Mid-Term

Isaiah Thompson Ocansey

2022-07-18

```
 rm(list=ls())
```

```
library(stringr)
library(data.table)
data <- fread("C:\\Users\\thomo\\Downloads\\winequality-red.csv")
names(data)<-str_replace_all(names(data), c(" " = "." , "," = "" ))
head(data);dim(data)
```

```
##    fixed.acidity volatile.acidity citric.acid residual.sugar chlorides
## 1:           7.4             0.70        0.00            1.9     0.076
## 2:           7.8             0.88        0.00            2.6     0.098
## 3:           7.8             0.76        0.04            2.3     0.092
## 4:          11.2             0.28        0.56            1.9     0.075
## 5:           7.4             0.70        0.00            1.9     0.076
## 6:           7.4             0.66        0.00            1.8     0.075
##    free.sulfur.dioxide total.sulfur.dioxide density   pH sulphates alcohol
## 1:                  11                   34  0.9978 3.51      0.56     9.4
## 2:                  25                   67  0.9968 3.20      0.68     9.8
## 3:                  15                   54  0.9970 3.26      0.65     9.8
## 4:                  17                   60  0.9980 3.16      0.58     9.8
## 5:                  11                   34  0.9978 3.51      0.56     9.4
## 6:                  13                   40  0.9978 3.51      0.56     9.4
##    quality
## 1:       5
## 2:       5
## 3:       5
## 4:       6
## 5:       5
## 6:       5
```

```
## [1] 1599   12
```

The data has 12 variables and 1599 observations

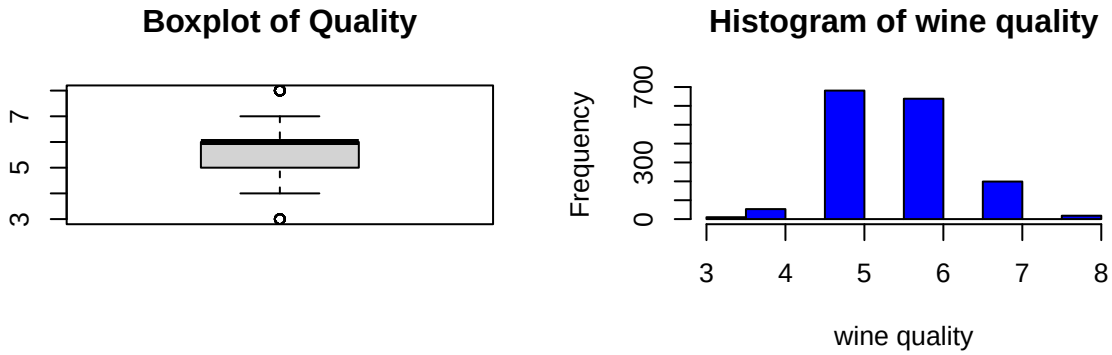
Question(1)

a) Using graphical methods, can you say something about the distribution of the wine quality? Do many of them have low/high quality?

```
par(mfrow = c(2, 2))
boxplot(data$quality,main="Boxplot of Quality")
hist(data$quality, col = "blue",main = "Histogram of wine quality",xlab = "wine quality")

table(data$quality)
```

```
##
##   3   4   5   6   7   8
##  10  53 681 638 199  18
```

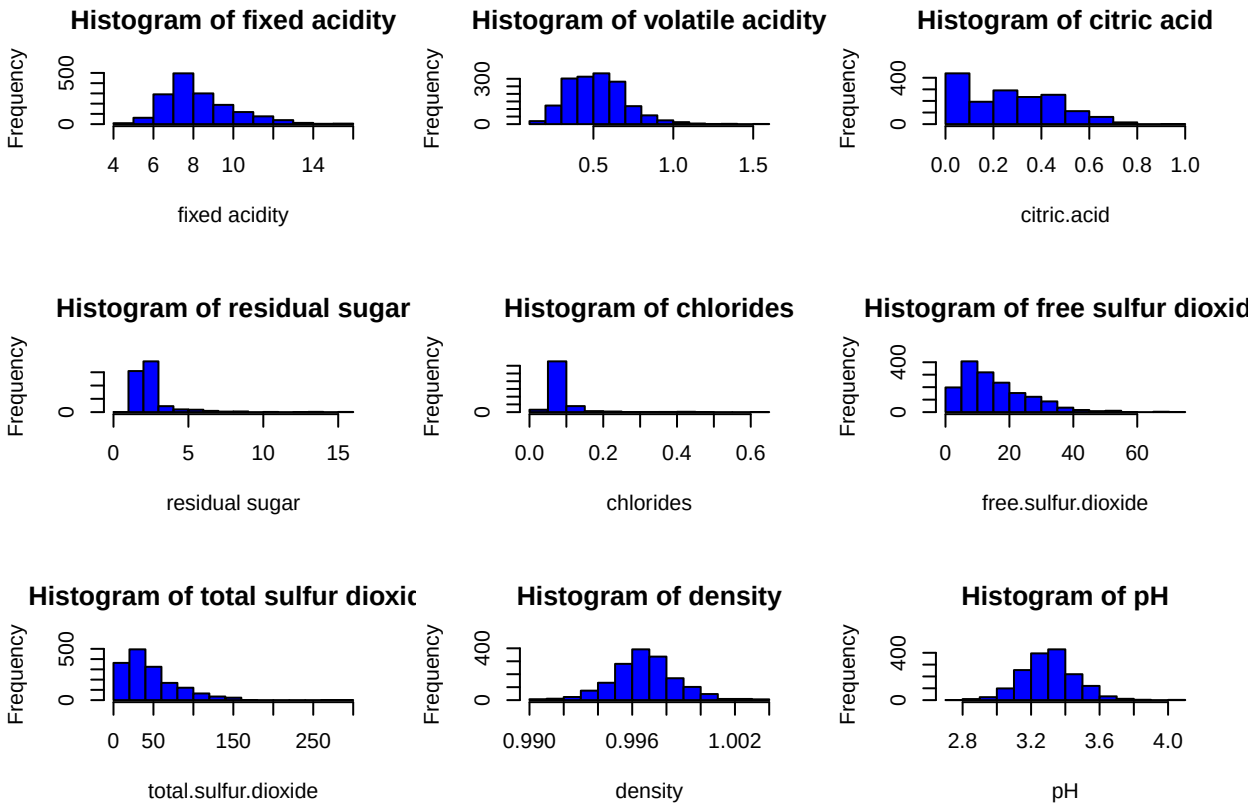


We can observe from the above table that about 681 and 638 of the clients gave 5 and 6 ratings respectively to the quality of wine. Also, 199 and 18 clients gave 7 and 8 ratings respectively. Interestingly, none of the clients gave ratings of 1,2,9 and 10 and this is because the sample was random.So we can confirm that many clients gave a high rating. Graphically, we can observe that the distribution of the wine quality is fairly normal as evident from the histogram plot. We can equally observe from the boxplot that about about 75% of the clients gave a high ratings.

Furthermore, we can observe the distribution of the individual variables using the histogram below

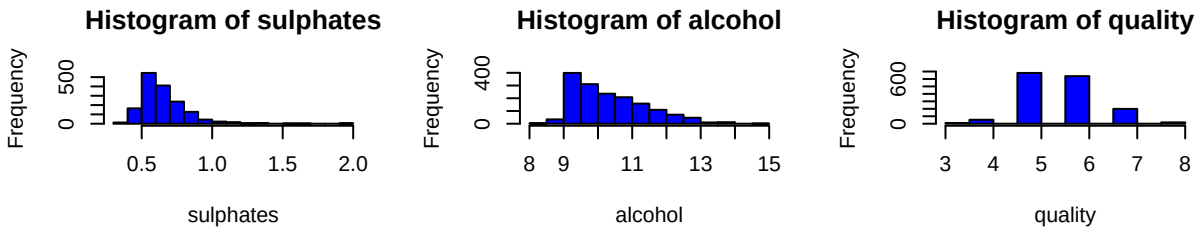
```
attach(data)
par(mfrow = c(3, 3))
hist(fixed.acidity, col = "blue",main = "Histogram of fixed acidity",xlab = "fixed acidity")
hist(volatile.acidity, col = "blue",main = "Histogram of volatile acidity",xlab = "")
hist(citric.acid, col = "blue",main = "Histogram of citric acid",xlab = "citric.acid")
hist(residual.sugar, col = "blue",main = "Histogram of residual sugar",xlab = "residual sugar")
hist(chlorides, col = "blue",main = "Histogram of chlorides",xlab = "chlorides")
```

```r
hist(free.sulfur.dioxide, col = "blue",main = "Histogram of free sulfur dioxide",xlab = "free.sulfur.di
hist(total.sulfur.dioxide, col = "blue",main = "Histogram of total sulfur dioxide",xlab = "total.sulfur
hist(density, col = "blue",main = "Histogram of density",xlab = "density")
hist(pH, col = "blue",main = "Histogram of pH",xlab = "pH")
```




```r
hist(sulphates, col = "blue",main = "Histogram of sulphates",xlab = "sulphates")
hist(alcohol, col = "blue",main = "Histogram of alcohol",xlab = "alcohol")

hist(quality, col = "blue",main = "Histogram of quality",xlab = "quality")
```



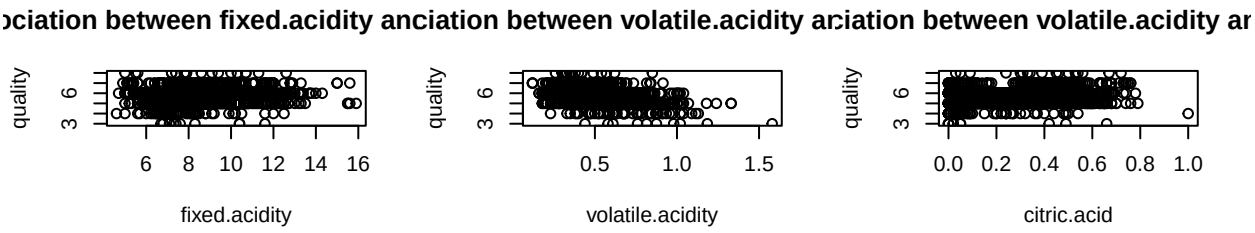
From the above histogram plots, we can observe that fixed acidity, volatile acidity,pH, density,chlorides and suphates are fairly normal whiles citric acid,sidual sugar,free sulphur dioxide,total sulphur dioxide and alcohol are skewed.

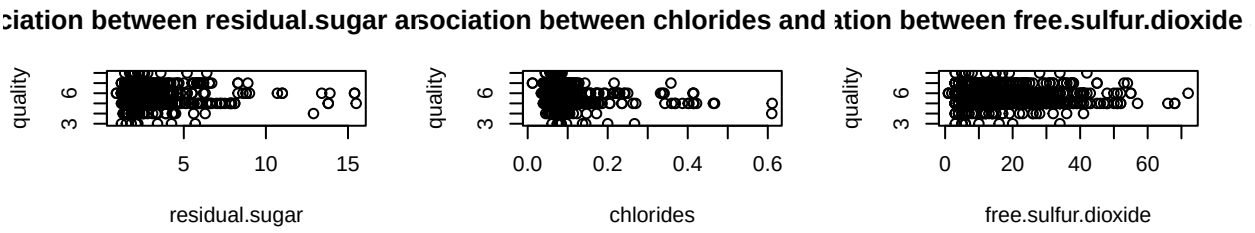
b) Which are the independent variables have either positive or negative association with the output variable "quality"? Do you observe any nonlinear association/ no association at all? Hint: scatter plots for each of variables 1-11 with variable 12.

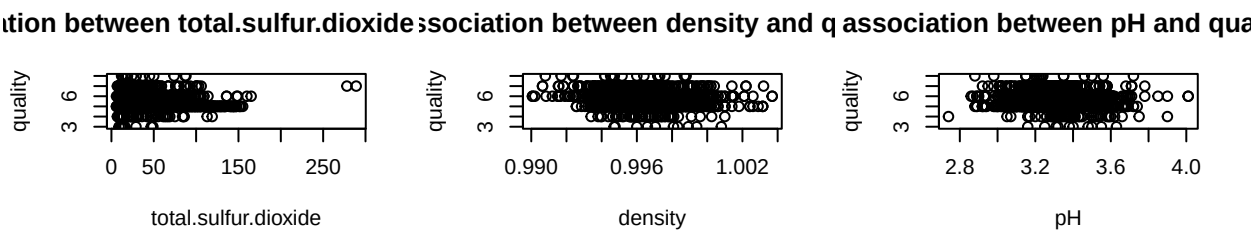
```
attach(data)
```

```
## The following objects are masked from data (pos = 3):
##
##     alcohol, chlorides, citric.acid, density, fixed.acidity,
##     free.sulfur.dioxide, pH, quality, residual.sugar, sulphates,
##     total.sulfur.dioxide, volatile.acidity
```

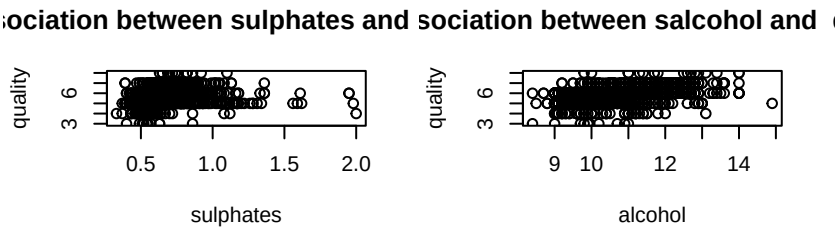
```
par(mfrow = c(3, 3))
plot(x=fixed.acidity,y=quality, main = "association between fixed.acidity and quality")
plot(x=volatile.acidity,y=quality, main = "association between volatile.acidity and quality")
plot(x=citric.acid,y=quality, main = "association between volatile.acidity and quality")
plot(x=residual.sugar,y=quality, main = "association between residual.sugar and quality")
plot(x=chlorides,y=quality, main = "association between chlorides and quality")
plot(x=free.sulfur.dioxide,y=quality, main = "association between free.sulfur.dioxide and quality")
plot(x=total.sulfur.dioxide,y=quality, main = "association between total.sulfur.dioxide and quality")
plot(x=density,y=quality, main = "association between density and quality")
plot(x=pH,y=quality, main = "association between pH and quality")
```








```
plot(x=sulphates,y=quality, main = "association between sulphates and  quality")
plot(x=alcohol,y=quality, main = "association between salcohol and  quality")
```

