MDA-710 Fall 2022

Homework 2

Intro to Programming II & Pandas & Handling Categorical Data

Due: September 29th @ 12AM
Submit on Blackboard

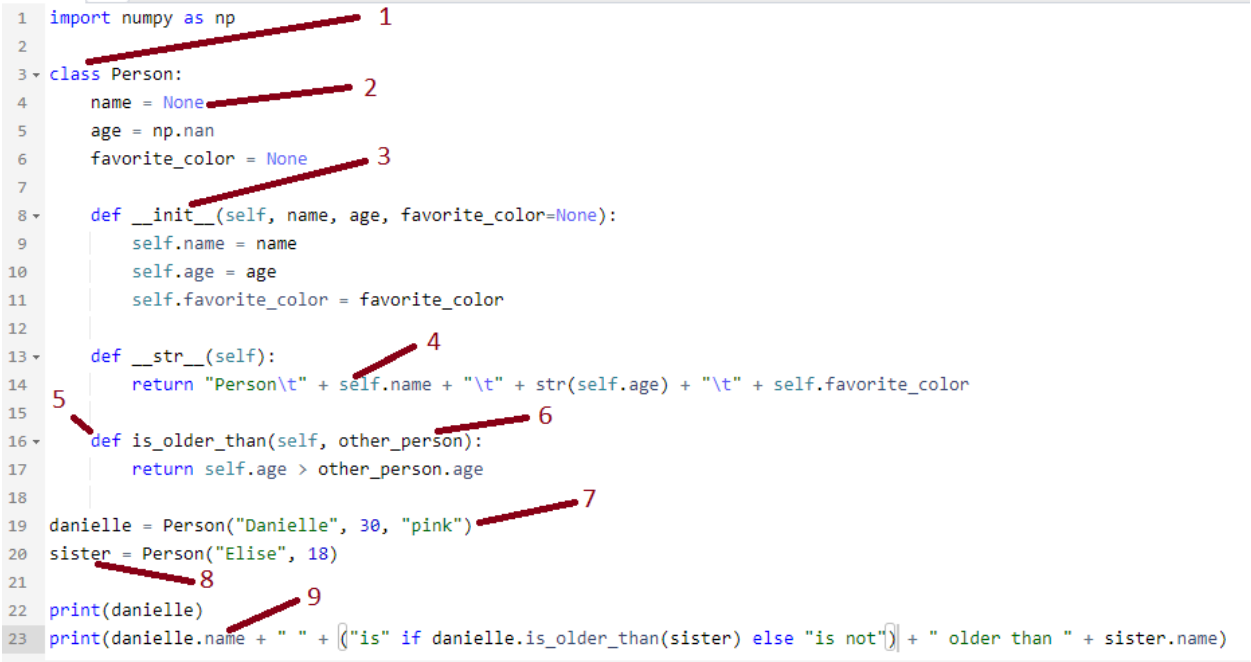
Earnable Points: 100pts
Base Points: 5ts
You can miss 5 points and still get a perfect score on this assignment

Programming II - Classes

Match the numbered parts in the image with the following terms/phrases. Only use each lettered term/phrase once. Each match is worth 2 points.

A. Keyword used to define a class
B. Parameter
C. Defining a Property
D. Used within a class to reference the current object
E. Object
F. Accessing a property
G. Creating an object/instance
H. Constructor
I. Keyword used to define a function

1. A
2. C
3. H
4. D
5. I
6. B
7. G
8. E
9. F




```
1  import numpy as np
2
3  class Person:
4      name = None
5      age = np.nan
6      favorite_color = None
7
8      def __init__(self, name, age, favorite_color=None):
9          self.name = name
10         self.age = age
11         self.favorite_color = favorite_color
12
13     def __str__(self):
14         return "Person\t" + self.name + "\t" + str(self.age) + "\t" + self.favorite_color
15
16     def is_older_than(self, other_person):
17         return self.age > other_person.age
18
19 danielle = Person("Danielle", 30, "pink")
20 sister = Person("Elise", 18)
21
22 print(danielle)
23 print(danielle.name + " " + ("is" if danielle.is_older_than(sister) else "is not") + " older than " + sister.name)
```

Pandas

For each question, answer with the code that would perform the given task. Add a screenshot of the output. Each question is worth 3 points except the last question, which is 5 points. (50 total points)

Using pandas, import the dataset “netflix_titles” as a dataframe and

1. Display the first 6 rows.

```
import pandas as pd
df = pd.read_csv('netflix_titles.csv')
```

0.5s Python

```
# 1.  Display the first 6 rows.
df.head(6)
```

0.2s Python

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, filmm...
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lor...
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	1 Season	Docuseries, Reality TV	Feuds, flirtations and toilet talk go down amo...
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, Romantic TV Shows, TV ...	In a city of coaching centers known to train I...
5	s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach Gilford, Hamish Linklater, H...	NaN	September 24, 2021	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries	The arrival of a charismatic young priest brin...

