MULTIPLE LINEAR REGRESSION

1. SELECTION OF DATASET:

The dataset used here is 'Wine Quality Dataset'. The dataset is related to red variants of the Portuguese "Vinho Verde" wine. This dataset is to model wine quality based on physicochemical tests.

Dimension: 1599 rows and 12 columns

Attributes used: fixed acidity, volatile acidity, citric acid, residual sugar, chlorides, free sulfur dioxide, total sulfur dioxide, density, pH, sulphates, alcohol, quality (score between 0 and 10).

```
> redwine<- read.csv("C:/Users/AISHIKA/Desktop/R datasets/redwine.csv", header=TRUE) #importing the dataset
                      #this command shows the first six rows of the dataset
> head(redwine)
  fixed.acidity volatile.acidity citric.acid residual.sugar chlorides free.sulfur.dioxide total.sulfur.dioxide density pH
                                    0.00
                                                  1.9
           7.4
                         0.70
                                                         0.076
                                                                             11
                                                                                                 34 0.9978 3.51
2
          7.8
                         0.88
                                    0.00
                                                  2.6
                                                         0.098
                                                                                                 67 0.9968 3.20
3
          7.8
                         0.76
                                    0.04
                                                  2.3
                                                         0.092
                                                                             15
                                                                                                 54 0.9970 3.26
                                   0.56
                                                  1.9
                                                                             17
4
          11.2
                         0.28
                                                         0.075
                                                                                                 60 0.9980 3.16
5
          7.4
                         0.70
                                   0.00
                                                                             11
                                                                                                 34 0.9978 3.51
                                                  1.9
                                                         0.076
6
          7.4
                                   0.00
                                                  1.8
                                                                             13
                         0.66
                                                         0.075
                                                                                                 40 0.9978 3.51
  sulphates alcohol quality
      0.56
              9.4
2
      0.68
                       5
              9.8
3
      0.65
             9.8
                       5
4
      0.58
             9.8
5
      0.56
             9.4
                       5
6
      0.56
              9.4
                       5
```

> summary(redwine) #this command gives the summary of the dataset						
fixed.acidity	volatile.acidity	citric.acid	residual.sugar	chlorides	free.sulfur.dioxid	de total.sulfur.dioxide
Min. : 4.60	Min. :0.1200	Min. :0.000	Min. : 0.900	Min. :0.01200	Min. : 1.00	Min. : 6.00
1st Qu.: 7,10	1st Qu.:0.3900	1st Qu.:0.090	1st Qu.: 1.900	1st Qu.:0.07000	1st Qu.: 7.00	1st Qu.: 22.00
Median: 7.90	Median :0.5200	Median :0.260	Median : 2.200	Median :0.07900	Median :14.00	Median : 38.00
Mean : 8.32	Mean :0.5278	Mean :0.271	Mean : 2.539	Mean :0.08747	Mean :15.87	Mean : 46.47
3rd Qu.: 9.20	3rd Qu.:0.6400	3rd Qu.:0.420	3rd Qu.: 2.600	3rd Qu.:0.09000	3rd Qu.:21.00	3rd Qu.: 62.00
Max. :15.90	Max. :1.5800	Max. :1.000	Max. :15.500	Max. :0.61100	Max. :72.00	Max. :289.00
density	pH	sulphates	alcohol	quality		
Min. :0.9901	Min. :2,740	Min. :0.3300	Min. : 8.40	Min. :3.000		
1st Qu.:0.9956	1st Qu.:3.210	1st Qu.:0.5500	1st Qu.: 9.50	1st Qu.:5.000		
Median :0.9968	Median :3.310	Median :0.6200	Median :10.20	Median :6.000		
Mean :0.9967	Mean :3.311	Mean :0.6581	Mean :10.42	Mean :5.636		
3rd Qu.:0.9978	3rd Qu.:3,400	3rd Qu.:0.7300	3rd Qu.:11.10	3rd Qu.:6.000		
Max. :1.0037	Max. :4.010	Max. :2.0000	Max. :14.90	Max. :8.000		

```
#this command gives the structure of the dataset
> str(redwine)
data.frame': 1599 obs. of 12 variables:
$ fixed.acidity : num 7.4 7.8 7.8 11.2 7.4 7.4 7.9 7.3 7.8 7.5 ..
$ volatile.acidity
                       : num 0.7 0.88 0.76 0.28 0.7 0.66 0.6 0.65 0.58 0.5 ...
                       : num 0 0 0.04 0.56 0 0 0.06 0 0.02 0.36 ...
$ citric.acid
$ residual.sugar
                      : num 1.9 2.6 2.3 1.9 1.9 1.8 1.6 1.2 2 6.1 .
$ chlorides : num 0.076 0.098 0.092 0.075 0.076 0.075 0.069 0.065 0.073 0.071 ... $ free.sulfur.dioxide : num 11 25 15 17 11 13 15 15 9 17 ...
$ total.sulfur.dioxide: num 34 67 54 60 34 40 59 21 18 102 ..
                      : num 0.998 0.997 0.997 0.998 0.998 .
$ density
                              3.51 3.2 3.26 3.16 3.51 3.51 3.3 3.39 3.36 3.35 ...
$ pH
                       : num
                      : num 0.56 0.68 0.65 0.58 0.56 0.56 0.46 0.47 0.57 0.8 ...
$ sulphates
$ alcohol
                       : num 9.4 9.8 9.8 9.8 9.4 9.4 9.4 10 9.5 10.5 ...
$ quality
                       : int 5556555775 ...
```

