

ACADGILD ASSIGNMENT 10

SESSION 10: Correlations

1. Import dataset from the following link:

<https://archive.ics.uci.edu/ml/machine-learning-databases/00360/> Perform the below written operations:

a. Read the file in Zip format and get it into R

Answer:

```
library(readxl)
AirQualityUCI <-
read_excel("C:/Users/manojChowdary/Downloads/AirQualityUCI.xlsx"
)
View(AirQualityUCI)
str(AirQualityUCI)
```

Output from R-console

```
> View(AirQualityUCI)
> str(AirQualityUCI)
Classes 'tbl_df', 'tbl' and 'data.frame': 9357 obs. of
15 variables:
 $ Date      : POSIXct, format: "2004-03-10" "2004-
03-10" "2004-03-10" ...
 $ Time      : POSIXct, format: "1899-12-31
18:00:00" "1899-12-31 19:00:00" "1899-12-31 20:00:00"
...
 $ CO(GT)    : num  2.6 2 2.2 2.2 1.6 1.2 1.2 1 0.9
0.6 ...
 $ PT08.S1(CO) : num  1360 1292 1402 1376 1272 ...
 $ NMHC(GT)   : num  150 112 88 80 51 38 31 31 24 19
...
 $ C6H6(GT)   : num  11.88 9.4 9 9.23 6.52 ...
 $ PT08.S2(NMHC): num  1046 955 939 948 836 ...
 $ NOx(GT)    : num  166 103 131 172 131 89 62 62 45
-200 ...
 $ PT08.S3(NOx) : num  1056 1174 1140 1092 1205 ...
 $ NO2(GT)     : num  113 92 114 122 116 96 77 76 60 -
200 ...
 $ PT08.S4(NO2) : num  1692 1559 1554 1584 1490 ...
 $ PT08.S5(O3) : num  1268 972 1074 1203 1110 ...
 $ T          : num  13.6 13.3 11.9 11 11.2 ...
 $ RH         : num  48.9 47.7 54 60 59.6 ...
```

```
$ AH          : num  0.758 0.725 0.75 0.787 0.789 ...
```

b.Create Univariate for all the columns.

Answer:

```
library(readxl)
AirQualityUCI <-
read_excel("C:/Users/manojChowdary/Downloads/AirQualityUCI.xlsx"
)
View(AirQualityUCI)
dim(AirQualityUCI)
summary(AirQualityUCI)
AirQualityUCI[AirQualityUCI== -200.0]<-NA
for(i in
1:ncol(AirQualityUCI)){ AirQualityUCI[is.na(AirQualityUCI[,i]),i] <-
mean(AirQualityUCI[,i], na.rm = TRUE)}
summary(AirQualityUCI)
AirQualityUCI[7:14,]
hist(AirQualityUCI$`NOx(GT)` ,col="red")
```

Output from R-console

```
> dim(AirQualityUCI)
[1] 9357  15

> summary(AirQualityUCI)
      Date
Min.   :2004-03-10 00:00:00
1st Qu.:2004-06-16 00:00:00
Median :2004-09-21 00:00:00
Mean    :2004-09-21 04:30:05
3rd Qu.:2004-12-28 00:00:00
Max.    :2005-04-04 00:00:00
      Time
Min.   :1899-12-31 00:00:00
1st Qu.:1899-12-31 05:00:00
Median :1899-12-31 11:00:00
Mean    :1899-12-31 11:29:55
3rd Qu.:1899-12-31 18:00:00
Max.    :1899-12-31 23:00:00
      CO(GT)      PT08.S1(CO)
Min.   : -200.00  Min.   : -200
1st Qu.:   0.60   1st Qu.:  921
```

Median :	1.50	Median :	1052
Mean :	-34.21	Mean :	1049
3rd Qu.:	2.60	3rd Qu.:	1221
Max. :	11.90	Max. :	2040

NMHC(GT)		C6H6(GT)	
Min. :	-200.0	Min. :	-200.000
1st Qu.:	-200.0	1st Qu.:	4.005
Median :	-200.0	Median :	7.887
Mean :	-159.1	Mean :	1.866
3rd Qu.:	-200.0	3rd Qu.:	13.636
Max. :	1189.0	Max. :	63.741

PT08.S2(NMHC)		NOx(GT)	
Min. :	-200.0	Min. :	-200.0
1st Qu.:	711.0	1st Qu.:	50.0
Median :	894.5	Median :	141.0
Mean :	894.5	Mean :	168.6
3rd Qu.:	1104.8	3rd Qu.:	284.2
Max. :	2214.0	Max. :	1479.0

PT08.S3(NOx)		NO2(GT)	
Min. :	-200.0	Min. :	-200.00
1st Qu.:	637.0	1st Qu.:	53.00
Median :	794.2	Median :	96.00
Mean :	794.9	Mean :	58.14
3rd Qu.:	960.2	3rd Qu.:	133.00
Max. :	2682.8	Max. :	339.70

PT08.S4(NO2)		PT08.S5(O3)	
Min. :	-200	Min. :	-200.0
1st Qu.:	1185	1st Qu.:	699.8
Median :	1446	Median :	942.0
Mean :	1391	Mean :	975.0
3rd Qu.:	1662	3rd Qu.:	1255.2
Max. :	2775	Max. :	2522.8

T		RH	
Min. :	-200.000	Min. :	-200.00
1st Qu.:	10.950	1st Qu.:	34.05
Median :	17.200	Median :	48.55
Mean :	9.777	Mean :	39.48
3rd Qu.:	24.075	3rd Qu.:	61.88
Max. :	44.600	Max. :	88.72

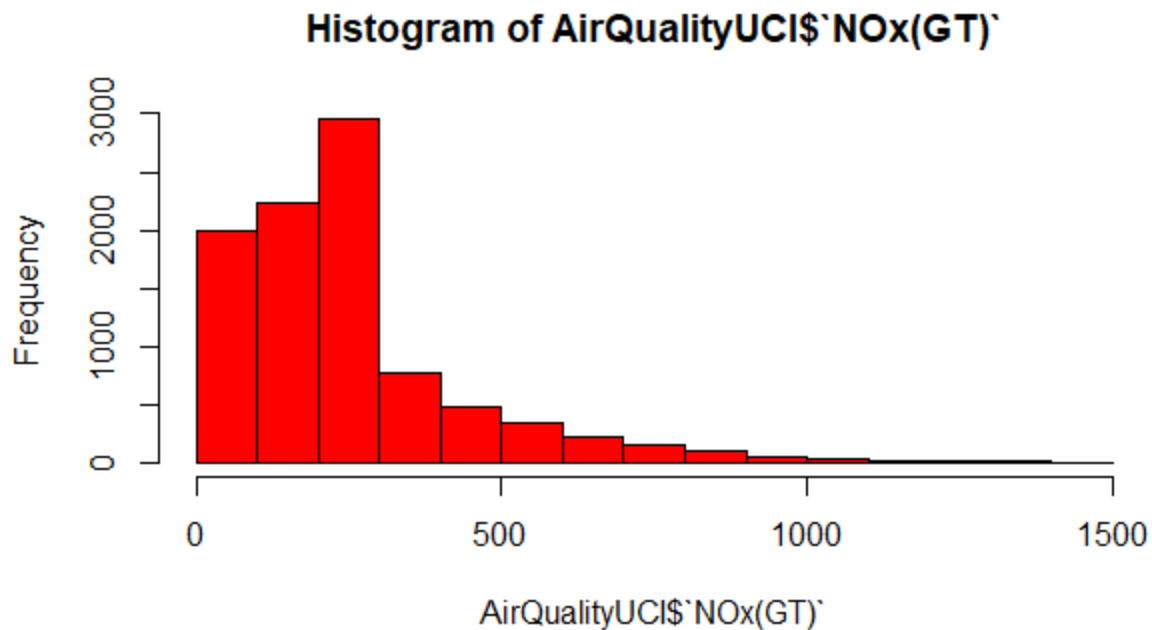
AH	
Min. :	-200.0000
1st Qu.:	0.6923
Median :	0.9768
Mean :	-6.8376
3rd Qu.:	1.2962

```

Max.      :    2.2310  > AirQualityUCI[7:14,]
      Date              Time CO(GT) PT08.S1(CO)
NMHC(GT) C6H6(GT) PT08.S2(NMHC) NOx(GT)
7  2004-03-11 1899-12-31 00:00:00 1.20000      1185.00
31 3.624399      689.50  62.0000
8  2004-03-11 1899-12-31 01:00:00 1.00000      1136.25
31 3.326677      672.00  62.0000
9  2004-03-11 1899-12-31 02:00:00 0.90000      1094.00
24 2.339416      608.50  45.0000
10 2004-03-11 1899-12-31 03:00:00 0.60000      1009.75
19 1.696658      560.75 246.8813
11 2004-03-11 1899-12-31 04:00:00 2.15275      1011.00
14 1.293620      526.75  21.0000
12 2004-03-11 1899-12-31 05:00:00 0.70000      1066.00
8  1.133431      512.00  16.0000
13 2004-03-11 1899-12-31 06:00:00 0.70000      1051.75
16 1.603768      553.25  34.0000
14 2004-03-11 1899-12-31 07:00:00 1.10000      1144.00
29 3.243618      667.00  98.0000
      PT08.S3(NOx) NO2(GT) PT08.S4(NO2) PT08.S5(O3)
T      RH      AH
7      1461.75  77.0000      1332.75      732.50
11.325 56.775 0.7603119
8      1453.25  76.0000      1332.75      729.50
10.675 60.000 0.7702385
9      1579.00  60.0000      1276.00      619.50
10.650 59.675 0.7648187
10      1705.00 113.0755      1234.75      501.25
10.250 60.200 0.7516572
11      1817.50  34.0000      1196.75      445.25
10.075 60.475 0.7464945
12      1918.00  28.0000      1182.00      421.75
11.000 56.175 0.7365596
13      1738.25  48.0000      1221.25      471.50
10.450 58.125 0.7352951
14      1489.75  82.0000      1339.00      729.75
10.200 59.600 0.7417362
> hist(AirQualityUCI$`NOx(GT)`,col="red")

