Intro_to_R_II

2023-05-18

Intro to R (Part II)

Section 1. Set working directory, install packages, training data sets

```
# Set working directory
# setwd(<working_dir>)

# Get working directory
getwd()
```

```
## [1] "C:/Users/slava/OneDrive - University of Lethbridge/R_training"
```

```
# Install R packages
# install.packages(<package_name>)
# example:
# install.packages("ape")
# install.packages(c("ape", "MASS"))

# Install Bioconductor packages:
# Go to Bioconductor package's web page and
# copy the installation code block into the R session

# Check Bioconductor packages's web site and package vignette

# Get help for functions
# ?t.test # or

# help(t.test)

# Pre-loaded data
# data()
# data(mtcars)
# head(mtcars)
#
# data("carnivora")
# head(carnivora)
#
# data("USArrests")
# head(USArrests)
#
# ?mtcars
# ?USArrests
```

Section 2. Import and export data

```
# Write tab delimited file
data("mtcars")
write.table(mtcars, file="mtcars.txt", sep="\t")
# ?write.table

# Write comma separated file
write.csv(mtcars, file="mtcars.csv")

# Read tab delimited file
cars <-
  read.table("mtcars.txt", sep = "\t", header = T, stringsAsFactors = F)
cars
```

```
##                      mpg cyl  disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4           21.0   6 160.0 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag       21.0   6 160.0 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710          22.8   4 108.0  93 3.85 2.320 18.61  1  1    4    1
## Hornet 4 Drive      21.4   6 258.0 110 3.08 3.215 19.44  1  0    3    1
## Hornet Sportabout   18.7   8 360.0 175 3.15 3.440 17.02  0  0    3    2
## Valiant             18.1   6 225.0 105 2.76 3.460 20.22  1  0    3    1
## Duster 360          14.3   8 360.0 245 3.21 3.570 15.84  0  0    3    4
## Merc 240D           24.4   4 146.7  62 3.69 3.190 20.00  1  0    4    2
## Merc 230            22.8   4 140.8  95 3.92 3.150 22.90  1  0    4    2
## Merc 280            19.2   6 167.6 123 3.92 3.440 18.30  1  0    4    4
## Merc 280C           17.8   6 167.6 123 3.92 3.440 18.90  1  0    4    4
## Merc 450SE          16.4   8 275.8 180 3.07 4.070 17.40  0  0    3    3
## Merc 450SL          17.3   8 275.8 180 3.07 3.730 17.60  0  0    3    3
## Merc 450SLC         15.2   8 275.8 180 3.07 3.780 18.00  0  0    3    3
## Cadillac Fleetwood  10.4   8 472.0 205 2.93 5.250 17.98  0  0    3    4
## Lincoln Continental 10.4   8 460.0 215 3.00 5.424 17.82  0  0    3    4
## Chrysler Imperial   14.7   8 440.0 230 3.23 5.345 17.42  0  0    3    4
## Fiat 128            32.4   4  78.7  66 4.08 2.200 19.47  1  1    4    1
## Honda Civic         30.4   4  75.7  52 4.93 1.615 18.52  1  1    4    2
## Toyota Corolla      33.9   4  71.1  65 4.22 1.835 19.90  1  1    4    1
## Toyota Corona       21.5   4 120.1  97 3.70 2.465 20.01  1  0    3    1
## Dodge Challenger    15.5   8 318.0 150 2.76 3.520 16.87  0  0    3    2
## AMC Javelin         15.2   8 304.0 150 3.15 3.435 17.30  0  0    3    2
## Camaro Z28          13.3   8 350.0 245 3.73 3.840 15.41  0  0    3    4
## Pontiac Firebird    19.2   8 400.0 175 3.08 3.845 17.05  0  0    3    2
## Fiat X1-9           27.3   4  79.0  66 4.08 1.935 18.90  1  1    4    1
## Porsche 914-2       26.0   4 120.3  91 4.43 2.140 16.70  0  1    5    2
## Lotus Europa        30.4   4  95.1 113 3.77 1.513 16.90  1  1    5    2
## Ford Pantera L      15.8   8 351.0 264 4.22 3.170 14.50  0  1    5    4
## Ferrari Dino        19.7   6 145.0 175 3.62 2.770 15.50  0  1    5    6
## Maserati Bora       15.0   8 301.0 335 3.54 3.570 14.60  0  1    5    8
## Volvo 142E          21.4   4 121.0 109 4.11 2.780 18.60  1  1    4    2
```

```
# Read csv file
```

```
cars <- read.csv("mtcars.csv", header = T)
cars
```

```
##                      X  mpg cyl  disp  hp drat    wt  qsec vs am gear carb
## 1            Mazda RX4 21.0   6 160.0 110 3.90 2.620 16.46  0  1    4    4
## 2        Mazda RX4 Wag 21.0   6 160.0 110 3.90 2.875 17.02  0  1    4    4
## 3           Datsun 710 22.8   4 108.0  93 3.85 2.320 18.61  1  1    4    1
## 4       Hornet 4 Drive 21.4   6 258.0 110 3.08 3.215 19.44  1  0    3    1
## 5    Hornet Sportabout 18.7   8 360.0 175 3.15 3.440 17.02  0  0    3    2
## 6              Valiant 18.1   6 225.0 105 2.76 3.460 20.22  1  0    3    1
## 7           Duster 360 14.3   8 360.0 245 3.21 3.570 15.84  0  0    3    4
## 8            Merc 240D 24.4   4 146.7  62 3.69 3.190 20.00  1  0    4    2
## 9             Merc 230 22.8   4 140.8  95 3.92 3.150 22.90  1  0    4    2
## 10            Merc 280 19.2   6 167.6 123 3.92 3.440 18.30  1  0    4    4
## 11           Merc 280C 17.8   6 167.6 123 3.92 3.440 18.90  1  0    4    4
## 12          Merc 450SE 16.4   8 275.8 180 3.07 4.070 17.40  0  0    3    3
## 13          Merc 450SL 17.3   8 275.8 180 3.07 3.730 17.60  0  0    3    3
## 14         Merc 450SLC 15.2   8 275.8 180 3.07 3.780 18.00  0  0    3    3
## 15  Cadillac Fleetwood 10.4   8 472.0 205 2.93 5.250 17.98  0  0    3    4
## 16 Lincoln Continental 10.4   8 460.0 215 3.00 5.424 17.82  0  0    3    4
## 17   Chrysler Imperial 14.7   8 440.0 230 3.23 5.345 17.42  0  0    3    4
## 18            Fiat 128 32.4   4  78.7  66 4.08 2.200 19.47  1  1    4    1
## 19         Honda Civic 30.4   4  75.7  52 4.93 1.615 18.52  1  1    4    2
## 20      Toyota Corolla 33.9   4  71.1  65 4.22 1.835 19.90  1  1    4    1
## 21       Toyota Corona 21.5   4 120.1  97 3.70 2.465 20.01  1  0    3    1
## 22    Dodge Challenger 15.5   8 318.0 150 2.76 3.520 16.87  0  0    3    2
## 23         AMC Javelin 15.2   8 304.0 150 3.15 3.435 17.30  0  0    3    2
## 24          Camaro Z28 13.3   8 350.0 245 3.73 3.840 15.41  0  0    3    4
## 25    Pontiac Firebird 19.2   8 400.0 175 3.08 3.845 17.05  0  0    3    2
## 26           Fiat X1-9 27.3   4  79.0  66 4.08 1.935 18.90  1  1    4    1
## 27       Porsche 914-2 26.0   4 120.3  91 4.43 2.140 16.70  0  1    5    2
## 28        Lotus Europa 30.4   4  95.1 113 3.77 1.513 16.90  1  1    5    2
## 29      Ford Pantera L 15.8   8 351.0 264 4.22 3.170 14.50  0  1    5    4
## 30        Ferrari Dino 19.7   6 145.0 175 3.62 2.770 15.50  0  1    5    6
## 31       Maserati Bora 15.0   8 301.0 335 3.54 3.570 14.60  0  1    5    8
## 32          Volvo 142E 21.4   4 121.0 109 4.11 2.780 18.60  1  1    4    2
```

