

```
In [294]: import pandas as pd
import numpy as np
import requests
import bs4 as bs
import urllib.request
```

Movie 2020

```
In [295]: link = "https://en.wikipedia.org/wiki/List_of_American_films_of_2020"
```

```
In [296]: source = urllib.request.urlopen(link).read()
soup = bs.BeautifulSoup(source, 'lxml')
```

```
In [297]: tables = soup.find_all('table', class_='wikitable sortable')
```

```
In [298]: df1 = pd.read_html(str(tables[0]))[0]
df2 = pd.read_html(str(tables[1]))[0]
df3 = pd.read_html(str(tables[2]))[0]
df4 = pd.read_html(str(tables[3]).replace("'1\\\"'", "'1'"))[0]
```

```
In [299]: df = df1.append(df2.append(df3.append(df4, ignore_index=True), ignore_index=True), ignore_index=True)
del df['.mw-parser-output .tooltip-dotted{border-bottom:1px dotted;cursor:help}Re
```

In [300]: df

Out[300]:

	Opening	Opening.1	Title	Production company	Cast and crew	Ref.
0	JANUARY	3	The Grudge	Screen Gems / Stage 6 Films / Ghost House Pict...	Nicolas Pesce (director/screenplay); Andrea Ri...	NaN
1	JANUARY	10	Underwater	20th Century Fox / TSG Entertainment / Chernin...	William Eubank (director); Brian Duffield, Ada...	NaN
2	JANUARY	10	Like a Boss	Paramount Pictures	Miguel Arteta (director); Sam Pitman, Adam Col...	NaN
3	JANUARY	10	Three Christs	IFC Films	Jon Avnet (director/screenplay); Eric Nazarian...	NaN
4	JANUARY	10	Inherit the Viper	Barry Films / Tycor International Film Company	Anthony Jerjen (director); Andrew Crabtree (sc...	NaN
...
267	DECEMBER	25	We Can Be Heroes	Netflix / Troublemaker Studios	Robert Rodriguez (director/screenplay); Priyan...	[240]
268	DECEMBER	25	News of the World	Universal Pictures / Perfect World Pictures	Paul Greengrass (director/screenplay); Luke Da...	[241]
269	DECEMBER	25	One Night in Miami...	Amazon Studios	Regina King (director); Kemp Powers (screenpla...	[242]
270	DECEMBER	25	Promising Young Woman	Focus Features / FilmNation Entertainment	Emerald Fennell (director/screenplay); Carey M...	[243]
271	DECEMBER	30	Pieces of a Woman	Netflix / Bron Studios	Kornél Mundruczó (director); Kata Wéber (scree...	[244]

272 rows × 6 columns

In [301]: df_2020 = df[['Title', 'Cast and crew']]

```
In [302]: #df_2020.to_csv('check1.csv',index=False)
df_2020=pd.read_csv('check1.csv')
df_2020
```

Out[302]:

	Title	Cast and crew	genres
0	The Grudge	Nicolas Pesce (director/screenplay); Andrea Ri...	NaN
1	Underwater	William Eubank (director); Brian Duffield, Ada...	NaN
2	Like a Boss	Miguel Arteta (director); Sam Pitman, Adam Col...	NaN
3	Three Christs	Jon Avnet (director/screenplay); Eric Nazarian...	NaN
4	Inherit the Viper	Anthony Jerjen (director); Andrew Crabtree (sc...	NaN
...
267	We Can Be Heroes	Robert Rodriguez (director/screenplay); Priyan...	NaN
268	News of the World	Paul Greengrass (director/screenplay); Luke Da...	NaN
269	One Night in Miami...	Regina King (director); Kemp Powers (screenpla...	NaN
270	Promising Young Woman	Emerald Fennell (director/screenplay); Carey M...	NaN
271	Pieces of a Woman	Kornél Mundruczó (director); Kata Wéber (scree...	NaN

