# Requirements: the employees data from Exercise 17.

library(tidyr)

# Load the employees data and convert it into a tibble. Then, make the gender and title variables into a factor.

employees <- read.csv("employee-data.csv", skip=23, stringsAsFactors = FALSE)

A screenshot of a computer

Description automatically generated

# Hint: the commands are analogous to the as.data.frame() function.

employees <- as\_tibble(employees)



# Check if the data has any missing values.

any(is.na(employees))



# Arrange the dataset by gender, then last name, and make first\_name, last\_name, and gender the first three variables in the data frame. Show only the employees whose salary is higher than 70,000.

filter\_employees <- employees %>%

arrange(gender, last\_name) %>%

filter(salary > 70000) %>%

select(first\_name, last\_name, gender, everything())

A screenshot of a computer

Description automatically generated

# Use the ends\_with() function with select.

# Create a data frame containing only the employee numbers of those earning more than 70,000.

employees\_no\_more\_than\_70000 <- filter\_employees %>%

select(emp\_no)

A screenshot of a computer

Description automatically generated

# Find out how much each position (title) earns on average by gender;

# create a new variable that represents the average monthly figures

# and arrange your result by gender and average monthly salary, from largest to smallest.

result <- employees %>%

group\_by(gender, title) %>%

summarise(avg\_monthly = mean(salary)) %>%

arrange(gender, desc(avg\_monthly))

A screenshot of a computer

Description automatically generated