

**ΜΑΘΗΜΑ:** ΣΤΑΤΙΣΤΙΚΗ & ΘΕΩΡΙΑ ΑΠΟΦΑΣΕΩΝ

**ΔΙΔΑΣΚΟΝΤΕΣ:** Ε. ΓΑΚΗ –Β. ΜΕΝΝΗΣ

## Ασκήσεις για τα Χριστούγεννα

### Α ΜΕΡΟΣ

Η Chamari Atarathne, νέα διευθύνουσα σύμβουλος της εταιρείας SriLankan Tea, μιας από τις μεγαλύτερες εταιρείες τσαγιού της Σρι Λάνκα, που χρησιμοποιεί ως κύριο μέσο προώθησης το ηλεκτρονικό ταχυδρομείο για να πουλήσει τα προϊόντα της, θέλει να αναπροσαρμόσει την στρατηγική πωλήσεων και παράλληλα να κάνει αναδιατάξεις στο ανθρώπινο προσωπικό. Πριν προχωρήσει σε οποιεσδήποτε αλλαγές, ζήτησε να συγκεντρωθούν στοιχεία που αφορούν, μεταξύ άλλων, τα ηλεκτρονικά μηνύματα που στέλνουν οι εργαζόμενοι στον τομέα των πωλήσεων και τις αντίστοιχες πωλήσεις τους σε χιλιάδες κιλά. Στον παρακάτω πίνακα καταγράφονται τα στοιχεία για ένα τυχαίο δείγμα 10 υπαλλήλων της εταιρείας κατά τη διάρκεια του τελευταίου μήνα

Αριθμός ηλεκτρονικών μηνυμάτων που εστάλησαν (X)	20	40	20	30	10	10	20	20	20	30
Πωλήσεις σε χιλιάδες κιλά (Y)	30	60	40	60	30	40	40	50	30	70

Χρησιμοποιώντας τα δεδομένα αυτά:

#### Θέμα 1ο

Η Chamari δίνει bonus σε εκείνους τους υπαλλήλους που έχουν το υψηλότερο 10% των πωλήσεων. Πόσα κιλά θα πρέπει να πουλήσει ένας υπάλληλος προκειμένου να πάρει το bonus; Η Chamari πιστεύει ότι οι υπάλληλοι δεν στέλνουν πολλά ηλεκτρονικά μηνύματα. Συμφωνείτε με την άποψη αυτή; Αιτιολογήστε χρησιμοποιώντας το κατάλληλο στατιστικό μέτρο (Για διευκόλυνσή σας δίνεται ότι η τυπική απόκλιση του αριθμού των ηλεκτρονικών μηνυμάτων είναι 9,2).

#### Θέμα 2ο

- Η Chamari θέλει να χρησιμοποιήσει ένα μοντέλο πρόβλεψης των πωλήσεων. Ποιο θα ήταν αυτό; Πόσο βέβαιοι είστε για την καταλληλότητα του μοντέλου;
- Ο πρόεδρος του συλλόγου εργαζομένων ισχυρίζεται ότι η πίεση προς τους εργαζόμενους να στέλνουν πολλά ηλεκτρονικά μηνύματα για να αυξηθούν οι πωλήσεις είναι υπερβολική καθώς πιστεύει ότι τα ηλεκτρονικά μηνύματα δεν επηρεάζουν τις πωλήσεις. Συμφωνείτε με τον ισχυρισμό του; Αιτιολογήστε χρησιμοποιώντας το κατάλληλο στατιστικό μέτρο.
- Ας υποθέσουμε ότι η σχέση που συνδέει τις δύο μεταβλητές είναι μη γραμμική και μάλιστα της μορφής  $Y = \gamma + \delta x^2 + \varepsilon$ . Ο πίνακας που ακολουθεί περιέχει την ανάλυση του μοντέλου αυτού. Ποιο από τα δύο μοντέλα πρέπει να επιλέξει η Chamari; Αιτιολογήστε την απόφασή σας.

