



ESCUELA POLITÉCNICA NACIONAL
ESCUELA DE FORMACIÓN DE TECNÓLOGOS



ANÁLISIS DE DATOS
TSDS

ASIGNATURA: ANÁLISIS DE DATOS
PROFESOR: Ing. Lorena Chulde / Ing. Juan Pablo Zaldumbide
FECHA: 13 – 08 - 2024
PERÍODO ACADÉMICO: 2024-A

PPROYECTO FINAL – BIMESTRE 2



Integrantes:

Jonathan Ramírez
Andrés Tufiño
Evelyn Guachamin

2024-A

Link de github del proyecto

[https://github.com/Andrespipe1/Proyecto Analisis](https://github.com/Andrespipe1/Proyecto_Analisis)

<https://www.youtube.com/watch?v=X2rgjO88Kj8>

<https://www.youtube.com/watch?v=0iS8tnuOyOE>

INFORME

DEFINICIÓN DEL CASO DE ESTUDIO

Este proyecto se centra en el análisis de datos de diversas fuentes para evaluar y visualizar una amplia gama de temas, como eventos deportivos a nivel mundial, actividades recreativas, conciertos, películas y restaurantes. El objetivo principal es integrar y analizar esta información utilizando herramientas de bases de datos tanto relacionales como NoSQL, y presentar los resultados a través de visualizaciones en Power BI. En la actualidad, el análisis de datos se ha vuelto una práctica fundamental tanto en el ámbito empresarial como en el personal, ya que permite obtener información valiosa que facilita la toma de decisiones. Un análisis exhaustivo proporciona una visión completa, ayudando a elegir desde la mejor película para ver hasta la estrategia más adecuada para un negocio.

OBJETIVO GENERAL

- Combinar y analizar datos de distintas fuentes para ofrecer una perspectiva completa sobre eventos deportivos, actividades, conciertos, películas y restaurantes, utilizando Power BI para visualizar los resultados de manera sencilla y comprensible.

OBJETIVOS ESPECÍFICOS

- Reunir datos de al menos 10 fuentes diferentes que cubren una variedad de temas, como deportes, conciertos, películas y restaurantes.
- Utilizar dos tipos de bases de datos relacionales y dos bases de datos NoSQL para almacenar y gestionar la información.
- Convertir los datos entre formatos como CSV y JSON para asegurar que se puedan unir y utilizar sin problemas.
- Visualizar todos los datos combinados en Power BI para obtener información clara y útil que ayude a tomar decisiones más informadas.

Descripción del equipo de trabajo y actividades realizadas por cada uno.

Equipo de trabajo:

El equipo de trabajo consta de los siguientes estudiantes:

Jonathan Ramírez: Coordinador del Proyecto y Encargado de Bases de Datos.

Andrés Tufiño: Especialista en Análisis y Transformación de Datos.

Evelyn Guachamin: Responsable de la Visualización de Datos en Power BI.

Actividades realizadas:

- **Coordinador del Proyecto:** Organiza y supervisa las tareas del equipo, asegurando que todos estén alineados con los objetivos. También se encarga de gestionar las bases de datos, asegurando que la información esté bien almacenada.
- **Encargado del Análisis de Datos:** Recolecta y limpia los datos, además de convertirlos entre formatos como CSV y JSON para que sean compatibles con las diferentes herramientas que usamos.
- **Responsable de Visualización:** Crea los gráficos y paneles en Power BI, presentando los datos de manera visual para que sea más fácil entender la información y sacar conclusiones útiles.

Cronograma de actividades incluido el diagrama de gantt (project)

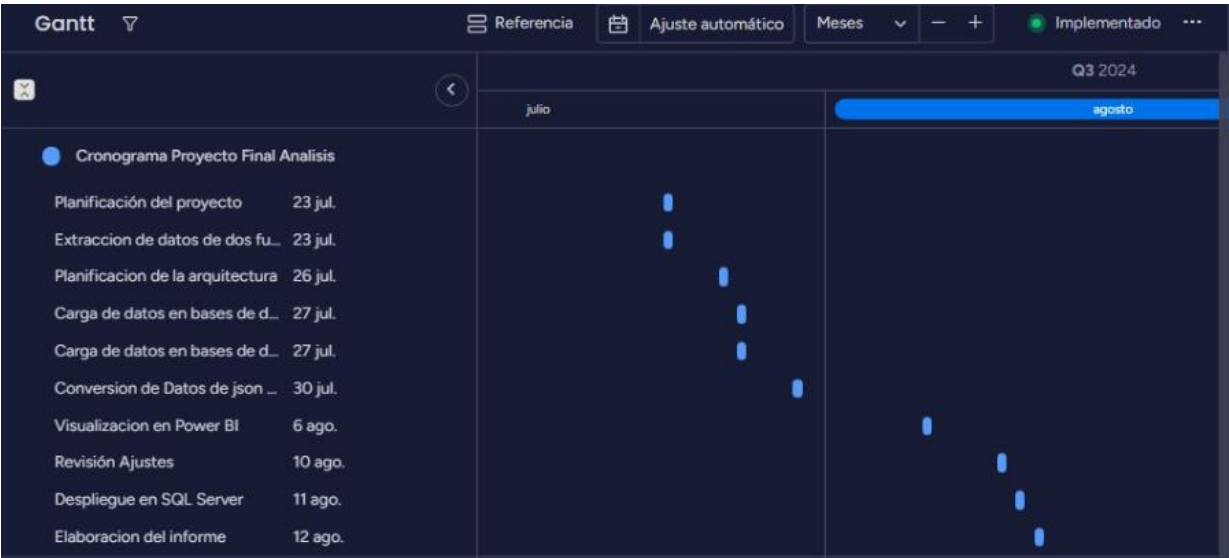


Diagrama Gantt Proyecto Final Analisis					
<input type="checkbox"/>	Elemento		Persona	Estado	Fecha
<input type="checkbox"/>	Planificación del proyecto	⊕	E +2	Listo	23 jul.
<input type="checkbox"/>	Extraccion de datos de dos fuentes distintas	⊕	JJ +2	Listo	23 jul.
<input type="checkbox"/>	Planificacion de la arquitectura	⊕	AT +2	Listo	26 jul.
<input type="checkbox"/>	Carga de datos en bases de datos Relacionales	⊕	AT +2	Listo	27 jul.
<input type="checkbox"/>	Carga de datos en bases de datos no SQL	⊕	AT	Listo	27 jul.
<input type="checkbox"/>	Conversion de Datos de json a csv y viceversa	⊕	JJ +2	Listo	30 jul.
<input type="checkbox"/>	Visualizacion en Power BI	⊕	E +2	Listo	6 ago.
<input type="checkbox"/>	Revisión Ajustes	⊕	JJ +2	Listo	10 ago.
<input type="checkbox"/>	Despliegue en SQL Server	⊕	E +2	En Proceso	11 ago.
<input type="checkbox"/>	Elaboracion del informe	⊕	JJ +2	En Proceso	12 ago.

- **Recurso y herramientas utilizadas.**

- Bases de Datos Relacionales: MySQL- SQL Server
- Bases de Datos NoSQL: MongoDB
- Herramientas de conversión de datos: CSV - JSON
- Herramientas de Visualización: Power BI
- Otras Herramientas: Microsoft Project, Microsoft Excel

- **Arquitectura de la solución.**



- **Extracción de datos.**

La extracción de los datos se obtuvo de los siguientes links, de acuerdo con la categoría en mención.

Evento Deportivos a nivel mundial

- <https://www.kaggle.com/datasets/amirba/olympic-sports-and-medals-18962021>
- https://figshare.com/articles/dataset/Olympic_history_longitudinal_data_scraped_from_www_sports-reference_com/6121274

Actividades y Hobbies

- <https://www.kaggle.com/datasets/mfrancis23/museum-of-modern-art-collection?select=Artworks.csv>
- <https://www.kaggle.com/datasets/drahulsingh/largest-art-museums>

Conciertos y Evento Públicos

- <https://data.world/ben-pfeifer/nirvana-live-performances>
- <https://www.kaggle.com/datasets/dmytroievdyokymov/metallica-concerts-data>

Restaurantes y Sitios de Esparcimientos











- <https://www.kaggle.com/datasets/phiitm/chennai-zomato-restaurants-data>
- <https://www.kaggle.com/datasets/shrutimehta/zomato-restaurants-data?select=zomato.csv>

Películas

- <https://www.kaggle.com/datasets/coltonbarger/rotten-tomatoes-reviews-for-online-streaming-shows>
- <https://www.kaggle.com/datasets/priyamchoksi/rotten-tomato-movie-reviews-1-44m-rows>

La descarga se realizó en un archivo zip. Para usar los archivos, primero se descomprime el zip, luego se renombran los datasets y se guardan en una carpeta organizada para facilitar el trabajo con ellos.

Archivos Descomprimidos:

Nombre
 Artworks.csv
 athlete_events.csv
 Largest-art-museums.csv
 Metallica_gigs.csv
 netflix_titles.csv
 Peliculas.csv
 rotten_tomatoes_movie_reviews.csv
 summer.csv
 Zomato Chennai Listing 2020.csv
 zomato.csv

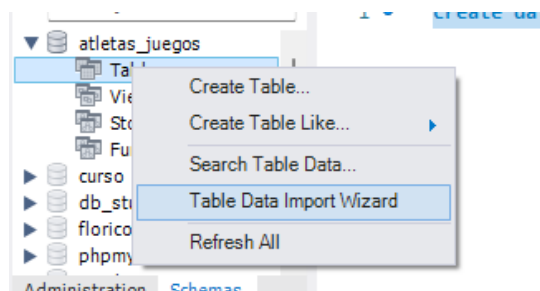
ANÁLISIS DE INFORMACIÓN.

Cambio de un archivo CSV a MYSQL

1. Creación de una Base de Datos en el programa MYSQL

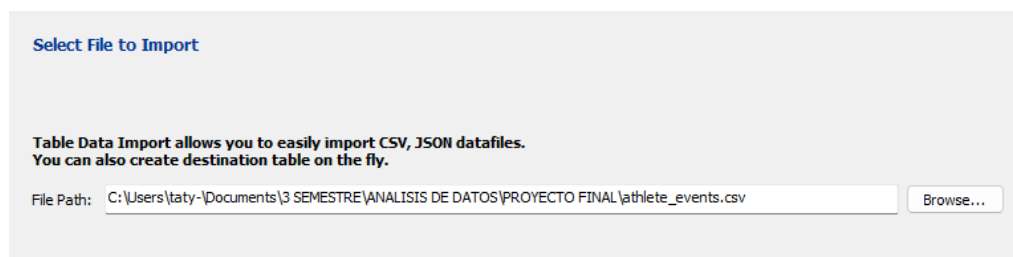


2. Dar clic derecho en la opción tables de la base de datos



3. Seguimos los paso requeridos para importar de un archivo CSV a nuestra base de datos

- Escogemos el archivo



- Elegimos la opción crear una nueva tabla

Select Destination

Select destination table and additional options.

☐ Use existing table:

☒ Create new table: ,

☐ Drop table if exists

- Elegimos el formato de nuestra tabla

Configure Import Settings

Detected file format: csv

Encoding:

Columns:

<input checked="" type="checkbox"/> Source Column	Field Type
<input checked="" type="checkbox"/> ID	<input type="text" value="int"/>
<input checked="" type="checkbox"/> Name	<input type="text" value="text"/>
<input checked="" type="checkbox"/> Sex	<input type="text" value="text"/>
<input checked="" type="checkbox"/> Age	<input type="text" value="int"/>
<input checked="" type="checkbox"/> Height	<input type="text" value="text"/>
<input checked="" type="checkbox"/> Weight	<input type="text" value="text"/>

ID	Name	Sex	Age	Height	Weight	Team	NOC	Games	Year
1	A Dijiang	M	24	180	80	China	CHN	1992 Sum...	1992
2	A Lamusi	M	23	170	60	China	CHN	2012 Sum...	2012
3	Gunnar Niel...	M	24	NA	NA	Denmark	DEN	1920 Sum...	1920
4	Edgar Linde...	M	34	NA	NA	Denmark/S...	DEN	1900 Sum...	1900

- Importamos la tabla desde nuestro archivo CSV

The following tasks will now be performed. Please monitor the execution.

☒ Prepare Import

☒ Import data file

Data import

File C:\Users\taty-\Documents\3 SEMESTRE\ANALISIS DE DATOS\PROYECTO FINAL\athlete_events.csv was imported in 195.274 s

Table atletas_juegos.athlete_events was created

19928 records imported

- Visualizamos la importación del archivo.

