

Assignment

March 18, 2022

1 PROBLEM STATEMENT

Amazon is a multinational company which focusses on ecommerce, cloud computing, digital streaming, and artificial intelligence. Amazon prime is a subscription service offered by Amazon where it deliver music, gaming, video's etc for the customer's. As a datascientist at Amazon, you are assigned with a task to bring some insights from the historical data that includes information such as show_id, type of the show, title of the show, director of the show, cast, country, year, duration etc..

Let's start exploring the data set

2 IMPORTING LIBRARIES

```
[1]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px
```

3 LOADING DATA

```
[4]: prime = pd.read_csv("../..../datasets/data.csv")
prime
```

```
[4]:
```

	show_id	type	title	director	\
0	s1	Movie	The Grand Seduction	Don McKellar	
1	s2	Movie	Take Care Good Night	Girish Joshi	
2	s3	Movie	Secrets of Deception	Josh Webber	
3	s4	Movie	Pink: Staying True	Sonia Anderson	
4	s5	Movie	Monster Maker	Giles Foster	
...	
9663	s9664	Movie	Pride Of The Bowery	Joseph H. Lewis	
9664	s9665	TV Show	Planet Patrol	NaN	
9665	s9666	Movie	Outpost	Steve Barker	
9666	s9667	TV Show	Maradona: Blessed Dream	NaN	

```
9667    s9668    Movie                Harry Brown    Daniel Barber
```

```

                                cast                country \
0      Brendan Gleeson, Taylor Kitsch, Gordon Pinsent      Canada
1      Mahesh Manjrekar, Abhay Mahajan, Sachin Khedekar      India
2      Tom Sizemore, Lorenzo Lamas, Robert LaSardo, R...    United States
3      Interviews with: Pink, Adele, Beyoncé, Britney...    United States
4      Harry Dean Stanton, Kieran O'Brien, George Cos...  United Kingdom
...
9663                                Leo Gorcey, Bobby Jordan      NaN
9664    DICK VOSBURGH, RONNIE STEVENS, LIBBY MORRIS, M...      NaN
9665    Ray Stevenson, Julian Wadham, Richard Brake, M...      NaN
9666    Esteban Recagno, Ezequiel Stremiz, Luciano Vit...      NaN
9667    Michael Caine, Emily Mortimer, Joseph Gilgun, ...      NaN

    date_added  release_year  rating    duration                listed_in \
0    March 30, 2021          2014     NaN    113 min      Comedy, Drama
1    March 30, 2021          2018    13+    110 min      Drama, International
2    March 30, 2021          2017     NaN    74 min    Action, Drama, Suspense
3    March 30, 2021          2014     NaN    69 min      Documentary
4    March 30, 2021          1989     NaN    45 min      Drama, Fantasy
...
9663            NaN            1940     7+     60 min      Comedy
9664            NaN            2018    13+  4 Seasons      TV Shows
9665            NaN            2008     R     90 min      Action
9666            NaN            2021  TV-MA    1 Season      Drama, Sports
9667            NaN            2010     R    103 min    Action, Drama, Suspense

```

```

                                description
0      A small fishing village must procure a local d...
1      A Metro Family decides to fight a Cyber Crimin...
2      After a man discovers his wife is cheating on ...
3      Pink breaks the mold once again, bringing her ...
4      Teenage Matt Banting wants to work with a famo...
...
9663    New York City street principles get an East Si...
9664    This is Earth, 2100AD - and these are the adve...
9665    In war-torn Eastern Europe, a world-weary grou...
9666    The series tells the story of Diego Maradona, ...
9667    Harry Brown, starring two-time Academy Award w...

```

```
[9668 rows x 12 columns]
```

The dataset contains 9668 rows and 12 columns. Each row demonstrates information about a unique movie/show.

```
[6]: prime.describe()
```

```
[6]:      release_year
count    9668.000000
mean      2008.341849
std        18.922482
min       1920.000000
25%       2007.000000
50%       2016.000000
75%       2019.000000
max        2021.000000
```

The dataset has records from 1920 to 2021. In this dataset, there are 9668 rows and 12 columns. Also, there are many missing values in this dataset so let's focus on data cleaning.

