

Data Analytics Final Project

CR7

A COMPREHENSIVE ANALYSIS OF CRISTIANO RONALDO'S GOAL-SCORING CAREER

A project by:
Filipa Barroqueiro
Joana Correia
Luis Nicolau

PROJECT OVERVIEW

This project provides a full analytics-driven breakdown of Cristiano Ronaldo's **goal-scoring career**, combining multiple data sources into a validated and enriched dataset covering 954 official goals across clubs and the Portuguese National Team.

MOTIVATION & OBJECTIVES

Our aim was to go beyond commonly available statistics and rebuild CR7's scoring history through a clean, transparent, and fully traceable data pipeline.

By combining Python, Power Query, and Power BI, we transformed fragmented datasets into a **unified analytical model** that captures his evolution, efficiency, and performance across competitions and stages of his career.



Build a Complete & Accurate Dataset

Reconstruct CR7's full goal history by merging the Kaggle dataset with manually validated updates: missing goals, competition corrections, match dates, and the latest goals up to 23 November 2025.



Explore Career-Long Performance Patterns

Analyze how Ronaldo's goals distribute across clubs, competitions, and seasons, including age at each goal, scoring streaks, and efficiency ratios such as "matches with goals vs. total matches".



Create Interactive Dashboards for Insights

Develop Power BI dashboards embedded in PowerPoint, enabling the audience to explore trends, compare clubs, filter competitions, and interact with detailed statistics across Ronaldo's entire career.

DATA SOURCES FEATURE OVERVIEW

Our analysis combines multiple verified sources to rebuild Cristiano Ronaldo's complete goal history with accuracy and transparency.



Kaggle Base Dataset

Initial dataset containing 710 official goals, used as the foundation for cleaning, structuring and validating all records.



Manual Goal Additions

Enrichment of missing goals across seasons, including competitions not listed in the original Kaggle file.



Latest Goal Updates

Added all goals up to 23/11/2025, including the most recent bicycle kick and additional Al-Nassr records (UAFA).



National Team Data

Integrated Portugal National Team goals to complete the official tally across all competitions.



Injury Data

Standardized injury dates, durations and recovery details for reliable performance impact analysis.



Discrepancy Resolution

Reconciled differences between public stats and raw dataset values — expanding the total from 710 → 954 goals after full data validation.

DATA ENGINEERING PROCESS



STEP ONE

Python Data Cleaning

Handling missing values

Standardizing columns

Merging datasets

Matching goal dates & competitions

01

CR7 Last Checkpoint: 2 minutes ago

Run Kernel Settings Help

Code

JupyterLab Python [conda env:base] *

```
Goals_Opponent"] = df["Result_clean"].str.split(":").str[1]  
inverter para número inteiro  
Goals_Opponent"] = df["Goals_Opponent"].astype(int)
```

read()

| Season | Competition | Matchday | Date | Venue | Club | Opponent | Result | Playing_Position | Minute | ... | Type | Goal_assist | Category | Org | | |
|--------|--------------------------|----------|----------|-------|-------------|-------------------------|--------|------------------|--------|------|------|-------------|-------------------|------------------|------------------|---------------|
| 02/03 | Liga Portugal | 6 | 10-07-02 | H | Sporting CP | Moreirense FC | 3:00 | | LW | 34 | ... | Solo run | NaN | National Leagues | Portugal Fed. | |
| 02/03 | Liga Portugal | 6 | 10-07-02 | H | Sporting CP | Moreirense FC | 3:00 | | LW | 90+5 | ... | Header | Rui Jorge | National Leagues | Portugal Fed. | |
| 02/03 | Liga Portugal | 8 | 10/26/02 | A | Sporting CP | Boavista FC | 1:02 | | | NaN | 88 | ... | Right-footed shot | Carlos Martins | National Leagues | Portugal Fed. |
| 02/03 | Taca de Portugal Placard | Round 4 | 11/24/02 | H | Sporting CP | CD Estarreja | 4:01 | | | NaN | 67 | ... | Left-footed shot | Cesar Prates | National Cups | Portugal Fed. |
| 02/03 | Taca de Portugal Placard | Round 5 | 12/18/02 | H | Sporting CP | FC Oliveira do Hospital | 8:01 | | | NaN | 13 | ... | NaN | NaN | National Cups | Portugal Fed. |



DATA ENGINEERING PROCESS



STEP TWO



Data Transformation in Power Query

- Importing & shaping
- Removing duplicates
- Standardizing competition names
- Normalizing clubs over years



