**Loan Analysis Report**

**Validation and Cleaning**

**emp\_length:**

**Employment length in years. Possible values are between 0 and 10 where 0 means less than one year and 10 means ten or more years.**

* Has blanks and n/a values
* Categorical and can be changed to numerical with values from 0 to 10.

**emp\_title:**

**The job title supplied by the Borrower when applying for the loan.**

* Has blanks (many)
* Has a number that is not a string or indicate a title

**Grade:**

**LC assigned loan grade, loan grading is a classification system that involves assigning a quality score to a loan based on a borrower's** [**credit history**](https://www.investopedia.com/terms/c/credit-history.asp)**, quality of the** [**collateral**](https://www.investopedia.com/terms/c/collateral.asp)**, and the likelihood of repayment of the principal and interest. A score can also be applied to a portfolio of loans.**

* No blanks or outliers.

**Subgrade**:

**LC assigned loan subgrade.**

* No blanks or outliers.

**funded\_amnt:**

**The total amount committed to that loan at that point in time.**

* No blanks or outliers.

**Loan\_amnt:**

**Total amount for the loan, The listed amount of the loan applied for by the borrower. If at some point in time, the credit department reduces the loan amount, then it will be reflected in this value.**

* No blanks or outliers.

**Id:**

**A unique LC assigned ID for the loan listing**.

**Member\_id:**

**A unique LC assigned Id for the borrower member.**

**Term:**

**The number of payments on the loan. Values are in months and can be either 36 or 60.**

**Int\_rate:**

**Interest Rate on the loan.**

**installment:**

**The monthly payment owed by the borrower if the loan originates.**

**Objectives**

List of customers eligible for loans.

Minimize the risk for loan default.

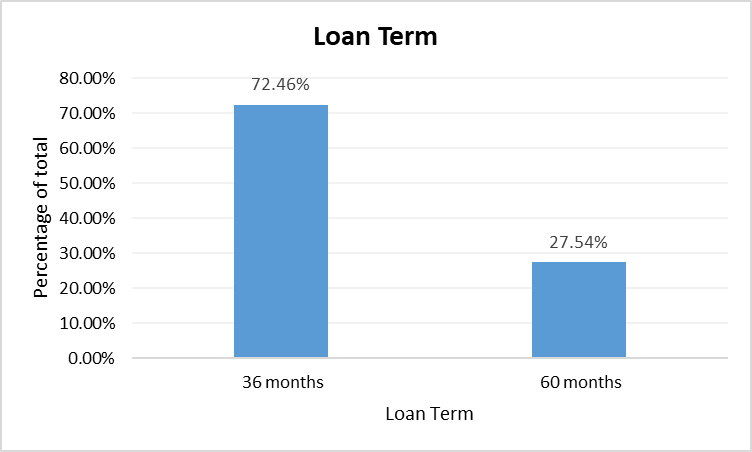
**Questions**

1. Which loan term decreases the risk of loan default? S
2. Does Home ownership indicate an eligible customer segment? S
3. Interest Rate vs loan status S,A
4. Loan Amount vs loan status S
5. Should the bank target specific states where the risk is min?S
6. Should the bank avoid specific states where the risk is max?S
7. Loan status vs issue date. S
8. Top 5 Titles who can afford a loan. S
9. Relation between loan grade and interest rate A
10. What are grades associated with loans charge-off and full payment A
11. Loan grade vs. annual income verification A
12. Is it more likely for borrowers with specific purposes to default? A
13. Term vs loan amount of charged off loans\*\*\*
14. Does Annual income and dti affect loan status? K
15. Is there a relationship between loan amount with term and status? M
16. Certain combinations of interest rate and loan amount that affect loan status for each purpose? M
17. Expected paid amount vs actual paid for each purpose M

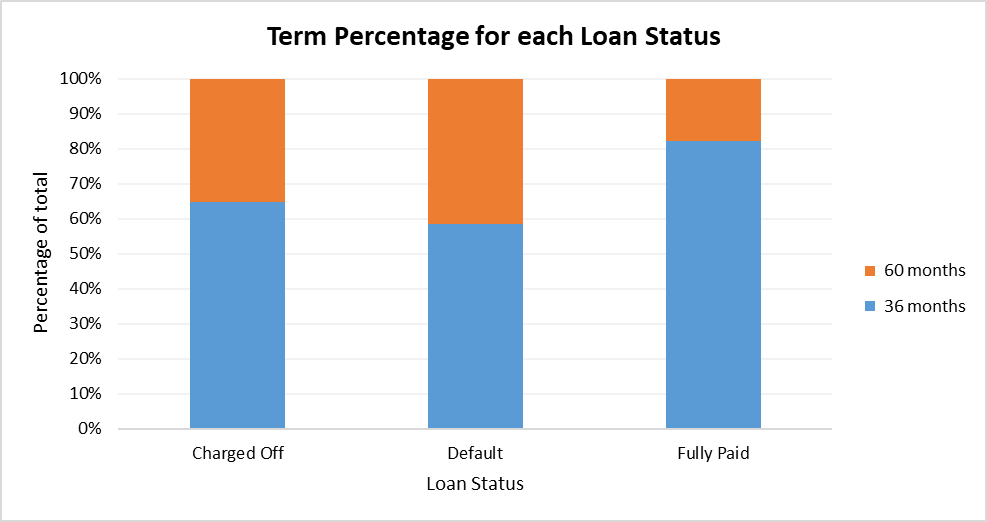
**Insights and Analysis Steps**

1. The percentage of charged off or default in our data is already low (9.47%)
2. **Loan Term Vs Loan Status**

We checked the **loan term**, it appears that the number of loans with the term 36 months is way more than those with the 60 month term.

****

This in our mind, let's check whether the term has an influence on the loan status.

****

Taking into account that there is a small amount of 60 months term loans compared to the 36 months ones, but that it takes up to 40% in the charged off and default columns, while nearly 80% of the fully paid loans has 36 months term, We can say that **Clients who are offered loans with a higher loan term are more likely to default on their loans.**

**Recommendation:**

The bank should consider offering loans with smaller loan terms to reduce the risk of default.

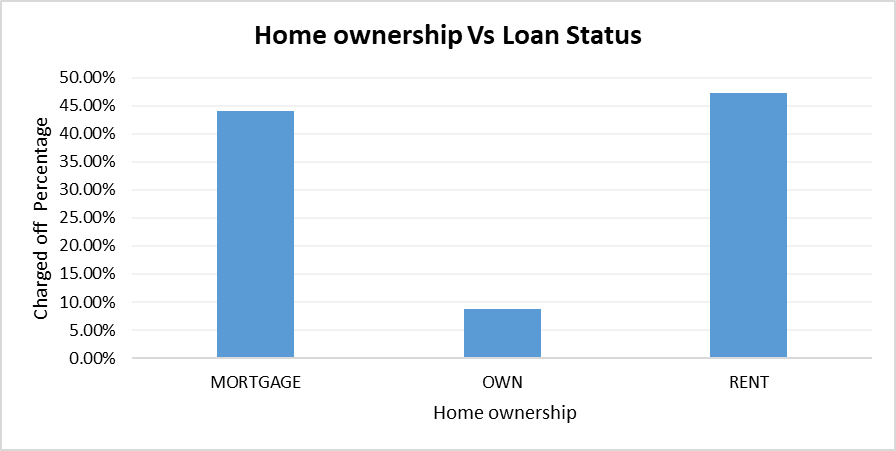
**Notes**:

**Longer loan terms may increase the risk of default**: Loans with longer terms typically have lower monthly payments, which can make them more affordable for borrowers in the short term. However, longer loan terms also mean that borrowers will be making payments for a longer period of time, which increases the likelihood of life events (such as job loss, illness, or unexpected expenses) that could disrupt their ability to make payments on time. This can increase the risk of default, particularly for borrowers who are already struggling financially.

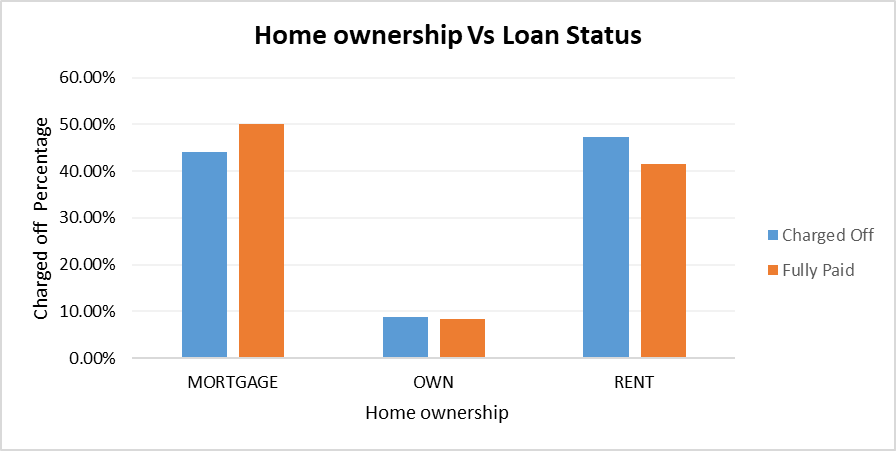
**Shorter loan terms may reduce the risk of default:** Loans with shorter terms typically have higher monthly payments, which can make them less affordable for borrowers in the short term. However, shorter loan terms also mean that borrowers will be making payments for a shorter period of time, which can reduce the likelihood of life events that could disrupt their ability to make payments on time. This can reduce the risk of default, particularly for borrowers who are more financially stable.

