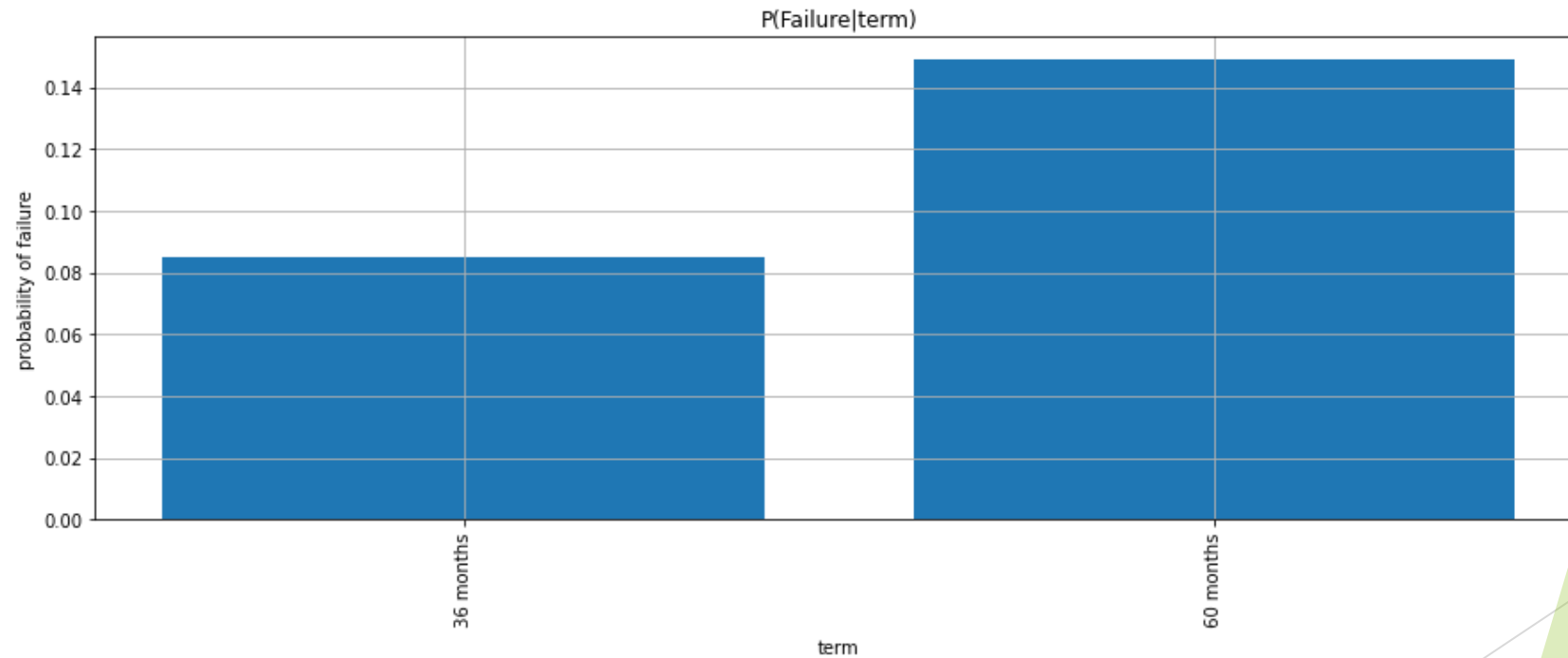# LENDING CLUB CASE STUDY

V S S ANIRUDH SHARMA

PROBLEM STATEMENT

- In this assignment, we are trying to analyse the impact of various factors on loan repayment defaulting.

