Peer2Peer Lending

HOMER KAY

What is peer-to-peer lending?

Traditional Bank Financing

High credit costs
Lack of available credit
Hidden fees

High cost base
(branches, ATM, etc.)

Peer-to-Peer Lending

Competitive rates
Access to credit
Transparent terms

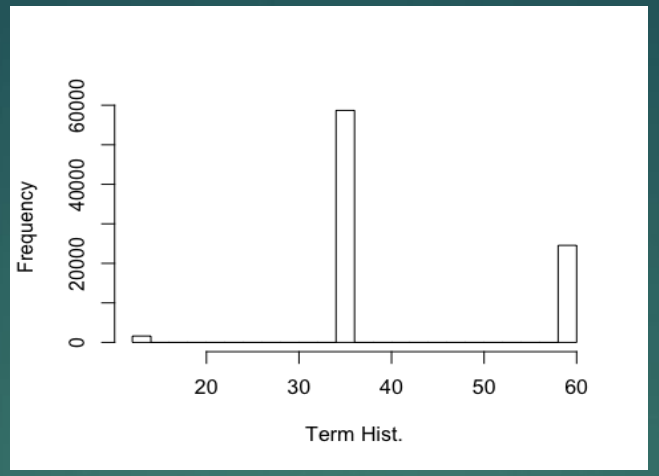
Low cost base

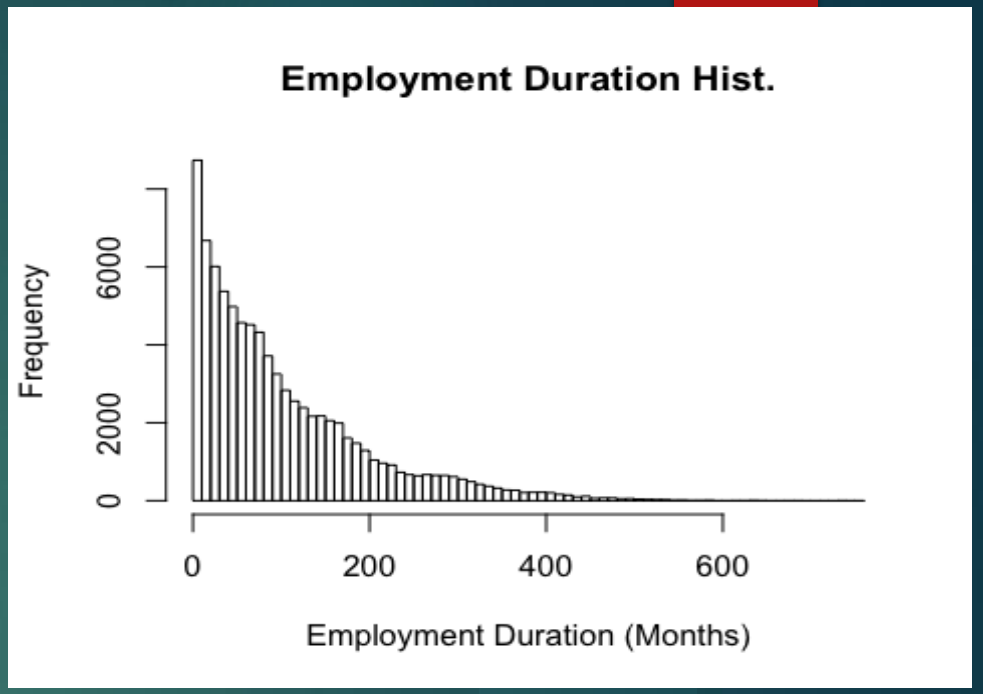
Business Case(s)

