1. **Make use of all the modeling techniques that you know to build models to forecast mortgage delinquency.**

Answer:

- Two Class Logistic Regression
- Two Class Boosted Decision Tree
- Two Class Decision Forest
-

2. **Show model performance metrics: FP, FN, Overall error, Sensitivity, Specificity, F1, AUC. What is the best metric to evaluate model performance?**

Answer:

1. **Two Class Logistic Regression**

True Positive	False Negative	Accuracy	Precision	Threshold	AUC
1063	6056	0.837	0.559	0.5	0.793
False Positive	**True Negative**	**Recall**	**F1 Score**		
838	34261	0.149	0.236		

2. **Two Class Boosted Decision Tree**

True Positive	False Negative	Accuracy	Precision	Threshold	AUC
1686	5433	0.838	0.545	0.5	0.804
False Positive	**True Negative**	**Recall**	**F1 Score**		
1409	33690	0.237	0.330		

3. Two Class Decision Forest

Model	FP	FN	Overall Error	Sensitivity	Specificity	F1	AUC	FNR	FPR
Two Class Logistic Regression	838	6056	17.1%	14.9%	97.6%	0.236	0.793	76.3%	4%
Two Class Boosted Decision Tree	1409	5433	17.5%	23.7%	96%	0.330	0.804	83.1%	2.4%
Two Class Boosted Decision Tree	1496	5784	18.5%	18.8%	95.7%	0.268	0.750	81.2%	4.3%

I think the best metric to evaluate model performance is **Sensitivity**, because this tells us about the accuracy of how many positives were identified correctly.

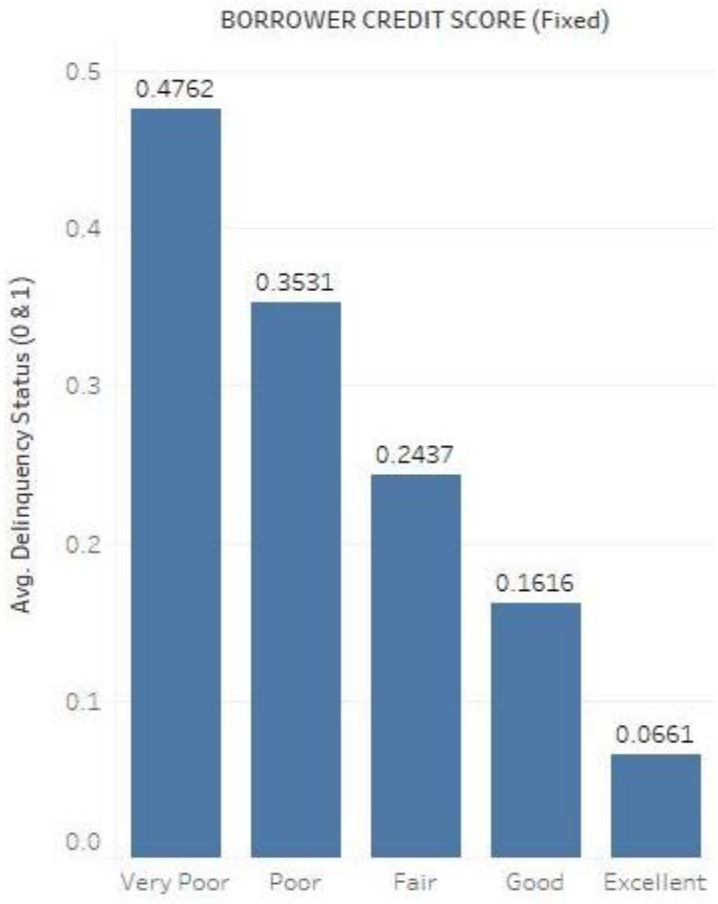
4. Did Fannie Mae have information that could have accurately predicted defaults among mortgages issued in Q1 2007?

Answer: Yes, Fannie Mae have the information that could have accurately predicted defaults among mortgages issued in Q1 2007. However, there are still false negatives that was considered while prediction, so the outcome of the model may vary.

5. **Visualize the effects of the top 4 predictors of mortgage defaults. Are there clear emergent patterns?**

Answer:

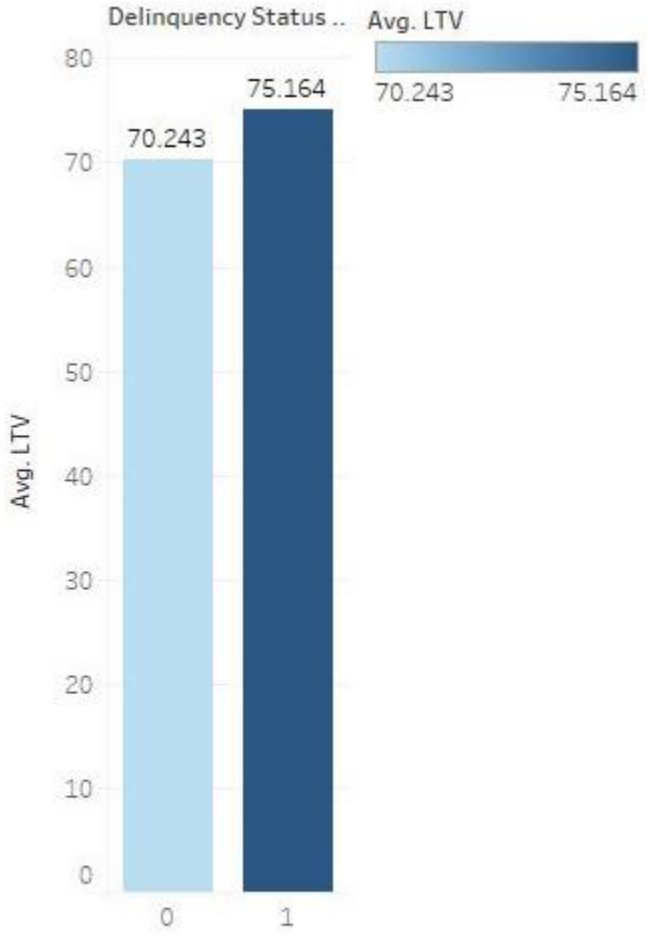
Borrower Credit Score vs. Avg. Delinquency Status



