

PART 1

Research and summarize what your client's needs are going to be.

Our client MAR is a regulator whose work is to regulate how the model can be used or cannot be used.

His primary function is to check the parameters for both- lenders and borrowers and training different models for predicting the interest rate and finally, choosing the most profitable model for Lending Club.

For performing the above-mentioned tasks effectively, my client, MAR is going to need the following data (apart from the normal mandatory data fields like loan amount, term, interest rate)-

1. Grade and Sub Grade
2. **Employment status (Should be included)**
3. Employment Length
4. Home Ownership
5. Annual Income
6. Loan status
7. Purpose
8. Delinquency – last 2 years
9. **FICO Score (Should be included)**
10. DTI ratio
11. Number of open credit lines
12. Application type

Explore the data and comment on data quality, features and get a feel for the data.

Lending Club Dataset consists of loan.csv that contains full-fledged loan data for all loans issued throughout the years 2007-2015, including the amount of loans funded, purpose of loan, debt to income ratios of borrowers, amount of current loan status (Current, Late, Fully Paid, etc.) and latest payment information.

Additional features include credit scores, number of finance inquiries, number of delinquencies in past 2 years, address including zip codes and state, and collections among others.

The file is a matrix of about 890 thousand observations and 75 variables. A data dictionary is provided in a separate file that has the description about all the variables used in the dataset.

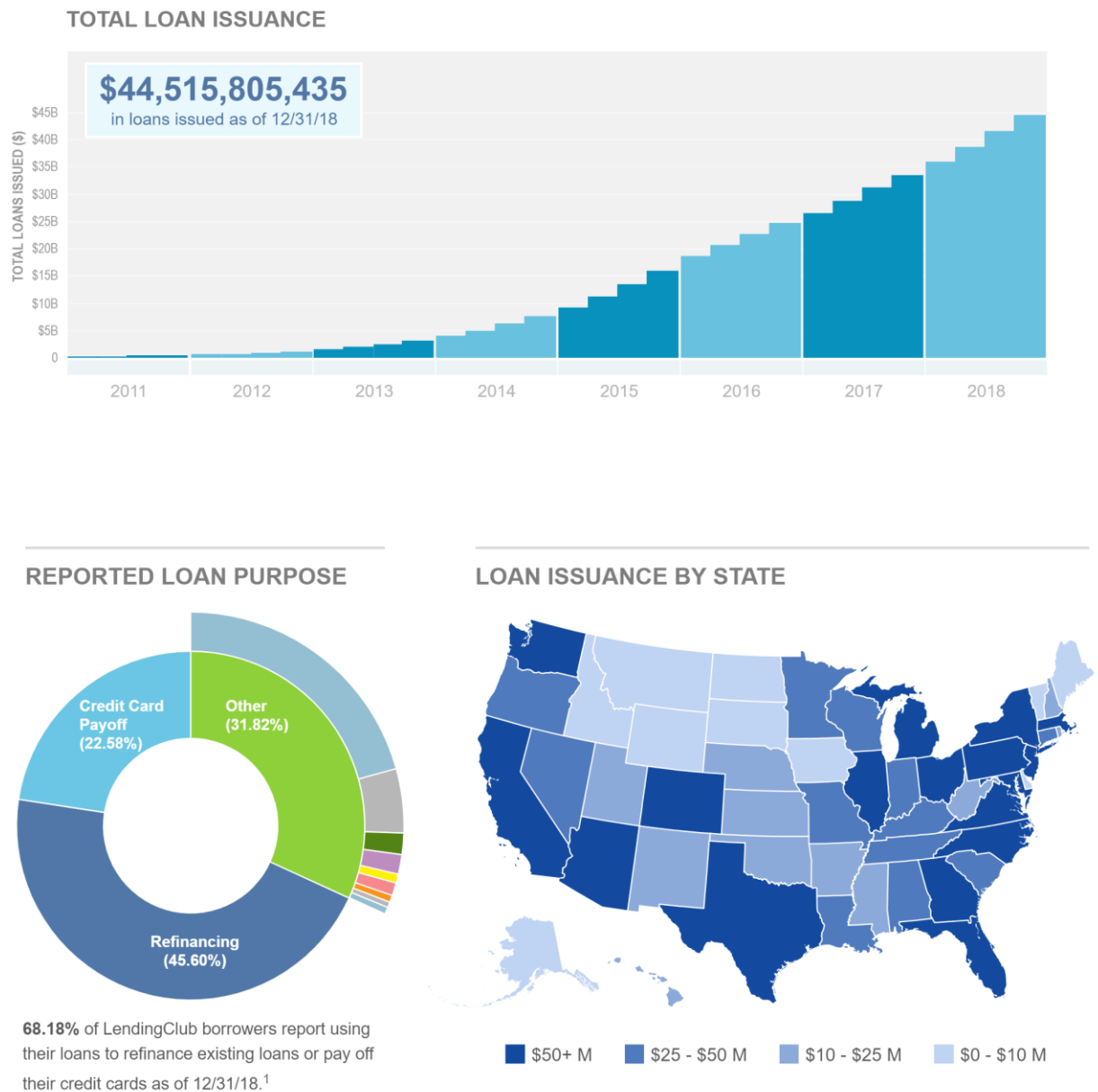
GRAPH ANALYSIS

Lending Club is a US peer to peer Company where investors can search and browse the loan listings on its website and select loans that they want to invest in based on information supplied about the borrower, amount of loan, loan grade and loan purpose. Lending club has generated many graphs for analysis which includes Loan Issuance and its purpose, Investor account performance, average interest rate, grade mix

over time, loan performance details, net annualized return by vintage, loan status migration over 9 months.

Graph 1:

Platform: Highlights | Public Note Offering: [Investor Performance](#) | [Loan Statistics](#) | [Download Data](#)



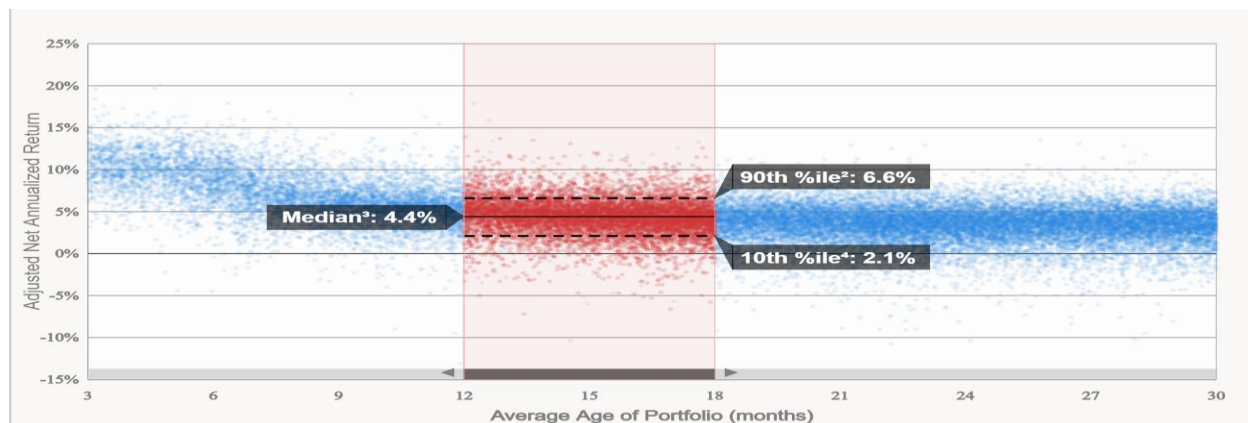
The first chart is about Total Loan Issuance. The total loan issued till 31st December 2018 was \$44,515,805,435. The graph shows the total loan issued in \$ for that particular year from 2011-2018. There is an increasing trend in total loan issued from 2011-2018. From the graph it is clear that 68.18% of Lending club borrowers used their loan to refinance existing loans or pay off their credit cards. Total 22.58% of the borrowers took loan for credit card pay off and 45.60% of the borrowers took loan for

refinancing. There are other loan purposes too which are not mentioned and total 31.82% of the borrowers fall under these categories. The graph of loan issuance by state depicts that the darkest shade of blue shows the highest amount of loan issued by those states, while the lightest shade shows the lowest amount of loan issued by those states. In 2011 the purpose of loan issued by borrowers was due to refinancing while in 2012 the purpose of the loan issued by the borrowers was due to credit card payoff. The graph just shows the issuance of loan by the borrowers and does not show the actual usage of the loan.

