Cairo University

Faculty of Computers and Artificial Intelligence

European Football Leagues

Data Warehouse & ETL Project

Course:
IS341 - DataWarehousing

Supervised:
Dr. Ali Zidan

Team Members:

Name	ID	Email
Mohammed Atef Abd El Kader	20231143	20231143@stud.fci-cu.edu.eg
Mohammed Ahmed Mohammed	20231134	20231134@stud.fci-cu.edu.eg
Mostafa Ehab Mostafa Akl	20231167	20231167@stud.fci-cu.edu.eg
Mostafa Mahmoud Fathy	20231244	20231244@stud.fci-cu.edu.eg

May. 2026

Table Of Contents

1. Source Database 4
- **1.1. Source Database Overview** 4
  - **1.1.1. Source Tables** 4
  - **1.1.2. Relationships Table** 4
- **1.2. Source DDL** 5
- **1.3. Source ERD** 8

2. Staging Area 9
- **2.1. STG DDL** 9
- **2.2. STG ERD** 12

3. Data Warehouse 13
- **3.1. Dimensional Model** 13
- **3.2. Fact Tables** 14
  - **3.2.1. `Fact_Team_Monthly_Stats`** 14
  - **3.2.2. `Fact_Scoring_Efficiency`** 16
  - **3.2.3. `Fact_Coach_Stadium_Performance`** 18
- **3.3. DWH DDL** 20
- **3.4. DWH ERD** 24

4. ETL Process 25
- **4.1. ETL Architecture** 25
- **4.2. Extraction Phase** 25
- **4.3. Transformation Phase** 26
- **4.4. Loading Phase** 26
- **4.5. ETL Workflow Steps** 27
- **4.6. Incremental Load Process** 27
- **4.7. SSIS Control / Data Flow Screenshots** 28
  - **4.7.1. Load `Dim_Coach`** 28
  - **4.7.2. Load `Dim_Leagues`** 28
  - **4.7.3. Load `Dim_Score_junk`** 29
  - **4.7.4. Load `Dim_Stadium`** 29
  - **4.7.5. Load `Dim_Team`** 30
  - **4.7.6. Load `Fact_Team_Monthly_Stat`** 31
  - **4.7.7. Load `Fact_Scoring_Efficiency`** 32
  - **4.7.8. Load `Fact_Coach_Stadium_Performance`** 33

5. KPIs 34
- **5.1. Team Performance KPIs (Season / League / Team Level)** 34
- **5.2. League Level KPIs (League + Season)** 34
- **5.3. Match Performance KPIs (Match + Date Level)** 35
- **5.4. Team Form & Momentum KPIs** 35
- **5.5. Coach & Stadium KPIs** 35

6. SSIS Packages & Scheduling Screenshots 36
6.1. Schedule Packages Screenshots 36
6.2. After First Run Screenshots 37
6.3. After Second Run Screenshots 39

7. Queries & Insights 40
7.1. Team Performance Analysis 40
7.2. Goals & Offensive Performance 40
7.3. Team Efficiency Metrics 41
7.4. League Performance Analysis 41
7.5. Home Advantage Analysis 42
7.6. First Half vs Second Half Goals 42
7.7. Comeback & Clean Sheet Analysis 43
7.8. Coach Performance Analysis 43
7.9. Stadium Performance Analysis 44
7.10. Monthly Performance Trend 45
7.11. Second Half Specialist Teams 45
7.12. Coach Impact Score 46
7.13. Stadium Capacity Impact 46
7.14. Average Goals per Match by League 47
7.15. Competitive Balance Analysis 47
7.16. Stadium Tier Performance Analysis 48
7.17. Second Half Goal Ratio Analysis 48
7.18. Away Match Coach Analysis 49
7.19. Lead Blowing Analysis 49

8. Power BI Dashboard (Bonus) 50
8.1. Dashboard Data Source 50
8.2. Dashboard Screenshot 51

9. Conclusion 54

GitHub Repo: *https://github.com/Mohammed-3tef/European_Football_Leagues_DataWarehouse*

1. Source Database

The source system represents an OLTP database for managing European football league data.
The database stores operational data related to teams, matches, leagues, coaches, stadiums, players, and standings.

1.1. Source Database Overview

1.1.1. Source Tables

Table Name	Description
leagues	Stores league information such as name and country
seasons	Stores season information for each league
teams	Stores football team details
stadiums	Stores stadium information and capacity
coaches	Stores coach information
players	Stores player information and team assignment
matches	Stores match schedules and results
scores	Stores match score details
standings	Stores league standings and team statistics
referees	Stores referee information

1.1.2. Relationships Table

Parent Table	Child Table	Relationship
leagues	seasons	One League → Many Seasons
leagues	teams	One League → Many Teams
stadiums	teams	One Stadium → Many Teams
coaches	teams	One Coach → Many Teams
teams	players	One Team → Many Players
seasons	matches	One Season → Many Matches
teams	matches	One Team → Many Matches
matches	scores	One Match → One Score Record
seasons	standings	One Season → Many Standings
teams	standings	One Team → Many Standing Records
leagues	standings	One League → Many Standings

1.2. Source DDL

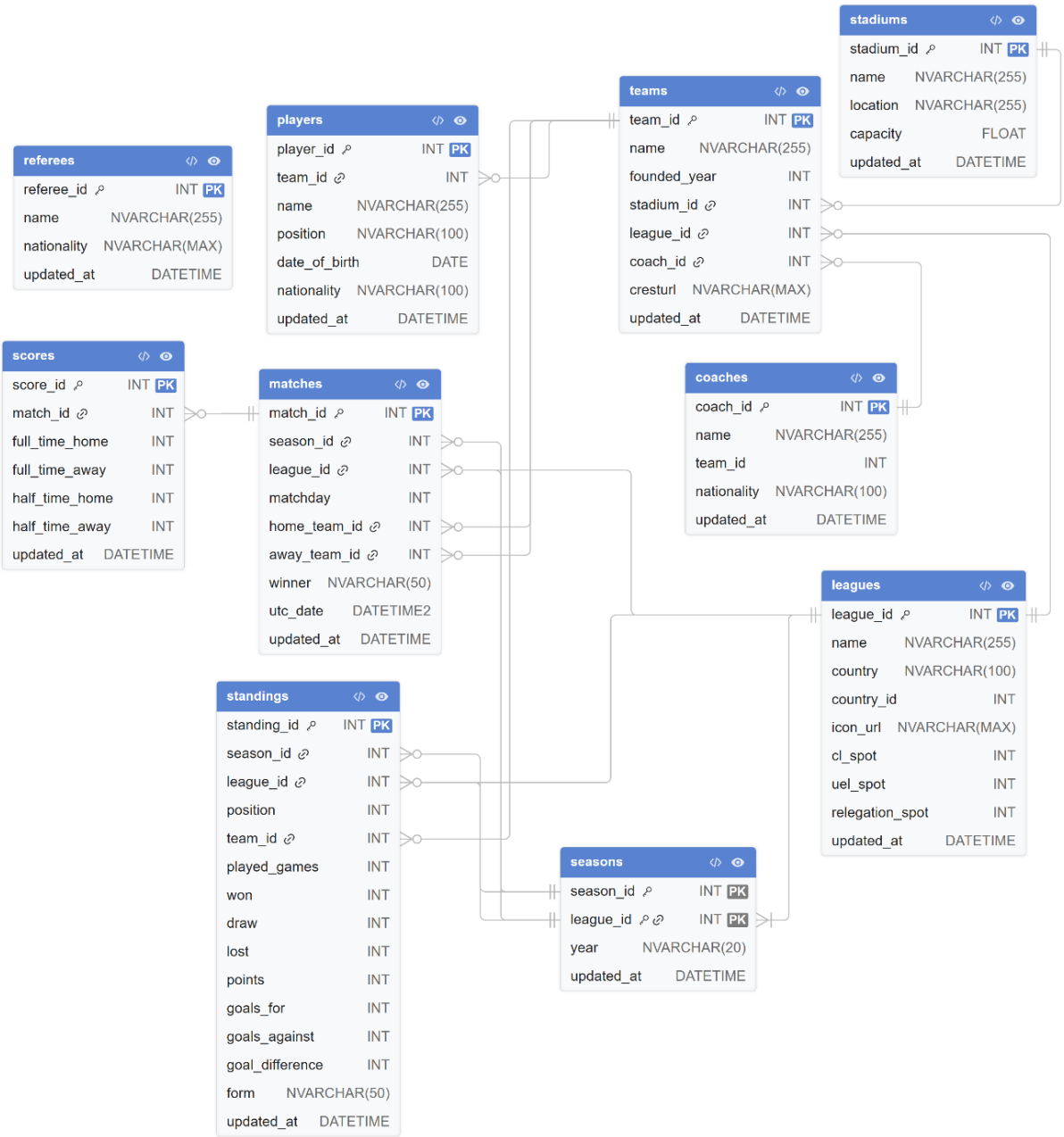
```sql
CREATE DATABASE European_Football_Leagues;
GO

ALTER DATABASE European_Football_Leagues
COLLATE SQL_Latin1_General_CP1_CI_AS;
GO

USE European_Football_Leagues;
GO

CREATE TABLE stadiums (
    stadium_id INT PRIMARY KEY,
    name NVARCHAR(255),
    location NVARCHAR(255),
    capacity FLOAT,
    updated_at DATETIME DEFAULT ('2024-08-01')
);

CREATE TABLE leagues (
    league_id INT PRIMARY KEY,
    name NVARCHAR(255),
    country NVARCHAR(100),
    country_id INT,
    icon_url NVARCHAR(MAX),
    cl_spot INT,
    uel_spot INT,
    relegation_spot INT,
    updated_at DATETIME DEFAULT ('2024-08-01')
);

CREATE TABLE coaches (
    coach_id INT PRIMARY KEY,
    name NVARCHAR(255),
    team_id INT NULL,
    nationality NVARCHAR(100),
    updated_at DATETIME DEFAULT ('2024-08-01')
);

CREATE TABLE referees (
    referee_id INT PRIMARY KEY,
    name NVARCHAR(255),
    nationality NVARCHAR(MAX),
    updated_at DATETIME DEFAULT ('2024-08-01')
);
```

```sql
CREATE TABLE teams (
    team_id INT PRIMARY KEY,
    name NVARCHAR(255),
    founded_year INT,
    stadium_id INT,
    league_id INT,
    coach_id INT,
    cresturl NVARCHAR(MAX),
    updated_at DATETIME DEFAULT ('2024-08-01'),

    CONSTRAINT FK_teams_stadium
        FOREIGN KEY (stadium_id) REFERENCES stadiums(stadium_id),

    CONSTRAINT FK_teams_league
        FOREIGN KEY (league_id) REFERENCES leagues(league_id),

    CONSTRAINT FK_teams_coach
        FOREIGN KEY (coach_id) REFERENCES coaches(coach_id)
);

CREATE TABLE players (
    player_id INT PRIMARY KEY,
    team_id INT,
    name NVARCHAR(255),
    position NVARCHAR(100),
    date_of_birth DATE,
    nationality NVARCHAR(100),
    updated_at DATETIME DEFAULT ('2024-08-01'),

    CONSTRAINT FK_players_team
        FOREIGN KEY (team_id) REFERENCES teams(team_id)
);

CREATE TABLE seasons (
    season_id INT,
    league_id INT,
    year NVARCHAR(20),
    updated_at DATETIME DEFAULT ('2024-08-01'),

    CONSTRAINT PK_seasons
        PRIMARY KEY (season_id, league_id),

    CONSTRAINT FK_seasons_league
        FOREIGN KEY (league_id) REFERENCES leagues(league_id)
);

CREATE TABLE scores (
    score_id INT PRIMARY KEY,
    match_id INT,
    full_time_home INT,   -- goals scored by home team at full time
    full_time_away INT,   -- goals scored by away team at full time
    half_time_home INT,   -- goals scored by home team at half time
    half_time_away INT,   -- goals scored by away team at half time
    updated_at DATETIME DEFAULT ('2024-08-01'),

    CONSTRAINT FK_scores_match
        FOREIGN KEY (match_id) REFERENCES matches(match_id)
);
```

```sql
CREATE TABLE matches (
    match_id INT PRIMARY KEY,
    season_id INT,
    league_id INT,
    matchday INT,
    home_team_id INT,
    away_team_id INT,
    winner NVARCHAR(50),
    utc_date DATETIME2,
    updated_at DATETIME DEFAULT ('2024-08-01'),

    CONSTRAINT FK_matches_season
        FOREIGN KEY (season_id, league_id)
        REFERENCES seasons(season_id, league_id),

    CONSTRAINT FK_matches_league
        FOREIGN KEY (league_id) REFERENCES leagues(league_id),

    CONSTRAINT FK_matches_home_team
        FOREIGN KEY (home_team_id) REFERENCES teams(team_id),

    CONSTRAINT FK_matches_away_team
        FOREIGN KEY (away_team_id) REFERENCES teams(team_id)
);

CREATE TABLE standings (
    standing_id INT PRIMARY KEY,
    season_id INT,
    league_id INT,
    position INT,
    team_id INT,
    played_games INT,
    won INT,
    draw INT,
    lost INT,
    points INT,
    goals_for INT,
    goals_against INT,
    goal_difference INT,
    form NVARCHAR(50),
    updated_at DATETIME DEFAULT ('2024-08-01'),

    CONSTRAINT FK_standings_season
        FOREIGN KEY (season_id, league_id)
        REFERENCES seasons(season_id, league_id),

    CONSTRAINT FK_standings_league
        FOREIGN KEY (league_id) REFERENCES leagues(league_id),

    CONSTRAINT FK_standings_team
        FOREIGN KEY (team_id) REFERENCES teams(team_id)
);
```

Click Here to Download Source DML.

1.3. Source ERD



2. Staging Area

The staging area acts as an intermediate layer between the source system and the Data Warehouse. It is responsible for temporary storage, cleansing, transformation, and integration of extracted data before loading it into the dimensional model.

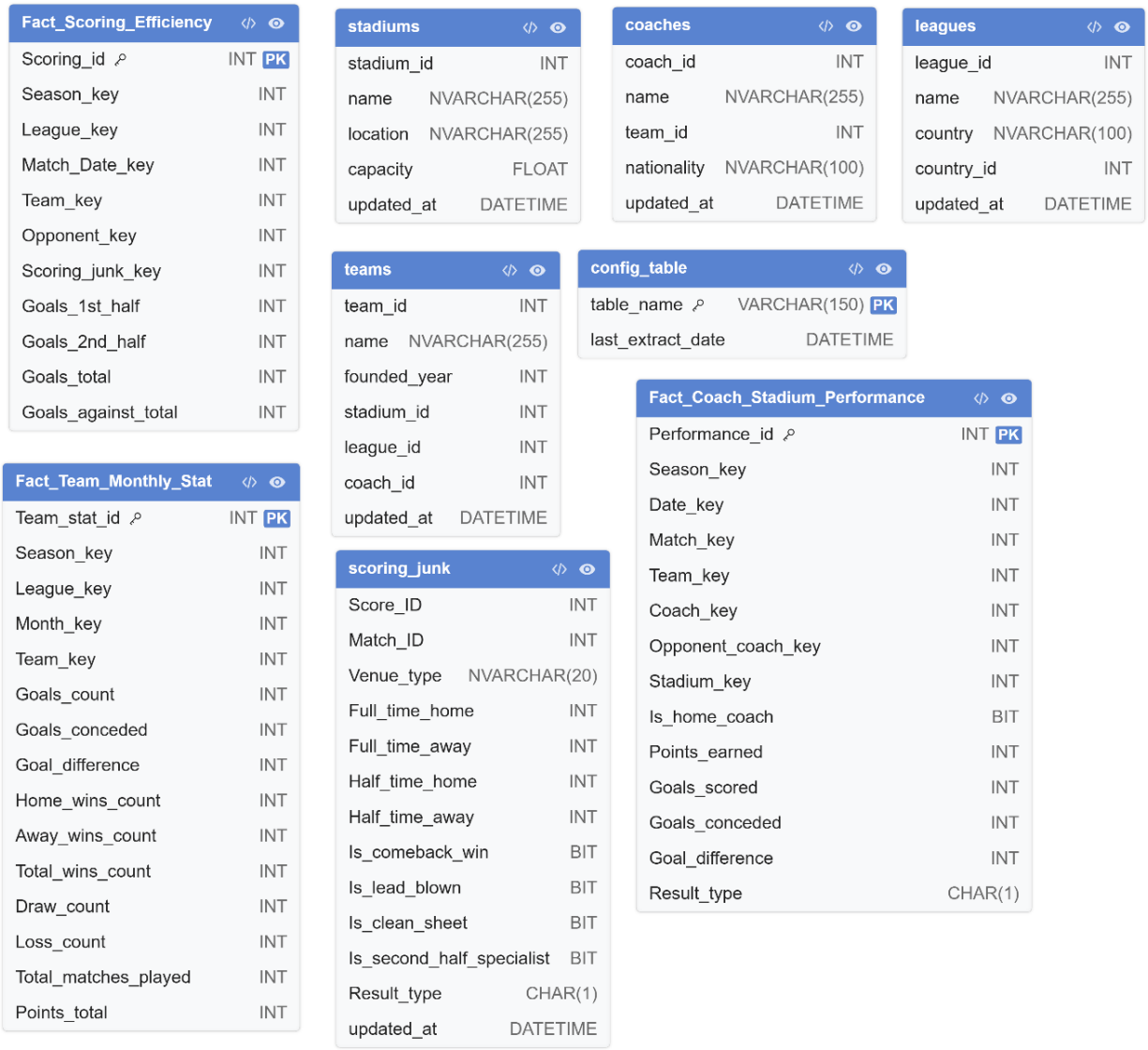
2.1. STG DDL

```sql
CREATE DATABASE European_Football_Leagues_DWH;
GO

ALTER DATABASE European_Football_Leagues_DWH
COLLATE SQL_Latin1_General_CP1_CI_AS;
GO

USE European_Football_Leagues_DWH;
GO

CREATE SCHEMA STG;
GO

CREATE TABLE STG.stadiums (
    stadium_id INT,
    name NVARCHAR(255),
    location NVARCHAR(255),
    capacity FLOAT,
    updated_at DATETIME
);

CREATE TABLE STG.leagues (
    league_id INT,
    name NVARCHAR(255),
    country NVARCHAR(100),
    country_id INT,
    updated_at DATETIME
);

CREATE TABLE STG.coaches (
    coach_id INT,
    name NVARCHAR(255),
    team_id INT NULL,
    nationality NVARCHAR(100),
    updated_at DATETIME
);

CREATE TABLE STG.teams (
    team_id INT,
    name NVARCHAR(255),
    founded_year INT,
    stadium_id INT,
    league_id INT,
    coach_id INT,
    updated_at DATETIME
);
```

```sql
CREATE TABLE STG.scoring_junk (
    Score_ID INT,
    Match_ID INT,

    Venue_type NVARCHAR(20),
    Full_time_home INT,  -- goals scored by home team at full time
    Full_time_away INT,  -- goals scored by away team at full time
    Half_time_home INT,  -- goals scored by home team at half time
    Half_time_away INT,  -- goals scored by away team at half time

    Is_comeback_win BIT DEFAULT 0,
    Is_lead_blown BIT DEFAULT 0,
    Is_clean_sheet BIT DEFAULT 0,
    Is_second_half_specialist BIT DEFAULT 0,
    Result_type CHAR(1),

    updated_at DATETIME
);

CREATE TABLE STG.Fact_Team_Monthly_Stat(
    Team_stat_id INT IDENTITY PRIMARY KEY,

    Season_key INT,  -- DEGNERATE DIM (20232024)
    League_key INT,
    Month_key INT,
    Team_key INT,

    Goals_count INT,
    Goals_conceded INT,
    Goal_difference INT,

    Home_wins_count INT,
    Away_wins_count INT,
    Total_wins_count INT,

    Draw_count INT,
    Loss_count INT,

    Total_matches_played INT,
    Points_total INT
);

CREATE TABLE STG.Fact_Scoring_Efficiency(
    Scoring_id INT IDENTITY PRIMARY KEY,

    Season_key INT, -- DEGNERATE DIM (20232024)
    League_key INT,
    Match_Date_key INT,
    Team_key INT,
    Opponent_key INT,

    Scoring_junk_key INT,

    Goals_1st_half INT,
    Goals_2nd_half INT,
    Goals_total INT,
    Goals_against_total INT
);
```

```sql
CREATE TABLE STG.Fact_Coach_Stadium_Performance(
    Performance_id INT IDENTITY PRIMARY KEY,

    Season_key INT, -- DEGNERATE DIM (20232024)
    Date_key INT,
    Match_key INT,
    Team_key INT,
    Coach_key INT,
    Opponent_coach_key INT,
    Stadium_key INT,

    Is_home_coach BIT,

    Points_earned INT,

    Goals_scored INT,
    Goals_conceded INT,
    Goal_difference INT,

    Result_type CHAR(1)  -- [W, D, L]
);


CREATE TABLE STG.config_table(
    table_name VARCHAR(150) PRIMARY KEY,
    last_extract_date DATETIME
);

INSERT INTO STG.config_table VALUES
('stadiums', '1900-01-01'),
('leagues', '1900-01-01'),
('teams', '1900-01-01'),
('coaches', '1900-01-01'),
('scores', '1900-01-01');
```

2.2. STG ERD


Fact_Scoring_Efficiency
Scoring_id INT PK
Season_key INT
League_key INT
Match_Date_key INT
Team_key INT
Opponent_key INT
Scoring_junk_key INT
Goals_1st_half INT
Goals_2nd_half INT
Goals_total INT
Goals_against_total INT
stadiums
stadium_id INT
name NVARCHAR(255)
location NVARCHAR(255)
capacity FLOAT
updated_at DATETIME
coaches
coach_id INT
name NVARCHAR(255)
team_id INT
nationality NVARCHAR(100)
updated_at DATETIME
leagues
league_id INT
name NVARCHAR(255)
country NVARCHAR(100)
country_id INT
updated_at DATETIME
teams
team_id INT
name NVARCHAR(255)
founded_year INT
stadium_id INT
league_id INT
coach_id INT
updated_at DATETIME
config_table
table_name VARCHAR(150) PK
last_extract_date DATETIME
Fact_Team_Monthly_Stat
Team_stat_id INT PK
Season_key INT
League_key INT
Month_key INT
Team_key INT
Goals_count INT
Goals_conceded INT
Goal_difference INT
Home_wins_count INT
Away_wins_count INT
Total_wins_count INT
Draw_count INT
Loss_count INT
Total_matches_played INT
Points_total INT
scoring_junk
Score_ID INT
Match_ID INT
Venue_type NVARCHAR(20)
Full_time_home INT
Full_time_away INT
Half_time_home INT
Half_time_away INT
Is_comeback_win BIT
Is_lead_blown BIT
Is_clean_sheet BIT
Is_second_half_specialist BIT
Result_type CHAR(1)
updated_at DATETIME
Fact_Coach_Stadium_Performance
Performance_id INT PK
Season_key INT
Date_key INT
Match_key INT
Team_key INT
Coach_key INT
Opponent_coach_key INT
Stadium_key INT
Is_home_coach BIT
Points_earned INT
Goals_scored INT
Goals_conceded INT
Goal_difference INT
Result_type CHAR(1)

3. Data Warehouse

3.1. Dimensional Model

The data warehouse follows a **Star Schema** design where fact tables are connected to shared dimensions.
The dimensional model is optimized for analytical queries, KPI calculations, and Power BI reporting.

Dimension Tables

Dimension Table	Description	Type	Key
`Dim_Date`	Stores full calendar and time attributes for analysis by date, month, quarter, and year	Time Dimension	`DateKey`
`Dim_League`	Stores league information such as league name and country	Standard Dimension	`League_key`
`Dim_Team`	Stores team details with Slowly Changing Dimension (SCD Type 2) support	SCD Type 2	`Team_key`
`Dim_Stadium`	Stores stadium information with historical tracking support	SCD Type 2	`Stadium_key`
`Dim_Coach`	Stores coach details and nationality information	Standard Dimension	`Coach_key`
`Dim_Scoring_Junk`	Junk dimension that stores match behavior flags and categorical indicators	Junk Dimension	`Scoring_junk_key`

Fact Tables

Fact Table	Grain	Description
`Fact_Team_Monthly_Stat`	One Team + One League + One Month	Stores aggregated monthly performance statistics for teams
`Fact_Scoring_Efficiency`	One Team + One Match	Stores detailed scoring and efficiency metrics for every match
`Fact_Coach_Stadium_Performance`	One Team + One Match	Stores coach and stadium related performance analytics

3.2. Fact Tables

The Data Warehouse contains multiple fact tables designed to support different analytical perspectives related to football performance and competition analysis.

3.2.1. `Fact_Team_Monthly_Stats`

Section	Details
Grain	One Team + One League + One Month
Source Tables	`matches, scores`
Dimensions Used	`Dim_Team, Dim_League, Dim_Date`
Main Purpose	Monthly team performance analytics

Foreign Keys

Fact Column	Dimension	Lookup Condition
`Team_Key`	`Dim_Team`	Team_ID = source team_id
`League_Key`	`Dim_League`	League_ID = source league_id
`Season_Key`	`Degenerate Dimension`	season_year (e.g. 20232024)
`Month_Key`	`Dim_Date`	First day of month from utc_date

Measures

Measure	Source	Calculation
`Goals_Count`	`scores`	`SUM(full_time_home / full_time_away)`
`Home_Wins_Count`	`matches`	`winner = 'HOME_TEAM'`
`Away_Wins_Count`	`matches`	`winner = 'AWAY_TEAM'`
`Total_Wins_Count`	`Calculated`	`Home + Away wins`
`Total_Matches_Played`	`matches`	`COUNT(match_id)`
`Draw_Count`	`matches`	`winner = 'DRAW'`
`Loss_Count`	`Calculated`	`Matches - Wins - Draws`
`Goals_Conceded`	`scores`	`Opponent goals`
`Goal_Difference`	`Calculated`	`Goals_Count - Goals_Conceded`
`Points_Total`	`Calculated`	`Win=3, Draw=1, Loss=0`

ETL Flow

Step	Action
1	Read matches + scores
2	Generate Home & Away rows
3	Calculate KPIs
4	Group by Team + League + Month + Season
5	Lookup Dimension Keys
6	Insert into `Fact_Team_Monthly_Stats`

3.2.2. `Fact_Scoring_Efficiency`

Section	Details
Grain	One Team + One Match
Source Tables	`matches, scores`
Dimensions Used	`Dim_Team, Dim_League, Dim_Date, Dim_Scoring_Junk`
Main Purpose	Scoring efficiency & match analysis

Foreign Keys

Fact Column	Dimension	Lookup Condition
`Team_Key`	`Dim_Team`	Team_ID = current team
`Opponent_Key`	`Dim_Team`	Team_ID = opponent team
`League_Key`	`Dim_League`	League_ID = source league_id
`Season_Key`	`Degenerate Dimension`	season_year (e.g. 20232024)
`Season_Key`	`Dim_Season`	Season_ID = source season_id
`Match_Date_Key`	`Dim_Date`	utc_date
`Scoring_Junk_Key`	`Dim_Scoring_Junk`	Matching junk flag combination

Measures

Measure	Source	Calculation
`Goals_1st_Half`	`scores`	`half_time_home / half_time_away`
`Goals_2nd_Half`	`scores`	`full_time - half_time`
`Goals_Total`	`scores`	`full_time_home / full_time_away`
`Goals_Against_Total`	`scores`	`Opponent full time goals`

Junk Dimension Flags

Flag	Logic
`Venue_Type`	Home / Away
`Is_Comeback_Win`	Losing halftime AND won fulltime
`Is_Lead_Blown`	Winning halftime AND did not win final
`Is_Clean_Sheet`	Opponent goals = 0
`Is_Second_Half_Specialist`	2nd half goals > 1st half goals
`Result_Type`	W / D / L

ETL Flow

Step	Action
1	Read matches + scores
2	Create Home Team row
3	Create Away Team row
4	Calculate goals metrics
5	Calculate junk flags
6	Lookup dimension surrogate keys
7	Insert into `Fact_Scoring_Efficiency`

3.2.3. `Fact_Coach_Stadium_Performance`

Section	Details
Grain	One Team + One Match
Source Tables	`matches, scores, teams, stadiums, coach history`
Dimensions Used	`Dim_Team, Dim_Coach, Dim_Stadium, Dim_League, Dim_Date`
Main Purpose	Coach and stadium performance analytics

Foreign Keys

Fact Column	Dimension	Lookup Condition
`Match_Key`	`Degenerate Dimension`	Source Match_id
`Team_Key`	`Dim_Team`	Team_ID = current team
`Coach_Key`	`Dim_Coach`	Active coach during match date
`Opponent_Coach_Key`	`Dim_Coach`	Opponent active coach
`Stadium_Key`	`Dim_Stadium`	Stadium_ID = source stadium_id
`League_Key`	`Dim_League`	League_ID = source league_id
`Season_Key`	`Degenerate Dimension`	season_year (e.g. 20232024)
`Date_Key`	`Dim_Date`	utc_date

Measures

Measure	Source	Calculation
`Points_Earned`	`Calculated`	`Win=3, Draw=1, Loss=0`
`Goals_Scored`	`scores`	`Team full time goals`
`Goals_Conceded`	`scores`	`Opponent full time goals`
`Goal_Difference`	`Calculated`	`Goals_Scored - Goals_Conceded`
`Result_type`	`Calculated`	`W / D / L`
`Is_Home_Coach`	`matches`	`Team = home_team_id`

ETL Flow

Step	Action
1	Read matches + scores
2	Find active coach from bridge table
3	Generate Home Team row
4	Generate Away Team row
5	Calculate performance KPIs
6	Lookup surrogate keys
7	Insert into `Fact_Coach_Stadium_Performance`

3.3. DWH DDL

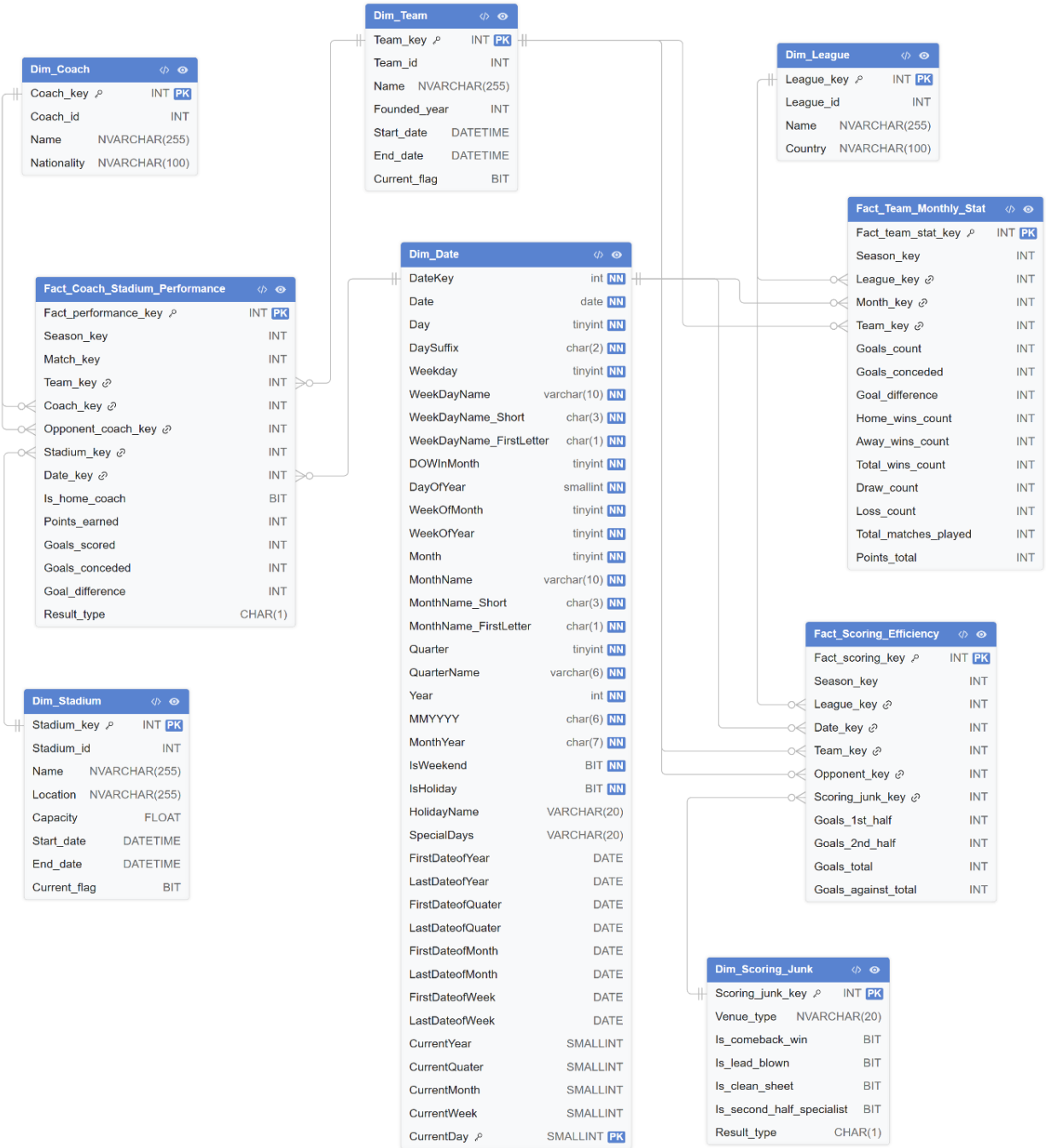
```
USE European_Football_Leagues_DWH;
GO

CREATE SCHEMA DWH;
GO

CREATE TABLE DWH.Dim_Date (
   [DateKey] [int] NOT NULL,
   [Date] [date] NOT NULL,
   [Day] [tinyint] NOT NULL,
   [DaySuffix] [char](2) NOT NULL,
   [Weekday] [tinyint] NOT NULL,
   [WeekDayName] [varchar](10) NOT NULL,
   [WeekDayName_Short] [char](3) NOT NULL,
   [WeekDayName_FirstLetter] [char](1) NOT NULL,
   [DOWInMonth] [tinyint] NOT NULL,
   [DayOfYear] [smallint] NOT NULL,
   [WeekOfMonth] [tinyint] NOT NULL,
   [WeekOfYear] [tinyint] NOT NULL,
   [Month] [tinyint] NOT NULL,
   [MonthName] [varchar](10) NOT NULL,
   [MonthName_Short] [char](3) NOT NULL,
   [MonthName_FirstLetter] [char](1) NOT NULL,
   [Quarter] [tinyint] NOT NULL,
   [QuarterName] [varchar](6) NOT NULL,
   [Year] [int] NOT NULL,
   [MMYYYY] [char](6) NOT NULL,
   [MonthYear] [char](7) NOT NULL,
   [IsWeekend] BIT NOT NULL,
   [IsHoliday] BIT NOT NULL,
   [HolidayName] VARCHAR(20) NULL,
   [SpecialDays] VARCHAR(20) NULL,
   [FirstDateofYear] DATE NULL,
   [LastDateofYear] DATE NULL,
   [FirstDateofQuater] DATE NULL,
   [LastDateofQuater] DATE NULL,
   [FirstDateofMonth] DATE NULL,
   [LastDateofMonth] DATE NULL,
   [FirstDateofWeek] DATE NULL,
   [LastDateofWeek] DATE NULL,
   [CurrentYear] SMALLINT NULL,
   [CurrentQuater] SMALLINT NULL,
   [CurrentMonth] SMALLINT NULL,
   [CurrentWeek] SMALLINT NULL,
   [CurrentDay] SMALLINT NULL,
   PRIMARY KEY CLUSTERED ([DateKey] ASC)
);
```

```sql
CREATE TABLE DWH.Dim_League (
    League_key INT IDENTITY PRIMARY KEY,
    League_id INT,
    Name NVARCHAR(255),
    Country NVARCHAR(100)
);

CREATE TABLE DWH.Dim_Team (
    Team_key INT IDENTITY PRIMARY KEY,
    Team_id INT,
    Name NVARCHAR(255),
    Founded_year INT,
    Start_date DATETIME,
    End_date DATETIME,
    Current_flag BIT
);

CREATE TABLE DWH.Dim_Stadium (
    Stadium_key INT IDENTITY PRIMARY KEY,
    Stadium_id INT,
    Name NVARCHAR(255),
    Location NVARCHAR(255),
    Capacity FLOAT,
    Start_date DATETIME,
    End_date DATETIME,
    Current_flag BIT
);

CREATE TABLE DWH.Dim_Coach (
    Coach_key INT IDENTITY PRIMARY KEY,
    Coach_id INT,
    Name NVARCHAR(255),
    Nationality NVARCHAR(100)
);

CREATE TABLE DWH.Dim_Scoring_Junk (
    Scoring_junk_key INT IDENTITY PRIMARY KEY,
    Venue_type NVARCHAR(20),
    Is_comeback_win BIT,
    Is_lead_blown BIT,
    Is_clean_sheet BIT,
    Is_second_half_specialist BIT,
    Result_type CHAR(1)
);
```

```sql
CREATE TABLE DWH.Fact_Team_Monthly_Stat(
    Fact_team_stat_key INT IDENTITY PRIMARY KEY,

    Season_key INT,  -- DEGNERATE DIM (20232024)
    League_key INT,
    Month_key INT,
    Team_key INT,

    Goals_count INT,
    Goals_conceded INT,
    Goal_difference INT,

    Home_wins_count INT,
    Away_wins_count INT,
    Total_wins_count INT,

    Draw_count INT,
    Loss_count INT,

    Total_matches_played INT,
    Points_total INT,

    FOREIGN KEY (Month_key) REFERENCES DWH.Dim_Date(DateKey),
    FOREIGN KEY (League_key) REFERENCES DWH.Dim_League(League_key),
    FOREIGN KEY (Team_key) REFERENCES DWH.Dim_Team(Team_key)
);

CREATE TABLE DWH.Fact_Scoring_Efficiency(
    Fact_scoring_key INT IDENTITY PRIMARY KEY,

    Season_key INT, -- DEGNERATE DIM (20232024)
    League_key INT,
    Date_key INT,
    Team_key INT,
    Opponent_key INT,

    Scoring_junk_key INT,

    Goals_1st_half INT,
    Goals_2nd_half INT,
    Goals_total INT,
    Goals_against_total INT,

    FOREIGN KEY (Date_key) REFERENCES DWH.Dim_Date(DateKey),
    FOREIGN KEY (League_key) REFERENCES DWH.Dim_League(League_key),
    FOREIGN KEY (Team_key) REFERENCES DWH.Dim_Team(Team_key),
    FOREIGN KEY (Opponent_key) REFERENCES DWH.Dim_Team(Team_key),
    FOREIGN KEY (Scoring_junk_key) REFERENCES DWH.Dim_Scoring_Junk(Scoring_junk_key)
);
```

```sql
CREATE TABLE DWH.Fact_Coach_Stadium_Performance(
    Fact_performance_key INT IDENTITY PRIMARY KEY,

    Season_key INT, -- DEGNERATE DIM (20232024)
    Match_key INT,
    Team_key INT,
    Coach_key INT,
    Opponent_coach_key INT,
    Stadium_key INT,
    Date_key INT,

    Is_home_coach BIT,

    Points_earned INT,

    Goals_scored INT,
    Goals_conceded INT,
    Goal_difference INT,

    Result_type CHAR(1), -- [W, D, L]

    FOREIGN KEY (Date_key) REFERENCES DWH.Dim_Date(DateKey),
    FOREIGN KEY (Team_key) REFERENCES DWH.Dim_Team(Team_key),
    FOREIGN KEY (Coach_key) REFERENCES DWH.Dim_Coach(Coach_key),
    FOREIGN KEY (Opponent_coach_key) REFERENCES DWH.Dim_Coach(Coach_key),
    FOREIGN KEY (Stadium_key) REFERENCES DWH.Dim_Stadium(Stadium_key)
);
```

3.4. DWH ERD


Dim_Coach
Coach_key INT PK
Coach_id INT
Name NVARCHAR(255)
Nationality NVARCHAR(100)
Dim_Team
Team_key INT PK
Team_id INT
Name NVARCHAR(255)
Founded_year INT
Start_date DATETIME
End_date DATETIME
Current_flag BIT
Dim_League
League_key INT PK
League_id INT
Name NVARCHAR(255)
Country NVARCHAR(100)
Fact_Team_Monthly_Stat
Fact_team_stat_key INT PK
Season_key INT
League_key INT
Month_key INT
Team_key INT
Goals_count INT
Goals_conceded INT
Goal_difference INT
Home_wins_count INT
Away_wins_count INT
Total_wins_count INT
Draw_count INT
Loss_count INT
Total_matches_played INT
Points_total INT
Dim_Date
DateKey int NN
Date date NN
Day tinyint NN
DaySuffix char(2) NN
Weekday tinyint NN
WeekDayName varchar(10) NN
WeekDayName_Short char(3) NN
WeekDayName_FirstLetter char(1) NN
DOWInMonth tinyint NN
DayOfYear smallint NN
WeekOfMonth tinyint NN
WeekOfYear tinyint NN
Month tinyint NN
MonthName varchar(10) NN
MonthName_Short char(3) NN
MonthName_FirstLetter char(1) NN
Quarter tinyint NN
QuarterName varchar(6) NN
Year int NN
MMYYYY char(6) NN
MonthYear char(7) NN
IsWeekend BIT NN
IsHoliday BIT NN
HolidayName VARCHAR(20)
SpecialDays VARCHAR(20)
FirstDateofYear DATE
LastDateofYear DATE
FirstDateofQuater DATE
LastDateofQuater DATE
FirstDateofMonth DATE
LastDateofMonth DATE
FirstDateofWeek DATE
LastDateofWeek DATE
CurrentYear SMALLINT
CurrentQuater SMALLINT
CurrentMonth SMALLINT
CurrentWeek SMALLINT
CurrentDay SMALLINT PK
Fact_Coach_Stadium_Performance
Fact_performance_key INT PK
Season_key INT
Match_key INT
Team_key INT
Coach_key INT
Opponent_coach_key INT
Stadium_key INT
Date_key INT
Is_home_coach BIT
Points_earned INT
Goals_scored INT
Goals_conceded INT
Goal_difference INT
Result_type CHAR(1)
Fact_Scoring_Efficiency
Fact_scoring_key INT PK
Season_key INT
League_key INT
Date_key INT
Team_key INT
Opponent_key INT
Scoring_junk_key INT
Goals_1st_half INT
Goals_2nd_half INT
Goals_total INT
Goals_against_total INT
Dim_Stadium
Stadium_key INT PK
Stadium_id INT
Name NVARCHAR(255)
Location NVARCHAR(255)
Capacity FLOAT
Start_date DATETIME
End_date DATETIME
Current_flag BIT
Dim_Scoring_Junk
Scoring_junk_key INT PK
Venue_type NVARCHAR(20)
Is_comeback_win BIT
Is_lead_blown BIT
Is_clean_sheet BIT
Is_second_half_specialist BIT
Result_type CHAR(1)

4. ETL Process

4.1. ETL Architecture

The ETL process follows a three-layer architecture:

1. Source Database (OLTP)
2. Staging Area (STG)
3. Data Warehouse (DWH)

The source system contains normalized operational football data such as teams, matches, leagues, coaches, and scores. Data is extracted from the source database and loaded into the staging area for cleaning and transformation before loading into the dimensional warehouse.

ETL Workflow

```
Source Database → STG Tables → Dimension Tables → Fact Tables → Power BI
```

The ETL pipeline was implemented using SQL Server Integration Services (SSIS).

4.2. Extraction Phase

During the extraction phase, data is collected from the source OLTP database.

Extraction Features

Feature	Description
Source System	`European_Football_Leagues`
Extraction Type	Incremental Load
Extraction Tool	SSIS
Change Tracking	`updated_at` column
Metadata Tracking	`STG.config_table`

Incremental Loading Strategy

The ETL process uses the `updated_at` column to extract only new or modified records.
The `STG.config_table` stores the last successful extraction date for each table.

Example:

```
SELECT *
FROM teams
WHERE updated_at > @LastExtractDate;
```

This strategy improves performance and reduces ETL execution time.

4.3. Transformation Phase

The transformation phase prepares data for analytical processing.

Main Transformations

Transformation	Description
Data Cleaning	Remove NULLs and invalid values
Data Standardization	Normalize text and formats
KPI Calculations	Calculate goals, points, wins, losses
Aggregation	Monthly team statistics aggregation
Derived Columns	Goal difference and result type
Lookup Operations	Generate surrogate keys
SCD Type 2	Historical tracking for dimensions
Junk Dimension Generation	Match behavior flags

4.4. Loading Phase

After transformations are completed, data is loaded into the warehouse.

Loading Order:

1. Dimention Tables
2. Junk Dimention
3. Fact Tables

Dimension Loading: Dimensions are loaded first to ensure surrogate keys are available for fact tables.

Fact Table Loading: Fact tables are populated after performing lookup operations to retrieve surrogate keys from dimensions.

Error Handling

- Lookup error redirection
- Duplicate prevention
- NULL handling
- Data validation
- SSIS logging

4.5. ETL Workflow Steps

1. Extract source data
2. Load data into STG tables
3. Clean and transform data
4. Populate dimension tables
5. Populate fact tables
6. Update config table

4.6. Incremental Load Process

The warehouse supports incremental refreshes.

Benefits

- Faster ETL execution
- Reduced server load
- Improved scalability
- Better scheduling efficiency

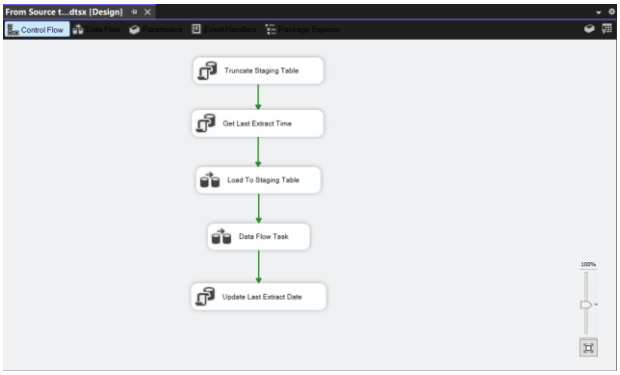
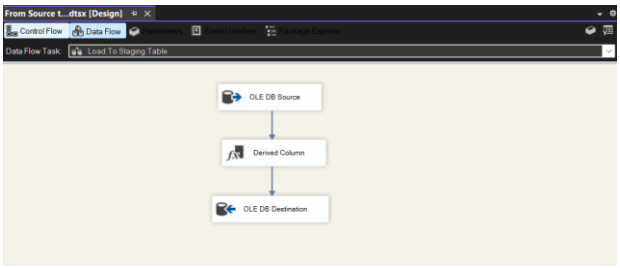
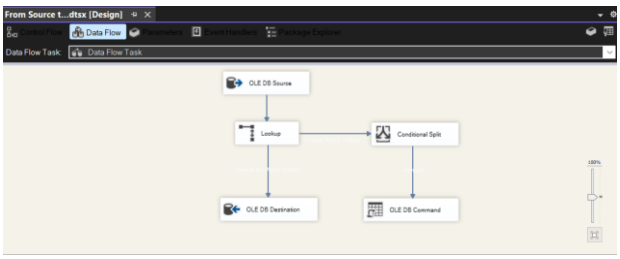
Incremental Load Logic

```
WHERE updated_at > last_extract_date
```

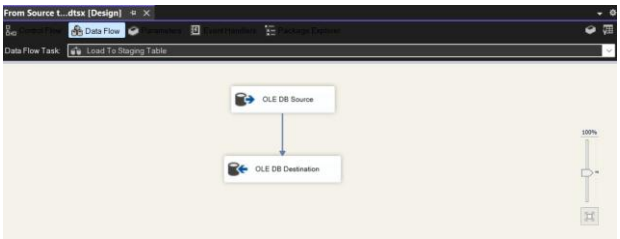
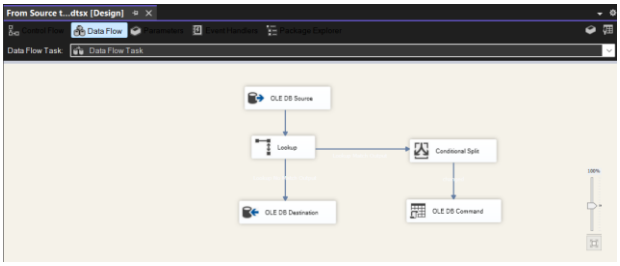
The extraction date is updated after each successful ETL execution.

4.7. SSIS Control / Data Flow Screenshots

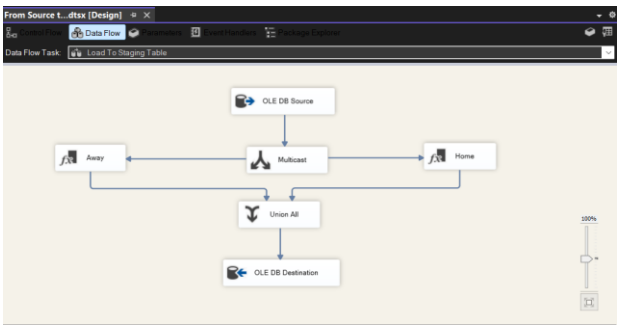
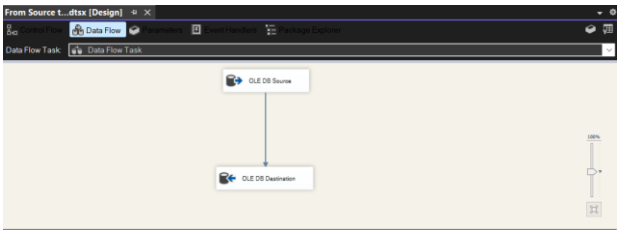
4.7.1. Load `Dim_Coach`

Control Flow	Data Flow
	 Load to STG  Load to DWH

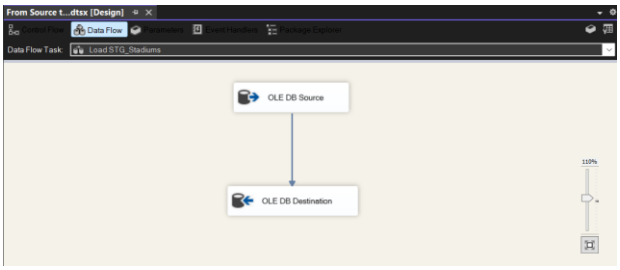
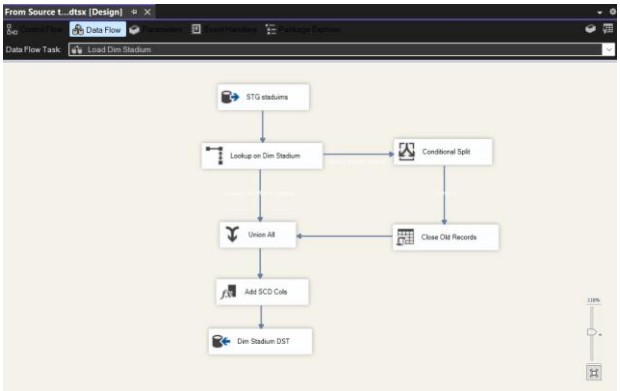
4.7.2. Load `Dim_Leagues`

Control Flow	Data Flow
	 Load to STG  Load to DWH

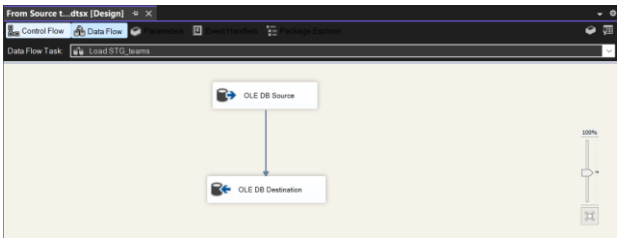
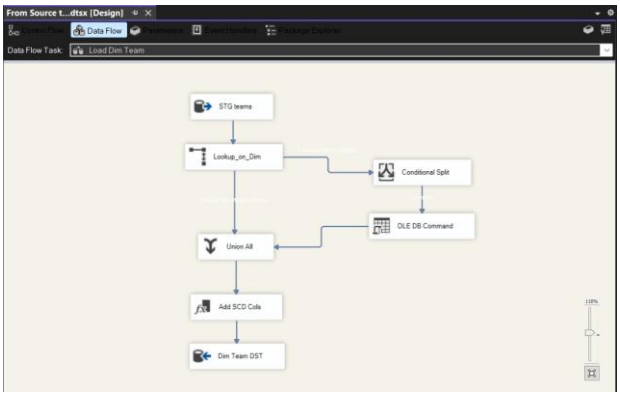
4.7.3. Load `Dim_Score_junk`

Control Flow	Data Flow
	 Load to STG  Load to DWH

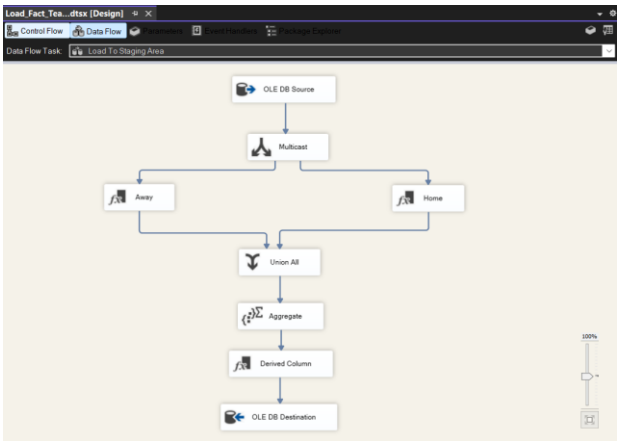
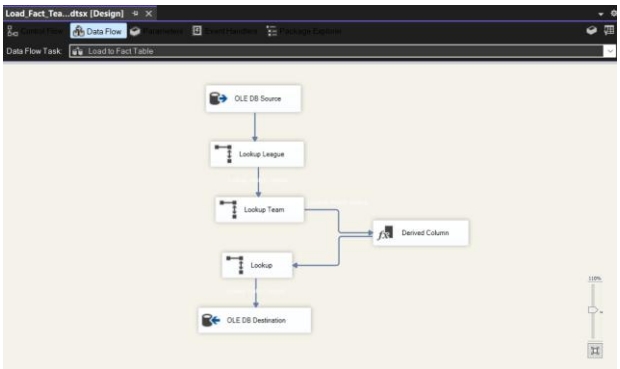
4.7.4. Load `Dim_Stadium`

Control Flow	Data Flow
	 Load to STG  Load to DWH

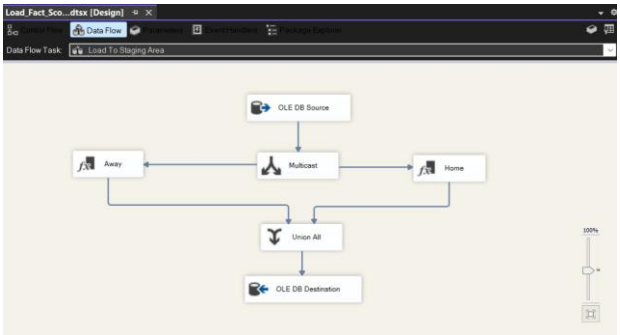
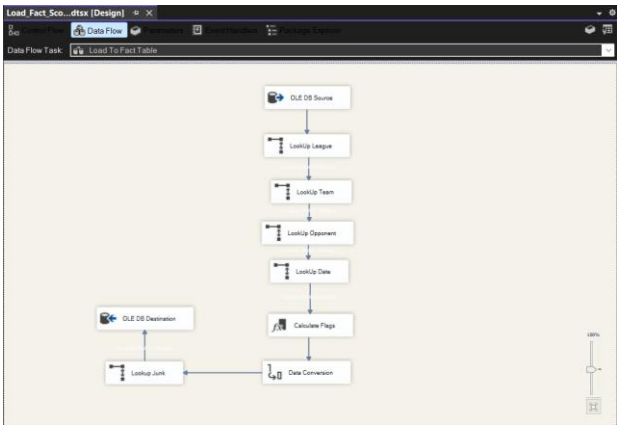
4.7.5. Load `Dim_Team`

Control Flow	Data Flow
	 Load to STG  Load to DWH

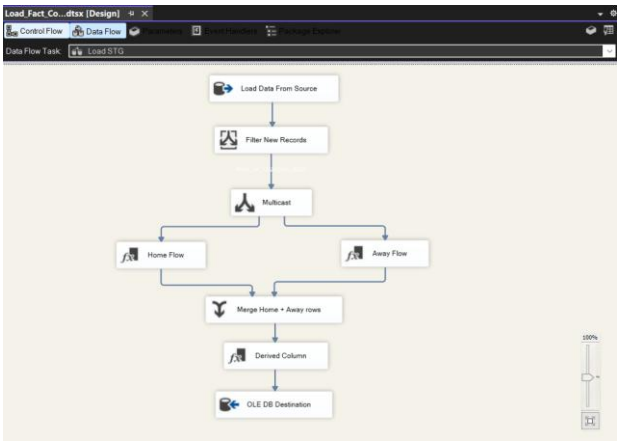
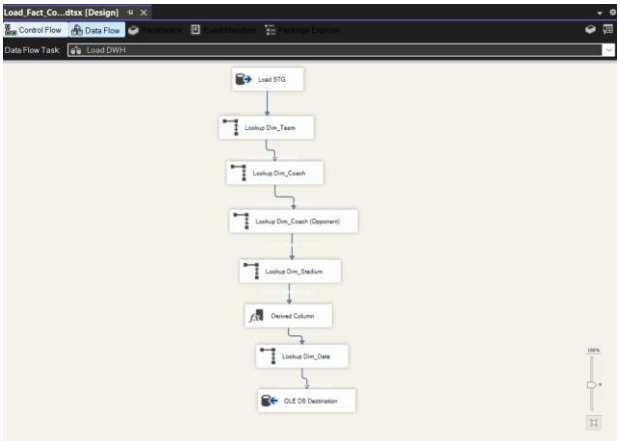
4.7.6. Load `Fact_Team_Monthly_Stat`

Control Flow	Data Flow
	 Load to STG  Load to DWH

4.7.7. Load `Fact_Scoring_Efficiency`

Control Flow	Data Flow
	 Load to STG  Load to DWH

4.7.8. Load `Fact_Coach_Stadium_Performance`

Control Flow	Data Flow
	 Load to STG  Load to DWH

5. KPIs

5.1. Team Performance KPIs (Season / League / Team Level)

KPI	Definition	Scope	Fact Table
1. Total Points	مجموع النقاط	Team + Season + League	`Fact_Team_Monthly_Stat`
2. Total Wins	عدد الانتصارات	Team + Season	`Fact_Team_Monthly_Stat`
3. Total Draws	عدد التعادلات	Team + Season	`Fact_Team_Monthly_Stat`
4. Total Losses	عدد الهزائم	Team + Season	`Fact_Team_Monthly_Stat`
5. Goals Scored	إجمالي الأهداف	Team + Season	`Both Facts`
6. Goals Conceded	أهداف مستقبلة	Team + Season	`Both Facts`
7. Goal Difference	فرق الأهداف	Team + Season	`Both Facts`
8. Matches Played	عدد المباريات	Team + Season	`Fact_Team_Monthly_Stat`
9. Win Rate	نسبة الفوز	Team + Season	`Derived`
10. Points per Match	متوسط النقاط	Team + Season	`Derived`

5.2. League Level KPIs (League + Season)

KPI	Definition	Scope	Fact Table
11. Total League Goals	إجمالي أهداف الدوري	League + Season	`Fact_Team_Monthly_Stat`
12. Average Goals per Match	متوسط الأهداف	League + Season	`Derived`
13. Total Matches	عدد مباريات الدوري	League + Season	`Fact_Team_Monthly_Stat`
14. Average Points per Team	متوسط النقاط لكل فريق	League + Season	`Derived`
15. Competitive Balance Index	توازن الدوري	League + Season	`Derived`
16. Home Advantage Ratio	أفضلية الأرض	League + Season	`Fact_Team_Monthly_Stat`

5.3. Match Performance KPIs (Match + Date Level)

KPI	Definition	Scope	Fact Table
17. Total Goals per Match	أهداف المباراة	Match + Date	`Fact_Scoring_Efficiency`
18. First Half Goals	أهداف الشوط الأول	Match	`Fact_Scoring_Efficiency`
19. Second Half Goals	أهداف الشوط الثاني	Match	`Fact_Scoring_Efficiency`
20. Comeback Matches Count	مباريات العودة	Season + Date	`Dim_Scoring_Junk`
21. Clean Sheet Matches	مباريات بدون أهداف	Season	`Dim_Scoring_Junk`
22. Draw Rate	نسبة التعادل	Season + League	`Fact_Team_Monthly_Stat`

