Credit Risk and Expected Loss

A detailed analysis of Peer to Peer loans at Lending Club.

Lending Club

- Lending Club is a Peer to Peer Lender that matches lenders and borrowers
- It offers an alternative to traditional banking for investors and borrowers
- It is innovative and data focused with greenfield systems
- It wants to increase the number of loans it makes

Capital and Regulation

- Banks are strategically important and are strongly regulated
- Capital must be set aside as a buffer against shock
- The Capital Requirement is based on the Risk of loans
- Banks can use their own models to calculate Credit Risk and Expected Loss

The Questions to Answer

- What is the Probability of Default for a new loan applicant
- What is the Expected Loss on a new loan
- How much Capital should Lending Club hold against losses

Target Variables: 3 Components of Expected Loss

- PD = Probability of Default - (likelihood of Good or Bad borrower)
- EAD = Exposure at Default - (percentage of loan outstanding)
- LGD = Loss Given Default - (will anything be recovered after default)

$$EL = PD * EAD * LGD$$

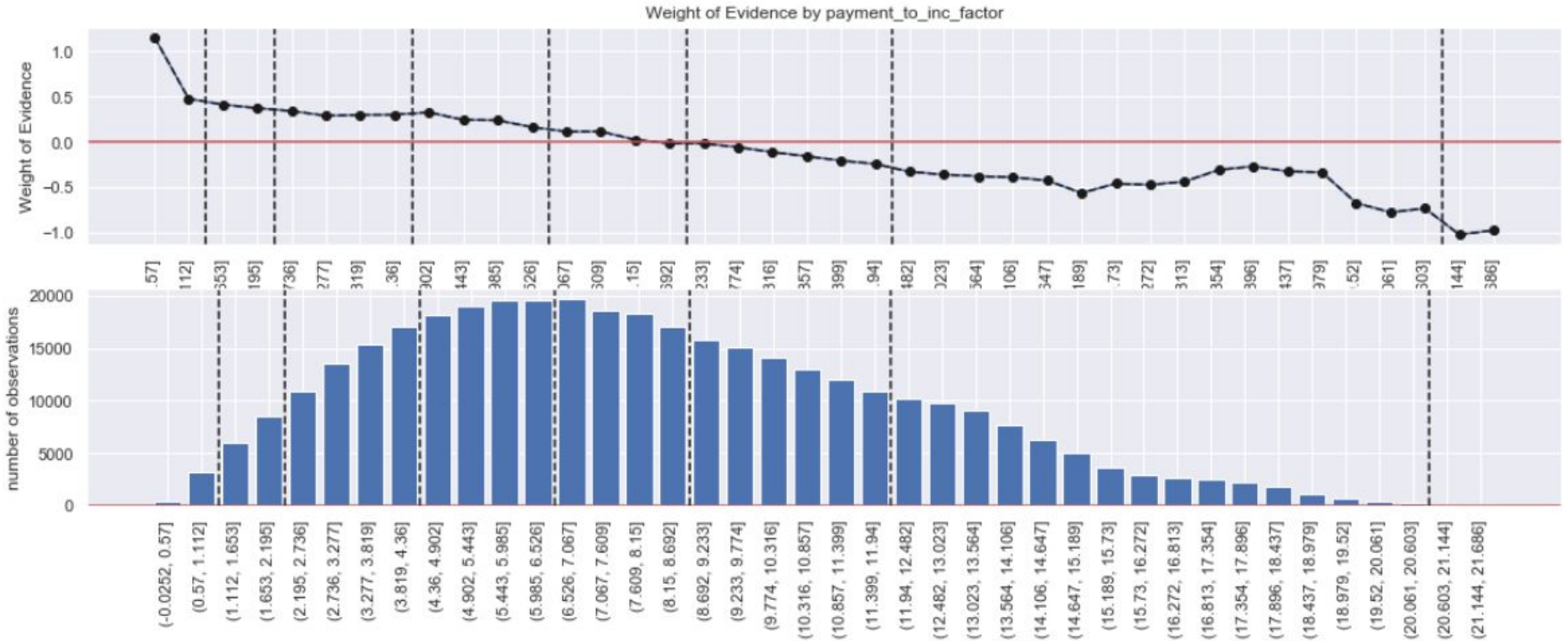
The Dataset

- The Lending Club Loans Data 2007-2014 was sourced from Kaggle.
- It contains customer information, loan amount, loan purpose, credit history, current debts, home state, income, loan purpose, home ownership, etc
- It also contains an external ratings Grade used by Lending Club
- Economic Data sourced from macrotrends.net: Unemployment and Retail Sales, plus market indicators Ted Spread & VIX (volatility)

