```
pip install pandas
```

```
Requirement already satisfied: pandas in /usr/local/lib/python3.7/dist-packages (1.3.5)
Requirement already satisfied: pytz>=2017.3 in /usr/local/lib/python3.7/dist-packages (
Requirement already satisfied: numpy>=1.17.3 in /usr/local/lib/python3.7/dist-packages
Requirement already satisfied: python-dateutil>=2.7.3 in /usr/local/lib/python3.7/dist-
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.7/dist-packages (from
```

```
import os
```

```
import pandas as pd
```

```
pd.read_csv('/content/CreditRisk.csv')
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIn
0	LP001002	Male	No	0	Graduate	No	
1	LP001003	Male	Yes	1	Graduate	No	
2	LP001005	Male	Yes	0	Graduate	Yes	
3	LP001006	Male	Yes	0	Not Graduate	No	
4	LP001008	Male	No	0	Graduate	No	
...	...	...	...	...	...	...	
609	LP002978	Female	No	0	Graduate	No	
610	LP002979	Male	Yes	3+	Graduate	No	
611	LP002983	Male	Yes	1	Graduate	No	
612	LP002984	Male	Yes	2	Graduate	No	
613	LP002990	Female	No	0	Graduate	Yes	

614 rows × 13 columns

```
data1=pd.read_csv('/content/CreditRisk.csv')
```

```
data1.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         614 non-null    int64  
 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    int64  
dtypes: float64(3), int64(3), object(7)
memory usage: 62.5+ KB
```

```
data1.head()
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantInco
0	LP001002	Male	No	0	Graduate	No	58
1	LP001003	Male	Yes	1	Graduate	No	45
2	LP001005	Male	Yes	0	Graduate	Yes	30
3	LP001006	Male	Yes	0	Not Graduate	No	25
4	LP001008	Male	No	0	Graduate	No	60

```
data1.size
```

```
7982
```

```
data1.shape
```

```
(614, 13)
```

```
data1.ndim
```

```
2
```

```
data1.at[4,'Education']
```

```
'Graduate'
```

```
data1.loc[:,'Education']
```

```
0          Graduate
1          Graduate
2          Graduate
3      Not Graduate
4          Graduate
           ...
609        Graduate
610        Graduate
611        Graduate
612        Graduate
613        Graduate
Name: Education, Length: 614, dtype: object
```

```
data1.loc[0:5,'Education']
```

```
0        Graduate
1        Graduate
2        Graduate
3    Not Graduate
4        Graduate
5        Graduate
Name: Education, dtype: object
```

```
data1.iloc[0:5,0:2]
```

	Loan_ID	Gender
0	LP001002	Male
1	LP001003	Male
2	LP001005	Male
3	LP001006	Male
4	LP001008	Male

```
data1.iloc[0:10,0:10]
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantInco
0	LP001002	Male	No	0	Graduate	No	58
1	LP001003	Male	Yes	1	Graduate	No	45
2	LP001005	Male	Yes	0	Graduate	Yes	30
3	LP001006	Male	Yes	0	Not Graduate	No	25
4	LP001008	Male	No	0	Graduate	No	60
5	LP001011	Male	Yes	2	Graduate	Yes	54
6	LP001013	Male	Yes	0	Not Graduate	No	23
7	LP001014	Male	Yes	3+	Graduate	No	30

```
data1.iloc[10:8,:]
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome

```
data1.iloc[22:32,:]
```

Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantInc

```
data1.dtypes
```

```
Loan_ID              object
Gender               object
Married              object
Dependents           object
Education            object
Self_Employed        object
ApplicantIncome       int64
CoapplicantIncome   float64
LoanAmount            int64
Loan_Amount_Term    float64
Credit_History      float64
Property_Area        object
Loan_Status           int64
dtype: object
```

```
data1['Education'].dtype
```

```
dtype('O')
```

```
data1.axes
```

```
[RangeIndex(start=0, stop=614, step=1),
 Index(['Loan_ID', 'Gender', 'Married', 'Dependents', 'Education',
        'Self_Employed', 'ApplicantIncome', 'CoapplicantIncome', 'LoanAmount',
        'Loan_Amount_Term', 'Credit_History', 'Property_Area', 'Loan_Status'],
       dtype='object')]
```

```
data1.columns
```

