Assignment 8(1): Visualization of Air quality dataset

33140 (Sahil Naphade)

17/04/2020

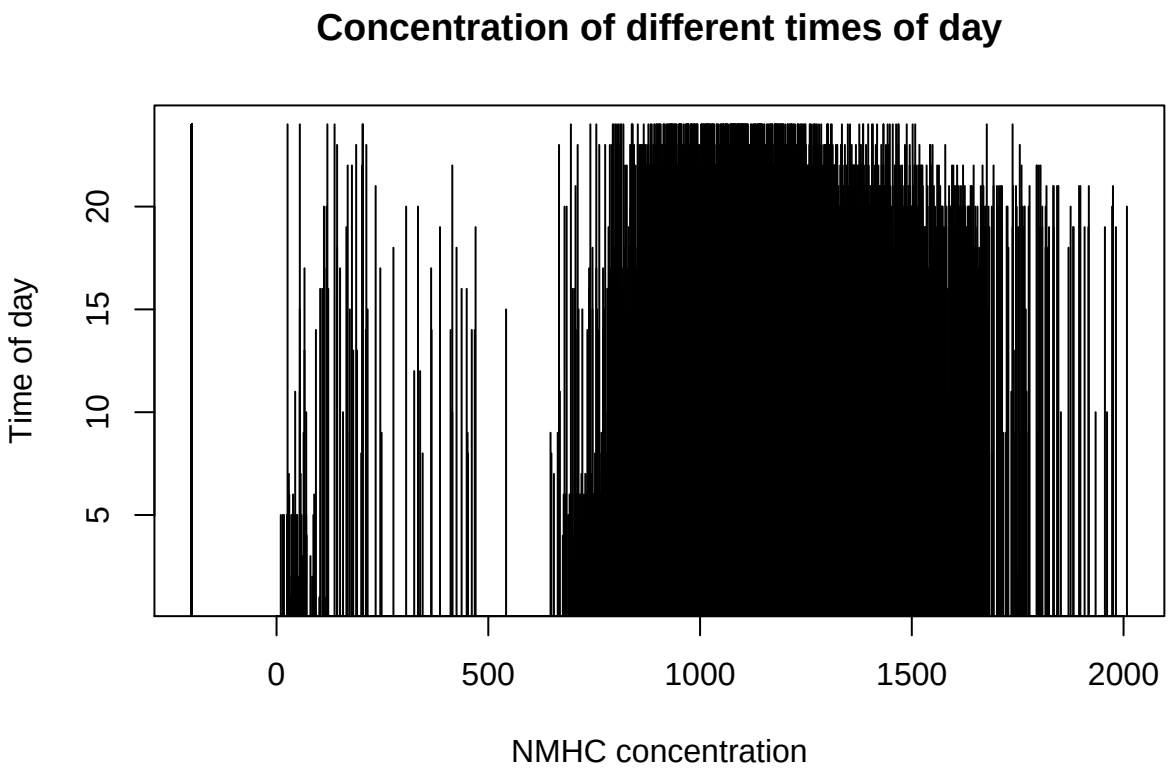
```
dataset <- read.csv2("G:/College/Sl-VI DataSets/AirQualityUCI.csv", header = T, sep =',')
View(dataset)
names(dataset)[13] <- 'Temp'
head(dataset)
```

