



ESCUELA POLITÉCNICA NACIONAL

ESCUELA DE FORMACIÓN DE TECNÓLOGOS



ANÁLISIS DE DATOS

ASIGNATURA:

Análisis de Datos

PROFESOR:

Ing. Lorena Chulde / Ing. Juan Pablo Zaldumbide

PERÍODO ACADÉMICO:

2023-B

TALLER DE REFUERZO

ESTUDIANTES

Chaluisa Silvia

Pinzón Marcelo



	BrandName	MaxWeight(Mg)	WorkingWeight(Mg)
0	CastAluminum	450	430
1	Blue	510	450
2	Green	580	530
3	Red	570	540
4	Standard	550	550
5	ShortHaulTruck_1	1950	1570
6	ShortHaulTruck_2	2260	2110
7	Crane	1300	1360

2023 - B

Objetivo:

Consultar y manipular varios data set mediante bases de datos SQL y NoSQL locales y en la nube para la descripción y análisis de la información de los clientes de todas las bases de la empresa inmobiliaria “El Búho”.

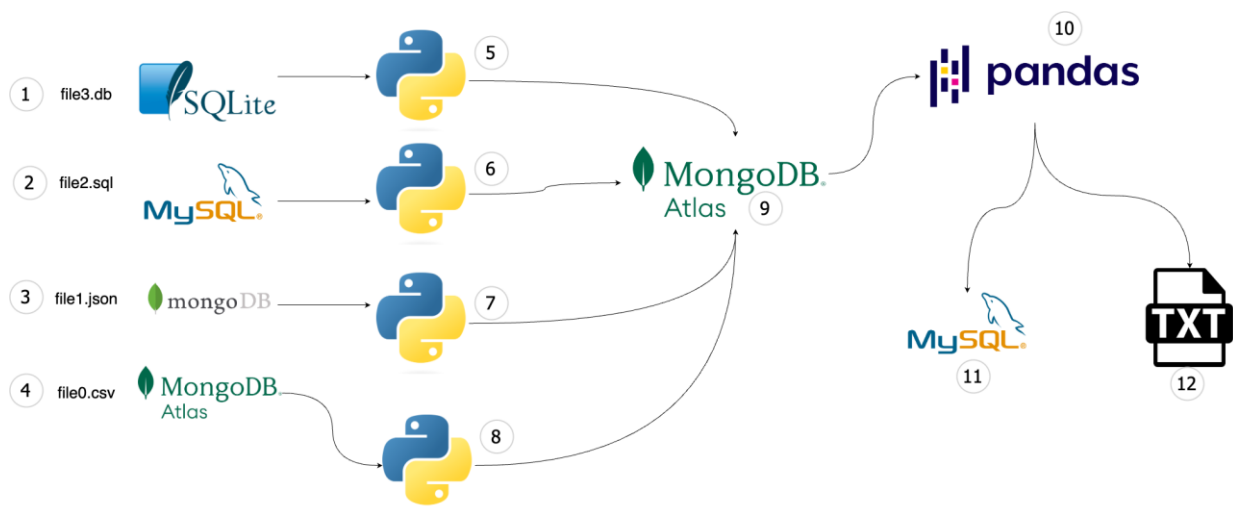
Objetivos específicos:

- Importar y exportar archivos desde varias Bases de Datos
- Integrar los data set de las cuatro sucursales mediante MongoDB Atlas para la accesibilidad, manipulación y análisis de los datos.

Planteamiento del problema:

La empresa inmobiliaria “El Búho” tiene como objetivo averiguar la información de sus clientes, lastimosamente sus datos están divididos en las bases de datos de sus cuatro sucursales.

Dada la siguiente arquitectura:



Indicaciones:

1. SQLite

- Importar file3.db desde la base de datos SQLite.
- Exportar con la extensión que crea conveniente

```
[8]: import sqlite3
import pandas as pd
import os

# Rutas de los archivos
archivo_origen = r'C:\Users\USER\Desktop\TALLER ANALISIS\file3.db'
directorio_destino = r'C:\Users\USER\Desktop\TALLER ANALISIS\Archivos generados'
archivo_destino_csv = 'file3_exportado.csv'

# a. Importar file3.db desde la base de datos SQLite.
conn = sqlite3.connect(archivo_origen)
cursor = conn.cursor()

#-----

# b. Exportar con la extensión CSV
df = pd.read_sql_query("SELECT * FROM file3", conn) # Aquí, estoy haciendo una consulta de ejemplo para exportar toda la tabla 'file3'

# Utilizando pandas para exportar a CSV
ruta_archivo_destino_csv = os.path.join(directorio_destino, archivo_destino_csv)
df.to_csv(ruta_archivo_destino_csv, index=False)

# Cerrar la conexión
conn.close()

# Imprimir mensaje de confirmación
print(f"La exportación CSV se ha guardado en: {ruta_archivo_destino_csv}")
```

La exportación CSV se ha guardado en: C:\Users\USER\Desktop\TALLER ANALISIS\Archivos generados\file3_exportado.csv

Iniciar copia de seguridad > Escritorio > TALLER ANALISIS > Archivos generados

Nombre	Fecha de modificación	Tipo	Tamaño
el_buho.json	28/12/2023 9:06	Archivo de origen ...	5 KB
file3_exportado.csv	28/12/2023 20:47	Archivo de valores...	27 KB
repaso.repaso1.csv	28/12/2023 9:49	Archivo de valores...	51 KB
taller.taller.1.json	28/12/2023 12:47	Archivo de origen ...	89 KB

2. MySQL

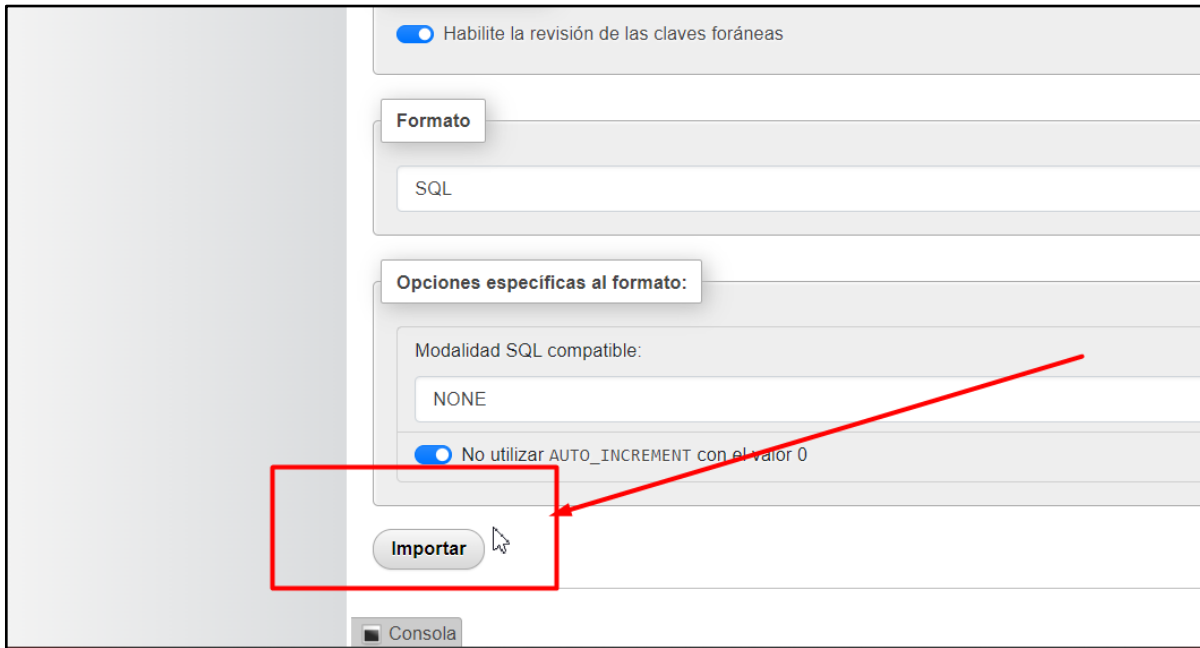
- Importar file.sql desde la base de datos MySQL. (Corregir el script en el caso que tenga errores)

The screenshot shows the XAMPP Control Panel v3.3.0 on the left and the phpMyAdmin interface on the right. In the XAMPP panel, the Apache and MySQL services are highlighted with a red box. The phpMyAdmin interface shows the 'Bases de datos' (Databases) section with a table listing databases like information_schema, mysql, performance_schema, phpmyadmin, and test. The 'Crear base de datos' (Create database) form is visible, and the 'Importar' (Import) button is highlighted with a red box.

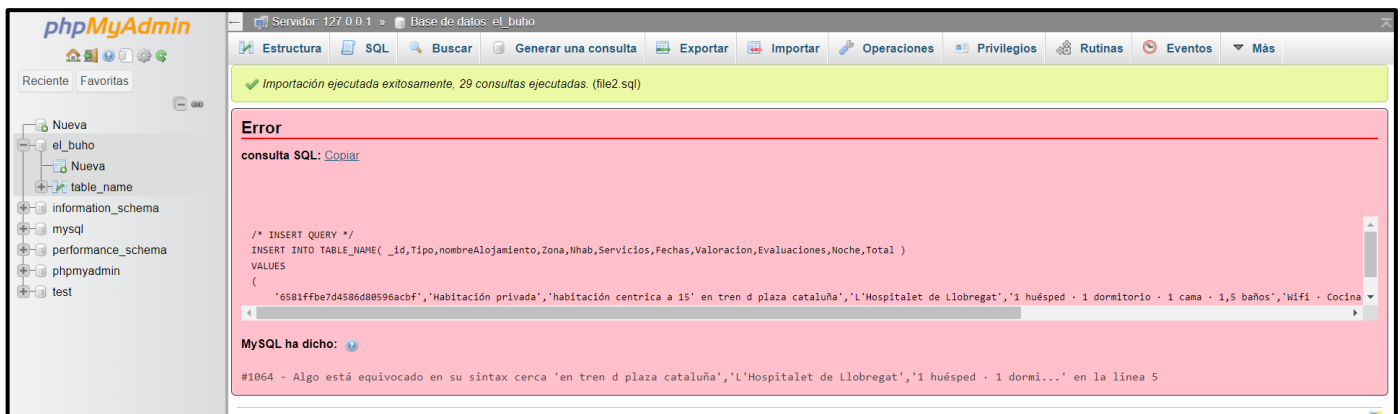
The screenshot shows the phpMyAdmin 'Importando en la base de datos "el_buho"' (Importing into the database "el_buho") screen. The 'Importar' (Import) button is highlighted with a red box. The 'Archivo a importar:' (File to import) section shows a text input field with 'file2.sql' and a 'Seleccionar archivo' (Select file) button, both highlighted with red boxes. The 'Conjunto de caracteres del archivo:' (File character set) dropdown is set to 'utf-8'.

