# Factored
# Data Visualization Project

**Charly Moreno**
**September 12th, 2022**

Factored

# AGENDA

1. Step-by-Step elaboration process

2. Presentation of the Report: Club Soccer Performance

2.1 Best / Worst performers per league

2.2 League's competitiveness

2.3 What if scenario evaluation

3. Conclusions / Ending

Factored

# Step by Step

1. Importing data to SQL Server

2. Problem abstraction
   a. Review of questions to be answered
   b. EDA of the tables to establish the analysis path
      i. National leagues analysis (last season)
      ii. Exclusion of leagues w/o info in last season
      (Women's leagues, UEFA international leagues)

3. Definition / Calculation of display tables
   a. Summary by league / best teams /what-if scenario /KPIs

4. Elaboration of Data Model for reporting

5. Creation of final report (PBI Desktop)

7. Summary of conclusions

## BI Tools used

## Languages used

1. SQL
2. Lenguaje M

Factored

FiveThirtyEight

# Club Soccer Performance

Charly Moreno
September 12th, 2022

# Factored | Leagues Performance during the last Season

LEAGUE

Chinese Super League

Season

2019

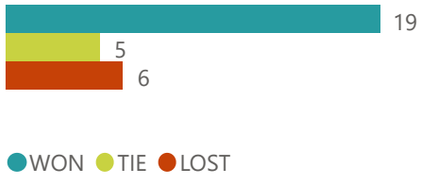
Please, select a **LEAGUE** to see the indicators of the Top and Worst performers teams of the last session.

## Top performers

| TEAM | POINTS | MATCHES | WON | T |
|---|---|---|---|---|
| Guangzhou Evergrande | 72 | 30 | 23 | |
| Beijing Guoan | 70 | 30 | 23 | |
| Shanghai SIPG | 66 | 30 | 20 | |
| Jiangsu Suning FC | 53 | 30 | 15 | |
| Shandong Luneng | 51 | 30 | 15 | |
| **Total** | **312** | **150** | **96** | |

## Average of Top performers

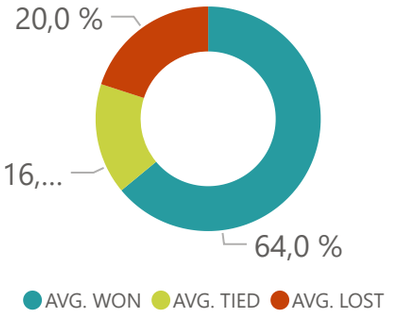
Count of match results




## Average points of Top performers

62,4

## Match results of group (%)




## Max of points

72,0

## Min of points

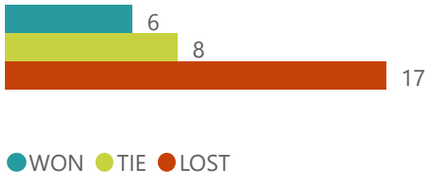
51,0

## Worst performers

| TEAM | POINTS | MATCHES | WON | TI |
|---|---|---|---|---|
| Guangzhou RF | 32 | 30 | 9 | |
| Shanghai Greenland | 30 | 30 | 8 | |
| Tianjin Quanujian | 25 | 30 | 4 | 1 |
| Shenzhen FC | 21 | 30 | 4 | |
| Guizhou Renhe | 14 | 30 | 3 | |
| **Total** | **122** | **150** | **28** | **3** |

## Average of Worst performers

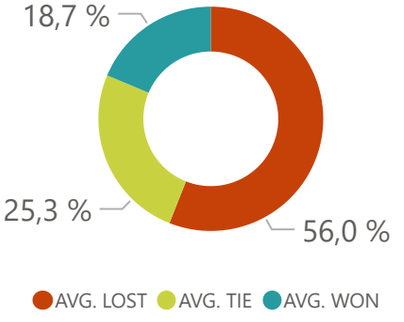
Count of match results




## Average points of Worst performers

24,4

## Match results of group (%)




## Max of points

32.0

## Min of points

14,0

# Factored | Leagues Competitiveness during the last Session

LEAGUE

Chinese Super League

Season

2019

Please, select a **LEAGUE** to see the competitiveness indicators of the last session.

Total teams

16

Total matches

30

Competitiveness Rank

39

SD. Tie (%)

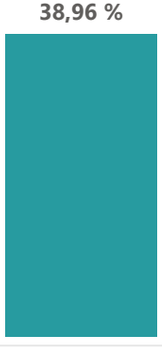
9,18 %

SD. Won (%)

20,72 %

Avg. Tie (%)


Avg. Won (%)




Statistics of % Won per team

**76,67 %**
Max.

**50,00 %**
Percentile 75%

**33,33 %**
Percentile 50%

**26,67 %**
Percentile 25%

**10,00 %**
Min.

Measuring the competitiveness of each league as the smallest dispersion of the average winning percentage that teams can have (because anyone can win), we have:

1. **Spanish Segunda Division** is the **BEST (R°1) LEAGUE** in **COMPETITIVENESS** (with 4.22% of average winning and 2.20% of standard deviation around it by the teams)
2. **Chinese Super League** is the **WORST (R°39) LEAGUE** in **COMPETITIVENESS** (with 38.96% of average winning and 20.72% of standard deviation around it by the teams)

# What if scenario evaluation - Leagues

## Methodology

To compare leagues internationally, an Adjustment Factor (AF) was calculated to assess each team's performance when switching.

