In [1]:

```
import pandas as pd
```

In [3]:

```
df = pd.read_csv("C:\\Users\\bujji\\sai\\Ex-01_DS_Data_Cleansing\\Data_set.csv")
```

In [4]:

```
df
```

Out[4]:

	show_name	country	num_episodes	aired_on	original_network	rating	current_overal
0	NaN	South Korea	16	Friday, Saturday	tvN	8.9	
1	NaN	South Korea	16	Friday, Saturday	jTBC	8.7	
2	Descendants of the Sun	South Korea	16	Wednesday, Thursday	KBS2	8.7	
3	Boys Over Flowers	South Korea	25	Monday, Tuesday	KBS2	7.7	2
4	W	South Korea	16	Wednesday, Thursday	MBC	8.5	
...	...	...	...	...	...	...	
95	Shut Up: Flower Boy Band	South Korea	16	Monday, Tuesday	tvN	8.1	
96	Blood	South Korea	20	Monday, Tuesday	KBS2	7.4	3
97	Chicago Typewriter	South Korea	16	Friday, Saturday	tvN	8.8	
98	Sungkyunkwan Scandal	South Korea	20	Monday, Tuesday	KBS2	8.2	
99	Vagabond	South Korea	16	Friday, Saturday	SBS, Netflix	8.5	

100 rows × 9 columns

In [7]:

```
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
```

In [9]:

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

Out[9]:

```
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
```

In [13]:

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

In [ ]:

```

```

In [20]:

```
df
```

Out[20]:

	show_name	country	num_episodes	aired_on	original_network	rating	current_overal
0	A Korean Odyssey	South Korea	16	Friday, Saturday	tvN	8.9	
1	A Korean Odyssey	South Korea	16	Friday, Saturday	jTBC	8.7	
2	Descendants of the Sun	South Korea	16	Wednesday, Thursday	KBS2	8.7	
3	Boys Over Flowers	South Korea	25	Monday, Tuesday	KBS2	7.7	2
4	W	South Korea	16	Wednesday, Thursday	MBC	8.5	
...	...	...	...	...	...	...	
95	Shut Up: Flower Boy Band	South Korea	16	Monday, Tuesday	tvN	8.1	
96	Blood	South Korea	20	Monday, Tuesday	KBS2	7.4	3
97	Chicago Typewriter	South Korea	16	Friday, Saturday	tvN	8.8	
98	Sungkyunkwan Scandal	South Korea	20	Monday, Tuesday	KBS2	8.2	
99	Vagabond	South Korea	16	Friday, Saturday	SBS, Netflix	8.5	

100 rows × 9 columns

In [18]:

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

In [19]:

```
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                 99 non-null     object 
 4   original_network         99 non-null     object 
 5   rating                   100 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
```