File Explorer view showing the contents of a folder named 'Tareas'.

Nombre	Estado	Fecha de modificación	Tipo	Tamaño
file0.csv	✓	27/12/2023 19:45	Archivo de valores...	51 KB
file1.json	✓	27/12/2023 19:45	Archivo de origen ...	91 KB
file2.sql	✓	27/12/2023 19:45	SQL Text File	85 KB
file3.db	✓	27/12/2023 19:45	Data Base File	280 KB
Taller de repaso.docx	✓	27/12/2023 19:45	Documento de Mi...	511 KB



Aparece error en el script, para lo cual se procede a corregirlo



La selección actual no contiene una columna única. La edición de la grilla y los enlaces de copiado, eliminación y edición no están disponibles.

Mostrando filas 0 - 9 (total de 10, La consulta tardó 0,0003 segundos.)

SELECT * FROM `table_name`

Perfilando [Editar en línea] [Editar] [Explicar SQL] [Crear código PHP] [Actualizar]

Mostrar todo | Número de filas: 25 | Filtrar filas: Buscar en esta tabla

Opciones extra

_id	Tipo	nombreAlojamiento	Zona	Nhab	Servicios	Fechas	Valoracion	Evaluaciones	Noche	Total
6581ffb7d4586d80596acba	Habitación privada	Gran Habitación en el Centro de Barcelona	Eixample	2 huéspedes - 1 dormitorio - 1 cama - 1 baño compa...	Wifi - Cocina - Lavadora	17 dic. - 19 dic.	4,75	(414 evaluaciones)	76 €	151€ en total
6581ffb7d4586d80596acbb	Alojamiento entero - loft	Un espacio con encanto en Barcelona	Districto de Sants-Montjuic	1 huésped - Estudio - 2 camas - 1 baño	Wifi - Cocina - Lavadora - Aire acondicionado	24 dic. - 26 dic.	4,81	(442 evaluaciones)	107 €	213€ en total
6581ffb7d4586d80596acbc	Habitación privada	Twin Room Central + Close to the beach	Sant Marti	2 huéspedes - 1 dormitorio - 2 camas - 1,5 baños	Wifi - Cocina	24 dic. - 26 dic.	4,5	(20 evaluaciones)	48 €	95€ en total
6581ffb7d4586d80596acbd	Habitación privada	Habitación acogedora en el Centro de Barcelona	Ciutat Vella	2 huéspedes - 1 dormitorio - 1 baño compa...	Wifi	10 dic. - 12 dic.	4,31	(26 evaluaciones)	53 €	106€ en total
6581ffb7d4586d80596acbe	Habitación privada	Charming private 1 bedroom w/ balcony in Poble Sec	El Poble Sec	2 huéspedes - 1 dormitorio - 1 cama - 1 baño compa...	Wifi - Cocina - Lavadora	26 nov. - 28 nov.	4,86	(101 evaluaciones)	53 €	106€ en total
6581ffb7d4586d80596acba	Habitación privada	Gran Habitación en el Centro de Barcelona	Eixample	2 huéspedes - 1 dormitorio - 1 cama - 1 baño compa...	Wifi - Cocina - Lavadora	17 dic. - 19 dic.	4,75	(414 evaluaciones)	76 €	151€ en total

b. Exportar con la extensión que crea conveniente.

Exportando tablas de la base de datos "el_buho"

Exportar plantillas:

Nueva plantilla:

Nombre de plantilla: el_buho2 [Crear]

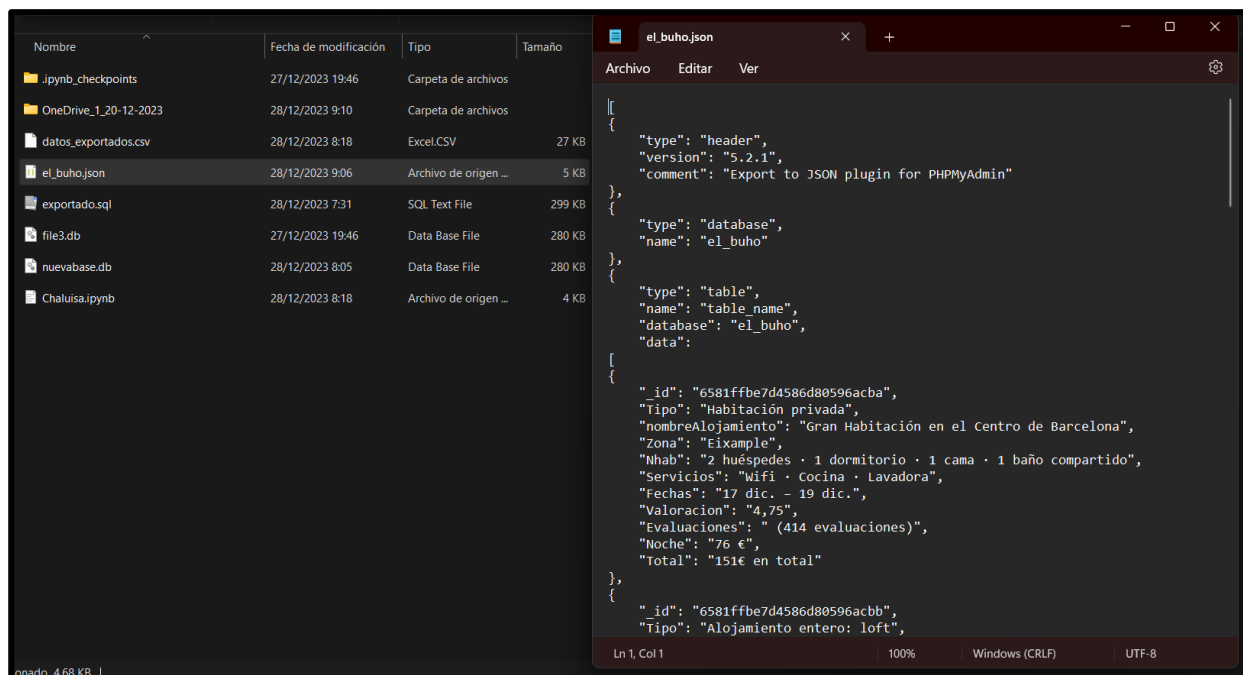
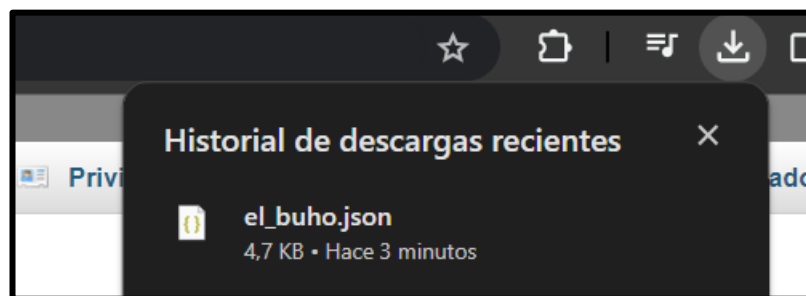
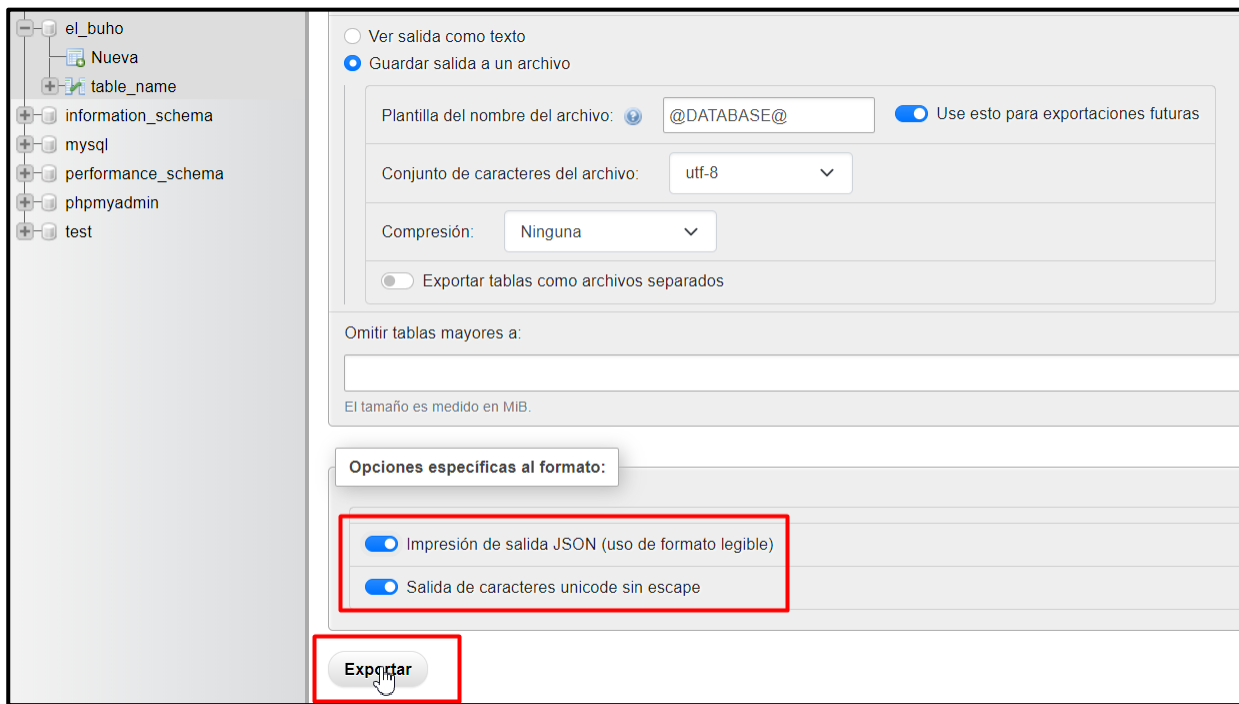
Método de exportación:

☐ Rápido - mostrar sólo el mínimo de opciones de configuración

☒ Personalizado - mostrar todas las opciones de configuración posibles

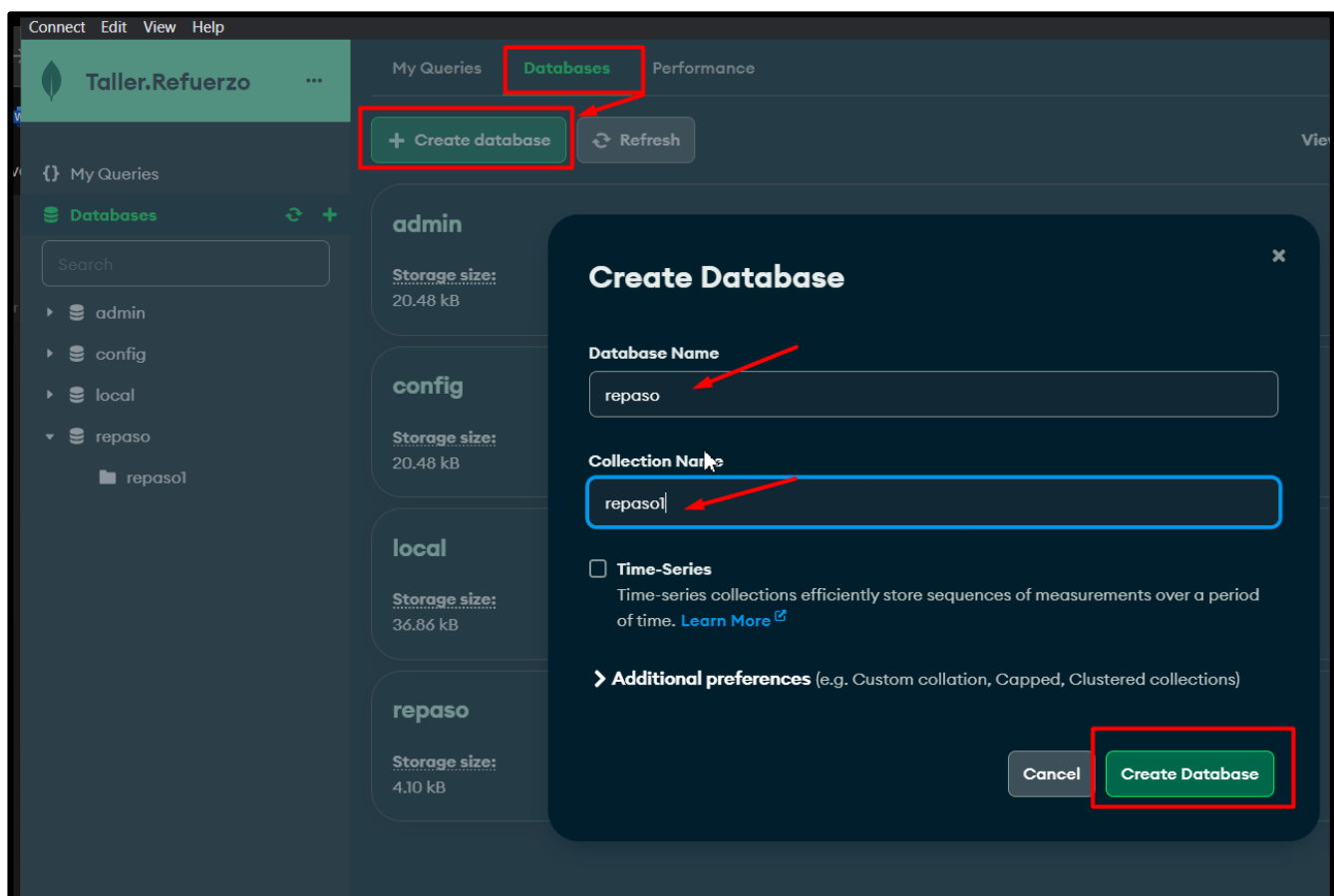
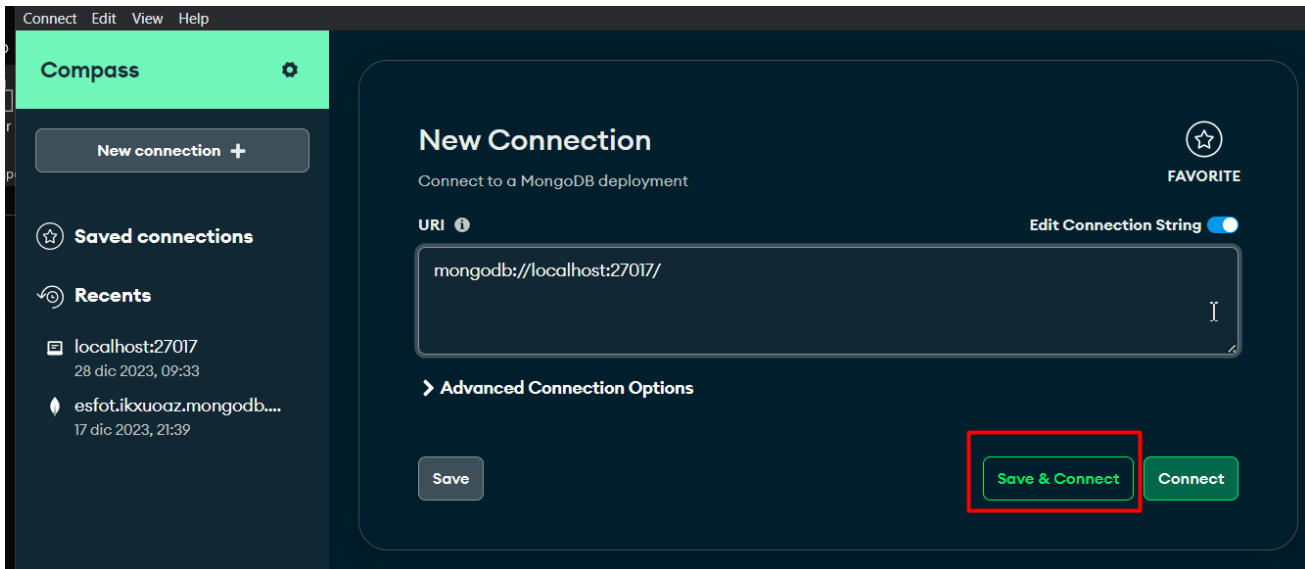
Formato:

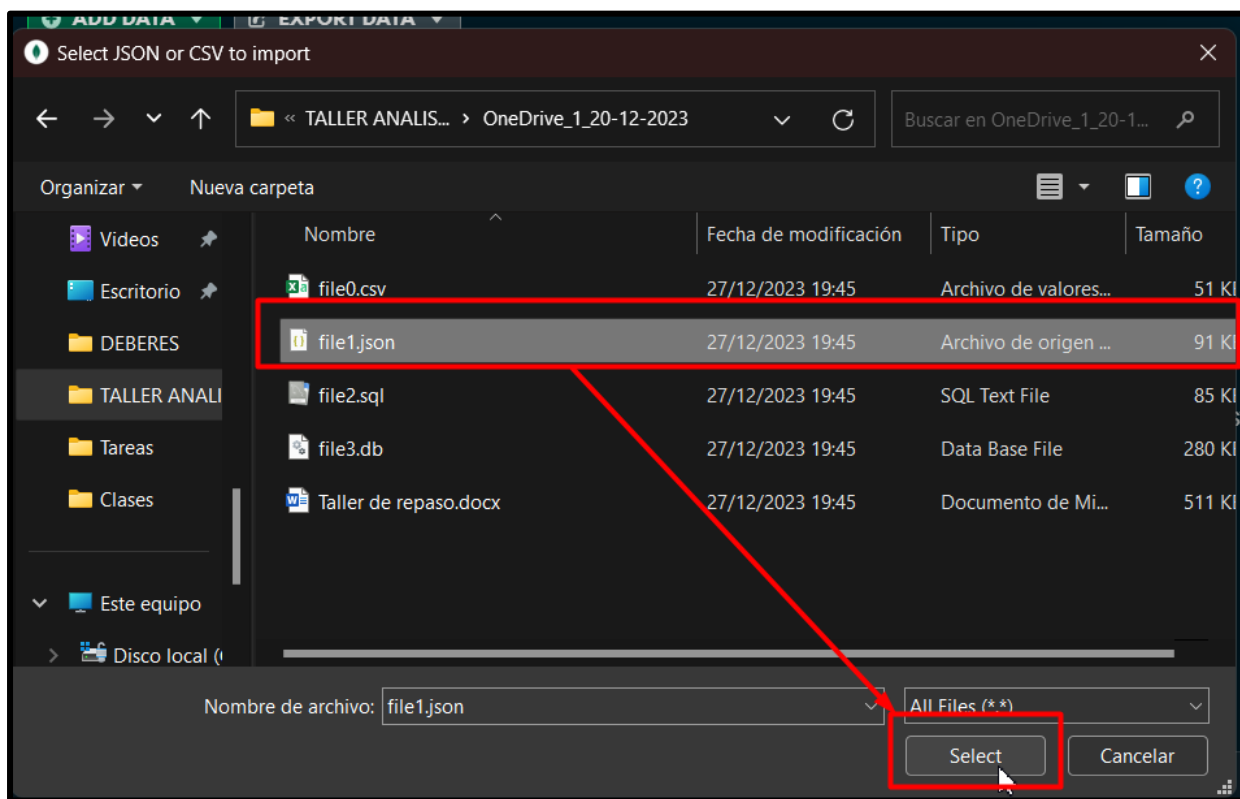
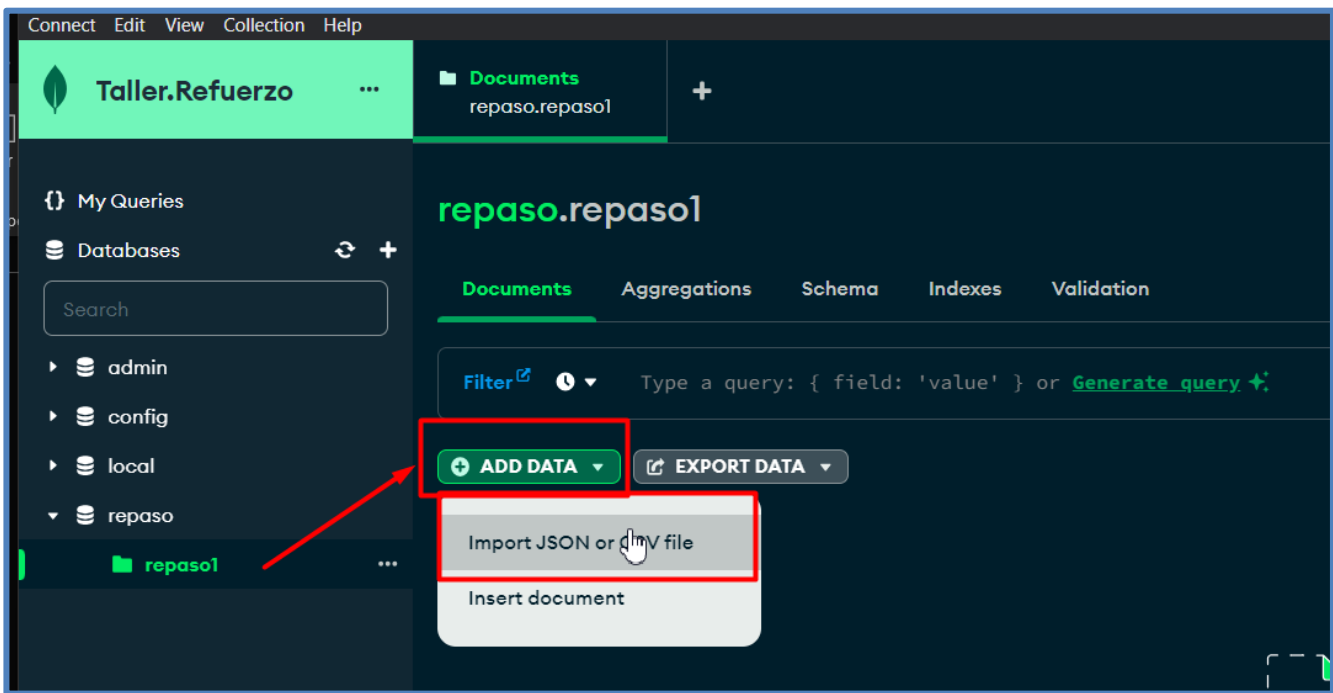
JSON

