



# **ANÁLISIS DE CASO DE NEGOCIO**

**GRUPO 5**

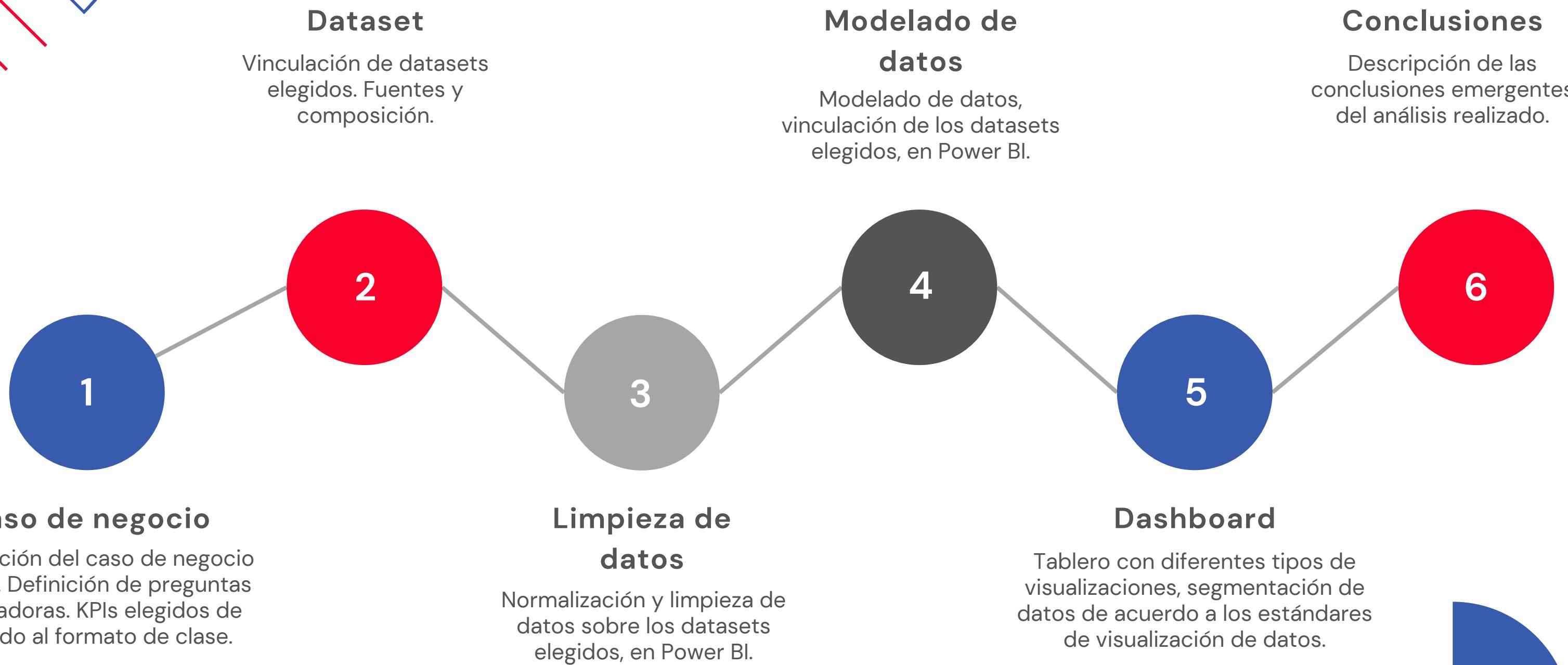
**LABORATORIO DE RECOPILACIÓN  
E INTEGRACIÓN DE DATOS**

"El básquet es más grande que anotar puntos y ganar campeonatos. Es más de si uno ha sido capaz de inspirar a la gente por un instante a que quieran ser una mejor versión de sí mismos"

**KOBE BRYANT**



# LÍNEA DE TIEMPO ANÁLISIS DE CASO



# 1 | CASO DE NEGOCIO

Se eligió trabajar como caso de negocio la NBA (National Basketball Association), nacida en 1946. El desarrollo de la historia del básquet llevó a que hoy en día la analítica sea una parte esencial de su desarrollo.

*"El Big Data y la analítica están detrás de la forma en que se juega al baloncesto de alto nivel hoy en día. La intuición y las sensaciones han dado paso al estudio pormenorizado de los datos para ganar más partidos. ¿Pueden los números y el talento convivir y hacer mejor al baloncesto?"*

Albert Vázquez | Cómo el Big Data cambió la NBA

## PROBLEMÁTICA ANALIZADA:

Optimizar la gestión de un entrenador al frente de un equipo de la NBA, optimizando así su estrategia de juego y la toma de decisiones tácticas.

## PREGUNTAS DISPARADORAS

1. ¿Qué jugadores tienen mayor impacto en el resultado de un juego?
2. ¿Cómo evoluciona el rendimiento de los equipos a lo largo de las temporadas?
3. ¿Existe una correlación entre la antigüedad de un equipo y su éxito a largo plazo?

