

HW4 – Univariate and Multivariate Analysis

Σας δίνεται το αρχείο **results.csv** με τα αποτελέσματα των αγωνισμάτων στίβου των όλων των διοργανώσεων των Ολυμπιακών Αγώνων. Η σημασιολογία των στηλών είναι προφανής: Gender, Event, Location, Year, Medal (G, S, B σημαίνει χρυσό, αργυρό, χάλκινο), Name, Nationality, Result.

(α) Διορθώστε και/ή καθαρίστε τη στήλη Result που έχει αρκετές προβληματικές εγγραφές. Εννοείται ότι κάποιες διορθώσεις μπορείτε να τις κάνετε και χειροκίνητα. Χειριστείτε κατάλληλα τις τιμές None που εμφανίζονται στη στήλη αυτή.

(β) Χρησιμοποιήστε τεχνικές univariate και multivariate analysis και ανάλυσης χρονοσειρών και δώστε ενδιαφέροντα γραφήματα για τα παραπάνω δεδομένα μαζί με τον κατάλληλο σχολιασμό.

Καταθέστε τις απαντήσεις σας με όλες τις ενέργειές σας πάνω στα δεδομένα σε αρχείο R Notebook.

Για όσες/ους ενδιαφέρονται, παραθέτω κάποιες πηγές στις οποίες μπορείτε να βρείτε μια σύντομη και ουσιαστική εισαγωγή στα R Notebooks:

https://rstudio-pubs-static.s3.amazonaws.com/256225_63ebef4029dd40ef8e3679f6cf200a5a.html
<https://bookdown.org/yihui/rmarkdown/> (Section 3.2)