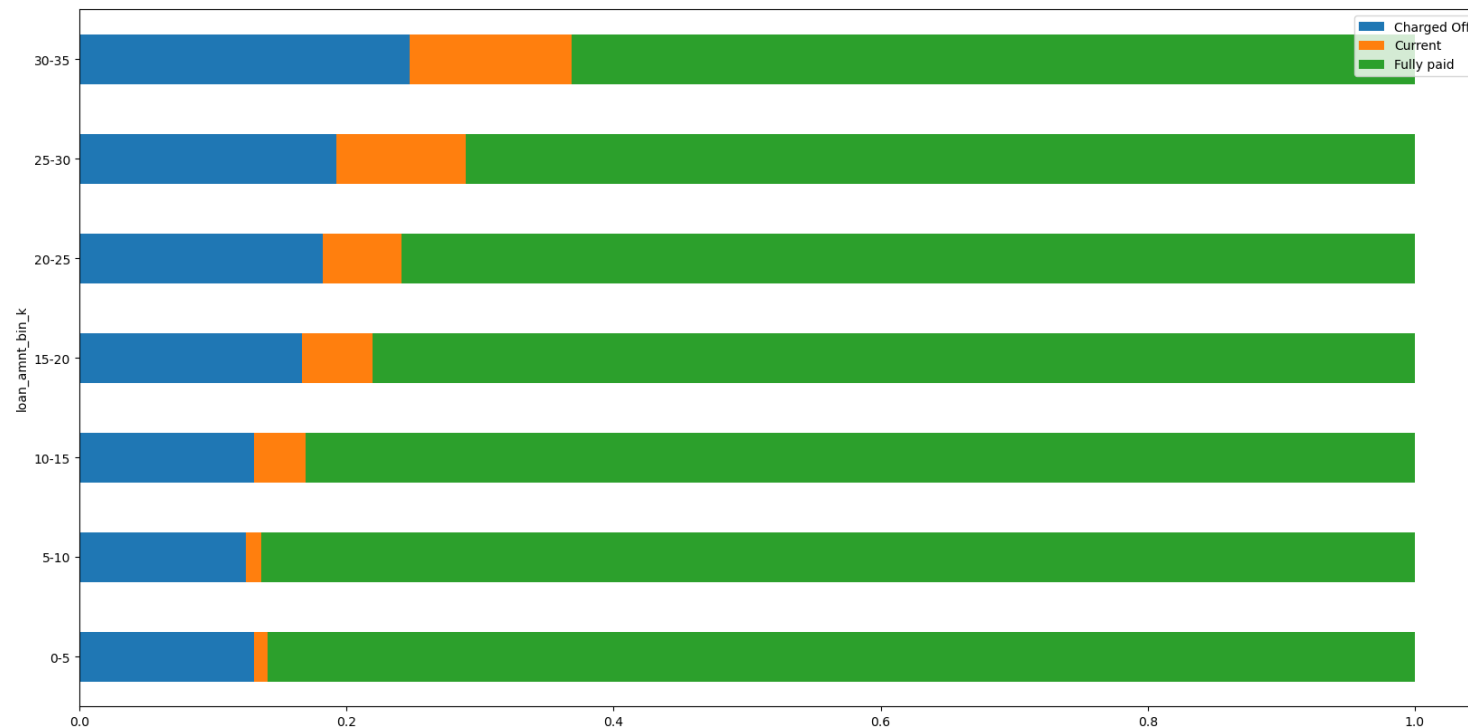


Lending Club Case Study

Univariate analysis: Segmented univariate analysis

Loan amount:

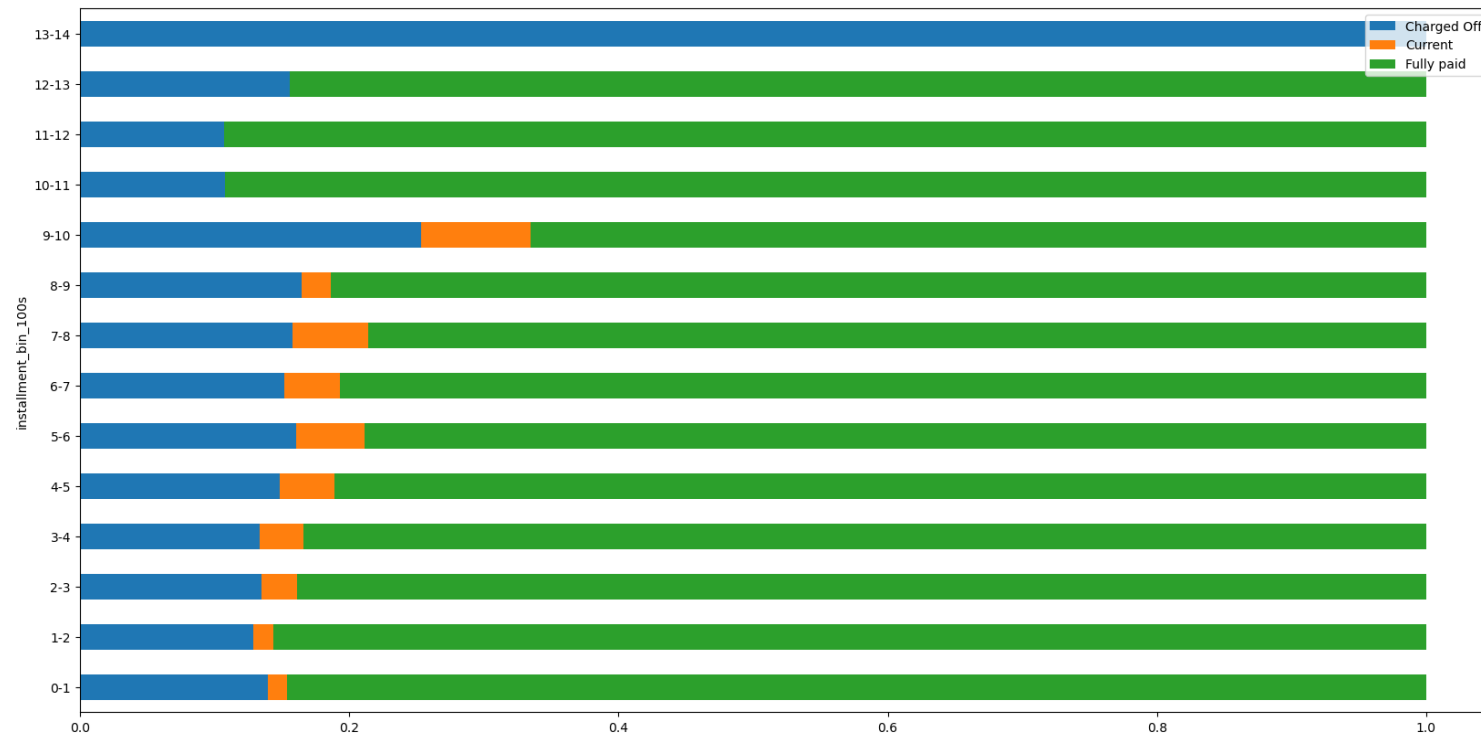
- Highest percentage of people defaulted are taking loan amount between 30,000 to 35000.



Univariate analysis: Segmented univariate analysis

Installment amount:

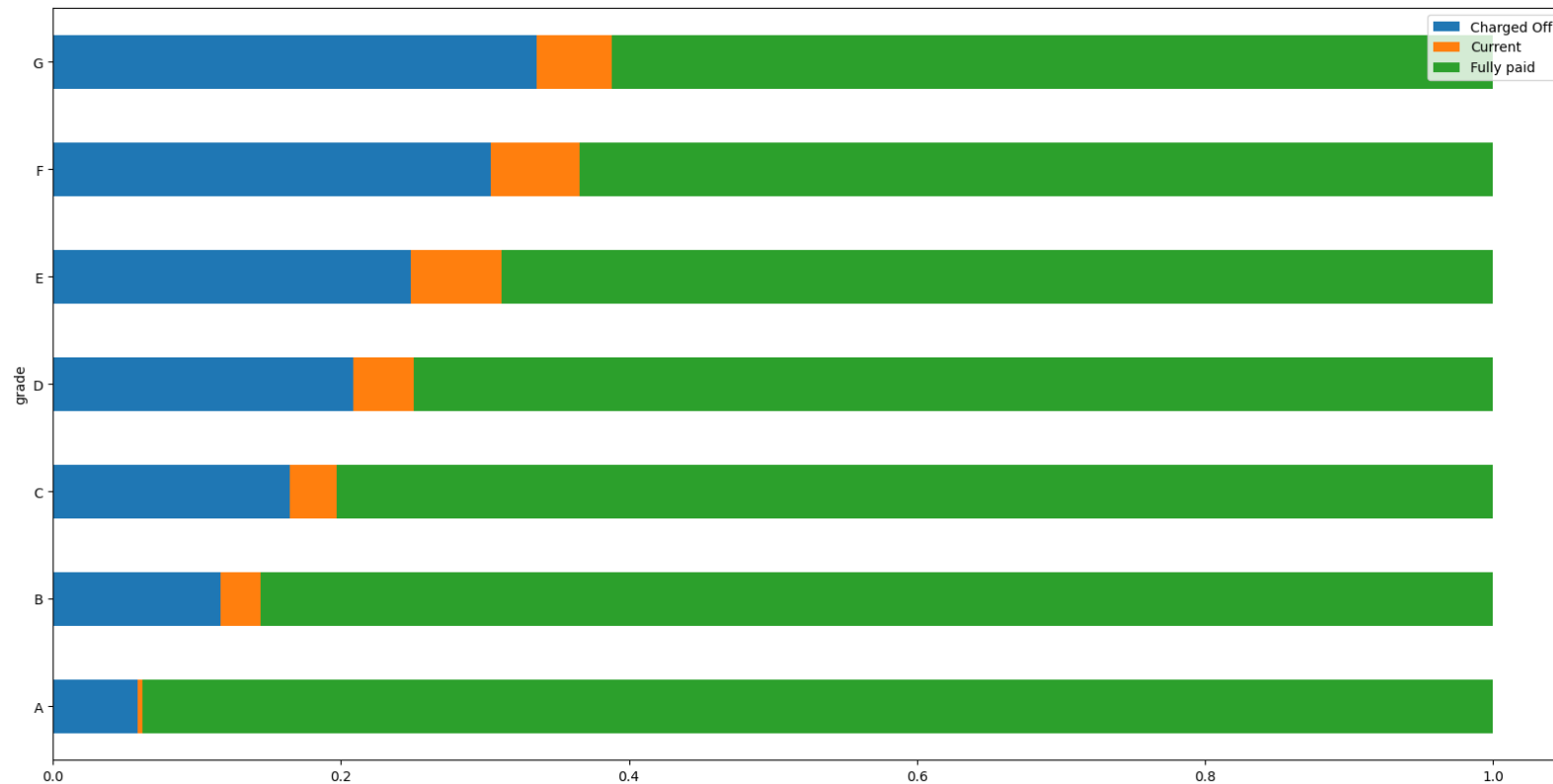
- Highest percentage of people likely defaulted are from installment paying between 1300 to 1400, the next most defaulted group is paying 900 to 1000 installment.



