Module 2 Project

Name: Parth Sawant

Course Name: ALY6000

Due Date: 29/01/2022

Title: Executive summary report 2

Professor: Mohammad Shafiqul Islam

NUID: 002123792

INTRODUCTION

This is a project which has been demonstrated in R and R studio, getting some hands-on experience for the data set provided. How to install and load some new libraries. We get to analyze the structure and summary of the filtered dataset, that is the overview of the dataset. We get to know some different types of graphs namely scatterplot, histogram, over dense plot. In addition to this we also learn about regression line and legend.

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 shots to support your observations and conclusions. Below is a sample excerpt of an analysis of Harrison Lake fish from the BullTroutRML2 Dataset. 1]Code &Output: -

1]Code &Output: -

```
> str(df)
'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 ...
>
```

Explanation: -

>>After applying the filter function on the original dataset that is BullTroutRml2, it can be seen that there are 61 observations and 4 variables namely age, fork length, lake, era.

>> The age of the lake is of integer data type.

>> fl is the fork length of the lake which is also of integer data type.

>> There are two types of lakes namely Harrison and Osprey which are of factor data types, having 2 levels.

>> There are 2 era of the lakes as seen “1977-80”, “1997-01” which are also of factor data type

2]Code & Output: -

```
> summary(df)
      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
>
```

Explanation: -

>> This is the summary of the new dataset after filtering out the osprey lake which is why it is been seen 0.

>>There are in total 61 observations for Harrison Lake.

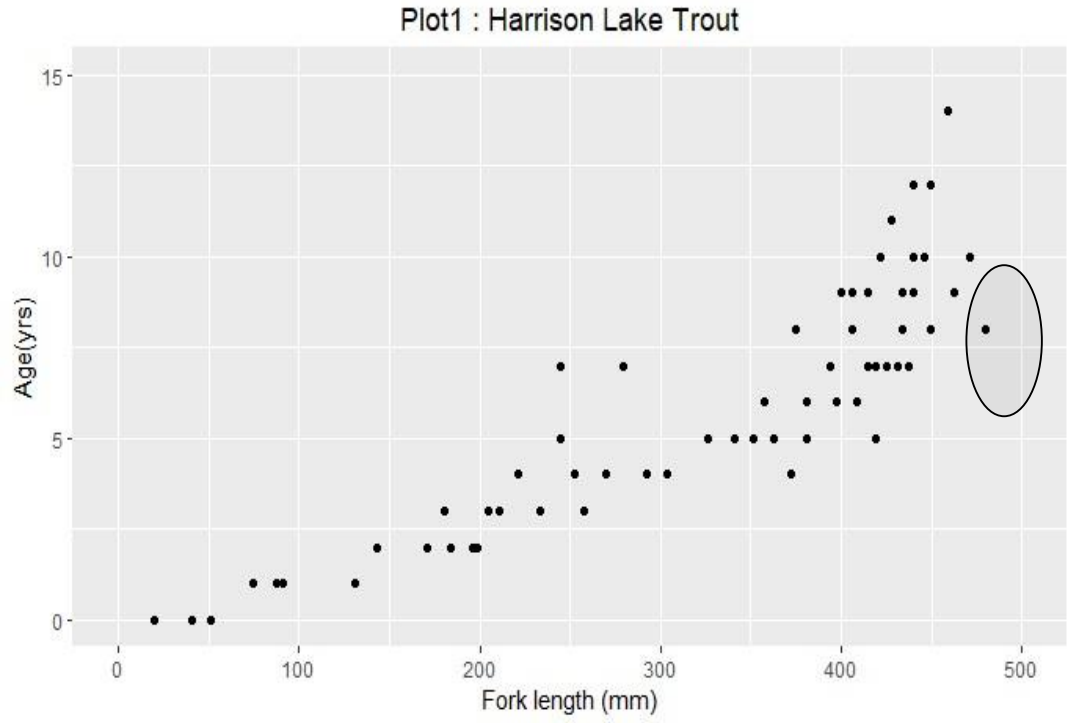
>> The fork length has a minimum length of 20mm and a maximum of 480mm.

>>The age of Harrison Lake has a median of 6 years.

>> The are 2 era's "1977-80" and "1997-01" with 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, bar plots (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.

