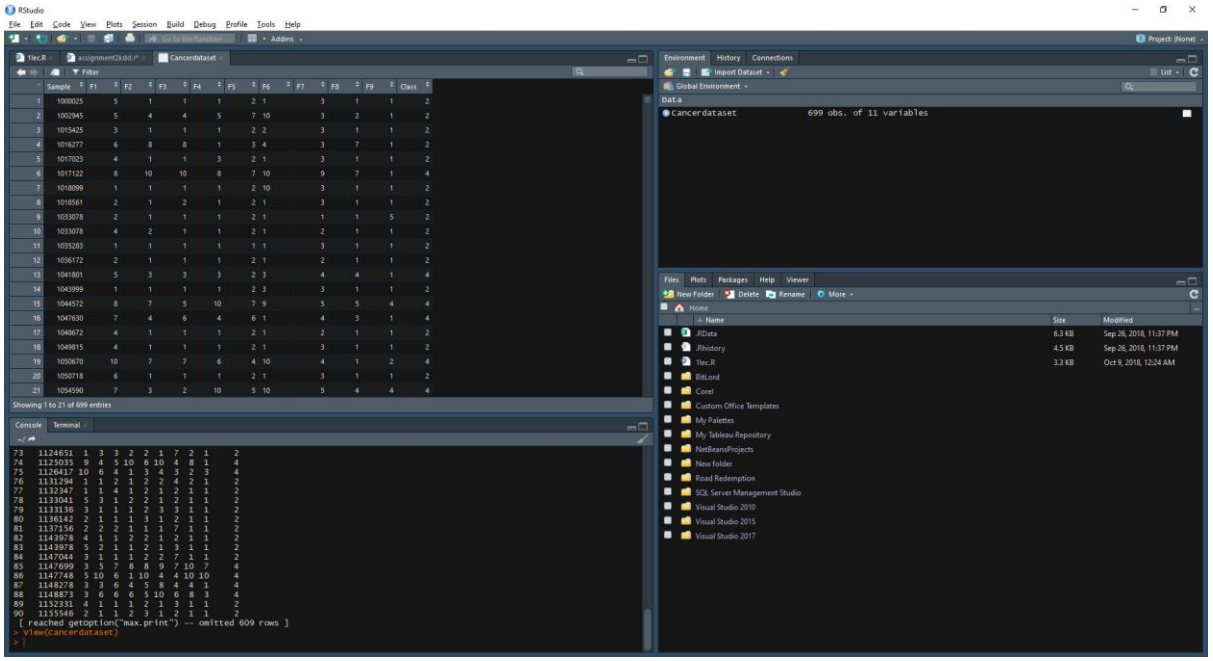
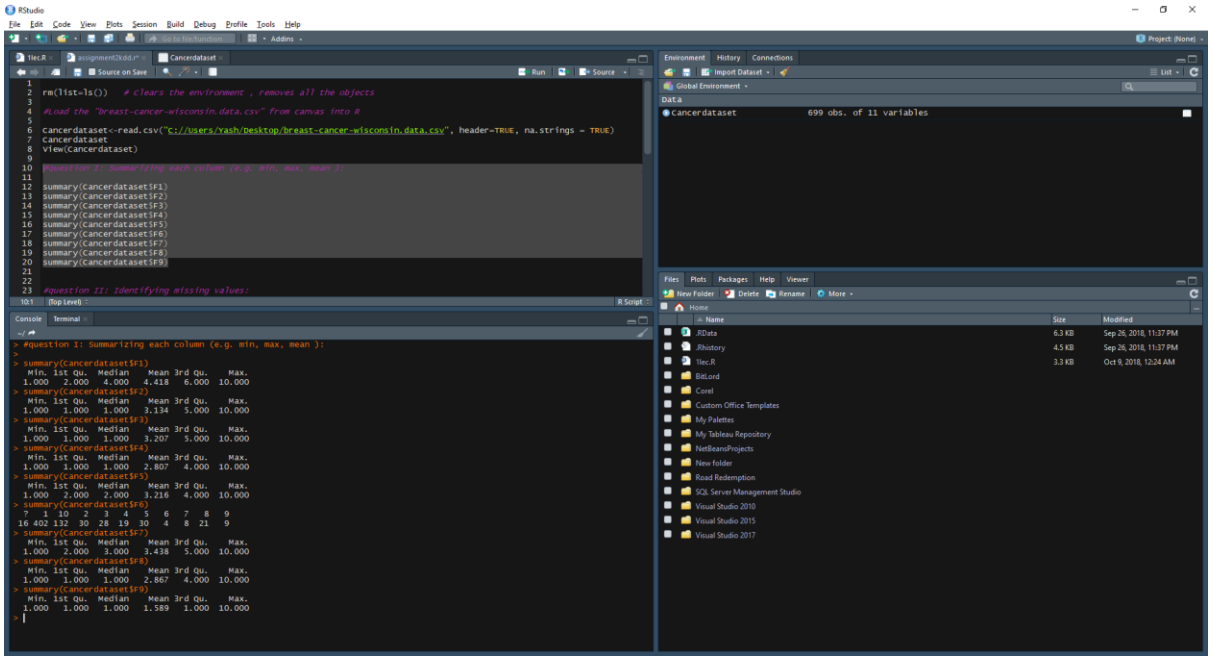
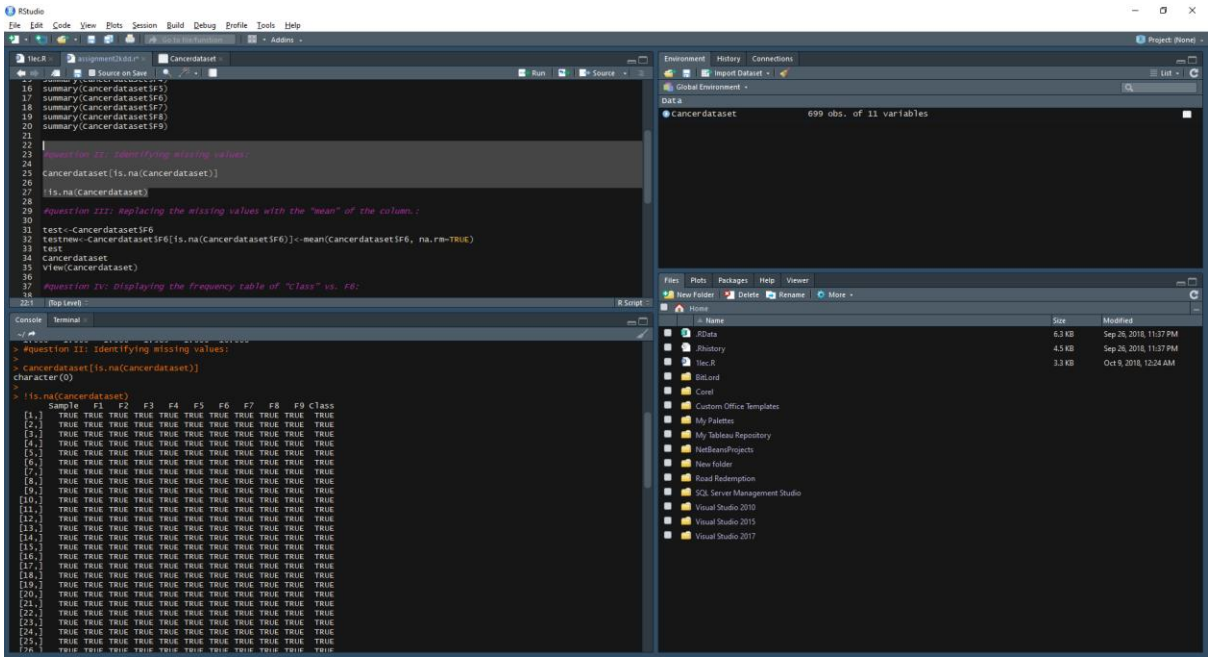
