

Probing the quality of football leagues through player's foot laterality: A data analytics approach

by Dicky Silitonga

Submission date: 17-Mar-2023 01:05PM (UTC+0700)

Submission ID: 2039168333

File name: 22369-Article_Text-75242-1-2-20230301.docx (695.74K)

Word count: 3840

Character count: 20793



6
7

8 Probing the quality of football leagues through player's foot laterality: 9 A data analytics approach

10
11 Received: date month xxxx Accepted: date month xxxx Published: date month xxxx

13 Abstract

14 The ability to use both sides of feet in association football, or so-called soccer in some countries, is highly
15 valued, and therefore professional athletes rigorously train their non-dominant hand to perform effectively on
16 the pitch. However, a subject that relates that players' laterality aspect to the overall quality of the league of a
17 given country has not yet been widely explored. Despite being seemingly trivial, scrutiny on that subject reveals
18 the details and exceptions pertaining to the correlation between those variables. This paper discusses the subject
19 through a data analytics approach, leveraging the abundance of publicly available football datasets. In the
20 present study, the foot laterality measure is represented by the percentage of left-footed scored goals and
21 league's overall quality by the market values of players. Results show that in general, leagues with higher market
22 values have a higher rate of left-footed goals. Interestingly, this tendency does not always hold true, especially
23 in leagues of countries where their high-performing native talents are pursuing careers in foreign competitions.
24 That exception leads to a conclusion that the performance of national teams in countries where many of their
25 native players are pursuing careers abroad is not necessarily linear with the quality indicator of their domestic
26 league. In other words, it is not unusual that a country's domestic league seems poor or mediocre, yet the
27 national team excels in the international stage since the premium players are competing in the top leagues
28 abroad and only joining domestic clubs during their youth and near-retirement periods. The findings in this
29 work agree with those in past research conducted with other methods, hence confirming the validity of data
30 analytics based on open-source football data. This research is also part of endeavors to amplify the potential
31 of data-driven analysis in sport science, promoting the multidisciplinary nature of the field itself and motivating
32 the practitioners to use data analytics to fine-tune their physical training methods and sport psychology
33 approaches.

34
35 **Keywords:** Left-footedness; Foot laterality; Football analytics; Data-driven approach; Crowd-sourced sport
36 data

39 1. Introduction

40 Studies revealed that a vast majority of the world population is right-handed, comprising a significantly
41 unbalanced ratio between right- and left-handed humans of nearly 9:1 (Corballis, 2003; Peters et al.,
42 2006). Furthermore, this laterality, or preference for one side of the body, works in the same way in
43 most of the cases for handedness and footedness (Tran & Voracek, 2016). Research shows that the
44 right-footers are also heavily dominant in human population (Porac & Coren, 1981) although at a
45 lower percentage than the right-handers.

46 In normal life for regular activities, this side preference does not really matter, or one can adapt for
47 certain purposes without being severely handicapped. Indeed, in societies, right-handedness
48 sometimes becomes culturally induced (Alexandre Jehan Marcori, Victor Hugo Alves Okazaki, 2020)

49 where failing to act with that side of hand may be considered a non-compliance to traditional or
50 common values. Thus, in general, one can still manage to live and thrive without any restrictions
51 regardless of his or her lateral preference, however unbalanced it is.

52 Nevertheless, in some sports, well-balanced laterality is valuable and highly sought after (Stöckel &
53 Carey, 2016). In football, the ability to utilize both feet is beneficial for players to excel in their game.
54 Having both feet work at nearly the same level of control and strength allows a player to have
55 flexibility to operate across the width of the field, agility to anticipate the incoming ball and freedom
56 to take a shooting decision as the momentum arises without the need to first position the ball towards
57 the strong foot. As a remark, the term football that will be used for the rest of this paper refers to the
58 association football or soccer, not to be confused with American football.

