Mathematics and Sports

Money Bull: Analyzing the Application of Ranking Methods to Rodeo

M. Adler[1,*], F. McDill[2], T. Ng[3], W. Paz[4], I. Horng[5], S. Thomas[6], A. Harsy[7], A. Schultze[7]

[1]Pomona College, Claremont, CA, [2]Wesleyan University, Middletown, CT, [3]Kenyon College, Gambier, OH, [4]Miami University, Oxford, OH, [5]University of Pennsylvania, Philadelphia, PA, [6]Brown University, Providence, RI, [7]Lewis University, Romeoville, IL, [7]Lewis University, Romeoville, IL

[*]Corresponding Author E-mail:
mrad2022@mymail.pomona.edu

Abstract

Drawing millions of fans each year and surpassing even golf and tennis in sporting event attendance in the United States, the rodeo stands as one of North America's most unique and iconic sports. Despite having vast numbers of participants and spectators, little mathematical work has been published on the rodeo's ranking system, strategy, or other common topics covered in sports analytics. In this research, we examine the current ranking system for bareback riding in the largest rodeo organization in America, the Professional Rodeo Cowboys Association (PRCA). Due to a wide range of rodeo prize pools with no observable pattern for how they are set, the PRCA's use of total earnings as the primary measure of ranking may not accurately represent rider skill. We explore alternative methods of comparing bareback riders by extending classical linear algebraic ranking methods (specifically, the Colley, Massey, Keener, and PageRank methods) to rank bareback riders based on their PRCA performance data. We assess the effectiveness and predictive power of these standard methods. Ultimately, we find that each of these linear algebra models favors a different aspect of rider performance—such as average earnings, average total score, and average rider score—together comprising a more holistic ranking system.

Keywords: ranking, rodeo, mathematical modeling, linear algebra

1 Introduction

The rise in popularity of sports analytics is often attributed to Michael Lewis' best-selling book *Moneyball: The Art of Winning an Unfair Game* [40], inspired by the analytical approach to performance and management-based decision-making found in Bill James' *Historical Baseball Abstract* [23, 30]. Following the success of its use in baseball, increased funding and resources have been dedicated to analytics programs in other popular sports, such as American football [8], soccer [64], and basketball [27]. Despite the field's surge in popularity, not-as-mainstream sports, such as NASCAR [3], e-sports [56], and lacrosse [45], have been neglected by the analytical community [67]. In this paper, we focus on applying these analytical tools to a sport that mathematics has left in the dust: rodeo.

Even though it offers a unique setting for analysis, to the best of our knowledge, the rodeo suffers a general lack of attention from the analytical community. The Professional Rodeo Cowboys Association (PRCA), the world's largest rodeo organization, uses only total prize earnings as its metric to calculate the world ranking riders [46]. These top-ranking riders have the best chances of receiving bids to larger rodeos with exponentially larger prize pools. Having the most lucrative rodeos only available to riders earning in the top percentile increases the disparity between the winnings of riders who earn enough to rank and those who do not. This isn't simply a matter of recognition as, unlike most professional athletes, professional riders do not have a standard salary [46]. This ranking system implicitly equates total earnings and skill, and investigating alternative ranking methods could open up new opportunities for riders whose abilities are not accurately represented.

We begin our investigation by providing a general background on the sport of rodeo and its influence as a pillar of American culture. We consider ranking systems that more explicitly account for athlete skill and performance, modifying commonly-used algorithms originally developed or adapted to rank match-based team sports [15, 24, 31, 41]. We apply the methods to rank 71 PRCA bareback riders based on their performance data in the 2024 season and finish with a discussion of our results, comparing the different rankings with that of the PRCA's official world ranking and providing insight into the individual strengths of each ranking system. In this paper,

Figure 1: PRCA circuit map

we aim to compare our findings to the official world rankings of bareback riders in the PRCA and use our findings to consider different ranking metrics for the sport. The code utilized for data collection and in the implementation of ranking methods and additional processes is available on GitHub [2].

2 Background

The rodeo has been a staple in North America since the 1800s when ranch hands, known as *vaqueros*, would compete in horse riding and roping competitions similar to rodeo events today. The growing influence of *vaquero* culture in the American West increased the popularity of such sports, leading to the emergence of the sensationalized form of rodeo we see today, such as in *Buffalo Bill's Wild West* [68]. The rodeo remains well-loved to this day, with rodeos nationwide attracting millions of attendees annually and distributing millions of earnings to their contestants [46]. For example, during the 2024 season, the PRCA rodeos had allocated over $30 million to their winners by mid-July [46].

2.1 Rodeo Structure

Thousands of rodeos are held each year across North America by numerous different rodeo organizations. For the purposes of this paper, we'll focus on the rodeos of the 13 circuits sanctioned by the PRCA. Circuits are the regional groupings of competitions that each cowboy competes in based on their geographic location (see Figure 1). Performing well in circuits can earn riders a bid to the more lucrative national rodeos, tours, and series and gives riders a chance to qualify for the largest event of the year, the National Finals Rodeo (NFR), where the financial payouts are the highest across all PRCA events [12, 46].

The rodeo generally consists of a core group of ten events that can be split into timed events (such as barrel racing, steer wrestling, and steer roping) and judged *roughstock* events (such as saddle bronc riding, bull riding, and bareback riding) [52]. Though events within the roughstock category differ in rules and types of stock (bulls and horses), riders in each event attempt to stay on the stock for eight seconds and are scored by a pair of judges based on their evaluation of rider technique and the performance of the stock [5].

Rodeos source their animals from stock contractors, who breed and train them to perform the specified task for each event. In roughstock events, the stock selected are bred for their extreme jumping and kicking characteristics [11]. Each stock is then assigned to a rider before the event through a lottery held at the PRCA headquarters in a process known as "the draw" to ensure fairness in assigning stock [50].

2.2 Bareback Riding

The focus of our work is on the sport of bareback riding, also referred to as bareback bronc riding. During this roughstock event, participants ride horses (called *broncs*) using only a rigging resembling a suitcase handle rather than a saddle [11].

The challenge of the game lies in ensuring a rider's raised hand does not touch the stock and that they do not get thrown off (or *bucked off*) before the eight seconds are up. Failure to do so results in no score. Upon leaving the chute (a small, fenced-in area attached to the arena where riders mount the stock) and until the horse's front feet hit the ground from its first buck, the cowboy must also *mark out* the bronc—that is, keep their spurs at or above the horse's shoulders—or else receive no score [5].

If the rider is able to last the eight seconds without being bucked off, two judges give them two scores each, judging their technique and the stock's performance according to the rubric outlined in the PRCA Judge's Handbook [4]. Each judge scores the rider from 0 to 25 (with a step size of 0.5) on their ability to maintain correct spurring techniques as well as how well they kept rhythm and timing according to the bronc's movements. They also each score the stock out of 25 based on the difficulty of its bucking style and bucking strength, with high kicking and consistent movements earning the stock more points. The two judges then sum up their scores for the rider (the *rider score*) and their scores for the stock (the *stock score*) to give the rider a final score out of 100 points for that ride.

2.3 Ranking System of the PRCA

As it currently stands, the PRCA's world ranking of bareback riders is based on total earnings in bareback events in a given rodeo season [49]. This method has limitations. For one, the payout of each event varies greatly between different rodeos (even those within the same circuits), and to the best of our knowledge, there does not seem to exist a standard correlation between prize pool size and the competitive level of a given rodeo. For example, if a rider does not participate in a large rodeo like RodeoHouston, they will fall tens of thousands behind those who have, as the prizes offered in RodeoHouston can be twice that of rodeos considered to be of the same notability. In this paper, we explore alternative methods of comparing bareback riders to build more holistic ranking systems and examine their benefits.

3 Classic Ranking Methods

Individual sports and judged sports have had their fair share of analysis in the math community, with a large field of research dedicated to examining various ways of ranking athletes [6, 22, 25, 71]. Rankings can offer athletes, coaches, team managers, sportsbooks, and other professionals in sports useful data for important decision-making in both individual sports (see [9, 32, 69]) and judged sports (see [53, 55, 70]). Despite being an incredibly popular individual judged sport, we found little to no existing work studying the rodeo or comparing riders' performances.

In this research, we examine four commonly used sports ranking algorithms: the Colley method [15], the Massey method [41], Keener's method [31], and PageRank [54]. These methods are typically used for ranking team sports and predicting end-of-season win percentages [9, 15, 26, 28, 31, 38, 41, 54, 61], but in the following sections, we adapt each algorithm to the sport of bareback riding and use them to compare 71 riders in the PRCA. In order to utilize these methods of ranking for the individual sport of bareback riding, we follow the approach used by Beggs, Shepherd, Emmonds, and Jones [9]. We define a "game" as the comparison between two riders in the same round of the same rodeo. So, for example, a rodeo with four competitors and one round would have each participating rider competing in three games. Within a game, the rider with the higher score is given a "win," no matter their place in the whole event. So, in the same hypothetical rodeo with four competitors, the rider in first place is given three wins, the rider in second is given two wins, and so on. Losses are determined similarly.

3.1 The Colley Method

Published by Wesley N. Colley in 2002 to study college football rankings, the Colley method builds upon the winning percentage model and utilizes a simple "who beat who" matrix and Laplace's rule of succession, adding one win and one loss to each team's record before calculating winning percentage [15, 20]. This adjustment gives each untried team or player a rating of $\frac{1}{2}$. Thus, the proposed rating for each team, r_i using the Colley method is given by [15]:

$$r_i = \frac{w_i + 1}{t_i + 2}$$

where w_i is the total number of wins of team i and t_i is the total number of games played by team i. Colley then rearranged this equation to incorporate strength of schedule (see [13] and [28] for details):

$$(2 + t_i)r_i = 1 + \frac{w_i - l_i}{2} + S_i \tag{1}$$

where l_i is the total number of losses of team i and S_i, representing the sum of the ratings of the teams that played against i, incorporates strength of schedule [15]. Following this method for n teams creates a system of equations which can then be represented as an [n x 1]

vector, **b** where:

$$b_i = 1 + \frac{w_i - l_i}{2}$$

and a Colley coefficient matrix **C** defined as:

$$C_{ij} = \begin{cases} 2 + p_i & i = j \\ -p_{ij} & i \neq j \end{cases}$$

where p_i is the total number of games played by team i and p_{ij} is the number of games team i and team j played against each other [26]. We then set up and solve a system of equations, represented as the matrix system $\mathbf{Cr} = \mathbf{b}$, for a ratings vector $\mathbf{r}$. Next, we order the team rankings by the value of the ratings with the largest rating corresponding with the top-ranked team.

