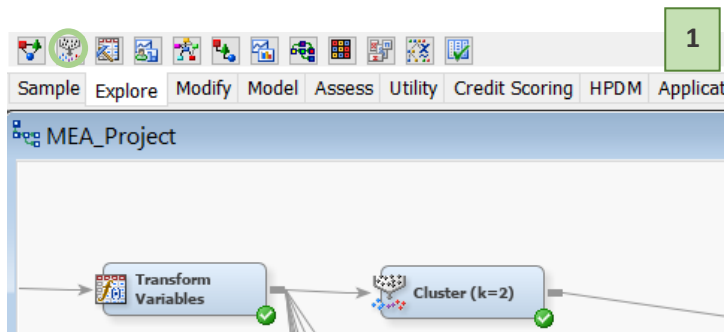
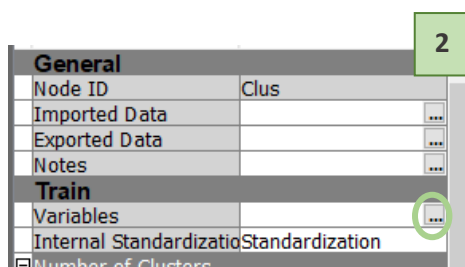


10 – Cluster Node (to perform cluster observation)

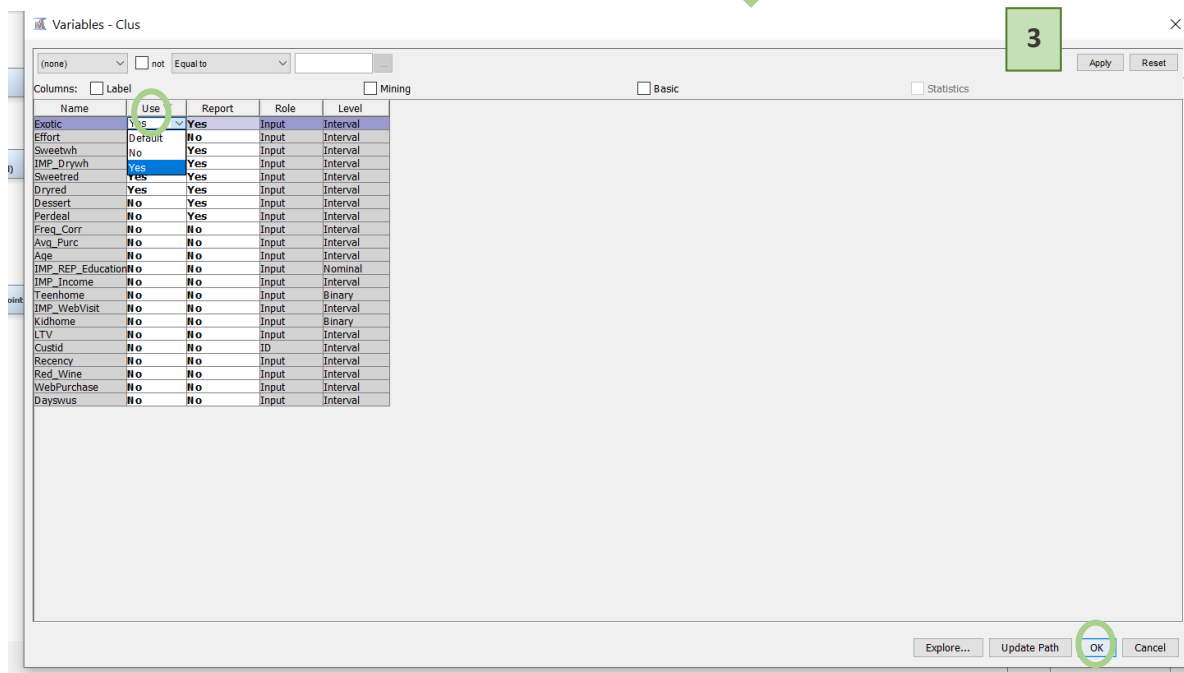


Go to the Explore tab and drag and drop the Cluster node. Connect it to the previous node.



