

Τα Είδη των Δεδομένων

- 1) Ποσοτικά και Ποιοτικά
- 2) Πρωτογενή και Δευτερογενή
- 3) Διαστρωματικά, Χρονολογικά και Πάνελ

✓ Everything

Audio Recordings

Books/Printed Material

Films, Videos

Legislation

Manuscripts/Mixed Material

Maps

Notated Music

Newspapers

Periodicals

Personal Narratives

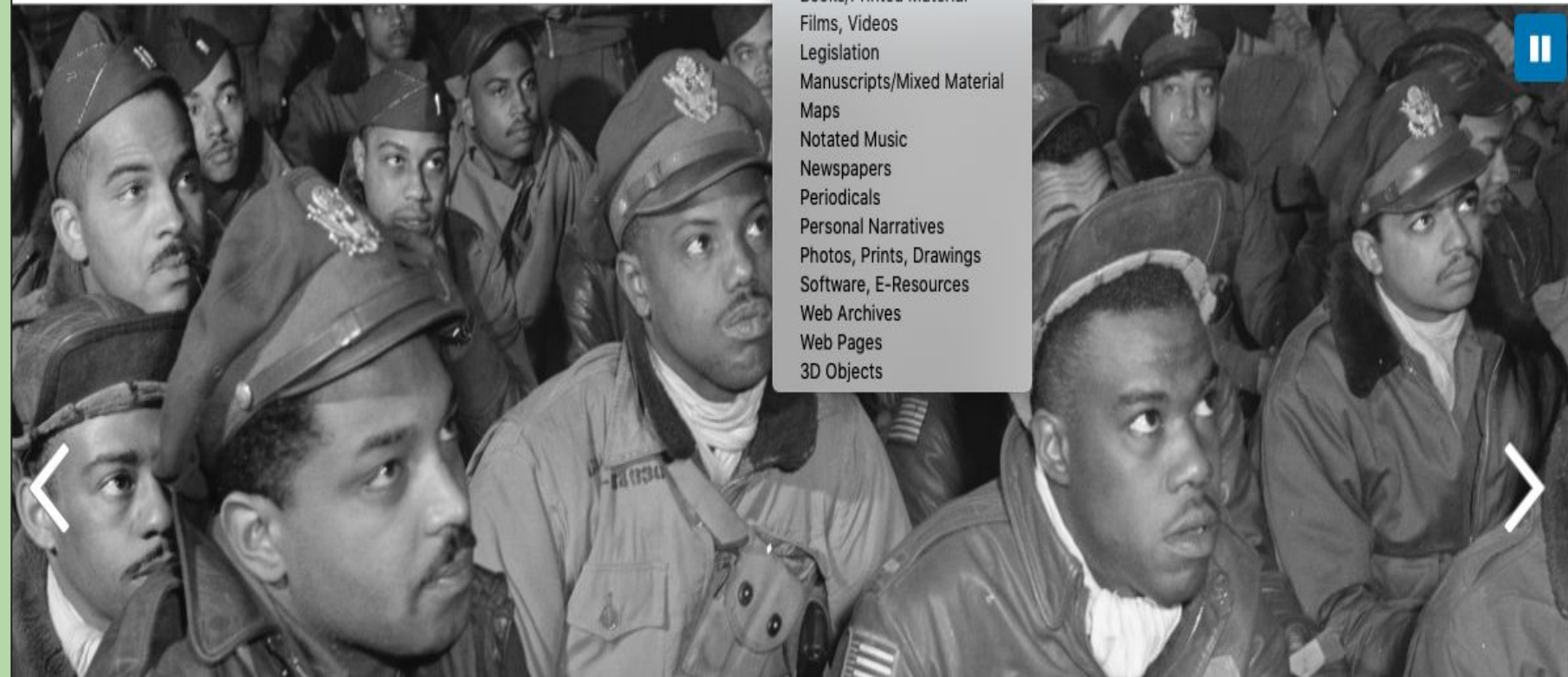
Photos, Prints, Drawings

Software, E-Resources

Web Archives

Web Pages

3D Objects



COLLECTION

Voices Remembering Slavery: Freed People Tell Their Stories

[About this Collection](#)

[Collection Items](#)

[Articles and Essays](#)

Ως ήχος ως μεταβλητή

Featured Content



Interview with Uncle Billy McCrea, Jasper, Texas, 1940 (part 2 of 2)

[Interview with Uncle Billy McCrea, Jasper, Texas, 1940 \(part 2 ...](#)



AUDIO

Interview with Fountain Hughes, Baltimore, Maryland, June 11, ...



AUDIO

Interview with Alice Gaston, Gee's Bend, Alabama, 1941



AUDIO

Interview with Dave White, St. Simons Island, Georgia, July 26, ...



AUDIO

Interview with Susan A. Quall, Johns Island, South Carolina, May...