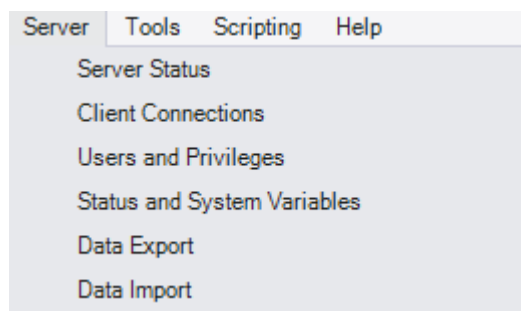
```

1 • create database Atletas_Juegos;
2 • use Atletas_Juegos;
3 • select * from athlete_events
4

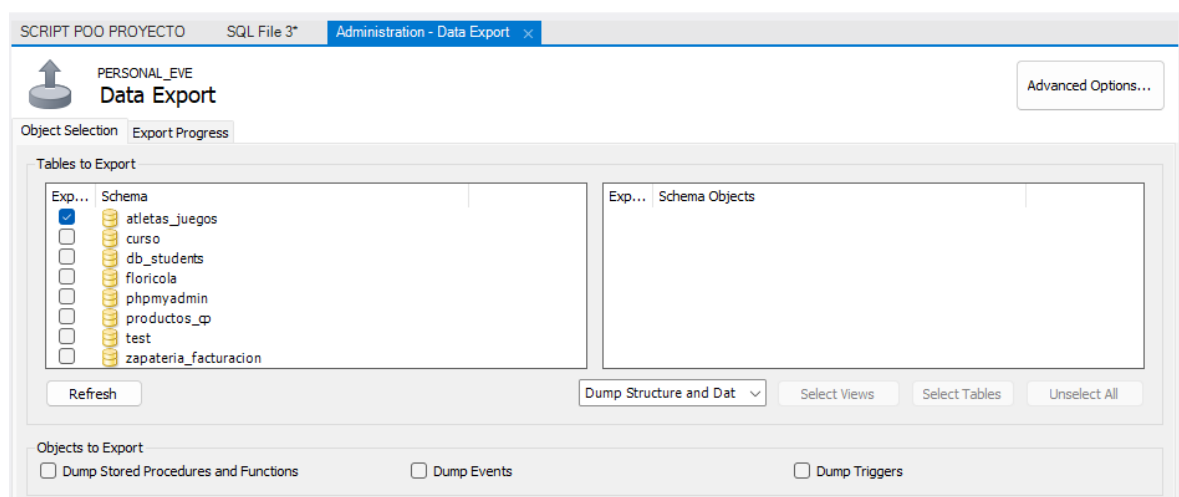
```

Result Grid															
Filter Rows: <input type="text"/> Export: <input type="text"/> Wrap Cell Content: <input type="text"/> Fetch rows: <input type="text"/>															
	ID	Name	Sex	Age	Height	Weight	Team	NOC	Games	Year	Season	City	Sport	E	
▶	1	A Dijiang	M	24	180	80	China	CHN	1992 Summer	NULL	Summer	Barcelona	Basketball	B	
	2	A Lamusi	M	23	170	60	China	CHN	2012 Summer	NULL	Summer	London	Judo	Ju	
	3	Gunnar Nielsen Aaby	M	24	NA	NA	Denmark	DEN	1920 Summer	NULL	Summer	Antwerpen	Football	Fi	
	4	Edgar Lindenau Aabye	M	34	NA	NA	Denmark/Sweden	DEN	1900 Summer	NULL	Summer	Paris	Tug-Of-War	Ti	
	5	Christine Jacoba Aaftink	F	21	185	82	Netherlands	NED	1988 Winter	NULL	Winter	Calgary	Speed Skating	Sj	
	5	Christine Jacoba Aaftink	F	21	185	82	Netherlands	NED	1988 Winter	NULL	Winter	Calgary	Speed Skating	Sj	
	5	Christine Jacoba Aaftink	F	25	185	82	Netherlands	NED	1992 Winter	NULL	Winter	Albertville	Speed Skating	Sj	
	5	Christine Jacoba Aaftink	F	25	185	82	Netherlands	NED	1992 Winter	NULL	Winter	Albertville	Speed Skating	Sj	
	5	Christine Jacoba Aaftink	F	27	185	82	Netherlands	NED	1994 Winter	NULL	Winter	Lillehammer	Speed Skating	Sj	
	5	Christine Jacoba Aaftink	F	27	185	82	Netherlands	NED	1994 Winter	NULL	Winter	Lillehammer	Speed Skating	Sj	
	6	Per Knut Aaland	M	31	188	75	United States	USA	1992 Winter	NULL	Winter	Albertville	Cross Countr...	C	
	6	Per Knut Aaland	M	31	188	75	United States	USA	1992 Winter	NULL	Winter	Albertville	Cross Countr...	C	
	6	Per Knut Aaland	M	31	188	75	United States	USA	1992 Winter	NULL	Winter	Albertville	Cross Countr...	C	
	6	Per Knut Aaland	M	31	188	75	United States	USA	1992 Winter	NULL	Winter	Albertville	Cross Countr...	C	
	6	Per Knut Aaland	M	33	188	75	United States	USA	1994 Winter	NULL	Winter	Lillehammer	Cross Countr...	C	
	6	Per Knut Aaland	M	33	188	75	United States	USA	1994 Winter	NULL	Winter	Lillehammer	Cross Countr...	C	
	6	Per Knut Aaland	M	33	188	75	United States	USA	1994 Winter	NULL	Winter	Lillehammer	Cross Countr...	C	

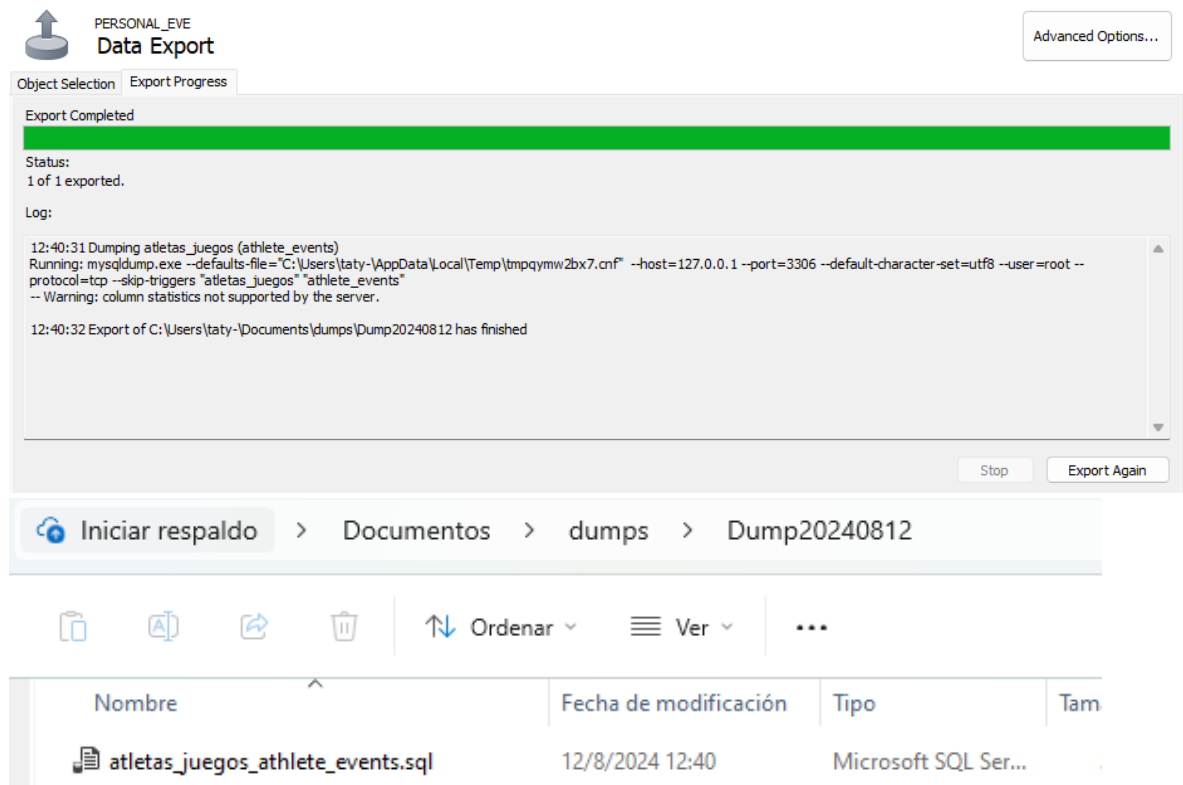
MIGRACION DE UN ARCHIVO DE MYSQL A SQL SERVER



- Seleccionamos la base de datos

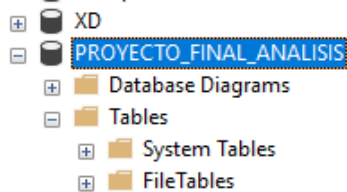


- Exportamos la base de datos en un archivo SQL, el cual se guarda en la carpeta dumps.

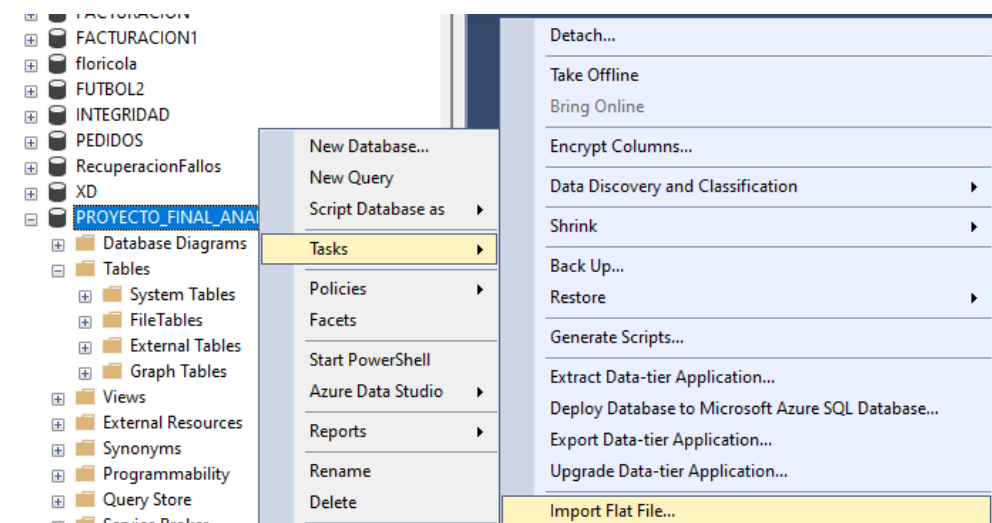


El archivo descargado abrimos en SQL Server

- Creamos una base de datos donde contendrá todas nuestras fuentes de datos



- En la base de datos creada en SQL Server procedemos a importar los archivos SQL.



- Seleccionamos el archivo SQL

Specify Input File

This operation will create a table from your input file.

Location of file to be imported

C:\Users\taty-\Documents\dumps\Dump20240812\atletas_juegos_athlete_events.sql

Browse...

New table name:

atletas_juegos_athlete_events

Table schema:

dbo

**Preview Data**

Introduction

Specify Input File

Preview Data

Modify Columns

Summary

Results

Help

Preview Data

This operation analyzed the input file structure to generate the preview below for up to the first 50 rows.

ID	Name	Sex	Age	Height	Weight
1	A Dijiang	M	24	180	80
2	A Lamusi	M	23	170	60
3	Gunnar Nielsen...	M	24	NA	NA
4	Edgar Lindenu...	M	34	NA	NA
5	Christine Jacob...	F	21	185	82
5	Christine Jacob...	F	21	185	82
5	Christine Jacob...	F	25	185	82
5	Christine Jacob...	F	25	185	82
5	Christine Jacob...	F	27	185	82
5	Christine Jacob...	F	27	185	82
6	Per Knut Aaland	M	31	188	75
6	Per Knut Aaland	M	31	188	75
6	Per Knut Aaland	M	31	188	75
6	Per Knut Aaland	M	31	188	75
6	Per Knut Aaland	M	33	188	75
6	Per Knut Aaland	M	32	188	75

☒ Use Rich Data Type Detection - may provide a closer type fit. However, cells with anomalous values may be dropped.

- Seleccionamos y modificamos los tipos de datos y configuraciones de cada columna.

Modify Columns

This operation generated the following table schema. Please verify if schema is accurate, and if not, please make any changes.

