

# **Inside the Octagon: Examining Objectivity in MMA Judging Through Data**

## **Abstract**

Our analysis of over 500 UFC split-decision fights shows inconsistencies in judging due to subjective preferences for certain fighting styles. Implementing comprehensive training, employing judges with diverse expertise, and increasing transparency can improve reliability. More objective judging is crucial for maintaining fan trust and enhancing the sport's marketability.

15.285: Sports Strategy and Analytics  
Ethan Fahimi, Zack Horton, Tanner Street  
fahimi@mit.edu, zhhorton@mit.edu, street24@mit.edu

## **Introduction**

Imagine a championship fight where the verdict leaves fans debating, not celebrating. This was the reality after the UFC 227 bout between champion Demetrious Johnson and challenger Henry Cejudo. In a split decision, the title was awarded to elite wrestler Cejudo, despite many believing that Johnson was the rightful winner due to his better striking. This decision divided the MMA community and sparked a wider conversation about the objectivity and consistency of MMA judging. Such controversial outcomes highlight the need for a reassessment of how fights are scored and judges are trained.

## **Background**

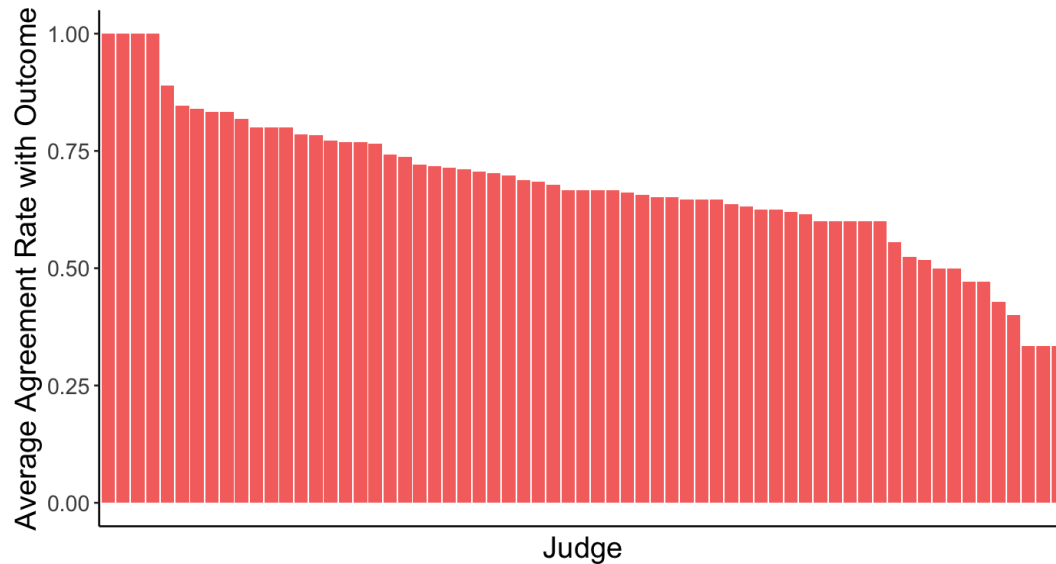
Mixed Martial Arts (MMA) is a complex sport that combines various fighting styles, from striking to grappling. Judges score fights based on a 10-9 must system, where the winner of each round receives ten points and the loser receives at most nine. However, the subjectiveness of this system and the multifaceted nature of the sport often lead to inconsistent and controversial decisions. One of the most controversial outcomes is a split-decision, which occurs when one judge disagrees with the other two.

## **Data**

We collected data from over 500 UFC fights between 2003 and 2021 that were split-decisions. The data includes statistics from each fight, such as total strikes landed, knockdowns, control time, and more. We also compiled judge scores and media scores sourced from [mmadecisions.com](http://mmadecisions.com).

## **Analysis**

To observe if significant disparities exist among MMA judges, we focused on the "agreement rate" between judges' decisions and fight outcomes. The agreement variable equals one if the judge agrees with the outcome and zero if the judge dissents. Figure 1 displays the average agreement rate for each judge, revealing significant variation in their decisions. Indeed, some judges rarely dissent while others dissent for most of the split-decisions they judge. Figure 2 in the appendix presents a comparable graph, but it plots agreement rates with media decisions, not the actual fight outcomes.



*Figure 1: Average agreement rate with fight outcome for all judges with minimum 5 fights judged.*

We hypothesized that the subjectivity in MMA judging stems from judges' personal preferences for certain fighting styles, a pertinent issue given MMA's wide array of techniques. To analyze this, we employed k-means clustering to categorize fights based on performance statistics, focusing on the differences between winners and losers. For example, we considered the control time difference, where a negative value indicates less control by the winner.

We created six distinct categories, with Table 2 in the appendix showing the average statistics for each category. Using this, we created Table 1, which summarizes the categories based on three general characteristics: power, striking, and grappling. The table categorizes these characteristics from significant advantages to disadvantages, clearly indicating the areas where the winner excelled and where they fell short in each fight. For each category, we created Figures 3-15 in the appendix to examine average agreement rate at a more granular level.

Table 1: Summary of fight categories obtained from clustering algorithm.

Category	Power Differential	Striking Differential	Grapple Differential
1	XX	✓	✓
2	X	—	X
3	—	X	✓✓
4	X	XX	—
5	✓✓	—	XX
6	✓	✓✓	X

Weighting Scheme:

- *Power Differential*: heavily weighing knockdown differential, also considering significant strikes landed differential
- *Striking Differential*: heaving weighting total strikes landed and significant strikes landed, also considering knockdown differential and attempted strike differentials
- *Grapple Differential*: equally weighing control time differential and takedown success differential

(XX=significantly negative), (X=negative), (—=neutral), (✓=positive), and (✓✓=significantly positive)

For example, the UFC 227 bout between Johnson and Cejudo was classified into category 3. This placement aligns with the fight dynamics, as Cejudo was outstruck but achieved a significant grappling advantage. Examples such as Johnson-Cejudo underline the effectiveness of the clustering algorithm. By grouping each fight into a specific category, we can analyze a judge's style preferences.

