



Homework Assignment 4 [30 pts]

STAT430 Unsupervised Learning - Fall 2020

Due: Friday, September 24 on Compass at 11:59pm CST.

Questions 1-11: [20.5 pts] Complete questions 1-11 in Assignment_04.ipynb.

Question 12: [2 pt] Calculate the rand index between the two clusterings shown below by hand. Show your work.

	Clustering 1
Object 1	9
Object 2	9
Object 3	8
Object 4	10
Object 5	9

	Clustering 2
Object 1	9
Object 2	9
Object 3	9
Object 4	10
Object 5	10

$$f_{00} = \{(1,4), (2,4), (3,4), (3,5)\} = 4$$

$$f_{11} = \{(1,2)\}$$

$$f_{01} = \{(2,3), (4,5), (1,3)\}$$

$$f_{10} = \{(1,5), (2,5)\}$$

Rand Statistic = ...

$$\dots \frac{f_{00} + f_{11}}{f_{00} + f_{01} + f_{10} + f_{11}}$$

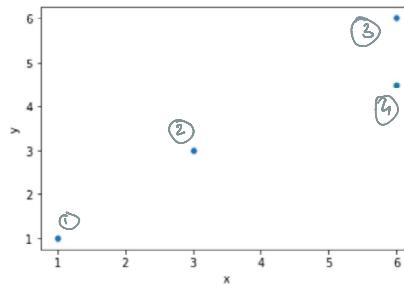
$$= \frac{4+1}{4+3+2+1}$$

$$= \frac{5}{10}$$

$$= 0.5$$

Question 13: [7.5 pts] For the dataset below, sketch three dendograms: a.) one found with agglomerative hierarchical clustering with single linkage proximity [2.5 pt], b.) one found with agglomerative hierarchical clustering with complete linkage proximity [2.5 pt], and c.) one found with agglomerative hierarchical clustering with average linkage proximity [2.5 pt]. Show your work and make sure your dendograms are drawn to scale and labeled appropriately.

	Data	
	x	y
Object 1	1	1
Object 2	3	3
Object 3	6	6
Object 4	6	4.5



	Distance Matrix			
	Object 1	Object 2	Object 3	Object 4
Object 1	0	2.83	7.07	6.1
Object 2	2.83	0	4.24	3.35
Object 3	7.07	4.24	0	1.5
Object 4	6.1	3.35	1.5	0

1) With Agglomerative hierarchical clustering with

1) With Agglomerative hierarchical clustering with Single linkage proximity.

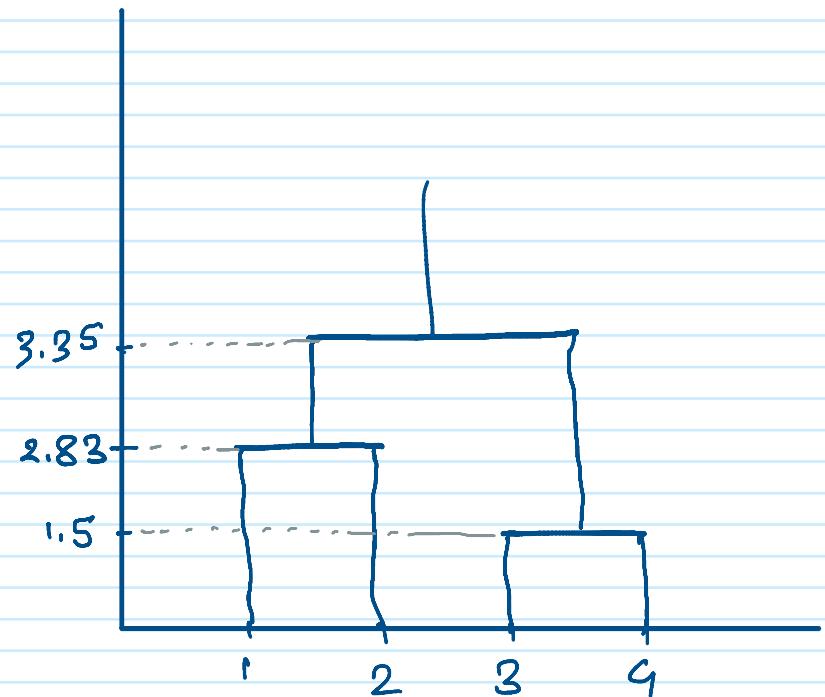
Iteration 1:

	1	2	3,4
1	0	2.83	6.1
2		0	3.35
3,4			0

Iteration 2:

	1,2	3,4
1,2	0	3.35
3,4		0

	1,2,3,4
1,2,3,4	0



2) With agglomerative hierarchical clustering with complete linkage

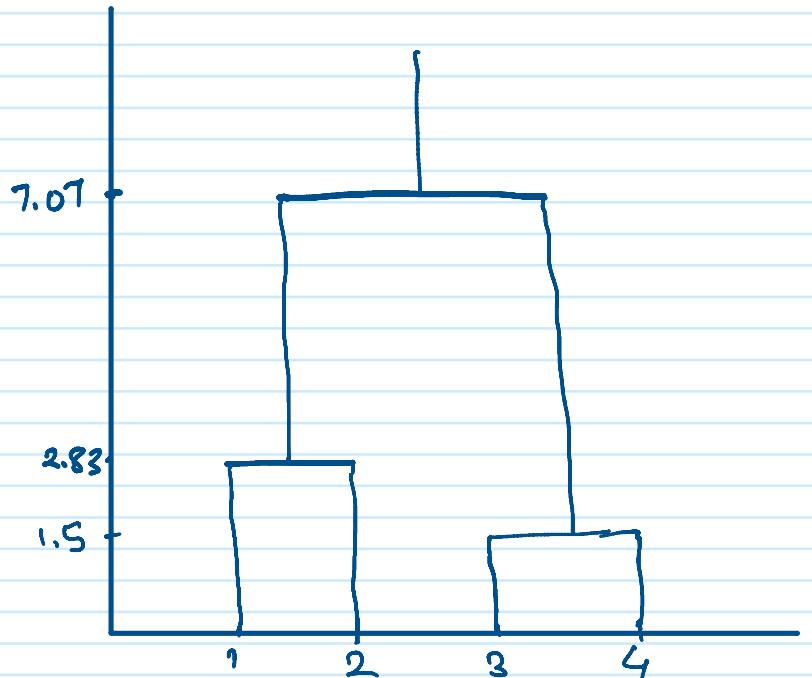
Iteration 1:

	1	2	3,4
1	0	2.83	7.07
2		0	4.28
3,4			0

Iteration 2:

	1,2	3,4
1,2	0	7.07
3,4		0

	1,2,3,4
1,2,3,4	0



3) With agglomerative hierarchical clustering with avg. linkage

Iteration 1:

		1	2	3,4
1	0	2.83	6.585	
2		0	3.195	
3,4			0	

Iteration 2:

		1,2	3,4
1,2	0	5.19	
3,4		0	

Iteration 3:

		1,2,3,4
1,2,3,4	0	

