

Ανάλυση ταξινόμησης

Cluster analysis

Γούργουλης Βασίλειος, Ph.D.
Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

1

Ταξινόμηση:

τεχνική ομαδοποίησης των **ατόμων** ενός συνόλου



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

2

Ταξινόμηση:

Δεν είναι γνωστός (εκ των προτέρων) ο αριθμός των ομάδων
και **ποια άτομα ανήκουν** σε κάθε ομάδα



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

3

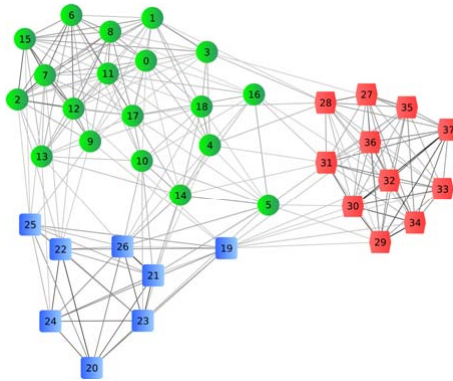
Ταξινόμηση (ομαδοποίηση) των ατόμων:
τα άτομα που βρίσκονται «κοντά» το ένα στο άλλο
στον αρχικό χώρο των μεταβλητών
θα ομαδοποιηθούν (ταξινομηθούν) στις **ίδιες κλάσεις**



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

4

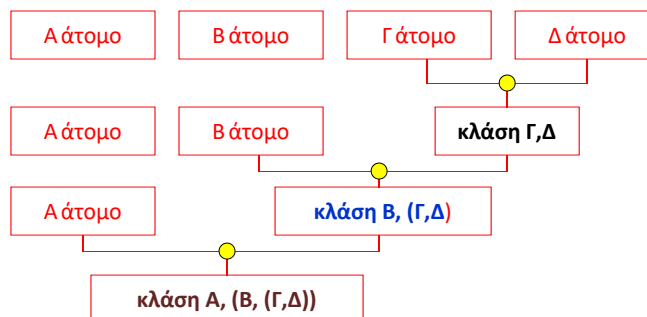
Ταξινόμηση (ομαδοποίηση) των ατόμων:
 τα άτομα που βρίσκονται «κοντά» το ένα στο άλλο
 στον αρχικό χώρο των μεταβλητών
 θα ομαδοποιηθούν (ταξινομηθούν) στις **ίδιες κλάσεις**



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

5

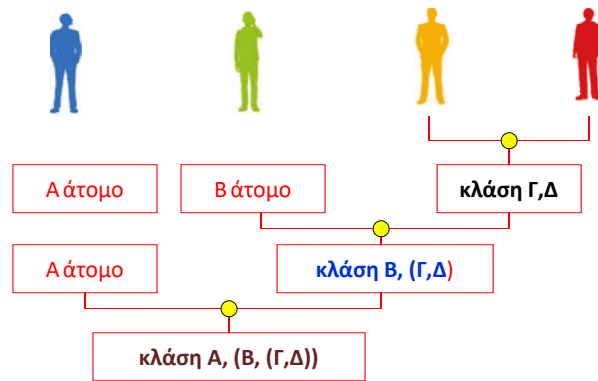
Ταξινόμηση (ομαδοποίηση) των ατόμων:
 τα άτομα που βρίσκονται «κοντά» το ένα στο άλλο
 στον αρχικό χώρο των μεταβλητών
 θα ομαδοποιηθούν (ταξινομηθούν) στις **ίδιες κλάσεις**



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

6

Ταξινόμηση (ομαδοποίηση) των ατόμων:
 τα άτομα που βρίσκονται «κοντά» το ένα στο άλλο
 στον αρχικό χώρο των μεταβλητών
 θα ομαδοποιηθούν (ταξινομηθούν) στις **ίδιες κλάσεις**



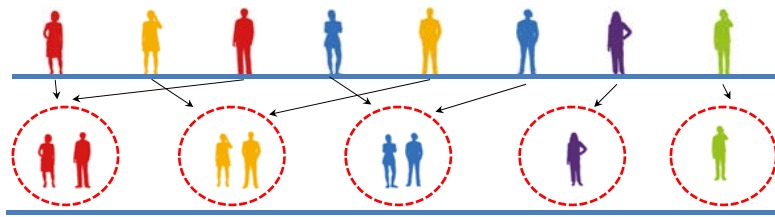
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

7



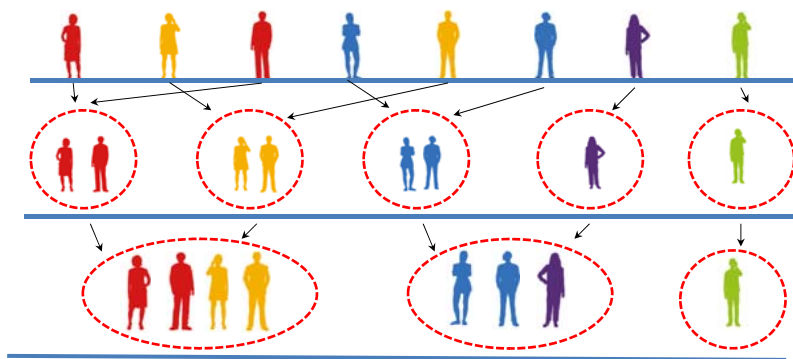
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

8



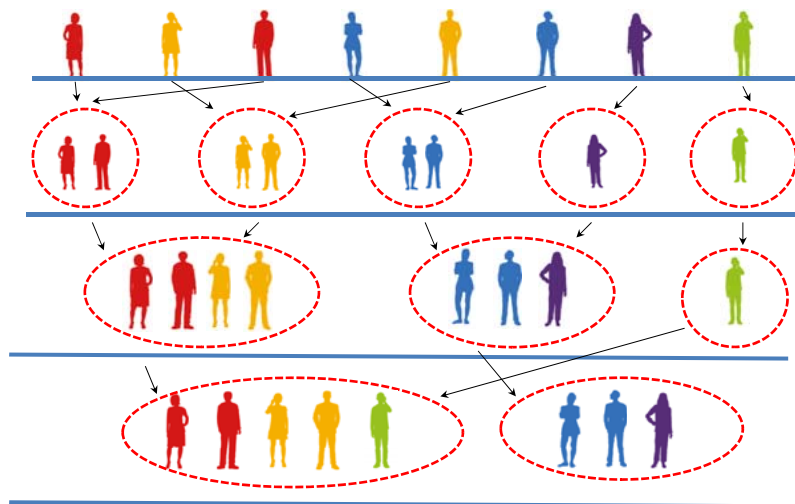
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

9



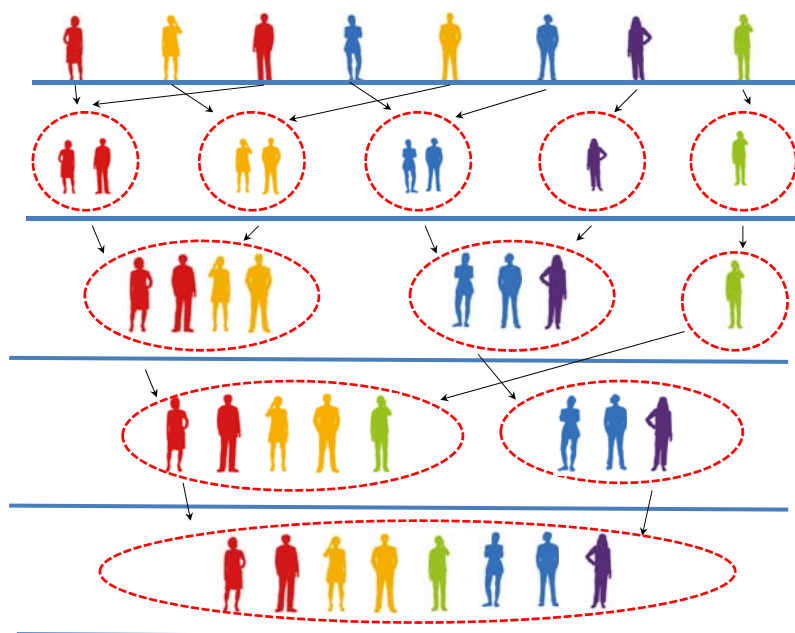
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

10



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

11

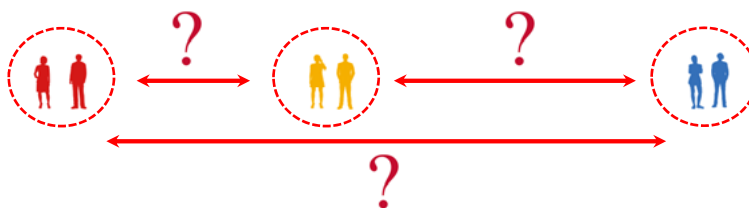


Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

12

τι σημαίνει «**κοντά**»; ?

για την ένταξη (ομαδοποίηση) δύο ατόμων στην ίδια κλάση
ή για την ένωση δύο κλάσεων σε μια νέα κλάση;



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

13

Ποιος θα πρέπει να είναι ο **αριθμός**
των **νέων κλάσεων** που θα προκύψουν; ?

