

Learning Objectives: Box Plots and Histograms

- **Create a box plot**
- **Calculate and visualize the box plot median and quartiles**
- **Create a histogram**
- **Calculate and visualize histogram frequency densities**

definition

Assumptions

- Learners are comfortable reading and importing CSV data sets, extracting relevant data into data frames, and printing that data to the console.

Limitations

- This section will cover distribution charts in brief details only and will offer practical visualization functions for learners to start creating charts right away.

Box Plots

Creating Box Plots

Follow the directions below to open up the `box.r` file in RStudio.

info

Open the `box.r` file

Within RStudio, open the `box.r` file by selecting: File → Open File...
→ code → dist → `box.r`

Data Import

```
# The mtcars dataset is natively available  
print(mtcars)
```

The basic syntax is:

```
ggplot($data, aes(x=$x, y=$y)) +  
  geom_boxplot() +  
  xlab("cyl")
```

Add on the following code into the text editor and then click the Source button to see the result.

```
plot <- ggplot(mtcars, aes(x=as.factor(cyl), y=mpg)) +  
  geom_boxplot(fill="slateblue", alpha=0.2) +  
  xlab("cyl")  
print(plot)
```

Histograms

Creating Histograms

Follow the directions below to open up the `histogram.r` file in RStudio.

info

Open the `histogram.r` file

Within RStudio, open the `histogram.r` file by selecting: File → Open File... → code → dist → `histogram.r`

Data Import

```
set.seed(1234)
df <- data.frame(sex=factor(rep(c("F", "M"), each=200)),
  weight=round(c(rnorm(200, mean=55, sd=5), rnorm(200, mean=65,
    sd=5))))
```

The basic syntax is:

```
ggplot($data, aes(x=$x)) +
  geom_histogram()
```

Add on the following code into the text editor and then click the Source button to see the result.

```
plot <- ggplot(df, aes(x=weight)) +
  geom_histogram()
print(plot)
```