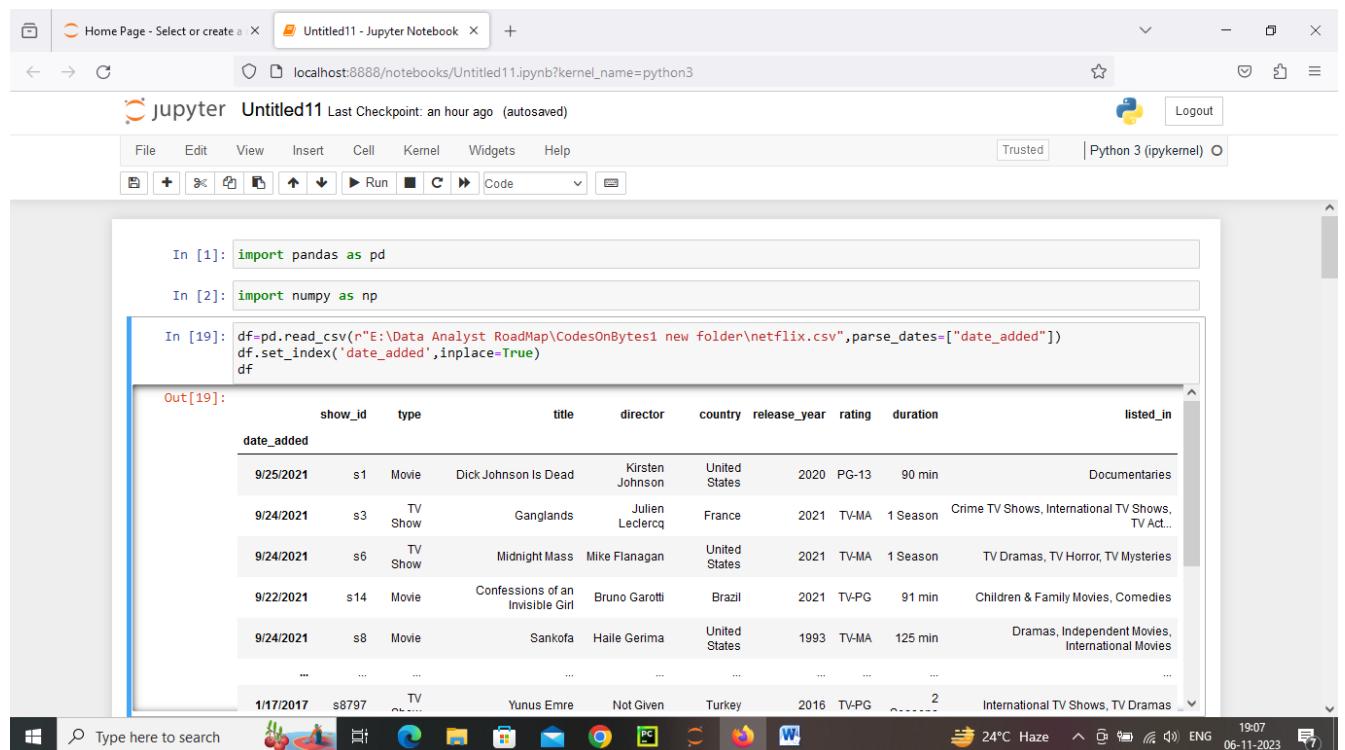


## Exploring and Cleaning Given DataSet

Dataset :- <https://docs.google.com/spreadsheets/d/e/2PACX-1vTSS-TcErkXNk8KB0AljihitwetxeHD2M3R0HJI2QPMAYFq0fxFX4PFKnzAWLDnratIz67DNL6GsZnV/pub?output=csv>

```
import pandas as pd  
  
import numpy as np  
  
df=pd.read_csv(r"E:\Data Analyst RoadMap\CodesOnBytes1 new folder\netflix.csv",parse_dates=["date_added"])  
  
df.set_index('date_added',inplace=True)  
  
df
```



The screenshot shows a Jupyter Notebook interface running on a local host. The top bar includes tabs for 'Home Page - Select or create a' and 'Untitled11 - Jupyter Notebook'. The URL is 'localhost:8888/notebooks/Untitled11.ipynb?kernel\_name=python3'. The notebook title is 'Untitled11' with a note 'Last Checkpoint: an hour ago (autosaved)'. The toolbar includes File, Edit, View, Insert, Cell, Kernel, Widgets, Help, Run, and Code. A Python 3 (ipykernel) icon is also present.

The code cells show:

```
In [1]: import pandas as pd  
In [2]: import numpy as np  
In [19]: df=pd.read_csv(r"E:\Data Analyst RoadMap\CodesOnBytes1 new folder\netflix.csv",parse_dates=["date_added"])  
df.set_index('date_added',inplace=True)  
df
```

The output cell (Out[19]) displays a table of movie data:

date_added	show_id	type	title	director	country	release_year	rating	duration	listed_in
9/25/2021	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	2020	PG-13	90 min	Documentaries
9/24/2021	s3	TV Show	Ganglands	Julien Lericq	France	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...
9/24/2021	s6	TV Show	Midnight Mass	Mike Flanagan	United States	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries
9/22/2021	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Brazil	2021	TV-PG	91 min	Children & Family Movies, Comedies
9/24/2021	s8	Movie	Sankofa	Haile Gerima	United States	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies
...	...	...	...	...	...	...	...	...	...
1/17/2017	s8797	TV	Yunus Emre	Not Given	Turkey	2016	TV-PG	2	International TV Shows, TV Dramas

```
new_df=df.fillna(method="bfill")
```

```
new_df
```

The screenshot shows a Jupyter Notebook interface with the following details:

- Header:** Home Page - Select or create a × Untitled11 - Jupyter Notebook × +
- Toolbar:** File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel) ○ Logout
- In [17]:** new\_df=df.fillna(method="bfill")  
new\_df
- Out[17]:** A table showing TV show data. The columns are: show\_id, type, title, director, country, release\_year, rating, duration, and listed\_in. The table includes rows for various shows like "Dick Johnson Is Dead", "Ganglands", "Midnight Mass", etc.
- System Tray:** Shows icons for battery, signal, and volume, along with the date and time (19:12 06-11-2023).

```
new_df=df.fillna(method="ffill")
```

```
new_df
```

Jupyter Untitled11 Last Checkpoint: an hour ago (unsaved changes)

In [29]: new\_df=df.fillna(method="ffill")  
new\_df

Out[29]:

	show_id	type	title	director	country	release_year	rating	duration	listed_in
	date_added								
9/25/2021	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	2020	PG-13	90 min	Documentaries
9/24/2021	s3	TV Show	Ganglands	Julien Ledercq	France	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...
9/24/2021	s6	TV Show	Midnight Mass	Mike Flanagan	United States	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries
9/22/2021	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Brazil	2021	TV-PG	91 min	Children & Family Movies, Comedies
9/24/2021	s8	Movie	Sankofa	Halle Gerima	United States	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies
...	...	...	...	...	...	...	...	...	...
1/17/2017	s8797	TV Show	Yunus Emre	Not Given	Turkey	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas
9/13/2018	s8798	TV Show	Zak Storm	Not Given	United States	2016	TV-Y7	3 Seasons	Kids' TV
12/15/2016	s8801	TV Show	Zindagi Gulzar Hai	Not Given	Pakistan	2012	TV-PG	1 Season	International TV Shows, Romantic TV Shows, TV ...
6/23/2018	s8784	TV Show	Yoko	Not Given	Pakistan	2016	TV-Y	1 Season	Kids' TV

new\_df.replace('Not Given','N.A.')

Jupyter Untitled11 Last Checkpoint: an hour ago (autosaved)

In [18]: new\_df.replace('Not Given','N.A.')

Out[18]:

	show_id	type	title	director	country	release_year	rating	duration	listed_in
	date_added								
9/25/2021	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	2020	PG-13	90 min	Documentaries
9/24/2021	s3	TV Show	Ganglands	Julien Ledercq	France	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...
9/24/2021	s6	TV Show	Midnight Mass	Mike Flanagan	United States	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries
9/22/2021	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Brazil	2021	TV-PG	91 min	Children & Family Movies, Comedies
9/24/2021	s8	Movie	Sankofa	Halle Gerima	United States	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies
...	...	...	...	...	...	...	...	...	...
1/17/2017	s8797	TV Show	Yunus Emre	N.A.	Turkey	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas
9/13/2018	s8798	TV Show	Zak Storm	N.A.	United States	2016	TV-Y7	3 Seasons	Kids' TV
12/15/2016	s8801	TV Show	Zindagi Gulzar Hai	N.A.	Pakistan	2012	TV-PG	1 Season	International TV Shows, Romantic TV Shows, TV ...
6/23/2018	s8784	TV Show	Yoko	N.A.	Pakistan	2016	TV-Y	1 Season	Kids' TV

new\_df.replace('1 Season','90 min')

In [30]: `new_df.replace('1 Season','90 min')`

Out[30]:

date_added	show_id	type	title	director	country	release_year	rating	duration	listed_in
9/25/2021	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	2020	PG-13	90 min	Documentaries
9/24/2021	s3	TV Show	Ganglands	Julien Leclercq	France	2021	TV-MA	90 min	Crime TV Shows, International TV Shows, TV Act...
9/24/2021	s6	TV Show	Midnight Mass	Mike Flanagan	United States	2021	TV-MA	90 min	TV Dramas, TV Horror, TV Mysteries
9/22/2021	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Brazil	2021	TV-PG	91 min	Children & Family Movies, Comedies
9/24/2021	s8	Movie	Sankofa	Halle Gerima	United States	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies
...	...	...	...	...	...	...	...	...	...
1/17/2017	s8797	TV Show	Yunus Emre	Not Given	Turkey	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas
9/13/2018	s8798	TV Show	Zak Storm	Not Given	United States	2016	TV-Y7	3 Seasons	Kids' TV
12/15/2016	s8801	TV Show	Zindagi Gulzar Hai	Not Given	Pakistan	2012	TV-PG	90 min	International TV Shows, Romantic TV Shows, TV ...
6/23/2018	s8784	TV Show	Yoko	Not Given	Pakistan	2016	TV-Y	90 min	Kids' TV
		TV							

`new_df.replace('Dramas, Independent Movies, International Movies','TV Dramas, TV Horror, TV Mysteries')`

In [25]: `new_df.replace('Dramas, Independent Movies, International Movies','TV Dramas, TV Horror, TV Mysteries')`

Out[25]:

date_added	show_id	type	title	director	country	release_year	rating	duration	listed_in
9/25/2021	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	2020	PG-13	90 min	Documentaries
9/24/2021	s3	TV Show	Ganglands	Julien Leclercq	France	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...
9/24/2021	s6	TV Show	Midnight Mass	Mike Flanagan	United States	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries
9/22/2021	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Brazil	2021	TV-PG	91 min	Children & Family Movies, Comedies
9/24/2021	s8	Movie	Sankofa	Halle Gerima	United States	1993	TV-MA	125 min	TV Dramas, TV Horror, TV Mysteries
...	...	...	...	...	...	...	...	...	...
1/17/2017	s8797	TV Show	Yunus Emre	Not Given	Turkey	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas
9/13/2018	s8798	TV Show	Zak Storm	Not Given	United States	2016	TV-Y7	3 Seasons	Kids' TV
12/15/2016	s8801	TV Show	Zindagi Gulzar Hai	Not Given	Pakistan	2012	TV-PG	1 Season	International TV Shows, Romantic TV Shows, TV ...
6/23/2018	s8784	TV Show	Yoko	Not Given	Pakistan	2016	TV-Y	1 Season	Kids' TV
		TV							