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

```
data1['ApplicantIncome'].std()
```

```
6109.041673387174
```

```
data1['LoanAmount'].mean()
```

```
141.16612377850163
```

```
data1['LoanAmount'].median()
```

```
125.0
```

```
data1['ApplicantIncome'].describe()
```

```
count      614.000000
mean      5403.459283
std       6109.041673
min        150.000000
25%       2877.500000
50%       3812.500000
75%       5795.000000
max      81000.000000
Name: ApplicantIncome, dtype: float64
```

```
data1.head(15)
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantInc
0	LP001002	Male	No	0	Graduate	No	5
1	LP001003	Male	Yes	1	Graduate	No	4
2	LP001005	Male	Yes	0	Graduate	Yes	3
3	LP001006	Male	Yes	0	Not Graduate	No	2
4	LP001008	Male	No	0	Graduate	No	6
5	LP001011	Male	Yes	2	Graduate	Yes	5
6	LP001013	Male	Yes	0	Not Graduate	No	2
7	LP001014	Male	Yes	3+	Graduate	No	3
8	LP001018	Male	Yes	2	Graduate	No	4
9	LP001020	Male	Yes	1	Graduate	No	12
10	LP001024	Male	Yes	2	Graduate	No	3
11	LP001027	Male	Yes	2	Graduate	NaN	2
12	LP001028	Male	Yes	2	Graduate	No	3
13	LP001029	Male	No	0	Graduate	No	1
14	LP001030	Male	Yes	2	Graduate	No	1

```
data1.iloc[1]
```

```
Loan_ID              LP001003
Gender                   Male
Married                   Yes
Dependents                  1
Education            Graduate
Self_Employed              No
ApplicantIncome          4583
CoapplicantIncome      1508.0
LoanAmount                128
Loan_Amount_Term        360.0
Credit_History            1.0
Property_Area           Rural
Loan_Status                 0
Name: 1, dtype: object
```

```
data1.iloc[:,-1]
```

```
0      1
1      0
2      1
3      1
4      1
      ..
609    1
610    1
611    1
612    1
613    0
Name: Loan_Status, Length: 614, dtype: int64
```

```
data1.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         614 non-null    int64  
 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    int64  
dtypes: float64(3), int64(3), object(7)
memory usage: 62.5+ KB
```

```
data1.iloc[-1]
```

```
Loan_ID               LP002990
Gender                  Female
Married                     No
Dependents                   0
Education             Graduate
Self_Employed              Yes
ApplicantIncome           4583
CoapplicantIncome          0.0
LoanAmount                 133
Loan_Amount_Term         360.0
Credit_History             0.0
Property_Area        Semiurban
Loan_Status                  0
Name: 613, dtype: object
```

```
data1.iloc[1]
```

```
Loan_ID              LP001003
Gender                   Male
Married                   Yes
Dependents                  1
Education            Graduate
Self_Employed              No
ApplicantIncome          4583
CoapplicantIncome      1508.0
LoanAmount                128
Loan_Amount_Term        360.0
Credit_History            1.0
Property_Area           Rural
Loan_Status                 0
Name: 1, dtype: object
```

```
data1_sorted=data1.sort_values(by='LoanAmount')
```

```
data1_sorted.head()
```

```
data1[data1['Loan_Status']==1]
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIn
0	LP001002	Male	No	0	Graduate	No	
2	LP001005	Male	Yes	0	Graduate	Yes	
3	LP001006	Male	Yes	0	Not Graduate	No	
4	LP001008	Male	No	0	Graduate	No	
5	LP001011	Male	Yes	2	Graduate	Yes	
...	...	...	...	...	...	...	
608	LP002974	Male	Yes	0	Graduate	No	
609	LP002978	Female	No	0	Graduate	No	
610	LP002979	Male	Yes	3+	Graduate	No	
611	LP002983	Male	Yes	1	Graduate	No	
612	LP002984	Male	Yes	2	Graduate	No	

422 rows × 13 columns

```
data1[data1['Loan_Status']==1].count()
```

```
Loan_ID              422
Gender               414
Married              419
Dependents           413
Education            422
Self_Employed        399
ApplicantIncome      422
CoapplicantIncome    422
LoanAmount           422
Loan_Amount_Term     414
Credit_History       385
Property_Area        422
Loan_Status          422
dtype: int64
```

```
data1.describe()
```

	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History
count	614.000000	614.000000	614.000000	600.00000	564.000000
mean	5403.459283	1621.245798	141.166124	342.00000	0.842199
std	6109.041673	2926.248369	88.340630	65.12041	0.364878
min	150.000000	0.000000	0.000000	12.00000	0.000000
25%	2877.500000	0.000000	98.000000	360.00000	1.000000
50%	3812.500000	1188.500000	125.000000	360.00000	1.000000
75%	5795.000000	2297.250000	164.750000	360.00000	1.000000
max	81000.000000	41667.000000	700.000000	480.00000	1.000000

✓ 0s completed at 10:43 AM