```

```
>
```



C. Check for missing values in all columns.

Answer:

```
col1<- mapply(anyNA,AirQualityUCI)
col1
summary(AirQualityUCI)
is.na(AirQualityUCI)
```

Output from R-console

```
> colSums(is.na(AirQualityUCI))
```

Date	Time	CO(GT)	PT08.S1(CO)
NMHC(GT)	C6H6(GT)		
0	0	1683	366
8443	366		
PT08.S2(NMHC)	NOx(GT)	PT08.S3(NOx)	NO2(GT)
PT08.S4(NO2)	PT08.S5(O3)		
366	1639	366	1642
366	366		
T	RH	AH	
366	366	366	

```
col1<- mapply(anyNA,AirQualityUCI)
> col1
```

```
is.na(AirQualityUCI)
```

```
[46.] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
```

[47,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[48,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[49,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[50,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[51,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[52,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[53,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[54,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[55,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[56,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[57,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[58,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[59,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[60,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[61,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[62,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[63,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[64,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[65,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE
[66,]	FALSE	FALSE	FALSE		FALSE	FALSE	FALSE	FALSE	FALSE

```

[44,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[45,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[46,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[47,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[48,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[49,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[50,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[51,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[52,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[53,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[54,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[55,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[56,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[57,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[58,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[59,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[60,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[61,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[62,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[63,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[64,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[65,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[66,] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[ reached getOption("max.print") -- omitted 9291 rows ]

```

>

D. Impute the missing values using appropriate methods.

Answer:

```

colSums(is.na(AirQualityUCI))
library(plyr)
AirQualityUCI[AirQualityUCI==-200.0]<-NA
for(i in 1:ncol(AirQualityUCI)){
  AirQualityUCI[is.na(AirQualityUCI[,i]),i] <- mean(AirQualityUCI[,i],
na.rm = TRUE)}
summary(AirQualityUCI)

```

Output from R-console

```

> AirQualityUCI[AirQualityUCI==-200.0]<-NA
> for(i in 1:ncol(AirQualityUCI)){
+   AirQualityUCI[is.na(AirQualityUCI[,i]),i] <- mean(AirQualityUCI[,i], na.rm = TRUE)}
> summary(AirQualityUCI)

```

Date		Time		CO(GT)	
Min.	:2004-03-10 00:00:00	Min.	:1899-12-31 00:00:00	Min.	: 0.100
1st Qu.	:2004-06-16 00:00:00	1st Qu.	:1899-12-31 05:00:00	1st Qu.	: 1.200
Median	:2004-09-21 00:00:00	Median	:1899-12-31 11:00:00	Median	: 2.153
Mean	:2004-09-21 04:30:05	Mean	:1899-12-31 11:29:55	Mean	: 2.153
3rd Qu.	:2004-12-28 00:00:00	3rd Qu.	:1899-12-31 18:00:00	3rd Qu.	: 2.600
Max.	:2005-04-04 00:00:00	Max.	:1899-12-31 23:00:00	Max.	:11.900

PT08.S1(CO)		NMHC(GT)		C6H6(GT)		PT08.S2(NMHC)	
Min.	: 647.2	Min.	: 7.0	Min.	: 0.149	Min.	: 383.2
1st Qu.	: 941.2	1st Qu.	: 218.8	1st Qu.	: 4.591	1st Qu.	: 742.5
Median	:1074.5	Median	: 218.8	Median	: 8.593	Median	: 923.2
Mean	:1099.7	Mean	: 218.8	Mean	:10.083	Mean	: 939.0
3rd Qu.	:1221.2	3rd Qu.	: 218.8	3rd Qu.	:13.636	3rd Qu.	:1104.8
Max.	:2039.8	Max.	:1189.0	Max.	:63.741	Max.	:2214.0

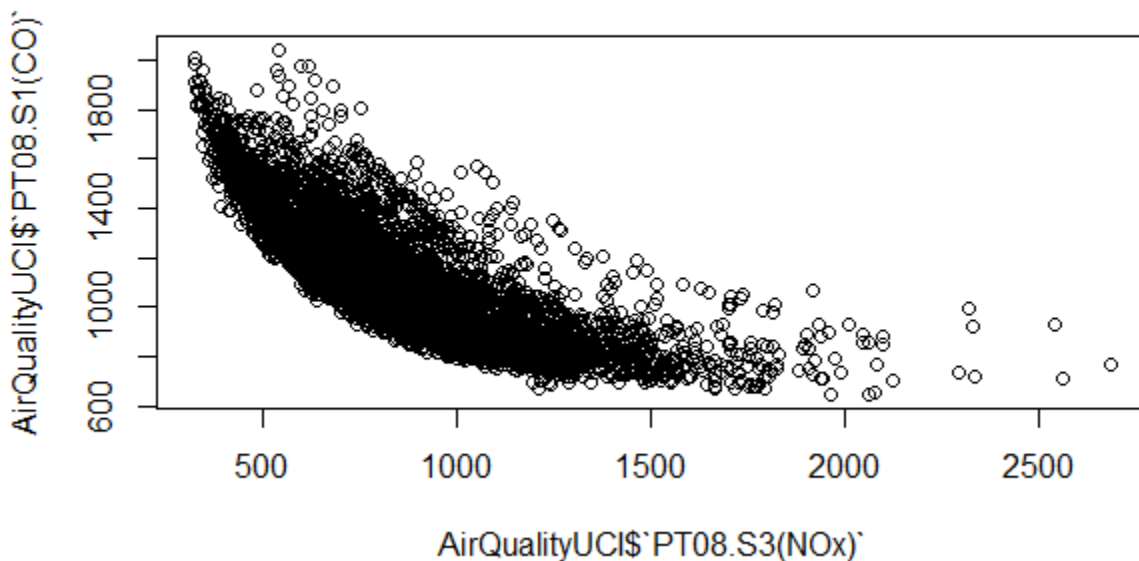
NOx(GT)	PT08.S3(NOx)	NO2(GT)	PT08.S4(NO2)	PT08.S5(O3)
Min. : 2.0	Min. : 322.0	Min. : 2.0	Min. : 551	Min. : 221.0
1st Qu.: 112.0	1st Qu.: 665.5	1st Qu.: 85.9	1st Qu.: 1242	1st Qu.: 741.8
Median : 229.0	Median : 817.5	Median : 113.1	Median : 1456	Median : 982.5
Mean : 246.9	Mean : 835.4	Mean : 113.1	Mean : 1456	Mean : 1022.8
3rd Qu.: 284.2	3rd Qu.: 960.2	3rd Qu.: 133.0	3rd Qu.: 1662	3rd Qu.: 1255.2
Max. : 1479.0	Max. : 2682.8	Max. : 339.7	Max. : 2775	Max. : 2522.8
T	RH	AH		
Min. : -1.90	Min. : 9.175	Min. : 0.1847		
1st Qu.: 12.03	1st Qu.: 36.550	1st Qu.: 0.7461		
Median : 18.27	Median : 49.232	Median : 1.0154		
Mean : 18.32	Mean : 49.232	Mean : 1.0255		
3rd Qu.: 24.07	3rd Qu.: 61.875	3rd Qu.: 1.2962		
Max. : 44.60	Max. : 88.725	Max. : 2.2310		

