

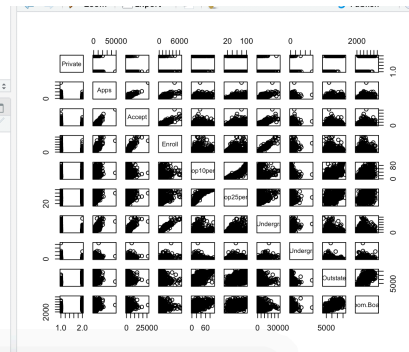
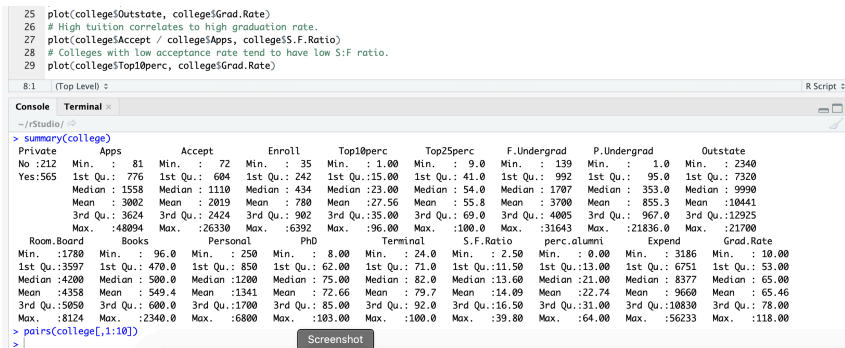
Homework 1

Ankit Chaudhary
ac957 [31470815]

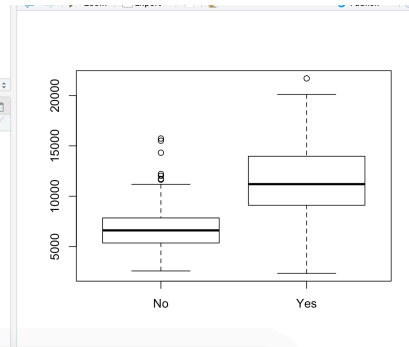
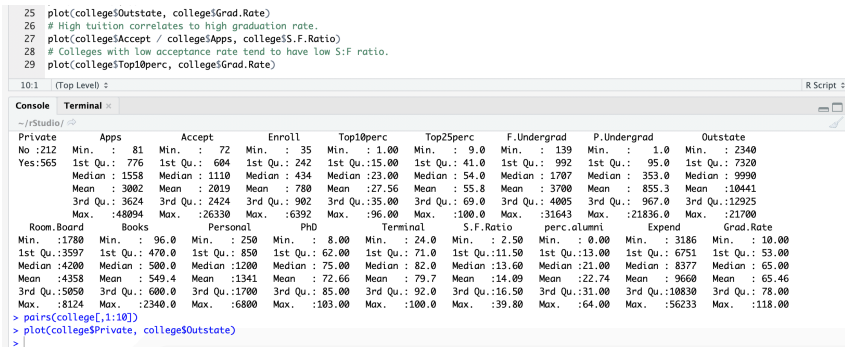
```
> college =  
read.csv("/Users/ankitchaudhary/Documents/semester2_spring19/cs678_SMDS/week2/College.c  
sv")  
  
> fix(college)  
Error in edit.data.frame(get(subx, envir = parent), title = subx, ...) :  
  X11 is not available  
  
> rownames(college) = college[,1]  
  
> college = college[,-1]  
  
> summary(college)  
Private      Apps      Accept      Enroll      Top10perc      Top25perc      F.Undergrad  
P.Undergrad      Outstate  
No:212 Min. : 81 Min. : 72 Min. : 35 Min. : 1.00 Min. : 9.0 Min. : 139  
Min. : 1.0 Min. : 2340  
Yes:565 1st Qu.: 776 1st Qu.: 604 1st Qu.: 242 1st Qu.:15.00 1st Qu.: 41.0 1st Qu.:  
992 1st Qu.: 95.0 1st Qu.: 7320  
Median : 1558 Median : 1110 Median : 434 Median :23.00 Median : 54.0 Median :  
1707 Median : 353.0 Median : 9990  
Mean : 3002 Mean : 2019 Mean : 780 Mean :27.56 Mean : 55.8 Mean :  
3700 Mean : 855.3 Mean :10441  
3rd Qu.: 3624 3rd Qu.: 2424 3rd Qu.: 902 3rd Qu.:35.00 3rd Qu.: 69.0 3rd Qu.:  
4005 3rd Qu.: 967.0 3rd Qu.:12925  
Max. :48094 Max. :26330 Max. :6392 Max. :96.00 Max. :100.0 Max. :  
:31643 Max. :21836.0 Max. :21700  
Room.Board      Books      Personal      PhD      Terminal      S.F.Ratio      perc.alumni  
Expend      Grad.Rate  
Min. :1780 Min. : 96.0 Min. : 250 Min. : 8.00 Min. : 24.0 Min. : 2.50 Min. :  
0.00 Min. : 3186 Min. : 10.00  
1st Qu.:3597 1st Qu.: 470.0 1st Qu.: 850 1st Qu.: 62.00 1st Qu.: 71.0 1st Qu.:11.50 1st  
Qu.:13.00 1st Qu.: 6751 1st Qu.: 53.00  
Median :4200 Median : 500.0 Median :1200 Median : 75.00 Median : 82.0 Median :13.60  
Median :21.00 Median : 8377 Median : 65.00  
Mean :4358 Mean : 549.4 Mean :1341 Mean : 72.66 Mean : 79.7 Mean :14.09  
Mean :22.74 Mean : 9660 Mean : 65.46  
3rd Qu.:5050 3rd Qu.: 600.0 3rd Qu.:1700 3rd Qu.: 85.00 3rd Qu.: 92.0 3rd Qu.:16.50  
3rd Qu.:31.00 3rd Qu.:10830 3rd Qu.: 78.00
```

