

# **upGrad and IIITB's ACPDL: November 2021**

**LENDING CLUB CASE STUDY**

# Team

## **Group Facilitator :**

**Name:** Nitish Pandey

**Email ID:** [nitishpandey2117@gmail.com](mailto:nitishpandey2117@gmail.com)

**Phone No:** 919818659521

## **Team Member Detail:**

**Name:** Nitish Pandey

**Email ID:** [nitishpandey2117@gmail.com](mailto:nitishpandey2117@gmail.com)

**Phone no:** 919818659521

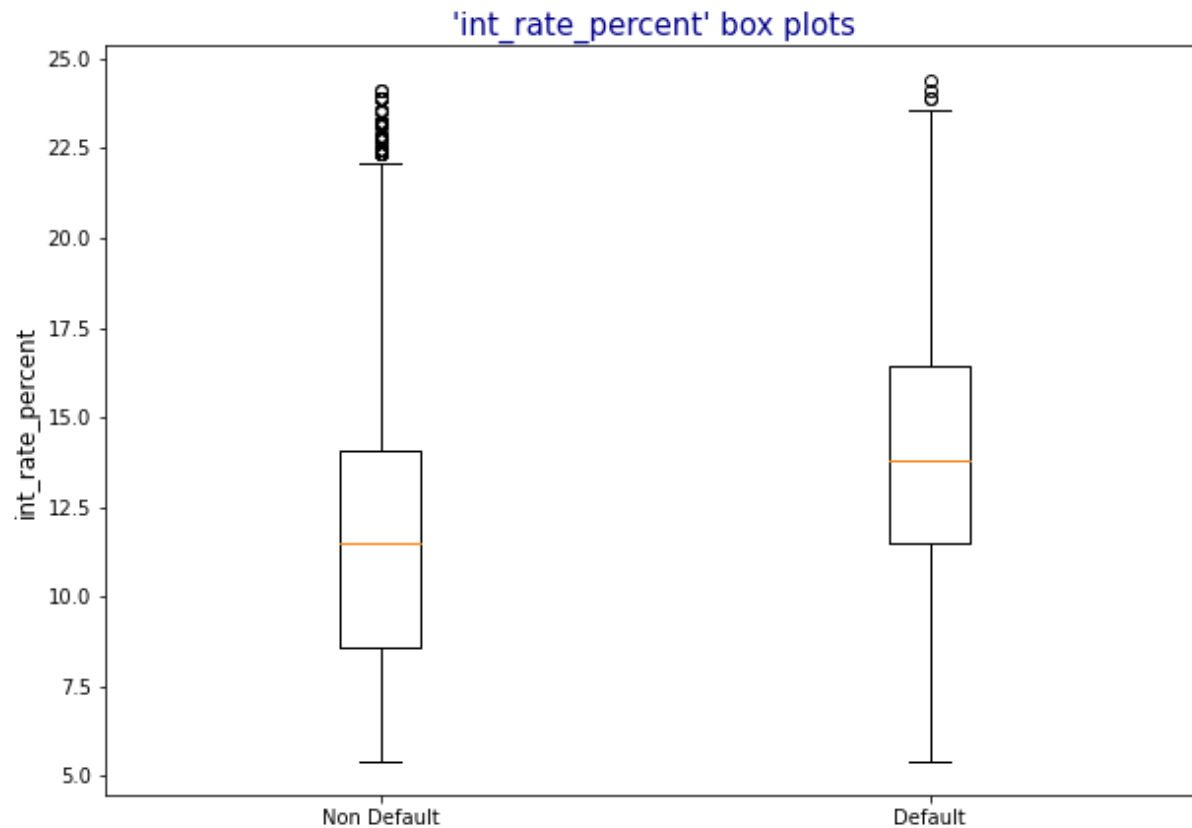
# Problem Statement

- I work for a consumer finance company which provides loans to urban customers
- Borrowers who default (run away with or refuse to pay the money) cause largest amount of loss.
- To prevent credit loss, I have to use EDA on historical loan dataset to analyze which type of customers are more likely to default
- This will help the company to take better decisions

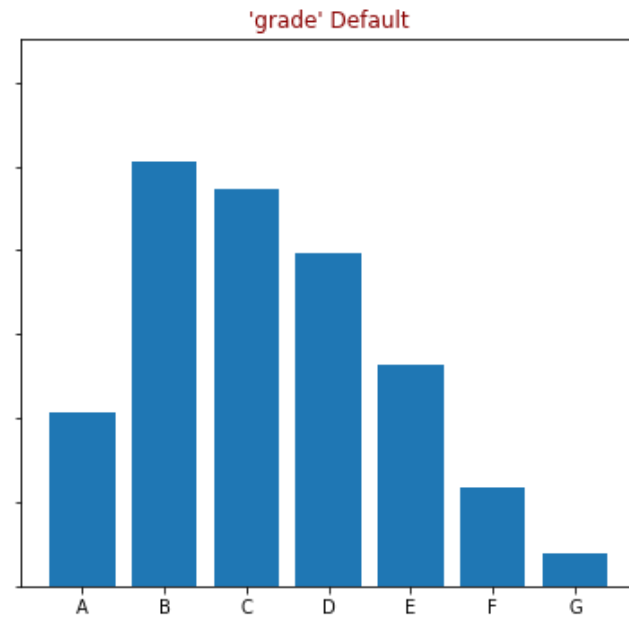
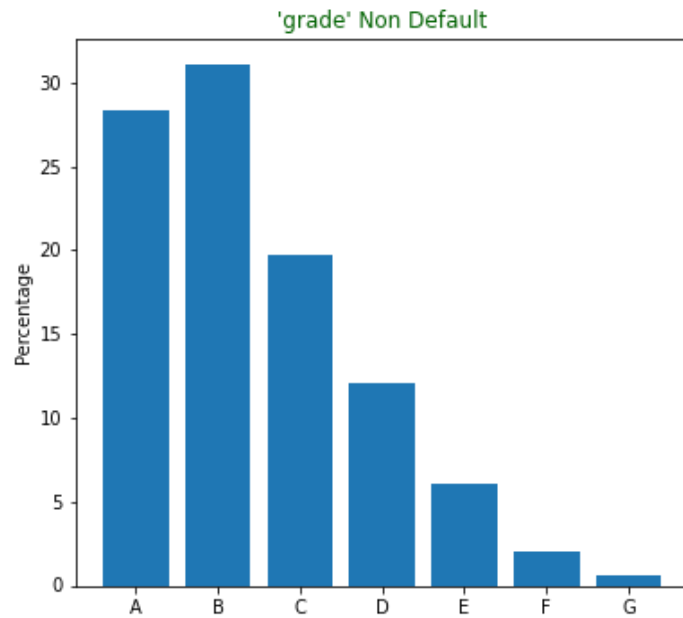
# Exploratory Data Analysis will include:-

- Data Sourcing
- Metadata/Sanity Check
- Data Cleaning
- Univariate Analysis
- Segmented Univariate Analysis
- Bivariate Analysis

