Problem Statement : Based on the data provided, identify risky loan applicants to reduce credit loss to the lending company.

Lending Club Case Study

Solution Approach

- Identify null values in the dataset
- Replace / Drop null values from the dataset
- Remove unwanted columns which are not relevant for the analysis
- Univariate / Bivariate Analysis to find risky propositions

Identifying Null Values

```
[14]  #Identifying null values
      loan_df_nulls = loan_df.isna().sum()
      loan_df_nulls

      id                              0
      member_id                       0
      loan_amnt                       0
      funded_amnt                     0
      funded_amnt_inv                 0
                                    ...
      tax_liens                      39
      tot_hi_cred_lim             39717
      total_bal_ex_mort           39717
      total_bc_limit              39717
      total_il_high_credit_limit  39717
      Length: 111, dtype: int64
```

```
[17]  #Drop columns having null values greater than 30000
      loan_df.drop(loan_df_nulls[loan_df_nulls > 30000].index, axis=1, inplace=True)
```

Drop Null Values

Remove unwanted columns

```
[22]  #Removing columns as per data dictionary which wont make sense from analysis perspective
      cols_del = ['member_id', 'url', 'title', 'emp_title', 'out_prncp_inv', 'zip_code',
                  'last_credit_pull_d', 'total_rec_late_fee', 'delinq_2yrs',
                  'last_pymnt_d', 'revol_bal', 'out_prncp', 'total_pymnt',
                  'total_rec_int', 'total_rec_prncp', 'collection_recovery_fee',
                  'last_pymnt_amnt','recoveries', 'earliest_cr_line', 'addr_state']

      loan_df.drop(cols_del, axis = 1, inplace = True)
      #loan_df['chargeoff_within_12_mths']
      #loan_df['collections_12_mths_ex_med']
      #loan_df['tax_liens']
```

```
[26] loan_df = loan_df[loan_df.loan_status != "Current"]

[27] loan_df.shape

     (38577, 26)
```
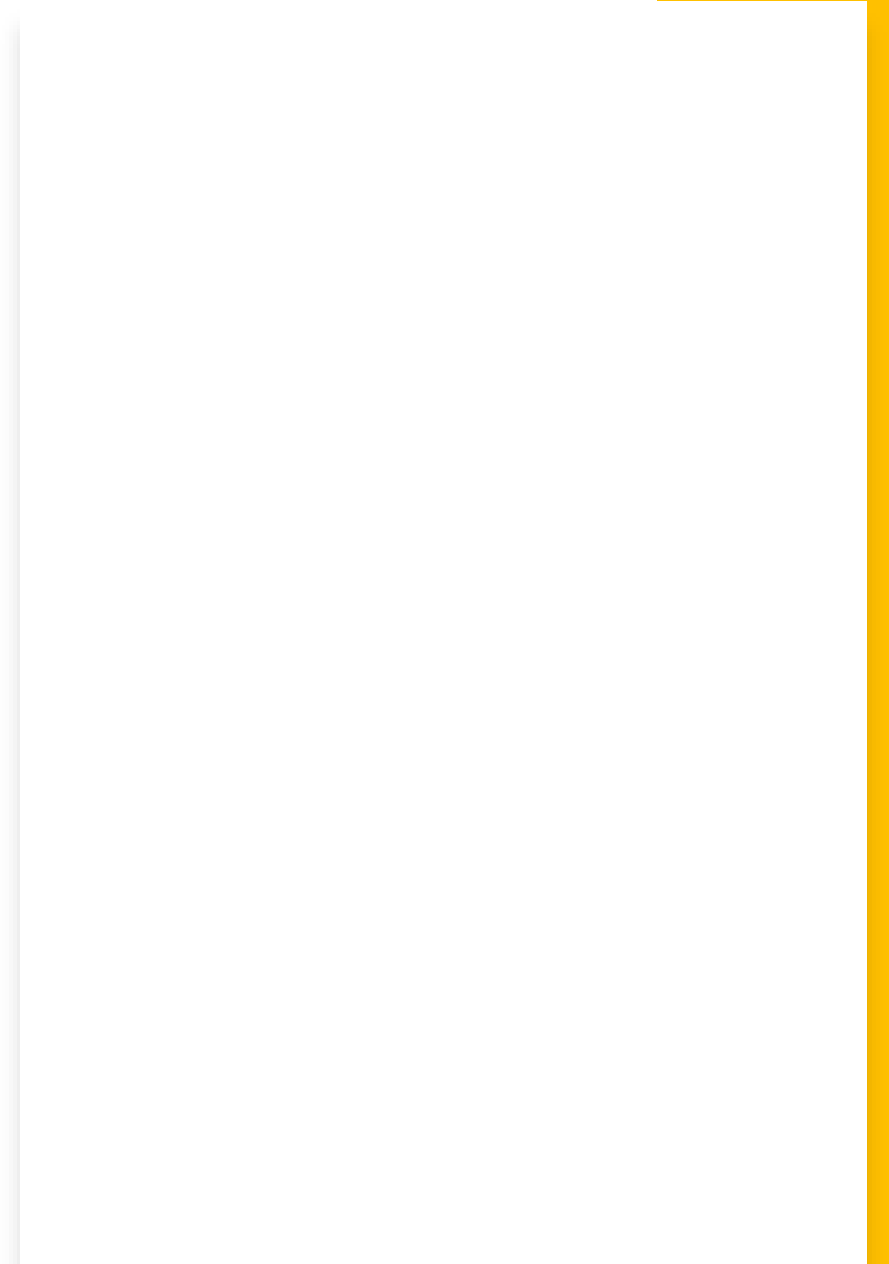
Drop Current Status records

Outcome of the analysis

"Small business" is a risky proposition

loan_status purpose	Charged Off	Fully Paid	All	Charged Off Percentage
car	160	1339	1499	10.674
credit_card	542	4485	5027	10.782
debt_consolidation	2767	15288	18055	15.325
educational	56	269	325	17.231
home_improvement	347	2528	2875	12.070
house	59	308	367	16.076
major_purchase	222	1928	2150	10.326
medical	106	575	681	15.565
moving	92	484	576	15.972
other	633	3232	3865	16.378
renewable_energy	19	83	102	18.627
small_business	475	1279	1754	27.081
vacation	53	322	375	14.133
wedding	96	830	926	10.367

Loans with "Higher Interest Rates" are a risky proposition

int_rate_buckets	loan_status	Charged Off	Fully Paid	All	Charged Off Percentage
5-10		830	11486	12316	6.739
10-15		2707	15558	18265	14.821
15-20		1794	5432	7226	24.827
20 and above		296	474	770	38.442

Loan amounts greater than 30K are a risky proposition

loan_status loan_amt_buckets	Charged Off	Fully Paid	All	Charged Off Percentage
0-5K	1314	8158	9472	13.872
5K-10K	1642	11160	12802	12.826
10K-15K	1055	6628	7683	13.732
15K-20K	751	3598	4349	17.268
20K-25K	542	2294	2836	19.111
25K-30K	143	557	700	20.429
30K and above	180	555	735	24.490

Loan Periods greater than 36 months are a risky proposition

loan_status	Charged Off	Fully Paid	All	Charged Off Percentage
term				
36 months	3227	25869	29096	11.091
60 months	2400	7081	9481	25.314

Thank You