

Model building:

Logistic regression score (logistic_model): 86.52482269503547

Decision tree score (tree_model): 79.83870967741935

Random forest score (forest_model): 84.78260869565219

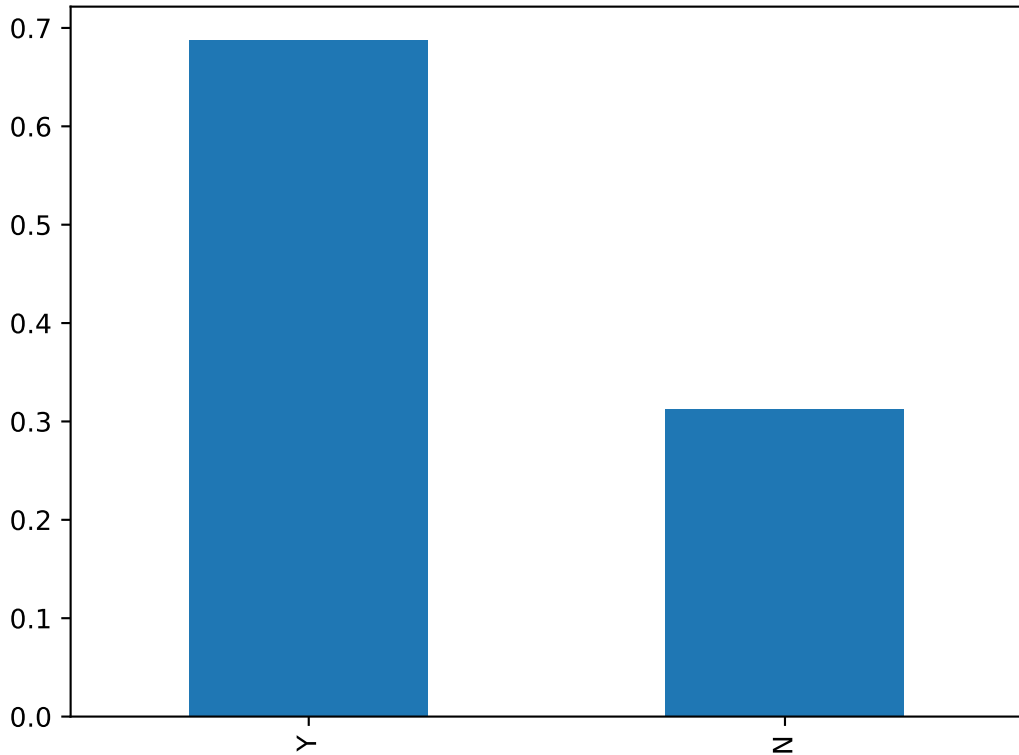
XGBoost score (xgb_model): 81.203007518797

Model with highest accuracy is: logistic_model

Using this model to predict the test data:

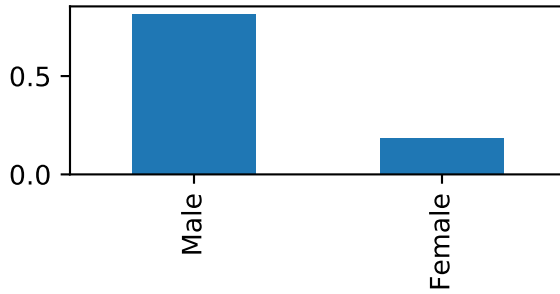
```
[1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 0 1  
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 0 1 1 1 1 0 1 1 0 0 1 0 1 1 1 1  
1 1 1 1 1 1 1 0 1 0 1 0 1 0 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 0 1 1 1 1 1  
1 1 1 1 1 1 1 0 0 0 1 1 1 0 0 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 0 1 1 1 1 0  
1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 0 1 1 0 0 0 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 1 1 1  
1 1 1 1 1 1 1 1 0 0 1 1 0 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1  
1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 0 0 1 1 1 1 0 1 0 1 0 1 1 1 1 1 0 1 1 1 1 0 1 1 1  
1 1 1 1 1 1 1 1 0 1 0 1 1 1 1 1 0 0 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1  
1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1  
1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1]
```

