MATH1309 – Practice Problems 10

This week we will explore the solutions to the Examples in the lecture notes using the procedure PROC CLUSTER

Question 1

Example: The distances between pairs of five item are given below:

	1	2	3	4	5
1	0				
2	4	0			
3	6	9	0		
4	1	7	10	0	
5	6	3	5	8	0

Cluster the five items using the

(a) Single linkage

(b) Complete linkage and

(c) Average linkage

methods. Draw a dendograms.

You can complete this by hand, or using SAS

Question 2

The datafile on Canvas, and in SAS Studio, Example17.dat gives the data collected at 25 hotels:

where X_1 = average daily occupancy

X_2 = monthly average number of check-ins

X_3 = number of hours per week service desk is in operations

X_4 = total common use area in square feet

X_5 = number of building wings

X_6 = capacity

X_7 = number of rooms

X_8 = monthly work-hours required.

Identify the hotel-clusters using:

a) single linkage
b) complete linkage and
c) average linkage

d) Use the k-means method to cluster the hotels into three groups.

e) Compare the results obtained using the different clustering methods.

Question 3

By hand

Example: Certain characteristics associated with six workers in a factory are listed below:

Person	Married	School	Sex
1	yes	Private	Female
2	no	state	Female
3	yes	state	Male
4	yes	state	Male
5	no	Private	Female
6	no	Private	Male

Find the similarity of the workers based on the above observations

Calculate the distances and construct a distance matrix