Lending Club Loan Analysis

Karl L. & Yusuf C.

Data Description

- **42.538 observations**
- **52 features**
- **5 years: 2007-2011**
- **https://www.lendingclub.com/**

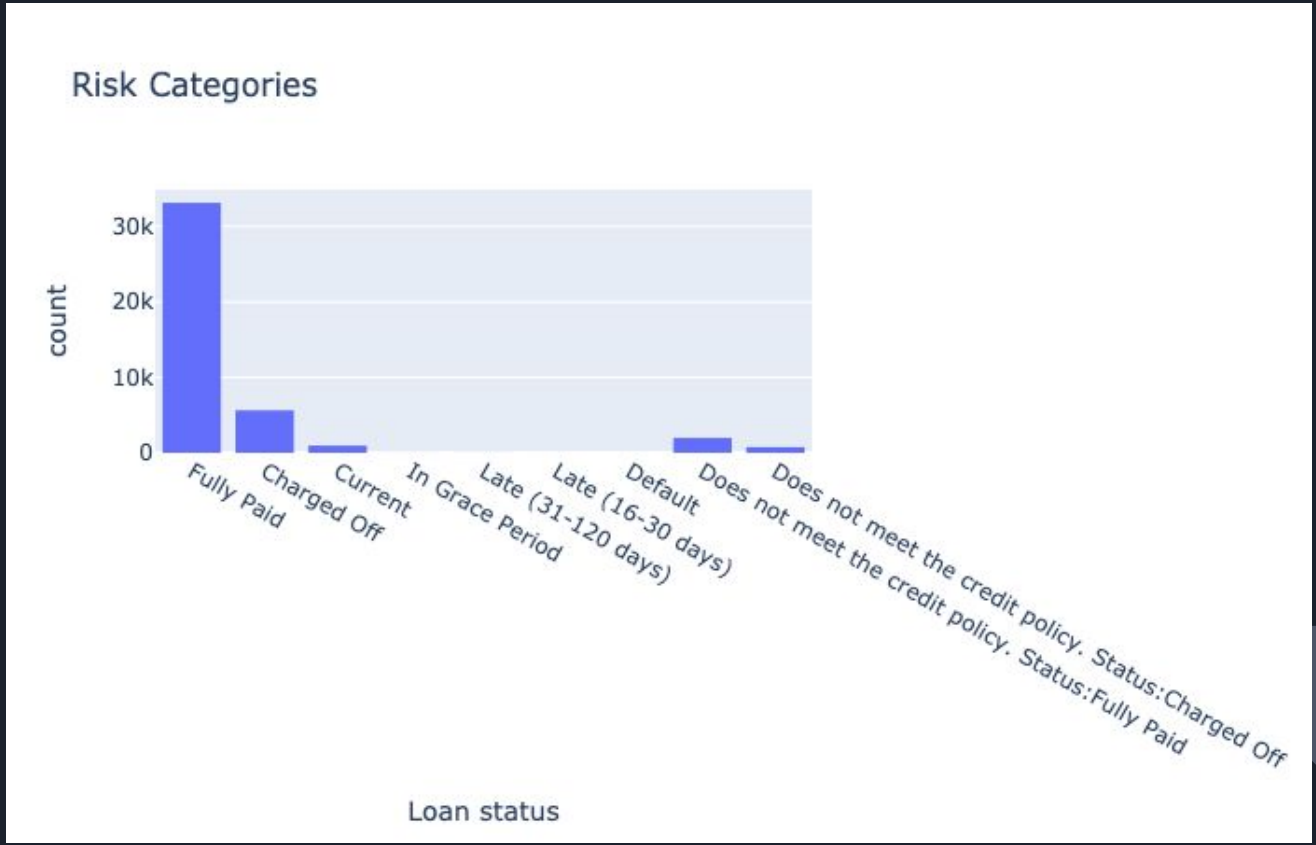
Identifying the business problem

- Lending Club is a Peer-to-Peer lending company
- Lending Club's model for risk assessment categorizes by assigning them grades

