In [1]:
```
import pandas as pd
import numpy as np
```

In [2]:
```
df = pd.read_csv("Dataset.csv")
```

In [3]:
```
df.head()
```

Out[3]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area	Loan_Status
0	LP001002	Male	No	0	Graduate	No	5849	0.0	NaN	360.0	1.0	Urban	Y
1	LP001003	Male	Yes	1	Graduate	No	4583	1508.0	128.0	360.0	1.0	Rural	N
2	LP001005	Male	Yes	0	Graduate	Yes	3000	0.0	66.0	360.0	1.0	Urban	Y
3	LP001006	Male	Yes	0	Not Graduate	No	2583	2358.0	120.0	360.0	1.0	Urban	Y
4	LP001008	Male	No	0	Graduate	No	6000	0.0	141.0	360.0	1.0	Urban	Y

In [4]:
```
df.shape
```

Out[4]:
```
(614, 13)
```

In [5]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 614 entries, 0 to 613
Data columns (total 13 columns):
 #   Column             Non-Null Count  Dtype  
---  ------             --------------  -----  
 0   Loan_ID            614 non-null    object 
 1   Gender             601 non-null    object 
 2   Married            611 non-null    object 
 3   Dependents         599 non-null    object 
 4   Education          614 non-null    object 
 5   Self_Employed      582 non-null    object 
 6   ApplicantIncome    614 non-null    int64  
 7   CoapplicantIncome  614 non-null    float64
 8   LoanAmount         592 non-null    float64
 9   Loan_Amount_Term   600 non-null    float64
 10  Credit_History     564 non-null    float64
 11  Property_Area      614 non-null    object 
 12  Loan_Status        614 non-null    object 
dtypes: float64(4), int64(1), object(8)
memory usage: 62.5+ KB
```

In [6]:
```
df.describe()
```

Out[6]:

	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History
count	614.000000	614.000000	592.000000	600.00000	564.000000
mean	5403.459283	1621.245798	146.412162	342.00000	0.842199
std	6109.041673	2926.248369	85.587325	65.12041	0.364878
min	150.000000	0.000000	9.000000	12.00000	0.000000
25%	2877.500000	0.000000	100.000000	360.00000	1.000000
50%	3812.500000	1188.500000	128.000000	360.00000	1.000000
75%	5795.000000	2297.250000	168.000000	360.00000	1.000000
max	81000.000000	41667.000000	700.000000	480.00000	1.000000

In [7]:
```
df['Gender'].value_counts()
```

Out[7]:
```
Male      489
Female    112
Name: Gender, dtype: int64
```

In [8]:
```
df['Married'].value_counts()
```

Out[8]:
```
Yes    398
No     213
Name: Married, dtype: int64
```

In [9]:
```
df['Dependents'].value_counts()
```

Out[9]:
```
0     345
1     102
2     101
3+     51
Name: Dependents, dtype: int64
```

In [10]:
```
df['Education'].value_counts()
```

Out[10]:
```
Graduate        480
Not Graduate    134
Name: Education, dtype: int64
```

In [11]:
```
df['Self_Employed'].value_counts()
```

Out[11]:
```
No     500
Yes     82
Name: Self_Employed, dtype: int64
```

In [12]:
```
df.columns
```

Out[12]:
```
Index(['Loan_ID', 'Gender', 'Married', 'Dependents', 'Education',
       'Self_Employed', 'ApplicantIncome', 'CoapplicantIncome', 'LoanAmount',
       'Loan_Amount_Term', 'Credit_History', 'Property_Area', 'Loan_Status'],
      dtype='object')
```

In [13]:
```
df.isnull().sum()
```

Out[13]:
```
Loan_ID               0
Gender               13
Married               3
Dependents           15
Education             0
Self_Employed        32
ApplicantIncome       0
CoapplicantIncome     0