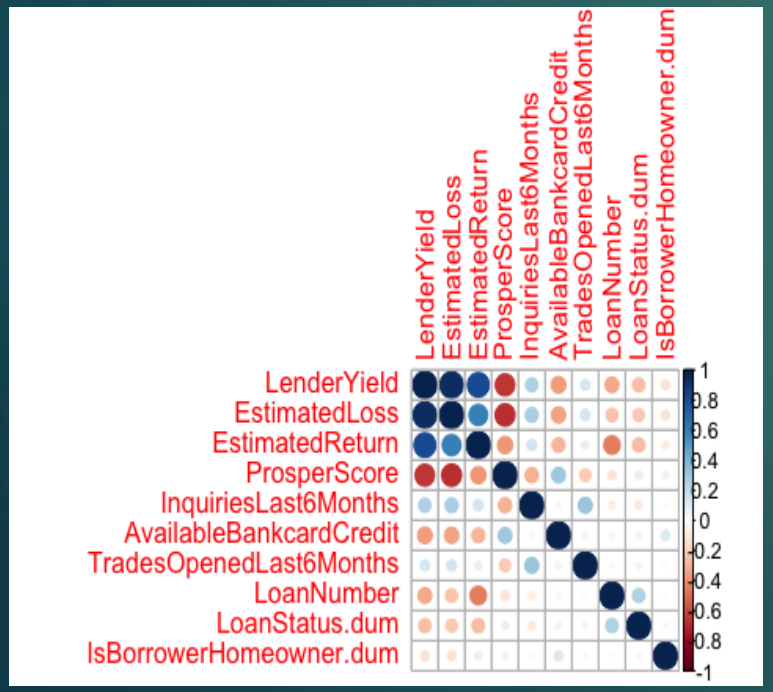
- 1. Predict Numerical "Prosper Scores" so that buyers and sellers may be better informed of the risk placed on themselves and determine if their estimated yield (how much they make/spend) is worth that burden.
  - Regression Techniques
  - Calculations of this are currently proprietary information.
- 2. Understand what characteristics lead a lessee to defaulting on their loan.
  - Classification Techniques