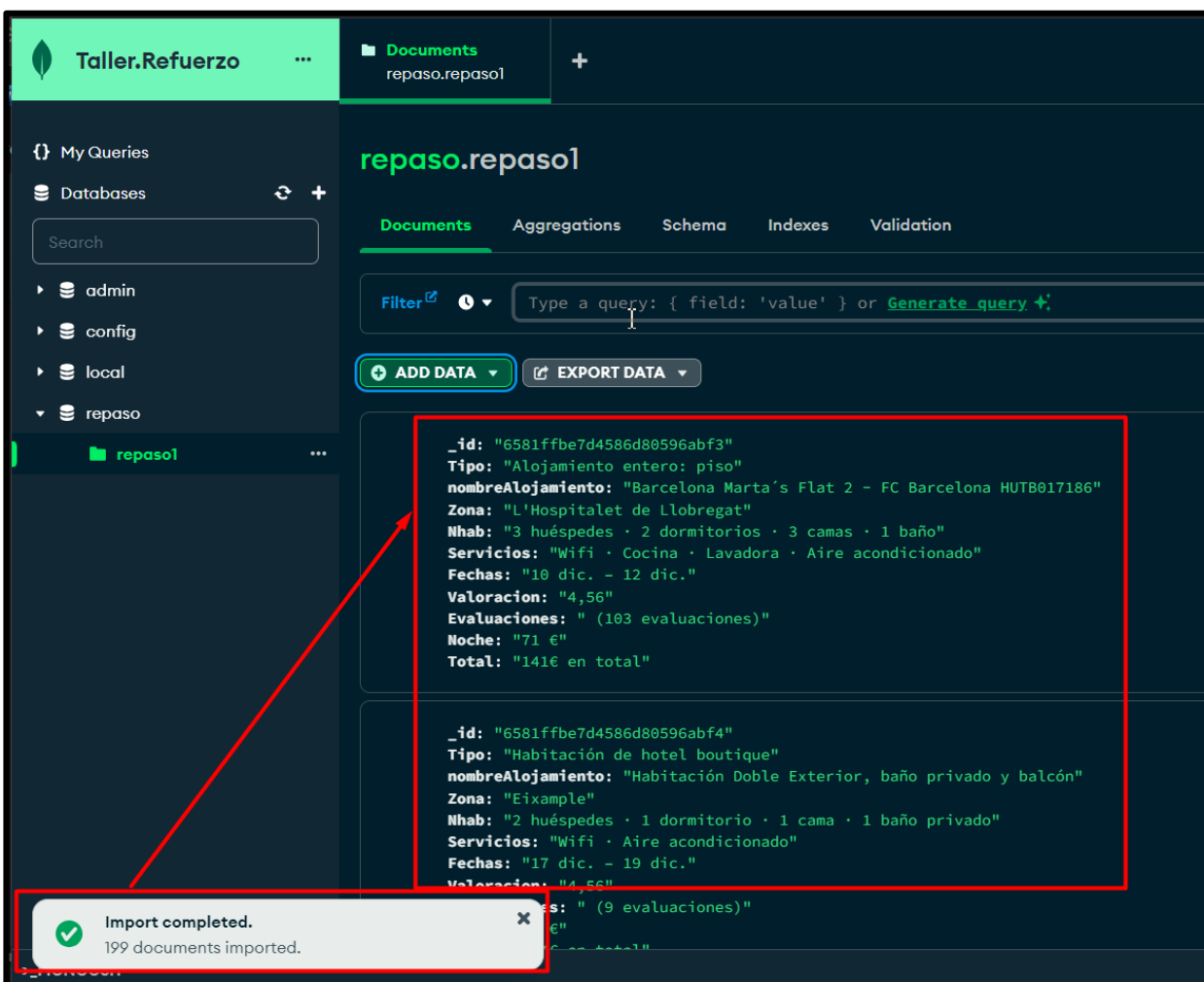
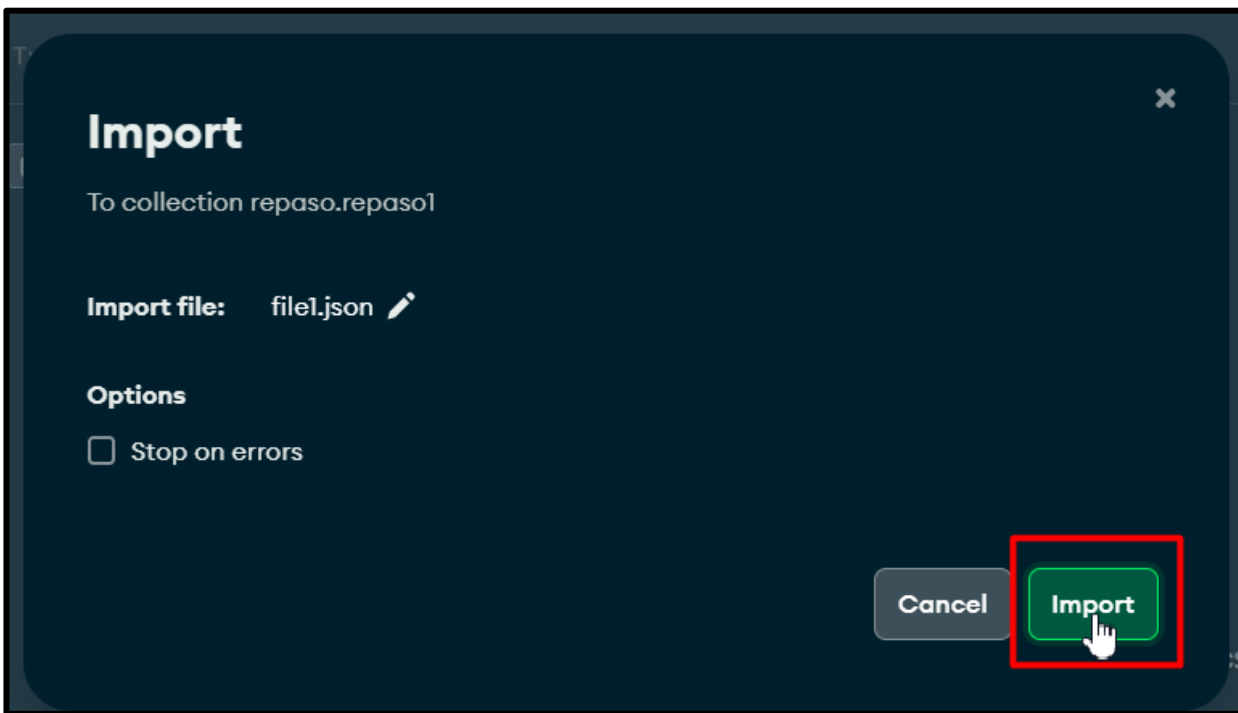