272 rows × 3 columns

```
In [303]: !pip install tmdbv3api
```

```
Requirement already satisfied: tmdbv3api in c:\users\vohan\anaconda3\lib\site-p
ackages (1.7.6)
Requirement already satisfied: requests in c:\users\vohan\anaconda3\lib\site-pa
ckages (from tmdbv3api) (2.25.1)
Requirement already satisfied: chardet<5,>=3.0.2 in c:\users\vohan\anaconda3\li
b\site-packages (from requests->tmdbv3api) (4.0.0)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\vohan\anaconda3\l
ib\site-packages (from requests->tmdbv3api) (2020.12.5)
Requirement already satisfied: idna<3,>=2.5 in c:\users\vohan\anaconda3\lib\sit
e-packages (from requests->tmdbv3api) (2.10)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in c:\users\vohan\anaconda
3\lib\site-packages (from requests->tmdbv3api) (1.26.4)
```

```
In [304]: from tmdbv3api import TMDb
import json
import requests
tmdb = TMDb()
tmdb.api_key = 'b7e7aeea608257af93a032205bfada32'
```

```
In [305]: i=0
result=[]
from tmdbv3api import Movie
tmdb_movie = Movie()
for x in df_2020['Title']:
    result.append(tmdb_movie.search(x))
    #print(result[i])
for i in range(0,272):
    print(result[i][0].id)
    i=i+1
    print(i)
```

```
759902
263
774372
264
584288
265
464052
266
508442
267
615677
268
581032
269
661914
270
582014
271
641662
272
```

```
In [306]: from tmdbv3api import Movie
tmdb_movie = Movie()
def get_genre(x):
    genres = []
    result = tmdb_movie.search(x)
    movie_id = result[0].id
    response = requests.get('https://api.themoviedb.org/3/movie/{}?api_key={}'.format(movie_id, api_key))
    data_json = response.json()
    if data_json['genres']:
        genre_str = " "
        for i in range(0, len(data_json['genres'])):
            genres.append(data_json['genres'][i]['name'])
        return genre_str.join(genres)
    else:
        return np.NaN
```

```
In [307]: df_2020['genres'] = df_2020['Title'].map(lambda x: get_genre(str(x)))
```

In [308]: df_2020

Out[308]:

	Title	Cast and crew	genres
0	The Grudge	Nicolas Pesce (director/screenplay); Andrea Ri...	Horror Mystery Thriller
1	Underwater	William Eubank (director); Brian Duffield, Ada...	Action Horror Science Fiction Thriller
2	Like a Boss	Miguel Arteta (director); Sam Pitman, Adam Col...	Comedy
3	Three Christs	Jon Avnet (director/screenplay); Eric Nazarian...	Drama
4	Inherit the Viper	Anthony Jerjen (director); Andrew Crabtree (sc...	Crime Thriller Drama
...
267	We Can Be Heroes	Robert Rodriguez (director/screenplay); Priyan...	Action Fantasy Family Comedy
268	News of the World	Paul Greengrass (director/screenplay); Luke Da...	Adventure Western
269	One Night in Miami...	Regina King (director); Kemp Powers (screenpla...	Drama
270	Promising Young Woman	Emerald Fennell (director/screenplay); Carey M...	Thriller Crime Drama
271	Pieces of a Woman	Kornél Mundruczó (director); Kata Wéber (scree...	Drama