Section 3. Visualization with ggplot2

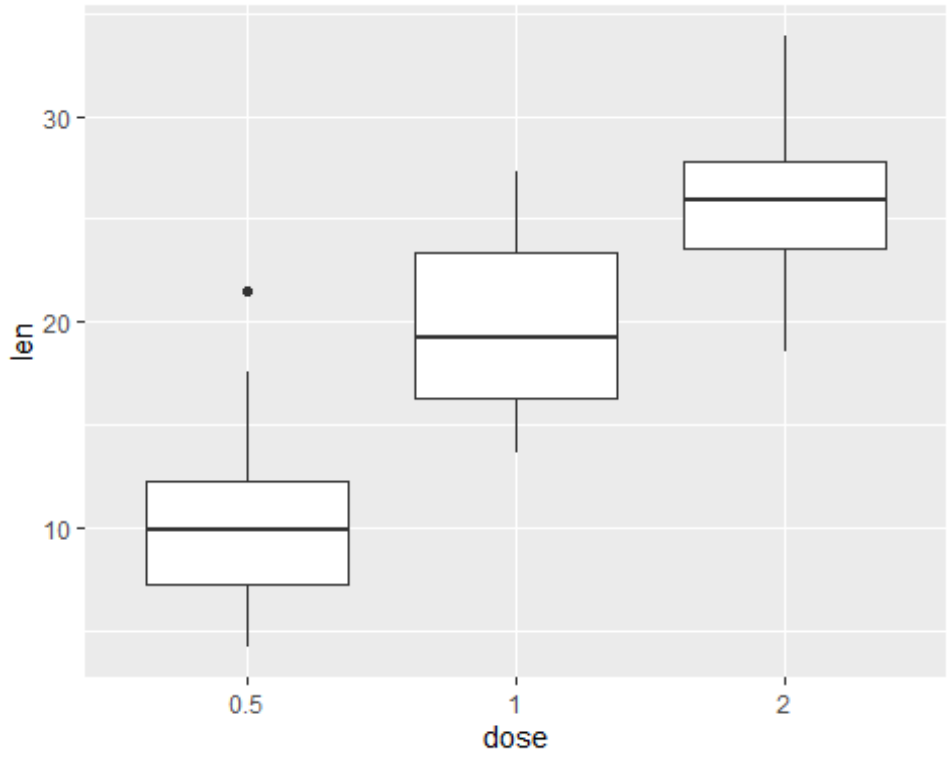
```
library(ggplot2)

# Load the data
data("ToothGrowth")
ToothGrowth
```

```
##     len supp dose
## 1   4.2   VC  0.5
## 2  11.5   VC  0.5
## 3   7.3   VC  0.5
## 4   5.8   VC  0.5
## 5   6.4   VC  0.5
```