Loan_Status

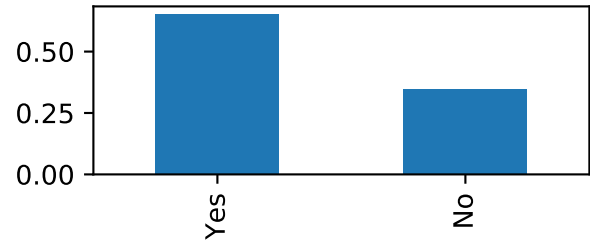


Analysing categorical variables

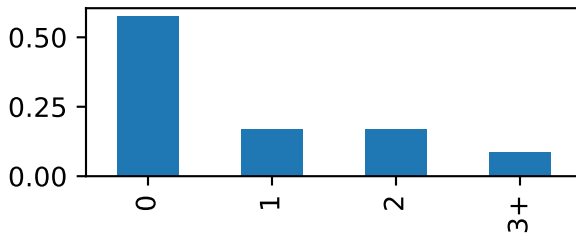
Gender



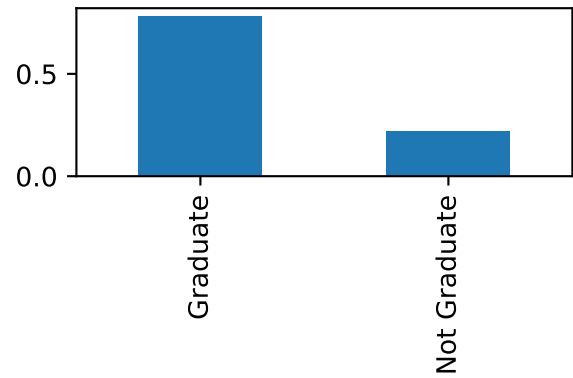
Married



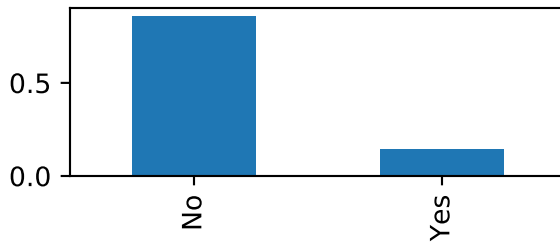
Dependents



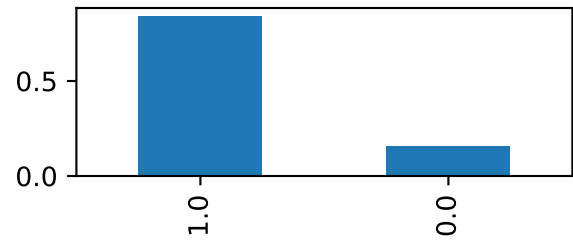
Education



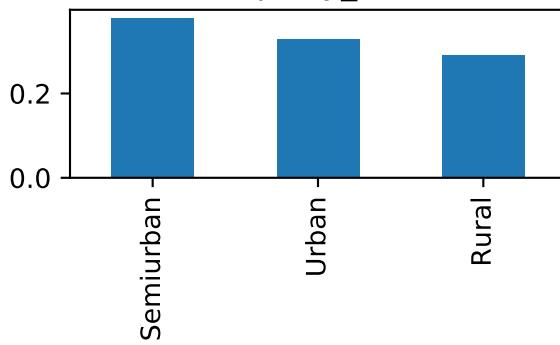
Self_Employed