Column Name	Data Type	Primary Key	Allow Nulls
ID	tinyint	<input type="checkbox"/>	<input type="checkbox"/>
Name	nvarchar(50)	<input type="checkbox"/>	<input type="checkbox"/>
Sex	nvarchar(50)	<input type="checkbox"/>	<input type="checkbox"/>
Age	tinyint	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Height	tinyint	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Weight	float	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Team	nvarchar(50)	<input type="checkbox"/>	<input type="checkbox"/>
NOC	nvarchar(50)	<input type="checkbox"/>	<input type="checkbox"/>
Games	nvarchar(50)	<input type="checkbox"/>	<input type="checkbox"/>
Year	smallint	<input type="checkbox"/>	<input type="checkbox"/>
Season	nvarchar(50)	<input type="checkbox"/>	<input type="checkbox"/>
City	nvarchar(50)	<input type="checkbox"/>	<input type="checkbox"/>
Sport	nvarchar(50)	<input type="checkbox"/>	<input type="checkbox"/>
Event	nvarchar(100)	<input type="checkbox"/>	<input type="checkbox"/>
Medal	nvarchar(50)	<input type="checkbox"/>	<input type="checkbox"/>

Row granularity of error reporting (performance impact with smaller ranges)

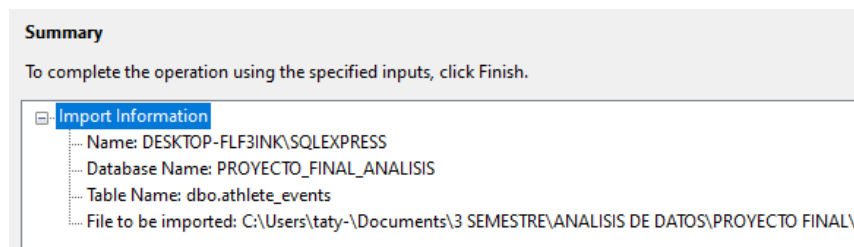
No Range

< Previous

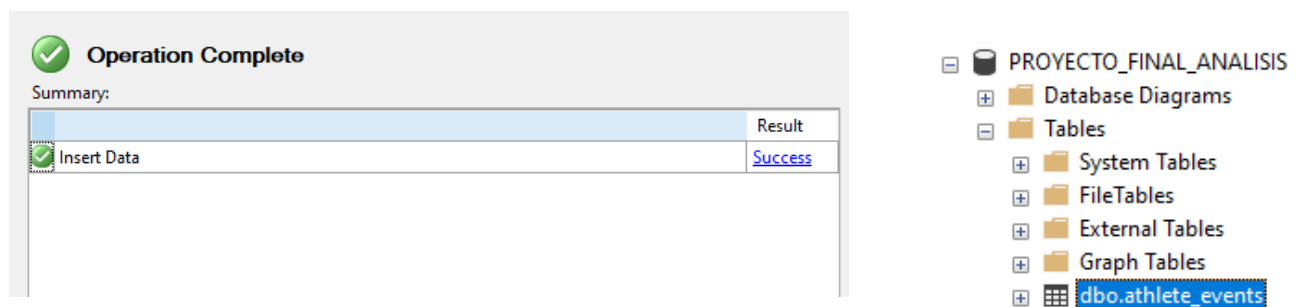
Next >

Cancel

- Verificamos la información.



- Verificamos que la data se encuentre en la base de datos que hemo creado.

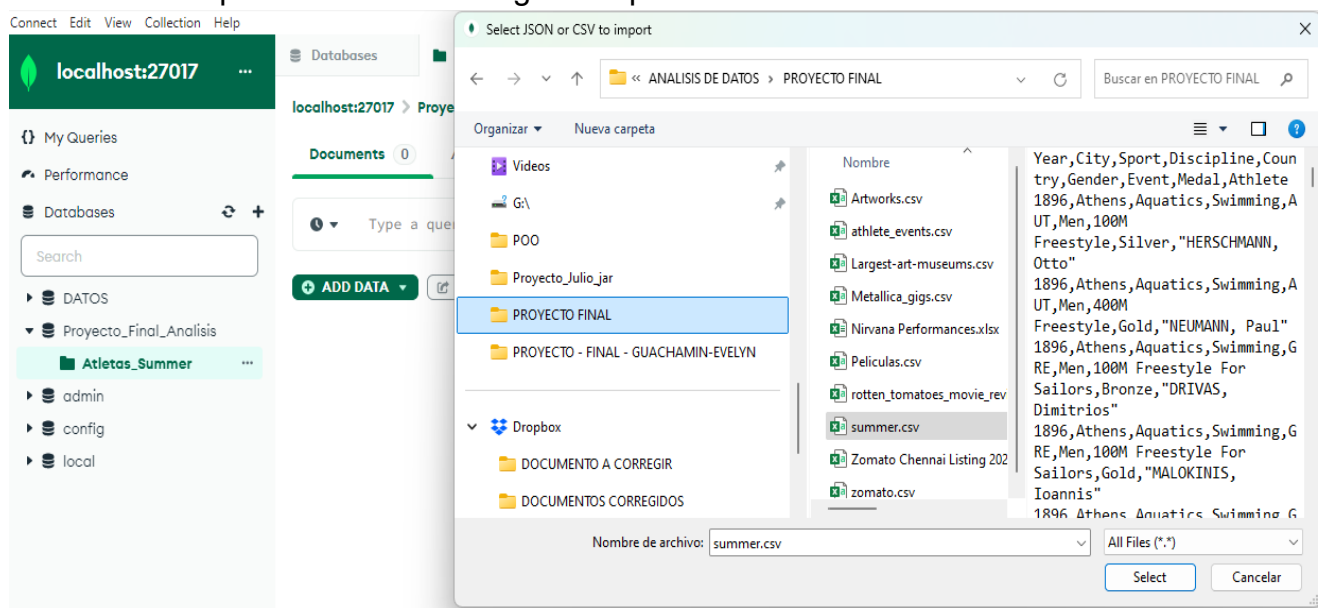


Conexión desde MongoDB a MYSQL Y SQL SERVER.

La importación o migración de los archivos se puede realizar a partir de un programa. Esta herramienta permite realizar migración desde SQL – MONGODB o viceversa.

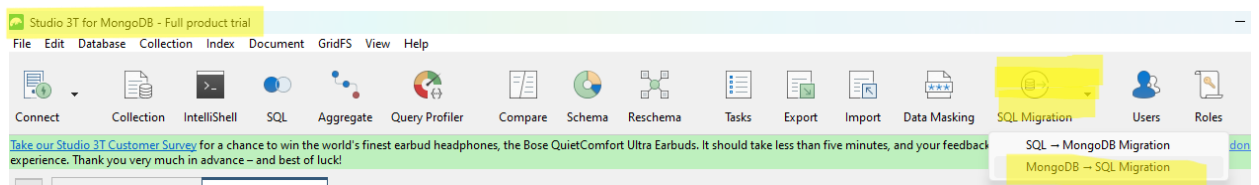
La Herramienta se llama Studio 3T, para utilizarla se realiza los siguientes pasos:

- 1- Se debe importar la base a MongoDB a partir de un formato CSV o JSON.

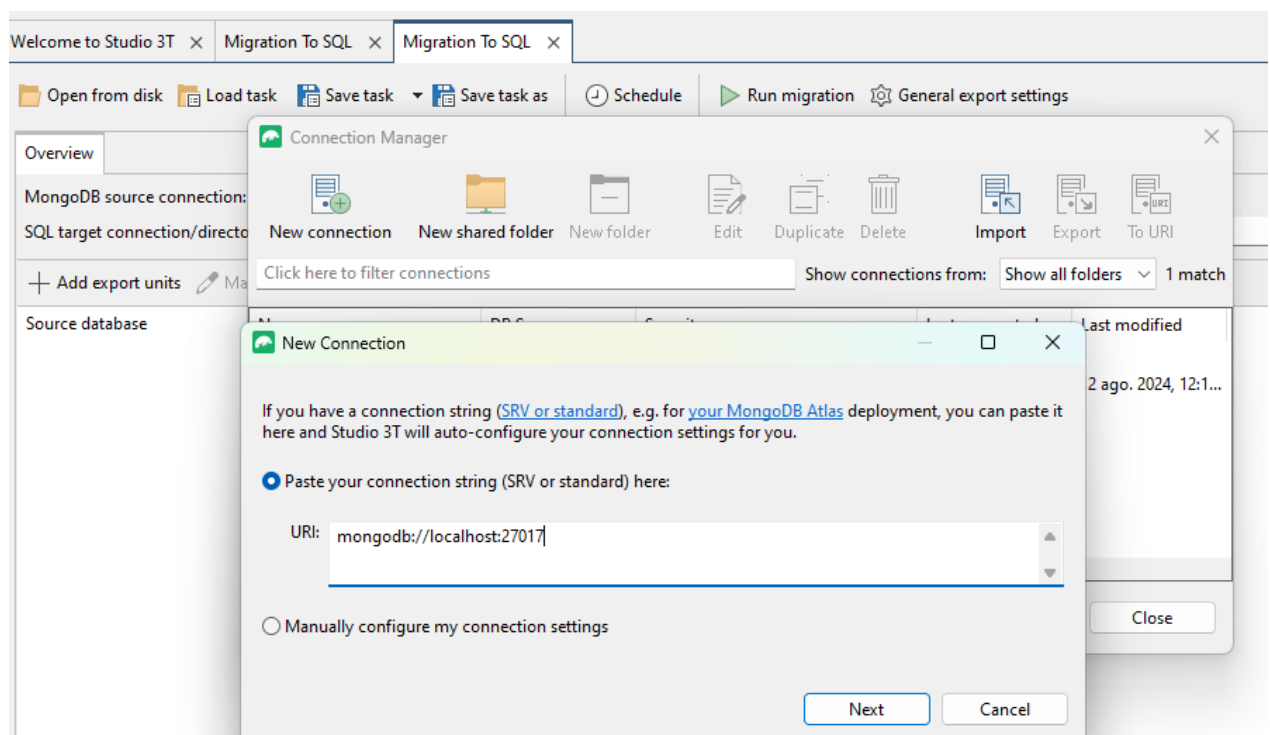




2- Con el programa descargado procedemos a realizar la migración



3- Al ser la primera conexión procedemos a guardar la conexión que tendremos con nuestro MongoDB



- La Uri encontramos en la página principal de MongoDB

New Connection

Connect to a MongoDB deployment



FAVORITE

URI ⓘ

Edit Connection String ☒

mongodb://localhost:27017/

- Agregamos un nombre a la conexión que usaremos

Name	DB Server	Security	Last connected	Last modified
Local resources				
Eve_Proyecto	localhost:27017		1 hour ago	12 ago. 2024, 12:1...

Overview

MongoDB source connection: Eve_Proyecto [Switch connection](#)

SQL target connection/directory: Director Eve_Proyecto (localhost:27017) ts\migration

- Añadimos la base de datos y la tabla que requerimos migrar

+ Add export units

Source database

Select a Collection or View

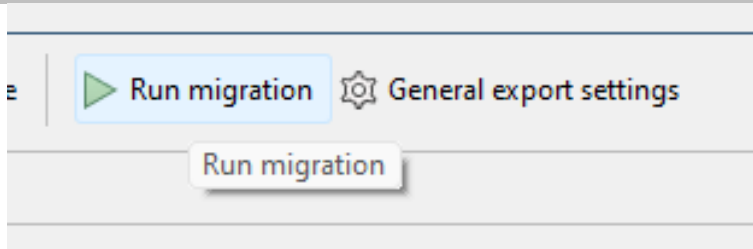
Please select a collection or view:

Refresh

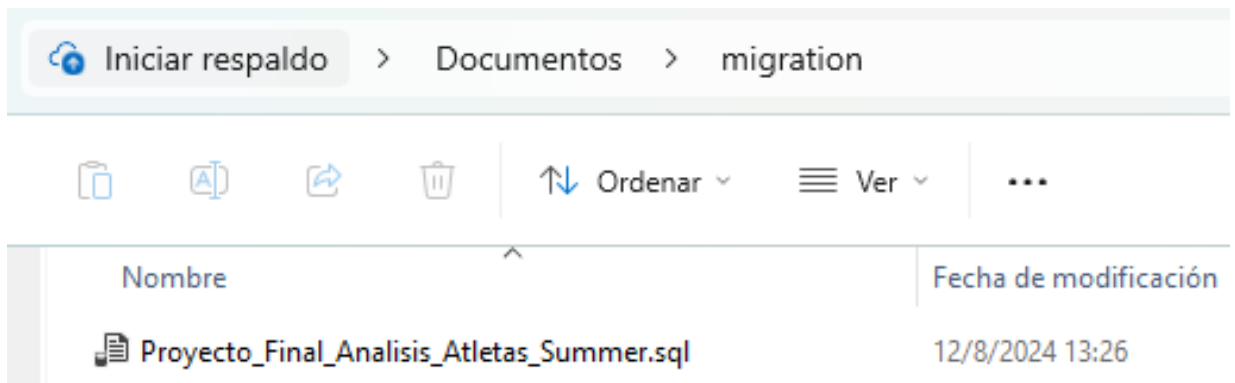
- ✓ Eve_Proyecto localhost:27017 [direct]
 - > DATOS
 - ✓ Proyecto_Final_Analisis
 - ✓ Collections (1)
 - Atletas_Summer
 - Views (0)
 - > admin
 - > config
 - > local

OK Cancel

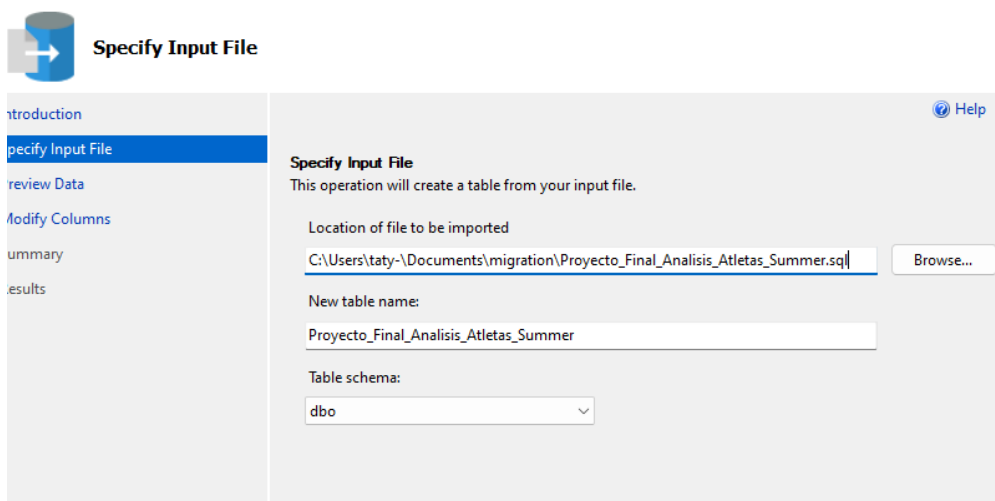
- Colocamos en un directorio nuestra migración y ejecutamos la migración



- De esta manera en nuestro directorio ya disponemos nuestro archivo SQL para la importación en SQLSERVER



- Realizamos las importaciones de nuestras bases de datos en SQL SERVER.