`new_df.dropna()`

Jupyter Untitled11 Last Checkpoint: an hour ago (autosaved)

In [31]: new\_df.dropna()

Out[31]:

	show_id	type	title	director	country	release_year	rating	duration	listed_in
	date_added								
9/25/2021	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	2020	PG-13	90 min	Documentaries
9/24/2021	s3	TV Show	Ganglands	Julien Leclercq	France	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...
9/24/2021	s6	TV Show	Midnight Mass	Mike Flanagan	United States	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries
9/22/2021	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Brazil	2021	TV-PG	91 min	Children & Family Movies, Comedies
9/24/2021	s8	Movie	Sankofa	Halle Gerima	United States	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies
...	...	...	...	...	...	...	...	...	...
1/17/2017	s8797	TV Show	Yunus Emre	Not Given	Turkey	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas
9/13/2018	s8798	TV Show	Zak Storm	Not Given	United States	2016	TV-Y7	3 Seasons	Kids' TV
12/15/2016	s8801	TV Show	Zindagi Gulzar Hai	Not Given	Pakistan	2012	TV-PG	1 Season	International TV Shows, Romantic TV Shows, TV ...
6/23/2018	s8784	TV Show	Yoko	Not Given	Pakistan	2016	TV-Y	1 Season	Kids' TV

Type here to search 24°C 19:24 06-11-2023

new\_df.interpolate()

Jupyter Untitled11 Last Checkpoint: an hour ago (unsaved changes)

In [32]: new\_df.interpolate()

Out[32]:

	show_id	type	title	director	country	release_year	rating	duration	listed_in
	date_added								
9/25/2021	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	2020	PG-13	90 min	Documentaries
9/24/2021	s3	TV Show	Ganglands	Julien Leclercq	France	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...
9/24/2021	s6	TV Show	Midnight Mass	Mike Flanagan	United States	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries
9/22/2021	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Brazil	2021	TV-PG	91 min	Children & Family Movies, Comedies
9/24/2021	s8	Movie	Sankofa	Halle Gerima	United States	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies
...	...	...	...	...	...	...	...	...	...
1/17/2017	s8797	TV Show	Yunus Emre	Not Given	Turkey	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas
9/13/2018	s8798	TV Show	Zak Storm	Not Given	United States	2016	TV-Y7	3 Seasons	Kids' TV
12/15/2016	s8801	TV Show	Zindagi Gulzar Hai	Not Given	Pakistan	2012	TV-PG	1 Season	International TV Shows, Romantic TV Shows, TV ...
6/23/2018	s8784	TV Show	Yoko	Not Given	Pakistan	2016	TV-Y	1 Season	Kids' TV

Type here to search 24°C 19:26 06-11-2023

new\_df.replace('05-01-2021','9/23/2021')

In [33]: `new_df.replace('05-01-2021', '9/23/2021')`

Out[33]:

	show_id	type	title	director	country	release_year	rating	duration	listed_in
	date_added								
9/25/2021	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	2020	PG-13	90 min	Documentaries
9/24/2021	s3	TV Show	Ganglands	Julien Leclercq	France	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...
9/24/2021	s6	TV Show	Midnight Mass	Mike Flanagan	United States	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries
9/22/2021	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Brazil	2021	TV-PG	91 min	Children & Family Movies, Comedies
9/24/2021	s8	Movie	Sankofa	Haile Gerima	United States	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies
...	...	...	...	...	...	...	...	...	...
1/17/2017	s8797	TV Show	Yunus Emre	Not Given	Turkey	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas
9/13/2018	s8798	TV Show	Zak Storm	Not Given	United States	2016	TV-Y7	3 Seasons	Kids' TV
12/15/2016	s8801	TV Show	Zindagi Gulzar Hai	Not Given	Pakistan	2012	TV-PG	1 Season	International TV Shows, Romantic TV Shows, TV ...
6/23/2018	s8784	TV Show	Yoko	Not Given	Pakistan	2016	TV-Y	1 Season	Kids' TV

```
import pandas as pd
```

```
pd.set_option('display.max_columns',None)
```

```
import numpy as np
```

```
df=pd.read_csv(r"E:\Data Analyst RoadMap\CodesOnBytes1 new folder\netflix.csv",parse_dates=["date_added"])
```

```
df.set_index('date_added',inplace=True)
```

```
df
```

In [10]: `import pandas as pd  
pd.set_option('display.max_columns',None)`

In [11]: `import numpy as np`

In [14]: `df=pd.read_csv(r"E:\Data Analyst RoadMap\CodesOnBytes1 new folder\netflix.csv",parse_dates=["date_added"])  
df.set_index('date_added',inplace=True)  
df`

Out[14]:

date_added	show_id	type	title	director	country	release_year	rating	duration	listed_in
9/25/2021	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	2020	PG-13	90 min	Documentaries
9/24/2021	s3	TV Show	Ganglands	Julien Ledercq	France	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...
9/24/2021	s6	TV Show	Midnight Mass	Mike Flanagan	United States	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries
9/22/2021	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Brazil	2021	TV-PG	91 min	Children & Family Movies, Comedies
9/24/2021	s8	Movie	Sankofa	Halle Gerima	United States	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies
...	...	...	...	...	...	...	...	...	...
1/17/2017	s8797	TV Show	Yunus Emre	Not Given	Turkey	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas

df.columns

In [15]: `df.columns`

Out[15]: `Index(['show_id', 'type', 'title', 'director', 'country', 'release_year',  
'rating', 'duration', 'listed_in'],  
 dtype='object')`

## OUTPUT

`Index(['show_id', 'type', 'title', 'director', 'country', 'release_year',  
'rating', 'duration', 'listed_in'],`

```
dtype='object')
```

```
df.isna().sum()
```

The screenshot shows a Jupyter Notebook interface with the title "Replace Values". The notebook has a Python 3 kernel. In cell 15, the command `df.columns` is run, showing the column names: `show_id`, `type`, `title`, `director`, `country`, `release_year`, `rating`, `duration`, `listed_in`, and `dtype='object'`. In cell 19, the command `df.isna().sum()` is run, displaying the count of missing values for each column: show\_id (0), type (0), title (0), director (0), country (0), release\_year (0), rating (0), duration (0), listed\_in (0), and dtype: int64.

```
In [15]: df.columns
Out[15]: Index(['show_id', 'type', 'title', 'director', 'country', 'release_year',
   'rating', 'duration', 'listed_in'],
              dtype='object')

In [19]: df.isna().sum()
Out[19]: show_id      0
         type        0
         title       0
         director    0
         country     0
         release_year 0
         rating      0
         duration     0
         listed_in    0
         dtype: int64
```

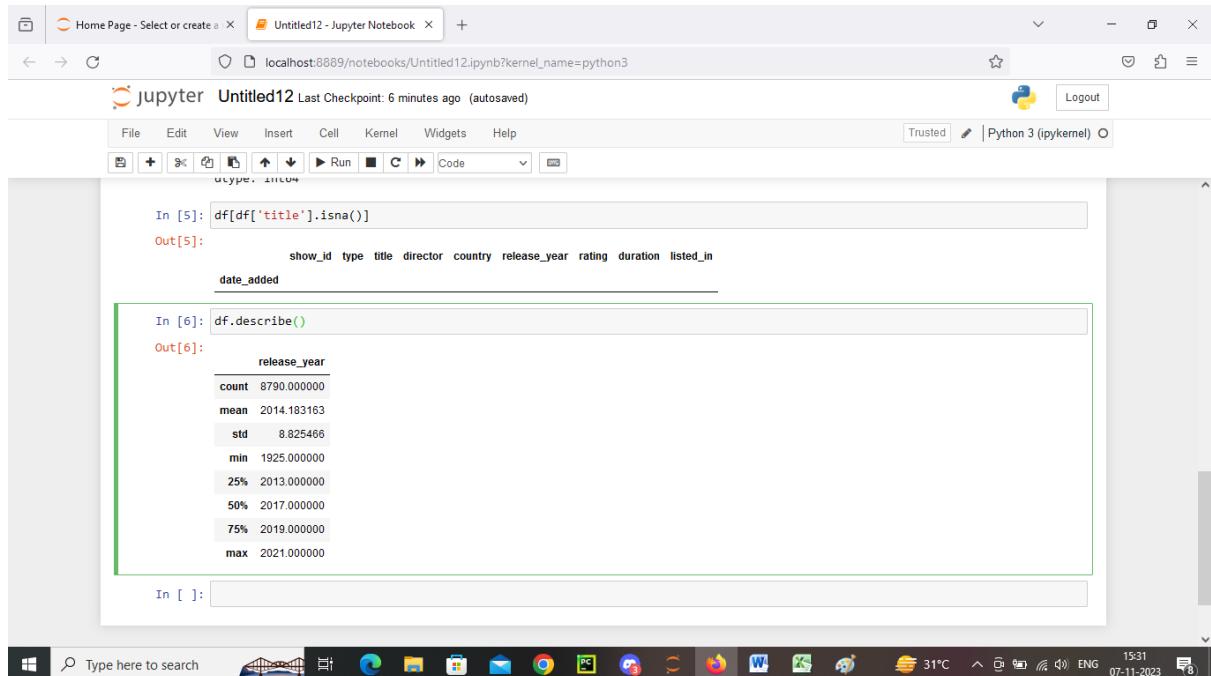
```
df[df['title'].isna()]
```

The screenshot shows a Jupyter Notebook interface with the title "Untitled12". The notebook has a Python 3 kernel. In cell 5, the command `df[df['title'].isna()]` is run, showing the rows where the title column contains missing values. In cell 6, the command `df.describe()` is run, providing a summary statistics for the dataset, including the count, mean, standard deviation, minimum, and quartiles for each column. The `release_year` column is highlighted.

```
In [5]: df[df['title'].isna()]
Out[5]:
   show_id  type  title  director  country  release_year  rating  duration  listed_in
   date_added

In [6]: df.describe()
Out[6]:
   release_year
   count    8790.000000
   mean    2014.183163
   std     8.825466
   min    1925.000000
   25%    2013.000000
   50%    2017.000000
   75%    2019.000000
   max    2021.000000
```

