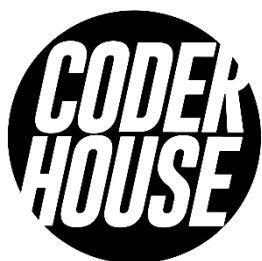




PROYECTO DATA ANALYTICS



Autor: Joaquín Perusia

Institución: Coderhouse

Fecha entrega: 11/2/2025



Índice

Glosario palabras clave	4
Introducción	5
Temática del proyecto	5
Objetivo	5
Hipótesis	6
Alcance	6
Nivel de aplicación del análisis	7
Diagramas	7
Diagrama Entidad Relación	8
Diagrama de relaciones	8
Listado de tablas	9
Tabla Equipos	9
Tabla Estadios	9
Tabla Partidos	9
Tabla Estadísticas	10

Glosario palabras clave

Asistencias → Pase que conduce directamente a una anotación.

Conferencia → Sistema geográfico en que se divide la NBA. Dos conferencias, Este y Oeste, formadas por 15 equipos cada una.

Foreign key (FK) → clave foránea. Columna o conjunto de columnas que se usa para vincular dos tablas en una base de datos.

Localía → Ventaja de jugar en el propio estadio.

Modalidad burbuja → Medida especial durante la temporada 2019 por la aparición del COVID-19, en la que todos los equipos jugaban en un único estadio sin público.

Modelo predictivo → Algoritmo o método utilizado para hacer predicciones basadas en datos históricos.

Porcentaje de tiros de campo → Relación entre tiros anotados y tiros intentados.

Porcentaje de tiros libres → Relación entre tiros libres convertidos y tiros intentados desde la línea de tiros libres.

Porcentaje de triples → Relación entre triples anotados y triples intentados.

Primary key (PK) → clave primaria de las tablas de base de datos.

Rebotes → Captura del balón después de un intento fallido de tiro.

Introducción

Cuando alguien quiere expresar comodidad y seguridad, suele decir: “*me siento como en casa*”. En el deporte, esta sensación puede ser un factor determinante en el rendimiento de los equipos. Jugar en casa no solo significa competir en un entorno conocido, sino también contar con el apoyo del público, evitar largos viajes y estar habituado a las condiciones del estadio. Pero, ¿realmente esta ventaja se traduce en un mayor porcentaje de victorias?

El análisis de datos en el ámbito deportivo permite obtener información valiosa sobre el rendimiento de los equipos y los factores que influyen en los resultados de los partidos. En este proyecto, se busca predecir la probabilidad de que un equipo local gane un partido en función de sus estadísticas. Para ello, se analizarán métricas clave del rendimiento. Además, se estudiará si la ventaja de jugar en casa tiene un impacto significativo en los resultados.

Temática del proyecto

El análisis del rendimiento en los deportes ha evolucionado con el uso de la analítica de datos, proporcionando información clave sobre el desempeño de los equipos y factores que influyen en los resultados. En este proyecto, se estudiará la influencia de la localía en los partidos de la NBA, explorando si los equipos que juegan en casa tienen una ventaja estadística sobre los visitantes.

Objetivo

El objetivo de este análisis es evaluar el impacto de la localía en los resultados de los partidos de la NBA y desarrollar un modelo predictivo que estime la probabilidad de victoria del equipo local en función de sus estadísticas de rendimiento.

Para ello, se analizarán variables clave como puntos anotados, porcentajes de tiros de campo, triples y tiros libres, asistencias y rebotes de cada equipo. A partir de estos datos, se buscará identificar patrones y tendencias que permitan determinar si jugar en casa representa una ventaja significativa para los equipos y qué factores influyen en el desenlace de los encuentros.

Además, el análisis permitirá responder preguntas como:

- ¿Cuánto influye la localía en la probabilidad de ganar un partido?
- ¿Existen métricas específicas que tengan un mayor peso en la victoria del equipo local?
- ¿Ha cambiado la ventaja de la localía a lo largo de las temporadas?

