


DATA MINING PROJECT

TITLE :- HOME LOAN PREDICTION

A wooden house-shaped block and several stacks of coins on a wooden surface.

DONE BY :-
E Hemasundar Yadav

AGENDA :-

- ❖ Find whether the person is eligible for loan or not.
- ❖ Import data set and Finding whether the null values present in the data set.
- ❖ Data preprocessing.
- ❖ Variable analysis.
- ❖ Log transformation.
- ❖ Label encoding.
- ❖ Model training.

Problem Statement :-

A finance company wants to classify whether the person is eligible to get loan or not eligible. So when a customer fills online application, company wants to automate the loan based on some variables. Here let's help the company and make their work easy.

Taken columns :- Loan_id, Gender, married, Dependents, Education, Self_Employed, ApplicationIncome, CoapplicationIncome, LoanAmount, Loan_Amount_Term, Credit_History, Property_Area, Loan_status

Data preprocessing :-

- ❖ There are “149” Null Values.
- ❖ For Categorical Columns :-

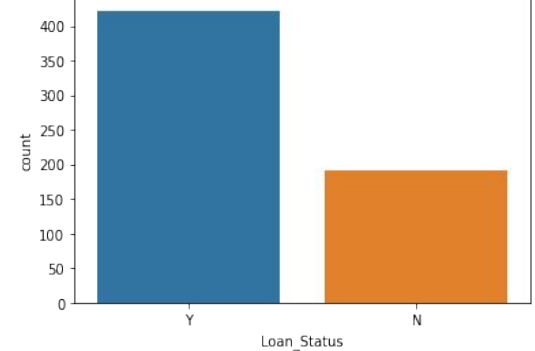
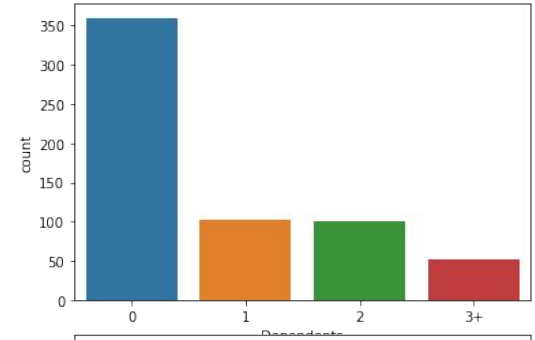
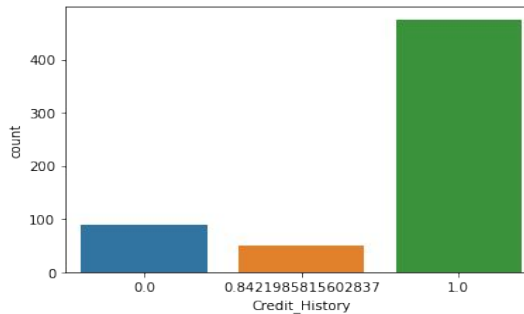
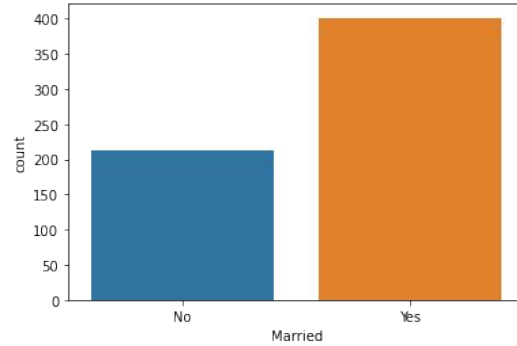
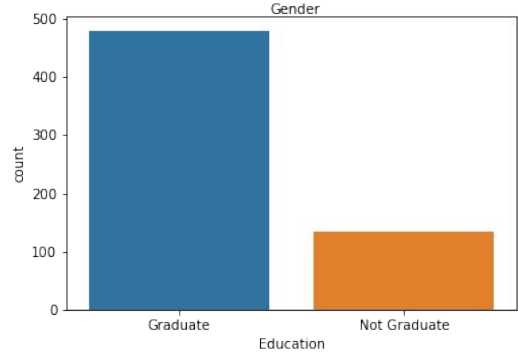
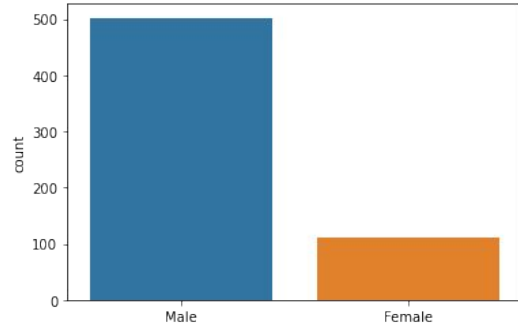
Filling null values for the categorical columns with mode values.