3. MongoDB

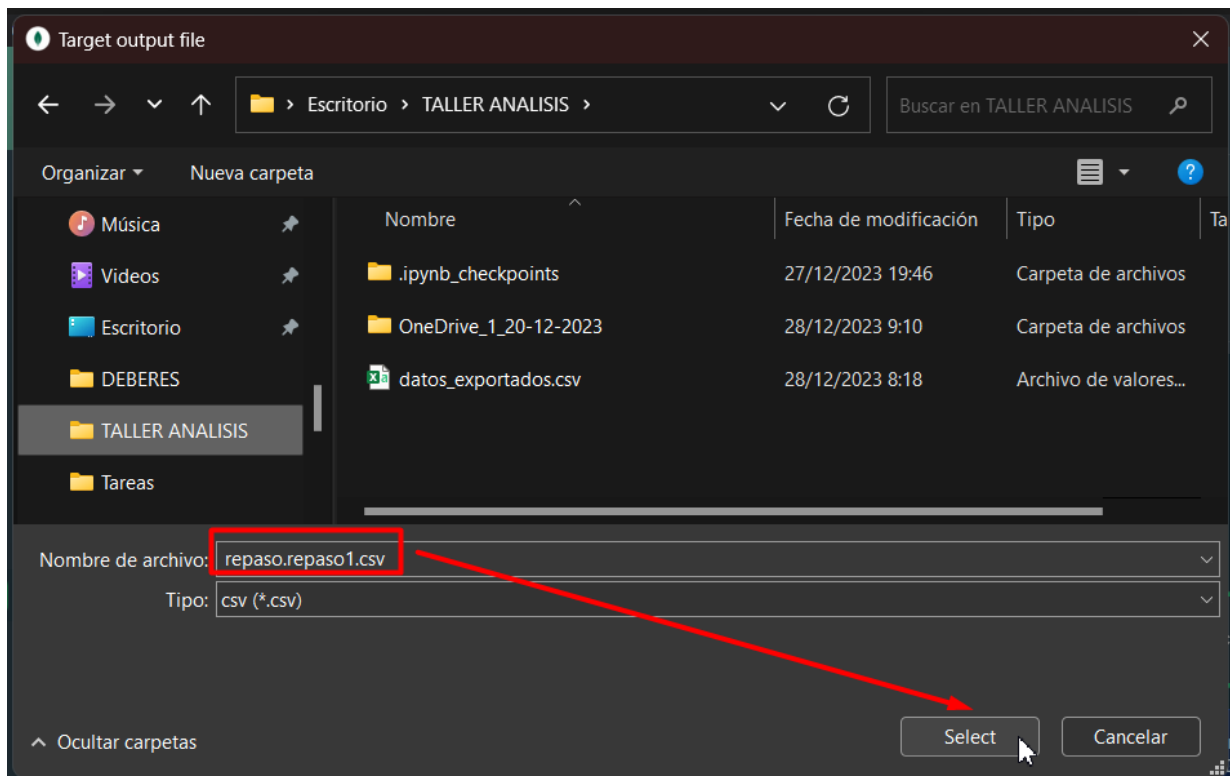
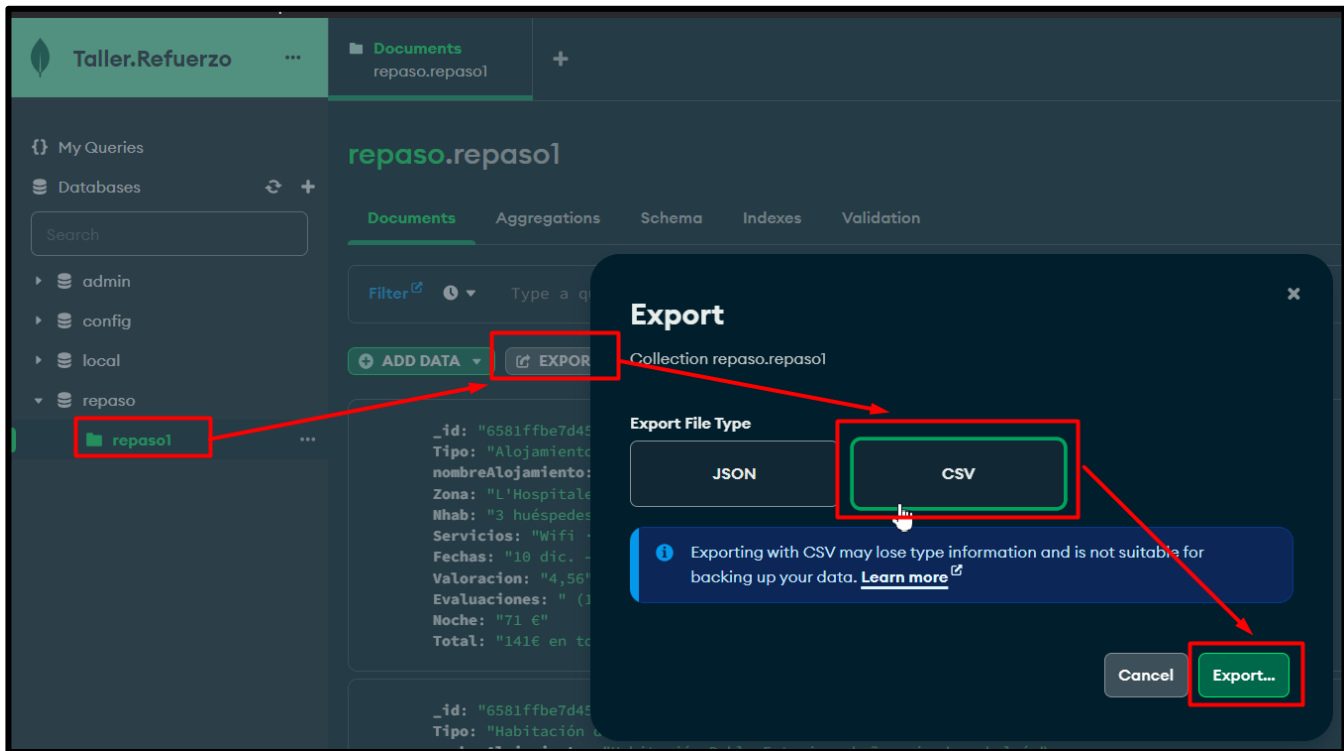
- a. Importar file1.json desde la base de datos MongoDB.

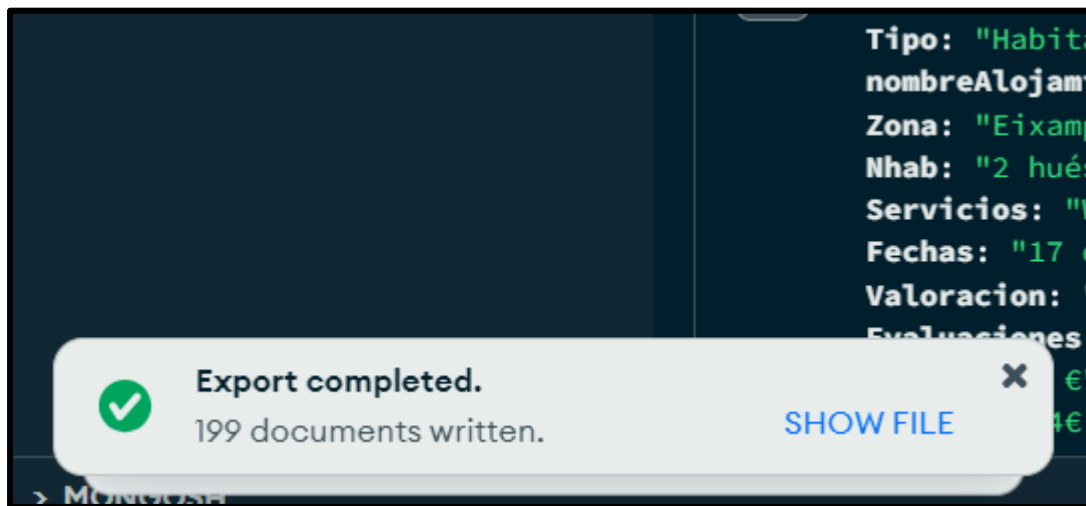






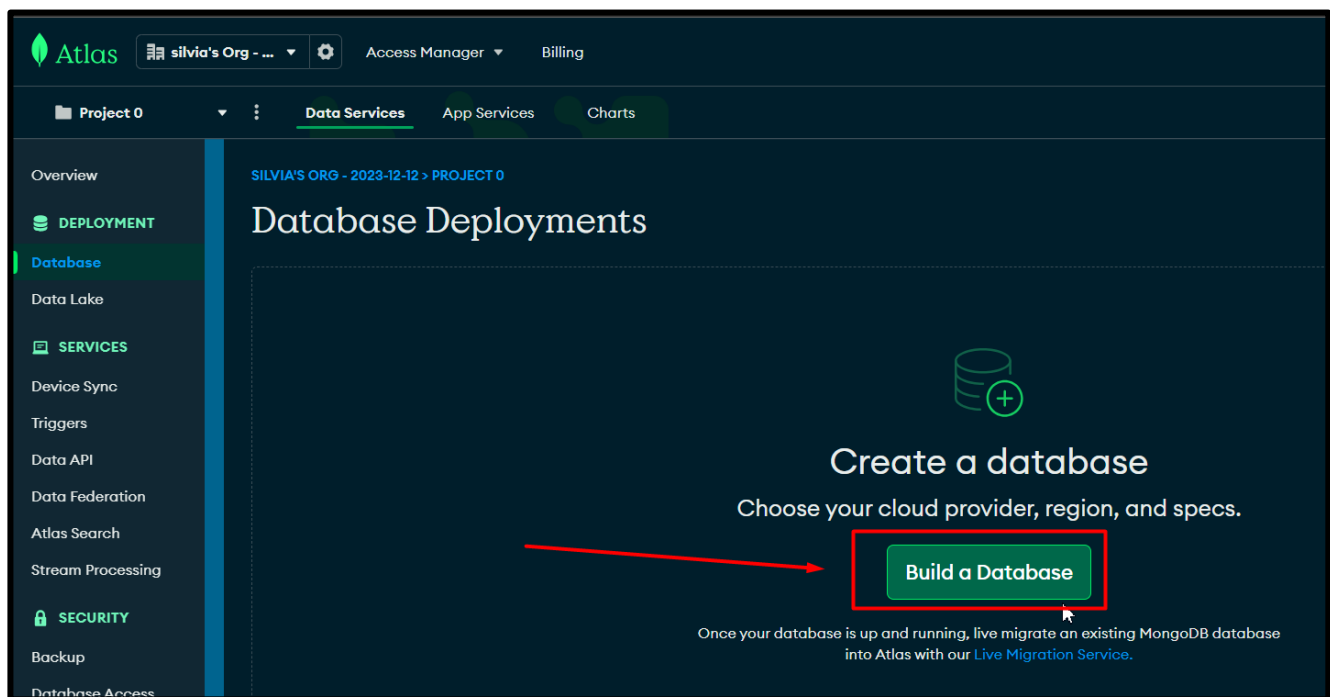
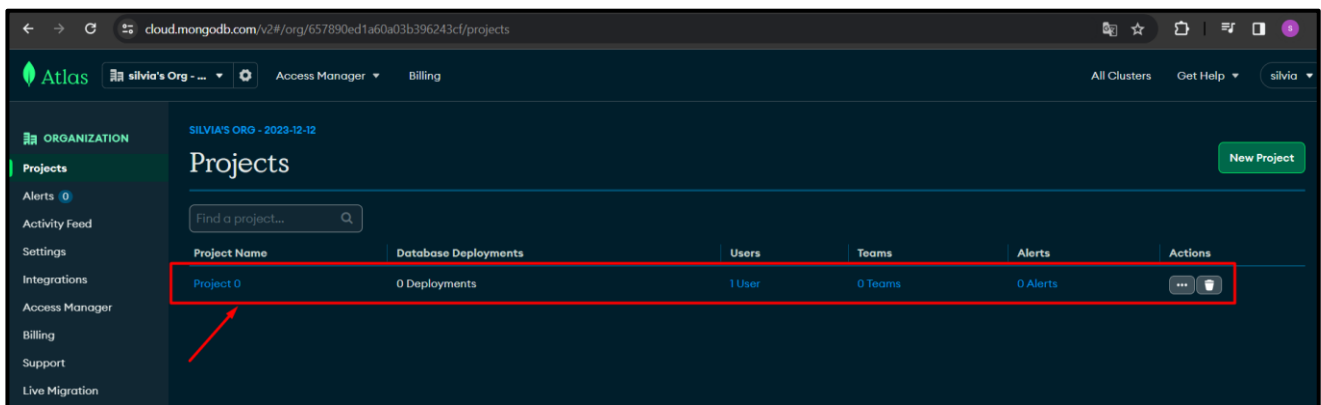
b. Exportar con la extensión que crea conveniente.





