1. Read both the datasets
2. Create Interest Rate Variable in Test set
3. Create a New column representing whether the row coming from train set or test set
4. Check the column names matching or not
5. Concatenate the data frames
6. Deal with Interest rate, Debt to income ratio
7. Create two columns with Fico Range and calculate the mean, create that as new fico column.
8. Remove amount funded by investors because that variable is a future variable.
9. Create dummy variables for Loan purpose and state columns, make sure that there are not too many dummies created, check the frequency and control the dummies based on frequency.
10. Clean the employment length variable
11. Check the missing values and impute them with mean
12. Separate the train and test datasets
13. Split the train dataset into train1 and train2 datasets, bui9ld the model on train1 and validate the model on train2.
14. Perform regularization techniques like Ridge and Lasso
15. Find out the MAE and select the best estimators, fix the model.