In [1]:

*#Project:*

*#TMDb movies Dataset*

*# Research Qustion :*

*#created 5 research questions associated with the director as a target to be answered:*

*#1- Which most popular genres movies? Drama &Comedy.*

*#2- Which director has most popularity?Colin*

*#3- Which director is having the highest budget? Sngmoo Lee*

*#4- Which director has most vote count? Pop &Pet*

*#5- Which director has the most vote average?Mark Cousins*

*#Investigation and Wrangling Documentation:*

*#• Data wrangling and cleaning Data*

*#1. Importing the packages and uploading the data file into Jupiter. (tmdb-movies(8).csv).*

*#2. Importing statements for packages to be used.*

*#3. Counting the number of unique values in our data for each data type*

*#4. One duplicate record is present and deleted.*

*#5. Dropping 10 missing values*

*#conclusion:*

*#Summary statistics and plots communicating the final results.*

*#1-from below analaysis i conclued the the 5 variable showing wide range between year Releases, popularity, budget, vote count and vote average.*

*#2-Isee correlation between the popularity and the revenue for the directors espacily for lower value .*

*#3-Mark Cousins got the most vote average and not leaner with popularity, release, budget and vote count .*

*#4-the vote count goes to Bob Peterson|Pete Docter and again not leaner with popularity , releases and budget.*

*#5-Sngmoo Lee showing the highest budget was refer to but not leaner with popularity and releases.*

*#6-Colin Trevorrow is most popular but he didn't have the most realeases.so the popularity and biger anumber of realease dindnt go in leanear direction*

*#7-from this data found most realease is belong to woody Allen but that amount of releases is not linear with popularity for Colin Trevorrow .*

*#limitations.*

*# Not able to show all the directors names on the x-axis labeldue to the high count of director names.*

*# There are no missing values for any of the factors under consideration, but the data did not consider all elements and aspect for example the data set is not showing every single release for every director*

*# I have applied the dollar currency because the data is not showing which currency applied to the revenue or the budget for the movies that created in different regions.*

*# Usually for the movie revenue there are more sources to get more revenue like ticket sales in theaters, distributors, selling pay-TV, overseas rights and creating video games and so on.*

*# Some famous directors are having more releases within one year and over the years than new directors which will result of suddenly change direction or position.*

*#Resources:*

*#1- https://pandas.pydata.org/pandas-docs/stable/10min.html*

*#2- https://www.datacamp.com/community/blog/python-numpy-cheat-sheet*

*#3- http://www.developintelligence.com/blog/2017/08/data-cleaning-pandas-python/*

*#4- https://s3.amazonaws.com/assets.datacamp.com/production/course\_3485/slides/ch1\_slides.p*

*#5- https://github.com/nirupamaprv/Investigate\_Dataset/blob/master/Investigate\_a\_Dataset\_TMDb\_Directors\_NirupamaPV.ipynb*

*#6- https://docs.google.com/document/d/e/2PACX-1vTlVmknRRnfy\_4eTrjw5hYGaiQim5ctr9naaRd4V9du2B5bxpd8FEH3KtDgp8qVekw7Cj1GLk1IXdZi/pub?embedded=True*

In [2]:

*#import statements for packages to be used*

%matplotlib inline

**import** **pandas** **as** **pd**

**import** **csv** *# read CSV (Comma Separated Values) format*

**from** **datetime** **import** datetime *# only importing the datetime class within that file.*

**import** **matplotlib.pyplot** **as** **plt** *#importing the module "matplotlib.pyplot" and binds that to the name "plt".*

In [3]:

*#Importing the packages and uploading the data file into Jupiter. (tmdb-movies(8).csv).*

df = pd.read\_csv('tmdb-movies.csv', encoding='ISO-8859-1')

df.info()

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 10866 entries, 0 to 10865

Data columns (total 21 columns):

id 10866 non-null int64

imdb\_id 10856 non-null object

popularity 10866 non-null float64

budget 10866 non-null int64

revenue 10866 non-null int64

original\_title 10866 non-null object

cast 10790 non-null object

homepage 2936 non-null object

director 10822 non-null object

tagline 8042 non-null object

keywords 9373 non-null object

overview 10862 non-null object

runtime 10866 non-null int64

genres 10843 non-null object

production\_companies 9836 non-null object

release\_date 10866 non-null object

vote\_count 10866 non-null int64

vote\_average 10866 non-null float64

release\_year 10866 non-null int64

budget\_adj 10866 non-null float64

revenue\_adj 10866 non-null float64

dtypes: float64(4), int64(6), object(11)

memory usage: 1.7+ MB

In [4]:

*# reading the csv data set file.*

df

Out[4]:

|  | **id** | **imdb\_id** | **popularity** | **budget** | **revenue** | **original\_title** | **cast** | **homepage** | **director** | **tagline** | **...** | **overview** | **runtime** | **genres** | **production\_companies** | **release\_date** | **vote\_count** | **vote\_average** | **release\_year** | **budget\_adj** | **revenue\_adj** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | 135397 | tt0369610 | 32.985763 | 150000000 | 1513528810 | Jurassic World | Chris Pratt|Bryce Dallas Howard|Irrfan Khan|Vi... | http://www.jurassicworld.com/ | Colin Trevorrow | The park is open. | ... | Twenty-two years after the events of Jurassic ... | 124 | Action|Adventure|Science Fiction|Thriller | Universal Studios|Amblin Entertainment|Legenda... | 6/9/2015 | 5562 | 6.5 | 2015 | 1.379999e+08 | 1.392446e+09 |
| **1** | 76341 | tt1392190 | 28.419936 | 150000000 | 378436354 | Mad Max: Fury Road | Tom Hardy|Charlize Theron|Hugh Keays-Byrne|Nic... | http://www.madmaxmovie.com/ | George Miller | What a Lovely Day. | ... | An apocalyptic story set in the furthest reach... | 120 | Action|Adventure|Science Fiction|Thriller | Village Roadshow Pictures|Kennedy Miller Produ... | 5/13/2015 | 6185 | 7.1 | 2015 | 1.379999e+08 | 3.481613e+08 |
| **2** | 262500 | tt2908446 | 13.112507 | 110000000 | 295238201 | Insurgent | Shailene Woodley|Theo James|Kate Winslet|Ansel... | http://www.thedivergentseries.movie/#insurgent | Robert Schwentke | One Choice Can Destroy You | ... | Beatrice Prior must confront her inner demons ... | 119 | Adventure|Science Fiction|Thriller | Summit Entertainment|Mandeville Films|Red Wago... | 3/18/2015 | 2480 | 6.3 | 2015 | 1.012000e+08 | 2.716190e+08 |
| **3** | 140607 | tt2488496 | 11.173104 | 200000000 | 2068178225 | Star Wars: The Force Awakens | Harrison Ford|Mark Hamill|Carrie Fisher|Adam D... | http://www.starwars.com/films/star-wars-episod... | J.J. Abrams | Every generation has a story. | ... | Thirty years after defeating the Galactic Empi... | 136 | Action|Adventure|Science Fiction|Fantasy | Lucasfilm|Truenorth Productions|Bad Robot | 12/15/2015 | 5292 | 7.5 | 2015 | 1.839999e+08 | 1.902723e+09 |
| **4** | 168259 | tt2820852 | 9.335014 | 190000000 | 1506249360 | Furious 7 | Vin Diesel|Paul Walker|Jason Statham|Michelle ... | http://www.furious7.com/ | James Wan | Vengeance Hits Home | ... | Deckard Shaw seeks revenge against Dominic Tor... | 137 | Action|Crime|Thriller | Universal Pictures|Original Film|Media Rights ... | 4/1/2015 | 2947 | 7.3 | 2015 | 1.747999e+08 | 1.385749e+09 |
| **5** | 281957 | tt1663202 | 9.110700 | 135000000 | 532950503 | The Revenant | Leonardo DiCaprio|Tom Hardy|Will Poulter|Domhn... | http://www.foxmovies.com/movies/the-revenant | Alejandro GonzÃÂ¡lez IÃÂ±ÃÂ¡rritu | (n. One who has returned, as if from the dead.) | ... | In the 1820s, a frontiersman, Hugh Glass, sets... | 156 | Western|Drama|Adventure|Thriller | Regency Enterprises|Appian Way|CatchPlay|Anony... | 12/25/2015 | 3929 | 7.2 | 2015 | 1.241999e+08 | 4.903142e+08 |
| **6** | 87101 | tt1340138 | 8.654359 | 155000000 | 440603537 | Terminator Genisys | Arnold Schwarzenegger|Jason Clarke|Emilia Clar... | http://www.terminatormovie.com/ | Alan Taylor | Reset the future | ... | The year is 2029. John Connor, leader of the r... | 125 | Science Fiction|Action|Thriller|Adventure | Paramount Pictures|Skydance Productions | 6/23/2015 | 2598 | 5.8 | 2015 | 1.425999e+08 | 4.053551e+08 |
| **7** | 286217 | tt3659388 | 7.667400 | 108000000 | 595380321 | The Martian | Matt Damon|Jessica Chastain|Kristen Wiig|Jeff ... | http://www.foxmovies.com/movies/the-martian | Ridley Scott | Bring Him Home | ... | During a manned mission to Mars, Astronaut Mar... | 141 | Drama|Adventure|Science Fiction | Twentieth Century Fox Film Corporation|Scott F... | 9/30/2015 | 4572 | 7.6 | 2015 | 9.935996e+07 | 5.477497e+08 |
| **8** | 211672 | tt2293640 | 7.404165 | 74000000 | 1156730962 | Minions | Sandra Bullock|Jon Hamm|Michael Keaton|Allison... | http://www.minionsmovie.com/ | Kyle Balda|Pierre Coffin | Before Gru, they had a history of bad bosses | ... | Minions Stuart, Kevin and Bob are recruited by... | 91 | Family|Animation|Adventure|Comedy | Universal Pictures|Illumination Entertainment | 6/17/2015 | 2893 | 6.5 | 2015 | 6.807997e+07 | 1.064192e+09 |
| **9** | 150540 | tt2096673 | 6.326804 | 175000000 | 853708609 | Inside Out | Amy Poehler|Phyllis Smith|Richard Kind|Bill Ha... | http://movies.disney.com/inside-out | Pete Docter | Meet the little voices inside your head. | ... | Growing up can be a bumpy road, and it's no ex... | 94 | Comedy|Animation|Family | Walt Disney Pictures|Pixar Animation Studios|W... | 6/9/2015 | 3935 | 8.0 | 2015 | 1.609999e+08 | 7.854116e+08 |
| **10** | 206647 | tt2379713 | 6.200282 | 245000000 | 880674609 | Spectre | Daniel Craig|Christoph Waltz|LÃÂ©a Seydoux|Ra... | http://www.sonypictures.com/movies/spectre/ | Sam Mendes | A Plan No One Escapes | ... | A cryptic message from BondÃ¢â¬â¢s past send... | 148 | Action|Adventure|Crime | Columbia Pictures|Danjaq|B24 | 10/26/2015 | 3254 | 6.2 | 2015 | 2.253999e+08 | 8.102203e+08 |
| **11** | 76757 | tt1617661 | 6.189369 | 176000003 | 183987723 | Jupiter Ascending | Mila Kunis|Channing Tatum|Sean Bean|Eddie Redm... | http://www.jupiterascending.com | Lana Wachowski|Lilly Wachowski | Expand your universe. | ... | In a universe where human genetic material is ... | 124 | Science Fiction|Fantasy|Action|Adventure | Village Roadshow Pictures|Dune Entertainment|A... | 2/4/2015 | 1937 | 5.2 | 2015 | 1.619199e+08 | 1.692686e+08 |
| **12** | 264660 | tt0470752 | 6.118847 | 15000000 | 36869414 | Ex Machina | Domhnall Gleeson|Alicia Vikander|Oscar Isaac|S... | http://exmachina-movie.com/ | Alex Garland | There is nothing more human than the will to s... | ... | Caleb, a 26 year old coder at the world's larg... | 108 | Drama|Science Fiction | DNA Films|Universal Pictures International (UP... | 1/21/2015 | 2854 | 7.6 | 2015 | 1.379999e+07 | 3.391985e+07 |
| **13** | 257344 | tt2120120 | 5.984995 | 88000000 | 243637091 | Pixels | Adam Sandler|Michelle Monaghan|Peter Dinklage|... | http://www.pixels-movie.com/ | Chris Columbus | Game On. | ... | Video game experts are recruited by the milita... | 105 | Action|Comedy|Science Fiction | Columbia Pictures|Happy Madison Productions | 7/16/2015 | 1575 | 5.8 | 2015 | 8.095996e+07 | 2.241460e+08 |
| **14** | 99861 | tt2395427 | 5.944927 | 280000000 | 1405035767 | Avengers: Age of Ultron | Robert Downey Jr.|Chris Hemsworth|Mark Ruffalo... | http://marvel.com/movies/movie/193/avengers\_ag... | Joss Whedon | A New Age Has Come. | ... | When Tony Stark tries to jumpstart a dormant p... | 141 | Action|Adventure|Science Fiction | Marvel Studios|Prime Focus|Revolution Sun Studios | 4/22/2015 | 4304 | 7.4 | 2015 | 2.575999e+08 | 1.292632e+09 |
| **15** | 273248 | tt3460252 | 5.898400 | 44000000 | 155760117 | The Hateful Eight | Samuel L. Jackson|Kurt Russell|Jennifer Jason ... | http://thehatefuleight.com/ | Quentin Tarantino | No one comes up here without a damn good reason. | ... | Bounty hunters seek shelter from a raging bliz... | 167 | Crime|Drama|Mystery|Western | Double Feature Films|The Weinstein Company|Fil... | 12/25/2015 | 2389 | 7.4 | 2015 | 4.047998e+07 | 1.432992e+08 |
| **16** | 260346 | tt2446042 | 5.749758 | 48000000 | 325771424 | Taken 3 | Liam Neeson|Forest Whitaker|Maggie Grace|Famke... | http://www.taken3movie.com/ | Olivier Megaton | It Ends Here | ... | Ex-government operative Bryan Mills finds his ... | 109 | Crime|Action|Thriller | Twentieth Century Fox Film Corporation|M6 Film... | 1/1/2015 | 1578 | 6.1 | 2015 | 4.415998e+07 | 2.997096e+08 |
| **17** | 102899 | tt0478970 | 5.573184 | 130000000 | 518602163 | Ant-Man | Paul Rudd|Michael Douglas|Evangeline Lilly|Cor... | http://marvel.com/movies/movie/180/ant-man | Peyton Reed | Heroes Don't Get Any Bigger | ... | Armed with the astonishing ability to shrink i... | 115 | Science Fiction|Action|Adventure | Marvel Studios | 7/14/2015 | 3779 | 7.0 | 2015 | 1.195999e+08 | 4.771138e+08 |
| **18** | 150689 | tt1661199 | 5.556818 | 95000000 | 542351353 | Cinderella | Lily James|Cate Blanchett|Richard Madden|Helen... | NaN | Kenneth Branagh | Midnight is just the beginning. | ... | When her father unexpectedly passes away, youn... | 112 | Romance|Fantasy|Family|Drama | Walt Disney Pictures|Genre Films|Beagle Pug Fi... | 3/12/2015 | 1495 | 6.8 | 2015 | 8.739996e+07 | 4.989630e+08 |
| **19** | 131634 | tt1951266 | 5.476958 | 160000000 | 650523427 | The Hunger Games: Mockingjay - Part 2 | Jennifer Lawrence|Josh Hutcherson|Liam Hemswor... | http://www.thehungergames.movie/ | Francis Lawrence | The fire will burn forever. | ... | With the nation of Panem in a full scale war, ... | 136 | War|Adventure|Science Fiction | Studio Babelsberg|StudioCanal|Lionsgate|Walt D... | 11/18/2015 | 2380 | 6.5 | 2015 | 1.471999e+08 | 5.984813e+08 |
| **20** | 158852 | tt1964418 | 5.462138 | 190000000 | 209035668 | Tomorrowland | Britt Robertson|George Clooney|Raffey Cassidy|... | http://movies.disney.com/tomorrowland | Brad Bird | Imagine a world where nothing is impossible. | ... | Bound by a shared destiny, a bright, optimisti... | 130 | Action|Family|Science Fiction|Adventure|Mystery | Walt Disney Pictures|Babieka|A113 | 5/19/2015 | 1899 | 6.2 | 2015 | 1.747999e+08 | 1.923127e+08 |
| **21** | 307081 | tt1798684 | 5.337064 | 30000000 | 91709827 | Southpaw | Jake Gyllenhaal|Rachel McAdams|Forest Whitaker... | NaN | Antoine Fuqua | Believe in Hope. | ... | Billy "The Great" Hope, the reigning junior mi... | 123 | Action|Drama | Escape Artists|Riche-Ludwig Productions | 6/15/2015 | 1386 | 7.3 | 2015 | 2.759999e+07 | 8.437300e+07 |
| **22** | 254128 | tt2126355 | 4.907832 | 110000000 | 470490832 | San Andreas | Dwayne Johnson|Alexandra Daddario|Carla Gugino... | http://www.sanandreasmovie.com/ | Brad Peyton | A rescue pilot survived an earthquake, this is... | ... | In the aftermath of a massive earthquake in Ca... | 114 | Action|Drama|Thriller | New Line Cinema|Village Roadshow Pictures|Warn... | 5/27/2015 | 2060 | 6.1 | 2015 | 1.012000e+08 | 4.328514e+08 |
| **23** | 216015 | tt2322441 | 4.710402 | 40000000 | 569651467 | Fifty Shades of Grey | Dakota Johnson|Jamie Dornan|Jennifer Ehle|Eloi... | https://www.facebook.com/fiftyshadesofgreymovie | Sam Taylor-Johnson | Are you curious? | ... | When college senior Anastasia Steele steps in ... | 125 | Drama|Romance | Focus Features|Trigger Street Productions|Mich... | 2/11/2015 | 1865 | 5.3 | 2015 | 3.679998e+07 | 5.240791e+08 |
| **24** | 318846 | tt1596363 | 4.648046 | 28000000 | 133346506 | The Big Short | Christian Bale|Steve Carell|Ryan Gosling|Brad ... | http://www.thebigshortmovie.com/ | Adam McKay | This is a true story. | ... | The men who made millions from a global econom... | 130 | Comedy|Drama | Paramount Pictures|Plan B Entertainment|Regenc... | 12/11/2015 | 1545 | 7.3 | 2015 | 2.575999e+07 | 1.226787e+08 |
| **25** | 177677 | tt2381249 | 4.566713 | 150000000 | 682330139 | Mission: Impossible - Rogue Nation | Tom Cruise|Jeremy Renner|Simon Pegg|Rebecca Fe... | http://www.missionimpossible.com | Christopher McQuarrie | Desperate Times. Desperate Measures. | ... | Ethan and team take on their most impossible m... | 131 | Action | Paramount Pictures|Skydance Productions|China ... | 7/23/2015 | 2349 | 7.1 | 2015 | 1.379999e+08 | 6.277435e+08 |
| **26** | 214756 | tt2637276 | 4.564549 | 68000000 | 215863606 | Ted 2 | Mark Wahlberg|Seth MacFarlane|Amanda Seyfried|... | NaN | Seth MacFarlane | Ted is Coming, Again. | ... | Newlywed couple Ted and Tami-Lynn want to have... | 115 | Comedy | Universal Pictures|Media Rights Capital|Fuzzy ... | 6/25/2015 | 1666 | 6.3 | 2015 | 6.255997e+07 | 1.985944e+08 |
| **27** | 207703 | tt2802144 | 4.503789 | 81000000 | 403802136 | Kingsman: The Secret Service | Taron Egerton|Colin Firth|Samuel L. Jackson|Mi... | http://www.kingsmanmovie.com/ | Matthew Vaughn | Manners maketh man. | ... | The story of a super-secret spy organization t... | 130 | Crime|Comedy|Action|Adventure | Twentieth Century Fox Film Corporation|Marv Fi... | 1/24/2015 | 3833 | 7.6 | 2015 | 7.451997e+07 | 3.714978e+08 |
