Date: 2nd February,2022

Executive Summary Report 2

Course Name: ALY6000 Introduction to Analytics [CRN 22279]

Student Name - Poonam Dighe [NUID 002965616]

Professor - Mohammad Shafiqul Islam

Introduction

The following project gives a brief understanding of R and RStudio as well as some hands-on experience.

A. Provide an analysis of descriptive characteristics of the data set provided by your instructor. This includes pertinent statistics including mean, median, quartiles, variance, standard deviation, skew, kurtosis, outliers etc. Include R console screen snipet to support your observations and conclusions. Below is a sample excerpt of an analysis of Harrison Lake fish from the BullTroutRML2 dataset.

Ans-

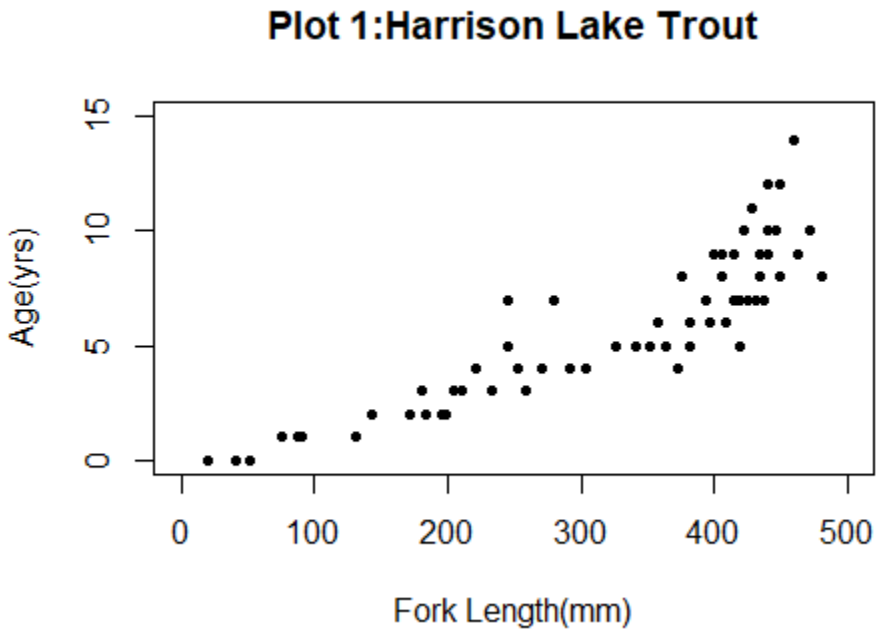
```
> str(New_Harrison)
'data.frame':   61 obs. of  4 variables:
 $ age : int  14 12 10 10 9 9 9 8 8 7 ...
 $ fl  : int  459 449 471 446 400 440 462 480 449 437 ...
 $ lake: Factor w/ 2 levels "Harrison","Osprey": 1 1 1 1 1 1 1 1 1 1 ...
 $ era : Factor w/ 2 levels "1977-80","1997-01": 1 1 1 1 1 1 1 1 1 1 ...
>
```

- After applying a **filter [lake == Harrison]** to the original dataset i.e. **BullTroutRML2** , we get a new dataset which is named **New_Harrison**.
- The new dataset shows **61** observations and **4** variables: namely – **age, fl, lake, era**. Data type of age and fl is integer whereas lake and era factor data structure.

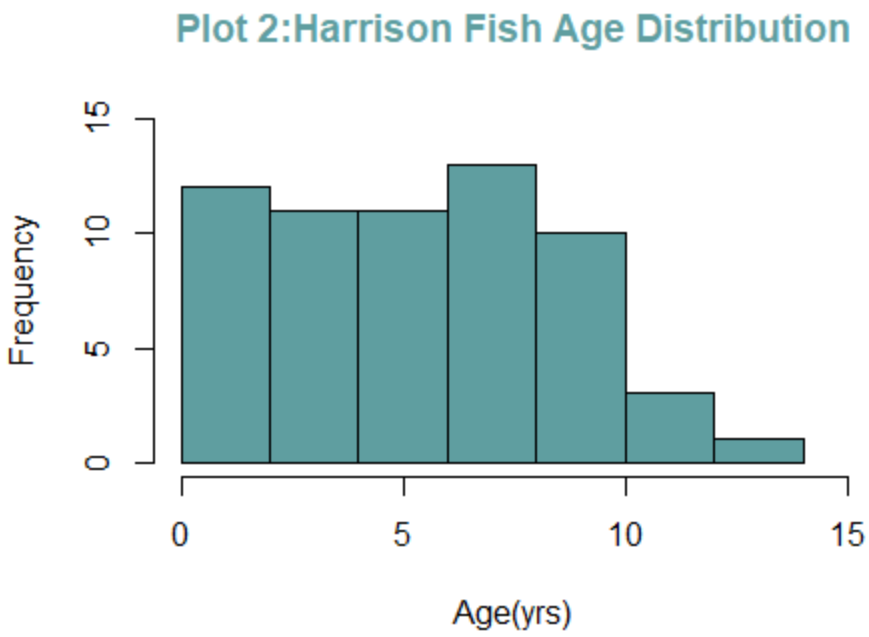
```
> summary(New_Harrison)
      age              fl            lake         era
 Min.   : 0.000   Min.   : 20   Harrison:61   1977-80:23
 1st Qu.: 3.000   1st Qu.:221   Osprey  : 0   1997-01:38
 Median : 6.000   Median :372
 Mean   : 5.754   Mean   :319
 3rd Qu.: 8.000   3rd Qu.:425
 Max.   :14.000   Max.   :480
```

- Above data shows the summary of the filtered dataset **New_Harrison**. Harrison Lake has a minimum age of 0 years and a maximum of 14 years. The mean age of the lake is 5.754 years and its median is 6.
- The fork length has a minimum of 20mm and a maximum of 480mm. Its median is 372mm and mean is 319mm. The era has two levels, 1977-1980 and 1997-2001 has 23 and 38 observations respectively.