Βάση ποιου κριτηρίου θα γίνει η επιλογή; ?

Ποια μέθοδος ταξινόμησης θα επιλεγεί; ?

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

14 14

Μέθοδοι ταξινόμησης



«ιεραρχημένη» ... (**hierarchical**):

- ξεκινώντας από τις ειδικές καταλήγουμε σε γενικότερες κλάσεις
- δεν είναι προαποφασισμένος ο αριθμός των κλάσεων
- όλες ποσοτικές μεταβλητές
- σχετικά μικρός αριθμός δεδομένων (< 250 άτομα)



«μη ιεραρχημένη - διαμερισματική» ... (**k - means**):

- προαποφασισμένος αριθμός των κλάσεων
(εκ των προτέρων επιλογή των κέντρων των κλάσεων και όχι σχηματισμός τους από τις μικρότερες προς τις μεγαλύτερες ή αντίστροφα)
- χρησιμοποιείται σε πολύ μεγάλο αριθμό δεδομένων (> 1000 άτομα)



«ταξινόμηση 2 βημάτων» ... (**Two Steps cluster**):

- ποσοτικές και ποιοτικές μεταβλητές

15 15

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

Μέθοδοι ταξινόμησης



Μετά την εφαρμογή της **ιεραρχημένης** ανάλυσης ταξινόμησης και αφού καθοριστεί ο επιθυμητός αριθμός κλάσεων (ο οποίος **δεν είναι προαποφασισμένος**)

ενδεχομένως

ο ερευνητής να θέλει να αναλύσει όλα τα δεδομένα με **μη ιεραρχημένη ανάλυση ταξινόμησης (k – means)** με **προκαθορισμένο** αριθμό κλάσεων

16 16

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

Μοντέλο ανάλυσης ιεραρχημένης ταξινόμησης

άτομο	X_1	X_2	X_3	...	X_j	...	X_p
A_1	X_{11}	X_{12}	X_{13}	...	X_{1j}	...	X_{1p}
A_2	X_{21}	X_{22}	X_{23}	...	X_{2j}	...	X_{2p}
A_3	X_{31}	X_{32}	X_{33}	...	X_{3j}	...	X_{3p}
...	...	...	...	...	...	...	...
A_N	X_{N1}	X_{N2}	X_{N3}	...	X_{nj}	...	X_{np}

X_j = ποσοτικές μεταβλητές (χαρακτηριστικά)

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

17

Αρχικά δεδομένων

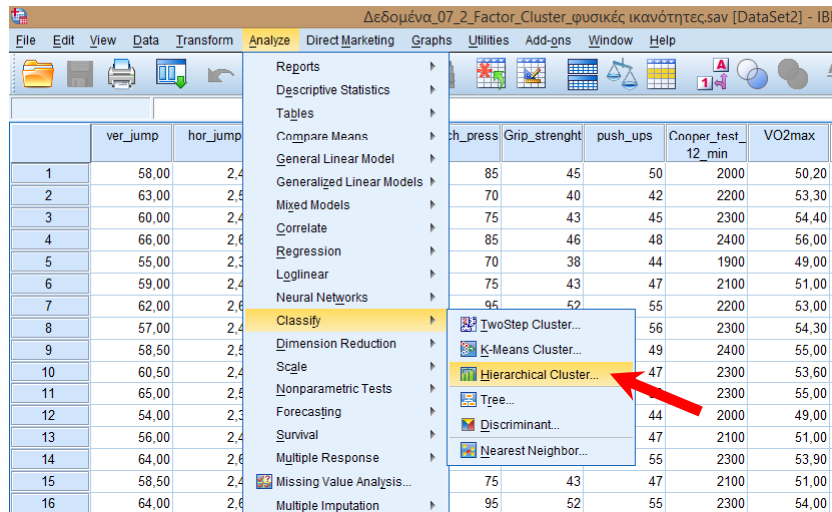


	ver_jump	hor_jump	m30	squat	Bench_press	Grip_strength	push_ups	Cooper_test_12_min	VO2max
1	58,00	2,40	4,90	32	85	45	50	2000	50,20
2	63,00	2,55	4,75	33	70	40	42	2200	53,30
3	60,00	2,45	4,90	32	75	43	45	2300	54,40
4	66,00	2,60	4,55	35	85	46	48	2400	56,00
5	55,00	2,35	5,15	29	70	38	44	1900	49,00
6	58,00	2,45	5,05	31	75	43	47	2100	51,00
7	65,00	2,60	4,70	35	95	52	55	2200	53,00
8	57,00	2,40	4,85	32	90	53	56	2300	54,30
9	58,00	2,50	4,90	33	80	42	49	2400	55,00
10	60,00	2,45	4,90	32	85	44	47	2300	53,60
11	65,00	2,55	4,60	36	80	45	50	2300	55,00
12	54,00	2,35	5,05	30	70	39	44	2000	49,00
13	58,00	2,45	5,05	31	75	40	47	2100	51,00
14	64,00	2,60	4,80	34	95	55	55	2300	53,90
15	58,00	2,45	5,05	31	75	43	47	2100	51,00
16	64,00	2,60	4,60	34	95	52	55	2300	54,00
17	57,00	2,40	4,85	32	90	53	56	2300	54,30
18	59,00	2,50	4,90	33	80	42	49	2400	55,00
19	61,00	2,50	4,95	33	85	44	47	2200	53,60
20	65,00	2,55	4,60	36	80	45	50	2300	55,00

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

18

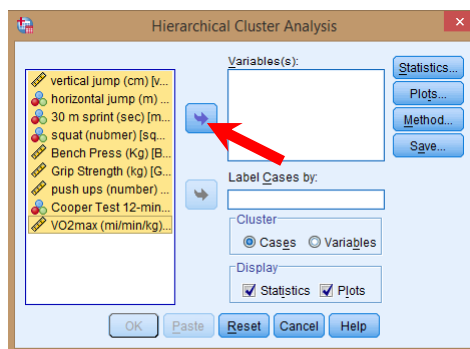
Analyze
Classify
Hierarchical Cluster ...



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

19

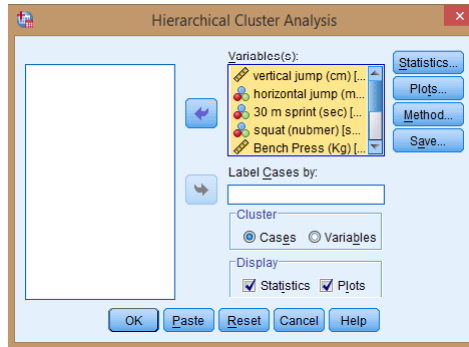
Καθορισμός μεταβλητών
βάσει των οποίων θα πραγματοποιηθεί
η ομαδοποίηση των ατόμων



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

20

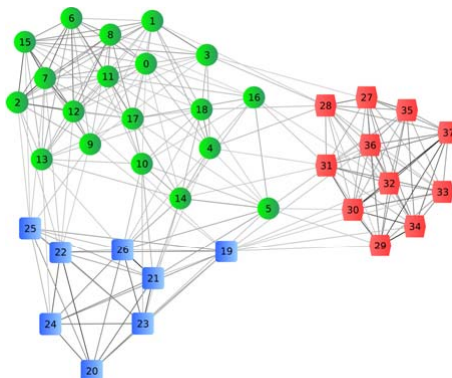
Καθορισμός μεταβλητών
βάσει των οποίων θα πραγματοποιηθεί
η ομαδοποίηση των ατόμων



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

21

τα άτομα που βρίσκονται «**κοντά**» το ένα στο άλλο
στον αρχικό χώρο των μεταβλητών
θα **ομαδοποιηθούν** (ταξινομηθούν) στις **ίδιες κλάσεις**



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

22

τι σημαίνει «**κοντά**»;



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

23

τι σημαίνει «**κοντά**»;



Υπολογισμός των «αποστάσεων» μεταξύ των ατόμων



Ευκλείδεια απόσταση: $\gamma = \sqrt{\alpha^2 + \beta^2}$
(*Euclidean distance*)

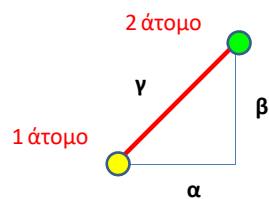


Τετράγωνο Ευκλείδειας απόστασης: $\gamma^2 = \alpha^2 + \beta^2$
(*Squared Euclidean distance*)

(απομακρύνει το πρόσημο και
αυξάνει την επίδραση των ακραίων τιμών)



... Κ.Α.