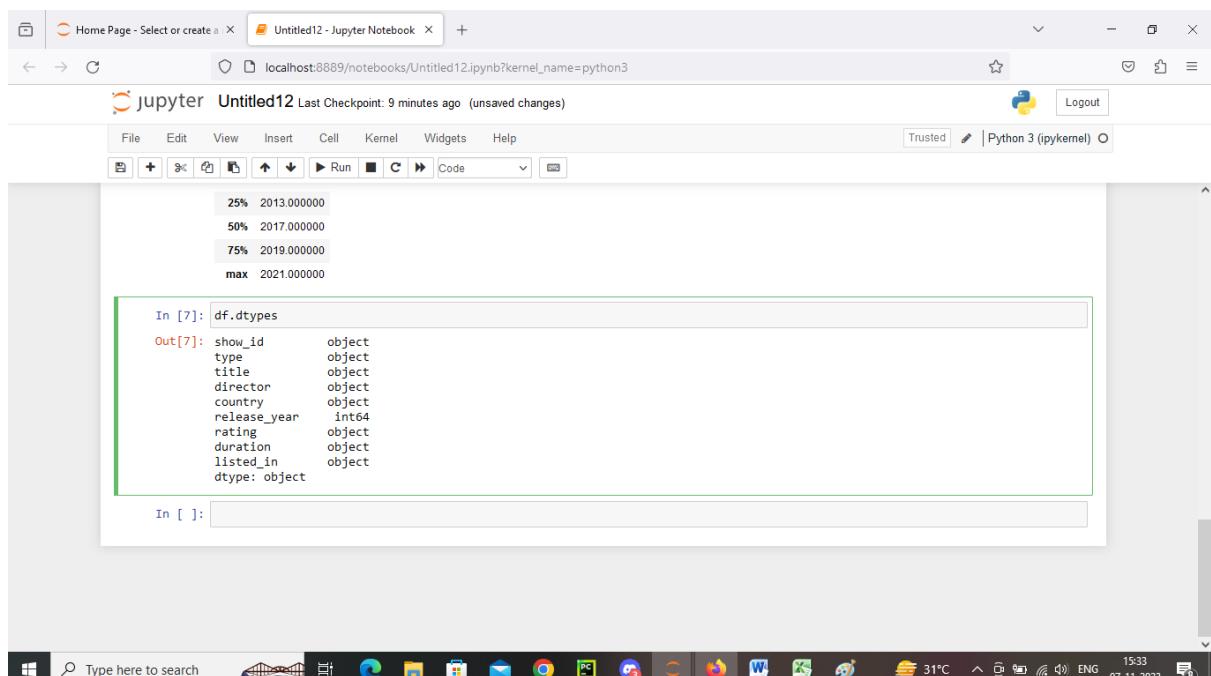
`df.describe()`



```
In [5]: df['title'].isna()
Out[5]:
show_id type title director country release_year rating duration listed_in
date_added

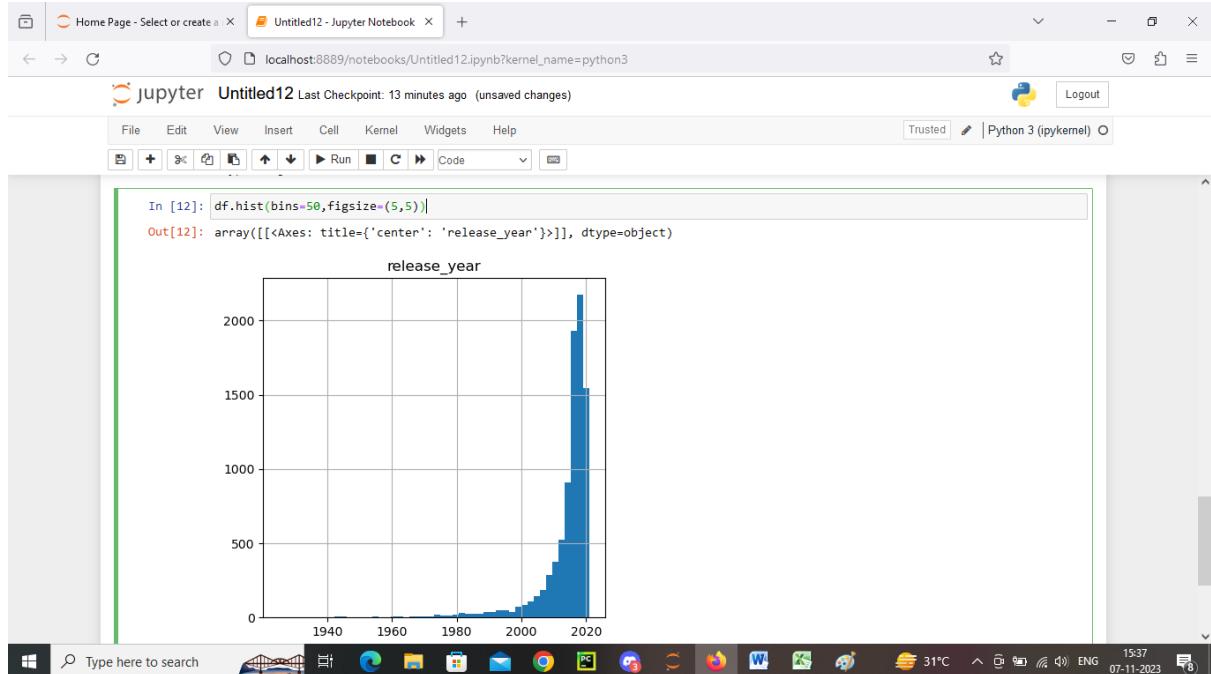
In [6]: df.describe()
Out[6]:
release_year
count    8790.000000
mean    2014.183163
std     8.825466
min    1925.000000
25%   2013.000000
50%   2017.000000
75%   2019.000000
max    2021.000000
```

`df.dtypes`



```
In [7]: df.dtypes
Out[7]:
show_id      object
type        object
title       object
director    object
country     object
release_year int64
rating      object
duration    object
listed_in   object
dtype: object
```

```
df.hist(bins=50,figsize=(5,5))
```



```
new_df=df[df['release_year']<2000]
```

```
new_df
```

A screenshot of a Jupyter Notebook interface. The code cell In [10] contains the command `new_df=df[df['release_year']<2000]`. The output cell Out[10] displays a table of movie data for releases before 2000. The table includes columns for show\_id, type, title, director, country, release\_year, rating, duration, and listed\_in. Notable entries include "Jaws" (1975), "The Super Mario Bros. Super Show!" (1989), and "The Twilight Zone (Original Series)" (1963).

show_id	type	title	director	country	release_year	rating	duration	listed_in
date_added								
9/24/2021	s8	Movie	Sankofa	Haile Gerima	United States	1993	TV-MA	125 min
9/21/2021	s25	Movie	Jeans	S. Shankar	India	1998	TV-14	166 min
9/21/2021	s27	Movie	Minsara Kanavu	Rajiv Menon	India	1997	TV-PG	147 min
9/21/2021	s23	Movie	Avai Shanmughi	K.S. Ravikumar	Not Given	1996	TV-PG	161 min
9/16/2021	s42	Movie	Jaws	Steven Spielberg	United States	1975	PG	124 min
...	...	...	...	...	...	...	...	...
9/20/2018	s8232	Movie	The Bund	Not Given	Hong Kong	1983	TV-14	103 min
5/22/2016	s8524	TV Show	The Super Mario Bros. Super Show!	Not Given	United States	1989	TV-Y7	1 Season
07-01-2017	s8542	TV Show	The Twilight Zone (Original Series)	Not Given	United States	1963	TV-14	4 Seasons
07-01-2017	s8645	TV Show	Twin Peaks	Not Given	United States	1990	TV-14	2 Seasons

```
new_df=df[df['release_year']>2000]
```

```
new_df
```

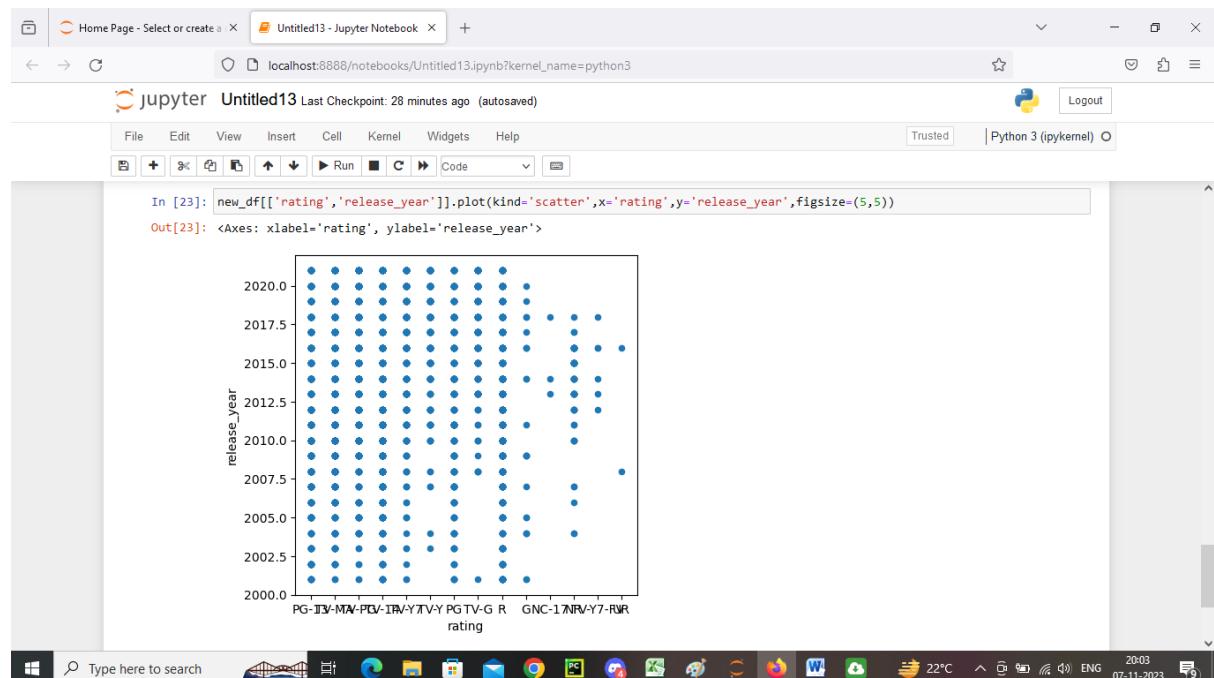
Jupyter Notebook Screenshot showing the execution of the code to filter the dataset by release year and display the resulting DataFrame.