This system of equations considers the winning percentages of a team's opponents, allowing teams with the same record to have different ratings, given they played different opponents [15]. Though primarily used for match-based sports, the Colley method has been adapted for use in individual sports such as track [9], whose competition structures are most similar to roughstock events.

As mentioned earlier, in order to utilize Colley's method of ranking for bareback riding, we use the approach in [9] and define a game as the comparison between two riders in the same round of the same rodeo. The total number of games that a given rider i participates in is entered as t_i. Recall, within a game, the rider with the higher score is given a win, no matter their place in the event. The total wins and losses of a given rider i is entered as w_i and l_i, respectively.

To further illustrate how we applied the Colley method to bareback riding, we consider the simple example provided in Table 1 and Figure 2.

Rodeo	Rider	Total Score
Rodeo 1	A	87
	B	79
	C	76
	D	80
Rodeo 2	B	85
	C	82
	D	75
Rodeo 3	A	84
	B	83
	D	86

Table 1: A simple example with four riders and three rodeos

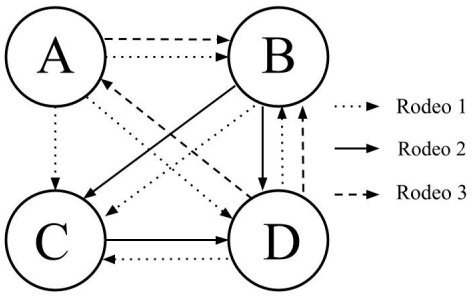

Figure 2: Visualization of Table 1. *An arrow points from the rider with the higher score to the rider with the lower score in a given rodeo round.*

Using (1), our initial system of equations would be as follows:

$$(2+5)r_A = 1+\frac{4-1}{2}+2r_B+r_C+2r_D$$
$$(2+7)r_B = 1+\frac{3-4}{2}+2r_A+2r_C+3r_D$$
$$(2+5)r_C = 1+\frac{1-4}{2}+r_A+2r_B+2r_D$$
$$(2+7)r_D = 1+\frac{4-3}{2}+2r_A+3r_B+2r_C$$

Rearranging this system of equations allows us to solve for the ratings vector $\mathbf{r}$:

$$\begin{bmatrix} r_A \\ r_B \\ r_C \\ r_D \end{bmatrix} = \begin{bmatrix} 0.688 \\ 0.458 \\ 0.313 \\ 0.542 \end{bmatrix} \tag{2}$$

We see that Rider A has the largest corresponding rating, followed by D, then B, then C. Therefore, the Colley method places Rider A in first, Rider D in second, Rider B in third, and Rider C in fourth.

3.2 The Massey Method

The Massey method introduced by Kenneth Massey in 1997 provides an alternative method for ranking teams, using point differentials as its metric [41]. When considering team sports, this method works by aggregating data from all games played and combining them into a large system of equations where each row represents a particular game from the data set. A row has the form of:

$$r_i - r_j = k \tag{3}$$

where, for a given game, Team i beat Team j by k points and r_i and r_j represent the rating of Team i and Team j, respectively. Thus, using Massey's method, we would assume that the difference in the rating of Team i and Team j differ by k [28, 61]. We represent this system of equations as $\mathbf{Xr} = \mathbf{b}$, where $\mathbf{X}$ is a matrix with m rows (one for each game) and n columns (one for each team) and $\mathbf{b}$ is the point differential vector [41].

However, since each row represents a unique game, in practice, we usually have more rows than columns. Furthermore, if two teams play each other more than once, it is rare that they will have the same point differential, creating an inconsistent system. To account for this, Massey uses the least squares method to estimate a solution that produces the matrix equation $\mathbf{Mr} = \mathbf{p}$. However, the solution to this system is not necessarily unique. To address this, Massey replaced one of the rows with an equation that is not in the span of the other rows of the system. Although Massey could replace any row, his original work replaced the last row of the system with $\sum_{i=1}^{m} r_i = 0$ which forces the system to have a unique solution (see [41], [13], [28], and [38] for additional details).

As with the Colley method, when applying the Massey method to bareback riding, we treat each pair of riders in one round of a rodeo as one game. Referring to the example in Table 1 and Figure 2, Rider A would play three games in Rodeo 1, as would Rider B, C, and D. From (3), we can construct our matrix system $\mathbf{Xr} = \mathbf{b}$ to be the following:

$$\begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_A \\ r_B \\ r_C \\ r_D \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 11 \\ 1 \\ 8 \\ 7 \\ 7 \\ 10 \\ 3 \\ 1 \\ 3 \\ 2 \end{bmatrix}.$$

Solving this system using least squares and using Massey's steps to ensure a unique solution, we get the following ratings vector:

$$\begin{bmatrix} r_A \\ r_B \\ r_C \\ r_D \end{bmatrix} = \begin{bmatrix} 3.771 \\ 0.163 \\ -1.538 \\ -2.396 \end{bmatrix}.$$

Thus, the Massey method gives us Rider A in first place, Rider B in second, Rider C in third, and Rider D in fourth.

3.3 Keener's Method

Keener's method was published in 1993 by James Keener to rank college football teams with the aim to alleviate the challenges that come when teams or players in a league do not play in a round-robin style, resulting in "uneven paired competition" [31]. Keener's method offers its users flexibility in the statistic they choose for performance analysis [38]. In the case of point differentials, Keener uses Laplace's rule of succession so that the a_{ij} entry of the Keener matrix is [20, 31, 38]:

$$a_{ij} = \frac{S_{ij}+1}{S_{ij}+S_{ji}+1} \tag{4}$$

where a_{ij} is the statistic produced when team i competes against team j, S_{ij} is the cumulative points team i scored against team j over the investigated period and S_{ji} is the cumulative points team j scored against team i. This statistic can also be thought of as an approximation of the probability that team i will beat team j in their next match [26]. Keener's method then uses the Perron-Frobenius theorem [42] to find the ratings of the team using the *keystone equation* [38],

$$\mathbf{Ar} = \mathbf{r}\lambda \tag{5}$$

where $\mathbf{A}$ is the Keener matrix, λ is the proportionality constant, and $\mathbf{r}$ is the ratings vector. It is important to note that the Keener matrix must be non-negative to apply the Perron-Frobenius theorem [42]. Skewing and weighting functions can be applied to the entries of the Keener matrix to address this constraint or minimize the effects of bias or uneven distribution. Additionally, the matrix must be irreducible to ensure there is enough competition between teams so that any pair of teams can be ranked against one another [38]. This can be addressed by adding the Keener matrix to another matrix of the same size populated with some small number, relative to the smallest non-negative entry of the Keener matrix [38]. The teams can then be ranked based on their ratings. While there have been papers using Keener's method for individual sports [9], there is limited existing work on applying Keener's method to individual judged sports that have point differentials, such as roughstock events, gymnastics, and diving. [1]

Using the same example from Table 1 and Figure 2, our Keener's matrix using point differentials is as follows:

$$\mathbf{A} = \begin{bmatrix} 0.01 & 0.919 & 0.927 & 0.737 \\ 0.101 & 0.01 & 0.885 & 0.698 \\ 0.093 & 0.135 & 0.01 & 0.625 \\ 0.283 & 0.323 & 0.395 & 0.01 \end{bmatrix},$$

which is used to find the ratings of the riders in its Perron-Frobenius vector $\mathbf{r}$ from (5):

$$\begin{bmatrix} r_A \\ r_B \\ r_C \\ r_D \end{bmatrix} = \begin{bmatrix} 0.403 \\ 0.246 \\ 0.153 \\ 0.199 \end{bmatrix}.$$

Notice that Keener's method places Rider A in first, Rider B in second, Rider D in third, and Rider C in fourth.

3.4 PageRank

Developed by Larry Page and Sergey Brin in 1996, PageRank is now incorporated into Google's algorithm to sort websites by their popularity [13, 54]. This method works by creating a directed graph where an edge points from one website to another—or one website "votes" for another—if the first site links to the second. The rank of a website is determined both by how many other sites "vote" for it and by the rank of the sites that "vote" for it [9]. This voting process can be modeled using a directed graph like the one in Figure 2 and enables us to create an adjacency matrix, $\mathbf{S}$ of the directed graph. The goal of PageRank is to use the values of the steady-state vector of the stochastic adjacency matrix as the ratings for each website. Thus, to make the matrix stochastic, the entries in row i of the matrix $\mathbf{S}$ are divided by $\frac{1}{s_i}$, where s_i is the sum of row i, such that all row sums are 1. Then, to ensure that the matrix is interconnected and

[1] To find an unweighted Keener ranking from our full PRCA dataset (see Section 4), we constructed a Keener matrix for the 71 riders using the definition of a_{ij} above. We perturbed our matrix by adding a 71 by 71 matrix populated with 0.01 in each entry to ensure irreducibility. The Perron-Frobenius vector was obtained and riders were ranked.

irreducible so that we can find the steady state vector, the matrix is dampened to create the Google matrix $\mathbf{G}$:

$$\mathbf{G} = \alpha\mathbf{S} + (1-\alpha)\mathbf{E} \tag{6}$$

where, following the methodology of Beggs et al., the dampening factor α is typically 0.85, and $\mathbf{E}$ is an [n x n] matrix of entirely $\frac{1}{n}$ where n is the number of websites being ranked [9]. Finally, we find the steady-state vector of the stochastic adjacency matrix to estimate the ratings of each website [37].

Originally adapted for the National Football League by Anjela Govan and Carl Meyer [24], PageRank is now a popular ranking method in sports analytics [9, 33, 36]. Again following the steps taken by Beggs et al., we adapted the PageRank method such that riders give a "vote" to all the riders who scored higher than them in a given round of a given rodeo [9].