1. **Home Ownership vs Loan Status**

Inspecting **home ownership**’s effect on loan status, it appears for a while that rent or mortgage in home ownership indicates a big risk or high probability of default.



But, further analysis showed that home ownership doesn’t have a significant effect on loan status as for fully paid loans mortgage and rent also appears to have the biggest shares.

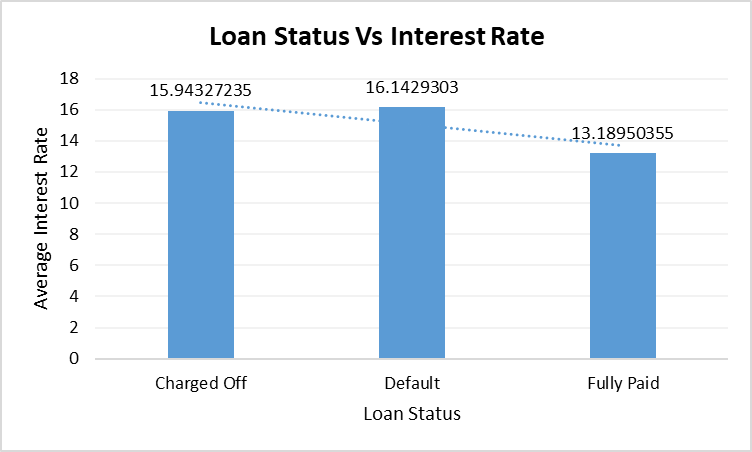


**Recommendation:**

Home ownership is not an indicator for loan default, so bank can use the campaign for all people whether they have a home or not without increasing the risk of default.

1. **Loan Status Vs Interest rate**

Now, Inspecting the interest rate for the fully paid and default or charged off loans, it appears that fully paid loans have on average smaller interest rates compared to defaulted loans.



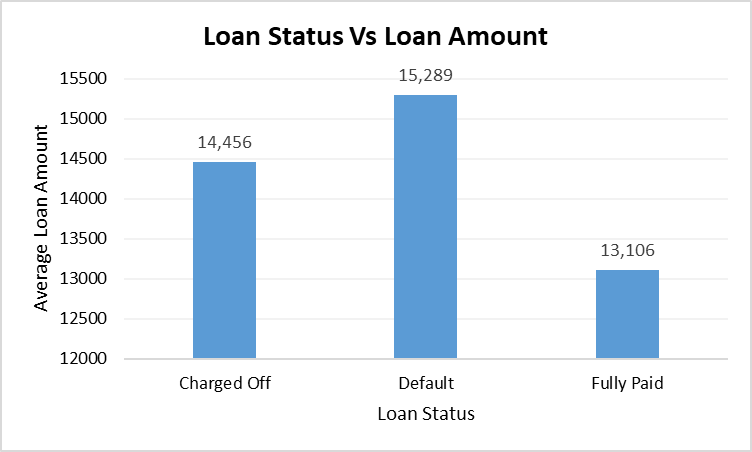
**Clients who are offered loans at a higher interest rate are more likely to default on their loans.**

**Recommendation:**

Therefore, the bank should consider offering loans at a reasonable interest rate to reduce the risk of default.

1. **Loan Status Vs Loan Amount**

By investigating the loan amount with the loan status we find that the defaulted loans have on average higher loan amounts than the fully paid loans.

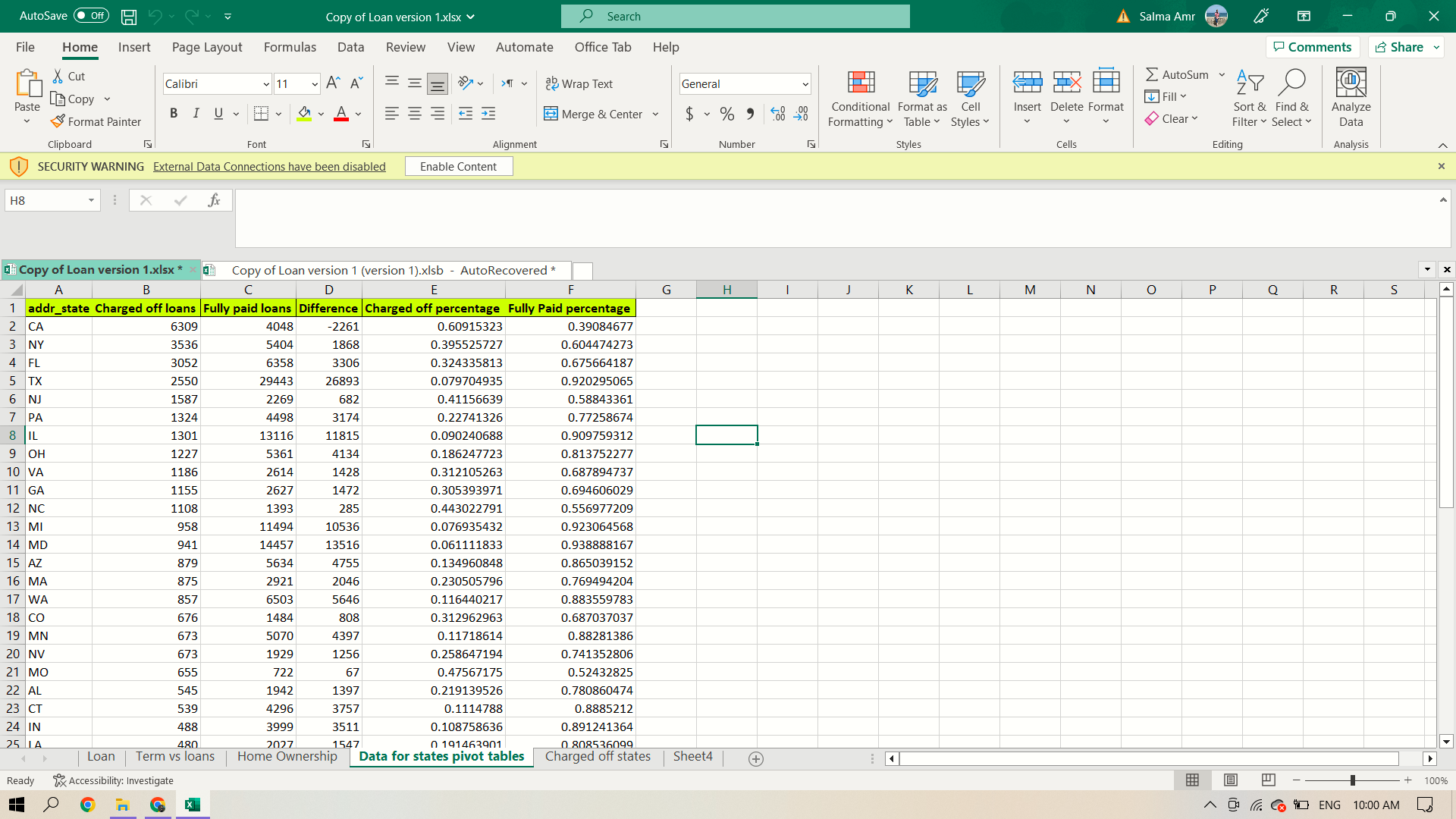


**Recommendation:**

The bank should consider offering smaller loans to clients who have a higher risk of default. This reduces the bank's risk exposure and improves the chances of repayment.

