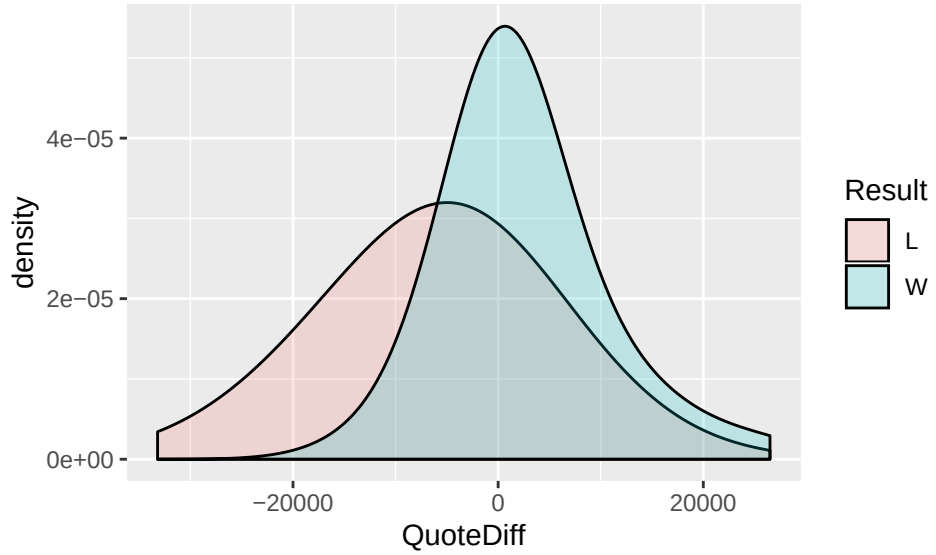
DA2 Homework Classification 1

Get *[Quote_ID],[Quote], [Competitor_Quote] and [Result] FROM [dbo].[Quote]* in the Accouting database. Then, create a column named QuoteDiff. Review a density plot of QuoteDiff based on Result.



Build a model to predict Result (W or L) based on the QuoteDiff only. Use LDA. Create a confusition matrix, and show. Should be similar to the following:

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction   L   W
##          L  66  17
##          W  71 163
##
##                Accuracy : 0.7224
##                  95% CI : (0.6696, 0.771)
##     No Information Rate : 0.5678
##     P-Value [Acc > NIR] : 9.285e-09
##
##                   Kappa : 0.4064
##
##  Mcnemar's Test P-Value : 1.606e-08
##
##             Sensitivity : 0.9056
##             Specificity : 0.4818
##          Pos Pred Value : 0.6966
##          Neg Pred Value : 0.7952
##              Prevalence : 0.5678
##          Detection Rate : 0.5142
##    Detection Prevalence : 0.7382
```

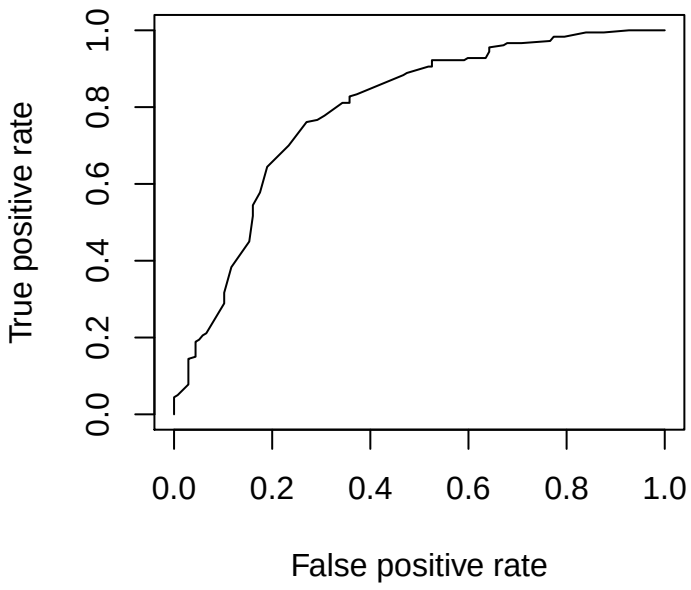
```
##       Balanced Accuracy : 0.6937
##
##        'Positive' Class : W
##
```

Now plot out the ROC curve.

```
##          Actual
## Predicted   L   W Sum
##       L    66  17  83
##       W    71 163 234
##       Sum 137 180 317
```

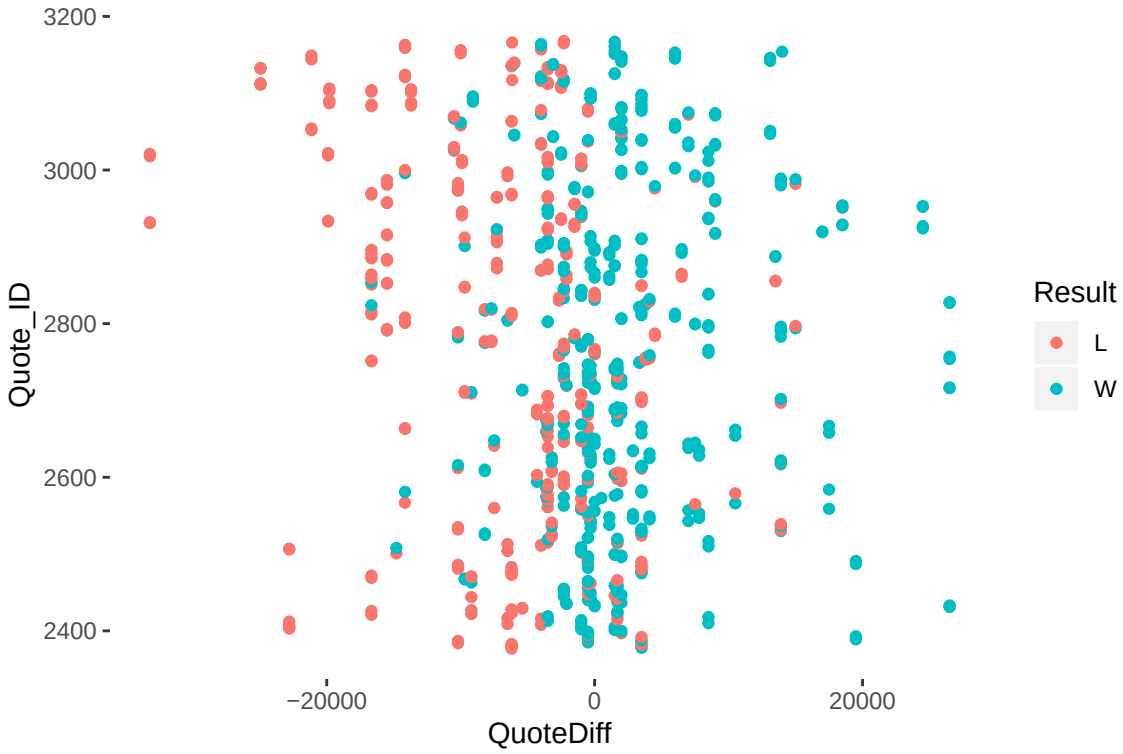
```
## [1] "Sensitivity = 0.905555555555556"
```

```
## [1] "Specificity = 0.481751824817518"
```




```
## [1] "AUC = 0.789334955393349"
```

Now, show model predictions on a plot as shown below:


3200
3000
2800
2600
2400
Quote_ID
−20000
0
20000
QuoteDiff
Result
L
W