The first thing that you have to do is to define the variables that will be included in the segmentation.

You do that clicking on the three dots option and a pop up window appears.



In the Use column by selecting YES or NO you are choosing/excluding the variables to consider in the segmentation.

4

Train	
Variables	
Internal Standardization	Standardization
Number of Clusters	
Specification Method	User Specify
Maximum Number of Clusters	2
Selection Criterion	
Clustering Method	Ward
Preliminary Maximum	50
Minimum	2
Final Maximum	20
CCC Cutoff	3
Encoding of Class Variables	
Ordinal Encoding	Rank
Nominal Encoding	GLM
Initial Cluster Seeds	
Seed Initialization Method	Princomp
Minimum Radius	0.0
Drift During Training	No
Training Options	
Use Defaults	Yes

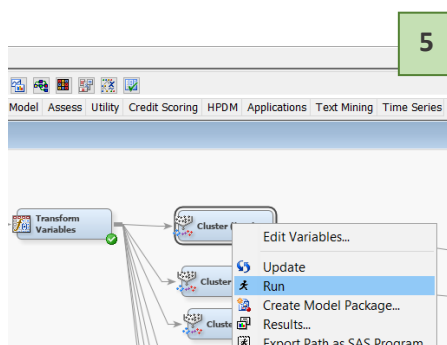
In the properties panel first you must guarantee that the Standardization options is defined for the Internal Standardization:

Internal Standardization - the variable values are divided by the standard deviation.

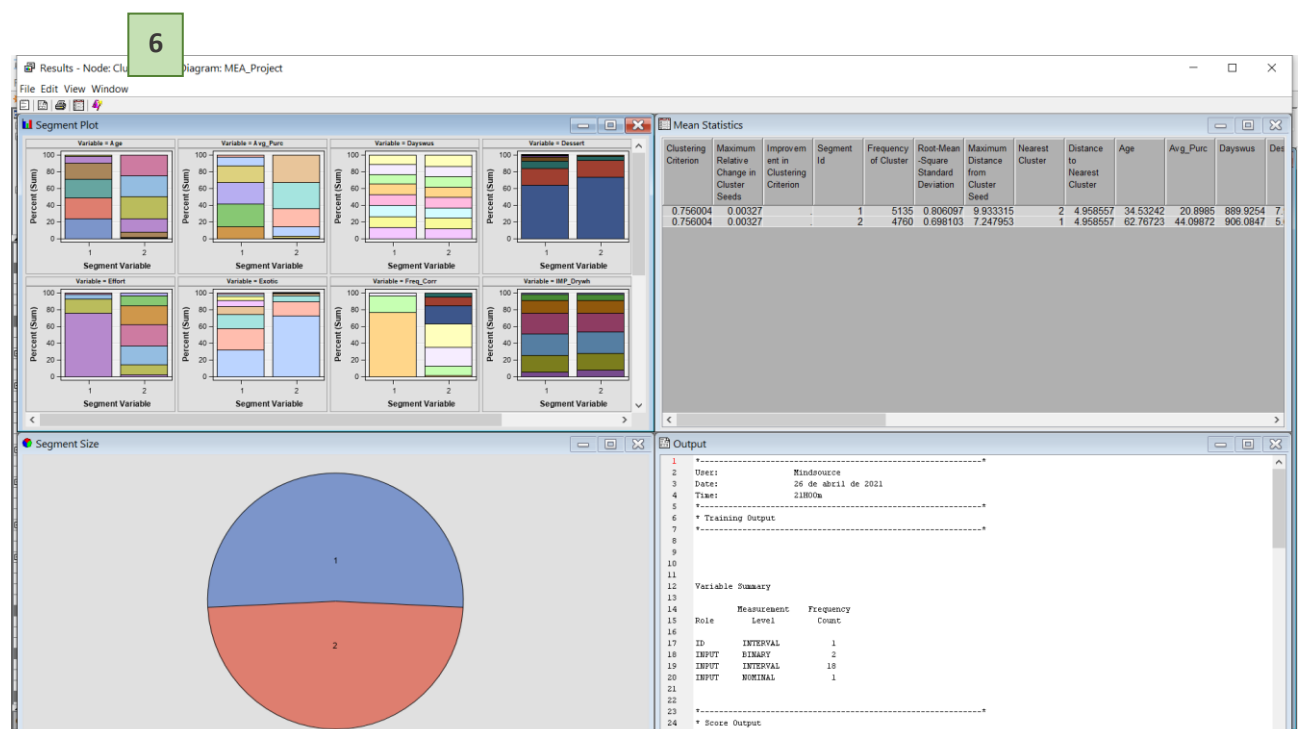
To define the **k-means method** you have to select User Specify in the Specification Method and for the Maximum Number of Clusters you must define the number of groups that you want.