Univariate analysis: Segmented univariate analysis

Loan grade:

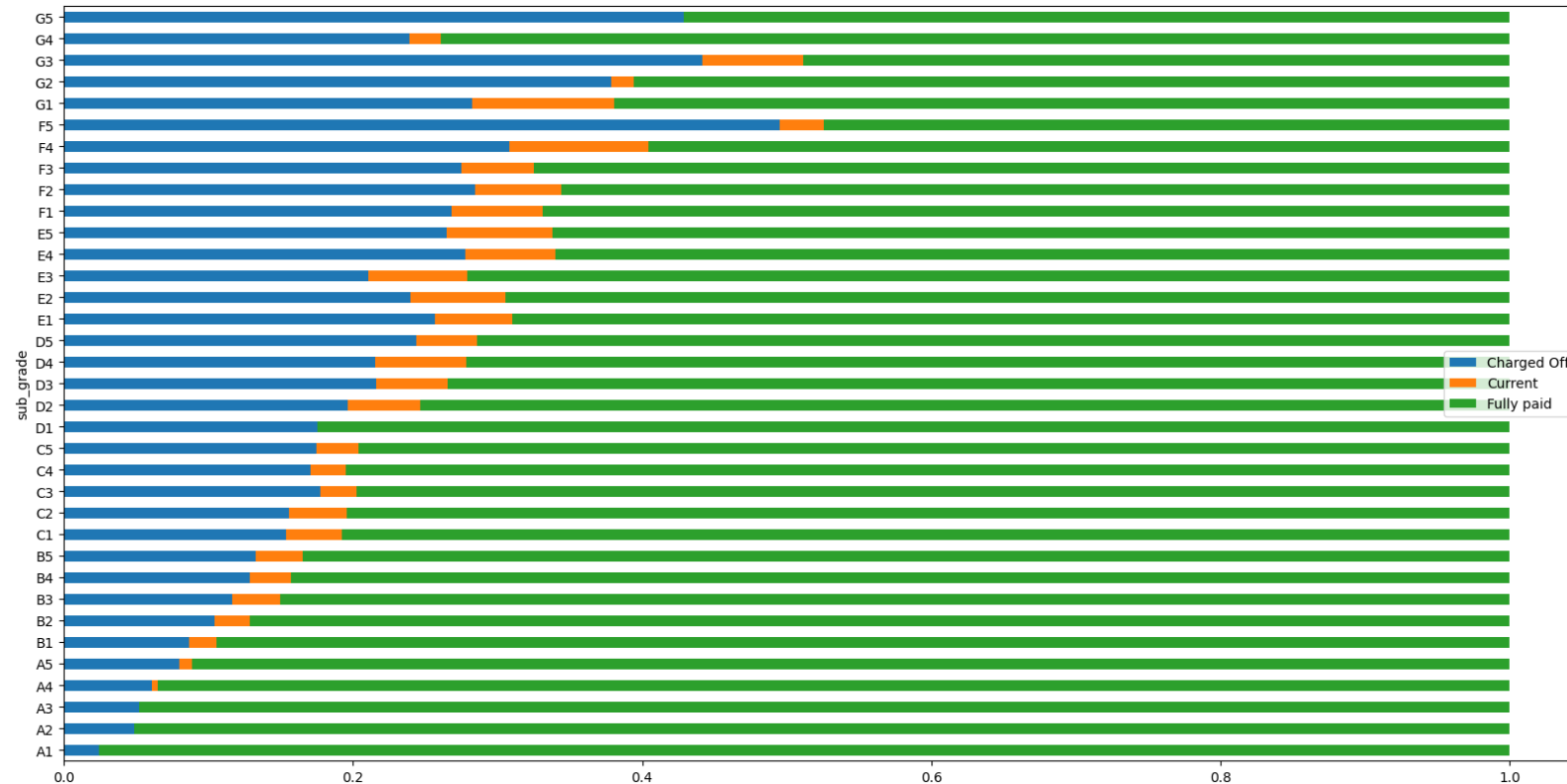
- Highest percentage of people are defaulted from grade G, and grade A people are most fully paid members.



Univariate analysis: Segmented univariate analysis

Sub grade:

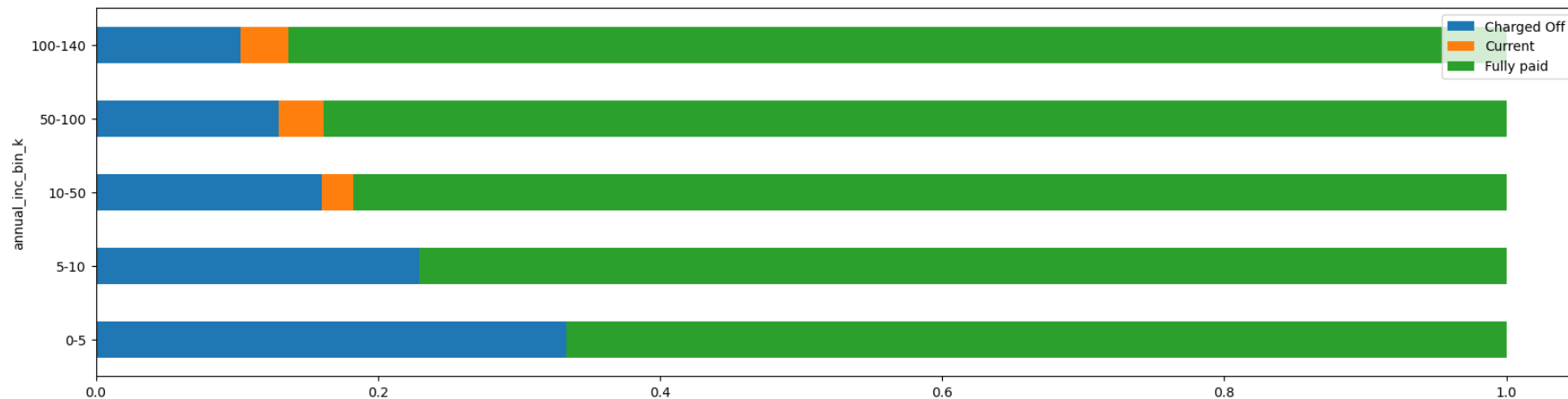
- Highest percentage of people defaulted from sub_grade members are from **F5** and fully paid members are from **A1**.



Univariate analysis: Segmented univariate analysis

Annual income:

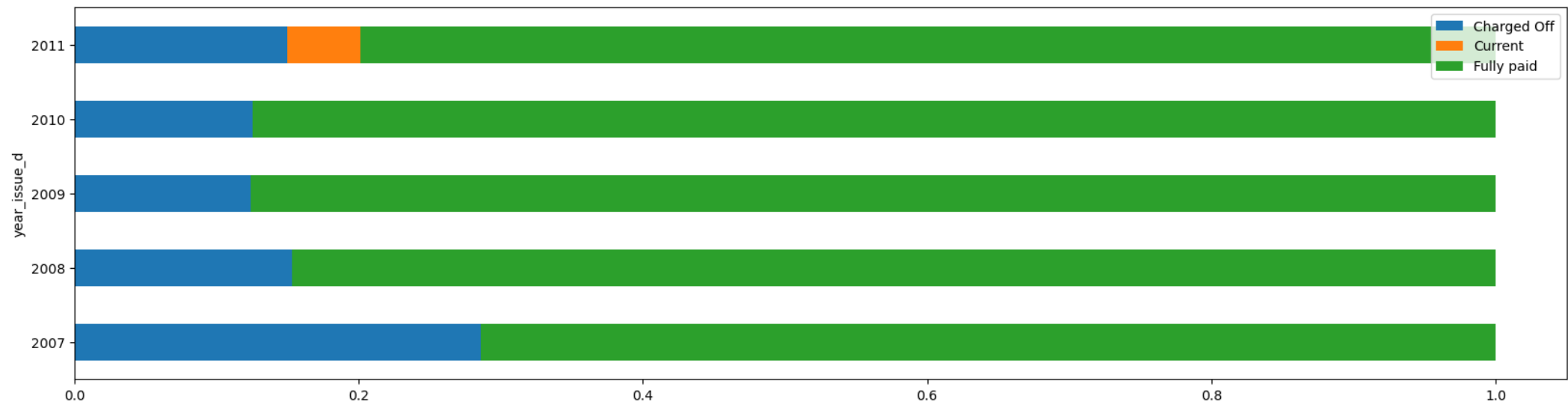
- Highest percentage of charged off members earning below 5000 annual income and highest percentage of member with annual income more then 100000 are fully paid.



Univariate analysis: Segmented univariate analysis

Loan issue year:

- Loans issued in 2007 have high percentage of charged off loans. And there is no active loans which issued before 2011



Univariate analysis: Segmented univariate analysis

Loan issue month:

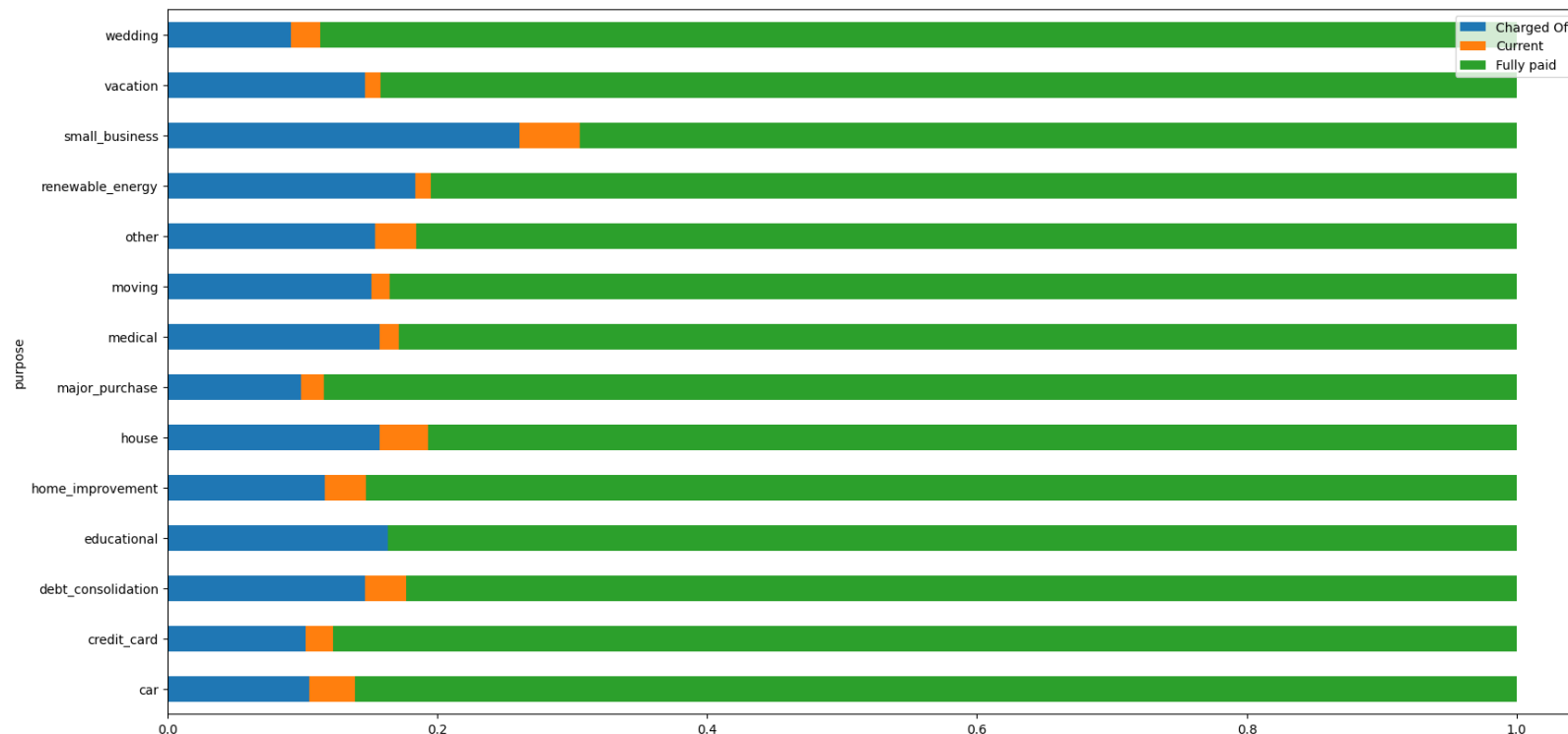
- As per observation from month of loan issued there is not much difference in percentage of defaulters. But loans issued in January, February, March, April months are not currently active.



