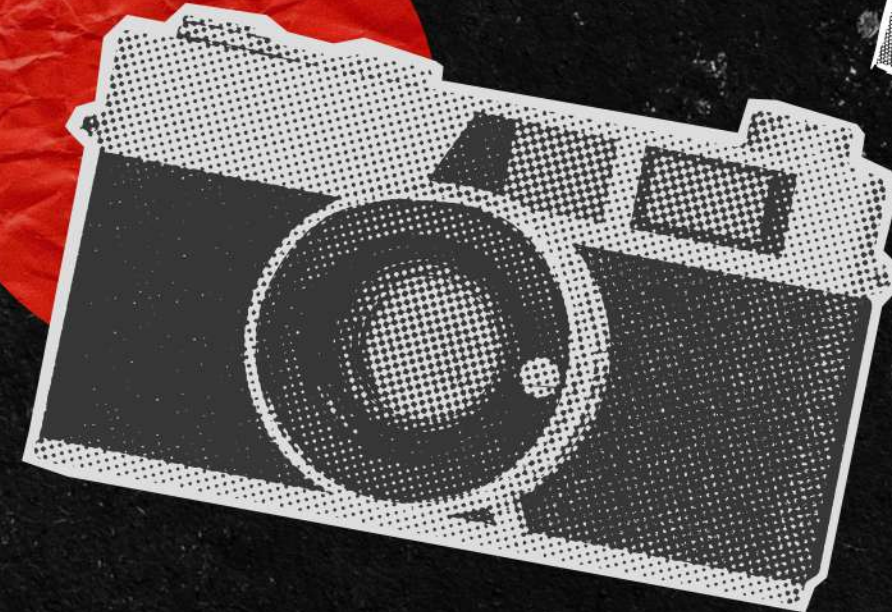


TOP DIRECTOR



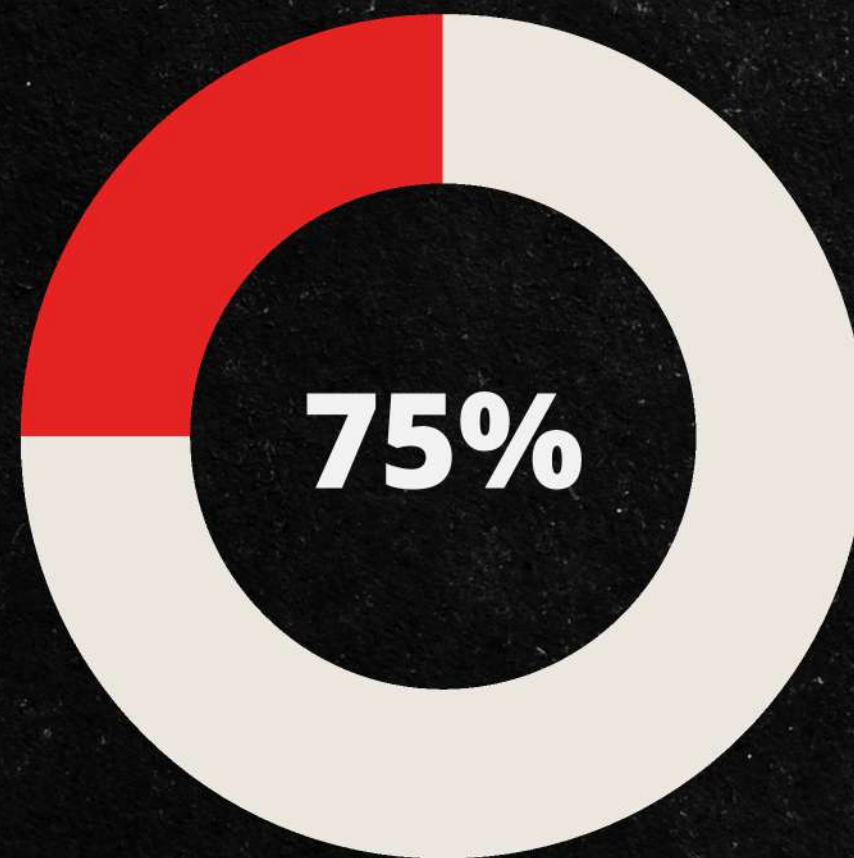
IMDB MOVIES INDIA

HW6

Punramon Ratchakom

653020213-2

เกณฑ์การพิจารณา



- *Rating*
- *Votes*

หมายเหตุ

คำนึงถึงจำนวนหนังของผู้กำกับ

HOW?

- นำเข้าข้อมูล
- แปลง datatype ของ Year, Duration, Rating, Votes เป็นตัวเลข
- ทำความสะอาดข้อมูล (column ที่สนใจ)

```
movie_df = movie_df.dropna(subset=['Rating', 'Votes'])
```
- หาผู้กำกับ(Director)ชั้นนำของอินเดีย 3 คน

- ดูรายละเอียดของข้อมูลที่เป็นตัวเลข

```
1 movie_df.describe()
```

	Year	Duration	Rating	Votes
count	5659.000000	5659.000000	5659.000000	5659.000000
mean	1996.247570	133.439124	5.898533	2697.649585
std	19.741839	25.319939	1.381165	13651.503584
min	1931.000000	21.000000	1.100000	5.000000
25%	1983.000000	119.000000	5.000000	30.000000
50%	2002.000000	135.000000	6.100000	131.000000
75%	2013.000000	150.000000	6.900000	922.500000
max	2021.000000	321.000000	10.000000	591417.000000

prompt: หาผู้กำกับ(Director)ชั้นนำของอินเดีย 3 คน ว่าทำไมเราถึงคิดว่าเป็นผู้กำกับชั้นนำ
#(อาจจะดูจาก rating ที่มากกว่า 75% , Votes ที่มากกว่า 75%) พร้อมแสดงตัวอย่างหนัง 3 เรื่องที่ rating และ votes มากสุด ของผู้กำกับแต่ละคน

ทำความสะอาดข้อมูล

```
movie_df = movie_df.dropna(subset=['Rating', 'Votes']) # ลบแถวที่มี missing values
```

คำนวณค่า quantile ที่ 75% สำหรับ Rating และ Votes

```
rating_q3 = movie_df['Rating'].quantile(0.75) #6.9
```

```
tes_q3 = movie_df['Votes'].quantile(0.75) #922.5
```

เลือกผู้กำกับที่มี Rating และ Votes มากกว่า 75%

```
top_directors = movie_df[(movie_df['Rating'] >= rating_q3) & (movie_df['Votes'] >= tes_q3)]['Director'].value_counts().head(3)
```

แสดงผล

```
print("Top 3 directors in India:")
```

```
for director, count in top_directors.items():
```

```
    print(f"- {director} ({count} movies)")
```

แสดงข้อมูลเพิ่มเติมเกี่ยวกับผู้กำกับแต่ละคน

```
director_movies = movie_df[movie_df['Director'] == director]
```

```
print(f" - Average rating: {director_movies['Rating'].mean():.2f}")
```

```
print(f" - Average votes: {director_movies['Votes'].mean():.0f}")
```

```
print(f" - Example movies:")
```

```
director_movies = movie_df[movie_df['Director'] == director].sort_values(by=['Rating', 'Votes'], ascending=[False, False]).head(3)
```

```
for _, movie in director_movies.head(3).iterrows():
```

```
    print(f" - {movie['Name']} ({movie['Year']}) - Rating: {movie['Rating']} - Votes: {movie['Votes']} ")
```

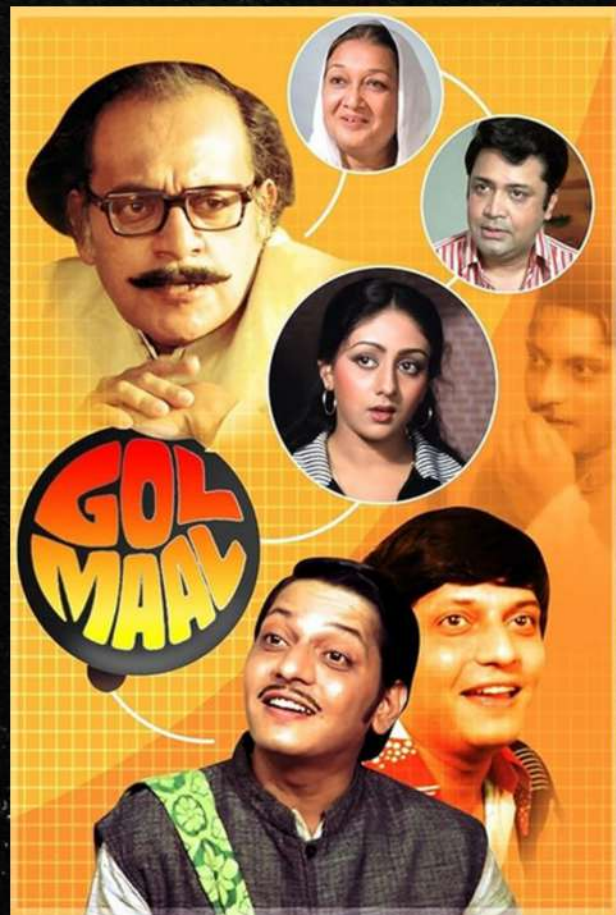
04

TOP 3 Directors in INDIA

Hrishikesh Mukherjee (10 movies)

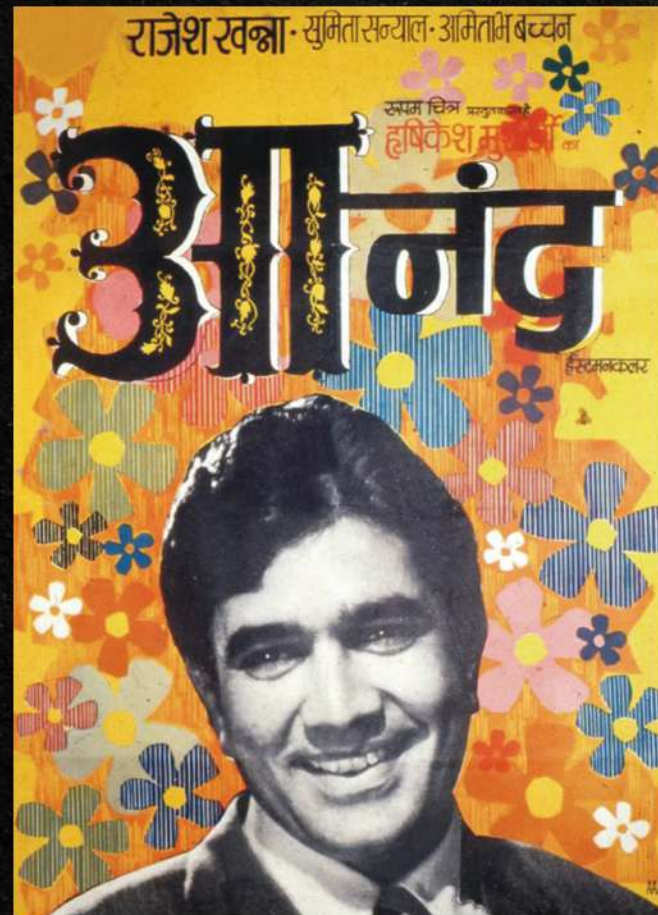
- Average rating: 7.27
- Average votes: 2,464

Example movies:



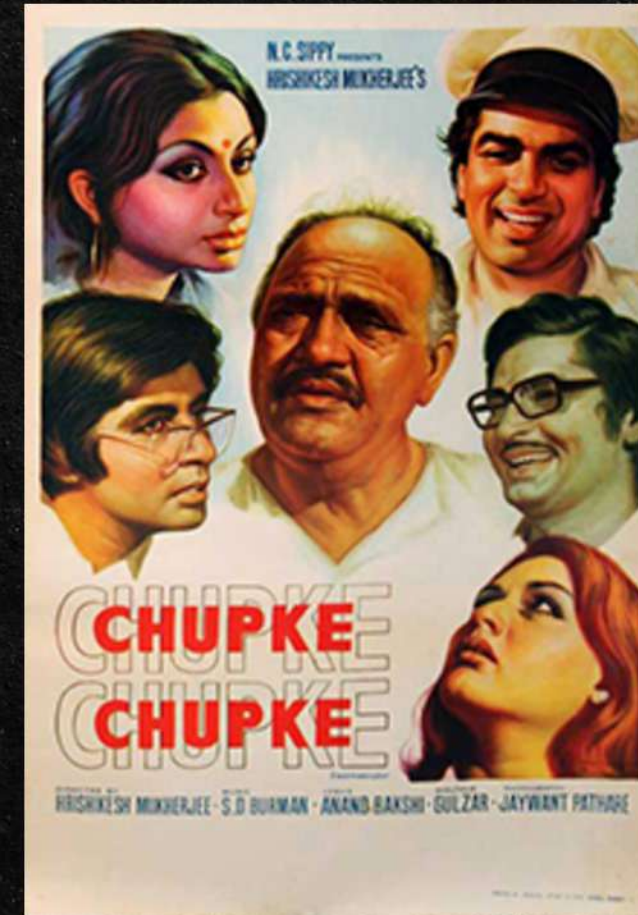
Hanky Panky (1979)

- Rating: 8.5
- Votes: 18376



Anand (1971)

- Rating: 8.3
- Votes: 31937



Chupke Chupke (1975)

- Rating: 8.3
- Votes: 11923

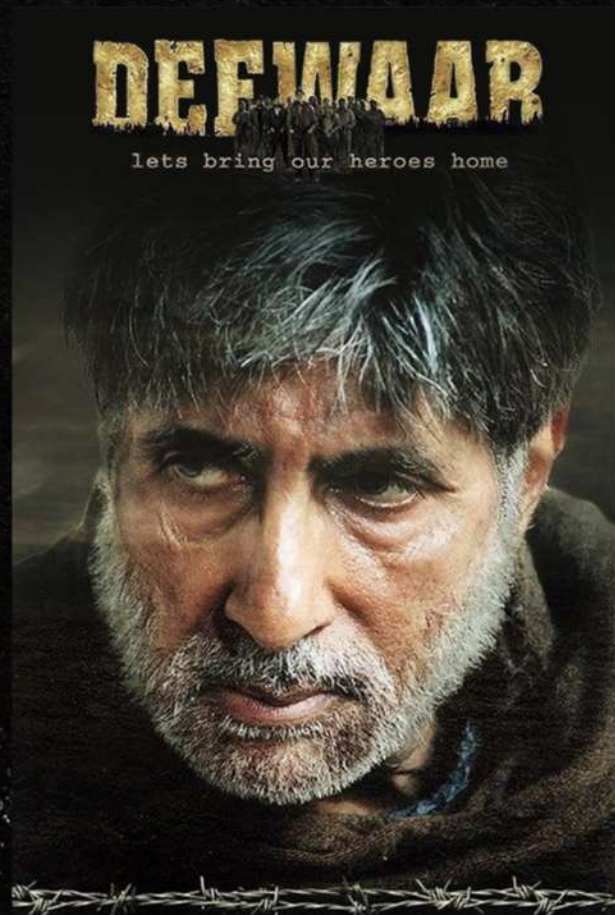


TOP 3 Directors in INDIA

Yash Chopra (11 movies)

- Average rating: 6.97
- Average votes: 7,948

Example movies:



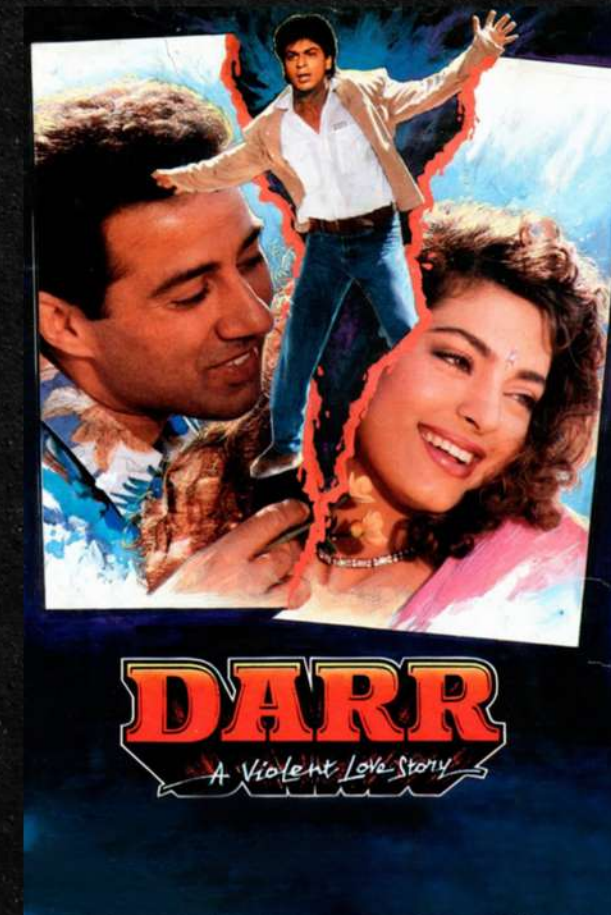
Deewaar (1975)

- Rating: 8.1
- Votes: 10361



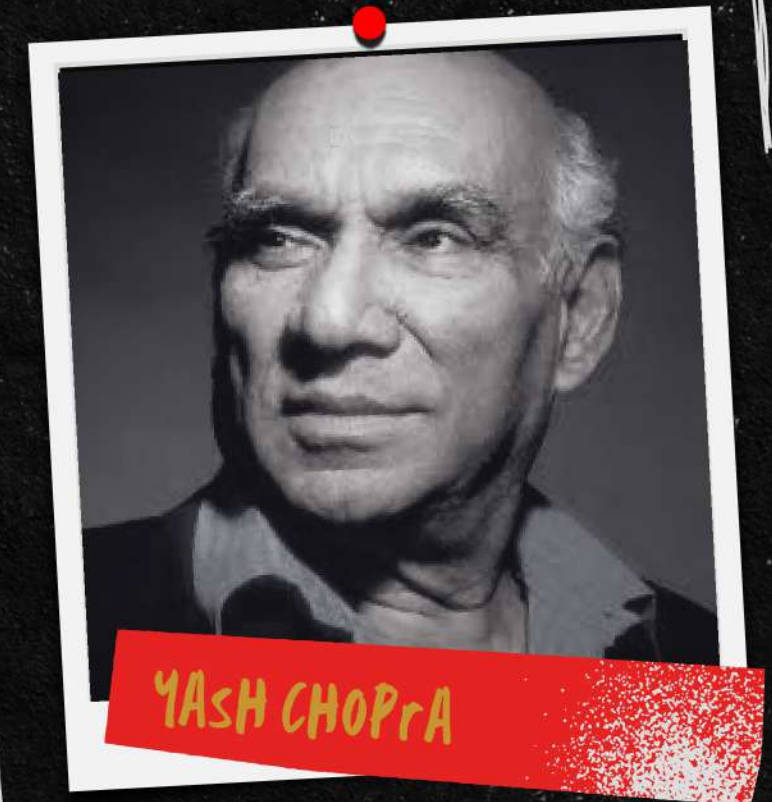
Veer-Zaara (2004)

- Rating: 7.8
- Votes: 50101



Darr (1993)

