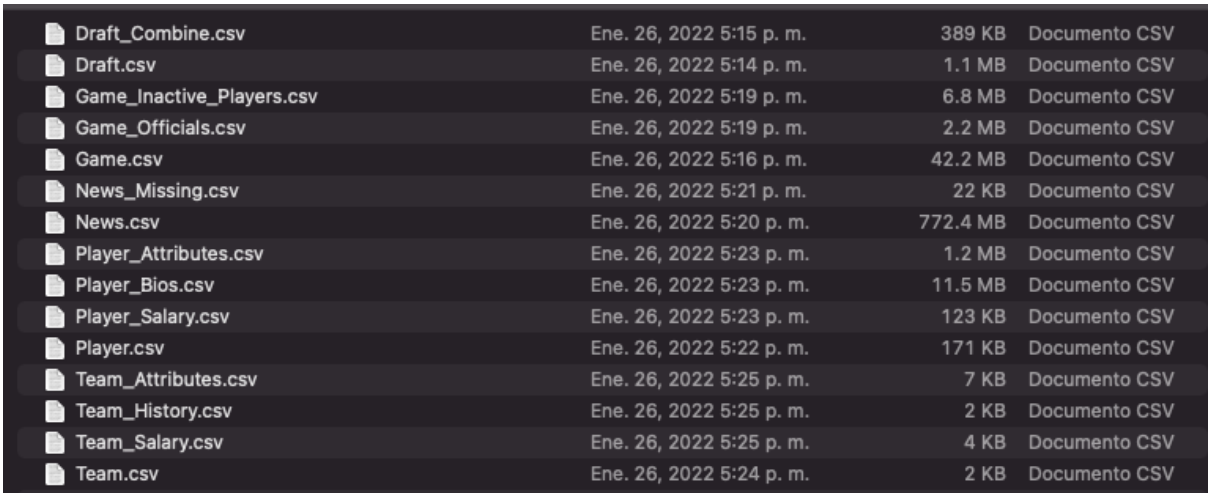
Sebastián Franco -21484
Mario Puente -21290

# Proyecto No. 1

## Desarrollo y consulta de Base de Datos

Etapa 1:

1. Descargar los archivos CSV y verificar qué información se incluye en cada uno.



| | | | |
|---|---|---|---|
| Draft_Combine.csv | Ene. 26, 2022 5:15 p. m. | 389 KB | Documento CSV |
| Draft.csv | Ene. 26, 2022 5:14 p. m. | 1.1 MB | Documento CSV |
| Game_Inactive_Players.csv | Ene. 26, 2022 5:19 p. m. | 6.8 MB | Documento CSV |
| Game_Officials.csv | Ene. 26, 2022 5:19 p. m. | 2.2 MB | Documento CSV |
| Game.csv | Ene. 26, 2022 5:16 p. m. | 42.2 MB | Documento CSV |
| News_Missing.csv | Ene. 26, 2022 5:21 p. m. | 22 KB | Documento CSV |
| News.csv | Ene. 26, 2022 5:20 p. m. | 772.4 MB | Documento CSV |
| Player_Attributes.csv | Ene. 26, 2022 5:23 p. m. | 1.2 MB | Documento CSV |
| Player_Bios.csv | Ene. 26, 2022 5:23 p. m. | 11.5 MB | Documento CSV |
| Player_Salary.csv | Ene. 26, 2022 5:23 p. m. | 123 KB | Documento CSV |
| Player.csv | Ene. 26, 2022 5:22 p. m. | 171 KB | Documento CSV |
| Team_Attributes.csv | Ene. 26, 2022 5:25 p. m. | 7 KB | Documento CSV |
| Team_History.csv | Ene. 26, 2022 5:25 p. m. | 2 KB | Documento CSV |
| Team_Salary.csv | Ene. 26, 2022 5:25 p. m. | 4 KB | Documento CSV |
| Team.csv | Ene. 26, 2022 5:24 p. m. | 2 KB | Documento CSV |

2. Crear la base de datos y las tablas en donde almacenará los datos.

De las tablas elegidas anteriormente, se seleccionan solamente las que ayudarán para la investigación. Entre estas se incluyen: Draft, Game_officials, Game, Player_Attributes, Player_salary, Player, Team_attributes, Team_salary, Team.

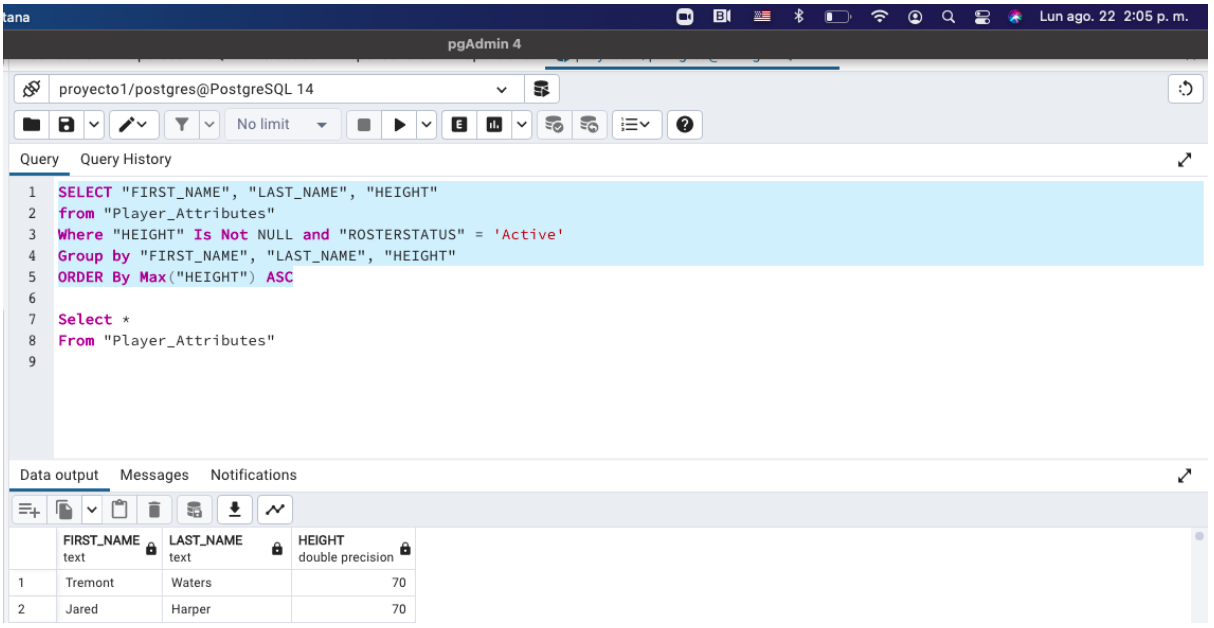
3. Desarrollar un script de Python (o lenguaje a su elección) que pueda leer la información de los archivos CSV, conectarse a su base de datos y alimentar la información en las tablas creadas.