4. MongoDB Atlas

- Importar file0.csv desde la base de datos MongoDBAtlas.



M10

\$0.08/hour

For production applications with sophisticated workload requirements.

STORAGE

10 GB

RAM

2 GB

vCPU

2 vCPUs

SERVERLESS

\$0.10/1M reads

For application development and testing, or workloads with variable traffic.

STORAGE

Up to 1 TB

RAM

Auto-scale

vCPU

Auto-scale

M0

FREE

For learning and exploring MongoDB in a cloud environment.

STORAGE

512 MB

RAM

Shared

vCPU

Shared

Provider

aws

Google Cloud

A

Region

🇺🇸

N. Virginia (us-east-1)

★

🌿

★ Recommended ⓘ

🌿 Low carbon emissions ⓘ

Name

You cannot change the name once the cluster is

FREE

Create

Free forever! Your M0 cluster is ideal for experimenting in a limited sandbox. You can upgrade to a production cluster anytime.

[I'll deploy my database later](#)

[Access Advanced Configuration](#)

M10

\$0.08/hour

For production applications with sophisticated workload requirements.

STORAGE

10 GB

RAM

2 GB

vCPU

2 vCPUs

SERVERLESS

\$0.10/1M reads

For application development and testing, or workloads with variable traffic.

STORAGE

Up to 1 TB

RAM

Auto-scale

vCPU

Auto-scale

M0

FREE

For learning and exploring MongoDB in a cloud environment.

STORAGE

512 MB

RAM

Shared

vCPU

Shared

Provider

aws

Google Cloud

A

Region

🇺🇸

N. Virginia (us-east-1)

★

🌿

★ Recommended ⓘ

🌿 Low carbon emissions ⓘ

Name

You cannot change the name once the cluster is

FREE

Create

Free forever! Your M0 cluster is ideal for experimenting in a limited sandbox. You can upgrade to a production cluster anytime.

[I'll deploy my database later](#)

[Access Advanced Configuration](#)

Atlas silvia's Org - ... Access Manager Billing

Project 0 Data Services App Services Charts

Overview

DEPLOYMENT

Database

Data Lake

SERVICES

Device Sync

Triggers

Data API

Data Federation

Atlas Search

Stream Processing

SECURITY

Quickstart

Backup

Database Access

Network Access

Advanced

1 How would you like to authenticate your connection?

Your first user will have permission to read and write any data in your project.

Username and Password Certificate

i We autogenerated a username and password for your first database user in this project using your MongoDB Cloud registration information.

Create a database user using a username and password. Users will be given the *read and write to any database privilege* by default. You can update these permissions and/or create additional users later. Ensure these credentials are different to your MongoDB Cloud username and password.

Username

silviaveronica141998

Password

Silvia123

Autogenerate Secure Password Copy

Create User

Atlas silvia's Org - ... Access Manager Billing

Project 0 Data Services App Services Charts

Overview

DEPLOYMENT

Database

Data Lake

SERVICES

Device Sync

Triggers

Data API

Data Federation

Atlas Search

Stream Processing

SECURITY

Backup

Database Access

Network Access

Advanced

My Local Environment

Use this to add network IP addresses to the IP Access List. This can be modified at any time.

Cloud Environment

Use this to configure network access between Atlas and your cloud or on-premise environment. Specifically, set up IP Access Lists, Network Peering, and Private Endpoints.

Set your network security with any of the following options

Only an IP address you add to your Access List will be able to connect to your project's clusters. You can manage existing IP entries via the [Network Access Page](#).

IP Address	Description
192.168.1.6	Silvia

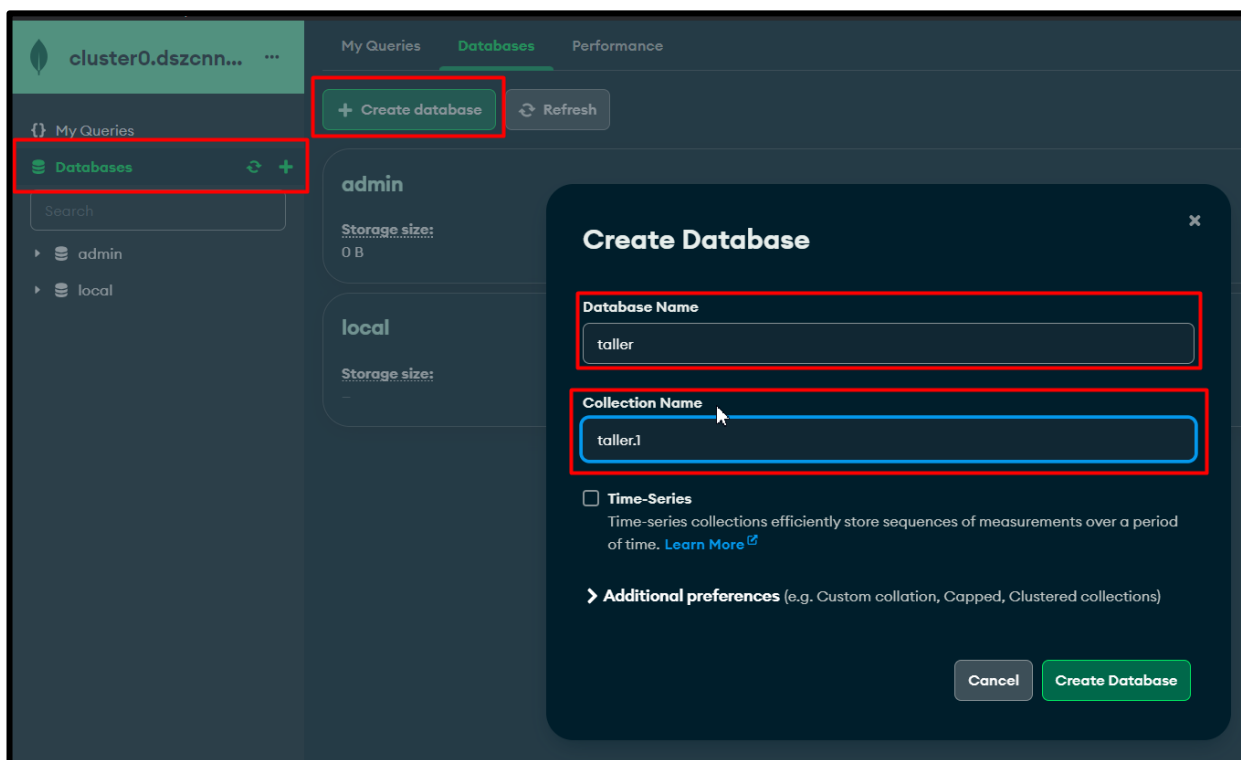
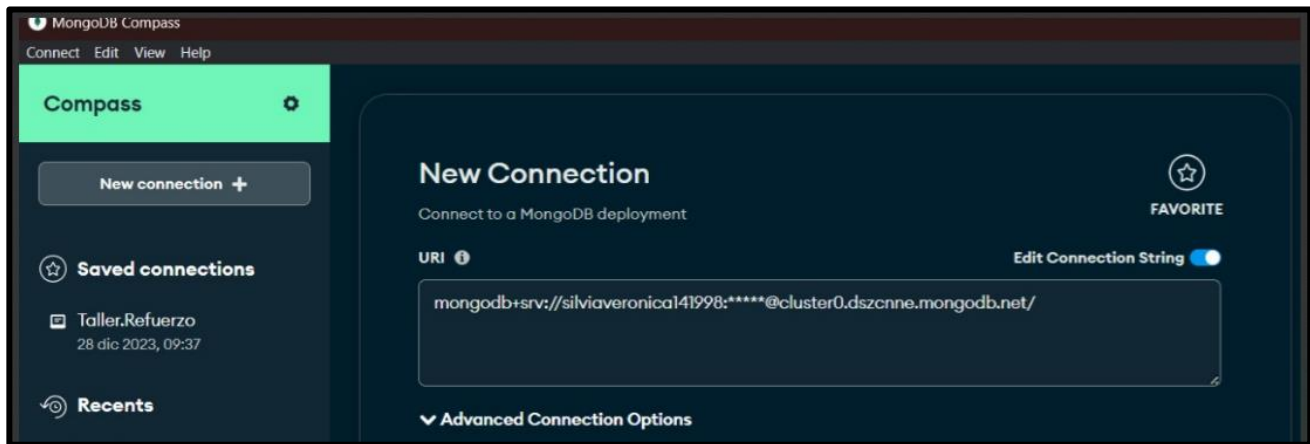
Add My Current IP Address