59 To improve lateral balance, one should train so as to develop the strength and control of his or her
60 genetically non-preferred side (Akpınar, 2022). That being said, high-performing players are those
61 who, besides other skills and instincts, strive and somehow achieve a certain level of lateral balance in
62 terms of their feet. High-performing footballers commonly play in the top-tier leagues, where players
63 with genetic physical advantages, hard-working ethics, and focused mentality are sharpened by the
64 skill-intensive, well-managed competition. With that concept in mind, then one may draw a hypothesis
65 that the overall quality of a group of players in a competition, represented by a league, can be reflected
66 by the footedness contribution of players in decisive actions or events. Specifically, this paper
67 discusses the footedness ratio of goals, i.e., whether they are scored by left or right foot. The study
68 takes into account the data from multiple leagues with the objective of observing the characteristics
69 of those leagues or the corresponding countries' football atmosphere.

70 This work demonstrates a proof of concept of employing a subset of overwhelmingly available
71 football data open to the public through the world wide web to deduce valuable information with the
72 help of computational tools to visualize and uncover the obscured correlations between variables.
73 Further, this paper is aimed towards a diverse spectrum of audience: for professionals such as players,
74 coaches and scouts to gain insights not only on the substantial subject covered in this paper, but also
75 on the methods that may inspire them to apply a similar approach to scrutinize other topics of interest;
76 for football enthusiasts to expand their horizon of knowledge as they follow their favorite leagues,
77 teams, and players; and for data analytics and machine learning experts to even motivate them in
78 finding easter eggs in football statistics that advance the roles of big data and artificial intelligence in
79 sports.

80 **2. Method**

81 Today's football is a grand industry sector where business and financial consciousness are part of the
82 game itself (Alaminos et al., 2020). Therefore, it is reasonable to consider the monetary factor, in this
83 case the players' prices which in turn accumulate into team valuation, as parameters that determine
84 the quality or competitiveness. In this research, the foot laterality figures are correlated with those
85 valuations to decipher the meaningful information. Note that the team value here refers to the players'
86 total market values instead of the club's valuation as a company or business entity.

87 The data is obtained from transfermarkt.com, an online source that provides extensive data analytics
88 related to football from around the world. It is a part of the German-based publishing group, Axel
89 Springer SE, that owns several major media with extensive circulation in Europe. Practically,
90 transfermarkt.com, and its other country specific domains, are accessed by worldwide users in real-
91 life for various purposes from entertainment and gaming, up to financially impacted activities such as
92 player scouting and, undoubtedly, betting. Considering the scale of users and reputation of the
93 platform, the author believes that the data available on transfermarkt.com is valid to a greater extent.

94 In fact, a previous study endorsed the use of information from crowdsourcing platforms like
 95 transfermarkt.com for research purposes (Prockl & Frick, 2018).

96 For the investigation, top-tier leagues are picked from leading countries representing all regional
 97 confederations, excluding the Oceania Football Confederation (OFC) due to its small size and least
 98 coverage. In addition, to provide richer knowledge for Indonesian readers as the main audience of
 99 this paper, the author includes Indonesia and Thailand, although those countries occupy the lower
 100 part in FIFA country ranking. The author is aware that it would be more interesting and
 101 comprehensive if all countries in ASEAN are included, so as to obtain comparative insights of regional
 102 football situation. However, only those two mentioned countries have good records of league goals
 103 dataset in transfermarkt.com. Table 1 lists the leagues selected for analysis in this paper.

104 **Table 1. Leagues included in the present study**

Country	League	No. of Teams	League market value (USD mio.)	Average team market value (USD mio.)
England	Premier League	20	10320	516
Spain	La Liga	20	4830	241.5
Germany	Bundesliga	18	4290	238.3
Italy	Serie A	20	4520	226
France	Ligue 1	20	3350	167.5
USA	Major League Soccer	29	1130	38.8
Mexico	Liga MX Clausura	18	812	45.1
Brazil	Camp. Bras. Série A	20	1330	66.26
Argentina	Liga Profesional de Fútbol	28	783.7	27.9
Morocco	Botola Pro Inwi	16	126.4	7.9
Japan	J1 League	18	272.7	15.2
South Korea	K League 1	12	159.9	13.3
Australia	A-League Men	12	98.7	8.2
Indonesia	Liga 1 Indonesia	18	71.7	3.9
Thailand	Thai League	16	73.6	4.6

105 Datasets extracted for footedness analysis are statistics of goals from 2018 to 2022 that contain the
 106 number of goals scored with the right and left foot for all competing teams. The underlying reasons
 107 for the selected time frame are that within a relatively short period, the prominent players and overall
 108 style of play across leagues have not undergone a major change. Hence, the analysis will be relevant
 109 to global modern, state-of-the-art, football games. That period also considers the chance of blank
 110 dataset due to halted competitions during COVID-19 outbreak.

