**Lab 2C**

**R.Harini**

**18BCE1010**

---

title: "LabEx2-Data frame Basics & Indexing"

output:

pdf\_document: default

html\_document:

fig\_height: 4

highlight: pygments

theme: spacelab

---

### Reg. No: <your reg.no.here>

18BCE1010

### Name:<your name here>

R. Harini

## Setup

```{r}

#install.packages("rmarkdown")

```

### Load packages if any using library(packagename)

```{r}

library(MASS)

```

### Load data

```{r}

rm(list=ls())

df=na.omit(survey)

df

```

\* \* \*

**## Part 1: Data (Chosen Dataset, Description of dataset, basic commands to describe dataset)**

```{r}

#Using the Survey dataset from MASS package.

#Description: Responses of Students at a univeristy to a number of questions. The components are Sex of the student, Writing Hand Span, Non-writing Hand Span, Writing hand of the student, Fold, Pulse Rate, Clap, Exercise, Smoke, Height of the student in cm and Age of the student.

str(df)

class(df)

typeof(df)

nrow(df)

ncol(df)

dim(df)

summary(df)

colnames(df)

head(df)

```

**Output:**

Male :84 1st Qu.:17.5 1st Qu.:17.50 Right:156

Median :18.5 Median :18.50

Mean :18.8 Mean :18.73

3rd Qu.:20.0 3rd Qu.:20.00

Max. :23.2 Max. :23.50

Fold Pulse Clap Exer

L on R :72 Min. : 35.00 Left : 28 Freq:85

Neither: 8 1st Qu.: 66.75 Neither: 33 None:14

R on L :88 Median : 72.00 Right :107 Some:69

Mean : 74.02

3rd Qu.: 80.00

Max. :104.00

Smoke Height M.I

Heavy: 7 Min. :152.0 Imperial: 58

Never:134 1st Qu.:165.0 Metric :110

Occas: 13 Median :170.6

Regul: 14 Mean :172.5

3rd Qu.:180.0

Max. :200.0

Age

Min. :16.92

1st Qu.:17.67

Median :18.58

Mean :20.43

3rd Qu.:20.17

Max. :70.42

[1] "Sex" "Wr.Hnd" "NW.Hnd" "W.Hnd" "Fold" "Pulse"

[7] "Clap" "Exer" "Smoke" "Height" "M.I" "Age"

\* \* \*

**## Part 2: Selected Questions and Corresponding Code & Output**

```{r}

#1. Calculating the avg of all rows for Wr.Hnd Span and NW.Hnd

colMeans(df[,2:3])

```

```{r}

**Output:**

Wr.Hnd NW.Hnd

18.80238 18.73155

#2. Finding the min and max of Wr. Hnd Span

min(df[,2])

max(df[,2])

```

**Output:**

[1] 13

[1] 23.2

```{r}

#3. Finding the min and max of NW. Hnd Span

min(df[,3])

max(df[,3])

```

**Output:**

[1] 12.5

[1] 23.5

```{r}

#4. Finding the average pulse rate of students

colMeans(df[6])

```

**Output:**

Pulse

74.02381

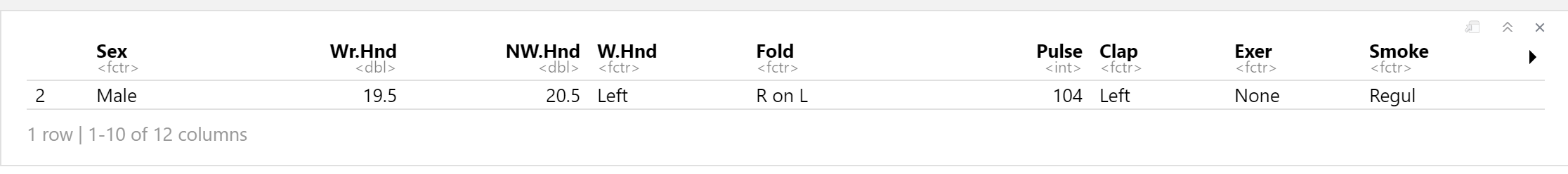
```{r}

#5. Display the Student row with highest pulse rate

df[which.max(df$Pulse),]

