

TRABAJO FINAL SQL SERVER

Tema: 50 años de estadísticas de jugadores mundialistas en todos los Mundiales masculinos desde 1966 hasta 2018

Estudiante: Peña Peña, Alvaro Rodrigo

Base de datos

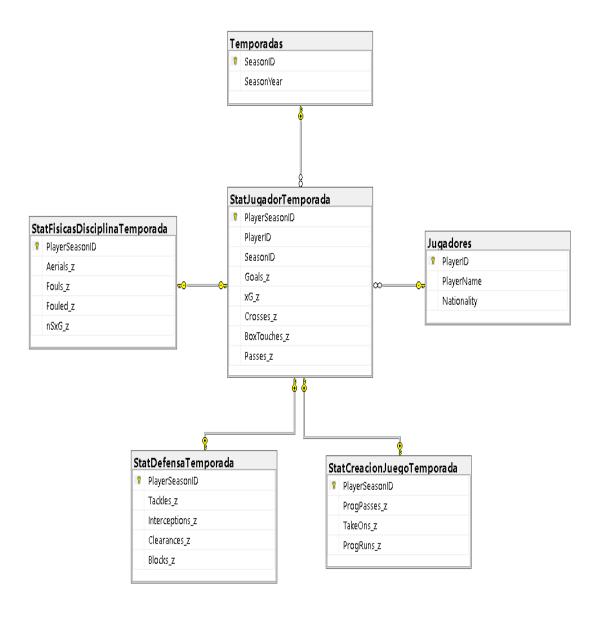
La base de datos fue nombrada "worldcup" y esta compuesta de tablas que provienen del dataset "FIFA World Cups"

URL: https://www.kaggle.com/datasets/joebeachcapital/fifa-world-cups

La fuente contiene data de 5886 estadisticas de jugadores. El anexo uno muestra el trabajo realizado para generar las tablas. A partir de esta data, se generaron las siguientes tablas:

- Jugadores: Contiene información detallada sobre los jugadores que participaron en la Copa del Mundo. Incluye datos como el nombre del jugador, su nacionalidad y otra información relevante.
- Temporadas: En esta tabla se registran las temporadas de la Copa del Mundo a lo largo de los años. Cada registro representa un año específico en el que se llevó a cabo el torneo.
- StatJugadorTemporada: Esta tabla almacena estadísticas individuales de los jugadores en cada temporada de la Copa del Mundo. Incluye métricas relacionadas con el rendimiento de los jugadores en el torneo, como goles, pases, regates y más.
- StatCreacionJuegoTemporada: En esta tabla se registran estadísticas específicas relacionadas con la creación de juego de los jugadores en cada temporada de la Copa del Mundo. Esto puede incluir métricas como pases progresivos y regates.
- StatDefensaTemporada: Contiene estadísticas relacionadas con la defensa de los jugadores en cada temporada de la Copa del Mundo. Esto podría incluir métricas como tackles, bloqueos e intercepciones.
- StatFisicasDisciplinaTemporada: Esta tabla almacena estadísticas físicas y disciplinarias de los jugadores en cada temporada del torneo. Puede incluir métricas como duelos ganados, faltas cometidas y tarjetas recibidas.

Diagrama relacional



Vistas

Vista 1: Rendimiento Promedio de los Jugadores por Temporada

Esta vista calcula el rendimiento promedio de los jugadores en términos de goles, pases y regates por temporada de la Copa del Mundo. Esto puede ayudar a analizar la evolución del rendimiento de los jugadores a lo largo de diferentes temporadas.

```
CREATE VIEW RendimientoPromedioJugadoresPorTemporada AS

SELECT
    T.SeasonYear,
    AVG(SJT.Goals_z) AS PromedioGoles,
    AVG(SJT.Passes_z) AS PromedioPases,
    AVG(SCT.TakeOns_z) AS PromedioRegates

FROM Temporadas T

INNER JOIN StatJugadorTemporada SJT ON T.SeasonID = SJT.SeasonID

INNER JOIN StatCreacionJuegoTemporada SCT ON SJT.PlayerSeasonID = SCT.PlayerSeasonID

GROUP BY T.SeasonYear;
```

	SeasonYear	PromedioGoles	PromedioPases	PromedioRegates
1	1998	0.003677	-0.000018	-0.003602
2	1978	0.003082	-0.000225	-0.000563
3	1990	-0.001568	-0.000102	0.001362
4	2010	-0.001981	0.000000	-0.001945
5	1970	-0.002037	-0.000377	-0.000830
6	2002	0.002718	0.000037	-0.001247
7	1982	0.002540	0.000080	-0.000213
8	1994	-0.002342	0.000125	0.001788
9	2014	-0.000842	0.000053	-0.001594
10	1974	-0.001860	0.000193	-0.000465
11	1986	0.002130	0.000100	-0.000676
12	2006	0.001610	-0.000017	0.000769
13	1966	-0.003700	0.000078	0.001220
14	2018	-0.004387	0.000274	-0.001736

Vista 2: Estadísticas de Defensa de los Jugadores por Nacionalidad

Esta vista proporciona estadísticas de defensa promedio de los jugadores, incluyendo tackles, intercepciones, clearance y blocks, agrupados por nacionalidad. Esto permite identificar qué nacionalidades tienen jugadores con un mejor desempeño defensivo.

```
CREATE VIEW EstadisticasDefensaPorNacionalidad AS
SELECT
    J.Nationality,
    AVG(SDT.Tackles_z) AS PromedioTackles,
    AVG(SDT.Interceptions_z) AS PromedioIntercepciones,
    AVG(SDT.Clearances_z) AS PromedioClearances,
    AVG(SDT.Blocks_z) AS PromedioBlocks
FROM Jugadores J
INNER JOIN StatDefensaTemporada SDT ON J.PlayerID = SDT.PlayerSeasonID
GROUP BY J.Nationality;
SELECT * FROM EstadisticasDefensaPorNacionalidad;
```

	Nationality	PromedioTackles	PromedioIntercepciones	PromedioClearances	PromedioBlocks
1	South Korea	0.006711	-0.074765	0.041476	0.053892
2	Czechoslovakia	-0.196851	-0.140925	-0.113703	-0.013888
3	USA	-0.100775	-0.187844	0.066293	0.021724
4	New Zealand	-0.590740	-0.464074	0.411481	-0.108148
5	Angola	-0.445625	-0.245000	0.191250	-0.170000
6	Egypt	-0.149032	0.103225	0.478709	-0.171290
7	Scotland	-0.086836	-0.247653	-0.118061	-0.017653
8	Italy	-0.097719	-0.033464	-0.052807	0.061096
9	Rep of Ireland	-0.256666	-0.045333	0.025111	-0.196222
10	Brazil	0.339316	0.338846	-0.107521	0.002179
11	Netherlands	0.240000	0.268540	-0.101532	-0.029854
12	USSR	0.128863	0.140909	-0.147386	-0.064431
13	Cameroon	-0.115752	-0.041415	-0.074778	0.129380
14	Bolivia	-0.033750	-0.162500	0.158750	0.112500
15	Jamaica	-0.182941	-0.118235	0.054117	-0.174705
16	Germany	0.362015	0.136511	-0.130155	0.057364
17	England	-0.136178	-0.159842	-0.089109	0.075078
18	Iceland	-0.629333	-0.490000	0.377333	-0.156000
19	Senegal	-0.368571	-0.340857	-0.042285	-0.258857
20	Kuwait	0.217142	0.170714	-0.078571	-0.124285
21	Hungary	-0.048787	-0.220606	-0.057121	-0.223636
22	Switzerland	0.193894	0.286000	0.017789	0.048421
23	Algeria	-0.169076	-0.098923	0.232000	0.045538
24	Ecuador	0.122553	0.095744	0.160425	0.141489
25	Nigeria	-0.190388	-0.176019	0.059417	0.072038
26	El Salvador	-0.319677	-0.071612	0.431290	0.110000

Preguntas

Pregunta 1. Rank de los mejores promedios goleadores en cada temporada de la Copa del Mundo

Esta pregunta busca identificar a los máximos goleadores en cada temporada de la Copa del Mundo y clasificarlos en función de la cantidad de goles que anotaron.

La consulta utiliza dos estructuras de tabla común de expresión (CTE) para calcular el ranking de goleadores en cada temporada. Primero, se suman los goles de cada jugador por temporada y luego se clasifican los jugadores en función de sus goles, identificando así a los mejores goleadores de cada edición del torneo.

```
WITH MaxGoleadoresPorTemporada AS (
    SELECT
        T.SeasonYear AS Temporada,
        J.PlayerName AS Jugador,
        SUM(E.Goals_z) AS Goles
    FROM StatJugadorTemporada E
    INNER JOIN Jugadores J ON E.PlayerID = J.PlayerID
    INNER JOIN Temporadas T ON E.SeasonID = T.SeasonID
    GROUP BY T.SeasonYear, J.PlayerName
MaxGoleadorPorTemporada AS (
    SELECT
        Temporada,
        Jugador,
        Goles,
        RANK() OVER (PARTITION BY Temporada ORDER BY Goles DESC) AS Ranking
    FROM MaxGoleadoresPorTemporada
SELECT
    Temporada,
    Jugador,
   Goles
FROM MaxGoleadorPorTemporada
WHERE Ranking = 1
ORDER BY Goles DESC;
```

	Temporada	Jugador	Goles
1	1994	Oleg Salenko	10.12
2	2018	Harry Kane	8.87
3	1966	Eusébio	7.50
4	2014	James RodrÃ-guez	7.45
5	1970	Gerd Müller	7.13
6	2002	Ronaldo	6.85
7	1986	Gary Lineker	6.84
8	1990	TomÃjs SkuhravÃ1⁄2	6.53
9	2010	Rðbert Vittek	6.51
10	1974	Grzegorz Lato	5.97
11	1998	Luis Hernández	5.76
12	1998	Marcelo Salas	5.76
13	1982	Laszlo Kiss	5.47
14	2006	David Villa	5.22
15	2006	Fernando Torres	5.22
16	1978	Teúfilo Cubillas	5.00

Pregunta 2. Top 5 selecciones con el mayor promedio de gol en todas las ediciones del mundial

Esta pregunta tiene como objetivo identificar las 5 selecciones nacionales que tienen el mayor promedio de goles en todas las ediciones de la Copa del Mundo.

La consulta SQL calcula el total de goles (Goals_z) anotados por cada jugador y luego agrupa los resultados por nacionalidad del equipo (J.Nationality). A continuación, se ordenan los equipos en orden descendente según la cantidad total de goles y se seleccionan los 5 equipos principales utilizando la cláusula TOP 5. Esto proporciona una lista de las 5 selecciones con el mayor promedio de goles en todas las ediciones del torneo.

SELECT TOP 5 J.Nationality AS Equipo, SUM(E.Goals_z) AS TotalGoles FROM StatJugadorTemporada E
INNER JOIN Jugadores J ON E.PlayerID = J.PlayerID
GROUP BY J.Nationality
ORDER BY TotalGoles DESC

	Equipo	TotalGoles
1	Brazil	57.32
2	West Germany	30.44
3	Germany	25.97
4	Argentina	24.96
5	Spain	24.61

Pregunta 3. ¿Cuáles son los 10 jugadores con el mejor promedio de bloqueos en la historia de la Copa del Mundo?

Esta pregunta busca identificar a los 10 jugadores con el mejor promedio de bloqueos en todas las ediciones de la Copa del Mundo, lo que demuestra su habilidad en la defensa.

La consulta SQL calcula el total de bloqueos (Blocks_z) realizados por cada jugador y luego agrupa los resultados por nombre de jugador (J.PlayerName). Los jugadores se ordenan en orden descendente según la cantidad total de bloqueos, y se seleccionan los 10 mejores jugadores utilizando la cláusula TOP 10. Esto proporciona una lista de los 10 jugadores con el mejor promedio de bloqueos en la historia de la Copa del Mundo.

```
SELECT TOP 10 J.PlayerName AS Jugador, SUM(E.Blocks_z) AS TotalBloqueos
FROM StatDefensaTemporada E
INNER JOIN Jugadores J ON E.PlayerSeasonID = J.PlayerID
GROUP BY J.PlayerName
ORDER BY TotalBloqueos DESC;
```

	Jugador	TotalBloqueos
1	John Terry	14.50
2	Elias Figueroa	13.00
3	Myung-Bo Hong	12.90
4	Carlos Salcedo	11.57
5	Gaetano Scirea	9.63
6	Javier Mascherano	9.12
7	Carlos Gamarra	9.07
8	Alberto Quintano	9.02
9	Steven Gerrard	8.92
10	Franco Baresi	8.60

Pregunta 4. ¿Cuáles son las 10 nacionalidad cuyos jugadores han recibido la mayor cantidad de faltas en todas las ediciones de la Copa del Mundo?

Esta pregunta tiene como objetivo determinar la nacionalidad cuyos jugadores han recibido la mayor cantidad de faltas en todas las ediciones de la Copa del Mundo, lo que puede indicar la agresividad de los oponentes contra los jugadores de esa nacionalidad.

La consulta SQL calcula la suma de faltas recibidas (Fouled_z) por cada jugador y luego agrupa los resultados por la nacionalidad del equipo (J.Nationality). Los equipos se ordenan en orden descendente según la cantidad total de faltas recibidas, y se selecciona la nacionalidad con la mayor cantidad utilizando la cláusula TOP 10. Esto proporciona la nacionalidad cuyos jugadores han recibido la mayor cantidad de faltas en la historia de la Copa del Mundo.

```
SELECT TOP 10 J.Nationality AS Nacionalidad, SUM(E.Fouled_z) AS TotalFaltasRecibidas
FROM StatFisicasDisciplinaTemporada E
INNER JOIN Jugadores J ON E.PlayerSeasonID = J.PlayerID
GROUP BY J.Nationality
ORDER BY TotalFaltasRecibidas DESC;
```

	Nacionalidad	TotalFaltasRecibidas
1	Argentina	47.13
2	Italy	32.40
3	Uruguay	29.70
4	Japan	24.63
5	Mexico	18.33
6	Chile	17.08
7	West Germany	16.86
8	Yugoslavia	14.57
9	Iran	11.12
10	Bolivia	9.81

Pregunta 5. ¿Cuál es el jugador que ha tenido el mejor desempeño en términos de regates y pases progresivos combinados en la historia de la Copa del Mundo, considerando solo a los jugadores de equipos que llegaron a la final del torneo?

Esta pregunta tiene como objetivo identificar al jugador que ha tenido el mejor desempeño en términos de regates y pases progresivos combinados en la historia de la Copa del Mundo, pero considerando solo a los jugadores que formaron parte de equipos que llegaron a la final del torneo. Esto puede revelar quién destacó en términos de creatividad y avance del juego en las ediciones finales del torneo.

La consulta SQL utiliza varias estructuras de tabla común de expresión (CTE) para calcular el total de regates y pases progresivos para cada jugador en cada temporada de la Copa del Mundo desde 1990, y luego clasifica a los jugadores en función de su desempeño en estas métricas. Aquí se consideran solo las ediciones de la Copa del Mundo a partir de 1990.

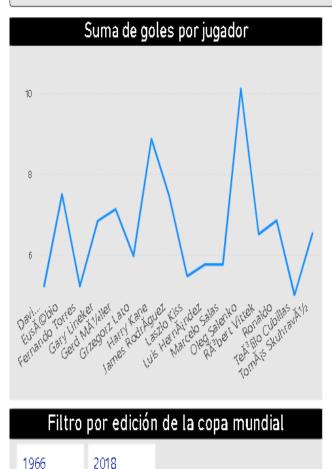
La CTE EquiposFinalistas identifica los equipos finalistas en cada temporada desde 1990, y la CTE DesempenoJugadores calcula el desempeño de regates y pases progresivos para cada jugador en esas temporadas. La CTE RankingDesempeno asigna un ranking a los jugadores en función de su desempeño en cada temporada.

```
WITH EquiposFinalistas AS (
   SELECT DISTINCT T.SeasonYear, J.Nationality
   FROM StatJugadorTemporada E
   INNER JOIN Jugadores J ON E.PlayerID = J.PlayerID
   INNER JOIN Temporadas T ON E.SeasonID = T.SeasonID
   WHERE T.SeasonYear >= 1990
DesempenoJugadores AS (
   SELECT
       J.PlayerName,
       T.SeasonYear,
       SUM(E.Takeons_z + E.ProgPasses_z) AS TotalRegatesYProgresivos
    FROM StatCreacionJuegoTemporada E
   INNER JOIN Jugadores J ON E.PlayerSeasonID = J.PlayerID
   INNER JOIN Temporadas T ON E.PlayerSeasonID = T.SeasonID
   WHERE T.SeasonYear >= 1990
   GROUP BY J.PlayerName, T.SeasonYear
RankingDesempeno AS (
   SELECT
       DJ.PlaverName.
       DJ.SeasonYear,
       DJ.TotalRegatesYProgresivos,
       RANK() OVER (PARTITION BY DJ.SeasonYear ORDER BY DJ.TotalRegatesYProgresivos DESC) AS Ranking
   FROM DesempenoJugadores DJ
   INNER JOIN EquiposFinalistas EF ON DJ.SeasonYear = EF.SeasonYear AND DJ.PlayerName
   IN (SELECT J.PlayerName FROM Jugadores J WHERE J.Nationality = EF.Nationality)
SELECT
   RD.PlayerName AS MejorJugador,
   RD.SeasonYear AS Temporada,
   RD.TotalRegatesYProgresivos AS RegatesYProgresivos
FROM RankingDesempeno RD
WHERE RD.Ranking = 1
ORDER BY RD.TotalRegatesYProgresivos DESC;
```

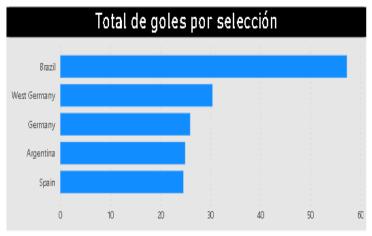
	MejorJugador	Temporada	RegatesYProgresivos
1	Lionel Messi	2010	9.56
2	Isco	2018	8.13
3	Hatem Trabelsi	2006	8.11
4	Jay-Jay Okocha	2002	7.34
5	Lionel Messi	2014	6.99
6	Harold Lozano	1998	6.57
7	Rafael Martin-Vasquez	1990	6.34
8	Fernando Redondo	1994	6.27

DASHBOARD

50 años de estadísticas de jugadores mundialistas en todos los Mundiales masculinos desde 1966 hasta 2018 Número de registros 5886







ANEXO 1 – Creación de tablas y carga de registros

Dado que el dataset original no incluía las tablas necesarias para el análisis, y toda la información estaba consolidada en una sola tabla, decidí en primer lugar crear mis propias tablas después de analizar el contenido de la tabla fuente. Una vez que tenía un modelo listo, agregué un identificador a la tabla principal. Este identificador me sería de utilidad posteriormente para extraer los datos de la tabla fuente y transferirlos a las demás tablas que había creado.

Creación de las tablas:

```
-- Crear la tabla Jugadores
CREATE TABLE Jugadores (
    PlayerID INT PRIMARY KEY,
    PlayerName VARCHAR(255),
    Nationality VARCHAR(255)
-- Crear la tabla Temporadas
CREATE TABLE Temporadas (
    SeasonID INT PRIMARY KEY,
    SeasonYear INT
);
-- Crear la tabla StatJugadorTemporada
CREATE TABLE StatJugadorTemporada (
    PlayerSeasonID INT PRIMARY KEY,
    PlayerID INT,
    SeasonID INT,
    Goals z DECIMAL(10, 2),
    xG_z DECIMAL(10, 2),
    Crosses_z DECIMAL(10, 2),
    BoxTouches_z DECIMAL(10, 2),
    Passes_z DECIMAL(10, 2),
    FOREIGN KEY (PlayerID) REFERENCES Jugadores(PlayerID),
    FOREIGN KEY (SeasonID) REFERENCES Temporadas(SeasonID)
-- Crear la tabla StatCreacionJuegoTemporada
CREATE TABLE StatCreacionJuegoTemporada (
    PlayerSeasonID INT PRIMARY KEY,
    ProgPasses_z DECIMAL(10, 2),
    TakeOns_z DECIMAL(10, 2),
    ProgRuns_z DECIMAL(10, 2),
    FOREIGN KEY (PlayerSeasonID) REFERENCES StatJugadorTemporada(PlayerSeasonID)
-- Crear la tabla StatDefensaTemporada
CREATE TABLE StatDefensaTemporada (
   PlayerSeasonID INT PRIMARY KEY,
    Tackles_z DECIMAL(10, 2),
    Interceptions_z DECIMAL(10, 2),
    Clearances_z DECIMAL(10, 2),
    Blocks_z DECIMAL(10, 2),
    FOREIGN KEY (PlayerSeasonID) REFERENCES StatJugadorTemporada(PlayerSeasonID)
-- Crear la tabla StatFisicasDisciplinaTemporada
CREATE TABLE StatFisicasDisciplinaTemporada (
    PlayerSeasonID INT PRIMARY KEY,
    Aerials_z DECIMAL(10, 2),
    Fouls_z DECIMAL(10, 2),
   Fouled_z DECIMAL(10, 2),
   nSxG_z DECIMAL(10, 2),
    FOREIGN KEY (PlayerSeasonID) REFERENCES StatJugadorTemporada(PlayerSeasonID)
);
```

Cargando los registros a las tablas:

```
-- Creando columna ID para la tabla fuente
ALTER TABLE dbo.world cup comparisons
ADD wcID INT;
G0
--Agregendo IDs para la tabla dbo.world cup comparisons
WITH NumberedWC AS (
    SELECT wcID.
           ROW_NUMBER() OVER (ORDER BY (SELECT NULL)) AS NewCovidDataID
    FROM dbo.world_cup_comparisons
UPDATE NumberedWC
SET wcID = NewCovidDataID;
--Agregano datos de los jugadores
INSERT INTO Jugadores (PlayerID, PlayerName, Nationality)
SELECT dbo.world_cup_comparisons.wcID, dbo.world_cup_comparisons.player, dbo.world_cup_comparisons.team
FROM dbo.world_cup_comparisons;
-- Agregando datos Temporadas
INSERT INTO Temporadas (SeasonID , SeasonYear)
SELECT dbo.world_cup_comparisons.wcID, dbo.world_cup_comparisons.season
FROM dbo.world cup comparisons;
-- Agregando datos StatJugadorTemporada
INSERT INTO
StatJugadorTemporada (PlayerSeasonID , PlayerID, SeasonID, Goals_z, xG_z, Crosses_z, BoxTouches_z, Passes_z)
dbo.world_cup_comparisons.wcID, dbo.world_cup_comparisons.wcID, dbo.world_cup_comparisons.wcID,
dbo.world cup comparisons.goals z, dbo.world cup comparisons.xg z, dbo.world cup comparisons.crosses z,
dbo.world_cup_comparisons.boxtouches_z, dbo.world_cup_comparisons.passes_z
FROM
dbo.world cup comparisons;
-- Agregando datos Temporadas
INSERT INTO StatCreacionJuegoTemporada (PlayerSeasonID , ProgPasses_z, TakeOns_z, ProgRuns_z)
SELECT dbo.world_cup_comparisons.wcID, dbo.world_cup_comparisons.progpasses_z, dbo.world_cup_comparisons.takeons_z,
dbo.world_cup_comparisons.progruns_z
FROM dbo.world_cup_comparisons;
-- Agregando datos StatDefensaTemporada
INSERT INTO StatDefensaTemporada (PlayerSeasonID , Tackles_z, Interceptions_z, Clearances_z, Blocks_z)
SELECT dbo.world_cup_comparisons.wcID, dbo.world_cup_comparisons.tackles_z, dbo.world_cup_comparisons.interceptions_z,
dbo.world_cup_comparisons.clearances_z, dbo.world_cup_comparisons.blocks_z
FROM dbo.world cup comparisons;
-- Agregando datos StatFisicasDisciplinaTemporada
INSERT INTO StatFisicasDisciplinaTemporada (PlayerSeasonID , Aerials_z, Fouls_z, Fouled_z, nSxG_z)
SELECT dbo.world_cup_comparisons.wcID, dbo.world_cup_comparisons.aerials_z, dbo.world_cup_comparisons.fouls_z,
dbo.world_cup_comparisons.fouled_z, dbo.world_cup_comparisons.nsxg_z
FROM dbo.world_cup_comparisons;
```