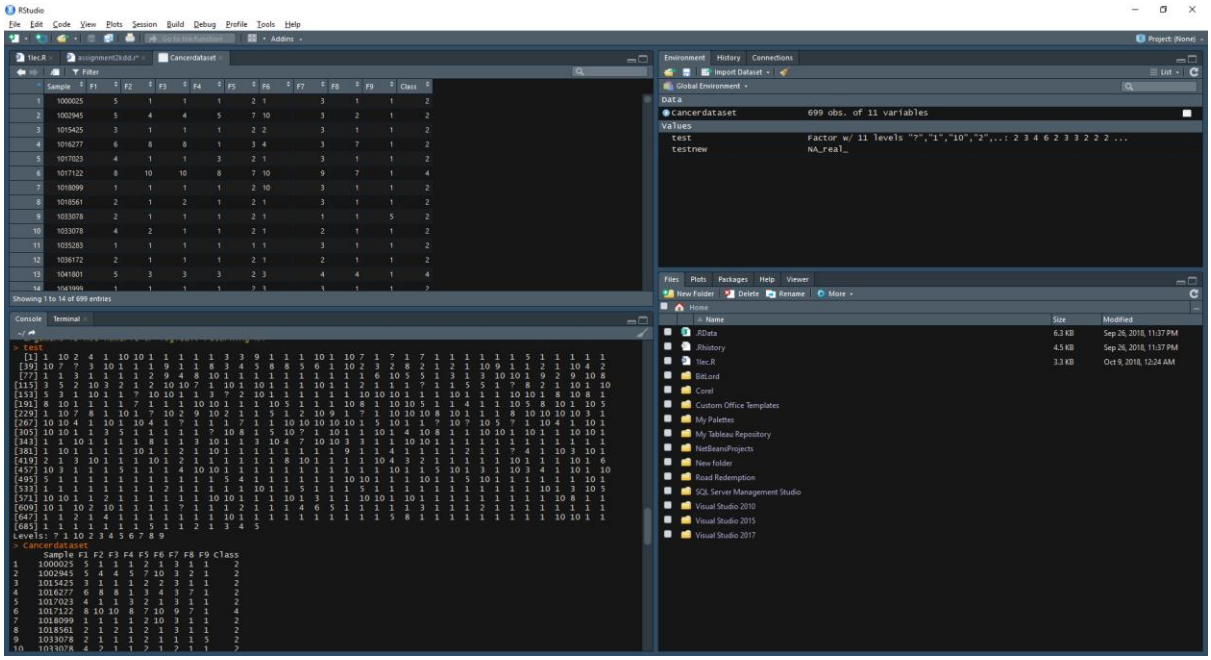
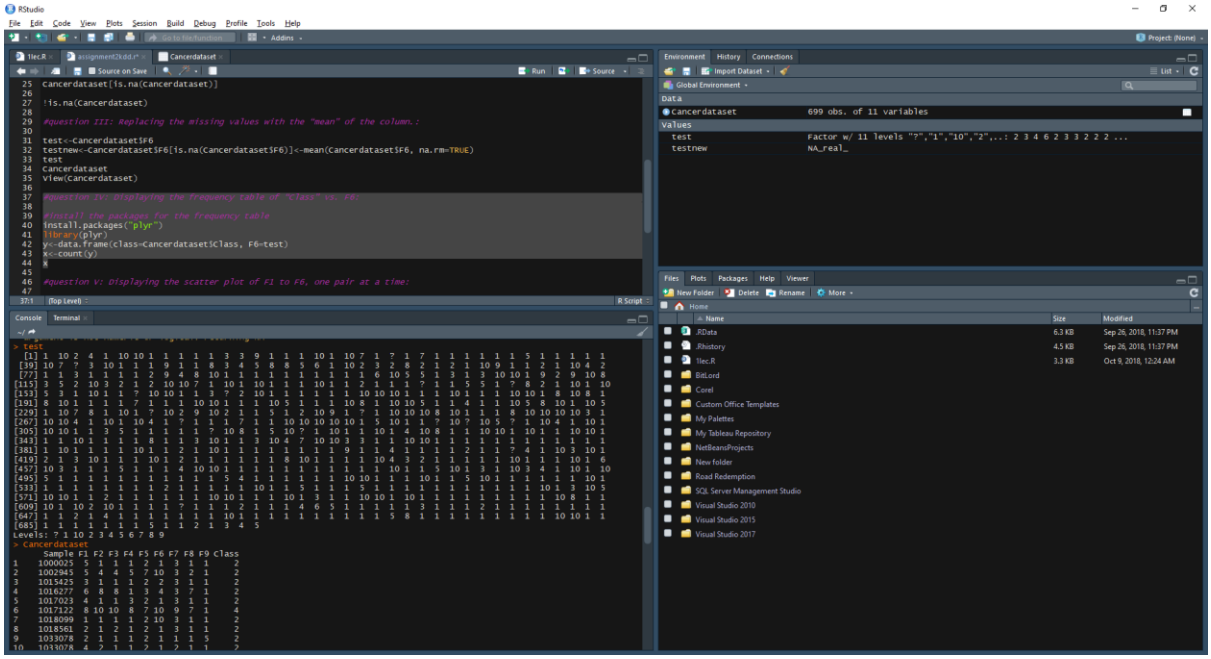
Results:

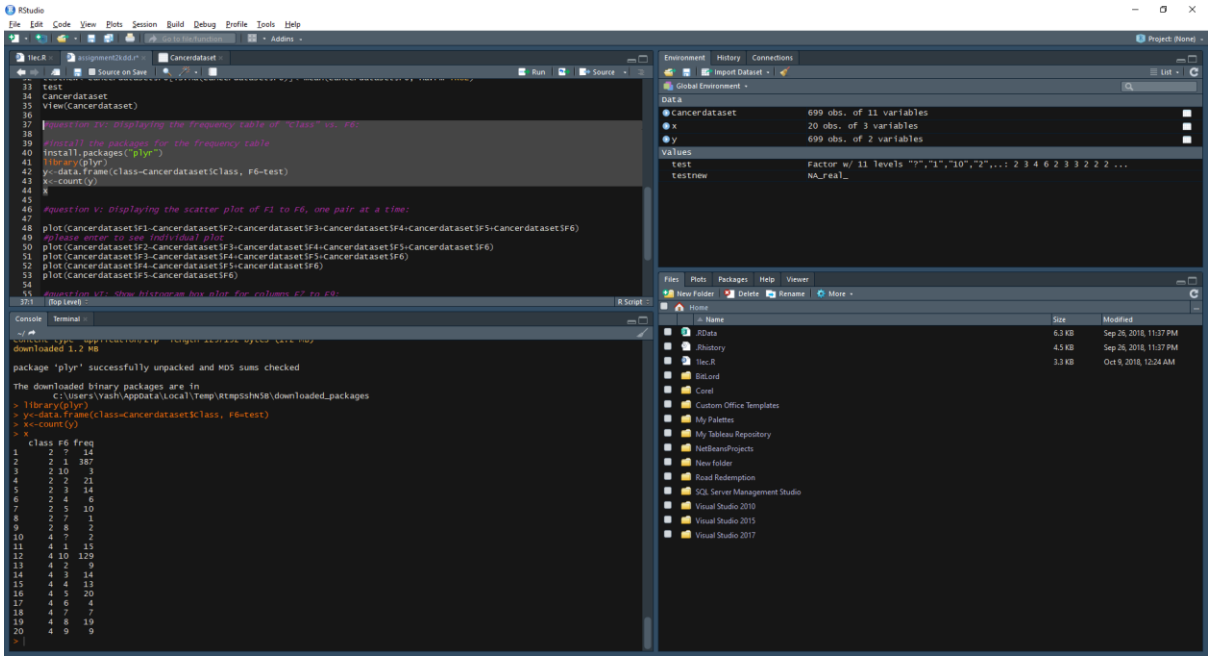


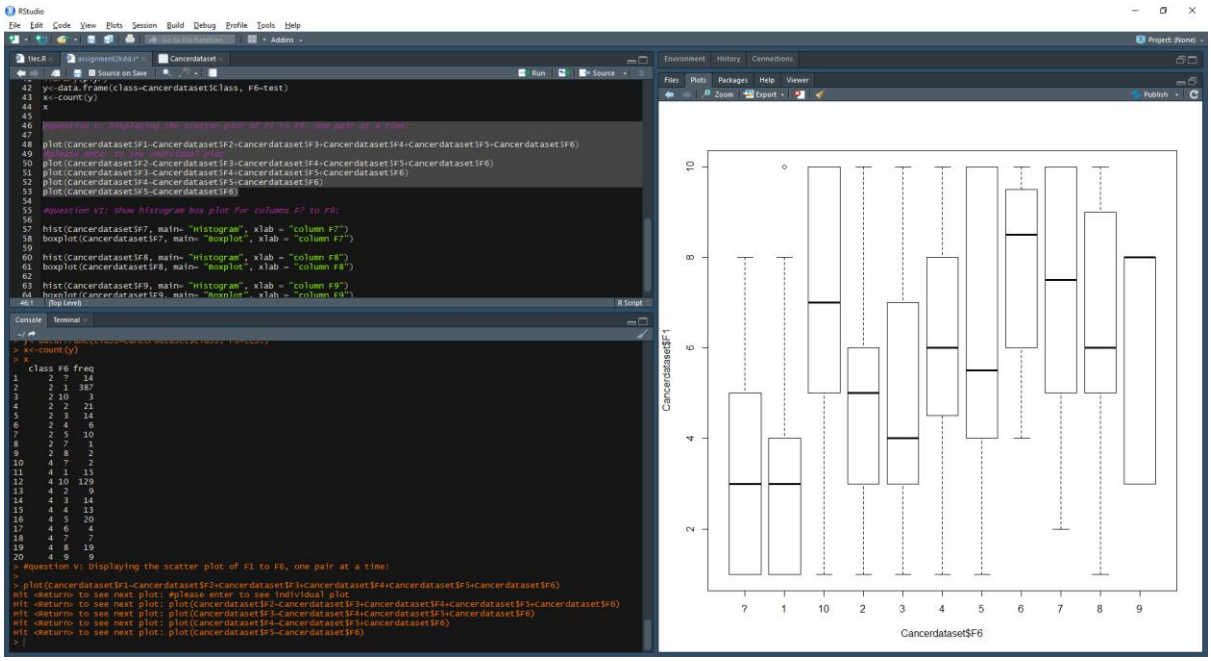


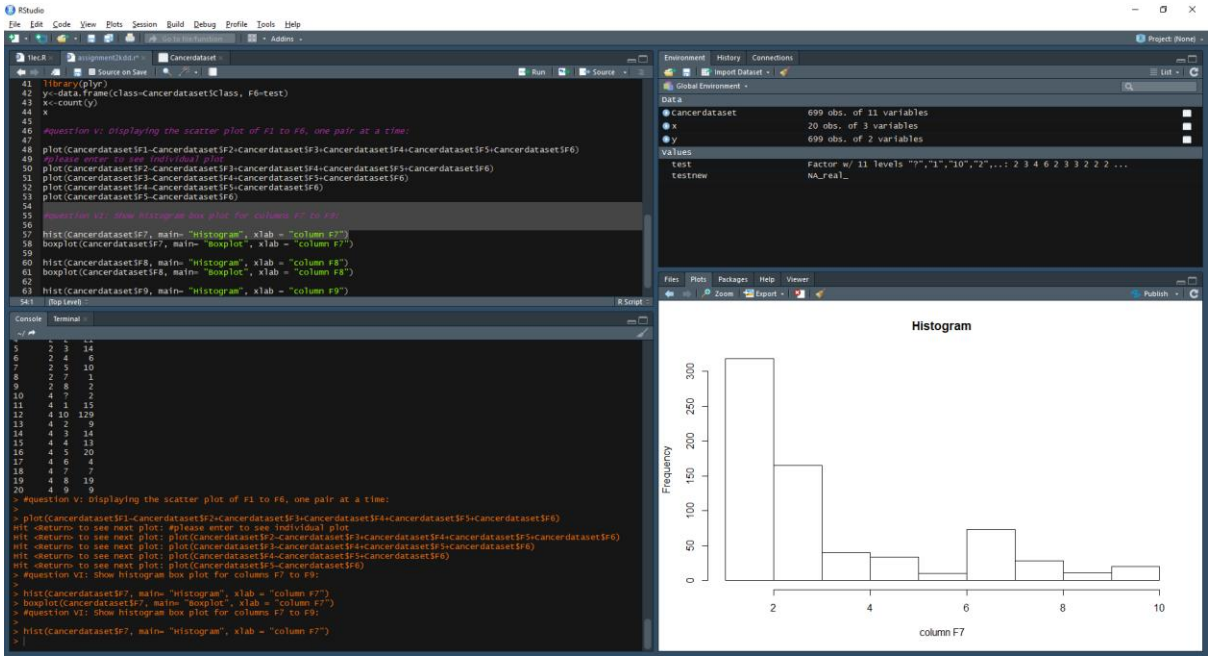

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
summary(Cancerdataset$F5)
summary(Cancerdataset$F6)
summary(Cancerdataset$F7)
summary(Cancerdataset$F8)
summary(Cancerdataset$F9)
#question II: Identifying missing values:
Cancerdataset[is.na(Cancerdataset)]
!is.na(Cancerdataset)
#question III: Replacing the missing values with the "mean" of the column.:
test<-Cancerdataset$F6
testnew<-Cancerdataset$F6[is.na(Cancerdataset$F6)]<-mean(Cancerdataset$F6, na.rm=TRUE)
test
Cancerdataset
View(Cancerdataset)
#question IV: Displaying the frequency table of "Class" vs. F6:
> #question II: Identifying missing values:
> Cancerdataset[is.na(Cancerdataset)]
character(0)
> !is.na(Cancerdataset)
Sample F1 F2 F3 F4 F5 F6 F7 F8 F9 Class
Environment History Connections
Import Dataset
Global Environment
Data
Cancerdataset 699 obs. of 11 variables
Files Plots Packages Help Viewer
New Folder Delete Rename More
Home
Name Size Modified
.RData 6.3 KB Sep 26, 2018, 11:37 PM
.Rhistory 4.5 KB Sep 26, 2018, 11:37 PM
1lec.R 3.3 KB Oct 9, 2018, 12:24 AM
BitLord
Corel
Custom Office Templates
My Palettes
My Tableau Repository
NetBeansProjects
New folder
Road Redemption
SQL Server Management Studio
Visual Studio 2010
Visual Studio 2015
Visual Studio 2017

