UpGrad Submission

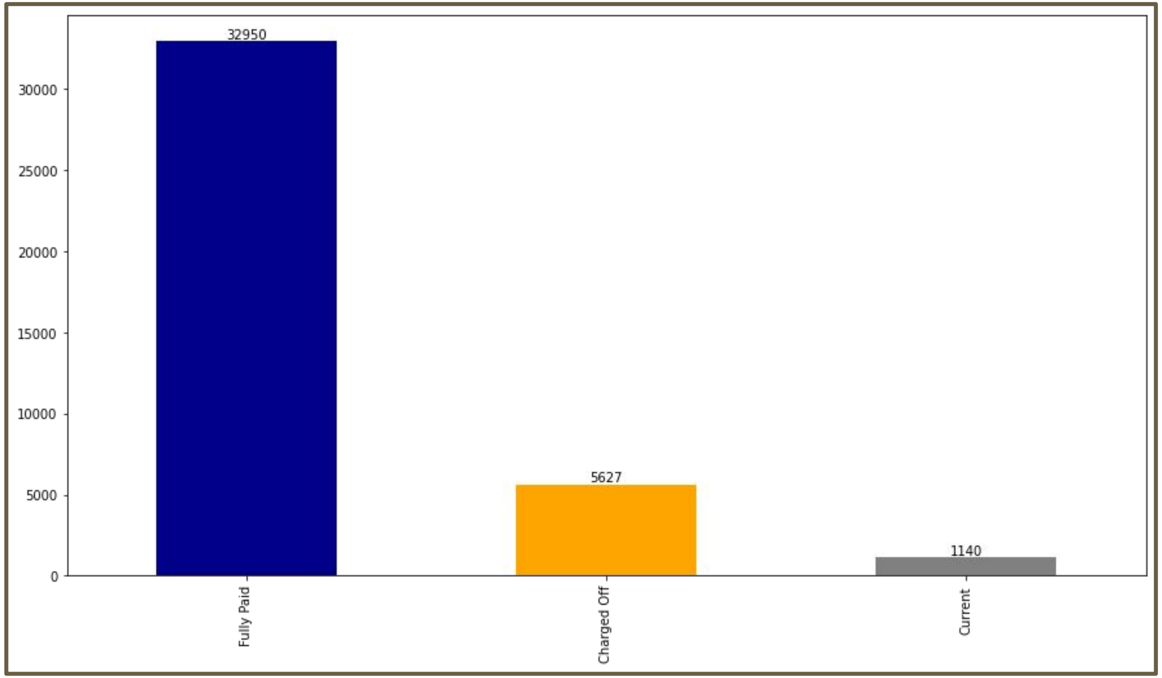
Lending Case Study

Problem Overview

- Based on the lending data of company, we analyze and exploit all factors that have strong relationship with the status of loans.
- There are 3 main types of loan status
    - Fully Paid: Non-default (Good label)
    - Charged Off: Default (Bad label)
    - Current: N/A

