Using Pattern Recognition Techniques to Analyze Educational Data

49th Annual Frontiers in Education (FIE) Conference

Handout

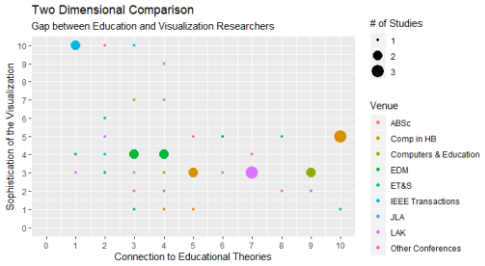
Camilo Vieira, PhD cvieira@uninorte.edu.co
Alejandra J. Magana , PhD admagana@purdue.edu
Mireille Boutin, PhD mboutin@purdue.edu

For this part of the special session, we have designed different R scripts that will allow you to use the methods we have discussed today. Let's start by opening the R programming environment **RStudio**, and select what you want to do from the alternatives below.

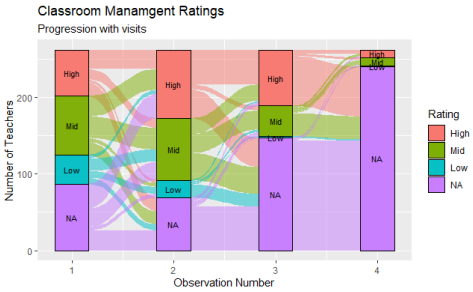
From an early introduction to R, to creating visualizations and conducting cluster analysis, it all depends on your previous experiences and your interest in these methods. All the files are available at: **C:\workshop**

If you know nothing about R programming, we suggest starting by working on the *IntrodutinoToR.r* file. If you already know a little bit of R programming, feel free to start in any of the files we list below:

- **IntroductionToR.R**: Basics of R, you will learn how to program in R, create and manipulate variables, data structures and functions.
- **IntroToBasicStats.R:** Descriptive and inferential statistics with R
- **introToPlots.R**: Create simple plots with ggplot such as scatter plot, line plot, and histograms. Here we will create a scatter plot like the one we used in our Literature Review study:



- **observationPlots.R**: Create parallel coordinate plots that describe how individuals move from one category to another one between classroom observations.



- **clusteringPlusPlots.R**: Analyzing think-aloud data using clustering and plotting.