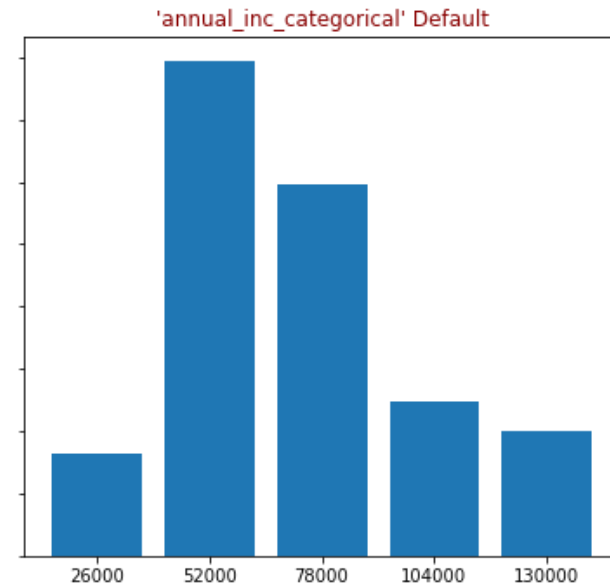
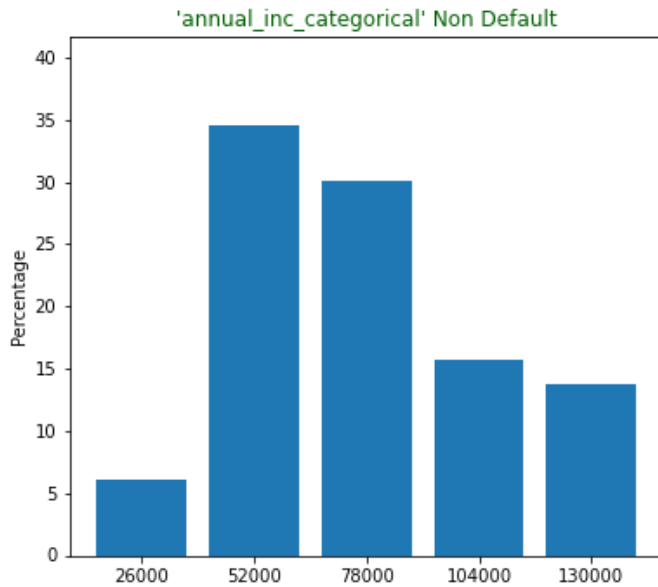
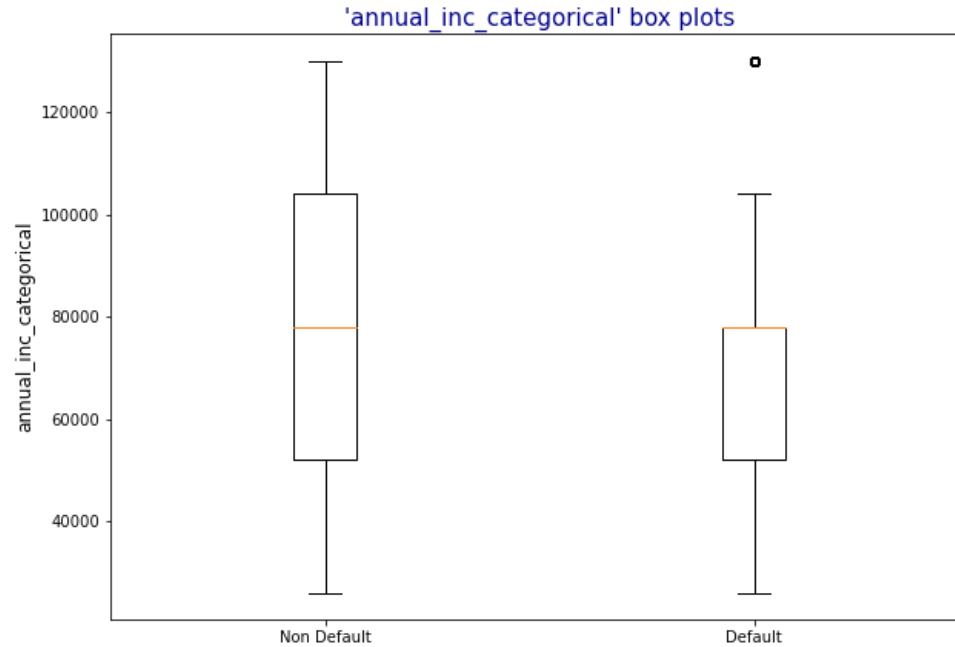
- Interest rate percentage is higher for default as compared to non default customers



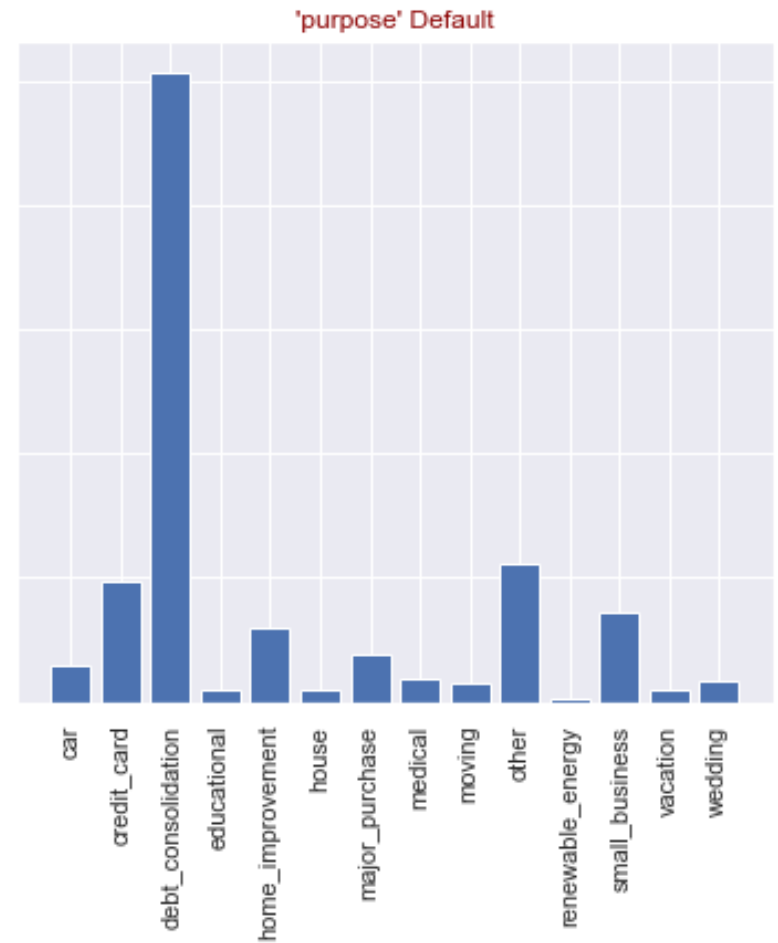
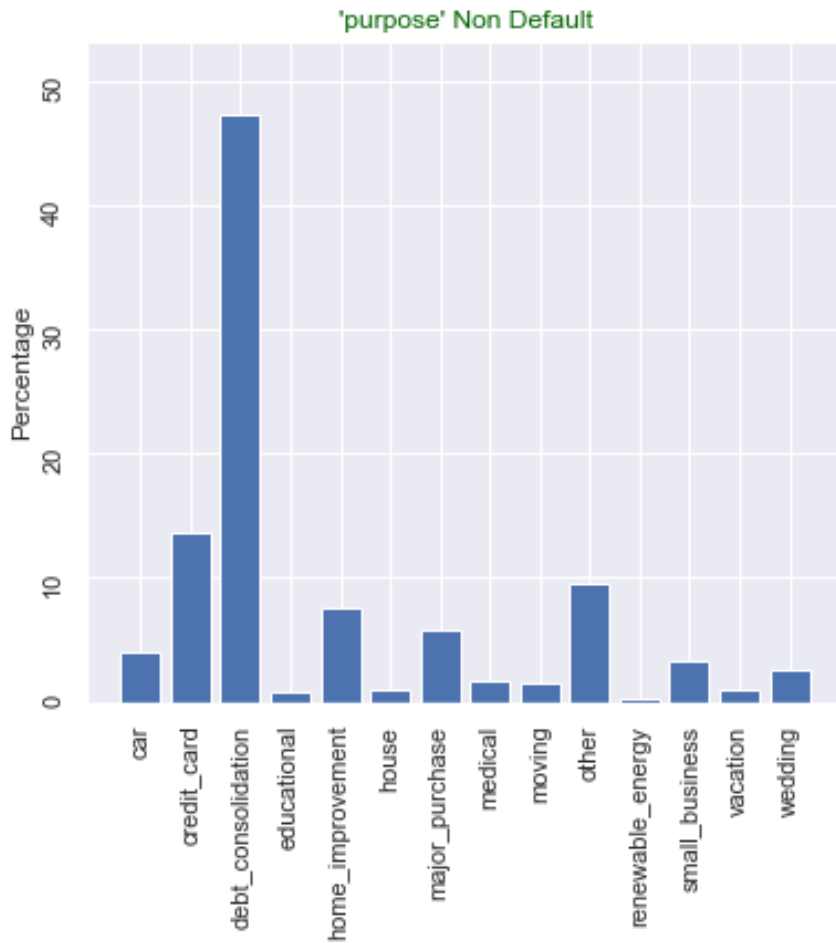
- Default has more percentage of borrowers in higher grades as compared to Non default



- Non default has more percentage of applicants in higher income slabs as compared to Default

