

# LENDING CLUB ASSIGNMENT SUBMISSION

Name: Monish Mathpal

## **Abstract: Lending Club Investment**

- The company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. The company can utilise this knowledge for its portfolio and risk assessment.

# Problem solving methodology



Grade change impact on determination of a consumer going to default or not

	Defaulter_prct	Non_Defaulter_prct	Diff_Defaulter_prct_Non_Defaulter_prct
A	10.70	28.66	-17.96
B	25.32	31.11	-5.79
C	23.94	19.69	4.25
D	19.87	12.04	7.83
E	12.71	5.91	6.80
F	5.67	1.99	3.68
G	1.79	0.60	1.19

Weight of evidence analysis shows mths\_since\_last\_delinq is a strong predictor as the **information value 0.36** falls between **0.3 to 0.5 range**.

Information Value (IV)	Predictive Power
< 0.02	useless for prediction
0.02 to 0.1	weak predictor
0.1 to 0.3	medium predictor
0.3 to 0.5	strong predictor
> 0.5	suspicious or too good to be true

min_val_	max_val_	count_	count_events_	count_non_events_	%_count_events	%_non_count_events	WOE	diff_event_%_non_event	IV
0.0	20.0	3955	647	3308	30.21	28.69	0.051624	1.52	0.08
21.0	50.0	6132	907	5225	42.34	45.32	-0.068016	-2.98	0.20
51.0	120.0	3585	588	2997	27.45	25.99	0.054654	1.46	0.08

## **CHI SQUARE analysis**

### **title and is defaulter derv:**

- Looking at the chi\_2 result, it becomes apparent that there is no significant influence of title on the chances of getting default or not.
- Chi\_Square Result:
  - P\_value for chi\_2 test between is\_defaulter\_derv and title 0.96

### **purpose and is defaulter derv:**

- Looking at the chi\_2 result, it becomes apparent that there is a significant influence of purpose on the chances of getting default or not.
- Chi\_Square Result:
  - P\_value for chi\_2 test between is\_defaulter\_derv and purpose 0.0

### **home\_ownership and is\_defaulter\_derv:**

- Looking at the chi\_2 result, it becomes apparent that there is a significant influence of home\_ownership on the chances of getting default or not.
- Chi\_Square Result:
  - P\_value for chi\_2 test between home\_ownership and is\_defaulter\_derv 0.0

### **pub\_rec and is\_defaulter\_derv:**

- Looking at the chi\_2 result, it becomes apparent that there is a significant influence of pub\_rec on the chances of getting default or not.
- Chi\_Square Result:
  - P\_value for chi\_2 test between pub\_rec and is\_defaulter\_derv 0.0

### **emp\_length and is\_defaulter\_derv**

- Looking at the chi\_2 result, it becomes apparent that there is a significant influence of emp\_length on the chances of getting default or not.
- Chi\_Square Result:
  - P\_value for chi\_2 test between emp\_length and is\_defaulter\_derv 0.01

