

LENDING CLUB CASE STUDY

Hemanth
Pratik

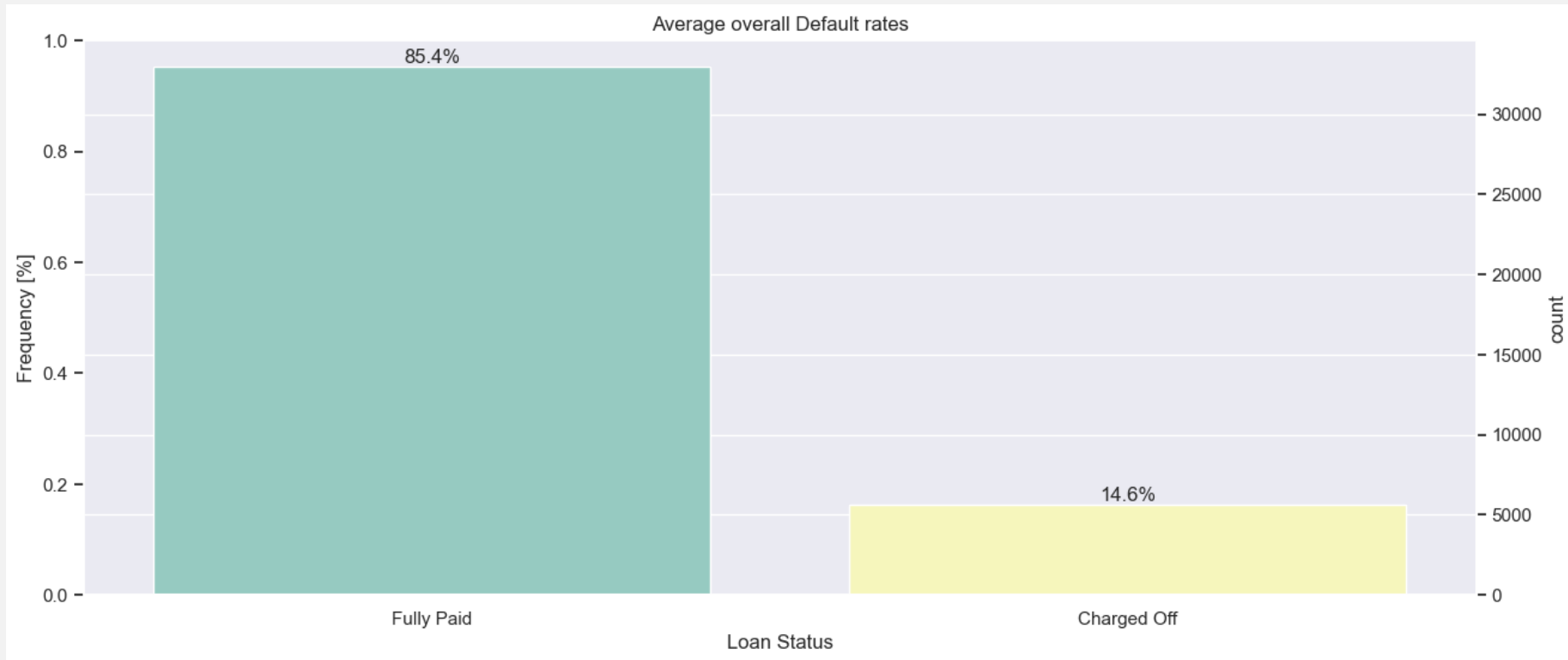
PROBLEM STATEMENT

- Given a banking loan related data, to find out the prominent factors which can be used to identify and pinpoint the people which are likely to default on their respective loans.

