**Introduction:**

We will solve the Loan Approval Prediction in this article. We must classify whether the loan will be granted or not in this classification issue. A predictive modeling task in which a class label is anticipated for a given example of input data is referred to as classification.

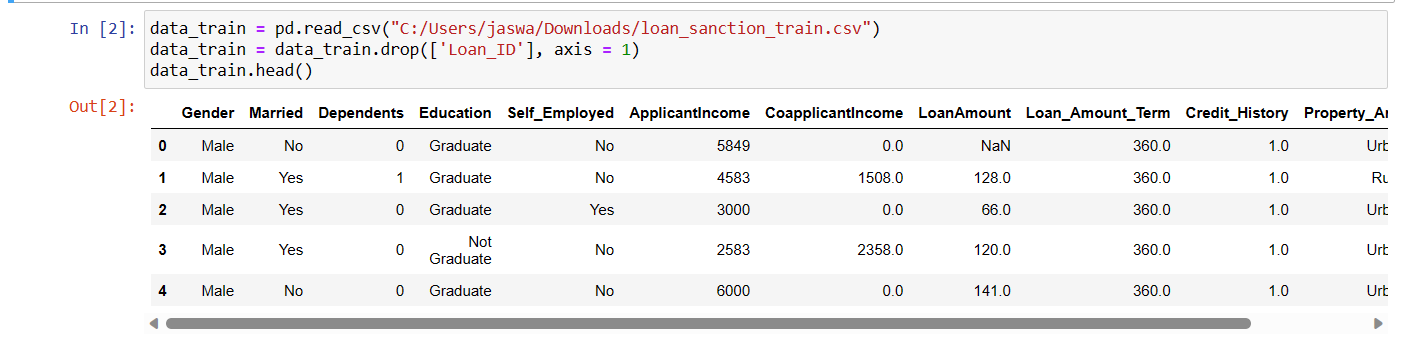
**Data attributes:**

Data columns (total 12 columns):

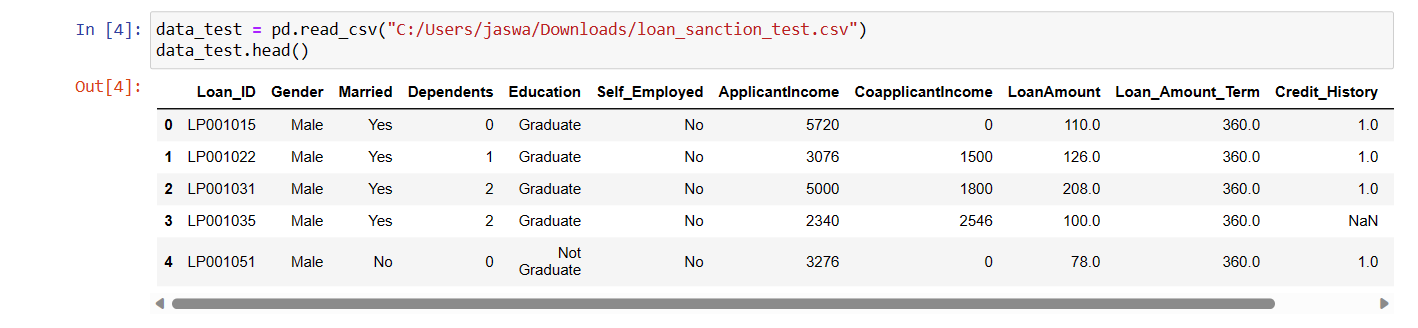
# Column Non-Null Count Dtype

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 0 Gender | 601 | non-null |  | object |
| 1 Married | 611 | non-null |  | object |
| 2 Dependents | 599 | non-null |  | object |
| 3 Education | 614 | non-null |  | object |
| 4 Self\_Employed | 582 | non-null |  | object |
| 5 ApplicantIncome | 614 | non-null |  | int64 |
| 6 CoapplicantIncome | 614 | non-null |  | float64 |
| 7 LoanAmount | 592 | non-null |  | float64 |
| 8 Loan\_Amount\_Term | 600 | non-null |  | float64 |
| 9 Credit\_History | 564 | non-null |  | float64 |
| 10 Property\_Area | 614 | non-null |  | object |
| 11 Loan\_Status | 614 | non-null |  | object |

