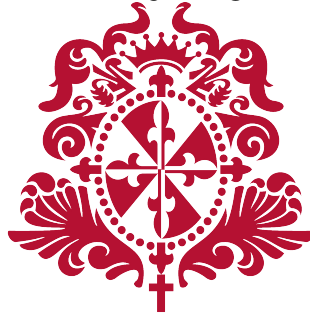


LIGA COLOMBIANA
DE FÚTBOL



UNIVERSIDAD DEL ROSARIO
MATEMÁTICAS APLICADAS Y CIENCIAS DE LA
COMPUTACIÓN

INGENIERIA DE DATOS

Miguel Salamanca,
Juan Simbaqueba,
Camila Pataroyo

Proyecto Final

November 19, 2022

TABLA DE CONTENIDO

1	Problemática	1
2	Diseño	1
2.1	Diseño Conceptual(Diagrama entidad relación)	1
2.2	Diseño Lógico(Diagrama relacional)	2
3	Normalización	3
3.1	Reglas para 1NF	3
3.2	Regla para 2NF	3
3.3	Reglas para 3NF	4
4	Inserción Datos	4
5	Consultas	6
5.1	Consulta 1 - Información de los clubes	6
5.2	Consulta 2 - Información de los jugadores	7
5.3	Consulta 3 - Información de las estadísticas de cada uno de los jugadores.	7
5.4	Consulta 4 - Los 13 máximos goleadores del torneo.	8
5.5	Consulta 5 - Los 13 máximos asistentes del torneo.	9
5.6	Consulta 6 - Los colores característicos de cada club.	10
5.7	Consulta 7 - Presupuesto disponible para cada uno de los direc- tores técnicos.	10
5.8	Consulta 8 - Jugadores más influyentes.	11
6	Visualización	11
6.1	Goleadores	11
6.2	Asistentes	12
6.3	Presupuesto director técnico	12
7	Análisis	12
8	Conclusión	13

