

Tugas Mata Kuliah Manajemen dan Analisis Data dengan R

Disusun oleh:

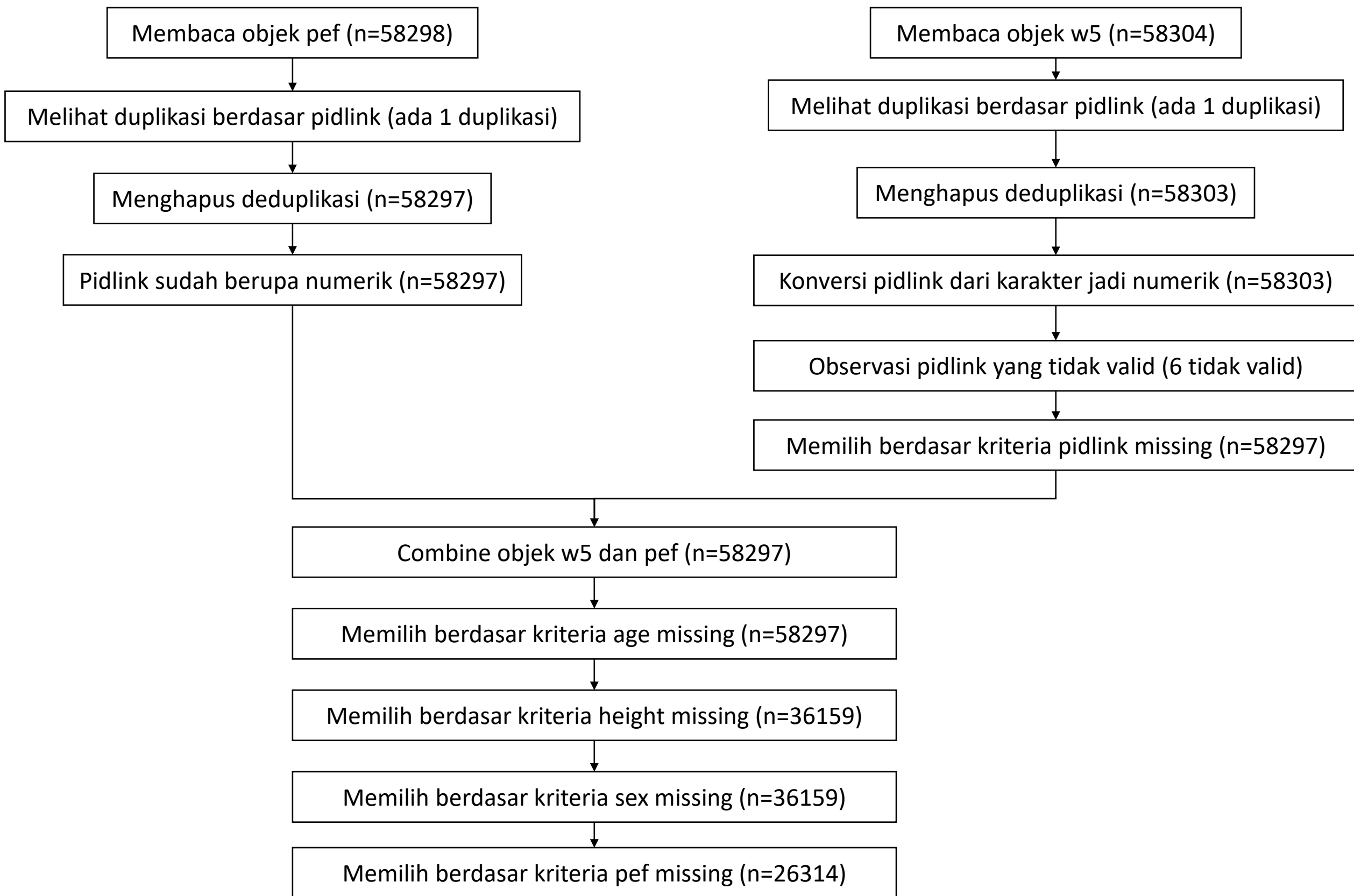
Syndi Nurmawati



Program Studi Magister S2 Epidemiologi

Fakultas Kedokteran

Universitas Padjadjaran



Kode

#1. mengaktifkan package dplyr

```
library(dplyr)
```

#2. mengaktifkan package readr

```
library(readr)
```

#3. mengaktifkan package janitor

```
library(janitor)
```

#4. membaca objek pef (n=58298)

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

#5. melihat adanya duplikasi pada objek pef (ada 1 duplikasi)

```
get_dupes(pef, pidlink)
```

#6. menghapus duplikasi/deduplikasi(n=58297)

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

#7. membaca objek w5 (n=58304)

```
w5 = read_csv('https://raw.githubusercontent.com/dwi-agustian/biostat/main/w5.csv')
```

#8. melihat deduplikasi w5 (ada duplikasi)

```
get_dupes(w5, pidlink)
```

#9. menghapus duplikasi/deduplikasi(n=58303)

```
w5 = w5[!duplicated(w5$pidlink),]
```

#10. konversi pidlink di w5 dari karakter menjadi numerik

```
w5$pidlink_c = as.numeric(w5$pidlink)
```

#11. menampilkan summary objek w5

```
summary(w5$pidlink_c)
```

#12. observasi pidlink yang tidak valid (ada 6 yang tidak valid)

```
str(w5$pidlink_c)
```

#13. memilih berdasarkan kriteria pidlink_c missing (NA) (n=58297)

```
w5 = filter(w5,!is.na(pidlink_c))
```

#14. ganti pidlink_c dengan pidlink

```
w5$pidlink = as.numeric(w5$pidlink_c)
```

#15. combining dataset (menggabungkan data) w5 (58297) vs pef (58297)

#menggabungkan variables dengan common variable

#variabel digabungkan berdasarkan data yang sebelah kiri

```
w5_pef_lj = left_join(w5, pef, by = "pidlink")
```

#16. memilih berdasarkan kriteria age missing (NA) (n=58297)

```
w5_age = filter(w5_ds_lj,!is.na(age))
```

#17. memilih berdasarkan kriteria height missing (NA) (n=36159)

```
w5_ah = filter(w5_age,!is.na(height))
```

#18. memilih berdasarkan kriteria sex missing (NA) (n=36159)

```
w5_ahs = filter(w5_ah,!is.na(sex))
```

#19. memilih berdasarkan kriteria pef missing (NA) (n=26314)

```
w5_ahsp = filter(w5_ahs,!is.na(pef))
```

#20. checking data

```
summary(w5_ahsp)
```