272 rows × 3 columns

```
In [309]: def get_director(x):
           if " (director)" in x:
               return x.split(" (director)")[0]
           elif " (directors)" in x:
               return x.split(" (directors)")[0]
           else:
               return x.split(" (director/screenplay)")[0]
```

```
In [310]: df_2020['director_name'] = df_2020['Cast and crew'].map(lambda x: get_director(st
```

```
In [311]: def get_actor1(x):
           return ((x.split("screenplay); ")[-1]).split(", ")[0])
```

```
In [312]: df_2020['actor_1_name'] = df_2020['Cast and crew'].map(lambda x: get_actor1(str(>
```

```
In [313]: def get_actor2(x):  
          if len((x.split("screenplay); ")[-1]).split(", ")) < 2:  
              return np.NaN  
          else:  
              return ((x.split("screenplay); ")[-1]).split(", ")[1])
```

```
In [314]: df_2020['actor_2_name'] = df_2020['Cast and crew'].map(lambda x: get_actor2(str(x)))
```

```
In [315]: def get_actor3(x):  
          if len((x.split("screenplay); ")[-1]).split(", ")) < 3:  
              return np.NaN  
          else:  
              return ((x.split("screenplay); ")[-1]).split(", ")[2])
```

```
In [316]: df_2020['actor_3_name'] = df_2020['Cast and crew'].map(lambda x: get_actor3(str(x)))
```

In [317]: df_2020

Out[317]:

	Title	Cast and crew	genres	director_name	actor_1_name	actor_2_name	actor_3_name
0	The Grudge	Nicolas Pesce (director/screenplay); Andrea Ri...	Horror Mystery Thriller	Nicolas Pesce	Andrea Riseborough	Demián Bichir	
1	Underwater	William Eubank (director); Brian Duffield, Ada...	Action Horror Science Fiction Thriller	William Eubank	Kristen Stewart	Vincent Cassel	
2	Like a Boss	Miguel Arteta (director); Sam Pitman, Adam Col...	Comedy	Miguel Arteta	Tiffany Haddish	Rose Byrne	Saln
3	Three Christs	Jon Avnet (director/screenplay); Eric Nazarian...	Drama	Jon Avnet	Richard Gere	Peter Dinklage	
4	Inherit the Viper	Anthony Jerjen (director); Andrew Crabtree (sc...	Crime Thriller Drama	Anthony Jerjen	Josh Hartnett	Margarita Levieva	
...
267	We Can Be Heroes	Robert Rodriguez (director/screenplay); Priyan...	Action Fantasy Family Comedy	Robert Rodriguez	Priyanka Chopra Jonas	Pedro Pascal	YaYa
268	News of the World	Paul Greengrass (director/screenplay); Luke Da...	Adventure Western	Paul Greengrass	Tom Hanks	Helena Zengel	
269	One Night in Miami...	Regina King (director); Kemp Powers (screenpla...	Drama	Regina King	Kingsley Ben-Adir	Eli Goree	Ald
270	Promising Young Woman	Emerald Fennell (director/screenplay); Carey M...	Thriller Crime Drama	Emerald Fennell	Carey Mulligan	Bo Burnham	Al
271	Pieces of a Woman	Kornél Mundruczó (director); Kata Wéber (scree...	Drama	Kornél Mundruczó	Vanessa Kirby	Shia LaBeouf	Mol

272 rows × 7 columns



In [318]: df_2020 = df_2020.rename(columns={'Title':'movie_title'})

In [319]: new_df20 = df_2020.loc[:,['director_name','actor_1_name','actor_2_name','actor_3_

In [320]: new_df20

Out[320]:

	director_name	actor_1_name	actor_2_name	actor_3_name	genres	movie_title
0	Nicolas Pesce	Andrea Riseborough	Demián Bichir	John Cho	Horror Mystery Thriller	The Grudge
1	William Eubank	Kristen Stewart	Vincent Cassel	Jessica Henwick	Action Horror Science Fiction Thriller	Underwater
2	Miguel Arteta	Tiffany Haddish	Rose Byrne	Salma Hayek	Comedy	Like a Boss
3	Jon Avnet	Richard Gere	Peter Dinklage	Walton Goggins	Drama	Three Christs
4	Anthony Jerjen	Josh Hartnett	Margarita Levieva	Chandler Riggs	Crime Thriller Drama	Inherit the Viper
...
267	Robert Rodriguez	Priyanka Chopra Jonas	Pedro Pascal	YaYa Gosselin	Action Fantasy Family Comedy	We Can Be Heroes
268	Paul Greengrass	Tom Hanks	Helena Zengel	NaN	Adventure Western	News of the World
269	Regina King	Kingsley Ben-Adir	Eli Goree	Aldis Hodge	Drama	One Night in Miami...
270	Emerald Fennell	Carey Mulligan	Bo Burnham	Alison Brie	Thriller Crime Drama	Promising Young Woman
271	Kornél Mundruczó	Vanessa Kirby	Shia LaBeouf	Molly Parker	Drama	Pieces of a Woman