1]



Explanation: -

>> This is a scatter plot of the filtered dataset of Harrison Lake

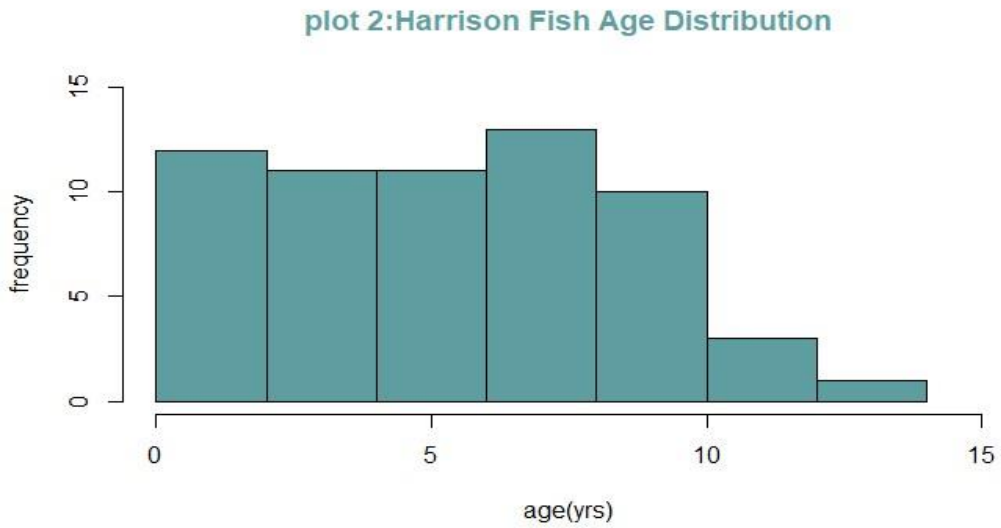
>> The X axis represents the fork length of the lake while the Y axis represents the age of the lake.

>> It can be depicted that the length of the lake is gradually increasing over the years.

>>Between the years 6-7, there were times when the lake was shrunk.

>> Approximately at the age of 8, the lake reaches a maximum of 480mm highlighted by a black circle.

[2]



Explanation: -

>> The above histogram shows the frequency of fish over the years.

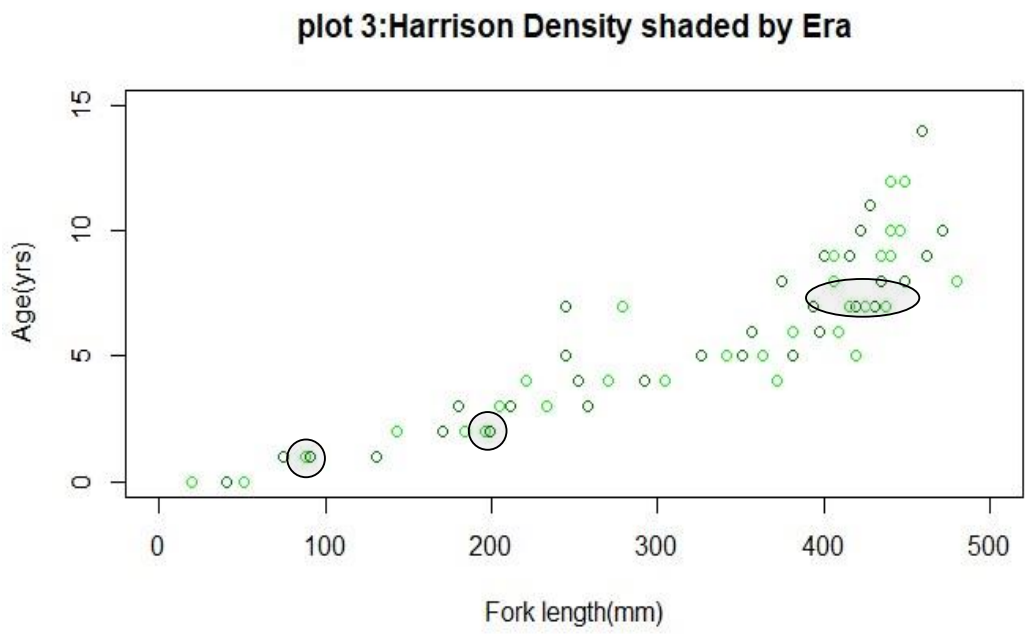
>>There are around 12 fishes in the first 2 years.