Table 3 in the appendix displays the average agreement rates with fight outcomes and media scores for judges across six categories. We will spotlight Derek Cleary, who has judged 71 total fights in our dataset. Cleary agreed with category 3 (significant grappling advantage) outcomes 100% of the time but category 6 (significant striking advantage) outcomes only 50% of the time. On average, other judges agreed with category 3 fights 66% of the time and category 6 fights 68% of the time, so Cleary's scores indicate that he favors grappling more and striking less than most judges.

For example, in the Yair Rodriguez vs. Alex Caceres fight, categorized under category 6, Rodriguez dominated in striking and was favored by all media and 93.7% of fans. Yet, Cleary controversially scored the bout 49-46 for Caceres, seemingly preferring Caceres' grappling. This analysis isn't to single out Cleary but to illustrate that such biases and inconsistencies could be common among judges, as Table 3 demonstrates.

**Calls to Action**

We recommend several key reforms. First, implement training programs for judges that emphasize the diverse aspects of MMA and the recognition of biases. Second, employ more judges with expertise in MMA instead of judges only specialized in one discipline. Third, reevaluate the 10-9 must system to allow for more nuanced scoring that reflects the complexities unique to MMA bouts. Additionally, establish regular performance reviews and public transparency by releasing detailed judges' scorecards and decision rationales. Engaging all stakeholders in dialogue will ensure these reforms are effective and responsive to the sport's needs.

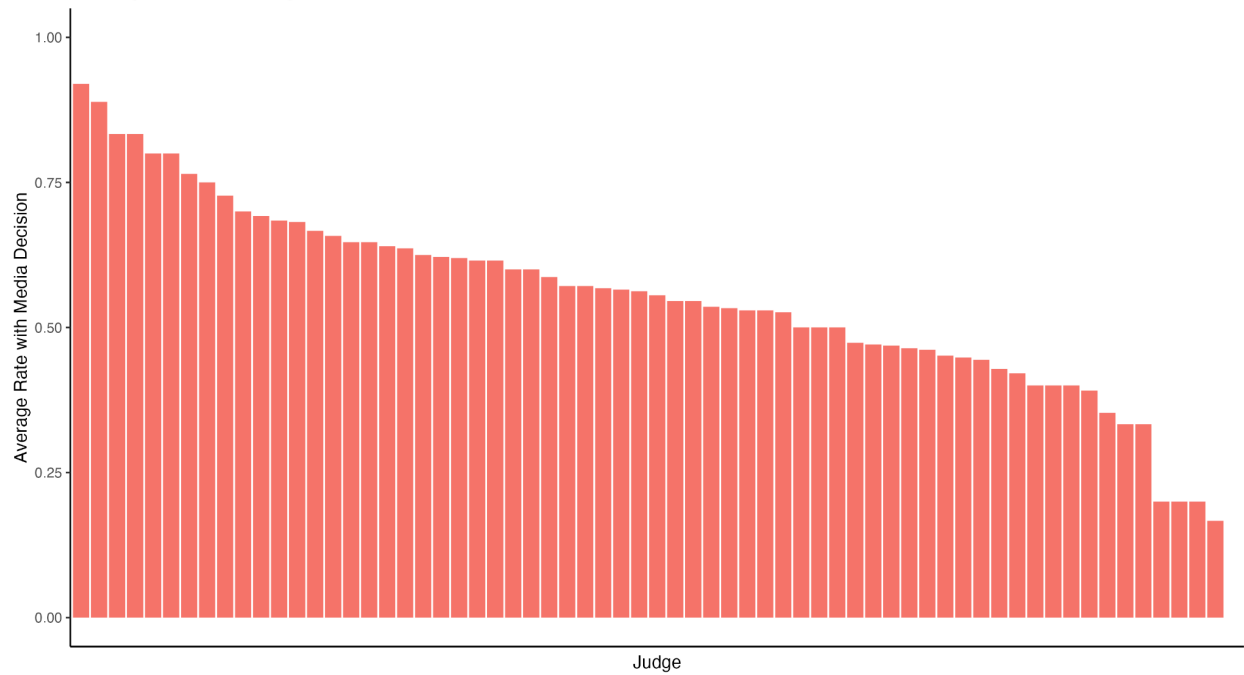
**Business Impact**

From a business perspective, fair judging is crucial. Transparent and reliable decision-making can significantly enhance MMA's marketability. When fans trust the outcomes, they are more likely to invest their time and money into the sport. This is especially pertinent given the persistent accusations of corruption in combat sports. Additionally, a reputation for fairness will attract more potential fighters, increasing the talent pool and elevating the sport's competitive standards.

**Conclusion**

The integrity of MMA judging isn't just a matter of fairness—it's a business imperative. Indeed, improving judging standards and methods will uphold the sport's credibility and enhance its appeal and profitability.

## Appendix



*Figure 2: Average agreement rate with media decision for all judges with minimum 5 fights judged.*

