**Milestone Report**

Lending Club is a peer to peer lending company based in the United States, in which investors provide funds for potential borrowers and investors earn a profit depending on the risk they take (the borrowers credit score). Lending Club provides the "bridge" between investors and borrowers. For more basic information about the company please check out the wikipedia article about the company.

Primary focus will be to study different patterns in data where we can gain insights on how much loans has been disbursed, what is general tenure of loans, how many of them have been defaulted, types of loans, interest rates, major purpose of loans, etc

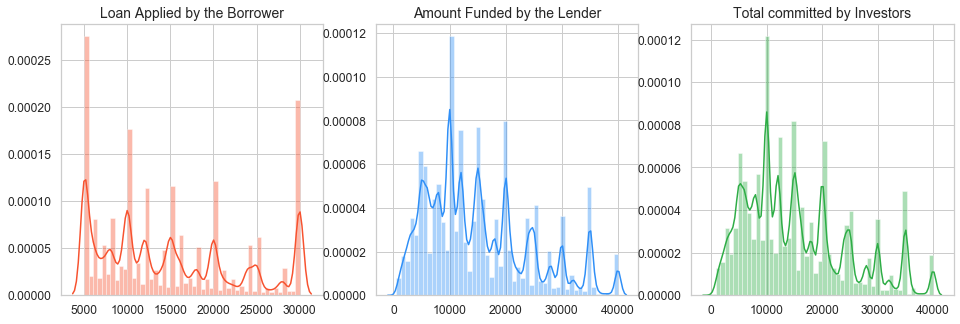
We will start by exploring the distribution of the loan amounts and see when did the loan amount issued increased significantly.

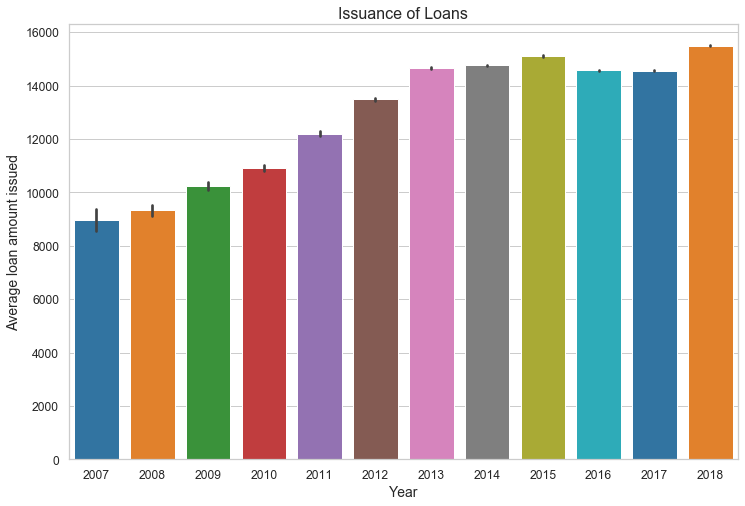
What we need to know:

* Understand what amount was mostly issued to borrowers.
* Which year issued the most loans.
* The distribution of loan amounts is a multinomial distribution.

What we can conclude:

* Most of the loans issued were in the range of 10,000 to 20,000 USD.
* The year of 2015 was the year were most loans were issued.
* Loans were issued in an incremental manner. (Possible due to a recovery in the U.S economy)
* The loans applied by potential borrowers, the amount issued to the borrowers and the amount funded by investors are similarly distributed, meaning that it is most likely that qualified borrowers are going to get the loan they had applied for.





## **Types of Loans:**

In this section, we will see what the amount of bad loans Lending Club has declared so far.

#### What we need to know:

* The amount of bad loans could increment as the days pass by, since we still have a great amount of current loans.
* Average annual income is an important key metric for finding possible opportunities of investments in a specific region.

#### What we can conclude:

* Currently, bad loans consist 13.14% of total loans but remember that we still have current loans which have the risk of becoming bad loans. (So this percentage is subjected to possible changes.)





