1. Imported the python libraries
2. Read the data
3. Checking of the dataset has any null values.
4. Calculation natural logarithm of LoanAmount. **Why?-> Values seemed to be right-skewed(most values are small while some are very large). np.log() compresses large values and spreads out small values, making the distribution closer to normal.**
5. Finding the total income(application income+coapplication income) and finding its logarithm.
6. Filling the null values.
7. Checking the number of people by groups of gender, marital status, dependents, self-employed, loan amount and credit history.
8. Selecting specific columns from the dataset for feature selection(x) and target selection(y).
9. Use LabelEncoder to convert categorical data into numerical data.
10. Training the model using random forest, decision tree, naïve bayes and kneighbours to check which has highest frequesncy.