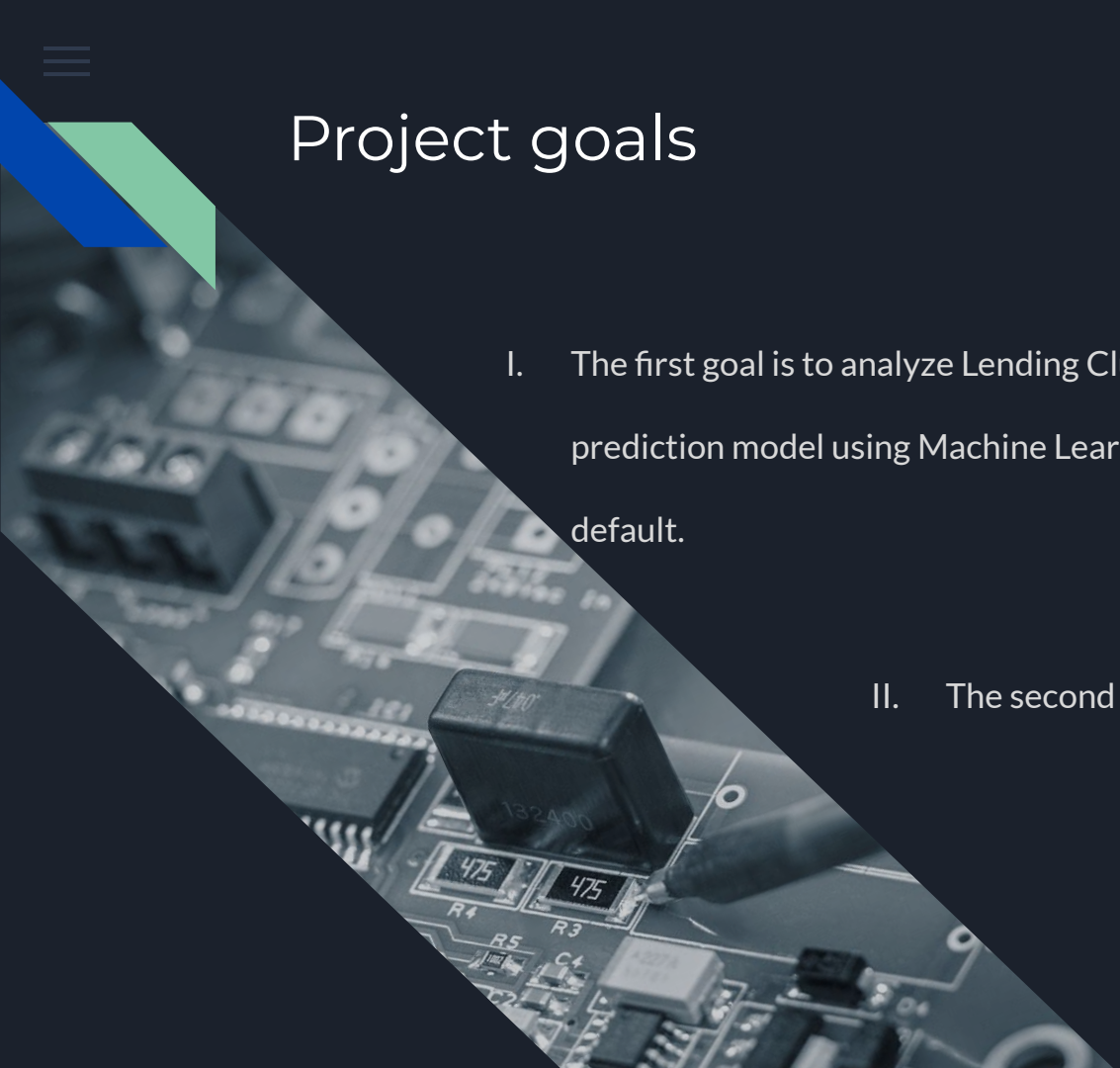
- Investors are presented with a list of borrowers
- They have the opportunity to choose which borrowers they will fund

- ***Investors require more comprehensive assessment of borrowers***
- ***Objective is to identify the risky loan applicants so that cutting down the amount of credit loss***

Different Risk Categories (y)