In [11]: new\_df=df[df['release\_year']>2000]  
new\_df

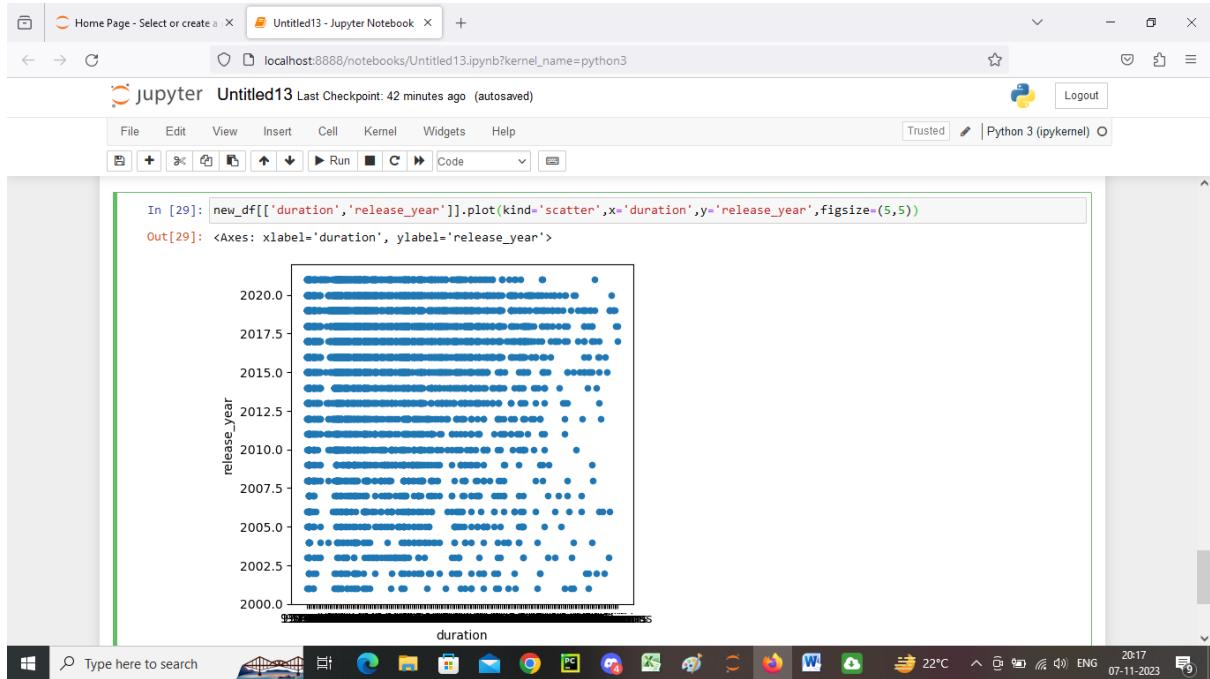
Out[11]:

	show_id	type	title	director	country	release_year	rating	duration	listed_in
		date_added							
9/25/2021	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	2020	PG-13	90 min	Documentaries
9/24/2021	s3	TV Show	Ganglands	Julien Lelercq	France	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...
9/24/2021	s6	TV Show	Midnight Mass	Mike Flanagan	United States	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries
9/22/2021	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Brazil	2021	TV-PG	91 min	Children & Family Movies, Comedies
9/24/2021	s9	TV Show	The Great British Baking Show	Andy Devonshire	United Kingdom	2021	TV-14	9 Seasons	British TV Shows, Reality TV
...	...	...	...	...	...	...	...	...	...
1/17/2017	s8797	TV Show	Yunus Emre	Not Given	Turkey	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas
9/13/2018	s8798	TV Show	Zak Storm	Not Given	United States	2016	TV-Y7	3 Seasons	Kids' TV
12/15/2016	s8801	TV Show	Zindagi Gutzar Hai	Not Given	Pakistan	2012	TV-PG	1 Season	International TV Shows, Romantic TV Shows, TV ...
6/23/2018	s8784	TV Show	Yoko	Not Given	Pakistan	2016	TV-Y	1 Season	Kids' TV

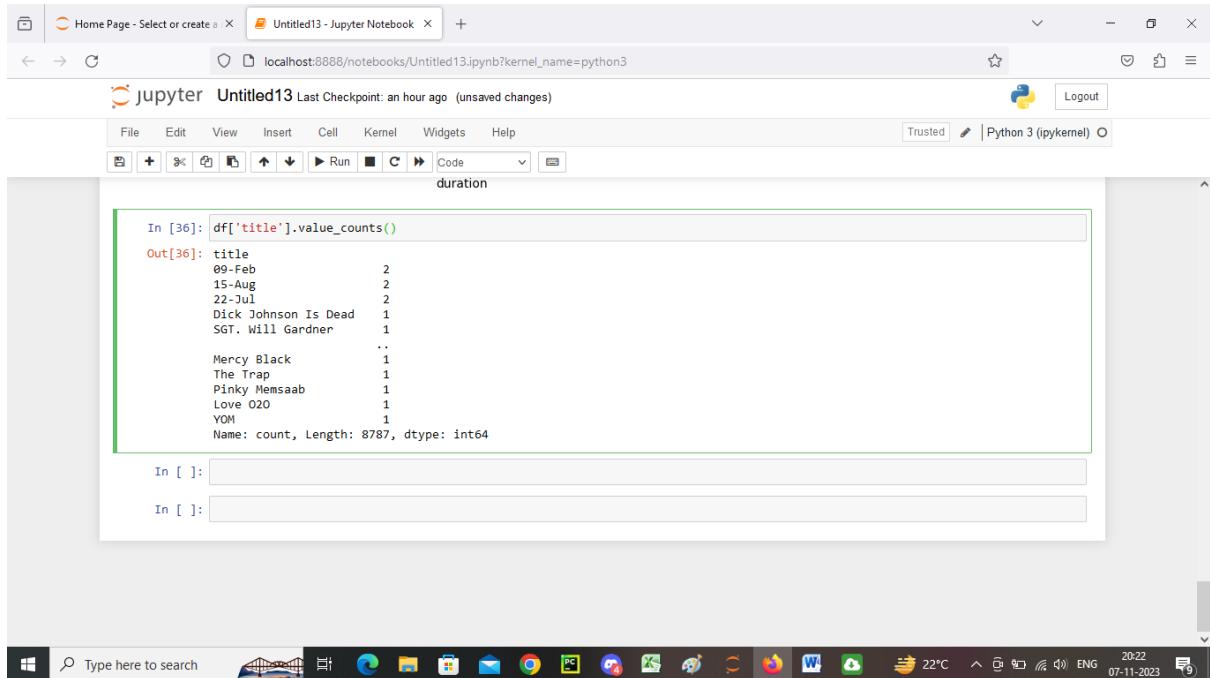
```
new_df[['rating','release_year']].plot(kind='scatter',x='rating',y='release_year',figsize=(5,5))
```



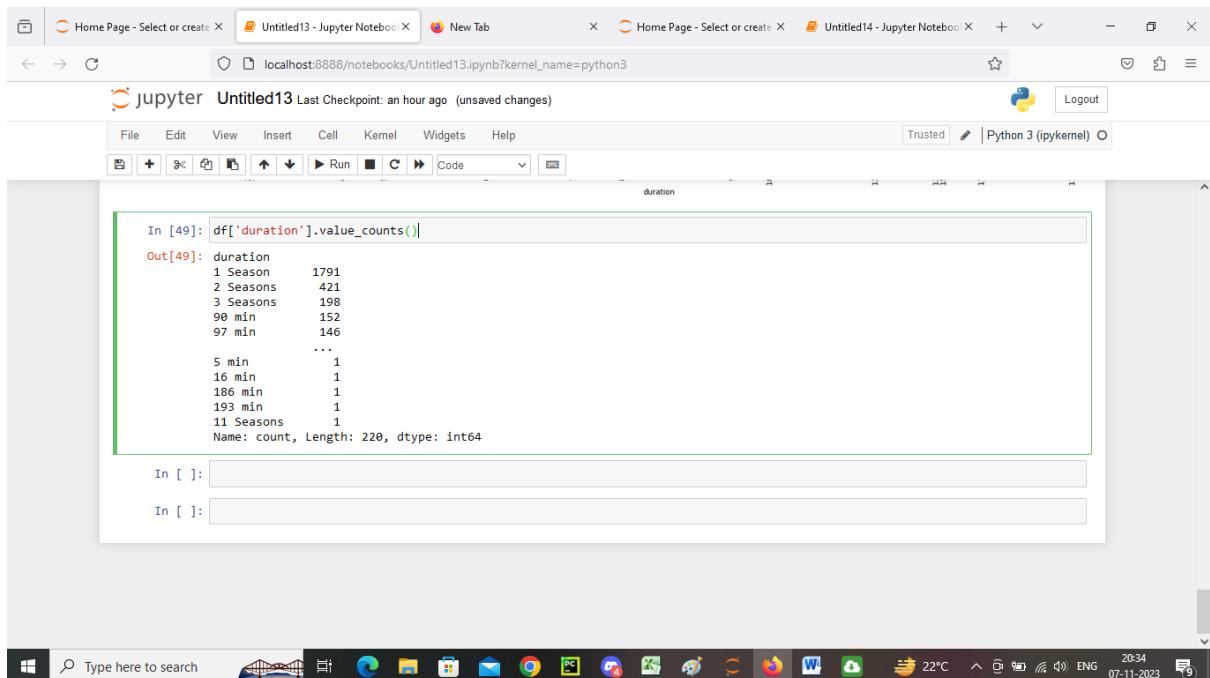
```
new_df[['duration','release_year']].plot(kind='scatter',x='duration',y='release_year',figsize=(5,5))
```



```
df['title'].value_counts()
```

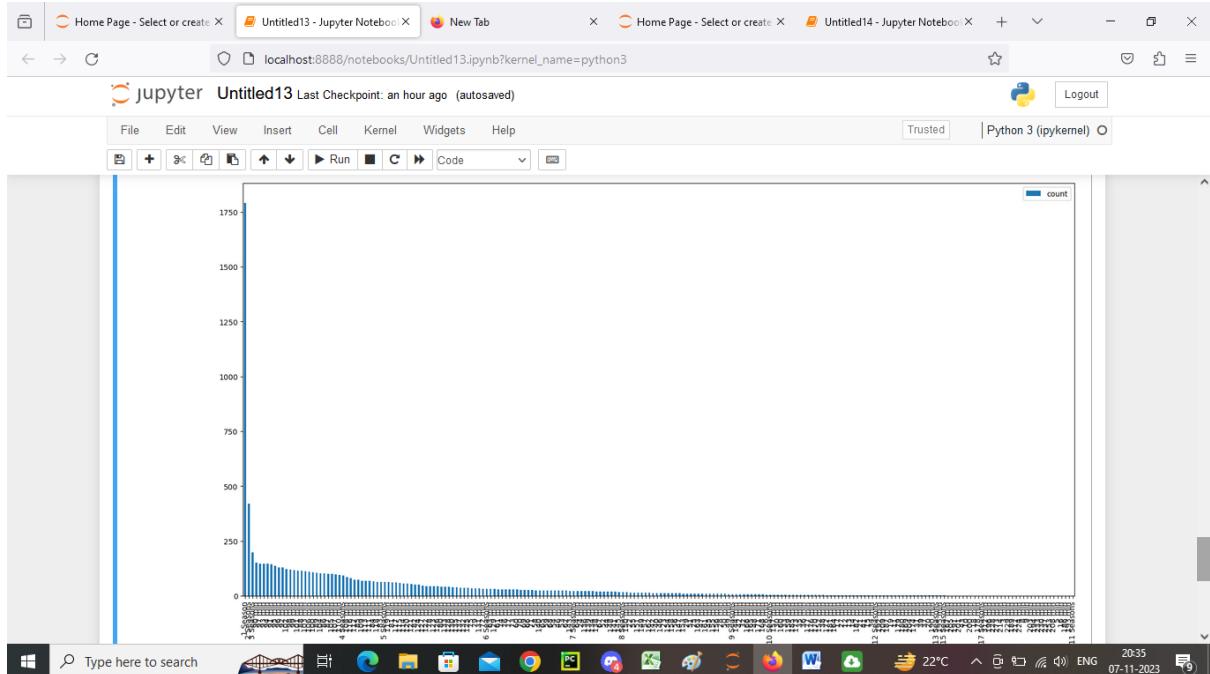


```
df['duration'].value_counts()
```



```
In [49]: df['duration'].value_counts()
Out[49]: duration
1 Season      1791
2 Seasons     421
3 Seasons     198
90 min        152
97 min        146
...
5 min          1
16 min         1
186 min        1
193 min        1
11 Seasons    1
Name: count, Length: 220, dtype: int64
```

```
pd.DataFrame(df['duration'].value_counts()).plot(kind='bar', figsize=(20,10))
```



```
df['title'].value_counts()
```