LoanAmount           22
Loan_Amount_Term     14
Credit_History       50
Property_Area         0
Loan_Status           0
dtype: int64
```

In [14]:
```
df1=df.copy()
```

In [15]:
```
df1= df1.drop(columns=['Loan_ID'])
```

In [16]:
```
df1.shape
```

Out[16]:
```
(614, 12)
```

In [17]:
```
df1['LoanAmount']=df1['LoanAmount'].fillna(df1['LoanAmount'].mean())
```

In [18]:
```
df1['Loan_Amount_Term']=df1['Loan_Amount_Term'].fillna(df1['Loan_Amount_Term'].mean())
```

In [19]:
```
df1['Credit_History']=df1['Credit_History'].fillna(df1['Credit_History'].mean())
```

In [20]:
```
df1['Gender']=df1['Gender'].fillna(df1['Gender'].mode()[0])
```

In [21]:
```
df1['Married']=df1['Married'].fillna(df1['Married'].mode()[0])
```

In [22]:
```
df1['Dependents']=df1['Dependents'].fillna(df1['Dependents'].mode()[0])
```

In [23]:
```
df1['Self_Employed']=df1['Self_Employed'].fillna(df1['Self_Employed'].mode()[0])
```

In [24]:
```
df1.isnull().sum()
```

Out[24]:
```
Gender               0
Married              0
Dependents           0
Education            0
Self_Employed        0
ApplicantIncome      0
CoapplicantIncome    0
LoanAmount           0
Loan_Amount_Term     0
Credit_History       0
Property_Area        0
Loan_Status          0
dtype: int64
```

In [25]:
```
df1.dtypes
```

Out[25]:
```
Gender               object
Married              object
Dependents           object
Education            object
Self_Employed        object
ApplicantIncome       int64
CoapplicantIncome   float64
LoanAmount          float64
Loan_Amount_Term    float64
Credit_History      float64
Property_Area        object
Loan_Status          object
dtype: object
```

In [26]:
```
from sklearn import preprocessing
label_encoder = preprocessing.LabelEncoder()
```

In [27]:
```
df1['Gender']= label_encoder.fit_transform(df1['Gender'])
```

In [28]:
```
df1['Married']= label_encoder.fit_transform(df1['Married'])
```

In [29]:
```
df1['Education']= label_encoder.fit_transform(df1['Education'])
```

In [30]:
```
df1['Self_Employed']= label_encoder.fit_transform(df1['Self_Employed'])
```

In [31]:
```
df1['Property_Area']= label_encoder.fit_transform(df1['Property_Area'])
```

In [32]:
```
df1['Loan_Status']= label_encoder.fit_transform(df1['Loan_Status'])
```

In [33]:
```
df1.dtypes
```

Out[33]:
```
Gender               int32
Married              int32
Dependents          object
Education            int32
Self_Employed        int32
ApplicantIncome      int64
CoapplicantIncome    float64
LoanAmount           float64
Loan_Amount_Term     float64
Credit_History       float64
Property_Area          int32
Loan_Status            int32
dtype: object
```

In [34]:
```
df1.head(10)
```

Out[34]:

	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area	Loan_Status
0	1	0	0	0	0	5849	0.0	146.412162	360.0	1.0	2	1
1	1	1	1	0	0	4583	1508.0	128.000000	360.0	1.0	0	0
2	1	1	0	0	1	3000	0.0	66.000000	360.0	1.0	2	1
3	1	1	0	1	0	2583	2358.0	120.000000	360.0	1.0	2	1
4	1	0	0	0	0	6000	0.0	141.000000	360.0	1.0	2	1
5	1	1	2	0	1	5417	4196.0	267.000000	360.0	1.0	2	1
6	1	1	0	1	0	2333	1516.0	95.000000	360.0	1.0	2	1
7	1	1	3+	0	0	3036	2504.0	158.000000	360.0	0.0	1	0
8	1	1	2	0	0	4006	1526.0	168.000000	360.0	1.0	2	1
9	1	1	1	0	0	12841	10968.0	349.000000	360.0	1.0	1	0