272 rows × 6 columns

In [321]: ' + new_df20['actor_2_name'] + ' ' + new_df20['actor_3_name'] + ' ' + new_df20['dir

In [322]: new_df20 = new_df20.dropna(how='any')

In [323]: new_df20['movie_title'] = new_df20['movie_title'].str.lower()

<ipython-input-323-83e83e91073b>:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy (https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)

```
new_df20['movie_title'] = new_df20['movie_title'].str.lower()
```


In [324]:

new_df20

Out[324]:

	director_name	actor_1_name	actor_2_name	actor_3_name	genres	movie_title	c
0	Nicolas Pesce	Andrea Riseborough	Demián Bichir	John Cho	Horror Mystery Thriller	the grudge	An Risebor Dei Bichir, Cho Ni
1	William Eubank	Kristen Stewart	Vincent Cassel	Jessica Henwick	Action Horror Science Fiction Thriller	underwater	Kri Ste Vin Ca Jes Henw
2	Miguel Arteta	Tiffany Haddish	Rose Byrne	Salma Hayek	Comedy	like a boss	Til Hac Rose B Si Hi Migu
3	Jon Avnet	Richard Gere	Peter Dinklage	Walton Goggins	Drama	three christs	Ric Gere F Dink Wi Goc J
4	Anthony Jerjen	Josh Hartnett	Margarita Levieva	Chandler Riggs	Crime Thriller Drama	inherit the viper	Har Marg Lev Char Rig
...	
266	Pete Docter	Jamie Foxx	Tina Fey	Graham Norton	Family Animation Music Comedy Fantasy	soul	Jamie I Tina Gra Norton Doct
267	Robert Rodriguez	Priyanka Chopra Jonas	Pedro Pascal	YaYa Gosselin	Action Fantasy Family Comedy	we can be heroes	Priy: Ch Jonas P Pascal) Gos:
269	Regina King	Kingsley Ben-Adir	Eli Goree	Aldis Hodge	Drama	one night in miami...	King Ben-Ad Goree) He Regi
270	Emerald Fennell	Carey Mulligan	Bo Burnham	Alison Brie	Thriller Crime Drama	promising young woman	C Mulliga Burn Alison Emera

	director_name	actor_1_name	actor_2_name	actor_3_name	genres	movie_title	c
271	Kornél Mundruczó	Vanessa Kirby	Shia LaBeouf	Molly Parker	Drama	pieces of a woman	Van Kirby LaB Molly P Kori

245 rows × 7 columns

Movie 2021

```
In [362]: link = "https://en.wikipedia.org/wiki/List_of_American_films_of_2021"
```

```
In [363]: source = urllib.request.urlopen(link).read()
soup = bs.BeautifulSoup(source, 'lxml')
```

```
In [364]: tables = soup.find_all('table', class_='wikitable sortable')
```

```
In [365]: df1 = pd.read_html(str(tables[0]))[0]
df2 = pd.read_html(str(tables[1]))[0]
df3 = pd.read_html(str(tables[2]))[0]
df4 = pd.read_html(str(tables[3]).replace("'1\"'\",'1'))[0]
```

```
In [366]: df = df1.append(df2.append(df3.append(df4, ignore_index=True), ignore_index=True), i
del df['.mw-parser-output .tooltip-dotted{border-bottom:1px dotted;cursor:help}Re
```

