EXP.NO: 01 **DATA CLEANING**

DATE:

AIM:

Aim is to clean the data with pandas library.

ALGORITHM:

1. Removal of unwanted observations.
2. Fixing structural errors.
3. Managing Unwanted Outliers.
4. Handling Missing Data.

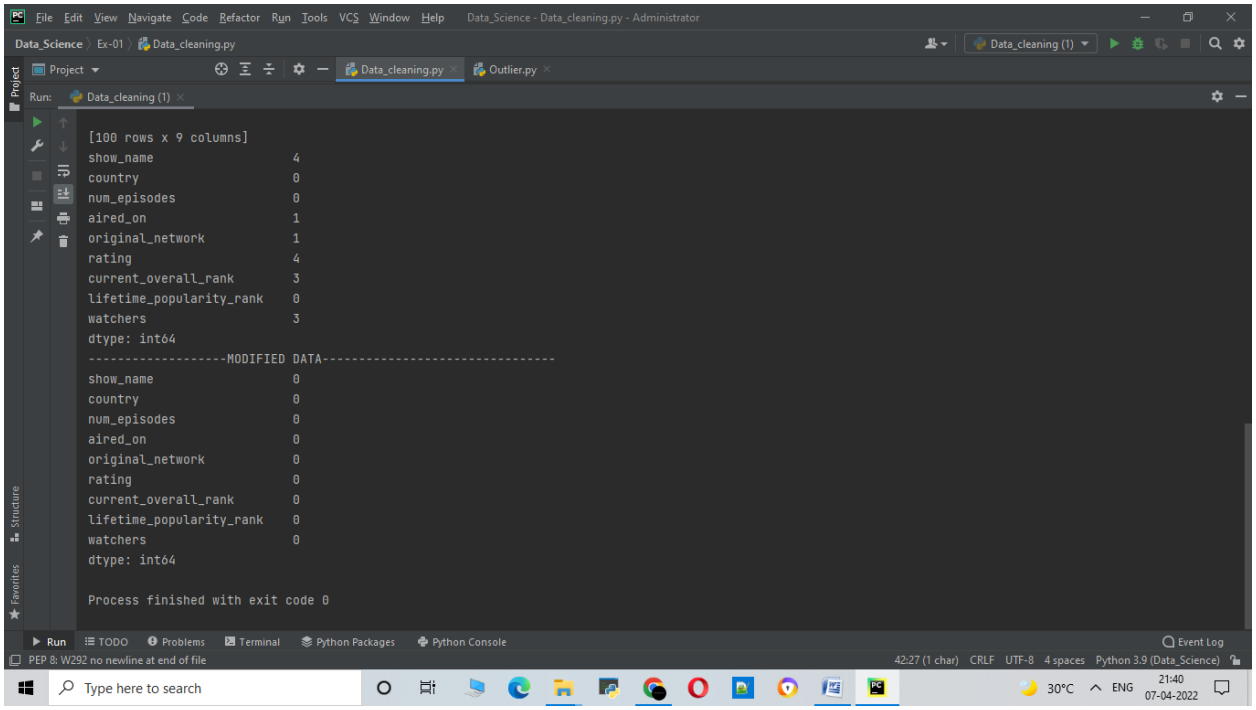
PROGRAM:

```
#importing packages

import pandas as pd

#reading CSV file

file=pd.read_csv('Data_set.csv')

#viewing first five rows

print(file.head())

#checking datatypes and background object
```

```
print(file.info())
#checking total no of rows and columns
print(file.shape)
#Locate Missing Data
print(file.isnull())
#checking datas
print(file.isnull().sum())
#input missing data
'''
df['col_name']=df['col_name'].fillna(df['col_name'].mode()[0])
df['col_name']=df['col_name'].fillna(df['col_name'].mean())
df['col_name']=df['col_name'].fillna(df['col_name'].median())
'''
file['show_name']=file['show_name'].fillna(file['show_name'].mode()[0])
file['aired_on']=file['aired_on'].fillna(file['aired_on'].mode()[0])
file['original_network']=file['original_network'].fillna(file['original_network'].mode()[0])
file['rating']=file['rating'].fillna(file['rating'].mean())
```

```
file['current_overall_rank']=file['current_overall_rank'].fillna(file['current_overall_rank'].median())

file['watchers']=file['watchers'].fillna(file['watchers'].median())

#CHEKING DATA AGAIN:

print('------------------MODIFIED DATA------------------------------')

print(file.isnull().sum())
```

OUTPUT:



RESULT:

Thus the program to clean the null values in data set has been executed and the output was verified successfully.