

In [392]:

```
#import librairies:
import networkx as nx
import matplotlib.pyplot as plt
import pandas as pd
import csv
```

In [393]:

```
csv_file = pd.read_csv('netflix_titles.csv')
```

In [394]:

```
csv_file.shape
```

Out[394]:

```
(6234, 12)
```

In [395]:

```
#1.1. Check duplicates values in the dataset
```

```
csv_file.loc[csv_file.duplicated(), :]
```

Out[395]:

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in
										

In [396]:

```
#1.2. Find the missing rows in each column.
```

```
csv_file.isnull().sum()
```

Out[396]:

```
show_id      0
type         0
title        0
director    1969
cast        570
country     476
date_added   11
release_year  0
rating       10
duration     0
listed_in    0
description  0
dtype: int64
```

In [397]:

```
director = csv_file[csv_file['director'].isnull()]
director
```

Out[397]:

	show_id	type	title	director	cast	country	date_added	release_year
1	80117401	Movie	Jandino: Whatever it Takes	NaN	Jandino Asporaat	United Kingdom	September 9, 2016	2016
2	70234439	TV Show	Transformers Prime	NaN	Peter Cullen, Sumalee Montano, Frank Welker, J...	United States	September 8, 2018	2013
3	80058654	TV Show	Transformers: Robots in Disguise	NaN	Will Friedle, Darren Criss, Constance Zimmer, ...	United States	September 8, 2018	2016
5	80163890	TV Show	Apaches	NaN	Alberto Ammann, Eloy Azorín, Verónica Echegui,...	Spain	September 8, 2017	2016
8	80117902	TV Show	Fire Chasers	NaN	NaN	United States	September 8, 2017	2017
...
6229	80000063	TV Show	Red vs. Blue	NaN	Burnie Burns, Jason Saldaña, Gustavo Sorola, G...	United States	NaN	2015
6230	70286564	TV Show	Maron	NaN	Marc Maron, Judd Hirsch, Josh Brener, Nora Zeh...	United States	NaN	2016
6231	80116008	Movie	Little Baby Bum: Nursery Rhyme Friends	NaN	NaN	NaN	NaN	2016
6232	70281022	TV Show	A Young Doctor's Notebook and Other Stories	NaN	Daniel Radcliffe, Jon Hamm, Adam Godley, Chris...	United Kingdom	NaN	2013

	show_id	type	title	director	cast	country	date_added	release_year
6233	70153404	TV Show	Friends	NaN	Jennifer Aniston, Courteney Cox, Lisa Kudrow, ...	United States	NaN	2003

1969 rows × 12 columns

In [398]:

```
#missing rows in cast
cast = csv_file[csv_file['cast'].isnull()]

cast
```

Out[398]:

	show_id	type	title	director	cast	country	date_added	release_year	rating
8	80117902	TV Show	Fire Chasers	NaN	NaN	United States	September 8, 2017	2017	TV-MA
22	80046727	Movie	Rolling Papers	Mitch Dickman	NaN	United States, Uruguay	September 8, 2017	2015	TV-MA
27	80203094	Movie	City of Joy	Madeleine Gavin	NaN	United States,	September 7, 2018	2018	TV-MA
28	80190843	TV Show	First and Last	NaN	NaN	NaN	September 7, 2018	2018	TV-MA
66	81011682	TV Show	Christiane Amanpour: Sex & Love Around the World	NaN	NaN	United States	September 30, 2018	2018	TV-MA
...
6205	80036747	TV Show	Pororo - The Little Penguin	NaN	NaN	South Korea	April 19, 2019	2013	TV-Y
6207	80190407	TV Show	Murderous Affairs	NaN	NaN	United States	April 17, 2018	2017	TV-14
6211	80231523	TV Show	The Minimighty Kids	NaN	NaN	France	April 15, 2018	2012	TV-G
6214	80049872	TV Show	Chelsea	NaN	NaN	United States	April 14, 2017	2017	TV-MA
6231	80116008	Movie	Little Baby Bum: Nursery Rhyme Friends	NaN	NaN	NaN	NaN	2016	NaN

570 rows × 12 columns



In [399]:

```
#missing rows in country
country = csv_file[csv_file['country'].isnull()]

country
```

