

Πολλαπλή Παλινδρόμηση

με τη βοήθεια της εφαρμογής
<https://www.statskingdom.com/index.html>

<https://www.statskingdom.com/index.html>

Πηγαίνετε στην επιλογή πάνω δεξιά “**regression**” (παλινδρόμηση).

Επιλέξτε την πολλαπλή παλινδρόμηση → “**multiple linear regression**”

Στην στήλη **Y** εισάγετε τις τιμές της **εξαρτημένης** μεταβλητής σας.

Στην στήλη **X1** εισάγετε τις τιμές της **ανεξάρτητης** μεταβλητής X1.

Στην στήλη **X2** εισάγετε τις τιμές της **ανεξάρτητης** μεταβλητής X2.

Αν θέλετε να προσθέσετε επιπλέον ανεξάρτητες μεταβλητές

πατήστε “**insert column**”

<https://www.statskingdom.com/index.html>

Σαν παράδειγμα, καταχωρήσαμε τιμές που αφορούν την ελληνική οικονομία.

year	cons	debt	wag
2026	148	372	69
2025	145	366	66
2024	143	364	63
2023	141	369	60
2022	138	368	56
2021	127	364	53
2020	121	351	49
2019	129	331	52
2018	126	335	50
2017	123	318	48
2016	120	315	48
2015	119	312	47

cons: τελική ιδιωτική κατανάλωση (σε δις.)

wag: ακαθάριστα εισοδήματα και μισθοί (σε δις.)

debt: ακαθάριστο δημόσιο χρέος

*Αν υπάρχουν (,) αντικαταστήστε με (.)
η στρογγυλοποιήστε στον πιο κοντινό
ακέραιο*

*Include	✓	✓	
Transform			
Groups	X1	X2	Y
Data	69	372	148
	66	366	145
	63	364	143
	60	369	141
	56	368	138
	53	364	127
	49	351	121
	52	331	129
	50	335	126
	48	318	123
	48	315	120
	47	312	119
πατήστε “calculate”			

Calculate

Insert column

Delete column

Clear

<https://www.statskingdom.com/index.html>

Ερμηνεία αποτελεσμάτων:

$$\hat{Y} = 56.927544 + 1.356837 X_1$$

Η εφαρμογή έχει πετάξει τον παράγοντα X_2 (“χρέος”) έξω από το μοντέλο ως μη στατιστικά σημαντικό.

Αν αυξηθούν οι μισθοί κατά 1 μονάδα (δηλ κατά 1 δις.) τότε θα αυξηθεί η κατανάλωση κατά περίπου 1,35 μονάδες (δηλ κατά 1,35 δις.)

<https://www.statskingdom.com/index.html>

Ερμηνεία αποτελεσμάτων:

$$(F(1, 10) = 187.72, p < .001, R^2 = 0.95, R^2_{adj} = 0.94).$$

Η εφαρμογή μας δίνει το αποτέλεσμα της στατιστικής τιμής ελέγχου F για τον έλεγχο της από κοινού σημαντικότητας στο μοντέλο που κρατάει.

F(1,10) → $k=1$ (αριθμός ανεξάρτητων μεταβλητών) , $n-k-1 = 12-2= 10$

Η στατιστική τιμή $F = 187,72$ είναι πάρα πολύ μεγάλη.

Μεγάλη F-stat σημαίνει την απόρριψη της H_0 (μηδενική υπόθεσης) του ελέγχου από κοινού σημαντικότητας. Το ίδιο συμπέρασμα προκύπτει και αν κοιτάξουμε το αντίστοιχο p-value του ελέγχου (είναι μικρότερο από το 0,001).

Ο συντελεστής προσδιορισμού είναι πολύ υψηλός. Το 95% της συνολικής μεταβλητότητας της κατανάλωσης εξηγείται από τη μεταβλητότητα των τιμών της ανεξάρτητης μεταβλητής (βλ. μισθοί).

