

Υπολογιστική Εργασία στο IBM SPSS

Ανάλυση Δεδομένων Ημερήσιας
Θνητότητας στην Μολδαβία και το
Ηνωμένο Βασίλειο

Τσαμουρίδης Αναστάσιος Αθανάσιος

Ιούλιος 2021, 4^ο Εξάμηνο
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών
Υπολογιστών
Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

Περιεχόμενα

1	Εισαγωγή	2
2	Συλλογή δεδομένων	2
2.1	Επιλογή δείγματος	2
3	Μελέτη Α	6
3.1	Σχολιασμός της Κατανομής του Ημερήσιου Ποσοστού Θνητότητας	6
3.2	Μελέτη της Μέσης Τιμής του Ημερήσιου Ποσοστού Θνητότητας .	8
3.3	Σύγκριση της Μέσης Τιμής του Ημερήσιου Ποσοστού Θνητότητας	10
4	Μελέτη Β	11
4.1	Συσχέτιση Ημερήσιων Κρουσμάτων και Θανάτων με τον Αύξοντα Αριθμό Ημέρας μετά την Κορύφωση του Κύματος	11
4.2	Μοντέλο Γραμμικής Παλινδρόμησης Ημερήσιων Κρουσμάτων και Θανάτων με τον Αύξοντα Αριθμό Ημέρας μετά την Κορύφωση του Κύματος	14
4.3	Πρόβλεψη Θανάτων κάποια Ημέρα t , γνωρίζοντας τα Ημερήσια Κρούσματα τ Ημέρες Νωρίτερα	17
4.4	Εκτίμηση Μοντέλου Παλινδρόμησης για την εξάρτηση του αριθμού ημερήσιων νέων θανάτων από τον αριθμό των ημερήσιων νέων κρουσμάτων τ ημέρες	21
5	Παράρτημα	23
5.1	Κατάλογοι Σχημάτων και Πινάκων	23
6	Πηγές	25

1 Εισαγωγή

Στο παρόν έγγραφο, γίνεται ανάλυση της θνητότητας στις χώρες της Μολδαβίας και του Ηνωμένου Βασιλείου και τα στατιστικά των δύο χωρών συγκρίνονται.

2 Συλλογή δεδομένων

Για την ανάλυση του ποσοστού θνητότητας της νόσου Covid-19 επιλέχθηκαν στατιστικά δεδομένα από την Μολδαβία και το Ηνωμένο Βασίλειο. Τα δεδομένα των δύο χωρών αναφέρονται, συχνά, ως δείγμα A και δείγμα B, αντίστοιχα. Τα δεδομένα που χρησιμοποιούνται προέρχονται από το: <https://www.worldometers.info/coronavirus/>.

2.1 Επιλογή δείγματος

Για τις δύο χώρες επιλέχθηκε ένα δείγμα μεγέθους 45 ημερών (ένας και μισός μήνας, θεωρώντας ότι ο μήνας έχει, κατά μέσο όρο, 30 ημέρες). Ως αρχή του δείγματος για τους καθημερινούς θανάτους επιλέχθηκε η κορύφωση του κύματος για το 2021. Ομοίως, για τα ημερήσια κρούσματα. Συγκεκριμένα:

1. Για την Μολδαβία:

- Ως πρώτη ημέρα για την καταγραφή ημερήσιων κρουσμάτων ορίζεται η 24^η Μαρτίου (22073 κρούσματα).
- Ως πρώτη ημέρα για την καταγραφή ημερήσιων θανάτων ορίζεται η 30^η Μαρτίου (46 θάνατοι). Εδώ πρέπει να σημειωθεί ότι ο ίδιος αριθμός ημερήσιων θανάτων παρουσιάστηκε και στις 6 Απριλίου, ωστόσο επιλέχθηκε ως κορύφωση η πρώτη εμφάνιση του μεγίστου αριθμού θανάτων.

2. Για το Ηνωμένο Βασίλειο:

- Ως πρώτη ημέρα για την καταγραφή των ημερήσιων κρουσμάτων ορίζεται η 8^η ημέρα του Ιανουαρίου (67803 κρούσματα).
- Ως πρώτη ημέρα για την καταγραφή ημερήσιων θανάτων ορίζεται η 20^η Ιανουαρίου (1823 θάνατοι).

Για τις δύο χώρες, στα δεδομένα δεν εμφανίζεται κάποια προβληματική ή ακραία τιμή. Στον Πίνακα 1 και στον Πίνακα 2 παρατίθενται τα δεδομένα που συλλέχθηκαν.

Πίνακας 1: Δεδομένα Μολδαβίας

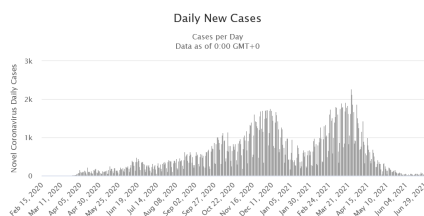
Ημ. Κρουσμάτων	Κρούσματα	Ημ. Θανάτων	Θάνατοι	Ποσοστό Θνητότητας
24-Μαρ-21	2273	30-Μαρ-21	46	2,023757
25-Μαρ-21	2132	31-Μαρ-21	45	2,110694
26-Μαρ-21	1866	1-Απρ-21	43	2,304394
27-Μαρ-21	1674	2-Απρ-21	45	2,688172
28-Μαρ-21	861	3-Απρ-21	44	5,110337
29-Μαρ-21	932	4-Απρ-21	44	4,72103
30-Μαρ-21	917	5-Απρ-21	42	4,580153
31-Μαρ-21	1871	6-Απρ-21	46	2,458578
1-Απρ-21	1515	7-Απρ-21	44	2,90429
2-Απρ-21	1380	8-Απρ-21	39	2,826087
3-Απρ-21	1258	9-Απρ-21	32	2,54372
4-Απρ-21	693	10-Απρ-21	30	4,329004
5-Απρ-21	703	11-Απρ-21	19	2,702703
6-Απρ-21	773	12-Απρ-21	21	2,716688
7-Απρ-21	1155	13-Απρ-21	29	2,510823
8-Απρ-21	1428	14-Απρ-21	28	1,960784
9-Απρ-21	910	15-Απρ-21	27	2,967033
10-Απρ-21	830	16-Απρ-21	28	3,373494
11-Απρ-21	331	17-Απρ-21	27	8,1571
12-Απρ-21	603	18-Απρ-21	23	3,814262
13-Απρ-21	544	19-Απρ-21	20	3,676471
14-Απρ-21	1001	20-Απρ-21	25	2,497502
15-Απρ-21	812	21-Απρ-21	27	3,325123
16-Απρ-21	698	22-Απρ-21	25	3,581662
17-Απρ-21	628	23-Απρ-21	23	3,66242
18-Απρ-21	403	24-Απρ-21	18	4,466501
19-Απρ-21	341	25-Απρ-21	21	6,158358
20-Απρ-21	453	26-Απρ-21	15	3,311258
21-Απρ-21	700	27-Απρ-21	17	2,428571
22-Απρ-21	618	28-Απρ-21	18	2,912621
23-Απρ-21	509	29-Απρ-21	15	2,946955
24-Απρ-21	380	30-Απρ-21	17	4,473684
25-Απρ-21	241	01-Μάιος-2021	14	5,809129
26-Απρ-21	246	02-Μάιος-2021	12	4,878049
27-Απρ-21	329	03-Μάιος-2021	12	3,647416
28-Απρ-21	424	04-Μάιος-2021	19	4,481132
29-Απρ-21	370	05-Μάιος-2021	23	6,216216
30-Απρ-21	329	06-Μάιος-2021	20	6,079027
01-Μάιος-2021	323	07-Μάιος-2021	17	5,263158
02-Μάιος-2021	175	08-Μάιος-2021	14	8
03-Μάιος-2021	43	09-Μάιος-2021	9	20,930233
04-Μάιος-2021	126	10-Μάιος-2021	6	4,761905
05-Μάιος-2021	316	11-Μάιος-2021	12	3,797468
06-Μάιος-2021	333	12-Μάιος-2021	11	3,303303
07-Μάιος-2021	260	13-Μάιος-2021	14	5,384615