When the Target variable changes from Non\_Defaulter to Defaulter

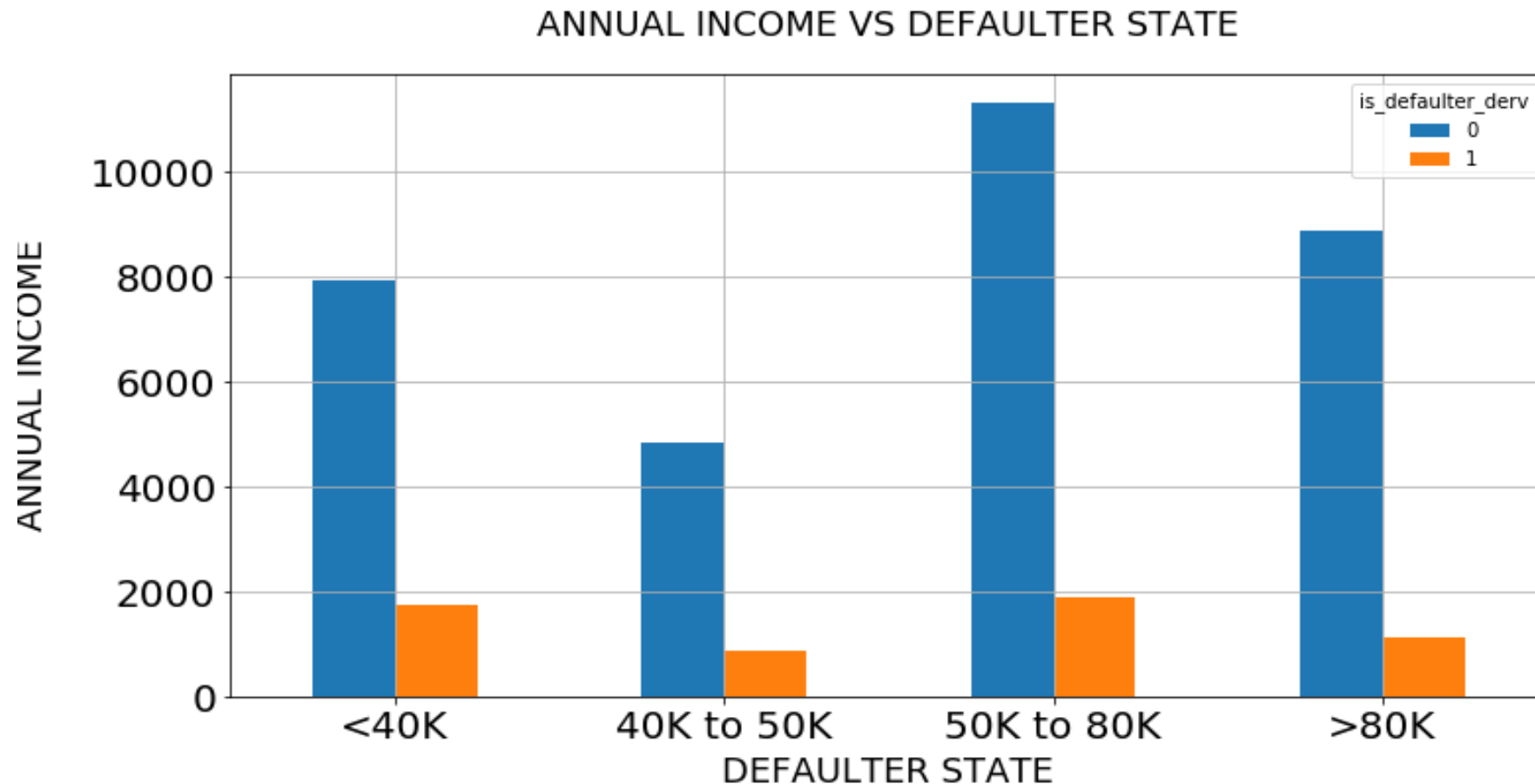
- There is a significant **negative change of (-17.96)%** in GRADE A.
- There is a significant **negative change of (-5.79)%** in GRADE B.
- There is a significant **positive change of 4.25%** in GRADE C.
- There is a significant **positive change of 7.83%** in GRADE D.
- There is a significant **positive change of 6.80%** in GRADE E.
- There is a significant **positive change of 3.68%** in GRADE F.
- There is a slight **positive change of 1.19%** in GRADE G.



