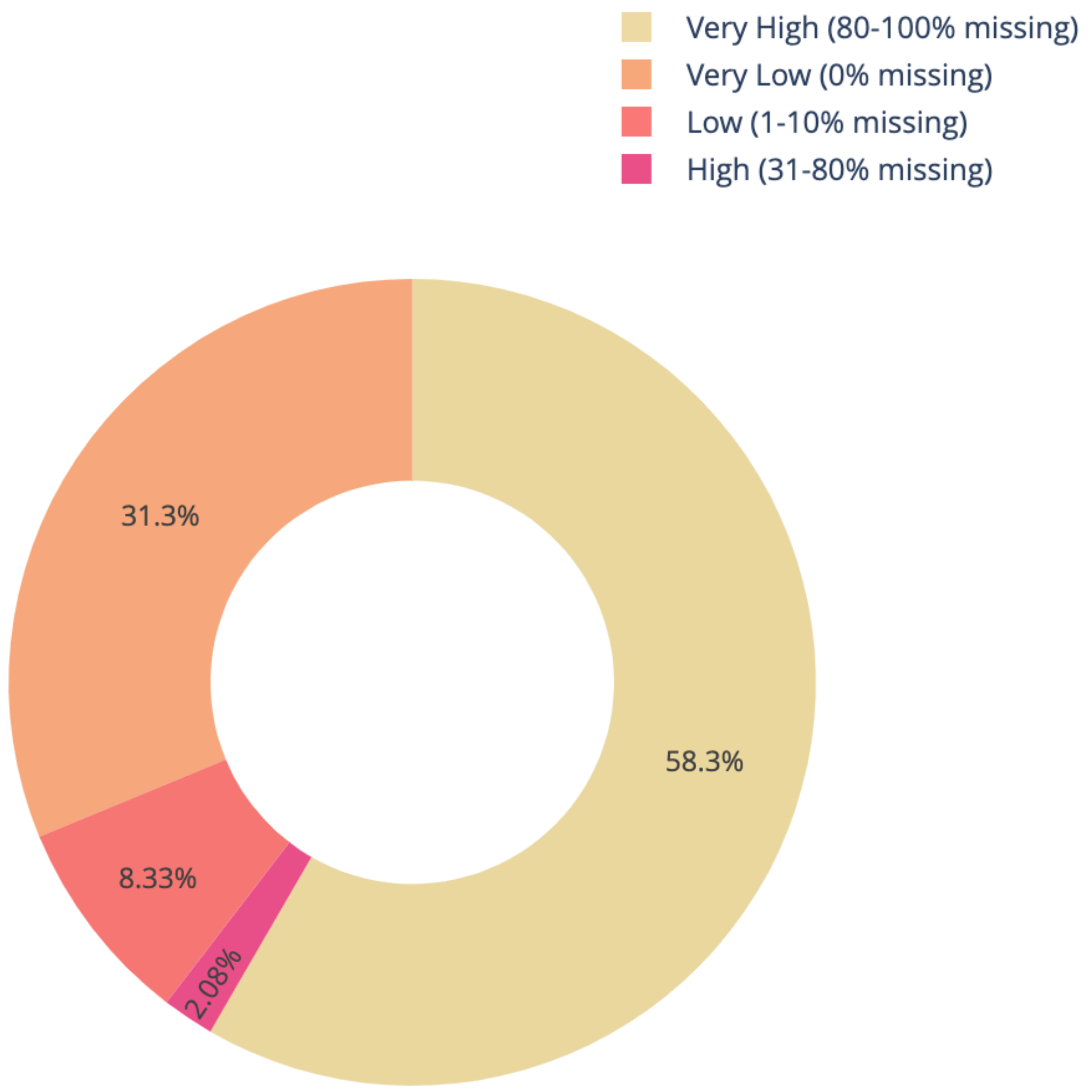
Lending Club Case Study

Sonal Gupta & Shubham Srivastava
ML-C36

Missing Values

- 58% of columns have 100% missing value
- 2% of columns have 31 - 80% missing values
- Only 39.6% of columns have less then 30% missing values

So decided to go with only those rows which have less then or equal to 30% missing value