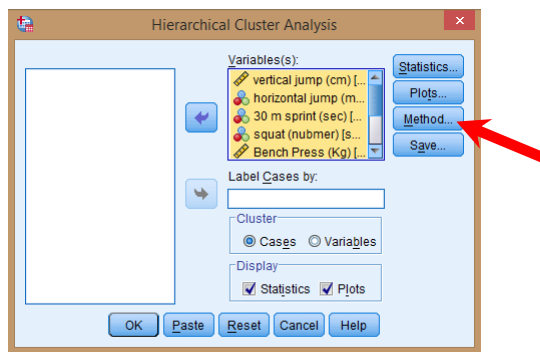
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

24

τι σημαίνει «**κοντά**»;



Υπολογισμός των «αποστάσεων» μεταξύ των ατόμων



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

25

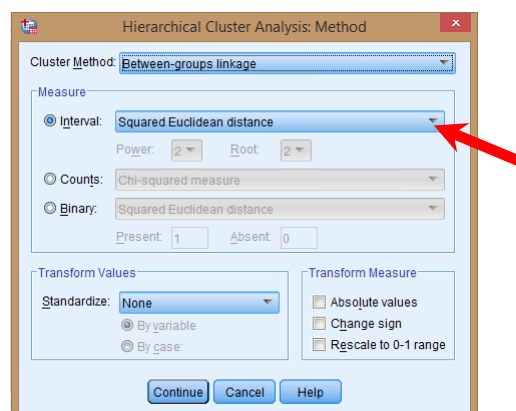
τι σημαίνει «**κοντά**»;



Υπολογισμός των «αποστάσεων» μεταξύ των ατόμων

Measure

Interval



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

26

τι σημαίνει «**κοντά**»;

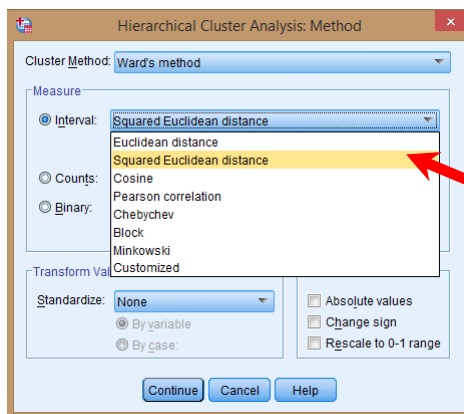


Υπολογισμός των «αποστάσεων» μεταξύ των ατόμων

Measure

Interval

**Squared
Euclidian distance**



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

27

τι σημαίνει «**κοντά**»;

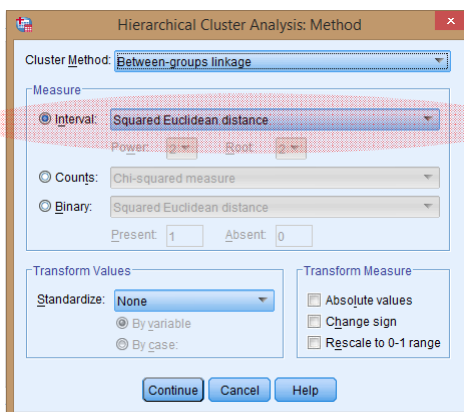


Υπολογισμός των «αποστάσεων» μεταξύ των ατόμων

Measure

Interval

**Squared
Euclidian distance**



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

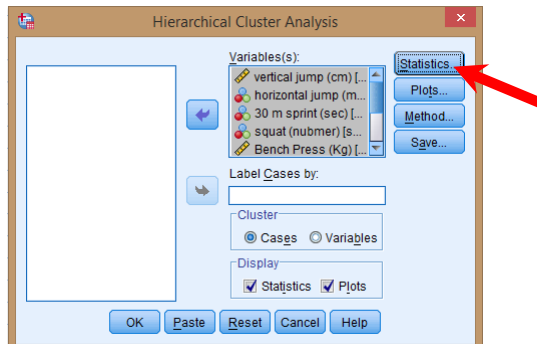
28

τι σημαίνει «**κοντά**»;



Ποιες είναι οι «αποστάσεις» μεταξύ των ατόμων
βάσει της επιλεγμένης μεθόδου (... "Τετράγωνο Ευκλείδειας απόστασης"...);

Statistics



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

29

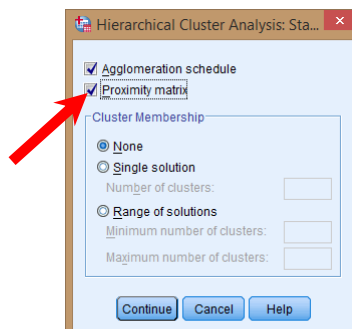
τι σημαίνει «**κοντά**»;



Ποιες είναι οι «αποστάσεις» μεταξύ των ατόμων
βάσει της επιλεγμένης μεθόδου (... "Τετράγωνο Ευκλείδειας απόστασης"...);

Statistics

Proximity matrix



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

30

τι σημαίνει «κοντά»;



Ποιες είναι οι «αποστάσεις» μεταξύ των ατόμων
βάσει της επιλεγμένης μεθόδου (... “Τετράγωνο Ευκλείδειας απόστασης”...);

Proximity matrix

City	Proximity Matrix																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0.00	40349.655	90102.643	102111.803	10326.505	10115.685	40208.820	90742.813	100059.300	90207.813	90113.153	918.465	10138.685	90276.740	10114.915	90228.510	90142.813	100061.300	40031.572	90120.483
2	40349.655	0.00	10004.242	40017.333	90106.690	10084.390	943.595	10003.033	40176.165	10273.373	10204.912	40113.820	10108.390	11021.365	10086.640	10040.515	10003.033	40168.165	270.330	10207.142
3	90102.643	10004.242	0.00	10165.705	10014.233	40017.583	10084.023	405.015	10045.613	105.860	95.460	90111.163	40041.583	664.283	40018.833	601.272	405.015	10040.513	10107.645	10017.910
4	102111.803	40017.333	10165.705	0.00	20011.423	90206.273	40210.023	10231.020	103.363	10050.165	10033.005	10008.313	10236.473	90207.523	10194.003	10231.020	90.363	40039.900	10032.295	10032.295
5	10326.505	90106.690	10014.233	20011.423	0.00	40003.020	91042.265	100015.163	200205.335	100306.483	100370.343	10003.010	40047.020	101165.165	40079.270	101073.360	100015.163	200213.335	100342.233	100302.593
6	10115.685	10004.390	90017.583	90206.273	40003.020	0.00	10074.145	40421.933	90050.275	40111.033	40115.213	10000.010	10.000	40050.495	.250	40088.205	40421.933	90050.275	10115.773	40121.463
7	40208.820	943.595	10084.023	40210.023	91042.265	10074.145	.000	10082.753	40391.300	10239.673	10313.013	41028.195	10084.145	10014.820	10077.365	10008.010	10082.753	40371.300	233.433	10116.263
8	90742.813	10003.033	405.015	10231.020	10014.163	40421.933	10082.753	0.00	10272.753	190.740	200.575	90701.132	40447.903	93.202	40426.163	901.160	0.00	10272.753	10204.910	206.825
9	100059.300	40176.165	10045.613	103.363	200205.335	90050.275	40081.300	10039.963	0.00	10039.963	10081.343	90701.945	90050.275	10042.480	90050.025	10393.350	10273.753	1.000	40041.213	10088.093
10	90207.813	10273.373	105.860	10005.165	100306.483	40111.033	10239.673	190.740	10039.963	0.00	73.310	90328.443	40144.033	301.373	40112.783	244.523	190.740	10039.963	10001.255	76.960
11	90113.153	10204.912	95.460	10023.005	100370.343	40115.213	10313.013	200.575	10091.343	73.310	0.00	90301.243	40101.213	306.253	40121.463	305.003	200.575	10044.343	10062.065	.250
12	918.465	40113.820	90111.163	10108.390	10003.010	10000.010	41028.195	90701.132	100196.295	90306.443	90306.243	0.00	91142.135	10075.260	91056.265	90701.132	100209.265	40038.163	90176.493	90176.493
13	10138.685	10108.390	40041.583	90276.740	10014.233	40047.020	10.000	10044.145	40487.933	90059.275	40144.033	40181.213	10044.010	0.00	40770.495	15.250	40040.225	40487.933	90065.275	10151.773
14	90276.740	11021.365	664.283	10236.473	101073.360	101073.360	40487.933	90059.275	10042.480	301.373	206.253	91142.135	10075.260	91056.265	90701.132	100209.265	40038.163	90176.493	90176.493	90176.493
15	10114.915	10086.640	40018.833	90207.523	40079.270	.250	10077.365	40426.163	90050.025	40112.783	40121.463	10075.260	15.250	40055.745	.000	40059.475	.000	40059.475	10116.023	10116.023
16	90228.510	10040.515	601.272	10194.003	101073.360	40081.300	40081.300	40081.300	40081.300	244.523	305.003	91056.265	40055.745	9.000	40049.475	.000	40049.475	.000	40049.475	10116.023
17	90142.813	10003.033	405.015	10231.020	10014.163	40421.933	10082.753	0.00	10272.753	190.740	200.575	90701.132	40447.903	93.202	40426.163	901.160	0.00	10272.753	10204.910	206.825
18	100061.300	40176.165	10045.613	103.363	200205.335	90050.275	40081.300	10039.963	0.00	10039.963	10081.343	90701.945	90050.275	10042.480	90050.025	10393.350	10273.753	1.000	40041.213	10088.093
19	40031.572	270.330	10107.645	10032.295	100306.483	40115.773	233.433	10204.910	40041.213	10001.255	10002.085	40338.163	10151.773	10296.123	10118.023	10236.263	10204.910	40037.213	.000	10066.335
20	90120.483	10207.142	10017.910	10032.295	100306.483	40121.463	10316.263	206.825	10088.093	76.960	.250	90376.493	40190.493	307.503	40126.213	309.253	206.825	10055.093	10066.335	.000

Πηγή: W. B. Bonaventura et al.