1 Problemática

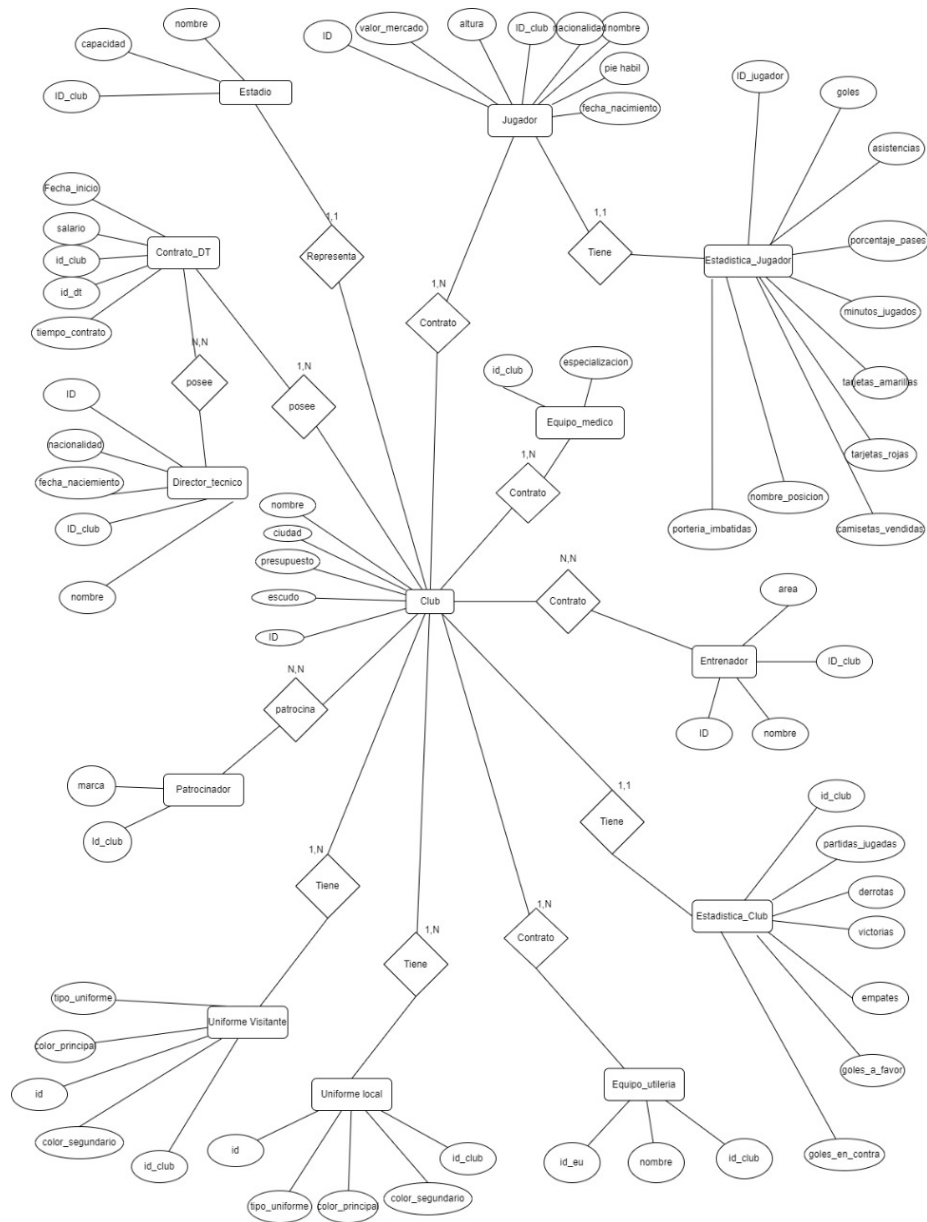
Esta propuesta busca facilitar la búsqueda de información por parte de los usuarios que necesiten o quieran por su propia cuenta, obtener datos acerca de su liga de preferencia. Hoy en día hay muchos usuarios dentro del mundo del futbol que quisieran obtener una gran cantidad de datos para realizar un buen análisis sobre su equipo o sobre algún jugador, ya sea solo para adquirir conocimiento o para algún otro tema que incluya un objetivo monetario para esta persona, como lo es el mundo de las apuestas, o también si se necesitan sacar datos para los mismos funcionarios del club para manejar el tema de las finanzas e igualmente para los funcionarios de las propias ligas. En cuanto al mundo de las apuestas, existen muchas personas que basan su economía en esta actividad, ya sea por parte de las casas de apuestas como también de sus usuarios, el hecho de tener una base de datos que contenga cada estadística del club y sus jugadores permitirá una mejor predicción y facilitara el hecho de sacar un indicador de multiplicación para las apuestas de cada partido. Como segunda problemática a superar, actualmente cada equipo de futbol mueve mucho dinero fuera de sus partidos, la realización de esta propuesta busca facilitar para los usuarios la búsqueda de cualquier producto que quieran adquirir que esté relacionado con el club, ya sean sus camisetas, sus sudaderas, sus equipaciones, o también las boletas de los partidos, sin tener que salir del lugar donde encuentran las estadísticas y resultados del club. Todo al alcance de un clic desde la comodidad ya sea de su celular, computador o tableta.

2 Diseño

En las siguientes dos subsecciones se encuentran los modelos de datos que definen como se modela la estructura lógica y conceptual de la base de datos de la liga de futbol colombiana. Estos modelos son entidades importantes para introducir la abstracción en una base de datos, como se conectan entre sí, como se procesan y almacenan dentro del sistema. También ayudan a evitar la redundancia de datos porque se implementará la normalización 3NF.

2.1 Diseño Conceptual(Diagrama entidad relación)

En el modelo conceptual fue diseñado en la plataforma ERDPlus, el modelo de datos identifica las relaciones de más alto nivel entre las diferentes entidades. Incluye las entidades importantes y las relaciones entre ellas, como lo son club, jugador, estadísticas jugador, estadísticas club, entrenador, director técnico entre otras.



2.2 Diseño Lógico(Diagrama relacional)

En el modelo de datos lógico fue diseñado en PGModeler, el modelo describe los datos con el mayor detalle posible, independientemente de cómo se implementarán físicamente en la base de datos. Incluye todas las tablas y atributos que se relacionan entre ellos, cada atributo está especificado con su tipo y en

[illegible]

Para la edición del plano inicial de las tablas tuvimos en cuenta la normalización 3NF que se estructura de la siguiente forma:

1. Cada registro debe ser único, sin filas duplicadas
2. Cada celda debe contener un valor, debe verificar que no existan atributos multivalor

3.2 Regla para 2NF

1. Debe satisfacer 1NF
2. Si la clave principal (PK) de una tabla consta de uno y solo una columna, automáticamente satisface 2NF
3. Si hay una clave principal compuesta, cada columna no clave debe depender de toda la llave. Se deben eliminar dependencias parciales de la llave.

