R Notebook

Code

UNIVARIANT ANALYSIS

Import data

Hide

```
#data separated by commas, replace missing values with NA, use first row as headers
data <- read.csv(file = "C:/Users/16133/OneDrive/Desktop/CIND820/Data sets/PRSA_Data_Shunyi_20130301-20170228.csv", sep =
",", header = TRUE, na.strings = c("","NA"))


#check data types and first 6 rows of data
str(data)
```

```
'data.frame':   35064 obs. of  18 variables:
 $ No     : int  1 2 3 4 5 6 7 8 9 10 ...
 $ year   : int  2013 2013 2013 2013 2013 2013 2013 2013 2013 2013 ...
 $ month  : int  3 3 3 3 3 3 3 3 3 3 ...
 $ day    : int  1 1 1 1 1 1 1 1 1 1 ...
 $ hour   : int  0 1 2 3 4 5 6 7 8 9 ...
 $ PM2.5  : num  3 12 14 12 12 11 12 13 8 3 ...
 $ PM10   : num  6 12 14 12 12 11 12 13 8 6 ...
 $ SO2    : num  3 3 NA 3 3 3 3 3 3 3 ...
 $ NO2    : num  8 7 7 5 NA 7 9 23 19 21 ...
 $ CO     : int  300 300 200 NA 200 200 300 300 400 400 ...
 $ O3     : num  44 47 22 NA 11 45 74 59 66 60 ...
 $ TEMP   : num  -0.9 -1.1 -1.7 -2.1 -2.4 -2.8 -4 -2.4 -1 0 ...
 $ PRES   : num  1026 1026 1026 1027 1028 ...
 $ DEWP   : num  -20.5 -21.3 -23 -23.3 -22.9 -22.1 -21.2 -21.3 -21.8 -22.9 ...
 $ RAIN   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ wd     : chr  "NW" "NW" "NW" "NW" ...
 $ WSPM   : num  9.3 9.4 8.6 6.6 4.5 1.7 1.6 1.7 2.7 0.8 ...
 $ station: chr  "Shunyi" "Shunyi" "Shunyi" "Shunyi" ...
```

Hide

```
head(data)
```

	No <int>	year <int>	month <int>	day <int>	hour <int>	PM2.5 <dbl>	PM10 <dbl>	SO2 <dbl>	NO2 <dbl>
1	1	2013	3	1	0	3	6	3	8
2	2	2013	3	1	1	12	12	3	7
3	3	2013	3	1	2	14	14	*NA*	7
4	4	2013	3	1	3	12	12	3	5
5	5	2013	3	1	4	12	12	3	*NA*
6	6	2013	3	1	5	11	11	3	7

6 rows | 1-10 of 18 columns

#INDIVIDUAL ANALYSIS

Complete individual analysis of each attribute (except the first and last column which just gives number and station respectively)

Hide

```
#install.packages("viridis")
library("viridis")
```

```
Warning: package ‘viridis’ was built under R version 4.1.3
Loading required package: viridisLite
```

#Year,Month,Day,Hour

Hide

```
#check for missing values
sum(is.na(data$year))
```

```
[1] 0
```

Hide

```
sum(is.na(data$month))
```

```
[1] 0
```

Hide

```
sum(is.na(data$day))
```

```
[1] 0
```

Hide

```
sum(is.na(data$hour))
```

```
[1] 0
```

Hide

```
#there are no missing values
```

##PM2.5

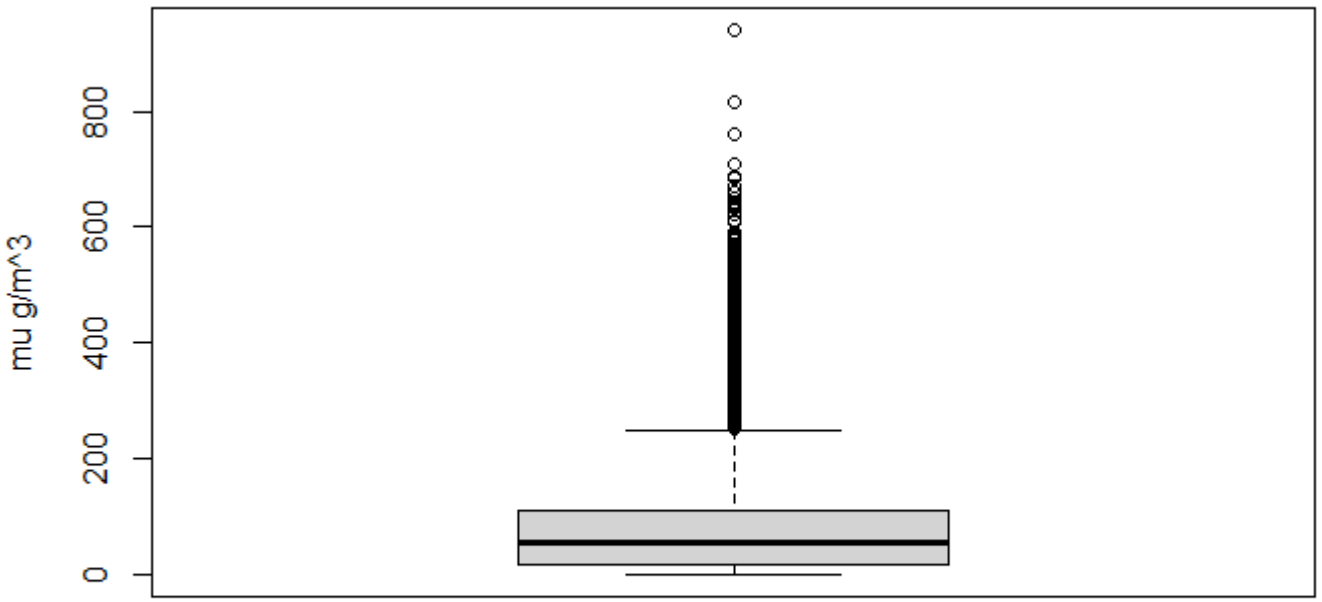
Hide

```
#summary statistics
summary(data$PM2.5)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
   2.00   19.00   55.00   79.49  112.00  941.00     913
```

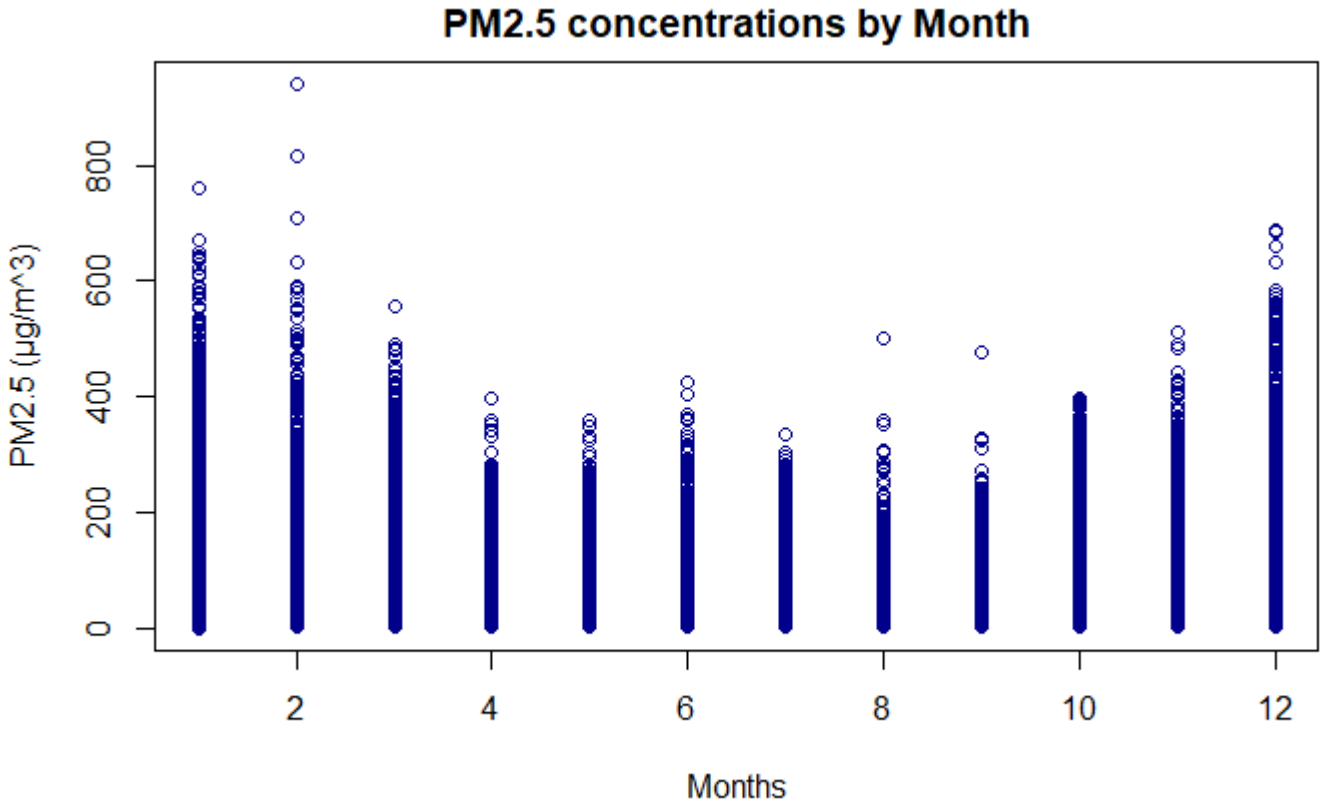
Hide

```
#boxplot (looking for outliers)
boxplot(data$PM2.5,ylab = "mu g/m^3")
```



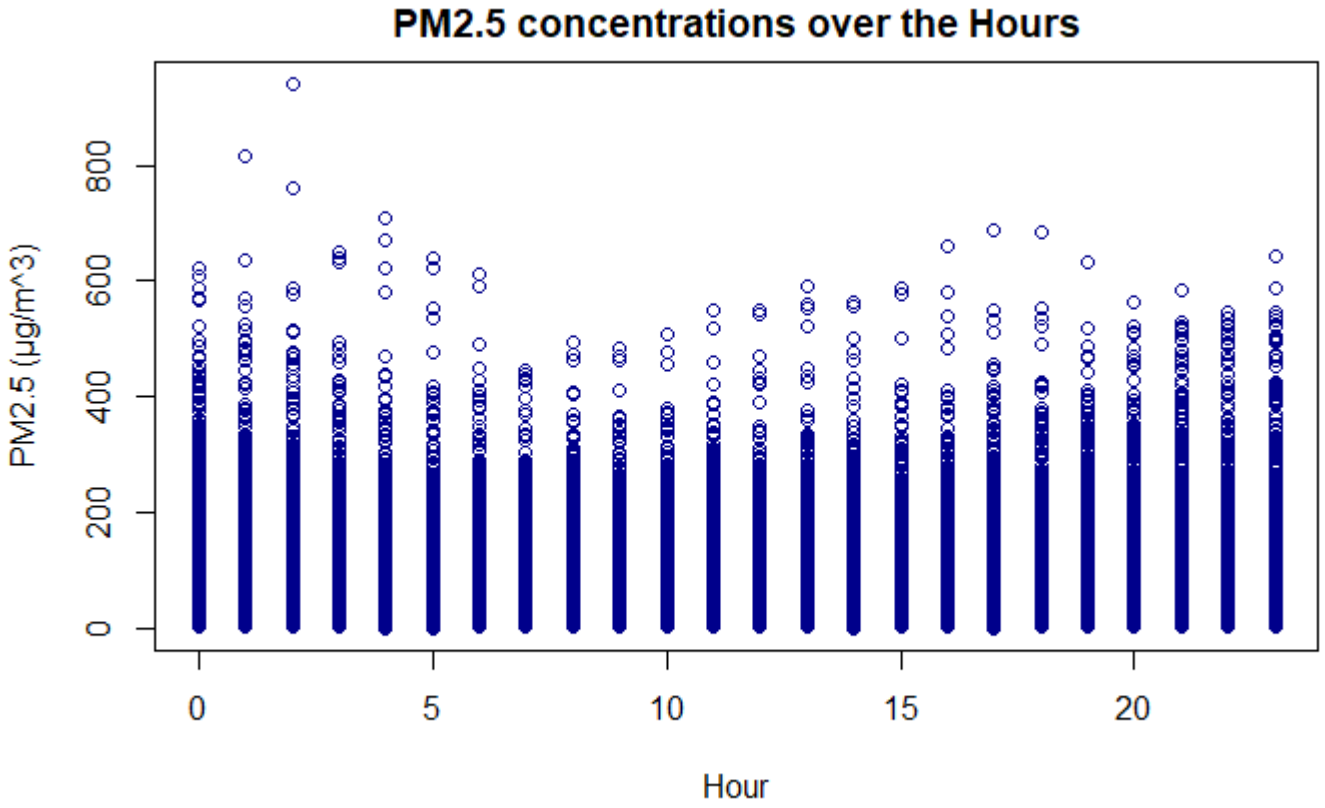
Hide

```
#plot PM2.5 data against time variables
plot(data$month, data$PM2.5, col = "blue4", xlab="Months",ylab="PM2.5 (µg/m^3)",main="PM2.5 concentrations by Month")
```



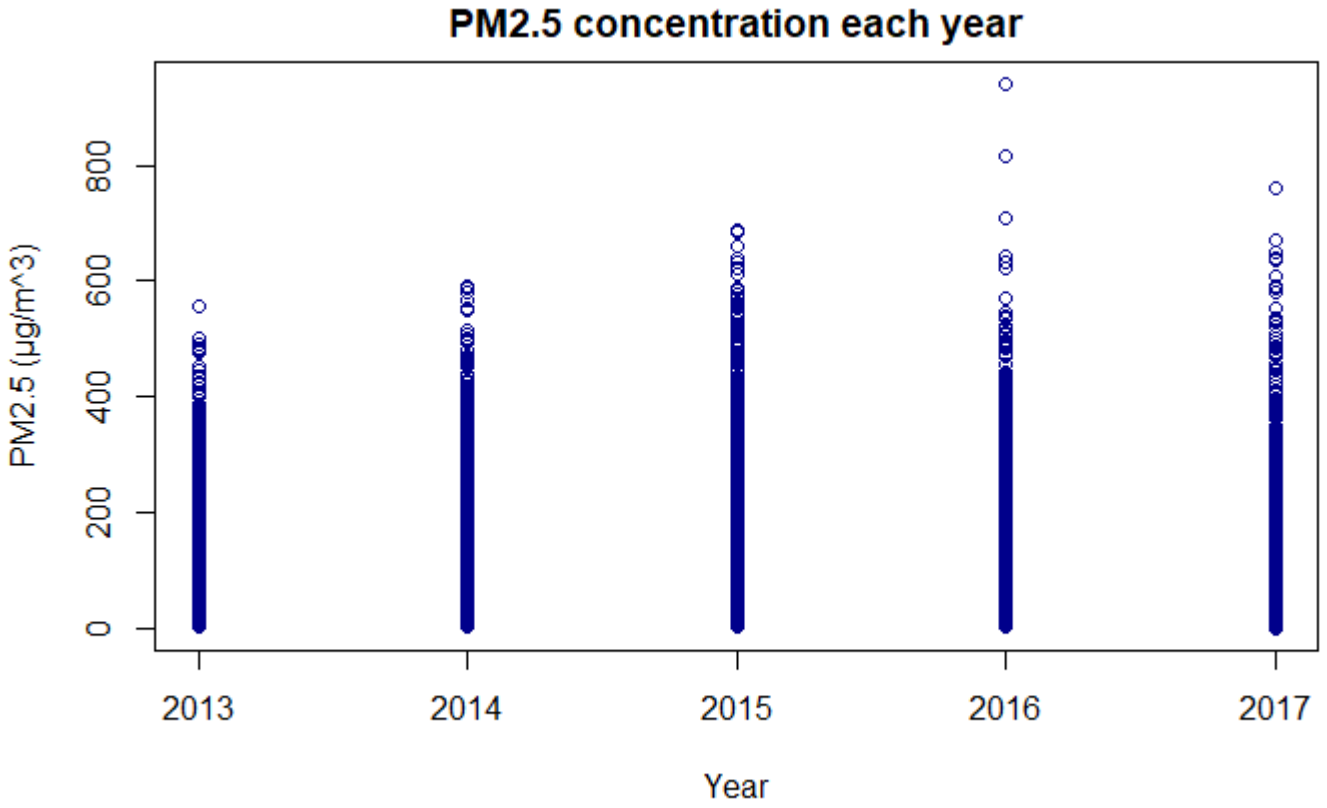
Hide

```
plot(data$hour, data$PM2.5, col="blue4", xlab="Hour",ylab="PM2.5 (µg/m^3)",main="PM2.5 concentrations over the Hours")
```



Hide

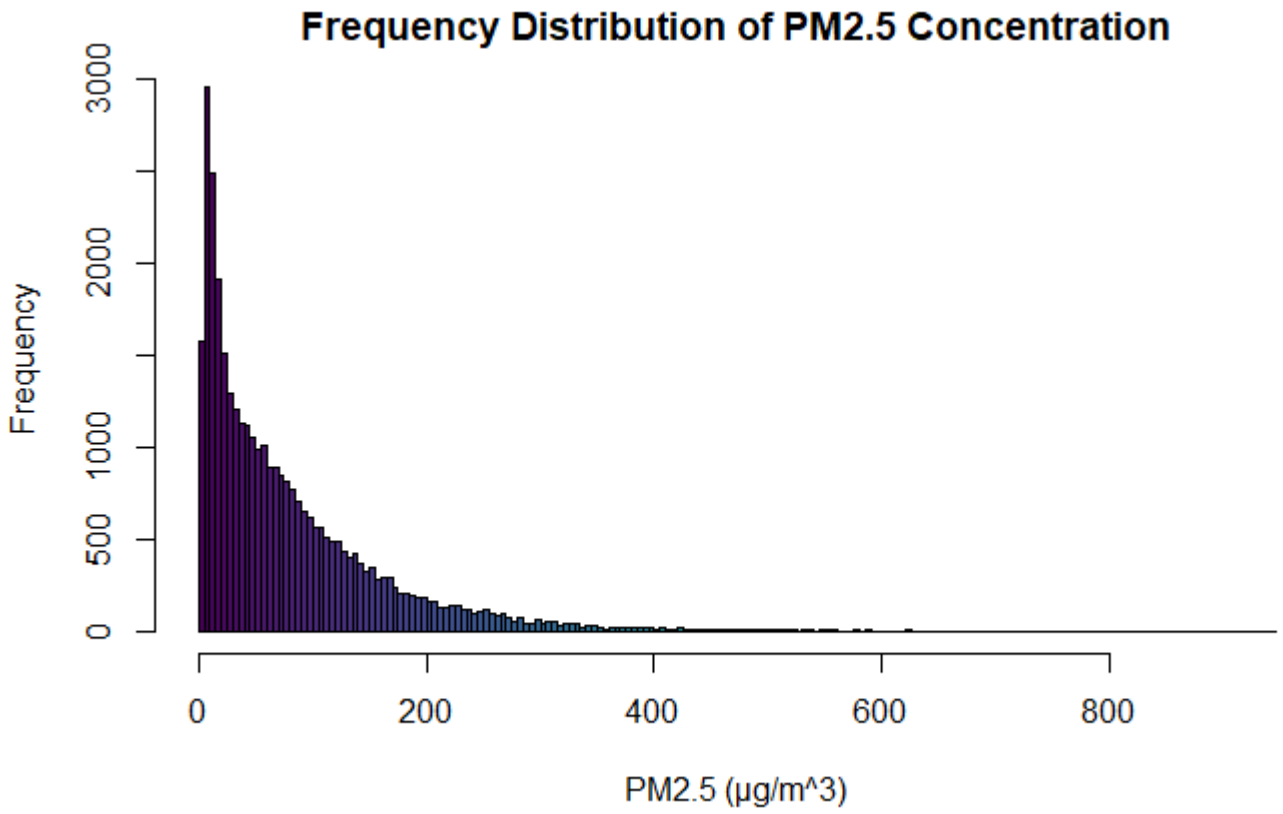
```
plot(data$year, data$PM2.5, col="blue4", xlab="Year", ylab="PM2.5 (µg/m^3)", main="PM2.5 concentration each year")
```



Hide

```
PM2.5 <- data$PM2.5

#create a histogram to visualize frequency of values
hist(PM2.5, breaks = sqrt(nrow(data)), xlab = "PM2.5 (µg/m^3)", main = "Frequency Distribution of PM2.5 Concentration" , col
=viridis(sqrt(nrow(data))))
```



Hide

```
NA
NA
NA
NA
NA
```

##PM10

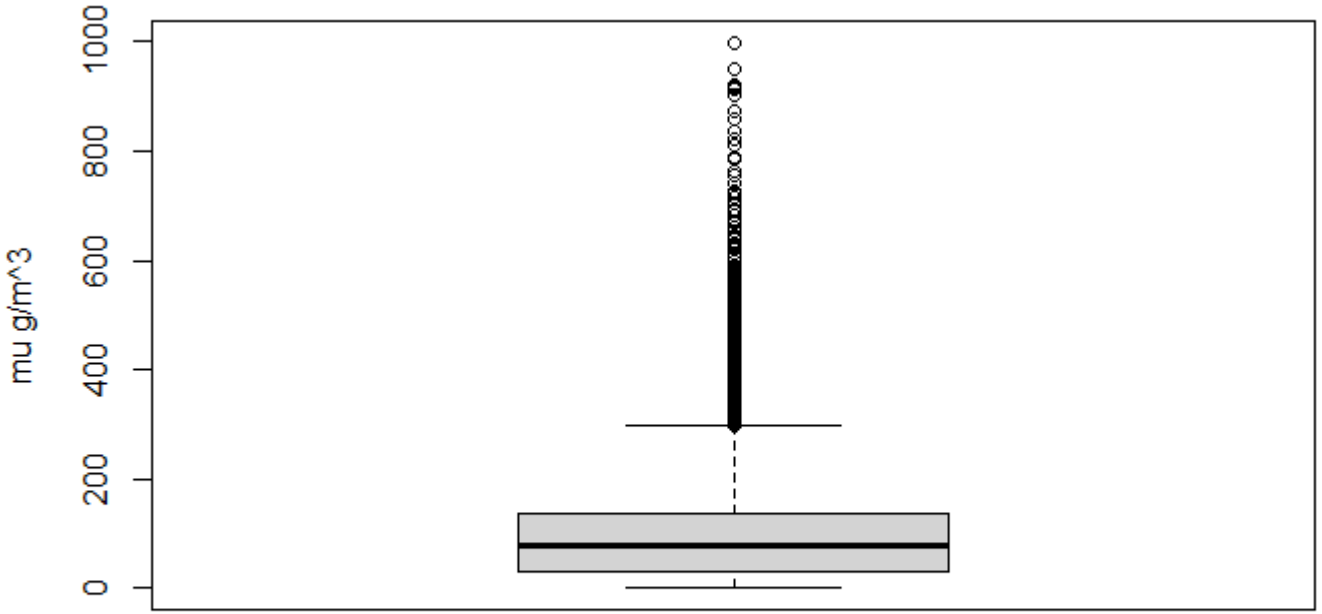
Hide

```
#summary stats
summary(data$PM10)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
   2.00   31.00   77.00   98.74  138.00  999.00     548
```