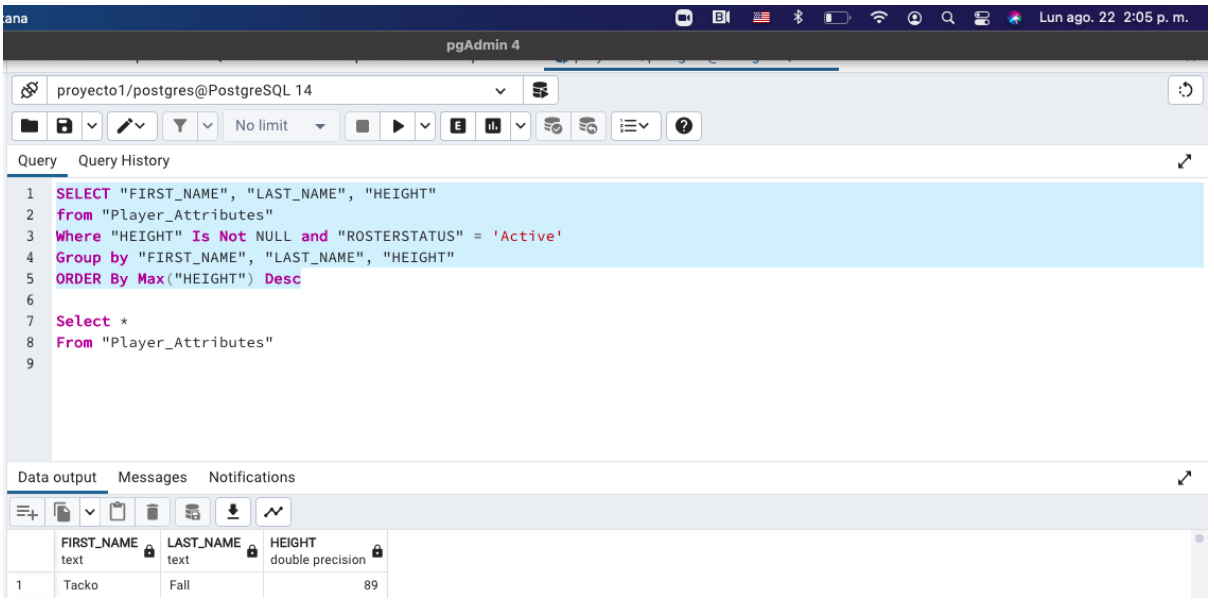
*El script de Python está en el zip, al igual que el diagrama Entidad Relación.
Link al repositorio de GitHub: https://github.com/sebastianf232/Proyecto-Bases-de-Datos

Etapa 2:

1. ¿Quién es el jugador activo más alto? ¿Y el más bajo?

Sebastián Franco -21484
Mario Puente -21290

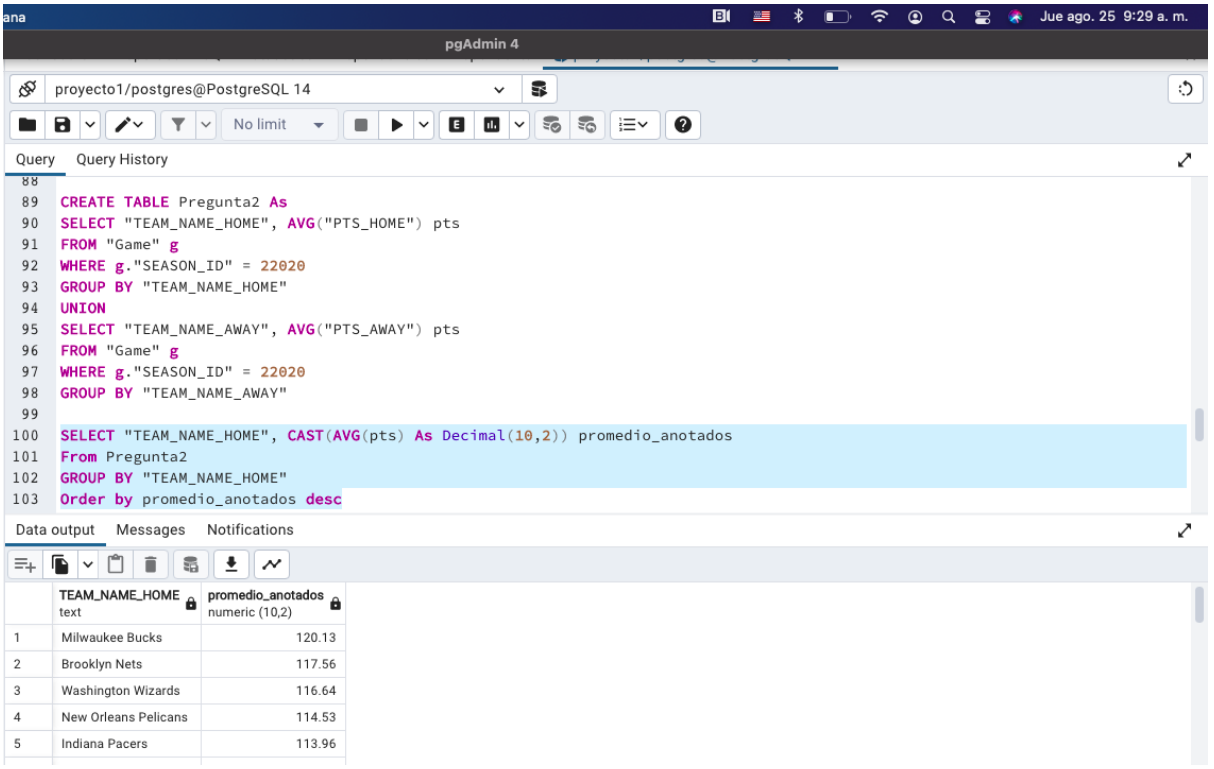


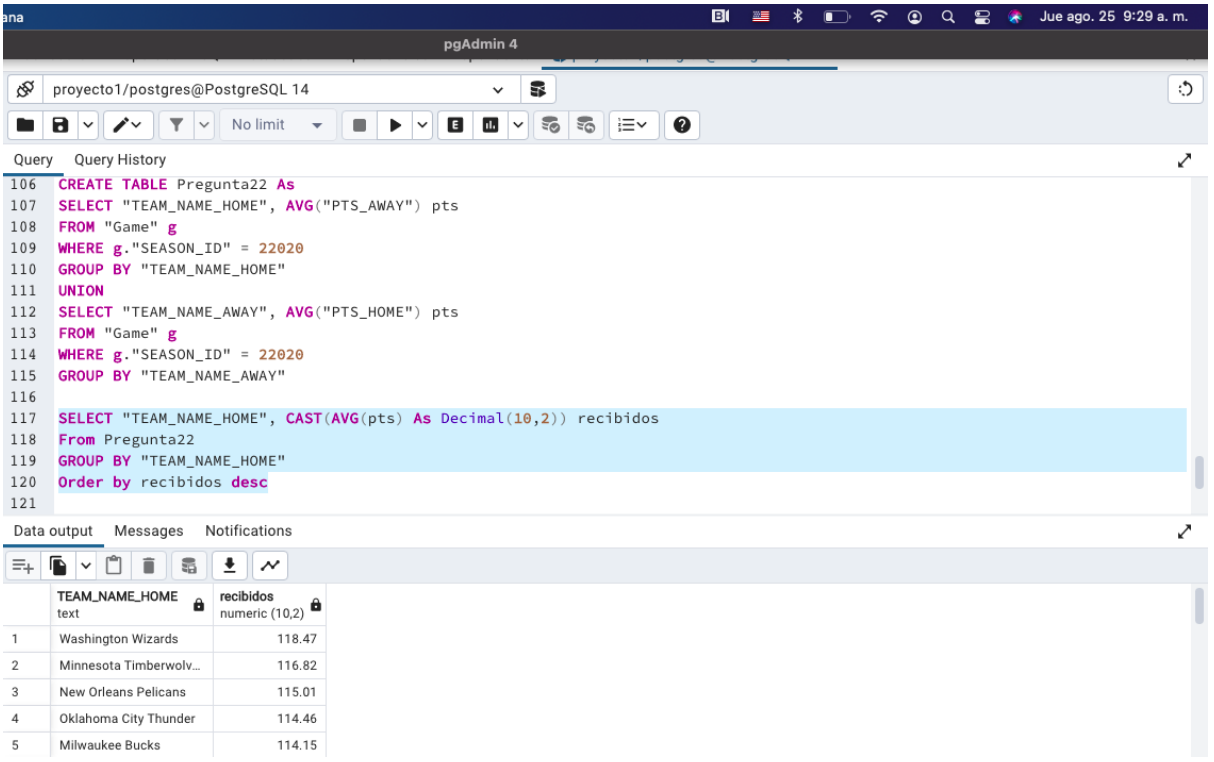

El Jugador más alto es Tacko Fall, y en el más bajo es un empate entre Tremont Waters y Jared Harper.

2. ¿Cuál fue el promedio de puntos anotados y recibidos por cada equipo en cada una de las temporadas relevantes?
   Queries Generales:
   Query que indica promedio de puntos anotados, se cambia la temporada para averiguar por cada temporada.

Sebastián Franco -21484
Mario Puente -21290

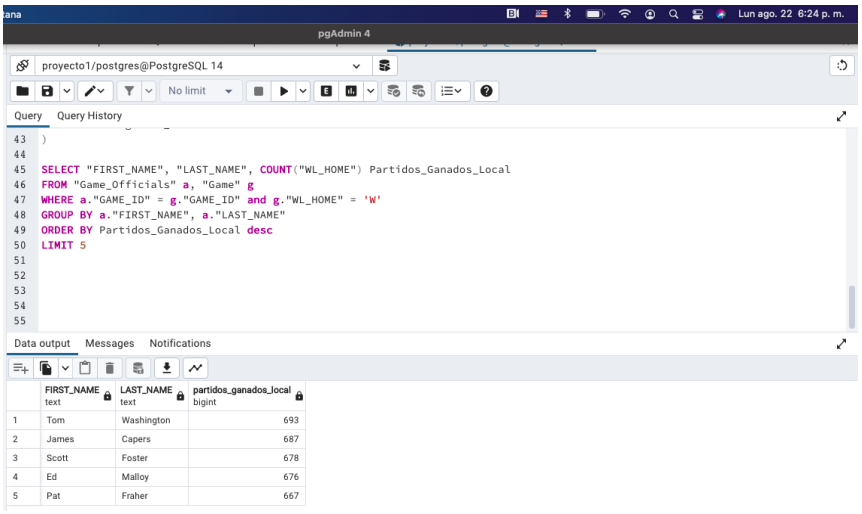



Query que indica promedio de puntos recibidos, se cambia la temporada.

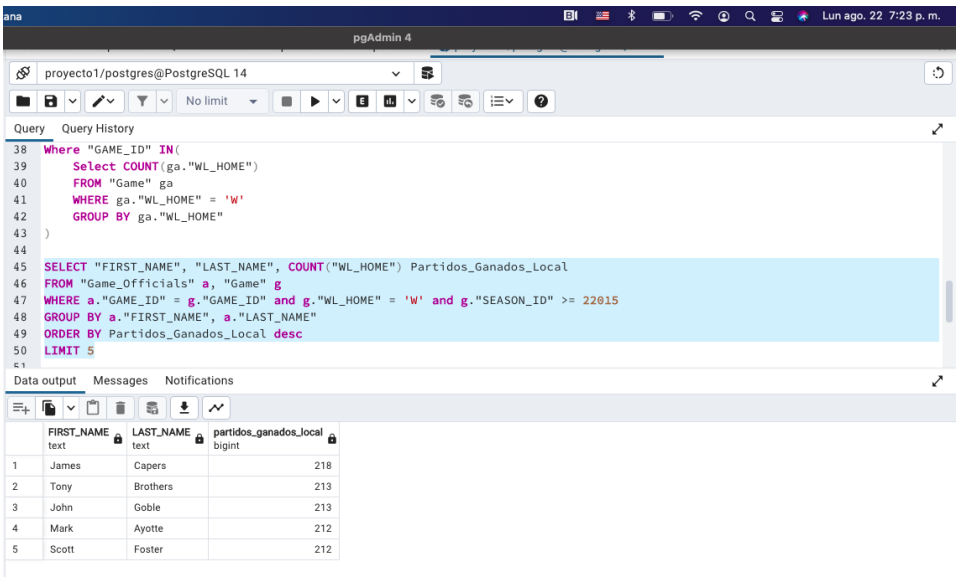


Sebastián Franco -21484
Mario Puente -21290

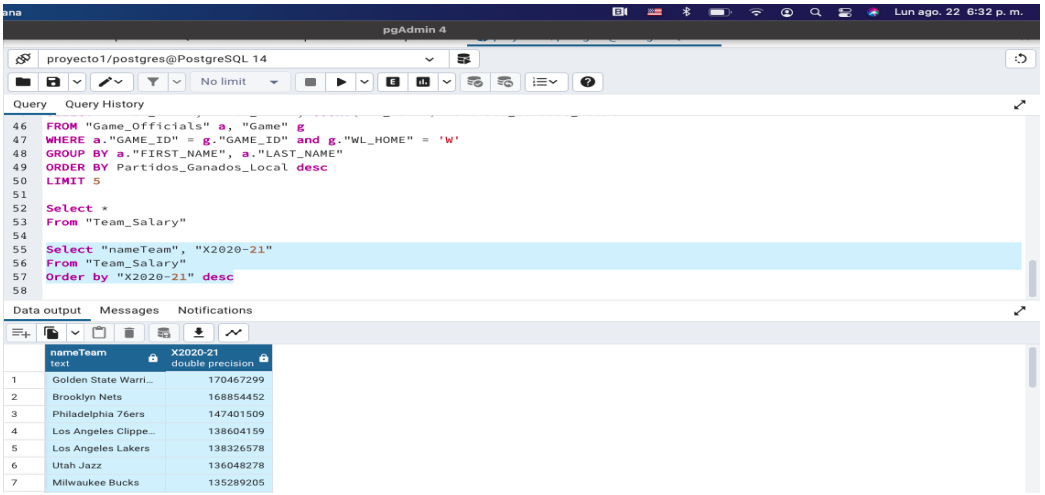
3. Muestre el top 5 de árbitros en cuyos juegos pitados el equipo visitante pierde.




Si se toma en cuenta solamente desde la temporada 2015/16




4. ¿Qué equipo maneja los salarios más altos actualmente?




5. ¿Cuál fue la temporada con más partidos en la historia de la NBA? ¿Cuál fue la temporada que más se ha prolongado en fechas?

Sebastián Franco -21484
Mario Puente -21290

```
SELECT COUNT("GAME_ID") as cuenta, "SEASON"
FROM "Game"
GROUP BY "SEASON"
ORDER BY cuenta desc
LIMIT 1
```

output Messages Notifications

| cuenta bigint | SEASON bigint |
|---|---|
| 1286 | 2013 |

```
SELECT MAX("GAME_DATE")-MIN("GAME_DATE") AS DateDifference, "SEASON"
FROM "Game"
GROUP BY "SEASON"
ORDER BY DateDifference desc
LIMIT 1
```

output Messages Notifications

| datedifference interval | SEASON bigint |
|---|---|
| 297 days | 2019 |

6. ¿Cuál es el equipo que más diferencia de puntos a favor obtuvo en promedio por partido para la temporada 2017? ¿Y para la temporada 2018?

Sebastián Franco -21484
Mario Puente -21290

```sql
SELECT AVG(promedio) as prom, nombre_equipo
FROM (
    SELECT SUM("PLUS_MINUS_AWAY") AS promedio, "TEAM_NAME_AWAY" AS nombre_equipo
    FROM "Game"
    WHERE "SEASON" = 2017
    GROUP BY "TEAM_NAME_AWAY"
    UNION
    SELECT SUM("PLUS_MINUS_HOME"), "TEAM_NAME_HOME"
    FROM "Game"
    WHERE "SEASON" = 2017
    GROUP BY "TEAM_NAME_HOME"
    ) AS subq
GROUP BY nombre_equipo
ORDER BY prom desc
LIMIT 1
```

output Messages Notifications

| prom numeric | nombre_equipo text |
|---|---|
| 347.5000000 | Houston Rockets |

```sql
SELECT AVG(promedio) as prom, nombre_equipo
FROM (
    SELECT SUM("PLUS_MINUS_AWAY") AS promedio, "TEAM_NAME_AWAY" AS nombre_equipo
    FROM "Game"
    WHERE "SEASON" = 2018
    GROUP BY "TEAM_NAME_AWAY"
    UNION
    SELECT SUM("PLUS_MINUS_HOME"), "TEAM_NAME_HOME"
    FROM "Game"
    WHERE "SEASON" = 2018
    GROUP BY "TEAM_NAME_HOME"
    ) AS subq
GROUP BY nombre_equipo
ORDER BY prom desc
LIMIT 1
```

output Messages Notifications

| prom numeric | nombre_equipo text |
|---|---|
| 363.5000000000000000 | Milwaukee Bucks |

7. ¿Quién es el jugador más valioso del draft del 2018 hoy en día?

Sebastián Franco -21484
Mario Puente -21290

```
SELECT "DISPLAY_FIRST_LAST" AS nombre, "DRAFT_YEAR", "PIE"
FROM "Player_Attributes"
WHERE "DRAFT_YEAR" = '2018' AND "PIE" IS NOT NULL
ORDER BY "PIE" desc
LIMIT 1
```

output Messages Notifications

| nombre text | DRAFT_YEAR text | PIE double precision |
|---|---|---|
| Luka Doncic | 2018 | 0.187 |

8. Calcule el top 5 de los estados que más salarios pagaron durante las temporadas 2020/2021 y 2021/2022.

```
SELECT "state", SUM("X2020-21") as salario
FROM "Team_Salary" JOIN "Team" ON "Team_Salary"."slugTeam" = "Team"."abbreviation"
GROUP BY "state"
ORDER BY salario DESC
LIMIT 5
```

output Messages Notifications

| state text | salario double precision |
|---|---|
| California | 553350641 |
| Texas | 388247885 |
| New York | 266468144 |
| Florida | 255954783 |
| Pennsylv... | 147401509 |

Sebastián Franco -21484
Mario Puente -21290

```
SELECT "state", SUM("X2021-22") as salario
FROM "Team_Salary" JOIN "Team" ON "Team_Salary"."slugTeam" = "Team"."abbreviation"
GROUP BY "state"
ORDER BY salario DESC
LIMIT 5
```

output Messages Notifications

| state text | salario double precision |
|---|---|
| California | 524437877 |
| Texas | 235941792 |
| New York | 213999731 |
| Florida | 179376763 |
| Wisconsin | 143531830 |

Etapa 3:

1. De los jugadores actuales, ¿Quién tiene el mejor PIE de la NBA?

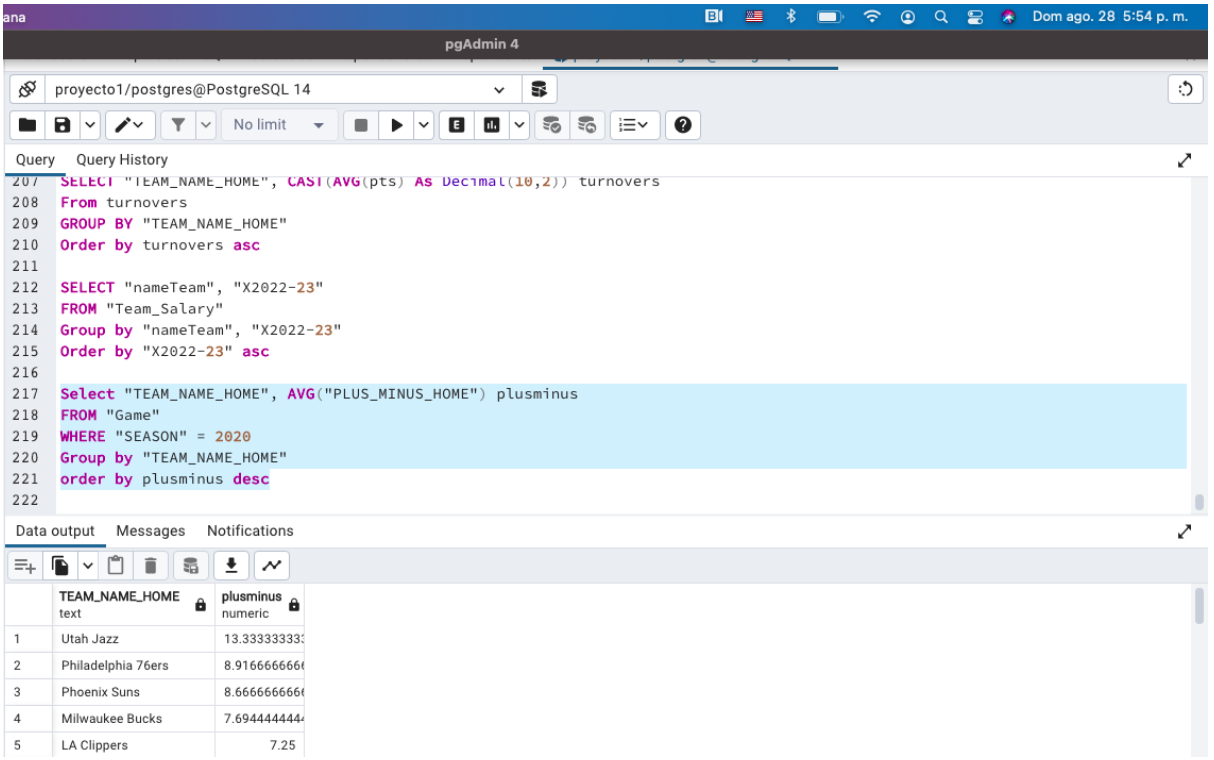
```
223 Select "FIRST_NAME", "LAST_NAME", "TEAM_NAME", "PIE", "PTS", "AST", "REB"
224 From "Player_Attributes"
225 Where "ROSTERSTATUS" = 'Active'
226 Group By "FIRST_NAME", "LAST_NAME", "TEAM_NAME", "PIE", "PTS", "AST", "REB"
227 Order By "PIE" desc
228 LIMIT 15
```

