

WEEKLY PROJECT 20 :

IMDB - MOVIE DATASET ANALYSIS :

In [4]:

```
import pandas as pd
import numpy as np
```

In [5]:

```
df = pd.read_csv("IMDB-Movie-Data - IMDB-Movie-Data.csv")
```

In [6]:

```
df.head()
```

Out[6]:

| | 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 |

1.Top 5 directors having ratings more than average?

In [7]:

```
df.isnull().sum()
```

Out[7]:

```
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
```

In [15]:

```
len_ = len(df["Rating"])
sum_ = sum(df["Rating"])
avg_ = (sum_/len_)
print("The Average rating is","-", "{0:.2f}".format(avg_))
```

The Average rating is - 6.72

In [29]:

```
len_ = len(df["Title"])
rating_ = []
director_ = []
for i in range(len_):
    if(df["Rating"][i]>avg_):
        rating_.append(df["Rating"][i])
        director_.append(df["Director"][i])
```

In [30]:

```
print(rating_)
```

[8.1, 7.0, 7.3, 7.2, 8.3, 7.1, 7.0, 7.5, 7.8, 7.9, 7.7, 8.2, 8.1, 8.0, 7.9, 6.8, 8.3, 7.5, 7.1, 8.0, 7.9, 8.6, 7.6, 6.9, 7.1, 7.5, 7.2, 7.3, 7.1, 8.1, 7.2, 9.0, 7.3, 7.2, 7.4, 7.0, 7.5, 6.8, 8.5, 7.7, 7.4, 8.1, 7.5, 7.2, 7.1, 7.5, 6.8, 8.1, 7.1, 8.1, 8.3, 7.3, 8.8, 7.9, 8.2, 8.1, 7.2, 7.0, 7.8, 7.8, 7.4, 8.1, 7.0, 8.1, 7.1, 7.4, 7.4, 8.6, 8.5, 7.0, 7.0, 8.0, 7.9, 7.3, 7.7, 7.7, 8.1, 7.7, 8.1, 7.8, 7.2, 7.4, 7.4, 8.5, 7.5, 8.0, 7.3, 8.5, 7.1, 7.3, 8.1, 7.3, 8.1, 7.1, 8.0, 7.8, 8.2, 8.4, 8.1, 7.4, 7.6, 7.6, 7.2, 7.6, 8.1, 7.0, 7.4, 7.5, 7.9, 7.0, 8.0, 8.0, 7.3, 7.3, 6.8, 7.9, 7.9, 8.0, 7.5, 7.6, 7.0, 7.4, 7.3, 6.8, 7.0, 8.0, 7.8, 7.2, 8.1, 8.2, 6.9, 8.1, 8.1, 7.9, 7.3, 7.9, 7.8, 7.5, 6.8, 7.1, 7.0, 7.8, 7.7, 7.1, 6.9, 7.8, 7.0, 7.0, 7.1, 7.0, 8.2, 7.8, 7.4, 8.0, 6.8, 8.1, 7.6, 8.2, 7.6, 7.9, 7.4, 7.1, 8.6, 7.3, 7.9, 7.5, 7.8, 7.6, 7.1, 7.8, 8.0, 7.0, 7.5, 7.8, 7.9, 7.2, 8.0, 6.8, 7.0, 7.1, 7.7, 7.0, 7.2, 7.3, 7.6, 7.1, 7.0, 7.5, 7.2, 7.7, 7.1, 6.8, 7.1, 8.1, 7.2, 7.5, 7.0, 7.7, 7.8, 6.8, 7.4, 7.5, 7.0, 7.3, 7.3, 7.7, 7.4, 7.0, 6.9, 7.3, 8.0, 7.4, 8.1, 7.8, 7.8, 7.4, 6.8, 7.7, 7.3, 7.7, 7.3, 7.1, 7.4, 7.7, 7.8, 6.9, 7.8, 7.7, 8.0, 6.9, 7.0, 6.8, 6.9, 6.9, 7.6, 7.1, 7.6, 7.4, 7.6, 7.8, 7.1, 7.2, 7.2, 6.8, 7.1, 7.7, 7.4, 7.8, 7.8, 7.3, 7.2, 6.9, 7.8, 7.7, 7.2, 6.8, 7.8, 8.1, 8.4, 7.0, 7.4, 7.1, 7.6, 7.5, 7.7, 7.4, 7.1, 7.9, 7.8, 7.0, 7.0, 6.8, 8.3, 7.3, 7.6, 7.5, 7.5, 6.9, 8.0, 7.8, 7.5, 7.7, 7.4, 8.5, 8.3, 7.2, 7.7, 7.8, 8.0, 6.9, 7.0, 7.2, 6.8, 7.1, 7.4, 8.3, 7.2, 7.3, 7.9, 7.7, 7.8, 7.8, 7.2, 7.1, 7.1, 6.9, 7.9, 7.2, 7.0, 6.9, 7.0, 7.2, 7.0, 7.3, 7.2, 7.4, 7.1, 6.8, 6.8, 7.9, 7.1, 7.1, 7.1, 6.9, 6.8, 7.8, 7.5, 7.1, 7.1, 7.8, 6.8, 6.8, 6.8, 7.6, 7.6, 7.2, 7.3, 7.2, 7.8, 7.1, 7.2, 7.9, 7.3, 7.2, 7.4, 7.2, 6.9, 8.1, 7.0, 7.6, 7.5, 6.9, 7.3, 7.3, 7.0, 8.4, 7.0, 7.5, 7.3, 8.0, 8.1, 7.5, 8.2, 7.4, 6.8, 7.0, 7.9, 7.7, 7.0, 6.9, 7.7, 7.0, 6.8, 8.3, 7.6, 7.2, 7.2, 7.5, 7.7, 8.0, 7.6, 8.3, 7.1, 7.8, 7.9, 8.1, 7.3, 7.0, 2, 7.0, 7.1, 7.1, 7.4, 7.6, 7.3, 8.2, 7.1, 7.6, 6.8, 7.0, 6.9, 7.7, 7.5, 7.1, 8.2, 7.6, 7.2, 7.9, 7.3, 7.8, 7.5, 7.3, 7.3, 8.2, 8.1, 7.3, 7.2, 6.9, 7.3, 6.9, 7.8, 6.8, 7.8, 7.3, 7.4, 7.0, 7.2, 6.8, 7.1, 6.8, 7.0, 7.8, 7.6, 7.6, 6.9, 6.8, 7.0, 7.6, 7.2, 7.4, 7.5, 7.8, 7.0, 7.6, 6.9, 7.5, 6.9, 7.3, 6.9, 7.7, 7.0, 7.3, 6.9, 7.7, 7.0, 7.3, 6.9, 7.4, 6.8, 7.1, 7.6, 7.7, 7.1, 6.8, 6.8, 6.8, 7.3, 7.4, 6.9, 7.2, 7.8, 7.3, 7.2, 7.7, 7.8, 7.0, 6.9, 7.0, 7.3, 6.9, 7.4, 6.8, 7.1, 7.5, 8.5]

