



IMDB Movie Analysis

Santhosh R



Contents

- Introduction
- Data Cleaning Process
- Analysis Data

Here, we are analysis the Movies
related data the data we are get from
the IMDB

IMDB_Movies - Excel (Product Activation Failed)																
FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW DEVELOPER																
<div> <div> Cut Copy Paste </div> <div> Format Painter </div> <div> Clipboard </div> </div> <div> <div>Calibri 10 A</div> <div> B I U </div> <div> Font </div> </div> <div> <div> Merge & Center </div> <div> Alignment </div> </div> <div> <div> \$ % </div> <div> Number </div> </div> <div> <div> Conditional Formatting </div> <div> Styles </div> </div> <div> <div> Cell Styles </div> <div> Cells </div> </div> <div> <div> AutoSum </div> <div> Fill </div> <div> Clear </div> </div> <div> <div> Sort & Find </div> <div> Filter & Select </div> </div>																

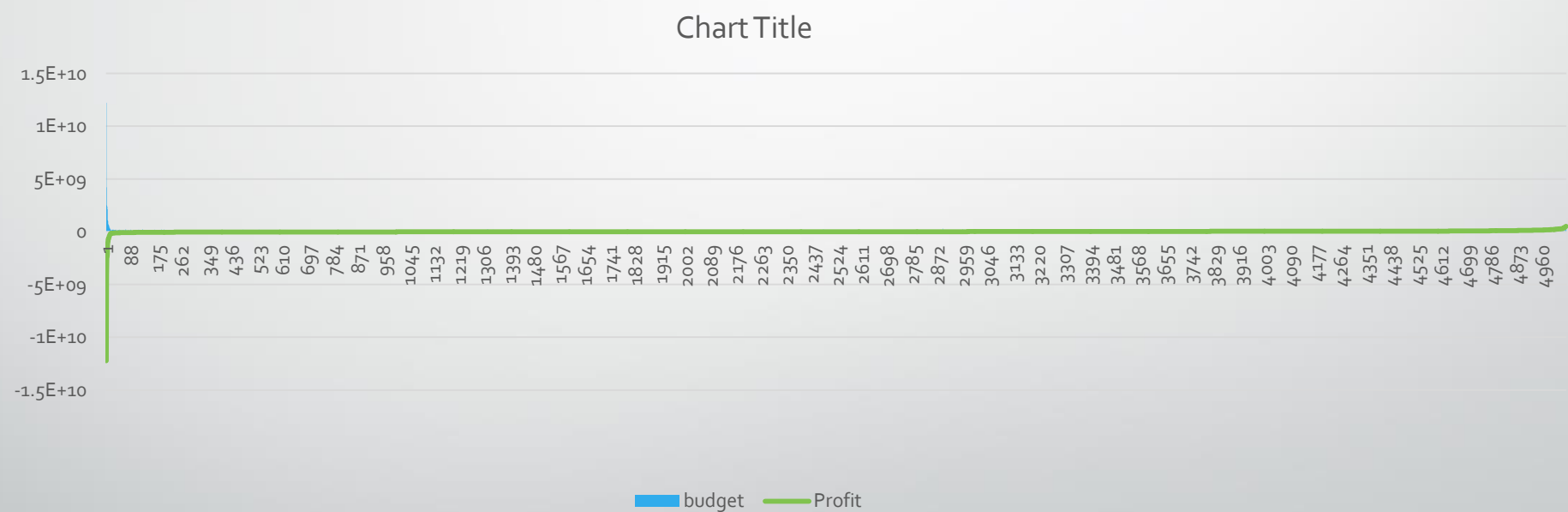
																P30																
	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V																
	actor_2_name	actor_1_name	genres	actor_k_name	movie_title	num_voted_us	east_total_1act	actor_1_name	facenumber_1	plot_keywords	movie_imdb_link	user_for_r	language	country	content_rating	budget																
1	actor_2_name	actor_1_name	genres	actor_k_name	movie_title	num_voted_us	east_total_1act	actor_1_name	facenumber_1	plot_keywords	movie_imdb_link	user_for_r	language	country	content_rating	budget																
2	Joel David Moore	9000	76050547	Action/Adventure	CCH Pounder	Avatar	886204	4534	Yes Studi	0	avatarurtema	<http://www.imdb.com/title/tt0417757>	3054	English	USA	PG-13	237000000	2009														
3	Orlando Bloom	40000	30944462	Action/Adventure	Johnny Depp	Pirates of the C	471220	45360	Jack Davenport	0	godessmarina	<http://www.imdb.com/title/tt0417757>	1238	English	UK	PG-13	300000000	2007														
4	Poo Kishor	10000	20004175	Action/Adventure	Christophe Yau	275986	1700	Stephanie Sign	0	deceptionimpr	<http://www.imdb.com/title/tt0417757>	2704	English	UK	PG-13	245000000	2016															
5	Christian Bale	27000	44813062	Action/Thriller	Tom Hardy	The Dark Knight	1144337	106759	Joseph Gordon	0	deceptionimpr	<http://www.imdb.com/title/tt0417757>	2704	English	USA	PG-13	250000000	2012														
6	Rob Walker	131		Documentary	Doug Walker	Star Wars: Episo	8	143	0	http://www.imdb.com/title/tt2383954?ref=fn_tt_1																						
7	Samantha Morton	640	73059879	Action/Adventure	Daryl Sabara	John Carter	212204	773	Polly Walker	0	alienamerica	<http://www.imdb.com/title/tt0417757>	738	English	USA	PG-13	263700000	2012														
8	James Franco	24000	33653003	Action/Adventure	Ed J. Simmons	Spider-Man 3	283056	46055	Kirsten Dunst	0	sandmanpider	<http://www.imdb.com/title/tt0417757>	1902	English	USA	PG-13	258000000	2007														
9	Donna Murphy	799	200807262	Adventure/Exam	Brad Garrett	Tangled	349094	2036	M.C. Gainey	1	17th centurybas	<http://www.imdb.com/title/tt0417757>	387	English	USA	PG	260000000	2010														
10	Robert Downey Jr.	26000	45899599	Action/Adventure	Chris Hemsworth	The Avengers	462669	92000	Scarlett Johans	4	artificial intellig	<http://www.imdb.com/title/tt0417757>	1117	English	USA	PG-13	250000000	2015														
11	Daniel Radcliffe	25000	30195680	Adventure/Famil	Alan Rickman	Harry Potter and	321795	58753	Rupert Grint	3	bookbooklove	<http://www.imdb.com/title/tt0417757>	973	English	UK	PG	250000000	2009														
12	Leven Coaker	50000	33024362	Action/Adventure	Henry Cavill	Batman v Super	376383	24450	Alan D. Forwin	0	based on comic	<http://www.imdb.com/title/tt0417757>	3018	English	USA	PG-13	250000000	2016														
13	Malin Akerman	80000	200030406	Action/Adventure	Kevin Spacey	Superman Rem	240296	23991	Frank Langella	0	originalseries	<http://www.imdb.com/title/tt0417757>	2367	English	USA	PG-13	209000000	2006														
14	Manju Manjula	451	86369427	Action/Adventure	Giancarlo Giam	Quantum of Sol	220794	2032	Rory Kinnear	1	action heroact	<http://www.imdb.com/title/tt0417757>	1248	English	UK	PG-13	200000000	2008														
15	Orlando Bloom	40000	423032628	Action/Adventure	Johnny Depp	Pirates of the C	522040	48406	Jack Davenport	2	box office hitra	[http://www.im](http://www.imdb.com/title/tt0417757)																				

Data Cleaning

- Before using given Raw Data we should clean. Because may be the given Data set having empty cell values. Sometime values not present.
- Here, I am clean the numerical data set by using the method of Average, Mode, Mean.
- Categorical Data column if null values present I drop that entire row from the data set

[illegible]

Movie Profit graph



Top 250

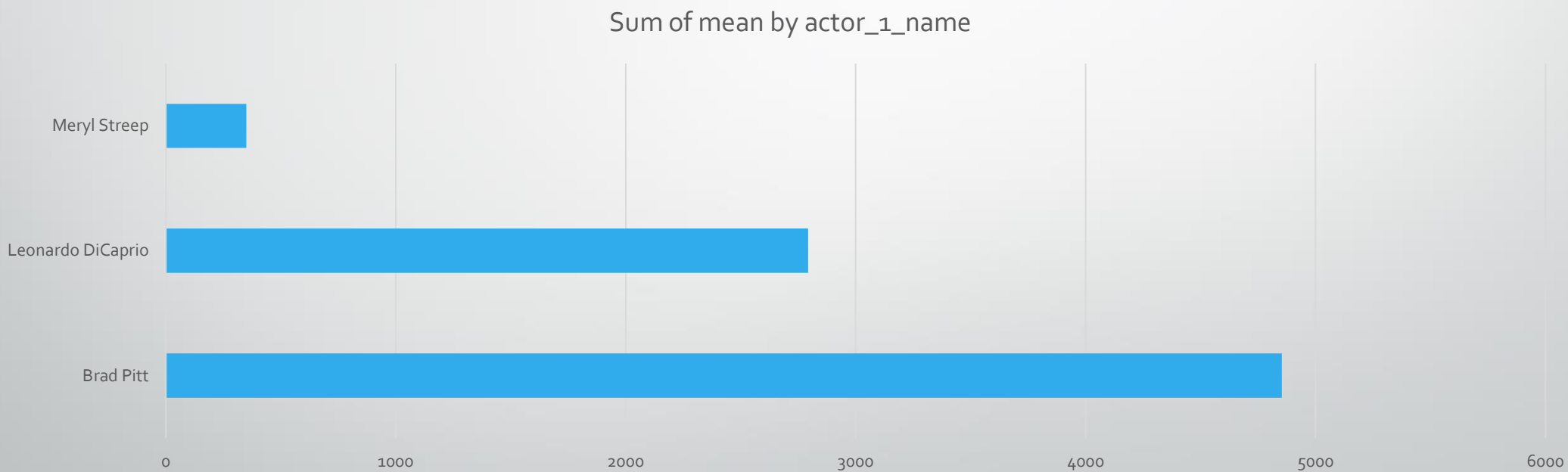
IMDb_Top_250	Rank
The Shawshank RedemptionÂ	1
The GodfatherÂ	2
FargoÂ	3
The Godfather: Part IIÂ	4
The Dark KnightÂ	5
12 Angry MenÂ	6
The Good, the Bad and the UglyÂ	7
Schindler's ListÂ	8
The Lord of the Rings: The Return of the KingÂ	9
Pulp FictionÂ	10
It's Always Sunny in PhiladelphiaÂ	11
DaredevilÂ	12
Star Wars: Episode V - The Empire Strikes BackÂ	13
The Lord of the Rings: The Fellowship of the RingÂ	14
Forrest GumpÂ	15
Fight ClubÂ	16
InceptionÂ	17
Friday Night LightsÂ	18
Seven SamuraiÂ	19
City of GodÂ	20
One Flew Over the Cuckoo's NestÂ	21
GoodfellasÂ	22
Star Wars: Episode IV - A New HopeÂ	23
The Lord of the Rings: The Two TowersÂ	24
The MatrixÂ	25
LutherÂ	26
Modern TimesÂ	27
HannibalÂ	28
Spartacus: War of the DamnedÂ	29
Once Upon a Time in the WestÂ	30

Best Directors

Top_10_Director_name
John Blanchard
Frank Darabont
Francis Ford Coppola
John Stockwell
Christopher Nolan
Francis Ford Coppola
Peter Jackson
Quentin Tarantino
Sergio Leone
Steven Spielberg

- By using the Pivot chart I got the top 10 Best Directors base on rating point
- Action|Adventure|Crime|Drama|Sci-Fi|Thriller - Popular Genres

Charts





Thank You