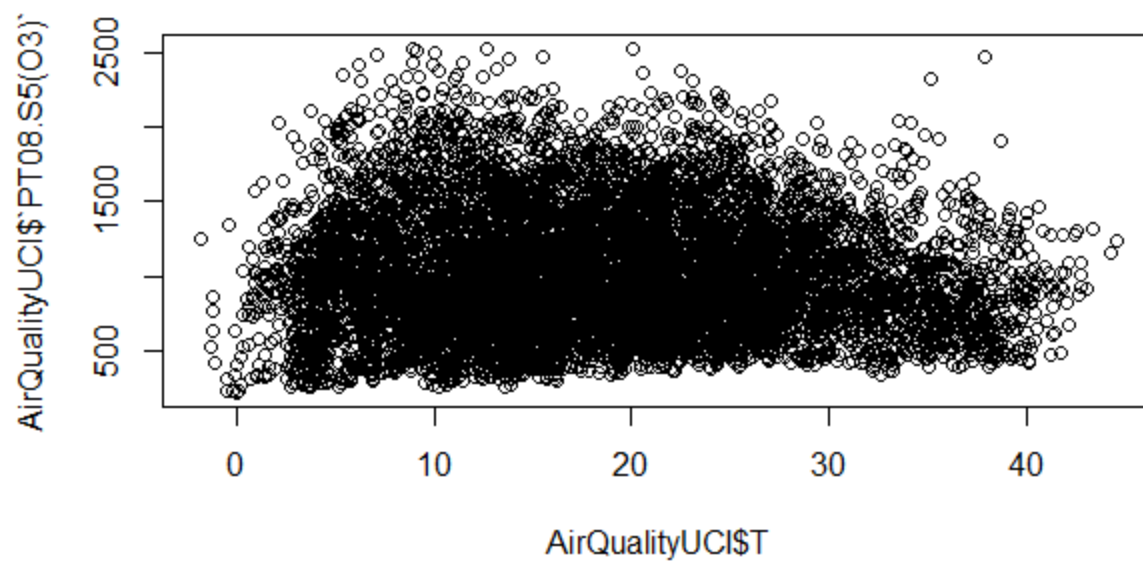
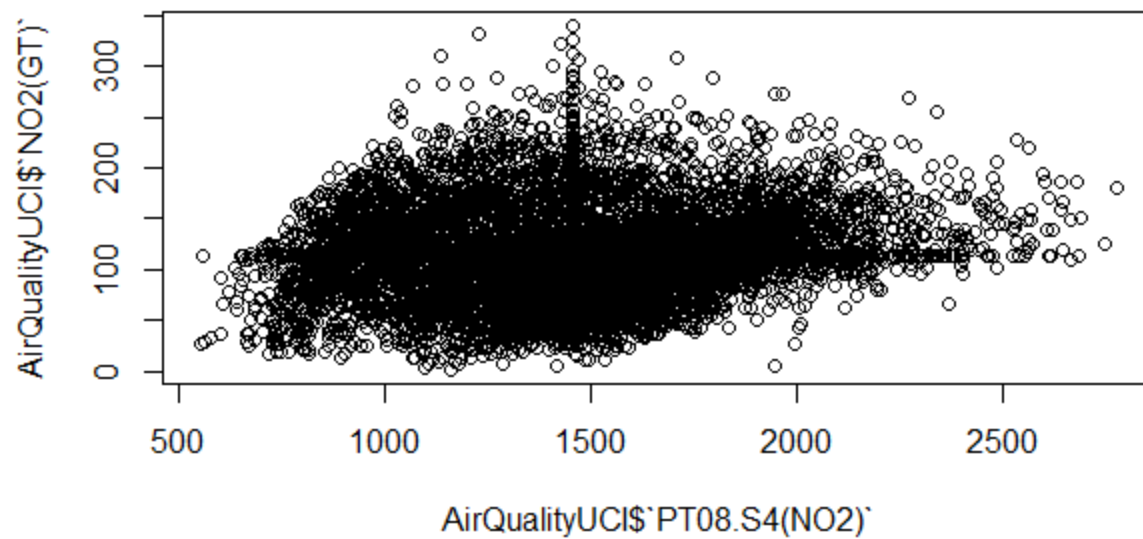
E. Create bi-variate analysis for all relationships

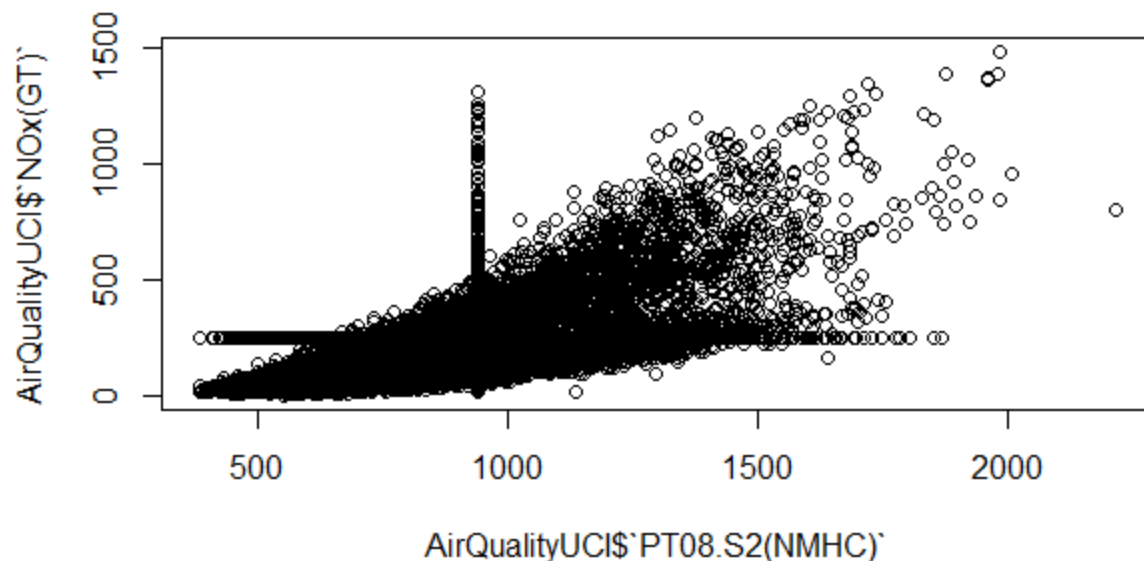
Answer:

```
summary(AirQualityUCI)
plot(AirQualityUCI$`NOx(GT)`~AirQualityUCI$`PT08.S2(NMHC)` )
plot(AirQualityUCI$`PT08.S1(CO)`~AirQualityUCI$`PT08.S3(NOx)` )
plot(AirQualityUCI$`NO2(GT)`~AirQualityUCI$`PT08.S4(NO2)` )
plot(AirQualityUCI$`PT08.S5(O3)`~AirQualityUCI$`T` )
```

Output from Rconsole







F. Test relevant hypothesis for valid relations

Answer:

```
plot(AirQualityUCI$`PT08.S1(CO)` ,AirQualityUCI$T)
lm(formula=AirQualityUCI$`PT08.S3(NOx)`~AirQualityUCI$`NOx(GT)` )
lm(formula = AirQualityUCI$`PT08.S1(CO)`~AirQualityUCI$T)
lm(formula =
AirQualityUCI$`NMHC(GT)`~AirQualityUCI$`PT08.S2(NMHC)` )
plot(AirQualityUCI$`PT08.S5(O3)` ,AirQualityUCI$`NOx(GT)` )
lm(formula
=AirQualityUCI$`PT08.S5(O3)`~AirQualityUCI$`NOx(GT)` )
pnorm(1.49)
pnorm(1.097)
qnorm(0.9318879)
qnorm(0.8636793)
library(car)
mod=lm(AirQualityUCI$`PT08.S5(O3)` ~ AirQualityUCI$`NOx(GT)` )
summary(mod)
predict(mod)
```