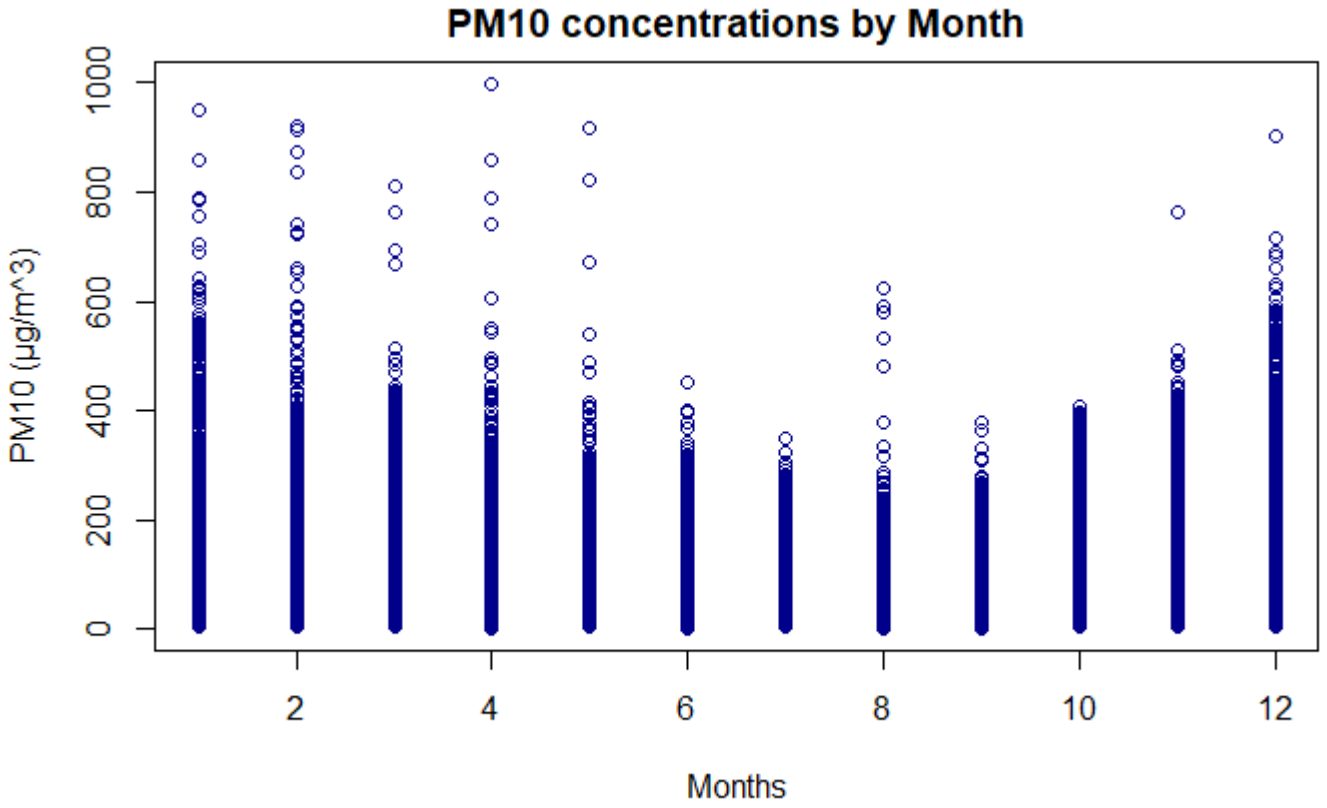
Hide

```
#looking at outliers
boxplot(data$PM10,ylab = "mu g/m^3")
```



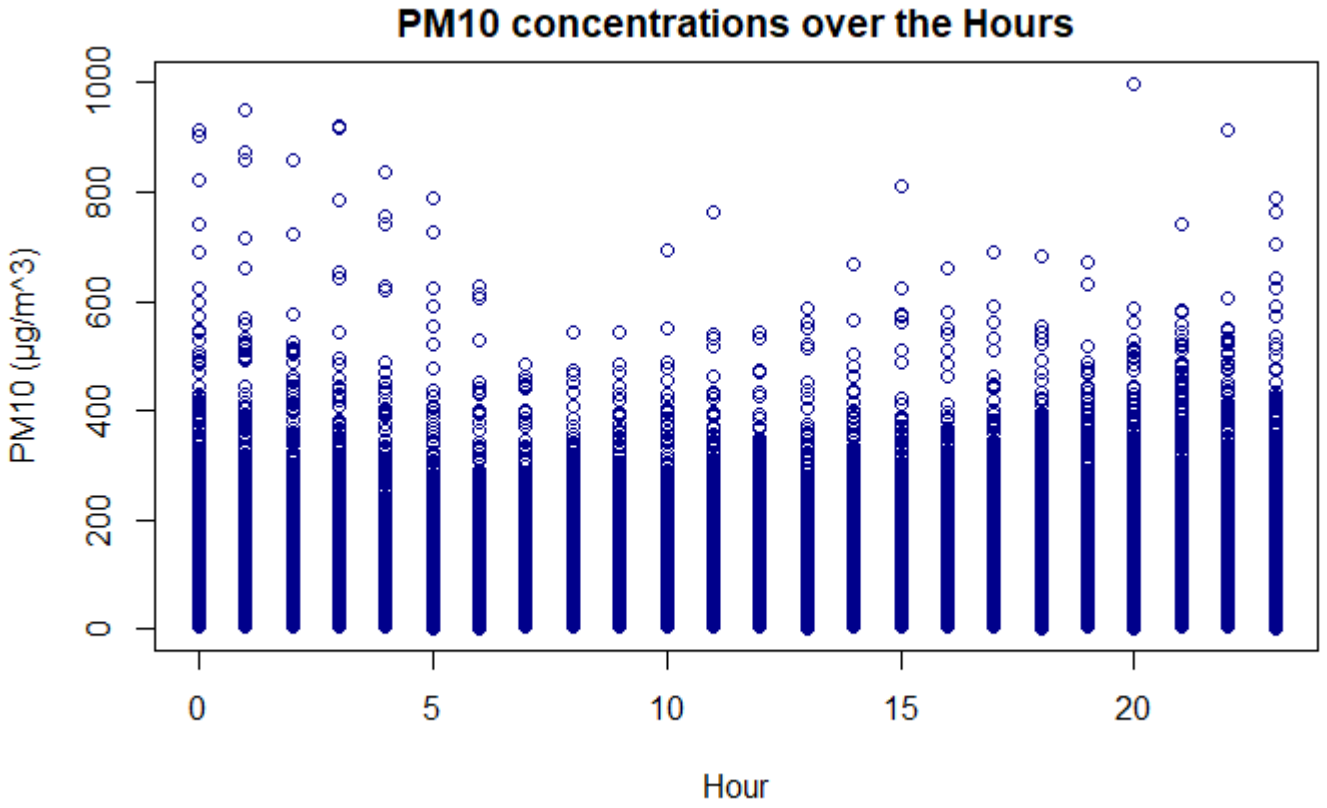
Hide

```
#plot values against time variables
plot(data$month, data$PM10, col="blue4", xlab="Months",ylab="PM10 (µg/m^3)",main="PM10 concentrations by Month")
```



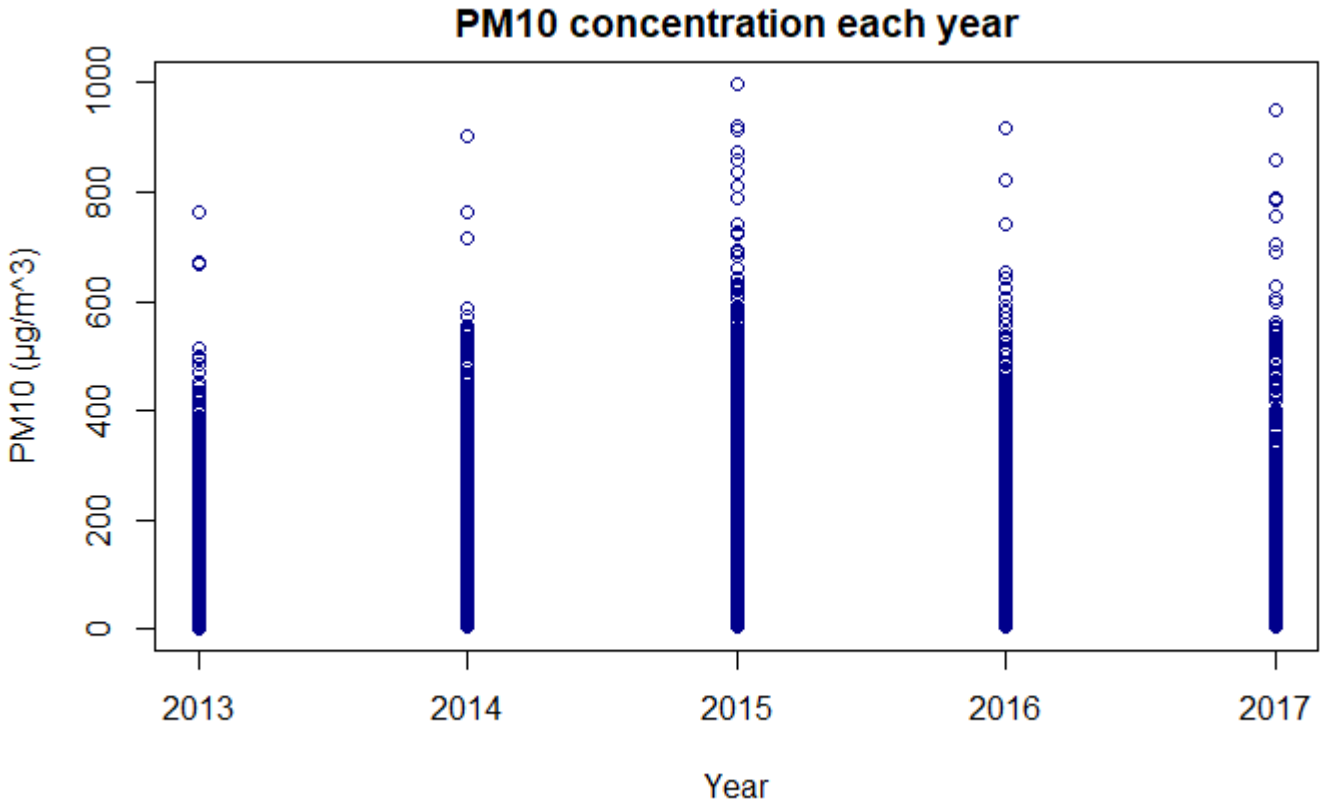
Hide

```
plot(data$hour, data$PM10, col="blue4", xlab="Hour",ylab="PM10 (µg/m^3)",main="PM10 concentrations over the Hours")
```



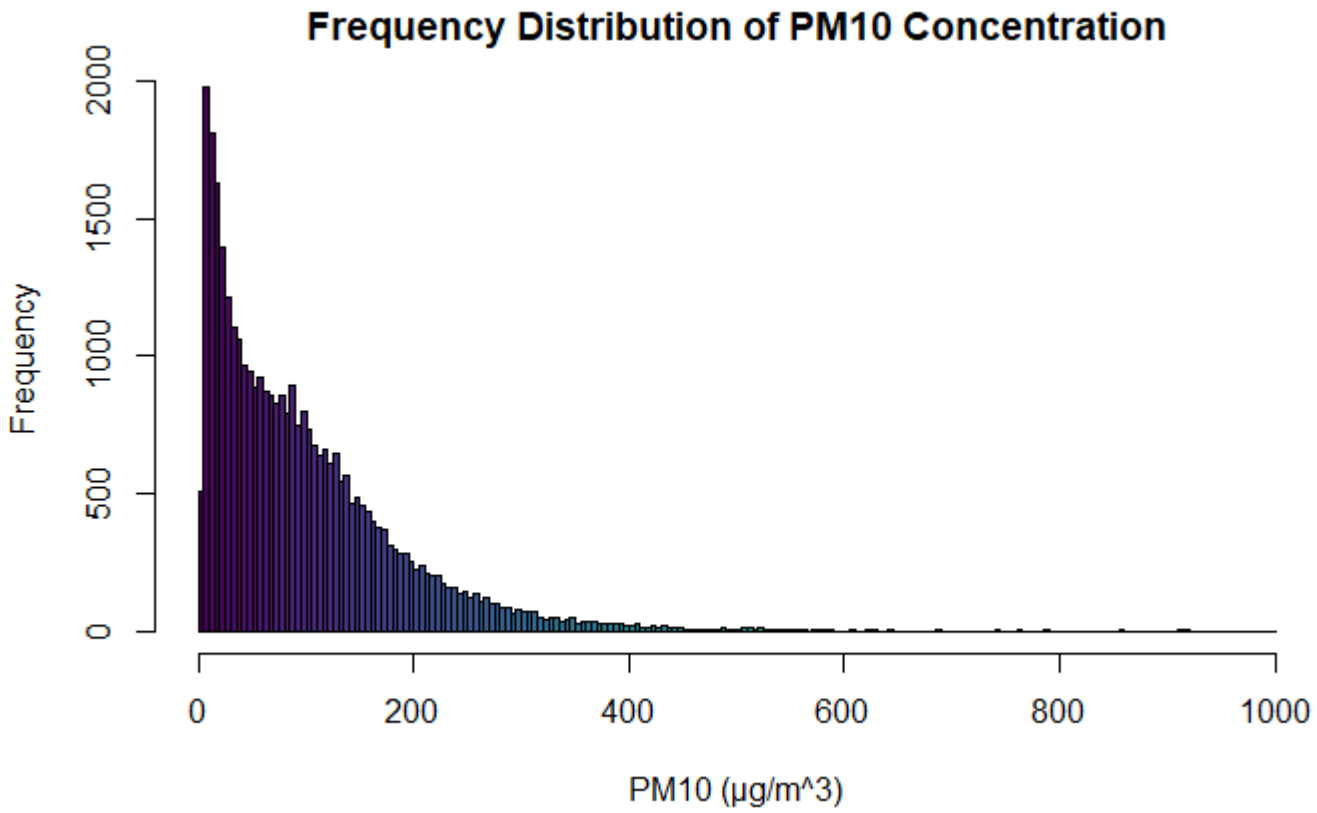
Hide

```
plot(data$year, data$PM10, col="blue4", xlab="Year", ylab="PM10 (µg/m^3)", main="PM10 concentration each year")
```



Hide

```
#check distribution of values
hist(data$PM10, breaks = sqrt(nrow(data)), xlab="PM10 (µg/m^3)", main = "Frequency Distribution of PM10 Concentration" , col
=viridis(sqrt(nrow(data))))
```



##SO2

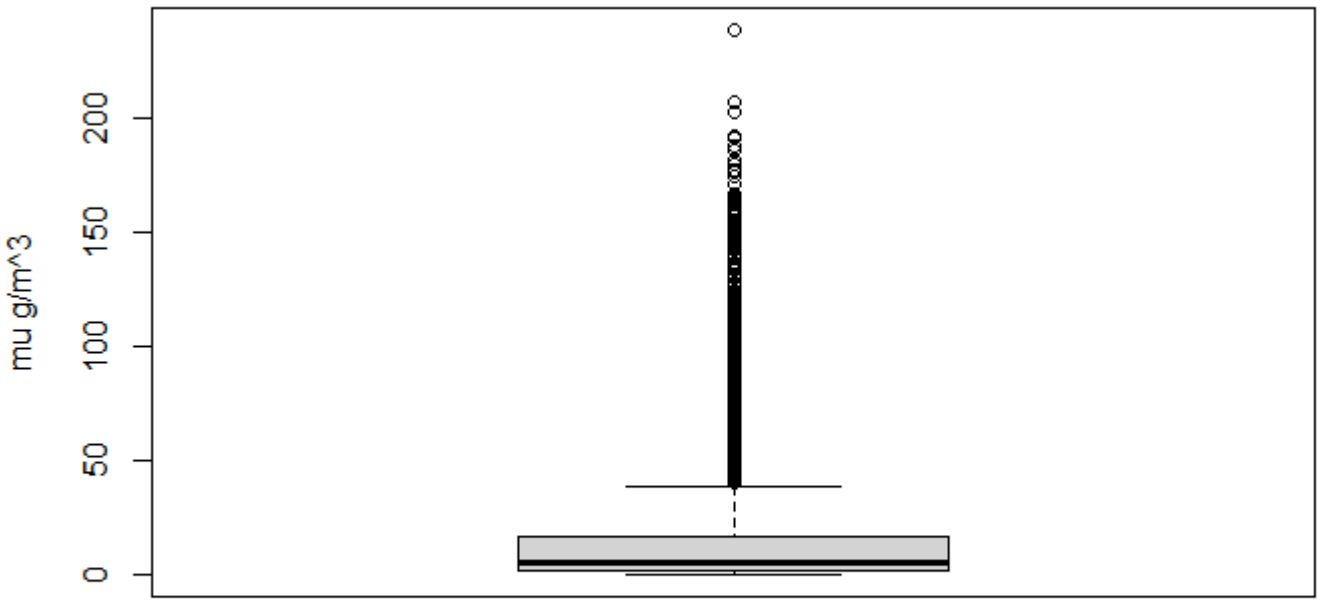
Hide

```
#summary stats
summary(data$SO2)
```

```
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.     NA's
  0.2856   2.0000   5.0000  13.5720  17.0000 239.0000     1296
```

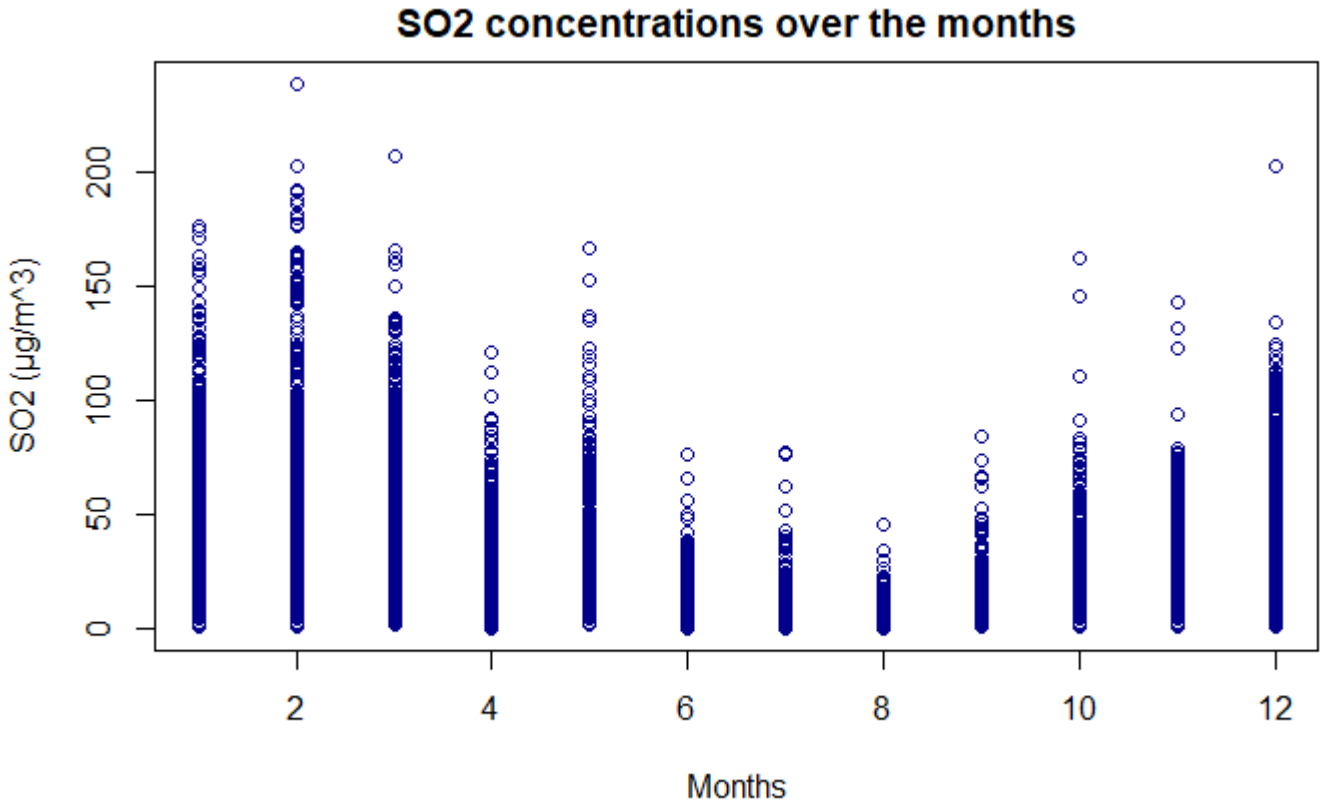
Hide

```
#outliers
boxplot(data$SO2,ylab = "mu g/m^3")
```



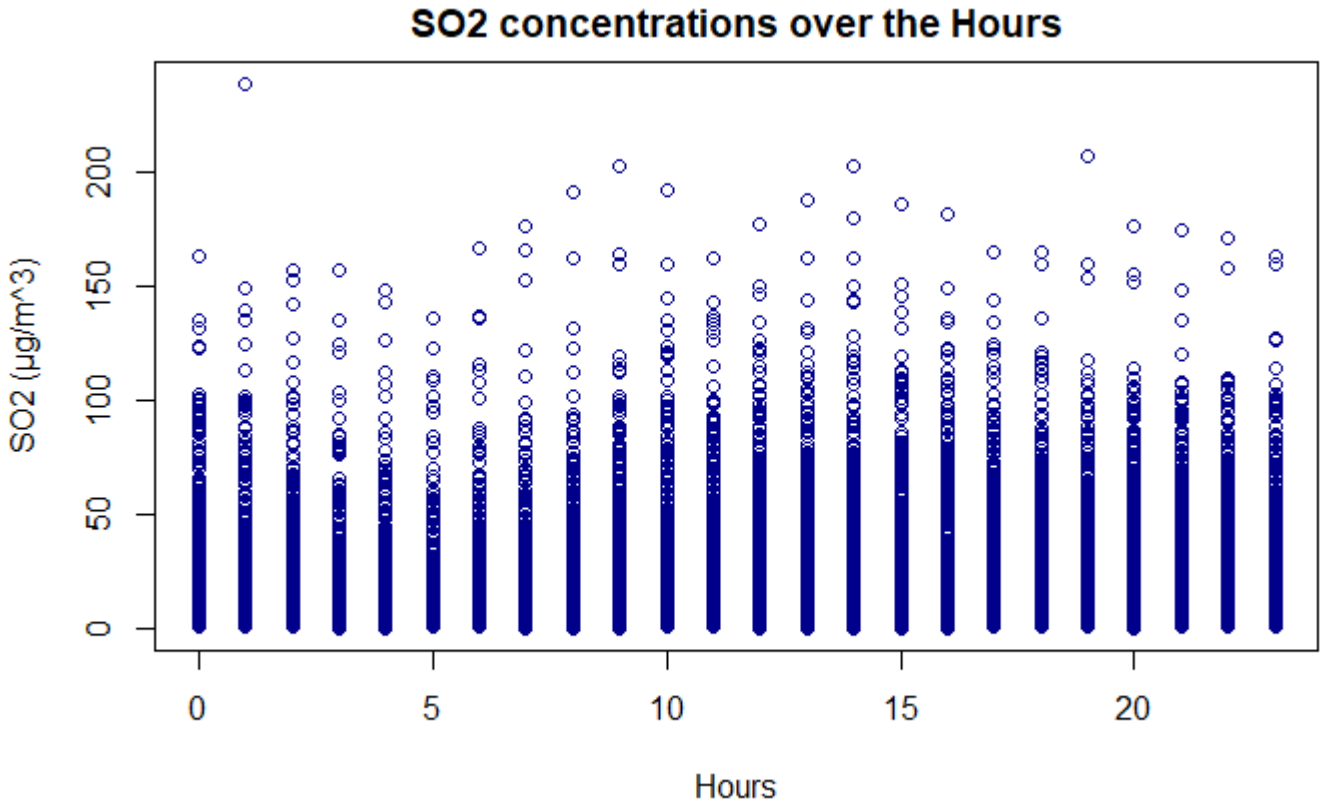
Hide

```
#plot values against time variables
plot(data$month, data$SO2, col="blue4", xlab="Months",ylab="SO2 (µg/m^3)",main="SO2 concentrations over the months")
```



