



# ANÁLISIS ESTADÍSTICO DE LA NBA

Academia Coderhouse

Comisión 31460

Grupo 56

Maldonado, Santiago

Salazar garzón, Facundo

Docente: Cucho, Cesar

Tutor: Benegas, Mauricio

Carrera: Data Scientist

Curso: Data Analytics

Fecha de entrega: 10/06/2022

## Contenido

Descripción del proyecto .....	1
Alcance.....	1
Base de datos.....	1
Diagrama Entidad-Relación inicial .....	2
Listado de tablas y columnas iniciales .....	3
Diagrama entidad-relación final .....	5
Columnas calculadas.....	5
Medidas calculadas .....	6
Gráficos y segmentadores .....	6

## Descripción del proyecto

La temática central para analizar en el proyecto será examinar la influencia de diferentes variables en el rendimiento de los jugadores y equipos de la NBA, con el objetivo de determinar los factores clave.

Se analizarán las diferentes variables de desempeño y rendimiento de los jugadores y equipos de la NBA con el objetivo de identificar la influencia del dinero (sueldos de jugadores y presupuesto de organización de la franquicia) principalmente y otras variables en los resultados obtenidos, tanto individuales como colectivos.

Además, se analizarán diferentes estadísticas con respecto a jugadores, entrenadores y equipos que resultan de mucho interés para la comunidad de la NBA.

## Alcance

Se determinó como banda temporal de análisis las temporadas comenzadas en la década del 2010, teniendo en cuenta que las temporadas están comprendidas entre años calendario diferentes.

## Base de datos

<https://docs.google.com/spreadsheets/d/1hjSXZeGUq2easNJCipX6GeW857Ijv2PZ/edit?usp=sharing&oid=104779111586528444744&rtpof=true&sd=true>

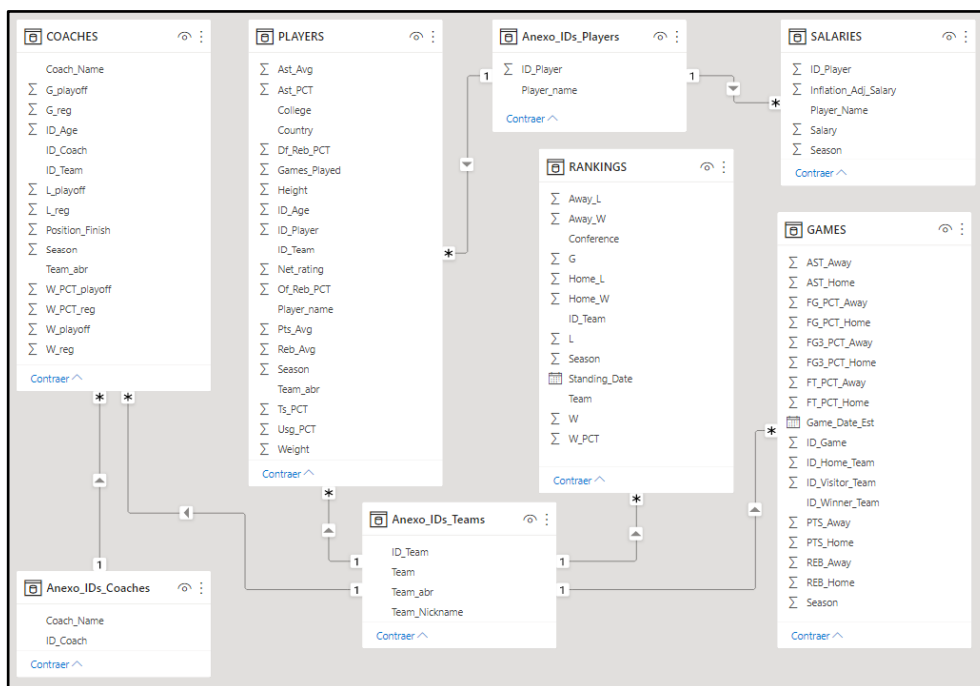
Aclaración: dado que es un archivo realizado en Excel y no en hojas de cálculo de Google, se recomienda descargarlo y abrirlo en Excel para evitar la pérdida de formato.