In [31]:

```
print(director_)
```

['James Gunn', 'Ridley Scott', 'M. Night Shyamalan', 'Christophe Lourdelet', 'Damien Chazelle', 'James Gray', 'Morten Tyldum', 'David Yates', 'Theodore Melfi', 'Gareth Edwards', 'Ron Clements', 'Mel Gibson', 'Garth Davis', 'Denis Villeneuve', 'Kenneth Lonergan', 'Dominique Abel', 'S.S. Rajamouli', 'Tom Ford', 'Bryan Singer', 'Tim Miller', 'Anthony Russo', 'Christopher Nolan', 'Scott Derrickson', 'Antoine Fuqua', 'Patrick Read Johnson', 'Barry Jenkins', 'John Lee Hancock', 'John Madden', 'Justin Lin', 'J.J. Abrams', 'Chad Stahelski', 'Christopher Nolan', 'Martin Scorsese', 'Fede Alvarez', 'Thea Sharrock', 'Lone Scherfig', 'Clint Eastwood', 'André Øvredal', 'Christopher Nolan', 'Matthew Vaughn', 'Peter Berg', 'George Miller', 'Robin Swicord', 'Peter Berg', 'Robert Zemeckis', 'J.A. Bayona', 'David Frankel', 'Byron Howard', 'Gore Verbinski', 'Joss Whedon', 'Quentin Tarantino', 'Gore Verbinski', 'Christopher Nolan', 'Matt Ross', 'Martin Scorsese', 'David Fincher', 'James Wan', 'Colin Trevorrow', 'James Cameron', 'Quentin Tarantino', 'Gavin O'Connor', 'Denis Villeneuve', 'Duncan Jones', 'Tate Taylor', 'Todd Phillips', 'Joss Whedon', 'Shane Black', 'Makoto Shinkai', 'Martin Scorsese', 'Brian Helgeland', 'Kenneth Branagh', 'Ridley Scott', 'Oriol Paulo', 'Guy Ritchie', 'David Mackenzie', 'Gary Garland', 'Steve McQueen', 'Zack Snyder', 'David Yates', 'Nitesh Tiwari', 'Dan Trachtenberg', 'Andrew Stanton', 'Todor Chapkanov', 'Christopher Nolan', 'Jon Favreau', 'Alejandro González Iñárritu', 'Denzel Washington', 'Damien Chazelle', 'Justin Lin', 'Derek Cianfrance', 'Ethan Coen', 'Baz Luhrmann', 'Martin Scorsese', 'Martin Scorsese', 'J.J. Abrams', 'Adam McKay', 'Lenny Abrahamson', 'Quentin Tarantino', 'Chan-wook Park', 'Kelly Fremon Craig', 'Zack Snyder', 'Greg Mottola', 'Gary Ross', 'Denis Villeneuve', 'David Lynch', 'Guillermo del Toro', 'Glenn Ficarra', 'Edgar Wright', 'Edgar Wright', 'Ben Wheatley', 'Bryan Singer', 'Martin Campbell', 'Simon Curtis', 'Michael Bay', 'Sam Mendes', 'Dan Gilroy', 'Travis Knight', 'Spike Jonze', 'Chris Buck', 'Matt Reeves', 'Ben Stiller', 'James Wan', 'Peyton Reed', 'Robert Eggers', 'Kenneth Branagh', 'Asghar Farhadi', 'Nicolas Winding Refn', 'Derek Cianfrance', 'Tom McCarthy', 'Gavin O'Connor', 'Joe Johnston', 'Sean Penn', 'Morten Tyldum', 'Doug Liman', 'Oliver Stone', 'Jon Favreau', 'Matthew Vaughn', 'Julia Ducournau', 'Jim Jarmusch', 'Paul Feig', 'Michael Bay', 'Alan Taylor', 'Anthony Russo', 'Matthew Vaughn', 'John Lasseter', 'David Robert Mitchell', 'David Fincher', 'Rawson Marshall Thurber', 'Andrea Arnold', 'Yorgos Lanthimos', 'Robert Stromberg', 'Guillermo del Toro', 'Mel Gibson', 'Christopher McQuarrie', 'Stephen Chbosky', 'Pablo Larraín', 'Wes Anderson', 'David Ayer', 'Pete Docter', 'Tom Hooper', 'Alfonso Cuarón', 'Mike Mills', 'Paul Feig', 'Olivier Nakache', 'Yorgos Lanthimos', 'F. Gary Gray', 'James Wan', 'Todd Phillips', 'Rupert Wyatt', 'Joseph Cedar', 'Alejandro González Iñárritu', 'Darren Aronofsky', 'Lars von Trier', 'Tom Tykwer', 'Sam Mendes', 'Peter Jackson', 'Phil Lord', 'John Carney', 'Eric Summer', 'Joseph Kosinski', 'Phil Lord', 'David Fincher', 'Richard Linklater', 'Shane Black', 'Louis Leterrier', 'Guy Ritchie', 'Quentin Tarantino', 'Tom Hooper', 'Sang-ho Yeon', 'Jason Moore', 'David Leveaux', 'Zack Snyder', 'David Frankel', 'Brad Furman', 'Paul Thomas Anderson', 'Antoine Fuqua', 'Peter Berg', 'Drew Goddard', 'John Lee Hancock', 'Abdellatif Kechiche', 'Jennifer Kent', 'Peter Jackson', 'David Yates', 'Bong Joon Ho', 'Justin Lin', 'Guy Ritchie', 'David Fincher', 'Roger Spottiswoode', 'Jeremy Saulnier', 'Nicholas Stoller', 'Clint Eastwood', 'Jean-Marc Vallée', 'Steven Spielberg', 'Ron Howard', 'David O. Russell', 'Josh Boone', 'Rian Johnson', 'David Lowery', 'Matthew Vaughn', 'David O. Russell', 'Woody Allen', 'William Oldroyd', 'Jocelyn Moorhouse', 'Hugo Gelin', 'James Marsh', 'Richard Curtis', 'Adam McKay', 'J.J. Abrams', 'Ruben Fleischer', 'Danny Boyle', 'Alex Ranarivelo', 'Marc Webb', 'Wes Ball', 'Stephen Frears', 'Scott Cooper', 'Ryan Coogler', 'Dan Kwan', 'Jemaine Clement', 'Antoine Fuqua', 'Steven Spielberg', 'Phil Lord', 'Baltasar Kormákur', 'Danny Boyle', 'Lee Toland Krieger', 'Louis Leterrier', 'Nicolas Winding Refn', 'Pierre Coffin', 'Kathryn Bigelow', 'Nathan Greno', 'Pierre Morel', 'Dan Scanlon', 'Paul Verhoeven', 'Mark Williams', 'Joe Wright', 'David Yates', 'Steve McQueen', 'Joe Wright', 'David Fincher', 'Paul Greengrass', 'Rajkumar Hirani', 'Marc Forster', 'Brad Bird', 'Jon S. Baird', 'George Tillman Jr.', 'J.A. Bayona', 'John Crowley', 'Henry Selick', 'Derek Cianfrance', 'Pete Travis', 'Taika Waititi', 'Don Hall', 'Jon Favreau', 'Jean-Marc Vallée', 'Christian Alvart', 'Thomas Vinterberg', 'Sam Mendes', 'Ben Affleck', 'Michael Spierig', 'Guy Ritchie', 'Denis Villeneuve', 'Neill Blomkamp', 'Ridley Scott', 'David Yates', 'Ben Affleck', 'Dexter Fletcher', 'Florian Henckel von Donnersmarck', 'Amber Tamblyn', 'Nicholas Stoller', 'Paolo Genovese', 'Wes Anderson', 'Brad Bird', 'Albert Hughes', 'Matt Reeves', 'Francis Lawrence', 'Barry Sonnenfeld', 'J.J. Abrams', 'F. Gary Gray', 'Pete Docter', 'Todd Haynes', 'Paolo Sorrentino', 'Jaco Van Dormael', 'Marc Webb', 'Alfonso Cuarón', 'Mark Herman', 'Antoine Fuqua', 'S. Craig Zahler', 'Will Gluck', 'Neill Blomkamp', 'Peter Jackson', 'Sean Ellis', 'David Schwimmer', 'Ridley Scott', 'Tony Scott', 'Terence Davies', 'Christopher McQuarrie', 'Alexandros Avranas', 'Roman Polanski', 'Neil Burger', 'Richard LaGravenese', 'Mike Birbiglia', 'Philippe Falardeau', 'Ken Loach', 'Jeff Nichols', 'Florian Gallenberger', 'Joel Edgerton', 'Ron Howard', 'Scott Derrickson', 'Rich Moore', 'Jason Reitman', 'Kevin Lima', 'Nancy Meyers', 'Jonathan Dayton', 'Ben Younger', 'Craig Gillespie', 'Joseph Kosinski', 'Francis Lawrence', 'Bennett Miller', 'Angelina Jolie', 'Danny Boyle', 'Mark Andrews', 'Niels Arden Oplev', 'Michaël R. Roskam', 'Matt Reeves', 'Duncan Jones', 'Tom McGrath', 'Alessandro Carboni', 'John Butler', 'Debra Granik', 'Seth Gordon', 'Xavier Dolan', 'Guillermo del Toro', 'Maren Ade', 'Andrew Dominik', 'Gary Ross', 'Ben Stiller', 'Rick Famuyiwa', 'Bong Joon Ho', 'Andrew Stanton', 'Judd Apatow', 'Duncan Jones', 'John Hillcoat', 'Gabriele Muccino', 'Damián Szifron', 'François Ozon', 'Clint Eastwood', 