Will Smith Versus Denzel Washington
Which is the better actor?

Jessica Smith *Washington State University*

This project explored multiple variables from the IMDB database in order to determine which actor is better - Denzel Washington or Will Smith. A mathematical analysis was performed to quantitatively evaluate which actor is better using appropriate data features selected by the author. An adjacency matrix was constructed from data on 50 contemporary actors, and the eigen ratings of each actor were calculated. The results of the multivariate analysis showed that Denzel Washington can be considered the better actor.

Keywords: Will Smith, Denzel Washington, IMDB, Multivariate Analysis

December 08, 2020

1 Introduction

The IMDB database is a public repository containing information related to movies and actors, including cast, crew, ratings, and earnings data. This project attempted to investigate, compare, and draw conclusions regarding two well-known actors by performing multivariate analysis on the data gleaned from the IMDB website. Previous work into the subject evaluated the actors by considering box office sales and IMDB movie rankings. It was shown that the median box office sales, when adjusted for inflation, were about the same for the two stars. The median Metacritic ratings for each actors films were also similar and did not enable a meaningful distinction. As the results were inconclusive, further investigation is needed to provide deterministic insights into which actor is "better". This project will look at multiple variables and perform a mathematical analysis to attempt to quantitatively evaluate which actor is better using the available data features. This an important demonstration of how data analysis can be used to empirically inform a response to an otherwise hard to answer and somewhat subjective question. The author has no predisposition toward either actor, and the information presented is intended to provide an unbiased and data driven perspective.

2 Overview

Denzel Washington and Will Smith are two well-known movie stars with comparable ratings and extensive filmographies. Denzel is 66 years old and made his first movie in 1981, while Will Smith is 52 and made his first movies in 1992. Will Smith has 111 films listed on IMDB with a mean rating of 6.2, while Denzel has 61 films listed with a mean of 6.8. Both actors are listed in the top 500 on IMDB, and differentiating an advantage between the two on the basis of published statistics has proven challenging.

The IMDB data contains metrics on movies in which the actors have appeared including ratings, box office sales, and Metacritic scores, as well as info on the cast and crew. For actors, their rating within each movie is available, allowing for the determination of whether their role was a lead or whether they played a less prominent character. Demographic data and star meter rank was also available for each actor.

The concept of which actor is "better" is an inherently subjective measure, and the choice of which factors to consider in the analysis was given careful consideration. Calculating a diversity score from an analysis of the gender breakdown of the cast and crew for each movie was suggested. However, different genres can have different casting requirements, and the outcome of any analysis may be more indicative of which genre an actor prefers than anything else. A preliminary assessment of the diversity index of Will Smith movies, for example, showed a wide variation in score across genres. A strong correlation between gender diversity and acting acumen appears unlikely, and gender/diversity score was not selected as a component of the analysis.

Another available factor was the "star meter" ranking, a measure of popularity created by IMDB that is a function of the number of credits a person has, popularity of the work a star appears in, and traffic to the celebrity's profile. While this initially seemed promising, after closer inspection, many of the actors used to build the data set had identical ratings. This rating also appears as a static snapshot that reflects the score today, rather than at the time a particular movie was released, and it's unclear how these ratings might change over time. This metric was deemed unreliable and non-deterministic, and was not included in the final analysis.

The analysis required building a dataframe for each actor with the desired covariates. For each film an actor was associated with, the movie ratings, Metacritic score, box office sales, and actor rank were selected. The decision was made to only consider movies where the star had been a headliner, defined as having an actor rank of 1, 2, or 3 for a given movie. This was done in an attempt to make sure the covariates were a mainly a function of the actor being considered, not some one else who may have been in the movie and had a larger role. The assumption was made that a "better" actor would be appear as the lead more often than in a supporting role, so the ratio of leading roles to total movies was also considered. Box office sales were adjusted to 2020 dollars, and the mean of the ratings, scores, and sales for each movie were calculated.

To obtain a more robust analysis, Will and Denzel should be compared to more actors than just each other. A pool of 48 contemporaries was drawn from the actorRank2000.rds provided by the instructor, ensuring only candidates from the modern era were considered. Actor ID's were randomly selected and screened for quality. The criteria were at least 10 movies where the actor ranked 3 or lower, and complete data for ratings, Metacritic scores, and box office sales. Table 1 shows an example of the first ten actors in the pool and the final selection of covariates that were used in the analysis.

	actor	lead.ratio	avg.ratings	avg.metacritic	avg.box.office
1	Will Smith	0.26	6.57	51.12	186.15
2	Denzel Washington	0.61	6.97	61.42	85.52
3	Adam Beach	0.24	6.28	59.20	29.66
4	Juliane Kohler	0.41	6.59	68.67	3.49
5	Emmanuelle Chriqui	0.32	5.08	45.00	38.88
6	Helen Hunt	0.41	6.37	57.07	91.78
7	Demi Moore	0.43	6.09	48.75	79.66
8	Keira Knightley	0.34	6.74	58.50	66.76
9	Kristen Wiig	0.33	6.26	61.47	92.57
10	Sam Worthington	0.33	6.18	42.91	161.27

Table 1: A sample of the actors and covariates used in the analysis.

Using the entire pool of 50 actors, the data was standardized by dividing each value for a covariate by the maximum value for that feature found in the pool. The result was a 50 X 4 matrix where each value was a positive integer between 0 and 1. An adjacency matrix was calculated where the value of the eigenvector in each row/column is the rating of the actor evaluated against the other actors in the pool for that factor. The eigenvector ranking quantified the approximate importance of each actor in the pool. From these eigen-rankings, an empirical evaluation of which actor is "better", Will or Denzel, was able to be obtained.

3 Key Findings

The results of the analysis show that Denzel Washington has a higher eigen-rank than Will Smith, indicating that given the metrics evaluated, Denzel can be considered the better actor (.20 to .16). These results are in agreement with the rankings returned from the instructors Actor-Actor matrix that showed Denzel Washington scoring slightly higher than Will Smith (57.11 to 50.79). Given the authors lack of preference

	eigen.rank
Will Smith	0.17
Denzel Washington	0.20

Table 2: The eigenvector rankings show Denzel Washington is the better actor.

at the outset, the relative 'closeness' of the results, and the alignment with existing research outcomes, the results can be concluded to be reasonable and acceptable.

4 Conclusion

Denzel has a higher eigen-rank than Will Smith when considering the average ratings of the movies in which he starred, the average box office sales of those movies, and the ratio of leading to non leading roles. The results of the multivariate analysis have shown that Denzel Washington is the better actor.

TABLE OF CONTENTS