

Θεωρία Πιθανοτήτων και Στατιστική

Εργασία Στατιστικής 2020

Ρόντος Έκτορας – Θωμάς

AEM:9477

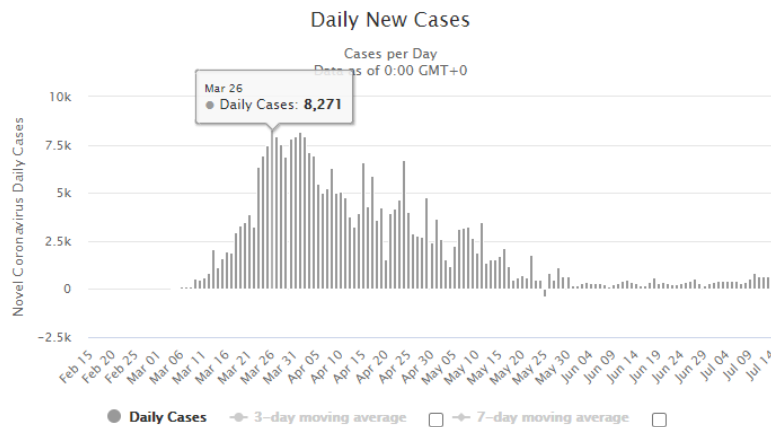
$77 \bmod 37 = 3$: Ισπανία

$94 \bmod 37 = 20$: Τουρκία

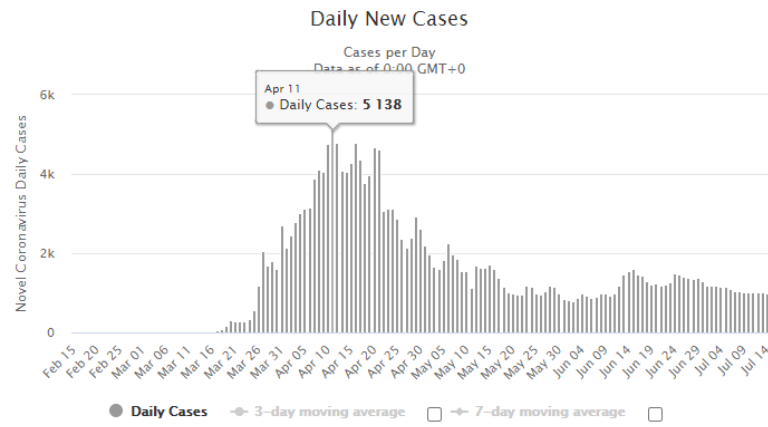
Γενικά

Θα πραγματοποιήσουμε στατιστική ανάλυση στα καταγεγραμμένα στοιχεία που αφορούν τα κρούσματα και τους θανάτους του κορονοϊού στην Ισπανία και στην Τουρκία.