Σαν Σήμερα... (Πρωτοσέλιδα 16ης Φεβρουαρίου)



Μακεδονία
16/02/1925



Ελευθερία
16/02/1949



Σκριπ
16/02/1912



Μακεδονία
16/02/1928



Ελευθερία
16/02/1951



Συνάντηση: 6η

Βάρος (σε κιλά)	Έτη σπουδών	Γεωγραφία
i=1 65,8	3	Νότος
i=2 70,3	5	Νότος
i=3 93,2	5	Δύση
i=4 69,7	4	Βορράς
...		
i=99 105,3	4	Ανατολή
i=100 74,1	4	Ανατολή
		
παίρνει οποιαδήποτε τιμή μέσα σε ένα διάστημα τιμών	λαμβάνει συγκεκριμένες τιμές	ποιοτική μεταβλητή
ποσοτική + συνεχής	ποσοτική + διακριτή	ποιοτική

Συνάντηση: 6η

Επίπεδα διάστασης της εργασίας
μας με δευτερογενή δεδομένα:

1) Διαστρωματικά δεδομένα

2) Δεδομένα Χρονολογικών Σειρών

3) Δεδομένα Panel

Διαστρωματικά Δεδομένα (cross-sectional data)

ποιοτικός
παράγοντας

ποσοτικές μεταβλητές

Time	Sections	ΑΕΠ (Ψ)	ΑΠΑΣΧΟΛΗΣΗ (X1)	ΕΚΠΑΙΔΕΥΣΗ (X2)	ΚΥΒΕΡΝΗΣΗ (D)	...
2009	Greece	...	...	...	1	...
2009	Italy	...	...	...	0	...
2009	Spain	...	...	...	0	...
2009	UK	...	...	...	1	...
...	...	...	...	...	1	...
2009	Germany	...	...	...	0	...

Ο χρόνος (**t**) είναι «παγωμένος».

Τα «στρώματα» (**i**) μπορεί να είναι χώρες, πόλεις, άτομα, επιχειρήσεις κλπ

n: μέγεθος του δείγματός μας $\rightarrow i = 1, 2, 3, \dots, n$

N: το μέγεθος του πληθυσμού

Συνήθως καλούμε μεγάλα δείγματα αυτά που έχουν πάνω από 25 παρατηρήσεις (**n>25**)

Δεδομένα Χρονοσειρών (time-series data)

ποιοτικός
παράγοντας

ποσοτικές μεταβλητές



Time	Sections	Τιμή Μετοχής (Ψ)	Γενικός Δείκτης (Χ1)	Επενδύσεις (Χ2)	ΒΡΟΧΗ (D)	...
1/1/22	OTE	2,02	...	...	1	...
2/1/22	OTE	1,68	...	...	0	...
3/1/22	OTE	2,01	...	...	0	...
4/1/22	OTE	1,90	...	...	1	...
...	...	...	...	...	1	...
20/10/22	OTE	1,32	...	...	0	...

Το «στρώμα» **(i)** είναι «παγωμένο». Ο χρόνος (t) αλλάζει!!!

Τα «στρώματα» **(i)** μπορεί να είναι χώρες, πόλεις, άτομα, επιχειρήσεις κλπ

n: μέγεθος του δείγματός μας $\rightarrow i = 1, 2, 3, \dots, n$

N: το μέγεθος του πληθυσμού

Συνήθως καλούμε μεγάλα δείγματα αυτά που έχουν πάνω από 25 παρατηρήσεις (**n>25**)

Συνδυασμός Διαστρωματικών
Δεδομένων
& Δεδομένων Χρονολογικών Σειρών
(panel data)

ποιοτικός
παράγοντας

Time	Sections	Γονιμότητα (Ψ)	Ανισότητα (X1)	Προσδόκιμο Ζωής (X2)	Θρησκεία (D)	...
1998	Ευρώπη	2,02	...	...	1	...
1999	Ευρώπη	1,68	...	...	0	...
2000	Ευρώπη	2,01	...	...	0	...
1998	Ασία	2,90	...	...	1	...
1999	Ασία	2,80	...	...	1	...
2000	Ασία	2,72	...	...	0	...

Τόσο τα «στρώματα» (**i**) όσο και ο χρόνος (**t**) αλλάζουν!!!

