Course: Data Analysis with SPSS (task №4)

Students' Name and Surname __Rashid Ali_____

Please do the task below.

1. Open **fact+clust.sav** file. Do the factor analysis. Interpret the factors and save them as new variables. Use saved factors for cluster analysis. Define the number of clusters. Describe the clusters' characteristics.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.812
Bartlett's Test of Sphericity	Approx. Chi-Square	176.560
	df	28
	Sig.	.000

KMO= 0,812>0,6

H0: All the correlations are not significant
H1: There are significant correlations

Sig=0,000 => H1 accepted.
So, these variables could be used for factor analysis.

Communalities

	Initial	Extraction
Percentage of urban population	1.000	.668
Average life expectancy for men	1.000	.769
Average life expectancy for women	1.000	.900
Infant mortality per 1,000 newborns	1.000	.872
Number of sunshine days per year	1.000	.876
Number of dull days per year	1.000	.731
The average daily temperature in January	1.000	.728
The average daily temperature in July	1.000	.808

Extraction Method: Principal Component Analysis.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.944	61.801	61.801	4.944	61.801	61.801
2	1.408	17.604	79.406	1.408	17.604	79.406
3	.698	8.721	88.127			
4	.389	4.868	92.994			
5	.193	2.414	95.408			
6	.175	2.187	97.595			
7	.123	1.536	99.131			
8	.069	.869	100.000			

Extraction Method: Principal Component Analysis.

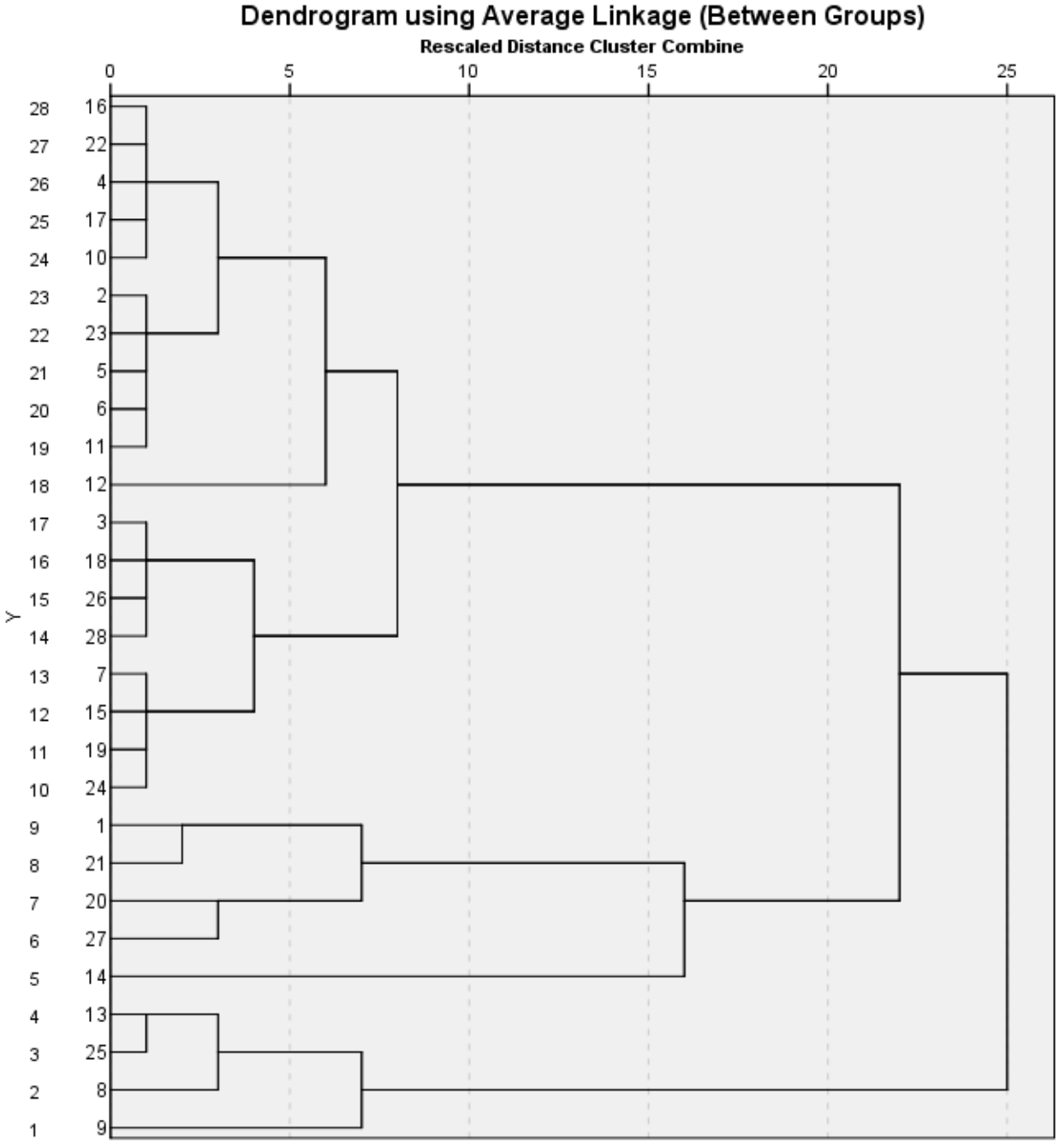
We have 2 factors and they explain 79,406% of the initial variables' variation.

Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	16	22	.015	0	0	8
2	2	23	.016	0	0	10
3	5	6	.018	0	0	5
4	4	17	.055	0	0	8
5	5	11	.085	3	0	10
6	3	18	.086	0	0	12
7	7	15	.108	0	0	15
8	4	16	.118	4	1	13
9	26	28	.129	0	0	12
10	2	5	.148	2	5	18
11	19	24	.164	0	0	15
12	3	26	.183	6	9	20
13	4	10	.228	8	0	18
14	13	25	.231	0	0	19
15	7	19	.254	7	11	20
16	1	21	.438	0	0	22
17	20	27	.645	0	0	22
18	2	4	.648	10	13	21
19	8	13	.810	0	14	23
20	3	7	.939	12	15	24
21	2	12	1.665	18	0	24
22	1	20	1.793	16	17	25
23	8	9	1.839	19	0	27
24	2	3	2.229	21	20	26
25	1	14	4.220	22	0	26
26	1	2	5.925	25	24	27
27	1	8	6.957	26	23	0

23-28= 5

So, there will be 5 clusters.


Dendrogram using Average Linkage (Between Groups)
Rescaled Distance Cluster Combine
0
5
10
15
20
25
Y
28 16
27 22
26 4
25 17
24 10
23 2
22 23
21 5
20 6
19 11
18 12
17 3
16 18
15 26
14 28
13 7
12 15
11 19
10 24
9 1
8 21
7 20
6 27
5 14
4 13
3 25
2 8
1 9