Lab2-R

Clayton Glenn

January 22, 2018

TASKS

Task 1

Show Working Directory

```
getwd()
```

```
## [1] "C:/Users/cglen/Documents/Stat Methods/Labs/LAB2"
```

Task 2

Read Data File

```
mpg.df=read.table("EPAGAS.csv",header=TRUE,sep=",")
head(mpg.df)
```

```
##    MPG
## 1 36.3
## 2 41.0
## 3 36.9
## 4 37.1
## 5 44.9
## 6 36.8
```

Task 3

Create Z vector from MPG

```
mpg=mpg.df$MPG
z=scale(mpg)
class(z)
```

```
## [1] "matrix"
```

Verify Z-BAR and S^2vZ

```
apply(z,2,mean)
```

```
## [1] 9.706208e-17
```

```
apply(z,2,sd)
```

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

Find the values of mpg that are possible outliers

```
mpg[abs(z)>=2 & abs(z)<=3]
```

```
## [1] 30.0 42.1 31.8
```

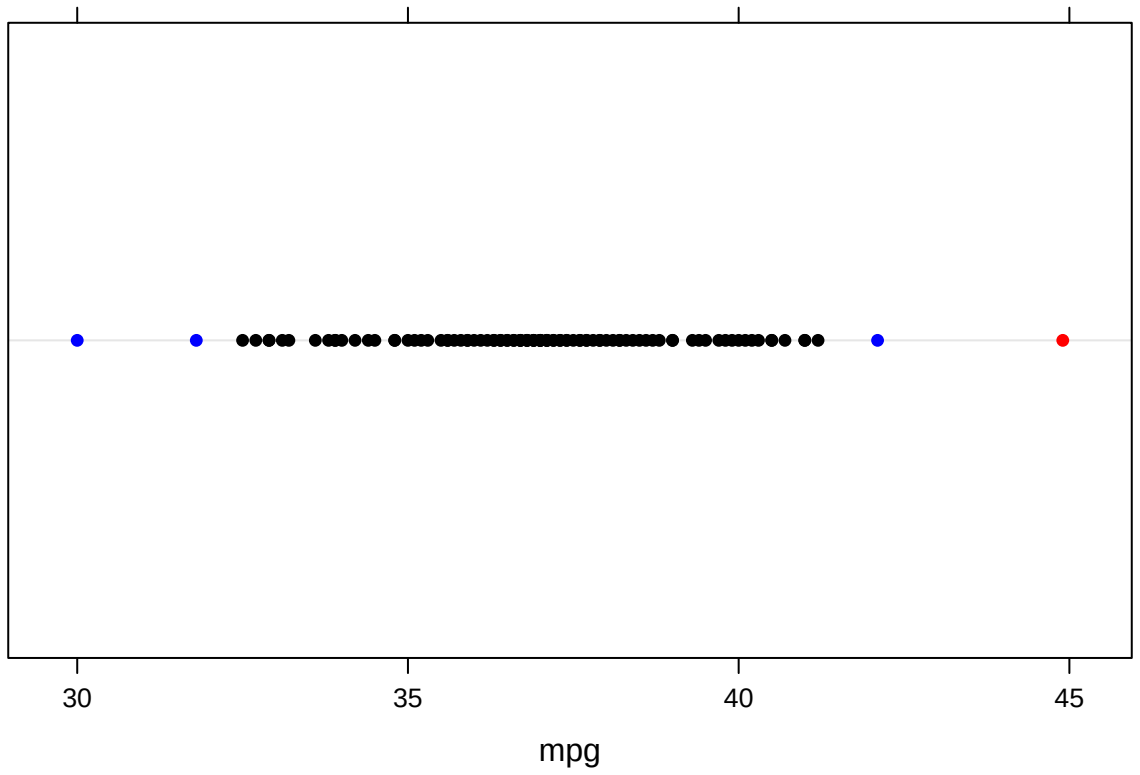
Find the values of mpg that defined as outliers

```
mpg[abs(z)>3]
```

```
## [1] 44.9
```