>>From 2-6 years there were constant no of fishes that is 11.

>>From the age 6-8 years, largest number of fishes are observed.

>>There was extreme fall in number of fishes from the years 10-15, with approximately 1-2 fish in the last year.

[3]



Explanation: -

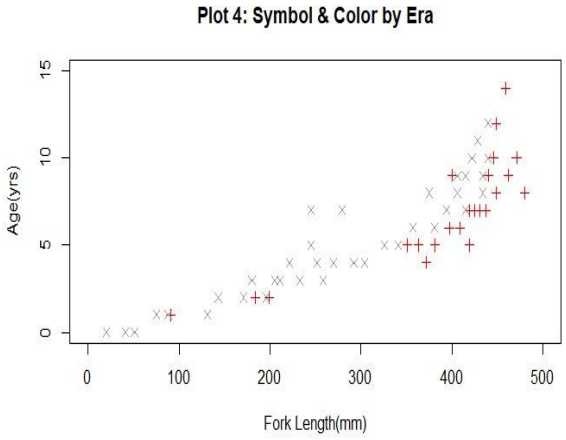
>>This is an over dense plot of the filtered dataset where fork length and age of the lake is related.

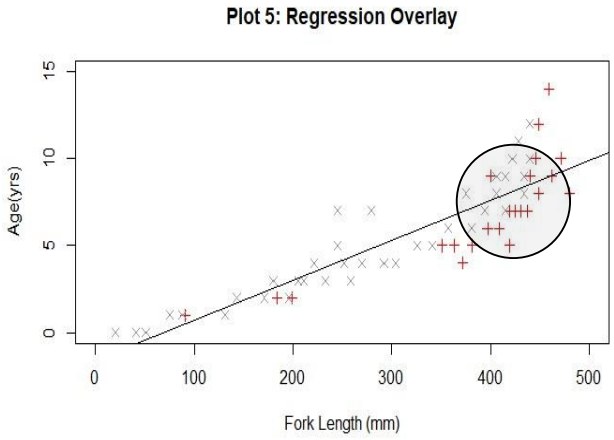
>> Since there two levels of era, we have used two shades of green to compare them through a graph.

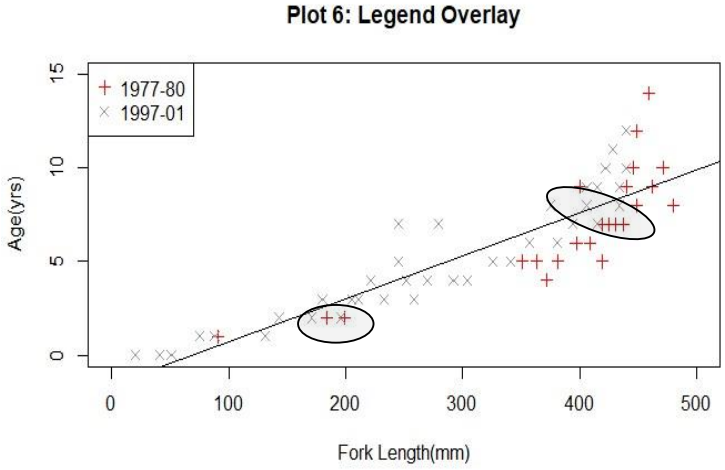
>> The maximum fork length was found to be 480mm and minimum to be 20mm.

>> The black circles show the intersection of the two levels of era.

[4]







Explanation: -

>> All the three graphs show the age and Fork length of Harrison Lake over an era of 1977-80 and 1997-01

>> The red + sign is for the time period of 1977-80 whereas 1977-01 is shown through grey x sign.

>>There are some intersections of era's as observed in the graph highlighted through grey color.

>> It can be observed that some points are above the regression line and some below it, few are on the line, from this we can say that it is pretty much accurate except one + point which is very far from the regression line.