The screenshot shows a Jupyter Notebook interface running in a browser window. The title bar indicates the notebook is titled "Untitled13 - Jupyter Notebook". The code cell In [55] contains the command `df['title'].value_counts()`. The output cell Out[55] displays the following frequency distribution:

Title	Count
09-Feb	2
15-Aug	2
22-Jul	2
Dick Johnson Is Dead	1
SGT. Will Gardner	1
..	
Mercy Black	1
The Trap	1
Pinky Mensaab	1
Love O2O	1
YOM	1

Name: count, Length: 8787, dtype: int64

```
df['duration'].value_counts()
```

The screenshot shows a Jupyter Notebook interface running in a browser window. The title bar indicates the notebook is titled "Untitled13 - Jupyter Notebook". The code cell In [56] contains the command `df['duration'].value_counts()`. The output cell Out[56] displays the following frequency distribution:

Duration	Count
1 Season	1791
2 Seasons	421
3 Seasons	198
90 min	152
97 min	146
..	
5 min	1
16 min	1
186 min	1
193 min	1
11 Seasons	1

Name: count, Length: 220, dtype: int64

```
df['show_id'].value_counts()
```

The screenshot shows a Jupyter Notebook interface running on a Windows desktop. The notebook title is "Clean Data - Jupyter Notebook". The code cell In [10] contains the command `df['show_id'].value_counts()`. The output cell Out[10] displays the following frequency distribution:

```
Out[10]: show_id
s1           1
s7990        1
s7982        1
s7984        1
s7986        1
..
s3960        1
s3958        1
s3956        1
s3955        1
s8786        1
Name: count, Length: 8790, dtype: int64
```

```
df['type'].value_counts()
```

The screenshot shows a Jupyter Notebook interface running on a Windows desktop. The notebook title is "Clean Data - Jupyter Notebook". The code cell In [11] contains the command `df['type'].value_counts()`. The output cell Out[11] displays the following frequency distribution:

```
Out[11]: type
Movie       6126
TV Show    2664
Name: count, dtype: int64
```

```
df[df['title'].isna()]
```

The screenshot shows a Jupyter Notebook interface with the following code and its output:

```
In [4]: df[df['title'].isna()]
```

```
Out[4]:
```

show_id	type	title	director	country	release_year	rating	duration	listed_in
								date_added

```
In [5]: df['title'].value_counts()
```

```
Out[5]:
```

title	count
09-Feb	2
15-Aug	2
22-Jul	2
Dick Johnson Is Dead	1
SGT. Will Gardner	1
..	
Mercy Black	1
The Trap	1
Pinky Mensaab	1
Love O2O	1
YOM	1
Name: count, Length: 8787, dtype: int64	

```
In [6]: df.replace('09-Feb','Not Applicable' ,inplace=True)
```

```
Out[6]:
```

show_id	type	title	director	country	release_year	rating	duration	listed_in
								date_added

```
df.replace('09-Feb','Not Applicable' ,inplace=True)
```

```
df
```

The screenshot shows a Jupyter Notebook interface with the following code and its output:

```
In [6]: df.replace('09-Feb','Not Applicable' ,inplace=True)
```

```
Out[6]:
```

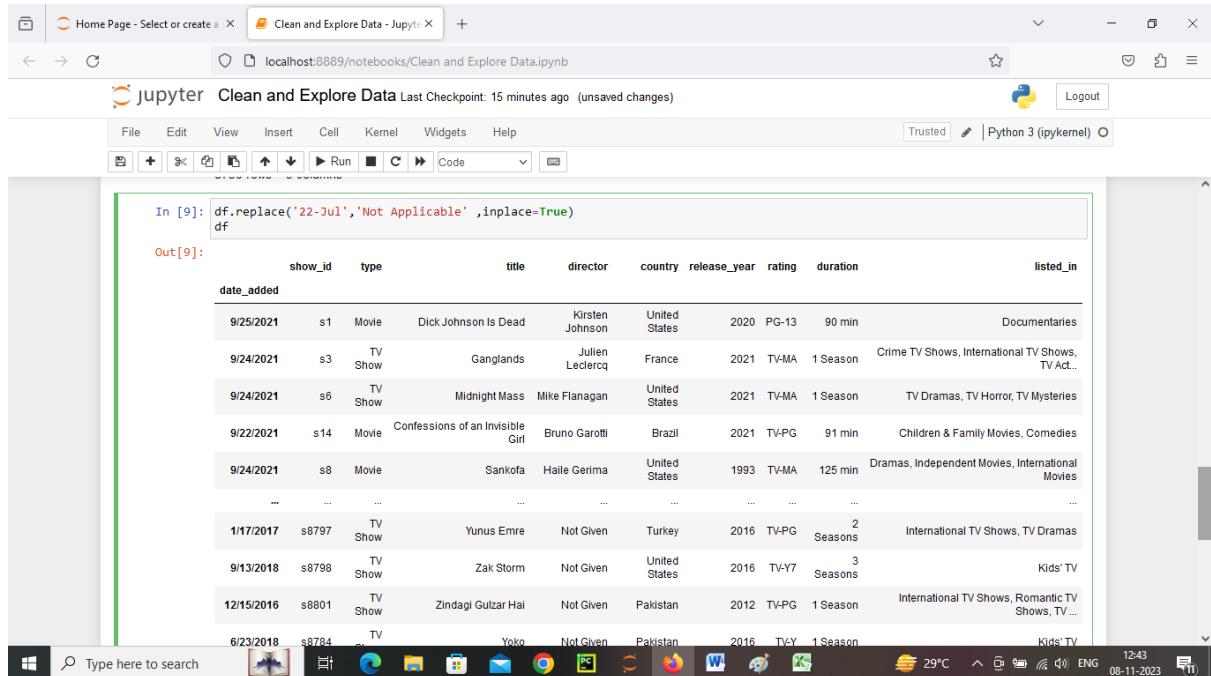
show_id	type	title	director	country	release_year	rating	duration	listed_in	
9/25/2021	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	2020	PG-13	90 min	Documentaries
9/24/2021	s3	TV Show	Ganglands	Julien Leclercq	France	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...
9/24/2021	s6	TV Show	Midnight Mass	Mike Flanagan	United States	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries
9/22/2021	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Brazil	2021	TV-PG	91 min	Children & Family Movies, Comedies
9/24/2021	s8	Movie	Sankofa	Halle Gerima	United States	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies
...	...	...	...	...	...	...	...	...	...
1/17/2017	s8797	TV Show	Yunus Emre	Not Given	Turkey	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas
9/13/2018	s8798	TV Show	Zak Storm	Not Given	United States	2016	TV-Y7	3 Seasons	Kids' TV
12/15/2016	s8801	TV Show	Zindagi Gulzar Hai	Not Given	Pakistan	2012	TV-PG	1 Season	International TV Shows, Romantic TV Shows, TV ...
6/23/2018	s8784	TV Show	Yoko	Not Given	Pakistan	2016	TV-Y	1 Season	Kids' TV

```
df.replace('15-Aug','Not Applicable' ,inplace=True)
```

```
df
```

```
df.replace('22-Jul','Not Applicable' ,inplace=True)
```

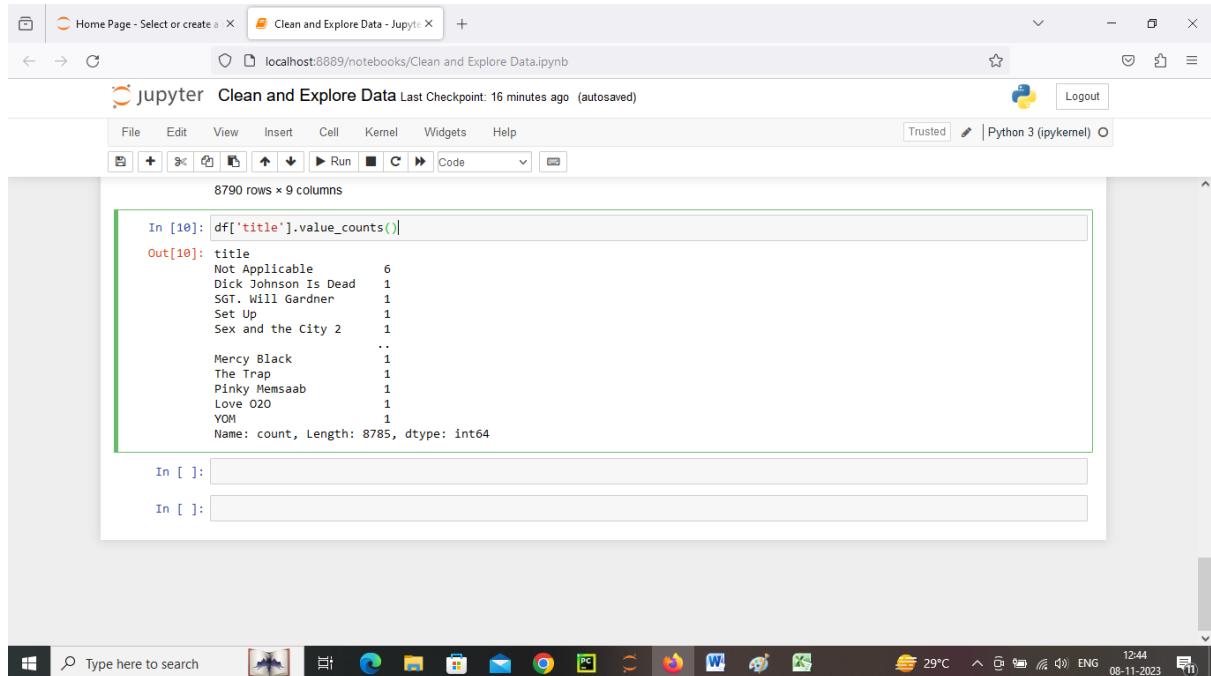
```
df
```



In [9]: df.replace('22-Jul','Not Applicable' ,inplace=True)  
Out[9]:

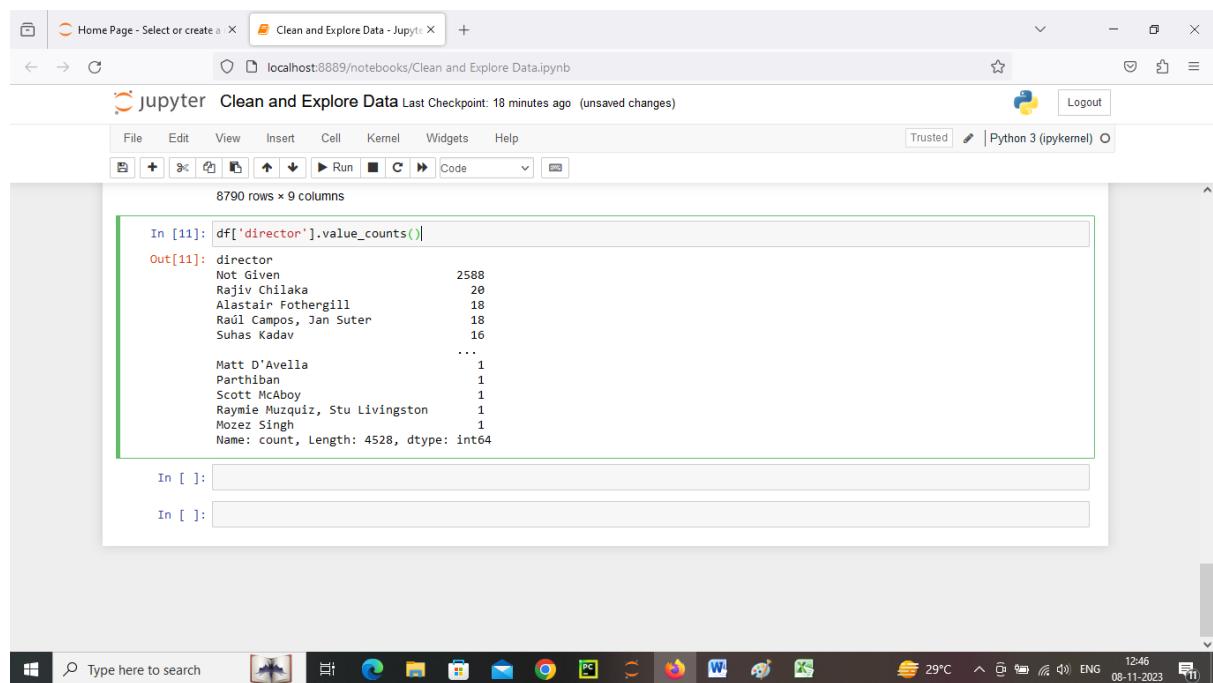
show_id	type	title	director	country	release_year	rating	duration	listed_in	
date_added									
9/25/2021	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	2020	PG-13	90 min	Documentaries
9/24/2021	s3	TV Show	Ganglands	Julien Ledercq	France	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...
9/24/2021	s6	TV Show	Midnight Mass	Mike Flanagan	United States	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries
9/22/2021	s14	Movie	Confessions of an Invisible Girl	Bruno Gattati	Brazil	2021	TV-PG	91 min	Children & Family Movies, Comedies
9/24/2021	s8	Movie	Sankofa	Halle Gerima	United States	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies
...	...	...	...	...	...	...	...	...	...
1/17/2017	s8797	TV Show	Yunus Emre	Not Given	Turkey	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas
9/13/2018	s8798	TV Show	Zak Storm	Not Given	United States	2016	TV-Y7	3 Seasons	Kids' TV
12/15/2016	s8801	TV Show	Zindagi Gulzar Hai	Not Given	Pakistan	2012	TV-PG	1 Season	International TV Shows, Romantic TV Shows, TV ...
6/23/2018	s8784	TV	Yoko	Not Given	Pakistan	2016	TV-Y	1 Season	Kids' TV

```
df['title'].value_counts()
```



In [10]: df['title'].value\_counts()  
Out[10]: title  
Not Applicable 6  
Dick Johnson Is Dead 1  
SGT. Will Gardner 1  
Set Up 1  
Sex and the City 2 1  
..  
Mercy Black 1  
The Trap 1  
Pinky Memsaab 1  
Love O2O 1  
YOM 1  
Name: count, Length: 8785, dtype: int64

```
df['director'].value_counts()
```



In [11]: `df['director'].value_counts()`

Out[11]:

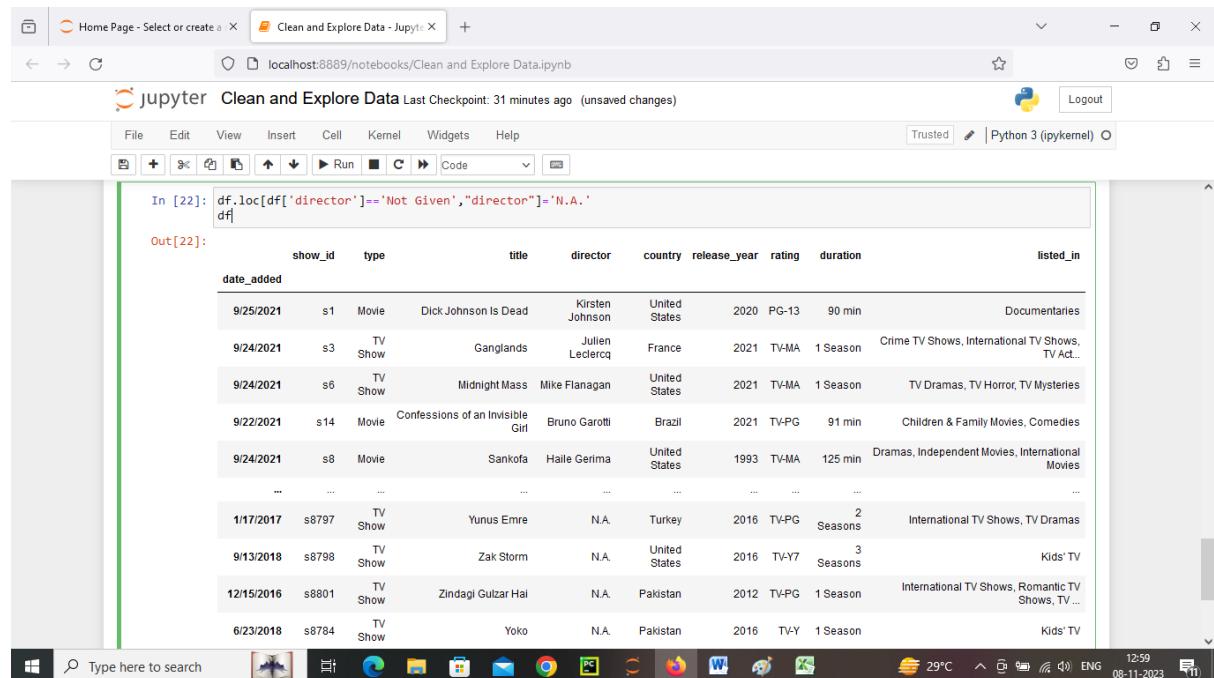
director	count
Not Given	2588
Rajiv Chilaka	20
Alastair Fothergill	18
Raul Campos, Jan Suter	18
Suhas Kadav	16
...	
Matt D'Avella	1
Parthiban	1
Scott McAboy	1
Raymie Muquiz, Stu Livingston	1
Mozez Singh	1

Name: count, Length: 4528, dtype: int64



```
df.loc[df['director']=='Not Given','director']='N.A.'
```

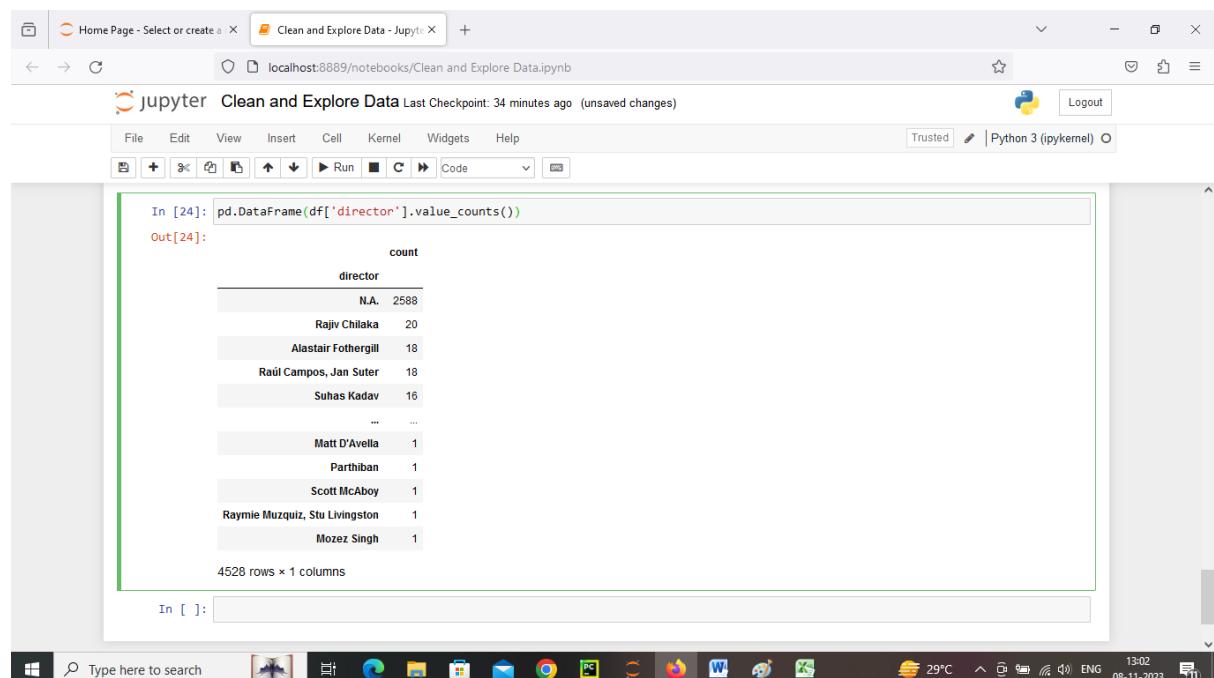
```
df
```



In [22]: `df.loc[df['director']=='Not Given','director']='N.A.'`  
Out[22]:

date_added	show_id	type	title	director	country	release_year	rating	duration	listed_in
9/25/2021	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	2020	PG-13	90 min	Documentaries
9/24/2021	s3	TV Show	Ganglands	Julien Ledercq	France	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...
9/24/2021	s6	TV Show	Midnight Mass	Mike Flanagan	United States	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries
9/22/2021	s14	Movie	Confessions of an Invisible Girl	Bruno Gattotti	Brazil	2021	TV-PG	91 min	Children & Family Movies, Comedies
9/24/2021	s8	Movie	Sankofa	Haille Gerima	United States	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies
...	...	...	...	...	...	...	...	...	...
1/17/2017	s8797	TV Show	Yunus Emre	N.A.	Turkey	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas
9/13/2018	s8798	TV Show	Zak Storm	N.A.	United States	2016	TV-Y7	3 Seasons	Kids' TV
12/15/2016	s8801	TV Show	Zindagi Gulzar Hai	N.A.	Pakistan	2012	TV-PG	1 Season	International TV Shows, Romantic TV Shows, TV...
6/23/2018	s8784	TV Show	Yoko	N.A.	Pakistan	2016	TV-Y	1 Season	Kids' TV

```
pd.DataFrame(df['director'].value_counts())
```



In [24]: `pd.DataFrame(df['director'].value_counts())`  
Out[24]:

director	count
N.A.	2588
Rajiv Chilaka	20
Alastair Fothergill	18
Raúl Campos, Jan Suter	18
Suhdas Kadav	16
...	...
Matt D'Avella	1
Parthibnan	1
Scott McAboy	1
Raymie Muzquiz, Stu Livingston	1
Mozez Singh	1

4528 rows × 1 columns

```
pd.DataFrame(df['show_id'].value_counts())
```

The screenshot shows a Jupyter Notebook interface running on a Windows desktop. The notebook is titled "Clean and Explore Data". In the code cell (In [25]), the command `pd.DataFrame(df['show_id'].value_counts())` is entered. The output (Out[25]) displays a DataFrame with one column, "count", showing the frequency of each show ID. The data is as follows:

show_id	count
s1	1
s7990	1
s7982	1
s7984	1
s7986	1
...	...
s3960	1
s3958	1
s3956	1
s3955	1
s8786	1