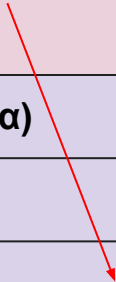
Τα «στρώματα» (**i**) μπορεί να είναι χώρες, πόλεις, άτομα, επιχειρήσεις κλπ

n: μέγεθος του δείγματός μας → **i = 1,2,3, ... , n**

N: το μέγεθος του πληθυσμού

Συνήθως καλούμε μεγάλα δείγματα αυτά που έχουν πάνω από 25 παρατηρήσεις (**n>25**)

παγωμένος
χρόνος



Διαστρωματικά Δεδομένα

i (στρώμα)	t (έτη)	X1	X2	X3
Ελλάδα	2020	X(1,1)	X(2,1)	X(3,1)
Ιταλία	2020	X(1,2)	X(2,2)	X(3,2)
Ισπανία	2020	X(1,3)	X(2,3)	X(3,3)
Γερμανία	2020	X(1,4)	X(2,4)	X(3,4)
Γαλλία	2020	X(1,5)	X(2,5)	X(3,5)
...	...	...	...	...
Ολλανδία	2020	X(1,n)	X(2,n)	X(3,n)

παγωμένο
στρώμα

Χρονοσειρές

i (στρώμα)	t (έτη)	X1	X2	X3
Ελλάδα	1999	X(1,1)	X(2,1)	X(3,1)
Ελλάδα	2000	X(1,2)	X(2,2)	X(3,2)
Ελλάδα	2001	X(1,3)	X(2,3)	X(3,3)
Ελλάδα	2002	X(1,4)	X(2,4)	X(3,4)
Ελλάδα	2003	X(1,5)	X(2,5)	X(3,5)
...	...	...	...	...
Ελλάδα	2024	X(1,n)	X(2,n)	X(3,n)

στρώματα και χρόνος
μεταβάλλονται

Δεδομένα Πάνελ

i	t	X1	Z2
Ελλάδα	1990	X1	Z1
Ελλάδα	1991	X2	Z2
Γαλλία	1990	X3	Z3
Γαλλία	1991	X4	Z4
Ιταλία	1990	X5	Z5
Ιταλία	1991	X6	Z6
Ολλανδία	1990	X7	Z7
Ολλανδία	1991	X8	Z8

Year	Area	Municipality	males	females	non-greeks	literate males	literate females	male students	female students	male teachers	female teachers	landowners	owners of manufacturing	merchants	#
1870	south	Abia	2376	2647	0	869	11	229	2	6	0	277	113	43	
1870	central	Acharne	1523	1392	5	210	14	72	20	2	1	31	84	17	
1870	island	acrolifites	1201	1109	6	504	14	102	0	2	0	153	65	41	
1870	island	adamas	347	366	41	180	88	56	24	0	0	4	23	29	
1870	island	Aedipos	1157	1149	16	148	0	110	0	3	0	62	36	17	
1870	south	Aegire	1248	1146	0	663	24	183	2	4	0	148	58	14	
1870	central	Aegition	1603	1514	0	359	12	92	4	2	0	15	29	9	
1870	central	Agria	2912	2865	44	702	3	164	3	5	0	85	68	254	
1870	central	Agrinio	3215	2890	524	1185	308	293	62	5	1	154	238	170	
1870	south	Aigio	4568	4286	26	1776	204	556	61	13	2	743	401	108	
1870	central	Aitoliko	1719	1370	100	603	105	203	71	1	1	190	224	18	
1870	central	Akraephinio	928	765	46	50	1	14	0	3	0	1	33	0	
1870	south	Alagonia	1858	1982	0	894	16	196	2	5	0	451	17	25	
1870	south	Alea	1332	1203	0	127	0	131	0	3	0	6	4	0	
1870	south	Alifeira	1957	1702	0	64	9	54	0	2	0	3	21	2	
1870	island	Alonisos	190	170	0	19	3	20	0	1	0	2	2	1	
1870	south	Amfeia	1719	1730	0	449	37	99	9	2	0	188	15	16	
1870	island	amfipagites	2230	2130	0	464	22	49	0	3	0	307	31	35	
1870	central	Amfissa	3168	3066	0	1332	134	445	44	5	1	147	433	196	
1870	island	amorgos	1058	1140	1	311	102	130	46	3	1	52	59	15	
1870	central	Amvrakia Tri	893	838	0	160	0	15	0	1	0	0	2	1	
1870	central	Amvrakia Va	2056	1908	90	457	21	95	41	2	1	50	162	54	
1870	central	Anactorio	3072	2792	414	667	34	189	36	5	1	70	305	99	
1870	island	anaphi	326	403	4	144	32	68	12	1	0	31	28	4	
1870	south	Andanias	2637	2479	0	557	67	211	3	2	0	118	22	10	
1870	south	Andritsaina	3644	3701	0	775	157	253	52	11	1	120	292	73	
1870	island	andros	3619	3821	48	1200	383	392	68	6	4	115	236	105	
1870	island	Angistri	216	204	0	74	0	24	0	1	0	8	1	0	
1870	central	Antikyra	849	916	0	169	2	74	0	1	0	35	9	8	
1870	island	aperanthia	880	840	6	164	1	49	0	2	0	7	30	3	
1870	central	Aperantia	2096	1906	82	308	9	120	1	3	0	71	33	9	
1870	island	apiliotes	2818	2775	0	965	29	238	3	6	0	205	127	98	
1870	central	Apoditias	2275	2549	0	656	5	152	0	5	0	139	56	17	
1870	island	apollones	1519	1468	2	494	0	80	0	2	0	82	79	2	
1870	central	Arachova	1394	1337	0	250	42	110	0	2	0	37	71	13	
1870	central	Arakynthia	1581	1594	0	520	4	141	0	4	0	91	50	21	
1870	island	arcadies	1674	1492	2	257	33	67	3	4	1	53	34	4	

