Comparing Tiki-Taka and Long Ball Strategies in Modern Soccer

Josh Zhang, Chris Tran, Chloe Headland, and Matteo Deambrogio

2025-04-25

Research Question

For this project, we seek to analyze how certain playstyles influence team performance outcomes in professional soccer matches. Specifically, we seek to see the difference in performance between short-passing, possession-based tiki-taka playstyles in comparison to direct, long ball strategies. With that, we ultimately formulate our research question:

How do different playing styles—specifically the short-passing, possession-based tiki-taka versus the direct, long ball approach—influence team performance outcomes in professional soccer matches?

Dataset

Within this project, we plan on utilizing the **European Soccer Database** by Hugo Mathien, a comprehensive dataset containing match-level, team-level, and player-level information across several major European soccer leagues from 2008 to 2016.

This dataset is particularly well-suited to our analysis for a multitude of reasons. First, it provides detailed match outcomes including goals scored by the home and away teams, which enables the construction of team-level seasonal summaries needed to analyze playing styles. Second, it contains team identifiers and league affiliations, allowing us to group matches by league and season to observe stylistic patterns over time and across competitions. Third, it offers broad coverage across multiple seasons and top-tier leagues (e.g., England, Spain, Italy, Germany, and France), which is essential for conducting a comparative study of different tactical strategies, particularly when talking of long-ball vs tiki-taka strategies.

Given its breadth and public availability, the European Soccer Database provides a strong foundation for extracting team-level features necessary for our exploratory and inferential analyses regarding soccer playing styles. For more information about this dataset, please click here!

Variables Measured

Passing Metrics: Total number of passes, average pass distance, pass completion rate

Possession Metrics: Percentage of ball possession

Attacking Indicators: Number of shots, shots on target, goal conversion rates

Defensive Indicators: Interceptions and defensive clearances (to control for counter-play)

Exploratory Data Analysis

To start with our exploratory data analysis, let us analyze what datasets are comprised within the European Soccer Database.

```r
library(DBI)
```

```
## Warning: package 'DBI' was built under R version 4.3.3
```

```r
library(RSQLite)
```

```
## Warning: package 'RSQLite' was built under R version 4.3.3
```

```r
# Connect to the SQL file we seek to access
con <- dbConnect(RSQLite::SQLite(), "data/database.sqlite")

# See what tables exist
dbListTables(con)
```

```
## [1] "Country"          "League"            "Match"
## [4] "Player"           "Player_Attributes" "Team"
## [7] "Team_Attributes"  "sqlite_sequence"
```