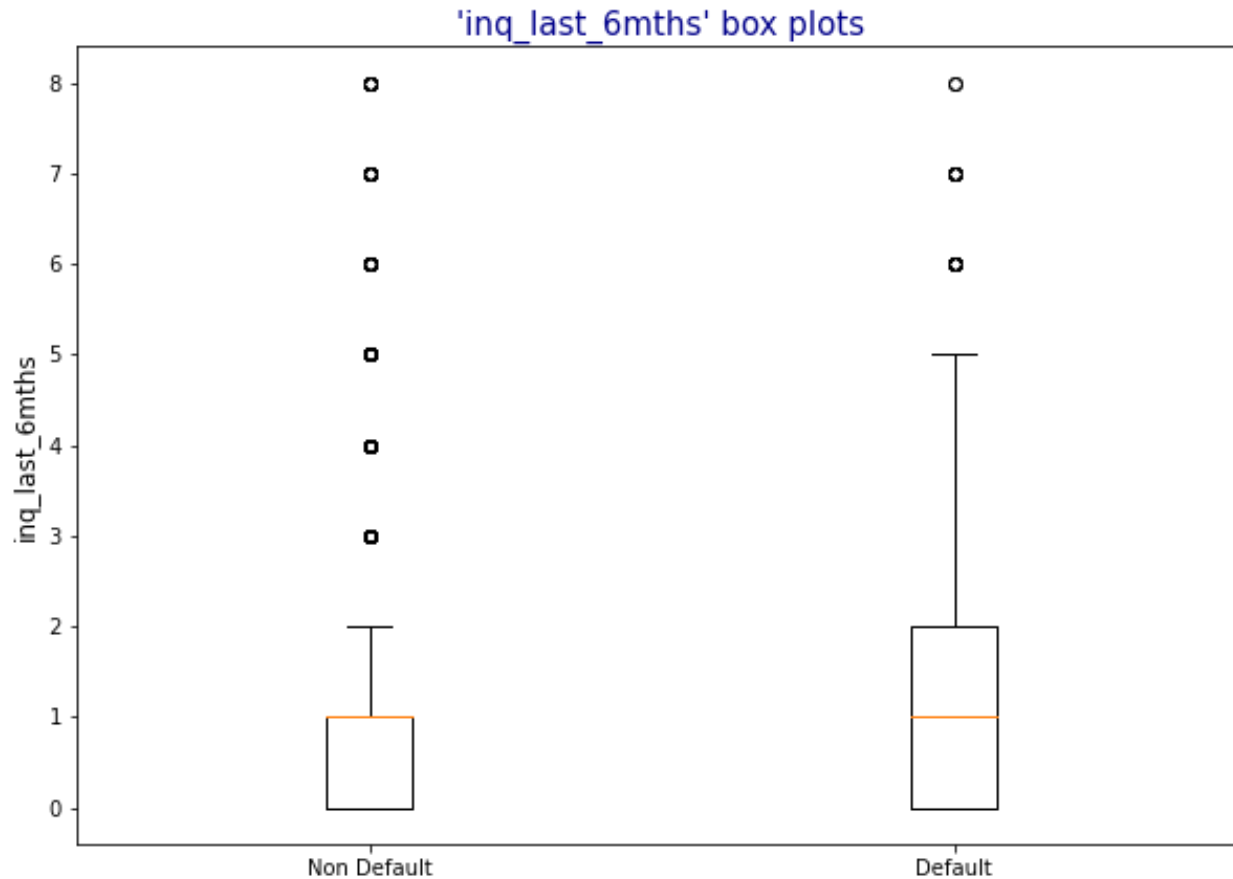


- In purpose column, 'debt\_consolidation', 'credit\_card', 'other', 'small\_business' have higher percentage in default as compared to non default

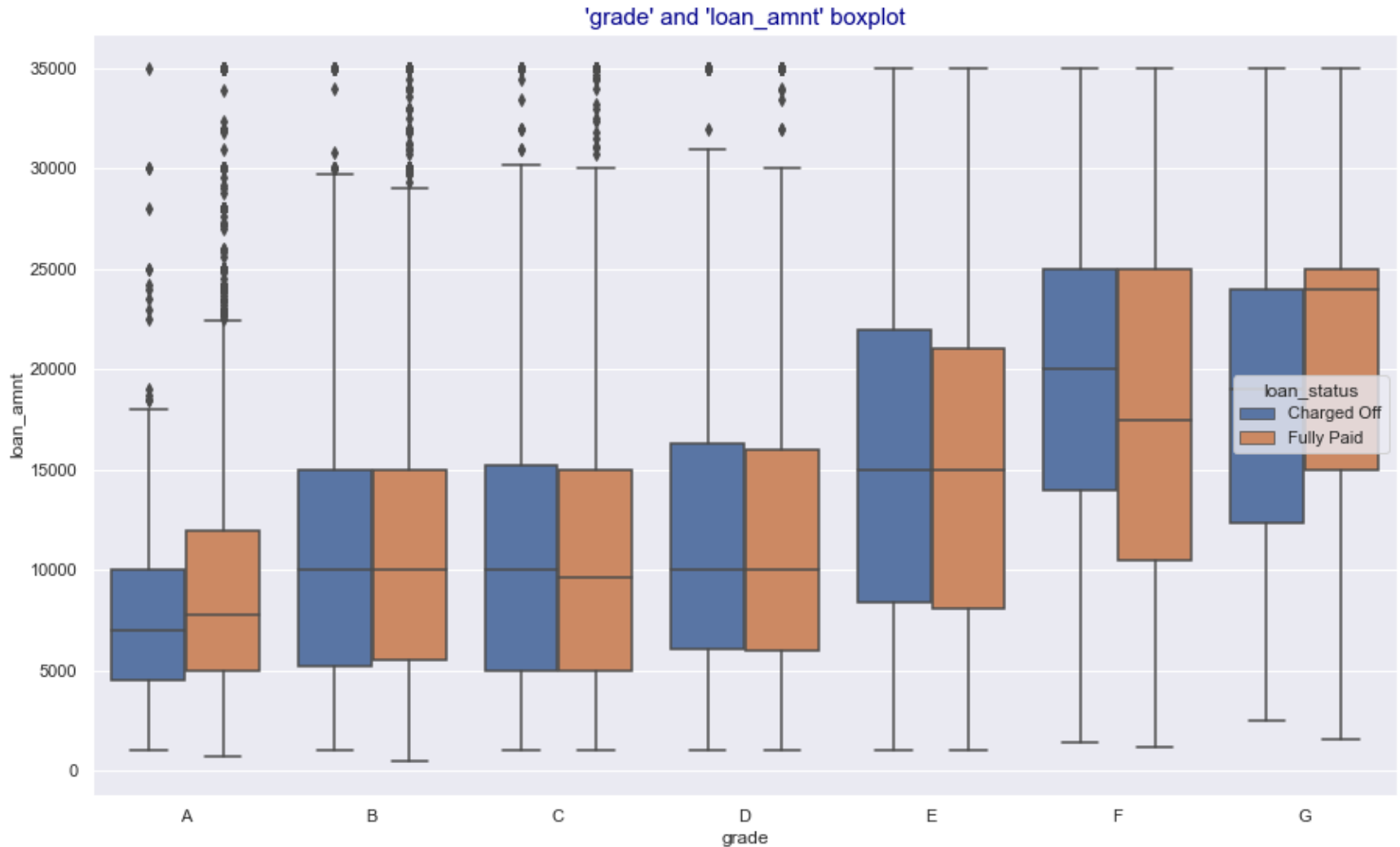




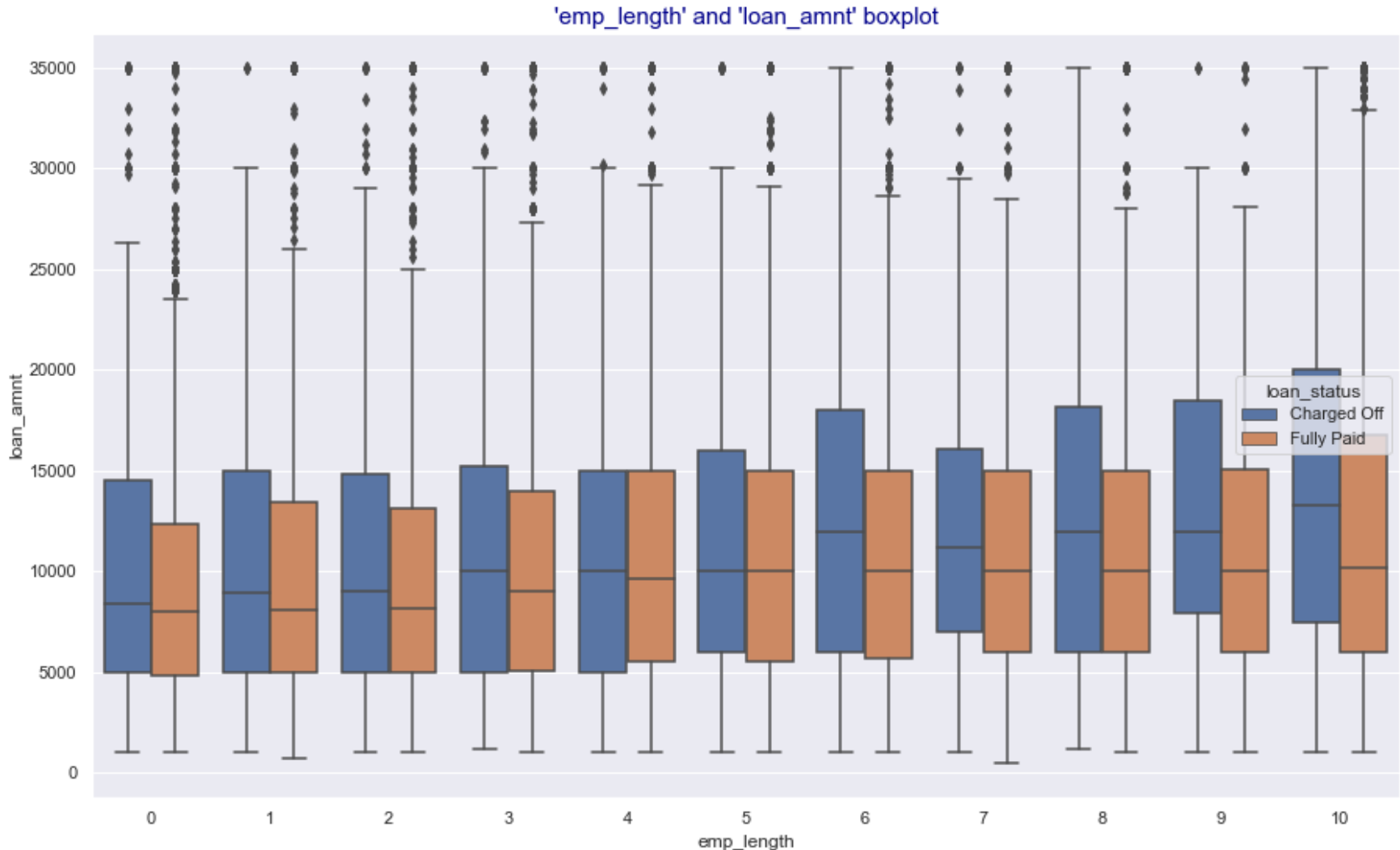
- Defaulters have more number of inquiries in last 6 months of approval as compared to non defaulters



