

Data Analytics

**CODER HOUSE**

**PROYECTO FINAL:**  
**Análisis de la importancia**  
**del tiro de 3 puntos en el**  
**basketball moderno.**



**María Mercedes Vazzano**

20 de Enero de 2022

## **Resumen**

El presente contiene la documentación que acompaña el proyecto final realizado para el curso de Data Analytics de CoderHouse.

El objetivo general de este proyecto es integrar la totalidad de los conocimientos adquiridos durante el curso, mientras que los objetivos específicos consisten en desarrollar un tablero de control y documentar la iniciativa. La presente documentación que acompaña el proyecto está compuesta por cuatro capítulos.

En el Capítulo 1 se da una introducción al proyecto, mencionando cuál fue la motivación del mismo y lo que me llevó a definir la temática. Además se incluyen los tipos de análisis que se pueden llevar adelante.

En el Capítulo 2 se presenta la base de datos explicitando cuál fue la fuente de información utilizada, y se describe brevemente el contenido de las tablas que conforman el set de datos. Además se incluye el diagrama entidad-relación y las tablas que especifican los tipos de datos.

En el Capítulo 3 se describe cómo fue implementado el tablero, desde la extracción de la información y transformación de los datos hasta el diseño del mismo. Se incluyen además todas las fórmulas utilizadas para construir las tablas y medidas necesarias.

Finalmente en el Capítulo 4 se mencionan brevemente dos aspectos que no he llegado a resolver en este proyecto pero que considero que serían interesantes para explorar en el futuro.

# Índice general

<b>1. Introducción</b>	<b>2</b>
1.1. Motivación . . . . .	2
1.2. Definición de la temática . . . . .	2
1.3. Tipos de Análisis . . . . .	2
<b>2. Base de datos</b>	<b>4</b>
2.1.  Base de datos utilizada . . . . .	4
2.2. Diagrama Entidad-Relación . . . . .	5
2.3. Tablas . . . . .	6
<b>3. Implementación del tablero en Power BI</b>	<b>8</b>
3.1. Extracción de la información . . . . .	8
3.2. Transformación de datos . . . . .	8
3.3. Diseño del dashboard . . . . .	9
3.4. Tablas, columnas y medidas calculadas . . . . .	13
3.5. Medidas claculadas avanzadas . . . . .	16
<b>4. Futuras líneas</b>	<b>18</b>
<b>A. Glosario</b>	<b>19</b>
<b>B. Pestañas del tablero</b>	<b>21</b>

# Capítulo 1

## Introducción

### 1.1. Motivación

La estadística juega un rol importante en el análisis deportivo. La aplicación de modelos estadísticos permite obtener información objetiva sobre la caracterización de los atletas, y el rol de cada deportista y del equipo frente a sus adversarios. Además, colabora a mejorar la exactitud en el pronóstico del rendimiento deportivo y aumenta la eficiencia en la detección de talentos.

No cabe duda que entre todos los deportes el *basketball* es el rey de las estadísticas. A lo largo de los últimos años la forma de entender el juego en la NBA ha cambiado enormemente en comparación a los 10 años anteriores, algo que se debe, en parte, al auge del uso, comprensión y expansión de la estadística avanzada. Las estadísticas convencionales (puntos, rebotes, asistencias) se fueron quedando atrás conforme el análisis del deporte evolucionaba y precisaba de herramientas más complejas para entender todo lo que sucedía en la cancha. Esto dio lugar al desarrollo de parámetros que aportan más información, como por ejemplo el plus/minus, que evalúa el impacto que tiene un jugador sobre el resultado del partido, más allá de su performance según las estadísticas clásicas.

Por lo dicho anteriormente considero que los datos estadísticos sobre este deporte pueden proporcionar un interesante campo de exploración.

### 1.2. Definición de la temática

Dentro de la inmensa complejidad que exhiben los análisis estadísticos del *basketball* he optado por concentrarme en el análisis de un factor en particular. En este trabajo me propongo analizar cómo ha cambiado la forma de juego en los partidos de la NBA a lo largo de los últimos años. Para esto basaré el análisis en la cantidad de puntos discriminando entre dobles, triples y tiros libres, tanto intentados como convertidos.

Este aspecto en particular me pareció de interés, ya que es sabido que la cantidad de puntos obtenidos por tiros de tres puntos ha ido en aumento en los últimos años. A modo de ejemplo, recientemente Stephen Curry consiguió convertirse en el mayor anotador de triples de la historia de la NBA. El día 14 de diciembre de 2021 el Madison Square Garden fue escenario de la victoria de los Golden State Warriors frente a New York Knicks por 105-96. Curry sólo necesitaba dos más para superar los 2973 de Ray Allen y acabó el encuentro con cinco de 14 intentos, lo que lo colocó con 2977 tiros de tres en su carrera.

Con el fin de llevar adelante este análisis me propuse diseñar un tablero que permitiera evaluar la importancia de los tiros de tres puntos en el *basketball* analizando estadísticas con tres diferentes enfoques: por equipos, jugadores y enfrentamientos.

### 1.3. Tipos de Análisis

A continuación se describen los tipos de análisis aplicados al proyecto, tales como fueron presentados en el desafío 'Tipos de Análisis' (fecha de entrega 11/10/21).

### **Análisis descriptivo ¿Qué pasa?**

Previo al comienzo de la 75° temporada de la NBA (2021-2022) se ha observado que los últimos años los equipos que lideran la competencia se destacan por tener mejor promedio de lanzadores de triples.

### **Análisis diagnóstico ¿Por qué pasa?**

A través de la creación de un dashboard interactivo se pretende investigar y argumentar por qué los equipos y jugadores obtienen mejor promedio de triples, y utilizar esto para analizar cómo fue cambiando la forma de juego en la NBA a lo largo de los años.

### **Análisis predictivo ¿Qué pasará?**

En base a los reportes obtenidos se busca identificar cuáles serían los encuentros con mayor cantidad de puntos, determinar cuales son los jugadores y equipos que se destacan, y predecir posibles resultados de encuentros.

### **Análisis prescriptivo ¿Qué se puede hacer?**

El análisis realizado puede resultar de util en varias áreas. Por ejemplo para ser usado en casas de apuestas o como materia prima para los buscadores de talentos. También podría ayudar en el análisis de los partidos a periodistas deportivos. Teniendo esto en cuenta los niveles de aplicación de data analytics serían operativo y/o táctico.