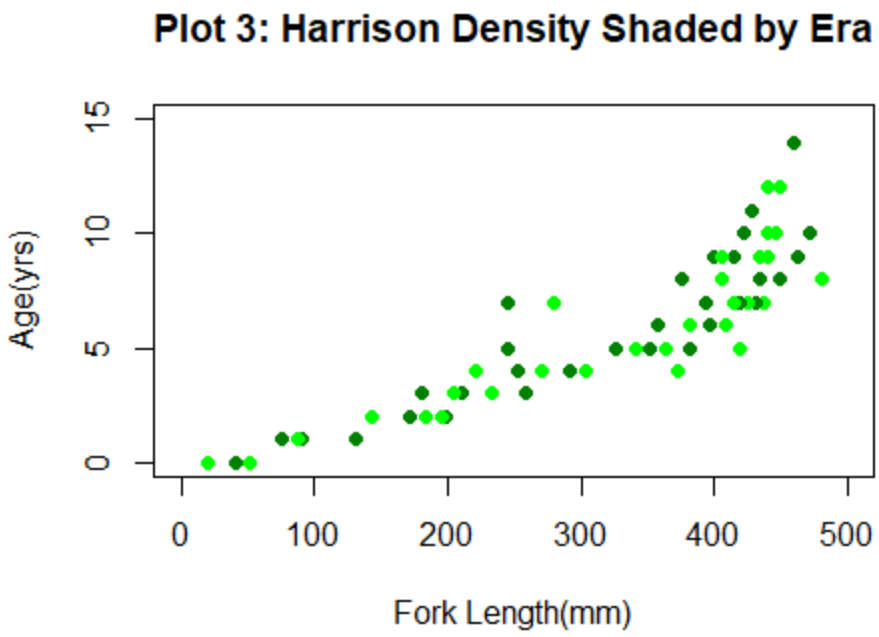
B. Provide the executive with visualizations (at least 6) in that help them see the key characteristics you want to highlight. They can be boxplots, histograms, frequency and probability distributions, barplots (bar charts) or pareto. Not only is the goal to present your visual results, but also to explain the significance of what the visuals are displaying.



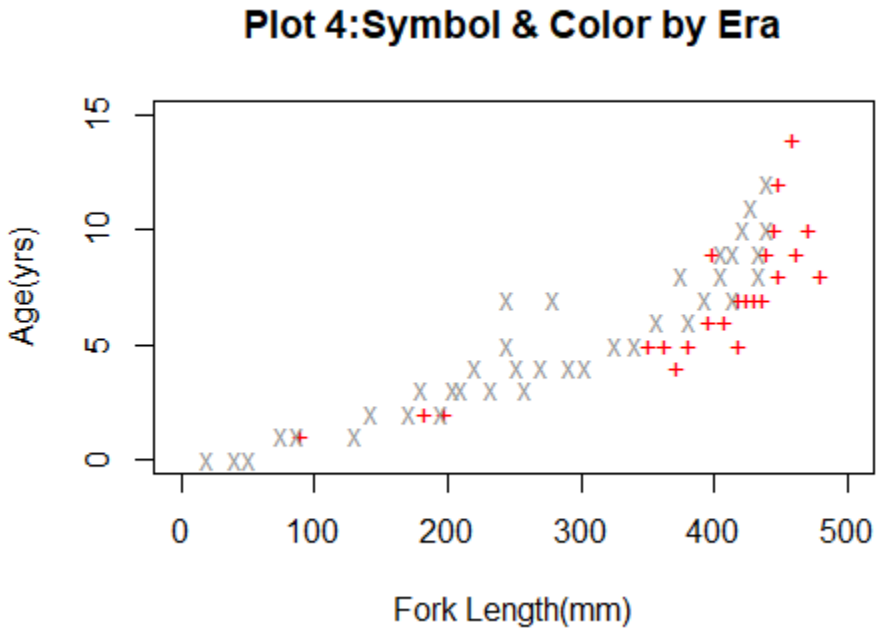
- The harrison data new dataset was used to construct the Harrison Lake Trout Scatter plot. Fork length is recorded in millimetres on the x-axis, while age is measured in years on the y-axis.
- The lake's length has gradually risen over time.The lake grew from 20mm to 480mm in 15 years, according to summary(harrison data new).



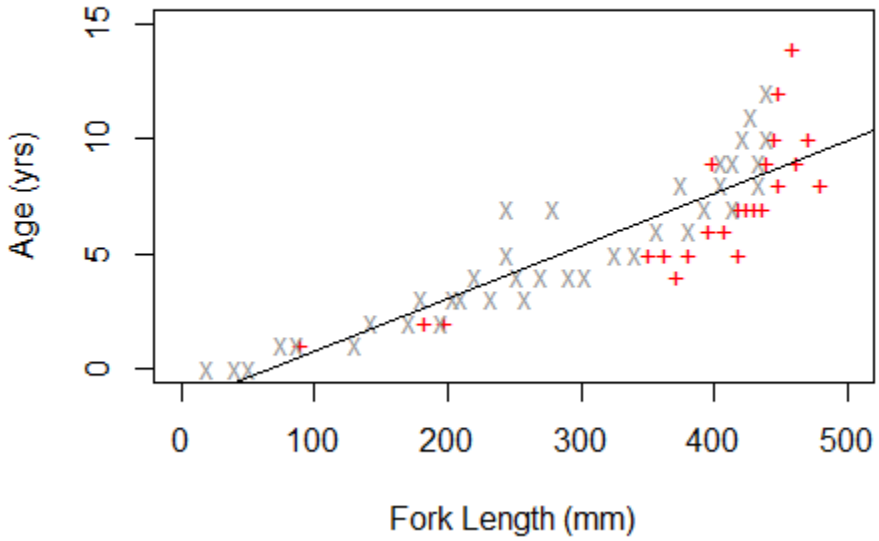
- The histogram depicts the age-related frequency of fish in Harrison Lake. We can observe from the graph that the maximum age of the fish is around 13-14 years. There are at least 12 fish between the ages of 0 and 2.
- Their frequency decreases from 13 to 10 years after crossing the 8-year mark. At the age of ten years, the frequency of fish drops dramatically, dropping to 2-3 by the age of twelve years, and 1-2 by the age of fourteen years.



