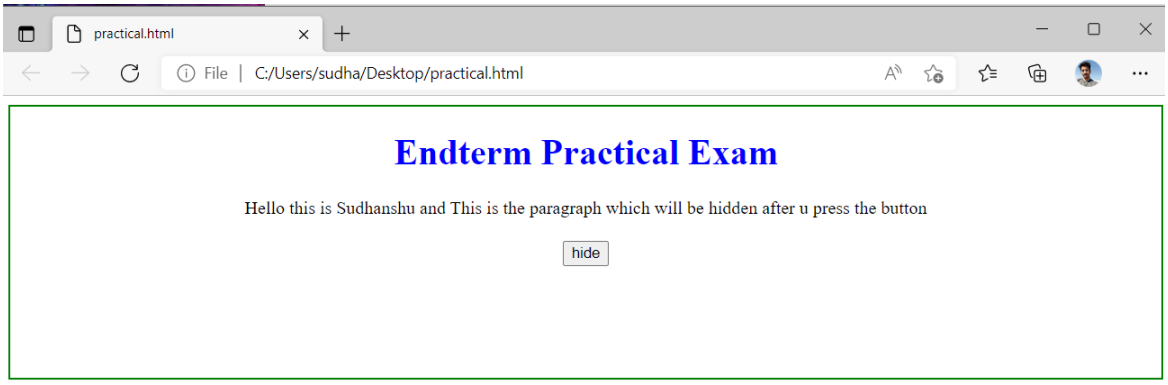
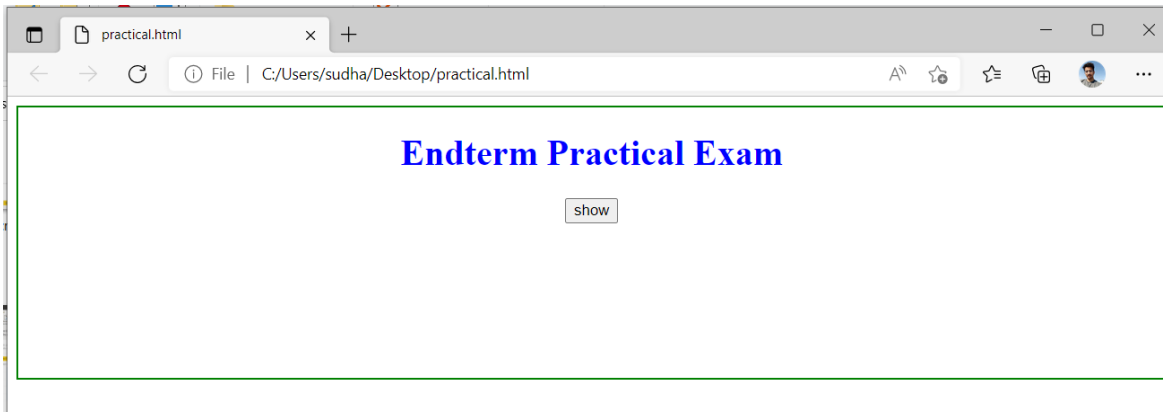
Outputs

