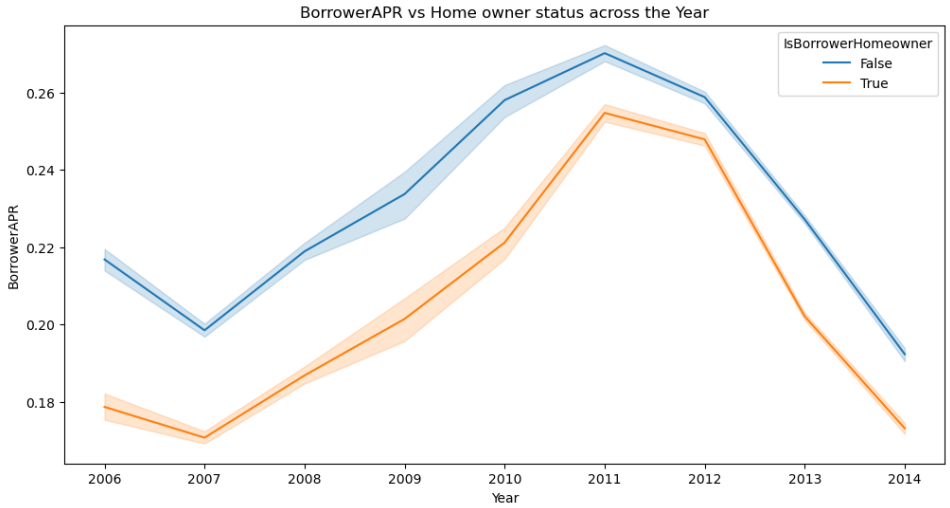
Prosper Loan dataset

Explanatory Analysis

This report aims to communicate the findings gotten from the exploratory analysis.

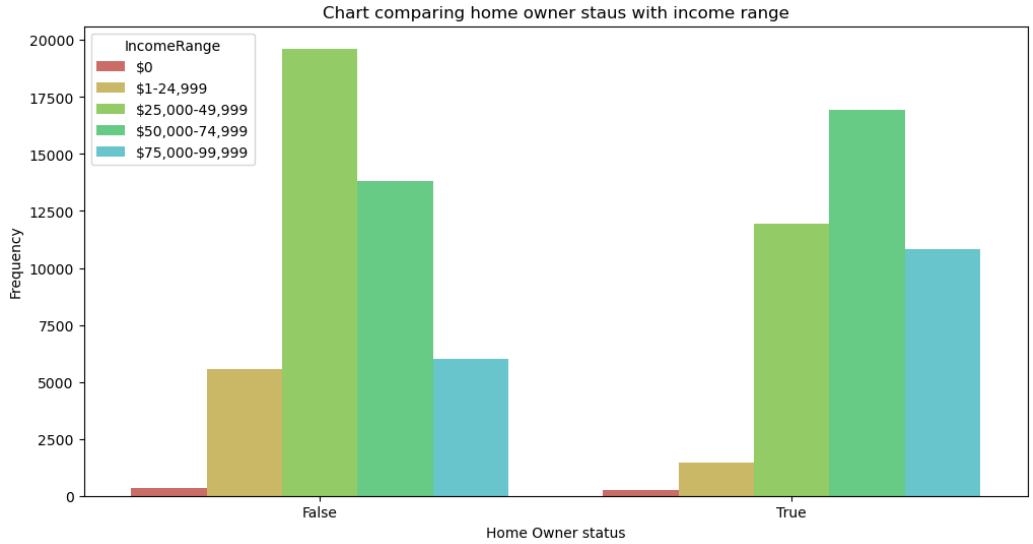
Observation 1: The cost of borrowing is higher for non-homeowners across the year



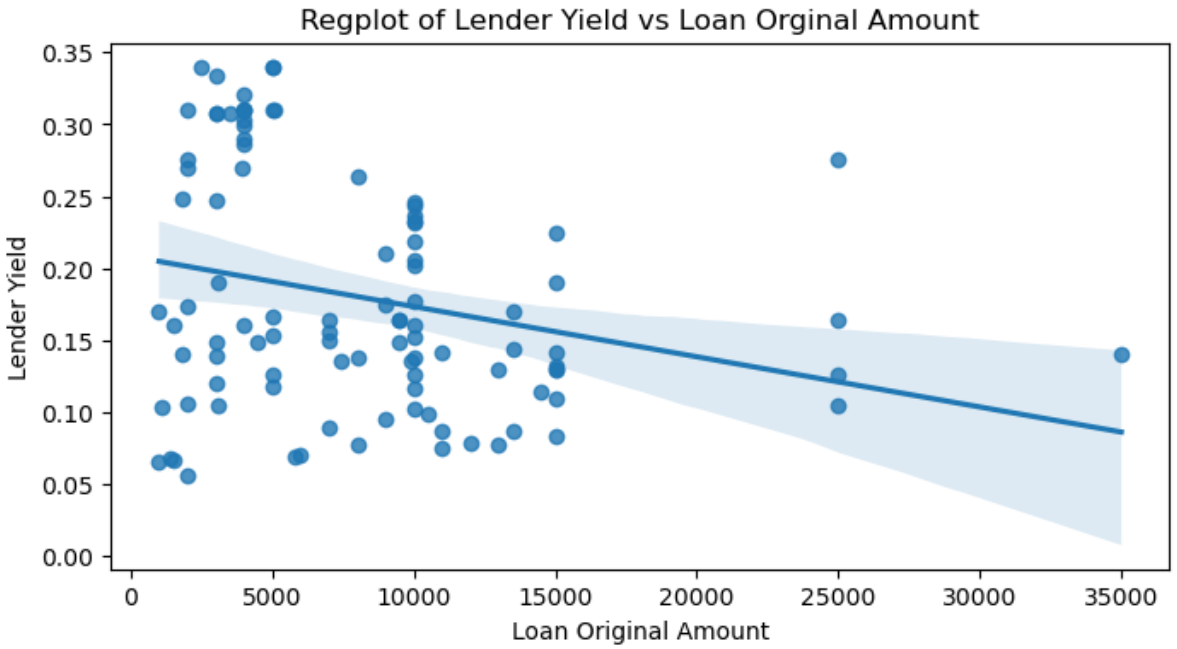
The line plot above shows the status of homeowners across the year. It can be observed that borrower APR is higher for borrowers without homes, with the highest value of borrower APR at 0.42. This means that the cost of borrowing is higher for non-homeowners across the year.

Observation 2: Most homeowners earn higher than non-homeowners

Similarly, the income range of homeowner status indicates that most homeowners earn higher than non-homeowners with the income range of $50,000 to $74,999 being the highest. Furthermore, most non-homeowners fall within the $25,000 to $49,999 income range.



Observation 3: There is a negative correlation between these lender yields and loan status.



This plot above shows a negative correlation between these two variables. as loan amount increases, lender yield decreases.

Conclusion

The analysis of the prosper loan data began with cleaning the dataset to make it tidy and improve the quality of insights. Subsequently, an exploratory analysis was conducted to determine the key insight from the dataset. It can be concluded that Borrower APR significantly affects most variables. For borrowers registered as homeowners, the borrower APR appears to be less than that of non-homeowners, indicating that the cost of borrowing is higher for non-homeowners. Similarly, as the loan amount increases, the lender yield tender reduces. It is recommended that the loan amount should range between $10,000 to $15,000 to maintain optimal lender yield.