**EXPERIMENT:01 DATA CLEANING**

**AIM:**

To read the given data and perform data cleaning and save the cleaned data to a file.

**EXPLANATION:**

Data cleaning is the process of preparing data for analysis by removing or modifying data that is incorrect ,incompleted , irrelevant , duplicated or improperly formatted. Data cleaning is not simply about erasing data ,but rather finding a way to maximize datasets accuracy without necessarily deleting the information.

**ALGORITHM:**

**STEP 1:**

Read the given Data.

**STEP 2:**

Get the information about the data.

**STEP 3:**

Remove the null values from the data.

**STEP 4:**

Save the Clean data to the file.

**CODE:**

import pandas as pd

df=pd.read\_csv("Data\_set.csv")

df.head(10)

df.tail()

df.info()

df.isnull().sum()

df['show\_name']=df['show\_name'].fillna(df['show\_name'].mode()[0])

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

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

df['aired\_on']=df['aired\_on'].fillna(df['aired\_on'].mode()[0])

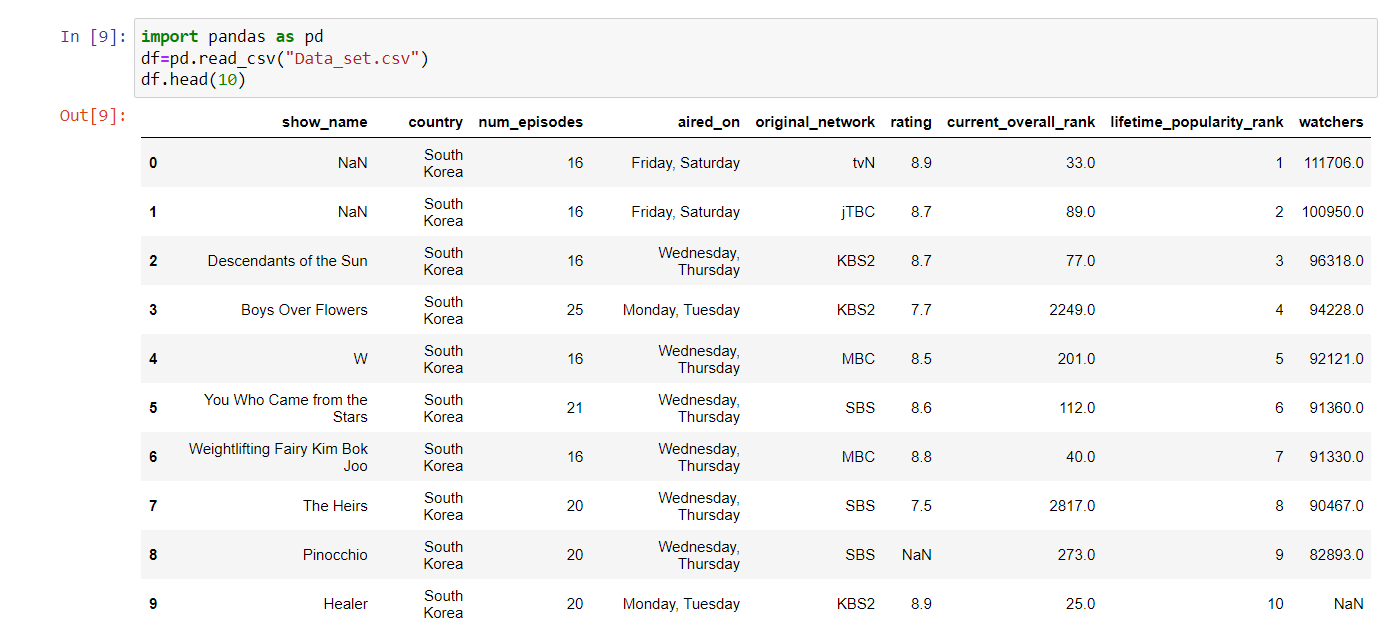
df['original\_network']=df['original\_network'].fillna(df['original\_network'].mode()[0])

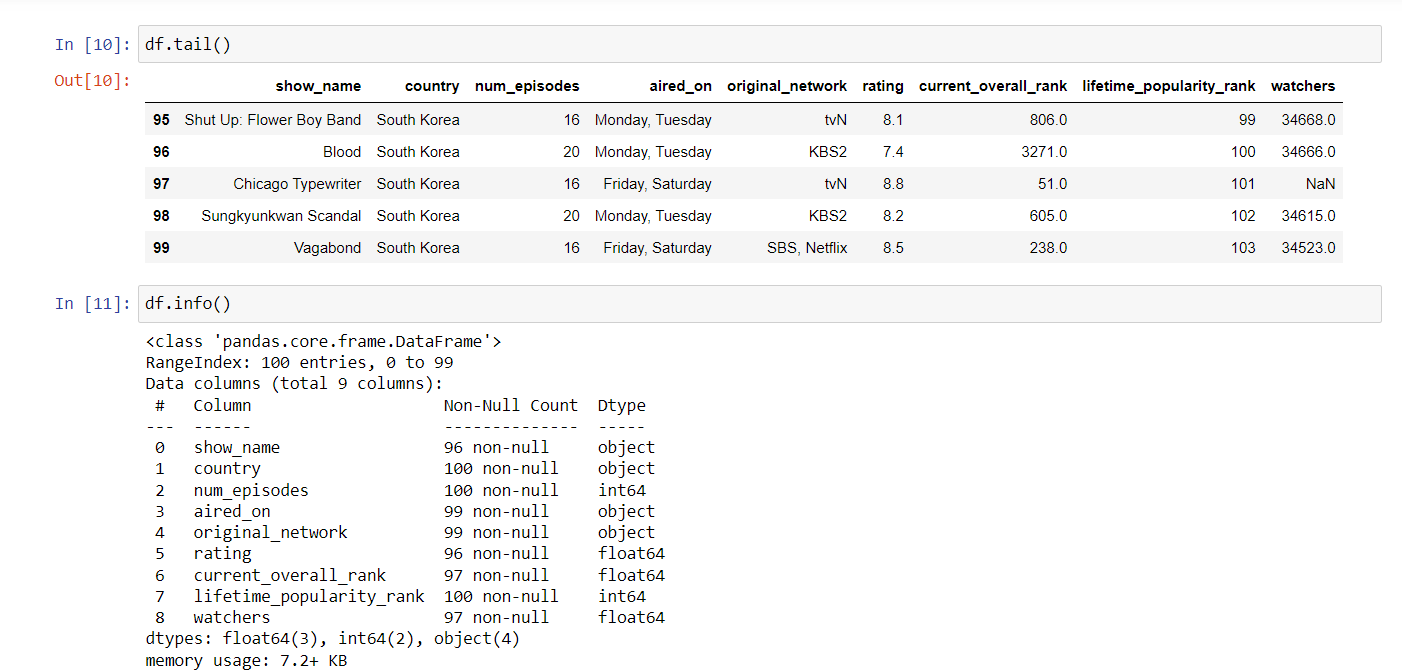
df['current\_overall\_rank']=df['current\_overall\_rank'].fillna(df['current\_overall\_rank'].median())

df.info()