31

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

τι σημαίνει «κοντά»;



Ποιες είναι οι «αποστάσεις» μεταξύ των ατόμων
βάσει της επιλεγμένης μεθόδου (... “Τετράγωνο Ευκλείδειας απόστασης”...);

Proximity matrix

City	Proximity Matrix																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0.00	40349.655	90102.643	102111.803	10326.505	10115.685	40208.820	90742.813	100059.300	90207.813	90113.153	918.465	10138.685	90276.740	10114.915	90228.510	90142.813	100061.300	40031.572	90120.483
2	40349.655	0.00	10004.242	40017.333	90106.690	10084.390	943.595	10003.033	40176.165	10273.373	10204.912	40113.820	10108.390	11021.365	10086.640	10040.515	10003.033	40168.165	270.330	10207.142
3	90102.643	10004.242	0.00	10165.705	10014.233	40017.583	10084.023	405.015	10045.613	105.860	95.460	90111.163	40041.583	664.283	40018.833	601.272	405.015	10040.513	10107.645	10017.910
4	102111.803	40017.333	10165.705	0.00	20011.423	90206.273	40210.023	10231.020	103.363	10050.165	10033.005	10008.313	10236.473	90207.523	10194.003	10231.020	90.363	40039.900	10032.295	10032.295
5	10326.505	90106.690	10014.233	20011.423	0.00	40003.020	91042.265	100015.163	200205.335	100306.483	100370.343	10003.010	40047.020	101165.165	40079.270	101073.360	100015.163	200213.335	100342.233	100302.593
6	10115.685	10004.390	90017.583	90206.273	40003.020	0.00	10074.145	40421.933	90050.275	40111.033	40115.213	10000.010	10.000	40050.495	.250	40088.205	40421.933	90050.275	10115.773	40121.463
7	40208.820	943.595	10084.023	40210.023	91042.265	10074.145	.000	10082.753	40391.300	10239.673	10313.013	41028.195	10084.145	10014.820	10077.365	10008.010	10082.753	40371.300	233.433	10116.263
8	90742.813	10003.033	405.015	10231.020	10014.163	40421.933	10082.753	0.00	10272.753	190.740	200.575	90701.132	40447.903	93.202	40426.163	901.160	0.00	10272.753	10204.910	206.825
9	100059.300	40176.165	10045.613	103.363	200205.335	90050.275	40081.300	10039.963	0.00	10039.963	10081.343	90701.945	90050.275	10042.480	90050.025	10393.350	10273.753	1.000	40041.213	10088.093
10	90207.813	10273.373	105.860	10005.165	100306.483	40111.033	10239.673	190.740	10039.963	0.00	73.310	90328.443	40144.033	301.373	40112.783	244.523	190.740	10039.963	10001.255	76.960
11	90113.153	10204.912	95.460	10023.005	100370.343	40115.213	10313.013	200.575	10091.343	73.310	0.00	90301.243	40101.213	306.253	40121.463	305.003	200.575	10044.343	10062.065	.250
12	918.465	40113.820	90111.163	10108.390	10003.010	10000.010	41028.195	90701.132	100196.295	90306.443	90306.243	0.00	91142.135	10075.260	91056.265	90701.132	100209.265	40038.163	90176.493	90176.493
13	10138.685	10108.390	40041.583	90276.740	10014.233	40047.020	10.000	10044.145	40487.933	90059.275	40144.033	40181.213	10044.010	0.00	40770.495	15.250	40040.225	40487.933	90065.275	10151.773
14	90276.740	11021.365	664.283	10236.473	101073.360	101073.360	40487.933	90059.275	10042.480	301.373	206.253	91142.135	10075.260	91056.265	90701.132	100209.265	40038.163	90176.493	90176.493	90176.493
15	10114.915	10086.640	40018.833	90207.523	40079.270	.250	10077.365	40426.163	90050.025	40112.783	40121.463	10075.260	15.250	40055.745	.000	40059.475	.000	40059.475	10116.023	10116.023
16	90228.510	10040.515	601.272	10194.003	101073.360	40081.300	40081.300	40081.300	40081.300	244.523	305.003	91056.265	40055.745	9.000	40049.475	.000	40049.475	.000	40049.475	10116.023
17	90142.813	10003.033	405.015	10231.020	10014.163	40421.933	10082.753	0.00	10272.753	190.740	200.575	90701.132	40447.903	93.202	40426.163	901.160	0.00	10272.753	10204.910	206.825
18	100061.300	40176.165	10045.613	103.363	200205.335	90050.275	40081.300	10039.963	0.00	10039.963	10081.343	90701.945	90050.275	10042.480	90050.025	10393.350	10273.753	1.000	40041.213	10088.093
19	40031.572	270.330	10107.645	10032.295	100306.483	40115.773	233.433	10204.910	40041.213	10001.255	10002.085	40338.163	10151.773	10296.123	10118.023	10236.263	10204.910	40037.213	.000	10066.335
20	90120.483	10207.142	10017.910	10032.295	100306.483	40121.463	10316.263	206.825	10088.093	76.960	.250	90376.493	40190.493	307.503	40126.213	309.253	206.825	10055.093	10066.335	.000

Ποια άτομα θα ενταχθούν στην ίδια κλάση;

Ποια άτομα ή κλάσεις θα ενωθούν σε κάθε βήμα;



Ward's method

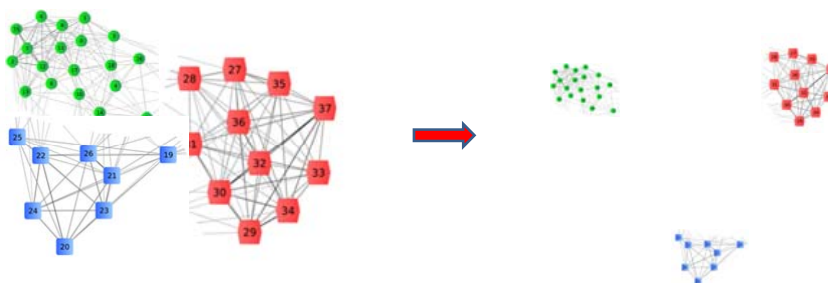
- 👉 **ελαχιστοποίηση** της «**εντός**» της κλάσης διακύμανσης και
- 👉 **αύξηση** της «**μεταξύ**» των κλάσεων διακύμανσης

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

33

ελαχιστοποίηση της διακύμανσης της νέας κλάσης
που προκύπτει από την ένωση δύο άλλων κλάσεων

- 👉 **ελαχιστοποίηση** της «**εντός**» της κλάσης διακύμανσης και
- 👉 **αύξηση** της «**μεταξύ**» των κλάσεων διακύμανσης

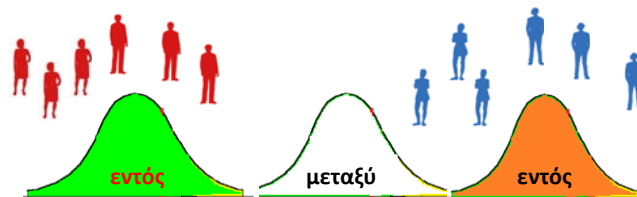


Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

34

ελαχιστοποίηση της διακύμανσης της νέας κλάσης
που προκύπτει από την ένωση δύο άλλων κλάσεων

-  ελαχιστοποίηση της «εντός» της κλάσης διακύμανσης
- και
-  αύξηση της «μεταξύ» των κλάσεων διακύμανσης

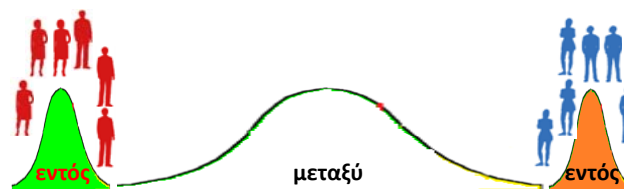


Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

35

ελαχιστοποίηση της διακύμανσης της νέας κλάσης
που προκύπτει από την ένωση δύο άλλων κλάσεων

-  ελαχιστοποίηση της «εντός» της κλάσης διακύμανσης
- και
-  αύξηση της «μεταξύ» των κλάσεων διακύμανσης

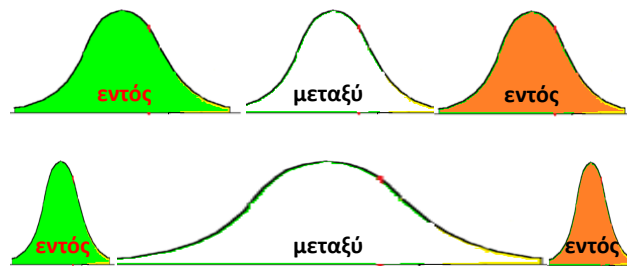


Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

36

ελαχιστοποίηση της διακύμανσης της νέας κλάσης
που προκύπτει από την ένωση δύο άλλων κλάσεων

