CptS575 Hw2

Mengxiao

Hello world

Part 1

a. Read the data into R

```
data = read.csv('https://scads.eecs.wsu.edu/wp-content/uploads/2017/09/College.csv')
```

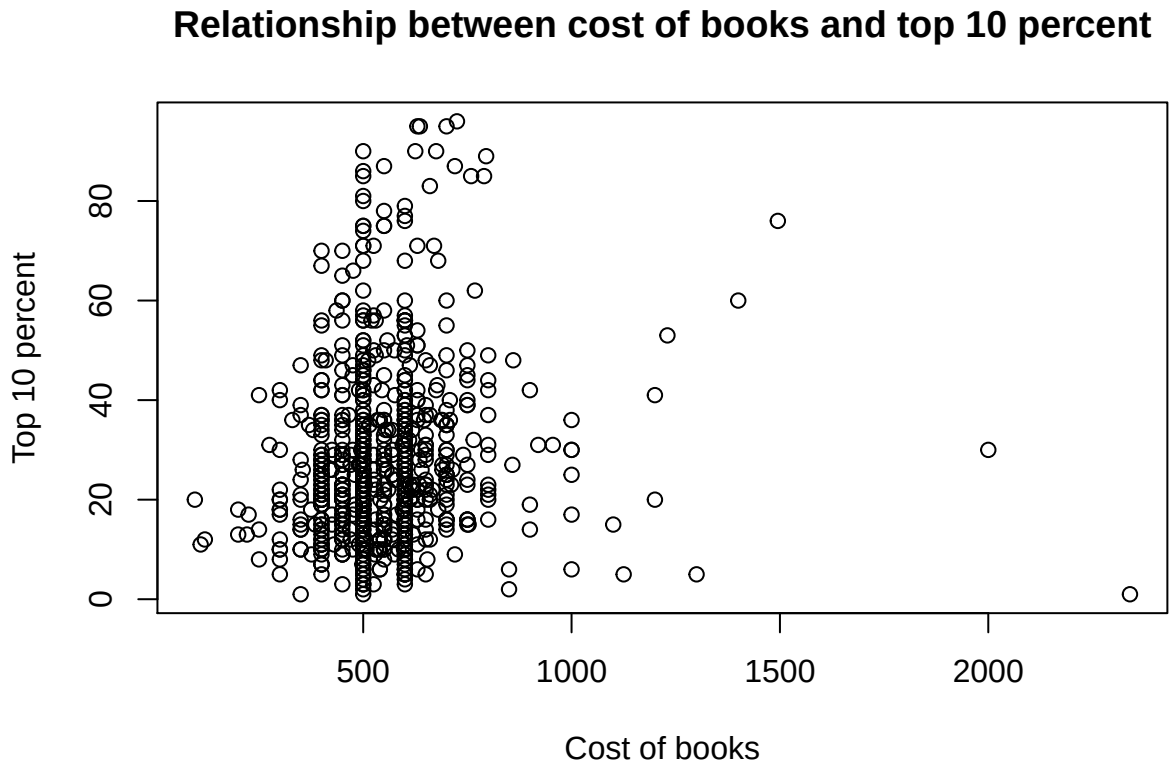
b. Find the median cost of books

```
books_median = median(data['Books'][, 1])
books_median
```

```
## [1] 500
```

c. Produce a scatterplot to show the relationship between the cost of books and Top 10 percent students.

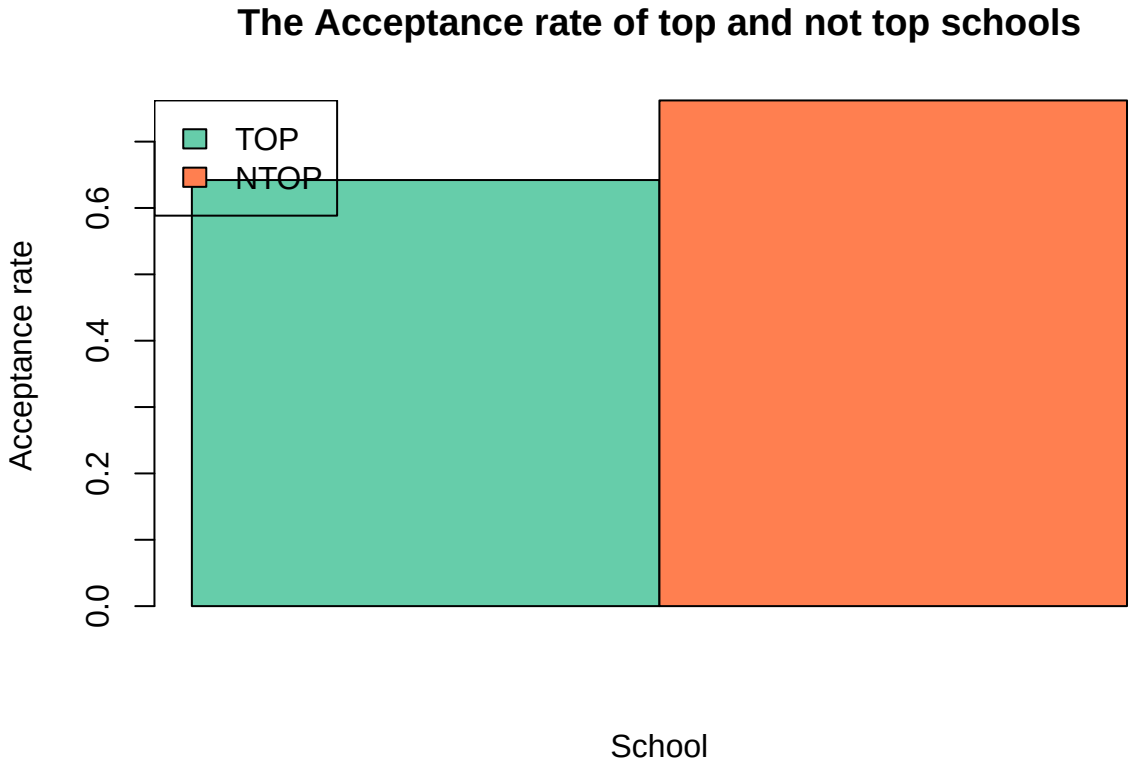
```
plot(x = data$Books,
     y = data$Top10perc,
     xlab = "Cost of books",
     ylab = "Top 10 percent",
     main = "Relationship between cost of books and top 10 percent"
     )
```



d. Produce a histogram showing the overall enrollment numbers for both public and private schools.

e. Seperate the schools to two parts, top and ntop(not top)

```
top <- sum(data[data$Top25perc>50,]$Accept)/sum(data[data$Top25perc>50,]$Apps)
ntop<- sum(data[data$Top25perc<=50,]$Accept)/sum(data[data$Top25perc<=50,]$Apps)
barplot(matrix(c(top, ntop)), ylab="Acceptance rate", xlab="School", col=c("aquamarine3", "coral"),
    main="The Acceptance rate of top and not top schools", beside=TRUE, width=0.2)
legend("topleft", legend=c("TOP","NTOP"), fill=c("aquamarine3", "coral"))
```


The Acceptance rate of top and not top schools
TOP
NTOP
Acceptance rate
0.0
0.2
0.4
0.6
School