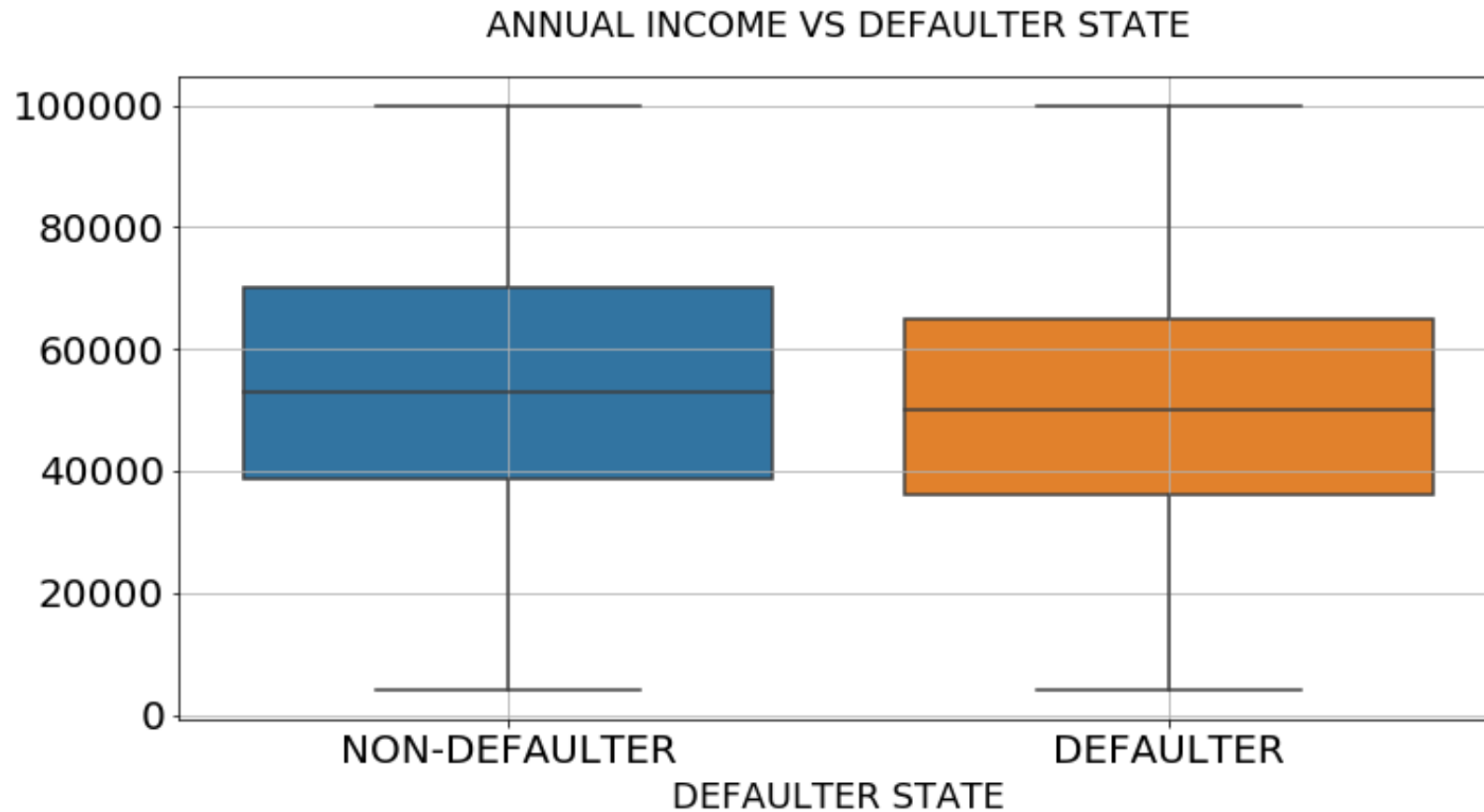
## <Analysis>

	Defaulter_prct	Non_Defaulter_prct	Diff_Defaulter_prct_Non_Defaulter_prct
0	91.70	95.14	-3.44
1	8.12	4.72	3.40
2	0.18	0.12	0.06
3	NaN	0.02	NaN
4	NaN	0.01	NaN

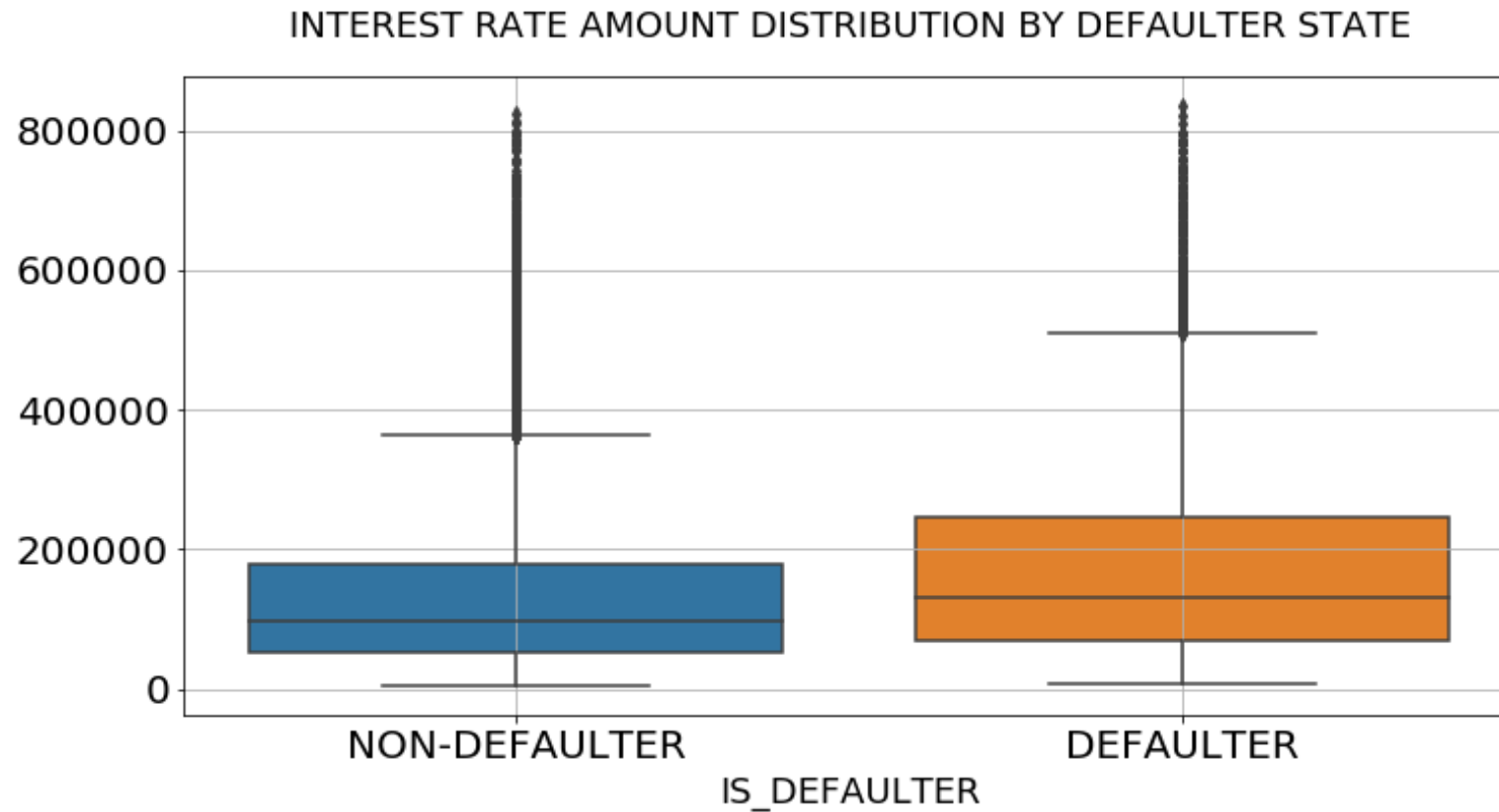
Annual Salary less than 40K and between 50K to 80K have got more chances to default.