Output from R console

```
> plot(AirQualityUCI$`PT08.S1(CO)` , AirQualityUCI$T)
> lm(formula=AirQualityUCI$`PT08.S3(NOx)` ~ AirQualityUCI$`NOx(GT)` )

Call:
lm(formula = AirQualityUCI$`PT08.S3(NOx)` ~ AirQualityUCI$`NOx(GT)` )

Coefficients:
            (Intercept)  AirQualityUCI$`NOx(GT)` 
            1016.3598                -0.7331 

> lm(formula = AirQualityUCI$`PT08.S1(CO)` ~ AirQualityUCI$T)

Call:
lm(formula = AirQualityUCI$`PT08.S1(CO)` ~ AirQualityUCI$T)

Coefficients:
            (Intercept)  AirQualityUCI$T 
            1077.818                1.195 

> lm(formula = AirQualityUCI$`NMHC(GT)` ~ AirQualityUCI$`PT08.S2(NMHC)` )

Call:
lm(formula = AirQualityUCI$`NMHC(GT)` ~ AirQualityUCI$`PT08.S2(NMHC)` )

Coefficients:
            (Intercept)  AirQualityUCI$`PT08.S2(NMHC)` 
            154.66404                0.06831 

> plot(AirQualityUCI$`PT08.S5(O3)` , AirQualityUCI$`NOx(GT)` )
> lm(formula = AirQualityUCI$`PT08.S5(O3)` ~ AirQualityUCI$`NOx(GT)` )

Call:
lm(formula = AirQualityUCI$`PT08.S5(O3)` ~ AirQualityUCI$`NOx(GT)` )

Coefficients:
            (Intercept)  AirQualityUCI$`NOx(GT)` 
            679.65                1.39 

> pnorm(1.49)
[1] 0.9318879
> pnorm(1.097)
[1] 0.8636793
> qnorm(0.9318879)
[1] 1.49
> qnorm(0.8636793)
[1] 1.097

> library(car)
Loading required package: carData
Warning messages:
```

```

1: package 'car' was built under R version 3.5.2
2: package 'carData' was built under R version 3.5.2
> mod=lm(AirQualityUCI$`PT08.S5(O3)` ~ AirQualityUCI$`NOx(GT)` )
> summary(mod)

```

```

Call:
lm(formula = AirQualityUCI$`PT08.S5(O3)` ~ AirQualityUCI$`NOx(GT)` )

```

```

Residuals:
    Min       1Q   Median       3Q      Max
-1477.57  -182.85   -10.98   168.81  1369.83

```

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    679.65470     4.75079   143.06  <2e-16 ***
AirQualityUCI$`NOx(GT)`  1.38984     0.01515    91.75  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

Residual standard error: 283.4 on 9355 degrees of freedom
Multiple R-squared:  0.4736,    Adjusted R-squared:  0.4736
F-statistic: 8418 on 1 and 9355 DF,  p-value: < 2.2e-16

```