```

**Output:**



```{r}

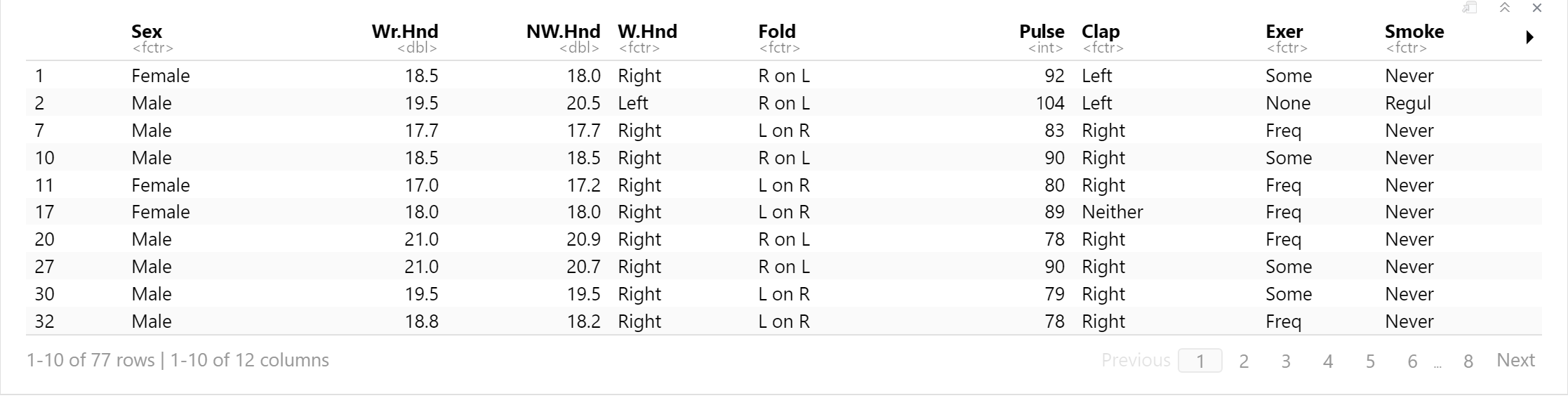
#6. Display all Student rows with pulse rate higher than the average

mean(df$Pulse) #74.02381

df[which(df$Pulse>mean(df$Pulse)),]

```

**Output:**



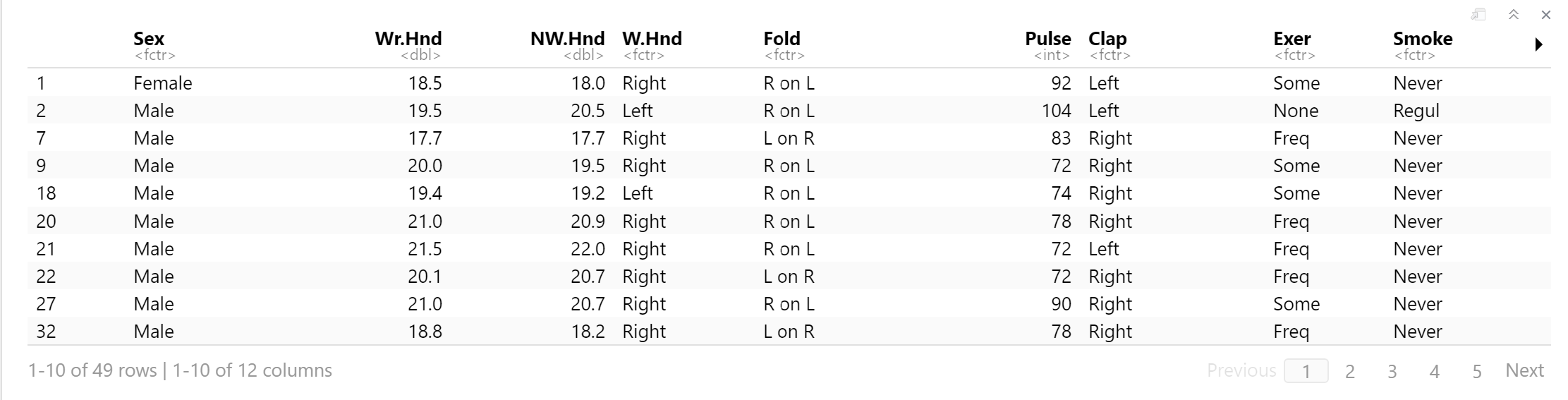
```{r}

#7. Display all student rows with height above 170cm and pulse above 70

df[which((df$Height>170) & (df$Pulse>70)),]

