



ITU

Assignment 4

Xin Huang

Answers to Question 1

1a)

```
> myData <- read.csv("HW01pb1data.csv", header = FALSE)
> #exam all the columns
> class(myData$V1)
```

```
[1] "integer"
```

```
> class(myData$V2)
```

```
[1] "integer"
```

```
> class(myData$V3)
```

```
[1] "integer"
```

```
> class(myData$V4)
```

```
[1] "factor"
```

```
> class(myData$V5)
```

```
[1] "factor"
```

Given the results, we can see that column V1, V2, V3 are quantitative V4 and V5 are qualitative.

1b)

```
> #print out the levels
> levels(myData$V4)
```

```

[1] "0"          "10"         "100"        "110"        "120"
[6] "140"        "15"         "150"        "160"        "20"
[11] "200"        "25"         "30"         "35"         "40"
[16] "5"          "50"         "55"         "60"         "65"
[21] "70"         "80"         "85"         "90"         "thirty five"

```

```
> levels(myData$V5)
```

```

[1] "0"          "10"         "120"        "140"        "15"
[6] "20"         "25"         "255"        "30"         "35"
[11] "40"         "45"         "5"          "50"         "55"
[16] "60"         "70"         "80"         "twenty five"

```

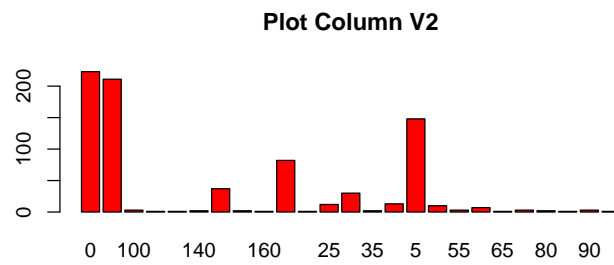
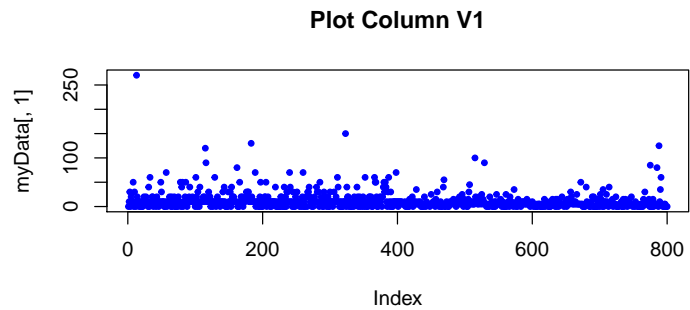
By printing out all the levels of column V4 and V5, we can see that they both contain data different data type. So when the data was read in, they can not be treated as numeric but as factors.

1c)

```

> mat <- matrix(1 : 2, nrow = 2)
> layout(mat)
> plot(myData[, 1], pch = 20, col = "blue", cex = 0.9,
+       main = "Plot Column V1")
> plot(myData[, 4], col = "red",
+       main = "Plot Column V2")

```



In the first pic, it plots column 1 scatters data on a x-y axis. It uses index and value as a x-y values In the second pic, it plots column 4 as a histogram graph. It uses factors to count how many element are in each factors.