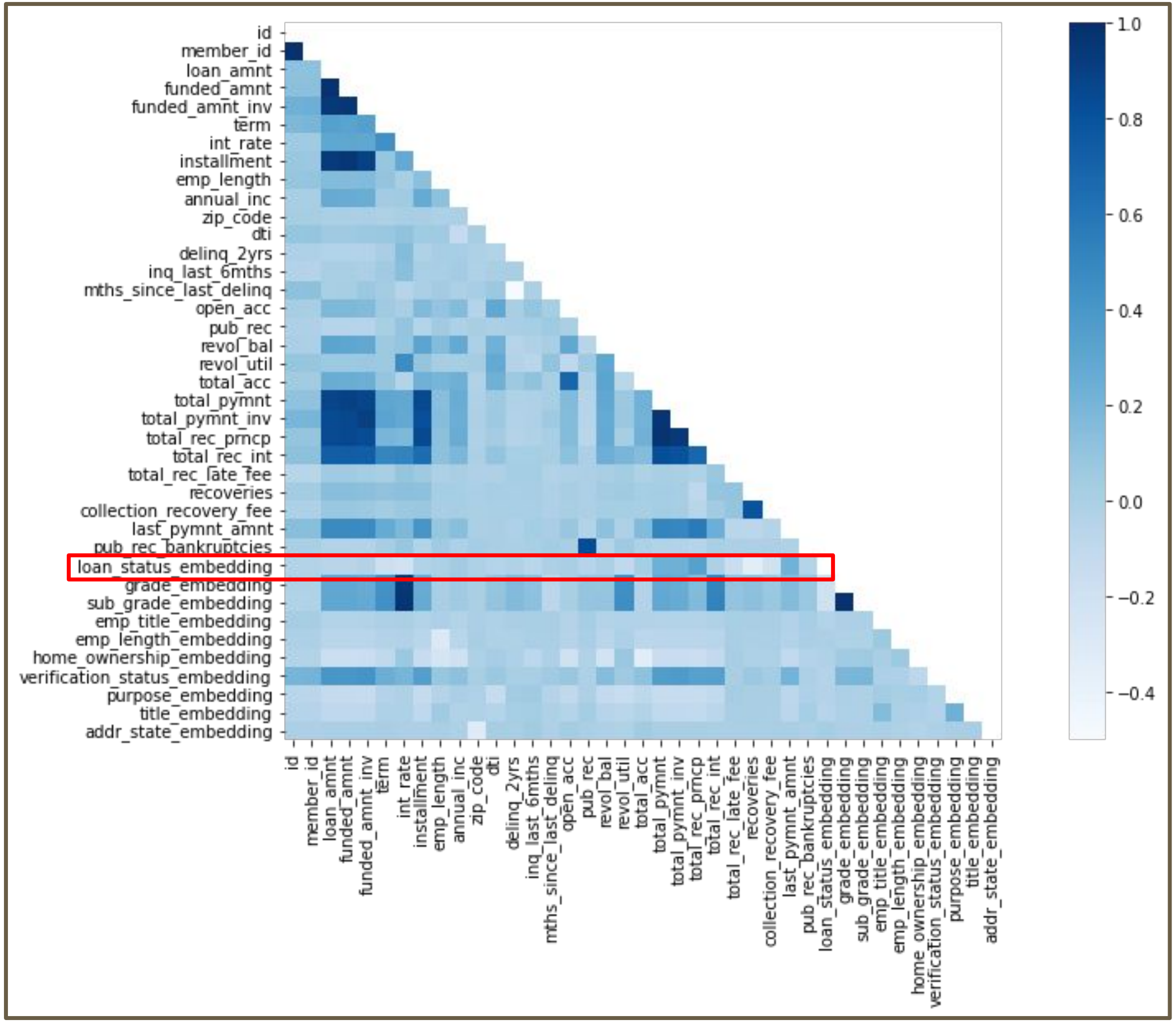
Dataset Overview

- Datashape:
  - Total Records: 39717 records
  - Total Columns: 111 columns
- Loan Status:
  - Fully Paid: 32950 records
  - Charged Off: 5627 records
  - Current: 1140 records



Correlation Map

- Most of features do not have high correlated with loan status column
- Some features have very high correlation