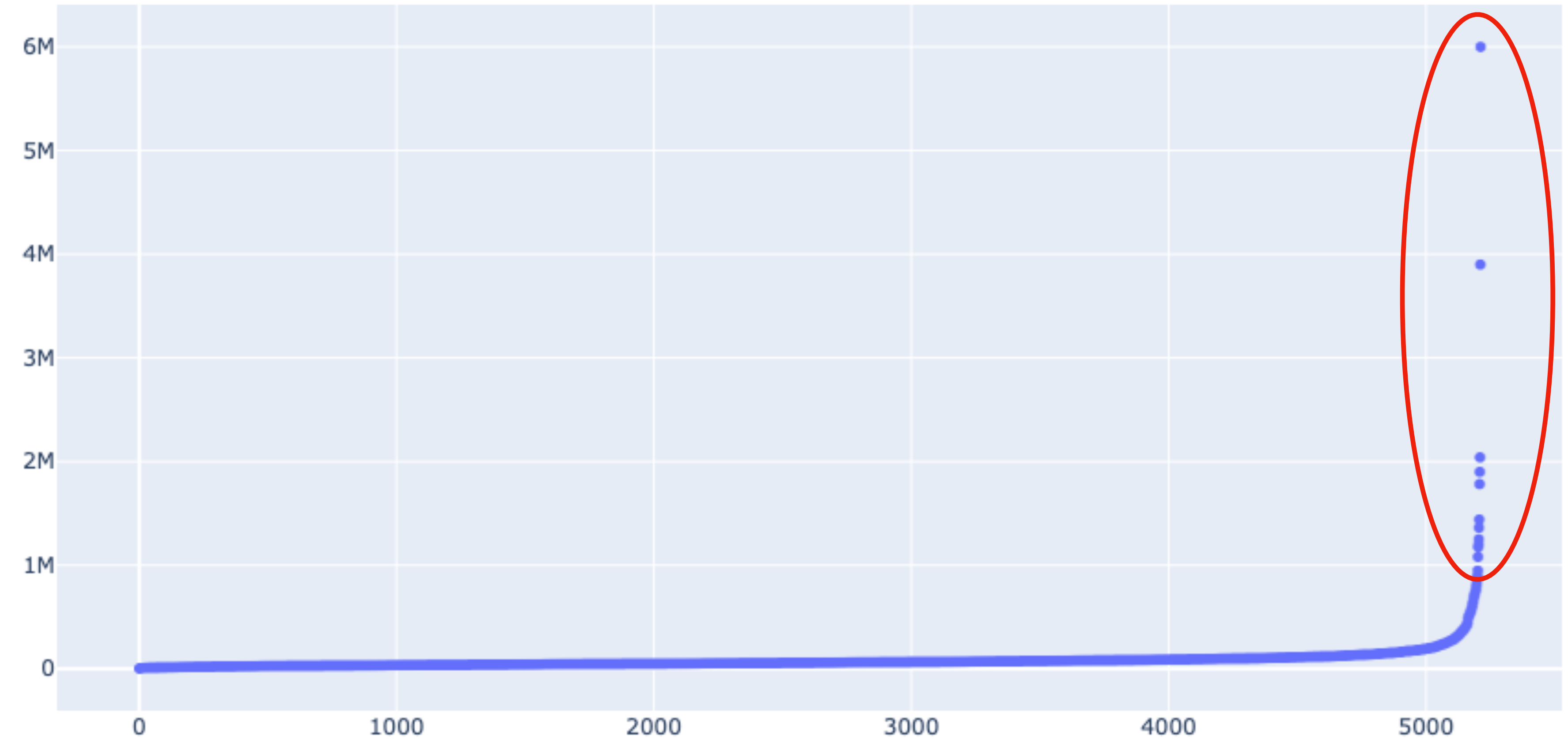


Percentage of column under missing categories(depends on % on value missing)

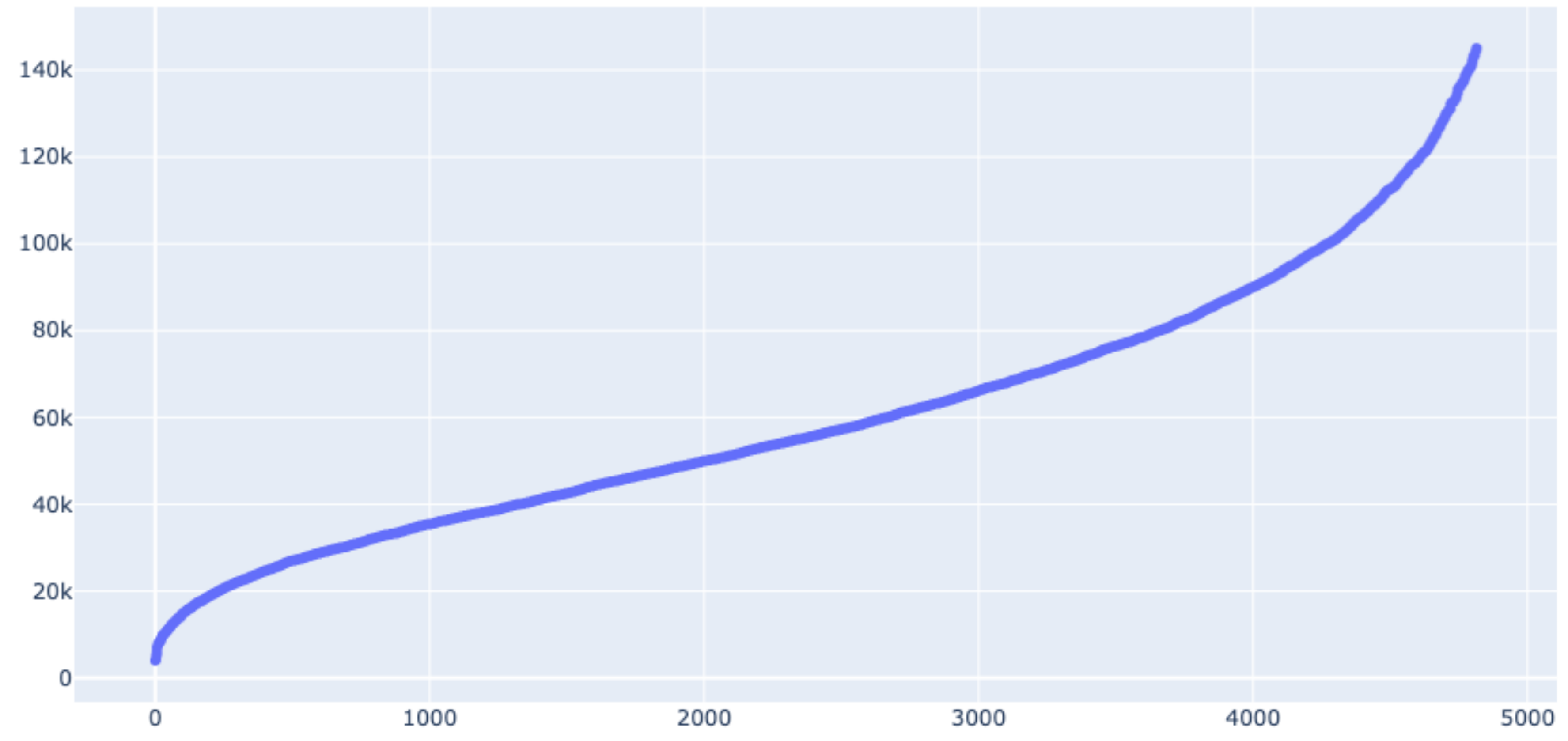
Outliers

- Only annual income data have significant outliers, other column outliers are negligible
- Due to few very high annual income data it is very difficult find some insight

By just dropping few very high value (red circle), able to find that data can be binned into different ranges



Annual Income (with outliers)

