Tugas Mata Kuliah Manajemen dan Analisis Data Dengan R

(Membuat Diagram Data Profile)

Disusun oleh:

Isni Nurul Aini NPM 131520220003

Prodi Magister Epidemiologi

Fakultas Kedokteran Universitas Padjadjaran

2023

1. Langkah-langkah

Script:

#Mengaktifkan package

- library(readr)
- library(dplyr)
- library(janitor)

#1. Reading data pef

- pef = read_csv("https://raw.githubusercontent.com/dwi-agustian/biostat/main/pef.csv")

#1a. deduplikasi

- pef = pef[!duplicated(pef$pidlink),]

#2. Reading data w5

- w5 = read_csv("https://raw.githubusercontent.com/dwi-agustian/biostat/main/w5.csv")
- length(unique(w5$pidlink))
- get_dupes(w5,pidlink)

#2a. deduplikasi

- w5 <- w5[!duplicated(w5$pidlink),]

#3. Konversi pidlink di w5 dari character jadi angka(numeric)

- w5$pidlink = as.numeric(w5$pidlink)
- summary(w5$pidlink)

#4. Observasi pidlink tidak valid

- str(w5$pidlink)

#5. Memilih berdasarkan kritera pidlink missing

- w5 = filter(w5,!is.na(pidlink))

#6. Combining dataset w5 (58.297) vs pef (58.297)

- w5_pef_lj = left_join(w5, pef, by = "pidlink")
- w5_pef_rj = right_join(w5, pef, by = "pidlink")
- w5_pef_ij = inner_join(w5, pef, by = "pidlink")
- w5_pef_fj = full_join(w5, pef, by = "pidlink")
- names (w5_pef_lj)

#7. Mengaktifkan packages

- library(readr)
- library(dplyr)

#2. Reading data (58.297 observasi)

- pefisni <- dplyr::select(w5_pef_lj, sex, age, height, pef)

#3.Melihat missing data (N/A) (22.138 observasi N/A)

- summary(pefisni$age)
- summary(pefisni$height)
- summary(pefisni$sex)
- summary(pefisni$pef)

#4. Memilih yang tidak missing dari variabel height dan pef

- pef_NA_Height = filter(pefisni,!is.na(height))
- pef_NA_pef = filter(pef_NA_Height,!is.na(pef))

2. Diagram Data Profil

```
┌─────────────┐        ┌─────────────┐
│  Data pef   │        │  Data w5    │
│  58.298     │        │  58.304     │
└─────────────┘        └─────────────┘
      │                      │
      ├──┌──────────────┐    ├──┌──────────────┐
      │  │ Exclude      │    │  │ Exclude      │
      │  │ duplicated(1)│    │  │ duplicated(1)│
      │  │ = 58.297     │    │  │ = 58.303     │
      │  └──────────────┘    │  └──────────────┘
      │                      │
      └──┌──────────────┐    └──┌──────────────┐
         │ Exclude      │       │ Exclude      │
         │ missing (0) =│       │ missing (6) =│
         │ 58.297       │       │ 58.297       │
         └──────────────┘       └──────────────┘
```

Compile

Variables age, height, pef, sex =

58.297

Exclude missing data =

26.314

Variables:

Age 0

Height 22.138

Sex 0

Pef 20.313