SADS-Problem-13
01-05-2024

The data on weight (kg) and height (cm) of 10 PG students are given below.

Weight	67 65 63 66 70 63 62 45 54 55 65
Height	167 166 165 166 168 162 156 153 160 152 163

Find the clusters of objects of the above data using agglomerative hierarchical clustering methods and also plot the dendrogram.

SOLUTION:-

Step:-

- Go to analyze
- Choose classify
- Select K- Means cluster
- Select Variable
- Iterative by Default & Save in Cluster membership & Distance from cluster center
- Go to options – Initial cluster centers & missing value Exclude cases listwise
- OK

LINK:-

```
QUICK CLUSTER Weight Heght
 /MISSING=LISTWISE
 /CRITERIA=CLUSTER(3) MXITER(10) CONVERGE(0)
 /METHOD=KMEANS(NOUPDATE)
 /SAVE CLUSTER
 /PRINT INITIAL CLUSTER DISTAN.
```

Initial Cluster Centers

	Cluster		
	1	2	3
Weight	70	45	62
Heght	168	153	156

Intial clusture

Devide in three cluster of 11 items

Iteration History[a]

Iteration	Change in Cluster Centers		
	1	2	3
1	4.549	.000	3.808
2	.695	.000	2.121
3	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 3. The minimum distance between initial centers is 14.422.

Itreration see the minumum distance

Cluster Membership

Case Number	Cluster	Distance
1	1	2.231
2	1	.915
3	1	2.587
4	1	.833
5	1	5.194
6	1	4.172
7	3	5.000
8	2	.000
9	3	5.000
10	3	4.472
11	1	2.356

In Cluster Membership in first cluster with 1,2,3,4,5,6 and 11

2nd only 8

3rd 7,9 and 10

Final Cluster Centers

	Cluster		
	1	2	3
Weight	66	45	57
Heght	165	153	156

#Final cluster centers

Distances between Final Cluster Centers

Cluster	1	2	3
1		23.961	12.637
2	23.961		12.369
3	12.637	12.369	

Number of Cases in each Cluster

Cluster	1	7.000
	2	1.000
	3	3.000
Valid		11.000
Missing		.000

Distances between Final Cluster Centers

Cluster	1	2	3
1		23.961	12.637
2	23.961		12.369
3	12.637	12.369	

Hierarchical Cluster

Steps :-

» Go to analyze and then classify

» Go to Hierarchical cluster select the variable.

» Go to statistics to choose agglomeration schedule and Proximity matrix use single solution of 3 cluster.

»Go to plot and set Dendrogram.

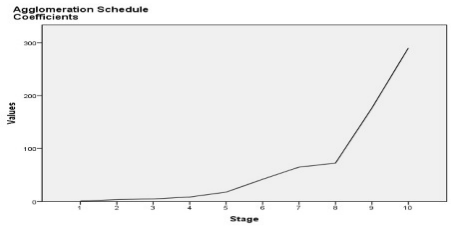
»Go to method and choose Between -groups linkage

»Square Euclidean distance & also in save for single solution.

Average Linkage (Between Groups)

Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	2	4	1.000	0	0	2
2	1	2	3.500	0	1	5
3	6	11	5.000	0	0	4
4	3	6	8.500	0	3	5
5	1	3	17.778	2	4	6
6	1	5	42.000	5	0	10
7	9	10	65.000	0	0	8
8	7	9	72.500	0	7	9
9	7	8	176.333	8	0	10
10	1	7	289.821	6	9	0



Cluster Membership

Case	3 Clusters
1	1
2	1
3	1
4	1
5	1
6	1
7	2
8	3
9	2
10	2
11	1

Quick Cluster

Case Processing Summary[a]

Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
11	100.0	0	.0	11	100.0

a. Average Linkage (Between Groups)

Proximity Matrix

Case	Squared Euclidean Distance								
	1	2	3	4	5	6	7	8	9
1	.000	5.000	20.000	2.000	10.000	41.000	146.000	680.000	218.000
2	5.000	.000	5.000	1.000	29.000	20.000	109.000	569.000	157.000
3	20.000	5.000	.000	10.000	58.000	9.000	82.000	468.000	106.000
4	2.000	1.000	10.000	.000	20.000	25.000	116.000	610.000	180.000
5	10.000	29.000	58.000	20.000	.000	85.000	208.000	850.000	320.000
6	41.000	20.000	9.000	25.000	85.000	.000	37.000	405.000	85.000
7	146.000	109.000	82.000	116.000	208.000	37.000	.000	298.000	80.000
8	680.000	569.000	468.000	610.000	850.000	405.000	298.000	.000	130.000
9	218.000	157.000	106.000	180.000	320.000	85.000	80.000	130.000	.000
10	369.000	296.000	233.000	317.000	481.000	164.000	65.000	101.000	65.000
11	20.000	9.000	8.000	10.000	50.000	5.000	58.000	500.000	130.000

This is a dissimilarity matrix

Number of Cases in each Cluster

Cluster	1	7.000
	2	1.000
	3	3.000
Valid		11.000
Missing		.000

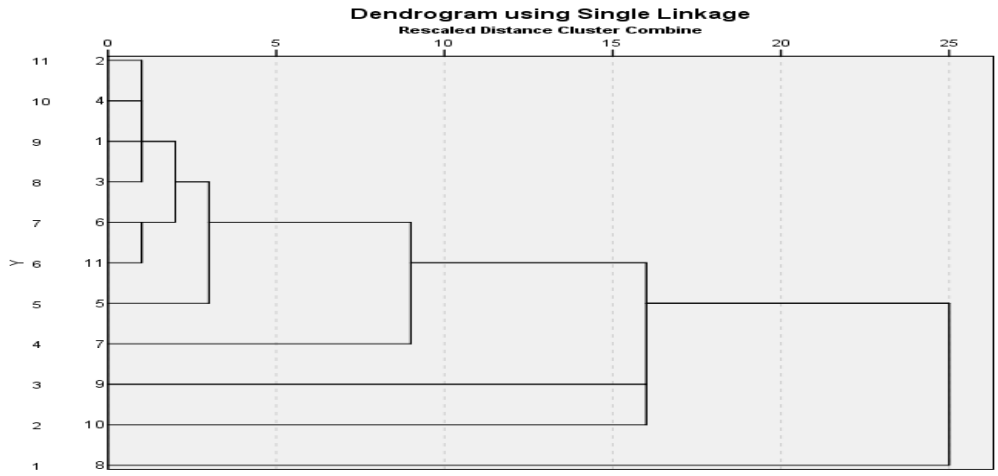
Number of Cases in each Cluster

Cluster	1	7.000
	2	1.000
	3	3.000
Valid		11.000
Missing		.000

Single Linkage

Cluster Membership

Case	3 Clusters
1	1
2	1
3	1
4	1
5	1
6	1
7	1
8	2
9	3
10	3
11	1



Conclusion:- Plot Dendrogram Average linkage method