In [367]: df2

Out[367]:

	Opening	Opening.1	Title	Production company	Cast and crew	Ref.
0	APRIL	2	Concrete Cowboy	Netflix	Ricky Staub (director/screenplay); Dan Walsler ...	[53]
1	APRIL	2	The Unholy	Screen Gems / Ghost House Pictures	Evan Spiliotopoulos (director/screenplay); Jef...	[54]
2	APRIL	2	Every Breath You Take	Vertical Entertainment	Vaughn Stein (director); David Murray (screenp...	[55]
3	APRIL	2	Shiva Baby	Utopia	Emma Seligman (director/screenplay); Rachel Se...	NaN
4	APRIL	2	The Girl Who Believes in Miracles	Atlas Distribution Company	Richard Correll (director/screenplay); G.M. Me...	NaN
...
81	JUNE	25	Mary J. Blige's My Life	Amazon Studios	Vanessa Roth (director/screenplay); Mary J. Bl...	NaN
82	JUNE	25	Fathom	Apple TV+	Drew Xanthopoulos (director/screenplay); Miche...	NaN
83	JUNE	25	Rollers	Level 33 Entertainment	Isaiah Smallman (director); Johnny Ray Gill, K...	NaN
84	JUNE	30	America: The Motion Picture	Netflix / Floyd County Productions	Matt Thompson (director); Dave Callahan (scree...	NaN
85	JUNE	30	Zola	A24	Janicza Bravo (director/screenplay); Jeremy O....	[104]

86 rows × 6 columns

In [368]: df_2021 = df[['Title', 'Cast and crew']]

In [369]: df_2021

Out[369]:

	Title	Cast and crew
0	Shadow in the Cloud	Roseanne Liang (director/screenplay); Max Land...
1	The White Tiger	Ramin Bahrani (director/screenplay); Adarsh Go...
2	Locked Down	Doug Liman (director); Steven Knight (screenpl...
3	The Dig	Simon Stone (director); Moira Buffini (screenp...
4	Outside the Wire	Mikael Håfström (director); Rob Yescombe, Rowa...
...
319	The Matrix Resurrections	Lana Wachowski (director/screenplay); David Mi...
320	Sing 2	Garth Jennings (director/screenplay); Matthew ...
321	The King's Man	Matthew Vaughn (director/screenplay); Karl Gaj...
322	American Underdog: The Kurt Warner Story	Erwin brothers (directors); David Aaron, Jon E...
323	Cyrano	Joe Wright (director); Erica Schmidt (screenpl...

324 rows × 2 columns

In [374]: df_2021['genres'] = df_2021['Title'].map(lambda x: get_genre(str(x)))

<ipython-input-374-252c1cda1e49>:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy (https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)

```
df_2021['genres'] = df_2021['Title'].map(lambda x: get_genre(str(x)))
```

In [375]: df_2021

Out[375]:

	Title	Cast and crew	genres
0	Shadow in the Cloud	Roseanne Liang (director/screenplay); Max Land...	Horror Action War
1	The White Tiger	Ramin Bahrani (director/screenplay); Adarsh Go...	Crime Drama
2	Locked Down	Doug Liman (director); Steven Knight (screenpl...	Comedy Crime Drama
3	The Dig	Simon Stone (director); Moira Buffini (screenp...	Drama History
4	Outside the Wire	Mikael Håfström (director); Rob Yescombe, Rowa...	Thriller Action Science Fiction
...
319	The Matrix Resurrections	Lana Wachowski (director/screenplay); David Mi...	Action Adventure Thriller Science Fiction
320	Sing 2	Garth Jennings (director/screenplay); Matthew ...	Family Comedy Animation Drama Music
321	The King's Man	Matthew Vaughn (director/screenplay); Karl Gaj...	War Action Thriller Adventure
322	American Underdog: The Kurt Warner Story	Erwin brothers (directors); David Aaron, Jon E...	Drama
323	Cyrano	Joe Wright (director); Erica Schmidt (screenpl...	Drama Comedy History Romance

