but meanwhile the whole class is watching. So I think that works actually pretty well and I don't have to sort of go around and do the same thing, 20 times.

I: Yeah. So, we found that for our workshops too. I'm coming to prefer to share the screen on zoom.

UCSB10: Yeah. So, there are things that I think we will retain from our remote instruction.

I: Yeah, that's interesting. Um, let's see the next question I have, I think you've touched on this a little already but maybe you can expand on it. The question is, you know, how do the ways in which you teach with data relate to the learning goals in sociology. And I don't know if there's anything more that you want to say about that other than that you're trying to integrate those in.

UCSB10: Yeah, I try to replicate the research process and give them a sense of what sociologists do so that's that's the part of it. And I try to sort of allow them to take some, you know, address their kind of sociological imagination in a sense or their curiosity about sociological relationships and to actually be able to address the question using real data. The other thing that I think is important is the writing process. So I try to, they do a final paper, and I try to teach them the formula that sociologists use when they write a research paper which is, you know, they introduce a problem, they review the literature on that problem so what do we know so far and what don't we know, and so how they're going to address a gap, and then present their data and methods. And then present their results, and then come back to the original question and say what does this tell us about this question, or how does this help us move the literature forward. So I think that's a very useful skill to have and to be able to sort of communicate sociological information. And, yeah, communicating to others what you've discovered through your analysis.

I: And then the other question was on, what kind of data skills... Are there any data skills that you teach that you imagine will prepare them for a future career? And I'm not sure with sociology if you're more focused on preparing the students to do research like graduate research or if there are like specific careers that people get into right after the degree.

UCSB10: I mean, No, I think, I mean, the minority will go to grad school in sociology, but it would certainly be useful for someone who did. Having that knowledge when they come to grad school. But, yeah, a lot of students will go into the you know things like teaching or marketing or, you know, some other kind of thing. So yeah, it teaches some communication skills and teaches them data management skills, it teaches them analytical skills. And I've actually had undergrad students tell me they got a job, some kind of internship, because they knew how to use the software. So I think, Yeah, I mean it is somewhat practical.

I: Yeah, yeah, so they're so using SPSS specifically?

UCSB10: So, yeah, in that case it was SPSS.

I: Yeah, that's interesting. I don't know how widespread SPSS is outside of academia, they must use it in some places. But I'm not very familiar with it myself, to be honest. And the next question here is, are there any ethical challenges then with teaching with data in these courses that you need to confront. And to what extent do those affect your teaching?

UCSB10: Oh I don't really have issues of confidentiality. The data I'm using are not, you know, they are very anonymized. I mean, the only thing is you know students' self-esteem or, you know, students’ feelings of frustration. I deal with students in all kinds of delicate states as I think all instructors do. And, you know, sometimes when people lack self confidence in a particular type of learning, it may be exacerbated. But yeah, I don't think it's any different from any other type of class.

I: Yeah, yeah. And, have there been any like policy or cultural changes here at UCSB specifically that influenced the way that you teach? With this course, or other courses.

UCSB10: Well, I think that, you know, in general, there's an awareness of, you know, racial justice, gender and issues of gender and sexuality, and sort of economic inequalities I mean all of these things are sort of rising to the, to the forefront of our of our cultural discourse right now and in recent years. I think in general, you know, social scientists are pretty aware of these anyway because a lot of the social sciences have that in mind, or because of those concerns so I don't know how much it changes the way we teach these courses. I hope, you know, I hope that people are aware of them.

I: I imagine since you know gender is your research area this is something that's already baked into the course.

UCSB10: Yeah. Right.

I: And then the next questions about the training for, that you have for students in the course. So it sounds like you have TAs, but are there other people that assist?

UCSB10: Yeah, no I mean just a TAs and it's fine. I feel like I have enough teaching support as I said, it's more that I would like to have more technical support. But you don't always know when you're going to need it and I can see that you can't just have someone standing there all the time. But it’s very distracting when things don't work. And when you have to, you have, you're sitting there in front of 60 students and you're, you can't get things to work.

I: Yeah, you know the technical part is difficult, especially with software updates and yeah you have all the technical bugs worked out and then

UCSB10: Yeah exactly they have to reboot the computers and that takes like 10 minutes and, yeah. So anyway, it's just a kind of a, that's been a bit of a struggle.

I: Yeah, I haven't talked to a lot of people that have had sections with TAs for more help for their students. Do you have a different focus between those and lectures like, do they work more with SPSS in the sections or do they do other activities?

UCSB10: I think sections, they answer questions, they allow students to work on their projects and they just kind of go around and help. So they meet in a computer lab as well.

