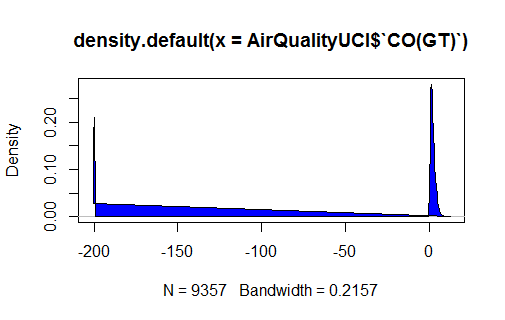
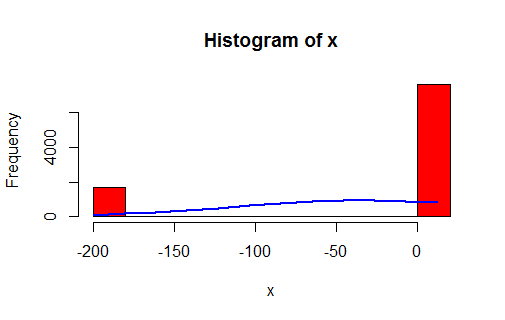
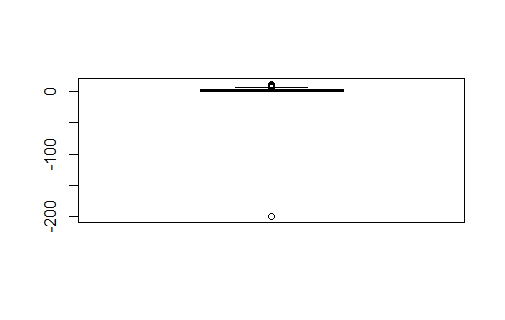
1: hist(AirQualityUCI$`PT08.S1(CO)`)

> summary(AirQualityUCI$`CO(GT)`)

Min. 1st Qu. Median Mean 3rd Qu. Max.

-200.00 0.60 1.50 -34.21 2.60 11.90





1b)

|  |
| --- |
| x<-AirQualityUCI$`PT08.S1(CO)`  > h<-hist(x,breaks=10,col="red")  > xfit<-seq(min(x),max(x),length=2000)  > yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))  > yfit<-yfit\*diff(h$mids[1:2]\*length(x))  > lines(xfit,yfit) |
|  |
| |  | | --- | | > | |

|  |
| --- |
| summary(AirQualityUCI$`PT08.S1(CO)`)  Min. 1st Qu. Median Mean 3rd Qu. Max.  -200 921 1052 1049 1221 2040 |
|  |
| |  | | --- | |  | |

3)

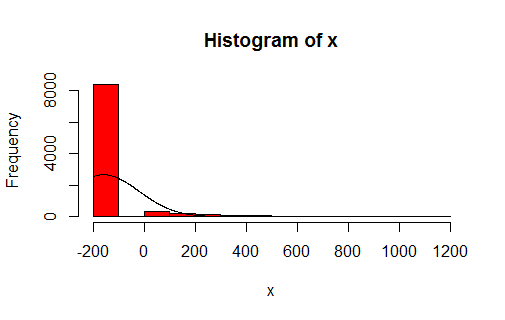
1c)

|  |
| --- |
| xfit<-seq(min(x),max(x),length=64)  > yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))  > yfit<-yfit\*diff(h$mids[1:2]\*length(x))  > lines(xfit,yfit) |
|  |
| |  | | --- | | > | |

1d) summary(AirQualityUCI$`NMHC(GT)`)

Min. 1st Qu. Median Mean 3rd Qu. Max.

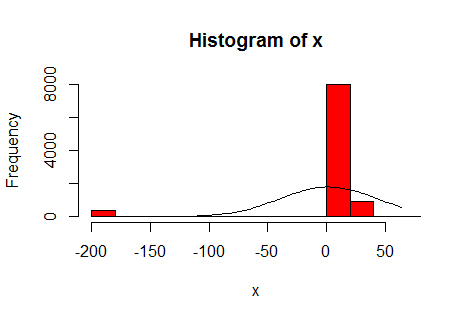
-200.0 -200.0 -200.0 -159.1 -200.0 1189.0



1e) summary(AirQualityUCI$`C6H6(GT)`)

Min. 1st Qu. Median Mean 3rd Qu. Max.

-200.000 4.005 7.887 1.866 13.636 63.741



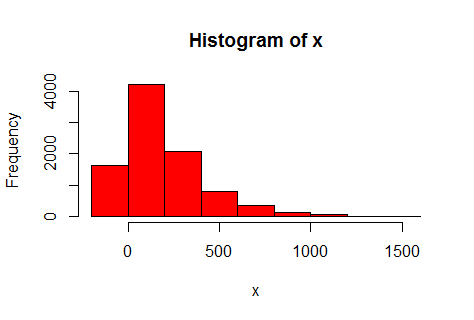
1f

|  |
| --- |
| summary(AirQualityUCI$`PT08.S2(NMHC)`)  Min. 1st Qu. Median Mean 3rd Qu. Max.  -200.0 711.0 894.5 894.5 1104.8 2214.0 |
|  |
| |  | | --- | | > | |

1g summary(AirQualityUCI$`NOx(GT)`)

Min. 1st Qu. Median Mean 3rd Qu. Max.

