

Principal Component Analysis Assignment

Anahita Zarei

Create a reproducible document containing your analysis for the following problem and submit your R markdown and html or PDF files on Canvas. (25 points)

Write a function in R called *myPCA* with the following properties:

1. The function must have 2 inputs: (2 points)
 - a data frame that contains the data set
 - a scaler that denotes the desired percentage of variance preserved in the transformed data set
2. The function must have 3 outputs returned in a list, called *number*, *PCA*, and *transformed*: (3 points)
 - the minimum number of principal components that is necessary to retain the desired variance specified by the user (5 points)
 - the principal components corresponding to previous part, in a matrix (5 points)
 - the transformed data set in a matrix (5 points)

Run your function on the first 5 features of R *longley* data set with the desired retained variance of 0.98. Longley data set is a data frame with 7 features related to economics. You will use the following features for this assignment: GNP.deflator, GNP, Unemployed, Armed.Forces, and Population. (5 points)