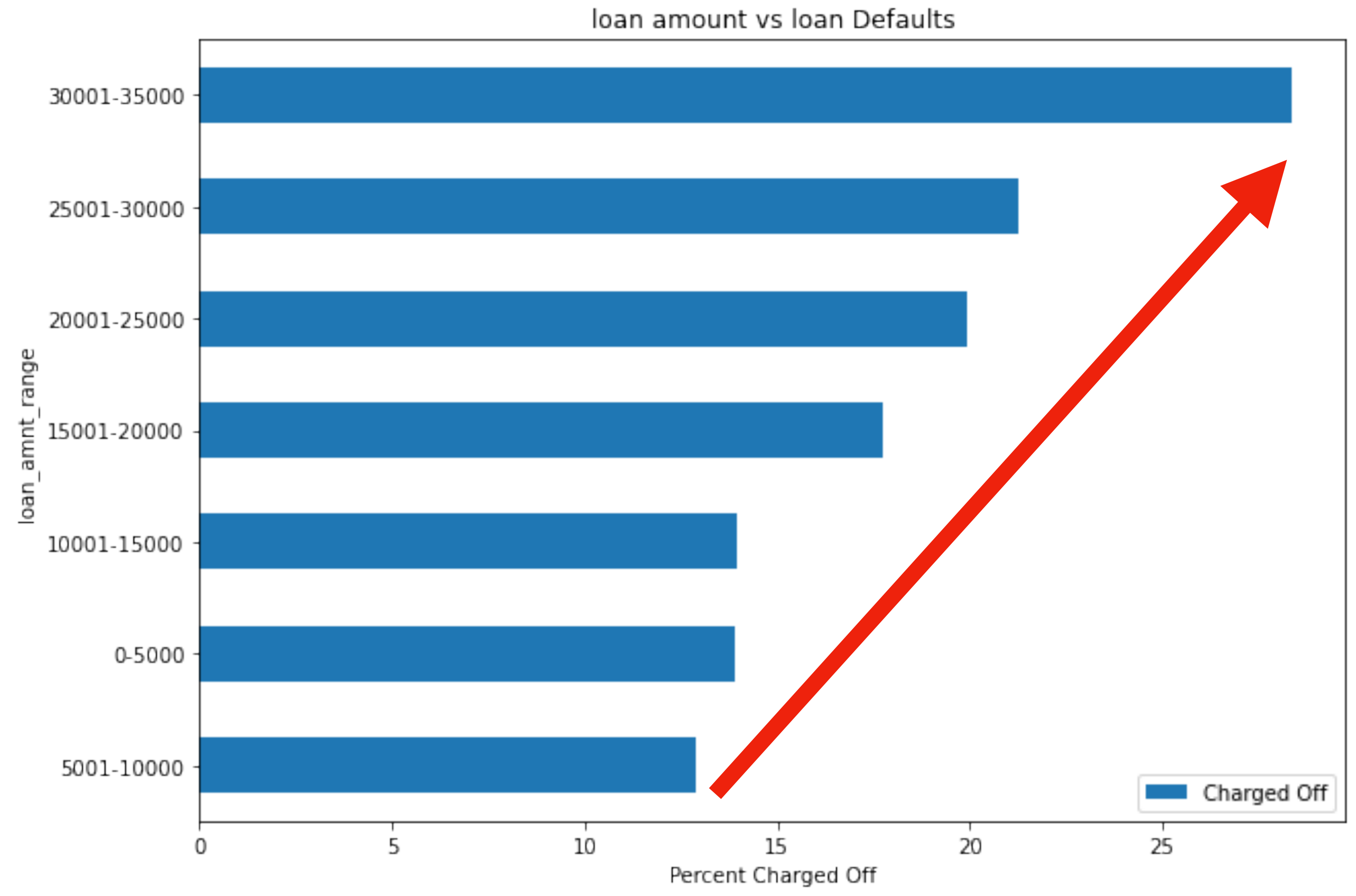


Annual Income (without outliers)

Analysis & Recommendation

Loan Amount

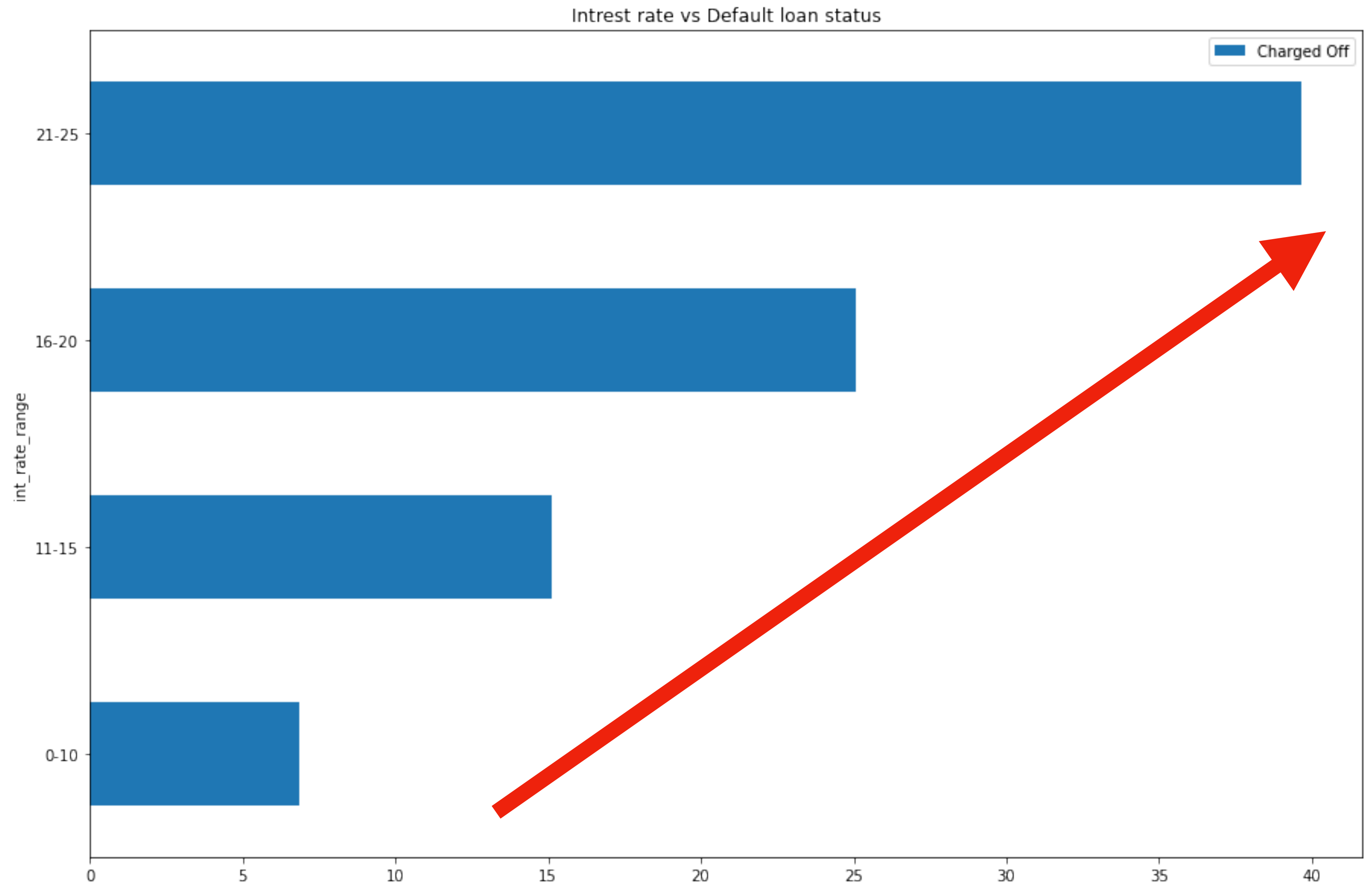
- Higher loan amount tend to default more than lower loan amount
- Charged off percent is as high as 28% in case of loan amount lying between 30-25k making them risky loans