<https://www.statskingdom.com/index.html>

Correlation matrix (pearson)

	Y	X ₁	X ₂
Y	1	0.974384	0.822786
X ₁	0.974384	1	0.806286
X ₂	0.822786	0.806286	1

Η εφαρμογή δίνει τις τιμές των συντελεστών γραμμικής συσχέτισης.

Τόσο οι μισθοί όσο και το χρέος φαίνεται πως σχετίζονται **έντονα θετικά** (οι μισθοί **σχεδόν τέλεια**) με την κατανάλωση. Έντονη είναι και η **θετική συσχέτιση** ανάμεσα στους μισθούς και το χρέος.

<https://www.statskingdom.com/index.html>

ANOVA table						
Source	DF	Sum of Square	Mean Square	F Statistic	P-value	
Regression (between $\hat{y}_i$ and $\bar{y}$)	1 k	1183.614431 SSR	1183.614431 SSR/k	187.71966	8.32168e-8	
Residual (between y_i and $\hat{y}_i$)	10 n-k-1	63.052236 SSE	6.305224 SSE/(n-k-1)			
Total (between y_i and $\bar{y}$)	11 n-1	1246.666667 SST	113.333333			

Ο πίνακας δείχνει τη διαδικασία της εξαγωγής της στατιστικής τιμής ελέγχου F για την από κοινού σημαντικότητα των συντελεστών (γίνεται πάνω στο υπόδειγμα που τελικά κράτησε η εφαρμογή, δηλ έχοντας πετάξει έξω το “χρέος”).

Υψηλή τιμή της F και πολύ χαμηλή του p-value (<0,05) συνεπάγεται την απόρριψη της μηδενικής υπόθεσης της από κοινού μη σημαντικότητα των β.

<https://www.statskingdom.com/index.html>

Coefficient Table Iteration 1 (adjusted R-squared = 0.943)								
	Coeff	SE	t-stat	lower t _{0.025} (9)	upper t _{0.975} (9)	Stand Coeff	p-value	VIF
b	46.687956	12.989219	3.594362	17.304303	76.07161	0	0.00579913	
X1	1.237617	0.16945	7.30372	0.854294	1.62094	0.888768	0.0000454648	2.857936
X2	0.0484225	0.0554917	0.872608	-0.0771084	0.173953	0.106185	0.405555	2.857936

Ο πίνακας δείχνει τα αποτελέσματα του ελέγχου της στατιστικής σημαντικότητας των β_1 και β_2 (καθώς και του σταθερού όρου). Η τιμή των t-stat δίνει την τιμή στατιστικού ελέγχου της σημαντικότητας. Πιο δίπλα, φαίνονται οι αντίστοιχες τιμές των p-value. **Μεγάλη τιμή των t-stat και πολύ μικρή τιμή των p-value (<0,05) σημαίνει την απόρριψη της μηδενικής υπόθεσης της μη στατιστικής σημαντικότητας.** Τόσο για τον σταθερό όρο όσο και για τον παράγοντα X1 (βλ. μισθοί) απορρίπτουμε την H_0 . Όχι όμως και για τον παράγοντα X2 (βλ. χρέος).

<https://www.statskingdom.com/index.html>

	Coeff	SE	t-stat	lower $t_{0.025}(10)$	upper $t_{0.975}(10)$	Stand Coeff	p-value	VIF
b	56.927544	5.50293	10.344952	44.666252	69.188835	0	0.00000116387	
X1	1.356837	0.0990314	13.701082	1.136182	1.577493	0.974384	8.32168e-8	1

Τέλος, η εφαρμογή ξανακάνει τους ελέγχους t για την στατιστική σημαντικότητα των συντελεστών β και για τον σταθερό όρο, αυτή τη φορά έχοντας πετάξει έξω τον μη στατιστικά σημαντικό παράγοντα “χρέος”.

Η τιμή του β_1 καπελάκι (coeff) ελαφρώς άλλαξε όταν πετάξαμε έξω τον παράγοντα X2. Από 1,23 έγινε 1,35.