-  ελαχιστοποίηση της «εντός» της κλάσης διακύμανσης και
-  αύξηση της «μεταξύ» των κλάσεων διακύμανσης



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

37



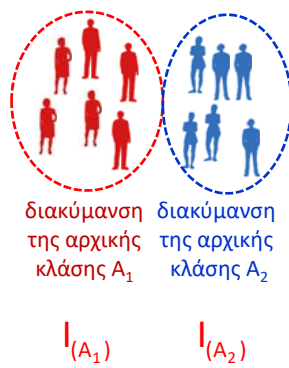
διακύμανση
της αρχικής
κλάσης A_1

$I_{(A_1)}$

Κατά την ένωση

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

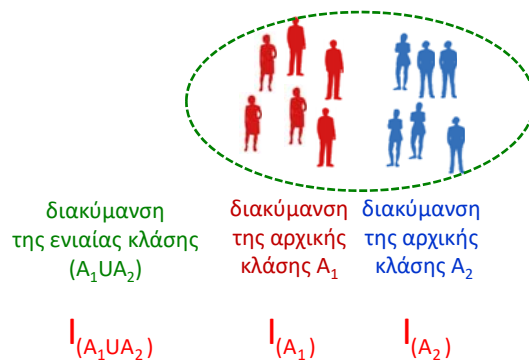
38



Κατά την ένωση δύο κλάσεων

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

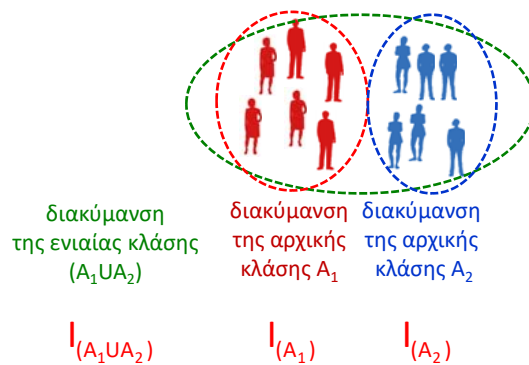
39



Κατά την ένωση δύο κλάσεων
η διακύμανση (αδράνεια) της ενιαίας κλάσης $(A_1 \cup A_2)$

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

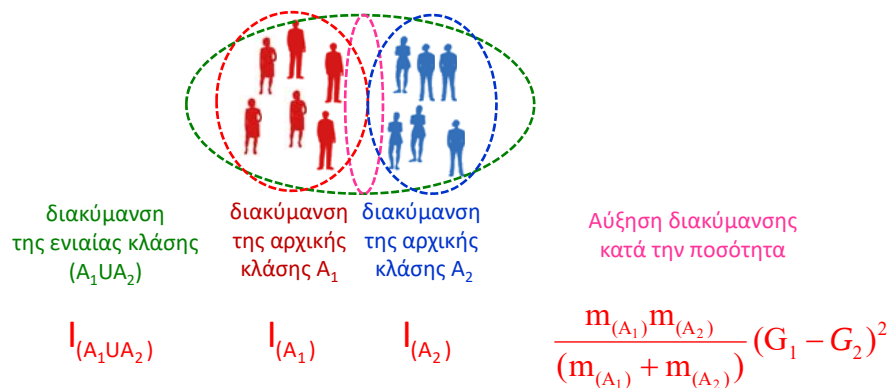
40



Κατά την ένωση δύο κλάσεων
 η διακύμανση (αδράνεια) της ενιαίας κλάσης (A_1UA_2)
 εκτός από το άθροισμα των διακυμάνσεων των δύο αρχικών μεμονωμένων κλάσεων

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

41



Κατά την ένωση δύο κλάσεων
 η διακύμανση (αδράνεια) της ενιαίας κλάσης (A_1UA_2)
 εκτός από το άθροισμα των διακυμάνσεων των δύο αρχικών μεμονωμένων κλάσεων
αυξάνεται και κατά την ποσότητα ...

όπου

G: ο μέσος όρος της κάθε κλάσης

m: ο αριθμός των ατόμων της κάθε κλάσης

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

42

$$I_{(A_1 \cup A_2)} = I_{(A_1)} + I_{(A_2)} + \frac{m_{(A_1)} \cdot m_{(A_2)}}{(m_{(A_1)} + m_{(A_2)})} (G_1 - G_2)^2$$

διακύμανση της ενιαίας κλάσης ($A_1 \cup A_2$) διακύμανση της αρχικής κλάσης A_1 διακύμανση της αρχικής κλάσης A_2 **Κριτήριο του Ward**

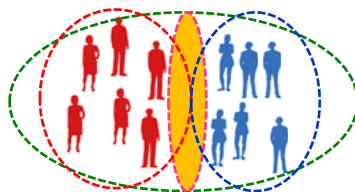
Κατά την ένωση δύο κλάσεων
 η διακύμανση (αδράνεια) της ενιαίας κλάσης ($A_1 \cup A_2$)
 εκτός από το άθροισμα των διακυμάνσεων των δύο αρχικών μεμονωμένων κλάσεων
αυξάνεται και κατά την ποσότητα ...

$$\frac{m_{(A_1)} m_{(A_2)}}{(m_{(A_1)} + m_{(A_2)})} (G_1 - G_2)^2 \quad (\text{Κριτήριο του Ward})$$

43

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

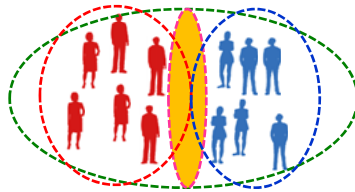
Δηλαδή
 αυξάνεται η διακύμανση «**εντός**» (**within**) της νέας κλάσης
 η οποία «χάνεται» από την διακύμανση «**μεταξύ**» (**between**) των αρχικών κλάσεων



44

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

Δηλαδή
 αυξάνεται η διακύμανση «**εντός**» (**within**) της νέας κλάσης
 η οποία «χάνεται» από την διακύμανση «**μεταξύ**» (**between**) των αρχικών κλάσεων



Κατά την ομαδοποίηση
 επιθυμούμε
 να χάνεται το λιγότερο δυνατόν
 από την «**μεταξύ**» των κλάσεων διακύμανση,
 ώστε να είναι καλά διαχωρισμένες.

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

45

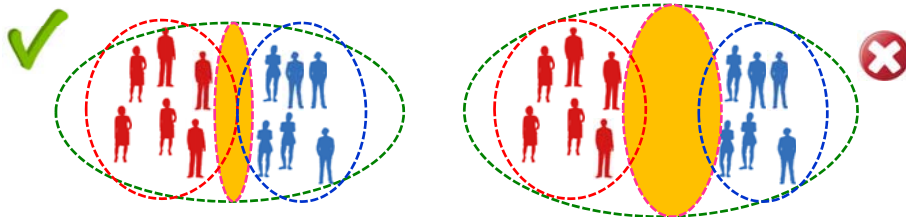
Άρα
 η μείωση της «**μεταξύ**» των κλάσεων διακύμανσης,
 θα πρέπει να είναι όσο το δυνατόν **μικρότερη**.



Το κριτήριο του Ward

$$\frac{m_{(A_1)} m_{(A_2)}}{(m_{(A_1)} + m_{(A_2)})} (G_1 - G_2)^2$$

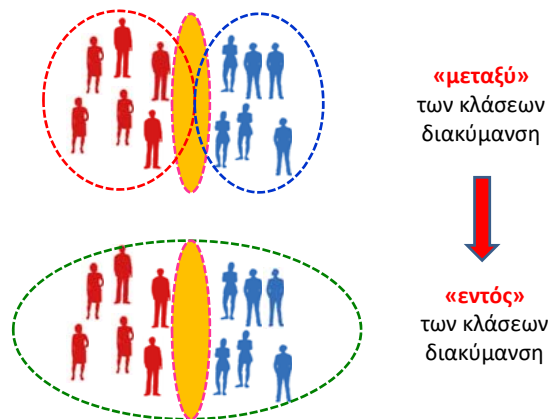
θα πρέπει να έχει την **ελάχιστη** τιμή του για να ενωθούν δύο κλάσεις



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

46

Καθ' όλη τη διάρκεια της ταξινόμησης,
δηλαδή του σχηματισμού των κλάσεων,
πραγματοποιείται ένας **μετασχηματισμός**
της «**μεταξύ**» των κλάσεων διακύμανσης (αδράνειας),
σε «**εντός**» των κλάσεων διακύμανση (αδράνεια).

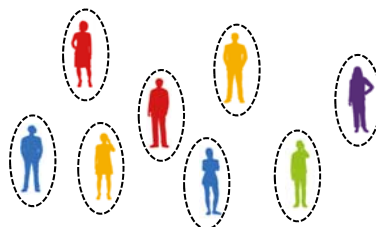


Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

47

Καθ' όλη τη διάρκεια της ταξινόμησης,
δηλαδή του σχηματισμού των κλάσεων,
πραγματοποιείται ένας **μετασχηματισμός**
της «**μεταξύ**» των κλάσεων διακύμανσης (αδράνειας),
σε «**εντός**» των κλάσεων διακύμανση (αδράνεια).