2. Display rows 10 through 20

```
# 2.  Display rows 10 through 20
df[10:21]
```

0.2s Python

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
10	s11	TV Show	Vendetta: Truth, Lies and The Mafia	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, Docuseries, International TV S...	Sicily boasts a bold "Anti-Mafia" coalition. B...
11	s12	TV Show	Bangkok Breaking	Kongkiat Komesiri	Sukollawat Kanarot, Sushar Manaying, Pavarit M...	NaN	September 23, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...	Struggling to earn a living in Bangkok, a man ...
12	s13	Movie	Je Suis Karl	Christian Schwochow	Luna Wedler, Jannis Niewöhner, Milan Peschel, ...	Germany, Czech Republic	September 23, 2021	2021	TV-MA	127 min	Dramas, International Movies	After most of her family is murdered in a terr...
13	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Klara Castanho, Lucca Picon, Júlia Gomes, Marc...	NaN	September 22, 2021	2021	TV-PG	91 min	Children & Family Movies, Comedies	When the clever but socially-awkward Tetê join...
14	s15	TV Show	Crime Stories: India Detectives	NaN	NaN	NaN	September 22, 2021	2021	TV-MA	1 Season	British TV Shows, Crime TV Shows, Docuseries	Cameras following Bengaluru police on the job ...
15	s16	TV Show	Dear White People	NaN	Logan Browning, Brandon P. Bell, DeRon Horton,...	United States	September 22, 2021	2021	TV-MA	4 Seasons	TV Comedies, TV Dramas	Students of color navigate the daily slights a...
16	s17	Movie	Europe's Most Dangerous Man: Otto Skorzeny in ...	Pedro de Echave García, Pablo Azorín Williams	NaN	NaN	September 22, 2021	2020	TV-MA	67 min	Documentaries, International Movies	Declassified documents reveal the post-WWII li...
17	s18	TV Show	Falsa identidad	NaN	Luis Ernesto Franco, Camila Sodi, Sergio Goyri...	Mexico	September 22, 2021	2020	TV-MA	2 Seasons	Crime TV Shows, Spanish-Language TV Shows, TV ...	Strangers Diego and Isabel flee their home in ...
18	s19	Movie	Intrusion	Adam Salky	Freida Pinto, Logan Marshall-Green, Robert Joh...	NaN	September 22, 2021	2021	TV-14	94 min	Thrillers	After a deadly home invasion at a couple's new...
19	s20	TV Show	Jaguar	NaN	Blanca Suárez, Iván Marcos, Óscar Casas, Adriá...	NaN	September 22, 2021	2021	TV-MA	1 Season	International TV Shows, Spanish-Language TV Sh...	In the 1960s, a Holocaust survivor joins a gro...
20	s21	TV Show	Monsters Inside: The 24 Faces of Billy Milligan	Olivier Megaton	NaN	NaN	September 22, 2021	2021	TV-14	1 Season	Crime TV Shows, Docuseries, International TV S...	In the late 1970s, an accused serial rapist cl...

3. Display rows 10 through 20 and columns 1 through 2

```
# 3.  Display rows 10 through 20 and columns 1 through 2
df[10:21][["show_id", "type"]]
```

✓ 0.2s

	show_id	type
10	s11	TV Show
11	s12	TV Show
12	s13	Movie
13	s14	Movie
14	s15	TV Show
15	s16	TV Show
16	s17	Movie
17	s18	TV Show
18	s19	Movie
19	s20	TV Show
20	s21	TV Show

4. Display the names of all the columns

```
# 4.  Display the names of all the columns
for col in df.columns:
    print(col)
```

✓ 0.1s

```
show_id
type
title
director
cast
country
date_added
release_year
rating
duration
listed_in
description
```

5. Display the values only for the column ‘title’ for rows 10 through 20

```
# 5.  Display the values only for the column ‘title’ for rows 10 through 20
df[10:21]["title"]
```