STEP THREE

Final Data Model

- Date dimension
- Clubs dimension
- Matches table
- Goals table
- Relationships diagram



CLUBS OVERVIEW

Goal Distribution

01

- Strongest scoring share in the Spanish league
- High consistency across multiple competitions
- Significant impact in Champions League performance

Club Performance

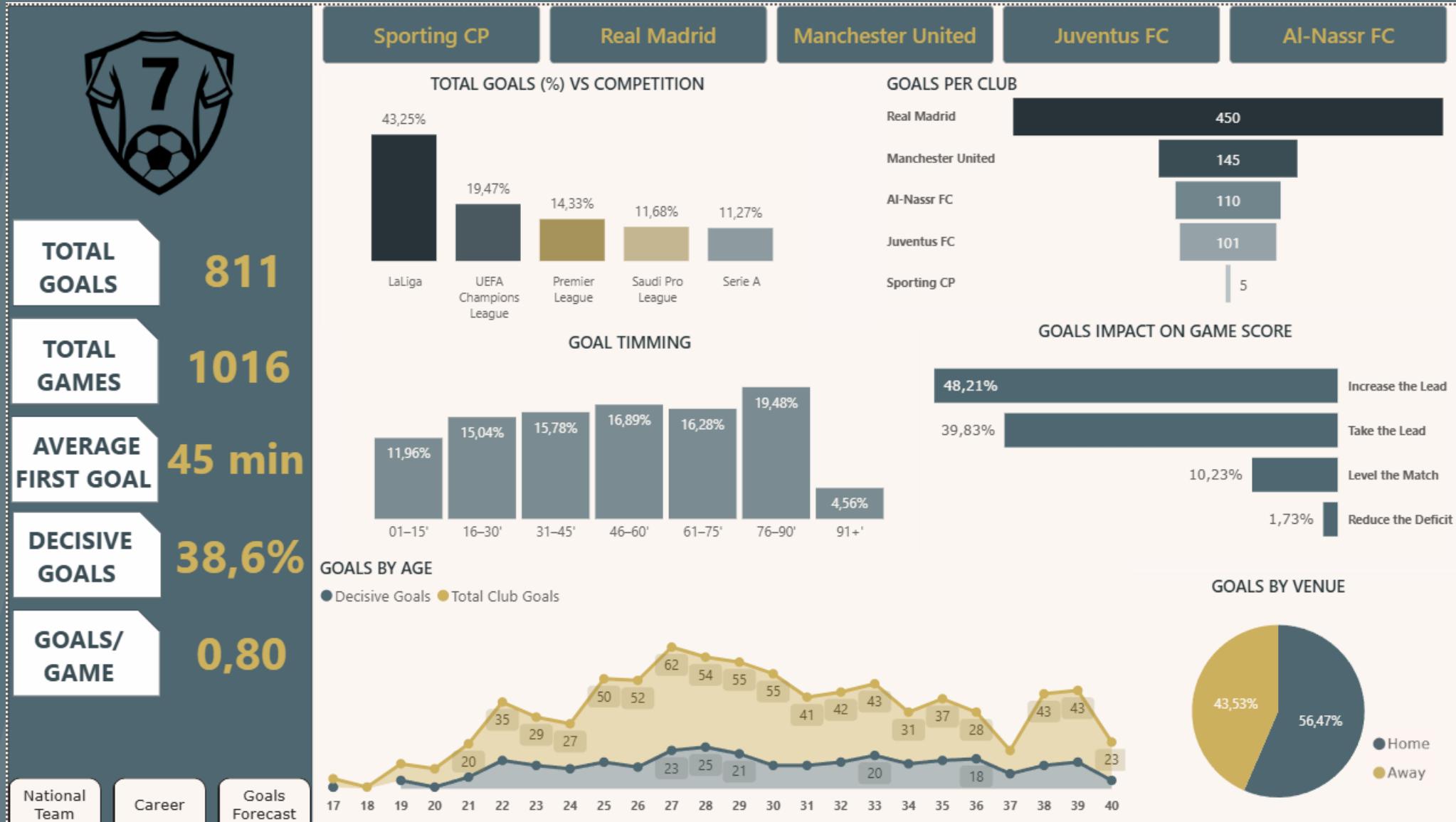
02

- Highest goal tally achieved at the second club
- Solid contribution at all clubs with decisive goals in key matches
- Proven adaptability across different leagues and playing styles