# 1 | CASO DE NEGOCIO

## KPIs INDICADORES CLAVES:

Se eligieron dos indicadores claves, que ayudarán a trabajar con la problemática desarrollada en el punto anterior.

Para ello que se analizó que los puntos por partido y la eficiencia por jugador, serán determinantes a la hora de analizar el impacto de la gestión del entrenador.

KPI	UNIDAD	DIMENSIONES	FÓRMULA
Puntos por partido	Puntos	Partido	Total puntos / total partidos jugados
Eficiencia por jugador	Puntos de eficiencia	Partido	$((\text{pts} + \text{reb} + \text{ast} + \text{stl} + \text{blk}) - ((\text{fga} - \text{fgm}) + (\text{fta} - \text{ftm}) + \text{to})) / \text{partidos}$

# 2 | DATASET

## DESCRIPCIÓN

Se utilizaron cinco (5) dataset con todos los juegos de NBA de todas las temporadas desde 2004 a diciembre 2020. Se encuentran completos y vinculados entre sí,

1. Detalle de juegos
2. Jugadores
3. Ranking
4. Equipos
5. Juegos

Fuente: NBA

Enlace a los set de datos: [set de datos de kaggle](#).

The image displays five Excel spreadsheets representing NBA datasets:

- TEAM\_ID:** Contains columns for TEAM\_ID, TEAM\_ABBREVIATION, TEAM\_CITY, and various player statistics.
- PLAYER\_ID:** Contains columns for PLAYER\_ID, PLAYER\_NAME, NICKNAME, and game statistics.
- TEAM\_ID:** Contains columns for TEAM\_ID, LEAGUE\_ID, SEASON\_ID, STANDINGSDATE, CONFERENCE, TEAM, and game statistics.
- LEAGUE\_ID:** Contains columns for LEAGUE\_ID, TEAM\_ID, MIN\_YEAR, MAX\_YEAR, ABBREVIATION, NICKNAME, YEARFOUNDED, CITY, ARENA, ARENACAPACITY, and OWNER.
- GAME\_DATE\_EST:** Contains columns for GAME\_DATE\_EST, GAME\_ID, GAME\_STATUS\_TEXT, HOME\_TEAM\_ID, VISITOR\_TEAM\_ID, SEASON, TEAM\_ID\_home, PTS\_home, FG\_PCT\_home, FT\_PCT\_home, and FG3\_PCT\_home.

The second TEAM\_ID sheet has a green header row and highlights the Denver Nuggets (row 1610612743). The LEAGUE\_ID sheet also highlights the Denver Nuggets (row 1610612742).

# 3 | LIMPIEZA Y TRANSFORMACIÓN DE DATOS

## REALIZADO EN POWER BI:

De acuerdo a las preguntas disparadoras, se normalizaron algunos datos que presentaban errores o anomalías, se modificaron nombres de campos claves, se eliminaron columnas que no serían de utilidad para el posterior análisis, se modificaron columnas que presentaban errores de formato y otras para mejorar la legibilidad para su posterior uso (división de nombres, fechas, entre otros).

Dentro de las transformaciones, principalmente se estandarizaron los tipos de datos.

No se hallaron cantidades significativas de datos vacíos o nulos en el set de datos analizado.

```
1 let
2     Origen = Csv.Document(File.Contents("D:\OneDrive\Escritorio\Carrera\NBA\games_details.csv"),[Delimiter=",", Columns=29,
3         Encoding=1252, QuoteStyle=QuoteStyle.None]),
4     #"Cambiar tipo" = Table.TransformColumnTypes(Origen,{{"Column1", type text}, {"Column2", type text}, {"Column3", type text},
5         {"Column4", type text}, {"Column5", type text}, {"Column6", type text}, {"Column7", type text}, {"Column8", type text},
6         {"Column9", type text}, {"Column10", type text}, {"Column11", type text}, {"Column12", type text}, {"Column13", type text},
7         {"Column14", type text}, {"Column15", type text}, {"Column16", type text}, {"Column17", type text}, {"Column18", type text},
8         {"Column19", type text}, {"Column20", type text}, {"Column21", type text}, {"Column22", type text}, {"Column23", type text},
9         {"Column24", type text}, {"Column25", type text}, {"Column26", type text}, {"Column27", type text}, {"Column28", type text},
10        {"Column29", type text}}),
11    #"Encabezados promovidos" = Table.PromoteHeaders(#"Cambiar tipo", [PromoteAllScalars=true]),
12    #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",{"GAME_ID", Int64.Type}, {"TEAM_ID", Int64.Type},
13        {"PLAYER_ID", Int64.Type}),
14    #"Valor reemplazado" = Table.ReplaceValue(#"Tipo cambiado",".",",",Replacer.ReplaceText,{"FGM", "FGA", "FG_PCT", "FG3M", "FG3A",
15        "FG3_PCT", "FTM", "FTA", "FT_PCT", "OREB", "REB", "AST", "STL", "BLK", "TO", "PF", "PTS", "PLUS_MINUS"}),
16    #"Tipo cambiado1" = Table.TransformColumnTypes(#"Valor reemplazado",{"FGM", Int64.Type}, {"FGA", Int64.Type}, {"FG3M",
17        Int64.Type}, {"FG3A", Int64.Type}, {"FTM", Int64.Type}, {"FTA", Int64.Type}, {"OREB", Int64.Type}, {"REB", Int64.Type},
18        {"AST", Int64.Type}, {"STL", Int64.Type}, {"BLK", Int64.Type}, {"TO", Int64.Type}, {"PF", Int64.Type},
19        {"PTS", Int64.Type}, {"PLUS_MINUS", Int64.Type}, {"FG_PCT", Percentage.Type}, {"FG3_PCT", Percentage.Type}, {"FT_PCT",
20        Percentage.Type}),
21    #"Columnas quitadas" = Table.RemoveColumns(#"Tipo cambiado1","NICKNAME"),
22    #"Dividir columna por delimitador" = Table.SplitColumn(#"Columnas quitadas", "PLAYER_NAME", Splitter.SplitTextByEachDelimiter({".",
23        }), QuoteStyle.Csv, false), {"PLAYER_NAME.1", "PLAYER_NAME.2"},
24    #"Tipo cambiado2" = Table.TransformColumnTypes(#"Dividir columna por delimitador",{"PLAYER_NAME.1", type text},
25        {"PLAYER_NAME.2", type text}),
26    #"Columnas con nombre cambiado" = Table.RenameColumns(#"Tipo cambiado2",{"PLAYER_NAME.1", "PLAYER_NAME"}, {"PLAYER_NAME.2",
27        "PLAYER_LASTNAME"}),
28    #"Personalizada agregada" = Table.AddColumn(#"Columnas con nombre cambiado", "START_POSITION_ADJ", each if [START_POSITION] = "".
29        and [MIN] <> "" then "Substitute" else if [MIN] = "" then "DNP" else [START_POSITION]),
30    #"Columnas reordenadas" = Table.ReorderColumns(#"Personalizada agregada","GAME_ID", "TEAM_ID", "TEAM_ABBREVIATION", "TEAM_CITY",
31        "PLAYER_ID", "PLAYER_NAME", "PLAYER_LASTNAME", "START_POSITION", "START_POSITION_ADJ", "COMMENT", "MIN", "FGM", "FGA",
32        "FG_PCT", "FG3M", "FG3A", "FG3_PCT", "FTM", "FTA", "FT_PCT", "OREB", "REB", "AST", "STL", "BLK", "TO", "PF", "PTS",
33        "PLUS_MINUS"),
34    #"Filas filtradas" = Table.SelectRows(#"Columnas reordenadas", each true),
35    #"Columnas quitadas1" = Table.RemoveColumns(#"Filas filtradas","START_POSITION"),
36    #"Columnas con nombre cambiado1" = Table.RenameColumns(#"Columnas quitadas1",{"START_POSITION_ADJ", "START_POSITION"}),
37    #"Valor reemplazado1" = Table.ReplaceValue(#"Columnas con nombre cambiado1","F","Forward",Replacer.ReplaceText,"START_POSITION"),
38        ),
39    #"Valor reemplazado2" = Table.ReplaceValue(#"Valor reemplazado1","C","Center",Replacer.ReplaceText,"START_POSITION"),
40    #"Valor reemplazado3" = Table.ReplaceValue(#"Valor reemplazado2","G","Guard",Replacer.ReplaceText,"START_POSITION"),
41    #"Texto extraído después del delimitador" = Table.TransformColumns(#"Valor reemplazado3",{"COMMENT", each Text.AfterDelimiter
42        (",", "-"), type text}),
43    #"Valor reemplazado4" = Table.ReplaceValue(#"Texto extraído después del delimitador","","","0",Replacer.ReplaceValue,"MIN"),
44    #"Valor reemplazado5" = Table.ReplaceValue(#"Valor reemplazado4",":",",",Replacer.ReplaceText,"MIN"),
45    #"Tipo cambiado3" = Table.TransformColumnTypes(#"Valor reemplazado5",{"MIN", type number}),
46    #"Texto extraído antes del delimitador" = Table.TransformColumns(#"Tipo cambiado3",{"MIN", each Text.BeforeDelimiter(Text.From
47        (",", "es-AR"), ","), type text}),
48    #"Tipo cambiado4" = Table.TransformColumnTypes(#"Texto extraído antes del delimitador",{"MIN", Int64.Type}),
49    #"Filas filtradas1" = Table.SelectRows(#"Tipo cambiado4", each true)
50 in
51    #"Filas filtradas1"
```