Modify Columns

This operation generated the following table schema. Please verify if schema is accurate, and if not, please make any changes.

Column Name	Data Type	Primary Key	<input checked="" type="checkbox"/> Allow Nulls	
Year	nvarchar(50) ▾	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
City	nvarchar(50) ▾	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
Sport	nvarchar(50) ▾	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
Discipline	nvarchar(50) ▾	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
Country	nvarchar(50) ▾	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
Gender	nvarchar(50) ▾	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
Event	nvarchar(50) ▾	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
Medal	nvarchar(50) ▾	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
Athlete	nvarchar(50) ▾	<input type="checkbox"/>	<input checked="" type="checkbox"/>	

Summary

To complete the operation using the specified inputs, click Finish.

[-] Import Information

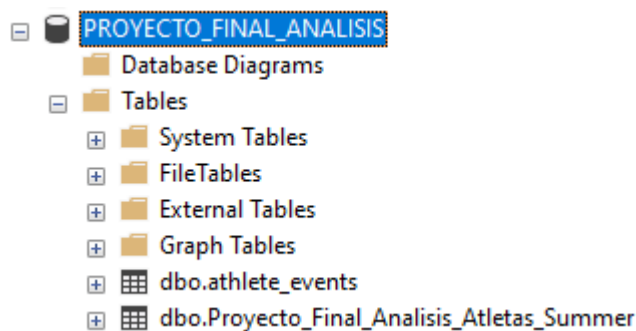
Name: DESKTOP-FLF3INK\SQLEXPRESS
 Database Name: PROYECTO_FINAL_ANALISIS
 Table Name: dbo.Proyecto_Final_Analisis_Atletas_Summer
 File to be imported: C:\Users\taty-\Documents\3 SEMESTRE\ANALISIS DE DATOS\PROYECTO FINAL\

**Operation Complete**

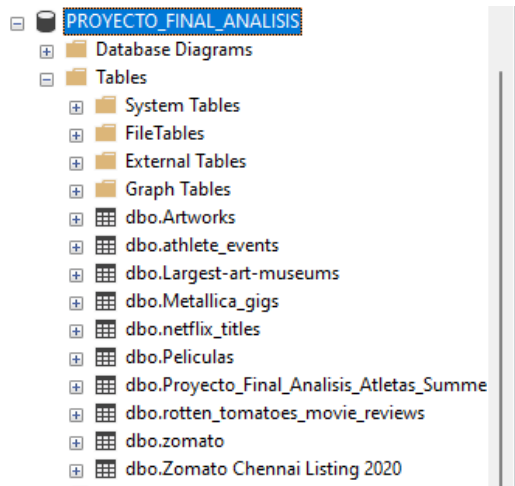
Summary:

Name	Result
Insert Data	Success

- Al finalizar de realizar la importación procedemos a verificar que la importación se realice de manera correcta.

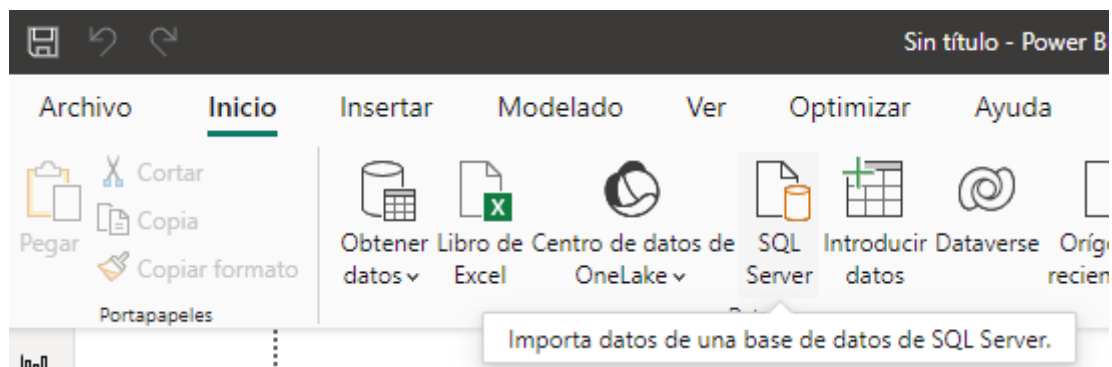


El proceso anterior se realiza para todos los dataset descargados para empezar a trabajar en POWER BI.



CONEXIÓN DE SQL SERVER CON POWER BI

- En POWER BI procedemos a escoger la opción SQL SERVER para establecer la conexión.



- Colocamos el nombre del servidor

Base de datos SQL Server

Servidor ①

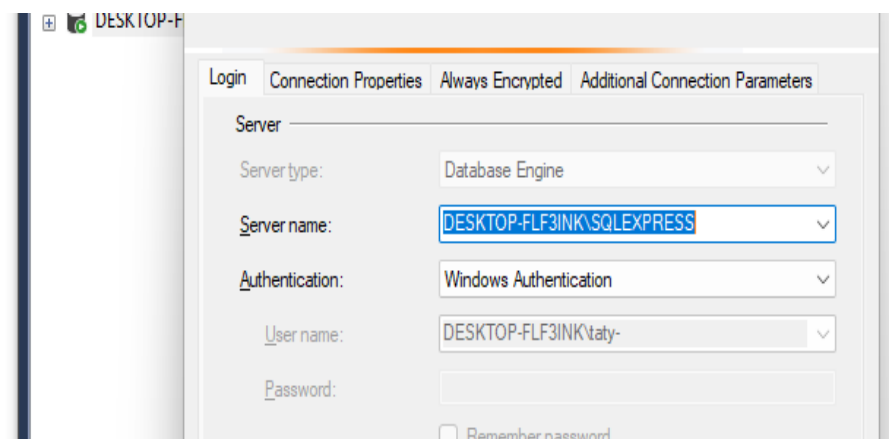
DESKTOP-FLF3INK\SQLEXPRESS

Base de datos (opcional)

Modo Conectividad de datos ①

☒ Importar

☐ DirectQuery



- Con ellos procedemos a aceptar para cargar la información seleccionamos las tablas y subimos la información.

Navegador

Opciones de presentación ▾

- EMPRESA_INMOBILIARIA
- FACTURACION
- FACTURACION1
- floricola
- FUTBOL2
- INTEGRIDAD
- PEDIDOS
- PROYECTO_FINAL_ANALISIS [10]
 - ☒ Artworks
 - ☒ athlete_events
 - ☒ Largest-art-museums
 - ☒ Metallica_gigs
 - ☒ netflix_titles
 - ☒ Peliculas
 - ☒ Proyecto_Final_Analisis_Atlas_Sum...
 - ☒ rotten_tomatoes_movie_reviews
 - ☒ zomato
 - ☒ Zomato Chennai Listing 2020
- RecuperacionFallos
- XD

Zomato Chennai Listing 2020

Zomato_URL	Name_of_Restaurant
https://www.zomato.com/chennai/yaa-mohaideen-briyani-1-pallavaram	Yaa Mohaideen Briyani
https://www.zomato.com/chennai/sukubhai-biriyani-alandur	Sukubhai Biriyani
https://www.zomato.com/chennai/ss-hyderabad-biryani-kodambakkam	SS Hyderabad Biryani
https://www.zomato.com/chennai/kfc-perambur	KFC
https://www.zomato.com/chennai/tasty-kitchen-perambur	Tasty Kitchen
https://www.zomato.com/chennai/dine-n-fun-medavakkam	Dine N Fun
https://www.zomato.com/chennai/bai-veetu-kalyanam-navallur	Bai Veetu Kalyanam
https://www.zomato.com/chennai/cafe-arabica-anna-nagar-east	Cafe Arabica
https://www.zomato.com/chennai/savoury-sea-shell-anna-nagar-east	Savoury Sea Shell
https://www.zomato.com/chennai/sangeetha-veg-restaurant-t-nagar	Sangeetha Veg Restaurant
https://www.zomato.com/chennai/mcdonalds-velachery	McDonald's
https://www.zomato.com/chennai/kovai-alankar-vilas-anna-nagar-east	Kovai Alankar Vilas
https://www.zomato.com/chennai/hotel-al-buhari-vadapalani	Hotel Al Buhari
https://www.zomato.com/chennai/dindigul-thalappakatti-1-porur	Dindigul Thalappakatti
https://www.zomato.com/chennai/hotel-paramount-kilpauk	Hotel Paramount
https://www.zomato.com/chennai/welcome-hotel-purasavakkam	Welcome Hotel
https://www.zomato.com/chennai/arab-street-ashok-nagar	Arab Street
https://www.zomato.com/chennai/abids-chetpet	Abid's
https://www.zomato.com/chennai/wire-room-bar-kitchen-velachery	Wire Room Bar & Kitchen
https://www.zomato.com/chennai/malabar-biriyani-center-aminijikarai	Malabar Biriyani Center
https://www.zomato.com/chennai/velidis-ramapuram	Velidis
https://www.zomato.com/chennai/mami-tiffen-stall-mylapore	Mami Tiffen Stall
https://www.zomato.com/chennai/supriya-andhra-restaurant-1-thuraimangalam	Supriya Andhra Restaurant

Seleccionar tablas relacionadas

Cargar

Transformar datos

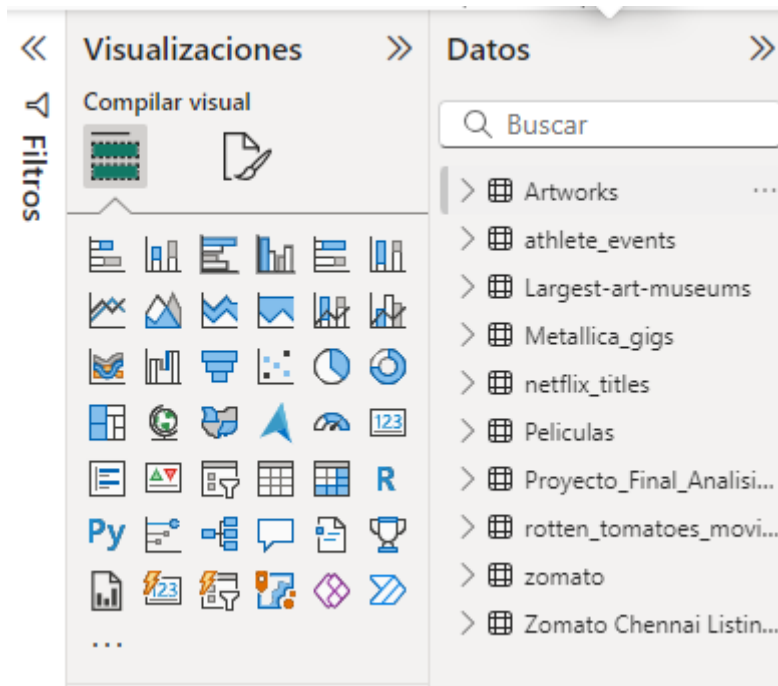
Cancelar

Cargar

- Artworks
138.161 filas de DESKTOP-
FLF3INK\SQLEXPRESS\PROYECTO_FINAL_ANALISIS.
- athlete_events
271.116 filas de DESKTOP-
FLF3INK\SQLEXPRESS\PROYECTO_FINAL_ANALISIS.
- Largest-art-museums
112 filas de DESKTOP-
FLF3INK\SQLEXPRESS\PROYECTO_FINAL_ANALISIS.
- Metallica_gigs
32.448 filas de DESKTOP-
- rotten_tomatoes_movie_reviews
1.444.963 filas de DESKTOP-
FLF3INK\SQLEXPRESS\PROYECTO_FINAL_ANALISIS.
- Zomato Chennai Listing 2020
12.032 filas de DESKTOP-
FLF3INK\SQLEXPRESS\PROYECTO_FINAL_ANALISIS.

Cancelar

- Con la información cargada procedemos a realizar los análisis requeridos



CONEXIÓN JSON - MONGODB ATLAS – MONGO DB COMPAS - MYSQL - SQL SERVER - POWER BI

- 1) **JSON:** Se obtuvo el dataset de **KAGGLE**, el cual se lo descargo en un formato csv.

Posterior a esto se utilizó una pagina online para convertir el archivo csv a un archivo json y así cumplir con la arquitectura planteada.