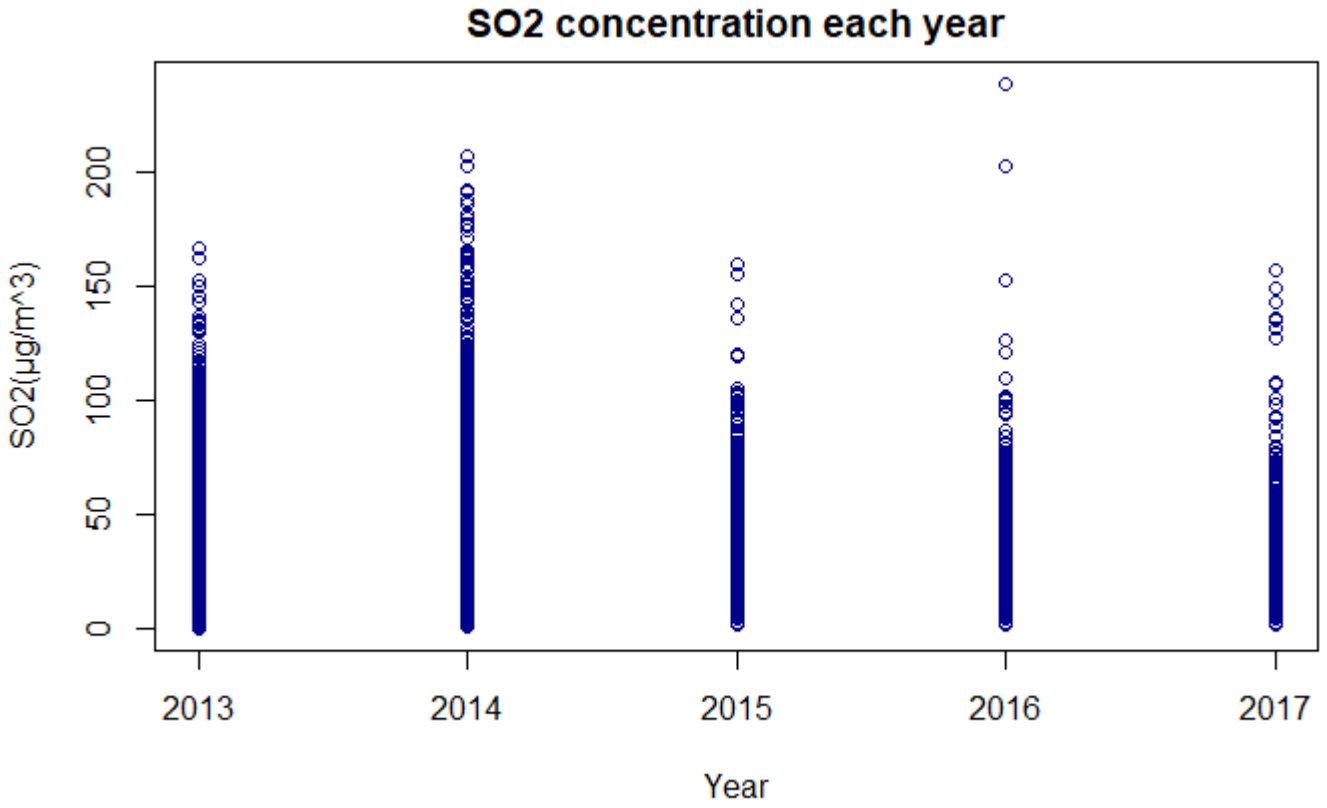
Hide

```
plot(data$hour, data$SO2, col="blue4", xlab="Hours",ylab="SO2 (µg/m^3)",main="SO2 concentrations over the Hours")
```



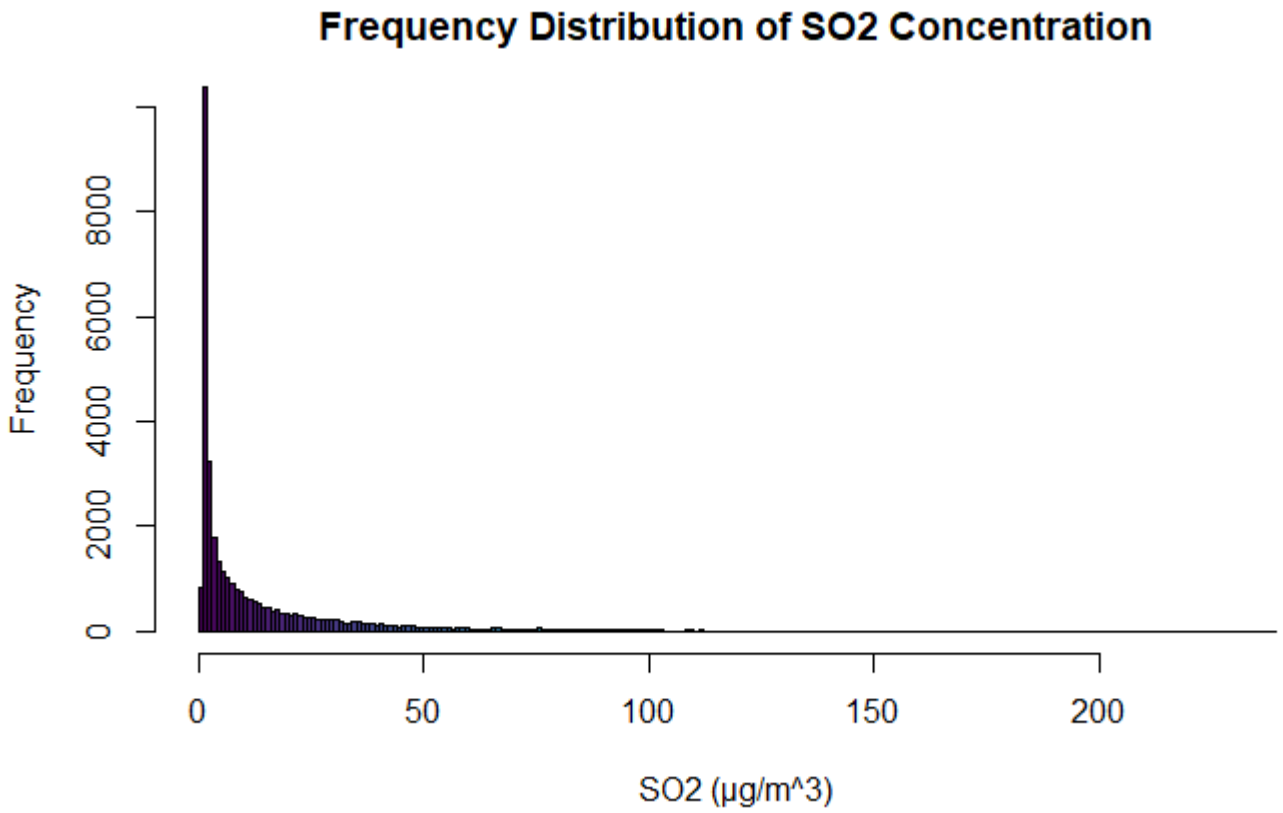
Hide

```
plot(data$year, data$SO2, col="blue4", xlab="Year", ylab="SO2(µg/m^3)", main="SO2 concentration each year")
```



Hide

```
#check frequency of values
hist(data$SO2, breaks = sqrt(nrow(data)), xlab="SO2 (µg/m^3)", main = "Frequency Distribution of SO2 Concentration" , col=vi
ridis(sqrt(nrow(data))))
```



##NO2

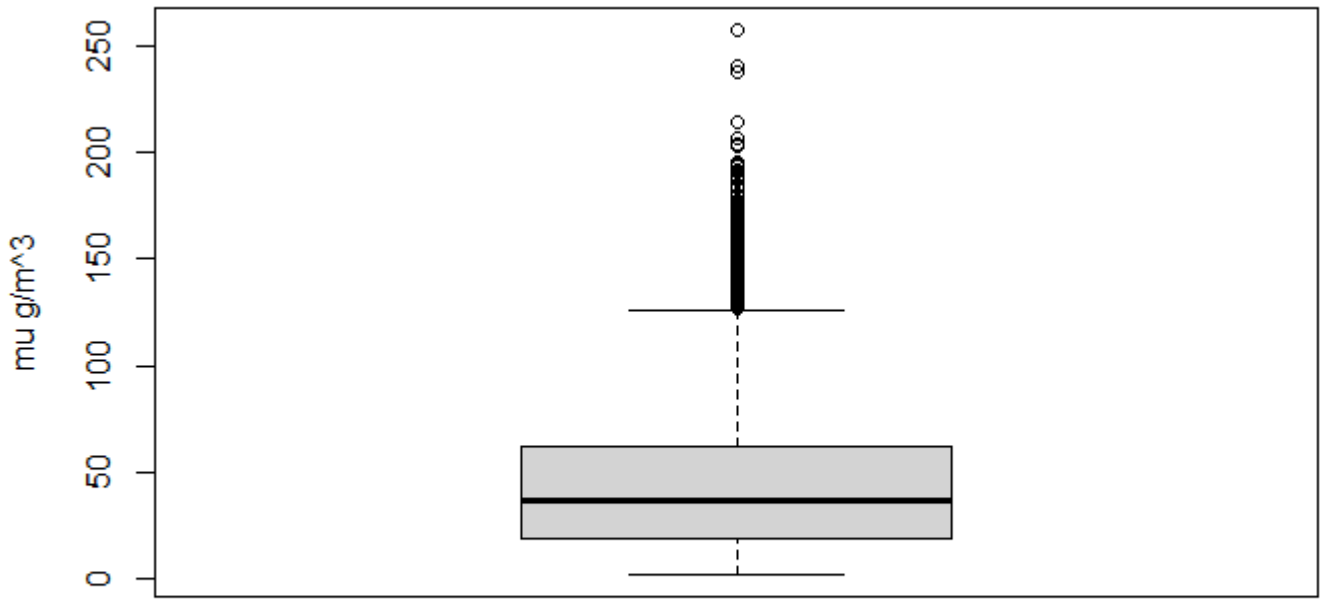
Hide

```
#summary stats
summary(data$NO2)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
   2.00   19.00   37.00   43.91   62.00  258.00    1365
```

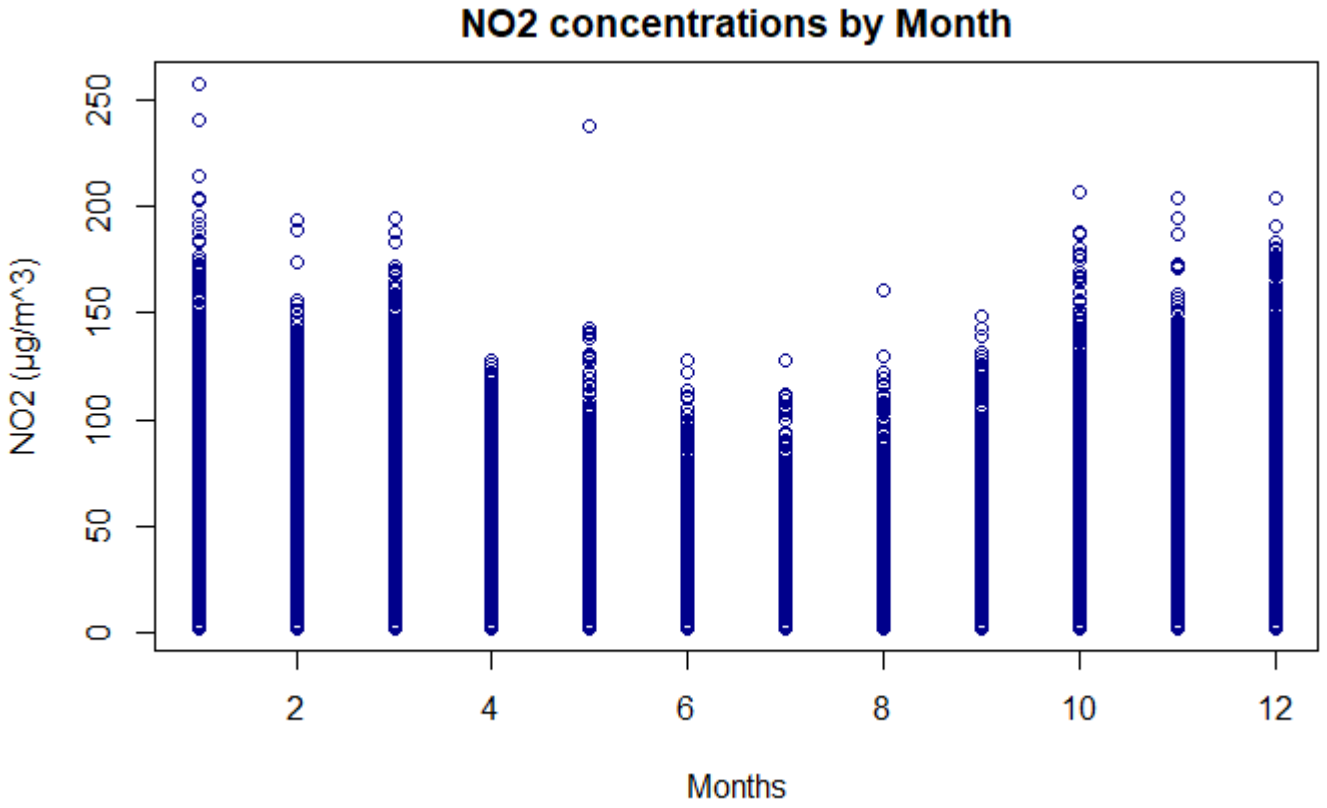
Hide

```
#outliers
boxplot(data$NO2,ylab = "mu g/m^3")
```



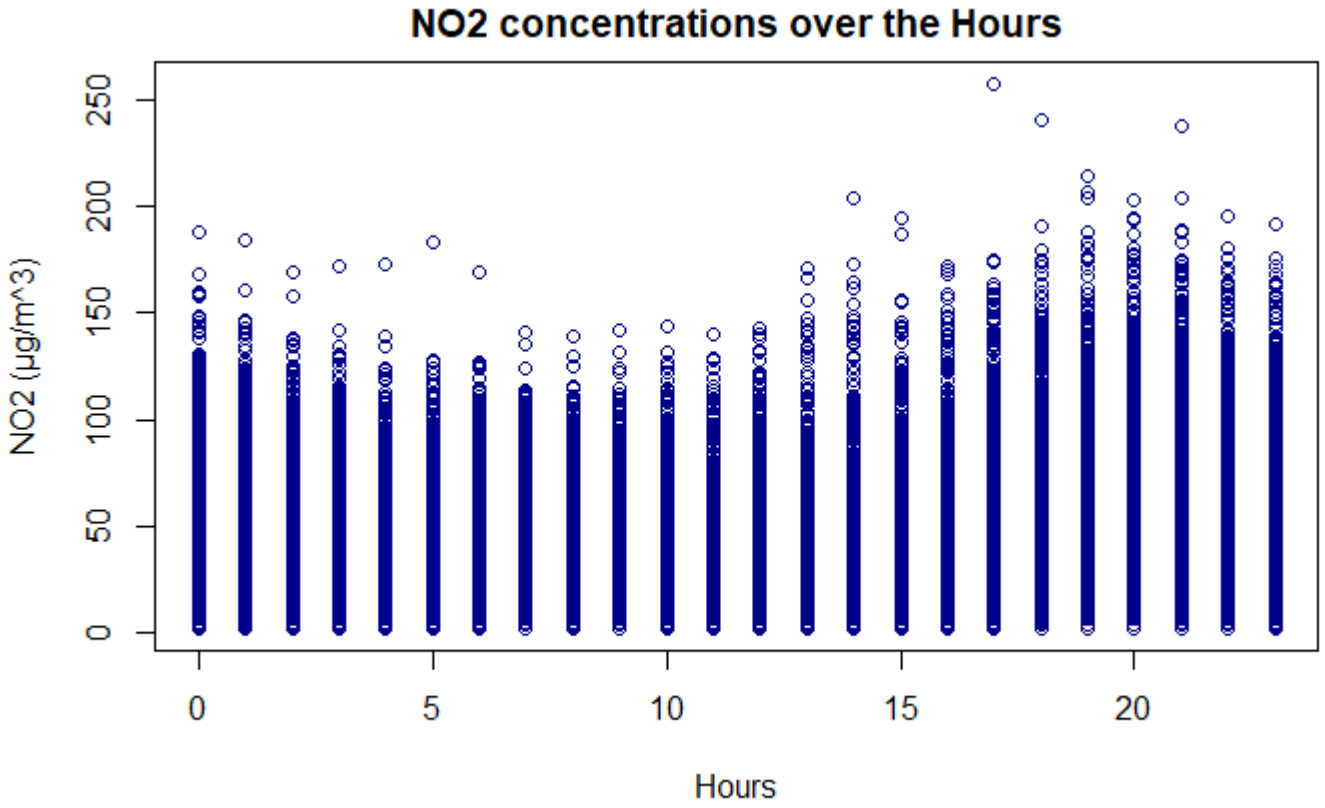
Hide

```
#plot values over time variables
plot(data$month, data$NO2, col="blue4", xlab="Months",ylab="NO2 (µg/m^3)",main="NO2 concentrations by Month")
```



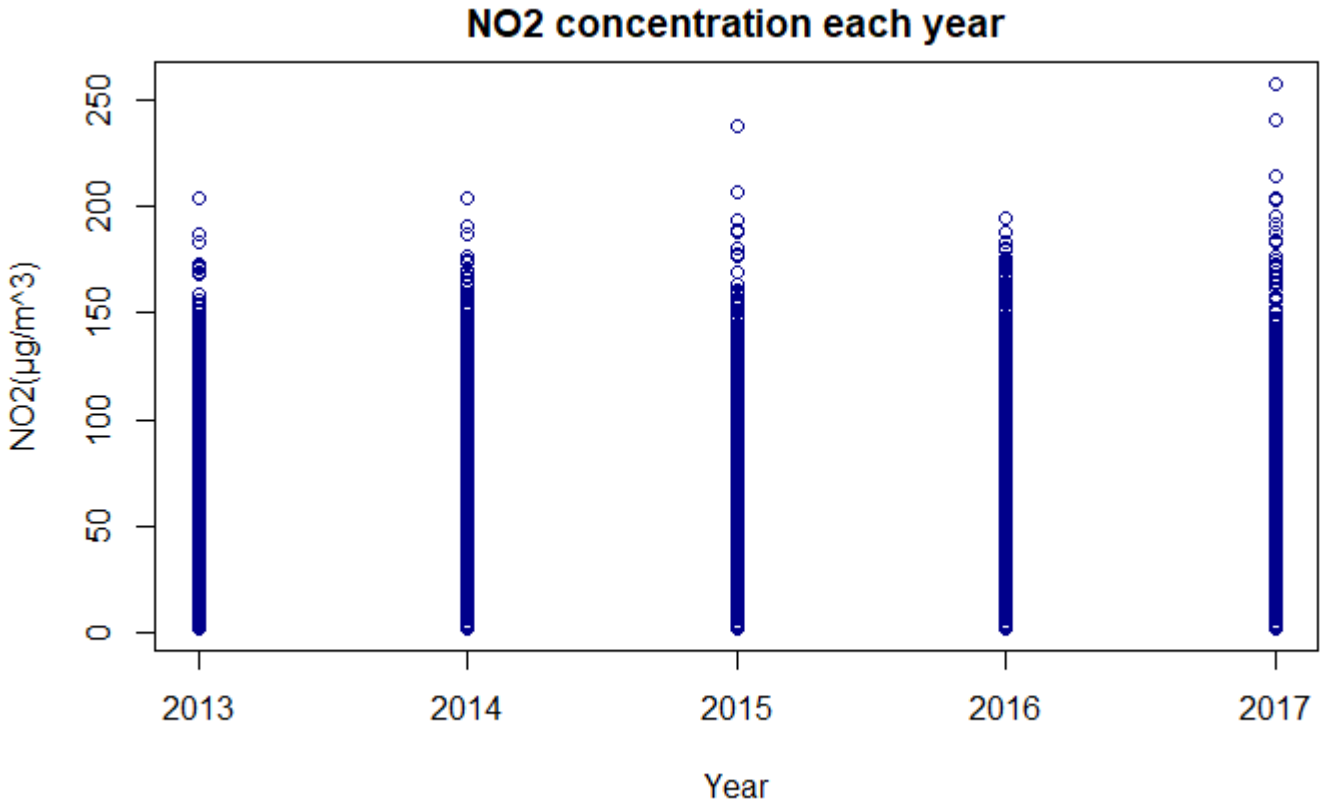
Hide

```
plot(data$hour, data$NO2, col="blue4", xlab="Hours",ylab="NO2 (µg/m^3)",main="NO2 concentrations over the Hours")
```



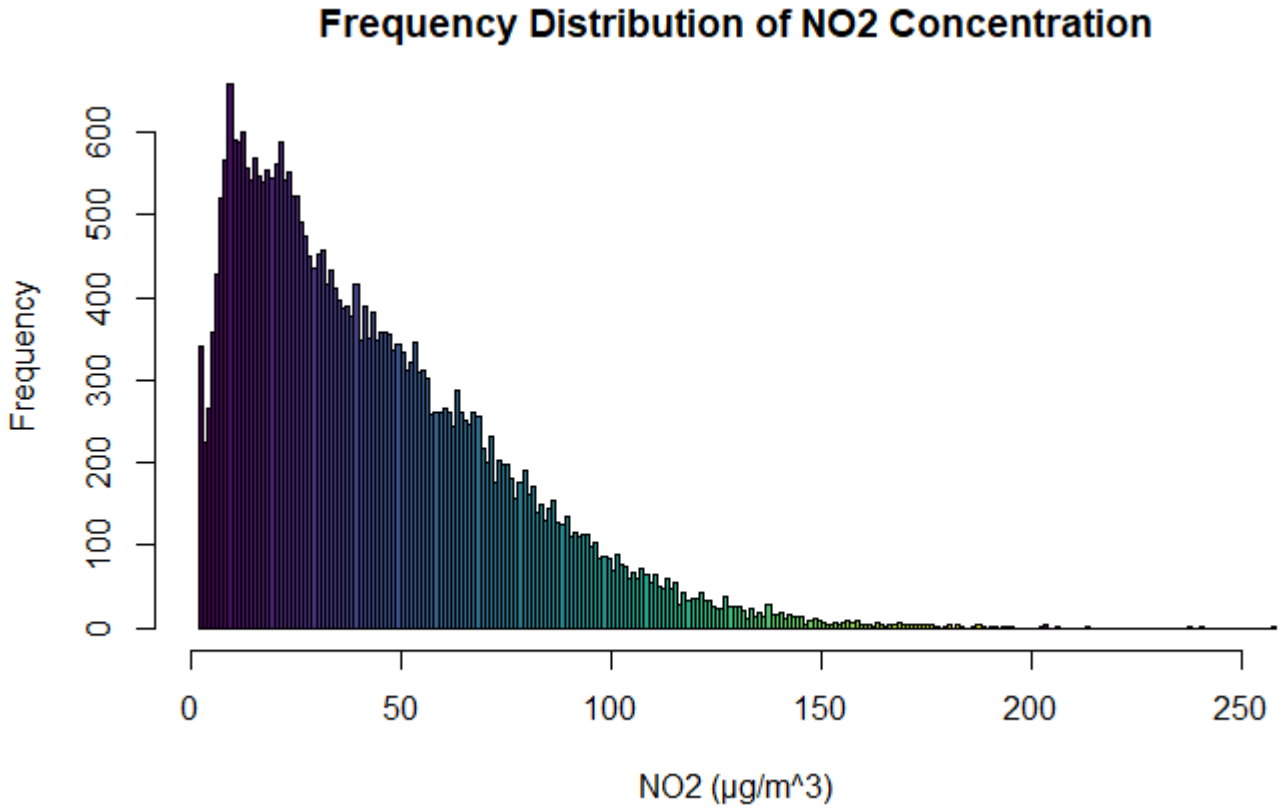
Hide

```
plot(data$year, data$NO2, col="blue4", xlab="Year", ylab="NO2(µg/m^3)", main="NO2 concentration each year")
```



Hide

```
#frequency of values
hist(data$NO2, breaks = sqrt(nrow(data)), xlab="NO2 (µg/m^3)", main = "Frequency Distribution of NO2 Concentration" , col=vi
ridis(sqrt(nrow(data))))
```



##CO

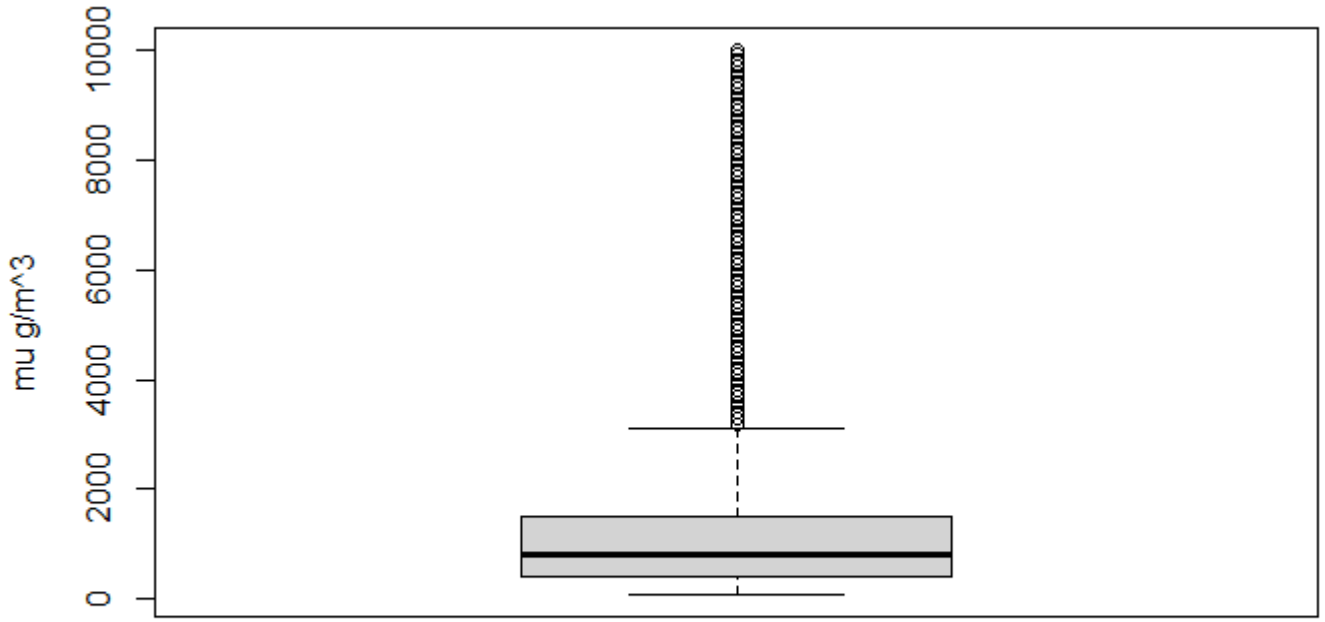
Hide

```
#summary stats
summary(data$CO)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
    100     400     800    1187    1500   10000    2178
```