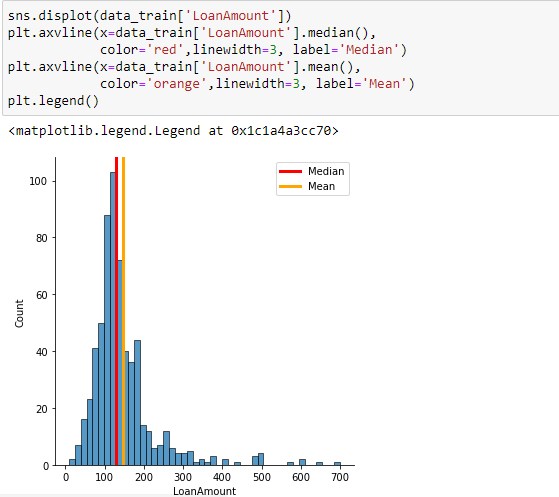
**Training data:**



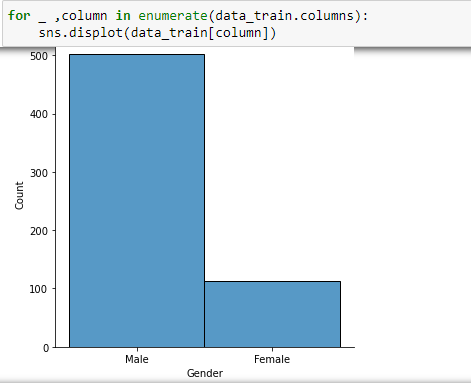
**Testing data:**

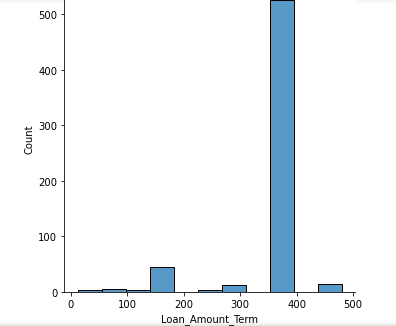
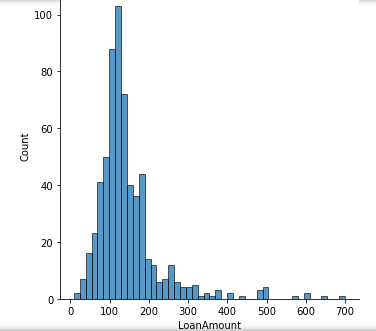
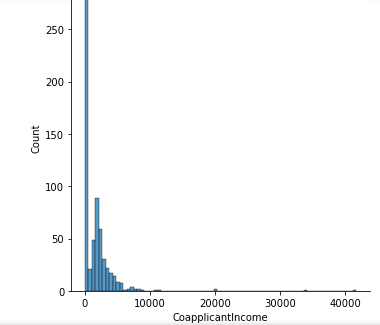
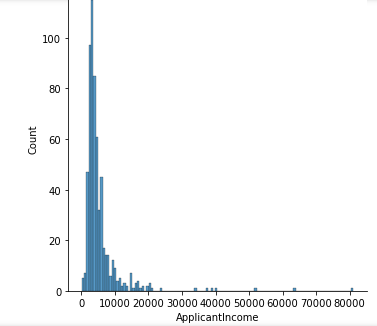
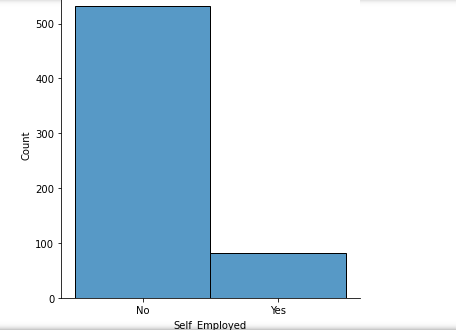
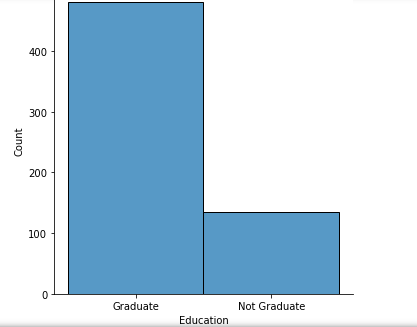
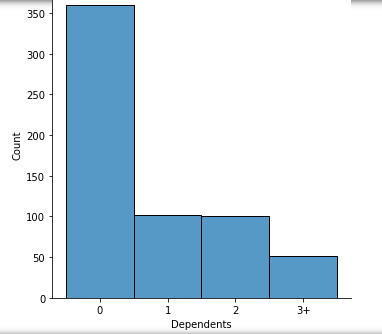
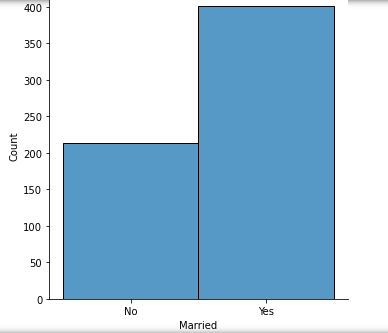
****