✓ 0.2s

```
10                  Vendetta: Truth, Lies and The Mafia
11                                      Bangkok Breaking
12                                          Je Suis Karl
13                     Confessions of an Invisible Girl
14                      Crime Stories: India Detectives
15                                     Dear White People
16    Europe's Most Dangerous Man: Otto Skorzeny in ...
17                                      Falsa identidad
18                                            Intrusion
19                                               Jaguar
20      Monsters Inside: The 24 Faces of Billy Milligan
Name: title, dtype: object
```

6. Display the standard deviation for the column ‘release_year’

```
# 6.  Display the standard deviation for the column ‘release_year’
df.std()["release_year"]
```

✓ 0.3s

```
/var/folders/57/h9hkt3s54t1cljt_qy48dj540000gn/T/ipykernel_41947/348894
version, it will default to False. In addition, specifying 'numeric_onl
warning.
  df.std()["release_year"]

8.819312130833966
```

7. Remove the ‘description’ column and show the results

```
# 7. Remove the ‘description’ column and show the results
df.drop(columns=["description"], inplace=False)
```

0.3s Python

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90 min	Documentaries
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	1 Season	Docuseries, Reality TV
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, Romantic TV Shows, TV ...
...	...	...	...	...	...	...	...	...	...	...	...
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	R	158 min	Cult Movies, Dramas, Thrillers
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	July 1, 2019	2018	TV-Y7	2 Seasons	Kids' TV, Korean TV Shows, TV Comedies
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009	R	88 min	Comedies, Horror Movies
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	PG	88 min	Children & Family Movies, Comedies
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	2015	TV-14	111 min	Dramas, International Movies, Music & Musicals

8807 rows × 11 columns

8. Display all the unique values for ‘type’

```
# 8.  Display all the unique values for 'type'
df["type"].unique()
```
✓ 0.2s

```
array(['Movie', 'TV Show'], dtype=object)
```

9. Display all the unique values for 'rating'

```
# 9.  Display all the unique values for 'rating'
df["rating"].unique()
```
✓ 0.6s

```
array(['PG-13', 'TV-MA', 'PG', 'TV-14', 'TV-PG', 'TV-Y', 'TV-Y7', 'R',
       'TV-G', 'G', 'NC-17', '74 min', '84 min', '66 min', 'NR', nan,
       'TV-Y7-FV', 'UR'], dtype=object)
```

10. Display all rows for which the type is 'Movie'

```
# 10. Display all rows for which the type is 'Movie'
df[df["type"] == "Movie"]
```
[22] ✓ 0.4s Python

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, filmm...
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	NaN	September 24, 2021	2021	PG	91 min	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D...	United States, Ghana, Burkina Faso, United Kin...	September 24, 2021	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies	On a photo shoot in Ghana, an American model s...
9	s10	Movie	The Starling	Theodore Melfi	Melissa McCarthy, Chris O'Dowd, Kevin Kline, T...	United States	September 24, 2021	2021	PG-13	104 min	Comedies, Dramas	A woman adjusting to life after a loss contend...
12	s13	Movie	Je Suis Karl	Christian Schwochow	Luna Wedler, Jannis Niewöhner, Milan Peschel, ...	Germany, Czech Republic	September 23, 2021	2021	TV-MA	127 min	Dramas, International Movies	After most of her family is murdered in a terr...
...	...	...	...	...	...	...	...	...	...	...	...	...
8801	s8802	Movie	Zinzana	Majid Al Ansari	Ali Suliman, Saleh Bakri, Yasa, Ali Al-Jabri, ...	United Arab Emirates, Jordan	March 9, 2016	2015	TV-MA	96 min	Dramas, International Movies, Thrillers	Recovering alcoholic Talal wakes up inside a s...
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	R	158 min	Cult Movies, Dramas, Thrillers	A political cartoonist, a crime reporter and a...
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009	R	88 min	Comedies, Horror Movies	Looking to survive in a world taken over by zo...
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	PG	88 min	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	2015	TV-14	111 min	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...