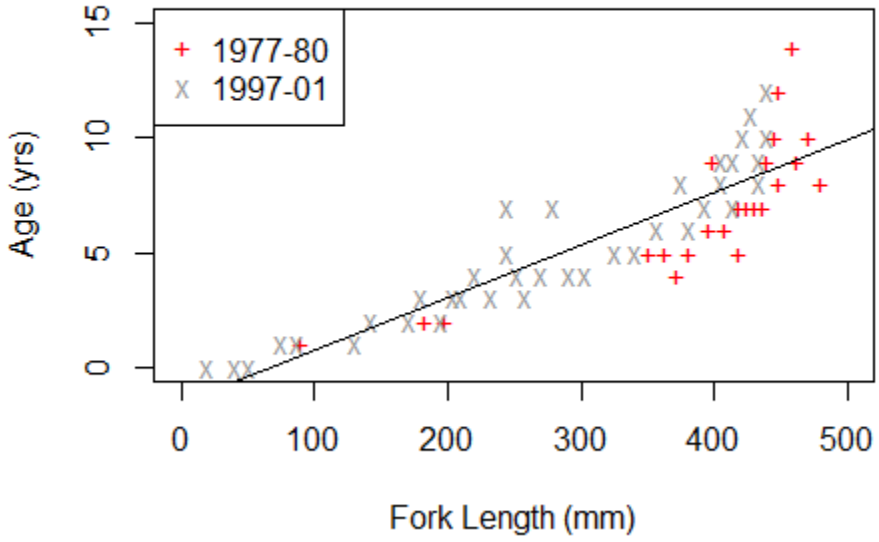
- The above Overdense plot of the **New_Harrison** dataset, which shows fork length vs. lake age. We have two different coloured points since there are two stages of the era.
- The shortest fork length is roughly 20mm, while the longest fork length is approximately 480mm. Three black circles can be seen on the graph if you look closely. The overlapping of the two eras is depicted by these circles.



Plot 5: Regression Overlay



Plot 6: Legend Overlay



- The fork length versus age graph of the filtered data set, **New_Harrison**, is shown in panels 4,5 and 6. The length of the fork is measured in millimetres, while the age is measured in years.

- The eras in all three graphs are of two levels. One is symbolised by a red cross, while the other is symbolised by grey Xs.
- 1977-80, is represented by the red cross, while the second era, 1997-01, is represented by the grey X.

C. Finally, provide a clear two to three sentence paragraph summary of the key points that you want the audience to walk away with regarding your analysis. This summary should present accurate analysis and be supported by the data presented in the rest of the report.

- New Harrison was the dataset that was used to create this report. This dataset was constructed after applying a filter to the dataset BullTroutRML2.
- Harrison Lake and Osprey Lake are both included in the original dataset, but only Harrison Lake data is included in **New_Harrison**.
- There are 61 observations in this new dataset, with four variables: age, fl, lake, and era. The length of the fork and the age of the lake are found to be proportionate.
- The data was visually portrayed in such a way that two eras were clearly visible. We noticed that not all the points were on the regression line when we generated it. Some can be found on both sides of the regression line. All graphs are thoroughly explained.

Appendix

1. Book R in Action by Robert I. Kabacoff Chap 6 – Basic Graphs.

Last Accessed: 1st February 2022

2. You Tube - Data visualization in R

https://www.youtube.com/watch?v=_WyUme_H2ZQ

Last Accessed: 1st February 2022

3. Histograms in R Language(Geeksforgeeks December 2021)

https://www.geeksforgeeks.org/histograms-in-r-language/

Last Accessed: 1st February 2022

4. How to check the data structure of an object in R (November 2021)

str in R: How to Check Data Structure of Object in R (r-lang.com)