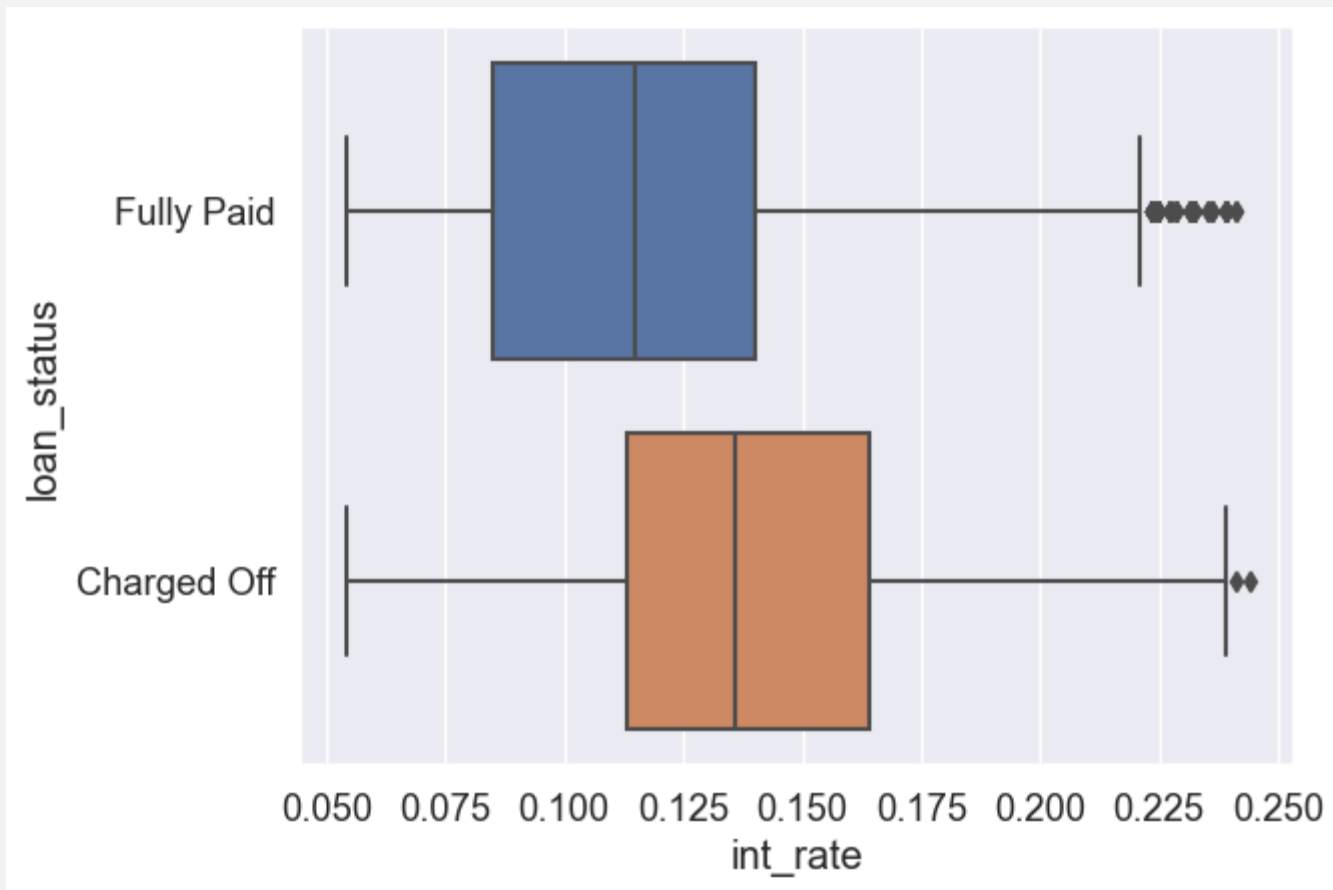
DATA DESCRIPTION

- 111 columns and ~36000 rows of data containing all the loan related information
- Broadly split into 3 bins:
 - 1. Customer Details: Emp Title, Annual income, Verification status...
 - 2. Loan Details: Instalment, Interest Rate, Loan Status, Grade...
 - 3. Customer Behavioural Metrics: Open accounts, Public Records, Instalments in last 6 months...