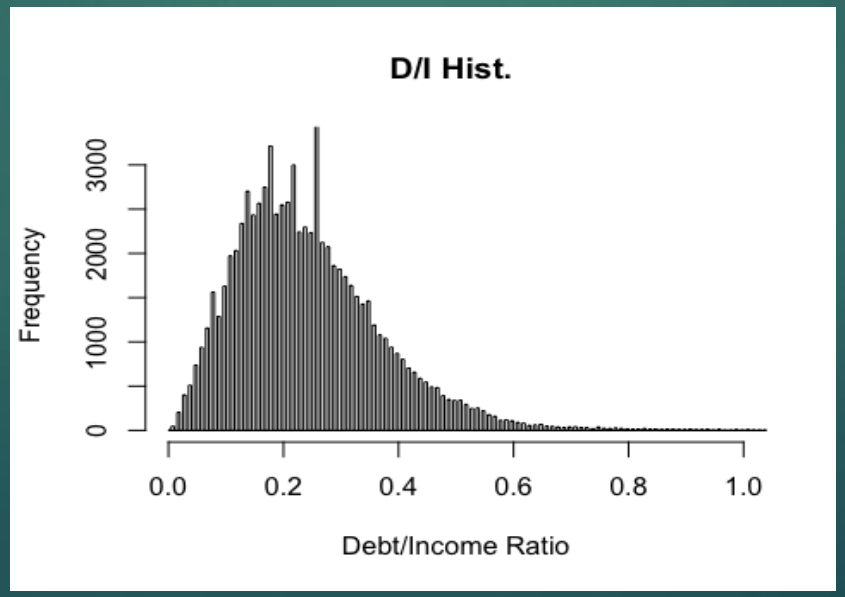
Data Overview

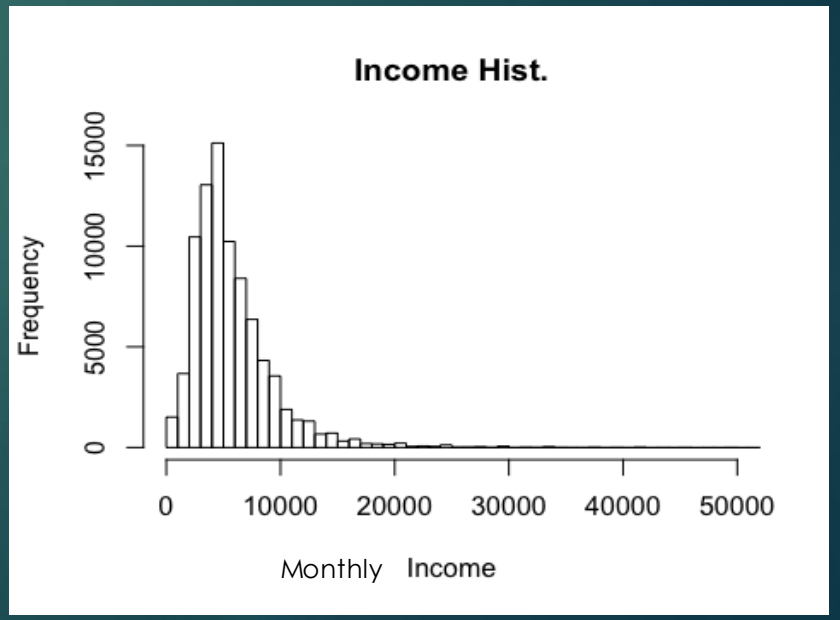
- 114k rows, 81 variables.
- Categorical and Numerical