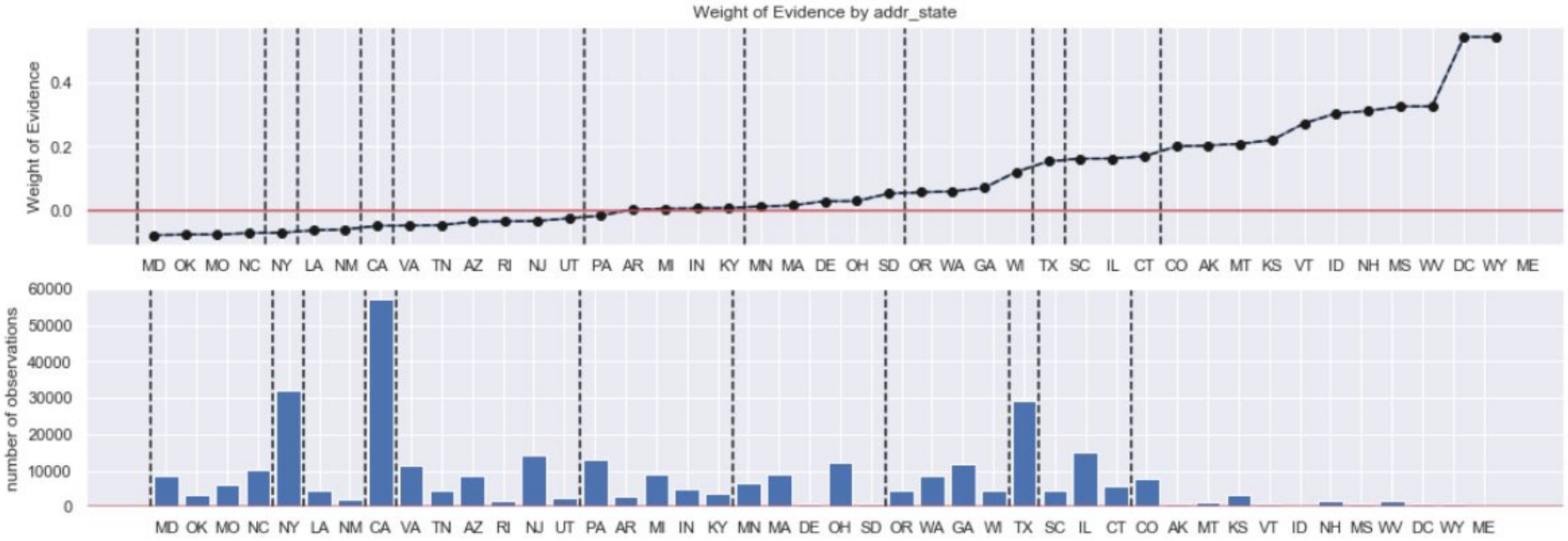
Probability of Default

- Probability of Default is the most important component of Expected Loss
- PD is used to distinguish between good and bad loan applications
- Easy to use transparent Credit Scorecards can constructed using PD smoothing the loans assessment process
- Let's look at the factors relating to a high Probability of Default

