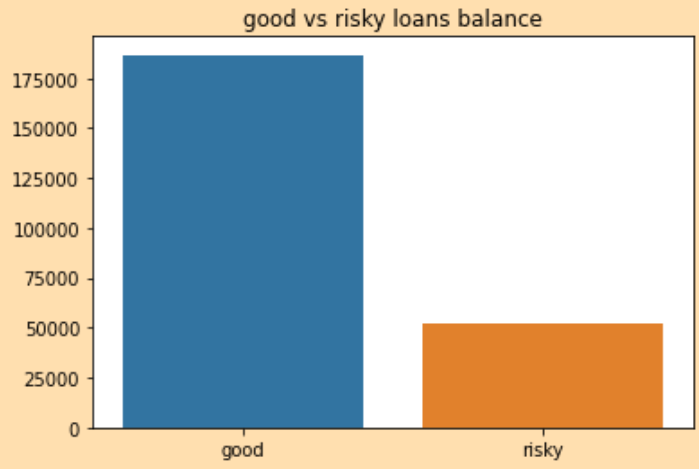
CREDIT RISK & SCORECARD MODEL

a lending company loan data issued from 2007 - 2014 with almost 75 features related to the borrowers. We want to develop a credit risk model in Python to predict the PD model and assign scorecard.

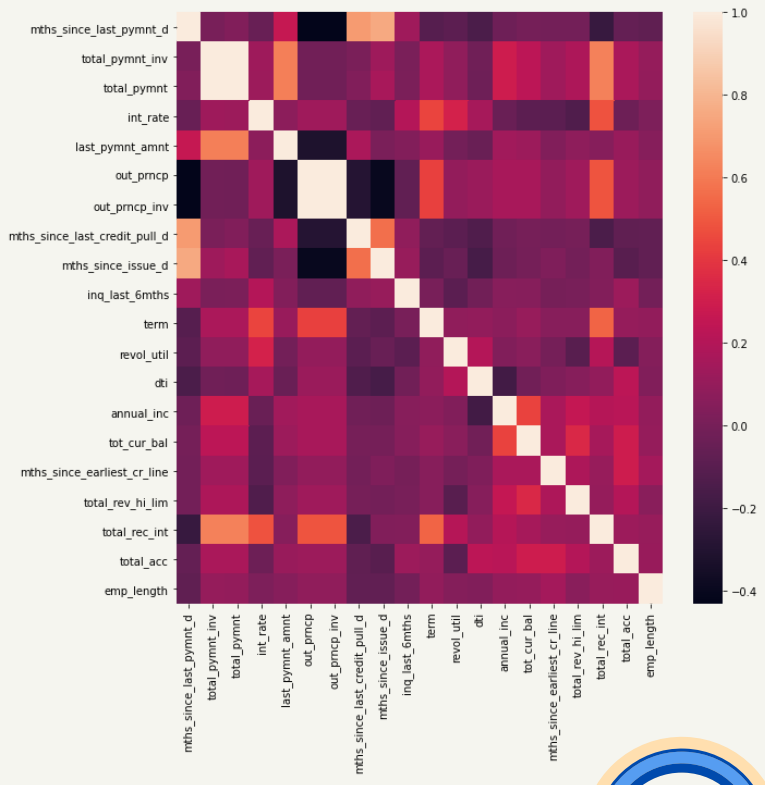
GOOD & RISKY LOANS

Our target variable is loan_status features. It contains of various values which represents by good (high values) and risky loans balance.



FEATURES SELECTION

A heat-map of these pair-wise correlations identifies two columns (out_prncp_inv and total_pymnt_inv) as highly correlated. It means that features should be removed.



WOE BY GRADE

There is a continuous increase in WoE by grades. Therefore, we do not need to combine it with any features together.

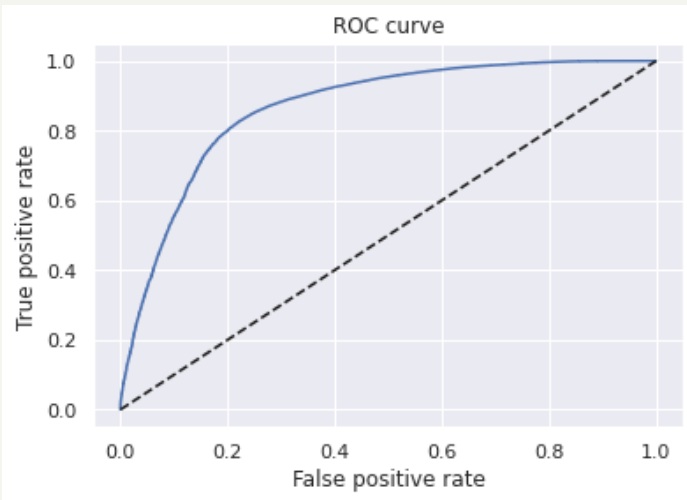
CREDIT RISK SCORECARD

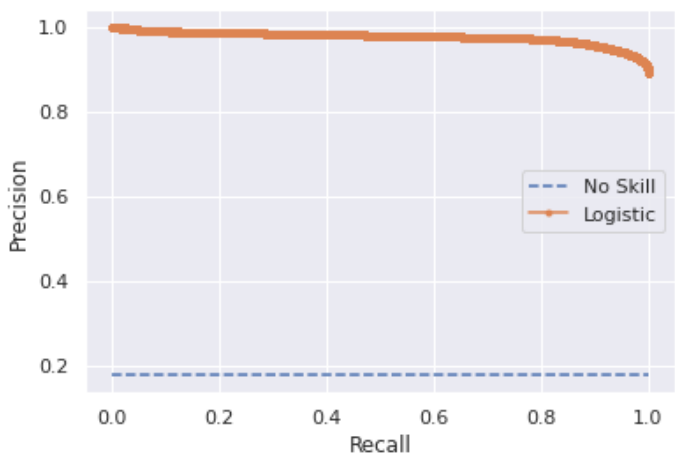
For calculate an individual's credit score we used summary table train test and scaling future categories.

Feature name	Coefficients	Original feature name	Score - Calculation	Score - Preliminary
Intercept	2.948892	Intercept	598.515609	599.0
grade:A	0.960200	grade	24.463996	24.0
grade:B	0.792926	grade	19.789993	20.0
grade:C	0.609305	grade	15.207145	15.0
grade:D	0.487386	grade	12.164269	12.0
grade:E	0.331929	grade	8.284354	8.0
grade:F	0.190615	grade	4.757402	5.0
home_ownership:OWN	-0.048374	home_ownership	-1.207327	-1.0
home_ownership:OTHER_NONE_RENT	-0.104195	home_ownership	-2.600515	-3.0
verification_status:Source Verified	-0.281427	verification_status	-7.023912	-7.0
verification_status:Verified	-0.459417	verification_status	-11.466206	-11.0
purpose:debt_consolidation	-0.315579	purpose	-7.876281	-8.0
purpose:credit_card	-0.222064	purpose	-5.542316	-6.0
purpose:educ__ren_en__sm_b__mov	-0.481318	purpose	-12.012820	-12.0
purpose:vacation__house__wedding__med__oth	0.003981	purpose	0.099355	0.0
term:36	-0.094122	term	-2.349118	-2.0
int_rate:<7.071	0.955408	int_rate	23.845255	24.0
int_rate:7.071-10.374	0.285426	int_rate	7.123724	7.0
int_rate:10.374-13.676	0.056631	int_rate	1.413405	1.0
int_rate:13.676-15.74	0.039281	int_rate	0.980393	1.0

AUROC ON TEST SET

Based on the graph below, AUROC on test set calculated comes out to 0.866 with a Gini of 0.732, both being considered as quite acceptable evaluation scores.





References
How to Develop a Credit Risk Model and Scorecard
https://towardsdatascience.com/how-to-develop-a-credit-risk-model-and-scorecard-91335fc01f03