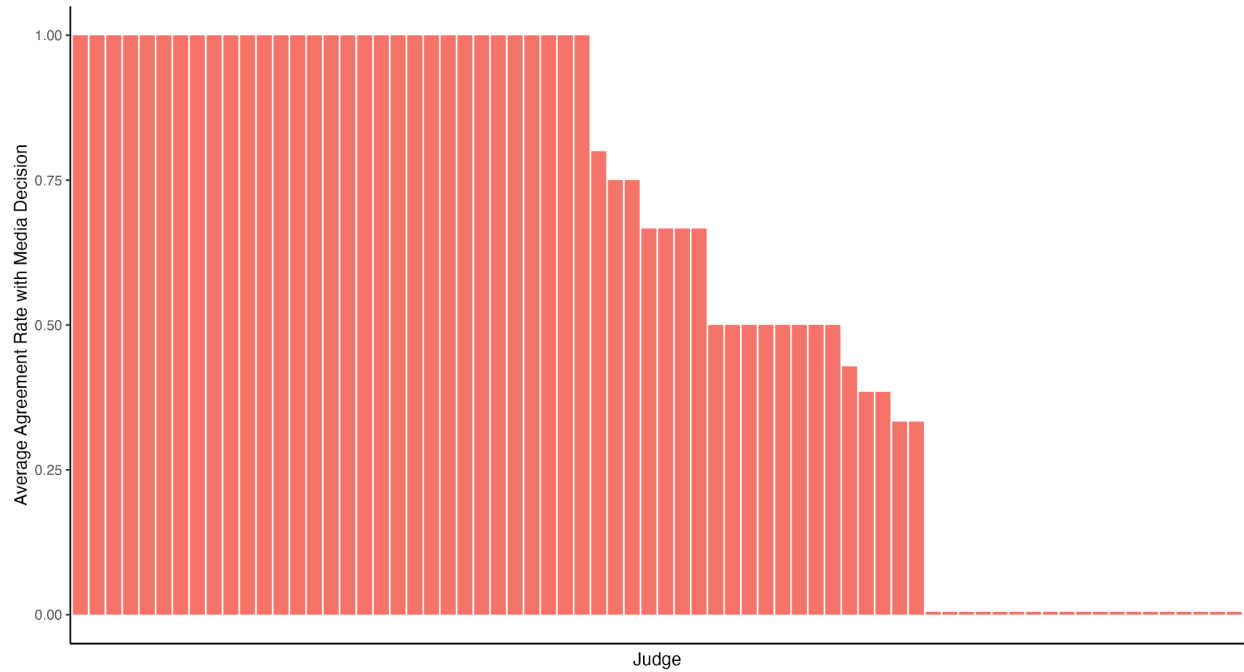
Table 2: Average differential statistics across each fight category.

Category	Knockdown Differential	Significant Strikes Landed Differential	Significant Strikes Attempted Differential	Total Strikes Landed Differential	Total Strikes Attempted Differential	Takedown Success Differential	Control Time Differential [secs]
1	-0.170	13.0	25.5	55.2	72.6	1.68	236
2	-0.103	11.0	17.6	12.4	18.8	-1.14	-77
3	-0.064	-6.2	-18.0	-2.9	-12.1	3.44	280
4	-0.088	-11.0	-32.3	-28.6	-51.3	0.37	-1
5	1.190	6.7	13.6	3.2	9.0	-1.60	-110
6	-0.043	33.0	88.1	44.9	101.0	-1.40	-66

Table 3: Judge agreement rates across fight categories in UFC split decisions; Top nine judges by total fights and media are represented.

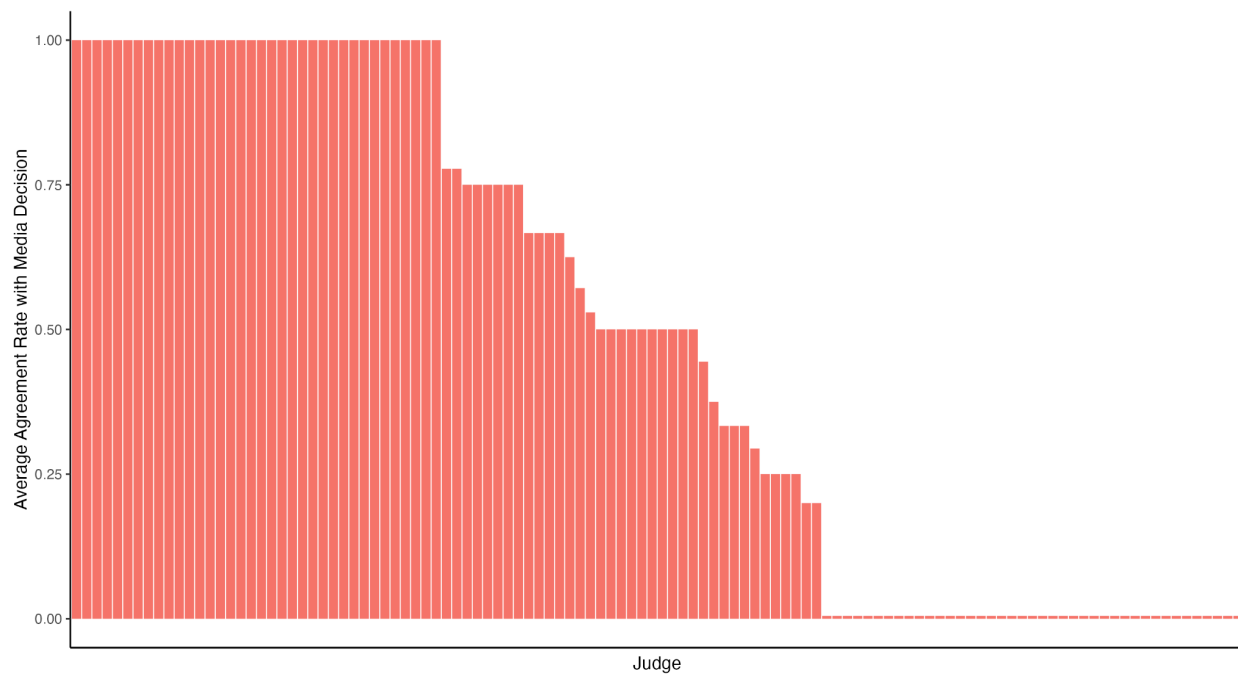
Judge \ Category	Count						Average Agreement Rate with Outcome						Average Agreement Rate with Media					
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
Media	53	184	110	114	42	47	0.75	0.58	0.79	0.51	0.79	0.74	1	1	1	1	1	1
Sal D'Amato	13	44	27	20	4	13	0.77	0.59	0.74	0.6	0.25	1	0.38	0.43	0.67	0.75	0.25	0.62
Chris Lee	13	28	18	16	6	11	0.46	0.68	0.78	0.56	0.5	0.55	0.38	0.68	0.67	0.5	0.5	0.45
Derek Cleary	9	23	8	17	4	10	0.89	0.61	1	0.71	1	0.5	0.67	0.65	0.75	0.53	0.75	0.5
Marcos Rosales	7	18	13	10	4	5	0.57	0.78	0.62	0.6	0.75	0.6	0.43	0.44	0.46	0.5	0.5	0.6
Tony Weeks	3	17	11	17	5	3	0.67	0.82	0.64	0.59	0.4	0.67	0.67	0.41	0.73	0.29	0.2	1
Glenn Trowbridge	1	18	8	9	4	6	1	0.56	0.5	0.78	0.75	0.83	1	0.5	0.63	0.78	0.25	0.67
Dave Hagen	3	15	9	9	0	2	1	0.8	0.56	0.67	-	1	0.67	0.73	0.67	0.44	-	1
Cecil Peoples	8	8	8	8	4	2	0.5	0.88	0.75	0.63	0.75	1	0.75	0.5	0.63	0.25	0.5	0.5
Junichiro Kamijo	5	14	6	8	0	4	1	0.86	0.83	0.63	-	0.5	0.8	0.64	0.5	0.75	-	0.25