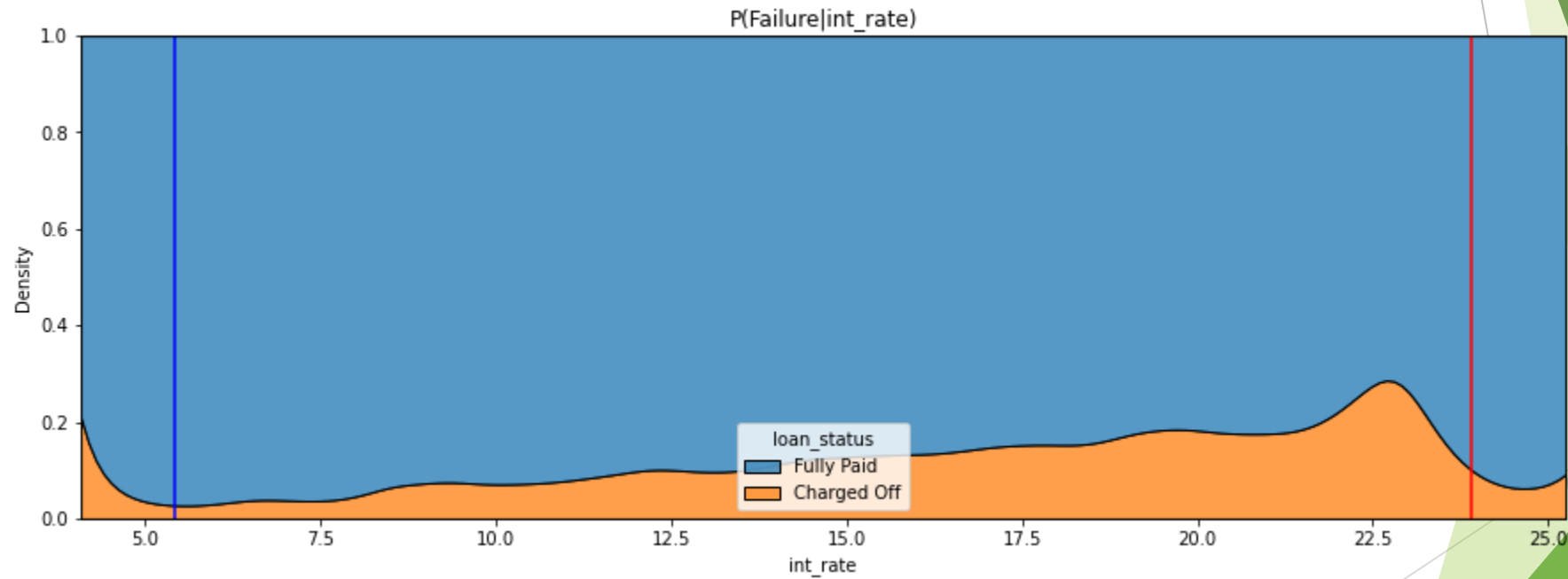- We shall perform Exploratory data analysis and draw observations and provide suggestions to minimize risk and loss