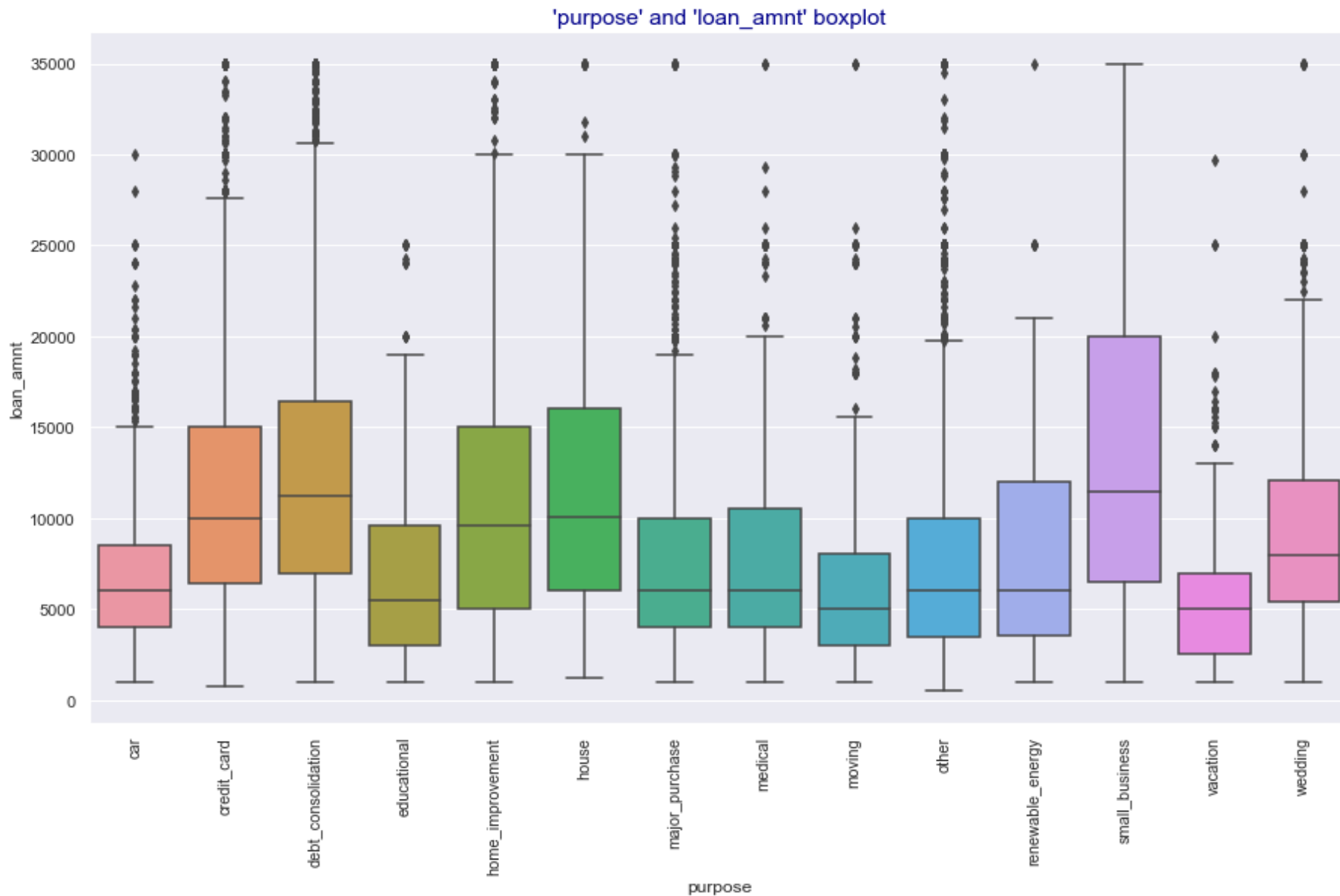
- Higher the grade, more the loan amount and defaulter's loan amount is always greater than or equal to non defaulters



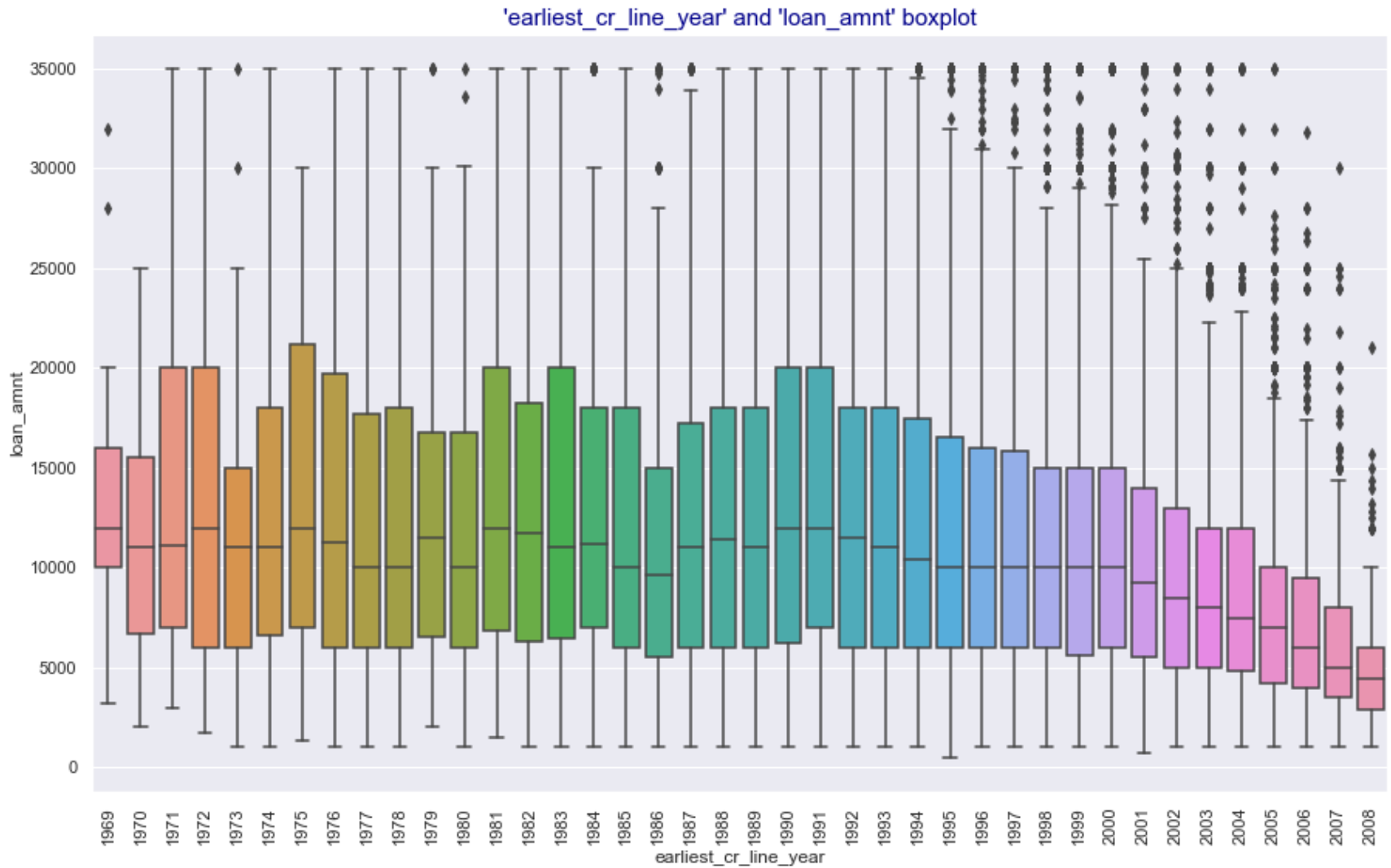
- At every emp\_length value, the loan\_amnt of defaulters is greater than or equal to that of non defaulters



