

CSE3505 – LAB

20BPS1022 PREYASH _ 07 Oct 2022

Name: Preyash

Date: 07th October, 2022

Reg No. 20BPS1022

```
library(missForest)
## Warning: package 'missForest' was built under R version 4.1.3
library(mice)
## Warning: package 'mice' was built under R version 4.1.3
##
## Attaching package: 'mice'
## The following object is masked from 'package:stats':
##
##   filter
## The following objects are masked from 'package:base':
##
##   cbind, rbind
library(MASS)
## Warning: package 'MASS' was built under R version 4.1.3
library("dplyr")
## Warning: package 'dplyr' was built under R version 4.1.3
##
## Attaching package: 'dplyr'
## The following object is masked from 'package:MASS':
##
##   select
## The following objects are masked from 'package:stats':
##
##   filter, lag
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

PART-01

1. Consider the iris dataset, artificially induce some percentage of missing data, use mice package and associated methods to replace missing values and generate a complete dataset. Show the step by step process.

#Q1

#installing missForest package for using prodNA function

#function to induce missing values to a dataset

```
x=head(iris,n=10)
```

```
y=prodNA(x,noNA=0.15)
```

y

##	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
## 1	5.1	3.5	NA	0.2	setosa
## 2	NA	NA	1.4	0.2	setosa
## 3	4.7	3.2	1.3	0.2	setosa
## 4	NA	3.1	NA	0.2	setosa
## 5	5.0	3.6	1.4	0.2	setosa
## 6	5.4	3.9	1.7	0.4	setosa
## 7	4.6	3.4	NA	0.3	setosa
## 8	NA	3.4	1.5	0.2	setosa
## 9	4.4	2.9	1.4	0.2	setosa
## 10	4.9	3.1	1.5	0.1	setosa

#removing categorical data

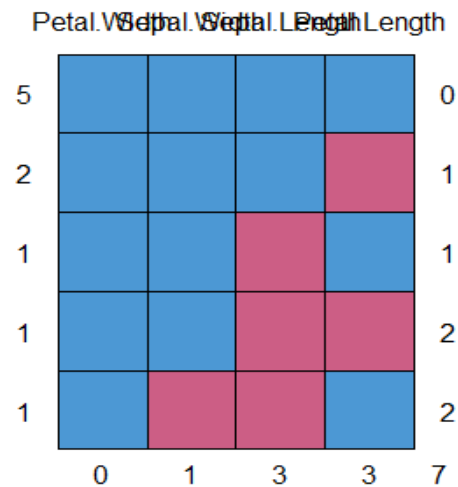
```
y=y[,1:4]
```

y

##	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
## 1	5.1	3.5	NA	0.2
## 2	NA	NA	1.4	0.2
## 3	4.7	3.2	1.3	0.2
## 4	NA	3.1	NA	0.2
## 5	5.0	3.6	1.4	0.2
## 6	5.4	3.9	1.7	0.4
## 7	4.6	3.4	NA	0.3
## 8	NA	3.4	1.5	0.2
## 9	4.4	2.9	1.4	0.2
## 10	4.9	3.1	1.5	0.1

#using mice package now

```
md.pattern(y)
```



```
## Petal.Width Sepal.Width Sepal.Length Petal.Length
## 5          1          1          1          1 0
## 2          1          1          1          0 1
## 1          1          1          0          1 1
## 1          1          1          0          0 2
## 1          1          0          0          1 2
##           0          1          3          3 7
```

#imputing data

```
imputed_data=mice(y, m=5, maxit = 50, method = 'pmm', seed = 500)
```

```
##
## iter imp variable
## 1 1 Sepal.Length Sepal.Width Petal.Length
## 1 2 Sepal.Length Sepal.Width Petal.Length
## 1 3 Sepal.Length Sepal.Width Petal.Length
## 1 4 Sepal.Length Sepal.Width Petal.Length
## 1 5 Sepal.Length Sepal.Width Petal.Length
## 2 1 Sepal.Length Sepal.Width Petal.Length
## 2 2 Sepal.Length Sepal.Width Petal.Length
## 2 3 Sepal.Length Sepal.Width Petal.Length
## 2 4 Sepal.Length Sepal.Width Petal.Length
## 2 5 Sepal.Length Sepal.Width Petal.Length
## 3 1 Sepal.Length Sepal.Width Petal.Length
## 3 2 Sepal.Length Sepal.Width Petal.Length
## 3 3 Sepal.Length Sepal.Width Petal.Length
## 3 4 Sepal.Length Sepal.Width Petal.Length
## 3 5 Sepal.Length Sepal.Width Petal.Length
## 4 1 Sepal.Length Sepal.Width Petal.Length
## 4 2 Sepal.Length Sepal.Width Petal.Length
## 4 3 Sepal.Length Sepal.Width Petal.Length
## 4 4 Sepal.Length Sepal.Width Petal.Length
```

[illegible]

[illegible]

[illegible]

[illegible]