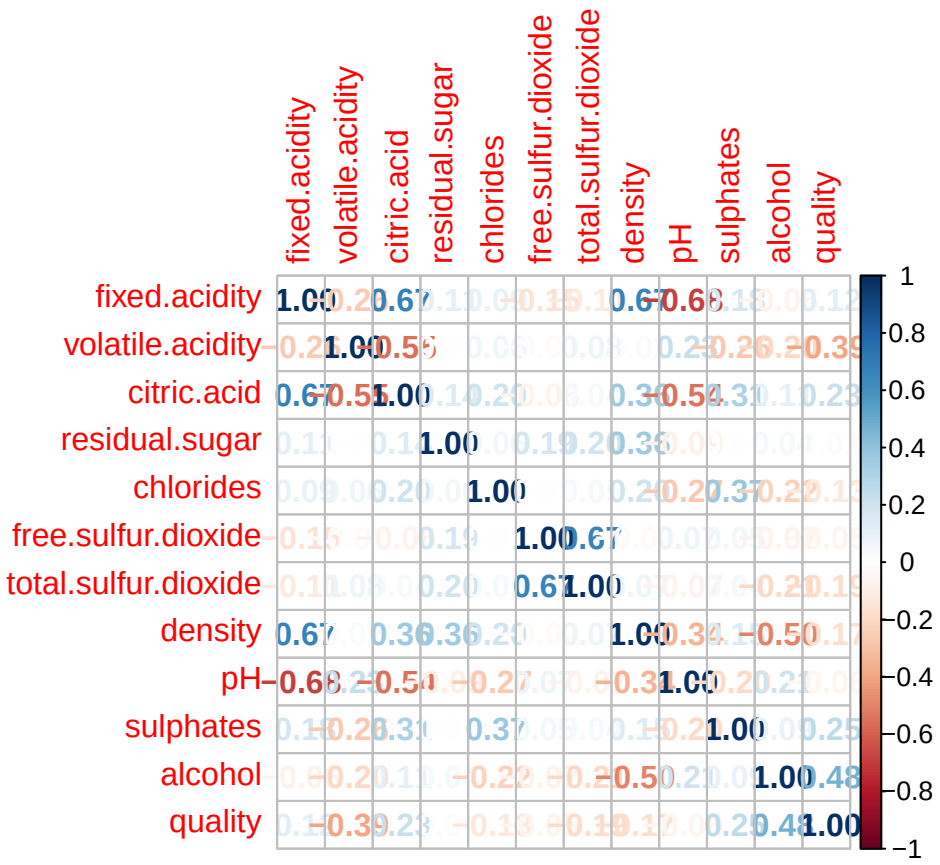


From the above scatter plot, we can observe there's some relationships between the response variable,quality and the independent variables however, we would be able to observe in details the kind of relationship using the corplot and the nlcor plot below.

```
library(corrplot)
```

```
## corrplot 0.92 loaded
```

```
x = cor(data)
corrplot(x, method = 'number') # colorful number
```




```
options(warn=-1)
library(nlcor)
```

```
##
## Attaching package: 'nlcor'
```

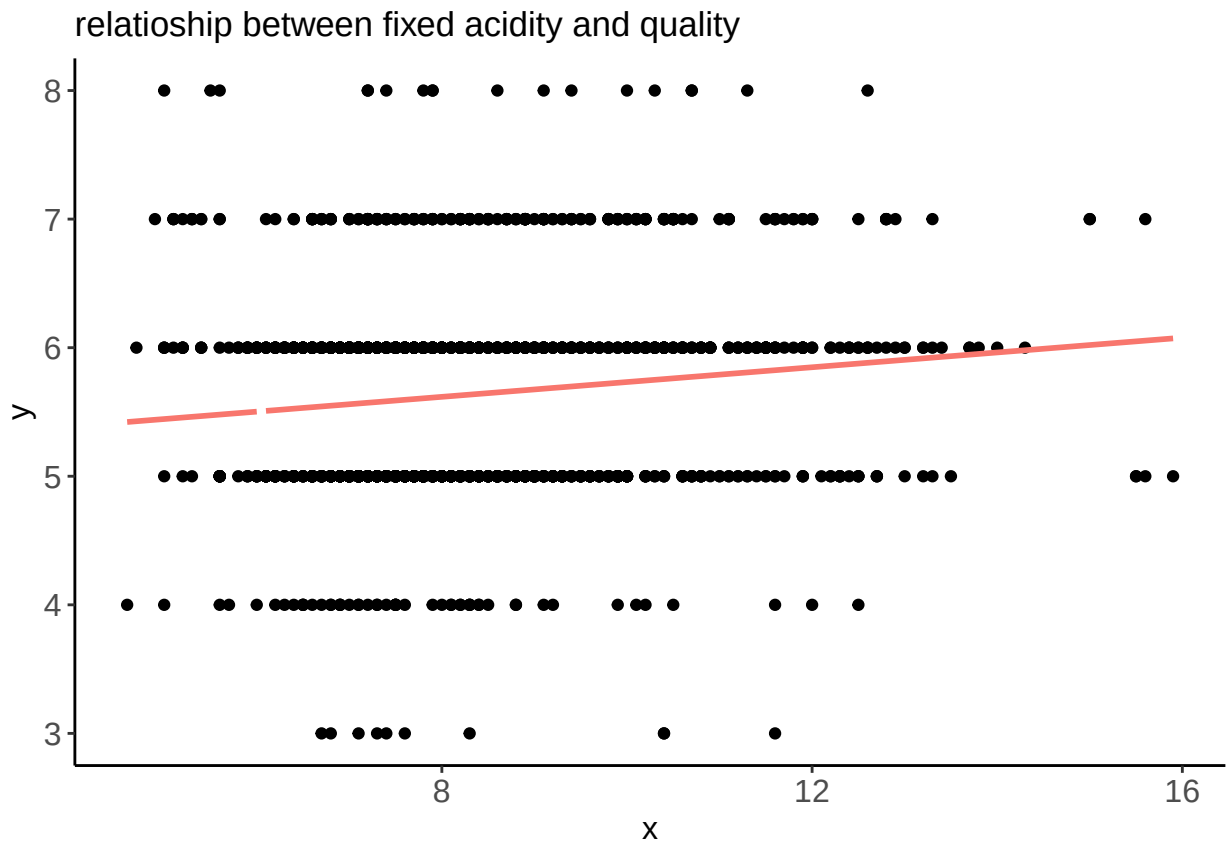
```
## The following object is masked _by_ '.GlobalEnv':
##
##     x
```

```
attach(data)
```

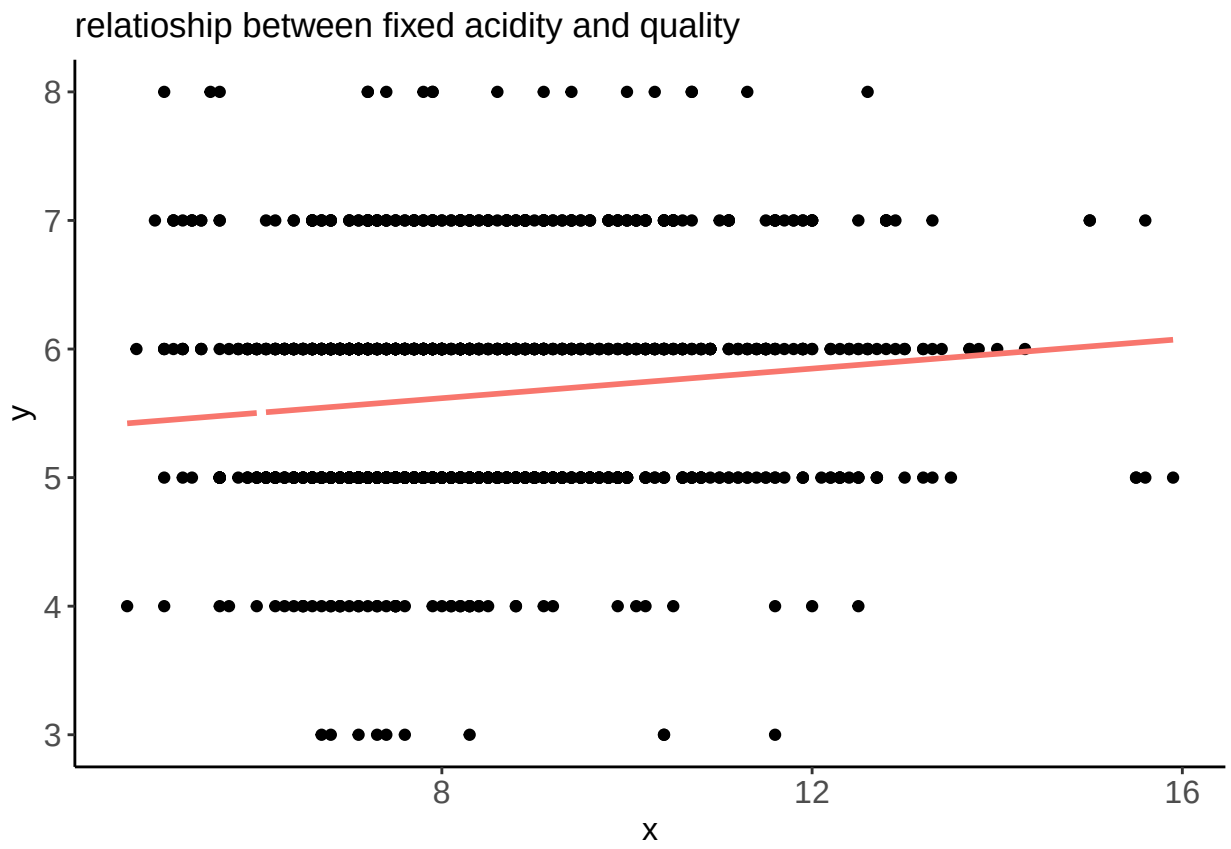
```
## The following objects are masked from data (pos = 5):
##
##     alcohol, chlorides, citric.acid, density, fixed.acidity,
##     free.sulfur.dioxide, pH, quality, residual.sugar, sulphates,
##     total.sulfur.dioxide, volatile.acidity
```

```
## The following objects are masked from data (pos = 6):
##
##     alcohol, chlorides, citric.acid, density, fixed.acidity,
##     free.sulfur.dioxide, pH, quality, residual.sugar, sulphates,
##     total.sulfur.dioxide, volatile.acidity
```

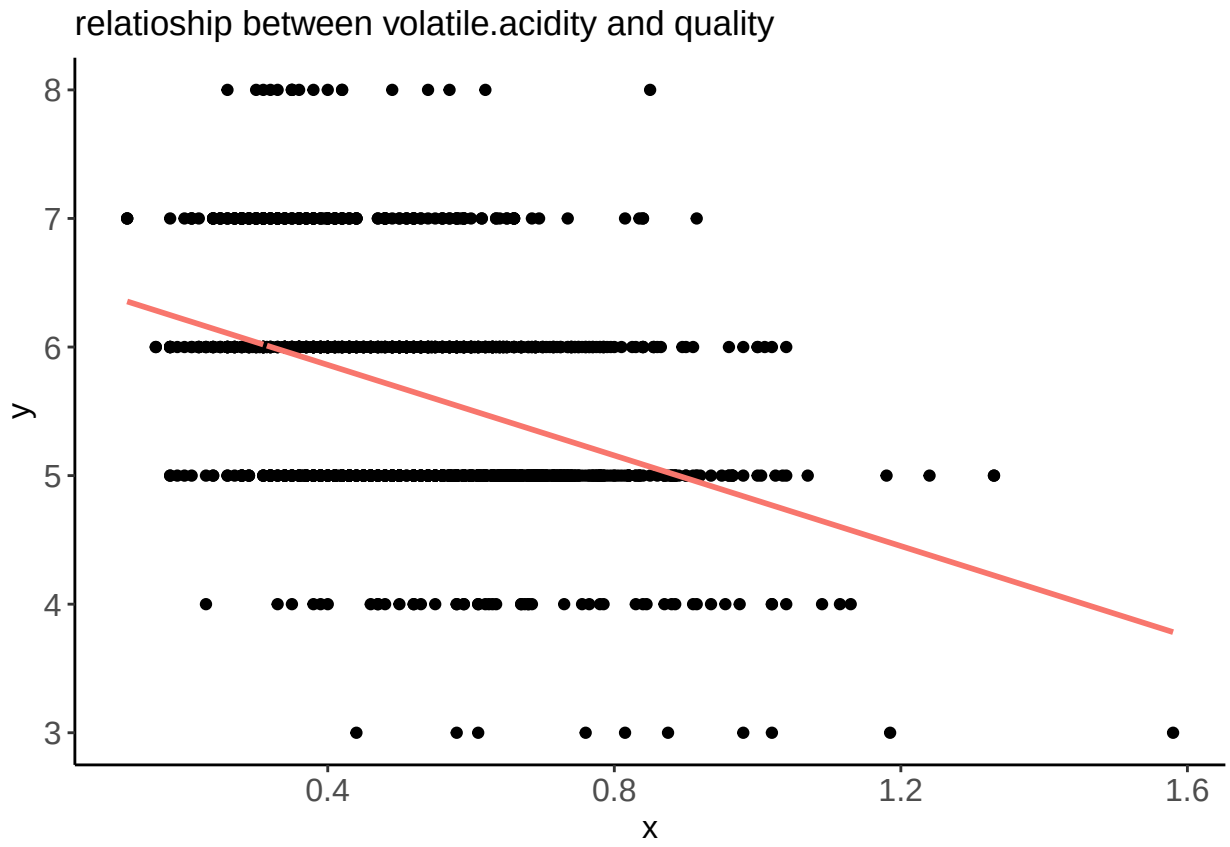
```
par(mfrow = c(3, 3))
nlcor(fixed.acidity,quality,chart_title= "relatioship between fixed acidity and quality", plt = T)
```



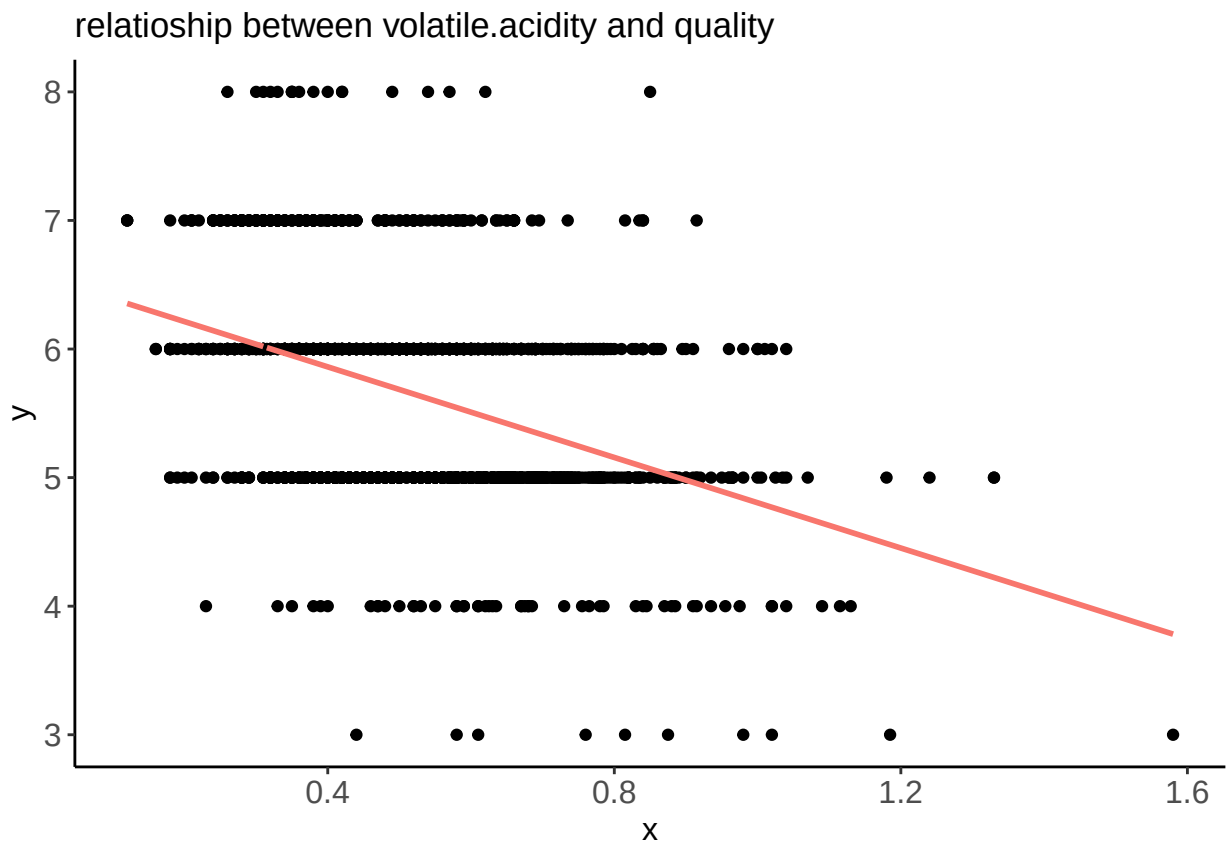

```
## $cor.estimate
## [1] 0.1240516
##
## $adjusted.p.value
## [1] 0
##
## $cor.plot
```