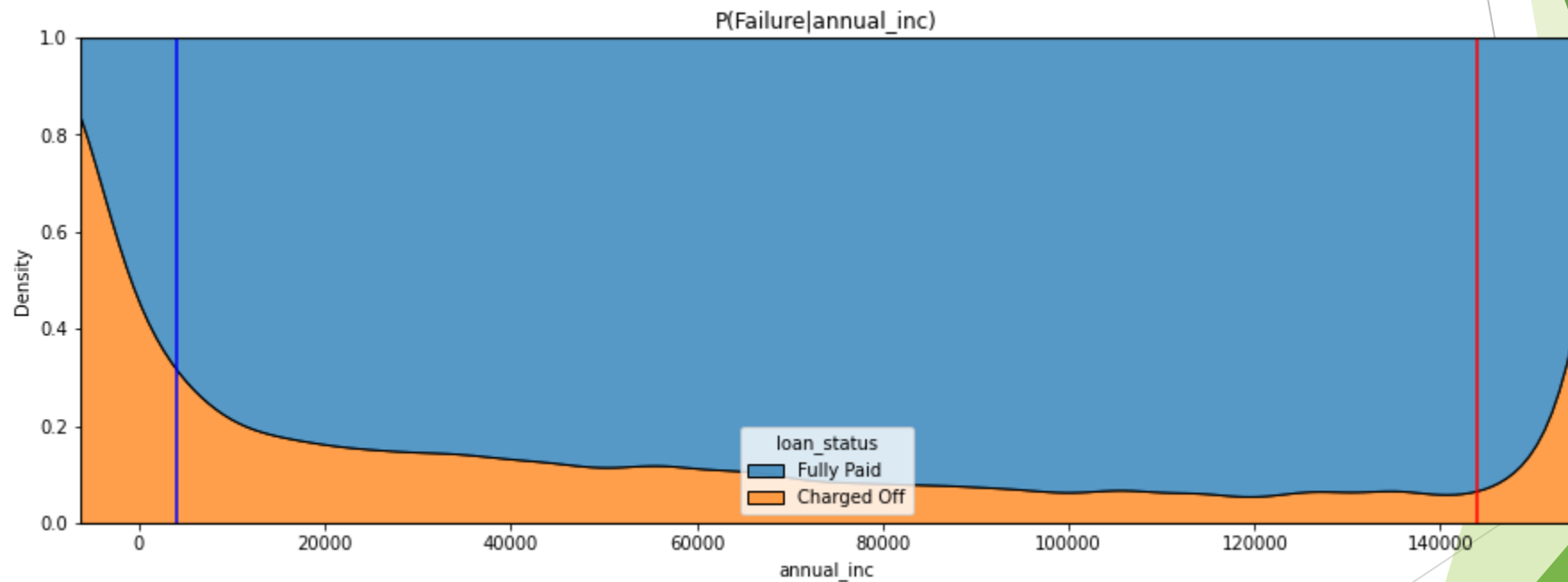# 60 month terms are significantly much riskier than 36 month terms