Univariate analysis: Segmented univariate analysis

Purpose of loan:

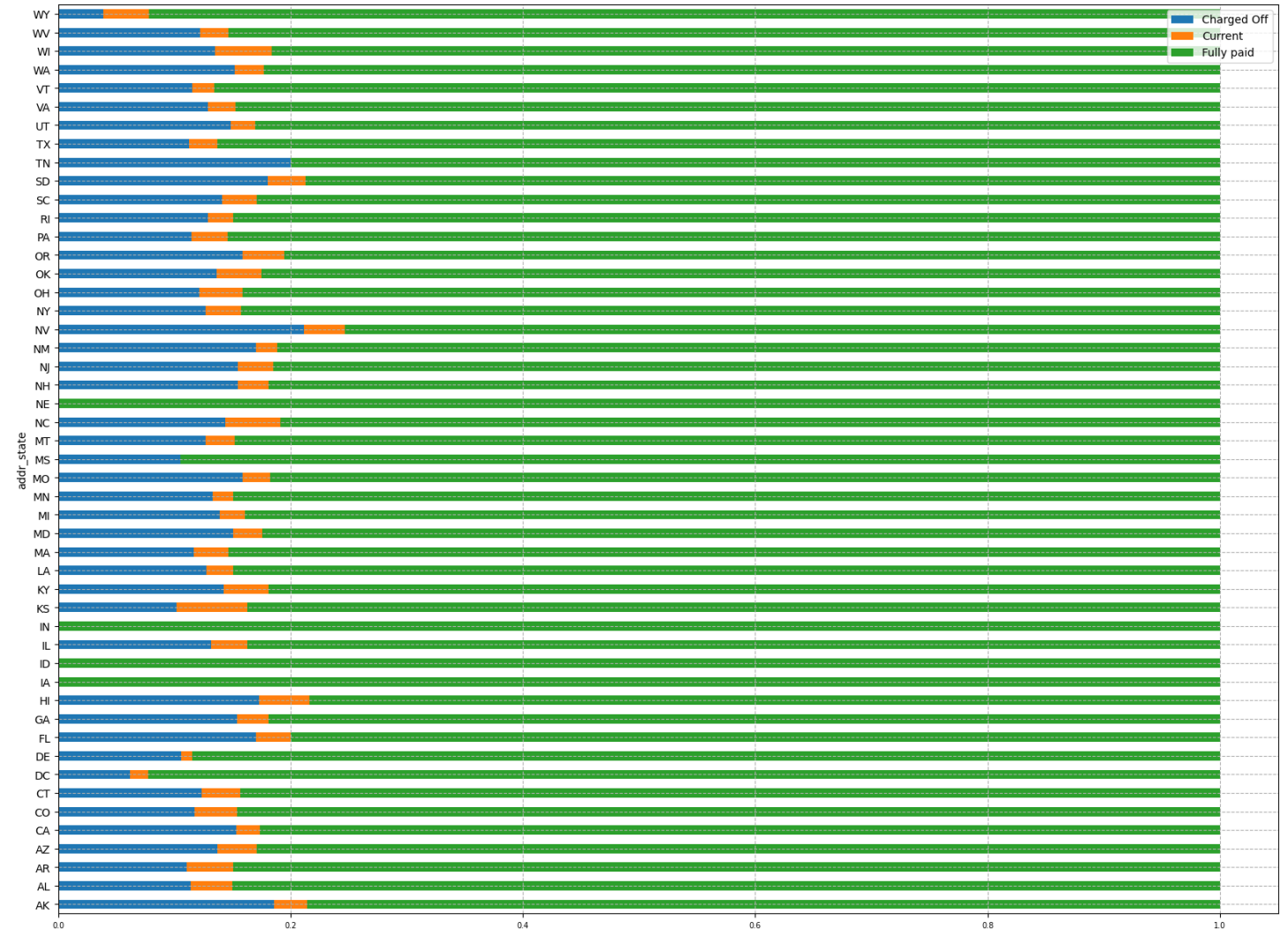
- Loan taken for small business members have high percentage of Charged off category, this might be due to loss in business. And highest percentage of fully paid people are taken loan for wedding expenses, this might be due to they are well planned for wedding.



Univariate analysis: Segmented univariate analysis

State address:

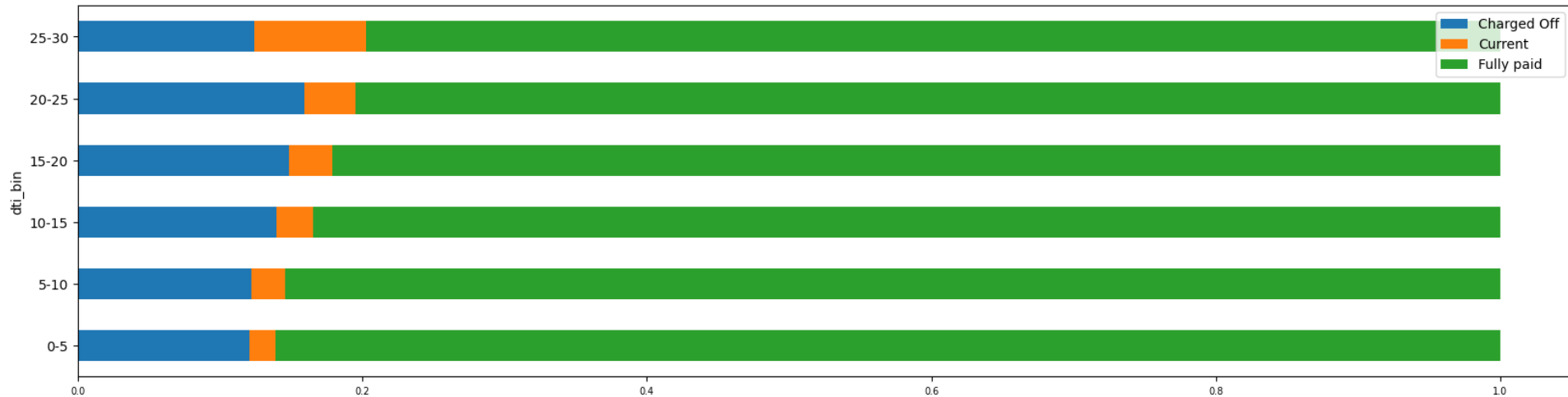
- Highest percentage of people defaulted from state **NV**. And highest percentage of members are from NE,IN,ID,IA state are fully paid.



Univariate analysis: Segmented univariate analysis

Debit-To-Income ratio:

- DTI range 0-5 group members have highest percentage of fully paid members, And DTI range 20-25 group have Highest percentage of Charged off members.



Univariate analysis: Segmented univariate analysis

Revolving line utilization rate:

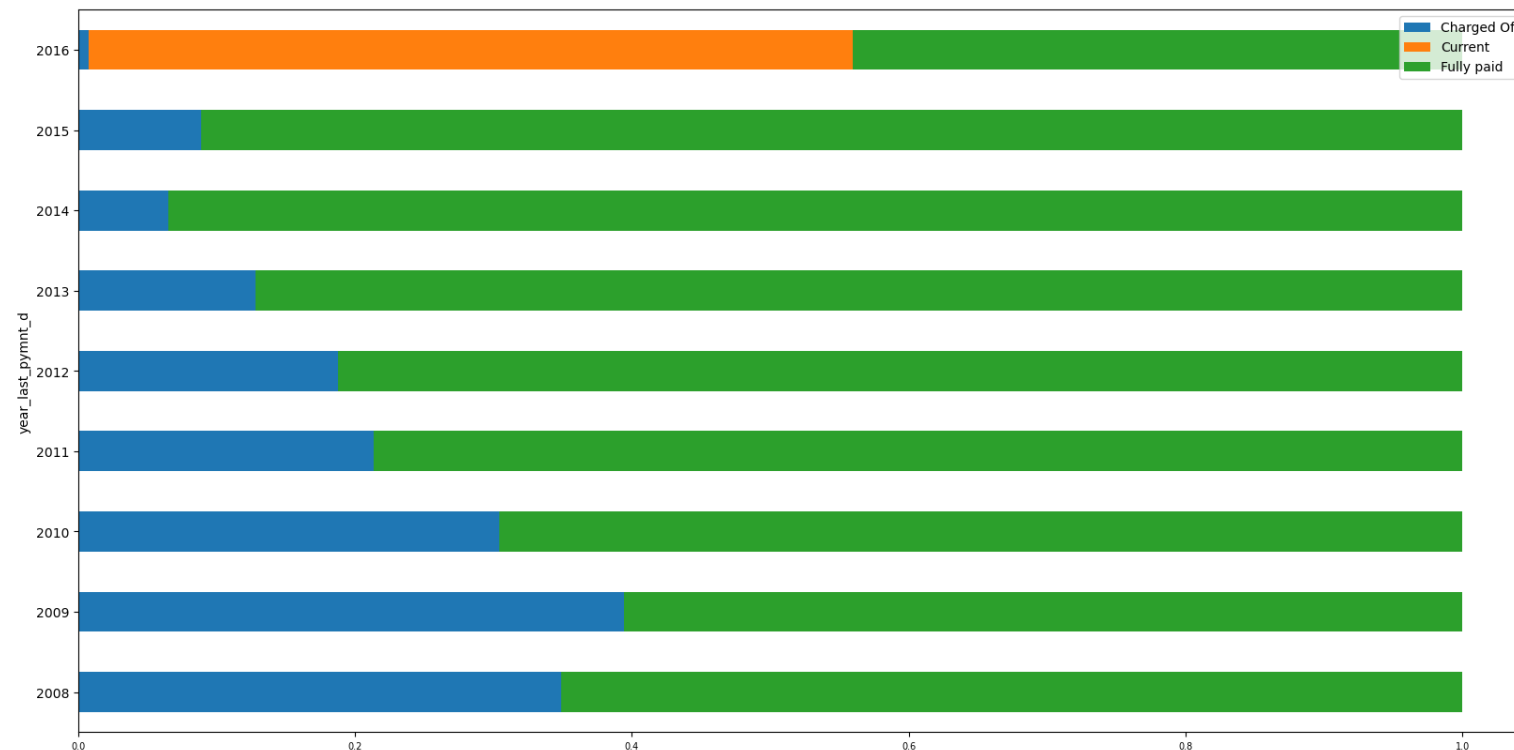
- Highest percentage of defaulters having revolving line utilization rate 90% to 100%,
- Highest percentage of fully paid members having revolving line utilization rate rate 0 to 10%.



