Survey for PhD Students in CSD – Using R and RMarkdown

The goal of this survey is to identify the needs of incoming PhD students in Communication Sciences and Disorders (i.e., Speech-Language Pathologist and Audiology) with respect to data management, organization, and statistical programming. Our goal is to develop an R-focused workshop for incoming PhD students in CSD. This workshop was initially intended to address needs identified by CSD PhD students at the University of Pittsburgh, but we hope to make it widely available to CSD PhD students.

The potential goals of the workshop are:

* Teach incoming CSD PhD students the basics and best-practices of using R for data management. The intention is not that students would be strong R-users after the workshop, but would be comfortable with basic use of R, RStudio, and statistical programming.
  + Many, if not most incoming PhD students in CSD have very little exposure to programming and using a programming language for data management, statistics, or data visualization can be intimidating. After completing the workshop, we hope students will be more comfortable and confident expanding their R skills independently.
  + Because students will be comfortable with basic R use, they can focus more on the statistical content of PhD courses taught in R, rather than getting bogged down by R programming.
* Promote social connections between PhD students in CSD. One of the consequences of the COVID-19 pandemic is that conference travel and networking has been limited. We hope that by bringing together CSD PhD students at the start of their programs, they will expand their network of peer supports for their PhD programs and beyond.

Right now, we’d like to ask you (the experienced PhD student or post-doc or junior faculty) about your experiences as you entered your PhD program and transitioned to the world of research, so that we can design this workshop to best address the needs of incoming PhD students. First, we will ask you about your current knowledge and skills in R and statistical programming. Then we’ll ask you about what you wish you’d had access to when you first entered your program. Finally, we want to know what you’d like to see more of now, so that you can continue your training.

Survey Questions

Section 1: About participant & their program:

Demographics

* Age
* Gender
* Year in PhD program
* Background – SLP/Aud clinical degree? Other job? Right out of school? Older phd student?
* Country? If international
* In a dedicated CSD program, in a multidisciplinary program (e.g. rehab sciences), CSD PhD student in related field (psychology, neuroscience)
* Roughly, how many other PhD students are in your program
* Carnegie rating of the university (R1, R2) and provide link.

Current knowledge & skills

* Do you use R or another programming language for data analysis/management (e.g. python, Julia, matlab)?
* Do you use other tools like Excel, SPSS, etc.
* What is your **primary** software tool?
* Self-rate your ability with all programming languages and software selected (some kind of likert…)
* Have you used Rmarkdown for generating reproducible reports? (in a class, in my lab, for a submitted or published manuscript)
* What tool do you use for data visualization (and is it R?)

Current PhD Program

* Where did you learn these skills? (independently, from class, from labmates/advisors/peers)
* What stats software have your stats & research methods classes used?
* What software does your primary advisor/lab use (primarily)?
* What % of your peers use a programming language like R/python/Julia/matlab as their primary tool?
* If you were going to learn R from scratch right now, would you have peer support at your institution to learn it?

Section 2: Hypotheticals.

Think back to when you first entered your PhD program (i.e., your first 1-2 semesters). [Paint a picture here of a potential workshop with a of flexibility]. Imagine a 1 or 2 day (or half day) online workshop with other incoming phd students in your field right before your first semester. The workshop might cover content such as [list a few of our potential content areas]. And it would be in R.

* How helpful would this workshop potentially have been for you during your 1st year.
* How helpful would this workshop have been over the course of your doctoral degree?
* How would this “hypothetical” workshop have helped you in your program?
* What would be the most important content of this workshop?
* What are the necessary qualities of a successful workshop like this?
* What qualities would make this workshop unsuccessful?
* What things might keep you from participating in the workshop?

Section 3: Current needs

* If there were workshops for current PhD students, what would be the most important content?
* What skills (related to data) do you feel you still need to work on before graduating?

Section 4:

Do you want to add anything else, have strong opinions, did we miss anything we should know. Open ended.

Questions we need to think about:

* How many leaders do we need?
* Lots of small groups