The probabilty of defaulting seems to increase down the grades, A being the safest and G being the riskiest
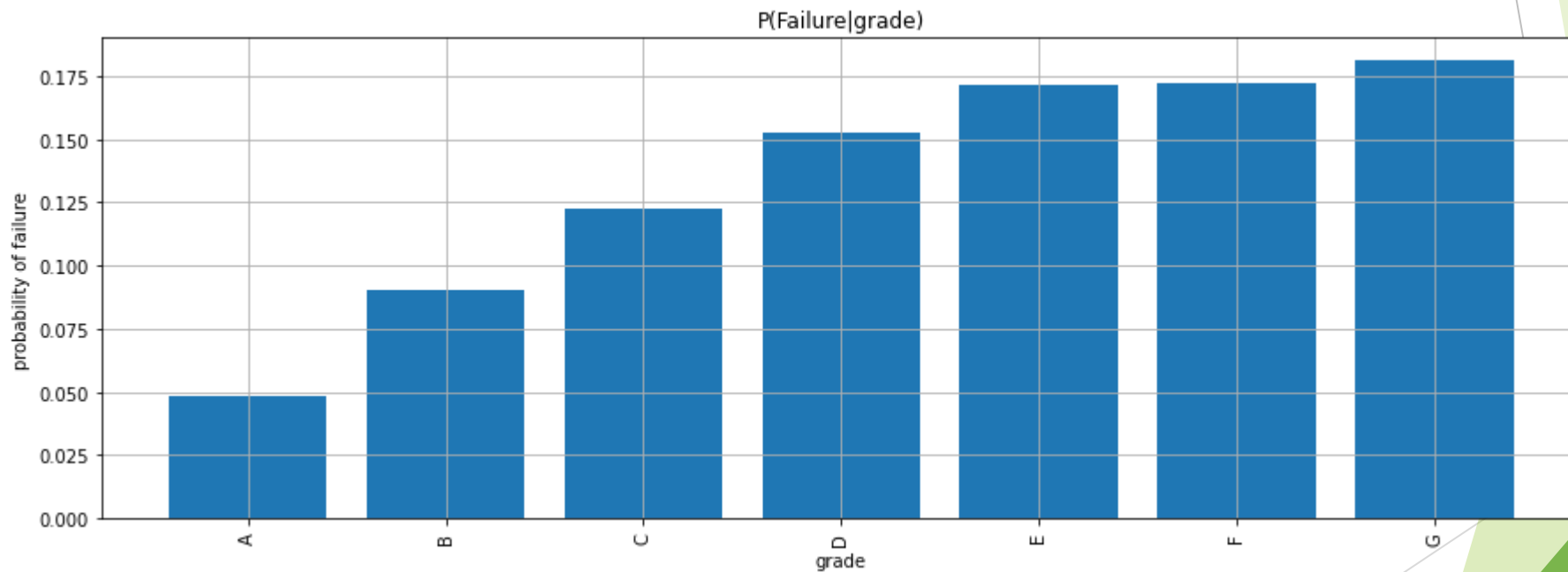