Cross-Sectional (Secondary) Data

year	(per 1,000 people)			status before marriage % (men)			status before marriage % (women)			educational level at marriage %						
	marriages	births	deaths	single men	widowers	divorced men	single women	widows	divorced women	illiterate men	literate men	high sch men	university men	illiterate women	literate women	high women
1956	6,9	19,7	7,4	93,9	3,4	2,6	95,9	2	2	3,4	83,9	8,7	3,3	9,2	82,4	6,3
1957	8,5	19,3	7,6	94,2	3	2,5	96,2	1,7	1,8	2,8	83,2	9	4	7,4	82,9	7,2
1958	8,5	19	7,1	94,1	3,1	2,4	96,4	1,5	1,8	2,6	83,8	9,7	4	7	83,3	7,8
1959	8,9	19,4	7,4	94,6	2,7	2,5	96,7	1,4	1,8	2,2	82,4	10,7	4,5	6,1	83	9,2
1960	7	18,9	7,3	94,2	3	2,7	96,4	1,4	2,1	2,6	82,4	10,3	4,6	6,6	82,9	8,6
1961	8,4	17,9	7,6	94,4	2,8	2,7	96,5	1,4	2,1	2	80,9	11,7	5,3	5,4	82,5	10,1
1962	8,4	18	7,9	94,4	2,7	2,8	96,6	1,3	2,1	1,9	80,9	11,8	5,3	5	82,4	10,6
1963	9,2	17,5	7,9	94,7	2,5	2,7	96,8	1,2	2	1,6	79,5	13,2	5,6	4	82,2	11,6
1964	8,9	18	8,2	94,5	2,7	2,8	96,7	1,2	2,1	1,4	80,1	12,7	5,6	3,9	82,6	11,4
1965	9,4	17,7	7,9	94,6	2,4	3	96,8	1,1	2,1	1,3	78,2	14,2	6,3	3,4	81,5	12,7
1966	8,3	17,9	7,9	94,2	2,8	3	96,5	1,2	2,3	1,2	77,9	14,7	6,2	2,9	81,3	13,2
1967	9,4	18,7	8,3	94,1	2,7	3,2	96,5	1,2	2,3	1	75,5	16,5	7	2,5	79,3	15,4
1968	7,4	18,2	8,3	93,6	2,8	3,6	96,3	1,2	2,5	1,1	76,8	15,4	6,7	2,5	80,4	14,3
1969	8,2	17,4	8,1	94	2,6	3,4	96,7	1	2,3	0,8	75	16,9	7,3	1,9	78,4	16,6
1970	7,7	16,5	8,4	93,6	2,9	3,5	96,2	1,2	2,6	0,9	73,6	17,7	7,8	1,8	76,9	17,7
1971	8,3	16	8,4	93,9	2,6	3,5	96,4	1,1	2,5	0,8	71,4	19,2	8,6	1,5	74,8	19,7
1972	6,8	15,9	8,6	93,2	2,9	3,9	96,1	1,2	2,7	0,9	72	18,3	8,8	1,6	75,2	19
1973	8,3	15,4	8,7	93,8	2,5	3,7	96,2	1,2	2,6	0,7	69,2	20,2	9,9	1,1	72,4	21,2
1974	7,6	16,1	8,5	93,4	2,6	4	96	1,1	2,9	0,7	67,3	21,6	10,4	1,1	69,7	23,4
1975	8,4	15,7	8,8	93,5	2,5	4	96,2	1,1	2,7	0,6	65,2	22,6	11,6	0,9	67,4	25,4
1976	6,9	16	8,9	92,9	2,6	4,5	95,8	1,1	3,1	0,6	65,7	22,9	10,8	0,9	67,7	25,3
1977	8,2	15,5	9	93,5	2,3	4,2	96,1	1	2,9	0,5	63,1	24,7	11,7	0,8	64,6	28
1978	7,7	15,7	8,7	93,4	2,2	4,5	96	1	3	0,6	62,8	24,5	12,1	0,7	63,7	28,6
1979	8,4	15,7	8,7	93,5	2	4,5	96	0,9	3,1	0,9	61	25,8	12,3	1,1	60,4	31,2
1980	6,5	15,4	9,1	92,5	2,4	5,1	95,3	1,2	3,5	1,2	60,4	25,7	12,7	1,4	60	30,7
1981	7,3	14,5	8,9	92,3	2,1	5,6	95,1	1,1	3,8	0,7	57,4	28,4	13,5	0,9	56,9	33,5
1982	6,9	14	8,8	91,5	2,2	6,3	94,7	1	4,3	0,8	54,6	31,2	13,4	1	53,6	37,2
1983	7,2	13,5	9,2	91,4	2,2	6,4	94,3	1	4,7	0,5	50,4	34,9	14,2	0,8	50,5	40
1984	5,5	12,7	8,9	90	2,6	7,4	93,1	1,5	5,4	1	54,5	30,5	14	1,2	52,5	37,3
1985	6,4	11,7	9,4	90,1	2	7,9	93,1	1,1	5,8	0,7	50	34,2	15,1	0,8	48,6	40,2
1986	5,8	11,3	9,2	89,6	2,3	8,1	92,5	1,3	6,2	0,7	49	35,1	15,2	0,8	46,8	41,4
1987	6,6	10,6	9,6	89,6	2,1	8,3	92,4	1,2	6,4	0,8	44,7	38,8	15,7	1	41,8	45,2
1988	4,8	10,7	9,2	88,4	2,2	9,4	91,5	1,2	7,3	0,9	36,5	46,2	16,4	1	32,1	52,7
1989	6,1	10,1	9,2	89,5	1,8	8,7	91,9	1,1	7	0,6	30	51,7	17,7	0,8	25,4	58
1990	5,8	10,1	9,3	89,3	1,8	8,9	91,8	1	7,2	0,7	27,5	53,5	18,3	0,8	23,3	58,7
1991	6,4	10	9,3	89,7	1,7	8,6	91,9	1,1	7	0,8	25,2	56,2	17,8	1,1	21	60,8