'Tim Burton', 'James Wan', 'Michael Mann', 'Richard Linklater', 'David Cronenberg', 'David Gordon Green', 'Jon Kasdan', 'Ben Affleck', 'Mia Hansen-Løve', 'Edward Zwick', 'Kathryn Bigelow', 'Frank Darabont', 'Christian Ditter', 'Jee-woon Kim', 'Lynne Ramsay', 'Gabriele Muccino', 'Tom Hooper', 'Steve McQueen', 'Lee Unkrich', 'Tomas Alfredson', 'Gauri Shinde', 'Ang Lee', 'Lasse Hallström', 'Mike Cahill', 'Len Wiseman', 'Seth MacFarlane', 'Robert Schwentke', 'George Nolfi', 'David Dobkin', 'Michael Dudok de Wit', 'Jorge R. Gutiérrez', 'Denis Villeneuve', 'Jean-Marc Vallée', 'Neil Burger', 'Joe Carnahan', 'Greg Mottola', 'Simon Aboud', 'Bruce Beresford', 'Martin Scorsese', 'Ridley Scott', 'Juan José Campanella', 'Ethan Coen', 'Gregory Hoblit', 'Martin McDonagh', 'John Hillcoat', 'David O. Russell', 'Kirsten Sheridan', 'Jon Favreau', 'Gavin Hood', 'Rajkumar Hirani', 'Dean DeBlois', 'Darren Aronofsky', 'Matt Brown', 'Jaume Collet-Serra', 'Bruce A. Evans', 'Chris Evans', 'Paul Greengrass', 'Jim Sheridan', 'Claude Barras', 'Werner Herzog', 'Pierre Coffin', 'Edgar Wright', 'Martin McDonagh', 'Mikael Häfström', 'James Ponsoldt', 'Amma Asante', 'Jaume Collet-Serra', 'Wes Anderson', 'Spike Lee', 'Danny Boyle', 'Saul Dibb', 'Terry Gilliam', 'Antonio Campos', 'Robert Rodriguez', 'Sylvester Stallone', 'Cary Joji Fukunaga', 'Christian Mungiu', 'Aisling Walsh', 'Lars von Trier', 'Ron Clements', 'Tony Scott', 'Robert Zemeckis', 'Giuseppe Tornatore', 'Brian Helgeland', 'Robert Luketic', 'John Carney', 'Scott Cooper', 'Woody Allen', 'Mark Osborne', 'Richard LaGravenese', 'Clint Eastwood', 'Allen Coulter', 'Naoko Yamada', 'Steven Knight', 'James Schamus', 'Kyle Patrick Alvarez', 'J.J. Abrams', 'Hannes Holm', 'Lone Scherfig', 'Juan Carlos Fresnadillo', 'Jonathan Levine', 'Woody Allen', 'Hideaki Anno', 'John Lee Hancock', 'Carlos Saldanha', 'Alejandro González Iñárritu', 'Terrence Malick', 'Jonathan Levine', 'Shawn Levy', 'Sylvester Stallone', 'Robert Rodriguez', 'Peter Landessman', 'Tarsem Singh', 'Michael Dowse', 'Dean DeBlois', 'Steven Soderbergh', 'Gabor Csupo', 'James Ward Byrkit', 'Hong-jin Na', 'Tom Tykwer', 'John Erick Dowdle', 'Lee Daniels', 'Paul Thomas Anderson', 'R.J. Cutler', 'James Gunn', 'Richie Smyth', 'Jason Reitman', 'Christopher Smith', 'Lee Daniels', 'Alfonso Gomez-Rejon', 'Alexander Payne', 'Jordan Vogt-Roberts', 'David Ayer', 'Paul McGuigan', 'Danny Boyle', 'James Lapine', 'David Lynch', 'Robert Zemeckis', 'D.J. Caruso', 'Mira Nair', 'Xavier Dolan', 'Pedro Almodóvar', 'Patricia Riggen', 'Julie Taymor', 'Boaz Yakin', 'Pascal Laugier', 'Ava DuVernay', 'Aamir Khan']

In [46]:

```
df_ = pd.DataFrame(rating_, columns = ["rating"])
df_["director"] = director_
a = df_.sort_values(by = 'rating', ascending = False).head(5)["rating"].values
b = df_.sort_values(by = 'rating', ascending = False).head(5)["director"].values
df1 = pd.DataFrame(a, columns = ["RATING"])
df1["DIRECTOR"] = b
(df1.head())
```

Out[46]:

| | RATING | DIRECTOR |
|---|--------|-------------------|
| 0 | 9.0 | Christopher Nolan |
| 1 | 8.8 | Christopher Nolan |
| 2 | 8.8 | Nitesh Tiwari |
| 3 | 8.6 | Makoto Shinkai |
| 4 | 8.6 | Olivier Nakache |