111 Although being dependent of many random variables, the total number of scored goals in a league
 112 has a close relationship with the number of teams competing in that league. Since the sample leagues
 113 have different numbers of teams, as listed in Table 1, thus the figure taken for comparison is the
 114 percentage of goals scored with either side of feet. That way, the nominal number of goals which may
 115 vary a lot from top leagues to minor leagues is normalized into percentage level, allowing facile
 116 comparison among them. Furthermore, since left-footedness is the one being rare and thus becoming
 117 the main interest, the laterality parameter taken here is in terms of left-footed goal percentage.
 118 Choosing the right-footed percentage would work as well for the analysis yet showing the left one is
 119 more intuitive for the reader to focus on the unique skill pursued by high performing players. The
 120 left-footed goal percentage is calculated through the following straightforward formula:

121
$$\text{Left-foot goals \%} = \frac{\text{Left-foot goals}}{\text{Left-foot goals} + \text{Right-foot goals}} \times 100 \% \quad (1)$$

122 The percentages for all the studied leagues are plotted against the corresponding average team market
 123 values. A league's market value, in this case, is the sum of the prices of players in the league in question,
 124 not the value of the team as a business entity or company that may include other non-player assets.
 125 The average team market values are taken here instead of the total league market values of the league
 126 since, again, the number of teams playing in a league is different, as shown in Table 1.

127 To add a broader picture on the comparison of foot utilization of players, an evaluation on the
 128 distribution of passes attempted with either foot is included in our analysis. The data is limited to the
 129 top 5 European leagues because only those leagues have extensive open source big data availability.
 130 The data, however, takes into account all of the passes attempted by players, not only the decisive
 131 ones. Therefore, it is more appropriate that this indicator be evaluated in terms of the frequency to
 132 provide evidence that normally players do prefer to use their right foot, even in those top-class
 133 competition.

134 The percentage of passing attempts by each foot is obtained from FBref.com for 2021-2022 [Big 5](#)
 135 [European Leagues Stats](#). FBref.com was created by Sports Reference (sports-reference.com), a team
 136 behind popular sports statistics websites like Baseball-Reference.com and Basketball-Reference.com,
 137 launched in June 2018 with league coverage for six nations: England, France, Spain, Italy, Germany,
 138 and the USA. The pass attempts data is extracted from detailed player statistics throughout one full
 139 season, with the percentage of each foot is given by the following formulas:

140
$$\text{Left-foot Pass \%} = (\text{Left-foot Pass} / \text{Total Pass}) \times 100 \% \quad (2)$$

141
$$\text{Right-foot Pass \%} = (\text{Right-foot Pass} / \text{Total Pass}) \times 100 \% \quad (3)$$

142 The percentages for all players are plotted with violin chart method to show the distribution of the
 143 number of players attempting passes at a particular percentage level. This plot represents the spread
 144 of footedness in terms of passing attempts for top 5 European leagues in one full season 2021-2022.

145 The data is drawn from multiple pages in transfermarkt.com and FBref (*2021-2022 Big 5 European*
 146 *Leagues Stats, 2022*) by either manual copying or scraping. Then, the raw data was processed using
 147 Python and R programming language on a shareable notebook on Google Colab
 148 (colab.research.google.com). Visualization of those processed data helps the reader to understand and
 149 perceive the characteristics of the observed variables and their relation to each other.

150 **3. Results**

151 The recapitulation of data drawn from transfermarkt.com pertaining to the relation of left-footedness
 152 and market values of leagues is tabulated in Table 2. For ease of interpretation, however, it is visualized
 153 in Figure 1 as a scatter plot. Presenting the table as well as the scatter plot in this section does not pose
 154 any redundancy as both of them will be useful for the readers to digest the data analytics logic
 155 discussed in the subsequent section.