C. Summary: -

In this report we have used BullTroutRML 2 dataset from the FSA package which was already available in it. Using the filter function, we filtered the dataset taking into consideration only the Harrison Lake. It had 4 variables age, fl, lake and era. From all the graphs taking into consideration “era”, it can be understood that as the age of the lakes is increased, the length of the lake is also increased. Since all the points were close to the regression line the dataset was pretty much accurate. Key points of every graph are explained below it.

D. Bibliography

[1] Filter: Return rows with matching conditions (R documentation)

Source: https://www.rdocumentation.org/packages/dplyr/versions/0.7.8/topics/filter

Last Accessed: 28th January,2022

[2] Scatter plot in R using ggplot2(guru99, December 2021)

Source- https://www.guru99.com/r-scatter-plot-ggplot2.html

Last Accessed: 28th January,2022

[3] ggplot – centre plot title with total width of image. (Stack overflow October,2019)

Source- https://stackoverflow.com/questions/58545686/ggplot-centre-plot-title-with-total-width-ofimage

Last Accessed: 28th January,2022

[4] R histograms(tutorialspoint)

Source- https://www.tutorialspoint.com/r/r_histograms.htm

Last Accessed: 28th January,2022

[5] How to create a scatterplot with a regression line in R(geeksforgeeks February ,2021)

Source- https://www.geeksforgeeks.org/how-to-create-a-scatterplot-with-a-regression-line-in-r/

Last Accessed: 28th January,2022

My Github repository link:

https://github.com/parthh03/ALY6000_module2.git

Appendix

module2.R

Parth

2022-01-28

```
#1. print "Plotting Basics:Lastname"
print("Plotting Basics: Sawant")
```

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

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

#2.Installing and importing required packages
install.packages('FSA')
```

```
## Installing package into 'C:/Users/Parth/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\Parth\AppData\Local\Temp\Rtmpa6C3ag\downloaded_packages
```

```
install.packages('FSAdata')
```

```
## Installing package into 'C:/Users/Parth/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\Parth\AppData\Local\Temp\Rtmpa6C3ag\downloaded_packages
```

```
install.packages('magrittr')
```

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

```
##
##   There is a binary version available but the source version is later:
##          binary source needs_compilation
## magrittr  2.0.1  2.0.2              TRUE
##
##   Binaries will be installed
## 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\Parth\Documents\R\win-library\4.1\00LOCK\magrittr\libs\x64\magrittr
.dll
## to C:\Users\Parth\Documents\R\win-library\4.1\magrittr\libs\x64\magrittr.d
ll:
## Permission denied

## Warning: restored 'magrittr'

##
## The downloaded binary packages are in
##  C:\Users\Parth\AppData\Local\Temp\Rtmpa6C3ag\downloaded_packages

install.packages('dplyr')

## Installing package into 'C:/Users/Parth/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\Parth\Documents\R\win-library\4.1\00LOCK\dplyr\libs\x64\dplyr.dll t
o C:
## \Users\Parth\Documents\R\win-library\4.1\dplyr\libs\x64\dplyr.dll: Permiss
ion
## denied

## Warning: restored 'dplyr'

##
## The downloaded binary packages are in
##  C:\Users\Parth\AppData\Local\Temp\Rtmpa6C3ag\downloaded_packages

install.packages('plotrix')

## Installing package into 'C:/Users/Parth/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\Parth\AppData\Local\Temp\Rtmpa6C3ag\downloaded_packages

install.packages('ggplot2')

## Installing package into 'C:/Users/Parth/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\Parth\AppData\Local\Temp\Rtmpa6C3ag\downloaded_packages

install.packages('moments')

## Installing package into 'C:/Users/Parth/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\Parth\AppData\Local\Temp\Rtmpa6C3ag\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)

#3.Loading the data
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

head(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
```

```r
#4. Print the first and last 3 records from the BullTroutRML2 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
```

```r
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
#5.Remove all records except those from Harisson lake
df <- filter(BullTroutRML2, lake=="Harrison")
df
```