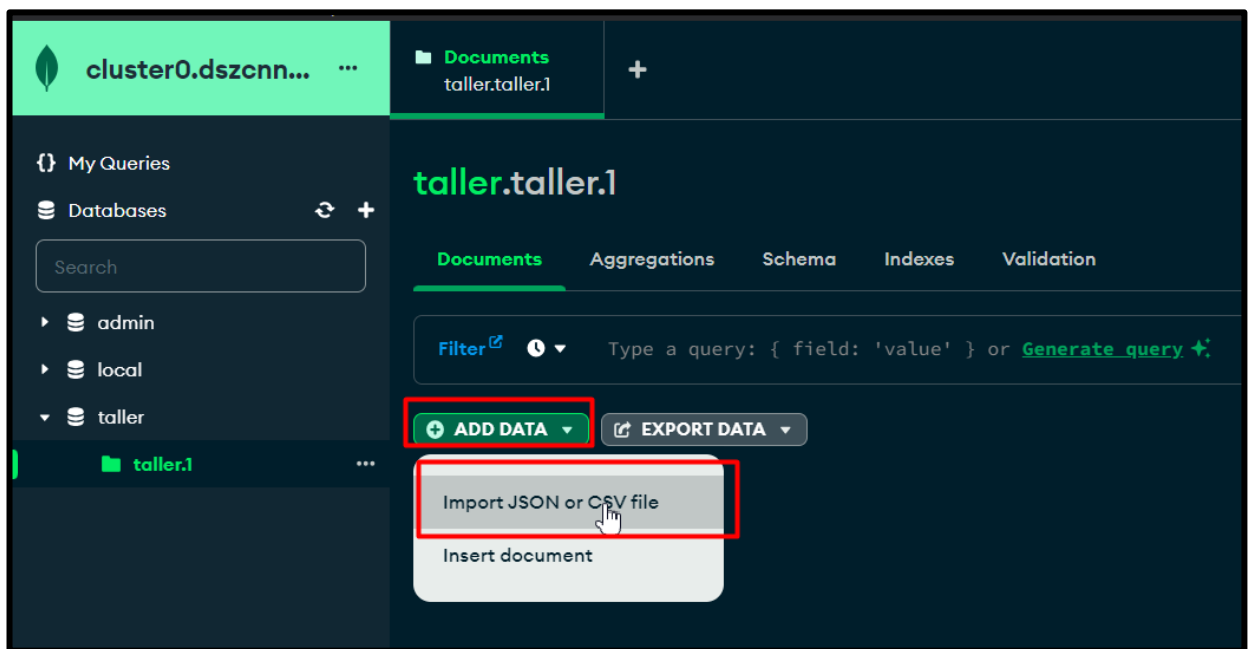
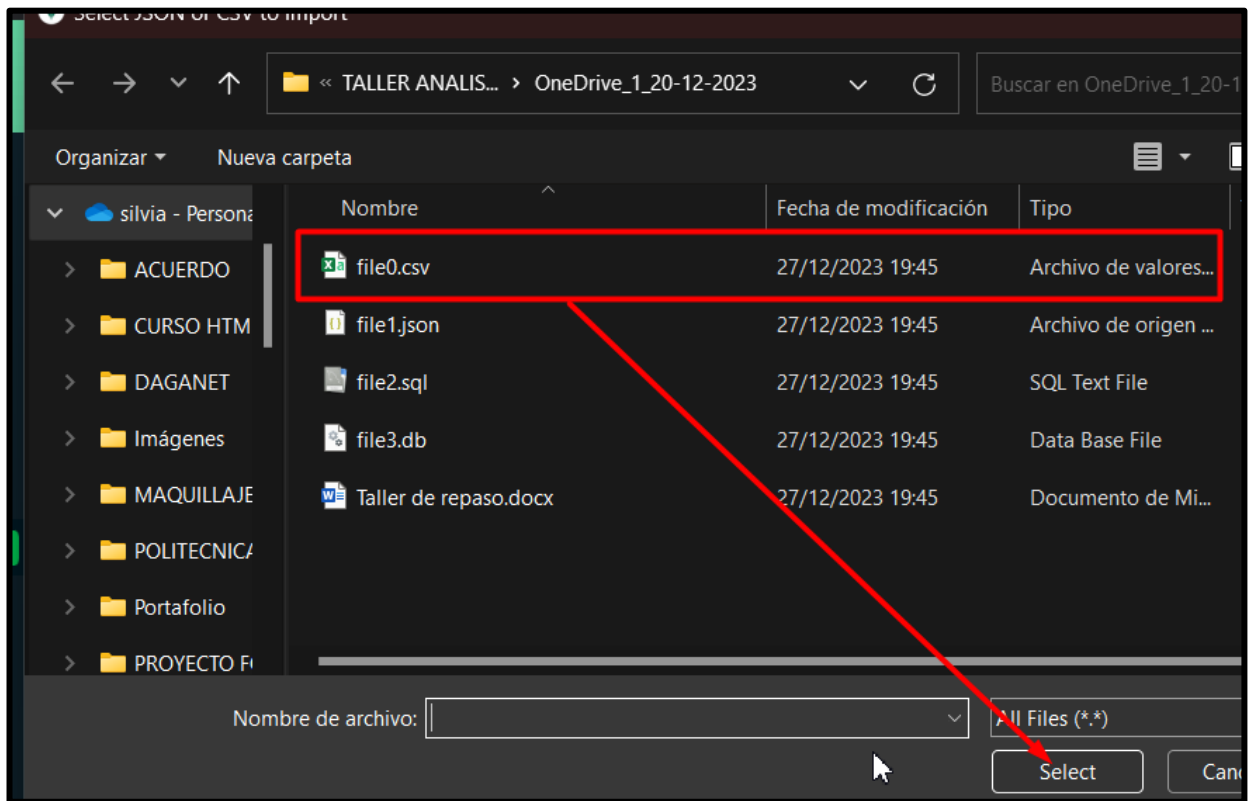
Add Entry

IP Access List	Description
190.152.177.7/32	My IP Address

EDIT REMOVE

Conectarse a la base de datos Mongo Atlas





Import

To collection taller.taller.1

Import file: file0.csv

Options

Select delimiter: Comma

☒ Ignore empty strings

☐ Stop on errors

Specify Fields and Types [Learn more about data types](#)

	<input checked="" type="checkbox"/> _id ObjectId	<input checked="" type="checkbox"/> Tipo String	<input checked="" type="checkbox"/> nombreAlojamiento String	<input checked="" type="checkbox"/> Zona String
1	6581ffbe7d4586d80596ab2c	Habitación privada	Habitacion centro Barcelona, Ciutat Vell...	Ciutat Vella
2	6581ffbe7d4586d80596ab2d	Habitación privada	Agradable habitación en pleno centro de...	La Font de la Guatlla
3	6581ffbe7d4586d80596ab2e	Habitación privada	Room in the city center (Gothic,Rambla,C...	Ciutat Vella
4	6581ffbe7d4586d80596ab2f	Habitación privada	Comfortable single room	Eixample
5	6581ffbe7d4586d80596ab30	Alojamiento entero: piso	Outstanding apartment &Amazing locati...	Ciutat Vella
6	6581ffbe7d4586d80596ab31	Habitación privada	Habitación en Gracia cerca del Park Güell	El Camp d'en Graciet i Gracia

Cancel Import

cluster0.dszone.mongodb.net

Documents taller.taller.1

taller.taller.1

Documents Aggregations Schema Indexes Validation

Filter Type a query: { field: 'value' } or [Generate query](#)

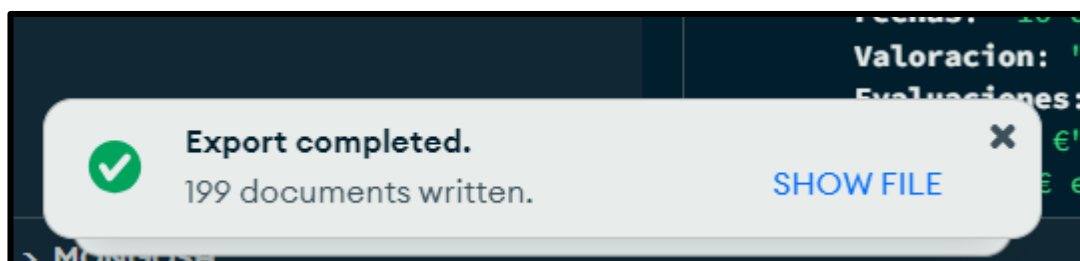
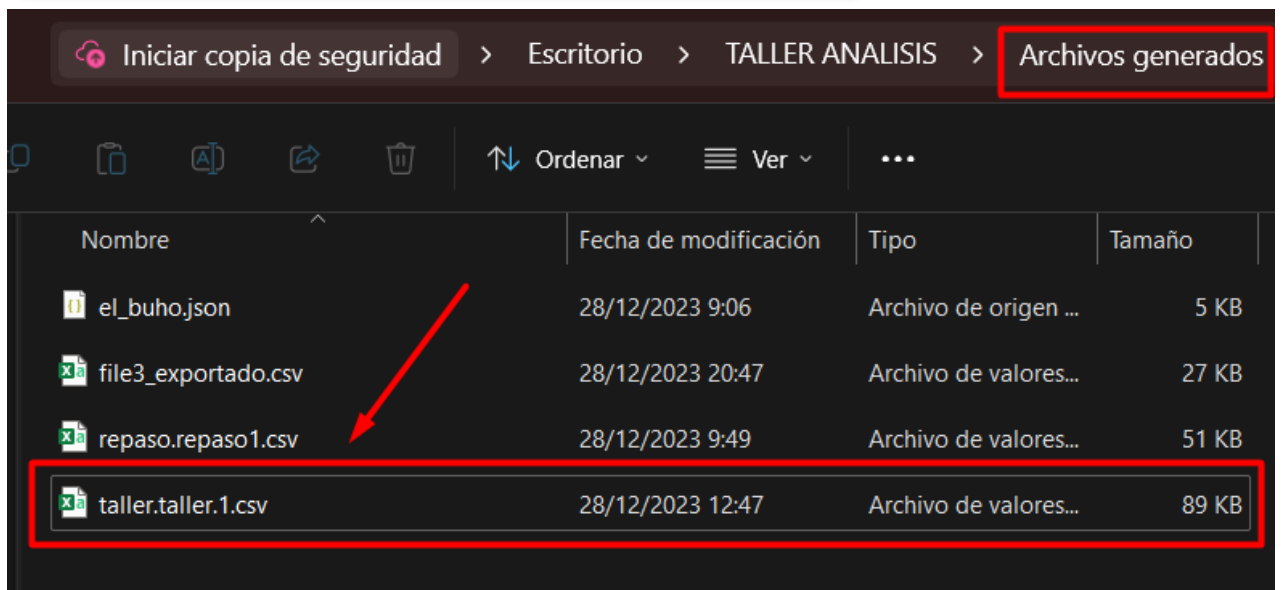
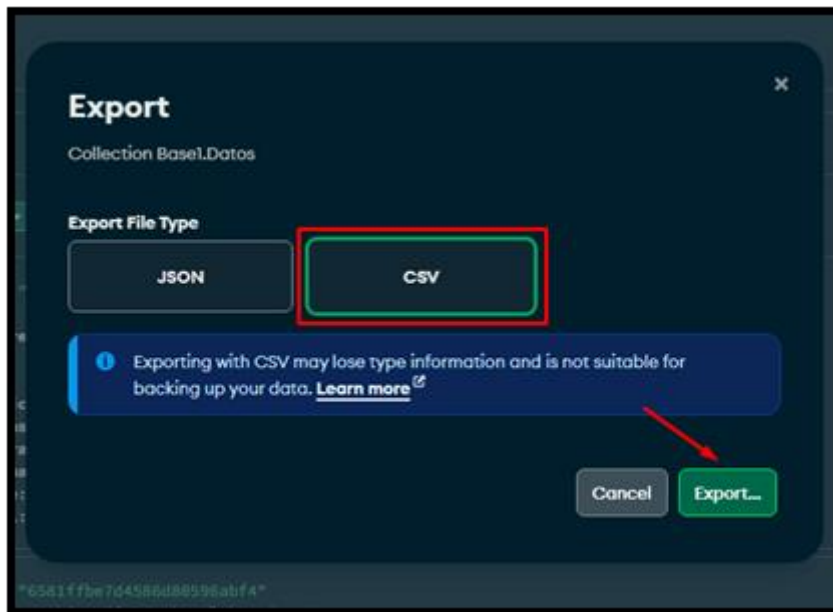
ADD DATA EXPORT DATA

```
_id: ObjectId('6581ffbe7d4586d80596ab2c')
Tipo: "Habitación privada"
nombreAlojamiento: "Habitacion centro Barcelona, Ciutat Vella room."
Zona: "Ciutat Vella"
Nhab: "2 huéspedes · 1 dormitorio · 1 cama · 1 baño compartido"
Servicios: "Wifi · Cocina"
Fechas: "17 dic. - 19 dic."
Valoracion: "4,6"
Evaluaciones: " (25 evaluaciones)"
Noche: "24 €"
Total: "47€ en total"
```

```
_id: ObjectId('6581ffbe7d4586d80596ab2d')
Tipo: "Habitación privada"
nombreAlojamiento: "Agradable habitación en pleno centro de Barcelona"
Zona: "La Font de la Guatlla"
Nhab: "1 huésped · 1 dormitorio · 1 cama · 1,5 baños compartidos"
Servicios: "Wifi · Cocina · Aire acondicionado"
Fechas: "10 dic. - 12 dic."
Valoracion: "4,57"
Evaluaciones: " (449 evaluaciones)"
Noche: "24 €"
Total: "47€ en total"
```

