

Causality

Claire Liow

University of Tokyo

5/26/2022

Table of Contents

- Subset data
- Summarize data
- Add new variable

Load packages and data

```
## load packages  
library(tidyverse)  
library(qss)  
  
## load data  
resume <- read_csv("data/resume.csv")
```

Subset data

- `select`: Return columns by name/number/etc.
- `filter`: Return rows by name/number/etc.

```
## subset data with first name
```

```
resumeN <- select(resume, firstname)
```

```
## subset data with black names
```

```
resumeB <- filter(resume, race == "black")
```

```
## subset data with black, female-sounding names
```

```
resumeBf <- filter(resume, race == "black" & sex == "female")
```

Summarize data

```
## callback rate for black female names  
Bf_callback <- filter(resume, race == "black" & sex == "female")  
  summarize(callback_rate = mean(call, na.rm = TRUE))  
  
## callback rate for white female names  
Wf_callback <- filter(resume, race == "white" & sex == "female")  
  summarize(callback_rate = mean(call, na.rm = TRUE))  
  
## difference between white and black women  
Wf_callback - Bf_callback
```

Add new variable

- calculate target values
- create factor variable

```
## calculate race gap in callback rate
```

```
race_gap_by_sex <- resume %>%  
  group_by(race, sex) %>%  
  summarize(mean(callback)) %>%  
  pivot_wider(names_from = race,  
              values_from = callback) %>%  
  mutate(race_gap = white - black)
```

```
## create a factor variable that takes one of the four values
```

```
resume <- resume %>%  
  mutate(type = case_when(race == "black" & sex == "female" ~  
                           race == "black" & sex == "male" ~ "F  
                           race == "white" & sex == "female" ~  
                           race == "white" & sex == "male" ~ "W
```