Πίνακας 2: Δεδομένα Ηνωμένου Βασιλείου

Ημ. Κρουσμάτων	Κρούσματα	Ημ. Θανάτων	Θάνατοι	Ποσοστό Θνητότητας
8-Ιαν-21	67803	20-Ιαν-21	1823	2,688672
9-Ιαν-21	59717	21-Ιαν-21	1292	2,163538
10-Ιαν-21	54739	22-Ιαν-21	1403	2,563072
11-Ιαν-21	45999	23-Ιαν-21	1350	2,934846
12-Ιαν-21	45364	24-Ιαν-21	611	1,346883
13-Ιαν-21	47350	25-Ιαν-21	593	1,252376
14-Ιαν-21	48502	26-Ιαν-21	1633	3,366871
15-Ιαν-21	55553	27-Ιαν-21	1728	3,110543
16-Ιαν-21	41192	28-Ιαν-21	1241	3,012721
17-Ιαν-21	38455	29-Ιαν-21	1247	3,242751
18-Ιαν-21	37395	30-Ιαν-21	1202	3,214333
19-Ιαν-21	33235	31-Ιαν-21	588	1,769219
20-Ιαν-21	38726	1-Φεβ-21	407	1,050974
21-Ιαν-21	37752	2-Φεβ-21	1451	3,843505
22-Ιαν-21	40111	3-Φεβ-21	1324	3,30084
23-Ιαν-21	33431	4-Φεβ-21	916	2,739972
24-Ιαν-21	29897	5-Φεβ-21	1015	3,394989
25-Ιαν-21	22112	6-Φεβ-21	829	3,749096
26-Ιαν-21	20016	7-Φεβ-21	374	1,868505
27-Ιαν-21	25217	8-Φεβ-21	333	1,320538
28-Ιαν-21	28627	9-Φεβ-21	1054	3,681839
29-Ιαν-21	28910	10-Φεβ-21	1002	3,465929
30-Ιαν-21	23138	11-Φεβ-21	679	2,934567
31-Ιαν-21	20964	12-Φεβ-21	759	3,620492
1-Φεβ-21	18495	13-Φεβ-21	622	3,363071
2-Φεβ-21	16738	14-Φεβ-21	258	1,541403
3-Φεβ-21	19088	15-Φεβ-21	230	1,204946
4-Φεβ-21	20513	16-Φεβ-21	800	3,899966
5-Φεβ-21	19000	17-Φεβ-21	739	3,889474
6-Φεβ-21	18153	18-Φεβ-21	455	2,506473
7-Φεβ-21	15754	19-Φεβ-21	534	3,389615
8-Φεβ-21	14022	20-Φεβ-21	446	3,180716
9-Φεβ-21	12292	21-Φεβ-21	214	1,74097
10-Φεβ-21	12938	22-Φεβ-21	178	1,375792
11-Φεβ-21	13417	23-Φεβ-21	549	4,091824
12-Φεβ-21	15056	24-Φεβ-21	443	2,942349
13-Φεβ-21	13233	25-Φεβ-21	323	2,440868
14-Φεβ-21	10908	26-Φεβ-21	346	3,171984
15-Φεβ-21	9708	27-Φεβ-21	290	2,987227
16-Φεβ-21	10561	28-Φεβ-21	144	1,363507
17-Φεβ-21	12646	1-Μαρ-21	104	0,822394
18-Φεβ-21	11988	2-Μαρ-21	343	2,861195
19-Φεβ-21	11958	3-Μαρ-21	315	2,63422
20-Φεβ-21	10346	4-Μαρ-21	242	2,339068
21-Φεβ-21	9777	5-Μαρ-21	236	2,413828

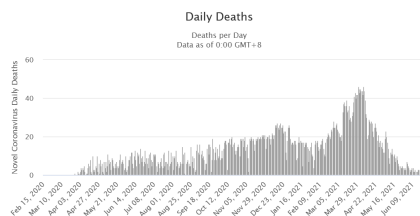
Στο Σχήμα 1 παρατίθενται τα γραφήματα όλων των ημερήσιων κρουσμάτων και θανάτων για τις δύο χώρες, όπως αυτές εμφανίζονται στο <https://www.worldometers.info/coronavirus/>.

Daily New Cases in Moldova



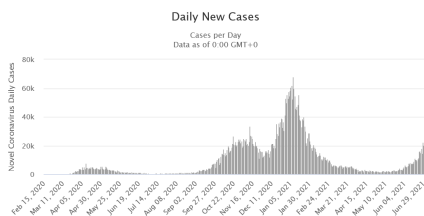
(α') Κρούσματα Μολδαβίας

Daily New Deaths in Moldova



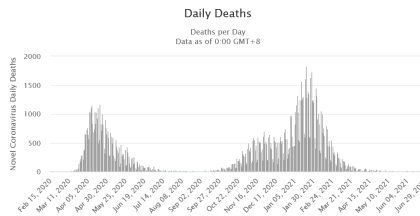
(β') Θάνατοι Μολδαβίας

Daily New Cases in the United Kingdom



(γ') Κρούσματα Ηνωμένου Βασιλείου

Daily New Deaths in the United Kingdom



(δ') Θάνατοι Ηνωμένου Βασιλείου

Σχήμα 1: Ημερήσια Κρούσματα και Ημερήσιοι Θάνατοι στην Μολδαβία και το Ηνωμένο Βασίλειο, όπως αυτά εμφανίζονται στο [worldometers.info](https://www.worldometers.info/coronavirus/)

3 Μελέτη Α

Μελέτη της κατανομής και της Μέσης Τιμής του Ημερήσιου Ποσοστού Θνητότητας στις δύο Χώρες την Περίοδο μετά την Κορύφωση του Κύματος Κορωνοϊού.

3.1 Σχολιασμός της Κατανομής του Ημερήσιου Ποσοστού Θνητότητας

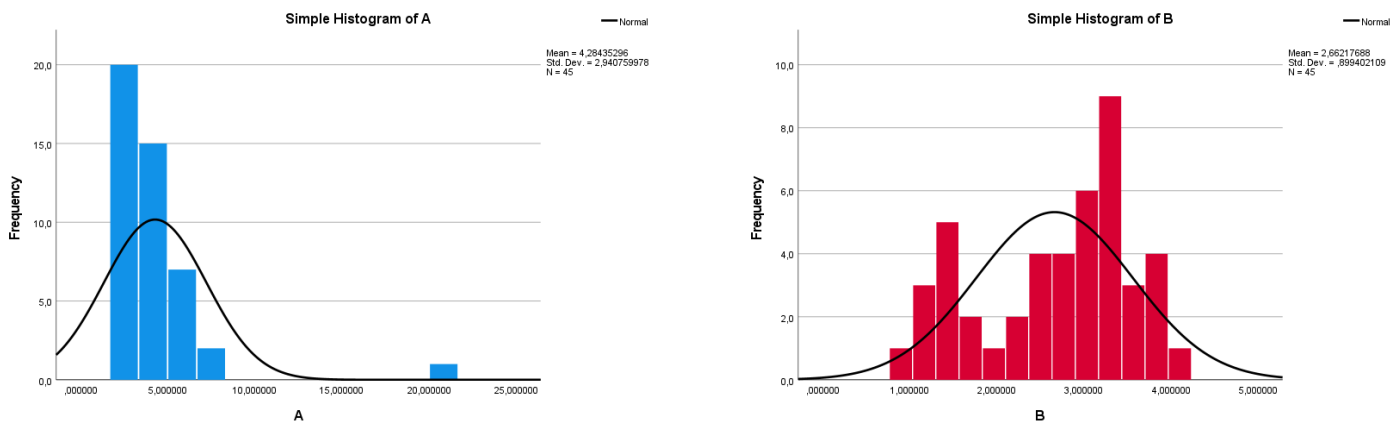
Αρχικά, παραθέτουμε το Σχήμα 2, στο οποίο παρουσιάζονται τα βασικά μέτρα κεντρικής τάσης και μεταβλητότητας για τα ποσοστά θνητότητας στις δύο χώρες.