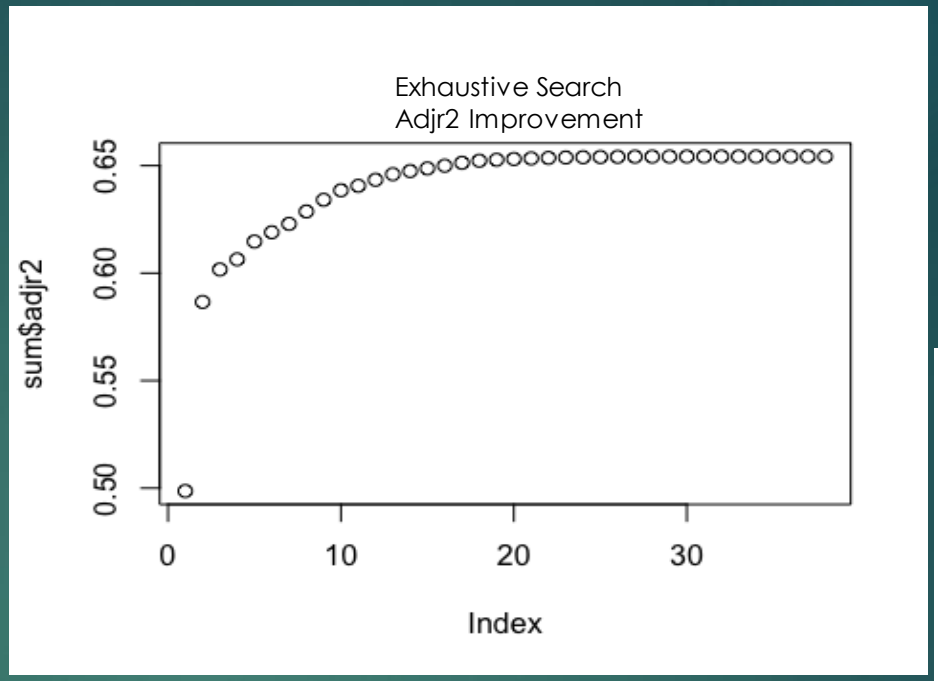


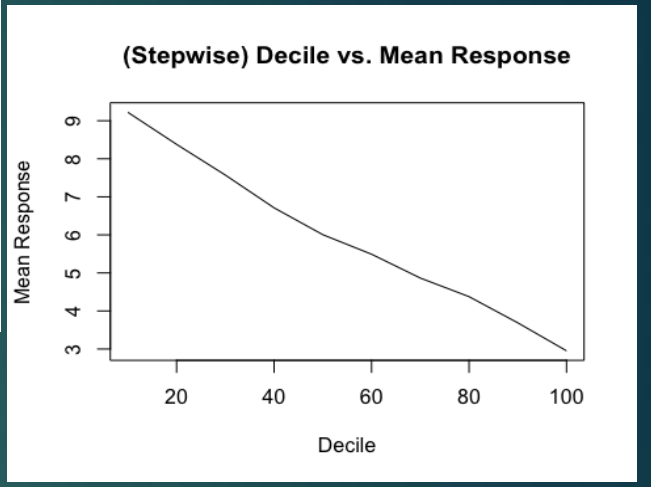


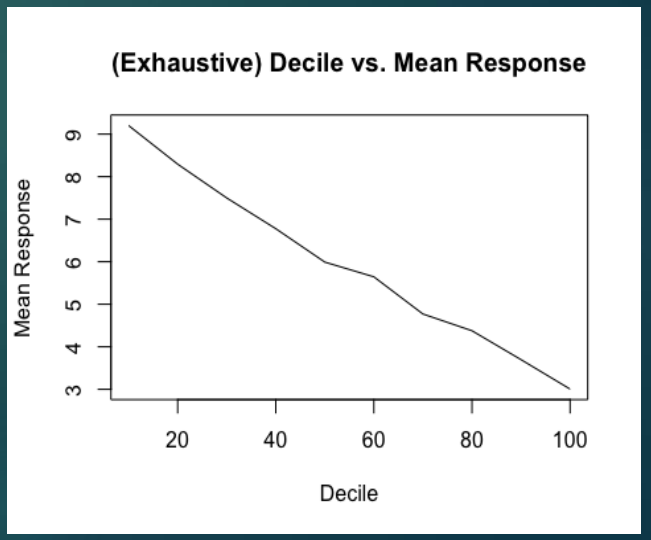


Prosper Score

- Stepwise reg. resulted in lowest Errors. With 13 variables.
- Exhaustive was close with 10 variables.



