Informe del Proyecto de Base de Datos

Integrantes:

- Morales Bautista Luis Miguel
- Riofrio Jaramillo Santiago Ismael

Introducción

Este proyecto tuvo como objetivo brindar una experiencia práctica sobre el manejo del ciclo de vida de los datos utilizando archivos CSV. En este proyecto, se aplicaron conceptos teóricos y técnicas de modelado de datos, diseño de bases de datos, y manipulación de datos.

1. Análisis de los Datos y Modelo Conceptual

Análisis de los Datos

Como punto de partida inicial, se comenzó explorando los archivos CSV proporcionados para entender la naturaleza de los datos, identificando los elementos clave como entidades, atributos y posibles relaciones.

Para la exploración de los archivos se hizo uso del sistema de gestión de bases de datos (DBMS) denominado MySQL Workbench para importar los archivos CSV y a su vez se trabajó con en el entorno de desarrollo DataGrip, el cual nos permitirá ejecutar sentencias SQL para poder así analizar de mejor manera los datasets y sus datos correspondientes.

A continuación, se presentan las sentencias usadas con su respectiva salida generada, las cuales nos permitieron, de manera eficaz, obtener más información acerca de los datasets con los que se trabajo.

Dataset: PartidosYGoles (dspartidosygoles)

Sentencia SQL:

```sql
# Numero de Registros
SELECT
    'PartidosYGoles' AS Tabla,
    'Indicador General Del Datset PartidosYGoles' AS TipoIndicador,
    'Cantidad de Registros' AS Indicador,
    COUNT(*) AS Valor
FROM
    dspartidosygoles;
```

Salida:

Tabla	TipoIndicador	Indicador	Valor
PartidosYGoles	Indicador General Del Datset PartidosYGoles	Cantidad de Registros	3723

Sentencia SQL:

```sql
# Exploracion de Nulls

# Minute Regulation
SELECT DISTINCT
    goals_minute_regulation AS "Minute Regulation"
FROM
    dspartidosygoles
ORDER BY
    1;


# Playe Id
SELECT DISTINCT
    goals_player_id AS "Player Id"
FROM
    dspartidosygoles
ORDER BY
    1;
```

Salidas:

	`Minute Regulation`
1	<null>
2	1
3	2
4	3
5	4
6	5
7	6
8	7

	`Player ID`
1	NA
2	P-00042
3	P-00064
4	P-00065
5	P-00105
6	P-00119
7	P-00152

Dataset: AlineacionesXTorneo (alineacionesxtorneo-2)

Sentencias:

```
#                    Numero                         de                     Registros
SELECT
    'PartidosYGoles'                                      AS                        Tabla,
    'Indicador General Del Datset AlineacionesXTorneo' AS TipoIndicador,
    'Cantidad           de           Registros'          AS              Indicador,
    COUNT(*)                                            AS                        Valor
FROM
    `dsalineacionesxtorneo-2` ;
```

Salida:

Tabla	TipoIndicador	Indicador	Valor
PartidosYGoles	Indicador General Del Datset AlineacionesXTorneo	Cantidad de Registros	13843

Sentencias:

```
#       Identifiacion       de       Relaciones       entre       Tablas

SELECT

    squads_player_id       AS       'ID       de       los       Jugador',
    squads_tournament_id         AS         'ID         del         Torneo',
    CONCAT(players_given_name,  ' ',  players_family_name)  AS  'Nombre  deL
Jugador',
    squads_position_name'Posicion               deL               Jugador'

FROM
    `dsalineacionesxtorneo-2`

ORDER                                                                   BY
    squads_player_id;
```

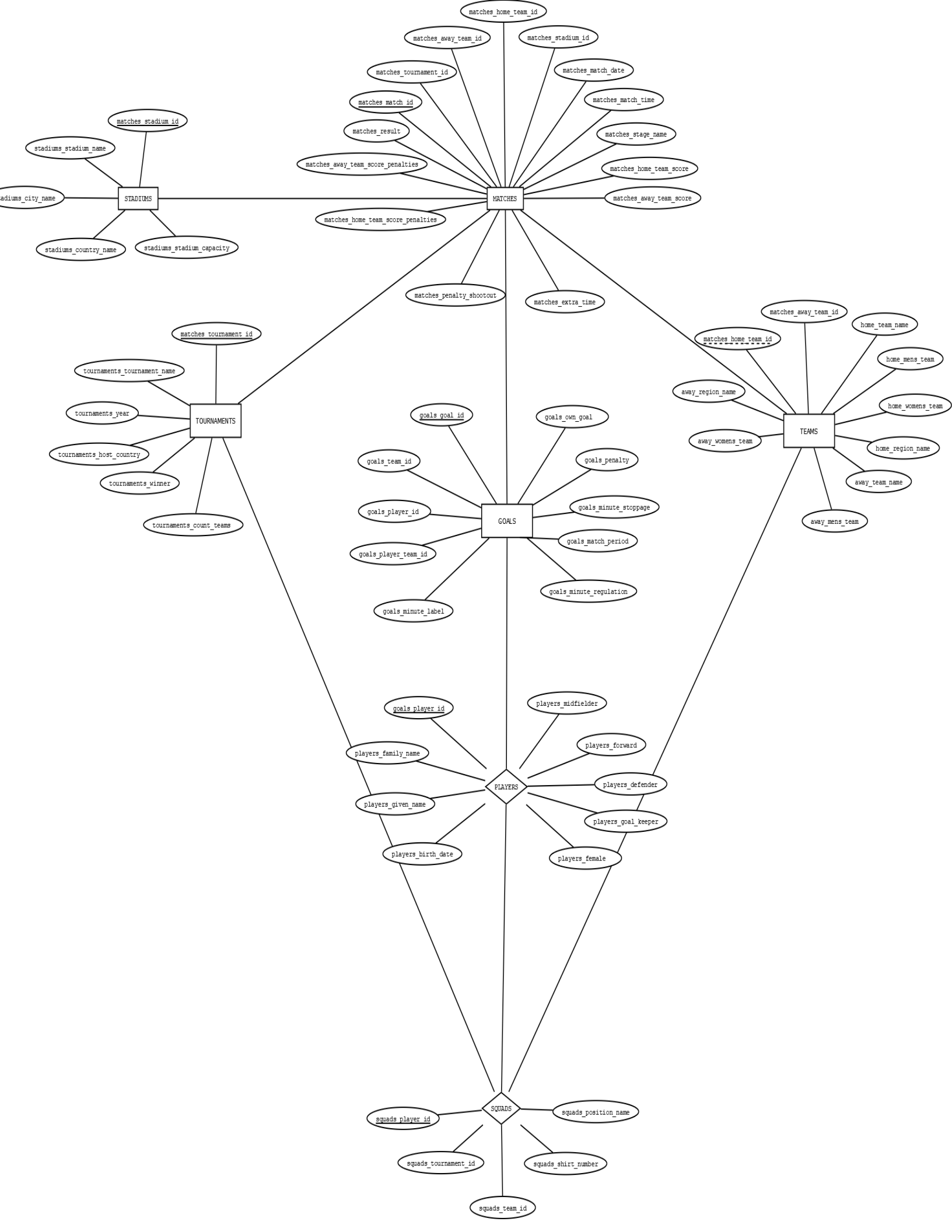
Salida:

`ID de los Jugador`	`ID del Torneo`	`Nombre deL Jugador`	`Posicion deL Jugador`
P-00020	WC-2007	Randee Hermus	defender
P-00020	WC-2003	Randee Hermus	defender
P-00034	WC-1998	not applicable Alfonso	forward
P-00036	WC-1974	Pleun Strik	defender
P-00042	WC-1954	Miloš Milutinović	forward
P-00042	WC-1958	Miloš Milutinović	forward
P-00052	WC-2022	not applicable Ró-Ró	defender

Modelo Conceptual Entidad-Relación (E/R)

Para la diagramación del modelo (E/R) se optó por utilizar una herramienta que nos permita representar las entidades identificadas, sus atributos, y las relaciones entre ellas. En este caso se utilizó la herramienta DIA para capturar la estructura lógica de los datos sin preocuparnos por detalles de implementación.