In the Seed Initial Cluster Seeds option you must select Princomp:

Princomp: The principal components setting initializes seeds on an evenly spaced grid in the plane of the first two principal components.

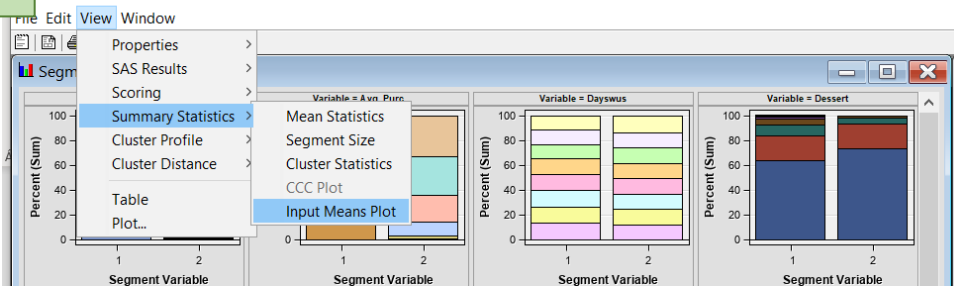


After running, check the results.

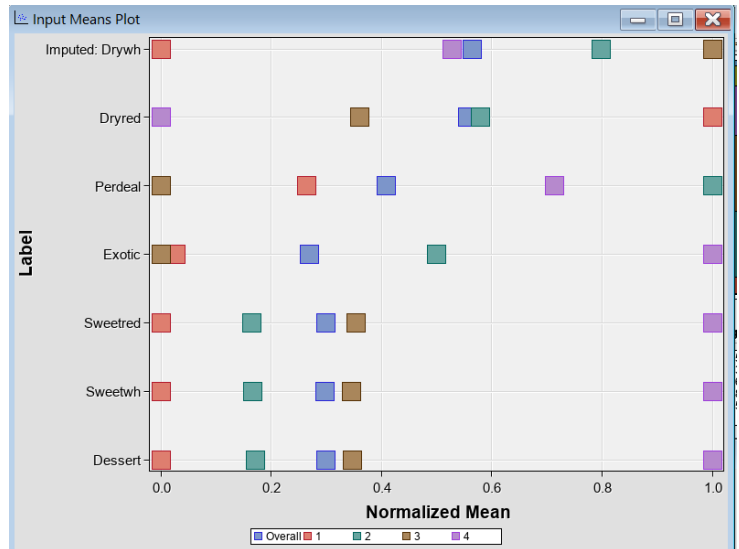


7

Results - Node: Cluster (k=2) Diagram: MEA_Project

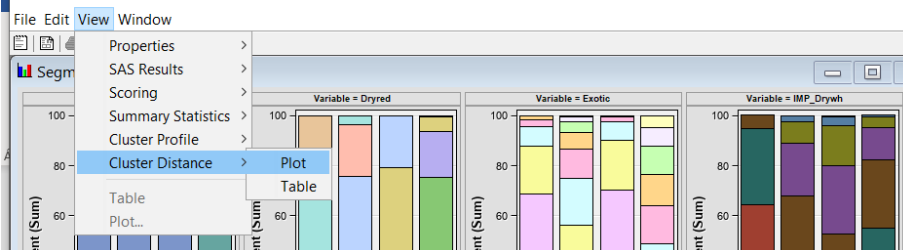


Input Means Plot: displays the normalized mean value for each variable, both inside each cluster and for the complete data set.