```
##    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
```

```
#6. Display the first and last 5 records from the filtered BullTroutRML2 Data
set
head(df,n=3)
```

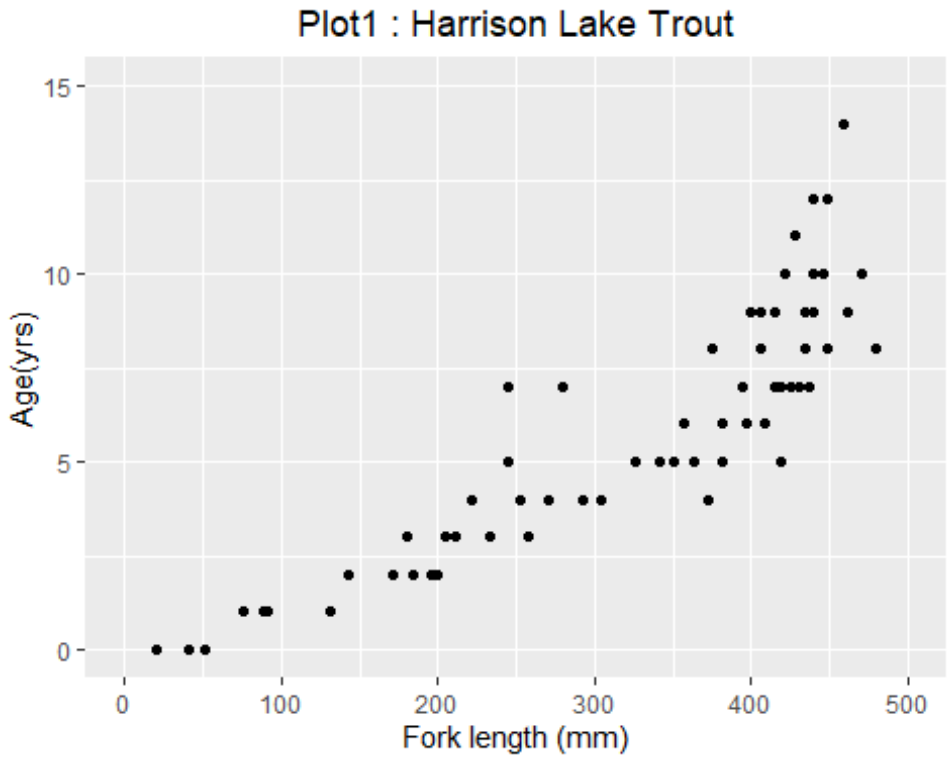
```
##   age  fl     lake     era
## 1  14 459 Harrison 1977-80
## 2  12 449 Harrison 1977-80
## 3  10 471 Harrison 1977-80
```

```
tail(df,n=3)
```

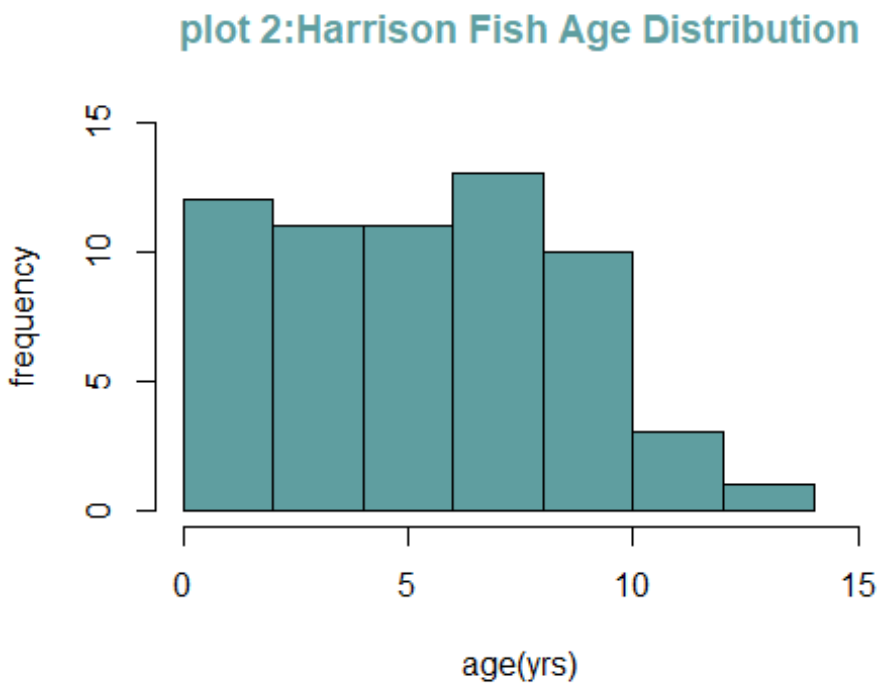
```
##    age  fl     lake     era
## 59   7 245 Harrison 1997-01
## 60   7 279 Harrison 1997-01
## 61   5 245 Harrison 1997-01
```