```
## 44 5 Sepal.Length Sepal.Width Petal.Length
## 45 1 Sepal.Length Sepal.Width Petal.Length
## 45 2 Sepal.Length Sepal.Width Petal.Length
## 45 3 Sepal.Length Sepal.Width Petal.Length
## 45 4 Sepal.Length Sepal.Width Petal.Length
## 45 5 Sepal.Length Sepal.Width Petal.Length
## 46 1 Sepal.Length Sepal.Width Petal.Length
## 46 2 Sepal.Length Sepal.Width Petal.Length
## 46 3 Sepal.Length Sepal.Width Petal.Length
## 46 4 Sepal.Length Sepal.Width Petal.Length
## 46 5 Sepal.Length Sepal.Width Petal.Length
## 47 1 Sepal.Length Sepal.Width Petal.Length
## 47 2 Sepal.Length Sepal.Width Petal.Length
## 47 3 Sepal.Length Sepal.Width Petal.Length
## 47 4 Sepal.Length Sepal.Width Petal.Length
## 47 5 Sepal.Length Sepal.Width Petal.Length
## 48 1 Sepal.Length Sepal.Width Petal.Length
## 48 2 Sepal.Length Sepal.Width Petal.Length
## 48 3 Sepal.Length Sepal.Width Petal.Length
## 48 4 Sepal.Length Sepal.Width Petal.Length
## 48 5 Sepal.Length Sepal.Width Petal.Length
## 49 1 Sepal.Length Sepal.Width Petal.Length
## 49 2 Sepal.Length Sepal.Width Petal.Length
## 49 3 Sepal.Length Sepal.Width Petal.Length
## 49 4 Sepal.Length Sepal.Width Petal.Length
## 49 5 Sepal.Length Sepal.Width Petal.Length
## 50 1 Sepal.Length Sepal.Width Petal.Length
## 50 2 Sepal.Length Sepal.Width Petal.Length
## 50 3 Sepal.Length Sepal.Width Petal.Length
## 50 4 Sepal.Length Sepal.Width Petal.Length
## 50 5 Sepal.Length Sepal.Width Petal.Length
```

```
completedata=complete(imputed_data,2)
completedata
```

```
## Sepal.Length Sepal.Width Petal.Length Petal.Width
## 1 5.1 3.5 1.4 0.2
## 2 4.6 3.1 1.4 0.2
## 3 4.7 3.2 1.3 0.2
## 4 4.7 3.1 1.3 0.2
## 5 5.0 3.6 1.4 0.2
## 6 5.4 3.9 1.7 0.4
## 7 4.6 3.4 1.4 0.3
## 8 5.0 3.4 1.5 0.2
## 9 4.4 2.9 1.4 0.2
## 10 4.9 3.1 1.5 0.1
```


2. Consider data set of your choice, Using R try to find the correlation coefficient between any two variables using the correlation methods.

```
#Q2
#using first 10 rows of iris dataset
x=x[,1:4]
cor(x$Sepal.Length,x$Petal.Length)

## [1] 0.600216

cor(x)
```

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
Sepal.Length	1.0000000	0.7872066	0.6002160	0.3770977
Sepal.Width	0.7872066	1.0000000	0.5191385	0.6787563
Petal.Length	0.6002160	0.5191385	1.0000000	0.5216405
Petal.Width	0.3770977	0.6787563	0.5216405	1.0000000

PART-02

Part-2

Questions

Statistical Analysis of qualitative data

Use the survey data in MASS package to do the following:

1. Import the package MASS.
2. List the rows of data that has missing values.
3. Create a data frame 'newsurvey' that contains the survey data after removing the na values. Use it for answering further queries
4. How many male and female students participated in the survey?
5. How many the left and right handers are there?
6. Find the relative frequency distribution of left and right handers and display them with the precision of two decimal places.
7. Display the male left hander and female left hander in the column format.
8. What percentage of male left handers never smokes?

#1

```
library(MASS)
```

#2

```
dim(survey)
```

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

```
which(!complete.cases(survey))
```

```
## [1] 3 4 12 13 15 16 19 25 26 29 31 35 37 40 41 43 45  
46 56
```

```
## [20] 58 60 64 66 67 68 69 70 72 78 80 81 83 84 90 92 94  
96 99
```

```
## [39] 101 103 107 108 121 126 133 137 139 142 157 159 162 165 169 171 173 1  
79 195
```

