Pertemuan 4 - Tidyverse

ryankny

2024-09-25

Import Library

```
library(dslabs)
library(tidyverse)
```

```
## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
## ✓ dplyr     1.1.4     ✓ readr     2.1.5
## ✓ forcats   1.0.0     ✓ stringr   1.5.1
## ✓ ggplot2   3.5.1     ✓ tibble    3.2.1
## ✓ lubridate 1.9.3     ✓ tidyr     1.3.1
## ✓ purrr     1.0.2
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to be
come errors
```

```
data(murders)
# View(murders)
```

mutate

```
# tambah kolom
murders = mutate(murders, rate = total / population * 10^5)
## head(murders)

# ubah kolom
murders = mutate(murders, rate = 0)

# hapus kolom
murders = mutate(murders, rate = NULL)
```

filter

```
#filter(murders, region == 'West')
#filter(murders, 100 < total & total < 500)
filter(murders, region == 'West' | region == 'East')
```

select

```
new_table = select(murders, state, total)
new_table
```

Operator Pipe

```
sqrt(16) # tanpa operator pipe
```

```
## [1] 4
```

```
16 %>% sqrt() # menggunakan operator pipe
```

```
## [1] 4
```

```
filter(murders, total > 100) # tanpa operator pipe
```

```
murders %>% filter(total > 100) # menggunakan operator pipe
```

```
# menggunakan operator pipe
murders %>% mutate(rate = total / population * 10^5) %>% filter(rate <= 0.71)%>% select(stat
e, region, rate)
```

```
# tanpa operator pipe
select(filter(mutate(murders, rate = total / population * 10^5), rate <= 0.71),
state, region, rate)
```

sumarize

```
murders %>% summarize(jumlah = sum(total), rata_rata = mean(total))
```

group_by

```
murders %>% group_by(region) %>% summarize(jumlah = sum(total),
rata_rata = mean(total))
```

Sorting Data Frame

arrange

```
murders %>% arrange(desc(total))
```

nested_sorting

```
murders %>% arrange(region, desc(total))
```

top_n

```
murders %>% top_n(5, total)
```

Latihan Soal

```
murders = mutate(murders, rate = total / population * 10^5)
murders = mutate(murders, peringkat = rank(desc(rate)))
murders %>% filter(peringkat<=5)
```

```
murders %>% filter((region == 'Northeast' | region == 'West') & rate < 1) %>% select(state, r
ate, peringkat)
```