Last Accessed: 2nd February 2022

M2_Assignment_2.R

poonam

2022-02-03

```
r = getOption("repos")
r["CRAN"]="http://cran.us.r-project.org"
options(repos=r)

#Q.1.Print your name at the top of the script.

print("Plotting Basics: Dighe")
```

```
## [1] "Plotting Basics: Dighe"
```

```
#Q.2.Import libraries including: FSA, FSAdata, magrittr, dplyr, plotrix, ggpl
ot2, and moments

install.packages("FSA")
```

```
## Installing package into 'C:/Users/poonam/Documents/R/win-library/4.1'
## (as 'lib' is unspecified)
```

```
## package 'FSA' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##  C:\Users\poonam\AppData\Local\Temp\RtmpATQlkv\downloaded_packages
```

```
install.packages("FSAdata")
```

```
## Installing package into 'C:/Users/poonam/Documents/R/win-library/4.1'
## (as 'lib' is unspecified)
```

```
## package 'FSAdata' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##  C:\Users\poonam\AppData\Local\Temp\RtmpATQlkv\downloaded_packages
```

```
install.packages("magrittr")
```

```
## Installing package into 'C:/Users/poonam/Documents/R/win-library/4.1'
## (as 'lib' is unspecified)
```

```
## package 'magrittr' successfully unpacked and MD5 sums checked

## Warning: cannot remove prior installation of package 'magrittr'
```

```
## Warning in file.copy(savedcopy, lib, recursive = TRUE): problem copying C:
## \Users\poonam\Documents\R\win-library\4.1\00LOCK\magrittr\libs\x64\magritt
r.dll
## to C:\Users\poonam\Documents\R\win-library\4.1\magrittr\libs\x64\magrittr.
```

```
dll:
## Permission denied

## Warning: restored 'magrittr'

##
## The downloaded binary packages are in
##  C:\Users\poonam\AppData\Local\Temp\RtmpATQlkv\downloaded_packages

install.packages("dplyr")

## Installing package into 'C:/Users/poonam/Documents/R/win-library/4.1'
## (as 'lib' is unspecified)

## package 'dplyr' successfully unpacked and MD5 sums checked

## Warning: cannot remove prior installation of package 'dplyr'

## Warning in file.copy(savedcopy, lib, recursive = TRUE): problem copying C:
## \Users\poonam\Documents\R\win-library\4.1\00LOCK\dplyr\libs\x64\dplyr.dll
to C:
## \Users\poonam\Documents\R\win-library\4.1\dplyr\libs\x64\dplyr.dll: Permis
sion
## denied

## Warning: restored 'dplyr'

##
## The downloaded binary packages are in
##  C:\Users\poonam\AppData\Local\Temp\RtmpATQlkv\downloaded_packages

install.packages("plotrix")

## Installing package into 'C:/Users/poonam/Documents/R/win-library/4.1'
## (as 'lib' is unspecified)

## package 'plotrix' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##  C:\Users\poonam\AppData\Local\Temp\RtmpATQlkv\downloaded_packages

install.packages("ggplot2")

## Installing package into 'C:/Users/poonam/Documents/R/win-library/4.1'
## (as 'lib' is unspecified)

## package 'ggplot2' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##  C:\Users\poonam\AppData\Local\Temp\RtmpATQlkv\downloaded_packages

install.packages("moments")
```

```
## Installing package into 'C:/Users/poonam/Documents/R/win-library/4.1'
## (as 'lib' is unspecified)

## package 'moments' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##  C:\Users\poonam\AppData\Local\Temp\RtmpATQlkv\downloaded_packages

library("FSA")

## ## FSA v0.9.1. See citation('FSA') if used in publication.
## ## Run fishR() for related website and fishR('IFAR') for related book.

library("FSAdata")

## ## FSAdata v0.3.8. See ?FSAdata to find data for specific fisheries analys
es.

library("magrittr")
library("dplyr")

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

library("plotrix")
library("ggplot2")
library("moments")

#Q.3.Load the BullTroutRML2 dataset

BullTroutRML2

##    age  fl     lake     era
## 1   14 459 Harrison 1977-80
## 2   12 449 Harrison 1977-80
## 3   10 471 Harrison 1977-80
## 4   10 446 Harrison 1977-80
## 5    9 400 Harrison 1977-80
## 6    9 440 Harrison 1977-80
## 7    9 462 Harrison 1977-80
## 8    8 480 Harrison 1977-80
## 9    8 449 Harrison 1977-80
## 10   7 437 Harrison 1977-80
## 11   7 431 Harrison 1977-80
```

```
## 12   7 425 Harrison 1977-80
## 13   7 419 Harrison 1977-80
## 14   6 409 Harrison 1977-80
## 15   6 397 Harrison 1977-80
## 16   5 419 Harrison 1977-80
## 17   5 381 Harrison 1977-80
## 18   5 363 Harrison 1977-80
## 19   5 351 Harrison 1977-80
## 20   4 372 Harrison 1977-80
## 21   2 199 Harrison 1977-80
## 22   2 184 Harrison 1977-80
## 23   1  91 Harrison 1977-80
## 24  12 440 Harrison 1997-01
## 25  11 428 Harrison 1997-01
## 26  10 440 Harrison 1997-01
## 27  10 422 Harrison 1997-01
## 28   9 434 Harrison 1997-01
## 29   9 415 Harrison 1997-01
## 30   9 406 Harrison 1997-01
## 31   8 434 Harrison 1997-01
## 32   8 406 Harrison 1997-01
## 33   8 375 Harrison 1997-01
## 34   7 415 Harrison 1997-01
## 35   7 394 Harrison 1997-01
## 36   6 381 Harrison 1997-01
## 37   6 357 Harrison 1997-01
## 38   5 341 Harrison 1997-01
## 39   5 326 Harrison 1997-01
## 40   4 304 Harrison 1997-01
## 41   4 292 Harrison 1997-01
## 42   4 270 Harrison 1997-01
## 43   4 252 Harrison 1997-01
## 44   4 221 Harrison 1997-01
## 45   3 258 Harrison 1997-01
## 46   3 233 Harrison 1997-01
## 47   3 211 Harrison 1997-01
## 48   3 205 Harrison 1997-01
## 49   3 180 Harrison 1997-01
## 50   2 196 Harrison 1997-01
## 51   2 171 Harrison 1997-01
## 52   2 143 Harrison 1997-01
## 53   1 131 Harrison 1997-01
## 54   1  88 Harrison 1997-01
## 55   1  75 Harrison 1997-01
## 56   0  51 Harrison 1997-01
## 57   0  41 Harrison 1997-01
## 58   0  20 Harrison 1997-01
## 59   7 245 Harrison 1997-01
## 60   7 279 Harrison 1997-01
## 61   5 245 Harrison 1997-01
```

```
## 62   8 360   Osprey 1977-80
## 63   8 357   Osprey 1977-80
## 64   7 357   Osprey 1977-80
## 65   7 329   Osprey 1977-80
## 66   6 385   Osprey 1977-80
## 67   6 323   Osprey 1977-80
## 68   5 369   Osprey 1977-80
## 69   5 326   Osprey 1977-80
## 70   4 357   Osprey 1977-80
## 71   4 326   Osprey 1977-80
## 72   4 258   Osprey 1977-80
## 73   4 239   Osprey 1977-80
## 74   3 221   Osprey 1977-80
## 75   3 258   Osprey 1977-80
## 76   3 276   Osprey 1977-80
## 77  11 688   Osprey 1997-01
## 78  10 369   Osprey 1997-01
## 79   9 400   Osprey 1997-01
## 80   8 381   Osprey 1997-01
## 81   8 332   Osprey 1997-01
## 82   7 394   Osprey 1997-01
## 83   7 388   Osprey 1997-01
## 84   7 354   Osprey 1997-01
## 85   7 320   Osprey 1997-01
## 86   6 320   Osprey 1997-01
## 87   6 347   Osprey 1997-01
## 88   6 360   Osprey 1997-01
## 89   5 354   Osprey 1997-01
## 90   5 335   Osprey 1997-01
## 91   5 313   Osprey 1997-01
## 92   5 289   Osprey 1997-01
## 93   4 313   Osprey 1997-01
## 94   4 298   Osprey 1997-01
## 95   3 279   Osprey 1997-01
## 96   3 273   Osprey 1997-01
```

```
#Q.4.Print the first and last 3 records from the BullTroutRMS2 dataset

head(BullTroutRML2,n=3)

##   age  fl     lake     era
## 1  14 459 Harrison 1977-80
## 2  12 449 Harrison 1977-80
## 3  10 471 Harrison 1977-80

tail(BullTroutRML2,n=3)

##    age  fl   lake     era
## 94   4 298 Osprey 1997-01
## 95   3 279 Osprey 1997-01
## 96   3 273 Osprey 1997-01
```

```r
#Q.5.Remove all records except those from Harrison Lake

New_Harrison <- filter(BullTroutRML2, lake=="Harrison")
New_Harrison
```

```
##    age  fl     lake     era
## 1   14 459 Harrison 1977-80
## 2   12 449 Harrison 1977-80
## 3   10 471 Harrison 1977-80
## 4   10 446 Harrison 1977-80
## 5    9 400 Harrison 1977-80
## 6    9 440 Harrison 1977-80
## 7    9 462 Harrison 1977-80
## 8    8 480 Harrison 1977-80
## 9    8 449 Harrison 1977-80
## 10   7 437 Harrison 1977-80
## 11   7 431 Harrison 1977-80
## 12   7 425 Harrison 1977-80
## 13   7 419 Harrison 1977-80
## 14   6 409 Harrison 1977-80
## 15   6 397 Harrison 1977-80
## 16   5 419 Harrison 1977-80
## 17   5 381 Harrison 1977-80
## 18   5 363 Harrison 1977-80
## 19   5 351 Harrison 1977-80
## 20   4 372 Harrison 1977-80
## 21   2 199 Harrison 1977-80
## 22   2 184 Harrison 1977-80
## 23   1  91 Harrison 1977-80
## 24  12 440 Harrison 1997-01
## 25  11 428 Harrison 1997-01
## 26  10 440 Harrison 1997-01
## 27  10 422 Harrison 1997-01
## 28   9 434 Harrison 1997-01
## 29   9 415 Harrison 1997-01
## 30   9 406 Harrison 1997-01
## 31   8 434 Harrison 1997-01
## 32   8 406 Harrison 1997-01
## 33   8 375 Harrison 1997-01
## 34   7 415 Harrison 1997-01
## 35   7 394 Harrison 1997-01
## 36   6 381 Harrison 1997-01
## 37   6 357 Harrison 1997-01
## 38   5 341 Harrison 1997-01
## 39   5 326 Harrison 1997-01
## 40   4 304 Harrison 1997-01
## 41   4 292 Harrison 1997-01
## 42   4 270 Harrison 1997-01
## 43   4 252 Harrison 1997-01
## 44   4 221 Harrison 1997-01
```

```
## 45   3 258 Harrison 1997-01
## 46   3 233 Harrison 1997-01
## 47   3 211 Harrison 1997-01
## 48   3 205 Harrison 1997-01
## 49   3 180 Harrison 1997-01
## 50   2 196 Harrison 1997-01
## 51   2 171 Harrison 1997-01
## 52   2 143 Harrison 1997-01
## 53   1 131 Harrison 1997-01
## 54   1  88 Harrison 1997-01
## 55   1  75 Harrison 1997-01
## 56   0  51 Harrison 1997-01
## 57   0  41 Harrison 1997-01
## 58   0  20 Harrison 1997-01
## 59   7 245 Harrison 1997-01
## 60   7 279 Harrison 1997-01
## 61   5 245 Harrison 1997-01
```

```
#Q.6.Display the first and last 5 records from the filtered BullTroutRML2 dat
aset
```

```
head(New_Harrison,n=5)
```

```
##   age  fl     lake     era
## 1  14 459 Harrison 1977-80
## 2  12 449 Harrison 1977-80
## 3  10 471 Harrison 1977-80
## 4  10 446 Harrison 1977-80
## 5   9 400 Harrison 1977-80
```

```
tail(New_Harrison,n=5)
```

```
##    age  fl     lake     era
## 57   0  41 Harrison 1997-01
## 58   0  20 Harrison 1997-01
## 59   7 245 Harrison 1997-01
## 60   7 279 Harrison 1997-01
## 61   5 245 Harrison 1997-01
```

```
#Q.7.Display the structure of the filtered BullTroutRML2dataset
```

```
structure(New_Harrison)
```

```
##    age  fl     lake     era
## 1   14 459 Harrison 1977-80
## 2   12 449 Harrison 1977-80
## 3   10 471 Harrison 1977-80
## 4   10 446 Harrison 1977-80
## 5    9 400 Harrison 1977-80
## 6    9 440 Harrison 1977-80
## 7    9 462 Harrison 1977-80
```

