Lending Club Case Study using EDA

- N VENKATESHAN

- ABHISHEK MISHRA

Background

- Lending Club, a US based finance company facilitates different type of loans to customers

- Decision is made weather to approve or reject the loan based on different parameters.

- It operates on peer to peer lending

- Based on the consumer given data the investor will decide weather to invest in the customer.

- After the investors approval the lending club will provide interest rates based on the different loan grades.

- The company can run into loss under the circumstance that the customer does not pay the loan amount.

- The company can also loose business if a good customer is denied loan by the investor.

Business problem

- Lending Club categories their customers into grade and sub grade based on credit history.

- Investors are free to chose which borrower they want to invest in.

- If the borrower does not repay the loan or run away with out paying the company and investors loose money

- If a good applicant is rejected in the process of loan approval the company looses business.

- More reliable ways of identifying bad customers from the good one are needed.

Approach

Domain Knowledge
- Understanding how the business works

Data Understanding
- Understanding the given data

Data Cleaning
- Removing irrelevant, null & outlying values

Data Analysis
- Conducting Univariate & Bivariate analysis

Summary & Recomendation
- Learn from the data & Recommend the company to take decisions

Interest Comparison based on Grade

It can be seen that the loan interest is based on the loan grade.

The loan interest rate of the charged off customers is higher than the fully paid customer in every case.

Status of loan based on Interest and Grade
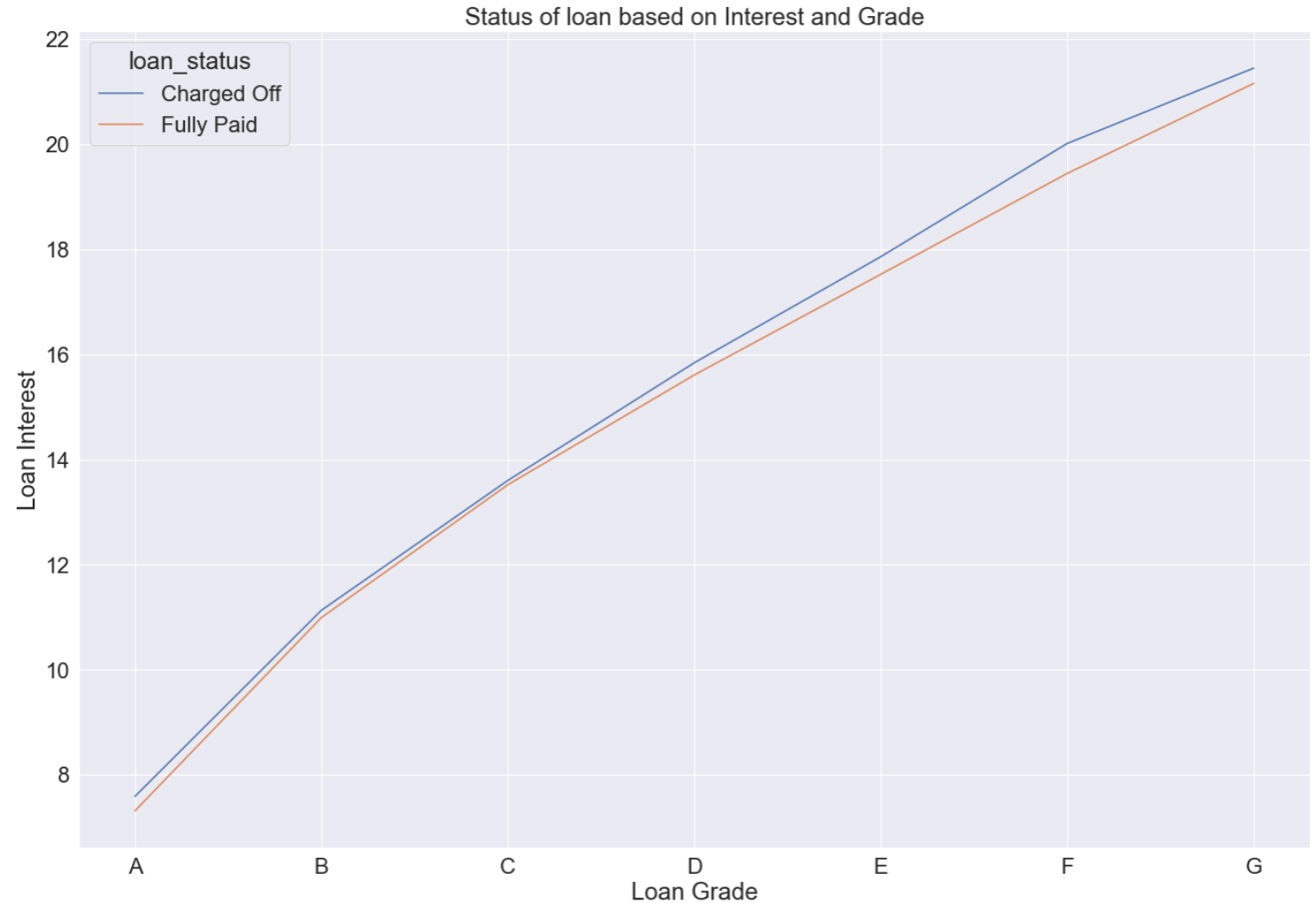
Loan Status based on Grade

