STAT 53533: R Assignments 01
Upload your answers**, as a single file,** in Blackboard by
11:59 PM on Wednesday, November 06, 2024

Separate your answer in two parts (within the same file) as follows:

1. Part 1 must contain complete answers for all questions **without** any R code. It must have all numbers/values/diagrams as asked in the questions and should be accompanied by necessary explanations.

2. Part 2 contains only the R code and nothing else. **Any output/numbers/explanations written inside the R Code will NOT be graded**.

3. **You must type all your answers and explanations for submission, as a professional report.** Scanned copy of any **handwritten answers/explanations will not be graded.**

Refer to the compressed file attached with this assignment. It contains a folder named "yalefaces_train" and another folder "yalefaces_test". [1+2+1+2+1+2+1=10 points]

(a) Use the yalefaces_train folder as the training dataset and report the values of n and p.

(b) Do a PCA on these images. Report the ELOV values for different choices of the number of principal components. What should the value of k be if we allow up to 20% loss?

(c) Now, choose <u>any two</u> test images from yalefaces_test folder (**Every student's choice should be independent of any other student's choice. Mention your chosen image names clearly in your written report**). Display your chosen images.

(d) Reconstruct your chosen images using the value of k from part (b). Plot the reconstructed images.

(e) Now, choose k again allowing up to 4% loss.

(f) Repeat (d) for the two previously chosen test images using the new value of k from part (e). Plot the reconstructed images.

(g) Comment about the the qualities of two reconstructions that you get in part (f) relative to the ones you got in part (d).