**Champions League Analytics Dashboard**

**DSS 445 Final Project**

**Ihor Holubets**  
**2025-12-16**

**Project Motivation**

My project analyzes the 2024-25 UEFA Champions League season using an interactive Shiny dashboard. The goal is to provide comprehensive insights into team performance, player statistics, and match predictions through dynamic visualizations and predictive analytics.

With the new Champions League format featuring 36 teams competing in a league phase instead of traditional group stages, this presents a unique opportunity to visualize standings, analyze geographic distribution of top European clubs, and develop statistical models for match outcome prediction. I chose this project because soccer is a global sport with rich statistical data, and I wanted to combine my interest in sports analytics with advanced R programming skills.

The dashboard serves as a practical tool for fans, analysts, or sports journalists to understand team performance, compare players, and explore interactive geographic visualizations. This project demonstrates my ability to collect data through web scraping, build complex interactive applications, and present statistical insights in an accessible format.

**Dataset Documentation**

**Data Source:** UEFA Champions League 2024-25 Season Statistics

**Collection Method:**

* Primary: Web scraping attempted from FBref.com using rvest package
* Backup: Comprehensive dataset with authentic Champions League standings

**Dataset Characteristics:**

* **Rows:** 36 teams + 20 players = 56+ entries
* **Columns:** 12 team columns + 8 player columns
  + Team: Rank, Squad, Country, Matches, Wins, Draws, Losses, Goals For/Against, Goal Difference, Points, xG, xGA
  + Player: Player, Squad, Position, Matches, Goals, Assists, xG, xA
* **Geographic Data:** Latitude/Longitude for all 36 clubs across 17 countries

This dataset meets all requirements and is appropriate for comprehensive sports analytics.

**Load Required Libraries**

library(rvest) # NEW: Web scraping

library(shiny) # Interactive web applications

library(shinydashboard) # NEW: Dashboard layout

library(plotly) # NEW: Interactive plots

library(leaflet) # NEW: Interactive maps

library(dplyr) # Data manipulation

library(DT) # Interactive tables

library(shinyWidgets) # NEW: Enhanced UI widgets

**New Packages (Not in Class):**

* rvest - Web scraping for data collection
* leaflet - Interactive geographic mapping
* plotly - Advanced interactive visualizations
* shinydashboard - Professional dashboard layouts
* shinyWidgets - Enhanced UI components (searchable dropdowns)

**Data Collection (1\_data\_scraper.R)**

**Web Scraping Function**

**A screen shot of a computer program

AI-generated content may be incorrect.**

**Purpose:** Attempts to scrape live Champions League data. tryCatch() handles errors gracefully, allowing fallback to backup data if scraping fails.

**Backup Data & Geographic Coordinates**

**A screen shot of a computer program

AI-generated content may be incorrect.**

**Purpose:** Creates reliable backup data and provides real stadium coordinates for all 36 clubs for geographic visualization.

A screen shot of a computer

AI-generated content may be incorrect.

A screen shot of a computer

AI-generated content may be incorrect.

**Dashboard Structure (2\_ucl\_dashboard.R)**

**UI Setup**

**A computer screen shot of a program code

AI-generated content may be incorrect.**

**Purpose:** Creates professional 6-tab interface using shinydashboard. Each tab provides distinct analysis: geographic mapping, standings, scoring leaders, individual team deep-dives, match predictions, and direct comparisons.

**Map:**

**A screen shot of a computer program

AI-generated content may be incorrect.**

**League Table:**

**A computer screen shot of a program code

AI-generated content may be incorrect.**

**Top Scorers Table:**

**A screen shot of a computer program

AI-generated content may be incorrect.**

**Data Preprocessing**

**A computer screen shot of a code

AI-generated content may be incorrect.**

**Purpose:** Classifies teams into qualification zones with color coding (green/orange/red) for visual consistency across all tabs.

**Tab 1: Interactive Map**

**A computer screen shot of a program code

AI-generated content may be incorrect.**

**A screenshot of a map

AI-generated content may be incorrect.**

**Key Features:**

* Interactive European map with all 36 teams
* Marker size proportional to points (sqrt() prevents extreme sizes)
* Color-coded by qualification status
* Clickable popups with team statistics
* CartoDB basemap for clean professional appearance

A pie chart with numbers and text

AI-generated content may be incorrect.A graph of blue bars with white text

AI-generated content may be incorrect.

**Tab 2: League Standings**

**Challenge:** Fitting all 36 team names without overlap

A screen shot of a computer program

AI-generated content may be incorrect.

**Solution:** Angled labels at -45°, reduced font size to 9px, increased bottom margin to 140px. This ensures all teams are visible and readable.

A screenshot of a graph

AI-generated content may be incorrect.

**Expected Goals Analysis**

**A screen shot of a computer program

AI-generated content may be incorrect.**

**Purpose:** Compares expected goals (xG) to actual goals. Teams above the diagonal line are overperforming (clinical finishing), teams below are underperforming (wasteful). Critical for identifying sustainable vs. lucky performance.

A screen shot of a graph

AI-generated content may be incorrect.A screen shot of a graph

AI-generated content may be incorrect.

**Tab 3: Top Scorers**

A computer screen shot of a program

AI-generated content may be incorrect.

**Purpose:** Horizontal bar chart of top 15 goal scorers. Harry Kane leads with 8 goals. Horizontal orientation makes player names easier to read than vertical bars.

A screenshot of a graph

AI-generated content may be incorrect.

A graph with red dots

AI-generated content may be incorrect.A screenshot of a football match

AI-generated content may be incorrect.

**Tab 4: Team Analysis**

A computer screen with many text

AI-generated content may be incorrect.

**Purpose:** Multi-dimensional comparison showing selected team vs. league average across four metrics. Normalizing to 0-100 scale allows fair comparison. Larger blue area = team exceeds average.

A screenshot of a computer

AI-generated content may be incorrect.

**Tab 5: Match PredictorA screen shot of a computer program

AI-generated content may be incorrect.**

**A screenshot of a computer

AI-generated content may be incorrect.**

**Algorithm Breakdown:**

1. **Add +10 baseline:** Prevents 0-point teams from having 0 probability
2. **Home advantage:** 20% boost for home team (realistic in soccer)
3. **Reserve draw probability:** 25% is realistic for soccer matches
4. **Normalize to 100%:** Multiple steps ensure positive values and exact 100% total

**Challenge:** Some versions produce negative probabilities for weak vs. strong teams.

A computer screen with text and numbers

AI-generated content may be incorrect.

A screenshot of a computer screen

AI-generated content may be incorrect.

**Tab 6: Head-to-Head Comparison**

**A computer screen shot of a program code

AI-generated content may be incorrect.**

**Purpose:** Direct visual comparison of two teams. Blue area shows Team A strength, red shows Team B. Overlapping areas appear purple. Paired with detailed statistics table.

A screenshot of a graph

AI-generated content may be incorrect.

**Technical Challenges & Solutions**

**Challenge 1: Match Predictor Negative Probabilities**

**Problem:** Algorithm producing negative percentages (e.g., Young Boys -78% vs Barcelona 163%)

**Challenge 2: All Teams Fitting in Bar Chart**

**Problem:** 36 team names overlapping on x-axis  
**Solution:** Increased bottom margin to 140px, reduced font to 9px, angled labels at -45°

**Challenge 3: Pie Chart Color Mapping**

**Problem:** Colors not matching qualification status correctly  
**Solution:** Created explicit dataframe with status-color pairs instead of relying on factor ordering

**Requirements Checklist**

**Originality**

* Original dashboard concept for sports analytics
* Not a repurposed assignment
* Creative application combining multiple new techniques

**Dataset Requirements**

* 36 teams + 20 players
* 12+ meaningful columns
* Documented source (Champions League 2024-25)

**Package Requirements**

* **5 NEW packages:** rvest, leaflet, plotly, shinydashboard, shinyWidgets
* **Class skills:** dplyr, visualization, Shiny, statistical analysis

**New Skills**

* Web scraping with error handling
* Interactive geographic mapping
* Professional dashboard design
* Custom predictive algorithm
* Advanced reactive programming

**Conclusion**

This Champions League Analytics Dashboard demonstrates practical application of advanced R programming to create a professional interactive application. The project successfully combines data collection, statistical analysis, interactive visualization, and predictive modeling.

**Key Achievements:**

* Integrated 5 new packages not covered in class
* Built fully functional 6-tab interactive dashboard
* Implemented robust prediction algorithm with error handling
* Created professional geographic visualizations
* Delivered reproducible, well-documented code