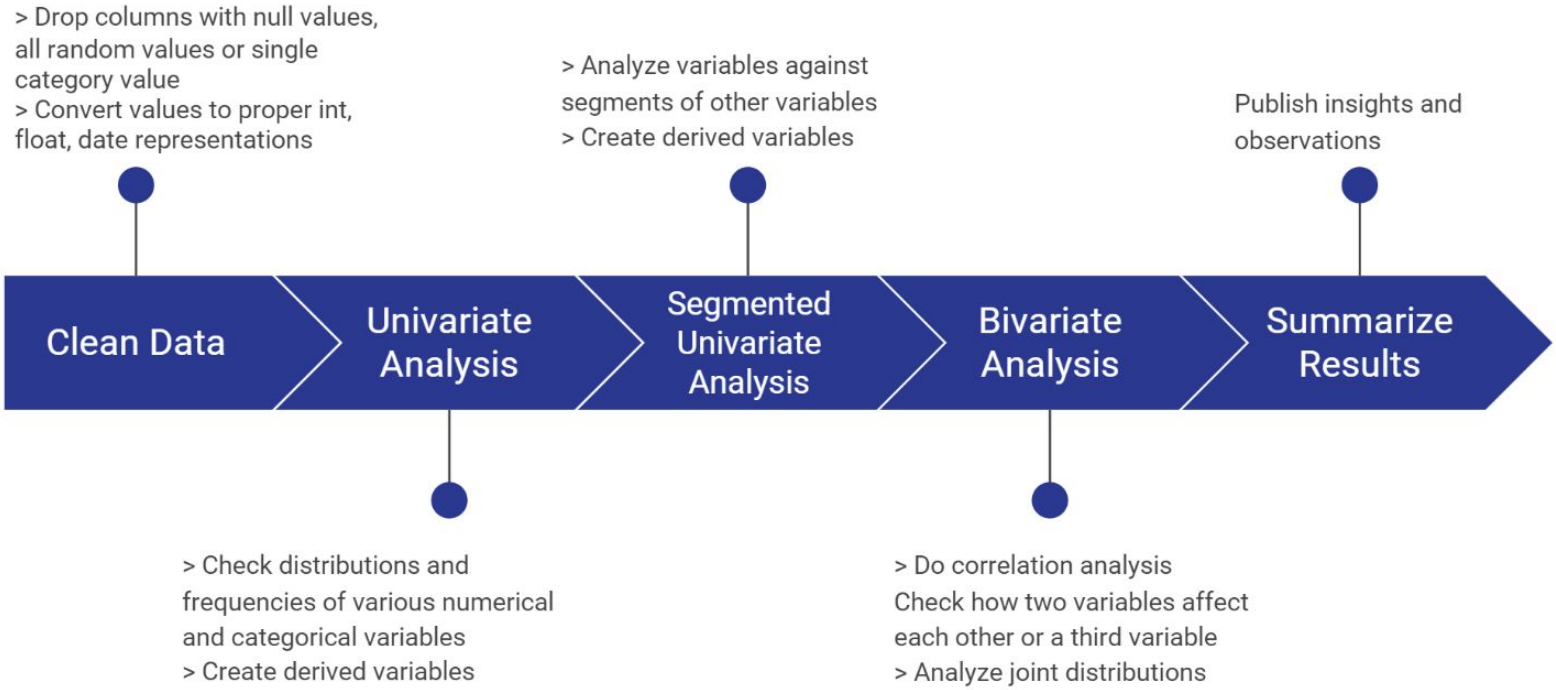
Project goals



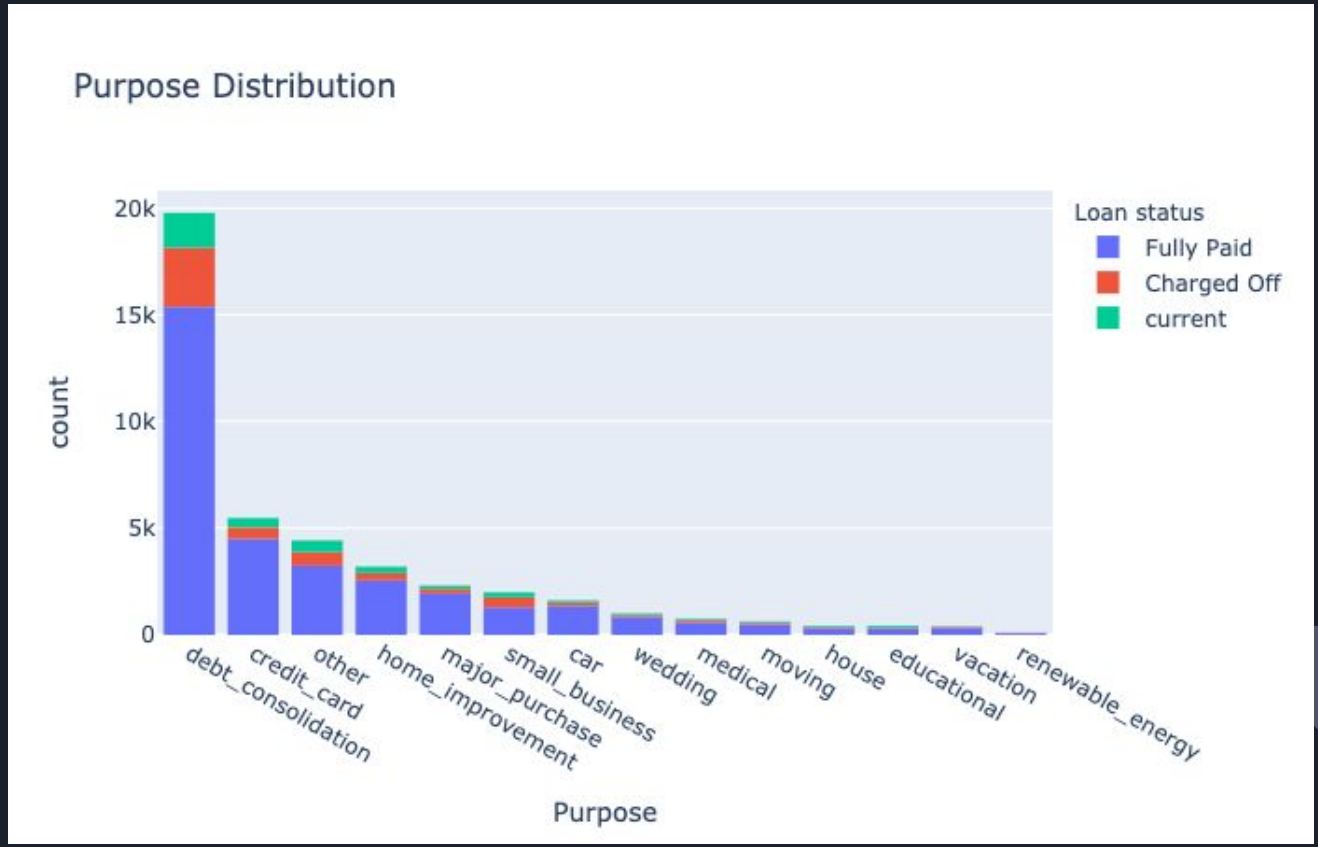
I. The first goal is to analyze Lending Club' s issued loans and to create prediction model using Machine Learning algorithms to predict who might default.