Question 1

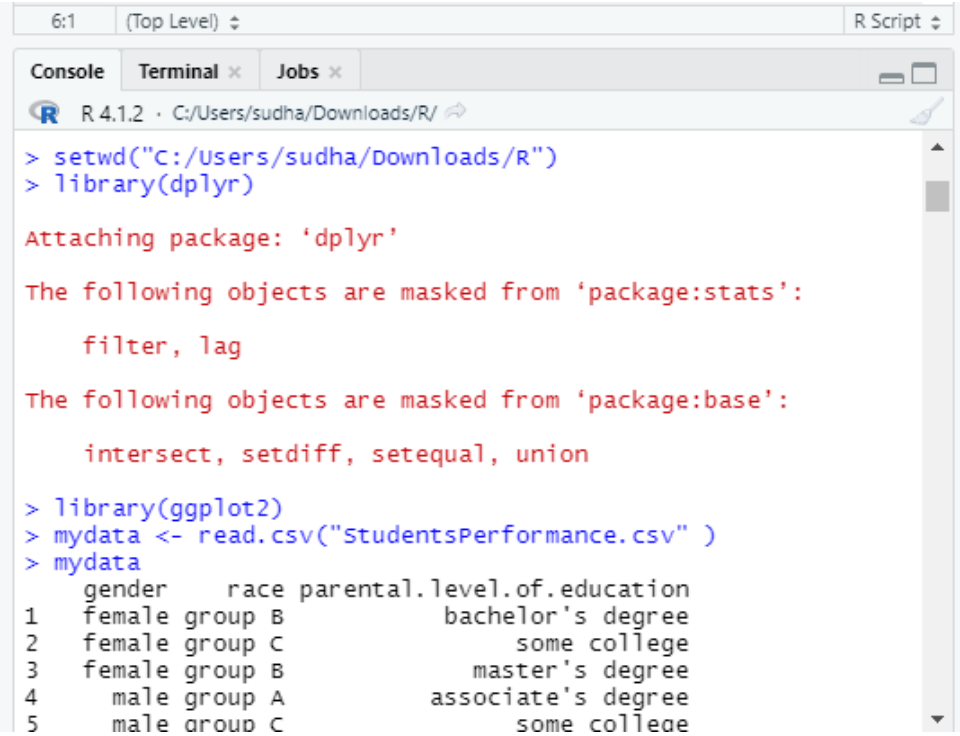
c_no	c_name	item_purchased	mob_no
1	dhondu	phone	349249
2	gauri	lappy	890949
3	raju	tablet	234244

Question 2





Question 3


6:1
(Top Level)
R Script
Console
Terminal
Jobs
R 4.1.2 · C:/Users/sudha/Downloads/R/
> setwd("C:/Users/sudha/Downloads/R")
> library(dplyr)
Attaching package: 'dplyr'
The following objects are masked from 'package:stats':
filter, lag
The following objects are masked from 'package:base':
intersect, setdiff, setequal, union
> library(ggplot2)
> mydata <- read.csv("StudentsPerformance.csv")
> mydata
gender race parental.level.of.education
1 female group B bachelor's degree
2 female group C some college
3 female group B master's degree
4 male group A associate's degree
5 male group C some college

```
R 4.1.2 · C:/Users/sudha/Downloads/R/
> names(mydata)
[1] "gender"
[2] "race"
[3] "parental.level.of.education"
[4] "lunch"
[5] "test.preparation.course"
[6] "math.score"
[7] "reading.score"
[8] "writing.score"
> str(mydata)
'data.frame':   1000 obs. of  8 variables:
 $ gender                     : chr  "female" "female" "femal
e" "male" ...
 $ race                       : chr  "group B" "group C" "grou
p B" "group A" ...
 $ parental.level.of.education: chr  "bachelor's degree" "some
 college" "master's degree" "associate's degree" ...
 $ lunch                      : chr  "standard" "standard" "st
andard" "free/reduced" ...
 $ test.preparation.course    : chr  "none" "completed" "none"
 "none" ...
 $ math.score                 : int  72 69 90 47 76 71 88 40 6
4 38 ...
 $ reading.score              : int  72 90 95 57 78 83 95 43 6
4 60 ...
 $ writing.score              : int  74 88 93 44 75 78 92 39 6
7 50 ...
> dim(mydata)
[1] 1000    8
> head(mydata)
  gender    race parental.level.of.education        lunch
1 female group B           bachelor's degree     standard
2 female group C                some college     standard
3 female group B             master's degree     standard
4   male group A          associate's degree free/reduced
5   male group C                some college     standard
6 female group B          associate's degree     standard
  test.preparation.course math.score reading.score
1                    none         72            72
2               completed         69            90
3                    none         90            95
4                    none         47            57
5                    none         76            78
6                    none         71            83
  writing.score
1            74
2            88
3            93
4            44
5            75
6            78
> tail(mydata)
     gender    race parental.level.of.education
995    male group A                high school
996  female group E             master's degree
997    male group C                 high school
998  female group C                 high school
999  female group D                some college
1000 female group D                some college
          lunch test.preparation.course math.score
```