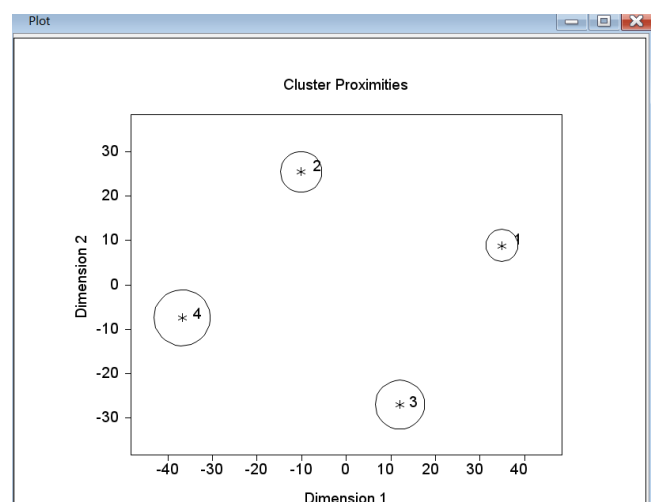


9

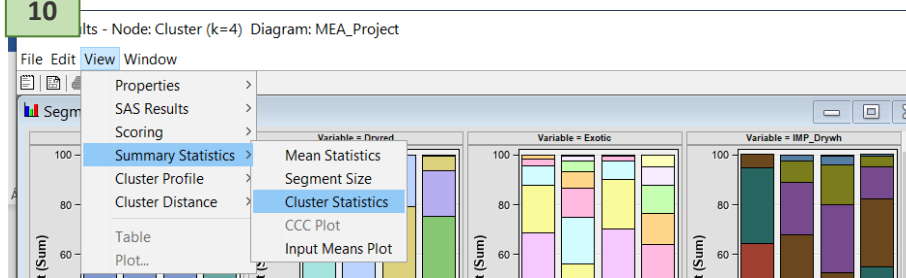
Results - Node: Cluster (k=4) Diagram: MEA_Project



Cluster Plot: provides a graphical representation of the size of each cluster and the relationship among clusters.