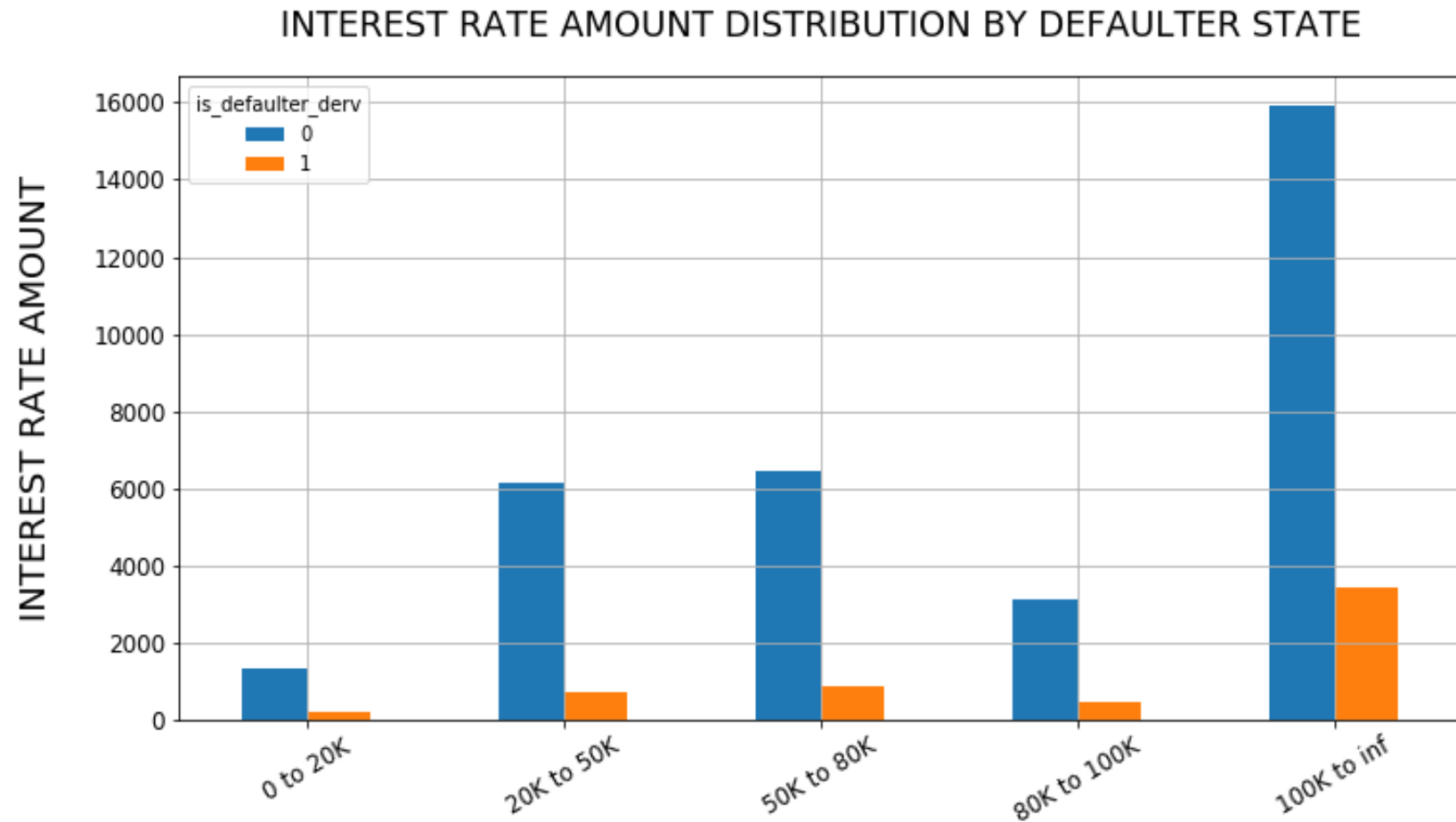
The chances of getting default gets influenced by annual income range.