Αρχικά
η **συνολική διακύμανση** του νέφους των ατόμων
(εφόσον το κάθε άτομο θεωρείται ξεχωριστή κλάση)
ισούται με την «**μεταξύ**» των κλάσεων διακύμανση.



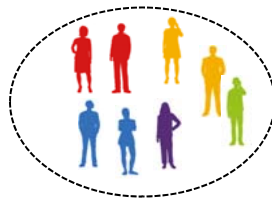
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

48

Καθ' όλη τη διάρκεια της ταξινόμησης,
δηλαδή του σχηματισμού των κλάσεων,
πραγματοποιείται ένας **μετασχηματισμός**
της «**μεταξύ**» των κλάσεων διακύμανσης (αδράνειας),
σε «**εντός**» των κλάσεων διακύμανση (αδράνεια).

Στο τέλος

(εφόσον υπάρχει μόνο μία κλάση)
η **συνολική διακύμανση** του νέφους των ατόμων
ισούται με την «**εντός**» των κλάσεων διακύμανση.



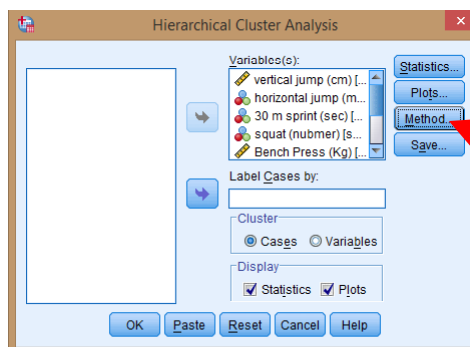
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

49

Ποια άτομα θα ενταχθούν στην **ίδια κλάση**;



Ποια άτομα ή κλάσεις θα **ενωθούν** σε κάθε βήμα;

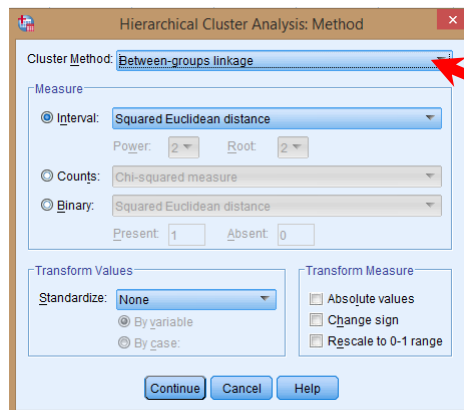


Cluster method

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

50

Ποια άτομα θα ενταχθούν στην ίδια κλάση;
 Ποια άτομα ή κλάσεις θα ενωθούν σε κάθε βήμα;

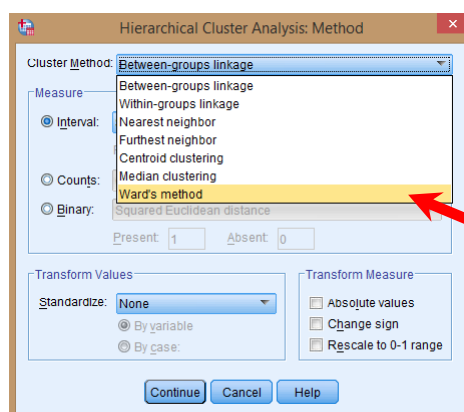


Cluster method

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

51

Ποια άτομα θα ενταχθούν στην ίδια κλάση;
 Ποια άτομα ή κλάσεις θα ενωθούν σε κάθε βήμα;



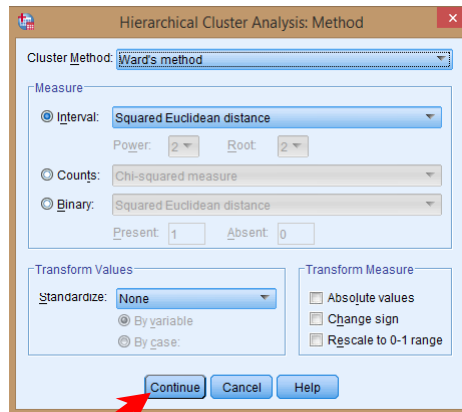
Cluster method

Ward's method

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

52

Ποια άτομα θα ενταχθούν στην **ίδια κλάση**;
 Ποια άτομα ή κλάσεις **θα ενωθούν** σε κάθε βήμα;



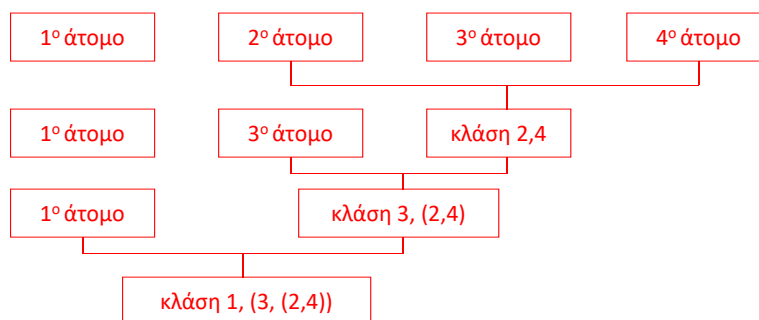
Cluster method
Ward's method

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

53

Αρχικά

η **συνολική διακύμανση** του νέφους των ατόμων
 (εφόσον το κάθε άτομο θεωρείται ξεχωριστή κλάση)
 ισούται με την **«μεταξύ»** των κλάσεων διακύμανση.



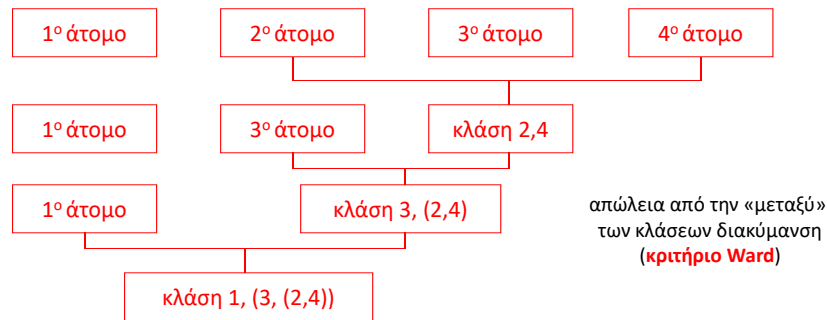
Στο τέλος

(εφόσον υπάρχει μόνο μία κλάση)
 η **συνολική διακύμανση** του νέφους των ατόμων
 ισούται με την **«εντός»** των κλάσεων διακύμανση.

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

54

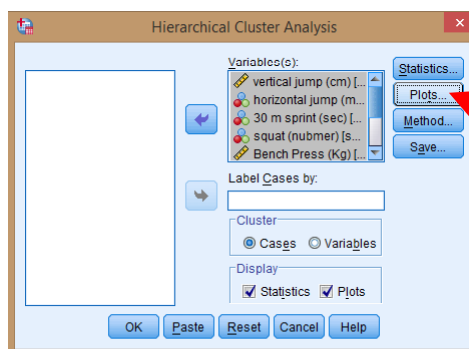
Κάθε φορά που σχηματίζεται μια νέα κλάση έχουμε **απώλεια μιας ποσότητας από την «μεταξύ»** των κλάσεων διακύμανση. Το άθροισμα των διαδοχικών απωλειών ισούται με την συνολική διακύμανση.



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

55

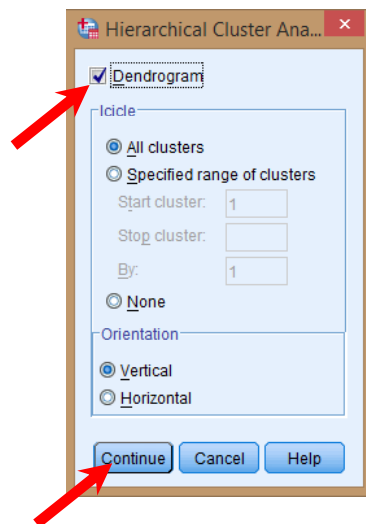
Απεικόνιση κλάσεων



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

56

Απεικόνιση κλάσεων

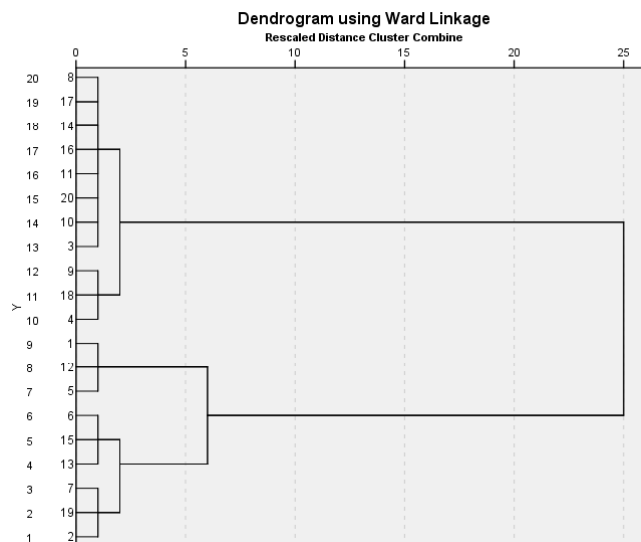


Δενδρόγραμμα

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

57

Απεικόνιση κλάσεων - Δενδρόγραμμα



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

58

Ποιος θα πρέπει να είναι ο αριθμός
των νέων κλάσεων που θα προκύψουν;