Subir un archivo CSV

Seleccionar archivo

Ningún archivo seleccionado

O pegue su CSV aquí

Zomato URL,Name of Restaurant,Address,Location,Cuisine,Top Dishes,Price for 2,Dining Rating,Dining Rating Count,Delivery Rating,Delivery Rating Count,Features
https://www.zomato.com/chennai/yaa-mohaideen-briyani-1-pallavaram,Yaa Mohaideen Briyani," 336 & 338, Main Road, Pallavaram, Chennai",Pallavaram, ['Biryani'],['Bread Halwa', ' Chicken 65', ' Mutton Biryani', ' Chicken Biryani', ' Brinjal Gravy', ' Basmati Rice', ' Raita'],500.0,4.3,1500,4.3,9306,['Home Delivery', 'Indoor Seating']
https://www.zomato.com/chennai/sukkubhai-biriyani-alandur,Sukkubhai Biriyani," New 14, Old 11/30, Railway Station Road, MKN Road, Near Antony Church, Alandur, Chennai",Alandur,['Biryani', ' North Indian', ' Mughlai', ' Desserts', ' Beverages'],['Beef Biryani', ' Beef Fry', ' Paratha', ' Paya', ' Brinjal Curry', ' Chicken 65', ' Chicken

Convertir

Claro

Separador

Detección automática

☒ Analizar números
☒ Analizar JSON
☐ Transponer

Producción:
☒ Formación
☐ Picadillo
☐ Minimizar

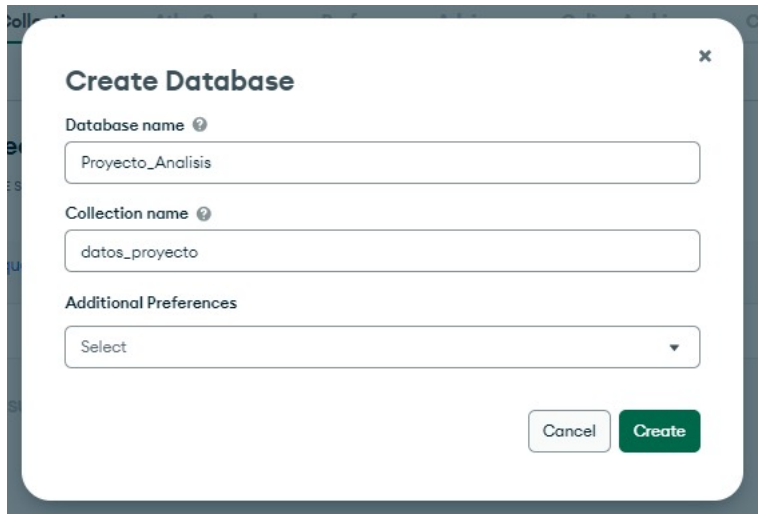
JSON

[
{
"Zomato URL": "https://www.zomato.com/chennai/yaa-mohaideen-briyani-1-pallavaram",
"Name of Restaurant": "Yaa Mohaideen Briyani",
"Address": "336 & 338, Main Road, Pallavaram, Chennai",
"Location": "Pallavaram",
"Cuisine": "['Biryani']",
"Top Dishes": "['Bread Halwa', ' Chicken 65', ' Mutton Biryani', ' Chicken Biryani', ' Brinjal Gravy', ' Basmati Rice', ' Raita']",
"Price for 2": 500,
"Dining Rating": 4.3,
"Dining Rating Count": 1500,
"Delivery Rating": 4.3,
"Delivery Rating Count": 9306,
"Features": "['Home Delivery', 'Indoor Seating']"
},
{

Descargar

Copiar al portapapeles

- 2) **MongoDB Atlas – MongoDB Compas:** Se inicio sesión en MongoDB Atlas para después crear un clúster en donde se creó una base de datos (Proyecto_Análisis) y una colección (datos_proyecto) en la nube.



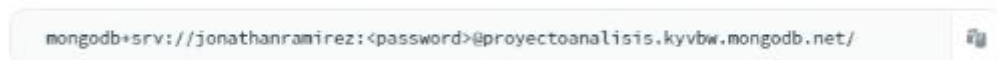
- Luego para hacer una conexión debemos añadir un usuario y contraseña



- Y se nos da una url para hacer la respectiva conexión

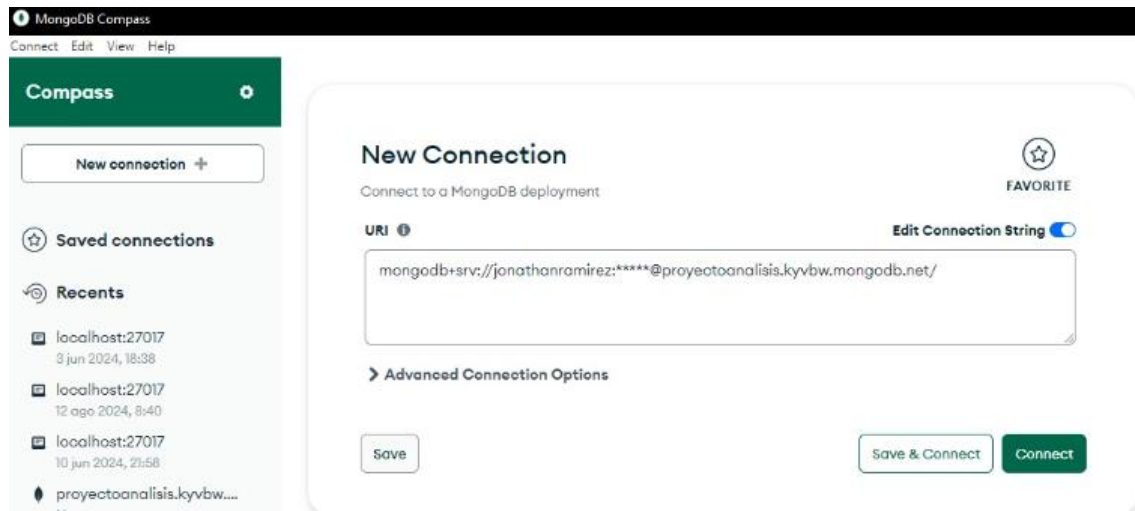
2. Copiela cadena de conexión y luego abra MongoDB Compass

Utilice esta cadena de conexión en su aplicación

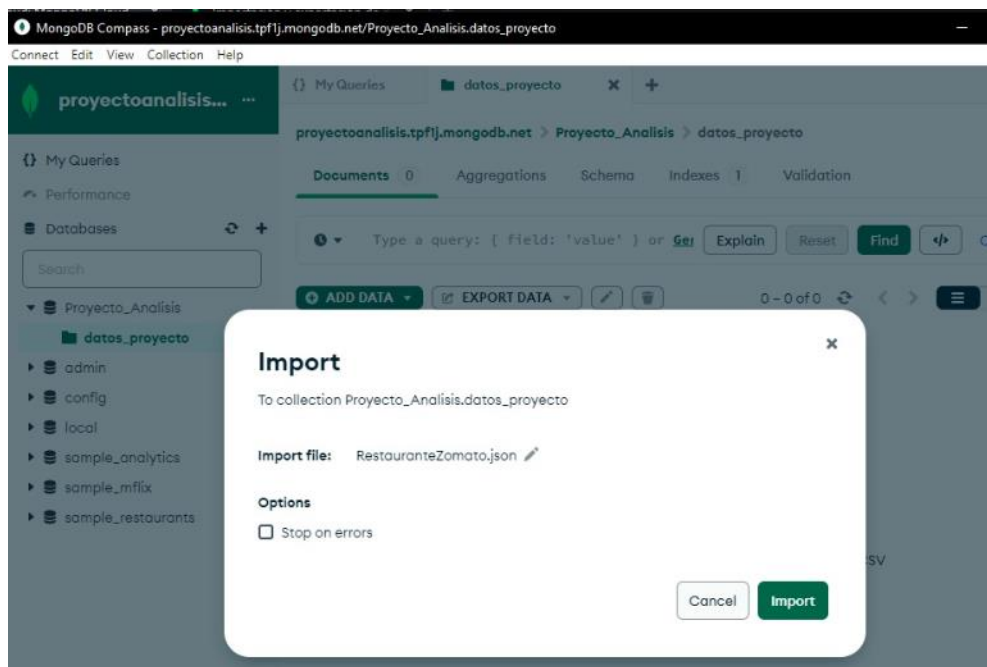


Reemplace **<password>** con la contraseña del usuario **jonathanramirez**. Asegúrese de que todas las opciones estén [codificadas en URL](#).

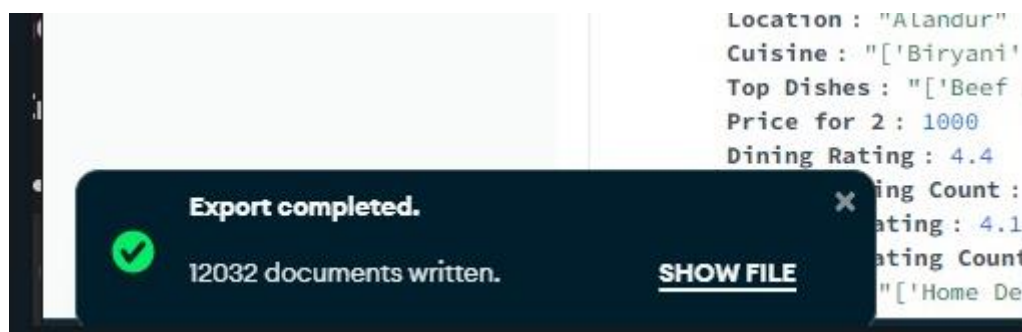
- Nos dirigimos al MongoDB Compass



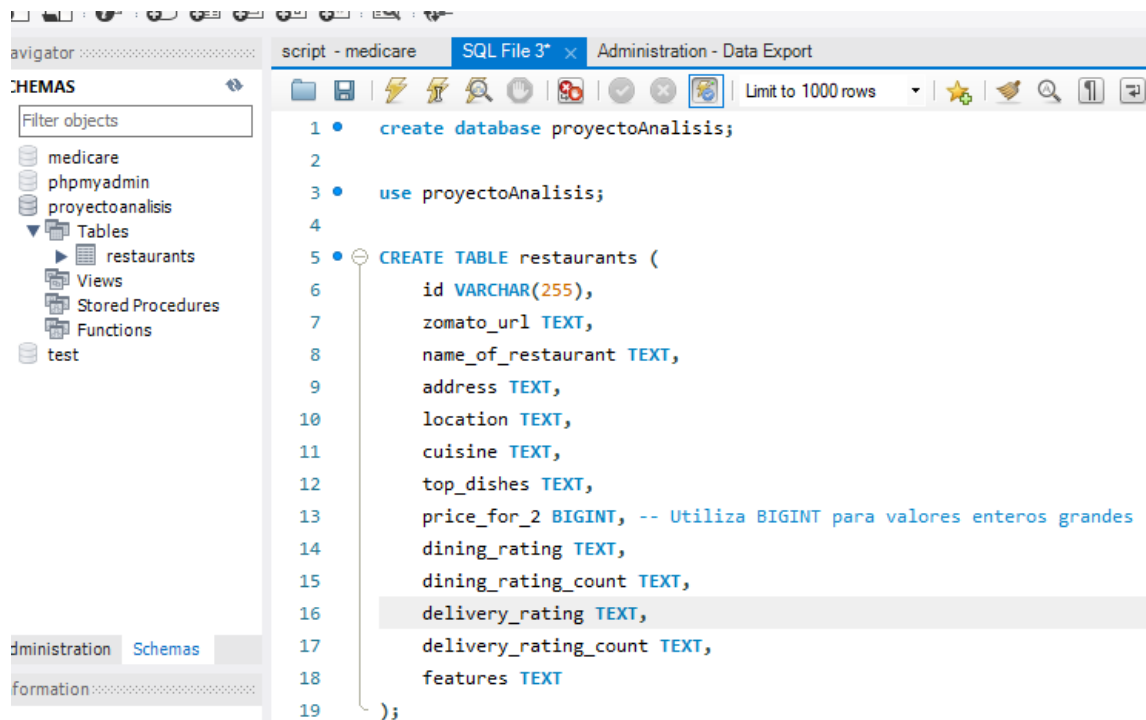
- Ya lista la conexión se importa el archivo JSON en la base de datos ya creada