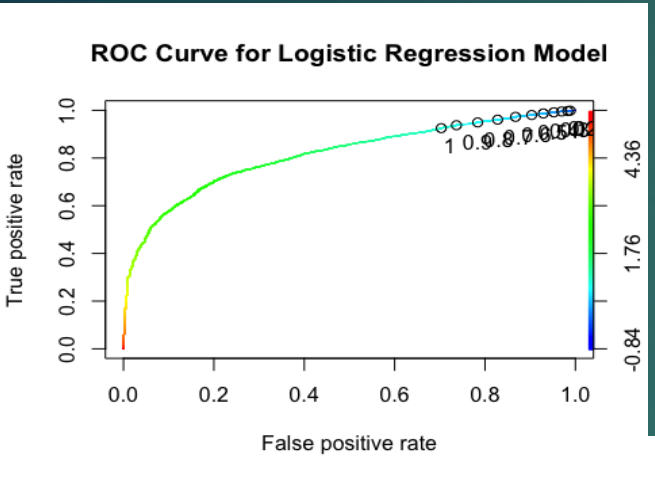


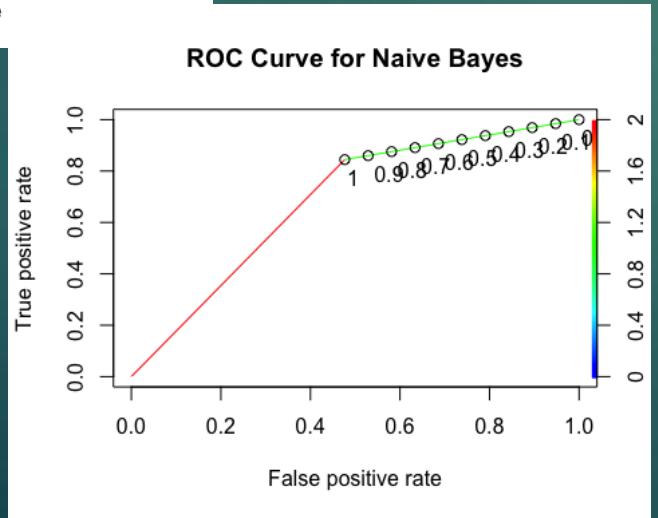


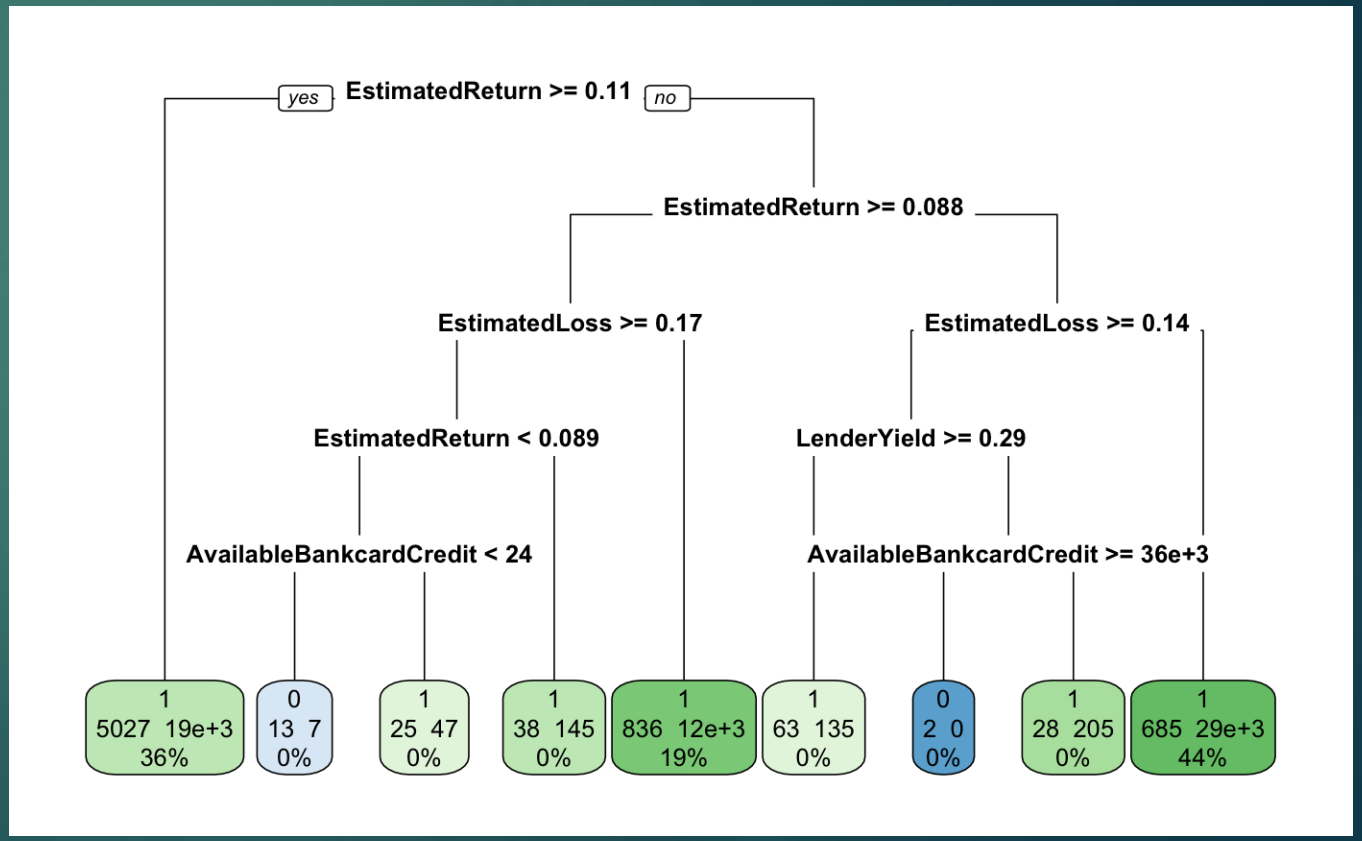
	Train.LM	Train.LM.2	Stepreg.train	Train.LM.3	Exh.lm
AdjR2	0.6579	0.6576	0.658	0.5099	0.6385
Multi R2	0.6582	0.6578	0.6582	0.5099	0.6386
ME			-0.023	-0.028	-0.029
RMSE			1.396	1.68	1.439
MAE			1.089	1.388	1.120
MPE			-10.176	-14.34	-10.86
MAPE			25.76	33.14	26.6
# of Variables used	48	40	13	8	10

