

DATA ANALYST INTERNSHIP

Task 1: Data Cleaning and Preprocessing

Selected Dataset: Netflix Movies and TV Shows

Raw Data:

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
s1	Movie	Dick	Johr	Kirsten Johnson	United States	25-Sep-21	2020	PG-13	90 min	Documentary	As her father nears the end of his life, filmmaker Kirsten Johnson stars
s2	TV Show	Blood & Water	Ama	Qamata, Khosi Ngema, Gail Mabalane, Thabang Molaba, Dillo	South Africa	24-Sep-21	2021	TV-MA	2 Seasons	International	After crossing paths at a party, a Cape Town teen sets out to prove w
s3	TV Show	Gangland	Julien Le Sami	Bouajila, Tracy Gotoas, Samuel Jouy, Nabih Akkari, Sofia Lesaffre, Salin	France	24-Sep-21	2021	TV-MA	1 Season	Crime TV	To protect his family from a powerful drug lord, skilled thief Mehdi a
s4	TV Show	Jailbirds New Orleans	Mayur	More, Jitendra Kumar, Ranjan Raj, Alam Khan, Ahsaas Channa	India	24-Sep-21	2021	TV-MA	1 Season	Docuseries	Feuds, flirtations and toilet talk go down among the incarcerated wo
s5	TV Show	Kota Factory				24-Sep-21	2021	TV-MA	2 Seasons	International	In a city of coaching centers known to train Indian students, the finest collegia
s6	TV Show	Midnight Mike	Fla	Kate Siegel, Zach Gilford, Hamish Linklater, Henry Thomas, Kristin Lehman, Sa	United States	24-Sep-21	2021	TV-MA	1 Season	TV Drama	The arrival of a charismatic young priest brings glorious miracles, on
s7	Movie	My Little	Robert Cl	Vanessa Hudgens, Kimiko Glenn, James Marsden, Sofia Carson, Liza Koshy, Ke	United States	24-Sep-21	2021	PG	91 min	Children	Equestria's divided. But a bright-eyed hero believes Earth Ponies, Pe
s8	Movie	Sankofa	Haile Ge	Korfi Ghanaba, Oyafunmike Ogunlano, Alexandra Duah, Nick Medley	United States	24-Sep-21	1993	TV-MA	125 min	Dramas	On a photo shoot in Ghana, an American model slips back in time, br
s9	TV Show	The Great	Andy Dev	Mel Giedroyc, Sue Perkins, Mary Berry, Paul Hollywood	United Kingdom	24-Sep-21	2021	TV-14	9 Seasons	British TV	A talented batch of amateur bakers face off in a 10-week competitio
s10	Movie	The Star	Theodor	Melissa McCarthy, Chris O'Dowd, Kevin Kline, Timothy Olyphant, Davi	United States	24-Sep-21	2021	PG-13	104 min	Comedies	A woman adjusting to life after a loss contends with a feisty bird tha
s11	TV Show	Vendetta	Truth, Lies and The Mafia			24-Sep-21	2021	TV-MA	1 Season	Crime TV	Sicily boasts a bold "Anti-Mafia" coalition. But what happens when t
s12	TV Show	Bangkok	Kongkiat	Sukollawat Kanarot, Sushar Manaying, Pavarit Mongkolpisit, Sahajak Boontha	Thailand	23-Sep-21	2021	TV-MA	1 Season	Crime TV	Struggling to earn a living in Bangkok, a man joins an emergency rescu
s13	Movie	Je Suis Ki	Christian Luna	Wedler, Jannis Niewöhner, Milan Peschel, Edin Hasanović, A Germany	Germany	23-Sep-21	2021	TV-MA	127 min	Dramas	After most of her family is murdered in a terrorist bombing, a young v
s14	Movie	Confessio	Bruno G	Klara Castanho, Lucca Picon, Jârlia Gomes, Marcus Bessa, Kiria Malheiros, Fer	Brazil	22-Sep-21	2021	TV-PG	91 min	Children	When the clever but socially-awkward Tetê joins a new school, she'
s15	TV Show	Crime Stories	India	Detectives		22-Sep-21	2021	TV-MA	1 Season	British TV	Cameras following Bengaliu police on the job offer a rare glimpse i
s16	TV Show	Dear White People	Logan Browning, Brandon P. Bell, DeRon Horton, Antoinette Roberts	United States	22-Sep-21	2021	TV-MA	4 Seasons	TV Comedy	Students of color navigate the daily slights and slippery politics of H	
s17	Movie	Europe's	Pedro de	Enache Garc�a, Pablo Ator�n Williams	Spain	22-Sep-21	2020	TV-MA	67 min	Documentary	Declassified documents reveal the post-WWII life of Otto Skorzeny, a
s18	TV Show	Falsa	Identidad	Luis Ernesto Franco, Camila Sodi, Sereio Govi, Samadhi Zendeias	Brazil	22-Sep-21	2020	TV-MA	2 Seasons	Crime TV	Strangers Diego and Isabel flee their home in Mexico and pretend to

Variables:

1. show_id: Identify the id name of particular movie/tv show
2. type: Is it Movie or TV show
3. title: Title of movie/tv show
4. director: Name of Director
5. cast: who are part of acting
6. country: Release in which countries
7. date_added: date of streaming
8. release_year: release year of that movie/tv show
9. rating: Likewise of that show
10. duration: duration of show in mins/seasons
11. listed_in: category of the TV show/Movie
12. description: short summary

Data Cleaning:

Tool: Excel and Power Query and Pandas

Brief Description:

- Check Null Values using filters

Findings

director: 2,634

cast: 825

country: 831 → Filled with title-cased placeholder 'Nan' (as string)

date_added: 10

rating: 4

duration: 3

Replace by Director, Cast, Country by “Unknown”

And date_added by 01/01/2020 as per release year

- Checking Duplicate Value

No Duplicate Values

- Convert Date-time in dd/mm/yyyy format using change type of power query
- Standardize the string element by proper type casting
- Uniform the Headers

Bold and standardize with first Letter Uppercase and rest of Lower case and space by “ _ ”

Cleaned dataset View:

Show_id	Type	Title	Director	Cast	Country	Date_added	Release	Rating	Duration	Listed_in	Description
s1	Movie	Dick Johnson's Dead	Kristen Johnson	Documentary	United States	9/25/2021	2020 PG-13		30 min	Documentaries	As her f
s2	TV Show	Blood & Water	NONE	Amá Damata, Kho	South Africa	9/24/2021	2021 TV-MA	2 Seasons		International TV Shows, TV Dramas, TV Mys	Alter or
s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tra	Unknown	9/24/2021	2021 TV-MA	1 Season		Crime TV Shows, International TV Shows, T	To pose
s4	TV Show	Jaibards New Orleans	NONE	Documentary	Unknown	9/24/2021	2021 TV-MA	1 Season		Docuseries, Reality TV	Feuds, f
s5	TV Show	Kota Factory	NONE	Mayur More, Jitenc	India	9/24/2021	2021 TV-MA	2 Seasons		International TV Shows, Romantic TV Show	In a city,
s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach	Unknown	9/24/2021	2021 TV-MA	1 Season		TV Dramas, TV Horror, TV Mysteries	The am,
s7	Movie	My Little Pony: A New Generation	Robert Cullen, Jos	Robert Cullen, Jos	USA, Luis Ucha	9/24/2021	2021 PG		31 min	Children & Family Movies	Equestri
s8	Movie	Sanjida	Hale Geima	Kofi Ghanaba, Oyi	United States, Ghana, Burkina Fi	9/24/2021	1993 TV-MA		125 min	Dramas, Independent Movies, International	On a ph
s9	TV Show	The Great British Baking Show	Andy Devonshire	Mel Giedroyc, Sue	United Kingdom	9/24/2021	2021 TV-14	9 Seasons		British TV Shows, Reality TV	A talent
s10	Movie	The Starling	Theodore Mell	Melissa McCarthy	United States	9/24/2021	2021 PG-13		104 min	Comedies, Dramas	A verna
s11	TV Show	Vendetta: Truth, Lies and The Mafia	NONE	Documentary	Unknown	9/24/2021	2021 TV-MA	1 Season		Crime TV Shows, Docuseries, International	Sickly bo
s12	TV Show	Bangkok Breaking	Kongkiat Komestit	Sukollawat Kanar	Unknown	9/23/2021	2021 TV-MA	1 Season		Crime TV Shows, International TV Shows, T	Struggl
s13	Movie	Je Suis Karl	Christian Schwochow	Luna Medler, Jan	Germany, Czech Republi	9/23/2021	2021 TV-MA		127 min	Dramas, International Movies	Alter mo
s14	Movie	Confessions of an Invisible Girl	Bruno Garcia	Klara Cernáňová, Li	Unknown	9/22/2021	2021 TV-PG		31 min	Children & Family Movies, Comedies	When th
s15	TV Show	Crime Stories: India Detectives	NONE	Documentary	Unknown	9/22/2021	2021 TV-MA	1 Season		British TV Shows, Crime TV Shows, Docuse	Camera
s16	TV Show	Dear White People	NONE	Logan Browning, E	United States	9/22/2021	2021 TV-MA	4 Seasons		TV Comedies, TV Dramas	Student
s17	Movie	Europe's Most Dangerous Man: Otto	Pedro de Echave Gar	Álvaro, Pablo Azor	Unknown	9/22/2021	2020 TV-MA		67 min	Documentaries, International Movies	Declass
s18	TV Show	Falsa identidad	NONE	Luis Ernesto Franc	Mexico	9/22/2021	2020 TV-MA	2 Seasons		Crime TV Shows, Spanish-Language TV Sh	Strange
s19	Movie	Intrusion	Adam Salky	Freida Pinto, Loga	Unknown	9/22/2021	2021 TV-14		94 min	Thrillers	Alter a d
s20	TV Show	Jaguar	NONE	Blanca Suárez, Iv	Unknown	9/22/2021	2021 TV-MA	1 Season		International TV Shows, Spanish-Languag	In the 19
s21	TV Show	Monsters Inside: The 24 Faces of B	Oliver Hegaton	Documentary	Unknown	9/22/2021	2021 TV-14	1 Season		Crime TV Shows, Docuseries, International	In the lat
s22	TV Show	Resurrection: Enayut	NONE	Ergin Altun Dikici	Turkey	9/22/2021	2018 TV-14	5 Seasons		International TV Shows, TV Action & Adven	When a
s23	Movie	Awaj Shamnugh	K. S. Ravikumar	Kamal Hassan, Me	Unknown	9/21/2021	1996 TV-PG		161 min	Comedies, International Movies	Newly d
s24	Movie	Gol Gol Cony Carson: Christy Takes	Alex Woo, Stanley Moore	Maïse Benson, Pa	Unknown	9/21/2021	2021 TV-Y		61 min	Children & Family Movies	From arc
s25	Movie	Jeans	S. Shankar	Prashanth, Aishw	India	9/21/2021	1998 TV-14		166 min	Comedies, International Movies, Romantic	When th
s26	TV Show	Love on the Spectrum	NONE	Brooke Satchwell	Australia	9/21/2021	2021 TV-14	2 Seasons		Docuseries, International TV Shows, Realit	Findin

Thank You