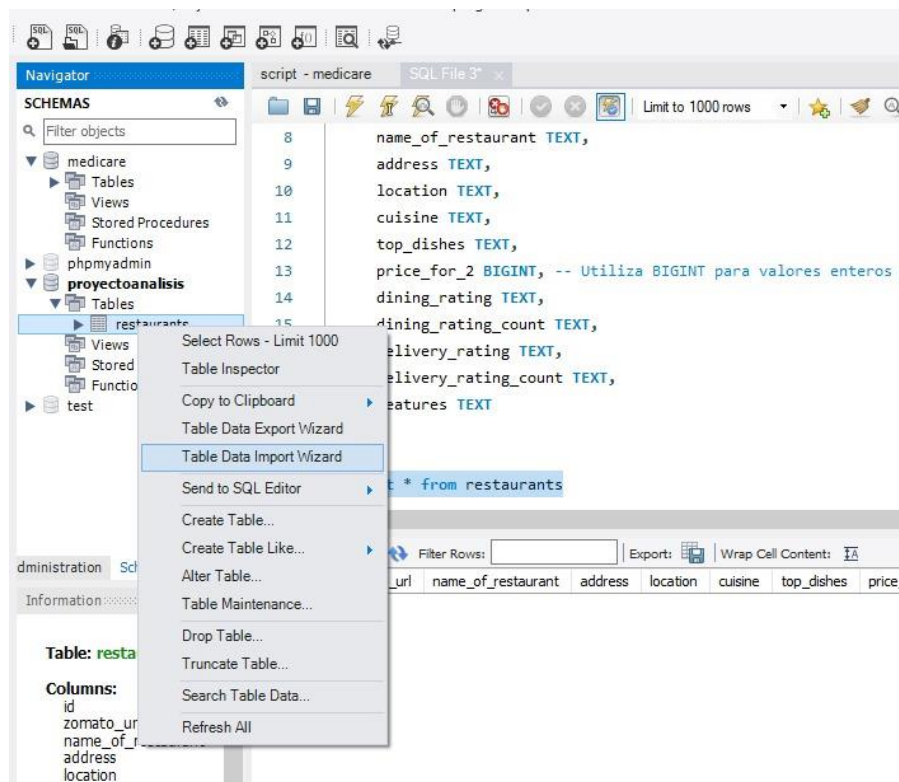
- Finalmente se exporta el archivo **JSON** en **CSV** para poder pasarlo a MySQL

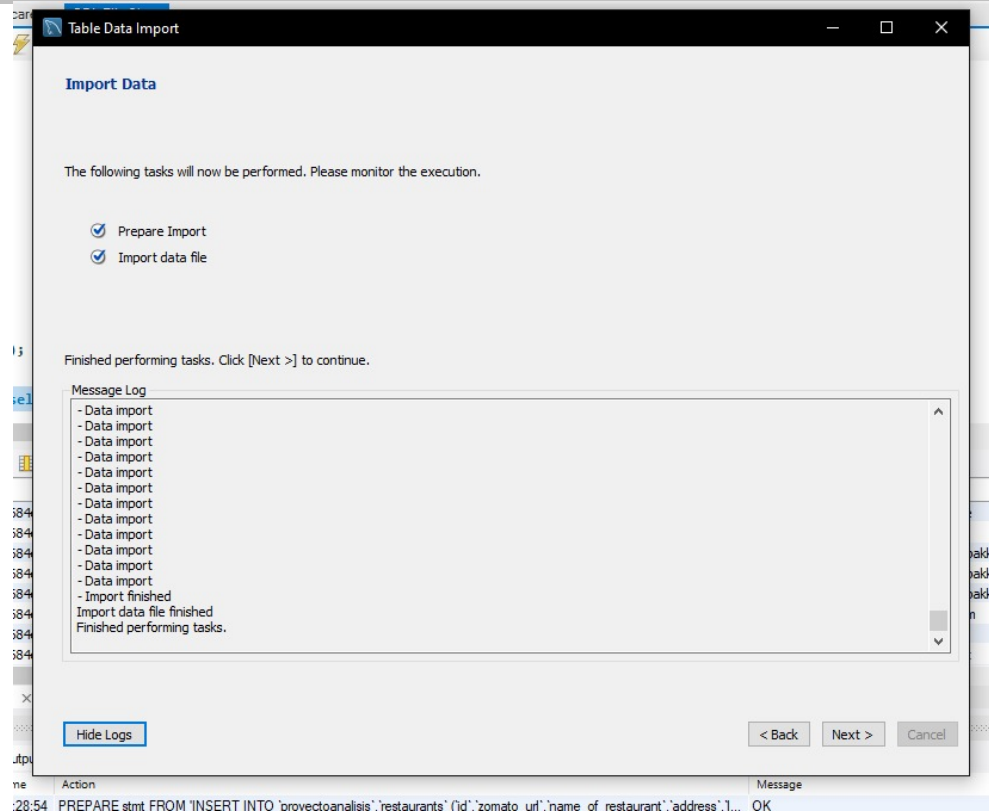


3) MYSQL: Creamos una base de datos y la tabla en donde se van a almacenar los datos del CSV.

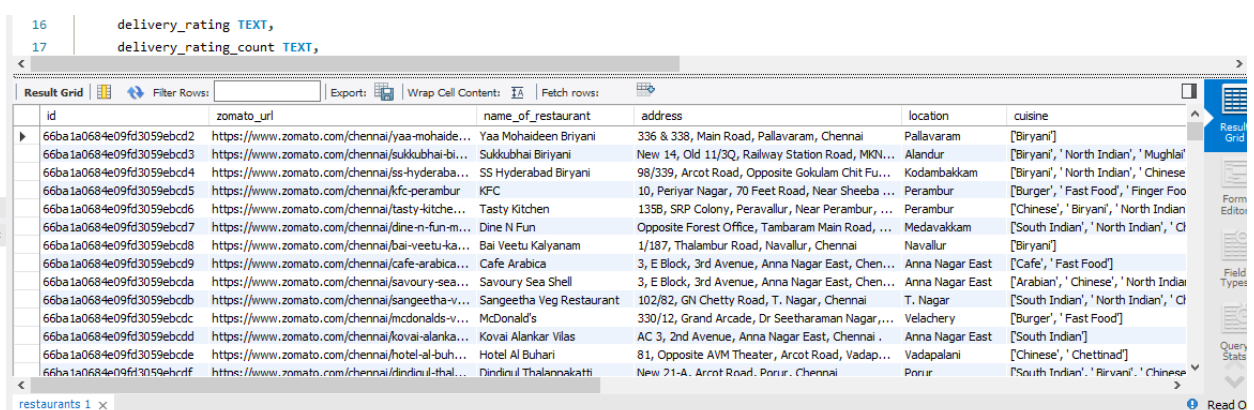


- Y exportamos nuestro csv en la tabla correspondiente

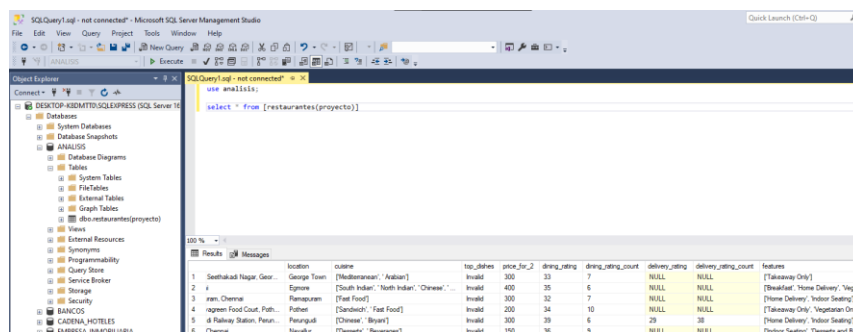




- Comprobamos los datos



- Y luego exportamos el archivo para poder abrirlo en SQL Server



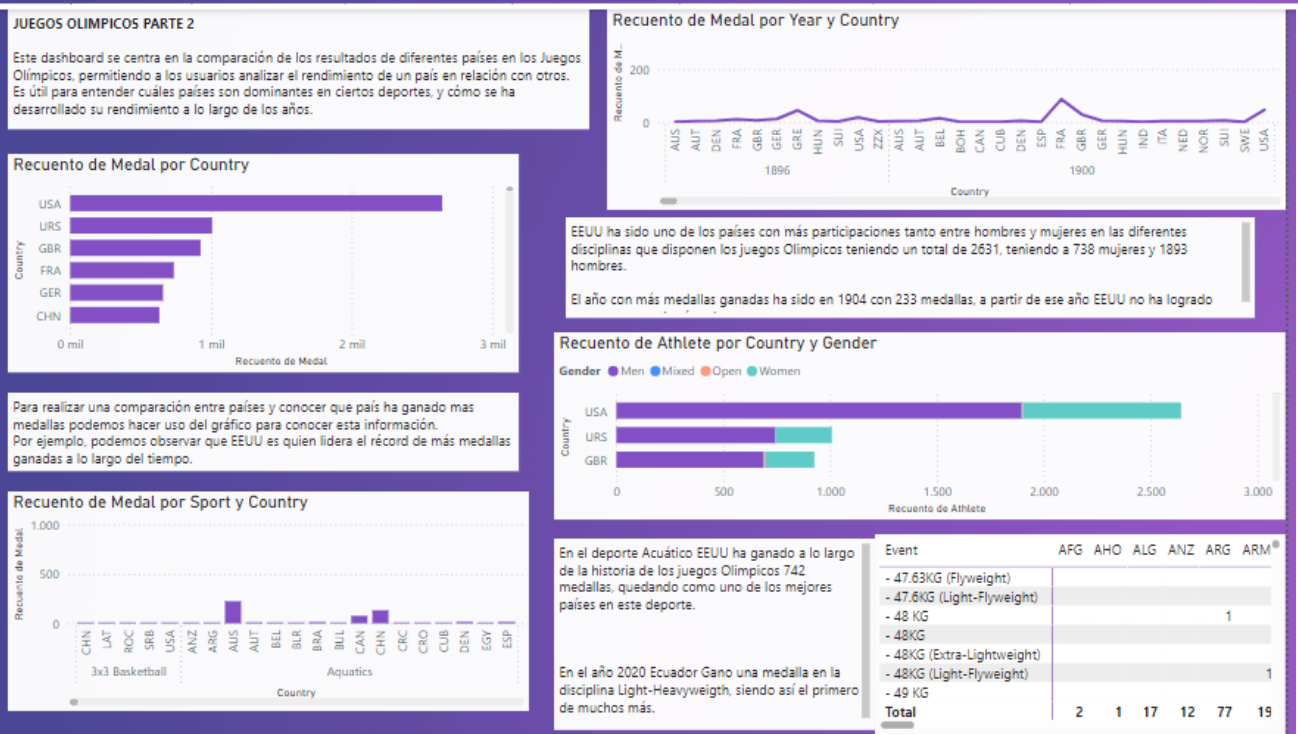
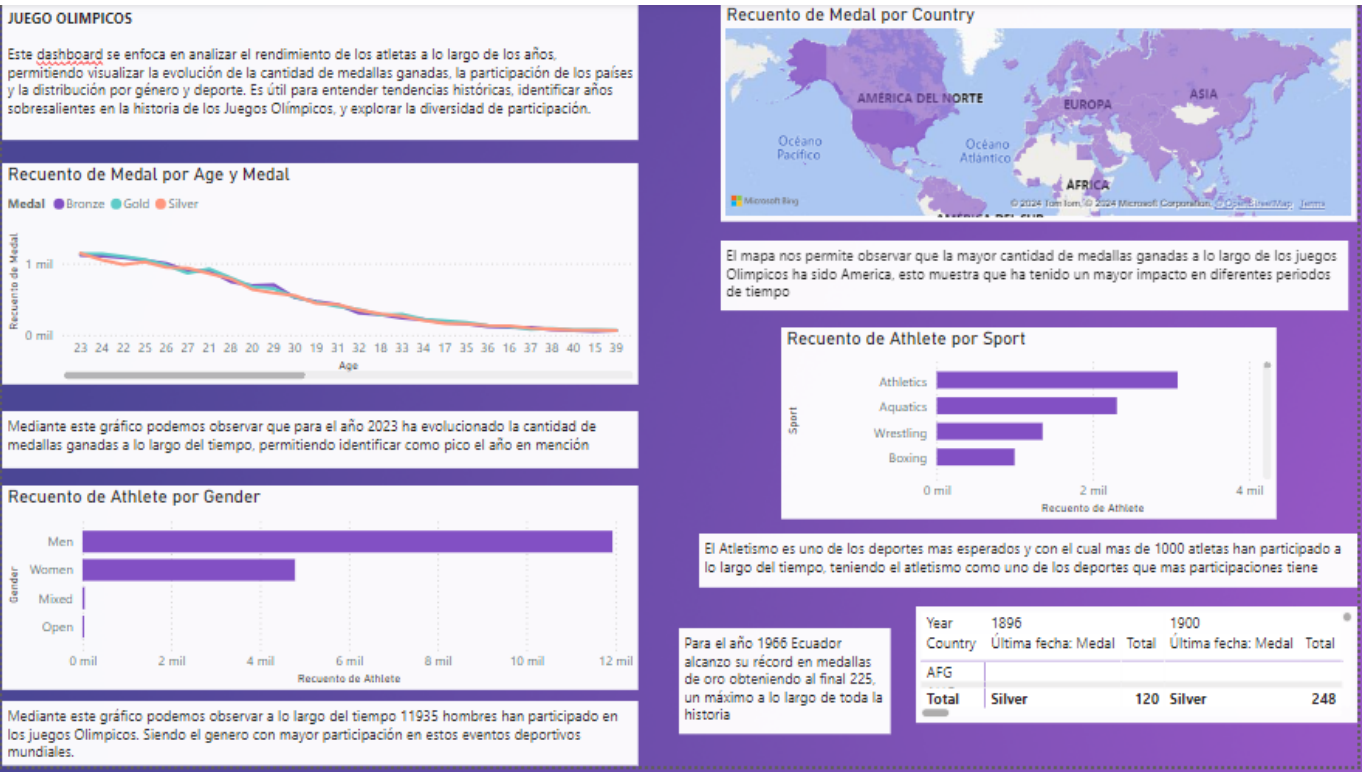
- Se importan los datos en el SQL Server para posteriormente hacer la conexión con el POWER BI

VISUALIZACIÓN Y RESULTADOS DE LA INFORMACIÓN

Presentación de DASHBOARD Y CONSULTAS realizados en POWER-BI

- Eventos deportivos a nivel mundial.

