NAME – ANSH GOEL

REG NO. – 20BCE1798

COURSE NAME – FOUNDATION OF DATA ANALYTICS (FDA)

COURSE CODE: 3505

DATE – 3rd November, 2022

LAB 11

Data Manipulation using dplyr package

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

1. Install the dplyr package and import it.
2. Filter all male left handers.
3. Display all female right handers who keep right on top while clapping.
4. Display all students who never exercise.
5. Display only the gender, age and writing hand of the students.
6. Display the age, pulse rate and writing hand span of female left handers.

Note: use dplyr::select while using select operator as dplyr package conflicts with MASS package

---

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

#1.Install the dplyr package and import it.

```
install.packages("dplyr")

library("dplyr")

library("MASS")

data("survey")

newsurvey <- na.omit(survey)

dplyr::select(newsurvey)
```

```
> library("dplyr")
> library("MASS")
> data("survey")
> newsurvey <- na.omit(survey)
> dplyr::select(newsurvey)
data frame with 0 columns and 168 rows
>
```

#2.Filter all male left handers.

head(filter(newsurvey,Sex=="Male"&W.Hnd=="Left"),3)

```
> #2.Filter all male left handers.
> head(filter(newsurvey,Sex=="Male"&W.Hnd=="Left"),3)
   Sex Wr.Hnd NW.Hnd W.Hnd   Fold Pulse  Clap Exer Smoke Height      M.I    Age
1 Male   19.5   20.5  Left R on L   104  Left None Regul 177.80 Imperial 17.583
2 Male   19.4   19.2  Left R on L    74 Right Some Never 182.88 Imperial 18.333
3 Male   22.0   21.5  Left R on L    55  Left Freq Never 200.00   Metric 18.500
```

#3.Display all female right handers who keep right on top while clapping.

head(filter(newsurvey,Sex=="Male"&W.Hnd=="Left"),3)

```
> #3.Display all female right handers who keep right on top while clapping.
> head(filter(newsurvey,Sex=="Male"&W.Hnd=="Left"),3)
   Sex Wr.Hnd NW.Hnd W.Hnd   Fold Pulse  Clap Exer Smoke Height      M.I    Age
1 Male   19.5   20.5  Left R on L   104  Left None Regul 177.80 Imperial 17.583
2 Male   19.4   19.2  Left R on L    74 Right Some Never 182.88 Imperial 18.333
3 Male   22.0   21.5  Left R on L    55  Left Freq Never 200.00   Metric 18.500
```

#4.Display all students who never exercise.

head(filter(newsurvey,Exer=="None"),3)

```
> #4.Display all students who never exercise.
> head(filter(newsurvey,Exer=="None"),3)
     Sex Wr.Hnd NW.Hnd W.Hnd   Fold Pulse  Clap Exer Smoke Height      M.I    Age
1   Male   19.5   20.5  Left R on L   104  Left None Regul  177.8 Imperial 17.583
2   Male   22.5   23.0 Right R on L    96 Right None Never  170.0   Metric 19.417
3 Female   18.0   17.9 Right R on L    50  Left None Never  165.0   Metric 30.750
```

#5.Display only the gender, age and writing hand of the students.

head(filter(newsurvey[,c("Sex","Age","W.Hnd")]),3)

```
> #5.Display only the gender, age and writing hand of the students.\
> head(filter(newsurvey[,c("Sex","Age","W.Hnd")]),3)
     Sex    Age W.Hnd
1 Female 18.250 Right
2   Male 17.583  Left
3   Male 23.667 Right
```

#6.Display the age, pulse rate and writing hand span of female left handers.

head(filter(newsurvey,Sex=="Female"&W.Hnd=="Left")[,c("Age","Pulse","W.Hnd")],3)

```
> #6.Display the age, pulse rate and writing hand span of female left handers.
> head(filter(newsurvey,Sex=="Female"&W.Hnd=="Left")[,c("Age","Pulse","W.Hnd")],3)
     Age Pulse W.Hnd
1 18.917   100  Left
2 18.500    80  Left
3 19.167    68  Left
```

More functions in dplyr package

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

1. Install the dplyr package and import it.
2. Arrange all male left handers according to descending order of their heights.
3. Introduce a new column hand_span which contains the value as difference between the span of writing hand and non-writing hand and display it along with gender, writing hand and non-writing hand span.
4. Display the average writing span of male and female left handers.
5. Find the maximum pulse rate of male left and right handers.