```
## [58] 203 210 213 216 217 219 221 224 225 226 232 235
```

#3

```
good=complete.cases(survey)
```

```
new_survey=survey[good,]
```

```
new_survey
```

```
##           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 173.00  M
```

etric ## 2 erial	Male	19.5	20.5	Left	R on L	104	Left	None	Regul	177.80	Imp
etric ## 5 erial	Male	20.0	20.0	Right	Neither	35	Right	Some	Never	165.00	M
etric ## 6 erial	Female	18.0	17.7	Right	L on R	64	Right	Some	Never	172.72	Imp
etric ## 7 erial	Male	17.7	17.7	Right	L on R	83	Right	Freq	Never	182.88	Imp
etric ## 8 erial	Female	17.0	17.3	Right	R on L	74	Right	Freq	Never	157.00	M
etric ## 9 erial	Male	20.0	19.5	Right	R on L	72	Right	Some	Never	175.00	M
etric ## 10 erial	Male	18.5	18.5	Right	R on L	90	Right	Some	Never	167.00	M
etric ## 11 erial	Female	17.0	17.2	Right	L on R	80	Right	Freq	Never	156.20	Imp
etric ## 14 erial	Female	19.5	20.2	Right	L on R	66	Neither	Some	Never	155.00	M
etric ## 17 erial	Female	18.0	18.0	Right	L on R	89	Neither	Freq	Never	157.00	M
etric ## 18 erial	Male	19.4	19.2	Left	R on L	74	Right	Some	Never	182.88	Imp
etric ## 20 erial	Male	21.0	20.9	Right	R on L	78	Right	Freq	Never	177.00	M
etric ## 21 erial	Male	21.5	22.0	Right	R on L	72	Left	Freq	Never	190.50	Imp
etric ## 22 erial	Male	20.1	20.7	Right	L on R	72	Right	Freq	Never	180.34	Imp
etric ## 23 erial	Male	18.5	18.0	Right	L on R	64	Right	Freq	Never	180.34	Imp
etric ## 24 erial	Male	21.5	21.2	Right	R on L	62	Right	Some	Never	184.00	M
etric ## 27 erial	Male	21.0	20.7	Right	R on L	90	Right	Some	Never	172.72	Imp
etric ## 28 erial	Male	20.8	21.4	Right	R on L	62	Neither	Freq	Never	175.26	Imp
etric ## 30 erial	Male	19.5	19.5	Right	L on R	79	Right	Some	Never	167.00	M
etric ## 32 erial	Male	18.8	18.2	Right	L on R	78	Right	Freq	Never	180.00	M
etric ## 33 erial	Female	17.1	17.5	Right	R on L	72	Right	Freq	Heavy	166.40	Imp
etric ## 34 erial	Male	20.1	20.0	Right	R on L	70	Right	Some	Never	180.00	M
etric ## 36 erial	Male	22.2	21.0	Right	L on R	66	Right	Freq	Occas	190.00	M
etric ## 38 erial	Male	19.4	18.5	Right	R on L	72	Neither	Freq	Never	182.50	M
etric ## 39 erial	Male	22.0	22.0	Right	R on L	80	Right	Some	Never	185.00	M

etric											
## 42	Female	17.8	18.0	Right	R on L	72	Right	Some	Never	154.94	Imp
erial											
## 44	Female	20.1	20.2	Right	L on R	80	Right	Some	Never	176.50	Imp
erial											
## 47	Male	23.2	22.7	Right	L on R	84	Left	Freq	Regul	180.00	M
etric											
## 48	Male	22.5	23.0	Right	R on L	96	Right	None	Never	170.00	M
etric											
## 49	Female	18.0	17.6	Right	R on L	60	Right	Some	Occas	168.00	M
etric											
## 50	Female	18.0	17.9	Right	R on L	50	Left	None	Never	165.00	M
etric											
## 51	Male	22.0	21.5	Left	R on L	55	Left	Freq	Never	200.00	M
etric											
## 52	Male	20.5	20.0	Right	L on R	68	Right	Freq	Never	190.00	M
etric											
## 53	Male	17.0	18.0	Right	L on R	78	Left	Some	Never	170.18	Imp
erial											
## 54	Male	20.5	19.5	Right	L on R	56	Right	Freq	Never	179.00	M
etric											
## 55	Male	22.5	22.5	Right	R on L	65	Right	Freq	Regul	182.00	M
etric											
## 57	Female	15.5	15.4	Right	R on L	70	Neither	None	Never	157.48	Imp
erial											
## 59	Male	19.5	19.0	Right	L on R	62	Right	Freq	Never	177.80	Imp
erial											
## 61	Male	22.8	23.2	Right	R on L	66	Neither	Freq	Never	187.00	M
etric											
## 62	Female	18.5	18.2	Right	R on L	72	Neither	Freq	Never	167.64	Imp
erial											
## 63	Female	19.6	19.7	Right	L on R	70	Right	Freq	Never	178.00	M
etric											
## 65	Female	17.3	18.0	Right	L on R	64	Neither	Freq	Never	164.00	M
etric											
## 71	Female	18.0	17.5	Right	L on R	64	Left	Freq	Never	170.00	M
etric											
## 73	Female	17.0	16.6	Right	R on L	68	Right	Some	Never	171.00	M
etric											
## 74	Female	16.5	17.0	Right	L on R	40	Left	Freq	Never	167.64	Imp
erial											
## 75	Female	15.6	15.8	Right	R on L	88	Left	Some	Never	165.00	M
etric											
## 76	Female	17.5	17.5	Right	Neither	68	Right	Freq	Heavy	170.00	M
etric											
## 77	Female	17.0	17.6	Right	L on R	76	Right	Some	Never	165.00	M
etric											
## 79	Female	18.3	18.5	Right	R on L	68	Neither	Some	Never	165.10	Imp
erial											
## 82	Male	19.2	18.9	Right	R on L	76	Right	Freq	Never	176.50	Imp

erial											