Max. :8124 Max. :2340.0 Max. :6800 Max. :103.00 Max. :100.0 Max. :39.80
 Max. :64.00 Max. :56233 Max. :118.00

```
> pairs(college[,1:10])
```



```
> plot(college$Private, college$Outstate)
```



```
> Elite = rep("No", nrow(college))
> Elite[college$Top10perc>50] = "Yes"
> Elite = as.factor(Elite)
> college = data.frame(college, Elite)
> summary(college$Elite)
No Yes
699 78
> plot(college$Elite, college$Outstate)
```

```

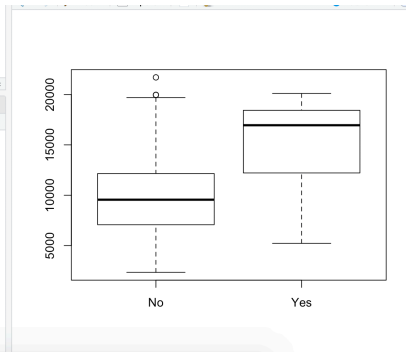
25 par(mfrow=c(1,1))
26 plot(college$Outstate, college$Grad.Rate)
27 # High tuition correlates to high graduation rate.
28 plot(college$Accept / college$Apps, college$S.F.Ratio)
29 # Colleges with low acceptance rate tend to have low S:F ratio.
30 plot(college$Top10perc, college$Grad.Rate)
31 (Top Level)

```

```

Console Terminal
~/RStudio/
Min. :1780 Min. : 96.0 Min. : 250 Min. : 8.00 Min. : 24.0 Min. : 2.50 Min. : 0.00 Min. : 3186 Min. : 10.00
1st Qu.:3597 1st Qu.: 470.0 1st Qu.: 850 1st Qu.: 62.00 1st Qu.: 71.0 1st Qu.:11.50 1st Qu.:13.00 1st Qu.: 6751 1st Qu.: 53.00
Median :4200 Median : 500.0 Median :1200 Median : 75.00 Median : 82.0 Median :13.60 Median :21.00 Median : 8377 Median : 65.00
Mean :4358 Mean : 549.4 Mean :1341 Mean : 72.66 Mean : 79.7 Mean :14.09 Mean :22.74 Mean : 9660 Mean : 65.46
3rd Qu.:5050 3rd Qu.: 600.0 3rd Qu.:1700 3rd Qu.: 85.00 3rd Qu.: 92.0 3rd Qu.:16.50 3rd Qu.:31.00 3rd Qu.:10830 3rd Qu.: 78.00
Max. :8124 Max. :2340.0 Max. :6800 Max. :103.00 Max. :100.0 Max. :39.80 Max. :64.00 Max. :56233 Max. :118.00
> pairs(college[,1:10])
> plot(college$Private, college$Outstate)
> Elite = rep("No", nrow(college))
> Elite[college$Top10perc>50] = "Yes"
> Elite = as.factor(Elite)
> college = data.frame(college, Elite)
> summary(college$Elite)
No Yes
699 78
> plot(college$Elite, college$Outstate)
>

```



- > par(mfrow=c(2,2))
- > hist(college\$Apps)
- > hist(college\$perc.alumni, col=2)
- > hist(college\$S.F.Ratio, col=3, breaks=10)
- > hist(college\$Expend, breaks=100)

```

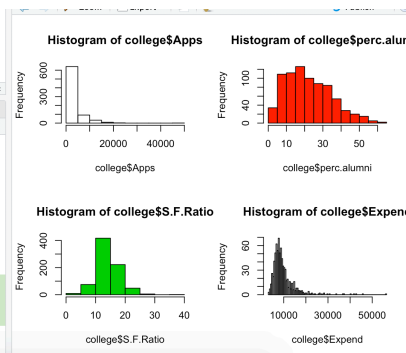
25 par(mfrow=c(1,1))
26 plot(college$Outstate, college$Grad.Rate)
27 # High tuition correlates to high graduation rate.
28 plot(college$Accept / college$Apps, college$S.F.Ratio)
29 # Colleges with low acceptance rate tend to have low S:F ratio.
30 plot(college$Top10perc, college$Grad.Rate)
31 (Top Level)
24:1