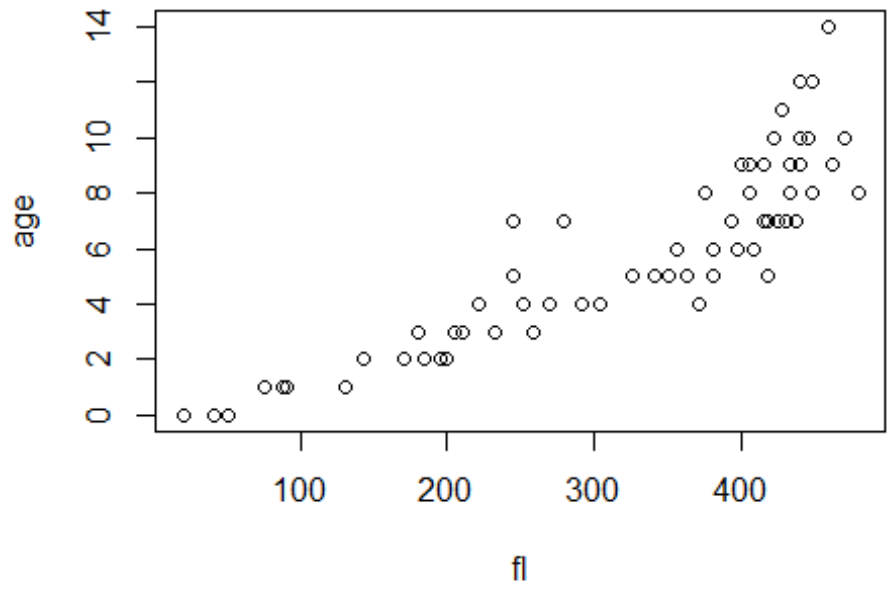
```
## 8    8 480 Harrison 1977-80
## 9    8 449 Harrison 1977-80
## 10   7 437 Harrison 1977-80
## 11   7 431 Harrison 1977-80
## 12   7 425 Harrison 1977-80
## 13   7 419 Harrison 1977-80
## 14   6 409 Harrison 1977-80
## 15   6 397 Harrison 1977-80
## 16   5 419 Harrison 1977-80
## 17   5 381 Harrison 1977-80
## 18   5 363 Harrison 1977-80
## 19   5 351 Harrison 1977-80
## 20   4 372 Harrison 1977-80
## 21   2 199 Harrison 1977-80
## 22   2 184 Harrison 1977-80
## 23   1  91 Harrison 1977-80
## 24  12 440 Harrison 1997-01
## 25  11 428 Harrison 1997-01
## 26  10 440 Harrison 1997-01
## 27  10 422 Harrison 1997-01
## 28   9 434 Harrison 1997-01
## 29   9 415 Harrison 1997-01
## 30   9 406 Harrison 1997-01
## 31   8 434 Harrison 1997-01
## 32   8 406 Harrison 1997-01
## 33   8 375 Harrison 1997-01
## 34   7 415 Harrison 1997-01
## 35   7 394 Harrison 1997-01
## 36   6 381 Harrison 1997-01
## 37   6 357 Harrison 1997-01
## 38   5 341 Harrison 1997-01
## 39   5 326 Harrison 1997-01
## 40   4 304 Harrison 1997-01
## 41   4 292 Harrison 1997-01
## 42   4 270 Harrison 1997-01
## 43   4 252 Harrison 1997-01
## 44   4 221 Harrison 1997-01
## 45   3 258 Harrison 1997-01
## 46   3 233 Harrison 1997-01
## 47   3 211 Harrison 1997-01
## 48   3 205 Harrison 1997-01
## 49   3 180 Harrison 1997-01
## 50   2 196 Harrison 1997-01
## 51   2 171 Harrison 1997-01
## 52   2 143 Harrison 1997-01
## 53   1 131 Harrison 1997-01
## 54   1  88 Harrison 1997-01
## 55   1  75 Harrison 1997-01
## 56   0  51 Harrison 1997-01
## 57   0  41 Harrison 1997-01
```

```
## 58   0  20 Harrison 1997-01
## 59   7 245 Harrison 1997-01
## 60   7 279 Harrison 1997-01
## 61   5 245 Harrison 1997-01
```

