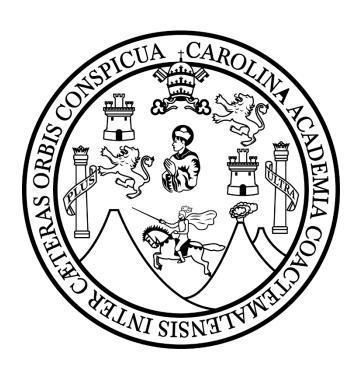
UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA
ESCUELA DE CIENCIAS Y SISTEMAS
SISTEMAS DE BASES DE DATOS 2
SECCIÓN A



Grupo 3 Proyecto 1	
201800476	Marvin Alexis Estrada Florian
201902781	Rodrigo Antonio Porón De León
201800524	Eduardo Tomás Ixén Rucuch

Cambios

Durante el proceso de extracción de datos de la API de IGDB se notaron cambios que debían realizarse en el modelo inicialmente propuesto en la *fase 1* del proyecto. Dichos cambios se muestran a continuación.

Para poder aceptar cualquier carácter unicode, se cambió el VARCHAR2 a NVARCHAR

En la tabla videojuego se realizaron los siguientes cambios:

Videojuego

- Descripcion VARCHAR2 -> NVARCHAR(5000)
- Historia VARCHAR2 -> NVARCHAR(5000)
- Calificacion general -> Double
- Calificación_profesional -> Double

En general, se eliminaron las siguientes tablas por el hecho que algunos datos de IGDB eran externos, (crítica, el tiempo para completar un juego):

- Critica, Videojuego_Critica
- Tiempo, Videojuego_Tiempo

En la tabla *Titulo*, la API no proveía el tipo, por lo que se debió eliminar la relación muchos a muchos y relacionarla con la tabla *Videojuegos*, de la misma como están relacionadas las tablas *franquicia*, *serie*, etc.

Con el anterior cambios se notó lo siguiente: como un juego puede tener varios títulos, a la tabla Titulo se le debió agregar los siguientes campos:

- Id (de cada fila de la tabla, debe ser autoincrementable)
- Videojuego Id (es la id de la tabla videojuego)
- Nombre
- Comentario

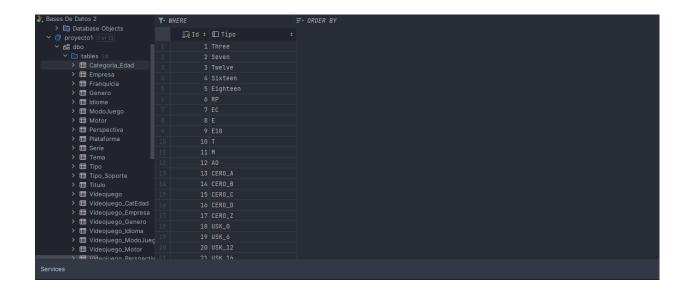
A la tabla *Empresa*, se le quitó la columna "Fundacion" y "Estado" puesto que la API no los proporciona por ningún lado, el país algunos lo traen y solo devuelve este tipo: ISO 3166-1 country code, por lo que a country se le puso el tipo de dato int.

Se le debió agregar a la columna Id de las siguientes tablas la propiedad para que sean autoincrementables:



Tablas finales

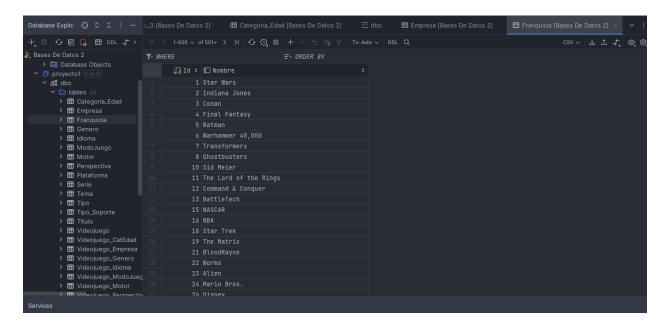
Categoria_Edad



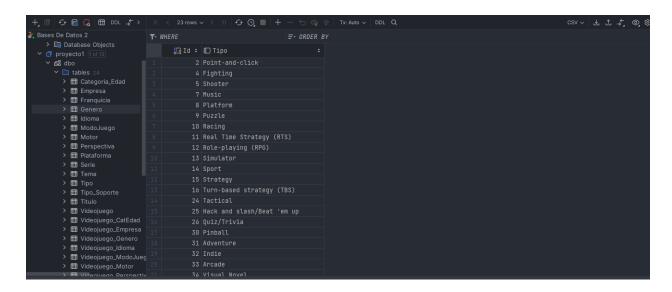
• Empresa



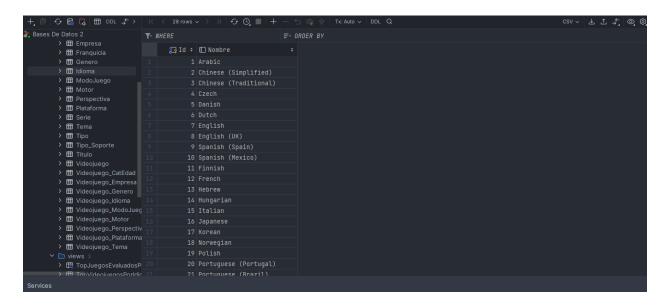
• Franquicia



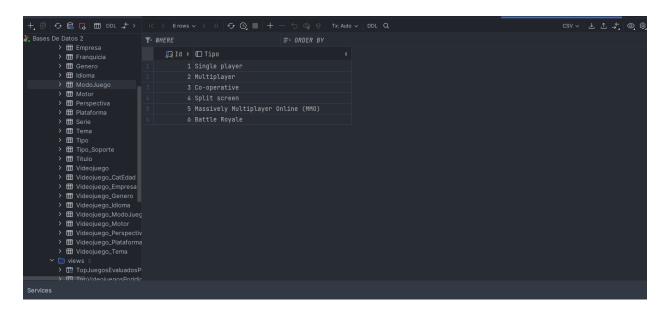
Género



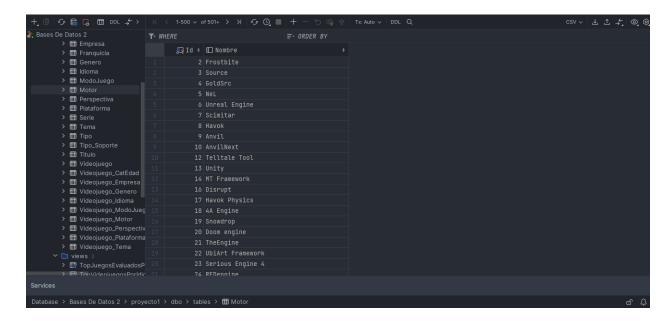
Idioma



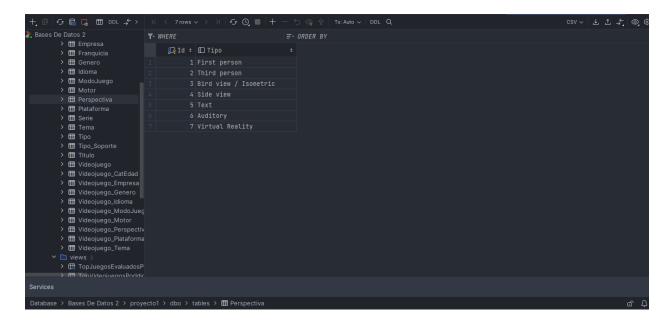
ModoJuego



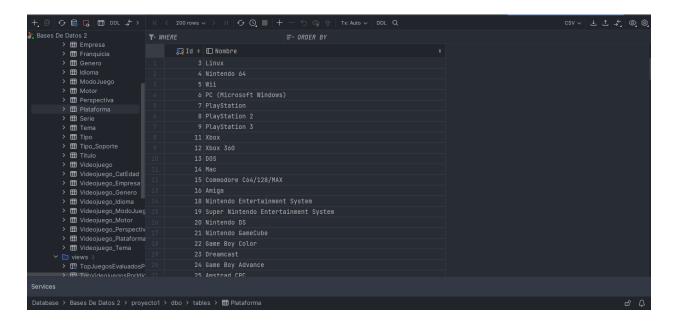
Motor



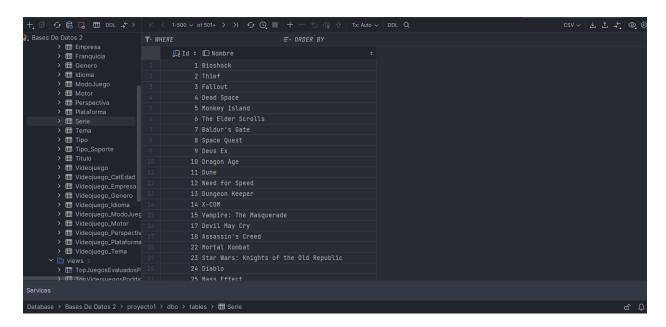
Perspectiva



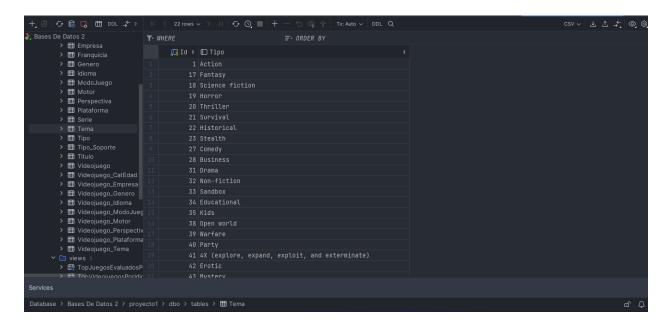
Plataforma



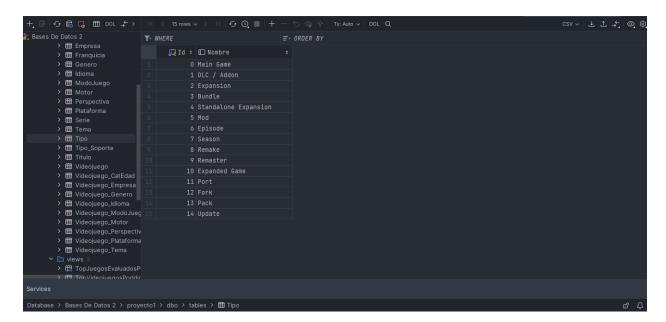
Serie



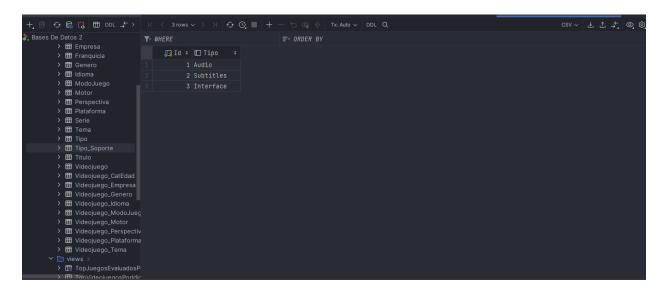
Tema



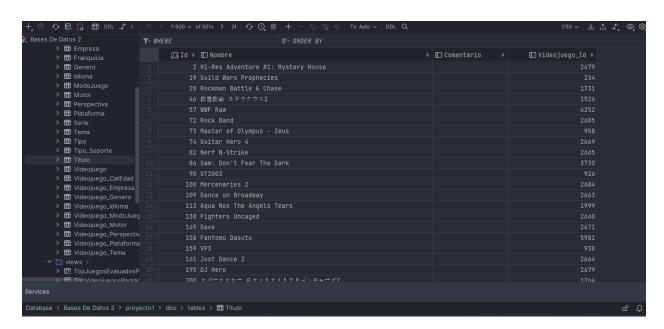
Tipo



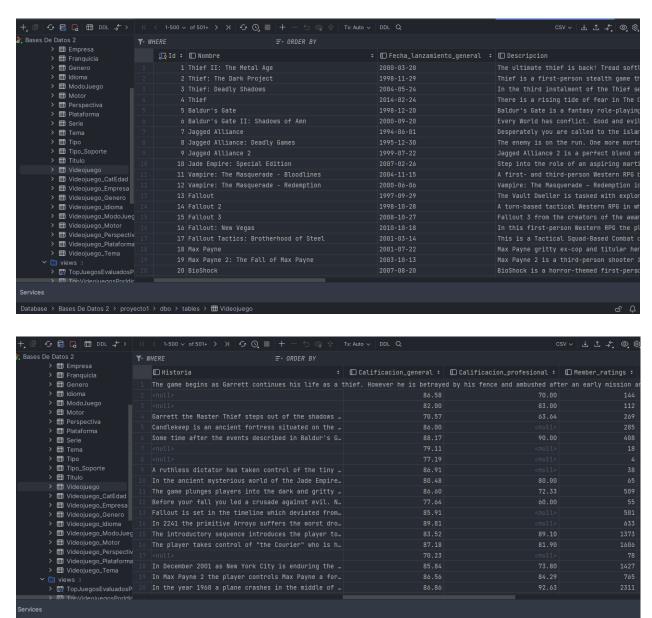
• Tipo_Soporte

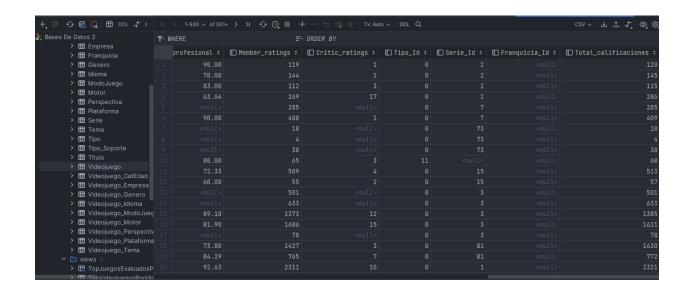


Titulo

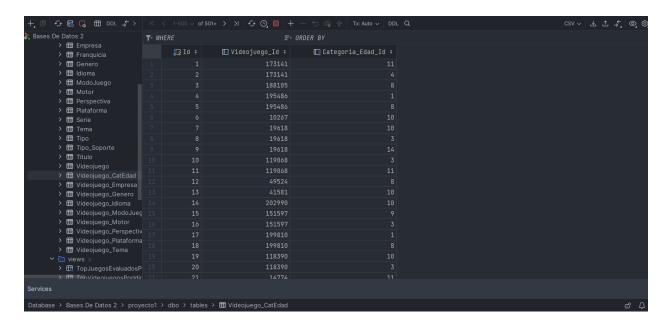


Videojuego

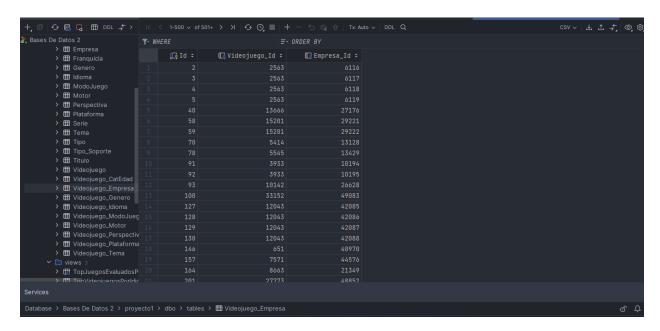




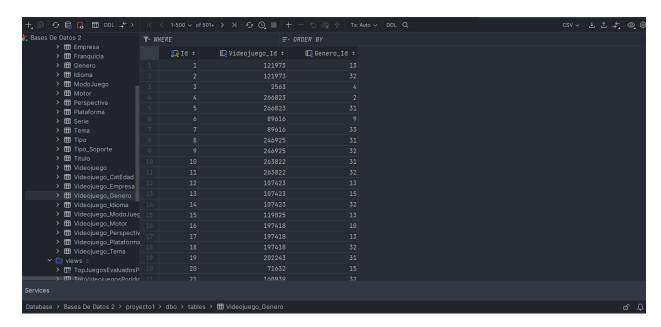
Videojuego_CatEdad



