

PHASE 1 :

CLEAN THE DATASET BY REPLACING VALUES

STEP 1: import pandas as pd

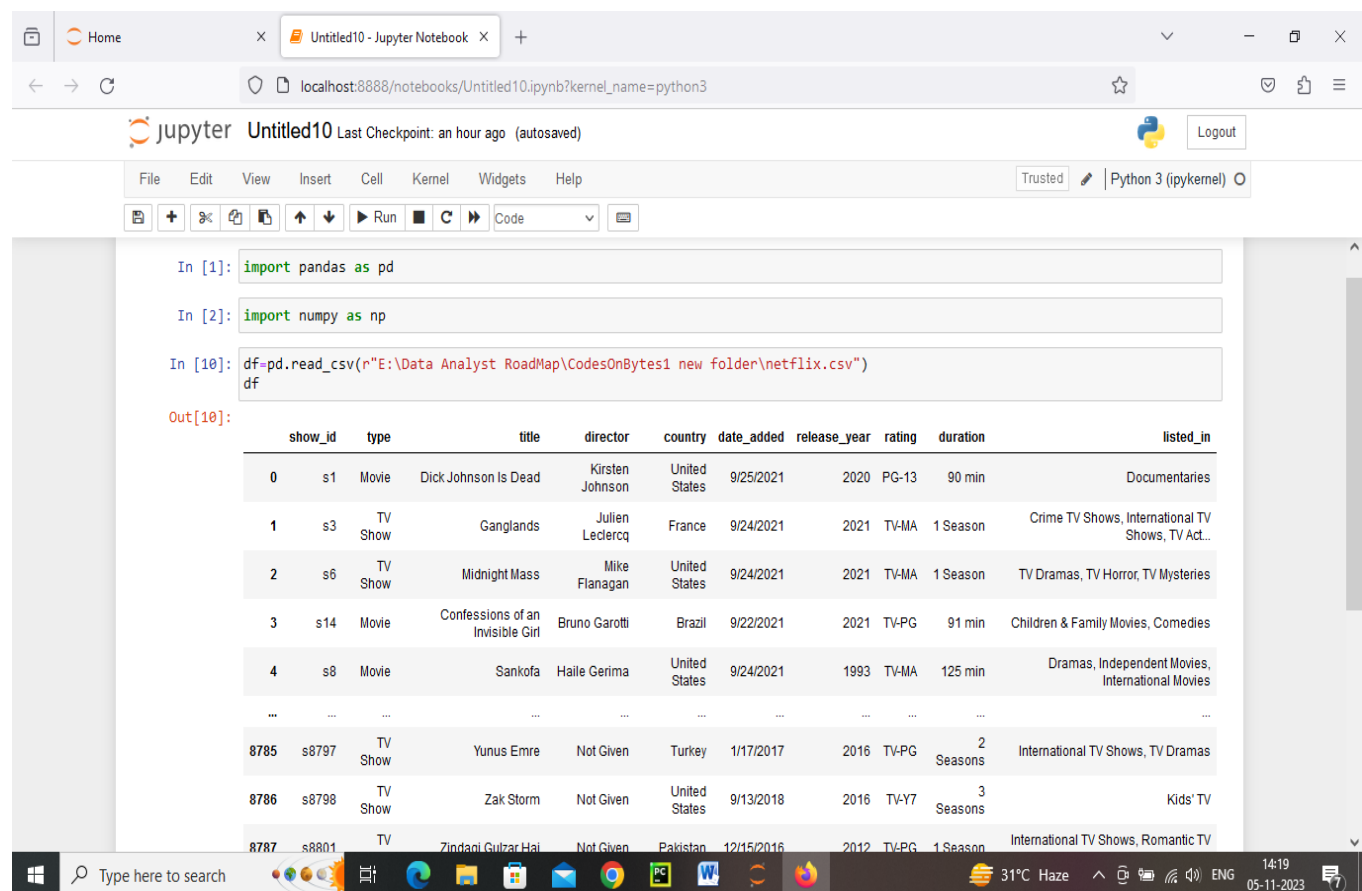
STEP 2: import numpy as np

STEP 3: Read CSV File with following below syntax

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

df

OUTPUT:



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

```
In [1]: import pandas as pd
```

```
In [2]: import numpy as np
```

```
In [10]: df=pd.read_csv(r"E:\Data Analyst RoadMap\CodesOnBytes1 new folder\netflix.csv")
df
```

Out[10]:

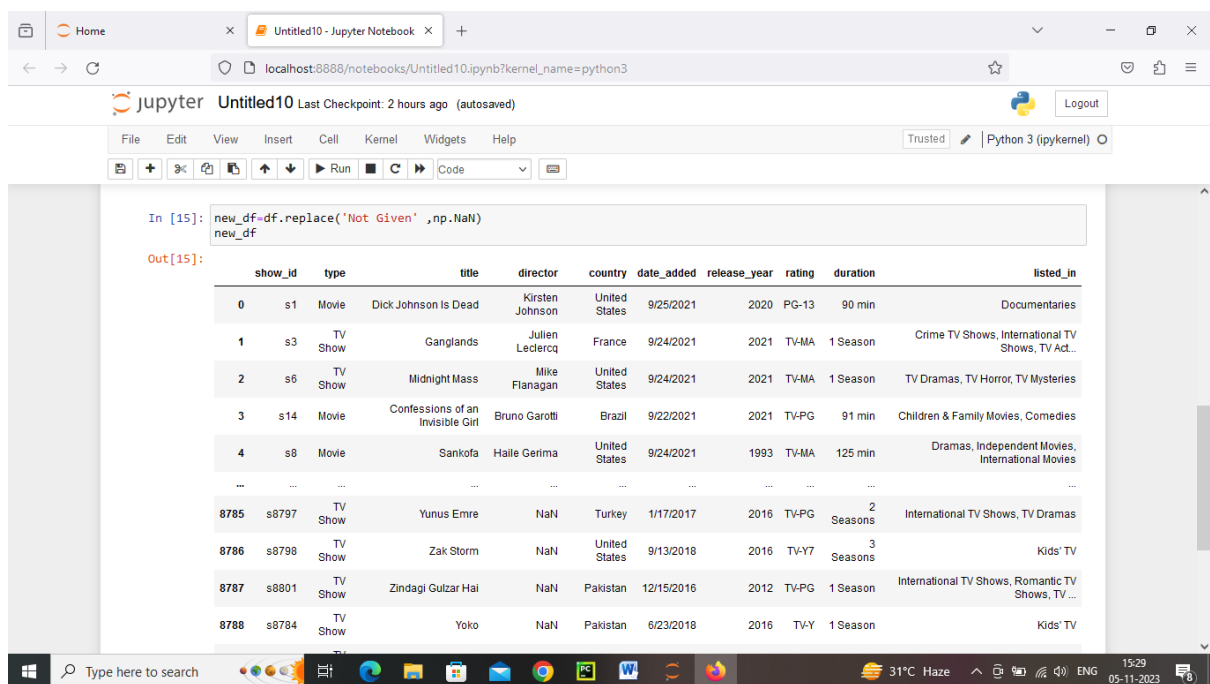
	show_id	type	title	director	country	date_added	release_year	rating	duration	listed_in
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	9/25/2021	2020	PG-13	90 min	Documentaries
1	s3	TV Show	Ganglands	Julien Leclercq	France	9/24/2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...
2	s6	TV Show	Midnight Mass	Mike Flanagan	United States	9/24/2021	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries
3	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Brazil	9/22/2021	2021	TV-PG	91 min	Children & Family Movies, Comedies
4	s8	Movie	Sankofa	Haile Gerima	United States	9/24/2021	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies
...
8785	s8797	TV Show	Yunus Emre	Not Given	Turkey	1/17/2017	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas
8786	s8798	TV Show	Zak Storm	Not Given	United States	9/13/2018	2016	TV-Y7	3 Seasons	Kids' TV
8787	s8801	TV	Zindagi Gulzar Hai	Not Given	Pakistan	12/15/2016	2012	TV-PG	1 Season	International TV Shows, Romantic TV

STEP 4: Replace Not Given values to NaN by Syntax

```
new_df=df.replace("Not Given",np.NaN)
```

```
new_df
```

OUTPUT:



In [15]: new_df=df.replace('Not Given', np.NaN)
new_df

Out[15]:

	show_id	type	title	director	country	date_added	release_year	rating	duration	listed_in
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	9/25/2021	2020	PG-13	90 min	Documentaries
1	s3	TV Show	Ganglands	Julien Leclercq	France	9/24/2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...
2	s6	TV Show	Midnight Mass	Mike Flanagan	United States	9/24/2021	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries
3	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Brazil	9/22/2021	2021	TV-PG	91 min	Children & Family Movies, Comedies
4	s8	Movie	Sankofa	Haile Gerima	United States	9/24/2021	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies
...
8785	s8797	TV Show	Yunus Emre	NaN	Turkey	1/17/2017	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas
8786	s8798	TV Show	Zak Storm	NaN	United States	9/13/2018	2016	TV-Y7	3 Seasons	Kids' TV
8787	s8801	TV Show	Zindagi Gulzar Hai	NaN	Pakistan	12/15/2016	2012	TV-PG	1 Season	International TV Shows, Romantic TV Shows, TV ...
8788	s8784	TV Show	Yoko	NaN	Pakistan	6/23/2018	2016	TV-Y	1 Season	Kids' TV
...

STEP 5: Set Indexing Date column at first

```
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
```

OUTPUT:

The screenshot shows a Jupyter Notebook window titled 'Untitled10 - Jupyter Notebook' running on a local host. The code in the notebook imports pandas and numpy, then reads a CSV file from a local path and sets the 'date_added' column as the index. The output is a table with columns: date_added, show_id, type, title, director, country, release_year, rating, duration, and listed_in.

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	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