3

primaria de esta tabla.







public.Estadisticas_club		
 id_club	smallint	« i »
 partidos_jugados	smallint	
 victorias	smallint	
 empates	smallint	
 derrotas	smallint	
 goles_a_favor	smallint	
 goles_en_contra	smallint	
 id_estadisticas_club	smallint	« p »
 id_club	constraint	« fk »
 Estadisticas_club_pk	constraint	« pk »







3.3 Reglas para 3NF

1. Satisfacer 2NF
2. Sin dependencias transitivas: Las columnas no claves (que no forman parte de la PK), no pueden depender de otras columnas que tampoco sean no clave.

En nuestro caso en específico inicialmente teníamos una tabla llamada uniforme con los siguientes atributos: id club, tipo de uniforme, color primario, color secundario

Note que hay una dependencia transitiva entre los colores, el tipo y el id del club, ya que un club tiene dos tipos de uniformes y varios colores pueden ser de muchos clubs y tipos de uniforme. Por lo que decidimos crear dos tablas, uniforme visitante y uniforme local el cual solucionó el problema de transitividad.

public.uniforme_visitante		
 id_uv	smallint	« p »
 color_principal	varchar(30)	
 color_secundario	varchar(30)	
 id_club	smallint	« i »
 id	constraint	« pk »
 id_club	constraint	« fk »

public.uniforme_local		
 id_lo	smallint	« »
 color_principal	varchar(30)	
 color_secundario	varchar(30)	
 id_club	smallint	« »
 id_club	constraint	« fk »
 uniforme_local_pk	constraint	« pk »

4 Inserción Datos

Web scraping, es una técnica utilizada mediante Python para extraer información de sitios web, en nuestro caso en específico extrajimos los datos de TransferMarket. El web scraping está muy relacionado con la indexación de la web, la cual indexa la información de la web utilizando la librería requests,

se enfoca más en la transformación de datos sin estructura en la web (como el formato HTML) en datos estructurados que pueden ser almacenados y analizados en la base de datos liga, en una archivo CSV. El uso que le dimos a esta herramienta fue la extracción de los datos de las tablas principales y con mayor cantidad de datos que era muy complicado e ineficiente extraerlos a mano ya que gracias a su capacidad de generar grandes cantidades de datos para crear contenidos de calidad pudimos extraer con éxito todos los datos para poder hacer un mayor análisis.

```

1  import requests
2  from bs4 import BeautifulSoup
3  import pandas as pd
4
5  """
6  To make the request to the page we have to inform the site that we are a browser
7  and that's why we use the variable headers
8  """
9  headers = {'User-Agent': 'Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/47.0.2526.106 Safari/537.36'}
10
11 # page_address represents the link that directs to the page with the data
12 page_address = "https://www.transfermarkt.com/liga-dinamo-1/startseite/wettbewerb/COLP"
13
14 # in object_response we will download the web page
15 object_response = requests.get(page_address, headers=headers)
16
17 """
18 Now we will create a BeautifulSoup object from our object_response.
19 The 'html.parser' parameter represents which parser we will use when creating our object,
20 a parser is a software responsible for performing the conversion of an input to a data structure.
21 """
22 page_bs = BeautifulSoup(object_response.content, 'html.parser')
23
24 # First we will get the names of each country in the competition
25 names_clubs = [] # Sorted list of all clubs' name
26
27 # The find_all() method is able to return all tags that comply with the constraints inside the parentheses
28 tags_clubs = page_bs.find_all("td", {"class": "zentriert no-border-rechts"})
29 # In our case we are finding all the anchors with the class 'vereinprofil_tooltip'
30
31 # Dictionaries
32 goalkeepers = ["Partidos jugados", "En alineación", "Puntos por juego", "Goles", "Autogoles", "Entra por sustitución",
33               "Sale por sustitución", "Amarillas", "Segundas amarillas", "Rojas", "Goles concedidos", "Porterías a cero", "Minutos jugados"]
34 players = ["Partidos jugados", "En alineación", "Puntos por juego", "Goles", "Asistencias", "Autogoles", "Entra por sustitución",
35           "Sale por sustitución", "Amarillas", "Segundas amarillas", "Rojas", "Goles de penalty", "Goles por minuto", "Minutos jugados"]
36
37 # Header
38 info = ["Name in home country", "Date of birth", "Age", "Height", "Citizenship", "Position", "Foot", "Player agent", "Joined",
39        "Contract expires", "Contract option", "Outfitter", "Social-Media"]
40
41 # File headers
42 info_headers = info + goalkeepers + players
43
44 # Delete duplicate
45 my_headers = []
46 for i in info_headers:
47     if i not in my_headers:
48         my_headers.append(i)
49
50

