## **Homework Assignment 9 [30 points]**

STAT430 Unsupervised Learning – Fall 2021

<u>Due</u>: Friday, October 29 on Compass at 11:59pm CST.

- Answer questions 1-5 in case study 1 (assignment\_09\_fashion\_mnist.ipynb)
- Answer questions 1-4 in case study 2 (assignment\_09\_shill\_bidding.ipynb)
- Answer questions 1-2 in case study 3 (in this pdf below)

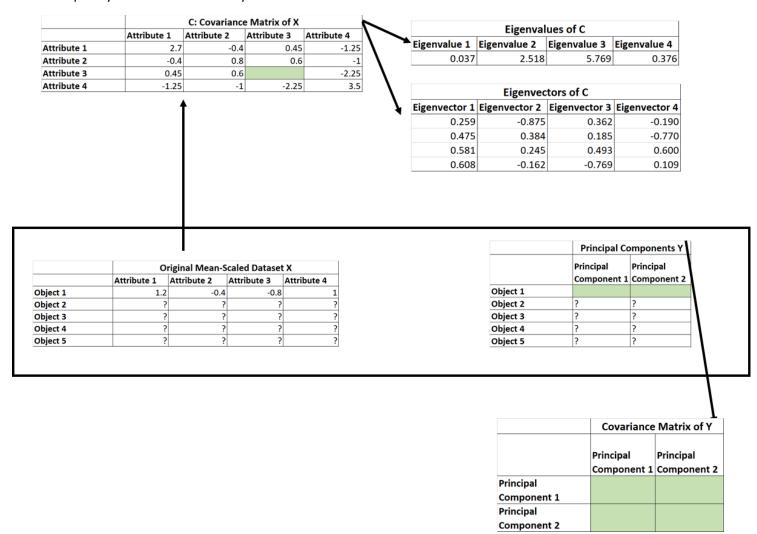
Problems	Points
Case Study 1	
1.1.1	0.25
1.1.2	0.5
1.2	0.5
2.1	0.5
2.2	0.25
3.1	2
3.2	0.75
4.1	1.25
4.2.1	1.25
4.2.2	1.25
4.2.3	1
4.3.1	0.5
4.3.2	0.5
4.3.3	0.75
4.4	1.25
4.5	1.25
4.6	1
4.7.1	1
4.7.2	1
5.1	0.75
5.2	0.25
5.3	0.5
Case Study 2	
1	0.25
2	0.25
3	0.5
4.1.	1
4.2.1	0.5
4.2.3	0.25
3.1	0.5
3.2	0.5
3.3	0.25
3.3	0.5
4.1	0.5
4.2	0.25
4.3	0.5
5.1	0.5
5.2	0.25
5.3	0.5
6.1	0.5
6.2	0.75
Case Study 3	
1 3	
2	0.5
_	0.5

## Case Study 3: Artificial Dataset

## Question 1

A partially completed, mean-centered,  $X_{5\times4}$  matrix is given below. It's covariance matrix  $C_{4\times4}$  is also given below. Finally, the four eigenvectors of  $C_{4\times4}$  and their corresponding eigenvalues are given below. Note that these eigenvalues are not in order.

We decide to use PCA to project  $X_{5\times4}$  onto the matrix  $Y_{5\times2}$  (ie. two principal components) also shown below and partially completed. Use the information given to fill in the blanks for the 7 green boxes shown below. Explain your answers or show your work.



## Question 2

What percent of total original attribute variability is preserved in the two principal components above?