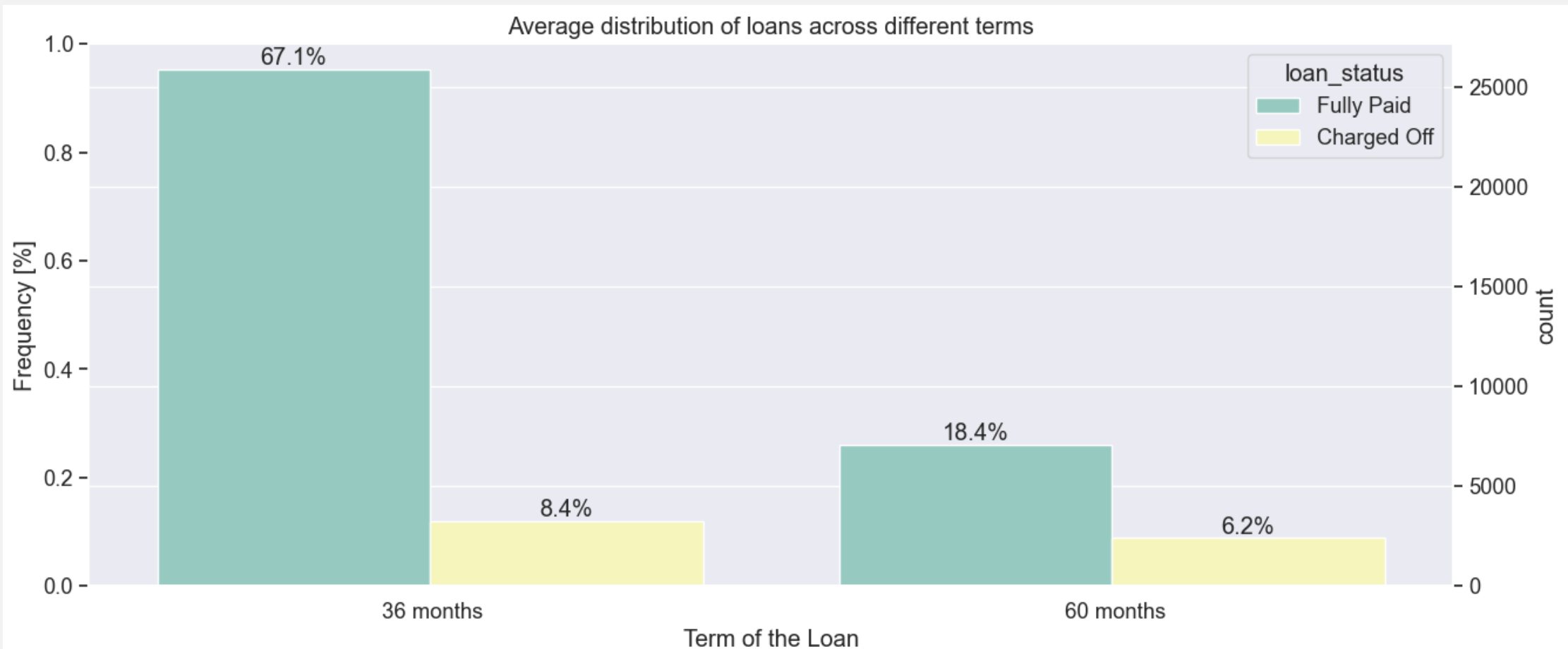
OVERALL DEFAULT RATE – 14.6%



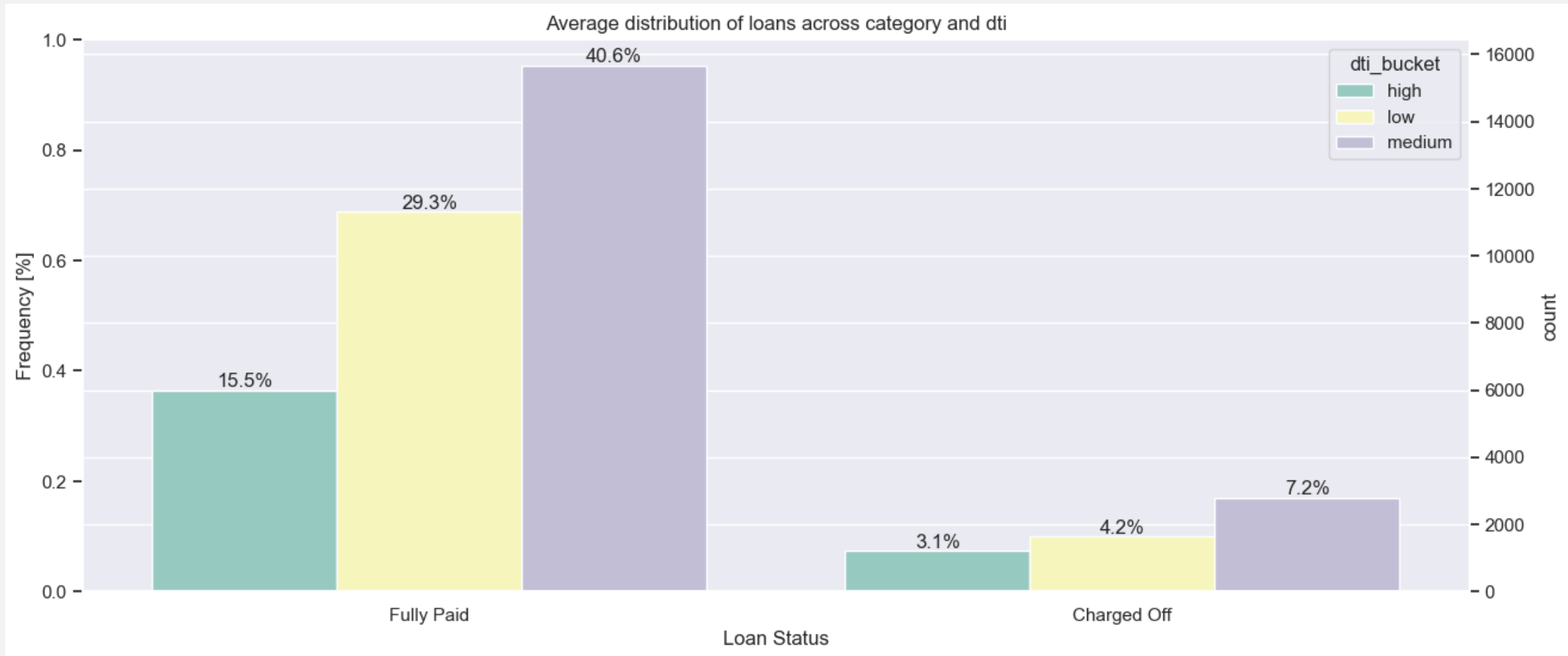
PEOPLE WITH HIGH INTEREST RATE LOANS LIKELY TO DEFAULT



