My methodology

There are eight different features about students’ confidence in Mathematics in my dataset, since I only want one feature about students’ confidence in Mathematics as the final feature, so I need to combine the eight different features to one feature. And because I don’t know if each feature has any relationship with other features, so I will choose the PCA as my methodology to analysis this eight features, and combine them as one final feature.

I will use RStudio as the tool to do the PCA for the eight features in my dataset, and the important thing for me is to figure out how to write the code in RStudio for PCA. After doing the PCA, I need to consider the methodology for my next step.

In the result of PCA, I will get the proportion of each feature in the component, and some features’ proportion may higher than other features. There should be a latent construct between this higher proportion features. And I will drop features which have the low variance based on the result of PCA, but the dropped feature may important for the students’ confidence.

For the validation part, I don’t know which validation method can be used for my dataset, I need to understand more different validation methods first.