Time-Series (Secondary) Data

B	C	D	E	F	G	H	I
geographical area	region	urbanization %	births outside marriage %	child's death on birth (per 1.000 live births)	crude birth rate (per 1.000 people)	crude death rate (per 1.000 people)	
South	Achaia & Elis	1922	0,055931631	0,004192244	15,57119329	23,91652313	14,4272
South	Achaia & Elis	1923	0,055451885	0,011900139	18,86269071	23,83187455	17,8325
South	Achaia & Elis	1924	0,051917027	0,027587952	38,21817261	12,80298121	15,2106
South	Achaia & Elis	1925	0,054043394	0,012316384	9,039548023	28,57973261	16,0563
South	Achaia & Elis	1926	0,054025665	0,013080523	8,040321613	26,56020909	14,0371
South	Achaia & Elis	1927	0,052898132	0,012831077	7,769276045	26,76264886	17,0027
South	Achaia & Elis	1928	0,053727164	0,012791926	11,03156906	26,39405137	17,9
South	Achaia & Elis	1929	0,053556503	0,012458567	6,743627843	27,09088094	18,3650
South	Achaia & Elis	1930	0,053498792	0,011068237	7,920389927	30,16787158	15,8865
South	Achaia & Elis	1931	0,053437212	0,011598212	8,831666312	28,42108446	17,3526
South	Achaia & Elis	1932	0,053395189	0,012319168	7,911392405	26,47159617	16,2817
South	Achaia & Elis	1933	0,053319983	0,012328922	6,899671983	26,18634489	16,6340
South	Achaia & Elis	1934	0,053258371	0,021863404	5,917159763	29,14133739	13,4498
South	Achaia & Elis	1935	0,053888456	0,011876226	7,844846372	26,46558436	13,584
South	Achaia & Elis	1936	0,0539121	0,011945205	4,493150685	25,94835921	11,4428
South	Achaia & Elis	1937	0,05403624	0,012278923	8,11084826	24,94940978	13,4935
South	Achaia & Elis	1938	0,05435106	0,012533057	5,749108888	24,1677319	12,4326
Central	Aetolia and Acarnania	1922	0,038793015	0,007577352	12,20795622	24,52875973	12,4992
Central	Aetolia and Acarnania	1923	0,021209609	0,006395907	23,58464982	27,02305645	19,2799
Central	Aetolia and Acarnania	1924	0,035631971	0,004653869	21,23327516	16,23229462	10,1888
Central	Aetolia and Acarnania	1925	0,036917261	0,007664234	19,16058394	25,50649081	13,6860
Central	Aetolia and Acarnania	1926	0,036824722	0,00470984	11,26997477	27,79985971	12,1019
Central	Aetolia and Acarnania	1927	0,036184775	0,006108066	10,3367267	29,40634643	14,668
Central	Aetolia and Acarnania	1928	0,036878422	0,006733479	11,27466333	29,03651162	15,6959
Central	Aetolia and Acarnania	1929	0,036876791	0,009583134	17,08992174	28,13577641	17,9025
Central	Aetolia and Acarnania	1930	0,036898827	0,008950484	6,92940667	30,7661559	14,6968
Central	Aetolia and Acarnania	1931	0,036948874	0,009940067	10,57859962	29,92039888	15,3472
Central	Aetolia and Acarnania	1932	0,037020707	0,011106108	7,50412727	28,75204971	16,1474
Central	Aetolia and Acarnania	1933	0,03702645	0,008799879	7,586102261	28,11260397	17,0612
Central	Aetolia and Acarnania	1934	0,037078199	0,01077903	6,73689368	34,27228076	13,5426
Central	Aetolia and Acarnania	1935	0,037701252	0,011355768	6,785763745	29,76259171	13,7169
Central	Aetolia and Acarnania	1936	0,037761159	0,010248268	5,629330254	28,12715683	13,9539
Central	Aetolia and Acarnania	1937	0,037975722	0,01051129	4,931222424	30,81783643	14,901
Central	Aetolia and Acarnania	1938	0,03835506	0,008879781	8,606557377	28,82457177	13,7743
South	Arcadia	1922	0,026557125	0,009247028	11,22853369	22,83594021	13,2430
South	Arcadia	1923	0,028962891	0,010877193	25,61403509	18,09857053	11,8116

Panel (Secondary) Data

Συνάντηση: 6η

Πήραμε δεδομένα από τη βάση δεδομένων της Ευρωπαϊκής Επιτροπής:

https://dashboard.tech.ec.europa.eu/qs_digit_dashboard_mt/public/sense/app/667e9fba-eea7-4d17-abf0-ef20f6994336/sheet/f38b3b42-402c-44a8-9264-9d422233add2/state/analysis/

παρατηρήσεις	διαστρωματικά		χώρες	μισθοί
	έτη	ανεργία (%)		
1	2023	4	Bulgaria	-
2	2023	3	Czechia	105
3	2023	5	Denmark	174
4	2023	3	Germany	1846
5	2023	6	Estonia	15
6	2023	4	Ireland	121
7	2023	11	Greece	60
8	2023	12	Spain	558
9	2023	7	France	1097
10	2023	6	Croatia	34

ανεργία: *Percentage of active population (%)*

μισθοί: *Gross wages and salaries (Mrd ECU/EUR)*

n=10 παρατηρήσεις (**n=9** μετά την αφαίρεση της Βουλγαρία)

Συνάντηση: 6η

Προσπάθησε να υπολογίσεις την τιμή του συντελεστή γραμμικής συσχέτισης !!!

Πρέπει να βγει μια τιμή κοντά στο $\rho = -0.16$

	διαστρωματικά			
παρατηρήσεις	έτη	ανεργία (%)	χώρες	μισθοί
1	2023	4	Bulgaria	-
2	2023	3	Czechia	105
3	2023	5	Denmark	174
4	2023	3	Germany	1846
5	2023	6	Estonia	15
6	2023	4	Ireland	121
7	2023	11	Greece	60
8	2023	12	Spain	558
9	2023	7	France	1097
10	2023	6	Croatia	34

Το πρόσημο του συντελεστή συσχέτισης δείχνει μια **αρνητική σχέση**.

Ωστόσο, η τιμή του είναι **αρκετά χαμηλή !!!**

Συνάντηση: 6η

Δοκίμασε να βάλεις τις τιμές του πίνακα στις στήλες της εφαρμογής:

<https://www.statskingdom.com/correlation-calculator.html>

Results	
Parameter	Value
Pearson correlation coefficient (r)	-0.1746
r^2	0.03048
P-value	0.6533
Covariance	-357.5833
Sample size (n)	9
Statistic	-0.4691

→ συντελεστής συσχέτισης

→ συντελεστής προσδιορισμού

→ $p\text{-value} > 0,05$ (5%)

Άρα, **ΔΕΝ** μπορώ να απορρίψω την H_0
πως ο συντελεστής γραμμικής συσχέτισης
είναι **στατιστικά μηδέν !!!**

Συνάντηση: 6η

ΣΥΜΠΕΡΑΣΜΑ:

Αν και το πρόσημο της σχέσης που συνδέει την ανεργία με τους μισθούς φαίνεται να είναι **αρνητικό**, η ένταση αυτής της σχέσης δεν είναι αρκετά έντονη (βλ. **χαμηλή τιμή** του συντελεστή συσχέτισης).

