




Exploratory Data Analysis






INTRODUCTION

- The presentation explains the loan approval data analysis and the prediction
 - The dataset contains all the details of applicants(age, income, credit score & loan status etc)
 - The goal is to understand the approval patterns and identify the most influential factors
 - The dataset provides a balanced mix of applicant demographics and financial indicators.
- 



PROBLEM STATEMENT

- Loan approval prediction
 - Income Vs Loan Approval
 - Credit score analysis
 - Age factor in Loan Decisions
 - Occupation and Loan Approval
 - Education level impact
 - Martial Status Insight
 - Multi-Factor Risk Analysis
- 



TOOLS AND DATASET USED

Tools Used

- Python programming language
- Pandas for data handling and implementation
- Numpy to do the mathematical operations
- Matplotlib to do the data visualization
- Google Colab to implement

Dataset

- Dataset:-Loan Dataset
- Source:- Kaggle



ROLE OF TOOLS


Pandas

- Loaded the dataset into data frame
- Displayed the rows to get an overview
- Analyzed the dataset info(column names, data types)
- Checked the shape of dataset(rows and columns)
- Generated statistical summary of data
- Counted the categorical values such as loan approval status

Numpy

- Performed correlation calculation and lambda functions

Matplotlib

- Created bar charts and plots
- 

IMPLEMENTATION

Step A:

- Load and read the dataset, check shape, datatype, and summary statistics
- Calculated mean, median, standard deviation for numeric features(Age, income, Credit score)
- Frequency counts for categorical features(Gender, Occupation, Education, Martial status)
- Values counts and visualizations for loan status and credit score distribution

Step B:

- Grouped by loan status to compare average Age, Income and Credit score
- Approval rate analysis across gender, martial status, occupation, and education
- Bar plots for distribution and approval rates
- Credit score bins and approval rates

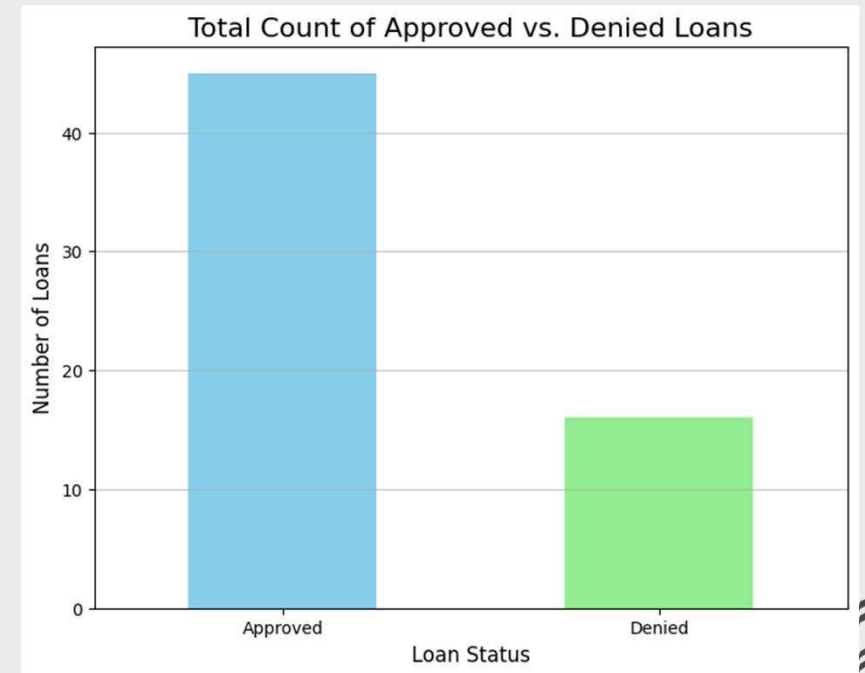
Step C:

- Correlation of numeric features (Age, income, credit score) with loan approval

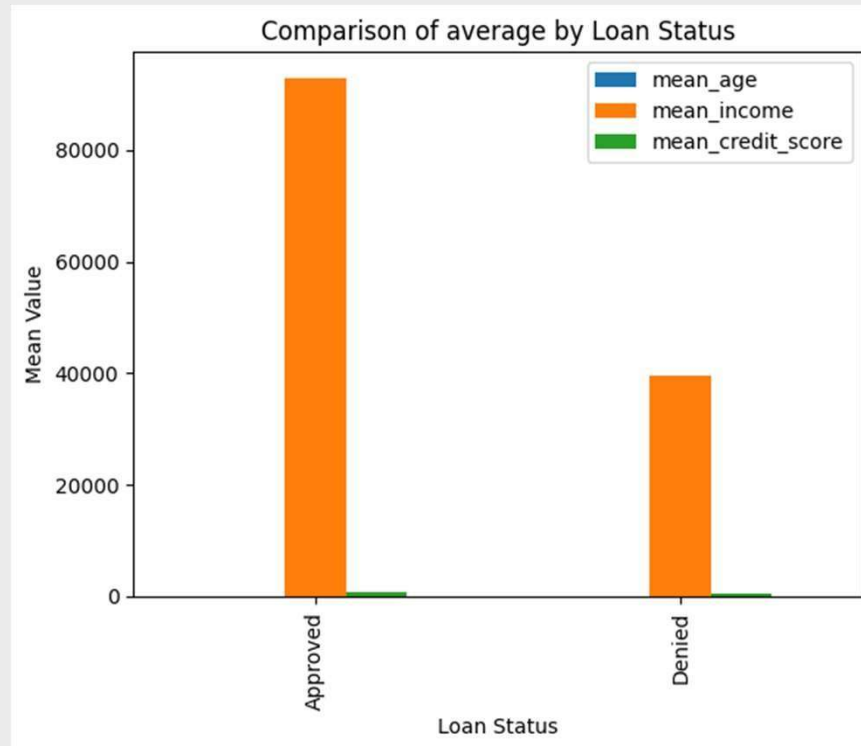
RESULTS

TARGET BALANCE

- Loan status shows the medium imbalance of the class towards the approval of loans prediction is based on the approved and denial of loan
- Approved loans must have high chance that can relate to credit score and income