| **28** | 314365 | tt1895587 | 4.062293 | 20000000 | 88346473 | Spotlight | Mark Ruffalo|Michael Keaton|Rachel McAdams|Lie... | http://www.spotlightthefilm.com | Tom McCarthy | Break the story. Break the silence. | ... | The true story of how The Boston Globe uncover... | 128 | Drama|Thriller|History | Participant Media|Open Road Films|Anonymous Co... | 11/6/2015 | 1559 | 7.8 | 2015 | 1.839999e+07 | 8.127872e+07 |
| **29** | 294254 | tt4046784 | 3.968891 | 61000000 | 311256926 | Maze Runner: The Scorch Trials | Dylan O'Brien|Kaya Scodelario|Thomas Brodie-Sa... | http://mazerunnermovies.com | Wes Ball | The Maze Was Just the Beginning. | ... | Thomas and his fellow Gladers face their great... | 132 | Action|Science Fiction|Thriller | Gotham Group|Temple Hill Entertainment|TSG Ent... | 9/9/2015 | 1849 | 6.4 | 2015 | 5.611998e+07 | 2.863562e+08 |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| **10836** | 38720 | tt0061170 | 0.239435 | 0 | 0 | Walk Don't Run | Cary Grant|Samantha Eggar|Jim Hutton|John Stan... | NaN | Charles Walters | Run, don't walk to see Walk, Don't Run. | ... | British industrialist Sir William Rutland - "B... | 114 | Comedy|Romance | Columbia Pictures Corporation | 1/1/1966 | 11 | 5.8 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10837** | 19728 | tt0060177 | 0.291704 | 0 | 0 | The Blue Max | George Peppard|James Mason|Ursula Andress|Jere... | NaN | John Guillermin | There was no quiet on the Western Front! | ... | A young pilot in the German air force of 1918,... | 156 | War|Action|Adventure|Drama | Twentieth Century Fox Film Corporation | 6/21/1966 | 12 | 5.5 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10838** | 22383 | tt0060862 | 0.151845 | 0 | 0 | The Professionals | Burt Lancaster|Lee Marvin|Robert Ryan|Woody St... | NaN | Richard Brooks | Rough, tough and ready. | ... | The Professionals is a 1966 American Western f... | 117 | Action|Adventure|Western | Columbia Pictures | 11/1/1966 | 21 | 6.0 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10839** | 13353 | tt0060550 | 0.276133 | 0 | 0 | It's the Great Pumpkin, Charlie Brown | Christopher Shea|Sally Dryer|Kathy Steinberg|A... | NaN | Bill Melendez | Every year he rises from the pumpkin patch... | ... | This classic "Peanuts" tale focuses on the thu... | 25 | Family|Animation | Warner Bros. Home Video | 10/27/1966 | 49 | 7.2 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10840** | 34388 | tt0060437 | 0.102530 | 0 | 0 | Funeral in Berlin | Michael Caine|Paul Hubschmid|Oskar Homolka|Eva... | NaN | Guy Hamilton | NaN | ... | Colonel Stok, a Soviet intelligence officer re... | 102 | Thriller | Lowndes Productions Limited | 12/22/1966 | 13 | 5.7 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10841** | 42701 | tt0062262 | 0.264925 | 75000 | 0 | The Shooting | Will Hutchins|Millie Perkins|Jack Nicholson|Wa... | NaN | Monte Hellman | Suspenseful desert pursuit in the "High Noon" ... | ... | A hired gun seeks to enact revenge on a group ... | 82 | Western | Proteus Films | 10/23/1966 | 12 | 5.5 | 1966 | 5.038511e+05 | 0.000000e+00 |
| **10842** | 36540 | tt0061199 | 0.253437 | 0 | 0 | Winnie the Pooh and the Honey Tree | Sterling Holloway|Junius Matthews|Sebastian Ca... | NaN | Wolfgang Reitherman | NaN | ... | Christopher Robin's bear attempts to raid a be... | 25 | Animation|Family | NaN | 1/1/1966 | 12 | 7.9 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10843** | 29710 | tt0060588 | 0.252399 | 0 | 0 | Khartoum | Charlton Heston|Laurence Olivier|Richard Johns... | NaN | Basil Dearden|Eliot Elisofon | Where the Nile divides, the great Cinerama adv... | ... | English General Charles George Gordon, a devou... | 134 | Adventure|Drama|War|History|Action | Julian Blaustein Productions Ltd. | 6/9/1966 | 12 | 5.8 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10844** | 23728 | tt0059557 | 0.236098 | 0 | 0 | Our Man Flint | James Coburn|Lee J. Cobb|Gila Golan|Edward Mul... | NaN | Daniel Mann | The ORIGINAL man of mystery! | ... | When scientists use eco-terrorism to impose th... | 108 | Adventure|Comedy|Fantasy|Science Fiction | 20th Century Fox | 1/16/1966 | 13 | 5.6 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10845** | 5065 | tt0059014 | 0.230873 | 0 | 0 | Carry On Cowboy | Sid James|Jim Dale|Angela Douglas|Kenneth Will... | NaN | Gerald Thomas | How the west was lost! | ... | Stodge City is in the grip of the Rumpo Kid an... | 93 | Comedy|Western | Peter Rogers Productions | 3/1/1966 | 15 | 5.9 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10846** | 17102 | tt0059127 | 0.212716 | 0 | 0 | Dracula: Prince of Darkness | Christopher Lee|Barbara Shelley|Andrew Keir|Fr... | NaN | Terence Fisher | DEAD for Ten Years DRACULA, Prince of Darkness... | ... | Whilst vacationing in the Carpathian Mountain,... | 90 | Horror | Seven Arts Productions|Hammer Film Productions | 1/9/1966 | 16 | 5.7 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10847** | 28763 | tt0060548 | 0.034555 | 0 | 0 | Island of Terror | Peter Cushing|Edward Judd|Carole Gray|Eddie By... | NaN | Terence Fisher | How could they stop the devouring death...that... | ... | A small island community is overrun with creep... | 89 | Science Fiction|Horror | Planet Film Productions|Protelco | 6/20/1966 | 13 | 5.3 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10848** | 2161 | tt0060397 | 0.207257 | 5115000 | 12000000 | Fantastic Voyage | Stephen Boyd|Raquel Welch|Edmond O'Brien|Donal... | NaN | Richard Fleischer | A Fantastic and Spectacular Voyage... Through ... | ... | The science of miniaturization has been unlock... | 100 | Adventure|Science Fiction | Twentieth Century Fox Film Corporation | 8/24/1966 | 42 | 6.7 | 1966 | 3.436265e+07 | 8.061618e+07 |
| **10849** | 28270 | tt0060445 | 0.206537 | 0 | 0 | Gambit | Michael Caine|Shirley MacLaine|Herbert Lom|Joh... | NaN | Ronald Neame | Shirley MacLaine raises Michael Caine! | ... | Harry Dean (Michael Caine) has a perfect plan ... | 109 | Action|Comedy|Crime | Universal Pictures | 12/16/1966 | 14 | 6.1 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10850** | 26268 | tt0060490 | 0.202473 | 0 | 0 | Harper | Paul Newman|Lauren Bacall|Julie Harris|Arthur ... | NaN | Jack Smight | Harper takes a case - and the payoff is murder. | ... | Harper is a cynical private eye in the best tr... | 121 | Action|Drama|Thriller|Crime|Mystery | Warner Bros. | 2/23/1966 | 14 | 6.0 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10851** | 15347 | tt0060182 | 0.342791 | 0 | 0 | Born Free | Virginia McKenna|Bill Travers|Geoffrey Keen|Pe... | NaN | James Hill | From The Pages Of The Beloved Best Seller... A... | ... | Born Free (1966) is an Open Road Films Ltd./Co... | 95 | Adventure|Drama|Action|Family|Foreign | High Road | 6/22/1966 | 15 | 6.6 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10852** | 37301 | tt0060165 | 0.227220 | 0 | 0 | A Big Hand for the Little Lady | Henry Fonda|Joanne Woodward|Jason Robards|Paul... | NaN | Fielder Cook | All the action you can take...all the adventur... | ... | A naive traveler in Laredo gets involved in a ... | 95 | Western | Eden Productions Inc. | 5/31/1966 | 11 | 6.0 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10853** | 15598 | tt0060086 | 0.163592 | 0 | 0 | Alfie | Michael Caine|Shelley Winters|Millicent Martin... | NaN | Lewis Gilbert | Is any man an Alfie? Ask any girl! | ... | The film tells the story of a young man who le... | 114 | Comedy|Drama|Romance | NaN | 3/29/1966 | 26 | 6.2 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10854** | 31602 | tt0060232 | 0.146402 | 0 | 0 | The Chase | Marlon Brando|Jane Fonda|Robert Redford|E.G. M... | NaN | Arthur Penn | The chase is on! | ... | Most everyone in town thinks that Sheriff Cald... | 135 | Thriller|Drama|Crime | Horizon Pictures|Columbia Pictures Corporation | 2/17/1966 | 17 | 6.0 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10855** | 13343 | tt0059221 | 0.141026 | 700000 | 0 | The Ghost & Mr. Chicken | Don Knotts|Joan Staley|Liam Redmond|Dick Sarge... | NaN | Alan Rafkin | G-G-GUARANTEED! YOU'LL BE SCARED UNTIL YOU LAU... | ... | Luther Heggs aspires to being a reporter for h... | 90 | Comedy|Family|Mystery|Romance | Universal Pictures | 1/20/1966 | 14 | 6.1 | 1966 | 4.702610e+06 | 0.000000e+00 |
| **10856** | 20277 | tt0061135 | 0.140934 | 0 | 0 | The Ugly Dachshund | Dean Jones|Suzanne Pleshette|Charles Ruggles|K... | NaN | Norman Tokar | A HAPPY HONEYMOON GOES TO THE DOGS!...When a G... | ... | The Garrisons (Dean Jones and Suzanne Pleshett... | 93 | Comedy|Drama|Family | Walt Disney Pictures | 2/16/1966 | 14 | 5.7 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10857** | 5921 | tt0060748 | 0.131378 | 0 | 0 | Nevada Smith | Steve McQueen|Karl Malden|Brian Keith|Arthur K... | NaN | Henry Hathaway | Some called him savage- and some called him sa... | ... | Nevada Smith is the young son of an Indian mot... | 128 | Action|Western | Paramount Pictures|Solar Productions|Embassy P... | 6/10/1966 | 10 | 5.9 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10858** | 31918 | tt0060921 | 0.317824 | 0 | 0 | The Russians Are Coming, The Russians Are Coming | Carl Reiner|Eva Marie Saint|Alan Arkin|Brian K... | NaN | Norman Jewison | IT'S A PLOT! ...to make the world die laughing!! | ... | Without hostile intent, a Soviet sub runs agro... | 126 | Comedy|War | The Mirisch Corporation | 5/25/1966 | 11 | 5.5 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10859** | 20620 | tt0060955 | 0.089072 | 0 | 0 | Seconds | Rock Hudson|Salome Jens|John Randolph|Will Gee... | NaN | John Frankenheimer | NaN | ... | A secret organisation offers wealthy people a ... | 100 | Mystery|Science Fiction|Thriller|Drama | Gibraltar Productions|Joel Productions|John Fr... | 10/5/1966 | 22 | 6.6 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10860** | 5060 | tt0060214 | 0.087034 | 0 | 0 | Carry On Screaming! | Kenneth Williams|Jim Dale|Harry H. Corbett|Joa... | NaN | Gerald Thomas | Carry On Screaming with the Hilarious CARRY ON... | ... | The sinister Dr Watt has an evil scheme going.... | 87 | Comedy | Peter Rogers Productions|Anglo-Amalgamated Fil... | 5/20/1966 | 13 | 7.0 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10861** | 21 | tt0060371 | 0.080598 | 0 | 0 | The Endless Summer | Michael Hynson|Robert August|Lord 'Tally Ho' B... | NaN | Bruce Brown | NaN | ... | The Endless Summer, by Bruce Brown, is one of ... | 95 | Documentary | Bruce Brown Films | 6/15/1966 | 11 | 7.4 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10862** | 20379 | tt0060472 | 0.065543 | 0 | 0 | Grand Prix | James Garner|Eva Marie Saint|Yves Montand|Tosh... | NaN | John Frankenheimer | Cinerama sweeps YOU into a drama of speed and ... | ... | Grand Prix driver Pete Aron is fired by his te... | 176 | Action|Adventure|Drama | Cherokee Productions|Joel Productions|Douglas ... | 12/21/1966 | 20 | 5.7 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10863** | 39768 | tt0060161 | 0.065141 | 0 | 0 | Beregis Avtomobilya | Innokentiy Smoktunovskiy|Oleg Efremov|Georgi Z... | NaN | Eldar Ryazanov | NaN | ... | An insurance agent who moonlights as a carthie... | 94 | Mystery|Comedy | Mosfilm | 1/1/1966 | 11 | 6.5 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10864** | 21449 | tt0061177 | 0.064317 | 0 | 0 | What's Up, Tiger Lily? | Tatsuya Mihashi|Akiko Wakabayashi|Mie Hama|Joh... | NaN | Woody Allen | WOODY ALLEN STRIKES BACK! | ... | In comic Woody Allen's film debut, he took the... | 80 | Action|Comedy | Benedict Pictures Corp. | 11/2/1966 | 22 | 5.4 | 1966 | 0.000000e+00 | 0.000000e+00 |
| **10865** | 22293 | tt0060666 | 0.035919 | 19000 | 0 | Manos: The Hands of Fate | Harold P. Warren|Tom Neyman|John Reynolds|Dian... | NaN | Harold P. Warren | It's Shocking! It's Beyond Your Imagination! | ... | A family gets lost on the road and stumbles up... | 74 | Horror | Norm-Iris | 11/15/1966 | 15 | 1.5 | 1966 | 1.276423e+05 | 0.000000e+00 |