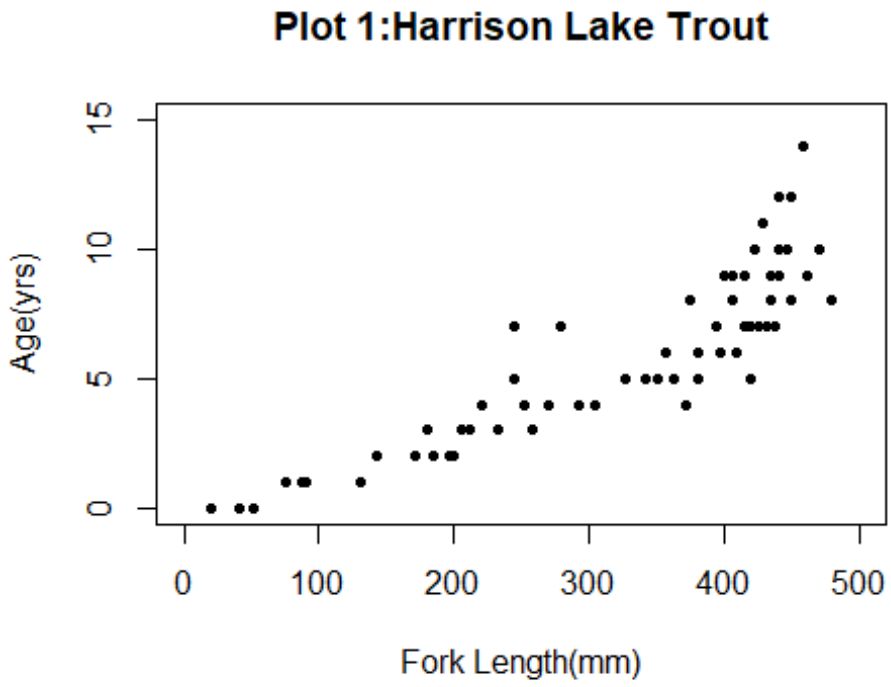
```
#Q.8.Display the summary of the filtered BullTroutRML2dataset

summary(New_Harrison)
```

```
##       age               fl            lake         era
##  Min.   : 0.000   Min.   : 20   Harrison:61   1977-80:23
##  1st Qu.: 3.000   1st Qu.:221   Osprey  : 0   1997-01:38
##  Median : 6.000   Median :372
##  Mean   : 5.754   Mean   :319
##  3rd Qu.: 8.000   3rd Qu.:425
##  Max.   :14.000   Max.   :480
```

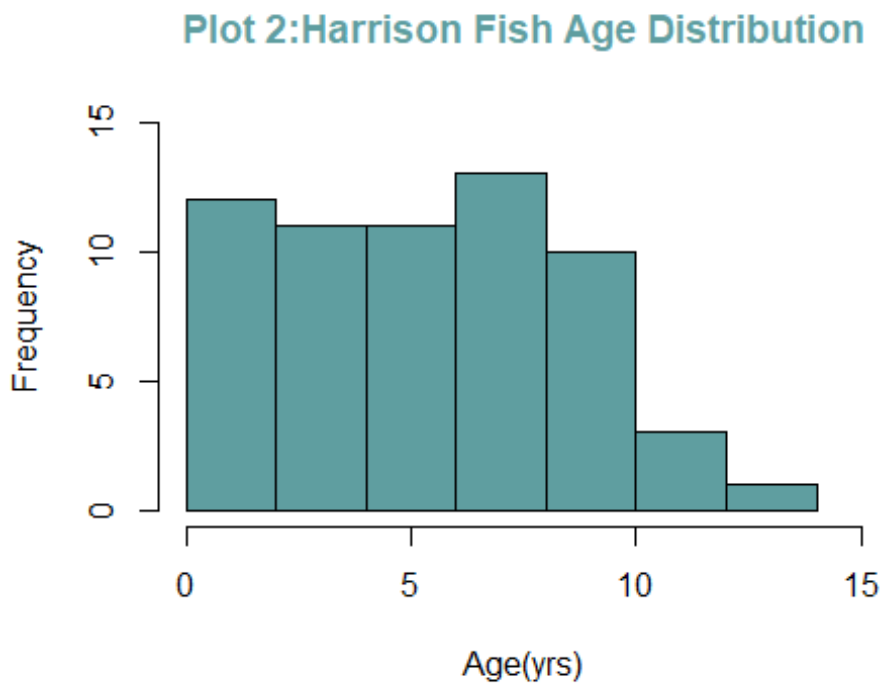
```
#Q.9.Create a scatterplot for "age" (y variable) and "fl" (x variable) with t
he given specifications
#Limit of x axis is (0,500)
#Limit of y axis is (0,15)
#Title of graph is "Plot 1: Harrison Lake Trout
#Y axis label is "Age (yrs)"
#X axis label is "Fork Length (mm)"
#Use a small filled circle for the plotted data points

plot(age~fl,data=New_Harrison)
```



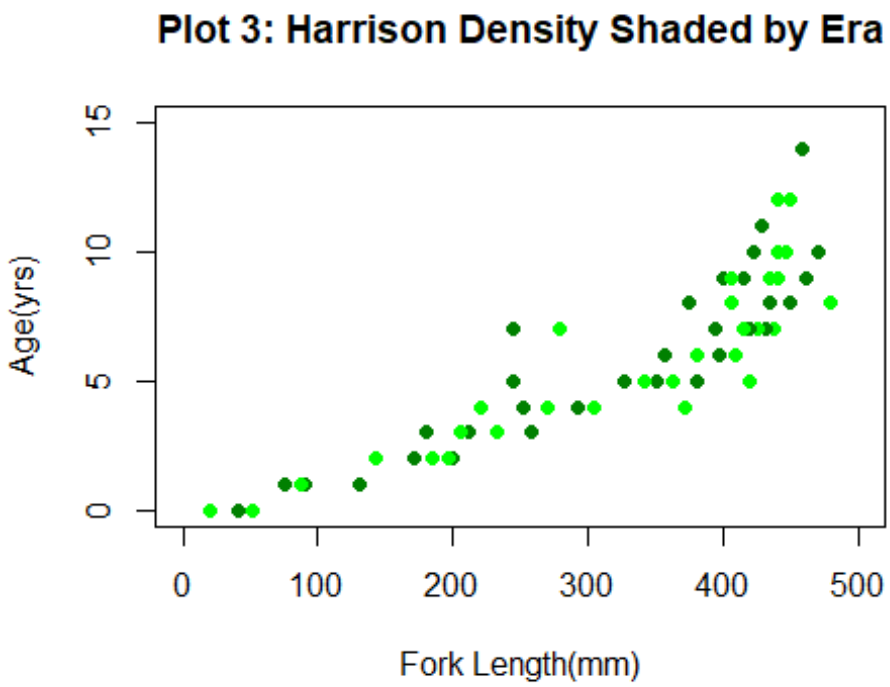
```
ScatterPlot <- plot(age~fl,data=New_Harrison,
                    xlim=c(0,500),ylim=c(0,15),
                    main="Plot 1:Harrison Lake Trout",
                    xlab="Fork Length(mm)",ylab="Age(yrs)",
                    pch=20)
```




```
#Q.10.Plot an "Age" histogram with the given specifications

hist(New_Harrison$age,xlab="Age(yrs)",ylab="Frequency",main="Plot 2:Harrison
Fish Age Distribution",xlim = c(0,15),
     ylim = c(0,15),col="cadetblue",col.main="cadetblue")
```

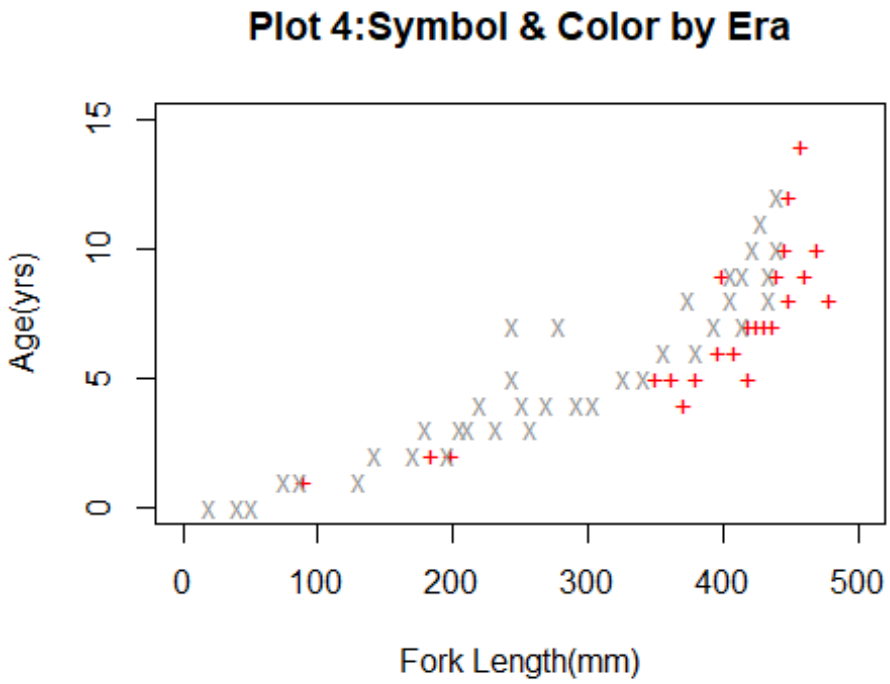



```
fl <- New_Harrison$fl
age <- New_Harrison$age

#Q.11.Create an overdense plot using the same specifications as the previous
scatterplot.

OverdensePlot <- plot(age~fl,
                      xlim=c(0,500),ylim=c(0,15),
                      main="Plot 3: Harrison Density Shaded by Era",
                      xlab="Fork Length(mm)",ylab="Age(yrs)",
                      pch=19,col=rgb(0,(1:2)/2,0))
```