Univariate analysis: Segmented univariate analysis

Last payment year:

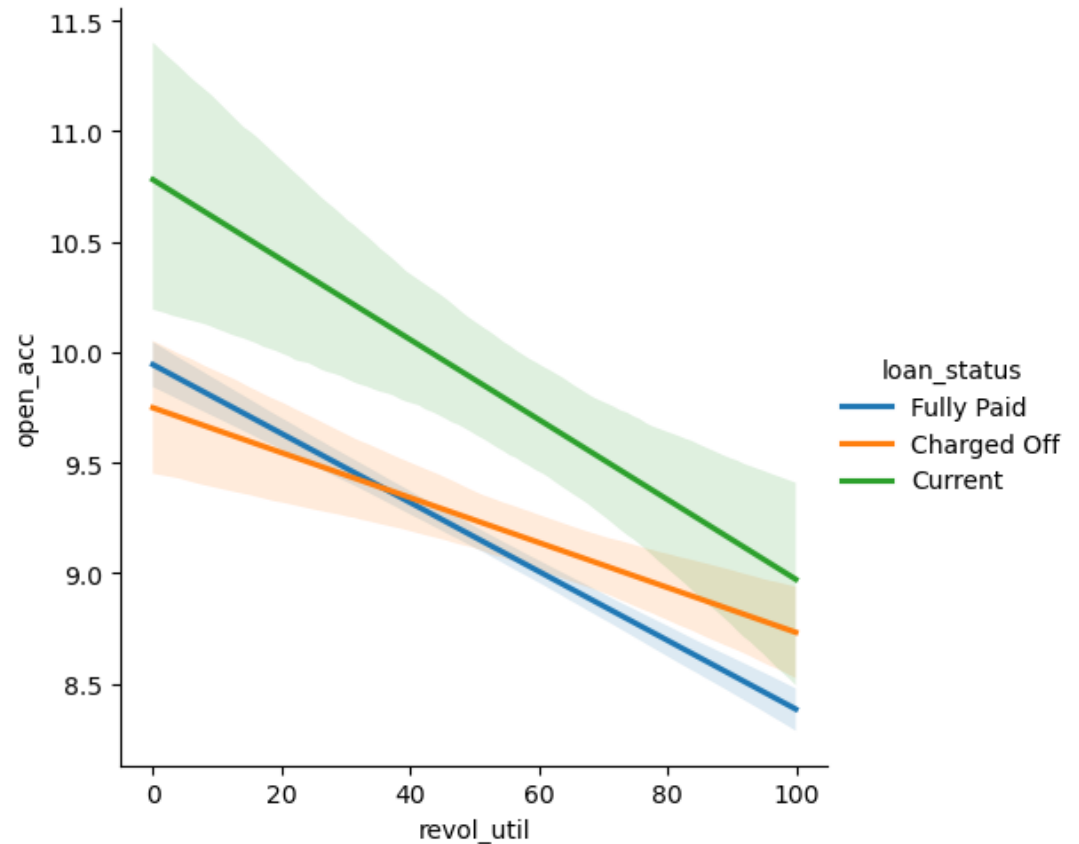
- The members last payment date was in 2009 are most defaulters. And members whose last payment year was in 2016 are having least percentage of defaulters.



Bivariate analysis: continuous variables:

Revolving line utilization rate and open credit lines:

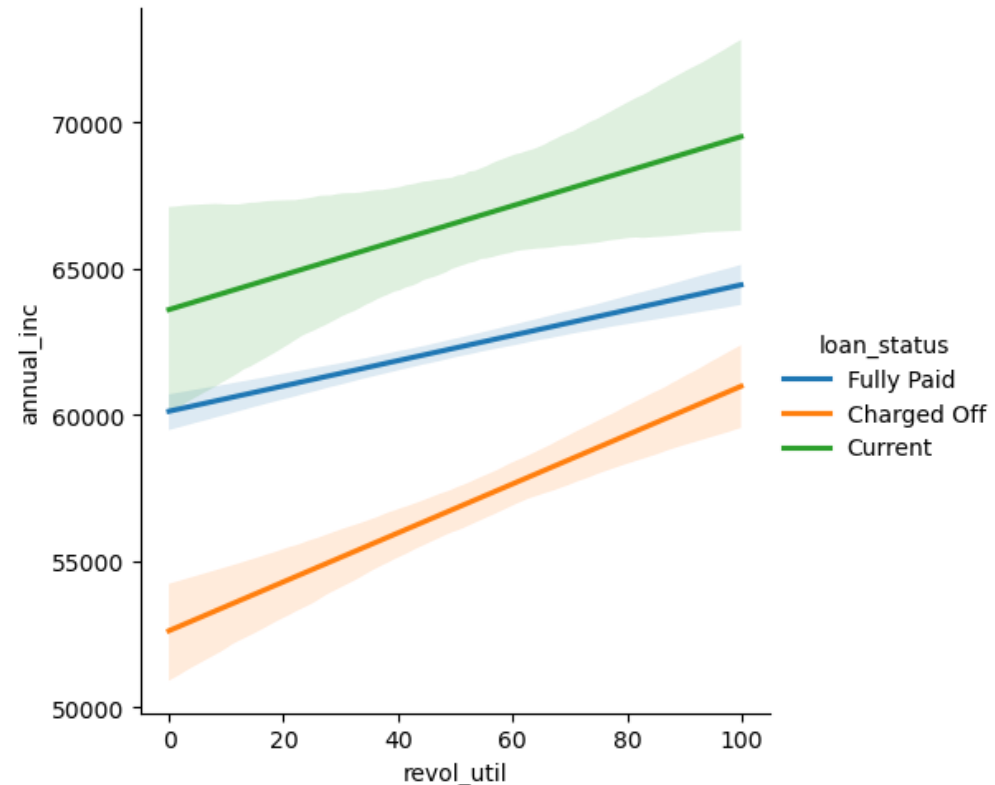
- The correlation is negative between revolving line utilization rate and open credit lines is **high** for **fully paid** members and **low** for **charged off** members.
- Correlation on Charged off : -0.063
- Correlation on Fully paid : -0.102



Bivariate analysis: continuous variables:

Revolving line utilization rate and annual income:

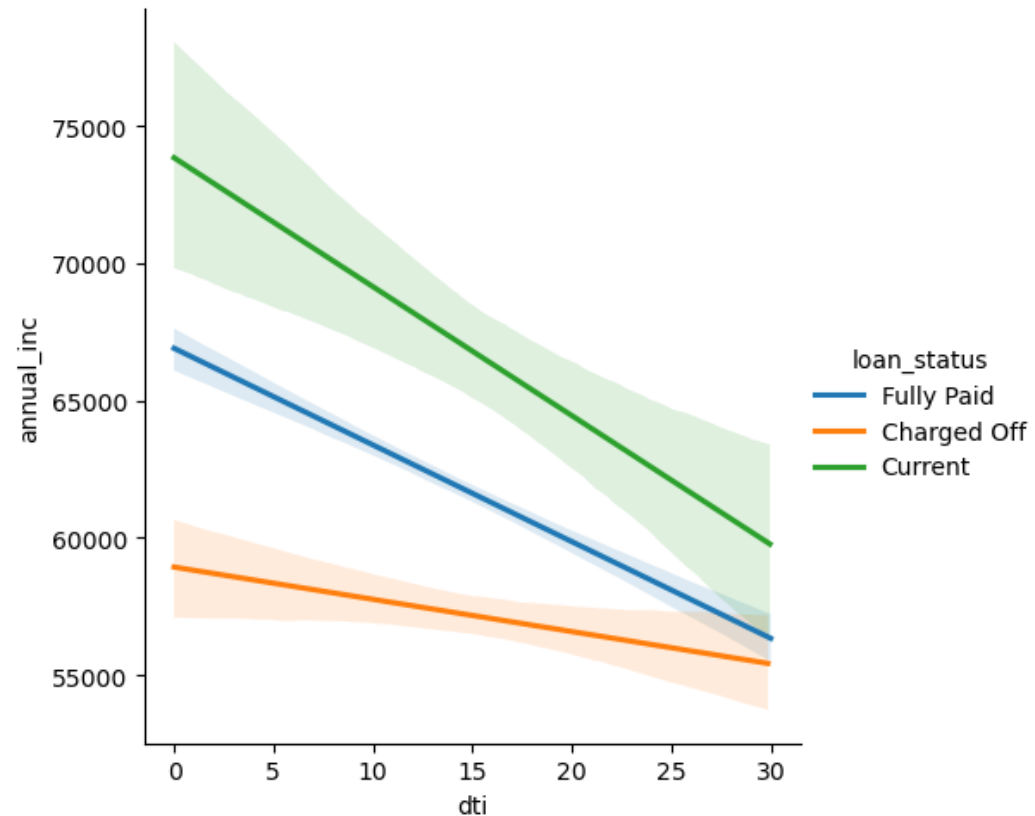
- The correlation is positive between revolving line utilization rate and annual income, is **high** for **charged off** members and **low** for **fully paid** members.
- Correlation on Charged off : 0.088
- Correlation on Fully paid : 0.044



Bivariate analysis: continuous variables:

Annual income and Debit-To-Income ratio:

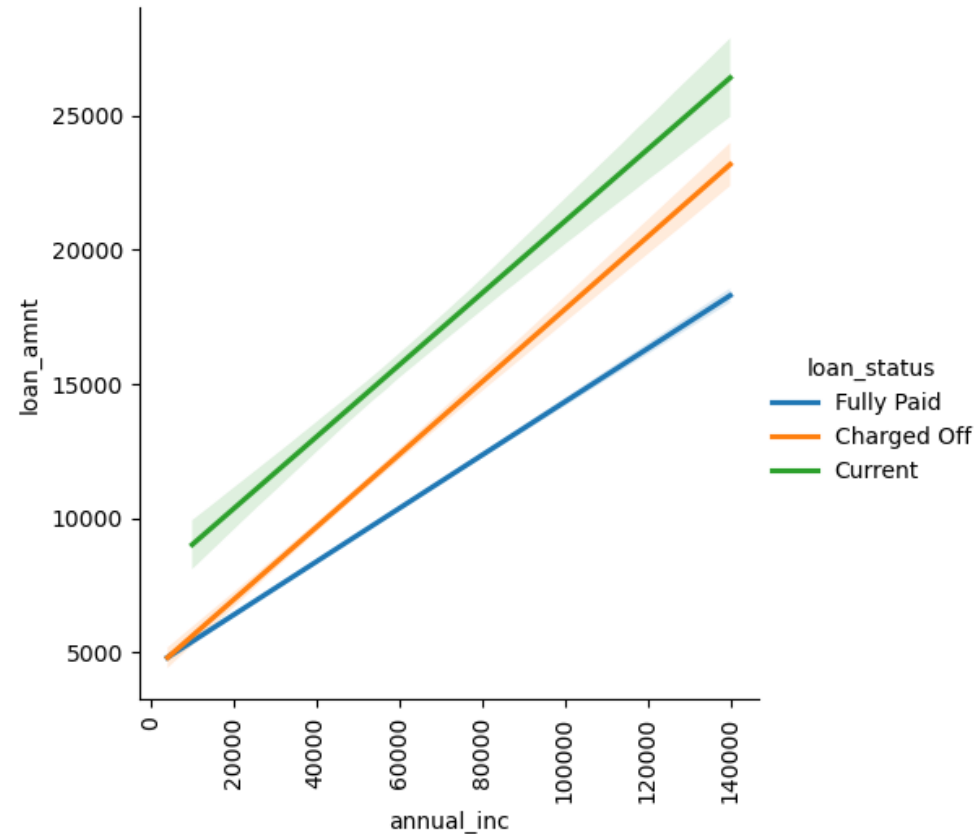
- The correlation is negative between annual income and Debit-To-Income ratio, is **high** for **fully paid** members and **low** for **charged off** members.



