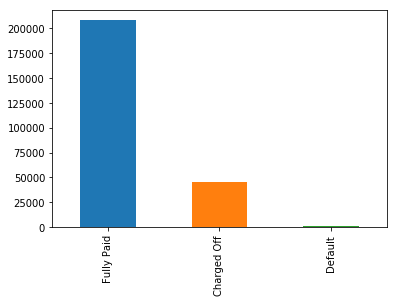
**Exploratory Data Analysis**

**Lending Club Loan Data: Analyze Lending Club Issued Loan’s**



My exploratory data analysis is split in into two parts. The first part consists of looking at the loan characteristics and demographics to get a general idea of the dataset and trends. The second part consists of looking at the loan characteristics and demographics for people that default or are charged off.

Part 1:



A loan becomes “Charged Off” when there is no longer a reasonable expectation of further payments.  Charge Off typically occurs when a loan is 120 days or more past due and there is no reasonable expectation of sufficient payment to prevent the charge off. Thus, to predict whether or not a person will default we can merge charged off and default.

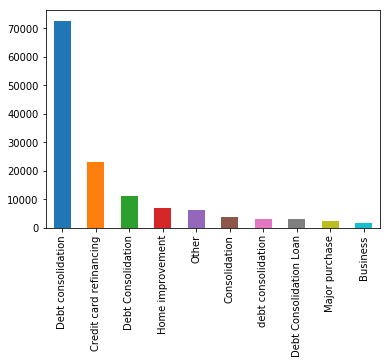
The first thing to count is the number of people that fully paid off their loan vs the number of people that defaults.

Fully Paid 148052

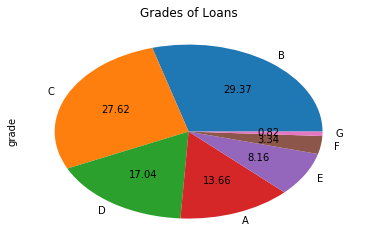
Default 33833

The ratio of people that default in our dataset for the timeframe 2007 - 2015 is 22.85%. This is an alarmingly high rate just to start off with.

The top 10 title provided by borrowers. As we can see the main reason that people using Lending Club is for debt consolidation. A reason for that can be that people that resort to Lending Club is because they have exhausted the traditional means of borrowing.



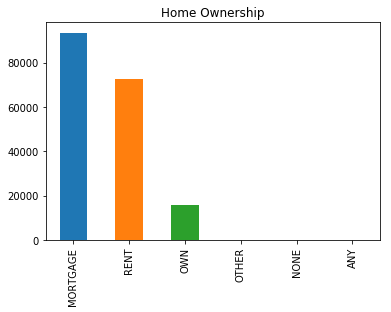
Let’s look at the grades of loans issued by Lending Club.



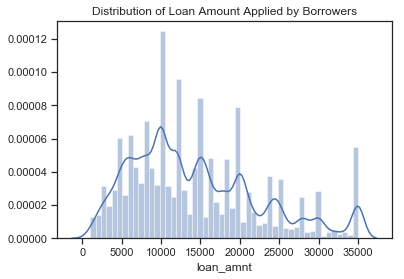
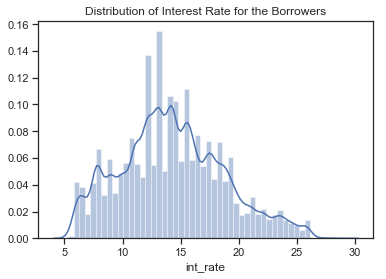
More than 50% of the loans that are issued in Lending Club is B and C grade. A quick look at the pie chart shows the grade distribution of the loans.



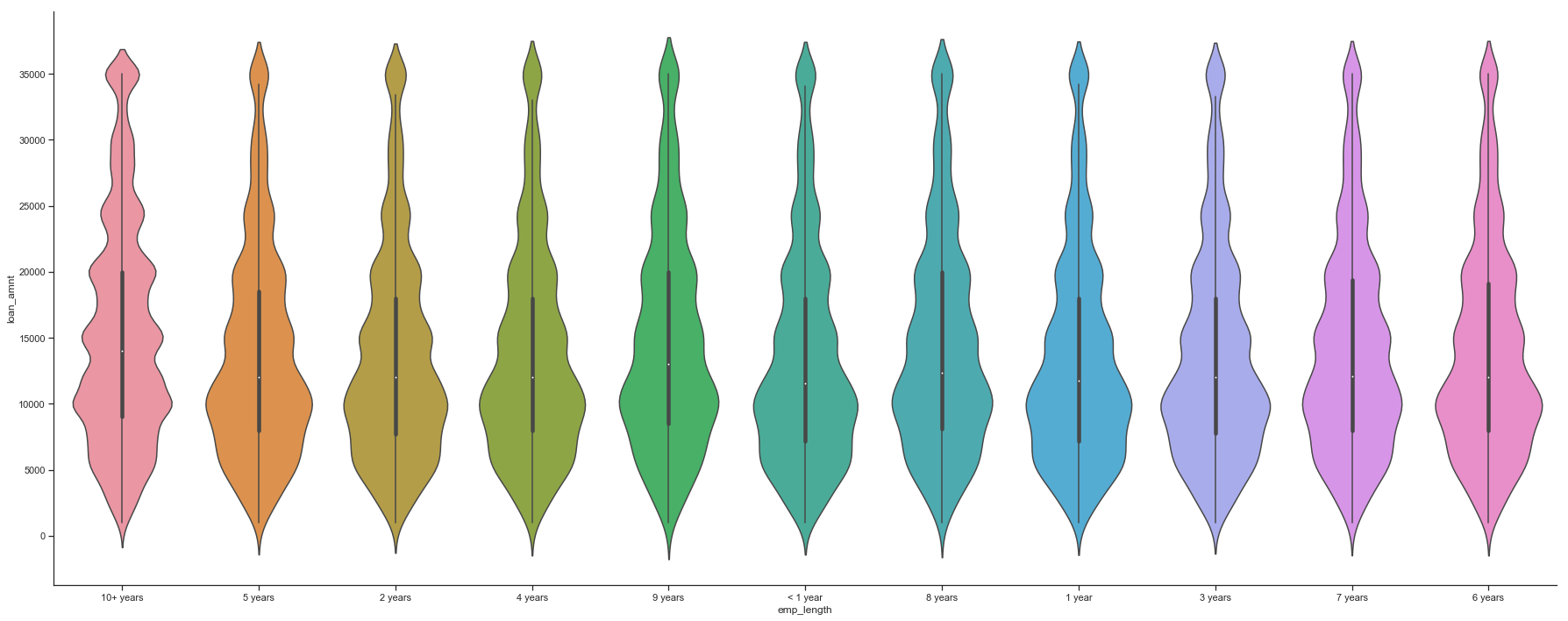
Due to most of the loans being of B and C grade the majority of the loans are of the shorter term of 36 months



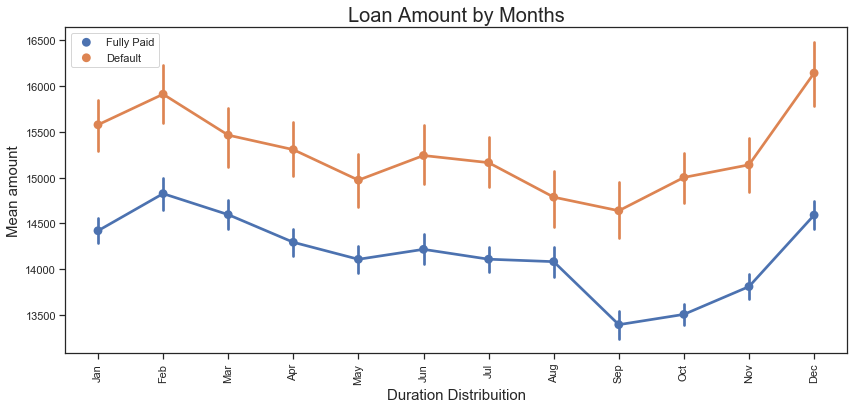
Mortgage and Rent are the two main categories of the people who use lending club. It will be interesting to see how many of the people who own vs mortgage vs rent default on their loans.



These two distribution plots show the distribution of loan amount and interest rate of the people borrowing from Lending Club. We can see that most of the loan amount range from 5k to 15k where 10k is the most commonly taken loan amount. While the most common distribution of interest rate ranges from 10% to 20% with 13% being the most common.

Also, interesting to see the distribution of the amount that is borrowed compared to how long someone has worked. People who have worked more than 10+ years tend to have fatter tails meaning they borrow more because they have more income. People who have been employed for lesser time tend to have fatter bottoms for the same reason.

Part II:



Very interesting. It is surprising that people who default borrow more than people who pay off their loans. Another thing to look for is the borrowing increases the months of October, November and December. A reason for the spike at the end of the year could be because of the holiday season and people consolidating their debt heading in the new year.

