



UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

MAESTRÍA EN CIENCIAS DE DATOS

BASES DE DATOS RELACIONALES

MAESTRO: JOSÉ ALBERTO BENAVIDES VÁZQUEZ

TAREA #5

ALUMNO: EDWIN MARTÍN ROMERO SILVA

MATRÍCULA: 1731276

1.-Agregar datos ficticios o de otras fuentes de manera automática mediante funciones como las vistas en esta tarea.

select count(\*) as conteo from master;

Resultados 1 X

select count(\*) as conteo from master Enter a SQL expression to

Grilla	123 conteo
1	18.846

select count(\*) as conteo from teams;

Resultados 1 X

select count(\*) as conteo from teams Enter a SQL expression to

Grilla	123 conteo
1	2.805

select count(\*) as conteo from pitching;

Resultados 1 X

select count(\*) as conteo from pitching Enter a SQL expression to

Grilla	123 conteo
1	44.139

select count(\*) as conteo from SeriesPost;

Resultados 4 X

select count(\*) as conteo from SeriesPost Enter a SQL expression to

Grilla	123 conteo
1	300

select count(\*) as conteo from batting;

Resultados 1 X

select count(\*) as conteo from batting Enter a SQL expression to

Grilla	123 conteo
1	101.332

select count(\*) as conteo from managers;

Resultados 1 X

select count(\*) as conteo from managers Enter a SQL expression to

Grilla	123 conteo
1	3.405

select count(\*) as conteo from fielding;

Resultados 1 X

select count(\*) as conteo from fielding Enter a SQL expression to

Grilla	123 conteo
1	170.526

select count(\*) as conteo from AwardManagers;

Resultados 1 X

select count(\*) as conteo from AwardManagers Enter a SQL expression to

Grilla	123 conteo
1	176

select count(\*) as conteo from AwardPlayers;

Resultados 1 X

select count(\*) as conteo from AwardPlayers Enter a SQL expression to

Grilla	123 conteo
1	6.078

localhost - localhost:3306

Databases

baseball\_databank

Tables

> awardmanagers	32K
> awardplayers	32K
> batting	31M
> fielding	46M
> managers	48K
> master	4,5M
> pitching	17M
> seriespost	48K
> teams	1,6M

## 2.-Reportar en menos de 5 minutos hallazgos, dificultades, recomendaciones o recursos que seal relevantes.

Para insertar valores en mis bases de datos no utilicé la sentencia ***mysql -u root -p baseball\_databank <archivo\_insert\_into.sql*** ya que no generé valores aleatorios desde la página que revisamos en clase, la cual los exporta en formato sql. Utilicé los csv's originales que encontré en la página de Kaggle.

Tuve dificultades al tratar de insertar los datos con la sentencia **LOAD DATA INFILE 'C:\Directorio seguro\archivo.csv'**, por lo que decidí importar manualmente los csv's desde la plataforma de dbeaver.

Surgieron algunos problemas al tratar de importar estos csv's:

- Al tratar de usar la sentencia **LOAD DATA INFILE**, los archivos debían estar depositados en una 'ruta segura', la cual desconocía. Pude obtener la ruta al correr la sentencia: **SHOW VARIABLES LIKE 'secure\_file\_priv';**
- Al especificar el número de caracteres de las variables tipo string, me refiero a la sentencia **columna VARCHAR(n)**, surgió un error. Algunos valores de la columna superaban el número límite de caracteres. Por lo que tuve que incrementar el tamaño de n, incluso truncar algunos de los valores del csv.
- Al tratar de importar una de las tablas con una columna referenciada a otra tabla, surgió otro error, esto debido a que algunos valores de la foreign key no se encontraban en la tabla referenciada, eran pocos registros, así que los eliminé, preferiblemente eso a perder toda la tabla.
- También me surgió un error al tratar de establecer 2 foreign keys en una tabla, lo solucioné agregando un índice en la tabla de referencia.
- También obtuve errores al tener 2 formatos distintos de fecha en una sola columna, tuve que solucionarlo desde el csv, mantuve el siguiente formato de fecha: AAAA-MM-DD.

## 3.- Autoevaluación.

Me pongo un 8.5/10, ya que logré importar todos los csv's y pude crear correctamente las relaciones con las llaves foráneas, además logré resolver los errores que surgieron, incluso el error al tratar de establecer 2 llaves foráneas, el cual me pareció complicado de entender.

No importé desde el cmd el archivo .sql descargado de la página que genera valores aleatorios (la forma difícil) ni utilicé la sentencia para insertar csv's con código, en lugar de eso utilicé una importación tradicional desde la interfaz de dbeaver, pero se cumplió el objetivo de tener las bases de datos listas para continuar trabajando.