Loan Mount VS Charged Off %

Interest rate

- Charged off loans share linear relationship with interest rate
- Interest rate increase so does the charge off rate
- Most Charged off loans had interest rate between 20-25%



Interest Rate VS Changed Off

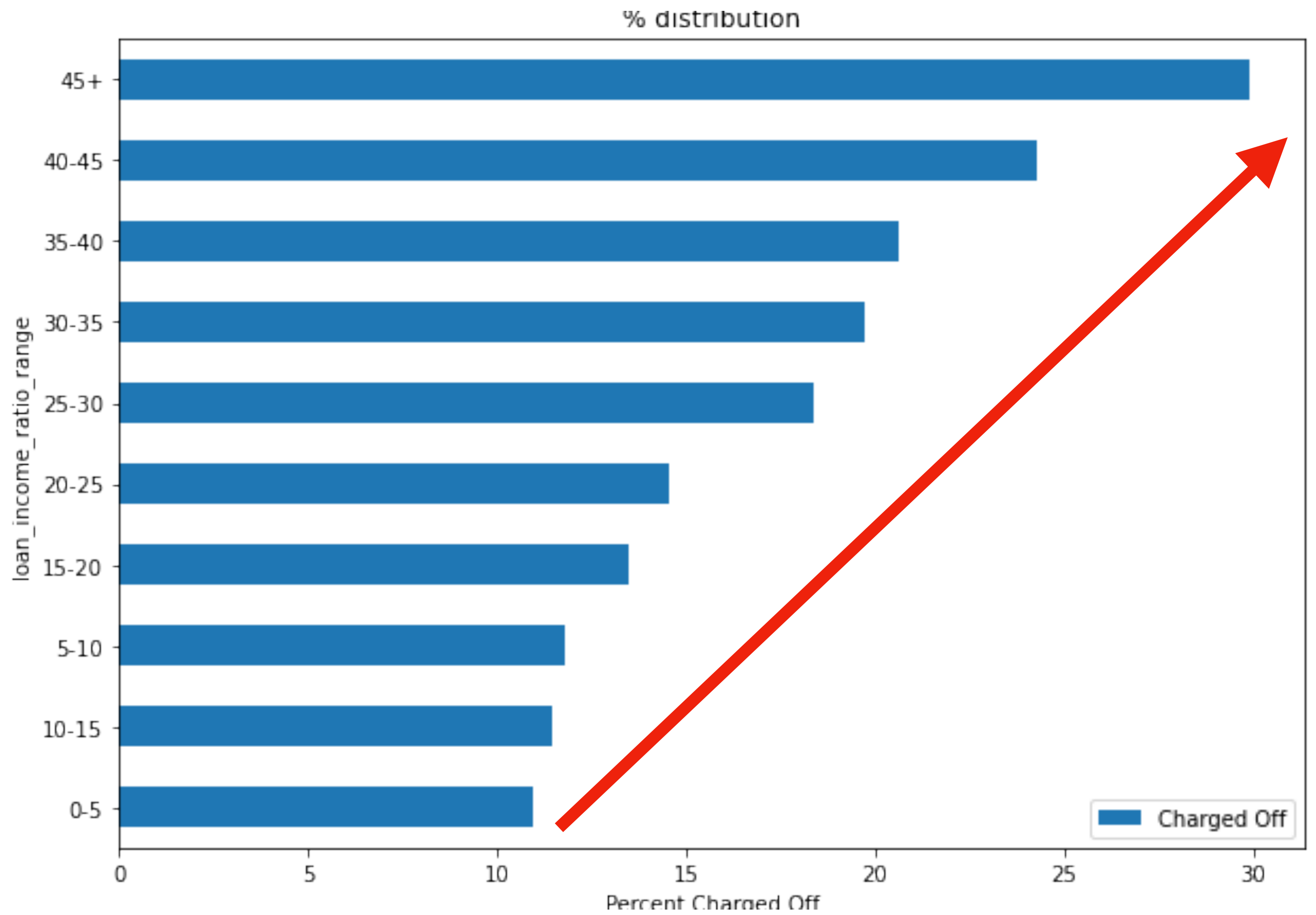
Analysis & Recommendation

Loan Income ratio

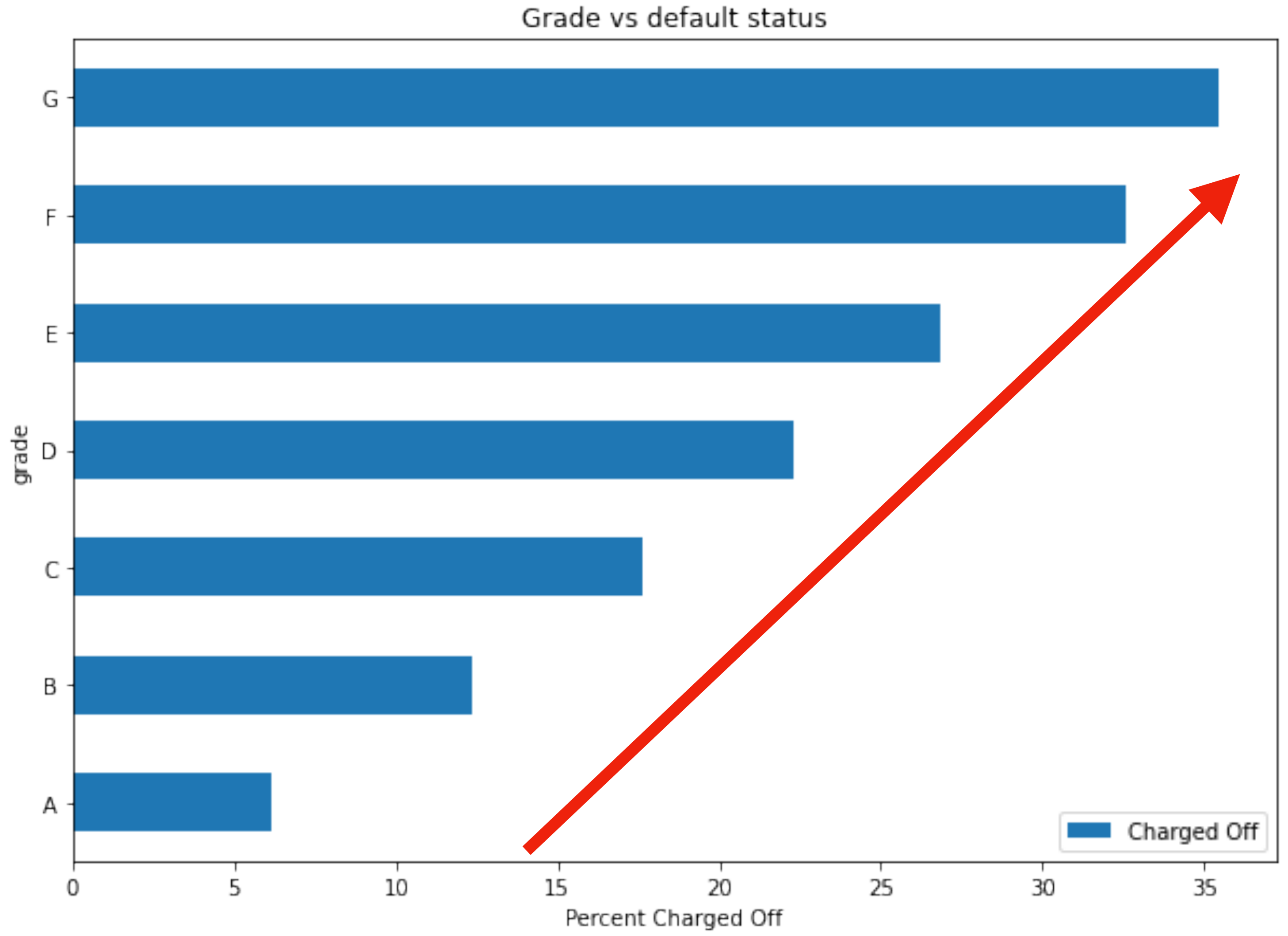
- Loan income ratio tends to be higher for loan defaulters
- Around 30% tend to default when loan income ratio is 45+