Let us now extract each dataset and also examine them as well.

```r
# Check all tables
country_df <- dbReadTable(con, "Country")
league_df <- dbReadTable(con, "League")
match_df <- dbReadTable(con, "Match")
player_df <- dbReadTable(con, "Player")
player_at_df <- dbReadTable(con, "Player_Attributes")
team_df <- dbReadTable(con, "Team")
team_at_df <- dbReadTable(con, "Team_Attributes")
what_is_this_df <- dbReadTable(con, "sqlite_sequence")
```

```r
head(country_df, 5)
```

```
##       id    name
## 1      1 Belgium
## 2   1729 England
## 3   4769  France
## 4   7809 Germany
## 5  10257   Italy
```

```r
head(league_df, 5)
```

```
##       id country_id                    name
## 1      1          1 Belgium Jupiler League
## 2   1729       1729 England Premier League
## 3   4769       4769          France Ligue 1
## 4   7809       7809  Germany 1. Bundesliga
## 5  10257      10257           Italy Serie A
```

```r
head(match_df, 5)
```

```
##   id country_id league_id    season stage                date match_api_id
## 1  1          1         1 2008/2009     1 2008-08-17 00:00:00       492473
## 2  2          1         1 2008/2009     1 2008-08-16 00:00:00       492474
## 3  3          1         1 2008/2009     1 2008-08-16 00:00:00       492475
## 4  4          1         1 2008/2009     1 2008-08-17 00:00:00       492476
## 5  5          1         1 2008/2009     1 2008-08-16 00:00:00       492477
##   home_team_api_id away_team_api_id home_team_goal away_team_goal
## 1             9987             9993              1              1
```

```
## 2            10000           9994            0            0
## 3             9984           8635            0            3
## 4             9991           9998            5            0
## 5             7947           9985            1            3
##   home_player_X1 home_player_X2 home_player_X3 home_player_X4 home_player_X5
## 1             NA             NA             NA             NA             NA
## 2             NA             NA             NA             NA             NA
## 3             NA             NA             NA             NA             NA
## 4             NA             NA             NA             NA             NA
## 5             NA             NA             NA             NA             NA
##   home_player_X6 home_player_X7 home_player_X8 home_player_X9 home_player_X10
## 1             NA             NA             NA             NA              NA
## 2             NA             NA             NA             NA              NA
## 3             NA             NA             NA             NA              NA
## 4             NA             NA             NA             NA              NA
## 5             NA             NA             NA             NA              NA
##   home_player_X11 away_player_X1 away_player_X2 away_player_X3 away_player_X4
## 1              NA             NA             NA             NA             NA
## 2              NA             NA             NA             NA             NA
## 3              NA             NA             NA             NA             NA
## 4              NA             NA             NA             NA             NA
## 5              NA             NA             NA             NA             NA
##   away_player_X5 away_player_X6 away_player_X7 away_player_X8 away_player_X9
## 1             NA             NA             NA             NA             NA
## 2             NA             NA             NA             NA             NA
## 3             NA             NA             NA             NA             NA
## 4             NA             NA             NA             NA             NA
## 5             NA             NA             NA             NA             NA
##   away_player_X10 away_player_X11 home_player_Y1 home_player_Y2 home_player_Y3
## 1              NA              NA             NA             NA             NA
## 2              NA              NA             NA             NA             NA
## 3              NA              NA             NA             NA             NA
## 4              NA              NA             NA             NA             NA
## 5              NA              NA             NA             NA             NA
##   home_player_Y4 home_player_Y5 home_player_Y6 home_player_Y7 home_player_Y8
## 1             NA             NA             NA             NA             NA
## 2             NA             NA             NA             NA             NA
## 3             NA             NA             NA             NA             NA
## 4             NA             NA             NA             NA             NA
## 5             NA             NA             NA             NA             NA
##   home_player_Y9 home_player_Y10 home_player_Y11 away_player_Y1 away_player_Y2
## 1             NA              NA              NA             NA             NA
## 2             NA              NA              NA             NA             NA
## 3             NA              NA              NA             NA             NA
## 4             NA              NA              NA             NA             NA
## 5             NA              NA              NA             NA             NA
##   away_player_Y3 away_player_Y4 away_player_Y5 away_player_Y6 away_player_Y7
## 1             NA             NA             NA             NA             NA
## 2             NA             NA             NA             NA             NA
## 3             NA             NA             NA             NA             NA
## 4             NA             NA             NA             NA             NA
## 5             NA             NA             NA             NA             NA
##   away_player_Y8 away_player_Y9 away_player_Y10 away_player_Y11 home_player_1
## 1             NA             NA              NA              NA             NA
```

```
##  2            NA            NA            NA            NA            NA
##  3            NA            NA            NA            NA            NA
##  4            NA            NA            NA            NA            NA
##  5            NA            NA            NA            NA            NA
##    home_player_2 home_player_3 home_player_4 home_player_5 home_player_6
##  1            NA            NA            NA            NA            NA
##  2            NA            NA            NA            NA            NA
##  3            NA            NA            NA            NA            NA
##  4            NA            NA            NA            NA            NA
##  5            NA            NA            NA            NA            NA
##    home_player_7 home_player_8 home_player_9 home_player_10 home_player_11
##  1            NA            NA            NA             NA             NA
##  2            NA            NA            NA             NA             NA
##  3            NA            NA            NA             NA             NA
##  4            NA            NA            NA             NA             NA
##  5            NA            NA            NA             NA             NA
##    away_player_1 away_player_2 away_player_3 away_player_4 away_player_5
##  1            NA            NA            NA            NA            NA
##  2            NA            NA            NA            NA            NA
##  3            NA            NA            NA            NA            NA
##  4            NA            NA            NA            NA            NA
##  5            NA            NA            NA            NA            NA
##    away_player_6 away_player_7 away_player_8 away_player_9 away_player_10
##  1            NA            NA            NA            NA             NA
##  2            NA            NA            NA            NA             NA
##  3            NA            NA            NA            NA             NA
##  4            NA            NA            NA            NA             NA
##  5            NA            NA            NA            NA             NA
##    away_player_11 goal shoton shotoff foulcommit card cross corner possession
##  1             NA <NA>   <NA>    <NA>       <NA> <NA>  <NA>   <NA>       <NA>
##  2             NA <NA>   <NA>    <NA>       <NA> <NA>  <NA>   <NA>       <NA>
##  3             NA <NA>   <NA>    <NA>       <NA> <NA>  <NA>   <NA>       <NA>
##  4             NA <NA>   <NA>    <NA>       <NA> <NA>  <NA>   <NA>       <NA>
##  5             NA <NA>   <NA>    <NA>       <NA> <NA>  <NA>   <NA>       <NA>
##    B365H B365D B365A  BWH  BWD  BWA  IWH IWD IWA  LBH LBD  LBA PSH PSD PSA  WHH
##  1  1.73  3.40  5.00 1.75 3.35 4.20 1.85 3.2 3.5 1.80 3.3 3.75  NA  NA  NA 1.70
##  2  1.95  3.20  3.60 1.80 3.30 3.95 1.90 3.2 3.5 1.90 3.2 3.50  NA  NA  NA 1.83
##  3  2.38  3.30  2.75 2.40 3.30 2.55 2.60 3.1 2.3 2.50 3.2 2.50  NA  NA  NA 2.50
##  4  1.44  3.75  7.50 1.40 4.00 6.80 1.40 3.9 6.0 1.44 3.6 6.50  NA  NA  NA 1.44
##  5  5.00  3.50  1.65 5.00 3.50 1.60 4.00 3.3 1.7 4.00 3.4 1.72  NA  NA  NA 4.20
##     WHD  WHA  SJH SJD  SJA  VCH  VCD  VCA  GBH  GBD  GBA  BSH  BSD  BSA
##  1 3.30 4.33 1.90 3.3 4.00 1.65 3.40 4.50 1.78 3.25 4.00 1.73 3.40 4.20
##  2 3.30 3.60 1.95 3.3 3.80 2.00 3.25 3.25 1.85 3.25 3.75 1.91 3.25 3.60
##  3 3.25 2.40 2.63 3.3 2.50 2.35 3.25 2.65 2.50 3.20 2.50 2.30 3.20 2.75
##  4 3.75 6.00 1.44 4.0 7.50 1.45 3.75 6.50 1.50 3.75 5.50 1.44 3.75 6.50
##  5 3.40 1.70 4.50 3.5 1.73 4.50 3.40 1.65 4.50 3.50 1.65 4.75 3.30 1.67
```