```
##         Date     Time CO.GT. PT08.S1.CO. NMHC.GT. C6H6.GT. PT08.S2.NMHC.
## 1 10/03/2004 18.00.00      2           6     1360      150            11
## 2 10/03/2004 19.00.00      2        1292      112        9             4
## 3 10/03/2004 20.00.00      2           2     1402       88             9
## 4 10/03/2004 21.00.00      2           2     1376       80             9
## 5 10/03/2004 22.00.00      1           6     1272       51             6
## 6 10/03/2004 23.00.00      1           2     1197       38             4
##   NOx.GT. PT08.S3.NOx. NO2.GT. PT08.S4.NO2. PT08.S5.O3. Temp   RH AH R1 R2 R3
## 1       9         1046     166         1056         113 1692 1268 13  6 48  9
## 2     955          103    1174           92        1559  972   13  3 47  7  0
## 3       0          939     131         1140         114 1555 1074 11  9 54  0
## 4       2          948     172         1092         122 1584 1203 11  0 60  0
## 5       5          836     131         1205         116 1490 1110 11  2 59  6
## 6       7          750      89         1337          96 1393  949 11  2 59  2
##     R4   R5
## 1    0 7578
## 2 7255   NA
## 3    0 7502
## 4    0 7867
## 5    0 7888
## 6    0 7848
```