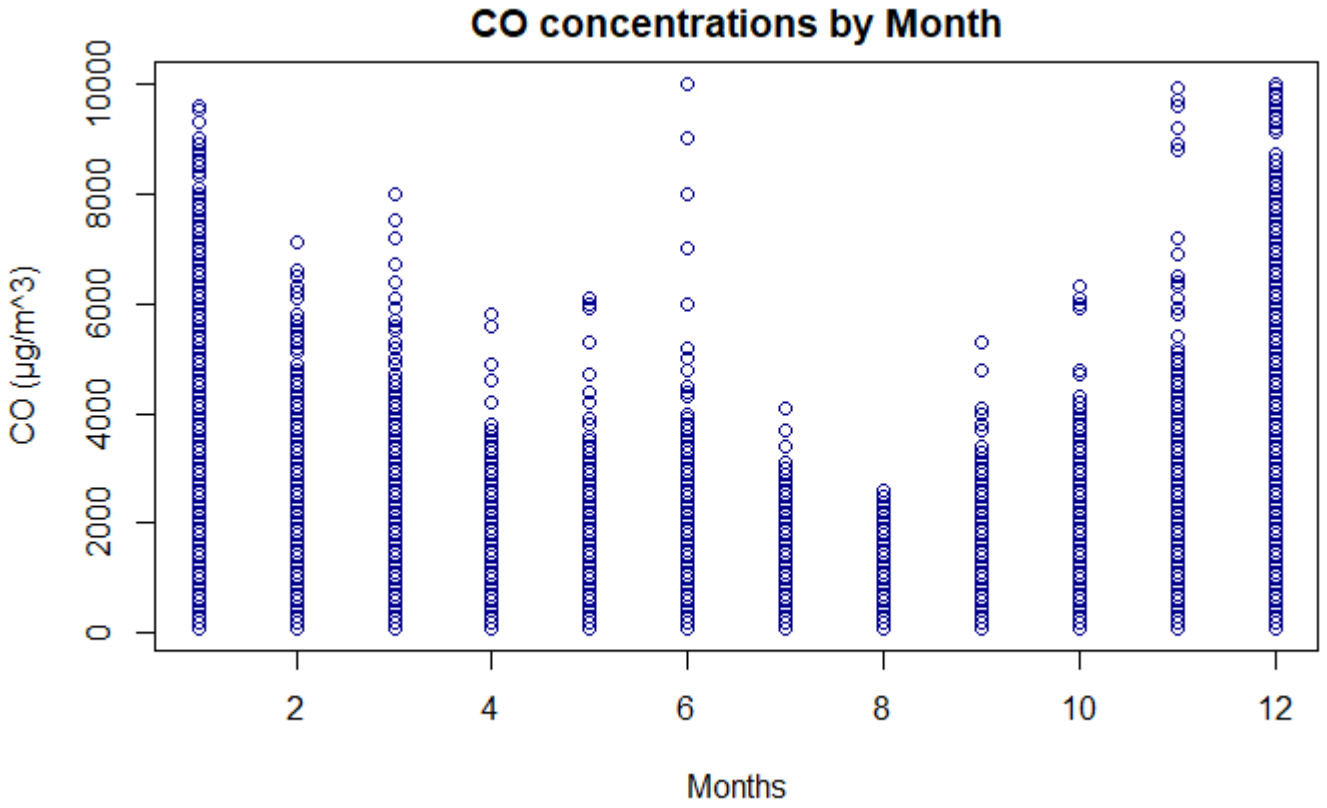
Hide

```
#outliers
boxplot(data$CO,ylab = "mu g/m^3")
```



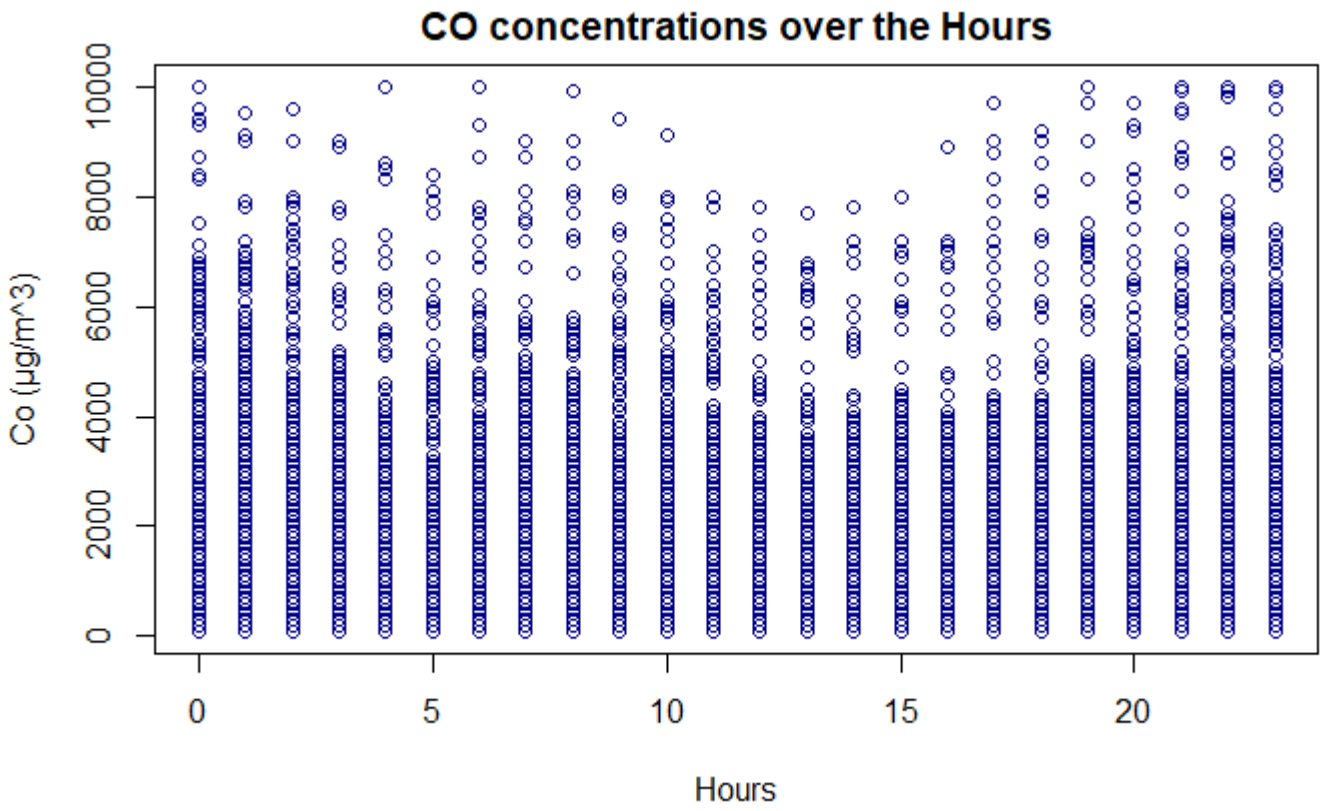
Hide

```
#plot values over time variables
plot(data$month, data$CO, col="blue4", xlab="Months",ylab="CO (µg/m^3)",main="CO concentrations by Month")
```



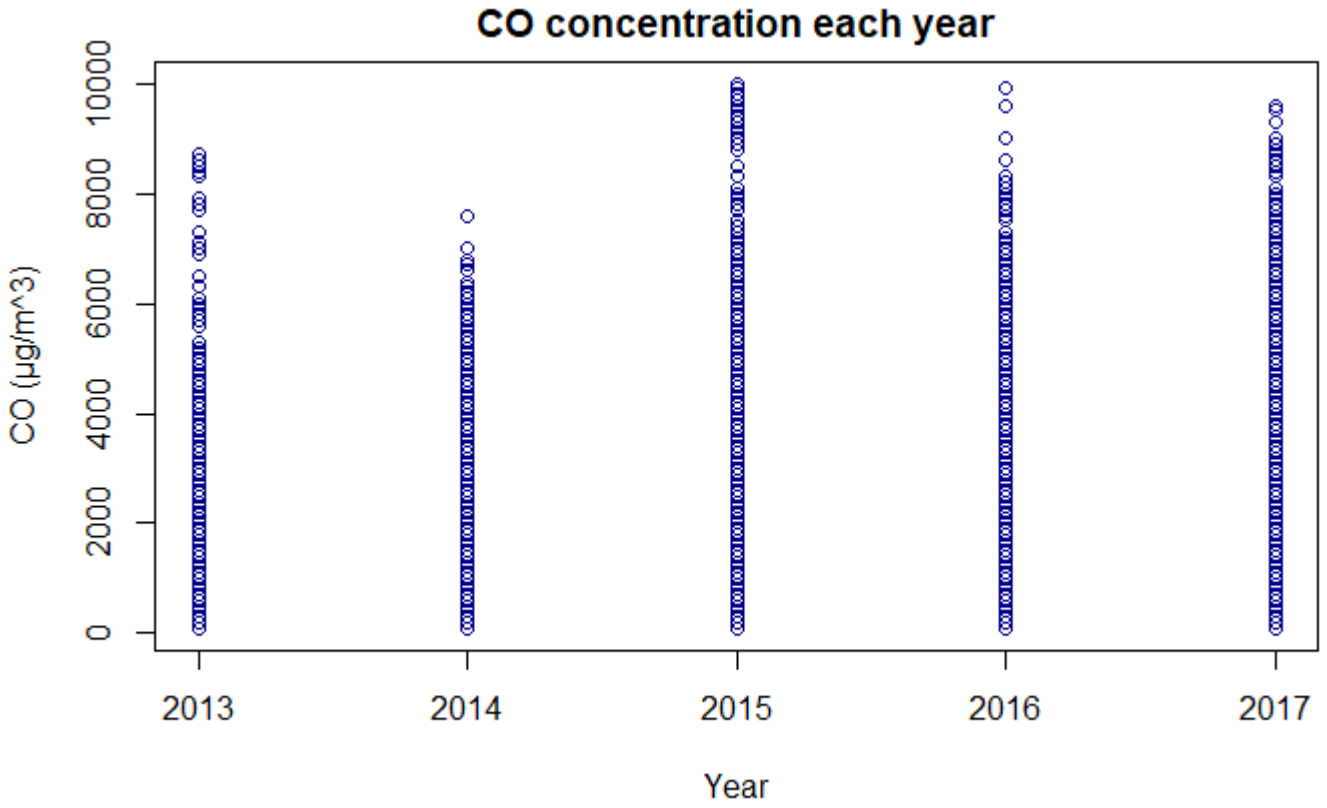
Hide

```
plot(data$hour, data$CO, col="blue4", xlab="Hours",ylab="Co (µg/m^3)",main="CO concentrations over the Hours")
```



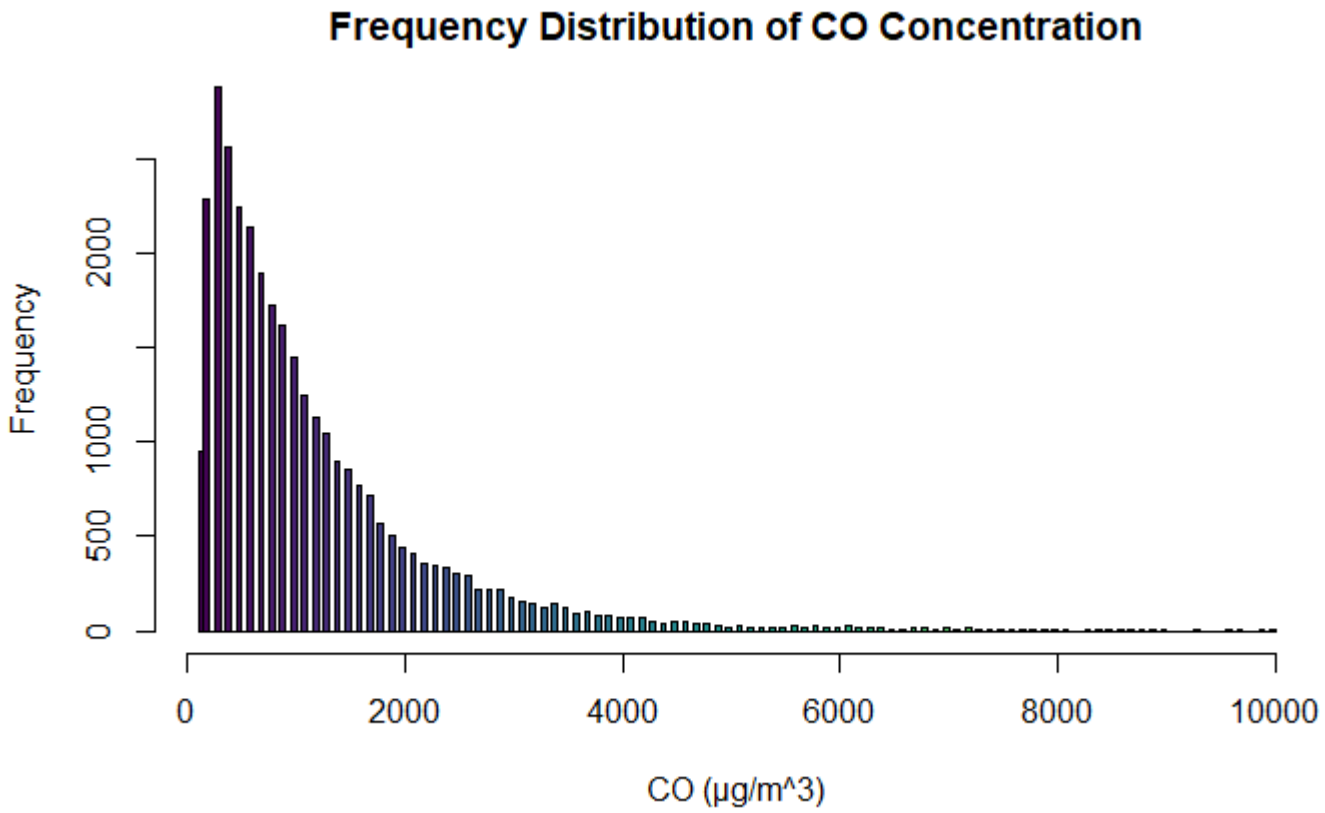
Hide

```
plot(data$year, data$CO, col="blue4", xlab="Year", ylab="CO (µg/m^3)", main="CO concentration each year")
```



Hide

```
#check frequency of values
hist(data$CO, breaks = sqrt(nrow(data)), xlab="CO (µg/m^3)", main = "Frequency Distribution of CO Concentration" , col=viridis(sqrt(nrow(data))))
```



##O3

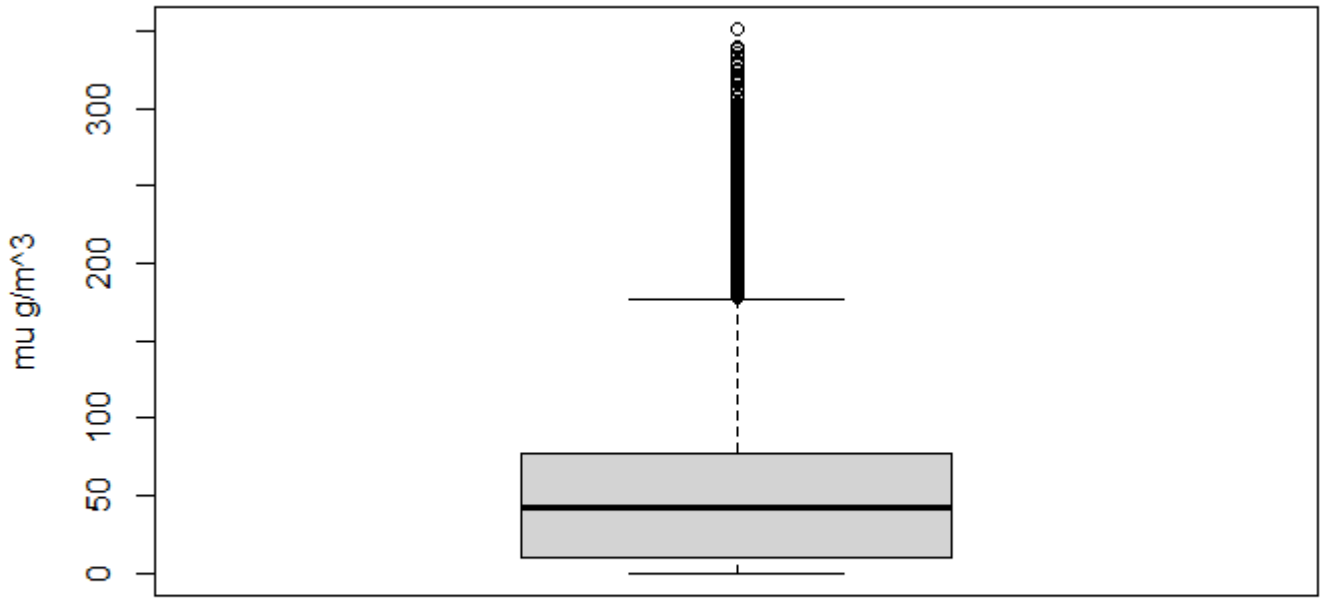
Hide

```
#summary statistics
summary(data$O3)
```

```
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.     NA's
  0.2142  10.0000  43.0000  55.2013  77.0000 351.7164     1489
```

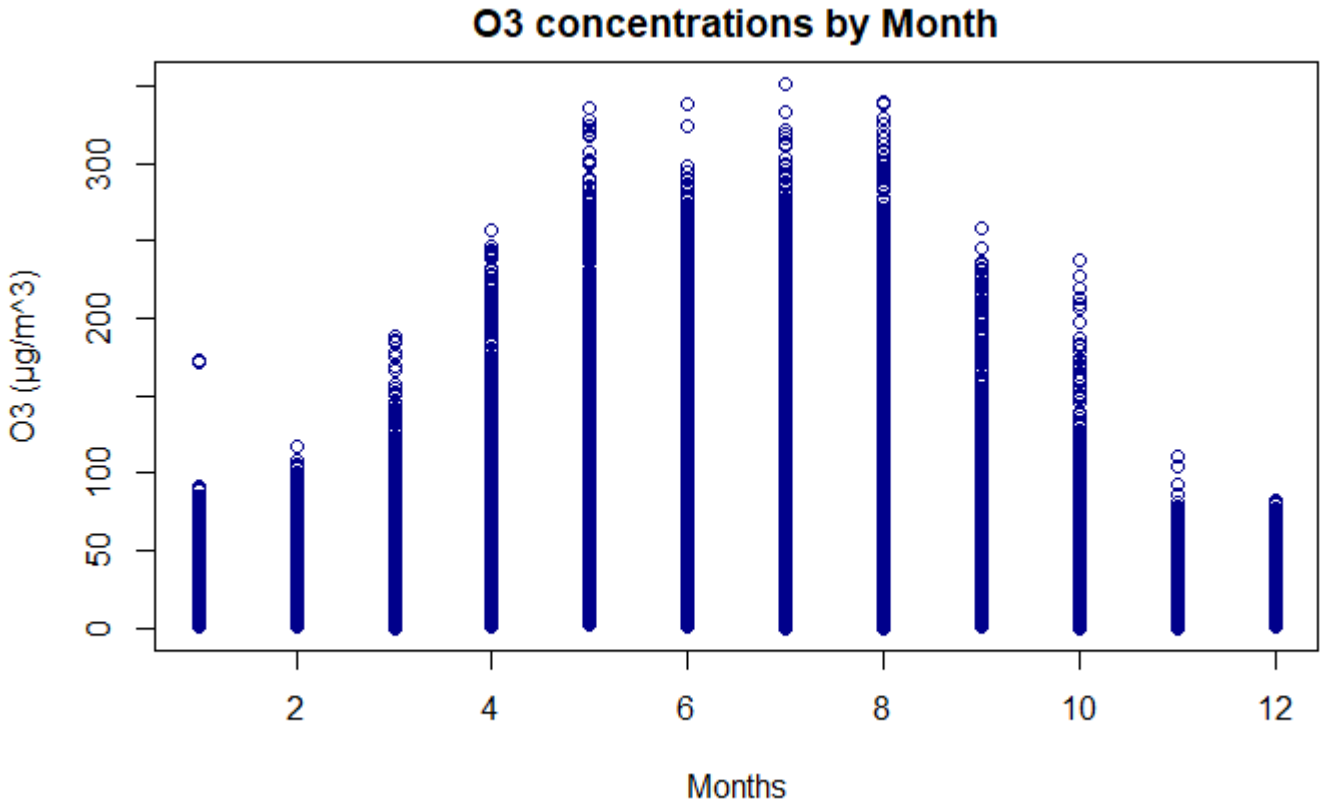
Hide

```
#outliers
boxplot(data$O3,ylab = "mu g/m^3")
```



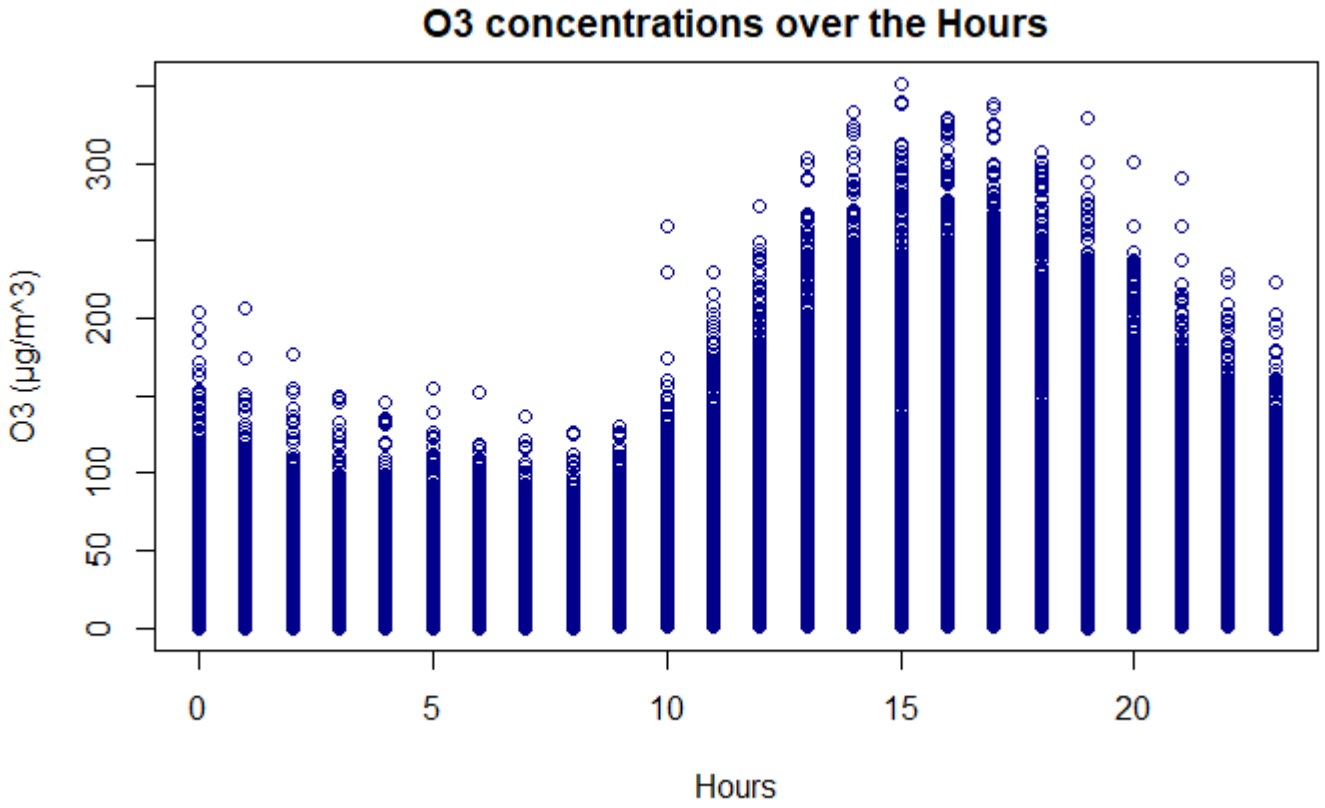
Hide

```
#plot values over time variables
plot(data$month, data$O3, col="blue4", xlab="Months",ylab="O3 (µg/m^3)",main="O3 concentrations by Month")
```