156

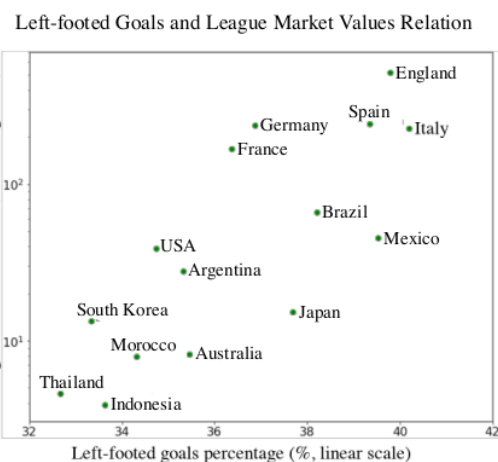
157

Table 2. Left-footed goals percentages leagues

Country	Average team market value (USD mio.)	Left-footed goals (%)
England	516	39.8
Spain	241.5	39.4
Germany	238.3	36.9

Italy	226	40.2
France	167.5	36.4
USA	38.8	34.7
Mexico	45.1	39.5
Brazil	66.26	38.2
Argentina	27.9	35.3
Morocco	7.9	34.3
Japan	15.2	37.7
South Korea	13.3	33.3
Australia	8.2	35.4
Indonesia	3.9	33.6
Thailand	4.6	32.7

158 In Figure 1, the horizontal and vertical axes represent the level of left-footedness and the indicative
 159 value of league, respectively. Note that the vertical axis is in logarithmic scale, while the horizontal
 160 one is linear. This logarithmic scaling style in the vertical axis is chosen considering the significant
 161 disparity between the lowest and highest valued league. Meanwhile keeping it linear would describe
 162 the extent of the disparity, it is not favorable for visualization as some of the low-valued points would
 163 be obscured due to their closeness to each other. For the interpretation of the chart in Figure 1, one
 164 may start with the understanding that the more to the right a data point is, the better-skilled player the
 165 corresponding league is, at least from the perspective of foot dominance balance. Likewise, the higher
 166 the point means the more valuable players are in a given league. Both variables then converge into the
 167 overall insights of player skill level, since the highly skilled players would expectedly have a well-
 168 balanced footedness as well as high market price.



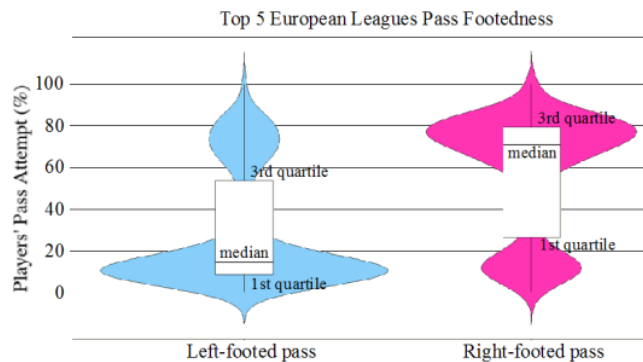
169

170 **Figure 1. Scatter plot correlating left-foot goal percentage and average market value of**
 171 **leagues**

172

173 The distribution of passing attempts percentage with left- and right-foot, summarized from
 174 FBref.com, is visualized in Figure 2. The violin plots show that right-footed pass is at significantly
 175 higher rate with most of players have approximately 75% pass attempts with their right foot, shown
 176 in Figure 2. The thick part indicates that the number of players with that rate of passing is heavily
 177 concentrated around that value of percentage. For example, the violin chart for Left-footed pass
 178 implies that most of the players (median value) passes with their left foot at around 15% rate and the

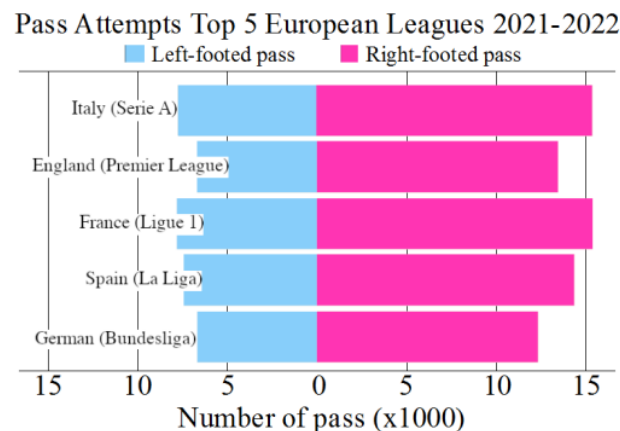
179 thickest part around that median means that the majority of all sampling point is indeed at around that
 180 median value. To evaluate the nominal number of pass attempts, the data is visualized in Figure 3.