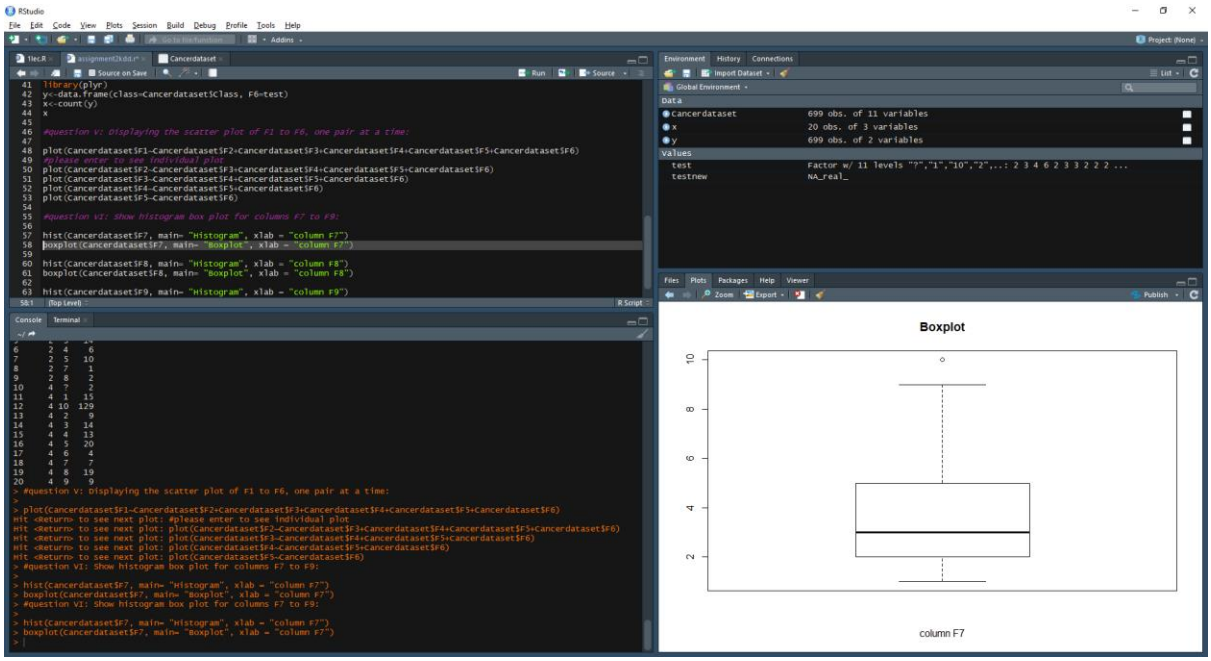

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Showing 1 to 14 of 699 entries
> test
Levels: ? 1 10 2 3 4 5 6 7 8 9
> Cancerdataset
Environment History Connections
Import Dataset
Global Environment
Data
Cancerdataset 699 obs. of 11 variables
Values
test Factor w/ 11 levels "?","1","10","2",..: 2 3 4 6 2 3 3 2 2 2 ...
testnew NA_real_
Files Plots Packages Help Viewer
New Folder Delete Rename More
Home
Name Size Modified
.RData 6.3 KB Sep 26, 2018, 11:37 PM
.Rhistory 4.5 KB Sep 26, 2018, 11:37 PM
1lec.R 3.3 KB Oct 9, 2018, 12:24 AM
BitLord
Corel
Custom Office Templates
My Palettes
My Tableau Repository
NetBeansProjects
New folder
Road Redemption
SQL Server Management Studio
Visual Studio 2010
Visual Studio 2015
Visual Studio 2017

