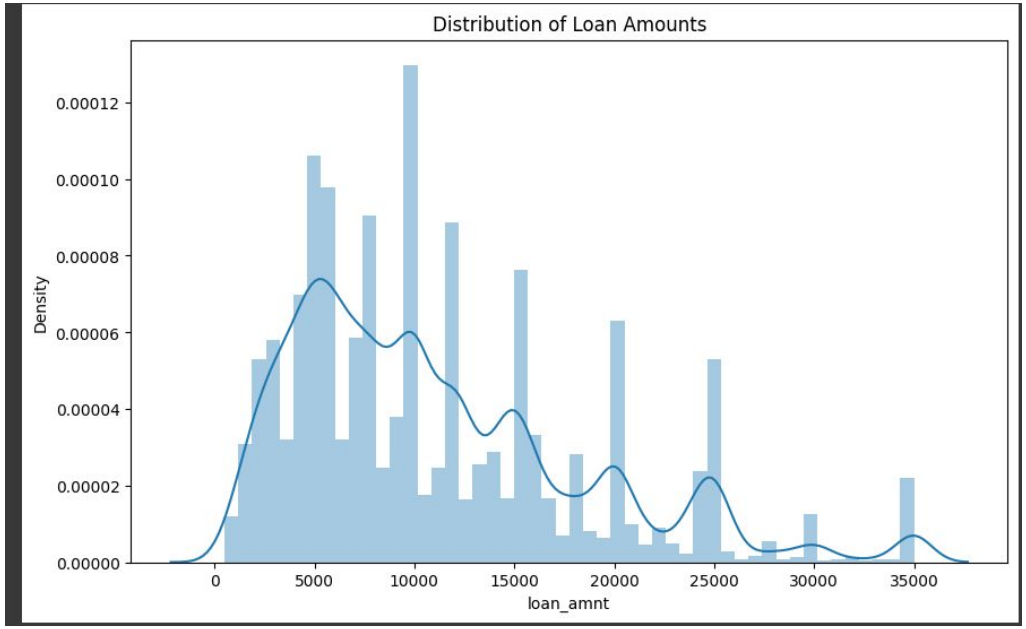
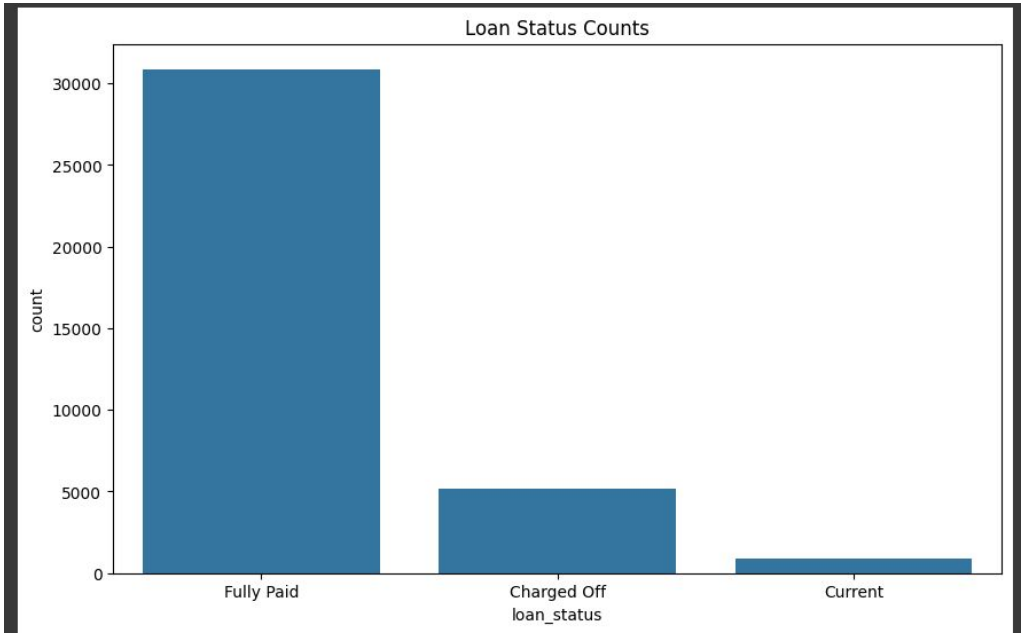
Loan Approval Analysis