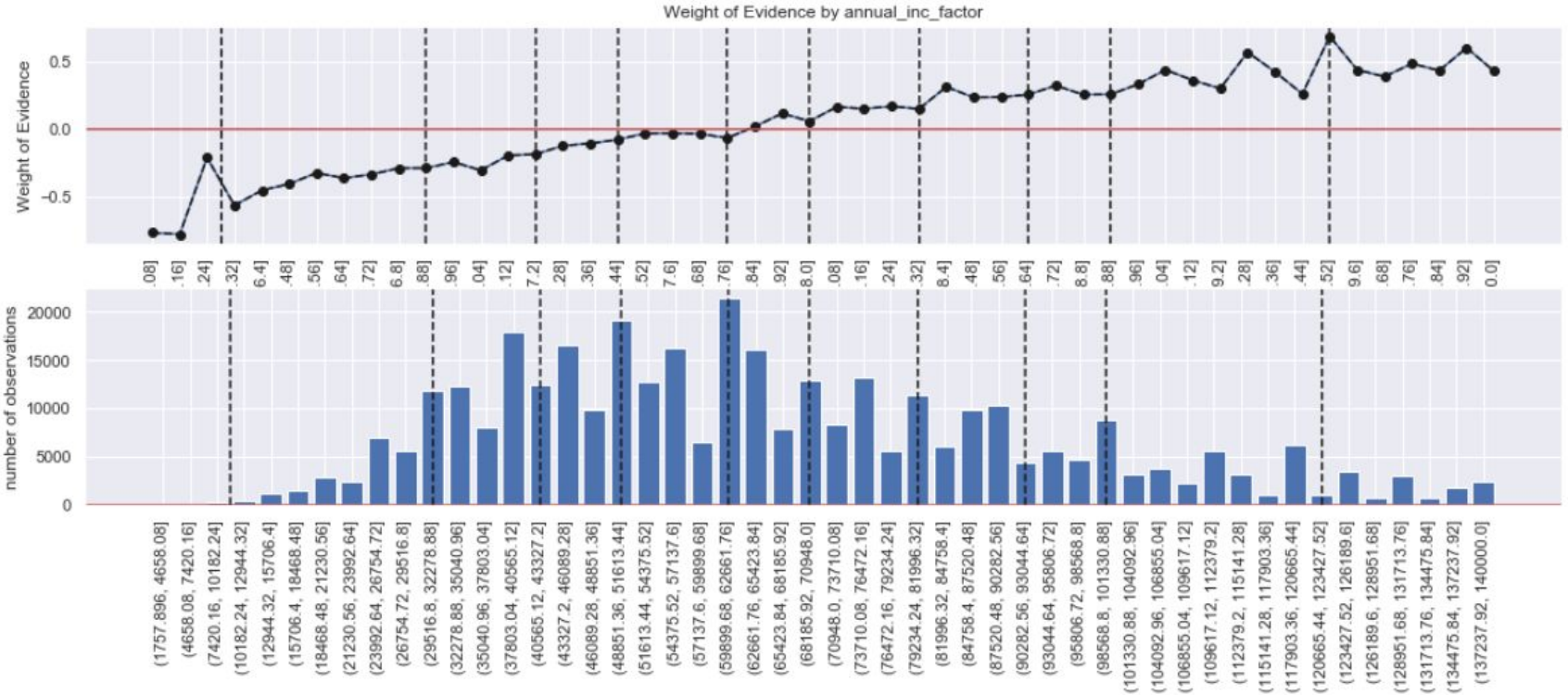
PD: Weight of Evidence: Payment to Income Ratio vs Repayment



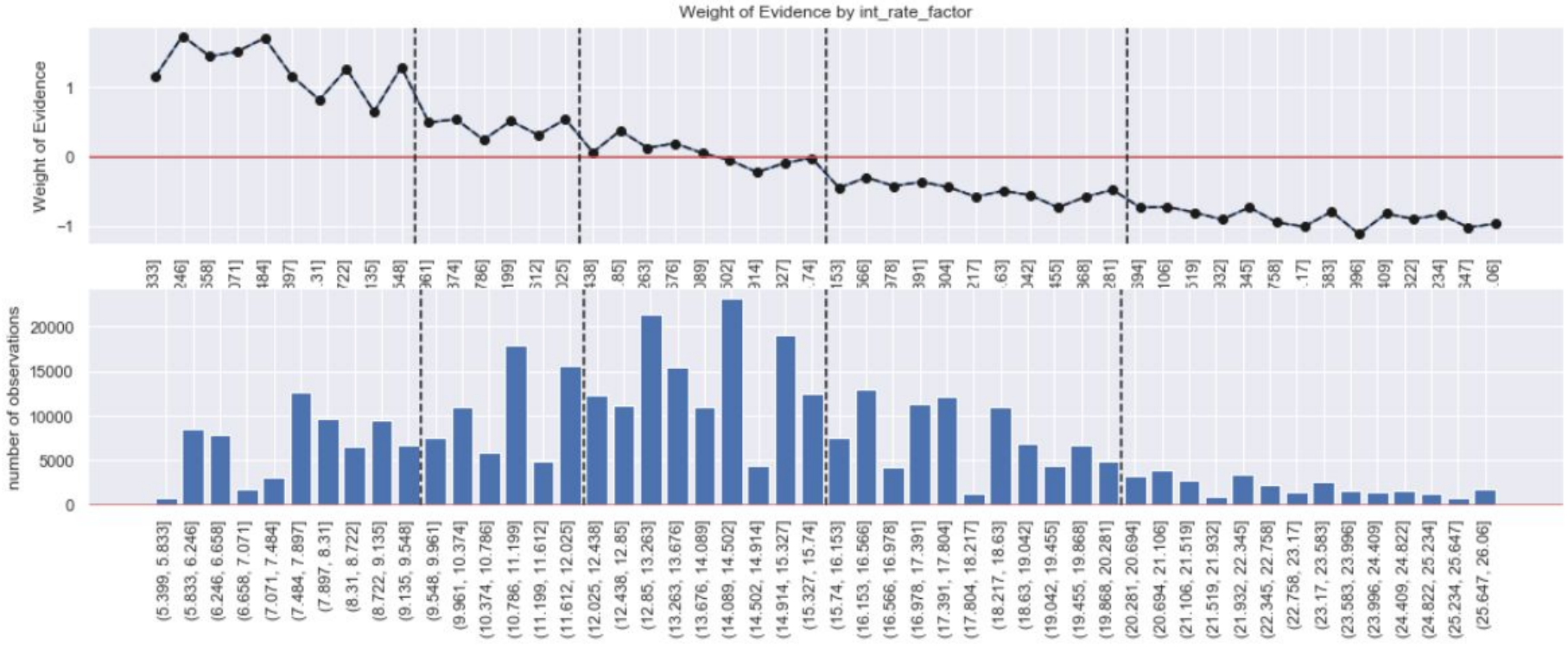
PD: Weight of Evidence: State of Residence vs Repayment