10866 rows × 21 columns

In [5]:

*#finding duplication*

sum(df.duplicated())

Out[5]:

1

In [6]:

*#Remove duplication and Confirming the duplication*

df.drop\_duplicates(inplace=**True**)

df.info()

<class 'pandas.core.frame.DataFrame'>

Int64Index: 10865 entries, 0 to 10865

Data columns (total 21 columns):

id 10865 non-null int64

imdb\_id 10855 non-null object

popularity 10865 non-null float64

budget 10865 non-null int64

revenue 10865 non-null int64

original\_title 10865 non-null object

cast 10789 non-null object

homepage 2936 non-null object

director 10821 non-null object

tagline 8041 non-null object

keywords 9372 non-null object

overview 10861 non-null object

runtime 10865 non-null int64

genres 10842 non-null object

production\_companies 9835 non-null object

release\_date 10865 non-null object

vote\_count 10865 non-null int64

vote\_average 10865 non-null float64

release\_year 10865 non-null int64

budget\_adj 10865 non-null float64

revenue\_adj 10865 non-null float64

dtypes: float64(4), int64(6), object(11)

memory usage: 1.8+ MB

In [7]:

*#Listing the columns with null value*

df.isnull().sum()

Out[7]:

id 0

imdb\_id 10

popularity 0

budget 0

revenue 0

original\_title 0

cast 76

homepage 7929

director 44

tagline 2824

keywords 1493

overview 4

runtime 0

genres 23

production\_companies 1030

release\_date 0

vote\_count 0

vote\_average 0

release\_year 0

budget\_adj 0

revenue\_adj 0

dtype: int64

In [8]:

*#After removing duplicate value, we Remove rows with null value for the target variable.*

df.dropna(subset=['director'], inplace=**True**)

df.dropna(subset=['popularity'], inplace=**True**)

df.dropna(subset=['runtime'], inplace=**True**)

df.dropna(subset=['budget'], inplace=**True**)

df.dropna(subset=['vote\_count'], inplace=**True**)

df.dropna(subset=['revenue'], inplace=**True**)

df.info()

<class 'pandas.core.frame.DataFrame'>

Int64Index: 10821 entries, 0 to 10865

Data columns (total 21 columns):

id 10821 non-null int64

imdb\_id 10815 non-null object

popularity 10821 non-null float64

budget 10821 non-null int64

revenue 10821 non-null int64

original\_title 10821 non-null object

cast 10751 non-null object

homepage 2926 non-null object

director 10821 non-null object

tagline 8031 non-null object

keywords 9356 non-null object

overview 10818 non-null object

runtime 10821 non-null int64

genres 10800 non-null object

production\_companies 9814 non-null object

release\_date 10821 non-null object

vote\_count 10821 non-null int64

vote\_average 10821 non-null float64

release\_year 10821 non-null int64

budget\_adj 10821 non-null float64

revenue\_adj 10821 non-null float64

dtypes: float64(4), int64(6), object(11)

memory usage: 1.8+ MB

In [9]:

*#Visually inspecting the data and circumstantially display a certain number of rows and columns.*

df.head()

Out[9]:

|  | **id** | **imdb\_id** | **popularity** | **budget** | **revenue** | **original\_title** | **cast** | **homepage** | **director** | **tagline** | **...** | **overview** | **runtime** | **genres** | **production\_companies** | **release\_date** | **vote\_count** | **vote\_average** | **release\_year** | **budget\_adj** | **revenue\_adj** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | 135397 | tt0369610 | 32.985763 | 150000000 | 1513528810 | Jurassic World | Chris Pratt|Bryce Dallas Howard|Irrfan Khan|Vi... | http://www.jurassicworld.com/ | Colin Trevorrow | The park is open. | ... | Twenty-two years after the events of Jurassic ... | 124 | Action|Adventure|Science Fiction|Thriller | Universal Studios|Amblin Entertainment|Legenda... | 6/9/2015 | 5562 | 6.5 | 2015 | 137999939.3 | 1.392446e+09 |
| **1** | 76341 | tt1392190 | 28.419936 | 150000000 | 378436354 | Mad Max: Fury Road | Tom Hardy|Charlize Theron|Hugh Keays-Byrne|Nic... | http://www.madmaxmovie.com/ | George Miller | What a Lovely Day. | ... | An apocalyptic story set in the furthest reach... | 120 | Action|Adventure|Science Fiction|Thriller | Village Roadshow Pictures|Kennedy Miller Produ... | 5/13/2015 | 6185 | 7.1 | 2015 | 137999939.3 | 3.481613e+08 |
| **2** | 262500 | tt2908446 | 13.112507 | 110000000 | 295238201 | Insurgent | Shailene Woodley|Theo James|Kate Winslet|Ansel... | http://www.thedivergentseries.movie/#insurgent | Robert Schwentke | One Choice Can Destroy You | ... | Beatrice Prior must confront her inner demons ... | 119 | Adventure|Science Fiction|Thriller | Summit Entertainment|Mandeville Films|Red Wago... | 3/18/2015 | 2480 | 6.3 | 2015 | 101199955.5 | 2.716190e+08 |
| **3** | 140607 | tt2488496 | 11.173104 | 200000000 | 2068178225 | Star Wars: The Force Awakens | Harrison Ford|Mark Hamill|Carrie Fisher|Adam D... | http://www.starwars.com/films/star-wars-episod... | J.J. Abrams | Every generation has a story. | ... | Thirty years after defeating the Galactic Empi... | 136 | Action|Adventure|Science Fiction|Fantasy | Lucasfilm|Truenorth Productions|Bad Robot | 12/15/2015 | 5292 | 7.5 | 2015 | 183999919.0 | 1.902723e+09 |
| **4** | 168259 | tt2820852 | 9.335014 | 190000000 | 1506249360 | Furious 7 | Vin Diesel|Paul Walker|Jason Statham|Michelle ... | http://www.furious7.com/ | James Wan | Vengeance Hits Home | ... | Deckard Shaw seeks revenge against Dominic Tor... | 137 | Action|Crime|Thriller | Universal Pictures|Original Film|Media Rights ... | 4/1/2015 | 2947 | 7.3 | 2015 | 174799923.1 | 1.385749e+09 |

