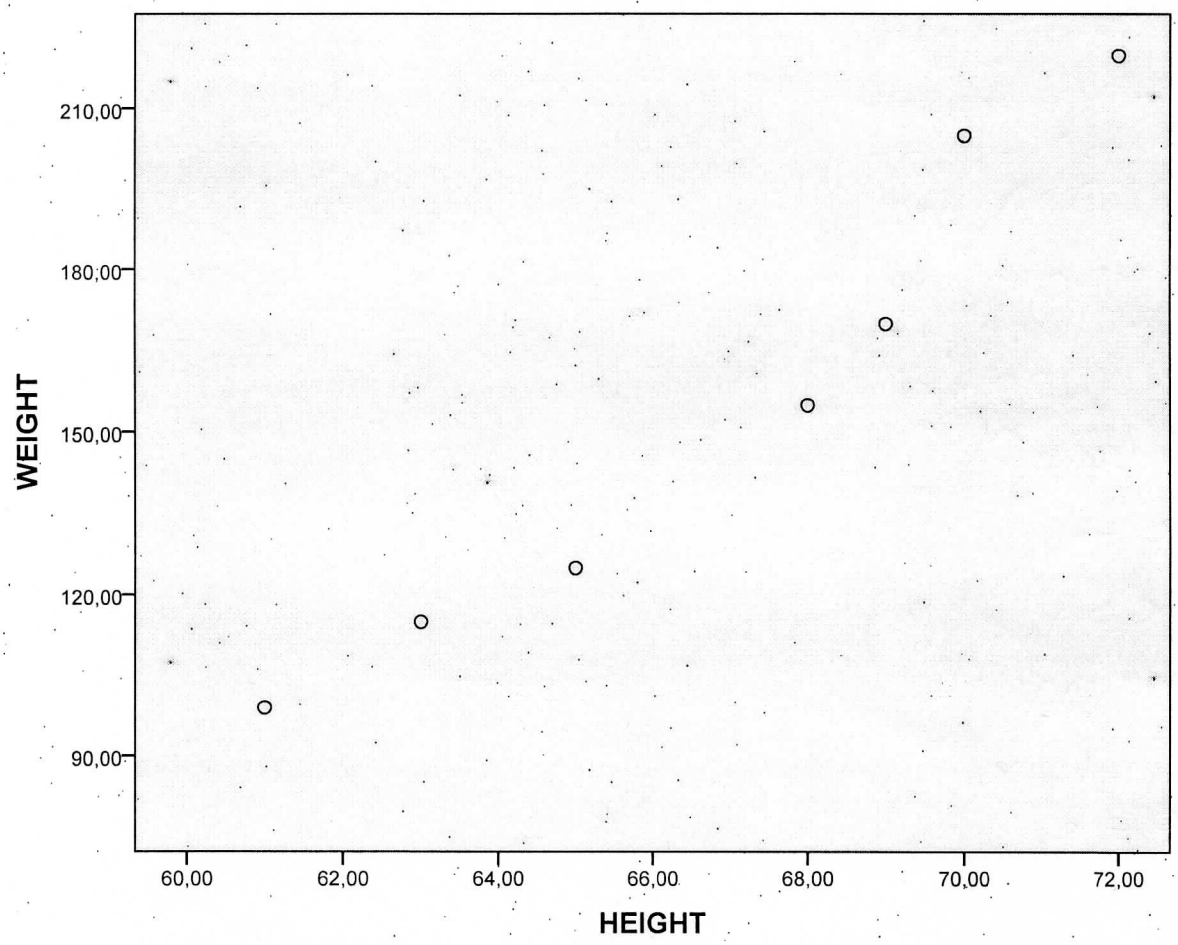
```
DATASET ACTIVATE DataSet2.
DATASET CLOSE DataSet1.
GRAPH
  /SCATTERPLOT(BIVAR)=HEIGHT WITH WEIGHT
  /MISSING=LISTWISE.
```

Graph

[DataSet2] C:\BACKUP\APPLIED STATISTICS\DATA_SET_12.sav



Scatterplot Y WEIGHT X HEIGHT

```
GET
  FILE='C:\BACKUP\APPLIED STATISTICS\DATA_SET_12.sav'.
DATASET NAME DataSet1 WINDOW=FRONT.
REGRESSION
  /DESCRIPTIVES MEAN STDDEV CORR SIG N
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT WEIGHT
  /METHOD=ENTER HEIGHT
  /SCATTERPLOT=(WEIGHT ,*ZRESID)
  /RESIDUALS HISTOGRAM(ZRESID) NORMPROB(ZRESID)
  /CASEWISE PLOT(ZRESID) ALL
  /SAVE PRED RESID.
```