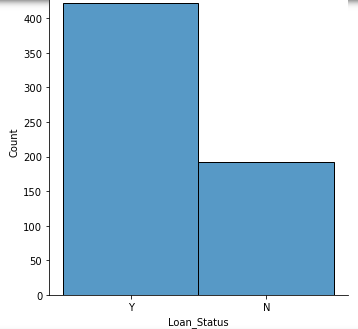
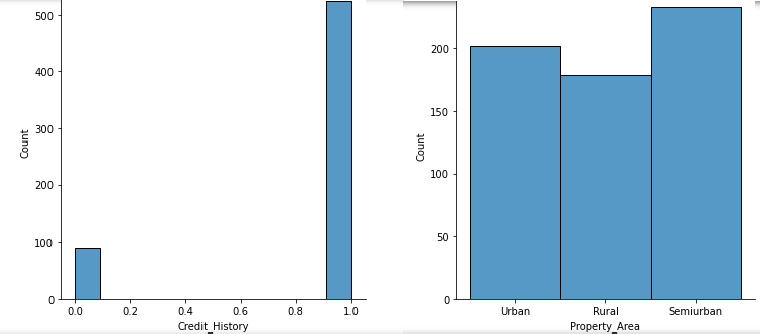
**Loan amount Mean and median:**



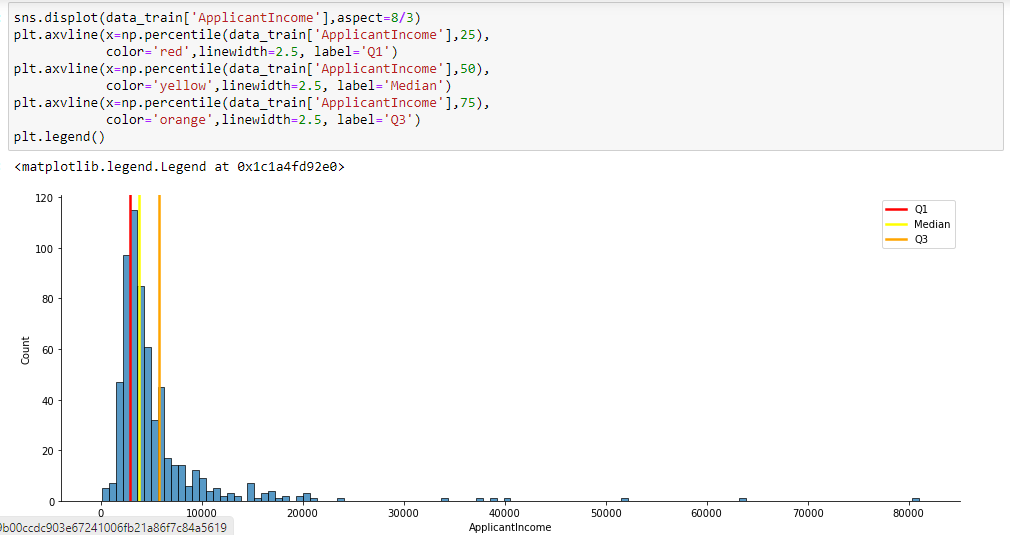
**Information about values for attributes:**