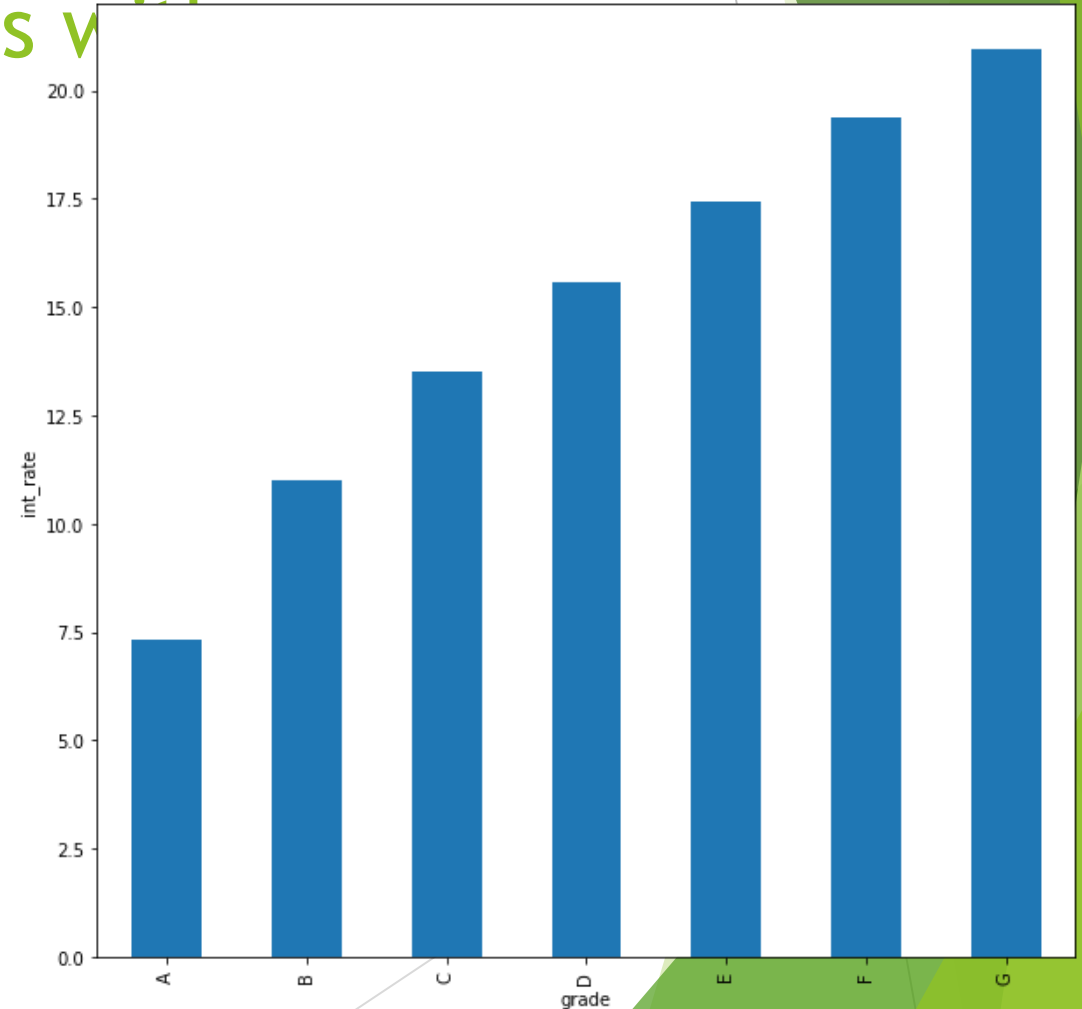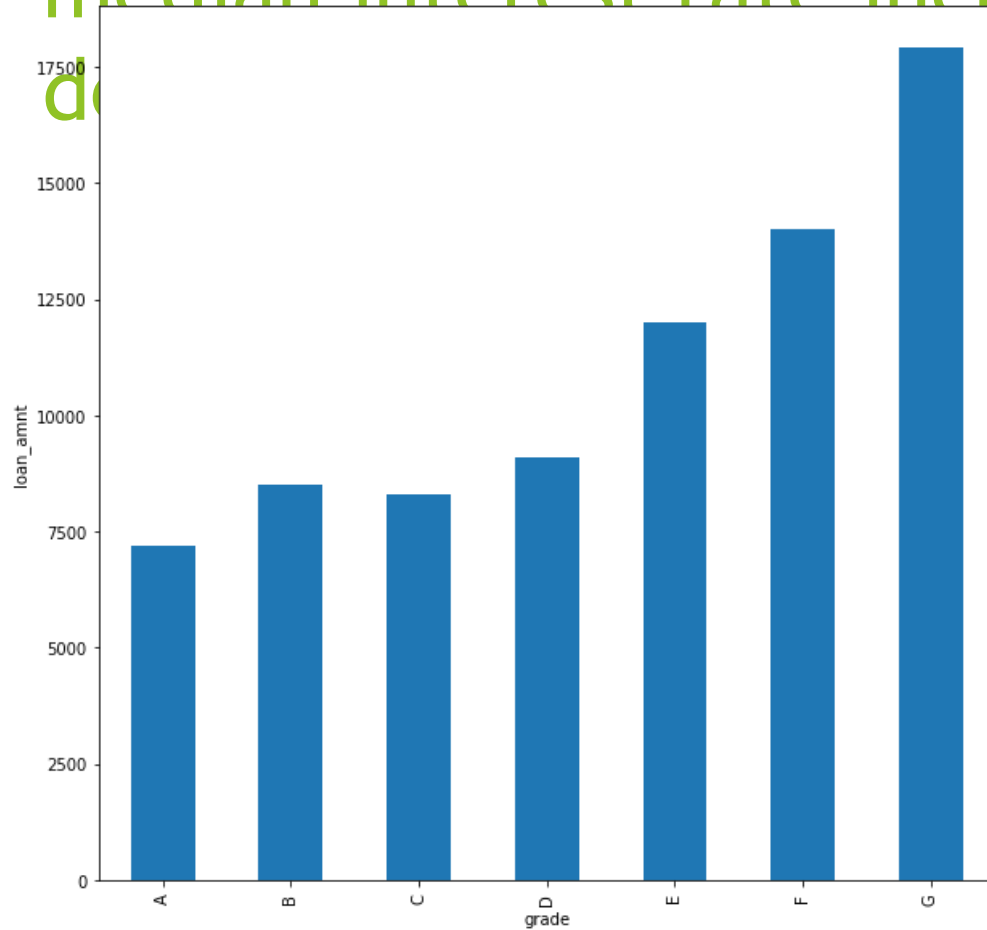# The probability of defaulting seems to increase with int_rate