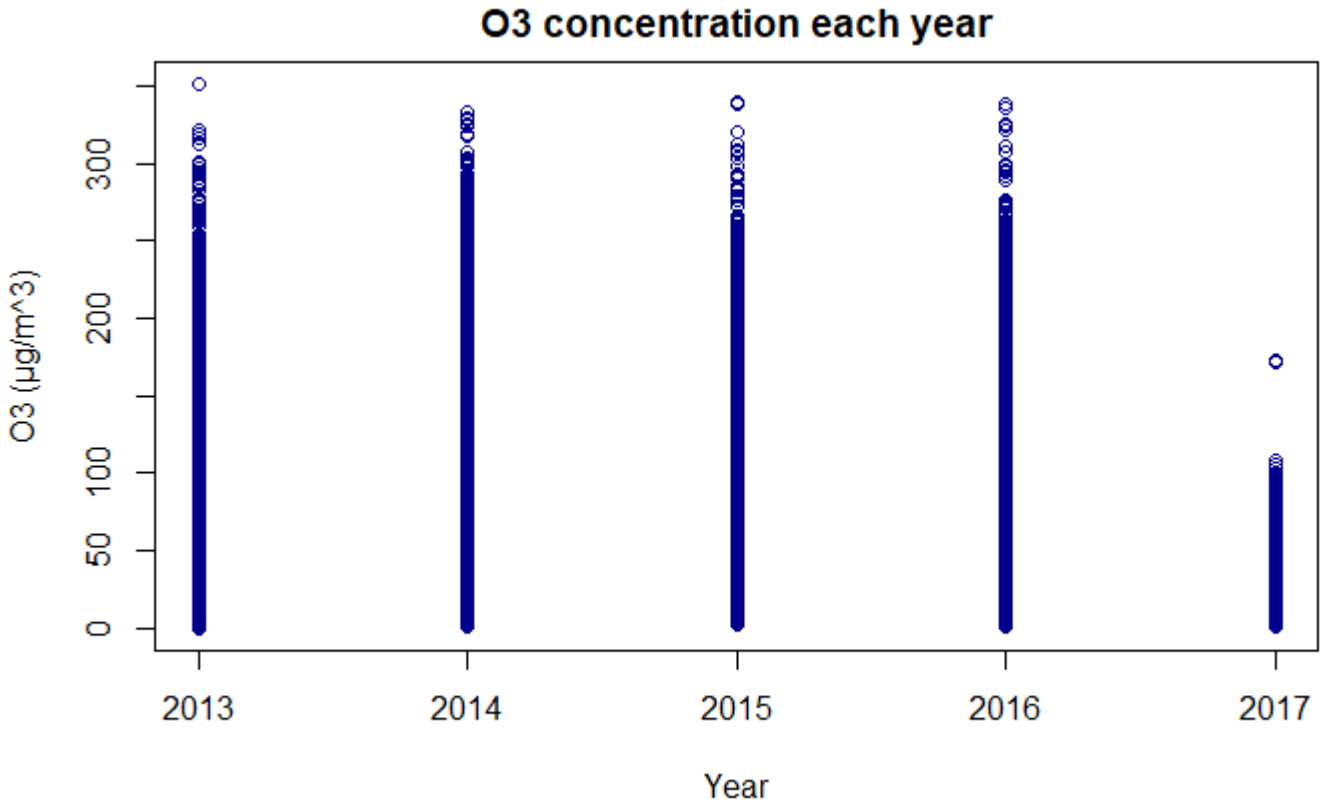
Hide

```
plot(data$hour, data$O3, col="blue4", xlab="Hours",ylab="O3 (µg/m^3)",main="O3 concentrations over the Hours")
```



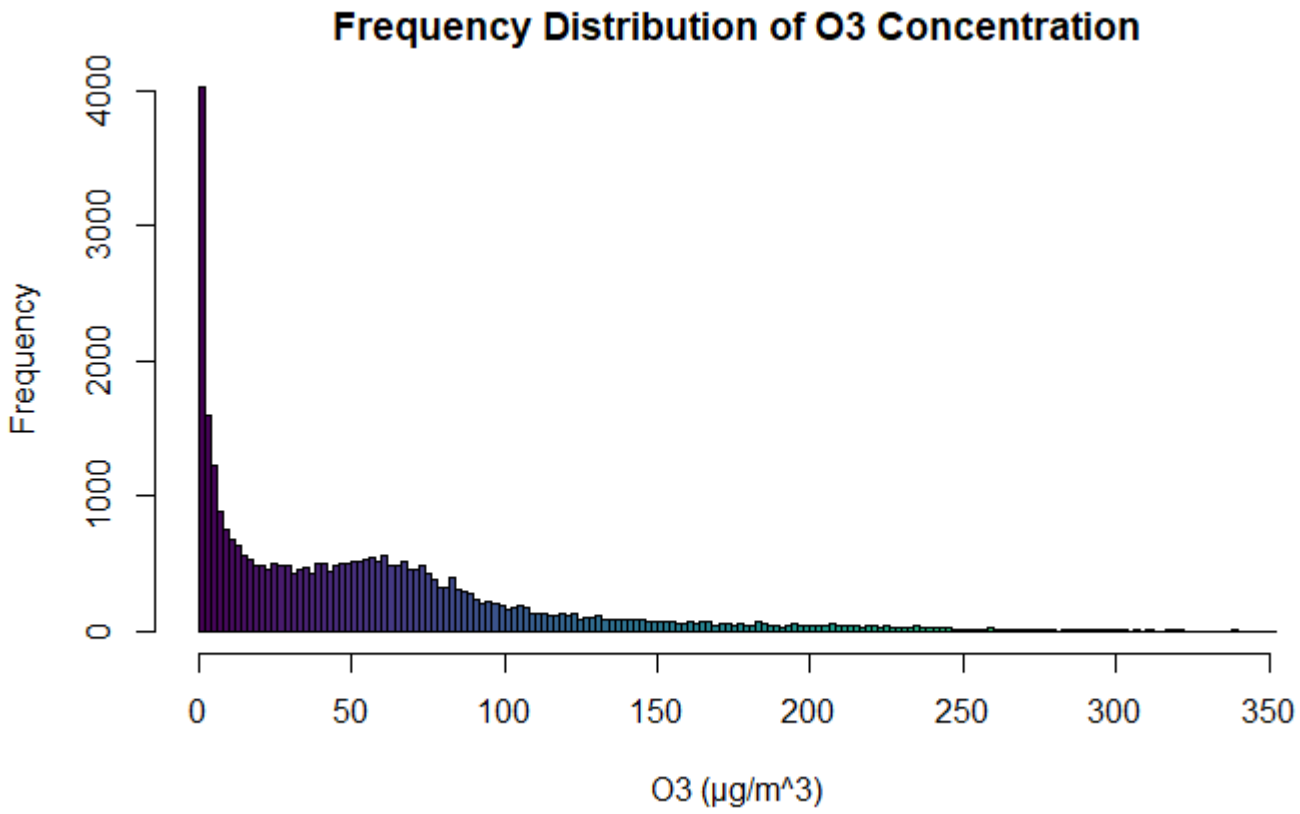
Hide

```
plot(data$year, data$O3, col="blue4", xlab="Year", ylab="O3 (µg/m^3)", main="O3 concentration each year")
```



Hide

```
#frequency of values
hist(data$O3, breaks = sqrt(nrow(data)), xlab="O3 (µg/m^3)", main = "Frequency Distribution of O3 Concentration" , col=viridis(sqrt(nrow(data))))
```



##Temperature

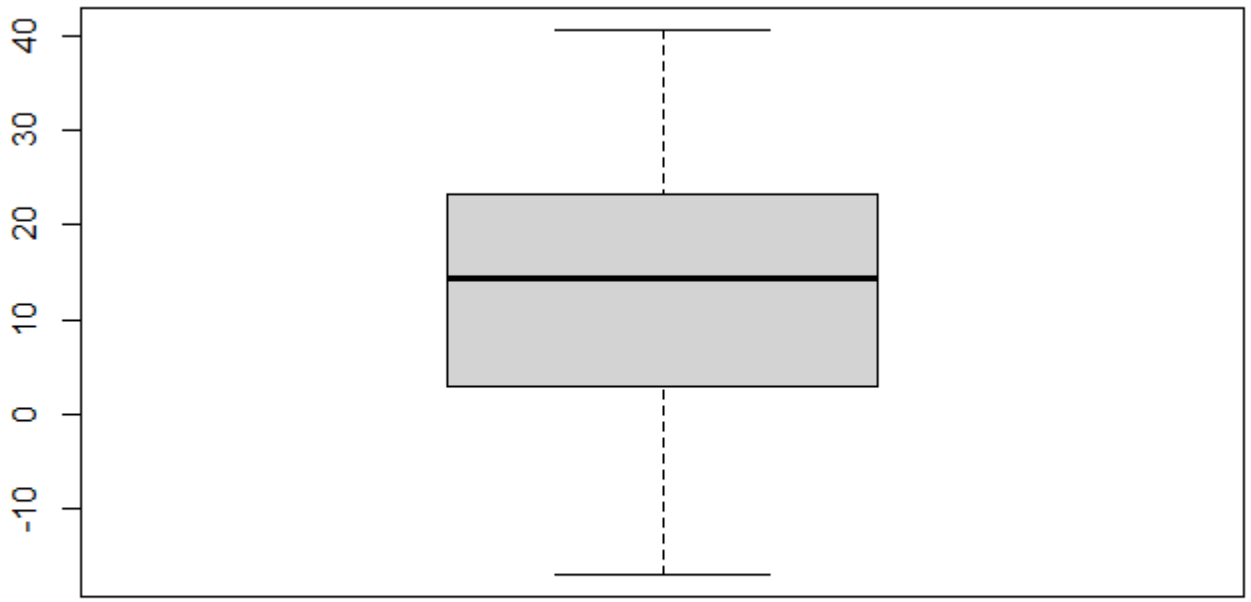
Hide

```
#summery statistics
summary(data$TEMP)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
 -16.80    3.00   14.40   13.39   23.20   40.60      51
```

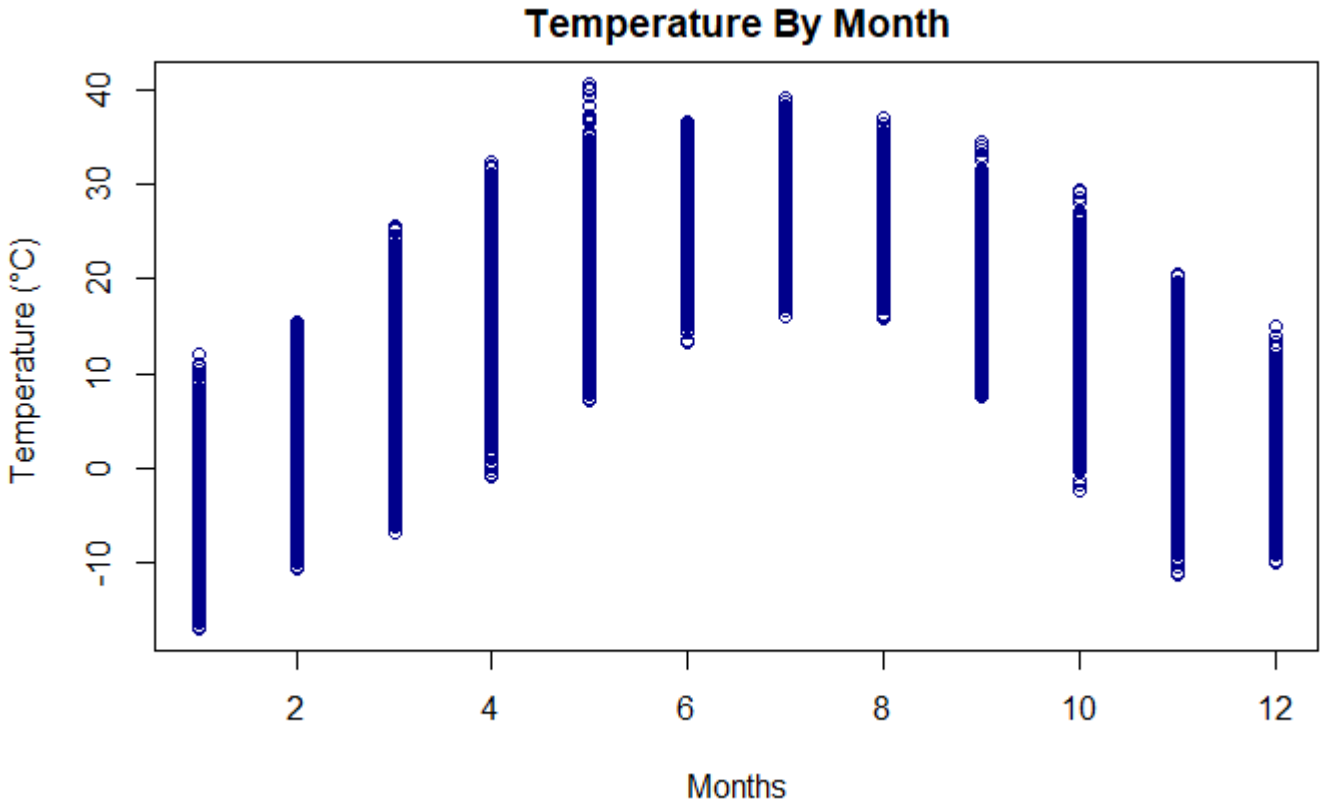
Hide

```
#outlines
boxplot(data$TEMP)
```



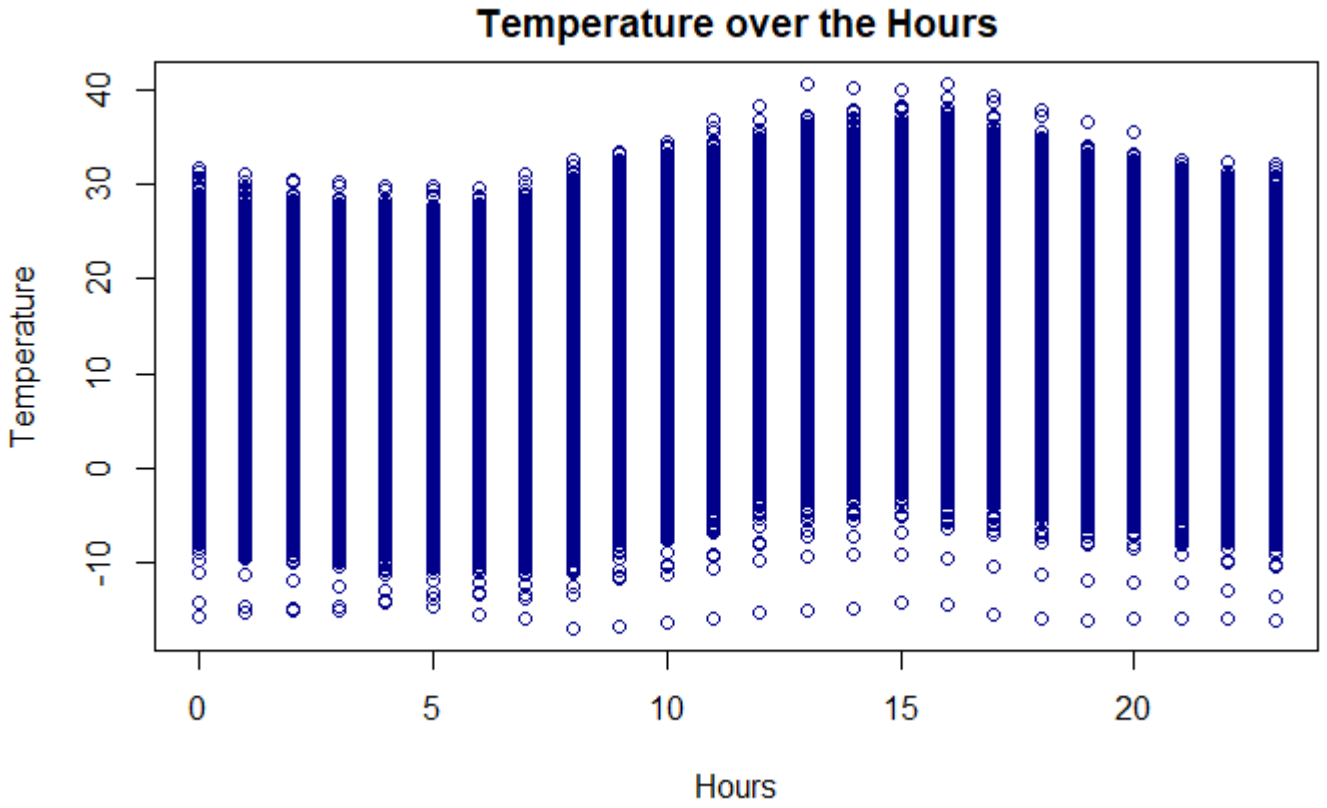
Hide

```
#plot values over time variables
plot(data$month, data$TEMP, col="blue4", xlab="Months",ylab="Temperature (°C) ",main="Temperature By Month")
```



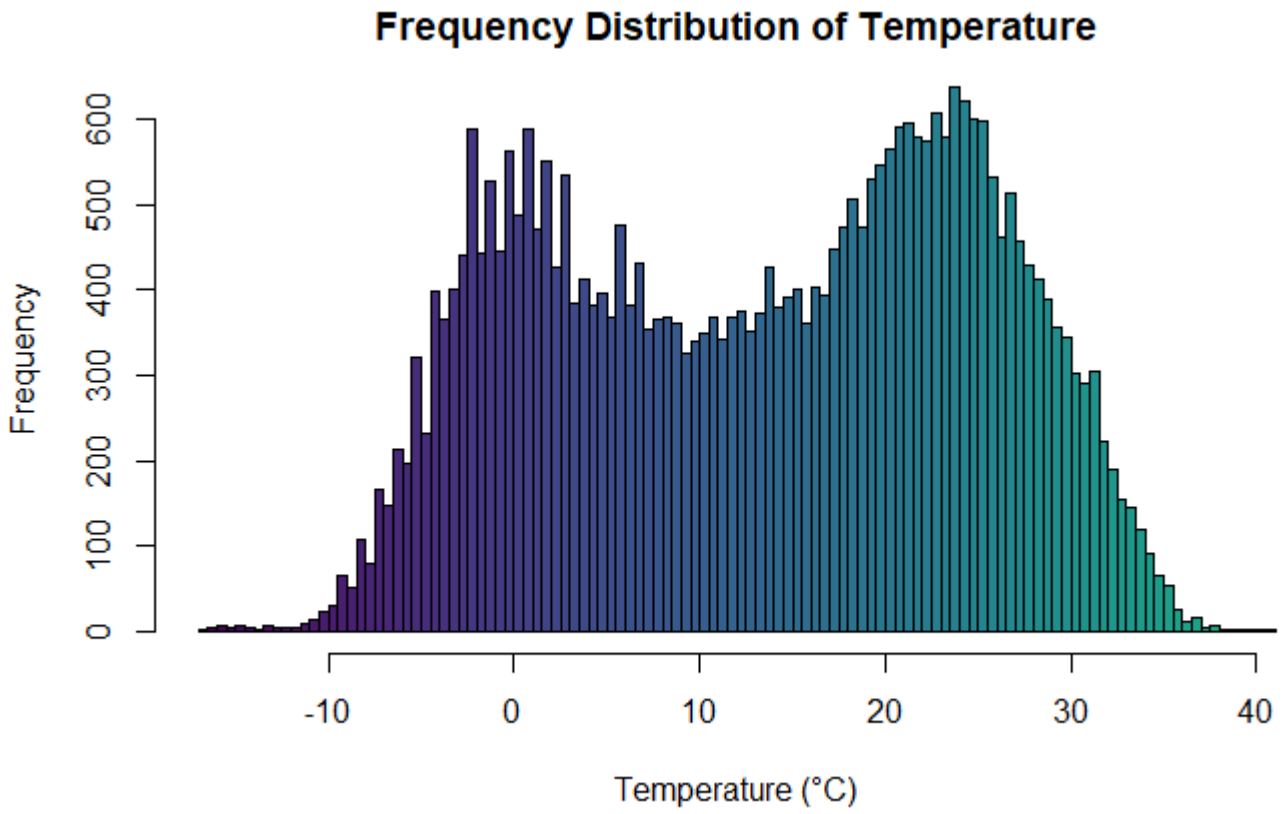
Hide

```
plot(data$hour, data$TEMP, col="blue4", xlab="Hours",ylab="Temperature",main="Temperature over the Hours")
```



Hide

```
#distribution of values
hist(data$TEMP, breaks = sqrt(nrow(data)), xlab="Temperature (°C)", main = "Frequency Distribution of Temperature" , col=vir
idis(sqrt(nrow(data))))
```



##Pressure

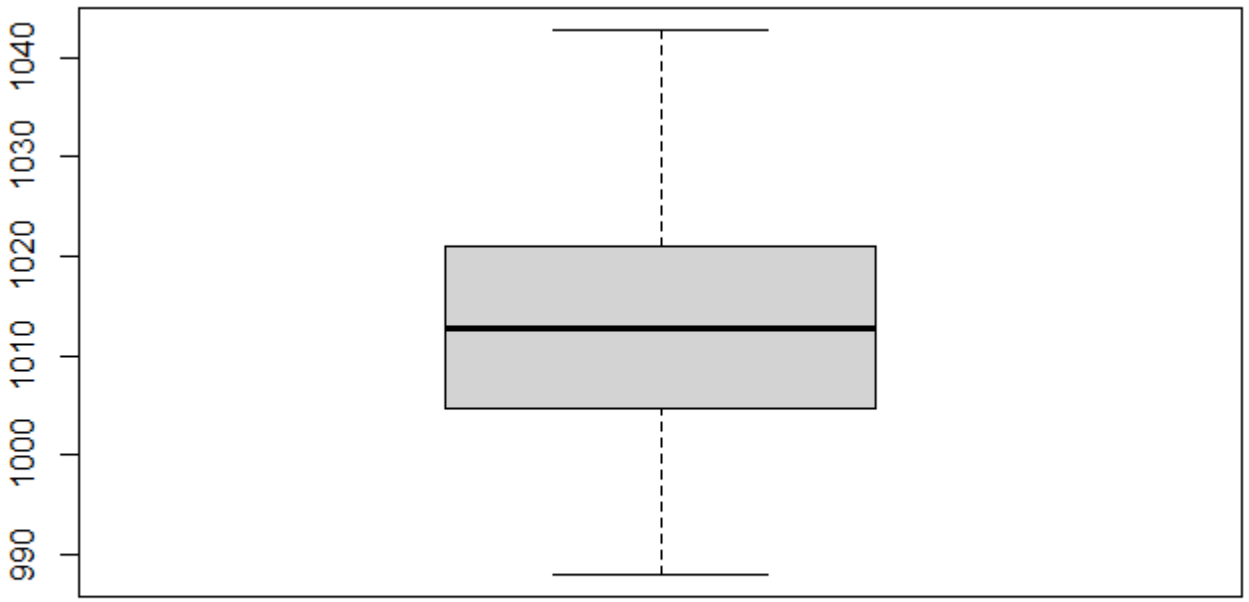
Hide

```
#pressure
#Summary statistics
summary(data$PRES)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
    988    1005    1013    1013    1021    1043      51
```