## Objetivo

El tablero busca representar el impacto del factor económico en el éxito individual y colectivo de los integrantes de la NBA.

Este tablero puede ser utilizado por jugador para ver sus propias estadísticas y la relación con su remuneración percibida y por los manager (director deportivo) de las franquicias de la NBA.

## Diagrama Entidad-Relación inicial



Activo	Desde: tabla (columna)	A: tabla (columna)
✓	COACHES (ID_Coach)	Puente_IDS_Coaches (ID_Coach)
✓	COACHES (ID_Team)	Puente_IDS_Teams (ID_Team)
✓	GAMES (ID_Winner_Team)	Puente_IDS_Teams (ID_Team)
✓	PLAYERS (ID_Player)	Puente_IDS_Players (ID_Player)
✓	PLAYERS (ID_Team)	Puente_IDS_Teams (ID_Team)
✓	RANKINGS (ID_Team)	Puente_IDS_Teams (ID_Team)
✓	SALARIES (ID_Player)	Puente_IDS_Players (ID_Player)

## Listado de tablas y columnas iniciales

Tabla Games					
Tipo de clave	Campo	Tipo de dato	Restriccion	Unique	Descripcion
<b>PK</b>	Game_Date_Est	DATE	<b>Not null</b>	<b>Unique</b>	Fecha en la que se jugo el partido
	<b>ID_Game</b>	<b>INT</b>			<b>ID unico de cada juego</b>
	ID_Home_Team	INT			ID unico de equipo local
	ID_Visitor_Team	INT			ID unico de equipo visitante
	Season	INT			Temporada de juego
	PTS_Home	INT			Puntos realizados por el equipo local
	FG_PCT_Home	DECIMAL			Tiros de campo realizados por equipo local
	FT_PCT_Home	DECIMAL			Tiros libres realizados por equipo local
	FG3_PCT_Home	DECIMAL			Tiros de campo de 3 puntos realizados por equipo local
	AST_Home	DECIMAL			Asistencias realizadas por equipo local
	REB_Home	DECIMAL			Rebotes realizados por equipo local
	PTS_Away	DECIMAL			Puntos realizaos por el equipo visitante
	FG_PCT_Away	DECIMAL			Tiros de campo realizados por equipo visitante
	FT_PCT_Away	DECIMAL			Tiros libres realizados por equipo visitante
	FG3_PCT_Away	DECIMAL			Tiros de campo de 3 puntos realizados por equipo visitante
<b>FK</b>	AST_Away	INT			Asistencias realizadas por equipo visitante
	REB_Away	INT			Rebotes realizados por equipo visitante
	<b>ID_Winner_Team</b>	<b>INT</b>			<b>ID del equipo ganador de dicho partido</b>

Tabla Rankings					
Tipo de clave	Campo	Tipo de dato	Restriccion	Unique	Descripcion
<b>PK</b>	<b>ID_Team</b>	<b>INT</b>	<b>Not null</b>	<b>Unique</b>	<b>ID unico de equipo</b>
<b>FK</b>	<b>ID_Team</b>	<b>INT</b>	<b>Not null</b>	<b>Unique</b>	<b>ID unico de equipo</b>
	Season	INT			Temporada de juego
	Standing_Date	DATE			Fecha de clasificaciones
	Conference	VARCHAR			Conferencia de equipo
	Team	VARCHAR			Nombre de equipo
	G	INT			Juegos
	W	INT			Juegos ganados
	L	INT			Juegos perdidos
	W_PCT	DECIMAL			Porcentaje de victoria
	Home_W	INT			Cantidad de partidos ganados del equipo local en la temporada
	Home_L	INT			Cantidad de partidos perdidos del equipo local en la temporada
	Away_W	INT			Cantidad de partidos ganados del equipo visitante en la temporada
	Away_L	INT			Cantidad de partidos perdidos del equipo visitante en la temporada

