

## Homework Assignment 5 [30 pts]

STAT437 Unsupervised Learning – Spring 2025

Due: Friday, February 28 on Canvas at 11:59pm CST.

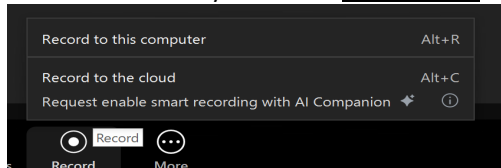
**Video Question:** Make a 3-4 minute video explaining your answer to **question #12 in this pdf**.

### IMPORTANT Video Element of ALL Homework Assignments:

- In order to receive points for each video submission, you need to do **ALL** of the following.
  - Have your camera on.
  - Show your FULL screen in Zoom (not just a particular application).
  - We should be able to hear the audio. Make sure to turn your mic on.
  - You should give a good faith attempt to answer the prompt.
  - Your video meet the minimum time requirement.
  - It should not sound like you are just reading off a script.
  - It's ok if your video recording is not the most eloquent. What's important is that you are putting together YOUR authentic thoughts on your particular understanding of the assignment and the lecture content.

### How to Submit Videos:

- You should record your videos in your UIUC Zoom client.
- You should record your videos To the Cloud.



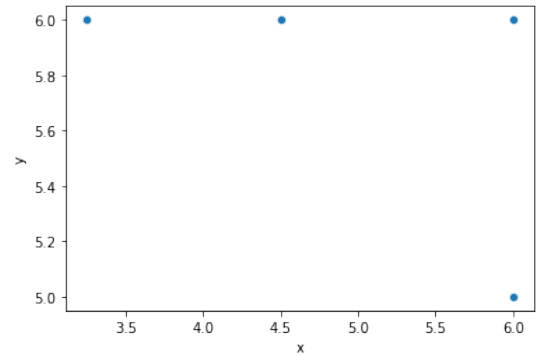
- You can find your recording link at <https://illinois.zoom.us/recording/>.
- Click on the corresponding video and Copy shareable link to paste the link in Canvas.

Problem	Points
1.1	0.25
1.2	0.75
2	1
3.1	1
3.2	1.25
3.3	1.25
4.1	1.5
4.2	1
5.1	0.75
5.2	1
6.1	0.75
6.2	1
7.1	0.75
7.2	1
8.1	0.75
8.2	1
9.1	1
9.2	1
9.3	1
10.1	1.5
10.2	1.5
11	4.5
12	3
Video Question	1.5

**Questions 1-10:** See the Jupyter notebook

**Question 11:** For the dataset below, sketch three dendrograms: a.) one found with agglomerative hierarchical clustering with single linkage proximity, b.) one found with agglomerative hierarchical clustering with complete linkage proximity, and c.) one found with agglomerative hierarchical clustering with average linkage proximity. Show your work and make sure your dendrograms are drawn to scale and labeled appropriately.

	Data	
	x	y
Object 1	4.5	6
Object 2	3.25	6
Object 3	6	6
Object 4	6	5



	Distance Matrix			
	Object 1	Object 2	Object 3	Object 4
Object 1	0	1.25	1.5	1.802776
Object 2	1.25	0	2.75	2.926175
Object 3	1.5	2.75	0	1
Object 4	1.802776	2.926175	1	0