The notebook interface includes a toolbar with various icons for file operations, a status bar at the bottom showing the date and time (08-11-2023, 13:05), and a Python 3 (ipykernel) kernel indicator.

```
pd.DataFrame(df['type'].value_counts())
```

The screenshot shows a Jupyter Notebook interface running on a Windows desktop. The notebook is titled "Clean and Explore Data". In the code cell (In [26]), the command `pd.DataFrame(df['type'].value_counts())` is entered. The output (Out[26]) displays a DataFrame with one column, "count", showing the frequency of each show type. The data is as follows:

type	count
Movie	6126
TV Show	2664

The notebook interface includes a toolbar with various icons for file operations, a status bar at the bottom showing the date and time (08-11-2023, 13:06), and a Python 3 (ipykernel) kernel indicator.

```
pd.DataFrame(df['director'].value_counts())
```

The screenshot shows a Jupyter Notebook interface running on a Windows desktop. The code cell In [28] contains the command `pd.DataFrame(df['director'].value_counts())`. The output Out[28] displays a DataFrame with the following data:

director	count
N.A.	2588
Rajiv Chilaka	20
Alastair Fothergill	18
Raúl Campos, Jan Suter	18
Suhas Kadav	16
...	...
Matt D'Avella	1
Parthiban	1
Scott McAboy	1
Raymie Muzquiz, Stu Livingston	1
Mozez Singh	1

The notebook interface includes a toolbar with various icons, a menu bar with File, Edit, View, Insert, Cell, Kernel, Widgets, Help, and a status bar at the bottom showing the date and time.

```
pd.DataFrame(df['country'].value_counts())
```

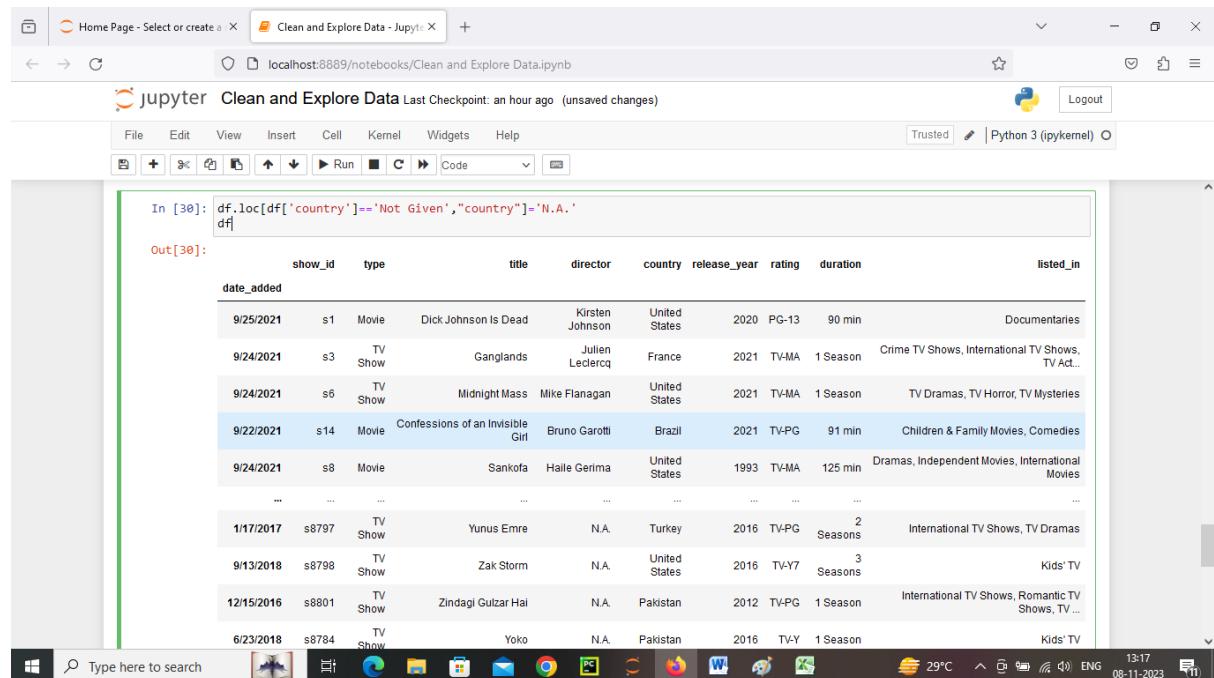
The screenshot shows a Jupyter Notebook interface running on a Windows desktop. The code cell In [28] contains the command `pd.DataFrame(df['country'].value_counts())`. The output Out[28] displays a DataFrame with the following data:

country	count
N.A.	2588
Rajiv Chilaka	20
Alastair Fothergill	18
Raúl Campos, Jan Suter	18
Suhas Kadav	16
...	...
Matt D'Avella	1
Parthiban	1
Scott McAboy	1
Raymie Muzquiz, Stu Livingston	1
Mozez Singh	1

The notebook interface includes a toolbar with various icons, a menu bar with File, Edit, View, Insert, Cell, Kernel, Widgets, Help, and a status bar at the bottom showing the date and time.

```
df.loc[df['country']=='Not Given',"country"]='N.A.'
```

```
df
```

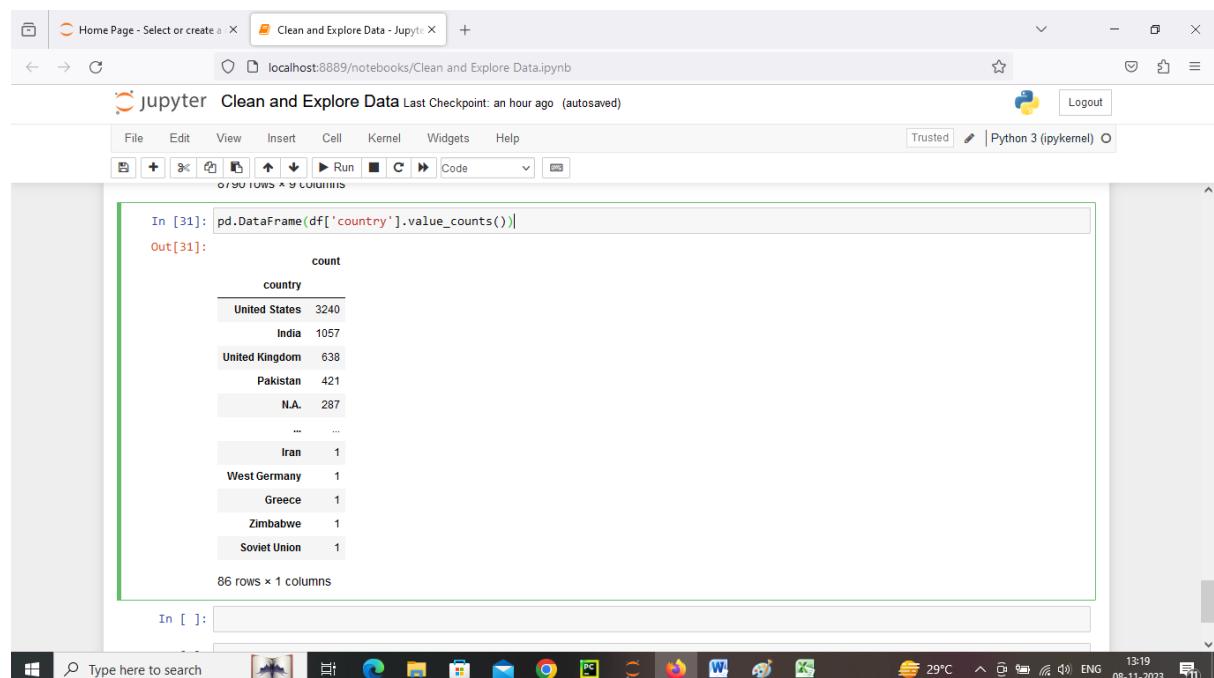


In [30]: df.loc[df['country']=='Not Given',"country"]='N.A.'  
df

Out[30]:

show_id	type	title	director	country	release_year	rating	duration	listed_in	
9/25/2021	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	2020	PG-13	90 min	Documentaries
9/24/2021	s3	TV Show	Ganglands	Julien Leclercq	France	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...
9/24/2021	s6	TV Show	Midnight Mass	Mike Flanagan	United States	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries
9/22/2021	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Brazil	2021	TV-PG	91 min	Children & Family Movies, Comedies
9/24/2021	s8	Movie	Sankofa	Halle Gerima	United States	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies
...	...	...	...	...	...	...	...	...	...
1/17/2017	s8797	TV Show	Yunus Emre	N.A.	Turkey	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas
9/13/2018	s8798	TV Show	Zak Storm	N.A.	United States	2016	TV-Y7	3 Seasons	Kids' TV
12/15/2016	s8801	TV Show	Zindagi Gulzar Hai	N.A.	Pakistan	2012	TV-PG	1 Season	International TV Shows, Romantic TV Shows, TV ...
6/23/2018	s8784	TV Show	Yoko	N.A.	Pakistan	2016	TV-Y	1 Season	Kids' TV

```
pd.DataFrame(df['country'].value_counts())
```



In [31]: pd.DataFrame(df['country'].value\_counts())  
Out[31]:

country	count
United States	3240
India	1057
United Kingdom	638
Pakistan	421
N.A.	287
...	...
Iran	1
West Germany	1
Greece	1
Zimbabwe	1
Soviet Union	1

86 rows × 1 columns

```
pd.DataFrame(df['release_year'].value_counts())
```

The screenshot shows a Jupyter Notebook interface running on a Windows desktop. The notebook title is "Clean and Explore Data - Jupyter". The code cell In [32] contains the command `pd.DataFrame(df['release_year'].value_counts())`. The output cell Out[32] displays a DataFrame with the following data:

release_year	count
2018	1146
2017	1030
2019	1030
2020	953
2016	901
...	...
1966	1
1959	1
1925	1
1947	1
1961	1

The output indicates 74 rows × 1 columns. Below this, another code cell In [33] is partially visible with the command `df.loc[df['country']=='Not Given','country']='N.A.'`.

```
pd.DataFrame(df['rating'].value_counts())
```

The screenshot shows a Jupyter Notebook interface running on a Windows desktop. The notebook title is "Clean and Explore Data - Jupyter". The code cell In [34] contains the command `pd.DataFrame(df['rating'].value_counts())`. The output cell Out[34] displays a DataFrame with the following data:

rating	count
TV-MA	3205
TV-14	2157
TV-PG	861
R	799
PG-13	490
TV-Y7	333
TV-Y	306
PG	287
TV-G	220
NR	79
G	41
TV-Y7-FV	6
NC-17	3
UR	3

```
pd.DataFrame(df['duration'].value_counts())
```

A screenshot of a Jupyter Notebook interface. The top bar shows tabs for 'Home Page - Select or create a...' and 'Clean and Explore Data - Jupyter'. The URL is 'localhost:8889/notebooks/Clean and Explore Data.ipynb'. The title bar says 'jupyter Clean and Explore Data Last Checkpoint: an hour ago (autosaved)'. The menu bar includes File, Edit, View, Insert, Cell, Kernel, Widgets, Help. The toolbar has icons for New, Open, Save, Run, Cell, Help, Code, and Cell Kernel. A status bar at the bottom right shows 'Notebook saved' and 'Trusted' with 'Python 3 (ipykernel)'.

