

WEEKLY PROJECT 20

IMDB Movie Dataset

Importing libraries and dataset

```
In [31]: import numpy as np
import pandas as pd
```

```
In [32]: df=pd.read_csv("IMDB-Movie-Data.csv")
df.head()
```

Out[32]:

	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

Top 5 directors having rating more than average

```
In [33]: l=len(df['Rank'])
c=0
for i in range(l):
    c+=df['Rating'][i]
a=c/l
avg=round(a,1)
r=[]
n=[]
for i in range(l):
    if(df['Rating'][i]>avg):
        r.append(df['Rating'][i])
        n.append(df['Director'][i])
df_=pd.DataFrame(r,columns=['Rate'])
df_['Direct']=n
x=df_.sort_values(by='Rate',ascending=False).head(5)['Rate'].values
y=df_.sort_values(by='Rate',ascending=False).head(5)['Direct'].values
df1=pd.DataFrame(x,columns=['Rating'])
df1['Director']=y
df1.head()
```

Out[33]:

	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

```
In [ ]:
```