Looking at Figure 2, we can construct our initial adjacency matrix as:

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 2 & 0 & 0 & 2 \\ 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

where a_{ij} is the amount of "votes" rider i gave to rider j, or the number of times rider i scored fewer points than rider j while both riders were in the same rodeo. We then adjust the matrix so that each row sum equals one:

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \end{bmatrix}$$

and apply (6) using $\alpha = 0.85$ to make our matrix irreducible and interconnected:

$$\begin{bmatrix} 0.0375 & 0.0375 & 0.0375 & 0.8875 \\ 0.4625 & 0.0375 & 0.0375 & 0.4625 \\ 0.25 & 0.4625 & 0.0375 & 0.25 \\ 0.3208 & 0.3208 & 0.3208 & 0.0375 \end{bmatrix}.$$

After solving for our steady-state vector, we get our PageRank ratings to be:

$$\begin{bmatrix} r_A \\ r_B \\ r_C \\ r_D \end{bmatrix} = \begin{bmatrix} 0.265 \\ 0.208 \\ 0.146 \\ 0.382 \end{bmatrix}.$$

Hence, PageRank gives us that Rider D is ranked first, Rider A is second, Rider B is third, and Rider C is fourth.

3.5 Point Caps

In ranking methods that incorporate score differentials, we can lessen the impact of blowouts by introducing point differential caps [13, 28]. Blowouts are a common occurrence in bareback riding, typically occurring when a rider gets bucked off and their opponents do not. In order to reduce their impact on the rankings, we added a point cap to our models. To determine the point differential cap, we examined the distribution of total scores across our dataset (see Figure 4). The interquartile range of the total score distribution is 7.5 points, while, as there are no upper outliers, the difference between the maximum score and the median score was 12.5 points. We decided that a point differential cap of 10 would accurately represent non-blowout score differences and applied this cap to our adapted Massey and Keener models.

3.6 Review of Examples

It is interesting to observe how each classic method ranks the riders from our example. Rider D is ranked differently by each method, yet the remaining riders are always ordered so that A is ranked above B and B is ranked above C (see Table 2). Each ranking method seems to have some underlying similarity when applied to bareback riding, but produces different results overall. We begin our investigation on how these methodologies behave when applied to real bareback riding data in order to examine the source and extent of their differences.

Rank	Colley	Adj. Colley	Massey	Keener	PageRank
1	A	A	A	A	D
2	D	D	B	B	A
3	B	B	C	D	B
4	C	C	D	C	C

Table 2: Riders from Table 1 ranked by classic unweighted ranking methods

4 Data Collection

We selected riders for data collection by consolidating all bareback riding contestants in the 2023 and 2024 world rankings and all bareback riding contestants in the 2023 and 2024 PRCA playoff series rankings as posted on the PRCA site on July 12th, 2024 [47, 48, 57, 58]. We then web-scraped 2024 season performance data from the PRCA profiles of all but four riders: Kaycee Feild, Yance Day, and Pascal Isabelle, who did not have active PRCA profiles, and Trenten Montero who passed away due to injuries incurred during a ride in August 2023. In addition to these riders, we removed any riders who competed in less than five rodeos. This removed Tim O'Connell from our dataset, as he could not compete after being injured in his first rodeo of the PRCA 2024 season. Once these athletes were omitted from the final data collection, we were left with data from 71 unique riders across 244 unique bareback riding events. The collected data includes the riders that participated, the stock they rode, the total score for each of their rides, where the rider placed in the round, the rider's round earnings, and the total available prize pool for the round.

5 Independence Testing, Bias

In sports with a subjective scoring system, it is important to have unbiased judges to ensure the sport's fairness [29] Research has shown that there are known biases based on gender, competition order, difficulty of routine, the nationality of the judge, and country of origin of the judge [10, 19, 21, 29, 44, 62]. In [44], Morgan & Rothoff, found strong evidence of what the authors coined as difficulty bias in gymnastics whereas athletes who attempt more difficult tasks are rewarded with higher execution scores. We investigated whether there was a similar *difficulty bias* in bareback riding. That is, we wondered if there was a correlation between the stock score (determined by bucking difficulty) [50] and rider performance score (see Section 2.2). To obtain some understanding of potential bias in the difficulty level of the stock score, we used regression methods to examine the extent to which stock scores are predictors of rider performance scores. It is important to note that being *bucked off* yields a default total score of zero, so the judges cannot show bias. Thus, in the regression and correlation, we removed all scores equal to zero.

We used Pearson's correlation test for the difficulty bias inquiry in bareback riding. Since both scores are between 0 and 50, are normally distributed, and contextually irrelevant outlier values were removed, we found that Pearson's correlation was the appropriate test. As seen by the trend line shown in Figure 3, there is a minimal positive, linear trend between rider and stock score. Quantitatively, the resulting Pearson correlation coefficient was $r = 0.056$, which verifies there is little to no correlation between the two scores. As a result, we found no strong evidence of difficulty bias or correlation between rider and stock score. Although the scores are not truly statistically independent, there is enough evidence to allow us to assume that a rider's total score is two independent scores. Using this assumption, we can add different weights to the two in order to conduct further analysis of the impact of each score on predictability.

In general, analyzing other factors such as environmental elements, the rider's condition, or the presence of judging bias in the entire realm of the rodeo becomes a little trickier, as each rodeo is scored relative to itself. This is further explored in Section 7.5.

6 Normalization and Weighting

Beyond the standard models, we can apply weights to our classic ranking methods to account for how different elements of a sport impact an athlete's skill (as perceived by the ranking method). In our application of classic ranking methods to bareback riding, we investigated two types of weighting.

Firstly, we decided to add weight with respect to the total prize pot for a given rodeo's bareback event, or *potential earnings*. Rodeos that have larger prize pots tend to have stricter qualifications required to participate, hence making these higher-paying events more competitive and allowing potential earnings to serve as a possibly successful metric for measuring rider skill [12, 51].

Secondly, we chose to weigh rider score more than stock score. Although there is no true statistical independence, it is reasonable to treat rider score and stock score as individual variables within a specific event due to the uncorrelated nature discussed earlier (see Section 5). In addition, there is a low recurrence of stock within a season, as a stock cannot be ridden more than once per day [59]. Furthermore, "the draw" ensures randomness in the assignment of stock to a rider, as mentioned in Section 2.1 [50]. With all of the above considered, we conclude that stock score does not indicate rider skill and weight rider score to diminish the effect of stock score on a rider's ratings. We will now delve into how we assigned weights and how they were implemented in each ranking method.

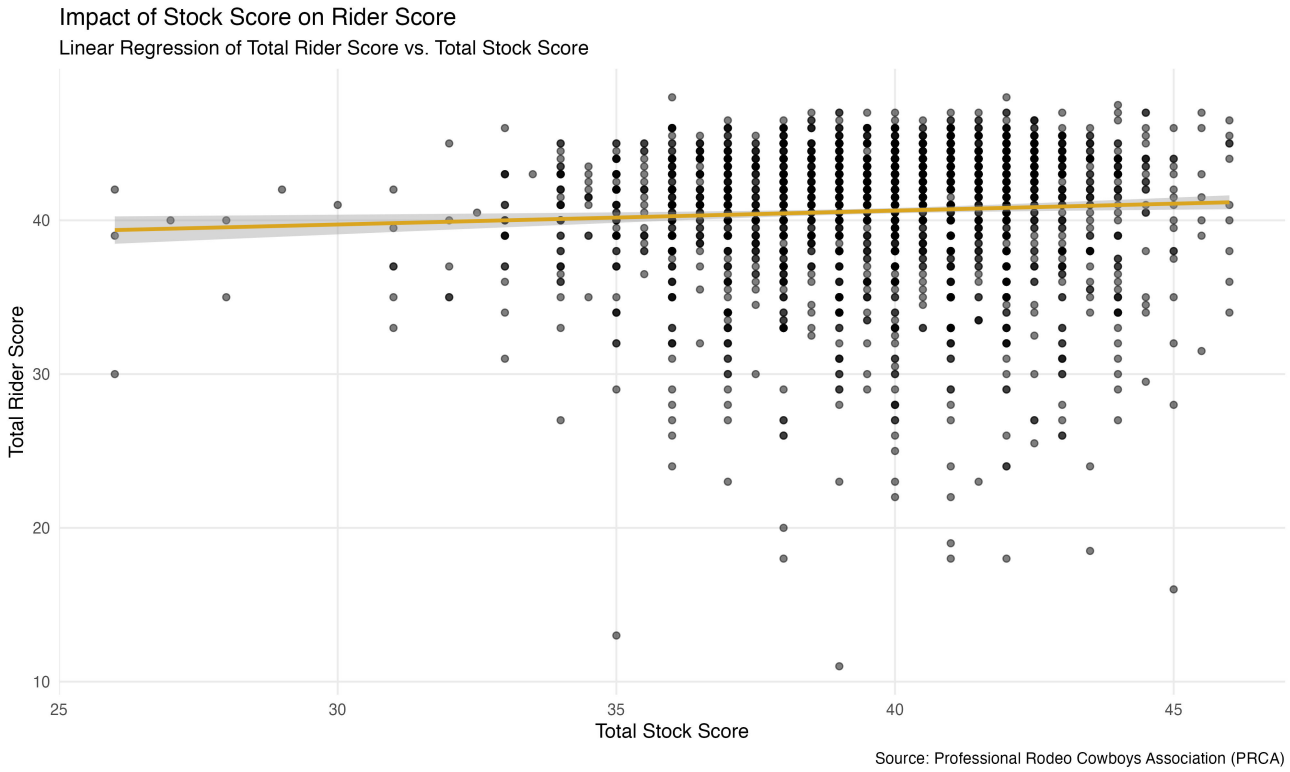

Figure 3: Linear regression of total rider score vs. total stock score

6.1 Weighting by Potential Earnings

In the distribution of prize pools across all of the rodeos in the outlined dataset (see Figure 5), there is an abundance of upper outliers, causing a skew that requires normalization. Without normalization, there would be skewed results within our weighted ranking methods favoring those who participated in higher-paying events too heavily, which in turn would create an adjusted ranking method similar to the PRCA's current rankings. As shown in Figure 6 we tried using log, square root, inverse, and Box-Cox transformations to normalize the potential earnings. A Shapiro-Wilk normality test was run on each transformation (see Table 10), which indicated that the log transformation most effectively normalized our data. Thus, the log transformation values were used to weight the classic ranking methods by potential earnings. This adjustment ensures the ranking reflects the normalized prize pool distribution, thereby minimizing bias towards higher-paying events.

