Homework Assignment 4 [30 pts]

STAT430 Unsupervised Learning - Fall 2020

Due: Friday, September 24 on Compass at 11:59pm CST.
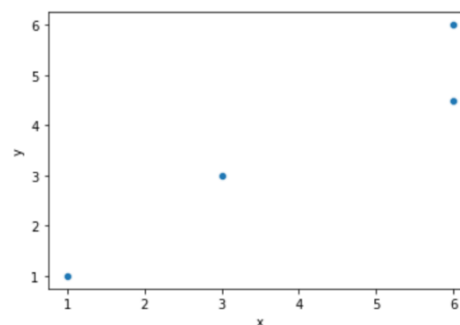
Questions 1-11: [20.5 pts] Complete questions 1-11 in Assignment_04.ipynb.

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

Question 13: [7.5 pts] For the dataset below, sketch three dendrograms: 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 dendrograms 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