Tabla Salaries					
Tipo de clave	Campo	Tipo de dato	Restriccion	Unique	Descripcion
<b>PK</b>	<b>ID_Player</b>	<b>INT</b>	<b>Not null</b>	<b>Unique</b>	<b>ID unico de cada jugador</b>
<b>FK</b>	<b>ID_Player</b>	<b>INT</b>	<b>Not null</b>	<b>Unique</b>	<b>ID unico de cada jugador</b>
	Player_Name	VARCHAR			Nombre de cada jugador
	Season	INT			Temporada de juego
	Salary	DECIMAL			Salario
	Inflation_Adj_Salary	DECIMAL			Inflacion adjunta al salario

Tabla Coaches					
Tipo de clave	Campo	Tipo de dato	Restriccion	Unique	Descripcion
<b>PK</b>	<b>ID_Coach</b>	<b>INT</b>	<b>Not null</b>	<b>Unique</b>	<b>ID unico de cada entrenador</b>
<b>FK</b>	<b>ID_Coach</b>	<b>INT</b>	<b>Not null</b>	<b>Unique</b>	<b>ID unico de cada entrenador</b>
	Coach_Name	VARCHAR			Nombre de cada entrenador
	ID_Age	INT	Not null	Unique	ID unico de edad
<b>FK</b>	<b>ID_Team</b>	<b>INT</b>	<b>Not null</b>	<b>Unique</b>	<b>ID unico de cada equipo</b>
	Team	VARCHAR			Nombre de equipo
	G_Reg	INT			Juegos en temporada regular
	W_Reg	INT			Juegos ganados en temporada regular
	L_Reg	INT			Juegos perdidos en temporada regular
	W_PCT_Reg	DECIMAL			Porcentaje de juegos ganados en Temp. Reg.
	Position_Finish	INT			Posicion en la que termino el equipo
	G_Playoff	INT			Juegos en temporada play off
	W_Playoff	INT			Juegos ganados en temporada play off
	L_Playoff	INT			Juegos perdidos en temporada play off
	W_PCT_Playoff	DECIMAL			Porcentaje de juegos ganados en Temp. Play off
	Season	INT			Temporada de juego

