

ΕΙΣΑΓΩΓΗ ΣΤΗ ΒΙΟΪΑΤΡΙΚΗ ΜΗΧΑΝΙΚΗ



Δωροθέα Κουμίδου
ΑΜ: 03119712

Σχολή Ηλεκτρολόγων μηχανικών &
Μηχανικών Υπολογιστών

ΤΕΧΝΙΚΗ ΑΝΑΦΟΡΑ ΤΡΙΤΗΣ ΕΡΓΑΣΤΗΡΙΑΚΗΣ ΆΣΚΗΣΗΣ: ΗΛΕΚΤΡΟΝΙΚΟΣ ΦΑΚΕΛΟΣ ΥΓΕΙΑΣ & ΗΛΕΚΤΡΟΝΙΚΗ ΣΥΝΤΑΓΟΓΡΑΦΗΣΗ

ΜΕΡΟΣ Α

Στα πλαίσια της τρίτης εργαστηριακής άσκησης, μας ζητήθηκε να πάρουμε μέσω της βιβλιογραφικής μηχανής αναζήτησης, PubMed, 50 άρθρα που σχετίζονται με τον όρο 'e-prescription' κατά το έτος 2012 (201χ και ΑΜ= 03119712 άρα χ=2). Στην συνέχεια μέσω των παρακάτω εντολών εμφανίζονται στην οθόνη με την σειρά:

- α)summary(): με βάση ποιων λέξεων και περιορισμών έγινε η αναζήτηση
- β)QueryId(): τον αναγνωριστικό αριθμό που αντιστοιχεί το κάθε άρθρο

Για την εμφάνιση τίτλων και περιλήψεων χρειάστηκαν οι εντολές:

#Μέρος Α'

```
search_topic = 'e-prescription'
search_query<- EUtilsSummary(search_topic, retmax=50, mindate=2016, maxdate=2016)
summary(search_query) #βάσει ποιων λέξεων και περιορισμών έγινε η αναζήτηση
QueryId(search_query) #τον αναγνωριστικό αριθμό που αντιστοιχεί σε καθένα από τα 50 άρθρα.

records<-EUtilsGet(search_query) #περιέχονται διάφορες πληροφορίες σχετικά με τα άρθρα
data<- data.frame('Title'=ArticleTitle(records), 'Abstract'=AbstractText(records)) #τοποθετούμε τους τίτλους
view(data)
head(data$Title,10) #τίτλος των 10 πρωτων
tail(data$Abstract,2) #περιληψη 2 τελευταιων
```

Δηλαδή δημιουργούμε ένα data.frame όπου τοποθετούμε τους τίτλους των 10 πρώτων άρθρων και τις περιλήψεις των 2 τελευταίων έπειτα για την εμφάνιση τους χρησιμοποιούμε συγκεκριμένα την εντολή head() για τον τίτλο ενώ μέσω της tail() οι περιλήψεις.

ΜΕΡΟΣ Β

Διαβάζουμε μέσω της R-studio το αρχείο που μας δίνεται όπου έχει μέσα αποθηκευμένες ασθένειες με κάποιους κωδικούς. Στην συνέχεια πρέπει να αναζητήσουμε 3 ασθένειες όπου ο κωδικός του αποτελείται με

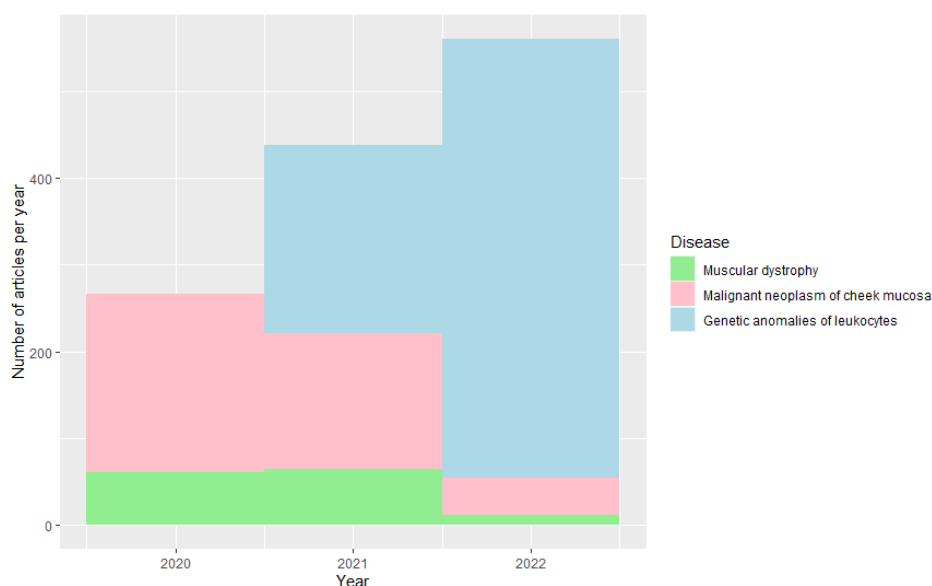
1)ασθένεια1=αρχικό γράμμα του επιθέτου μας και τα 3 τελευταία ψηφία του AM= **K712**

2)ασθένεια2= ο αρχικό γράμμα του ονόματός μας και τα 3 τελευταία ψηφία του AM=D712, όμως επειδή δεν υπάρχει ασθένεια αποθηκευμένη με τέτοιο κωδικό ως δεύτερη ασθένεια θεωρώ αυτή με κωδικό **D702**.

3)ασθένεια3= **K30** επιλέγεται τυχαία

Εφόσον λοιπόν αντιστοιχιστούν οι ασθένειες, αναζητούμε ξανά στην PubMed τις 3 αυτές ασθένειες και επιστρέφουμε τον αριθμό 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 και χρησιμοποιήσαμε ακόμα μία εντολή «na.omit» για την δημιουργία ενός πίνακα χωρίς κενά (NA) στοιχεία.

Στο επόμενο ερώτημα μας ζητείται ένα κοινό 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, θα πρέπει αρχικά να δημιουργήσουμε τρεις πίνακες-ένα για κάθε ασθένεια-που στοιχεί τους θα είναι η χρονιά που δημοσιεύτηκαν τα άρθρα και οι τίτλων τους. Η εντολή data[is.na(data)]<-0 αποσκοπεί στο να μηδενιστούν οι τιμές NA. Αφού λοιπόν μας ενδιαφέρουν μόνο τα τελευταία τρία χρόνια, δημιουργούμε άλλους αντίστοιχους πίνακες που να περιέχουν μόνο τα συγκεκριμένα άρθρα, δηλαδή από το 2020 και μετά. Για την δημιουργία κοινού ιστογράμματος απαιτείται κοινή μεταβλητή ως παράμετρος, έτσι αποθηκεύουμε σε ένα νέο data.frame τους τρεις προηγούμενους πίνακες και παράλληλα προσθέτουμε μια νέα στήλη ID για να ξεχωρίζουμε τις 3 ασθένειες.



Τέλος, δημιουργούμε ένα treemap για τον αριθμό των άρθρων για τα 3 τελευταία χρόνια και για τις 3 ασθένειες. Δεδομένου ότι έχουμε ήδη τους 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.