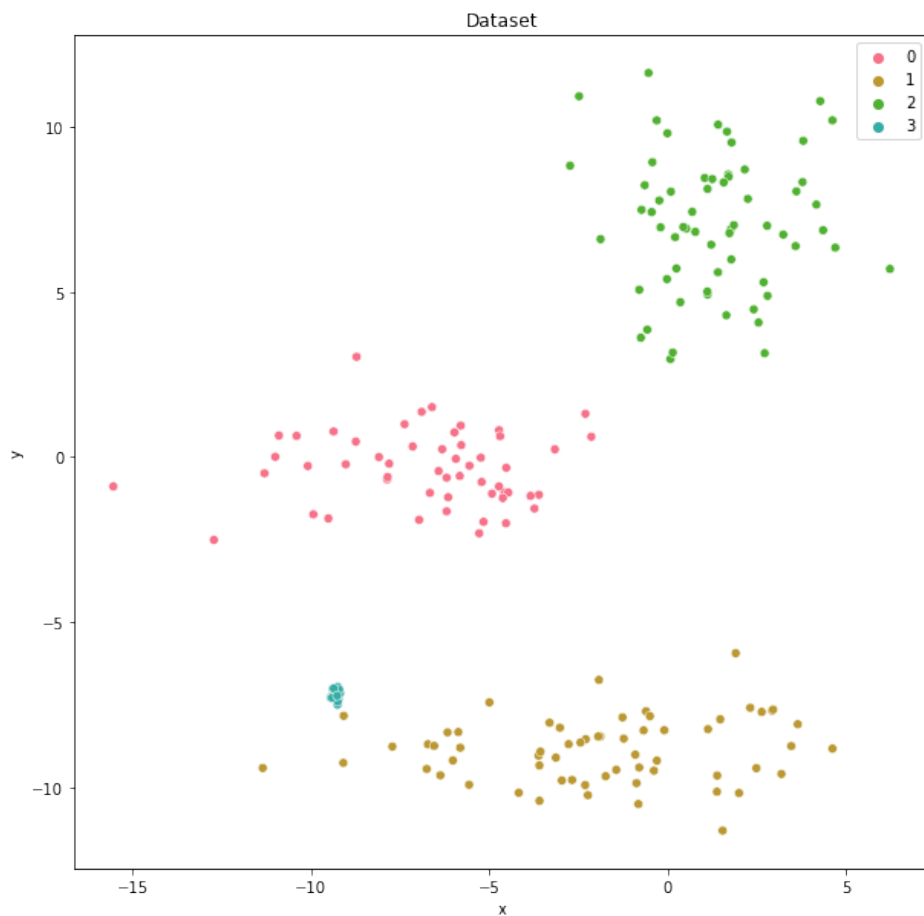
### Question 12:

**12.1.** We have clustered the dataset below using hierarchical agglomerative clustering using each of the following linkage functions:

- Single linkage
- Complete linkage
- Average linkage
- Ward's linkage

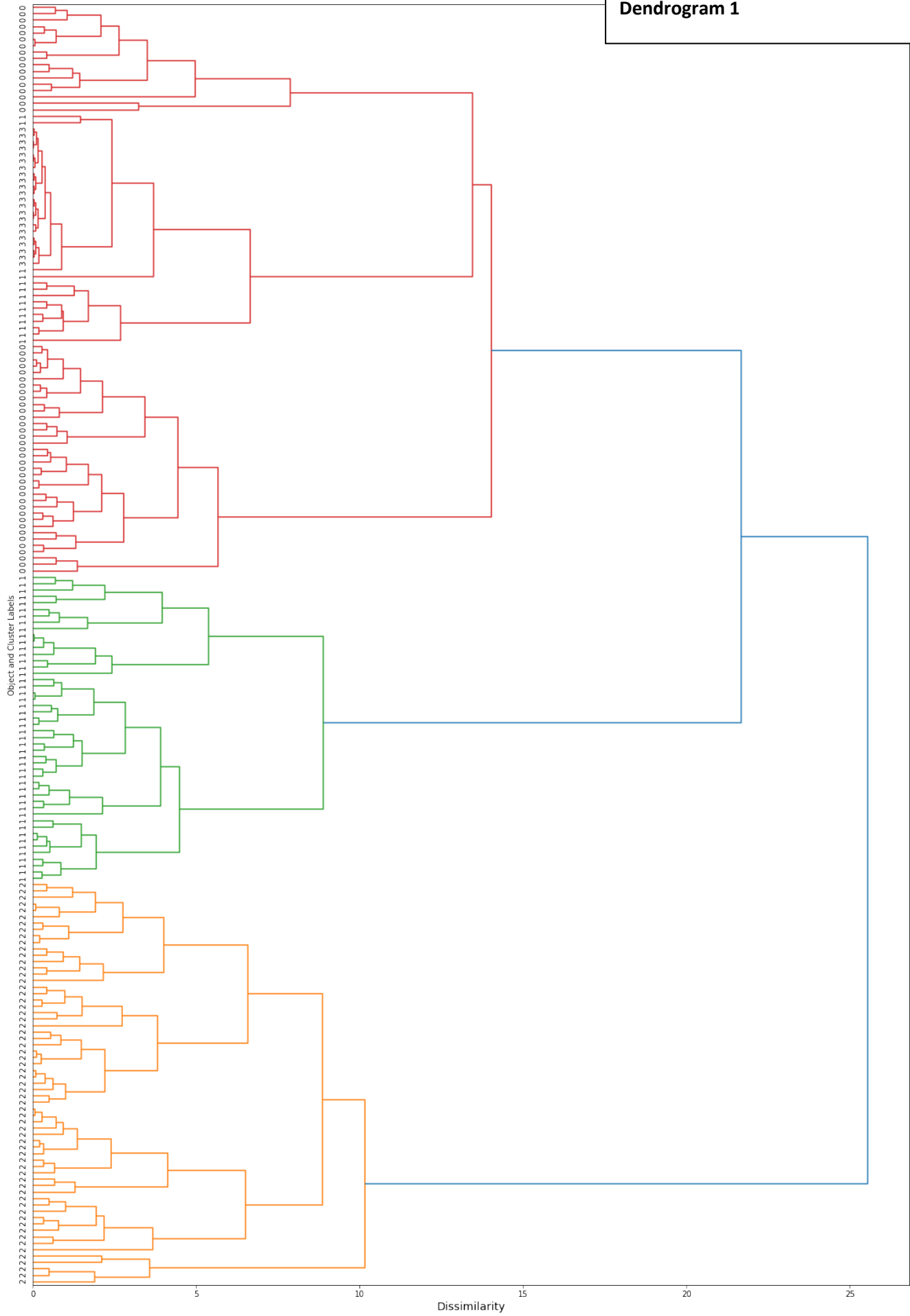
Each of the dendrograms for these linkage functions are also given below. State which dendrogram corresponds to each linkage function.

