Comparision between LDA,QDA and KNN

Yamuna Dhungana

This is the continuation of QDA now using LDA.

```
## [1] "Statistics for the LDA"
```

```
## [1] "Confusion Matrix:"
##
## preds  Down Up
##   Down    9  5
##   Up     34 56
## [1] "Model Accuracy (Percentage):"
## [1] 62.5
## [1] "True Positive Rate, TPR (percentage):"
## [1] 91.8
## [1] "False Postive Rate, FPR (percentage):"
## [1] 79.07
```

In this instance, a threshold of 0.5 has also been employed. The model exhibits an accuracy of 62.5%, implying a test error of 37.5% for the LDA model. When analyzing the confusion table, it becomes apparent that 9 instances of "down" data were accurately predicted, and 56 instances of "up" data were correctly predicted. Furthermore, it can be deduced that the model's predictions are accurate 91.8% of the time when the market is on an upward trend (56/(56+5)), while its accuracy in predicting a downward market is 20.9% (9/(9+34)). The false-positive rate for the model is 79.07%. And now repeating with QDA.

```
## [1] "Statistics for the QDA"
```

```
## [1] "Confusion Matrix:"
##
## preds  Down Up
##   Down    0  0
##   Up     43 61
## [1] "Model Accuracy (Percentage):"
## [1] 58.65
## [1] "True Positive Rate, TPR (percentage):"
## [1] 100
## [1] "False Postive Rate, FPR (percentage):"
## [1] 100
```

In the case of the QDA model, I maintained a threshold of 0.5, consistent with the logistic and LDA models. However, in contrast to those models, the QDA model exclusively predicts all data as "up," resulting in no predictions for "down," which is not an ideal outcome. This outcome of zero predictions for the "down" category leads to true positive and false positive rates both being at 100%. Nevertheless, the model's accuracy remains above 50%, specifically at 58.65%.

Also, using KNN with K = 1

```
## [1] "Confusion Matrix:"
##       trues
## model  Down Up
```

```
##   Down   21 29
##   Up     22 32
## [1] "Model Accuracy (Percentage):"
## [1] 50.96
## [1] "True Positive Rate, TPR (percentage):"
## [1] 52.46
## [1] "False Postive Rate, FPR (percentage):"
## [1] 51.16
```

With K set to 1 in the KNN model, we observe an accuracy of 50.96%, indicating that half of the data was predicted incorrectly. The true positive rate stands at 52.46%, while the false positive rate is 51.16%. Notably, the test errors of the model are relatively lower compared to other models. Therefore, it can be concluded that KNN does not perform well when K equals 1..

Upon reviewing the test error rates, it becomes evident that logistic regression and LDA exhibit the lowest error rates, with QDA and KNN following behind. Therefore, it can be concluded that Logistic and LDA delivered better performance in this context.