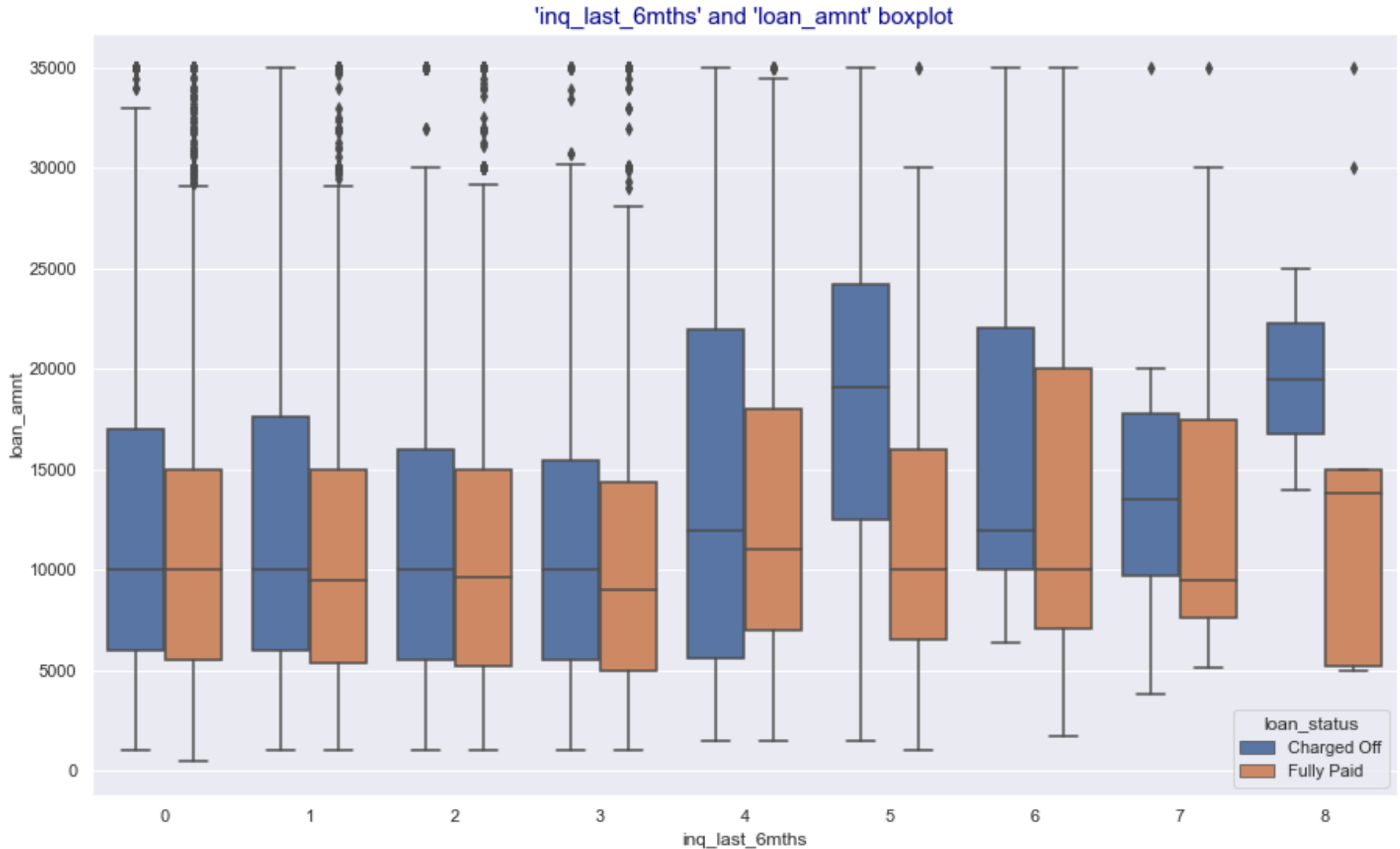
- 'debt\_consolidation' and 'small\_business' purposes account for highest amount of loans on average



- Customers having earliest credit line after 2000 are getting lower amount of loans

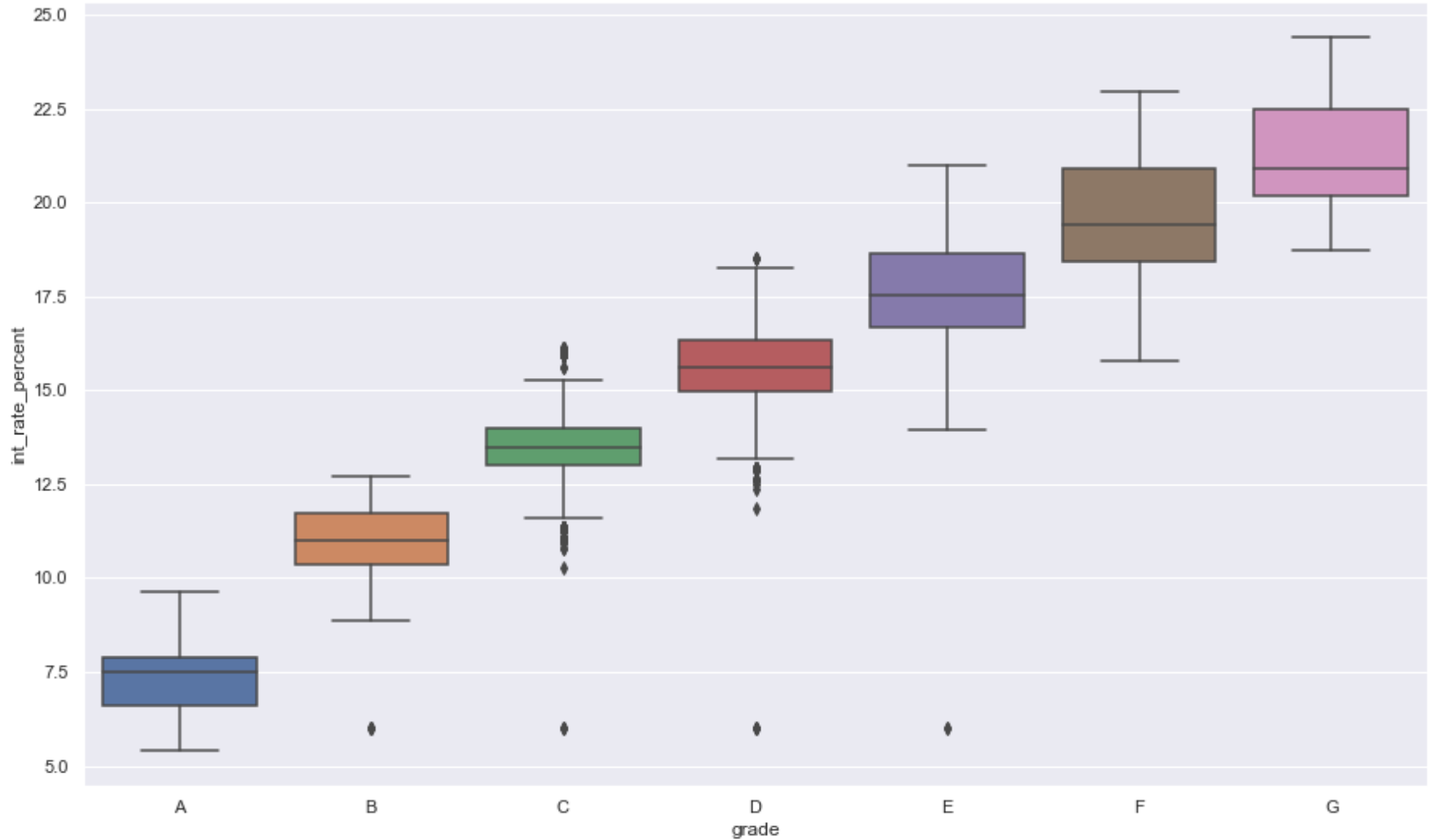


- More the inquiries in last 6 months, more the loan amount given to them, more the inquiries, more the chances of default and the difference is huge at 5 and 8 inquiries