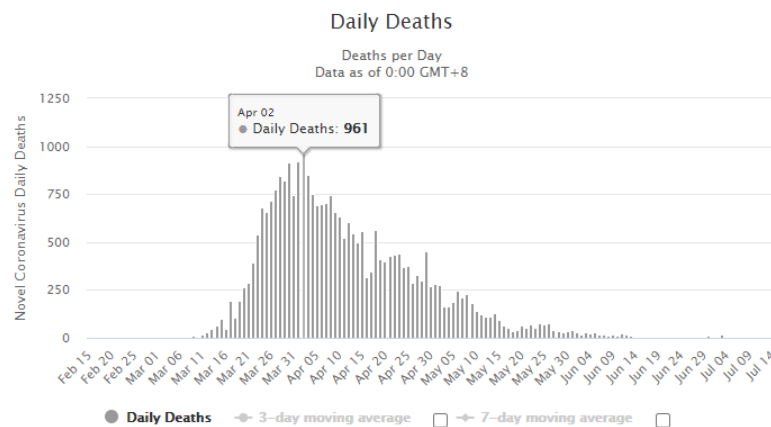
Daily New Cases in Spain



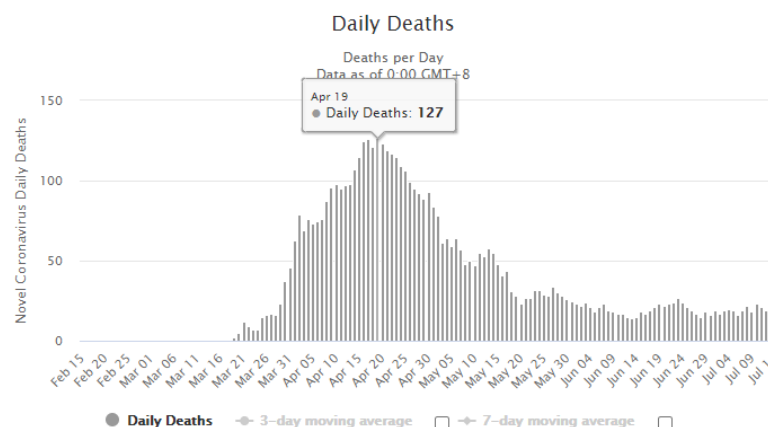
Daily New Cases in Turkey



Daily New Deaths in Spain



Daily New Deaths in Turkey



Ορίζουμε ως κορύφωση των νέων κρουσμάτων τα 8.271 νέα κρούσματα στις 26/3 στην Ισπανία και τα 5.138 νέα κρούσματα στις 11/4 στην Τουρκία. Ομοίως, έχουμε 961 στις 2/4 και 127 στις 19/4 για τους θανάτους στην Ισπανία και στην Τουρκία αντίστοιχα. Η επιλογή αυτών των στοιχείων ήταν σχετικά εύκολη καθώς παρατηρείται μείωση πριν και μετά από αυτές τις ημέρες και δεν υπήρχε αλλού κάποια άλλη ακραία τιμή.

Από εδώ και πέρα, θα αναφερόμαστε στα δείγματα για την Ισπανία και την Τουρκία με τα γράμματα Α και Β αντίστοιχα.

Μελέτη Α

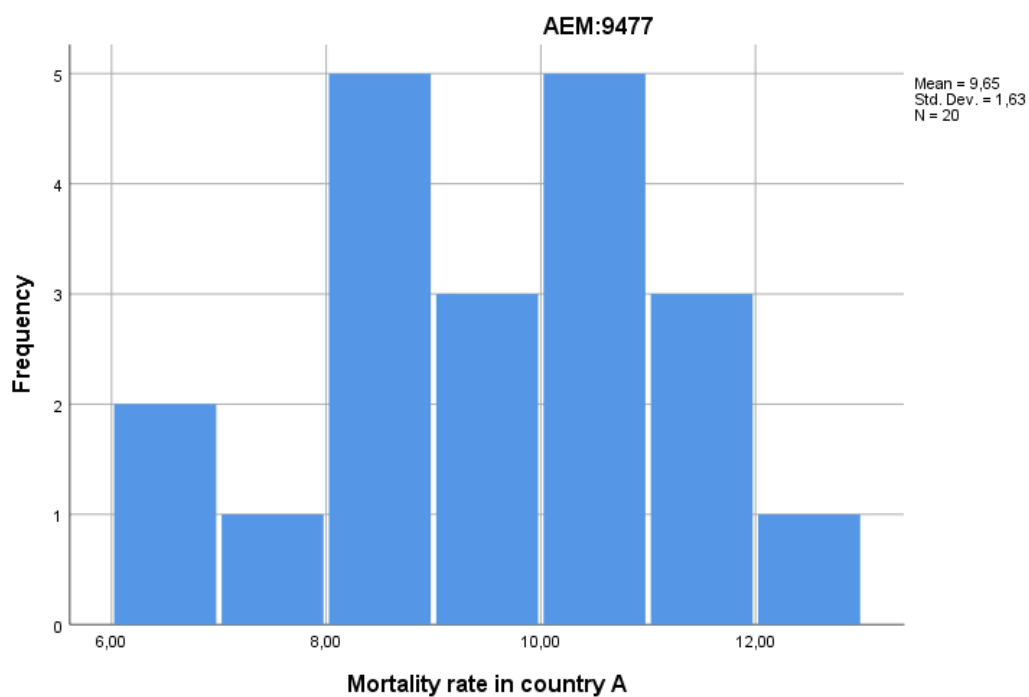
Θα εξετάσουμε το ημερήσιο ποσοστό θνητότητας στις δύο χώρες με βάση τα δείγματα που πήραμε έως και 20 μέρες μετά την κορύφωση.