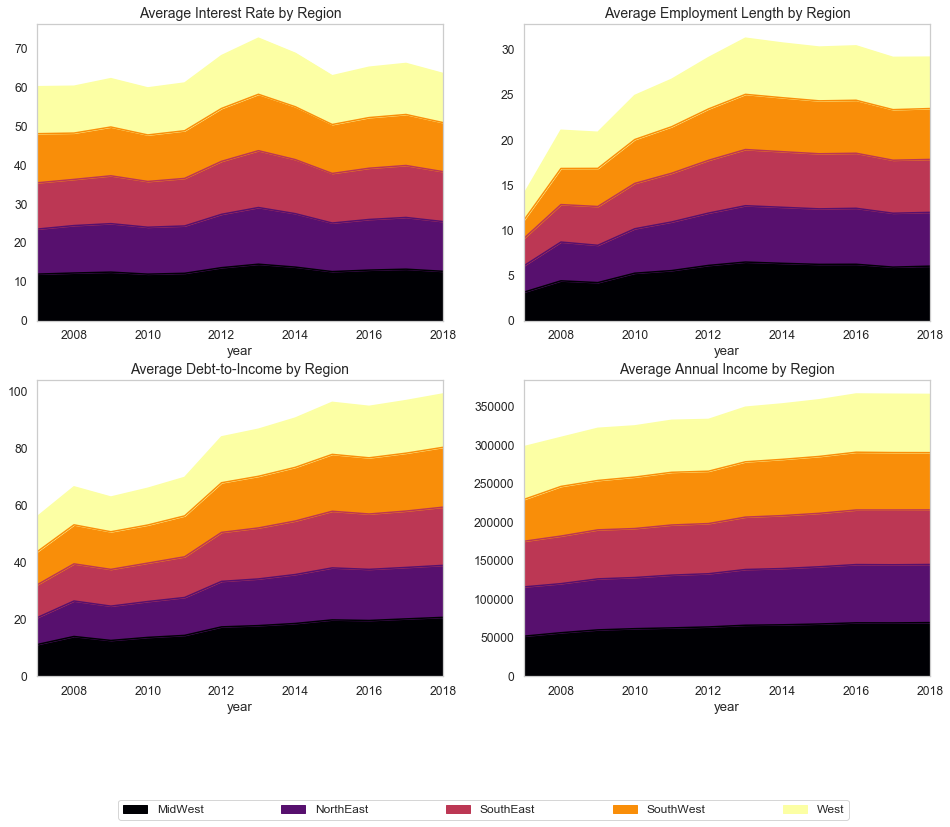
## **Analysis of the Region**

In this section we want to analyze loans issued by region in order to see region patters that will allow us to understand which region gives Lending Club.

#### What we can conclude:

* SouthEast , West and NorthEast regions had the highest amount lof loans issued.
* West and SouthWest had a rapid increase in debt-to-income starting in 2012.
* West and SouthWest had a rapid decrease in interest rates.
* The NorthEast region seems to be the most attractive in term of funding loans to borrowers.
* The SouthWest and West regions have experienced a slight increase in the "median income" in the past years.
* Employment Length tends to be greater in the regions of the SouthWest and West
* Clients located in the regions of NorthEast and MidWest have not experienced a drastic increase in debt-to-income(dti) as compared to the other regions





## **More Analysis on Bad loans**

#### What we need to know:

* The number of loans that were classified as bad loans for each region by its loan status.
* This won't give us the exact reasons why a loan is categorized as a bad loan (other variables that might have influence the condition of the loan) but it will give us a deeper insight on the level of risk in a particular region.

#### What we can conclude:

* The regions of the West and SouthEast had a higher percentage in most of the b "bad" loan statuses.
* The NorthEast region had a higher percentage in Grace Period and Does not meet Credit Policyloan status. However, both of these are not considered as bad as default for instance.
* Based on this small and brief summary we can conclude that the West and SouthEast regions have the most undesirable loan status, but just by a slightly higher percentage compared to the NorthEast region.
* Again, this does not tell us what causes a loan to be a bad loan , but it gives us some idea about the level of risk within the regions across the United States.





## **Analysis by Income:**

In this section we will create different income categories in order to detect important patters and go more into depth in our analysis.

**What we need to know:**

* Low income category: Borrowers that have an annual income lower or equal to 100,000 usd.
* Medium income category: Borrowers that have an annual income higher than 100,000 usd but lower or equal to 200,000 usd.
* High income category: Borrowers that have an annual income higher the 200,000 usd.

#### What we can conclude:

* Borrowers that made part of the high-income category took higher loan amounts than people from low and medium income categories. Of course, people with higher annual incomes are more likely to pay loans with a higher amount. (First row to the left of the subplots)
* Loans that were borrowed by the Low-income category had a slightly higher change of becoming a bad loan. (First row to the right of the subplots)
* Borrowers with High and Medium annual incomes had a longer employment length than people with lower incomes. (Second row to the left of the subplots)
* Borrowers with a lower income had on average higher interest rates while people with a higher annual income had lower interest rates on their loans. (Second row to the right of the subplots)

# **/var/folders/bc/k9gl254x609bq_zh_08jmhkm0000gn/T/com.microsoft.Word/Content.MSO/5FD49CAE.tmp**

## **Analysis of Credit Scores:**

Credit scores are important metrics for assessing the overall level of risk. In this section we will analyze the level of risk as a whole and how many loans were bad loans by the type of grade received in the credit score of the customer.

#### What we need to know:

* The lower the grade of the credit score, the higher the risk for investors.
* There are different factors that influence on the level of risk of the loan.

#### What we can conclude:

* The scores that has a lower grade received a larger amount of loans (which might had contributed to a higher level of risk).
* Logically, the lower the grade the higher the interest the customer had to pay back to investors.
* Interestingly, customers with a grade of "C" were more likely to default on the loan



## **/var/folders/bc/k9gl254x609bq_zh_08jmhkm0000gn/T/com.microsoft.Word/Content.MSO/93D35CEC.tmp**

My main aim in this section is to find the main factors that causes for a loan to be considered a "Bad Loan". Logically, we could assume that factors such as a low credit grade or a high debt to income could be possible contributors in determining whether a loan is at a high risk of being defaulted. 

#### What we need to know:

* Factors that increase risk include: low annual income, high debt to income, high interest rates, low grade, among others.

#### What we can conclude:

* The types of bad loans in the last year are having a tendency to decline, except for late payments.
* Mortgage was the variable from the home ownership column that used the highest amount borrowed within loans that were considered to be bad.
* There is a slight increase on people who have mortgages that are applying for a loan.
* People who have a mortgage are more likely to ask for

