



## Introduction to Data Management PROJECT REPORT

(Project Semester August-December 2020)

PROJECT REPORT ON

Analysing IMDB data from 2006 to 2016

Submitted by

Rounak Agarwal

12009266

Programme: Bachelor of Technology

Section: KM100

Course Code: INT217

Under the Guidance of Veerpal

Kaur: 25909

Assistant Professor

Discipline of CSE/IT

Lovely School of Computer Science & Engineering

Lovely Professional University, Phagwara

## DECLARATION

I, Rounak Agarwal, student of Computer Science & Engineering under CSE/IT Discipline at, Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date: 05-11-2022

Rounak Agarwal

Registration No: 12009266

## ACKNOWLEDGEMENT

Primarily I'd thank god for being able to complete my project with success. Then I'd like to thank my mentor Ms. Veerpal Kaur, whose valuable guidance has been the ones that helped me patch this project and make it full proof success in contribution towards the completion of this project.

Last but not least I'd rather thanks to Lovely Professional University, and my parent's inspiration, who gave me this golden opportunity to learn many new things, to learn another aspects of life.

Rounak Agarwal

## CONTENTS:

Sr No.	Title	Page No.
1	Introduction	5
2	Objectives/Scope of the Analysis	6
3	Source of dataset	7
4	ETL Process	9
5	Analysis of dataset	18
6	List of Analysis with results	29
7	Final Dashboard	34
8	Bibliography	35

## INTRODUCTION

- ☐ Data management is important because the data your organization creates is a very valuable resource.
- ☐ The last thing you want to do is spend time and resources collecting data and business intelligence, only to lose or misplace that information.
- ☐ In that case, you would then have to spend time and resources again to get that same business intelligence you already had.
- ☐ And on that data analysis is carried out which show visualization of our problems in efficient way.
- ☐ Data Analysis is a process of inspecting, cleansing, transforming, and modeling data with the goal of discovering useful information, informing conclusions, and supporting decision- making.
- ☐ This project is based on such data analysis on IMDb data from 2006 to 2016
- ☐ IMDb is an online database of information related to films, television programs, home videos, video games, and streaming content online – including cast, production crew and personal biographies, plot summaries, trivia, ratings, and fan and critical reviews.
- ☐ This IMDb dataset contains 12 data fields.

## OBJECTIVES/SCOPE OF ANALYSIS

After analysis of the dataset, the aim of this project is to give answer of given

objectives in easy way:

- Total revenues earned by movies in each year
- Total Movies released in each year
- Top 5 IMDb rating movies
- Top 7 IMDb metascore movies
- Top 5 revenue movies from whole dataset or each year
- Most voted director from whole dataset or in each year
- Top 5 longest movies from whole dataset or in each year

## SOURCE OF DATASET:

Source of dataset: <https://www.kaggle.com/PromptCloudHQ/imdb-data>

The dataset is based on 1000 popular movies on IMDb from 2006 to 2016.

The columns included in the dataset are given below:

- ☐ Rank
  - ☐ Movie rank order
- ☐ Title
  - ☐ The title of the film
- ☐ Genre
  - ☐ Category of the film
- ☐ Description
  - ☐ Summary of the film
- ☐ Director
  - ☐ The name of the film's director
- ☐ Actors
  - ☐ Name of the Cast
- ☐ Year
  - ☐ Released year of the film
- ☐ Runtime
  - ☐ Duration of film in minutes
- ☐ Rating
  - ☐ IMDB rating given by the user ☐

Votes

- ☐ Number of votes
- ☐ Revenue

- ☐ Total revenues in millions
  - ☐ Metascores
    - ☐ An aggregated average of critic scores. Values are between 0 and 100.
- Higher scores represent positive reviews.

Sample of dataset with data fields is given below:

	A	B	C	D	E	F	G	H	I	J	K	L
1	Rank	Title	Genre	Description	Director	Actors	Year	Runtime	Rating	Votes	Revenue	Metascore
2	1	Guardians of the Galaxy	Action,Adventure,Sci-Fi	A group of intergalactic criminal	James Gunn	Chris Pratt, Vin Di	2014	121	8.1	757074	333.13	76
3	2	Prometheus	Adventure,Mystery,Sci-Fi	Following clues to the origin of	Ridley Scott	Noomi Rapace, Lo	2012	124	7	485820	126.46	65
4	3	Split	Horror,Thriller	Three girls are kidnapped by a m	M. Night Shyamalan	James McAvoy, Ai	2016	117	7.3	157606	138.12	62
5	4	Sing	Animation,Comedy,Family	In a city of humanoid animals, a	Christophe Lourdelet	Matthew McCona	2016	108	7.2	60545	270.32	59
6	5	Suicide Squad	Action,Adventure,Fantasy	A secret government agency rec	David Ayer	Will Smith, Jared	2016	123	6.2	393727	325.02	40
7	6	The Great Wall	Action,Adventure,Fantasy	European mercenaries searching	Yimou Zhang	Matt Damon, Tian	2016	103	6.1	56036	45.13	42
8	7	La La Land	Comedy,Drama,Music	A jazz pianist falls for an aspiri	ng Damien Chazelle	Ryan Gosling, Emr	2016	128	8.3	258682	151.06	93
9	8	Mindhorn	Comedy	A has-been actor best known for	Sean Foley	Essie Davis, Andre	2016	89	6.4	2490		71
10	9	The Lost City of Z	Action,Adventure,Biography	A true-life drama, centering on	James Gray	Charlie Hunnam, f	2016	141	7.1	7188	8.01	78
11	10	Passengers	Adventure,Drama,Romance	A spacecraft traveling to a distar	Morten Tyldum	Jennifer Lawrence	2016	116	7	192177	100.01	41
12	11	Fantastic Beasts and Where	Adventure,Family,Fantasy	The adventures of writer Newt	David Yates	Eddie Redmayne,	2016	133	7.5	232072	234.02	66
13	12	Hidden Figures	Biography,Drama,History	The story of a team of female Af	Theodore Melfi	Taraji P. Henson, e	2016	127	7.8	93103	169.27	74
14	13	Rogue One	Action,Adventure,Sci-Fi	The Rebel Alliance makes a risky	Gareth Edwards	Felicity Jones, Die	2016	133	7.9	323118	532.17	65
15	14	Moana	Animation,Adventure,Comedy	In Ancient Polynesia, when a tei	Ron Clements	Auli'i Cravalho, Dv	2016	107	7.7	118151	248.75	81
16	15	Colossal	Action,Comedy,Drama	Gloria is an out-of-work party gi	Nacho Vigalondo	Anne Hathaway, J	2016	109	6.4	8612	2.87	70
17	16	The Secret Life of Pets	Animation,Adventure,Comedy	The quiet life of a terrier named	Chris Renaud	Louis C.K., Eric Stc	2016	87	6.6	120259	368.31	61
18	17	Hacksaw Ridge	Biography,Drama,History	WWII American Army Medic Des	Mel Gibson	Andrew Garfield,	2016	139	8.2	211760	67.12	71
19	18	Jason Bourne	Action,Thriller	The CIA's most dangerous forme	Paul Greengrass	Matt Damon, Tom	2016	123	6.7	150823	162.16	58
20	19	Lion	Bioerathv.Drama	A five-year-old Indian boy gets	Garth Davis	Dev Patel, Nicole	2016	118	8.1	102061	51.69	69

## ETL PROCESS:



- ETL is a process that extracts the data from different source systems, then transforms the data (like applying calculations, concatenations, etc.) and finally loads the data into the Data Warehouse system.
- Full form of ETL is Extract, Transform and Load.
- The triple combination of ETL provides crucial functions that are many times combined into a single application or suite of tools that help in the following areas:
  - Enhances Business Intelligence solutions for decision making.
  - Allows verification of data transformation, aggregation and calculations rules.
  - Allows sample data comparison between source and target system.
  - Helps to improve productivity as it codifies and reuses without additional technical skills.
- Initially, the raw dataset was arranged as shown in given picture:

Rank	Title	Genre	Descriptio	Director	Actors	Year	Runtime (	Rating	Votes	Revenue (	Metascore
1	Guardians of the Galaxy	Action, Adventure	A group of	James Gunn	Chris Pratt	2014	121	8.1	757074	333.13	76
2	Prometheus	Adventure, Sci-Fi	Following	Ridley Scott	Noomi Rapace	2012	124	7	485820	126.46	65
3	Split	Horror, Thriller	Three girls	M. Night Shyamalan	James McAvoy	2016	117	7.3	157606	138.12	62
4	Sing	Animation, Comedy	In a city of	Christopher L. Butler	Matthew McConaughey	2016	108	7.2	60545	270.32	59
5	Suicide Squad	Action, Adventure	A secret g	David Ayer	Will Smith	2016	123	6.2	393727	325.02	40
6	The Great Wall	Action, Adventure	European	Yimou Zhang	Matt Damon	2016	103	6.1	56036	45.13	42
7	La La Land	Comedy, Drama, Romance	A jazz pianist	Damien Chazelle	Ryan Gosling	2016	128	8.3	258682	151.06	93
8	Mindhorn	Comedy	A has-bee	Sean Foley	Essie Davis	2016	89	6.4	2490		71
9	The Lost City of Z	Action, Adventure	A true-life	James Gray	Charlie Hulse	2016	141	7.1	7188	8.01	78
10	Passengers	Adventure, Sci-Fi	A spacecr	Morten Tyldum	Jennifer Lawrence	2016	116	7	192177	100.01	41
11	Fantastic Beasts and Where to Find Them	Adventure, Fantasy	The advent	David Yates	Eddie Redmayne	2016	133	7.5	232072	234.02	66
12	Hidden Figures	Biography	The story	Theodore Melfi	Taraji P. Henson	2016	127	7.8	93103	169.27	74
13	Rogue One: A Star Wars Story	Action, Adventure	The Rebel	Gareth Edwards	Felicity Jones	2016	133	7.9	323118	532.17	65
14	Moana	Animation	In Ancient	Ron Clement	Auli'i Cravalho	2016	107	7.7	118151	248.75	81
15	Colossal	Action, Comedy	Gloria is a	Nacho Vigalondo	Anne Hathaway	2016	109	6.4	8612	2.87	70
16	The Secret Life of Pets	Animation	The quiet	Chris Renaud	Louis C.K.	2016	87	6.6	120259	368.31	61
17	Hacksaw Ridge	Biography	WWII Am	Mel Gibson	Andrew Garfield	2016	139	8.2	211760	67.12	71
18	Jason Bourne	Action, Thriller	The CIA's	Paul Greengrass	Matt Damon	2016	123	6.7	150823	162.16	58
19	Lion	Biography	A five-yea	Garth Davis	Dev Patel	2016	118	8.1	102061	51.69	69
20	Arrival	Drama, Mystery	When twe	Denis Villeneuve	Amy Adams	2016	116	8	340798	100.5	81
21	Gold	Adventure	Kenny We	Stephen G. Soderbergh	Matthew McConaughey	2016	120	6.7	19053	7.22	49
22	Manchester by the Sea	Drama	A depress	Kenneth Lonergan	Casey Affleck	2016	137	7.9	134213	47.7	96

Steps taken to clean dataset thorough ETL process

## Step 1:

- Firstly, open blank excel file and go to Data from Tab and select New query and entered data csv file from it.

Column1	Column2	Column3	Column4	Column5	Column6
1	Rank	Title	Genre	Description	Director
2	1	Guardians of the Galaxy	Action,Adventure,Sci-Fi	A group of intergalactic criminals are forced to work together to stop a far	James Gunn
3	2	Prometheus	Adventure,Mystery,Sci-Fi	Following clues to the origin of mankind, a team finds a structure on a dist	Ridley Scott
4	3	Split	Horror,Thriller	Three girls are kidnapped by a man with a diagnosed 23 distinct personali	M. Night Shyamalan
5	4	Sing	Animation,Comedy,Family	In a city of humanoid animals, a hustling theater impresario's attempt to s	Christophe Lourdelet
6	5	Suicide Squad	Action,Adventure,Fantasy	A secret government agency recruits some of the most dangerous incarce	David Ayer
7	6	The Great Wall	Action,Adventure,Fantasy	European mercenaries searching for black powder become embroiled in tl	Yimou Zhang
8	7	La La Land	Comedy,Drama,Music	A jazz pianist falls for an aspiring actress in Los Angeles.	Damien Chazelle
9	8	Mindhorn	Comedy	A has-been actor best known for playing the title character in the 1980s d	Sean Foley
10	9	The Lost City of Z	Action,Adventure,Biography	A true-life drama, centering on British explorer Col. Percival Fawcett, who	James Gray
11	10	Passengers	Adventure,Drama,Romance	A spacecraft traveling to a distant colony planet and transporting thousan	Morten Tyldum
12	11	Fantastic Beasts and Where to Find Them	Adventure,Family,Fantasy	The adventures of writer Newt Scamander in New York's secret communit	David Yates
13	12	Hidden Figures	Biography,Drama,History	The story of a team of female African-American mathematicians who serv	Theodore Melfi
14	13	Rogue One	Action,Adventure,Sci-Fi	The Rebel Alliance makes a risky move to steal the plans for the Death Sta	Gareth Edwards
15	14	Moana	Animation,Adventure,Comedy	In Ancient Polynesia, when a terrible curse incurred by the Demigod Maui	Ron Clements
16	15	Colossal	Action,Comedy,Drama	Gloria is an out-of-work party girl forced to leave her life in New York City,	Nacho Vigalondo
17	16	The Secret Life of Pets	Animation,Adventure,Comedy	The quiet life of a terrier named Max is upended when his owner takes in	Chris Renaud
18	17	Hacksaw Ridge	Biography,Drama,History	WWII American Army Medic Desmond T. Doss, who served during the Bat	Mel Gibson
19	18	Jason Bourne	Action,Thriller	The CIA's most dangerous former operative is drawn out of hiding to unco	Paul Greengrass
20	19	Lion	Biography,Drama	A five-year-old Indian boy gets lost on the streets of Calcutta, thousands o	Garth Davis
21	20	Arrival	Drama,Mystery,Sci-Fi	When twelve mysterious spacecraft appear around the world, linguistics p	Denis Villeneuve
22	21	Gold	Adventure,Drama,Thriller	Kenny Wells, a prospector desperate for a lucky break, teams up with a sir	Stephen Gaghan
23	22	Manchester by the Sea	Drama	A depressed uncle is asked to take care of his teenage nephew after the b	Kenneth Lonergan
24	23	Hounds of Love	Crime,Drama,Horror	A cold-blooded predatory couple while cruising the streets in search of th	Ben Young
25	24	Troll	Animation,Adventure,Comedy	After the Bergens invade Troll Village, Donnu, the happiest Troll ever born	Walt Dohrn

Query Settings

PROPERTIES

Name

IMDB-Movie-Data

APPLIED STEPS

Source

Promoted Headers

Changed Type

Removed Columns

Sorted Rows

Filtered Rows

Sorted Rows1

Removed Duplicates

10MNS, 999+ ROWS

PREVIEW DOWNLOADED AT 10:10 PM

## Step 2:

- After entering the query editor, data fields title will promote into header

	Rank	Title	Genre	Description	Director
1	1	Guardians of the Galaxy	Action,Adventure,Sci-Fi	A group of intergalactic criminals are forced to work together to stop a far	James Gunn
2	2	Prometheus	Adventure,Mystery,Sci-Fi	Following clues to the origin of mankind, a team finds a structure on a dist	Ridley Scott
3	3	Split	Horror,Thriller	Three girls are kidnapped by a man with a diagnosed 23 distinct personali	M. Night Shyamalan
4	4	Sing	Animation,Comedy,Family	In a city of humanoid animals, a hustling theater impresario's attempt to s	Christophe Lourdelet
5	5	Suicide Squad	Action,Adventure,Fantasy	A secret government agency recruits some of the most dangerous incarce	David Ayer
6	6	The Great Wall	Action,Adventure,Fantasy	European mercenaries searching for black powder become embroiled in tl	Yimou Zhang
7	7	La La Land	Comedy,Drama,Music	A jazz pianist falls for an aspiring actress in Los Angeles.	Damien Chazelle
8	8	Mindhorn	Comedy	A has-been actor best known for playing the title character in the 1980s di	Sean Foley
9	9	The Lost City of Z	Action,Adventure,Biography	A true-life drama, centering on British explorer Col. Percival Fawcett, who	James Gray
10	10	Passengers	Adventure,Drama,Romance	A spacecraft traveling to a distant colony planet and transporting thousan	Morten Tyldum
11	11	Fantastic Beasts and Where to Find Them	Adventure,Family,Fantasy	The adventures of writer Newt Scamander in New York's secret communit	David Yates
12	12	Hidden Figures	Biography,Drama,History	The story of a team of female African-American mathematicians who serv	Theodore Melfi
13	13	Rogue One	Action,Adventure,Sci-Fi	The Rebel Alliance makes a risky move to steal the plans for the Death Sta	Gareth Edwards
14	14	Moana	Animation,Adventure,Comedy	In Ancient Polynesia, when a terrible curse incurred by the Demigod Maui	Ron Clements
15	15	Colossal	Action,Comedy,Drama	Gloria is an out-of-work party girl forced to leave her life in New York City,	Nacho Vigalondo
16	16	The Secret Life of Pets	Animation,Adventure,Comedy	The quiet life of a terrier named Max is upended when his owner takes in	Chris Renaud
17	17	Hacksaw Ridge	Biography,Drama,History	WWII American Army Medic Desmond T. Doss, who served during the Bat	Mel Gibson
18	18	Jason Bourne	Action,Thriller	The CIA's most dangerous former operative is drawn out of hiding to unco	Paul Greengrass
19	19	Lion	Biography,Drama	A five-year-old Indian boy gets lost on the streets of Calcutta, thousands o	Garth Davis
20	20	Arrival	Drama,Mystery,Sci-Fi	When twelve mysterious spacecraft appear around the world, linguistics p	Denis Villeneuve
21	21	Gold	Adventure,Drama,Thriller	Kenny Wells, a prospector desperate for a lucky break, teams up with a si	Stephen Gaghan
22	22	Manchester by the Sea	Drama	A depressed uncle is asked to take care of his teenage nephew after the b	Kenneth Lonergan
23	23	Hounds of Love	Crime,Drama,Horror	A cold-blooded predatory couple while cruising the streets in search of th	Ben Young
24	24	Trolls	Animation,Adventure,Comedy	After the Bergens invade Troll Village, Poppy, the happiest Troll ever born,	Walt Dohrn
25	25	Independence Day: Resurgence	Action,Adventure,Sci-Fi	Twelve decades after the first Independence Day invasion, Earth is faced with	Roland Emmerich

COLUMNS: 999+ ROWS

PREVIEW DOWNLOADED AT 10:10 PM

## Query Settings

### PROPERTIES

Name

IMDB-Movie-Data

[All Properties](#)

### APPLIED STEPS

Source

☒ Promoted Headers

Changed Type

Removed Columns

Sorted Rows

Filtered Rows

Sorted Rows1

Removed Duplicates

### Step 3:

- ❑ Changed data type will change the type of data fields such as metascores converted from text into number, Revenue converted from text into currency

	Rank	Title	Genre	Description	Director
1	1	Guardians of the Galaxy	Action,Adventure,Sci-Fi	A group of intergalactic criminals are forced to work together to stop a far	James Gunn
2	2	Prometheus	Adventure,Mystery,Sci-Fi	Following clues to the origin of mankind, a team finds a structure on a dist	Ridley Scott
3	3	Split	Horror,Thriller	Three girls are kidnapped by a man with a diagnosed 23 distinct personalit	M. Night Shyamalan
4	4	Sing	Animation,Comedy,Family	In a city of humanoid animals, a hustling theater impresario's attempt to s	Christophe Lourdelet
5	5	Suicide Squad	Action,Adventure,Fantasy	A secret government agency recruits some of the most dangerous incarce	David Ayer
6	6	The Great Wall	Action,Adventure,Fantasy	European mercenaries searching for black powder become embroiled in th	Yimou Zhang
7	7	La La Land	Comedy,Drama,Music	A jazz pianist falls for an aspiring actress in Los Angeles.	Damien Chazelle
8	8	Mindhorn	Comedy	A has-been actor best known for playing the title character in the 1960s d	Sean Foley
9	9	The Lost City of Z	Action,Adventure,Biography	A true-life drama, centering on British explorer Col. Percival Fawcett, who	James Gray
10	10	Passengers	Adventure,Drama,Romance	A spacecraft traveling to a distant colony planet and transporting thousa	Morten Tyldum
11	11	Fantastic Beasts and Where to Find Them	Adventure,Family,Fantasy	The adventures of writer Newt Scamander in New York's secret communit	David Yates
12	12	Hidden Figures	Biography,Drama,History	The story of a team of female African-American mathematicians who serv	Theodore Melfi
13	13	Rogue One	Action,Adventure,Sci-Fi	The Rebel Alliance makes a risky move to steal the plans for the Death Sta	Gareth Edwards
14	14	Moana	Animation,Adventure,Comedy	In Ancient Polynesia, when a terrible curse incurred by the Demigod Maui	Ron Clements
15	15	Colossal	Action,Comedy,Drama	Gloria is an out-of-work party girl forced to leave her life in New York City,	Nacho Vigalondo
16	16	The Secret Life of Pets	Animation,Adventure,Comedy	The quiet life of a terrier named Max is upended when his owner takes in	Chris Renaud
17	17	Hacksaw Ridge	Biography,Drama,History	WWII American Army Medic Desmond T. Doss, who served during the Bat	Mel Gibson
18	18	Jason Bourne	Action,Thriller	The CIA's most dangerous former operative is drawn out of hiding to unco	Paul Greengrass
19	19	Lion	Biography,Drama	A five-year-old Indian boy gets lost on the streets of Calcutta, thousands o	Garth Davis
20	20	Arrival	Drama,Mystery,Sci-Fi	When twelve mysterious spacecraft appear around the world, linguistics p	Denis Villeneuve
21	21	Gold	Adventure,Drama,Thriller	Kenny Wells, a prospector desperate for a lucky break, teams up with a sir	Stephen Gaghan
22	22	Manchester by the Sea	Drama	A depressed uncle is asked to take care of his teenage nephew after the b	Kenneth Lonergan
23	23	Hounds of Love	Crime,Drama,Horror	A cold-blooded predatory couple while cruising the streets in search of th	Ben Young
24	24	Trolls	Animation,Adventure,Comedy	After the Bergens invade Troll Village, Poppy, the happiest Troll ever born,	Walt Dohrn
25	25	Indiana Jones and the Kingdom of the Crystal Skull	Action,Adventure,Sci-Fi	Twelve decades after the first Indiana Jones film, the world is faced with	Steven Spielberg

#### Step 4:

- Now, from query editor select Home tab and remove column which are not in required for analysis. Column which are removed from dataset are given below:

- Rank
- Genre
- Description
- Actor

Rank	Title	Director	Year	Runtime (Minutes)	Rating	Votes	Revenue (Millions)	Metascore	
1	Guardians of the Galaxy	James Gunn	2014	121	8.1	757074	333.13	76	
2	Prometheus	Ridley Scott	2012	124	7	485820	126.46	65	
3	Split	M. Night Shyamalan	2016	117	7.3	157606	138.12	62	
4	Sing	Christophe Lourdelet	2016	108	7.2	60545	270.32	59	
5	Suicide Squad	David Ayer	2016	123	6.2	393727	325.02	40	
6	The Great Wall	Yimou Zhang	2016	103	6.1	56036	45.13	42	
7	La La Land	Damien Chazelle	2016	128	8.3	258682	151.06	93	
8	Mindhorn	Sean Foley	2016	89	6.4	2490	null	71	
9	The Lost City of Z	James Gray	2016	141	7.1	7188	8.01	78	
10	Passengers	Morten Tyldum	2016	116	7	192177	100.01	41	
11	Fantastic Beasts and Where to Find Them	David Yates	2016	133	7.5	232072	234.02	66	
12	Hidden Figures	Theodore Melfi	2016	127	7.8	99103	169.27	74	
13	Rogue One	Gareth Edwards	2016	133	7.9	323118	532.17	65	
14	Moana	Ron Clements	2016	107	7.7	118151	248.75	81	
15	Colossal	Nacho Vigalondo	2016	109	6.4	8612	2.87	70	
16	The Secret Life of Pets	Chris Renaud	2016	87	6.6	120259	368.31	61	
17	Hacksaw Ridge	Mel Gibson	2016	139	8.2	211760	67.12	71	
18	Jason Bourne	Paul Greengrass	2016	123	6.7	150823	162.16	58	
19	Lion	Garth Davis	2016	118	8.1	102061	51.69	69	
20	Arrival	Denis Villeneuve	2016	116	8	340798	100.5	81	
21	Gold	Stephen Gaghan	2016	120	6.7	19053	7.22	49	
22	Manchester by the Sea	Kenneth Lonergan	2016	137	7.9	134213	47.7	96	
23	Hounds of Love	Ben Young	2016	108	6.7	1115	null	72	
24	Trolls	Walt Dohrn	2016	92	6.5	38552	153.69	56	
25	Independence Day: Resurgence	Roland Emmerich	2016	120	5.3	127553	103.14	32	

Query Settings

PROPERTIES

Name

IMDB-Movie-Data

All Properties

APPLIED STEPS

Source

Promoted Headers

Changed Type

Removed Columns

Sorted Rows

Filtered Rows

Sorted Rows1

Removed Duplicates

8 COLUMNS, 999+ ROWS

PREVIEW DOWNLOADED AT 10:10 PM



## Step 5:

- Now to remove null values we first sort the dataset show that null values comes on the top and after that we can delete it through Delete rows from Home tab.

	Title	Director	Year	Runtime (Minutes)	Rating	Votes	Revenue (Millions)	Metascore
1	The Master	Paul Thomas Anderson	2012	144	7.1	112902	16.38	null
2	5- 25- 77	Patrick Read Johnson	2007	113	7.1	241	null	null
3	Grindhouse	Robert Rodriguez	2007	191	7.6	160350	25.03	null
4	Silent Hill	Christophe Gans	2006	125	6.6	184152	46.98	null
5	Disturbia	D.J. Caruso	2007	105	6.9	193491	80.05	null
6	Bronson	Nicolas Winding Refn	2008	92	7.1	93972	0.1	null
7	Up in the Air	Jason Reitman	2009	109	7.4	279694	83.81	null
8	Boyka: Undisputed IV	Todor Chapkanov	2016	86	7.4	10428	null	null
9	The Road	John Hillcoat	2009	111	7.3	187302	0.06	null
10	American Pastoral	Ewan McGregor	2016	108	6.1	7115	null	null
11	Selma	Ava DuVernay	2014	128	7.5	67637	52.07	null
12	P.S. I Love You	Richard LaGravenese	2007	126	7.1	177247	53.68	null
13	Handsome Devil	John Butler	2016	95	7.4	338	null	null
14	Rules Don't Apply	Warren Beatty	2016	127	5.7	3731	3.65	null
15	Free State of Jones	Gary Ross	2016	139	6.9	29895	null	null
16	Chalk It Up	Hisonni Johnson	2016	90	4.8	499	null	null
17	10 Years	Jamie Linden	2011	100	6.1	19636	0.2	null
18	Ender's Game	Gavin Hood	2013	114	6.7	194236	61.66	null
19	Elysium	Neill Blomkamp	2013	109	6.6	358932	null	null
20	The Stakelander	Dan Berk	2016	81	5.3	1263	null	null
21	Sleeping Beauty	Julia Leigh	2011	101	5.3	27006	0.03	null
22	The Walk	Robert Zemeckis	2015	123	7.3	92378	10.14	null
23	Vincent N Roxxy	Gary Michael Schultz	2016	110	5.5	403	null	null
24	Angry Birds	Clay Kaytis	2016	97	6.3	55781	107.51	null
25	Contratiempo	Oriol Paulo	2016	106	7.9	7204	null	null

	Title	Director	Year	Runtime (Minutes)	Rating	Votes	Revenue (Millions)	Metascore
1	Nine Lives	Barry Sonnenfeld	2016	87	5.3	12435	19.64	11
2	Disaster Movie	Jason Friedberg	2008	87	1.9	77207	14.17	15
3	Dirty Grandpa	Dan Mazer	2016	102	6	75137	35.54	18
4	Movie 43	Elizabeth Banks	2013	94	4.3	83625	8.83	18
5	Mother's Day	Garry Marshall	2016	118	5.6	20221	32.46	18
6	Grown Ups 2	Dennis Dugan	2013	101	5.4	114482	133.67	19
7	The Do-Over	Steven Brill	2016	108	5.7	24761	0.54	22
8	Max Steel	Stewart Hendler	2016	92	4.6	11555	3.77	22
9	Collateral Beauty	David Frankel	2016	97	6.8	43977	30.98	23
10	Underworld: Blood Wars	Anna Foerster	2016	91	5.8	41362	30.35	23
11	The Sea of Trees	Gus Van Sant	2015	110	5.9	7475	0.02	23
12	Couples Retreat	Peter Billingsley	2009	113	5.5	86417	109.18	23
13	Bride Wars	Gary Winick	2009	89	5.5	83976	58.72	24
14	The Loft	Erik Van Looy	2014	108	6.3	38804	5.98	24
15	Gods of Egypt	Alex Proyas	2016	126	5.5	73568	31.14	25
16	Taken 3	Olivier Megaton	2014	109	6	144715	89.25	26
17	The Choice	Ross Katz	2016	111	6.6	20514	18.71	26
18	Fantastic Four	Josh Trank	2015	100	4.3	121847	56.11	27
19	Sex and the City 2	Michael Patrick King	2010	146	4.3	62403	95.33	27
20	I Spit on Your Grave	Steven R. Monroe	2010	108	6.3	60133	0.09	27
21	Pixels	Chris Columbus	2015	105	5.6	101092	78.75	27
22	The Ugly Truth	Robert Luketic	2009	96	6.5	172558	88.92	28
23	Pandorum	Christian Alvart	2009	108	6.8	126656	10.33	28
24	London Has Fallen	Babak Najafi	2016	99	5.9	100702	62.4	28
25	Fool's Gold	Andy Tennant	2008	112	5.7	62719	70.22	29

JMNS, 838 ROWS

## Step 6:

- Finally, sort the dataset through filter of rating data fields

	Title	Director	Year	Runtime (Minutes)	Rating	Votes	Revenue (Millions)	Metascore
1	The Dark Knight	Christopher Nolan	2008	152	9	1791916	533.32	82
2	Inception	Christopher Nolan	2010	148	8.8	1583625	292.57	74
3	Kimi no na wa	Makoto Shinkai	2016	106	8.6	34110	4.68	79
4	The Intouchables	Olivier Nakache	2011	112	8.6	557965	13.18	57
5	Interstellar	Christopher Nolan	2014	169	8.6	1047747	187.99	74
6	The Departed	Martin Scorsese	2006	151	8.5	937414	132.37	85
7	Whiplash	Damien Chazelle	2014	107	8.5	477276	13.09	88
8	The Prestige	Christopher Nolan	2006	130	8.5	913152	53.08	66
9	Taare Zameen Par	Aamir Khan	2007	165	8.5	102697	1.2	42
10	The Lives of Others	Florian Henckel von Donnersmarck	2006	137	8.5	278103	11.28	89
11	The Dark Knight Rises	Christopher Nolan	2012	164	8.5	1222645	448.13	78
12	Django Unchained	Quentin Tarantino	2012	165	8.4	1039115	162.8	81
13	3 Idiots	Rajkumar Hirani	2009	170	8.4	238789	6.52	67
14	Toy Story 3	Lee Unkrich	2010	103	8.3	586669	414.98	92
15	Inglourious Basterds	Quentin Tarantino	2009	153	8.3	959065	120.52	69
16	Jagten	Thomas Vinterberg	2012	115	8.3	192263	0.61	76
17	La La Land	Damien Chazelle	2016	128	8.3	258682	151.06	93
18	Up	Pete Docter	2009	96	8.3	722203	292.98	88
19	Warrior	Gavin O'Connor	2011	140	8.2	355722	13.65	71
20	PK	Rajkumar Hirani	2014	153	8.2	103279	10.57	51
21	El secreto de sus ojos	Juan José Campanella	2009	129	8.2	144524	20.17	80
22	The Wolf of Wall Street	Martin Scorsese	2013	180	8.2	865134	116.87	75
23	Room	Lenny Abrahamson	2015	118	8.2	224132	14.68	86
24	Incendies	Denis Villeneuve	2010	131	8.2	92863	6.86	80
25	Inside Out	Pete Docter	2015	95	8.2	416689	356.45	94

## Step 7:

- Remove duplicates from the data using duplicate option from home tab

	Query	Manage Columns	Reduce Rows	Sort	Transform	Combine	New Query
Properties	Advanced Editor	Choose Columns	Remove Columns	Keep Rows	Remove Rows	Remove Duplicates	Remove Errors
Split Column	Group By	Replace Values	Use First Row As Headers	Merge Queries	Append Queries	Combine Binaries	New Source
Refresh Preview							Recent Sources

	Title	Director	Year	Runtime (Minutes)	Rating	Votes	Revenue (Millions)	Metascore
1	The Dark Knight	Christopher Nolan	2008	152	9	1791916	533.32	82
2	Inception	Christopher Nolan	2010	148	8.8	1583625	292.57	74
3	The Intouchables	Olivier Nakache	2011	112	8.6	557965	13.18	57
4	Kimi no na wa	Makoto Shinkai	2016	106	8.6	34110	4.68	79
5	Interstellar	Christopher Nolan	2014	169	8.6	1047747	187.99	74
6	Taare Zameen Par	Aamir Khan	2007	165	8.5	102697	1.2	42
7	The Prestige	Christopher Nolan	2006	130	8.5	913152	53.08	66
8	The Departed	Martin Scorsese	2006	151	8.5	937414	132.37	85
9	The Lives of Others	Florian Henckel von Donnersmarck	2006	137	8.5	278103	11.28	89
10	Whiplash	Damien Chazelle	2014	107	8.5	477276	13.09	88
11	The Dark Knight Rises	Christopher Nolan	2012	164	8.5	1222645	448.13	78
12	3 Idiots	Rajkumar Hirani	2009	170	8.4	238789	6.52	67
13	Django Unchained	Quentin Tarantino	2012	165	8.4	1039115	162.8	81
14	Jagten	Thomas Vinterberg	2012	115	8.3	192263	0.61	76
15	Toy Story 3	Lee Unkrich	2010	103	8.3	586669	414.98	92
16	Inglourious Basterds	Quentin Tarantino	2009	153	8.3	959065	120.52	69
17	La La Land	Damien Chazelle	2016	128	8.3	258682	151.06	93
18	Up	Pete Docter	2009	96	8.3	722203	292.98	88
19	Incendies	Denis Villeneuve	2010	131	8.2	92863	6.86	80
20	El secreto de sus ojos	Juan Jos�� Campanella	2009	129	8.2	144524	20.17	80
21	Pan's Labyrinth	Guillermo del Toro	2006	118	8.2	498879	37.62	98
22	PK	Rajkumar Hirani	2014	153	8.2	103279	10.57	51
23	Warrior	Gavin O'Connor	2011	140	8.2	355722	13.65	71
24	Hacksaw Ridge	Mel Gibson	2016	139	8.2	211760	67.12	71
25	The Wolf of Wall Street	Martin Scorsese	2013	180	8.2	865134	116.87	75

COLUMNS: 9, 837 ROWS

Finally, after cleaning the data, the final dataset sample is shown below:



Title	Director	Year	Runtime (Minutes)	Rating	Votes	Revenue (Millions)	Metascore
The Dark Knight	Christopher Nolan	2008	31.00:00:00	9	1791916	533.32	82
Inception	Christopher Nolan	2010	27.00:00:00	8.8	1583625	292.57	74
Interstellar	Christopher Nolan	2014	17.00:00:00	8.6	1047747	187.99	74
Kimi no na wa	Makoto Shinkai	2016	15.00:00:00	8.6	34110	4.68	79
The Intouchables	Olivier Nakache	2011	21.00:00:00	8.6	557965	13.18	57
The Prestige	Christopher Nolan	2006	9.00:00:00	8.5	913152	53.08	66
Taare Zameen Par	Aamir Khan	2007	13.00:00:00	8.5	102697	1.2	42
The Dark Knight Rises	Christopher Nolan	2012	12.00:00:00	8.5	1222645	448.13	78
Whiplash	Damien Chazelle	2014	16.00:00:00	8.5	477276	13.09	88
The Lives of Others	Florian Henckel von Donnersmarck	2006	16.00:00:00	8.5	278103	11.28	89
The Departed	Martin Scorsese	2006	30.00:00:00	8.5	937414	132.37	85
Django Unchained	Quentin Tarantino	2012	13.00:00:00	8.4	1039115	162.8	81
3 Idiots	Rajkumar Hirani	2009	18.00:00:00	8.4	238789	6.52	67
Koe no katachi	Naoko Yamada	2016	8.00:00:00	8.4	2421	6.98	80
Up	Pete Docter	2009	5.00:00:00	8.3	722203	292.98	88
Inglourious Basterds	Quentin Tarantino	2009	1.00:00:00	8.3	959065	120.52	69
Toy Story 3	Lee Unkrich	2010	12.00:00:00	8.3	586669	414.98	92
Jagten	Thomas Vinterberg	2012	24.00:00:00	8.3	192263	0.61	76
La La Land	Damien Chazelle	2016	7.00:00:00	8.3	258682	151.06	93
Paint It Black	Amber Tamblyn	2016	5.00:00:00	8.3	61	131.1	71
PK	Rajkumar Hirani	2014	1.00:00:00	8.2	103279	10.57	51
Room	Lenny Abrahamson	2015	27.00:00:00	8.2	224132	14.68	86
Pan's Labyrinth	Guillermo del Toro	2006	27.00:00:00	8.2	498879	37.62	98
Warrior	Gavin O'Connor	2011	19.00:00:00	8.2	355722	13.65	71
Hacksaw Ridge	Mel Gibson	2016	18.00:00:00	8.2	211760	67.12	71
Incendies	Denis Villeneuve	2010	10.00:00:00	8.2	92863	6.86	80
El secreto de sus ojos	Juan José Campanella	2009	8.00:00:00	8.2	144524	20.17	80
The Wolf of Wall Street	Martin Scorsese	2013	28.00:00:00	8.2	865134	116.87	75

## Analysis on dataset

### 1. Most Rating Movies

#### Introduction

- By performing this analysis, we will Sum of Rating earn by movies from 2006 to 2016.

#### Description

- The analysis based on the year, revenue of the dataset

#### Specific requirements, functions and formulas

- Pivot table is used for the analysis.
- Sum function is used in pivot table for the sum of the rating in the pivot table

#### Analysis results

Row Labels	Sum of Rating
Inception	8.8
Interstellar	8.6
Kimi no na wa	8.6
The Dark Knight	9
The Intouchables	8.6
<b>Grand Total</b>	<b>43.6</b>

#### Visualization



## 1. Most Metascore Movies

### Introduction

- ☐ By performing this analysis, we will Sum of Rating earn by movies from 2006 to 2016.

### Description

- ☐ The analysis based on the year, revenue of the dataset

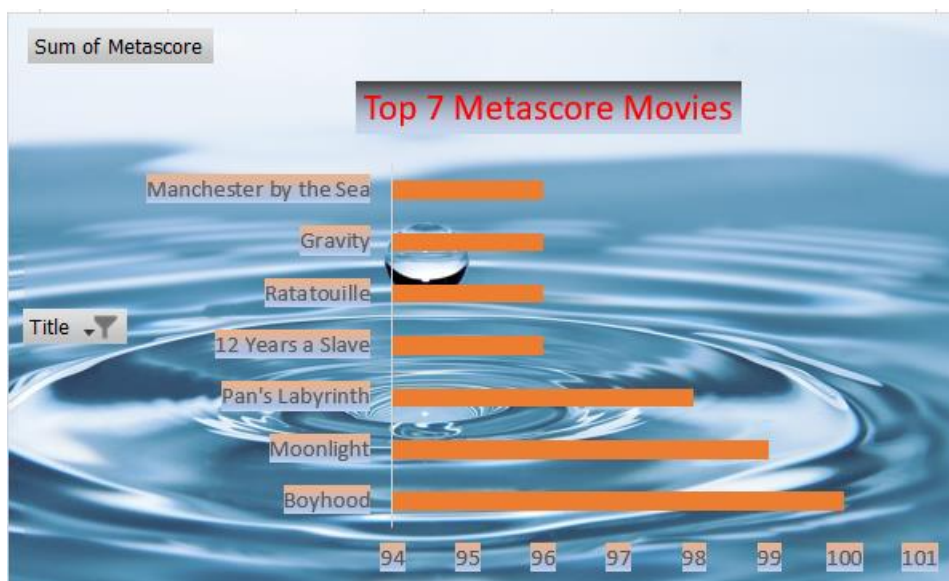
### Specific requirements, functions and formulas

- ☐ Pivot table is used for the analysis.
- ☐ Sum function is used in pivot table for the sum of the rating in the pivot table

### Analysis results

Row Labels	Sum of Metascore
Boyhood	100
Moonlight	99
Pan's Labyrinth	98
12 Years a Slave	96
Ratatouille	96
Gravity	96
Manchester by the Sea	96
<b>Grand Total</b>	<b>681</b>

## Visualization



### 3. Total revenues(millions) in each year

#### Introduction

- ☐ By performing this analysis, we will Total revenues earn by movies from 2006 to 2016.

#### Description

- ☐ The analysis based on the year, revenue of the dataset

#### Specific requirements, functions and formulas

- ☐ Pivot table is used for the analysis.
- ☐ Sum function is used in pivot table for the sum of the revenues in the pivot table

#### Analysis results

Row Labels	Sum of Revenue (Millions)
2006	3575.28
2007	4095.26
2008	4681.22
2009	5208.39
2010	5989.65
2011	5341.09
2012	6826.39
2013	7399.11
2014	7945.33
2015	8799.09
2016	11002.09
<b>Grand Total</b>	<b>70862.9</b>

## Visualization



### 4. Total Movies released in each year

#### Introduction

- ☐ By performing this analysis, we will get Number of movies released in each year

#### Description

- ☐ The analysis based on the year, title of the dataset

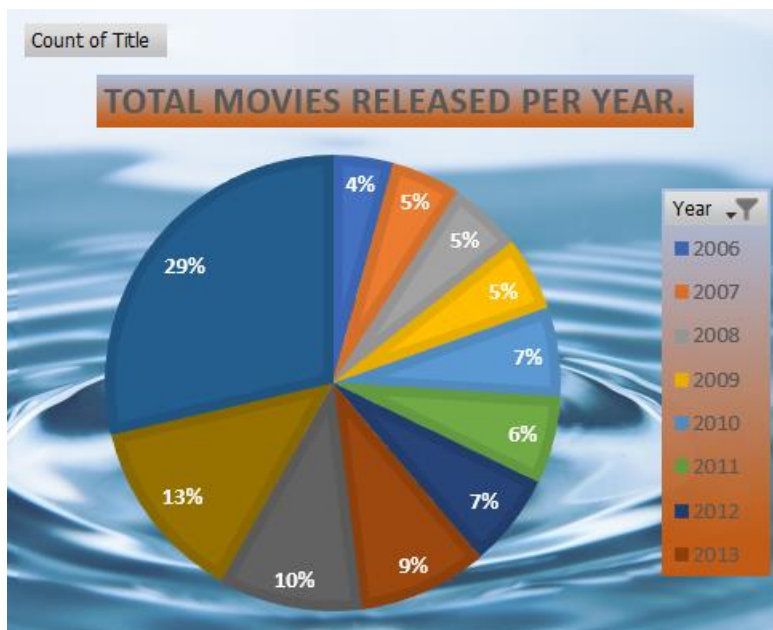
## Specific requirements, functions and formulas

- ☐ Pivot table is used for the analysis.
- ☐ Count function is used in pivot table for the counting of the title in the pivot table in their respective year

## Analysis results

Row Labels	Count of Title
2006	40
2007	46
2008	49
2009	49
2010	59
2011	58
2012	62
2013	86
2014	95
2015	123
2016	268
<b>Grand Total</b>	<b>935</b>

## Visualization



## 5. Top 5 Revenue Movies

### Introduction

- ☐ By performing this analysis, we will get Top 5 revenue movies of all time or selected year in the slicer

### Description

- ☐ The analysis is based on title, revenue of the dataset
- ☐ Filter is used to show top 5 movies
- ☐ Slicer is connected with the graph to get each year top 5 revenue movies

### Specific requirements, functions and formulas

- ☐ Pivot table is used for the analysis.
- ☐ Sum function is used in pivot table for the sum of revenue of the title

### Analysis results

Row Labels	Sum of Revenue (Millions)
The Dark Knight	533.32
The Avengers	623.28
Star Wars: Episode VII - The Force Awakens	936.63
Jurassic World	652.18
Avatar	760.51
<b>Grand Total</b>	<b>3505.92</b>



## Visualization



## 6. Most voted Director

### Introduction

- ☐ By performing this analysis, we will get Top 5 Most voted director from the dataset

### Description

- ☐ The analysis is based on Director, Votes of the dataset
- ☐ Filter is used to show top 5 most voted director
- ☐ Slicer is connected with the graph to get each year top 5 most voted director

## Specific requirements, functions and formulas

- ☐ Pivot table is used for the analysis.
- ☐ Sum function is used in pivot table for the sum of votes for the correspondence director.

## Analysis results

Row Labels	Sum of Votes
Christopher Nolan	6559085
David Fincher	2309652
David Yates	2033854
J.J. Abrams	2174937
Martin Scorsese	2966524
Michael Bay	1706531
Peter Jackson	1698695
Quentin Tarantino	2339350
Ridley Scott	2080074
Zack Snyder	2301544
<b>Grand Total</b>	<b>26170246</b>

## Visualization



## 7. Top 5 Longest Movies

### Introduction

- ☐ By performing this analysis, we will get Top 5 longest movies of all time or selected year in the slicer

### Description

- ☐ The analysis is based on title, runtimes of the dataset
- ☐ Filter is used to show top 5 movies
- ☐ Slicer is connected with the graph to get each year top 5 revenue movies

### Specific requirements, functions and formulas

- ☐ Pivot table is used for the analysis.
- ☐ Sum function is used in pivot table for the sum of runtimes of the title

### Analysis results

Row Labels	Sum of Runtime (Minutes)
The Hateful Eight	187
The Wolf of Wall Street	180
La vie d'Adèle	180
Cloud Atlas	172
3 Idiots	170
<b>Grand Total</b>	<b>889</b>

Visualization



## List of Analysis with results

### 1. Top 5 rating movies

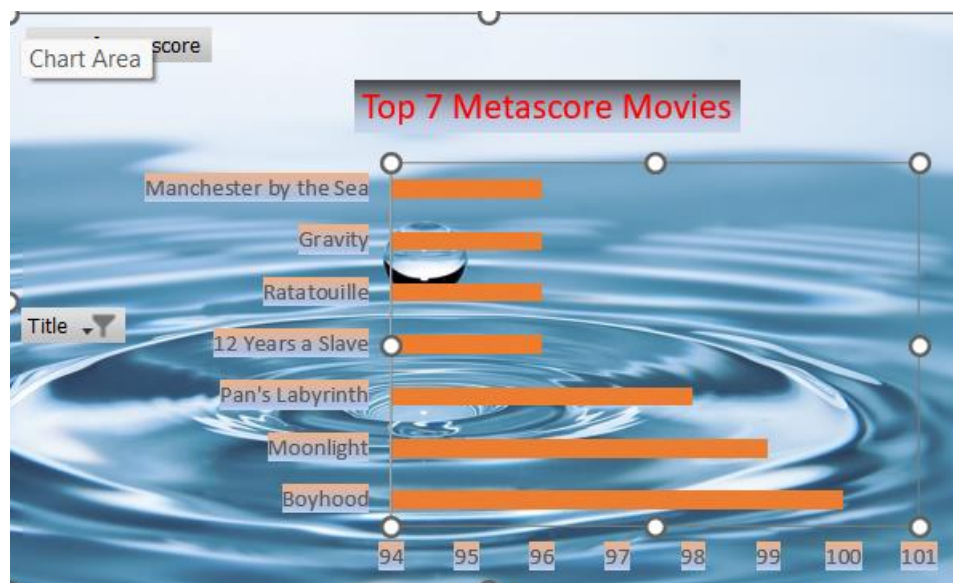
□ Top 5 rating movies:

- 1) The Dark knight
- 2) Inception
- 3) The Intouchables
- 4) Interstellar
- 5) Kimi no na wa



## 2. □ Top 7 Metascores movies:

- 1) Boyhood
- 2) Moonlight
- 3) Pan's Labyrinth
- 4) 12 Years a slave
- 5) Ratatouille
- 6) Manchester by the Sea
- 7) Gravity



### 3. Total revenues(millions) in each year

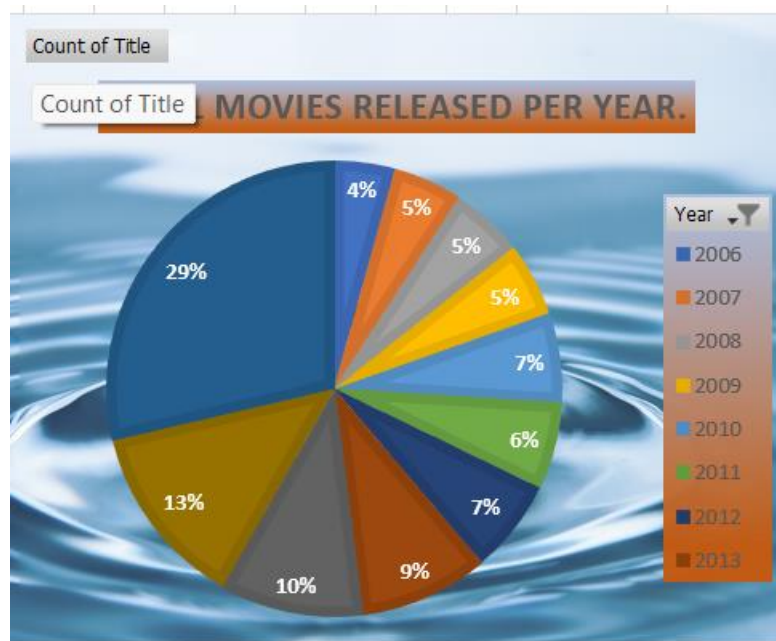
- The Highest Revenue was recorded in 2016 with total \$11002.09 million. Whereas the lowest Revenue was recorded in 2006 with total \$3575.28 million.





#### 4. Total Movies released in each year

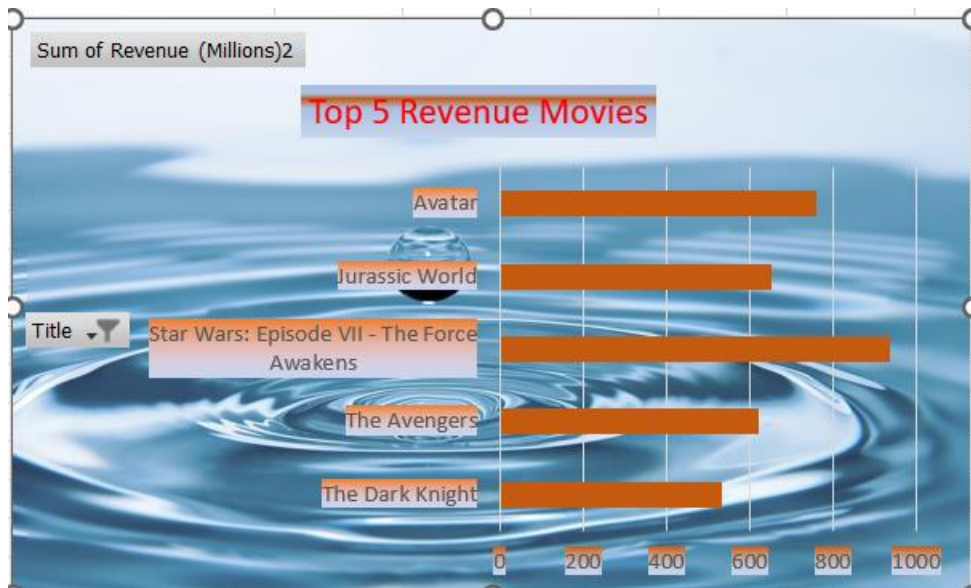
- ☐ The highest number of movies was released in 2016 with total 198 movies.
- ☐ While, the lowest number of movies was released in 2006 with total 40 movies.



#### 5. Top 5 Revenue Movies

- ☐ The top 5 revenues are given below:
  - 1) Star Wars: Episode VII – The Force Awakens
  - 2) Avatar
  - 3) Jurassic World
  - 4) The Avengers
  - 5) The Dark Knight





## 6. Most voted Director

- ☐ The Most Voted Director in the dataset is Christopher Nolan. ☐
- After that, the second most voted Director is Martin Scorsese



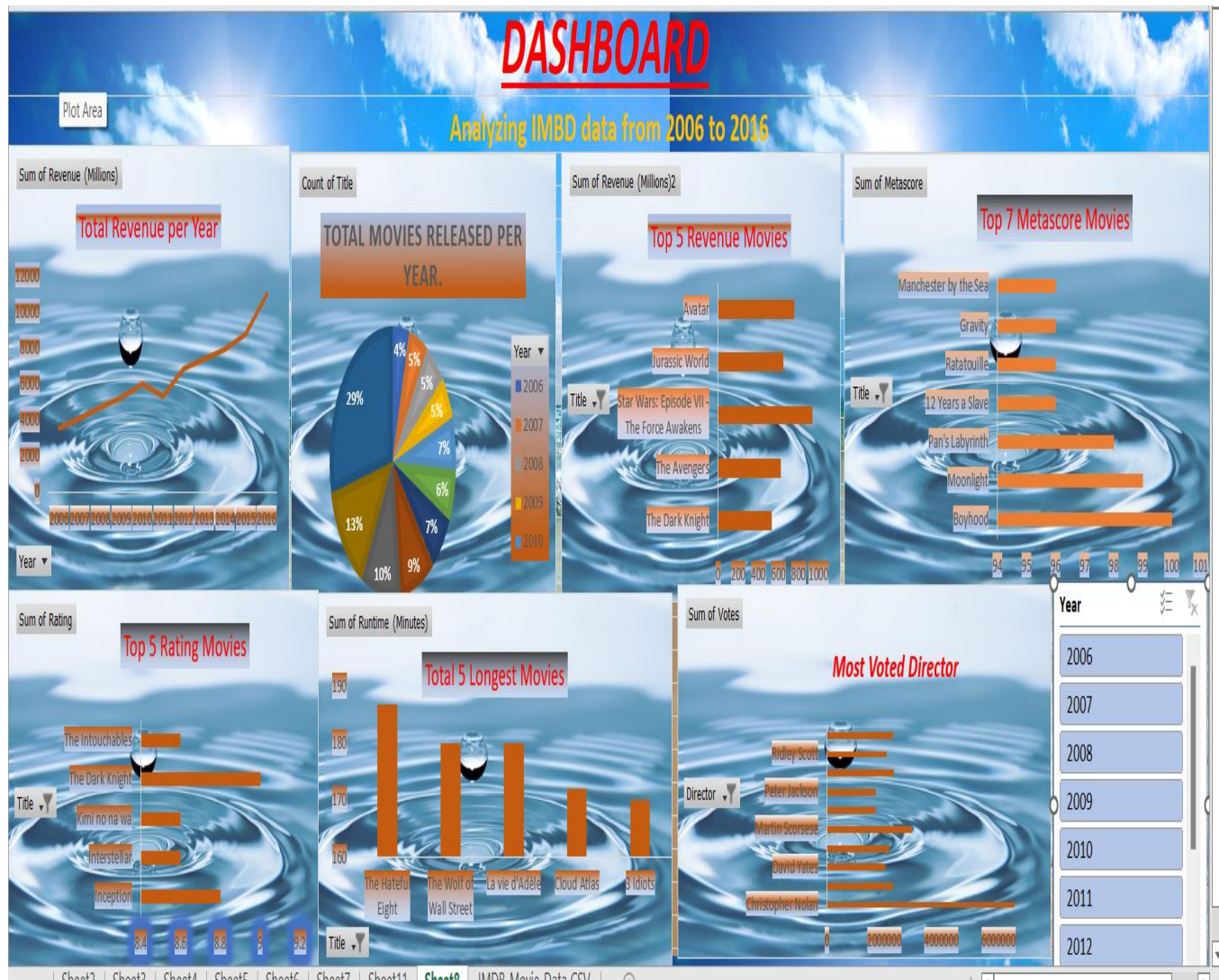
## 7. Top 5 Longest Movies

□ The top 5 longest movies are given below:

- 1) The Hateful Eight
- 2) The Wolf of Wall Street
- 3) La Vie d'Adèle
- 4) Cloud Atlas
- 5) 3 Idiots



FINAL DASHBOARD:



## BIBLIOGRAPHY:

□ Dataset source:

<https://www.kaggle.com/PromptCloudHQ/imdb-data>

## □ Dashboard Background Image:

[https://hubblecontent.osi.office.net/contentsvc/ui/index.html?picker=1&fb=1&p=0&t=0&a=1&idp=live&lid=2&taskpane=0&ht=0&license=Perpetual2019&lcid=1033&syslcid=16393&uilcid=1033&app=1&ver=16&tl=2&build=16.0.15726&searchtype=Image&cfd=1&moss=0&ins=1&albm=0&eurl=0&mset=0&sl=all&ontforcehomeurl=1&setlang=en-US&pi=1&style=0&correlationid=b7e15125-ee12-40c8-8291-c036c6356f48&uid=eee4cd33305f92e0\\_LiveId&cw=683&ch=402&cp=bing&comcall=1&fparty=1&mkt=en-IN](https://hubblecontent.osi.office.net/contentsvc/ui/index.html?picker=1&fb=1&p=0&t=0&a=1&idp=live&lid=2&taskpane=0&ht=0&license=Perpetual2019&lcid=1033&syslcid=16393&uilcid=1033&app=1&ver=16&tl=2&build=16.0.15726&searchtype=Image&cfd=1&moss=0&ins=1&albm=0&eurl=0&mset=0&sl=all&ontforcehomeurl=1&setlang=en-US&pi=1&style=0&correlationid=b7e15125-ee12-40c8-8291-c036c6356f48&uid=eee4cd33305f92e0_LiveId&cw=683&ch=402&cp=bing&comcall=1&fparty=1&mkt=en-IN)