```
#7. Display the structure of the filtered BullTroutRML2 Dataset
str(df)
```

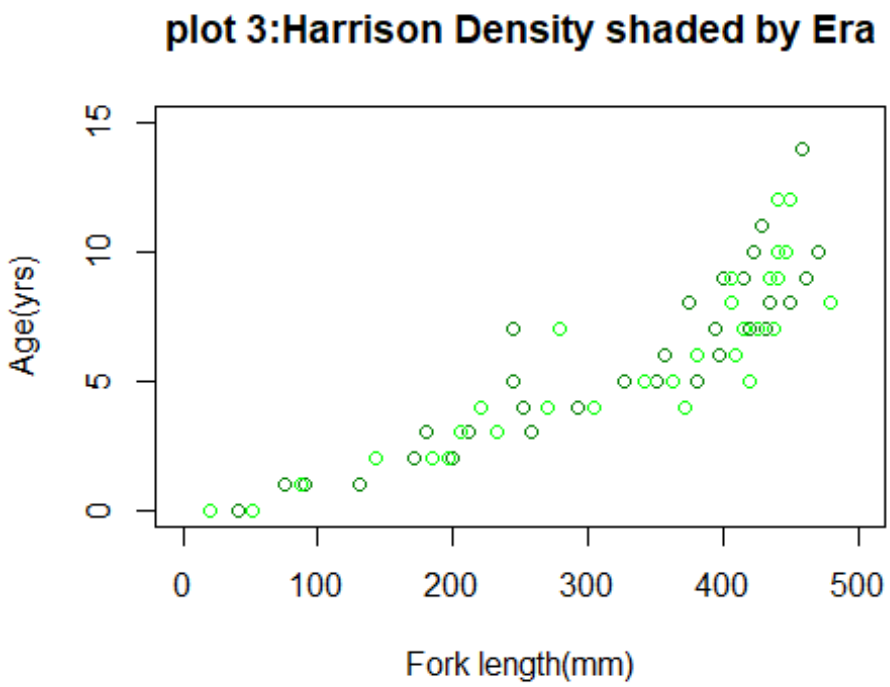
```
## 'data.frame':    61 obs. of  4 variables:
##  $ age : int  14 12 10 10 9 9 9 8 8 7 ...
##  $ fl  : int  459 449 471 446 400 440 462 480 449 437 ...
##  $ lake: Factor w/ 2 levels "Harrison","Osprey": 1 1 1 1 1 1 1 1 1 1 ...
##  $ era : Factor w/ 2 levels "1977-80","1997-01": 1 1 1 1 1 1 1 1 1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

#8. Display the summary of the filtered BullTroutRML2 Dataset
summary(df)

##       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

#9. create a scatterplot for "age" (y variable) and "fl"(x variable) with the
following specifications
# limit of x axix is (0,500)
# limit of y axix 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

scatterplot <- ggplot(df, aes(x=fl,y=age))+geom_point()+xlim(0,500)+ylim(0,15
)+
labs(title="Plot1 : Harrison Lake Trout", x="Fork length (mm)", y="Age(yrs)")
+
theme(plot.title = element_text(hjust=0.5))
scatterplot
```




```
#10. Plot an "Age histogram with the following specifications
# y axis label is "frequency"
# x axis label is "age (yrs)"
# title of the histogram is "plot 2: Harrison Fish Age Distribution"
# X and Y limits is 0,15
# The color of the frequency plots is "cadetblue"
# The color of the title is "cadetblue"

hist(df$age,xlab = "age(yrs)", ylab="frequency",
main = "plot 2:Harrison Fish Age Distribution", xlim=c(0,15), ylim=c(0,15), c
ol="cadetblue",
col.main="cadetblue")
```

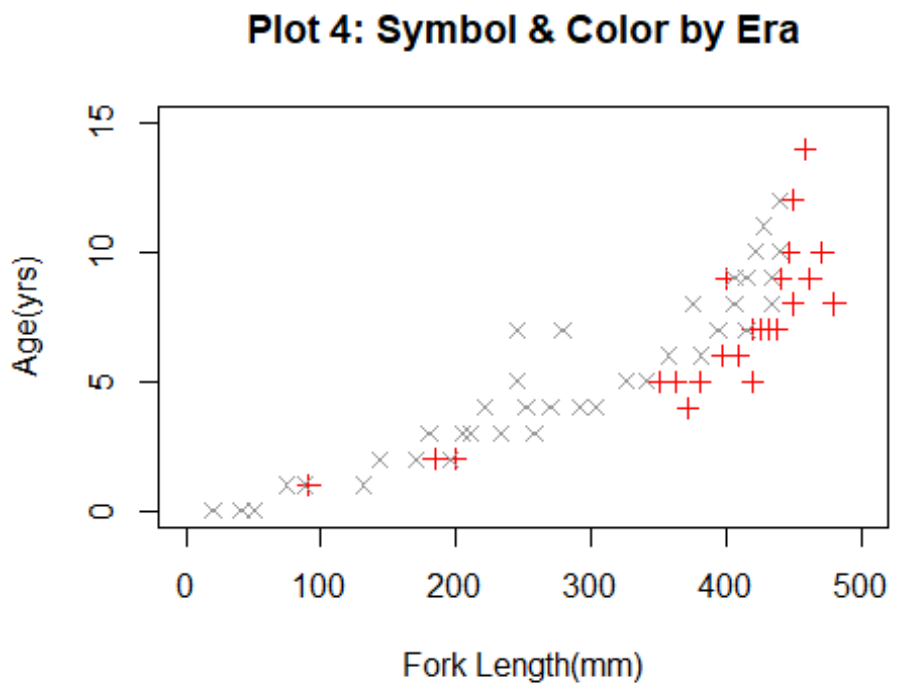



```
#11. create an overdense plot using the same specifications as the previous s
catterplot. But,
# Title the plot "plot 3:Harrison Density shaded by Era"
# y axis label is "Age(yrs)
# y axis limits are 0 to 15
# x axis label is "Fork length (mm)"
# x axis limits are 0 to 500
# include two levels of shading for the "green" data points
# plot solid circles as data points

fl <- df$fl
age <- df$age
overdense_plot <- plot(age~fl, main = "plot 3:Harrison Density shaded by Era"
,
xlab="Fork length(mm)", ylab="Age(yrs)", xlim=c(0,500), ylim=c(0,15), col=rgb
(0,(1:2)/2,0))
```



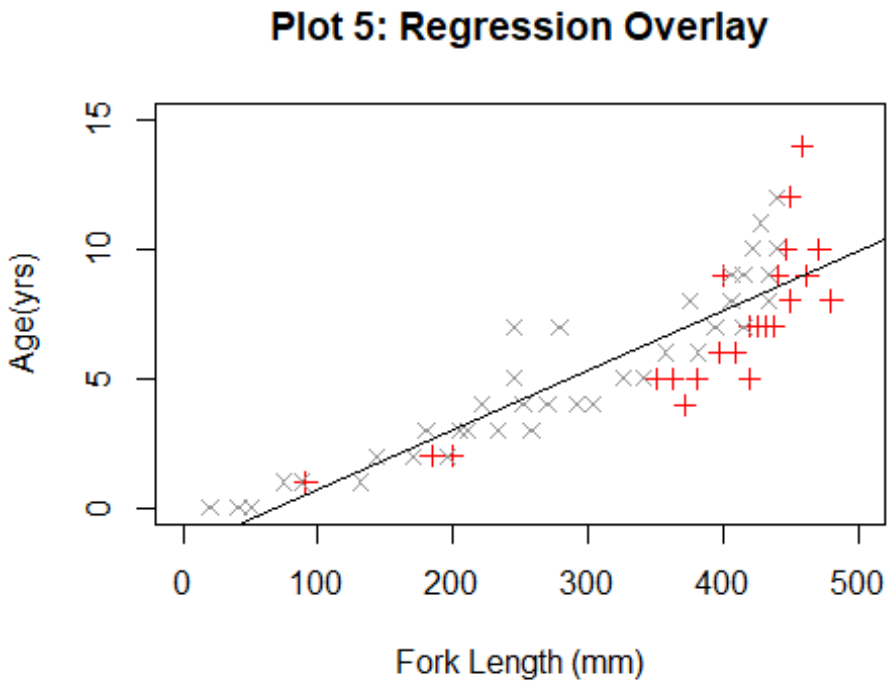