II. The second goal is to learn which loans meet the policy code.

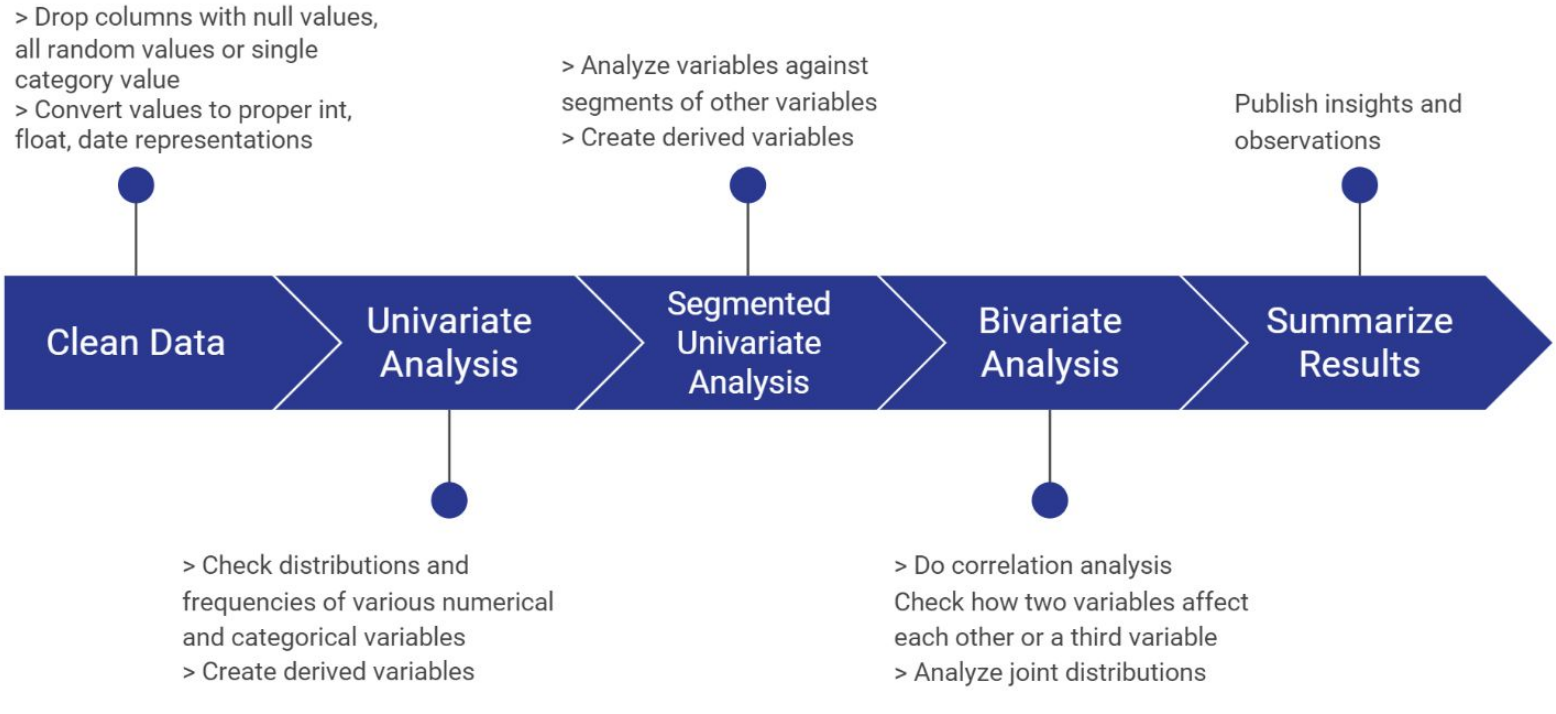
Analysis Approach



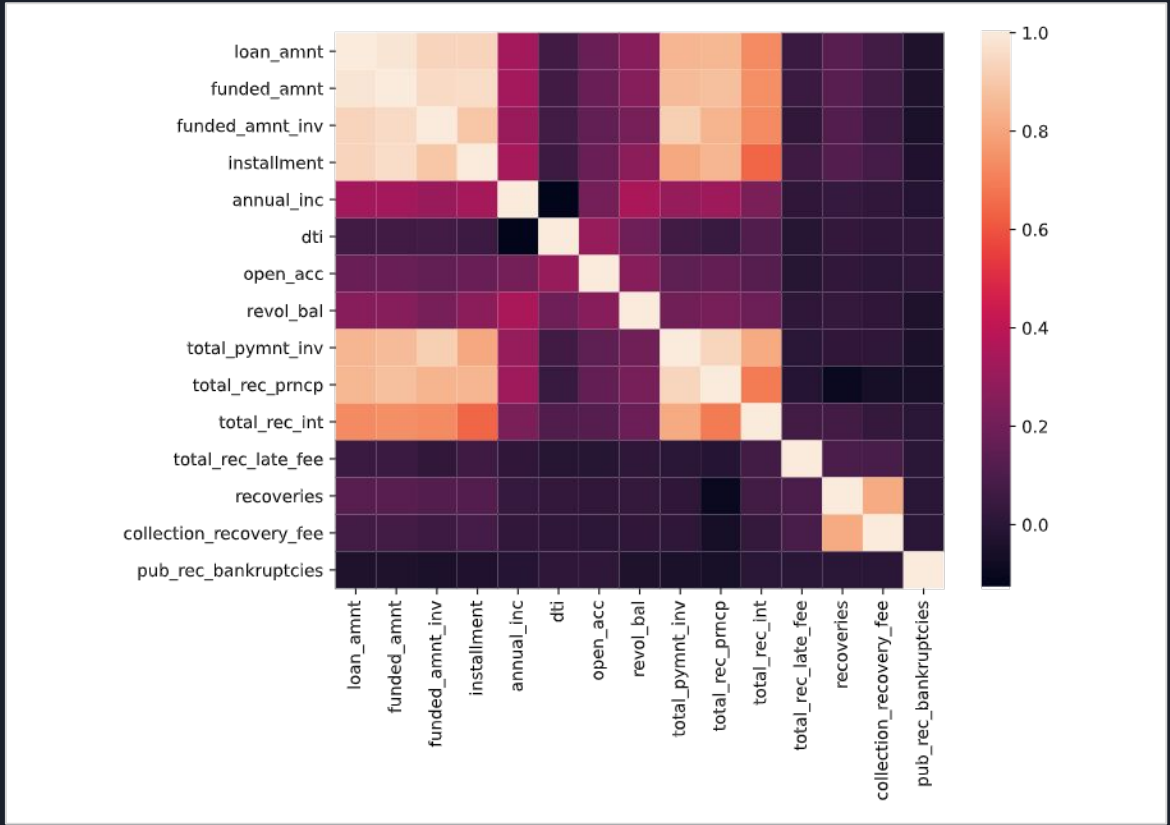
Check Purpose Distribution



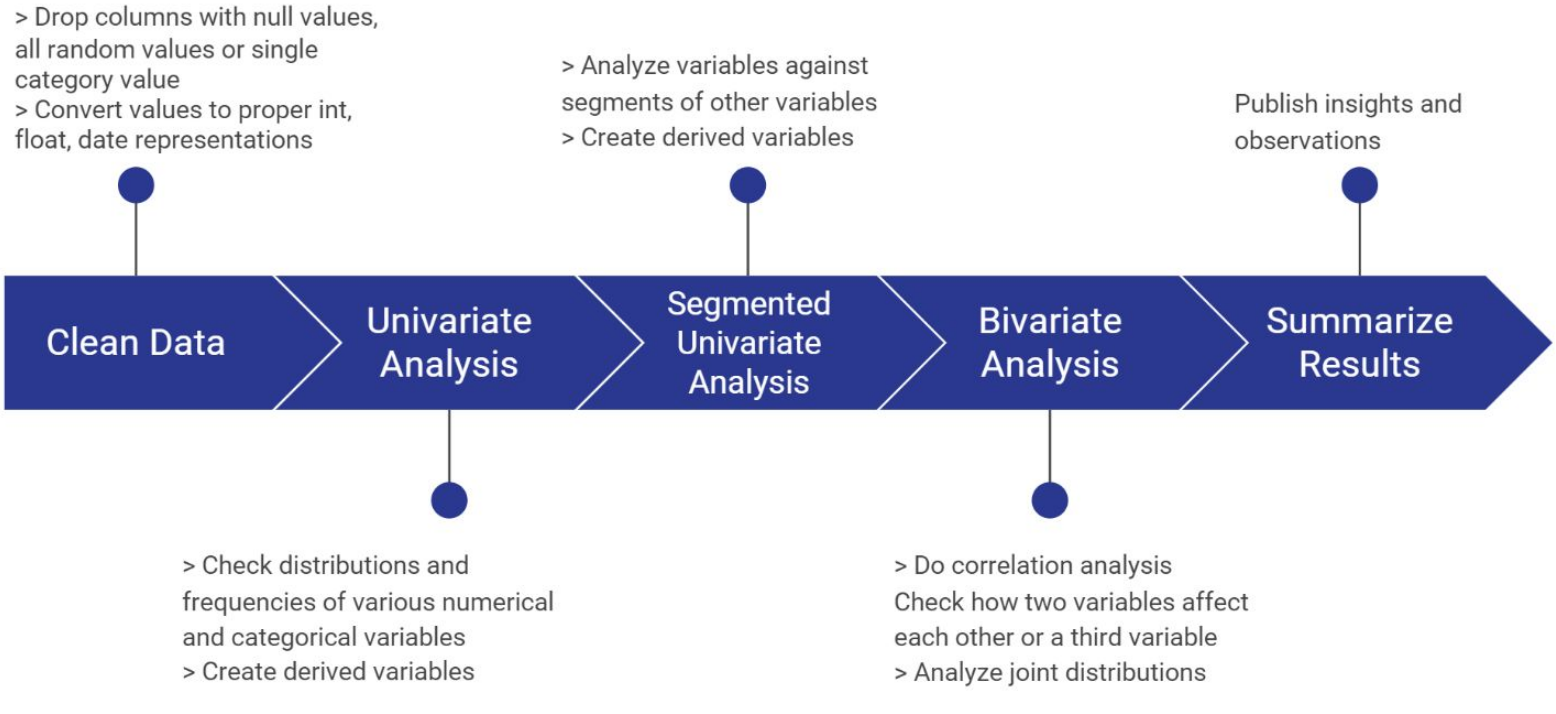
Analysis Approach



Check Feature Correlations



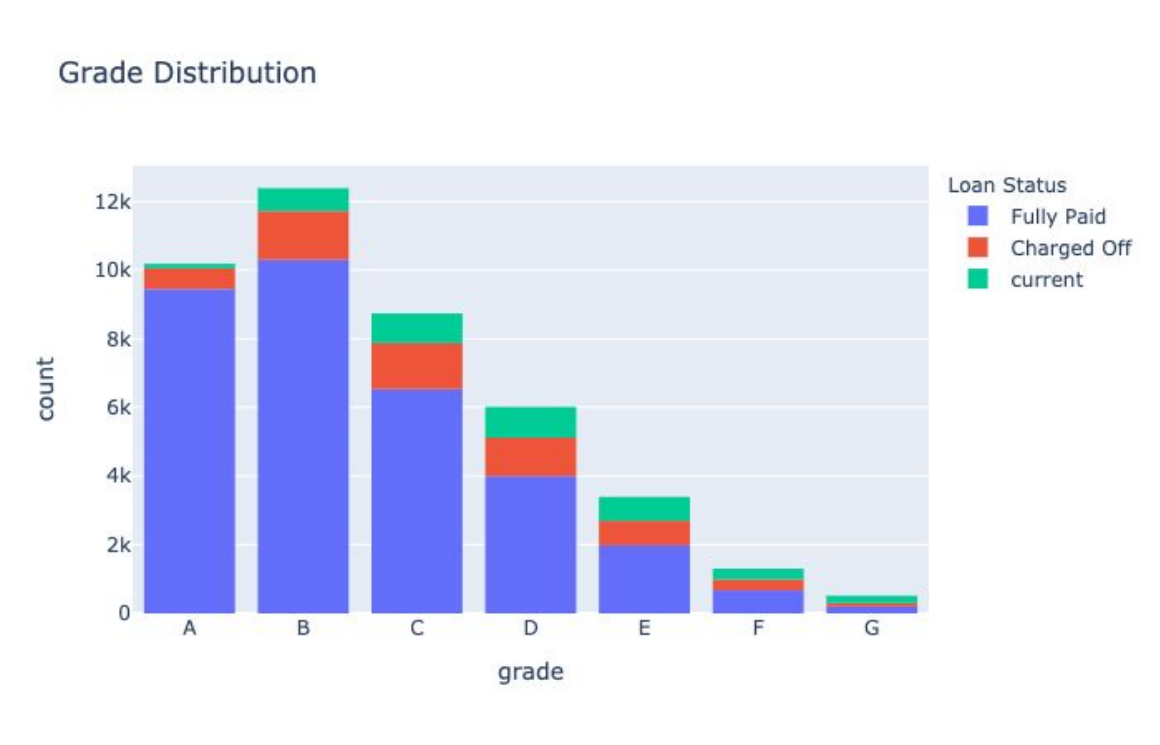
Analysis Approach



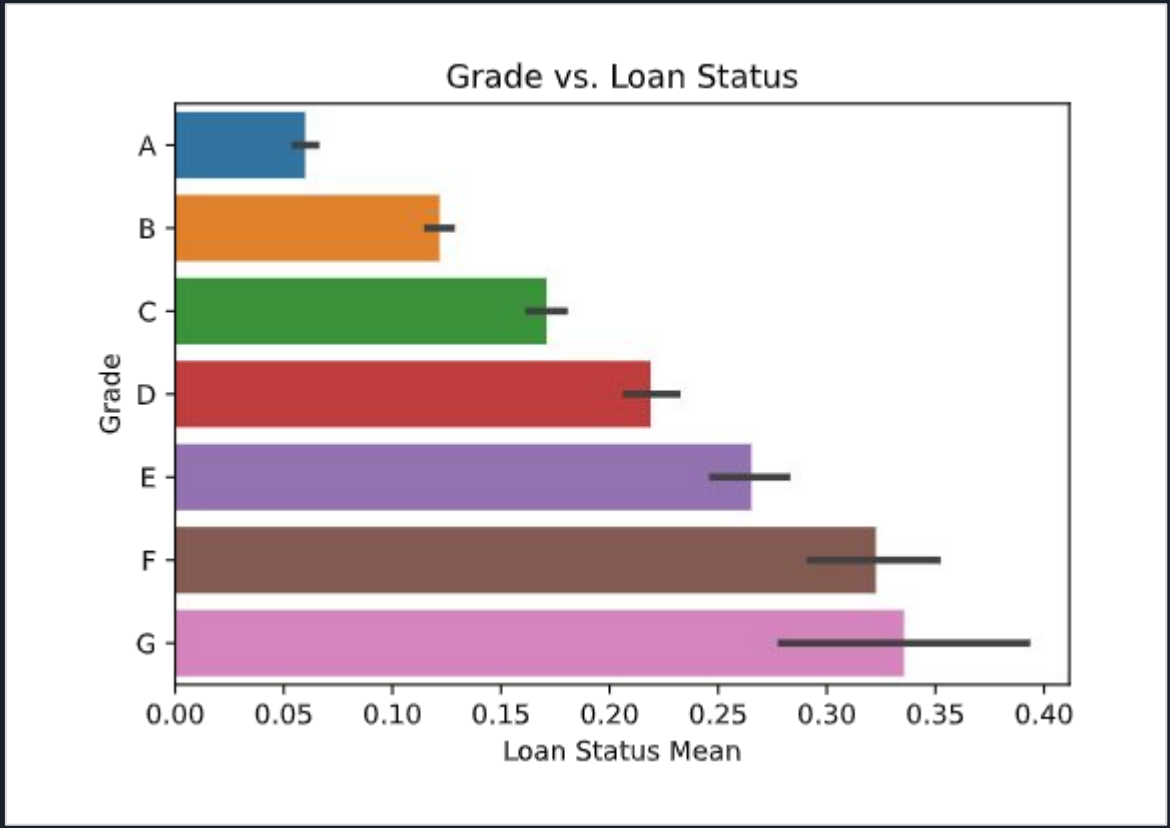
Machine Learning Models

- Logistic Regression
- Decision Tree
- Random Forest

Grade Distribution