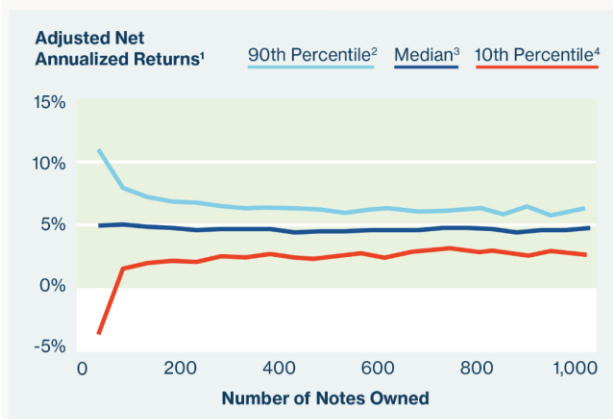
Key takeaways for client:

From this graph it is clear that for the regulator (client) loan purpose will be an important parameter for analysis and second important is state. Regulator gets a clear picture of what factors are leading to more loan issuance and what needs to be analyzed more for company's benefit.

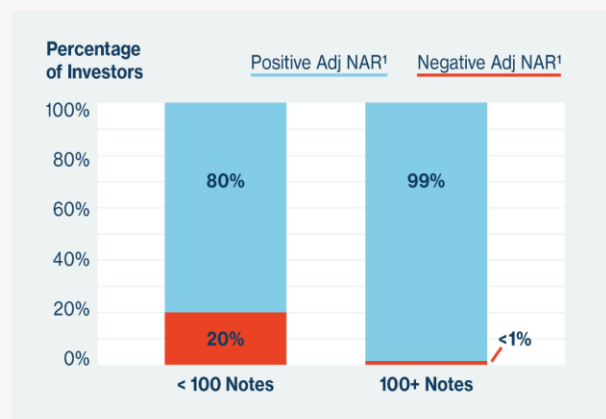
Graph 2:



Diversification Can Reduce Volatility of Returns



Owning 100+ Notes Reduces Risk in Your Returns



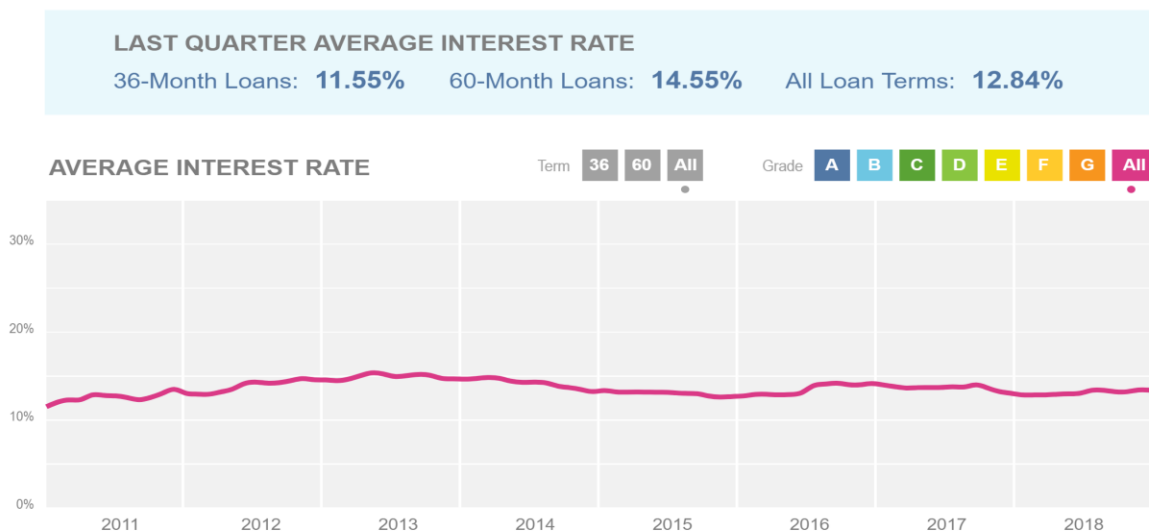
The second chart is about Investor account performance based on Investor account returns by average age of portfolio. The y axis of the chart has Adjusted Net Annualized Return while the x axis is average age of portfolio. Adjusted Net Annualized Return is a measure of return on investment which allows investors