-200.0 50.0 141.0 168.6 284.2 1479.0



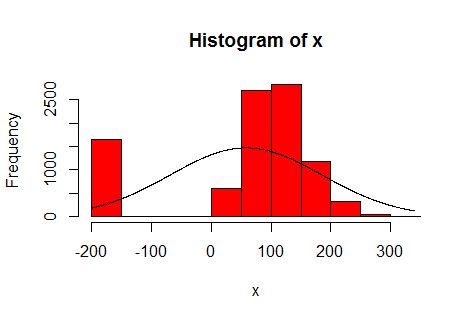
1h)

|  |
| --- |
| summary(AirQualityUCI$`PT08.S3(NOx)`)  Min. 1st Qu. Median Mean 3rd Qu. Max.  -200.0 637.0 794.2 794.9 960.2 2682.8 |
|  |
| |  | | --- | | > | |

1i) summary(AirQualityUCI$`NOx(GT)`)

Min. 1st Qu. Median Mean 3rd Qu. Max.

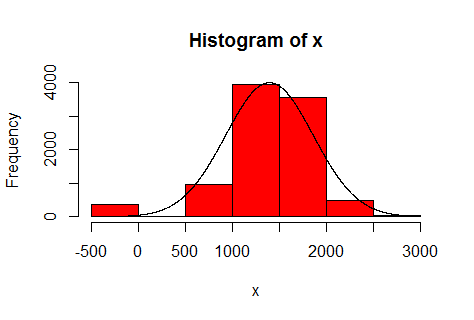
-200.0 50.0 141.0 168.6 284.2 1479.0

> 

1j) summary(AirQualityUCI$`PT08.S4(NO2)`)

Min. 1st Qu. Median Mean 3rd Qu. Max.

-200 1185 1446 1391 1662 2775



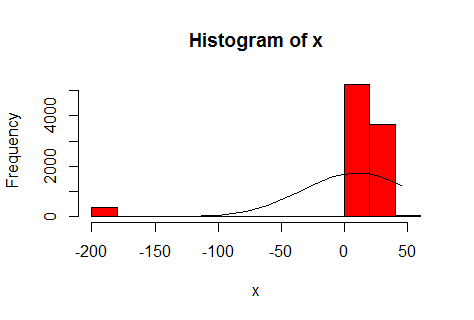
1K)

|  |
| --- |
| summary(AirQualityUCI$`PT08.S5(O3)`)  Min. 1st Qu. Median Mean 3rd Qu. Max.  -200.0 699.8 942.0 975.0 1255.2 2522.8 |
|  |
| |  | | --- | | > | |

1L) 1l) summary(AirQualityUCI$T)

Min. 1st Qu. Median Mean 3rd Qu. Max.

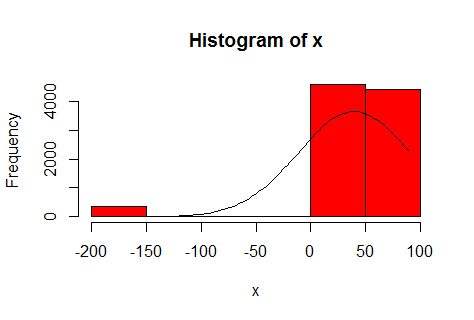
-200.000 10.950 17.200 9.777 24.075 44.600



1m) summary(AirQualityUCI$RH)

Min. 1st Qu. Median Mean 3rd Qu. Max.

-200.00 34.05 48.55 39.48 61.88 88.72

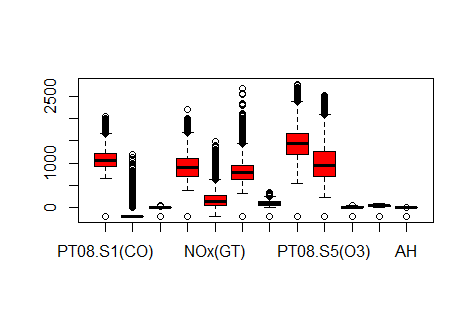


1n)

|  |
| --- |
| summary(AirQualityUCI$AH)  Min. 1st Qu. Median Mean 3rd Qu. Max.  -200.0000 0.6923 0.9768 -6.8376 1.2962 2.2310 |
|  |
| |  | | --- | | > | |

BOX plots for univariate analysis of each columns

boxplot(AirQualityUCI[,-1:-3],col="red")



**3. Check for missing values in all columns.**

**Ans: No missing values in any columns as given by R command below**

x3<-is.na(AirQualityUCI)

> which(x3==TRUE)

integer(0)

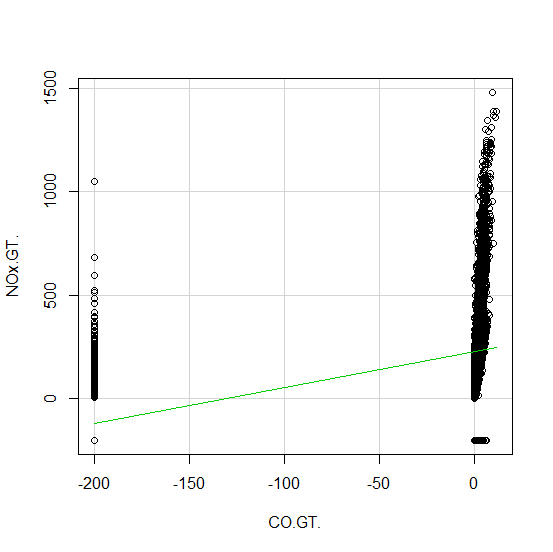
5. Create bi-variate analysis for all relationships.

Coefficient of Co-relation

x4<-AirQualityUCI[,-1:-2]

> cor(x4)

|  |
| --- |
| CO(GT) PT08.S1(CO) NMHC(GT) C6H6(GT) PT08.S2(NMHC) NOx(GT)  CO(GT) 1.00000000 0.04141486 1.283512e-01 -0.031377278 0.02993938 0.526450315  PT08.S1(CO) 0.04141486 1.00000000 1.700092e-01 0.852658739 0.93310134 0.278028873  NMHC(GT) 0.12835117 0.17000923 1.000000e+00 0.037328920 0.11009729 -0.004412891  C6H6(GT) -0.03137728 0.85265874 3.732892e-02 1.000000000 0.76740055 -0.001162976  PT08.S2(NMHC) 0.02993938 0.93310134 1.100973e-01 0.767400548 1.00000000 0.331330778  NOx(GT) 0.52645032 0.27802887 -4.412891e-03 -0.001162976 0.33133078 1.000000000  PT08.S3(NOx) -0.08998059 0.08693118 4.883223e-02 0.512154487 -0.07374833 -0.436082800  NO2(GT) 0.67113955 0.15405777 1.033448e-01 -0.010971438 0.17656903 0.817137607  PT08.S4(NO2) -0.07372055 0.84513309 1.626894e-01 0.774649129 0.87476071 0.035579549  PT08.S5(O3) 0.08031560 0.89243569 1.011892e-01 0.641305687 0.90990914 0.461915624  T -0.06895175 0.75480579 -7.577685e-06 0.971370209 0.66898434 -0.138457415  RH -0.04823058 0.74534354 8.287924e-03 0.925067595 0.58577531 -0.053008136  AH -0.04589222 0.76486567 1.250011e-02 0.984555673 0.64653471 -0.095840633  PT08.S3(NOx) NO2(GT) PT08.S4(NO2) PT08.S5(O3) T RH  CO(GT) -0.08998059 0.67113955 -0.07372055 0.0803156 -6.895175e-02 -0.048230582  PT08.S1(CO) 0.08693118 0.15405777 0.84513309 0.8924357 7.548058e-01 0.745343536  NMHC(GT) 0.04883223 0.10334476 0.16268939 0.1011892 -7.577685e-06 0.008287924  C6H6(GT) 0.51215449 -0.01097144 0.77464913 0.6413057 9.713702e-01 0.925067595  PT08.S2(NMHC) -0.07374833 0.17656903 0.87476071 0.9099091 6.689843e-01 0.585775310  NOx(GT) -0.43608280 0.81713761 0.03557955 0.4619156 -1.384574e-01 -0.053008136  PT08.S3(NOx) 1.00000000 -0.25621745 0.12267209 -0.2089346 5.880610e-01 0.573513202  NO2(GT) -0.25621745 1.00000000 -0.02209248 0.2534693 -8.408449e-02 -0.081299722  PT08.S4(NO2) 0.12267209 -0.02209248 1.00000000 0.7236700 7.550529e-01 0.640685489  PT08.S5(O3) -0.20893460 0.25346934 0.72367002 1.0000000 5.036591e-01 0.524921583  T 0.58806103 -0.08408449 0.75505285 0.5036591 1.000000e+00 0.885909850  RH 0.57351320 -0.08129972 0.64068549 0.5249216 8.859098e-01 1.000000000  AH 0.62157636 -0.06042311 0.69188886 0.5194315 9.809977e-01 0.943997165  AH  CO(GT) -0.04589222  PT08.S1(CO) 0.76486567  NMHC(GT) 0.01250011  C6H6(GT) 0.98455567  PT08.S2(NMHC) 0.64653471  NOx(GT) -0.09584063  PT08.S3(NOx) 0.62157636  NO2(GT) -0.06042311  PT08.S4(NO2) 0.69188886  PT08.S5(O3) 0.51943155  T 0.98099773  RH 0.94399717  AH 1.00000000 |
|  |
| |  | | --- | |  | |

 SCATTER PLOT of CO(GT) vs NOX(GT)

SCATTER PLOT OF VARAIABLES with CO RELATION