Out[399]:

	show_id	type	title	director	cast	country	date_added	release_year	r
10	80169755	Movie	Joaquín Reyes: Una y no más	José Miguel Contreras	Joaquín Reyes	NaN	September 8, 2017	2017	
12	80182480	Movie	Krish Trish and Baltiboy	NaN	Damandeep Singh Baggan, Smita Malhotra, Baba S...	NaN	September 8, 2017	2009	1
13	80182483	Movie	Krish Trish and Baltiboy: Battle of Wits	Munjal Shroff, Tilak Shetty	Damandeep Singh Baggan, Smita Malhotra, Baba S...	NaN	September 8, 2017	2013	1
14	80182596	Movie	Krish Trish and Baltiboy: Best Friends Forever	Munjal Shroff, Tilak Shetty	Damandeep Singh Baggan, Smita Malhotra, Deepak...	NaN	September 8, 2017	2016	
15	80182482	Movie	Krish Trish and Baltiboy: Comics of India	Tilak Shetty	Damandeep Singh Baggan, Smita Malhotra, Baba S...	NaN	September 8, 2017	2012	1
...	
6170	80136456	TV Show	Age of Glory	NaN	Debbie Goh, Aenie Wong, Frederick Lee, Leslie ...	NaN	August 14, 2017	2010	
6195	80136790	TV Show	Goodnight DJ 1	NaN	Chan Fong, Z-Chen, Frederick Lee, Melvin Sia, ...	NaN	April 27, 2018	2016	-
6199	80244228	TV Show	Beauty and the Bitches	NaN	Methinee Kingpayome, Son Songpaisarn, Yingying...	NaN	April 23, 2019	2016	-
6228	80159925	TV Show	Kikoriki	NaN	Igor Dmitriev	NaN	NaN	2010	

	show_id	type	title	director	cast	country	date_added	release_year	r
6231	80116008	Movie	Little Baby Bum: Nursery Rhyme Friends	NaN	NaN	NaN	NaN	2016	

476 rows × 12 columns



In [400]:

```
date_added = csv_file[csv_file['date_added'].isnull()]
```

date_added

Out[400]:

	show_id	type	title	director	cast	country	date_added	release_year
6223	70204989	TV Show	Gunslinger Girl	NaN	Yuuka Nanri, Kanako Mitsuhashi, Eri Sendai, Am...	Japan	NaN	2008
6224	70304979	TV Show	Anthony Bourdain: Parts Unknown	NaN	Anthony Bourdain	United States	NaN	2018
6225	70153412	TV Show	Frasier	NaN	Kelsey Grammer, Jane Leeves, David Hyde Pierce...	United States	NaN	2003
6226	70243132	TV Show	La Familia P. Luche	NaN	Eugenio Derbez, Consuelo Duval, Luis Manuel Áv...	United States	NaN	2012
6227	80005756	TV Show	The Adventures of Figaro Pho	NaN	Luke Jurevicius, Craig Behenna, Charlotte Haml...	Australia	NaN	2015
6228	80159925	TV Show	Kikoriki	NaN	Igor Dmitriev	NaN	NaN	2010
6229	80000063	TV Show	Red vs. Blue	NaN	Burnie Burns, Jason Saldaña, Gustavo Sorola, G...	United States	NaN	2015
6230	70286564	TV Show	Maron	NaN	Marc Maron, Judd Hirsch, Josh Brener, Nora Zeh...	United States	NaN	2016
6231	80116008	Movie	Little Baby Bum: Nursery Rhyme Friends	NaN	NaN	NaN	NaN	2016

	show_id	type	title	director	cast	country	date_added	release_year
6232	70281022	TV Show	A Young Doctor's Notebook and Other Stories	NaN	Daniel Radcliffe, Jon Hamm, Adam Godley, Chris...	United Kingdom	NaN	2013
6233	70153404	TV Show	Friends	NaN	Jennifer Aniston, Courteney Cox, Lisa Kudrow, ...	United States	NaN	2003

