```
import pandas as pd
df=pd.read_csv("/content/Data_set.csv")
```

```
print(df)
df.head(10)
```

```
                    show_name      country  num_episodes             aired_on  \
0                         NaN  South Korea            16     Friday, Saturday
1                         NaN  South Korea            16     Friday, Saturday
2     Descendants of the Sun   South Korea            16  Wednesday, Thursday
3          Boys Over Flowers   South Korea            25      Monday, Tuesday
4                           W  South Korea            16  Wednesday, Thursday
..                        ...          ...           ...                 ...
95  Shut Up: Flower Boy Band   South Korea            16      Monday, Tuesday
96                     Blood   South Korea            20      Monday, Tuesday
97        Chicago Typewriter   South Korea            16     Friday, Saturday
98      Sungkyunkwan Scandal   South Korea            20      Monday, Tuesday
99                  Vagabond   South Korea            16     Friday, Saturday

   original_network  rating  current_overall_rank  lifetime_popularity_rank  \
0               tvN     8.9                  33.0                         1
1              jTBC     8.7                  89.0                         2
2              KBS2     8.7                  77.0                         3
3              KBS2     7.7                2249.0                         4
4               MBC     8.5                 201.0                         5
..              ...     ...                   ...                       ...
95              tvN     8.1                 806.0                        99
96             KBS2     7.4                3271.0                       100
97              tvN     8.8                  51.0                       101
98             KBS2     8.2                 605.0                       102
99     SBS, Netflix     8.5                 238.0                       103

    watchers
0   111706.0
1   100950.0
2    96318.0
3    94228.0
4    92121.0
..       ...
95   34668.0
96   34666.0
97       NaN
98   34615.0
99   34523.0

[100 rows x 9 columns]
```

	show_name	country	num_episodes	aired_on	original_network	rating	current_overall_rank	li
0	NaN	South Korea	16	Friday, Saturday	tvN	8.9	33.0	
1	NaN	South Korea	16	Friday, Saturday	jTBC	8.7	89.0	
2	Descendants of the Sun	South Korea	16	Wednesday, Thursday	KBS2	8.7	77.0	
3	Boys Over Flowers	South Korea	25	Monday, Tuesday	KBS2	7.7	2249.0	
4	W	South Korea	16	Wednesday, Thursday	MBC	8.5	201.0	
5	You Who Came from the Stars	South Korea	21	Wednesday, Thursday	SBS	8.6	112.0	
	Weightlifting							

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100 entries, 0 to 99
Data columns (total 9 columns):
 #   Column                   Non-Null Count  Dtype
---  ------                   --------------  -----
```

```
 0   show_name                96 non-null     object
 1   country                  100 non-null    object
 2   num_episodes             100 non-null    int64
 3   aired_on                 99 non-null     object
 4   original_network         99 non-null     object
 5   rating                   96 non-null     float64
 6   current_overall_rank     97 non-null     float64
 7   lifetime_popularity_rank  100 non-null    int64
 8   watchers                 97 non-null     float64
dtypes: float64(3), int64(2), object(4)
memory usage: 7.2+ KB
```

```
df.isnull()
```

	show_name	country	num_episodes	aired_on	original_network	rating	current_overall_rank	lifetim
0	True	False	False	False	False	False	False	
1	True	False	False	False	False	False	False	
2	False	False	False	False	False	False	False	
3	False	False	False	False	False	False	False	
4	False	False	False	False	False	False	False	
...	...	...	...	...	...	...	...	
95	False	False	False	False	False	False	False	
96	False	False	False	False	False	False	False	
97	False	False	False	False	False	False	False	
98	False	False	False	False	False	False	False	
99	False	False	False	False	False	False	False	

100 rows × 9 columns

```
df.isnull().sum()
```

```
show_name                  4
country                    0
num_episodes               0
aired_on                   1
original_network           1
rating                     4
current_overall_rank       3
lifetime_popularity_rank   0
watchers                   3
dtype: int64
```