```

Luego de extraer los datos continuamos con la inserción de ellos mediante la siguiente imagen, la cual leía archivos csv donde se almacenaban gran cantidad de datos.

```

18 -----INSERCCIONES-----
11 COPY PUBLIC.club(id_cl,nombre,ciudad,presupuesto_en_miles,id_estadio)
12 FROM 'C:\Users\Laboratorios EICT\Documents\Liga\club.csv'
13 DELIMITER ','
14 CSV HEADER;
15
16 COPY PUBLIC.estadio(capacidad,nombre,id_es)
17 FROM 'C:\Users\Laboratorios EICT\Documents\Liga\estadio.csv'
18 DELIMITER ','
19 CSV HEADER;
20
21 COPY PUBLIC.entrenador(nombre,id_e,area,id_club)
22 FROM 'C:\Users\Laboratorios EICT\Documents\Liga\entrenador.csv'
23 DELIMITER ','
24 CSV HEADER;

```

5 Consultas

5.1 Consulta 1 - Informaci3n de los clubes

	id_cl [PK] smallint	nombre character varying (30)	ciudad character varying (30)	presupuesto_en_miles smallint	id_estadio smallint
1	101	Junior de Barranquilla	Barranquilla	2413	200
2	102	Atletico Nacional	Medellin	2238	201
3	103	Deportes Tolima	Ibague	2148	202
4	104	Millonarios FC	Bogota	1910	203
5	105	CD America de Cali	Cali	1505	204
6	106	Independiente Medellin	Medellin	1503	205
7	107	Deportivo Cali	Cali	147	206
8	108	Independiente Santa Fe	Bogota	1450	207
9	109	Once Caldas	Manizales	1055	208
10	110	Alianza Petrolera	Barrancabermeja	96	209
11	111	CD La Equidad Seguros SA	Bogota	93	210
12	112	Atletico Bucaramanga	Bucaramanga	895	211
13	113	Aguilas Doradas	Rionegro	86	212
14	114	Deportivo Pereira	Pereira	818	213
15	115	Jaguares de Cordoba	Cordoba	773	214
16	116	Asociacion Deportivo Pasto	Pasto	753	215
17	117	Envigado FC	Envigado	67	216
18	118	Boyaca Patriotas FC	Boyaca	610	217
19	119	Cortulua	Tulua	593	218
20	120	AD Union Magdalena	Santa Marta	455	219