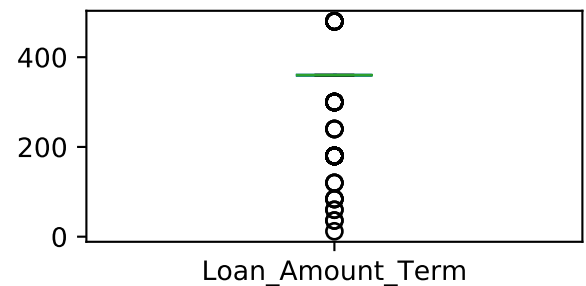
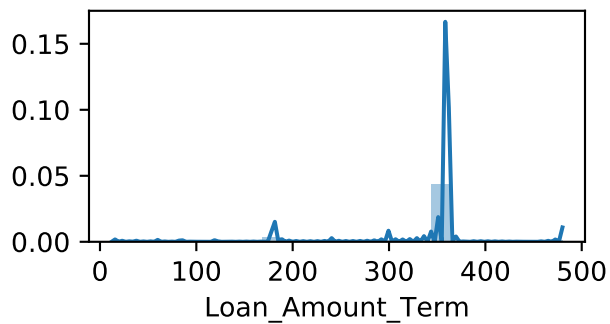
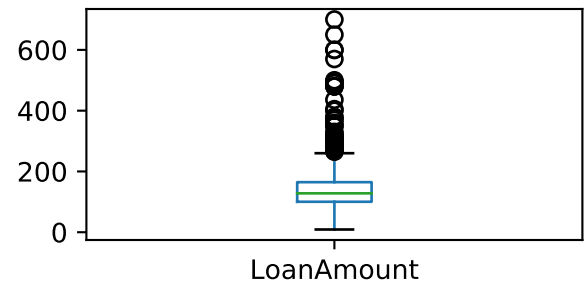
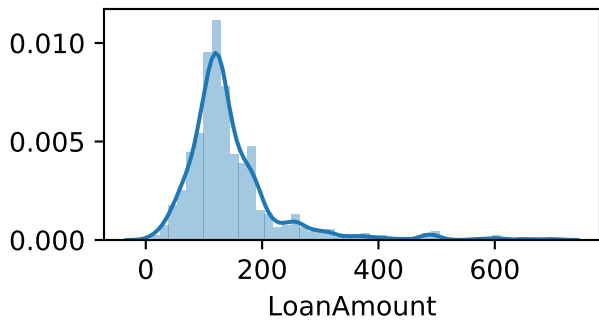
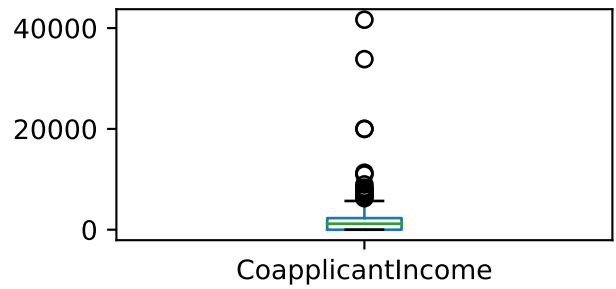
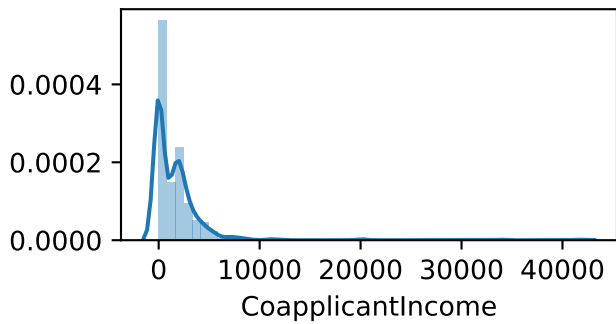
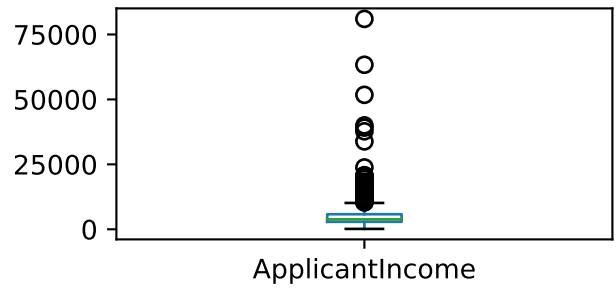
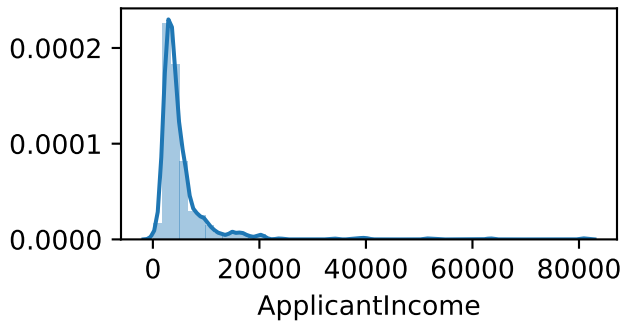
Credit_History