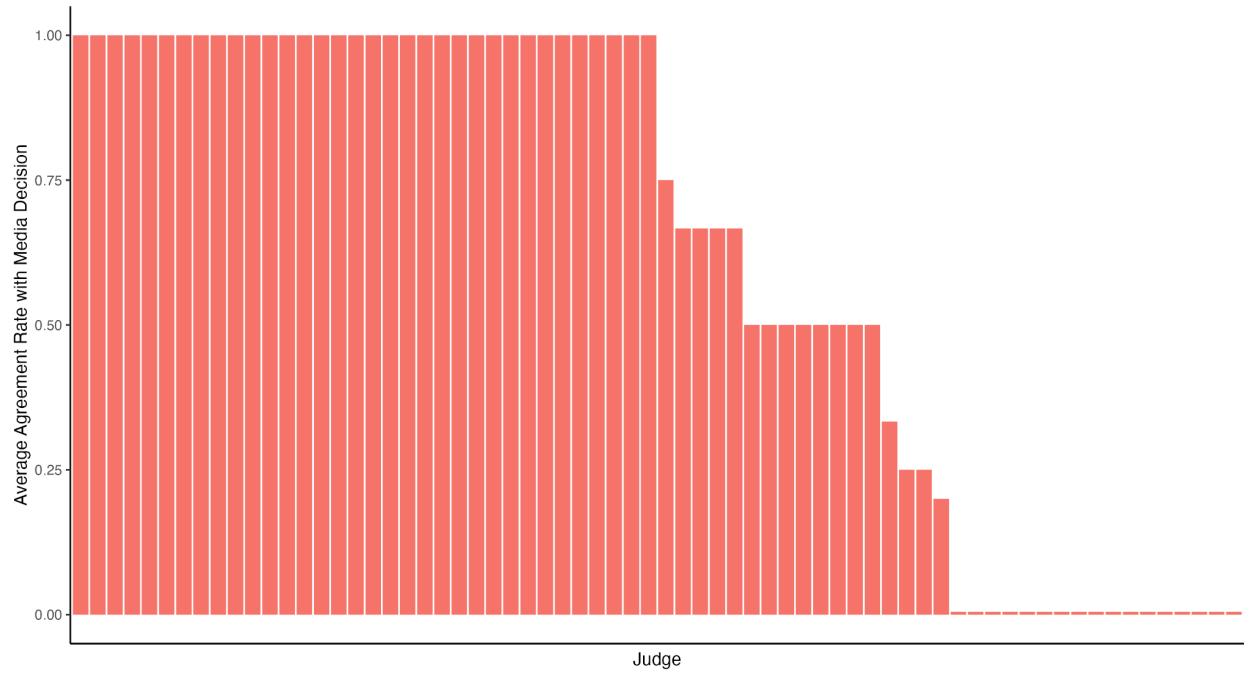




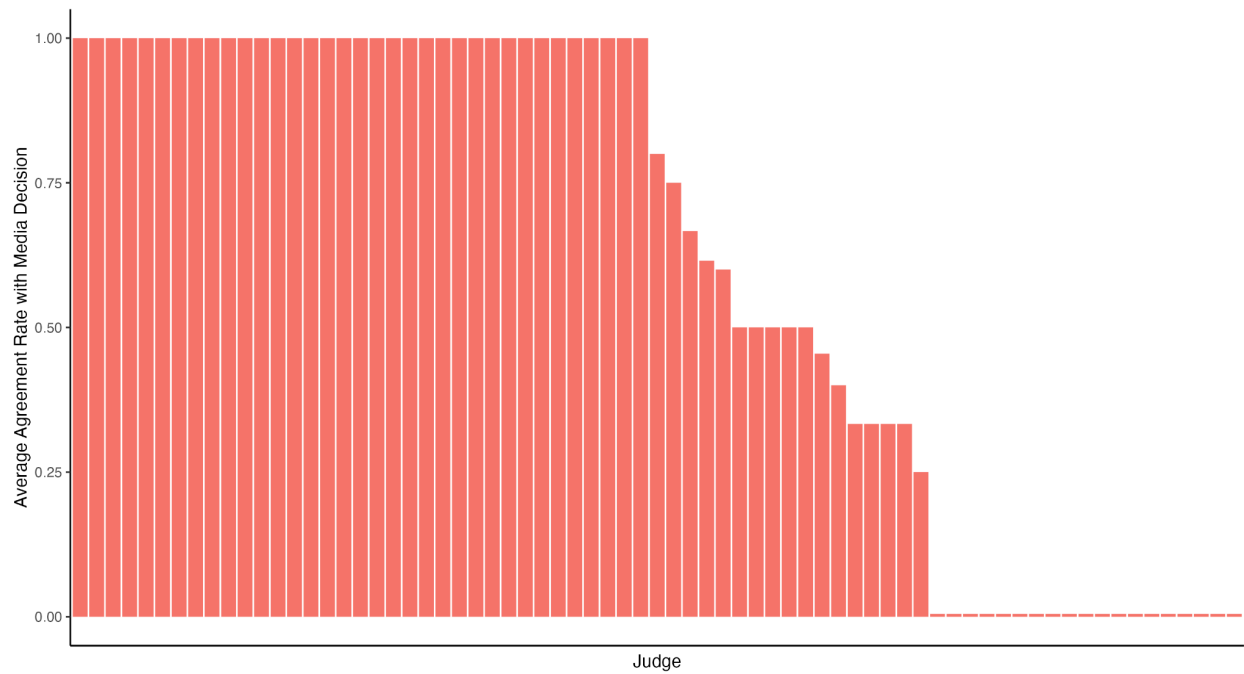
*Figure 5: Average agreement rate with media decision for judges with at least 1 fight within cluster 3.*