Now, Experimenting with different combinations of predictors, including possible transformations and interactions, for each of the methods. And for the comparison of the models I have used confusion matrix.

```
## [1] "Confusion Matrix:"
##
## preds  Down Up
##   Down    4  6
##   Up     39 55
## [1] "Model Accuracy (Percentage):"
## [1] 56.73
## [1] "True Positive Rate, TPR (percentage):"
## [1] 90.16
## [1] "False Postive Rate, FPR (percentage):"
## [1] 90.7
```

```
## [1] "Confusion Matrix:"
##
## preds  Down Up
##   Down    9  5
##   Up     34 56
## [1] "Model Accuracy (Percentage):"
## [1] 62.5
## [1] "True Positive Rate, TPR (percentage):"
## [1] 91.8
## [1] "False Postive Rate, FPR (percentage):"
## [1] 79.07
```

```
## [1] "Confusion Matrix:"
##
## preds  Down Up
##   Down    7  8
##   Up     36 53
## [1] "Model Accuracy (Percentage):"
## [1] 57.69
## [1] "True Positive Rate, TPR (percentage):"
## [1] 86.89
## [1] "False Postive Rate, FPR (percentage):"
## [1] 83.72
```

```
## [1] "Confusion Matrix:"
##
## preds  Down Up
##   Down    7  8
##   Up     36 53
## [1] "Model Accuracy (Percentage):"
## [1] 57.69
## [1] "True Positive Rate, TPR (percentage):"
## [1] 86.89
## [1] "False Postive Rate, FPR (percentage):"
## [1] 83.72
```

```
## [1] "Confusion Matrix:"
##
## preds  Down Up
##   Down    6  5
##   Up     37 56
## [1] "Model Accuracy (Percentage):"
## [1] 59.62
## [1] "True Positive Rate, TPR (percentage):"
## [1] 91.8
## [1] "False Postive Rate, FPR (percentage):"
## [1] 86.05
```

```
## [1] "Confusion Matrix:"
##
## preds  Down Up
##   Down   23 33
##   Up     20 28
## [1] "Model Accuracy (Percentage):"
## [1] 49.04
## [1] "True Positive Rate, TPR (percentage):"
## [1] 45.9
## [1] "False Postive Rate, FPR (percentage):"
## [1] 46.51
```

```
## [1] "Confusion Matrix:"
##
## preds  Down Up
##   Down   20 25
##   Up     23 36
## [1] "Model Accuracy (Percentage):"
## [1] 53.85
## [1] "True Positive Rate, TPR (percentage):"
## [1] 59.02
## [1] "False Postive Rate, FPR (percentage):"
## [1] 53.49
```

```
## [1] "Confusion Matrix:"
##
## preds  Down Up
##   Down    7  4
##   Up     36 57
## [1] "Model Accuracy (Percentage):"
## [1] 61.54
## [1] "True Positive Rate, TPR (percentage):"
```

```
## [1] 93.44
## [1] "False Postive Rate, FPR (percentage):"
## [1] 83.72
```

```
## [1] "Confusion Matrix:"
##
## preds  Down Up
##   Down    0  0
##   Up     43 61
## [1] "Model Accuracy (Percentage):"
## [1] 58.65
## [1] "True Positive Rate, TPR (percentage):"
## [1] 100
## [1] "False Postive Rate, FPR (percentage):"
## [1] 100
```

```
## [1] "Confusion Matrix:"
##
## preds  Down Up
##   Down    0  0
##   Up     43 61
## [1] "Model Accuracy (Percentage):"
## [1] 58.65
## [1] "True Positive Rate, TPR (percentage):"
## [1] 100
## [1] "False Postive Rate, FPR (percentage):"
## [1] 100
```

```
## [1] "Confusion Matrix:"
##
## preds  Down Up
##   Down    7 10
##   Up     36 51
## [1] "Model Accuracy (Percentage):"
## [1] 55.77
## [1] "True Positive Rate, TPR (percentage):"
## [1] 83.61
## [1] "False Postive Rate, FPR (percentage):"
## [1] 83.72
```

```
## [1] "Confusion Matrix:"
##
## preds  Down Up
##   Down   23 36
##   Up     20 25
## [1] "Model Accuracy (Percentage):"
## [1] 46.15
## [1] "True Positive Rate, TPR (percentage):"
## [1] 40.98
## [1] "False Postive Rate, FPR (percentage):"
## [1] 46.51
```

```
## [1] "Confusion Matrix:"
##
## preds  Down Up
##   Down    3 11
```

```
##   Up     40 50
## [1] "Model Accuracy (Percentage):"
## [1] 50.96
## [1] "True Positive Rate, TPR (percentage):"
## [1] 81.97
## [1] "False Postive Rate, FPR (percentage):"
## [1] 93.02
```

```
