20BCE1025_Abhishek_N_N_Ex-4(b) Basic Operations in data frame

20BCE1025_Abhishek_N_N

2022-09-12

1. Install the package MASS.

```
install.packages("MASS")
```

```
## Installing package into '/home/abhishek_n_n_20bce1025/R/x86_64-pc-linux-gnu-library/4.1'
## (as 'lib' is unspecified)
```

2. Import the package MASS.

```
library("MASS")
```

3. Display the structure of the data survey.

```
str(survey)
```

```
## 'data.frame':    237 obs. of  12 variables:
##  $ Sex   : Factor w/ 2 levels "Female","Male": 1 2 2 2 2 1 2 1 2 2 ...
##  $ Wr.Hnd: num  18.5 19.5 18 18.8 20 18 17.7 17 20 18.5 ...
##  $ NW.Hnd: num  18 20.5 13.3 18.9 20 17.7 17.7 17.3 19.5 18.5 ...
##  $ W.Hnd : Factor w/ 2 levels "Left","Right": 2 1 2 2 2 2 2 2 2 2 ...
##  $ Fold  : Factor w/ 3 levels "L on R","Neither",..: 3 3 1 3 2 1 1 3 3 3 ...
##  $ Pulse : int  92 104 87 NA 35 64 83 74 72 90 ...
##  $ Clap  : Factor w/ 3 levels "Left","Neither",..: 1 1 2 2 3 3 3 3 3 3 ...
##  $ Exer  : Factor w/ 3 levels "Freq","None",..: 3 2 2 2 3 3 1 1 3 3 ...
##  $ Smoke : Factor w/ 4 levels "Heavy","Never",..: 2 4 3 2 2 2 2 2 2 2 ...
##  $ Height: num  173 178 NA 160 165 ...
##  $ M.I   : Factor w/ 2 levels "Imperial","Metric": 2 1 NA 2 2 1 1 2 2 2 ...
##  $ Age   : num  18.2 17.6 16.9 20.3 23.7 ...
```

4. Check the class and type of the data set survey in MASS.

```
class(survey)
```

```
## [1] "data.frame"
```

```
typeof(survey)
```

```
## [1] "list"
```

5. Get the number of rows and columns of the survey data frame.

```
nrow(survey)
```

```
## [1] 237
```

```
ncol(survey)
```

```
## [1] 12
```

6. Get the dimension of the survey data frame.

```
dim(survey)
```

```
## [1] 237  12
```

7. Provide the statistical summary of the data frame.

```
summary(survey)
```

```
##      Sex          Wr.Hnd          NW.Hnd        W.Hnd          Fold
##  Female:118   Min.   :13.00   Min.   :12.50   Left : 18   L on R : 99
##  Male  :118   1st Qu.:17.50   1st Qu.:17.50   Right:218   Neither: 18
##  NA's  :  1   Median :18.50   Median :18.50   NA's :  1   R on L :120
##               Mean   :18.67   Mean   :18.58
##               3rd Qu.:19.80   3rd Qu.:19.73
##               Max.   :23.20   Max.   :23.50
##               NA's   :1       NA's   :1
##      Pulse             Clap       Exer       Smoke         Height
##  Min.   : 35.00   Left   : 39   Freq:115   Heavy: 11   Min.   :150.0
##  1st Qu.: 66.00   Neither: 50   None: 24   Never:189   1st Qu.:165.0
##  Median : 72.50   Right  :147   Some: 98   Occas: 19   Median :171.0
##  Mean   : 74.15   NA's   :  1              Regul: 17   Mean   :172.4
##  3rd Qu.: 80.00                            NA's :  1   3rd Qu.:180.0
##  Max.   :104.00                                        Max.   :200.0
##  NA's   :45                                            NA's   :28
##        M.I           Age
##  Imperial: 68   Min.   :16.75
##  Metric  :141   1st Qu.:17.67
##  NA's    : 28   Median :18.58
##                 Mean   :20.37
##                 3rd Qu.:20.17
##                 Max.   :73.00
##
```

8. Display the column names of the survey data frame

```
names(survey)
```

```
##  [1] "Sex"    "Wr.Hnd" "NW.Hnd" "W.Hnd"  "Fold"   "Pulse"  "Clap"   "Exer"
##  [9] "Smoke"  "Height" "M.I"    "Age"
```

9. Retrieve the top 3 rows from the data frame.

```
head(survey,3)
```

```
##      Sex Wr.Hnd NW.Hnd W.Hnd   Fold Pulse    Clap Exer Smoke Height      M.I
## 1 Female   18.5   18.0 Right R on L    92    Left Some Never  173.0   Metric
## 2   Male   19.5   20.5  Left R on L   104    Left None Regul  177.8 Imperial
## 3   Male   18.0   13.3 Right L on R    87 Neither None Occas     NA     <NA>
##      Age
## 1 18.250
## 2 17.583
## 3 16.917
```

10. Extract the bottom 2 rows from the data frame.

```
tail(survey,2)
```

```
##        Sex Wr.Hnd NW.Hnd W.Hnd   Fold Pulse  Clap Exer Smoke Height    M.I
## 236   Male   21.0   21.5 Right R on L    90 Right Some Never  183.0 Metric
## 237 Female   17.6   17.3 Right R on L    85 Right Freq Never  168.5 Metric
##        Age
## 236 17.167
## 237 17.750
```