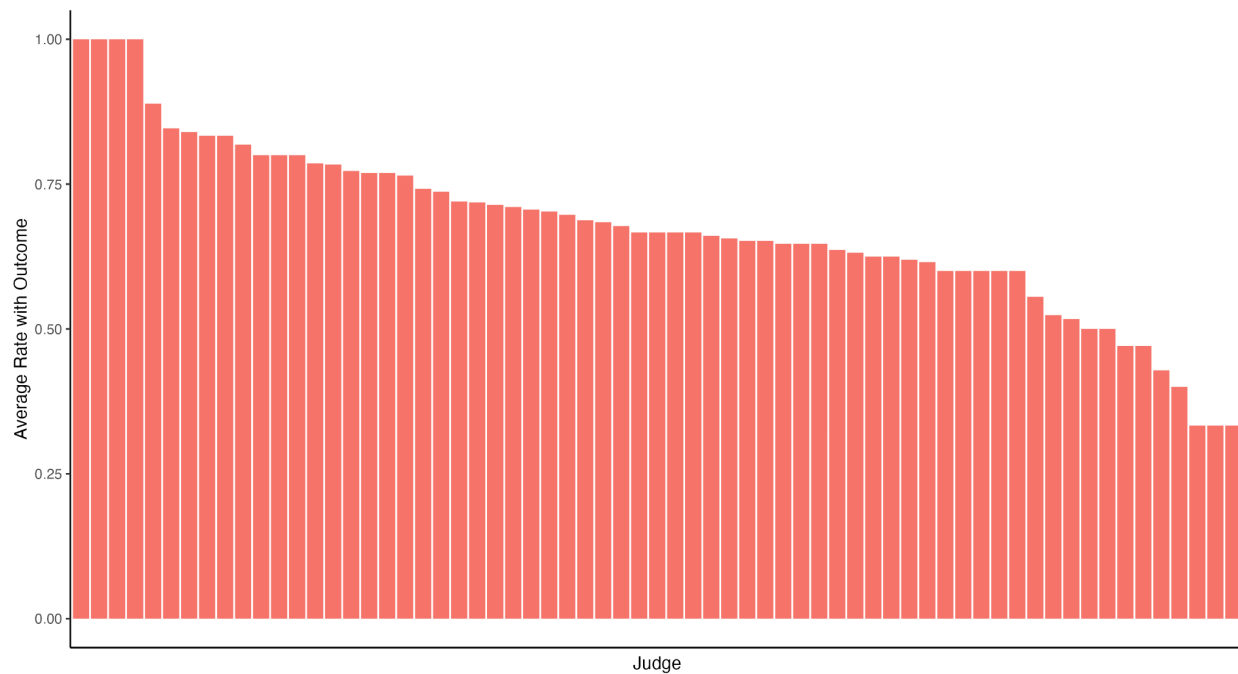
*Figure 6: Average agreement rate with media decision for judges with at least 1 fight within cluster 4.*



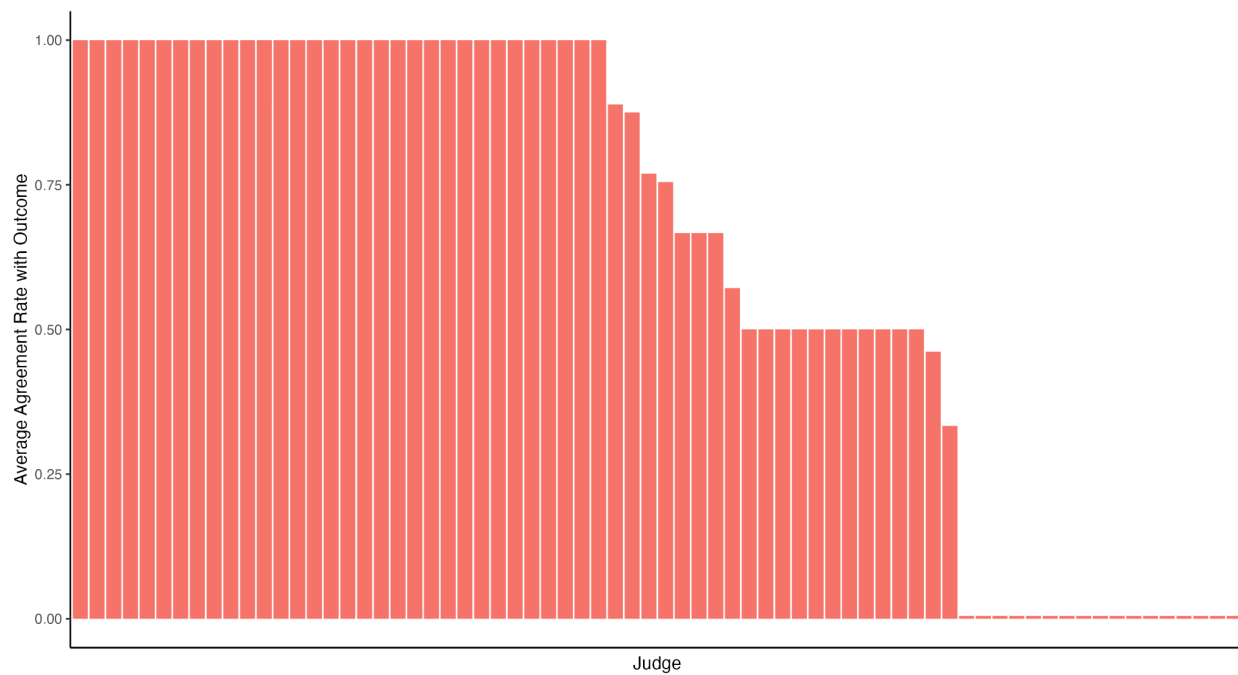
*Figure 7: Average agreement rate with media decision for judges with at least 1 fight within cluster 5.*