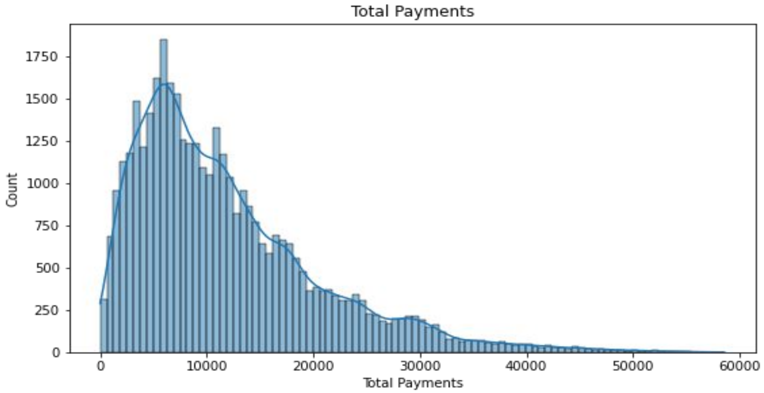


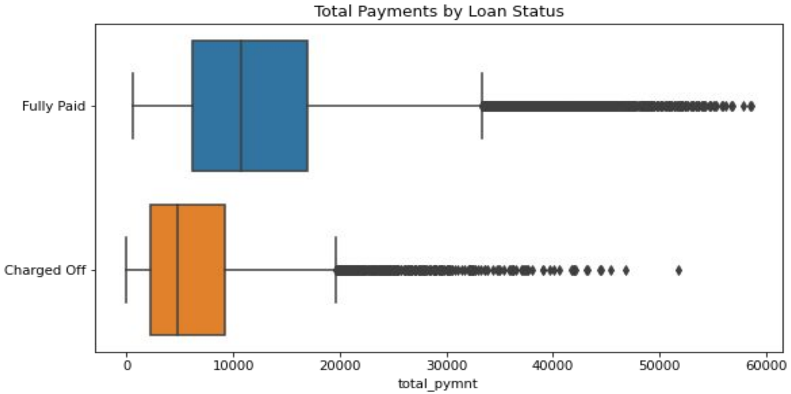
Conclusions

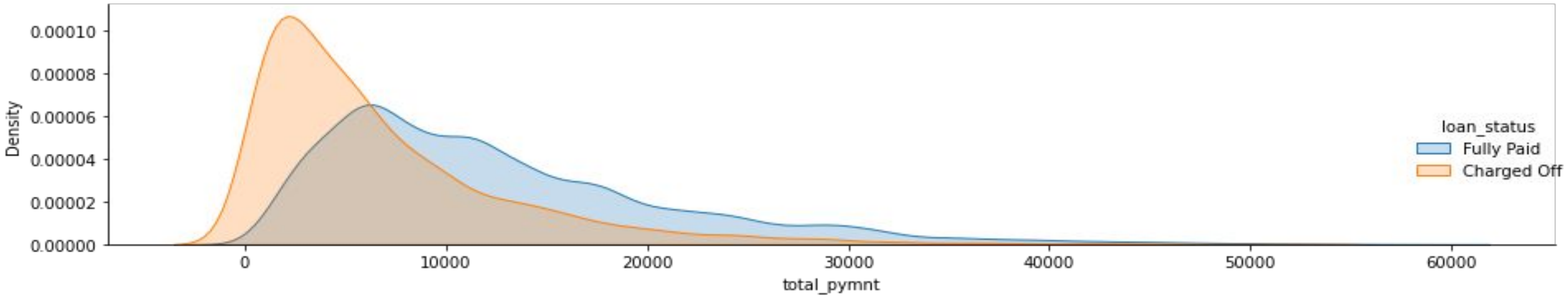
- The **total payment of less than 10000$,** which has **high default rate**, might be considered as the threshold for business to use in their rules-based approach.
- The company should develop some specific rules that target just **higher grades (A->C) in order to promote sales** or focus on **lower grades (D->G) in order to decrease risks**
- **Interest rate above 12%** is the point at which businesses should increase their efforts to evaluate and assess the possible **risk of lending cases** in order to reduce the bad rates.
- **<5000$ in Principal Received to Date** is the package that **75% charged off cases** can not attain
- Most of lending cases have **Lending Purposes** orginated from **Debt, Credit Card, Car and Home Improvement.**

Total Payments

- Most default loans have total payment of less than 10000$ (the highest is around 5000$), which might be used as a benchmark for businesses to use in their rules-based procedure.
- When the 75th percentile of Charged Off cases is less than $10,000, the median total payment of Fully Paid cases is greater than $10,000.