Grade

- Moving from A to G grade charged off % increase linearly
- Loans for Grade - G are 35% charged off hence they are most risky loans`
- Loans for Grade - A are 6% charged off hence they are most safe loans



Loan Income Ratio VS Changed Off %

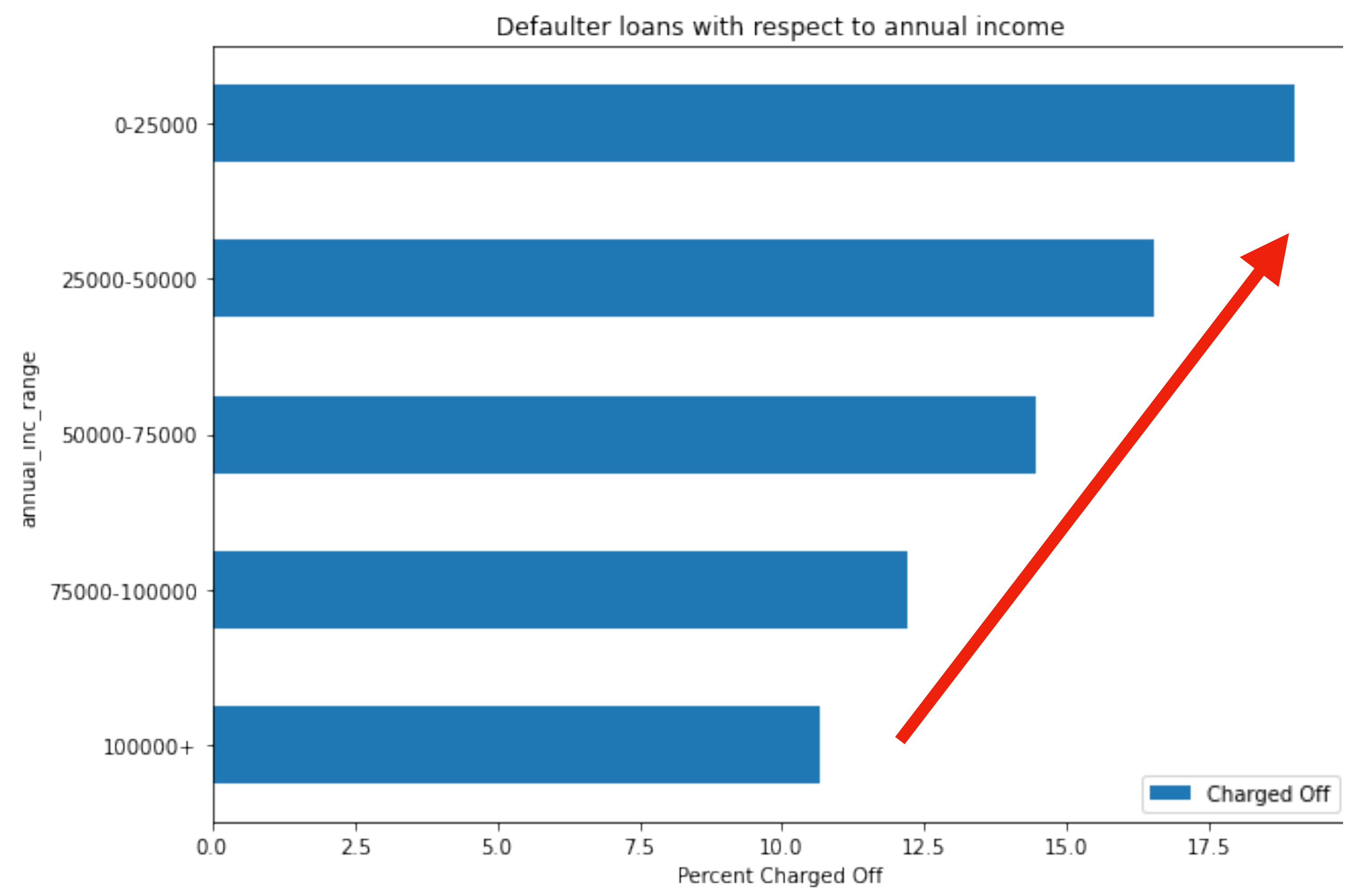


Grade VS Charged Off %

Analysis & Recommendation

Annual Income

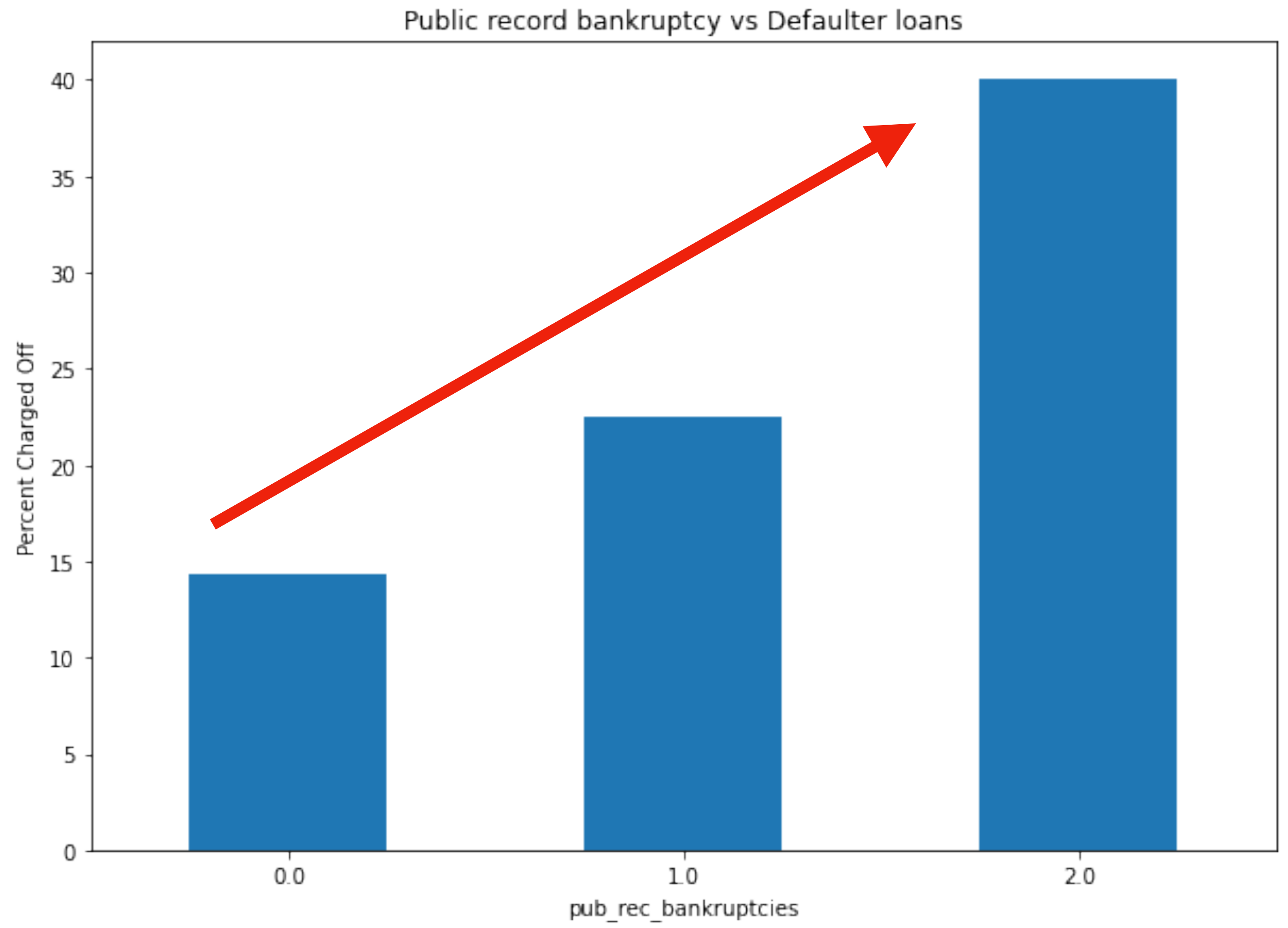
- Charged off is directly inversely related to annual income
- We see an upward trend, ie when the annual income increases the percent of loan default decreases



Annual Income VS Charged Off %

Bankruptcies

- This column is related to loan status,
- The more public recorded bankruptcies more are the chances of a person defaulting the loan
- Its risky to give loans to people with reported bankruptcies