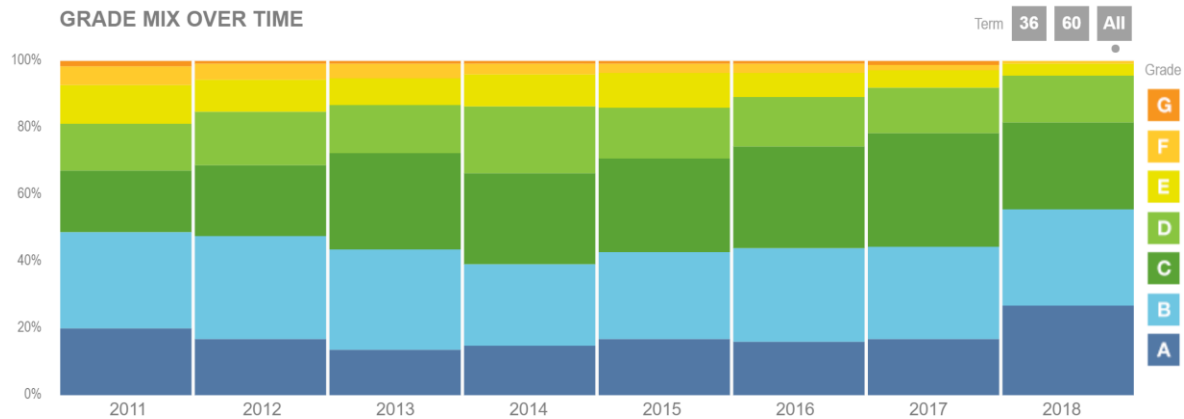
to model the impact of potential losses on notes in their portfolio. Adjusted NAR is similar to unadjusted NAR but it incorporates an estimate of future losses in any loans that are in past due status but have not been charged off. The purpose of the chart is to illustrate how the returns change over the life of an investment and how different factors can influence volatility returns. Population of the account that are tightly grouped around the median line may have fewer volatile returns while those more spread out vertically have more volatile returns. So, from the graph investors whose average age of portfolio is between 3-9 months have more volatile returns and those whose average age of portfolio is between 21-30 have less volatile return. The adjusted net annualized return is inversely proportional to average age of portfolio. More the age less is the returns. Owning small number of notes leads to more volatile returns. If the number of nodes is more than 100 then the positive adj NAR percentage is more. Also, concentration of investment may lead to more volatile returns. Returns are more stable with notes having lower interest rate as the corresponding loan have lower charge off rates. Diversification chart demonstrates that as we move from left to right the number of notes increases and lines in the chart gets closer. This shows that the returns get stable and we move from left to right to more the number of notes more the returns are stable.

Key takeaways for client:

From this graph Regulator gets a clear picture that the graph of adjusted NAR vs Age of portfolio is very important for investors as they can predict their future losses. Also, from the diversification charts Regulator can understand that as the number of notes increases the adjusted NAR stabilizes and which reduces the volatility of the returns which is beneficial for investor and hence even this chart is of utmost importance.

Graph 3:





The third graph is about the average interest rate. The graph depicts that from 2011 to 2013 the average interest rate percentage increased and from 2013 to 2016 it gradually decreased. From 2016 to 2018 the interest rate was stable. The average interest for 36-months loan was 11.55% while for the 60 months loan it was 14.55% and overall the rate was 12.84%. From the analysis we can say the overall graph gets C grade. The next graph is grade mix over time. This graph shows the mix of grades of loans each year. The distribution of grade B and C is more as compared to other grades. The lowest distribution is of grade F and G. The next graph is loan performance details. The table shows detailed historical performance of loans by grade. The Adj Net Annualized Return is directly proportional to interest payments received. As the interest payment received increases the adj net annualized return also increases. Grade B and C are the one which has maximum value of total loan issued while grade F and G are the lowest. The next chart is net annualized return by vintage. The graph shows that in 2008 the net annualized return majorly decreased in 36 months. There was not much decrease in net annualized return in 2013 as compared to other years or vintage. The next chart is loan status migration over 9 months. This chart refers to average 9-month recovery rates by loan status for the months of October 2017 through December 2017. In grace period the net charged off is minimum. Whereas between 30-120 days or 120+ days the net charged off is high.

Key takeaways for client:

The graph of grade mix over is very important for the Regulator as it depicts the grades for average rate interest and loans and gives better picture of which loans are better and have better interest rate.