Scoring Profile

03

- Peaks in scoring frequency between ages 24-29
- Majority of goals scored at home venues
- Most goals increase or secure the team's lead



NATIONAL TEAM OVERVIEW

Goal Volume

01

- 143 total goals across 226 matches
- Average of 0.63 goals per match

Goal Distribution by Competition

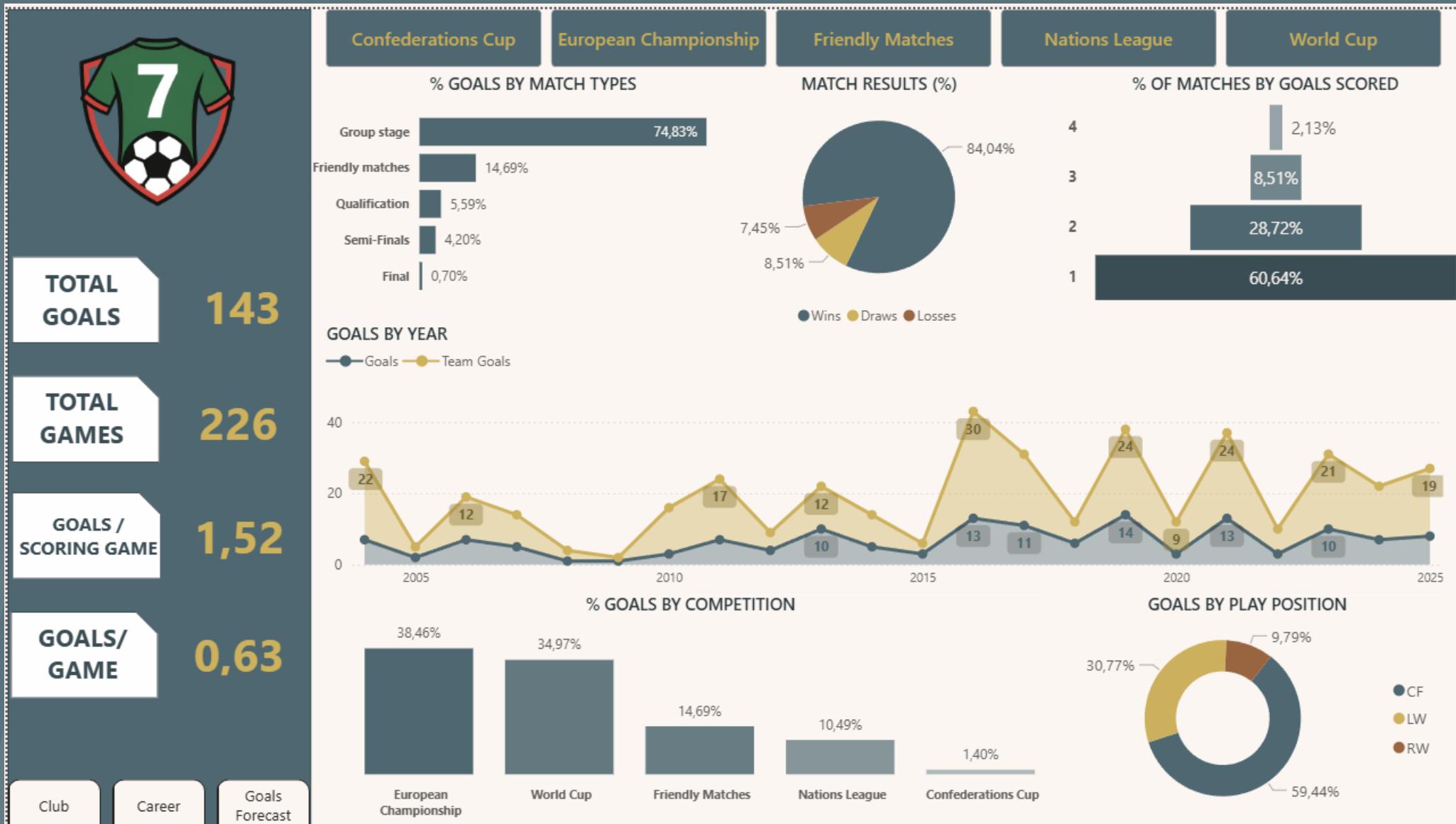
02

- Highest goal share from European Championship (37.06%)
- Significant contributions from World Cup (34.97%)
- Goals from Friendly Matches and Nations League among other competitions

Goal Scoring Positions

03

- Predominantly scores from CF (59.45%)
- Significant scoring contributions from RW and LW positions



CAREER OVERVIEW

Career Volume

01

- 954 total goals across 1242 matches
- 642 games with goals scored
- Career average of 0.77 goals per match

Seasonal Consistency

02

- Sustained goal production from early career (<20) through late 30s
- Peak seasons exceeding 50+ goals
- Maintained double-digit goals for nearly all seasons

Longevity & Projection

03

- Clear upward career curve toward the 1000-goal milestone
- Long-term productivity with minimal decline over time
- Strong performance forecast into late-career seasons

CAREER OVERVIEW

Goals vs. Injuries Profile

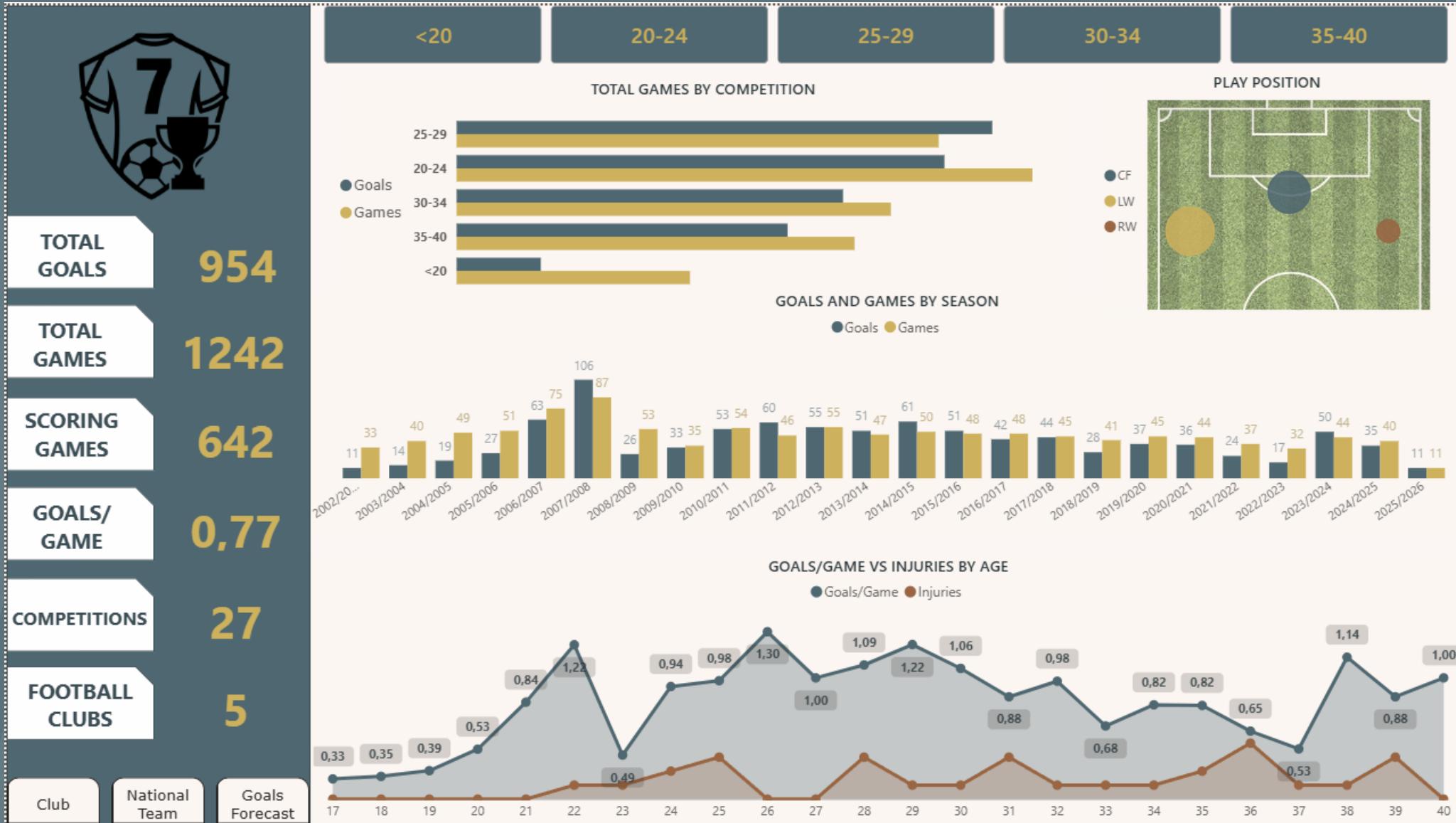
04

- More than 0.8 goals/game from ages 20 to 35
- Multiple age peaks above 1 goals/game
- Injuries present but generally not performance-limiting

Tactical Role

05

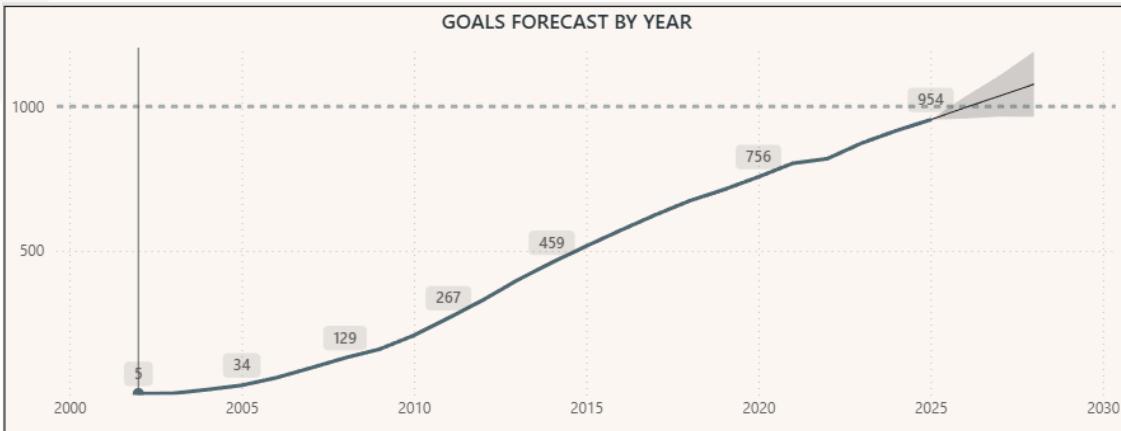
- Versatile positioning: primarily center forward, with significant impact as left winger
- Effective across multiple tactical systems
- Goal contribution unaffected by positional shifts



1K GOALS FORECAST

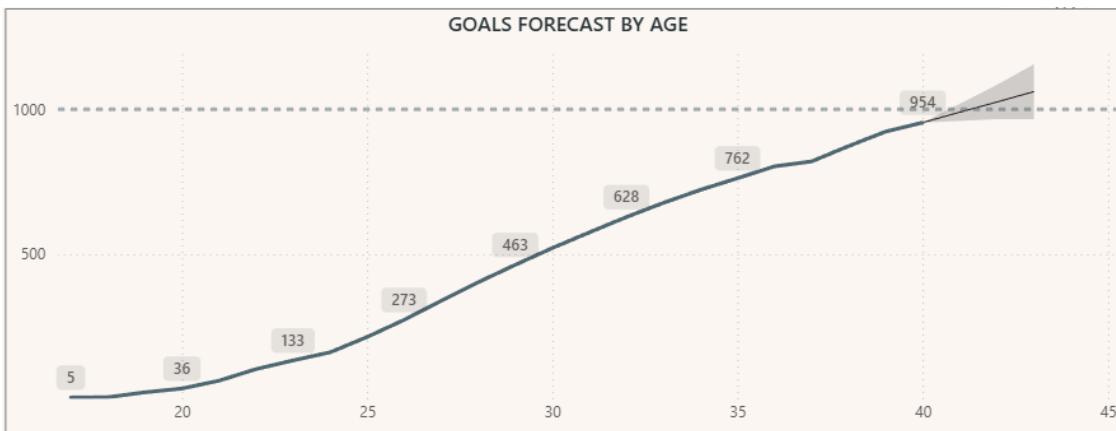
$$EM = \frac{1.000 - CG}{GpM} = \frac{1.000 - 954}{0,77} \simeq 60$$

*EM → Expected Matches
CG → Current Goals
GpM → Goals per Match*



Year
2026

Age
41



Games
60



SUMMARY

All insights supported by a clean, structured, enriched, and fully validated dataset.

Data confirms consistency across continents, competitions, and tactical roles.

954 official goals

recorded after full data
cleaning and validation

Scored across 20+ different competitions

from domestic leagues to
international tournaments

Five clubs + Portugal National Team

delivering consistently in every
environment

Goals scored from age 17 to 40

showing unmatched career
longevity

One of the greatest and longest-lasting goal scorers in football history.

**THANK
YOU**