Exercise2\_answers.R

2022-12-11

#Exercise 2 Answer Key  
library(tidyr)  
library(ggplot2)  
  
#1. Load the “Iris data” using the following code  
data("iris")  
  
#2. Convert the wide iris data to long.   
# a. Columns to convert to long include: Sepal.Length, Sepal.Width, Petal.Length, and Petal.Width  
# b. Name the new names column “measurement”  
# c. Name the new value column “length”  
  
iris\_long <- iris %>% pivot\_longer(cols = c("Sepal.Length", "Sepal.Width", "Petal.Length", "Petal.Width"),  
 names\_to = "measurement",  
 values\_to = "length")  
  
#3. Return a list of unique species names in the “iris\_long” data frame  
  
unique(iris\_long$Species)

## [1] setosa versicolor virginica   
## Levels: setosa versicolor virginica

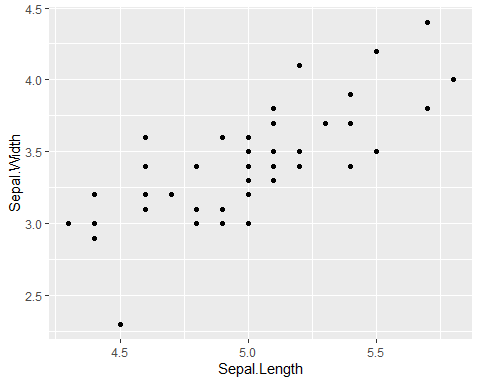
#4. Using the “iris\_long” data frame, filter only Sepal Lengths that are greater than 5.0 and   
#assign the results to a new data frame named “filtered\_iris’  
filtered\_iris <- iris\_long %>%  
 dplyr::filter(measurement == "Sepal.Length" & length > 5.0)  
filtered\_iris

## # A tibble: 118 × 3  
## Species measurement length  
## <fct> <chr> <dbl>  
## 1 setosa Sepal.Length 5.1  
## 2 setosa Sepal.Length 5.4  
## 3 setosa Sepal.Length 5.4  
## 4 setosa Sepal.Length 5.8  
## 5 setosa Sepal.Length 5.7  
## 6 setosa Sepal.Length 5.4  
## 7 setosa Sepal.Length 5.1  
## 8 setosa Sepal.Length 5.7  
## 9 setosa Sepal.Length 5.1  
## 10 setosa Sepal.Length 5.4  
## # … with 108 more rows

#5. Using the "iris\_long" data frame, select only the setosa species, group the data by measurement,   
#calculate the mean length for each measurement, and sort by the mean length for each measurement.  
  
iris\_long %>%  
 dplyr::filter(Species == "setosa") %>%  
 dplyr::group\_by(measurement) %>%  
 dplyr::summarise(Mean\_length = mean (length, na.rm = TRUE)) %>%  
 dplyr::arrange(Mean\_length)

## # A tibble: 4 × 2  
## measurement Mean\_length  
## <chr> <dbl>  
## 1 Petal.Width 0.246  
## 2 Petal.Length 1.46   
## 3 Sepal.Width 3.43   
## 4 Sepal.Length 5.01

#6. Create a scatterplot of setosa Sepal. Length (x-axis) and Sepal.Width (y-axis).   
#Note you will have to filter the original iris data set.  
  
irissub <- iris %>%   
 dplyr::filter(Species == "setosa")  
  
ggplot(irissub, aes(x = Sepal.Length, y = Sepal.Width)) +  
 geom\_point()



#7. Adjust the scatterplot created in 6 to the following: use the shape number 9,   
#increase the size to 3, change the color to anything you would like,   
#change the x-axis label to "Sepal length (cm)" and the x-aix label to "Sepal width (cm).  
  
ggplot(irissub, aes(x = Sepal.Length, y = Sepal.Width)) +  
 geom\_point(shape = 9, size = 3, color = "darkgreen") +  
 xlab("Sepal length (cm)") +   
 ylab("Sepal width (cm)")

