9.5

library(tidyverse)

library(ggthemes)

# Using the employees data, plot the salary frequencies. Limit the salaries

# under consideration to only those above 45,000 per year.

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

employees\_over\_45000 <- employees %>%

filter(salary > 45000)

# Choose an appropriate binwidth, and appearance.

# Title the x axis as "Salary", and the y axis as "Number of employees in the salary bracket".

# Title the plot as "Title distribution in the employee data".

# Install the ggthemes package and chose a theme from there. Look at the descriptions

# of each and choose one that has modifiable size and font family.

# Increase the size and set the font family to serif.

ggplot(employees\_over\_45000, aes(x = salary)) +

geom\_histogram(binwidth = 2000, fill = "#69b3a2", color = "#404040", alpha = 0.8) +

labs(title = "Title distribution in the employee data",

x = "Salary",

y = "Number of employees in the salary bracket") +

theme\_fivethirtyeight() +

theme(plot.title = element\_text(size = 20, face = "bold", family = "serif"),

axis.title = element\_text(size = 16, family = "serif"),

axis.text = element\_text(size = 14, family = "serif"),

panel.grid.major = element\_blank(),

panel.grid.minor = element\_blank(),

panel.background = element\_rect(fill = "#f5f5f5"),

panel.border = element\_blank())

A graph of a number of employees

Description automatically generated

9.6

library(tidyverse)

library(ggthemes)

# Create a bar chart representation of the number of employees in

# the different positions in the employees data by gender.

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

position\_gender\_counts <- employees %>%

group\_by(gender, title) %>%

summarise(count = n())

# Set the theme to fivethirtyeight.

# Add plot title "Job positions by gender",

# and axes titles: "Job position", and "Employee count". Can you do that?

ggplot(position\_gender\_counts, aes(x = title, y = count, fill = gender)) +

geom\_bar(stat = "identity", position = "dodge") +

labs(title = "Job positions by gender", x = "Job position", y = "Employee count") +

theme\_fivethirtyeight()

A graph of different colored columns

Description automatically generated

🡪 We can not name x and y title

# Try setting the theme to something different and give it another go.

ggplot(position\_gender\_counts, aes(x = title, y = count, fill = gender)) +

geom\_bar(stat = "identity", position = "dodge") +

labs(title = "Job positions by gender", x = "Job position", y = "Employee count")

A graph of different colored columns

Description automatically generated

🡪 Now we can name x and y title

# Look up the scale\_fill\_manual() and scale\_color\_manual() functions.

# Try setting custom colors to your plot.

ggplot(position\_gender\_counts, aes(x = title, y = count, fill = gender)) +

geom\_bar(stat = "identity", position = "dodge") +

labs(title = "Job positions by gender", x = "Job position", y = "Employee count") +

scale\_fill\_manual(values = c("#FF9999", "#66CCFF")) + # Custom fill colors for gender

scale\_color\_manual(values = c("#FF0000", "#0000FF")) # Custom outline colors for bars

A graph of different colored columns

Description automatically generated

# Try to change the position on the canvas the legend is in.

# Use the theme() function to do that.

ggplot(position\_gender\_counts, aes(x = title, y = count, fill = gender)) +

geom\_bar(stat = "identity", position = "dodge") +

labs(title = "Job positions by gender", x = "Job position", y = "Employee count") +

# theme\_fivethirtyeight() +

theme(legend.position = "bottom")

A graph of different colored bars

Description automatically generated

# Try to reverse the aesthetic mappings.

ggplot(position\_gender\_counts, aes(x = count, y = title, fill = gender)) +

geom\_bar(stat = "identity", position = "dodge") +

labs(title = "Job positions by gender", x = "Employee count", y = "Job position")

# theme\_fivethirtyeight()

A graph of different colored squares

Description automatically generated

# Does this graph give you a better idea of your data? Is it easier to read? (Shouldn't be ????)

🡪 it is same