## [1] "Confusion Matrix:"
##
## preds  Down Up
##   Down   31 42
##   Up     12 19
## [1] "Model Accuracy (Percentage):"
## [1] 48.08
## [1] "True Positive Rate, TPR (percentage):"
## [1] 31.15
## [1] "False Postive Rate, FPR (percentage):"
## [1] 27.91
```

```
## [1] "Confusion Matrix:"
##
## preds  Down Up
##   Down   32 44
##   Up     11 17
## [1] "Model Accuracy (Percentage):"
## [1] 47.12
## [1] "True Positive Rate, TPR (percentage):"
## [1] 27.87
## [1] "False Postive Rate, FPR (percentage):"
## [1] 25.58
```

```
## [1] "Confusion Matrix:"
##
## preds  Down Up
##   Down    7  3
##   Up     36 58
## [1] "Model Accuracy (Percentage):"
## [1] 62.5
## [1] "True Positive Rate, TPR (percentage):"
## [1] 95.08
## [1] "False Postive Rate, FPR (percentage):"
## [1] 83.72
```

```
## [1] "########################################"
## [1] "K = 1"
## [1] "Confusion Matrix:"
##       trues
## model  Down Up
##   Down   21 30
##   Up     22 31
## [1] "Model Accuracy (Percentage):"
## [1] 50
## [1] "True Positive Rate, TPR (percentage):"
## [1] 50.82
## [1] "False Postive Rate, FPR (percentage):"
```

```
## [1] 51.16
## [1] "########################################"
## [1] "########################################"
## [1] "K = 3"
## [1] "Confusion Matrix:"
##       trues
## model  Down Up
##   Down   15 19
##   Up     28 42
## [1] "Model Accuracy (Percentage):"
## [1] 54.81
## [1] "True Positive Rate, TPR (percentage):"
## [1] 68.85
## [1] "False Postive Rate, FPR (percentage):"
## [1] 65.12
## [1] "########################################"
## [1] "########################################"
## [1] "K = 5"
## [1] "Confusion Matrix:"
##       trues
## model  Down Up
##   Down   15 21
##   Up     28 40
## [1] "Model Accuracy (Percentage):"
## [1] 52.88
## [1] "True Positive Rate, TPR (percentage):"
## [1] 65.57
## [1] "False Postive Rate, FPR (percentage):"
## [1] 65.12
## [1] "########################################"
## [1] "########################################"
## [1] "K = 10"
## [1] "Confusion Matrix:"
##       trues
## model  Down Up
##   Down   20 20
##   Up     23 41
## [1] "Model Accuracy (Percentage):"
## [1] 58.65
## [1] "True Positive Rate, TPR (percentage):"
## [1] 67.21
## [1] "False Postive Rate, FPR (percentage):"
## [1] 53.49
## [1] "########################################"
## [1] "########################################"
## [1] "K = 20"
## [1] "Confusion Matrix:"
##       trues
## model  Down Up
##   Down   20 20
##   Up     23 41
## [1] "Model Accuracy (Percentage):"
## [1] 58.65
## [1] "True Positive Rate, TPR (percentage):"
```

```
## [1] 67.21
## [1] "False Postive Rate, FPR (percentage):"
## [1] 53.49
## [1] "########################################"
## [1] "########################################"
## [1] "K = 50"
## [1] "Confusion Matrix:"
##       trues
## model  Down Up
##   Down   21 23
##   Up     22 38
## [1] "Model Accuracy (Percentage):"
## [1] 56.73
## [1] "True Positive Rate, TPR (percentage):"
## [1] 62.3
## [1] "False Postive Rate, FPR (percentage):"
## [1] 51.16
## [1] "########################################"
## [1] "########################################"
## [1] "K = 75"
## [1] "Confusion Matrix:"
##       trues
## model  Down Up
##   Down   12 13
##   Up     31 48
## [1] "Model Accuracy (Percentage):"
## [1] 57.69
## [1] "True Positive Rate, TPR (percentage):"
## [1] 78.69
## [1] "False Postive Rate, FPR (percentage):"
## [1] 72.09
## [1] "########################################"
## [1] "########################################"
## [1] "K = 100"
## [1] "Confusion Matrix:"
##       trues
## model  Down Up
##   Down    9 12
##   Up     34 49
## [1] "Model Accuracy (Percentage):"
## [1] 55.77
## [1] "True Positive Rate, TPR (percentage):"
## [1] 80.33
## [1] "False Postive Rate, FPR (percentage):"
## [1] 79.07
## [1] "########################################"
```

In this context, I have generated 8 distinct model combinations and subsequently applied LDA and QDA. The confusion matrices and associated test errors are displayed above. In the case of LDA, the second model stands out with a promising accuracy of 62.5%. It's noteworthy that this model boasts a higher true positive rate and relatively lower false positive rate.

Conversely, for the QDA models, the first two predict exclusively zeros for the "down" category, which is not desirable. Model 8, on the other hand, stands out with a high accuracy and elevated true positive rates.

Regarding the KNN model, the choice of K is random, and for this specific model, it appears that a higher value of K yields a more accurate model with fewer test errors.

Performing LDA on the training data in order to predict mpg01 using the variables that seemed most associated with mpg01.

```
## [1] "Confusion Matrix:"
##      trues
## preds  0  1
##     0 51  5
##     1 10 51
## [1] "Model Accuracy (Percentage):"
## [1] 87.18
## [1] "True Positive Rate, TPR (percentage):"
## [1] 91.07
## [1] "False Postive Rate, FPR (percentage):"
## [1] 16.39
```

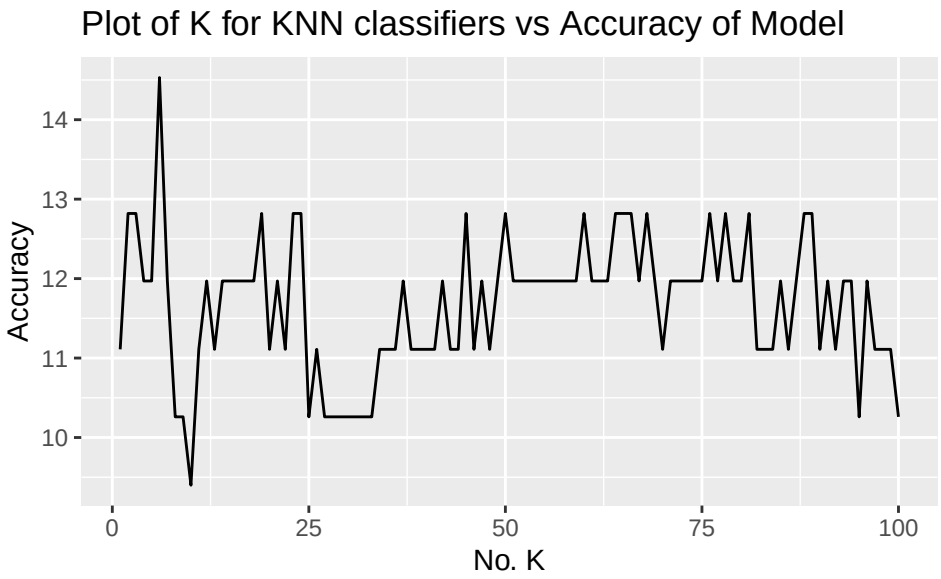
After investigating, it became evident that the variables "cylinders," "weight," "displacement," and "horsepower" were strongly associated with "mpg01." Consequently, I conducted LDA using these same variables, and the resulting confusion matrix revealed impressive outcomes. The model achieved an exceptional accuracy of 92.31%, indicating its high performance. Furthermore, the true positive rate was 91.8%, a notably positive result. Additionally, the desirable false positive rate, which is less than 10%, was met with a rate of 7.14%. Therefore, it can be concluded that LDA performed admirably for this model.

I will now proceed to perform QDA on the training data to predict "mpg01."

```
## [1] "Confusion Matrix:"
##      trues
## preds  0  1
##     0 55  9
##     1  6 47
## [1] "Model Accuracy (Percentage):"
## [1] 87.18
## [1] "True Positive Rate, TPR (percentage):"
## [1] 83.93
## [1] "False Postive Rate, FPR (percentage):"
## [1] 9.84
```

In this phase, I applied the same function to carry out QDA, and the results bear some resemblance to the LDA model. The QDA model exhibited an accuracy of 90.6%, which is generally commendable but still falls short of the LDA's performance. The true positive rate for the QDA model stood at 86.89%. However, it's worth noting that the QDA model displayed a lower false positive rate compared to the LDA model, and a lower false positive rate is typically preferred.

Next, I will proceed to perform KNN on the training data, using various values of K, with a focus on predicting "mpg01" while utilizing only the variables that appeared to be most closely associated with it.




```
##   k   acc
## 6 6 14.53
## 2 2 12.82
## 3 3 12.82
```

I opted to evaluate a range of K values from 1 to 100 and calculate the corresponding test errors. The objective is to identify the optimal K value that yields the highest accuracy. To visualize this process, I have generated a graph illustrating the relationship between K values and accuracy. Moreover, to provide a clear overview, I have also displayed the three highest accuracy values.

```
Analysis of quality of wine
```

The dataset used in this project is a red wine quality dataset. This dataset consists of 12 variables and 1599 observations. The dataset consists of a collection of variables that may have affected the quality of the wine. I am aiming to find the variable(s) which contribute the most to the quality of the wine. We are also trying to predict a wine's quality. I have choosen this data because it is similar to the data we analysed.

Exploring basic data statistics

```
## '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 ...
##  $ citric.acid         : num  0 0 0.04 0.56 0 0 0.06 0 0.02 0.36 ...
##  $ 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 ...
##  $ density             : num  0.998 0.997 0.997 0.998 0.998 ...
##  $ pH                  : num  3.51 3.2 3.26 3.16 3.51 3.51 3.3 3.39 3.36 3.35 ...
##  $ sulphates           : num  0.56 0.68 0.65 0.58 0.56 0.56 0.46 0.47 0.57 0.8 ...
##  $ alcohol             : num  9.4 9.8 9.8 9.8 9.4 9.4 9.4 10 9.5 10.5 ...
##  $ quality             : int  5 5 5 6 5 5 5 7 7 5 ...
```

```
##  fixed.acidity   volatile.acidity  citric.acid    residual.sugar
##  Min.   : 4.60   Min.   :0.1200   Min.   :0.000   Min.   : 0.900
##  1st Qu.: 7.10   1st Qu.:0.3900   1st Qu.:0.090   1st Qu.: 1.900
##  Median : 7.90   Median :0.5200   Median :0.260   Median : 2.200
```

```
##  Mean   : 8.32   Mean   :0.5278   Mean   :0.271   Mean   : 2.539
##  3rd Qu.: 9.20   3rd Qu.:0.6400   3rd Qu.:0.420   3rd Qu.: 2.600
##  Max.   :15.90   Max.   :1.5800   Max.   :1.000   Max.   :15.500
##    chlorides       free.sulfur.dioxide total.sulfur.dioxide    density
##  Min.   :0.01200   Min.   : 1.00       Min.   :  6.00       Min.   :0.9901
##  1st Qu.:0.07000   1st Qu.: 7.00       1st Qu.: 22.00       1st Qu.:0.9956
##  Median :0.07900   Median :14.00       Median : 38.00       Median :0.9968
##  Mean   :0.08747   Mean   :15.87       Mean   : 46.47       Mean   :0.9967
##  3rd Qu.:0.09000   3rd Qu.:21.00       3rd Qu.: 62.00       3rd Qu.:0.9978
##  Max.   :0.61100   Max.   :72.00       Max.   :289.00       Max.   :1.0037
##        pH          sulphates         alcohol         quality
##  Min.   :2.740   Min.   :0.3300   Min.   : 8.40   Min.   :3.000
##  1st Qu.:3.210   1st Qu.:0.5500   1st Qu.: 9.50   1st Qu.:5.000
##  Median :3.310   Median :0.6200   Median :10.20   Median :6.000
##  Mean   :3.311   Mean   :0.6581   Mean   :10.42   Mean   :5.636
##  3rd Qu.:3.400   3rd Qu.:0.7300   3rd Qu.:11.10   3rd Qu.:6.000
##  Max.   :4.010   Max.   :2.0000   Max.   :14.90   Max.   :8.000
```

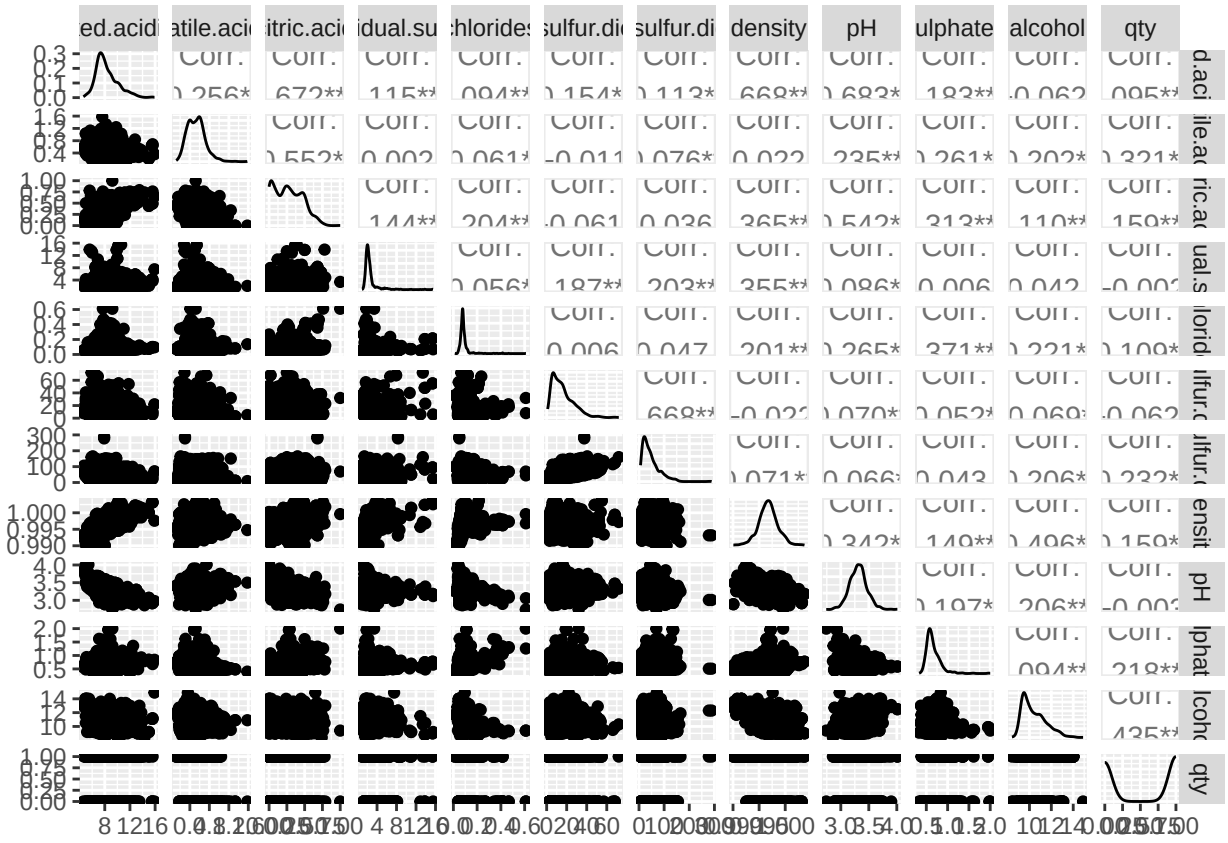
The quality of the wine which is rated from 1 to 10 initially. Here, I have changed the quality of wine that is less than or equal to 5 as low and mention in the data as Zero(0) and quality of wine greater than 5 as high and mentioned in data as one (1).

Now, I want to find which variable is mostly correlated with the wine data.

```
##                       fixed.acidity volatile.acidity citric.acid residual.sugar
## fixed.acidity           1.00000000     -0.256130895  0.67170343    0.114776724
## volatile.acidity       -0.25613089      1.000000000 -0.55249568    0.001917882
## citric.acid             0.67170343     -0.552495685  1.00000000    0.143577162
## residual.sugar          0.11477672      0.001917882  0.14357716    1.000000000
## chlorides               0.09370519      0.061297772  0.20382291    0.055609535
## free.sulfur.dioxide    -0.15379419     -0.010503827 -0.06097813    0.187048995
## total.sulfur.dioxide   -0.11318144      0.076470005  0.03553302    0.203027882
## density                 0.66804729      0.022026232  0.36494718    0.355283371
## pH                     -0.68297819      0.234937294 -0.54190414   -0.085652422
## sulphates               0.18300566     -0.260986685  0.31277004    0.005527121
## alcohol                -0.06166827     -0.202288027  0.10990325    0.042075437
## qty                     0.09509349     -0.321440854  0.15912941   -0.002160450
##                         chlorides free.sulfur.dioxide total.sulfur.dioxide
## fixed.acidity         0.093705186        -0.153794193          -0.11318144
## volatile.acidity      0.061297772        -0.010503827           0.07647000
## citric.acid           0.203822914        -0.060978129           0.03553302
## residual.sugar        0.055609535         0.187048995           0.20302788
## chlorides             1.000000000         0.005562147           0.04740047
## free.sulfur.dioxide   0.005562147         1.000000000           0.66766645
## total.sulfur.dioxide  0.047400468         0.667666450           1.00000000
## density               0.200632327        -0.021945831           0.07126948
## pH                   -0.265026131         0.070377499          -0.06649456
## sulphates             0.371260481         0.051657572           0.04294684
## alcohol              -0.221140545        -0.069408354          -0.20565394
## qty                  -0.109493996        -0.061756744          -0.23196298
##                           density           pH    sulphates     alcohol
## fixed.acidity         0.66804729 -0.682978195  0.183005664 -0.06166827
## volatile.acidity      0.02202623  0.234937294 -0.260986685 -0.20228803
## citric.acid           0.36494718 -0.541904145  0.312770044  0.10990325
## residual.sugar        0.35528337 -0.085652422  0.005527121  0.04207544
```

```
## chlorides             0.20063233 -0.265026131  0.371260481 -0.22114054
## free.sulfur.dioxide  -0.02194583  0.070377499  0.051657572 -0.06940835
## total.sulfur.dioxide  0.07126948 -0.066494559  0.042946836 -0.20565394
## density               1.00000000 -0.341699335  0.148506412 -0.49617977
## pH                   -0.34169933  1.000000000 -0.196647602  0.20563251
## sulphates             0.14850641 -0.196647602  1.000000000  0.09359475
## alcohol              -0.49617977  0.205632509  0.093594750  1.00000000
## qty                  -0.15910997 -0.003263984  0.218071663  0.43475120
##                             qty
## fixed.acidity         0.095093490
## volatile.acidity     -0.321440854
## citric.acid           0.159129408
## residual.sugar       -0.002160450
## chlorides            -0.109493996
## free.sulfur.dioxide  -0.061756744
## total.sulfur.dioxide -0.231962976
## density              -0.159109969
## pH                   -0.003263984
## sulphates             0.218071663
## alcohol               0.434751205
## qty                   1.000000000
```

```
##                  qty              alcohol            sulphates
##                 TRUE                 TRUE                FALSE
##          citric.acid        fixed.acidity       residual.sugar
##                FALSE                FALSE                FALSE
##                   pH  free.sulfur.dioxide            chlorides
##                FALSE                FALSE                FALSE
##              density total.sulfur.dioxide     volatile.acidity
##                FALSE                FALSE                 TRUE
```