```
#12. create a new object called "temp"that includes the first 3 and last 3 re
cords of the BulltroutRML2 data set

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

##    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

#13. Display the "era" column (variable) in the new "temp" object
temp$era

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

#14. create a pchs vector with the argument values for + and x
pchs <- c(3,4)

#15. create a cols vector with the two elements "red" and "gray60"
cols <- c("red","grey60")
cols

## [1] "red"    "grey60"
```

```
#16. convert the temp era values to numeric values
converted_temp <- as.numeric(temp$era)
converted_temp

## [1] 1 1 1 2 2 2

#17. Initialize the cols vector with the temp era values
cols[temp$era]

## [1] "red"    "red"    "red"    "grey60" "grey60" "grey60"

#18.  Create a plot of "Age (yrs)" (y variable) versus "Fork Length (mm)" (x
variable) with the following 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=df, main="Plot 4: Symbol & Color by Era",xlim = c(0,500),
ylim=c(0,15),xlab="Fork Length(mm)", ylab="Age(yrs)",
pch=pchs[df$era], col=cols[df$era])
```



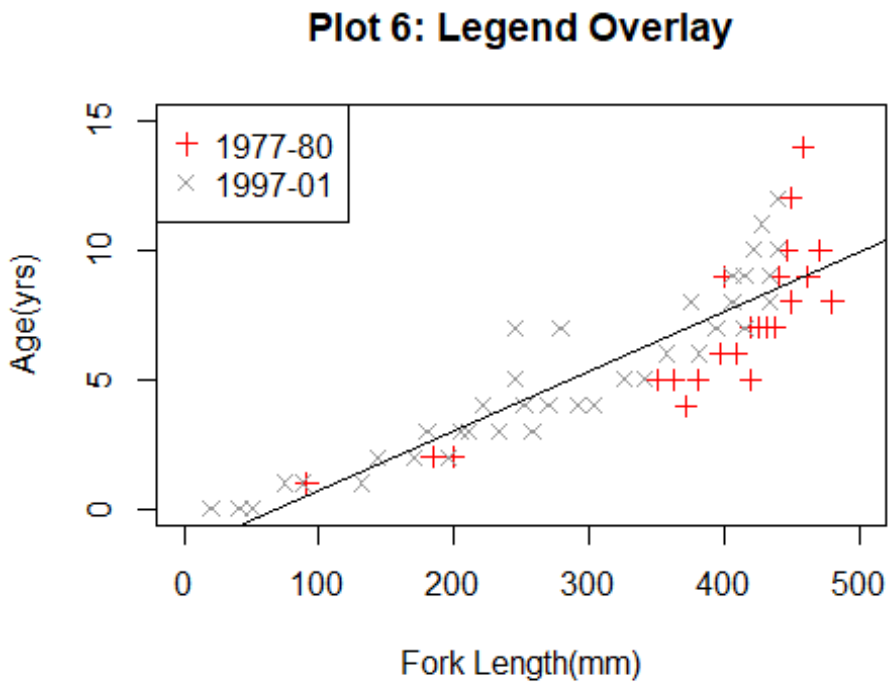

```
#19. Plot a regression line overlay on Plot 4 and title the new graph "Plot 5
: Regression Overlay".
```

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




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

df$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

plot(age~fl,data=df, main="Plot 6: Legend Overlay",xlim = c(0,500),
ylim=c(0,15), xlab="Fork Length(mm)", ylab="Age(yrs)",
```

```
pch=pchs[df$era], col=cols[df$era])
abline(lm(age~fl,data=df))
legend("topleft", legend=c("1977-80","1997-01"),pch = pchs,col = cols)
```




```
install.packages('tinytex')

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

## package 'tinytex' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##  C:\Users\Parth\AppData\Local\Temp\Rtmpa6C3ag\downloaded_packages

tinytex::install_tinytex()
```