Regression

[DataSet1] C:\BACKUP\APPLIED STATISTICS\DATA_SET_12.sav

Descriptive Statistics

	Mean	Std. Deviation	N
WEIGHT	155,5714	45,79613	7
HEIGHT	66,8571	3,97612	7

ΠΕΡΙΓΡΑΦΙΚΑ ΓΙΑ WEIGHT-HEIGHT

Correlations

		WEIGHT	HEIGHT
Pearson Correlation	WEIGHT	1,000	,972
	HEIGHT	,972	1,000
Sig. (1-tailed)	WEIGHT	.	,000
	HEIGHT	,000	.
N	WEIGHT	7	7
	HEIGHT	7	7

ΠΙΝΑΚΑΣ ΣΥΣΧΕΤΙΣΕΩΝ

$\hat{\alpha} = 0$

Ισχυρότατη συσχέτιση

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	HEIGHT[b]	.	Enter

→ ΜΕΘΟΔΟΣ ENTER

a. Dependent Variable: WEIGHT

b. All requested variables entered.

$SST = SSR + SSE$ $\quad \hat{\alpha} = 0$ (ισχυρή απόρριψη της H_0)

$MSR = \frac{SSR}{K}$ $\quad MSE = \frac{SSE}{n-K-1}$

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	11880,327	1	11880,327	84,451	,000[b]
	Residual	703,387	5	140,677		
	Total	12583,714	6			

a. Dependent Variable: WEIGHT

b. Predictors: (Constant), HEIGHT

K: = # ανεξ. μεταβ.

$H_0: b_0 = b_1 = 0$

H_1: κάποιο $b_i \neq 0$

$F = \frac{MSR}{MSE}$ $\quad F \sim F_{K, n-K-1}$

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-592,645	81,542		-7,268	,001
	HEIGHT	11,191	1,218	,972	9,190	,000

a. Dependent Variable: WEIGHT

$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i$

$11,191 \frac{S_{HEIGHT}}{S_{WEIGHT}} = 0,972$

Casewise Diagnostics[a]

Case Number	Std. Residual	X_i WEIGHT	Predicted Value	Residual
1	-1,127	155,00	168,3614	-13,36145
2	,757	99,00	90,0226	8,97741
3	,219	115,00	112,4051	2,59488
4	1,202	205,00	190,7440	14,25602
5	-,805	170,00	179,5527	-9,55271
6	-,825	125,00	134,7877	-9,78765
7	,580	220,00	213,1265	6,87349

a. Dependent Variable: WEIGHT

$\hat{\varepsilon}_i = Y_i - \hat{Y}_i$

$-1,127 \simeq \frac{-13,36145}{\sqrt{140,677}} = \frac{\text{Residual}_i}{\sqrt{MSE}}$

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	90,0226	213,1265	155,5714	44,49780	7
Residual	-13,36145	14,25602	,00000	10,82733	7
Std. Predicted Value	-1,473	1,293	,000	1,000	7
Std. Residual	-1,127	1,202	,000	,913	7

a. Dependent Variable: WEIGHT

$n = 7$

Charts

$\text{Std. Deviation}_{\text{Predicted value}} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\hat{Y}_i - \text{Mean})^2}$