Classification Results

	Tru Po	Tru Neg	Fal Po	Fal Neg	Accuracy rate	Sensitivity	Specificity	Precision	# of Variables
Logit Reg 1	15241	23	1668	39	89.94%	99.74%	1.36%	90.14%	13
Logit Reg 2	15240	21	1670	40	89.92%	99.74%	1.24%	90.12%	10
Naïve Bayes	13088	806	885	2192	81.87%	85.65%	47.66%	93.67%	10
Decision Tree	15278	1	1690	2	90.03%	99.99%	0.06%	90.04%	10

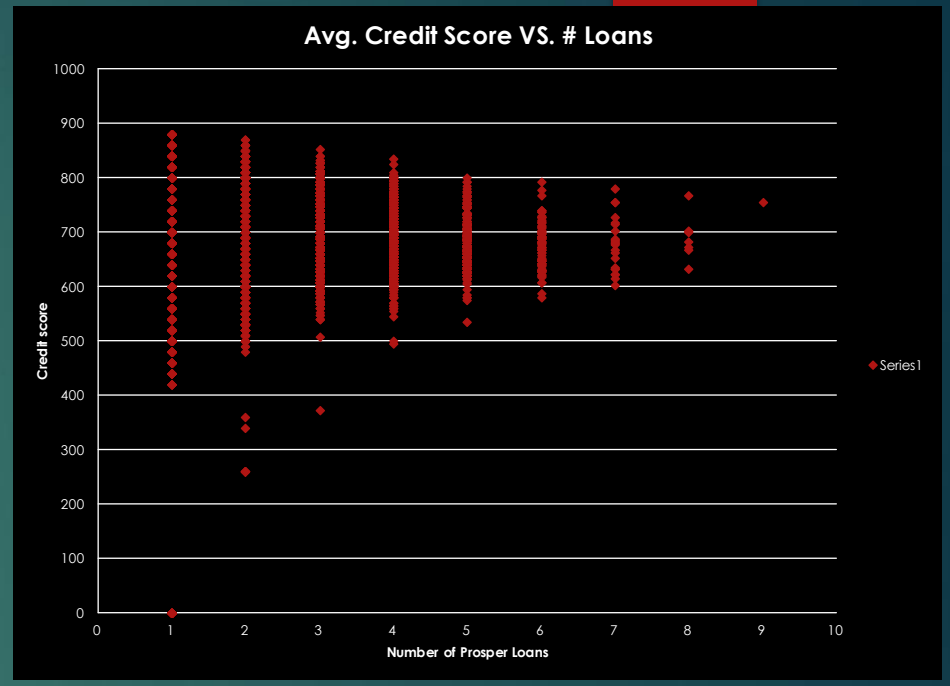






Conclusion

	Train.LM	Train.LM.2	Stepreg.train	Train.LM.3	Exh.lm
AdjR2	0.6579	0.6576	0.658	0.5099	0.6385
Multi R2	0.6582	0.6578	0.6582	0.5099	0.6386
ME			-0.023	-0.028	-0.029
RMSE			1.396	1.68	1.439
MAE			1.089	1.388	1.120
MPE			-10.176	-14.34	-10.86
MAPE			25.76	33.14	26.6
# of Variables used	48	40	13	8	10



	Tru Po	Tru Neg	Fal Po	Fal Neg	Accuracy rate	Sensitivity	Specificity	Precision	# of Variables
Logit Reg 1	15241	23	1668	39	89.94%	99.74%	1.36%	90.14%	13
Logit Reg 2	15240	21	1670	40	89.92%	99.74%	1.24%	90.12%	10
Naïve Bayes	13088	806	885	2192	81.87%	85.65%	47.66%	93.67%	10
Decision Tree	15278	1	1690	2	90.03%	99.99%	0.06%	90.04%	10