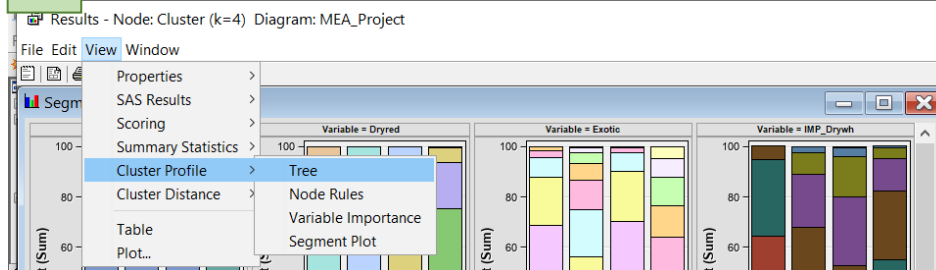
10



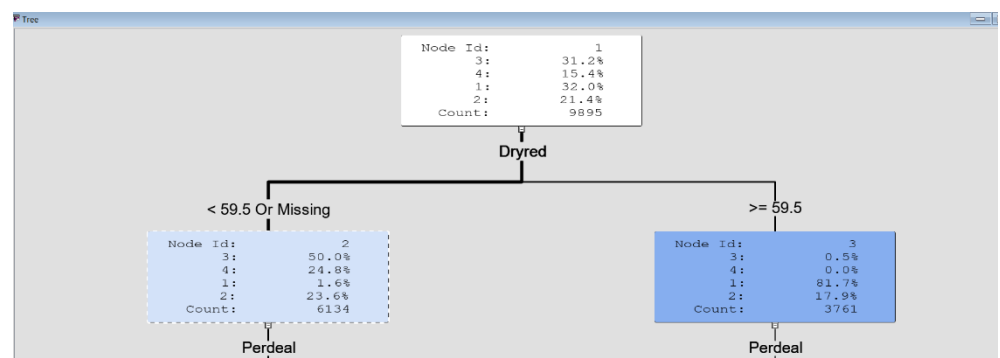
Cluster Statistics: contains summary columns for each input variable as well as columns for segment values.

Type of Observation	Segment Id	Statistic Applying Over All Variables	Dessert	Dryred	Exotic	Imputed: Drywh	Perdeal	Sweetred	Sweetwh
DMDB F...			9895	9895	9895	9895	9895	9895	9895
DMDB ...			9895	9895	9895	9895	9895	9895	9895
DMDB ...			6.857908	50.69955	16.07317	28.57672	32.20677	6.917938	6.922082
DMDB ...			7.725128	23.27379	16.34082	12.57697	27.83239	7.561575	7.694522
LOCATI...			6.857908	50.69955	16.07317	28.57672	32.20677	6.917938	6.922082
SCALE ...			7.725128	23.27379	16.34082	12.57697	27.83239	7.561575	7.694522
DMDB ...			0	1	0	1	0	0	0
DMDB ...			62	99	83	74	97	48	49
CRITERI...		0.710462							
PSEUD...		3236.748							
ERSQ		0.258912							
CCC		235.8686							
TOTAL ...		1	7.725128	23.27379	16.34082	12.57697	27.83239	7.561575	7.694522
WITHIN ...		0.710467	5.776967	11.62861	13.09689	9.246225	18.71076	5.554842	5.706537
RSQ		0.495389	0.440942	0.750431	0.35782	0.459687	0.548196	0.460505	0.450142
RSQ RA...		0.981725	0.788724	3.006915	0.557195	0.85078	1.213349	0.853586	0.818652
SEED	1	3115	2.194222	76.2565	9.45297	16.95148	23.76918	2.262921	2.276083

11



Tree: shows the decision tree that was used to form the individual clusters. The decision tree is based on your clustered data.

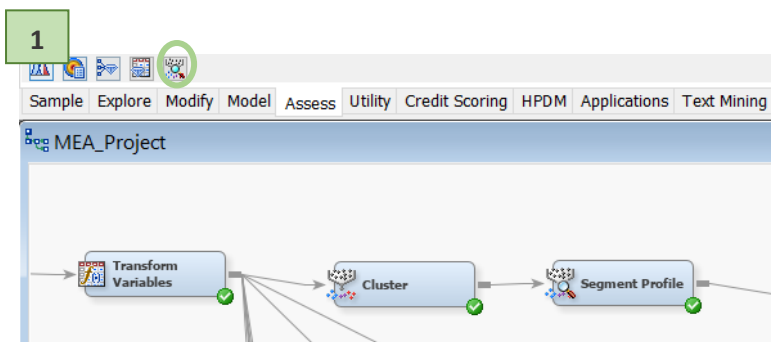


In the properties panel it is also possible to rename each segment name. Just click on the three dots option and a pop up window appears.