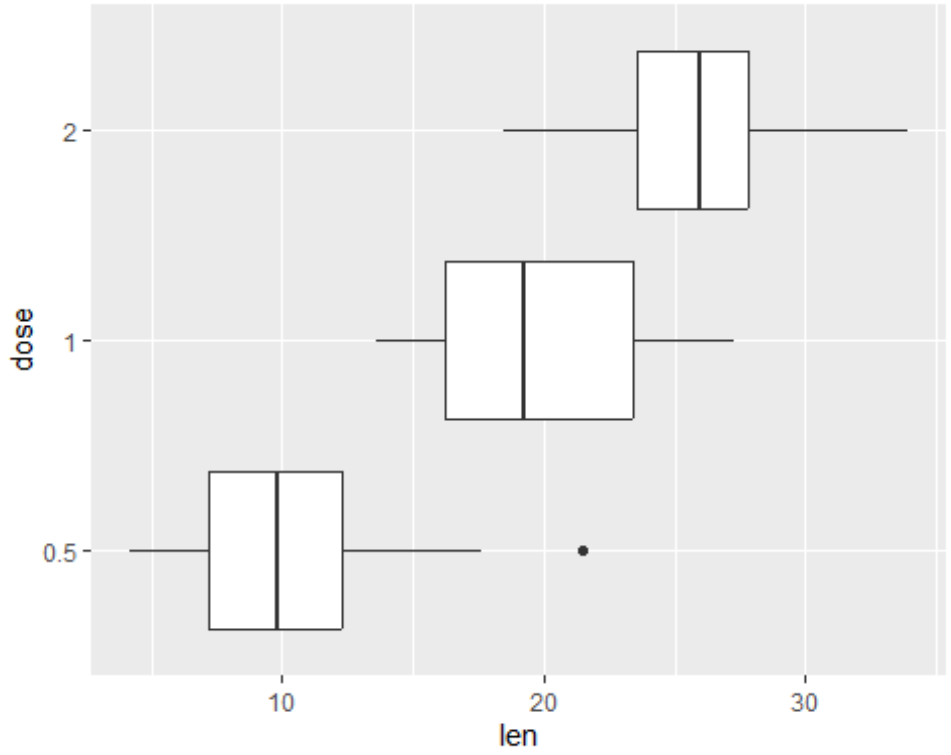
```
## 6  10.0   VC  0.5
## 7  11.2   VC  0.5
## 8  11.2   VC  0.5
## 9   5.2   VC  0.5
## 10  7.0   VC  0.5
## 11 16.5   VC  1.0
## 12 16.5   VC  1.0
## 13 15.2   VC  1.0
## 14 17.3   VC  1.0
## 15 22.5   VC  1.0
## 16 17.3   VC  1.0
## 17 13.6   VC  1.0
## 18 14.5   VC  1.0
## 19 18.8   VC  1.0
## 20 15.5   VC  1.0
## 21 23.6   VC  2.0
## 22 18.5   VC  2.0
## 23 33.9   VC  2.0
## 24 25.5   VC  2.0
## 25 26.4   VC  2.0
## 26 32.5   VC  2.0
## 27 26.7   VC  2.0
## 28 21.5   VC  2.0
## 29 23.3   VC  2.0
## 30 29.5   VC  2.0
## 31 15.2   OJ  0.5
## 32 21.5   OJ  0.5
## 33 17.6   OJ  0.5
## 34  9.7   OJ  0.5
## 35 14.5   OJ  0.5
## 36 10.0   OJ  0.5
## 37  8.2   OJ  0.5
## 38  9.4   OJ  0.5
## 39 16.5   OJ  0.5
## 40  9.7   OJ  0.5
## 41 19.7   OJ  1.0
## 42 23.3   OJ  1.0
## 43 23.6   OJ  1.0
## 44 26.4   OJ  1.0
## 45 20.0   OJ  1.0
## 46 25.2   OJ  1.0
## 47 25.8   OJ  1.0
## 48 21.2   OJ  1.0
## 49 14.5   OJ  1.0
## 50 27.3   OJ  1.0
## 51 25.5   OJ  2.0
## 52 26.4   OJ  2.0
## 53 22.4   OJ  2.0
## 54 24.5   OJ  2.0
## 55 24.8   OJ  2.0
```

```
## 56 30.9   OJ  2.0
## 57 26.4   OJ  2.0
## 58 27.3   OJ  2.0
## 59 29.4   OJ  2.0
## 60 23.0   OJ  2.0
```

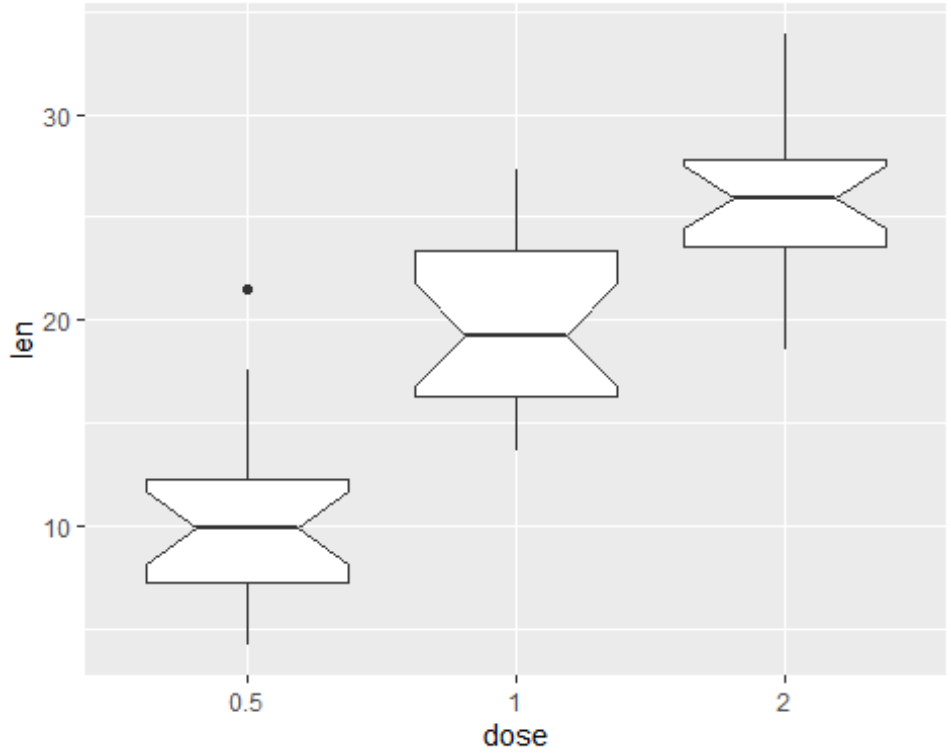
```
# Convert dose to factor
ToothGrowth$dose <- as.factor(ToothGrowth$dose)

p <- ggplot(ToothGrowth, aes(x=dose, y=len)) +
  geom_boxplot()
p
```



