

HW3 – Data Cleaning

Μελετήστε το discussion paper [An introduction to data cleaning with R](#) των Edwin de Jonge και Mark van der Loo. Σας ενδιαφέρει το Section 3 “From technically correct data to consistent data” καθώς **το παραδοτέο της εργασίας είναι οι λύσεις των Ασκήσεων 3.1 ως 3.4 (σελίδα 51)**. Παρόλα αυτά και το υπόλοιπο περιεχόμενο είναι χρήσιμο και έχει μεγάλο ενδιαφέρον.

Για όσους/ες ενδιαφέρονται για ένα επιπλέον βιβλίο για την ηλεκτρονική βιβλιοθήκη τους, οι ίδιοι συγγραφείς εξέδωσαν το 2018 στην Wiley το βιβλίο “Statistical Data Cleaning with Applications in R” το οποίο θεωρώ ως τη βίβλο του Data Cleaning με R. Το ευχάριστο είναι ότι το συγκεκριμένο βιβλίο μπορείτε να το “κατεβάσετε” δωρεάν και νόμιμα μέσω του HEAL-Link ως φοιτητές του ΠαΜακ (<https://onlinelibrary.wiley.com/doi/book/10.1002/9781118897126>).

Καθώς το discussion paper είναι του 2013, δεν χρησιμοποιεί κάποιες βιβλιοθήκες που έχουμε δει στο μάθημα (π.χ., tidyverse). Είστε ελεύθεροι/ες να χρησιμοποιήσετε και/ή άλλες βιβλιοθήκες και συναρτήσεις για να απαντήσετε στις Ασκήσεις 3.1 ως 3.4.