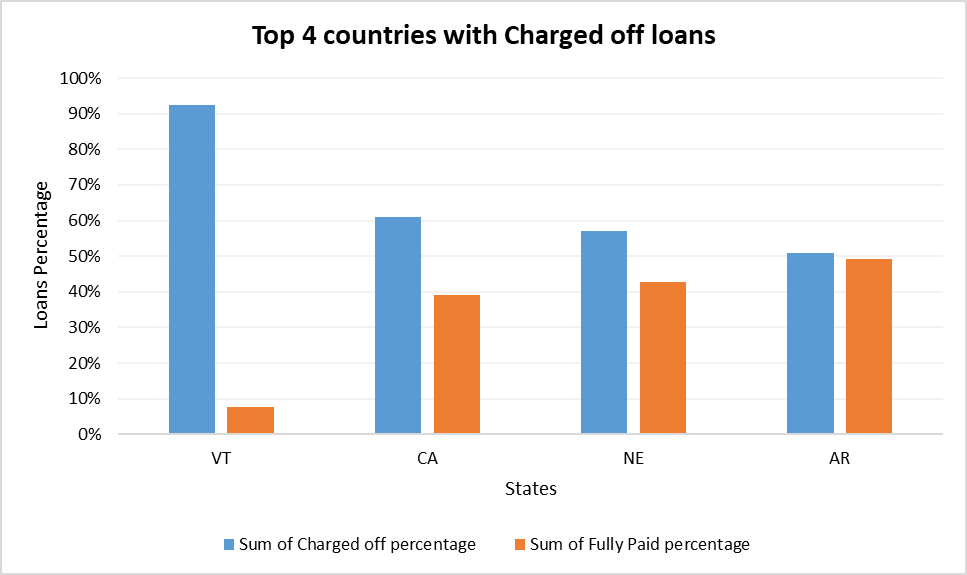
1. **Address State Vs Loan Status**

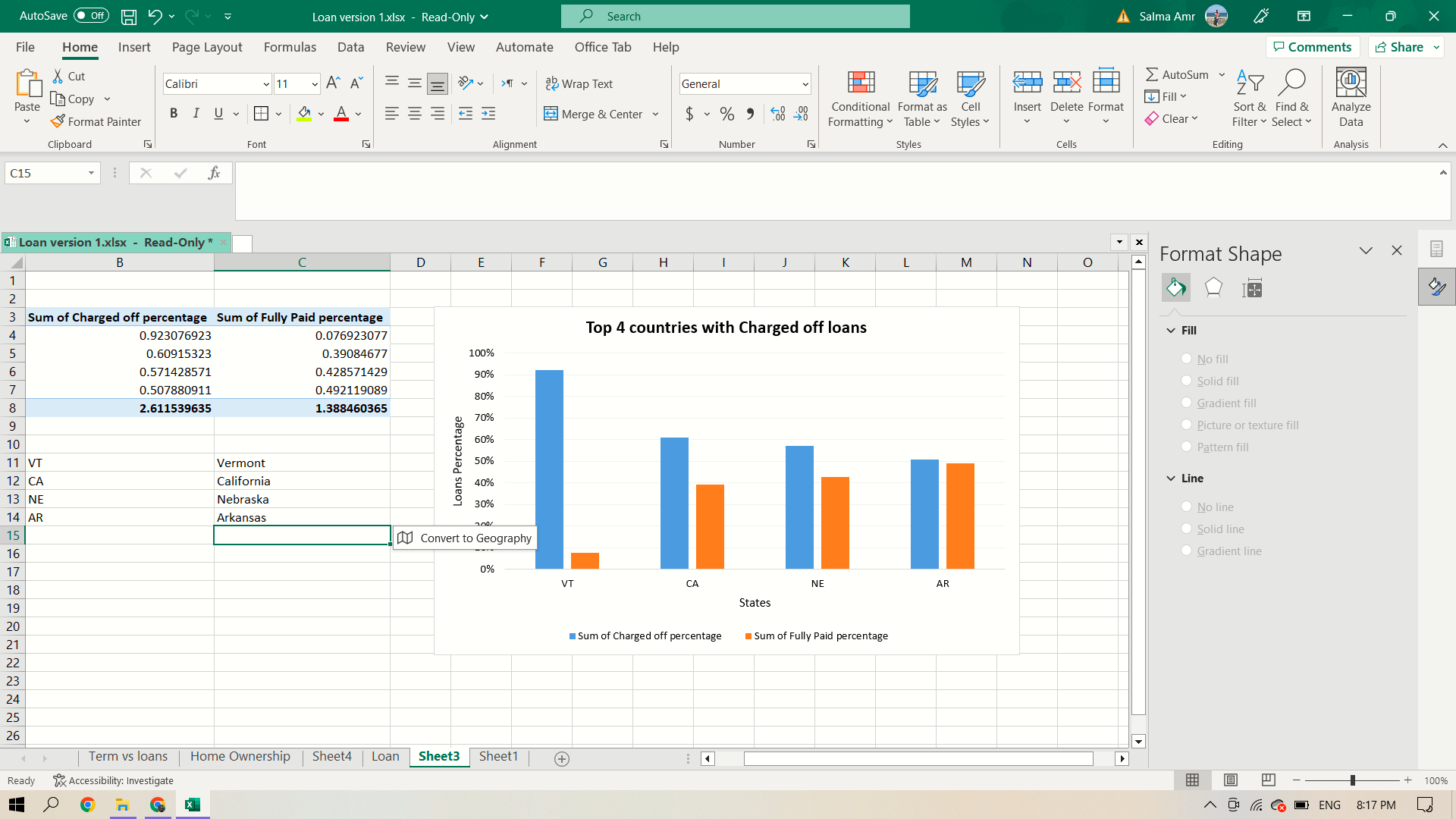
After calculating the sum of charged off loans and the sum of fully paid loans for each state, we calculated the percentage of fully paid and percentage of charged off loans for each state from the total of charged off + fully paid loans.

****

**From that we found:**

**The 4 countries that have the highest percentage of charged off loans** and also the highest difference in count between the charged off and fully paid loans, These were vermont, california, nebraska and arkansas.



****

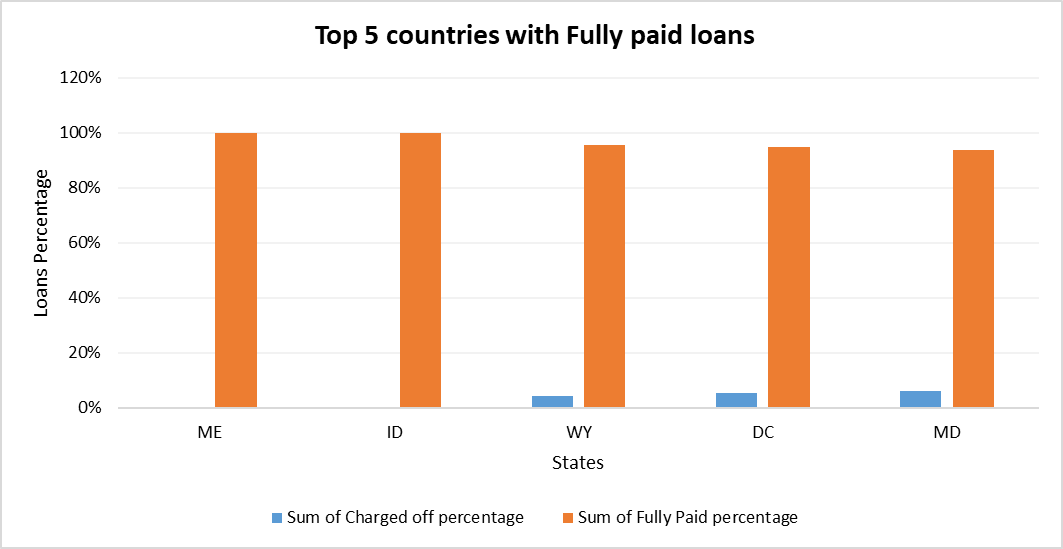
**Recommendation:**

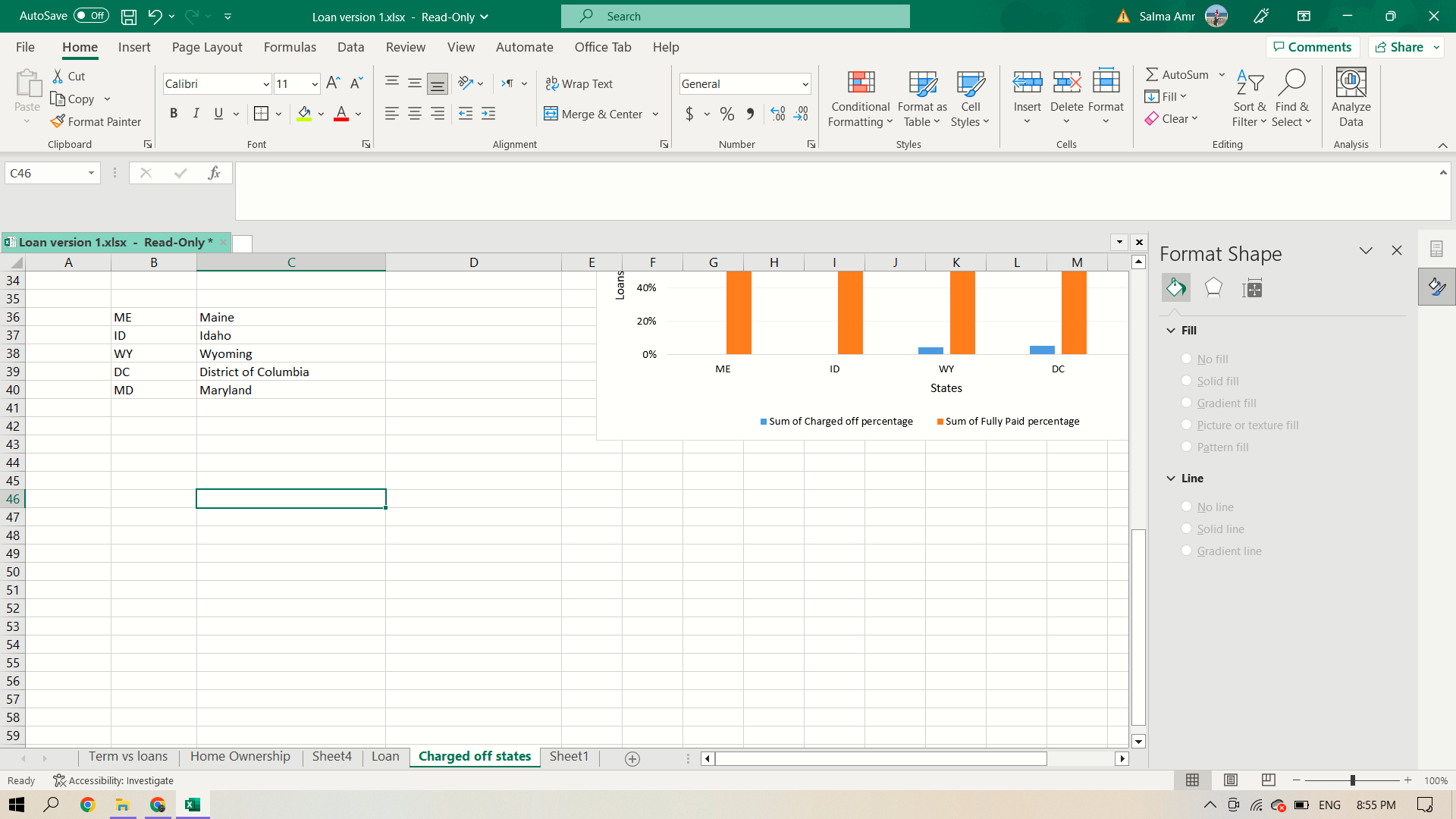
The bank may want to consider targeting states with a higher percentage of fully paid loans in its marketing campaign.