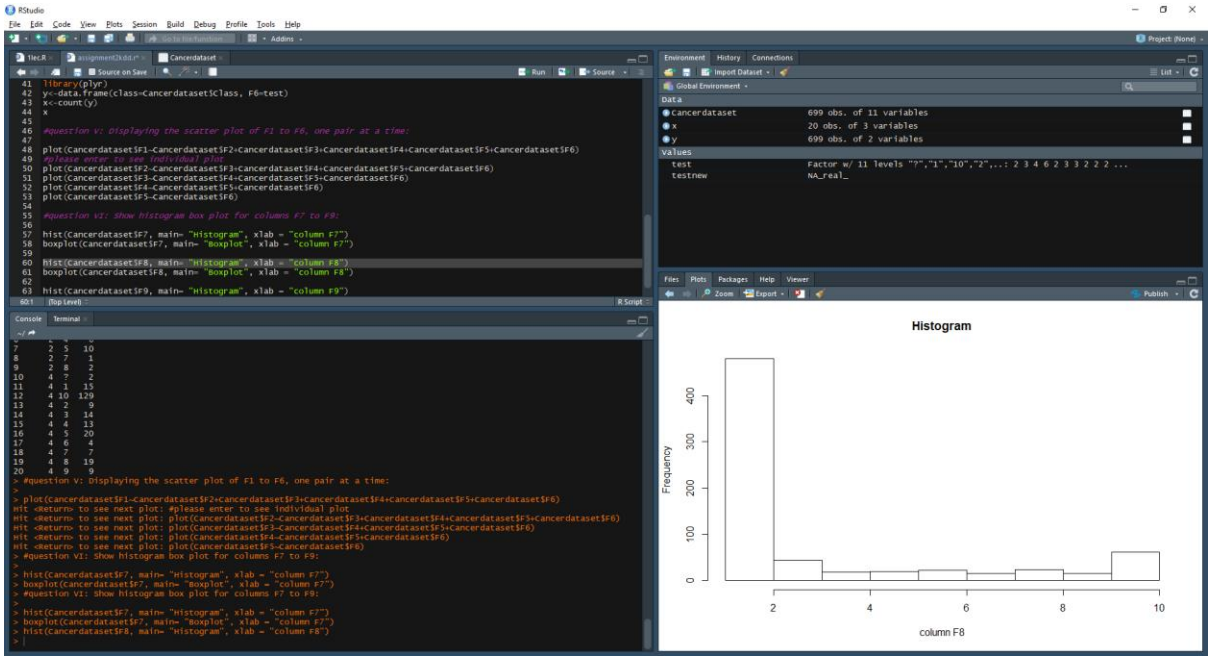

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
1lec.R
assignment2k.rd.r*
Cancerdataset
Source on Save
Run
Source
cancerdataset[is.na(cancerdataset)]
!is.na(cancerdataset)
#question III: Replacing the missing values with the "mean" of the column.:
test<-cancerdataset$F6
testnew<-cancerdataset$F6[!is.na(cancerdataset$F6)]<-mean(cancerdataset$F6, na.rm=TRUE)
test
cancerdataset
view(cancerdataset)
#question IV: Displaying the frequency table of "Class" vs. F6:
#install the packages for the frequency table
install.packages("plyr")
library(plyr)
y<-data.frame(class=cancerdataset$Class, F6=test)
x<-count(y)
x
#question V: Displaying the scatter plot of F1 to F6, one pair at a time:
(Top Level)
R Script
Console
Terminal
> test
[1] 1 10 2 4 1 10 10 1 1 1 1 1 3 3 9 1 1 1 10 1 10 7 1 7 1 7 1 1 1 1 1 5 1 1 1 1 1
[39] 10 7 7 3 10 1 1 1 9 1 1 8 3 4 5 8 8 5 6 1 10 2 3 2 8 2 1 2 1 10 9 1 1 2 1 10 4 2
[77] 1 1 3 1 1 1 1 2 9 4 8 10 1 1 1 1 1 1 1 1 1 6 10 5 5 1 3 1 3 10 10 1 9 2 9 10 8
[115] 3 5 2 10 2 2 1 2 10 10 7 1 10 1 10 1 1 1 10 1 1 2 1 1 1 7 1 1 5 5 1 7 8 2 1 10 1 10
[153] 5 3 1 10 1 2 1 7 10 10 1 1 3 7 2 10 1 1 1 1 1 1 1 10 1 1 1 10 10 1 1 1 10 10 1 8 10 8 1
[191] 8 10 1 1 1 1 7 1 1 1 10 10 1 1 1 10 5 1 1 1 10 8 1 10 10 5 1 1 4 1 1 10 5 8 10 1 10 5
[229] 1 10 7 8 1 10 1 7 10 2 9 10 2 1 1 5 1 2 10 9 1 7 1 10 10 10 8 10 1 1 1 8 10 10 10 10 3 1
[267] 10 10 4 1 10 1 10 4 1 7 1 1 1 7 1 1 10 10 10 10 10 1 5 10 1 1 7 10 7 10 5 7 1 10 4 1 10 1
[305] 10 10 1 1 3 5 1 1 1 1 1 7 10 8 1 5 10 7 1 10 1 1 10 1 4 10 8 1 1 10 10 1 10 1 1 10 10 1
[343] 1 1 10 1 1 1 1 8 1 1 3 10 1 1 3 10 4 7 10 10 3 1 1 10 10 1 1 1 1 1 1 1 1 1 1 1
[381] 1 10 1 1 1 1 10 1 1 2 1 10 1 1 1 1 1 1 1 1 9 1 1 4 1 1 1 1 2 1 1 7 4 1 10 3 10 1
[419] 2 1 3 10 1 1 1 10 1 2 1 1 1 1 1 1 8 10 1 1 1 1 10 4 3 2 1 1 1 1 1 10 1 1 1 10 1 6
[457] 10 3 1 1 1 5 1 1 1 1 4 10 10 1 1 1 1 1 1 1 1 1 1 10 1 1 5 10 1 3 1 10 3 4 1 10 1 10
[495] 5 1 1 1 1 1 1 1 1 1 1 1 5 4 1 1 1 1 1 1 1 1 10 10 1 1 1 10 1 1 5 10 1 1 1 1 1 10 1
[533] 1 1 1 1 1 1 1 1 2 1 1 1 1 1 10 1 1 5 1 1 1 1 1 1 1 1 1 1 1 10 1 3 10 5
[571] 10 10 1 1 2 1 1 1 1 1 1 1 10 10 1 1 1 10 1 3 1 1 10 10 1 10 1 1 1 1 1 1 1 1 1 10 8 1 1
[609] 10 1 10 2 10 1 1 1 1 7 1 1 1 2 1 1 1 4 6 5 1 1 1 1 1 3 1 1 1 2 1 1 1 1 1 1 1 1 1
[647] 1 1 2 1 4 1 1 1 1 1 1 1 1 10 1 1 1 1 1 1 1 1 1 5 8 1 1 1 1 1 1 1 1 1 10 10 1 1
[685] 1 1 1 1 3 1 1 1 5 1 1 2 1 3 4 5
Levels: ? 1 10 2 3 4 5 6 7 8 9
> cancerdataset
Sample F1 F2 F3 F4 F5 F6 F7 F8 F9 Class
1 1000025 5 1 1 1 2 1 3 1 1 2
2 1002945 5 4 4 5 7 10 3 2 1 2
3 1015425 3 1 1 1 2 2 3 1 1 2
4 1016277 6 8 8 1 3 4 3 7 1 2
5 1017023 4 1 1 3 2 1 3 1 1 2
6 1017122 8 10 10 8 7 10 9 7 1 4
7 1018099 1 1 1 1 2 10 3 1 1 2
8 1018561 2 1 2 1 2 1 3 1 1 2
9 1033078 2 1 1 1 2 1 1 1 5 2
10 1033078 4 2 1 1 2 1 2 1 1 2
Environment
History
Connections
Import Dataset
Global Environment
Data
Cancerdataset 699 obs. of 11 variables
Values
test Factor w/ 11 levels "?","1","10","2",..: 2 3 4 6 2 3 3 2 2 2 ...
testnew NA_real_
Files
Plots
Packages
Help
Viewer
New Folder
Delete
Rename
More
Home
Name
Size
Modified
.RData 6.3 KB Sep 26, 2018, 11:37 PM
.Rhistory 4.5 KB Sep 26, 2018, 11:37 PM
1lec.R 3.3 KB Oct 9, 2018, 12:24 AM
BitLord
Corel
Custom Office Templates
My Palettes
My Tableau Repository
NetBeansProjects
New folder
Road Redemption
SQL Server Management Studio
Visual Studio 2010
Visual Studio 2015
Visual Studio 2017

