Ethics Pledge

Consistent with the above statements, all homework exercises, tests and exams that are designated as individual assignments MUST contain the following signed statement before they can be accepted for grading.

I pledge on my honor that I have not given or received any unauthorized assistance on this assignment/examination. I further pledge that I have not copied any material from a book, article, the Internet or any other source except where I have expressly cited the source.
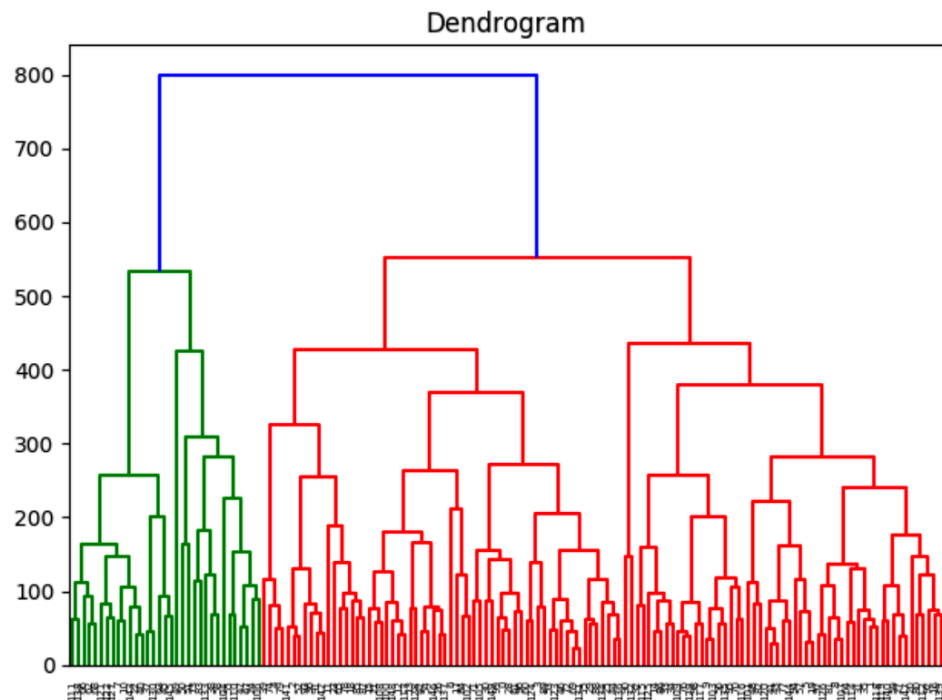
Signature: Haodong Zhao   Date: Mar 19th. 2019

Please note that assignments in this class may be submitted to www.turnitin.com, a web- based anti-plagiarism system, for an evaluation of their originality.

Question1: Program 16.6: For the family lung function data (Lung Function.xls **), perform a hierarchical cluster analysis using the variables FEV1 and FVC for mothers, fathers, and oldest children. Compare the distribution of area residence in the resulting groups.**

Answer:

By using 'complete' method:

Dendrogram


```
Following is Q1:

 Cluster    0   1   2  All
Area
1          13   8   3   24
2          19  14  16   49
3          10   7   2   19
4          20  26  12   58
All        62  55  33  150
```

From the result, we can find in cluster0, area4 has the most proportion and area3 has the least proportion. In cluster1, area4 occupy the most and area 3 occupy the least. And in cluster2, area2 has the most proportion and cluster3 has the least proportion.

Question2: Program 16.8: Repeat Problem 16.6, using the K-means method for K = 4. Compare the results with the four clusters produced in Problem 16.6.

Answer:

```
Following is Q2:

 Cluster    0   1   2   3  All
Area
1           5   9   7   3   24
2          10  12  16  11   49
3           4   7   6   2   19
4          13  24  12   9   58
All        32  52  41  25  150
```

By using k-means method and k = 4, we can get above result. We can find area4 has the most proportion in cluster0 and cluster1, and area2 has the most proportion in cluster2 and cluster3.