

Learning *R*: Needs Analysis, Learning Taxonomies, Methodology, and Visualization

Alon, Friedman PhD¹, Edd Schneider, PhD²

1 & 2 School of Information, University of South Florida

*Contact: alonfriedman@usf.edu

Keywords: New user, College students, curriculum, statistics, taxonomies, visualization.

The growing popularity of *R* has presented new challenges for educators as well as for new users. This paper addresses *R* from an educational standpoint, providing a methodological approach for teaching statistics and programming in any arts and science college curriculum. For many years, the main feature of the development of *R* focused on the growth of its popularity and the widening accessibility of the technology across platforms. However, with this growth more attention is needed to address specific educational concerns of people wanting to teach *R*. The open source nature of the application presents unique problems for educators, since the mechanics and functionality of *R* are constantly evolving. This in turn makes *R* a challenge in terms of instructional design.

In order to address this challenge, this presentation will share the results of the first year of an experiment teaching the *R* to undergraduate students in the University of South Florida's School of Information. Faculty taught the subject in both face-to-face and online class environments. The educational context of the lessons utilizing *R* focused on the power of the application to produce visualization; specifically visualization of large data sets. While the results of measuring student feedback from the class are still being measured and analyzed, early results indicate that using visualization can be guiding a factor for new users to understand why they are learning about *R*, while simultaneously improving their understanding of statistics by using *R*. Essentially, the classroom model discussed here uses the visualization aspects of the application to motivate students to further their own understanding of statistics through a better command of *R*'s data analysis functionality.

During the spring of 2013, we taught two different courses in *R* in our undergraduate program. A pedagogical framework was developed which encouraged students to become more personally involved in statistical production and control by creating visualization. Learning in an academic course on data analysis is more complex than merely getting students to learn a language for its own sake, and faculty have found that students do not necessarily learn through having an example explained step-by-step through the basic procedure of manipulating data. Many authors discussed the subject of "delivery" of statistics in the classroom and many point out the lack of standardized "visualization delivery". Further more, there are numerous educational taxonomies that could potentially be applied in the context of improving *R* education. However, a survey of materials found little in the way of attempts to explore a theoretical framework to teach statistics, programming and visualization using an open source platform.

We conducted a methodology study in the context of the specific statistics and programming languages courses taken prior to the students' course in *R*. Our framework blends Constructivist and Cognitivist approaches to instruction. Details of the results of this methodology will be presented, along with an analysis of how these results will influence teaching in the future.