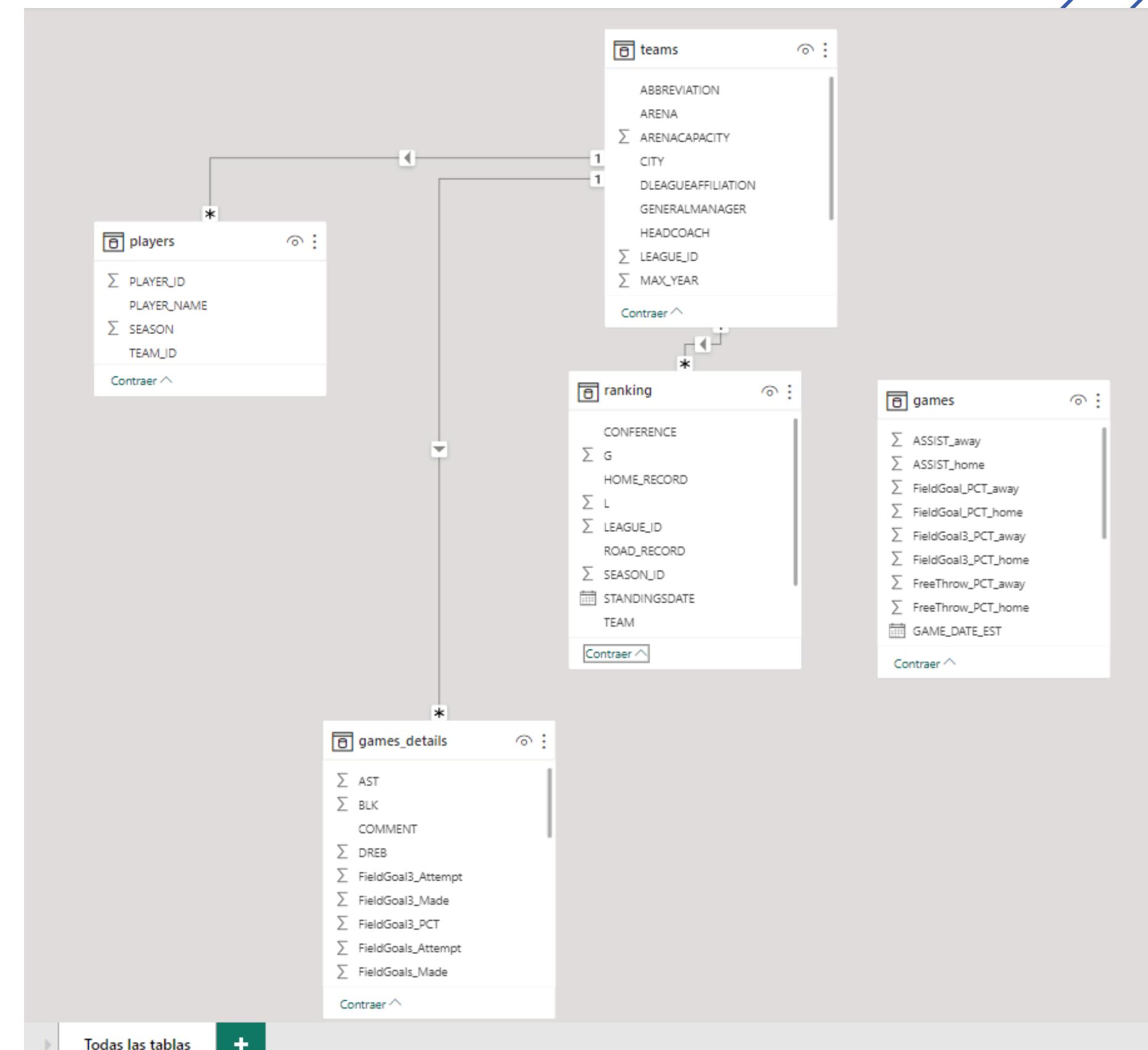
✓ No se han detectado errores de sintaxis.

# 4 | MODELADO DE DATOS

REALIZADO EN POWER BI:

De acuerdo a las preguntas disparadoras, posterior a la limpieza de datos, se realizaron las vinculaciones mediante claves primarias y foráneas para la vinculación entre las cinco tablas trabajadas.

Se realizaron modificaciones mediante el lenguaje DAX para la modelación de algunos campos importantes para el posterior análisis.