Vishal Loke

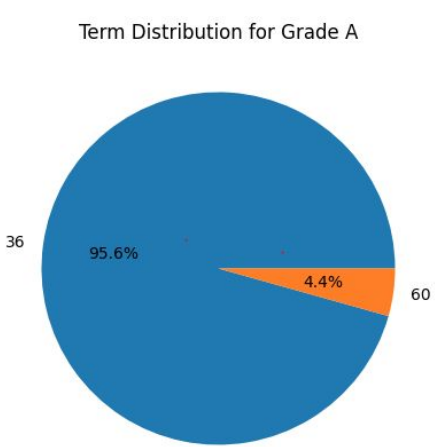
- Given graph shows distribution of loans where we can see majority of curve is left sided and majority of are between 5000-10000

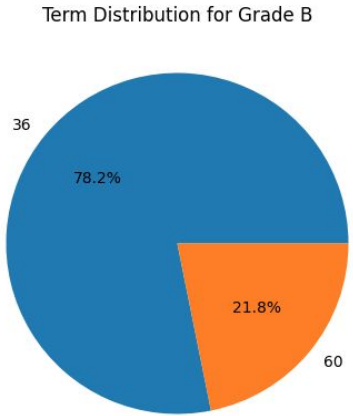


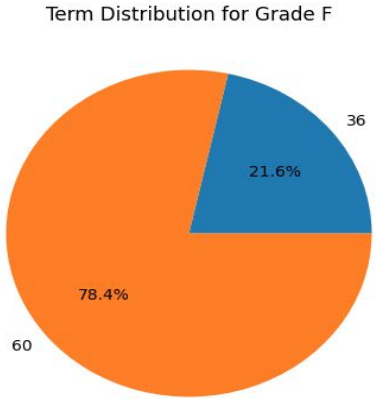
- Majority of loans are already being paid

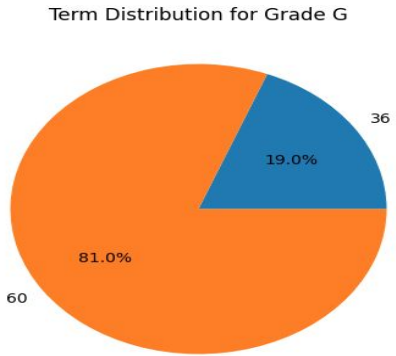


- Lower the grades goes more the length of the loan going to be , which does mean more interest rate









- Provided code

```python
def should_bank_give_loan(df):
  df_copy = df.copy()
  def risk_level(grade):
    if grade in ['A', 'B', 'C']:
      return 'Low'
    elif grade in ['D', 'E', 'F', 'G']:
      return 'Medium'
    else:
      return 'High'

  df_copy['risk_level'] = df_copy['grade'].apply(risk_level)

  avg_loan_amount_by_risk = df_copy.groupby('risk_level')['loan_amnt'].mean()

  avg_interest_rate_by_risk = df_copy.groupby('risk_level')['int_rate'].mean()

  df_copy['dti_ratio'] = df_copy['dti'] / 100

  df_copy['loan_term_years'] = df_copy['term'] / 12

  df_copy['loan_decision'] = 'Yes'

  df_copy.loc[(df_copy['risk_level'] == 'High') & ((df_copy['dti_ratio'] > 0.5) | (df_copy['loan_term_years'] > 5)), 'loan_decision'] = 'No'
  # Print details
  print(f"Average loan amount by risk level:\n{avg_loan_amount_by_risk}")
  print(f"\nAverage interest rate by risk level:\n{avg_interest_rate_by_risk}")
  print(f"\nNumber of borrowers approved for loan: {df_copy['loan_decision'].value_counts()['Yes']}")

should_bank_give_loan(working_df)
```

Output:-

```
Average loan amount by risk level:
risk_level
Low       10144.956734
Medium    14084.323367
Name: loan_amnt, dtype: float64

Average interest rate by risk level:
risk_level
Low       10.461333
Medium    16.824035
Name: int_rate, dtype: float64

Number of borrowers approved for loan: 36979
```

Code is suppose to go through classified grades and define risk based on the grade , I created smaller dataframe so i can go through data better and then provided same data frame while calling defined function