181

182 **Figure 2. Violin plots describing the distribution of passing attempts footedness percentage**
 183 **of players in top 5 European leagues.**

184 The difference number passing attempted around top 5 European league based on strategy or
 185 formation approach on each league, which Serie A and Ligue 1 are Possession tactical than Bundesliga
 186 and English Premier league which most of teams are using Direct Passing Tactical. Left-footed passing
 187 is fewer than right-footed passing in top 5 European League although proportion at Bundesliga is
 188 quite different which mean Player on Bundesliga has Left and Right foot preferences on passing
 189 attempted.



190

191 **Figure 3. Comparison of the number of pass attempts in top 5 European leagues (2021-2022)**

192 **4. Discussions**

193 The tendency shown in Figure 1 is self-explanatory in general, in the sense that the giant leagues are
 194 occupied with highly valued players, hence superior skills that include the ability to control and exert
 195 power with either foot (Bryson et al., 2013). World's highest-tier leagues, which are themselves the
 196 top 5 leagues in Europe, are in the upper-right regime of the chart. This position implies that they are

197 a group where players score at a high rate with the normally-less-dominant foot, i.e., the left foot.
 198 Recall that the vertical (market value) axis is logarithmic, so the top-5-leagues group have a substantial
 199 gap in terms of valuation with the lower groups, see Table 2 for nominal values comparison.

200 A closer observation of the chart shows that there are countries' leagues with left-footed goals
 201 percentage comparable to the top 5 leagues but at considerably low indicative market values, namely
 202 Brazil, Mexico, and Japan. Those countries are considered talent feeders to European leagues, notably
 203 Brazil (de Vasconcellos Ribeiro & Dimeo, 2009), as a long-standing fountain of exceptionally talented
 204 players. In the past decades, Japanese football has also substantially risen, with players aspiring to play
 205 in Europe (Yoshio & Horne, 2004). The serious development of football in Japan in the past 50 years
 206 (Taylor, 2006) is proven to be fruitful, as observed today, where many Japanese players have made
 207 their way to Europe (Orlowitz, 2022). As for Mexicans, their presence in European leagues is
 208 predominantly in Spain, where they speak their native tongue. However, when looking at the players
 209 migration distribution, a huge number of Mexican players are pursuing careers in the USA (Armstrong
 210 & Rosbrook-Thompson, 2010; Jewell & Molina, 2005; Martinez, 2008). It can be linked to the
 211 socioeconomic and geographic proximity factors that are beyond the scope of this study.

212 Interestingly, Argentina, the 2022 World Champion does not show any outstanding tendency. An
 213 argument can be constructed by keeping in mind that the chart does not describe the national team,
 214 but the local league. That being said, it is understandable that a league is not necessarily reflecting the
 215 skills of players holding citizenship of the corresponding country. There main cause is that the
 216 country's best players have left home for Europe. Speaking of Argentina, the FIFA World Cup 2022
 217 champion squad consists of almost entirely European major leagues players. The players spend their
 218 prime form in Europe and only play at home during their junior and near-retirement period. Hence,
 219 European leagues benefit from those players' peak performance during their golden age.

