Rotten Tomatoes and IMDB Ratings Analysis

(Brad Pitt vs. Leonardo DiCaprio)

Brigham Young University - STAT 340 – Dr. Tass

Tyler Hanks

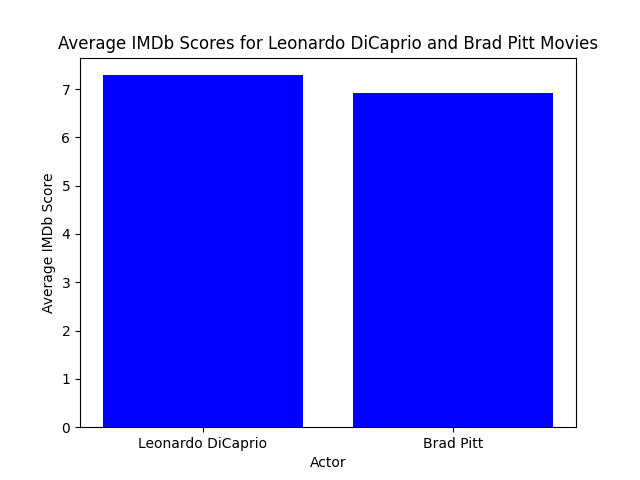
**Purpose and Background:**

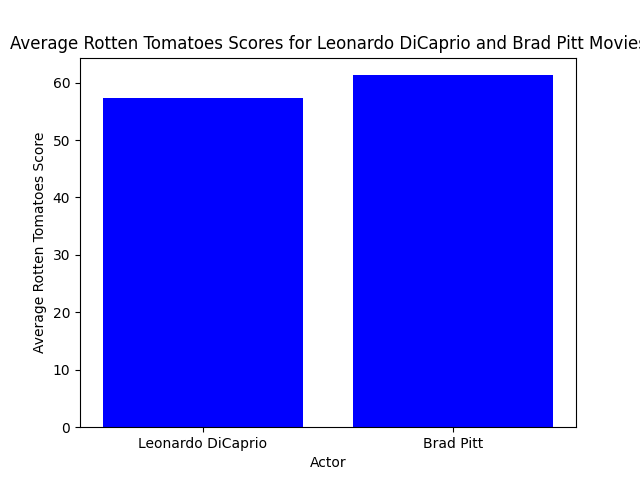
The purpose of this analysis is to answer the question; Which actor performing in a starring role, between Brad Pitt or Leonardo DiCaprio, leads to higher IMDb and Rotten Tomato ratings on average? Brad Pitt and Leonardo DiCaprio are two of the most famous and well respected actors and have been since the 1990’s. They have both been nominated for numerous Oscars and have each won at least one and are well known for their ability to take on diverse roles in films. With all these accolades, many have debated who is the better actor between the two. Although that debate is a matter of opinion, this analysis will dive into the hard facts of which actor has starred in higher rated films. To do so we will use IMDb and Rotten Tomatoes rating data. These two popular businesses are online movie rating databases where critics and fans can give ratings of any film.

**Data Collection:**

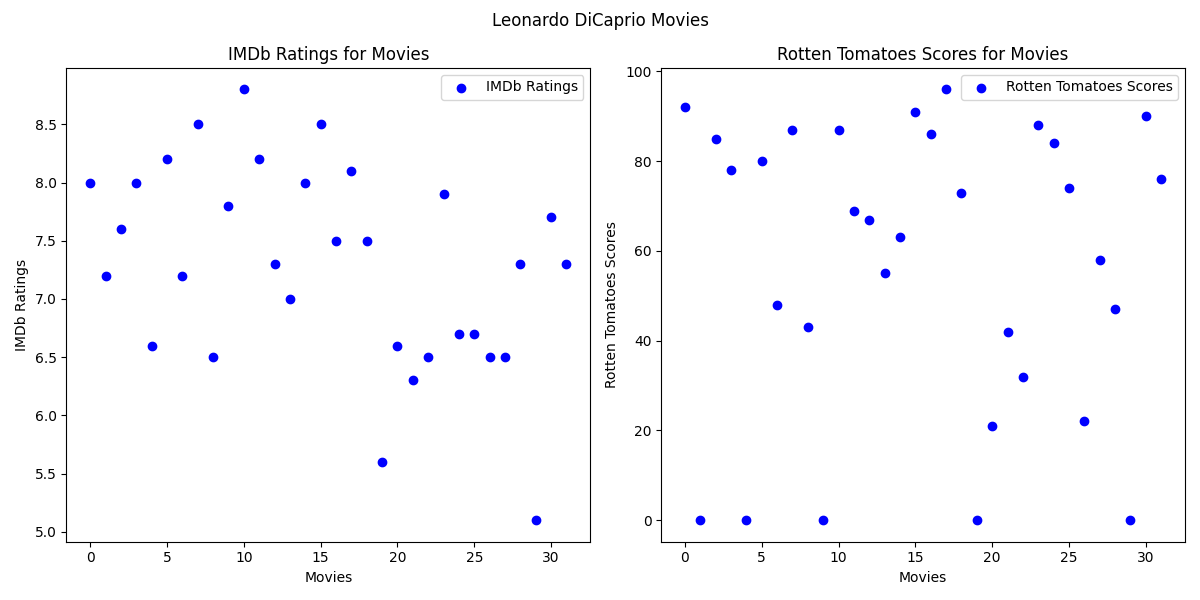
For this analysis I utilized IMDb and Rotten Tomatoes data available by scraping and querying the OMDb API which can be found at [www.omdbapi.com](http://www.omdbapi.com). To refine my focus, I filtered through the data and found the IMDb ratings and Rotten Tomatoes scores for every movie Leonardo DiCaprio and Brad Pitt starred in. It is important to note that Brad Pitt has starred in more movies (64) than Leonardo DiCaprio (33) which impacts our analysis as he has more ratings and scores to effect his average.

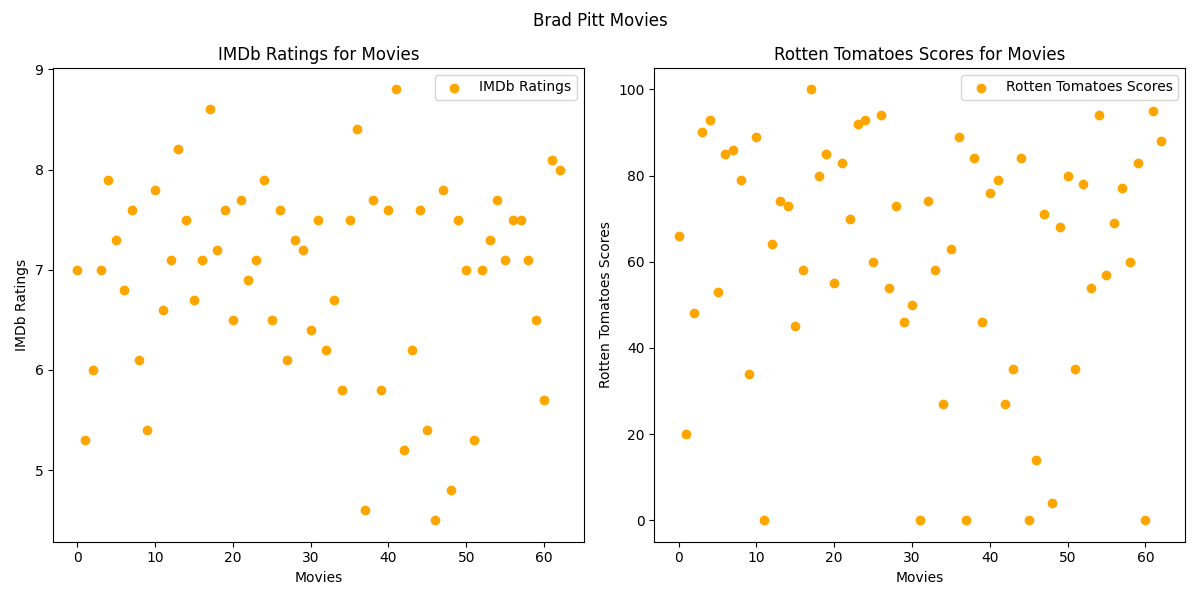
**Results:**

The bar graph below displays the average IMDb ratings for every movie the two actor’s have starred in. IMDb ratings are on a 1-10 scale. The analysis shows that, on average, that the film’s Leonardo DiCaprio starred in had the slightly better average (7.27) compared to those where Brad Pitt was featured with an average rating of (6.86). 

Our second bar graph below displays the average Rotten Tomatoes ratings for every movie the two actor’s have starred in. Rotten Tomatoes scores are on a 0-100 scale. The analysis shows that, on average, the films Leonardo DiCaprio played a role in had a slightly worse average (56.78) compared to those of Brad Pitt (60.97). 

Lastly, the scatter plots below show the IMDb ratings and Rotten Tomatoes scores for each film the actors have featured in. As displayed below, we can see that Brad Pitt has starred in more films (64) than Leonardo DiCaprio (33).





**Limitations:**

While our study provides valuable insights, it is essential to acknowledge its limitations. Bias from different user demographics on IMDb and Rotten Tomatoes, external influences, and the subjective nature of movie ratings may contribute to variations in our findings. Movie preferences are inherently subjective, and individual tastes can greatly influence ratings.

**Conclusion:**

In conclusion, our analysis sought to determine whether Brad Pitt or Leonardo DiCaprio, as lead actors, had higher average ratings on IMDb and Rotten Tomatoes. While the results indicate that Leonardo DiCaprio's movies had a slightly higher average IMDb rating, Brad Pitt's movies obtained a slightly higher average Rotten Tomatoes score.

Despite the aforementioned limitations, our analysis provides a starting point for understanding how these acclaimed actors' movies are received by audiences and critics alike. Moving forward, a more nuanced exploration of viewer demographics, directorial influences, and genre-specific considerations could enhance the depth and accuracy of our understanding.

In the realm of film appreciation, the debate between Brad Pitt and Leonardo DiCaprio's performances will likely persist, as personal preferences and opinions remain integral to the cinematic experience.

**Python Code:**

**import requests**

**from bs4 import BeautifulSoup**

**import csv**

**import matplotlib.pyplot as plt**

**def get\_movie\_data(api\_key, movie\_title):**

**base\_url = "http://www.omdbapi.com/"**

**params = {**

**"apikey": api\_key,**

**"t": movie\_title,**

**"type": "movie",**

**"plot": "short",**

**"tomatoes": "true",**

**"r": "json",**

**}**

**response = requests.get(base\_url, params=params)**

**data = response.json()**

**if data.get("Response") == "True":**

**return data**

**else:**

**return None**

**def extract\_movie\_info(movie):**

**return {**

**"Title": movie.get("Title", ""),**

**"Year": movie.get("Year", ""),**

**"IMDbRating": movie.get("imdbRating", ""),**

**"TomatoRating": next(**

**(rating["Value"] for rating in movie.get("Ratings", []) if rating["Source"] == "Rotten Tomatoes"),**

**"N/A"**

**),**

**}**

**def save\_to\_csv(data, filename):**

**with open(filename, mode="w", newline="", encoding="utf-8") as file:**

**fieldnames = ["Title", "Year", "IMDbRating", "TomatoRating"]**

**writer = csv.DictWriter(file, fieldnames=fieldnames)**

**writer.writeheader()**

**writer.writerows(data)**

**with open("api\_key\_omdb.txt", "r") as key\_file:**

**api\_key = key\_file.read().strip()**

**# Movies for Leonardo DiCaprio and Brad Pitt**

**movies\_leo = [**

**"Killers of the Flower Moon", "Don't Look Up", "Once Upon a Time in Hollywood",**

**"The Revenant", "The Audition", "The Wolf of Wall Street", "The Great Gatsby",**

**"Titanic: Deleted Scenes", "Django Unchained", "J. Edgar", "Inception: The Cobol Job",**

**"Inception", "Shutter Island", "Revolutionary Road", "Body of Lies", "Blood Diamond",**

**"The Departed", "The Aviator", "Catch Me If You Can", "Gangs of New York", "Don's Plum",**

**"The Beach", "Celebrity", "The Man in the Iron Mask", "Titanic", "Marvin's Room",**

**"Romeo + Juliet", "Total Eclipse", "The Quick and the Dead", "The Basketball Diaries",**

**"The Foot Shooting Party", "What's Eating Gilbert Grape", "This Boy's Life"**

**]**

**movies\_brad = [**

**"World War Z", "Wolves", "War Machine", "Voyage of Time", "Untitled Joseph Kosinski film",**

**"True Romance", "Troy", "The Tree of Life", "Thelma & Louise", "The Lost City",**

**"The Counselor", "The Big Short", "The Audition", "Spy Game", "Snatch",**

**"Sleepers", "Sinbad: Legend of the Seven Seas", "Seven Years in Tibet", "Seven",**

**"A River Runs Through It", "Once Upon a Time in Hollywood", "Ocean's Twelve",**

**"Ocean's Eleven", "Ocean's Thirteen", "No Way Out", "No Man's Land", "Mr. & Mrs. Smith",**

**"Moneyball", "The Mexican", "Megamind", "Meet Joe Black", "Less than Zero",**

**"Legends of the Fall", "Killing Them Softly", "Kalifornia", "Johnny Suede",**

**"Interview with the Vampire", "Inglourious Basterds", "Hunk", "Happy Together",**

**"Happy Feet Two", "Fury", "Fight Club", "The Favor", "The Devil's Own",**

**"Deadpool 2", "The Dark Side of the Sun", "Cutting Class", "The Curious Case of Benjamin Button",**

**"Cool World", "Contact", "Confessions of a Dangerous Mind", "By the Sea",**

**"Burn After Reading", "Bullet Train", "Being John Malkovich", "Babylon",**

**"Babel", "The Assassination of Jesse James by the Coward Robert Ford", "Allied",**

**"Ad Astra", "Across the Tracks", "12 Years a Slave", "12 Monkeys"**

**]**

**# Leo**

**leo\_movies\_info = []**

**for movie\_title in movies\_leo:**

**movie\_data = get\_movie\_data(api\_key, movie\_title)**

**if movie\_data:**

**leo\_movies\_info.append(extract\_movie\_info(movie\_data))**

**# Brad**

**brad\_movies\_info = []**

**for movie\_title in movies\_brad:**

**movie\_data = get\_movie\_data(api\_key, movie\_title)**

**if movie\_data:**

**brad\_movies\_info.append(extract\_movie\_info(movie\_data))**

**all\_movies\_info = leo\_movies\_info + brad\_movies\_info**

**save\_to\_csv(all\_movies\_info, "movies\_ratings\_comparison.csv")**

**print("Data successfully saved to movies\_ratings\_comparison.csv")**

**#finding avergaes**

**def calculate\_average\_scores(data):**

**imdb\_scores = [float(movie["IMDbRating"]) if movie["IMDbRating"] != "N/A" else 0 for movie in data]**

**tomato\_scores = [float(movie["TomatoRating"].rstrip('%')) if movie["TomatoRating"] != "N/A" else 0 for movie in data]**

**average\_imdb = sum(imdb\_scores) / len(imdb\_scores) if imdb\_scores else 0**

**average\_tomato = sum(tomato\_scores) / len(tomato\_scores) if tomato\_scores else 0**

**return average\_imdb, average\_tomato**

**#calculate averages**

**average\_imdb\_leo, average\_tomato\_leo = calculate\_average\_scores(leo\_movies\_info)**

**average\_imdb\_brad, average\_tomato\_brad = calculate\_average\_scores(brad\_movies\_info)**

**# Plotting bar graphs**

**def create\_bar\_graph(labels, scores, xlabel, ylabel, title):**

**x = range(len(labels))**

**plt.bar(x, scores, color='blue')**

**plt.xlabel(xlabel)**

**plt.ylabel(ylabel)**

**plt.title(title)**

**plt.xticks(x, labels)**

**plt.show()**

**# IMDb Scores**

**labels\_imdb = ['Leonardo DiCaprio', 'Brad Pitt']**

**average\_imdb\_scores = [average\_imdb\_leo, average\_imdb\_brad]**

**create\_bar\_graph(labels\_imdb, average\_imdb\_scores, 'Actor', 'Average IMDb Score', 'Average IMDb Scores for Leonardo DiCaprio and Brad Pitt Movies')**

**# Rotten Tomatoes Scores**

**labels\_tomato = ['Leonardo DiCaprio', 'Brad Pitt']**

**average\_tomato\_scores = [average\_tomato\_leo, average\_tomato\_brad]**

**create\_bar\_graph(labels\_tomato, average\_tomato\_scores, 'Actor', 'Average Rotten Tomatoes Score', 'Average Rotten Tomatoes Scores for Leonardo DiCaprio and Brad Pitt Movies')**