In [401]:

```
rating = csv_file[csv_file['rating'].isnull()]
```

rating

Out[401]:

	show_id	type	title	director	cast	country	date_added	release_year
211	70129452	Movie	Louis C.K.: Hilarious	Louis C.K.	Louis C.K.	United States	September 16, 2016	2010
2411	80144119	Movie	My Honor Was Loyalty	Alessandro Pepe	Leone Frisa, Paolo Vaccarino, Francesco Miglio...	Italy	March 1, 2017	2015
3288	80169801	Movie	13TH: A Conversation with Oprah Winfrey & Ava ...	NaN	Oprah Winfrey, Ava DuVernay	NaN	January 26, 2017	2017
4056	80078037	TV Show	Little Lunch	NaN	Flynn Curry, Olivia Deeble, Madison Lu, Oisín ...	Australia	February 1, 2018	2015
4402	80092839	Movie	Fireplace 4K: Classic Crackling Fireplace from...	George Ford	NaN	NaN	December 21, 2015	2015
4403	80092835	Movie	Fireplace 4K: Crackling Birchwood from Firepla...	George Ford	NaN	NaN	December 21, 2015	2015
4706	80039789	TV Show	Gargantia on the Verdurous Planet	NaN	Kaito Ishikawa, Hisako Kanemoto, Ai Kayano, Ka...	Japan	December 1, 2016	2013
5015	80114111	Movie	Louis C.K.: Live at the Comedy Store	Louis C.K.	Louis C.K.	United States	August 15, 2016	2015
5234	80161109	Movie	Louis C.K. 2017	Louis C.K.	Louis C.K.	United States	April 4, 2017	2017
6231	80116008	Movie	Little Baby Bum: Nursery Rhyme Friends	NaN	NaN	NaN	NaN	2016

In [402]:

#1.3. Remove the NaN values from the dataset.

no_nan = csv_file.dropna()

no_nan

Out[402]:

	show_id	type	title	director	cast	country	date_added	release
0	81145628	Movie	Norm of the North: King Sized Adventure	Richard Finn, Tim Maltby	Alan Marriott, Andrew Toth, Brian Dobson, Cole...	United States, India, South Korea, China	September 9, 2019	
4	80125979	Movie	#realityhigh	Fernando Lebrija	Nesta Cooper, Kate Walsh, John Michael Higgins...	United States	September 8, 2017	
6	70304989	Movie	Automata	Gabe Ibáñez	Antonio Banderas, Dylan McDermott, Melanie Gri...	Bulgaria, United States, Spain, Canada	September 8, 2017	
7	80164077	Movie	Fabrizio Copano: Solo pienso en mi	Rodrigo Toro, Francisco Schultz	Fabrizio Copano	Chile	September 8, 2017	
9	70304990	Movie	Good People	Henrik Ruben Genz	James Franco, Kate Hudson, Tom Wilkinson, Omar...	United States, United Kingdom, Denmark, Sweden	September 8, 2017	
...
6142	80063224	TV Show	The Great British Baking Show	Andy Devonshire	Mel Giedroyc, Sue Perkins, Mary Berry, Paul Ho...	United Kingdom	August 30, 2019	
6158	80164216	TV Show	Miraculous: Tales of Ladybug & Cat Noir	Thomas Astruc	Cristina Vee, Bryce Papenbrook, Keith Silverst...	France, South Korea, Japan	August 2, 2019	
6167	80115328	TV Show	Sacred Games	Vikramaditya Motwane, Anurag Kashyap	Saif Ali Khan, Nawazuddin Siddiqui, Radhika Ap...	India, United States	August 15, 2019	
6182	80176842	TV Show	Men on a Mission	Jung-ah Im	Ho-dong Kang, Soo-geun Lee, Sang-min Lee, Youn...	South Korea	April 9, 2019	

	show_id	type	title	director	cast	country	date_added	release_
6213	80126599	TV Show	Leyla and Mecnun	Onur Ünlü	Ali Atay, Melis Birkan, Serkan Keskin, Ahmet M...	Turkey	April 15, 2017	

3774 rows × 12 columns

In [403]:

```
#1.4.change the `date added` values to the only year. And rename the column: `added year`.
csv_cc = no_nan.rename(columns = {"date_added": "added_year"})

new_col = csv_cc['added_year'].str.split(',')

# code to split it into 2 lists
year_list = [i[1] for i in new_col]

csv_cc['added_year'] = year_list

csv_cc
```