This is because these states may indicate a more financially stable customer base with a lower risk of loan default. By targeting these states, the bank can increase its chances of attracting more creditworthy customers and reducing the risk of loan defaults.

We also found **the 5 countries that have the highest rate of fully paid loans**.

Those countries are Maine, Idaho, Wyoming, District of Columbia and Maryland.

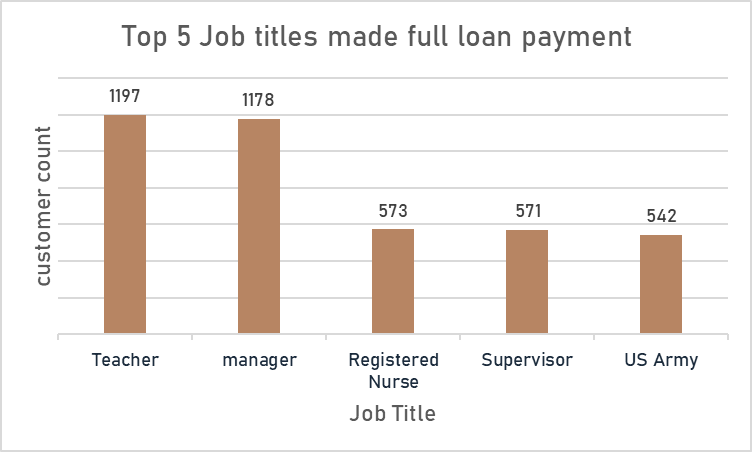
****

****

**Recommendation:**

The bank may want to be more careful with loans from states with higher percentages of loan defaults . This is because a higher rate of loan defaults in a state may indicate a higher risk of lending to customers in that state.

1. **Top 5 Titles who can afford a loan**



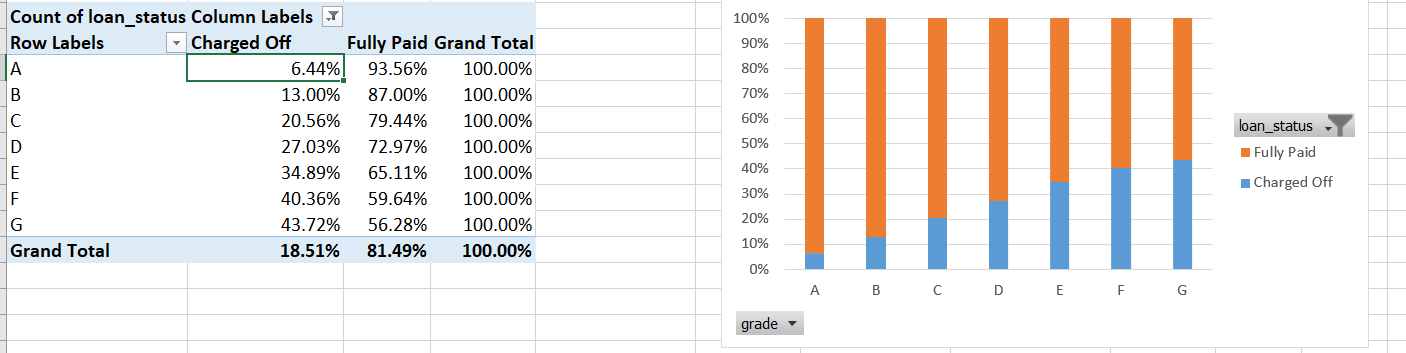
**Recommendation:**

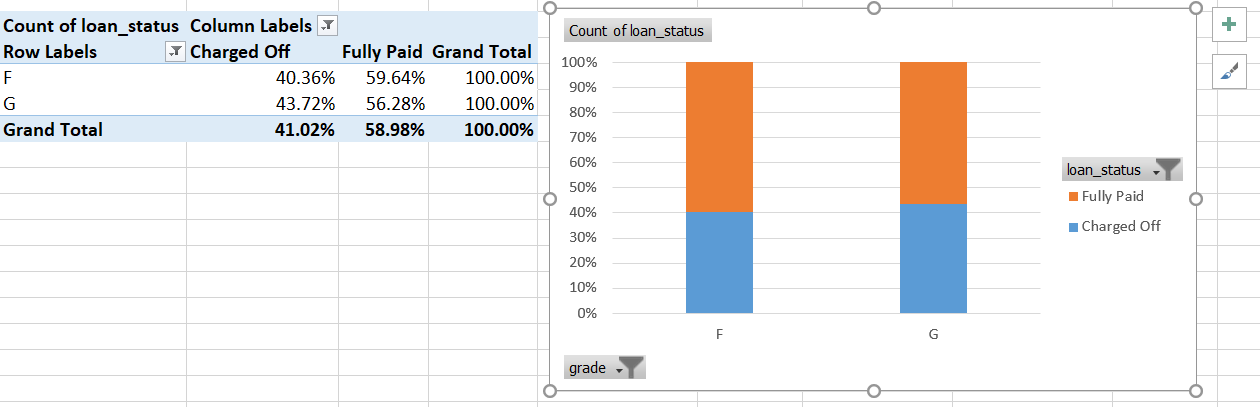
The bank may target specific job roles that proved to be able to fully pay their loans, these job roles may indicate a higher income or greater financial stability, which could make these individuals more attractive to the bank as potential borrowers as it can minimize the risk of loan default.

