

Predicting Lending Club Defaults: Can we do Better?

Will Sorenson

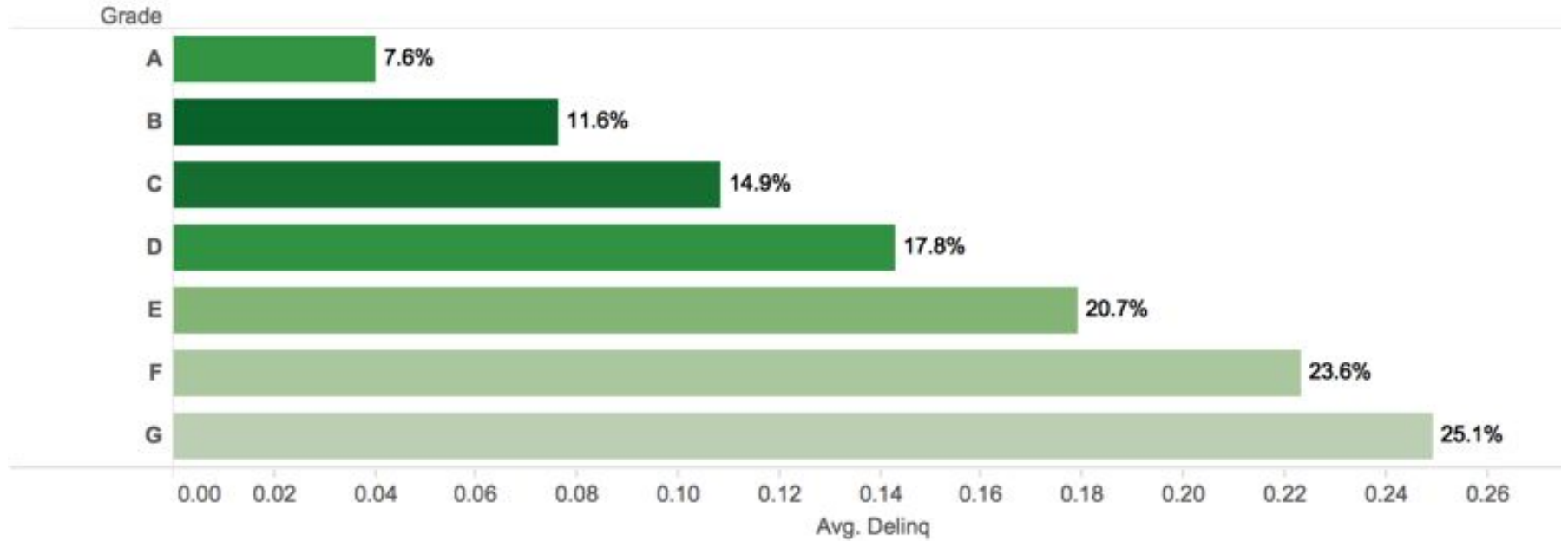
Basics of Lending Club

- Peer to Peer Lending Platform
- Median Returns for Investors ~ 7%
- 99.9% has positive nominal returns when diversified among 100 loans (costs \$2500)
- Extensive Access to data: State, employment, FICO, income, housing status, other debt, and much more.

