



## 1.3.3: Data Visualization

### (Asynchronous-Online)

#### Session Objectives

1. To visualize data using different R packages.

Key points to cover:

1. Introduce to ggplot2 and other R packages.
  2. Visualize one, two, or more variables at a time.
  3. Introduce other resources (e.g., books, blogs, or websites) trainees can refer to.
- 

#### 0. Pework - Before You Begin

If you do not have them already, install the following packages from CRAN:

- `ggplot2`
- `ggthemes`
- `readr`

Download the small training dataset:

- [mydata.csv](#) - right click and “SaveAs” to download this datafile to your computer - see “[Module 1.3.2: Data Wrangling](#)”

Import the data into your R computing session:

- Here is the code to read the data into your computing session:

```
library(readr)
readr::read_csv("mydata.csv")
```



```
# A tibble: 21 x 14
  SubjectID Age WeightPRE WeightPOST Height SES GenderSTR GenderCoded q1
    <dbl> <dbl>    <dbl>    <dbl>  <dbl> <dbl> <chr>          <dbl> <dbl>
1         1   45        68      145    5.6    9 m             1      4
2         2   50       167      166    5.4    2 f             2      3
3         3   35       143      135    5.6    2 <NA>          NA      3
4         4   44       216      201    5.6    2 m             1      4
5         5   32       243      223    6      2 m             1      5
6         6   48       165      145    5.2    2 f             2      2
7         8   50        60      132    3.3    2 m             1      3
8         9   51       110      108    5.1    3 f             2      1
9        12   46       167      158    5.5    2 F             2      1
10       14   35       190      200    5.8    1 Male          1      4
# i 11 more rows
# i 5 more variables: q2 <dbl>, q3 <dbl>, q4 <dbl>, q5 <dbl>, q6 <dbl>
```

- Get Inspired at [The R Graph Gallery](#)



## 1. Base R graphical functions

The base R `graphics` package is very powerful on its own. As you saw in [1.3.1: Introduction to R and R Studio](#), we can make a simple 2-dimensional scatterplot with the `plot()` function.

For example, let's make a plot of `Height` on the X-axis (horizontal) and `WeightPRE` on the Y-axis (vertical) from the `mydata` dataset. Since we are using base R function, we have to use the `$`selector to identify the variables we want inside the `mydata` dataset.

Learn more about the `plot()` function and arguments by running `help(plot, package = "graphics")`.

```
plot(x = mydata$Height,  
     y = mydata$WeightPRE)
```

```
Error in eval(expr, envir, enclos): object 'mydata' not found
```

The plot does look a little odd - this is due to some data errors in the `mydata` dataset. We will fix these below. But for now, you can “see” that these data may have some issues that need to be addressed. For example:

- There are 2 people with heights < 5 feet tall which may be suspect
- There are 2 people with a weight < 100 pounds which may be data entry errors or incorrect units

For now, let's add some additional graphical elements:

- a better label for the x-axis
- a better label for the y-axis
- a title for the graph
- a subtitle for the graph

```
plot(x = mydata$Height,  
     y = mydata$WeightPRE,  
     xlab = "Height (in decimal inches)",  
     ylab = "",  
     main = "",  
     sub = "")
```

```
Error in eval(expr, envir, enclos): object 'mydata' not found
```



## 2. ggplot2 package

---



### 3. Get boilerplate code to start

**R Gallery**

**R Graphics Cookbook**

---



## References

R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.

Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2024. *Readr: Read Rectangular Text Data*. <https://readr.tidyverse.org>.

## Other Helpful Resources

[Other Helpful Resources](#)