```

**Output:**



```{r}

#8. Display categories in Smoke column

unique(df$Smoke)

```

**Output:**

[1] Never Regul Heavy Occas

Levels: Heavy Never Occas Regul

```{r}

#9. Finding the average of Wr.Hnd span and NW.Hnd span

avg=c(rowMeans(df[,2:3]))

avg

```

**Output:**

1 2 5 6 7 8 9 10 11

18.25 20.00 20.00 17.85 17.70 17.15 19.75 18.50 17.10

14 17 18 20 21 22 23 24 27

19.85 18.00 19.30 20.95 21.75 20.40 18.25 21.35 20.85

28 30 32 33 34 36 38 39 42

21.10 19.50 18.50 17.30 20.05 21.60 18.95 22.00 17.90

44 47 48 49 50 51 52 53 54

20.15 22.95 22.75 17.80 17.95 21.75 20.25 17.50 20.00

55 57 59 61 62 63 65 71 73

22.50 15.45 19.25 23.00 18.35 19.65 17.65 17.75 16.80

74 75 76 77 79 82 85 86 87

16.75 15.70 17.50 17.30 18.40 19.05 23.25 17.35 18.10

88 89 91 93 95 97 98 100 102

18.40 18.00 20.25 17.85 21.05 19.75 17.50 19.50 19.00

104 105 106 109 110 111 112 113 114

17.40 17.25 19.00 17.50 19.90 18.50 19.40 16.95 20.75

115 116 117 118 119 120 122 123 124

15.75 16.45 16.85 22.50 18.25 20.70 22.50 18.25 19.90

125 127 128 129 130 131 132 134 135

18.30 16.00 18.95 17.25 16.45 21.75 19.25 15.90 17.85

136 138 140 141 143 144 145 146 147

22.80 22.00 19.00 18.30 18.90 21.20 19.75 18.50 22.55

148 149 150 151 152 153 154 155 156

19.85 18.00 18.25 22.05 12.75 16.25 21.55 19.00 20.25

158 160 161 163 164 166 167 168 170

19.05 18.75 17.30 20.25 16.70 17.40 19.35 15.75 18.75

172 174 175 176 177 178 180 181 182

20.00 17.75 17.75 18.75 20.50 16.85 16.75 19.25 13.75

183 184 185 186 187 188 189 190 191

17.55 18.75 18.25 20.60 17.00 18.50 18.25 18.25 19.75

```{r}

#10. Adding the avg column to the dataset

cbind(df,avg)

