Quiz 5

Due No due date	**Points** 20	**Questions** 10	**Time Limit** None

Attempt History

	Attempt	Time	Score
LATEST	Attempt 1	32 minutes	16 out of 20

Score for this quiz: **16** out of 20
Submitted Nov 19 at 1:17pm
This attempt took 32 minutes.

Question 1

2 / 2 pts

(1) True/False: Clustering is an unsupervised learning problem.

Correct!

- (•) True

  True – When we use cluster analysis we typically do not have a label or a response variable for each observation, hence clustering is an unsupervised learning problem.

- ( ) False

Question 2

2 / 2 pts

(2) True/False: Principal Components Analysis can be used to create a low dimensional projection of the data for use with clustering.

Correct!

- (•) True

True – Cluster analysis is better performed in lower dimensions than higher dimensions. One way to reduce the dimension (and get orthogonal variables) is to use Principal Components Analysis to transform your data.

False

Question 3

0 / 2 pts

(3) True/False: Common factors estimated using maximum likelihood estimation with a PROMAX rotation are orthogonal.

ou Answered

True

False – The PROMAX rotation is an oblique rotation, and hence the rotated common factors will be correlated with each other.

orrect Answer

False

Question 4

0 / 2 pts

(4) True/False: Common factors estimated using Iterated Principal Factor Analysis with a VARIMAX rotation are orthogonal.

orrect Answer

True

ou Answered

False

True – The VARIMAX rotation is an orthogonal rotation, and hence the rotated common factors will be orthogonal to each other.

Question 5

2 / 2 pts

(5) True/False: In cluster analysis the choice of similarity measure will affect the cluster assignments.

Correct!

- (•) True

True – The most common similarity measure is Euclidean distance. Other distance measures, or metrics, will yield different results. Some measures are more preferred over other metrics given the type of data or other data properties.

- ( ) False

Question 6

2 / 2 pts

(6) True/False: When computing principal components the data should be standardized, i.e. the data should be centered and scaled to a (0,1) distribution.

Correct!

- (•) True

True – Since principal components analysis aims to explain the variance in the data, all of the variables should be on the same scale. If the variables are not on the same scale, then PCA will consider the variables with the larger scale to be more important.

False

Question 7

2 / 2 pts

(7) True/False: Cluster analysis can only be performed on continuous variables.

True

Correct!

False

False – Cluster analysis can be performed on discrete variables, but the correct similarity measure must be used.

Question 8

2 / 2 pts

(8) True/False: Hierarchical clustering requires that the number of clusters be specified in advance.

True

Correct!

False

False – Hierarchical clustering techniques fall into either: (1) agglomerative methods or (2) divisive methods. Both subdivisions provide multiple partitions of the data that can be visualized through a dendrogram.

Question 9

2 / 2 pts

(9) True/False: Factor Analysis and Principal Components Analysis have the same objective of modeling the correlation structure in multivariate data.

- True
- False

Correct!

False – Factor Analysis is designed to explain the correlations between a set of multivariate data while Principal Components Analysis is designed to explain the variance in a set of multivariate data.

Question 10

2 / 2 pts

(10) True/False: Since cluster analysis is an unsupervised learning method, two different cluster partitions cannot be compared.

- True
- False

Correct!

False – Two cluster partitions can be compared using a variety of metrics. One such metric is the Rand Index. SAS will compute the Pseudo F, the Pseudo T-Squared, and the CCC metrics for cluster comparison.

Quiz Score: **16** out of 20