5 rows × 21 columns

In [10]:

*#displaying the rows*

df.tail()

Out[10]:

|  | **id** | **imdb\_id** | **popularity** | **budget** | **revenue** | **original\_title** | **cast** | **homepage** | **director** | **tagline** | **...** | **overview** | **runtime** | **genres** | **production\_companies** | **release\_date** | **vote\_count** | **vote\_average** | **release\_year** | **budget\_adj** | **revenue\_adj** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **10861** | 21 | tt0060371 | 0.080598 | 0 | 0 | The Endless Summer | Michael Hynson|Robert August|Lord 'Tally Ho' B... | NaN | Bruce Brown | NaN | ... | The Endless Summer, by Bruce Brown, is one of ... | 95 | Documentary | Bruce Brown Films | 6/15/1966 | 11 | 7.4 | 1966 | 0.0000 | 0.0 |
| **10862** | 20379 | tt0060472 | 0.065543 | 0 | 0 | Grand Prix | James Garner|Eva Marie Saint|Yves Montand|Tosh... | NaN | John Frankenheimer | Cinerama sweeps YOU into a drama of speed and ... | ... | Grand Prix driver Pete Aron is fired by his te... | 176 | Action|Adventure|Drama | Cherokee Productions|Joel Productions|Douglas ... | 12/21/1966 | 20 | 5.7 | 1966 | 0.0000 | 0.0 |
| **10863** | 39768 | tt0060161 | 0.065141 | 0 | 0 | Beregis Avtomobilya | Innokentiy Smoktunovskiy|Oleg Efremov|Georgi Z... | NaN | Eldar Ryazanov | NaN | ... | An insurance agent who moonlights as a carthie... | 94 | Mystery|Comedy | Mosfilm | 1/1/1966 | 11 | 6.5 | 1966 | 0.0000 | 0.0 |
| **10864** | 21449 | tt0061177 | 0.064317 | 0 | 0 | What's Up, Tiger Lily? | Tatsuya Mihashi|Akiko Wakabayashi|Mie Hama|Joh... | NaN | Woody Allen | WOODY ALLEN STRIKES BACK! | ... | In comic Woody Allen's film debut, he took the... | 80 | Action|Comedy | Benedict Pictures Corp. | 11/2/1966 | 22 | 5.4 | 1966 | 0.0000 | 0.0 |
| **10865** | 22293 | tt0060666 | 0.035919 | 19000 | 0 | Manos: The Hands of Fate | Harold P. Warren|Tom Neyman|John Reynolds|Dian... | NaN | Harold P. Warren | It's Shocking! It's Beyond Your Imagination! | ... | A family gets lost on the road and stumbles up... | 74 | Horror | Norm-Iris | 11/15/1966 | 15 | 1.5 | 1966 | 127642.2792 | 0.0 |

5 rows × 21 columns

In [11]:

*#listing available columns on a DataFrame*

df.columns

Out[11]:

Index(['id', 'imdb\_id', 'popularity', 'budget', 'revenue', 'original\_title',

'cast', 'homepage', 'director', 'tagline', 'keywords', 'overview',

'runtime', 'genres', 'production\_companies', 'release\_date',

'vote\_count', 'vote\_average', 'release\_year', 'budget\_adj',

'revenue\_adj'],

dtype='object')

In [12]:

*#Finding the frequency count for all variable and Droping the columns where all elements are nan((“not available”))*

df.director.value\_counts(dropna=**False**)

Out[12]:

Woody Allen 45

Clint Eastwood 34

Martin Scorsese 29

Steven Spielberg 29

Ridley Scott 23

Steven Soderbergh 22

Ron Howard 22

Joel Schumacher 21

Brian De Palma 20

Tim Burton 19

Barry Levinson 19

Wes Craven 19

David Cronenberg 18

Mike Nichols 18

John Carpenter 18

Rob Reiner 18

Sidney Lumet 17

Francis Ford Coppola 17

Norman Jewison 17

Robert Zemeckis 17

Spike Lee 17

Peter Hyams 17

Walter Hill 17

Renny Harlin 17

Oliver Stone 17

Stephen Frears 17

Tyler Perry 17

Richard Linklater 16

John Landis 16

Roman Polanski 16

..

Michael Showalter 1

Laurent Malaquais 1

Scott Walker 1

Robert Shaye 1

Wei Tung 1

Ross Duffer|Matt Duffer 1

Chris Messina 1

K.C. Bascombe 1

Morgan Matthews 1

Trey Nelson 1

Sean Mullin 1

Christian Robinson 1

Mitchell Lichtenstein 1

Alberto Marini 1

Dave Filoni 1

Fernando Barreda Luna 1

Aaron Seltzer|Jason Friedberg 1

Aki KaurismÃÂ¤ki 1

Joel Allen Schroeder 1

Gerado Olivares|Otmar Penker 1

Steven E. de Souza 1

Amy Rice 1

Y.K. Kim|Woo-sang Park 1

Inon Shampanier 1

Mark Mori 1

Marcus Warren 1

Justin Dix 1

Philip Sansom|Olly Williams 1

Shane Salerno 1

Jason Momoa 1

Name: director, Length: 5067, dtype: int64

In [13]:

*#Finding numeric data (statistics)*

df.describe()

Out[13]:

|  | **id** | **popularity** | **budget** | **revenue** | **runtime** | **vote\_count** | **vote\_average** | **release\_year** | **budget\_adj** | **revenue\_adj** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **count** | 10821.000000 | 10821.000000 | 1.082100e+04 | 1.082100e+04 | 10821.000000 | 10821.000000 | 10821.000000 | 10821.000000 | 1.082100e+04 | 1.082100e+04 |
| **mean** | 65722.143055 | 0.648530 | 1.468259e+07 | 3.998782e+07 | 102.130302 | 218.211810 | 5.971509 | 2001.294890 | 1.762011e+07 | 5.157685e+07 |
| **std** | 91841.687567 | 1.001704 | 3.096345e+07 | 1.172186e+08 | 30.872755 | 576.672383 | 0.933671 | 12.819687 | 3.435943e+07 | 1.448953e+08 |
| **min** | 5.000000 | 0.000188 | 0.000000e+00 | 0.000000e+00 | 0.000000 | 10.000000 | 1.500000 | 1960.000000 | 0.000000e+00 | 0.000000e+00 |
| **25%** | 10578.000000 | 0.209036 | 0.000000e+00 | 0.000000e+00 | 90.000000 | 17.000000 | 5.400000 | 1995.000000 | 0.000000e+00 | 0.000000e+00 |
| **50%** | 20533.000000 | 0.385121 | 0.000000e+00 | 0.000000e+00 | 99.000000 | 38.000000 | 6.000000 | 2006.000000 | 0.000000e+00 | 0.000000e+00 |
| **75%** | 74998.000000 | 0.716336 | 1.500000e+07 | 2.433232e+07 | 111.000000 | 146.000000 | 6.600000 | 2011.000000 | 2.099042e+07 | 3.398486e+07 |
| **max** | 417859.000000 | 32.985763 | 4.250000e+08 | 2.781506e+09 | 900.000000 | 9767.000000 | 9.200000 | 2015.000000 | 4.250000e+08 | 2.827124e+09 |

In [14]:

*# Research quistion:*

*#1-Which is the most popular genres movies ? Drama &Comedy.*

*#the graph below showing the most popular genres movies which are Drama &Comedy.*

temp1 = df[['genres','original\_title' ]]

group1 = temp1.groupby(['genres']).count()

group1 = group1.nlargest(20,'original\_title')

plt.figure()

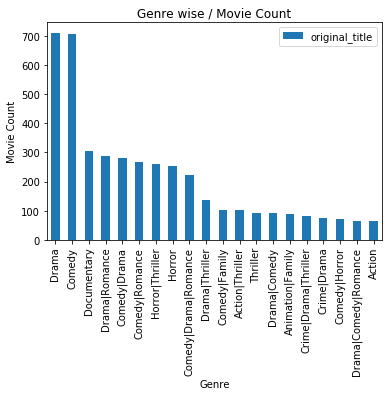
group1.plot(kind='bar', title="Genre wise / Movie Count")

plt.xlabel("Genre")

plt.ylabel("Movie Count")

plt.show()

<matplotlib.figure.Figure at 0x14ef8479d68>

In [15]:

*#copying the original file'tmdb-movies.csv' data*

df\_dir = df.copy()

In [16]:

*#2-Which director has most popularity? Colin*

*# from the data below showing Colin Trevorrow is most popular but he didn't have the most realeases.*

