

Tugas Mata Kuliah Manajemen dan Analisis Data dengan R

Eksklusi Missing Data dan Data Profile Diagram

Dibuat untuk memenuhi Tugas 2 Mata Kuliah Manajemen dan Analisis Data dengan R

Oleh:

Nurul Hidayah Chairunnisa

NPM: 131520220001



Program Studi Magister Epidemiologi

Fakultas Kedokteran

Universitas Padjadjaran

2023

Skrip R Tugas 2

#1. mengaktifkan paket readr
library(readr)

#2. mengaktifkan paket dplyr
library(dplyr)

#3. mengaktifkan paket janitor
library(janitor)

#4. membaca data pef (n=58.298)
pef = read_csv("https://raw.githubusercontent.com/dwi-agustian/biostat/main/pef.csv")

#5. identifikasi elemen duplikasi pada pef
get_dupes(pef, pidlink)

#6. menghapus elemen yang terduplikasi pada pef (n=58.297)
pef = pef[!duplicated(pef\$pidlink),]

#7. membaca data w5 (n=58.304)
w5 = read_csv("https://raw.githubusercontent.com/dwi-agustian/biostat/main/w5.csv")

#8. identifikasi elemen duplikasi pada w5
get_dupes(w5, pidlink)

#9. menghapus elemen yang terduplikasi pada w5 (n=58.303)
w5 = w5[!duplicated(w5\$pidlink),]

#10. menggabungkan data set pef dengan data set w5 berdasarkan variabel pidlink --> tidak bisa karena pidlink w5 berupa karakter
w5peflj = left_join(w5, pef, by = "pidlink")

#11. konversi pidlink di w5 menjadi angka (numerik) --> tidak bisa karena ada NA
w5\$pidlink = as.numeric(w5\$pidlink)

#12. membuat data set w5 yang tidak ada missing data pidlink
w5 = filter(w5, !is.na(pidlink))

#13. konversi pidlink di w5 menjadi angka (numerik)
w5\$pidlink = as.numeric(w5\$pidlink)

#14. menggabungkan data set pef (n=58.297) dengan data set w5 (n=58.303) berdasarkan variabel pidlink (n=58.297)
w5peflj = left_join(w5, pef, by = "pidlink")

#15. membuat data set w5peflj yang tidak ada missing data age (n=58.297)
w5peflj_a = filter(w5peflj, !is.na(age))

#16. membuat data set w5peflj yang tidak ada missing data age dan sex (n=58.297)

```
w5peflj_as = filter(w5peflj_a,!is.na(sex))
```

##17. membuat data set w5peflj yang tidak ada missing data age, sex, dan height (n=36.159)

```
w5peflj_ash = filter(w5peflj_as,!is.na(height))
```

#18. membuat data set w5peflj yang tidak ada missing data age, sex, height, dan pef
(n=26.314)

```
w5peflj_ashp = filter(w5peflj_ash,!is.na(pef))
```

Data Profile Diagram