5.4. Team Form & Momentum KPIs

KPI	Definition	Scope	Fact Table
23. Last 5 Matches Form Score	آخر 5 مباريات	Team + Season	`standings.form`
24. Winning Streak	أطول سلسلة انتصارات	Team + Season	`Derived`
25. Losing Streak	أطول سلسلة خسائر	Team + Season	`Derived`
26. Monthly Performance Trend	الأداء الشهري	Team + Season + Month	`Fact_Team_Monthly_Stat`

5.5. Coach & Stadium KPIs

KPI	Definition	Scope	Fact Table
27. Coach Win Rate	نسبة فوز المدرب	Coach + Season	`Fact_Coach_Stadium_Performance`
28. Coach Points per Match	نقاط لكل مباراة	Coach + Season	`Fact_Coach_Stadium_Performance`
29. Home Stadium Win Rate	نسبة الفوز في ملعب الفريق	Stadium + Team	`Fact_Coach_Stadium_Performance`
30. Coach Impact Score	تأثير المدرب على الأداء	Coach + Team + Season	`Derived`

6. SSIS Packages & Scheduling Screenshots

6.1. Schedule Packages Screenshots

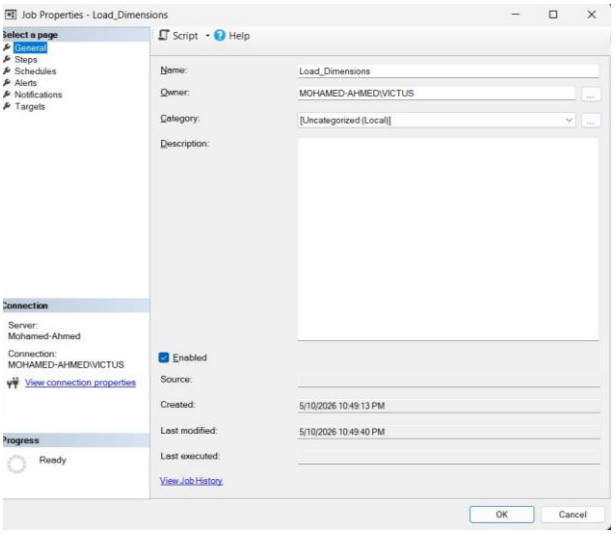


Fig. 6.1.1

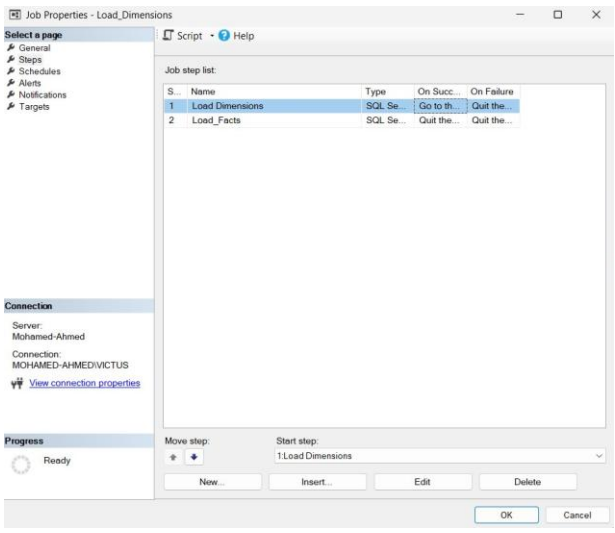


Fig. 6.1.2

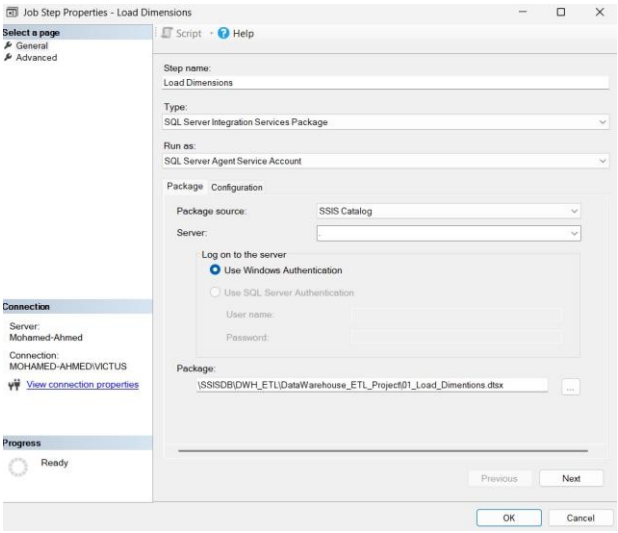


Fig. 6.1.3

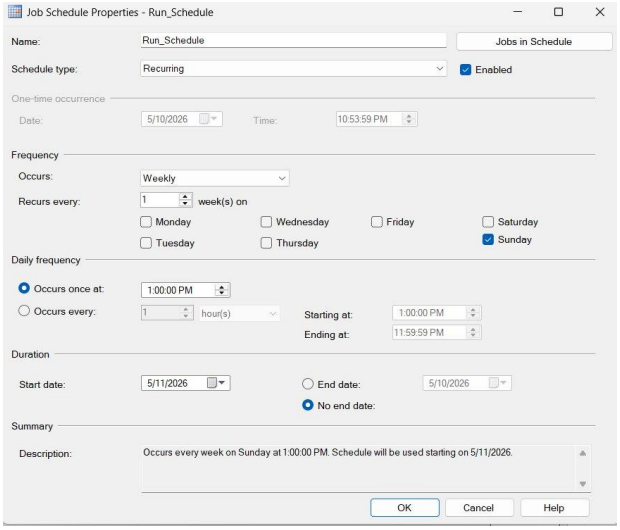


Fig. 6.1.4

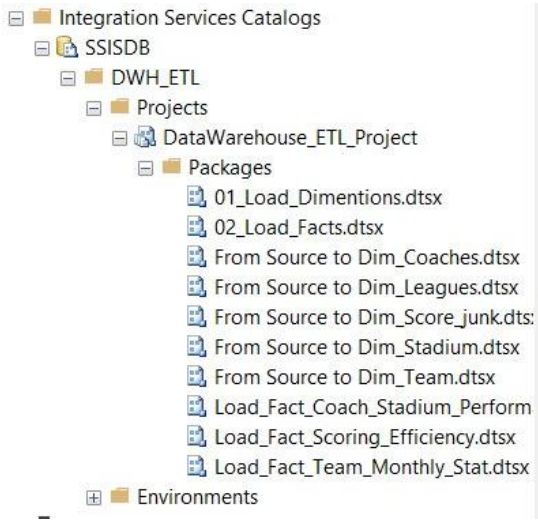


Fig. 6.1.5

6.2. After First Run Screenshots

Overview

on Mohamed-Ahmed at 5/10/2026 10:18:07 PM

This report provides an overview of the package tasks and parameters, including execution or validation information.

View Messages

View Performance

Execution Information

Operation ID	2
Package	DWH_ETL\DataWarehouse_ETL_Project\01_Load_Dimentions.dtsx
Environment	-
Status	Succeeded
Machine	Mohamed-Ahmed

Execution Overview

Filter: Result: All; (3 more)

Result	Duration (sec)	Package Name	Task Name	Execution Path
Succeeded	1419.422	01_Load_Dimentions.dtsx	01_Load_Dimentions	\01_Load_Dimentions
Succeeded	1419.313	01_Load_Dimentions.dtsx	Load_Dimentions	\01_Load_Dimentions \Load_Dimentions
Succeeded	303.578	01_Load_Dimentions.dtsx	Load_from_source_to_Dim_Coaches	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Coaches
Succeeded	303.282	01_Load_Dimentions.dtsx	Load_from_source_to_Dim_Leagues	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Leagues
Succeeded	239.047	01_Load_Dimentions.dtsx	Load_from_source_to_Dim_Score_Junk	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Score_Junk
Succeeded	286.625	01_Load_Dimentions.dtsx	Load_from_source_to_Dim_Stadium	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Stadium
Succeeded	286.781	01_Load_Dimentions.dtsx	Load_from_source_to_Dim_Team	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Team
Succeeded	171.547	From Source to Dim_Coaches.dtsx	Package1	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Coaches\Package1

1 / 5 Use scroll wheel or PAGE UP & PAGE DOWN to move through pages.

Fig. 6.2.1

SQL Server

Duration (sec)	1424.405
Start Time	5/10/2026 9:52:36 PM
End Time	5/10/2026 10:16:20 PM
Caller	MOHAMED-AHMED\VICTUS

Parameters Used

Name	Value	Data Type
CALLER_INFO		String
DUMP_EVENT_CODE	0	String
DUMP_ON_ERROR	False	Boolean
DUMP_ON_EVENT	False	Boolean
LOGGING_LEVEL	1	Int32
SYNCHRONIZED	False	Boolean

Property Overrides

Property Path	Property Value

2 / 5 Use scroll wheel or PAGE UP & PAGE DOWN to move through pages.

Fig. 6.2.2

Succeeded	89.75	From Source to Dim_Coaches.dtsx	Data Flow Task	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Coaches\Package1\Data Flow Task
Succeeded	8.188	From Source to Dim_Coaches.dtsx	Get Last Extract Time	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Coaches\Package1\Get Last Extract Time
Succeeded	57.172	From Source to Dim_Coaches.dtsx	Load To Staging Table	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Coaches\Package1\Load To Staging Table
Succeeded	8.219	From Source to Dim_Coaches.dtsx	Truncate Staging Table	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Coaches\Package1\Truncate Staging Table
Succeeded	8.203	From Source to Dim_Coaches.dtsx	Update Last Extract Date	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Coaches\Package1\Update Last Extract Date
Succeeded	171.969	From Source to Dim_Leagues.dtsx	From Source to Dim_Leagues	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Leagues\From Source to Dim_Leagues
Succeeded	90.047	From Source to Dim_Leagues.dtsx	Data Flow Task	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Leagues\From Source to Dim_Leagues\Data Flow Task
Succeeded	8.188	From Source to Dim_Leagues.dtsx	Get Last Extract Time	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Leagues\From Source to Dim_Leagues\Get Last Extract Time
Succeeded	57.344	From Source to Dim_Leagues.dtsx	Load To Staging Table	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Leagues\From Source to Dim_Leagues\Load To Staging Table
Succeeded	8.187	From Source to Dim_Leagues.dtsx	Truncate Staging Table	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Leagues\From Source to Dim_Leagues\Truncate Staging Table
Succeeded	8.203	From Source to Dim_Leagues.dtsx	Update Last Extract Date	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Leagues\From Source to Dim_Leagues\Update Last Extract Date
Succeeded	140.484	From Source to Dim_Score_junk.dtsx	Package1	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Score_Junk\Package1

3 / 5 Use scroll wheel or PAGE UP & PAGE DOWN to move through pages.

Fig. 6.2.3

Succeeded	57.422	From Source to Dim_Score_junk.dtsx	Data Flow Task	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Score_Junk\Package1\Data Flow Task
Succeeded	8.172	From Source to Dim_Score_junk.dtsx	Get Last Extract Time	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Score_Junk\Package1\Get Last Extract Time
Succeeded	58.437	From Source to Dim_Score_junk.dtsx	Load To Staging Table	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Score_Junk\Package1\Load To Staging Table
Succeeded	8.219	From Source to Dim_Score_junk.dtsx	Truncate Staging Table	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Score_Junk\Package1\Truncate Staging Table
Succeeded	8.219	From Source to Dim_Score_junk.dtsx	Update Last Extract Date	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Score_Junk\Package1\Update Last Extract Date
Succeeded	155.516	From Source to Dim_Stadium.dtsx	From Source to Dim_Stadium	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Stadium\From Source to Dim_Stadium
Succeeded	8.203	From Source to Dim_Stadium.dtsx	Get Last Extract Time	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Stadium\From Source to Dim_Stadium\Get Last Extract Time
Succeeded	81.875	From Source to Dim_Stadium.dtsx	Load Dim Stadium	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Stadium\From Source to Dim_Stadium\Load Dim Stadium
Succeeded	49.047	From Source to Dim_Stadium.dtsx	Load STG_Stadiums	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Stadium\From Source to Dim_Stadium\Load STG_Stadiums
Succeeded	8.187	From Source to Dim_Stadium.dtsx	Truncate STG Table	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Stadium\From Source to Dim_Stadium\Truncate STG Table
Succeeded	8.172	From Source to Dim_Stadium.dtsx	Update Last Extract Date	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Stadium\From Source to Dim_Stadium\Update Last Extract Date
Succeeded	155.703	From Source to Dim_Team.dtsx	Package1	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Team\Package1

4 / 5 Use scroll wheel or PAGE UP & PAGE DOWN to move through pages.

Fig. 6.2.4

Succeeded	8.219	From Source to Dim_Team.dtsx	Get Last Extract Time	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Team\Package1\Get Last Extract Time
Succeeded	81.875	From Source to Dim_Team.dtsx	Load Dim Team	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Team\Package1\Load Dim Team
Succeeded	49.187	From Source to Dim_Team.dtsx	Load STG_teams	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Team\Package1\Load STG_teams
Succeeded	8.203	From Source to Dim_Team.dtsx	Truncate STG Table	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Team\Package1\Truncate STG Table
Succeeded	8.203	From Source to Dim_Team.dtsx	Update Last Extract Date	\01_Load_Dimentions \Load_Dimentions \Load_from_source_to_Dim_Team\Package1\Update Last Extract Date

5 / 5 Use scroll wheel or PAGE UP & PAGE DOWN to move through pages.

Fig. 6.2.5

6.3. After Second Run Screenshots

Overview

on Mohamed-Ahmed at 5/10/2026 10:39:38 PM

This report provides an overview of the package tasks and parameters, including execution or validation information.

View Messages

View Performance

Execution Information

Operation ID	3
Package	DWH_ETL\DataWarehouse_ETL_Project\02_Load_Facts.dtsx
Environment	-
Status	Succeeded
Machine	Mohamed-Ahmed

Execution Overview

Filter: Result: All; (3 more)

Result	Duration (sec)	Package Name	Task Name	Execution Path
Succeeded	1093.64	02_Load_Facts.dtsx	02_Load_Facts	\02_Load_Facts
Succeeded	1093.593	02_Load_Facts.dtsx	Load_Facts	\02_Load_Facts\Load_Facts
Succeeded	386.437	02_Load_Facts.dtsx	Load_Fact_Coach_Stadium_Performance	\02_Load_Facts\Load_Facts\Load_Fact_Coach_Stadium_Performance
Succeeded	386.572	02_Load_Facts.dtsx	Load_Fact_Scoring_Efficiency	\02_Load_Facts\Load_Facts\Load_Fact_Scoring_Efficiency
Succeeded	320.484	02_Load_Facts.dtsx	Load_Fact_Team_Monthly_Stat1	\02_Load_Facts\Load_Facts\Load_Fact_Team_Monthly_Stat1
Succeeded	205.625	Load_Fact_Coach_Stadium_Performance.dtsx	Load_Fact_Coach_Stadium_Performance	\02_Load_Facts\Load_Facts\Load_Fact_Coach_Stadium_Performance\Load_Fact_Coach_Stadium_Performance
Succeeded	8.204	Load_Fact_Coach_Stadium_Performance.dtsx	Get Last Extract Date	\02_Load_Facts\Load_Facts\Load_Fact_Coach_Stadium_Performance\Load_Fact_Coach_Stadium_Performance\Get Last Extract Date
Succeeded	131.625	Load_Fact_Coach_Stadium_Performance.dtsx	Load DWH	\02_Load_Facts\Load_Facts\Load_Fact_Coach_Stadium_Performance\Load_Fact_Coach_Stadium_Performance\Load DWH

1 / 3 Use scroll wheel or PAGE UP & PAGE DOWN to move through pages.

Fig. 6.3.1

SQL Server

Duration (sec)	1094.793
Start Time	5/10/2026 10:20:16 PM
End Time	5/10/2026 10:38:31 PM
Caller	MOHAMED-AHMED\VICTUS

Parameters Used

Name	Value	Data Type
CALLER_INFO		String
DUMP_EVENT_CODE	0	String
DUMP_ON_ERROR	False	Boolean
DUMP_ON_EVENT	False	Boolean
LOGGING_LEVEL	1	Int32
SYNCHRONIZED	False	Boolean

Property Overrides

Property Path	Property Value

2 / 3 Use scroll wheel or PAGE UP & PAGE DOWN to move through pages.

Fig. 6.3.2

Succeeded	49.39	Load_Fact_Coach_Stadium_Performance.dtsx	Load STG	\02_Load_Facts\Load_Facts\Load_Fact_Coach_Stadium_Performance\Load_Fact_Coach_Stadium_Performance\Load STG
Succeeded	8.203	Load_Fact_Coach_Stadium_Performance.dtsx	Truncate STG Fact_Coach_Stadium_Performance	\02_Load_Facts\Load_Facts\Load_Fact_Coach_Stadium_Performance\Load_Fact_Coach_Stadium_Performance\Truncate STG Fact_Coach_Stadium_Performance
Succeeded	8.188	Load_Fact_Coach_Stadium_Performance.dtsx	Update Last Extract Date	\02_Load_Facts\Load_Facts\Load_Fact_Coach_Stadium_Performance\Load_Fact_Coach_Stadium_Performance\Update Last Extract Date
Succeeded	205.625	Load_Fact_Scoring_Efficiency.dtsx	Load_Fact_Scoring_Efficiency	\02_Load_Facts\Load_Facts\Load_Fact_Scoring_Efficiency\Load_Fact_Scoring_Efficiency
Succeeded	8.171	Load_Fact_Scoring_Efficiency.dtsx	Get Last Extract Time	\02_Load_Facts\Load_Facts\Load_Fact_Scoring_Efficiency\Load_Fact_Scoring_Efficiency\Get Last Extract Time
Succeeded	131.547	Load_Fact_Scoring_Efficiency.dtsx	Load To Fact Table	\02_Load_Facts\Load_Facts\Load_Fact_Scoring_Efficiency\Load_Fact_Scoring_Efficiency\Load To Fact Table
Succeeded	49.469	Load_Fact_Scoring_Efficiency.dtsx	Load To Staging Area	\02_Load_Facts\Load_Facts\Load_Fact_Scoring_Efficiency\Load_Fact_Scoring_Efficiency\Load To Staging Area
Succeeded	8.219	Load_Fact_Scoring_Efficiency.dtsx	Truncate Staging Table	\02_Load_Facts\Load_Facts\Load_Fact_Scoring_Efficiency\Load_Fact_Scoring_Efficiency\Truncate Staging Table
Succeeded	8.203	Load_Fact_Scoring_Efficiency.dtsx	Update Last Extract Date	\02_Load_Facts\Load_Facts\Load_Fact_Scoring_Efficiency\Load_Fact_Scoring_Efficiency\Update Last Extract Date
Succeeded	172.593	Load_Fact_Team_Monthly_Stat.dtsx	Package1	\02_Load_Facts\Load_Facts\Load_Fact_Team_Monthly_Stat\Package1
Succeeded	8.187	Load_Fact_Team_Monthly_Stat.dtsx	Get Last Extract Date	\02_Load_Facts\Load_Facts\Load_Fact_Team_Monthly_Stat\Package1\Get Last Extract Date
Succeeded	98.704	Load_Fact_Team_Monthly_Stat.dtsx	Load to Fact Table	\02_Load_Facts\Load_Facts\Load_Fact_Team_Monthly_Stat\Package1\Load to Fact Table
Succeeded	49.25	Load_Fact_Team_Monthly_Stat.dtsx	Load To Staging Area	\02_Load_Facts\Load_Facts\Load_Fact_Team_Monthly_Stat\Package1\Load To Staging Area
Succeeded	8.234	Load_Fact_Team_Monthly_Stat.dtsx	Truncate Staging Table	\02_Load_Facts\Load_Facts\Load_Fact_Team_Monthly_Stat\Package1\Truncate Staging Table
Succeeded	8.218	Load_Fact_Team_Monthly_Stat.dtsx	Update Last Extract Date	\02_Load_Facts\Load_Facts\Load_Fact_Team_Monthly_Stat\Package1\Update Last Extract Date

3 / 3 Use scroll wheel or PAGE UP & PAGE DOWN to move through pages.

Fig. 6.3.3

7. Queries & Insights

This section presents analytical SQL queries used to generate insights from the warehouse.

7.1. Team Performance Analysis

This query identifies the strongest teams based on total points, wins, draws, and losses across all seasons.

```
SELECT
    T.Name AS Team_Name,
    SUM(F.Points_total) AS Total_Points,
    SUM(F.Total_wins_count) AS
Total_Wins,
    SUM(F.Draw_count) AS Total_Draws,
    SUM(F.Loss_count) AS Total_Losses
FROM DWH.Fact_Team_Monthly_Stat F
JOIN DWH.Dim_Team T ON F.Team_key =
T.Team_key
GROUP BY T.Name;
```

Results | Messages

	Team_Name	Total_Points	Total_Wins	Total_Draws	Total_Losses
1	1. FC Heidenheim 1846	42	10	12	12
2	1. FC Köln	27	5	12	17
3	1. FC Union Berlin	33	9	6	19
4	1. FSV Mainz 05	35	7	14	13
5	AC Milan	75	22	9	7
6	AC Monza	45	11	12	15
7	ACF Fiorentina	60	17	9	12
8	AFC Bournemouth	48	13	9	16
9	Arsenal	89	28	5	6
10	AS Monaco FC	67	20	7	7
11	AS Roma	63	18	9	11
12	Aston Villa	68	20	8	10
13	Atalanta BC	69	21	6	11
14	Athletic Club	68	19	11	8
15	Bayer 04 Leverkusen	90	28	6	0
16	Bologna FC 1909	68	18	14	6
17	Borussia Dortmund	63	18	9	7
18	Borussia Mönchengla...	34	7	13	14
19	Brentford	39	10	9	19
20	Brighton & Hove Albion	48	12	12	14
21	Burnley	24	5	9	24
22	CA Osasuna	45	12	9	17
23	Cádiz CF	33	6	15	17
24	Cagliari Calcio	36	8	12	18
25	Chelsea	63	18	9	11

7.2. Goals & Offensive Performance

Teams with high goal differences generally maintained stronger attacking and defensive balance.

```
SELECT
    T.Name AS Team_Name,
    SUM(F.Goals_count) AS
Total_Goals_Scored,
    SUM(F.Goals_conceded) AS
Total_Goals_Conceded,
    SUM(F.Goal_difference) AS
Total_Goal_Difference
FROM DWH.Fact_Team_Monthly_Stat F
JOIN DWH.Dim_Team T ON F.Team_key =
T.Team_key
GROUP BY T.Name;
```

Results | Messages

	Team_Name	Total_Goals_Scored	Total_Goals_Conceded	Total_Goal_Difference
1	1. FC Heidenheim 1846	50	55	-5
2	1. FC Köln	28	60	-32
3	1. FC Union Berlin	33	58	-25
4	1. FSV Mainz 05	39	51	-12
5	AC Milan	76	49	27
6	AC Monza	39	51	-12
7	ACF Fiorentina	61	46	15
8	AFC Bournemouth	54	67	-13
9	Arsenal	94	32	62
10	AS Monaco FC	68	42	26
11	AS Roma	65	46	19
12	Aston Villa	76	61	15
13	Atalanta BC	72	42	30
14	Athletic Club	61	37	24
15	Bayer 04 Leverkusen	89	24	65
16	Bologna FC 1909	54	32	22
17	Borussia Dortmund	68	43	25
18	Borussia Mönchengla...	56	67	-11
19	Brentford	56	65	-9
20	Brighton & Hove Albion	55	62	-7
21	Burnley	41	78	-37
22	CA Osasuna	45	56	-11
23	Cádiz CF	26	55	-29
24	Cagliari Calcio	42	68	-26
25	Chelsea	77	63	14

7.3. Team Efficiency Metrics

This query measures team consistency and efficiency using win percentage and average points per match.

```sql
SELECT
    T.Name AS Team_Name,
    SUM(F.Total_matches_played) AS
Matches_Played,
    (CAST(SUM(F.Total_wins_count) AS
FLOAT) /
NULLIF(SUM(F.Total_matches_played), 0)) *
100 AS Win_Rate,
    CAST(SUM(F.Points_total) AS FLOAT) /
NULLIF(SUM(F.Total_matches_played), 0) AS
Avg_Points_Per_Match
FROM DWH.Fact_Team_Monthly_Stat F
JOIN DWH.Dim_Team T ON F.Team_key =
T.Team_key
GROUP BY T.Name;
```

Results | Messages

	Team_Name	Matches_Played	Win_Rate	Avg_Points_Per_Match
1	1. FC Heidenheim 1846	34	29.4117647058824	1.23529411764706
2	1. FC Köln	34	14.7058823529412	0.794117647058823
3	1. FC Union Berlin	34	26.4705882352941	0.970588235294118
4	1. FSV Mainz 05	34	20.5882352941176	1.02941176470588
5	AC Milan	38	57.8947368421053	1.97368421052632
6	AC Monza	38	28.9473684210526	1.18421052631579
7	ACF Fiorentina	38	44.7368421052632	1.57894736842105
8	AFC Bournemouth	38	34.2105263157895	1.26315789473684
9	Arsenal	39	71.7948717948718	2.28205128205128
10	AS Monaco FC	34	58.8235294117647	1.97058823529412
11	AS Roma	38	47.3684210526316	1.65789473684211
12	Aston Villa	38	52.6315789473684	1.78947368421053
13	Atalanta BC	38	55.2631578947368	1.81578947368421
14	Athletic Club	38	50	1.78947368421053
15	Bayer 04 Leverkusen	34	82.3529411764706	2.64705882352941
16	Bologna FC 1909	38	47.3684210526316	1.78947368421053
17	Borussia Dortmund	34	52.9411764705882	1.85294117647059
18	Borussia Mönchengladbach	34	20.5882352941176	1
19	Brentford	38	26.3157894736842	1.02631578947368
20	Brighton & Hove Albion	38	31.5789473684211	1.26315789473684
21	Burnley	38	13.1578947368421	0.631578947368421
22	CA Osasuna	38	31.5789473684211	1.18421052631579
23	Cádiz CF	38	15.7894736842105	0.868421052631579
24	Cagliari Calcio	38	21.0526315789474	0.947368421052632
25	Chelsea	38	47.3684210526316	1.65789473684211

7.4. League Performance Analysis

This query compares leagues based on total scoring activity and match competitiveness.

```sql
SELECT
    L.Name AS League_Name,
    SUM(F.Goals_count) AS
Total_League_Goals,
    CAST(SUM(F.Goals_count) AS FLOAT) /
SUM(F.Total_matches_played) AS
Avg_Goals_Per_Match,
    SUM(F.Total_matches_played) AS
Total_League_Matches
FROM DWH.Fact_Team_Monthly_Stat F
JOIN DWH.Dim_League L ON F.League_key =
L.League_key
GROUP BY L.Name;
```

Results | Messages

	League_Name	Total_League_Goals	Avg_Goals_Per_Match	Total_League_Matches
1	Bundesliga	985	1.60947712418301	612
2	La Liga	1005	1.32236842105263	760
3	Ligue 1	826	1.34967320261438	612
4	Premier League	1252	1.64304461942257	762
5	Serie A	992	1.30526315789474	760

7.5. Home Advantage Analysis

The results show how much playing at home impacts winning probability in each league.

```
SELECT
    L.Name AS League_Name,
    CAST(SUM(F.Home_wins_count) AS FLOAT)
/ NULLIF(SUM(F.Total_wins_count), 0) AS
Home_Advantage_Ratio
FROM DWH.Fact_Team_Monthly_Stat F
JOIN DWH.Dim_League L ON F.League_key =
L.League_key
GROUP BY L.Name;
```

Results | Messages

	League_Name	Home_Advantage_Ratio
1	Bundesliga	0.595555555555556
2	La Liga	0.611721611721612
3	Ligue 1	0.533333333333333
4	Premier League	0.588628762541806
5	Serie A	0.593283582089552

7.6. First Half vs Second Half Goals

This query analyzes whether teams perform better offensively during the first or second half.

```
SELECT
    T.Name AS Team_Name,
    SUM(F.Goals_1st_half) AS
Total_1st_Half_Goals,
    SUM(F.Goals_2nd_half) AS
Total_2nd_Half_Goals
FROM DWH.Fact_Scoring_Efficiency F
JOIN DWH.Dim_Team T ON F.Team_key =
T.Team_key
GROUP BY T.Name;
```

Results | Messages

	Team_Name	Total_1st_Half_Goals	Total_2nd_Half_Goals
1	1. FC Heidenheim 1846	18	28
2	1. FC Köln	10	15
3	1. FC Union Berlin	15	14
4	1. FSV Mainz 05	19	14
5	AC Milan	36	35
6	AC Monza	17	18
7	ACF Fiorentina	30	27
8	AFC Bournemouth	16	26
9	Arsenal	34	25
10	AS Monaco FC	34	30
11	AS Roma	20	30
12	Aston Villa	36	36
13	Atalanta BC	34	28
14	Athletic Club	30	17
15	Bayer 04 Leverkusen	30	35
16	Bologna FC 1909	25	21
17	Borussia Dortmund	35	23
18	Borussia Mönchengla...	30	26
19	Brentford	19	27
20	Brighton & Hove Albion	21	28
21	Burnley	16	18
22	CA Osasuna	19	21
23	Cádiz CF	15	9
24	Cagliari Calcio	12	27
25	Chelsea	36	36

7.7. Comeback & Clean Sheet Analysis

Teams with higher comeback wins and clean sheets demonstrated strong mentality and defensive stability.

```sql
SELECT
    T.Name AS Team_Name,
    COUNT(CASE WHEN J.Is_comeback_win = 1
THEN 1 END) AS Comeback_Count,
    COUNT(CASE WHEN J.Is_clean_sheet = 1
THEN 1 END) AS Clean_Sheet_Count
FROM DWH.Fact_Scoring_Efficiency F
JOIN DWH.Dim_Team T ON F.Team_key =
T.Team_key
JOIN DWH.Dim_Scoring_Junk J ON
F.Scoring_junk_key = J.Scoring_junk_key
GROUP BY T.Name;
```

Results | Messages

	Team_Name	Comeback_Count	Clean_Sheet_Count
1	FC Barcelona	7	6
2	US Sassuolo Calcio	2	2
3	Cagliari Calcio	3	3
4	Club Atlético de Madrid	4	6
5	Valencia CF	2	9
6	Clermont Foot 63	1	4
7	Chelsea	5	6
8	UD Almería	1	6
9	Paris Saint-Germain FC	4	8
10	Fulham	3	6
11	Sheffield United	1	1
12	Luton Town	3	2
13	1. FC Heidenheim 1846	3	3
14	SV Darmstadt 98	1	2
15	Nottingham Forest	3	3
16	Liverpool	10	6
17	Brentford	5	3
18	Manchester United	5	3
19	VfL Bochum 1848	0	2
20	Deportivo Alavés	0	6
21	FC Bayern München	6	5
22	AC Milan	7	9
23	RB Leipzig	6	7
24	Athletic Club	1	12
25	SV Werder Bremen	2	3

7.8. Coach Performance Analysis

This query evaluates coach effectiveness using win rate and average points earned.

```sql
SELECT
    C.Name AS Coach_Name,
    (CAST(COUNT(CASE WHEN F.Result_type =
'W' THEN 1 END) AS FLOAT) / COUNT(*)) *
100 AS Coach_Win_Rate,
    AVG(CAST(F.Points_earned AS FLOAT))
AS Coach_Points_Per_Match
FROM DWH.Fact_Coach_Stadium_Performance F
JOIN DWH.Dim_Coach C ON F.Coach_key =
C.Coach_key
GROUP BY C.Name;
```

Results | Messages

	Coach_Name	Coach_Win_Rate	Coach_Points_Per_Match
1	Adi Hütter	58.8235294117647	1.97058823529412
2	Albert Peris	26.3157894736842	1.05263157894737
3	Alberto Gilardino	31.5789473684211	1.28947368421053
4	Ange Postecoglou	52.6315789473684	1.73684210526316
5	Antoine Kombouare	26.4705882352941	0.970588235294118
6	Bo Henriksen	20.5882352941176	1.02941176470588
7	Bouziane Benaraïbi	35.2941176470588	1.35294117647059
8	Carles Martínez	32.3529411764706	1.26470588235294
9	Carlo Ancelotti	76.3157894736842	2.5
10	Christian Streich	32.3529411764706	1.23529411764706
11	Claudio	26.3157894736842	1.07894736842105
12	Claudio Ranieri	21.0526315789474	0.947368421052632
13	Damien Della Santa	47.0588235294118	1.55882352941176
14	Daniele De Rossi	47.3684210526316	1.65789473684211
15	Davide Ballardini	18.4210526315789	0.789473684210526
16	Davide Nicola	23.6842105263158	0.947368421052632
17	Diego Simeone	63.1578947368421	2
18	Dino Toppmöller	32.3529411764706	1.38235294117647
19	Eddie Howe	47.3684210526316	1.57894736842105
20	Edin Terzić	52.9411764705882	1.85294117647059
21	Enzo Maresca	47.3684210526316	1.65789473684211
22	Eric Roy	50	1.79411764705882
23	Erik ten Hag	47.3684210526316	1.57894736842105
24	Ernesto Valverde	50	1.78947368421053
25	Eusebio Di France...	21.0526315789474	0.921052631578947

7.9. Stadium Performance Analysis

Certain stadiums provided a stronger home advantage and positively influenced team performance.

```sql
SELECT
    S.Name AS Stadium_Name,
    (CAST(COUNT(CASE WHEN F.Result_type =
'W' AND F.Is_home_coach = 1 THEN 1 END)
AS FLOAT) /
     NULLIF(COUNT(CASE WHEN
F.Is_home_coach = 1 THEN 1 END), 0)) *
100 AS Home_Stadium_Win_Rate
FROM DWH.Fact_Coach_Stadium_Performance F
JOIN DWH.Dim_Stadium S ON F.Stadium_key =
S.Stadium_key
GROUP BY S.Name;
```

Results | Messages

	Stadium_Name	Home_Stadium_Win_Rate
1	Allianz Arena	82.3529411764706
2	Allianz Stadium	57.8947368421053
3	Anfield	78.9473684210526
4	BayArena	82.3529411764706
5	Bramall Lane	10.5263157894737
6	Camp Nou	78.9473684210526
7	Coliseum Alfonso Pérez	42.1052631578947
8	Craven Cottage	47.3684210526316
9	Dacia Arena	5.26315789473684
10	Deutsche Bank Park	41.1764705882353
11	Emirates Stadium	78.9473684210526
12	Estadi Municipal de Montilivi	78.9473684210526
13	Estadio Benito Villamarín	47.3684210526316
14	Estadio de Balaídos	31.5789473684211
15	Estadio de Gran Canaria	31.5789473684211
16	Estadio de la Cerámica	36.8421052631579
17	Estadio de los Juegos Med...	5.26315789473684
18	Estadio de Mendizorroza	47.3684210526316
19	Estadio de Mestalla	42.1052631578947
20	Estadio del Rayo Vallecano	21.0526315789474
21	Estadio El Sadar	31.5789473684211
22	Estadio Municipal de Anoeta	42.1052631578947
23	Estadio Nuevo Los Cármen...	21.0526315789474
24	Estadio Ramón de Carranza	26.3157894736842
25	Estadio Ramón Sánchez P...	31.5789473684211

7.10. Monthly Performance Trend

This query tracks seasonal momentum and monthly performance fluctuations for each team.

```sql
SELECT
    T.Name AS Team_Name,
    D.MonthName,
    SUM(F.Goals_count) AS Monthly_Goals,
    SUM(F.Points_total) AS Monthly_Points
FROM DWH.Fact_Team_Monthly_Stat F
JOIN DWH.Dim_Team T ON F.Team_key =
T.Team_key
JOIN DWH.Dim_Date D ON F.Month_key =
D.DateKey
GROUP BY T.Name, D.MonthName, D.Month
ORDER BY T.Name, D.Month;
```

Results | Messages

	Team_Name	MonthName	Monthly_Goals	Monthly_Points
1	1. FC Heidenheim 1846	January	3	3
2	1. FC Heidenheim 1846	February	5	5
3	1. FC Heidenheim 1846	March	5	2
4	1. FC Heidenheim 1846	April	6	7
5	1. FC Heidenheim 1846	May	6	5
6	1. FC Heidenheim 1846	August	2	0
7	1. FC Heidenheim 1846	September	8	7
8	1. FC Heidenheim 1846	October	3	0
9	1. FC Heidenheim 1846	November	4	4
10	1. FC Heidenheim 1846	December	8	9
11	1. FC Köln	January	2	2
12	1. FC Köln	February	4	5
13	1. FC Köln	March	5	2
14	1. FC Köln	April	3	4
15	1. FC Köln	May	4	4
16	1. FC Köln	August	1	0
17	1. FC Köln	September	3	1
18	1. FC Köln	October	3	3
19	1. FC Köln	November	2	2
20	1. FC Köln	December	1	4
21	1. FC Union Berlin	January	1	4
22	1. FC Union Berlin	February	5	8
23	1. FC Union Berlin	March	2	4
24	1. FC Union Berlin	April	1	1
25	1. FC Union Berlin	May	7	3

7.11. Second Half Specialist Teams

Some teams perform significantly better during the second half of matches.

```sql
SELECT
    T.Name AS Team_Name,
    SUM(F.Goals_1st_half) AS
Total_1st_Half,
    SUM(F.Goals_2nd_half) AS
Total_2nd_Half,
    COUNT(CASE WHEN
J.Is_second_half_specialist = 1 THEN 1
END) AS Specialist_Match_Count
FROM DWH.Fact_Scoring_Efficiency F
JOIN DWH.Dim_Team T ON F.Team_key =
T.Team_key
JOIN DWH.Dim_Scoring_Junk J ON
F.Scoring_junk_key = J.Scoring_junk_key
GROUP BY T.Name
ORDER BY Specialist_Match_Count DESC;
```

Results | Messages

	Team_Name	Total_1st_Half	Total_2nd_Half	Specialist_Match_Count
1	Luton Town	20	32	21
2	Villarreal CF	23	36	18
3	Cagliari Calcio	12	27	15
4	RC Strasbourg Alsace	9	24	15
5	1. FC Heidenheim 1846	18	28	14
6	Liverpool	31	45	14
7	AS Roma	20	30	14
8	Tottenham Hotspur	25	39	14
9	Brighton & Hove Albion	21	28	14
10	Wolverhampton Wanderers	17	29	13
11	TSG 1899 Hoffenheim	34	32	13
12	Olympique Lyonnais	15	28	13
13	AC Milan	36	35	13
14	Stade Rennais FC 1901	21	28	13
15	Granada CF	15	20	13
16	Aston Villa	36	36	12
17	West Ham United	25	33	12
18	Crystal Palace	22	24	12
19	Sheffield United	14	21	12
20	Frosinone Calcio	19	25	12
21	Brentford	19	27	12
22	Manchester City	38	46	12
23	Toulouse FC	17	22	12
24	Hellas Verona FC	15	18	12
25	AFC Bournemouth	16	26	12

7.12. Coach Impact Score

This query evaluates coach performance against different opponents and tactical matchups.

```
SELECT
    C.Name AS Coach_Name,
    OC.Name AS Opponent_Coach_Name,
    COUNT(*) AS Matches_Played,
    SUM(F.Points_earned) AS
Points_Gained_In_Matchup
FROM DWH.Fact_Coach_Stadium_Performance F
JOIN DWH.Dim_Coach C ON F.Coach_key =
C.Coach_key
JOIN DWH.Dim_Coach OC ON
F.Opponent_coach_key = OC.Coach_key
GROUP BY C.Name, OC.Name
ORDER BY Points_Gained_In_Matchup DESC;
```

Results | Messages

	Coach_Name	Opponent_Coach_Name	Matches_Played	Points_Gained_In_Matchup
1	Unknown	Unknown	84	114
2	Pep Guardiola	Unknown	14	40
3	Mikel Arteta	Unknown	14	39
4	Jürgen Klopp	Unknown	14	33
5	Ange Postecoglou	Unknown	14	32
6	Enzo Maresca	Unknown	14	30
7	Unai Emery	Unknown	14	28
8	Marco Silva	Unknown	14	26
9	Unknown	Gary O'Neil	14	26
10	Erik ten Hag	Unknown	14	23
11	Sean Dyche	Unknown	14	23
12	Unknown	Vincent Kompany	14	22
13	Eddie Howe	Unknown	14	22
14	Unknown	Nuno Espirito Santo	14	21
15	Unknown	Erik ten Hag	14	17
16	Unknown	Eddie Howe	14	16
17	Vincent Kompany	Unknown	14	16
18	Nuno Espirito Sa...	Unknown	14	15
19	Gary O'Neil	Unknown	14	14
20	Unknown	Sean Dyche	14	14
21	Unknown	Marco Silva	14	11
22	Unknown	Unai Emery	14	10
23	Unknown	Enzo Maresca	14	9
24	Unknown	Ange Postecoglou	14	8
25	Bouziane Benar...	Antoine Kombouare	2	6

7.13. Stadium Capacity Impact

Larger stadiums may positively influence team morale and overall performance.

```
SELECT
    S.Name AS Stadium_Name,
    S.Capacity,
    AVG(CAST(F.Points_earned AS FLOAT))
AS Avg_Points_At_Venue
FROM DWH.Fact_Coach_Stadium_Performance F
JOIN DWH.Dim_Stadium S ON F.Stadium_key =
S.Stadium_key
GROUP BY S.Name, S.Capacity
ORDER BY Avg_Points_At_Venue DESC;
```

Results | Messages

	Stadium_Name	Capacity	Avg_Points_At_Venue
1	Tottenham Hotspur Stadium	62062	1.5
2	Estadio Wanda Metropolitano	68456	1.47368421052632
3	Camp Nou	99354	1.47368421052632
4	Allianz Arena	75000	1.47058823529412
5	Emirates Stadium	60704	1.44736842105263
6	Estadi Municipal de Montilivi	14500	1.44736842105263
7	Stadio Atleti Azzurri d'Italia	24950	1.44736842105263
8	Craven Cottage	25700	1.44736842105263
9	Stadion An der Alten Försterei	22000	1.44117647058824
10	Stade Océane	25178	1.44117647058824
11	Molineux Stadium	32050	1.42105263157895
12	Anfield	54074	1.42105263157895
13	Old Trafford	74879	1.42105263157895
14	Estadio Santiago Bernabéu	78297	1.42105263157895
15	Signal Iduna Park	81365	1.41176470588235
16	Mercedes-Benz Arena	60449	1.41176470588235
17	Stade de la Beaujoire - Loui...	35200	1.41176470588235
18	Stade de la Meinau	29000	1.41176470588235
19	Stade Saint-Symphorien	30000	1.41176470588235
20	BayArena	30210	1.41176470588235
21	Stade Auguste-Delaune II	21000	1.41176470588235
22	Merck-Stadion am Böllenfall...	17800	1.41176470588235
23	Stade Louis II.	18500	1.41176470588235
24	Estadio de Mendizorroza	19800	1.39473684210526
25	Kenilworth Road Stadium	10356	1.39473684210526

7.14. Average Goals per Match by League

This query identifies high-scoring leagues and offensive playing styles.

```
SELECT
    L.Name AS League_Name,
    F.Season_key,
    AVG(CAST(F.Goals_total AS FLOAT)) AS
Avg_Goals_Per_Match
FROM DWH.Fact_Scoring_Efficiency F
JOIN DWH.Dim_League L ON F.League_key =
L.League_key
GROUP BY L.Name, F.Season_key
ORDER BY Avg_Goals_Per_Match DESC;
```

Results | Messages

	League_Name	Season_key	Avg_Goals_Per_Match
1	Premier League	1	1.56168359941945
2	Bundesliga	1	1.54545454545455
3	Ligue 1	1	1.3001808318264
4	Serie A	1	1.2561505065123
5	La Liga	1	1.25513196480938

7.15. Competitive Balance Analysis

Lower volatility indicates more stable and balanced team performance across the season.

```
SELECT
    T.Name AS Team_Name,
    L.Name AS League_Name,
    STDEV(F.Points_total) AS
Points_Volatility,
    AVG(F.Points_total) AS
Average_Monthly_Points
FROM DWH.Fact_Team_Monthly_Stat F
JOIN DWH.Dim_Team T ON F.Team_key =
T.Team_key
JOIN DWH.Dim_League L ON F.League_key =
L.League_key
GROUP BY T.Name, L.Name
ORDER BY Points_Volatility ASC;
```

Results | Messages

	Team_Name	League_Name	Points_Volatility	Average_Monthly_Points
1	Borussia Mönchengladbach	Bundesliga	1.07496769977314	3
2	US Salernitana 1919	Serie A	1.33749350984926	1
3	1. FC Köln	Bundesliga	1.56702123647242	2
4	Sheffield United	Premier League	1.57762127549323	1
5	Real Madrid CF	La Liga	1.64991582276861	9
6	SSC Napoli	Serie A	1.70293863659264	5
7	Le Havre AC	Ligue 1	1.75119007154183	3
8	VfL Bochum 1848	Bundesliga	1.76698110409314	3
9	SV Darmstadt 98	Bundesliga	1.76698110409314	1
10	Cagliari Calcio	Serie A	1.7763883459299	3
11	Clermont Foot 63	Ligue 1	1.84089350286454	2
12	Valencia CF	La Liga	1.85292561462497	4
13	Rayo Vallecano de Madrid	La Liga	1.87379590967403	3
14	VfL Wolfsburg	Bundesliga	1.88856206322871	3
15	Eintracht Frankfurt	Bundesliga	1.88856206322871	4
16	Stade de Reims	Ligue 1	1.88856206322871	4
17	Luton Town	Premier League	1.89736659610103	2
18	Bayer 04 Leverkusen	Bundesliga	1.9436506316151	9
19	Nottingham Forest	Premier League	1.95505043981536	3
20	Racing Club de Lens	Ligue 1	1.96920739836559	5
21	FC Bayern München	Bundesliga	1.98885785202351	7
22	Genoa CFC	Serie A	2.02484567313166	4
23	Getafe CF	La Liga	2.05750658160146	4
24	Paris Saint-Germain FC	Ligue 1	2.06559111797729	7
25	UD Almería	La Liga	2.07899548393502	2

7.16. Stadium Tier Performance Analysis

Teams playing in larger stadiums often achieve better attacking statistics and point averages.

```
SELECT
    CASE
        WHEN S.Capacity >= 50000 THEN 'Elite
(50k+)'
        WHEN S.Capacity >= 30000 THEN 'Large
(30k-50k)'
        ELSE 'Medium/Small (<30k)'
    END AS Stadium_Tier,
    AVG(CAST(F.Points_earned AS FLOAT)) AS
Avg_Points_Per_Match,
    SUM(F.Goals_scored) AS Total_Goals_At_Tier
FROM DWH.Fact_Coach_Stadium_Performance F
JOIN DWH.Dim_Stadium S ON F.Stadium_key =
S.Stadium_key
GROUP BY
    CASE
        WHEN S.Capacity >= 50000 THEN 'Elite
(50k+)'
        WHEN S.Capacity >= 30000 THEN 'Large
(30k-50k)'
        ELSE 'Medium/Small (<30k)'
    END;
```

Results | Messages

	Stadium_Tier	Avg_Points_Per_Match	Total_Goals_At_Tier
1	Elite (50k+)	1.38970588235294	1478
2	Large (30k-50k)	1.36589698046181	1592
3	Medium/Small (<30k)	1.35483870967742	1984

7.17. Second Half Goal Ratio Analysis

Teams with high second-half goal ratios usually demonstrate strong fitness and tactical adaptability.

```
SELECT
    T.Name AS Team_Name,
    SUM(F.Goals_1st_half) AS FH_Goals,
    SUM(F.Goals_2nd_half) AS SH_Goals,
    CAST(SUM(F.Goals_2nd_half) AS FLOAT)
/ NULLIF(SUM(F.Goals_total), 0) AS
SH_Goal_Ratio
FROM DWH.Fact_Scoring_Efficiency F
JOIN DWH.Dim_Team T ON F.Team_key =
T.Team_key
GROUP BY T.Name
HAVING CAST(SUM(F.Goals_2nd_half) AS
FLOAT) / NULLIF(SUM(F.Goals_total), 0) >
0.6
ORDER BY SH_Goal_Ratio DESC;
```

Results | Messages

	Team_Name	FH_Goals	SH_Goals	SH_Goal_Ratio
1	RC Strasbourg Alsace	9	24	0.727272727272727
2	Cagliari Calcio	12	27	0.692307692307692
3	US Lecce	10	20	0.666666666666667
4	Olympique Lyonnais	15	28	0.651162790697674
5	SSC Napoli	18	33	0.647058823529412
6	Wolverhampton Wanderers	17	29	0.630434782608696
7	FC Barcelona	21	35	0.625
8	AFC Bournemouth	16	26	0.619047619047619
9	SV Werder Bremen	15	24	0.615384615384615
10	Luton Town	20	32	0.615384615384615
11	US Salernitana 1919	12	19	0.612903225806452
12	Villarreal CF	23	36	0.610169491525424
13	Tottenham Hotspur	25	39	0.609375
14	1. FC Heidenheim 1846	18	28	0.608695652173913
15	Empoli FC	11	17	0.607142857142857

7.18. Away Match Coach Analysis

This query identifies coaches who maintain strong performance in away matches.

```
SELECT
    C.Name AS Coach_Name,
    COUNT(F.Fact_performance_key) AS
Away_Matches,
    AVG(CAST(F.Points_earned AS FLOAT))
AS Avg_Away_Points,
    SUM(SE.Goals_2nd_half) AS
Total_SH_Goals_Away
FROM DWH.Fact_Coach_Stadium_Performance F
JOIN DWH.Dim_Coach C ON F.Coach_key =
C.Coach_key
JOIN DWH.Fact_Scoring_Efficiency SE ON
F.Team_key = SE.Team_key
    AND F.Date_key = SE.Date_key
WHERE F.Is_home_coach = 0
GROUP BY C.Name
ORDER BY Avg_Away_Points DESC;
```

Results | Messages

	Coach_Name	Away_Matches	Avg_Away_Points	Total_SH_Goals_Away
1	Xabi Alonso	14	2.57142857142857	17
2	Luis Enrique	14	2.42857142857143	13
3	Gianluca Zappalà	13	2.30769230769231	11
4	Pep Guardiola	16	2.1875	23
5	Carlo Ancelotti	15	2.13333333333333	15
6	Xavi Hernandez	17	1.94117647058824	18
7	Mikel Arteta	14	1.92857142857143	12
8	Adi Hütter	15	1.86666666666667	17
9	Michel	18	1.72222222222222	16
10	Stefano Pioli	17	1.70588235294118	19
11	Jürgen Klopp	17	1.64705882352941	20
12	Thomas Tuchel	16	1.625	16
13	Edin Terzić	15	1.6	9
14	Carles Martínez	16	1.5	8
15	Imanol Alguacil	18	1.5	10
16	Paolo Montero	17	1.47058823529412	13
17	Marco Rose	15	1.46666666666667	19
18	Sebastian Hoeneß	13	1.46153846153846	12
19	Francesco Calzo...	18	1.38888888888889	17
20	Eric Roy	13	1.38461538461538	16
21	Enzo Maresca	19	1.36842105263158	19
22	Pellegrino Matar...	17	1.35294117647059	15
23	Thiago Motta	18	1.33333333333333	7
24	Unai Emery	17	1.29411764705882	10
25	Gian Piero Gasp...	17	1.29411764705882	13

7.19. Lead Blowing Analysis

This query analyzes tactical weaknesses and defensive instability after leading at halftime.

```
SELECT
    C.Name AS Coach_Name,
    COUNT(F.Fact_performance_key) AS
Matches_Led_At_HT,
    COUNT(CASE WHEN J.Is_lead_blown = 1
THEN 1 END) AS Leads_Blown,
    (CAST(COUNT(CASE WHEN J.Is_lead_blown
= 1 THEN 1 END) AS FLOAT) / COUNT(*)) *
100 AS Lead_Blowing_Rate
FROM DWH.Fact_Coach_Stadium_Performance F
JOIN DWH.Dim_Coach C ON F.Coach_key =
C.Coach_key
JOIN DWH.Fact_Scoring_Efficiency SE ON
F.Team_key = SE.Team_key AND F.Date_key =
SE.Date_key
JOIN DWH.Dim_Scoring_Junk J ON
SE.Scoring_junk_key = J.Scoring_junk_key
GROUP BY C.Name
ORDER BY Lead_Blowing_Rate DESC;
```

Results | Messages

	Coach_Name	Matches_Led_At_HT	Leads_Blown	Lead_Blowing_Rate
1	Ralph Hasenhüttl	33	13	39.3939393939394
2	Davide Ballardini	37	14	37.8378378378378
3	Bo Henriksen	32	12	37.5
4	Christian Streich	29	10	34.4827586206897
5	Gerardo Seoane	34	10	29.4117647058824
6	Edin Terzić	29	8	27.5862068965517
7	Pellegrino Matarazzo	34	9	26.4705882352941
8	Antoine Kombouare	31	8	25.8064516129032
9	Pepe Bordalás	35	9	25.7142857142857
10	Samba Diawara	32	8	25
11	Vincent Kompany	36	9	25
12	Stefano Colantuono	37	9	24.3243243243243
13	Dino Toppmöller	33	8	24.2424242424242
14	Torsten Lieberknecht	33	8	24.2424242424242
15	Michel Der Zakarian	29	7	24.1379310344828
16	Pepe Mel	38	9	23.6842105263158
17	Fabio Cannavaro	35	8	22.8571428571429
18	László Bölöni	31	7	22.5806451612903
19	Igor Tudor	31	7	22.5806451612903
20	Mauricio Pellegrino	36	8	22.2222222222222
21	Kersten Kuhl	32	7	21.875
22	Heiko Butscher	32	7	21.875
23	Jess Thorup	33	7	21.2121212121212
24	Gianluca Zappalà	29	6	20.6896551724138
25	Ange Postecoglou	34	7	20.5882352941176

8. Power BI Dashboard (Bonus)

An interactive dashboard was developed using Microsoft Power BI to visualize and analyze the football Data Warehouse data.

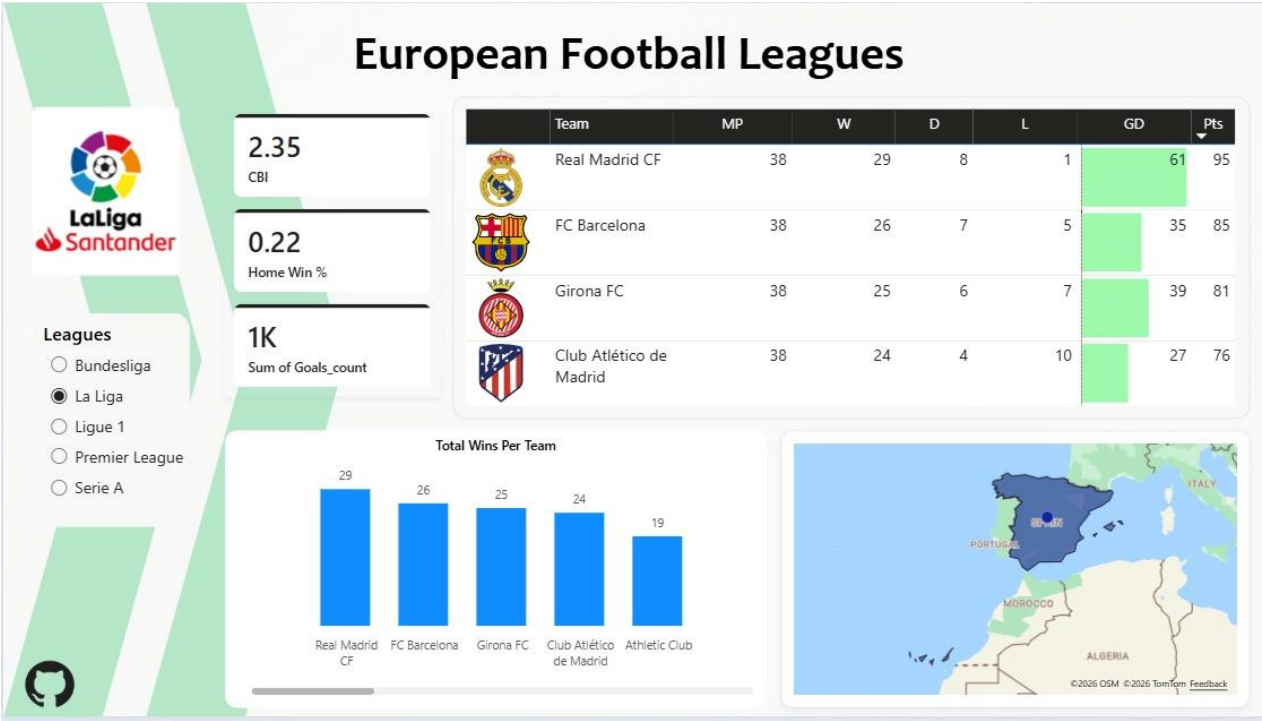
The dashboard provides dynamic insights into team performance, league statistics, coach analysis, scoring efficiency, and stadium performance across multiple seasons.

8.1. Dashboard Data Source

The dashboard is connected directly to the `European_Football_Leagues_DWH` database and retrieves data from:

- `Fact_Team_Monthly_Stat`
- All related dimension tables

8.2. Dashboard Screenshot



European Football Leagues
LIGUE 1
1.54
CBI
0.20
Home Win %
826
Sum of Goals_count
Leagues
Bundesliga
La Liga
Ligue 1
Premier League
Serie A
Team MP W D L GD Pts
Paris Saint-Germain FC 34 22 10 2 48 76
AS Monaco FC 34 20 7 7 26 67
Stade Brestois 29 34 17 10 7 19 61
Lille OSC 34 16 11 7 18 59
Total Wins Per Team
22
20
17
16
16
Paris Saint-Germ... FC
AS Monaco FC
Stade Brestois 29
Lille OSC
Olympique Lyonnais
London
GERMANY
Paris
CZECHIA
AUSTRIA
FRANCE
Bay of Biscay
ITALY
Rome
SPAIN

European Football Leagues
Premier League
2.19
CBI
0.23
Home Win %
1K
Sum of Goals_count
Leagues
Bundesliga
La Liga
Ligue 1
Premier League
Serie A
Team MP W D L GD Pts
Manchester City 38 28 7 3 62 91
Arsenal 38 28 5 5 62 89
Liverpool 38 24 10 4 45 82
Aston Villa 38 20 8 10 15 68
Total Wins Per Team
28
28
24
20
20
Arsenal
Manchester City
Liverpool
Aston Villa
Tottenham Hotspur
NORWAY
North Sea
DENMARK
UNITED KINGDOM
IRELAND

European Football Leagues
BUNDESLIGA
2.30
CBI
0.22
Home Win %
985
Sum of Goals_count
Leagues
Bundesliga
La Liga
Ligue 1
Premier League
Serie A
Team MP W D L GD Pts
Bayer 04 Leverkusen 34 28 6 0 65 90
VfB Stuttgart 34 23 4 7 39 73
FC Bayern München 34 23 3 8 49 72
RB Leipzig 34 19 8 7 38 65
Total Wins Per Team
28
23
23
19
18
Bayer 04 Leverkusen
FC Bayern München
VfB Stuttgart
RB Leipzig
Borussia Dortmund
UNITED KINGDOM
NETHERLANDS
Berlin
London
POLAND
GERMANY
CZECHIA
Paris
SLOVAKIA
FRANCE
AUSTRIA
HUNGARY

9. Conclusion

This project successfully designed and implemented a complete Football Data Warehouse solution using SQL Server and SSIS.

The system integrates data from multiple operational tables into a centralized analytical warehouse optimized for reporting and business intelligence.

The project included:

- Source database design
- Staging area implementation
- Dimensional modeling using Star Schema
- ETL pipeline development
- KPI generation
- Analytical SQL queries
- Power BI dashboard

The warehouse supports advanced football analytics including team performance, scoring efficiency, coach impact analysis, and stadium performance insights.

The implementation of incremental loading and dimensional modeling improved scalability, query performance, and reporting efficiency.

Future Enhancements

Potential future improvements include:

- Real-time streaming data integration
- Machine learning prediction models
- Player-level analytics
- Advanced Power BI dashboards
- Automated alert systems
- Cloud-based warehouse deployment