Import completed.
199 documents imported.

- b. Exportar con la extensión que crea conveniente.



5. Leer file3 desde Jupyter con Python

Jupyter Chaluisa Last Checkpoint: 1 hour ago

File Edit View Run Kernel Git Settings Help Trusted

JupyterLab Python 3 (ipykernel)

```
[12]: # Ruta del archivo CSV
ruta_archivo_csv = r'C:\Users\USER\Desktop\TALLER ANALISIS\Archivos generados\file3_exportado.csv'

# Leer el archivo CSV en un DataFrame
df = pd.read_csv(ruta_archivo_csv)
df
```

	field1	field2	field3	field4	field5	field6	field7	field8	field9	field10	field11
0	_id	Tipo	nombreAlojamiento	Zona	Nhab	Servicios	Fechas	Valoracion	Evaluaciones	Noche	Total
1	6581ffbe7d4586d80596ad81	Habitación privada	Beautiful and sunny apartment at Eixample!	El Fort Pienc	2 huéspedes · 1 dormitorio · 1 cama · 1 baño	Wifi · Cocina · Lavadora	17 dic. – 19 dic.	NaN	No hay evaluaciones	58 €	116€ en total
2	6581ffbe7d4586d80596ad82	Habitación privada	Room near the center	Distrito de Sants-Montjuïc	2 huéspedes · 1 dormitorio · 1 cama · 1 baño	Wifi	17 dic. – 19 dic.	4,21	(185 evaluaciones)	36 €	71€ en total
3	6581ffbe7d4586d80596ad83	Alojamiento entero: piso	Luminoso y coqueto ático. - License HUTB-016889	L'Hospitalet de Llobregat	4 huéspedes · 2 dormitorios · 4 camas · 1 baño	Tu anfitrión es una empresa · Wifi · Cocina · ...	26 nov. – 28 nov.	4,46	(216 evaluaciones)	93 €	185€ en total
4	6581ffbe7d4586d80596ad84	Habitación privada	Habitación en Gracia cerca del Park Güell	El Poble Sec	2 huéspedes · 1 dormitorio · 1 cama · 1 baño	Wifi · Cocina · Lavadora · Aire acondicionado	10 dic. – 12 dic.	4,78	(201 evaluaciones)	29 €	57€ en total
...
100	6581ffbe7d4586d80596ade4	Habitación privada	Habitación acogedora en el Centro de	Hostafrans	2 huéspedes · 1 dormitorio · 1 cama · 1 baño	Tu anfitrión es una empresa · Wifi · Cocina · ...	10 dic. – 12 dic.	4,73	(213 evaluaciones)	98 €	195€ en total

6. Leer file2 desde Jupyter con Python

7. Leer file1 desde Jupyter con Python

```
[15]: df2=pd.read_json('C:/Users/USER/Desktop/TALLER ANALISIS/file1.json')
df2
```

	_id	Tipo	nombreAlojamiento	Zona	Nhab	Servicios	Fechas	Valoracion	Evaluaciones	Noche	Total
0	6581ffbe7d4586d80596abf3	Alojamiento entero: piso	Barcelona Marta's Flat 2 - FC Barcelona HUTB01...	L'Hospitalet de Llobregat	3 huéspedes · 2 dormitorios · 3 camas · 1 baño	Wifi · Cocina · Lavadora · Aire acondicionado	10 dic. – 12 dic.	4,56	(103 evaluaciones)	71 €	141€ en total
1	6581ffbe7d4586d80596abf4	Habitación de hotel boutique	Habitación Doble Exterior, baño privado y balcón	Eixample	2 huéspedes · 1 dormitorio · 1 cama · 1 baño p...	Wifi · Aire acondicionado	17 dic. – 19 dic.	4,56	(9 evaluaciones)	87 €	174€ en total
2	6581ffbe7d4586d80596abf5	Habitación privada	Luxury room with private bathroom 2	Sant Pere-Santa Caterina i la Ribera	3 huéspedes · 1 dormitorio · 1 baño privado	Wifi · Cocina · Lavadora	10 dic. – 12 dic.	4,53	(15 evaluaciones)	76 €	151€ en total
3	6581ffbe7d4586d80596abf6	Habitación privada	Hostal Ola Living Diagonal 11	Sarrià-Sant Gervasi	2 huéspedes · Estudio · 1 baño	Wifi · Aire acondicionado	3 dic. – 5 dic.	4,5	(4 evaluaciones)	74 €	147€ en total
4	6581ffbe7d4586d80596abf7	Alojamiento entero: piso	Apartamento SAGRADA FAMILIA - free WIFI - NETFLIX	Sant Marti	5 huéspedes · 2 dormitorios · 2 camas · 1 baño	Wifi · Cocina · Lavadora · Aire acondicionado	17 dic. – 19 dic.	4,23	(344 evaluaciones)	110 €	220€ en total
...
194	6581ffbe7d4586d80596acb5	Alojamiento entero: piso	Precioso Estudio en Barcelona *Bcn* E04	Eixample	3 huéspedes · 1 dormitorio · 1 cama · 1 baño	Tu anfitrión es una empresa · Wifi · Cocina · ...	10 dic. – 12 dic.	4,62	(63 evaluaciones)	103 €	206€ en total

8. Leer file0 desde Jupyter con Python

```
[16]: df3=pd.read_csv('C:/Users/USER/Desktop/TALLER ANALISIS/file0.csv')
df3
```

	_id	Tipo	nombreAlojamiento	Zona	Nhab	Servicios	Fechas	Valoracion	Evaluaciones	Noche	Total
0	6581ffbe7d4586d80596ab2c	Habitación privada	Habitacion centro Barcelona, Ciutat Vella room.	Ciutat Vella	2 huéspedes · 1 dormitorio · 1 cama · 1 baño c...	Wifi · Cocina	17 dic. – 19 dic.	4,6	(25 evaluaciones)	24 €	47€ en total
1	6581ffbe7d4586d80596ab2d	Habitación privada	Agradable habitación en pleno centro de Barcelona	La Font de la Guatlla	1 huésped · 1 dormitorio · 1 cama · 1,5 baños ...	Wifi · Cocina · Aire acondicionado	10 dic. – 12 dic.	4,57	(449 evaluaciones)	29 €	57€ en total
2	6581ffbe7d4586d80596ab2e	Habitación privada	Room in the city center (Gothic,Rambla,Catedral)	Ciutat Vella	1 huésped · 1 dormitorio · 1 cama · 1 baño com...	Wifi · Cocina · Lavadora	24 dic. – 26 dic.	4,48	(120 evaluaciones)	25 €	50€ en total
3	6581ffbe7d4586d80596ab2f	Habitación privada	Comfortable single room	Eixample	1 huésped · 1 dormitorio · 1 cama · 1 baño com...	Wifi · Cocina	31 dic. – 2 ene.	4,57	(35 evaluaciones)	37 €	73€ en total
4	6581ffbe7d4586d80596ab30	Alojamiento entero: piso	Outstanding apartment &Amazing location :)	Ciutat Vella	2 huéspedes · 1 dormitorio · 1 cama · 1 baño	Wifi · Cocina · Aire acondicionado	17 dic. – 19 dic.	4,77	(35 evaluaciones)	85 €	170€ en total
...
					3 huéspedes						

9. Integrar los cuatro data set y enviar a MongoDB Atlas (puede usar concat)

10. Importar el archivo desde MongoDB Atlas hasta Jupyter haciendo uso de las librerías numpy y pandas y realizar lo siguiente:
 - a. Verificar el tiempo que se demora en cargar el data set.
 - b. Verificar el número de registros.
 - c. Verificar si hay registros duplicados, en el caso de haberlos, deberá eliminar los duplicados.
 - d. Verificar si hay celdas nulas
 - e. En caso de haber celdas nulas, se deberá sacar el promedio y completar la información con la media.
 - f. Describir el data set
 - g. Describir usando campos definidos, mostrando 10 filas.
 - h. Indique el alojamiento más costoso por noche.
 - i. Indique el total cancelado más costoso por un cliente.
 - j. Indique el tipo de alojamiento que hay para 6 personas.
 - k. Verifique cuál es la valoración más alta y cuál es la valoración más baja de los tipos d alojamientos.
11. Exportar a MySQL el archivo unificado.
12. Además, exporte el archivo unificado con extensión .txt.

PRESENTACIÓN

Al finalizar tu práctica deberás subir:

1.- Código fuente <apellido.py> o <apellido.ipynb>

2.- pdf con capturas de pantalla de la ejecución de cada ejercicio.

3.- Archivos .txt generados

Poner todo en un archivo comprimido con el nombre: (Taller-Refuerzo-AnalisisDatos-NApellido) y entregarlo en la tarea indicada.

RECURSOS NECESARIOS

- Material de clase
- Internet
- MySQL
- MongoDB
- MongoDB Atlas
- SQLite
- Json, CSV
- Jupyter
- Pandas, numpy