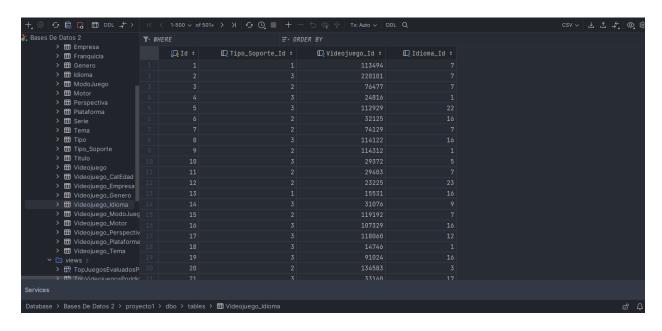
Videojuego_Empresa



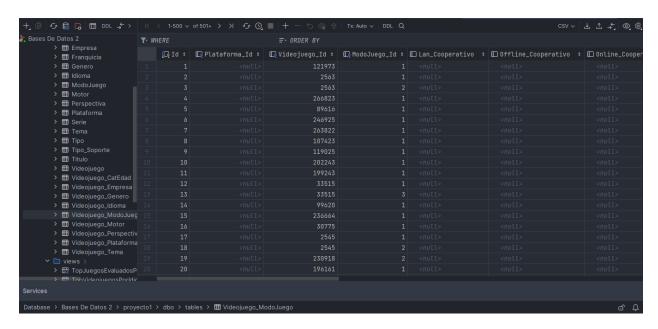
Videojuego_Genero



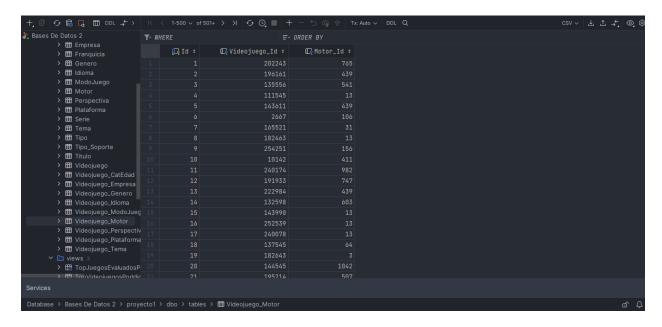
Videojuego_Idioma



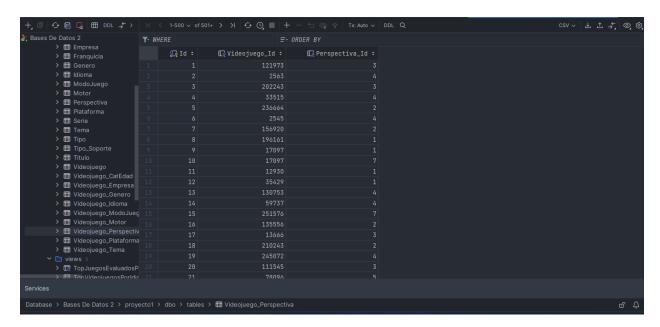
Videojuego_ModoJuego



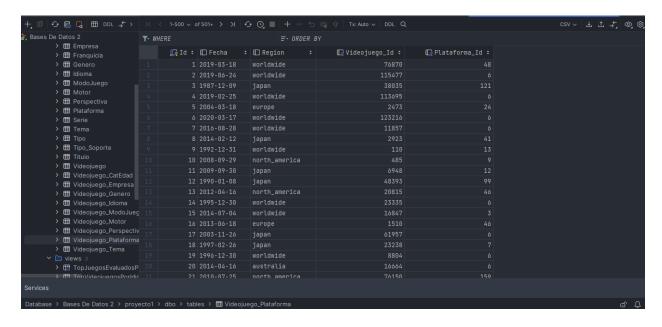
Videojuego_Motor



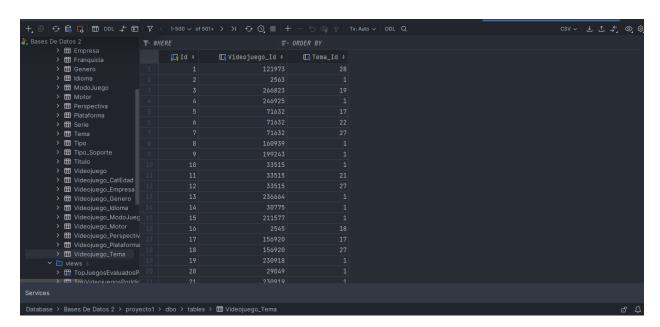
Videojuego_Perspectiva



Videojuego_Plataforma



• Videojuego_Tema



Carga de datos

Se llevó a cabo un proceso de extracción de datos desde la API de IGDB (Internet Game Database) utilizando sus diversos endpoints para recopilar información detallada sobre videojuegos. Estos datos fueron posteriormente almacenados en archivos CSV como parte de una etapa de preparación para su posterior carga en una base de datos.

```
def obtain data():
   total_data = []
   while True:
       games = obtain offset(params['offset'])
       if not games:
           break
        for game in games:
            if 'date' in game:
                game['date'] = convert unix timestamp(game['date'])
            if 'region' in game and game['region'] is not None:
                region_number = int(game['region'])
                region_name = Region(region_number).name
                game['region'] = region_name
        total data.extend(games)
        params['offset'] += params['limit']
        time.sleep(1)
        print(len(total data))
       if len(total_data) >= 326328:
           break
   return total_data
```

Durante la obtención de datos desde la API de IGDB, se realizaron solicitudes a los endpoints relevantes para adquirir información específica, como detalles de videojuegos, información sobre plataformas, calificaciones, géneros y otros atributos relacionados con los videojuegos.

Una vez que los datos fueron recuperados de la API, se procesaron y formatearon según las necesidades del proyecto y se almacenaron en archivos CSV. Estos archivos CSV se utilizarán como fuente de datos para la posterior carga en una base de datos, lo que permitirá un acceso más eficiente y organizado a la información de los videojuegos.

```
def obtain_offset(offset):
    params['offset'] = offset

response = requests.post('https://api.igdb.com/v4/release_dates', headers={'Client-ID': 'kkl3orxbx5oyw1ryu192wc4xqkejrp', 'Authorization': 'Bearer qqb6y6txmuiyy6v3tobundjj57

if response.status_code == 200:
    return response.json()
    else:
        print(f"Error: {response.status_code} - {response.text}")
        return []
```

Este proceso de extracción y almacenamiento en archivos CSV es una práctica común en la preparación de datos antes de su incorporación en una base de datos, ya que los archivos CSV son fáciles de gestionar y pueden ser importados con facilidad en una variedad de sistemas de gestión de bases de datos para su posterior análisis y consulta.

```
def data_count():
    response = requests.post('https://api.igdb.com/v4/release_dates/count', headers={'Client-ID': 'kk13orxbx5oyw1ryu192wc4xqkejrp',
    if response.status_code == 200:
        return response.json()
    else:
        print(f"Error: {response.status_code} - {response.text}")
        return []
```

Almacenar los datos primero en archivos CSV antes de cargarlos en una base de datos puede tener varias ventajas significativas en comparación con cargarlos directamente mediante consultas SQL. Aquí hay algunas razones por las cuales esta práctica puede ser preferible en ciertos casos: Los archivos son un formato de datos ampliamente aceptado y simple de manejar. Al guardar los datos en archivos primero, se tiene un control más preciso sobre el formato y la estructura de los datos antes de cargarlos en la base de datos. Puedes realizar transformaciones, limpiar los datos y aplicar cambios de formato de manera más sencilla.