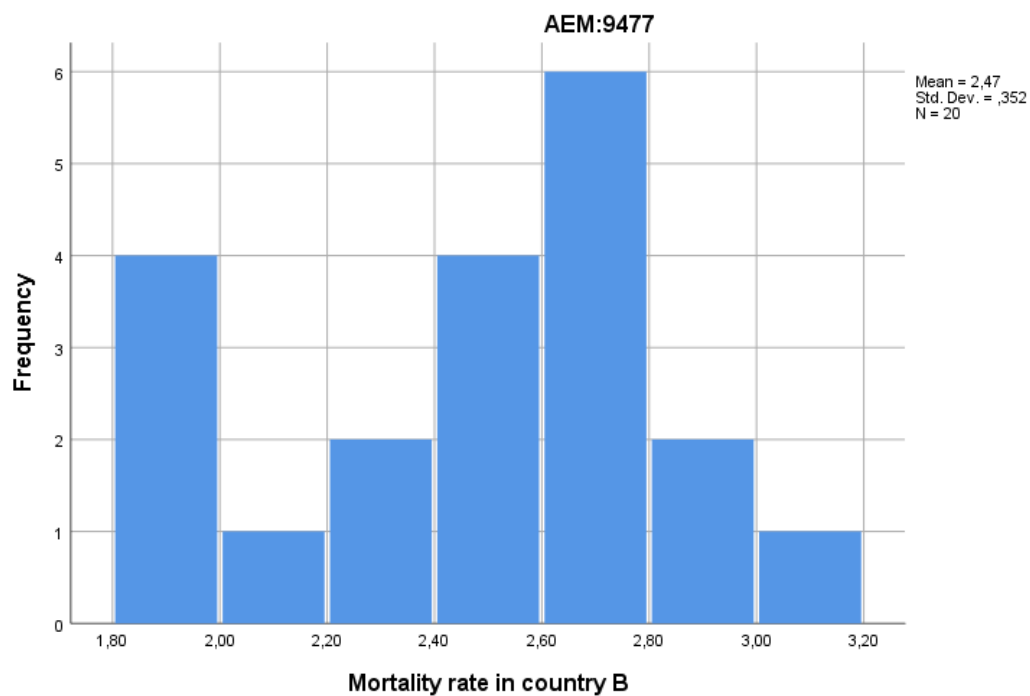
1.