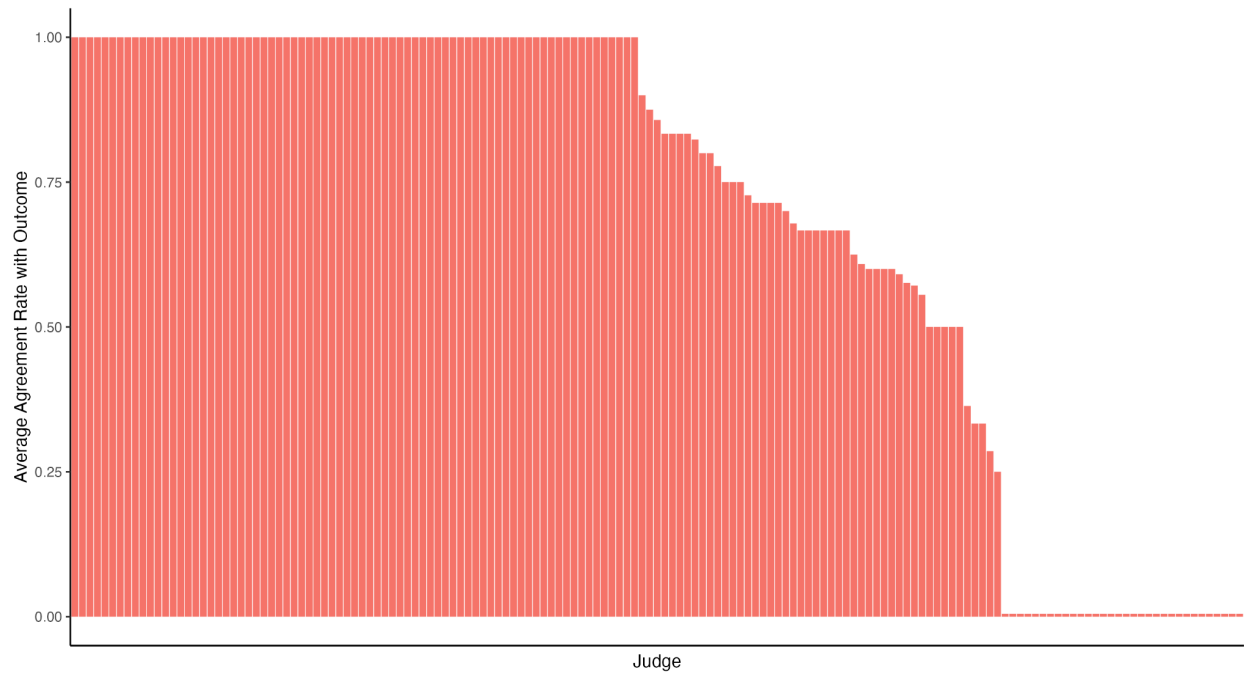
*Figure 8: Average agreement rate with media decision for judges with at least 1 fight within cluster 6.*



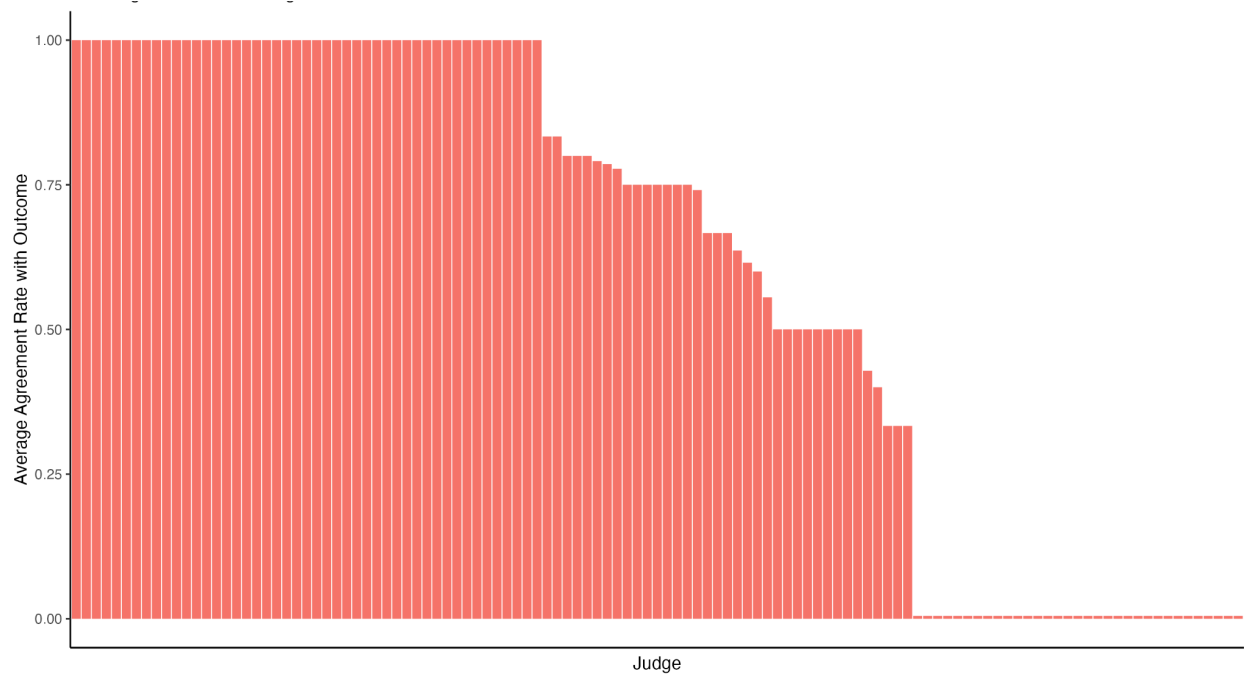
*Figure 9: Average agreement rate with fight outcome for judges with at least 5 fights judged.*