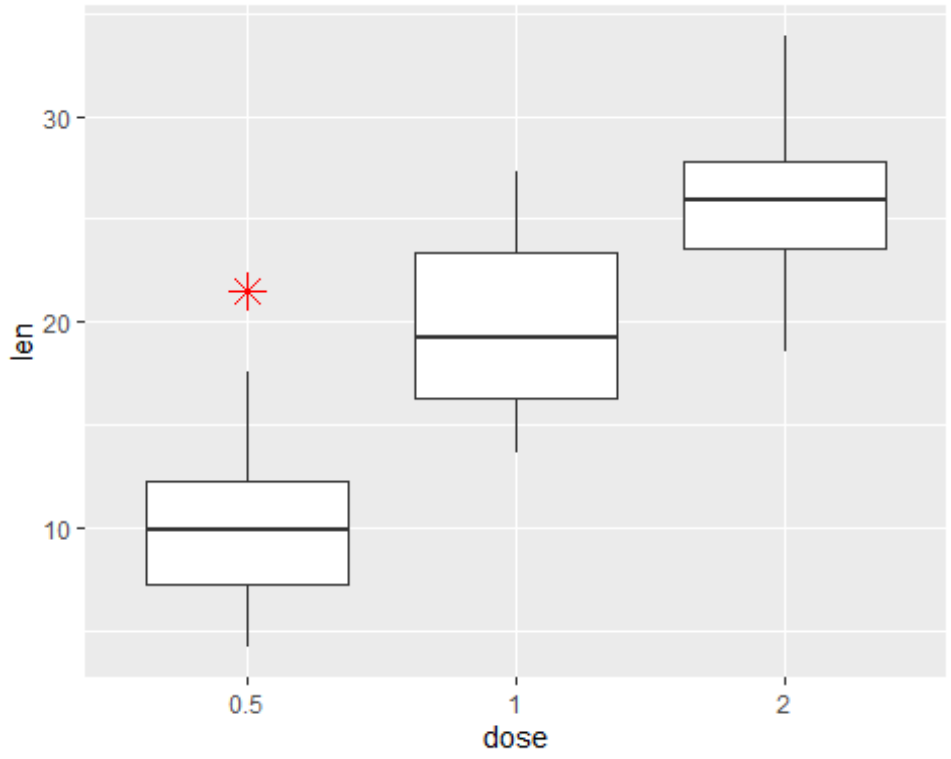

```
# Rotate the box plot
p + coord_flip()
```



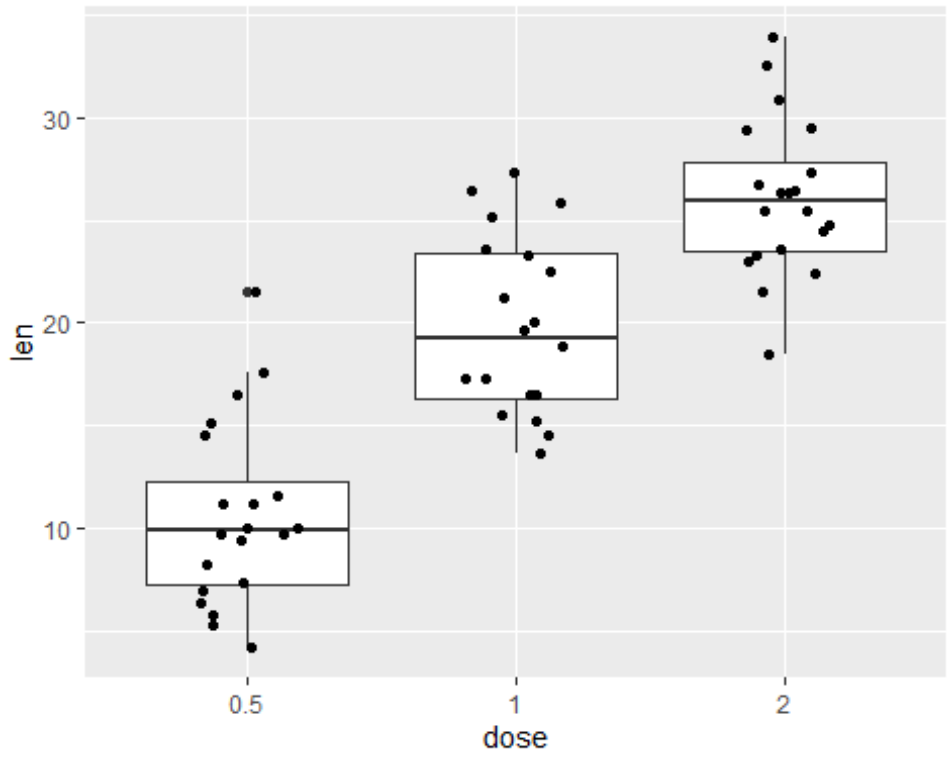

```
# Notched box plot
ggplot(ToothGrowth, aes(x=dose, y=len)) +
  geom_boxplot(notch=TRUE)
```



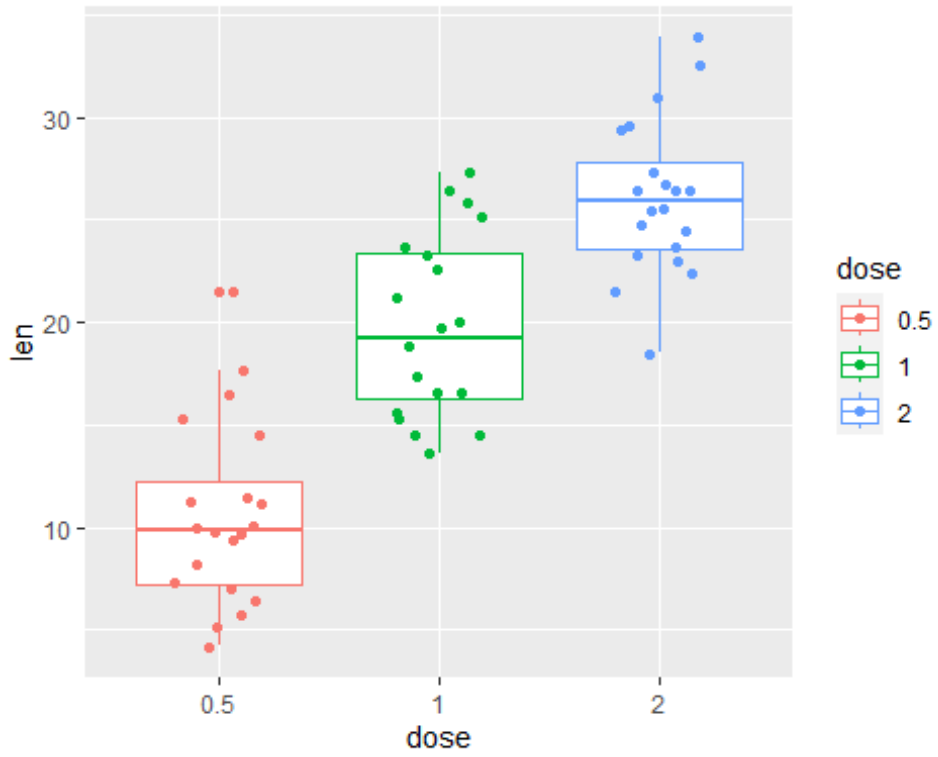
```r
# Change outlier, color, shape and size
ggplot(ToothGrowth, aes(x=dose, y=len)) +
  geom_boxplot(outlier.colour="red", outlier.shape=8,
               outlier.size=4)
```