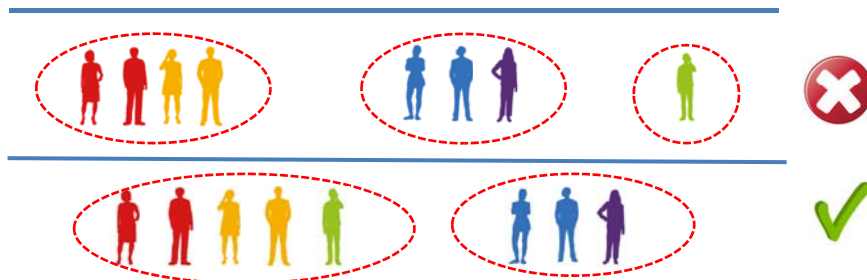


- ☞ • Όλες οι κλάσεις θα πρέπει να έχουν ικανοποιητικό αριθμό περιπτώσεων (ατόμων)
- ☞ • Μία ή περισσότερες πολύ μικρές κλάσεις (με μικρό αριθμό περιπτώσεων - ατόμων) δηλώνουν ότι έχουν επιλεγεί πολλές κλάσεις (περισσότερες από ότι θα έπρεπε)

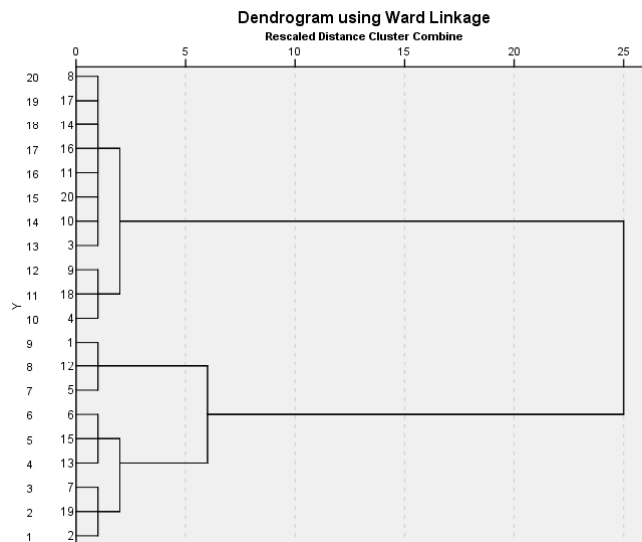
Ποιος θα πρέπει να είναι ο αριθμός
των νέων κλάσεων που θα προκύψουν;



- ☞ • Όλες οι κλάσεις θα πρέπει να έχουν ικανοποιητικό αριθμό περιπτώσεων (ατόμων)
- ☞ • Μία ή περισσότερες πολύ μικρές κλάσεις (με μικρό αριθμό περιπτώσεων - ατόμων) δηλώνουν ότι έχουν επιλεγεί πολλές κλάσεις (περισσότερες από ότι θα έπρεπε)



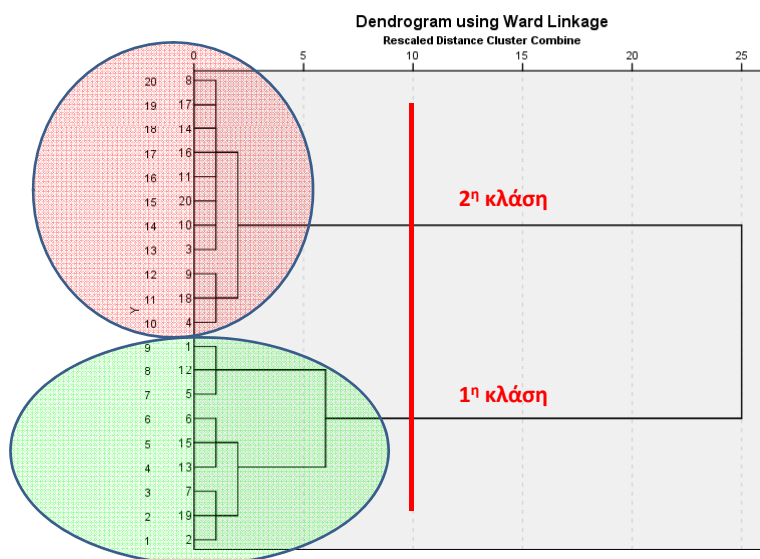
Απεικόνιση κλάσεων - Δενδρόγραμμα



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

61

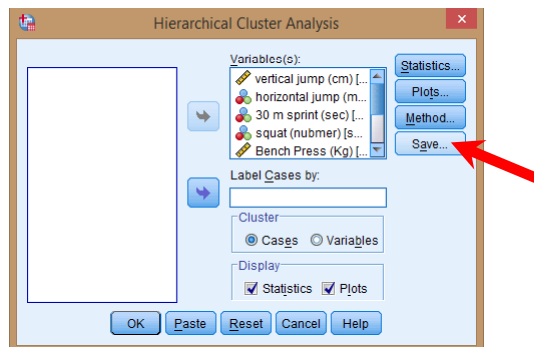
Απεικόνιση κλάσεων - Δενδρόγραμμα



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

62

Ένταξη των ατόμων σε κλάσεις



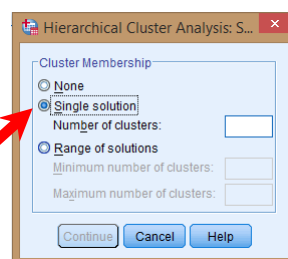
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

63

Ένταξη των ατόμων σε κλάσεις

Single solution

Number of clusters
(καθορισμός
του αριθμού
των κλάσεων)



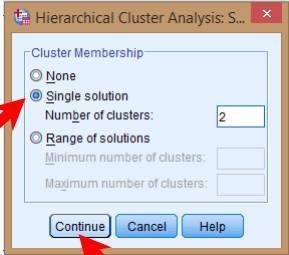
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

64

Ένταξη των ατόμων σε κλάσεις

Single solution

Number of clusters
(καθορισμός
του αριθμού
των κλάσεων)



2

(καθορισμός
2 κλάσεων
βάσει της εκτίμησής μας
από το δένδrogramma)

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

65

Ένταξη των ατόμων σε κλάσεις

	ver_jump	hor_jump	m30	squat	Bench_press	Grip_strenght	push_ups	Cooper_test_12_min	VO2max
1	58,00	2,40	4,90	32	85	45	50	2000	50,20
2	63,00	2,55	4,75	33	70	40	42	2200	53,30
3	60,00	2,45	4,90	32	75	43	45	2300	54,40
4	66,00	2,60	4,55	35	85	46	48	2400	56,00
5	55,00	2,35	5,15	29	70	38	44	1900	49,00
6	59,00	2,45	5,05	31	75	43	47	2100	51,00
7	62,00	2,60	4,70	35	95	52	55	2200	53,00
8	57,00	2,40	4,85	32	90	53	56	2300	54,30
9	58,50	2,50	4,90	33	80	42	49	2400	55,00
10	60,50	2,45	4,90	32	85	44	47	2300	53,60
11	65,00	2,55	4,60	36	80	45	50	2300	55,00
12	54,00	2,35	5,05	30	70	39	44	2000	49,00
13	56,00	2,45	5,05	31	75	40	47	2100	51,00
14	64,00	2,60	4,80	34	95	55	55	2300	53,90
15	58,50	2,45	5,05	31	75	43	47	2100	51,00
16	64,00	2,60	4,60	34	95	52	55	2300	54,00
17	57,00	2,40	4,85	32	90	53	56	2300	54,30
18	59,50	2,50	4,90	33	80	42	49	2400	55,00
19	61,00	2,50	4,95	33	85	44	47	2200	53,60
20	65,50	2,55	4,60	36	80	45	50	2300	55,00

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

66

Ένταξη των ατόμων σε κλάσεις

	ver_jump	hor_jump	m30	squat	Bench_press	Grip_strength	push_ups	Cooper_test_12_min	VO2max	CLU2_1
1	58,00	2,40	4,90	32	85	45	50	2000	50,20	1
2	63,00	2,55	4,75	33	70	40	42	2200	53,30	1
3	60,00	2,45	4,90	32	75	43	45	2300	54,40	2
4	66,00	2,60	4,55	35	85	46	48	2400	56,00	2
5	55,00	2,35	5,15	29	70	38	44	1900	49,00	1
6	59,00	2,45	5,05	31						1
7	62,00	2,60	4,70	35						1
8	57,00	2,40	4,85	32						2
9	58,50	2,50	4,90	33						2
10	60,50	2,45	4,90	32						2
11	65,00	2,55	4,60	36						2
12	54,00	2,35	5,05	30						1
13	56,00	2,45	5,05	31	75	40	47	2100	51,00	1
14	64,00	2,60	4,80	34	95	55	55	2300	53,90	2
15	58,50	2,45	5,05	31	75	43	47	2100	51,00	1
16	64,00	2,60	4,60	34	95	52	55	2300	54,00	2
17	57,00	2,40	4,85	32	90	53	56	2300	54,30	2
18	59,50	2,50	4,90	33	80	42	49	2400	55,00	2
19	61,00	2,50	4,95	33	85	44	47	2200	53,60	1
20	65,50	2,55	4,60	36	80	45	50	2300	55,00	2