Hipótesis

La hipótesis central de este proyecto es que los equipos que juegan en condición de local tienen una mayor probabilidad de ganar debido a diversos factores, como el apoyo del público, la familiaridad con la cancha y la reducción del desgaste físico causado por los viajes.

Para probar esta hipótesis, se aplicarán distintos métodos de análisis de datos. Se realizará una exploración descriptiva de los datos para identificar tendencias en las tasas de victorias locales a lo largo de varias temporadas.

Además, se analizará si esta ventaja ha cambiado con el tiempo, tomando como referencia las temporadas 2017, 2018, 2020, 2021 (ignorando la 2019 debido a que se jugó en una modalidad de burbuja) y si existen otros factores que puedan influir en los resultados, como el rendimiento del equipo visitante o la calidad del plantel del equipo local.

Alcance

El proyecto tiene como objetivo analizar la probabilidad de victoria del equipo local en los partidos de la NBA. Este análisis está dirigido a un público diverso que incluye a analistas deportivos, equipos técnicos, periodistas y fanáticos del deporte.

Para los analistas deportivos y estadísticos, el proyecto busca proporcionar información valiosa sobre las tendencias de victorias locales, ayudando a identificar patrones de rendimiento y mejorar pronósticos en base a datos históricos. Además, los equipos técnicos y entrenadores podrán utilizar estos datos para ajustar tácticas y estrategias,

aprovechando las métricas de desempeño para tomar decisiones más informadas tanto en partidos locales como de visitante.

Por otro lado, el análisis también apunta a periodistas y medios deportivos, quienes podrán enriquecer sus coberturas con información fundamentada sobre el impacto de la localía en los resultados de los partidos. Finalmente, se espera que el proyecto despierte el interés de los aficionados al deporte, permitiéndoles comprender mejor los factores que influyen en las victorias de sus equipos favoritos.

Nivel de aplicación del análisis

Este proyecto se sitúa en un nivel táctico, ya que proporciona información clave para la toma de decisiones a corto y mediano plazo. El análisis de las estadísticas de rendimiento de los equipos permite ajustar alineaciones y estrategias de juego según las tendencias observadas en los partidos locales y de visitante.

El análisis no solo ofrece una comprensión más profunda del juego, sino que también se convierte en una herramienta para entrenadores y analistas, pudiendo planificar partidos de manera más eficiente.

Diagramas

Se presentan los diagramas del proyecto en los que se establecen las relaciones correspondientes para el entendimiento del objeto de estudio.

Los diagramas se podrán visualizar de mejor manera en:



DER-NBA.drawio

Diagrama Entidad Relación

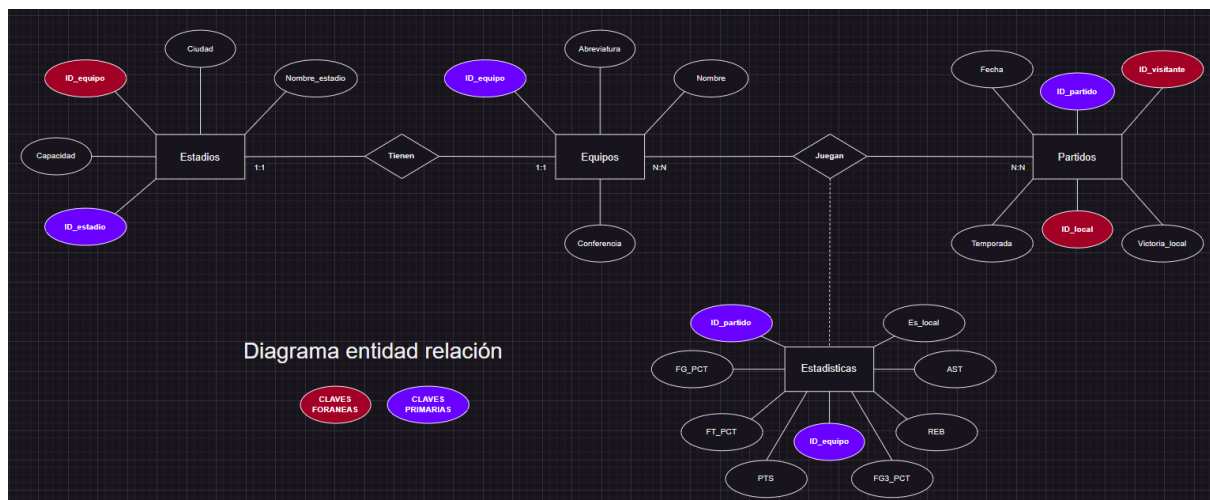
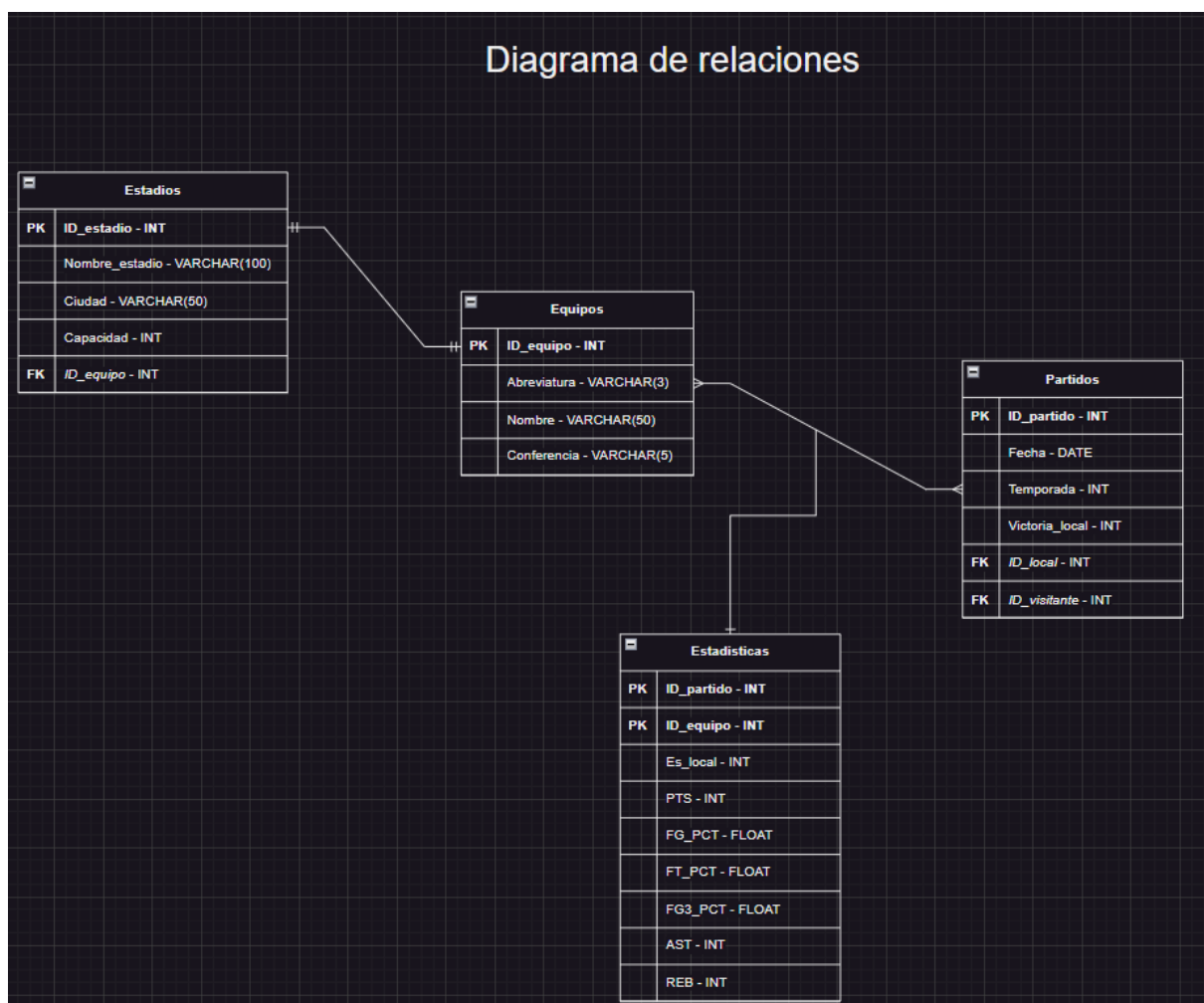


Diagrama de relaciones



Listado de tablas

En las siguientes páginas presentamos las tablas del proyecto final, con sus Columnas, Descripción, Tipo de Dato, Primary key y Foreign key.

Tabla Equipos

Columna	Descripción	Tipo de dato	Es PK	Es FK
ID_equipo	Identificador numérico de cada equipo	INT	SI	NO
Abreviatura	Abreviación (3 letras) del nombre del equipo	VARCHAR(3)	NO	NO
Nombre	Nombre del equipo	VARCHAR(50)	NO	NO
Conferencia	En que división (Este u Oeste) se encuentra el equipo	VARCHAR (5)	NO	NO

Esta tabla almacena información sobre los equipos de la NBA, incluyendo su nombre, abreviatura y conferencia (Este u Oeste). Es fundamental para identificar a los equipos en los partidos y relacionarlos con sus estadísticas y estadios. La clave primaria (PK) es el identificador numérico de cada equipo.

Tabla Estadios

Columna	Descripción	Tipo de dato	Es PK	Es FK
ID_estadio	Identificador numérico del estadio del equipo	INT	SI	NO
Capacidad	Cantidad de personas que entran en el estadio	INT	NO	NO
Ciudad	Ciudad en la que se encuentra el estadio	VARCHAR(50)	NO	NO
Nombre	Nombre del estadio	VARCHAR(100)	NO	NO
ID_equipo	Identificador numérico del equipo	INT	NO	SI

Incluye información sobre los estadios de los equipos, como nombre, ciudad y capacidad. Se utiliza para analizar la influencia de jugar en casa y su relación con las victorias locales. La clave primaria es el identificador numérico del estadio. La clave foránea es el identificador numérico de los equipos, para así poder establecer la conexión apropiada.

Tabla Partidos

Columna	Descripción	Tipo de dato	Es PK	Es FK
ID_partido	Identificador numérico del partido	INT	SI	NO
Fecha	Fecha en que se disputó el partido	DATE	NO	NO
Temporada	Temporada en que se disputó el partido	INT	NO	NO
Victoria_local	Columna que indica si el equipo local ganó el partido	INT	NO	NO
ID_local	Identificador numérico del equipo local	INT	NO	SI
ID_visitante	Identificador numérico del equipo visitante	INT	NO	SI

Contiene datos de cada partido jugado, como la fecha, temporada, equipos participantes (local y visitante) y el resultado del equipo local. Es esencial para analizar el

rendimiento de los equipos a lo largo del tiempo. La clave primaria es el identificador numérico de cada partido. Las claves foráneas son los identificadores numéricos de los equipos que disputaron ese partido, uno para el local y uno para el visitante.

Tabla Estadísticas

Columna	Descripción	Tipo de dato	Es PK	Es FK
ID_partido	Identificador numérico del partido	INT	SI	NO
ID_equipo	Identificador numérico del equipo	INT	SI	NO
Es_local	Columna que indica si el equipo era local o visitante	INT	NO	NO
PTS	Puntos anotados por el equipo en un determinado partido	INT	NO	NO
FG_PCT	Porcentaje de tiros de campo del equipo en un determinado partido	FLOAT	NO	NO
FT_PCT	Porcentaje de tiros libres del equipo en un determinado partido	FLOAT	NO	NO
FG3_PCT	Porcentaje de tiros de 3 puntos del equipo en un determinado partido	FLOAT	NO	NO
AST	Asistencia totales del equipo en un determinado partido	INT	NO	NO
REB	Rebotes totales del equipo en un determinado partido	INT	NO	NO

Almacena las estadísticas individuales de cada equipo en cada partido, como puntos, asistencias, rebotes y porcentajes de tiro. Permite evaluar el desempeño de los equipos y comparar su rendimiento en partidos locales y de visitante. La clave primaria es una compuesta, formada por el identificador numérico del partido y el del equipo.