Statistics			
		A	B
N	Valid	45	45
	Missing	0	0
Mean		4,28435296	2,66217688
Median		3,64741641	2,93456651
Std. Deviation		2,940759978	,899402109
Variance		8,648	,809
Range		18,969448	3,269429
Percentiles	25	2,70969547	1,81886220
	50	3,64741641	2,93456651
	75	4,81997677	3,36497128

Σχήμα 2: Βασικά Μέτρα Κεντρικής Τάσης και Μεταβλητότητας για τα ποσοστά θνητότητας στην Μολδαβία και το Ηνωμένο Βασίλειο

Όπως βλέπουμε, ο μέσος όρος και η διάμεσος της θνητότητας εμφανίζονται μεγαλύτερα για την χώρα της Μολδαβίας, υποδηλώνοντας ότι μεγαλύτερο ποσοστό κρουσμάτων οδηγήθηκαν στον θάνατο, σε σύγκριση με την Μεγάλη Βρετανία. Το ίδιο συμβαίνει με τα μέτρα μεταβλητότητας, κάτι που δείχνει μεγαλύτερη μεταβολή των ποσοστών θνητότητας στην Μολδαβία.

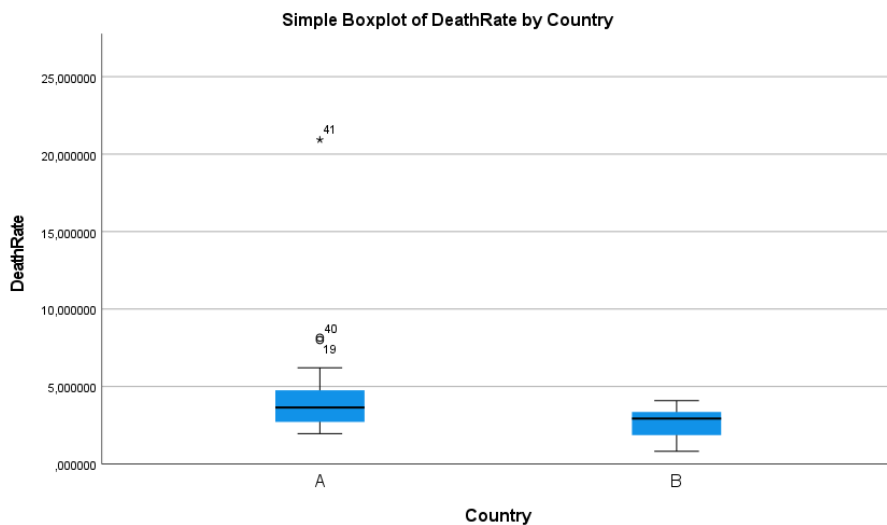
Στο Σχήμα 3 παρατίθενται τα ιστογράμματα θνητότητας των δύο χωρών και σημειώνεται και η μορφή της κανονικής κατανομής



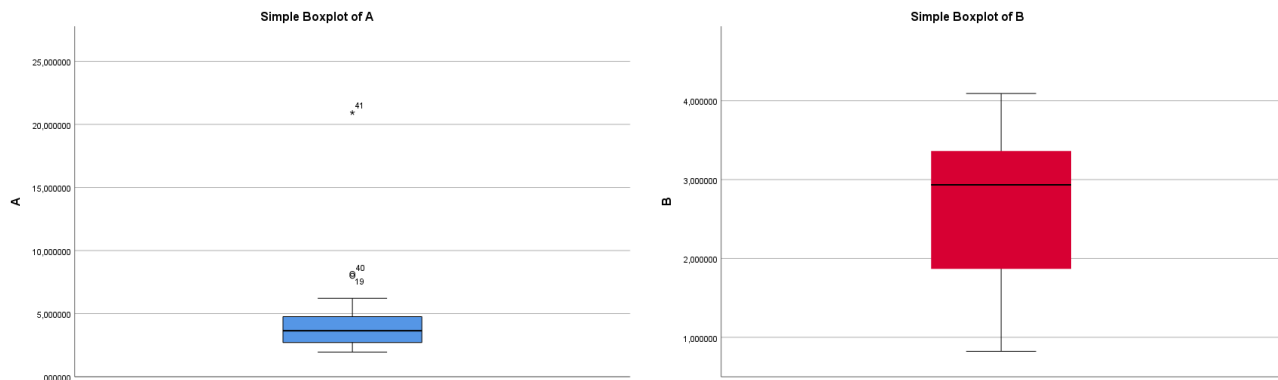
Σχήμα 3: Ιστογράμματα ποσοστού θνητότητας στην Μολδαβία και την Μεγάλη Βρετανία

Στο ιστόγραμμα του δείγματος A φαίνεται να έχουμε συγκέντρωση των ποσοστών θνητότητας σε μικρότερες τιμές. Αντίθετα στο δείγμα B, το ιστόγραμμα δείχνει να πλησιάζει περισσότερο την κανονική κατανομή.

Σχηματίζουμε τα θηκογράμματα για τα δεδομένα για να εξετάσουμε καλύτερα την κατανομή που ακολουθούν τα δείγματα A και B. Στο Σχήμα 4 φαίνονται τα θηκογράμματα και για τα δύο δείγματα. Για λόγους ευκρίνειας, παραθέτουμε και τα θηκογράμματα σε ξεχωριστή κλίμακα στο Σχήμα 5.



Σχήμα 4: Θηκογράμματα Ποσοστών Θνητότητας στην Μολδαβία και την Μεγάλη Βρετανία



Σχήμα 5: Θηκογράμματα Ποσοστών Θνητότητας στην Μολδαβία και την Μεγάλη Βρετανία σε Ξεχωριστή Κλίμακα

Από τα παραπάνω σχήματα, μπορούμε να εξάγουμε τα εξής συμπεράσματα:

1. Για την Μολδαβία (Δείγμα A):

- Η διάμεσος βρίσκεται περίπου στο μέσον των Q_1 και Q_3 , αν και τείνει να πλησιάζει το πρώτο τεταρτομόριο.
- Οι μύστακες φαίνεται να έχουν το ίδιο μήκος.
- Ωστόσο, σημειώνονται τρεις ακραίες τιμές.

Συνεπώς, για το δείγμα A, δεν μπορούμε να αποδεχτούμε ότι ακολουθεί κανονική κατανομή.