The probability of defaulting seems to decrease with annual_inc

median loan amount increases with decreasing grade
median interest rate increases with decreasing grade

- 1. Higher the grade (A is highest), lower the interest rate
- 2. Higher the interest rate, higher the revolving utility
- 3. annual_income is mildly postively correlated with loan/funded amounts

Number of loans issue increased month by month, almost doubling each year
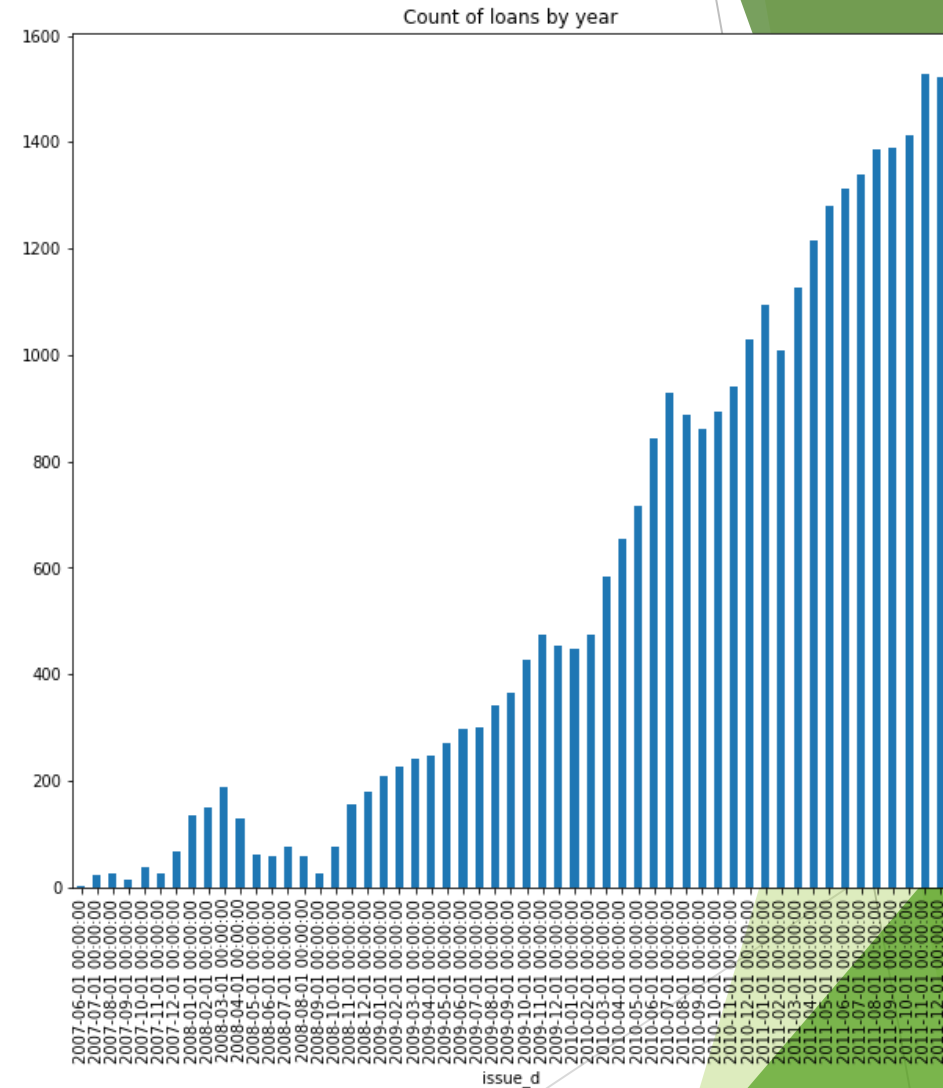
Each year saw an increase with year average loan
But annual incomes didn't change much over the

