Predicting Loan Defaults

Ahsan Ahmed Raunak Sharma Scott Silverstein Rylan Tribush

The Heart of Home Credit

Serving the unbanked

Empowering people to enter the broader global economy

Financial inclusion

Focusing on delivering greater access to financial services through financial inclusion efforts across our markets

Global partnerships

Working with dynamic e-commerce and manufacturing partners across our markets

Enabling growth

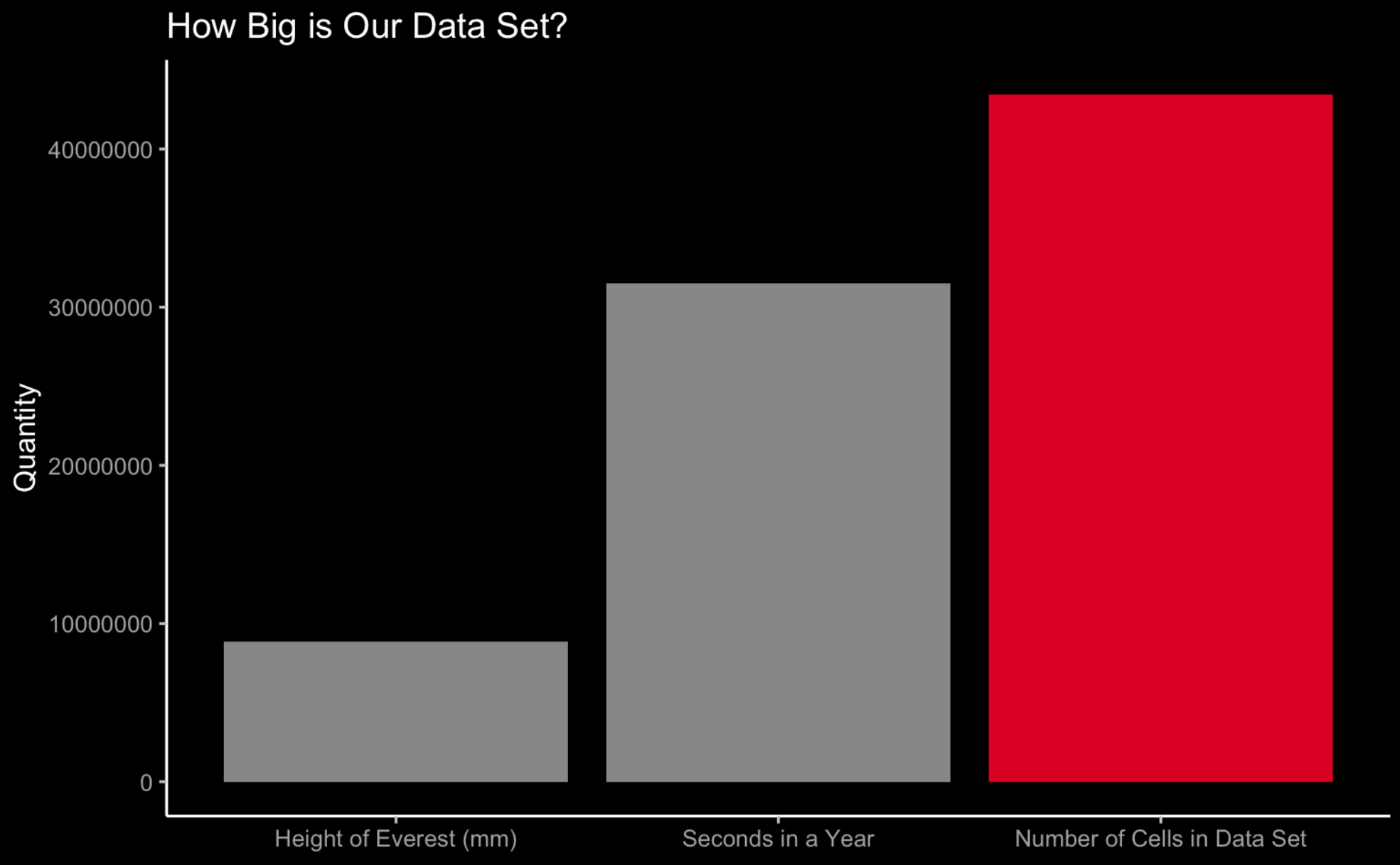
We have been offering affordable, accessible financial products and services for decades

Our Goal: Improve Financial Inclusion Through Access to Credit

How Do We Achieve that?