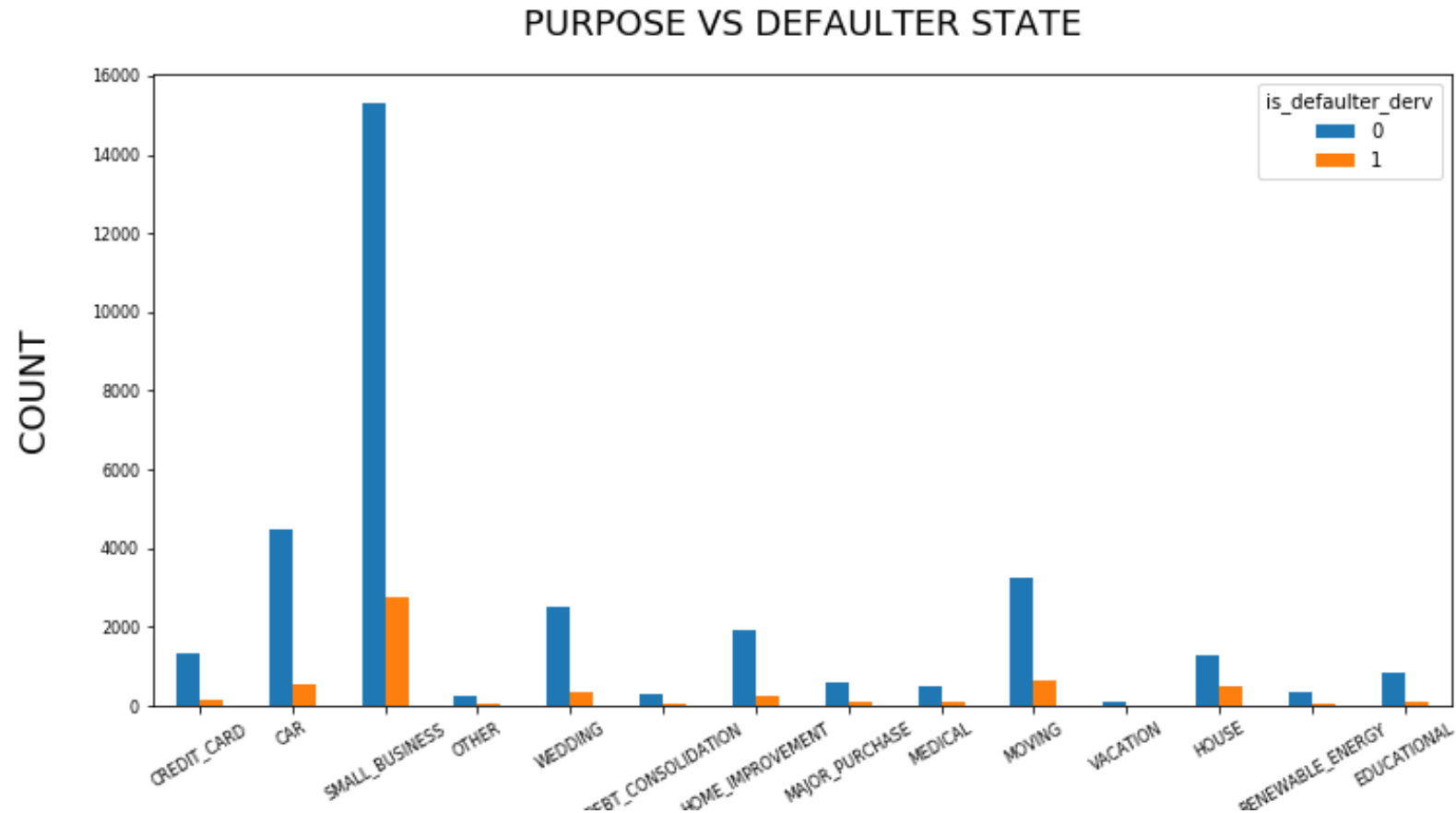
The higher interest rate amount increases the chances of getting default.