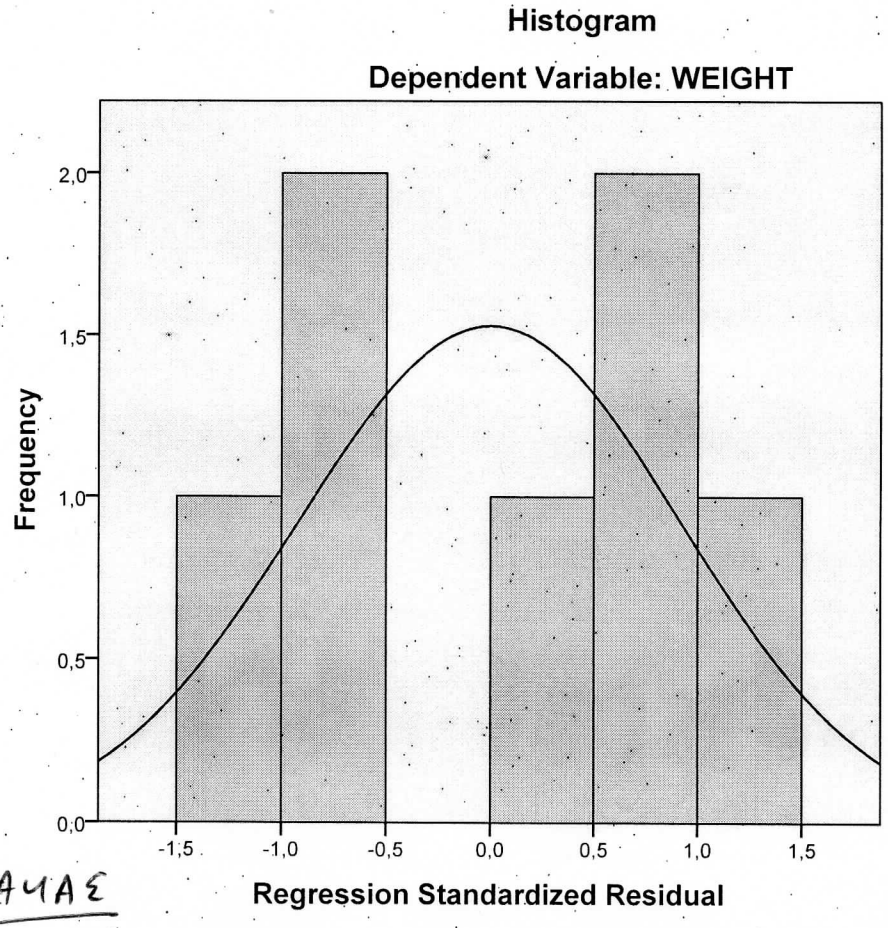
$\text{Std. Deviation}_{\text{residual}} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\hat{\varepsilon}_i - \text{Mean})^2}$

η αντίστοιχα

$\text{Std. Predicted}_{\pi X} = \frac{\text{Minimum} - \text{Mean}}{\text{Std. Deviation}}$ $\quad$ Std. Deviation

Αν οι τιμές είναι εντός των Std. Residuals εντός $[-2, 2]$ ενδείξη για καινούρια παρατήρηση

Charts



Εμπειρική ένδειξη (όχι τόσο ισχυρή) για κανονικότητα των καταλοίπων

ΠΙΝΑΚΑΣ

Model Summary

$R = r = 0,972$ (ισχυρότατη θετική γραμμική σχέση)

↑ Pearson

$R^2 = 0,972^2 = r^2 = R^2 = 1 - \frac{SSE}{SST} = \frac{SSR}{SST} = 0,944$

$R^2_{adj} = 1 - \left(\frac{n-1}{n-k-1}\right)(1-R^2)$

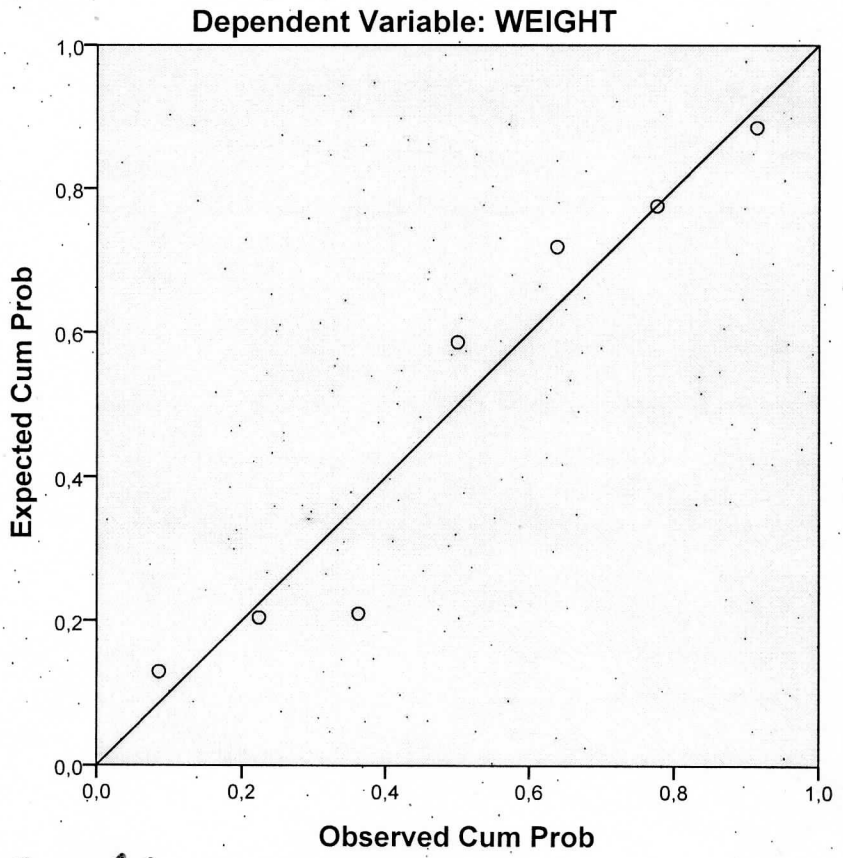
μικρή διαφορά R^2_{adj} με R^2 η οποία γίνεται μεγαλύτερη και σημαντική όταν έχουμε πολλές ανεξάρτητες μεταβλητές