# Capítulo 2

## Base de datos

### 2.1. Base de datos utilizada

He seleccionado para este trabajo la base de datos **NBA games data**<sup>1</sup> que se encuentra en [kaggle](#). El dataset fue creado a partir de datos disponibles en la página de la NBA<sup>2</sup> y está compuesto por 5 tablas con datos estadísticos tomados entre los años 2003 y 2020. A continuación se describe brevemente el contenido de cuatro de estas tablas, que son las que se utilizaron en este trabajo.

- ⊗ `games.csv` : contiene datos de todos los partidos desde la temporada 2003 hasta 2020, los equipos que disputaron cada partido y algunos detalles como el número de puntos, asistencias y rebotes.
- ⊗ `games_details.csv` : contiene datos referidos al desempeño de cada jugador en cada partido.
- ⊗ `players.csv` : muestra detalles de los distintos jugadores, tales como en qué equipo jugó en cada temporada.
- ⊗ `teams.csv` : contiene datos referidos a cada equipo, tales como nickname, ciudad, estadio, etc.

Las tablas `gamas.csv` y `games_detail.csv` son las que proporcionan información cuantitativa relevante para el proyecto, mientras que las tablas `players.csv` y `teams.csv` son complementarias. No se uso para el análisis la tabla `ranking.csv` que se encuentra en el dataset original, ya que no proporciona información relevante para el objetivo propuesto en este proyecto.

Las tablas `players.csv` y `games_details.csv` en el set de datos original no poseen una columna que pueda usarse como *primary key*, por lo que se ha agregado una columna ID en cada tabla. Las tablas modificadas utilizadas en este trabajo pueden encontrarse en [google drive](#)<sup>3</sup>.

Dado que los nombres de las tablas y las columnas de estas están en inglés en el dataset original, en el [Apéndice A](#) se incluye un glosario con el nombre o descripción en español de cada campo.

---

<sup>1</sup><https://www.kaggle.com/nathanlauga/nba-games>

<sup>2</sup><https://www.nba.com/stats/>

<sup>3</sup>[https://drive.google.com/drive/folders/1dsiAKnmL3pA\\_hb9RGZVIERYPU-Q6IbpW?usp=sharing](https://drive.google.com/drive/folders/1dsiAKnmL3pA_hb9RGZVIERYPU-Q6IbpW?usp=sharing)

## 2.2. Diagrama Entidad-Relación

En esta sección se incluye el diagrama entidad relación para las cuatro tablas utilizadas en el trabajo, y que fue parte del desafío ‘Diagrama Entidad-Relación’ (fecha de entrega 18/10/21). Este diagrama está elaborado en base a las tablas originales y no incluye las tablas punte, calendario y otras tablas auxiliares que se describirán más adelante.

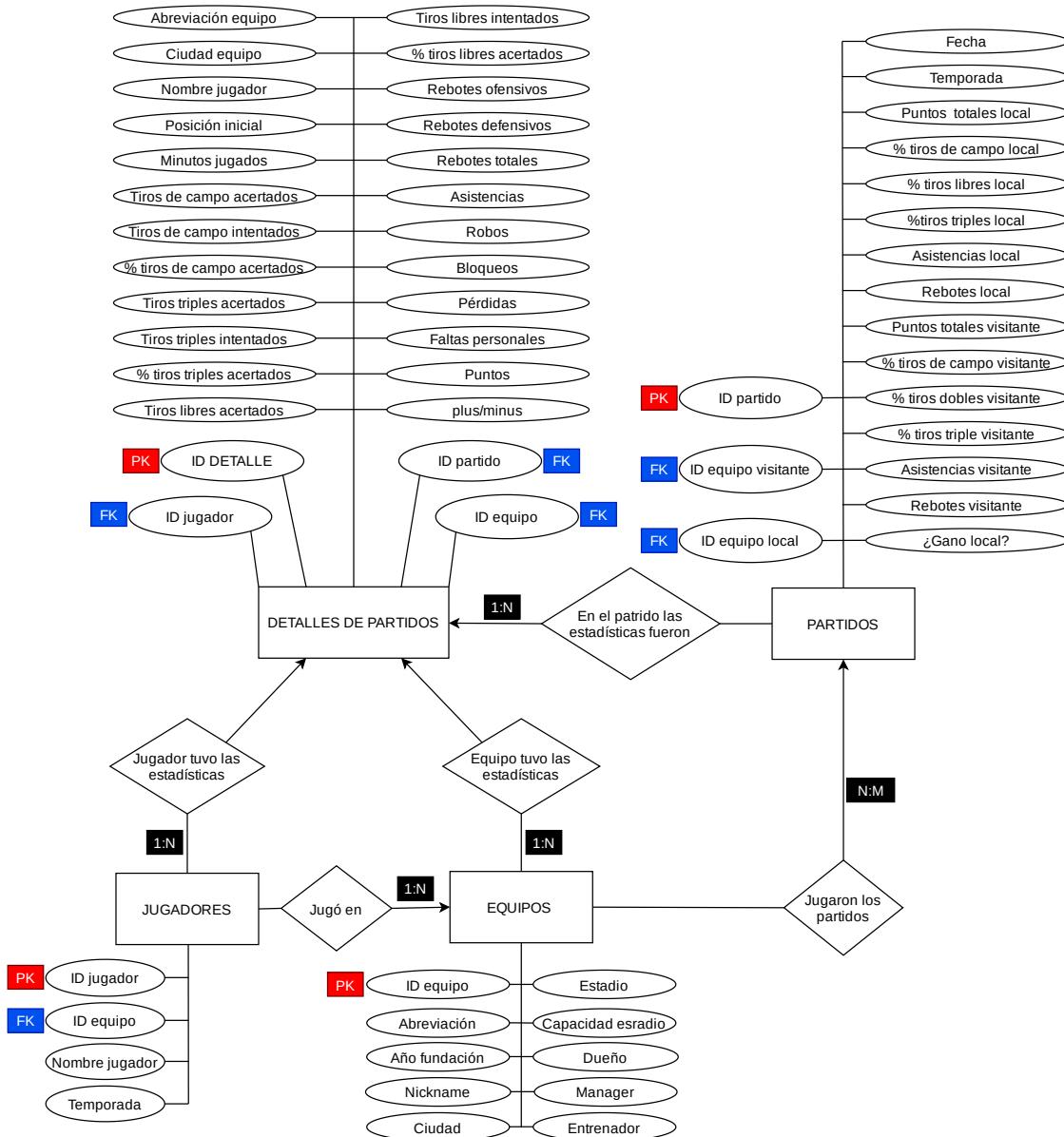


Figura 2.1: Diagrama entidad-relación

## 2.3. Tablas

En esta sección se incluye la información referida a cada una de las tablas del set de datos, y que formaron parte del desafío 'Definición de Tablas del Proyecto Final' (entregado el 20/10/21). Se detalla para cada una de ellas la *primary key* (PK), las *foreign keys* (FK) y los tipos de campo para cada columna.