In [36]: pd.DataFrame(df['duration'].value\_counts())

Out[36]:

duration	count
1 Season	1791
2 Seasons	421
3 Seasons	198
90 min	152
97 min	146
...	...
5 min	1
16 min	1
186 min	1
193 min	1
11 Seasons	1

220 rows × 1 columns

In [33]: df.loc[df['country']=='Not Given','country']='N.A.'

df

```
df.loc[df['duration']=='1 Season',"duration"]='90 min'
```

df

A screenshot of a Jupyter Notebook interface, identical to the one above but with a different code execution history. The top bar, title bar, and menu bar are the same. The status bar at the bottom right shows '30°C', '13:28 08-11-2023', and battery level.

In [37]: df.loc[df['duration']=='1 Season','duration']='90 min'

df

Out[37]:

show_id	type	title	director	country	release_year	rating	duration	listed_in	
9/25/2021	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	2020	PG-13	90 min	Documentaries
9/24/2021	s3	TV Show	Ganglands	Julien Leclercq	France	2021	TV-MA	90 min	Crime TV Shows, International TV Shows, TV Action & Adventure
9/24/2021	s6	TV Show	Midnight Mass	Mike Flanagan	United States	2021	TV-MA	90 min	TV Dramas, TV Horror, TV Mysteries
9/22/2021	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Brazil	2021	TV-PG	91 min	Children & Family Movies, Comedies
9/24/2021	s8	Movie	Sankofa	Halle Gerima	United States	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies
...	...	...	...	...	...	...	...	...	...
1/17/2017	s8797	TV Show	Yunus Emre	N.A.	Turkey	2016	TV-PG	2 seasons	International TV Shows, TV Dramas
9/13/2018	s8798	TV Show	Zak Storm	N.A.	United States	2016	TV-Y7	3 seasons	Kids' TV
12/15/2016	s8801	TV Show	Zindagi Gulzar Hai	N.A.	Pakistan	2012	TV-PG	90 min	International TV Shows, Romantic TV Shows, TV Dramas
...	...	...	...	...	...	...	...	...	...

```
pd.DataFrame(df['duration'].value_counts())
```

Jupyter Clean and Explore Data Last Checkpoint: an hour ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel) O

In [39]: pd.DataFrame(df['duration'].value\_counts())

Out[39]:

duration	count
90 min	1943
2 Seasons	421
3 Seasons	198
94 min	146
97 min	146
...	...
5 min	1
16 min	1
186 min	1
193 min	1
11 Seasons	1

219 rows x 1 columns

Type here to search 30°C 13:31 08-11-2023

df.loc[df['duration']=='2 Seasons',"duration"]='90 min'

df

Jupyter Clean and Explore Data Last Checkpoint: an hour ago (unsaved changes)

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel) O

In [41]: df.loc[df['duration']=='2 Seasons',"duration"]='90 min'  
df

Out[41]:

show_id	type	title	director	country	release_year	rating	duration	listed_in	
9/25/2021	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	2020	PG-13	90 min	Documentaries
9/24/2021	s3	TV Show	Ganglands	Julien Ledercq	France	2021	TV-MA	90 min	Crime TV Shows, International TV Shows, TV Act...
9/24/2021	s6	TV Show	Midnight Mass	Mike Flanagan	United States	2021	TV-MA	90 min	TV Dramas, TV Horror, TV Mysteries
9/22/2021	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Brazil	2021	TV-PG	91 min	Children & Family Movies, Comedies
9/24/2021	s8	Movie	Sankofa	Halle Gerima	United States	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies
...	...	...	...	...	...	...	...	...	...
1/17/2017	s8797	TV Show	Yunus Emre	N.A.	Turkey	2016	TV-PG	90 min	International TV Shows, TV Dramas
9/13/2018	s8798	TV Show	Zak Storm	N.A.	United States	2016	TV-Y7	3 Seasons	Kids' TV
12/15/2016	s8801	TV Show	Zindagi Gulzar Hai	N.A.	Pakistan	2012	TV-PG	90 min	International TV Shows, Romantic TV Shows, TV ...

Type here to search 30°C 13:34 08-11-2023

```
df.loc[df['duration']=='3 Seasons',"duration"]='97 min'
```

df

The screenshot shows a Jupyter Notebook interface with the following details:

- In [43]:** df.loc[df['duration']=='3 Seasons',"duration"]='97 min'  
df
- Out[43]:** A DataFrame table with the following columns: show\_id, type, title, director, country, release\_year, rating, duration, and listed\_in.
- The table contains 10 rows of data, each representing a media item with its details like title, director, and genre listed in the listed\_in column.
- The bottom status bar shows the date as 08-11-2023 and the time as 13:36.

```
df.loc[df['duration']=='11 Seasons',"duration"]='193 min'
```

df

The screenshot shows a Jupyter Notebook interface with the following details:

- In [44]:** df.loc[df['duration']=='11 Seasons',"duration"]='193 min'  
df
- Out[44]:** A DataFrame table with the same structure as Out[43], showing the same 10 rows of data.
- The bottom status bar shows the date as 08-11-2023 and the time as 13:37.

```
pd.DataFrame(df['duration'].value_counts())
```

In [45]: `pd.DataFrame(df['duration'].value_counts())`

Out[45]:

duration	count
90 min	2364
97 min	344
93 min	146
94 min	146
91 min	144
...	...
230 min	1
237 min	1
253 min	1
5 min	1
233 min	1

216 rows × 1 columns

In [44]: `df.loc[df['duration']=='11 Seasons','duration']='193 min'`  
df

```
df.loc[df['listed_in']=='Dramas, Independent Movies, International  
Movies','listed_in']='Dramas, International Movies'
```

```
df
```

In [49]: `df.loc[df['listed_in']=='Dramas, Independent Movies, International Movies','listed_in']='Dramas, International Movies'`  
df

Out[49]:

show_id	type	title	director	country	release_year	rating	duration	listed_in	
date_added									
9/25/2021	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	2020	PG-13	90 min	Documentaries
9/24/2021	s3	TV Show	Ganglands	Julien Ledercq	France	2021	TV-MA	90 min	Crime TV Shows, International TV Shows, TV Act...
9/24/2021	s6	TV Show	Midnight Mass	Mike Flanagan	United States	2021	TV-MA	90 min	TV Dramas, TV Horror, TV Mysteries
9/22/2021	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Brazil	2021	TV-PG	91 min	Children & Family Movies, Comedies
9/24/2021	s8	Movie	Sankofa	Haile Gerima	United States	1993	TV-MA	125 min	Dramas, International Movies
...	...	...	...	...	...	...	...	...	...
1/17/2017	s8797	TV Show	Yunus Emre	N.A.	Turkey	2016	TV-PG	90 min	International TV Shows, TV Dramas
9/13/2018	s8798	TV Show	Zak Storm	N.A.	United States	2016	TV-Y7	97 min	Kids' TV
12/15/2016	s8801	TV Show	Zindagi Gulzar Hai	N.A.	Pakistan	2012	TV-PG	90 min	International TV Shows, Romantic TV Shows, TV ...
6/23/2018	s8784	TV Show	Yoko	N.A.	Pakistan	2016	TV-Y	90 min	Kids' TV

```
pd.DataFrame(df['listed_in'].value_counts())
```

A screenshot of a Jupyter Notebook interface. The title bar shows "Home Page - Select or create a..." and "Clean and Explore Data - Jupyter". The notebook tab is titled "localhost:8889/notebooks/Clean and Explore Data.ipynb". The toolbar includes File, Edit, View, Insert, Cell, Kernel, Widgets, Help, and a Trusted Python 3 (ipykernel) button. Below the toolbar is a toolbar with various icons for cell operations like Run, Insert, Delete, and Cell Type. The main area shows the output of a cell:

```
In [51]: pd.DataFrame(df['listed_in'].value_counts())
Out[51]:
   listed_in    count
0  Dramas, International Movies    888
1  Documentaries                 359
2  Stand-Up Comedy                334
3    Kids' TV                      219
4 Children & Family Movies        215
.. ...
9 Docuseries, Reality TV Shows      1
10 Classic Movies, Independent Movies, Thrillers 1
11 Children & Family Movies, Faith & Spirituality 1
12 Classic Movies, Comedies, Sports Movies 1
13 Classic & Cult TV, Crime TV Shows, TV Dramas 1
```

The output indicates 511 rows by 1 column. Below the output, another cell is partially visible:

```
In [44]: df.loc[df['duration']=='11 Seasons','duration']='193 min'
```

```
df.loc[df['listed_in']=='Comedies, Cult Movies, Sports Movies',"listed_in"]='Stand-Up Comedy'
```

```
df
```

A screenshot of a Jupyter Notebook interface, identical to the one above but with a different cell content. The title bar and toolbar are the same. The main area shows the output of a cell:

```
In [66]: df.loc[df['listed_in']=='Comedies, Cult Movies, Sports Movies',"listed_in"]='Stand-Up Comedy'
```

The output displays the modified DataFrame 'df' with the 'listed\_in' column updated. The table includes columns: show\_id, type, title, director, country, release\_year, rating, duration, and listed\_in. The 'date\_added' column is also present. The 'listed\_in' column values are now either 'Documentaries', 'Crime TV Shows, International TV Shows, TV Act...', 'TV Dramas, TV Horror, TV Mysteries', 'Children & Family Movies, Comedies', 'Dramas, International Movies', 'International TV Shows, TV Dramas', 'Kids' TV', 'International TV Shows, Romantic TV Shows, TV ...', or 'Kids' TV'. The 'duration' column has been updated for some rows.

```
pd.DataFrame(df['listed_in'].value_counts())
```

The screenshot shows a Jupyter Notebook interface with the title "Clean and Explore Data". The code cell In [67] contains the command `pd.DataFrame(df['listed_in'].value_counts())`. The output Out[67] displays a DataFrame with the following data:

listed_in	count
Dramas, International Movies	889
Documentaries	359
Stand-Up Comedy	335
Kids' TV	219
Children & Family Movies	216
...	...
British TV Shows, Classic & Cult TV, Kids' TV	1
Action & Adventure, Sci-Fi & Fantasy, Sports Movies	1
Cult Movies, Dramas, International Movies	1
Crime TV Shows, TV Horror, TV Mysteries	1
Classic & Cult TV, Crime TV Shows, TV Dramas	1

The DataFrame has 505 rows by 1 column. Below this, another code cell In [68] shows the command `df.loc[df['duration']=='11 Seasons','duration']='193 min'`, and its output Out[68] shows the modified DataFrame.

```
df.isna().sum()
```

The screenshot shows a Jupyter Notebook interface with the title "Untitled15". The code cell In [2] contains the command `df.isna().sum()`. The output Out[2] displays the sum of missing values for each column:

	show_id	type	title	director	country	release_year	rating	duration	listed_in	dtype: int64
	0	0	0	0	0	0	0	0	0	0

The DataFrame has 8790 rows by 11 columns. Below this, another code cell In [ ] is present.