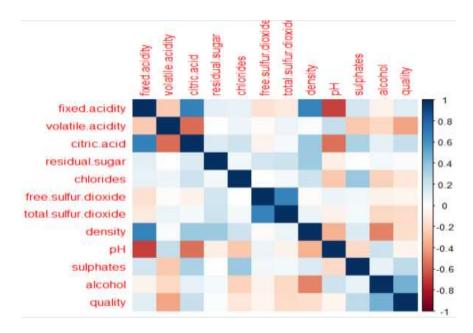
The structure of the data set gives a detailed description of the data types and the levels for the variables which have factors.

```
> any(is.na(redwine)) #to check whether there are any null values in the dataset
[1] FALSE
```

2. CORRELATION ANALYSIS OF THE DATASET:

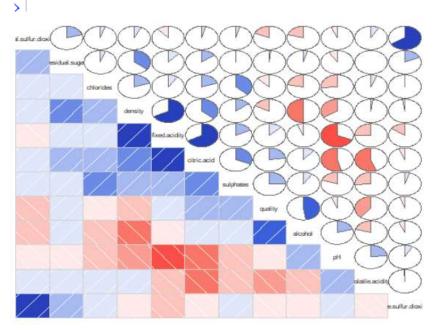
```
#The following libraries are used for Exploratory Data Analysis
  library(ggplot2)
library(ggthemes)
library(dplyr)
> #Grabbing only the numeric columns as we can't see correlation between the categorical variables
> num.cols<- sapply(redwine, is.numeric)
> #Filtering to numeric columns for correlation
> cor.data<-cor(redwine[ , num.cols])
> cor.data
                 fixed.acidity volatile.acidity citric.acid residual.sugar
                                                                    chlorides free.sulfur.dioxide
                               -0.256130895 0.67170343 0.114776724 0.093705186
fixed.acidity
                   1.00000000
                                                                                   -0.153794193
                  -0.25613089
volatile.acidity
                                -0.010503827
citric.acid
                    0.67170343
                               -0.552495685 1.00000000
                                                        0.143577162 0.203822914
                                                                                   -0.060978129
                   0.11477672
                                0.001917882 0.14357716 1.000000000 0.055609535
                                                                                   0.187048995
residual.sugar
chlorides
                   0.09370519
                                0.061297772  0.20382291  0.055609535  1.000000000
                                                                                   0.005562147
                                free.sulfur.dioxide -0.15379419
                                                                                   1 000000000
total.sulfur.dioxide -0.11318144
                                                                                   0.667666450
                                0.022026232 0.36494718 0.355283371 0.200632327
density
                   0.66804729
                                                                                  -0.021945831
                  -0.68297819
                                0.234937294 -0.54190414 -0.085652422 -0.265026131
                                                                                   0.070377499
                               0.18300566
-0.06166827
sulphates
                                                                                   0.051657572
alcohol
                                                                                  -0.069408354
                   0.12405165
                                -0.390557780 0.22637251 0.013731637 -0.128906560
                                                                                   -0.050656057
quality
                 total.sulfur.dioxide
                                      density
                                                    pH
                                                                               quality
                                                         sulphates
                                                                     alcohol
                       -0.11318144  0.66804729  -0.68297819  0.183005664  -0.06166827  0.12405165
fixed.acidity
volatile.acidity
                          0.07647000 0.02202623 0.23493729 -0.260986685 -0.20228803 -0.39055778
                          0.03553302  0.36494718 -0.54190414  0.312770044  0.10990325  0.22637251
citric.acid
                          residual.sugar
                         chlorides
free.sulfur.dioxide
                         0.66766645 -0.02194583 0.07037750 0.051657572 -0.06940835 -0.05065606
                         1.00000000 0.07126948 -0.06649456 0.042946836 -0.20565394 -0.18510029
total.sulfur.dioxide
                         0.07126948 1.00000000 -0.34169933 0.148506412 -0.49617977 -0.17491923
density
                         -0.06649456 -0.34169933 1.00000000 -0.196647602 0.20563251 -0.05773139
sulphates
                         0.04294684  0.14850641 -0.19664760  1.000000000  0.09359475  0.25139708
                         -0.20565394 -0.49617977 0.20563251 0.093594750 1.00000000 0.47616632
alcohol
                         -0.18510029 -0.17491923 -0.05773139 0.251397079 0.47616632 1.00000000
quality
```

- > #For proper data visualization we use the 'corrgram' package and the 'corrplot' package
- > library(corrgram)
- > library(corrplot)
- > #Now we perform the correlation plot and see what we can infer from that
- > corrplot(cor.data, method='color')



We can see from the above correlation plot that white indicates zero correlation, the shades from white to blue indicates positive correlation and from white to red indicates negative correlation. Positive correlation indicates that as the value of one variable increases, the value of the other variable also increases, i.e., directly proportional, while Negative correlation indicates that as the value of one variable increases the value of the other variable will decrease, i.e., indirectly proportional. We can see that citric acid and fixed acidity, density and fixed acidity are highly correlated to each other. This means that higher the citric acid, higher is the fixed acidity and higher the density, higher is the fixed acidity. We are going to predict the fixed acidity based on the variables citric acid and density. Here fixed acidity is a dependent variable and citric acid and density are independent variables.

- > #Now we perform the correlogram and see what we can infer from that
- > corrgram(redwine,order=TRUE, lower.panel=panel.shade, upper.panel=panel.pie, text.panel=panel.txt)



In R, correlograms are implemented using the 'corrgram' function. It shows the graph of the correlation matrix and is very useful for highlighting the most correlated values. The results of this plot can be interpreted in the same way as 'corrplot'. Here blue is positive correlation and pink is negative correlation.

```
> #Histogram of the variable fixed.acidity
> ggplot(redwine,aes(x=fixed.acidity))+ geom_histogram(bins=20, alpha=0.5, fill='blue') + theme_minimal()
```

As we have to predict fixed acidity so we draw a histogram of it. From the graph we can see that the higher amount of fixed acidity is below the mean of fixed acidity and there are much more values or observations beyond the mean value. So we can say that this graph is positively skewed.

3. MULTIPLE LINEAR REGRESSION:

```
> Awe need to split our data into a training set and a testing set in order to test our accuracy, so we can do this using the caTools library
> library(caTools)
> #Spliting up the sample for training and testing and assigns boolean values to a new column
> sample <- sample.split(redwine$fixed.acidity, SplitRatio = 0.70)
> #Training data
> train_data = subset(redwine, sample == TRUE)
> #Testing data
> test_data= subset(redwine, sample == FALSE)
> #Training the model
> model <- lm(fixed.acidity - citric.acid+density, train_data)
> summary(nodel)
ln(formula = fixed.acidity ~ citric.acid + density, data = train_data)
Residuals:
             10 Median
    Min
                              30
-4.9697 -0.5808 -0.0218 0.6785 5.5298
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
                         17.4941 -24.96
0.1712 25.58
                                              <2e-16 ***
(Intercept) -436.6703
citric.acid
               4.3776
                                              <2e-16 ***
                         17,5692 25.34
             445.2637
                                             <2e-16 ***
density
Signif, codes: 0 'nee' 0.001 'en' 0.01 'e' 0.05 '.' 0.1 ' ' 1
Residual standard error: 1.026 on 1117 degrees of freedom
Multiple R-squared: 0.656, Adjusted R-squared: 0.6554
F-statistic: 1065 on 2 and 1117 DF, p-value: < 2.2e-16
```

The residuals are the difference between the actual values of the variable we are predicting and predicted values from our regression.

The stars are for significance levels, with the number of asterisks displayed according to the p-value computed. *** indicates high significance. In this case, *** indicates that there is high significance between citric.acid and density with fixed.acidity.

The estimated coefficient is the value of slope calculated by the regression.

Standard Error of the Coefficient Estimate is the measure of the variability in the estimate for the coefficient.

t-value shows how many standard deviations the coefficient is far away from zero. Further it is away from zero, stronger the relationship between the variables.

p-value shows whether the overall model is significant or not. Coefficients having p-value less than the level of significance are said to be statistically significant. In this case we can say that the variables citric.acid and density are statistically significant for the prediction of fixed.acidity.

R-squared is an overall measure of the strength of association and adjusted R-squared gives a more proper estimate of the R-squared value for the population. Its value shows that 65.54 % of the variance in fixed acidity can be predicted from the variables citric acid and density.

We have to remove the negative values from here.

```
> #Predictions
> fixed.acidity.predictions <- predict(model,test_data)
> results<- cbind(fixed.acidity.predictions,test_data$fixed.acidity)
> colnames(results) <- c('predicted','real')
> results<- as.data.frame(results)
> results
```

```
predicted real
          7.613825
7.253109
                             7.4
7.9
1
12
13
14
16
18
           9.189746
                             7.5
                             5.6
7.8
           6.055401
           8.705211
          8.801772
8.394277
9.445642
7.290127
                             8.9
                             8.1
20
26
45
47
                             6.3
           6.988954
                             6.8
          9.139964
8.833535
54
                             8.6
          7.650092
7.120280
7.819939
67
70
73
75
77
                             8.0
                             7.7
9.7
         10.244865
10.771673
7.079508
10.500010
                             8.8
78
                             6.8
82
                             7.8
          9.315066
8.217673
7.034982
                             7.4
83
94
97
                             6.8
          8.134627
8.081091
100
102
                             7.8
          8.134627
103
                             8.1
          8.134627
7.822191
7.384436
7.428211
7.076505
7.118028
111
                             7.8
112
113
119
                             8.8
120
                             7.0
          8.056085
7.601812
                             9.0
126
129
                             8.0
132
           6.182223
                              5.6
          6.182223
7.733139
                             5.6
7.8
133
140
          6.895396
144
                             6.3
          4.853189
9.576218
6.498413
9.229767
145
                             5.2
146
                             8.1
                             7.5
7.1
153
156
          7.213087
159
                             7.1
                             7.6
161
           6.899150
164
          8.930095
                             7.4
          8.885569
6.765571
7.477993
7.213087
                             7.3
7.3
165
168
                             7.9
7.0
171
179
184
          8.177652
                             6.8
          9.797348
8.569380
186
                             8.9
7.9
190
          8.530109
7.909743
7.909743
192
                             6.4
194
                             7.6
195
                             7.6
          8.481828
8.455061
                             7.3
7.0
197
205
          9.372096
8.241438
7.785924
7.601812
9.033903
211
212
                             8.0
215
                             7.8
217
225
                             8.7
                             8.4
          9.179235
8.493841
8.285214
227
228
                             9.0
230
                             6.9
         8.578389 8.5
10.541794 15.0
233
245
        9.851635 10.8
7.346666 7.1
10.495506 7.7
10.495015 10.0
7.930504 7.9
251
255
259
260
261
263
           6.832361
                            8.0
          9.444891 11.8
8.189664 6.9
266
269
         10.647854 11.5
10.842967 10.9
272
273
277
          8.189664 6.9
```

```
10.647854 11.5
278
             11.358775 11.6
7.918752 8.7
290
291
                9.333085 10.4
293
             9.492421 11.1
10.164821 10.3
10.252372 10.3
8.517606 9.6
302
308
309
318
           9.620254 9.3

10.136812 10.0

9.891657 10.3

10.504515 10.2

10.504515 10.2

8.247935 8.9

10.823707 12.5

9.447895 10.6

8.326246 9.1

7.435719 7.7

11.482594 11.9

11.927857 12.6

11.126383 11.9

12.462174 12.8

7.567797 6.9

8.875058 7.9

9.221509 9.4
                9.620254
322
                                          9.3
326
328
331
332
337
340
342
350
353
359
360
363
367
371
372
            8.875058 /.9
9.221509 9.4
12.193514 13.7
9.274294 8.3
8.885569 7.8
8.003300 8.3
9.979208 9.5
378
382
384
387
388
393
            9.979208 9.5
10.641848 11.5
8.413798 6.6
7.654597 7.7
11.304488 12.0
9.796598 10.4
9.546707 9.1
398
401
405
408
409
412
               9.546/0/ 9.1
7.779917 7.1
9.280300 8.8
7.524772 6.8
7.465230 7.7
9.671277 10.5
9.743813 11.9
413
415
420
423
424
428
433
            11.165653 11.1
10.822957 10.4
7.199573 7.0
13.345193 15.6
438
439
440
443
            11.041835 12.5
11.346762 11.9
8.754993 8.3
10.234354 11.0
10.410957 8.1
447
451
462
463
464
            11.613920 11.4
6.543690 8.7
9.709797 9.9
9.109702 9.6
9.714302 10.4
8.362514 9.4
10.552795 10.6
9.062924 9.3
8.599641 8.9
7.462977 6.5
9.666772 10.7
8.480327 7.2
9.666772 70.7
8.975372 7.8
9.926423 10.2
             11.613920 11.4
469
470
474
476
478
480
483
491
493
495
496
498
499
501
506
                9.926423 10.2
             11.001062 11.2
11.253205 13.3
508
510
             10.192830 10.5
10.471250 10.4
514
518
521
                9.936934
                                           9.8
522
                9.847881
                                            7.6
                9.803\overline{355}
523
                                            8.2
                9.669776
9.224512
                                           9.2
7.3
525
527
                9.198755
531
                                           9.1
```

```
9.644019 10.0
535
                9.198755
8.397280
                                            9.1
536
                                            8.1
538
             10.204092
542
                                            9.5
           9.936934 9.1
10.159566 10.6
10.471250 12.4
6.767073 6.8
8.753491 9.4
12.140989 15.5
10.827461 10.9
10.471250 12.7
9.936934 7.6
11.673462 13.0
9.243282 8.7
8.664439 9.9
9.669776 10.6
10.827461 12.3
10.738409 11.7
8.753491 7.6
10.115039 11.1
5.948330 5.0
10.159566 9.0
10.299151 8.9
11.136894 13.2
8.450065 8.3
9.586729 8.8
10.200338 8.8
10.648605 9.6
10.200338 8.8
10.648605 9.6
10.024485 11.5
8.575386 8.3
9.586729 8.8
10.024485 11.5
8.575386 8.3
8.531610 8.2
8.459826 6.8
8.662937 10.162569 9.9
8.318739 7.8
7.565544 7.3
8.303723 8.7
9.677283 10.7
10.133058 9.8
9.515695 9.7
7.3451655 7.1
546
548
549
551
552
559
561
563
564
565
567
577
580
582
583
586
587
589
593
596
602
605
608
611
617
618
621
622
625
631
641
646
647
649
651
652
656
660
661
                7.697622
               9.449396 10.1
8.961858 10.0
7.818438 6.9
9.588231 9.3
664
669
671
676
                8.139882
678
                                            8.6
               9.059921
7.079508
9.554966
9.890906
683
                                           8.5
7.4
691
692
693
                                            8.6
694
                9.281802
                                            9.0
               9.261602
9.447895
8.928594
8.842544
7.161053
7.345165
704
715
                                            7.5
                                            9.9
722
                                            8.9
725
729
                                            7.5
                                            6.4
            7.345165
10.681119
8.682457
8.843295
8.700060
8.348353
8.205170
8.153991
                                           9.5
731
738
739
                                            9.0
742
                                            9.2
750
                                            7.3
                                            7.6
753
,
754
                                            8.3
             10.038438
755
                                            7.8
               8.703762
8.564979
8.214774
8.751990
761
                                            9.0
762
                                            9.3
767
772
                                            8.8
                                            9.4
774
                8.660685
782
                8.012956
                9.586729
786
                                            9.9
787
                9.586729 9.9
9.243282 10.0
789
                8.953808 10.8
796
```

```
7.345113
801
          7.807436
7.652345
                            8.6
7.7
802
804
          7.476930
                             8.2
806
812
         11.054494 12.9
          8.406029 10.8
6.954834 7.1
6.839066 7.0
818
819
821
          7.960275
7.960275
7.532071
8.369063
7.495001
823
                             6.7
824
                             6.7
828
                             7.1
                             8.8
835
836
                             7.6
          8.781501
845
                             9.9
846
          8.127922
                             6.4
848
          8.269708
                             7.4
          8.751135
8.751135
7.725088
                             9.3
854
857
861
          8.337842
7.725088
863
                             7.5
864
                             7.2
          7.097861
                             6.9
868
          6.596371
                             6.9
872
          9.110996 10.4
875
877
          6.442781
                             7.1
          7.633085
7.796873
7.580404
7.589309
7.824986
7.855664
880
                             7.2
883
                             8.4
894
                             7.2
895
897
                             8.3
903
                             7.4
          7.164056
8.172396
905
                             6.8
907
                             7.2
          7.321838
8.171541
6.731295
915
                             7.3
916
                             8.6
917
                             5.3
          8.391274
919
                             8.4
          8.391274 8.4

9.260237 9.6

8.391274 8.4

6.910204 8.4

9.166784 9.8

9.085135 9.4

7.515116 7.4

7.515116 7.4

7.506210 6.6

8.222022 8.8

9.851583 12.0

7.389379 9.6
921
922
923
924
927
932
934
935
936
938
939
941
          8.683490
                             9.6
          9.605290 10.1
8.058672 8.3
8.058672 8.3
943
948
952
          8.358552 8.3
7.513562 6.7
9.192593 10.1
8.900376 9.1
966
970
973
981
          9.358091 12.2
7.319846 7.7
8.456614 7.1
985
989
992
          7.045284
7.506909
                             6.5
993
996
          8.095898
7.882923
6.788325
8.095200
7.882923
999
                             8.9
1003
                             9.1
1004
                             6.8
1005
1007
                             8.2
                             9.1
          8.336289
7.486794
1009
                             8.9
1013
                             7.7
          8.817122 10.5
5.791633 8.0
9.888549 11.3
1016
1019
1021
          7.135088
7.800835
1023
                             7.0
                             8.6
1026
1029
          7.669300
                             7.2
                             7.7
7.3
          6.713641
6.473199
1030
1032
          7.126933
1034
```

```
1035
        8.219227
        7.707924
5.508426
                      7.1
5.6
1048
1053
1058
        8.726025
1064
        9.501921 10.2
        6.085016
8.789759
1067
                      6.6
1070
1071
                      8.0
        8.328082
                      9.3
1075
        8.988032
                      7.5
                      9.1
7.9
1076
        9.053216
1082
        8.524541
       8.504092
6.846417
9.177837
                      7.2
1085
                      6.8
1086
1087
                      8.5
        9.777338 11.6
9.777338 11.6
1089
1090
       6.667561
8.455061
7.451225
1093
                      7.0
1094
                      9.2
1095
                      6.6
1097
        7.451225
                      6.6
        8.769696
1098
                      8.6
        7.050332
1102
        6.324708
1103
                      6.1
        7.232087
7.698811
1105
                      8.0
1113
                      7.9
        6.418600
1115
                      5.0
        6.994105
6.675507
1117
                      7.0
                      7.9
1121
        5.708096
5.946828
1122
                      6.6
1127
                      5.8
       5.828317 6.3
9.325577 10.0
6.578404 5.9
1128
1129
1132
        7.098508
1133
                      7.4
1135
        8.006095
                      8.5
1136
        8.071331
                      8.0
        8.593196
1140
                      8.2
        8.320032
1142
                      8.2
       6.198532 6.9
7.765757 7.3
9.236420 10.0
1143
1145
1148
        6.563002
1155
                      6.6
1158
        5.997361
1159
        8.570026
                      6.7
       9.264638 10.2
8.476521 8.8
9.152519 8.5
1160
1162
1166
1168
        8.199008
                      8.2
1171
        8.599047
                      9.2
        5.556706
8.252492
8.151478
6.994648
                      7.1
1178
1180
                      8.2
1182
                      9.8
1188
                      7.0
1190
        7.102470
                      8.8
        6.594171
7.922350
1194
                      6.4
1196
                      6.2
1204
        8.881919
                      9.9
        8.532153
7.928846
1209
1210
                      6.2
        6.445732 6.8
8.950701 10.2
1211
1215
        8.512246
7.577140
1217
1218
1224
        9.039003 10.5
        6.908055
8.659235
1227
                      7.5
1228
                      9.0
        8.607253
1230
                      7.6
1232
        6.949631
                      7.8
1233
        8.607253
                      7.6
        5.262081
8.088495
1235
                      7.1
1236
                      6.0
        7.215882
8.799519
1241
                      7.5
1244
                      8.3
        6.682420
1247
                      7.4
        7.888981
7.558787
                      7.4
7.5
1248
1252
1255
        7.199677
```

```
5.822363
1256
       6.086570
1262
                    6.3
                    9.8
       9.036854
1263
1267
       6.870933
1277
       8.411984
                    8.5
1278
       6.972645
                    7.0
       8.167245
7.985281
1279
                    8.0
1280
                    9.8
1282
       7.509756
                    7.1
1283
       6.896950
1286
       9.597783 11.3
       7.527358
7.419306
1287
                    7.1
1292
                    8.2
       7.132940
                    7.5
8.2
1294
1295
       7.419306
       4.991221
6.117738
7.509005
1299
                    5.7
1301
                    5.2
7.6
1305
                    7.5
1311
       8.003947
1312
       6.088666
                    6.5
1314
       7.544626
                    7.0
       5.289548
7.328647
                    5.4
6.7
1317
1325
                    6.7
1328
       7.328647
1331
       8.021757
                    7.4
       6.135549
1338
                    6.0
1342
       6.623837
                    7.5
1349
       6.885792
1350
       6.447234
                    6.9
       6.801192
1355
1357
       7.301181
                    6.1
       8.076691
1358
                   7.3
       9.026448 11.6
6.454637 7.2
1363
1365
1366
       7.696120
                    7.8
       8.462516
                    8.0
1369
       9.147315
1373
                    8.7
       8.279259
1374
1383
       7.833297
                    8.0
       7.437815
1385
                    7.1
       8.694804
1389
                    6.6
       7.828844
1392
                    8.0
       7.895738
1396
                    8.6
1399
       7.604815
                    7.0
       7.820690
7.943007
1402
                    7.9
                    8.1
1409
                    8.2
7.7
       8.924089
1413
       7.553532
1420
1426
       9.055416
                    8.3
       6.901403
8.080340
1429
                    7.8
                    7.6
1432
       6.099928
7.599455
                    5.9
7.2
1433
1450
1451
       7.634326
                    7.2
       8.728924
5.454243
7.758457
7.097319
1454
1457
                    7.6
                    6.0
1461
                    6.4
1462
                    6.2
1469
       8.783106
                    7.3
1470 7.498755
1480 10.454086
                    7.3
                    8.2
      8.480223
7.539475
1484
                    8.1
1486
                    7.0
1494
       8.110810
                    7.7
       8.268853
1502
                    7.8
                    7.5
1505
       9.329822
1512
       7.486148
                    7.1
1513
       6.934020
                    6.4
1514
       6.106477
                    6.4
       8.809175
8.053677
                    6.9
1515
1517
                    6.1
1523
       8.053677
                    6.1
1524
       7.595055
                    6.8
1530
       7.198176
                    7.1
       6.511719
1531
                    6.6
1534
       8.646576
                    8.0
1535
       6.521323
```

```
1539
       6.930267
                   6.2
       8.472068
7.949712
                   8.4
1545
1546
1554
       7.547035
                   7.3
1556
       7.173608
                   7.0
1557
       7.147747
                   6.6
       9.089587
1559
                   6.9
1560
       8.048473
                   7.8
1566
       6.405658
                   6.7
1567
       9.142008
                   6.7
1569
       7.510559
                   7.0
       8.002341
1576
1578
                   7.5
6.2
       7.566942
1579
       7.201825
                   6.8
1582
       6.324708
                   6.2
       8.266496
                   7.2
1586
       7.101667
                   5.9
1598
1599
       8.642718
                   6.0
> #To remove negative predictions and replace it with 0 > to_zero <- function(x){ + if(x <0) [{\{}
       return(0)
     }else{
       return(x)
+
> results$predicted <- sapply(results$predicted,to_zero)</pre>
      predicted real
       7.613825
7.253109
1
                   7.4
12
       9.189746
13
       6.055401
                   5.6
       8.705211
8.801772
                   7.8
14
16
                   8.9
       8.394277
18
                   8.1
20
26
       9.445642
                   7.9
       7.290127
                   6.3
45
47
       6.988954
                   6.8
       9.139964
54
       8.833535
                   8.6
67
       7.650092
70
       7.120280
                   8.0
73
75
       7.819939
                   7.7
9.7
      10.244865
77
      10.771673
                   8.8
                   6.8
78
       7.079508
82
      10.500010
                   7.8
       9.315066
8.217673
83
                   7.4
94
                   7.7
97
       7.034982
                   6.8
100
       8.134627
       8.081091
                   7.8
102
       8.134627
7.822191
                   8.1
103
111
                   7.8
112
       7.384436
113
       7.428211
119
       7.076505
                   8.8
       7.118028
8.056085
120
                   7.0
                   9.0
126
129
       7.601812
                   8.0
132
       6.182223
                    5.6
       6.182223
                    5.6
133
       7.733139
6.895396
140
                   7.8
144
                   6.3
145
       4.853189
146
       9.576218
                   8.1
       6.498413
153
                   7.5
       9.229767
                   7.1
156
159
       7.213087
                   7.1
161
       6.899150
                   7.6
       8.930095
164
                   7.4
       8.885569
                   7.3
165
       6.765571
                   7.3
168
                   7.9
171
       7.477993
179
       7.213087
                   7.0
184
       8.177652
                   6.8
```

```
9.797348
186
                    8.569380
8.530109
                                                       7.9
6.4
190
192
                    7.909743
7.909743
194
                                                       7.6
                                                       7.6
7.3
195
                    8.481828
8.455061
9.372096
197
                                                       7.0
205
211
                    8.241438
7.785924
7.601812
212
                                                       8.0
215
                                                       7.8
217
                                                       8.7
                    9.033903
9.179235
225
227
                                                       8.9
                    8.493841
228
230
                    8.285214
                                                       6.9
               8.578389 8.5
10.541794 15.0
9.851635 10.8
7.346666 7.1
10.495506 7.7
10.495506 7.7
10.495506 7.9
6.832361 8.0
9.444891 11.8
8.189664 6.9
10.647854 11.5
10.842967 10.9
8.189664 6.9
10.647854 11.5
11.358775 11.6
7.918752 8.7
9.333085 10.4
9.492421 11.1
10.164821 10.3
10.252372 10.3
8.517606 9.6
9.620254 9.3
10.136812 10.0
1.36812 10.0
9.891657 10.3
10.504515 10.2
8.247935 8.9
10.823707 12.5
9.447895 10.6
8.326246 9.1
7.435719 7.1
482594 11.9
11.927857 12.6
11.126383 11.9
12.462174 12.8
7.567797 6.9
8.875058 7.9
9.21509 9.4
12.193514 13.7
9.274294 8.3
8.885569 7.8
8.003300 8.3
9.979208 9.5
10.641848 11.5
8.413798 6.6
7.654597 7.8
8.003300 8.3
9.979208 9.5
10.641848 12.5
9.796598 10.4
9.796598 10.4
9.796598 10.4
9.546707 9.1
7.779917 7.1
9.280300 8.8
233
245
251
255
259
260
261
263
266
269
272
273
277
278
290
291
293
302
 308
 309
318
322
326
328
331
332
337
340
 342
350
353
359
360
 363
367
371
372
378
 382
384
 387
388
 393
 398
401
405
408
409
412
413
415
                   7.524772 6.8
7.465230 7.7
9.671277 10.5
9.022151 9.5
9.743813 11.9
420
423
424
428
433
                 11.165653 11.1
10.822957 10.4
7.199573 7.0
438
439
440
                 13.345193 15.6
11.041835 12.5
443
447
                 11.346762 11.9
451
```

```
8.754993 8.3
10.234354 11.0
10.410957 8.1
462
463
464
                11.613920 11.4

6.543690 8.7

9.709797 9.9

9.109702 9.6

9.714302 10.4

8.362514 9.4

10.552795 10.6

9.062924 9.3

8.599641 8.9

7.462977 6.5

9.666772 10.7

8.480327 7.2

9.666772 10.7

8.480327 7.2

9.666772 10.7

8.975372 7.8

9.926423 10.2

11.001062 11.2

11.253205 13.3

10.192830 10.5

10.471250 10.4

9.936934 9.8

9.847881 7.6
                  11.613920 11.4
469
470
474
476
478
480
483
491
493
495
496
498
499
501
 506
 508
 510
 514
 518
 521
                     9.847881
9.803355
                9.803355 8.2

9.669776 9.2

9.224512 7.3

9.198755 9.1

9.644019 10.0

9.198755 9.1

8.397280 8.1

10.204092 9.5

9.936934 9.1

10.159566 10.6

10.471250 12.4

6.767073 6.8

8.753491 9.4

12.140989 15.5

10.827461 10.9

10.471250 12.7

9.936934 9.0

9.447144 7.6

11.673462 13.0
 523
                                                         8.2
 525
 527
 531
 535
 536
 538
 542
 546
 548
 549
 551
552
555
559
 561
 563
 564
                 11.673462 13.0
9.243282 8.7
565
               9.243282 8.7

8.664439 9.9

9.669776 10.6

10.827461 12.3

10.738409 11.7

8.753491 7.6

10.115039 11.1

5.948330 5.0

10.159566 9.0

10.299151 8.9

11.136894 13.2

8.450065 8.3

9.586729 8.8

10.200338 8.8

10.648605 9.6

10.024485 11.5

8.575386 8.3

8.575386 8.3

8.531610 8.2

8.459826 6.8
567
 577
 580
 582
583
 586
587
589
593
596
602
605
608
611
617
618
621
622
                8.531610 8.2
8.459826 6.8
8.662937 8.7
10.162569 9.9
8.318739 7.8
7.565544 7.3
8.303723 8.7
9.677283 10.7
10.133058 9.8
9.515695 9.7
7.345165 7.1
7.697622 7.2
625
631
641
646
647
649
651
652
656
660
661
                     7.697622
664
                      9.449396 10.1
                     8.961858 10.0
669
                     7.818438
671
                                                         6.9
676
                     9.588231
                                                         9.3
                     8.139882
                                                         8.6
```

```
683
             9.059921
            7.079508
9.554966
                                   7.4
9.2
691
692
693
             9.890906
                                   9.0
7.5
694
             9.281802
704
             9.447895
            8.928594
8.842544
                                   9.9
715
                                   8.9
722
          7.161053
7.345165
10.681119
8.682457
8.843295
8.700060
725
729
                                   9.5
731
738
739
                                   9.0
742
                                   9.2
          8.348353
8.205170
8.153991
10.038438
750
                                   7.3
                                   7.6
753
754
755
                                   8.3
                                   7.8
            8.703762
8.564979
8.214774
8.751990
8.660685
761
762
                                   9.3
767
                                   8.8
                                  9.4
7.9
772
774
             8.012956
9.586729
782
         9.586729 9.9

9.586729 9.9

9.243282 10.0

8.953808 10.8

7.345113 7.2

7.807436 8.6

7.652345 7.7

7.476930 8.2

11.054494 12.9

8.406029 10.8

6.954834 7.1

6.839066 7.1

6.839066 7.7

7.960275 6.7

7.960275 6.7

7.532071 7.1

8.369063 8.8

7.495001 7.6

8.781501 9.9
786
                                   9.9
787
789
796
801
802
804
806
812
818
819
821
823
824
828
835
836
845
            8.781501
                                  9.9
            8.127922
8.269708
846
                                   6.4
848
                                   7.4
            8.751135
8.751135
7.725088
854
857
                                   9.3
9.3
7.2
861
            8.337842 7.5
7.725088 7.2
7.097861 6.9
6.596371 6.9
9.110996 10.4
863
864
868
872
875
            9.110996
6.442781
7.633085
7.796873
7.580404
7.589309
7.824986
7.855664
7.164056
8.172396
877
880
                                   7.2
                                  8.4
7.2
883
894
895
                                   8.3
897
903
                                   7.4
905
                                   6.8
907
                                   7.2
            7.321838
8.171541
6.731295
915
916
                                   8.6
917
                                   5.3
            8.391274
9.260237
919
                                   8.4
921
                                   9.6
            8.391274
922
                                   8.4
            6.910204
9.166784
9.085135
7.515116
923
                                   8.4
924
                                   6.8
927
                                   9.4
                                   7.4
932
934
             7.515116
935
             7.506210
                                   6.6
936
             8.222022
                                   8.8
            9.851583 12.0
7.389379 7.2
938
939
             8.683490
                                  9.6
941
```

```
943
        9.605290 10.1
       8.058672
8.058672
                     8.3
948
952
966
        8.358552
                     8.3
970
        7.513562
973
        9.192593 10.1
       8.900376 9.1
9.358091 12.2
981
985
       7.319846
989
                     7.7
992
        8.456614
                      7.1
                     6.5
7.7
993
        7.045284
       7.506909
8.095898
996
999
                     8.9
1003
        7.882923
                     9.1
1004
        6.788325
                     6.8
        8.095200
1005
                     8.2
1007
        7.882923
                     9.1
1009
        8.336289
                     8.9
1013
        7.486794
1016
        8.817122 10.5
        5.791633
1019
                     8.0
       9.888549 11.3
7.135088 7.0
1021
1023
       7.800835
7.669300
1026
                     8.6
1029
                     7.2
       6.713641
                     7.7
1030
       6.473199
7.126933
1032
                     7.3
1034
       8.219227
7.707924
1035
                     8.9
1048
                     7.1
        5.508426
1053
                     5.6
       8.726025 7.6
9.501921 10.2
1058
1064
1067
        6.085016
                     6.6
1070
       8.789759
                     8.0
1071
        8.328082
                     9.3
        8.988032
1075
                     7.5
       9.053216
1076
                     9.1
       8.524541
8.504092
6.846417
1082
                     7.9
1085
                     6.8
1086
1087
       9.177837
                     8.5
       9.777338 11.6
1089
1090
       9.777338 11.6
       6.667561
8.455061
7.451225
7.451225
1093
                     7.0
                     9.2
1094
1095
                     6.6
1097
                     6.6
1098
       8.769696
                     8.6
       7.050332
6.324708
7.232087
7.698811
1102
                     7.4
1103
                     6.1
1105
                     8.0
1113
                     7.9
       6.418600
1115
                     5.0
       6.994105
6.675507
5.708096
5.946828
                     7.0
1117
1121
1122
1127
                     6.6
                      5.8
        5.828317
1128
                     6.3
       9.325577 10.0
6.578404 5.9
1129
1132
       7.098508
8.006095
1133
                     7.4
1135
1136
       8.071331
                     8.0
       8.593196
1140
                     8.2
        8.320032
                     8.2
1142
       6.198532 6.9
7.765757 7.3
9.236420 10.0
1143
1145
1148
       6.563002
5.997361
1155
                     6.6
1158
                     5.1
1159
        8.570026
1160
        9.264638 10.2
       8.476521
9.152519
                     8.8
1162
                     8.5
1166
        8.199008
1168
1171
        8.599047
                     9.2
```

```
5.556706
1178
1180
       8.252492
                    8.2
       8.151478
                    9.8
1182
1188
       6.994648
                    7.0
1190
       7.102470
                    8.8
       6.594171
7.922350
8.881919
1194
                    6.4
1196
                    6.2
                    9.9
1204
       8.532153
7.928846
1209
                    7.2
1210
                    6.2
       6.445732
1211
                    6.8
       8.950701 10.2
8.512246 7.9
1215
1217
       7.577140
1218
                    8.2
1224
       9.039003 10.5
1227
       6.908055
                    7.5
       8.659235
8.607253
                    9.0
1228
1230
       6.949631
                    7.8
1232
1233
       8.607253
                    7.6
       5.262081
1235
                    7.1
       8.088495
1236
                    6.0
1241
       7.215882
       8.799519
1244
                    8.3
1247
       6.682420
                    7.4
1248
       7.888981
                    7.4
       7.558787
7.199677
1252
1255
                    7.5
                    7.8
1256
       5.822363
                    6.1
1262
       6.086570
                    6.3
1263
       9.036854
                    9.8
       6.870933
1267
1277
       8.411984
                    8.5
       6.972645
8.167245
7.985281
1278
                    7.0
1279
                    8.0
1280
                    9.8
       7.509756
1282
                    7.1
       6.896950
1283
                    7.9
       9.597783 11.3
7.527358 7.1
7.419306 8.2
1286
1287
1292
1294
       7.132940
       7.419306
1295
                    8.2
1299
       4.991221
                    5.7
       6.117738
7.509005
                    5.2
7.6
1301
1305
1311
       8.003947
1312
       6.088666
                    6.5
1314
       7.544626
                    7.0
       5.289548
7.328647
1317
                    5.4
                    6.7
1325
       7.328647
8.021757
1328
                    6.7
1331
                    7.4
1338
       6.135549
                    6.0
       6.623837
1342
                    7.5
7.2
       6.885792
1349
       6.447234
1350
                    6.9
1355
       6.801192
1357
       7.301181
                    6.1
       8.076691 7.3
9.026448 11.6
1358
1363
       6.454637
7.696120
1365
                    7.2
1366
                    7.8
1369
       8.462516
                    8.0
       9.147315
8.279259
                    8.7
1373
1374
1383
       7.833297
                    8.0
1385
       7.437815
                    7.1
1389
       8.694804
                    6.6
       7.828844
7.895738
1392
                    8.0
1396
                    8.6
1399
       7.604815
                    7.0
       7.820690
1402
                    7.9
       7,943007
                    8.1
1409
       8.924089
7.553532
                    8.2
7.7
1413
1420
1426
       9.055416
                    8.3
```

```
1429 6.901403
      8.080340
1432
      6.099928
1433
                  7.2
7.2
7.6
      7.599455
1450
1451
       7.634326
      8.728924
1454
      5.454243
7.758457
                  6.0
1457
1461
1462
      7.097319
1469
      8.783106
      7.498755
1470
1480 10.454086
1484
      8.480223
1486
      7.539475
                  7.7
7.8
1494
      8.110810
      8.268853
1502
      9.329822
1505
      7.486148
1512
1513
      6.934020
1514
      6.106477
      8.809175
1515
      8.053677
1517
                  6.1
1523
      8.053677
      7.595055
                  6.8
1530
      7.198176
1531
      6.511719
                  6.6
      8.646576
                  8.0
1534
       6.521323
1539
       6.930267
1545
      8.472068
      7.949712
                  6.5
1546
      7.547035
                  7.3
7.0
1554
1556
1557
      7.173608
       7.147747
                  6.6
1559
       9.089587
                  7.8
1560
      8.048473
      6.405658
1566
1567
      9.142008
1569
      7.510559
                  7.5
6.2
1576
      8.002341
1578
      7.566942
1579
      7.201825
                  6.8
1582
       6.324708
                  7.2
1586
      8.266496
1598
      7.101667
1599
      8.642718
 > #Evaluating the prediction values by the method of MSE(Mean Squared Error)
 > mse <- mean((results$real - results$predicted)^2)
 > print(mse)
 [1] 1.008759
 > #Evaluating the prediction values by the method of RMSE(Root Mean Squared Error)
 > mse^0.5
 [1] 1.00437
 > #Or we can just use the R-Squared Value for the model which gives the accuracy of the model
 > SSE <- sum((results$predicted - results$real)^2)</pre>
 > TSS <- sum( (mean(redwine$fixed.acidity) - results$real)^2)
 > R2 <- 1-SSE/TSS
 > R2
 [1] 0.6607791
```

R-Squared (Coefficient of Determination) - This value lies between 0 and 1, and the higher it is, the better the model fits the data set.

Our main aim is to find those variables who give the lowest RMSE value and the highest R-Squared value.

The R-Squared for the training set is 66.07%. It means that the model can explain more than 66.07% of the variation.