GBA 464 - Assignment #1

Please upload your completed assignment to Blackboard. You do not have to upload all of the R code that you wrote, but please include your answer/output and enough of your code that I can see how you arrived at your answer. The work you submit must be your own original work. Please upload a Word doc or pdf of your work.

Load the data in the file "MR_data.csv" into R. The file contains data from the 2018 calendar year for 619 client projects of a marketing consulting firm that has four divisions.

Field	Data Definition
InvoiceNumber	Unique 6-digit number for project invoice
InvoiceDate	Date project was invoiced
Division	Division responsible for project
ProjectType	Description of project type
HoursRequired	Number of hours invoiced to client
ClientFeedbackScore	Rating of customer satisfaction with project on a scale of 100 (0 = lowest, 100 = highest) based on client feedback form
InvoiceAmount	Dollar amount of project invoice
RepeatClient	Boolean indicator: 0 = new client, 1 = client has had prior contract with firm

1) Which of the firm's divisions offered web design services in 2018 for their clients?
   ANS: Denver and Mobile

```
> table(mr$Division, mr$ProjectType)

           Customer_Survey Logo_Design Market_Research Social_Media_Analytics Web_Design
  Atlanta                2           0               5                     49          0
  Denver                 6          11              24                      2         88
  Mobile                12          37              35                      5         91
  Portland              20           0             222                     10          0
```

2) What is the average Client Feedback Score in the Atlanta division for projects which required more than 100 hours?
   ANS: 61

```
> mean(mr[mr$HoursRequired > 100 & mr$Division == "Atlanta", "ClientFeedbackScore"])
[1] 61
```

3) What is the range (max and min) and average hourly rate for market research projects across all divisions?

ANS: min = $65.75, max = $163.89, average = $132.73

```
> mr$HourlyRate <- mr$InvoiceAmount / mr$HoursRequired
> mean(mr[mr$ProjectType == "Market_Research", "HourlyRate"])
[1] 132.7266
> min(mr[mr$ProjectType == "Market_Research", "HourlyRate"])
[1] 65.74583
> max(mr[mr$ProjectType == "Market_Research", "HourlyRate"])
[1] 163.8851
```

4) What was the invoice number and the invoice date for the market research project that had the lowest hourly rate?

ANS: Invoice Number = 313125, Invoice Date = 11/2/2018

```
> mr[mr$HourlyRate == min(mr[mr$ProjectType == "Market_Research", "HourlyRate"]), c("InvoiceNumber", "InvoiceDate","HourlyRate")]
    InvoiceNumber InvoiceDate HourlyRate
551        313125   11/2/2018   65.74583
```

5) An analyst at corporate headquarters took 2018 projects across all divisions and divided them into quintiles (bottom 20%, ..., top 20%) by Client Feedback Score. Which division had more projects in the bottom 20% Client Feedback Scores than any other division?

ANS: Mobile

```
> mr$ClientFeedbackQuintile <- cut(mr$ClientFeedbackScore, quantile(mr$ClientFeedbackScore, seq(0,1,.2)), include.lowest = TRUE)
> table(mr$Division, mr$ClientFeedbackQuintile)

          [11,72] (72,84.2] (84.2,91] (91,97] (97,100]
Atlanta       44         9         2       1        0
Denver        22        56        24      20        9
Mobile        61        36        31      14       38
Portland       0        20        75      82       75
```

6) The analyst also created the table below, showing the distribution of each division's project Client Feedback Scores. For example, 30% of Portland's projects had scores between 97 and 100, which was the top 20% score group across all divisions. Submit the R code to recreate this table.

```
> t <- table(mr$Division, mr$ClientFeedbackQuintile)
> round(t / rowSums(t),2)

          [11,72] (72,84.2] (84.2,91] (91,97] (97,100]
Atlanta      0.79      0.16      0.04    0.02     0.00
Denver       0.17      0.43      0.18    0.15     0.07
Mobile       0.34      0.20      0.17    0.08     0.21
Portland     0.00      0.08      0.30    0.33     0.30
```

7) Create a conditional box plot of ClientFeedbackScore by RepeatClient. How does the customer satisfaction of new clients compare to repeat clients?
ANS: Repeat clients have an average client feedback score almost 8 points higher than new clients. The low outliers (feedback scores between 0 and 20) are all new clients.
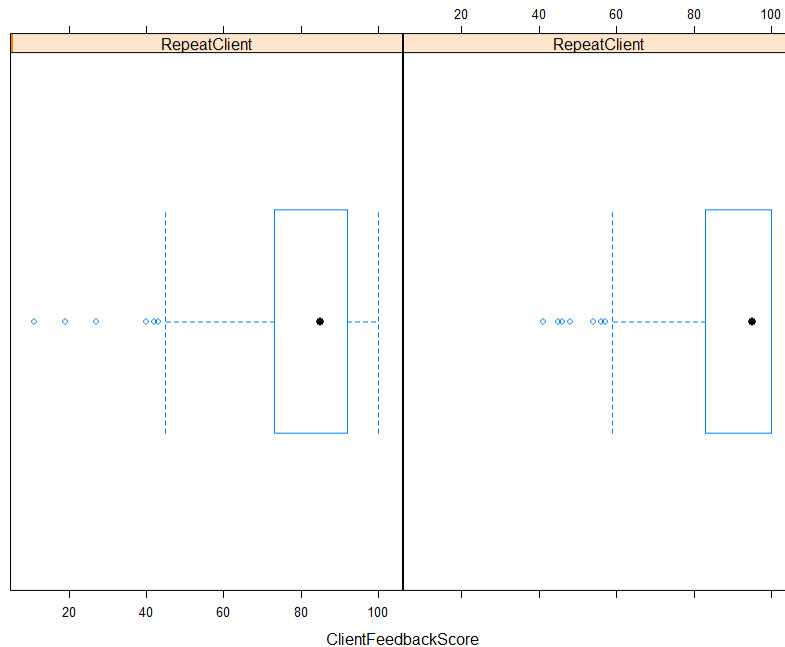
8) Create a visualization of Hours Required by Project Type. You may choose what type of chart to use, pick one that you think makes it easy to discern patterns in the data.