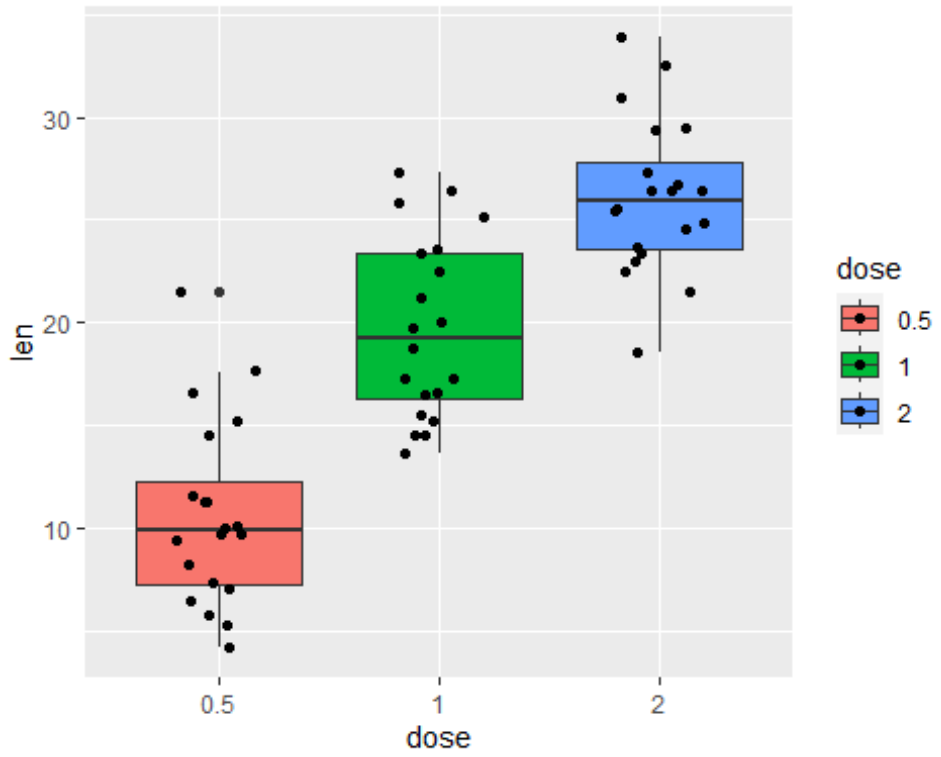

```r
# Add individual data points as jitter
p <- ggplot(ToothGrowth, aes(x=dose, y=len)) +
  geom_boxplot() + geom_jitter(shape=16, position=position_jitter(0.2))
p
```



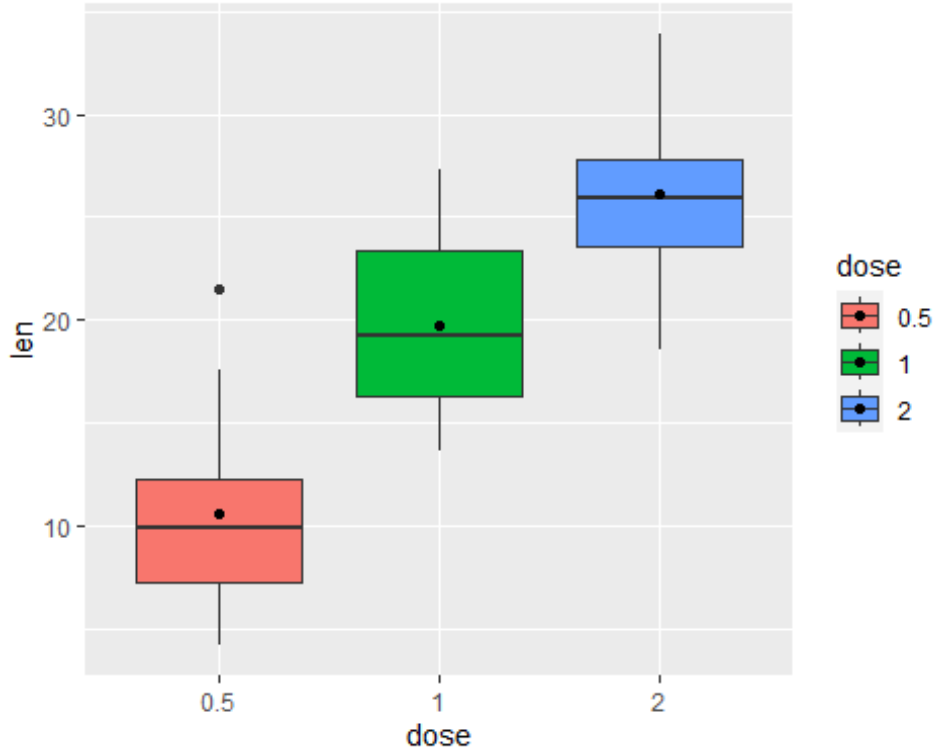

```
# Add color
p <- ggplot(ToothGrowth, aes(x=dose, y=len, color=dose)) +
  geom_boxplot() + geom_jitter(shape=16, position=position_jitter(0.2))
p
```