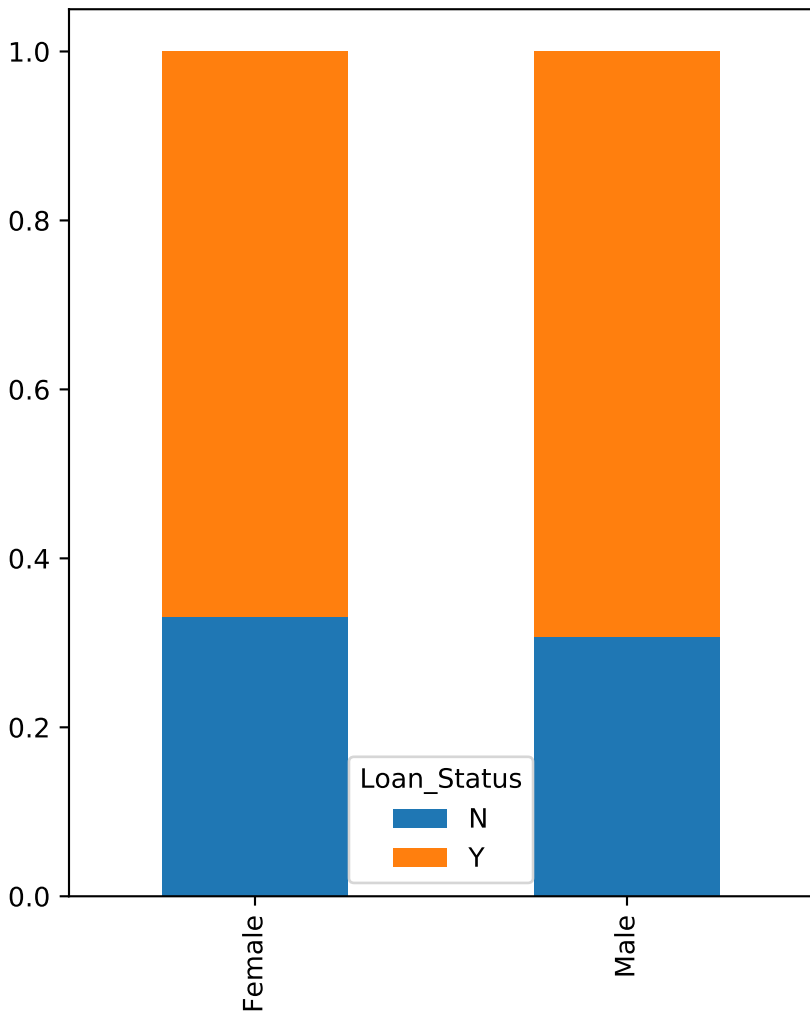
Property_Area



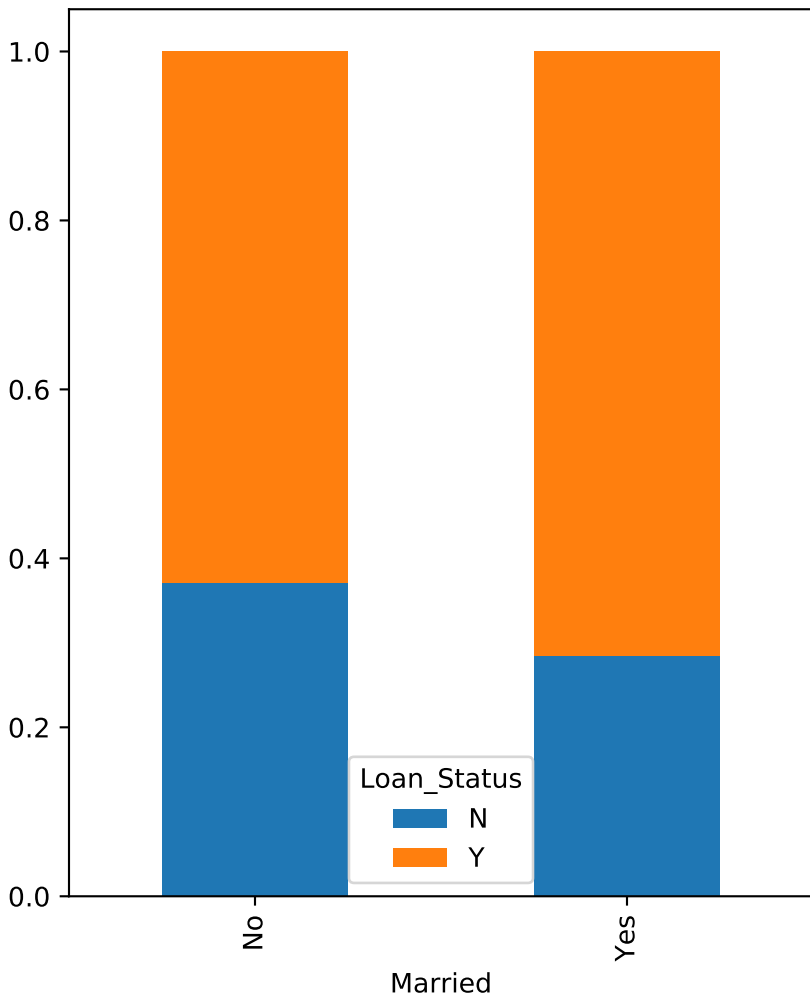