- ❖ For Numerical Columns :-

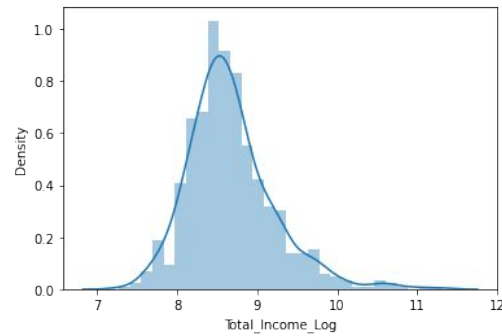
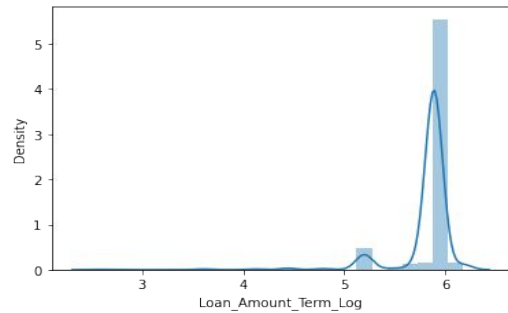
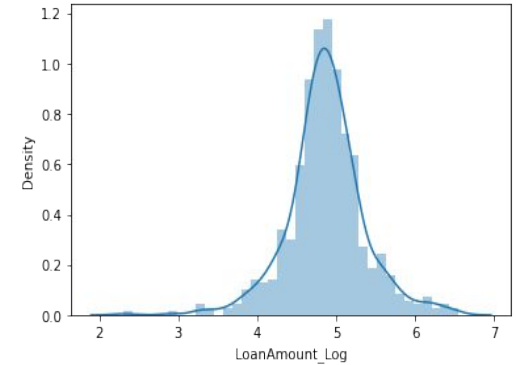
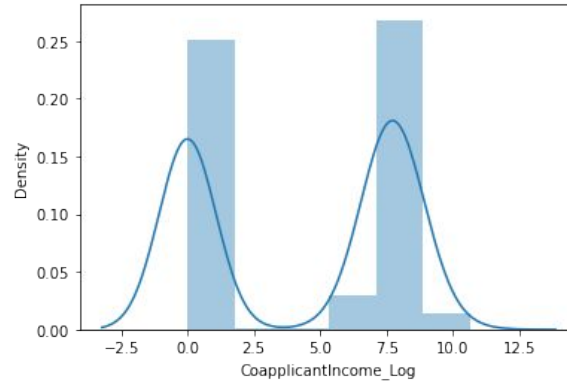
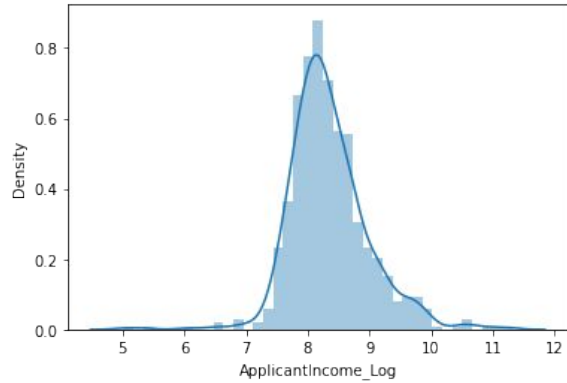
Filled the null values for the numerical columns with mean.