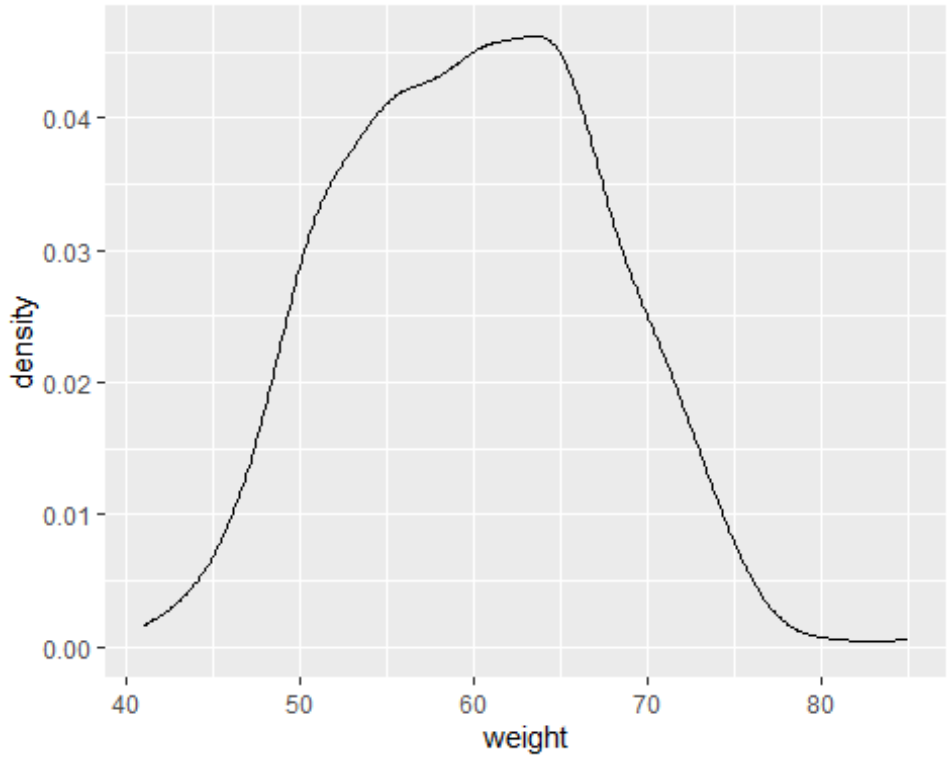

```
# Fill
p <- ggplot(ToothGrowth, aes(x=dose, y=len, fill=dose)) +
  geom_boxplot() + geom_jitter(shape=16, position=position_jitter(0.2))
p
```



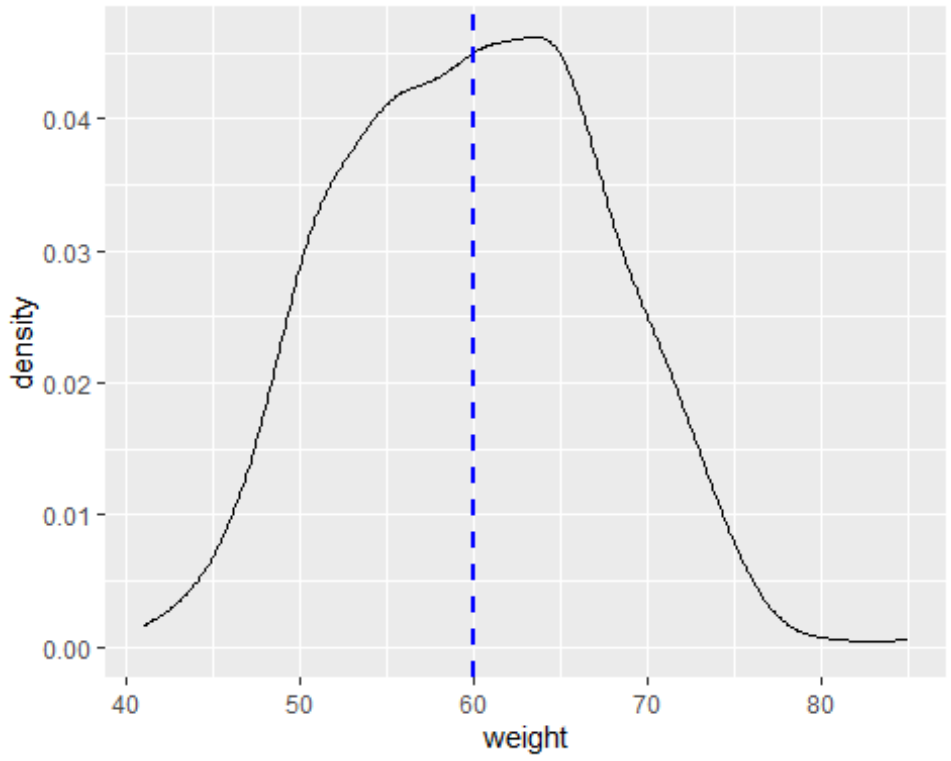

```
# Boxplot with multiple groups
ggplot(ToothGrowth, aes(x=dose, y=len, fill=dose)) +
  geom_boxplot() + stat_summary(fun="mean", geom="point")
```



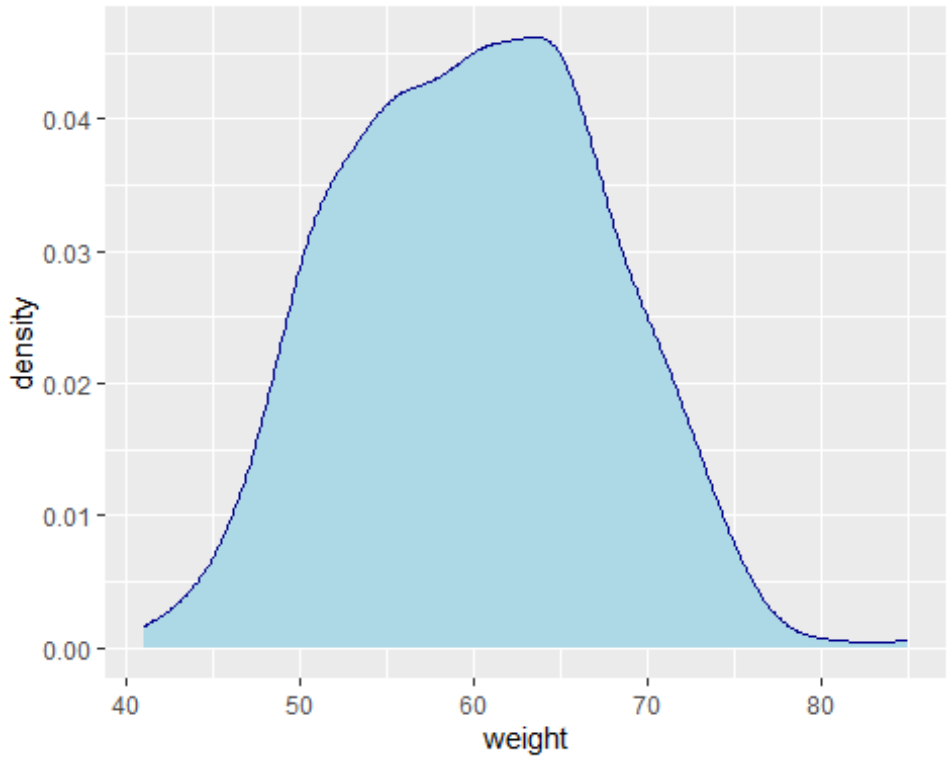
```
# ------------------------------------------------------------ #
# Density plots
df <- data.frame(
  sex=factor(rep(c("F", "M"), each=200)),
  weight=round(c(rnorm(200, mean=55, sd=5),
                 rnorm(200, mean=65, sd=5)))
  )
head(df)

##   sex weight
## 1   F     51
## 2   F     58
## 3   F     56
## 4   F     52
## 5   F     55
## 6   F     49

p <- ggplot(df, aes(x=weight)) +
  geom_density()
p
```