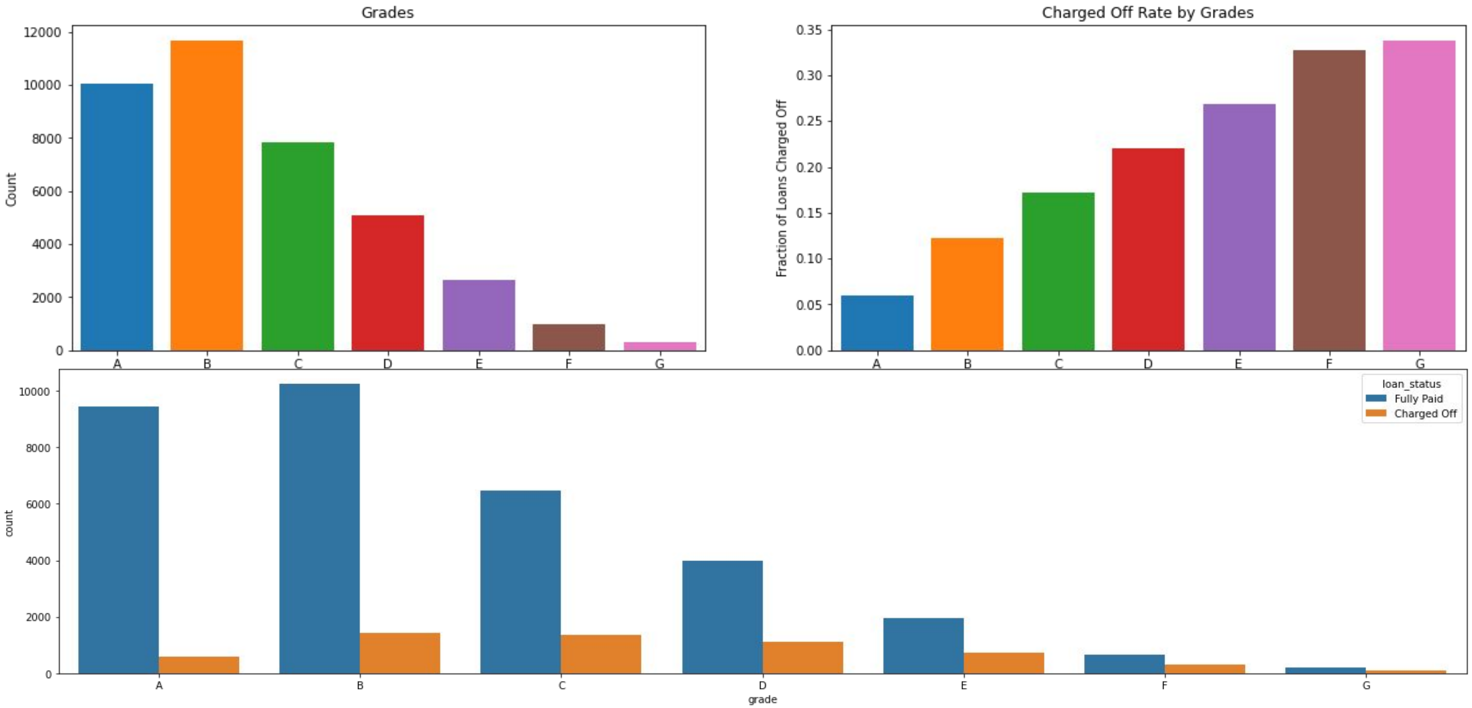






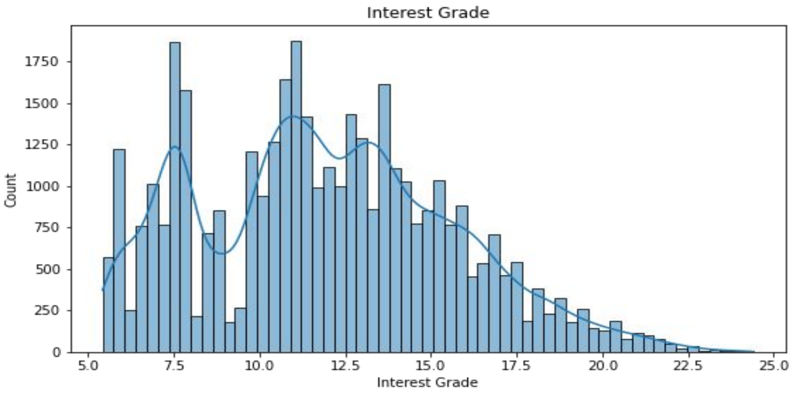
Grade

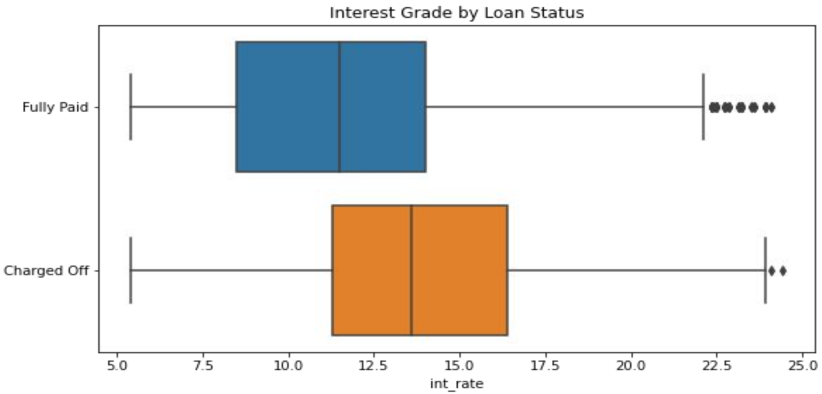
- The lower the grade, the higher the rate at which loan instances are charged off.
- The company can establish a rule that only targets better grades (A->C) to enhance sales or focus on lower grades (D->G) to decrease risks.
- The failure rate of Grade D->G is more than 20%.