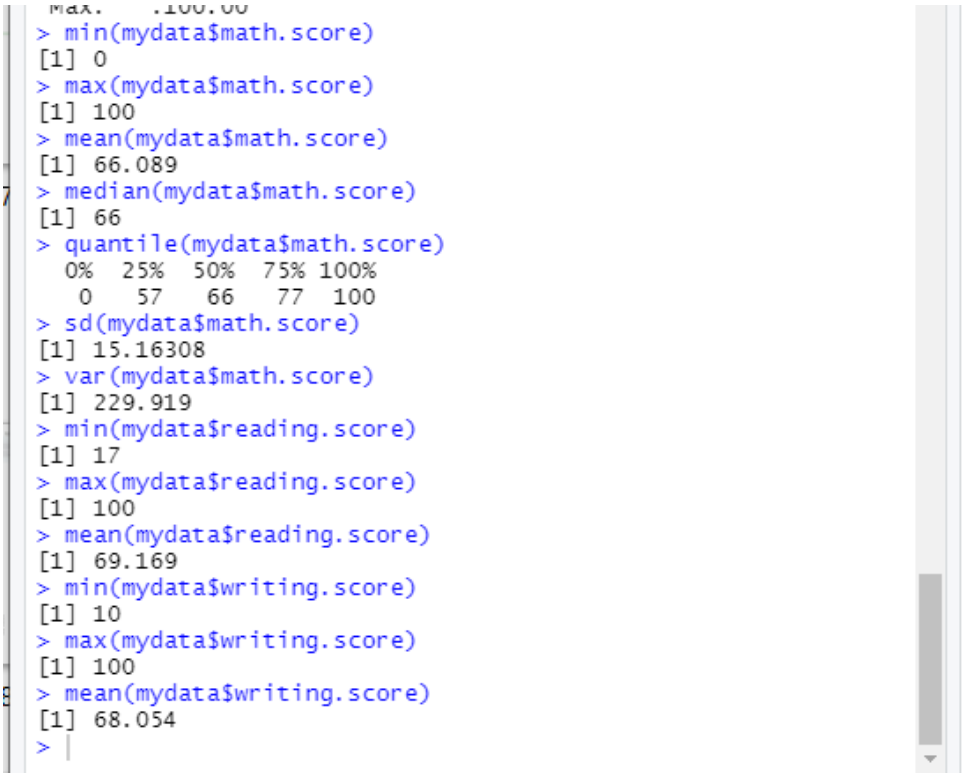
```
> summary(mydata)
    gender              race
 Length:1000        Length:1000
 Class :character   Class :character
 Mode  :character   Mode  :character


 parental.level.of.education    lunch
 Length:1000                 Length:1000
 Class :character            Class :character
 Mode  :character            Mode  :character


 test.preparation.course   math.score     reading.score
 Length:1000             Min.   :  0.00   Min.   : 17.00
 Class :character        1st Qu.: 57.00   1st Qu.: 59.00
 Mode  :character        Median : 66.00   Median : 70.00
                         Mean   : 66.09   Mean   : 69.17
                         3rd Qu.: 77.00   3rd Qu.: 79.00
                         Max.   :100.00   Max.   :100.00
 writing.score
 Min.   : 10.00
 1st Qu.: 57.75
 Median : 69.00
 Mean   : 68.05
 3rd Qu.: 79.00
 Max.   :100.00
> |
```

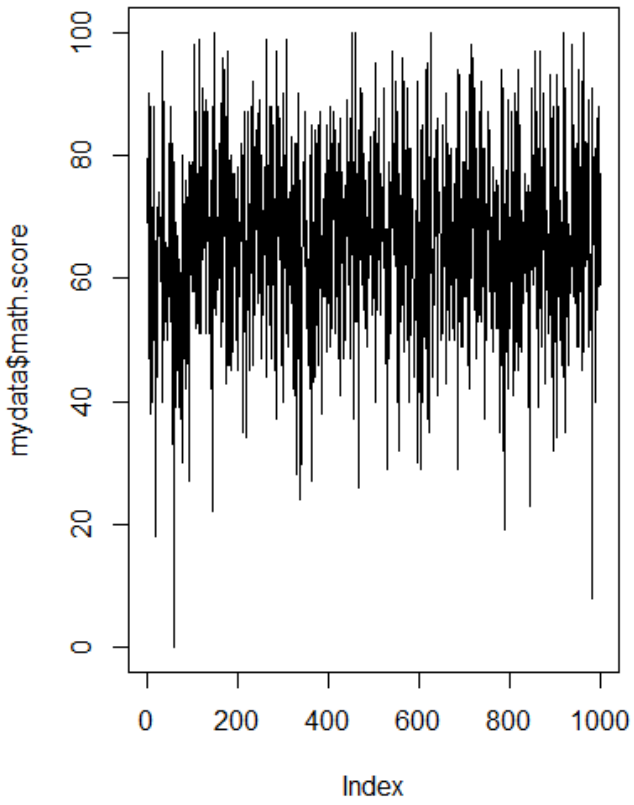
Question 4

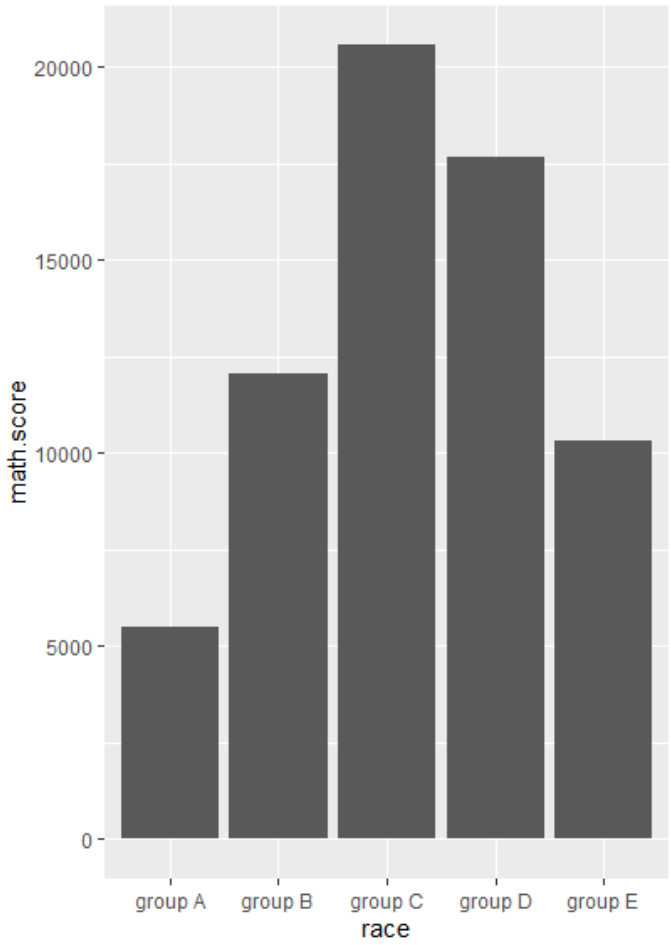
Inferential Data

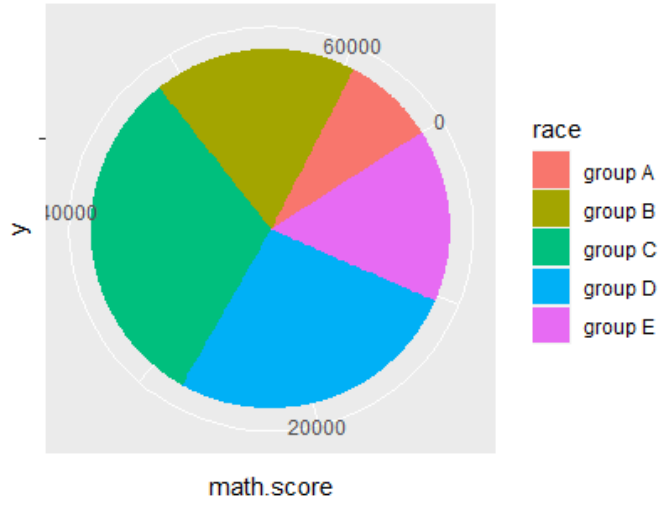


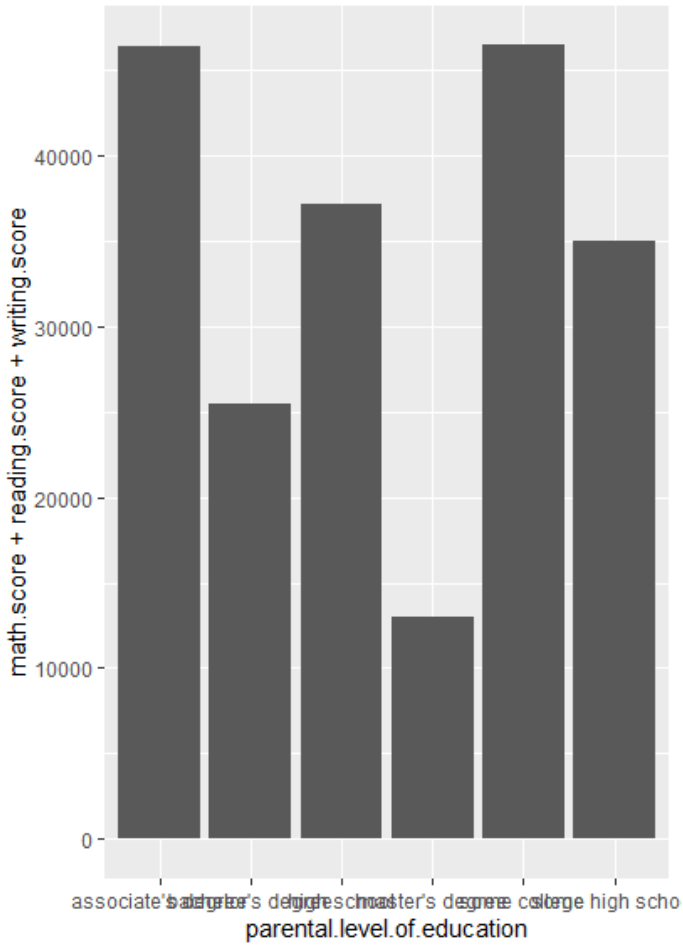
```
 Max.   :100.00
> min(mydata$math.score)
[1] 0
> max(mydata$math.score)
[1] 100
> mean(mydata$math.score)
[1] 66.089
> median(mydata$math.score)
[1] 66
> quantile(mydata$math.score)
  0%  25%  50%  75% 100%
   0   57   66   77  100
> sd(mydata$math.score)
[1] 15.16308
> var(mydata$math.score)
[1] 229.919
> min(mydata$reading.score)
[1] 17
> max(mydata$reading.score)
[1] 100
> mean(mydata$reading.score)
[1] 69.169
> min(mydata$writing.score)
[1] 10
> max(mydata$writing.score)
[1] 100
> mean(mydata$writing.score)
[1] 68.054
>
```

Descriptive


mydata$math.score
0
20
40
60
80
100
0
200
400
600
800
1000
Index


20000
15000
10000
5000
0
math.score
group A
group B
group C
group D
group E
race


60000
0
20000
y
race
group A
group B
group C
group D
group E
math.score


40000
30000
20000
10000
0
math.score + reading.score + writing.score
parental.level.of.education