JUEGOS OLIMPICOS



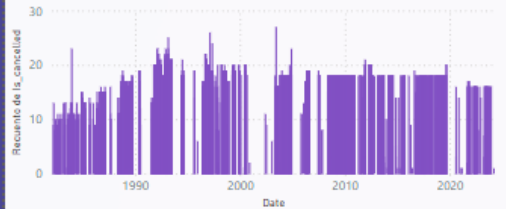
- Conciertos y eventos públicos.

Conciertos y eventos públicos

El siguiente DASHBOARD nos permite obtener informacion y comparación entre una banda de música y otra.

Numero de conciertos realizados por Año de la banda Metallica desde 1982 hasta el 2024

Recuento de Is_cancelled por Date



Numero de conciertos dados por la banda Metallica

2202

Recuento de Date

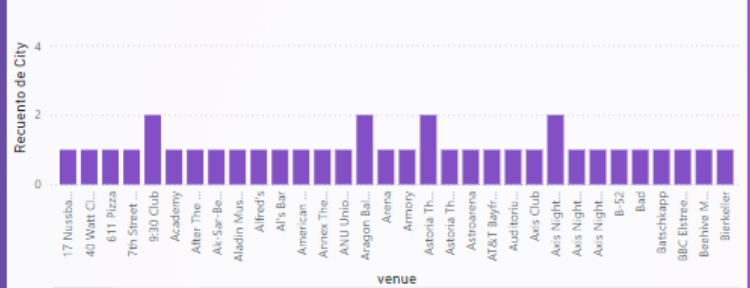
Numero de conciertos dados por la banda Nirvana

961

Recuento de Date

Numero de conciertos realizados por Nirvana desde 1987 hasta 1994

Recuento de City por venue



Como podemos ver hay una gran diferencia en la cantidad de conciertos dados por las dos bandas, y esto tiene sentido ya que el cantante de la banda Nirvana murió en 1994 por lo que el ultimo registro de sus conciertos fue en 1994 a diferencia de Metallica que a tenido sus últimos conciertos en 2024.

Conciertos y eventos públicos

El siguiente DASHBOARD nos permite obtener informacion y comparación entre una banda de música y otra.

Como se puede ver en los gráficos la banda Metallica a dado conciertos en todos los continentes del mundo a excepción de la banda Nirvana

Location y Location



Países donde la banda "Metallica" ha realizado conciertos.

Numero de conciertos que Metallica a cancelado por razones desconocidas

Yes

Recuento de Setlist_no

City and Country



Países donde la banda "Nirvana" ha realizado conciertos.

Location

- ☐ Aarhus,Denmark
- ☐ Aberdeen Township,New Je...
- ☐ Aberdeen,Washington,Unit...
- ☐ Abilene,Texas,United States

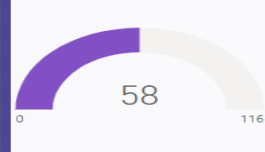
Country

- ☐ Argentina
- ☐ Australia
- ☐ Austria
- ☐ Belgium

Conciertos y eventos públicos

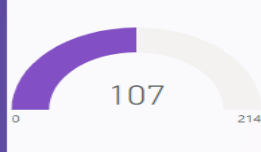
El siguiente DASHBOARD nos permite obtener informacion y comparación entre una banda de música y otra.

Recuento de Date



Numero de conciertos realizados por la banda Metallica en 1991

Recuento de Date



Numero de conciertos realizados por la banda Nirvana en 1991

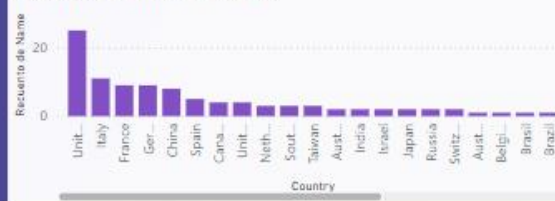
Como podemos ver en las graficas a pesar de que la banda Nirvana solo dio conciertos hasta el año de 1994 en el año 1991 Tuvo mas conciertos realizados que la banda Metallica esto es porque en el año de 1991 esta banda disparo su numero de discos vendidos y realizo un tour mundial logrando ganar mucha popularidad

- Actividades y hobbies.

Actividades y hobbies.

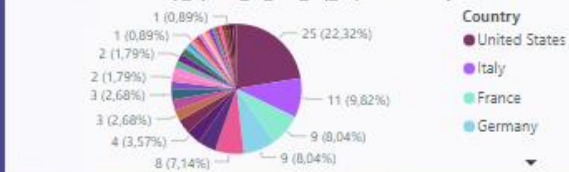
En el siguiente Dashboard se muestra las diferentes obras de arte que se realizan en los diferentes países, como un atractivo para visitar.

Recuento de Name por Country



Estados Unidos lidera con 25 museos, seguido por Italia con 11 y Francia con 9. Esto indica que estos países tienen una mayor concentración de museos, reflejando un mayor compromiso con la cultura y el patrimonio en comparación con otros países.

Recuento de Gallery_space_in_m2_sq_ft por Country



Se muestra la distribución del espacio de galería total entre diferentes países. Estados Unidos, Italia y Francia tienen la mayor proporción de espacio de exhibición. Esto indica que estos países poseen los museos más grandes en términos de espacio.

Recuento de Name por City and Country



Se muestra la distribución geográfica de los museos por ciudad y país en donde se permite identificar las áreas con mayor concentración de museos y su distribución global.

Name	Year_estal
Antonino Salinas Regional Archeological Museum	1866
Arsenal (Biennale)	1895
Art & History Museum	1835
Art Gallery of New South Wales	1874
Art Gallery of Ontario	1900

La tabla muestra los museos ordenados por año de establecimiento. El museo en la parte superior tiene el año más bajo, indicando que Capitoline Museums es el más antiguo.

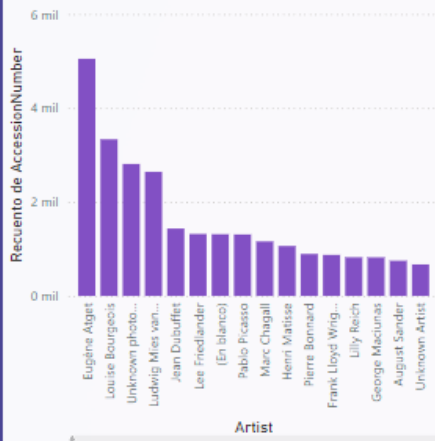
1471

Primera fecha: Year_established

Actividades y hobbies.

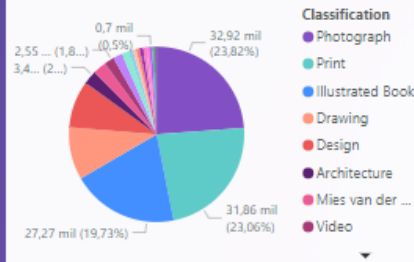
En el siguiente Dashboard se muestra las diferentes obras de arte que se realizan en los diferentes países, como un atractivo para visitar.

Recuento de AccessionNumber por Artist



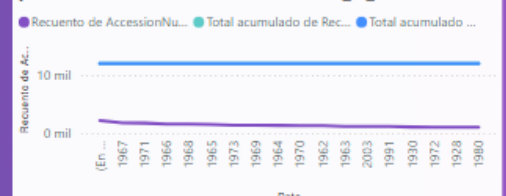
Se muestran cuántas obras ha creado cada artista. Los artistas con más obras aparecen con barras más largas, indicando su mayor producción.

Recuento de AccessionNumber por Classification



Se muestra cuáles son los tipos de obras más comunes en la colección. En donde la fotografía tiene un 23.82% casi a la par con el print que tiene un 23.06%

Recuento de AccessionNumber, Total acumulado de Recuento de Name_of_Restaurant en Name_of_Restaurant y Total acumulado de Recuento de Name_of_Restaurant e...

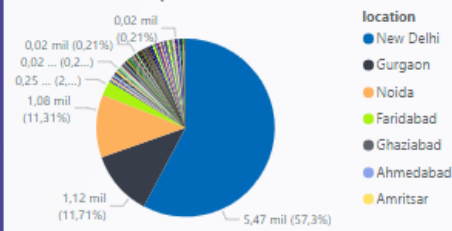


- Restaurantes y sitios de esparcimiento.

Restaurantes y sitios de esparcimiento.

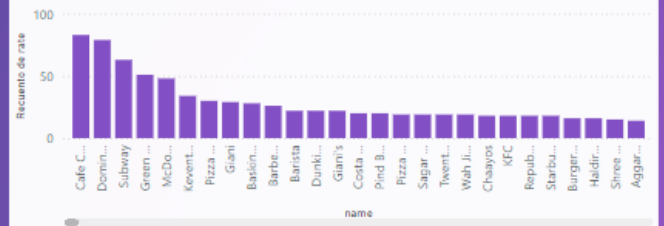
El Dashboard proporciona información de restaurantes de los cuales son mas cotizados y buscado en el mundo.

Recuento de name por location



Según el gráfico, área con mayor cantidad de restaurantes es New Delhi, lo cual podría indicar una alta competencia en esta zona.

Recuento de rate por name



Los restaurantes más valorados por los clientes en términos de cantidad de calificaciones son Cafe coffee day (96), onesta (85), just bake (74), estos restaurantes probablemente deben tener una clientela mas grande y mayor visibilidad en el mercado.

Locality_Verbose

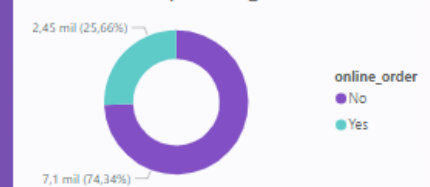
- ☐ ILD Trade Centre Mall...
- ☐ 12th Square Building,...
- ☐ A Hotel, Gurdev Nag...
- ☐ Aaya Nagar, New Delhi
- ☐ Abu Dhabi Mall, Touri...
- ☐ Abu Shagara, Sharjah
- ☐ Acropolis Mall, Kasba...

9551

Total acumulado de
Recuento de
Average_Cost_for_two en A...

Se observar una gran proporción de que los restaurantes NO ofrecen reservas en línea

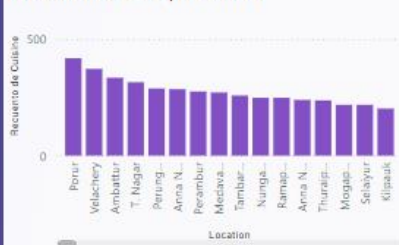
Recuento de name por online_order



Restaurantes y sitios de esparcimiento.

El Dashboard proporciona información de restaurantes de los cuales son mas cotizados y buscado en el mundo.

Recuento de Cuisine por Location

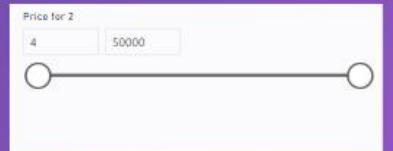


El recuento de Cocina fue el más alto para Porur con un total de 418, seguido de Velachery con 371, y Ambattur con 334

location



Se muestra la ubicación en la que se encuentran los diferentes establecimientos de cocina a nivel global



RANGO DE PRECIOS PARA 2

8328

Total de restaurantes

Domino's Pizza, Meat And Eat,
A2B - Adyar Ananda Bhavan, Cafe ...

Top 3 de mejores calificaciones de comedor de los restaurantes

- **Tema Libre:**

PELICULAS

ENTRENAMIENTO

PELICULAS

El siguiente DASHBOARD nos permite ver la cantidad de películas realizadas a partir de los años 80' en adelante hasta el año 2023 dando un total de mas 1 un millón de películas. Las películas son contadas ya sean reconocidas o no por el mundo.

1,44 mill.

Recuento de publicationName

Recuento de publicationName por originalScore y publicationName

La película Slant Magazine publicada a partir 2002 , es la película con mas numero de diferentes ediciones, versiones, o formatos en los que la película fue publicada a lo largo de los años, teniendo al final 347 publicaciones realizadas entre el periodo de 2002 hasta el 2022.

En interesante conocer que además las películas con una calificación inferior a lo esperado son películas publicadas entre el periodo de 2015 y 2023.

publicationName	Año	Trimestre	Mes	Día
(All (Parentheses))	2020	Trim. 4	octubre	7
(All (Parentheses))	2020	Trim. 4	octubre	19
(All (Parentheses))	2020	Trim. 4	noviembre	20
(All (Parentheses))	2020	Trim. 4	diciembre	31
(All (Parentheses))	2021	Trim. 1	marzo	26

Aunque por lo general la películas han tenido una buena aceptación por parte del publico no siempre es así y en el gráfico se puede observar que mas 954 películas han llegado a obtener una critica de 0/5o cero estrellas.