```

**Output:**



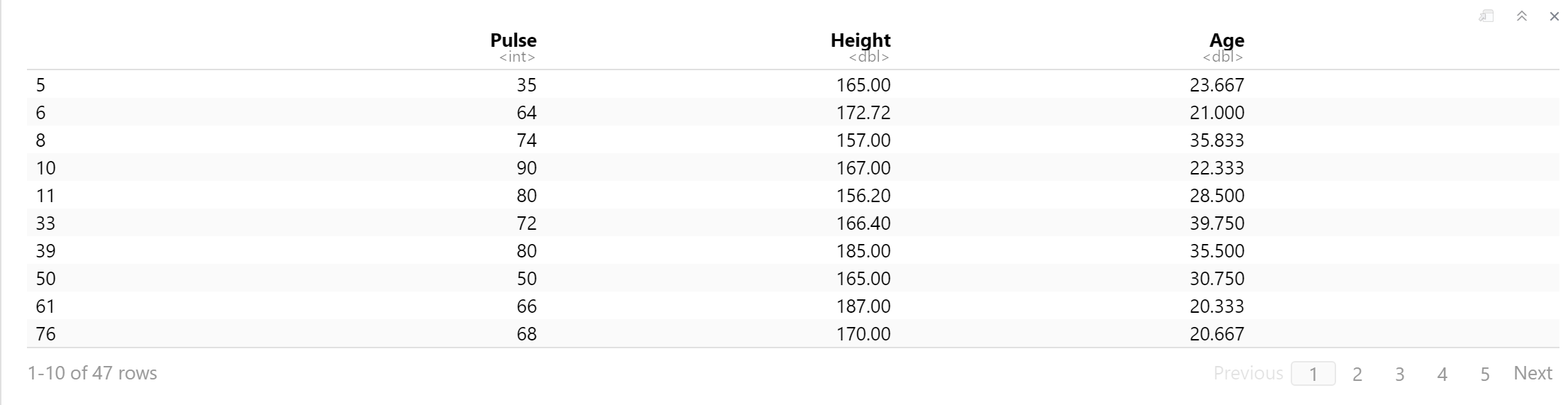
```{r}

#11. Display Pulse rate and height of Students above 20 years old

df[which(df$Age>20), c(6,10,12)]

```

**Output:**



```{r}

#12. Display Wr.Hnd Span of Right handed Students

df[which(df$W.Hnd=="Right"), 2]

```

**Output:**

[1] 18.5 20.0 18.0 17.7 17.0 20.0 18.5 17.0 19.5 18.0

[11] 21.0 21.5 20.1 18.5 21.5 21.0 20.8 19.5 18.8 17.1

[21] 20.1 22.2 19.4 22.0 17.8 20.1 23.2 22.5 18.0 18.0

[31] 20.5 17.0 20.5 22.5 15.5 19.5 22.8 18.5 19.6 17.3

[41] 18.0 17.0 16.5 15.6 17.5 17.0 18.3 19.2 23.0 17.7

[51] 18.2 18.3 18.0 20.5 18.2 21.3 20.0 17.5 19.4 18.9

[61] 17.5 17.5 19.5 17.5 19.7 18.5 19.2 17.2 20.5 16.0

[71] 16.9 17.0 21.0 22.5 18.5 18.5 16.0 18.8 17.5 16.4

[81] 22.0 19.0 17.9 23.1 22.0 19.5 18.0 19.0 21.4 18.5

[91] 22.5 19.5 18.0 18.0 21.8 13.0 16.3 21.5 18.9 20.5

[101] 18.9 18.5 17.5 20.2 16.5 17.6 19.5 16.5 19.0 18.0

[111] 17.5 20.5 16.7 17.0 19.0 14.0 17.5 18.5 18.0 20.5

[121] 17.0 18.5 18.0 18.5 20.0 22.0 17.9 17.6 17.0 15.0

[131] 16.0 19.1 17.5 16.2 21.0 18.5 17.0 17.5 17.5 17.5

[141] 18.6 17.0 18.0 18.2 23.2 15.9 17.5 18.8 20.0 18.6

[151] 18.6 18.8 18.0 18.5 21.0 17.6

```{r}

#13. Display the average of Wr.Hnd Span of all Left handed Students

mean(df[which(df$W.Hnd=="Left"), 2])

```

**Output:**

[1] 19.34167

\* \* \*