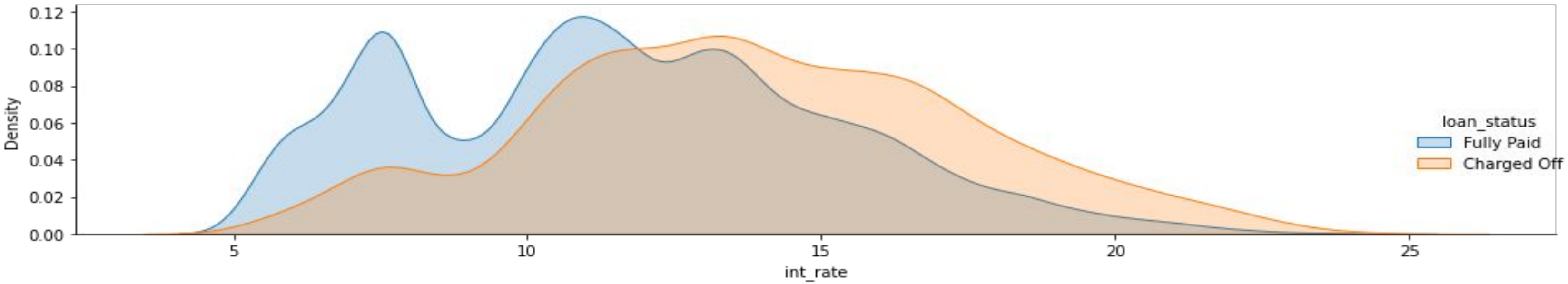


Interest Rate

- The greater the interest rate, the higher the bad loan rate. There are many examples that have been charged off, especially when the interest rate is near 12%.
- Perhaps 12% is where businesses should focus their efforts to assess and analyze the possible risk of lending cases in order to decrease bad rates.