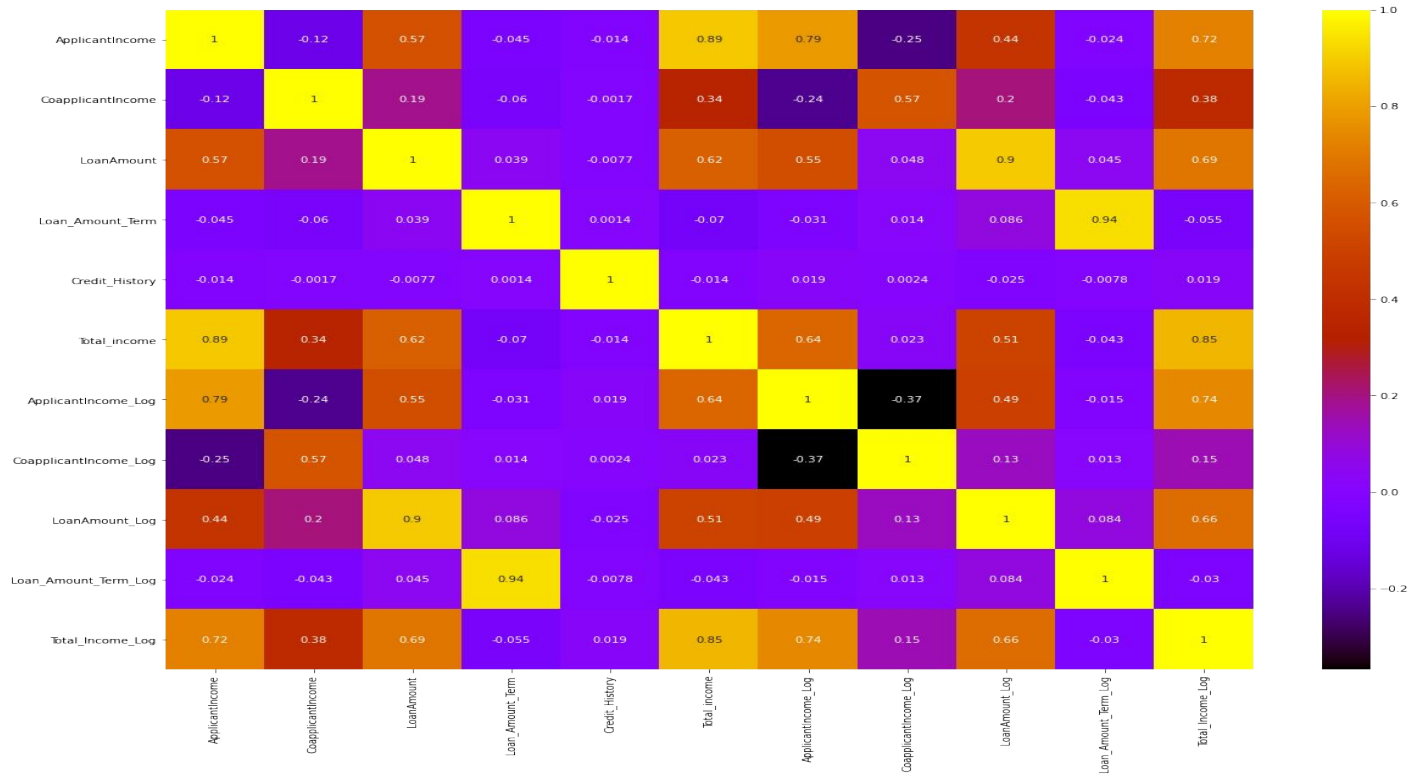
Variable analysis :-



Log transformation :-



HeatMap :-



Result :-

- In Logistic regression we got most accuracy score that is 81.16883116883116.
- We can see that logistic regression fits the model better than the other classification models.



THANK YOU :)