2. Για το Ηνωμένο Βασίλειο (Δείγμα B):

- Οι μύστακες έχουν σχεδόν το ίδιο μήκος.
- Η διάμεσος πλησιάζει ελάχιστα περισσότερο στο Q_3 . Ωστόσο, επειδή το δείγμα είναι σχετικά μικρό (45 τιμές) μπορούμε να αποδεχτούμε ότι βρίσκεται στο μέσον του ενδοτεταρτομοριακού εύρους.
- Δεν υπάρχουν ακραίες τιμές στο δείγμα.

Επομένως, μπορούμε να αποδεχτούμε την κανονική κατανομή ως την κατανομή του δείγματος B.

3.2 Μελέτη της Μέσης Τιμής του Ημερήσιου Ποσοστού Θνητότητας

Σε αυτό το εδάφιο θα γίνει μελέτη της μέσης τιμής των δύο δειγμάτων. Παράλληλα θα εξεταστεί για διάστημα εμπιστοσύνης 95% της μέσης τιμής αν η θνητότητα δύναται να φτάσει το 5%.

Στο Σχήμα Σχήμα 6 σημειώνονται τα αποτελέσματα για διάστημα εμπιστοσύνης 95% της μέσης τιμής.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
A	45	4,28435296	2,940759978	,438382614
B	45	2,66217688	,899402109	,134074950

One-Sample Test

Test Value = 0

	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
A	9,773	44	<,001	4,284352962	3,40085086	5,16785507
B	19,856	44	<,001	2,662176880	2,39196657	2,93238719

Σχήμα 6: Αποτελέσματα Ανάλυσης για 95% Διάστημα Εμπιστοσύνης για την Μέση τιμή της Θνητότητας σε Μολδαβία και Μεγάλη Βρετανία

Όπως προκύπτει από τα παραπάνω, το διάστημα εμπιστοσύνης προκύπτει ίσο με [3.40085086, 5.16785507] για το δείγμα A και [2.39196657, 2.93238719] για το δείγμα B. Παρατηρούμε ότι το εύρος του διαστήματος εμπιστοσύνης για την Μεγάλη Βρετανία είναι μικρότερο σε σύγκριση με αυτό της Μολδαβίας. Επίσης, αξίζει να σημειωθεί ότι το διάστημα εμπιστοσύνης της Μολδαβίας εμφανίζει μεγαλύτερες τιμές από αυτό του δείγματος B, επιβεβαιώνοντας τα όσα ειπώθηκαν έως τώρα.

Η μέση τιμή της θνητότητας μπορεί να φτάσει το 5% για την Μολδαβία, όχι όμως και για την Αγγλία. Αυτό προκύπτει εύκολα παρατηρώντας αν ανήκει, ή όχι, το 5 στο διάστημα εμπιστοσύνης κάθε δείγματος. Στο Σχήμα 7 επιβεβαιώνονται τα παραπάνω χρησιμοποιώντας ως Test Value το 5. Όπως βλέπουμε προκύπτουν τα διαστήματα [-1.59914914, 0.16788507] για το δείγμα A (στο οποίο ανήκει το μηδέν) και [-2.60803343, -2.06761281] για το δείγμα B (στο οποίο δεν ανήκει το μηδέν).

One-Sample Test

Test Value = 5

	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
A	-1,632	44	,110	-,715647038	-1,59914914	,16788507
B	-17,437	44	<,001	-2,337823120	-2,60803343	-2,06761281

Σχήμα 7: Διάστημα εμπιστοσύνης 95% για την Μέση τιμή της Θνητότητας σε Μολδαβία και Μεγάλη Βρετανία με Test Value 5

Τα παραπάνω ισχύουν με μία μικρή επιφύλαξη, καθώς το διάστημα εμπιστοσύνης είναι στο 95% (όχι 100%) και το δείγμα σχετικά μικρό, οπότε δεν μπορούμε να είμαστε απόλυτα βέβαιοι για τα αποτελέσματα μας.

3.3 Σύγκριση της Μέσης Τιμής του Ημερήσιου Ποσοστού Θνητότητας

Σε αυτό το εδάφιο θα εξεταστεί αν η μέση τιμή του ποσοστού θνητότητας μπορεί να είναι ίδια για την Μολδαβία και το Ηνωμένο Βασίλειο.

Μελετώντας το 95% διάστημα εμπιστοσύνης για το $\mu_A - \mu_B$ προκύπτουν τα αποτελέσματα που εμφανίζονται στο Σχήμα 8.

Group Statistics										
Country		N	Mean	Std. Deviation	Std. Error Mean					
DeathRate	A	45	4,28435296	2,940759978	,438382614					
	B	45	2,66217688	,899402109	,134074950					

Independent Samples Test										
Levene's Test for Equality of Variances				t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
DeathRate	Equal variances assumed	5,664	,019	3,539	88	<,001	1,622176082	,458427103	,711148546	2,533203618
	Equal variances not assumed			3,539	52,160	<,001	1,622176082	,458427103	,702341945	2,542010219

Σχήμα 8: Αποτελέσματα Ανάλυσης για 95% Διάστημα Εμπιστοσύνης για την Διαφορά των Μέσων Τιμών της Θνητότητας σε Μολδαβία και Μεγάλη Βρετανία

Όπως φαίνεται απο τον πρώτο πίνακα οι τυπικές αποκλίσεις, άρα και οι διακυμάνσεις των δύο δειγμάτων, δεν μπορούν να θεωρηθούν ίσες.

Από τον δεύτερο πίνακα βλέπουμε ότι το 95% διάστημα εμπιστοσύνης για το $\mu_A - \mu_B$ προκύπτει ως:

- [0.711148546, 2.533203618] αν θεωρήσουμε ότι οι διακυμάνσεις των δύο δειγμάτων είναι ίσες.
- [0.702341945, 2.542010219] αν θεωρήσουμε ότι οι διακυμάνσεις των δύο δειγμάτων δεν είναι ίσες, κάτι που ισχύει στην περίπτωση μας.

Σε κάθε μία από τις παραπάνω περιπτώσεις, το διάστημα εμπιστοσύνης δεν περιέχει την μηδενική τιμή, συνεπώς, η μέση τιμή του ημερήσιου ποσοστού θανάτων δεν μπορεί να είναι ίδια για τις δύο χώρες. Επιπρόσθετα, το διάστημα έχει θετικές τιμές, κάτι που υποδηλώνει ότι η μέση τιμή του δείγματος A είναι μεγαλύτερη από την μέση τιμή του δείγματος B. Εδώ πρέπει να σημειωθεί πως τα δείγματα είναι σχετικά μικρά, και το διάστημα εμπιστοσύνης ίσο με 95%, οπότε δεν μπορούμε να είμαστε απολύτως βέβαιοι ότι τα παραπάνω αποτελέσματα είναι απολύτως ακριβή.

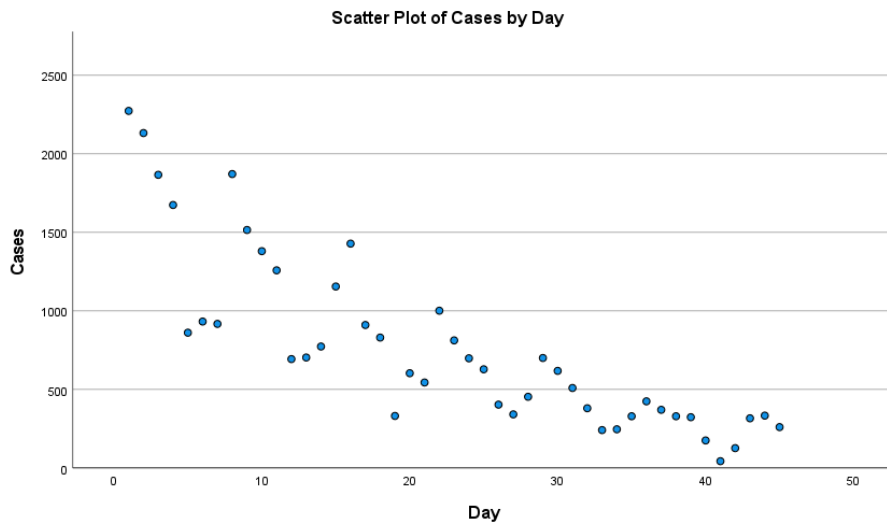
4 Μελέτη B

Σε αυτή την μελέτη εξετάζεται η συσχέτιση των ημερήσιων κρουσμάτων με τον χρόνο μετά την κορύφωση του κύματος

4.1 Συσχέτιση Ημερήσιων Κρουσμάτων και Θανάτων με τον Αύξοντα Αριθμό Ημέρας μετά την Κορύφωση του Κύματος

Για τη μελέτη αυτή θα γίνει χρήση των δεδομένων των ημερήσιων θανάτων και κρουσμάτων για την Μολδαβία (Δεήμα Α), όπως αυτά φαίνονται στον Πίνακα 1.

Σχηματίζουμε το διάγραμμα διασποράς που παρουσιάζεται στο Σχήμα 9.



Σχήμα 9: Διάγραμμα Διασποράς Ημερήσιων Κρουσμάτων στην Μολδαβία ως προς τον Αύξοντα Αριθμό Ημερών

Παρατηρώντας το σχήμα φαίνεται να υπάρχει ισχυρή αρνητική γραμμική συσχέτιση μεταξύ των ημερήσιων θανάτων και του αύξοντα αριθμού ημέρας. Ο πίνακας Σχήμα 10 επιβεβαιώνει τα παραπάνω.

Correlations

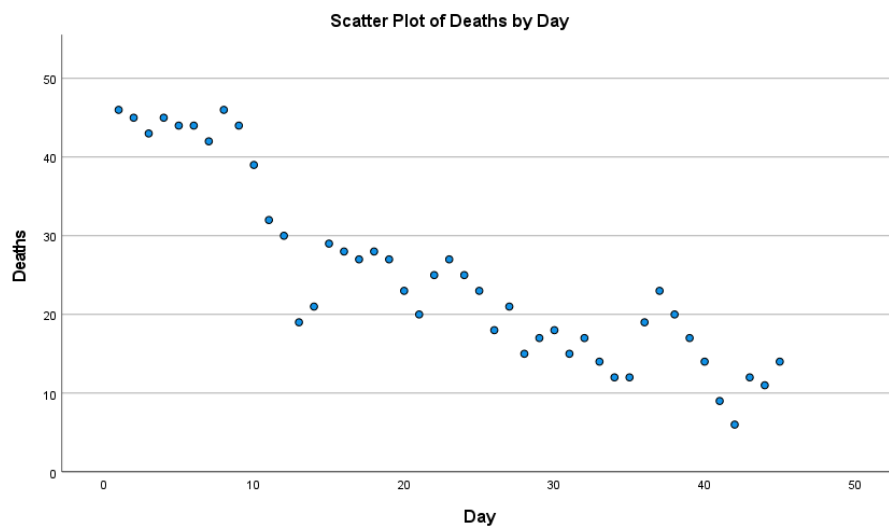
		Day	Cases
Day	Pearson Correlation	1	-,841 ^{**}
	Sig. (2-tailed)		<,001
	N	45	45
Cases	Pearson Correlation	-,841 ^{**}	1
	Sig. (2-tailed)	<,001	
	N	45	45

****.** Correlation is significant at the 0.01 level (2-tailed).

Σχήμα 10: Συσχέτιση ημερήσιων θανάτων και αύξοντα Αριθμού Ημέρας στην Μολδαβία

Ο συντελεστής συσχέτισης *Pearson* που προκύπτει για το δείγμα Α είναι ίσος με -0.841, έχουμε δηλαδή όπως ειπώθηκε, ισχυρή αρνητική συσχέτιση για τα δεδομένα μας.

Ακολουθούμε την ίδια διαδικασία για την εύρεση συσχέτισης, ή μη, μεταξύ των ημερήσιων θανάτων και του αύξοντα αριθμού ημερών. Το διάγραμμα διασποράς φαίνεται στο Σχήμα 11 και ο πίνακας για την εύρεση του συντελεστή συσχέτισης στο Σχήμα 12.



Σχήμα 11: Διάγραμμα Διασποράς Ημερήσιων Θανάτων στην Μολδαβία με τον Αύξοντα Αριθμό Ημέρας

Correlations

		Day	Deaths
Day	Pearson Correlation	1	-,905**
	Sig. (2-tailed)		<,001
	N	45	45
Deaths	Pearson Correlation	-,905**	1
	Sig. (2-tailed)	<,001	
	N	45	45

** . Correlation is significant at the 0.01 level (2-tailed).

Σχήμα 12: Συσχέτιση Ημερήσιων Θανάτων και Αύξοντα Αριθμού Ημέρας στην Μολδαβία

Και εδώ, υπάρχει αρνητική γραμμική συσχέτιση, όπως προκύπτει από τα παραπάνω. Ο συντελεστής *Pearson* είναι ίσος με -0.905, δηλαδή και εδώ, έχουμε ισχυρή αρνητική συσχέτιση.

4.2 Μοντέλο Γραμμικής Παλινδρόμησης Ημερήσιων Κρούσμάτων και Θανάτων με τον Αύξοντα Αριθμό Ημέρας μετά την Κορύφωση του Κύματος

Ημερήσια Κρούσματα

Για την εύρεση Μοντέλου Γραμμικής Παλινδρόμησης για τα ημερήσια κρούσματα χρησιμοποιούμε την μέθοδο ελαχίστων τετραγώνων και προκύπτει το Σχήμα 13 και το Σχήμα 14.

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,841 ^a	,708	,701	303,800

a. Predictors: (Constant), Day

Σχήμα 13: Ισχύς Συσχέτισης των Δεδομένων για τα Ημερήσια Κρούσματα και τον Αύξοντα Αριθμό Ημέρας για την Μολδαβία

Coefficients ^a					
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1	(Constant)	1589,636	92,107	17,259	<,001
	Day	-35,581	3,487	-,841	<,001

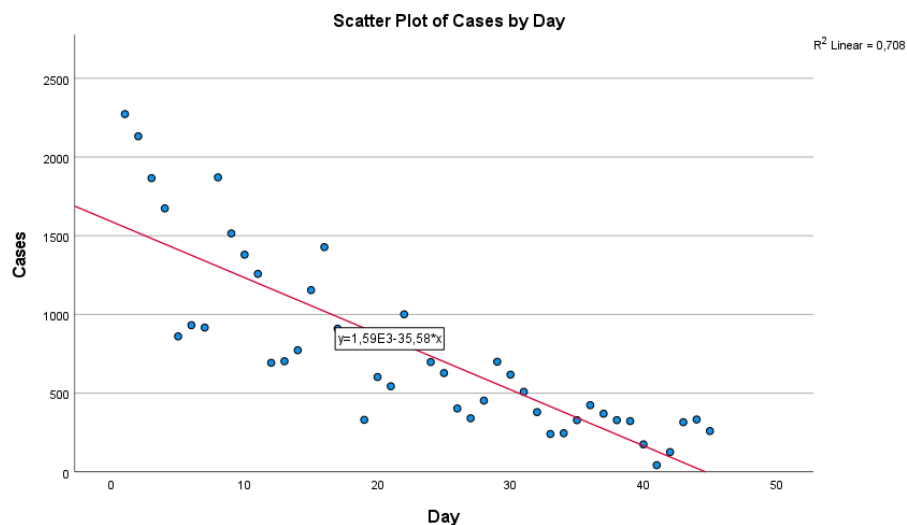
a. Dependent Variable: Cases

Σχήμα 14: Συντελεστές γραμμικής Παλινδρόμησης για τα Ημερήσια Κρούσματα και τον Αύξοντα Αριθμό Ημέρας για την Μολδαβία

Όπως φαίνεται από τα παραπάνω, η εξίσωση γραμμικής παλινδρόμησης ισούται με:

$$y = 1589.636 - 35.581x$$

με τυπική απόκλιση σφαλμάτων παλινδρόμησης ίση με 303.8 Η γραμμική συσχέτιση σημειώνεται και στο Σχήμα 15.



Σχήμα 15: Η Ευθεία Παλινδρόμησης στο Διάγραμμα Διασποράς Κρουσμάτων για την Μολδαβία

Ημερήσιοι Θάνατοι

Όσον αφορά το Μοντέλο Γραμμικής Παλινδρόμησης ημερήσιων θανάτων ακολουθούμε την ίδια μεθοδολογία με τα προηγούμενα και προκύπτει το Σχήμα 16 και το Σχήμα 17.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,905 ^a	,820	,816	5,028

a. Predictors: (Constant), Day

Σχήμα 16: Ισχύς Συσχέτισης των Σεδομένων για τους Ημερήσιους Θανάτους με τον Αύξοντα Αριθμό Ημέρας στην Μολδαβία

Coefficients ^a					
Model		Unstandardized Coefficients		Standardized Coefficients	
		B	Std. Error	Beta	t
1	(Constant)	43,589	1,524		28,595
	Day	-,807	,058	-,905	-13,988

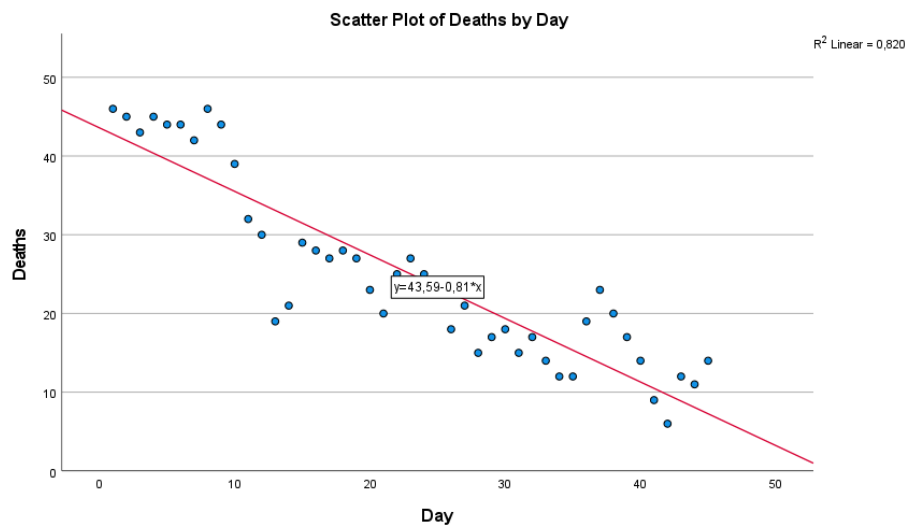
a. Dependent Variable: Deaths

Σχήμα 17: Συντελεστές γραμμικής Παλινδρόμησης για τους ημερήσιους θανάτους με τον αύξοντα αριθμό ημέρας στην Μολδαβία

Από τα παραπάνω, η εξίσωση γραμμικής παλινδρόμησης ισούται με:

$$y = 43.589 - 0.807x$$

με τυπική απόκλιση σφαλμάτων παλινδρόμησης ίση με 5.028 Η γραμμική συσχέτιση σημειώνεται και στο Σχήμα 18.



Σχήμα 18: Η Ευθεία Παλινδρόμησης στο Διάγραμμα Διασποράς Θανάτων για την Μολδαβία

Είναι εμφανές, παρατηρώντας τα γραφήματα, ότι η σχέση μεταξύ κρουσμάτων και θανάτων με τον αύξοντα αριθμό ημέρας προσεγγίζουν την γραμμική συσχέτιση. Τα δύο μοντέλα είναι κατάλληλα για πρόβλεψη, καθώς εμφανίζουν μεγάλο συντελεστή R^2 . Πρέπει, ωστόσο, να σημειωθεί ότι το μοντέλο πρόβλεψης θανάτων είναι καλύτερο από αυτό των κρουσμάτων λόγω της μικρότερης τυπικής απόκλισης σφαλμάτων παλινδρόμησης και του μεγαλύτερου συντελεστή R^2

4.3 Πρόβλεψη Θανάτων κάποια Ημέρα t , γνωρίζοντας τα Ημερήσια Κρούσματα τ Ημέρες Νωρίτερα

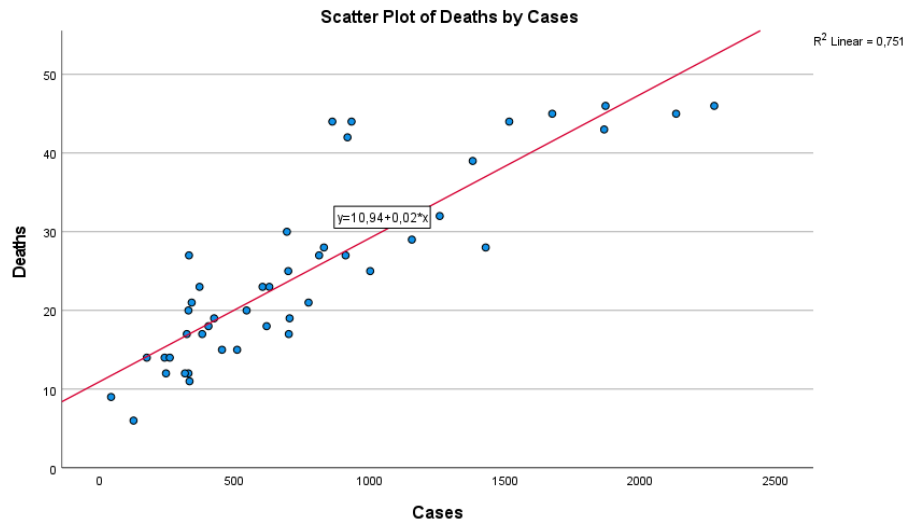
Σε αυτό το εδάφιο θα δημιουργηθεί ένα μοντέλο πρόβλεψης, ώστε να μπορούν να προβλεφθούν οι ημερήσιοι θάνατοι, δεδομένου του αριθμού νέων κρουσμάτων τ ημέρες πριν. Ως τ ορίζουμε την χρονική υστέρηση που μεσολαβεί μεταξύ της κορύφωσης των ημερήσιων θανάτων και της κορύφωσης των ημερήσιων κρουσμάτων στο κύμα που μελετήσαμε στα προηγούμενα εδάφια.

Χρησιμοποιούμε το δείγμα Α. Όπως ορίσαμε στην υποενότητα 2.1 το δείγμα Α, για την Μολδαβία περιλαμβάνει τις χρονικές περιόδους:

- 24/03/2021 με 07/05/2021 για τα ημερήσια κρούσματα.
- 30/03/2021 με 13/05/2021 για τους ημερήσιους θανάτους.

Άρα ως ημέρες υστέρησης ορίζεται: $\tau = 6$ ημέρες.

Στο Σχήμα 19 φαίνεται το διάγραμμα διασποράς μεταξύ των νέων θανάτων και νέων κρουσμάτων. Σημειώνεται και η ευθεία γραμμικής συσχέτισης που προκύπτει από την μέθοδο των ελαχίστων τετραγώνων. Η ανάλυση για την εύρεση της ευθείας γίνεται στην υποενότητα 4.4. Στο Σχήμα 20 φαίνεται και ο υπολογισμός του συντελεστή *Pearson* ο οποίος προκύπτει ίσος με 0.866.



Σχήμα 19: Διάγραμμα Διασποράς Ημερήσιων Θανάτων και Ημερήσιων Κρουσμάτων για την Μολδαβία

Correlations ($\tau = 6$)

		Deaths	Cases_t6
Deaths	Pearson Correlation	1	,866**
	Sig. (2-tailed)		<,001
	N	45	45
Cases_t6	Pearson Correlation	,866**	1
	Sig. (2-tailed)	<,001	
	N	45	45

** . Correlation is significant at the 0.01 level (2-tailed).

Σχήμα 20: Συντελεστής Pearson για $\tau = 6$

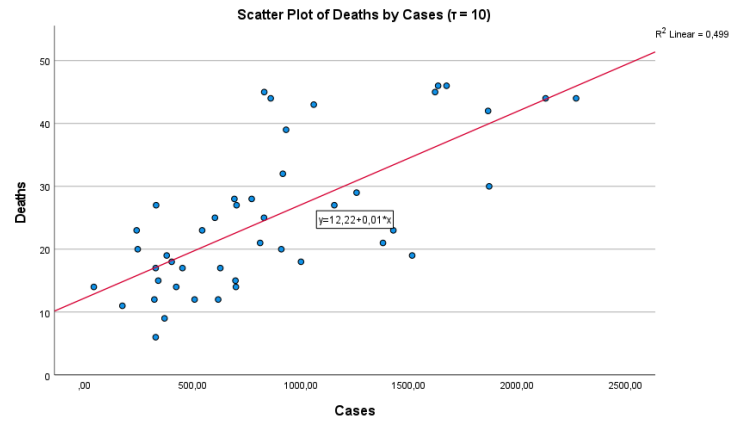
Ακολουθούμε την ίδια διαδικασία επιλέγοντας αυτή την φορά διαφορετικές τιμές για την υστέρηση τ , διατηρώντας σταθερά τα δεδομένα για τους θανάτους. Συγκεκριμένα θα επιλέξουμε $\tau = 10$ και $\tau = 20$. Με αυτά τα δεδομένα έχουμε:

- για $\tau = 10$, προκύπτουν τα δεδομένα για τις ημέρες των κρουσμάτων κρουσμάτων: 20/03/2021 έως 03/05/2021.
- για $\tau = 20$, προκύπτουν τα δεδομένα για τις ημέρες των κρουσμάτων κρουσμάτων: 10/03/2021 έως 23/04/2021.

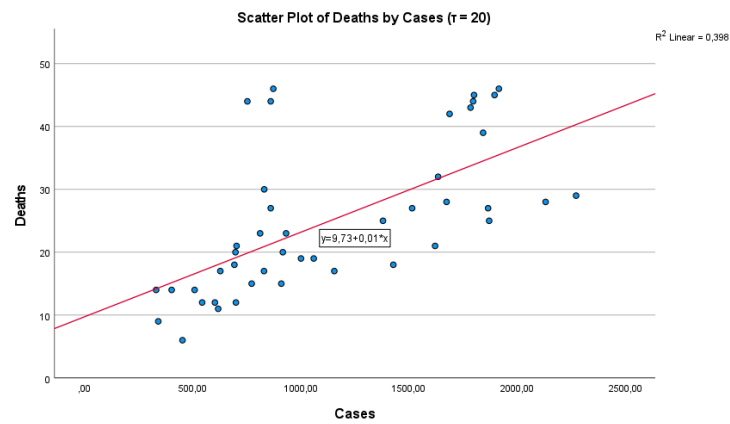
Στον Πίνακα 3 παρατίθενται τα δεδομένα για τα ημερήσια κρούσματα για τις ημερομηνίες 10/03/2021 έως 23/03/2021, τα οποία δεν συμπεριλήφθηκαν στον Πίνακα 1

Από την ανάλυση των δεδομένων για $\tau = 10$ και $\tau = 20$ προκύπτουν τα γραφήματα διασποράς στο Σχήμα 21. Σημειώνεται και η ευθεία παλινδρόμησης που προκύπτει με την μέθοδο ελαχίστων τετραγώνων. Η ανάλυση για την εύρεση των ευθειών γίνεται και αυτή στην υποενότητα 4.4. Ο συντελεστής Pearson εμφανίζεται στο Σχήμα 22 και προκύπτει ίσος με 0.707 και 0.631 για $\tau = 10$ και $\tau = 20$ αντίστοιχα.

Σε όλα τα γραφήματα φαίνεται καθαρά η θετική γραμμική σχέση που συνδέει τον αριθμό κρουσμάτων σε σχέση με τον μελλοντικό αριθμό θανάτων. Ωστόσο, θα πρέπει να σημειωθεί ότι η σχέση είναι ισχυρότερη όταν ως υστέρηση επιλέγεται η διαφορά της κορύφωσης ημερήσιων κρουσμάτων και ημερήσιων θανάτων ($\tau = 6$).



(α') Υστέρηση $\tau = 10$



(β') Υστέρηση $\tau = 20$

Σχήμα 21: Γραφήματα Διασποράς για Διαφορετικές Τιμές Υστέρησης

Correlations ($\tau = 10$)

		Deaths	Cases
Deaths	Pearson Correlation	1	,707**
	Sig. (2-tailed)		<,001
	N	45	45
Cases	Pearson Correlation	,707**	1
	Sig. (2-tailed)	<,001	
	N	45	45

** . Correlation is significant at the 0.01 level (2-tailed).

(α') $\tau = 10$

Correlations ($\tau = 20$)

		Deaths	Cases
Deaths	Pearson Correlation	1	,631**
	Sig. (2-tailed)		<,001
	N	45	45
Cases	Pearson Correlation	,631**	1
	Sig. (2-tailed)	<,001	
	N	45	45

** . Correlation is significant at the 0.01 level (2-tailed).

(β') $\tau=20$

Σχήμα 22: Συντελεστής Pearson για Διαφορετικές Τιμές του τ

Πίνακας 3: Ημερήσια Κρούσματα Μολδαβίας για την Περίοδο 10/03/21 έως 23/03/21

Ημερομηνία	Ημ. Κρούσματα
10-Μαρ-2021	873,00
11-Μαρ-2021	1896,00
12-Μαρ-2021	1785,00
13-Μαρ-2021	1801,00
14-Μαρ-2021	753,00
15-Μαρ-2021	861,00
16-Μαρ-2021	1688,00
17-Μαρ-2021	1916,00
18-Μαρ-2021	1797,00
19-Μαρ-2021	1843,00
20-Μαρ-2021	1635,00
21-Μαρ-2021	831,00
22-Μαρ-2021	1060,00
23-Μαρ-2021	1621,00

Επίσης, πρέπει να σημειωθεί ότι η συσχέτιση για $\tau = 10$ είναι ισχυρότερη από την συσχέτιση για $\tau = 20$. Αυτά είναι προφανή από το σκόρπισμα που εμφανίζουν οι τιμές γύρω από την ευθεία παλινδρόμησης, και αποδεικνύονται με τον συντελεστή *Pearson*.

4.4 Εκτίμηση Μοντέλου Παλινδρόμησης για την εξάρτηση του αριθμού ημερήσιων νέων θανάτων από τον αριθμό των ημερήσιων νέων κρουσμάτων τ η-μέρες

Σε αυτό το εδάφιο θα εκτιμηθούν οι ευθείες που εμφανίζονται στα προηγούμενα διαγράμματα διασποράς και οι οποίες αποτελούν ένα μοντέλο πρόβλεψης των ημερήσιων θανάτων, βάσει των ημερήσιων κρουσμάτων που είχαμε τ ημέρες νωρίτερα. Το τ , όπως και στην υποενότητα 4.3 ορίζεται ως 6 (Διαφορά κορύφωσης κύματος για κρούσματα και θανάτους), 10 και 20.

Με την ανάλυση των δεδομένων για τις διάφορες τιμές του τ προκύπτουν οι πίνακες που εμφανίζονται στο Σχήμα 23

Προκύπτουν δηλαδή ότι προκύπτουν οι ευθείες:

- Για $\tau = 6$: $y = 10.938 + 0.018x$ με τυπική απόκλιση σφάλματος 5.916 και $R^2 = 0.745$
- Για $\tau = 10$: $y = 12.221 + 0.015x$ με τυπική απόκλιση σφάλματος 8.381 και $R^2 = 0.488$
- Για $\tau = 20$: $y = 9.727 + 0.13x$ με τυπική απόκλιση σφάλματος 9.188 και $R^2 = 0.384$

Model Summary ($\tau = 6$)				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,866 ^a	,751	,745	5,916

a. Predictors: (Constant), Cases

Coefficients (r = 6) ^a						
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	
	B	Std. Error	Beta			
1	(Constant)	10,938	1,520	7,195	<,001	
	Cases	,018	,002	,866	11,375	<,001

a. Dependent Variable: Deaths

Model Summary ($\tau = 10$)				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,707 ^a	,499	,488	8,381

a. Predictors: (Constant), Cases

Coefficients (r = 10) ^a						
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	
	B	Std. Error	Beta			
1	(Constant)	12,221	2,320	5,268	<.,001	
	Cases	,015	,002	,707	6,549	<.,001

a. Dependent Variable: Deaths

Model Summary ($\tau = 20$)				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,631 ^a	,398	,384	9,188

a. Predictors: (Constant), Cases

Coefficients (r = 20) ^a						
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	
	B	Std. Error	Beta			
1	(Constant)	9,727	3,177	3,062	,004	
	Cases	,013	,003	,631	5,336	<,001

a. Dependent Variable: Deaths

Σχήμα 23: Ισχύς Παλινδρόμησης και Συντελεστές Παλινδρόμησης για τις διάφορες Τιμές του τ .

Είναι εμφανές ότι καταλληλότερο μοντέλο πρόβλεψης είναι για χρόνο υστέρησης $\tau = 6$, το οποίο έχει χαμηλότερη τυπική απόκλιση σφάλματος και υψηλότερο δείκτη R^2 . Ακολουθεί το μοντέλο για $\tau = 10$ και τέλος για $\tau = 20$.

5 Παράρτημα

5.1 Κατάλογοι Σχημάτων και Πινάκων

Κατάλογος Σχημάτων

1	Ημερήσια Κρούσματα και Ημερήσιοι Θάνατοι στην Μολδαβία και το Ηνωμένο Βασίλειο, όπως αυτά εμφανίζονται στο worldometers.info	5
2	Βασικά Μέτρα Κεντρικής Τάσης και Μεταβλητότητας για τα ποσοστά θνητότητας στην Μολδαβία και το Ηνωμένο Βασίλειο	6
3	Ιστογράμματα ποσοστού θνητότητας στην Μολδαβία και την Μεγάλη Βρετανία	7
4	Θηκογράμματα Ποσοστών Θνητότητας στην Μολδαβία και την Μεγάλη Βρετανία	7
5	Θηκογράμματα Ποσοστών Θνητότητας στην Μολδαβία και την Μεγάλη Βρετανία σε Ξεχωριστή Κλίμακα	8
6	Αποτελέσματα Ανάλυσης για 95% Διάστημα Εμπιστοσύνης για την Μέση τιμή της Θνητότητας σε Μολδαβία και Μεγάλη Βρετανία . .	9
7	Διάστημα εμπιστοσύνης 95% για την Μέση τιμή της Θνητότητας σε Μολδαβία και Μεγάλη Βρετανία με Test Value 5	9
8	Αποτελέσματα Ανάλυσης για 95% Διάστημα Εμπιστοσύνης για την Διαφορά των Μέσων Τιμών της Θνητότητας σε Μολδαβία και Μεγάλη Βρετανία	10
9	Διάγραμμα Διασποράς Ημερήσιων Κρουσμάτων στην Μολδαβία ως προς τον Αύξοντα Αριθμό Ημερών	11
10	Συσχέτιση ημερήσιων θανάτων και αύξοντα Αριθμού Ημέρας στην Μολδαβία	12
11	Διάγραμμα Διασποράς Ημερήσιων Θανάτων στην Μολδαβία με τον Αύξοντα Αριθμό Ημέρας	13
12	Συσχέτιση Ημερήσιων Θανάτων και Αύξοντα Αριθμού Ημέρας στην Μολδαβία	13
13	Ισχύς Συσχέτισης των Δεδομένων για τα Ημερήσια Κρούσματα και τον Αύξοντα Αριθμό Ημέρας για την Μολδαβία	14
14	Συντελεστές γραμμικής Παλινδρόμησης για τα Ημερήσια Κρούσματα και τον Αύξοντα Αριθμό Ημέρας για την Μολδαβία	14
15	Η Ευθεία Παλινδρόμησης στο Διάγραμμα Διασποράς Κρουσμάτων για την Μολδαβία	15
16	Ισχύς Συσχέτισης των Σεδομένων για τους Ημερήσιους Θανάτους με τον Αύξοντα Αριθμό Ημέρας στην Μολδαβία	15
17	Συντελεστές γραμμικής Παλινδρόμησης για τους ημερήσιους θανάτους με τον αύξοντα αριθμό ημέρας στην Μολδαβία	16
18	Η Ευθεία Παλινδρόμησης στο Διάγραμμα Διασποράς Θανάτων για την Μολδαβία	16
19	Διάγραμμα Διασποράς Ημερήσιων Θανάτων και Ημερήσιων Κρουσμάτων για την Μολδαβία	17
20	Συντελεστής Pearson για $\tau = 6$	18

21	Γραφήματα Διασποράς για Διαφορετικές Τιμές Υστέρησης	19
22	Συντελεστής Pearson για Διαφορετικές Τιμές του τ	20
23	Ισχύς Παλινδρόμησης και Συντελεστές Παλινδρόμησης για τις δι- άφορες Τιμές του τ	22

Κατάλογος Πινάκων

1	Δεδομένα Μολδαβίας	3
2	Δεδομένα Ηνωμένου Βασιλείου	4
3	Ημερήσια Κρούσματα Μολδαβίας για την Περίοδο 10/03/21 έως 23/03/21	21

6 Πηγές

- Όλα τα δεδομένα προέρχονται από το:
<https://www.worldometers.info/coronavirus/>.
- Τα γραφήματα και οι πίνακες που παρατίθενται στο έγγραφο έγιναν με χρήση του λογισμικού *IBM SPSS*.
- Για την συγγραφή χρησιμοποιήθηκε η γλώσσα δημιουργίας εγγράφων *L^AT_EX*.