Rodeo	Weight by Potential Earnings
Rodeo 1	3.8
Rodeo 2	4.0
Rodeo 3	4.4

Table 3: Example weights by potential earnings for the example in Table 1 and Figure 2

To integrate the weighted potential earnings into the Colley ranking method, we must adjust the coefficients in the Colley matrix equation accordingly. These changes can be made in the original system of equations. Thinking back to the example in Table 1, we can imagine that the three rodeos have the log-transformed potential earnings found in Table 3. The equation for Rider A would be:

$$(2+3.8+3.8+3.8+4.4+4.4)r_A = 1+\frac{3(3.8)+4.4-4.4}{2}+(3.8+4.4)r_B+3.8r_C+(3.8+4.4)r_D$$

$$22.2r_A-8.2r_B-3.8r_C-8.2r_D = 6.7.$$

The other equations in the system would be found similarly. The ratings would then be found following the steps detailed in Section 3.1.

For the Massey method, we can create a square *weight matrix* $\mathbf{W}$ with each diagonal entry filled with the weight for that given rodeo.

So, following the example, the *weight matrix* would be populated by:

$$\mathbf{W}_{ij} = \begin{cases} 3.8 & 1 \leq i = j \leq 6 \\ 4.0 & 7 \leq i = j \leq 9 \\ 4.4 & 10 \leq i = j \leq 12 \\ 0 & i \neq j \end{cases}.$$

In the Keener method, point differentials are adjusted by multiplying them by the weight given to the associated rodeo before applying Laplace's rule of succession. This approach allows the point differentials to reflect the normalized earnings, providing a more balanced and equitable basis for ranking the riders.

In a similar vein to the Colley method, the entry a_{ij} of the initial adjacency matrix for the PageRank method is determined by multiplying each "vote" from one rider to another by the corresponding weight and summing the resulting products. This ensures that the influence of each connection between riders is weighted by the normalized potential earnings, enhancing the empirical nature of the ranking.

6.2 Weighting by Rider Score

With our goal of emphasizing rider skill in mind, we also weighted rider score higher than stock score. We created a metric calculated as $(R_i * q) + S_j$, where q is the weight we give to the rider score, R_i is the rider score for rider i who rode a stock j with a stock score of S_j. Note that we have an unweighted rider score when $q = 1$. To find the most predictive q, we calculated PageRank and Massey rankings of the riders using varying values of k. We tested our new rankings and estimated their predictability based on cross-validation with the rest of the data (this process is laid out in detail in Section 7.4). Using rank-biased overlap, we also compared each ranking to a set of basic rankings (this process is laid out in detail in Section 7.3). We first ran the predictability tests on ranking the rider score q times as much, where $q \in \{1, 2, \ldots, 10\}$. Our rankings were the most accurate when q was between 1 and 4, with predictive percentages around 64.4% as opposed to around 64.2% for $q \geq 7$. From here, we tested $q \in \{1, 1.25, 1.5, \ldots, 3.75, 4\}$. Table 4 summarizes our results based on the value of q.

q	PageRank Cross-Validation Percentage	Massey Cross-Validation Percentage	Average RBO Comparison
1	60.95	62.3	0.539
1.25	62.5	63.34	0.527
1.5	62.56	63.49	0.52
1.75	62.73	63.58	0.518
2	62.71	63.54	0.495
2.25	62.8	63.58	0.494
2.5	62.69	63.52	0.494
2.75	62.76	63.46	0.492
3	62.8	63.47	0.492
3.25	62.87	63.46	0.492
3.5	62.86	63.47	0.483
3.75	62.95	63.46	0.482
4	62.85	63.46	0.476

Table 4: Rider score predictive power

From this information, we ultimately decided that the weighting rider score of 1.75 was the best value for q. Although it does not have the highest PageRank cross-validation percentage, the Massey cross-validation was the highest of any q-value, and had a large RBO comparison value. We then calculated the Colley, Massey, Keener, and PageRank ratings (as described in Section 3) using this metric instead of the total score.

7 Discussion and Conclusions

7.1 Introduction to Discussion

The results of our study show promising evidence for the use of classic ranking methods to construct more holistic ways of ranking riders. In the following sections, we explore how each ranking system behaves in the context of bareback riding, comparing them to

ranks based on simple rider metrics, investigating adjustments to the Colley method, and utilizing cross-validation techniques to examine the predictive power of our adapted ranking models. We look at the limitations and generalizations of the existing study, posing topics for future investigation.

7.2 Adjusted Colley

As discussed in Section 3.1, the Colley method is commonly applied to match-based sports. In our analysis, we found the non-adjusted unweighted (see (1)) and weighted Colley methods do not correlate well to any basic performance-based ranking systems (see Table 13). In an attempt to improve on the Colley method's application to bareback riding, we tested several variations in weights used for our Colley method. One variation, which we will call the adjusted Colley method, only considered wins in a rider's ranking and did not penalize their number of losses. As mention in Section 3.1, the Colley matrix $\mathbf{C}$ in the system $\mathbf{Cr} = \mathbf{b}$ is defined as:

$$\mathbf{C}_{ij} = \begin{cases} 2 + p_i & i = j \\ -p_{ij} & i \neq j \end{cases} \tag{7}$$

where p_{ij} is the number of games rider i and rider j played against one another and p_i is the total number of games played by rider i. In our Adjusted Colley method, the adjusted Colley coefficient matrix $\mathbf{C}_A$ is defined as:

$$\mathbf{C}_{A_{ij}} = \begin{cases} 2 + p_i & i = j \\ -w_{ij} & i \neq j \end{cases} \tag{8}$$

where $\mathbf{C}_A$ is our adjusted Colley matrix, w_{ij} is the number of winning games rider i played against rider j , and p_i remains the total number of games played by rider i. Note that, unlike the traditional Colley method, the adjusted Colley method does not create a symmetric matrix. For example, if some Player 1 and Player 2 compete against each other once and Player 1 wins, then the adjusted Colley coefficient matrix has $\mathbf{C}_{A_{12}} = -1$ and $\mathbf{C}_{A_{21}} = 0$, unlike the traditional Colley coefficient matrix which would have $\mathbf{C}_{12} = \mathbf{C}_{21} = -1$.

This adjustment gives victories a greater impact in determining the rating of a rider. The remainder of the Colley method follows the same as in Section 3.1. To exhibit this adjustment, we can use the adjusted Colley coefficient matrix $\mathbf{C}_A$ for our previous example (found in Table 1). In doing this, we obtain the rankings:

$$\begin{bmatrix} r_A \\ r_B \\ r_C \\ r_D \end{bmatrix} = \begin{bmatrix} 0.404 \\ 0.072 \\ -0.040 \\ 0.223 \end{bmatrix}.$$

In this case, the adjusted Colley method ranks the riders in the same order as the traditional Colley method (see (2)), but gives the riders different ratings. Applying the adjusted Colley method to the PRCA dataset obtained new rankings that better correlated to basic ranking metrics and to the current PRCA ranking system, as shown in Table 11. Beyond that, there is a lot of variation across all six implementations of the Colley method as displayed in Table 12. Across each of the unweighted and weighted Colley methods, the adjusted Colley has a stronger correlation to other ratings than the non-adjusted methods, thereby exhibiting the benefit in adjusting Colley's ranking method to better suit the sport of bareback riding.

7.3 Comparing Results

There are multiple methods for comparing two ordered lists of items. Since we want our comparison to emphasize differences in the people who are ranked higher rather than those ranked near the bottom, we used rank-biased overlap (RBO). Rank-biased overlap compares two lists by looking at prefixes of the list of increasing lengths. It compares how alike the prefixes are and calculates a range of how similar the whole lists can be. Then, the length of the prefix is increased and the process is repeated, but the possible similarity range is decreased. As the length of the prefix continues to increase, the length of the range approaches zero, meaning the similarity has a limit, which is the RBO value. A given RBO value output provides a measure of similarity (weighting similarity in top ranks heavier than similarity in bottom ranks) ranging from 0 to 1, with 0 being completely disjoint lists and 1 being exactly the same lists [72].

Kendall tau is another common method to compare ordered lists, which looks purely at the distances between the ranked items [1]. Using the Kendall tau method, if someone is ranked first in one ranking and second in another, those orderings are as similar to each other as two rankings where someone was ranked 50^{th} in one and 51^{st} in the other. Since we wanted to focus more on differences between people who were ranked higher versus people who were ranked near the bottom, we used RBO as described above.

One challenge in properly evaluating our ranking methods is that we have no baseline for comparison other than the PRCA rankings, which are based on total season earnings. As mentioned earlier, this is an incomplete ranking system. We created new baselines for comparison by ranking the riders based on four other statistics: total earnings, average earnings per rodeo, average rider score, and average total score. The top ten riders based on these methods of ranking are provided in Table 8.

We now use our RBO scores to compare our linear algebra methods to these more basic rankings (see Table 13). The following

PRCA Rankings	Ranking	Unweighted Colley	Unweighted Massey	Unweighted Keener	Unweighted PageRank
Keenan Hayes	1	Rocker Steiner	R.C. Landingham	R.C. Landingham	Keenan Hayes
Leighton Berry	2	Keenan Hayes	Rocker Steiner	Keenan Hayes	Jacob Lees
Dean Thompson	3	R.C. Landingham	Keenan Hayes	Clayton Biglow	Dean Thompson
Rocker Steiner	4	Ty Breuer	Ty Breuer	Jess Pope	Coop Cooke
Coop Cooke	5	Leighton Berry	Jess Pope	Dean Thompson	Tanner Aus
R.C. Landingham	6	Jess Pope	Clayton Biglow	Clayton Biglow	Cole Franks
Wacey Schalla	7	Clayton Biglow	Garrett Shadbolt	Tanner Aus	Rocker Steiner
Tanner Aus	8	Dean Thompson	Dean Thompson	Garrett Shadbolt	R.C. Landingham
Garrett Shadbolt	9	Garrett Shadbolt	Tanner Aus	Coop Cooke	Bradlee Miller
Jacob Lees	10	Tilden Hooper	Leighton Berry	Jacob Lees	Nick Pelke

Table 5: Top ten riders from the unweighted classic ranking methods

PRCA Rankings	Ranking	Earnings Weighted Colley	Earnings Weighted Massey	Earnings Weighted Keener	Earnings Weighted PageRank
Keenan Hayes	1	Rocker Steiner	Rocker Steiner	R.C. Landingham	Keenan Hayes
Leighton Berry	2	Clayton Biglow	R.C. Landingham	Jess Pope	Jacob Lees
Dean Thompson	3	Jess Pope	Keenan Hayes	Keenan Hayes	Dean Thompson
Rocker Steiner	4	Keenan Hayes	Ty Breuer	Clayton Biglow	Coop Cooke
Coop Cooke	5	Leighton Berry	Jess Pope	Tanner Aus	Tanner Aus
R.C. Landingham	6	R.C. Landingham	Garrett Shadbolt	Rocker Steiner	Cole Franks
Wacey Schalla	7	Garrett Shadbolt	Clayton Biglow	Dean Thompson	Rocker Steiner
Tanner Aus	8	Richmond Champion	Tanner Aus	Jacob Lees	R.C. Landingham
Garrett Shadbolt	9	Tanner Aus	Dean Thompson	Garrett Shadbolt	Jess Pope
Jacob Lees	10	Clint Laye	Coop Cooke	Coop Cooke	Bradlee Miller

Table 6: Top ten riders from the classic ranking methods weighted by earnings

analysis is exploratory and the statistical significance of RBO differences is beyond the scope of this paper, though it would be an interesting area for future study.

Interestingly, the unweighted Keener method resulted in the highest average RBO across all basic rankings of 0.574. This was followed by Keener weighted by rider score (0.55), Massey weighted by earning potential (0.542), unweighted Massey (0.54), and average earnings (0.519). Moreover, we can use RBO to stipulate which factors these weightings highlight when evaluating rider performance. Let's focus on unweighted Keener and PageRank weighted by rider score, our first and seventh most similar ranking methods to basic rider rankings, with average RBO scores of 0.574 and 0.49, respectively. Our unweighted Keener ranking was most similar to the average earnings per rodeo and average total score ranks, with RBO values of 0.678 and 0.603. On the other hand, PageRank was most similar to the total earnings rank, with an RBO value of 0.564 to total earnings and lower scores when compared to the other basic rankings.

In general, our weighted and unweighted Keener methods were most similar to average earnings (with an average RBO across the three methods of 0.657) and average total score (with an RBO of 0.637). On the other hand, the Keener method is not as similar to average rider score (0.421) and total earnings (0.476). Therefore, Keener also ranked riders who consistently won money at rodeos highly. However, Keener emphasized total score more than rider score, meaning Keener's method accentuated riders who constantly won some amount of money at rodeos and also got higher total scores. The higher total scores could be because these riders got higher stock scores or because they stayed on the animal more consistently and had fewer scoreless rides. Either way, it is still the ranking method closest to most of these basic rankings.

Thus, the Keener method aligns with characteristics that current rider rankings may overlook, while PageRank agrees most closely with the published PRCA rankings. In fact, PageRank is the advanced ranking system most similar to total earnings, which is surprising because unweighted PageRank is not weighted by earnings. This could possibly be explained by the fact that winners of rodeos receive more points in the PageRank method than in other methods. However, this method does not correlate very closely with either of the basic rankings based on score, which is interesting since in a rodeo, the cowboys with the highest score win. So, PageRank is a method that ranks riders with large total earnings (and not necessarily better scores) higher.

Massey was the second most similar method, aligning with average rider score and average earnings. All three Massey methods consistently had high RBO values when compared to these rankings (around 0.525 for average rider score, and 0.607 for average earnings), and lower RBO values when compared to average total score and total earnings (around 0.495 for average score and 0.454 for total earnings). Therefore, the Massey method ranked riders who perform consistently well better. In other words, cowboys who did well, no matter what stock they rode, and those who consistently won money at events were ranked highly by our Massey methods. In this respect, this method was good at finding consistent riders, who may not always come in first, but will usually win some amount of money in any given rodeo.

PRCA Rankings	Ranking	Rider Score Weighted Colley	Rider Score Weighted Massey	Rider Score Weighted Keener	Rider Score Weighted PageRank
Keenan Hayes	1	Keenan Hayes	Rocker Steiner	R.C. Landingham	Keenan Hayes
Leighton Berry	2	R.C. Landingham	Clayton Biglow	Jess Pope	Jacob Lees
Dean Thompson	3	Jess Pope	R.C. Landingham	Keenan Hayes	Dean Thompson
Rocker Steiner	4	Rocker Steiner	Jess Pope	Clayton Biglow	R.C. Landingham
Coop Cooke	5	Clayton Biglow	Keenan Hayes	Dean Thompson	Rocker Steiner
R.C. Landingham	6	Leighton Berry	Ty Breuer	Rocker Steiner	Tanner Aus
Wacey Schalla	7	Dean Thompson	Dean Thompson	Tanner Aus	Coop Cooke
Tanner Aus	8	Tilden Hooper	Leighton Berry	Garrett Shadbolt	Cole Franks
Garrett Shadbolt	9	Ty Breuer	Tilden Hooper	Mason Clements	Jess Pope
Jacob Lees	10	Garrett Shadbolt	Garrett Shadbolt	Jacob Lees	Bradlee Miller

Table 7: Top ten riders from the classic ranking methods weighted by rider score

Rank	Total Earnings	Average Earnings Per Rodeo	Average Rider Score	Average Total Score
1	Leighton Berry	Leighton Berry	Jess Pope	Cole Reiner
2	Keenan Hayes	Rocker Steiner	R.C. Landingham	Dean Thompson
3	Rocker Steiner	Coop Cooke	Cole Reiner	Keenan Hayes
4	Coop Cooke	Garrett Shadbolt	Keenan Hayes	R.C. Landingham
5	Dean Thompson	Tim Kent	Richmond Champion	Rocker Steiner
6	Garrett Shadbolt	Keenan Hayes	Dean Thompson	Leighton Berry
7	Tanner Aus	Cole Reiner	Jayco Roper	Jess Pope
8	Jacob Lees	Clayton Biglow	Waylon Bourgeois	Garrett Shadbolt
9	Taylor Broussard	Dean Thompson	Jacob Lees	Tanner Aus
10	Bradlee Miller	Tanner Aus	Kade Sonnier	Waylon Bourgeois

Table 8: Top ten riders based on various basic ranking metrics

Lastly, we noticed that both the adjusted and non-adjusted Colley methods did not align closely with any of our basic rankings. Unweighted Colley had an average RBO score of 0.505, while our adjusted Colley method had an average RBO score of 0.442. Both Colley methods did not highlight any specific attributes of the data.

7.4 A Note on Predictions

We also tested how well all of our ranking systems could predict which of two riders would do better in an event. To do this, we chose one of our rankings and examined all pairs of riders that competed in the same round of a rodeo. If the rider who scored better was ranked higher by the ranking, we considered that a success. Then, we calculated the percentage of correct guesses to get a predictive accuracy for that ranking. The percentages are provided in Table 9.

Furthermore, we also used MatLab's built-in classification learner to test the predictive power of our assorted models. We used multiple methods, including neural networks, linear mixed models, support vector machines, logistic regressions, k-nearest neighbors, and decision trees. These have built in cross-validation methods, so we used those to evaluate how effective our classification methods were. To use the classification learner, we built a matrix where each row contained the difference between the first and second cowboys' scores in unweighted and weighted Massey, Colley, Keener, and PageRank as well as which rider won. The classification learner used this data to predict who would win, and also provided data on which points were predicted incorrectly, which methods were most accurate in predicting, and whether or not there were any matchups that were especially hard to guess the winner in.

From these predictive tests, the unweighted adjusted Colley ranking was the best predictor, which predicted the winner 63.72% of the time. This was followed by unweighted Massey and Massey weighted by earning potential, which predicted the better rider 63.7% of the time and 63.66% of the time, respectively. Other than Colley and Massey, the next best predictive model was unweighted Keener, with a predictability of 63.43%. The weighted Keener methods had similar measures with 62.89% and 62.39% accuracy. This was followed by two of the basic rankings: ranking by average total score had an accuracy of 62.31% and by average earnings per rodeo had an accuracy of 62.01%. All of these measures were markedly better than the total earnings metric that the PRCA uses, which only predicted the better rider 60.13% of the time. Therefore, if we want to predict which rider will do better in an event it is around 3.6% better to use the unweighted adjusted Colley method than the PRCA's total earnings system.

After testing our methods individually, we wanted to determine if a mix of multiple methods would be able to predict more accurately than 63.72%. To test to see if we could do better, we used Matlab's classification learner as outlined earlier. After testing the six classification models listed above, we found that the most accurate was a generalized linear mixed model, which had an accuracy of 63.9%. This was followed by a linear regression model, which had an accuracy of 63.8%. Interestingly, all of the methods with a cross-validation accuracy of over 63% relied heavily on the unweighted Massey predictor, despite the unweighted adjusted Colley having the highest accuracy on its own. The generalized linear mixed model (GLMM) as well as the optimizable neural network had an accuracy of 63.9% and 63.3%, respectively, and used the unweighted Massey scores as the strongest predictor. The cap on our predictability using these more advanced methods was 63.9%. So, while it could be marginally more accurate to use these machine learning methods,

Ranking Method	Percentage of Correct Predictions
Unweighted Adjusted Colley	63.72
Unweighted Massey	63.70
Massey Weighted by Earning Potential	63.66
Unweighted Colley	63.62
Massey Weighted by Rider Score	63.56
Adjusted Colley Weighted by Rider Score	63.49
Colley Weighted by Rider Score	63.46
Unweighted Keener	63.43
Keener Weighted by Rider Score	62.89
Keener Weighted by Earning Potential	62.39
Average Total Score	62.31
Colley Weighted by Earning Potential	62.10
Average Earnings	62.01
Average Rider Score	61.21
PageRank Weighted by Rider Score	60.72
Total Earnings (used by the PRCA)	60.13
PageRank Weighted by Earning Potential	60.12
Unweighted PageRank	59.97
Average Place	59.67
Adjusted Colley Weighted by Earning Potential	37.84

Table 9: Ranking method along with its cross-validation accuracy

Massey's method gave a similar rate of prediction with less time commitment to training models and cleaning the data to be able to feed it to Matlab's classification learner.

Despite bareback riding's seemingly unpredictable nature (see Section 7.5), our top-performing models are comparable to other linear and statistical models deemed successful in other sports (see [6], [22], [34], and [39].

7.5 Limitations

The application of mathematical ranking methods to the rodeo is relatively unexplored territory in the field of sports analytics, and with the creation of novel work come obstacles and limitations. First and foremost was the lack of a central database for rodeo records. The impracticality of individually collecting data from the smaller, local events limited the scope of our analysis to only PRCA-sanctioned rodeos, which had performance data available on the PRCA site [49].

It is important to note that the PRCA's site still had some incomplete and inconsistent data, forcing us to make decisions in data selection to maintain uniformity. For example, the PRCA reports a rider's *total earnings* (their official ranking metric) and the breakdown of their winnings by rodeo, the sum of which did not always equal their *total earnings*. We proceeded by using the summed earnings in our ranking algorithms and, for our comparison to PRCA standings, using the reported earnings to rank riders in places 51-71 as the PRCA would, since the PRCA only publicly announces the top 50 riders at any given moment.

Beyond general data unavailability, the rodeo poses unique challenges for common focuses of sports analysis. In judged sports, analysts have conducted numerous studies on judging bias [14, 17, 18, 35, 43]. However, an extensive investigation on judging bias in bareback riding was inhibited, as we were unable to find consistent information on judges, their individual scores, or any sort of rider evaluation. Thus, we did not account for judging variability in our rankings.

The traditional ranking methods adapted in this paper have also been used to create predictive ranking models, another focus of the sports analytics community [7, 16]. However, the application of predictive modeling to timed events may not yet be feasible due to the lack of both rider and stock performance data, a product of the high turnover of both riders and stock. For riders, we see high rates of injury [63] and, since only the top 50 are ranked in the PRCA, there is not easily accessible performance data for the vast majority of riders. Moreover, stock are minimally ridden across the PRCA, and contractors are continuously improving and developing technologies and techniques to cultivate competitive stock. In addition, we found a limited availability of both rider and stock performance metrics. Judging breakdowns that delve into specific aspects of a cowboy's ride are not easily available, although judging guidelines are available in the PRCA judging handbook (see pages 13-16 and page 21 in [4]). On the other hand, the official judging handbook of the PRCA does not have a set rubric for scoring stock difficulty (see pages 14-18 in [4] for more information), making it challenging to empirically analyze the performance of the stock. We feel that a predictive model based purely on final judging scores would lack nuance and not accurately represent the dynamics of timed events. With that being said, there has been a rise in rodeo analytics, such as the rodeo

database, OMNi [60], and the barrel racing analysis firm Rodeo Analytics [65]. The PRCA has also expressed a hope to analyze stock performances in the future using movement tracking devices [66]. Despite all of these limitations, our model predictability of 63.7% is quite impressive as described in Section 7.4.

7.6 Future Studies

The present study leads us to many valuable future investigations in the pursuit of rodeo analysis. Foremost is a study using a similar procedure to investigate how classic ranking methods apply to saddle bronc riding and bull riding, the two other roughstock events. While data for saddle bronc riding would be collected similarly to bareback riding, those studying bull riding would be able to pull information from both the PRCA website and the Professional Bull Riders (PBR) website. The hope is that our methodologies can go beyond bareback riding and further the exploration of roughstock analysis.

Additionally, we believe a more extensive study on judging bias would be a great contribution to rodeo analytics. According to the PRCA handbook, judges are instructed to "use the full spread when at all possible and don't hesitate to mark the top of the spread when you see something outstanding, either rider or horse" [4]. This opens the door to potential new forms of bias caused by relative scoring. A more thorough investigation of judging dynamics can improve ranking and predictive models. Furthermore, if data can be found with additional information about the rider and rodeo (like rider experience and environmental conditions), it would be interesting to see if these influence rider scores. Also, more exploration could be done related to what conditions impact when a rider is bucked off.

Finally, as officially reported, circuits differ only in their location. However, we believe there is a possibility that a rodeo insider may have knowledge of the strength of different circuits or events. There are no quantitative studies that have been done to assess the strength of each circuit. In conducting this study, the results could give a better insight into the difficulty of circuit or rodeo, the caliber of competition between riders across circuits, and ultimately give a greater holistic view into the ranking of athletes within rodeo.

8 Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. DMS-1929284 while the authors were in residence at the Institute for Computational and Experimental Research in Mathematics in Providence, RI, during the Summer@ICERM 2024 program.

References

[1] Hervé Abdi. The kendall rank correlation coefficient. *Encyclopedia of measurement and statistics*, 2:508–510, 2007.

[2] Mia Alder, Ford McDill, Tiffanie Ng, and Will Paz. Github money bull code, 2024. `https://github.com/WillPaz16/moneybull`.

[3] Mary Allender et al. Predicting the outcome of NASCAR races: The role of driver experience. *Journal of Business & Economics Research (JBER)*, 6(3), 2008.

[4] Professional Rodeo Cowboys Association. PRCA 2023 judging handbook. Online; Accessed on July 25, 2024, `https://www.prorodeo.org/Documents/Portal/Home/PrcaBusiness/2023-PRCA-Judging-Handbook.pdf`.

[5] Professional Rodeo Cowboys Association. PRCA 2023 rule book. Online; Accessed on June 25, 2024, `https://www.prorodeo.org/Documents/Portal/Home/PrcaBusiness/2023-PRCA-Rule-Book.pdf`.

[6] Daniel Barrow, Ian Drayer, Peter Elliott, Garren Gaut, and Braxton Osting. Ranking rankings: an empirical comparison of the predictive power of sports ranking methods. *Journal of Quantitative Analysis in Sports*, 9(2):187–202, 2013.

[7] Daniel Barrow, Ian Drayer, Peter Elliott, Garren Gaut, and Braxton Osting. Ranking rankings: an empirical comparison of the predictive power of sports ranking methods. *Journal of Quantitative Analysis in Sports*, 9(2):187–202, 2013.

[8] Randy Bean. NFL extends moneyball to a new level of professional sports leadership, 2021. Online; Accessed on June 27, 2024, `https://www.forbes.com/sites/randybean/2021/11/26/nfl-extends-moneyball-to-a-new-level-of-professional-sports-leadership/?sh=183658542784`.

[9] Clive B Beggs, Simon J Shepherd, Stacey Emmonds, and Ben Jones. A novel application of pagerank and user preference algorithms for assessing the relative performance of track athletes in competition. *PLoS One*, 12(6):e0178458, 2017.

[10] F. Boen, K. Van Hoye, Y. Vanden Auweele, J. Feys, and T. Smits. Open feedback in gymnastic judging causes conformity bias based on informational influencing. *Journal of sports sciences*, 26(6):621–628, 2008.

[11] Matt Buckner. The evolution of rodeo: How the contemporary professional rodeo cowboy evolved from the working cowboy of the 1930s era. *Texas Tech University*, 2000.

[12] Alex Cawthon. PRCAcircuit system explained, 2023. Online; Accessed on June 25, 2024, `https://www.si.com/fannation/rodeo/pro-rodeo/prca-circuit-system-explained-alex9`.

[13] Timothy P Chartier. *When Life is Linear: From Computer Graphics to Bracketology*. Mathematical Association of America, 2015.

[14] Diana Cheng and John Gonzalez. Technical and program component scores frozen together: Difficulty bias and outcome prediction in international figure skating. *Mathematics and Sports*, 4(1), 2022.

[15] Wesley Colley. Colley's bias free college football ranking method. *Princeton, NJ, USA: Princeton University*, 2002.

[16] Shaunak S Dabadghao and Baback Vaziri. The predictive power of popular sports ranking methods in the NFL, NBA, and NHL. *Operational Research*, 22(3):2767–2783, 2022.

[17] Ana E Diaz, Mary S Johnston, Jennifer Lucitti, Wendi S Neckameyer, and Katy M Moran. Scoring variables and judge bias in united states dressage competitions. *Journal of Quantitative Analysis in Sports*, 6(3), 2010.

[18] John W Emerson and Silas Meredith. Nationalistic judging bias in the 2000 olympic diving competition. *Math Horizons*, 18(3):8–11, 2011.

[19] J.W. Emerson and S. Meredith. Nationalistic judging bias in the 2000 olympic diving competition. *Math Horizons*, 18(3):8–11, 2011.

[20] William Feller et al. *An Introduction to Probability Theory and Its Applications*. John Wiley, New York, 1971.

[21] L.C. Findlay and D.M. Ste-Marie. A reputation bias in figure skating judging. *Journal of Sport and Exercise Psychology*, 26(1):154–166, 2004.

[22] Joseph A Gallian. *Mathematics and sports*. Number 43 in 1. MAA Press, Washington DC, 2010.

[23] Bill Gerrard. Beyond moneyball: Using data analytics to improve performance in elite team sports. *REVIEW PAPERS*, page 3, 2016.

[24] Anjela Govan and Carl D Meyer Jr. Ranking national football league teams using google's pagerank. Technical report, North Carolina State University. Center for Research in Scientific Computation, 2006.

[25] Anjela Y Govan, Amy N Langville, and Carl D Meyer. Offense-defense approach to ranking team sports. *Journal of Quantitative Analysis in Sports*, 5(1), 2009.

[26] Anjela Yuryevna Govan et al. Ranking theory with application to popular sports. *NC State Repository*, 2008.

[27] Leila Hamdad, Karima Benatchba, Fella Belkham, and Nesrine Cherairi. Basketball analytics. data mining for acquiring performances. In *Computational Intelligence and Its Applications: 6th IFIP TC 5 International Conference, CIIA 2018, Oran, Algeria, May 8-10, 2018, Proceedings 6*, pages 13–24. Springer, 2018.

[28] Amanda Harsy and Michael Smith. *Teaching Mathematics Through Cross-curricular Projects*, chapter 26: Get in the Game with Linear Algebra, pages 339–350. MAA Press, Washington DC, 2024.

[29] S. Heiniger and H. Mercier. Judging the judges: evaluating the accuracy and national bias of international gymnastics judges. *Journal of Quantitative Analysis in Sports*, 17(4):289–305, 2021.

[30] Bill James. *The new Bill James historical baseball abstract*. Free Press, Michigan, 2010.

[31] James P Keener. The Perron–Frobenius theorem and the ranking of football teams. *SIAM review*, 35(1):80–93, 1993.

[32] Leonid Kholkine, Thomas Servotte, Arie-Willem De Leeuw, Tom De Schepper, Peter Hellinckx, Tim Verdonck, and Steven Latré. A learn-to-rank approach for predicting road cycling race outcomes. *Frontiers in sports and active living*, 3:714107, 2021.

[33] Jason Kolbush and Joel Sokol. A logistic regression/markov chain model for american college football. *International Journal of Computer Science in Sport*, 16(3):185–196, 2017.

[34] Eiji Konaka. A unified statistical rating method for team ball games and its application to predictions in the olympic games. *IEICE TRANSACTIONS on Information and Systems*, 102(6):1145–1153, 2019.

[35] Alex Krumer, Felix Otto, and Tim Pawlowski. Nationalistic bias among international experts: Evidence from professional ski jumping. *The Scandinavian Journal of Economics*, 124(1):278–300, 2022.

[36] Paul Kvam and Joel S Sokol. A logistic regression/markov chain model for NCAA basketball. *Naval Research Logistics (NrL)*, 53(8):788–803, 2006.

[37] Amy N Langville and Carl D Meyer. *Google's PageRank and beyond: The science of search engine rankings*. Princeton university press, New Jersey, 2006.

[38] Amy N Langville and Carl D Meyer. *Who's #1? The science of rating and ranking*. Princeton University Press, New Jersey, 2012.

[39] Amy N Langville and Carl D Meyer. *Who's# 1? The science of rating and ranking*. Princeton University Press, New Jersey, 2012.

[40] Michael Lewis. *Moneyball: The art of winning an unfair game*. WW Norton & Company, New York, 2004.

[41] Kenneth Massey. Statistical models applied to the rating of sports teams. *Bluefield College*, 1077, 1997.

[42] Carl Meyer. Perron–frobenius theory of nonnegative matrices. *Matrix analysis and applied linear algebra, SIAM*, 2000.

[43] Hillary N Morgan and Kurt W Rotthoff. The harder the task, the higher the score: Findings of a difficulty bias. *Economic Inquiry*, 52(3):1014–1026, 2014.

[44] H.N. Morgan and K.W. Rotthoff. The harder the task, the higher the score: Findings of a difficulty bias. *Economic Inquiry*, 52(3):1014–1026, 2014.

[45] Bret R Myers, Michael Burns, Brian Q Coughlin, and Edward Bolte. On the development and application of an expected goals model for lacrosse. *The Sport Journal*, 2021.

[46] PRCA Sports News. About the PRCA. Online; Accessed on July 12, 2024, `https://www.prorodeo.com/prorodeo/rodeo/about-the-prca#`.

[47] PRCA Sports News. PRCA playoff series bareback riding rodeo standings 2023. Online; Accessed on July 1, 2024, `https://www.prorodeo.com/standings?eventType=BB&standingType=tour&id=17&circuitId=&year=2023`.

[48] PRCA Sports News. PRCA playoff series Online; Accessed on July 1, 2024 `https://www.prorodeo.com/standings?eventType=BB&standingType=tour&id=17&circuitId=&year=2024`.

[49] PRCA Sports News. PRCA sports news. Online; Accessed July 1, 2024, `https://www.prorodeo.com`.

[50] PRCA Sports News. Rodeo terminology. Online; Accessed on June 25, 2024, `https://prorodeo.com/bareback-riding-101`.

[51] PRCA Sports News. Wrangler nfr. Online; Accessed on June 26, 2024, `https://www.prorodeo.com/nfr`.

[52] ProRodeo Hall of Fame. Rodeo events. Online; Accessed June 25, 2024, `https://www.prorodeohalloffame.com/rodeo-events/`.

[53] Berfin Serdil Örs. The effect of difficulty and execution scores on total ranking during 2019 rhythmic gymnastics world championships. *African Educational Research Journal*, 8(1):37–42, 2020.

[54] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford infolab, 1999.

[55] Sławomir Pietrzak, Tomasz Próchniak, and Katarzyna Osińska. The influence of certain factors on the results obtained by horses classified in dressage ranks of international federation for equestrian sports. *Annales Universitatis Mariae Curie-Skłodowska. Sectio EE: Zootechnica*, 31(3):27–33, 2013.

[56] Sean Pradhan and Yann Abdourazakou. "power ranking" professional circuit esports teams using multi-criteria decision-making (mcdm). *Journal of Sports Analytics*, 6(1):61–73, 2020.

[57] PRCA Sports News. World bareback riding rodeo standings 2023. Online; Accessed on July 1, 2024, `https://www.prorodeo.com/standings?eventType=BB&standingType=world&id=&circuitId=&year=2023`.

[58] PRCA Sports News. World bareback riding rodeo standings 2024. Online; Accessed on July 1, 2024, `https://www.prorodeo.com/standings?eventType=BB&standingType=world&id=&circuitId=&year=2024`.

[59] Caldwell Night Rodeo. Rodeo 101. https://caldwellnightrodeo.com/rodeo101/. Online; Accessed July 16, 2024, https://caldwellnightrodeo.com/rodeo101/.

[60] Rodeo Results. Revolutionizing data tracking & uniting western sports: The power of rodeo results & update omni, 2023. Online; Accessed July 23, 2024, https://www.rodeoresults.com/revolutionizing-data-tracking-connecting-western-sports-the-power-of-rodeo-results-software-update-omni/.

[61] James E Salzman and JB Ruhl. Who's number one? In *Environmental Forum*, volume 26, page 36, 2009.

[62] A. Sandberg. Competing biases: Effects of gender and nationality in sports judging. *Stockholm School of Economics, mimeo*, 2014.

[63] Clinton L Seifert, Mark Rogers, Stephen D Helmer, Jeanette G Ward, and James M Haan. Rodeo trauma: outcome data from 10 years of injuries. *Kansas journal of medicine*, 15:208, 2022.

[64] SoccerTake.com. Soccer analytics: How data is changing the game. Online; Accessed on June 27, 2024, https://www.soccertake.com/performance/soccer-analytics-how-data-is-changing-the-game.

[65] SPN Team. Rodeo analytics wants to shake up barrel racing, 2015. Online; Accessed July 25, 2024, https://siliconprairienews.com/2015/07/rodeo-analytics-wants-to-shake-up-barrel-racing/.

[66] Julia Stumbaugh. Money bull: Data analytics has potential to change rodeo, 2020. Online; Accessed on July 23, 2024, https://www.buffalobulletin.com/sports/article_7f6a50d6-e176-11ea-ab3d-73d9213ca2a6.html.

[67] Marta Teneva and Nikolay Ivanov. Sports analytics: A complete handbook for organizations, 2023. Online; Accessed on June 27, 2024, https://b-eye.com/blog/sports-analytics-a-complete-handbook-for-organizations/.

[68] Chris La Tray. A brief history of the rodeo: The humble origins and complex future of cowboy competition, 2022. Online; Accessed August 18, 2024, https://www.smithsonianmag.com/history/brief-history-rodeo-180980341.

[69] Cassie B Trewin, William G Hopkins, and David B Pyne. Relationship between world-ranking and olympic performance of swimmers. *Journal of sports sciences*, 22(4):339–345, 2004.

[70] Michel Truchon. Aggregation of rankings in figure skating. Cirpee working paper, Interuniversity Center on Risk, Economic Policies and Employment, Department of Economics Laval University, Laval University, 2004.

[71] Baback Vaziri, Shaunak Dabadghao, Yuehwern Yih, and Thomas L Morin. Properties of sports ranking methods. *Journal of the Operational Research Society*, 69(5):776–787, 2018.

[72] William Webber, Alistair Moffat, and Justin Zobel. A similarity measure for indefinite rankings. *ACM Transactions on Information Systems (TOIS)*, 28(4):1–38, 2010.

A Figures

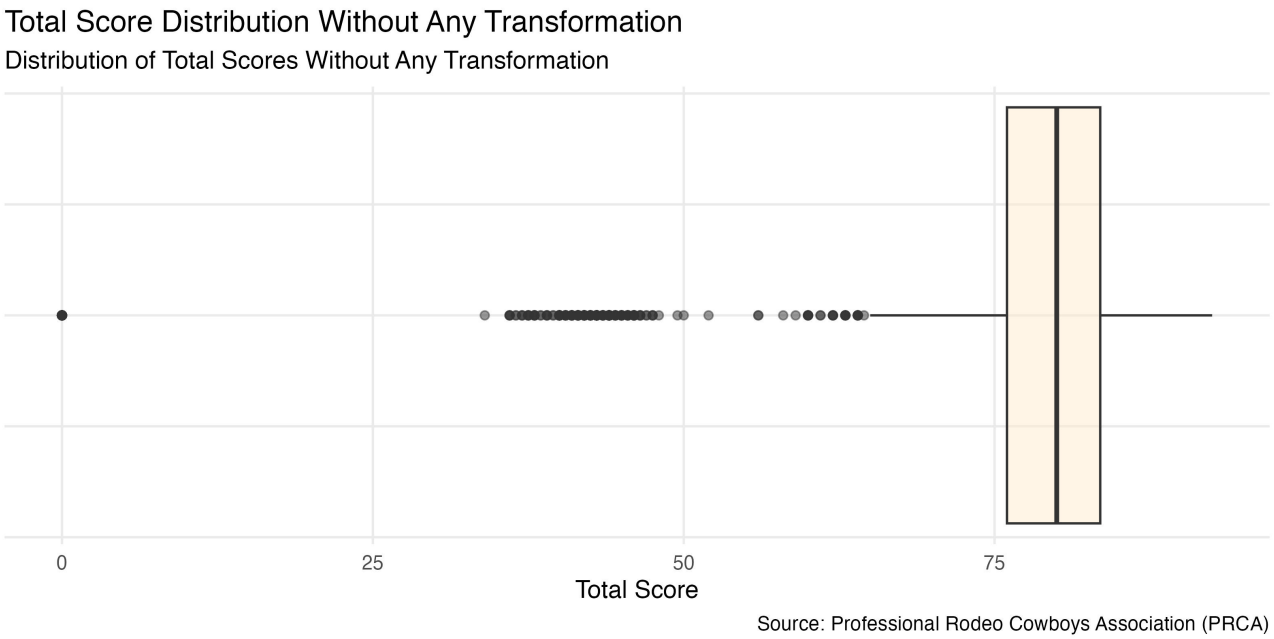

Figure 4: Boxplot of the distribution of total score across entire dataset

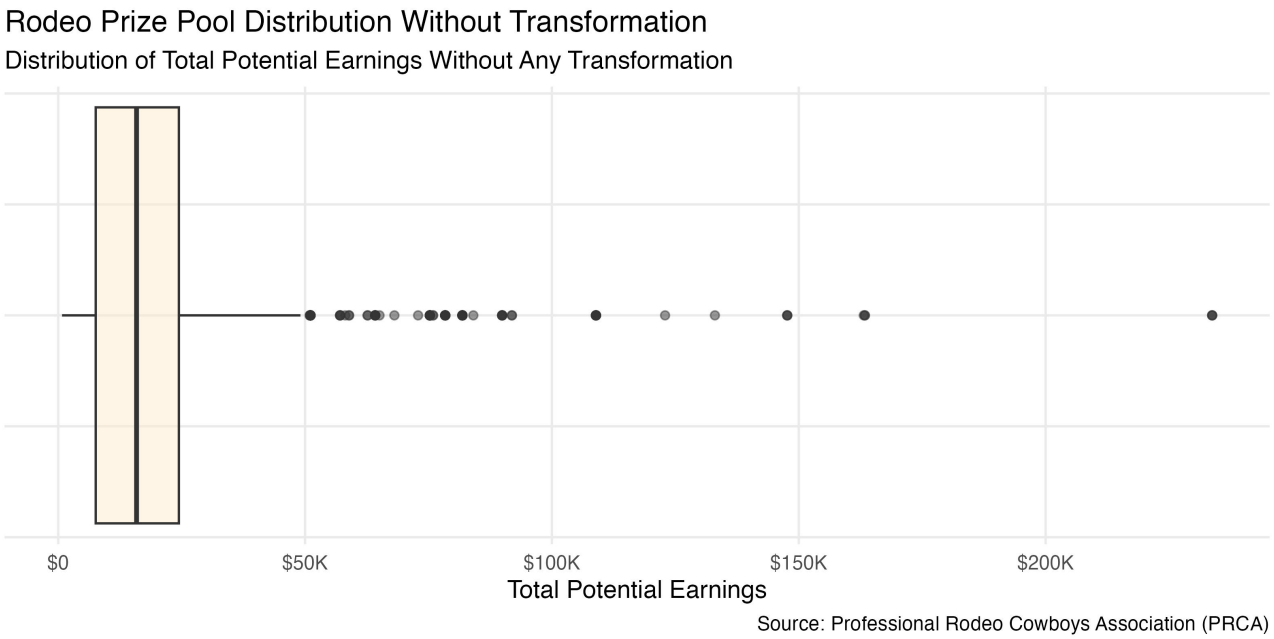

Figure 5: Boxplot of the raw distribution of total bareback prize pools

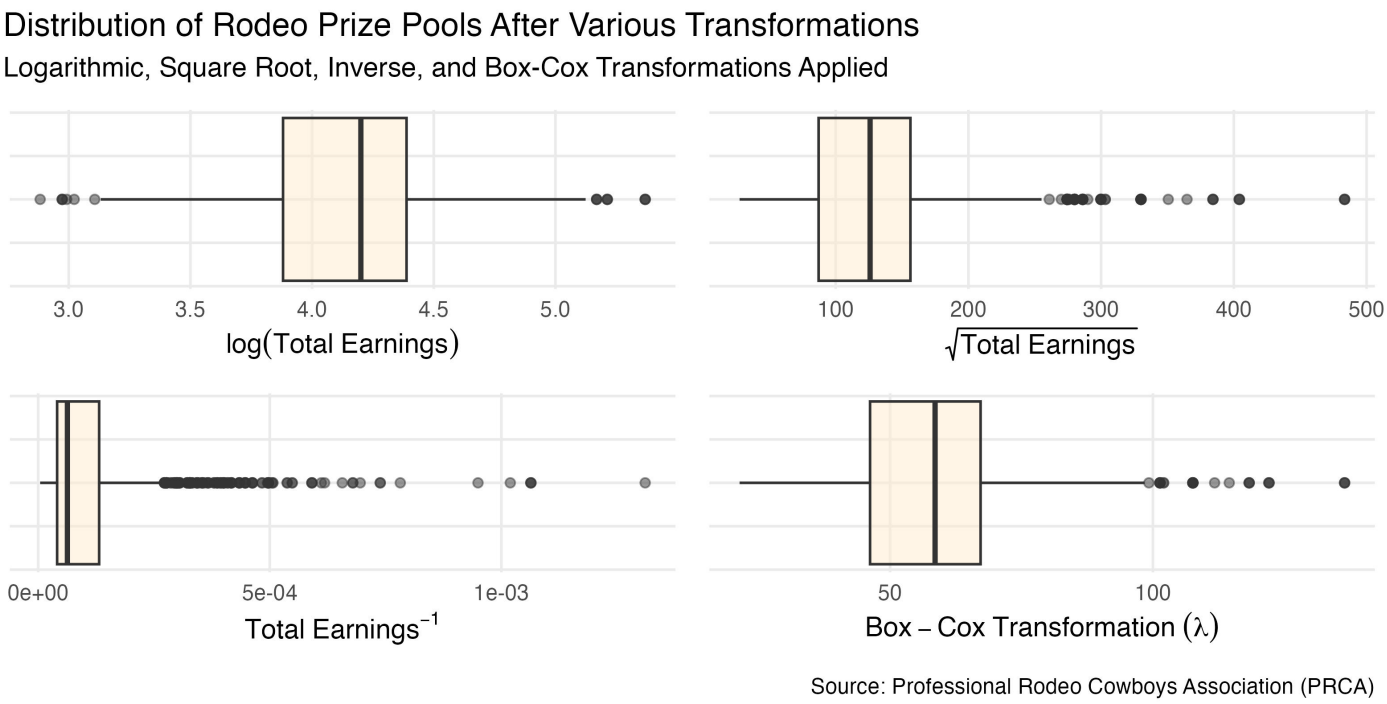

Figure 6: Boxplot of the logarithmic, squareroot, inverse, and Box-Cox transformations applied to total bareback prize pools (see Figure 5)

B Tables

Transformation	**W-score**	**p-value**
$\log(\text{Total Earnings})$	0.9883	5.663×10^{-12}
$\sqrt{\text{Total Earnings}}$	0.9235	$< 2.2 \times 10^{-16}$
$\text{Total Earnings}^{-1}$	0.6609	$< 2.2 \times 10^{-16}$
Box-Cox Transformation (λ)	0.9740	$< 2.2 \times 10^{-16}$

Table 10: Shapiro-Wilk normality test results

Statistic	**Value**
Correlation coefficient (r)	0.056
p-value	0.006

Table 11: Correlation test results

Ranking	Unweighted Colley	Earnings Weighted Colley	Rider Score Weighted Colley
1	Rocker Steiner	Rocker Steiner	Keenan Hayes
2	Keenan Hayes	Clayton Biglow	R.C. Landingham
3	R.C. Landingham	Jess Pope	Jess Pope
4	Ty Breuer	Keenan Hayes	Rocker Steiner
5	Leighton Berry	Leighton Berry	Clayton Biglow
6	Jess Pope	R.C. Landingham	Leighton Berry
7	Clayton Biglow	Garrett Shadbolt	Dean Thompson
8	Dean Thompson	Richmond Champion	Tilden Hooper
9	Garrett Shadbolt	Tanner Aus	Ty Breuer
10	Tilden Hooper	Clint Laye	Garrett Shadbolt
Ranking	**Adjusted Unweighted Colley**	**Adjusted Weighted Colley (Earnings)**	**Adjusted Weighted Colley (Rider Score)**
1	Rocker Steiner	Briar Dittmer	Keenan Hayes
2	Ty Breuer	Andy Gingerich	R.C. Landingham
3	Keenan Hayes	Ben Kramer	Jess Pope
4	R.C. Landingham	Ty Fast Taypotat	Rocker Steiner
5	Leighton Berry	Ethan Crouch	Clayton Biglow
6	Jess Pope	Jacob Raine	Ty Breuer
7	Clayton Biglow	Dylan Riggins	Leighton Berry
8	Dean Thompson	Jade Taton	Dean Thompson
9	Garrett Shadbolt	Tuker Carricato	Tilden Hooper
10	Tilden Hooper	Kody Lamb	Garrett Shadbolt

Table 12: Top ten riders from the unweighted, weighted, and adjusted colley ranking methods

Ranking Method	Average Rider Score	Average Earnings	Average Score	Total Earnings	Average RBO
Unweighted Keener	0.466	0.678	0.603	0.548	0.574
Keener Weighted by Earning Potential	0.398	0.618	0.622	0.439	0.519
Keener Weighted by Rider Score	0.398	0.674	0.687	0.441	0.55
Unweighted Colley	0.554	0.555	0.452	0.534	0.523
Unweighted Adjusted Colley	0.552	0.524	0.421	0.488	0.496
Colley Weighted by Earning Potential	0.514	0.488	0.405	0.426	0.458
Adjusted Colley Weighted by Earning Potential	0.27	0.372	0.29	0.298	0.308
Colley Weighted by Rider Score	0.45	0.624	0.583	0.484	0.535
Adjusted Colley Weighted by Rider Score	0.435	0.612	0.572	0.473	0.523
Unweighted Massey	0.529	0.641	0.516	0.472	0.54
Massey Weighted by Earning Potential	0.536	0.64	0.51	0.483	0.542
Massey Weighted by Rider Score	0.511	0.539	0.46	0.406	0.479
Unweighted PageRank	0.405	0.462	0.381	0.578	0.456
PageRank Weighted by Earning Potential	0.41	0.47	0.388	0.58	0.462
PageRank Weighted by Rider Score	0.392	0.555	0.449	0.564	0.49
Average Rider Score	N/A	0.426	0.288	0.762	0.492
Average Earnings	0.426	N/A	0.671	0.461	0.519
Average Total Score	0.288	0.671	N/A	0.319	0.426
Total Earnings (used by the PRCA)	0.762	0.461	0.319	N/A	0.514

Table 13: Comparisons based on rank-biased overlap.