Statistics AEM:9477

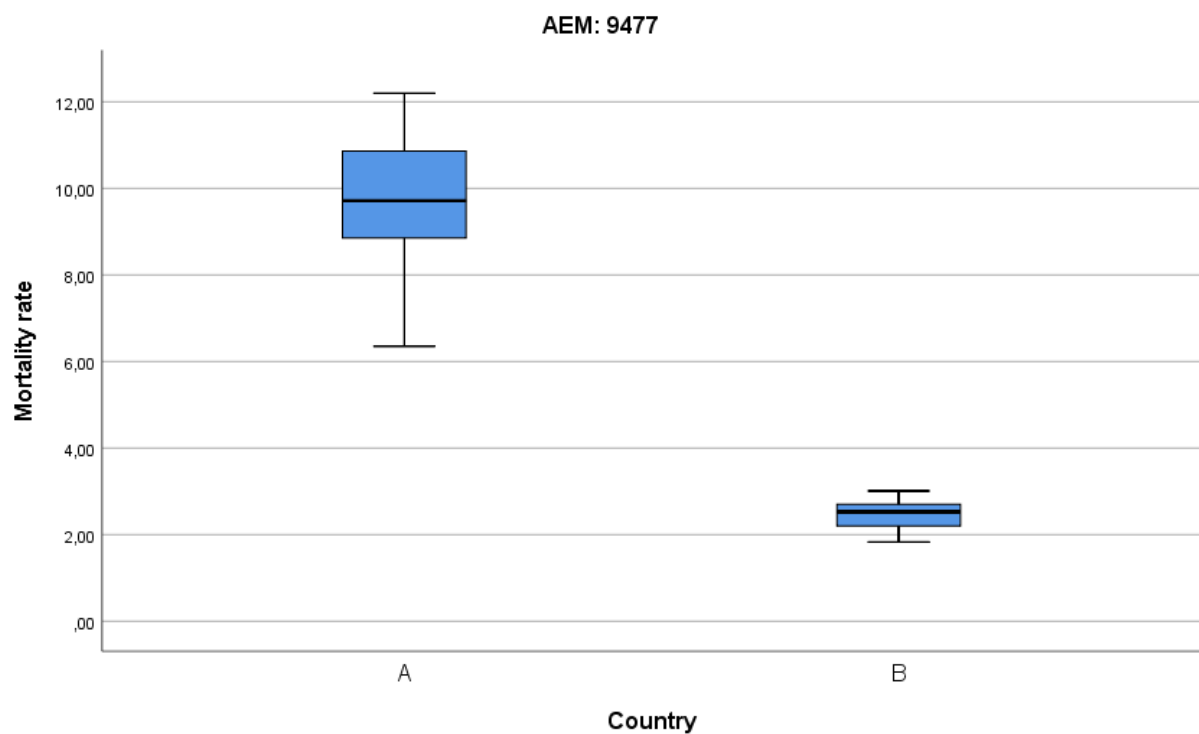
| | | Mortality rate in country A | Mortality rate in country B |
|----------------|---------|--------------------------------|--------------------------------|
| N | Valid | 20 | 20 |
| | Missing | 0 | 0 |
| Mean | | 9,6510 | 2,4650 |
| Median | | 9,7150 | 2,5250 |
| Std. Deviation | | 1,63049 | ,35198 |
| Variance | | 2,659 | ,124 |
| Range | | 5,85 | 1,18 |
| Percentiles | 25 | 8,8400 | 2,1650 |
| | 75 | 10,8650 | 2,7050 |

Η χώρα Α φαίνεται να έχει πολύ μεγαλύτερης τιμές και σε μεγαλύτερο εύρος σε σχέση με την Β.





Από τα ιστογράμματα παρατηρούμε ότι και στις 2 περιπτώσεις, έχουμε σε μεγαλύτερη συχνότητα μεσαίες τιμές αλλά και μία μικρή αύξηση στις χαμηλές.



Από τα θηκογράμματα συμπεραίνουμε ότι τα δύο δήγματα ακολουθούν κανονική κατανομή, με μία όμως μικρή επιφύλαξη λόγω της διαφοράς των μυστάκων στην A και της διαμέσου στην B

που φαίνεται να τείνει προς το 3^ο τεταρτημόριο. Είναι φανερό για άλλη μία φορά ότι στην χώρα Α οι τιμές είναι πολύ μεγαλύτερες.

2.

One-Sample Test AEM:9477

| Test Value = 10 | | | | | | |
|-----------------------------|---------|----|-----------------|-----------------|---|---------|
| | t | df | Sig. (2-tailed) | Mean Difference | 95% Confidence Interval of the Difference | |
| | | | | | Lower | Upper |
| Mortality rate in country A | -,957 | 19 | ,350 | -,34900 | -1,1121 | ,4141 |
| Mortality rate in country B | -95,737 | 19 | ,000 | -7,53500 | -7,6997 | -7,3703 |

Στην περίπτωση του Α, η μέση τιμή μπορεί να είναι 10% σε διάστημα εμπιστοσύνης 95%. Αντιθέτως, στην περίπτωση του Β παίρνει πολύ χαμηλότερες τιμές και δεν μπορεί σε καμία περίπτωση να φτάσει την τιμή 10%.