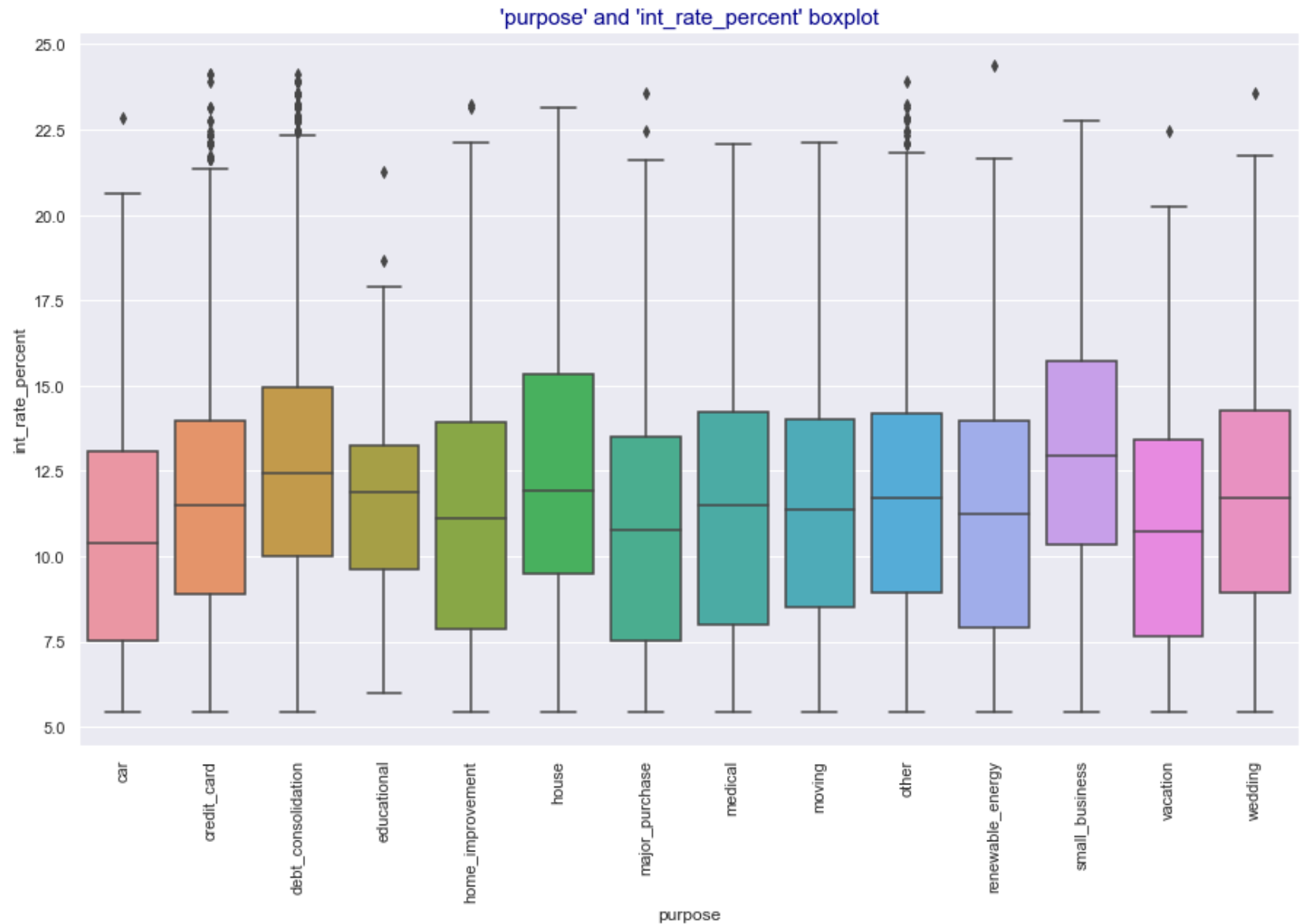


- Higher the grade, more the interest rate of the loan

'grade' and 'int\_rate\_percent' boxplot

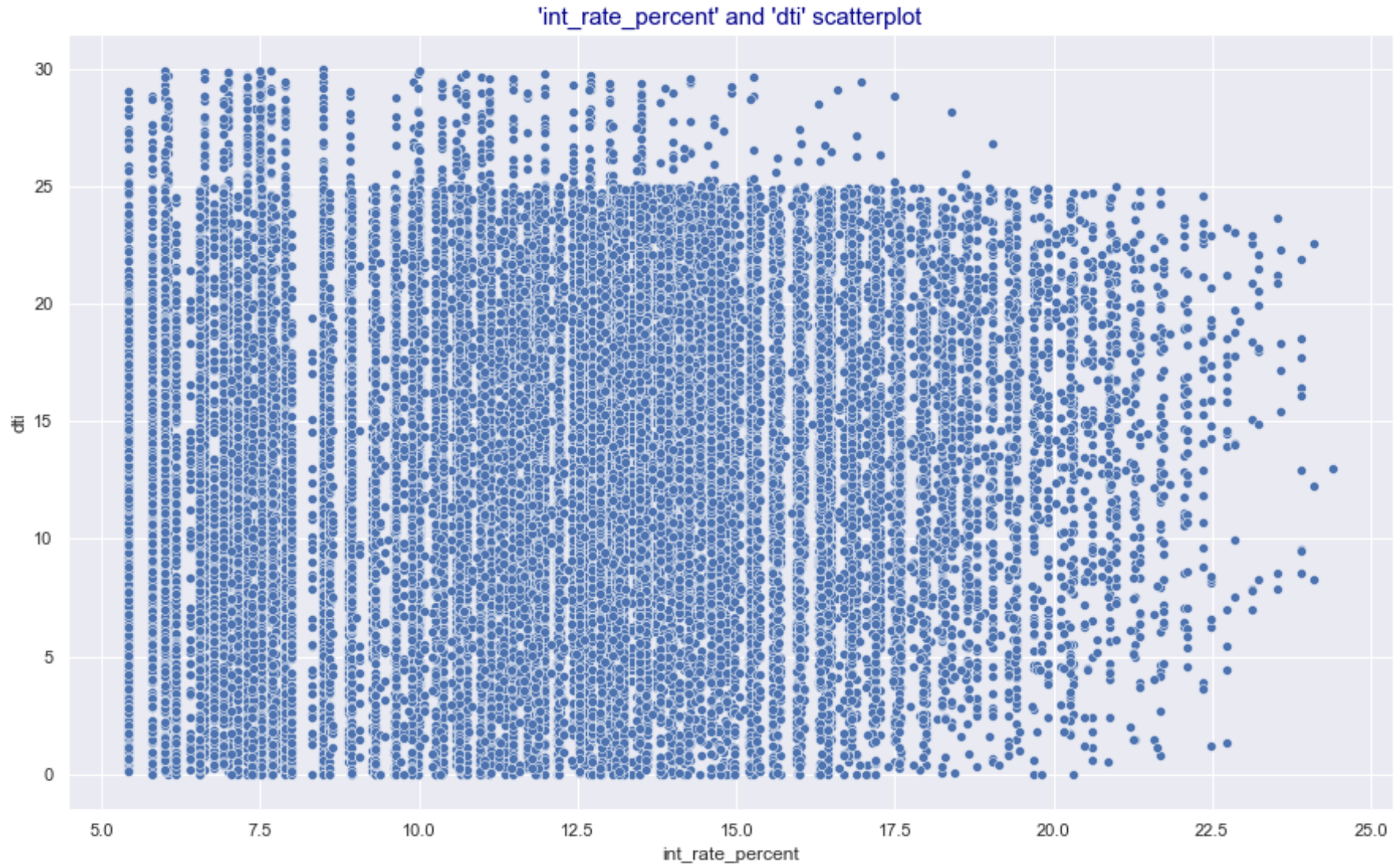


- 'debt\_consolidation', 'house' and 'small\_business' purposes have highest interest rates on average

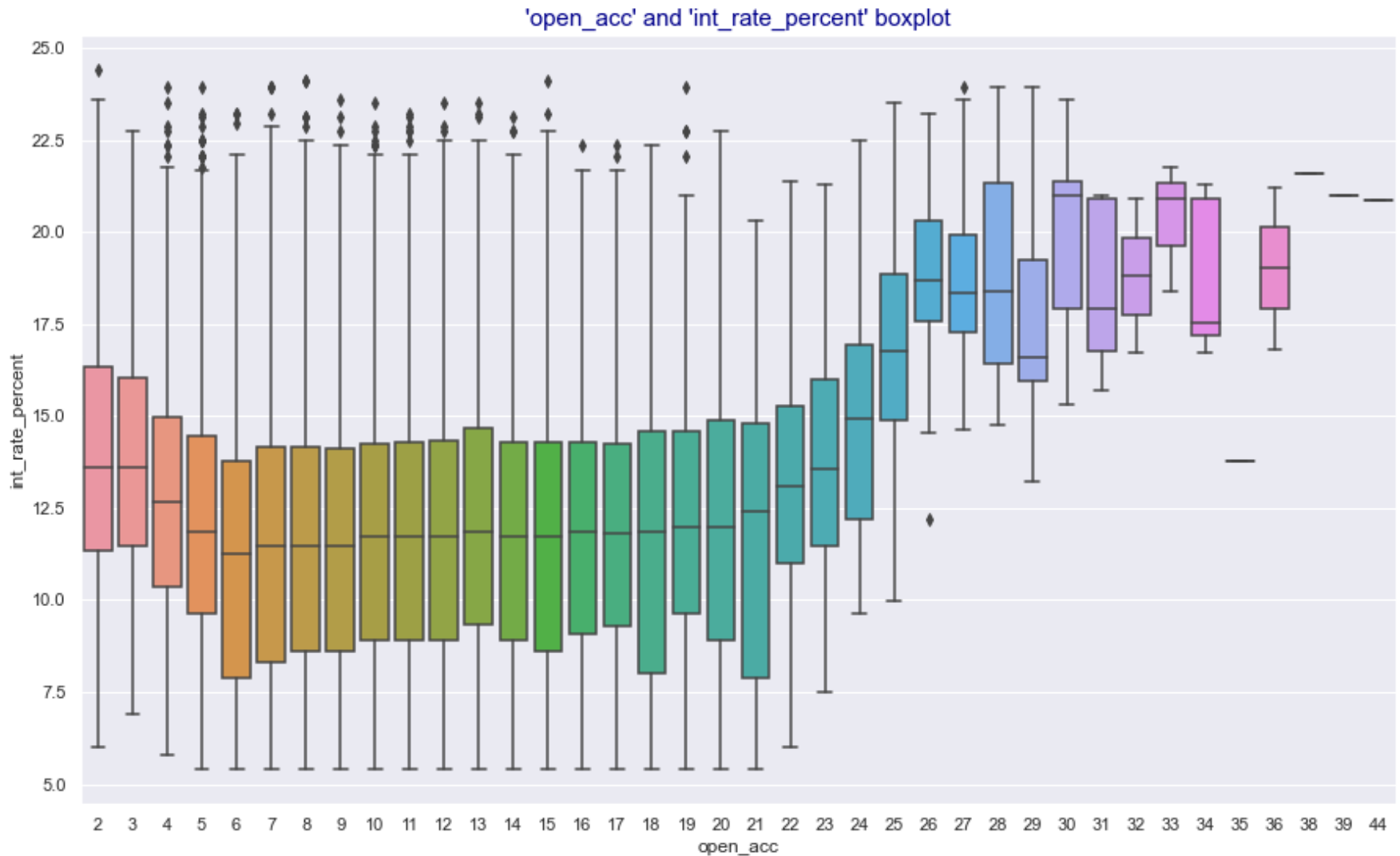




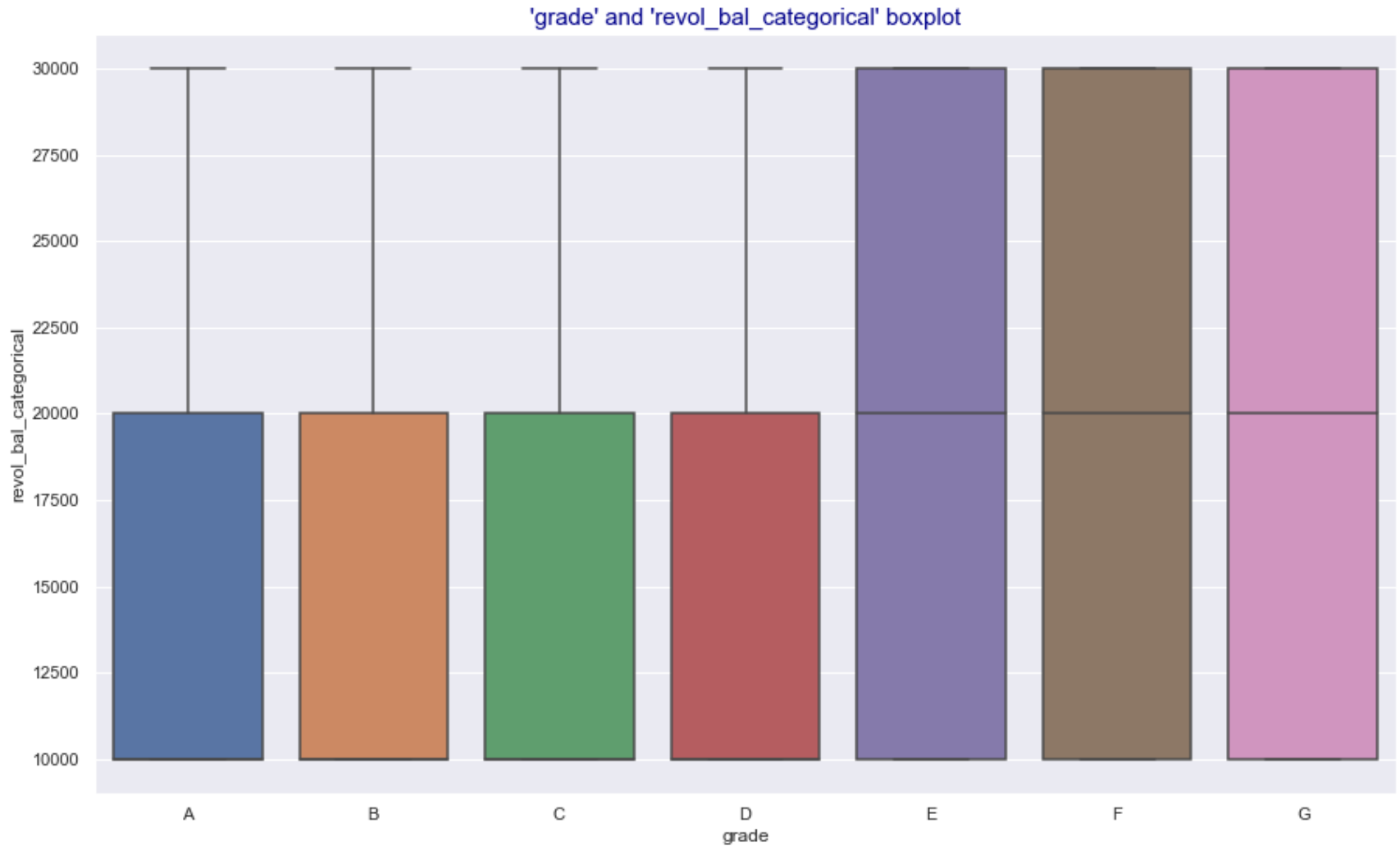
- dti above 25 have very few loans with interest rate above 13 percent



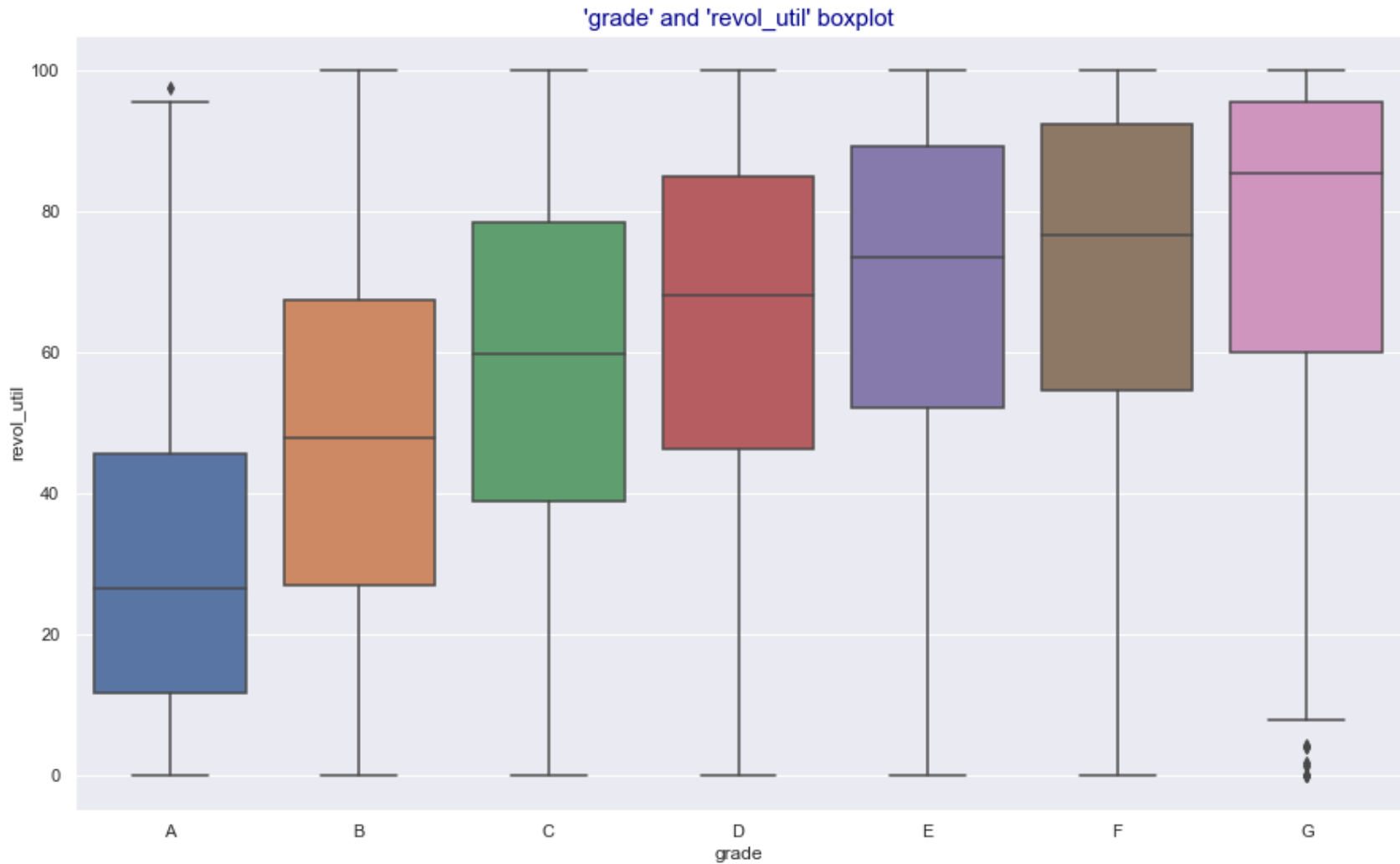
- the interest rate rise steadily for customers having more than 20 open accounts, approximately