Σημείωση: Για διευκόλυνση των υπολογισμών δίνονται  $S_{YY} = 1850$   $S_{XX} = 760$   $S_{Y/X} = 0,35$

### Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.749 <sup>a</sup>	.561	.507	10.071

a. Predictors: (Constant), Αρ. ηλ. μην/των (x2)

### Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	31.781	5.216		6.093	.000
	Αρ. ηλ. μην/των (x2)	.024	.007	.749	3.200	.013

a. Dependent Variable: Πωλήσεις σε χιλιάδες κιλά (Y)

### Θέμα 3<sup>ο</sup>

Η εταιρεία διαθέτει τηλεφωνική γραμμή υποστήριξης των πελατών της. Από ιστορικά στοιχεία είναι γνωστό ότι οι κλήσεις καταφθάνουν με ρυθμό 48 ανά ώρα. Έστω ότι δεν υπάρχει καμία κλήση αυτή τη στιγμή και ο υπάλληλος που έχει βάρδια ζητά άδεια για να απομακρυνθεί από τη θέση του για 2 λεπτά. Σύμφωνα με την πολιτική της εταιρίας δικαιούται να πάρει άδεια αν η πιθανότητα να υπάρχει τουλάχιστον 1 κλήση στο διάστημα της απουσίας του δεν υπερβαίνει το 60%. Μπορεί να φύγει από τη θέση του τελικά ο υπάλληλος για 2 λεπτά;

### Θέμα 4<sup>ο</sup>

Ο Chandika Hathurusanthe επικεφαλής του τμήματος μάρκετινγκ της εταιρείας, προτείνει την εντατικοποίηση των προσπαθειών προώθησης ενός νέου τσαγιού της εταιρείας που θα βοηθά σημαντικά στην απώλεια βάρους καθώς πιστεύει ότι σε αυτόν τον τομέα η εταιρεία υστερεί με αποτέλεσμα την απώλεια εσόδων. Μάλιστα, πρόσφατα η εταιρεία έβγαλε στην αγορά τρεις ποικιλίες τέτοιου τσαγιού (Π1, Π2, Π3). Ο Chandika πιστεύει ότι θα πρέπει να προωθηθούν και οι τρεις αυτές ποικιλίες καθώς έχουν τις ίδιες πωλήσεις, αντίθετα η Chamari υποστηρίζει ότι θα πρέπει να επενδύσουν μόνο σε μία ποικιλία. Για να καταλήξουν σε μια απόφαση, προχώρησαν στον ακόλουθο πειραματικό σχεδιασμό. Χρησιμοποίησαν 15 υπαλλήλους, τους οποίους χώρισαν σε 3 ισοπληθείς ομάδες, κάθε μία από τις οποίες προώθησε μία από τις τρεις διαφορετικές ποικιλίες τσαγιού και κατέγραψαν τις πωλήσεις τους. Τα αποτελέσματα συνοψίζονται στον πίνακα που ακολουθεί. Με ποιον συντάσσετε; Με την Chamari ή με τον Chandika; Γιατί; Αιτιολογήστε χρησιμοποιώντας την κατάλληλη στατιστική τεχνική. Για την απάντηση των ερωτημάτων μπορείτε να χρησιμοποιήσετε τους πίνακες αποτελεσμάτων που ακολουθούν. Εναλλακτικά μπορείτε να τρέξετε την ανάλυση που απαιτείται. Για διευκόλυνσή σας δίνεται ότι

$$\sum_{i=1}^k \sum_{j=1}^{n_i} Y_{ij} = 2338, \quad \alpha = 0,05$$

Πίνακας με τα δεδομένα

<b>Π1</b>	8	7	9	13	10
<b>Π2</b>	14	16	12	17	11
<b>Π3</b>	10	12	16	15	12

Πίνακες αποτελεσμάτων

### Descriptive Statistics

Dependent Variable: Πωλήσεις

Ποικιλία τσαγιού	Mean	Std. Deviation	N
Ποικιλία 1	9.40	2.302	5
Ποικιλία 2	14.00	2.550	5
Ποικιλία 3	13.00	2.449	5
Total	12.13	3.044	15

### Tests of Between-Subjects Effects

Dependent Variable: Πωλήσεις

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	58.533 <sup>a</sup>	2	29.267	4.933	.027
Intercept	2208.267	1	2208.267	372.180	.000
poikilia	58.533	2	29.267	4.933	.027
Error	71.200	12	5.933		
Total	2338.000	15			
Corrected Total	129.733	14			

a. R Squared = .451 (Adjusted R Squared = .360)

### Multiple Comparisons

Dependent Variable: Πωλήσεις

Tukey HSD

(I) Ποικιλία τσαγιού	(J) Ποικιλία τσαγιού	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Ποικιλία 1	Ποικιλία 2	-4.60*	1.541	.028	-8.71	-.49
	Ποικιλία 3	-3.60	1.541	.089	-7.71	.51
Ποικιλία 2	Ποικιλία 1	4.60*	1.541	.028	.49	8.71
	Ποικιλία 3	1.00	1.541	.796	-3.11	5.11
Ποικιλία 3	Ποικιλία 1	3.60	1.541	.089	-.51	7.71
	Ποικιλία 2	-1.00	1.541	.796	-5.11	3.11

Based on observed means.

The error term is Mean Square(Error) = 5.933.

\*. The mean difference is significant at the .05 level.

### Levene's Test of Equality of Error Variances<sup>a,b</sup>

		Levene Statistic	df1	df2	Sig.
Πωλήσεις	Based on Mean	.120	2	12	.888
	Based on Median	.086	2	12	.918
	Based on Median and with adjusted df	.086	2	11.024	.918
	Based on trimmed mean	.128	2	12	.881

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Dependent variable: Πωλήσεις

b. Design: Intercept + poikilia

### Tests of Normality

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Residual	,146	15	,200*	,927	15	,248

\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

### Test Statistics<sup>a,b</sup>

	Πωλήσεις
Kruskal-Wallis H	6.050
df	2
Asymp. Sig.	.049

a. Kruskal Wallis Test

b. Grouping Variable:  
Ποικιλία τσαγιού

## Β ΜΕΡΟΣ

Ο Διευθυντής της παθολογικής κλινικής ενός πανεπιστημιακού νοσοκομείου, διεξάγει μία έρευνα για τη σχέση που υπάρχει μεταξύ της ηλικίας και της αρτηριακής πίεσης των ασθενών. Για το σκοπό αυτό έχει συγκεντρώσει διάφορα στοιχεία που αφορούν την αρτηριακή πίεση των ασθενών, την ηλικία καθώς και το φύλο τους. Στη συνέχεια θα χρησιμοποιηθούν ορισμένα στοιχεία από την έρευνα αυτή.

### Θέμα 1<sup>ο</sup>

Ο πίνακας που ακολουθεί περιέχει στοιχεία που αφορούν την αρτηριακή πίεση για τριανταπέντε άνδρες ασθενείς. Ο διευθυντής πιστεύει ότι οι περισσότεροι άνδρες έχουν χαμηλή πίεση. Συμφωνείτε με τη άποψή του; Αποδείξτε το χρησιμοποιώντας το κατάλληλο στατιστικό μέτρο. Ποια είναι η αρτηριακή πίεση πάνω από την οποία έχει το 10% των ανδρών ασθενών; Για διευκόλυνσή σας, η τυπική απόκλιση των δεδομένων είναι ίση με 23,34.

Αρτηριακή πίεση	Άνδρες
100-120	4
120-140	10
140-160	12
160-180	5
180-200	4
<b>Σύνολο</b>	<b>35</b>

### Θέμα 2<sup>ο</sup> (2,5 μονάδες)

Ο πίνακας που ακολουθεί περιέχει την αρτηριακή πίεση για τρεις ηλικιακές κατηγορίες τυχαίου δείγματος 24 ασθενών. Υπάρχει στατιστικά σημαντική διαφορά στην αρτηριακή πίεση ανάλογα με την ηλικία; Αν ναι, μεταξύ ποιων ηλικιακών ομάδων; Για την απάντηση των ερωτημάτων μπορείτε να χρησιμοποιήσετε τους πίνακες αποτελεσμάτων που ακολουθούν. Εναλλακτικά μπορείτε να τρέξετε την ανάλυση που απαιτείται. Για διευκόλυνσή σας δίνεται ότι

$$\sum_{i=1}^k \sum_{j=1}^{ni} Y_{ij}^2 = 477984$$

Πίνακας με τα δεδομένα

Ηλικία ασθενών (agecat)		
Κάτω από 25	25-45	Πάνω από 45
130	135	160
118	149	155
150	145	160
125	125	155
130	125	160
118	149	178
130	125	150
110	125	155

Πίνακες αποτελεσμάτων

**Descriptive Statistics**

Dependent Variable: Αρτηριακή πίεση

Ηλικία ασθενών	Mean	Std. Deviation	N
Κάτω από 25	126,38	11,999	8
25-45	134,75	11,285	8
Πάνω από 45	159,13	8,391	8
Total	140,08	17,475	24

**Levene's Test of Equality of Error Variances<sup>a</sup>**

Dependent Variable: Αρτηριακή πίεση

F	df1	df2	Sig.
1,077	2	21	,359

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + agecat

**Tests of Between-Subjects Effects**

Dependent Variable: Αρτηριακή πίεση

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	4631,583 <sup>a</sup>	2	2315,792	20,329	,000
Intercept	470960,167	1	470960,167	4134,252	,000
agecat	4631,583	2	2315,792	20,329	,000
Error	2392,250	21	113,917		
Total	477984,000	24			
Corrected Total	7023,833	23			

a. R Squared = ,659 (Adjusted R Squared = ,627)

### Multiple Comparisons

Dependent Variable: Αρτηριακή πίεση

Tukey HSD

(I) Ηλικία ασθενών	(J) Ηλικία ασθενών	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Κάτω από 25	25-45	-8,38	5,337	,281	-21,83	5,08
	Πάνω από 45	-32,75*	5,337	,000	-46,20	-19,30
25-45	Κάτω από 25	8,38	5,337	,281	-5,08	21,83
	Πάνω από 45	-24,38*	5,337	,000	-37,83	-10,92
Πάνω από 45	Κάτω από 25	32,75*	5,337	,000	19,30	46,20
	25-45	24,38*	5,337	,000	10,92	37,83

Based on observed means.

The error term is Mean Square(Error) = 113,917.

\*. The mean difference is significant at the ,05 level.

### Tests of Normality

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Residual for pressureall	,153	24	,153	,931	24	,101

a. Lilliefors Significance Correction

### Θέμα 3° (3 μονάδες)

Ο Διευθυντής της παθολογικής κλινικής χρησιμοποίησε στοιχεία που αφορούσαν δείγμα δώδεκα γυναικών και κατέγραψε την ηλικία τους και την αρτηριακή πίεσή τους. Τα στοιχεία που συνέλεξε συνοψίζονται στον πίνακα που ακολουθεί.

<b>Ηλικία (X)</b>	56	42	72	36	63	47	55	49	38	42	68	60
<b>Αρτηριακή Πίεση (Y)</b>	147	125	160	118	149	128	150	145	115	140	152	155

Με βάση τα παραπάνω στοιχεία και αν επιπλέον δίνονται  $S_{xx} = 1550,67$ ,  $S_{yy} = 2500,67$ ,  $S_{y|x} = 7.018$ :

- Υπάρχει στατιστικά σημαντική σχέση μεταξύ ηλικίας και αρτηριακής πίεσης;
- Εάν χρειαζόταν να προβλέψετε την αρτηριακή πίεση μίας γυναίκας 45 ετών, ποια θα ήταν αυτή;
- Εάν είχατε να επιλέξετε μεταξύ ενός γραμμικού μοντέλου και ενός τετραγωνικού μοντέλου για το οποίο η ανάλυση έδωσε τα αποτελέσματα που ακολουθούν, ποιο θα επιλέγατε; Αιτιολογήστε την απάντησή σας.



### Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,870 <sup>a</sup>	,758	,733	7,787

a. Predictors: (Constant), Ηλικία γυναικών (x2)

### Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	110,814	5,740		19,306	,000
	Ηλικία γυναικών (x2)	,010	,002	,870	5,589	,000

a. Dependent Variable: Αρτηριακή Πίεση

#### Θέμα 4<sup>ο</sup>

Πρόσφατες έρευνες υποστηρίζουν ότι όσο αυξάνεται η αρτηριακή πίεση τόσο αυξάνεται ο αριθμός ημερών νοσηλείας στα νοσηλευτικά ιδρύματα. Εάν είναι γνωστό ότι η αρτηριακή πίεση ακολουθεί κανονική κατανομή με μέση τιμή 155 και τυπική απόκλιση 6:

- i. Να υπολογισθεί το ποσοστό των ασθενών που έχουν αρτηριακή πίεση μεταξύ 150 και 165.
- ii. Να υπολογιστεί η αρτηριακή πίεση κάτω από την οποία έχει το 35% των ασθενών.

#### Θέμα 5<sup>ο</sup>

Για την μέτρηση της αρτηριακής πίεσης μπορεί να χρησιμοποιηθεί είτε αναλογικό είτε ψηφιακό πιεσόμετρο. Στην δεύτερη περίπτωση, η πιθανότητα λάθος μέτρησης είναι 5%. Να υπολογισθεί η πιθανότητα σε 10 διαδοχικές μετρήσεις να γίνει λάθος μέτρηση σε 2 τουλάχιστον περιπτώσεις.