Average loan amount by year

Average income by year
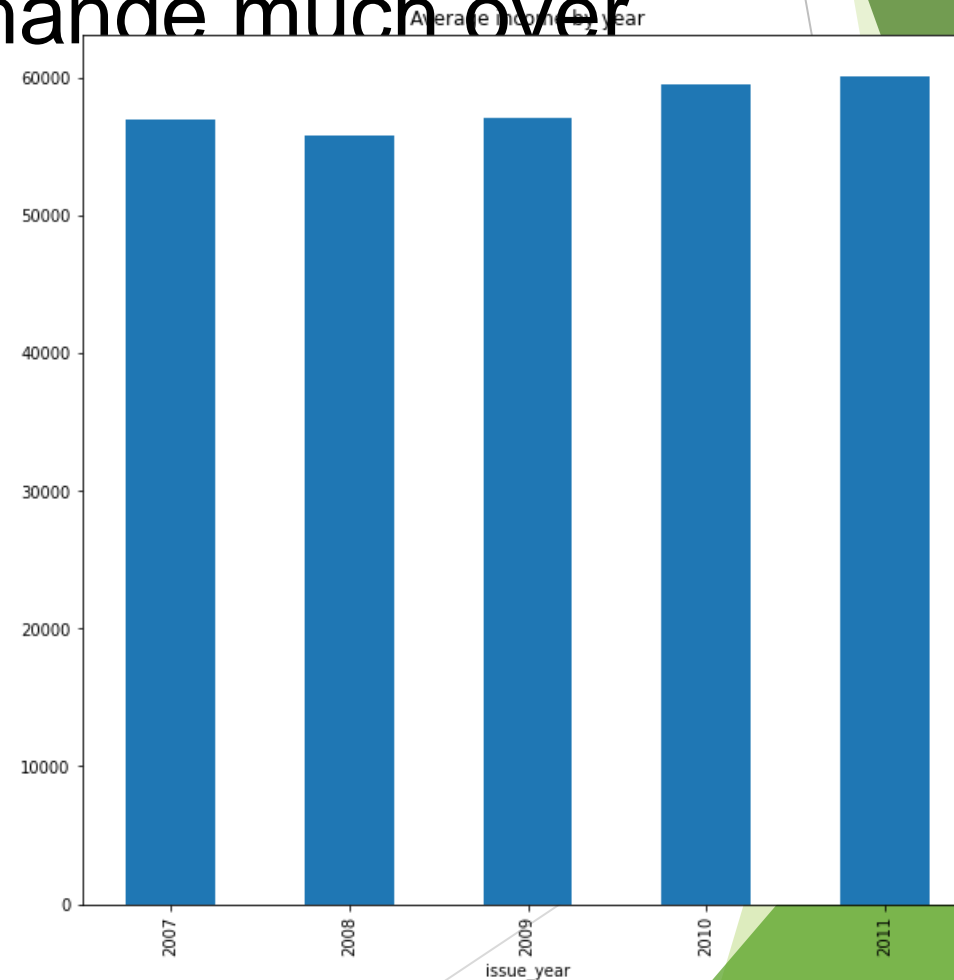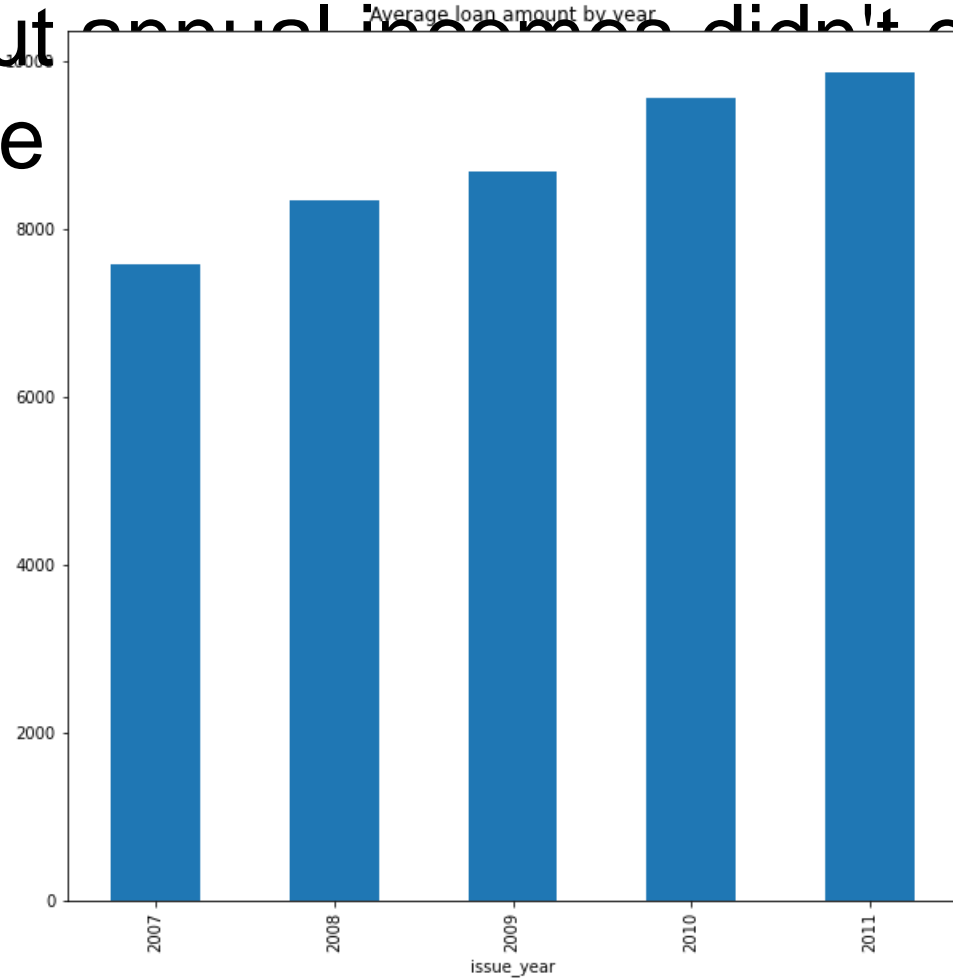
OBSERVATIONS