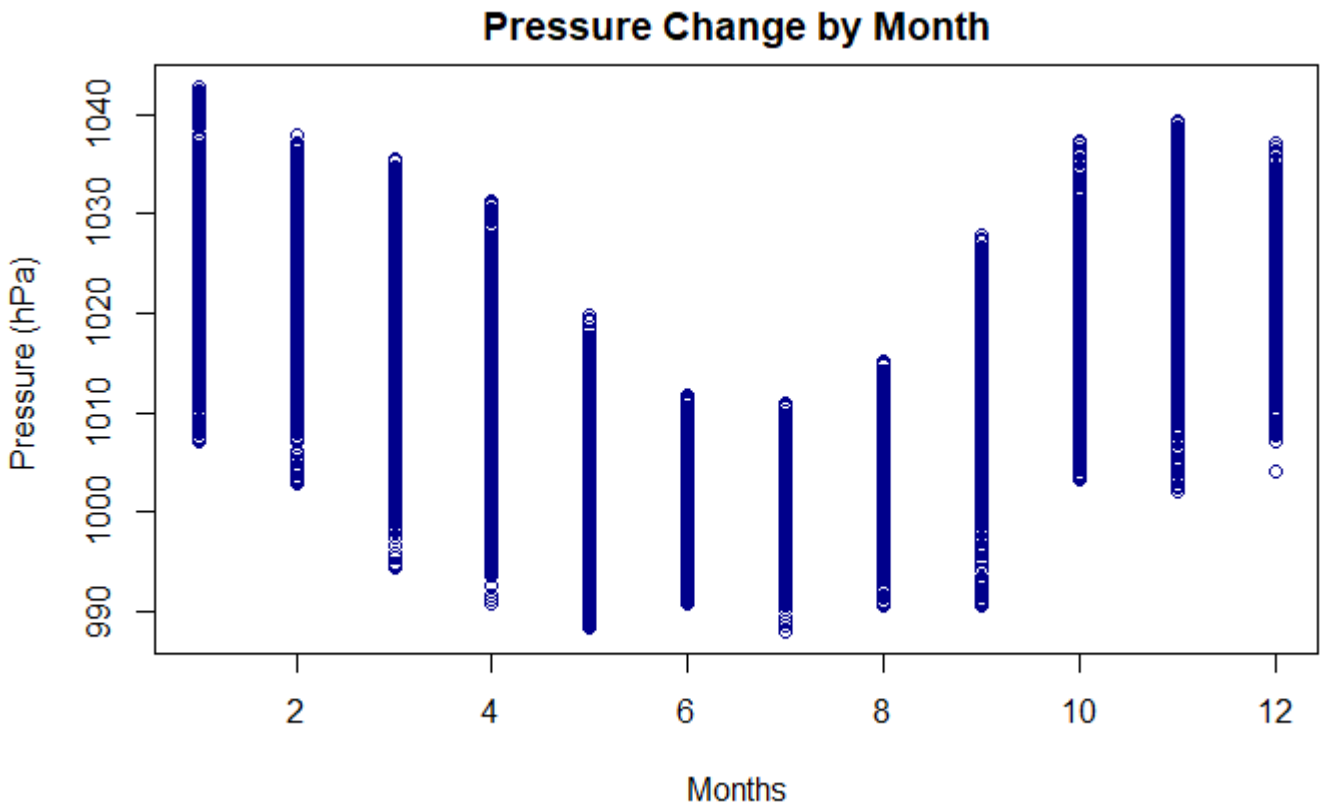
Hide

```
#outliers
boxplot(data$PRES)
```



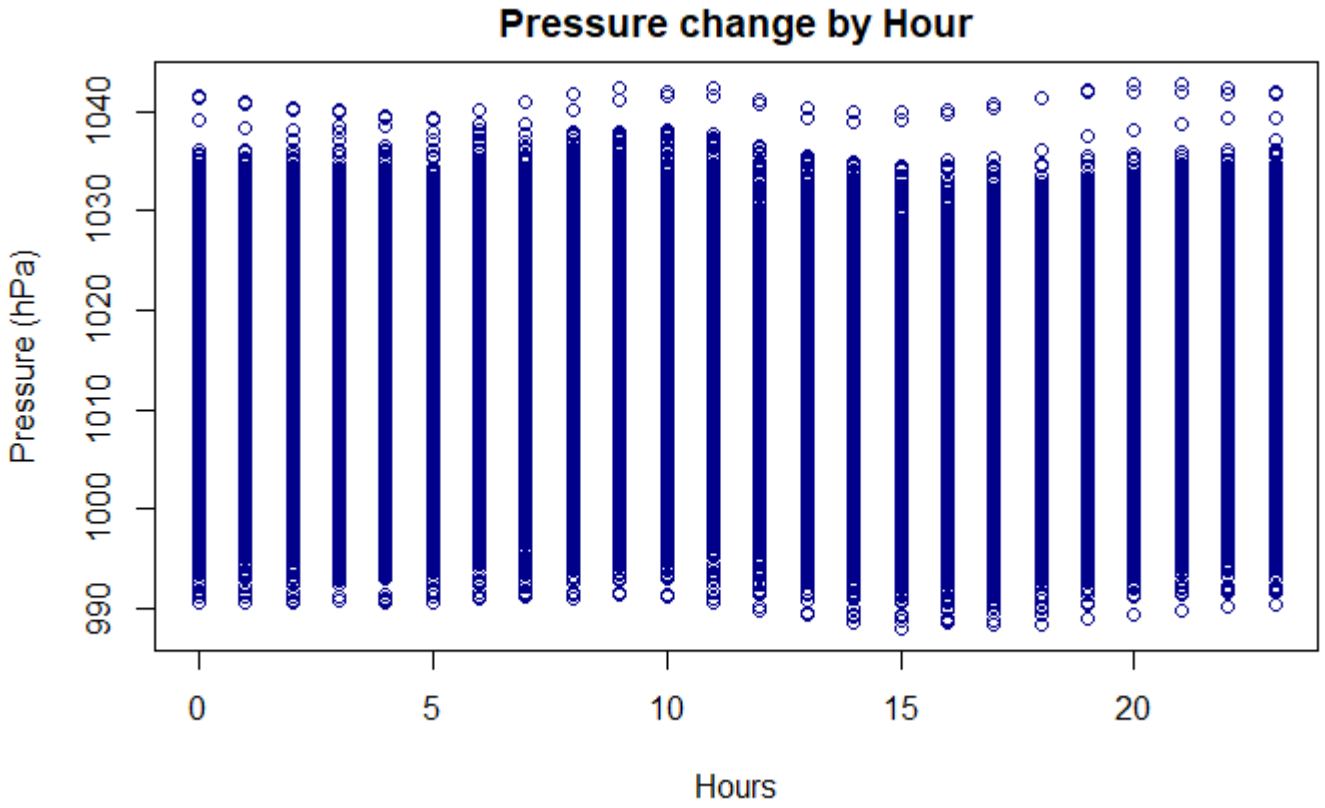
Hide

```
#plot values over time variables
plot(data$month, data$PRES, col="blue4", xlab="Months",ylab="Pressure (hPa) ",main="Pressure Change by Month")
```



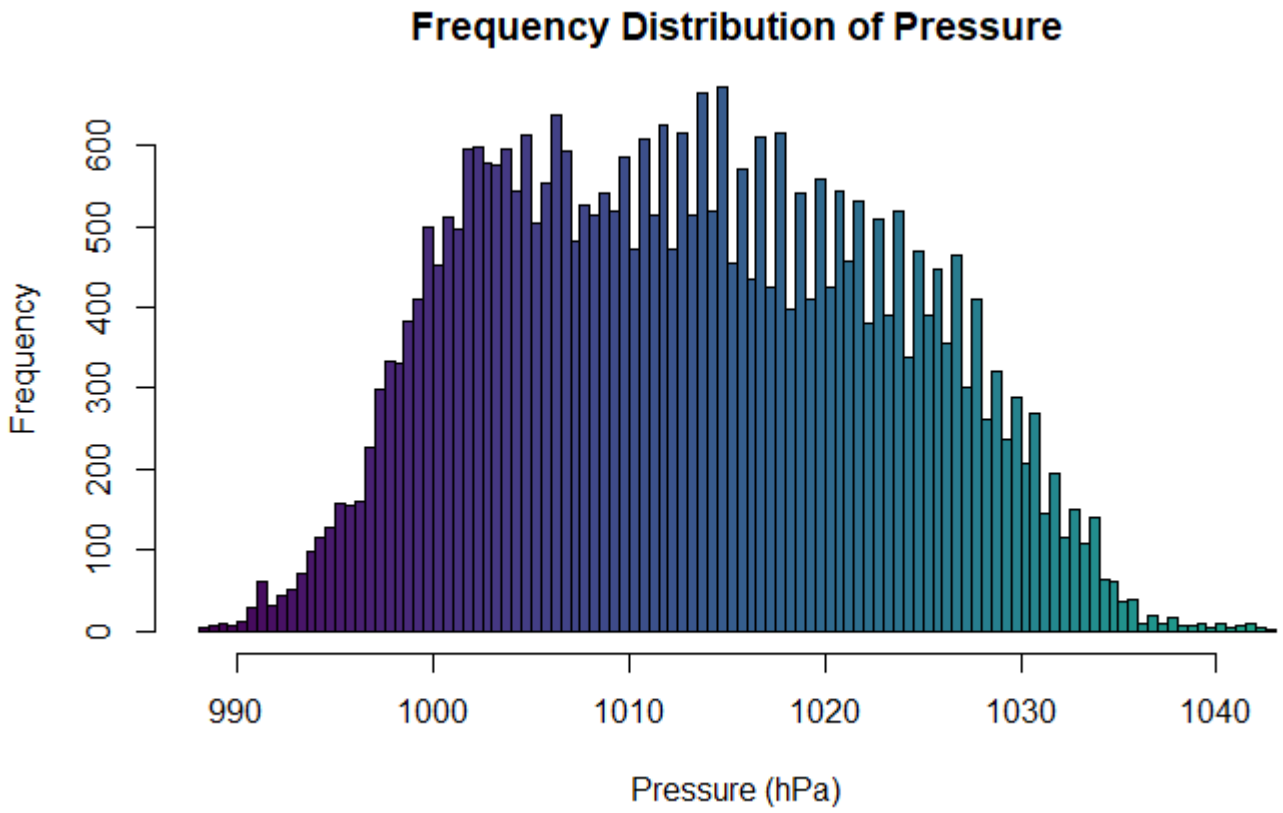
Hide

```
plot(data$hour, data$PRES, col="blue4", xlab="Hours",ylab="Pressure (hPa)",main="Pressure change by Hour")  #not much change in hour by hour data
```



Hide

```
#distribution of values
hist(data$PRES, breaks = sqrt(nrow(data)), xlab="Pressure (hPa)", main = "Frequency Distribution of Pressure" , col=viridis
(sqrt(nrow(data))))
```



##Dew Point

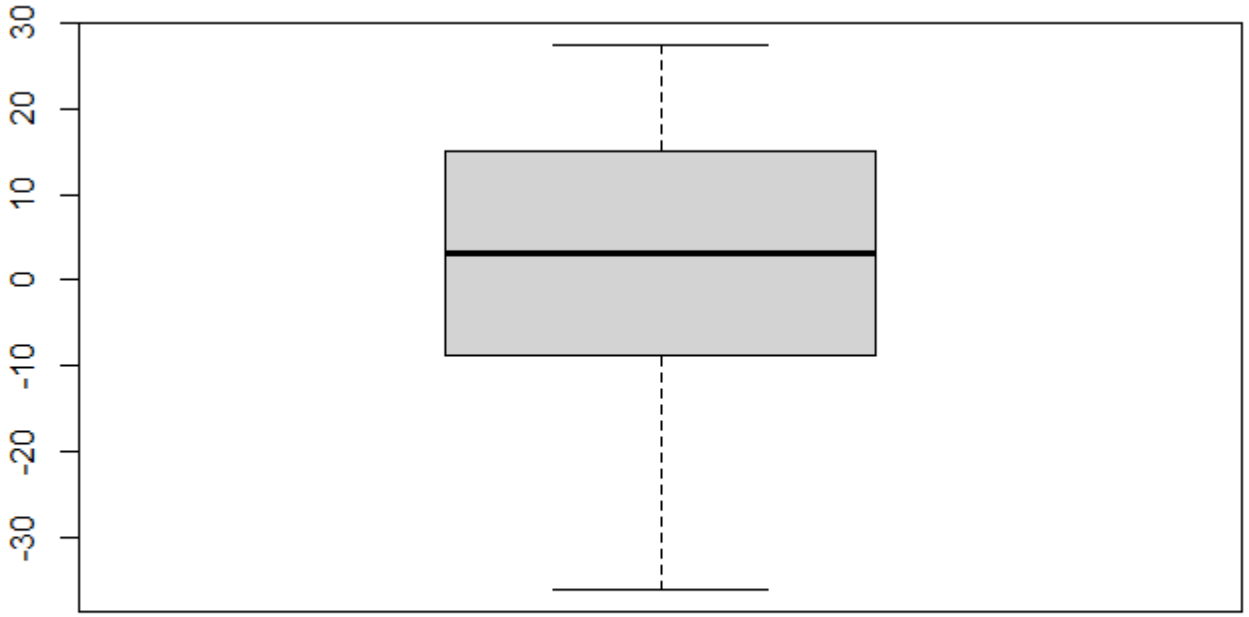
Hide

```
#summary stats
summary(data$DEWP)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
-36.000  -8.800   3.100   2.465  15.100  27.500      54
```

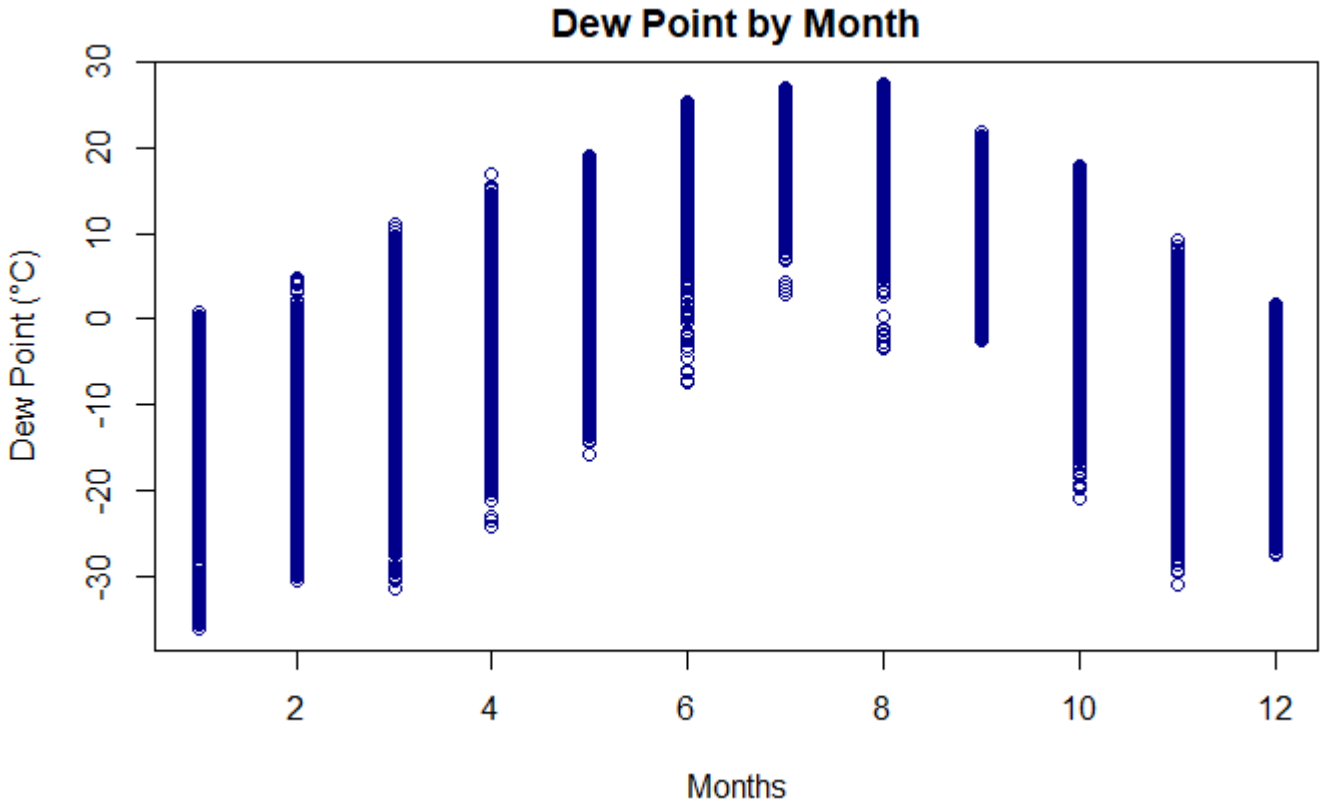
Hide

```
#outliers
boxplot(data$DEWP)
```



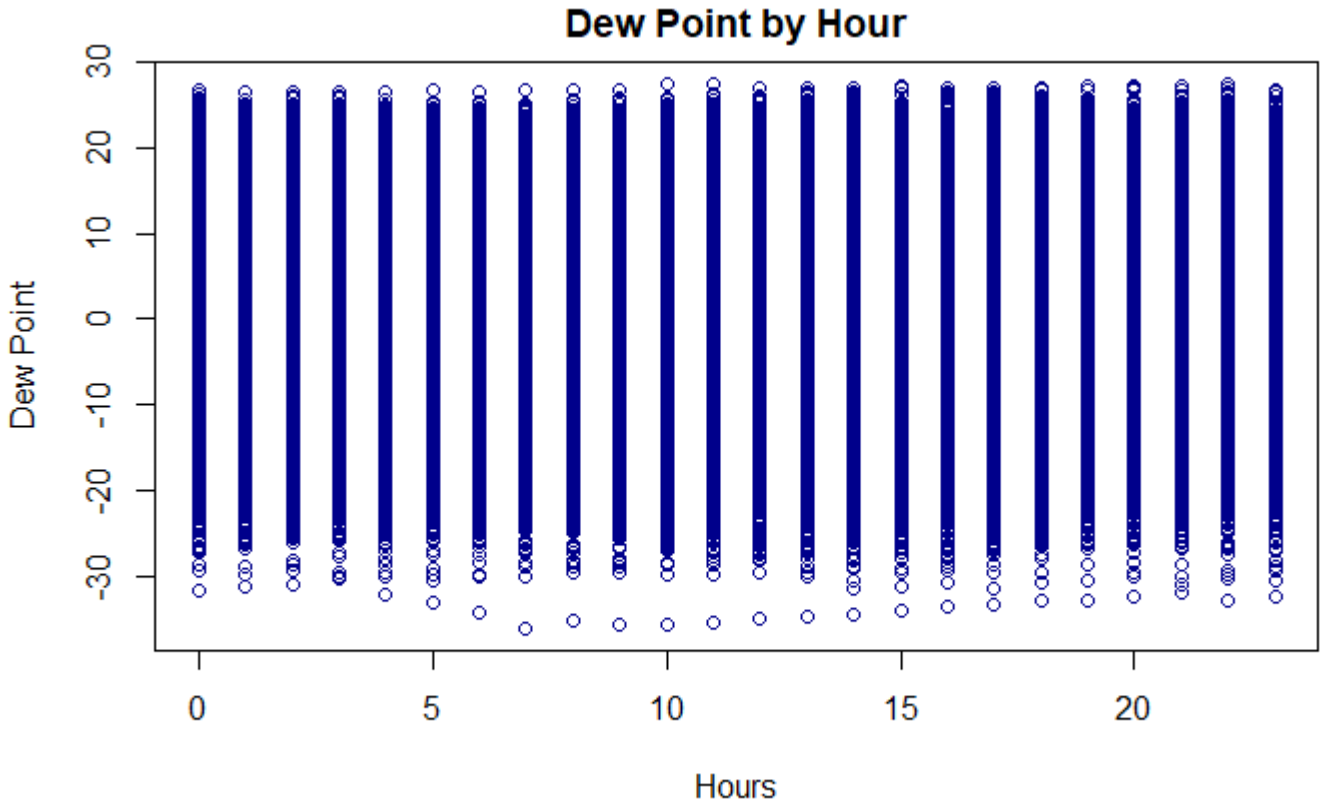
Hide

```
#plot values over time variables
plot(data$month, data$DEWP, col="blue4", xlab="Months",ylab="Dew Point (°C)",main="Dew Point by Month")
```



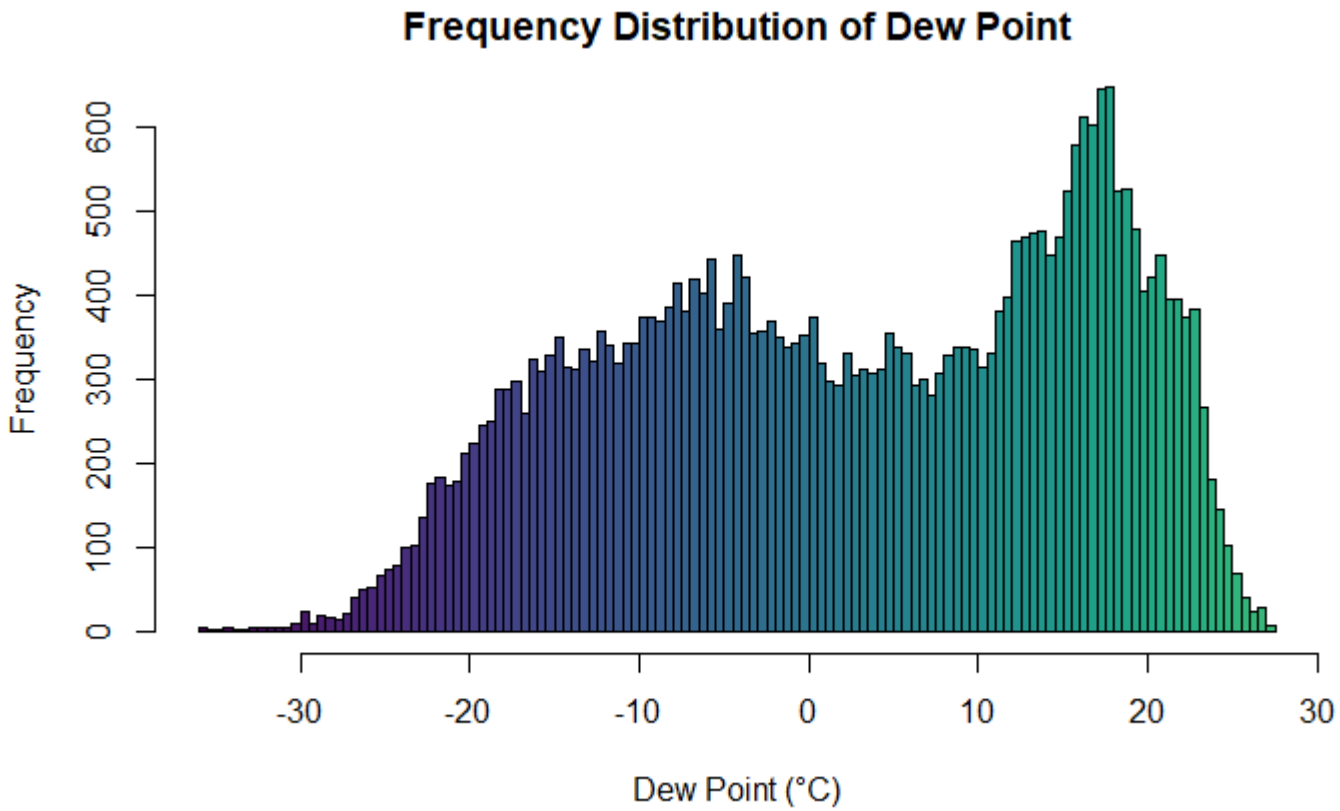
Hide

```
plot(data$hour, data$DEWP, col="blue4", xlab="Hours",ylab="Dew Point",main="Dew Point by Hour")  #not really any changes
```



Hide

```
#distribution of values
hist(data$DEWP, breaks = sqrt(nrow(data)), xlab="Dew Point (°C)", main = "Frequency Distribution of Dew Point" , col=viridis
(sqrt(nrow(data))))
```



##Rain

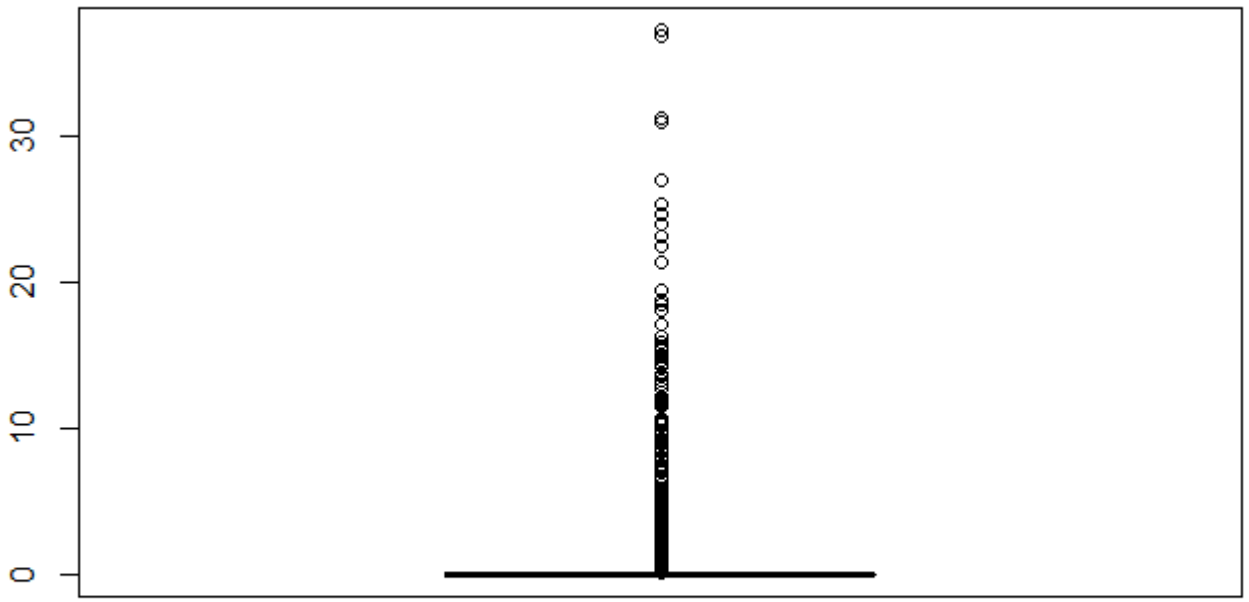
Hide

```
#summary statistics
summary(data$RAIN)
```

```
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.     NA's
 0.00000  0.00000  0.00000  0.06109  0.00000 37.30000       51
```

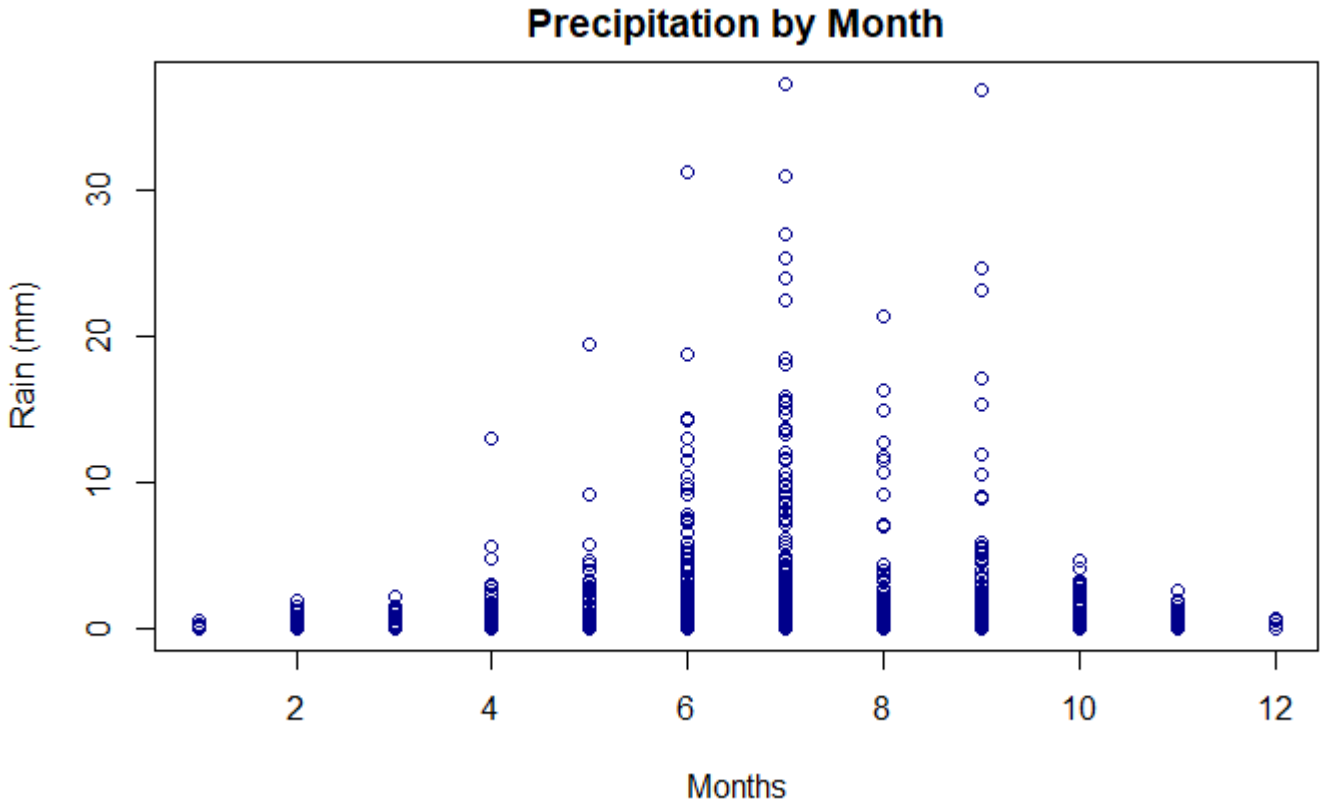
Hide

```
#outliers
boxplot(data$RAIN)
```



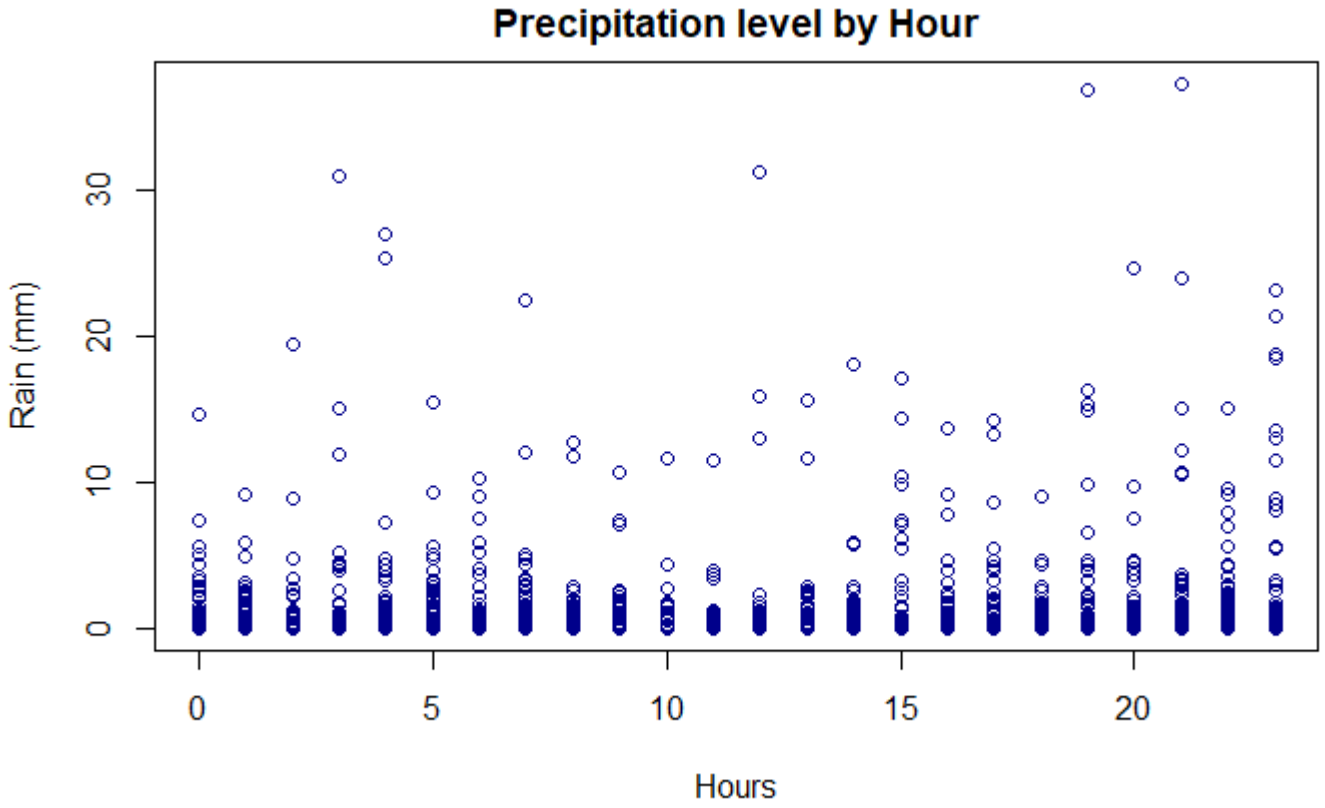
Hide

```
#plot values over time variables
plot(data$month, data$RAIN, col="blue4", xlab="Months",ylab="Rain (mm)",main="Precipitation by Month")
```



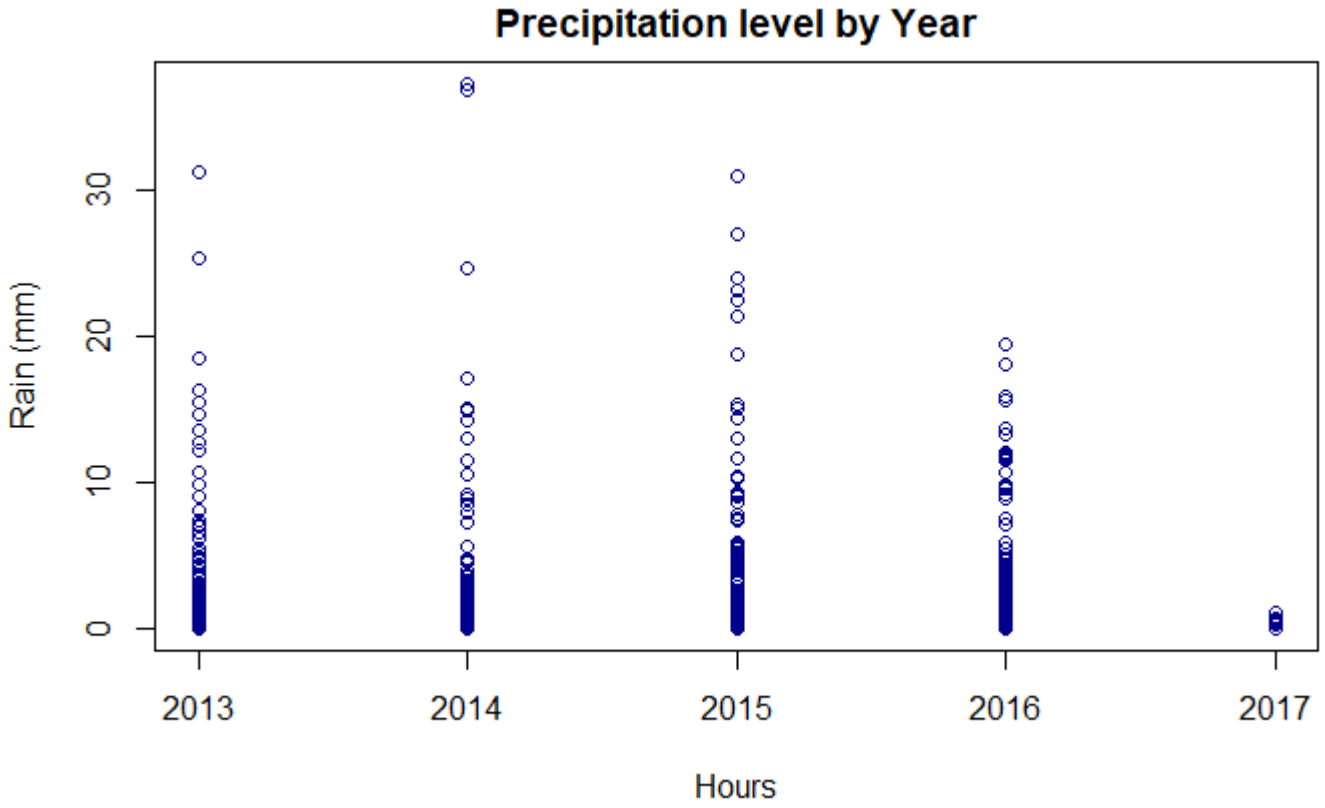
Hide

```
plot(data$hour, data$RAIN, col="blue4", xlab="Hours",ylab="Rain (mm)",main="Precipitation level by Hour")
```



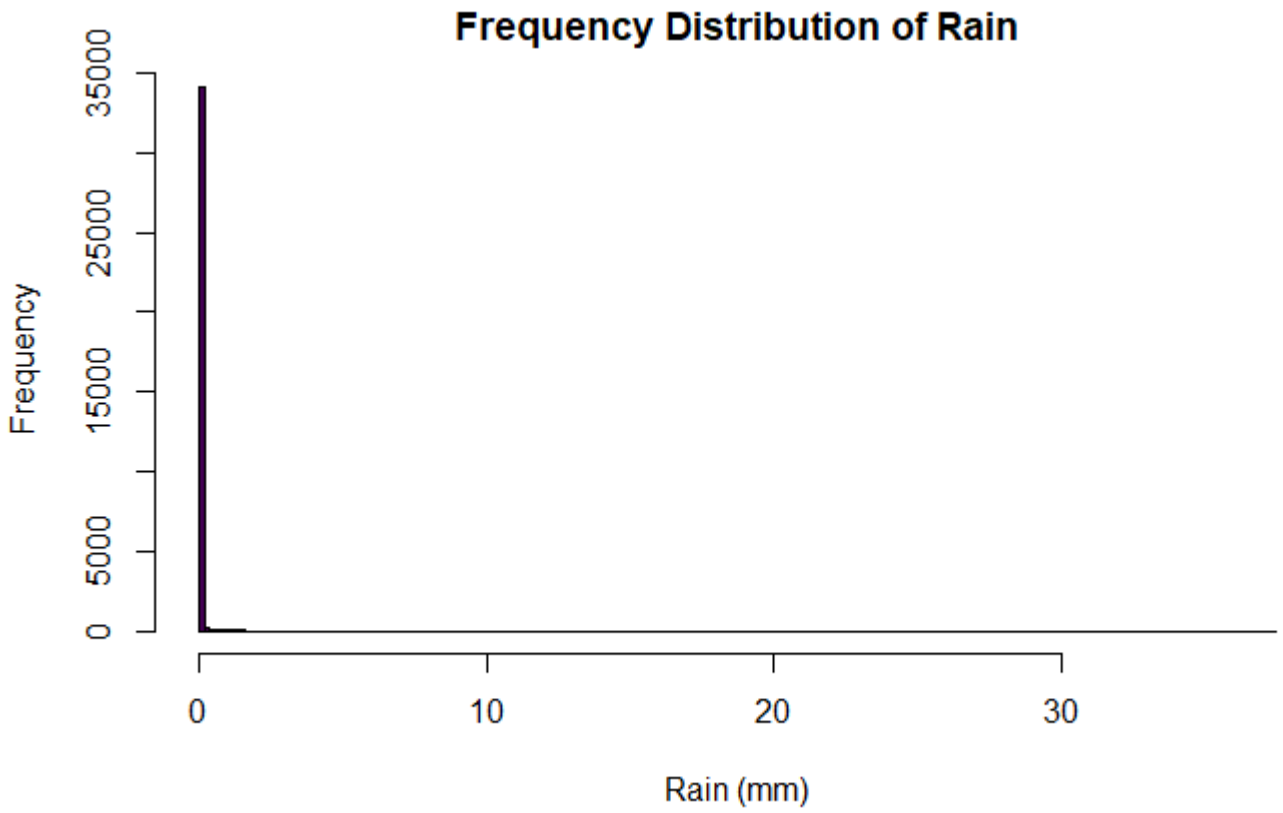
Hide

```
plot(data$year, data$RAIN, col="blue4", xlab="Hours",ylab="Rain (mm)",main="Precipitation level by Year")
```



Hide

```
#distribution
hist(data$RAIN, breaks = sqrt(nrow(data)), xlab="Rain (mm)", main = "Frequency Distribution of Rain" , col=viridis(sqrt(nrow
(data))))
```



##Wind Direction

Hide

```
summary(data$wd)
```

```
   Length     Class      Mode
    35064 character character
```

Hide

```
#already have package installed
#load package
library('dplyr')
```

```
Attaching package: ‘dplyr’

The following objects are masked from ‘package:stats’:

    filter, lag

The following objects are masked from ‘package:base’:

    intersect, setdiff, setequal, union
```

Hide

```
#count the number of values per wind direction
count_vals <- count(data, vars= wd)
names(count_vals) <- c("Wind Direction", "Total Number")

count_vals
```

Wind Direction	Total Number
<chr>	<int>
E	1153
ENE	1416
ESE	1377
N	3877
NE	3351
NNE	4540
NNW	1656
NW	2446
S	2613
SE	2665

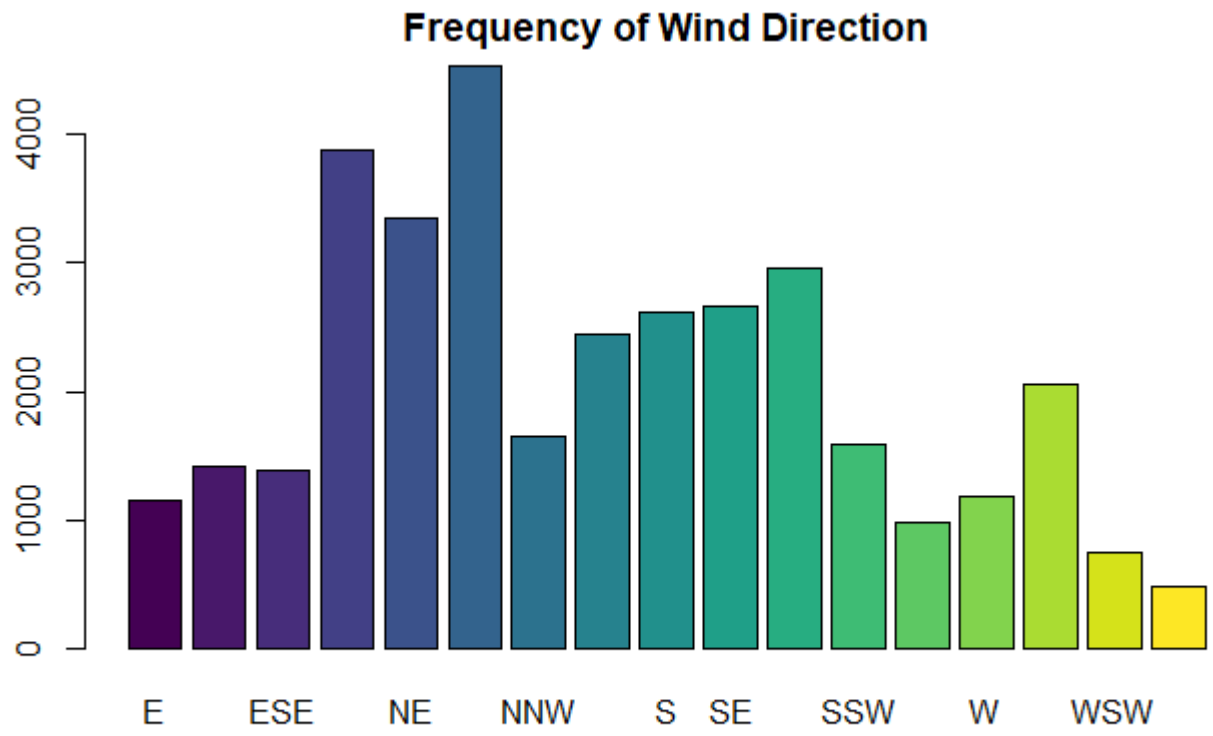
1-10 of 17 rows Previous **1** 2 Next

Hide

```
directions <- c("E","ENE","ESE","N","NE", "NNE","NNW","NW","S","SE", "SSE","SSW","SW","W","WNW","WSW","NA")

colours <- c("")

barplot(count_vals$`Total Number`, names.arg = directions , col = viridis(17), main = "Frequency of Wind Direction")
```



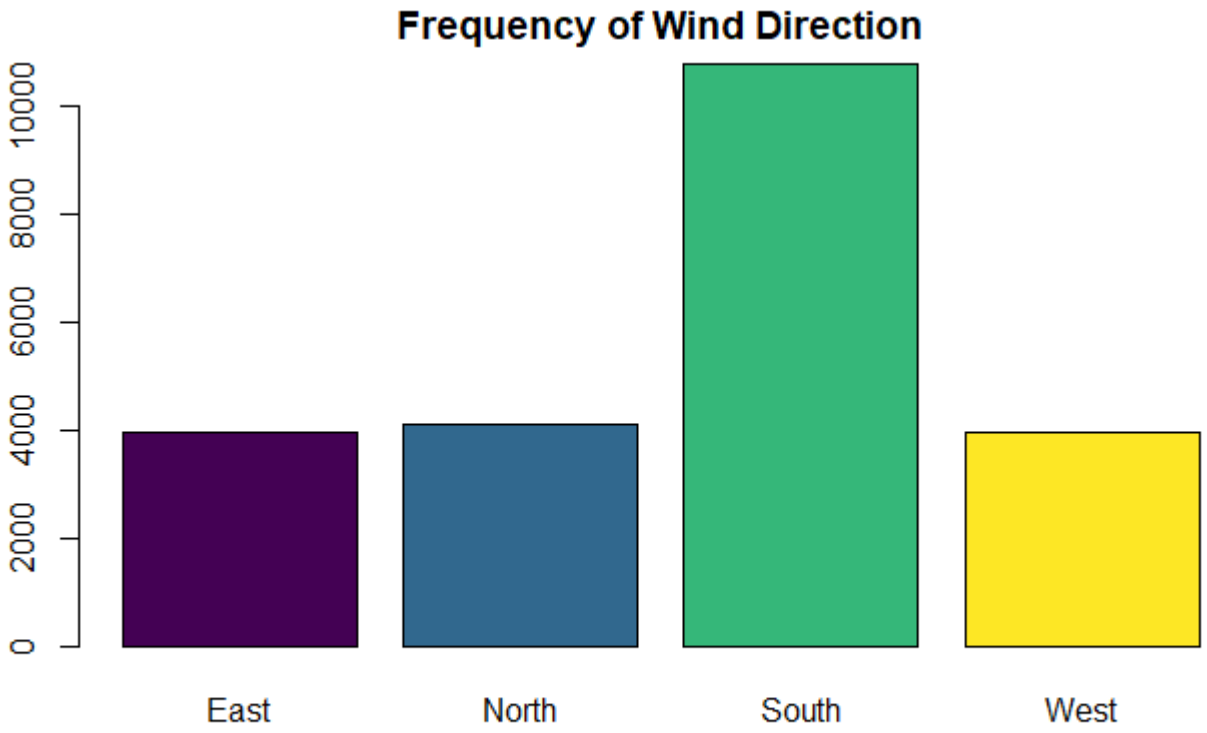
Hide

```
#group values into 4 major directions

east <- sum(count_vals[1:3,2])
north <- sum(count_vals[7:8,2])
south <- sum(count_vals[9:13,2])
west <- sum(count_vals[14:16,2])

NESW <- c(east, north, south, west)

barplot(NESW, names = c("East","North","South","West"), col = viridis(4), main = "Frequency of Wind Direction")
```