# 5 | DASHBOARD

**INFORME NBA:**  
POTENCIA TUS ESTRATEGIAS DE JUEGO  
CON DATOS HISTÓRICOS



Desempeño Jugadores

Desempeño Equipos

Evolución Equipos

Conclusión





## DESEMPEÑO JUGADORES

2004-2020

Buscar Temporada

Todas

Buscar Equipo

Todas

Buscar Jugador y posición

Todas

207,8

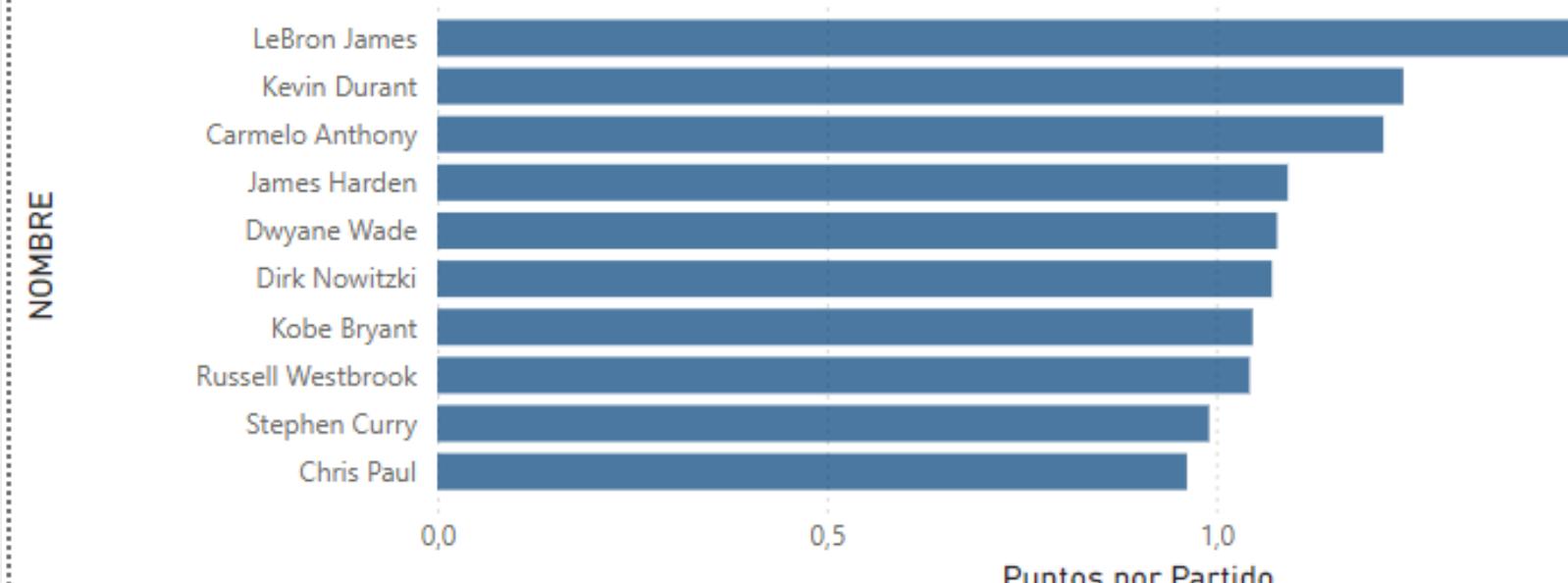
Puntos por Partido

0,13

Media Eficiencia

NOMBRE JUGADOR	POSICIÓN INICIAL	PUNTOS	ASISTENCIAS	REBOTE	ROBOS	TOMA DE BALÓN	EFICIENCIA
LeBron James	Delantero	43320	11613	12306	2490	5561	1,79
Chris Paul	Guardia	24412	12587	6098	2810	3268	1,24
Kevin Durant	Delantero	29719	4721	7806	1148	3360	1,16
Russell Westbrook	Guardia	26294	9629	8454	1908	4712	1,08
Stephen Curry	Guardia	25224	6575	4859	1664	3232	0,97
Dirk Nowitzki	Delantero	24448	2803	8512	829	1914	0,97
James Harden	Guardia	24596	7037	5420	1476	3841	0,95
Carmelo Anthony	Delantero	28680	3489	7938	1256	3190	0,94
Dwyane Wade	Guardia	25399	6065	5363	1774	3623	0,93
Total		5016658	1081866	2090126	369247	676582	215,70

Puntos por Partido por Jugador



Puntos por Partido y Eficiencia por Año





## DESEMPEÑO EQUIPOS 2004-2020

15 mil

Victorias por Temporada

Buscar Temporada

Todas

Buscar Equipo

Todas

### TOP 5 MAYORES VICTORIAS

Equipo	Victorias
Spurs	670
Heat	622
Warriors	596
Celtics	595
Mavericks	579
Total	3062

11 mil

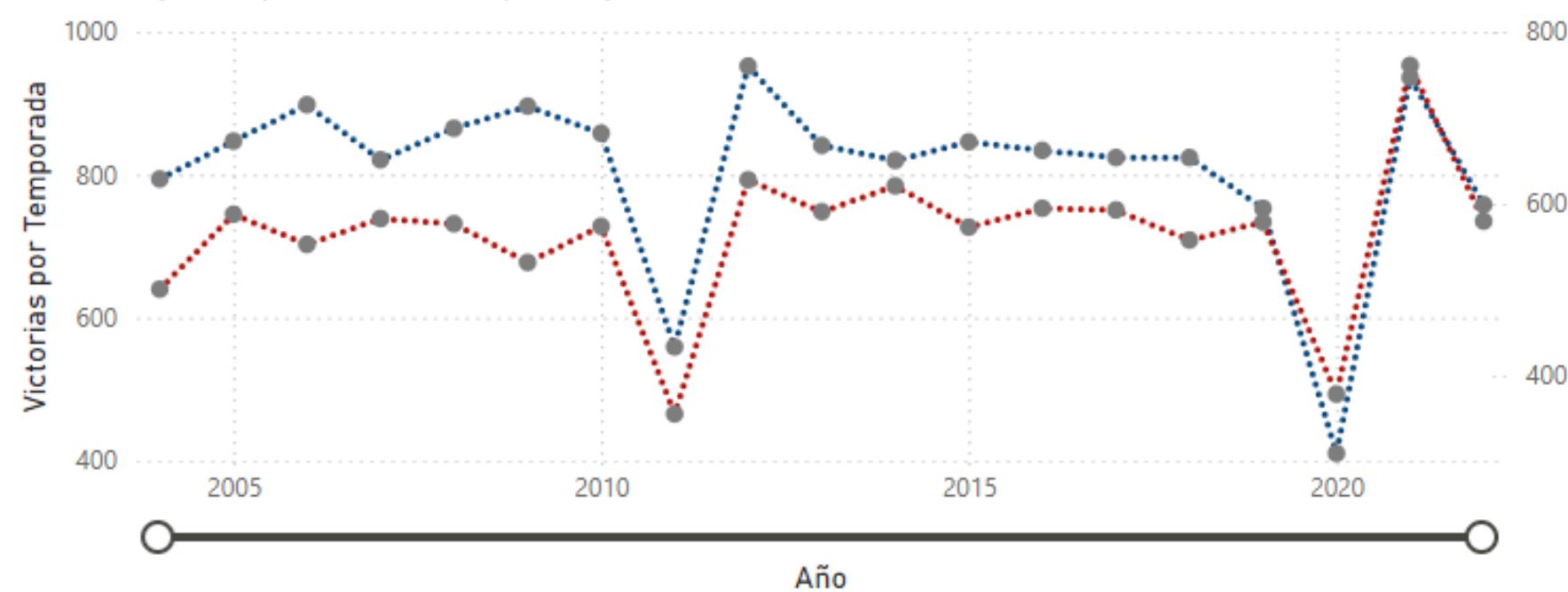
Derrotas por Temporada

### TOP 5 MAYORES DERROTAS

Equipo	Derrotas
Timberwolves	438
Knicks	433
Kings	418
Nets	418
Hornets	405
Total	2112

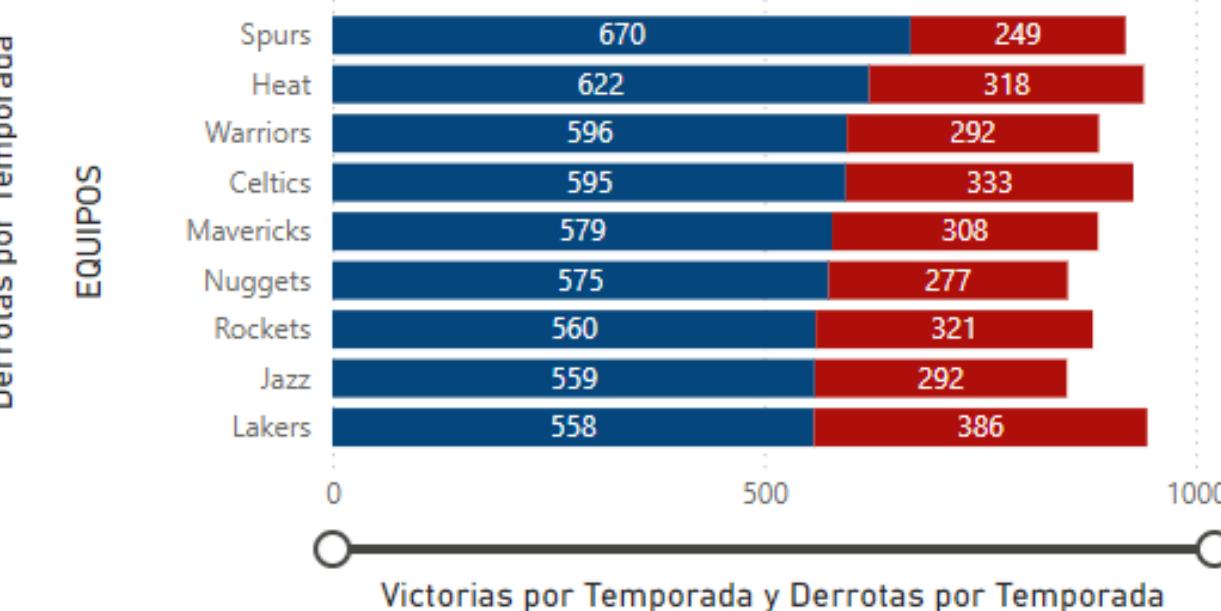
### Victorias por Temporada y Derrotas por Temporada por Año

● Victorias por Temporada ● Derrotas por Temporada



### Victorias y Derrotas por Temporada

● Victorias por Temporada ● Derrotas por Temporada





## EVOLUCIÓN EQUIPOS 2004-2020

78

Años de Existencia

15 mil

Victorias por Temporada

11 mil

Derrotas por Temporada

Buscar Temporada

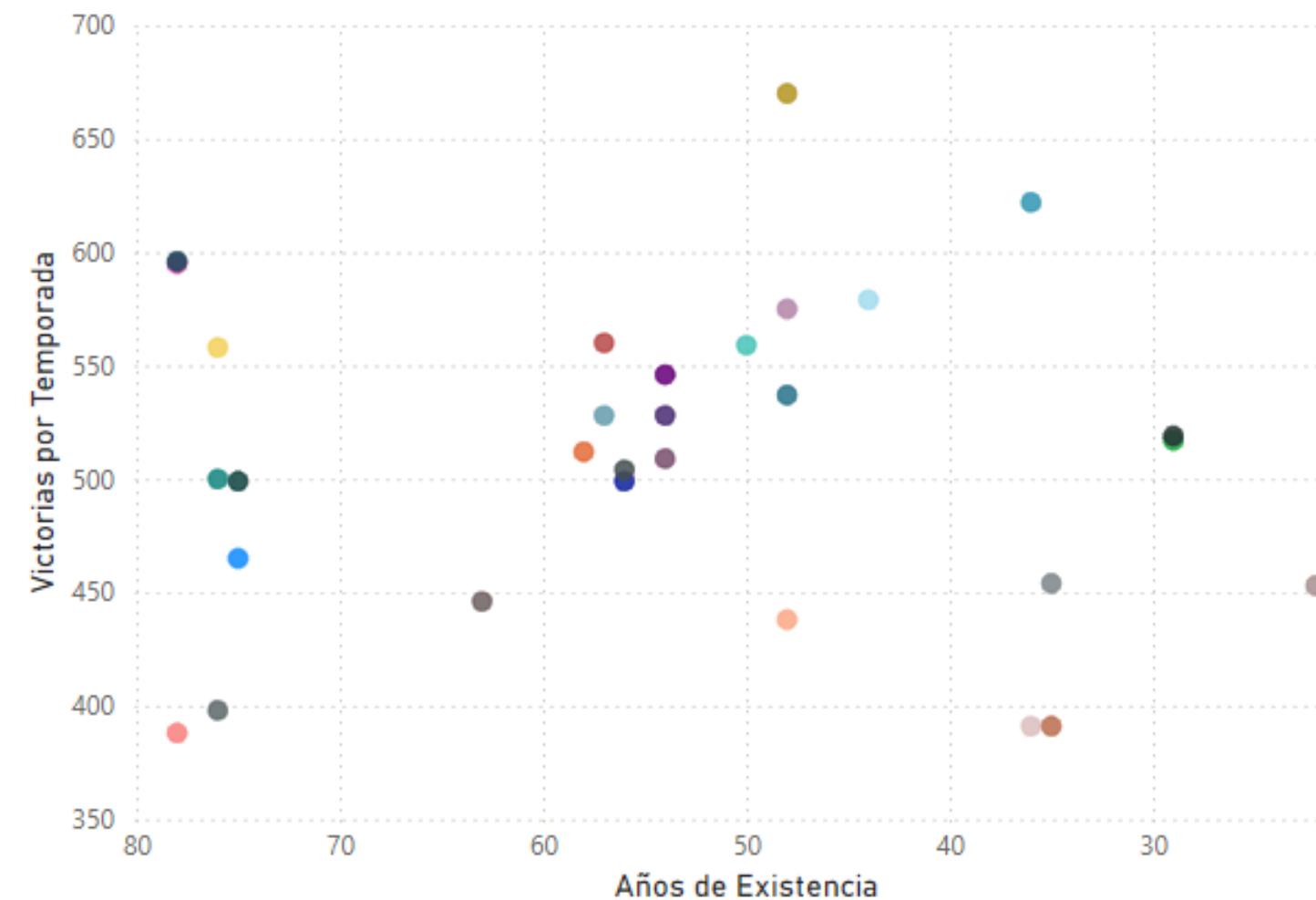
Todas

Buscar Equipo

Todas

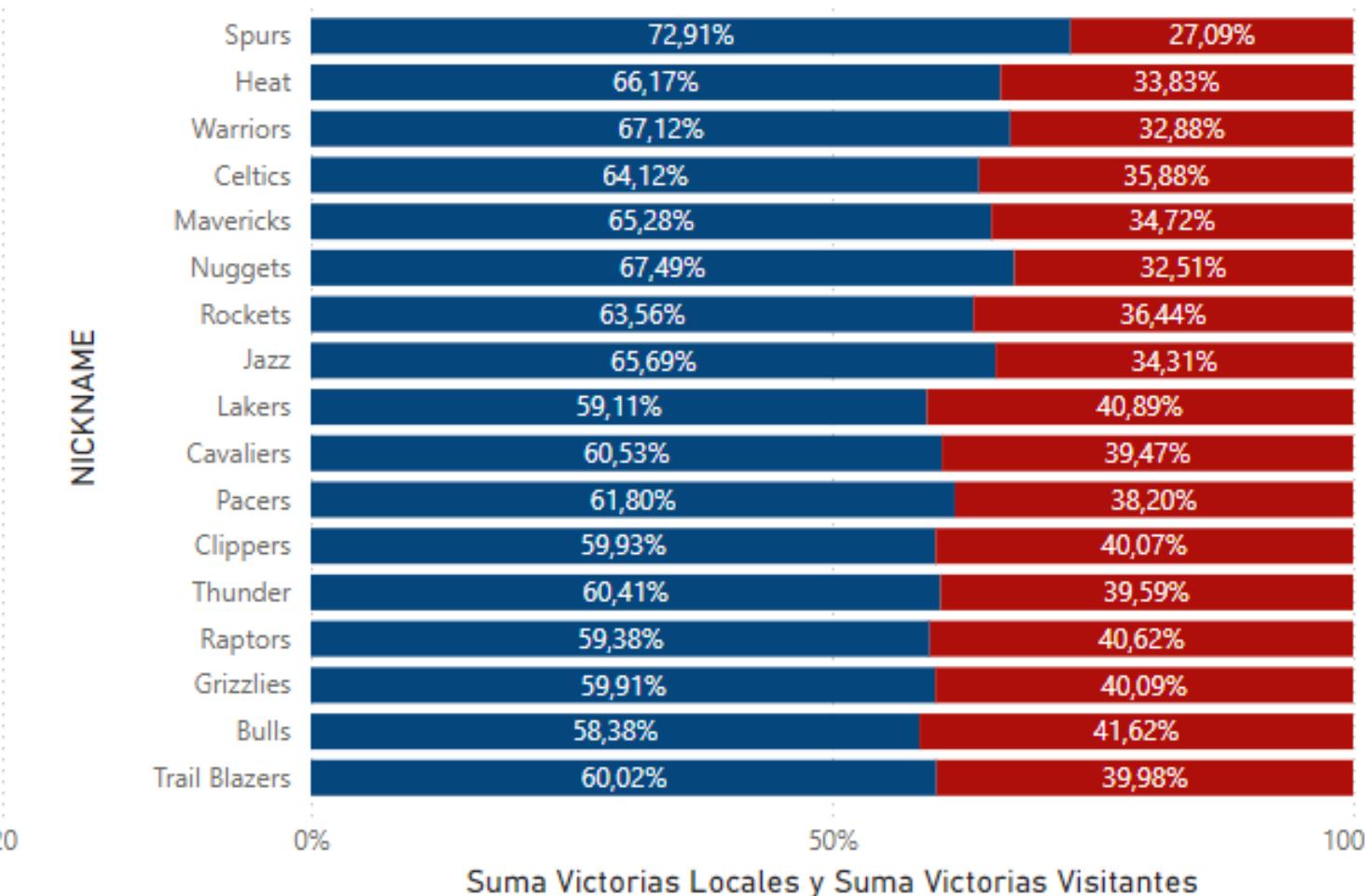
Años de Existencia y Victorias por Temporada por NICKNAME

NICKNAME ● 76ers ● Bucks ● Bulls ● Cavaliers ● Celtics ● Clippers ● Grizzlies ● Hawks ● Heat ►



Victorias locales y visitantes por Equipo

● Suma Victorias Locales ● Suma Victorias Visitantes



# 6 | CONCLUSIONES



La **eficiencia** es un valor esencial ya que permite identificar no solo a los mejores anotadores, sino también a los **jugadores** que contribuyen de manera más completa al equipo.

Según los datos analizados, dentro de este período histórico, podemos observar que la **media se ubica en el 0.13**, todos los jugadores que superen esa medida se consideran dentro del rango de eficiencia esperado.

1



En el desempeño de los **equipos** observados a un período de tiempo significativo, se puede analizar que los equipos con **mayor cantidad de victorias, se mantienen estables en el tiempo**.

Una explicación posible puede ser que los mismos tengan una base sólida de jugadores talentosos, una buena gestión, un manejo de lesiones más proactivo y estrategias efectivas, estos factores contribuyen a que mantengan un alto rendimiento a lo largo de las temporadas reduciendo la probabilidad de acumular muchas derrotas.

2



De acuerdo al gráfico de dispersión podemos observar **tres situaciones**:

Aquellos equipos que tienen **entre 80 y 60 años de antigüedad**, su desempeño fue muy variable, ya que encontramos a uno de los mejores equipos de la tabla histórica y a uno de los peores. Esto puede haberse debido a diversos factores de gestión e históricos.

Aquellos equipos que tienen **entre 60 y 40 años de antigüedad**, donde se encuentran la mayoría de los equipos, se observa un desempeño cercano a la media, es decir similar entre sí. Se observa además un único caso diferente (*outlier*) que es el de los Spurs, encontrándose muy por encima de la media, siendo el principal ganador histórico.

Aquellos equipos que tienen **entre 40 y 20 años de antigüedad**, se observa un rendimiento por debajo de la media, con excepción del equipo Heat que se ubica por encima de la misma convirtiéndose en el segundo de la lista de principales ganadores

3

ESTE INFORME FUE REALIZADO POR:

Tobias Pazos, Gabriel Perez,  
Ailén Iglesias Barrera, María Barreto,  
Santiago Alegre.

# **LABORATORIO DE RECOPILACIÓN E INTEGRACIÓN DE DATOS**

**PROF. AGUSTÍN CLAVEL**

## **EQUIPO 5**

Tobias Pazos  
Gabriel Perez  
Ailén Iglesias Barrera  
Maria Barreto  
Santiago Alegre