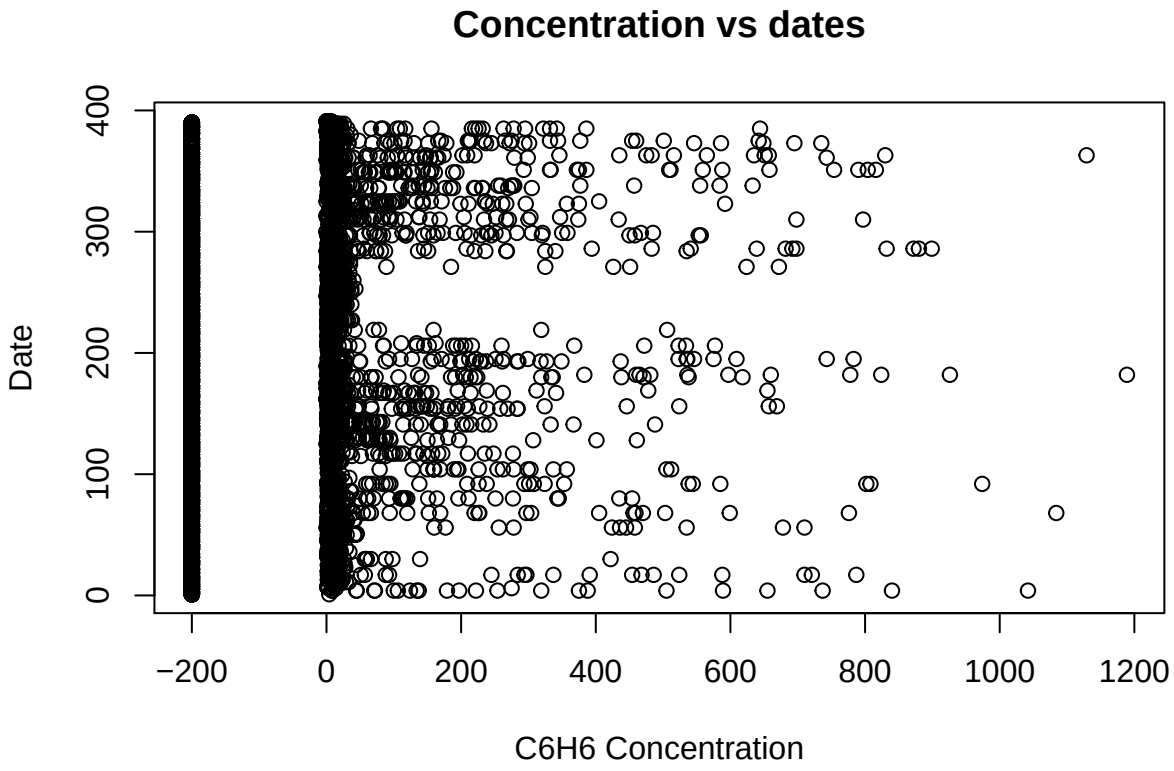
```
# Generic Plots
help(plot)
```

```
## starting httpd help server ... done
```

```
plot(dataset$NMHC.GT.,dataset$Time,main = "Concentration of different times of day",
     xlab = "NMHC concentration", ylab = "Time of day", type = "h") # Histogram type
```



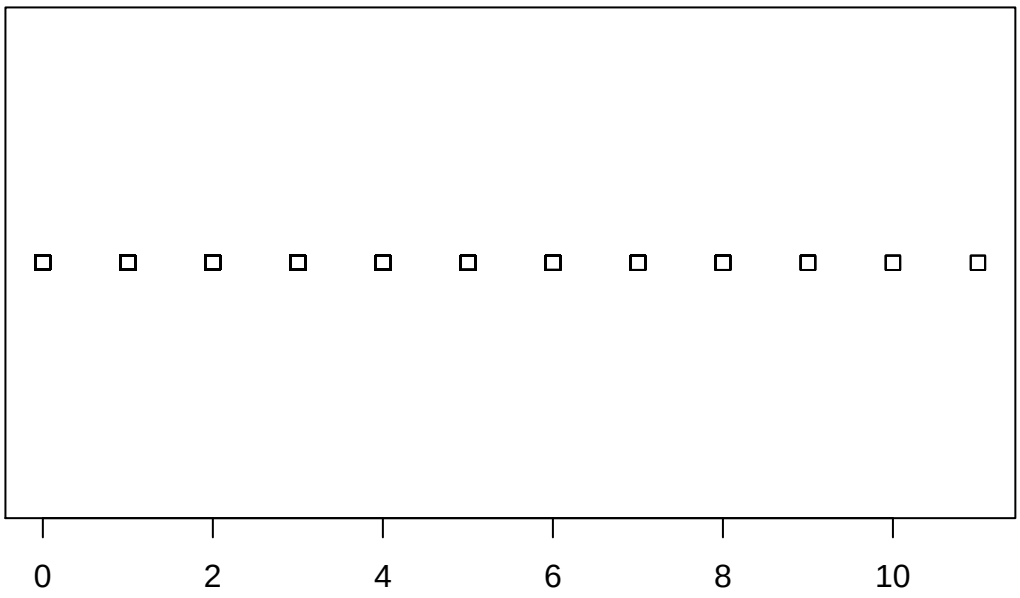

```
plot(dataset$C6H6.GT.,dataset$Date,main = "Concentration vs dates",
     xlab = "C6H6 Concentration", ylab = "Date",type = "p") # Point type graph
```




```
# Strip Charts
# 1. for CO.GT.
help(stripchart)
dataset$CO.GT. <- replace(dataset$CO.GT.,dataset$CO.GT. == -200,NA)
summary(dataset$CO.GT.)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   1.000   1.000   1.701   2.000  11.000    1683
```

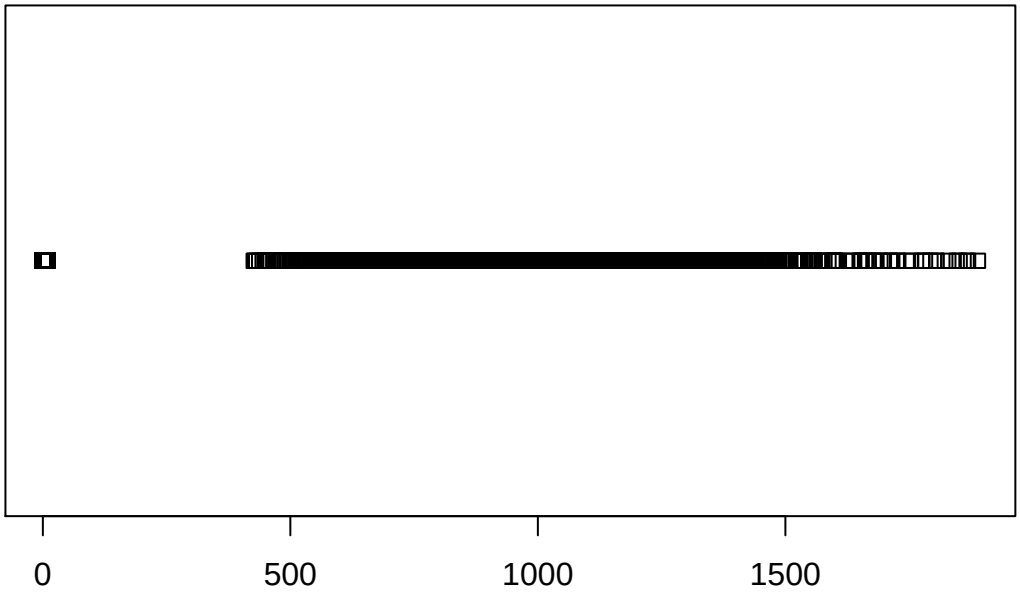
```
dataset$CO.GT. <- replace(dataset$CO.GT.,is.na(dataset$CO.GT.),1) # replace with the median value
stripchart(dataset$CO.GT.)