Campo	Tipo de clave	Tipo de campo
ID partido	PK	int
Fecha		datetime
ID equipo local	FK	int
ID equipo visitante	FK	int
Temporada		int
Puntos totales local		int
% tiros de campo local		decimal
% tiros libres local		decimal
% tiros triple local		decimal
Asistencias local		int
Rebotes local		int
Puntos totales visitante		int
% tiros de campo visitante		decimal
% tiros dobles visitante		decimal
% tiros triple visitante		decimal
Asistencias visitante		int
Rebotes visitante		int
¿Ganó el local?		bool

Tabla 2.1: PARTIDOS

Campo	Tipo de clave	Tipo de campo
ID detalle	<b>PK</b>	int
ID partido	<b>FK</b>	int
ID equipo	<b>FK</b>	int
Abreviación equipo		nvchar(50)
Ciudad equipo		nvchar(50)
ID jugador	<b>FK</b>	int
Nombre jugador		nvchar(50)
Posición inicial		nvchar(50)
Minutos jugados		timestamp
Tipos de campo acertados		int
Tiros de campo intentados		int
% tiros de campo intentados		decimal
Tiros triples acertados		int
Tiros triples intentados		int
% triples acertados		decimal
Tiros libres acertados		int
Tiros libres intentados		int
% tiros libres acertados		decimal
Rebotes ofensivos		int
Rebotes defensivos		int
Rebotes totales		int
Asistencias		int
Robos		int
Bloqueos		int
Perdidas		int
Faltas personales		int
Puntos		int
Plus/minus		int

Tabla 2.2: DETALLE PARTIDOS

Campo	Tipo de clave	Tipo de campo
ID equipo	<b>PK</b>	int
Año fundación		int
Abreviación		nvchar(50)
Nickname		nvchar(50)
Ciudad		nvchar(50)
Estadio		nvchar(50)
Capacidad estadio		int
Dueño		nvchar(50)
Manager		nvchar(50)
Entrenador		nvchar(50)

Tabla 2.3: EQUIPOS

Campo	Tipo de clave	Tipo de campo
ID jugador	<b>PK</b>	int
Nombre		nvchar(20)
ID equipo	<b>FK</b>	int
Temporada		int

Tabla 2.4: JUGADORES

## Capítulo 3

# Implementación del tablero en Power BI

### 3.1. Extracción de la información

Las cinco tablas utilizadas fueron incluidas como pestañas en un mismo archivo `desafio_NBA.xlsx`<sup>1</sup>. El dataset utilizado fue cargado en Power BI utilizando los botones Inicio>Obtener Datos>Libro de Excel, y el modelo relacional fue modificado para que las conexiones entre las tablas sean las definidas en el proyecto (ver diagrama entidad relación en al Figura 2.1).

### 3.2. Transformación de datos

Algunas tablas debieron ser modificadas para poder llevar adelante el proyecto. A continuación se listan las tablas modificadas y agregadas.

- A la tabla `teams` se le agregó una columna con la dirección URL en donde se encuentran las imágenes de los logos de cada equipo.
- En la tabla `teams` se incluyeron las columnas latitud y longitud, con la localización geográfica del estadio.
- En todas las tablas se generaron transformaciones sobre los campos ID seteando como ‘No resumir’ ya que no serán utilizadas para cálculos.
- Se agregaron dos tablas puente para evitar que en el modelo relacional aparecieran relaciones de muchos a muchos:
  - `players_puente`: tabla puente entre `players` y `games_details`.
  - `games_puente`: tabla puente entre `games` y `games_details`.
- Se agregaron las tablas `teams_home` y `teams_away` (para equipo local y visitante) que contiene el ID y nickname de cada equipo, y que hacen de puente entre las tablas `teams` y `game_details`.
- Se generó una `tabla_calendario` a partir de la tabla `games` que contiene las fechas de los partidos y temporadas. El campo `Season` (temporada) es necesario, ya que estas no son por año calendario sino de mitad a mitad de año y será necesario para las segmentaciones.

El modelo relacional resultante luego de todos los cambios es el que se muestra en la Figura 3.1

---

<sup>1</sup><https://docs.google.com/spreadsheets/d/1wKicDjVdPwqrJxvsooeBuGLIEBtbXVMn/edit?usp=sharing&ouid=102465342678283560019&rtpof=true&sd=true>

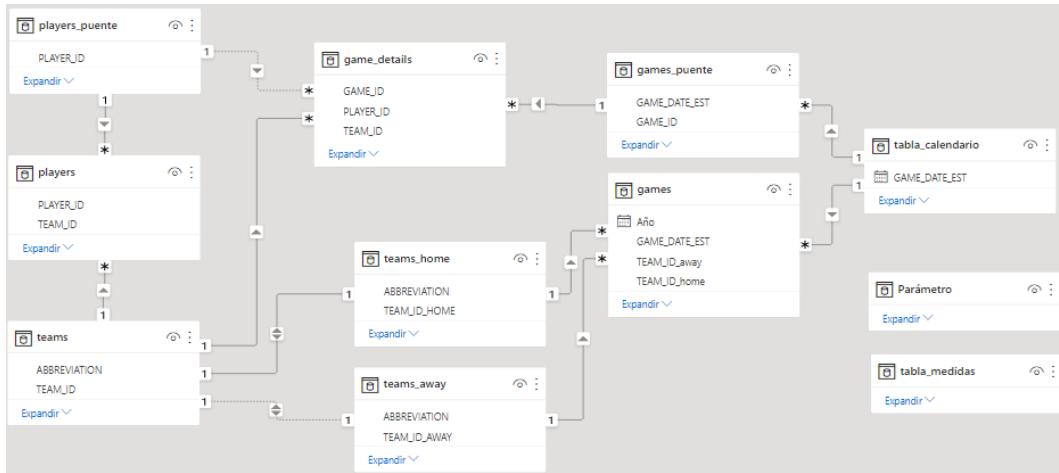


Figura 3.1: Modelo relacional

### 3.3. Diseño del dashboard

En esta sección describiré cómo fue diseñado el tablero. Para eso incluyo capturas de pantalla del mismo, en algunos casos de una pestaña completa y en otro de alguna parte de las mismas. Capturas de pantalla de las 4 pestañas completas se adjuntan en el [Apéndice B](#).

#### Portada y pestañas

Se incluyó en el dashboard una portada con el título del trabajo y nombre de la autora. Además se incluyeron tres pestañas: **Equipos**, **Jugadores** y **Enfrentamientos**. La imagen de fondo de la portada muestra imágenes de jugadores de la NBA de diferentes equipos y temporadas. Esta imagen se repite como fondo en el resto de las pestañas con transparencia para que no interfiera con la información allí mostrada.

#### Paleta de colores y tipografía

Se seleccionó una gama de colores con **azul** (#17408B), **rojo** (#C9082A), colores del logo de la NBA, y se incluyó el **gris** (#AAA3A3) para complementar. La tipografía elegida fue Segoe UI.

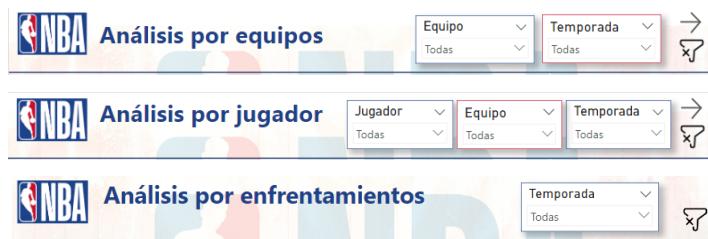
#### Navegación entre pestañas y filtros

La portada presenta cuatro botones de navegación: en la esquina superior derecha una flecha que permite pasar a la pestaña siguiente, y en la esquina inferior derecha tres íconos que redireccionan a cada una de las tres pestañas (Equipos, Jugadores y Enfrentamientos).



Las tres pestañas incluyen una barra en la parte superior del dashboard en las que se muestra:

- el logo de la NBA que funciona como botón para volver al inicio
- el título de la pestaña
- los filtros
- la flecha que direcciona a la pestaña siguiente (excepto en la última pestaña)
- el botón para limpiar los filtros



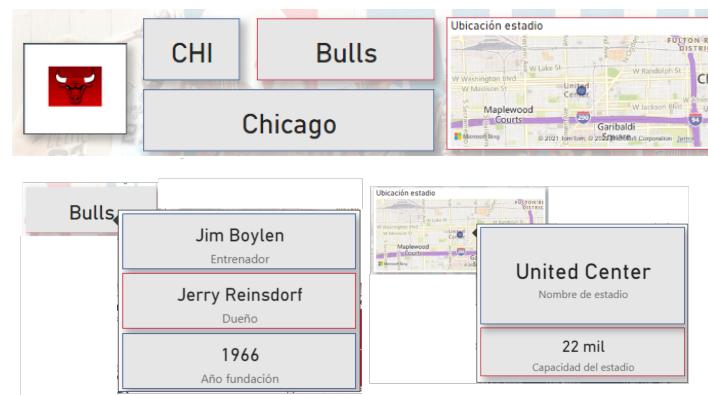
### Contenido de las pestañas

A continuación se describe el contenido de cada una de las pestañas del dashboard. Los detalles respecto a cómo se realizaron los cálculos se especifican en las siguientes secciones.

- **Análisis por equipos:** Se incluyen dos filtros (por equipo y por temporada).

Al seleccionar un equipo se muestra el logo (implementado a través de una tabla) y tres cards con la abreviación del equipo, el nickname y la ciudad. Para poder visualizar el logo del equipo fue necesario editar la columna que contiene la dirección URL en donde están localizadas las imágenes seleccionando Categoría de datos>Dirección URL de la imagen. Además se despliega en un mapa la localización geográfica del estadio. Para esto se utilizaron las columnas Latitud y Longitud que fueron agregadas manualmente al set de datos antes de ser cargado en Power BI.

La card con el nickname tiene asociado un tooltip que muestra información adicional (entrenador, propietario y año de fundación del equipo) y el mapa tiene asociado un tooltip que muestra información sobre el estadio (nombre y capacidad). Esto permite incluir más información sobre cada equipo sin que la pestaña se vea recargada. Toda esta información del equipo se despliega en la parte superior de la pestaña y es lo primero que debe ser leído por el usuario.

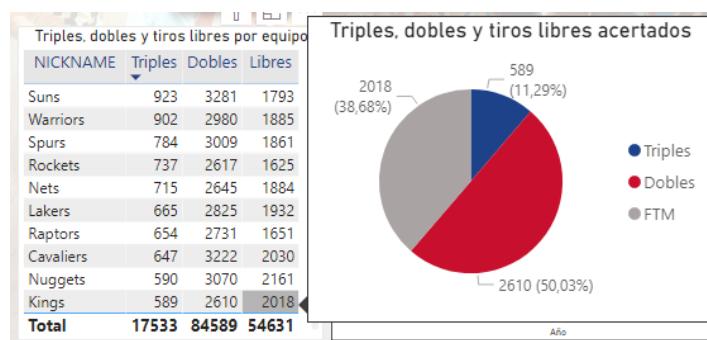


A la derecha de la información sobre el equipo y se incluye un KPI que muestra la primera información cuantitativa de la pestaña. Este compara los triples convertidos en las temporadas 2003 y 2000 (primera y última del dataset). A pesar de que este dato no nos muestra detalles sobre cómo evolucionó la conversión de triples a lo largo de los años, nos permite hacer una

primera comparación: en general, ¿los equipos hoy en día hace más o menos triples que hace 17 años? Se puede ver que aunque para algunos equipos la cantidad de triples convertidos aumentó más que para otros, en general la tendencia es que este valor va en aumento, ya que los valores mostrados en son siempre positivos.



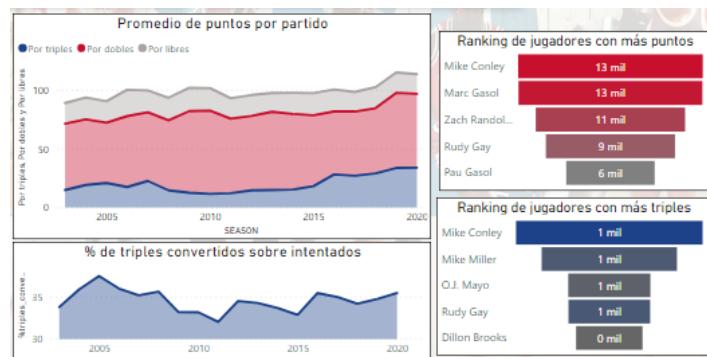
A continuación (siguiendo el patrón Z) el usuario puede ver una tabla (no afectada por el filtro de equipo pero si por temporada) que muestra los triples, dobles y de tiros libres convertidos por cada equipo. Esta tiene asociado un tooltip que muestra los mismos datos de la tabla pero representados en un gráfico de torta que resulta mucho más visual y fácil de interpretar a simple vista.



A la derecha de la tabla se incluyen dos gráficos de áreas. Uno de ellos (de áreas apiladas) muestra el promedio de puntos por partidos logrado por un equipo para cada temporada. Las áreas apiladas representan los puntos logrados por triples (azul) por dobles (rojo) y por tiros libres (gris). Este gráfico permite ver como evoluciona la cantidad promedio de puntos que hace un equipo a lo largo de los años, distinguiendo de que forma estos fueron obtenidos. Así es posible analizar si el aumento de puntos se debe o no a que se han convertido más tiros de tres puntos.

Debajo se muestra otro gráfico de áreas en el que se representa el porcentaje de triples convertidos sobre los intentados. Esto permite evaluar si el aumento de los puntos obtenidos por tiros de tres puntos se deben a que mejoró la efectividad de conversión o a que se intentaron más tiros.

Finalmente, y a modo de cierre de la pestaña, se presentan dos gráficos que muestran el top 5 de jugadores con más puntos y con más triples convertidos. Este permite evaluar si los jugadores que más puntos consiguen son también los que aciertan más tiros de tres puntos.

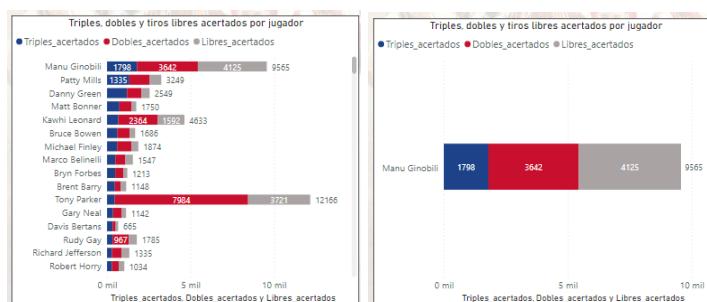


■ **Análisis por jugador:** Se incluyen tres filtros (por jugador, por equipo y por temporada).

Dos cards muestran el nombre del jugador y el equipo en el que jugó en la temporada seleccionada. Se incluyen 5 cards con estadísticas del jugador: promedio de minutos jugados, plus/minus, porcentaje de triples, dobles y tiros libres convertidos. Estas cards tienen fondo de color gris para que resalten por sobre los gráficos ya que muestran los principales parámetros que se quieren mostrar de un jugador.



Siguiendo el patrón Z el usuario se encuentra con un gráfico de barras apiladas muestra los triples, dobles y tiros libres logrados por cada jugador. Si se selecciona un equipo en el filtro se muestran sólo los que jugaron allí, si se selecciona además temporada se muestran los que jugaron en el equipo y temporada seleccionados. Si se selecciona un jugador muestra solamente la barra del jugador elegido.



Tres gráficos de torta muestran los triples, dobles y libres intentados (azul) y convertidos (rojo). Permite ver la efectividad de conversión de un cierto jugador.



Finalmente, y como cierre de la pestaña, una medida con parámetro muestra si el porcentaje de triples del jugador seleccionado es mayor o menor a cierto valor.

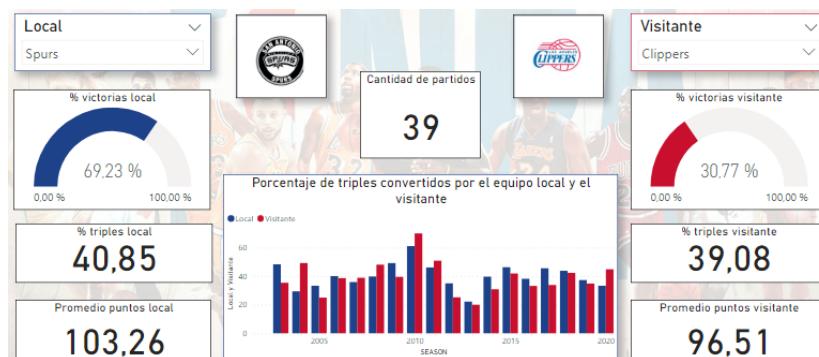


■ **Análisis por enfrentamientos:** Además del filtro por temporada ubicado en la barra superior, esta pestaña incluye un filtro para seleccionar equipo local y uno para visitante.

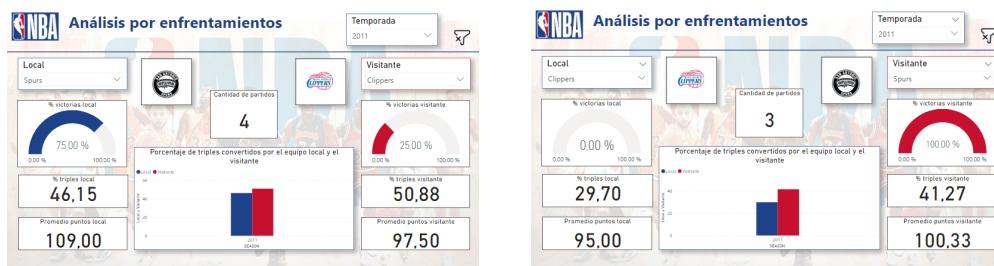
Esta pestaña no cumple estrictamente con el patrón Z ya que debe ser leída desde arriba hacia abajo, y cada lado muestra los datos de un equipo (local a la izquierda y visitante a la derecha). Al seleccionar los equipos se muestran las estadísticas de los partidos en los que se enfrentaron los dos equipos seleccionados. Esto permite evaluar si el equipo que hizo más cantidad de puntos

fue el que tuvo un porcentaje de triples mayor, y si es más probable que haga más triples un equipo cuando juega de local o de visitante.

En la parte central de la pestaña, se muestra una card con la cantidad de enfrentamientos analizados, y un gráfico de columnas agrupadas muestra la cantidad de triples convertidos por uno y otro equipo en cada año en enfrentamientos entre ellos (siendo local o visitante según se seleccionó). Permite ver como cambió la cantidad de triples convertidos por los locales y visitantes a lo largo de los años.



Es necesario aclarar que no se muestran todos los enfrentamientos entre un par de equipos, sino solamente en los que cada uno de ellos fue local o visitante según se haya seleccionado. Si la selección se invierte se mostrarán diferentes datos tal como ejemplifico en la siguiente figura:



### 3.4. Tablas, columnas y medidas calculadas

En esta sección se detalla cómo se llevaron a cabo los cálculos necesarios para el desarrollo del dashboard. Se incluyen todos los cálculos realizados, incluso los que se hicieron para presentar en algún desafío y no fueron utilizados en la versión final del trabajo.

#### Tabla calendario

A partir de la tabla games se generó la tabla\_calendario. Para esto se duplicó la tabla, se eliminaron todas las columnas excepto GAME\_DATE\_EST y Season, y se quitaron los duplicados. Luego esta tabla es vinculada en el diagrama relacional mediante una relación de uno a muchos con la tabla games, y con la tabla games\_details (en este segundo caso pasando por la tabla games\_puente). Fue necesario incluir en la tabla calendario la columna Season, ya que las temporadas no son por año calendario, sino de mitad a mitad de año. La columna Season es la que se usa para filtrar en todas las pestañas.

#### Columnas calculadas

Se generaron tres nuevas columnas en la tabla games\_details (presentadas en el desafío ‘Columnas y medidas calculadas’ el día 24/11/21 y no utilizadas en la versión final). En todos los casos se incluyó la función IFERROR para evitar que apareciera NaN al dividir por cero. Estas son:

- Porcentaje\_triples: indica qué porcentaje de los puntos anotados por un jugador fueron obtenidos por triples.
- ```
IFERROR(Porcentaje_triples = 100*(3*game_details[FG3M])/game_details[PTS],0)
```
- Porcentaje\_dobles: indica qué porcentaje de los puntos anotados por un jugador fueron obtenidos por dobles. Para obtenerlo debemos restar el número de triples al número de tiros de campo acertados.
- ```
IFERROR(Porcentaje_dobles = 100*2*(game_details[FGM])-game_details[FG3M])/game_details[PTS],0)
```
- Porcentaje\_libres: indica qué porcentaje de los puntos anotados por un jugador fueron obtenidos por tiros libres.
- ```
IFERROR(Porcentaje_libres = 100*(game_details[FTM])/game_details[PTS],0)
```

También se calcularon las columnas Minutos\_jugados y Segundos\_jugados en la tabla game\_details, en las que se calculan cuantos minutos (y segundos) jugó un cierto jugador en una partida. Esto fue necesario ya que el formato de la columna MIN no permitía graficar. Por ejemplo se pasó de 26:33 en la columna MIN (26 minutos y 33 segundos) a 26 en la columna Minutos\_jugados y 33 en la columna Segundos\_jugados. Para hacer esto fui a Trasnformar datos>Aregar columna y duplique la comununa MIN. Luego me posicioné sobre la columna duplicada y fui a Extraer>Texto antes del delimitador. Allí seleccioné que se quedara con el texto adelante de los dos puntos. Estos serán los munitos jugados. Repetí el proceso quedándome con el texto después de los dos puntos usando Extraer>Texto después del delimitador, y así cree la columna con los segundos.

Luego utilizando una función DAX se construyó la columna Tiempo\_jugado que es utilizada en la pestaña **Jugadores**:

- Tiempo\_jugado=DIVIDE(60\*game\_details[Minutos\_jugados]+game\_details[Segundos\_jugados],60)

## Medidas calculadas

A continuación se detallan las medidas calculadas que se encuentran en la tabla\_medidas. Algunas de ellas se calcularon para cumplir con la consigna del agún desafío pero no fueron utilizadas en la versión final del proyecto.

Las cinco medidas que se listan a continuación fueron creadas para la entrega de uno del desafío 'Columnas y medidas calculadas' (presentado el 24/11/21) y luego no fueron utilizadas.

- Promedio de puntos hechos por los equipos cuando juegan de local.
- ```
Promedio_puntos_local = AVERAGE(games[PTS_home])
```
- Promedio de puntos hechos por los equipos cuando juegan de visitante.
- ```
Promedio_puntos_visitante = AVERAGE(games[PTS_away])
```
- Número total de partidos que figuran en el dataset.
- ```
Nro_partidos = COUNT(games[HOME_TEAM_WINS])
```
- Número de partidos ganados por un equipo local (la columna HOME\_TEAM\_WINS contiene un 1 si ganó el local y un 0 si ganó el visitante).
- ```
Partidos_ganados_local = SUM(games[HOME_TEAM_WINS])
```
- Número de partidos ganados por un equipo visitante.
- ```
Partidos_ganados_visitante = COUNT(games[HOME_TEAM_WINS]) - SUM(games[HOME_TEAM_WINS])
```

Otras medidas calculadas que fueron parte de la segunda preentrega del proyecto (presentada el 20/12/21) son las siguientes:

- Dobles intentados. Se calcula como tiros de campo menos triples intentados (no se usó en la versión final).

```
Dobles_intentados = SUM(game_details[FGA])-SUM(game_details[FG3A])
```

- Dobles acertados. Se calcula como tiros de campo menos triples acertados (se utiliza en las pestañas **Equipos** y **Jugadores**).

`Dobles_acertados = SUM(game_details[FGM]) -SUM(game_details[FG3M])`

- Libres acertados. Se calcula simplemente como al suma de los tiros libres acertados. Se aplica en la tabla de la pestaña **Equipos** y en el gráfico de barras de la pestaña **Jugadores**. Para hacer estas visualizaciones no hubiese sido imprescindible generar esta medida, ya que podría haber utilizado directamente el campo `game_details[FTM]` y setearlo para que muestre la suma. Sin embargo me pareció más prolíjo calcular la medida para que aparezca en la `tabla_medidas` junto con `Dobles_acertados`.

`Libres_acertados = SUM(game_details[FTM])`

- Triples acertados. Análogamente a los libres se calcula simplemente como la suma de los triples acertados. Para este caso valen las mismas aclaraciones que para la medida descripta anteriormente. Se usa en las pestañas **Equipos** y **Jugadores**.

`Triples_acertados = SUM(game_details[FG3M])`

- Porcentaje de dobles convertidos (se utiliza en la pestaña **Jugadores**).

`Porcentaje_dobles_jugador = 100*DIVIDE(2*(SUM(game_details[FGM]) -SUM(game_details[FG3M])), SUM(game_details[PTS]))`

- Porcentaje de triples convertidos (se utiliza en la pestaña **Jugadores**).

`Porcentaje_libres_jugador = 100*DIVIDE(SUM(game_details[FTM]), SUM(game_details[PTS]))`

- Porcentaje de triples convertidos (se utiliza en la pestaña **Jugadores**).

`Porcentaje_triples_jugador = 100*DIVIDE(3*SUM(game_details[FG3M]), SUM(game_details[PTS]))`

Finalmente a continuación listo las medidas calculadas que no fueron presentadas en ningún desafío ni preentrega y aparecen en la versión final del proyecto:

- Porcentaje de triples convertidos. Esta medida calcula el cociente entre los triples convertidos y los triples intentados y lo multiplica por 100. Da información sobre la efectividad en la conversión y se representa en un gráfico de áreas en la pestaña **Equipos**.

`%triples_convertidos = 100*DIVIDE(SUM(game_details[FG3M]), SUM(game_details[FG3A]))`

- Promedio de puntos del local. Para calcular el promedio de los puntos logrados por un equipo de local sumo la cantidad de puntos y la divido por la cantidad de partidos jugados. Se usa en la pestaña **Enfrentamientos**.

`Puntos_local = DIVIDE(SUM(games[PTS_home]), COUNT(games[GAME_ID]))`

- Promedio de puntos del visitante. Analogamente a la medida anterior calculo los puntos logrados por un equipo jugado como visitante. Se usa en la pestaña **Enfrentamientos**.

`Puntos_visitante = DIVIDE(SUM(games[PTS_away]), COUNT(games[GAME_ID]))`

- Puntos por partido obtenidos mediante tiros de dos puntos. Para este cálculo se le resta la cantidad de tiros de tres puntos al total de tiros de campo y se multiplica por 2, que es lo que vale cada acierto. Para poder calcular el valor promedio por partido es necesario dividir por el numero total de partidos. Esto se hace multiplicando por dos la cuenta de cantidad de partidos. Es necesario este factor dos ya que la cantidad de ID de partidos en el dataset es la mitad de los partidos jugados por cada equipo, ya que en cada partido participan dos equipos. Esta medida se utiliza en el gráfico de áreas apiladas de la pestaña **Equipos**.

`Puntos_por_dobles_por_partido = DIVIDE(2*(SUM(game_details[FGM]) -SUM(game_details[FG3M])), 2*COUNT(games[game_ID]))`

- Puntos por partido obtenidos mediante tiros de tres puntos. Se multiplica la cantidad de aciertos por tres y se divide por el número de partidos. Se utiliza en el gráfico de áreas apiladas de la pestaña **Equipos**.

```
Puntos_por_triples_por_partido = DIVIDE(3*SUM(game_details[FGM]), 2*COUNT(games[game_ID]))
```

- Puntos por partido obtenidos mediante tiros libres. Directamente se divide la cantidad de aciertos por el número de partidos. Se utiliza en el gráfico de áreas apiladas de la pestaña **Equipos**.

```
Puntos_por_libres_por_partido = DIVIDE(SUM(game_details[FGM]), 2*COUNT(games[game_ID]))
```

- Porcentaje de triples del local. Se calcula como el promedio de la columna games[FG3\_PCT\_home] multiplicada por 100. Se muestra en una card en la pestaña **Enfrentamientos**.

```
Triples_local = 100*AVERAGE(games[FG3_PCT_home])
```

- Porcentaje de triples del visitante. Se calcula como el promedio de la columna games[FG3\_PCT\_away] multiplicada por 100. Se muestra en una card en la pestaña **Enfrentamientos**.

```
Triples_visitante = 100*AVERAGE(games[FG3_PCT_away])
```

### 3.5. Medidas claculadas avanzadas

#### KPI

Se generó un KPI que se visualiza en la pestaña **Equipos**. Este compara la cantidad de triples hechos por un equipo en la temporada 2003 (primera del set de datos) y la 2020 (última en el set de datos). Use la medida Triples\_2003 como Objetivo de destino, y Triples\_2020 como indicador. Este KPI permite ver como cambió el número de triples en el total de los años analizados.

- Triples\_2003 = CALCULATE(SUM(game\_details[FG3M]), tabl\_calendario\_2[Año] = 2003)
- Triples\_2020 = CALCULATE(SUM(game\_details[FG3M]), tabl\_calendario\_2[Año] = 2020)

#### Medidas calculadas con variables

Se calcularon dos medidas que incluyen variables y fueron parte del desafío ‘Medidas calculadas avanzadas’ (entregado el 8/12/21). Una de ellas es el porcentaje de victorias logradas por el equipo local y otra por el visitante. Ambas incluyen variables con funciones de agregación que se muestran en la pestaña **Enfrentamientos** en un gráfico Medidor:

- Victorias\_local\_% =
 

```
VAR victorias_local = SUM(games[HOME_TEAM_WINS])
VAR total_enfrentamientos = COUNT(games[HOME_TEAM_WINS])
RETURN
DIVIDE(victorias_local, total_enfrentamientos)
```
- Victorias\_visitante\_% =
 

```
VAR victorias_local = SUM(games[HOME_TEAM_WINS])
VAR total_enfrentamientos = COUNT(games[HOME_TEAM_WINS])
RETURN
DIVIDE(total_enfrentamientos-victorias_local, total_enfrentamientos)
```

## Medidas calculadas con función de inteligencia de tiempo

Para el desafío ‘Medidas claculadas avanzadas’ se generaron tres card (que se mostraban en la pestaña **Equipos**) con la cantidad de triples hechos por un equipo el año que selecciona en el filtro, y en lo años anterior y siguiente. Para ello se generaron las siguientes medidas utilizando funcion de inteligencia de tiempo:

- Cantidad\_triples\_año\_anterior = CALCULATE(SUM(game\_details[FG3M]), PARALLELPERIOD(tabla\_calendario[GAME\_DATE\_EST].[Date], -12, MONTH))
- Cantidad\_triples\_año\_actual = CALCULATE(SUM(game\_details[FG3M]), PARALLELPPARALLELPERIOD(tabla\_calendario[GAME\_DATE\_EST].[Date], +0, MONTH))
- Cantidad\_triples\_año\_siguiente = CALCULATE(SUM(game\_details[FG3M]), PARALLELPPARALLELPERIOD(tabla\_calendario[GAME\_DATE\_EST].[Date], +12, MONTH))

Estas cards fueron eliminadas de la versión final, ya que ocupaban una espacio muy grande en la pestaña complicando el diseño de la misma, y la información que mostraban se podía interpretar a partir del gráfico de áreas, por lo que la información estaba repetida.

## Medidas calculadas con parámetro

También como parte del desafío ‘Medidas calculadas avanzadas’ se generó un parámetro que varía entre 10 y 45 cada 5. Este parámetro se incluyó en la medida porcentaje\_triples\_parámetro definida como sigue:

- Porcentaje\_triples\_parametro = IF((100\*(3\*SUM(game\_details[FG3M])/SUM(game\_details[PTS])))>'Parámetro'[Valor Parámetro], CONCATENATE("Porcentaje triples mayor a ", [Valor Parámetro]), CONCATENATE("Porcentaje triples menor a ", [Valor Parámetro]))

Esta medida se usa para visualizar una card que nos muestra si un jugador tiene un porcentaje de triples mayor o menor que el valor del parámetros que se seleccione.

## Tooltips

Se generaron tres pestañas con tooltips:

- **tooltip:** Grafico de torta con cantidad de triples, dobles y tiros libres acertados por equipo. Relacionado con la tabla en la pestaña **Equipos**.
- **tooltip2:** Contiene dos cards con el nombre y la capacidad del estadio. Se relaciona con el mapa de la pestaña **Equipos**.
- **tooltip3:** Contiene tres cards con el nombre del entrenador, el dueño y el año de fundación de un equipo. Se relaciona con una de las cards la pestaña **Equipos**.

## **Capítulo 4**

### **Futuras líneas**

La cantidad de análisis que se pueden llevar adelante a partir de datos de la NBA es inmensa. Sin embargo hay dos aspectos en particular que están relacionados con la cantidad de tiros de tres puntos convertidos y que me resultaían interesantes de analizar.

- Dado que la cantidad de triples intentados/convertidos ha ido en aumento pero el tiempo de duración de los partidos sigue siendo el mismo existen dos posibilidades: o bien se tiran menos tiros de dos puntos, o el tiempo de las posesiones es menor. Resultaría entonces interesante evaluar cuál de estas dos situaciones es la que realmente ocurre. Para ello necesitaría nuevos datos que contengan los tiempos de posesión.
- Generalmente los jugadores más altos de la NBA son los que han jugado cerca del aro, y por lo tanto los que habitualmente convertían mayor cantidad de tiros de dos puntos. Resultaría interesante analizar la eficiencia en la conversión de dobles y triples en función de la posición que los jugadores ocupan en la cancha para evaluar si el cambio en la forma de juego está haciendo que todos los jugadores ejecuten más tiros de tres puntos, independientemente de la posición que ocupen.

# Apéndice A

## Glosario

En este apéndice incluimos una lista con los nombres de las columnas de las cuatro tablas que componen el dataset y el contenido de cada una de ellas.

🌐 **game\_details** (detalles del partido):

GAME_DETAIL_ID	ID del detalle
GAME_ID	ID del juego
TEAM_ID	ID del equipo
TEAM_ABBREVIATION	Abreviación del nombre del equipo
TEAM_CITY	Ciudad del equipo
PLAYER_ID	ID del jugador
PLAYER_NAME	Nombre del jugador
START_POSITION	Posición del jugador al comienzo del partido
COMMENT	Comentarios
MIN	Minutos jugador por el jugador
FGM	Tiros de campo acertados (Field Goal Made)
FGA	Tiros de campo intentados (Field Goal Attempts)
FG_PCT	Porcentaje de tiros de campo acertados
FG3M	Triples acertados
FG3A	Triples intentados
FG3_PCT	Porcentaje de triples acertados
FTM	Tiros libres acertados (Free Throw Made)
FTA	Tiros libres intentados (Free Throw Attempts)
FT_PCT	Porcentaje tiros libres acertados
OREB	Rebotes ofensivos (Offensive Rebounds)
DREB	Rebotes defensivos (Defensive Rebounds)
REB	Rebotes totales
AST	Asistencias
STL	Robos
BLK	Bloqueos
TO	Pérdidas (Turnovers)
PF	Faltas personales
PTS	Puntos anotados por el jugador
PLUS_MINUS	Más/menos de un jugador respecto al rendimiento del equipo por cada 100 posesiones

 **games** (partidos):

GAME_DATE_EST	Fecha del partidos
GAME_ID	ID del partido
GAME_STATUS_TEXT	Estado del partido
HOME_TEAM_ID	ID del equipo local
VISITOR_TEAM_ID	ID del equipo visitante
SEASON	Temporada
TEAM_ID_home	ID del equipo local
PTS_home	Puntos del equipo local
FG_PCT_home	Porcentaje de tiros de campo del local
FT_PCT_home	Porcentaje de tiros libres del local
FG3_PCT_home	Porcentaje de tiros triples del local
AST_home	Asistencias del local
REB_home	Rebotes del local
TEAM_ID_away	ID del equipo visitante
PTS_away	Puntos del equipo visitante
FG_PCT_away	Porcentaje de tiros de campo del visitante
FT_PCT_away	Porcentaje de tiros libres del visitante
FG3_PCT_away	Porcentaje de tiros triples del visitante
AST_away	Asistencias del visitante
REB_away	Rebotes del visitante
HOME_TEAM_WINS	¿Ganó el local?

 **players** (jugadores):

PLAYER_NAME	Nombre del jugador
TEAM_ID	ID del equipo
PLAYER_ID	ID del jugador
SEASON	Temporada

 **teams** (equipos):

TEAM_ID	ID equipo
ABBREVIATION	Abrevación ciudad
NICKNAME	Apodo del equipo
YEARFOUNDED	Año en que se fundó
CITY	Ciudad en que se encuentra el estadio
ARENA	Nombre del estadio
ARENACAPACITY	Capacidad del estadio
OWNER	Propietario
GENERALMANAGER	Manager
HEADCOACH	Entrenador

## Apéndice B

### Pestañas del tablero



**NBA** Análisis por equipos

Equipo: Hornets Temporada: Todas

CHA Hornets Charlotte

Ubicación estadio: Charlotte, NC

Triplazos convertidos entre 2003 y 2020: 1066 (+Infinity %)

NICKNAME	Triples	Dobles	Libres
76ers	11338	48387	28925
Bucks	12416	48844	28707
Bulls	11848	46813	28754
Cavaliers	13512	48689	30551
Celtics	13870	50673	31625
Clippers	12336	49285	31555
Grizzlies	10751	49431	29847
Hawks	12921	47663	29302
Heat	13796	50611	31284
Hornets	10914	42597	27418
Jazz	11599	46940	31409
Kings	11873	47870	30700
<b>Total</b>	<b>385496</b>	<b>1458155</b>	<b>899482</b>

Promedio de puntos por partido

% de triples convertidos sobre intentados

Ranking de jugadores con más puntos

Ranking de jugadores con más triples

