**HANDS-ON OUTPUT 1 (HOO1): Create a single scatter plot using different markers and colors for each iris species. Save the image and upload it as an output for today’s activity.**

# Create a scatter plot

plot(iris$Sepal.Length, iris$Sepal.Width,

xlab = "Sepal Length",

ylab = "Sepal Width",

main = "Scatter Plot: Sepal Length vs Sepal Width",

col=c("orange","purple","magenta")[as.integer(iris$Species)], pch = c(9, 10, 11)[as.integer(iris$Species)])

# Add grid lines

grid()

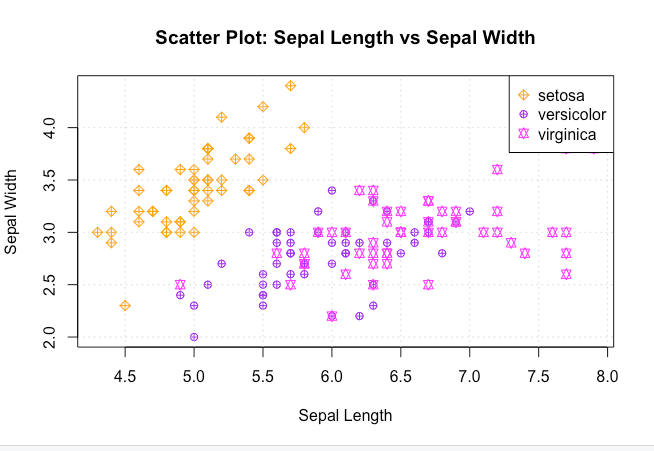
# Add a legend

legend(x="topright",

legend=c("setosa","versicolor","virginica"),

col=c("orange","purple","magenta"),

pch=c(9, 10, 11))



**HOO2: Create horizontal bar plots for the remaining continuous variables in the iris dataset. Use the command ?barplot to see which argument to include in the command.**

#Create the bar plot

barplot(iris$Petal.Length, col = c("pink","orange","lightblue")[iris$Species], ylab = "Iris", xlab = "Petal length", horiz = TRUE, main="Petal Length Bar Plot")

#Add the legend

legend("bottomright", c("setosa", "versicolor", "virginica"), col=c("lightblue","orange", "pink"), lwd=5)

A graph of different colored lines

Description automatically generated

#Create the bar plot

barplot(iris$Sepal.Width, col = c("pink","orange","lightblue")[iris$Species], ylab = "Iris", xlab = "Sepal width", horiz = TRUE, main="Sepal width Bar Plot")

#Add the legend

legend("right", c("setosa", "versicolor", "virginica"), col=c("lightblue","orange", "pink"), lwd=5)

A graph of different colors

Description automatically generated

**HOO3: Create histograms using the sepal length for the versicolor species using the following breaks: 5, 10, 20, 30.**

**A blue graph with white text

Description automatically generated**

**Code: > hist(iris$Sepal.Length[iris$Species == "versicolor"], col="blue", xlab = "Sepal Length of Versicolor", ylab = "Frequency", main = "Histogram of Sepal Length of Versicolor", breaks = 5)**

**A green bar graph with black text

Description automatically generated**

**Code: hist(iris$Sepal.Length[iris$Species == "versicolor"], col="green", xlab = "Sepal Length of Versicolor", ylab = "Frequency", main = "Histogram of Sepal Length of Versicolor", breaks = 10)**

**A graph of a number of different colored bars

Description automatically generated with medium confidence**

**Code: > hist(iris$Sepal.Length[iris$Species == "versicolor"], col="pink", xlab = "Sepal Length of Versicolor", ylab = "Frequency", main = "Histogram of Sepal Length of Versicolor", breaks = 20)**

**A graph of a number of objects

Description automatically generated with medium confidence**

**Code: > hist(iris$Sepal.Length[iris$Species == "versicolor"], col="magenta", xlab = "Sepal Length of Versicolor", ylab = "Frequency", main = "Histogram of Sepal Length of Versicolor", breaks = 30)**

**Exercises**

**Use the *mpg* dataset for the following exercises. You need to load the ggplot2 package first. Exert effort to customize and improve the aesthetic quality of your plots. Take a screenshot of your console showing the code, and the chart output. ggplots  are just for comparisons**

**1. Create a scatter plot where the x axis is “cty” and the y axis is “hwy”. Color the data points by class category.**

**A screenshot of a computer

Description automatically generated (ggplot2 plotting)**

**> ggplot(data=mpg)+aes(x=cty, y=hwy, color = class)+geom\_point()**

**A screenshot of a computer

Description automatically generated (R basic graphing)**

# Create a scatter plot

plot(mpg$cty, mpg$hwy,

xlab = "cty",

ylab = "hwy",

main = "Scatter Plot: cty vs hwy",

col=c("brown1","dodgerblue1","limegreen", "yellow", "orange", "purple", "pink")[as.factor(mpg$class)],

pch = 20)

# Add grid lines

grid()

# Add a legend

legend(x="topleft",

legend=c("2seater","compact","midzise", "minivan", "pickup", "subcompact", "suv"),

col=c("brown1","dodgerblue1","limegreen", "yellow", "orange", "purple", "pink"), pch=20)