Analysing categorical variables



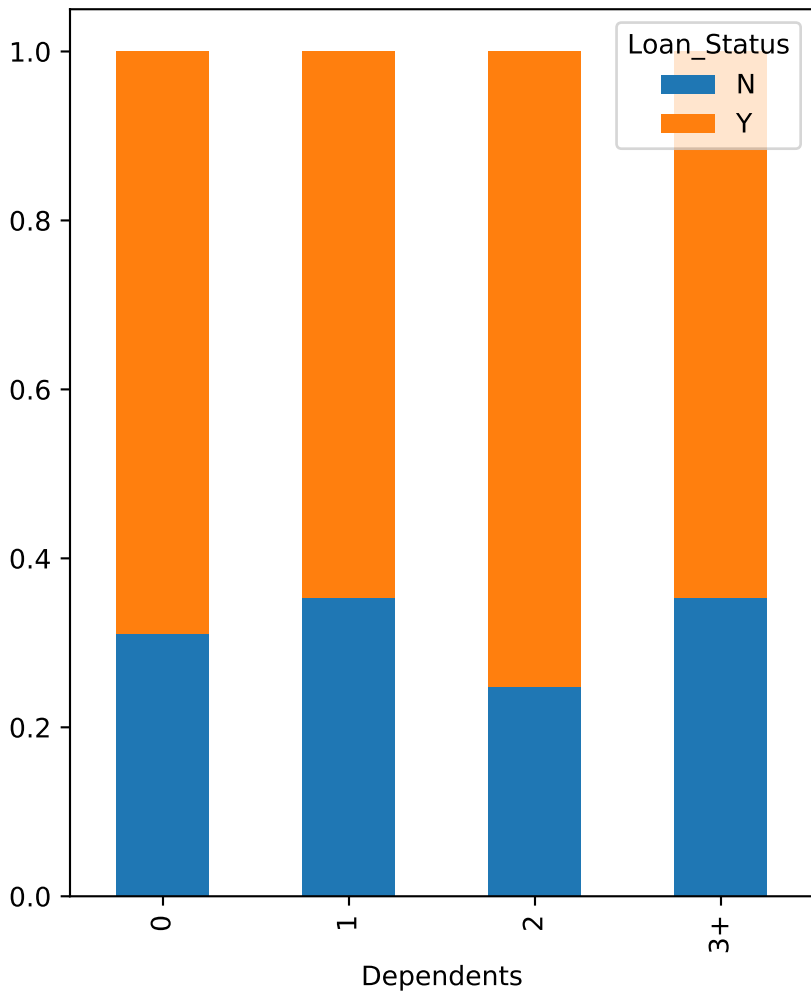
Gender