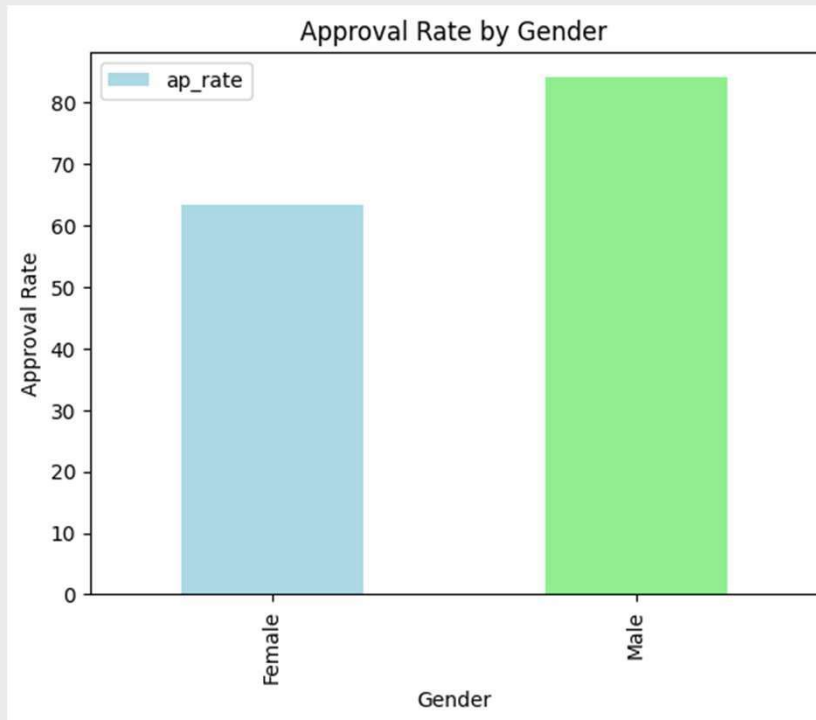
LOAN STATUS



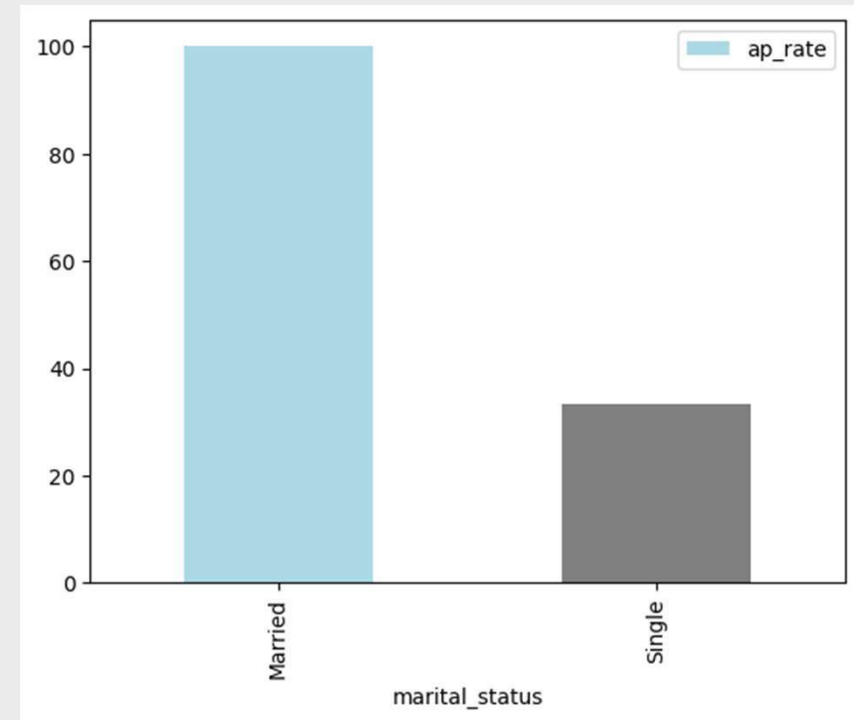
	mean_age	mean_income	mean_credit_score
loan_status			
Approved	40.311111	92955.555556	745.111111
Denied	28.000000	39687.500000	610.625000

- This numerical data shows the applicants with approvals are older, earns high income, and have higher credit score than denied application
- This shows another possibility for the approval of loans

CATEGORY APPROVALS

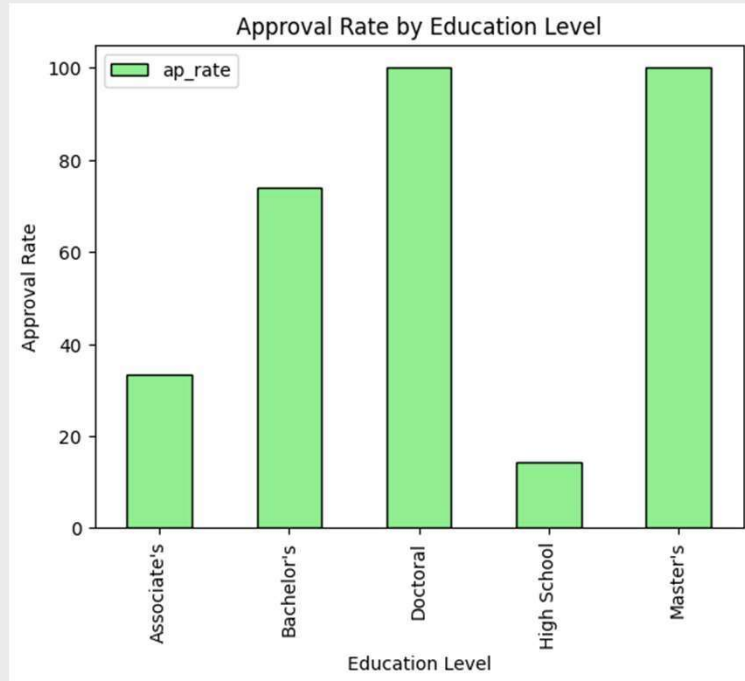


- By gender the approval rate of loans for male applicant are 83% and female 63%
- Which shows a greater influence of the loan approval for male applicants



- By marital status the married applicant have 100% approval rate and single applicants only has 33% approval rate
- This insight shows the marital stability to strongly influence approval

