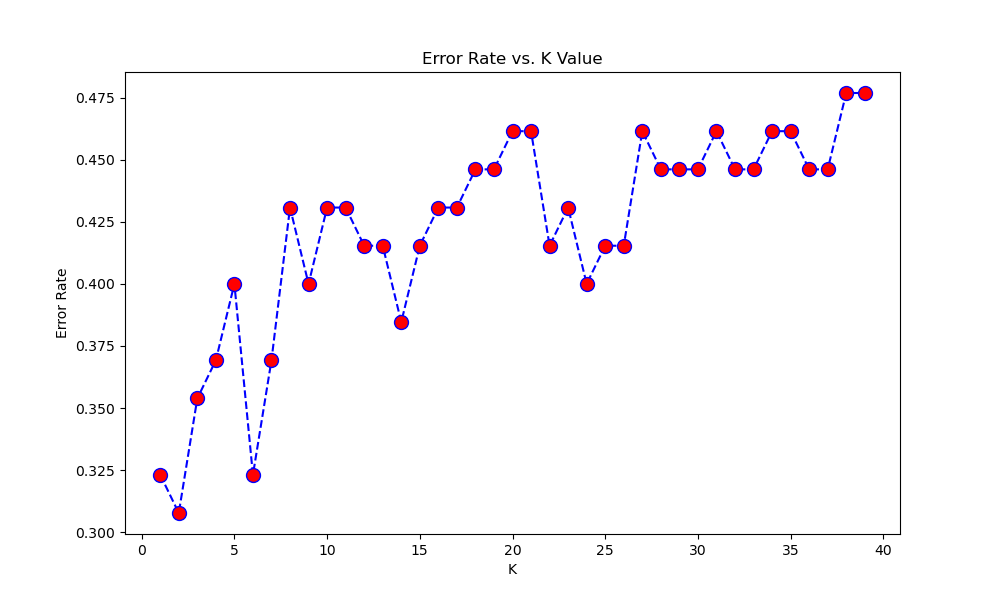
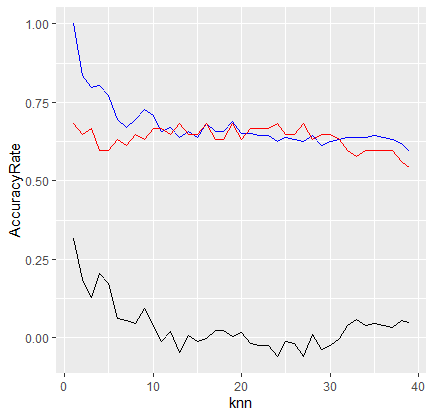
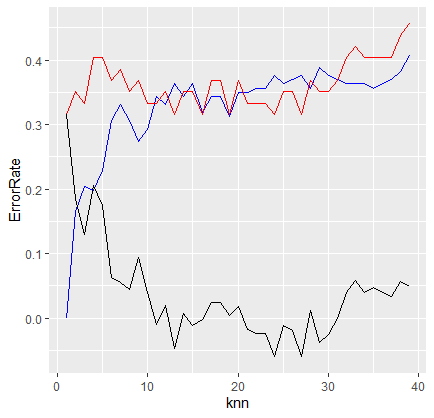
KNN Classification ML algo

* Glass Dataset:
* We are predicting the “type” col or feature based on the components in that glass.
* There are 7 types of glass with its freq.
* 1 2 3 5 6 7
* 70 76 17 13 9 29
* #1 2 3 5 6 7
* #32.710280 35.514019 7.943925 6.074766 4.205607 13.55140
* Type 3 Glass is the most frequently occurred in the dataset with 35% .
* Firstly k=1 was used and gave Accuracy 0.6842105.
* It was Overfitting as k value was too small and it literally predict based on 1 point as its neighbor and predict.
* In the Error rate v K value as the K value increases the error also rose up.
* The lowest dip in error was at k =7 or k=3.
* The other graph is the Error and Accuracy of prediction from Actual and black line is the Difference btw them.
* # Build a KNN model on dataset knn( X\_train , X\_test , y\_train) on k=13
* predictions <- knn(train = glass\_train, test = glass\_test, cl = glass\_train\_labels, k=13)
* # ----------- Model Evaluations
* table(predictions,glass\_test\_labels)
* #Accuracy 0.6842105
* mean(predictions==glass\_test\_labels)
* #0.6491228
* CrossTable(x=glass\_test\_labels,y=predictions,prop.chisq = FALSE)
* #Accuracy increased to 70 %







* **Zoo DataSet:**
* There are 7 types of animal class in this dataset.
* #1 2 3 4 5 6 7
* #40.6 19.8 5.0 12.9 4.0 7.9 9.9
* With Class Type 1 being the most repeated one with 40.6 % .
* With k =5 we got the of Accuracy 0.7619048.
* # knn lowest error rate occurs when k = 10
* With k =10 we got [1] 0.8095238 . Better accuracy then the last k value.

Accuracy plot : Acc of train data goes lower and lower with increase in K value.

Error Plot : Same with blue line .Error also goes slightly lower with increase in k value.