One-Sample Test AEM:9477

| Test Value = 0 | | | | | | |
|-----------------------------|--------|----|-----------------|-----------------|---|---------|
| | t | df | Sig. (2-tailed) | Mean Difference | 95% Confidence Interval of the Difference | |
| | | | | | Lower | Upper |
| Mortality rate in country A | 26,471 | 19 | ,000 | 9,65100 | 8,8879 | 10,4141 |
| Mortality rate in country B | 31,319 | 19 | ,000 | 2,46500 | 2,3003 | 2,6297 |

Για την χώρα Α, έχουμε δειγματική μέση τιμή 9,65 η οποία βρίσκεται μέσα στο διάστημα [8,88 , 10,41]. Ομοίως και για την χώρα Β, έχουμε 2,46 που βρίσκεται στο διάστημα [2,30 , 2,62]. Η εκτίμηση της μέσης τιμής φαίνεται να είναι σωστή. Έχουμε επιφύλαξη για την εγκυρότητα των αποτελεσμάτων λόγω του μικρού δείγματος (<30).

3.

Independent Samples Test AEM:9477

| | | 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 | |
| Mortality rate | Equal variances assumed | 28,043 | ,000 | 19,266 | 38 | ,000 | 7,18600 | ,37299 | 6,43093 | 7,94107 |
| | Equal variances not assumed | | | 19,266 | 20,767 | ,000 | 7,18600 | ,37299 | 6,40980 | 7,96220 |

Το διάστημα δεν περιλαμβάνει το 0 επομένως η μέση τιμή δεν μπορεί να είναι ίδια στις 2 χώρες.

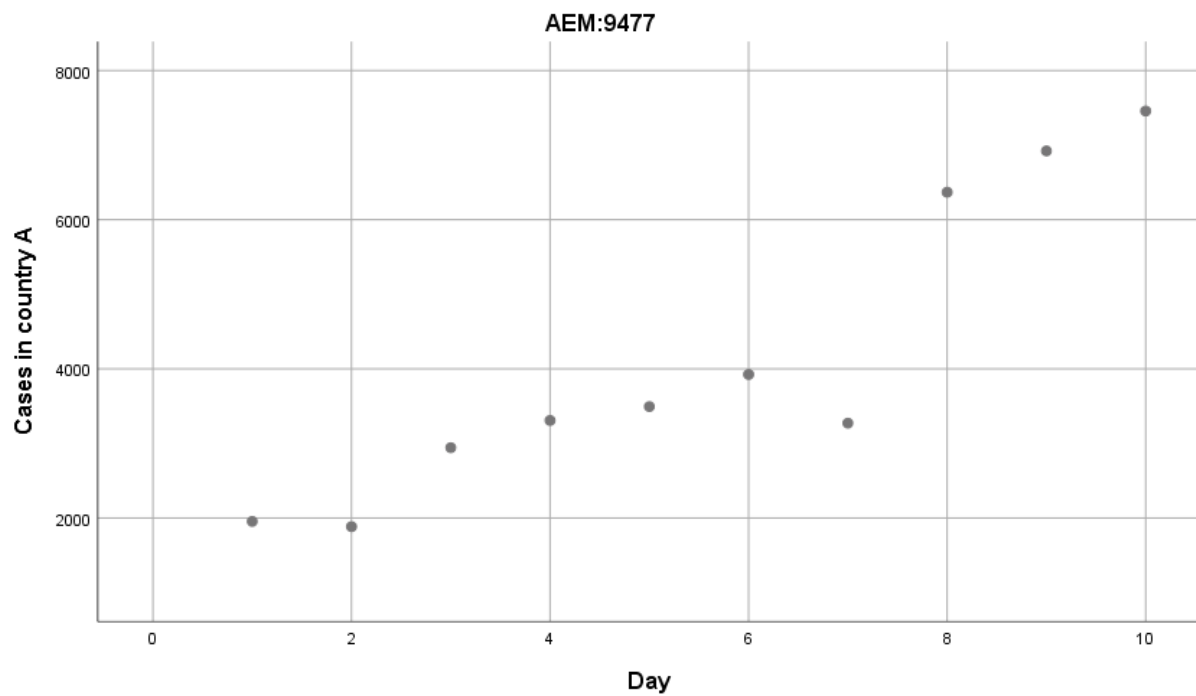
4.