Recuento de publicationName por Año

La producción de las películas en años modernos ha ido en incremento, el gráfico muestra que a pesar de que en el año 2019 , en el cual la mayoría de las personas nos encontrábamos en confinamiento, el cine hizo lo contrario ya que en este año se realizaron 107.053 publicaciones , versiones de películas, dando este año como el mayor año de producción cinematográfica

Recuento de publicationName por originalScore

ENTRENAMIENTO

PELICULAS

El siguiente DASHBOARD nos permite ver la cantidad de películas con la que cuenta el servicio de NETFLIX.

Países ordenados por el numero de películas que se encuentran en la plataforma Netflix.

Estados Unidos es el país con mayor cantidad de peliculas en la plataforma, lo que indica una fuerte producción nacional y una amplia distribución de su contenido a nivel global.

Recuento de title por country

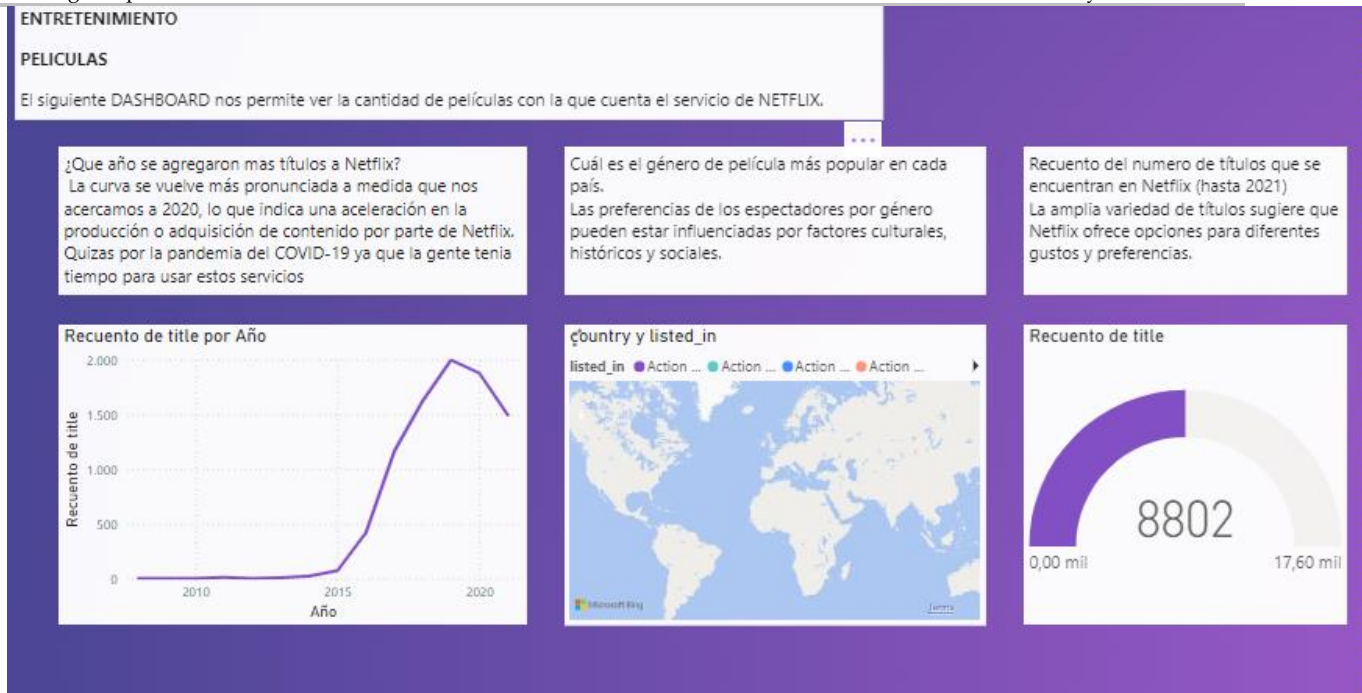
country	Recuento de title
United States	~2,800
India	~1,000
(En blanco)	~800
United Kingd...	~600
Japan	~400
South Korea	~300
Canada	~200
Spain	~150
France	~100
Mexico	~100
Egypt	~100
Turkey	~100
Nigeria	~100
Australia	~100
Taiwan	~100
Indonesia	~100
Brazil	~100
Philippines	~100
United Kingd...	~100
United States...	~100
Germany	~100
China	~100
Thailand	~100
Argentina	~100
Hong Kong	~100

Películas con un año de despliegue en Netflix menor a 2021

Aunque la distribución es relativamente uniforme, es posible que exista una ligera tendencia hacia el contenido más reciente, ya que las plataformas de streaming suelen priorizar las novedades para atraer a los suscriptores.

Suma de release_year por title y country

title	Suma de release_year
Death Note	2 mil (5,88%)
SAINT SEIYA: Knights of t...	2 mil (5,88%)
#Alive	2 mil (5,88%)
#blackAF	2 mil (5,88%)
#cats_the_mewvie	2 mil (5,88%)
#FriendButMarried 2	2 mil (5,88%)
(Un)Well	2 mil (5,88%)
100 Humans	2 mil (5,88%)
100% Halal	2 mil (5,88%)
13 Reasons Why	2 mil (5,88%)
18 Presents	2 mil (5,88%)
2 Hearts	2 mil (5,88%)



CONCLUSIONES Y RECOMENDACIONES

4- Conclusiones:

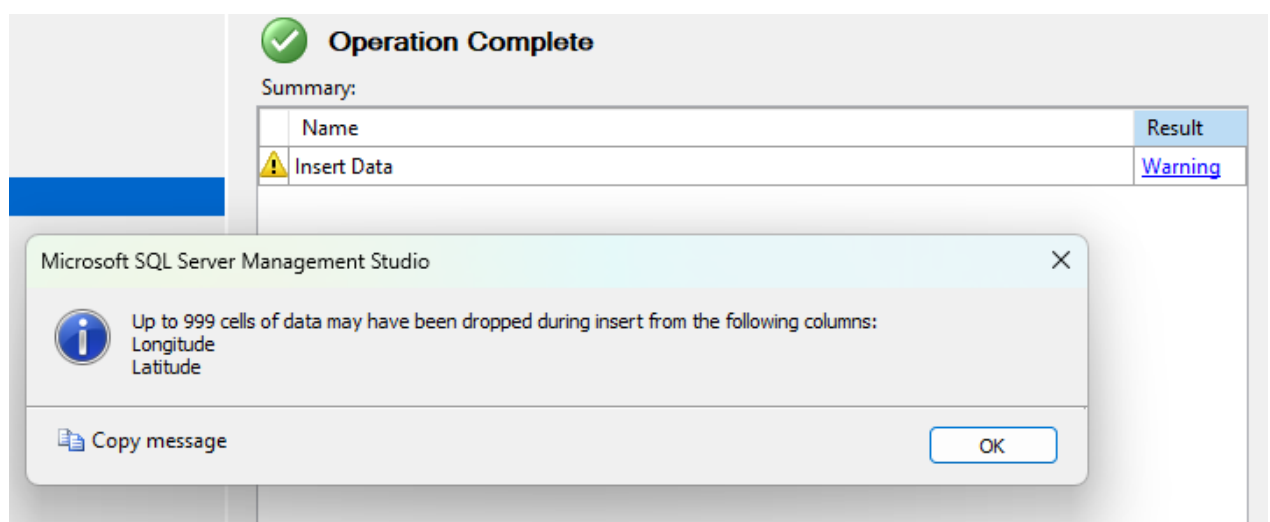
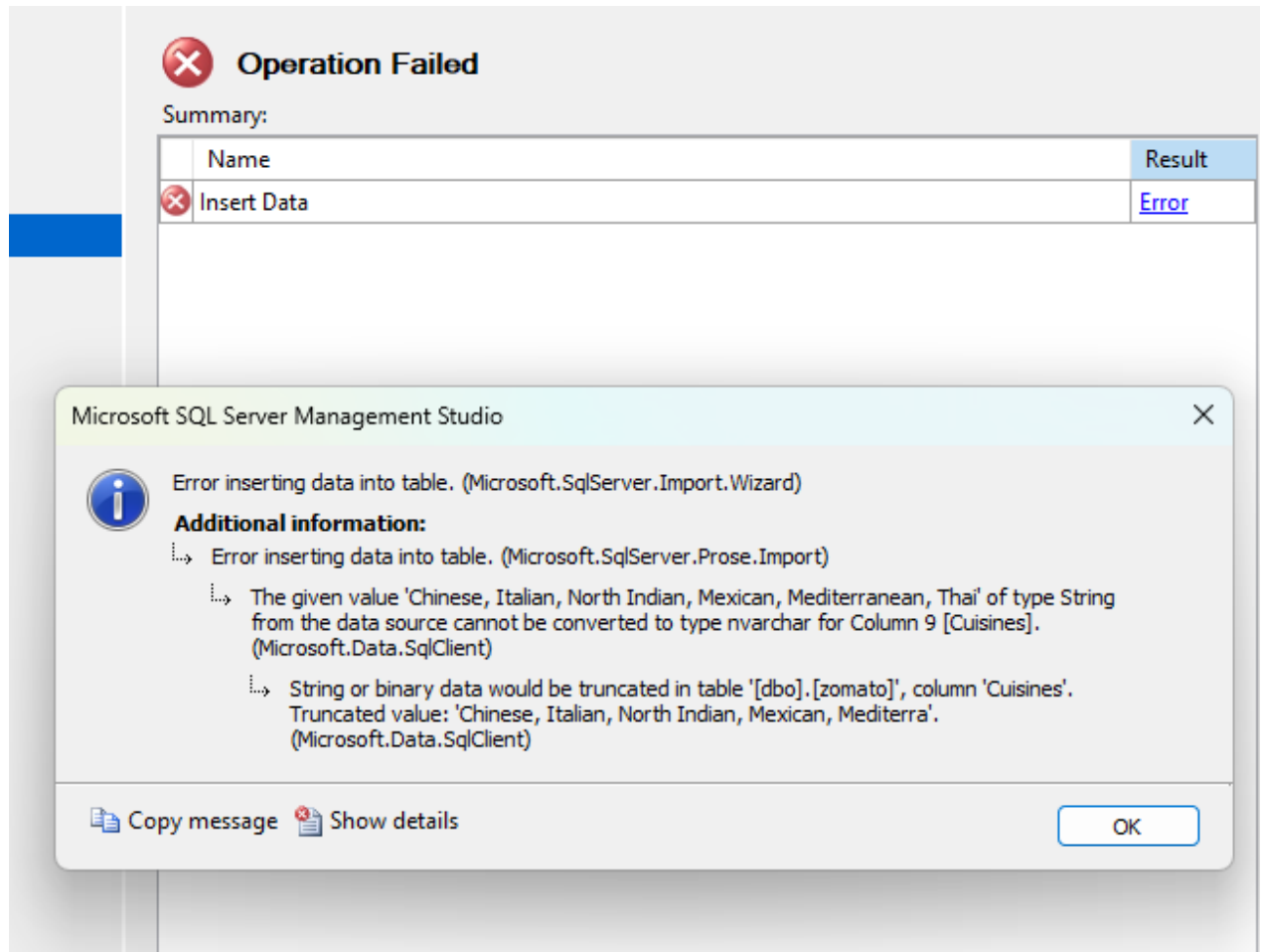
- Es importante buscar data sets que tengan una buena estructura caso contrario se dificultara a la hora de realizar dashboards como nos ocurrió a nosotros.
- Manejar diferentes formatos de datos (JSON, CSV) y estructuras de bases de datos (relacionales y NoSQL) requiere una cuidadosa planificación y herramientas adecuadas si no se realiza correctamente se pierde la integridad de los datos.
- Los dashboards bien diseñados permiten comunicar de manera clara y concisa los resultados del análisis a un público no técnico.

- Recomendaciones:

- Un cronograma detallado y realista es esencial para cumplir con los plazos del proyecto.
- La arquitectura de la solución debe ser diseñada cuidadosamente para garantizar la escalabilidad, la flexibilidad y la eficiencia.

DESAFÍOS Y PROBLEMAS ENCONTRADOS.

- **Importación de CSV a SQL Server:** Durante la importación de un archivo CSV a una base de datos en SQL Server, nos enfrentamos a un error que señalaba la necesidad de ajustar los datos de cada campo de acuerdo con la sintaxis de SQL Server. Este inconveniente requirió realizar modificaciones en el formato de los datos para cumplir con los requisitos del sistema.



- **Integración de Dashboards en Power BI:** La tarea de unificar los dashboards previamente desarrollados resultó compleja debido a la ausencia de una opción gratuita en Power BI que permitiera consolidar múltiples archivos en un solo proyecto. Esta limitación complicó el proceso de integración de todos los dashboards en un único archivo de manera efectiva.



5- Link de github del proyecto

https://github.com/Andrespipe1/Proyecto_Analisis

<https://www.youtube.com/watch?v=X2rgjO88Kj8>

<https://www.youtube.com/watch?v=0iS8tnuOyOE>