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

#1.Install the dplyr package and import it.

```
install.packages("dplyr")

library("dplyr")

library("MASS")

data("survey")

newsurvey <- na.omit(survey)

dplyr::select(newsurvey)
```

```
> library("dplyr")
> library("MASS")
> data("survey")
> newsurvey <- na.omit(survey)
> dplyr::select(newsurvey)
data frame with 0 columns and 168 rows
```

#2.Arrange all male left handers according to descending order of their heights.

```
filter(newsurvey,Sex=="Male"&W.Hnd=="Left") %>%

arrange(desc(Height))
```

```
> #2.Arrange all male left handers according to descending order of their heights.
> filter(newsurvey,Sex=="Male"&W.Hnd=="Left") %>%
+ arrange(desc(Height))
   Sex Wr.Hnd NW.Hnd W.Hnd   Fold Pulse    Clap Exer Smoke Height      M.I    Age
1 Male   22.0   21.5  Left R on L    55    Left Freq Never 200.00   Metric 18.500
2 Male   23.0   22.0  Left L on R    83    Left Some Heavy 193.04 Imperial 18.917
3 Male   19.4   19.2  Left R on L    74   Right Some Never 182.88 Imperial 18.333
4 Male   20.5   19.5  Left L on R    80   Right Some Occas 182.88 Imperial 18.667
5 Male   19.8   20.0  Left L on R    59   Right Freq Never 180.00   Metric 17.417
6 Male   19.5   20.5  Left R on L   104    Left None Regul 177.80 Imperial 17.583
7 Male   17.5   17.0  Left L on R    97 Neither None Never 165.00   Metric 19.500
```

#3.Introduce a new column hand_span which contains the value as difference between the span of writing hand and non-writing hand and display it along with gender, writing hand and non-writing hand span.

```
newsurvey$hand_span <- newsurvey$Wr.Hnd - newsurvey$NW.Hnd

head(newsurvey[,c("Sex","Wr.Hnd","NW.Hnd","hand_span")],3)
```

```
> #3.Introduce a new column hand_span which contains the value as difference between the span of writi
ng hand and non-writing hand and display it along with gender, writing hand and non-writing hand span.
> newsurvey$hand_span <- newsurvey$Wr.Hnd - newsurvey$NW.Hnd
> head(newsurvey[,c("Sex","Wr.Hnd","NW.Hnd","hand_span")],3)
     Sex Wr.Hnd NW.Hnd hand_span
1 Female   18.5   18.0       0.5
2   Male   19.5   20.5      -1.0
5   Male   20.0   20.0       0.0
```

#4.Display the average writing span of male and female left handers.

mean(filter(newsurvey,Sex=="Male"&W.Hnd=="Left")$Wr.Hnd)

mean(filter(newsurvey,Sex=="Female"&W.Hnd=="Left")$Wr.Hnd)

```
> #4.Display the average writing span of male and female left handers.
> mean(filter(newsurvey,Sex=="Male"&W.Hnd=="Left")$Wr.Hnd)
[1] 20.24286
```

#5.Find the maximum pulse rate of male left and right handers.

max(filter(newsurvey,Sex=="Male"&W.Hnd=="Left")[,"Pulse"])

max(filter(newsurvey,Sex=="Male"&W.Hnd=="Right")[,"Pulse"])

```
> #5.Find the maximum pulse rate of male left and right handers.
> max(filter(newsurvey,Sex=="Male"&W.Hnd=="Left")[,"Pulse"])
[1] 104
> max(filter(newsurvey,Sex=="Male"&W.Hnd=="Right")[,"Pulse"])
[1] 100
```