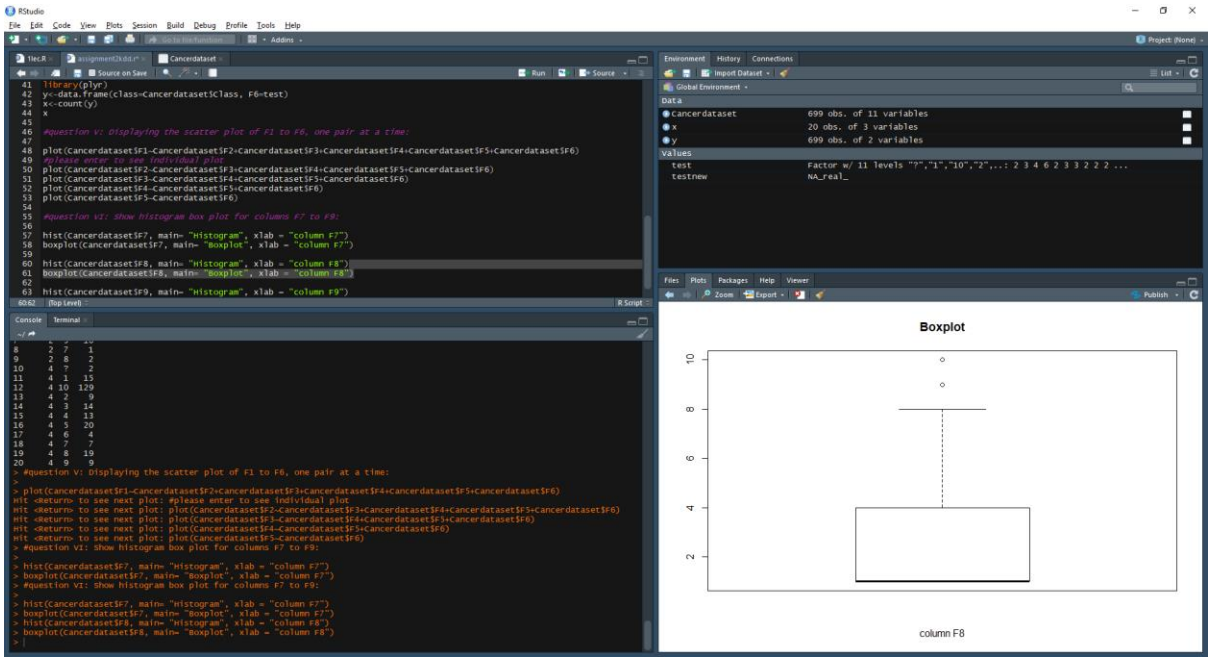

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
1lec.R
assignment2k.rd.r*
Cancerdataset
Source on Save
Run
Source
test
cancerdataset
view(cancerdataset)
#question IV: Displaying the frequency table of "Class" vs. F6:
#install the packages for the frequency table
install.packages("plyr")
library(plyr)
y<-data.frame(class=cancerdataset$Class, F6=test)
x<-count(y)
x
#question V: Displaying the scatter plot of F1 to F6, one pair at a time:
plot(cancerdataset$F1~cancerdataset$F2+cancerdataset$F3+cancerdataset$F4+cancerdataset$F5+cancerdataset$F6)
#Please enter to see individual plot :
plot(cancerdataset$F2~cancerdataset$F3+cancerdataset$F4+cancerdataset$F5+cancerdataset$F6)
plot(cancerdataset$F3~cancerdataset$F4+cancerdataset$F5+cancerdataset$F6)
plot(cancerdataset$F4~cancerdataset$F5+cancerdataset$F6)
plot(cancerdataset$F5~cancerdataset$F6)
#question VI: Show histogram box plot for columns F2 to F9:
(Top Level)
R Script
Console
Terminal
downloaded 1.2 MB
package 'plyr' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
C:\Users\Yash\AppData\Local\Temp\RtmpSshN5B\downloaded_packages
> library(plyr)
> y<-data.frame(class=cancerdataset$Class, F6=test)
> x<-count(y)
> x
class F6 freq
1 2 ? 14
2 2 1 387
3 2 10 3
4 2 2 21
5 2 3 14
6 2 4 6
7 2 5 10
8 2 7 1
9 2 8 2
10 4 ? 2
11 4 1 15
12 4 10 129
13 4 2 9
14 4 3 14
15 4 4 13
16 4 5 20
17 4 6 4
18 4 7 7
19 4 8 19
20 4 9 9
Environment
History
Connections
Import Dataset
Global Environment
Data
Cancerdataset 699 obs. of 11 variables
x 20 obs. of 3 variables
y 699 obs. of 2 variables
Values
test Factor w/ 11 levels "?","1","10","2",..: 2 3 4 6 2 3 3 2 2 2 ...
testnew NA_real_
Files
Plots
Packages
Help
Viewer
New Folder
Delete
Rename
More
Home
Name
Size
Modified
.RData 6.3 KB Sep 26, 2018, 11:37 PM
.Rhistory 4.5 KB Sep 26, 2018, 11:37 PM
1lec.R 3.3 KB Oct 9, 2018, 12:24 AM
BitLord
Corel
Custom Office Templates
My Palettes
My Tableau Repository
NetBeansProjects
New folder
Road Redemption
SQL Server Management Studio
Visual Studio 2010
Visual Studio 2015
Visual Studio 2017

