Applying a simple correlational approach to a crowdstormed dataset

Authors: Christopher R. Madan[1]*

Affiliations

[1] Department of Psychology, University of Alberta

*Correspondence to: cmadan@ualberta.ca

Abstract

Crowdstorming a dataset allows for an unprecedented approach to research: to see how others would analyze the same dataset as you, but with their own analytical methods. While open data is similar in the regard, usually researchers use the data to answer new questions with the existing data, rather than to provide the opportunity of comparing data analysis methods. Here I applied a simple correlational approach to the crowdstormed dataset on potential effects of skin tone on referees issuing soccer plays a red card. While this correlational approach may appear overly simplistic, comparisons with more theoretically sophisticated approaches will determine if this approach is either elegant or naïve in its simplicity. Results from this simple correlational approach suggest no meaningful effect of skin tone on the issuance of red cards.

One Sentence Summary

Results from the simple correlational approach suggest no meaningful effect of skin tone on the issuance of red cards.

Results

Here I applied a simple correlational analysis to the crowdstorming project dataset. While many of the other teams have used analyses methods that are more theoretically sophisticated, one of the primary goals of the project is to compare a variety of statistical approaches and determine if the methods lead to different conclusions. While this correlational approach may appear overly simplistic, comparisons with more sophisticated approaches will determine if this approach is either elegant or naïve in its simplicity.

Initial Approach

Given that the primary goal of the analyses is to measure the influence of skin tone on referees issuing red cards, dyads was excluded where the skin tone was rated by either of the two raters. If the skin tone was rated by both raters, the mean rating was used. Since red cards can only be issued in proportion to the number of games played, this proportion measure was instead of the raw frequency counts. In all cases, a Spearman non-parametric correlation was used.

This simple correlation between rated skin tone and red card proportion yielded a small but significant relationship [Spearman's $\rho(124619)=.010$, $p>.001$, Cohen's $d=.020$]. Given that this correlation is quite small, but is nonetheless significant due to the high number of degrees of freedom, I calculated a number of additional correlations to provide interpretational context for this correlation.

First I tested if this correlation between skin tone and red card proportion would be significant based on referees from any particular country. For this follow-up analysis, I only selected referee countries that would have yielded a sufficient amount of data, which was chosen as more than a degrees of freedom of 20. Of the 161 referee countries present in the dataset, only 58 referee countries had a sufficient amount of data. However, after a Bonferonni-correction was applied, none of these individual referee country correlations were statistically significant.

As an additional follow-up, I returned to the initial dataset and calculated the correlation with the red card proportion with the player's height and weight. As I predict these variables to not have a meaningful relationship with the likelihood of being issued a red card, these correlations can serve as control correlations that have similarly high number of degrees of freedom. For these height and weight correlations, if the player's height or weight was missing, respectively, the dyad was excluded. Both the height and weight correlations were significant, however, they were both nearly as large as the correlation with the rated skin tone [height: $\rho(124573)=.0083$, $p=.0035$, $d=.017$; weight: $\rho(123866)=.0086$, $p=.0024$, $d=.017$]. Given the magnitude of this correlation, it is unlikely that the player-referee dyad correlations are meaningfully significant.

To further test for a relationship between skin tone and the proportion of games where red cards were issued, I calculated the correlation after collapsing across referee. This correlation was significant, but markedly larger than that observed when data was analyzed as player-referee dyads [$\rho(1583)=.051$, $p=.041$, $d=.10$]. This correlation suggests that there may be a small effect of skin tone on the issuance of red cards.

For the secondary research question, I calculated a correlation between skin tone and proportion of games where red cards were received, factoring in the influence of implicit and explicit racial attitudes. Using partial correlation and multiplicative approaches to test for an

influence of racial attitudes, none of the correlations yielded meaningful relationships [all ρ's < .02, d's ≈ .02].

Final Approach

The feedback provided did not notably influence my approach. However, one interesting comment was to incorporate the player's position into the analysis. To accommodate this, I re-did the analysis where the referee collapsed out of the analysis, which also was the analysis that produced the largest effect size [d=.10; 95% CI=(0.1026, 0.1031)]. From here, I divided players into their positions, categorizing the positions into "back", "midfielder", "forward/winger", and "goalkeeper". Players with positions not available were excluded from this follow-up analysis. Given that some players had to be excluded in this analysis, I first re-calculated the correlation between skin tone and red card proportion and found the effect to be slightly stronger than now that players with positions not available were excluded [ρ(1431)=.062, p=.020, d=.12]. None of the position-specific correlations were significant after correcting for multiple comparisons. Nonetheless, for all positions other than the goalkeeper, correlations were similar, but higher, likely due to the variability due to player position being removed, particularly the goalkeeper [ρ's between .086 and .117, d's between .17 and .23]. This correlation was negative for the goalkeeper.

Conclusion

Results from the simple correlational approach suggest no meaningful effect of skin tone on the issuance of red cards. However, it is possible that this approach is not sophisticated enough to take advantage of the complexity of the dataset.

Data and Output

Code available at: http://osf.io/rkehg