Out[403]:

	show_id	type	title	director	cast	country	added_year	release
0	81145628	Movie	Norm of the North: King Sized Adventure	Richard Finn, Tim Maltby	Alan Marriott, Andrew Toth, Brian Dobson, Cole...	United States, India, South Korea, China	2019	
4	80125979	Movie	#realityhigh	Fernando Lebrija	Nesta Cooper, Kate Walsh, John Michael Higgins...	United States	2017	
6	70304989	Movie	Automata	Gabe Ibáñez	Antonio Banderas, Dylan McDermott, Melanie Gri...	Bulgaria, United States, Spain, Canada	2017	
7	80164077	Movie	Fabrizio Copano: Solo pienso en mi	Rodrigo Toro, Francisco Schultz	Fabrizio Copano	Chile	2017	
9	70304990	Movie	Good People	Henrik Ruben Genz	James Franco, Kate Hudson, Tom Wilkinson, Omar...	United States, United Kingdom, Denmark, Sweden	2017	
...
6142	80063224	TV Show	The Great British Baking Show	Andy Devonshire	Mel Giedroyc, Sue Perkins, Mary Berry, Paul Ho...	United Kingdom	2019	
6158	80164216	TV Show	Miraculous: Tales of Ladybug & Cat Noir	Thomas Astruc	Cristina Vee, Bryce Papenbrook, Keith Silverst...	France, South Korea, Japan	2019	

	show_id	type	title	director	cast	country	added_year	release_
6167	80115328	TV Show	Sacred Games	Vikramaditya Motwane, Anurag Kashyap	Saif Ali Khan, Nawazuddin Siddiqui, Radhika Ap...	India, United States	2019	
6182	80176842	TV Show	Men on a Mission	Jung-ah Im	Ho-dong Kang, Soo-geun Lee, Sang-min Lee, Youn...	South Korea	2019	
6213	80126599	TV Show	Leyla and Mecnun	Onur Ünlü	Ali Atay, Melis Birkan, Serkan Keskin, Ahmet M...	Turkey	2017	

3774 rows × 12 columns

In [404]:

```
#1.5. Make a copy of the dataset
new_df = csv_cc.copy()
```

In [410]:

```
#1.6. In the copied dataset, remove all the columns, except for the `show-id`, `type` and `c
new_df1 = new_df.drop(["title", "title", "director", "country", "added_year", "release_year"]

new_df1
```

Out[410]:

	show_id	type	cast
0	81145628	Movie	Alan Marriott, Andrew Toth, Brian Dobson, Cole...
4	80125979	Movie	Nesta Cooper, Kate Walsh, John Michael Higgins...
6	70304989	Movie	Antonio Banderas, Dylan McDermott, Melanie Gri...
7	80164077	Movie	Fabrizio Copano
9	70304990	Movie	James Franco, Kate Hudson, Tom Wilkinson, Omar...
...
6142	80063224	TV Show	Mel Giedroyc, Sue Perkins, Mary Berry, Paul Ho...
6158	80164216	TV Show	Cristina Vee, Bryce Papenbrook, Keith Silverst...
6167	80115328	TV Show	Saif Ali Khan, Nawazuddin Siddiqui, Radhika Ap...
6182	80176842	TV Show	Ho-dong Kang, Soo-geun Lee, Sang-min Lee, Youn...
6213	80126599	TV Show	Ali Atay, Melis Birkan, Serkan Keskin, Ahmet M...

3774 rows × 3 columns

In [411]:

```
#In_the_copied df, you make sure that each artist will be in a row with the showid that they
```

```
sp_actor = new_df1['cast'].str.split(',')  
sp_actor
```

Out[411]:

```
0      [Alan Marriott,  Andrew Toth,  Brian Dobson,  ...  
4      [Nesta Cooper,  Kate Walsh,  John Michael Higg...  
6      [Antonio Banderas,  Dylan McDermott,  Melanie ...  
7                                     [Fabrizio Copano]  
9      [James Franco,  Kate Hudson,  Tom Wilkinson,  ...  
                                     ...  
6142   [Mel Giedroyc,  Sue Perkins,  Mary Berry,  Pau...  
6158   [Cristina Vee,  Bryce Papenbrook,  Keith Silve...  
6167   [Saif Ali Khan,  Nawazuddin Siddiqui,  Radhika...  
6182   [Ho-dong Kang,  Soo-geun Lee,  Sang-min Lee,  ...  
6213   [Ali Atay,  Melis Birkan,  Serkan Keskin,  Ahm...  
Name: cast, Length: 3774, dtype: object
```

In []: