

ΜΑΘΗΜΑ: “Πληροφορική με Εφαρμογές Στατιστικής”

ΕΡΓΑΣΤΗΡΙΟ 10: Γραφήματα με το CALC – Γραμμική Παλινδρόμηση

Ασκήσεις

1. Τα παρακάτω δεδομένα αφορούν τις πωλήσεις μπίρας (σε εκατομμύρια βαρέλια), για 10 διαφορετικές ετικέτες, καθώς και την αντίστοιχη διαφημιστική δαπάνη (σε εκατομμύρια \$) γι' αυτές. Τα δεδομένα βρίσκονται στο αρχείο Lab10.ods, φύλλο εργασίας EX3. Χρησιμοποιώντας τα δεδομένα αυτά,

(i) να υπολογίσετε την τιμή του συντελεστή γραμμικής συσχέτισης r και να κατασκευάσετε το διάγραμμα διασποράς των δεδομένων. Να ερμηνεύσετε τα αποτελέσματα.

(ii) να βρεθεί η εξίσωση της προσαρμοσμένης ευθείας παλινδρόμησης $\hat{y} = \hat{a} + \hat{b}x$, όπου Y =πωλήσεις και X =διαφημιστική δαπάνη.

(iii) να απεικονίσετε στο διάγραμμα διασποράς στο (i) και την προσαρμοσμένη ευθεία παλινδρόμησης.

(iv) Υπολογίστε τις διαφορές $e = y - \hat{y}$ και να κατασκευάσετε το διάγραμμα διασποράς (e, \hat{y}) .

(v) Χρησιμοποιώντας ως 3^η μεταβλητή την τιμή του αλκοόλ κάθε μπίρας, να κατασκευάσετε το bubble chart των X =διαφημιστική δαπάνη και Y =πωλήσεις.

Ετικέτα	Διαφημιστική δαπάνη (εκατ. \$)	Πωλήσεις (σε εκατ. βαρέλια)	Αλκοόλ (%)
Budweiser	120,0	36,3	5
Bud Light	68,7	20,7	4.2
Miller Lite	100,1	15,9	4.2
Coors Light	76,6	13,2	4.2
Busch	8,7	8,1	5.9
Natural Light	0,1	7,1	4.2
Miller Genuine Draft	21,5	5,6	4.7
Miller High Lite	1,4	4,4	4.2
Busch Lite	5,3	4,3	4.2
Milwaukee's Best	1,7	4,3	4.5

2. Στο αρχείο Lab10.ods, φύλλο εργασίας Askisi2.

(i-hw) Να υπολογίσετε την τιμή του συντελεστή γραμμικής συσχέτισης r και να κατασκευάσετε το διάγραμμα διασποράς των δεδομένων. Να ερμηνεύσετε τα αποτελέσματα.

(ii-hw) Να βρεθεί η εξίσωση της προσαρμοσμένης ευθείας παλινδρόμησης $\hat{y} = \hat{a} + \hat{b}x$.

(iii-hw) Να απεικονίσετε στο διάγραμμα διασποράς στο (i) και την προσαρμοσμένη ευθεία παλινδρόμησης.

(iv) Να υπολογίσετε τις τιμές των ποσοτήτων $\hat{e} = y - \hat{y}$. Στη συνέχεια να βρεθούν οι τιμές MIN, Q1, δ, Q3, MAX. Να χαρακτηρίσετε τη συμπεριφορά του συνόλου των τιμών \hat{e} ως προς τη συμμετρία και ως προς την κύρτωση.

- (v) Να υπολογίσετε τις τιμές των παρακάτω αθροισμάτων $\sum_{i=1}^n (y_i - \bar{y})^2$, $\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ και $\sum_{i=1}^n (y_i - \hat{y}_i)^2$, όπου y_i είναι η i -οστή τιμή στη στήλη των Y , $i = 1, 2, \dots, n$ (εδώ το $n = 50$) και \hat{y}_i είναι η i -οστή προσαρμοσμένη τιμή.
- (vi) Υπολογίστε το πηλίκο $\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$ και στη συνέχεια βρείτε την τετραγωνική του ρίζα. Τι παρατηρείτε;
- (vii) Υπολογίστε τις ποσότητες $\sqrt{\frac{1}{n-2} \cdot \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$ και $\sqrt{\frac{1}{n(n-2)} \cdot \sum_{i=1}^n (y_i - \hat{y}_i)^2 \cdot \frac{\sum_{i=1}^n x_i^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$

Πριν φύγετε, να αποθηκεύσετε (σε δικό σας φάκελο, είτε στο σκληρό δίσκο του Η/Υ στο εργαστήριο, είτε σε “φλασάκι”) το βιβλίο εργασίας που δημιουργήσαμε στο 10ο εργαστήριο. Να δώσετε το όνομα LAB10.ods. Ενδέχεται να το χρησιμοποιήσουμε και σε επόμενη διάλεξη (οπότε φροντίστε να το πάρετε μαζί σας φεύγοντας).