Μπορούμε να πούμε ότι βλέπουμε ένα χαρακτηριστικό ποσοστό όπου όμως σε καμία περίπτωση δεν είναι ίδιο στις 2 χώρες. Για την χώρα Α είναι γύρω στο 10% ενώ για την Β στο 2,7%.

Μελέτη Β

Θα εξετάσουμε αν κατά την εξάπλωση του ιού στην χώρα, υπάρχει γραμμική συσχέτιση των ημερών με τα νέα κρούσματα. Χρησιμοποιούμε δείγματα από 10 μέρες πριν την κορύφωση έως και την μέρα πριν την κορύφωση.

5.

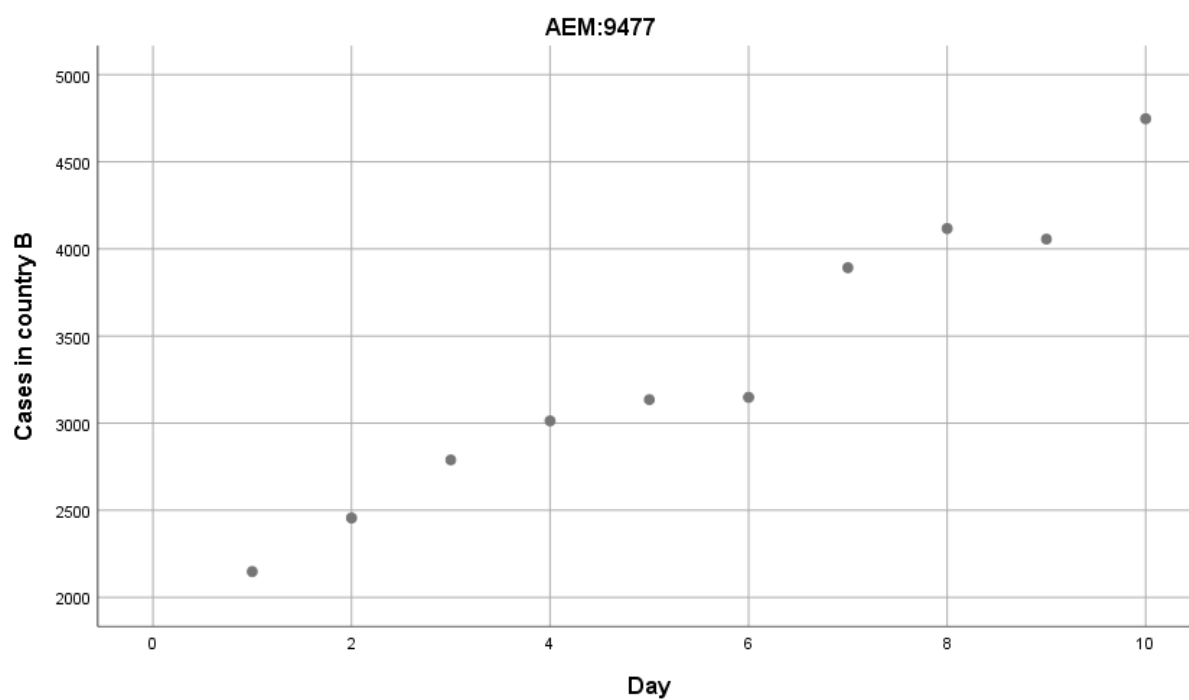


Correlations AEM:9477

| | | Day | Cases in country A |
|--------------------|---------------------|---------|--------------------|
| Day | Pearson Correlation | 1 | -,926** |
| | Sig. (2-tailed) | | ,000 |
| | N | 10 | 10 |
| Cases in country A | Pearson Correlation | -,926** | 1 |
| | Sig. (2-tailed) | ,000 | |
| | N | 10 | 10 |

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

Η συσχέτιση είναι ισχυρή (>0.9) επομένως υπάρχει εξάρτηση ανάμεσα στην ημέρα και τα ημερήσια κρούσματα.