Bivariate analysis: continuous variables:

Annual income and Loan amount:

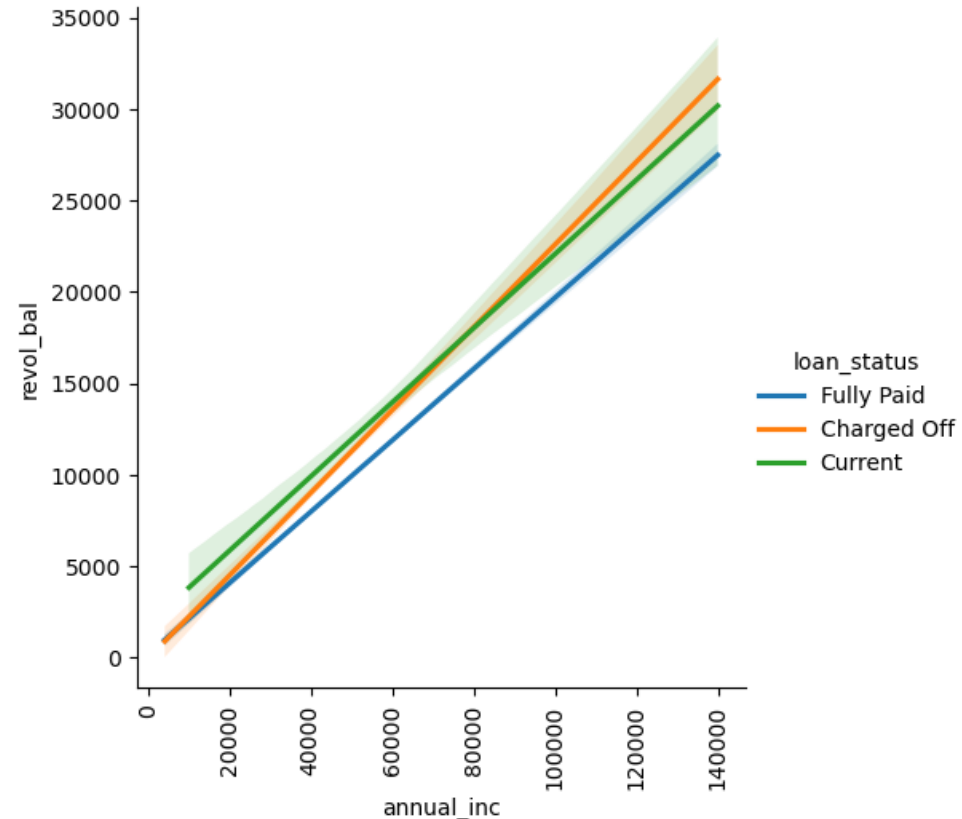
- The correlation is positive between annual income and loan amount, is **high** for **charged off** members and **low** for **fully paid** members.
- Correlation on Charged off : 0.45
- Correlation on Fully paid : 0.39



Bivariate analysis: continuous variables:

Annual income and Revolving line utilization rate:

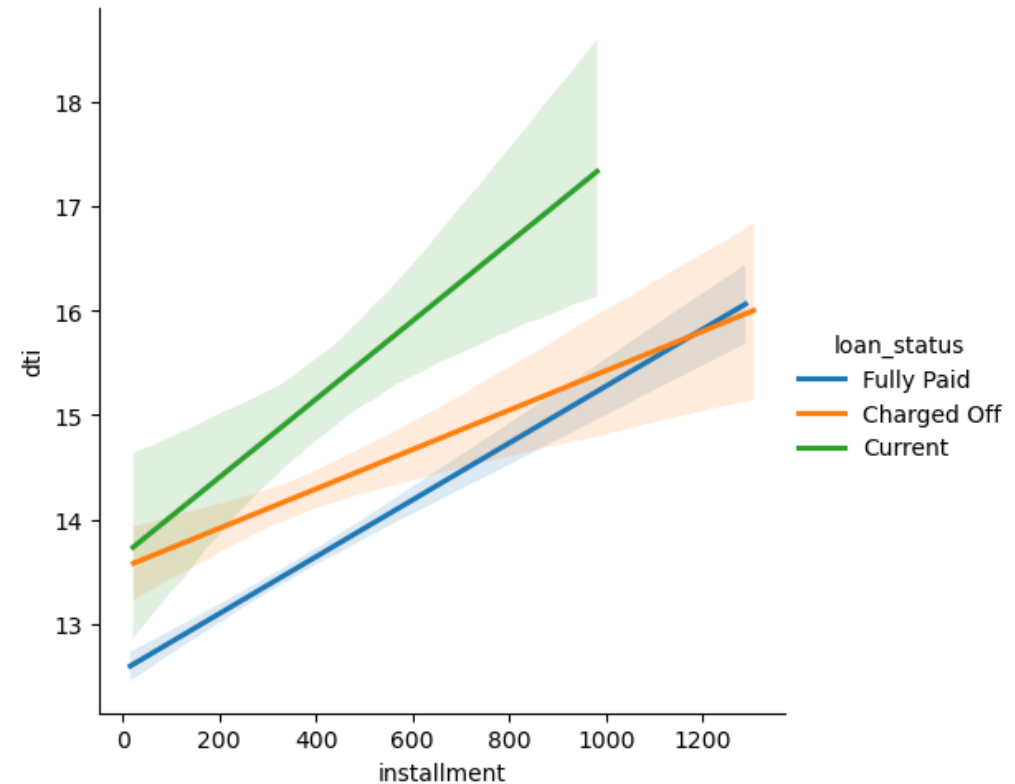
- The correlation is positive between annual income and revolving line utilization rate, is **high** for **charged off** members and **low** for **fully paid** members.
- Correlation on Charged off : 0.435
- Correlation on Fully paid : 0.393



Bivariate analysis: continuous variables:

Installment paid and Debit-To-Income ratio:

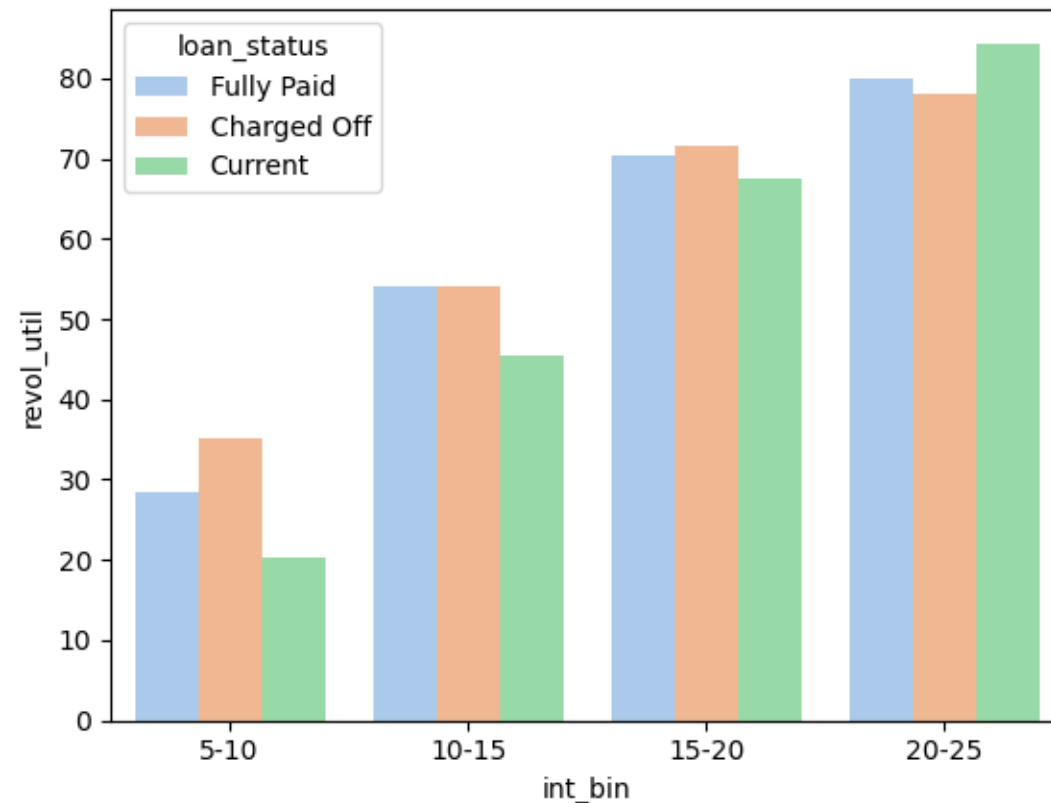
- The correlation is positive between installment paid and Debit-To-Income ratio, is **high** for **fully paid** members and **low** for **charged off** members.
- Correlation on Charged off : 0.06
- Correlation on Fully paid : 0.08



Bivariate analysis: continuous categorical

Interest and revolving utilization rate:

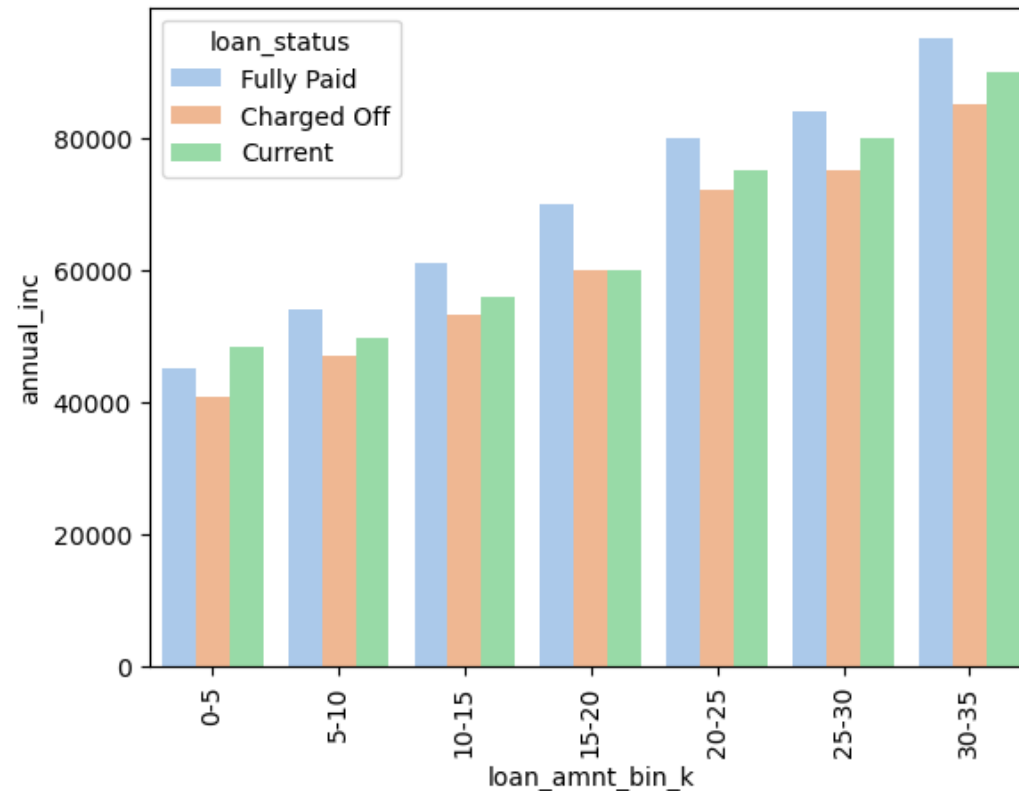
- Members with high Revolving line utilization rate and high interest rate are charged off and with low Revolving line utilization and low interest rate are fully paid.