6131 rows × 12 columns

11. Display all row for which the type is 'Movie' and the rating is 'R'

```
# 11. Display all row for which the type is 'Movie' and the rating is 'R'
df[(df["type"] == "Movie") & (df["rating"] == "R")]
```
[23] ✓ 0.3s Python

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
46	s47	Movie	Safe House	Daniel Espinosa	Denzel Washington, Ryan Reynolds, Vera Farmiga,...	South Africa, United States, Japan	September 16, 2021	2012	R	115 min	Action & Adventure	Young CIA operative Matt Weston must get a dan...
48	s49	Movie	Training Day	Antoine Fuqua	Denzel Washington, Ethan Hawke, Scott Glenn, T...	United States	September 16, 2021	2001	R	122 min	Dramas, Thrillers	A rookie cop with one day to prove himself to ...
81	s82	Movie	Kate	Cedric Nicolas-Troyan	Mary Elizabeth Winstead, Jun Kunimura, Woody H...	United States	September 10, 2021	2021	R	106 min	Action & Adventure	Slipped a fatal poison on her final job, a rut...
122	s123	Movie	In the Cut	Jane Campion	Meg Ryan, Mark Ruffalo, Jennifer Jason Leigh, ...	United Kingdom, Australia, France, United States	September 2, 2021	2003	R	118 min	Thrillers	After embarking on an affair with the cop prob...
131	s132	Movie	Blade Runner: The Final Cut	Ridley Scott	Harrison Ford, Rutger Hauer, Sean Young, Edwar...	United States	September 1, 2021	1982	R	117 min	Action & Adventure, Classic Movies, Cult Movies	In a smog-choked dystopian Los Angeles, blade ...
...	...	...	...	...	...	...	...	...	...	...	...	...
8765	s8766	Movie	XX	Jovanka Vuckovic, Annie Clark, Roxanne Benjami...	Natalie Brown, Jonathan Watton, Melanie Lynske...	Canada, United States	June 22, 2017	2017	R	81 min	Horror Movies	This four-part anthology of short horror films...
8768	s8769	Movie	Y Tu Mamá También	Alfonso Cuarón	Maribel Verdú, Gael García Bernal, Diego Luna,...	Mexico	June 1, 2017	2001	R	106 min	Dramas, Independent Movies, International Movies	When rich teens Tenoch and Julio meet the allu...
8791	s8792	Movie	Young Adult	Jason Reitman	Charlize Theron, Patton Oswalt, Patrick Wilson...	United States	November 20, 2019	2011	R	94 min	Comedies, Dramas, Independent Movies	When a divorced writer gets a letter from an o...
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	R	158 min	Cult Movies, Dramas, Thrillers	A political cartoonist, a crime reporter and a...
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009	R	88 min	Comedies, Horror Movies	Looking to survive in a world taken over by zo...

797 rows × 12 columns

12. Display all rows for which rating does not start with 'TV'

```
# 12. Display all rows for which rating does not start with 'TV'
df[~df["rating"].str.startswith("TV", na=False)]
```

0.3s — Python

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, filmm...
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	NaN	September 24, 2021	2021	PG	91 min	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...
9	s10	Movie	The Starling	Theodore Melfi	Melissa McCarthy, Chris O'Dowd, Kevin Kline, T...	United States	September 24, 2021	2021	PG-13	104 min	Comedies, Dramas	A woman adjusting to life after a loss contend...
27	s28	Movie	Grown Ups	Dennis Dugan	Adam Sandler, Kevin James, Chris Rock, David S...	United States	September 20, 2021	2010	PG-13	103 min	Comedies	Mourning the loss of their beloved junior high...
28	s29	Movie	Dark Skies	Scott Stewart	Keri Russell, Josh Hamilton, J.K. Simmons, Dak...	United States	September 19, 2021	2013	PG-13	97 min	Horror Movies, Sci-Fi & Fantasy	A family's idyllic suburban life shatters when...
...	...	...	...	...	...	...	...	...	...	...	...	...
8792	s8793	Movie	Young Tiger	Mu Chu	Qiu Yuen, Charlie Chin, Jackie Chan, Hu Chin, ...	Hong Kong	November 1, 2016	1973	NR	81 min	Action & Adventure, International Movies	Aided only by a tough female police officer, a...
8793	s8794	Movie	Yours, Mine and Ours	Raja Gosnell	Dennis Quaid, Rene Russo, Sean Faris, Katija P...	United States	November 20, 2019	2005	PG	88 min	Children & Family Movies, Comedies	When a father of eight and a mother of 10 prep...
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	R	158 min	Cult Movies, Dramas, Thrillers	A political cartoonist, a crime reporter and a...
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009	R	88 min	Comedies, Horror Movies	Looking to survive in a world taken over by zo...
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	PG	88 min	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...