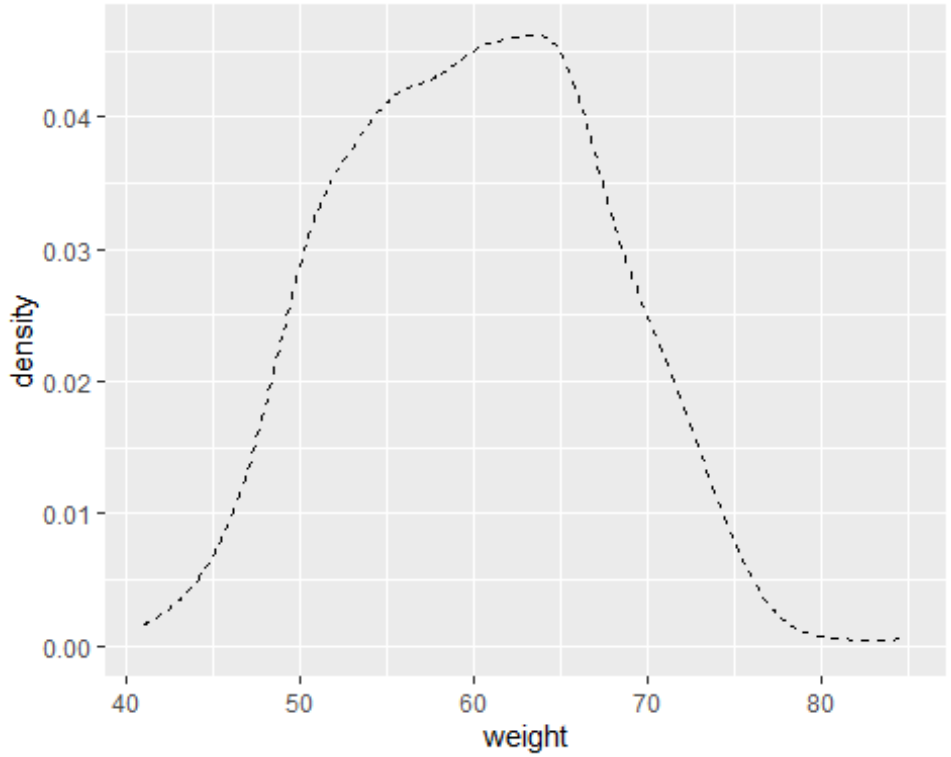

```
# Add mean line
p + geom_vline(aes(xintercept=mean(weight)),
               color="blue", linetype="dashed", size=1)
```



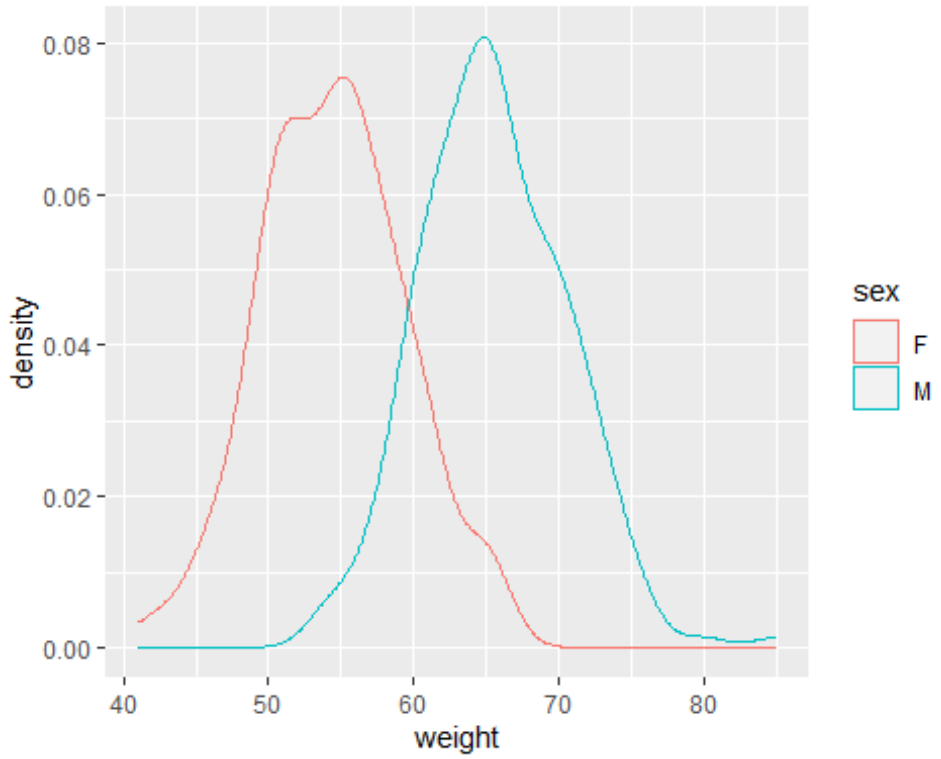

```r
# Change line color and fill color
ggplot(df, aes(x=weight)) +
  geom_density(color="darkblue", fill="lightblue")
```