### \*\*Relation between loan grade and interest rate

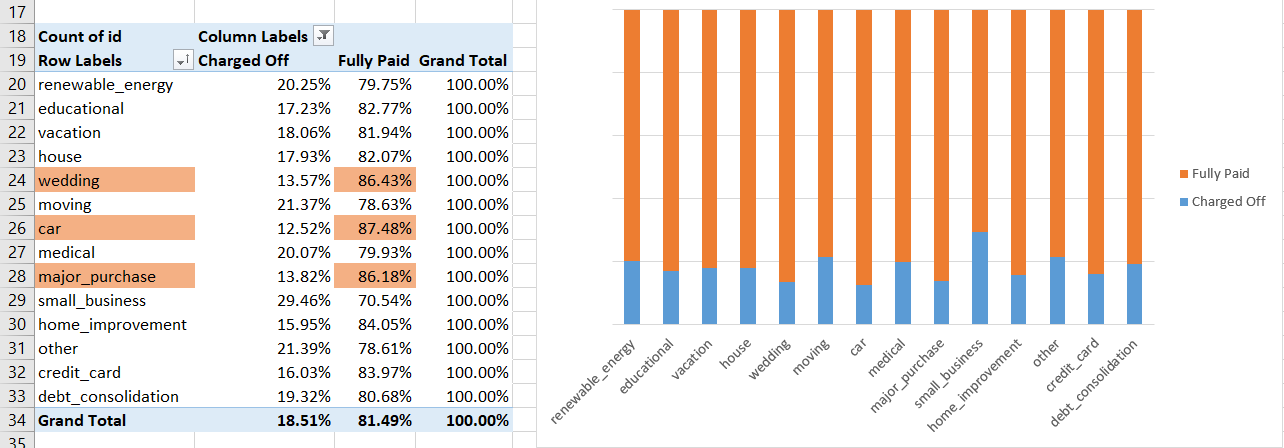
### 

### \*\*What are grades associated with loans charge-off and full payment





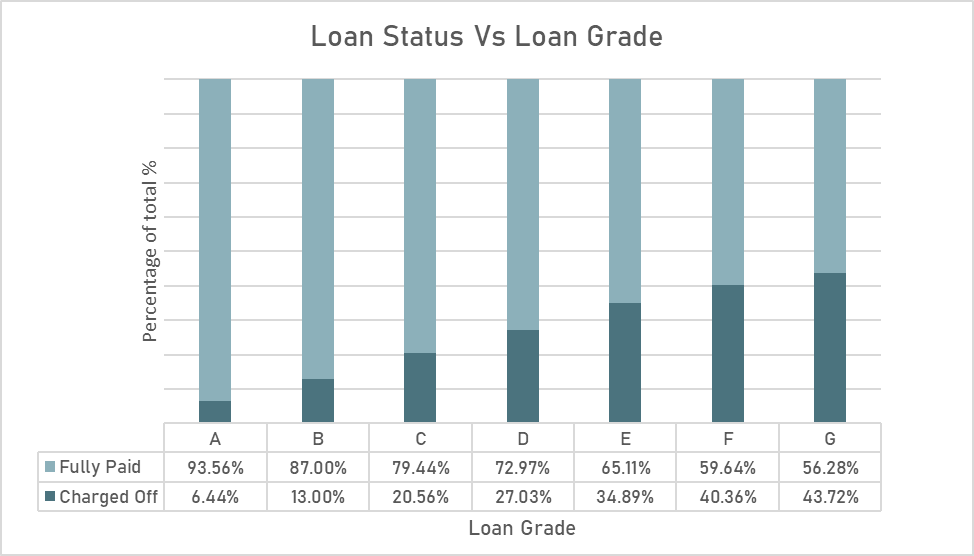
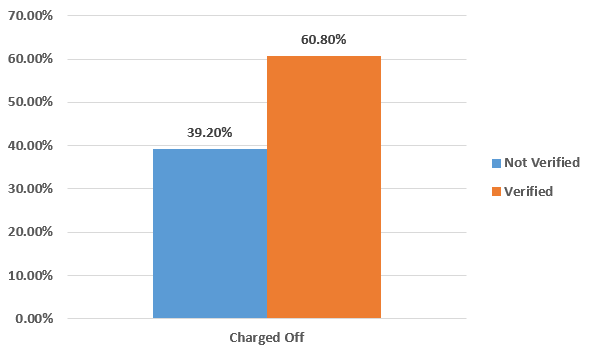
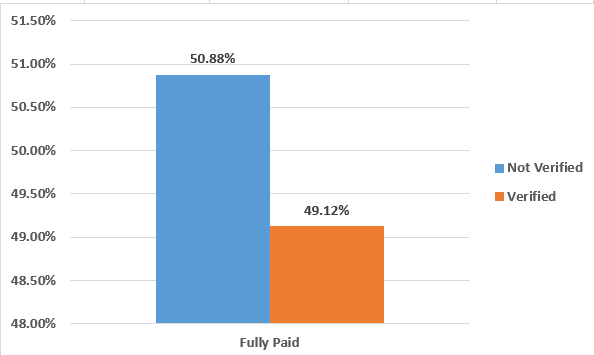
1. \*\*Loan status vs. purpose >> car, major purchases, wedding



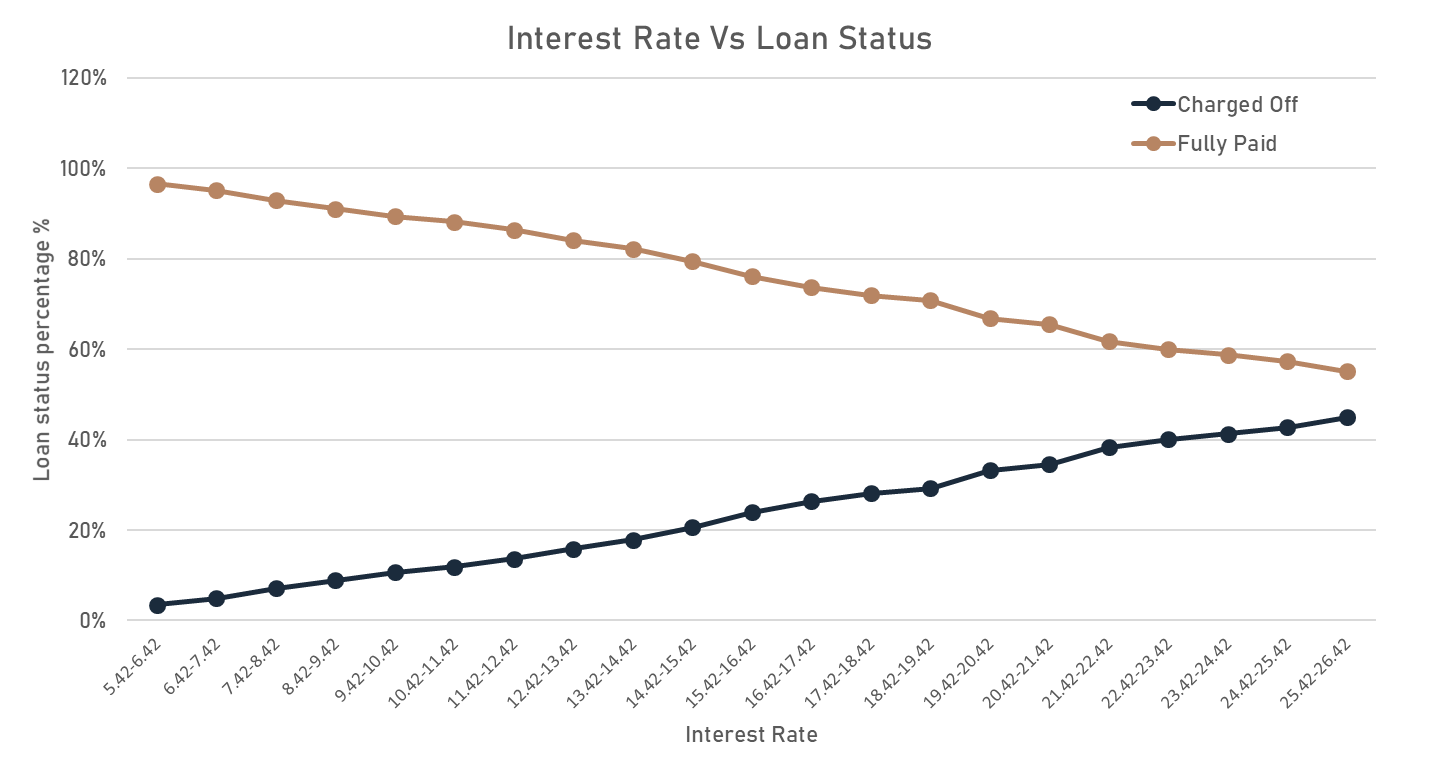
### Loan grade vs. annual income verification

### 

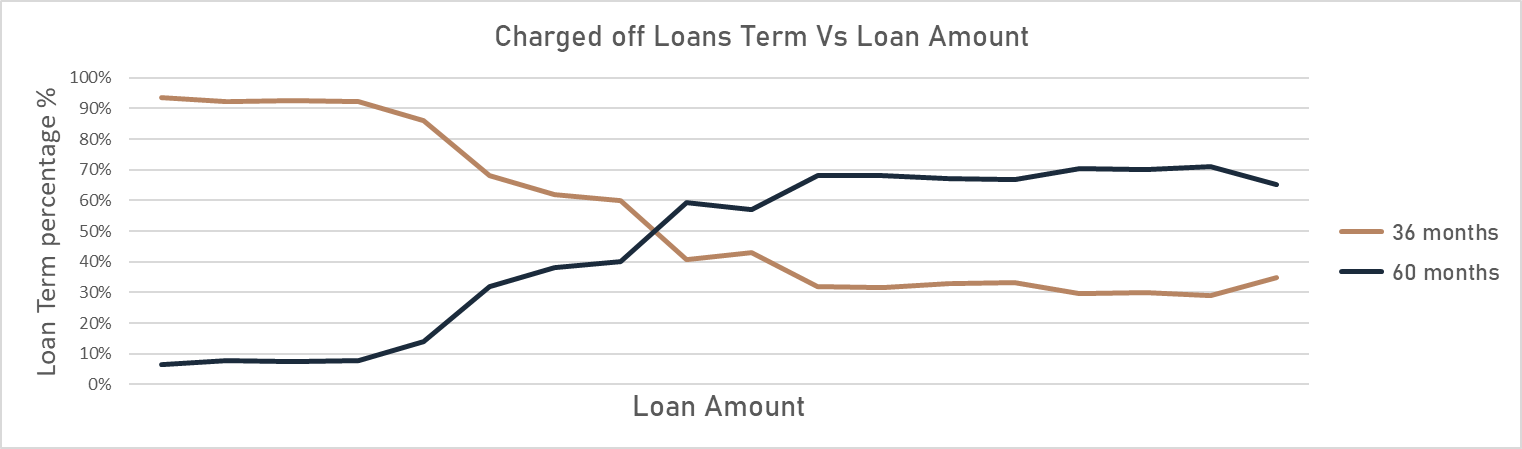
### Loan status vs annual income verification



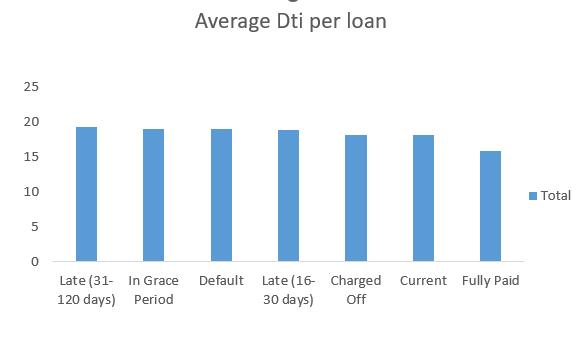
1. Interest rate vs loan status



1. Charged off Loans Term Vs Loan Amount



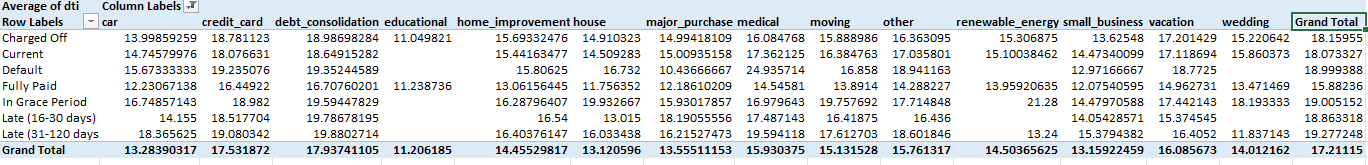
1. **Average Dti per loan**

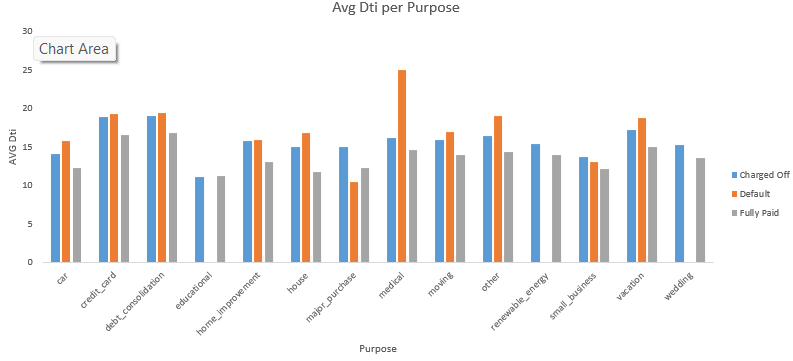


We see that fully paid and current has the lowest Dti ratio so we need to inspect the Dti more

To see if the behavior remains true for each purpose or not.

1. **Average Dti per purpose**





Most of the purposes avg for fully paid is less than charged off which is normal but for some purposes like educational renewable energy for all loan status they are close so they have higher risk for giving loans for these purposes.

And for recommendation here are some avg Dti to lower the risk for loans

wedding, purchases, educational need data to be sure

Small business 12 and below

Renewable energy 13-15 and below more confident will be <=13

Vacation 14-15 and below

Credit card 16-17 and below

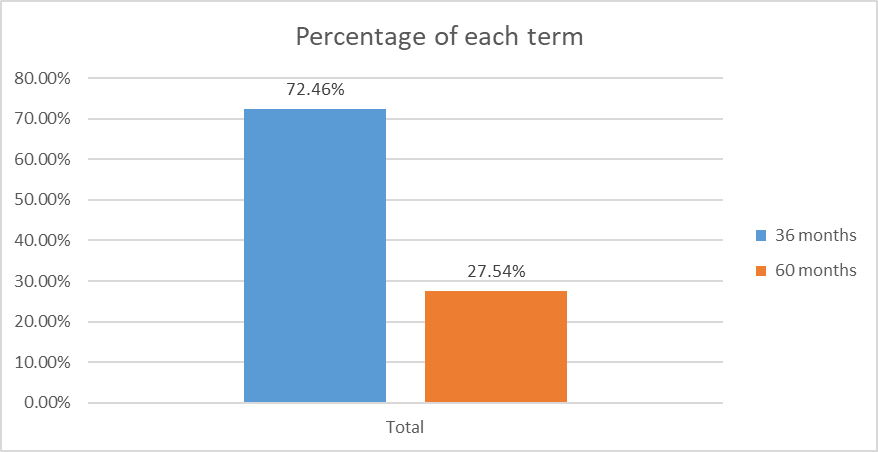
Debt consolidation 16-18 and below

Other 14-15 and below

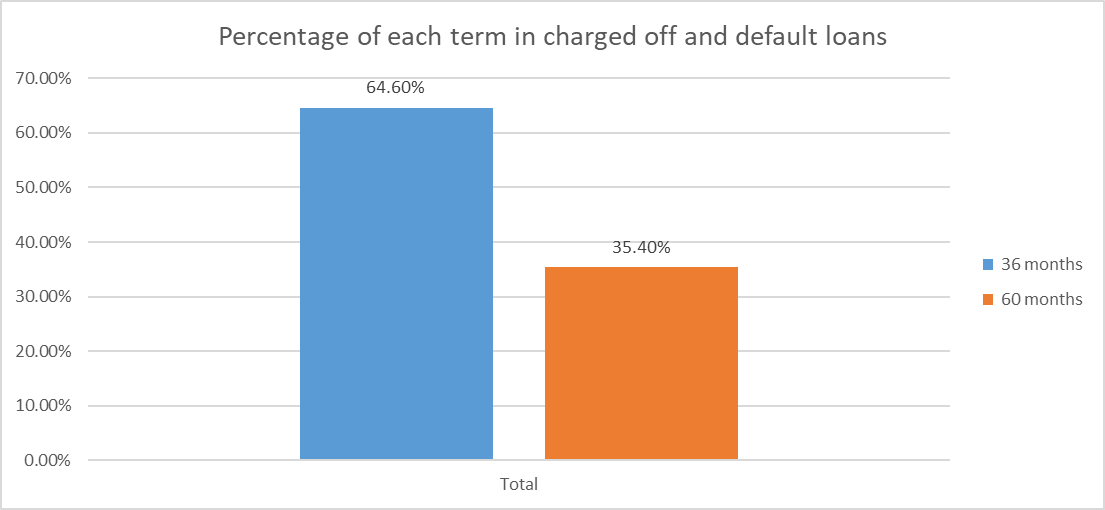
These number build on average Dti so there could be some tolerance with the recommended values

**15) Is there a relationship between loan amount with term and status? M**

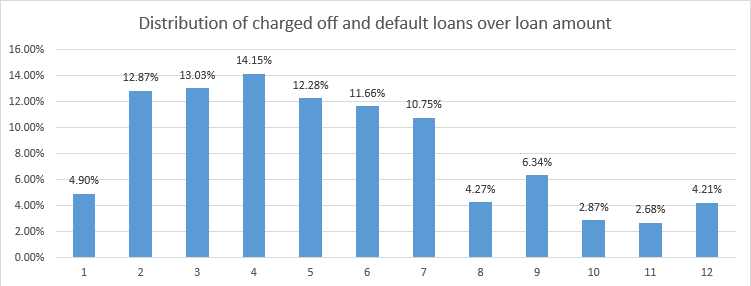
First we need to look at the distribution of terms



Then we need to look at the same distribution but in charged off and default loans

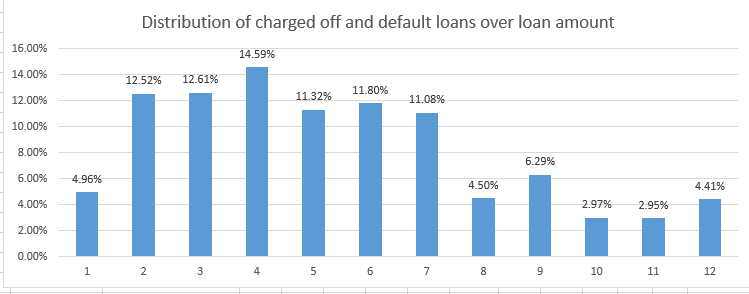


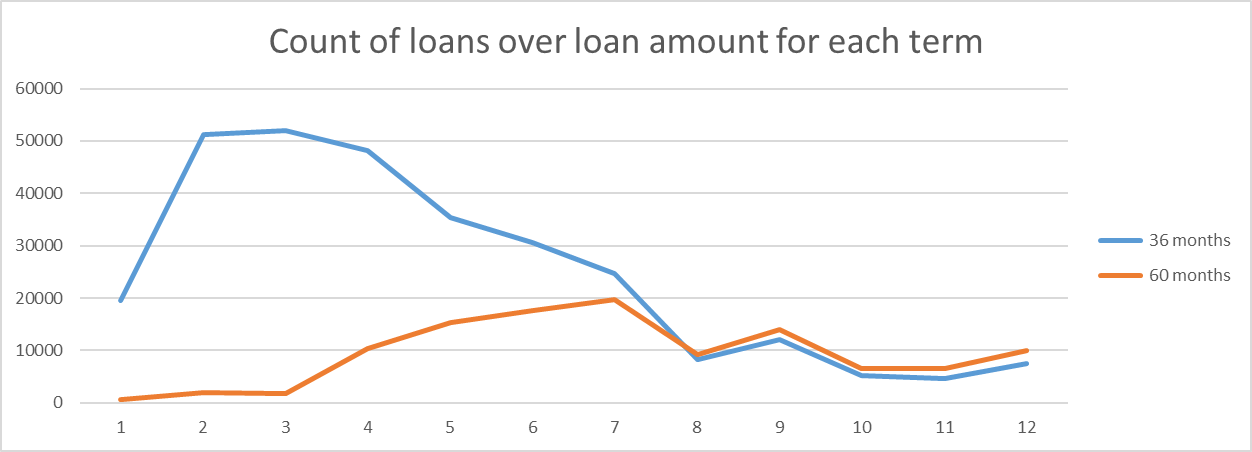
Here is the distribution of loans over loan amount