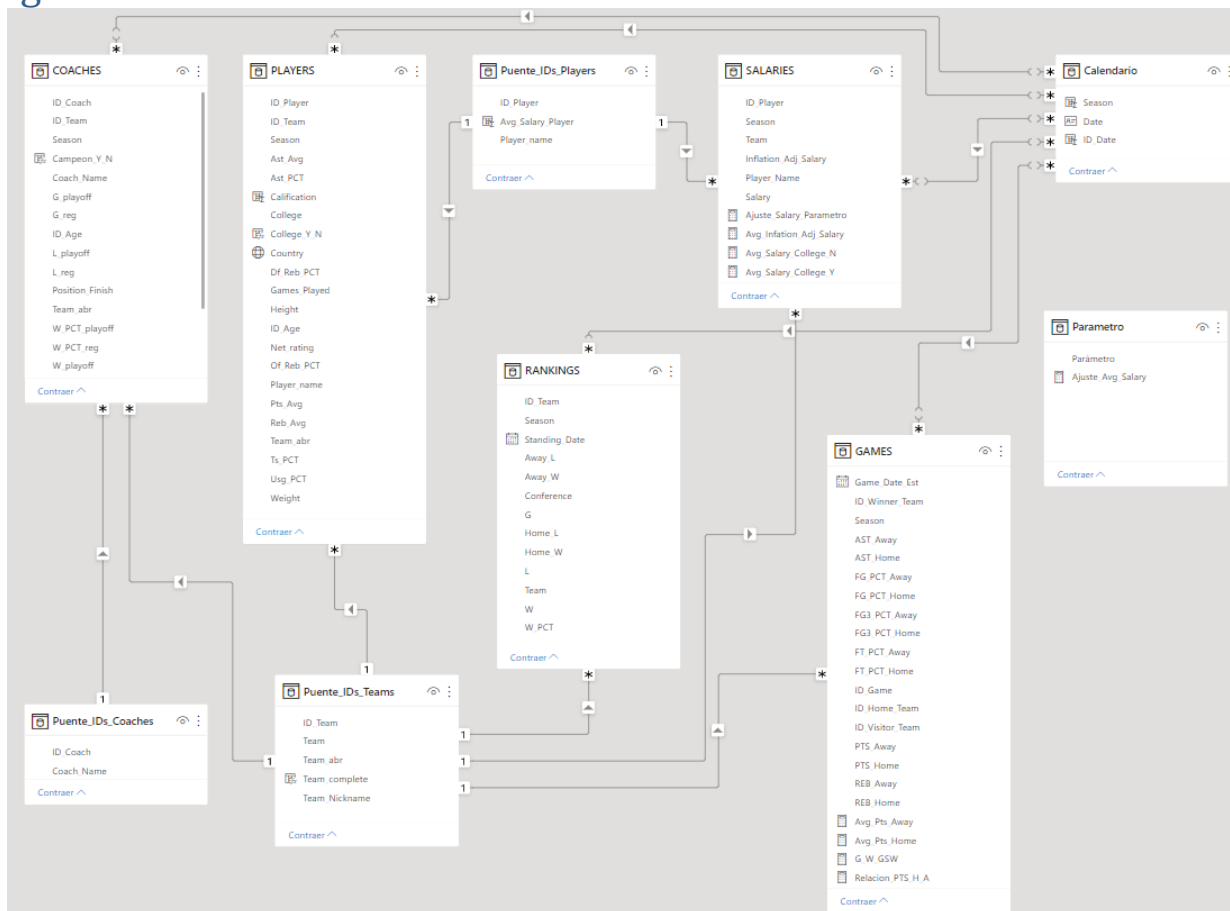
Tabla Players					
Tipo de clave	Campo	Tipo de dato	Restriccion	Unique	Descripcion
<b>PK</b>	<b>ID_Player</b>	<b>INT</b>	<b>Not null</b>	<b>Unique</b>	<b>ID unico de cada jugador</b>
<b>FK</b>	<b>ID_Player</b>	<b>INT</b>	<b>Not null</b>	<b>Unique</b>	<b>ID unico de cada jugador</b>
	Player_name	VARCHAR			Nombre de cada jugador
<b>FK</b>	<b>ID_Team</b>	<b>INT</b>	<b>Not null</b>	<b>Unique</b>	<b>ID unico de cada equipo</b>
	Team_abr	VARCHAR			Nombre de equipo abreviado
	ID_Age	INT			ID unico de edad
	Height	DECIMAL			Altura de jugador
	Weight	DECIMAL			Peso de jugador
	College	VARCHAR			Universidad en la que surgio cada jugador
	Country	VARCHAR			Nacionalidad de cada jugador
	Games_Played	INT			Partidos jugados
	Pts	DECIMAL			Puntos realizados por jugador - promedio por partido
	Reb	DECIMAL			Rebotes capturados por jugador - promedio por partido
	Ast	DECIMAL			Asistencias realizadas por jugador - promedio por partido
	Net_rating	DECIMAL			Diferencial de puntos del equipo por cada 100 posesiones mientras el jugador esta en la pista
	Of_Reb_PCT	DECIMAL			Porcentaje de rebotes ofensivos
	Df_Reb_PCT	DECIMAL			Porcentaje de rebotes defensivos
	Usg_PCT	DECIMAL			Volumen de uso de jugador en jugadas ofensivas
	Ts_PCT	DECIMAL			La cantidad de veces que el jugador toca la pelota en ataques
	Ast_PCT	DECIMAL			La cantidad de asistencias del jugador sobre el equipo
	Season	INT			Temporada de juego

