EXPLORATORY DATA ANALYSIS OF MOVIES DATASET

BY- MEGHA GOEL

• Link of Dataset: https://github.com/insaid2018/Term-1/blob/master/Data/Projects/1000%20movies%20data.csv

• Data:

[] data.head()												
F	Rank	Title	Genre	Description	Director	Actors	Year	Runtime (Minutes)	Rating	Votes	Revenue (Millions)	Metascore
0	1	Guardians of the Galaxy	Action,Adventure,Sci-Fi	A group of intergalactic criminals are forced	James Gunn	Chris Pratt, Vin Diesel, Bradley Cooper, Zoe S	2014	121	8.1	757074	333.13	76.0
1	2	Prometheus	Adventure, Mystery, Sci-Fi	Following clues to the origin of mankind, a te	Ridley Scott	Noomi Rapace, Logan Marshall-Green, Michael Fa	2012	124	7.0	485820	126.46	65.0
2	3	Split	Horror,Thriller	Three girls are kidnapped by a man with a diag	M. Night Shyamalan	James McAvoy, Anya Taylor-Joy, Haley Lu Richar	2016	117	7.3	157606	138.12	62.0
3	4	Sing	Animation,Comedy,Family	In a city of humanoid animals, a hustling thea	Christophe Lourdelet	Matthew McConaughey,Reese Witherspoon, Seth Ma	2016	108	7.2	60545	270.32	59.0
4	5	Suicide Squad	Action,Adventure,Fantasy	A secret government agency recruits some of th	David Ayer	Will Smith, Jared Leto, Margot Robbie, Viola D	2016	123	6.2	393727	325.02	40.0
	0 1 2	Rank 0 1 1 2 2 3 3 4	Rank Title O 1 Guardians of the Galaxy 1 2 Prometheus 2 3 Split 3 4 Sing	Rank Title Genre 0 1 Guardians of the Galaxy Action,Adventure,Sci-Fi 1 2 Prometheus Adventure,Mystery,Sci-Fi 2 3 Split Horror,Thriller 3 4 Sing Animation,Comedy,Family	Rank Title Genre Description 0 1 Guardians of the Galaxy Action,Adventure,Sci-Fi A group of intergalactic criminals are forced 1 2 Prometheus Adventure,Mystery,Sci-Fi Following clues to the origin of mankind, a te 2 3 Split Horror,Thriller Three girls are kidnapped by a man with a diag 3 4 Sing Animation,Comedy,Family In a city of humanoid animals, a hustling thea 4 5 Suicide Squad Action Adventure Fantasy A secret government agency	Rank Title Genre Description Director 0 1 Guardians of the Galaxy Action,Adventure,Sci-Fi A group of intergalactic criminals are forced James Gunn 1 2 Prometheus Adventure,Mystery,Sci-Fi Following clues to the origin of mankind, a te Ridley Scott 2 3 Split Horror,Thriller Three girls are kidnapped by a man with a diag M. Night Shyamalan 3 4 Sing Animation,Comedy,Family In a city of humanoid animals, a hustling thea Christophe Lourdelet 4 5 Suicide Squad Action Adventure Fantasy A secret government agency David Aver	Rank Title Genre Description Director Actors 0 1 Guardians of the Galaxy Action,Adventure,Sci-Fi A group of intergalactic criminals are forced James Gunn Chris Pratt, Vin Diesel, Bradley Cooper, Zoe S 1 2 Prometheus Adventure,Mystery,Sci-Fi Following clues to the origin of mankind, a te Ridley Scott Noomi Rapace, Logan Marshall-Green, Michael Fa 2 3 Split Horror,Thriller Three girls are kidnapped by a man with a diag M. Night Shyamalan James McAvoy, Anya Taylor-Joy, Haley Shyamalan 3 4 Sing Animation,Comedy,Family In a city of humanoid animals, a hustling thea Christophe Lourdelet Matthew McConaughey,Reese Witherspoon, Seth Ma 4 5 Suicide Stuad Action, Adventure, Esci-Fi A secret government agency David Aver Will Smith, Jared Leto, Margot Robbie,	Rank Title Genre Description Director Actors Year 1 Guardians of the Galaxy Action, Adventure, Sci-Fi Agroup of intergalactic criminals