Hint: This dataset also has preassigned class labels {0,1,2,3}. Rather than label each unique object id in the dendrograms, we instead give the pre-assigned class label of the object instead.

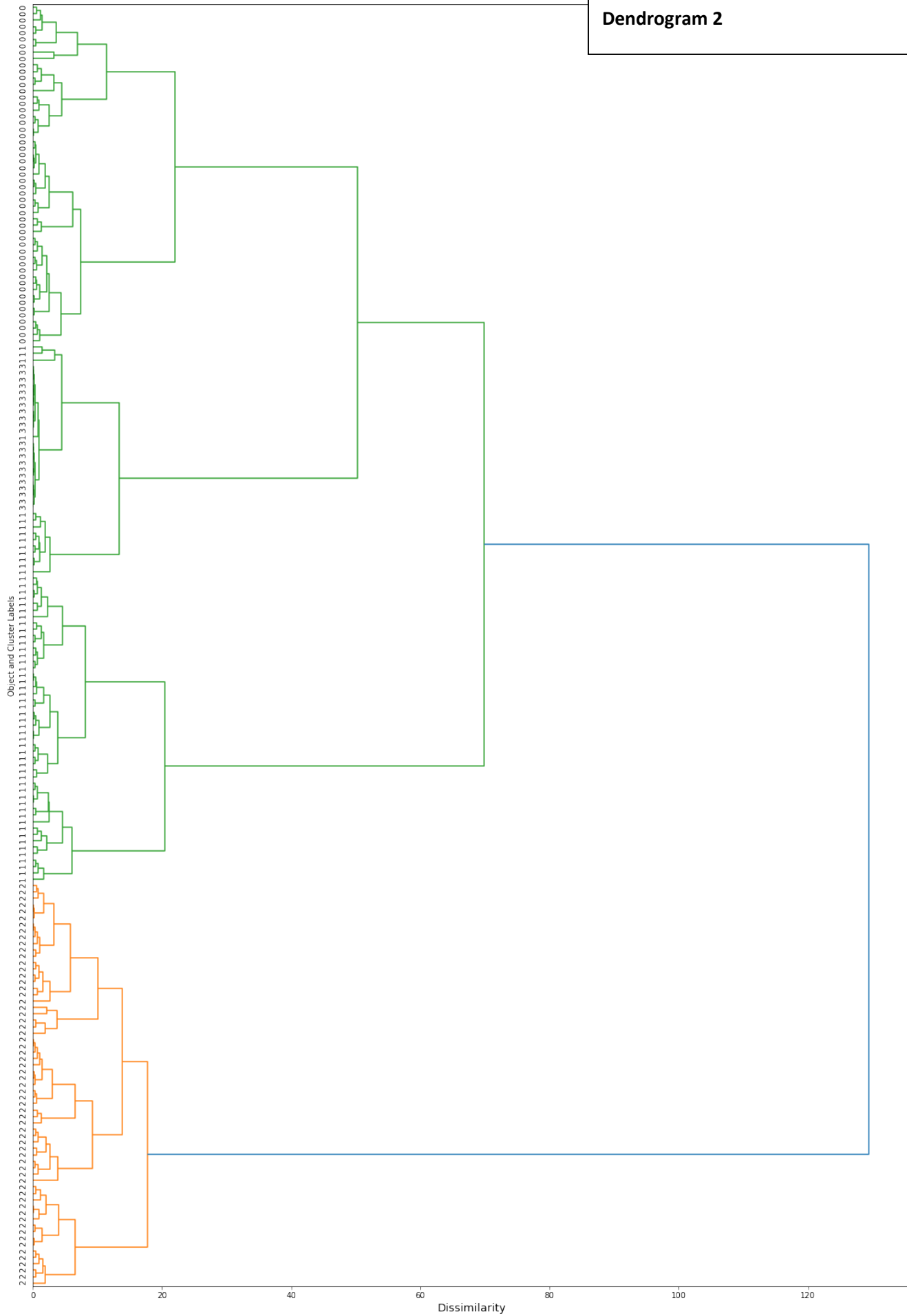


**12.2.** Which of these dendrograms successfully identified the 4 inherent clusters in this dataset (labeled by labels 0,1,2,3 above), if any?