##Wind Speed

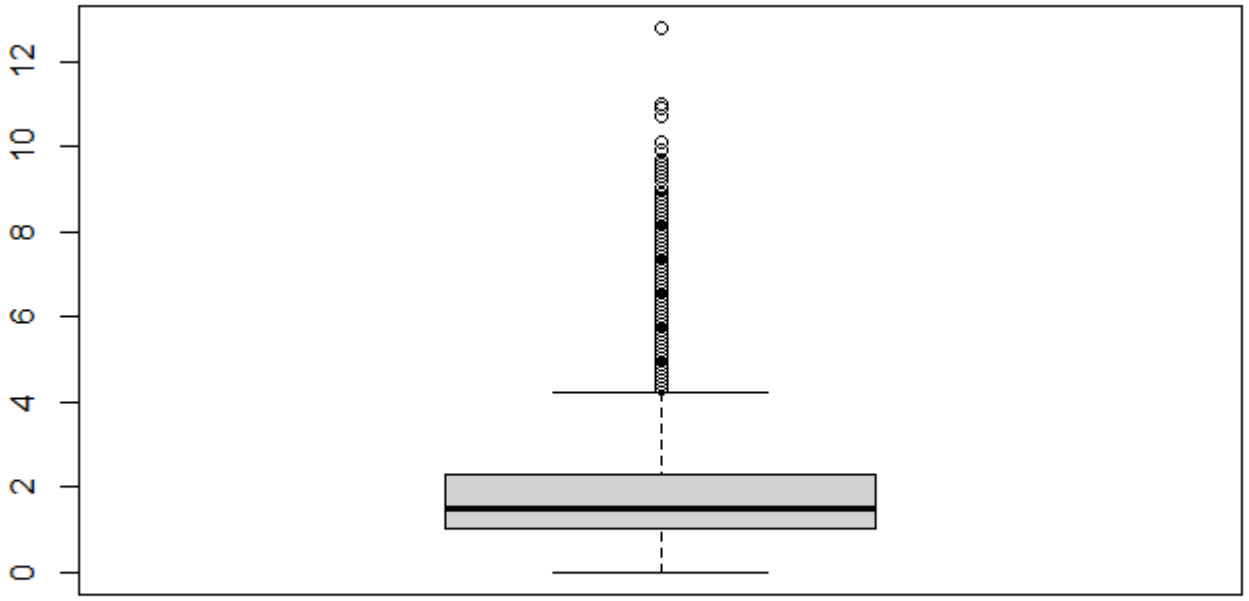
Hide

```
#summary statistics
summary(data$WSPM)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
  0.000   1.000   1.500   1.808   2.300  12.800      44
```

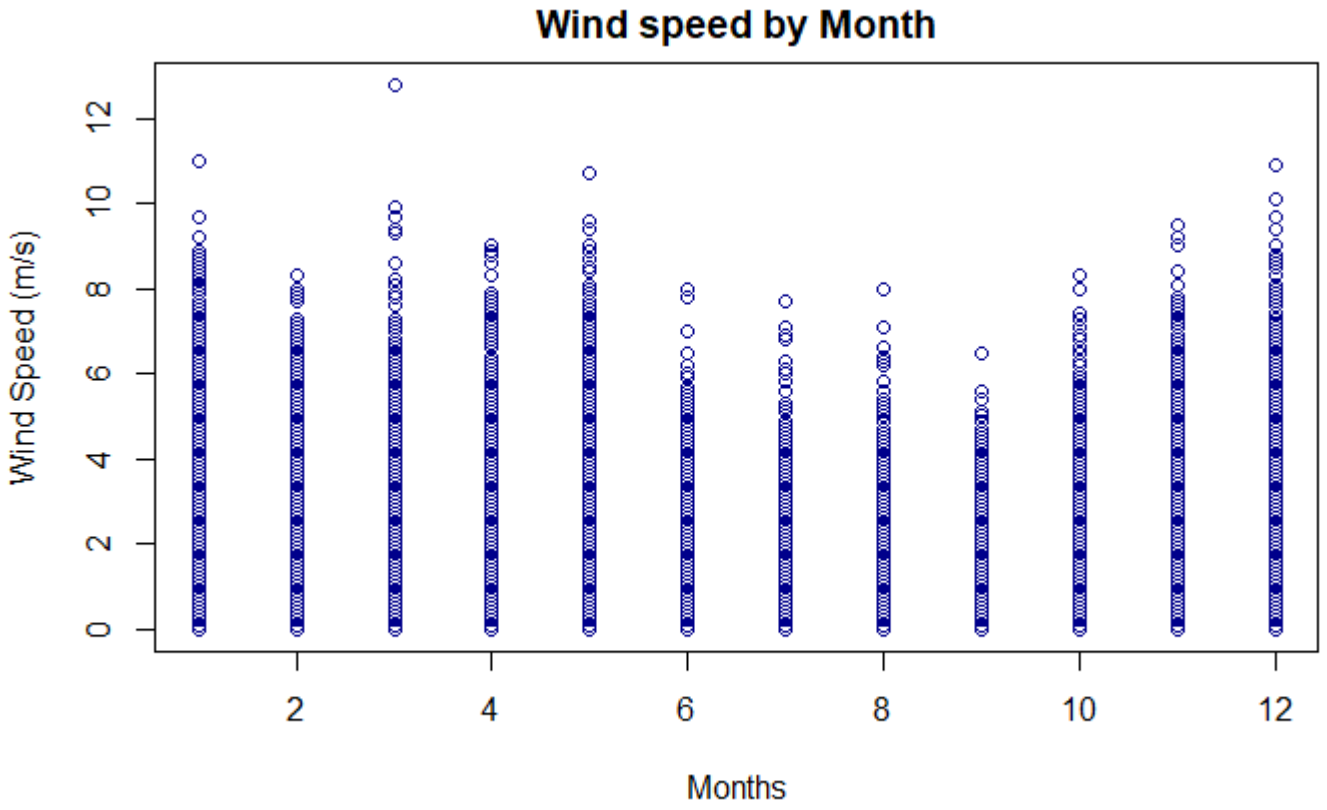
Hide

```
#boxplot for outliers
boxplot(data$WSPM)
```



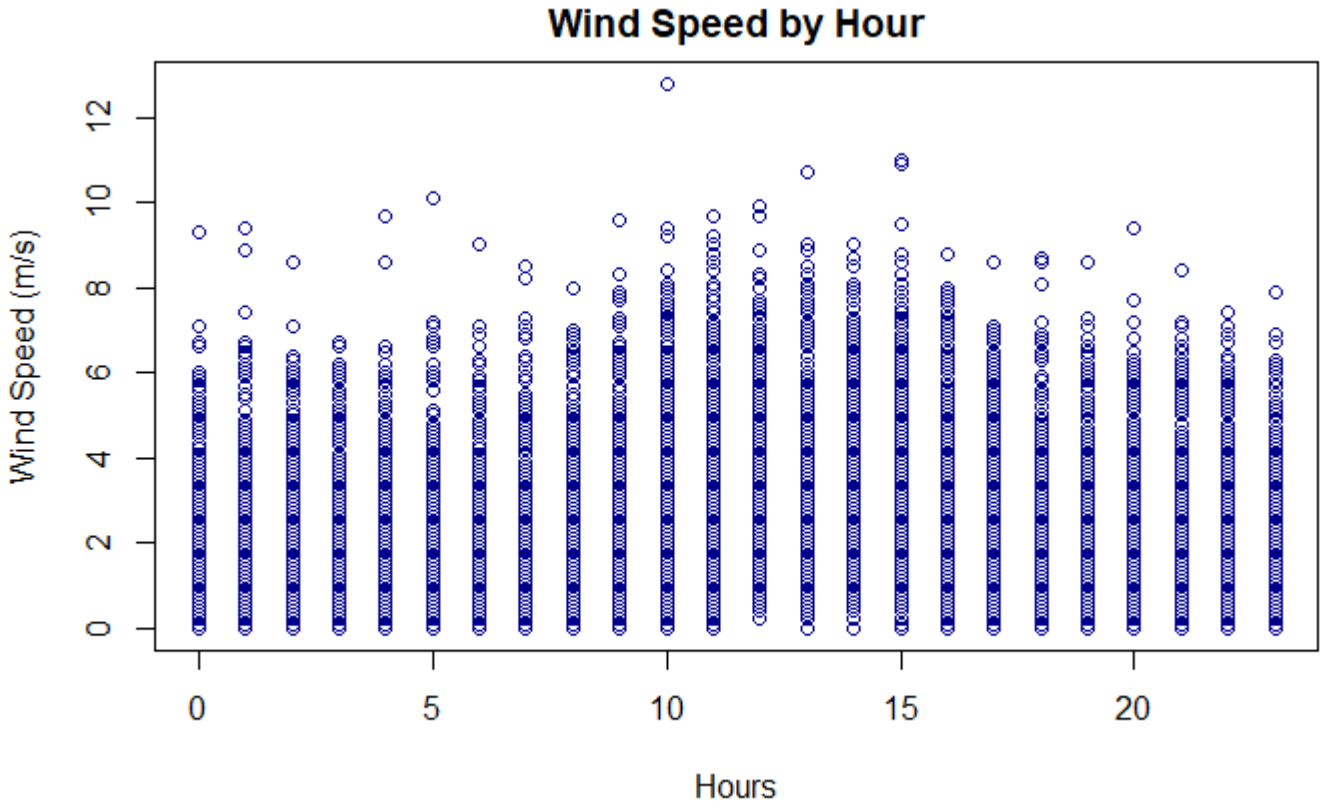
Hide

```
#plot values over time variables
plot(data$month, data$WSPM, col="blue4", xlab="Months",ylab="Wind Speed (m/s)",main="Wind speed by Month")
```



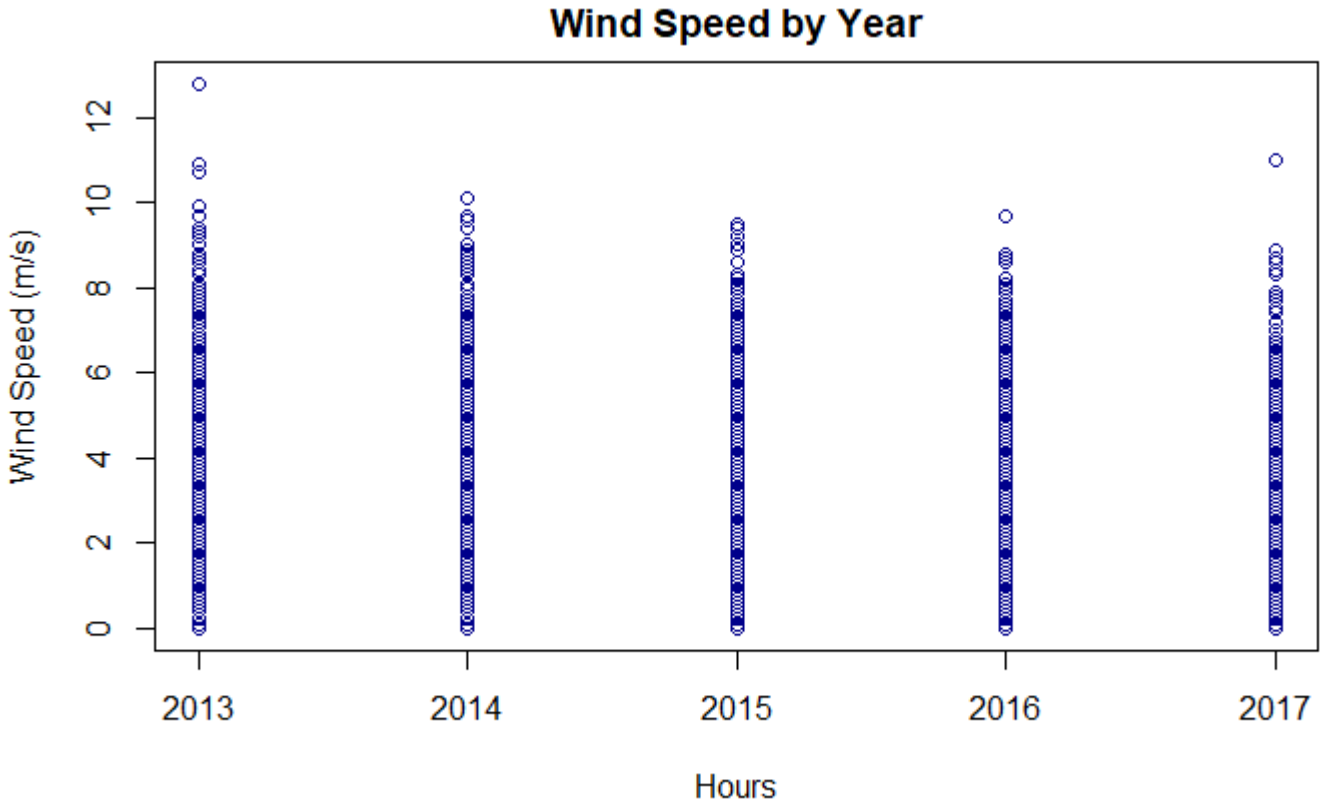
Hide

```
plot(data$hour, data$WSPM, col="blue4", xlab="Hours",ylab="Wind Speed (m/s)",main="Wind Speed by Hour")
```



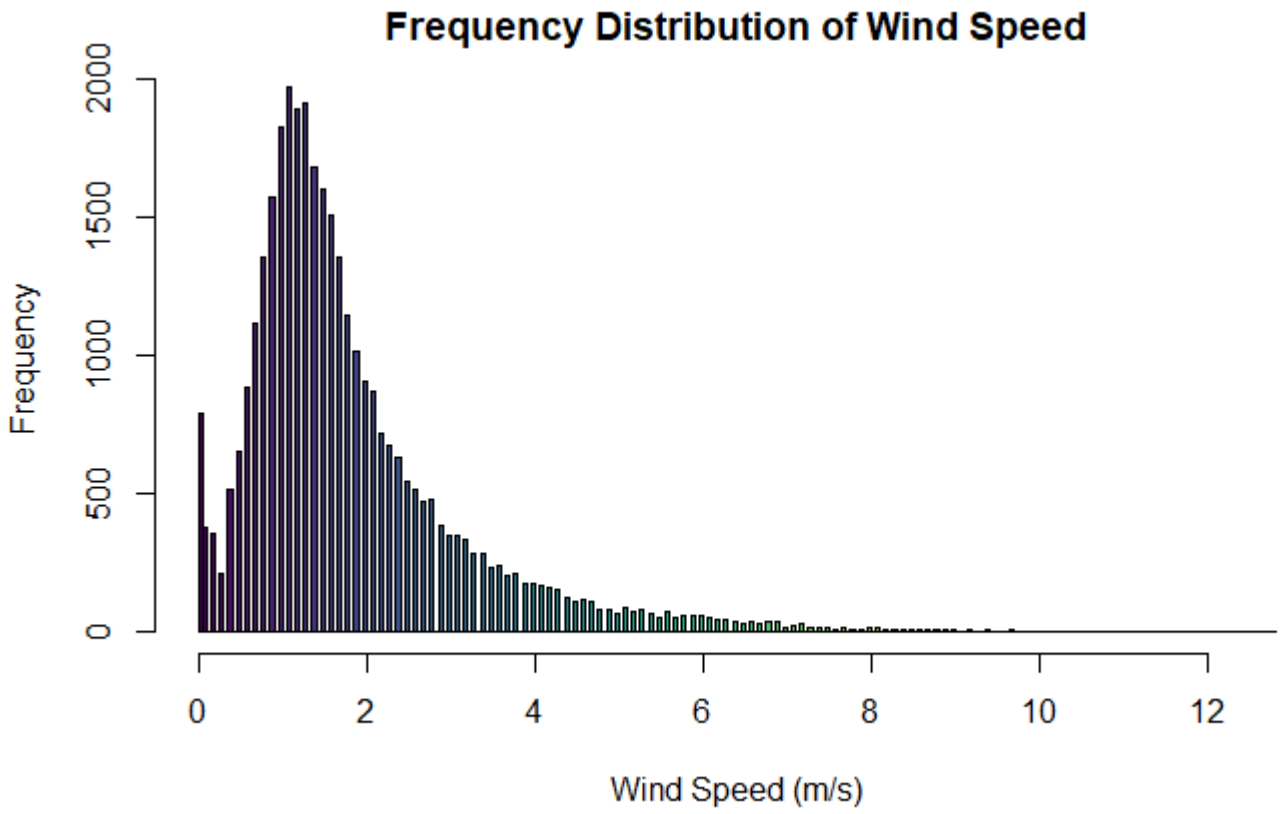
Hide

```
plot(data$year, data$WSPM, col="blue4", xlab="Hours",ylab="Wind Speed (m/s)",main="Wind Speed by Year")
```



Hide

```
#check distribution
hist(data$WSPM, breaks = sqrt(nrow(data)), xlab = "Wind Speed (m/s)", main = "Frequency Distribution of Wind Speed", col=vir
idis(sqrt(nrow(data))))
```



Hide

```
#install.packages("psych")
library(psych)
```

```
Warning: package ‘psych’ was built under R version 4.1.3
```

Hide

```
#create a summary table of all values
describe(data, fast = TRUE)
```

```
Warning in FUN(newX[, i], ...) :
  no non-missing arguments to min; returning Inf
Warning in FUN(newX[, i], ...) :
  no non-missing arguments to min; returning Inf
Warning in FUN(newX[, i], ...) :
  no non-missing arguments to max; returning -Inf
Warning in FUN(newX[, i], ...) :
  no non-missing arguments to max; returning -Inf
```

	vars <dbl>	n <dbl>	mean <dbl>	sd <dbl>	min <dbl>	max <dbl>	range <dbl>	se <dbl>
No	1	35064	17532.50	10122.25	1.00	35064.00	35063.00	54.06
year	2	35064	2014.66	1.18	2013.00	2017.00	4.00	0.01
month	3	35064	6.52	3.45	1.00	12.00	11.00	0.02
day	4	35064	15.73	8.80	1.00	31.00	30.00	0.05
hour	5	35064	11.50	6.92	0.00	23.00	23.00	0.04
PM2.5	6	34151	79.49	81.23	2.00	941.00	939.00	0.44
PM10	7	34516	98.74	89.14	2.00	999.00	997.00	0.48
SO2	8	33768	13.57	19.57	0.29	239.00	238.71	0.11
NO2	9	33699	43.91	31.00	2.00	258.00	256.00	0.17
CO	10	32886	1187.06	1156.37	100.00	10000.00	9900.00	6.38

1-10 of 18 rows Previous **1** 2 Next

Hide

```
#gives column number, number of valid cases, mean value,max and min values, the range and standard error
```

Hide

```
#create another summary table grouped by year
#only looking at rows 6 to 17, ignoring station and time values
describeBy(data[,6:17],group=data$year, fast = TRUE)
```

```
Warning in FUN(newX[, i], ...) :
  no non-missing arguments to min; returning Inf
Warning in FUN(newX[, i], ...) :
  no non-missing arguments to max; returning -Inf
Warning in FUN(newX[, i], ...) :
  no non-missing arguments to min; returning Inf
Warning in FUN(newX[, i], ...) :
  no non-missing arguments to max; returning -Inf
Warning in FUN(newX[, i], ...) :
  no non-missing arguments to min; returning Inf
Warning in FUN(newX[, i], ...) :
  no non-missing arguments to max; returning -Inf
Warning in FUN(newX[, i], ...) :
  no non-missing arguments to min; returning Inf
Warning in FUN(newX[, i], ...) :
  no non-missing arguments to max; returning -Inf
Warning in FUN(newX[, i], ...) :
  no non-missing arguments to min; returning Inf
Warning in FUN(newX[, i], ...) :
  no non-missing arguments to max; returning -Inf
```

```

 Descriptive statistics by group
group: 2013
```

	vars <dbl>	n <dbl>	mean <dbl>	sd <dbl>	min <dbl>	max <dbl>	range <dbl>	se <dbl>
PM2.5	1	7082	77.20	72.19	3.00	558.00	555.00	0.86
PM10	2	7279	94.63	77.18	2.00	764.00	762.00	0.90
SO2	3	6859	15.76	21.52	0.29	167.00	166.71	0.26

	vars <dbl>	n <dbl>	mean <dbl>	sd <dbl>	min <dbl>	max <dbl>	range <dbl>	se <dbl>
NO2	4	7281	41.62	28.72	2.00	204.00	202.00	0.34
CO	5	6220	1238.58	963.11	100.00	8700.00	8600.00	12.21
O3	6	7204	56.62	54.07	0.21	351.72	351.50	0.64
TEMP	7	7344	15.54	10.31	-9.90	37.80	47.70	0.12
PRES	8	7344	1010.49	9.63	988.00	1033.80	45.80	0.11
DEWP	9	7341	5.55	13.26	-27.50	27.40	54.90	0.15
RAIN	10	7344	0.06	0.77	0.00	31.20	31.20	0.01

1-10 of 12 rows Previous **1** 2 Next

```
------------------------------------------------------------------------------------------------------------
group: 2014
```

	vars <dbl>	n <dbl>	mean <dbl>	sd <dbl>	min <dbl>	max <dbl>	range <dbl>	se <dbl>
PM2.5	1	8569	84.58	79.43	3.0	592.0	589.0	0.86
PM10	2	8659	110.94	90.06	3.0	903.0	900.0	0.97
SO2	3	8595	17.45	24.87	1.0	207.0	206.0	0.27
NO2	4	8105	45.66	30.46	2.0	204.0	202.0	0.34
CO	5	8484	1166.18	993.71	100.0	7600.0	7500.0	10.79
O3	6	8557	58.42	59.20	1.0	333.0	332.0	0.64
TEMP	7	8760	13.84	11.52	-10.5	40.6	51.1	0.12
PRES	8	8760	1012.51	9.23	989.5	1035.9	46.4	0.10
DEWP	9	8760	2.44	13.30	-30.9	24.9	55.8	0.14

	vars	n	mean	sd	min	max	range	se
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
RAIN	10	8760	0.05	0.79	0.0	37.3	37.3	0.01

1-10 of 12 rows Previous **1** 2 Next

```
------------------------------------------------------------------------------------------------------------
group: 2015
```

	vars	n	mean	sd	min	max	range	se
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
PM2.5	1	8636	81.28	85.79	3.0	689.0	686.0	0.92
PM10	2	8660	99.70	95.70	3.0	999.0	996.0	1.03
SO2	3	8562	10.47	15.14	2.0	160.0	158.0	0.16
NO2	4	8571	43.21	30.17	2.0	238.0	236.0	0.33
CO	5	8447	1180.27	1259.98	100.0	10000.0	9900.0	13.71
O3	6	8554	52.09	52.47	2.0	340.0	338.0	0.57
TEMP	7	8727	13.29	11.33	-11.2	39.0	50.2	0.12
PRES	8	8727	1013.50	10.37	988.4	1039.4	51.0	0.11
DEWP	9	8727	2.74	13.13	-31.5	25.4	56.9	0.14
RAIN	10	8727	0.08	0.88	0.0	30.9	30.9	0.01

1-10 of 12 rows Previous **1** 2 Next

```
------------------------------------------------------------------------------------------------------------
group: 2016
```

	vars	n	mean	sd	min	max	range	se
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>

	vars	n	mean	sd	min	max	range	se
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
PM2.5	1	8463	71.77	77.27	3.0	941.0	938.0	0.84
PM10	2	8517	87.47	81.84	3.0	917.0	914.0	0.89
SO2	3	8350	9.68	13.26	2.0	239.0	237.0	0.15
NO2	4	8340	42.95	31.44	2.0	195.0	193.0	0.34
CO	5	8347	1081.36	1133.79	100.0	9900.0	9800.0	12.41
O3	6	7859	56.12	55.94	1.0	339.0	338.0	0.63
TEMP	7	8776	13.38	11.86	-16.8	36.6	53.4	0.13
PRES	8	8776	1013.54	10.58	990.5	1042.8	52.3	0.11
DEWP	9	8776	2.16	14.22	-36.0	27.5	63.5	0.15
RAIN	10	8776	0.06	0.65	0.0	19.5	19.5	0.01

1-10 of 12 rows Previous **1** 2 Next

```
------------------------------------------------------------------------------------------------------------
group: 2017
```

	vars	n	mean	sd	min	max	range	se
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
PM2.5	1	1401	95.59	116.96	2.0	762.0	760.0	3.12
PM10	2	1401	107.18	124.67	3.0	951.0	948.0	3.33
SO2	3	1402	21.20	20.06	2.0	157.0	155.0	0.54
NO2	4	1402	55.59	42.60	2.0	258.0	256.0	1.14
CO	5	1388	1760.88	1921.45	100.0	9600.0	9500.0	51.57
O3	6	1401	42.08	32.47	1.0	173.0	172.0	0.87

	vars	n	mean	sd	min	max	range	se
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
TEMP	7	1406	-0.01	4.68	-11.0	15.4	26.4	0.12
PRES	8	1406	1024.26	5.84	1002.8	1037.6	34.8	0.16
DEWP	9	1406	-13.26	6.40	-27.4	0.8	28.2	0.17
RAIN	10	1406	0.00	0.04	0.0	1.1	1.1	0.00

1-10 of 12 rows Previous **1** 2 Next