are forced 2 Prometheus Adventure, Mystery, Sci-Fi Following clues to the origin of mankind, a te 2 Prometheus Adventure, Mystery, Sci-Fi Following clues to the origin of mankind, a te 3 Split Horror, Thriller Three girls are kidnapped by a man with a diag 4 Sing Animation, Comedy, Family In a city of humanoid animals, a hustling thea A secret government agency Pavid Aver Will Smith, Jared Leto, Margot Robbie, 2016	Rank Title Genre Description Director Actors Year Runtime (Minutes) 0 1 Guardians of the Galaxy Action, Adventure, Sci-Fi A group of intergalactic criminals are forced James Gun Chris Pratt, Vin Diesel, Bradley Cooper, Zoe S 2014 121 1 2 Prometheus Adventure, Mystery, Sci-Fi Following clues to the origin of mankind, a te Ridley Scott Noomi Rapace, Logan Marshall-Green, Michael Fa 2012 124 2 3 Split Horror, Thriller Three girls are kidnapped by a man with a diag M. Night Shyamalan James McAvoy, Anya Taylor-Joy, Haley Lu Richar 2016 117 3 4 Sing Animation, Comedy, Family In a city of humanoid animals, a hustling thea Christophe Lourdelet Matthew McConaughey, Reese Witherspoon, Seth Ma 2016 108 4 5 Suicide Squad Action, Adventure, Esci-Fi A secret government agency David Aver Will Smith, Jared Leto, Margot Robbie, 2016 2016 123	Rank Title Genre Description Director Actors Year Runtime (Minutes) Rating 1 Guardians of the Galaxy Action, Adventure, Sci-Fi Agroup of intergalactic criminals are forced James Gun Chris Pratt, Vin Diesel, Bradley Cooper, Zoe S 2014 121 8.1 2 Prometheus Adventure, Mystery, Sci-Fi Following clues to the origin of mankind, a te Ridley Scott Moomi Rapace, Logan Marshall-Green, Michael Fa 2012 124 7.0 3 Split Horror, Thriller Three girls are kidnapped by a man with a diag Shyamalan Shyamalan Christophe Lu Richar 2016 117 7.3 4 Sing Animation, Comedy, Family In a city of humanoid animals, a hustling thea Christophe Lourdelet Witherspoon, Seth Ma 2016 123 6.2	Rank Title Genre Description Director Actors Year Runtime (Minutes) Rating Votes 1 Guardians of the Galaxy Action, Adventure, Sci-Fi Agroup of intergalactic criminals are forced James Gunn Chris Pratt, Vin Diesel, Bradley Cooper, Zoe S 2014 121 8.1 757074 1 2 Prometheus Adventure, Mystery, Sci-Fi Following clues to the origin of mankind, a te Ridley Scott Moomi Rapace, Logan Marshall-Green, Michael Fa 2012 124 7.0 485820 2 3 Split Horror, Thriller Three girls are kidnapped by a man with a diag M. Night Shyamalan Shyamalan Christophe Lu Richar 2016 117 7.3 157606 3 4 Sing Animation, Comedy, Family In a city of humanoid animals, a hustling thea Christophe Lourdelet Witherspoon, Seth Ma Will Smith, Jared Leto, Margot Robbie, 2016 123 6.2 393727	Rank Title Genre Description Director Actors Vear Runtime (Minutes) Rating Votes (Millions) 1 Guardians of the Galaxy Action, Adventure, Sci-Fi Agroup of intergalactic criminals are forced James Gun Galaxy Prometheus Adventure, Mystery, Sci-Fi Following clues to the origin of mankind, a te Ridley Scott Moomi Rapace, Logan Marshall-Green, Michael Fa 2012 124 7.0 485820 126.46 2 3 Split Horror, Thriller Three girls are kidnapped by a man with a diag Shyamalan Shyamalan Christophe Lu Richar 2016 117 7.3 157606 138.12 3 4 Sing Animation, Comedy, Family Asserted government agency Pavid Aver Will Smith, Jared Leto, Margot Robbie, 2016 123 6.2 393727 325.02