1710 rows × 12 columns

13. Display all rows for which the rating is missing (this will be all the ratings that have *nan* values)

```
# 13. Display all rows for which the rating is missing (this will be all the ratings that have nan values)
df[df["rating"].isna()]
```

0.4s — Python

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
5989	s5990	Movie	13TH: A Conversation with Oprah Winfrey & Ava ...	NaN	Oprah Winfrey, Ava DuVernay	NaN	January 26, 2017	2017	NaN	37 min	Movies	Oprah Winfrey sits down with director Ava DuVe...
6827	s6828	TV Show	Gargantia on the Verdurous Planet	NaN	Kaito Ishikawa, Hisako Kanemoto, Ai Kayano, Ka...	Japan	December 1, 2016	2013	NaN	1 Season	Anime Series, International TV Shows	After falling through a wormhole, a space-dwel...
7312	s7313	TV Show	Little Lunch	NaN	Flynn Curry, Olivia Deeble, Madison Lu, Oisín ...	Australia	February 1, 2018	2015	NaN	1 Season	Kids' TV, TV Comedies	Adopting a child's perspective, this show take...
7537	s7538	Movie	My Honor Was Loyalty	Alessandro Pepe	Leone Frisa, Paolo Vaccarino, Francesco Miglio...	Italy	March 1, 2017	2015	NaN	115 min	Dramas	Amid the chaos and horror of World War II, a c...

14. How many rows have missing ratings?

```
# 14. How many rows have missing ratings?
df[df["rating"].isna()].shape[0]
```

0.4s

15. Add a new column called 'isMature' and initialize all the values to False

```
# 15. Add a new column called 'isMature' and initialize all the values to False
df["isMature"] = False
df
```

0.3s Python

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	isMature
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, filmm...	False
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	False
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lor...	False
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	1 Season	Docuseries, Reality TV	Feuds, flirtations and toilet talk go down amo...	False
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, Romantic TV Shows, TV ...	In a city of coaching centers known to train I...	False
...	...	...	...	...	...	...	...	...	...	...	...	...	...
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	R	158 min	Cult Movies, Dramas, Thrillers	A political cartoonist, a crime reporter and a...	False
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	July 1, 2019	2018	TV-Y7	2 Seasons	Kids' TV, Korean TV Shows, TV Comedies	While living alone in a spooky town, a young g...	False
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009	R	88 min	Comedies, Horror Movies	Looking to survive in a world taken over by zo...	False
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	PG	88 min	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...	False
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	2015	TV-14	111 min	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	False

8807 rows × 13 columns

16. For each row, set the column value for 'isMature' to True if the rating is any of the following: 'PG-13', 'TV-MA', 'TV-14', 'R','NC-17'

```
# 16. For each row, set the column value for 'isMature' to True if the rating is any of the following: 'PG-13', 'TV-MA', 'TV-14', 'R','NC-17'
df.loc[df["rating"].isin(['PG-13', 'TV-MA', 'TV-14', 'R','NC-17']), "isMature"] = True
df
```

0.3s Python

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	isMature
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, filmm...	True
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	True
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lor...	True
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	1 Season	Docuseries, Reality TV	Feuds, flirtations and toilet talk go down amo...	True
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, Romantic TV Shows, TV ...	In a city of coaching centers known to train I...	True
...	...	...	...	...	...	...	...	...	...	...	...	...	...
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	R	158 min	Cult Movies, Dramas, Thrillers	A political cartoonist, a crime reporter and a...	True
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	July 1, 2019	2018	TV-Y7	2 Seasons	Kids' TV, Korean TV Shows, TV Comedies	While living alone in a spooky town, a young g...	False
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009	R	88 min	Comedies, Horror Movies	Looking to survive in a world taken over by zo...	True
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	PG	88 min	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...	False
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	2015	TV-14	111 min	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	True

8807 rows × 13 columns

MCAR, MNAR, and MAR

(9 points) Notice that one of the values for 'rating' is 'nan'. This means the value is missing. Match each scenario for why the rating is missing with the correct missing data type (each type can only be used once):

1. Movies that are shorter than an hour and are not released in the US or Europe will not have ratings.
   MNAR - Missing not at random

2. We have no idea why the ratings are missing
   MCAR - Missing completely at random

3. When constructing the database, entering the ratings were optional
   MAR - Missing at random

Handling Categorical Data

For each question, answer with the code that would perform the given task. Add a screenshot of the output.