```



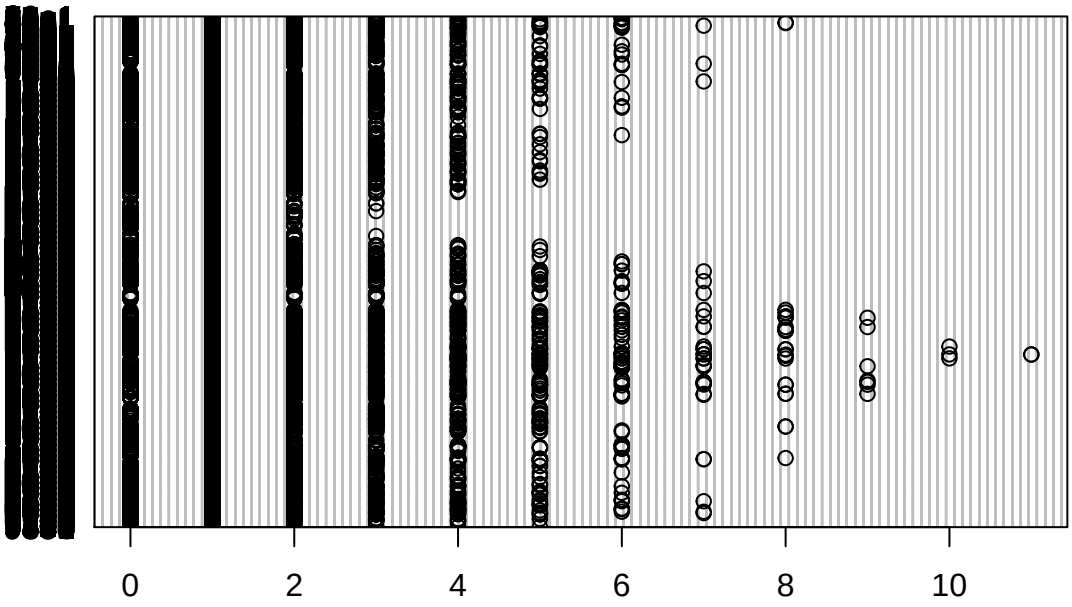

```
# 2. for NOX.GT.

dataset$NOx.GT. <- replace(dataset$NOx.GT.,dataset$NOx.GT. == -200.0,NA)
summary(dataset$NOx.GT.)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##     0.0     2.0     6.0   212.9     9.0  1889.0      61
```

```
dataset$NOx.GT. <- replace(dataset$NOx.GT.,is.na(dataset$NOx.GT.),6) # replace with median value
stripchart(dataset$NOx.GT.)