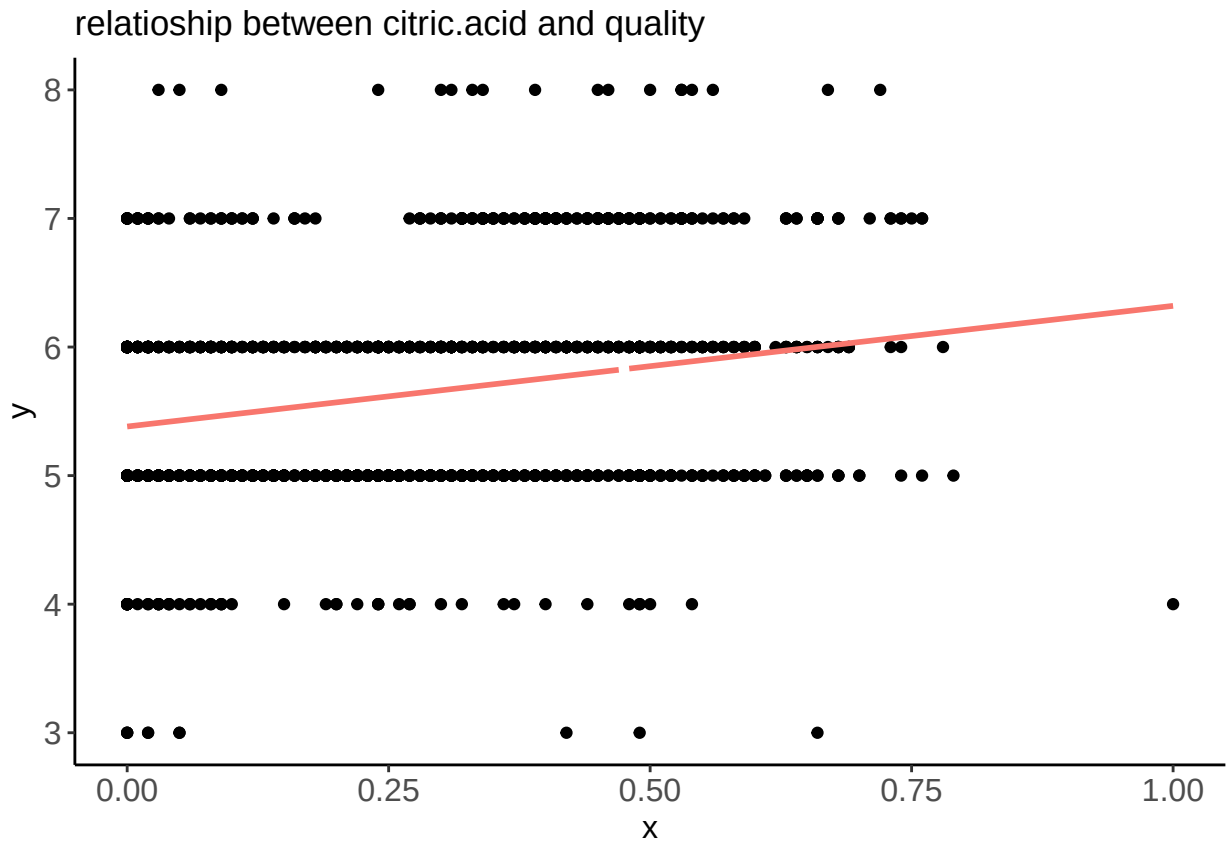

```
nlcor(volatile.acidity,quality, chart_title= "relatioship between volatile.acidity and quality",plt = T
```



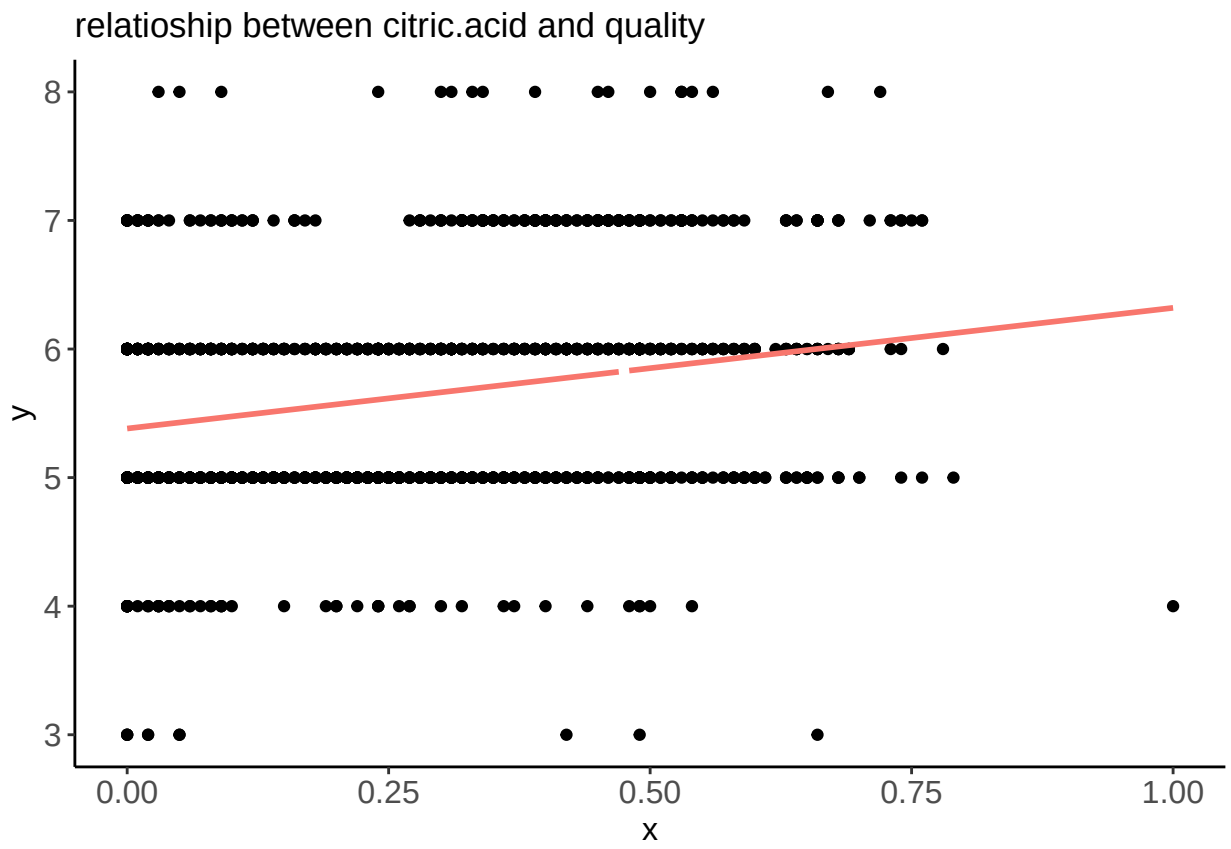

```
## $cor.estimate
## [1] 0.3905578
##
## $adjusted.p.value
## [1] 0
##
## $cor.plot
```



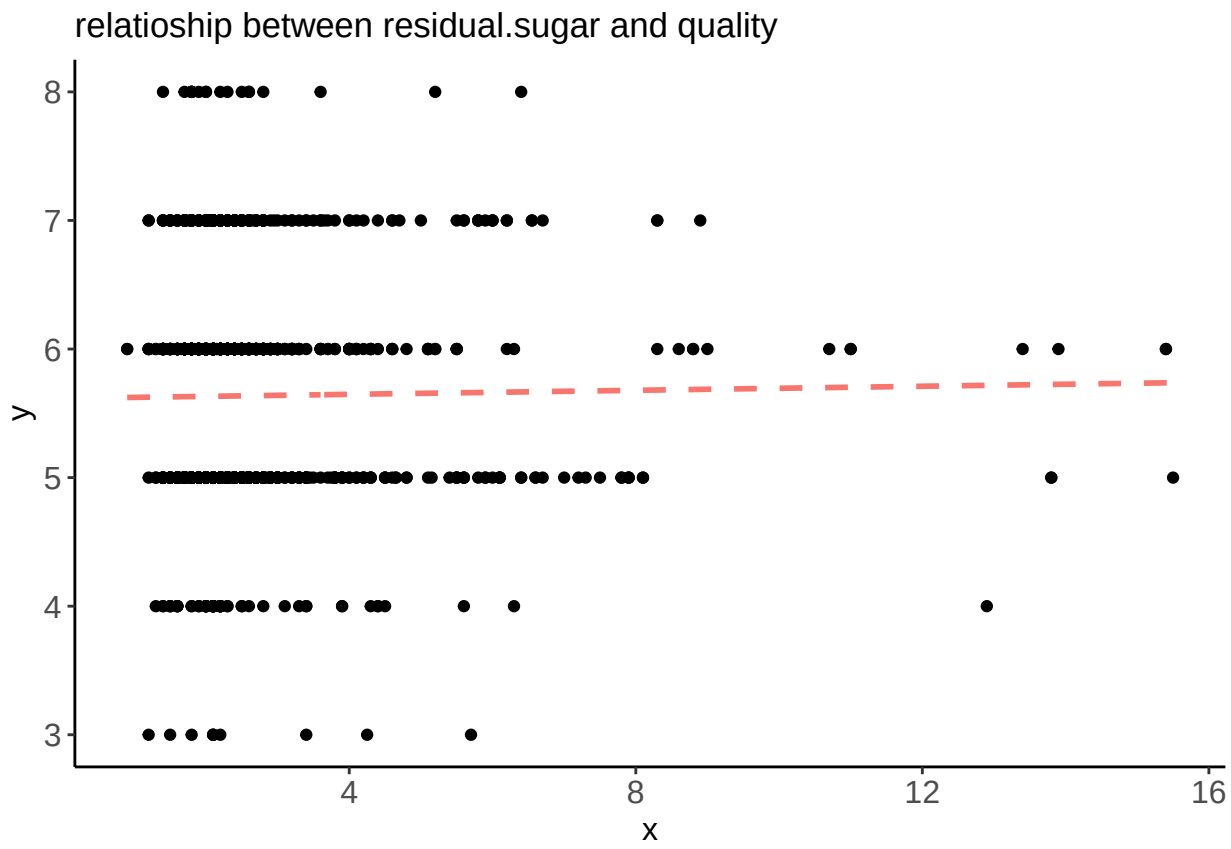

```
nlcor(citric.acid,quality, chart_title= "relatioship between citric.acid and quality",plt = T)
```



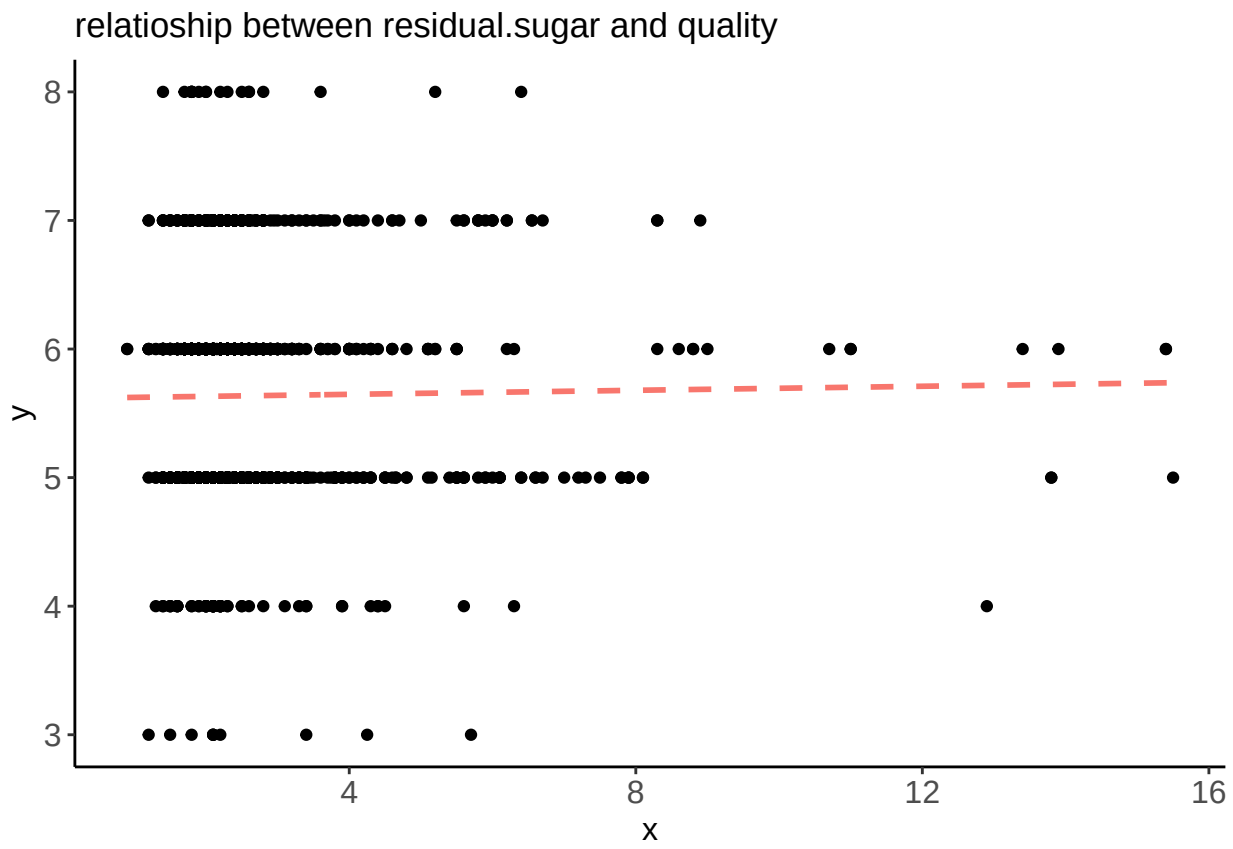

```
## $cor.estimate
## [1] 0.2263725
##
## $adjusted.p.value
## [1] 0
##
## $cor.plot
```