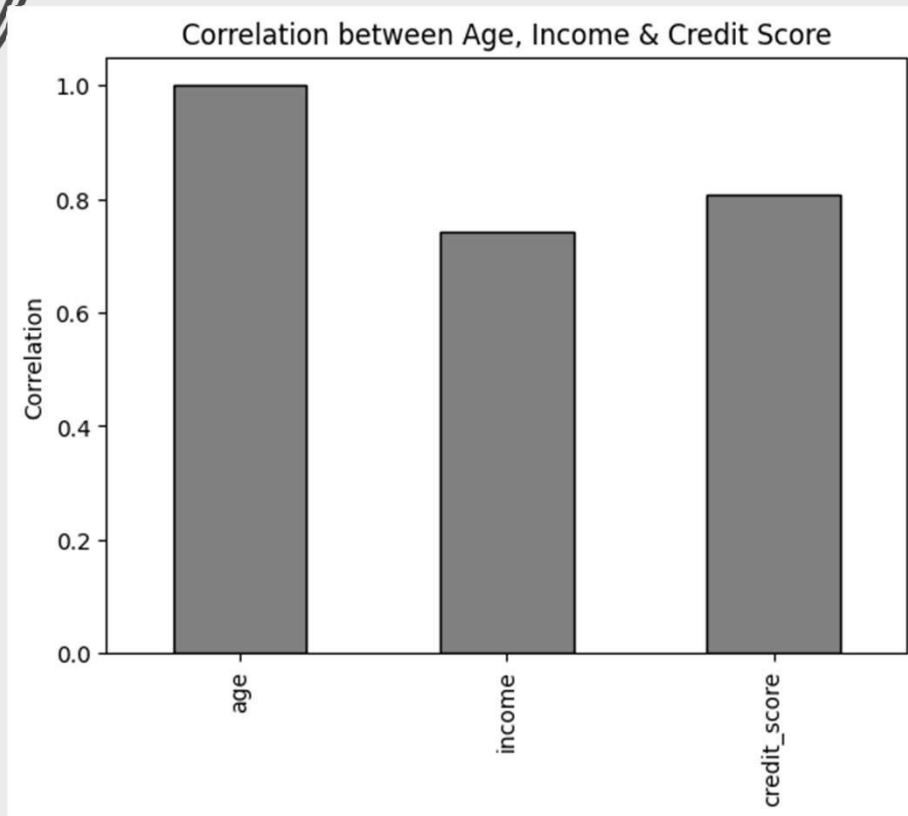
CATEGORY APPROVALS



- By Education the data shows that the applicant with Doctor's & Master's has 100% approvals and bachelors with 70% and for the other it is less than 50%
- And the insight shows that the higher education strongly correlates with approval of the loan applications

	education_level	ap_rate
0	Associate's	33.333333
1	Bachelor's	73.913043
2	Doctoral	100.000000
3	High School	14.285714
4	Master's	100.000000

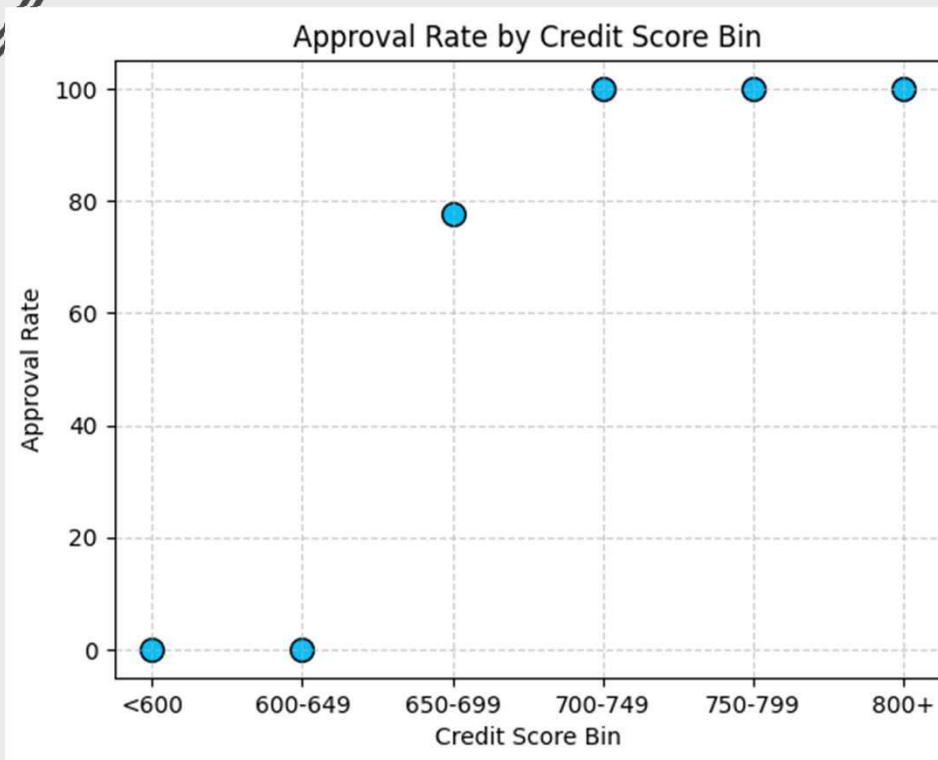
CORRELATION WITH APPROVALS



- This insight shows the main insight for the approval of loans where the approvals of loan approval is influenced by credit score followed by income and age
- The credit score an strong factor for loan approval and income were higher income improves approval chances and age has the chances as older the applicants they are more likely to get approved

```
age          1.000000
income       0.743184
credit_score  0.806042
Name: age, dtype: float64
```