## 85	Male	23.0	23.5	Right	L on R	90	Right	Freq	Never	167.64	Imp
erial											
## 86	Female	17.7	17.0	Right	R on L	76	Right	Some	Never	167.00	M
etric											
## 87	Female	18.2	18.0	Right	L on R	70	Right	Some	Never	162.56	Imp
erial											
## 88	Female	18.3	18.5	Right	R on L	75	Left	Freq	Never	170.00	M
etric											
## 89	Male	18.0	18.0	Right	Neither	60	Right	Freq	Never	179.00	M
etric											
## 91	Male	20.5	20.0	Right	R on L	75	Left	Some	Never	183.00	M
etric											
## 93	Female	18.2	17.5	Right	L on R	70	Right	Some	Never	165.00	M
etric											
## 95	Male	21.3	20.8	Right	R on L	65	Right	Freq	Heavy	179.00	M
etric											
## 97	Male	20.0	19.5	Right	R on L	68	Neither	Freq	Regul	190.00	M
etric											
## 98	Female	17.5	17.5	Right	R on L	60	Right	Freq	Never	166.50	M
etric											
## 100	Female	19.4	19.6	Right	R on L	68	Neither	Freq	Never	175.26	Imp
erial											
## 102	Male	18.9	19.1	Right	L on R	60	Neither	None	Never	170.00	M
etric											
## 104	Female	17.5	17.3	Right	R on L	72	Right	Freq	Never	175.00	M
etric											
## 105	Female	17.5	17.0	Right	R on L	80	Left	Some	Heavy	163.00	M
etric											
## 106	Female	19.5	18.5	Right	R on L	80	Right	Some	Never	170.00	M
etric											
## 109	Male	17.5	17.5	Right	L on R	64	Neither	Freq	Never	180.00	M
etric											
## 110	Male	19.7	20.1	Right	R on L	67	Left	Some	Regul	180.34	Imp
erial											
## 111	Female	18.5	18.5	Right	R on L	76	Left	Freq	Never	175.00	M
etric											
## 112	Male	19.2	19.6	Right	L on R	80	Right	None	Never	190.50	Imp
erial											
## 113	Female	17.2	16.7	Right	R on L	75	Right	Freq	Never	170.18	Imp
erial											
## 114	Male	20.5	21.0	Right	R on L	60	Right	Freq	Never	185.00	M
etric											
## 115	Female	16.0	15.5	Right	L on R	60	Left	Freq	Never	162.56	Imp
erial											
## 116	Female	16.9	16.0	Right	L on R	70	Right	None	Never	158.00	M
etric											
## 117	Female	17.0	16.7	Right	R on L	70	Right	Some	Never	159.00	M
etric											
## 118	Male	23.0	22.0	Left	L on R	83	Left	Some	Heavy	193.04	Imp

erial												
## 119	Female	18.5	18.0	Left	L on R	100	Neither	Some	Never	171.00	M	
etric												
## 120	Male	21.0	20.4	Right	L on R	100	Right	Freq	Heavy	184.00	M	
etric												
## 122	Male	22.5	22.5	Right	L on R	76	Right	Freq	Occas	177.00	M	
etric												
## 123	Female	18.5	18.0	Right	R on L	92	Right	Freq	Never	172.00	M	
etric												
## 124	Male	19.8	20.0	Left	L on R	59	Right	Freq	Never	180.00	M	
etric												
## 125	Male	18.5	18.1	Right	L on R	66	Left	Freq	Never	175.26	Imp	
erial												
## 127	Female	16.0	16.0	Right	R on L	68	Right	Freq	Never	172.72	Imp	
erial												
## 128	Male	18.8	19.1	Right	L on R	66	Neither	Freq	Regul	178.50	M	
etric												
## 129	Female	17.5	17.0	Right	R on L	74	Right	Freq	Never	157.00	M	
etric												
## 130	Female	16.4	16.5	Right	L on R	90	Right	Some	Never	152.00	M	
etric												
## 131	Male	22.0	21.5	Right	R on L	86	Right	Freq	Never	187.96	Imp	
erial												
## 132	Male	19.0	19.5	Right	L on R	60	Right	Some	Never	178.00	M	
etric												
## 134	Female	15.4	16.4	Left	L on R	80	Left	Freq	Occas	160.02	Imp	
erial												
## 135	Male	17.9	17.8	Right	R on L	85	Left	Some	Never	175.26	Imp	
erial												
## 136	Male	23.1	22.5	Right	L on R	90	Right	Some	Regul	189.00	M	
etric												
## 138	Male	22.0	22.0	Right	L on R	72	Right	Freq	Never	182.88	Imp	
erial												
## 140	Female	19.5	18.5	Right	L on R	68	Right	None	Never	167.00	M	
etric												
## 141	Female	18.0	18.6	Right	R on L	84	Right	Some	Never	175.00	M	
etric												
## 143	Female	19.0	18.8	Right	R on L	65	Right	Freq	Never	172.72	Imp	
erial												
## 144	Male	21.4	21.0	Right	L on R	96	Neither	Some	Never	180.00	M	
etric												
## 145	Female	20.0	19.5	Left	R on L	68	Neither	Freq	Never	172.00	M	
etric												
## 146	Male	18.5	18.5	Right	R on L	75	Neither	Some	Never	185.00	M	
etric												
## 147	Male	22.5	22.6	Right	L on R	64	Right	Freq	Regul	187.96	Imp	
erial												
## 148	Male	19.5	20.2	Right	R on L	60	Neither	Freq	Never	185.42	Imp	
erial												
## 149	Female	18.0	18.0	Right	L on R	92	Neither	Freq	Never	165.00	M	

etric											