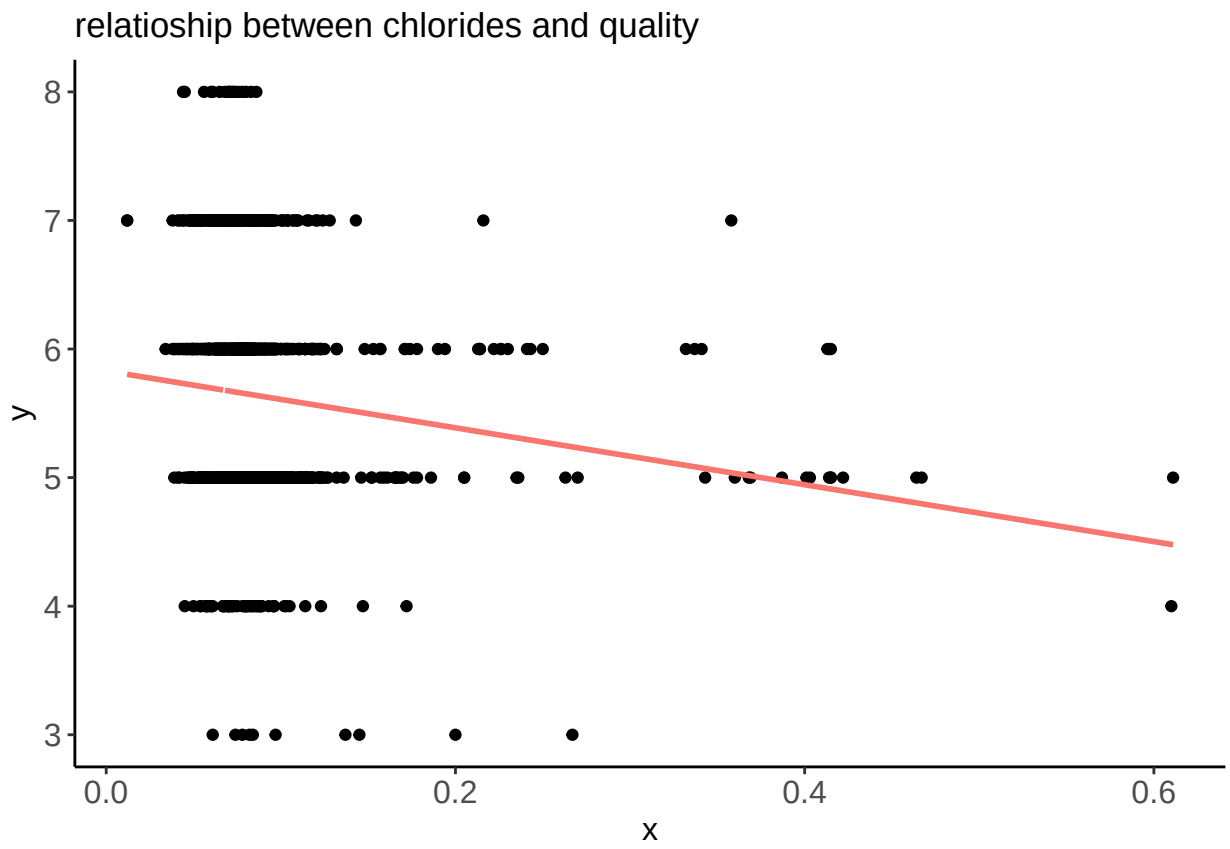

```
nlcor(residual.sugar,quality, chart_title= "relatioship between residual.sugar and quality",plt = T)
```



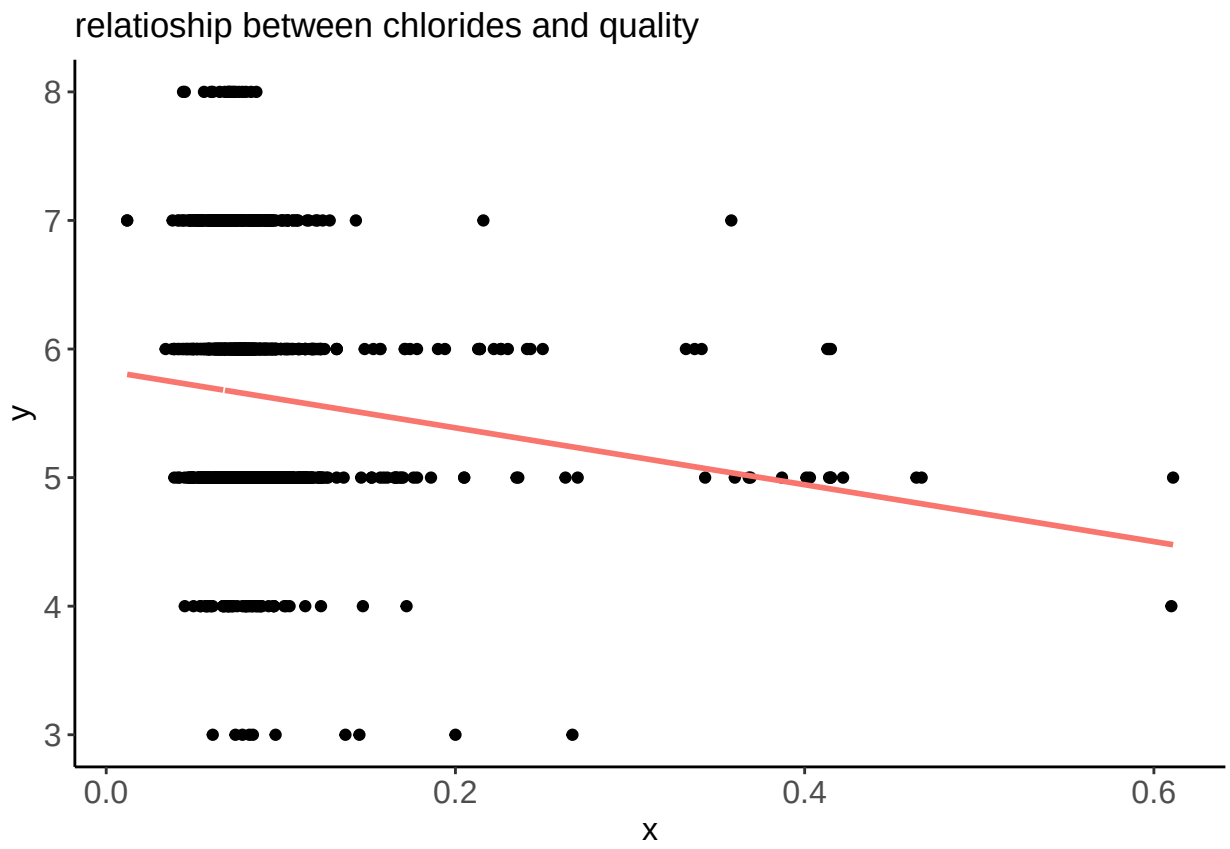

```
## $cor.estimate
## [1] 0.01373164
##
## $adjusted.p.value
## [1] 0.58
##
## $cor.plot
```



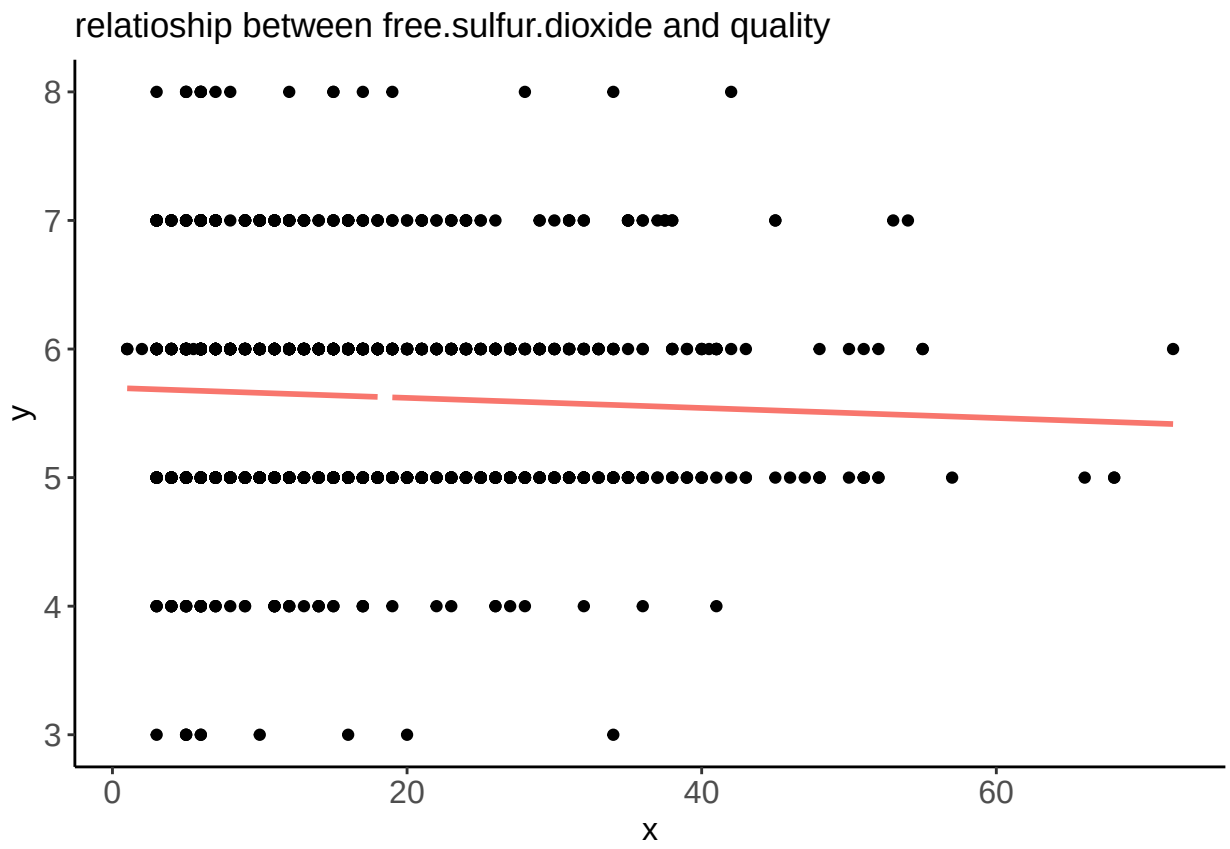

```
nlcor(chlorides,quality,chart_title= "relatioship between chlorides and quality", plt = T)
```



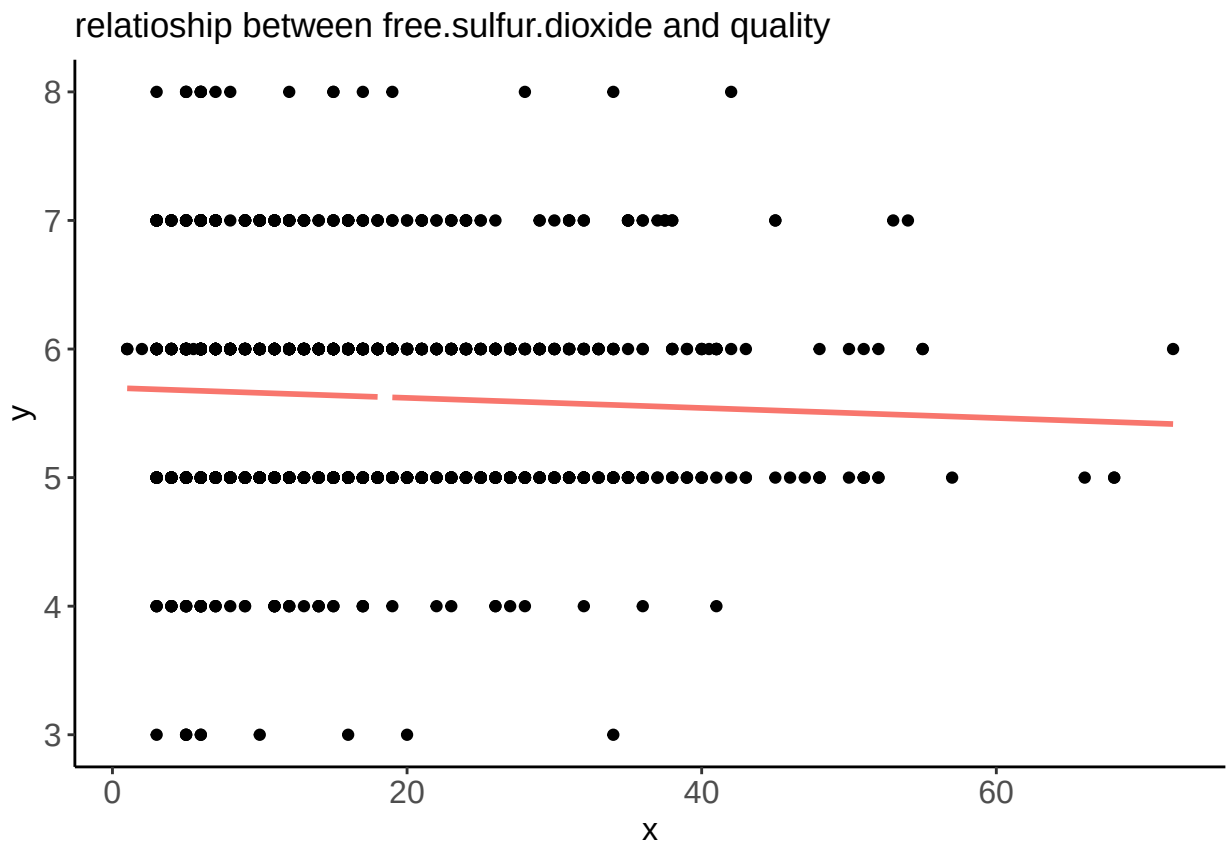

```
## $cor.estimate
## [1] 0.1289066
##
## $adjusted.p.value
## [1] 0
##
## $cor.plot
```



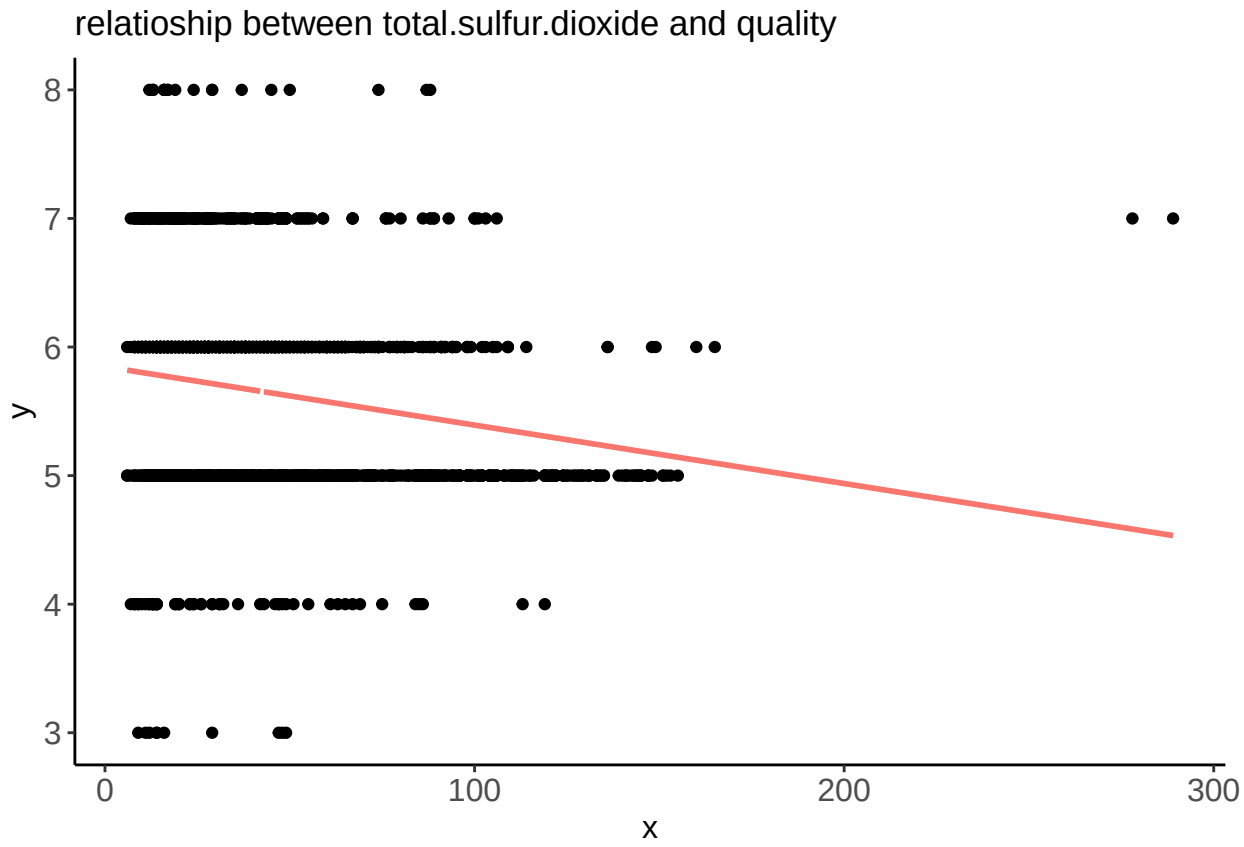

```
nlcor(free.sulfur.dioxide,quality, chart_title= "relatioship between free.sulfur.dioxide and quality",pl
```



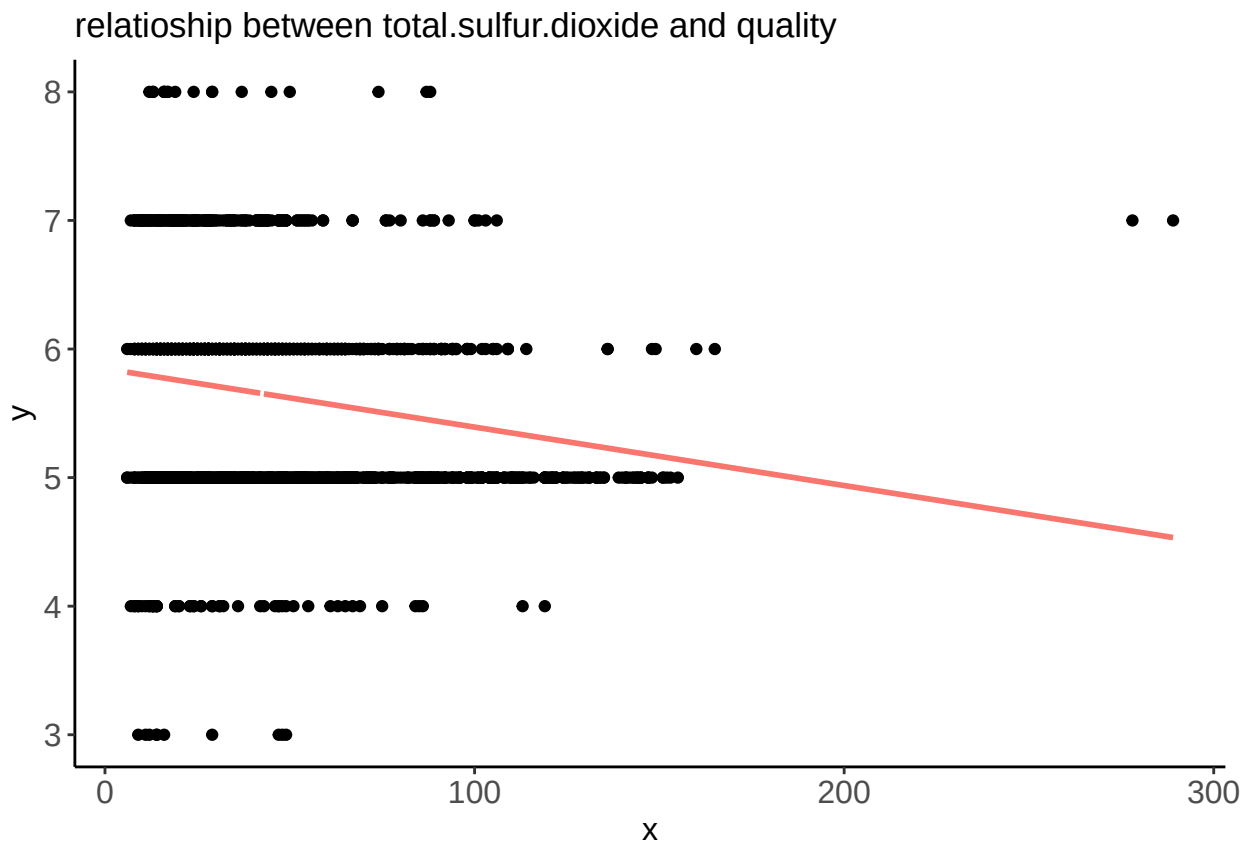

```
## $cor.estimate
## [1] 0.05065606
##
## $adjusted.p.value
## [1] 0.04
##
## $cor.plot
```



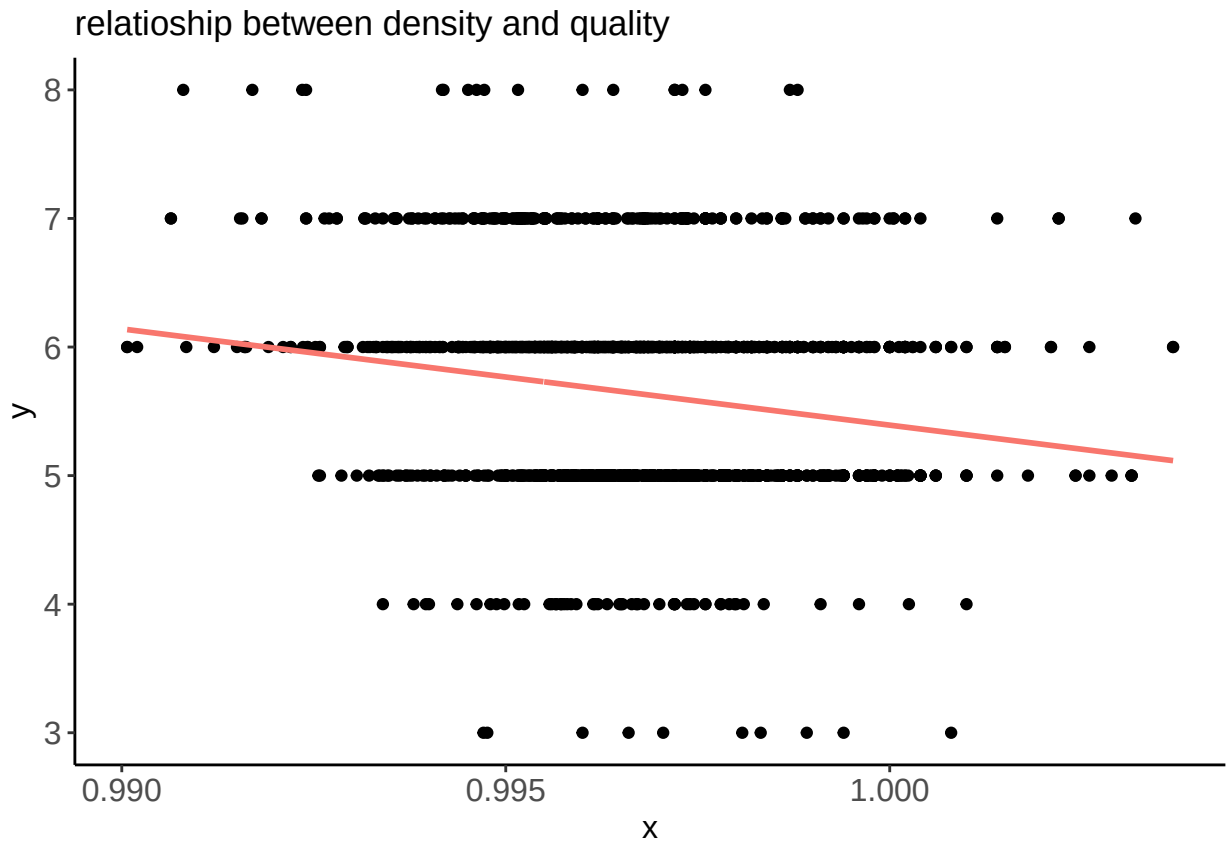

```
nlcor(total.sulfur.dioxide,quality,chart_title= "relatioship between total.sulfur.dioxide and quality",
```



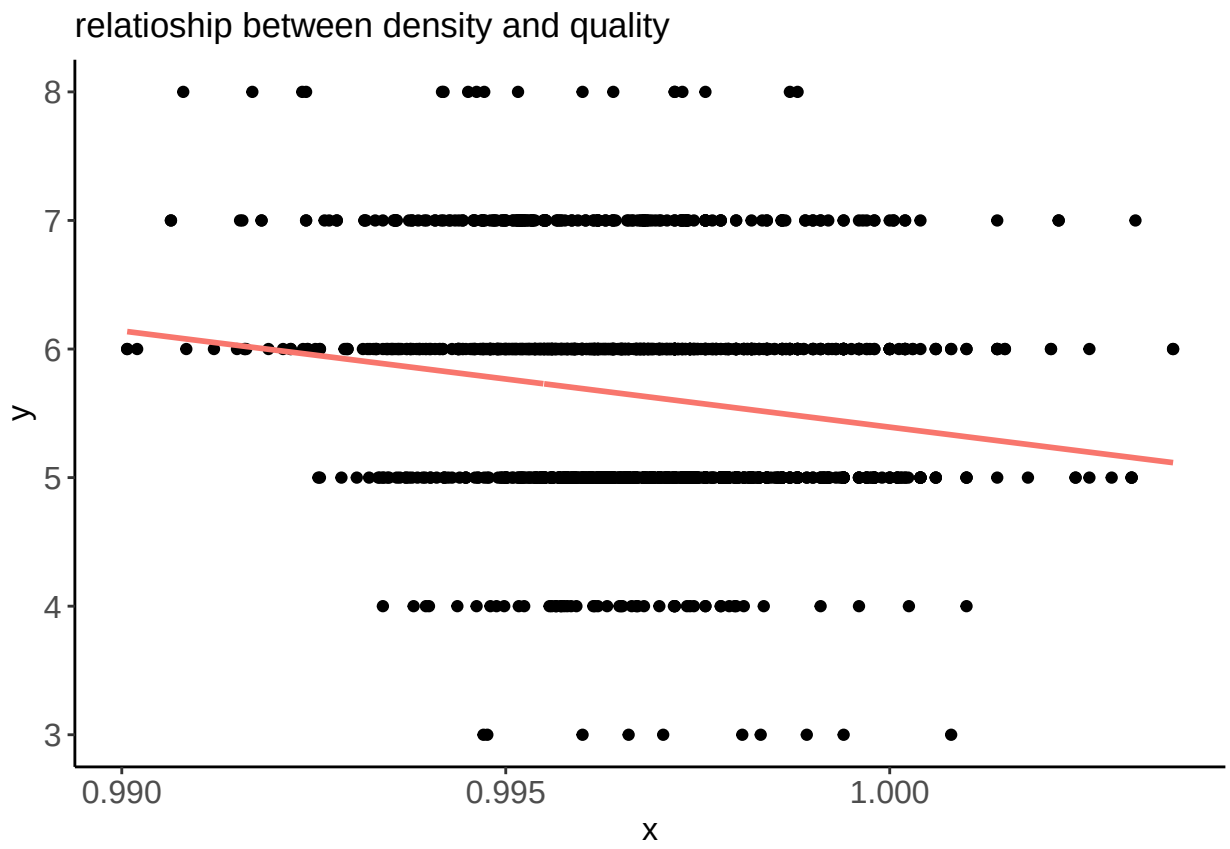

```
## $cor.estimate
## [1] 0.1851003
##
## $adjusted.p.value
## [1] 0
##
## $cor.plot
```



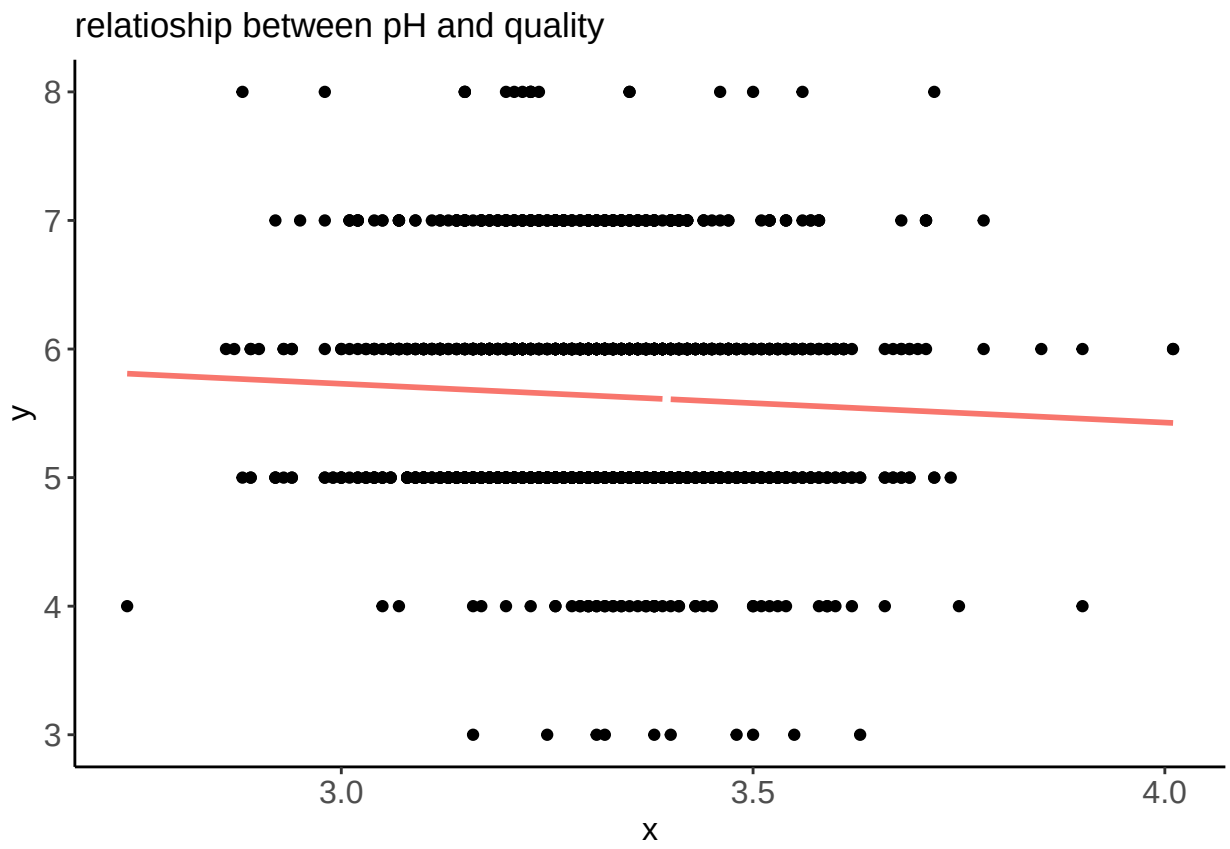

```
nlcor(density,quality, chart_title= "relatioship between density and quality", plt = T)
```



