



## SEGUNDA ENTREGA - GAMEVERSE

Flores Villarreal Samuel Fernando, Gómez Hernández Ana Camila, González Pinzón Nicolás David, Mendoza Gómez María Lucía

*Escuela de Ingeniería Ciencia y Tecnología, Universidad del Rosario,  
Bogotá-Colombia*

---

### Introducción

Continuando con la primera entrega de GAMEVERSE, para poder lograr el modelo relacional normalizado en tercera forma normal, se agregaron dos nuevas entidades y atributos:

- Región:
  - Code
  - Name
- Season:
  - Id
  - Name
  -

Y las relaciones:

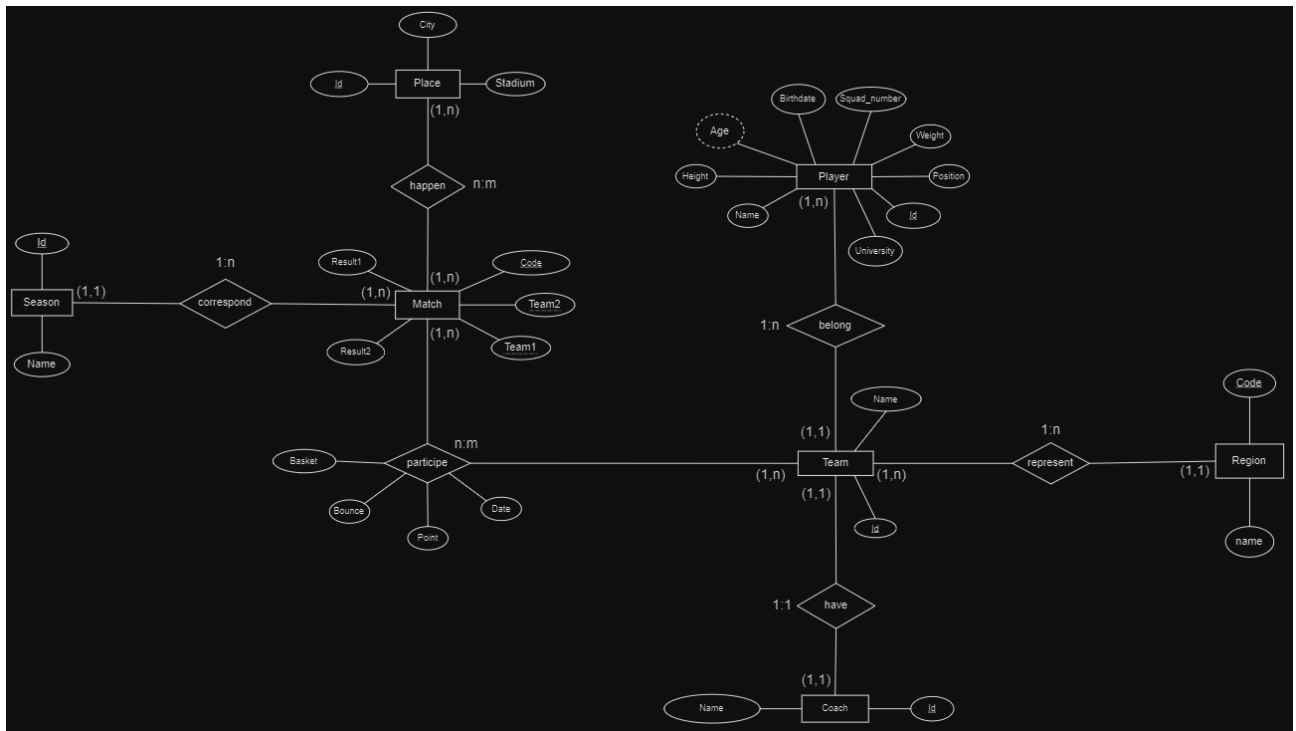
- correspond
- depend

Debido a lo anterior surgieron nuevas reglas del negocio:

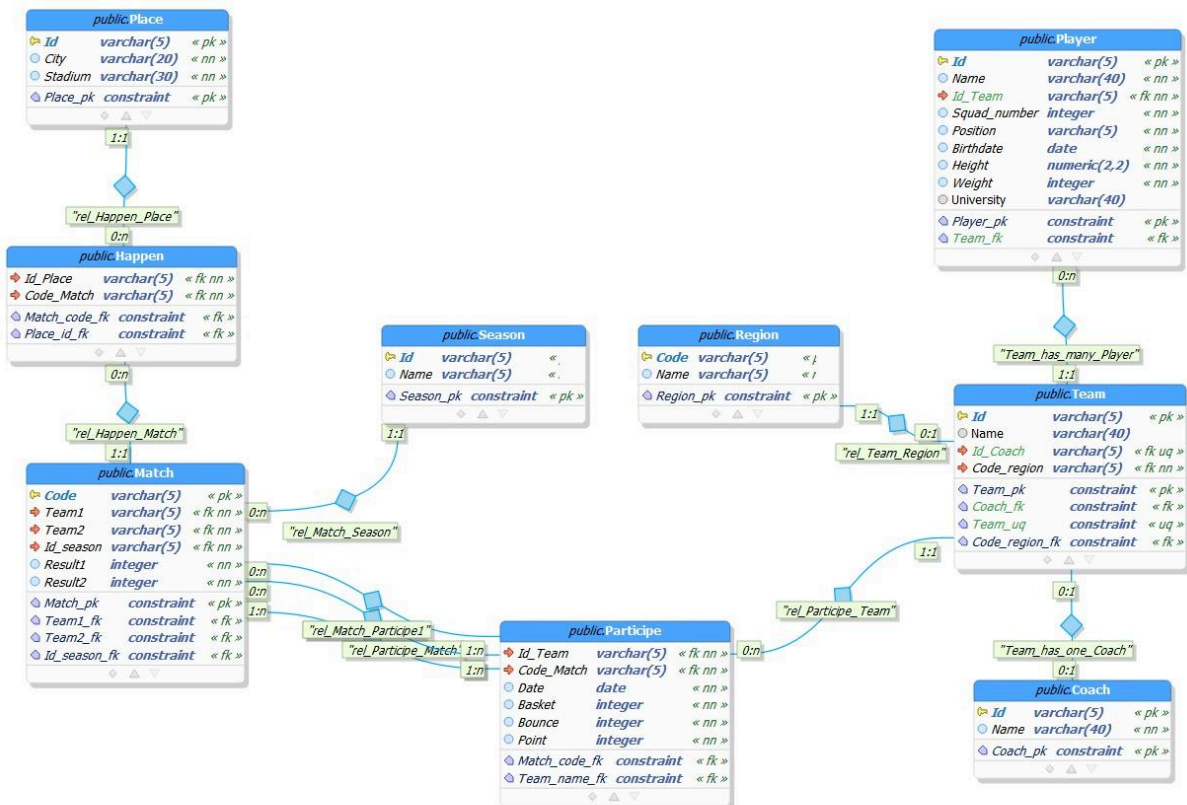
1. Una región(Region) representa(represent) unos equipos(Team)
2. Los equipos(Team) representan(represent) una región(Region)
3. Región(Region) tiene codigo(Code) que es una llave primaria y nombre(Name)
4. Equipo(Team) tiene como llave foránea codigo\_region(Code\_region)
5. Una temporada(Season) le corresponde(correspond) un partido(Match)
6. A los partidos(Match) le corresponde(correspond) una temporada(Season)
7. Temporada(Season) tiene id(Id) que es una llave primaria y nombre(Name)
8. Partido(Match) tiene como llave foránea id\_temporada(Id\_season)

Por lo cual al utilizar la base de datos original en formato CSV y emplearla en excel para convertirlas en entidades con sus respectivos atributos y relaciones se extrajeron como resultados:

### Modelo de entidad relacional



## Diagrama relacional normalizado en tercera forma normal



## Carga de información en la base de datos

Al realizar la carga masiva de datos, utilizamos la aplicación Excel, la cual proporciona herramientas avanzadas para la extracción de datos de diversas fuentes. Estas funciones están disponibles bajo la pestaña "Datos" en la barra de herramientas.

Dentro de esta sección, encontramos un menú desplegable llamado "Obtener datos", que ofrece distintos métodos para importar datos desde una variedad de archivos y fuentes, como páginas web.

Cuando seleccionamos la opción "De otras formas", se despliegan diversas alternativas para importar datos, incluyendo la posibilidad de extraer información desde archivos PDF o desde páginas web.

Para cargar datos desde una página web, copiamos la URL correspondiente y la pegamos en el campo designado bajo "Dirección URL". Una vez establecida la conexión con el servidor que aloja la página web, podemos explorar las distintas formas de visualizar y organizar las tablas de datos disponibles en ella.

Al momento de cargar la tabla, tenemos la opción de editarla utilizando la función "Consultar", seguida por la opción "Editar". Esto nos permite modificar las columnas y los datos utilizando las funciones preestablecidas o aplicando mejoras personalizadas según nuestras necesidades.

En ocasiones, este método de carga masiva puede enfrentar desafíos que requieren características específicas de la página web. Por ejemplo, al intentar extraer datos de partidos, nos encontramos con dificultades y no logramos obtener la información deseada.

Esto puede ocurrir debido a varios factores, como el formato complejo de la página web, la estructura de las tablas de datos o la manera en que se presentan los detalles de los partidos.

En tales situaciones, es importante tener en cuenta que la extracción de datos de manera automatizada puede no ser siempre directa. A menudo, se requiere un análisis más detallado de la estructura de la página web y, en algunos casos, la personalización de los métodos de extracción para adaptarse a las particularidades del sitio.

Adicionalmente cada una de las tablas resultantes del diagrama relacional normalizado en tercera forma normal fueron organizados en diferentes archivos en formato CSV UTF-8, con el fin de utilizarlos en pgAdmin4, junto a la función 'copy' para copiar los datos de los archivos y pegarlos en la base datos, realizando una instrucción para cada tabla, de las cuales al ejecutarlas permite cargar la información y obtener la base de datos con sus respectivos datos en SQL.

## **Análisis de escenarios**

Con base en los datos de la NBA y los datos que se extrajeron, se proponen los siguientes escenarios posibles para GAMEVERSE:

- Predecir el ganador de la NBA
  - Metodología:
    - Unir la entidad 'Season' con 'Match' utilizando el atributo 'Id' en 'Season' y 'Id\_season' en 'Match'. Lo que permitirá obtener todos los partidos de una temporada específica. Luego, unir la tabla resultante con la tabla 'Participe' utilizando el atributo 'Code' en 'Match' y 'Code\_Match' en 'Participe'. Lo que permitirá obtener información sobre qué equipos participaron en cada partido y los resultados de esos partidos.

- Utilidad:
  - Utilizando las entidades, atributos y relaciones creadas anteriormente, junto con la posible consulta propuesta, se busca analizar los resultados de 'Match', 'Team' y 'Season'; preferiblemente la posttemporada en donde se reducen a pocos equipos el ganador. Calculando el mayor número de 'Basket' y 'Point' por equipo. De esta manera lograr hacer una predicción cercana al posible equipo ganador de la NBA gracias a los datos obtenidos.
- El equipo(Team) con mayor número de rebotes(Bounce)
  - Metodología:
    - Uniendo la entidad 'Team' con 'Participe' mediante el atributo 'Id' en 'Team' y 'Id\_team' en 'Participe'. Luego, se relaciona con 'Match' utilizando el atributo 'Code\_Match' en 'Participe' y 'Code' en 'Match'. Esto permite obtener información sobre qué equipos participaron en cada partido y los resultados de esos partidos. Además, se incluye la tabla 'Season' para filtrar por temporada. Al sumar los 'Bounce' por equipo, se calcula el total de 'Bounce' realizados por cada uno.
  - Utilidad:
    - Utilizando las entidades, atributos y relaciones previamente establecidas, junto con la consulta propuesta, se busca identificar el equipo con el mayor número de 'Bounce'. Al ordenar los resultados de manera descendente, se puede identificar el equipo que se destaca en esta área del juego y lograr de manera asertiva nuestra propuesta de escenario.
- El equipo(Team) que mayor puntos(Point) logró
  - Metodología:
    - Se une la entidad 'Team' con 'Participe' mediante el atributo 'Id' en 'Team' y 'Id\_team' en 'Participe'. Posteriormente, se relaciona con 'Match' utilizando el atributo 'Code\_Match' en 'Participe' y 'Code' en 'Match'. Esto permite obtener información sobre qué equipos participaron en cada partido y los resultados de esos partidos. Además, se incluye la tabla 'Season' para filtrar por temporada. Se suman los 'Point' por equipo para calcular el total de 'Point' logrados por cada equipo.
  - Utilidad:
    - Utilizando las entidades, atributos y relaciones previamente establecidas, junto con la consulta propuesta, se busca identificar el equipo que logró el mayor número de puntos. Este indicador puede ser fundamental para evaluar el posible ganador de la NBA ya que esto corresponde a la efectividad en ataque de un equipo durante una temporada específica. Al calcular el total de puntos(Point) por equipo y ordenar los resultados de manera descendente, se puede identificar el equipo que mayor puntos(Point) logró.
- En qué estadio(Stadium) se repiten más partidos(Match)
  - Metodología:
    - Se unen las entidades 'Place' y 'Match' utilizando el atributo 'Id' en 'Place' y 'Id\_place' en 'Match', respectivamente. Esto permite obtener información sobre qué partidos ocurrieron en cada estadio. Se cuentan los

partidos agrupados por estadio para determinar cuál estadio ha albergado la mayor cantidad de partidos.

- Utilidad:
  - Utilizando las entidades, atributos y relaciones previamente establecidas, junto con la consulta propuesta, se busca identificar el estadio que ha sido sede de la mayor cantidad de partidos. Este análisis puede ser útil para entender la popularidad y la capacidad de los estadios para albergar eventos deportivos. Al contar los partidos por 'stadium' y ordenar los resultados de manera descendente, se puede identificar el 'stadium' que se ha utilizado con mayor frecuencia para celebrar partidos.
- Los jugadores(Player) más altos de cada equipo(Team)
  - Metodología:
    - Se unen las entidades 'Team' y 'Player' utilizando el atributo 'Id' en 'Team' y 'Id\_team' en 'Player', respectivamente. Esto permite obtener información sobre los jugadores de cada equipo. Se utiliza la función de agregación 'MAX' para encontrar la altura máxima de cada equipo.
  - Utilidad:
    - Utilizando las entidades, atributos y relaciones previamente establecidas, junto con la consulta propuesta, se busca identificar a los jugadores más altos de cada equipo. Este análisis puede ser útil para entender la composición física de los equipos y su posible estrategia de juego. Al encontrar la altura máxima de cada equipo, se puede identificar a los jugadores que sobresalen en este aspecto.
- Los jugadores(Player) más pesados de cada equipo(Team)
  - Metodología:
    - Se unen las entidades 'Team' y 'Player' utilizando el atributo 'Id' en 'Team' y 'Id\_team' en 'Player'. Esto permite obtener información sobre los jugadores de cada equipo. Se utiliza la función de agregación 'MAX' para encontrar el peso máximo de cada equipo.
  - Utilidad:
    - Utilizando las entidades, atributos y relaciones previamente establecidas, junto con la consulta propuesta, se busca identificar a los jugadores más pesados de cada equipo. Este análisis puede ser útil para entender la composición física de los equipos y su posible fortaleza en áreas como la defensa y el rebote. Al encontrar el peso máximo de cada equipo, se puede identificar a los jugadores que tienen un mayor impacto debido a su peso.
- El jugador(Player) más joven de la NBA
  - Metodología:
    - Se selecciona la entidad 'Player' y se utiliza la función de agregación 'MIN' para encontrar al jugador más joven de la NBA. Se extraen los datos de la tabla 'Player' y se ordenan por la fecha de nacimiento en orden ascendente.
  - Utilidad:
    - Utilizando las entidades, atributos y relaciones previamente establecidas, junto con la consulta propuesta, se busca identificar al jugador más joven de la NBA. Este análisis puede ser útil para destacar a los jugadores jóvenes con un gran potencial en la liga. Al encontrar el jugador con la

fecha de nacimiento más reciente, se puede identificar al jugador más joven de la NBA.

- El jugador(Player) con mayor edad de la NBA
  - Metodología:
    - Se selecciona la entidad 'Player' y se utiliza la función de agregación 'MAX' para encontrar al jugador con la mayor edad en la NBA. Se extraen los datos de la tabla 'Player' y se ordenan por la fecha de nacimiento en orden descendente.
  - Utilidad:
    - Utilizando las entidades, atributos y relaciones previamente establecidas, junto con la consulta propuesta, se busca identificar al jugador más veterano de la NBA. Este análisis puede ser útil para destacar a los jugadores con una larga trayectoria en la liga. Al encontrar el jugador con la fecha de nacimiento más antigua, se puede identificar al jugador más veterano de la NBA.
- La región(Region) con el mayor número de equipos(Team) ganadores.
  - Metodología:
    - Se selecciona la entidad 'Region' y se une con la entidad 'Team' utilizando el atributo 'Code' en 'Region' y 'Code\_region' en 'Team'. Luego, se une con la entidad 'Match' para obtener información sobre los equipos ganadores. Se utiliza la función de agregación 'COUNT' para contar el número de equipos ganadores por región. Finalmente, se ordenan los resultados en orden descendente para identificar la región con el mayor número de equipos ganadores.
  - Utilidad:
    - Utilizando las entidades, atributos y relaciones previamente establecidas, junto con la consulta propuesta, se busca identificar la región con el mayor número de equipos ganadores. Este análisis puede ser útil para comprender la distribución geográfica de los equipos exitosos en la liga. Al contar el número de equipos ganadores por región, se puede identificar la región que ha tenido más éxito en la NBA.
- Qué día se jugaron más partidos(Match) en la NBA
  - Metodología:
    - Se selecciona la entidad 'Match' y se une con la entidad 'Participe' mediante el atributo 'Code' en 'Match' y 'Code\_Match' en 'Participe'. Luego, se utiliza la función de agregación 'COUNT' para contar el número de partidos jugados en cada día. Se extraen los datos de la tabla 'Match' y se agrupan por la fecha del partido. Luego, se ordenan los resultados en orden descendente para identificar el día con la mayor cantidad de partidos jugados.
  - Utilidad:
    - Utilizando las entidades, atributos y relaciones previamente establecidas, junto con la consulta propuesta, se busca identificar el día con la mayor cantidad de partidos jugados en la NBA. Al contar el número de partidos por día, se puede identificar el día con la mayor actividad en la liga.

## Fuentes

- **Repositorio:** <https://github.com/camila-gomez/GAMEVERSE.git>
- **Página Principal de la NBA:** <https://www.nba.com/>
- **Equipos de la NBA:** <https://www.nba.com/teams>
- **Juegos de la NBA:** <https://www.nba.com/games>