## 150	Female	18.0	18.5	Right	R on L	64	Neither	Freq	Never	164.00	M
etric											
## 151	Male	21.8	22.3	Right	R on L	76	Left	Freq	Never	195.00	M
etric											
## 152	Female	13.0	12.5	Right	L on R	80	Right	Freq	Never	165.00	M
etric											
## 153	Female	16.3	16.2	Right	L on R	92	Right	Some	Regul	152.40	Imp
erial											
## 154	Male	21.5	21.6	Right	R on L	69	Right	Freq	Never	172.72	Imp
erial											
## 155	Male	18.9	19.1	Right	L on R	68	Right	None	Never	180.34	Imp
erial											
## 156	Male	20.5	20.0	Right	R on L	76	Right	Freq	Never	173.00	M
etric											
## 158	Female	18.9	19.2	Right	L on R	74	Right	Some	Never	167.64	Imp
erial											
## 160	Male	18.5	19.0	Right	L on R	84	Right	Freq	Regul	187.00	M
etric											
## 161	Female	17.5	17.1	Right	R on L	80	Left	None	Never	167.00	M
etric											
## 163	Male	20.2	20.3	Right	L on R	72	Neither	Some	Never	191.80	Imp
erial											
## 164	Female	16.5	16.9	Right	R on L	60	Neither	Freq	Occas	169.20	M
etric											
## 166	Female	17.6	17.2	Right	R on L	81	Left	Some	Never	168.00	M
etric											
## 167	Female	19.5	19.2	Right	R on L	70	Right	Some	Never	170.00	M
etric											
## 168	Female	16.5	15.0	Right	L on R	65	Right	Some	Regul	160.02	Imp
erial											
## 170	Male	19.0	18.5	Right	R on L	72	Right	Freq	Never	180.34	Imp
erial											
## 172	Male	20.5	19.5	Left	L on R	80	Right	Some	Occas	182.88	Imp
erial											
## 174	Female	18.0	17.5	Right	R on L	48	Neither	Freq	Never	165.00	M
etric											
## 175	Female	17.5	18.0	Right	R on L	68	Neither	Freq	Never	157.48	Imp
erial											
## 176	Female	19.0	18.5	Left	L on R	104	Left	Freq	Never	170.00	M
etric											
## 177	Male	20.5	20.5	Right	Neither	76	Right	Freq	Regul	172.72	Imp
erial											
## 178	Female	16.7	17.0	Right	L on R	84	Left	Freq	Never	164.00	M
etric											
## 180	Female	17.0	16.5	Right	R on L	70	Right	Some	Never	162.56	Imp
erial											
## 181	Male	19.0	19.5	Right	R on L	68	Right	Freq	Occas	172.00	M
etric											
## 182	Female	14.0	13.5	Right	R on L	87	Neither	Freq	Occas	165.10	Imp

erial												
## 183	Female	17.5	17.6	Right	L on R	79	Right	Some	Never	162.50	M	
etric												
## 184	Male	18.5	19.0	Right	L on R	70	Left	Freq	Never	170.00	M	
etric												
## 185	Male	18.0	18.5	Right	Neither	90	Right	Some	Never	175.00	M	
etric												
## 186	Male	20.5	20.7	Right	R on L	72	Right	Some	Never	168.00	M	
etric												
## 187	Female	17.0	17.0	Right	L on R	79	Right	Some	Never	163.00	M	
etric												
## 188	Male	18.5	18.5	Right	R on L	65	Right	None	Never	165.00	M	
etric												
## 189	Male	18.0	18.5	Right	R on L	62	Right	Freq	Never	173.00	M	
etric												
## 190	Male	18.5	18.0	Right	Neither	63	Neither	Freq	Never	196.00	M	
etric												
## 191	Male	20.0	19.5	Right	R on L	92	Right	Some	Never	179.10	Imp	
erial												
## 192	Male	22.0	22.5	Right	L on R	60	Right	Some	Never	180.00	M	
etric												
## 193	Male	17.9	18.4	Right	R on L	68	Left	None	Occas	176.00	M	
etric												
## 194	Female	17.6	17.8	Right	L on R	72	Left	Some	Never	160.02	Imp	
erial												
## 196	Female	17.0	17.6	Right	L on R	76	Right	Some	Never	165.00	M	
etric												
## 197	Female	15.0	13.0	Right	R on L	80	Neither	Freq	Never	170.18	Imp	
erial												
## 198	Male	16.0	15.5	Right	Neither	71	Right	Freq	Never	154.94	Imp	
erial												
## 199	Female	19.1	19.0	Right	R on L	80	Right	Some	Occas	170.00	M	
etric												
## 200	Female	17.5	16.5	Right	R on L	80	Neither	Some	Never	164.00	M	
etric												
## 201	Female	16.2	15.8	Right	R on L	61	Right	Some	Occas	167.00	M	
etric												
## 202	Male	21.0	21.0	Right	L on R	48	Neither	Freq	Never	174.00	M	
etric												
## 204	Female	18.5	18.0	Right	Neither	86	Right	None	Never	160.00	M	
etric												
## 205	Male	17.0	17.5	Right	R on L	80	Right	Some	Regul	179.10	M	
etric												
## 206	Female	17.5	17.0	Right	R on L	83	Neither	Freq	Occas	168.00	M	
etric												
## 207	Female	17.5	17.6	Right	L on R	76	Right	Some	Never	153.50	M	
etric												
## 208	Male	17.5	17.6	Right	R on L	84	Right	Some	Never	160.00	M	
etric												
## 209	Male	17.5	17.0	Left	L on R	97	Neither	None	Never	165.00	M	