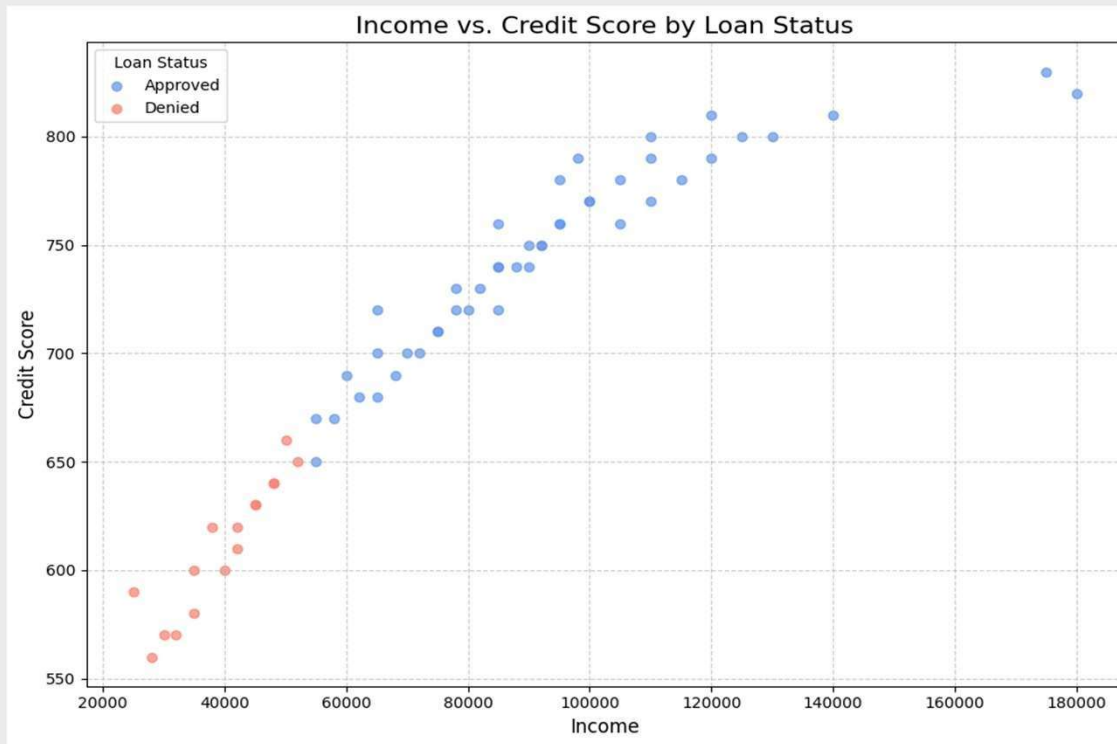
CREDIT SCORE AND APPROVAL



- The average credit score for the approved applicants are higher with the rate of 745+ for approved and 610 for the denied applicants
- The large gap shows a strong positive correlation with approval
- This aligns with practical criteria where higher credit score more chance for the loan approval

	cs_bin	ap_rate
0	<600	0.000000
1	600-649	0.000000
2	650-699	77.777778
3	700-749	100.000000
4	750-799	100.000000
5	800+	100.000000


CREDIT SCORE AND APPROVAL



- Approved applicant have much higher average income than denied applicants
- Approved income : 92k vs Denied income : 39k
- Approved credit score : 745 vs Denied credit score: 610
- Both income and credit score strongly influence approval decisions
- Applicants with higher earnings and better credits history are more likely to be approved




FUTURE LOAN APPROVAL PREDICTION

- Credit score will continue to be strongest predictor of approval
 - Applicants with higher income levels are expected to have greater approval chances
 - Higher education levels will consistently show stronger approval rates
 - Married application will likely remain a positive factor
- 



CONCLUSION

- Loan approval decisions are strongly influenced by credit score, income and age
 - Approved applicants are older, earn more, and have significantly higher credit scores than denied ones
 - Predictive models can be built using these features for automated approval systems
 - As overall financial credibility and stability factors (credit score, income, education, marital status) will drive future approvals
- 



THANK YOU