Data output Messages Notifications

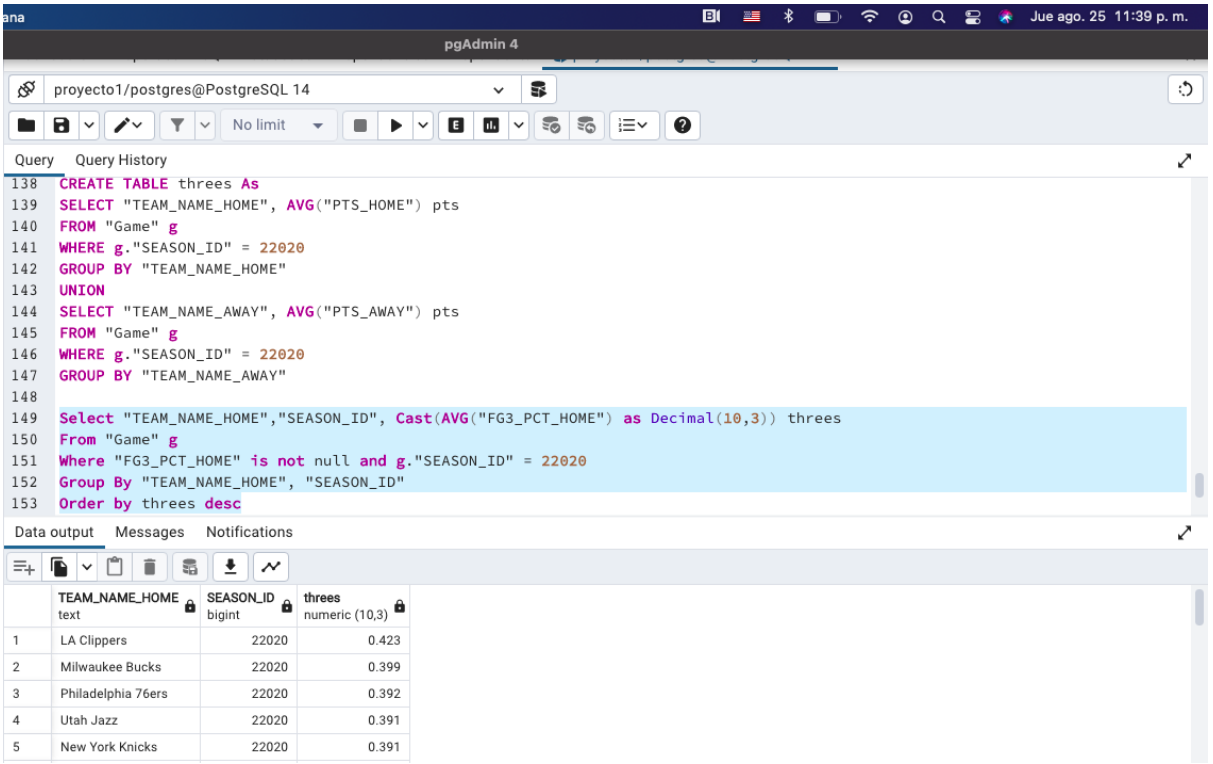
| | FIRST_NAME text | LAST_NAME text | TEAM_NAME text | PIE double precision | PTS double precision | AST double precision | REB double precision |
|---|---|---|---|---|---|---|---|
| 1 | Joel | Embiid | 76ers | 0.211 | 29.9 | 3.3 | 11.5 |
| 2 | Nikola | Jokic | Nuggets | 0.205 | 27.1 | 8.6 | 11.3 |
| 3 | Giannis | Antetokounm... | Bucks | 0.196 | 29 | 6.4 | 11.7 |
| 4 | LeBron | James | Lakers | 0.194 | 25.4 | 7.9 | 7.9 |
| 5 | Jimmy | Butler | Heat | 0.19 | 21.3 | 7.3 | 7.9 |

2. ¿Qué equipo tiene el mejor +- en casa de las temporadas de interés?

Sebastián Franco -21484
Mario Puente -21290

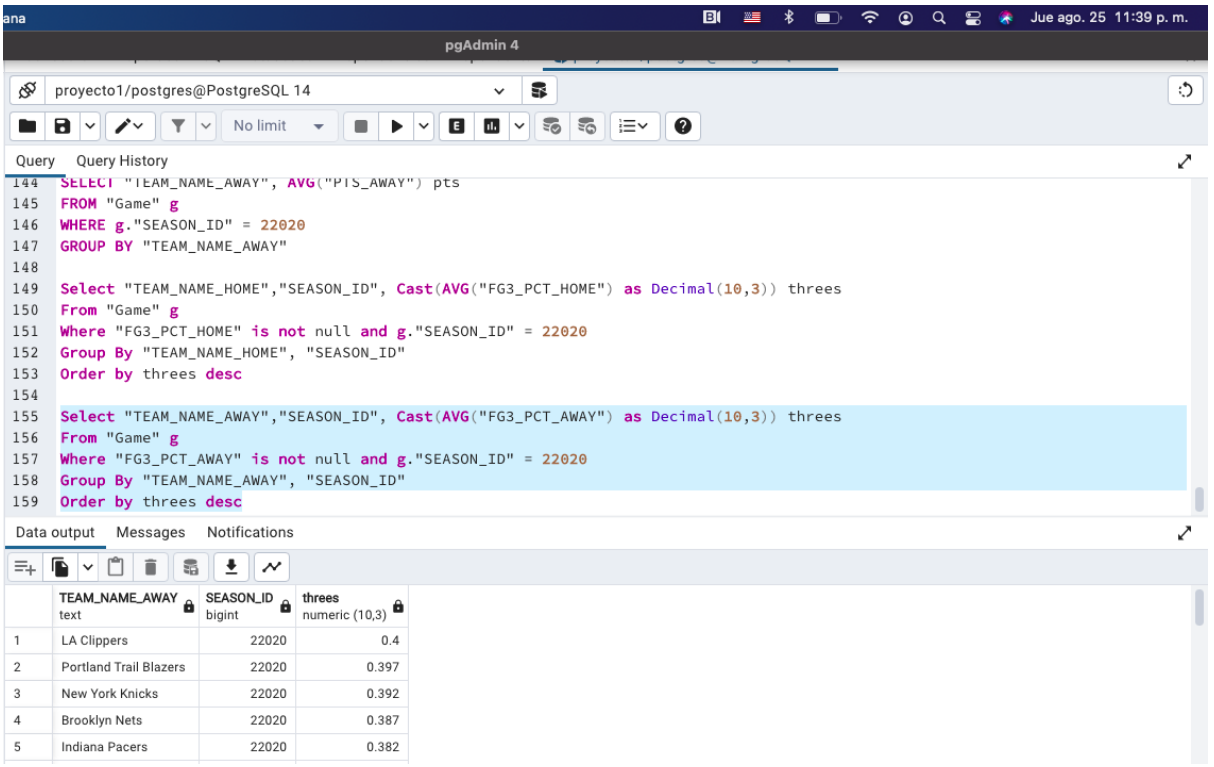



3. ¿Qué equipo tira mejor desde la línea de 3 puntos en la última temporada de interés? ?Desde la línea de tiros libres?
   Porcentaje de 3 en casa:



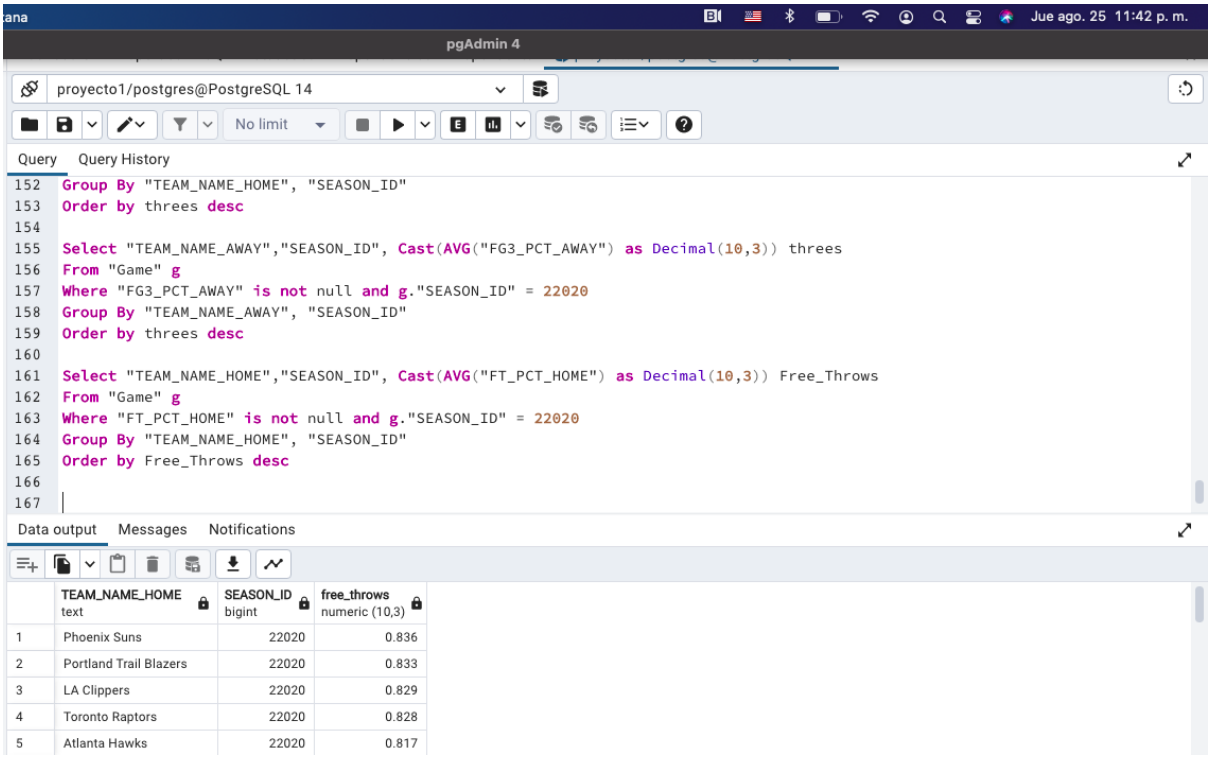
Sebastián Franco -21484
Mario Puente -21290

Porcentaje de 3 de visita:

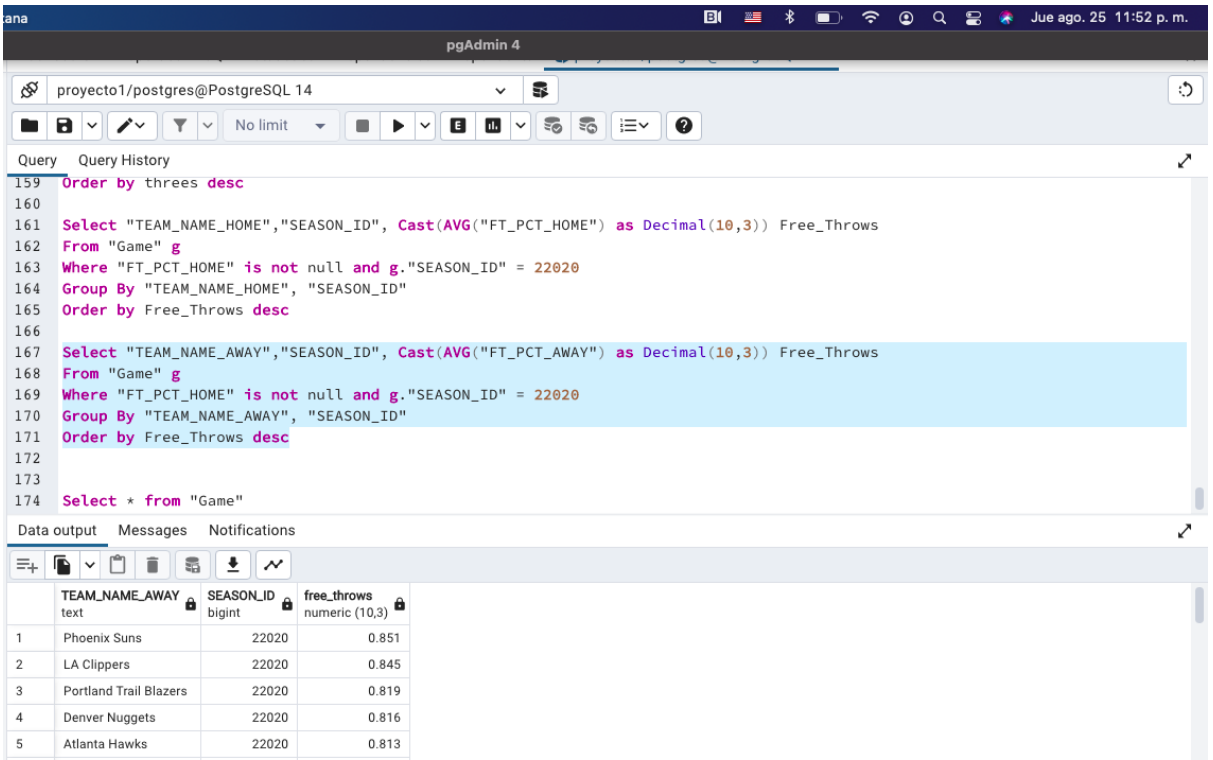



Porcentaje de FT en casa:

Sebastián Franco -21484
Mario Puente -21290

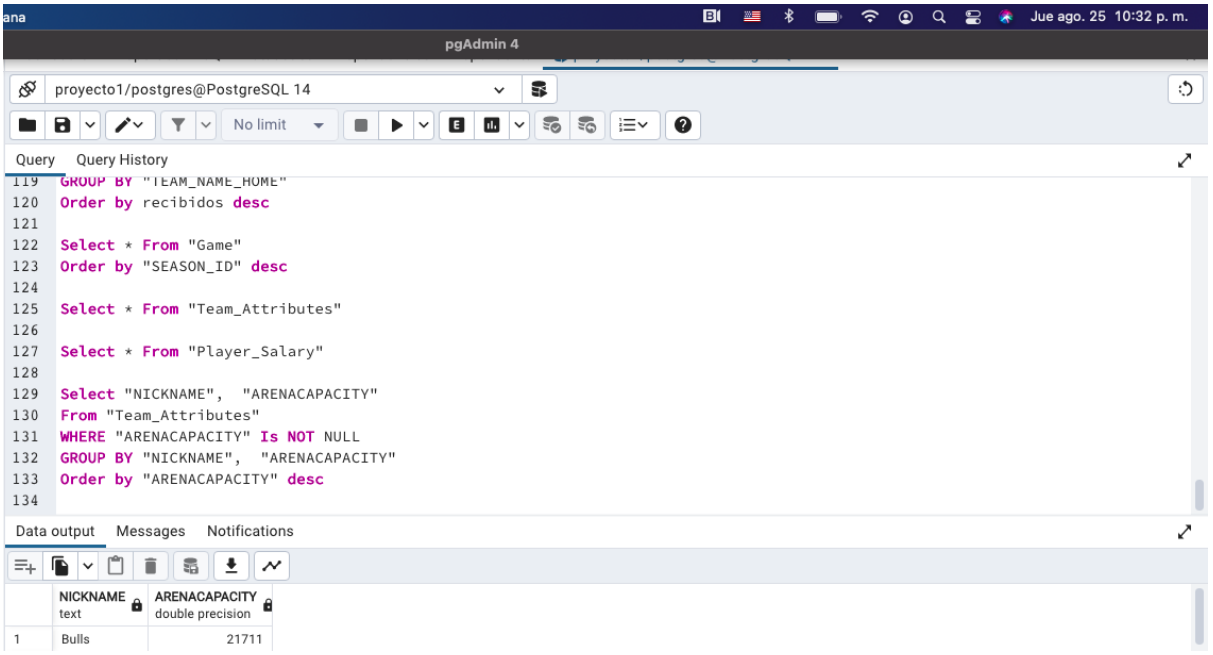



Porcentaje de Ft de visita:

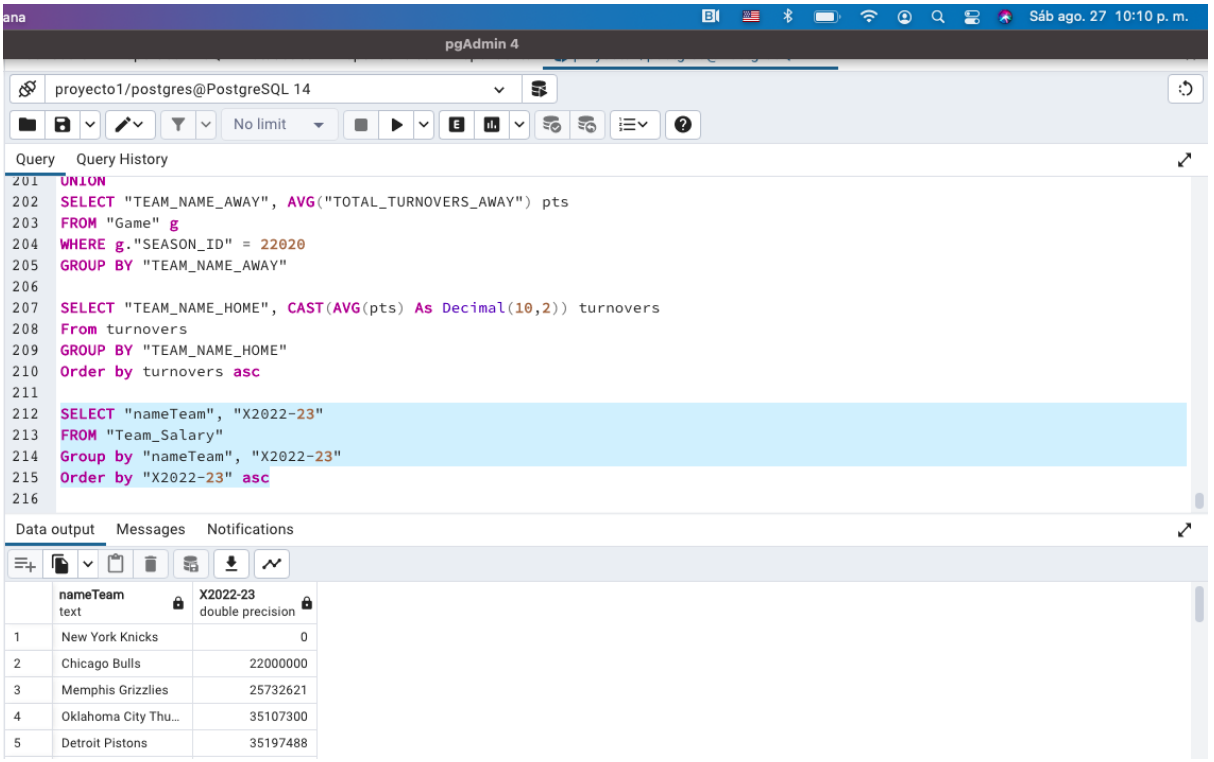



4. ¿Qué equipo tiene la mayor capacidad en su estadio?

Sebastián Franco -21484
Mario Puente -21290




5. ¿Qué equipos manejan los salarios más bajos para la siguiente temporada?




6. ¿Qué equipo pierde menos la pelota por partido?

Sebastián Franco -21484
Mario Puente -21290