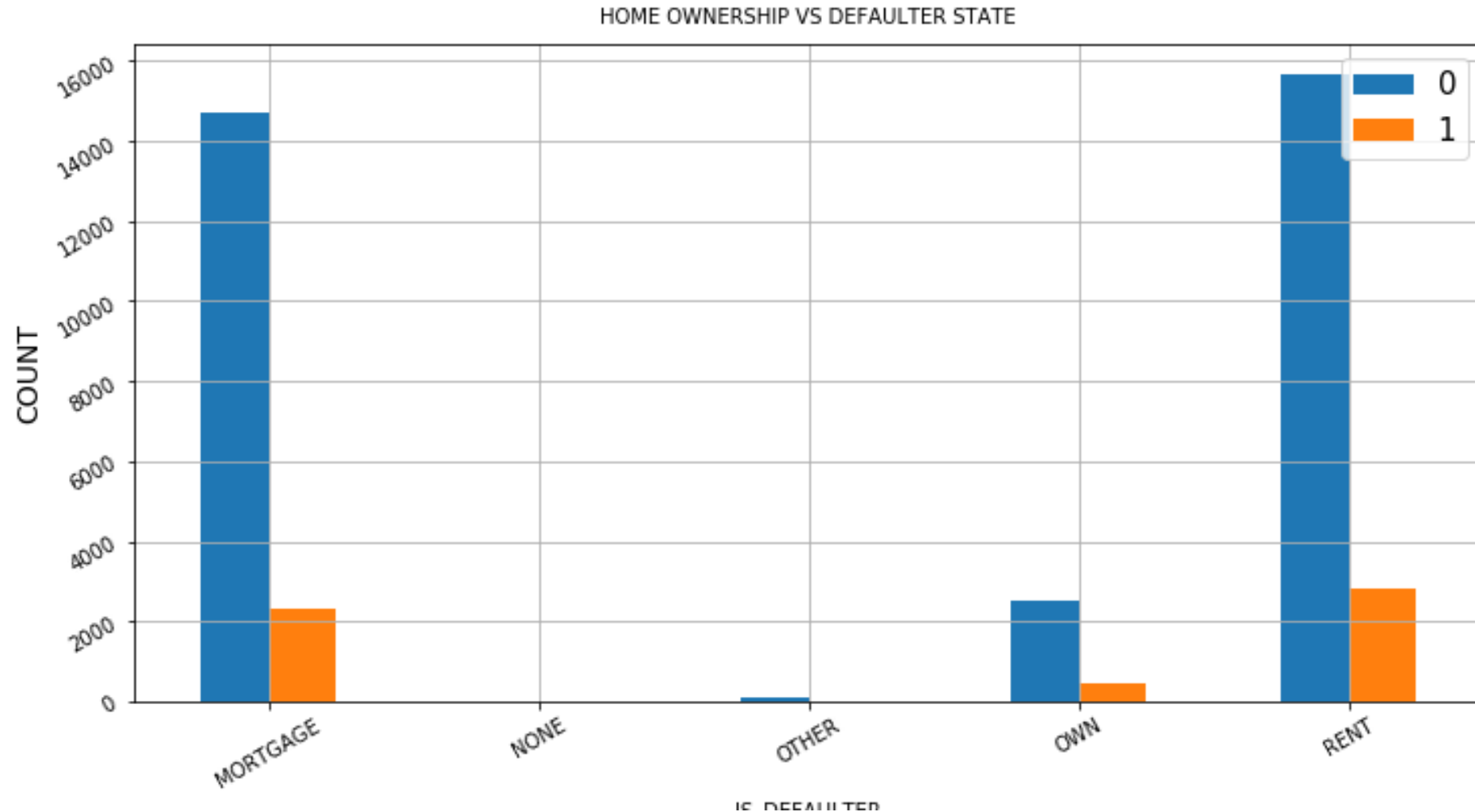
The high interest rate amount increases the chances of getting default.



