anova factorial design in R

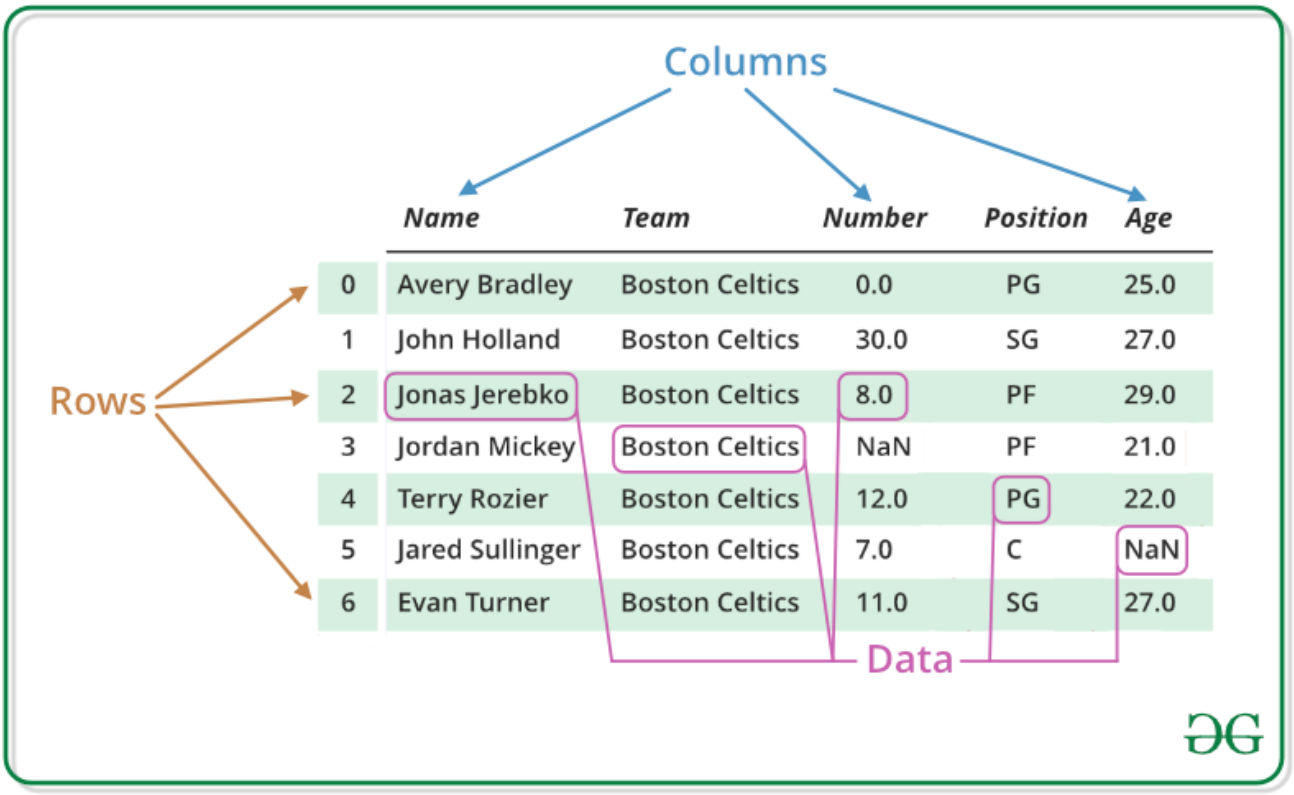
R

data frames in R

Data Frames are data displayed in a format as a table.

- character
- numeric
- logical
- `data.frame()`

make data frame




```
# R program to create dataframe
# creating a data frame
```

```
friend.data <- data.frame(
        friend_id = c(1:5),
        friend_name = c("Sachin", "Sourav",
                                        "Dravid",
"Sehwag",
                                        "Dhoni"),
        stringsAsFactors = FALSE
)
# print the data frame
print(friend.data)
#-------------------#
#     read csv      #
#-------------------#
df<-read.csv("G://R CLASS
practice/box/Boxplot_data.csv")
# it one two // two mean bacslash
```