```r
head(player_df, 5)
```

```
##   id player_api_id       player_name player_fifa_api_id            birthday
## 1  1        505942 Aaron Appindangoye            218353 1992-02-29 00:00:00
## 2  2        155782    Aaron Cresswell            189615 1989-12-15 00:00:00
## 3  3        162549        Aaron Doran            186170 1991-05-13 00:00:00
## 4  4         30572      Aaron Galindo            140161 1982-05-08 00:00:00
## 5  5         23780      Aaron Hughes             17725 1979-11-08 00:00:00
```

```
##   height weight
## 1 182.88    187
## 2 170.18    146
## 3 170.18    163
## 4 182.88    198
## 5 182.88    154
```

```r
head(player_at_df, 5)
```

```
##   id player_fifa_api_id player_api_id                date overall_rating
## 1  1             218353       505942 2016-02-18 00:00:00             67
## 2  2             218353       505942 2015-11-19 00:00:00             67
## 3  3             218353       505942 2015-09-21 00:00:00             62
## 4  4             218353       505942 2015-03-20 00:00:00             61
## 5  5             218353       505942 2007-02-22 00:00:00             61
##   potential preferred_foot attacking_work_rate defensive_work_rate crossing
## 1        71          right              medium              medium       49
## 2        71          right              medium              medium       49
## 3        66          right              medium              medium       49
## 4        65          right              medium              medium       48
## 5        65          right              medium              medium       48
##   finishing heading_accuracy short_passing volleys dribbling curve
## 1        44               71            61      44        51    45
## 2        44               71            61      44        51    45
## 3        44               71            61      44        51    45
## 4        43               70            60      43        50    44
## 5        43               70            60      43        50    44
##   free_kick_accuracy long_passing ball_control acceleration sprint_speed
## 1                 39           64           49           60           64
## 2                 39           64           49           60           64
## 3                 39           64           49           60           64
## 4                 38           63           48           60           64
## 5                 38           63           48           60           64
##   agility reactions balance shot_power jumping stamina strength long_shots
## 1      59        47      65         55      58      54       76         35
## 2      59        47      65         55      58      54       76         35
## 3      59        47      65         55      58      54       76         35
## 4      59        46      65         54      58      54       76         34
## 5      59        46      65         54      58      54       76         34
##   aggression interceptions positioning vision penalties marking standing_tackle
## 1         71            70          45     54        48      65              69
## 2         71            70          45     54        48      65              69
## 3         63            41          45     54        48      65              66
## 4         62            40          44     53        47      62              63
## 5         62            40          44     53        47      62              63
##   sliding_tackle gk_diving gk_handling gk_kicking gk_positioning gk_reflexes
## 1             69         6          11         10              8           8
## 2             69         6          11         10              8           8
## 3             69         6          11         10              8           8
## 4             66         5          10          9              7           7
## 5             66         5          10          9              7           7
```

```r
head(team_df, 5)
```

```
##   id team_api_id team_fifa_api_id    team_long_name team_short_name
```

```
## 1  1          9987          673        KRC Genk              GEN
## 2  2          9993          675        Beerschot AC          BAC
## 3  3         10000        15005  SV Zulte-Waregem            ZUL
## 4  4          9994         2007  Sporting Lokeren            LOK
## 5  5          9984         1750 KSV Cercle Brugge            CEB
```

```r
head(team_at_df, 5)
```

```
##   id team_fifa_api_id team_api_id                date buildUpPlaySpeed
## 1  1              434        9930 2010-02-22 00:00:00               60
## 2  2              434        9930 2014-09-19 00:00:00               52
## 3  3              434        9930 2015-09-10 00:00:00               47
## 4  4               77        8485 2010-02-22 00:00:00               70
## 5  5               77        8485 2011-02-22 00:00:00               47
##   buildUpPlaySpeedClass buildUpPlayDribbling buildUpPlayDribblingClass
## 1              Balanced                   NA                    Little
## 2              Balanced                   48                    Normal
## 3              Balanced                   41                    Normal
## 4                  Fast                   NA                    Little
## 5              Balanced                   NA                    Little
##   buildUpPlayPassing buildUpPlayPassingClass buildUpPlayPositioningClass
## 1                 50                   Mixed                   Organised
## 2                 56                   Mixed                   Organised
## 3                 54                   Mixed                   Organised
## 4                 70                    Long                   Organised
## 5                 52                   Mixed                   Organised
##   chanceCreationPassing chanceCreationPassingClass chanceCreationCrossing
## 1                    60                     Normal                     65
## 2                    54                     Normal                     63
## 3                    54                     Normal                     63
## 4                    70                      Risky                     70
## 5                    53                     Normal                     48
##   chanceCreationCrossingClass chanceCreationShooting
## 1                      Normal                     55
## 2                      Normal                     64
## 3                      Normal                     64
## 4                        Lots                     70
## 5                      Normal                     52
##   chanceCreationShootingClass chanceCreationPositioningClass defencePressure
## 1                      Normal                      Organised              50
## 2                      Normal                      Organised              47
## 3                      Normal                      Organised              47
## 4                        Lots                      Organised              60
## 5                      Normal                      Organised              47
##   defencePressureClass defenceAggression defenceAggressionClass
## 1               Medium                55                  Press
## 2               Medium                44                  Press
## 3               Medium                44                  Press
## 4               Medium                70                 Double
## 5               Medium                47                  Press
##   defenceTeamWidth defenceTeamWidthClass defenceDefenderLineClass
## 1               45                Normal                    Cover
## 2               54                Normal                    Cover
## 3               54                Normal                    Cover
## 4               70                  Wide                    Cover
```

```
## 5                    52                    Normal                    Cover
```

```
head(what_is_this_df, 5)
```

```
##      name     seq
## 1    Team 103916
## 2 Country  51958
## 3  League  51958
## 4   Match  51958
## 5  Player  11075
```