Std. Error of Estimates

$= s = \sqrt{MSE} = \sqrt{\frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$

(σχετικά μικρο ενδειξεις "καλής" προσαρμογής του μοντέλου)

$= 11,86075$



→ ΟΧΙ ΙΔΙΑΙΤΕΡΕΣ ΑΠΟΚΛΙΣΕΙΣ ΑΠΟ ΤΗΝ ΥΠΟΘΕΣΗ ΤΗΣ ΚΑΝΟΝΙΚΟΤΗΤΑΣ ΤΩΝ ΚΑΤΑΛΟΙΠΩΝ (ΕΜΠΕΙΡΙΚΕΣ ΕΝΔΕΙΞΕΙΣ ΚΑΝΟΝΙΚΟΤΗΤΑΣ)

ΠΙΝΑΚΑΣ Coefficients

β = Constant = $\hat{b}_0 = -592,645$

$\hat{b}_1 = 11,191$

$\sqrt{S^2_{\hat{b}_0}} = \sqrt{s^2\left(\frac{1}{n} + \frac{\bar{x}^2}{(n-1)s_x^2}\right)}$, $s_x^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)$

$= S_{\hat{b}_0} = 81,542$

$\sqrt{S^2_{\hat{b}_1}} = S_{\hat{b}_1} = \sqrt{\frac{s^2}{(n-1)s_x^2}} = 1,218$

$H_0: b_0 = 0$ vs $H_1: b_0 \neq 0$

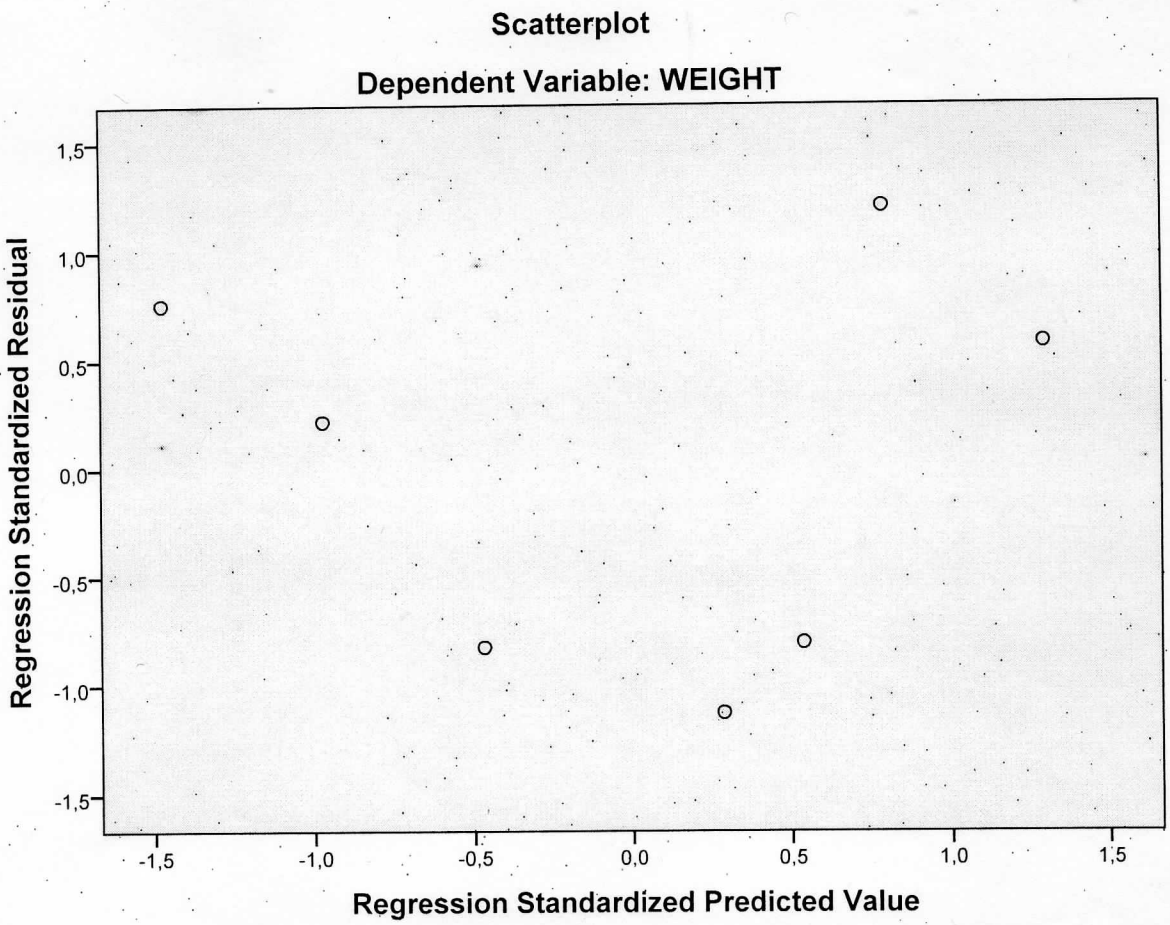
$T_{b_0} = \frac{|\hat{b}_0|}{S_{b_0}} \sim t_{n-2;\alpha/2}$

$= -7.268$ $\hat{\alpha} = 0.001$ σημαντική και η σταθερά και η μεταβλητή

$H_0: b_1 = 0$ vs $H_1: b_1 \neq 0$

$T_{b_1} = \frac{|\hat{b}_1|}{S_{b_1}}$

$= 9,190$ $\hat{\alpha} = 0$ HEIGHT.



Random pattern centred around the line of zero value. The points should have the same dispersion (διασπορά, διάχυση) γύρω από τη γραμμή Ο. Ένδειξη (εμπειρική) "τυχαιότητας" των καταλοίπων (Δεν φαίνεται κάποια "καθαρή" συστηματική σχέση των καταλοίπων (π.χ. συγγραμμικότητα των καταλοίπων).

Case Processing Summary

	N
Total Cases	7
Excluded Cases[a]	0
Forecasted Cases	0
Newly Created Cases	0

a. Cases with a missing value in any variable are excluded from the analysis.

Variable Processing Summary

		Variables	
		Dependent	Independent
		WEIGHT	HEIGHT
Number of Positive Values		7	7
Number of Zeros		0	0
Number of Negative Values		0	0
Number of Missing Values	User-Missing	0	0
	System-Missing	0	0

WEIGHT

Linear

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
,972	,944	,933	11,861

The independent variable is HEIGHT.