```



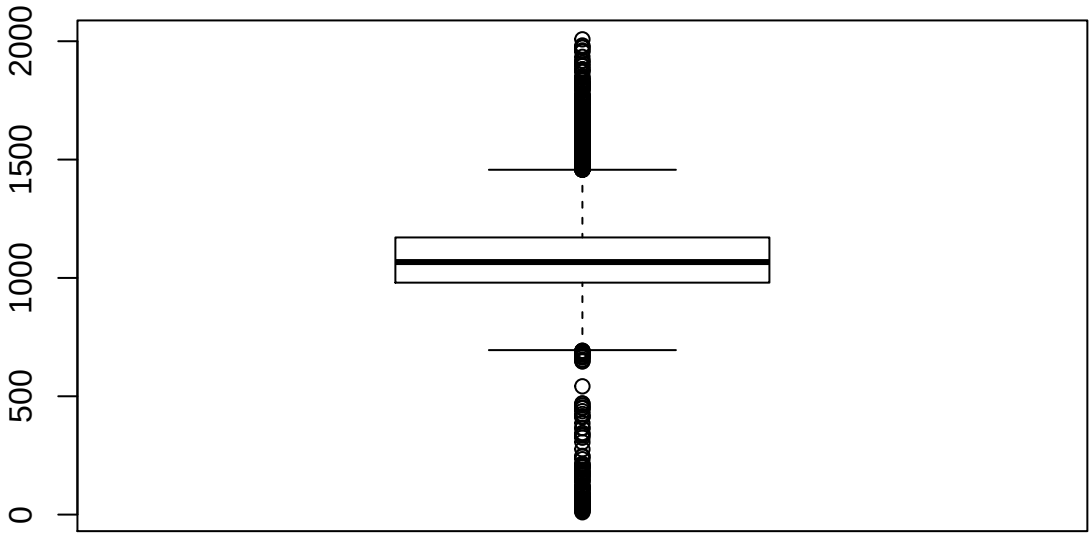

```
#Dotcharts
dotchart(t(dataset$CO.GT.))
```




```
#Boxplot
dataset$NMHC.GT. <- replace(dataset$NMHC.GT.,dataset$NMHC.GT. == -200.0,NA)
summary(dataset$NMHC.GT.)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##      10     936    1067    1092    1238    2008    2322
```

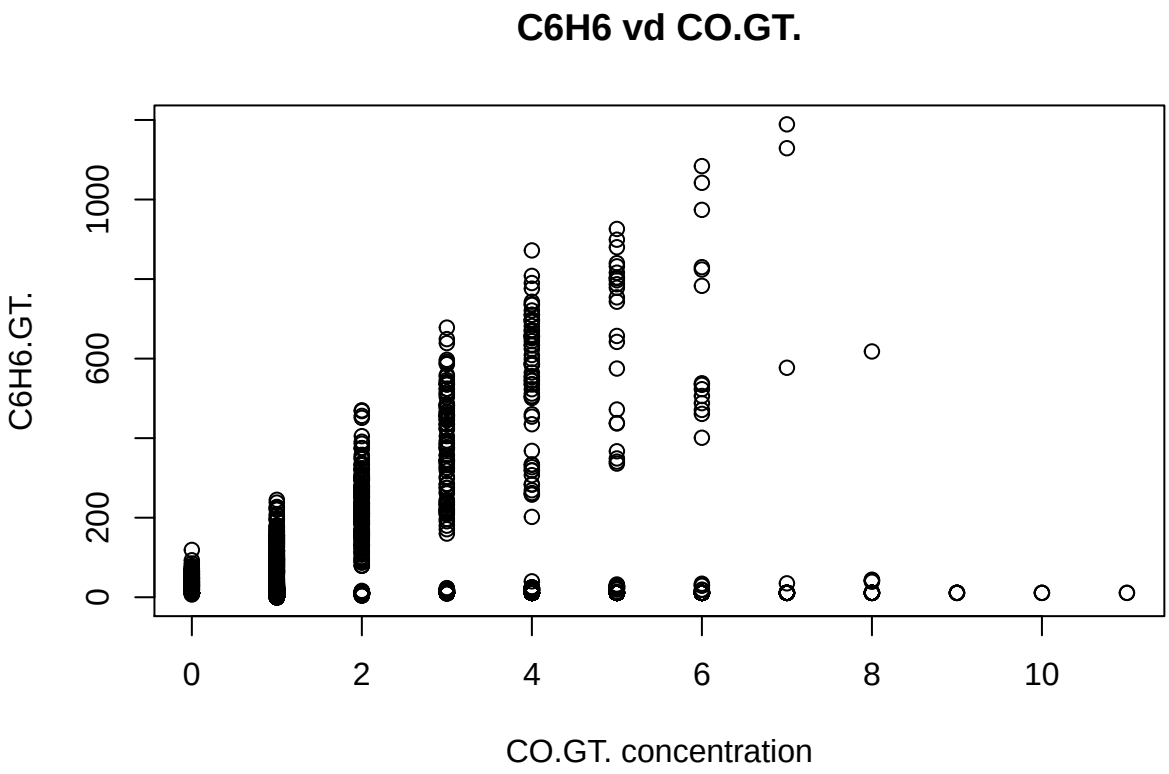
```
dataset$NMHC.GT. <- replace(dataset$NMHC.GT.,is.na(dataset$NMHC.GT.),1067) #replace with the median val
boxplot(dataset$NMHC.GT.)