Add Rows and Columns

```
Data_Frame <- data.frame (
  Training = c("Strength", "Stamina", "Other"),
  Pulse = c(100, 150, 120),
  Duration = c(60, 30, 45)
)

# Add a new row
New_row_DF <- rbind(Data_Frame, c("Strength", 110, 110))

# Print the new row
New_row_DF

#-------------------#
#    add new col    #
#-------------------#
```

```r
Data_Frame <- data.frame (
  Training = c("Strength", "Stamina", "Other"),
  Pulse = c(100, 150, 120),
  Duration = c(60, 30, 45)
)

# Add a new column
New_col_DF <- cbind(Data_Frame, Steps = c(1000, 6000,
2000))

# Print the new column
New_col_DF
```

working with data frame

```r
# make directory
Data_Frame <- data.frame (
  Training = c("Strength", "Stamina", "Other"),
  Pulse = c(100, 150, 120),
  Duration = c(60, 30, 45)
)

# Remove the first row and column
Data_Frame_New <- Data_Frame[-c(1), -c(1)]

# Print the new data frame
Data_Frame_New

#-------------------#
#      Remove       #
#-------------------#
library(dplyr)
```

```r
# Create a data frame
data <- data.frame(
friend_id = c(1, 2, 3, 4, 5),
friend_name = c("Sachin", "Sourav", "Dravid", "Sehwag",
"Dhoni"),
location = c("Kolkata", "Delhi", "Bangalore",
"Hyderabad", "Chennai")
)

data

# Remove a row with friend_id = 3
data <- subset(data, friend_id != 3)

data
#--------------------#
#    data frames     #
#--------------------#

library(dplyr)
# Create a data frame
data <- data.frame(
friend_id = c(1, 2, 3, 4, 5),
friend_name = c("Sachin", "Sourav", "Dravid", "Sehwag",
"Dhoni"),
location = c("Kolkata", "Delhi", "Bangalore",
"Hyderabad", "Chennai")
)
data

# Remove the 'location' column
data <- select(data, -location)

data
```

```r
# R program to extract
# data from the data frame

# creating a data frame
friend.data <- data.frame(
        friend_id = c(1:5),
        friend_name = c("Sachin", "Sourav",
                                        "Dravid",
"Sehwag",
                                        "Dhoni"),
        stringsAsFactors = FALSE
)

# Extracting friend_name column
result <- data.frame(friend.data$friend_name)
print(result)

#move of part of our data #p
datap<- data3[c(1:5,1:5),]
#this one repead 5 data two time
datap2<-data3[1,-3] #this one remove 3 col
datap3<-data3[1:10,] #like head 10 data 3
datap4<-data3[-1:-95,-3] #like tail 10 data 3


# R program to expand
# the data frame

# creating a data frame
friend.data <- data.frame(
        friend_id = c(1:5),
```

```
        friend_name = c("Sachin", "Sourav",
                                        "Dravid",
"Sehwag",
                                        "Dhoni"),
        stringsAsFactors = FALSE
)

# Expanding data frame
friend.data$location <- c("Kolkata", "Delhi",
                                        "Bangalore",
"Hyderabad",
                                        "Chennai")
resultant <- friend.data
# print the modified data frame
print(resultant)
```

```
  friend_id friend_name  location
1         1      Sachin   Kolkata
2         2      Sourav     Delhi
3         3      Dravid Bangalore
4         4      Sehwag Hyderabad
5         5       Dhoni   Chennai
```

```
# creating a data frame
friend.data <- data.frame(
friend_id = c(1:5),
friend_name = c("Sachin", "Sourav",
                                "Dravid", "Sehwag",
                                "Dhoni"),
stringsAsFactors = FALSE
)
```

```r
# Access Items using []
friend.data[1]

# Access Items using [[]]
friend.data[['friend_name']]

# Access Items using $
friend.data$friend_id
```

```r
Data_Frame <- data.frame (
  Training = c("Strength", "Stamina", "Other"),
  Pulse = c(100, 150, 120),
  Duration = c(60, 30, 45)
)
head(Data_Frame)
tail(Data_Frame)
dim(Data_Frame)
ncol(Data_Frame)
nrow(Data_Frame)
str(Data_Frame)
summary(Data_Frame)
#cbind
#rbind
combined_df <-rbind(df1, df2)
```

