

Overview

In this study, we conducted a comparative analysis of players' performance between international and club soccer games to assess whether their respective performances exhibit similarities. The motivation for this investigation stems from the common criticism players face from the public for not meeting expectations, especially in international competitions.

A player's performance is evaluated primarily by how many goals they scored. Hence, we have chosen the same approach to measure their performance. Their performance comparison between both (international and club) games by looking only at their goals will not be an ideal approach since their appearances between those games might be different. Hence, for our analysis, we have calculated their goals per appearance (GPA) in both games and compared them with each other.

Dataset

The dataset for this study comprises two CSV files. The first file encompasses club data, detailing player performances in club games, while the second file contains data from the same years, capturing player performances in international games. These datasets were sourced from Wikipedia's official website, www.wikipedia.org.

The CSV file dedicated to international games features twelve players listed in rows, with 8 columns each presenting data on their country, appearances, and goals over three years (2006, 2010, and 2014). Similarly, the second CSV file, focusing on club data, also has the same twelve players in rows, with 8 columns reflecting their respective clubs, appearances, and goals during the same three-year span.

Analysis

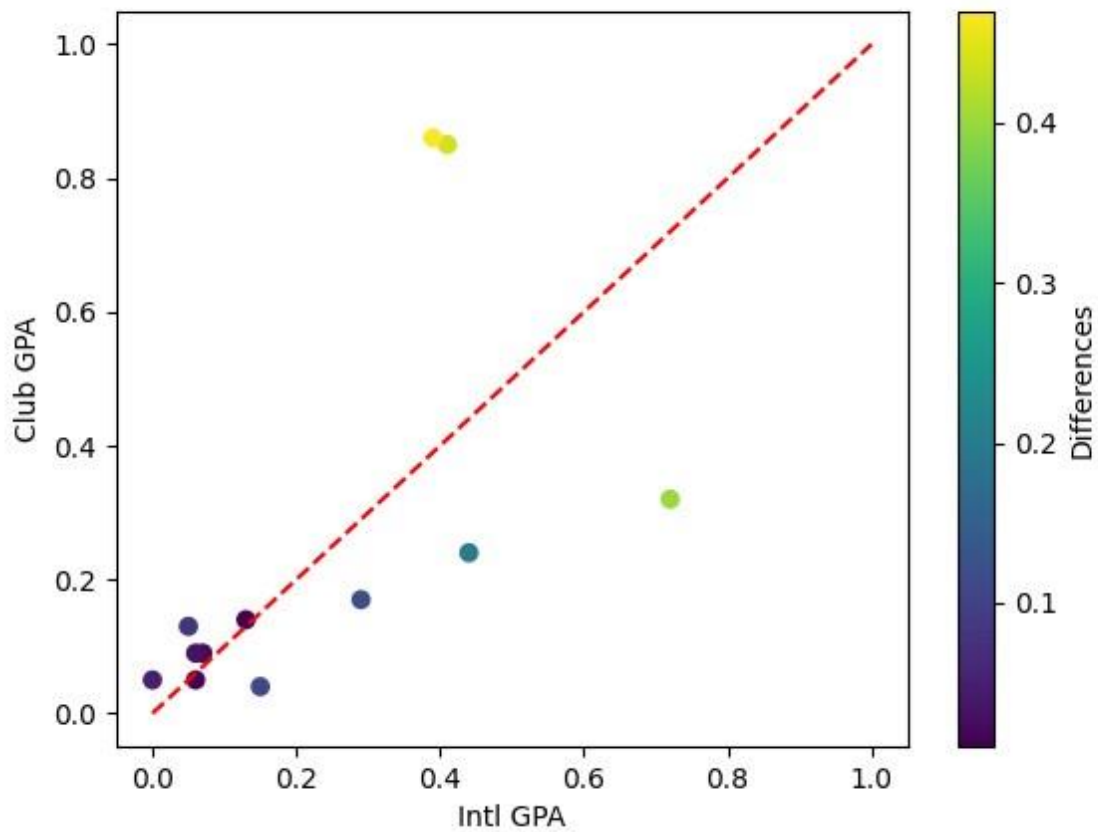
Goals per appearance (GPA) are calculated by dividing total goals scored by the total number of games they appeared in using the below formula:

$$\text{GPA} = \frac{\text{Total Goals}}{\text{Total Appearances}}$$

The results below were obtained from our analysis:

| Players | GPA _{Club} | GPA _{Intl} | Absolute Difference |
|------------------------|---------------------|---------------------|---------------------|
| Cristiano Ronaldo | 0.85 | 0.41 | 0.44 |
| Lionel Messi | 0.86 | 0.39 | 0.47 |
| Sergio Ramos | 0.13 | 0.05 | 0.08 |
| Rafael Márquez | 0.04 | 0.15 | 0.11 |
| Xavi | 0.09 | 0.07 | 0.02 |
| Andrés Iniesta | 0.14 | 0.13 | 0.01 |
| Andrés Guardado | 0.09 | 0.06 | 0.03 |
| Wesley Sneijder | 0.33 | 0.23 | 0.05 |
| Tim Cahill | 0.24 | 0.44 | 0.20 |
| Miroslav Klose | 0.32 | 0.72 | 0.40 |
| Philipp Lahm | 0.05 | 0.06 | 0.01 |
| Bastian Schweinsteiger | 0.17 | 0.29 | 0.12 |

Examining the above results, we can conclude that only three players out of twelve (Cristiano Ronaldo, Lionel Messi, and Miroslav Klose) exhibit significantly different performances. Other players have less significant differences.



The accompanying scatter plot, generated using Python, displays goals per appearance in international games on the x-axis and goals per appearance in club games on the y-axis, with dot colors denoting their differences. It highlights the same three notable outliers.