**ΕΡΓΑΣΙΑ 2023**

**MAT\_DATA**

Moderately abnormal cells are found on the surface of the cervix. CIN2 is usually caused by certain types of human papillomavirus (HPV) and is found when a cervical biopsy is done. CIN2 is not cancer, but may become cancer and spread to nearby normal tissue if not treated. Treatment for CIN2 may include cryotherapy, laser therapy, loop electrosurgical procedure (LEEP), or cone biopsy to remove or destroy the abnormal tissue. CIN2 is sometimes called high-grade or moderate dysplasia. Also called cervical squamous intraepithelial neoplasia 2.

**ΣΚΟΠΟΣ ΤΗΣ ΑΝΑΛΥΣΗΣ**

Πραγματοποιήστε την στατιστική ανάλυση με σκοπό να καταλήξετε στο βέλτιστο μοντέλο για τα διαθέσιμα δεδομένα. Θεωρήστε δύο διαφορετικά μοντέλα με δύο ξεχωριστές εξαρτημένες μεταβλητές (α) age\_pregn\_0 (ως συνεχή μεταβλητή), και (β) την cin2 (ως μεταβλητή με δύο στάθμες).

Χρησιμοποιήστε το μοντέλο (β) ως μοντέλο πρόβλεψης για νέα περιστατικά. Για να το μελετήσετε αυτό, χρησιμοποιήστε το 80% των διαθέσιμων δεδομένων (training set) για την εκτίμηση του μοντέλου και χρησιμοποιήστε το υπόλοιπο 20% των δεδομένων (test set) για να αξιολογήστε το μοντέλο. Τι συμπεράσματα βγάζετε?

**ΠΕΡΙΟΡΙΣΜΟΙ**

Max 3 σελίδες γραπτού κειμένου, ακολουθούμενες με όσα γραφήματα και πίνακες θέλετε να συμπεριλάβετε στην ανάλυση σας. Μην αμελήσετε να επισυνάψετε τον κώδικα στην R που χρησιμοποιήσατε σε ένα appendix στο τέλος.

**ΑΛΛΕΣ ΠΛΗΡΟΦΟΡΙΕΣ**

Βρείτε στα αρχεία και την σχετική εργασία που δημοσιεύτηκε για αυτά τα δεδομένα.