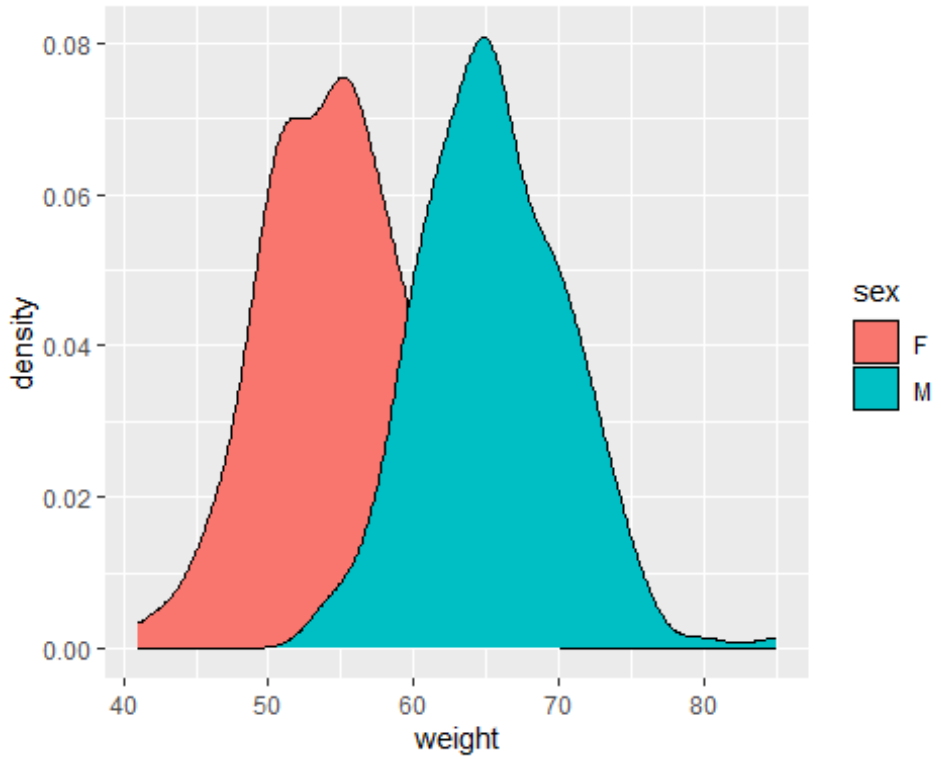
```
# Change line type
ggplot(df, aes(x=weight))+
  geom_density(linetype="dashed")
```



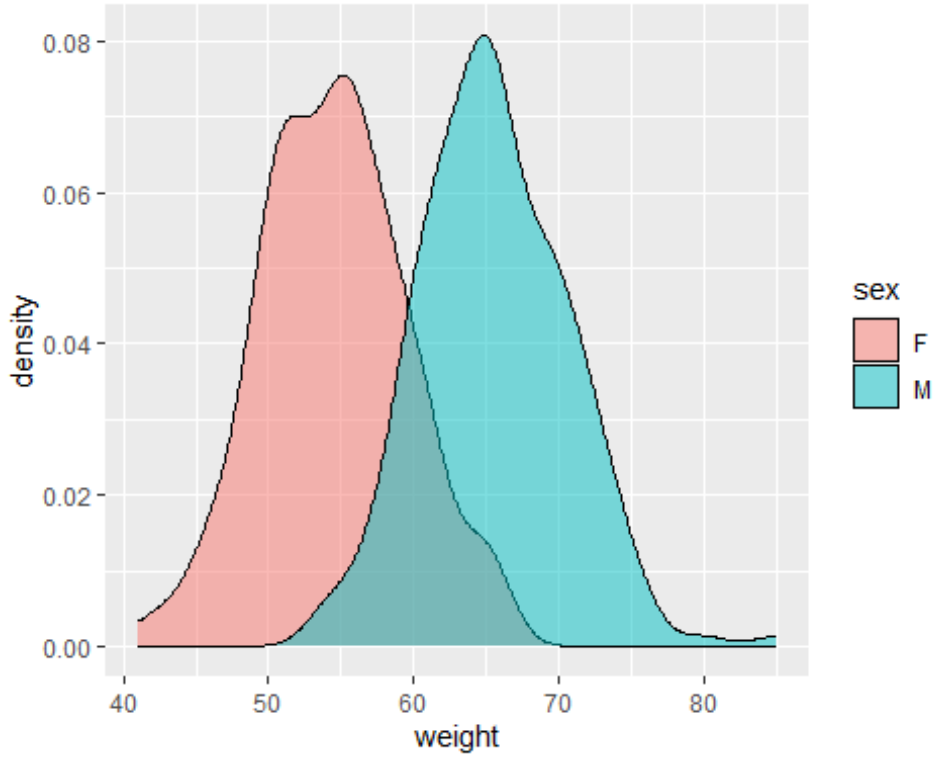

```
# Change density plot line colors by groups
ggplot(df, aes(x=weight, color=sex)) +
  geom_density()
```




```
# Change fill by groups
ggplot(df, aes(x=weight, fill=sex)) +
  geom_density()
```



```r
# Add transparency
ggplot(df, aes(x=weight, fill=sex)) +
  geom_density(alpha=0.5)
```



Section 4. R functions

```r
## Create user-defined functions
# name_func <- function(args) {
#   commands
#   return()
# }

# Create a function that will calculate a mean
# of 2 numbers
calcMeanTwo <- function(num1, num2) {
  return(num1 + num2 / 2)
}

calcMeanTwo(num1 = 2, num2 = 5)

## [1] 4.5

MeanAndSum <- function(num1, num2) {
  mean <- num1 + num2 / 2
  sum <- num1 + num2
  out_list <- list(Mean=mean, Sum=sum)
  return(out_list)
}
```

```
numList <- MeanAndSum(num1 = 10, num2 = 4)
numList
```

```
## $Mean
## [1] 12
##
## $Sum
## [1] 14
```

Learn R with swirl

Install package “swirl” and follow

the instructions

Run R script command line

Rscript script.R

R CMD BATCH script.R

Will create the file with output

cat script.Rout

Add shebang #!/usr/bin/env Rscript

to be able to run the script as ./script.R