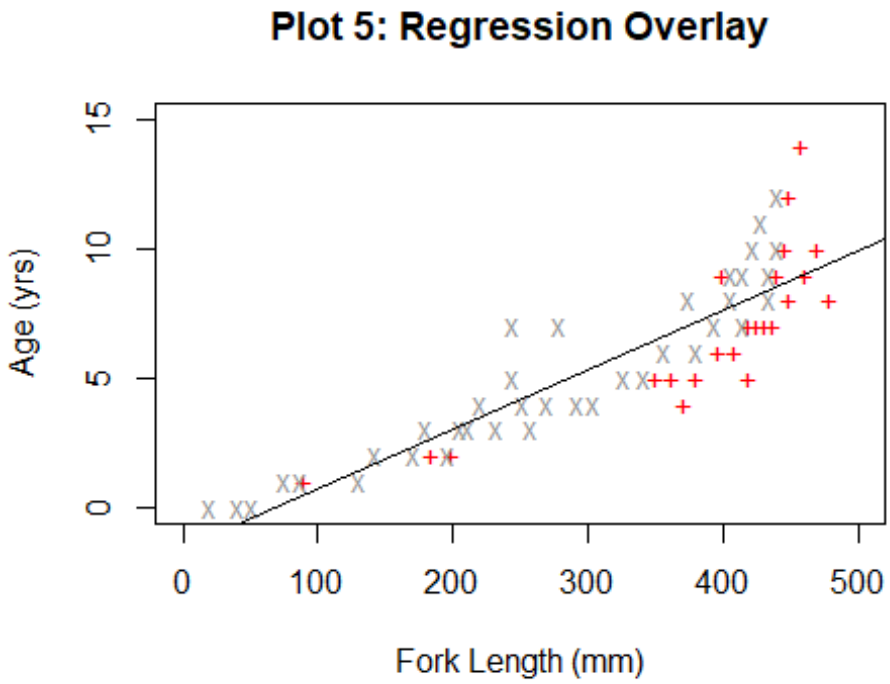



```
#Q.12. Create a new object called "tmp" that includes the first 3 and last 3
records of the BullTroutRML2 data set.

tmp <- headtail(BullTroutRML2,n=3)
tmp

##    age  fl     lake     era
## 1   14 459 Harrison 1977-80
## 2   12 449 Harrison 1977-80
## 3   10 471 Harrison 1977-80
## 94   4 298   Osprey 1997-01
## 95   3 279   Osprey 1997-01
## 96   3 273   Osprey 1997-01

#Q.13.Display the "era" column (variable) in the new "tmp" object

tmp$era

## [1] 1977-80 1977-80 1977-80 1997-01 1997-01 1997-01
## Levels: 1977-80 1997-01

#Q.14. Create a pchs vector with the argument values for + and x.
pchs <- c('+','x')
pchs

## [1] "+" "x"
```

```r
#Q.15.Create a cols vector with the two elements "red" and "gray60"

cols <- c("red","gray60")
cols

## [1] "red"    "gray60"

#Q.16.Convert the tmp era values to numeric values.

tmp$era

## [1] 1977-80 1977-80 1977-80 1997-01 1997-01 1997-01
## Levels: 1977-80 1997-01

as.numeric(tmp$era)

## [1] 1 1 1 2 2 2

#Q.17.Initialize the cols vector with the tmp era values

tmp$era <- cols

tmp$era

## [1] "red"    "gray60" "red"    "gray60" "red"    "gray60"

#Q.18.Create a plot of "Age (yrs)" (y variable) versus "Fork Length (mm)" (x
variable) with the given specifications:
#Title of graph is "Plot 4: Symbol & Color by Era"
#Limit of x axis is (0,500)
#Limit of y axis is (0,15)
#X axis label is "Age (yrs)"
#Y axis label is "Fork Length (mm)"
#Set pch equal to pchs era values
#Set col equal to cols era values

plot(age~fl,data=New_Harrison,
     xlim=c(0,500),ylim=c(0,15),
     main="Plot 4:Symbol & Color by Era",
     xlab="Fork Length(mm)",ylab="Age(yrs)",
     pch=pchs[New_Harrison$era],col=cols[New_Harrison$era])
```

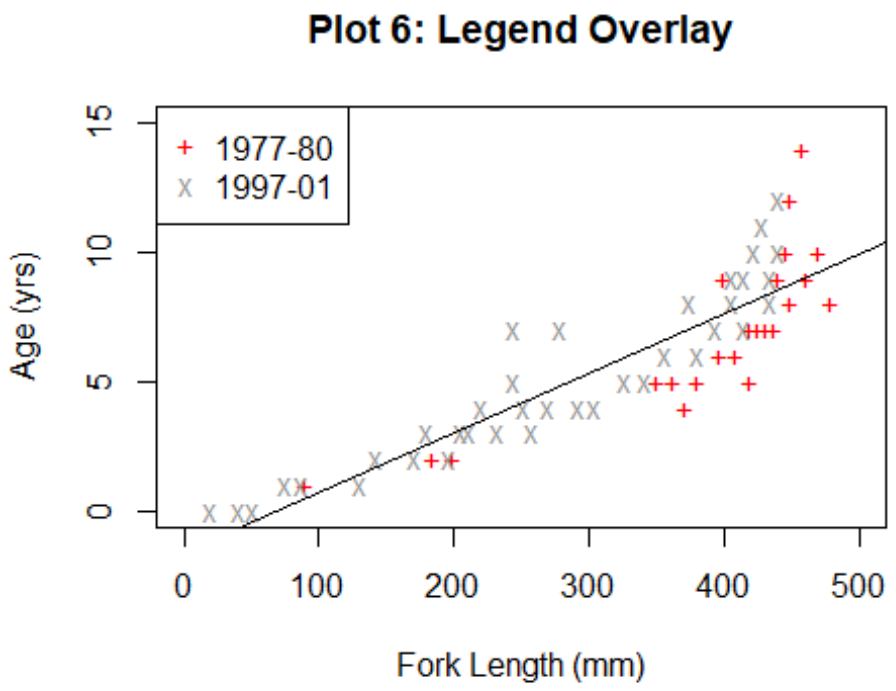



```
#Q.19.Plot a regression line overlay on Plot 4 and title the new graph "Plot
5: Regression Overlay".

plot(age~fl, data= New_Harrison,
     main = "Plot 5: Regression Overlay",
     xlab ="Fork Length (mm)", ylab = "Age (yrs)",
     xlim = c(0, 500), ylim = c(0,15),
     pch=pchs[New_Harrison$era] , col=cols[New_Harrison$era])

R <- lm(age~fl,data = New_Harrison)
abline(R)
```




```
#Q.20.Place a legend of on Plot 5 and call the new graph "Plot 6: :Legend Ove
rlay"

plot(age~fl, data= New_Harrison, main = "Plot 6: Legend Overlay",
     xlab ="Fork Length (mm)", ylab = "Age (yrs)",
     xlim = c(0,500),ylim = c(0,15),
     pch=pchs[New_Harrison$era] ,
     col=cols[New_Harrison$era])

a <- lm(age~fl,data = New_Harrison)

abline(a)

?legend

## starting httpd help server ...

##  done

New_Harrison$era

##  [1] 1977-80 1977-80 1977-80 1977-80 1977-80 1977-80 1977-80 1977-80 1977-
80
## [10] 1977-80 1977-80 1977-80 1977-80 1977-80 1977-80 1977-80 1977-80 1977-
80
## [19] 1977-80 1977-80 1977-80 1977-80 1977-80 1997-01 1997-01 1997-01 1997-
01
```

```
## [28] 1997-01 1997-01 1997-01 1997-01 1997-01 1997-01 1997-01 1997-01 1997-
01
## [37] 1997-01 1997-01 1997-01 1997-01 1997-01 1997-01 1997-01 1997-01 1997-
01
## [46] 1997-01 1997-01 1997-01 1997-01 1997-01 1997-01 1997-01 1997-01 1997-
01
## [55] 1997-01 1997-01 1997-01 1997-01 1997-01 1997-01 1997-01
## Levels: 1977-80 1997-01

legend("topleft",legend = c("1977-80","1997-01"), pch = pchs, col = cols)
```




```
#Q.21.Commit your code in your github/gitlab repo
```

https://github.com/PoonamDighe/ALY6000_MODULE2.git