

SQLDataScience_Movies_01

July 19, 2024

```
[1]: import pandas as pd
```

```
Movie = pd.read_csv('movies.csv')
```

```
[2]: Movie['ReleaseDate'] = pd.to_datetime(Movie['ReleaseDate'], format='%d-%b-%y')
Movie
```

```
[2]:
```

	Title	Rating	Genre	Year	ReleaseDate	Score	\
0	Superbabies: Baby Geniuses 2	PG	Comedy	2004	2004-08-27	1.9	
1	The Hottie & the Nottie	PG-13	Comedy	2008	2008-02-21	1.9	
2	Disaster Movie	PG-13	Comedy	2008	2008-08-29	1.9	
3	From Justin to Kelly	PG	Comedy	2003	2003-06-20	2.1	
4	House of the Dead	R	Action	2003	2003-10-10	2.1	
...	
7609	The Dark Knight	PG-13	Action	2008	2008-07-18	9.0	
7610	The Shawshank Redemption	R	Drama	1994	1994-10-14	9.3	
7611	It's Just Us	NaN	Drama	2020	2020-10-01	NaN	
7612	Love by Drowning	R	Drama	2020	2020-11-06	NaN	
7613	The Robinsons	NaN	Action	2020	2020-11-10	NaN	

	Votes	Director	Writer	Star	\
0	30000.0	Bob Clark	Robert Grasmere	Jon Voight	
1	36000.0	Tom Putnam	Heidi Ferrer	Paris Hilton	
2	88000.0	Jason Friedberg	Jason Friedberg	Carmen Electra	
3	25000.0	Robert Iscove	Kim Fuller	Kelly Clarkson	
4	36000.0	Uwe Boll	Mark A. Altman	Jonathan Cherry	
...	
7609	2400000.0	Christopher Nolan	Jonathan Nolan	Christian Bale	
7610	2400000.0	Frank Darabont	Stephen King	Tim Robbins	
7611	NaN	James Randall	James Randall	Christina Roz	
7612	NaN	Justin Kreinbrink	C.E. Poverman	Nicky Whelan	
7613	NaN	Directors	Aleks Alifirenko Jr.	Billy Hartmann	

	Country	Budget	Gross	\
0	Germany	20000000.0	9.448644e+06	
1	United States	NaN	1.596232e+06	
2	United States	20000000.0	3.481682e+07	

3	United States	12000000.0	4.928883e+06
4	Germany	12000000.0	1.381818e+07
...
7609	United States	185000000.0	1.005974e+09
7610	United States	25000000.0	2.881729e+07
7611	United States	15000.0	NaN
7612	United States	1000000.0	NaN
7613	United States	10000.0	NaN

	Company	Runtime
0	ApolloMedia Distribution	88.0
1	Purple Pictures	91.0
2	Lionsgate	87.0
3	19 Entertainment	81.0
4	Boll Kino Beteiligungs GmbH & Co. KG	90.0
...
7609	Warner Bros.	152.0
7610	Castle Rock Entertainment	142.0
7611	NaN	120.0
7612	Dow Jazz Films	121.0
7613	NaN	90.0

[7614 rows x 15 columns]

```
[3]: # SELECT Title
      # FROM Movie
      # WHERE Rating = "G"
```

```
[4]: # SELECT Title
      # FROM Movie
      # WHERE ReleaseDate > "2020-01-01"
```

```
[5]: # SELECT Title
      # FROM Movie
      # WHERE Rating In ("G", "PG")

Movie.query("Rating=='G' or Rating=='PG') [['Title']]
```

```
[5]:
      Title
0      Superbabies: Baby Geniuses 2
3      From Justin to Kelly
5      Leonard Part 6
6      Son of the Mask
15     Dragonball Evolution
...
7574    Back to the Future
7575    Cinema Paradiso
```

7577	The Lion King
7594	Spirited Away
7598	Star Wars: Episode V - The Empire Strikes Back

[1398 rows x 1 columns]

```
[6]: # SELECT Title  
# FROM Movie  
# WHERE Rating = "PG-13" and ReleaseDate > "2008-01-01"
```

```
[7]: # SELECT Title, ReleaseDate  
# FROM Movie  
# WHERE Budget > Gross
```

```
[8]: # SELECT Title  
# FROM Movie  
# WHERE Budget < Gross AND Rating = "G"
```

```
[9]: # SELECT Title, Year, Budget  
# FROM Movie  
# WHERE Budget > 300000000
```