Average of Delinquency Status (0 & 1) for each BORROWER CREDIT SCORE (Fixed). The marks are labeled by average of Delinquency Status (0 & 1). The view is filtered on BORROWER CREDIT SCORE (Fixed), which keeps Excellent, Good, Fair, Poor and Very Poor.

There is negative relation in between Borrower Credit Score vs. Avg. Delinquency Status. Borrowers with high credit score have low possibility of Delinquency status.

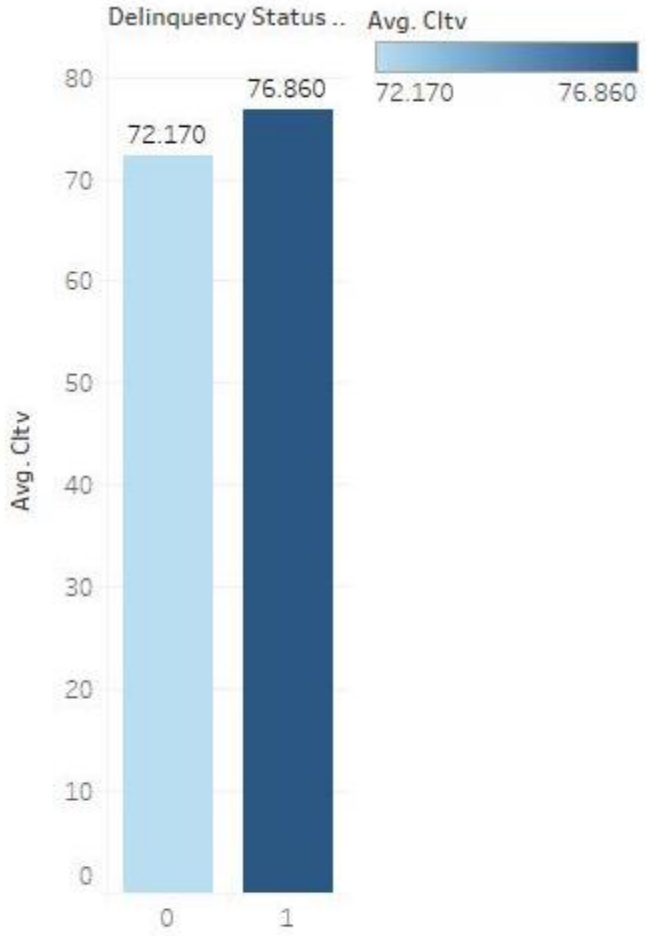
Average LTV vs. Delinquency Status



Average of LTV for each Delinquency Status (0 & 1). Color shows average of LTV. The marks are labeled by average of LTV.

As the average of LTV increases, the chance of Delinquency Status is increases.

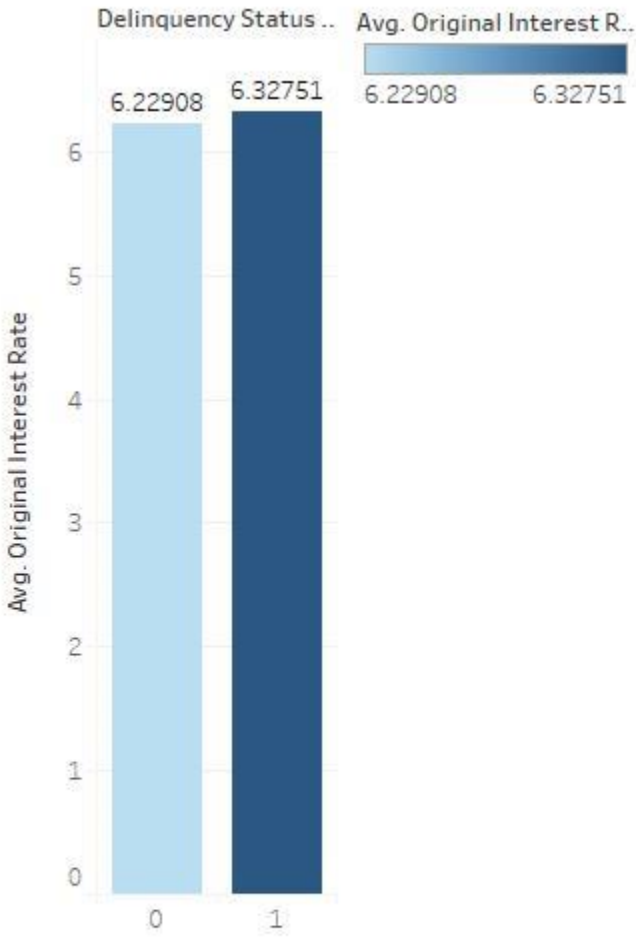
Average CLTV vs. Delinquency Status



Average of Cltv for each Delinquency Status (0 & 1). Color shows average of Cltv. The marks are labeled by average of Cltv.

As the average of CLTV increases, the chance of Delinquency Status is increases.

Average Original Interest Rate vs Delinquency Status



Average of Original Interest Rate for each Delinquency Status (0 & 1). Color shows average of Original Interest Rate. The marks are labeled by average of Original Interest Rate.