```
df['show_name']=df['show_name'].fillna (df['aired_on'].mode()[0])
df['aired_on']=df['aired_on'].fillna (df['aired_on'].mode()[0])
df[ 'original_network']=df[ 'original_network'].fillna (df['aired_on'].mode()[0])
df.head()
```

	show_name	country	num_episodes	aired_on	original_network	rating	current_overall_rank	lifet
0	Wednesday, Thursday	South Korea	16	Friday, Saturday	tvN	8.9	33.0	
1	Wednesday, Thursday	South Korea	16	Friday, Saturday	jTBC	8.7	89.0	
2	Descendants of the Sun	South Korea	16	Wednesday, Thursday	KBS2	8.7	77.0	
3	Boys Over	South	25	Monday,	KBS2	7.7	2240.0	

```
df['rating']=df['rating'].fillna (df['rating'].mean())
df['current_overall_rank']=df['current_overall_rank'].fillna(df['current_overall_rank'].mean())
df.head()
```

	show_name	country	num_episodes	aired_on	original_network	rating	current_overall_rank	lifet
0	Wednesday, Thursday	South Korea	16	Friday, Saturday	tvN	8.9	33.0	
1	Wednesday, Thursday	South Korea	16	Friday, Saturday	jTBC	8.7	89.0	

```
df[ 'watchers']=df['watchers'].fillna (df['watchers'].median())
df.head()
```

	show_name	country	num_episodes	aired_on	original_network	rating	current_overall_rank	lifet
0	Wednesday, Thursday	South Korea	16	Friday, Saturday	tvN	8.9	33.0	
1	Wednesday, Thursday	South Korea	16	Friday, Saturday	jTBC	8.7	89.0	
2	Descendants of the Sun	South Korea	16	Wednesday, Thursday	KBS2	8.7	77.0	
	Boys Over	South		Monday,				

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100 entries, 0 to 99
Data columns (total 9 columns):
 #   Column                   Non-Null Count  Dtype
---  ------                   --------------  -----
 0   show_name                100 non-null    object
 1   country                  100 non-null    object
 2   num_episodes             100 non-null    int64
 3   aired_on                 100 non-null    object
 4   original_network         100 non-null    object
 5   rating                   100 non-null    float64
 6   current_overall_rank     100 non-null    float64
 7   lifetime_popularity_rank 100 non-null    int64
 8   watchers                 100 non-null    float64
dtypes: float64(3), int64(2), object(4)
memory usage: 7.2+ KB
```

```
df.isnull().sum()
```

```
show_name                   0
country                     0
num_episodes                0
aired_on                    0
original_network            0
rating                      0
current_overall_rank        0
lifetime_popularity_rank    0
watchers                    0
dtype: int64
```

```
import pandas as pd
df=pd.read_csv("/content/Loan_data.csv")
```

```
print(df)
df.head(10)
```

```
      Loan_ID Gender Married Dependents     Education Self_Employed  \
0    LP001015   Male     Yes          0      Graduate            No
1    LP001022   Male     Yes          1      Graduate            No
2    LP001031   Male     Yes          2      Graduate            No
3    LP001035   Male     Yes          2      Graduate            No
4    LP001051   Male      No          0  Not Graduate            No
..        ...    ...     ...        ...           ...           ...
362  LP002971   Male     Yes         3+  Not Graduate           Yes
363  LP002975   Male     Yes          0      Graduate            No
364  LP002980   Male      No          0      Graduate            No
365  LP002986   Male     Yes          0      Graduate            No
366  LP002989   Male      No          0      Graduate           Yes

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

     Credit_History Property_Area
0               1.0         Urban
1               1.0         Urban
2               1.0         Urban
3               NaN         Urban
4               1.0         Urban
..              ...           ...
362             1.0         Urban
363             1.0         Urban
364             NaN     Semiurban
365             1.0         Rural
366             1.0         Rural

[367 rows x 12 columns]
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome
0	LP001015	Male	Yes	0	Graduate	No	5720	0
1	LP001022	Male	Yes	1	Graduate	No	3076	1500
2	LP001031	Male	Yes	2	Graduate	No	5000	1800