This factor was calculated and used as follows:

1. **Start**.
2. Using the Score Power Index (**SPI**) done by **FiveThirtyEight**
3. A General SPI (GSPI) statistic was generated calculating the average SPI per League
4. The best and worst Leagues were selected using the GSPI
5. An AF was calculated (difference between GSPI of Leagues)
6. The AF improves performance when downgrading (by dividing)
7. The AF decreases performance when upgrading (by multiplying)
8. Worst teams from the best League were selected to downgrading
9. Best teams from the worst League were selected to upgrading
10. The AF will recalculate the final points of the selected teams
11. Final results will be seen as a forecast of the teams' performance
12. **End**.

## Results - Best League (Barclays PL) - Worst League (English LT)

### Worst teams from Barclays PL at English LT

Added teams in worst league

| TEAM | POINTS | LEAGUE_RANK |
|---|---|---|
| Everton | 60,00 | 1 |
| West Ham United | 60,00 | 2 |
| Nottingham Forest | 60,00 | 3 |
| Aston Villa | 60,00 | 4 |
| Leicester City | 55,00 | 6 |
| **Total** | **295,00** | **16** |

| Avg. points added teams | Avg. points of other teams |
|---|---|
| 59.0 | 48.5 |

| Matches | AF (/) |
|---|---|
| 46 | 0.60255 |

### Best teams from English LT at Barclays PL

Added teams in worst league

| TEAM | POINTS | RANK_1 |
|---|---|---|
| Leyton Orient | 35,00 | 20 |
| Stevenage | 33,00 | 21 |
| Barrow | 33,00 | 22 |
| Doncaster Rovers | 32,00 | 24 |
| Northampton Town | 32,00 | 25 |
| Salford City | 32,00 | 26 |
| **Total** | **197,00** | **138** |

| Avg. points added teams | Avg. points of other teams |
|---|---|
| 32.8 | 40.1 |

| Matches | AF (*) |
|---|---|
| 38 | 0.60255 |

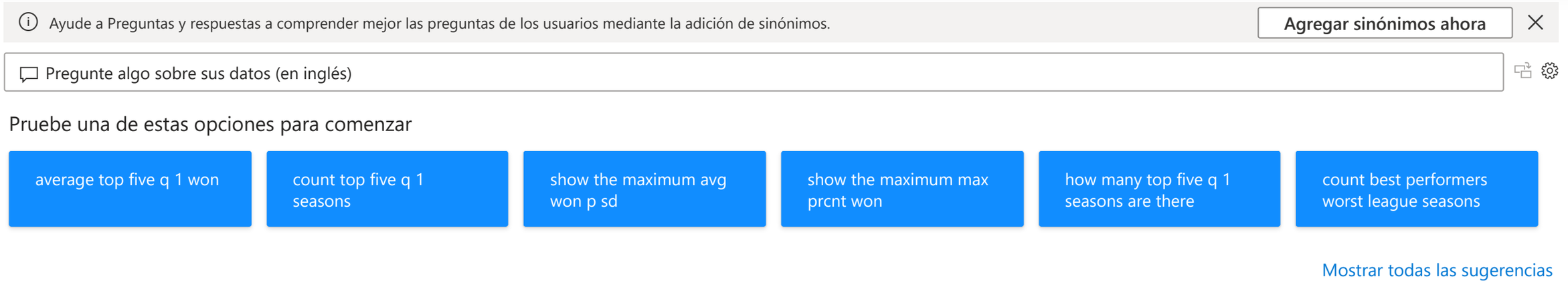
# What if scenario evaluation - Leagues

## Conclusions

- Initial analysis was conducted on the FiveThirtyEight Club Football Data. Conducting it, different statistics per team and League were calculated.
- Using the latest season for each League, the **Spanish Segunda Division** was identified as the best League in competitive performance (with a 4.22% of average winning and 2.20% of standard deviation around it by the teams). This League has a **difference** only of **6.6** in the points won between the best and worst groups of the tournament.
- Also, the **Chinese Super League** was identified as the worst League in competitive performance (with 38.96% of average winning and 20.72% of standard deviation around it by the teams). This League has a considerable **difference** of **38** in the points won between the best and worst groups of the tournament.
- To generate comparisons between different Leagues, a methodology for an Adjustment Factor was proposed. This factor is based on the SPI ranking generated by FiveThirtyEight and can help to forecast the performance of teams when switching Leagues.
- Using the methodology proposed, the best League was identified as the **Barclays Premier League** and the worst League identified was the **English League Two**. When using the Adjustment Factor, the result is that the worst teams of the best League performed the best in the worst League, and vice-versa.
- Finally, for further work, other comparisons can be made using historic data. For example, it is possible to generate dynamic indicators of competitiveness for the Leagues using different seasons. In addition to that, other Leagues not considered in this exercise can be studied using historical data (UEFA Leagues and Women's League).

# Factored | What if scenario evaluation - Leagues

## Any other questions?




# Thank you!

# Factored
# Data Visualization Project

**Charly Moreno**
**September 12th, 2022**