PD: Weight of Evidence: Annual Income vs Repayment



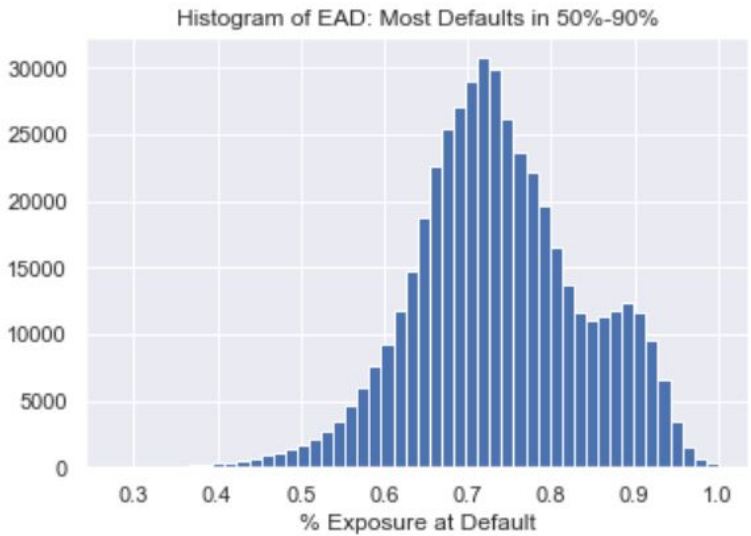
PD: Weight of Evidence: Interest Rate vs Repayment

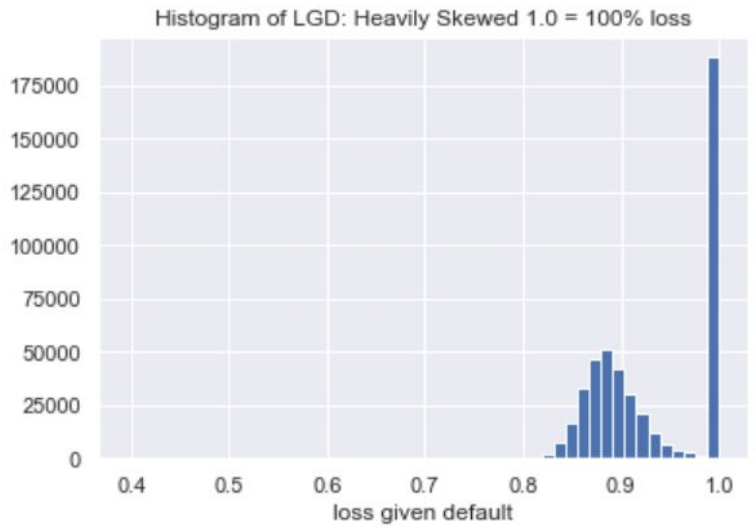


Exposure at Default and Loss Given Default

Exposure at Default and Loss Given Default are calculated next based on the defaulted loans population

- Both populations are skewed towards large losses: EAD has a mean of 75% with a min of 27.9%
- LGD has a mean of 93.4% and a minimum of 39.7%





Results

- Total Loans $6.66bn
- Total Capital Required $373.9m
- Capital Requirement of 5.611%
- Average Expected Loss = PD * EAD * LGD = 5.611%

Future Work:

- Comparison with the 2015 Lending Club dataset (430K+ new loans)
- Consider alternate models and methodologies for EAD and LGD estimates
- Consider other statistical packages: LGD would be suited to R
- Consider other classification approaches for PD
- Consider use of AWS for more computationally expensive techniques