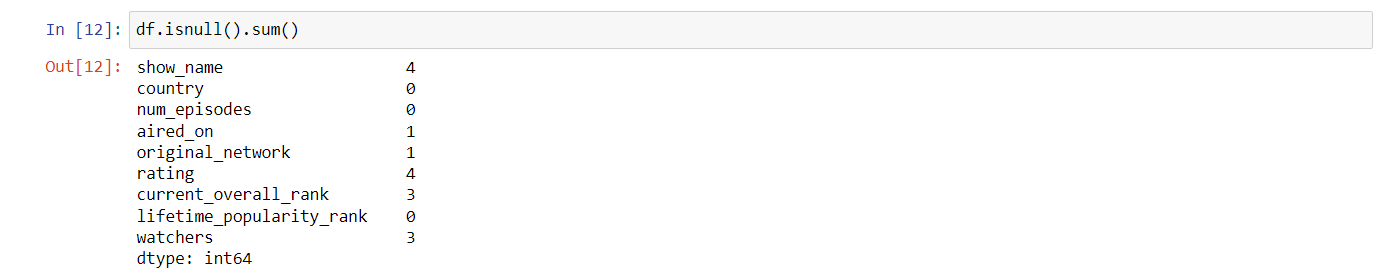
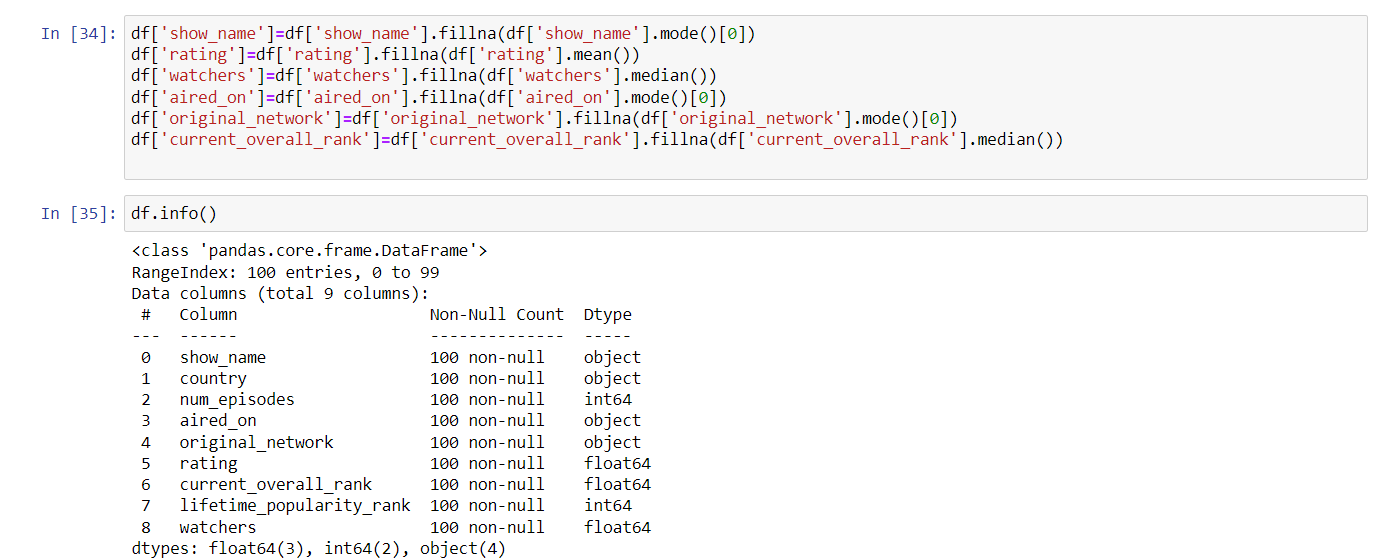
**OUTPUT:**

**BEFORE CLEANING THE DATA:**

[](https://github.com/Rithigasri/Ex-01_DS_Data_Cleansing/blob/main/out1.png)

[](https://github.com/Rithigasri/Ex-01_DS_Data_Cleansing/blob/main/out2.png)

**AFTER CLEANING THE DATA:**

[](https://github.com/Rithigasri/Ex-01_DS_Data_Cleansing/blob/main/out3.png) [](https://github.com/Rithigasri/Ex-01_DS_Data_Cleansing/blob/main/out4.png)

# RESULT:

Hence the given data is read and has undergone data cleaning.