324 rows × 3 columns

In [383]: df_2021['director_name'] = df_2021['Cast and crew'].map(lambda x: get_director(st

In [384]: df_2021['actor_1_name'] = df_2021['Cast and crew'].map(lambda x: get_actor1(str(>

In [385]: df_2021['actor_2_name'] = df_2021['Cast and crew'].map(lambda x: get_actor2(str(>

In [381]: df_2021['actor_3_name'] = df_2021['Cast and crew'].map(lambda x: get_actor3(str(>

In [386]: df_2021

Out[386]:

	Title	Cast and crew	genres	director_name	actor_1_name	actor_2_name	act
0	Shadow in the Cloud	Roseanne Liang (director/screenplay); Max Land...	Horror Action War	Roseanne Liang	Chloë Grace Moretz	Taylor John Smith	Be
1	The White Tiger	Ramin Bahrani (director/screenplay); Adarsh Go...	Crime Drama	Ramin Bahrani	Adarsh Gourav	Rajkummar Rao	Ch
2	Locked Down	Doug Liman (director); Steven Knight (screenpl...	Comedy Crime Drama	Doug Liman	Anne Hathaway	Chiwetel Ejiofor	
3	The Dig	Simon Stone (director); Moira Buffini (screenp...	Drama History	Simon Stone	Carey Mulligan	Ralph Fiennes	
4	Outside the Wire	Mikael Håfström (director); Rob Yescombe, Rowa...	Thriller Action Science Fiction	Mikael Håfström	Anthony Mackie	Damson Idris	
...	
319	The Matrix Resurrections	Lana Wachowski (director/screenplay); David Mi...	Action Adventure Thriller Science Fiction	Lana Wachowski	Keanu Reeves	Carrie-Anne Moss	Ya
320	Sing 2	Garth Jennings (director/screenplay); Matthew ...	Family Comedy Animation Drama Music	Garth Jennings	Matthew McConaughey	Reese Witherspoon	
321	The King's Man	Matthew Vaughn (director/screenplay); Karl Gaj...	War Action Thriller Adventure	Matthew Vaughn	Ralph Fiennes	Gemma Arterton	
322	American Underdog: The Kurt Warner Story	Erwin brothers (directors); David Aaron, Jon E...	Drama	Erwin brothers	Zachary Levi	Anna Paquin	
323	Cyrano	Joe Wright (director); Erica Schmidt (screenpl...	Drama Comedy History Romance	Joe Wright	Peter Dinklage	Haley Bennett	

324 rows × 7 columns



In [387]: df_2021 = df_2021.rename(columns={'Title': 'movie_title'})

```
In [389]: new_df21 = df_2021.loc[:, ['director_name', 'actor_1_name', 'actor_2_name', 'actor_3_
```

```
In [390]: new_df21
```

```
Out[390]:
```

	director_name	actor_1_name	actor_2_name	actor_3_name	genres	movie_title
0	Roseanne Liang	Chloë Grace Moretz	Taylor John Smith	Beulah Koale	Horror Action War	Shadow in the Cloud
1	Ramin Bahrani	Adarsh Gourav	Rajkummar Rao	Priyanka Chopra Jonas	Crime Drama	The White Tiger
2	Doug Liman	Anne Hathaway	Chiwetel Ejiofor	Stephen Merchant	Comedy Crime Drama	Locked Down
3	Simon Stone	Carey Mulligan	Ralph Fiennes	Lily James	Drama History	The Dig
4	Mikael Håfström	Anthony Mackie	Damson Idris	Emily Beecham	Thriller Action Science Fiction	Outside the Wire
...
319	Lana Wachowski	Keanu Reeves	Carrie-Anne Moss	Yahya Abdul-Mateen II	Action Adventure Thriller Science Fiction	The Matrix Resurrections
320	Garth Jennings	Matthew McConaughey	Reese Witherspoon	Scarlett Johansson	Family Comedy Animation Drama Music	Sing 2
321	Matthew Vaughn	Ralph Fiennes	Gemma Arterton	Rhys Ifans	War Action Thriller Adventure	The King's Man
322	Erwin brothers	Zachary Levi	Anna Paquin	Ser'Darius Blain	Drama	American Underdog: The Kurt Warner Story
323	Joe Wright	Peter Dinklage	Haley Bennett	Kelvin Harrison Jr.	Drama Comedy History Romance	Cyrano

324 rows × 6 columns

```
In [391]: new_df21['comb'] = new_df21['actor_1_name'] + ' ' + new_df21['actor_2_name'] + ' ' + new_df21['actor_3_name']
```

```
In [392]: new_df21 = new_df21.dropna(how='any')
```

```
In [393]: new_df21['movie_title'] = new_df21['movie_title'].str.lower()
```

```
<ipython-input-393-6e765124d575>:1: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy (https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)

```
new_df21['movie_title'] = new_df21['movie_title'].str.lower()
```



In [395]:

new_df21

Out[395]:

	director_name	actor_1_name	actor_2_name	actor_3_name	genres	movie_title	cc
0	Roseanne Liang	Chloë Grace Moretz	Taylor John Smith	Beulah Koale	Horror Action War	shadow in the cloud	Chloë Gr Moretz Ta John Sr Beulah K
1	Ramin Bahrani	Adarsh Gourav	Rajkummar Rao	Priyanka Chopra Jonas	Crime Drama	the white tiger	Adā Got Rajkum Rao Priya Chopra .
2	Doug Liman	Anne Hathaway	Chiwetel Ejiofor	Stephen Merchant	Comedy Crime Drama	locked down	A Hatha Chiw Eji Step Mercha
3	Simon Stone	Carey Mulligan	Ralph Fiennes	Lily James	Drama History	the dig	Cā Mulli Ralph Fien Lily Jar Simo
4	Mikael Håfström	Anthony Mackie	Damson Idris	Emily Beecham	Thriller Action Science Fiction	outside the wire	Anth Ma Damson I El Beech Mil
...	
319	Lana Wachowski	Keanu Reeves	Carrie-Anne Moss	Yahya Abdul-Mateen II	Action Adventure Thriller Science Fiction	the matrix resurrections	Ke Ree Carrie-A Moss Ya Abdul-Ma
320	Garth Jennings	Matthew McConaughey	Reese Witherspoon	Scarlett Johansson	Family Comedy Animation Drama Music	sing 2	Mattl McConaug Re Withersp Scarle
321	Matthew Vaughn	Ralph Fiennes	Gemma Arterton	Rhys Ifans	War Action Thriller Adventure	the king's man	Ralph Fien Gerr Arterton R Ifans Mattl
322	Erwin brothers	Zachary Levi	Anna Paquin	Ser'Darius Blain	Drama	american underdog: the kurt warner story	Zachary I Anna Pac Ser'Da Blain Er
323	Joe Wright	Peter Dinklage	Haley Bennett	Kelvin Harrison Jr.	Drama Comedy History Romance	cyrano	P Dinkl Haley Ben Ke Harrison

302 rows × 7 columns