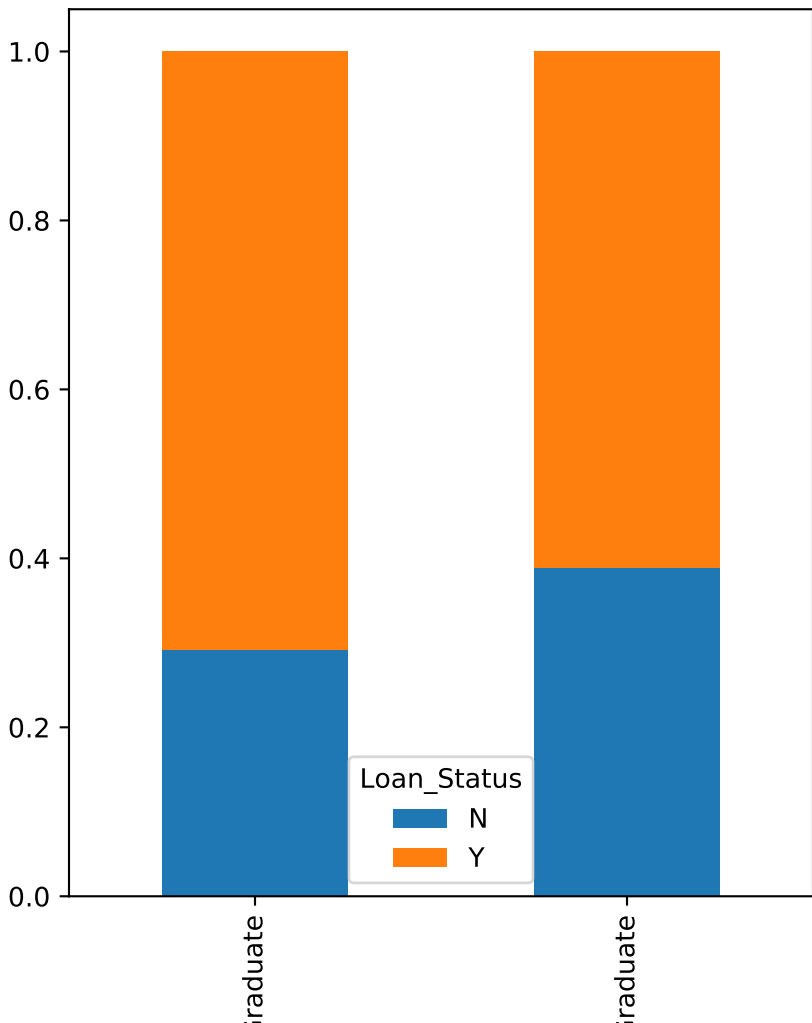
Married



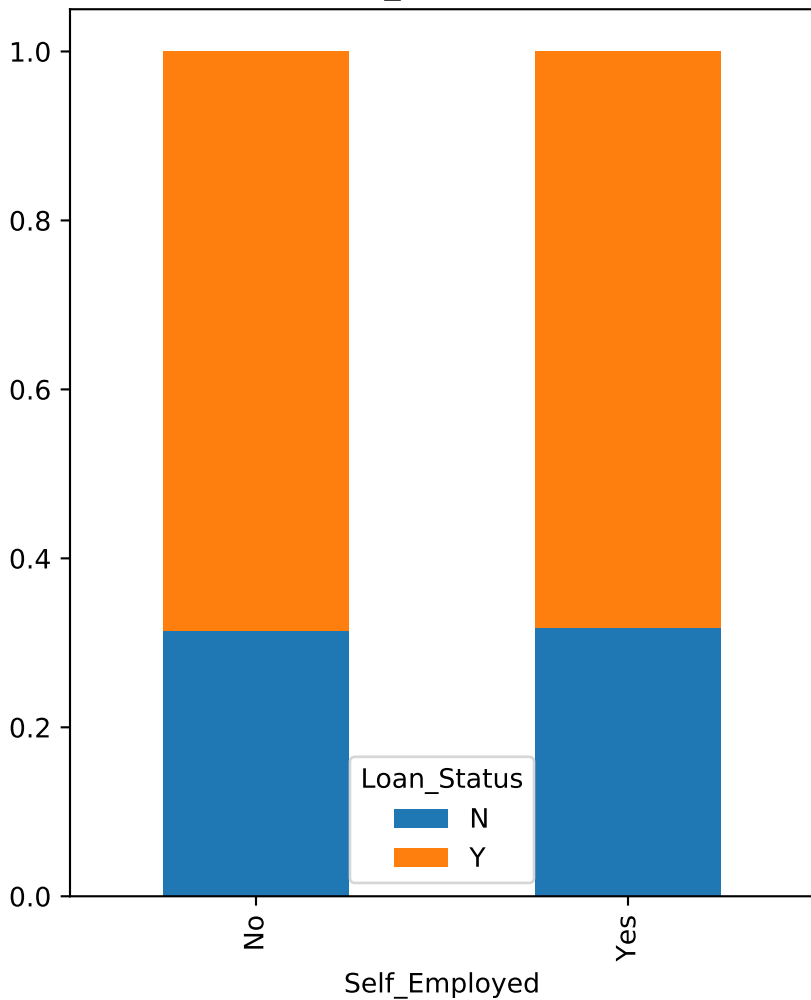
Dependents