Grade vs. Loan Status



Predict Grades

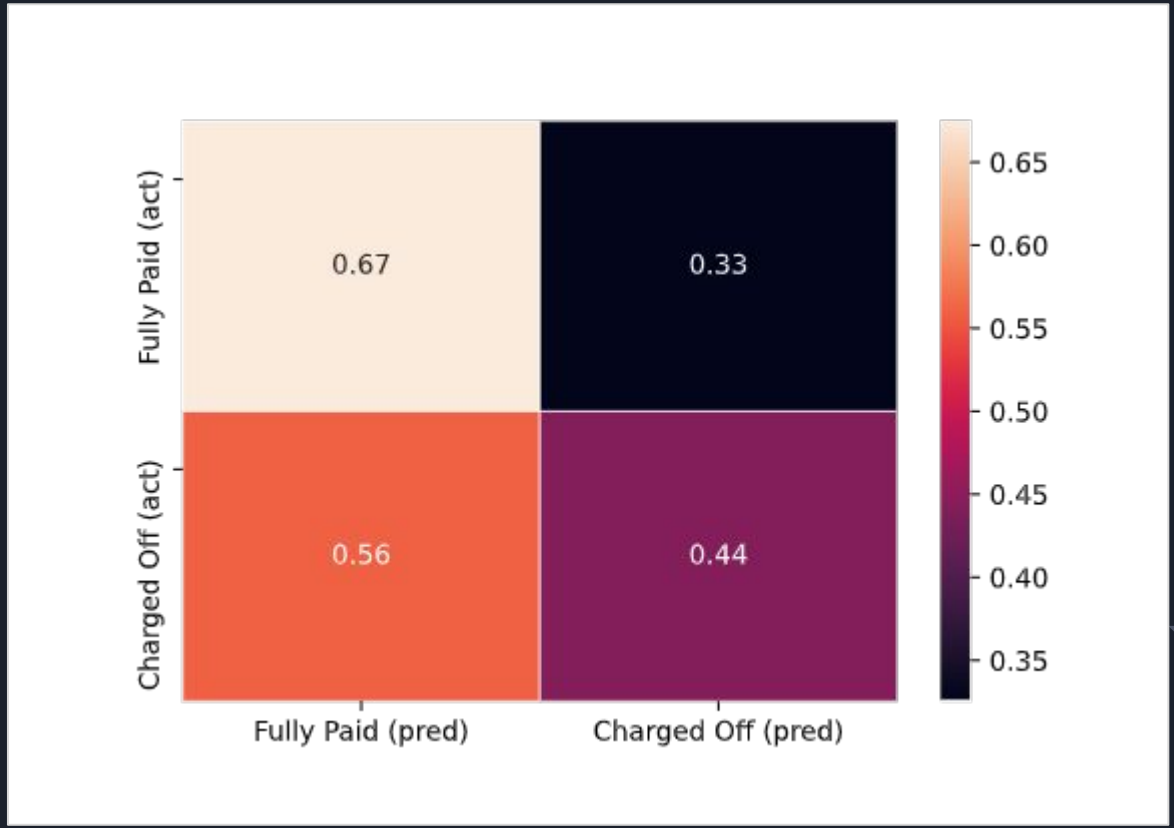
```
[[2512    0    1    0    0    0    0]
 [   0 2923    9    0    0    0    0]
 [   0    8 1916   45    0    0    0]
 [   0    1   11 1258   10    0    0]
 [   0    0    0   18  650    7    0]
 [   0    0    0    0    8  223   17]
 [   0    0    0    0    0    6   69]]

              precision    recall  f1-score   support

           A       1.00      1.00      1.00      2513
           B       1.00      1.00      1.00      2932
           C       0.99      0.97      0.98      1969
           D       0.95      0.98      0.97      1280
           E       0.97      0.96      0.97       675
           F       0.94      0.90      0.92       248
           G       0.80      0.92      0.86        75

    accuracy                           0.99      9692
   macro avg       0.95      0.96      0.96      9692
weighted avg       0.99      0.99      0.99      9692
```

Predict Loan Status

```
Fully Paid                                                  33133
Charged Off                                                  5634
Does not meet the credit policy. Status:Fully Paid           1988
Current                                                       961
Does not meet the credit policy. Status:Charged Off           761
Late (31-120 days)                                             24
In Grace Period                                                20
Late (16-30 days)                                               8
Default                                                         3
Name: loan_status, dtype: int64
```

Naive Estimator: 17 % = Charged Off

Predict Loan Status



Conclusion

I. Prediction the Loan Status was not successfully.

II. Prediction of the Grades was quite successfully.

With some more time:

- Improve the prediction for Loan Status
- Discover conditions for the distribution of the Grades
- Check interest rate if it correctly prices the capital

Vielen Dank für die Blumen!

Euer Karl L. & Yusuf C.