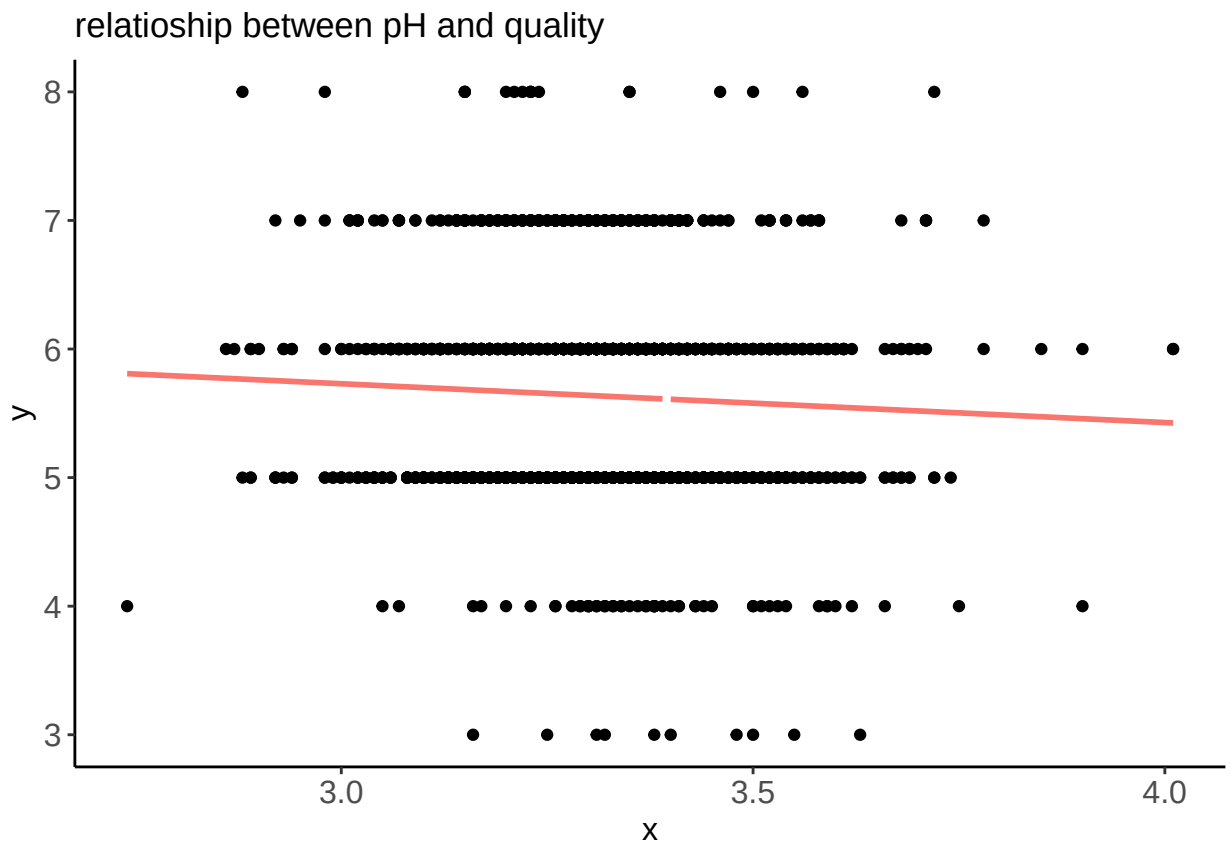

```
## $cor.estimate
## [1] 0.1749192
##
## $adjusted.p.value
## [1] 0
##
## $cor.plot
```



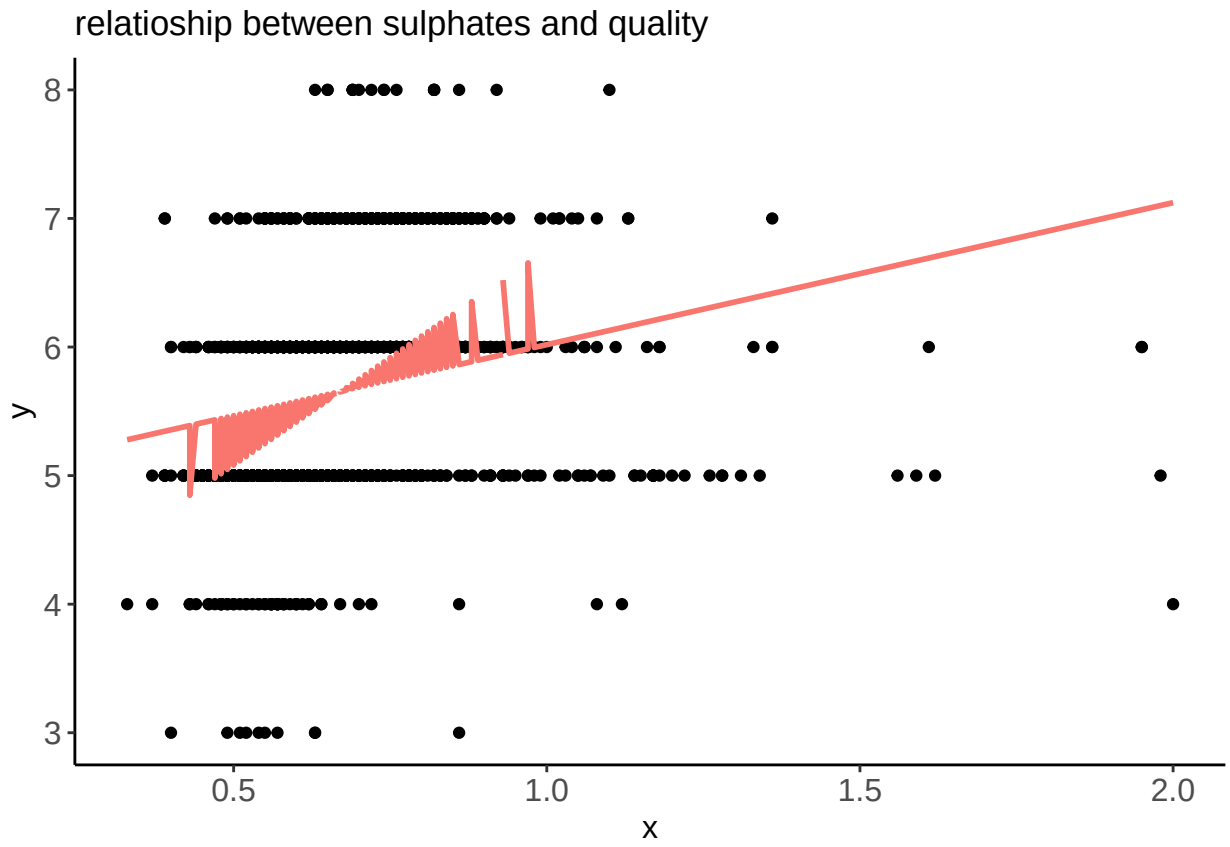

```
nlcor(pH,quality,chart_title= "relatioship between pH and quality", plt = T)
```



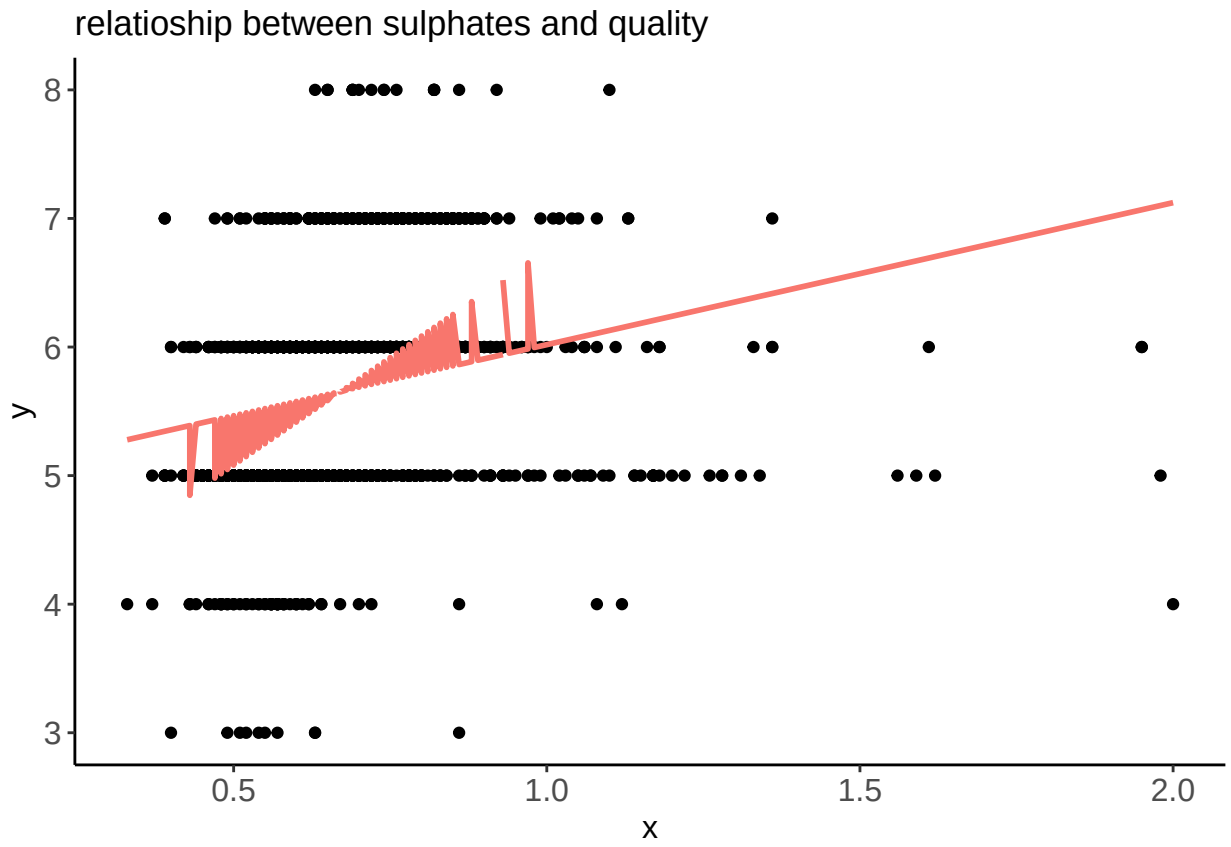

```
## $cor.estimate
## [1] 0.05773139
##
## $adjusted.p.value
## [1] 0.02
##
## $cor.plot
```



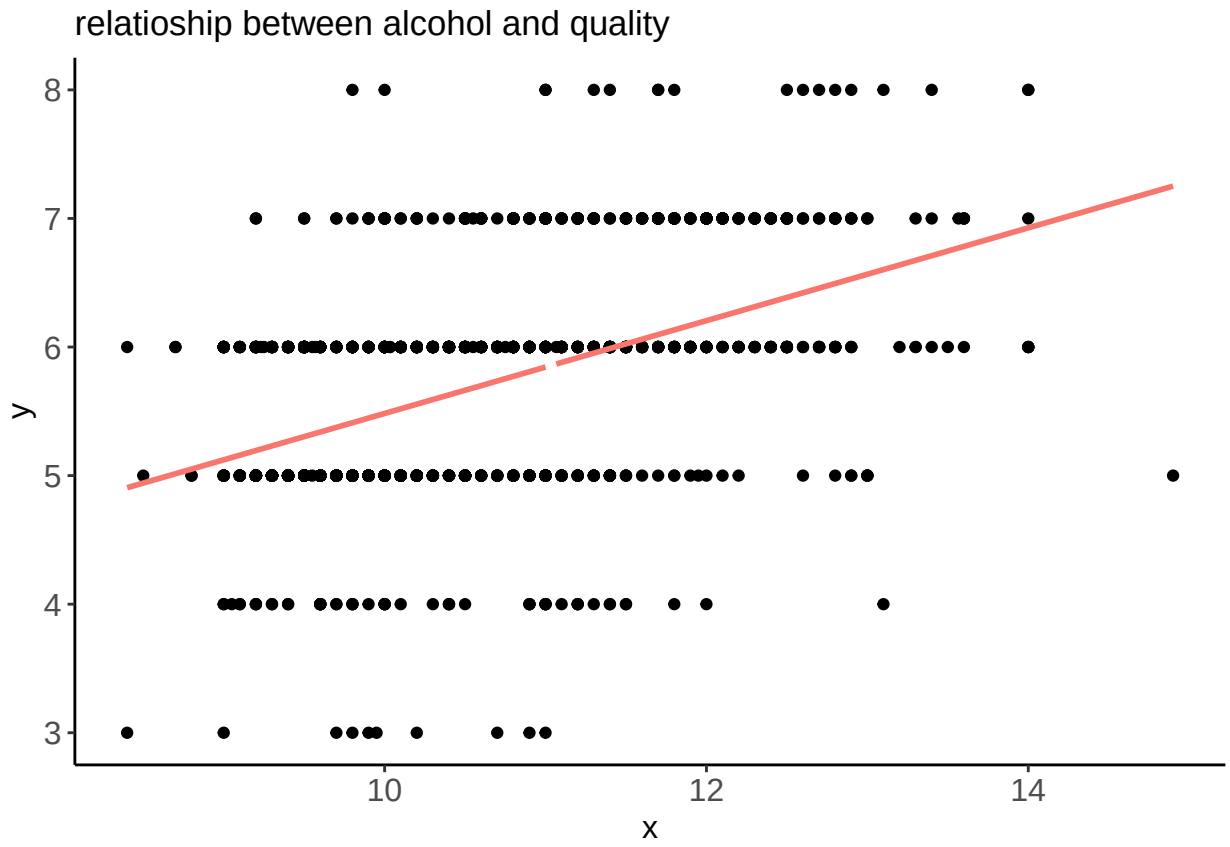

```
nlcor(sulphates,quality,chart_title= "relatioship between sulphates and quality", plt = T)
```