Public Records Bankruptcies VS Charged Off %

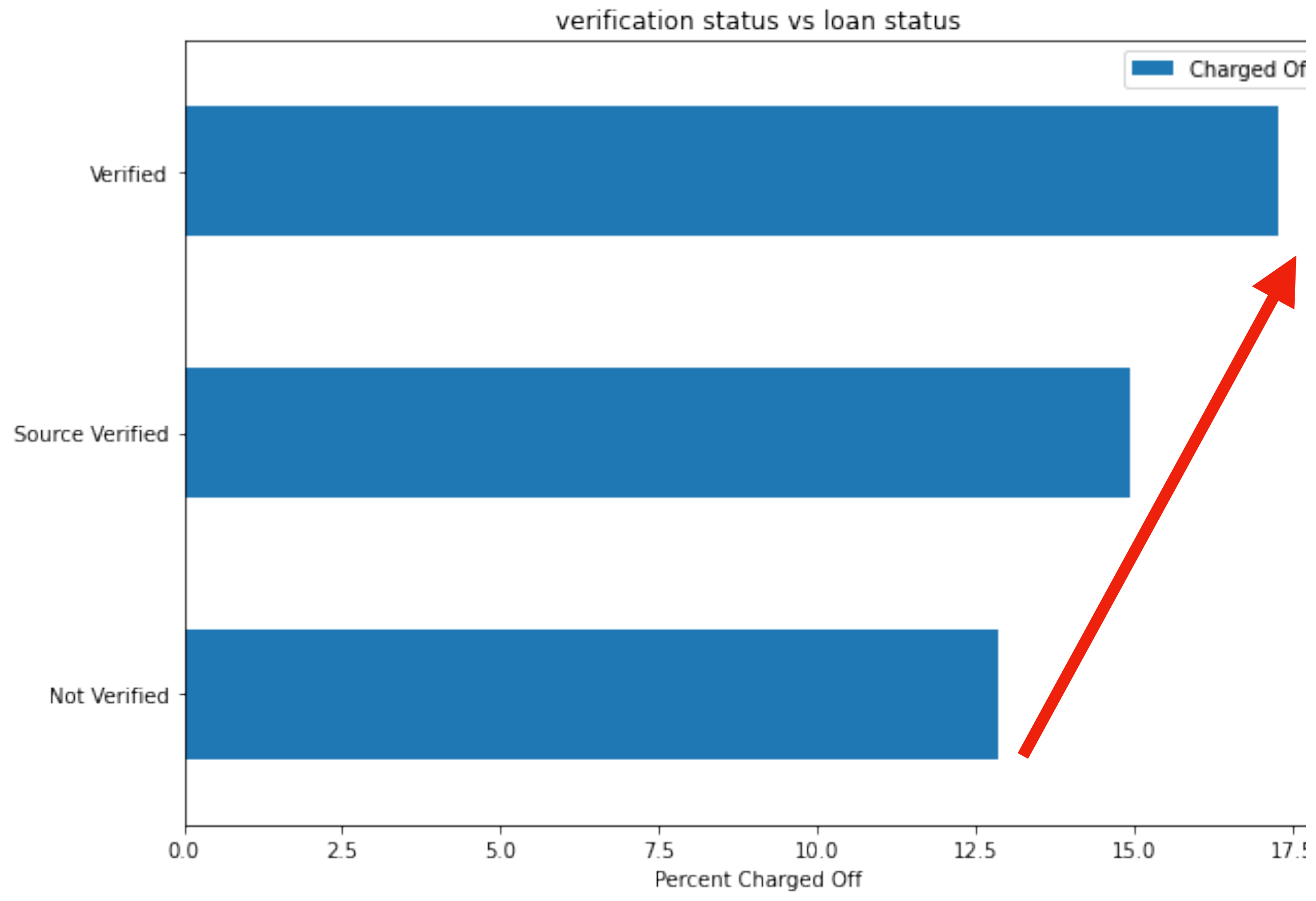
Analysis & Recommendation

Verification Status

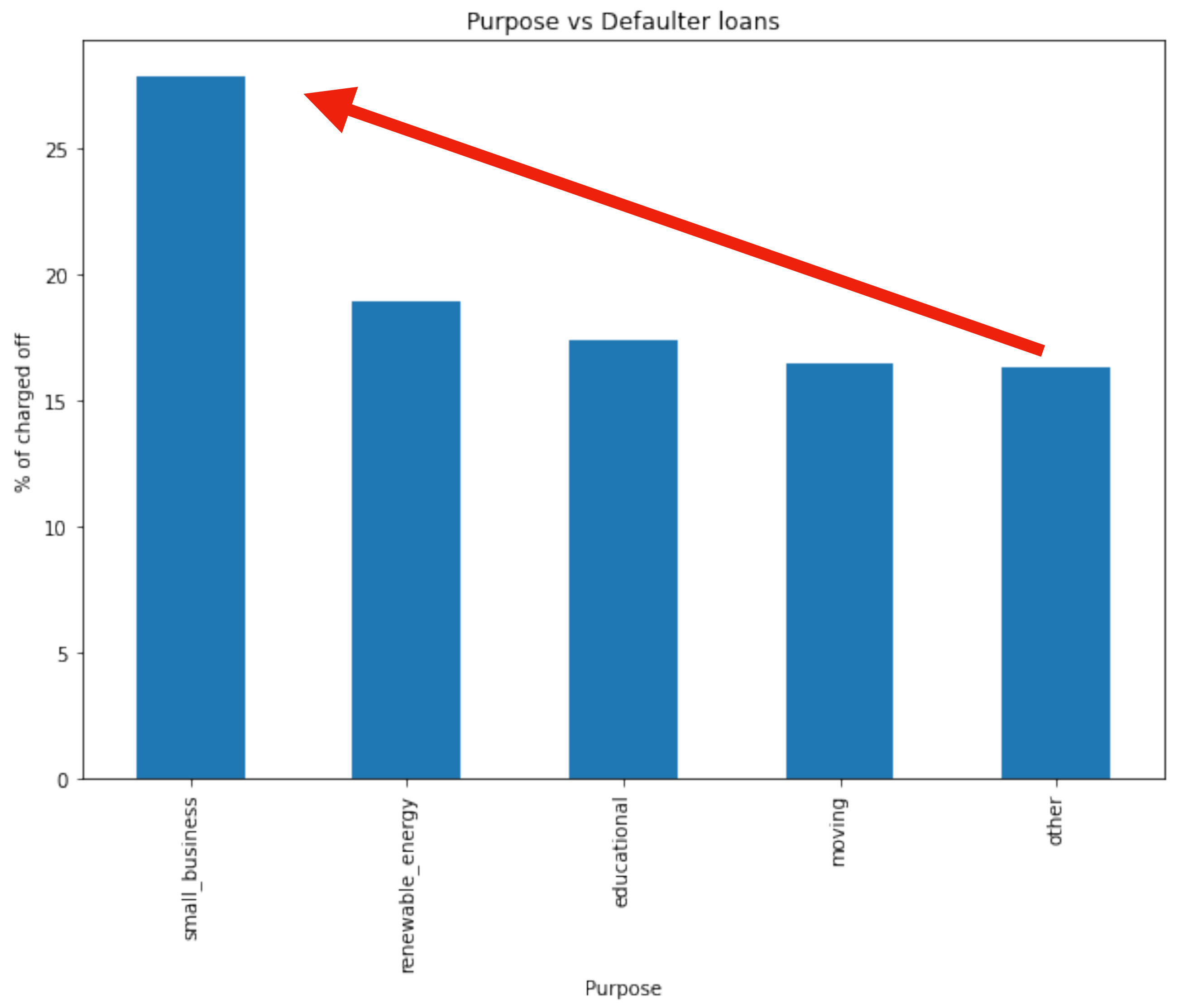
- Most of the loans belong to not verified category
- Loans verified by LC tend to default the most at 17%
- Higher charged off status is seen in loans which were **verified > verified by LC > not verified**

Purpose

- **Debt consolidation** has highest number of loan applications
- Loans for small business are 26% charged off hence investing in it tend to high risk
- Investing in medical is even better (10% lower risk then small businesses)