Better Identification of Creditworthy Customers!

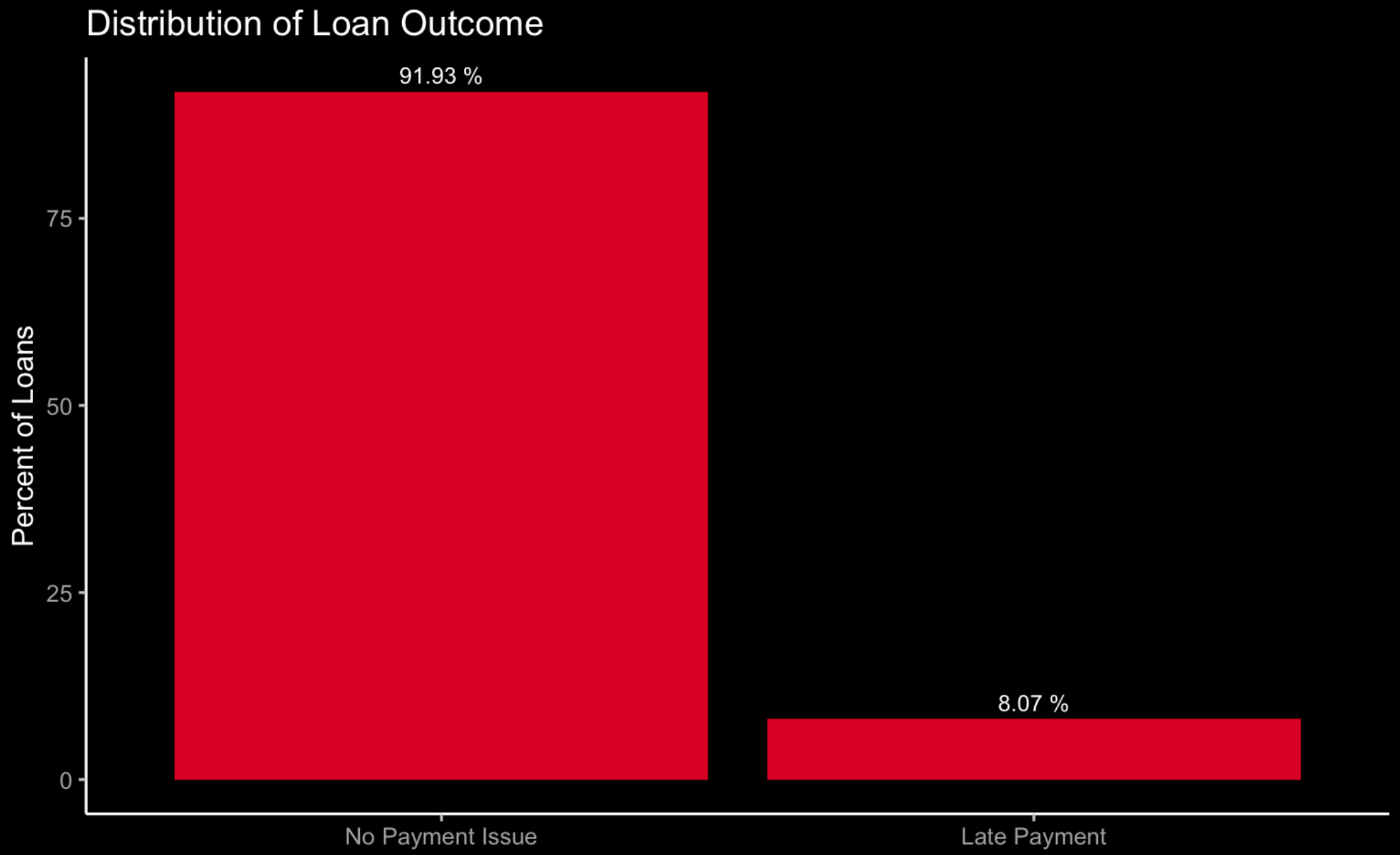
Finding #1: Quantity of Data



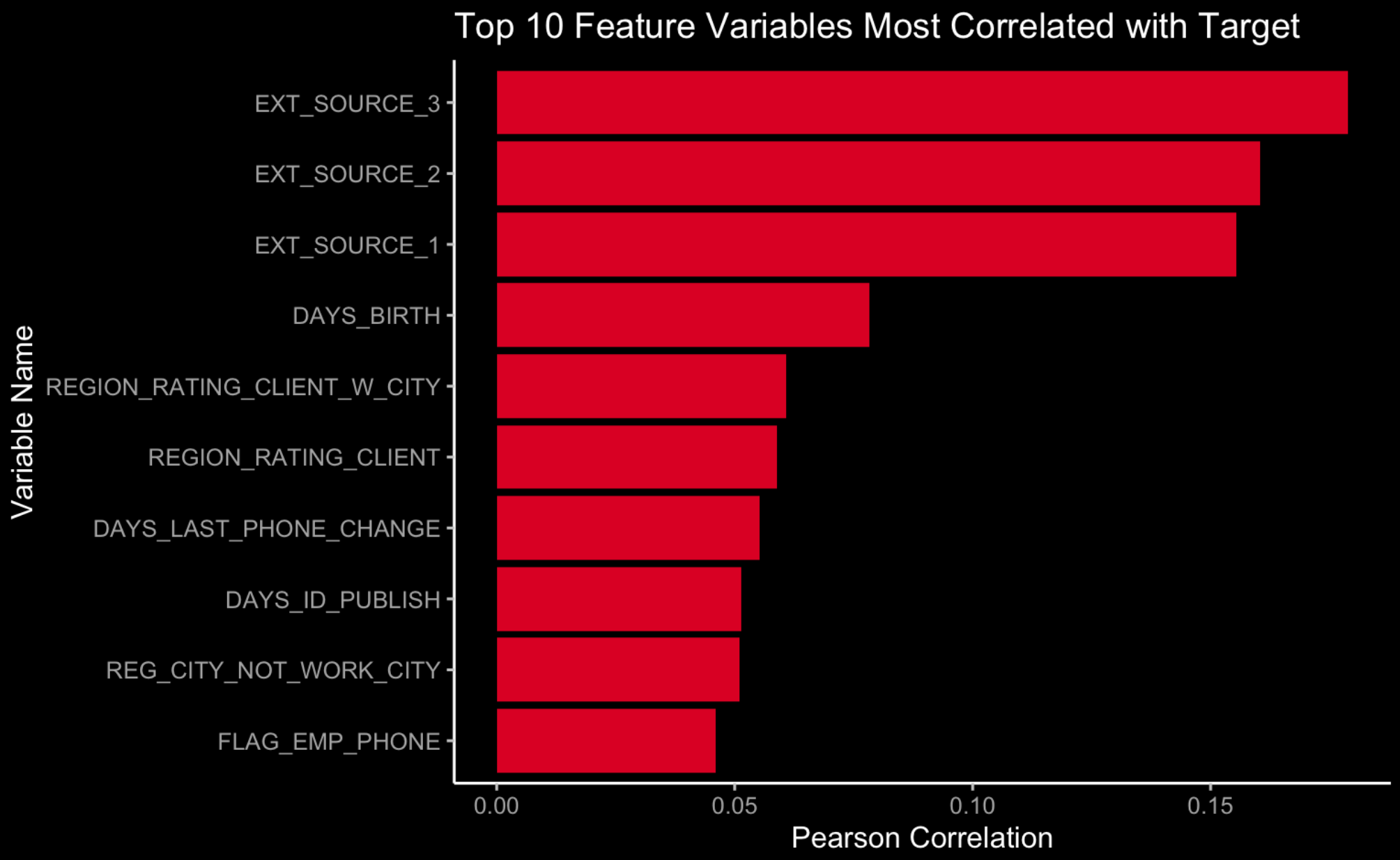
Data Set Summary

- Home Credit Default Risk Data Set
- Application Data: Core data with individual loan applications, personal and financial details.
- Credit Bureau Data: past credits from other financial institutions.
- Previous Loan Performance: with home credit.
- Repayment History: records of payments made and missed for previously disbursed credits