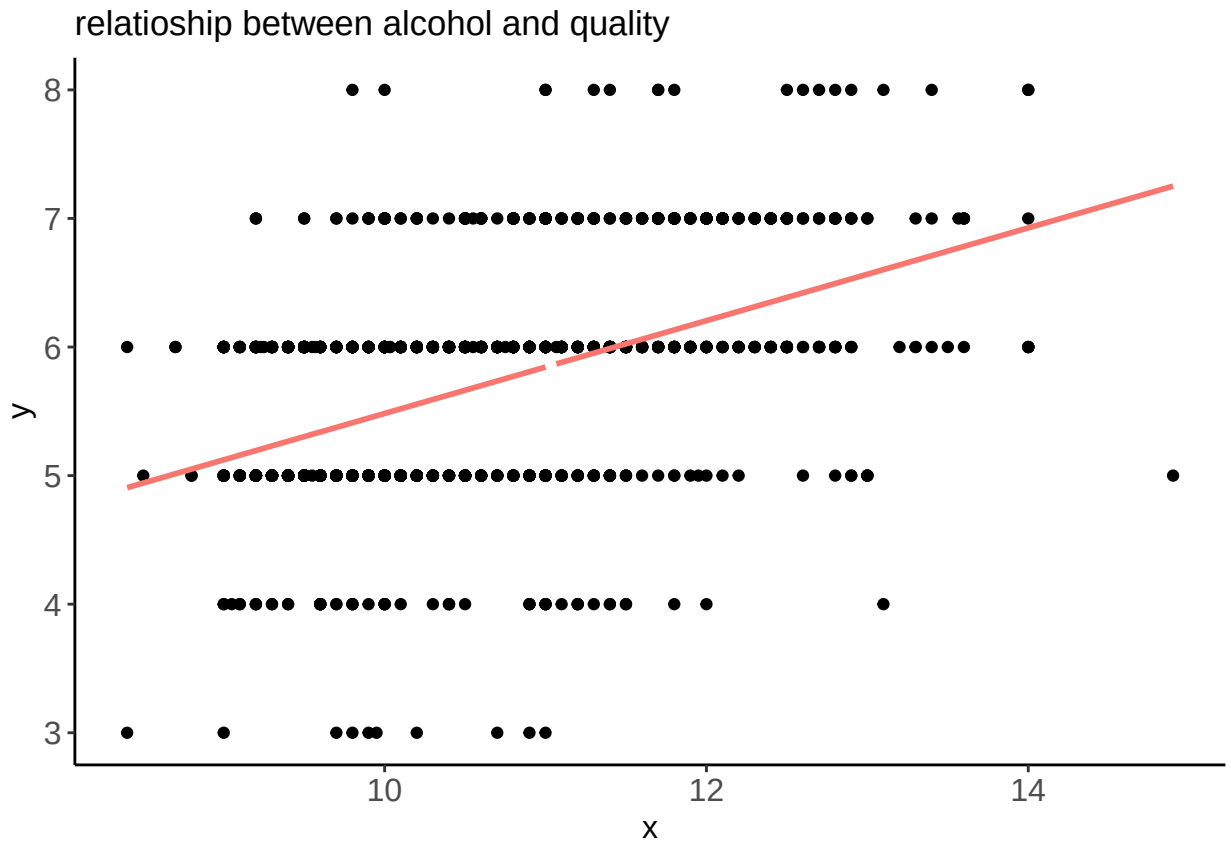

```
## $cor.estimate
## [1] 0.3562027
##
## $adjusted.p.value
## [1] 4.125641e-05
##
## $cor.plot
```




```
nlcor(alcohol,quality,chart_title= "relatioship between alcohol and quality", plt = T)
```



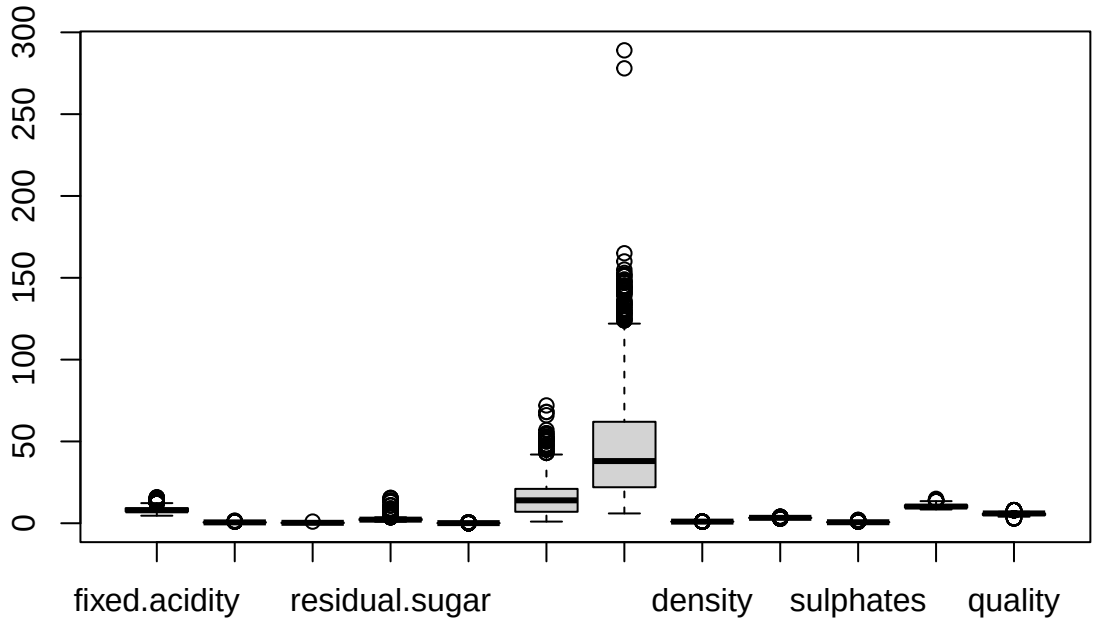

```
## $cor.estimate
## [1] 0.4761663
##
## $adjusted.p.value
## [1] 0
##
## $cor.plot
```



From the plots above, it can be observed that there's a fair positive linear relationship between fixed acidity,citric acid,alcohol and the output variable, Quality. Whiles there's a negative linear relationship between volatile acidity,chlorides,total sulphur dioxide, density, pH and the output variable, Quality, there seems to be a nonlinear relationship between sulphates and quality and almost no relationship between quality and residual sugar.

c)Construct the boxplots for all the variables in one figure (side by side)

```
boxplot(data)
```



The boxplot for all the variables are plotted above, there seems to be few potential outliers

d) Compute the descriptive statistics (mean, min, max, variance, median, standard deviation, skewness and Kurtosis) for all the variables.

```
library(psych)
describe(data)
```

```
##                       vars    n  mean    sd median trimmed   mad  min    max
## fixed.acidity            1 1599  8.32  1.74   7.90    8.15  1.48 4.60  15.90
## volatile.acidity         2 1599  0.53  0.18   0.52    0.52  0.18 0.12   1.58
## citric.acid              3 1599  0.27  0.19   0.26    0.26  0.25 0.00   1.00
## residual.sugar           4 1599  2.54  1.41   2.20    2.26  0.44 0.90  15.50
## chlorides                5 1599  0.09  0.05   0.08    0.08  0.01 0.01   0.61
## free.sulfur.dioxide      6 1599 15.87 10.46  14.00   14.58 10.38 1.00  72.00
## total.sulfur.dioxide     7 1599 46.47 32.90  38.00   41.84 26.69 6.00 289.00
## density                  8 1599  1.00  0.00   1.00    1.00  0.00 0.99   1.00
## pH                       9 1599  3.31  0.15   3.31    3.31  0.15 2.74   4.01
## sulphates               10 1599  0.66  0.17   0.62    0.64  0.12 0.33   2.00
## alcohol                 11 1599 10.42  1.07  10.20   10.31  1.04 8.40  14.90
## quality                 12 1599  5.64  0.81   6.00    5.59  1.48 3.00   8.00
##                        range skew kurtosis   se
## fixed.acidity          11.30 0.98     1.12 0.04
## volatile.acidity        1.46 0.67     1.21 0.00
## citric.acid             1.00 0.32    -0.79 0.00
## residual.sugar         14.60 4.53    28.49 0.04
```

```
## chlorides              0.60 5.67    41.53 0.00
## free.sulfur.dioxide   71.00 1.25     2.01 0.26
## total.sulfur.dioxide 283.00 1.51     3.79 0.82
## density                0.01 0.07     0.92 0.00
## pH                     1.27 0.19     0.80 0.00
## sulphates              1.67 2.42    11.66 0.00
## alcohol                6.50 0.86     0.19 0.03
## quality                5.00 0.22     0.29 0.02
```

```
var(data)
```

```
##                       fixed.acidity volatile.acidity   citric.acid
## fixed.acidity          3.031416389    -7.985142e-02  0.2278200037
## volatile.acidity      -0.079851417     3.206238e-02 -0.0192716208
## citric.acid            0.227820004    -1.927162e-02  0.0379474831
## residual.sugar         0.281756262     4.841910e-04  0.0394342700
## chlorides              0.007678692     5.165869e-04  0.0018687248
## free.sulfur.dioxide   -2.800921493    -1.967359e-02 -0.1242521139
## total.sulfur.dioxide  -6.482345858     4.504257e-01  0.2276972740
## density                0.002195224     7.443665e-06  0.0001341746
## pH                    -0.183585704     6.494699e-03 -0.0162975823
## sulphates              0.054010092    -7.921434e-03  0.0103277145
## alcohol               -0.114421153    -3.860022e-02  0.0228151729
## quality                0.174423588    -5.647588e-02  0.0356118929
##                       residual.sugar     chlorides free.sulfur.dioxide
## fixed.acidity          0.2817562623  7.678692e-03       -2.800921e+00
## volatile.acidity       0.0004841910  5.165869e-04       -1.967359e-02
## citric.acid            0.0394342700  1.868725e-03       -1.242521e-01
## residual.sugar         1.9878971330  3.690176e-03        2.758611e+00
## chlorides              0.0036901759  2.215143e-03        2.738303e-03
## free.sulfur.dioxide    2.7586114522  2.738303e-03        1.094149e+02
## total.sulfur.dioxide   9.4164414790  7.338675e-02        2.297375e+02
## density                0.0009454109  1.782176e-05       -4.332504e-04
## pH                    -0.0186442890 -1.925745e-03        1.136531e-01
## sulphates              0.0013209414  2.961878e-03        9.159247e-02
## alcohol                0.0632189598 -1.109152e-02       -7.736984e-01
## quality                0.0156350457 -4.899545e-03       -4.279071e-01
##                       total.sulfur.dioxide       density            pH
## fixed.acidity               -6.482346e+00  2.195224e-03 -1.835857e-01
## volatile.acidity             4.504257e-01  7.443665e-06  6.494699e-03
## citric.acid                  2.276973e-01  1.341746e-04 -1.629758e-02
## residual.sugar               9.416441e+00  9.454109e-04 -1.864429e-02
## chlorides                    7.338675e-02  1.782176e-05 -1.925745e-03
## free.sulfur.dioxide          2.297375e+02 -4.332504e-04  1.136531e-01
## total.sulfur.dioxide         1.082102e+03  4.424727e-03 -3.376988e-01
## density                      4.424727e-03  3.562029e-06 -9.956395e-05
## pH                          -3.376988e-01 -9.956395e-05  2.383518e-02
## sulphates                    2.394710e-01  4.750962e-05 -5.146186e-03
## alcohol                     -7.209298e+00 -9.979518e-04  3.383162e-02
## quality                     -4.917237e+00 -2.666037e-04 -7.197822e-03
##                          sulphates       alcohol       quality
## fixed.acidity         5.401009e-02 -0.1144211534  0.1744235876
## volatile.acidity     -7.921434e-03 -0.0386002214 -0.0564758833
## citric.acid           1.032771e-02  0.0228151729  0.0356118929
```

```
## residual.sugar        1.320941e-03  0.0632189598  0.0156350457
## chlorides             2.961878e-03 -0.0110915178 -0.0048995449
## free.sulfur.dioxide   9.159247e-02 -0.7736984004 -0.4279070696
## total.sulfur.dioxide  2.394710e-01 -7.2092978950 -4.9172370717
## density               4.750962e-05 -0.0009979518 -0.0002666037
## pH                   -5.146186e-03  0.0338316166 -0.0071978223
## sulphates             2.873262e-02  0.0169067772  0.0344134084
## alcohol               1.690678e-02  1.1356473950  0.4097890108
## quality               3.441341e-02  0.4097890108  0.6521684000
```

The codes above computes mean, min, max, variance, median, standard deviation, skewness and Kurtosis for all the variables

e)Create a categorical variable using the variable "sulphates" with the following coding scheme: new_variable1= (if sulphates <=0.4 then 1, if sulphates<=0.7 and >0.4 then 2, if sulphates >0.7 then 3). Construct boxplots of variable "quality" for each category of new_variable1. Similarly create another categorical variable for the variable "pH": new_variable2=(if pH <=3.15 then 1, if pH<=3.45 and >3.15 then 2, if pH >3.45 then 3). Construct boxplots of variable "quality" for each category of new_variable2.

```
data$new_variable1 <- as.factor(ifelse(data$sulphates<=0.4, 1,
                     ifelse(data$sulphates<=0.7 & data$sulphates >0.4,2,
                     ifelse(data$sulphates >0.7, 3,
                            ifelse(data$sulphates, 'NA')))))
```

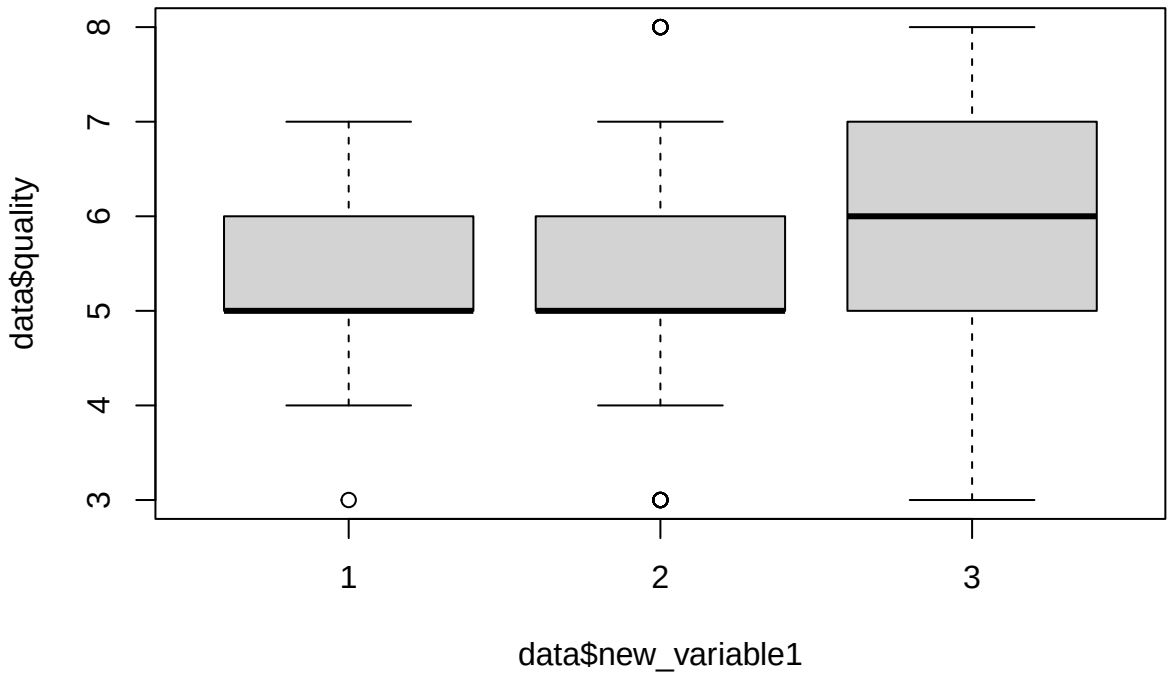
```
head(data);dim(data)
```

```
##    fixed.acidity volatile.acidity citric.acid residual.sugar chlorides
## 1:           7.4             0.70        0.00            1.9     0.076
## 2:           7.8             0.88        0.00            2.6     0.098
## 3:           7.8             0.76        0.04            2.3     0.092
## 4:          11.2             0.28        0.56            1.9     0.075
## 5:           7.4             0.70        0.00            1.9     0.076
## 6:           7.4             0.66        0.00            1.8     0.075
##    free.sulfur.dioxide total.sulfur.dioxide density   pH sulphates alcohol
## 1:                  11                   34  0.9978 3.51      0.56     9.4
## 2:                  25                   67  0.9968 3.20      0.68     9.8
## 3:                  15                   54  0.9970 3.26      0.65     9.8
## 4:                  17                   60  0.9980 3.16      0.58     9.8
## 5:                  11                   34  0.9978 3.51      0.56     9.4
## 6:                  13                   40  0.9978 3.51      0.56     9.4
##    quality new_variable1
## 1:       5             2
## 2:       5             2
## 3:       5             2
## 4:       6             2
## 5:       5             2
## 6:       5             2
```

```
## [1] 1599   13
```