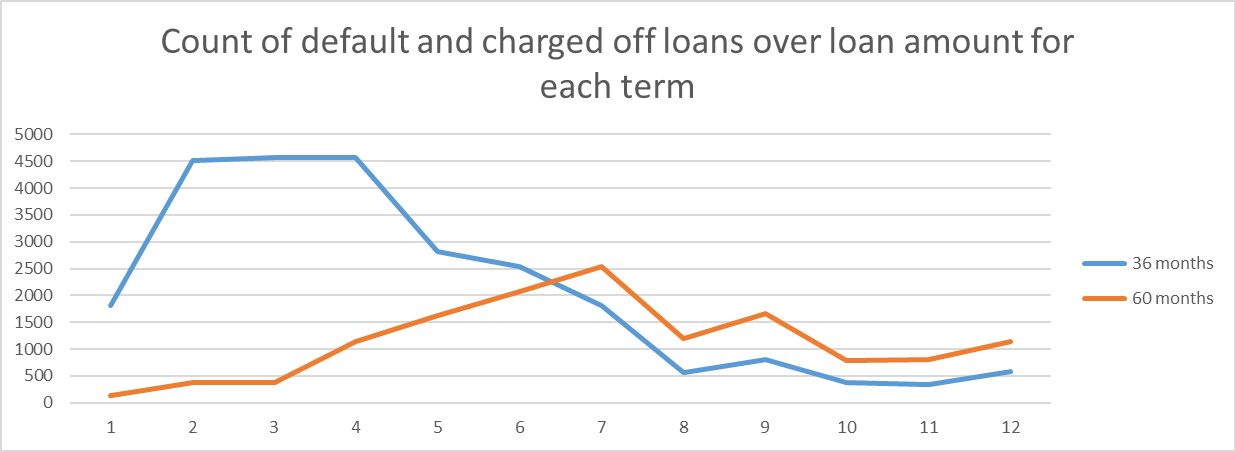


Look at the same distribution but in charged off and default loans

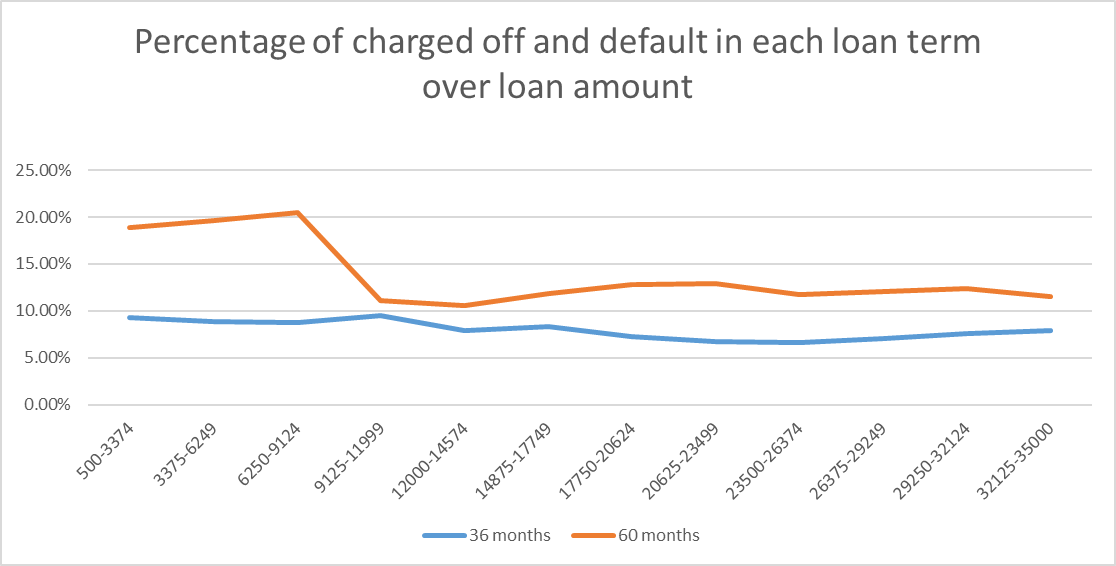
Now we need to split our analysis based on terms to get deeper, both in total loans and in charged off and default loans







Nothing is much surprising but, what is surprising is the percentage of charged off and default in each term over the loan amounts



**Insight**: Having a loan in 60-months term doesn’t mean its safer to be fully paid as 60-months term contains a higher percentage of charged off and default loans

**Recommendation**: although 36-months term loans are a lot more than 60-months loans thus 36-months term is more concerning but, in future decisions don’t consider 60-months term as a safer loan to be paid (based on history)

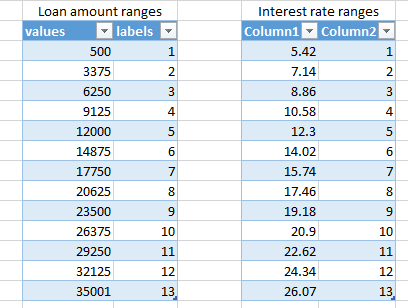
**16) Certain combinations of interest rate and loan amount that affect loan status for each purpose? M**

It is expected that the larger and higher the interest rate the more likely this loan will default or get charged off, but we need to dig deeper to find certain value of interest rate that loans tend to default and get charged off exceeding it

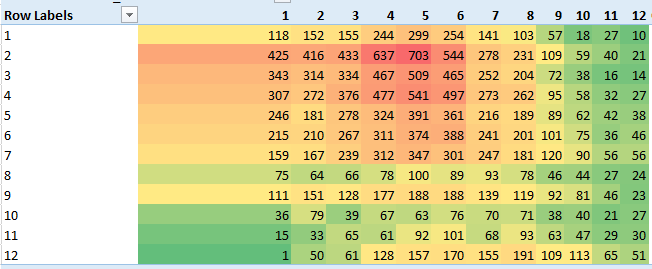
Let us look at the top 4 purposes in count that default or charged off. Debt consolidation, credit card and home improvement.

Loan amount and interest rate are divided into 12 ranges for this analysis

We will be splitting loan amount and interest rate as follows for easier plotting

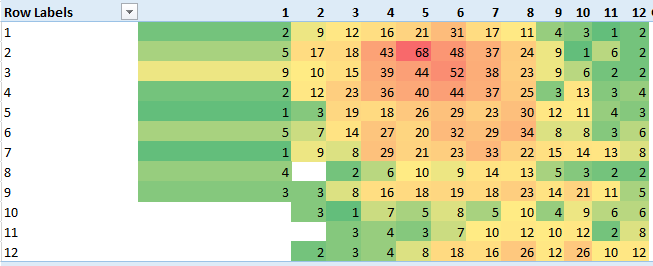


**Home improvement**

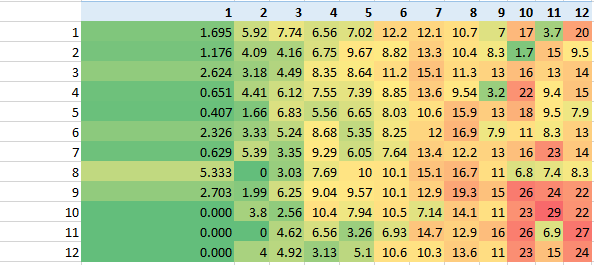


Here is the distribution of all home improvement loans with respect to loan amount on x-axis

And interest rate on y-axis



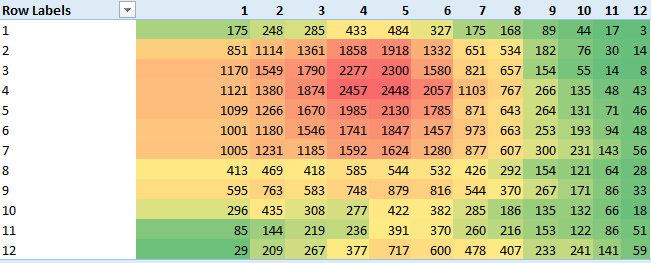
And here is the distribution of charged off and default home improvement loans



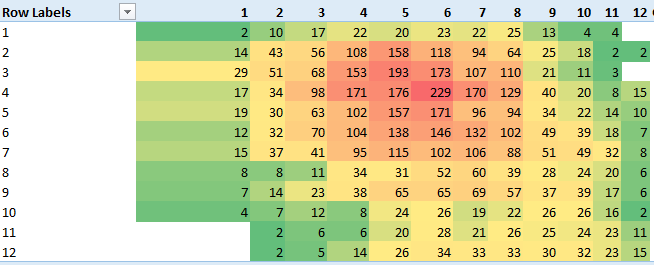
When looking at the percentage we can see that majority of high percentages of charged off and default is at the two extremes

**Insight**: by average, any loan with home improvement purpose, amount more than 23,500 and interest rate higher than 20.9% is by average 25% likely to charge off or default based on history

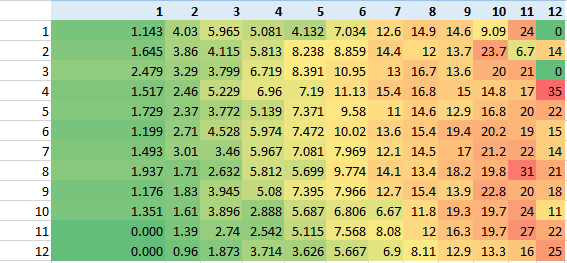
**Next, loans with purpose of credit cards**



Here is the credit card loans count



And here is the count of its charged off and default loans



And this is how charged off and default loans percentages looks like with respect to loan amount and interest rate

More analysis is needed regarding that matter

**Insight**: loans with credit card purpose looks healthy with average charged off and default percentage of 6.9% in the highest loan counts

( with loan amount range 3,375$-17,750$ and interest rate range 8.86%-14.02%

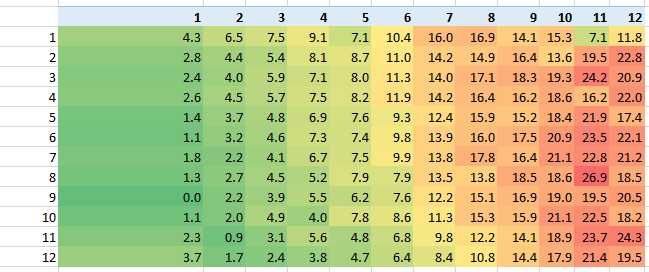
Having charged off and default loans is almost independent of loan amount and directly correlated with the increase of interest rate

Loans with amount range 9,125$-12,000$ and interest rate range 24.34%- 26.07% alone are with almost 35% defaults and charged off

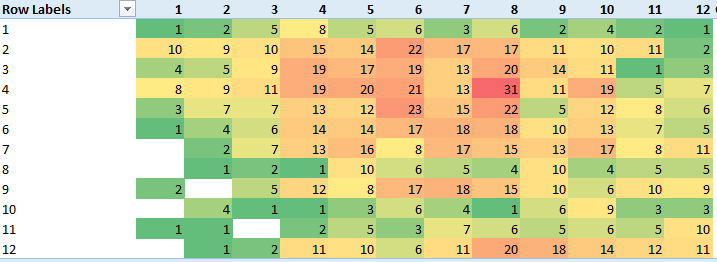
Charged off and default percentages start reaching 20% and higher at interest rate of 20.9% or greater

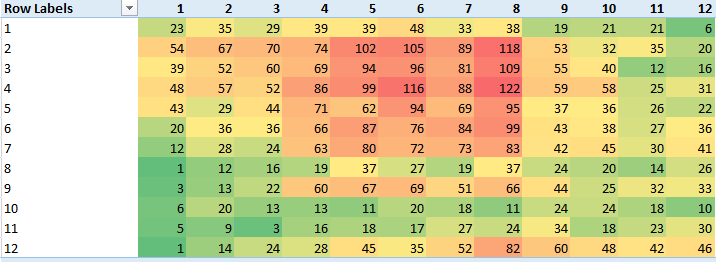
**Next, loans with debt consolidation purpose**

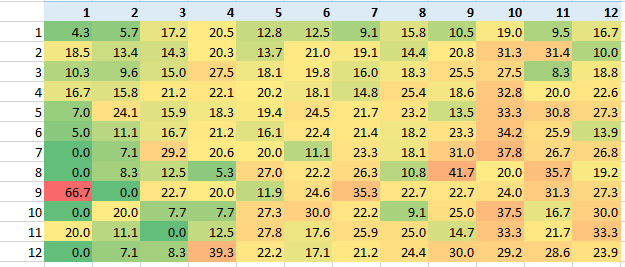
The same is for debt consolidation

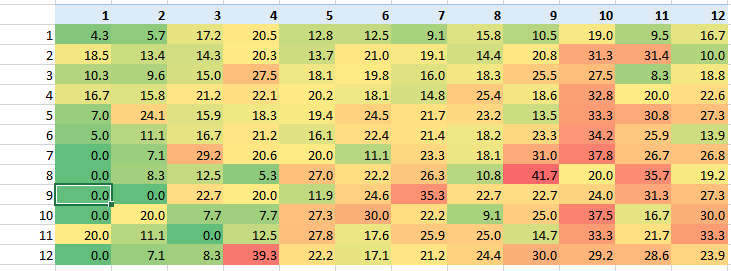


**Next, loans with small business purpose**

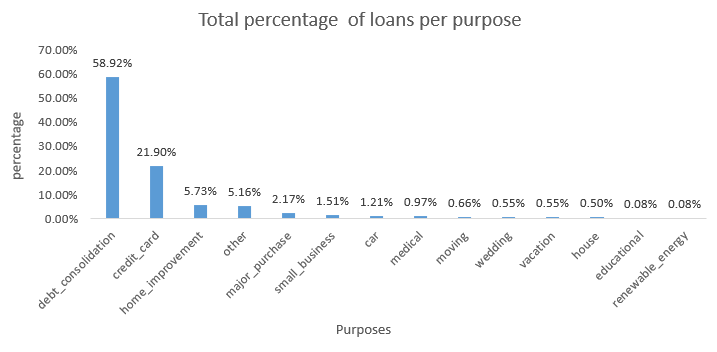








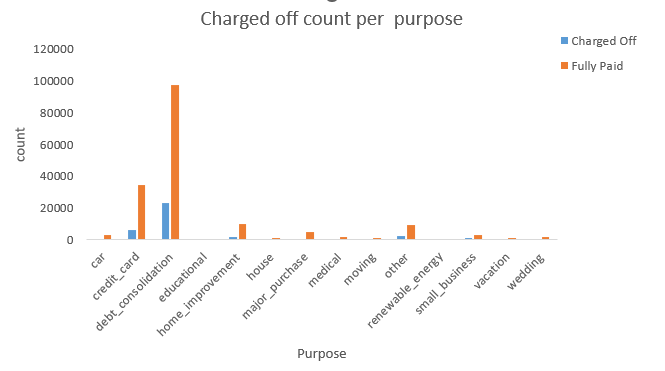
**Insight**: default and charged off cases are non uniform with respect to loan amount and interest rate

**17) what is the most frequent purpose** 

So debt consolidation has the higher percentage anthem credit card

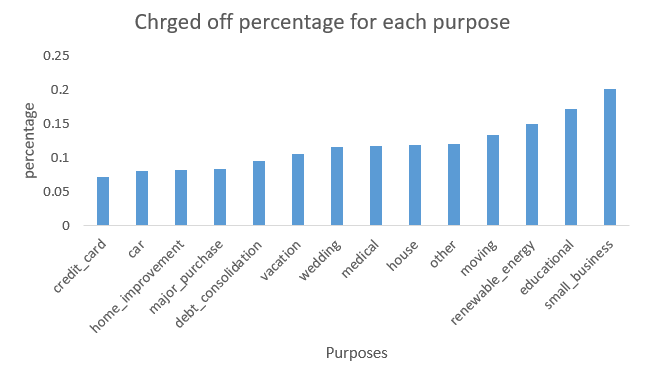
Let’s go deeper

By knowing the most frequent with charged off and default counts



From the first look we see that debt consolidation and credit card has the most Charged off but the real question would be are these values indicative or not

It is obvious that the most frequent purposes probably have the most values for fully paid or charged off so let’s see what is the most charged of the purpose itself not in general

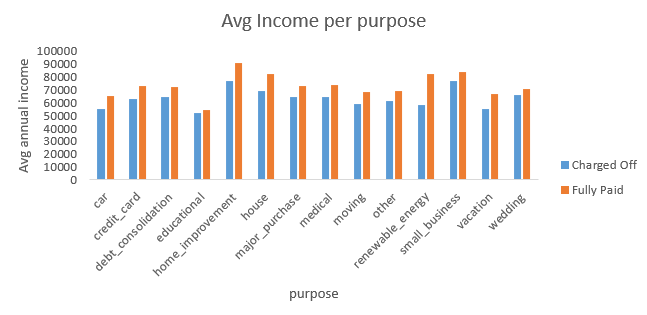


The most charged off percentage is small business then educational and renewable energy.

so as recommendation I would say these purposes has high risk and these purpose has low dti rate and even they defaulting from low dti

If there is a marketing campaign then the most appropriate choice would be credit card, car, home improvement, purchases and debt consolidation.

**18) what is the Average salary per purpose**



Here is the recommendation for the average annual salary based on the purpose.

Car >=65000

Medical >=70000

Home\_improvement >=80000

House >=70000

Moving >=66000

Small\_business >=84000

Renewable\_enregy >=78000

Vacation >=60000

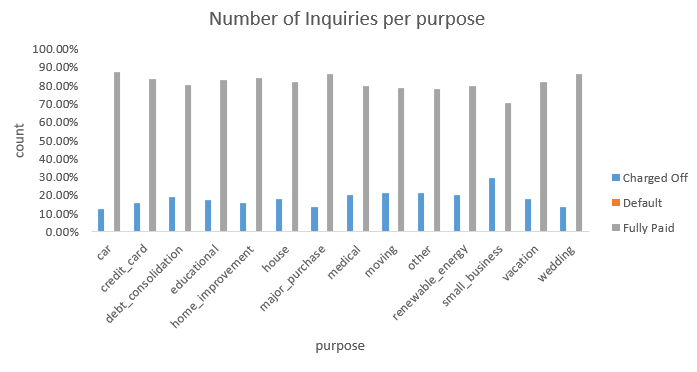
Other >=68000

credit\_Card >=69000

Debt\_consolidation >=72000

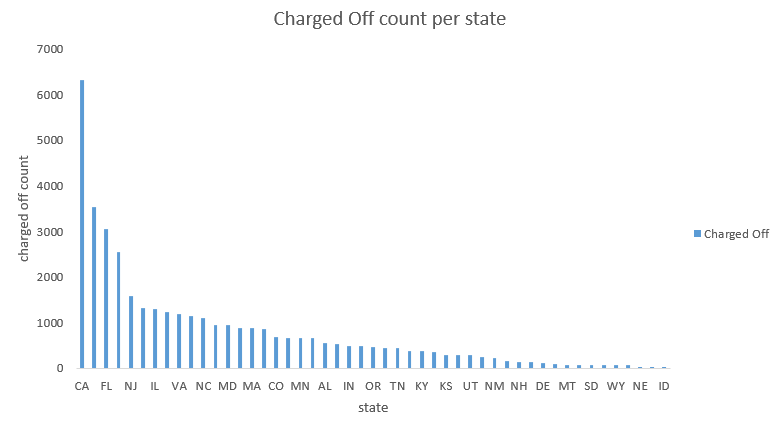
Wedding >=70000

**19) what is the number of Inquiries per purpose**



Higher number of Inquiries mean low probability to charge off.

**20) what is the most states with high charge off counts**

****

California, Florida, new Jersey, Illinois, Virginia these states has the most charged off number So there is high risk in the states due to some factors like high cost living

While on the other hand states like wyoming, south Dakota, Delaware have low risk

So for a marketing campaign we should target the low risk states more.