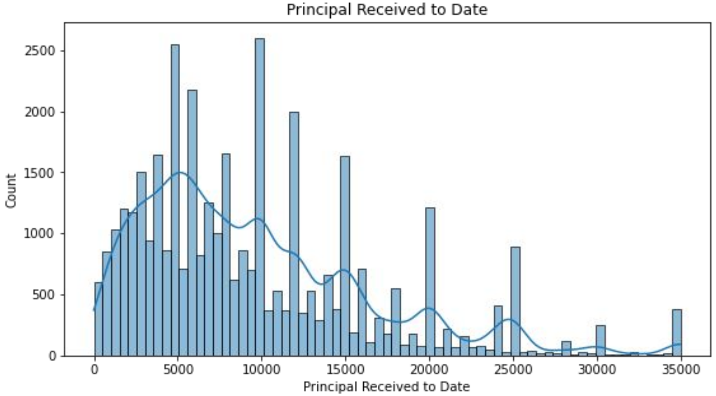


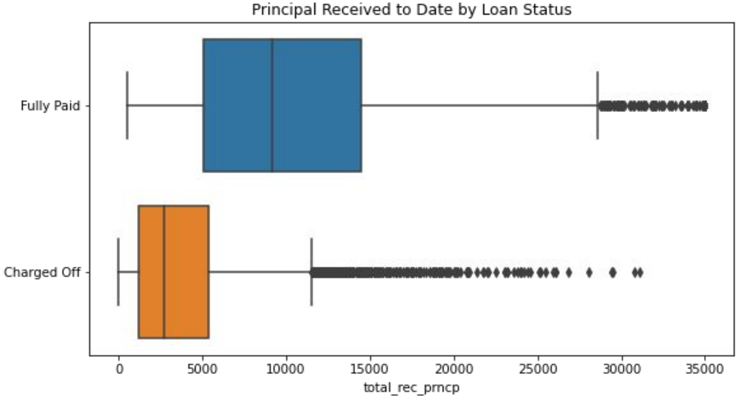


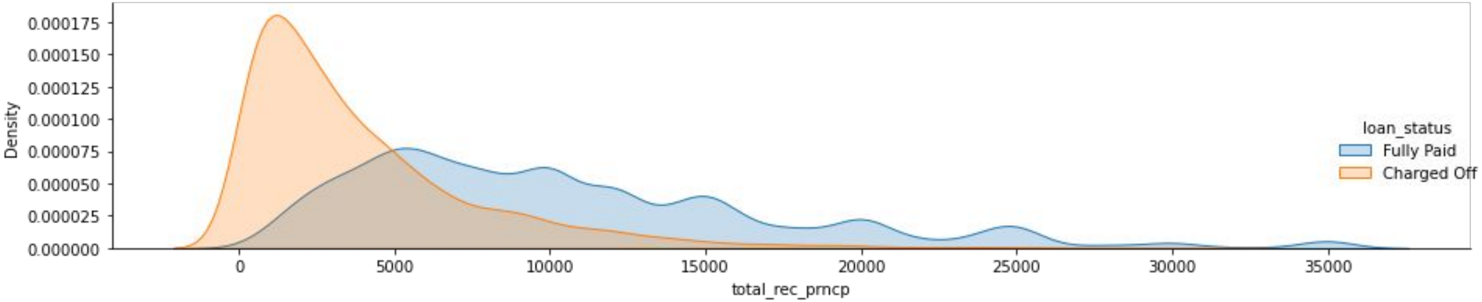


Principal Received to Date

- The highest peak of charged-off loan situations is $4000.
- The principle is usually rounded up to 5000, 10,000, or 15000 dollars. —> Perhaps it's the business default loan package.
- 5000$ is the bundle that is 75% charged off but cannot be reached.

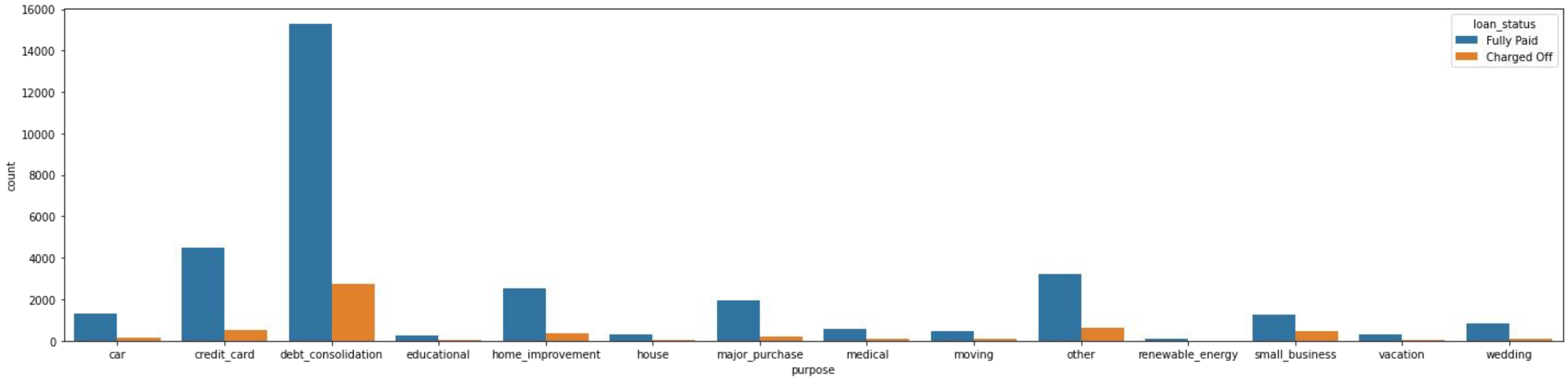






Lending Purpose

- Debt, credit card, car, and home improvement loans account for the majority of our lending instances.
- We might apply greater domain knowledge of heuristics to this industry and client behaviour.



Conclusions

- The **total payment of less than 10000$,** which has **high default rate**, might be considered as the threshold for business to use in their rules-based approach.
- The company should develop some specific rules that target just **higher grades (A->C) in order to promote sales** or focus on **lower grades (D->G) in order to decrease risks**
- **Interest rate above 12%** is the point at which businesses should increase their efforts to evaluate and assess the possible **risk of lending cases** in order to reduce the bad rates.
- **<5000$ in Principal Received to Date** is the package that **75% charged off cases** can not attain
- Most of lending cases have **Lending Purposes** orginated from **Debt, Credit Card, Car and Home Improvement.**