5.2 Consulta 2 - Información de los jugadores

	id_j [PK] smallint	edad smallint	altura numeric (4)	pie_habil character varying (15)	nacionalidad character varying (30)	nombre character varying (50)	valor_mercado_en_miles integer	id_club smallint	posicion character varying (30)
1	1001	28	185	derecho	Colombia	Jefferson Martinez	300	101	Portero
2	1002	39	184	derecho	Uruguay	Sebastian Viera	275	101	Portero
3	1003	24	183	derecho	Colombia	Sebastian Araujo	150	101	Portero
4	1004	20	16	derecho	Colombia	Jaime Acosta	280	101	Portero
5	1005	24	185	derecho	Colombia	Homer Martinez	950	101	Defenza central
6	1006	29	183	izquierdo	Colombia	Jorge Arias	800	101	Defenza central
7	1007	32	187	izquierdo	Colombia	German Mera	800	101	Defenza central
8	1008	28	187	derecho	Colombia	Dany Rosero	800	101	Defenza central
9	1009	21	19	izquierdo	Colombia	Alfonso Simarra	150	101	Defenza central
10	1010	25	18	izquierdo	Colombia	Gabriel Fuentes	990	101	Lateral izquierdo
11	1011	30	179	izquierdo	Colombia	Edwin Velasco	600	101	Lateral izquierdo
12	1012	27	172	derecho	Colombia	Walmer Pacheco	150	101	Lateral derecho
13	1013	30	167	derecho	Colombia	Fabian Viazara	280	101	Lateral derecho
14	1014	26	183	derecho	Colombia	Nilson Castrillon	950	101	Lateral derecho
15	1015	29	178	derecho	Colombia	Daniel Giraldo	800	101	Pivote
16	1016	30	181	izquierdo	Colombia	Didier Moreno	800	101	Pivote
17	1017	26	178	derecho	Colombia	Enrique Serje	800	101	Pivote
18	1018	21	175	derecho	Colombia	Fabian Angel	150	101	Mediocentro
19	1019	31	168	derecho	Colombia	Jesus Cabrera	990	101	Mediocentro
20	1020	23	179	derecho	Colombia	Carlos Esparragoza	600	101	Mediocentro
21	1021	22	182	derecho	Colombia	Neider Sanjuan	700	101	Interior derecho
22	1022	33	17	derecho	Argentina	Fabian Sambueza	600	101	Mediocentro ofensivo
23	1023	31	17	derecho	Venezuela	Luis Gonzalez	100	101	Extremo izquierdo
24	1024	32	178	izquierdo	Colombia	Fredy Hinestroza	200	101	Extremo izquierdo

5.3 Consulta 3 - Información de las estadísticas de cada uno de los jugadores.

	id_jugador smallint	partidos_jugados smallint	partidos_en_alineacion smallint	puntos_por_partido numeric	goles smallint	autogoles smallint	entra_sustituido smallint	sustituido smallint	amarillos smallint	doble_amarillo smallint	rojas smallint	goles_concedidos smallint	portadas_a_cero smallint	minutos_jugados numeric	asistencias smallint	id_estadista [PK] sma
1	1001	16	2	2	0	0	0	0	0	0	0	1	1	180	0	
2	1002	17	17	2	0	0	0	0	1	0	0	16	5	2	0	
3	1003	3	0	0	0	0	0	0	0	0	0	0	0	1000	0	
4	1004	2	0	1	0	0	0	0	0	0	0	0	0	1001	0	
5	1005	11	11	2	0	0	0	4	2	0	0	0	0	705	0	
6	1006	13	12	2	0	0	0	0	1	3	0	0	0	1	0	
7	1007	8	4	1	0	0	0	1	0	1	0	1	0	283	0	
8	1008	16	15	2	0	0	0	0	3	1	0	0	0	1	0	
9	1009	4	3	1	0	1	0	0	1	0	0	0	0	253	0	
10	1010	17	17	2	1	2	0	2	1	5	0	0	1	1	0	
11	1011	7	4	1	1	0	0	0	3	1	0	0	0	257	0	
12	1012	10	9	2	0	1	0	5	2	1	0	0	0	334	0	
13	1013	14	12	2	0	0	0	4	5	3	0	0	0	700	0	
14	1014	13	12	1	0	1	0	4	3	4	0	0	0	785	0	
15	1015	15	14	2	1	0	0	7	2	3	0	0	0	762	0	
16	1016	18	18	2	0	0	0	0	7	7	0	0	0	1	1	
17	1017	7	3	2	1	0	0	3	0	1	0	0	0	99	1	
18	1018	13	10	2	1	0	0	7	3	1	0	0	0	328	1	
19	1019	17	16	2	0	4	0	4	8	2	0	0	0	980	1	
20	1020	4	2	2	0	0	0	2	0	1	0	0	0	76	1	
21	1021	4	2	2	0	0	0	2	0	1	0	0	0	76	1	
22	1022	17	16	2	0	1	0	7	5	3	0	0	0	827	1	
23	1023	11	11	2	5	1	0	4	6	2	0	0	0	658	1	
24	1024	16	16	2	2	0	0	3	5	5	0	0	0	1	1	

5.4 Consulta 4 - Los 13 máximos goleadores del torneo.

```
116 CREATE VIEW Goleadores AS
117 SELECT nombre, id_jugador, goles
118 FROM jugador, estadisticas_jugador
119 WHERE jugador.id_j=estadisticas_jugador.id_jugador;
120
121 SELECT nombre, goles, RANK() OVER(ORDER BY goles DESC) AS goles_rank
122 FROM Goleadores;demo
```

	nombre character varying (50)	goles smallint	goles_rank bigint
1	Luis Carlos Ruiz	10	1
2	Luciano Pons	9	2
3	Leonardo Castro	9	2
4	Miguel Borja	8	4
5	Duvier Riascos	8	4
6	Dayro Moreno	8	4
7	Wilson Morelo	8	4
8	Andres Renteria	6	8
9	Diego Herazo	5	9
10	Mender Garcia	5	9
11	Adrian Ramos	5	9
12	Michael Rangel	5	9
13	Luis Gonzalez	5	9

Se crea una vista llamada Goleadores con el producto cartesiano entre las tablas jugador y estadísticas jugador, donde la llave foránea es id jugador . En la vista se puede ver los atributos nombre, id jugador y goles. Luego se hace un Rank y order by de los mejores goleadores de todos los clubs, es decir los jugadores que hallan hecho más goles y se ordena de forma decreciente. Gracias a la consulta podemos ver que el mejor goleador de toda los clubs es Luis Carlos Ruiz con el cual cuanta con 10 goles.

5.5 Consulta 5 - Los 13 máximos asistidores del torneo.

```
124 CREATE VIEW Asistidores AS
125 SELECT nombre, id_jugador, asistencias
126 FROM jugador, estadisticas_jugador
127 WHERE jugador.id_j=estadisticas_jugador.id_jugador;
128
129
130 SELECT nombre, asistencias, RANK() OVER(ORDER BY asistencias DESC) AS asistencias_rank
131 FROM Asistidores;
132
```

	nombre character varying (50)	asistencias smallint	asistencias_rank bigint
1	Juan Sebastian Villota	7	1
2	Edwin Laszo	4	2
3	Jonny Mosquera	4	2
4	Luis Caicedo	4	2
5	Maikoll Herrera	4	2
6	Lewis Sinisterra	4	2
7	Jhonatan Perez	4	2
8	Juan Camilo Mosquera	4	2
9	Kevin Palacios	4	2
10	Andres Rivera	4	2
11	Fabio Delgado	4	2
12	Jaison Mina	4	2
13	Jose Garcia	4	2

Se creó una vista llamada Asistidores con el producto cartesiano entre las tablas jugador y estadísticas jugador , donde la llave foránea es id jugador. En la vista se selecciona los atributos nombre, id jugador y asistencias .Luego se hace un Rank y order by de las mejores asistencias de todos los clubs, es decir los jugadores que hallan hecho más asistencias y se ordena de forma decreciente. Gracias a la consulta podemos ver que la mejor asistencia es de Juan Sebastián Villota con 7 asistencias y los siguientes 12 jugadores que siguen ocupan el segundo lugar en el ranking con 4 asistencias.

5.6 Consulta 6 - Los colores característicos de cada club.

```

144 SELECT nombre, color_principal, color_secundario
145 FROM club INNER JOIN uniforme_local ON club.id_cl = uniforme_local.id_club;
146

```

	nombre character varying (30)	color_principal character varying (30)	color_secundario character varying (30)
1	Junior de Barranquilla	rojo	blanco
2	Atletico Nacional	verde	blanco
3	Deportes Tolima	vinotinto	amarillo
4	Millonarios FC	azul	blanco
5	CD America de Cali	rojo	rojo oscuro
6	Independiente Medellin	rojo	azul
7	Deportivo Cali	blanco	verde
8	Independiente Santa Fe	rojo	blanco
9	Once Caldas	blanco	verde
10	Alianza Petrolera	amarillo	negro
11	CD La Equidad Seguros SA	verde	verde oscuro
12	Atletico Bucaramanga	amarillo	verde
13	Aguilas Doradas	dorado	negro

Creamos un inner join entre las tablas club y uniforme local, donde la llave foránea es club id. Esta consulta nos permite ver el nombre del club y los colores principales y secundarios del uniforme local de cada club de la liga.

5.7 Consulta 7 - Presupuesto disponible para cada uno de los directores técnicos.

```

151
152 CREATE VIEW Directores_Tecnicos_Presupuesto AS
153 SELECT D.nombre, D.nacionalidad, C.presupuesto_en_miles
154 FROM director_tecnico AS D INNER JOIN club AS C ON D.id_club=C.id_cl;
155 SELECT nombre, presupuesto_en_miles, RANK() OVER ORDER BY presupuesto_en_miles DESC AS dt_presupuesto_rank
156 FROM Directores_Tecnicos_Presupuesto;
157

```

	nombre character varying (30)	presupuesto_en_miles smallint	dt_presupuesto_rank bigint
1	Juan Cruz Real	2413	1
2	Hernan Herrera	2238	2
3	Hernan Torres	2148	3
4	Alberto Gamero	1910	4
5	Alexandre Guimaraes	1505	5
6	Julio Comesania	1503	6
7	Alfredo Arias	1450	7
8	Diego Corredor	1055	8
9	Armando Osma	895	9
10	Jose Fernando Santa	818	10
11	Cesar Torres	773	11
12	Flabio Torres	753	12
13	Arturo Boyaca Gamboa	610	13

Creamos una vista llamada Directores técnicos presupuesto en donde se creaba un inner join entre las tablas director técnico y club con llave foránea en el atributo id club. La vista nos permite ver el nombre, nacionalidad y presupuesto en miles de cada director técnico, luego haciendo un Rank y order by pudimos obtener el ranking de los directores técnicos que tienen mejores presupuestos, cómo lo es Juan Cruz Real con un presupuesto de 2413, Hernan Herrera con 2238 y Hernán Torres con 2148.

5.8 Consulta 8 - Jugadores más influyentes.

	nombre character varying (50)	puntos_por_partido numeric	influyentes_rank bigint
1	Carlos Arboleda	158	1
2	Javier Lopez	99	2
3	Cristian Arrieta	81	3
4	Juan Sebastian Villota	76	4
5	Mauricio Gomez	67	5
6	Juan Pablo Nieto	61	6
7	Cesar Quintero	60	7
8	George Saunders	56	8
9	Deivy Balanta	50	9
10	Diego Gomez	48	10
11	Feiver Mercado	47	11
12	Jairo Ditta	46	12
13	Juan Camilo Mosquera	45	13
14	Jose Garcia	44	14
15	Rodrigo Paillalef	43	15
16	Duvan Viafara	42	16
17	Manuel Arias	41	17
18	Juan Nazarit	40	18
19	Jacobo Escobar	39	19
20	Camilo Charris	38	20
21	Jose Guzman	37	21
22	Adam Ozeri	36	22
23	Arley Mosquera	35	23
24	Luis Hurtado	34	24

```
CREATE VIEW Jugadores_influyentes AS
SELECT nombre, id_jugador, puntos_por_partido
FROM jugador, estadisticas_jugador
WHERE jugador.id_j=estadisticas_jugador.id_jugador;
SELECT nombre, puntos_por_partido, RANK() OVER(ORDER BY puntos_por_partido DESC) AS Influyentes_rank
FROM Jugadores_influyentes;
```

Creamos una vista llamada jugadores influyentes en donde se hace un producto cartesiano entre las tablas jugador y estadisticas jugador, donde la llave foranea es id jugador y se selecciona los atributos nombre, id y puntos por partido de cada uno de los jugadores. Adicionalmente se creo un rank y orden by para que se seleccionaran los jugadores mas influyentes de todos los clubs.

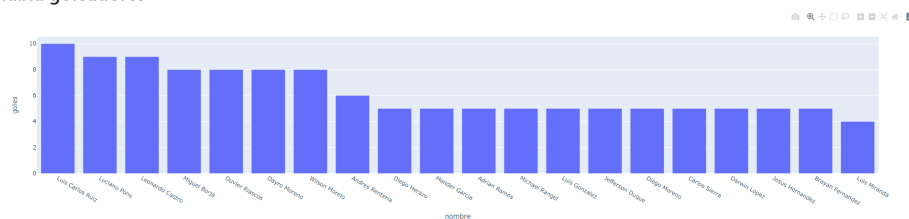
6 Visualización

Las siguientes visualizaciones fueron creadas con Dash y con las consultas de la sección anterior

6.1 Goleadores

Consulta 1, ranking de los mejores goleadores de todos los clubs de la liga colombiana de fútbol.

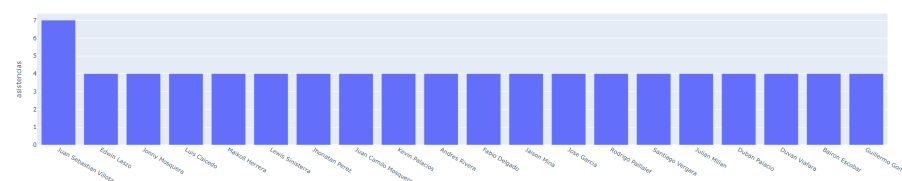
Tabla goleadores



6.2 Asistidores

Consulta 2, ranking de los mejores asistidores de todos los clubs de la liga colombiana de fútbol.

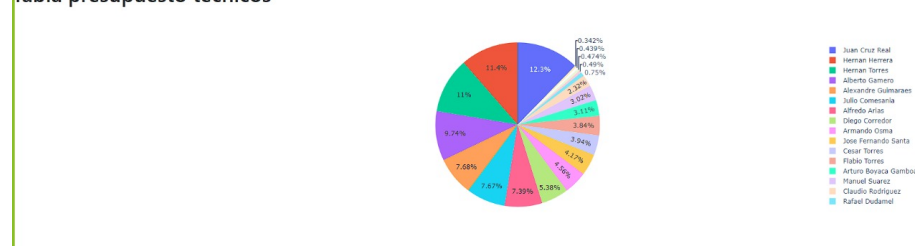
Tabla asistidores



6.3 Presupuesto director tecnico

Consulta 7, que muestra el presupuesto disponible de cada uno de los directores técnicos en gráfica de torta.

Tabla presupuesto técnicos



7 Análisis

Identificando las características de cada análisis desarrollado, mencionaremos las ventajas y desventajas de cada visualización.

La primera de las gráficas, la del pastel de presupuestos técnicos tiene la ventaja de contar con una representación visual muy clara, pudiendo entender de forma muy directa la relación de los valores, y su diferencia con respecto a otros. Pero con la desventaja de que, al tener varios datos de números muy pequeños, estos no tienen su representación, y se terminan perdiendo, sin saber

a qué jugador representa.

Por otra parte, la segunda y tercera en grafica tienen la ventaja de poseer todos los datos, pero con la desventaja de ocupar bastante espacio, además de poder verse feos en caso de gran variedad de datos.

Sirve en caso de tener pocos datos, ya que en caso de una gran cantidad se extendería horizontalmente donde resultaría imposible ver bien los datos. A diferencia de la grafica por pastel, donde da algo igual que tantos datos haya.

8 Conclusión

Usamos una excelente fuente de datos llamada TransferMarket la cual se actualiza bastante seguido, con mucha información del mundo del fútbol y de manera organizada, por lo que nos ayudo a hacer una base de datos grande en cuanto a tablas y datos.

El diseño en la base de datos presento cierta complejidad, al necesitar realizar varios cambios mientras se realizaba, pero una vez todo estuvo establecido en un papel, se pudo ordenar de una mejor forma cuando se paso a los modelos relacionales, facilitándonos el entendimiento y terminar de despejar varias dudas que había.

Presentamos bastantes inconvenientes con la inserción de datos , ya que la Liga de fútbol de Colombia cuenta con muchísimos datos, inicialmente creimos que podíamos ingresarlos a mano pero tuvimos que buscar otra forma más eficiente la cual fue Web scraping.

Sobre la conexión de base de datos, hubo ciertos problemas respecto al código en Python, pero nada serio, el tener apuntes de clase con el código de otro ejemplo facilito demasiado. Y ligado a eso, una vez con la conexión establecida, el desarrollo usando Dash tuvo cierta facilidad, teniendo ya básicamente todas las herramientas preparadas y solo necesitando realizar algunos cuantos cambios

Fue un proyecto que nos gasto mucho tiempo en la estructuración de la base de datos, en la búsqueda e inserción de datos, en la conexión de Python y demás objetivos que teníamos por la cantidad de tablas y datos. Fue un reto para nosotros que pudimos resolver de la mejor forma usando diferentes herramientas y aprendiendo muchas cosas más.