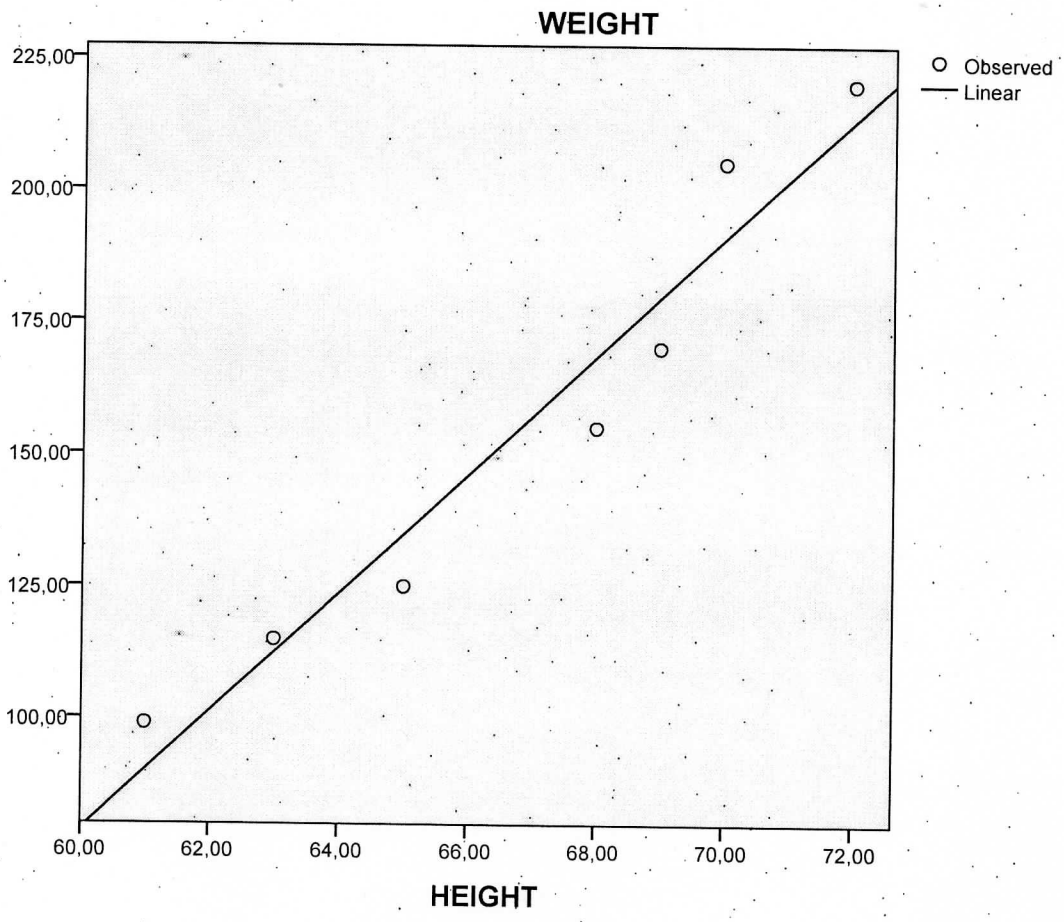
ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	11880,327	1	11880,327	84,451	,000
Residual	703,387	5	140,677		
Total	12583,714	6			

The independent variable is HEIGHT.

Coefficients

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
HEIGHT	11,191	1,218	,972	9,190	,000
(Constant)	-592,645	81,542		-7,268	,001



ΠΡΟΣΑΡΜΟΓΗ ΤΗΣ ΕΥΘΕΙΑΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ

```
ONEWAY mesi_katanalosi BY auto_mark
  /STATISTICS DESCRIPTIVES HOMOGENEITY
  /MISSING LISTWISE.
```

Oneway

[DataSet0] C:\BACKUP\APPLIED STATISTICS\DATA_SET_13.sav

Descriptives

mesi_katanalosi

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean	
					Lower Bound	Upper Bound
Toyota	6	26,9167	,74944	,30596	26,1302	27,7032
Datsun	5	26,3000	,57879	,25884	25,5813	27,0187
Mazda	4	23,7500	,59161	,29580	22,8086	24,6914
Total	15	25,8667	1,48019	,38218	25,0470	26,6864

ΠΕΡΙΓΡΑΦΙΚΑ ΜΕΤΡΑ

Descriptives

mesi_katanalosi

	Minimum	Maximum
Toyota	25,50	27,70
Datsun	25,30	26,80
Mazda	23,10	24,30
Total	23,10	27,70

(Ισχυρή ένδειξη μη-απόρριψης της H_0)

$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2$

Test of Homogeneity of Variances

mesi_katanalosi

Levene Statistic	df1	df2	Sig.
,078	2	12	,926

→ Levene Test

→ $\hat{\alpha} = 0.926$ ως προς τη διασπορά

Οι τρεις κάμπες είναι ομοιογενής ως προς τη διασπορά (ικανοποιώντας μία βασική υπόθεση της ANOVA.

ANOVA

mesi_katanalosi

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	25,475	$u-1$ 2	12,737	29,404	,000
Within Groups	5,198	n_+-k 12	,433		$\hat{\alpha}$
Total	30,673	n_+-1 14			

$SSTr \rightarrow$ Between Groups

$SSE \rightarrow$ Within Groups

$$F = \frac{MSSTr}{MSSE} \sim f_{u-1,\, n_+-k}$$

$$MSSTr = \frac{SSTr}{u-1}$$

$$MSSE = \frac{SSE}{n_+-u}$$