```
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
```

```
df.isnull()
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome
0	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False

```
df.isnull().sum()
```

```
Loan_ID               0
Gender               11
Married               0
Dependents           10
Education             0
Self_Employed        23
ApplicantIncome       0
CoapplicantIncome     0
LoanAmount            5
Loan_Amount_Term      6
Credit_History       29
Property_Area         0
dtype: int64
```

```
df['Loan_ID']=df['Loan_ID'].fillna(df['LoanAmount'].mode()[0])
df.head()
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome
0	LP001015	Male	Yes	0	Graduate	No	5720	0
1	LP001022	Male	Yes	1	Graduate	No	3076	1500
2	LP001031	Male	Yes	2	Graduate	No	5000	1800
3	LP001035	Male	Yes	2	Graduate	No	2340	2546
4	LP001051	Male	No	0	Not Graduate	No	3276	0

```
df['Gender']=df['Gender'].fillna(df['Gender'].mode()[0])
df['Self_Employed']=df['Self_Employed'].fillna(df['Self_Employed'].mode()[0])
df.head()
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome
0	LP001015	Male	Yes	0	Graduate	No	5720	0
1	LP001022	Male	Yes	1	Graduate	No	3076	1500
2	LP001031	Male	Yes	2	Graduate	No	5000	1800
3	LP001035	Male	Yes	2	Graduate	No	2340	2546
4	LP001051	Male	No	0	Not Graduate	No	3276	0

```
df['Dependents']=df['Dependents'].fillna(df['Education'].mode()[0])
df.head()
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome
0	LP001015	Male	Yes	0	Graduate	No	5720	0
1	LP001022	Male	Yes	1	Graduate	No	3076	1500
2	LP001031	Male	Yes	2	Graduate	No	5000	1800
3	LP001035	Male	Yes	2	Graduate	No	2340	2546
4	LP001051	Male	No	0	Not Graduate	No	3276	0

```
df['Education']=df['Education'].fillna(df['Dependents'].mode()[0])
df.head()
```

```
df['LoanAmount']=df['LoanAmount'].fillna(df['LoanAmount'].mode()[0])
df.head()
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome
0	LP001015	Male	Yes	0	Graduate	No	5720	0
1	LP001022	Male	Yes	1	Graduate	No	3076	1500
2	LP001031	Male	Yes	2	Graduate	No	5000	1800
3	LP001035	Male	Yes	2	Graduate	No	2340	2546
4	LP001051	Male	No	0	Not Graduate	No	3276	0

```
df['Loan_Amount_Term']=df['Loan_Amount_Term'].fillna(df['Loan_Amount_Term'].mode()[0])
df.head()
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome
0	LP001015	Male	Yes	0	Graduate	No	5720	0
1	LP001022	Male	Yes	1	Graduate	No	3076	1500
2	LP001031	Male	Yes	2	Graduate	No	5000	1800
3	LP001035	Male	Yes	2	Graduate	No	2340	2546
4	LP001051	Male	No	0	Not Graduate	No	3276	0

```
df['Credit_History']=df['Credit_History'].fillna(df['Credit_History'].median())
df.head()
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome
0	LP001015	Male	Yes	0	Graduate	No	5720	0
1	LP001022	Male	Yes	1	Graduate	No	3076	1500
2	LP001031	Male	Yes	2	Graduate	No	5000	1800
3	LP001035	Male	Yes	2	Graduate	No	2340	2546
4	LP001051	Male	No	0	Not Graduate	No	3276	0

```
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             367 non-null    object 
 2   Married            367 non-null    object 
 3   Dependents         367 non-null    object 
 4   Education          367 non-null    object 
 5   Self_Employed      367 non-null    object 
 6   ApplicantIncome    367 non-null    int64  
 7   CoapplicantIncome  367 non-null    int64  
 8   LoanAmount         367 non-null    float64
 9   Loan_Amount_Term   367 non-null    float64
 10  Credit_History     367 non-null    float64
 11  Property_Area      367 non-null    object 
dtypes: float64(3), int64(2), object(7)
memory usage: 34.5+ KB
```

```
df.isnull()
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome
0	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False

```
df.isnull().sum()
```

```
Loan_ID              0
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
dtype: int64
```