Verification status VS Charged Off %

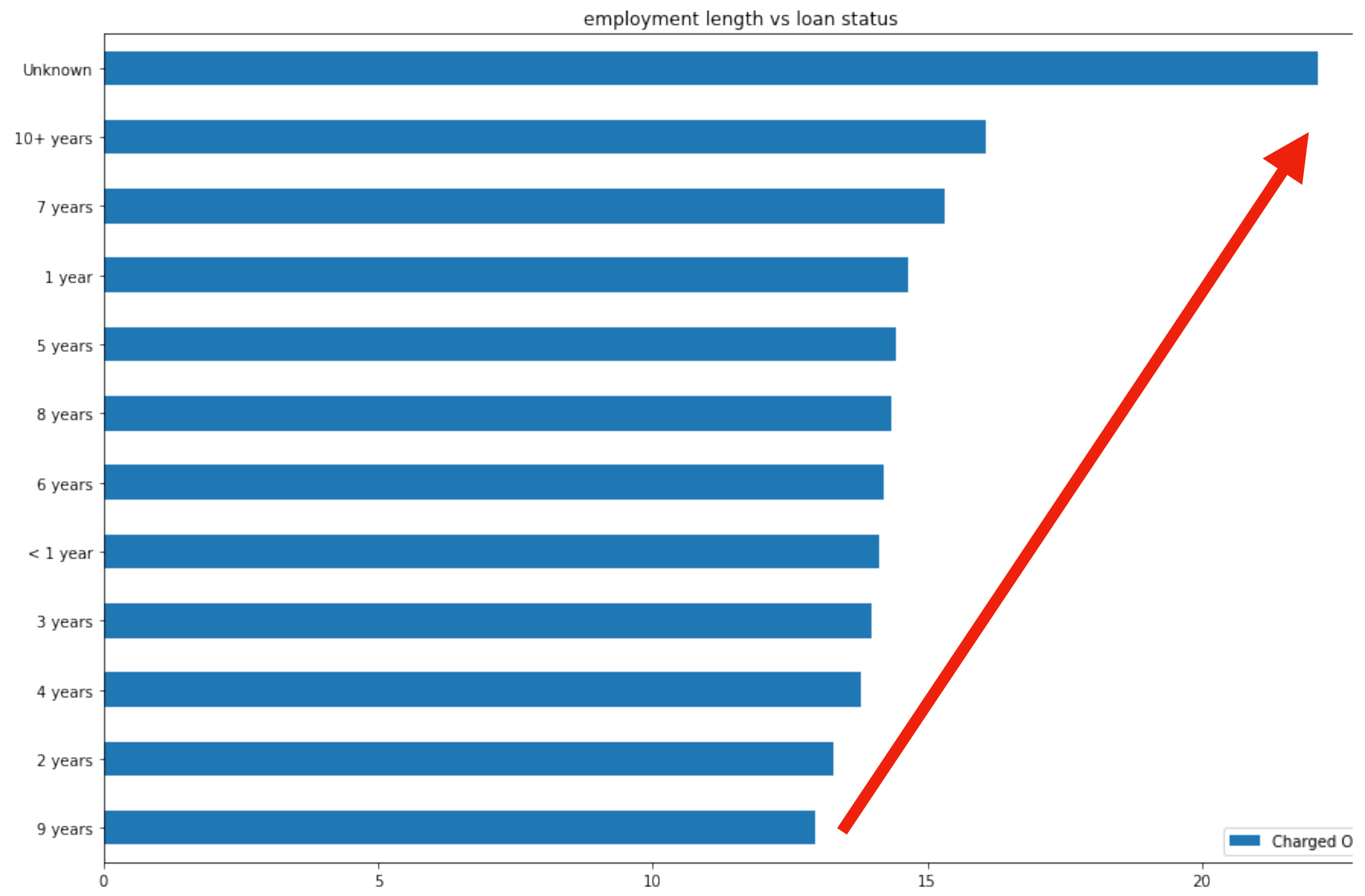


Purpose VS Charged Off %

Analysis & Recommendation

Employment length

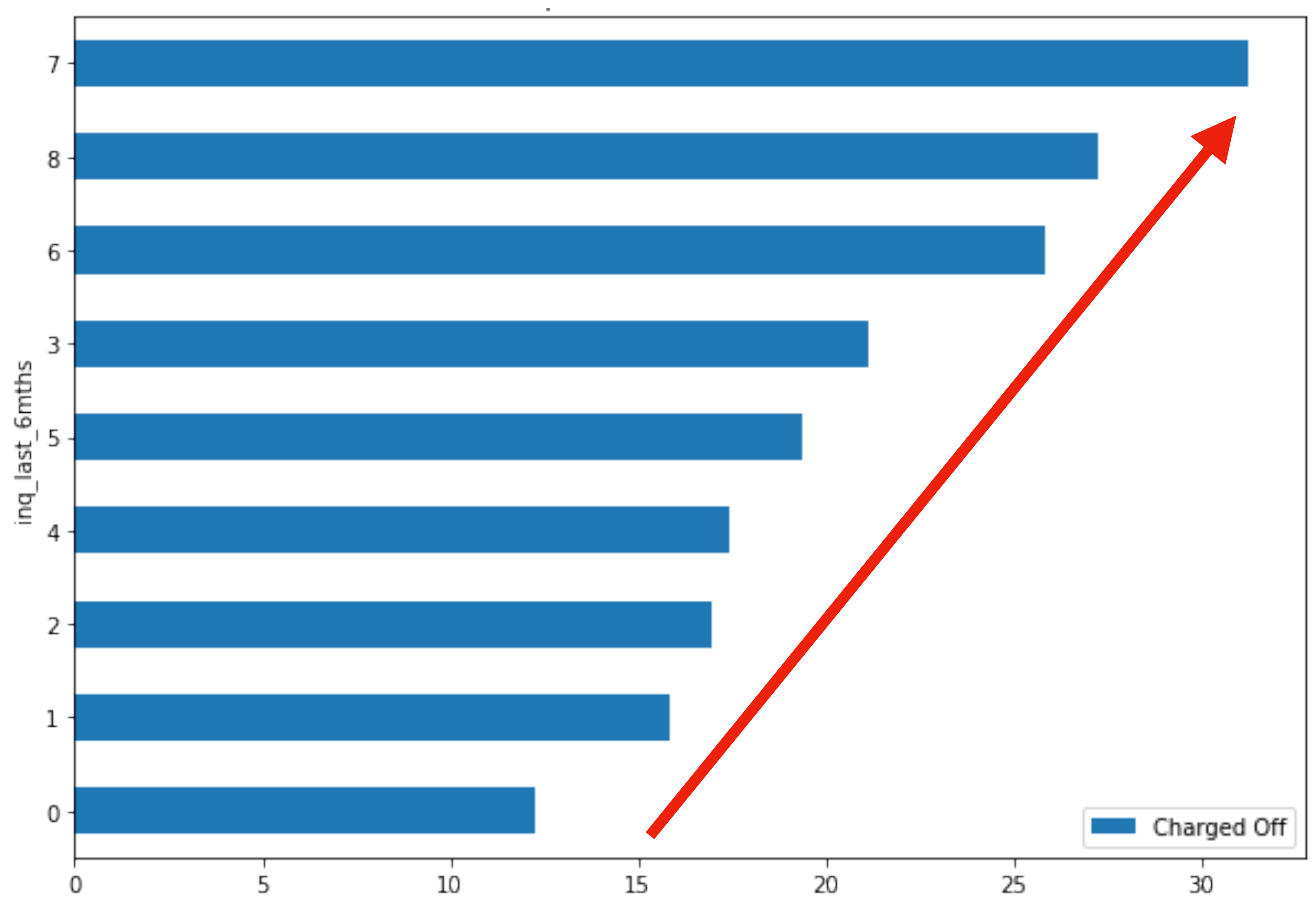
- Loan default was highest where employment length is missing followed by 10+ year employment
- For missing employment information data point 22% are charged off



Employment Length VS Charged Off

No. Last 6 Month Enquiries

- We see a slight upward trend enquires against loan status
- When enquires increase defaulter loans tend to increase



Last 8 Month Inquiry VS Charged Off %