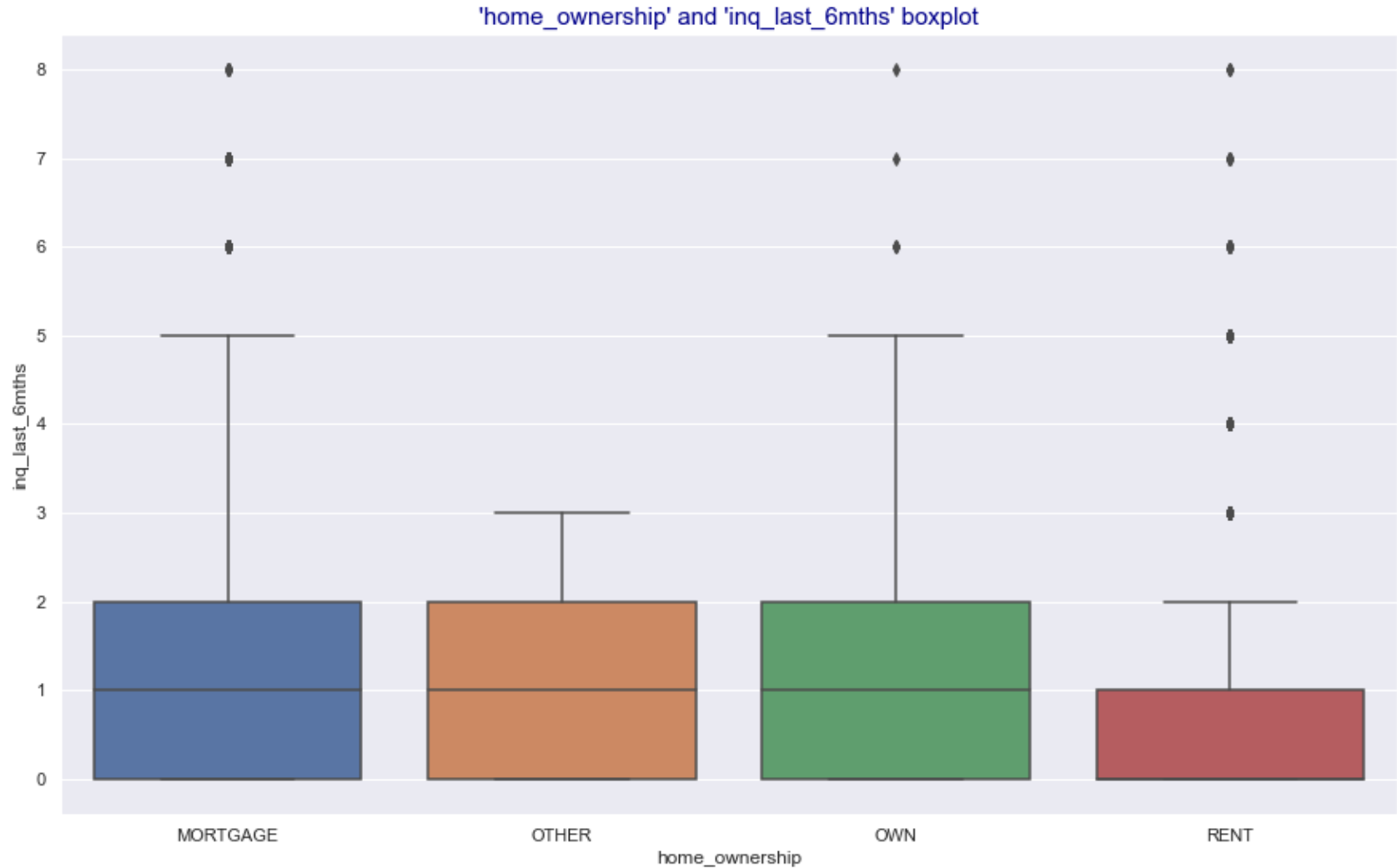
- Higher the grade, more is the revolving balance



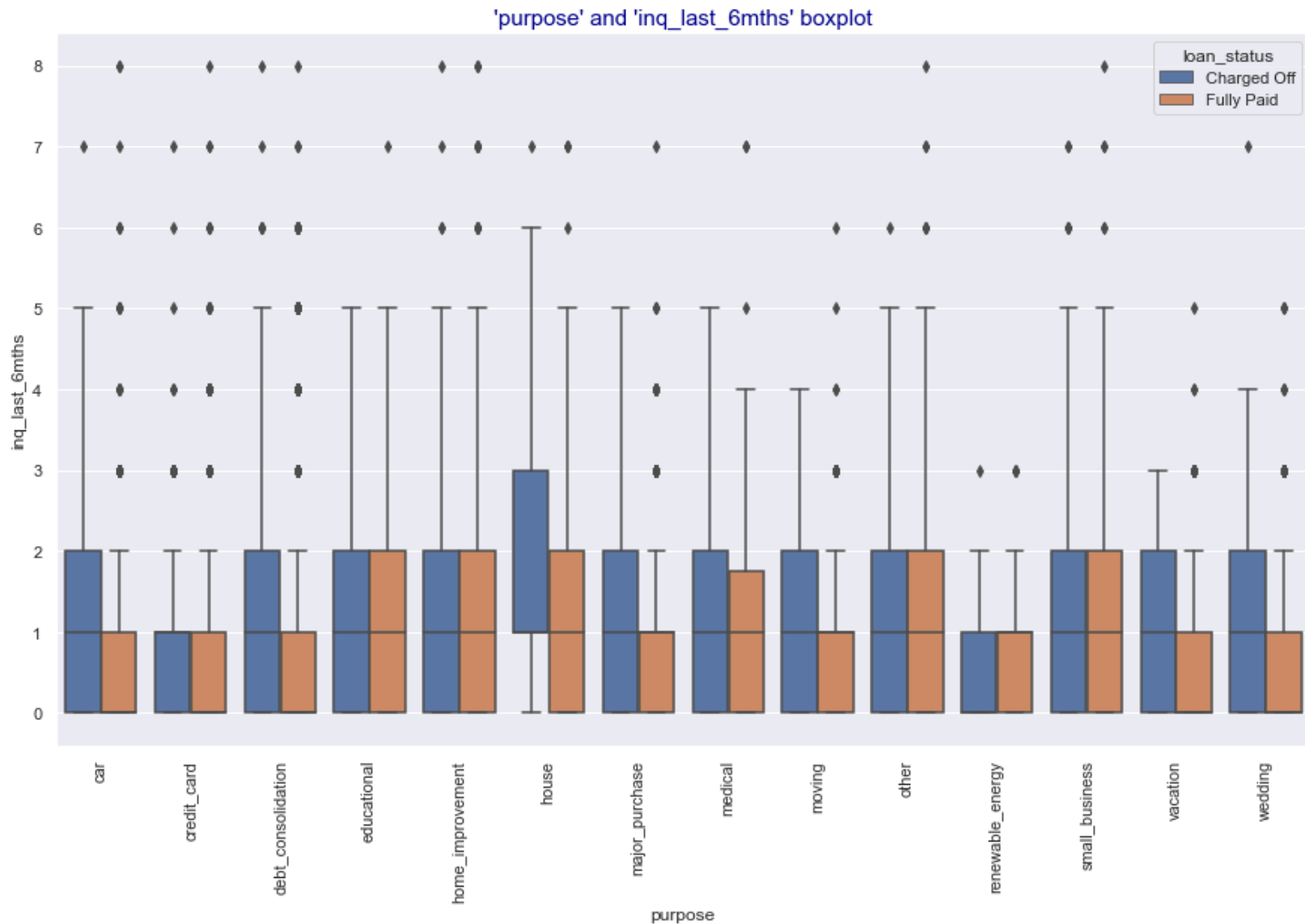
- Higher the grade, more is the revolving balance utilization rate



- In home\_ownership, 'OWN' and 'MORTGAGE' have more average inquiries, relatively



- In purpose, 'education', 'home\_improvement', 'house', 'other' and 'small\_business' have more number of inquiries. Inquiries for default is greater than or equal to non default for all purposes



# Conclusion

- 'interest\_rate' - higher the interest rate, more the chances of default
- 'loan\_amnt' - higher the loan amount, more the chances of default
- 'grade' - higher the grade, more the chances of default
- 'purpose' - Lender should pay special attention to 'debt\_consolidation', 'other' and 'small\_business' categories
- 'annual\_inc\_categorical' - lower the annual income, more likely to be a defaulter
- 'inq\_last\_6mths' - higher the inquiries, more the chances of default

If 3 or more of the above mentioned points are true for an applicant, he/she is very likely to be a defaulter.