```



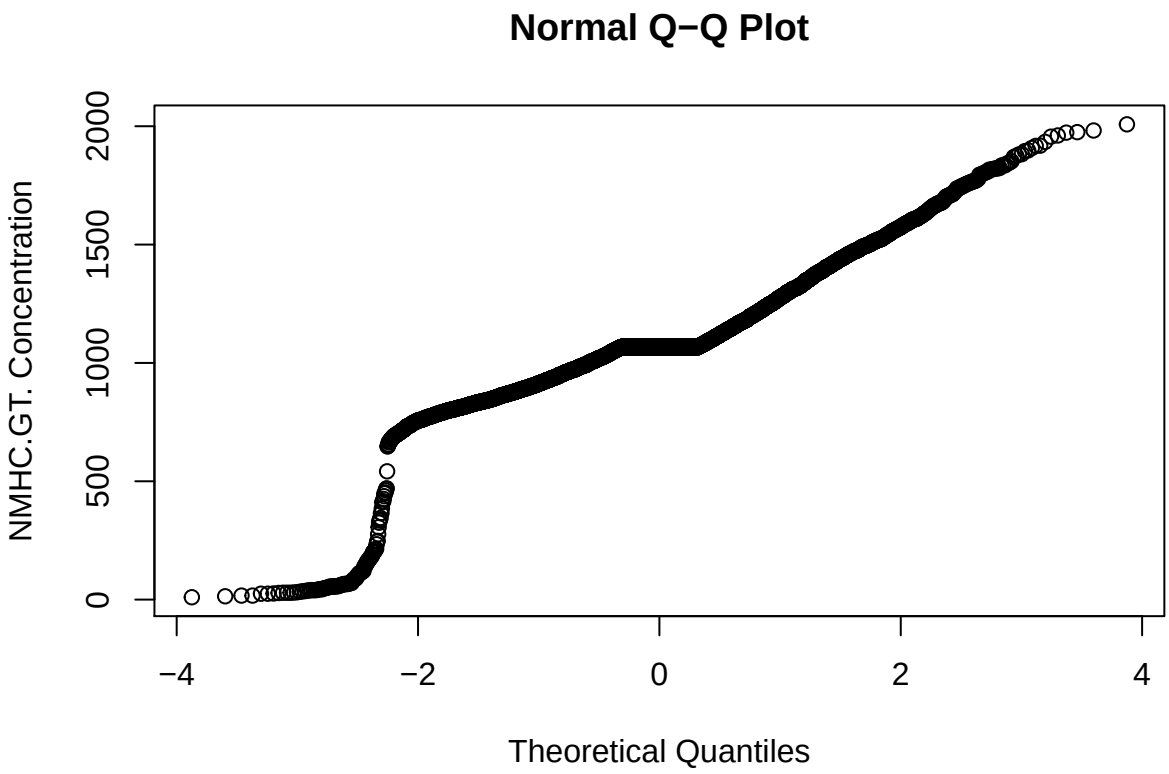

```
#Scatter Plots
dataset$C6H6.GT. <- replace(dataset$C6H6.GT.,dataset$C6H6.GT. == -200,NA)
summary(dataset$C6H6.GT.)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##    0.00    5.00   11.00   68.88   40.75 1189.00    6487
```

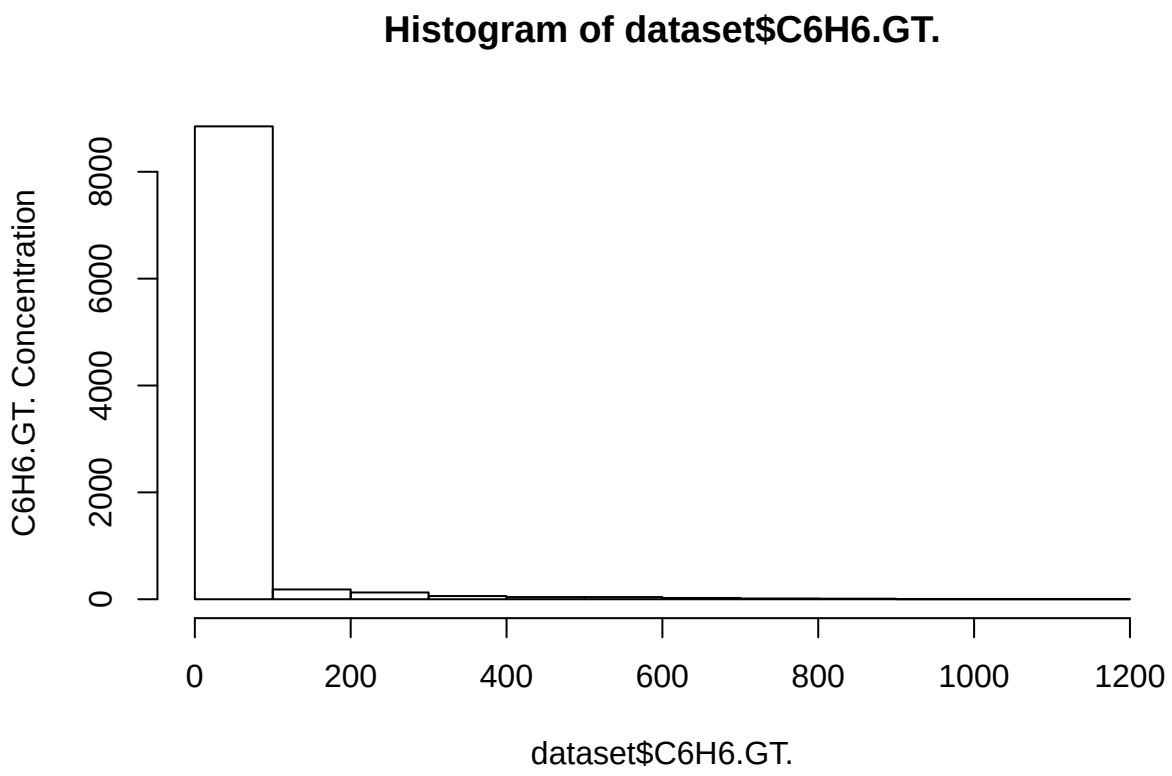
```
dataset$C6H6.GT. <- replace(dataset$C6H6.GT.,is.na(dataset$C6H6.GT.),11) # replace with median value
plot(dataset$CO.GT.,dataset$C6H6.GT.,xlab = "CO.GT. concentration", ylab = "C6H6.GT.",main = "C6H6 vd C