Using pandas, import the dataset “netflix_titles” as a dataframe and

1. (5pts) Create a reasonable mapping for the ‘ratings’ values replace the column values with the mappings

```
# 1.    (5pts) Create a reasonable mapping for the ‘ratings’ values replace the column values with the mappings
ratings_number_map = {
    "TV-Y": 0,
    "TV-Y7": 1,
    "TV-G": 2,
    "TV-PG": 3,
    "G": 4,
    "PG": 5,
    "PG-13": 6,
    "TV-MA": 7,
    "TV-14": 8,
    "R": 9,
    "NC-17": 10,
    "NR": 11,
    "UR": 12,
    "TV-Y7-FV": 13,
    "74 min": 14,
    "84 min": 15,
    "66 min": 16,
}

new_df = df.copy()
new_df["rating"] = new_df["rating"].map(ratings_number_map)
new_df
```

✓ 0.3s

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	6.0	90 min	Documentaries	As her father nears the end of his life, filmm...
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	7.0	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	7.0	1 Season	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lor...
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	7.0	1 Season	Docuseries, Reality TV	Feuds, flirtations and toilet talk go down amo...
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	7.0	2 Seasons	International TV Shows, Romantic TV Shows, TV ...	In a city of coaching centers known to train I...
...	...	...	...	...	...	...	...	...	...	...	...	...
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	9.0	158 min	Cult Movies, Dramas, Thrillers	A political cartoonist, a crime reporter and a...
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	July 1, 2019	2018	1.0	2 Seasons	Kids' TV, Korean TV Shows, TV Comedies	While living alone in a spooky town, a young g...
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009	9.0	88 min	Comedies, Horror Movies	Looking to survive in a world taken over by zo...
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	5.0	88 min	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	2015	8.0	111 min	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...

8807 rows × 12 columns

2. (5pt) Explain your mapping

   I just took a list of all the unique ratings that are present and created a object with serial numbers.

```
df["rating"].unique()
```
✓ 0.2s

```
array(['PG-13', 'TV-MA', 'PG', 'TV-14', 'TV-PG', 'TV-Y', 'TV-Y7', 'R',
       'TV-G', 'G', 'NC-17', '74 min', '84 min', '66 min', 'NR', nan,
       'TV-Y7-FV', 'UR'], dtype=object)
```

```
ratings_number_map = {
    "TV-Y": 0,
    "TV-Y7": 1,
    "TV-G": 2,
    "TV-PG": 3,
    "G": 4,
    "PG": 5,
    "PG-13": 6,
    "TV-MA": 7,
    "TV-14": 8,
    "R": 9,
    "NC-17": 10,
    "NR": 11,
    "UR": 12,
    "TV-Y7-FV": 13,
    "74 min": 14,
    "84 min": 15,
    "66 min": 16,
}
```

3. After replacing the ‘ratings’ values with your mapping
    a. (3pts) Drop all the ‘ratings’ with missing values (*dropna*) and find the average of the ‘rating’ column.

```
# a.  (3pts) Drop all the ‘ratings’ with missing values (dropna) and find the average of the ‘rating’ column.
new_df.dropna(subset=["rating"])["rating"].mean()
# new_df[new_df["rating"].isna()].shape[0]
```
✓ 0.7s

```
6.3490855390207885
```

    b. (3pts) Replace all the ‘ratings’ with missing values with 0 (*fillna*) and find the new average of the ‘rating’ column.

```
# b.  (3pts) Replace all the 'ratings' with missing values with 0 (fillna) and find the new average of the 'rating' column.
new_df["rating"].fillna(0).mean()
```
✓ 0.2s
```
6.346201884864312
```

c. (4pts) Replace all the 'ratings' with missing values with the average of the 'rating' column (*fillna*) and find the new average of the 'rating' column

```
# c.  (4pts) Replace all the 'ratings' with missing values with the average of the 'rating'
# column (fillna) and find the new average of the 'rating' column
new_df["rating"].fillna(new_df["rating"].mean()).mean()
```
✓ 0.4s
```
6.3490855390207885
```

4. (3pt) How does the average change based on how you handle it?
   Answer:
   Dropping all the missing value and filling na values with the average gave same output as calculating average using pandas doesn't consider NA values while doing so.

   Filling all na with 0 did reduce the average a bit.