Avances Proyecto Integrador

Integrantes:

- Luis Morales
- Santiago Riofrio

Estadística del Dataset Original

- Número total de jugadores: 10401
- Número total de equipos: 88
- Porcentaje de jugadores masculinos: 105.49947120469186%
- Porcentaje de jugadores femeninos: 27.59350062493991%
- Numero de Camisetas mas común y su posición:
  defender -> Camiseta = 3, repeticiones = 501
  midfielder -> Camiseta = 8, repeticiones = 370
  forward -> Camiseta = 0, repeticiones = 504
  goal keeper -> Camiseta = 1, repeticiones = 563

- Estadio con maxima capacidad: Estádio do Maracanã,20000
- Estadio con minima capacidad: Tingvalla IP,5000

Diccionario de Datos

Tabla: Players

Atributo	Tipo de Dato
player_id (PK)	Integer(100)
players_family_name	Varchar(MAX)
players_given_name	Varchar(MAX)
players_birth_date	Date
players_female	Boolean
players_goal_keeper	Boolean
players_defender	Boolean
players_midfielder	Boolean
players_forward	Boolean

Tabla: Teams

Atributo	Tipo de Dato
team_id (PK)	Integer
team_name	Varchar(MAX)

Tabla: Tournaments

Atributo	Tipo de Dato
tournament_id (PK)	Integer
tournament_name	Varchar(100)
tournament_name	Integer
tournament_name	Varchar(50)
tournament_name	Varchar(100)
tournament_name	Integer

Tabla: Stadiums

Atributo	Tipo de Dato
stadiums_stadium_id (PK)	Integer
stadiums_stadium_name	Varchar(MAX)
stadiums_city_name	Varchar(MAX)
stadiums_country_name	Varchar(MAX)
stadiums_stadium_capacity	Integer

Tabla: Squads

Atributo	Tipo de Dato
squad_id (PK)	Integer
squad_player_id (FK)	Integer
squad_team_id (FK)	Integer
squad_tournament_id (FK)	Integer
squad_shirt_number	Integer
squad_position_name	Varchar(MAX)

Tabla: Matches

Atributo	Tipo de Dato
matches_match_id (PK)	Integer
matches_tournament_id (FK)	Intege
matches_away_team_id (FK)	Integer
matches_home_team_id (FK)	Integer
matches_stadium_id (FK)	Integer
matches_match_date	Date
matches_match_time	Time

matches_stage_name	Varchar(MAX)
matches_home_team_score	Integer
matches_away_team_score	Integer
matches_extra_time	Boolean
matches_penalty_shootout	Boolean
matches_home_team_score_penalties	Integer
matches_away_team_score_penalties	Integer
matches_result	Varchar(MAX)

Tabla: Goals

Atributo	Tipo de Dato
goals_goal_id (PK)	Integer
goals_team_id (FK)	Integer
goals_	Integer
goals_player_team_id (FK)	Integer
goals_minute_label	INT
goals_minute_regulation	Double
goals_minute_stoppage	Integer
goals_match_period	Varchar(MAX)
goals_own_goal	Boolean
goals_penalty	Boolean

Esquema de Base de Datos

Para la esquematización del dataset se a empleado la herramienta: "Dia", para realizar el diagrama de Entidad Relación

Stadiums

*stadium_id	Integer(10)
*stadium_name	Varchar(100)
*city_name	Varchar(50)
*country_name	Varchar(50)
°stadium_capacity	Integer(5)

Matches

*match_id	Integer(10)
*tournament_id (FK)	Integer(10)
°away_team_id (FK)	Integer(10)
°home_team_id (FK)	Integer(10)
*stadium_id (FK)	Integer(10)
*match_date	Date
*match_time	Time
°stage_name	Varchar(50)
*home_team_score	Integer(3)
°away_team_score	Integer(3)
*extra_time	Boolean
*penalty_shootout	Boolean
*home_team_score_penalties	Integer(3)
*away_team_score_penalties	Integer(3)
*result	Varchar(20)

Goals

*goal_id	
*team_id (FK)	Integer(10)
*player_id (FK)	Integer(10)
*player_team_id (FK)	Integer(10)
*minute_label	Integer(10)
*minute_regulation	Double(3)
*minute_stoppage	Integer(3)
*match_period	Varchar(10)
*own_goal	Boolean
*penalty	Boolean

Tournaments

*tournament_id	Integer(10)
*tournament_name	Varchar(100)
*year	Integer(4)
*host_country	Varchar(50)
*winner	Varchar(100)
*count_teams	Integer(5)

Players

*player_id	Integer(10)
*family_name	Varchar(50)
*given_name	Varchar(50)
*birth_date	Date
*female	Boolean
*goal_keeper	Boolean
*defender	Boolean
*midfielder	Boolean
*forward	Boolean

Teams

*team_id	Integer(10)
*team_name	Varchar(100)

Squads

*squad_id	
*player_id (FK)	Integer(10)
*team_id (FK)	Integer(10)
*tournament_id (FK)	Integer(10)
*shirt_number	Integer(3)
*position_name	Varchar(50)
°	