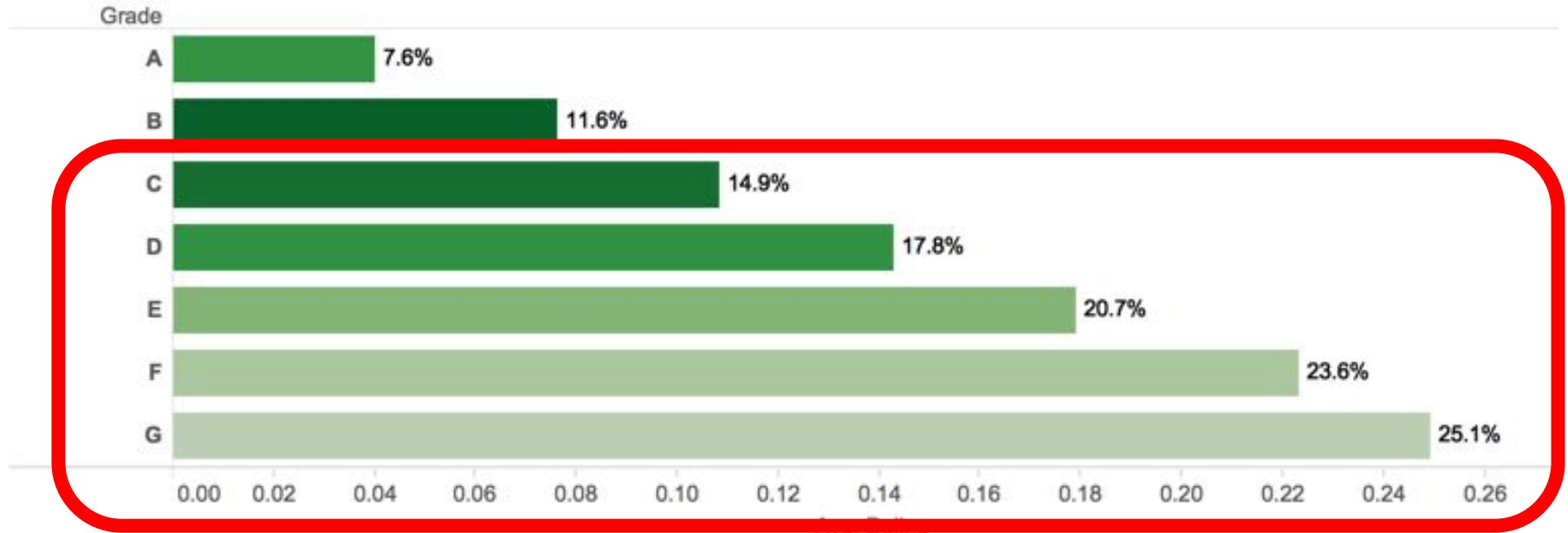
Predicting Defaults

- Predict Defaults better -> higher returns
- Lending Club Datasets
- Lending Club API

The Grade System



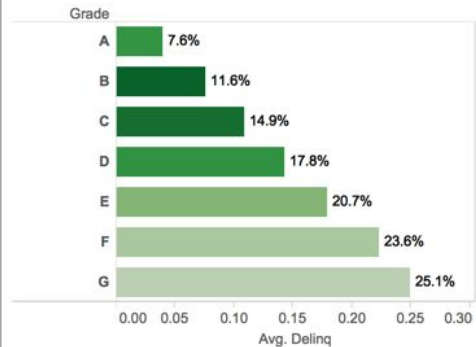
The Grade System



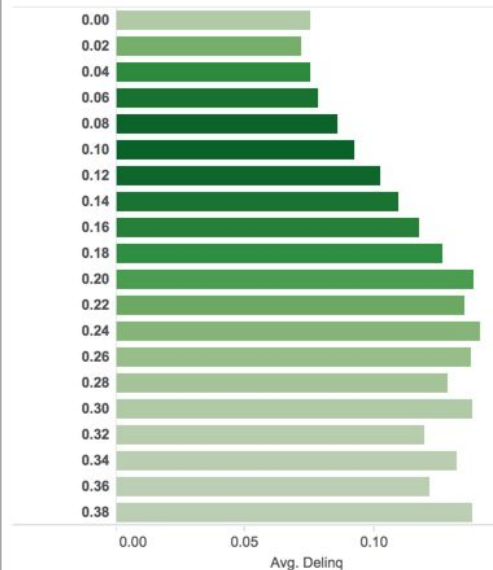
Can we do better?

1. Engineer New Features
2. Predict Defaults with our new models (Random Forests Won)
3. Pick out the loans (test set) from the riskier categories based on (2)
4. Evaluate performance