Bivariate analysis: continuous categorical

Annual income and loan amount:

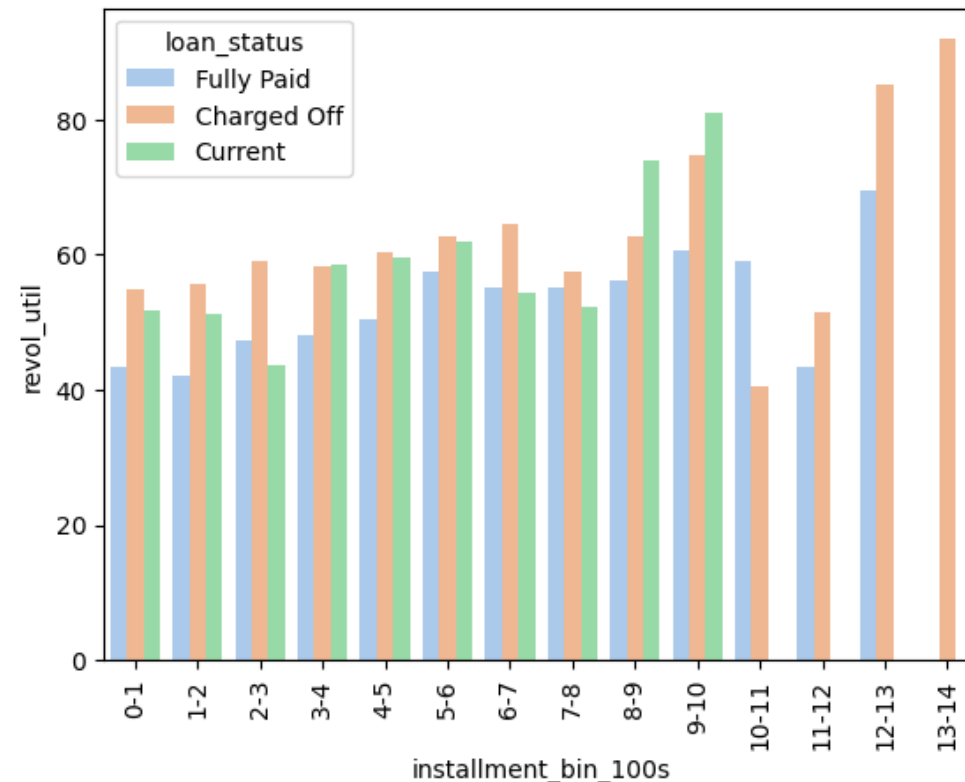
- The median value of annual income with respect to loan amount bins shows that Members with high annual income have opted for higher loan amount and charged off.



Bivariate analysis: continuous categorical

Installment and revolving utilization rate:

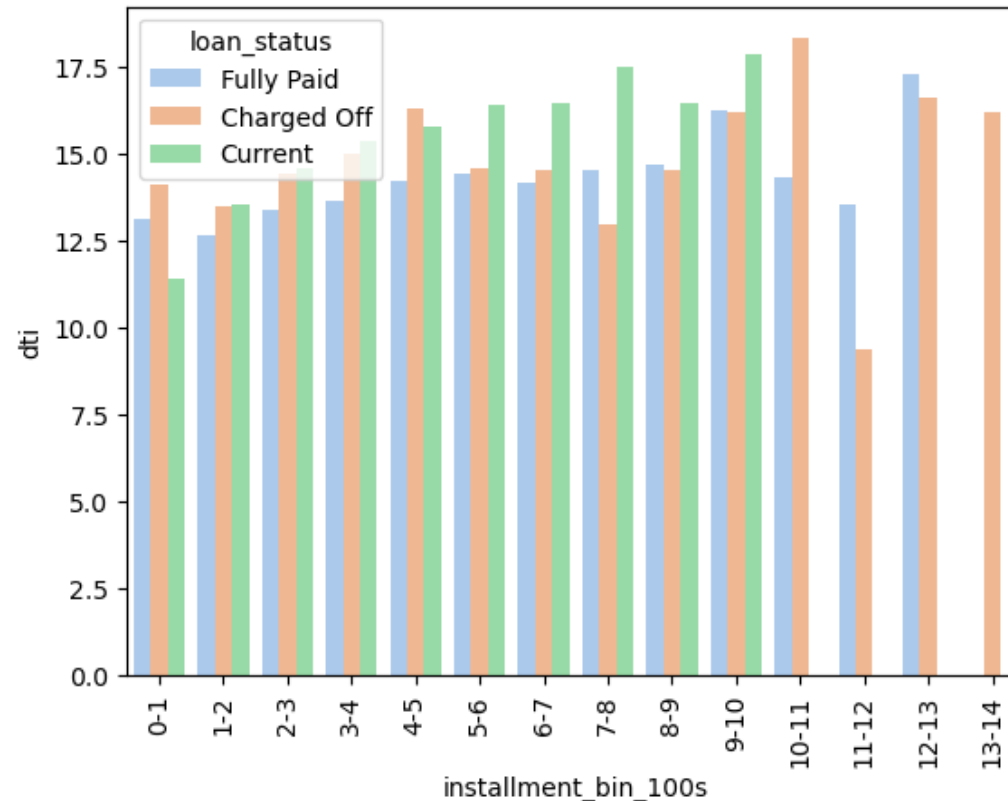
- The median value of revolving line utilization rate with respect to installment bins shows that members with higher revolving utilization rate and paying higher installment are charged off.



Bivariate analysis: continuous categorical

Debit-To-Income ratio and installment:

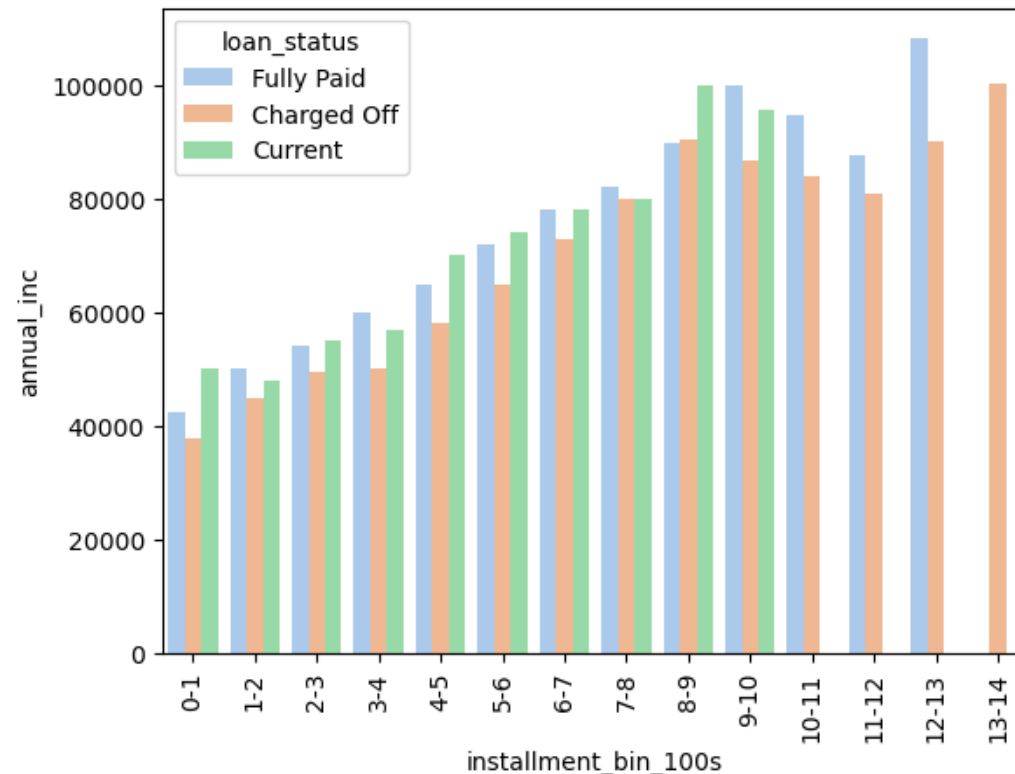
- The median value of Debit-To-Income ratio with respect to installment bins shows that, Members with higher DTI and paying higher installments are most charged off



Bivariate analysis: continuous categorical

Installment and annual income:

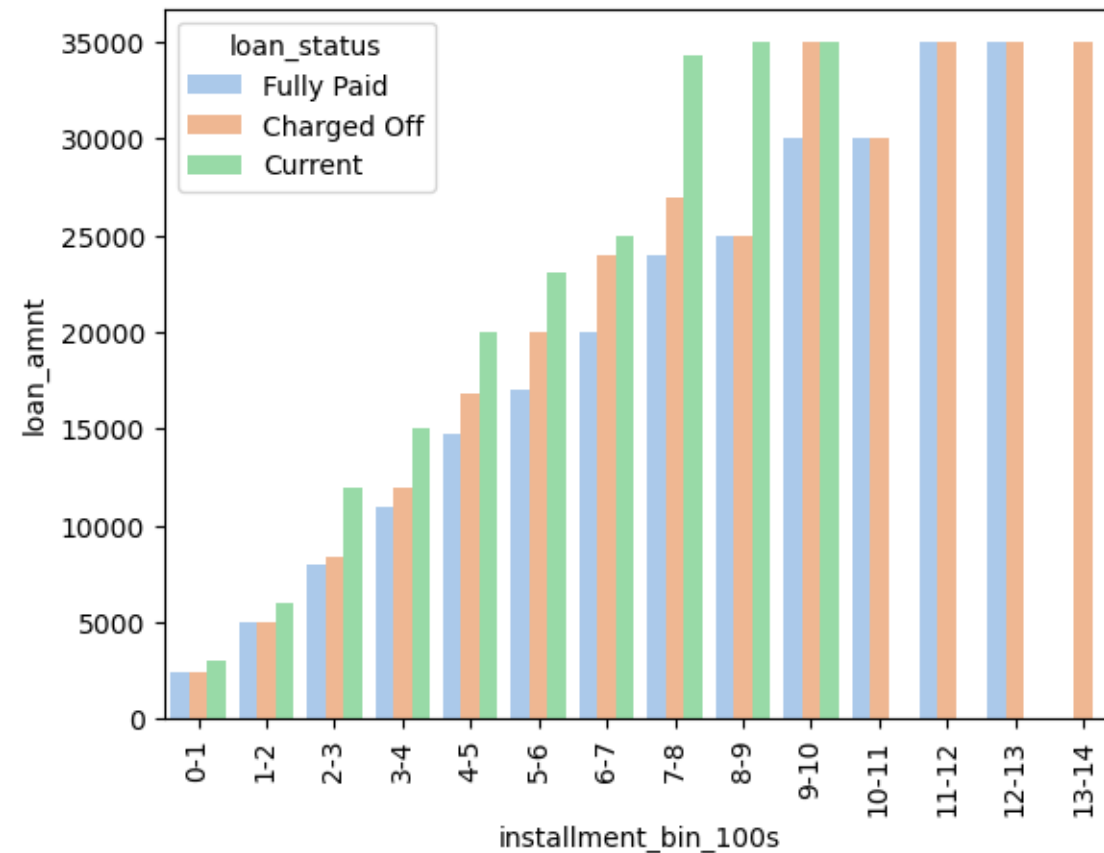
- The mean annual income is high for members who are paying installments between 1000 to 1100 and they are most Charged off. From installment bin 1300 to 1400 all members are defaulters with more annual income.