Tabla Puente_IDS_Coaches					
Tipo de clave	Campo	Tipo de dato	Restriccion	Unique	Descripcion
<b>PK</b>	<b>ID_Coach</b>	<b>INT</b>	<b>Not null</b>	<b>Unique</b>	<b>ID unico de cada entrenador</b>
<b>FK</b>	<b>ID_Coach</b>	<b>INT</b>	<b>Not null</b>	<b>Unique</b>	<b>ID unico de cada entrenador</b>
	Coach_Name	VARCHAR			Nombre de cada entrenador

Tabla Puente_IDS_Players					
Tipo de clave	Campo	Tipo de dato	Restriccion	Unique	Descripcion
<b>PK</b>	<b>ID_Player</b>	<b>INT</b>	<b>Not null</b>	<b>Unique</b>	<b>ID unico de cada jugador</b>
<b>FK</b>	<b>ID_Player</b>	<b>INT</b>	<b>Not null</b>	<b>Unique</b>	<b>ID unico de cada jugador</b>
	Player_name	VARCHAR			Nombre de cada jugador

Tabla Puente_IDS_Teams					
Tipo de clave	Campo	Tipo de dato	Restriccion	Unique	Descripcion
<b>PK</b>	<b>ID_Team</b>	<b>INT</b>	<b>Not null</b>	<b>Unique</b>	<b>ID unico de cada equipo</b>
<b>FK</b>	<b>ID_Team</b>	<b>INT</b>	<b>Not null</b>	<b>Unique</b>	<b>ID unico de cada equipo</b>
	Team_abr	VARCHAR			Nombre de equipo abreviado
	Team_Nickname	VARCHAR			Apodo de cada equipo
	Team	VARCHAR			Nombre de equipo

## Diagrama entidad-relación final



## Columnas calculadas

- Tabla Coaches:
  - Campeón\_Y\_N: indica si un determinado equipo fue campeón una determinada temporada.
- Tabla Players:
  - Calification: mediante un cálculo se determinó una calificación para cada jugador en función del promedio de puntos, asistencias y rebotes capturados a lo largo de una temporada.
  - College\_Y\_N: indica si un jugador asistió o no a la universidad.
- Tabla Puentes\_IDs\_Teams:
  - Team\_complete: representa el nombre completo de los equipos, es decir, nombre de la ciudad y su respectivo nickname.
- Tabla Calendario:
  - Season: dado que las temporadas no respetan los años calendario, se indicó un rango de tiempo donde las estadísticas de las diferentes tablas corresponden a una temporada u otra.

Todas las respectivas relaciones fueron configuradas.

## Medidas calculadas

- Tabla Coaches:
  - W\_Finals: representa la cantidad de partidos que necesita ganar un equipo para acceder las finales nacionales.
  - Min\_Posible\_Playoffs: Cantidad mínima de partidos ganados en playoffss. Utilizado para dar formato a KPI.
  - Max\_Posible\_Playoffs: Cantidad máxima de partidos ganados en playoffss. Utilizado para dar formato a KPI.
  - Count\_coaches: cuenta la cantidad de Coaches que cumplen los requisitos establecidos en los Segmentadores. Por ejemplo: cuantas veces ha sido campeón en un intervalo de tiempo.
- Tabla Salaries:
  - Ajuste\_Salary\_Parametro: Indica cuánto serían los salarios hoy en día en caso de la inflación ser la indicada en el parámetro.
  - Avg\_Salary\_College\_N: promedio de los salarios de jugadores que no han asistido a la universidad.
  - Avg\_Salary\_College\_Y: promedio de los salarios de jugadores que si han asistido a la universidad.
- Tabla Games:
  - Avg\_Pts\_Away: Promedio de puntos realizados en condición de visitante.
  - Avg\_Pts\_Home: Promedio de puntos realizados en condición de local.
  - G\_W\_GSW: cantidad de partidos ganados por un equipo específico.
  - Relación\_PTS\_H\_A: Establece un ratio entre los puntos anotados de local y de visitante.