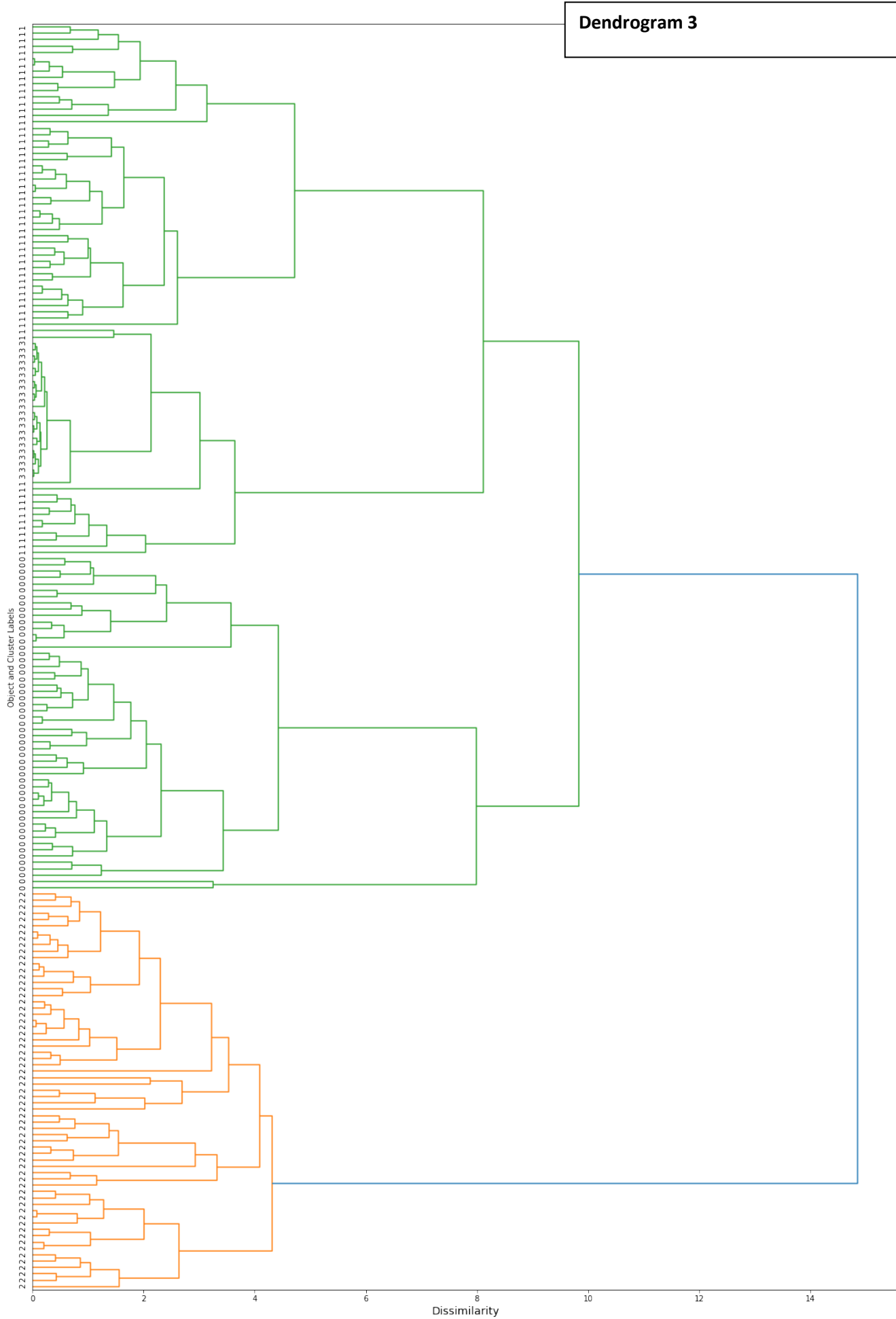
Dendrogram 1



## Dendrogram 2



### Dendrogram 3



### Dendrogram 4