A continuación, se presenta el diagrama (E/R) hecho en la herramienta mencionada.


matches_home_team_id
matches_away_team_id
matches_stadium_id
matches_tournament_id
matches_match_date
matches_match_id
matches_match_time
matches_result
matches_stage_name
matches_away_team_score_penalties
matches_home_team_score
MATCHES
matches_away_team_score
matches_home_team_score_penalties
matches_penalty_shootout
matches_extra_time
matches_stadium_id
stadiums_stadium_name
STADIUMS
adiums_city_name
stadiums_country_name
stadiums_stadium_capacity
matches_tournament_id
tournaments_tournament_name
tournaments_year
TOURNAMENTS
tournaments_host_country
tournaments_winner
tournaments_count_teams
matches_away_team_id
home_team_name
matches_home_team_id
home_mens_team
away_region_name
home_womens_team
TEAMS
away_womens_team
home_region_name
away_team_name
away_mens_team
goals_goal_id
goals_own_goal
goals_team_id
goals_penalty
goals_player_id
GOALS
goals_minute_stoppage
goals_match_period
goals_player_team_id
goals_minute_regulation
goals_minute_label
goals_player_id
players_midfielder
players_forward
players_family_name
PLAYERS
players_defender
players_given_name
players_goal_keeper
players_birth_date
players_female
SQUADS
squads_player_id
squads_position_name
squads_tournament_id
squads_shirt_number
squads_team_id

2. Normalización

Se analizaron los datos para identificar dependencias funcionales y transitivas entre atributos, lo que ayudó a entender las relaciones de unicidad y dependencia. Se aplicaron las reglas de normalización (hasta la tercera forma normal) para asegurar que la base de datos estuviera libre de redundancias innecesarias y anomalías de inserción, actualización o eliminación.

Primer Forma Normal (Datos Atómicos / Identificación Datos Repetitivos)

En esta primera forma de normalización se busca que los datos de cada atributo del dataset sean atómicos, además de analizarse los datos repetitivos; para esto usaremos la herramienta Excel que nos ayudara a representar de mejor manera dichos datos.

Dataset: AlineacionesXTorneo (alineacionesxtorneo-2)

				Identificacion de Datos Repetitivos								
squads_player	squads_tournam	squads_team	squads_shirt_	squads_posit	players_fami	players_give	players_birth	players_fema	players_goal	players_defe	players_midfielder	players_forward
P-00083	WC-2015	T-46	3	defender	Murillo	Christina	28/01/1993	1	0	1	0	0
P-00212	WC-2002	T-22	1	goal keeper	SÃ¸rensen	Thomas	12/06/1976	0	1	0	0	0
P-00212	WC-2010	T-22	1	goal keeper	SÃ¸rensen	Thomas	12/06/1976	0	1	0	0	0
P-00065	WC-2002	T-03	10	midfielder	Ortega	Ariel	04/03/1974	0	0	0	1	0
P-00065	WC-1994	T-03	17	midfielder	Ortega	Ariel	04/03/1974	0	0	0	1	0
P-00065	WC-1998	T-03	10	midfielder	Ortega	Ariel	04/03/1974	0	0	0	1	0
P-00064	WC-2015	T-30	12	midfielder	Thomis	Ã‰lodie	13/08/1986	1	0	0	1	1
P-00064	WC-2011	T-30	12	forward	Thomis	Ã‰lodie	13/08/1986	1	0	0	1	1
P-00471	WC-2015	T-83	3	defender	Rampone	Christie	24/06/1975	1	0	1	0	0
P-00471	WC-2003	T-83	3	defender	Rampone	Christie	24/06/1975	1	0	1	0	0
P-00471	WC-1999	T-83	3	defender	Rampone	Christie	24/06/1975	1	0	1	0	0
P-00471	WC-2011	T-83	3	defender	Rampone	Christie	24/06/1975	1	0	1	0	0
P-00471	WC-2007	T-83	3	defender	Rampone	Christie	24/06/1975	1	0	1	0	0

squads_player	squads_tournam	squads_team	squads_shirt_	squads_posit	players_fami	players_give	players_birth	players_fema	players_goal	players_defe	players_midfielder	players_forward
P-00083	WC-2015	T-46	3	defender	Murillo	Christina	28/01/1993	1	0	1	0	0
P-00212	WC-2002	T-22	1	goal keeper	SÃ¸rensen	Thomas	12/06/1976	0	1	0	0	0
	WC-2010		1	goal keeper								
P-00065	WC-2002	T-03	10	midfielder	Ortega	Ariel	04/03/1974	0	0	0	1	0
	WC-1994		17	midfielder								
	WC-1998		10	midfielder								
P-00064	WC-2015	T-30	12	midfielder	Thomis	Ã‰lodie	13/08/1986	1	0	0	1	1
	WC-2011		12	forward								
P-00471	WC-2015	T-83	3	defender	Rampone	Christie	24/06/1975	1	0	1	0	0
	WC-2003		3	defender								
	WC-1999		3	defender								
	WC-2011		3	defender								
	WC-2007		3	defender								

Dataset: PartidosYGoles (dspartidosygoles)

matches_tou	matches_mat	matches_awa	matches_hon	matches_stac	matches_match_date	matches_mat	matches_stag	matches_hon	matches_awa	matches_ext	matches_penalty_sh	matches_hon	matches_awa
WC-2022	M-2022-01	T-25	T-59	S-140	20/11/2022	19:00	group stage	0	2	0	0	0	0
WC-2022	M-2022-01	T-25	T-59	S-140	20/11/2022	19:00	group stage	0	2	0	0	0	0
WC-1930	M-1930-01	T-46	T-30	S-240	13/07/1930	15:00	group stage	4	1	0	0	0	0
WC-1930	M-1930-01	T-46	T-30	S-240	13/07/1930	15:00	group stage	4	1	0	0	0	0
WC-1930	M-1930-01	T-46	T-30	S-240	13/07/1930	15:00	group stage	4	1	0	0	0	0
WC-1930	M-1930-01	T-46	T-30	S-240	13/07/1930	15:00	group stage	4	1	0	0	0	0
WC-1930	M-1930-01	T-46	T-30	S-240	13/07/1930	15:00	group stage	4	1	0	0	0	0
WC-1930	M-1930-16	T-83	T-03	S-238	26/07/1930	14:45	semi-finals	6	1	0	0	0	0
WC-1930	M-1930-16	T-83	T-03	S-238	26/07/1930	14:45	semi-finals	6	1	0	0	0	0
WC-1930	M-1930-16	T-83	T-03	S-238	26/07/1930	14:45	semi-finals	6	1	0	0	0	0
WC-1930	M-1930-16	T-83	T-03	S-238	26/07/1930	14:45	semi-finals	6	1	0	0	0	0
WC-1930	M-1930-16	T-83	T-03	S-238	26/07/1930	14:45	semi-finals	6	1	0	0	0	0
WC-1930	M-1930-16	T-83	T-03	S-238	26/07/1930	14:45	semi-finals	6	1	0	0	0	0
WC-1930	M-1930-16	T-83	T-03	S-238	26/07/1930	14:45	semi-finals	6	1	0	0	0	0

