Data Article

Films nominated for Academy Award for Best Picture in the 2010s decade dataset

Faza Nanda Yudistira, Leo Harnadi Marlin, Muhammad Naufal Dzakki Rauf, Pandya Athallah Erlambang

Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada, Indonesia

A R T I C L E I N F O

A B S T R A C T

The dataset presented in this paper is a description of films nominated for the Academy Award for Best Picture from the year 2010 to 2019. The data were gathered by means of web scraping from the Academy Award for Best Picture article on Wikipedia. The tools used for the data collection are but not limited to Notebook, Python 3, Python BeautifulSoup library, and Python Requests library. The dataset consists of but is not limited to, the film's title, director, producer, score composer, actors, and etc. The dataset can provide insights as to how the movie industry landscape transforms over time and will provide means for people to do data analysis of films nominated for this specific category of the Oscars.

*Corresponding author(s).
E-mail address*: fazananda@mail.ugm.ac.id (Yudistira, F), leo.h.m@mail.ugm.ac.id (Marlin, L), mnaufaldr01@mail.ugm.ac.id (Rauf, M), pandya.ae@mail.ugm.ac.id (Erlambang, P)

Specifications Table

Subject	Film
Specific subject area	Films nominated for the Academy Awards, Academy Award, Academy Award for Best Picture, 2010s film
Type of data	Primary data and tables
How data were acquired	Data was gathered by means of web scraping from the Academy Award for Best Picture article on Wikipedia
Data format	Raw, analyzed
Parameter for data collection	The dataset was obtained from the Academy Award for Best Picture article. Only films released in the 2010s decade were collected.
Description of data collection	The data collection was done by scraping the article using Notebook (Python 3 IDE) with the BeautifulSoup and Requests library, the scraped data was then exported to JSON and cleaned. The resulting cleaning produced clean data with minimal anomalies and was finally exported to CSV.
Data source location	Source: World Wide Web Website: Wikipedia Article: Academy Award for Best Picture
Data accessibility	The dataset is available for use at the author's GitHub repository, link to the dataset: pandya-ae/SDE/Academy Award for Best Picture Dataset/dataset-2010s.csv

Value of the Data

- The dataset is important because it provides crucial data for data analysis for researchers and alike on this subject.
- The dataset will be useful for researchers and alike doing data analysis who wants to gain unique insights and useful information.
- The dataset will be valuable to researchers and alike wanting to examine the relationship between certain categorical variables of the dataset.
- The details of the analyzed dataset will shed a light to new insights and a unique perspective on the movie industry's ever-changing landscape over the years.

1. Data Description

This data set provides insightful information about films that were nominated for the Academy Awards for the best picture category in the time range from the year 2010 until 2019 based on the web scraping of information provided by Wikipedia. The data provided here includes details of the movies, such as the director, producers, actors and actresses, production companies, budget, box office gains, and scores given by movie critics. The code, written in Python, will also be provided as a supplementary file.

Table 1
Directors with the most nominated films.

Director	Count
Steven Spielberg	4
David O. Russell, Martin Scorsese	3
Adam McKay, Alejandro G. Iñárritu, Alexander Payne, Alfonso Cuarón, Christopher Nolan, Damien Chazelle, Greta Gerwig, Quentin Tarantino, Tom Hooper	2
Ang Lee, Ava DuVernay, Barry Jenkins, Ben Affleck, Benh Zeitlin, Bennett Miller, Bong Joon-ho, Bradley Cooper, Bryan Singer, Clint Eastwood, Danny Boyle, Darren Aronofsky, David Fincher, David Mackenzie, Debra Granik, Denis Villeneuve, Denzel Washington, Garth Davis, George Miller, Guillermo del Toro, James Mangold, James Marsh, Jean-Marc Vallée, Joe Wright, John Crowley, Jordan Peele, Kathryn Bigelow, Kenneth Lonergan, Lee Unkrich, Lenny Abrahamson, Lisa Cholodenko, Luca Guadagnino, Martin McDonagh, Mel Gibson, Michael Haneke, Michel Hazanavicius, Morten Tyldum, Noah Baumbach, Paul Greengrass, Paul Thomas Anderson, Peter Farrelly, Richard Linklater, Ridley Scott, Ryan Coogler, Sam Mendes, Spike Jonze, Spike Lee, Stephen Daldry, Stephen Frears, Steve McQueen, Taika Waititi, Tate Taylor, Terrence Malick, Theodore Melfi, Todd Phillips, Joel Coen & Ethan Coen	1

Table 2
Producers with the most nominated films.