The percent of loan defaulters is increasing with the loan grade.

Percent of Customers that are charged off are especially higher in F5 & G3 loan sub category.

Number of loan based on Interest and SubGrade
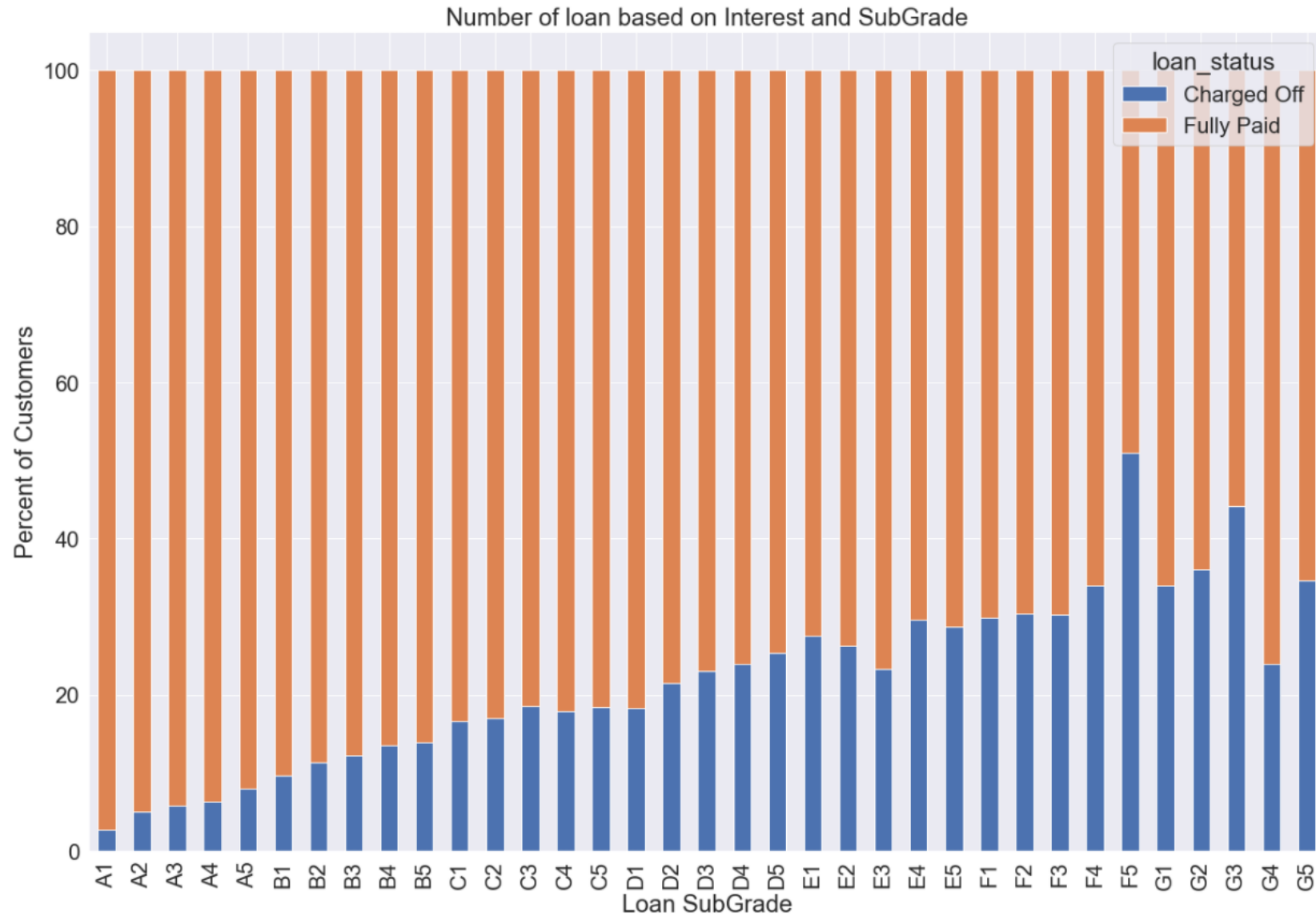
Distribution of DTI based on Purpose

Customers charged off for every purpose have a higher dti than the fully paid customers except for medical purposes.

Customers with the purpose of vacationing have a high difference in dti of charged off and fully paid

Status of loan based on dti and purpose
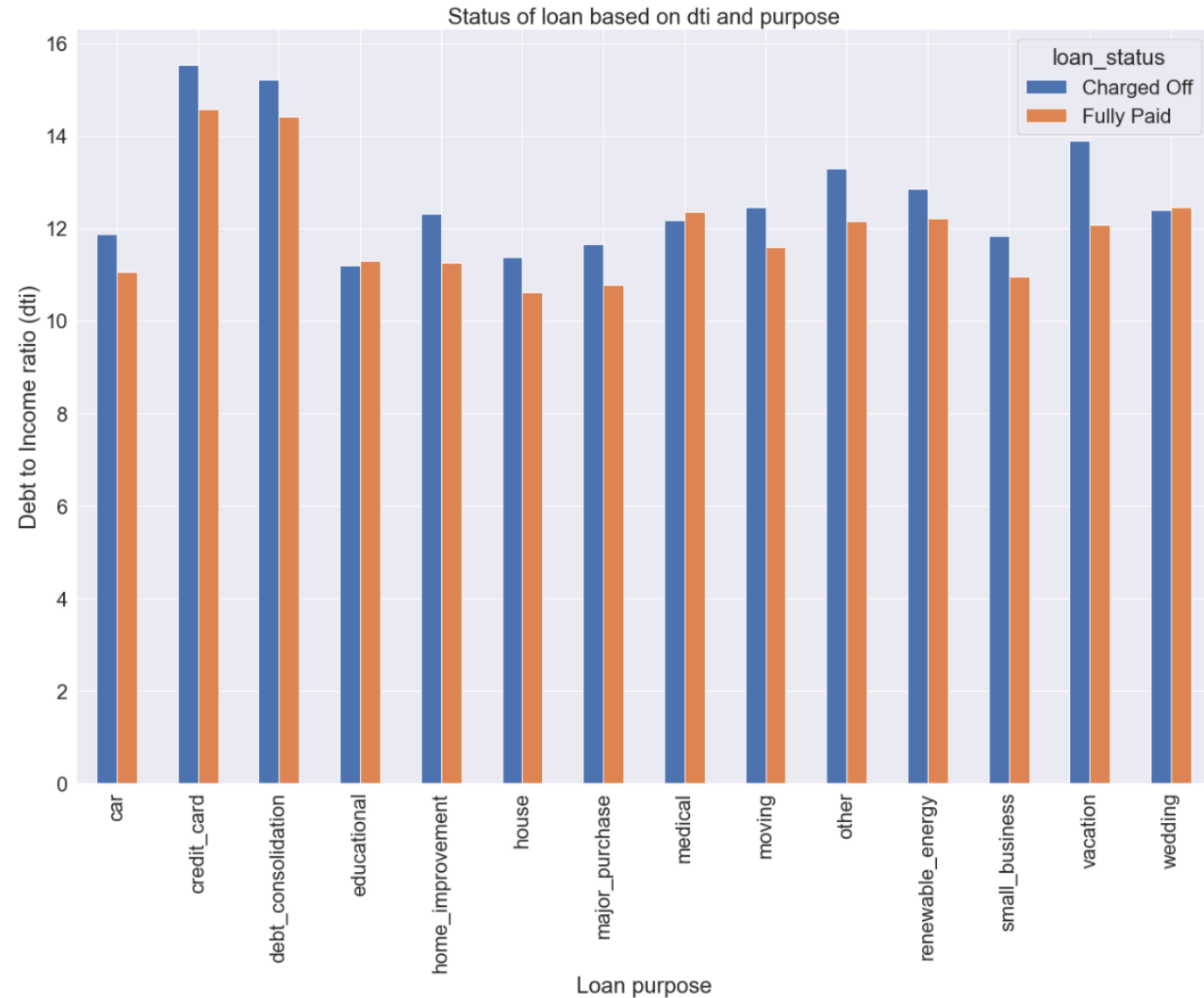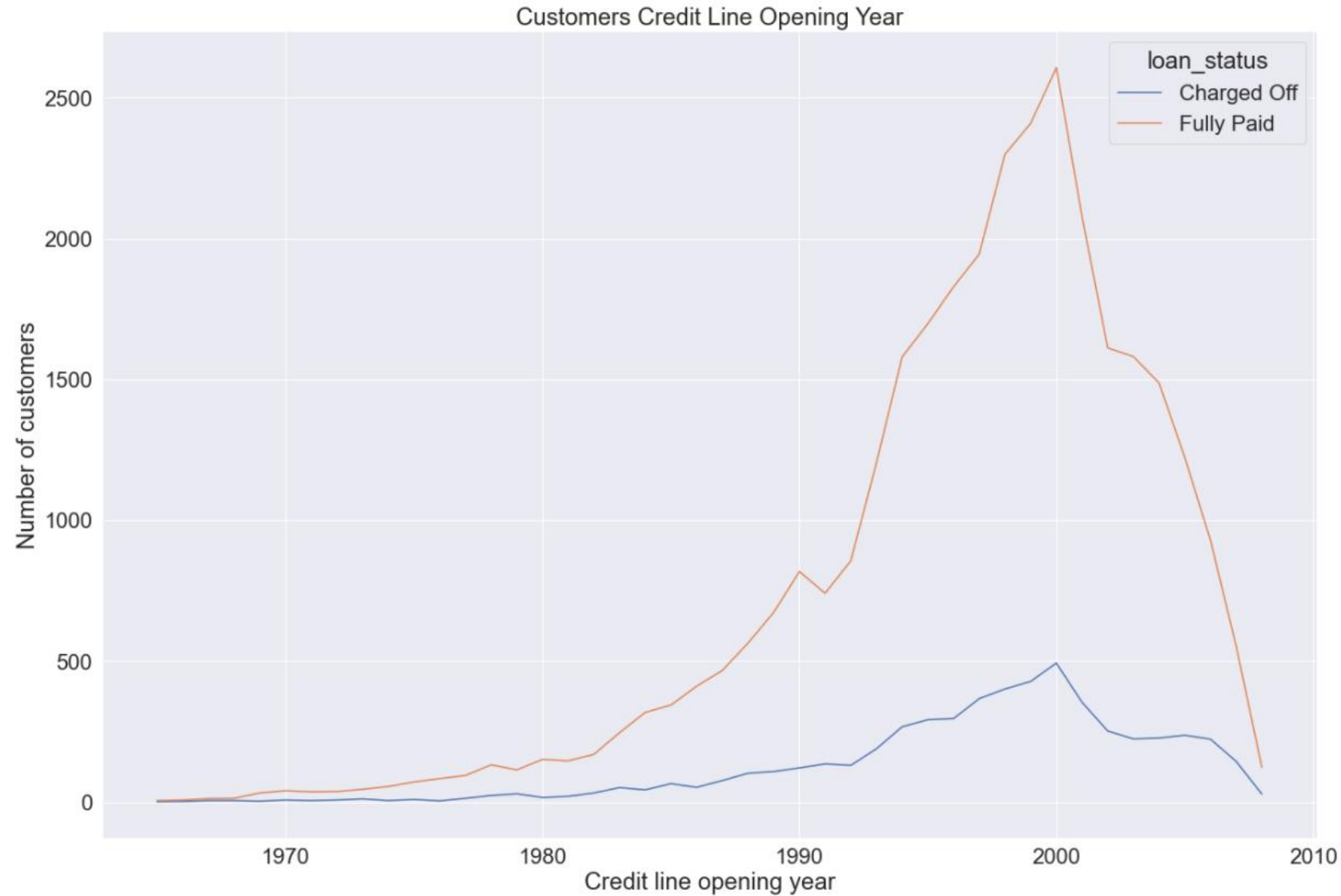
Interest Rate & Purpose of Loan

The interest rate of the charged off customers is consistently higher than that of the fully paid customers in every category of loan purpose.

Other than small business all of the customers above mean 12 % interest rate are charged off.

Loan interest by purpose
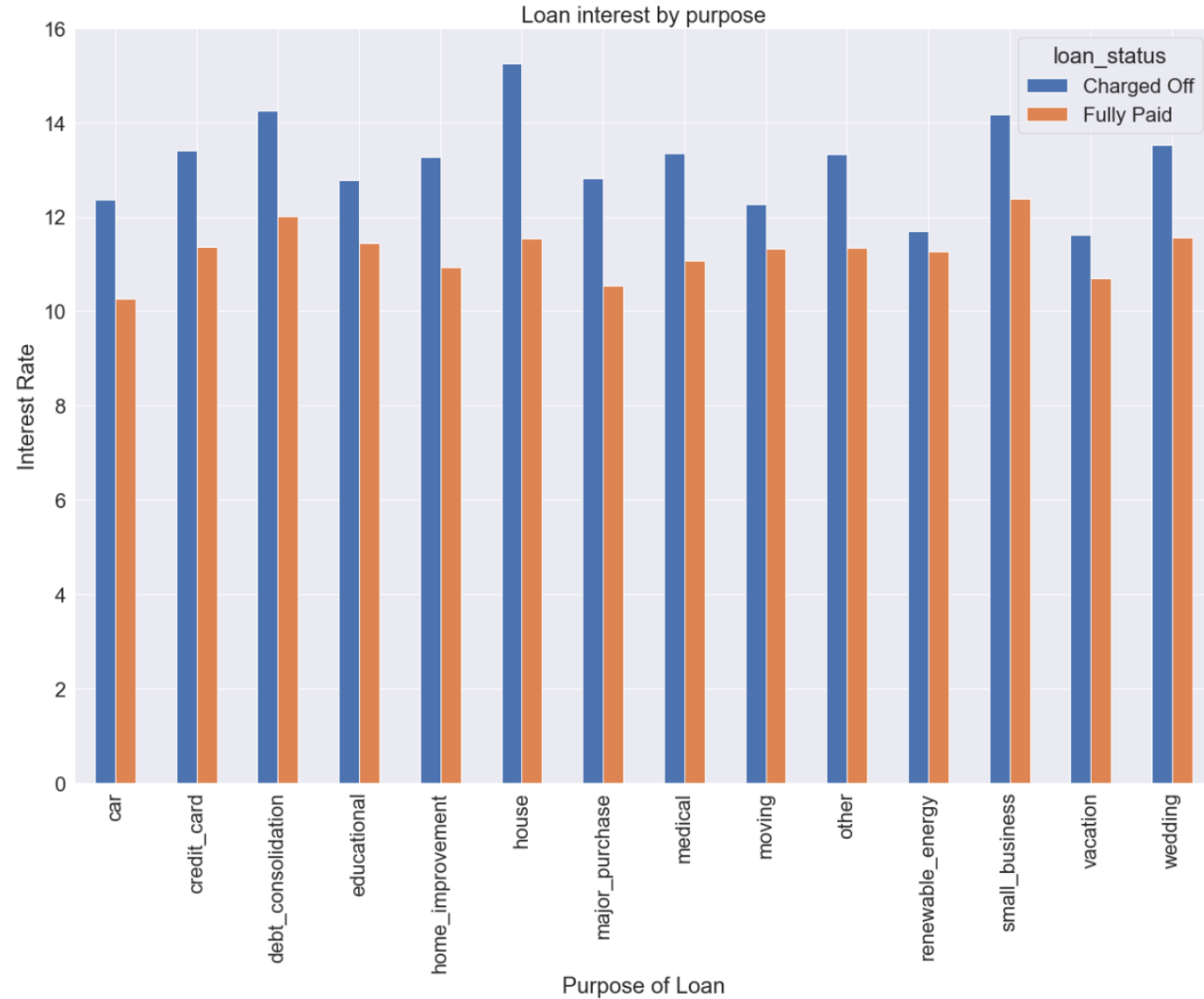
# Customer count by Credit line opening year

The number of customers that are charged off are significantly higher than the fully paid customers whose credit line opening year is between 1990's to early 2000's.

Customers Credit Line Opening Year

Customer distribution by state

Percentage of loan defaulters is very high in Neberaska (NE).

States like Alaska(AK),Florida (FL), Nevada(NV), South Dakota(SD) have a higher percent of charged off customers.

Percent of Customers by state
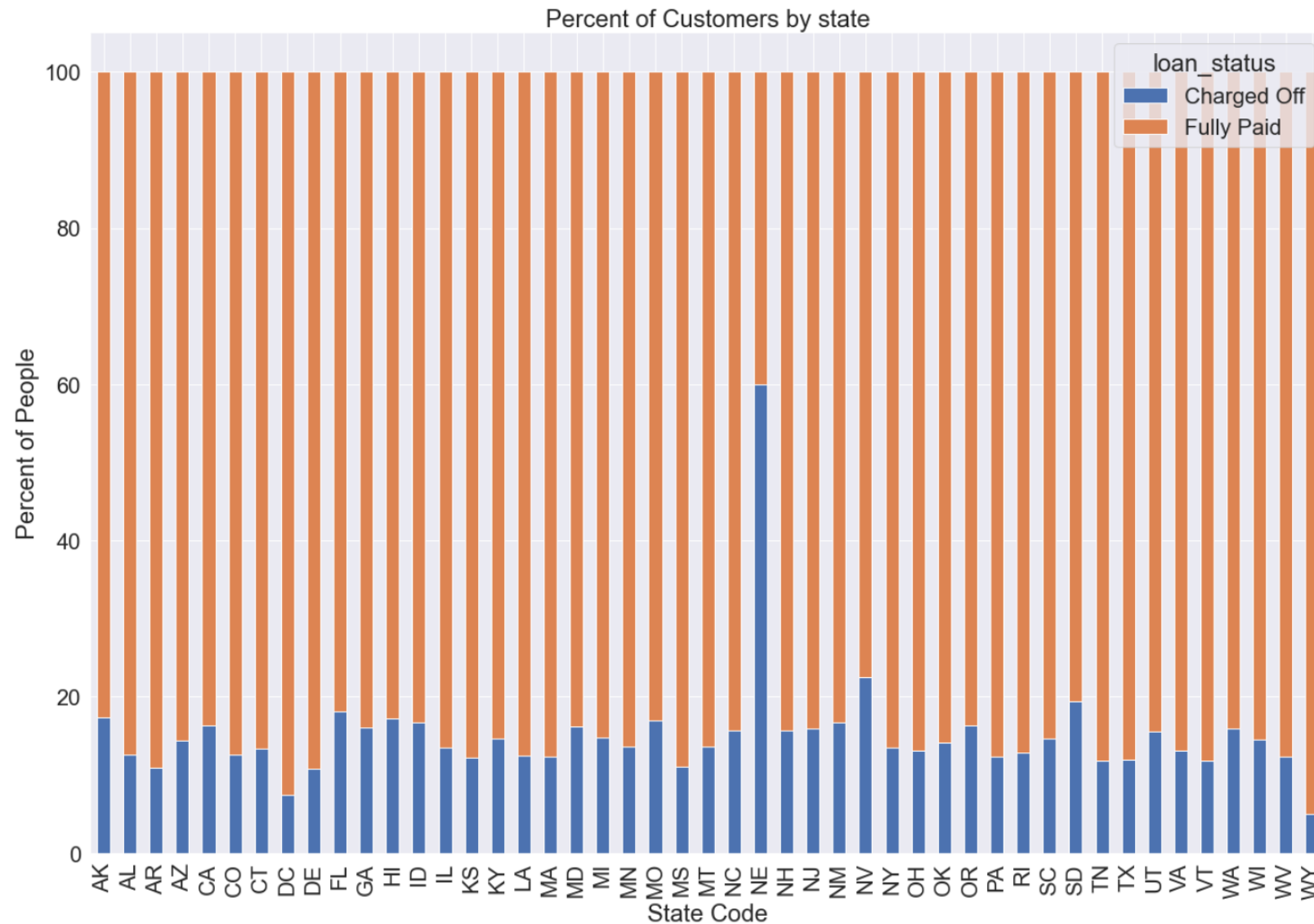
Interest on Home Ownership vs Purpose

Charged off customers in the purpose of house have highest interest rate followed by debt_consolidation and small business.

Again, Interest of charges off customers is higher than the fully paid customers in all categories

home_ownership	MORTGAGE		OWN		RENT	
loan_status	Charged Off	Fully Paid	Charged Off	Fully Paid	Charged Off	Fully Paid
purpose						
car	11.440857	9.567391	13.032000	10.482214	13.118649	11.054030
credit_card	13.416741	11.190741	14.264595	10.997959	13.261835	11.565919
debt_consolidation	14.228533	11.847061	14.220725	11.777846	14.262865	12.183286
educational	12.810909	11.014667	13.140000	11.226000	12.712895	11.656471
home_improvement	13.287331	10.816875	13.873000	11.043750	12.484474	11.687256
house	15.121765	11.539725	16.756667	12.546429	15.181944	11.363684
major_purchase	12.328356	9.775214	12.614615	10.404179	13.164202	11.279242
medical	13.611220	10.755339	10.830000	10.915472	13.614630	11.364093
moving	12.748750	10.005876	13.780000	11.169524	12.092000	11.691831
other	13.115143	10.691658	12.294407	11.272591	13.584974	11.812266
renewable_energy	12.071111	10.289091	6.540000	12.725556	11.922500	12.484000
small_business	13.999756	12.218309	14.909412	11.687463	14.212811	12.665328
vacation	12.509231	10.090342	11.945000	9.298077	11.248333	11.318150
wedding	12.937879	11.606059	10.130000	11.066122	14.026552	11.585878

Key Takeaways

• It is consistently seen that Charged off customers have higher interest rate than Fully paid customers.

• Number of Charges of customers is higher in Grade E, F, G, (especially F5 & G4), and the loan of such customers should be approved with caution & background check.

• Through checking must be done for customers, with dti more than 12 applying for vacation, other loan & home improvement, and customers with dti more than 11 applying for home, major purchases & small business.

• Cautious lone approval should be done for customers with creditline opening year after 1987, (especially from 1995 to 2005).

• Loan should be cautiously approved for customers from Nebraska (NE), Alaska(AK),Florida (FL), Nevada(NV), South Dakota(SD) .

• It is recommended that the Loan interest allotment system be re graded as interest of charged off customers is higher in every category.

Thank You !