matches_resu	tournaments_	tournaments_	tournaments_	tournaments_	tournaments_	stadiums_sta	stadiums_city	stadiums_cou	stadiums_sta	home_team_	home_mens_	home_wome	home_region	away_team_r
away team wi	2022 FIFA Me	2022	Qatar	Argentina	32	Al Bayt Stadiu	Al Khor	Qatar	69000	Qatar	1	0	Middle East	Ecuador
away team wi	2022 FIFA Me	2022	Qatar	Argentina	32	Al Bayt Stadiu	Al Khor	Qatar	69000	Qatar	1	0	Middle East	Ecuador
home team w	1930 FIFA Me	1930	Uruguay	Uruguay	13	Estadio Pocit	Montevideo	Uruguay	10000	France	1	1	Europe	Mexico
home team w	1930 FIFA Me	1930	Uruguay	Uruguay	13	Estadio Pocit	Montevideo	Uruguay	10000	France	1	1	Europe	Mexico
home team w	1930 FIFA Me	1930	Uruguay	Uruguay	13	Estadio Pocit	Montevideo	Uruguay	10000	France	1	1	Europe	Mexico
home team w	1930 FIFA Me	1930	Uruguay	Uruguay	13	Estadio Pocit	Montevideo	Uruguay	10000	France	1	1	Europe	Mexico
home team w	1930 FIFA Me	1930	Uruguay	Uruguay	13	Estadio Pocit	Montevideo	Uruguay	10000	France	1	1	Europe	Mexico
home team w	1930 FIFA Me	1930	Uruguay	Uruguay	13	Estadio Cente	Montevideo	Uruguay	90000	Argentina	1	1	South Americ	United States
home team w	1930 FIFA Me	1930	Uruguay	Uruguay	13	Estadio Cente	Montevideo	Uruguay	90000	Argentina	1	1	South Americ	United States
home team w	1930 FIFA Me	1930	Uruguay	Uruguay	13	Estadio Cente	Montevideo	Uruguay	90000	Argentina	1	1	South Americ	United States
home team w	1930 FIFA Me	1930	Uruguay	Uruguay	13	Estadio Cente	Montevideo	Uruguay	90000	Argentina	1	1	South Americ	United States
home team w	1930 FIFA Me	1930	Uruguay	Uruguay	13	Estadio Cente	Montevideo	Uruguay	90000	Argentina	1	1	South Americ	United States
home team w	1930 FIFA Me	1930	Uruguay	Uruguay	13	Estadio Cente	Montevideo	Uruguay	90000	Argentina	1	1	South Americ	United States
home team w	1930 FIFA Me	1930	Uruguay	Uruguay	13	Estadio Cente	Montevideo	Uruguay	90000	Argentina	1	1	South Americ	United States
home team w	1991 FIFA Wo	1991	China	United States	12	Tainhe Stadiu	Guangzhou	China	60000	China	1	1	East Asia	Norway
home team w	1991 FIFA Wo	1991	China	United States	12	Tainhe Stadiu	Guangzhou	China	60000	China	1	1	East Asia	Norway

away_mens_	away_womens_team	away_region	goals_goal_id	goals_team_i	goals_player_	goals_player_	goals_minute_label	goals_minute	goals_minute	goals_match_	goals_own_g	goals_penalty
1	1	South Americ	G-3467	T-25	P-75912	T-25	31'	31	0	first half	0	0
1	1	South Americ	G-3466	T-25	P-75912	T-25	16'	16	0	first half	0	1
1	1	North Americ	G-0003	T-30	P-60620	T-30	43'	43	0	first half	0	0
1	1	North Americ	G-0005	T-30	P-60620	T-30	87'	87	0	second half	0	0
1	1	North Americ	G-0001	T-30	P-05470	T-30	19'	19	0	first half	0	0
1	1	North Americ	G-0002	T-30	P-99087	T-30	40'	40	0	first half	0	0
1	1	North Americ	G-0004	T-46	P-94135	T-46	70'	70	0	second half	0	0
1	1	North Americ	G-0053	T-03	P-56486	T-03	69'	69	0	second half	0	0
1	1	North Americ	G-0055	T-03	P-70166	T-03	85'	85	0	second half	0	0
1	1	North Americ	G-0051	T-03	P-25760	T-03	20'	20	0	first half	0	0
1	1	North Americ	G-0057	T-83	P-06424	T-83	89'	89	0	second half	0	0
1	1	North Americ	G-0054	T-03	P-70166	T-03	80'	80	0	second half	0	0
1	1	North Americ	G-0056	T-03	P-56486	T-03	87'	87	0	second half	0	0

Segunda Forma Normal (Dependencias Funcionales)

En esta segunda forma de normalización se busca ubicar las dependencias funcionales que existen en cada tabla, así de esta forma poder identificar las posibles relaciones que existen entre los datasets.

Dataset: AlineacionesXTorneo (alineacionesxtorneo-2)

squads_player_	squads_tournam	squads_team	squads_shirt_	squads_position_name	players_famil	players_giver	players_birth	players_fema	players_goal_	players_defe	players_midfielder	players_forward
P-00083	WC-2015	T-46	3	defender	Murillo	Christina	28/01/1993	1	0	1	0	0
P-00212	WC-2002	T-22	1	goal keeper	SÃ¸rensen	Thomas	12/06/1976	0	1	0	0	0
	WC-2010		1	goal keeper								
P-00065	WC-2002	T-03	10	midfielder	Ortega	Ariel	04/03/1974	0	0	0	1	0
	WC-1994		17	midfielder								
	WC-1998		10	midfielder								
P-00064	WC-2015	T-30	12	midfielder	Thomis	Ã‰lodie	13/08/1986	1	0	0	1	1
	WC-2011		12	forward								
P-00471	WC-2015	T-83	3	defender	Rampone	Christie	24/06/1975	1	0	1	0	0
	WC-2003		3	defender								
	WC-1999		3	defender								
	WC-2011		3	defender								
	WC-2007		3	defender								

Dataset: PartidosYGoles (dspartidosygoles)

matches_tou	matches_mat	matches_awa	matches_hon	matches_sta	matches_match_date	matches_mat	matches_stag	matches_hon	matches_awa	matches_ext	matches_penalty_sh	matches_hon	matches_awa	matches_resu	tournaments
WC-1930	M-1930-01	T-46	T-30	S-240	13/07/1930	15:00	group stage	4	1	0	0	0	0	ome team wi	
	M-1930-16	T-83	T-03	S-238	26/07/1930	14:45	semi-finals	6	1	0	0	0	0	ome team wi	FA Men's Wo
WC-1991	M-1991-01	T-53	T-14	S-038	16/11/1991	20:45	group stage	4	0	0	0	0	0	ome team wi	
	M-1991-19	T-31	T-22	S-046	24/11/1991	15:30	quarter-final	1	2	1	0	0	0	away team wi	A Women's W

tournaments	tournaments	tournaments	tournaments	stadiums_sta	stadiums_city	stadiums_cou	stadiums_sta	home_team_	home_mens_	home_wome	home_region	away_team_	away_mens_	away_womens_team
1930	Uruguay	Uruguay	13	Estadio Pocito	Montevideo	Uruguay	10000	France	1	1	Europe	Mexico	1	1
				tadio Centena	Montevideo	Uruguay	90000	Argentina	1	1	South America	United States	1	1
1991	China	United States	12	Tainhe Stadiun	Guangzhou	China	60000	China	1	1	East Asia	Norway	1	1
				ongshan Stadi	Zhongshan	China	12000	Denmark	1	1	Europe	Germany	1	1

away_region_	goals_goal_id	goals_team_i	goals_player_	goals_player_	goals_minute_label	goals_minute	goals_minute	goals_match_	goals_own_g	goals_penalty
North America	G-0003	T-30	P-60620	T-30	43'	43	0	first half	0	0
	G-0005	T-30	P-60620	T-30	87'	87	0	second half	0	0
	G-0001	T-30	P-05470	T-30	19'	19	0	first half	0	0
	G-0002	T-30	P-99087	T-30	40'	40	0	first half	0	0
	G-0004	T-46	P-94135	T-46	70'	70	0	second half	0	0
North America	G-0053	T-03	P-56486	T-03	69'	69	0	second half	0	0
	G-0055	T-03	P-70166	T-03	85'	85	0	second half	0	0
	G-0051	T-03	P-25760	T-03	20'	20	0	first half	0	0
	G-0057	T-83	P-06424	T-83	89'	89	0	second half	0	0
	G-0054	T-03	P-70166	T-03	80'	80	0	second half	0	0
	G-0056	T-03	P-56486	T-03	87'	87	0	second half	0	0
	G-0052	T-03	P-44916	T-03	56'	56	0	second half	0	0
Europe	G-1445	T-14	P-10128	T-14	45'	45	0	first half	0	0
	G-1447	T-14	P-49949	T-14	75'	75	0	second half	0	0
	G-1444	T-14	P-84060	T-14	22'	22	0	first half	0	0
	G-1446	T-14	P-10128	T-14	50'	50	0	second half	0	0
Europe	G-1510	T-31	P-30335	T-31	98'	98	0	extra time, fir	0	0
	G-1509	T-22	P-25916	T-22	25'	25	0	first half	0	1
	G-1508	T-31	P-53366	T-31	17'	17	0	first half	0	1

Tercera Forma Normal (Dependencias Transitivas)

En esta tercera forma de normalización se busca emplear las dependencias transitivas que existen en cada tabla, así de esta forma poder ejemplificar de mejor manera las tablas así también como las relaciones existen dentro de los datasets.

Dataset: AlineacionesXTorneo (alineacionesxtorneo-2)

Players								
squads_player_	players_family_n	players_giver	players_birth	players_female	players_goal_	players_defe	players_midfi	players_forward
P-00083	Murillo	Christina	28/01/1993	1	0	1	0	0
P-00212	SÃ¸rensen	Thomas	12/06/1976	0	1	0	0	0
P-00065	Ortega	Ariel	04/03/1974	0	0	0	1	0
P-00064	Thomis	Ã‰lodie	13/08/1986	1	0	0	1	1
P-00471	Rampone	Christie	24/06/1975	1	0	1	0	0

Squads				
squads_player_	squads_tournam	squads_team	squads_shirt_	squads_position_name
P-00083	WC-2015	T-46	3	defender
P-00212	WC-2002	T-22	1	goal keeper
	WC-2010		1	goal keeper
P-00065	WC-2002	T-03	10	midfielder
	WC-1994		17	midfielder
	WC-1998		10	midfielder
P-00064	WC-2015	T-30	12	midfielder
	WC-2011		12	forward
P-00471	WC-2015	T-83	3	defender
	WC-2003		3	defender
	WC-1999		3	defender
	WC-2011		3	defender
	WC-2007		3	defender

Dataset: PartidosYGoles (dspartidosygoles)

Torunaments					
matches_tou	tournaments	tournaments	tournaments	tournaments	tournaments_count_teams
WC-1930	IFA Men's Wo	1930	Uruguay	Uruguay	13
WC-1991	A Women's W	1991	China	United States	12

Matches														
matches_tou	matches_match_id	matches_awa	matches_hon	matches_stad	matches_mat	matches_mat	matches_stag	matches_hon	matches_awa	matches_extr	matches_pen	matches_hon	matches_awa	matches_result
WC-1930	M-1930-01	T-46	T-30	S-240	13/07/1930	15:00	group stage	4	1	0	0	0	0	home team win
	M-1930-16	T-83	T-03	S-238	26/07/1930	14:45	semi-finals	6	1	0	0	0	0	home team win
WC-1991	M-1991-01	T-53	T-14	S-038	16/11/1991	20:45	group stage	4	0	0	0	0	0	home team win
	M-1991-19	T-31	T-22	S-046	24/11/1991	15:30	quarter-final	1	2	1	0	0	0	away team win

Stadiums

matches_stad	stadiums_sta	stadiums_city	stadiums_cou	stadiums_stadium_capacity
S-240	Estadio Pocitos	Montevideo	Uruguay	10000
S-238	tadio Centenar	Montevideo	Uruguay	90000
S-038	Tainhe Stadiun	Guangzhou	China	60000
S-046	ongshan Stadi	Zhongshan	China	12000

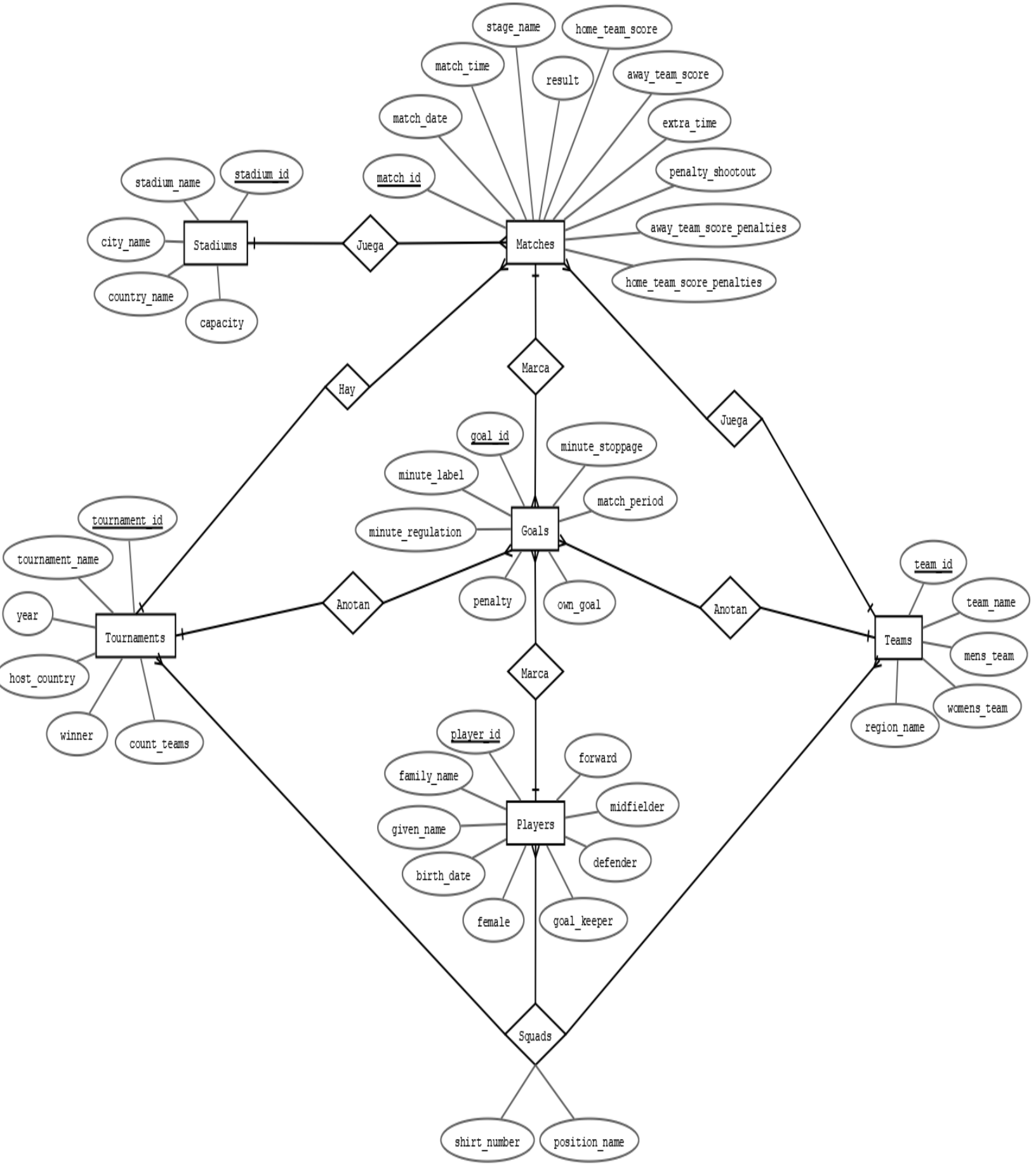
Teams

matches_hon	home_team_	home_mens_	home_wome	home_region_name
T-30	France	1	1	Europe
T-03	Argentina	1	1	South America
T-14	China	1	1	East Asia
T-22	Denmark	1	1	Europe

goals

goals_goal_id	goals_team_i	goals_player_	goals_player_	goals_minute	goals_minute_regulation	goals_minute	goals_match_	goals_own_g	goals_penalty	matches_tou	matches_match_id
G-0003	T-30	P-60620	T-30	43'	43	0	first half	0	0		
G-0005	T-30	P-60620	T-30	87'	87	0	second half	0	0		
G-0001	T-30	P-05470	T-30	19'	19	0	first half	0	0		M-1930-01
G-0002	T-30	P-99087	T-30	40'	40	0	first half	0	0		
G-0004	T-46	P-94135	T-46	70'	70	0	second half	0	0		
G-0053	T-03	P-56486	T-03	69'	69	0	second half	0	0		
G-0055	T-03	P-70166	T-03	85'	85	0	second half	0	0	WC-1930	
G-0051	T-03	P-25760	T-03	20'	20	0	first half	0	0		
G-0057	T-83	P-06424	T-83	89'	89	0	second half	0	0		M-1930-16
G-0054	T-03	P-70166	T-03	80'	80	0	second half	0	0		
G-0056	T-03	P-56486	T-03	87'	87	0	second half	0	0		
G-0052	T-03	P-44916	T-03	56'	56	0	second half	0	0		
G-1445	T-14	P-10128	T-14	45'	45	0	first half	0	0		
G-1447	T-14	P-49949	T-14	75'	75	0	second half	0	0		M-1991-01
G-1444	T-14	P-84060	T-14	22'	22	0	first half	0	0		
G-1446	T-14	P-10128	T-14	50'	50	0	second half	0	0	WC-1991	
G-1510	T-31	P-30335	T-31	98'	98	0	extra time, fir	0	0		
G-1509	T-22	P-25916	T-22	25'	25	0	first half	0	1		M-1991-19
G-1508	T-31	P-53366	T-31	17'	17	0	first half	0	1		

Modelo Entidad Relación Normalizado



Justificación Modelo Conceptual (Normalizado)

Después de haber aplicado las técnicas de normalización se opto por mejorar el modelo (E/R) optando por descartar atributos que resultaban redundantes dentro del modelo inicial, así también como cambiar los nombres de los atributos, esto debido a que sus etiquetas anteriores resultaban confusas y podrían haber errores tanto en el modelo de datos como al momento de construir un modelo físico.

3. Modelo Relacional

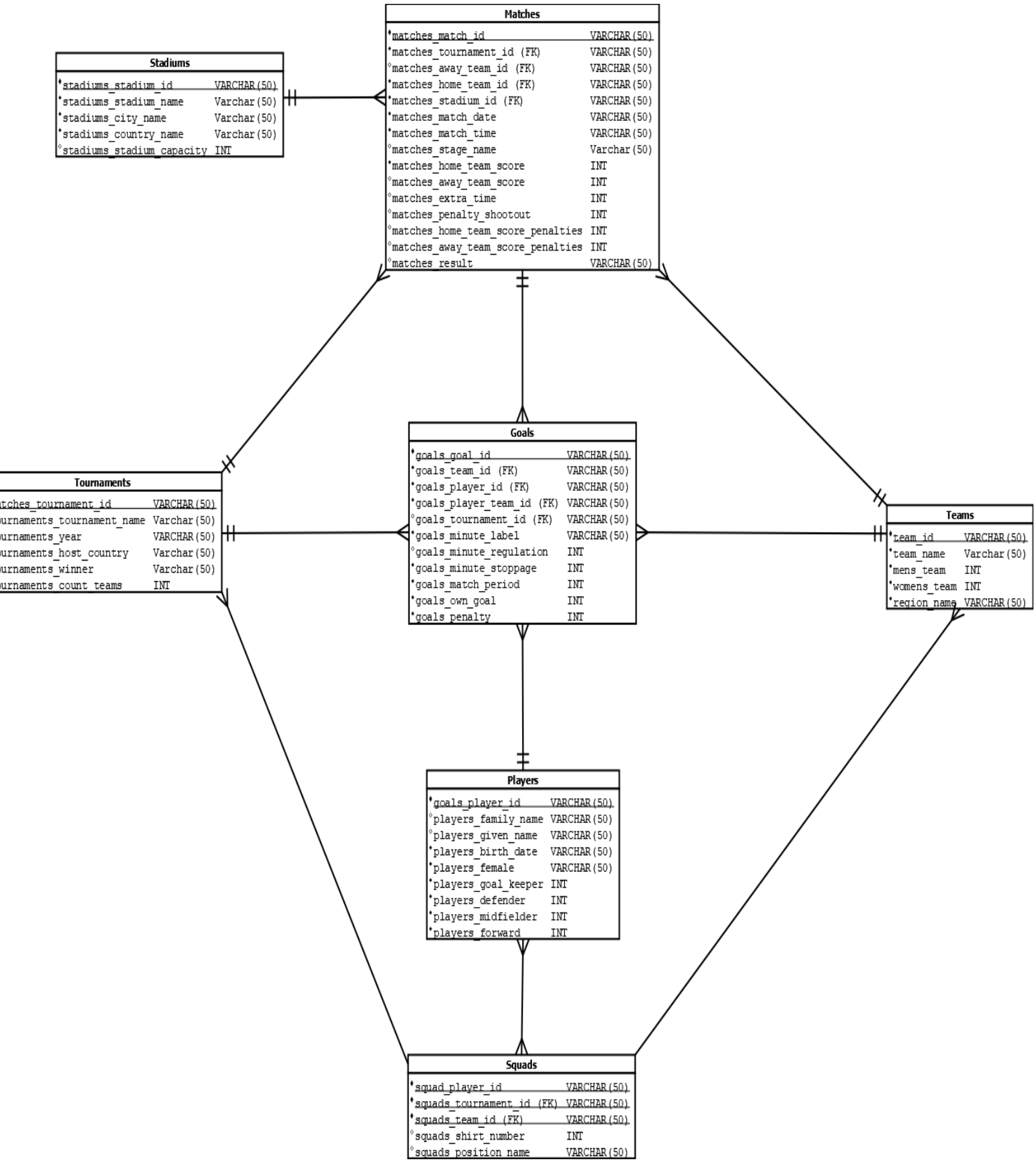
Para el siguiente punto se convirtió el modelo conceptual (E/R) en un modelo relacional.

Esto implicó definir tablas, relaciones para cada entidad y relación, estableciendo claves primarias y foráneas para mantener la integridad referencial.

Para cada tabla identificada, se definió un esquema detallado que incluyó nombres de columnas, tipos de datos y restricciones; como claves primarias y foráneas, condiciones, restricciones de unicidad, etc.

Al igual que el modelo anterior, la herramienta que se usó para la esquematizar dicho modelo es el programa de DIA, ya que también nos ofrece las herramientas apropiadas para construir este modelo.

Modelo Relacional



4. Implementación del Modelo Físico

Para la implementación del modelo físico, se creó la base de datos utilizando un sistema de gestión de bases de datos DBMS (MySQL Workbench), según el modelo relacional diseñado. Para un trabajo más eficiente se utilizó el entorno de desarrollo DataGrip en el cual se desarrollaron scripts y se utilizaron herramientas para importar los datos de los archivos CSV a las tablas correspondientes de la base de datos, asegurando que los datos se ajustaran a las restricciones definidas en el esquema.

A continuación, se presentan las sentencias DDL utilizadas para la creación del modelo físico

Script SQL:

```
# Creacion de la Tabla Stadiums
DROP TABLE IF EXISTS Stadiums;
CREATE TABLE
    Stadiums
SELECT DISTINCT
    matches_stadium_id AS stadium_id,
    stadiums_stadium_name AS stadium_name,
    stadiums_city_name AS city_name,
    stadiums_country_name AS country_name,
    stadiums_stadium_capacity AS stadium_capacity
FROM
    dspartidosygoles;

ALTER TABLE Stadiums
ADD PRIMARY KEY (stadium_id);


 # Creacion de la Tabla Teams
DROP TABLE IF EXISTS Teams;
CREATE TABLE
    Teams
SELECT DISTINCT
    matches_away_team_id AS team_id,
    away_team_name AS team_name,
    away_mens_team AS mens_team,
    away_womens_team AS womens_team,
    away_region_name AS region_name

FROM
    dspartidosygoles;

ALTER TABLE Teams
```

```
ADD PRIMARY KEY (team_id);


 # Creacion de la Tabla Tournaments
DROP TABLE IF EXISTS Tournaments;
CREATE TABLE
    Tournaments
SELECT DISTINCT
    matches_tournament_id AS tournament_id,
    tournaments_tournament_name AS tournament_name,
    tournaments_year,
    tournaments_host_country AS host_country,
    tournaments_winner AS winner,
    tournaments_count_teams AS count_teams
FROM
    dspartidosygoles;

ALTER TABLE Tournaments
ADD PRIMARY KEY (tournament_id);



 # Creacion de la Tabla Players
DROP TABLE IF EXISTS Players;
CREATE TABLE
    Players
SELECT DISTINCT
    squads_player_id AS player_id,
    players_family_name AS family_name,
    players_given_name AS given_name,
    players_birth_date AS birth_date,
    players_female AS female,
    players_goal_keeper AS goal_keeper,
    players_defender AS defender,
    players_midfielder AS midfielder,
    players_forward AS forward
FROM
    `dsalineacionesxtorneo-2`;

ALTER TABLE Players
ADD PRIMARY KEY (player_id);



   # Creacion de la Tabla Goals
DROP TABLE IF EXISTS Goals;
CREATE TABLE
    Goals
SELECT DISTINCT
    goals_goal_id AS goal_id,
    goals_team_id AS team_id,
    goals_player_id AS player_id,
    goals_player_team_id AS player_team_id,
    matches_tournament_id AS tournament_id,
    goals_minute_label AS minute_label,
    goals_minute_regulation AS minute_regulation,
    goals_minute_stoppage AS minute_stoppage,
```

```
    goals_match_period AS match_period,
    goals_own_goal AS own_goal,
    goals_penalty AS penalty
FROM
    dspartidosygoles

WHERE dspartidosygoles.goals_minute_regulation IS NOT NULL;

ALTER TABLE Goals
ADD PRIMARY KEY (goal_id),
ADD CONSTRAINT team_id_1 FOREIGN KEY (team_id) REFERENCES Teams(team_id),
ADD FOREIGN  KEY (player_id) REFERENCES Players(player_id),
ADD CONSTRAINT team_id_2 FOREIGN  KEY (player_team_id) REFERENCES
Teams(team_id),
ADD FOREIGN  KEY (tournament_id) REFERENCES Tournaments(tournament_id);




 # Creacion de la Tabla Matches
DROP TABLE IF EXISTS Matches;
CREATE TABLE
    Matches
SELECT DISTINCT
    matches_match_id AS match_id,
    matches_tournament_id AS tournament_id,
    matches_away_team_id AS away_team_id,
    matches_home_team_id AS home_team_id,
    matches_stadium_id AS stadium_id,
    matches_match_date AS match_date,
    matches_match_time AS match_time,
    matches_stage_name AS stage_name,
    matches_home_team_score AS home_team_score,
    matches_away_team_score AS away_team_score,
    matches_extra_time AS extra_time,
    matches_penalty_shootout AS penalty_shootout,
    matches_home_team_score_penalties AS home_team_score_penalties,
    matches_away_team_score_penalties AS away_team_score_penalties,
    matches_result AS result
FROM
    dspartidosygoles;

ALTER TABLE Matches
ADD PRIMARY KEY (match_id),
ADD FOREIGN  KEY (tournament_id) REFERENCES Tournaments(tournament_id),
ADD CONSTRAINT teamid_1 FOREIGN  KEY (away_team_id) REFERENCES
Teams(team_id),
ADD CONSTRAINT teamid_2 FOREIGN  KEY (home_team_id) REFERENCES
Teams(team_id),
ADD FOREIGN  KEY (stadium_id) REFERENCES Stadiums(stadium_id);



 # Creacion de la Tabla SQUADS
DROP TABLE IF EXISTS Squads;
CREATE TABLE
    Squads
```

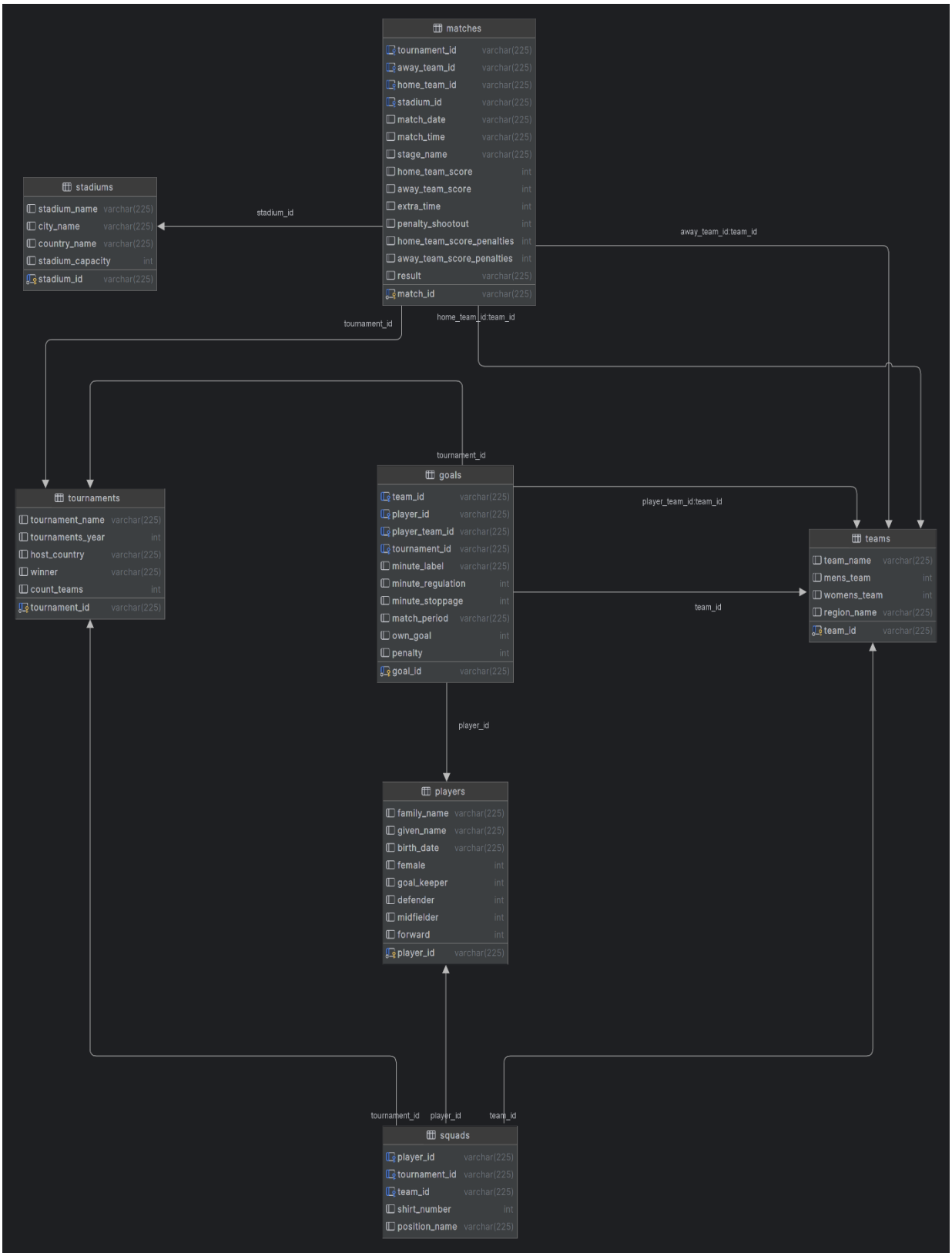
```sql
SELECT DISTINCT
    squads_player_id AS player_id,
    squads_tournament_id AS tournament_id,
    squads_team_id AS team_id,
    squads_shirt_number AS shirt_number,
    squads_position_name AS position_name
FROM
    `dsalineacionesxtorneo-2`;

ALTER TABLE Squads
ADD FOREIGN  KEY (player_id) REFERENCES Players(player_id),
ADD FOREIGN  KEY (tournament_id) REFERENCES Tournaments(tournament_id),
ADD FOREIGN  KEY (team_id) REFERENCES Teams(team_id);
```

Modelo Fisico

Como punto final del modelo físico y después de haber aplicado cada uno de los puntos anteriores a nuestra base de datos, se mostrará el diseño obtenido a través de la herramienta DataGrip la cual nos proporciona una funcionalidad para modelar nuestro diseño y así comprobar cualquier falencia en alguna relación entre tablas que se hayan podido ocasionar debido a errores de sintaxis, errores lógicos o de ejecución al haber creado nuestras sentencias.

Diseño Físico



Manipulación y Consulta de Datos

Se escribieron y ejecutaron consultas SQL para manipular y recuperar datos. Esto incluyó consultas para inserción, actualización, eliminación y selección de datos, así como consultas más avanzadas que involucraron operaciones como join, subconsultas, y funciones de agregación. Se realizó un análisis básico de los datos a través de consultas SQL.

Sentencias SQL

```
-- Selecciones que tienen equipo femenino
SELECT team_name
FROM teams
WHERE womens_team = 1;

-- Partidos jugados en cada mundial
SELECT tournament_id, COUNT(match_id) AS PartidosJugados
FROM matches
GROUP BY 1;

-- Cuantos partidos ganaron los equipos locales
SELECT COUNT(match_id) AS VictoriasLocales
FROM matches
WHERE result = 'home team win';
```

```
-- Total de goles marcados en los mundiales masculinos
SELECT t.tournament_name AS Mundiales, SUM(m.home_team_score + m.away_team_score +
                m.home_team_score_penalties + m.away_team_score_penalties) AS GolesAnotados
FROM matches m
   INNER JOIN tournaments t ON m.tournament_id = t.tournament_id
WHERE t.tournament_name LIKE '%Men\'s%'
GROUP BY 1;


-- Participaciones en mundiales de las jugadoras
SELECT CONCAT(p.given_name, ' ', p.family_name) AS NombreJugador,
      COUNT(s.tournament_id) AS ParticipacionesMundiales
FROM players p
   INNER JOIN squads s ON p.player_id = s.player_id
WHERE female = 1
GROUP BY 1;


-- Jugadores que han marcado goles en contra y el números de goles
SELECT CONCAT(p.given_name, ' ', p.family_name) AS NombreJugador, COUNT(g.own_goal) AS GolesEnContra
FROM goals g
   INNER JOIN players p ON g.player_id = p.player_id
WHERE g.own_goal != 0
GROUP BY 1;
```

```
-- Jugadores que marcaron en el mundial de Qatar con la cantidad que marcaron

SELECT CONCAT(p.given_name, ' ', p.family_name) AS NombreJugador, COUNT(g.goal_id) AS NumeroGoles

FROM Goals g

   INNER JOIN Players p ON g.player_id = p.player_id

WHERE g.tournament_id = (SELECT tournament_id FROM Tournaments WHERE host_country = 'Qatar')

GROUP BY 1

ORDER BY 1;



-- Convocatorias de Ecuador en los mundiales

SELECT t.tournament_name AS Mundial, CONCAT(p.given_name, ' ', p.family_name) AS NombreJugador,

     s.shirt_number AS Dorsal, s.position_name AS Posicion

FROM squads s

   INNER JOIN (SELECT player_id, given_name, family_name

          FROM players) p ON s.player_id = p.player_id

   INNER JOIN (SELECT tournament_id, tournament_name

          FROM tournaments) t ON s.tournament_id = t.tournament_id

WHERE s.team_id = (SELECT team_id FROM teams WHERE team_name = 'Ecuador')

ORDER BY 1, 3;



-- Inserción de datos

INSERT INTO stadiums (stadium_id, stadium_name, city_name, country_name, stadium_capacity)

VALUES ('S-241', 'Estadio Reina del Cisne', 'Loja', 'Ecuador', 15000),

     ('S-242', 'Estadio Monumental', 'Guayaquil', 'Ecuador', 59000),

     ('S-243', 'Estadio Olímpico Atahualpa', 'Quito', 'Ecuador', 41000),

     ('S-244', 'Estadio Rodrigo Paz Delgado', 'Quito', 'Ecuador', 55000)
```

```
-- Eliminación de datos
DELETE FROM stadiums WHERE city_name = 'Quito';
DELETE FROM stadiums WHERE city_name = 'Loja';
DELETE FROM stadiums WHERE city_name = 'Guayaquil';


-- Actualización de datos
UPDATE players
SET given_name = 'Richarlison', family_name = 'de Andrade'
WHERE player_id = 'P-68016';

UPDATE players
SET given_name = 'Marcelo', family_name = 'da Silva'
WHERE player_id = 'P-70650';

UPDATE players
SET given_name = 'Romário', family_name = 'de Souza'
WHERE player_id = 'P-61251';
```

Propuesta Base de Datos NOSQL

Se elaboró una propuesta borrador sobre cómo se verían los datos de la base de datos implementada, como un NOSQL con ayuda de la herramienta Neo4J el cual utiliza el lenguaje Cypher.

Creación de los nodos de ejemplo con sus propiedades:

```
CREATE (:Tournaments {
        tournament_id: "WC-2022",
        tournament_name: "2022 FIFA Men's World Cup",
        year: 2022,
        host_country: "Qatar",
        winner: "Argentina",
        count_teams: 32
})

CREATE (:Tournaments {
        tournament_id: "WC-2014",
        tournament_name: "2014 FIFA Men's World Cup",
        year: 2014,
        host_country: "Brazil",
        winner: "Germany",
        count_teams: 32
})
```

```
CREATE (:Stadiums {
        stadium_id: "S-140",
        stadium_name: "Al Bayt Stadium",
        city_name: "Al Khor",
        country_name: "Qatar",
        capacity: 69000
})

CREATE (:Teams {
        team_id: "T-59",
        team_name: "Qatar",
        mens_team: 1,
        womens_team: 0,
        region_name: "Middle East"
})

CREATE (:Teams {
        team_id: "T-25",
        team_name: "Ecuador",
        mens_team: 1,
        womens_team: 1,
        region_name: "South America"
})
```

```
CREATE (:Matches {
        match_id: "M-2022-01",
        match_date: "20/11/2022",
        match_time: "19:00:00",
        stage_name: "group stage",
        result: "away team win",
        home_team_score: 0,
        away_team_score: 2,
        extra_time: 0,
        penalty_shootout: 0,
        home_team_score_penalty: 0,
        away_team_score_penalty: 0
})

CREATE (:Players {
        player_id: "P-75912",
        family_name: "Valencia",
        given_name: "Enner",
        birth_date: "04/11/1989",
        female: 0,
        goal_keeper: 0,
        defender: 0,
        midfielder: 0,
        forward: 1})

CREATE (:Goals {
        goal_id: "G-3467",
        minute_label: "31'",
        minute_regulation: 31,
        minute_stoppage: 0,
```

```
        match_period: "first half",
        own_goal: 0,
        penalty: 0})

CREATE (:Goals {
        goal_id: "G-3466",
        minute_label: "16'",
        minute_regulation: 16,
        minute_stoppage: 0,
        match_period: "first half",
        own_goal: 0,
        penalty: 1})
```

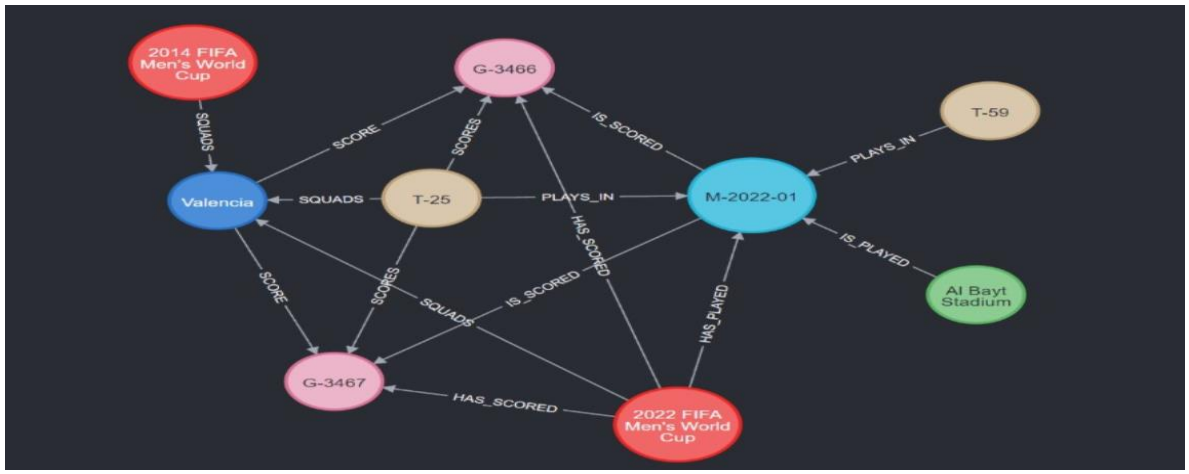
Creación de las relaciones entre los nodos:

```
MATCH (t1:Tournaments {tournament_id: "WC-2022"}),
        (s1:Stadiums {stadium_id: "S-140"}),
        (tm1:Teams {team_id: "T-25"}),
        (tm2:Teams {team_id: "T-59"}),
        (m1:Matches {match_id: "M-2022-01"})
CREATE (t1) - [:HAS_PLAYED] -> (m1),
        (s1) - [:IS_PLAYED] -> (m1),
        (tm1) - [:PLAYS_IN] -> (m1),
        (tm2) - [:PLAYS_IN] -> (m1);
```

```
MATCH (t1:Tournaments {tournament_id: "WC-2022"}),
        (tm1:Teams {team_id: "T-25"}),
        (m1:Matches {match_id: "M-2022-01"}),
        (p1:Players {player_id: "P-75912"}),
        (g1:Goals {goal_id: "G-3466"}),
        (g2:Goals {goal_id: "G-3467"})
CREATE (t1) - [:HAS_SCORED] -> (g1),
        (tm1) - [:SCORES] -> (g1),
        (m1) - [:IS_SCORED] -> (g1),
        (p1) - [:SCORE] -> (g1),
        (t1) - [:HAS_SCORED] -> (g2),
        (tm1) - [:SCORES] -> (g2),
        (m1) - [:IS_SCORED] -> (g2),
        (p1) - [:SCORE] -> (g2);


MATCH (t1:Tournaments {tournament_id: "WC-2022"}),
    (t2:Tournaments {tournament_id: "WC-2014"}),
    (p1:Players {player_id: "P-75912"}),
    (tm1:Teams {team_id: "T-25"})
CREATE (t1)-[:SQUADS]->(p1),
     (t2)-[:SQUADS]->(p1),
     (tm1)-[:SQUADS]->(p1);
```

Representación Grafica de Ejemplos de los Nodos y sus Conexiones:



Estructura de los Nodos y sus Conexiones: