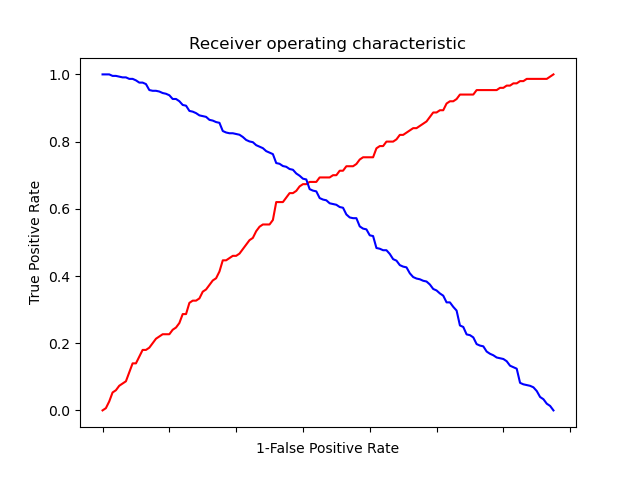
**Logistic Regression**

1. **Affairs Dataset**

In this dataset there were some multicollinearity present within dataset. Which makes the X -features have sensitivity among themselves. Every increase in X will have impact on y . But it should not have impact on within X itself.

vryhap,vryrel,yrsmarr6 having highest VIF (Variance Inflation Factor). 🡨- having multicollinearity between them.

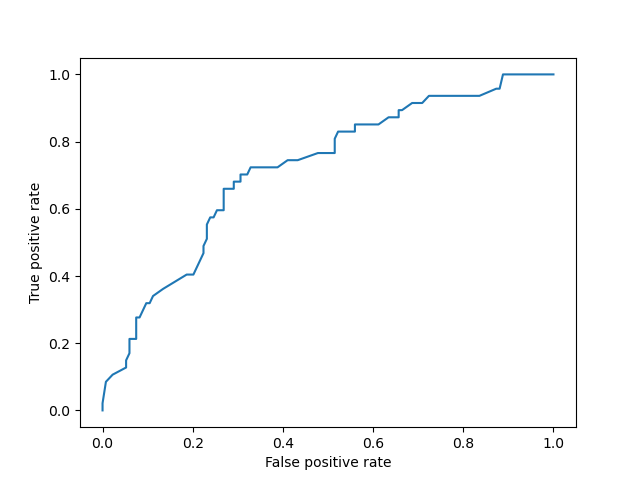


After logistic Regerssion , plotting ROC curve :

With ROCs you can assess some model's discrimination ability, independently of the response criteria, and also establish the optimal response criteria.

In this graph as the False Pos Rate is increasing we are farther away from TPR .

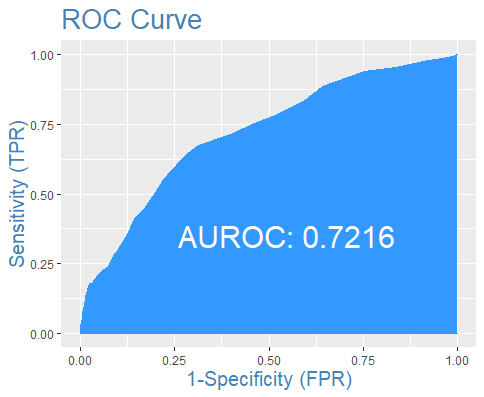
More values for FPR is inverse to TPR contribution . (more values are FP )



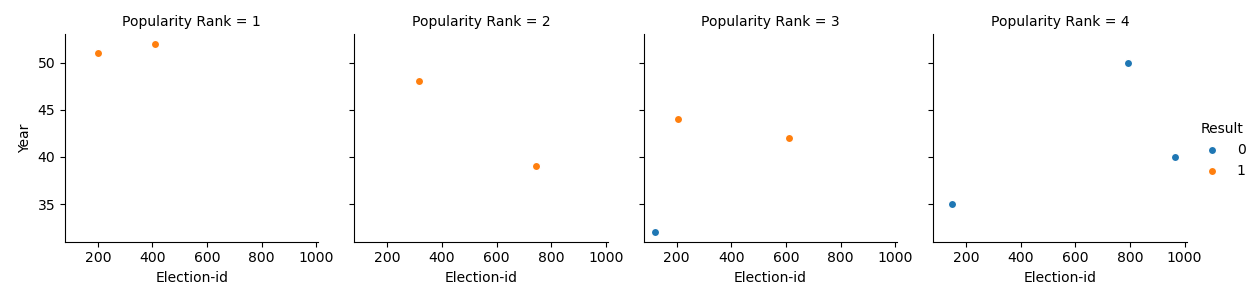
The Test Data is also showing signs of TPR . is also Diagonal indicating its useless test.

A perfect test is most touching y-axis and lower FPR.

Area Under ROC Curve .helps in indicating which models performed best after knowing the best threshold.

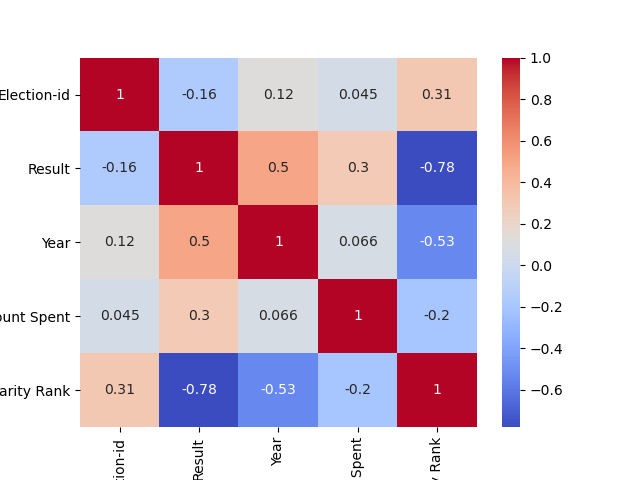


1. **Election Dataset**

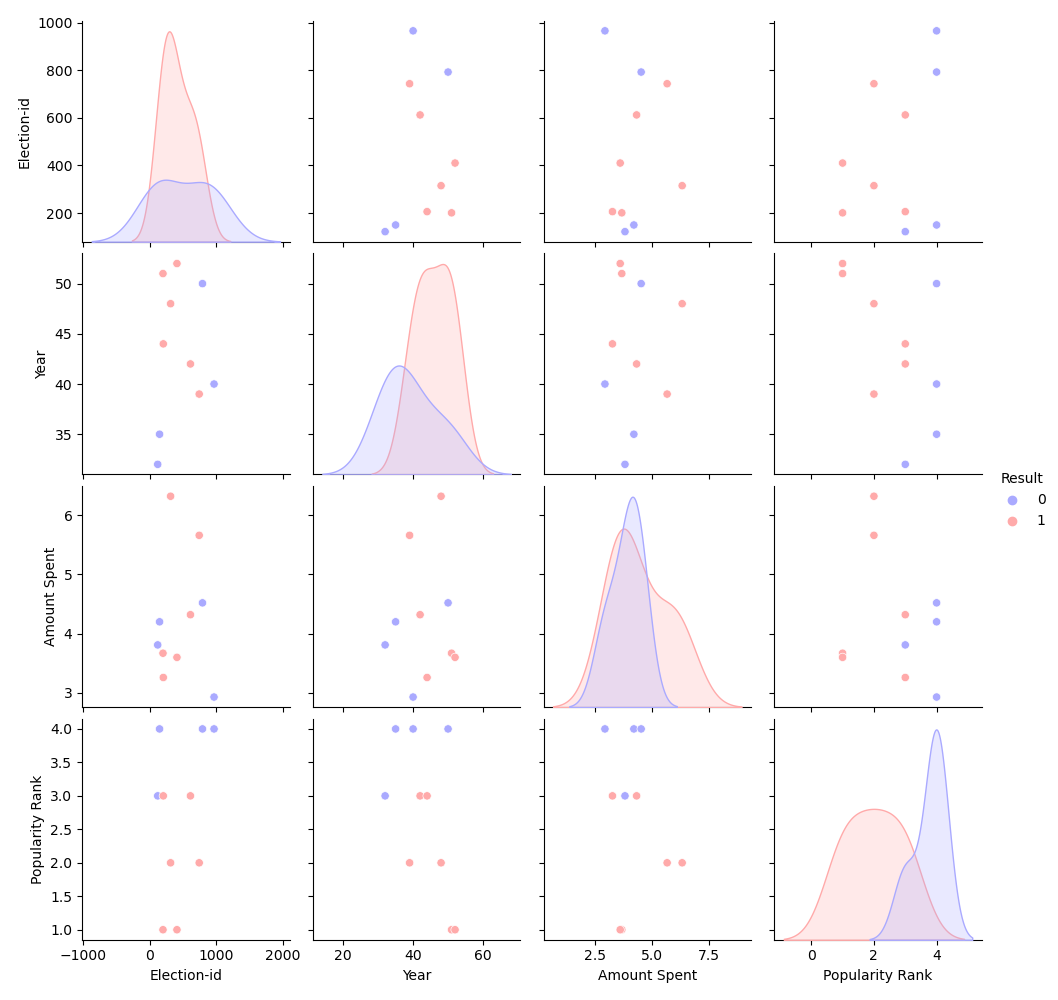


**This facetgrid indicates that more popularity Rank =4 .The result was differed as the year went by .**

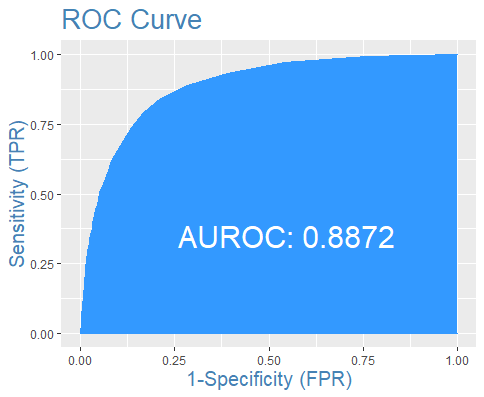
**Election id Represent different leader and by each year.**

****

**The only good correlation observed is for Popularity rank and Election ID which is 0.31**

****

1. **Bank Dataset**

****

**The Area underROC is 88% .**

0.8-0.9= Excellent discrimination of classification for Bank Dataset .

As in this curve at the start we can see the line of curve is straightened . almost closer to y -axis and at certain threshold it started to increase FPR.

That point can be around 0.25 as a threshold for classifying