Την ίδια στιγμή, ο **έλεγχος για την στατιστική σημαντικότητα του συντελεστή συσχέτισης** δίνει τιμή του **p-value > 0,05 (5%)**. Αυτό σημαίνει πως **δεν είμαστε σε θέση να απορρίψουμε την H_0 (μηδενική υπόθεση) της μη στατιστικής σημαντικότητας του συντελεστή γραμμικής συσχέτισης !!!**

ΠΡΟΣΟΧΗ: Αυτή είναι η εικόνα που προκύπτει από το δείγμα σου για **n=9** παρατηρήσεις ... (μικρό δείγμα)

Συνάντηση: 6η

		y	x
		cons	wages
1	2023	2462	105
2	2023	1127	174
3	2023	1772	1846
4	2023	14	15
5	2023	113	121
6	2023	141	60
7	2023	694	558
8	2023	1300	1097
9	2023	35	34
10	2023	1038	610
11	2023	17	11
12	2023	17	17
13	2023	28	34
14	2023	21	24
15	2023	22832	74
16	2023	8	7
17	2023	359	369
18	2023	190	192
19	2023	1376	242
20	2023	137	100
21	2023	656	121
22	2023	26	29
23	2023	53	40
24	2023	120	110
25	2023	2214	212

Τι συμβαίνει όταν εξετάζεις τη σχέση ανάμεσα στην **κατανάλωση** (Ψ) και τους **μισθούς** (X) για $n=25$ παρατηρήσεις ; ; ;

Δοκίμασε να εξάγεις τον συντελεστή συσχέτισης καταχωρώντας τις τιμές στις στήλες της εφαρμογής:

<https://www.statskingdom.com/correlation-calculator.html>

Περιμέναμε μια **έντονα θετική σχέση**, αλλά η τιμή του συντελεστή βγαίνει πάρα πολύ μικρή (δες στην επόμενη διαφάνεια)

ΑΙΤΙΑ ; ; ;

Συνάντηση: 6η

Τι συμβαίνει όταν εξετάζεις τη σχέση ανάμεσα στην **κατανάλωση** (Ψ) και τους **μισθούς** (X) για $n=25$ παρατηρήσεις ; ; ;

$$\rho = -0,003316$$

Αιτία ; ; ; (δε το είδαμε στην τάξη ...)

Η τιμή για την Ουγγαρία είναι στο εθνικό της νόμισμα !!! Αυτό την κάνει να αποκλίνει σημαντικά από τις άλλες τιμές.

Δείτε στην επόμενη διαφάνεια τι θα συμβαίνει αν βγάλω την τιμή για την Ουγγαρία ...

Results

Parameter	Value
Pearson correlation coefficient (r)	-0.003316
r^2	0.000011
P-value	0.9874
Covariance	-6214.0833
Sample size (n)	25
Statistic	-0.0159

Συνάντηση: 6η

Τι συμβαίνει όταν εξετάζεις τη σχέση ανάμεσα στην **κατανάλωση** (Ψ) και τους **μισθούς** (X) για $n=25$ παρατηρήσεις ; ; ;

Results	
Parameter	Value
Pearson correlation coefficient (r)	0.5074
r^2	0.2575
P-value	0.01137
Covariance	161934.9855
Sample size (n)	24
Statistic	2.762

$\rho = + 0,50$ (έντονα θετική συσχέτιση)

Τώρα, βλέπουμε το αναμενόμενο !!!

Μια έντονα θετική σχέση ανάμεσα στην κατανάλωση (Ψ) και το εισόδημα (X) !!!

Μάλιστα, αυτή τη φορά ο έλεγχος της στατιστικής σημαντικότητας του συντελεστή συσχέτισης πετυχαίνει να απορρίψει την μηδενική υπόθεση της μη στατιστικής σημαντικότητας του συντελεστή προσδιορισμού

Συνάντηση: 6η

	y		x1	x3	
	con	wag	inv	unemp	
1	2462		105	1561	3
2	1127		174	514	5
3	1772		1846	661	3
4	14		15	8	6
5	113		121	94	4
6	141		60	31	11
7	694		558	239	12
8	1300		1097	539	7
9	35		34	13	6
10	1038		610	421	8
11	17		11	6	6
12	17		17	7	7
13	28		34	13	7
14	21		24	11	5
15	22832		74	10016	4
16	8		7	3	4
17	359		369	170	4
18	190		192	91	5
19	1376		242	476	3
20	137		100	41	7
21	656		121	258	6
22	26		29	10	4
23	53		40	21	6
24	120		110	50	7
25	2214		212	1177	8

Τι συμβαίνει όταν περνάμε στο μοντέλο με **k=2** ανεξάρτητες μεταβλητές ; ; ;

Ψ: κατανάλωση

X1: επενδύσεις

X3: ανεργία

Δοκίμασε να τρέξεις την πολλαπλή παλινδρόμηση εδώ:

https://www.statskingdom.com/410_multi_linear_regression.html

Συνάντηση: 6η

Τι συμβαίνει όταν περνάμε στο μοντέλο με **k=2** ανεξάρτητες μεταβλητές ; ; ;

Δες τα αποτελέσματα που δίνει η εφαρμογή !!!

$$\text{κατανάλωση} = -18.545992 + 2.264844 \text{ επενδύσεις}$$

Ερμηνεία Συντελεστών Παλινδρόμησης:

Αν αυξηθούν οι επενδύσεις κατά 1 μονάδα τότε η κατανάλωση θα αυξηθεί κατά περίπου 2,26 μονάδες.