Producer	Count
Scott Rudin	7
Dede Gardner	6
Jeremy Kleiner, Brad Pitt	5
Steven Spielberg, Megan Ellison, Arnon Milchan	4

Tim Bevan, Michael De Luca, Martin Scorsese, Jason Blum, Eric Fellner, Emma Tillinger Koskoff, David Heyman, Brian Oliver, Bradley Cooper | 3

Todd Phillips, Steve Golin, Sean McKittrick, Peter Chernin, Marc Platt, Lisa Bruce, Kristie Macosko Krieger, Kathleen Kennedy, Jordan Peele, Jordan Horowitz, Jonathan Gordon, Jim Burke, Jenno Topping, James W. Skotchdopole, Iain Canning, Graham King, Gary Gilbert, Emma Thomas, Emile Sherman, Ed Guiney, Donna Gigliotti, Dana Brunetti, Christopher Nolan, Christian Colson, Bill Pohlad, Anthony McCarten, Amy Pascal, Alfonso Cuarón, Alejandro G. Iñárritu | 2

Yorgos Lanthimos, William D. Johnson, Will Ferrell, Wes Anderson, Vincent Landay, Veit Heiduschka, Tracey Seaward, Todd Lieberman, Todd Black, Timothy Headington, Thomas Langmann, Theodore Melfi, Terry Benedict, Teddy Schwarzman, Taika Waititi, Steven Rales, Steve McQueen, Steve Coogan, Stephen Tenenbaum, Stefan Arndt, Stacey Sher, Spike Lee, Spike Jonze, Simon Kinberg, Sidney Kimmel, Shawn Levy, Shaun Redick, Shannon McIntosh, Scott Franklin, Sarah Green, Sam Mendes, Ryan Kavanaugh, Ron Yerxa, Rodrigo Teixeira, Robin Swicord, Robert Lorenz, Robert De Niro, Robbie Brenner, Riza Aziz, Ridley Scott, Richard Suckle, Richard Linklater, Reginald Hudlin, Raymond Mansfield, Randall Emmett, Rachel Winter, Rachael Horovitz, Quentin Tarantino, PJ Voeten, Pippa Harris, Pilar Savone, Philippe Hellmann, Pharrell Williams, Peter Spears, Peter Morgan, Peter Farrelly, Peter Czernin, Peter Berg, Paul Thomas Anderson, Paul Tamasy, Paul Currie, Oprah Winfrey, Nora Grossman, Noah Baumbach, Nicole Rocklin, Nicolás Celis, Nick Vallelonga, Moon Yang-kwon, Mike Medavoy, Michel Litvak, Michael Sugar, Michael Schaefer, Michael Katz, Michael Gottwald, Michael Barnathan, Matt Damon, Mary Parent, Martin McDonagh, Mark Wahlberg, Mark Huffam, Mark Boal, Margaret Ménégoz, Marco Morabito, Lynette Howell Taylor, Luca Guadagnino, Letty Aronson, Leonardo DiCaprio, Lee Magiday, Lauren Beck, Kwak Sin-ae, Kimberly Steward, Kevin J. Walsh, Kevin J. Messick, Kevin Feige, Keith Redmon, Kathryn Bigelow, Julie Yorn, Josh Penn, Jonathan Sehring, Jon Peters, Johnny Depp, John Smithson, John Sloss, John Lesher, Joey McFarland, Joel Newton, Joel Coen, JoAnne Sellar, Jim Taylor, Jim Beach, Jeremy Dawson, Jeffrey Levy-Hinte, Jayne-Ann Tenggren, Jaume Roures, Jang Young-hwan, Jane Rosenthal, James Mangold, James Ivory, J. Miles Dale, Irwin Winkler, Ido Ostrowsky, Howard Rosenman, Helen Estabrook, Guillermo del Toro, Grant Hill, Grant Heslov, Graham Broadbent, Gil Netter, Gerald Chamales, George Miller, George Clooney, Gastón Pavlovich, Gareth Unwin, Gabrielle Tana, Gabriele Israilovici, Gabriela Rodríguez, Fred Berger, Finola Dwyer, Evelyn O'Neill, Ethan Coen, Emilie Georges, Eli Bush, Edward H. Hamm Jr., Douglas Urbanski, Doug Mitchell, Dorothy Aufiero, Denzel Washington, Denise Di Novi, Debra Hayward, David | 1

Womark, David   Permut, David   Linde, David   Lancaster, David Hoberman, David Gross, Darla K. Anderson, Danny Boyle, Daniela Taplin Lundberg, Daniel Lupi, Dan Levine, Dan Janvey, Clint Eastwood, Chris Moore, Chris Columbus, Chelsea Winstanley, Charles Roven, Charles B. Wessler, Celine   Rattray, Ceci   Dempsey, Ceán   Chaffin, Cathleen Sutherland, Carthew Neal, Carla Hacken, Cameron Mackintosh, Callum McDougall, Brunson Green, Bruce Davey, Bruce Cohen, Brian Hayes Currie, Bong Joon-ho, Blye Pagon Faust, Bill Mechanic, Bill Gerber, Ben Affleck, Arnold W. Messer, Anthony Katagas, Anne Rosellini, Angie Fielder, Ang   Lee, Andrew   Lazar, Amanda   Posey, Alix Madigan-Yorkin, Alexander Payne, Albert Berger, Adele Romanski, Adam McKay, Aaron Ryder

Table 3

Actors with the most nominated films.

Actor	Count
Tom Hanks, Leonardo DiCaprio, Christian Bale, Brad Pitt, Amy Adams	5
Tom Hardy, Saoirse Ronan, Jessica Chastain, Emma Stone, Bradley Cooper	4
Viola Davis, Tracy Letts, Timothée Chalamet, Scarlett Johansson, Sam Rockwell, Robert De Niro, Octavia Spencer, Michael Stuhlbarg, Michael Keaton, Matt Damon, John Goodman, Jennifer Lawrence, Bob Odenkirk, Benedict Cumberbatch	3
Tom Wilkinson, Steve Carell, Stephen McKinley Henderson, Sarah Paulson, Sandra Bullock, Sacha Baron Cohen, Ryan Gosling, Rooney Mara, Rachel McAdams, Owen Wilson, Michael Peña, Meryl Streep, Matthew McConaughey, Mark Strong, Mark Rylance, Mark Ruffalo, Marion Cotillard, Margot Robbie, Mahershala Ali, Lupita Nyongo, Lucas Hedges, Laura Dern, Kyle Chandler, Jude Law, Joseph Gordon-Levitt, Jonah Hill, John Hawkes, Joaquin Phoenix, Jeremy Renner, Jeff Bridges, Jean Dujardin, Janelle Monáe, Helena Bonham Carter, Harvey Keitel, George Clooney, Forest Whitaker, Emily Watson, Edward Norton, Eddie Redmayne, Domhnall Gleeson, Divian Ladwa, David Thewlis, Daniel Kaluuya, Daniel Day-Lewis, Colin Firth, Cillian Murphy, Chiwetel Ejiofor, Bruce Dern, Bradley Whitford, Armie Hammer, Andrew Garfield, Amy Ryan, Alessandro Nivola, Alan Alda, Al Pacino, Adrien Brody, Adam Driver	2
Zoë Kravitz, Zazie Beetz, Zach Galifianakis, Yalitza Aparicio, Woody Harrelson, Winston Duke, Winona Ryder, William H. Macy, Willem Dafoe, Will Poulter, Will Forte, Walton Goggins, Wallace Shawn, Vincent Cassel, Vince Vaughn, Viggo Mortensen, Victoire Du Bois, Vicky Krieps, Tzi Ma, Tyler Perry, Trevante Rhodes, Topher Grace, Tony Revolori, Tommy Lee Jones, Tom Hollander, Tom Glynn-Carney, Tom Berenger, Timothy Spall, Timothy Olyphant, Tim Roth, Tim Allen, Thomasin McKenzie, Thomas Horn, Teresa Palmer, Tate	1

Taylor, Taraji P. Henson, Tannishtha Chatterjee, Taika Waititi, Tabu, Suraj Sharma, Sunny Pawar, Steve Coogan, Stephen Root, Stephen Merchant, Stephen Graham, Stephen Dillane, Stanley Tucci, Stacy Keach, Song Kang-ho, Simon McBurney, Sienna Miller, Sheryl Lee, Shailene Woodley, Sebastian Stan, Sean Penn, Sean Bridgers, Sean Bean, Saniyya Sidney, Samuel L. Jackson, Sam Worthington, Sam Elliott, Sally Hawkins, Sally Field, Russell Hornsby, Russell Crowe, Rosie Huntington-Whiteley, Rosemarie DeWitt, Rory Kinnear, Ronald Pickup, Roman Griffin Davis, Rob Reiner, Riley Keough, Richard Madden, Richard Jenkins, Rebel Wilson, Ray Winstone, Ray Romano, Ray Liotta, Rami Malek, Ralph Fiennes, Rafe Spall, Rachel Weisz, Rachel Griffiths, Quvenzhané Wallis, Priyanka Bose, Philip Seymour Hoffman, Peter Mullan, Peter Dinklage, Penelope Ann Miller, Paul Reiser, Paul Giamatti, Paul Dano, Patricia Arquette, Park So-dam, Oprah Winfrey, Olivia Wilde, Olivia Colman, Odeya Rush, Niels Arestrup, Nicole Kidman, Nicholas Hoult, Ned Beatty, Nawazuddin Siddiqui, Natalie Portman, Naomie Harris, Naomi Watts, Mykelti Williamson, Missi Pyle, Miles Teller, Mila Kunis, Mike Myers, Michelle Williams, Michael Sheen, Michael Shannon, Michael Parks, Michael Gambon, Michael Fassbender, Michael Caine, Michael B. Jordan, Mia Wasikowska, Merritt Wever, Melissa Leo, Maxine Peake, Max von Sydow, Max Minghella, Matthew Rhys, Matthew Goode, Mathieu Amalric, Martin Freeman, Mark Wahlberg, Marina de Tavira, Margaret Qualley, Malcolm McDowell, Luke Bracey, Lucy Boynton, Louis Garrel, Louis C. K., Lorelei Linklater, Lois Smith, Linda Cardellini, Lily Rabe, Lily James, Liev Schreiber, Letitia Wright, Lesley Manville, Lee Sun-kyun, Lee Jung-eun, Léa Seydoux, Laurie Metcalf, Laura Harrier, Lakeith Stanfield, Lady Gaga, Kristin Scott Thomas, Kristen Wiig, Kirsten Dunst, Kevin Costner, Kevin Breznahan, Kerry Washington, Kenneth Branagh, Ken Watanabe, Ken Davitian, Keira Knightley, Kaushik Sen, Kathy Bates, Kate Mara, Justin Timberlake, June Squibb, Julie Walters, Julie Hagerty, Julianne Moore, Judy Greer, Judi Dench, Jovan Adepo, Josh Hutcherson, Josh Brolin, Jon Favreau, Jon Bernthal, John Slattery, John Ratzenberger, John Morris, John Legend, John David Washington, Joel Murray, Joel Edgerton, Joe Pesci, Joe Mazzello, Jodi Benson, Joan Cusack, Joan Allen, Jim Parsons, Jim Broadbent, Jharrel Jerome, Jesse Plemons, Jesse Eisenberg, Jeremy Irvine, Jennifer Garner, Jennifer Ehle, Jeffrey Wright, Jeff Goldblum, Jeff Daniels, Jean-Louis Trintignant, Jason Schwartzman, Jason Clarke, Jared Leto, Jang Hye-jin, Jamie Foxx, James Spader, James Remar, James Norton, James Franco, James DArcy, James Cromwell, Jag Huang, Jacob Tremblay, Jacki Weaver, Jack Lowden, J. K. Simmons, Isabelle Huppert, Irrfan Khan, Hugo Weaving, Hugh Keays-Byrne, Hugh Jackman, Harry Styles, Hal Holbrook, Hailee Steinfeld, Gwilym Lee, Guy Pearce, Giovanni Ribisi, Gil Birmingham, Gérard Depardieu, George MacKay, Geoffrey Rush, Gary Oldman, Garret Dillahunt, Frances McDormand, Frances Conroy, Florence Pugh, Fionn Whitehead, Felicity Jones, F. Murray Abraham, Ethan Hawke, Esther Garrel, Estelle Harris, Emory Cohen, Emmanuelle

Riva, Emma Watson, Emily Mortimer, Emile Hirsch, Elliot Page, Ellar Coltrane, Eliza Scanlen, Ed Lauter, Dwight Henry, Doug Jones, Don Rickles, Don Johnson, Dev Patel, Derek Jacobi, Denzel Washington, Dennis Christopher, Deepti Naval, Dean-Charles Chapman, David Wenham, David Strathairn, David Oyelowo, Dave Chappelle, Danai Gurira, Dale Dickey, Dakota Fanning, Cuba Gooding Jr., Courtney Eaton, Claire Duburcq, Christoph Waltz, Chris Tucker, Chris Pine, Choi Woo-shik, Cho Yeo-jeong, Chloë Grace Moretz, Charlize Theron, Charlie Cox, Charles Dance, Chadwick Boseman, Catherine Keener, Casey Affleck, Carmen Ejogo, Carla Bruni, Caleb Landry Jones, Bryce Dallas Howard, Bryan Cranston, Bruce Greenwood, Brie Larson, Bobby Cannavale, Bill Murray, Beth Grant, Bérénice Bejo, Ben Mendelsohn, Ben Kingsley, Ben Hardy, Ben Foster, Ben Affleck, Beau Bridges, Beanie Feldstein, Barry Pepper, Barry Keoghan, Barkhad Abdi, Barbara Hershey, Austin Butler, Ashton Sanders, Asa Butterfield, Anupam Kher, Annette Bening, Anne Hathaway, Anna Paquin, Angela Bassett, Aneurin Barnard, Andy Serkis, Andrew Scott, Andrew Dice Clay, Andrea Riseborough, André Holland, Amira Casar, Amanda Seyfried, Allison Williams, Allison Janney, Allen Leech, Alison Pill, Alfre Woodard, Alfie Allen, Alan Arkin, Aksel Hennie, Aidan Gillen, Abhishek Bharate, Abbie Cornish, Abbey Lee, Aaron McCusker

Table 4
Score Composers with the most nominated films.

Score Composer	Count
Alexandre Desplat	10
Hans Zimmer	4
Carter Burwell, John Williams, Nicholas Britell, Thomas Newman	3
Danny Elfman, Howard Shore, Jóhann Jóhannsson, Justin Hurwitz, Michael Brook, Mychael Danna, Randy Newman	2
A. R. Rahman, Antonio Sánchez, Arcade Fire, Claude-Michel Schönberg, Clint Mansell, Dario Marianelli, Dickon Hinchliffe, Harry Gregson-Williams, Henry Jackman, Hildur Guðnadóttir, Jason Moran, Jon Brion, Jonny Greenwood, Jung Jae-il, Junkie XL, Kris Bowers, Lesley Barber, Ludovic Bource, Ludwig Göransson, Marcelo Zarvos, Mark Orton, Michael Abels, Michael Giacchino, Queen, Rupert Gregson-Williams, Stephen Rennicks, Steven Price, Terence Blanchard, Dan Romer, Benh Zeitlin, Pharrell Williams, Benjamin Wallfisch, Marco Beltrami, Buck Sanders, Nick Cave, Warren Ellis, Ryuichi Sakamoto, Alva Noto, Trent Reznor, Atticus Ross, Dustin O'Halloran, Hauschka	1

Table 5
Directors of Photography (Cinematographer) with the most nominated films.