Because I have added a new_variable 1 to the data set, we now have 13 variables and 1599 observations.

```
boxplot(data$quality ~ data$new_variable1)
```



I have Constructed boxplots of the variable "quality" for each category of new_variable1 as shown above.

```
data$new_variable2 <- as.factor(ifelse(data$pH<=3.15, 1,
                      ifelse(data$pH<=3.45 & data$pH >3.15,2,
                      ifelse(data$sulphates >3.45, 3,
                              ifelse(data$sulphates, 3)))))

head(data);dim(data)
```

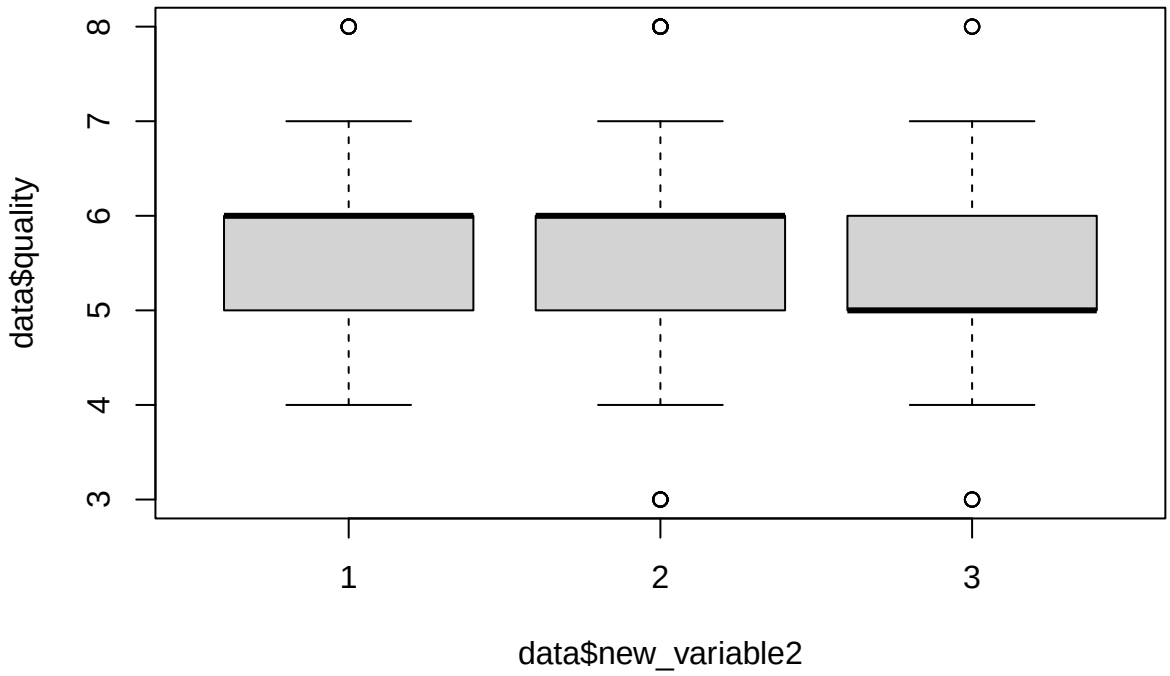
```
##    fixed.acidity volatile.acidity citric.acid residual.sugar chlorides
## 1:           7.4             0.70        0.00            1.9     0.076
## 2:           7.8             0.88        0.00            2.6     0.098
## 3:           7.8             0.76        0.04            2.3     0.092
## 4:          11.2             0.28        0.56            1.9     0.075
## 5:           7.4             0.70        0.00            1.9     0.076
## 6:           7.4             0.66        0.00            1.8     0.075
##    free.sulfur.dioxide total.sulfur.dioxide density   pH sulphates alcohol
## 1:                  11                   34  0.9978 3.51      0.56     9.4
## 2:                  25                   67  0.9968 3.20      0.68     9.8
## 3:                  15                   54  0.9970 3.26      0.65     9.8
## 4:                  17                   60  0.9980 3.16      0.58     9.8
## 5:                  11                   34  0.9978 3.51      0.56     9.4
## 6:                  13                   40  0.9978 3.51      0.56     9.4
```

```
##    quality new_variable1 new_variable2
## 1:       5             2             3
## 2:       5             2             2
## 3:       5             2             2
## 4:       6             2             2
## 5:       5             2             3
## 6:       5             2             3
```

```
## [1] 1599   14
```

After adding new_variable2, we now have 14 variables and 1599 observations.

```
boxplot(data$quality ~ data$new_variable2)
```



I have Constructed boxplots of the variable "quality" for each category of new_variable2 as shown above.

Question (2) a)Please write a program to print all primes smaller than 1000

```
prime = 0:1000
for(val in prime){
  if (val < 2)
    next
  else {
    f = FALSE
    for (temp in 2:sqrt(50))
       if (val %% temp == 0 && val > temp){
```

```
            f = TRUE
            break
        }
    if (f) next
  }
  print(val)
}
```

```
## [1] 2
## [1] 3
## [1] 5
## [1] 7
## [1] 11
## [1] 13
## [1] 17
## [1] 19
## [1] 23
## [1] 29
## [1] 31
## [1] 37
## [1] 41
## [1] 43
## [1] 47
## [1] 53
## [1] 59
## [1] 61
## [1] 67
## [1] 71
## [1] 73
## [1] 79
## [1] 83
## [1] 89
## [1] 97
## [1] 101
## [1] 103
## [1] 107
## [1] 109
## [1] 113
## [1] 121
## [1] 127
## [1] 131
## [1] 137
## [1] 139
## [1] 143
## [1] 149
## [1] 151
## [1] 157
## [1] 163
## [1] 167
## [1] 169
## [1] 173
## [1] 179
## [1] 181
```

```
## [1] 187
## [1] 191
## [1] 193
## [1] 197
## [1] 199
## [1] 209
## [1] 211
## [1] 221
## [1] 223
## [1] 227
## [1] 229
## [1] 233
## [1] 239
## [1] 241
## [1] 247
## [1] 251
## [1] 253
## [1] 257
## [1] 263
## [1] 269
## [1] 271
## [1] 277
## [1] 281
## [1] 283
## [1] 289
## [1] 293
## [1] 299
## [1] 307
## [1] 311
## [1] 313
## [1] 317
## [1] 319
## [1] 323
## [1] 331
## [1] 337
## [1] 341
## [1] 347
## [1] 349
## [1] 353
## [1] 359
## [1] 361
## [1] 367
## [1] 373
## [1] 377
## [1] 379
## [1] 383
## [1] 389
## [1] 391
## [1] 397
## [1] 401
## [1] 403
## [1] 407
## [1] 409
## [1] 419
```

```
## [1] 421
## [1] 431
## [1] 433
## [1] 437
## [1] 439
## [1] 443
## [1] 449
## [1] 451
## [1] 457
## [1] 461
## [1] 463
## [1] 467
## [1] 473
## [1] 479
## [1] 481
## [1] 487
## [1] 491
## [1] 493
## [1] 499
## [1] 503
## [1] 509
## [1] 517
## [1] 521
## [1] 523
## [1] 527
## [1] 529
## [1] 533
## [1] 541
## [1] 547
## [1] 551
## [1] 557
## [1] 559
## [1] 563
## [1] 569
## [1] 571
## [1] 577
## [1] 583
## [1] 587
## [1] 589
## [1] 593
## [1] 599
## [1] 601
## [1] 607
## [1] 611
## [1] 613
## [1] 617
## [1] 619
## [1] 629
## [1] 631
## [1] 641
## [1] 643
## [1] 647
## [1] 649
## [1] 653
```

```
## [1] 659
## [1] 661
## [1] 667
## [1] 671
## [1] 673
## [1] 677
## [1] 683
## [1] 689
## [1] 691
## [1] 697
## [1] 701
## [1] 703
## [1] 709
## [1] 713
## [1] 719
## [1] 727
## [1] 731
## [1] 733
## [1] 737
## [1] 739
## [1] 743
## [1] 751
## [1] 757
## [1] 761
## [1] 767
## [1] 769
## [1] 773
## [1] 779
## [1] 781
## [1] 787
## [1] 793
## [1] 797
## [1] 799
## [1] 803
## [1] 809
## [1] 811
## [1] 817
## [1] 821
## [1] 823
## [1] 827
## [1] 829
## [1] 839
## [1] 841
## [1] 851
## [1] 853
## [1] 857
## [1] 859
## [1] 863
## [1] 869
## [1] 871
## [1] 877
## [1] 881
## [1] 883
## [1] 887
```

```
## [1] 893
## [1] 899
## [1] 901
## [1] 907
## [1] 911
## [1] 913
## [1] 919
## [1] 923
## [1] 929
## [1] 937
## [1] 941
## [1] 943
## [1] 947
## [1] 949
## [1] 953
## [1] 961
## [1] 967
## [1] 971
## [1] 977
## [1] 979
## [1] 983
## [1] 989
## [1] 991
## [1] 997
```

The code above prints all primes smaller than 1000

Question(3) Write a function in R to compute the following: (a) Find an empirical cumulative distribution of a given array of length n at all the data points

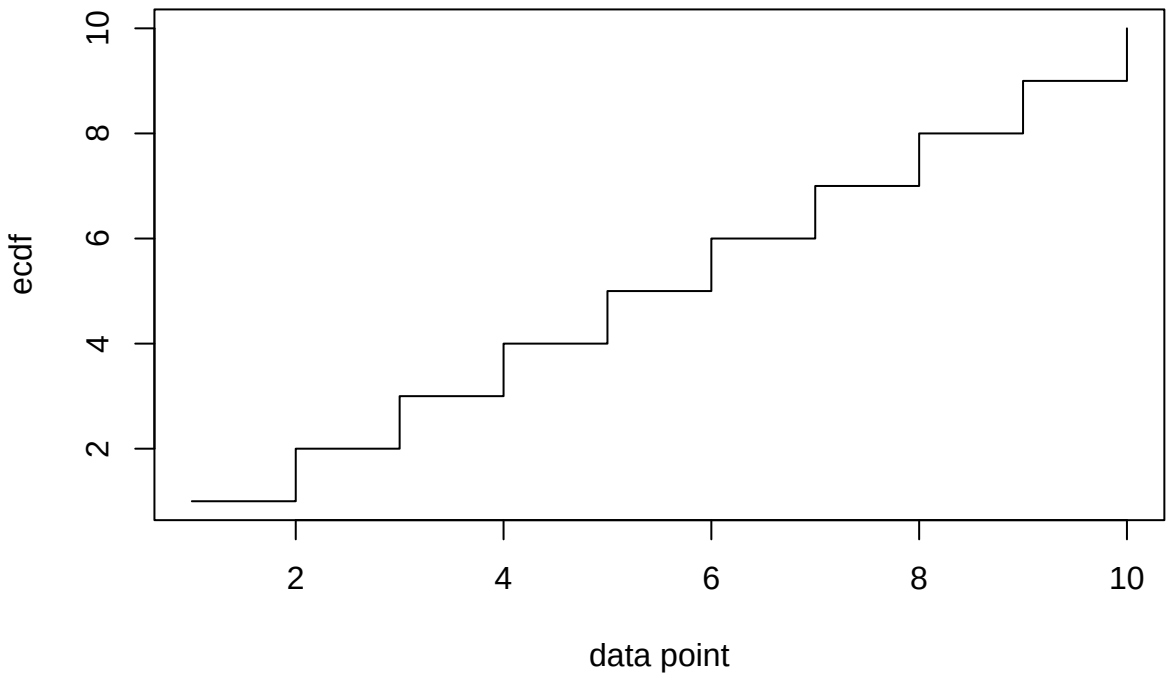
```
x<-rnorm(10)
ecd_func <-function(x){
  n<-length(x)
  x<-sort(x)
  result<-NULL
  for (t in x){
    f=sum(x<=t)/n
    result=c(result,f)
  }
  return(result)
}
ecd_func(1:10)
```

```
##  [1] 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
```

The code above finds an empirical cumulative distribution of a given array of length n at all the data points

b)Plot the empirical cumulative distribution function

```
plot_ecd_func <-function(x="",ecdf=""){
  y<-sort(x)
  plot(y,ecdf,xlab="data point",ylab="ecdf",type="s")
}
plot_ecd_func(1:10,1:10)
```



The code above plots the empirical cumulative distribution function

Question (4)Find bugs in the following the codes are written correctly below. a)

```
x<-0:9
if (x[1]!=1){
  print(x)
}
```

```
##  [1] 0 1 2 3 4 5 6 7 8 9
```

From line 288, the code is x[1]!=1 instead of x[1]=1

b)

```
myfactorial<-function(x){
  if (x==1)
    return(1)
  else
    return(x*myfactorial(x-1))
}
myfactorial(6)
```

```
## [1] 720
```

#the code on line 249 should have been return(x*myfactorial(x-1)) instead of return(x*myfactorial(x))**

c)

```
f<-function(n){
  if (n==1)
  return(1)
  else { if  (n%%2==0)
return(n/2)
else
  return(3*n)
  }

}

x<-1

n<-3
f(x);f(n)
```

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

```
## [1] 9
```

#in line 311, the code should be return(3*n) instead of return(3*x)

d)

```
f<-function(.)
{
  if (runif(1)>0.5)
    x<-1
  else
  return(x)
}
x<-0
sapply(1:10,f)
```

```
##  [1] 0 0 0 1 1 0 1 0 0 0
```

```
x
```

```
## [1] 0
```

in line 332, we should include the else statement as shown above.