

## Research Question

Are soccer players with a dark skin tone more likely than those with a light skin tone to receive red cards from referees?

## Dataset

The team collected data and profile photos of soccer players (N = 2053) playing in the first male divisions of England, Germany, France, and Spain in the 2012-2013 season and referees (N = 3147) that these players played under in their professional career were obtained from a company for sports statistics. A dataset of player-referee dyads was created, including the number of matches players and referees encountered each other and the number of red cards given to a player by a particular referee throughout all matches the two encountered each other.

Player photos were available for 1586 out of 2053 players from the source. Two independent raters, blind to the research question, coded players' skin tone based on their profile photos. They categorized players on a 5-point scale ranging from "very light skin" to "very dark skin," with "neither dark nor light skin" as the center value.

Additionally, implicit bias scores for each referee country were calculated using a race implicit association test (IAT), where higher values indicated faster white | good, black | bad associations. Explicit bias scores for each referee country were determined using a racial thermometer task, with higher values indicating greater feelings of warmth toward whites versus blacks. Both measures were derived from data aggregated from online users in referee countries who took these tests on Project Implicit (<http://projectimplicit.net>).

Overall, the dataset comprises 146,028 dyads of players and referees.

## Column Description

	field_name	description	data_type
0	playerShort	Short name of the football player.	string
1	player	Full name of the football player.	string

2	club	Name of the football club the player belongs to.	string
3	leagueCountry	Name of the country where the league is played.	string
4	birthday	Birthday of the player in dd.mm.yyyy format.	date
5	height	Height of the player in centimeters.	number
6	weight	Weight of the player in kilograms.	number
7	position	Position of the player in the field.	string
8	games	Number of games in the player-referee dyad.	number
9	victories	Number of victories in the player-referee dyad.	number
10	ties	Number of ties in the player-referee dyad.	number
11	defeats	Number of losses in the player-referee dyad.	number
12	goals	Number of goals scored by a player in the player-referee dyad.	number
13	yellowCards	Number of yellow cards player received from referee.	number
14	yellowReds	Number of yellow-red cards player received from referee.	number
15	redCards	Number of red cards player received from referee.	number
16	photoID	ID of player photo (if available).	string

17	rater1	Skin rating of photo by rater 1 (5-point scale ranging from 'very light skin' to 'very dark skin') normalized to 1.	number
18	rater2	Skin rating of photo by rater 2 (5-point scale ranging from 'very light skin' to 'very dark skin') normalized to 1.	number
19	refNum	Unique referee ID number (referee name removed for anonymizing purposes).	number
20	refCountry	Unique referee country ID number (country name removed for anonymizing purposes).	number
21	meanIAT	Mean implicit bias score (using the race IAT) for referee country, higher values correspond to faster white   good, black   bad associations.	number
22	nIAT	Sample size for race IAT in that particular country.	number
23	seIAT	Standard error for mean estimate of race IAT.	number
24	meanExp	Mean explicit bias score (using a racial thermometer task) for referee country, higher values correspond to greater feelings of warmth toward whites versus blacks.	number
25	nExp	Sample size for explicit bias in that particular country.	number
26	seExp	Standard error for mean estimate of explicit bias measure.	number