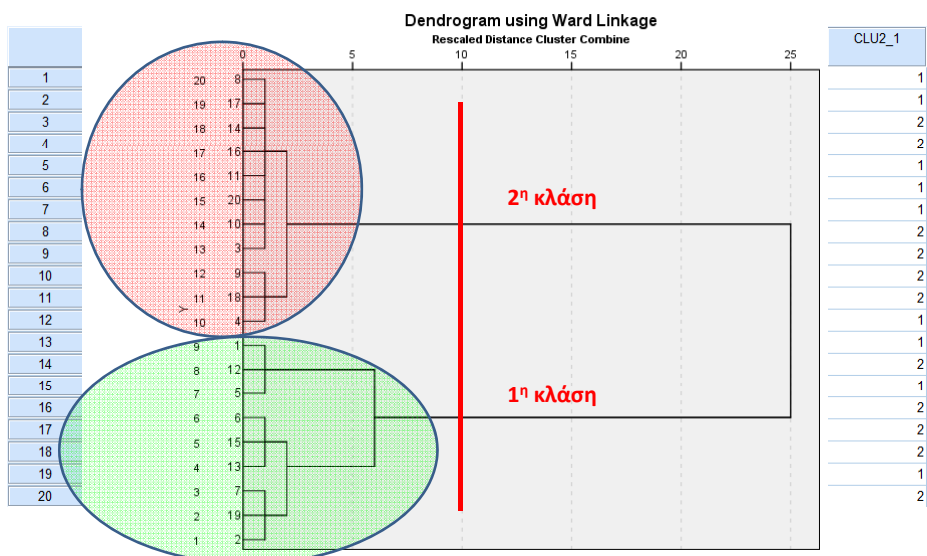
1 = το άτομο ανήκει στην 1^η κλάση

2 = το άτομο ανήκει στη 2^η κλάση

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

67

Ένταξη των ατόμων σε κλάσεις



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

68

Συνοψίζοντας ...

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

69

Συνοψίζοντας ...

Ποιος είναι ο στόχος της ανάλυσης ταξινόμησης:



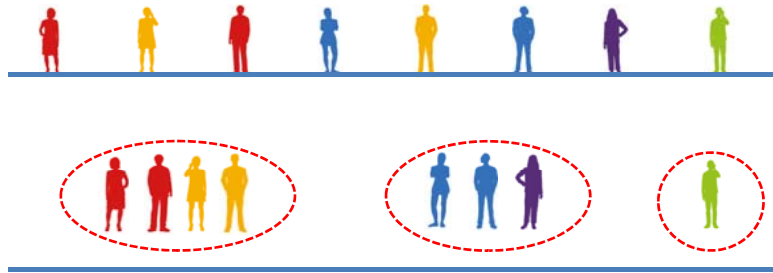
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

70

Συνοψίζοντας ...

Ποιος είναι ο στόχος της ανάλυσης ταξινόμησης:

η ομαδοποίηση των **ατόμων** (βάσει κοινών χαρακτηριστικών)



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

71

Συνοψίζοντας ...

Ποια είναι η κυρίαρχη διαφορά μεταξύ
«ιεραρχημένης» και «μη ιεραρχημένης» ταξινόμησης:



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

72

Συνοψίζοντας ...

Ποια είναι η κυρίαρχη διαφορά μεταξύ «ιεραρχημένης» και «μη ιεραρχημένης» ταξινόμησης:

«ιεραρχημένη» ... (*hierarchical*):



- δεν είναι προαποφασισμένος ο αριθμός των κλάσεων

«μη ιεραρχημένη - διαμερισματική» ... (*k - means*):



- είναι προαποφασισμένος ο αριθμός των κλάσεων

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

73

Συνοψίζοντας ...

Ποιες είναι



δύο (2) από τις βασικές μεθόδους ελέγχου της απόστασης
(του βαθμού συγγένειας – γειτνίασης) μεταξύ 2 ατόμων
στην ανάλυση ταξινόμησης:

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

74

Συνοψίζοντας ...

Ποιες είναι

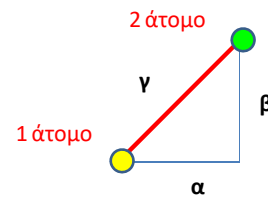
δύο (2) από τις βασικές μεθόδους ελέγχου της απόστασης

(του βαθμού συγγένειας – γειτνίασης) μεταξύ 2 ατόμων

στην ανάλυση ταξινόμησης:

 Ευκλείδεια απόσταση: $\gamma = \sqrt{\alpha^2 + \beta^2}$
(*Euclidean distance*)

 Τετράγωνο Ευκλείδειας απόστασης: $\gamma^2 = \alpha^2 + \beta^2$
(*Squared Euclidean distance*)



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

75

Συνοψίζοντας ...

Βάσει ποιού κριτηρίου

πραγματοποιείται η ένωση 2 ατόμων (ή κλάσεων) :



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

76

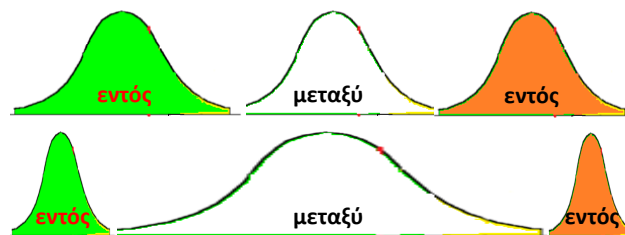
Συνοψίζοντας ...

Βάσει ποιού κριτηρίου

πραγματοποιείται η ένωση 2 ατόμων (ή κλάσεων) :

Βάσει του κριτηρίου του Ward

- 👉 ελαχιστοποίηση της «εντός» της κλάσης διακύμανσης και
- 👉 αύξηση της «μεταξύ» των κλάσεων διακύμανσης



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

77

Συνοψίζοντας ...

Με τι ισούται η συνολική διακύμανση όλων των ατόμων
στην ΑΡΧΗ και στο ΤΕΛΟΣ της ανάλυσης ταξινόμησης

(σε σχέση με τη διακύμανση «εντός» και «μεταξύ» των κλάσεων):



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

78

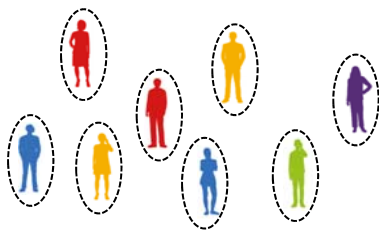
Συνοψίζοντας ...



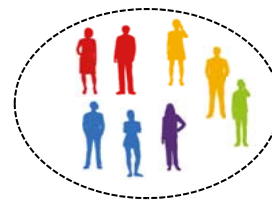
**Με τι ισούται η συνολική διακύμανση όλων των ατόμων
στην ΑΡΧΗ και στο ΤΕΛΟΣ της ανάλυσης ταξινόμησης**

(σε σχέση με τη διακύμανση εντός και μεταξύ των κλάσεων):

Στην ΑΡΧΗ
ισούται με την
«**μεταξύ**» των κλάσεων διακύμανση.



Στο ΤΕΛΟΣ
ισούται με την
«**εντός**» των κλάσεων διακύμανση.



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

79

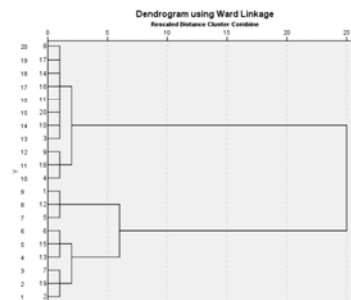
Συγγραφή αποτελεσμάτων ...

Από την εφαρμογή της **ανάλυσης ταξινόμησης**,
για την **ομαδοποίηση** των ατόμων,

χρησιμοποιώντας
το τετράγωνο της Ευκλείδειας απόστασης (Squared Euclidian distance)
για τον καθορισμό της απόστασης μεταξύ των ατόμων ή/και των επιμέρους κλάσεων
και
το κριτήριο του Ward
για την ένταξη των ατόμων ή/και των επιμέρους κλάσεων στην ίδια ομάδα

προέκυψαν
2 ομάδες ατόμων,

όπως απεικονίζεται στο δενδρόγραμμα:



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

80

Ερωτήσεις;



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

81

Youtube video_ Cluster Analysis

Cluster Analysis – SPSS : <https://www.youtube.com/watch?v=PBKxZkOzGxY>

Cluster Analysis – SPSS : <https://www.youtube.com/watch?v=bMH-aHNIhBA>

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

82