Score		12
Cluster Variable Role	Segment	
Hide Original Variables	Yes	
Cluster Label Editor		
Report		

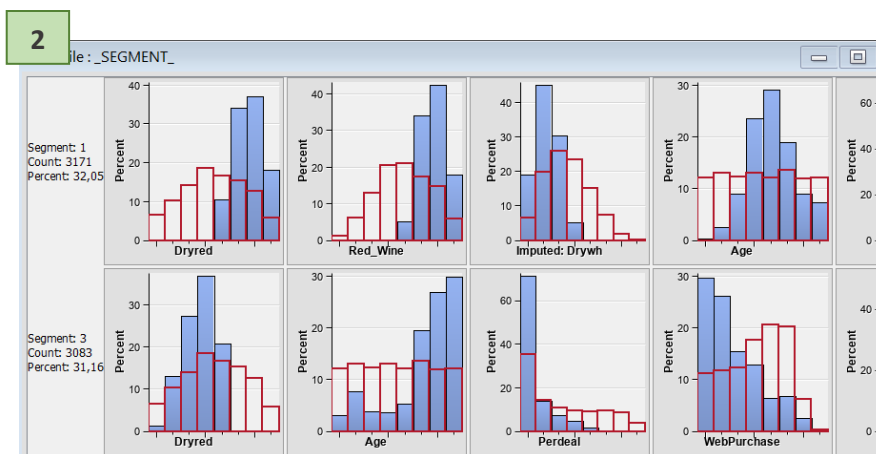
Segment Label	Segment Name	Dryred
Dryred lovers	1	76.19047619
	2	51.996222852
	3	39.432695427
	4	18.629678267

11 – Segment Profile (analyse clustered data)



Go to the Assess tab and drag and drop the Segment Profile node. Connect it to the previous node and run it.

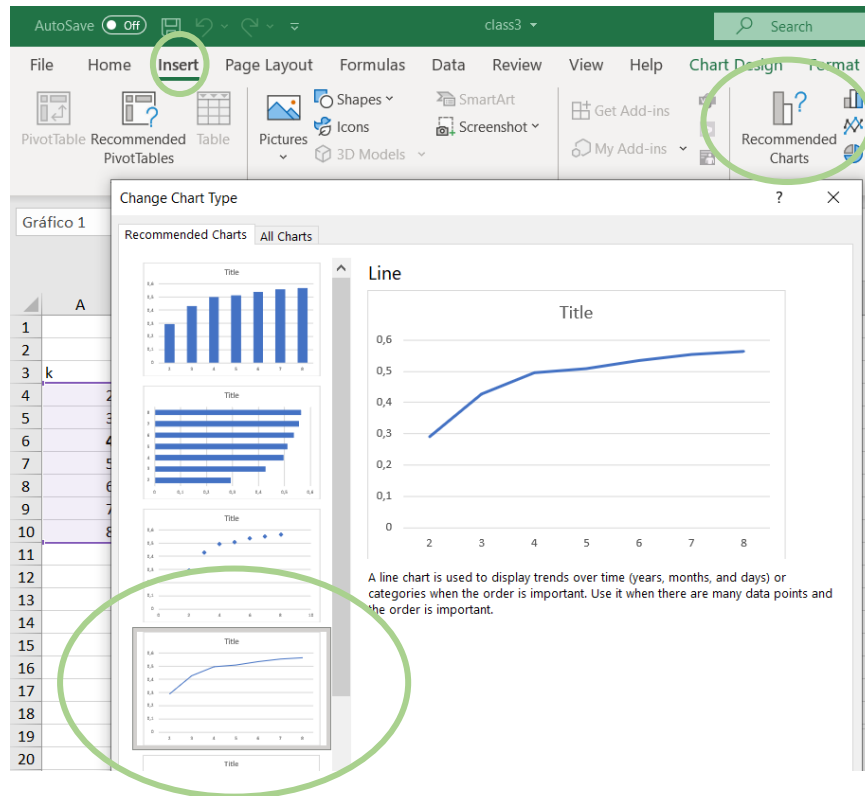
After execution, check the results.



Profile SEGMENT :The Profile window displays plots comparing the distribution for the identified and report variables for both the segment and the population.

EXTRA - EXCEL

Graphic 1:



Graphic 2:

