

# chpt2.p8

*Claudius Taylor*

7/13/2019

## (a) reading the data

```
college = College
```

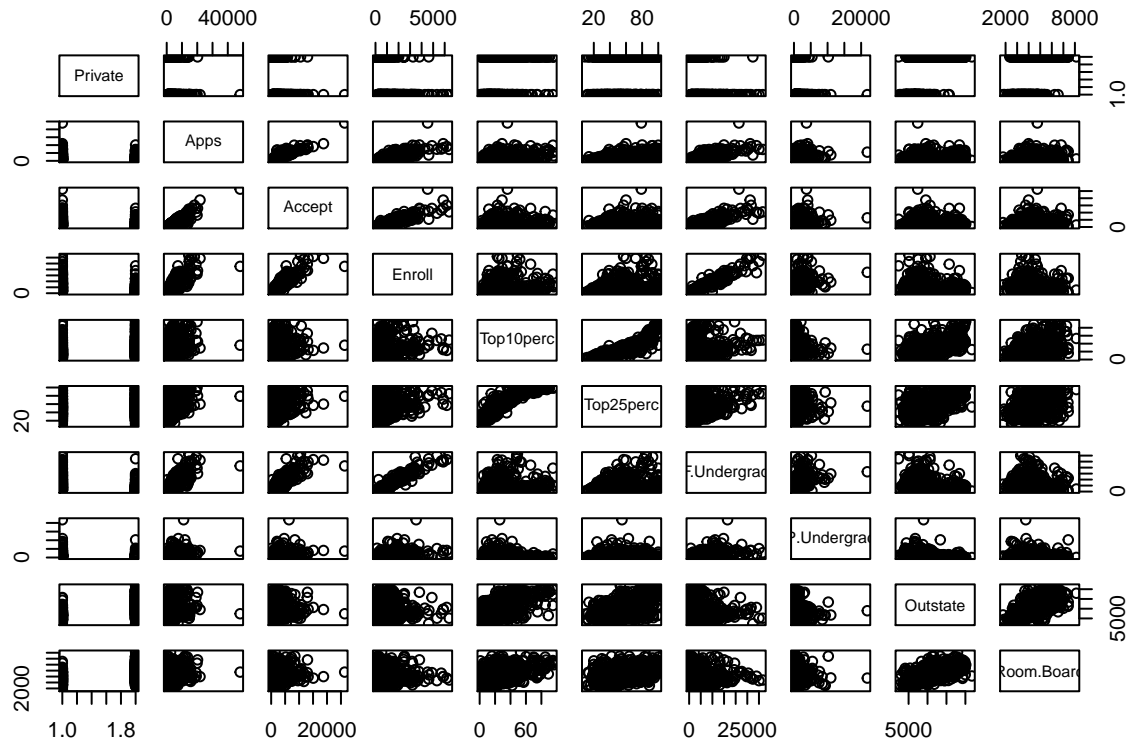
## (b) numerical summary of the variables in the data set

```
summary(college)
```

```
## Private      Apps      Accept      Enroll      Top10perc
## No :212      Min.   :   81      Min.   :   72      Min.   :   35      Min.   : 1.00
## Yes:565      1st Qu.:  776      1st Qu.:  604      1st Qu.:  242      1st Qu.:15.00
##           Median : 1558      Median : 1110      Median :  434      Median :23.00
##           Mean    : 3002      Mean    : 2019      Mean    :  780      Mean    :27.56
##           3rd Qu.: 3624      3rd Qu.: 2424      3rd Qu.:  902      3rd Qu.:35.00
##           Max.    :48094      Max.    :26330      Max.    :6392      Max.    :96.00
## Top25perc    F.Undergrad  P.Undergrad      Outstate
## Min.   : 9.0      Min.   : 139      Min.   :  1.0      Min.   : 2340
## 1st Qu.:41.0      1st Qu.: 992      1st Qu.: 95.0      1st Qu.: 7320
## Median :54.0      Median :1707      Median : 353.0      Median : 9990
## Mean    :55.8      Mean    :3700      Mean    : 855.3      Mean    :10441
## 3rd Qu.:69.0      3rd Qu.:4005      3rd Qu.: 967.0      3rd Qu.:12925
## Max.    :100.0      Max.    :31643      Max.    :21836.0      Max.    :21700
## Room.Board   Books      Personal      PhD
## Min.   :1780      Min.   : 96.0      Min.   : 250      Min.   : 8.00
## 1st Qu.:3597      1st Qu.:470.0      1st Qu.: 850      1st Qu.:62.00
## Median :4200      Median :500.0      Median :1200      Median :75.00
## Mean    :4358      Mean    :549.4      Mean    :1341      Mean    :72.66
## 3rd Qu.:5050      3rd Qu.:600.0      3rd Qu.:1700      3rd Qu.:85.00
## Max.    :8124      Max.    :2340.0      Max.    :6800      Max.    :103.00
## Terminal     S.F.Ratio    perc.alumni      Expend
## Min.   : 24.0      Min.   : 2.50      Min.   : 0.00      Min.   : 3186
## 1st Qu.: 71.0      1st Qu.:11.50      1st Qu.:13.00      1st Qu.:6751
## Median : 82.0      Median :13.60      Median :21.00      Median :8377
## Mean    : 79.7      Mean    :14.09      Mean    :22.74      Mean    :9660
## 3rd Qu.: 92.0      3rd Qu.:16.50      3rd Qu.:31.00      3rd Qu.:10830
## Max.    :100.0      Max.    :39.80      Max.    :64.00      Max.    :56233
## Grad.Rate
## Min.   : 10.00
## 1st Qu.: 53.00
## Median : 65.00
## Mean    : 65.46
## 3rd Qu.: 78.00
## Max.    :118.00
```

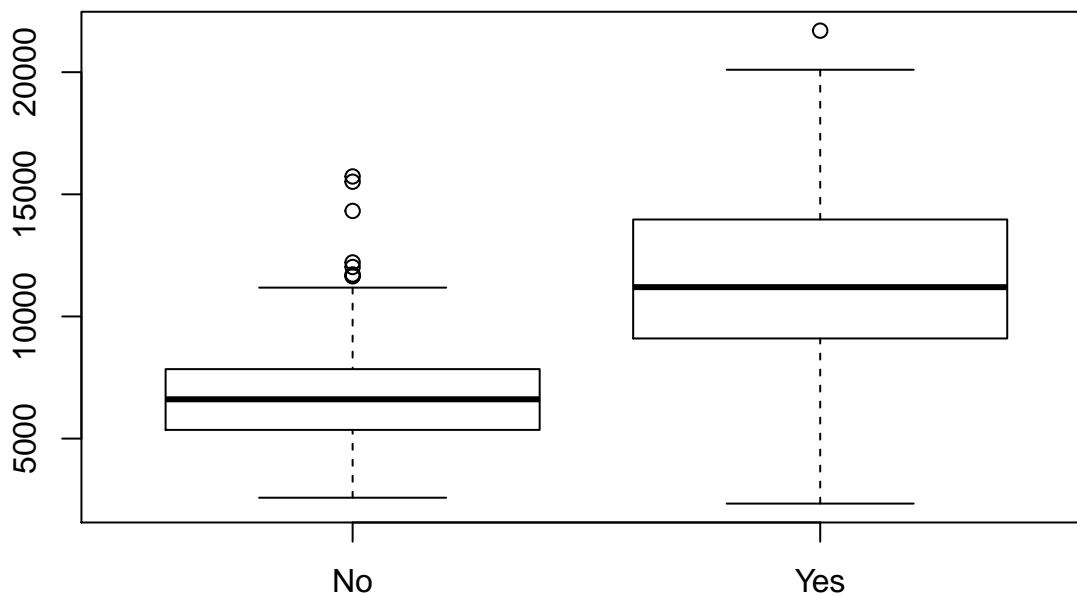
(c) scatterplot matrix of the first ten variables of the data

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



(d) produce a side-by-side boxplots of Outstate versus Private

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



(e) create a new qualitative variable, called Elite, by binning the Top10perc variable. Divide the universities into two groups based on whether or not the proportion of students coming from the Top10perc of the high school classes exceeds 50 %

```
Elite = rep("NO", nrow(college))
Elite[college$Top10perc > 50] = "YES"
Elite = as.factor(Elite)
college = data.frame(college, Elite)
```

(f) produce a side-by-side boxplots of Outstate versus Elite

```
plot(college$Elite, college$Outstate)
```