Finding #2: Imbalanced Target



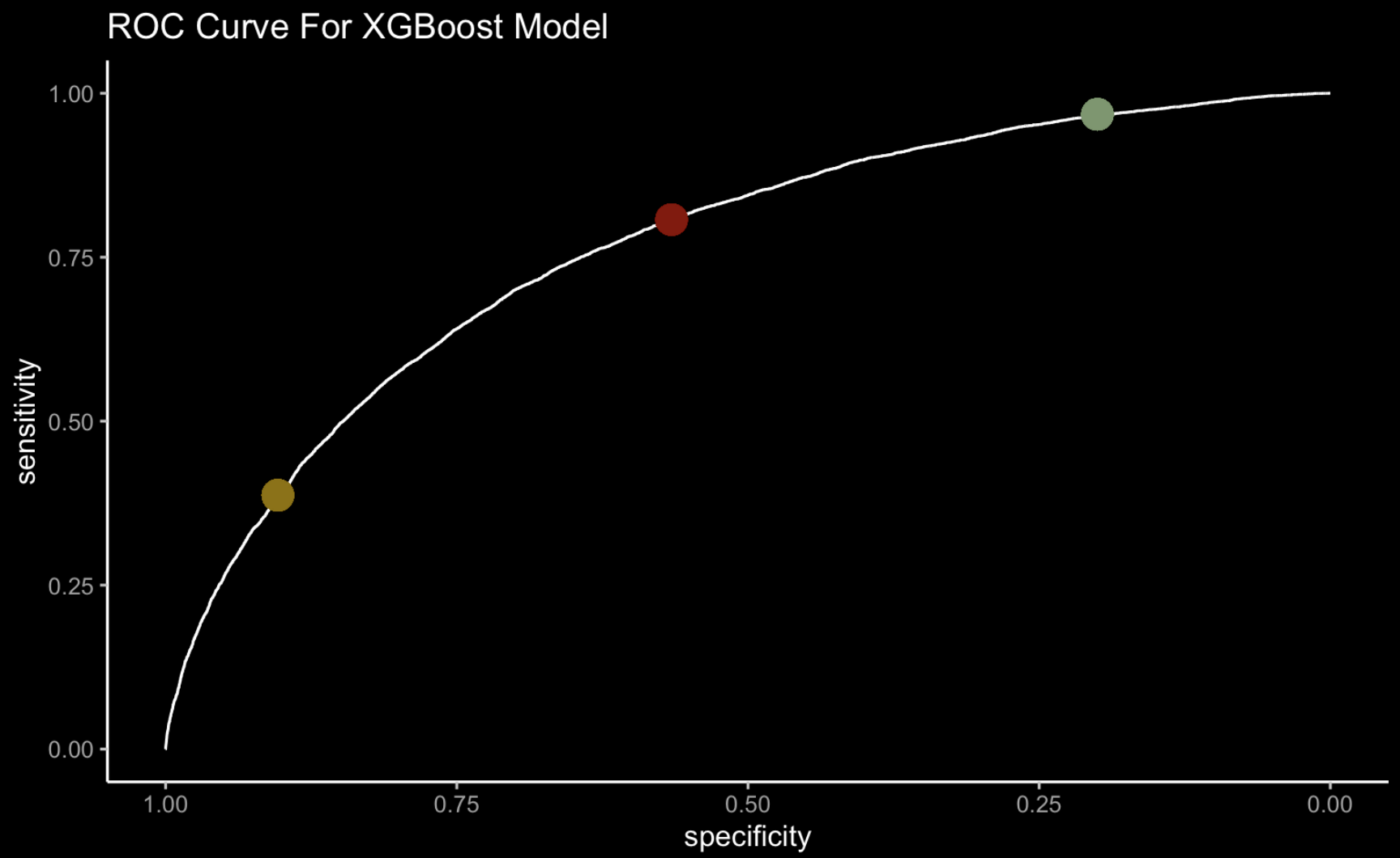
Finding #3: Some Features are More Important



Our Solution: XGBoost

- Usable with big data sets
- Handles imbalanced targets
- Automatically identifies important features
- Uses regularization to avoid overfitting
- High performance and customization
- Test Set AUC: 0.76