```

```

Console Terminal
~/RStudio/
Max. :8124 Max. :2340.0 Max. :6800 Max. :103.00 Max. :100.0 Max. :39.80 Max. :64.00 Max. :56233 Max. :118.00
> pairs(college[,1:10])
> plot(college$Private, college$Outstate)
> Elite = rep("No", nrow(college))
> Elite[college$Top10perc>50] = "Yes"
> Elite = as.factor(Elite)
> college = data.frame(college, Elite)
> summary(college$Elite)
No Yes
699 78
> plot(college$Elite, college$Outstate)
> par(mfrow=c(2,2))
> hist(college$Apps)
> hist(college$perc.alumni, col=2)
> hist(college$S.F.Ratio, col=3, breaks=10)
> hist(college$Expend, breaks=100)
>

```



- > par(mfrow=c(1,1))
- > plot(college\$Elite, college\$Grad.Rate)
- > # High tuition correlates to high graduation rate.

```

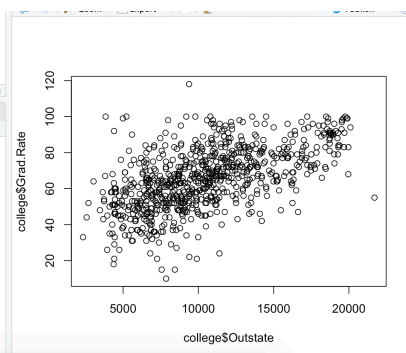
26 plot(college$Outstate, college$Grad.Rate)
27 # High tuition correlates to high graduation rate.
28 plot(college$Accept / college$Apps, college$S.F.Ratio)
29 # Colleges with low acceptance rate tend to have low S:F ratio.
30 plot(college$Top10perc, college$Grad.Rate)
31
26:1 (Top Level)

```

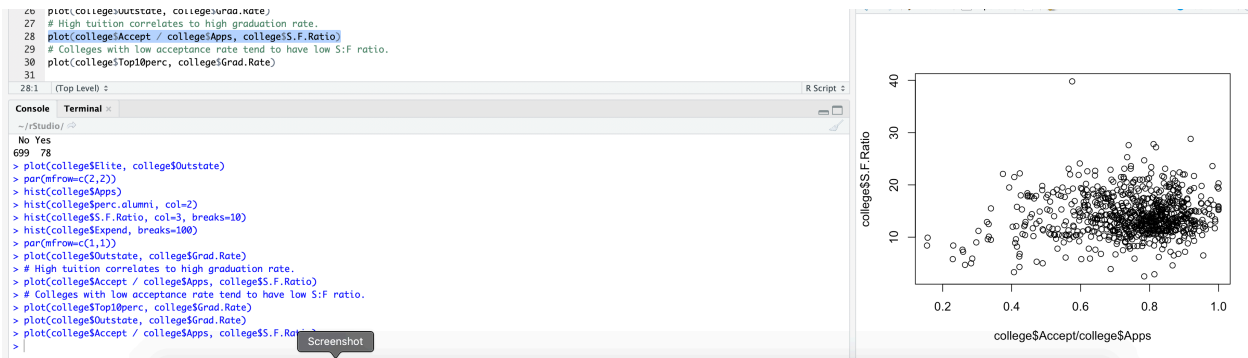
```

Console Terminal
~/RStudio/
> summary(college$Elite)
No Yes
699 78
> plot(college$Elite, college$Outstate)
> par(mfrow=c(2,2))
> hist(college$Apps)
> hist(college$perc.alumni, col=2)
> hist(college$S.F.Ratio, col=3, breaks=10)
> hist(college$Expend, breaks=100)
> par(mfrow=c(1,1))
> plot(college$Outstate, college$Grad.Rate)
> # High tuition correlates to high graduation rate.
> plot(college$Accept / college$Apps, college$S.F.Ratio)
> # Colleges with low acceptance rate tend to have low S:F ratio.
> plot(college$Top10perc, college$Grad.Rate)
> plot(college$Outstate, college$Grad.Rate)
>

```



- > plot(college\$Accept / college\$Apps, college\$S.F.Ratio)
- > # Colleges with low acceptance rate tend to have low S:F ratio.



> plot(college\$Top10perc, college\$Grad.Rate)