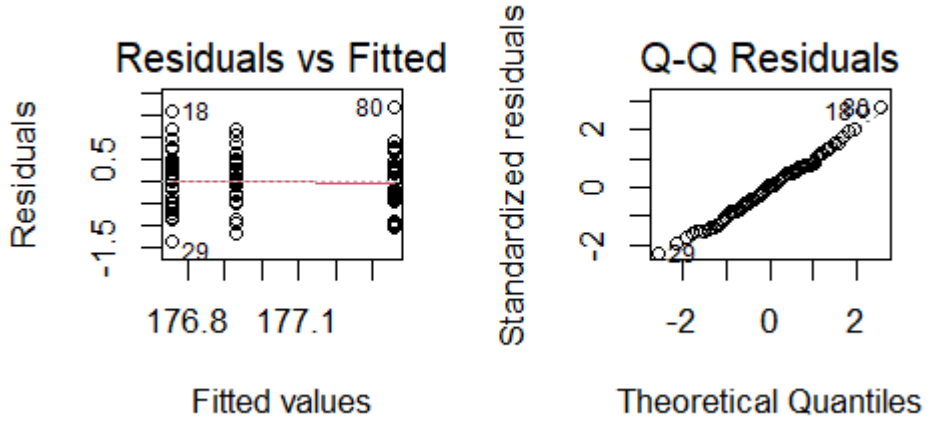
statistical

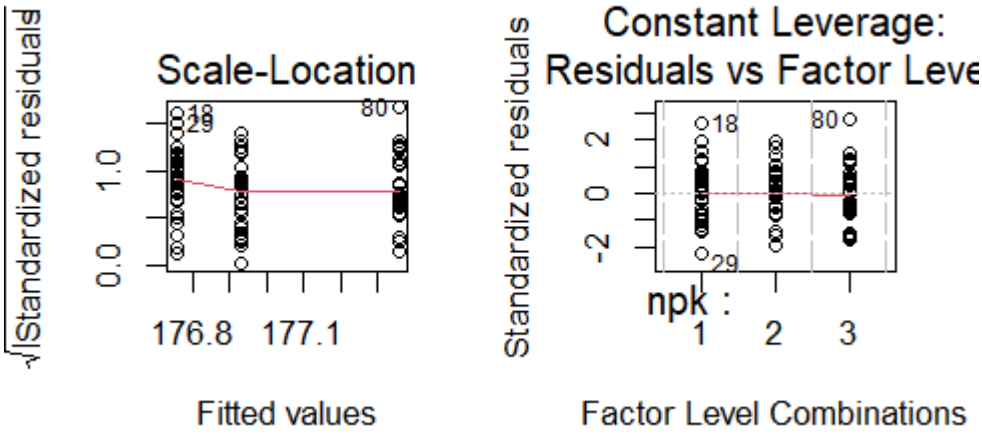
ANOVA

ANOVA is a statistical test for estimating how a quantitative dependent variable changes according to the levels of one or more categorical independent variables. ANOVA tests whether there is a difference in means of the groups at each level of the independent variable.

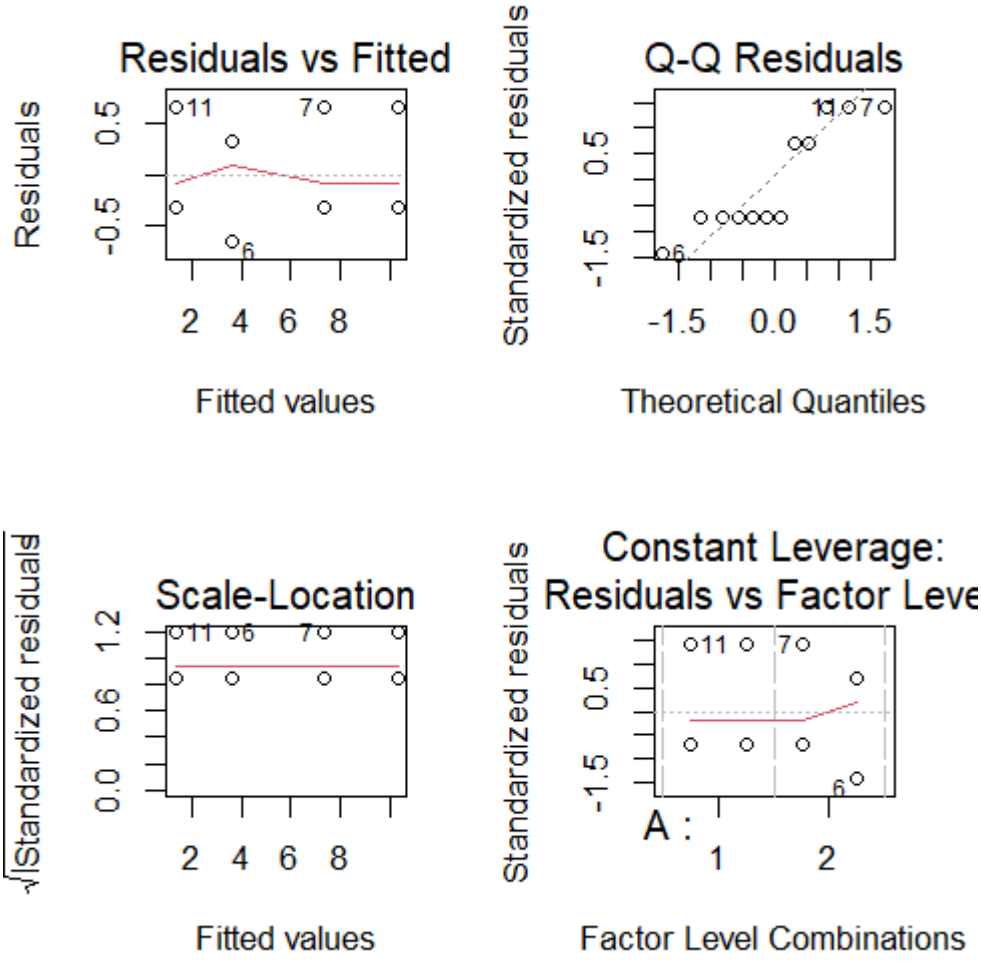
The null hypothesis (_H_0) of the ANOVA is no difference in means, and the alternative hypothesis (_H_a) is that the means are different from one another.

```
library(ggplot2)
library(ggpubr)
library(tidyverse)
library(broom)
library(AICcmodavg)
```