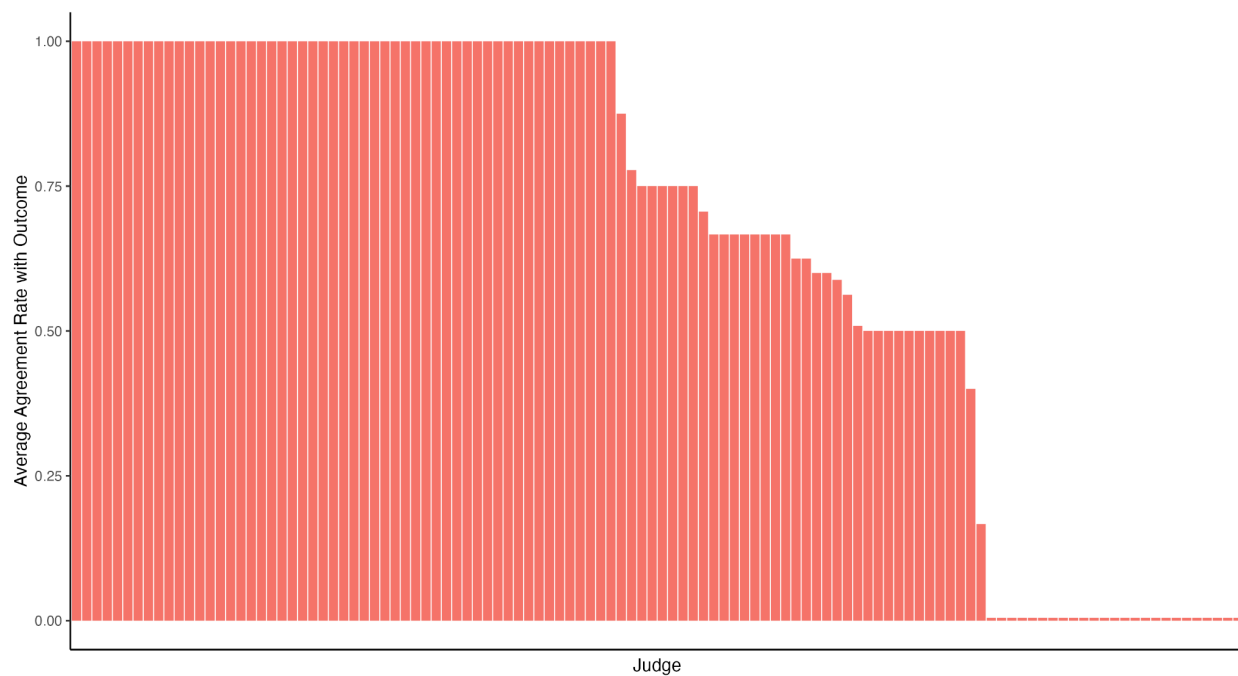
*Figure 10: Average agreement rate with fight outcome for judges with at least 1 fight judged within cluster 1.*