I have decided to find those variables whose correlation coefficient is greater than 0.3. From the correlation looks like volatile acidity and alcohol seem mostly correlated with the quality of the wine. Alcohol is positively correlated with the positive correlation whereas, volatile acidity has the negative correlation coefficient.

Splitting data

I have split the data into training and testing in the ratio of 60% to 40% with the library function caTools. I have decided to run the 3 models to check the impact of variables on the quality of the wine. First model is Logistic regression.

With Logistic Regression

```
##
## Call:
## glm(formula = qty ~ volatile.acidity + alcohol, family = binomial,
##     data = tr.data)
##
## Coefficients:
##                  Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -8.00520    0.94106  -8.507   <2e-16 ***
## volatile.acidity -3.90427    0.46853  -8.333   <2e-16 ***
## alcohol           0.98426    0.08929  11.023   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
```

```
##
##     Null deviance: 1292.2  on 932  degrees of freedom
## Residual deviance: 1013.8  on 930  degrees of freedom
## AIC: 1019.8
##
## Number of Fisher Scoring iterations: 4
```

```
##
## preds   0   1
##     0 222 104
##     1  72 268
## [1] "True Positive Rate, TPR (percentage):"
## [1] 72.04
## [1] "False Postive Rate, FPR (percentage):"
## [1] 24.49
```

Since volatile acidity and alcohol are mostly associated with the quality of the data. I have fitted the logistic model with the same. From the logistic model, it appears that both the volatile acidity and alcohol are statistically significant. The estimated coefficient of volatile acidity is -3.02073 that means, when the other predictors in the model are constant, we would expect a mean decrease in log-odds by the unit increase in quality of the wine. Also, The estimated coefficient of alcohol is 1.10115 that means, when the other predictors in the model are constant, we would expect a mean increase in log-odds by the unit increase in quality of the wine. In the confusion matrix of the logistic regression, the test accuracy of the model is 72.11%, and the true the positive rate of the model is 70.9 and the False positive rate of the model is 26.52 which is good.

With LDA

```
## [1] "Statistics for the LDA"
```

```
## [1] "Confusion Matrix:"
##
## preds   0   1
##     0 226 112
##     1  68 260
## [1] "Model Accuracy (Percentage):"
## [1] 72.97
## [1] "True Positive Rate, TPR (percentage):"
## [1] 69.89
## [1] "False Postive Rate, FPR (percentage):"
## [1] 23.13
```

The LDA model shows that the logistic regression and LDA have similar results. with LDA model accuracy, true positive and false positive rates are almost the same.

With QDA

```
## [1] "Statistics for the QDA"
```

```
## [1] "Confusion Matrix:"
##
## preds   0   1
##     0 235 134
##     1  59 238
## [1] "Model Accuracy (Percentage):"
## [1] 71.02
## [1] "True Positive Rate, TPR (percentage):"
```

```
## [1] 63.98
## [1] "False Postive Rate, FPR (percentage):"
## [1] 20.07
```

With the QDA model, The model accuracy is 71.06 which is a little less than the other models. The true positive rate is 65.54 which is also less than the other models however, the false positive rate is 22.68 which is 5% and 2% less than the other models and is considered better.