4 CHECKING MISSING VALUES

```
[7]: prime[prime.isnull()]
```

```
[7]:      show_id type title director cast country date_added  release_year rating \
0         NaN  NaN   NaN         NaN  NaN      NaN         NaN         NaN  NaN
1         NaN  NaN   NaN         NaN  NaN      NaN         NaN         NaN  NaN
2         NaN  NaN   NaN         NaN  NaN      NaN         NaN         NaN  NaN
3         NaN  NaN   NaN         NaN  NaN      NaN         NaN         NaN  NaN
4         NaN  NaN   NaN         NaN  NaN      NaN         NaN         NaN  NaN
...      ...  ...   ...         ...  ...      ...         ...         ...
9663      NaN  NaN   NaN         NaN  NaN      NaN         NaN         NaN  NaN
9664      NaN  NaN   NaN         NaN  NaN      NaN         NaN         NaN  NaN
9665      NaN  NaN   NaN         NaN  NaN      NaN         NaN         NaN  NaN
9666      NaN  NaN   NaN         NaN  NaN      NaN         NaN         NaN  NaN
9667      NaN  NaN   NaN         NaN  NaN      NaN         NaN         NaN  NaN

      duration listed_in description
0         NaN         NaN         NaN
1         NaN         NaN         NaN
2         NaN         NaN         NaN
3         NaN         NaN         NaN
4         NaN         NaN         NaN
...      ...         ...         ...
9663      NaN         NaN         NaN
9664      NaN         NaN         NaN
9665      NaN         NaN         NaN
9666      NaN         NaN         NaN
9667      NaN         NaN         NaN

[9668 rows x 12 columns]
```

```
[8]: #show the count of null values
prime.isnull().sum()
```

```
[8]: show_id      0
     type        0
     title       0
     director    2082
     cast        1233
     country     8996
     date_added  9513
     release_year 0
     rating      337
     duration    0
     listed_in   0
     description 0
     dtype: int64
```

From the above results, we find that there are many missing values in the columns such as "director", "cast", "country", "date added", and "ratings".

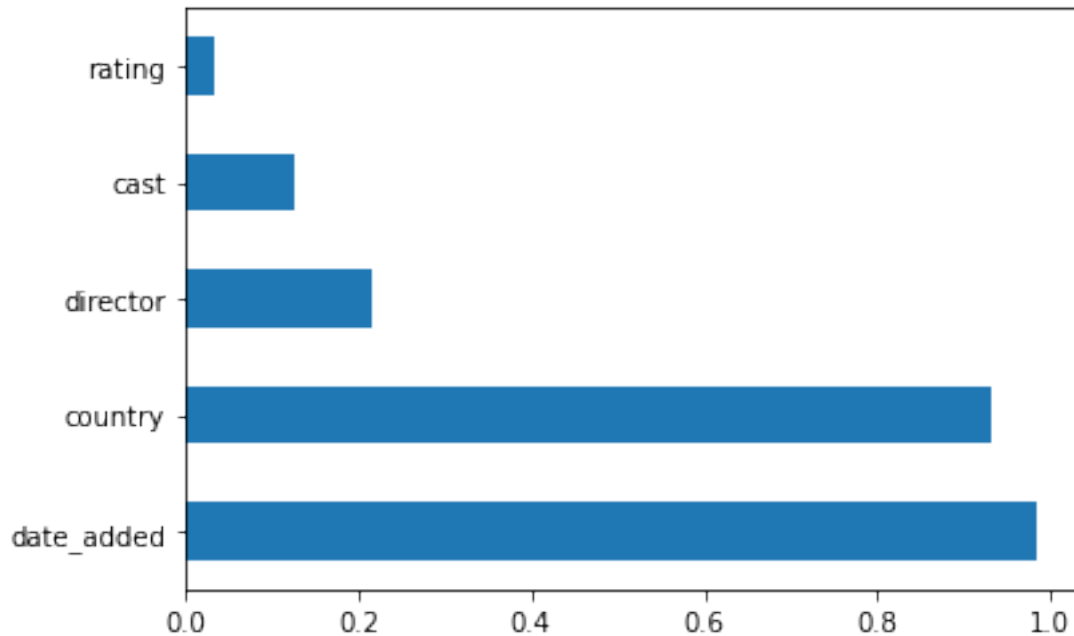
5 visualizing the percentage of the count of null values in the dataset.

```
[9]: values_missing_percent = prime.isnull().sum().sort_values(ascending=False)/
     ↪len(prime)
     values_missing_percent
```

```
[9]: date_added    0.983968
     country      0.930492
     director     0.215350
     cast         0.127534
     rating       0.034857
     show_id      0.000000
     type         0.000000
     title        0.000000
     release_year 0.000000
     duration     0.000000
     listed_in    0.000000
     description  0.000000
     dtype: float64
```

```
[10]: values_missing_percent[values_missing_percent!= 0].plot(kind='barh')
```

```
[10]: <AxesSubplot:>
```



From the above plots, it's clear that "date_added" and "country" have the highest number of null values.

6 Exploratory data analysis

Amazon would like to derive outputs for the below-mentioned business questions: 1. In which year was the highest number of tv shows or movies released? 2. Find out the total number of TV shows and movies that amazon prime has released so far? 3. Find out how many films were released in the year 2021 ? 4. What is the maximum duration of a Movie/ TV Show in Prime? 5. Find out the director who has given the highest number of TV Shows and Movies? 6. Find out the ratings given to all Movies/Tv Show directed by "Mark Knight", who has released the highest number of TV Shows/Movies? 7. Which country has released the highest number of Tv Shows and Movies? 8. Find out the maximum duration of TV Show/Movie released in the year 2021

7 We are going to find out which year does amazon prime released the highest number of TV shows and movies?

```
[11]: prime.head()
```

```
[11]:  show_id  type      title      director \
0      s1  Movie  The Grand Seduction  Don McKellar
1      s2  Movie  Take Care Good Night  Girish Joshi
```

```

2      s3  Movie  Secrets of Deception      Josh Webber
3      s4  Movie    Pink: Staying True  Sonia Anderson
4      s5  Movie          Monster Maker    Giles Foster

```

```

                                cast      country \
0      Brendan Gleeson, Taylor Kitsch, Gordon Pinsent      Canada
1      Mahesh Manjrekar, Abhay Mahajan, Sachin Khedekar      India
2      Tom Sizemore, Lorenzo Lamas, Robert LaSardo, R...  United States
3      Interviews with: Pink, Adele, Beyoncé, Britney...  United States
4      Harry Dean Stanton, Kieran O'Brien, George Cos...  United Kingdom

```

```

      date_added  release_year  rating  duration      listed_in \
0  March 30, 2021      2014    NaN  113 min      Comedy, Drama
1  March 30, 2021      2018    13+  110 min      Drama, International
2  March 30, 2021      2017    NaN   74 min  Action, Drama, Suspense
3  March 30, 2021      2014    NaN   69 min      Documentary
4  March 30, 2021      1989    NaN   45 min      Drama, Fantasy

```

```

                                description
0  A small fishing village must procure a local d...
1  A Metro Family decides to fight a Cyber Crimin...
2  After a man discovers his wife is cheating on ...
3  Pink breaks the mold once again, bringing her ...
4  Teenage Matt Banting wants to work with a famo...

```

```

[12]: count = prime["release_year"].value_counts()
      df = count.sort_values(ascending=False)
      df

```

```

[12]: 2021      1442
      2020       962
      2019       929
      2018       623
      2017       562
      ...
      1922         2
      1926         2
      1924         1
      1923         1
      1927         1
      Name: release_year, Length: 100, dtype: int64

```

```

[15]: df.head(5)

```

```

[15]: 2021      1442
      2020       962
      2019       929

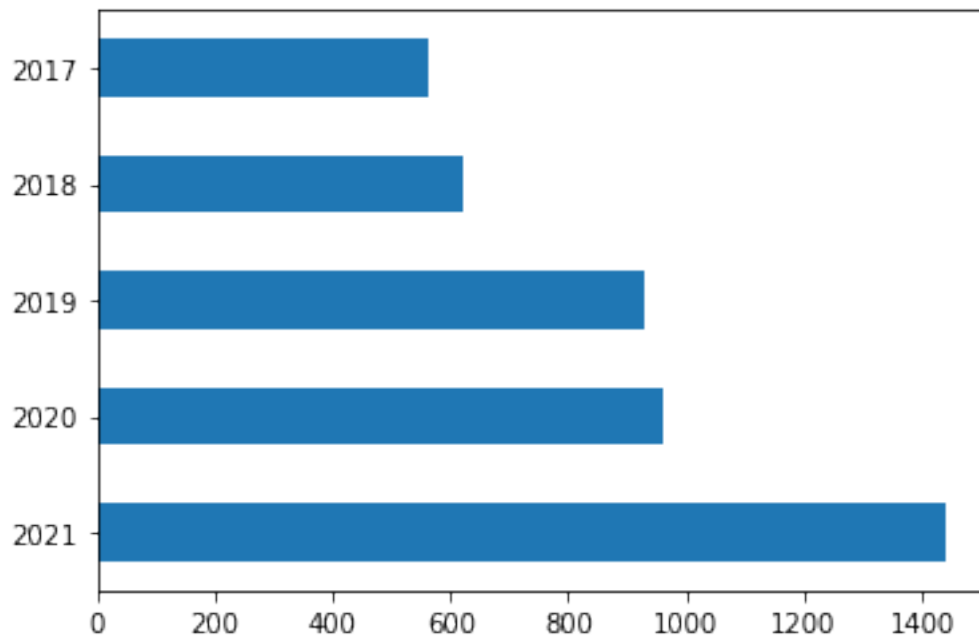
```

```
2018      623
2017      562
Name: release_year, dtype: int64
```

8 visualizing the results:

```
[14]: df[:5].plot(kind="barh")
```

```
[14]: <AxesSubplot:>
```



From the above results, it's clear that Amazon Prime has released the highest number of TVshows and movies in 2021.

9 Find out the total number of TV shows and movies that amazon prime has released so far?

```
[16]: prime.head()
```

```
[16]:  show_id  type      title      director \
0      s1  Movie  The Grand Seduction  Don McKellar
1      s2  Movie  Take Care Good Night  Girish Joshi
2      s3  Movie  Secrets of Deception  Josh Webber
```

```

3      s4  Movie      Pink: Staying True  Sonia Anderson
4      s5  Movie      Monster Maker      Giles Foster

```

```

                                cast      country \
0      Brendan Gleeson, Taylor Kitsch, Gordon Pinsent      Canada
1      Mahesh Manjrekar, Abhay Mahajan, Sachin Khedekar      India
2      Tom Sizemore, Lorenzo Lamas, Robert LaSardo, R...  United States
3      Interviews with: Pink, Adele, Beyoncé, Britney...  United States
4      Harry Dean Stanton, Kieran O'Brien, George Cos...  United Kingdom

```

```

      date_added  release_year  rating  duration      listed_in \
0  March 30, 2021      2014      NaN  113 min      Comedy, Drama
1  March 30, 2021      2018      13+  110 min      Drama, International
2  March 30, 2021      2017      NaN   74 min  Action, Drama, Suspense
3  March 30, 2021      2014      NaN   69 min      Documentary
4  March 30, 2021      1989      NaN   45 min      Drama, Fantasy

```

```

                                description
0  A small fishing village must procure a local d...
1  A Metro Family decides to fight a Cyber Crimin...
2  After a man discovers his wife is cheating on ...
3  Pink breaks the mold once again, bringing her ...
4  Teenage Matt Banting wants to work with a famo...

```

```
[17]: prime.groupby("type").type.count()
```

```

[17]: type
Movie      7814
TV Show    1854
Name: type, dtype: int64

```

```
[18]: import seaborn as sns
sns.countplot(prime["type"])
```

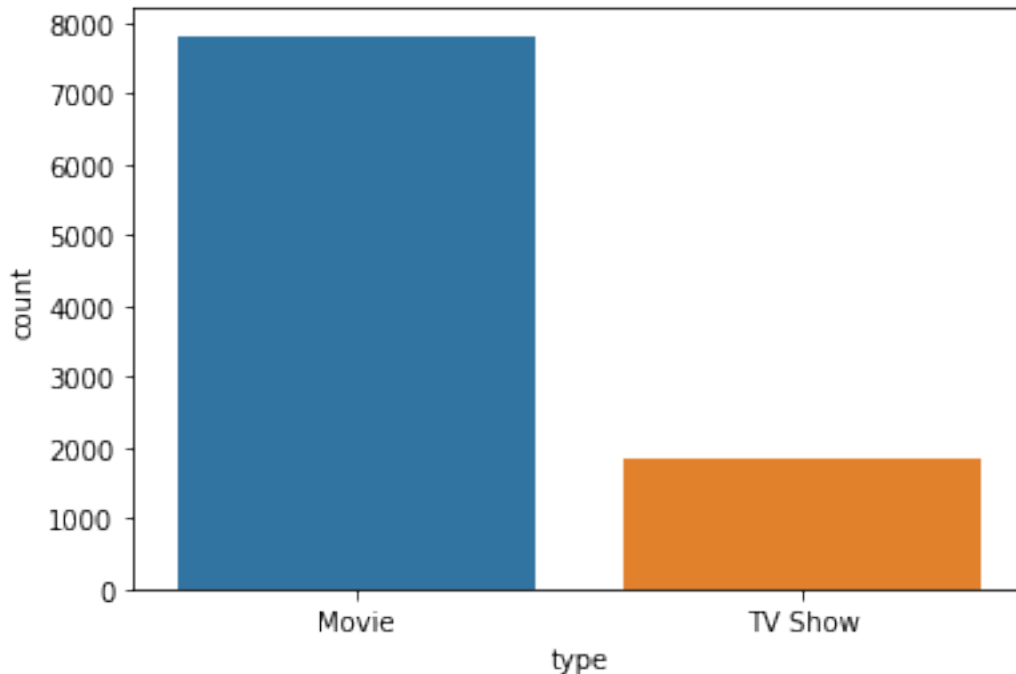
```

/usr/local/lib/python3.8/dist-packages/seaborn/_decorators.py:36: FutureWarning:
Pass the following variable as a keyword arg: x. From version 0.12, the only
valid positional argument will be `data`, and passing other arguments without an
explicit keyword will result in an error or misinterpretation.

```

```
warnings.warn(
```

```
[18]: <AxesSubplot:xlabel='type', ylabel='count'>
```

From the above results, it's clear that prime has released a total of 7814 Movies and 1854 Tv show's from the year 1920 - to 2021.

10 Find out how many movies were released in the year 2021 ?(Amazon prime has released the highest number of TVshows and filmsin the year 2021)

```
[19]: prime[(prime["type"] == "Movie") & (prime["release_year"] == 2021)]
```

```
[19]:   show_id  type      title \
96      s97  Movie      Wildlife
113     s114  Movie      White Tiger
151     s152  Movie      War of Likes
202     s203  Movie  V1 Murder Case (Telugu)
220     s221  Movie      Underplayed
...     ...   ...
9628    s9629  Movie      Thalaivii (Malayalam)
9629    s9630  Movie      Thalaivii (Kannada)
9656    s9657  Movie      Anaganaga Oka Nenu
9660    s9661  Movie      The Man in the Hat
9662    s9663  Movie      River
```

	director \
96	Frank Hall Green
113	Karen Shakhnazarov
151	María Ripoll
202	Pavel Navageethan
220	Stacey Lee
...	...
9628	A.L. Vijay
9629	A.L. Vijay
9656	Vamshi P
9660	John-Paul Davidson, Stephen Warbeck
9662	Emily Skye

	cast	country	date_added \
96	Bruce Greenwood, Ella Purnell, Brian Geraghty	NaN	NaN
113	Aleksey Vertkov, Vitaliy Kishchenko, Valeriy G...	NaN	NaN
151	Ludwika Paleta, Regina Blandón, Manolo Cardona...	NaN	NaN
202	Ram Arun Castro, Vishnupriya Pillai, Gayathri ...	NaN	NaN
220	TokiMonsta, Alison Wonderland, Nervo, Rezz, Ni...	NaN	NaN
...
9628	Kangana Ranaut, Arvind Swamy, Nassar	NaN	NaN
9629	Kangana Ranaut, Arvind Swamy, Nassar	NaN	NaN
9656	Vijay, Sweta chaudhary, NNR CHOWDARY, Hariya	NaN	NaN
9660	Ciaran Hinds, Stephen Dillane, Maiwenn	NaN	NaN
9662	Mary Cameron Rogers, Alexandra Rose, Rob Marsh...	NaN	NaN

	release_year	rating	duration	listed_in \
96	2021	16+	104 min	Action, Drama
113	2021	16+	109 min	Action, Science Fiction
151	2021	16+	104 min	Comedy
202	2021	13+	111 min	Action, Suspense
220	2021	13+	90 min	Documentary
...
9628	2021	ALL	151 min	Drama
9629	2021	ALL	151 min	Drama
9656	2021	18+	135 min	Action, Drama, Suspense
9660	2021	13+	96 min	Comedy
9662	2021	16+	93 min	Drama, Science Fiction, Suspense

	description
96	An unlikely friendship forms in the spectacula...
113	Great Patriotic War, early 1940s. After barely...
151	In order to advance her career in the dynamic ...
202	The film begins with the murder of a girl who ...
220	Filmed over the summer festival season, Underp...
...	...
9628	Thalaivi is the dramatic story of a woman who ...

```

9629 Thalaivi is the dramatic story of a woman who ...
9656 Peter(hero) Who went to the forest to find a c...
9660 The Man in the Hat journeys through France in ...
9662 River is a grounded Sci-Fi mystery Thriller, t...

```

```
[1139 rows x 12 columns]
```

```
[20]: prime[(prime["type"] == "TV Show") & (prime["release_year"] == 2021)]
```

```

[20]:      show_id      type      title director \
134      s135  TV Show  WGC-Dell Technologies Match Play Reveal Show      NaN
370      s371  TV Show                                     THEM      NaN
401      s402  TV Show                                     The Vineyard      NaN
404      s405  TV Show                      The Underground Railroad      NaN
470      s471  TV Show                                     The Record      NaN
...      ...      ...      ...      ...
9637     s9638  TV Show                      Paul Newman Trilogy      NaN
9642     s9643  TV Show                      Lunkerville      NaN
9645     s9646  TV Show      Let's Go Cozy Coupe! - Cartoons for Kids      NaN
9655     s9656  TV Show                      DaFUNdamentals Series      NaN
9666     s9667  TV Show                      Maradona: Blessed Dream      NaN

```

```

                                cast      country \
134                                NaN      NaN
370  Deborah Ayorinde, Ashley Thomas, Alison Pill, ...      NaN
401  Leonor Watling, Rafael Novoa, Nathaniel Parker...      NaN
404  Thuso Mbedu, Joel Edgerton, Chase W. Dillon, W...  United States
470                                NaN      NaN
...      ...      ...
9637                                NaN      NaN
9642                                Michael de Avila      NaN
9645                                NaN      NaN
9655      Emile Crawford, Randall Crawford, Jon Ward      NaN
9666  Esteban Recagno, Ezequiel Stremiz, Luciano Vit...      NaN

```

```

      date_added  release_year  rating  duration \
134            NaN          2021  TV-NR    1 Season
370            NaN          2021  TV-14    1 Season
401            NaN          2021    16+    1 Season
404            NaN          2021  TV-MA    1 Season
470            NaN          2021    NaN    1 Season
...      ...      ...      ...
9637            NaN          2021    13+    1 Season
9642  October 10, 2021          2021    13+    2 Seasons
9645            NaN          2021    ALL    1 Season
9655            NaN          2021  TV-Y    1 Season
9666            NaN          2021  TV-MA    1 Season

```

	listed_in \	
134	Sports	
370	Drama, Horror, Suspense	
401	Drama	
404	Drama	
470	Documentary, Sports	
...	...	
9637	TV Shows	
9642	Documentary, Special Interest, Sports	
9645	Kids	
9655	Animation, Kids	
9666	Drama, Sports	

	description
134	Host Jonathan Coachman and PGATOUR.COM Senior ...
370	THEM is a limited anthology series that explor...
401	The Vineyard is the story of Mauro Larrea and ...
404	Coming soon to Prime Video: May 14. The Underg...
470	The plan to smash the world record for attenda...
...	...
9637	The early days of television were filled with ...
9642	Lunkerville's launches "Get Folks Fishing", a ...
9645	Cozy Coupe is an online series that helps kids...
9655	Learn "daFundamentals" through song and fun ch...
9666	The series tells the story of Diego Maradona, ...

[303 rows x 12 columns]

From the above results, there are 1139 movies and 303 Tv Shows were released in the year 2021

11 What is the maximum duration of a Movie/ TV Show in Prime?

```
[22]: prime["duration"].unique()
```

```
[22]: array(['113 min', '110 min', '74 min', '69 min', '45 min', '52 min',
          '98 min', '131 min', '87 min', '92 min', '88 min', '93 min',
          '94 min', '46 min', '96 min', '1 Season', '104 min', '62 min',
          '50 min', '3 Seasons', '2 Seasons', '86 min', '36 min', '37 min',
          '103 min', '9 min', '18 min', '14 min', '20 min', '19 min',
          '22 min', '60 min', '6 min', '54 min', '5 min', '84 min',
          '126 min', '125 min', '109 min', '89 min', '85 min', '56 min',
          '40 min', '111 min', '33 min', '34 min', '95 min', '99 min',
          '78 min', '4 Seasons', '77 min', '55 min', '53 min', '115 min',
```

```
'58 min', '49 min', '135 min', '91 min', '64 min', '59 min',
'48 min', '122 min', '90 min', '102 min', '65 min', '114 min',
'136 min', '70 min', '138 min', '100 min', '480 min', '4 min',
'30 min', '152 min', '68 min', '57 min', '7 Seasons', '31 min',
'151 min', '149 min', '9 Seasons', '141 min', '121 min', '79 min',
'140 min', '51 min', '106 min', '75 min', '27 min', '107 min',
'108 min', '38 min', '157 min', '43 min', '118 min', '139 min',
'6 Seasons', '112 min', '15 min', '72 min', '5 Seasons', '116 min',
'142 min', '71 min', '42 min', '81 min', '32 min', '66 min',
'127 min', '159 min', '67 min', '29 min', '132 min', '101 min',
'164 min', '73 min', '61 min', '80 min', '83 min', '44 min',
'120 min', '26 min', '97 min', '23 min', '105 min', '82 min',
'11 min', '148 min', '161 min', '123 min', '29 Seasons', '0 min',
'124 min', '143 min', '35 min', '47 min', '170 min', '19 Seasons',
'3 min', '146 min', '601 min', '24 min', '21 Seasons', '154 min',
'128 min', '133 min', '153 min', '119 min', '63 min', '169 min',
'174 min', '144 min', '7 min', '137 min', '76 min', '39 min',
'8 Seasons', '12 Seasons', '134 min', '163 min', '1 min',
'145 min', '162 min', '41 min', '147 min', '155 min', '117 min',
'167 min', '11 Seasons', '28 min', '25 min', '180 min', '2 min',
'541 min', '240 min', '129 min', '178 min', '171 min', '21 min',
'172 min', '173 min', '10 min', '166 min', '160 min', '130 min',
'479 min', '13 min', '8 min', '10 Seasons', '17 min', '16 min',
'158 min', '183 min', '12 min', '14 Seasons', '150 min', '481 min',
'181 min', '156 min', '540 min', '177 min', '550 min', '485 min',
'176 min', '193 min', '165 min', '175 min', '188 min', '187 min',
'168 min', '190 min', '185 min', '209 min', '192 min', '182 min',
'207 min', '269 min', '15 Seasons', '191 min'], dtype=object)
```

```
[23]: prime["duration"].dtypes
```

```
[23]: dtype('O')
```

```
[24]: #splitting the duration column into mts and unit
prime[['mts', 'Unit']] = prime['duration'].str.split(' ', expand=True)
prime.head()
```

```
[24]:
```

	show_id	type	title	director \	cast	country \
0	s1	Movie	The Grand Seduction	Don McKellar	Brendan Gleeson, Taylor Kitsch, Gordon Pinsent	Canada
1	s2	Movie	Take Care Good Night	Girish Joshi	Mahesh Manjrekar, Abhay Mahajan, Sachin Khedekar	India
2	s3	Movie	Secrets of Deception	Josh Webber		
3	s4	Movie	Pink: Staying True	Sonia Anderson		
4	s5	Movie	Monster Maker	Giles Foster		

2	Tom Sizemore, Lorenzo Lamas, Robert LaSardo, R...	United States
3	Interviews with: Pink, Adele, Beyoncé, Britney...	United States
4	Harry Dean Stanton, Kieran O'Brien, George Cos...	United Kingdom

	date_added	release_year	rating	duration	listed_in \
0	March 30, 2021	2014	NaN	113 min	Comedy, Drama
1	March 30, 2021	2018	13+	110 min	Drama, International
2	March 30, 2021	2017	NaN	74 min	Action, Drama, Suspense
3	March 30, 2021	2014	NaN	69 min	Documentary
4	March 30, 2021	1989	NaN	45 min	Drama, Fantasy

	description	mts	Unit
0	A small fishing village must procure a local d...	113	min
1	A Metro Family decides to fight a Cyber Crimin...	110	min
2	After a man discovers his wife is cheating on ...	74	min
3	Pink breaks the mold once again, bringing her ...	69	min
4	Teenage Matt Banting wants to work with a famo...	45	min

```
[25]: prime["mts"].max()
```

```
[25]: '99'
```

The maximum duration of a TV Show/Movie is 99 mts.

12 Find out the director who has given highest number of TV Shows and Movies

```
[26]: prime["director"].value_counts().head(5)
```

```
[26]: Mark Knight          113
      Cannis Holder        61
      Moonbug Entertainment  37
      Jay Chapman          34
      Arthur van Merwijk    30
      Name: director, dtype: int64
```

From the above results "Mark Knight" has released highest number of Tv shows/Movies in Amazon prime

```
[27]: prime[(prime["release_year"] == 2021)& (prime["director"] == "Mark Knight")]
```

```
[27]: Empty DataFrame
      Columns: [show_id, type, title, director, cast, country, date_added,
      release_year, rating, duration, listed_in, description, mts, Unit]
      Index: []
```

As per the given data "Mark Knight" haven't released any TV Shows or Movies in the year 2021 whereas the highest number of TV Shows and Movies were released in the year 2021

13 Find out the ratings given to all Movies/Tv Show directed by "Mark Knight" who has released highest number of TV Shows/Movies

```
[28]: prime[prime["director"] == "Mark Knight"].rating.unique()
```

```
[28]: array(['ALL', nan], dtype=object)
```

```
[29]: prime["rating"].unique()
```

```
[29]: array([nan, '13+', 'ALL', '18+', 'R', 'TV-Y', 'TV-Y7', 'NR', '16+',  
          'TV-PG', '7+', 'TV-14', 'TV-NR', 'TV-G', 'PG-13', 'TV-MA', 'G',  
          'PG', 'NC-17', 'UNRATED', '16', 'AGES_16_', 'AGES_18_', 'ALL_AGES',  
          'NOT_RATE'], dtype=object)
```

It seems like he was given rating = "ALL". We need to investigate on the priority list for the rating since there 24 are unique values.

14 Which country has released highest number of Tv Shows and Movies?

```
[30]: prime[prime["type"] == "TV Show"].country.value_counts()
```

```
[30]: United States          59  
      India                 16  
      Spain                 5  
      United Kingdom        4  
      United Kingdom, United States 3  
      Italy                 2  
      Japan                 2  
      Canada, United States  2  
      Mexico                2  
      Germany               2  
      United States, United Kingdom, Canada 1  
      United Kingdom, Canada, United States 1  
      Chile, Argentina, United States 1  
      New Zealand           1  
      Czech Republic, Germany 1  
      France                1  
      Name: country, dtype: int64
```

```
[31]: prime[prime["type"] == "Movie"].country.value_counts()
```

```
[31]: India                213
      United States        194
      United Kingdom        24
      Canada                16
      United Kingdom, United States    9

      ...
      Australia, United States, Germany    1
      Germany, Austria, United States    1
      France, China, Belgium, Monaco    1
      Egypt    1
      Italy, Spain    1
      Name: country, Length: 79, dtype: int64
```

"United States" has released highest number of TV Shows and "India" has released highest number of Movies

15 Find out the maximum duration of TV Show/Movie released in the year 2021

```
[32]: prime[prime["release_year"] == 2021].duration.max()
```

```
[32]: '99 min'
```

```
[33]: dur = prime[(prime["release_year"] == 2021) & (prime["type"] == "Movie")].
      ↪duration.max()
      dur
```

```
[33]: '99 min'
```

The maximum duration of a Movie released in the year 2021 is 99 min

```
[34]: durr = prime[(prime["release_year"] == 2021) & (prime["type"] == "TV Show")].
      ↪duration.max()
      durr
```

```
[34]: '8 Seasons'
```

The maximum duration of a TV Show released in the year 2021 is 8 seasons

16 Insights plotted from the analysis:

1. Amazon prime has two main categories 1. TV Show 2. Movies

2. Amazon prime has released a total of 7814 Movies and 1854 Tv show
3. Amazon prime has released the highest number of TV Shows and Movies in the year 2021
4. Amazon prime has released 1139 movies, and 303 Tv Shows were released in the year 2021
5. The maximum duration of a Movie/ TV Show in Prime is 99 mts
6. The maximum duration of a movie released in the year 2021 is 99 min
7. The maximum duration of a TV Show released in the year 2021 is eight seasons
8. Mark Knight is the director who has contributed the most to amazon prime
9. United States and India is the best market place for upcoming years
10. United States is ideal for TV Show, and India is ideal for Movies

17 REFERENCES:

Dataset source: <https://www.kaggle.com/datasets/shivamb/amazon-prime-movies-and-tv-shows>

[]:



Assessment Submission Form

Student Number (If this is group work, please include the student numbers of all group participants)	GH1019707
Assessment Title	INDIVIDUAL FINAL PROJECT
Module Code	M504
Module Title	AI AND APPLICATIONS
Module Tutor	PROF MOHAMMAD MAHDAVI
Date Submitted	24 TH MARCH 2022

Declaration of Authorship

I declare that all material in this assessment is my own work except where there is clear acknowledgement and appropriate reference to the work of others.

I fully understand that the unacknowledged inclusion of another person's writings or ideas or works in this work may be considered plagiarism and that, should a formal investigation process confirms the allegation, I would be subject to the penalties associated with plagiarism, as per GISMA Business School, University of Applied Sciences' regulations for academic misconduct.

Signed.....ALINA SHAJI

Date24TH MARCH 2022...