- Rating: 7.8
- Votes: 22635

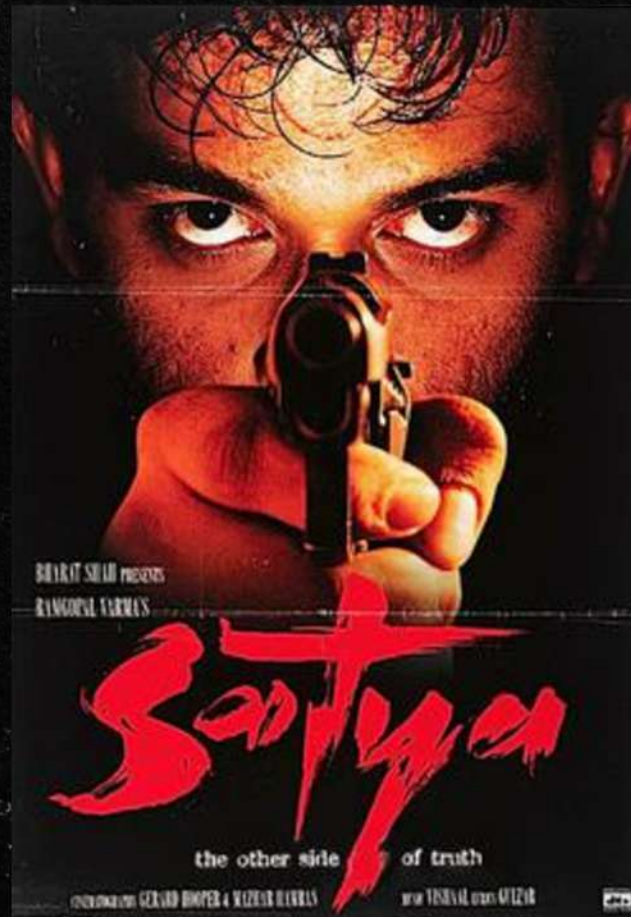


TOP 3 Directors in INDIA

Ram Gopal Varma (10 movies)

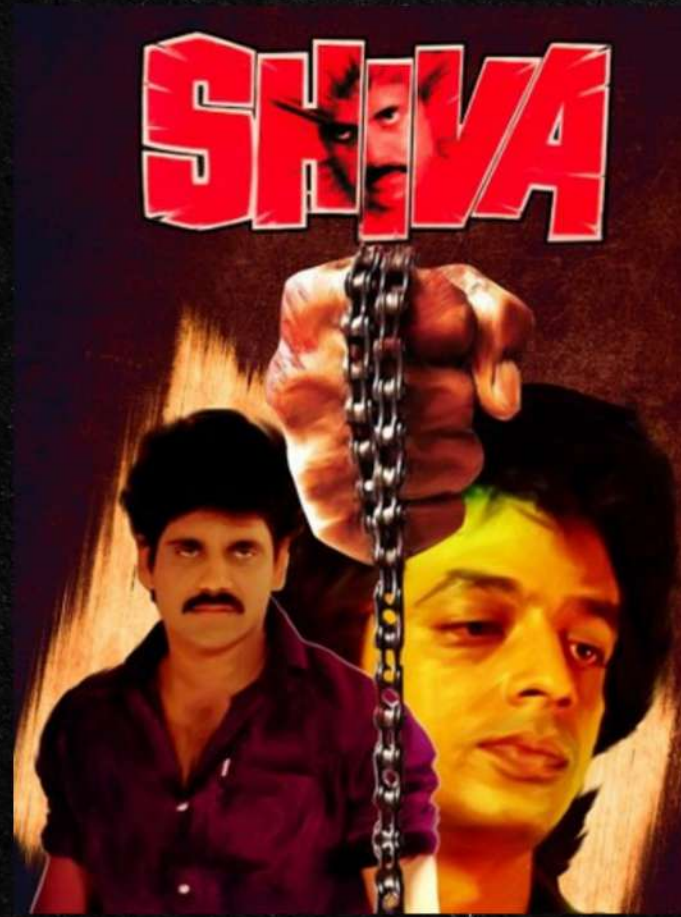
- Average rating: 5.60
- Average votes: 3,280

Example movies:



Satya (1998)

- Rating: 8.2
- Votes: 13918



Shiva (1989)

- Rating: 8.1
- Votes: 3399



Company (2002)

- Rating: 7.9
- Votes: 14795



ANY QUESTIONS?

