Εργασία SPSS

Τσαρναδέλης Αθανάσιος Γρηγόριος

10388

ΑΕΜ 10388 άρα επιλέγω την 13η χώρα, την Ισλανδία, και την Ουγγαρία σαν αλφαβητικά γειτονική.

Κάνω στο Excel την γραφική παράσταση του δείκτη θετικότητας για την Ισλανδία, την Ουγγαρία και την Ελλάδα .

Παρατηρώ ότι το μέγιστο της Ισλανδίας πολύ κοντά στην αρχή λήψης των δεδομένων, άρα δεν έχω δεδομένα για τις 12 εβδομάδες πριν το μέγιστο. Επομένως αλλάζω και επιλέγω την Ιρλανδία, αλφαβητικά επόμενη την Ισλανδίας, καθώς έχω ήδη επιλέξει την προηγούμενη, την Ουγγαρία. Το διάγραμμα του δείκτη θετικότητας της Ιρλανδίας φαίνεται παρακάτω.

Επιλέγω την εβδομάδα που γίνεται ο δείκτης μέγιστος. Για την Ιρλανδία είναι 2021-W01, άρα επιλέγω 2020-W42 έως 2020-W53. Αντίστοιχα για την Ουγγαρία είναι 2021-W11, άρα επιλέγω 2020-W52 έως 2021-W10. Για την Ελλάδα είναι 2020-W46 άρα επιλέγω 2020-W34 έως 2020-W45.

# Μελέτη Α

1. Κάνω τις αναλύσεις και παραθέτω παρακάτω τα αποτελέσματα.

* Μέση τιμή και διάμεσος, διασπορά, τυπική απόκλιση, εύρος δεδομένων, πρώτο και τρίτο τεταρτομόριο, σε ένα πίνακα και για τα τρία δείγματα. Table

  Description automatically generated
* Ιστογράμματα

Chart, bar chart, histogram

Description automatically generatedChart, bar chart

Description automatically generated

Chart, bar chart

Description automatically generated

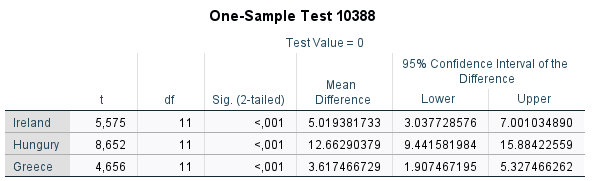
* Θηκόγραμμα

Chart, box and whisker chart

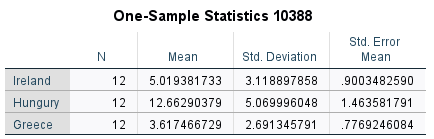
Description automatically generated

Με βάση τα ιστογράμματα, με οπτικό κριτήριο βλέπω ότι τα δείγματα δεν μοιάζει να ακολουθούν κανονική κατανομή. Τα θηκογράμματα της Ιρλανδίας και της Ελλάδας έχουν ακραίες τιμές, και στο θηκόγραμμα της Ουγγαρίας ο πάνω μύστακας είναι πολύ μεγαλύτερος από τον κάτω. Οι παραπάνω παρατηρήσεις με οδηγούν στο συμπέρασμα ότι παρότι έχω μικρό δείγμα(μόλις 12 παρατηρήσεις) και θα μπορούσα να είμαι ελαστικός, καμιά από τις τρεις περιπτώσεις δεν ακολουθούν κανονική κατανομή. Τέλος, βλέπω ότι ο δείκτης θετικότητας της Ελλάδας είναι συνολικά ο μικρότερος, έπεται η Ιρλανδία και μετά η Ουγγαρία, με τον μεγαλύτερο δείκτη θετικότητας από τις τρεις χώρες.

2.Βρίσκω τα διαστήματα εμπιστοσύνης.

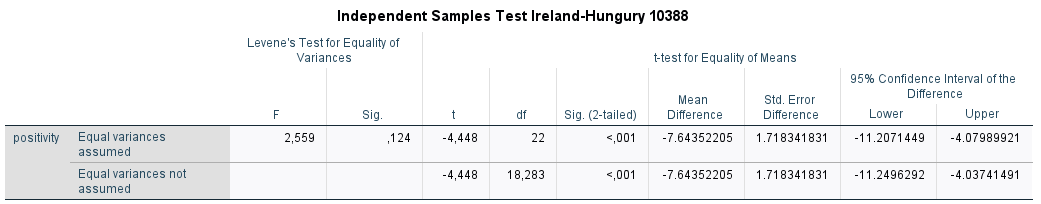


Με 95% διάστημα εμπιστοσύνης παρατηρώ ότι το 5 είναι μέσα στα διαστήματα εμπιστοσύνης της Ιρλανδίας, καθώς και της Ελλάδας. Το 5 δεν υπάρχει μέσα στο διάστημα εμπιστοσύνης της Ουγγαρίας, καθώς το τελευταίο είναι πολύ μεγαλύτερο.



Στον παραπάνω πίνακα φαίνεται το σφάλμα εκτίμησης της μέσης τιμής. Το μικρότερο σφάλμα το έχει η Ελλάδα, μετά η Ιρλανδία, και μετά η Ουγγαρία.

Συγκρίνω τα αποτελέσματα της Ιρλανδίας-Ουγγαρίας:



Ο πίνακας που με ενδιαφέρει είναι ο πάνω, καθώς υποθέτω ίσες διασπορές(Equal Variances Assumed). Παρατηρώ ότι το διάστημα εμπιστοσύνης είναι αυστηρά αρνητικό, που σημαίνει ότι η διαφορά δεν θα είναι 0 με εμπιστοσύνη 95%, δηλαδή με πιθανότητα 95% ο εβδομαδιαίος δείκτης θετικότητας δεν μπορεί να είναι ίδιος στις δύο χώρες.

Έπειτα συγκρίνω τα αποτελέσματα Ιρλανδίας-Ελλάδας:

Table

Description automatically generated with low confidence

Παρατηρώ ότι το 0 είναι μέσα στο διάστημα εμπιστοσύνης άρα η διαφορά μπορεί να είναι 0 με 95% εμπιστοσύνη.

Τελικά, με 95% πιθανότητα ο δείκτης θετικότητας δεν μπορεί να είναι ίδιος στη Ιρλανδία και την Ουγγαρία, ενώ με 95% πιθανότητα ο δείκτης θετικότητας μπορεί να είναι ίδιος στην Ιρλανδία και στην Ελλάδα.

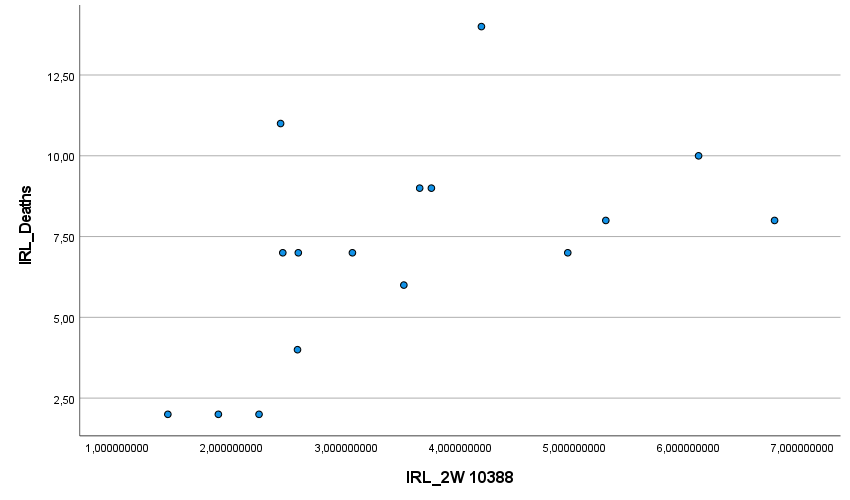
# Μελέτη Β

3) Επιλέγω για την Ιρλανδία το διάστημα 2020-W38 έως 2020-W53 (16 παρατηρήσεις) και αντίστοιχα τον εβδομαδιαίο αριθμό θανάτων από 1 έως 4 εβδομάδες πριν(16 παρατηρήσεις σε κάθε στήλη). Αντίστοιχα για την Ουγγαρία επιλέγω το διάστημα 2020-W46 έως 2021-W10 και τον αντίστοιχο εβδομαδιαίο αριθμό θανάτων από 1 έως 4 εβδομάδες πριν.

Παρακάτω φαίνονται τα διαγράμματα διασποράς για την χώρα Α, την Ιρλανδία, με 1,2,3,4 εβδομάδες υστέρηση αντίστοιχα.

Chart, scatter chart

Description automatically generated



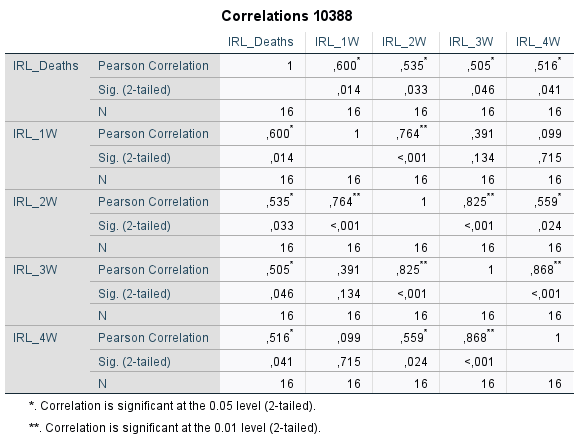
Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generated

Ο πίνακας συσχέτισης και για τις 4 περιπτώσεις φαίνεται παρακάτω:



Από τα διαγράμματα, καθώς και τον παραπάνω πίνακα, οδηγούμαι στο συμπέρασμα ότι υπάρχει ασθενής θετική εξάρτηση(r<0.9) των θανάτων από τα εβδομαδιαία κρούσματα. Επίσης, όσο μεγαλύτερη υστέρηση τόσο μικρότερη εξάρτηση, καθώς με υστέρηση μια εβδομάδας έχω μεγαλύτερο συντελεστή συσχέτισης(0.6) από ότι στις 3 άλλες περιπτώσεις(0.535, 0.505, 0.516).Το παραπάνω συμπέρασμα είναι λογικό, καθώς όσο μεγαλύτερη υστέρηση, τόσο λιγότερο εξαρτώνται οι θάνατοι από τον δείκτη θετικότητας, καθώς η χρονική απόσταση τους αυξάνεται.

Κάνω την ίδια μελέτη για την χώρα Β:

Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generated

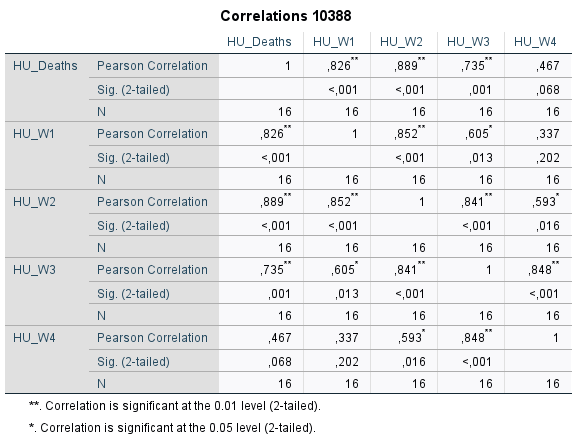
Chart, scatter chart

Description automatically generated

Chart, scatter chart

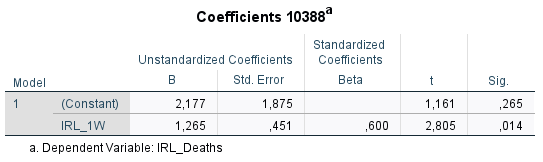
Description automatically generated

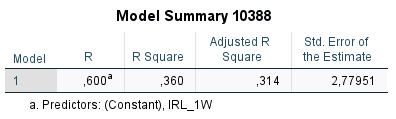
Παρακάτω φαίνεται ο πίνακας συσχέτισης για όλες τις περιπτώσεις:

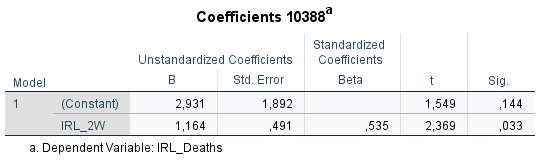


Αντίστοιχα συμπεράσματα προκύπτουν για την Ουγγαρία, καθώς όσο μεγαλύτερη η υστέρηση, τόσο μικρότερη η εξάρτηση. Σε αντίθεση με την Ιρλανδία, στην Ουγγαρία στις 2 πρώτες βδομάδες παρατηρώ σχεδόν ισχυρή εξάρτηση, καθώς r μικρότερο του 0.9 αλλά πολύ κοντά σε αυτό(0.826, 0.889). Επίσης, στις δύο πρώτες περιπτώσεις υστέρησης έχω μεγαλύτερη εξάρτηση από τις άλλες δύο περιπτώσεις. Τέλος, η μεγαλύτερη εξάρτηση φαίνεται αυτή με υστέρηση 2 εβδομάδων. Τα παραπάνω αποτελέσματα είναι λογικά, για τους ίδιους λόγους με τα προηγούμενα.

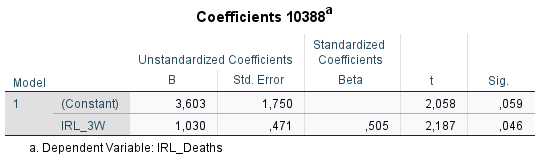
4) Παρακάτω φαίνονται οι πίνακες απλής γραμμικής παλινδρόμισης για 1,2,3,4 εβδομάδες υστέρηση αντίστοιχα:

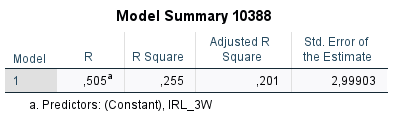


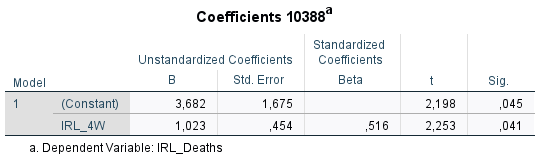


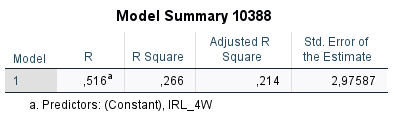
Table

Description automatically generated









Άρα τα 4 μοντέλα τα οποία ψάχνω είναι:

1)y=1.265x+2.177

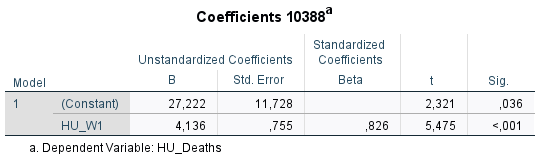
2) y=2.931x+1.164

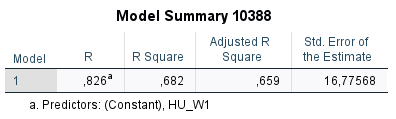
3) y=3.603x+1.030

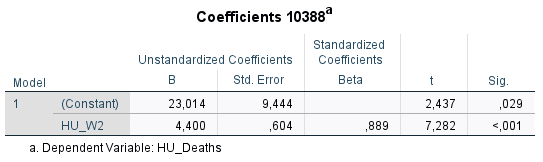
4) y=3.682x+1.023

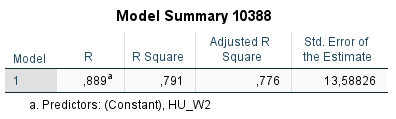
Το καταλληλότερο μοντέλο από τα 4 για προβλέψεις είναι το 1ο, καθώς έχω τον μεγαλύτερο συντελεστή προσδιορισμού r^2, καθώς και το μικρότερη τυπική απόκλιση σφάλματος παλινδρόμησης.

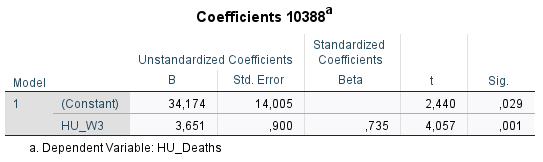
Κάνω την ίδια μελέτη για την Ουγγαρία:

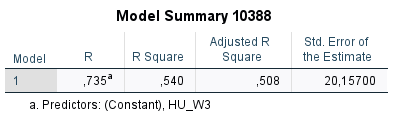


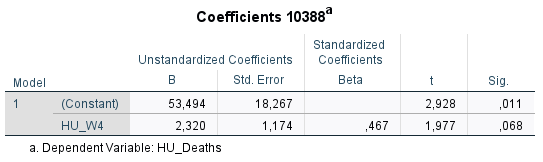


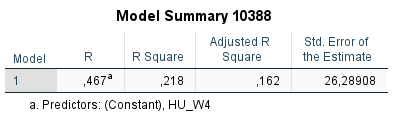












Άρα τα 4 μοντέλα τα οποία ψάχνω είναι:

1)y=4.136x+27.222

2) y=4.4x+23.014

3) y=3.651x+34.174

4) y=2.32x+53.494

Το καταλληλότερο μοντέλο από τα 4 για προβλέψεις είναι το 2ο, καθώς έχω τον μεγαλύτερο συντελεστή προσδιορισμού r^2, καθώς και το μικρότερη τυπική απόκλιση σφάλματος παλινδρόμησης.