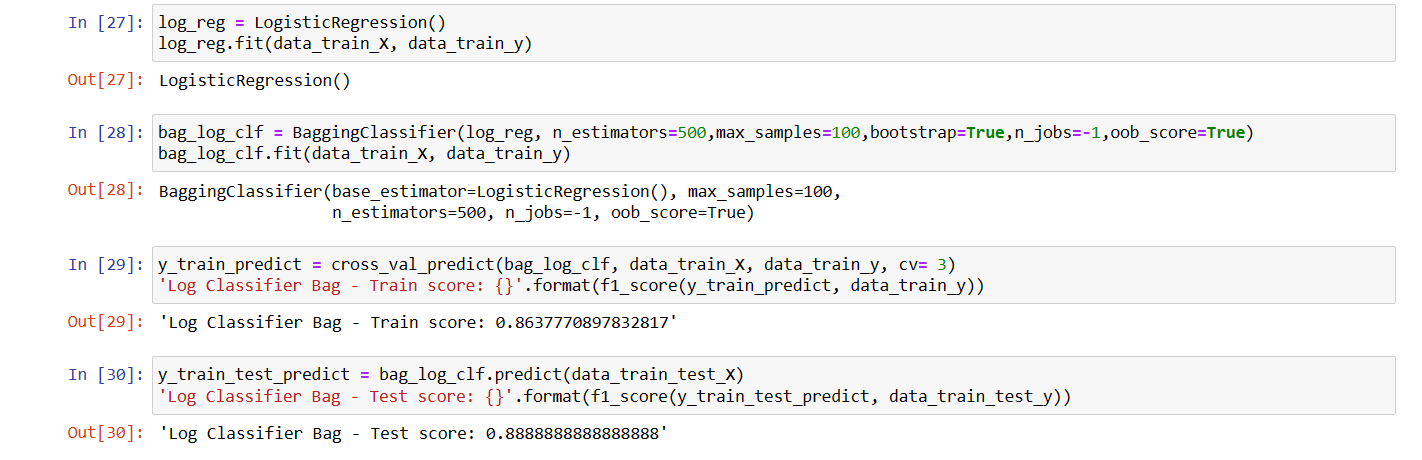


**Applicant income Median and quartiles:**



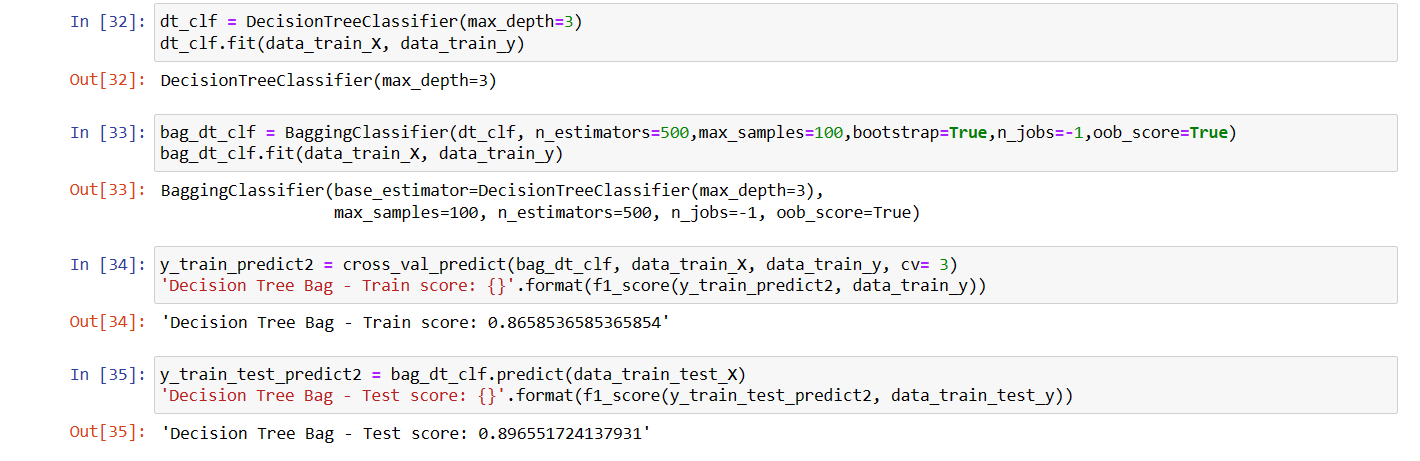
**Logistic regression:**

Predictive models built using this approach can make a positive difference in your business or organization. Because these models help you understand relationships and predict outcomes, you can act to improve decision-making.



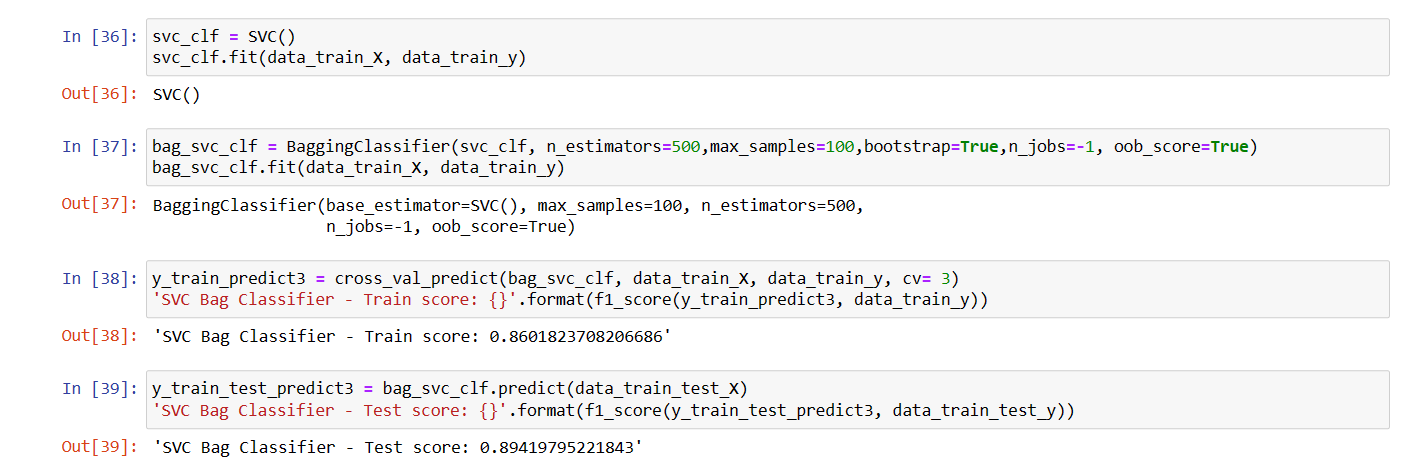
**Decision tree:**

A decision tree is a type of [supervised machine learning](https://www.mastersindatascience.org/resources/introduction-to-machine-learning-algorithms/) used to categorize or make predictions based on how a previous set of questions were answered. The model is a form of supervised learning, meaning that the model is trained and tested on a set of data that contains the desired categorization.



**Svc:**

A Linear SVC (Support Vector Classifier) is designed to fit to the data you provide and provide a "best fit" hyperplane that divides or categorizes your data. Following that, you may input some features to your classifier to check what the "predicted" class is after you've obtained the hyperplane.



**Results:**

**Accuracy table:**

|  |  |  |
| --- | --- | --- |
| model | train | test |
| Logistic regression | 86.3 | 88.8 |
| Decision tree | 86.5 | 89.6 |
| svc | 86.1 | 89.4 |

**Decision tree** gives the highest accuracy 86.5% on training data set and 89.6% on testing data set.

**Conclusion:**

From the Exploratory Data Analysis, we could generate insight from the data. How each of the features relates to the target. Also, it can be seen from the evaluation of three models that decision tree performed better th