```
In [396]: old_df = pd.read_csv('final_data.csv')
```

In [397]:

old_df

Out[397]:

	director_name	actor_1_name	actor_2_name	actor_3_name	genres	movie_title	comb
0	John Lasseter	Tom Hanks	Tim Allen	Don Rickles	Animation Comedy Family	toy story	Tom Hanks Tim Allen Don Rickles John Lasseter ...
1	Joe Johnston	Robin Williams	Jonathan Hyde	Kirsten Dunst	Adventure Fantasy Family	jumanji	Robin Williams Jonathan Hyde Kirsten Dunst Joe...
2	Howard Deutch	Walter Matthau	Jack Lemmon	Ann-Margret	Romance Comedy	grumpier old men	Walter Matthau Jack Lemmon Ann-Margret Howard ...
3	Forest Whitaker	Whitney Houston	Angela Bassett	Loretta Devine	Comedy Drama Romance	waiting to exhale	Whitney Houston Angela Bassett Loretta Devine ...
4	Charles Shyer	Steve Martin	Diane Keaton	Martin Short	Comedy	father of the bride part ii	Steve Martin Diane Keaton Martin Short Charles...
...
39656	Nick Bruno, Troy Quane	Will Smith	Tom Holland	Rashida Jones	Animation Action Adventure Comedy Family	spies in disguise	Will Smith Tom Holland Rashida Jones Nick Brun...
39657	Greta Gerwig	Saoirse Ronan	Emma Watson	Florence Pugh	Drama Romance	little women	Saoirse Ronan Emma Watson Florence Pugh Greta ...

	director_name	actor_1_name	actor_2_name	actor_3_name	genres	movie_title	comb
39658	Sam Mendes	George MacKay	Dean-Charles Chapman	Mark Strong	War Drama Action Thriller	1917	George MacKay Dean- Charles Chapman Mark Strong...
39659	Destin Daniel Cretton	Michael B. Jordan	Jamie Foxx	Brie Larson	Drama Crime History	just mercy	Michael B. Jordan Jamie Foxx Brie Larson Desti...
39660	Chinonye Chukwu	Alfre Woodard	Wendell Pierce	Aldis Hodge	Drama	clemency	Alfre Woodard Wendell Pierce Aldis Hodge Chino...

39661 rows × 7 columns



```
In [398]: final_df = old_df.append(new_df20.append(new_df21,ignore_index=True),ignore_index=True)
```

In [399]:

final_df

Out[399]:

	director_name	actor_1_name	actor_2_name	actor_3_name	genres	movie_title	
0	John Lasseter	Tom Hanks	Tim Allen	Don Rickles	Animation Comedy Family	toy story	Tom Tim Alle Rickles Lass
1	Joe Johnston	Robin Williams	Jonathan Hyde	Kirsten Dunst	Adventure Fantasy Family	jumanji	W Jor Hyde k Dunst
2	Howard Deutch	Walter Matthau	Jack Lemmon	Ann-Margret	Romance Comedy	grumpier old men	Matthau Lemmon M How
3	Forest Whitaker	Whitney Houston	Angela Bassett	Loretta Devine	Comedy Drama Romance	waiting to exhale	W Hc / B L Dev
4	Charles Shyer	Steve Martin	Diane Keaton	Martin Short	Comedy	father of the bride part ii	Steve Diane k Martin Ch
...	
40203	Lana Wachowski	Keanu Reeves	Carrie-Anne Moss	Yahya Abdul- Mateen II	Action Adventure Thriller Science Fiction	the matrix resurrections	I R Carrie Moss Abdul-M
40204	Garth Jennings	Matthew McConaughey	Reese Witherspoon	Scarlett Johansson	Family Comedy Animation Drama Music	sing 2	Ma McCona l Wither Sc
40205	Matthew Vaughn	Ralph Fiennes	Gemma Arterton	Rhys Ifans	War Action Thriller Adventure	the king's man	Ralph Fi G Arterton Ifans Ma
40206	Erwin brothers	Zachary Levi	Anna Paquin	Ser'Darius Blain	Drama	american underdog: the kurt warner story	Zachar Anna F Ser'l Blain
40207	Joe Wright	Peter Dinklage	Haley Bennett	Kelvin Harrison Jr.	Drama Comedy History Romance	cyrano	Dii Haley B Harris

40208 rows × 7 columns

```
In [400]: final_df.to_csv('main_data.csv', index=False)
```

```
In [ ]:
```