```

etric
## 211 Female 18.6 18.6 Right L on R 74 Right Some Never 160.00 M
etric
## 212 Female 17.5 17.5 Left R on L 83 Neither Some Never 163.00 M
etric
## 214 Male 17.0 17.5 Right R on L 65 Right Some Never 165.00 M
etric
## 215 Female 18.0 17.8 Right L on R 68 Right Some Never 168.90 Imp
erial
## 218 Male 18.2 19.8 Right R on L 88 Right Freq Never 185.00 M
etric
## 220 Male 23.2 23.2 Right L on R 75 Right Freq Never 188.00 M
etric
## 222 Female 15.9 16.5 Right R on L 70 Right Freq Never 167.64 Imp
erial
## 223 Female 17.5 18.4 Right R on L 88 Right Some Never 162.56 Imp
erial
## 227 Female 18.8 18.3 Right R on L 80 Right Some Heavy 170.18 Imp
erial
## 228 Male 20.0 19.8 Right L on R 68 Right Freq Never 185.00 M
etric
## 229 Female 18.6 18.8 Right L on R 70 Right Freq Regul 167.00 M
etric
## 230 Male 18.6 19.6 Right L on R 71 Right Freq Occas 185.00 M
etric
## 231 Female 18.8 18.5 Right R on L 80 Right Some Never 169.00 M
etric
## 233 Female 18.0 18.0 Right L on R 85 Right Some Never 165.10 Imp
erial
## 234 Female 18.5 18.0 Right L on R 88 Right Some Never 160.00 M
etric
## 236 Male 21.0 21.5 Right R on L 90 Right Some Never 183.00 M
etric
## 237 Female 17.6 17.3 Right R on L 85 Right Freq Never 168.50 M
etric
## Age
## 1 18.250
## 2 17.583
## 5 23.667
## 6 21.000
## 7 18.833
## 8 35.833
## 9 19.000
## 10 22.333
## 11 28.500
## 14 17.500
## 17 19.333
## 18 18.333
## 20 17.917
## 21 17.917

```

##	22	18.167
##	23	17.833
##	24	18.250
##	27	17.500
##	28	18.083
##	30	19.250
##	32	17.500
##	33	39.750
##	34	17.167
##	36	18.000
##	38	17.917
##	39	35.500
##	42	17.083
##	44	17.500
##	47	18.917
##	48	19.417
##	49	18.417
##	50	30.750
##	51	18.500
##	52	17.500
##	53	18.333
##	54	17.417
##	55	20.000
##	57	17.167
##	59	17.667
##	61	20.333
##	62	17.333
##	63	17.500
##	65	18.583
##	71	17.583
##	73	17.667
##	74	17.417
##	75	17.750
##	76	20.667
##	77	23.583
##	79	17.083
##	82	20.167
##	85	17.167
##	86	17.250
##	87	18.000
##	88	18.750
##	89	21.583
##	91	19.667
##	93	19.667
##	95	22.833
##	97	19.417
##	98	23.250
##	100	19.083
##	102	17.750
##	104	20.167

105 17.667
106 18.250
109 18.583
110 17.750
111 24.167
112 18.167
113 21.167
114 17.917
115 17.417
116 20.500
117 22.917
118 18.917
119 18.917
120 20.083
122 18.250
123 17.500
124 17.417
125 21.000
127 17.667
128 18.083
129 18.000
130 18.333
131 20.000
132 18.750
134 18.500
135 18.417
136 19.167
138 19.333
140 18.667
141 17.500
143 17.250
144 19.000
145 19.167
146 19.000
147 23.000
148 32.667
149 20.000
150 20.167
151 25.500
152 18.167
153 23.500
154 70.417
155 43.833
156 23.583
158 44.250
160 17.917
161 18.417
163 17.500
164 29.083
166 18.500

167 18.167
168 32.750
170 17.333
172 18.667
174 18.667
175 17.750
176 17.250
177 36.583
178 23.083
180 17.167
181 23.417
182 17.083
183 17.250
184 23.833
185 18.750
186 21.167
187 24.667
188 18.500
189 20.333
190 20.083
191 18.917
192 27.333
193 18.917
194 17.250
196 26.500
197 17.000
198 17.167
199 19.167
200 17.500
201 19.250
202 21.333
204 20.167
205 18.667
206 17.083
207 17.417
208 18.583
209 19.500
211 17.167
212 17.250
214 20.417
215 17.083
218 19.333
220 18.917
222 17.333
223 18.167
227 18.417
228 17.417
229 20.333
230 19.333
231 18.167

```

## 233 17.667
## 234 16.917
## 236 17.167
## 237 17.750

#4
table(new_survey['Sex'])

##
## Female    Male
##      84      84

#5
table(new_survey['W.Hnd'])

##
## Left Right
##    12   156

#
sprintf(table(new_survey$`W.Hnd`)/length(new_survey$`W.Hnd`), fmt = '%#.3f')

## [1] "0.071" "0.929"

#7
data1=new_survey[new_survey$`W.Hnd` == 'Left',c("Sex","W.Hnd")]
data1

##      Sex W.Hnd
## 2    Male Left
## 18    Male Left
## 51    Male Left
## 118   Male Left
## 119 Female Left
## 124   Male Left
## 134 Female Left
## 145 Female Left
## 172   Male Left
## 176 Female Left
## 209   Male Left
## 212 Female Left

#8
never_smoke_male=sum(new_survey$Sex=="Male" & new_survey$`W.Hnd`=="Left" & new_survey$Smoke=="Never")
total_male=sum(new_survey$Sex=="Male")
percentage=(never_smoke_male*100)/total_male
percentage

## [1] 4.761905

```

PART-03

Part-3

Questions

Statistical Analysis of quantitative data

Use the newsurvey data obtained by cleaning 'na' values in survey data of MASS package to do the following:

1. Find the range of students' age participated in the survey.
2. Break the age range into non-overlapping sub-intervals by defining a sequence of equal distance break points of 10 by rounding the range to nearest integer.
3. Find the distribution of the age range according to the sub-intervals with cut with its right boundary opened. Display it in column form.
4. Which age range of students has mostly participated in the survey.
5. Similarly, find the frequency distribution of Wr.Hnd span and display it in column format.
6. Find the relative frequency of Wr.Hnd and display it by correcting to 3 decimal places.

```
#1
#range of the age
age=new_survey$Age
range(age)

## [1] 16.917 70.417

#2
#nearest integer can be 16 70
breaks = seq(16, 71, by=6.1)
breaks

## [1] 16.0 22.1 28.2 34.3 40.4 46.5 52.6 58.7 64.8 70.9

#3
age.cut = cut(age, breaks, right=FALSE)
age.cut

## [1] [16,22.1) [16,22.1) [22.1,28.2) [16,22.1) [16,22.1) [34.3,40
.4)
## [7] [16,22.1) [22.1,28.2) [28.2,34.3) [16,22.1) [16,22.1) [16,22.1
)
## [13] [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1
)
```

```
## [19] [16,22.1) [16,22.1) [16,22.1) [34.3,40.4) [16,22.1) [16,22.1
)
## [25] [16,22.1) [34.3,40.4) [16,22.1) [16,22.1) [16,22.1) [16,22.1
)
## [31] [16,22.1) [28.2,34.3) [16,22.1) [16,22.1) [16,22.1) [16,22.1
)
## [37] [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1
)
## [43] [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1
)
## [49] [22.1,28.2) [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1
)
## [55] [16,22.1) [16,22.1) [16,22.1) [16,22.1) [22.1,28.2) [16,22.1
)
## [61] [22.1,28.2) [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1
)
## [67] [16,22.1) [16,22.1) [22.1,28.2) [16,22.1) [16,22.1) [16,22.1
)
## [73] [16,22.1) [16,22.1) [22.1,28.2) [16,22.1) [16,22.1) [16,22.1
)
## [79] [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1
)
## [85] [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1
)
## [91] [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1
)
## [97] [16,22.1) [16,22.1) [22.1,28.2) [28.2,34.3) [16,22.1) [16,22.1
)
## [103] [22.1,28.2) [16,22.1) [22.1,28.2) [64.8,70.9) [40.4,46.5) [22.1,28
.2)
## [109] [40.4,46.5) [16,22.1) [16,22.1) [16,22.1) [28.2,34.3) [16,22.1
)
## [115] [16,22.1) [28.2,34.3) [16,22.1) [16,22.1) [16,22.1) [16,22.1
)
## [121] [16,22.1) [34.3,40.4) [22.1,28.2) [16,22.1) [22.1,28.2) [16,22.1
)
## [127] [16,22.1) [22.1,28.2) [16,22.1) [16,22.1) [22.1,28.2) [16,22.1
)
## [133] [16,22.1) [16,22.1) [16,22.1) [22.1,28.2) [16,22.1) [16,22.1
)
## [139] [22.1,28.2) [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1
)
## [145] [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1
)
## [151] [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1
)
## [157] [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1
)
## [163] [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1) [16,22.1
```

```
)  
## 9 Levels: [16,22.1) [22.1,28.2) [28.2,34.3) [34.3,40.4) ... [64.8,70.9)
```

#4

```
age.freq=table(age.cut)  
cbind(age.freq)
```

```
##           age.freq  
## [16,22.1)      139  
## [22.1,28.2)     17  
## [28.2,34.3)      5  
## [34.3,40.4)      4  
## [40.4,46.5)      2  
## [46.5,52.6)      0  
## [52.6,58.7)      0  
## [58.7,64.8)      0  
## [64.8,70.9)      1
```

```
which(age.freq==max(age.freq))
```

```
## [16,22.1)  
##           1
```

#5

```
w=new_survey$Wr.Hnd  
range(w)
```

```
## [1] 13.0 23.2
```

```
break_w = seq(13, 25, by=2.0)  
w.cut = cut(w, break_w, right=FALSE)  
w.freq=table(w.cut)  
cbind(w.freq)
```

```
##           w.freq  
## [13,15)      2  
## [15,17)     16  
## [17,19)     82  
## [19,21)     41  
## [21,23)     22  
## [23,25)      5
```

#6

```
cbind(w.freq/length(w.cut))
```

```
##           [,1]  
## [13,15) 0.01190476  
## [15,17) 0.09523810  
## [17,19) 0.48809524  
## [19,21) 0.24404762  
## [21,23) 0.13095238  
## [23,25) 0.02976190
```


PART-04

Part-4

Questions

Numerical Measures on data

Use the newsurvey data obtained by cleaning 'na' values in survey data of MASS package to do the following:

1. Find the average age of the students participated in the survey.
2. Compute the standard deviation and variance of the height of the students participated in the survey.
3. Compute the quartile of write hand span.
4. Find the correlation between write hand span and pulse rate of the students.
5. Find the average age of the students based on how often the student exercises.
6. Find the standard deviation of height of the students under different categories of span of non-writing hand.
7. Obtain the summary statistics of pulse rate of the students.

```
#1
avg_age=mean(age)
avg_age

## [1] 20.43358

#2
h=new_survey$Height
sd(h)

## [1] 9.959542

var(h)

## [1] 99.19248

#3
hs=new_survey$Wr.Hnd
quantile(hs)

##  0%  25%  50%  75% 100%
## 13.0 17.5 18.5 20.0 23.2

#4
cor(new_survey$`Wr.Hnd`, new_survey$Pulse)

## [1] -0.01382144
```

```

#5
x=levels(new_survey$Exer)
for(i in x)
{
  m=mean(new_survey$Exer==i)
  print(paste("For ", i , " the mean is: ",m))
}

## [1] "For Freq the mean is: 0.505952380952381"
## [1] "For None the mean is: 0.0833333333333333"
## [1] "For Some the mean is: 0.410714285714286"

#6
nwh=new_survey$NW.Hnd
n=max(nwh)
i=min(nwh)
#dividing height based on the range of values of NW.Hnd (12.5-15.5,15.5-18.5,
18.5-21.5,21.5-23.5)
d1=new_survey[new_survey$`NW.Hnd` >=i & new_survey$`NW.Hnd` < (i+3),c("Height
")]
d1

## [1] 157.48 165.00 160.02 165.10 170.18

i=i+3
d2=new_survey[new_survey$`NW.Hnd` >=i & new_survey$`NW.Hnd` < (i+3),c("Height
")]
d2

## [1] 173.00 172.72 182.88 157.00 156.20 157.00 180.34 180.00 166.40 154.94
## [11] 168.00 165.00 170.18 167.64 164.00 170.00 171.00 167.64 165.00 170.00
## [21] 165.00 167.00 162.56 179.00 165.00 166.50 175.00 163.00 180.00 170.18
## [31] 162.56 158.00 159.00 171.00 172.00 175.26 172.72 157.00 152.00 160.02
## [41] 175.26 165.00 152.40 167.00 169.20 168.00 165.00 157.48 164.00 162.56
## [51] 162.50 163.00 196.00 176.00 160.02 165.00 154.94 164.00 167.00 160.00
## [61] 179.10 168.00 153.50 160.00 165.00 163.00 165.00 168.90 167.64 162.56
## [71] 170.18 165.10 160.00 168.50

i=i+3
d3=new_survey[new_survey$`NW.Hnd` >=i & new_survey$`NW.Hnd` < (i+3),c("Height
")]
d3

## [1] 177.80 165.00 175.00 167.00 155.00 182.88 177.00 180.34 184.00 172.72
## [11] 175.26 167.00 180.00 190.00 182.50 176.50 190.00 179.00 177.80 178.00
## [21] 165.10 176.50 170.00 183.00 179.00 190.00 175.26 170.00 170.00 180.34
## [31] 175.00 190.50 185.00 184.00 180.00 178.50 178.00 167.00 175.00 172.72
## [41] 180.00 172.00 185.00 185.42 164.00 180.34 173.00 167.64 187.00 191.80
## [51] 170.00 180.34 182.88 170.00 172.72 172.00 170.00 175.00 168.00 165.00
## [61] 173.00 179.10 170.00 174.00 160.00 185.00 185.00 167.00 185.00 169.00

```

```

i=i+3
d4=new_survey[new_survey$`NW.Hnd` >=i & new_survey$`NW.Hnd` < (i+3),c("Height
")]
d4

## [1] 190.50 185.00 180.00 170.00 200.00 182.00 187.00 167.64 193.04 177.00
## [11] 187.96 189.00 182.88 187.96 195.00 172.72 180.00 188.00 183.00

print("standard deviation")
## [1] "standard deviation"

sd(d1)
## [1] 4.943873

sd(d2)
## [1] 7.770754

sd(d3)
## [1] 7.799748

sd(d4)
## [1] 8.37023

#7
summary(new_survey$Pulse)
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   35.00   66.75   72.00   74.02   80.00  104.00

```