From univariate analysis

- ▶ 1. Most of the lendings were successful
- ▶ 2. Major purpose for applying for loan is debt_consolidation (almost 50%)
- ▶ 3. Second major purpose is for credit_card
- ▶ 4. Less than 1% of lendees have an education loan or renewable energy loan
- ▶ 4. Most income sources have not been verified by LC
- ▶ 5. CA has been the most active state where loans were give (about 18%)
- ▶ 6. About half the lendees rent and alomst another half have mortgaed their homes.
- ▶ 7. People with unmortgaged homes (7.5%) are least likely to apply for a loan
- ▶ 1. More than 3/4th of lendees went for 36 months term
- ▶ 2. More than 3/4th of lendees are graded C or above

From baysian analysis of P(Fail|factor)

- ▶ 1. OTHER type of home ownership is the riskiest albiet by a small margin
- ▶ 2. Small businesses are the riskiest to lend to
- ▶ 2. 60 month terms are significantly much riskier than 36 month terms
- ▶ 3. The probabilty of defaulting seems to increase down the grades, A being the safest and G being the riskiest
- ▶ 5. probability to default seems to increase with derogatory public records
- ▶ 6. In general, chances of defaulting seem to increase with inquiries in last 6 months
- ▶ 1. loan_amt shows up to have some impact on the default prbability: cases of fully paid are distribute around amounts greater than those in case of defaulting.
- ▶ 2. The probability of defaulting seems to increase with revol_util, and int_rate
- ▶ 4. The probability of defaulting seems to decrease with annual_inc

Bivariate analysis

- 1. median Small business, house related and credit/consolidation loans are higher than the rest

- 2. median loan amount increases with decreasing grade

- 3. median interest rate increases with decreasing grade

- 1. Higher the grade (A is highest), lower the interest rate

- 2. Higher the interest rate, higher the revolving utility

- 3. annual_income is mildly postively correlated with loan/funded amounts

Time series analysis

- 1. Number of loans issue increased month by month, almost doubling each year

- 2. Each year saw an increase with year average loan

- 3. But annual incomes didn't change much over the years

SUGGESTIONS

- Consider only cases with very few previous charge offs

- Reduce rate of interests. Interest rates might be higher for smaller grades, but the reason of defaulting could be high interest accumulation.

- Reduce loan amount sanctions for lower grade/more previous charge offs lendees to ensure minimal losses.