Biostatistics: Theory and Applications in R (Virtual)

Week3_Session2_R_training3

Contents

Prof Dr Mohammed Abu Sayed Arfin Khan
Department of Forestry and Environmental Science
Shahjalal University of Science and Technology, Sylhet
+8801917174537, khan-for@sust.edu, nobelarfin@yahoo.com
Homepage | Google scholar | Researchgate | ORCID | Publons | BayCEER

#Creating a subset
#Set the working directory- getwd()/ setwd("Y:/")
getwd()
setwd("C:/Users/Fahmida Sultana/Desktop/R training/R_training_Class_03")

#install openxlsx package or xlsx package
library(openxlsx)
library(readxl)

#####import data set from xlsx file
study_1 <- read.xlsx("Tree_height.xlsx")
str(study_1)

study_2 <- read_excel("Tree_height.xlsx", sheet = "study_2")
str(study_2)

###Import csv file
study_3 <- read.csv("C:/Users/Fahmida Sultana/Desktop/R training/R_training_Class_03/Tree_height1.csv")
str(study_3)

#read the data file
str(study_1)   #to see the data frame
unique(study_1$year)   # to see the the all year
range(study_1$year)   # to see the range of the year column
sort(study_1$year)   # to see the all input in year column
#study_1[1,]   # see the 1st row
study_1[,5]   # see the 5th column

 # create a subset from the main file study_1 with the year > 2007
unique(study_1$year)   # year- 2005 2006 2007 2008 2009 2010
A <- subset(study_1, study_1$year > 2007)
unique(A$year)   # check the data, all the year 2005, 2006, 2007 are not included in the subset
unique(study_1$year)# but all year are abailable in the main file

 # create a subset from the main file study_1 with the year > 2005 & year <2010
A2 <- subset(study_1, study_1$year > 2005 & study_1$year < 2010)

```
unique(A2$year)
```

#Task
 # please create a subset named A3 from the main file study_1 with the year <2009

 # create a subset from the main file study_1 with the control treatment
 # selecting a treatment
```
B<- subset(study_1, study_1$treatment== "Control") # selecting a treatment
unique(B$treatment)
```

 # delete a treatment
```
C<- subset(study_1, study_1$treatment!= "Control") # deleting a treatment
unique(C$treatment)
```

 # delete multiple treatments
```
D<- subset(study_1, study_1$treatment!= "Control" & study_1$treatment!= "Drought")
unique(D$treatment)
str(D)
```

#Task
#create a subset named E by deleting the "Drought" and "Rain" treatments

#Create scatter plot with trend line (lm)
```
library(plotrix)
??plotrix
study_2 <- read_excel("Tree_height.xlsx", sheet = "study_2")
str(study_2)

par(mfrow=c(1,1))
plot(study_2$Elevation_m, study_2$Diameter_cm, pch=9, col="blue")
abline(lm((Diameter_cm) ~ (Elevation_m), data =study_2 ))

plot(study_2$Height_m, study_2$Diameter_cm, pch=16, col="red")
abline(lm((Height_m) ~ (Elevation_m), data =study_2))
```

 # check the available symbols and colors for R figures
```
?pch
?colors

demo("colors")
Return    #Hit <Return> to start
```

Return   #Hit <Return> to see next plot: Return

#par() function
#Subsequent figures will be drawn in
#by rows (mfrow) or columns (mfcol) respectively.

#plot 2 figure in one row

```
par(mfrow=c(1,2)) # back to normal
plot(study_2$Elevation_m, study_2$Diameter_cm, pch=9, col="blue")
plot(study_2$Height_m, study_2$Diameter_cm, pch=16, col="red")
```

#plot 2 figure in one column
```
par(mfrow=c(2,1))
plot(study_2$Elevation_m, study_2$Diameter_cm, pch=9, col="blue")
plot(study_2$Height_m, study_2$Diameter_cm, pch=16, col="red")
```


#Create barplot
```
library(sciplot)
str(study_2)
par(mfrow = c(1,1))
bargraph.CI(x.factor = study_area, response=species_richness, ylim= c(0,50), ylab = "No of Species", xlab
= "Study area", las = 1, data =study_2,col = c("black", "red", "green"), main = "p=0.05")
```


#Create barplot by group
```
str(study_1)
bargraph.CI(x.factor = treatment, group= plant, response=biomass, ylim= c(0,200), ylab = "Biomass",
xlab = "Treatment", las = 1, data =study_1,legend = T, col = c("red", "green"), main = "p=0.05")
```

#Create line plot
```
str(study_1)

par(mfrow=c(1,1))
lineplot.CI((year), biomass,ylim=c(50,300), xlab="Year",ylab="Biomass",x.cont=T,legend=T,
data=study_1, main = "")
```

#Create line plot by group

par(mfrow=c(1,1))
lineplot.CI((year), biomass, group=treatment,ylim=c(50,300),
xlab="Year",ylab="Biomass",x.cont=T,legend=T, data=study_1, main = "")

#Week3_Session2_R Task

To do
#############
 # Use the study_2 data set to create the following scatter plot plot:
 # * plot diameter_cm against Height_m with filled red circles
 # * Add a background color greeen to a plot
 # * increase the size of the circle by 3
 # * change the text of the axislabels to include what was measured plus the units of measurement
 # * give the plot the heading "My Own Plot"
 # * export the plot to a PDF file

#######################Hints#####################

#library(plotrix), library
#pch=16, filled circle
#col="red", color
#cex= 2, size
#plot_bg(col="green"), background
#xlab="Height (m)", x axis label
#ylab="Diameter (cm)", y axis label
#main = "My Own Plot"), plot name