```




```
#Normal QQ Plots
qqnorm(dataset$NMHC.GT., ylab = "NMHC.GT. Concentration")
```

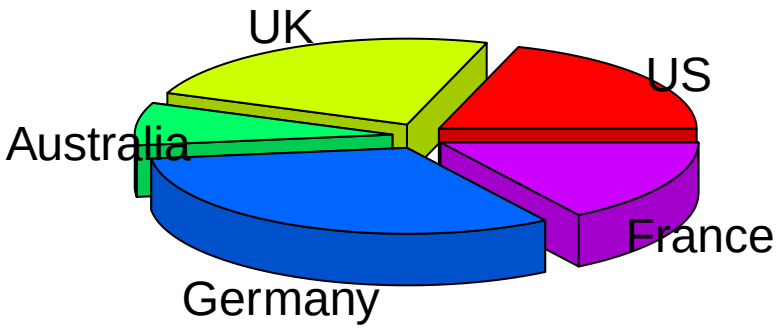



```
# Histograms
hist(dataset$C6H6.GT., ylab = "C6H6.GT. Concentration")
```




```
# Trying 3D Exploded Pie Chart
library(plotrix)
slices <- c(10, 12, 4, 16, 8)
lbls <- c("US", "UK", "Australia", "Germany", "France")
pie3D(slices,labels=lbls,explode=0.1,
      main="Pie Chart of Countries ")
```

Pie Chart of Countries




```
yaxis=c(10,20,30,40,50,60,70,80.90,100)
summary(dataset)
```

```
##          Date             Time          CO.GT.          PT08.S1.CO.
##  01/01/2005:  24   00.00.00: 390   Min.   : 0.000   Min.   :-200.0
##  01/02/2005:  24   01.00.00: 390   1st Qu.: 1.000   1st Qu.:   3.0
##  01/03/2005:  24   02.00.00: 390   Median : 1.000   Median :   6.0
##  01/04/2004:  24   03.00.00: 390   Mean   : 1.575   Mean   : 240.8
##  01/04/2005:  24   04.00.00: 390   3rd Qu.: 2.000   3rd Qu.:   9.0
##  01/05/2004:  24   05.00.00: 390   Max.   :11.000   Max.   :2040.0
##  (Other)   :9213   (Other) :7017
##     NMHC.GT.        C6H6.GT.        PT08.S2.NMHC.         NOx.GT.
##  Min.   :  10   Min.   :   0.00   Min.   :-200.000   Min.   :   0.0
##  1st Qu.: 980   1st Qu.:  11.00   1st Qu.:   3.000   1st Qu.:   2.0
##  Median :1067   Median :  11.00   Median :   6.000   Median :   6.0
##  Mean   :1086   Mean   :  28.75   Mean   :   1.601   Mean   : 211.5
##  3rd Qu.:1171   3rd Qu.:  11.00   3rd Qu.:  11.000   3rd Qu.:   9.0
##  Max.   :2008   Max.   :1189.00   Max.   :  63.000   Max.   :1889.0
##
##   PT08.S3.NOx.        NO2.GT.         PT08.S4.NO2.     PT08.S5.O3.
##  Min.   :-200.0   Min.   :-200.0   Min.   :-200.0   Min.   :-200.0
##  1st Qu.: 383.0   1st Qu.: 110.0   1st Qu.: 145.0   1st Qu.:  84.0
##  Median : 791.0   Median : 247.0   Median : 695.0   Median : 124.0
##  Mean   : 678.4   Mean   : 365.9   Mean   : 589.6   Mean   : 416.9
##  3rd Qu.:1040.0   3rd Qu.: 600.0   3rd Qu.: 891.0   3rd Qu.: 212.0
##  Max.   :2214.0   Max.   :2331.0   Max.   :2683.0   Max.   :2684.0
```

```
##
##       Temp             RH               AH                R1
##  Min.   :-200   Min.   :-200.0   Min.   :-200.000   Min.   :-200.000
##  1st Qu.: 967   1st Qu.:  35.0   1st Qu.:   6.000   1st Qu.:   2.000
##  Median :1312   Median : 816.0   Median :  12.000   Median :   6.000
##  Mean   :1265   Mean   : 763.9   Mean   :   5.715   Mean   :   6.279
##  3rd Qu.:1595   3rd Qu.:1177.0   3rd Qu.:  20.000   3rd Qu.:   9.000
##  Max.   :2775   Max.   :2523.0   Max.   :  44.000   Max.   :  87.000
##
##        R2                R3              R4             R5
##  Min.   :-200.00   Min.   :0.000   Min.   :   0   Min.   :   3
##  1st Qu.:   9.00   1st Qu.:1.000   1st Qu.:   0   1st Qu.:2990
##  Median :  40.00   Median :3.000   Median :   1   Median :5341
##  Mean   :  30.77   Mean   :3.679   Mean   :1207   Mean   :5245
##  3rd Qu.:  57.00   3rd Qu.:6.000   3rd Qu.:   1   3rd Qu.:7684
##  Max.   :  88.00   Max.   :9.000   Max.   :9996   Max.   :9998
##  NA's   :61        NA's   :366     NA's   :366    NA's   :2442
```

```r
barplot(dataset$Temp,main = "Temperature")
```