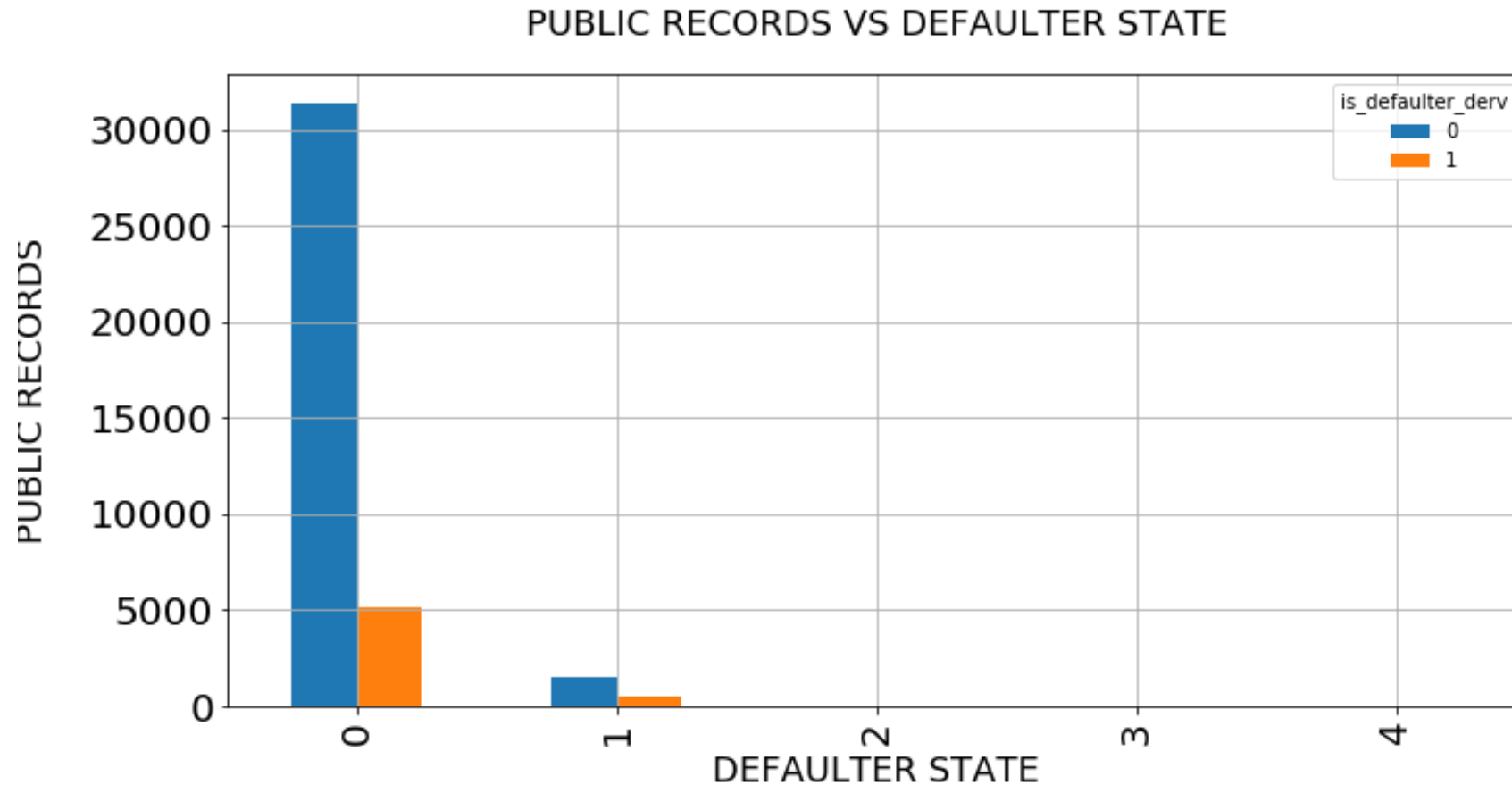
Purpose of the loan plays a significant role in determining the chances of getting default or not. As it can be observed that if the loan purpose is debt\_consolidation, home\_improvement, other, small\_business and credit\_card then there are chances of getting default is high unlike other loan purpose.