Lattice Dotplot with outliers and possible outliers

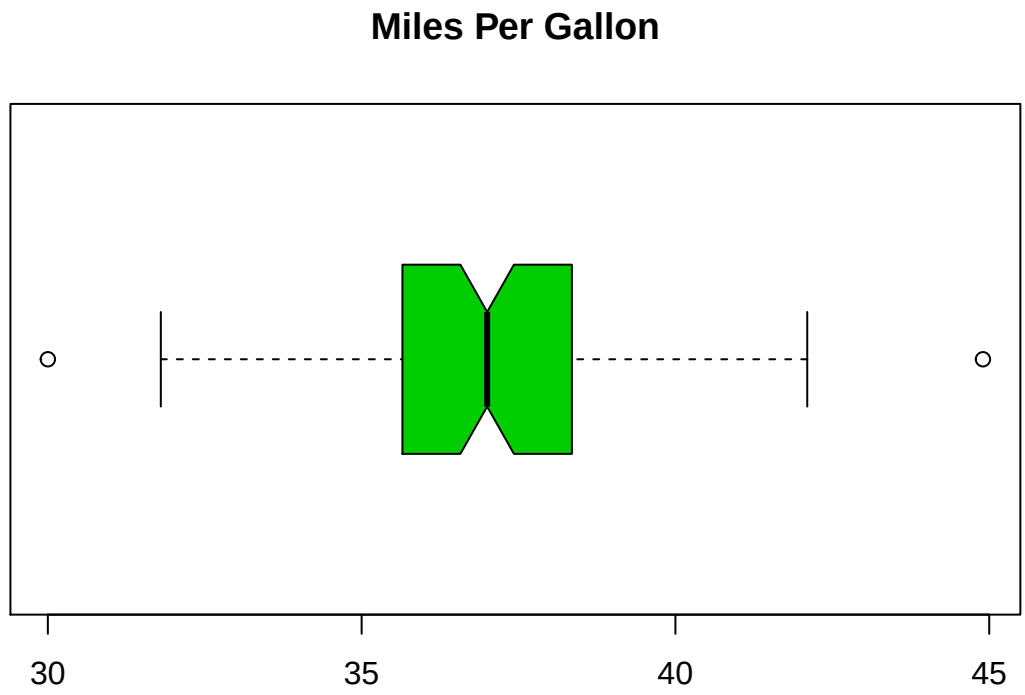
```
mycol = ifelse(abs(z)>3, "Red",
        ifelse(abs(z)>=2 &abs(z)<=3,"Blue", "Black"))
library(lattice)
dotplot(mpg,col=mycol)
```



Task 4

Boxplot of MPG

```
boxplot(mpg, data=mpg, notch=TRUE, col=(c("green3")), main="Miles Per Gallon", horizontal = TRUE)
```



Exact Proportion Within 2 STD's of the Mean.

```
sprintf("%.0f%%", length(mpg[abs(z)<2])/length(mpg) * 100)
```

```
## [1] "96%"
```

Chebyshev's Theorum

(1 - 1/(#)^2) of the Data lies within 2 STD's

```
sprintf("%.0f%%", (1 - 1/2^2)*100)
```

```
## [1] "75%"
```

According to the theorum, ATLEAST 75% of the data lies within 2 STD's. In this case 96% of the lies within 2 STD's, so Chevyshev's theorum is correct and does agree with the data.

Empirical Rule

Approximately 95% of Data lies within 2 STD's of the Mean

```
sprintf("%.0f%%", 95)
```

```
## [1] "95%"
```

The Empirical Rule responds very well to the data, due to the fact that the rule states that APPROXIMATELY 95% of the data lies within 2 STD's and according to the actual proportion of data, which is 96%, the Eprical Rule is a very close estimate to the real proportion.