*# so the popularity and biger anumber of realease dindnt go in leanear direction*

director\_df = df\_dir.groupby('director')

dir\_mean\_df = director\_df.mean()

dir\_mean\_df

Out[16]:

|  | **id** | **popularity** | **budget** | **revenue** | **runtime** | **vote\_count** | **vote\_average** | **release\_year** | **budget\_adj** | **revenue\_adj** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **director** |  |  |  |  |  |  |  |  |  |  |
| **FrÃÂ©dÃÂ©ric Jardin** | 79070.000000 | 0.434506 | 2.500000e+06 | 3.358000e+03 | 98.000000 | 28.000000 | 5.900000 | 2011.000000 | 2.423495e+06 | 3.255239e+03 |
| **A. Todd Smith** | 370687.000000 | 1.876037 | 0.000000e+00 | 0.000000e+00 | 0.000000 | 11.000000 | 5.400000 | 2015.000000 | 0.000000e+00 | 0.000000e+00 |
| **A.R. Murugadoss** | 14070.000000 | 0.215030 | 9.100000e+06 | 7.600000e+07 | 183.000000 | 53.000000 | 6.900000 | 2008.000000 | 9.216358e+06 | 7.697178e+07 |
| **Aaron Aites** | 36124.000000 | 0.108700 | 0.000000e+00 | 0.000000e+00 | 93.000000 | 13.000000 | 6.500000 | 2008.000000 | 0.000000e+00 | 0.000000e+00 |
| **Aaron Blaise|Robert Walker** | 10009.000000 | 1.653031 | 1.000000e+08 | 2.500000e+02 | 85.000000 | 753.000000 | 6.800000 | 2003.000000 | 1.185353e+08 | 2.963382e+02 |
| **Aaron Hann|Mario Miscione** | 335866.000000 | 0.350599 | 0.000000e+00 | 0.000000e+00 | 87.000000 | 167.000000 | 6.100000 | 2015.000000 | 0.000000e+00 | 0.000000e+00 |
| **Aaron Harvey** | 75736.000000 | 1.212003 | 7.000000e+06 | 0.000000e+00 | 93.000000 | 71.000000 | 4.900000 | 2011.000000 | 6.785786e+06 | 0.000000e+00 |
| **Aaron Katz|Martha Stephens** | 252102.000000 | 0.267446 | 0.000000e+00 | 0.000000e+00 | 96.000000 | 17.000000 | 5.800000 | 2014.000000 | 0.000000e+00 | 0.000000e+00 |
| **Aaron Keeling|Austin Keeling** | 326247.000000 | 0.078363 | 0.000000e+00 | 0.000000e+00 | 111.000000 | 10.000000 | 3.800000 | 2015.000000 | 0.000000e+00 | 0.000000e+00 |
| **Aaron Moorhead|Justin Benson** | 241855.000000 | 0.383406 | 0.000000e+00 | 4.997000e+04 | 109.000000 | 110.000000 | 6.500000 | 2015.000000 | 0.000000e+00 | 4.597238e+04 |
| **Aaron Norris** | 27370.000000 | 0.158442 | 1.200000e+06 | 4.528051e+06 | 98.800000 | 17.600000 | 5.120000 | 1991.200000 | 1.717160e+06 | 7.468127e+06 |
| **Aaron Schneider** | 44718.000000 | 0.661465 | 7.500000e+06 | 0.000000e+00 | 103.000000 | 41.000000 | 5.900000 | 2010.000000 | 7.500000e+06 | 0.000000e+00 |
| **Aaron Seltzer|Jason Friedberg** | 10073.000000 | 0.698297 | 2.000000e+07 | 4.854843e+07 | 83.000000 | 103.000000 | 3.900000 | 2006.000000 | 2.163338e+07 | 5.251334e+07 |
| **Aaron Woodley** | 73818.000000 | 0.285594 | 0.000000e+00 | 0.000000e+00 | 91.000000 | 19.000000 | 5.500000 | 2011.000000 | 0.000000e+00 | 0.000000e+00 |
| **Abbas Alibhai Burmawalla|Mastan Alibhai Burmawalla** | 63603.000000 | 0.091416 | 0.000000e+00 | 0.000000e+00 | 158.000000 | 14.000000 | 5.600000 | 2004.500000 | 0.000000e+00 | 0.000000e+00 |
| **Abbas Kiarostami** | 46693.500000 | 0.163712 | 0.000000e+00 | 0.000000e+00 | 100.500000 | 15.000000 | 6.450000 | 1993.000000 | 0.000000e+00 | 0.000000e+00 |
| **Abbas Tyrewala** | 14467.000000 | 0.193779 | 2.300000e+06 | 1.110000e+07 | 155.000000 | 18.000000 | 6.400000 | 2008.000000 | 2.329409e+06 | 1.124193e+07 |
| **Abe Levy|Silver Tree** | 180688.000000 | 0.091023 | 0.000000e+00 | 0.000000e+00 | 90.000000 | 12.000000 | 5.200000 | 2013.000000 | 0.000000e+00 | 0.000000e+00 |
| **Abe Sylvia** | 70695.000000 | 0.315527 | 4.000000e+06 | 5.512500e+04 | 90.000000 | 36.000000 | 6.200000 | 2010.000000 | 4.000000e+06 | 5.512500e+04 |
| **Abel Ferrara** | 48692.000000 | 0.408530 | 1.882750e+06 | 6.275768e+05 | 94.000000 | 29.875000 | 5.400000 | 1996.750000 | 2.874340e+06 | 1.008261e+06 |
| **Abel Ferry** | 30974.000000 | 0.097705 | 0.000000e+00 | 0.000000e+00 | 90.000000 | 38.000000 | 5.500000 | 2009.000000 | 0.000000e+00 | 0.000000e+00 |
| **Adam Brooks** | 8390.000000 | 1.341805 | 0.000000e+00 | 5.544797e+07 | 112.000000 | 389.000000 | 6.500000 | 2008.000000 | 0.000000e+00 | 5.615696e+07 |
| **Adam Brooks|Matthew Kennedy** | 285840.000000 | 0.203224 | 0.000000e+00 | 0.000000e+00 | 95.000000 | 18.000000 | 6.000000 | 2015.000000 | 0.000000e+00 | 0.000000e+00 |
| **Adam Curtis** | 140810.333333 | 0.127357 | 0.000000e+00 | 0.000000e+00 | 183.333333 | 15.666667 | 7.900000 | 2007.000000 | 0.000000e+00 | 0.000000e+00 |
| **Adam Deacon|Daniel Toland** | 68419.000000 | 0.174743 | 1.000000e+06 | 0.000000e+00 | 88.000000 | 17.000000 | 7.300000 | 2011.000000 | 9.693980e+05 | 0.000000e+00 |
| **Adam Egypt Mortimer** | 334028.000000 | 0.277613 | 0.000000e+00 | 0.000000e+00 | 82.000000 | 19.000000 | 3.900000 | 2015.000000 | 0.000000e+00 | 0.000000e+00 |
| **Adam Elliot** | 22684.500000 | 0.378397 | 4.000000e+06 | 0.000000e+00 | 57.500000 | 191.000000 | 7.150000 | 2006.000000 | 4.065602e+06 | 0.000000e+00 |
| **Adam Fields** | 24133.000000 | 0.130306 | 0.000000e+00 | 0.000000e+00 | 84.000000 | 11.000000 | 5.100000 | 2008.000000 | 0.000000e+00 | 0.000000e+00 |
| **Adam Gierasch** | 36473.333333 | 0.161406 | 8.333333e+06 | 0.000000e+00 | 90.666667 | 17.666667 | 4.233333 | 2009.333333 | 8.451935e+06 | 0.000000e+00 |
| **Adam Green** | 95040.500000 | 0.376700 | 6.250000e+05 | 7.796160e+05 | 90.750000 | 116.000000 | 5.400000 | 2010.250000 | 6.250000e+05 | 7.796160e+05 |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| **Zach Ingrasci|Sean Leonard|Chris Temple** | 229768.000000 | 0.268563 | 0.000000e+00 | 0.000000e+00 | 56.000000 | 18.000000 | 6.700000 | 2013.000000 | 0.000000e+00 | 0.000000e+00 |
| **Zach Lipovsky** | 209971.500000 | 0.384615 | 0.000000e+00 | 0.000000e+00 | 104.000000 | 58.000000 | 4.400000 | 2014.500000 | 0.000000e+00 | 0.000000e+00 |
| **Zachary Donohue** | 244610.000000 | 0.204989 | 0.000000e+00 | 0.000000e+00 | 81.000000 | 58.000000 | 5.900000 | 2013.000000 | 0.000000e+00 | 0.000000e+00 |
| **Zachary Heinzerling** | 159002.000000 | 0.079050 | 0.000000e+00 | 0.000000e+00 | 81.000000 | 20.000000 | 7.300000 | 2013.000000 | 0.000000e+00 | 0.000000e+00 |
| **Zachary Sluser** | 291189.000000 | 0.246609 | 0.000000e+00 | 0.000000e+00 | 95.000000 | 12.000000 | 5.000000 | 2015.000000 | 0.000000e+00 | 0.000000e+00 |
| **Zack Parker** | 211065.000000 | 0.187127 | 0.000000e+00 | 0.000000e+00 | 122.000000 | 19.000000 | 6.100000 | 2013.000000 | 0.000000e+00 | 0.000000e+00 |
| **Zack Snyder** | 21624.000000 | 2.199367 | 1.016667e+08 | 2.671562e+08 | 121.666667 | 1988.666667 | 6.566667 | 2009.000000 | 1.004855e+08 | 2.664120e+08 |
| **Zackary Adler** | 198630.000000 | 0.140723 | 1.250000e+06 | 0.000000e+00 | 101.500000 | 13.000000 | 5.350000 | 2011.000000 | 1.149999e+06 | 0.000000e+00 |
| **Zak Hilditch** | 199575.000000 | 0.276845 | 0.000000e+00 | 0.000000e+00 | 87.000000 | 86.000000 | 6.100000 | 2014.000000 | 0.000000e+00 | 0.000000e+00 |
| **Zak Knutson** | 299969.000000 | 0.293107 | 0.000000e+00 | 0.000000e+00 | 41.000000 | 15.000000 | 7.700000 | 2014.000000 | 0.000000e+00 | 0.000000e+00 |
| **Zak Levitt** | 114887.000000 | 0.004575 | 0.000000e+00 | 0.000000e+00 | 70.000000 | 11.000000 | 8.000000 | 2012.000000 | 0.000000e+00 | 0.000000e+00 |
| **Zak Penn** | 284470.000000 | 0.072514 | 0.000000e+00 | 0.000000e+00 | 66.000000 | 65.000000 | 6.700000 | 2014.000000 | 0.000000e+00 | 0.000000e+00 |
| **Zal Batmanglij** | 87155.500000 | 0.412176 | 3.250000e+06 | 1.200755e+06 | 100.500000 | 112.500000 | 6.000000 | 2012.000000 | 3.042108e+06 | 1.123947e+06 |
| **Zalman King** | 87409.000000 | 0.551235 | 2.333333e+06 | 5.157990e+05 | 103.666667 | 16.000000 | 4.433333 | 1996.666667 | 4.104298e+06 | 9.510771e+05 |
| **Zana Briski|Ross Kauffman** | 1392.000000 | 0.001117 | 3.500000e+05 | 3.515061e+06 | 85.000000 | 23.000000 | 6.400000 | 2004.000000 | 4.040559e+05 | 4.057946e+06 |
| **Zatella Beatty** | 266442.000000 | 0.173667 | 0.000000e+00 | 0.000000e+00 | 97.000000 | 14.000000 | 7.200000 | 2014.000000 | 0.000000e+00 | 0.000000e+00 |
| **Zeke Norton** | 194889.000000 | 0.549812 | 0.000000e+00 | 0.000000e+00 | 76.500000 | 30.000000 | 5.900000 | 2013.000000 | 0.000000e+00 | 0.000000e+00 |
| **Zev Berman** | 13548.000000 | 0.201639 | 0.000000e+00 | 0.000000e+00 | 105.000000 | 26.000000 | 5.300000 | 2007.000000 | 0.000000e+00 | 0.000000e+00 |
| **Zhao Lin-Shan** | 136701.000000 | 0.438730 | 0.000000e+00 | 0.000000e+00 | 107.000000 | 14.000000 | 5.600000 | 2012.000000 | 0.000000e+00 | 0.000000e+00 |
| **Ziad Doueiri** | 121791.000000 | 0.156925 | 0.000000e+00 | 0.000000e+00 | 102.000000 | 18.000000 | 5.800000 | 2012.000000 | 0.000000e+00 | 0.000000e+00 |
| **Zoe R. Cassavetes** | 13337.000000 | 0.266572 | 2.000000e+06 | 0.000000e+00 | 97.000000 | 11.000000 | 6.600000 | 2007.000000 | 2.103337e+06 | 0.000000e+00 |
| **ÃdÃÂ¡m CsÃÂ¡szi** | 254005.000000 | 0.109921 | 0.000000e+00 | 0.000000e+00 | 105.000000 | 12.000000 | 7.200000 | 2014.000000 | 0.000000e+00 | 0.000000e+00 |
| **Ãlex de la Iglesia** | 26014.250000 | 0.222866 | 6.775000e+06 | 8.046665e+05 | 101.750000 | 48.500000 | 6.025000 | 2003.750000 | 7.051741e+06 | 8.046665e+05 |
| **Ãlvaro de ArmiÃÂ±ÃÂ¡n** | 28499.000000 | 0.294762 | 6.500000e+06 | 0.000000e+00 | 88.000000 | 12.000000 | 4.600000 | 2009.000000 | 6.606603e+06 | 0.000000e+00 |
| **ÃâzgÃÂ¼r Yildirim** | 8064.000000 | 0.064319 | 0.000000e+00 | 0.000000e+00 | 92.000000 | 16.000000 | 6.300000 | 2008.000000 | 0.000000e+00 | 0.000000e+00 |
| **Ãâ¡aÃÅ¸an Irmak** | 13393.000000 | 0.265008 | 0.000000e+00 | 0.000000e+00 | 108.000000 | 12.000000 | 8.200000 | 2005.000000 | 0.000000e+00 | 0.000000e+00 |
| **Ãâ°mile Gaudreault** | 321.000000 | 0.276911 | 4.361898e+06 | 3.031801e+06 | 92.000000 | 13.000000 | 5.200000 | 2003.000000 | 5.170387e+06 | 3.593753e+06 |
| **Ãâ°ric Besnard** | 52349.000000 | 0.081724 | 0.000000e+00 | 0.000000e+00 | 100.000000 | 10.000000 | 5.100000 | 2010.000000 | 0.000000e+00 | 0.000000e+00 |
| **Ãâ°tienne Chatiliez** | 16637.333333 | 0.102346 | 4.455000e+06 | 0.000000e+00 | 101.333333 | 24.666667 | 6.700000 | 1994.666667 | 6.374957e+06 | 0.000000e+00 |
| **Ãâ¬lex Pastor|David Pastor** | 25769.000000 | 0.619369 | 0.000000e+00 | 5.802422e+06 | 84.000000 | 143.000000 | 5.700000 | 2009.000000 | 0.000000e+00 | 5.897584e+06 |

5067 rows × 10 columns

In [17]:

*#copying the original file'tmdb-movies.csv' data*

df\_popularity = df.copy()

In [18]:

*#list the most popular director.*

dir\_mean\_df['popularity'].nlargest(20)

Out[18]:

director

Colin Trevorrow 16.696886

Joe Russo|Anthony Russo 12.971027

Chad Stahelski|David Leitch 11.422751

Don Hall|Chris Williams 8.691294

Juno John Lee 8.411577

Kyle Balda|Pierre Coffin 7.404165

Alan Taylor 6.883129

Peter Richardson 6.668990

Pete Docter 6.326804

Christopher Nolan 6.195521

Alex Garland 6.118847

Patrick Tatopoulos 5.806897

Wes Ball 5.553082

Dan Gilroy 5.522641

Lilly Wachowski|Lana Wachowski 5.331930

James Gunn 5.225378

Bob Peterson|Pete Docter 4.908902

J.J. Abrams 4.800957

Alejandro GonzÃÂ¡lez IÃÂ±ÃÂ¡rritu 4.793536

Roger Allers|Rob Minkoff 4.782688

Name: popularity, dtype: float64

In [19]:

*#The graph is showing Colin Trevorrow is most popular but he didn't have the most realeases.*

*#df.plot(x='director', y='popularity')*

temp1 = dir\_mean\_df['popularity'].nlargest(10)

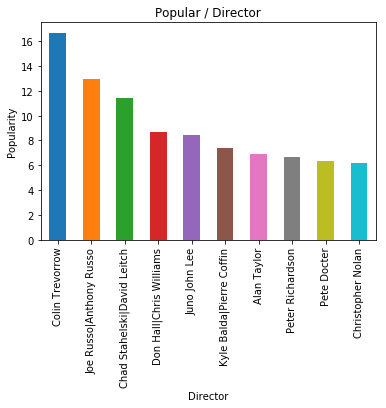
plt.figure()

temp1.plot(kind='bar', title="Popular / Director")

plt.xlabel("Director")

plt.ylabel("Popularity")

plt.show()

In [20]:

*#copying the original file'tmdb-movies.csv' data*

df\_budget = df.copy()

In [21]:

*#3-Which director is having the highest budget? Sngmoo Lee (scal is $(100.000000)*

*# The data below showing the highest budget was refer to Sngmoo Lee but not leaner with popularity and releases.*

dir\_mean\_df['budget'].nlargest(20)

Out[21]:

director

Sngmoo Lee 425000000.0

Andrew Stanton 220000000.0

Dan Scanlon 200000000.0

John Lasseter|Brad Lewis 200000000.0

Lee Unkrich 200000000.0

Mark Andrews|Steve Purcell|Brenda Chapman 185000000.0

Robert Stromberg 180000000.0

Bob Peterson|Pete Docter 175000000.0

Carl Rinsch 175000000.0

Pete Docter 175000000.0

Rob Letterman|Conrad Vernon 175000000.0

Joe Russo|Anthony Russo 170000000.0

Rupert Sanders 170000000.0

Don Hall|Chris Williams 165000000.0

Rich Moore 165000000.0

Alan Taylor 162500000.0

Chris Miller|Raman Hui 160000000.0

David Yates 155000000.0

Andrew Adamson|Kelly Asbury|Conrad Vernon 150000000.0

Brad Bird|Jan Pinkava 150000000.0

Name: budget, dtype: float64

In [22]:

*# The graph below is showing the highest budget was refer to Sngmoo Lee but not leaner with popularity and releases.(scal is $(100.000000)*

temp1 = dir\_mean\_df['budget'].nlargest(20)

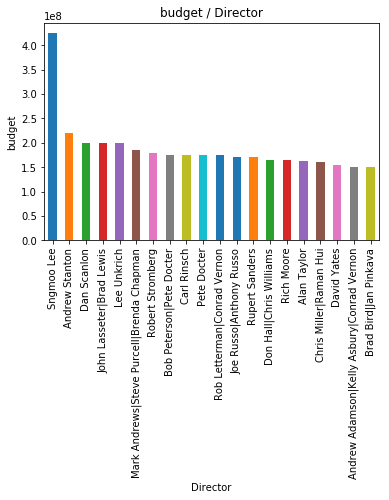
plt.figure()

temp1.plot(kind='bar', title="budget / Director")

plt.xlabel("Director")

plt.ylabel("budget")

plt.show()

In [23]:

*#copying the original file'tmdb-movies.csv' data*

df\_vote\_count = df.copy()

In [24]:

*#4-Which director has most vote count? Pop &Pet*

*# from the data below the vote count goes to Bob Peterson|Pete Docter and again not leaner with popularity , releases and budget.*

dir\_mean\_df['vote\_count'].nlargest(20)

Out[24]:

director

Bob Peterson|Pete Docter 4225.000000

Don Hall|Chris Williams 4185.000000

Christopher Nolan 4175.900000

Pierre Coffin|Chris Renaud 4143.500000

Pete Docter 3935.000000

Lee Unkrich|David Silverman|Pete Docter 3899.000000

Joe Russo|Anthony Russo 3848.000000

Shane Black 3704.500000

Andrew Stanton|Lee Unkrich 3692.000000

Lilly Wachowski|Lana Wachowski 3621.666667

Rich Moore 3492.000000

Roger Allers|Rob Minkoff 3489.000000

Mark Andrews|Steve Purcell|Brenda Chapman 3206.000000

Marc Webb 3142.666667

J.J. Abrams 2992.400000

Colin Trevorrow 2951.500000

Robert Stromberg 2928.000000

Lee Unkrich 2924.000000

Kyle Balda|Pierre Coffin 2893.000000

Joseph Kosinski 2877.500000

Name: vote\_count, dtype: float64

In [25]:

*#The chart below showing the vote count goes to Bob Peterson|Pete Docter and again not leaner with popularity , releases and budget.*

temp1 = dir\_mean\_df['vote\_count'].nlargest(20)

plt.figure()

temp1.plot(kind='bar', title="vote\_count /Director")

plt.xlabel("Director")

plt.ylabel("vote\_count")

plt.show()

In [26]:

*#copying the original file'tmdb-movies.csv' data*

df\_vote\_average = df.copy()

In [27]:

*#5-Which director has the most vote average? Mark Cousins*

*# the data and the chart below showing Mark Cousins got the most vote average and not leaner with popularity, release, budget and vote count .*

dir\_mean\_df['vote\_average'].nlargest(20)

Out[27]:

director

Mark Cousins 9.20

Derek Frankowski 8.80

David Mallet 8.70

Curt Morgan 8.50

James Payne 8.50

Martin Scorsese|Michael Henry Wilson 8.50

Saul Swimmer 8.50

Anthony Mandler 8.40

D.A. Pennebaker|David Dawkins|Chris Hegedus 8.40

Jorge RamÃÂ­rez SuÃÂ¡rez 8.40

Sam Dunn|Scot McFadyen 8.35

Jennifer Siebel Newsom 8.35

Andy Saunders 8.30

Claude Lanzmann 8.30

Dean Wellins 8.30

Jeff Feuerzeig 8.30

Al Reinert 8.20

Damien Chazelle 8.20

Evgeny Afineevsky 8.20

Ãâ¡aÃÅ¸an Irmak 8.20

Name: vote\_average, dtype: float64

In [28]:

*#the chart below showing Mark Cousins got the most vote average and not leaner with popularity, release, budget and vote count .*

temp1 = dir\_mean\_df['vote\_average'].nlargest(20)

plt.figure()

temp1.plot(kind='bar', title="vote\_average / Director")

plt.xlabel("Director")

plt.ylabel("vote\_average")

plt.show()

In [29]:

*# the chart below is showing the director Mark Cousine have the longest run time movie.*

temp1 = dir\_mean\_df['runtime'].nlargest(20)

plt.figure()

temp1.plot(kind='bar', title="runtime /Director")

plt.xlabel("Director")

plt.ylabel("runtime")

plt.show()

In [30]:

*#The hist plot below showing each distrebution frequency of score occurrences for popularity collected by directors.*

dir\_mean\_df['popularity'].hist();

plt.xlabel('Popularity bin range')

plt.ylabel('Counts')

plt.title('Popularity Over the Years');

In [31]:

*# the hist plot below showing higher frequencies in the middle and The distribution is skewed to the right of the value spectrum .*

*#and then the bar heights taper off as we move away from the middle.*

dir\_mean\_df['vote\_average'].hist(bins=100);

plt.xlabel('vote\_average bin range')

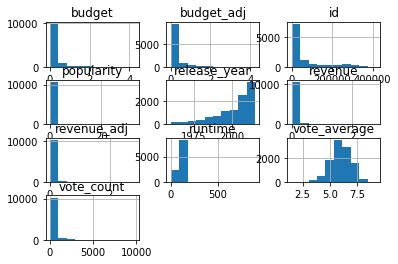
plt.ylabel('Counts')

plt.title('vote\_average Over the Years');

In [32]:

*#The hist below showing overviw picure for possible trends visually.*

df\_dir.hist();

In [33]:

*# The chart below showing the correlation between the popularity and the revenue for the directors espacily for lower value .*

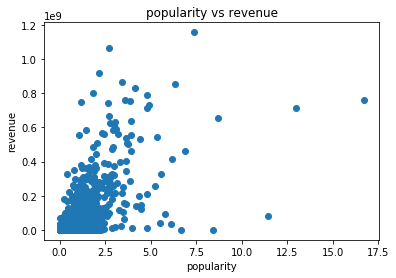
*# the relation is significant around 0.0- 2.5.*

plt.scatter(x=dir\_mean\_df['popularity'], y=dir\_mean\_df['revenue']);

plt.xlabel('popularity')

plt.ylabel('revenue')

plt.title('popularity vs revenue');

In [34]:

*# directors with maximum revenue ratings*

dir\_mean\_df['revenue'].nlargest(10)

Out[34]:

director

Kyle Balda|Pierre Coffin 1.156731e+09

Lee Unkrich 1.063172e+09

Andrew Adamson|Kelly Asbury|Conrad Vernon 9.198388e+08

Andrew Stanton|Lee Unkrich 8.646260e+08

Pete Docter 8.537086e+08

David Yates 8.308591e+08

Chris Miller|Raman Hui 7.989582e+08

Roger Allers|Rob Minkoff 7.882418e+08

Colin Trevorrow 7.587683e+08

Robert Stromberg 7.584104e+08

Name: revenue, dtype: float64

In [35]:

*# inspecting the statistics for the budget of directors as a whole*

dir\_mean\_df['budget'].describe()

Out[35]:

count 5.067000e+03

mean 9.132429e+06

std 2.222369e+07

min 0.000000e+00

25% 0.000000e+00

50% 0.000000e+00

75% 8.000000e+06

max 4.250000e+08

Name: budget, dtype: float64

In [36]:

*# cheking the value again.*

dir\_mean\_df[880:900]

Out[36]:

|  | **id** | **popularity** | **budget** | **revenue** | **runtime** | **vote\_count** | **vote\_average** | **release\_year** | **budget\_adj** | **revenue\_adj** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **director** |  |  |  |  |  |  |  |  |  |  |
| **Clive A. Smith** | 21356.00 | 0.195539 | 8000000.0 | 0.000000e+00 | 77.000000 | 14.000000 | 6.100000 | 1983.000000 | 1.751450e+07 | 0.000000e+00 |
| **Clive Barker** | 12819.00 | 0.321329 | 4000000.0 | 7.808794e+06 | 105.000000 | 86.666667 | 6.200000 | 1990.666667 | 6.758988e+06 | 1.424664e+07 |
| **Clive Donner** | 23138.25 | 0.250399 | 0.0 | 0.000000e+00 | 101.250000 | 19.500000 | 5.925000 | 1977.750000 | 0.000000e+00 | 0.000000e+00 |
| **Clyde Geronimi|Hamilton Luske|Wolfgang Reitherman** | 12230.00 | 2.631987 | 4000000.0 | 2.158800e+08 | 79.000000 | 913.000000 | 6.600000 | 1961.000000 | 2.917944e+07 | 1.574815e+09 |
| **ClÃÂ©ment Virgo** | 14174.00 | 1.203532 | 2200000.0 | 0.000000e+00 | 93.000000 | 36.000000 | 5.000000 | 2005.000000 | 2.456439e+06 | 0.000000e+00 |
| **Cody Calahan** | 207780.00 | 0.258887 | 0.0 | 0.000000e+00 | 90.000000 | 19.000000 | 4.300000 | 2013.000000 | 0.000000e+00 | 0.000000e+00 |
| **Cody Cameron** | 51170.00 | 0.484243 | 0.0 | 0.000000e+00 | 75.000000 | 64.000000 | 5.300000 | 2010.000000 | 0.000000e+00 | 0.000000e+00 |
| **Cody Cameron|Kris Pearn** | 109451.00 | 0.932599 | 78000000.0 | 2.483846e+08 | 95.000000 | 637.000000 | 6.400000 | 2013.000000 | 7.301060e+07 | 2.324963e+08 |
| **Coky Giedroyc** | 148110.50 | 0.131463 | 0.0 | 0.000000e+00 | 159.000000 | 18.500000 | 6.600000 | 2008.000000 | 0.000000e+00 | 0.000000e+00 |
| **Coley Sohn** | 138907.00 | 0.034628 | 0.0 | 0.000000e+00 | 87.000000 | 16.000000 | 5.300000 | 2012.000000 | 0.000000e+00 | 0.000000e+00 |
| **Colin Chilvers|Jerry Kramer|Jim Blashfield** | 13192.00 | 0.321068 | 22000000.0 | 0.000000e+00 | 93.000000 | 73.000000 | 6.100000 | 1988.000000 | 4.056560e+07 | 0.000000e+00 |
| **Colin Eggleston** | 39995.00 | 0.047645 | 0.0 | 0.000000e+00 | 92.000000 | 17.000000 | 6.100000 | 1978.000000 | 0.000000e+00 | 0.000000e+00 |
| **Colin Hanks** | 324253.00 | 0.143042 | 0.0 | 0.000000e+00 | 94.000000 | 22.000000 | 7.000000 | 2015.000000 | 0.000000e+00 | 0.000000e+00 |
| **Colin Higgins** | 17172.00 | 0.329755 | 0.0 | 0.000000e+00 | 113.333333 | 30.333333 | 6.266667 | 1980.000000 | 0.000000e+00 | 0.000000e+00 |
| **Colin Levy** | 45745.00 | 0.609669 | 550000.0 | 0.000000e+00 | 14.000000 | 51.000000 | 7.600000 | 2010.000000 | 5.500000e+05 | 0.000000e+00 |
| **Colin Minihan** | 265016.00 | 0.339749 | 0.0 | 0.000000e+00 | 101.000000 | 100.000000 | 4.600000 | 2014.000000 | 0.000000e+00 | 0.000000e+00 |
| **Colin Strause|Greg Strause** | 21562.00 | 0.997498 | 25000000.0 | 5.430905e+07 | 97.000000 | 363.500000 | 4.800000 | 2008.500000 | 2.603337e+07 | 5.538885e+07 |
| **Colin Teague** | 300803.00 | 1.144808 | 0.0 | 0.000000e+00 | 97.000000 | 59.000000 | 4.500000 | 2015.000000 | 0.000000e+00 | 0.000000e+00 |
| **Colin Theys** | 271665.00 | 0.171104 | 0.0 | 0.000000e+00 | 100.000000 | 10.000000 | 4.900000 | 2014.000000 | 0.000000e+00 | 0.000000e+00 |
| **Colin Trevorrow** | 109864.50 | 16.696886 | 75375000.0 | 7.587683e+08 | 104.500000 | 2951.500000 | 6.600000 | 2013.500000 | 6.935612e+07 | 6.981261e+08 |

In [37]:

*# inspecting the statistics for the runtime of directors as a whole.*

dir\_mean\_df['runtime'].describe()

Out[37]:

count 5067.000000

mean 97.899040

std 33.472968

min 0.000000

25% 89.000000

50% 96.000000

75% 106.000000

max 900.000000

Name: runtime, dtype: float64