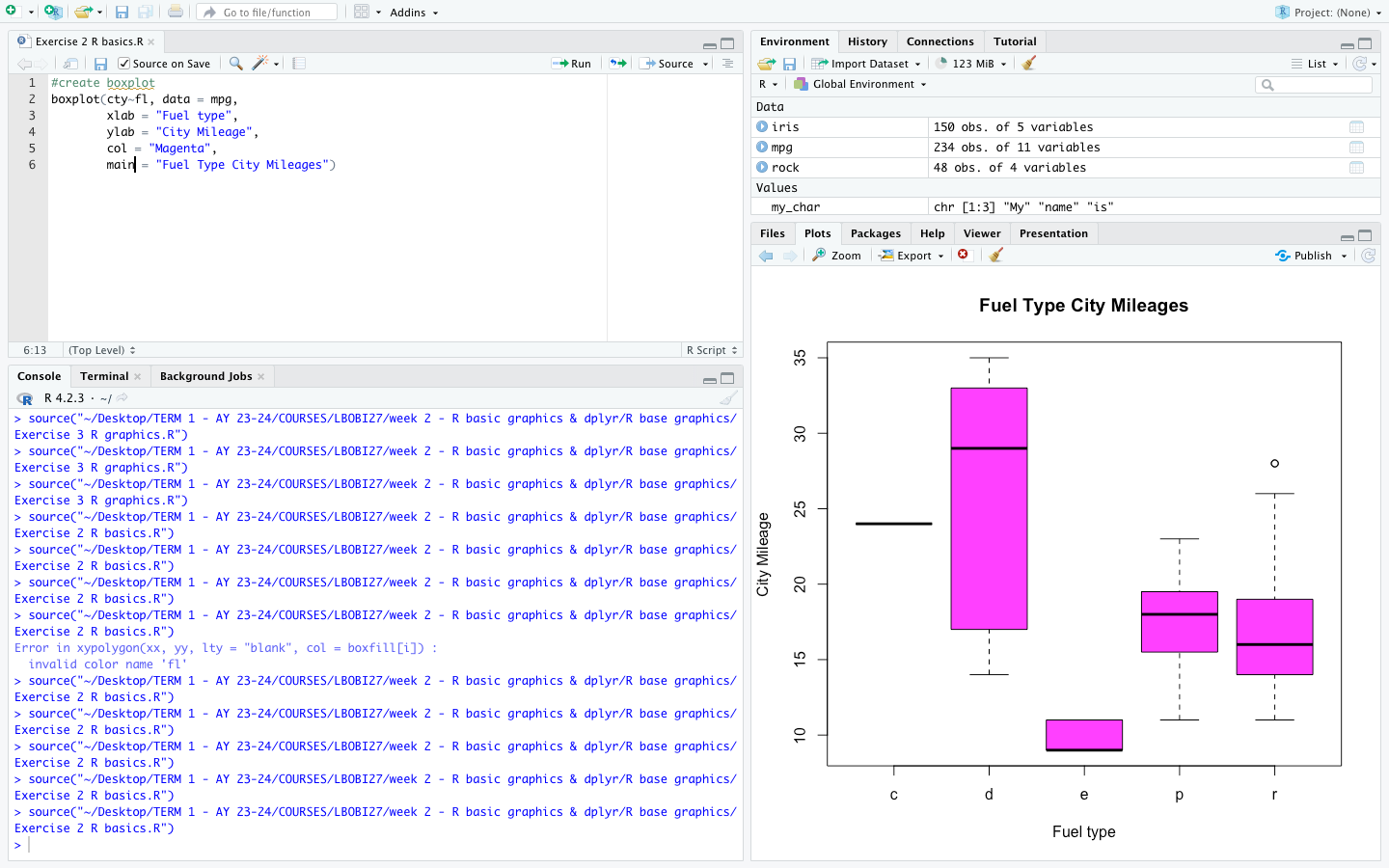
**2. Create the appropriate plot to determine which fuel type yields the best city driving mileage. Make sure that all axes are properly labeled. Why did you select your chosen plot?**

Boxplot allows a clear visualization and comparison as to what Fuel Type provides the best city driving mileage, since one variable is numeric and the other is categorical.  This plot will provide us with a clear distinction between the different fuel types with regard to the city mileages each fuel type can reach. From this plot, it is fuel “type d” that can provide the best city driving mileage.

**A screenshot of a computer

Description automatically generated(ggplot2)**

**> ggplot(mpg, aes(x=fl, y=cty, fill=fl)) + geom\_boxplot() + labs(title = “Comparison of Fuel Type’s City Mileages”, x = “Fuel Type”, y = “City Mileages”)**

**(R basics)**

**#create boxplot**

**boxplot(cty~fl, data = mpg,**

**xlab = "Fuel type",**

**ylab = "City Mileage",**

**col = "Magenta",**

**main = "Fuel Type City Mileages")**

**3. Create a vertical bar plot that shows the vehicle displacement (displ) and is colored based on the number of gears.**

**A screenshot of a computer

Description automatically generated(ggplot2)**

**>ggplot(data = mpg) + geom\_bar(mapping=aes(x=displ, fill=as.factor(cyl)))**

**A screenshot of a computer

Description automatically generated(R basic)**

#Create the bar plot

barplot(mpg$displ, col = c("yellow","blue","green", "purple"),

ylab = "Displ Value", xlab = "Displacement", main = "Displacement Values of Each Gear")

#Add the legend

legend("topright", c("4", "5", "6", "8"), title = ("Gear Number"),

col=c("yellow","blue","green", "purple"), lwd=10)