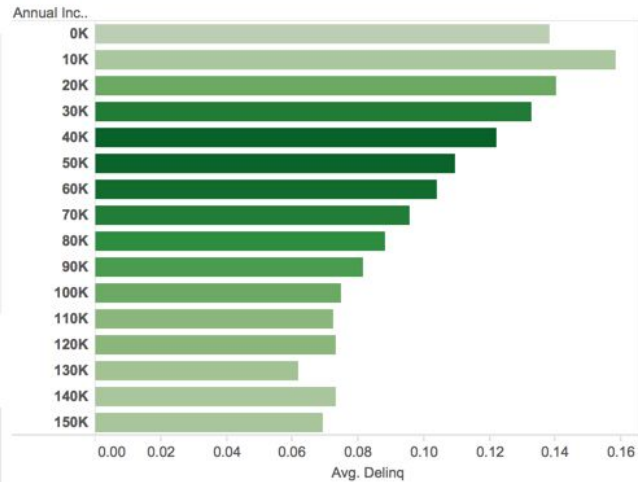
Grade



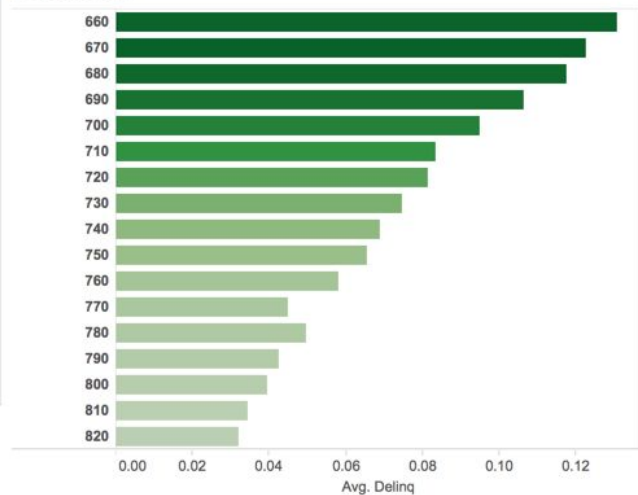
Debt Payments to Gross Income Ratio



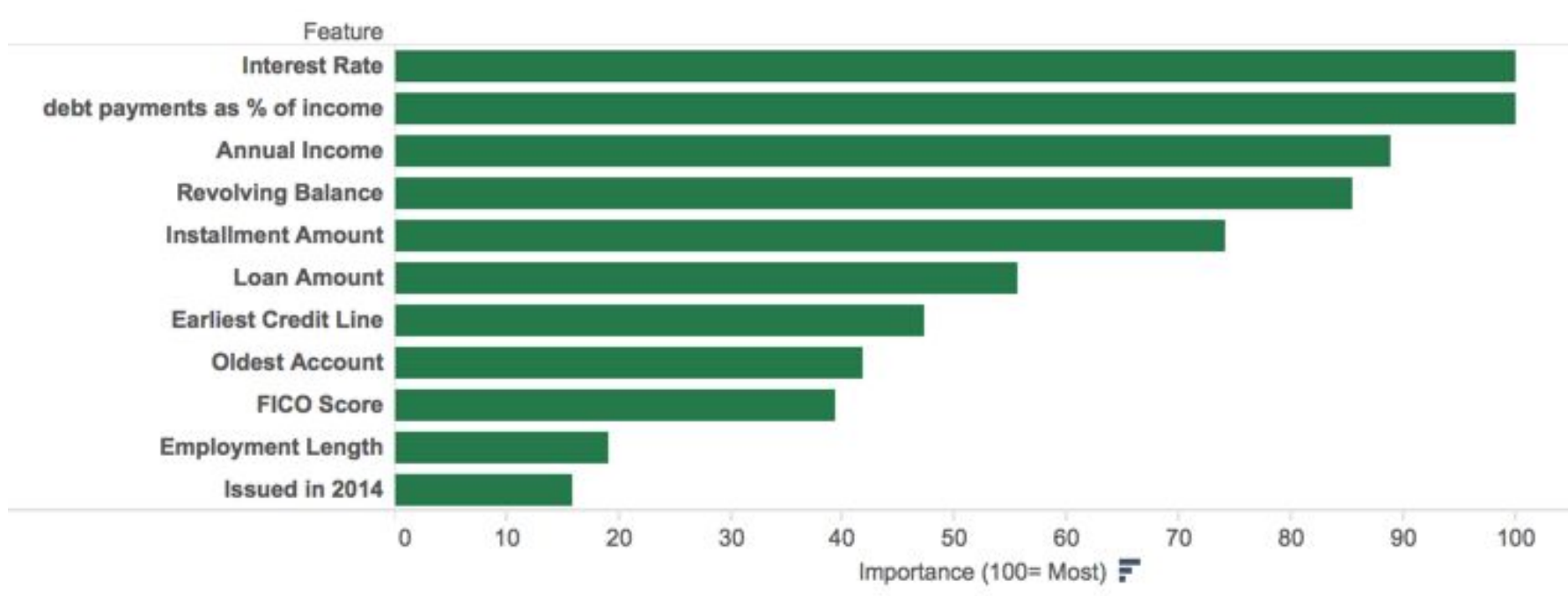
Income



FICO Scores



The Most Important Features



Yes, We Can Do Better

- Default rate decreases from 12% in the entire test set to 6% in our model's favorites (top 20%)
- Represents increase ROI from 9% to 13.5%

Backup