Country Dataframe: It appears that the country dataframe seems to map the country a league resides in within the dataset with a corresponding `id`. It appears that the `id` columns appears to be universal connector across datasets. Therefore it would be nice to keep around to ajoin datasets if needed.

League Dataframe: Similar to the country dataframe, every league is also given a corresponding unique `country_id` as well. We will likely use this dataframe to filter out datasets, on the basis of leagues. For the sake of simplicity, we will likely primarily analyze performance of short and long ball strategies within the context of the Big-5 Leagues.

Match Dataframe: The match dataframe appears to contain every single match within the dataset, giving each individual match a unique `match_api_id`. With each entry, goals for each team are shown, the `home_team_api_id` and `away_team_api_id` are shown, and a series of statistics are shown as well.

Player Dataframe: This is a rather funny dataset. It contains a unique player identifer id of `player_api_id`, their names, fifa game id, birthday, and physicals. This may be useful for tracking down certain players, but likely not for our analysis.

Player Attribute Dataframe: This is also a rather funny dataset. We believe that these are player's FIFA ratings. While quirky, not likely what we need for our dataset.

Team Dataframe: This appears to be a dataset housing the identifers for each team, along with the corresponding official team name.

Team Attribute Dataframe: Once again, same as the Player Attribute Dataframe, but for teams this time.

What Is This Dataframe: This appears to show the number of entries within each dataframe.