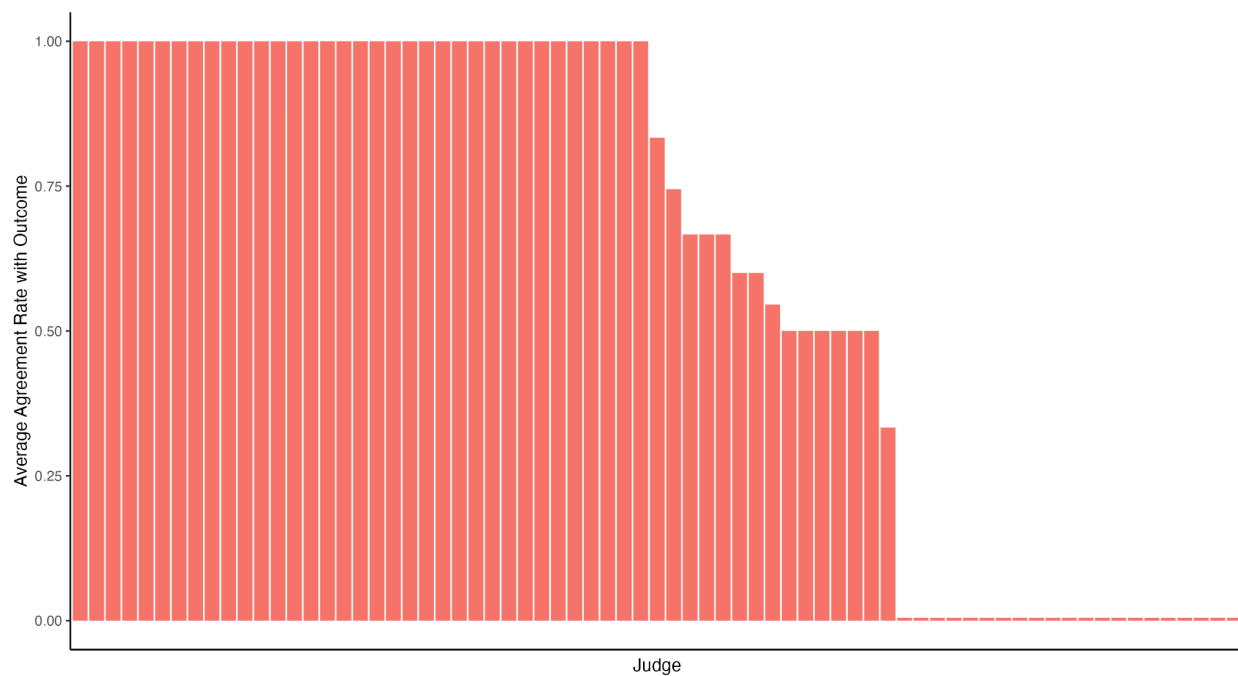


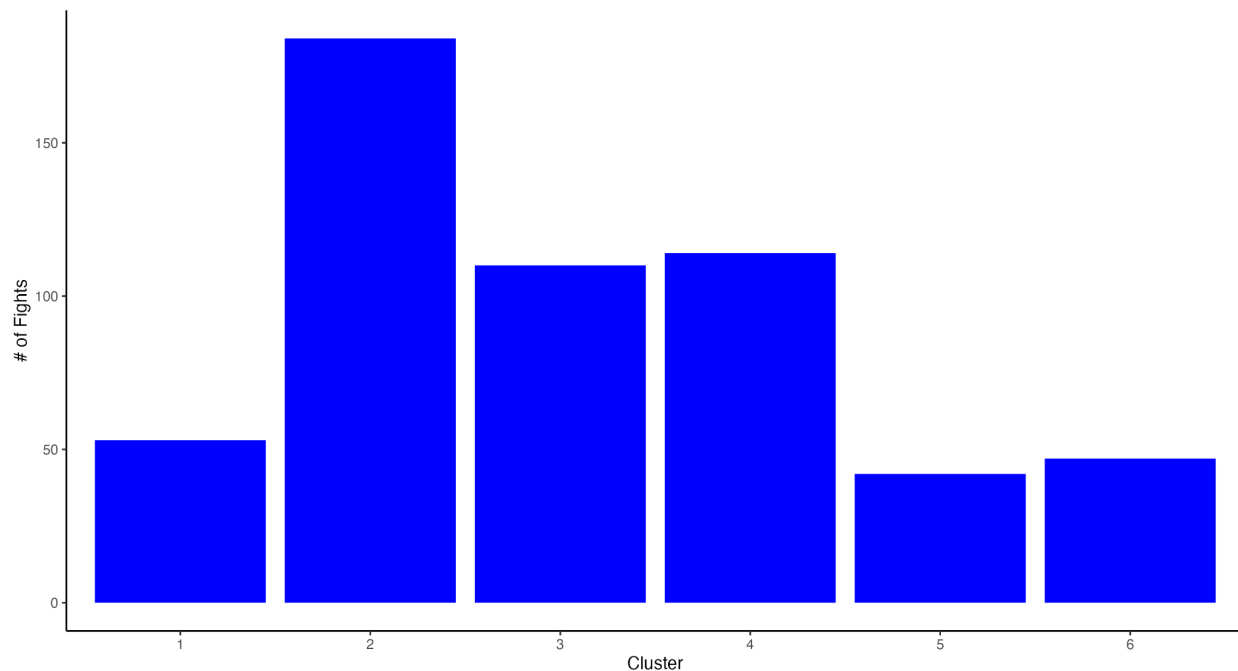
*Figure 11: Average agreement rate with fight outcome for judges with at least 1 fight judged within cluster 2.*



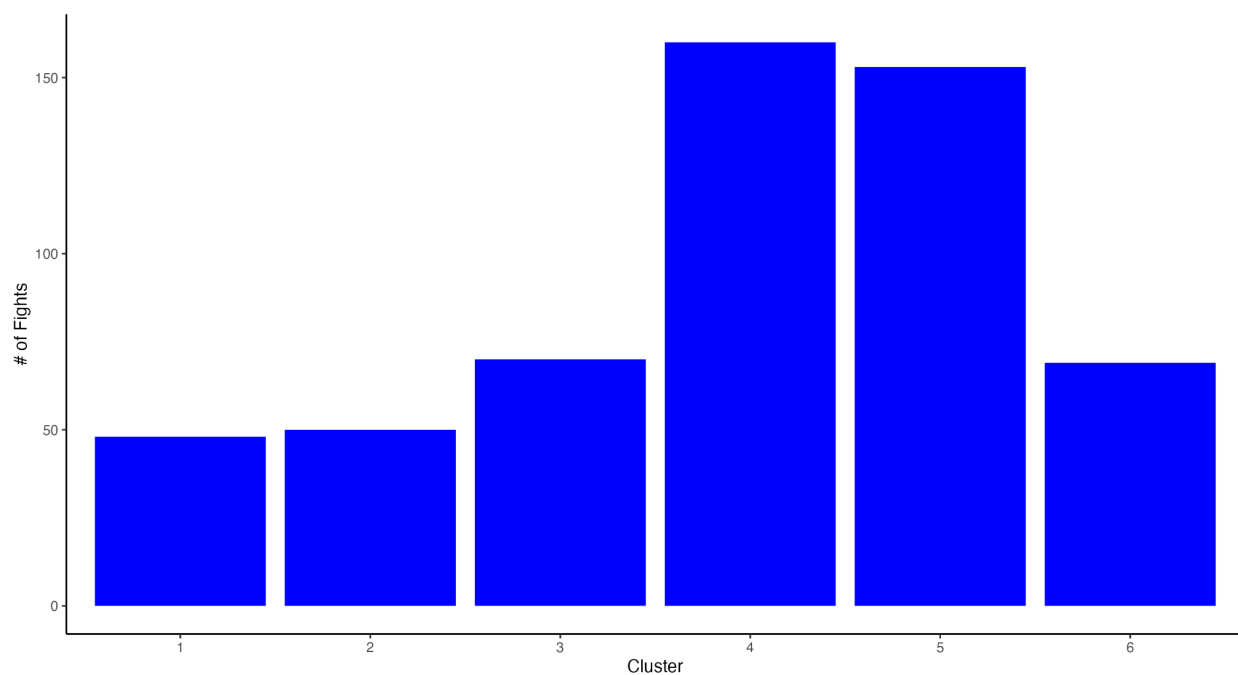
*Figure 12: Average agreement rate with fight outcome for judges with at least 1 fight judged within cluster 3.*







*Figure 17: Number of media votes for split-decision fights per cluster.*



*Figure 18: Number of split decision fights in each cluster based on official judge scores.*