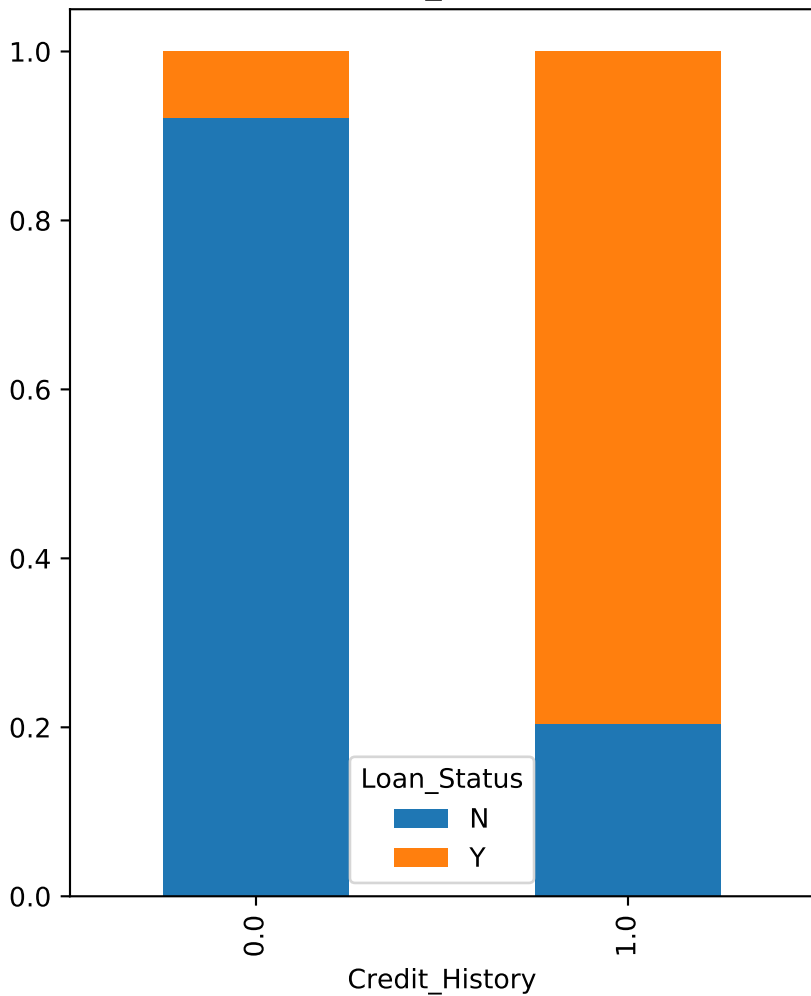
Education



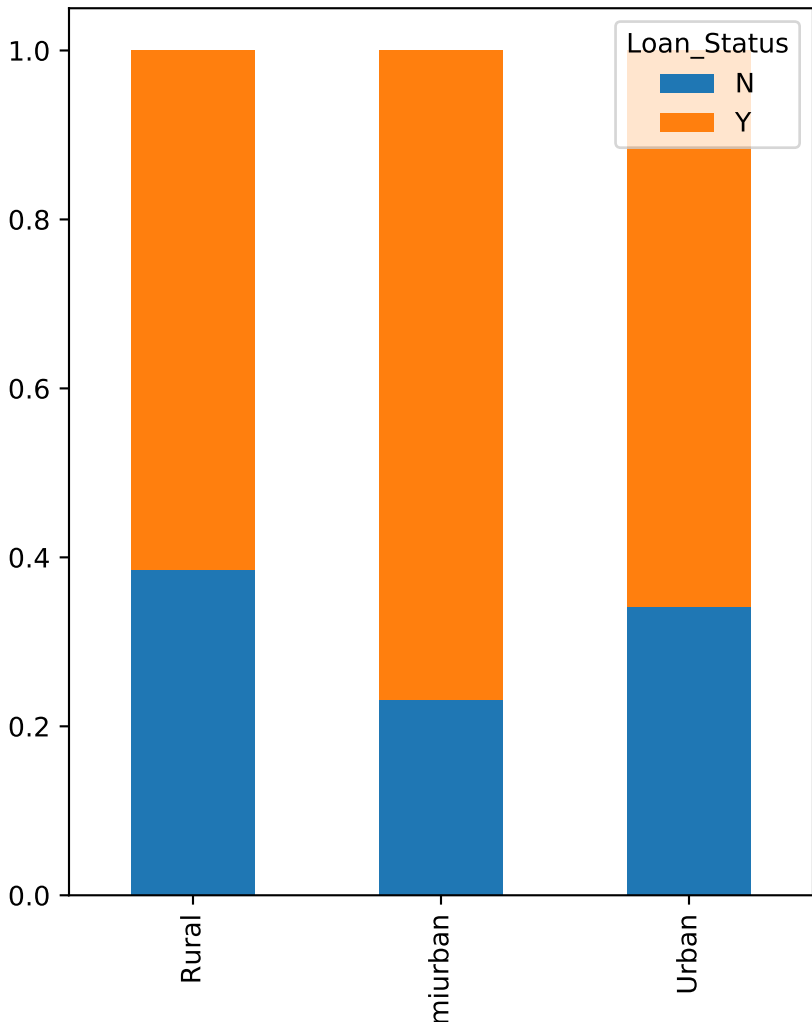
Self_Employed



Credit_History



Property_Area



Analysing numerical variables vs decision variable