Our Model Provides Performance and Adaptability



Prediction	Reference 0	1
0	68300	5048
1	2267	1262

Prediction	Reference 0	1
0	56968	2741
1	13599	3569

Prediction	Reference 0	1
0	27313	608
1	43254	5702

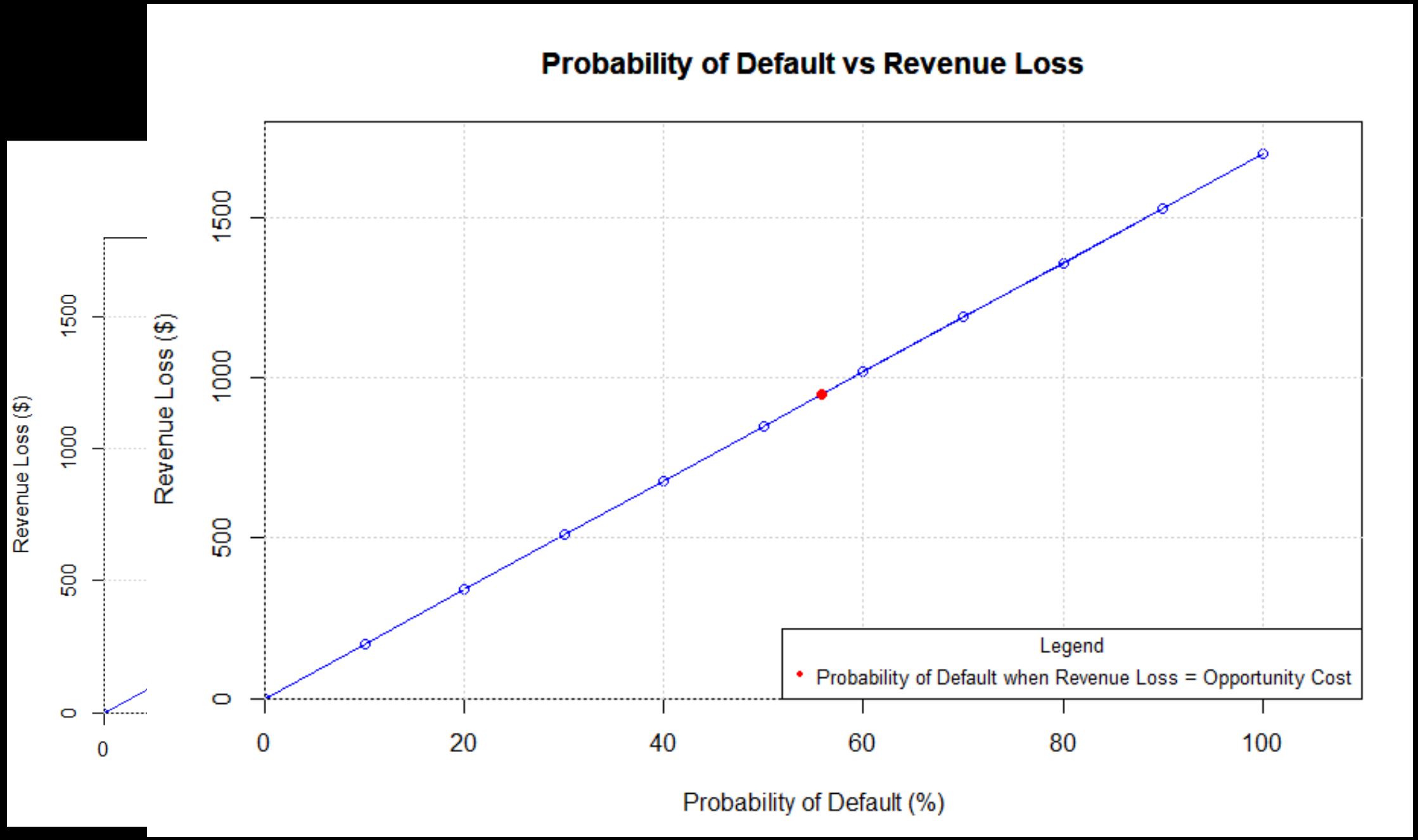
Model Value: A Cost-Benefit Analysis

Loss of Revenue due to Default =
(Loan Amount * APR) * P(Default)

Opportunity Cost of Refusing Loan =
Potential Profit from Interest + Goodwill +
Processing Fee

These values can change over time!

Example Analysis Under Current Market Conditions



Thank You!