220 Similar phenomena to Argentina exist in the case of Morocco, where its position in Figure 1 leads to
 221 a view of being a lower-mediocre league, meanwhile the national team performed excellently at the
 222 World Cup 2022. A deeper look into the players' profile, the line-up predominantly contains those
 223 who play abroad (*Morocco - Detailed Squad 2023*, 2022). Even more, several names that filled the regular
 224 starting squad were born and raised abroad without any experience in the domestic Moroccan
 225 professional league. Thanks to the strong diaspora tradition (Wagner, 2019) as well as linguistic
 226 advantage as a former French colony and a close tie with Spain, Moroccan-descent players can be
 227 found all over Europe bearing citizenship of European countries, Morocco, or dual. The dual
 228 citizenship policy is in their favor here (Seiberth et al., 2019), since it allows talented footballers bearing
 229 other countries' passports, mostly of European countries, to also have Moroccan citizenship without
 230 ditching their European citizenship. That way, they can continue to gain the benefit of being European
 231 players and at the same time opportunity in the national team when Morocco calls. To this end, again,
 232 national league quality does not necessarily represent the skill level of players with the corresponding
 233 nationality, if that country has many talents playing abroad in more competitive leagues. A past study
 234 confirms this argument, stating that the migration of players to foreign leagues positively affects the
 235 performance of their home country's national team (Allan & Moffat, 2014).

236 With a comparable player foot laterality percentage, the US league is seemingly much more expensive
 237 than the Australian is. In those countries, football, or soccer as they say, is not the most popular sport.
 238 Yet, the sport has stronger business and entertainment industry interest in the US (Kiuchi, 2013;
 239 Southall & Nagel, 2007). South Korea, although having an average league market value on a par with
 240 Japan, it is lagging in terms of laterality. Again, a quick internet survey reveals that many Korean talents
 241 compete abroad with a significant number of them joining Japanese clubs.

242 As for the ASEAN sides included in Figure 1, they are sitting at the low price, low foot balance, regime
 243 of the chart. A straightforward, or intuition based, reasoning can describe the fact that Indonesia and

244 Thailand are still requiring revolutionary advancement to catch up with the world class football
245 societies. Unlike the Morocco case, with the limited presence of ASEAN players in foreign top leagues,
246 the performance of the national teams mimics the league quality since the national squads are mostly
247 filled with those playing in national leagues.

248 Figure 2 and 3 both confirms the results of previous studies, as well as common knowledge, that
249 humans are predominantly right footed. Left-footed football players are rare, even in highly
250 competitive professional sports arena. But for decisive actions, such as goal scoring event, the left-
251 footedness rate is higher. It is also worth considering that players who are frequently taking decisive
252 actions are those high-performing names, which separates them from the overall descriptive statistics
253 of pass attempts in general.

254 5. Conclusion and Recommendation

255 This work has demonstrated the positive correlation between league quality and players' foot laterality.
256 As hypothesized, top leagues, with high monetary values, tend to have a high percentage of left-footed
257 goals. This fact would reiterate the common knowledge that professional-aspiring footballers should
258 work on their feet balance. The discussion also reveals that in the countries where many of their
259 football talents pursue career in major world-class leagues, the quality of the domestic league does not
260 necessarily reflect the performance of their national team in the international tournaments, since their
261 high-skilled players do not contribute to local league, but they do to the national team.

262 Finally, as an important supplementary remark, directly correlating that laterality with a financial-
263 related variable, as addressed in this paper, is expected to drive players' motivation to rigorously train
264 their less-dominant limbs by realizing the possible economic benefits of having both sides work in a
265 balanced manner. This tangible wealth-oriented point of view, not only said by their mentors but
266 proven by the data, hopefully can encourage ambitious talents to improve their skills in general as well
267 as to migrate abroad to play in highly competitive leagues abroad. In turn, it will be favorable not only
268 for the footballers themselves but also for the home country, especially in countries where football
269 has a massive impact on the local society, yet the national team is still struggling to enter the world's
270 elite playground, as in Indonesia.

271

272 References

273

Probing the quality of football leagues through player's foot laterality: A data analytics approach

ORIGINALITY REPORT

5%

SIMILARITY INDEX

4%

INTERNET SOURCES

2%

PUBLICATIONS

4%

STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Universitas Negeri Surabaya The State University of Surabaya Student Paper	3%
2	Submitted to Staffordshire University Student Paper	1%
3	www.nitwikfootball.com Internet Source	<1%
4	fbref.com Internet Source	<1%
5	escholarship.org Internet Source	<1%
6	citeseerx.ist.psu.edu Internet Source	<1%
7	"Machine Learning and Data Mining for Sports Analytics", Springer Science and Business Media LLC, 2019 Publication	<1%

Exclude quotes On

Exclude matches Off

Exclude bibliography On