I: Okay, so it's more it's more help for them to work on their coursework rather than like, there isn't any new material that's coming?

UCSB10: Oh no, not in sections. I know sometimes the TAs will kind of give a little mini lecture that kind of tries to re-explain what I explained in the lecture.

I: Okay. Yeah, yeah. Interesting. And do you know if there are other resources that students are using to learn how to work with data outside their formal coursework, you know like, online tools like data camp and things like that. I don't know if you hear about those or not.

UCSB10: Not that I know of, I don't know.

I: Yeah. Is that something that you encourage in your classes?

UCSB10: Well, if I knew about some good stuff I would.

I: Yeah, I imagine there isn't as much for SPSS and STATA as things like R and Python. That you know that seems like most of the free resources and like data camp and, you know, coding exercises that I see are usually centered around those rather than tools like SPSS. And then the next question is about your own training and have you, you know since doing your degree. Have you received any other training in order to teach with data, or to increase your own data analysis skills or learning other software?

UCSB10: Well, I certainly have, you know, learned other software. And I've sort of, I mean just through my research. But yeah I haven't taken courses in how to teach quantitative methods or anything. Yeah, I'm not aware of those but I'm sure they exist.

I: Yes, probably somewhere I'm not aware of them either. That would be interesting though. Do you use any data, or you obviously use data sets from other sources but are there any like syllabi or assignment plans, other instructional resources that you've received or adapted from other people and do you share your resources for other courses? I suppose, maybe outside the UCSB sociology department.

UCSB10: Yeah, when I've had other colleagues who are going to teach a similar course, I share my syllabus and I share my, you know some of my course materials. And I was gonna say something but I've forgotten what it was.

I: Shoot hopefully it comes back to you. Is that mostly within the sociology department here, or are there other colleagues at other universities that you share materials with too?

UCSB10: Well I guess I would say, I would share if people asked, I think I might have shared with some colleagues at UC San Diego, that we've exchanged syllabi. But yeah, I don't put them online or publish them or anything, I did publish some syllabi through the American Sociological Association for a course I taught on inequality. They have this collection of syllabi but I never was asked to share one on quantitative methods. I don't know if that is something that is being, you know, sort of collected through the sociological association but maybe it is.

I: Yeah, that's really interesting. I didn't know that there was a syllabi collection for courses, generally. I'm not sure about quantitative.

UCSB10: Yeah there is that, so there are sort of teaching resources that are collected by the American Sociological Association that people can, like if you're going to teach a course on X, you can sort of look at a collection of syllables.

I: That's really interesting. Have you found that, I don't know how much you've used it yourself but do you think that’s a good resource?

UCSB10: I haven’t used it myself that much but I think junior colleagues have found it helpful. Also the department collects syllabi and so if you’re going to teach a course like you can go to the department and say can I see some syllabi for this course that other colleagues have used.

I: Do you see any other types of training to be beneficial for instructors who are teaching with data, especially with quantitative methods?

UCSB10: Yeah, I mean, learning new methods, learning new software is always good and I mean I think through the Broom center I also familiarize myself with new methods and new data sets. There's another, another way that data sets are shared which is through Data Archive.

There's an archive service or center called ICPSR which is interuniversity... have you heard of it?

I: Yes, I'm familiar with it.

UCSB10: Yeah, so on Gauchospace, I have a link to that for, particularly my grad students, and I encourage them to search for data that way. So, if they're interested in, you know, incarceration or transgender people, you know they can, they can find surveys of people who are falling into, you know one or the other category that are very specialized.

I: Yeah, that's actually good for me to know that you point to that because I'm the point person in the library to advise on that.

UCSB10: Yeah, I use that a lot. I mean, both myself but also, you know, whenever grad students are asking and do I know of any data on X, I say well have you checked here. Right. So about half my students end up downloading data, my grad students end up downloading data from, from that for the, for the required course that I teach.

I: That's good to know that it's being used widely, we don't we don't hear about things until there's a problem.

UCSB10: That's an essential service I think. And I've used it too to get data.

I: Yeah, they definitely do some good work there, you know, I went to their summer institute last summer. You know they had to like to change it from in person to online just like within a month, so I have to imagine they were scrambling to do that, I can't imagine.

UCSB10: Yeah, I've had a lot of students go to that and the Broom center funds some grad students to go to that every year.

I: Okay. Good to know. Yeah, I think I spotted someone else from UCSB in one of the zoom sessions. So, is there anything else from your experience teaching qualitative methods that you want to share?

UCSB10: No, I don’t think so

I: Okay, I think that covers all of our questions then, thank you.