Bivariate analysis: continuous categorical

Installment and loan amount:

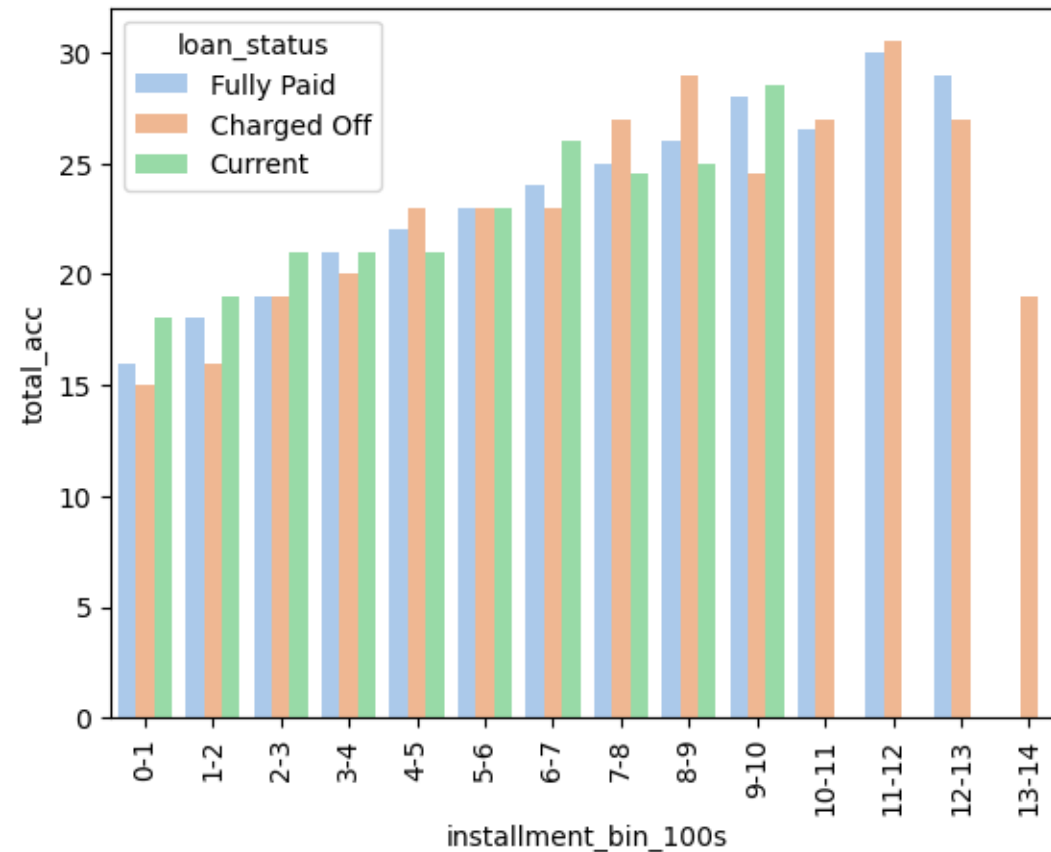
- Members opted for high loan amount and paying high installments are most defaulters/Charged off.



Bivariate analysis: continuous categorical

Number of credit lines and installment:

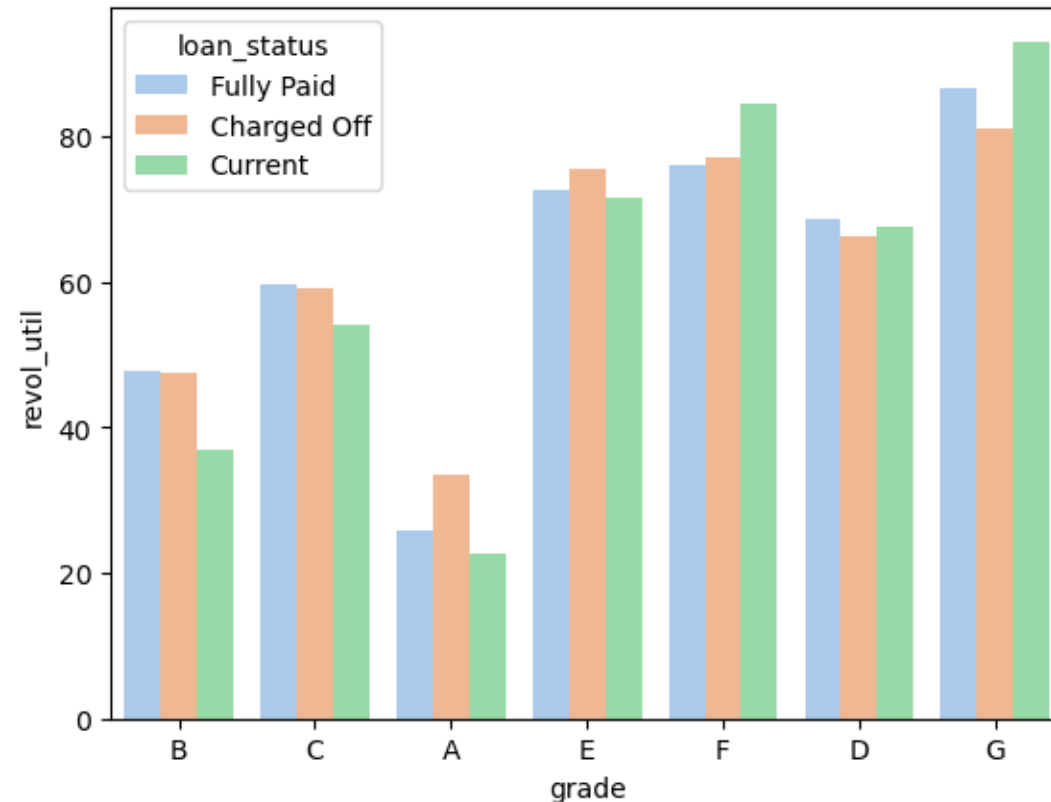
- Members with more number of credit lines(`total_acc`) and paying higher installment amount are most likely Charged off



Bivariate analysis: continuous categorical

Loan grade and revolving line utilization rate:

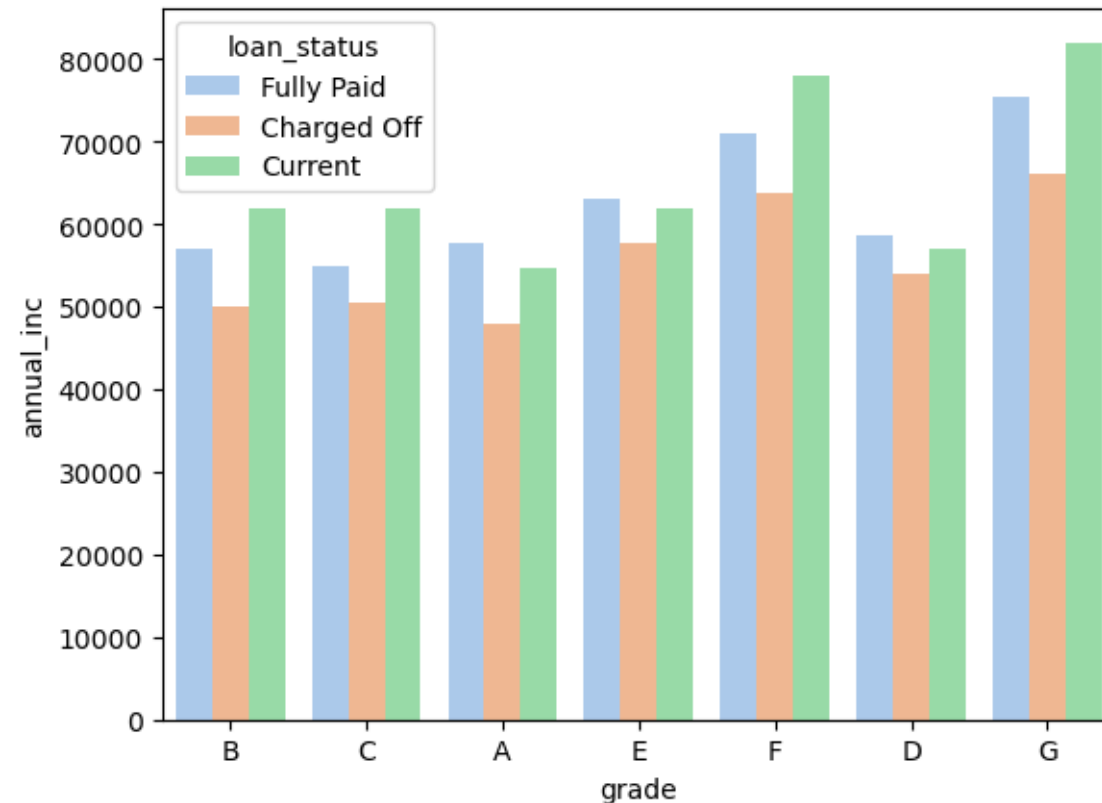
- Members belongs to Grade **A** are having lower revolving utilization rate are less Charged off and Grade **G** members are most Charged off with Higher rate.



Bivariate analysis: continuous categorical

Loan grade and annual income:

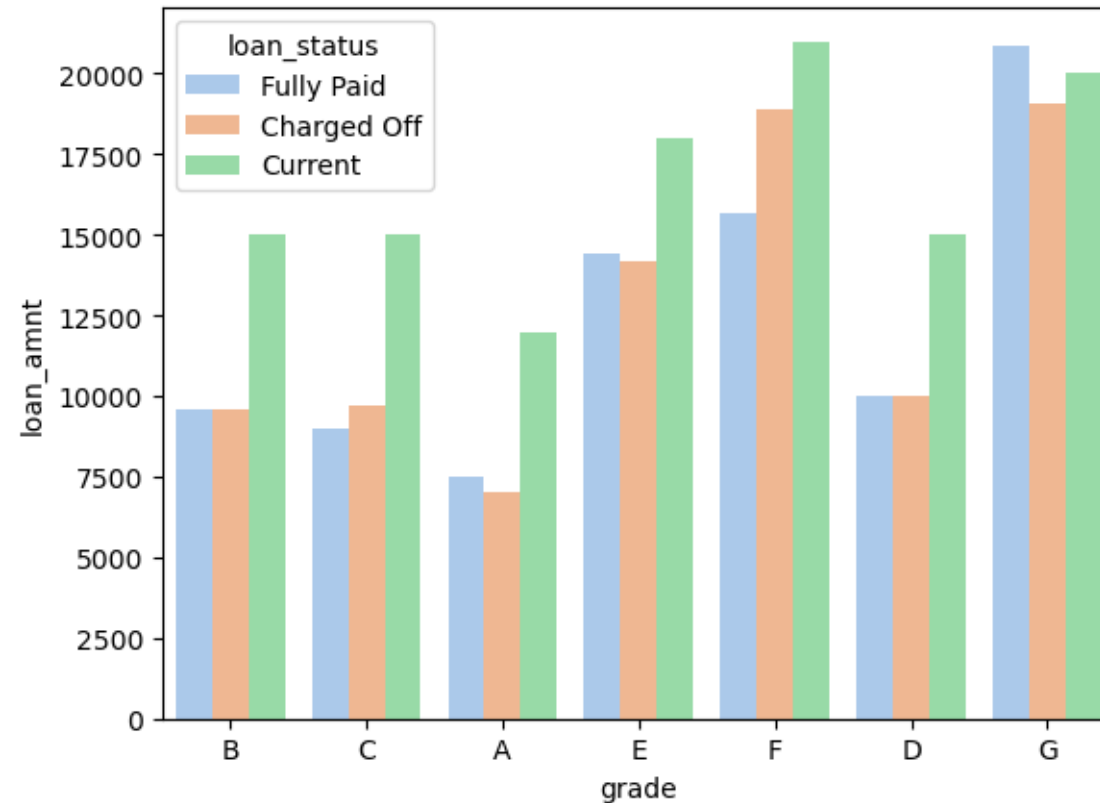
- members are most charged off people and they are highly paid employees.