• <u>Data Pre- Processing:</u>

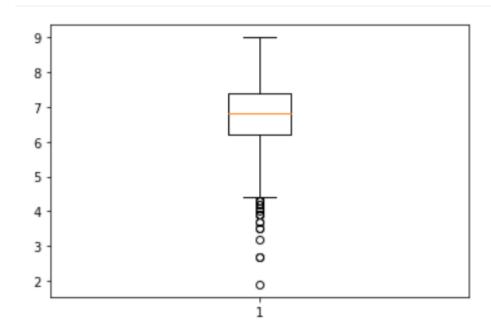
```
[46] data.isnull().sum()
 С⇒
     Rank
                               0
     Title
                               0
     Genre
                               0
     Description
                               0
     Director
                               0
     Actors
                               0
     Year
                               0
     Runtime (Minutes)
                               0
     Rating
                               0
     Votes
                               0
     Revenue (Millions)
                             128
     Metascore
                              64
     dtype: int64
[44] data.nunique()
 C→ Rank
                             1000
     Title
                              999
     Genre
                              207
     Description
                             1000
     Director
                              644
     Actors
                              996
     Year
                               11
     Runtime (Minutes)
                               94
     Rating
                               55
     Votes
                              997
     Revenue (Millions)
                              814
     Metascore
                               84
     dtype: int64
```

```
data = data.drop(['Metascore'], axis=1)
.
[48] movie.head()
```

[49] data["Genre"].value_counts()

₽	Action,Adventure,Sci-Fi	50	
_	Drama	48	
	Comedy,Drama,Romance	35	
	Comedy	32	
	Drama, Romance	31	
	Comedy,Western	1	
	Adventure, Comedy, Horror	1	
	Drama,Horror,Musical	1	
	Action,Crime,Sport	1	
	Comedy, Fantasy, Romance	1	
	Name: Genre, Length: 207	, dtype:	int64

• For identification of Outliers:



• Grouping Viewer Votes on Basis of Genre:

data.groupby(['Genre','Votes']).mean()

		Rank	Year	Runtime (Minutes)	Rating	Revenue (Millions)	Metascore
Genre	Votes						
Action	6809	581	2016	90	4.9	131.56	37.0
	10428	124	2016	86	7.4	NaN	NaN
Action,Adventure	122838	285	2014	98	6.0	72.66	47.0
	406219	345	2007	139	6.2	336.53	59.0
	474320	369	2012	136	7.0	262.03	66.0
Thriller	4204	912	2016	87	4.7	NaN	68.0
	4531	445	2016	81	6.0	NaN	31.0
	71069	730	2014	100	6.7	0.32	76.0
	220236	283	2007	113	7.1	NaN	NaN
Thriller,War	5926	161	2016	106	6.0	NaN	40.0

1000 rows × 6 columns

• Average Runtime of Movies:

```
[ ] # Average runtime of movies data['Runtime (Minutes)'].mean()
```

☐→ 113.172

• Average Revenue Of Movies:

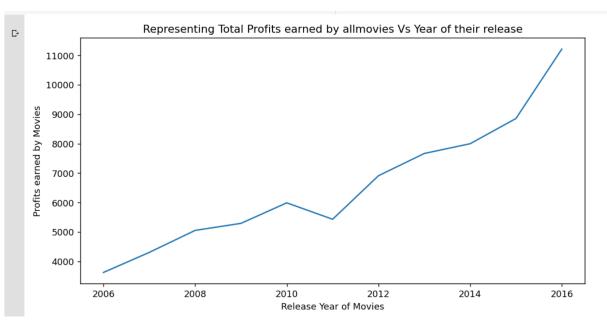
```
[ ] # Average revenue of movies
    data['Revenue (Millions)'].mean()
```

E 82.95637614678897

• Movies Rated above 7:

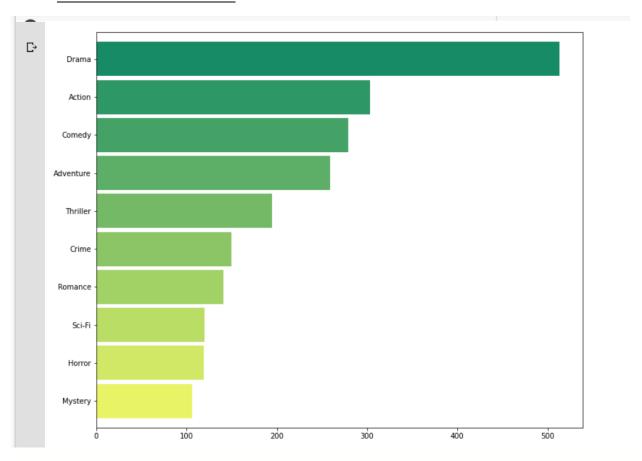
	#Movies rated above 7 data[data['Rating']>=7.0]											
	Rank	Title	Genre	Description	Director	Actors	Year	Runtime (Minutes)	Rating	Votes	Revenue (Millions)	Metascore
0	1	Guardians of the Galaxy	Action,Adventure,Sci-Fi	A group of intergalactic criminals are forced	James Gunn	Chris Pratt, Vin Diesel, Bradley Cooper, Zoe S	2014	121	8.1	757074	333.13	76.0
1	2	Prometheus	Adventure, Mystery, Sci-Fi	Following clues to the origin of mankind, a te	Ridley Scott	Noomi Rapace, Logan Marshall- Green, Michael Fa	2012	124	7.0	485820	126.46	65.0
2	3	Split	Horror,Thriller	Three girls are kidnapped by a man with a diag	M. Night Shyamalan	James McAvoy, Anya Taylor-Joy, Haley Lu Richar	2016	117	7.3	157606	138.12	62.0
3	4	Sing	Animation,Comedy,Family	In a city of humanoid animals, a hustling thea	Christophe Lourdelet	Matthew McConaughey,Reese Witherspoon, Seth Ma	2016	108	7.2	60545	270.32	59.0
6	7	La La Land	Comedy,Drama,Music	A jazz pianist falls for an aspiring actress i	Damien Chazelle	Ryan Gosling, Emma Stone, Rosemarie DeWitt. J	2016	128	8.3	258682	151.06	93.0

• Which years we had the most profitable movies?:



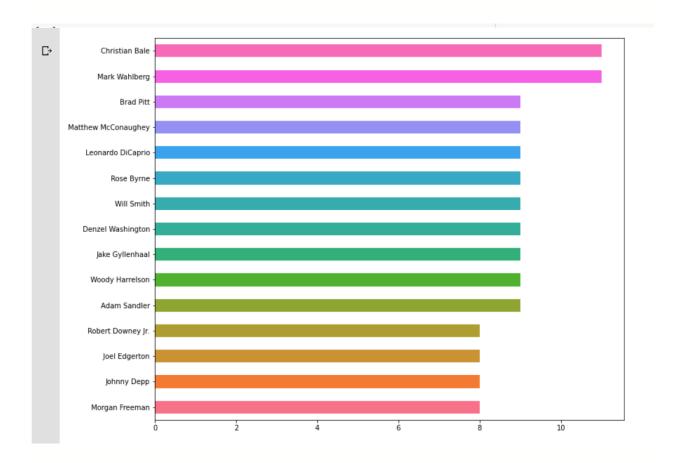
♣ So, the most profitable year of the given dataset was 2016

• Most Successful Genre:



→ From the bar plot above we can see that Drama is the most common genre among all the movies. This tells a bit about which genre to go for if you are trying to invest some money in the film industry.

• Frequent Cast in Movies:



♣ Looking at the bar plot, Christian Bale is behind most movies' success

• Correlation between no. of Votes a movie has gained and it's Revenue:

[64] #Using the .corr method from pandas, default is Pearson Corellation
 data[['Votes', 'Revenue (Millions)']].corr()

Votes Revenue (Millions)

Votes	1.000000	0.639661
Revenue (Millions)	0.639661	1.000000

♣ Shows positive correlation between Votes and Revenue

• CONCLUSION:

After this analysis we can conclude following:

For a Movie to be profitable for the producer:

- ♣ Average duration of the movie must be 114 minutes
- ♣ Any one of these should be in the cast: Christian Bale,
- ♣ Mark Wahlberg, Adam Sandler
- ♣ Genre must be: Drama, Action, Comedy, Adventure.

By doing all this the movie might be one of the hits and hence can earn an average revenue of around 83 million dollars.