```

> predict(mod)

```

1	2	3	4	5	6	7	8
910.3685	822.8085	861.7241	918.7076	861.7241	803.3507	765.8249	765.8249
9	10	11	12	13	14	15	16
742.1976	1022.7807	708.8414	701.8922	726.9093	815.8593	921.4873	858.9444
17	18	19	20	21	22	23	24
835.3170	811.6897	824.1983	882.5717	967.3521	935.3857	947.8943	1017.3864
25	26	27	28	29	30	31	32
1070.2004	1211.9643	1167.4894	1013.2169	810.2999	744.9773	849.2155	864.5037
33	34	35	36	37	38	39	40
793.6218	1022.7807	708.8414	721.3500	757.4859	831.1475	1106.3363	1022.7807
41	42	43	44	45	46	47	48
939.5552	979.8607	878.4022	902.0295	906.1990	943.7248	927.0466	888.1311
49	50	51	52	53	54	55	56
965.9622	960.4029	1152.2011	1060.4715	1031.2848	1096.6074	947.8943	872.8428
57	58	59	60	61	62	63	64
795.0116	1022.7807	831.1475	765.8249	753.3163	872.8428	1035.4543	1089.6582
65	66	67	68	69	70	71	72
1027.1153	1091.0480	1011.8270	892.3006	843.6561	845.0459	871.4529	885.3514
73	74	75	76	77	78	79	80
995.1489	1064.6410	1067.4207	1072.9801	977.0810	918.7076	943.7248	921.4873
81	82	83	84	85	86	87	88
857.5545	1022.7807	776.9437	753.3163	740.8078	782.5030	820.0288	843.6561
89	90	91	92	93	94	95	96
879.7920	924.2670	903.4193	836.7069	822.8085	835.3170	915.9279	929.8263
97	98	99	100	101	102	103	104
902.0295	1042.4036	1131.3535	981.2505	882.5717	890.9107	824.1983	826.9780
105	106	107	108	109	110	111	112
829.7577	1022.7807	771.3843	801.9608	871.4529	996.5388	1179.9980	1343.9994
113	114	115	116	117	118	119	120
1227.2526	1123.0144	959.0130	917.3177	900.6396	892.3006	878.4022	896.4701
121	122	123	124	125	126	127	128
1116.0652	1241.1510	1124.4043	1035.4543	947.8943	803.3507	856.1647	789.4522
129	130	131	132	133	134	135	136
739.4179	1022.7807	761.6554	732.4687	793.6218	817.2491	1045.1832	997.9286
137	138	139	140	141	142	143	144
1230.0323	1116.0652	981.2505	1029.8950	1045.1832	929.8263	935.3857	943.7248
145	146	147	148	149	150	151	152
935.3857	931.2162	910.3685	1120.2347	935.3857	860.3342	863.1139	811.6897
153	154	155	156	157	158	159	160
776.9437	1022.7807	754.7062	718.5703	743.5874	902.0295	1124.4043	1203.6253

161	162	163	164	165	166	167	168
1068.8106	986.8099	971.5216	963.1825	903.4193	867.2834	938.1654	883.9615
169	170	171	172	173	174	175	176
1095.2176	1250.8799	1161.9300	1146.6417	960.4029	927.0466	893.6904	807.5202
177	178	179	180	181	182	183	184
764.4351	1022.7807	763.0452	731.0789	774.1640	872.8428	1150.8113	1264.7783
185	186	187	188	189	190	191	192
1164.7097	990.9794	917.3177	886.7412	870.0631	886.7412	918.7076	910.3685
193	194	195	196	197	198	199	200
977.0810	1068.8106	1063.2512	927.0466	858.9444	811.6897	854.7748	822.8085
201	202	203	204	205	206	207	208
746.3671	1022.7807	707.4515	704.6719	757.4859	839.4866	1124.4043	1175.8284
209	210	211	212	213	214	215	216
1031.2848	982.6403	946.5044	924.2670	954.8435	974.3013	1025.7255	1054.9121
217	218	219	220	221	222	223	224
1000.7083	1028.5051	971.5216	1025.7255	878.4022	856.1647	824.1983	799.1811
225	226	227	228	229	230	231	232
782.5030	1022.7807	740.8078	738.0281	767.2148	839.4866	915.9279	883.9615
233	234	235	236	237	238	239	240
910.3685	950.6740	932.6060	913.1482	853.3850	811.6897	789.4522	845.0459
241	242	243	244	245	246	247	248
877.0123	978.4708	997.9286	965.9622	883.9615	849.2155	922.8771	864.5037
249	250	251	252	253	254	255	256
803.3507	1022.7807	797.7913	756.0960	713.0109	779.7234	849.2155	833.9272
257	258	259	260	261	262	263	264
850.6053	826.9780	822.8085	821.4186	746.3671	767.2148	815.8593	863.1139
265	266	267	268	269	270	271	272
920.0974	936.7755	1049.3528	940.9451	849.2155	835.3170	808.9100	782.5030
273	274	275	276	277	278	279	280
711.6211	1022.7807	703.2820	700.5023	736.6382	799.1811	993.7591	977.0810
281	282	283	284	285	286	287	288
839.4866	851.9952	849.2155	828.3678	864.5037	860.3342	888.1311	871.4529
289	290	291	292	293	294	295	296
975.6911	1002.0981	1010.4372	929.8263	857.5545	769.9945	753.3163	744.9773
297	298	299	300	301	302	303	304
750.5367	1022.7807	707.4515	697.7226	725.5195	822.8085	990.9794	1018.7762
305	306	307	308	309	310	311	312
1009.0473	985.4200	967.3521	968.7419	902.0295	861.7241	857.5545	924.2670
313	314	315	316	317	318	319	320
828.3678	861.7241	910.3685	803.3507	792.2319	788.0624	776.9437	785.2827
321	322	323	324	325	326	327	328
739.4179	1022.7807	740.8078	726.9093	739.4179	789.4522	921.4873	1006.2677
329	330	331	332	333	334	335	336
982.6403	868.6733	842.2663	845.0459	854.7748	818.6389	953.4536	950.6740
337	338	339	340	341	342	343	344
896.4701	1022.9458	943.7248	914.5381	843.6561	792.2319	790.8421	738.0281
345	346	347	348	349	350	351	352
732.4687	1022.7807	704.6719	696.3328	739.4179	796.4015	935.3857	960.4029
353	354	355	356	357	358	359	360
864.5037	860.3342	826.9780	863.1139	826.9780	853.3850	871.4529	846.4358
361	362	363	364	365	366	367	368
899.2498	1031.2848	1000.7083	888.1311	817.2491	783.8929	808.9100	789.4522
369	370	371	372	373	374	375	376
772.7741	1022.7807	742.1976	779.7234	822.8085	863.1139	1011.8270	1061.8614
377	378	379	380	381	382	383	384
1002.0981	929.8263	835.3170	840.8764	822.8085	803.3507	835.3170	856.1647
385	386	387	388	389	390	391	392
853.3850	846.4358	846.4358	817.2491	789.4522	771.3843	810.2999	776.9437
393	394	395	396	397	398	399	400
783.8929	1022.7807	758.8757	724.1297	754.7062	774.1640	815.8593	857.5545
401	402	403	404	405	406	407	408
858.9444	860.3342	833.9272	826.9780	870.0631	836.7069	849.2155	870.0631
409	410	411	412	413	414	415	416
908.9787	952.0638	921.4873	874.2326	821.4186	860.3342	847.8256	833.9272

417	418	419	420	421	422	423	424
800.5710	1022.7807	786.6726	733.8586	786.6726	806.1304	825.5881	864.5037
425	426	427	428	429	430	431	432
857.5545	854.7748	807.5202	771.3843	749.1468	776.9437	781.1132	786.6726
433	434	435	436	437	438	439	440
799.1811	808.9100	825.5881	772.7741	767.2148	767.2148	743.5874	708.8414
441	442	443	444	445	446	447	448
710.2312	1022.7807	733.8586	756.0960	785.2827	931.2162	1039.6239	843.6561
449	450	451	452	453	454	455	456
845.0459	850.6053	807.5202	806.1304	822.8085	821.4186	810.2999	815.8593
457	458	459	460	461	462	463	464
854.7748	829.7577	811.6897	788.0624	764.4351	756.0960	758.8757	736.6382
465	466	467	468	469	470	471	472
743.5874	1022.7807	736.6382	751.9265	833.9272	945.1146	1018.7762	1072.9801
473	474	475	476	477	478	479	480
1124.4043	1025.7255	984.0302	957.6232	871.4529	843.6561	817.2491	835.3170
481	482	483	484	485	486	487	488
842.2663	858.9444	863.1139	781.1132	774.1640	790.8421	774.1640	740.8078
489	490	491	492	493	494	495	496
744.9773	1022.7807	700.5023	740.8078	797.7913	967.3521	999.3184	931.2162
497	498	499	500	501	502	503	504
879.7920	874.2326	963.1825	886.7412	917.3177	950.6740	890.9107	857.5545
505	506	507	508	509	510	511	512
847.8256	877.0123	861.7241	792.2319	763.0452	758.8757	813.0796	797.7913
513	514	515	516	517	518	519	520
749.1468	1022.7807	758.8757	768.6046	892.3006	1154.9808	1205.0151	1102.1668
521	522	523	524	525	526	527	528
946.5044	860.3342	826.9780	821.4186	817.2491	829.7577	861.7241	906.1990
529	530	531	532	533	534	535	536
972.9114	1050.7426	1053.5223	893.6904	856.1647	849.2155	854.7748	796.4015
537	538	539	540	541	542	543	544
760.2656	1022.7807	731.0789	757.4859	806.1304	900.6396	943.7248	1025.7255
545	546	547	548	549	550	551	552
952.0638	929.8263	911.7584	900.6396	890.9107	886.7412	878.4022	936.7755
553	554	555	556	557	558	559	560
989.5895	978.4708	1056.3020	929.8263	839.4866	886.7412	815.8593	781.1132
561	562	563	564	565	566	567	568
775.5538	1022.7807	763.0452	757.4859	781.1132	900.6396	961.7927	972.9114
569	570	571	572	573	574	575	576
1022.7807	1022.7807	1022.7807	1022.7807	1022.7807	1022.7807	1022.7807	1022.7807
577	578	579	580	581	582	583	584
1022.7807	1022.7807	1022.7807	1022.7807	1022.7807	1022.7807	1022.7807	1022.7807
585	586	587	588	589	590	591	592
1022.7807	1022.7807	1022.7807	1022.7807	1022.7807	1022.7807	1022.7807	1022.7807
593	594	595	596	597	598	599	600
1022.7807	1022.7807	1022.7807	1022.7807	1022.7807	1022.7807	1022.7807	1022.7807
601	602	603	604	605	606	607	608
1022.7807	1022.7807	1022.7807	1022.7807	1022.7807	1022.7807	1022.7807	1022.7807
609	610	611	612	613	614	615	616
1022.7807	1022.7807	1022.7807	1022.7807	1022.7807	1022.7807	1022.7807	1022.7807
617	618	619	620	621	622	623	624
1007.6575	996.5388	965.9622	940.9451	899.2498	851.9952	815.8593	902.0295
625	626	627	628	629	630	631	632
949.2841	921.4873	897.8600	831.1475	797.7913	804.7405	764.4351	729.6890
633	634	635	636	637	638	639	640
714.4008	1022.7807	713.0109	781.1132	907.5888	1136.9128	1099.3871	1035.4543
641	642	643	644	645	646	647	648
932.6060	825.5881	868.6733	911.7584	910.3685	839.4866	831.1475	878.4022
649	650	651	652	653	654	655	656
900.6396	931.2162	927.0466	806.1304	785.2827	771.3843	757.4859	728.2992
657	658	659	660	661	662	663	664
708.8414	1022.7807	696.3328	710.2312	756.0960	890.9107	892.3006	890.9107
665	666	667	668	669	670	671	672
835.3170	845.0459	849.2155	839.4866	796.4015	882.5717	769.9945	846.4358

[illegible]


```

      929      930      931      932      933      934      935      936
1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807
      937      938      939      940      941      942      943      944
1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807
      945      946      947      948      949      950      951      952
1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807
      953      954      955      956      957      958      959      960
1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807
      961      962      963      964      965      966      967      968
1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807
      969      970      971      972      973      974      975      976
1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807
      977      978      979      980      981      982      983      984
1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807
      985      986      987      988      989      990      991      992
1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807
      993      994      995      996      997      998      999      1000
1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807 1022.7807
[ reached getOption("max.print") -- omitted 8357 entries ]

```

>

G. Create cross tabulations with derived variables

Answer:

```

mydata<-AirQualityUCI
View(mydata)           # 2-Way Frequency Table
attach(mydata)
mytable <- table(A,B)   # A will be rows, B will be columns
mytable                # print table
margin.table(mytable, 1) # A frequencies (summed over B)
margin.table(mytable, 2) # B frequencies (summed over A)
prop.table(mytable)     # cell percentages
prop.table(mytable, 1)  # row percentages
prop.table(mytable, 2)  # column percentages

```

```

data: mytable
x-squared = 2450, df = 2401, p-value = 0.2382

```

H. Check for trends and patterns in time series.

Answer:

```
ts (AirQualityUCI, frequency = 4, start = c(1959, 2)) # frequency 4 =>
Quarterly Data
ts (1:10, frequency = 12, start = 1990) # freq 12 => Monthly data.
ts (AirQualityUCI, start=c(2009), end=c(2014), frequency=1) # Yearly
Data
ts (1:1000, frequency = 365, start = 1990)# freq 365 => daily data.
tsAirqualityUCI <- EuStockMarkets[, 1] # ts data
copied some time series data as below
```

```
#plot time series
tsAirqualityUCI <- EuStockMarkets[, 1] # ts data
decomposedRes <- decompose(tsAirqualityUCI, type="mult") # use
type = "additive" for additive components
plot (decomposedRes) # see plot below
```

Output from Rconsole

```
> ts (AirQualityUCI, frequency = 4, start = c(1959, 2))# frequency 4 =>
Quarterly Data
```

	Date	Time	CO(GT)	PT08.S1(CO)	NMHC(GT)	C6H6(GT)	
1959	Q2	1078876800	-2209010400	2.60000	1360.0000	150.0000	11.8817235
1959	Q3	1078876800	-2209006800	2.00000	1292.2500	112.0000	9.3971649
1959	Q4	1078876800	-2209003200	2.20000	1402.0000	88.0000	8.9978169
1960	Q1	1078876800	-2208999600	2.20000	1375.5000	80.0000	9.2287964
1960	Q2	1078876800	-2208996000	1.60000	1272.2500	51.0000	6.5182237
1960	Q3	1078876800	-2208992400	1.20000	1197.0000	38.0000	4.7410124
1960	Q4	1078963200	-2209075200	1.20000	1185.0000	31.0000	3.6243992
1961	Q1	1078963200	-2209071600	1.00000	1136.2500	31.0000	3.3266770
1961	Q2	1078963200	-2209068000	0.90000	1094.0000	24.0000	2.3394162
1961	Q3	1078963200	-2209064400	0.60000	1009.7500	19.0000	1.6966583
1961	Q4	1078963200	-2209060800	2.15275	1011.0000	14.0000	1.2936198
1962	Q1	1078963200	-2209057200	0.70000	1066.0000	8.0000	1.1334306
1962	Q2	1078963200	-2209053600	0.70000	1051.7500	16.0000	1.6037679
1962	Q3	1078963200	-2209050000	1.10000	1144.0000	29.0000	3.2436181
1962	Q4	1078963200	-2209046400	2.00000	1333.2500	64.0000	8.0137730
1963	Q1	1078963200	-2209042800	2.20000	1351.0000	87.0000	9.5406429
1963	Q2	1078963200	-2209039200	1.70000	1233.2500	77.0000	6.3357824
1963	Q3	1078963200	-2209035600	1.50000	1178.7500	43.0000	4.9715838
1963	Q4	1078963200	-2209032000	1.60000	1236.0000	61.0000	5.2169190
1964	Q1	1078963200	-2209028400	1.90000	1285.5000	63.0000	7.2699334
1964	Q2	1078963200	-2209024800	2.90000	1371.0000	164.0000	11.5390072
1964	Q3	1078963200	-2209021200	2.20000	1310.0000	79.0000	8.8262227
1964	Q4	1078963200	-2209017600	2.20000	1291.7500	95.0000	8.3014134
1965	Q1	1078963200	-2209014000	2.90000	1383.0000	150.0000	11.1515812
1965	Q2	1078963200	-2209010400	4.80000	1580.7500	307.0000	20.7992169
1965	Q3	1078963200	-2209006800	6.90000	1775.5000	461.0000	27.3598075
1965	Q4	1078963200	-2209003200	6.10000	1640.0000	401.0000	24.0177569
1966	Q1	1078963200	-2208999600	3.90000	1312.7500	197.0000	12.7793682

1966	Q2	1078963200	-2208996000	1.50000	964.5000	61.0000	4.7070719
1966	Q3	1078963200	-2208992400	1.00000	912.7500	26.0000	2.6457215
1966	Q4	1079049600	-2209075200	1.70000	1080.2500	55.0000	5.8548015
1967	Q1	1079049600	-2209071600	1.90000	1043.7500	53.0000	6.3742975
1967	Q2	1079049600	-2209068000	1.40000	987.7500	40.0000	4.1323418
1967	Q3	1079049600	-2209064400	0.80000	888.7500	21.0000	1.8694446
1967	Q4	1079049600	-2209060800	2.15275	831.0000	10.0000	1.0682926
1968	Q1	1079049600	-2209057200	0.60000	847.2500	7.0000	1.0224146
1968	Q2	1079049600	-2209053600	0.80000	927.0000	17.0000	1.8304312
1968	Q3	1079049600	-2209050000	1.40000	1090.5000	33.0000	4.3593410
1968	Q4	1079049600	-2209046400	4.40000	1587.0000	202.0000	17.8655867
1969	Q1	1079049600	-2209042800	2.15275	1544.5000	218.8118	22.0741621
1969	Q2	1079049600	-2209039200	3.10000	1350.2500	208.0000	14.0270114
1969	Q3	1079049600	-2209035600	2.70000	1262.7500	166.0000	11.6456466
1969	Q4	1079049600	-2209032000	2.10000	1206.2500	114.0000	10.2246621
1970	Q1	1079049600	-2209028400	2.50000	1251.5000	140.0000	11.0399360
1970	Q2	1079049600	-2209024800	2.70000	1287.0000	169.0000	12.8164462
1970	Q3	1079049600	-2209021200	2.90000	1352.7500	185.0000	14.1738512
1970	Q4	1079049600	-2209017600	2.80000	1309.0000	165.0000	12.6905681
1971	Q1	1079049600	-2209014000	2.40000	1274.0000	133.0000	11.7384054
1971	Q2	1079049600	-2209010400	3.90000	1509.5000	233.0000	19.2909749
1971	Q3	1079049600	-2209006800	3.70000	1525.2500	242.0000	18.2261783
1971	Q4	1079049600	-2209003200	6.60000	1843.0000	488.0000	32.5562783
1972	Q1	1079049600	-2208999600	4.40000	1597.7500	333.0000	20.0929436
1972	Q2	1079049600	-2208996000	3.50000	1483.5000	215.0000	14.3213424
1972	Q3	1079049600	-2208992400	5.40000	1677.2500	367.0000	21.8128651
1972	Q4	1079136000	-2209075200	2.70000	1279.5000	122.0000	9.6389998
1973	Q1	1079136000	-2209071600	1.90000	1196.2500	67.0000	7.3751395
1973	Q2	1079136000	-2209068000	1.60000	1183.7500	43.0000	5.3696042
1973	Q3	1079136000	-2209064400	1.70000	1171.7500	46.0000	5.3901039
1973	Q4	1079136000	-2209060800	2.15275	1147.0000	56.0000	6.1990420
1974	Q1	1079136000	-2209057200	1.00000	978.2500	30.0000	2.5779322
1974	Q2	1079136000	-2209053600	1.20000	1099.5000	27.0000	2.9085480
1974	Q3	1079136000	-2209050000	1.50000	1112.2500	47.0000	5.1362558
1974	Q4	1079136000	-2209046400	2.70000	1335.5000	132.0000	11.8171386
1975	Q1	1079136000	-2209042800	3.70000	1408.3333	239.0000	15.1401612
1975	Q2	1079136000	-2209039200	3.20000	1447.0000	160.0000	12.9130631
1975	Q3	1079136000	-2209035600	4.10000	1541.5000	283.0000	16.1335088

	PT08.S2(NMHC)	NOx(GT)	PT08.S3(NOx)	NO2(GT)	PT08.S4(NO2)
PT08.S5(O3)					
1959 Q2	1045.5000	166.0000	1056.2500	113.0000	1692.0000
1267.5000					
1959 Q3	954.7500	103.0000	1173.7500	92.0000	1558.7500
972.2500					
1959 Q4	939.2500	131.0000	1140.0000	114.0000	1554.5000
1074.0000					
1960 Q1	948.2500	172.0000	1092.0000	122.0000	1583.7500
1203.2500					
1960 Q2	835.5000	131.0000	1205.0000	116.0000	1490.0000
1110.0000					
1960 Q3	750.2500	89.0000	1336.5000	96.0000	1393.0000
949.2500					
1960 Q4	689.5000	62.0000	1461.7500	77.0000	1332.7500
732.5000					
1961 Q1	672.0000	62.0000	1453.2500	76.0000	1332.7500
729.5000					

1961 Q2 619.5000	608.5000	45.0000	1579.0000	60.0000	1276.0000
1961 Q3 501.2500	560.7500	246.8813	1705.0000	113.0755	1234.7500
1961 Q4 445.2500	526.7500	21.0000	1817.5000	34.0000	1196.7500
1962 Q1 421.7500	512.0000	16.0000	1918.0000	28.0000	1182.0000
1962 Q2 471.5000	553.2500	34.0000	1738.2500	48.0000	1221.2500
1962 Q3 729.7500	667.0000	98.0000	1489.7500	82.0000	1339.0000
1962 Q4 1101.5000	899.7500	174.0000	1136.0000	112.0000	1517.0000
1963 Q1 1027.7500	960.2500	129.0000	1079.0000	101.0000	1582.7500
1963 Q2 859.7500	827.2500	112.0000	1218.0000	98.0000	1445.7500
1963 Q3 670.5000	762.0000	95.0000	1327.5000	92.0000	1361.7500
1963 Q4 664.0000	774.2500	104.0000	1301.2500	95.0000	1401.2500
1964 Q1 799.0000	868.5000	146.0000	1162.2500	112.0000	1536.7500
1964 Q2 1036.5000	1033.5000	207.0000	983.2500	128.0000	1730.2500
1964 Q3 946.2500	932.5000	184.0000	1081.7500	126.0000	1646.5000
1964 Q4 956.7500	911.5000	193.0000	1102.5000	131.0000	1590.7500
1965 Q1 1104.0000	1019.7500	243.0000	1008.0000	135.0000	1718.7500
1965 Q2 1408.5000	1318.5000	281.0000	798.5000	151.0000	2083.0000
1965 Q3 1704.0000	1487.7500	383.0000	702.2500	172.0000	2332.5000
1965 Q4 1653.7500	1404.0000	351.0000	742.7500	165.0000	2191.2500
1966 Q1 1284.7500	1076.2500	240.0000	957.2500	136.0000	1706.5000
1966 Q2 821.0000	748.5000	94.0000	1325.2500	85.0000	1332.5000
1966 Q3 551.7500	629.2500	47.0000	1564.5000	53.0000	1252.2500
1966 Q4 815.5000	805.0000	122.0000	1253.5000	97.0000	1375.0000
1967 Q1 831.5000	829.0000	133.0000	1247.2500	110.0000	1378.2500
1967 Q2 691.5000	718.0000	82.0000	1395.5000	91.0000	1303.5000
1967 Q3 512.0000	574.2500	246.8813	1680.2500	113.0755	1187.0000
1967 Q4 384.0000	505.7500	21.0000	1892.7500	32.0000	1133.7500
1968 Q1 394.0000	501.2500	30.0000	1894.5000	44.0000	1154.7500

1968 Q2	571.2500	56.0000	1684.7500	71.0000	1222.7500
486.5000					
1968 Q3	730.2500	109.0000	1387.0000	104.0000	1360.7500
748.2500					
1968 Q4	1235.5000	307.0000	896.5000	141.0000	1900.2500
1400.2500					
1969 Q1	1353.0000	246.8813	767.2500	113.0755	2058.0000
1587.7500					
1969 Q2	1117.5000	187.0000	912.0000	122.0000	1711.7500
1237.0000					
1969 Q3	1037.2500	216.0000	969.0000	143.0000	1598.2500
1166.5000					
1969 Q4	986.0000	143.0000	1034.5000	113.0000	1537.0000
959.0000					
1970 Q1	1015.7500	160.0000	1007.5000	116.0000	1592.7500
983.0000					
1970 Q2	1077.5000	163.0000	948.7500	123.0000	1660.2500
1060.7500					
1970 Q3	1122.2500	190.0000	921.7500	126.0000	1740.0000
1139.2500					
1970 Q4	1073.2500	178.0000	954.0000	120.0000	1657.2500
1112.2500					
1971 Q1	1040.5000	150.0000	1005.7500	119.0000	1609.7500
993.7500					
1971 Q2	1276.5000	206.0000	812.2500	149.0000	1909.7500
1409.5000					
1971 Q3	1246.0000	202.0000	821.0000	145.0000	1846.7500
1447.7500					
1971 Q4	1609.7500	340.0000	624.0000	170.0000	2390.2500
1886.5000					
1972 Q1	1299.0000	274.0000	752.0000	149.0000	1940.5000
1626.7500					
1972 Q2	1127.0000	253.0000	839.0000	139.0000	1723.0000
1491.0000					
1972 Q3	1346.0000	300.0000	740.5000	134.0000	2062.0000
1657.0000					
1972 Q4	964.0000	193.0000	962.5000	113.0000	1543.5000
1285.2500					
1973 Q1	873.0000	139.0000	1071.2500	97.0000	1463.2500
1144.2500					
1973 Q2	781.7500	83.0000	1176.2500	82.0000	1364.5000
1042.7500					
1973 Q3	782.7500	246.8813	1178.5000	113.0755	1379.7500
995.5000					
1973 Q4	821.0000	109.0000	1132.2500	83.0000	1411.7500
991.5000					
1974 Q1	624.7500	62.0000	1420.2500	65.0000	1274.2500
819.2500					
1974 Q2	646.2500	53.0000	1406.2500	60.0000	1267.5000
835.0000					
1974 Q3	770.2500	139.0000	1228.0000	77.0000	1408.5000
939.7500					
1974 Q4	1043.2500	256.0000	935.2500	96.0000	1678.0000
1191.7500					
1975 Q1	1153.0000	295.0000	830.3333	119.0000	1776.6667
1411.0000					

1975 Q2	1080.7500	250.0000	868.5000	126.0000	1666.7500
1465.0000					
1975 Q3	1183.7500	296.0000	808.2500	158.0000	1779.7500
1582.5000					

	T	RH	AH
1959 Q2	13.600000	48.87500	0.7577538
1959 Q3	13.300000	47.70000	0.7254874
1959 Q4	11.900000	53.97500	0.7502391
1960 Q1	11.000000	60.00000	0.7867125
1960 Q2	11.150000	59.57500	0.7887942
1960 Q3	11.175000	59.17500	0.7847717
1960 Q4	11.325000	56.77500	0.7603119
1961 Q1	10.675000	60.00000	0.7702385
1961 Q2	10.650000	59.67500	0.7648187
1961 Q3	10.250000	60.20000	0.7516572
1961 Q4	10.075000	60.47500	0.7464945
1962 Q1	11.000000	56.17500	0.7365596
1962 Q2	10.450000	58.12500	0.7352951
1962 Q3	10.200000	59.60000	0.7417362
1962 Q4	10.750000	57.42500	0.7407946
1963 Q1	10.500000	60.60000	0.7691108
1963 Q2	10.800000	58.35000	0.7551831
1963 Q3	10.500000	57.92500	0.7351608
1963 Q4	9.525000	66.77500	0.7950538
1964 Q1	8.300000	76.42500	0.8392681
1964 Q2	8.000000	81.15000	0.8735885
1964 Q3	8.325000	79.80000	0.8777844
1964 Q4	9.700000	71.15000	0.8569381
1965 Q1	9.775000	67.62500	0.8185012
1965 Q2	10.350000	64.17500	0.8065436
1965 Q3	9.650000	69.30000	0.8319211
1965 Q4	9.650000	67.75000	0.8133139
1966 Q1	9.125000	63.97500	0.7419242
1966 Q2	8.175000	63.40000	0.6904844
1966 Q3	8.250000	60.82500	0.6657444
1966 Q4	8.325000	58.52500	0.6437636
1967 Q1	7.725000	59.67500	0.6307661
1967 Q2	7.125000	61.80000	0.6275974
1967 Q3	6.975000	62.27500	0.6261075
1967 Q4	6.100000	65.90000	0.6247536
1968 Q1	6.275000	64.97500	0.6232823
1968 Q2	6.750000	62.95000	0.6234275
1968 Q3	6.450000	65.07500	0.6316281
1968 Q4	7.325000	63.15000	0.6499331
1969 Q1	9.225000	56.20000	0.6560651
1969 Q2	13.225000	41.75000	0.6319501
1969 Q3	14.325000	38.45000	0.6243043
1969 Q4	15.025000	36.50000	0.6195323
1970 Q1	16.100000	34.47500	0.6261647
1970 Q2	16.275001	35.72500	0.6560306
1970 Q3	15.825000	37.02500	0.6609611
1970 Q4	15.875000	37.17500	0.6657285
1971 Q1	16.875000	34.35000	0.6549085
1971 Q2	15.150000	39.55000	0.6766265
1971 Q3	14.400000	43.42500	0.7084498
1971 Q4	12.875000	50.52500	0.7478032
1972 Q1	12.150000	53.35000	0.7536202

```

1972 Q2 10.975000 59.12500 0.7739800
1972 Q3 9.675000 64.62500 0.7770739
1972 Q4 9.450000 64.12500 0.7597465
1973 Q1 9.150000 63.90000 0.7422764
1973 Q2 8.800000 63.92500 0.7256154
1973 Q3 7.800000 67.52500 0.7173121
1973 Q4 7.000000 71.07500 0.7157785
1974 Q1 8.300000 63.57500 0.6981546
1974 Q2 7.200000 67.47500 0.6886721
1974 Q3 6.350000 71.90000 0.6931986
1974 Q4 6.450000 71.55000 0.6944755
1975 Q1 9.566667 59.66667 0.7123666
1975 Q2 12.375000 51.17500 0.7334584
1975 Q3 15.650000 42.20000 0.7450938

```

[reached getOption("max.print") -- omitted 9291 rows]

```
> ts (1:10, frequency = 12, start = 1990) # freq 12 => Monthly data.
```

```

      Jan Feb Mar Apr May Jun Jul Aug Sep Oct
1990   1   2   3   4   5   6   7   8   9  10

```

```
> ts (AirQualityUCI, start=c(2009), end=c(2014), frequency=1) # Yearly Data
```

Time Series:

Start = 2009

End = 2014

Frequency = 1

	Date	Time	CO(GT)	PT08.S1(CO)	NMHC(GT)	C6H6(GT)
PT08.S2(NMHC)						
2009	1078876800	-2209010400	2.6	1360.00	150	11.881723
1045.50						
2010	1078876800	-2209006800	2.0	1292.25	112	9.397165
954.75						
2011	1078876800	-2209003200	2.2	1402.00	88	8.997817
939.25						
2012	1078876800	-2208999600	2.2	1375.50	80	9.228796
948.25						
2013	1078876800	-2208996000	1.6	1272.25	51	6.518224
835.50						
2014	1078876800	-2208992400	1.2	1197.00	38	4.741012
750.25						

	NOx(GT)	PT08.S3(NOx)	NO2(GT)	PT08.S4(NO2)	PT08.S5(O3)	T	RH
AH							
2009	166	1056.25	113	1692.00	1267.50	13.600	48.875
0.7577538							
2010	103	1173.75	92	1558.75	972.25	13.300	47.700
0.7254874							
2011	131	1140.00	114	1554.50	1074.00	11.900	53.975
0.7502391							
2012	172	1092.00	122	1583.75	1203.25	11.000	60.000
0.7867125							
2013	131	1205.00	116	1490.00	1110.00	11.150	59.575
0.7887942							
2014	89	1336.50	96	1393.00	949.25	11.175	59.175
0.7847717							

```
> ts (1:1000, frequency = 365, start = 1990) # freq 365 => daily data.
```

Time Series:

Start = c(1990, 1)

End = c(1992, 270)

Frequency = 365

15	[1]	1	2	3	4	5	6	7	8	9	10	11	12	13	14
30	[16]	16	17	18	19	20	21	22	23	24	25	26	27	28	29
45	[31]	31	32	33	34	35	36	37	38	39	40	41	42	43	44
60	[46]	46	47	48	49	50	51	52	53	54	55	56	57	58	59
75	[61]	61	62	63	64	65	66	67	68	69	70	71	72	73	74
90	[76]	76	77	78	79	80	81	82	83	84	85	86	87	88	89
105	[91]	91	92	93	94	95	96	97	98	99	100	101	102	103	104
120	[106]	106	107	108	109	110	111	112	113	114	115	116	117	118	119
135	[121]	121	122	123	124	125	126	127	128	129	130	131	132	133	134
150	[136]	136	137	138	139	140	141	142	143	144	145	146	147	148	149
165	[151]	151	152	153	154	155	156	157	158	159	160	161	162	163	164
180	[166]	166	167	168	169	170	171	172	173	174	175	176	177	178	179
195	[181]	181	182	183	184	185	186	187	188	189	190	191	192	193	194
210	[196]	196	197	198	199	200	201	202	203	204	205	206	207	208	209
225	[211]	211	212	213	214	215	216	217	218	219	220	221	222	223	224
240	[226]	226	227	228	229	230	231	232	233	234	235	236	237	238	239
255	[241]	241	242	243	244	245	246	247	248	249	250	251	252	253	254
270	[256]	256	257	258	259	260	261	262	263	264	265	266	267	268	269
285	[271]	271	272	273	274	275	276	277	278	279	280	281	282	283	284
300	[286]	286	287	288	289	290	291	292	293	294	295	296	297	298	299
315	[301]	301	302	303	304	305	306	307	308	309	310	311	312	313	314
330	[316]	316	317	318	319	320	321	322	323	324	325	326	327	328	329
345	[331]	331	332	333	334	335	336	337	338	339	340	341	342	343	344
360	[346]	346	347	348	349	350	351	352	353	354	355	356	357	358	359
375	[361]	361	362	363	364	365	366	367	368	369	370	371	372	373	374
390	[376]	376	377	378	379	380	381	382	383	384	385	386	387	388	389
405	[391]	391	392	393	394	395	396	397	398	399	400	401	402	403	404
420	[406]	406	407	408	409	410	411	412	413	414	415	416	417	418	419

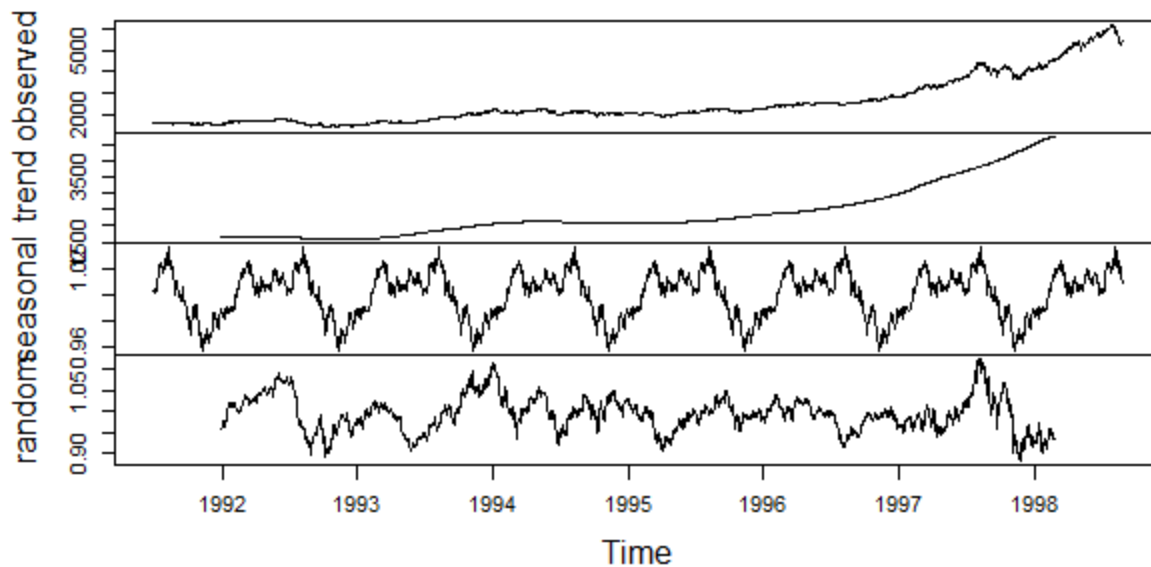
[illegible]

```

[841] 841 842 843 844 845 846 847 848 849 850 851 852 853 854
855
[856] 856 857 858 859 860 861 862 863 864 865 866 867 868 869
870
[871] 871 872 873 874 875 876 877 878 879 880 881 882 883 884
885
[886] 886 887 888 889 890 891 892 893 894 895 896 897 898 899
900
[901] 901 902 903 904 905 906 907 908 909 910 911 912 913 914
915
[916] 916 917 918 919 920 921 922 923 924 925 926 927 928 929
930
[931] 931 932 933 934 935 936 937 938 939 940 941 942 943 944
945
[946] 946 947 948 949 950 951 952 953 954 955 956 957 958 959
960
[961] 961 962 963 964 965 966 967 968 969 970 971 972 973 974
975
[976] 976 977 978 979 980 981 982 983 984 985 986 987 988 989
990
[991] 991 992 993 994 995 996 997 998 999 1000
> tsAirqualityUCI <- EuStockMarkets[, 1]

```

Decomposition of multiplicative time series



I. Find out the most polluted time of the day and the name of the chemical compound

Answer

#plot time series

tsAirqualityUCI <- EuStockMarkets[, 1] # ts data

```

decomposedRes <- decompose(tsAirqualityUCI, type="mult") # use
type = "additive" for additive components
plot(decomposedRes) # see plot below
stlRes <- stl(tsAirqualityUCI, s.window = "periodic")
plot(AirQualityUCI$T, type = "l")

```

Output from Rconsole

PT08.S4(NO2) is the highest pollution at 18.00 hr

Date	Time	NOx(GT)	PT08.S3(NOx)	NO2(GT)	PT08.S4(NO2)	PT08.S5(O3)
6/8/2004	8:00:00	376	525	125	2746	1708
6/9/2004	8:00:00	357	507	151	2691	2147
10/26/2004	18:00:00	952	325	180	2775	2372
max		1479.0	2682.8	339.7	2775.0	2522.8