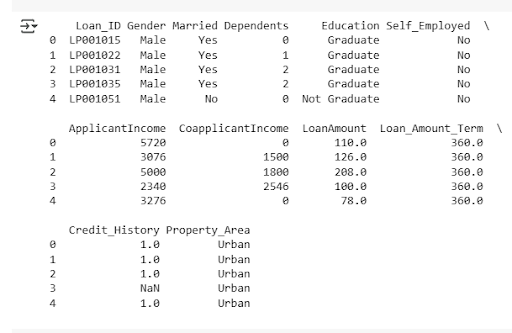
|  |  |  |
| --- | --- | --- |
| **EX:4** | **EDA-Data Inspection and Analysis** | **AD23632** |

**Aim:**

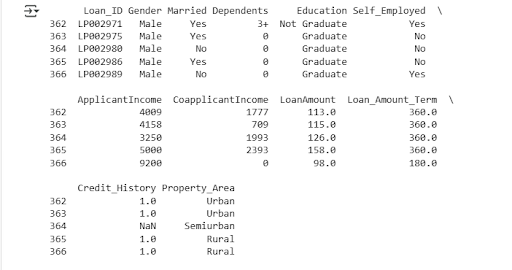
* Viewing and inspecting DataFrames
* Filtering and subsetting data using conditions
* Descriptive statistics: measures of central tendency (mean, median, mode) and measures of dispersion (range, variance, standard deviation)

**CODE:**

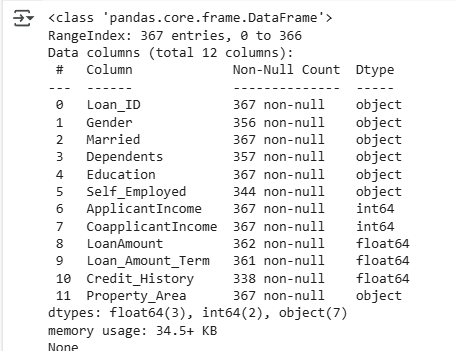
import pandas as pd  
df = pd.read\_csv('/content/test\_Y3wMUE5\_7gLdaTN.csv')  
print(df.head())



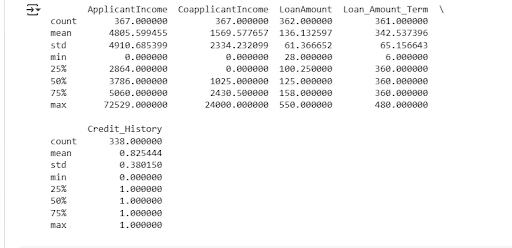
print(df.tail())



print([df.info](http://df.info" \t "_blank)())



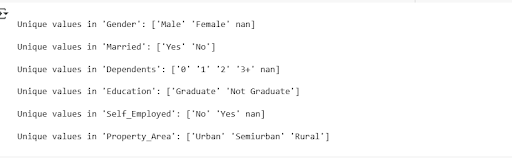
print(df.describe())



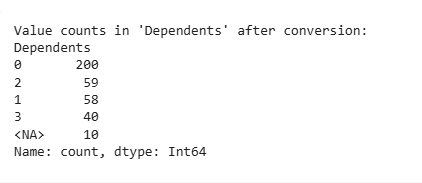
print(df.columns)

C:\Users\AI_LAB\Pictures\Saved Pictures\17.png

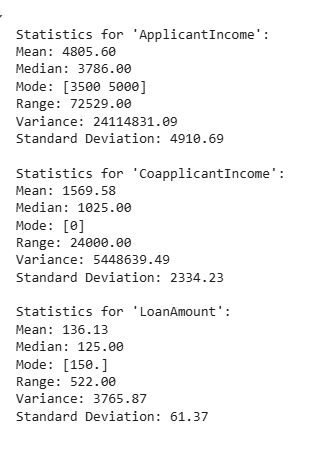
categorical\_cols = ['Gender', 'Married', 'Dependents', 'Education', 'Self\_Employed', 'Property\_Area']  
for col in categorical\_cols:  
    unique\_vals = df[col].unique()  
    print(f"\nUnique values in '{col}': {unique\_vals}")



df['Dependents'] = df['Dependents'].replace('3+', 3)  
df['Dependents'] = pd.to\_numeric(df['Dependents'], errors='coerce').astype('Int64')  
print("\nValue counts in 'Dependents' after conversion:")  
print(df['Dependents'].value\_counts(dropna=False))



cols = ['ApplicantIncome', 'CoapplicantIncome', 'LoanAmount']  
for col in cols:  
    print(f"\nStatistics for '{col}':")  
    print(f"Mean: {df[col].mean():.2f}")  
    print(f"Median: {df[col].median():.2f}")  
    print(f"Mode: {df[col].mode().values}")  
    print(f"Range: {df[col].max() - df[col].min():.2f}")  
    print(f"Variance: {df[col].var():.2f}")  
    print(f"Standard Deviation: {df[col].std():.2f}")



**Result:**

Thus the EDA-Data Inspection and Analysis is done successfully.