Bivariate analysis: continuous categorical

Loan grade and Loan amount:

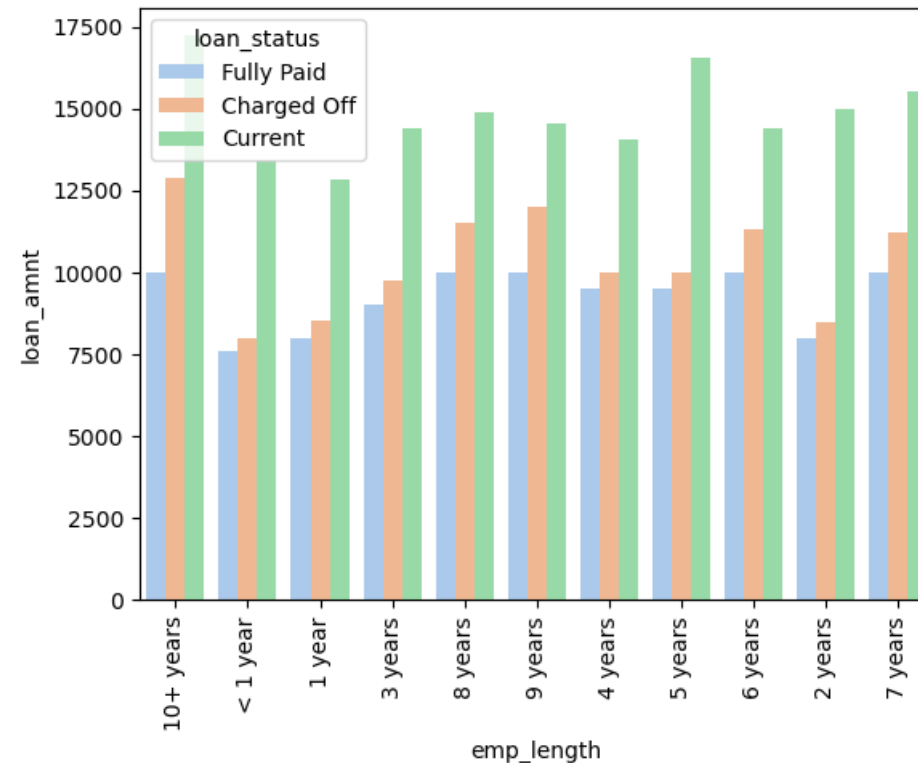
- Grade G members are opted for higher loans and they are most charged off.



Bivariate analysis: continuous categorical

Employment length and loan amount:

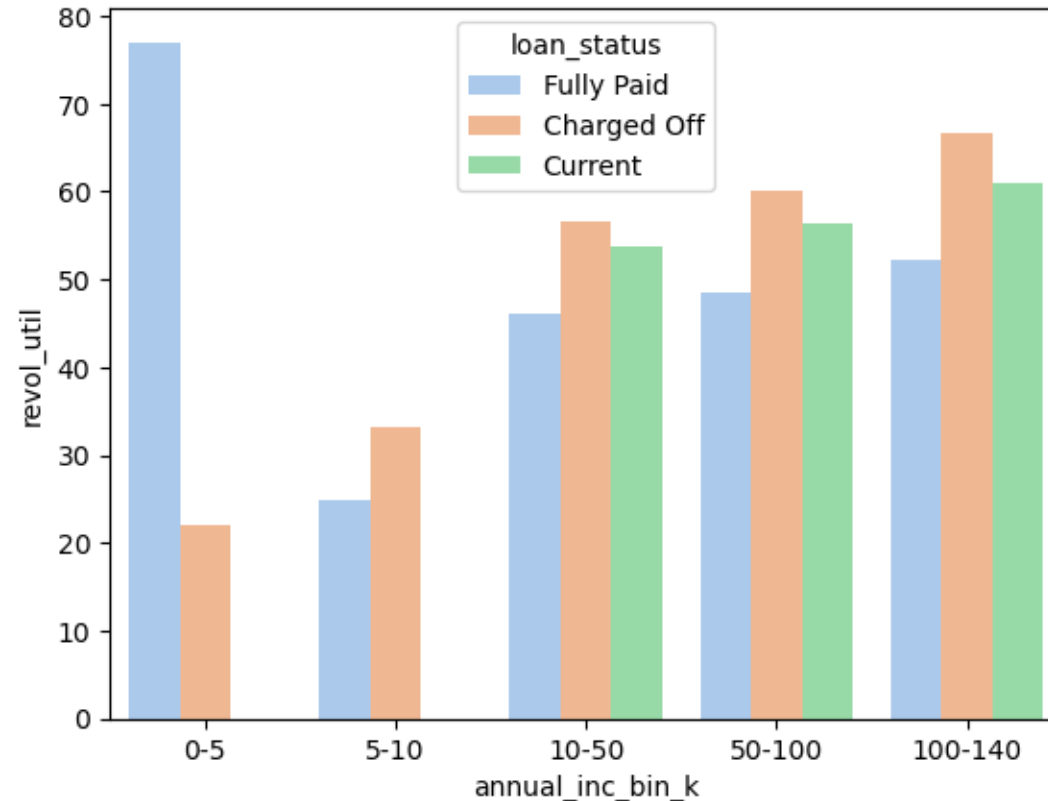
- Members with low employment length are opted for lower loan amount and they are least Charged off. And members with high employment length are opted for higher loan amount and they are defaulters.



Bivariate analysis: continuous categorical

Annual income and revolving line utilization rate:

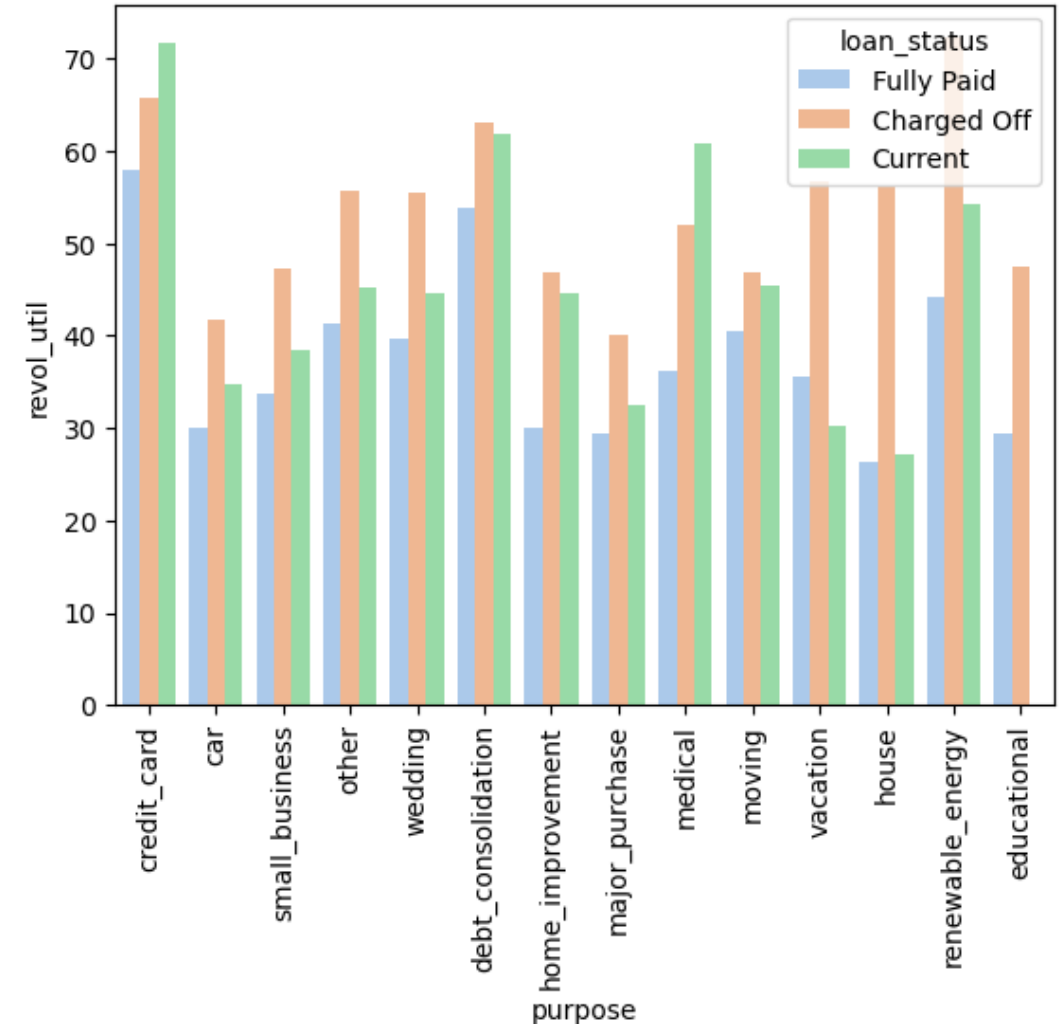
- Members with low annual income below 5000 have the high revolving utilization rate and they are fully paid. Members with revolving utilization rate between 40-50 have more than 10000 income but they are most Charged off.



Bivariate analysis: continuous categorical

Purpose and revolving line utilization rate:

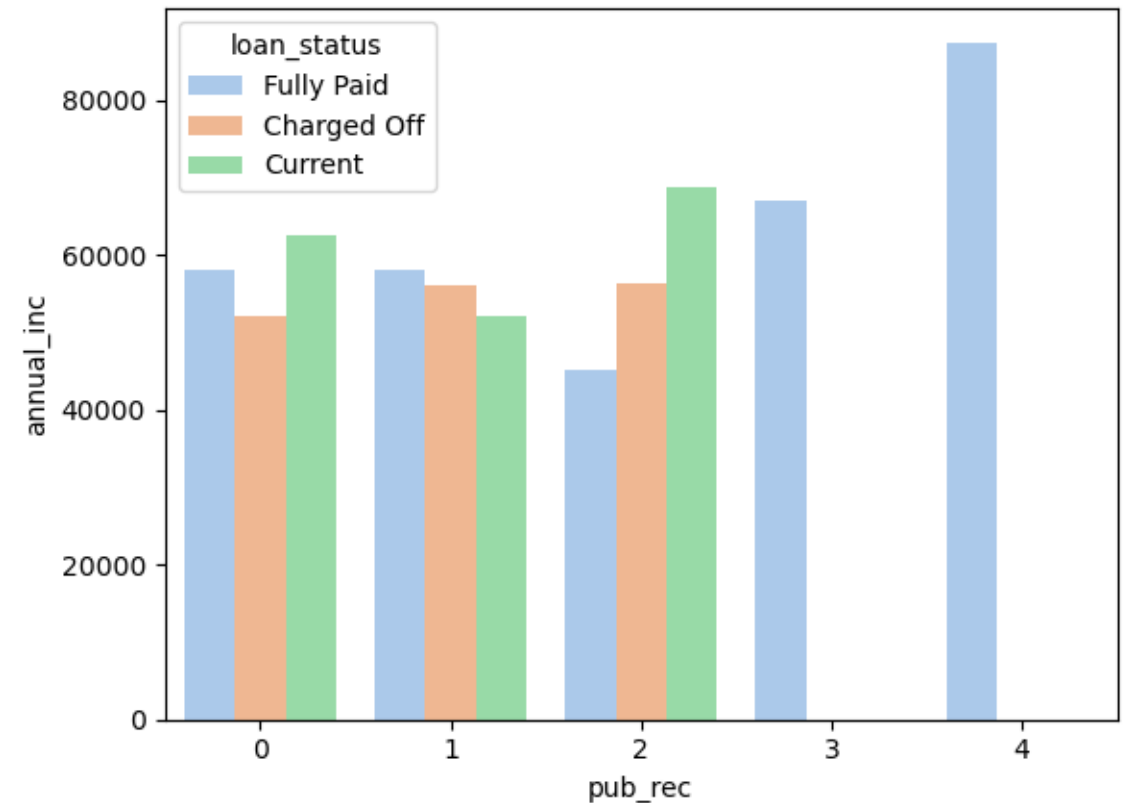
- Members with high revolving utilization rate and opting loan for credit_card, renewable_energy, debt_consolidation are most charged off and they are most fully paid.



Bivariate analysis: continuous categorical

Public derogatory records and annual income:

- The interesting factors from plot is members having more derogatory records with high annual income are most fully paid members and people who have less records are most charged off.



Thank you