



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

Τομέας Συστημάτων Μετάδοσης Πληροφορίας και Τεχνολογίας Υλικών

**Ηλεκτρονικός Φάκελος Υγείας & Ηλεκτρονική Συνταγογράφηση**

Τεχνική αναφορά 3ης εργαστηριακής άσκησης

στο μάθημα **«Eργαστήριο Βιοϊατρικής Τεχνολογίας»**

**Χαραλάμπους Γιώργος , Α.Μ.: 03119706**

# Μέρος Α

Στα πλαίσια αυτής της εργαστηριακής άσκησης, ζητείται μέσω μιας βιβλιογραφικής μηχανής αναζήτησης, PubΜed, να βρούμε 50 άρθρα που σχετίζονται με τον όρο *‘ e-prescription’* κατά το έτος 2016 (201χ και ΑΜ= 03119706 άρα χ=6) και στο επόμενο ερώτημα να εμφανιστούν: a) τα IDs και των 50 άρθρων που συλλέχθηκαν, b) οι τίτλοι των 10 πρώτων άρθρων, και c) οι περιλήψεις των 2 τελευταίων.

Χρησιμοποίηθηκαν με την σειρά, οι πιο κάτω συναρτήσεις για την επίλυση των ερωτημάτων:

1. *EUtilsSummary():* εύρεση των άρθρων που σχετίζονται με το ζητούμενο έτος
2. *summary()*: με βάση ποιών λέξεων και περιορισμών έγινε η αναζήτηση
3. *QueryId():* τον αναγνωριστικό αριθμό που αντιστοιχεί στο κάθε άρθρο
4. *EUtilsGet():* παρέχει έξτρα πληροφορίες για κάθε άρθρο
5. *Data.frame():* δημιουργία dataframe των τίτλων των άρθων με τις αντιστοιχες περιλήψεις
6. *Head(): εμφάνιση των τίτλων των άρθρων*
7. *Tail(): εμφάνιση των περιλήψεων των άρθρων*

*Graphical user interface, text

Description automatically generated*

# Μέρος Β

Μέσω της R-studio το αρχείο που μας δίνεται όπου έχει μέσα αποθηκευμένες ασθένειες με κάποιους κωδικούς. Στην συνέχεια αναζητούνται οι 3 ασθένειες όπου ο κωδικός του αποτελείται με

1)disease1=αρχικό γράμμα του επιθέτου μας και τα 3 τελευταία ψηφία του ΑΜ(03119706)-> G706, όμως δεν υπάρχει και θεωρώ κωδικό G710

2)disease2= αρχικό γράμμα του ονόματός μας και τα 3 τελευταία ψηφία του ΑΜ(03119706)->C706, όμως επειδή δεν υπάρχει ασθένεια αποθηκευμένη με τέτοιο κωδικό ως δεύτερη ασθένεια θεωρώ αυτή με κωδικό C060.

3)disease3= D720 επιλέγεται τυχαία



Εφόσον βρεθεί ο κωδικός για κάθε ασθένεια αναζητούνται ξανά στην PubMed οι 3 αυτές ασθένειες και επιστρέφεται ο αριθμός των Abstract που σχετίζονται με αυτές τα τελευταία 5 χρόνια. Ακολουθούμε την ίδια διαδικασία με το μέρος Α όμως τώρα ο όρος αναζήτησης είναι οι 3 ασθένειες. Τα αποτελέσματά μας αποθηκεύονται σε τρία διαφορετικά dataframes με ονόματα *data\_5\_last\_years\_disease1*, *data\_5\_last\_years\_disease2*, *data\_5\_last\_years\_disease3* αντίστοιχα

Στο επόμενο ερώτημα ζητείται ένα ιστόγραμμα για τον αριθμό άρθρων ανά έτος και για τις 3 ασθένειες μαζί για κάθε ένα από τα τελευταία 3 χρόνια που δημοσιεύτηκαν στο PubMed. Για την επίτευξη αυτού, δημιουργούνται τρεις πίνακες-ένα για κάθε ασθένεια-που στοιχεία τους θα είναι η χρονιά που δημοσιεύτηκαν, τα άρθρα και οι τίτλοι τους, δηλαδή από το 2020 και μετά, με τον ίδιο τρόπο το προηγούμενου ερωτήματος. Για την δημιουργία κοινού ιστογράμματος απαιτείται κοινή μεταβλητή ως παράμετρος, έτσι αποθηκεύονται σε ένα νέο data.frame οι τρεις προηγούμενοι πίνακες και παράλληλα προστείθεται μια νέα στήλη ID για να ξεχωρίζουμε τις 3 ασθένειες. Το τελικό ιστόγραμμα παρατίθεται πιο κάτω.

Chart

Description automatically generated

Τέλος, δημιουργείται ένα treemap για τον αριθμό των άρθρων για τα 3 τελευταία χρόνια και για τις 3 ασθένειες. Δεδομένου ότι έχουν ήδη βρεθεί οι πίνακες των 3 ασθενειών με τα άρθρα και τις χρονολογικές τους δημοσιεύσεις στο διάστημα που μας ενδιαφέρει, μένει μόνο να υπολογίσουμε τον αριθμό των άρθρων ανά έτος για κάθε ασθένεια και μετά με την σειρά τους να αποθηκευτούν σε νέο πίνακα.Αυτό θα γίνει με την χρήση της συνάρτησης sum() και το πλήθος των άρθων για κάθε ασθένεια και κάθε χρονολογία αντίστοιχα αποθηκεύονται σε ένα μονοδιάστατο πίνακα V. Το τελικό treemap παρουσιάζεται πιο κάτω.

Chart, treemap chart

Description automatically generated