It seems the one who owns home has lesser chance of becoming defaulter and greater for those who are on rent and mortgage

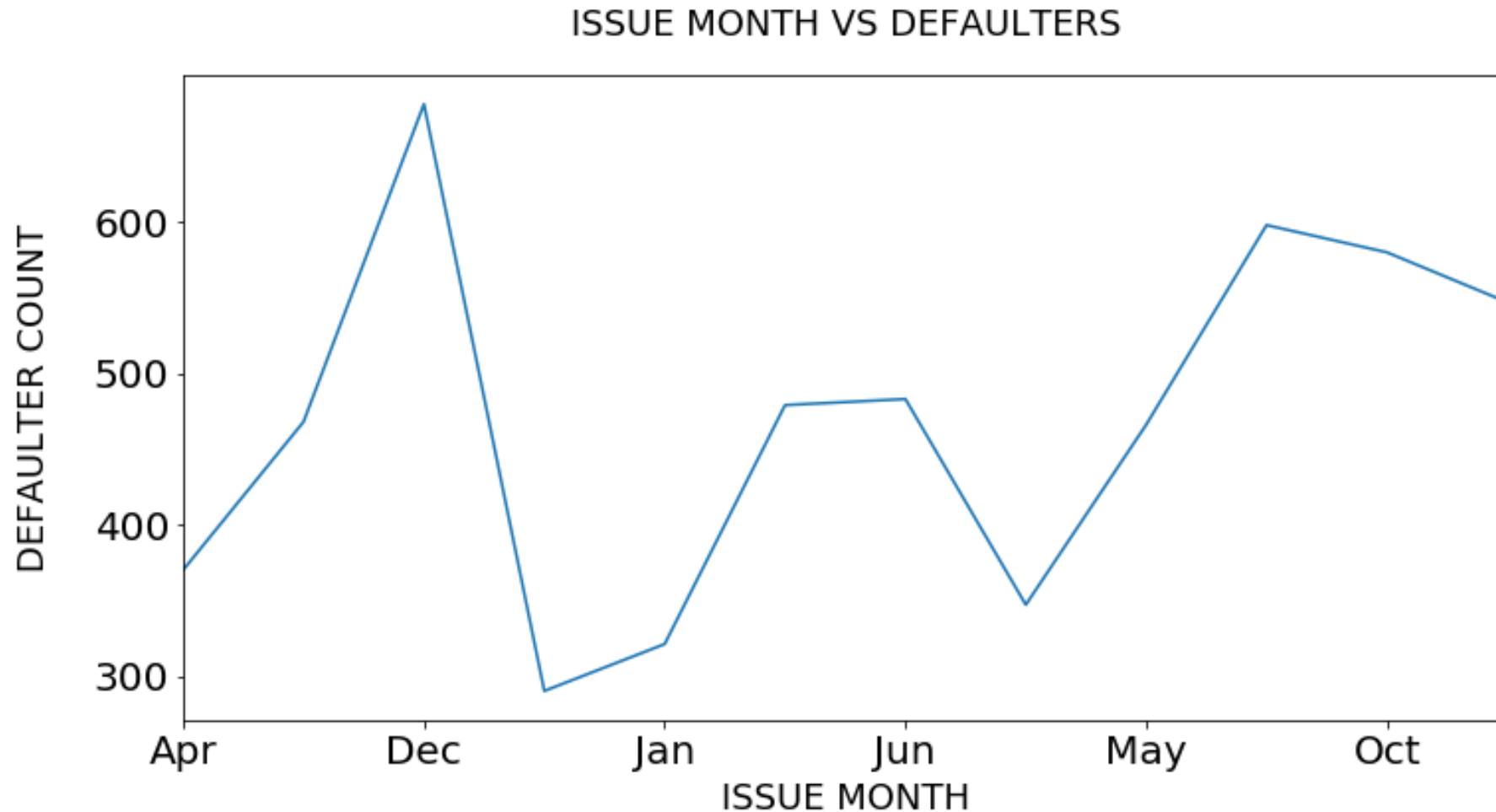


High number of Public records implies high risk, as the x-scale category number increases the Public records number decreases, which in turn reduces the chances of getting default.





The Month of December and November has got high number of defaulters unlike other months.



## Conclusions

- Exploratory analysis shows, loan issued month, no. of months since last delinquency, purpose of the loan, ownership status of a home, loan grade assigned to a consumer, no. of public records, annual income, interest amount on loan and job experience of the employee are the potential drivers in determination of whether loan is going to default or not.