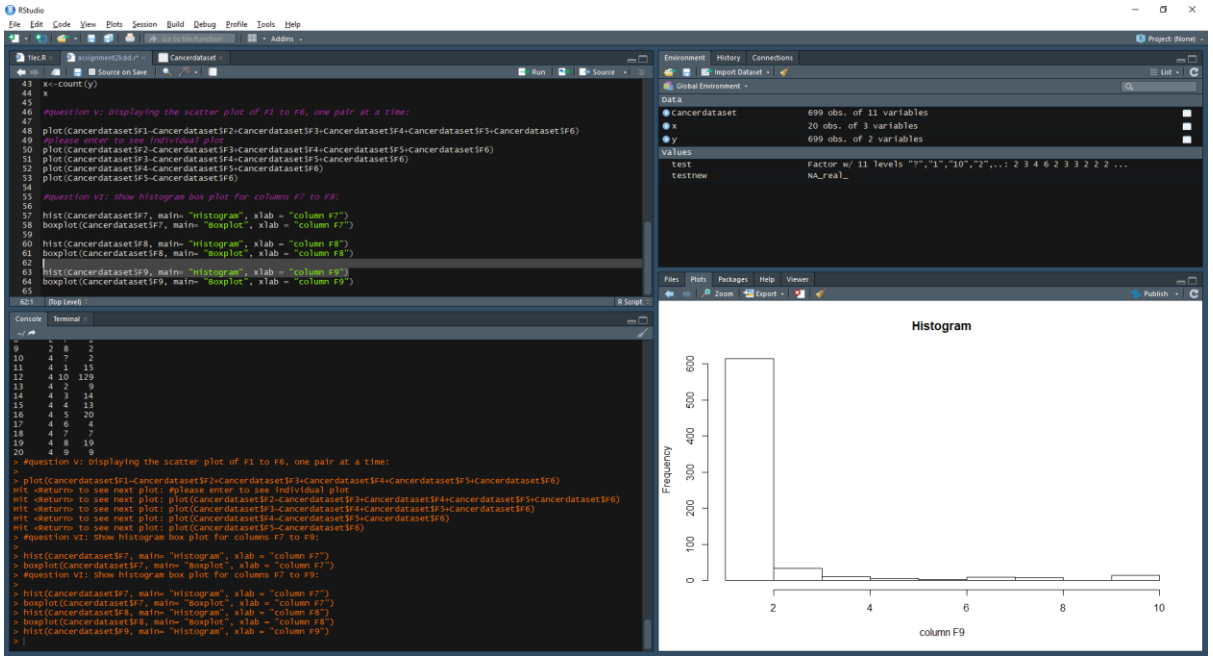

RStudio
Cancerdataset$F1
Cancerdataset$F6

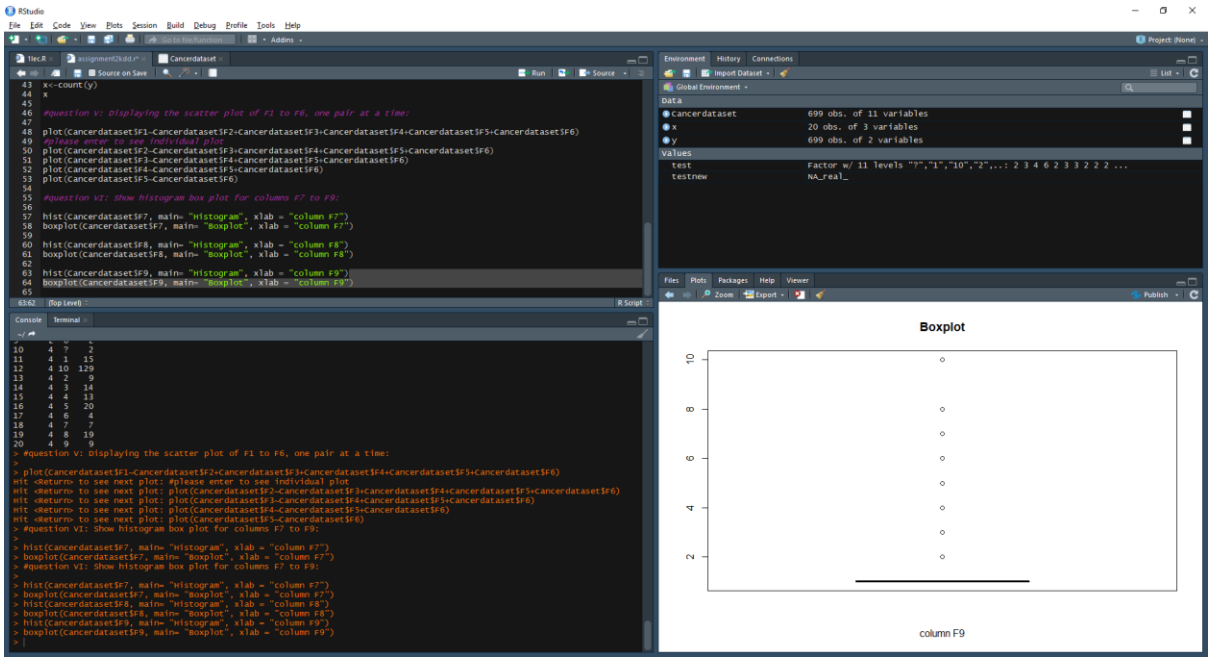

RStudio
Histogram
Frequency
column F7


Boxplot
column F7


Histogram
Frequency
column F8


Boxplot
column F8


Histogram
Frequency
column F9


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
1lec.R
assignment2kdd.rt
Cancerdataset
Source on Save
Run
Source
43 x<-count(y)
44 x
45
46 #question V: Displaying the scatter plot of F1 to F6, one pair at a time:
47
48 plot(Cancerdataset$F1~Cancerdataset$F2+Cancerdataset$F3+Cancerdataset$F4+Cancerdataset$F5+Cancerdataset$F6)
49 #please enter to see individual plot
50 plot(Cancerdataset$F2~Cancerdataset$F3+Cancerdataset$F4+Cancerdataset$F5+Cancerdataset$F6)
51 plot(Cancerdataset$F3~Cancerdataset$F4+Cancerdataset$F5+Cancerdataset$F6)
52 plot(Cancerdataset$F4~Cancerdataset$F5+Cancerdataset$F6)
53 plot(Cancerdataset$F5~Cancerdataset$F6)
54
55 #question VI: Show histogram box plot for columns F7 to F9:
56
57 hist(Cancerdataset$F7, main= "Histogram", xlab = "column F7")
58 boxplot(Cancerdataset$F7, main= "Boxplot", xlab = "column F7")
59
60 hist(Cancerdataset$F8, main= "Histogram", xlab = "column F8")
61 boxplot(Cancerdataset$F8, main= "Boxplot", xlab = "column F8")
62
63 hist(Cancerdataset$F9, main= "Histogram", xlab = "column F9")
64 boxplot(Cancerdataset$F9, main= "Boxplot", xlab = "column F9")
65
63:62 (Top Level) R Script
Console Terminal
10 4 ? 2
11 4 1 15
12 4 10 129
13 4 2 9
14 4 3 14
15 4 4 13
16 4 5 20
17 4 6 4
18 4 7 7
19 4 8 19
20 4 9 9
> #question V: Displaying the scatter plot of F1 to F6, one pair at a time:
>
> plot(Cancerdataset$F1~Cancerdataset$F2+Cancerdataset$F3+Cancerdataset$F4+Cancerdataset$F5+Cancerdataset$F6)
Hit <Return> to see next plot: #please enter to see individual plot
Hit <Return> to see next plot: plot(Cancerdataset$F2~Cancerdataset$F3+Cancerdataset$F4+Cancerdataset$F5+Cancerdataset$F6)
Hit <Return> to see next plot: plot(Cancerdataset$F3~Cancerdataset$F4+Cancerdataset$F5+Cancerdataset$F6)
Hit <Return> to see next plot: plot(Cancerdataset$F4~Cancerdataset$F5+Cancerdataset$F6)
Hit <Return> to see next plot: plot(Cancerdataset$F5~Cancerdataset$F6)
> #question VI: Show histogram box plot for columns F7 to F9:
>
> hist(Cancerdataset$F7, main= "Histogram", xlab = "column F7")
> boxplot(Cancerdataset$F7, main= "Boxplot", xlab = "column F7")
> #question VI: Show histogram box plot for columns F7 to F9:
>
> hist(Cancerdataset$F7, main= "Histogram", xlab = "column F7")
> boxplot(Cancerdataset$F7, main= "Boxplot", xlab = "column F7")
> hist(Cancerdataset$F8, main= "Histogram", xlab = "column F8")
> boxplot(Cancerdataset$F8, main= "Boxplot", xlab = "column F8")
> hist(Cancerdataset$F9, main= "Histogram", xlab = "column F9")
> boxplot(Cancerdataset$F9, main= "Boxplot", xlab = "column F9")
>
Environment History Connections
Import Dataset
Global Environment
Data
Cancerdataset 699 obs. of 11 variables
x 20 obs. of 3 variables
y 699 obs. of 2 variables
Values
test Factor w/ 11 levels "?","1","10","2",..: 2 3 4 6 2 3 3 2 2 2 ...
testnew NA_real_
Files Plots Packages Help Viewer
Zoom Export
Publish
Boxplot
10
8
6
4
2
column F9