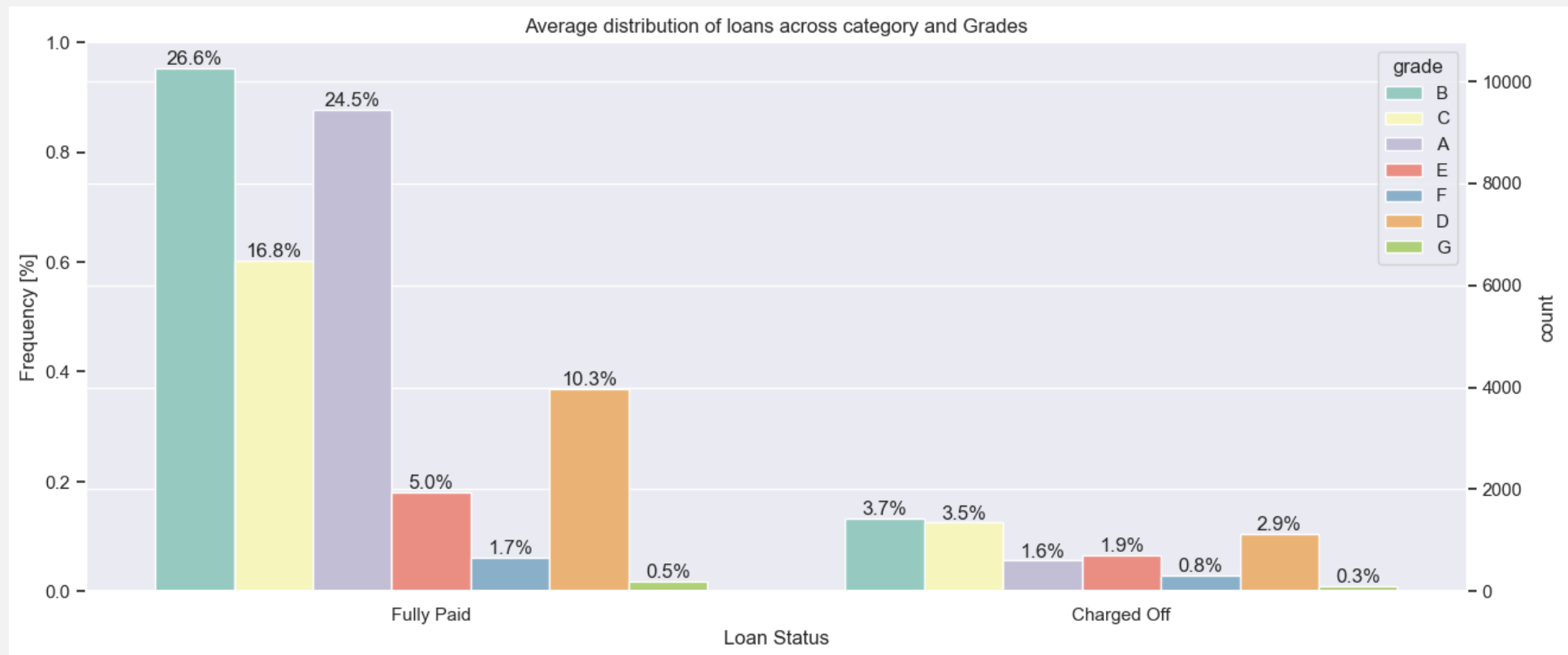
PEOPLE WITH A LOAN TERM OF 60 MONTHS ARE LIKELY TO DEFAULT



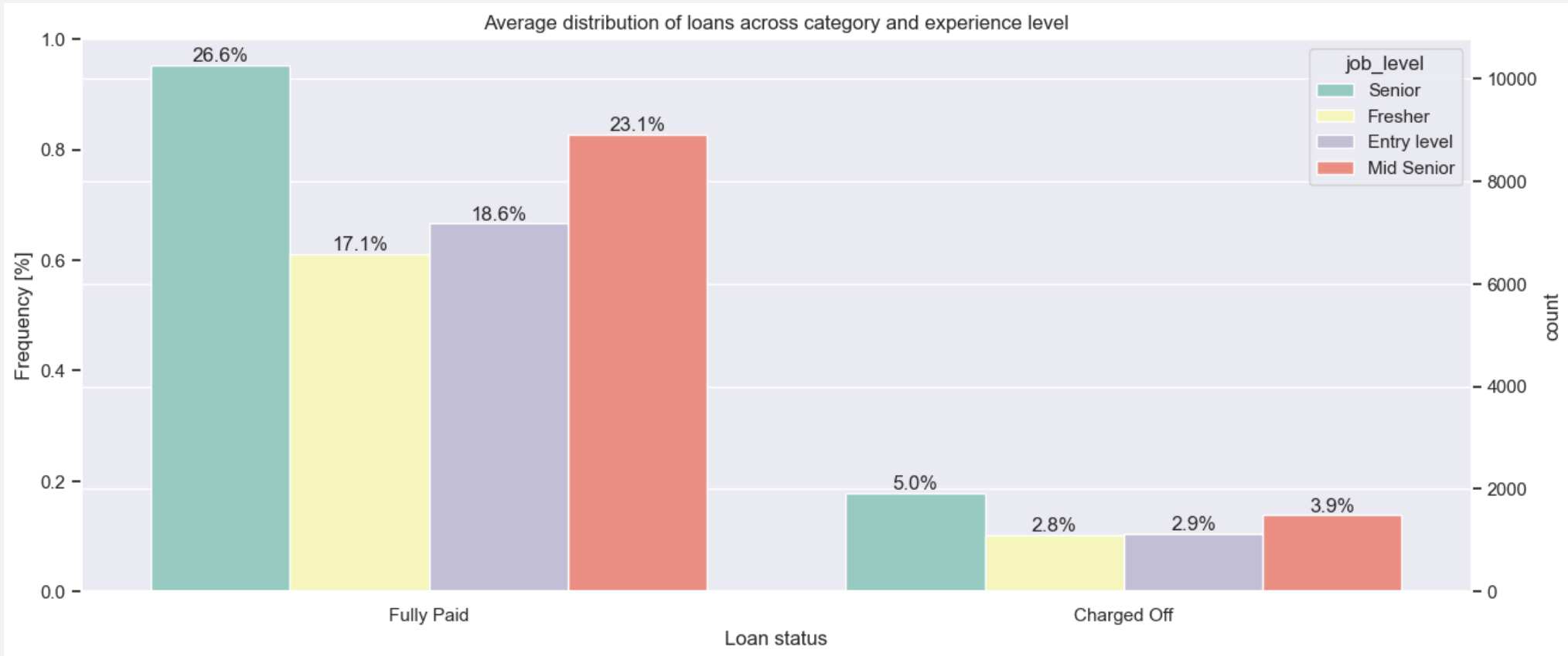
PEOPLE WITH MEDIUM DTI RATIO ARE LIKELY TO DEFAULT



PEOPLE WITH GRADE B AND GRADE C LOANS ARE LIKELY TO DEFAULT



PEOPLE WITH EXPERIENCE LEVEL OF MORE THAN 7 YEARS ARE LIKELY TO DEFAULT



PEOPLE WHO AVAIL LOANS FOR SMALL BUSINESSES ARE LIKELY TO DEFAULT