Director of Photography	Count
Emmanuel Lubezki, Janusz Kamiński	4
Danny Cohen, Greig Fraser, Hoyte van Hoytema, Phedon Papamichael, Robbie Ryan, Robert Richardson, Rodrigo Prieto	3
Barry Ackroyd, Bradford Young, Darius Khondji, Linus Sandgren, Masanobu Takayanagi, Matthew Libatique, Roger Deakins, Wally Pfister, Yves Bélanger	2
Alfonso Cuarón, Ben Davis, Ben Richardson, Benoît Delhomme, Bruno Delbonnel, Charlotte Bruus Christensen, Chayse Irvin, Chris Menges, Claudio Miranda, Dan Laustsen, Dariusz Wolski, Giles Nuttgens, Guillaume Schiffman, Hong Kyung-pyo, Igor Jadue-Lillo, James Laxton, Jeff Cronenweth, Jody Lee Lipes, John Seale, Lawrence Sher, Mandy Walker, Michael McDonough, Mihai Mălaimare Jr., Newton Thomas Sigel, Óscar Faura, Rachel Morrison, Robert Yeoman, Sam Levy, Sayombhu Mukdeeprom, Sean Bobbitt, Sean Porter, Sharone Meir, Simon Duggan, Stephen Goldblatt, Toby Oliver, Tom Stern, Yorick Le Saux	1

Table 6

Editors with the most nominated films.

Editor	Count
Michael Kahn	4
William Goldenberg, Thelma Schoonmaker, Lee Smith, Jay Cassidy, Hank Corwin	3
Valerio Bonelli, Tom Cross, Stephen Mirrione, Nick Houy, Kevin Tent, Joe Walker, Jennifer Lame, Jake Roberts, Hughes Winborne, Fred Raskin, Dylan Tichenor, Crispin Struthers, Alfonso Cuarón, Affonso Gonçalves	2
Yorgos Mavropsaridis, Yang Jin-mo, Walter Fasano, Tom McArdle, Tom Eagles, Tim Squyres, Tariq Anwar, Spencer Averick, Sidney Wolinsky, Sarah Broshar, Sandra Adair, Roderick Jaynes, Pietro Scalia, Peter Teschner, Patrick J. Don Vito, Pamela Martin, Nathan Nugent, Nat Sanders, Nadine Muse, Monika Willi, Michel Hazanavicius, Michael P. Shawver, Michael McCusker, Melanie Oliver, Martin Pensa, Mark Yoshikawa, Mark Sanger, Margaret Sixel, Kirk Baxter, Ken Schretzmann, Jon Harris, Jon Gregory, Joi McMillon, John Ottman, John Mac McMurphy, John Gilbert, Joel Cox, Jinx Godfrey, Jeffrey M. Werner, Jeff Groth, Jeff Buchanan, Jay Rabinowitz, Gregory Plotkin, Gary D. Roach, Eric Zumbrunnen, Douglas Crise, Debbie Berman, Daniel Rezende, Crockett Doob, Claire Simpson, Christopher Tellefsen, Christopher Rouse, Chris Dickens, Billy Weber, Barry Alexander Brown, Barney Pilling, Anne-Sophie Bion, Angus Wall, Andrew Weisblum, Andrew Buckland, Alisa Lepselter, Alexandre de	1

Franceschi, Alan Baumgarten, Adam Gough

Table 7
Distributors with the most nominated films.

Distributor	Count
20th Century Fox	12
Warner Bros. Pictures, Sony Pictures Releasing, Fox Searchlight Pictures	11
Universal Pictures	10
Paramount Pictures	9
Focus Features	7
Walt Disney Studios Motion Pictures, The Weinstein Company, Stage 6 Films, Lionsgate	4
Sony Pictures Classics, Netflix, A24	3
Walt Disney Studios, Roadside Attractions, Motion Pictures	2
X-Verleih, Warner Bros. Entertainment UK, Warner Bros., Village Roadshow Pictures, Transmission Films, StudioCanal, Pictureworks, Pathé Distribution, Paramount Vantage, Panorama Media, Open Road Films, Mongrel Media, Momentum Pictures, Mister Smith Entertainment, Mirror Releasing, Memento Films, Lionsgate Films, Les Films du Losange, IM Global, IFC Films, Icon Film Distribution, Filmladen, Espectáculos Fílmicos El Coyúl, Entertainment One Films, Entertainment One, Elevation Pictures Canada, Columbia Pictures, CJ Entertainment, CBS Films, Amazon Studios	1

Table 8
Nominated films with the longest running time.

Title	Running Time (minutes)
The Irishman	209
The Wolf of Wall Street	180
Django Unchained, Boyhood	165
Once Upon a Time in Hollywood	161
Les Misérables	158
Zero Dark Thirty	157
The Revenant	156
Ford v Ferrari	152

Lincoln	150
Inception	148
The Help, War Horse	146
Bridge of Spies, The Martian	141
The Tree of Life, Fences, Hacksaw Ridge	139
American Hustle	138
Manchester by the Sea, Marriage Story	137
A Star Is Born	136
Little Women, Roma	135
12 Years a Slave, Captain Phillips, Black Panther, Bohemian Rhapsody	134
Moneyball	133
American Sniper, Call Me by Your Name, Vice, Parasite	132
The Big Short, Phantom Thread, Green Book	130
Extremely Loud & Incredibly Close, Spotlight	129
La La Land, Selma, BlacKkKlansman	128
Life of Pi, Amour, Hidden Figures	127
Hugo, Her	126
Darkest Hour	125
The Theory of Everything, The Shape of Water	123
Silver Linings Playbook, Joker	122
The Social Network, Argo, Mad Max: Fury Road, The Favourite	120
Birdman or (The Unexpected Virtue of Ignorance), 1917, The King's Speech	119
Room, Lion	118
Dallas Buyers Club	117
The Fighter, Arrival, The Post	116
Three Billboards Outside Ebbing, Missouri; Nebraska, The Descendants	115
The Imitation Game	114
Brooklyn	112
Moonlight	111
True Grit	110

Black Swan, Jojo Rabbit	108
The Kids Are All Right, Whiplash	107
Dunkirk	106
Get Out	104
Toy Story 3	103
Hell or High Water	102
Winter's Bone, The Artist, The Grand Budapest Hotel	100
Philomena	98
127 Hours, Midnight in Paris, Lady Bird	94
Beasts of the Southern Wild	93
Gravity	91

Table 9
Nominated films with the highest budget.

Title	Budget
Black Panther, Toy Story 3	$200 million
Inception	$160 million
The Irishman	$159–250 million
Mad Max: Fury Road	$154.6–185.1 million
Hugo	$150–170 million
The Revenant	$135 million
Life of Pi	$120 million
The Martian	$108 million
The Wolf of Wall Street, Django Unchained	$100 million
Dunkirk	$100–150 million
Ford v Ferrari	$97.6 million
1917	$95 million
Once Upon a Time in Hollywood	$90–96 million
Gravity	$80–130 million
War Horse	$66–70 million
Lincoln	$65 million
Les Misérables	$61 million
Vice	$60 million
American Sniper	$59 million

Captain Phillips	$55 million
Joker	$55–70 million
Bohemian Rhapsody	$50–55 million
Moneyball, The Big Short, The Post	$50 million
Arrival	$47 million
Argo	$44.5 million
Zero Dark Thirty, Bridge of Spies, Hacksaw Ridge, American Hustle, Little Women, Extremely Loud & Incredibly Close, The Social Network	$40 million
True Grit	$38 million
A Star Is Born	$36 million
Phantom Thread	$35 million
The Tree of Life	$32 million
La La Land, Darkest Hour	$30 million
The Help, Hidden Figures, The Fighter, The Grand Budapest Hotel	$25 million
Fences	$24 million
Green Book	$23 million
Her	$23 million
Silver Linings Playbook	$21 million
12 Years a Slave	$20–22 million
Spotlight, Selma, The Descendants	$20 million
The Shape of Water	$19.5–20 million
Marriage Story, Birdman or (The Unexpected Virtue of Ignorance), 127 Hours	$18 million
Midnight in Paris	$17 million
Roma, BlacKkKlansman, The Theory of Everything, The Favourite, The King's Speech, The Artist	$15 million
Parasite	₩17.0 billion
The Imitation Game, Jojo Rabbit	$14 million
Nebraska	$13.5 million
Room, Black Swan	$13 million
Three Billboards Outside Ebbing, Missouri	$12–15 million
Lion, Hell or High Water, Philomena	$12 million
Brooklyn	$11 million

Lady Bird	$10 million
Manchester by the Sea	$9 million
Amour	$ 8.9 million
Dallas Buyers Club	$5 million
Get Out	$4.5 million
Boyhood, The Kids Are All Right	$4 million
Call Me by Your Name	$3.5 million
Whiplash	$3.3 million
Winter's Bone	$2 million
Beasts of the Southern Wild	$1.8 million
Moonlight	$1.5–4 million

Table 10
Highest grossing nominated films.

Title	Box Office
Black Panther	$1.348 billion
Joker	$1.074 billion
Toy Story 3	$1.067 billion
Bohemian Rhapsody	$905.2 million
Inception	$836.8 million
Gravity	$723.2 million
The Martian	$630.2 million
Life of Pi	$609 million
American Sniper	$547.4 million
The Revenant	$533 million
Dunkirk	$526.9 million
La La Land	$448.9 million
Les Misérables	$441.8 million
A Star Is Born	$436.2 million
The King's Speech	$427.4 million
Django Unchained	$425.4 million
The Wolf of Wall Street	$392 million
1917	$384.9 million
Mad Max: Fury Road	$375.4 million

Once Upon a Time in Hollywood	$374.3 million
Black Swan	$330 million
Green Book	$321.8 million
Lincoln	$275.3 million
Parasite	$258.8 million
Get Out	$255.4 million
True Grit	$252.3 million
American Hustle	$251.2 million
Silver Linings Playbook	$236.4 million
Hidden Figures	$236.2 million
The Imitation Game	$233.6 million
Argo	$232.3 million
Ford v Ferrari	$225.5 million
The Social Network	$224.9 million
Captain Phillips	$220.6 million
Little Women	$218.8 million
The Help	$216.6 million
Arrival	$203.4 million
The Shape of Water	$195.3 million
12 Years a Slave	$187.7 million
Hugo	$185.8 million
Hacksaw Ridge	$180.5 million
The Post	$179.8 million
War Horse	$177.6 million
The Descendants	$177.2 million
The Grand Budapest Hotel	$172.9 million
Bridge of Spies	$165.5 million
Three Billboards Outside Ebbing, Missouri	$160.2 million
Midnight in Paris	$154.1 million
Darkest Hour	$150.8 million
Lion	$140.3 million
The Big Short	$133.4 million
The Artist	$133.4 million
Zero Dark Thirty	$132.8 million

The Fighter	$129.2 million
The Theory of Everything	$123.7 million
Moneyball	$110.2 million
Birdman or (The Unexpected Virtue of Ignorance)	$103.2 million
Philomena	$100.1 million
Spotlight	$98.3 million
The Favourite	$96 million
BlacKkKlansman	$93.4 million
Jojo Rabbit	$90.3 million
Lady Bird	$79 million
Manchester by the Sea	$79 million
Vice	$76.1 million
Selma	$66.8 million
Moonlight	$65.3 million
Fences	$64.4 million
Brooklyn	$62.1 million
The Tree of Life	$61.7 million
127 Hours	$60.7 million
Boyhood	$57.3 million
Dallas Buyers Club	$55.7 million
Extremely Loud & Incredibly Close	$55.2 million
Whiplash	$49 million
Her	$48.3 million
Phantom Thread	$47.8 million
Call Me by Your Name	$41.9 million
Hell or High Water	$37.9 million
Room	$36.3 million
The Kids Are All Right	$34.7 million
Amour	$29.9 million
Nebraska	$27.7 million
Beasts of the Southern Wild	$23.3 million
Winter's Bone	$16.1 million
The Irishman	$8 million
Roma	$5.1 million

Marriage Story	$2.3 million

Table 11
Nominated films with the highest IMDb rating.

Title	IMDb Rating
Inception	8.8
Parasite	8.6
Whiplash	8.5
Django Unchained, Joker	8.4
1917	8.3
Toy Story 3, The Wolf of Wall Street, Green Book	8.2
Mad Max: Fury Road, Ford v Ferrari, 12 Years a Slave, Hacksaw Ridge, The Grand Budapest Hotel, Three Billboards Outside Ebbing, Missouri; Spotlight, Room	8.1
Bohemian Rhapsody, The Martian, The Revenant, La La Land, The King's Speech, Black Swan, The Imitation Game, The Help, Lion, Dallas Buyers Club, Her	8
Life of Pi, Arrival, The Artist, Jojo Rabbit, Boyhood, Call Me by Your Name, Amour, The Irishman, Marriage Story	7.9
Dunkirk, Hidden Figures, Captain Phillips, Little Women, The Big Short, The Fighter, Manchester by the Sea	7.8
Gravity, Get Out, Silver Linings Playbook, Argo, The Social Network, Midnight in Paris, The Theory of Everything, Birdman or (The Unexpected Virtue of Ignorance), Nebraska, Roma	7.7
Les Misérables, A Star Is Born, Once Upon a Time in Hollywood, True Grit, Bridge of Spies, Moneyball, Philomena, Hell or High Water	7.6
Hugo, The Favourite, BlacKkKlansman, Selma, Brooklyn, 127 Hours, Phantom Thread	7.5
Darkest Hour, Zero Dark Thirty, Lady Bird, Moonlight	7.4
Black Panther, American Sniper, Lincoln, The Shape of Water, The Descendants, Beasts of the Southern Wild	7.3
American Hustle, The Post, War Horse, Vice, Fences	7.2
Winter's Bone	7.1
The Kids Are All Right	7
Extremely Loud & Incredibly Close	6.9

The Tree of Life	6.8

Table 12
Nominated films with the highest Metascore rating.

Title	Metascore Rating
Boyhood	100
Moonlight	99
Parasite, 12 Years a Slave, Manchester by the Sea, Gravity, Roma	96
The Social Network, Zero Dark Thirty	95
La La Land, Amour, The Irishman, Marriage Story, Dunkirk, Lady Bird	94
Spotlight, Call Me by Your Name	93
Toy Story 3	92
Little Women, The Favourite	91
Mad Max: Fury Road, Her, Phantom Thread, American Hustle, Winter's Bone	90
The Artist	89
Whiplash, The Grand Budapest Hotel, Three Billboards Outside Ebbing, Missouri; The King's Speech, A Star Is Born, Hell or High Water, Brooklyn, Black Panther	88
Birdman or (The Unexpected Virtue of Ignorance), Nebraska, Moneyball, The Shape of Water	87
Room, Argo, Lincoln, Beasts of the Southern Wild, The Kids Are All Right	86
Get Out, The Tree of Life	85
The Descendants	84
Once Upon a Time in Hollywood, Hugo, BlacKkKlansman, The Post	83
Captain Phillips, 127 Hours	82
Django Unchained, Ford v Ferrari, Arrival, The Big Short, Silver Linings Playbook, Midnight in Paris, Bridge of Spies	81
The Martian, Dallas Buyers Club, True Grit, Selma	80
Black Swan, Life of Pi, The Fighter, Fences	79
1917	78
Philomena	77

The Revenant	76
The Wolf of Wall Street, Darkest Hour	75
Hidden Figures, Inception	74
The Imitation Game	73
The Theory of Everything, American Sniper, War Horse	72
Hacksaw Ridge	71
Green Book, Lion	69
Les Misérables	63
The Help	62
Vice	61
Joker	59
Jojo Rabbit	58
Bohemian Rhapsody	49
Extremely Loud & Incredibly Close	46

Table 13
Nominated films with the highest Rotten Tomatoes rating.

Title	Rotten Tomatoes Rating
Lady Bird, Selma	99%
Moonlight, Parasite, Toy Story 3, Get Out	98%
Boyhood, Spotlight, Mad Max: Fury Road, Hell or High Water, Brooklyn	97%
Manchester by the Sea, The Social Network, Black Panther, Argo, BlacKkKlansman	96%
12 Years a Slave, Gravity, The Irishman, Little Women, The Artist, True Grit, Roma	95%
Marriage Story, Call Me by Your Name, Her, Winter's Bone, Whiplash, The King's Speech, Moneyball, Arrival	94%
Amour, The Favourite, Room, Hugo, Captain Phillips, 127 Hours, Midnight in Paris, Dallas Buyers Club, Hidden Figures,	93%
Dunkirk, American Hustle, The Grand Budapest Hotel, The Shape of Water, The Kids Are All Right, Ford v Ferrari, Silver Linings Playbook, Fences	92%

Zero Dark Thirty, La La Land, Phantom Thread, Birdman or (The Unexpected Virtue of Ignorance), Nebraska, The Martian, Philomena	91%
Three Billboards Outside Ebbing, Missouri; A Star Is Born, Bridge of Spies, The Fighter	90%
Lincoln, 1917, The Imitation Game	89%
The Post, The Big Short	88%
The Descendants, Django Unchained, Inception	87%
Beasts of the Southern Wild, Life of Pi	86%
Once Upon a Time in Hollywood, Black Swan	85%
The Tree of Life, Darkest Hour, Hacksaw Ridge, Lion	84%
The Theory of Everything, Jojo Rabbit	80%
The Wolf of Wall Street	79%
The Revenant, Green Book	78%
The Help	76%
War Horse	74%
American Sniper	72%
Les Misérables	70%
Joker	68%
Vice	65%
Bohemian Rhapsody	61%
Extremely Loud & Incredibly Close	45%

Table 14
Genres with the most nominated films.

Genres	Count
Drama	81
Biography	29
Comedy	22
Thriller	17
Romance	16
History, Adventure	15
War	11
Crime	10
Action	9
Sci-Fi	8

Fantasy	7
Music	5
Western, Mystery	4
Sport	3
Musical, Family	2
Horror, Animation	1

2. Experimental design, materials and methods

The dataset includes all 88 movies that were nominated for the Academy Awards for the best picture category in the time range from 2010 until 2019, where it is acquired through web scraping the data from the list of all movies that were nominated for best picture, isolating to only scrape in the desired time range, then retrieve the data from the info box of the page of each movie. To achieve this we utilize the Python library BeautifulSoup and Requests, where BeautifulSoup plays an important part on the web scraping part, as it helps the process of retrieving the data from the HTML file and Requests helps to send a HTTP requests to get the additional data from the API.

The scraping process utilizes the BeautifulSoup library to extract the data from the HTML file of the webpage, in this case from Wikipedia, where it reads and retrieves the contents of the info box of the film page. For this process, it is very important to use the inspect element feature of the browser to target the desired class and body of information stored in the HTML file. The Requests library is also used here to send a HTTP request to the web page to get the information from each film's page.

After getting the desired data, we need to do data cleaning before moving on to the next step, as there might be some data that we need in a different format, in this case, the data of running time, budget, box office, and dates is converted into int, float, and datetime objects. In this step, the references to other pages are also removed, so that the final data will only show what is needed without the elements from the webpage.

Lastly, to get the ratings from IMDb, Rotten Tomatoes, and Metascore, an API is used to get those data., for this project we use the OMDb API. With the help of the Requests library, the data can be acquired and then classified as desired. The very last step is to save and store the data in the desired format.

Declaration of Competing Interest

This project did not receive financial support from any institutions. The authors declare that they have no known competing financial interests or personal relationships that could influence the work reported in this paper.

Ethics Statement

This project complies with the Wikipedia web scraping policy (see <u>robots.txt on Wikipedia</u>) and the OMDb API usage policy under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) copyright laws.

Acknowledgments

Supplementary Materials

Supplementary material associated with this paper can be found at the author's GitHub repository: <u>pandya-ae/SDE/Academy Award for Best Picture Dataset</u>.