

```
In [67]: import numpy as np
import pandas as pd
import seaborn as sns
from sklearn.linear_model import LogisticRegression
```

```
In [68]: df=pd.read_csv("C8_loan-test - C8_loan-test.csv")
df
```

Out[68]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	Coap
0	LP001015	Male	Yes	0	Graduate	No	5720	
1	LP001022	Male	Yes	1	Graduate	No	3076	
2	LP001031	Male	Yes	2	Graduate	No	5000	
3	LP001035	Male	Yes	2	Graduate	No	2340	
4	LP001051	Male	No	0	Not Graduate	No	3276	
...
362	LP002971	Male	Yes	3+	Not Graduate	Yes	4009	
363	LP002975	Male	Yes	0	Graduate	No	4158	
364	LP002980	Male	No	0	Graduate	No	3250	
365	LP002986	Male	Yes	0	Graduate	No	5000	
366	LP002989	Male	No	0	Graduate	Yes	9200	

367 rows × 12 columns



```
In [69]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 367 entries, 0 to 366
Data columns (total 12 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Loan_ID               367 non-null   object
1   Gender                356 non-null   object
2   Married               367 non-null   object
3   Dependents            357 non-null   object
4   Education             367 non-null   object
5   Self_Employed         344 non-null   object
6   ApplicantIncome       367 non-null   int64
7   CoapplicantIncome     367 non-null   int64
8   LoanAmount            362 non-null   float64
9   Loan_Amount_Term      361 non-null   float64
10  Credit_History         338 non-null   float64
11  Property_Area         367 non-null   object
dtypes: float64(3), int64(2), object(7)
memory usage: 34.5+ KB
```

```
In [70]: df1=df.fillna(value=0)
df1
```

Out[70]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome
0	LP001015	Male	Yes	0	Graduate	No	5720	
1	LP001022	Male	Yes	1	Graduate	No	3076	
2	LP001031	Male	Yes	2	Graduate	No	5000	
3	LP001035	Male	Yes	2	Graduate	No	2340	
4	LP001051	Male	No	0	Not Graduate	No	3276	
...
362	LP002971	Male	Yes	3+	Not Graduate	Yes	4009	
363	LP002975	Male	Yes	0	Graduate	No	4158	
364	LP002980	Male	No	0	Graduate	No	3250	
365	LP002986	Male	Yes	0	Graduate	No	5000	
366	LP002989	Male	No	0	Graduate	Yes	9200	

367 rows × 12 columns



```
In [71]: df1.columns
```

```
Out[71]: Index(['Loan_ID', 'Gender', 'Married', 'Dependents', 'Education',
               'Self_Employed', 'ApplicantIncome', 'CoapplicantIncome', 'LoanAmount',
               'Loan_Amount_Term', 'Credit_History', 'Property_Area'],
              dtype='object')
```

```
In [72]: df2=df1[['ApplicantIncome', 'CoapplicantIncome', 'LoanAmount', 'Loan_Amount_Term', 'Credit_History']]
df2
```

Out[72]:

	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History
0	5720	0	110.0	360.0	1.0
1	3076	1500	126.0	360.0	1.0
2	5000	1800	208.0	360.0	1.0
3	2340	2546	100.0	360.0	0.0
4	3276	0	78.0	360.0	1.0
...
362	4009	1777	113.0	360.0	1.0
363	4158	709	115.0	360.0	1.0
364	3250	1993	126.0	360.0	0.0
365	5000	2393	158.0	360.0	1.0
366	9200	0	98.0	180.0	1.0

367 rows × 5 columns

```
In [73]: feature_matrix=df2.iloc[:,0:11]
target_vector=df2.iloc[:,-1]
```

```
In [74]: feature_matrix.shape
```

Out[74]: (367, 5)

```
In [75]: target_vector.shape
```

Out[75]: (367,)

```
In [76]: from sklearn.preprocessing import StandardScaler
```

```
In [77]: fs=StandardScaler().fit_transform(feature_matrix)
```

```
In [78]: logr =LogisticRegression()
logr.fit(fs,target_vector)
```

Out[78]: LogisticRegression()

```
In [79]: observation=[[1.4,2.3,5.0,11,12]]
```

```
In [80]: prediction=logr.predict(observation)
print(prediction)
```

[1.]

```
In [81]: logr.classes_
```

```
Out[81]: array([0., 1.])
```

```
In [82]: logr.predict_proba(observation)[0][0]
```

```
Out[82]: 0.0
```

```
In [83]: logr.predict_proba(observation)[0][1]
```

```
Out[83]: 1.0
```

```
In [ ]:
```

```
In [ ]:
```