Correlations AEM:9477

| | | Day | Cases in country B |
|--------------------|---------------------|---------|--------------------|
| Day | Pearson Correlation | 1 | -,979** |
| | Sig. (2-tailed) | | ,000 |
| | N | 10 | 10 |
| Cases in country B | Pearson Correlation | -,979** | 1 |
| | Sig. (2-tailed) | ,000 | |
| | N | 10 | 10 |

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

Και εδώ ο συντελεστής συσχέτισης είναι υψηλός επομένως επαναλαμβάνονται τα προηγούμενα συμπεράσματα.

6.

Model Summary A AEM:9477

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|-------|-------------------|----------|-------------------|----------------------------|
| 1 | ,926 ^a | ,858 | ,840 | 809,173 |

a. Predictors: (Constant), DayTest

R square = 0,858 άρα κατά 85,8% κατάλληλο για προβλέψεις.

AEM:9477 Coefficients^a

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|-------|------------|-----------------------------|------------|---------------------------|-------|------|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 744,667 | 552,770 | | 1,347 | ,215 |
| | Day | 619,642 | 89,087 | ,926 | 6,955 | ,000 |

a. Dependent Variable: Cases in country A

Model Summary B AEM:9477

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|-------|-------------------|----------|-------------------|----------------------------|
| 1 | ,979 ^a | ,958 | ,953 | 179,053 |

a. Predictors: (Constant), DayTest

R square = 0,958 άρα κατά 95,8% κατάλληλο για προβλέψεις.

AEM:9477 Coefficients^a

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|-------|------------|-----------------------------|------------|---------------------------|--------|------|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 1887,400 | 122,317 | | 15,430 | ,000 |
| | Day | 265,945 | 19,713 | ,979 | 13,491 | ,000 |

a. Dependent Variable: Cases in country B

Και τα δύο μοντέλα είναι κατάλληλα για προβλέψεις όμως φαίνεται να διαφέρουν αρκετά. Το 1^ο έχει μεγαλύτερη κλίση και μικρότερο συντελεστή B κάτι που αντικατοπτρίζεται από το εύρος τιμών του.

7.

Προβλέπουμε την μέρα 11.

$$A: 619,642 \cdot (x=11) + 744,667 (+-) 809,173 = [6751,556, 8369,902]$$

Το 8271 ανήκει σε αυτό το διάστημα.



$$B: 265,945 \cdot (x=11) + 1887,400 (+-) 179,053 = [4633,742, 4991,848]$$

Το 5138 δεν ανήκει σε αυτό το διάστημα επομένως απέτυχε η πρόβλεψη παρόλο που το θεωρήσαμε ως πιο κατάλληλο το μοντέλο.




Επιπλέον

Παρατίθενται οι πίνακες με όλα τα δεδομένα και για τις 2 μελέτες:

A.

|  Mortality RateA |  Mortality RateB |
|---|---|
| 11,61 | 2,47 |
| 10,71 | 2,56 |
| 9,96 | 2,90 |
| 10,09 | 2,88 |
| 8,92 | 2,68 |
| 8,83 | 2,27 |
| 9,11 | 2,43 |
| 8,24 | 2,61 |
| 8,88 | 2,38 |
| 7,53 | 1,96 |
| 11,00 | 1,93 |
| 10,87 | 3,01 |
| 9,47 | 2,69 |
| 8,87 | 2,49 |
| 6,35 | 2,13 |
| 6,88 | 2,71 |
| 11,88 | 2,76 |
| 10,77 | 2,67 |
| 12,20 | 1,94 |
| 10,85 | 1,83 |

B.

|  CasesA |  CasesB |  Day |
|--|--|---|
| 7457 | 4747 | 10 |
| 6922 | 4056 | 9 |
| 6368 | 4117 | 8 |
| 3272 | 3892 | 7 |
| 3925 | 3148 | 6 |
| 3494 | 3135 | 5 |
| 3308 | 3013 | 4 |
| 2943 | 2789 | 3 |
| 1884 | 2456 | 2 |
| 1954 | 2148 | 1 |