Η εφαρμογή πετάει έξω τον παράγοντα “ανεργία” ως μη στατιστικά σημαντικό.

Συνάντηση: 6η

Τι συμβαίνει όταν περνάμε στο μοντέλο με **k=2** ανεξάρτητες μεταβλητές ; ; ;

Δες τα αποτελέσματα που δίνει η εφαρμογή !!!

Correlation matrix (pearson)			
	κατανάλωση	επενδύσεις	ανεργία
κατανάλωση	1	0.998366	-0.198939
επενδύσεις	0.998366	1	-0.207961
ανεργία	-0.198939	-0.207961	1

Πίνακας τιμών του συντελεστή γραμμικής συσχέτισης:

Φαίνεται η σχεδόν **ΤΕΛΕΙΑ ΘΕΤΙΚΗ** γραμμική σχέση ανάμεσα σε κατανάλωση και επενδύσεις ($\rho = + 0,99$)

Σε αντίθεση με τις επενδύσεις, η ανεργία φαίνεται να σχετίζεται αρνητικά με την κατανάλωση, αν και όχι ιδιαίτερα έντονα ($\rho = - 0,19$).

Συνάντηση: 6η

Τι συμβαίνει όταν περνάμε στο μοντέλο με **k=2** ανεξάρτητες μεταβλητές ; ; ;

Δες τα αποτελέσματα που δίνει η εφαρμογή !!!

ANOVA table

Source	DF	Sum of Square	Mean Square	F Statistic	P-value
Regression (between $\hat{y}_i$ and $\bar{y}$)	1	486866400	486866400	7020.459076	0
Residual (between y_i and $\hat{y}_i$)	23	1595042.016	69349.65288		
Total (between y_i and $\bar{y}$)	24	488461442	20352560.08		

Ο πίνακας αυτός αποδίδει τα αθροίσματα τετραγώνων (βλ. συνολικό-**SST**, παλινδρόμησης-**SSR**, σφάλματος-**SSE**).

Ο συντελεστής προσδιορισμού δίνεται από τον τύπο:

$$\mathbf{R-squared=SSR/SST=1-SSE/SST}$$

Συνάντηση: 6η

Τι συμβαίνει όταν περνάμε στο μοντέλο με **k=2** ανεξάρτητες μεταβλητές ; ; ;

Δες τα αποτελέσματα που δίνει η εφαρμογή !!!

ANOVA table

Source	DF	Sum of Square	Mean Square	F Statistic	P-value
Regression (between $\hat{y}_i$ and $\bar{y}$)	1	486866400	486866400	7020.459076	0
Residual (between y_i and $\hat{y}_i$)	23	1595042.016	69349.65288		
Total (between y_i and $\bar{y}$)	24	488461442	20352560.08		

Ακόμη, ο πίνακας αποδίδει την στατιστική τιμή F του ελέγχου της από κοινού σημαντικότητας των συντελεστών της παλινδρόμησης.

Το **p-value = 0 < 0,05** (5%). Αυτό σημαίνει πως απορρίπτουμε την μηδενική υπόθεση της από κοινού μη στατιστικής σημαντικότητας των συντελεστών β της παλινδρόμησης.

Συνάντηση: 6η

Τι συμβαίνει όταν περνάμε στο μοντέλο με **k=2** ανεξάρτητες μεταβλητές ; ; ;

Δες τα αποτελέσματα που δίνει η εφαρμογή !!!

Coefficient Table Iteration 1 (adjusted R-squared = 0.997)

	Coeff	SE	t-stat	lower $t_{0.025}(22)$	upper $t_{0.975}(22)$	Stand Coeff	p-value	VIF
b	-128.051231	158.740047	-0.806673	-457.25794	201.155478	0	0.42849	
επενδύσεις	2.269125	0.0279129	81.293048	2.211237	2.327013	1.000253	0	1.045203
ανεργία	18.022219	24.436622	0.737509	-32.656233	68.700671	0.00907452	0.468607	1.045203

Τέλος, η εφαρμογή αποδίδει τα αποτελέσματα των ελέγχων της στατιστικής σημαντικότητας **MEMONΩΜΕΝΑ** για τους συντελεστές β της παλινδρόμησης. Με βάση τις τιμές των **p-value** ΜΟΝΟ οι “επενδύσεις” φαίνεται να είναι **στατιστικά σημαντικός** παράγοντας.