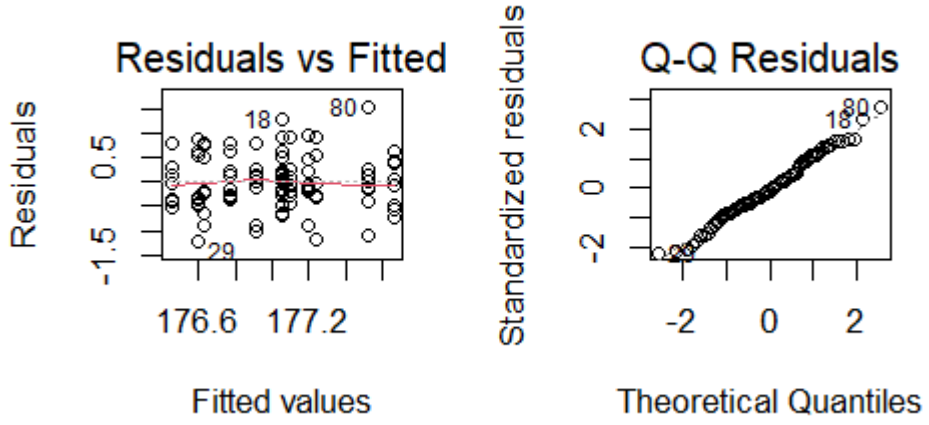

Residuals vs Fitted
Residuals
18
80
29
-1.5
0.5
176.8
177.1
Fitted values
Q-Q Residuals
Standardized residuals
18
80
29
-2
0
2
-2
0
2
Theoretical Quantiles

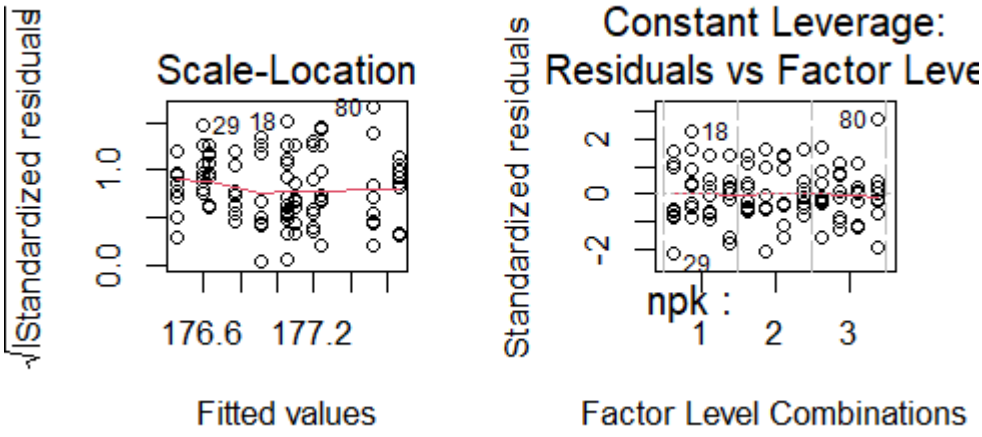

Scale-Location
√|Standardized residuals|
18
29
80
0.0
1.0
176.8
177.1
Fitted values
Constant Leverage:
Residuals vs Factor Leve
Standardized residuals
18
80
29
-2
0
2
npk :
1
2
3
Factor Level Combinations

CRD


Residuals vs Fitted
Residuals
Fitted values
Q-Q Residuals
Standardized residuals
Theoretical Quantiles
Scale-Location
√|Standardized residuals|
Fitted values
Constant Leverage:
Residuals vs Factor Leve
Standardized residuals
A :
Factor Level Combinations

CRD.fact





RBD

best model

```
Model selection based on AICc:

            K   AICc Delta_AICc AICcWt Cum.Wt     LL
two.way     5 173.86       0.00   0.71   0.71 -81.59
blocking    7 176.93       3.08   0.15   0.86 -80.83
interaction 7 177.12       3.26   0.14   1.00 -80.92
one.way     4 186.41      12.56   0.00   1.00 -88.99
```

https://www.scribbr.com/statistics/anova-in-r/