

Reed College

Summer Opportunity Fellowship Award 2021

Developing algorithm and dashboards for Yonsei University R and Python education program

Building an AI-based online adaptive learning program

Applicant Name: (Student name)

Email Address: (email)

Mentor Name: Professor Taeyoung Park

Email Address: tpark@yonsei.ac.kr

Faculty Name: Professor Kelly McConville

Email Address: mcconville@reed.ed

Project Proposal

On February 24th, I finalized my first ever R package for the Data Science class project. The project was an arduous task that required days of data wrangling and analysis, but I was further intrigued by the utility and intuitiveness of R during the process of making the package. I wanted more hands-on and real-world applications on R. Therefore, I went to professor Kelly McConville's office hours for her advice on finding summer research and internship opportunities. She is my current Data Science (MATH241) professor, and I took Introduction to Statistics with R (MATH141) from her during the Fall 2020 semester. Her passionate promotion of R throughout the courses attracted me to the R programming language, and I hoped that I could get some ideas on furthering my R journey from her. She directed me to the National Science Foundations (NSF) biostatistics Research Experiences for Undergraduates (REU) programs at Iowa University and the University of Southern California. In the search for more opportunities, I also asked whether cold-emailing non-Reed professors is acceptable, and she said it is okay as long as I am polite and concise. She also told me about Summer Opportunity Fellowship Awards when applying for out-of-Reed research. I started applications for REU programs mentioned by Kelly, and I started cold-emailing relevant statistics and biostatistics professors in Korea. My parents wanted me back in Korea for the summer break, so I sent a personalized email to the professors in Korea about my proficiency and achievements related to R. When I emailed twelve professors and was in the process of writing two more, Professor Taeyoung Park replied. Professor Park is a professor of applied statistics and data science at Yonsei University, one of the top universities in Korea. He wants me to assist in developing algorithms and question banks for his Python and R educational program. The professor plans to implement a program based on his algorithm to teach his students the R and Python programming languages. He also plans to release it for public use when it is deemed reliable. I replied that while I had limited experience in writing algorithms, I could help with generating R question banks. I drafted nine R questions with difficulties ranging from easy to hard and sent the questions to him in two days. He replied that a start-up that specializes in Python question-banks already generates the R question-banks. It was disheartening, but I replied to him to contact me for any other R-related research opportunities. He did and offered me a summer research position in his lab where I will create R shiny dashboards for his education program and help with the algorithm development. I gladly accepted his offer.

During my 10 weeks with Professor Park at Yonsei University, I will be working at the professor's lab at the university's department of applied statistics. My primary responsibility will be to write R code using the R Shiny package to generate dashboards for the Python and R programming education program. I will also assist in developing algorithms that give appropriate questions according to the program user's ability. R Shiny has two parts: User Interface (UI) and Server, to create a dashboard. In programming lingo, it is like building a frontend and backend of an application, respectively. The R Shiny UI generally works by taking inputs from the user and showing interactive outputs. The R Shiny Server works by controlling and adjusting the flow of

the information given to and from the user. Moreover, the general idea of the algorithm that I will assist in writing is to suggest R or Python questions according to the user's current ability rather than giving them programming questions with discrete levels of difficulties. I will be primarily working at home, but there will be weekly in-person meetings with the professor and his graduate students to update my progress and upload any data that cannot be transferred by Github.

This research opportunity would challenge me to write and learn about complex algorithms, and it will provide a chance for me to apply the R programming concepts that I learned in and out of Reed College into real-world applications. It also has the potential of generating thesis ideas because I will have obtained new sources of information from this research opportunity that is relevant to my expertise in R. I would also learn theoretical ideas of adaptive learning algorithms and time response theory from Professor Park and his graduate students. Moreover, I would like to be a data scientist in the future, and this usually requires a master's degree or a Ph.D. This research opportunity would allow me to experience the field of data science in a big university like Yonsei and better understand what graduate work in data science entails. Moreover, the connections that I build in Korea could allow me to further my journey as a data scientist. Lastly, when the education program is matured for outside use, I could also ask Professor Park to bring his programming education program to Reed College for Reedies who want to learn to code in R and Python.