

## 8ο ΕΡΓΑΣΤΗΡΙΟ

Σύγκριση μέσων τιμών σε περισσότερες από 2 ομάδες (ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ - ANOVA).

Επιλέγουμε από το Μενού:

Analyze → Compare Means → One Way ANOVA

Στην επιλογή “dependent variable list” τοποθετούμε πάντα την ποσοτική μεταβλητή, που θέλουμε να συγκρίνουμε τις μέσες τιμές της μεταξύ των ομάδων και ως “Factor” επιλέγουμε τη μεταβλητή που διαχωρίζει τις ομάδες.

Χρήσιμες επιλογές είναι επιπλέον:

A) Options – επιλέγουμε το Descriptive Statistics ώστε να πάρουμε στο output, έναν πίνακα με τις μέσες τιμές, τις τυπικές αποκλίσεις και άλλα χρήσιμα στοιχεία για τις διάφορες ομάδες που συγκρίνουμε.

B) Post Hoc – επιλέγουμε να κάνουμε πολλαπλές συγκρίσεις, μεταξύ των ομάδων με κάποια από τις μεθόδους που προτείνονται (μια πολύ καλή μέθοδος είναι αυτή του Tukey). Οι πολλαπλές συγκρίσεις έχουν νόημα, ΜΟΝΟ όταν το αποτέλεσμα της ANOVA είναι στατιστικά σημαντικό (δηλ. όταν  $p < 0,05$ )

### Παράδειγμα 1

Θέλουμε να συγκρίνουμε την Χοληστερόλη ανάμεσα στις 3 περιοχές (Αστικές, Ημιαστικές και Αγροτικές). Στην “dependent variable list” τοποθετούμε την ΧΟΛΗΣΤΕΡΟΛΗ, ενώ στο “Factor” την ΠΕΡΙΟΧΗ ΚΑΤΟΙΚΙΑΣ. Επιλέγοντας το Options (Descriptives) και το Post Hoc (Tukey), έχουμε τα αποτελέσματα

Αρχικά στο output, διαβάζουμε τον πίνακα με τις μέσες τιμές, τις τυπικές αποκλίσεις και τα τυπικά σφάλματα των μέσων τιμών ανά ομάδα (που θα μπορούσαμε να τα διαβάσουμε και με την εντολή EXPLORE, που ήδη γνωρίζουμε)

### Descriptives

Ολική χοληστερόλη

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Αγροτική	161	179,88	33,984	2,678	174,59	185,17	83	284
Ημιαστική	142	183,04	35,677	2,994	177,12	188,95	82	294
Αστική	178	183,55	33,788	2,532	178,55	188,55	106	290
Total	481	182,17	34,388	1,568	179,09	185,25	82	294

Από τον πίνακα μπορούμε να κρατήσουμε τα στοιχεία (με στρογγυλοποιήσεις)

Περιοχές	N	μέση τιμή	Τυπική απόκλιση
Αγροτική	161	179,9	33,98
Ημιαστική	142	183,0	35,68
Αστική	178	183,6	33,79

Στη συνέχεια έχουμε το F-test και την τιμή p (p-value)

#### ANOVA

Ολική χοληστερόλη

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1288,394	2	644,197	,544	,581
Within Groups	566319,627	478	1184,769		
Total	567608,021	480			

Οπότε διαβάζοντας τις παραπάνω τιμές, έχουμε

$F=0.544$  και η τιμή p (SIG στο SPSS), είναι 0.581 οπότε

$$p=0.581 > 0.05$$

συνεπώς δεν υπάρχει στατιστικά σημαντική διαφορά στα στις μέσες τιμές χοληστερόλης στις τρεις περιοχές. Οι διαφορές που παρουσιάζονται στον παραπάνω πίνακα (179,9 - 183 και 183,6) είναι πιθανότατα τυχαίες. Αφού δεν βρέθηκαν σημαντικές διαφορές, ο εκ των υστέρων έλεγχος (post hoc) του Tukey, δεν δείχνει διαφορές όπως φαίνεται και από τον επόμενο πίνακα, όπου όλες οι περιοχές ομαδοποιούνται στο ίδιο σύνολο (subset 1)

#### Ολική χοληστερόλη

Tukey HSD<sup>a,b</sup>

		Subset for alpha = 0.05
Περιοχή κατοικίας	N	1
Αγροτική	161	179,88
Ημιαστική	142	183,04
Αστική	178	183,55
Sig.		,609

#### Παράδειγμα 2

Θέλουμε να συγκρίνουμε την ΣΥΣΤΟΛΙΚΗ ΠΙΕΣΗ ανάμεσα σε τρεις ηλικιακές ομάδες, που έχουμε δημιουργήσει. Η πρώτη ομάδα είναι κάτω από 30 ετών, η δεύτερη μεταξύ 31 και 40 ετών και η τρίτη είναι πάνω από 41 ετών.

Αποτελέσματα

### Descriptives

Συστολική ΑΠ

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
<= 30	73	105,78	9,585	1,122	103,54	108,02	85	134
31 - 40	305	108,89	11,400	,653	107,61	110,18	85	148
41+	103	113,50	12,931	1,274	110,98	116,03	91	154
Total	481	109,41	11,721	,534	108,36	110,46	85	154

Από τον πίνακα μπορούμε να κρατήσουμε τα στοιχεία (με στρογγυλοποιήσεις)

Ηλικίες	N	μέση τιμή	Τυπική απόκλιση
Κάτω από 30	73	105,8	9,59
31-40 ετών	305	108,9	11,40
Πάνω από 41	103	113,5	12,93

Ο Πίνακας ANOVA

### ANOVA

Συστολική ΑΠ

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2770,463	2	1385,231	10,481	,000
Within Groups	63175,670	478	132,167		
Total	65946,133	480			

Διαβάζοντας τις παραπάνω τιμές, έχουμε

$F=10,481$  και η τιμή  $p$  (SIG στο SPSS), είναι πολύ κοντά στο 0, οπότε γράφουμε

$$P << 0.05$$

συνεπώς φαίνεται να υπάρχει στατιστικά σημαντική διαφορά στα στις μέσες τιμές της συστολικής πίεσης ανάμεσα στις ηλικιακές ομάδες. Ο εκ των υστέρων έλεγχος (post hoc) του Tukey, που εντοπίζει τις διαφορές ανάμεσα στις 3 ομάδες, δείχνει 2 σύνολα (subset 1 και subset 2) ως εξής

### Συστολική ΑΠ

Tukey HSD<sup>a,b</sup>

Ηλικιακές Ομάδες	N	Subset for alpha = 0.05	
		1	2
<= 30	73	105,78	
31 - 40	305	108,89	
41+	103		113,50
Sig.		,106	1,000

Από τον πίνακα βλέπουμε ότι στις 2 πρώτες ομάδες (κάτω από 30 και 31-40 ετών) ΔΕΝ υπάρχει διαφορά στις συστολικές πιέσεις. Όμως στην ηλικιακή ομάδα από 41 ετών και πάνω, η συστολική πίεση δείχνει να είναι διαφορετική (αυξημένη), σε σχέση με τις άλλες δύο ομάδες.