## Gráficos y Segmentadores

- La página de inicio cuenta con botones de navegación a las diferentes páginas, como a la página oficial de Coderhouse (logo de Coderhouse) y a la base de datos (logo NBA).
- Cada página cuenta con un acceso directo a la base de datos desde el logo de la NBA en el extremo superior izquierdo.
- Todas las hojas cuentan con botones de navegación hacia las otras páginas destacando en qué página nos encontramos actualmente.
- Página Players:
  - Segmentadores de temporada, nombre de jugador y valor del parámetro.
  - Gráfico de columnas agrupadas: se muestra el sueldo de los 15 jugadores con mayor sueldo de dicho equipo a lo largo de la década, en una temporada o en varias temporadas no continuas junto con la media de los sueldos.
  - Gráfico de dispersión: comparando el promedio de calificación obtenida, de los 100 jugadores con ingreso promedio más alto, con el promedio del sueldo del mismo jugador. De este modo, se da igual ponderación a los jugadores sin importar la cantidad de temporadas que han jugado a lo largo de la década. A su vez, se establece una línea de tendencia por lo que podemos interpretar que los jugadores por debajo de la misma justifican su sueldo con algo más que la calificación obtenida por su aporte al equipo. Ejemplo el marketing que le



- genera ese jugador a su equipo.
- Mapa: distribución de los jugadores con burbujas que deben su tamaño a la cantidad de jugadores nacidos en ese país y que han jugado en la liga a lo largo de la década.
- KPI que compara cual sería el presupuesto total de sueldos de los jugadores según cambie la inflación en la década (parámetro) estableciendo como objetivo la sumatoria total de los salarios ajustados por inflación de la base de datos original.
- Tarjetas con datos que profundizan las características de los jugadores.
- **Página Coaches:**
  - Segmentador de temporada.
  - Tarjetas con datos que profundizan las características de los jugadores.
  - Treemap: se representa la cantidad total de partidos ganados en Playoffs por los 6 entrenadores más exitosos de la década, discriminando si gana con 1 o más equipos. En caso de filtrar por temporada se muestran únicamente los entrenadores que hayan ganado más de 8 partidos en playoffs ya que eso representa que han llegado a las finales de conferencia (equivalente a la semifinal nacional).
  - Gráfico de áreas: comparación entre el porcentaje de victorias en la temporada regular y los playoffss. A su vez, se establece una línea de media para cada una de las medidas establecidas en este gráfico.
  - Gráfico de barras apiladas: se muestra el sueldo de los 12 equipos con mayor sueldo de dicho equipo a lo largo de la década, en una temporada o en varias temporadas no continuas.
- **Página Teams:**
  - Segmentadores de equipo y de temporada.
  - Tarjetas con datos que profundizan las características de los jugadores.
  - Gráfico de columnas apiladas y líneas comparando la suma de calificación obtenida por cada jugador a lo largo de la década (con posibilidad de filtrar por año debido al segmentador) con la suma de sus sueldos cobrados. Al ser la sumatoria de la década o de los años seleccionados, se da mayor ponderación a los jugadores que han jugado más años en la NBA.
  - KPI indicando cuanto es el máximo de partidos en playoffs que ha ganado un determinado equipo a lo largo de la década, en una temporada o en varias temporadas no continuas. Se establece como valor objetivo la medida W\_Finals que indica el mínimo de partidos que debe ganar un equipo para acceder a las finales nacionales. En caso de ser campeones alguna vez en la década o en las temporadas seleccionadas, se pinta de verde claro. Si dicho equipo llego a la final nacional pero perdió, se pinta de verde oscuro. Si su mejor resultado fue llegar a las finales de conferencia, se pinta de amarillo. Caso contrario, se pinta de rojo oscuro.
  - Gráfico de torta: se visualiza como se reparte la masa salarial del equipo entre los jugadores con los 5 sueldos más importantes de dicho año.
  - Mapa coroplético: nos indica y ubica la ciudad correspondiente al equipo por el cual se filtre.