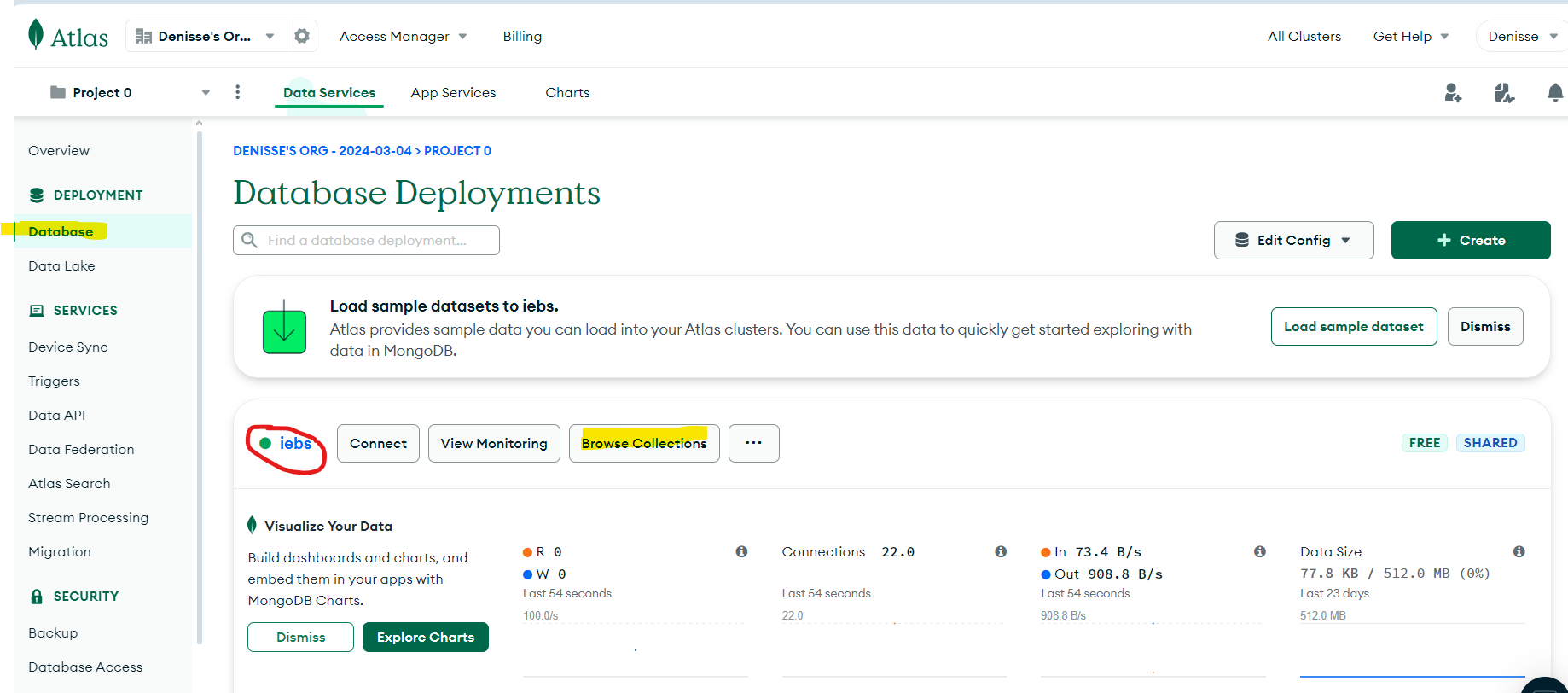
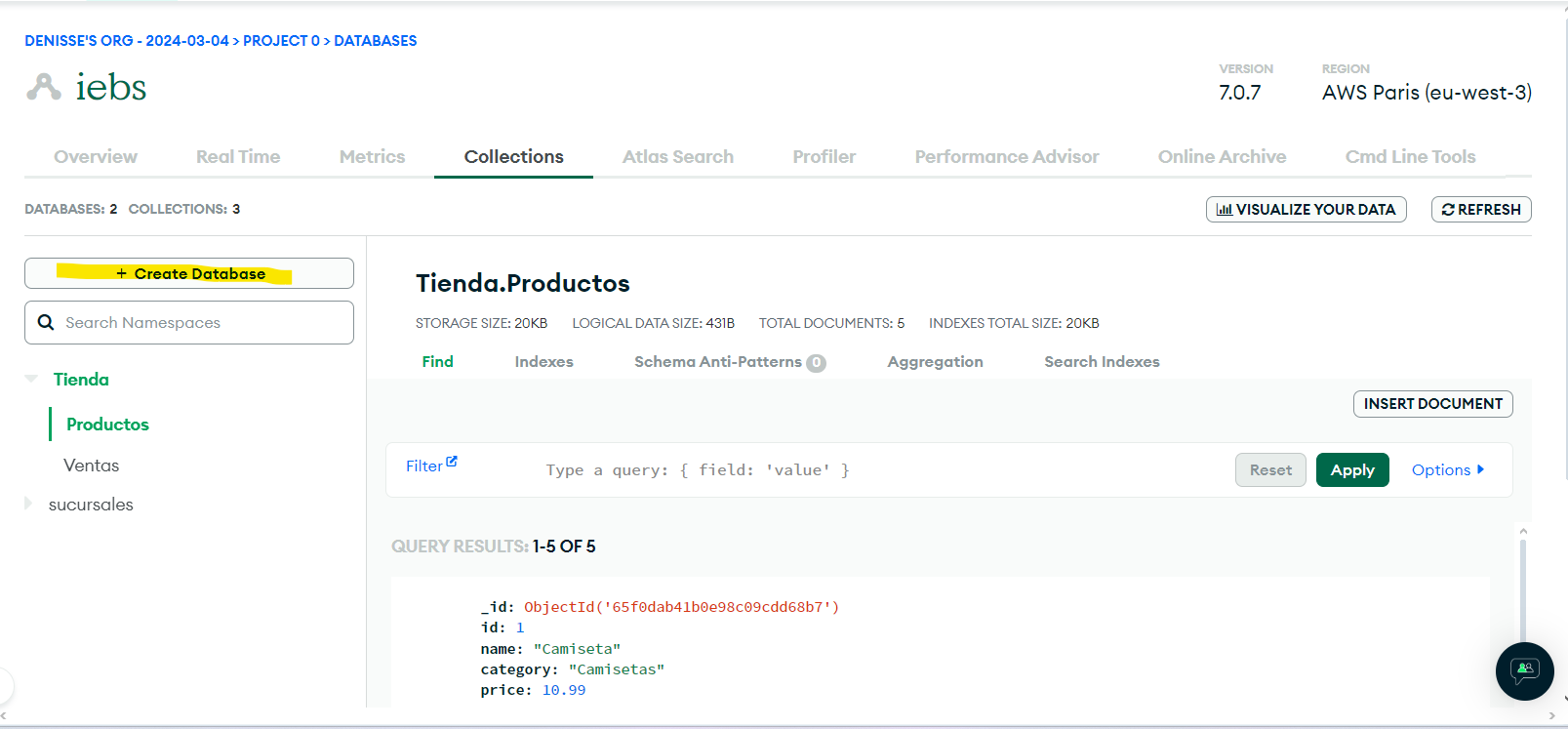
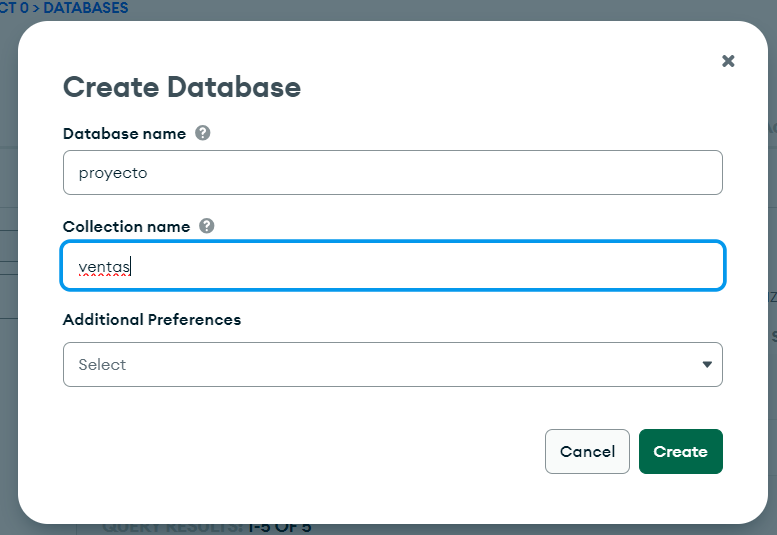
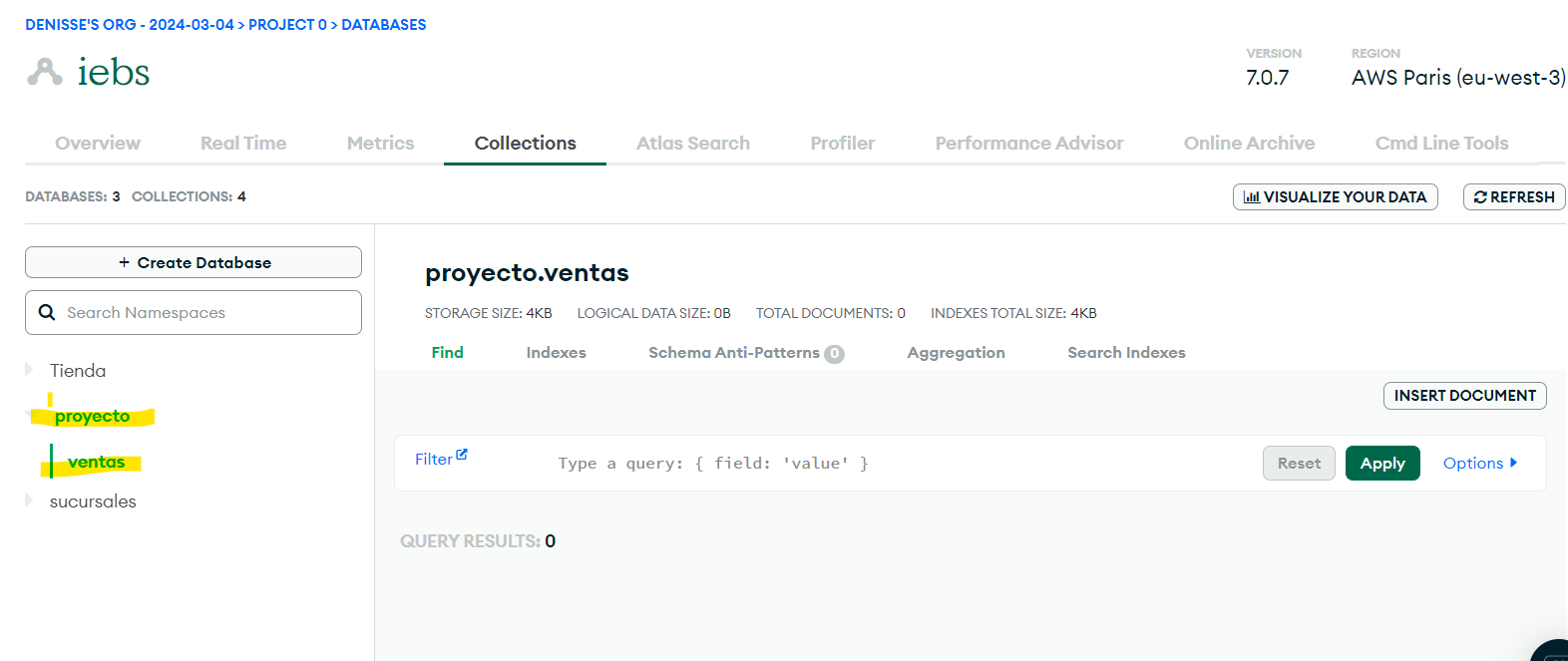
**Proyecto final**

1. **Crear la base de datos “proyecto” en nuestro cluster creado (si no se tiene habrá que crearlo).**

Accedemos a MongoDB Atlas, buscamos la opción de Database y vemos el cluster que ya había creado anteriormente llamado “IEBS”, seleccionamos “browse collections” para ver las bases de datos creadas.

Una vez ahí, seleccionamos “crear base de datos”.

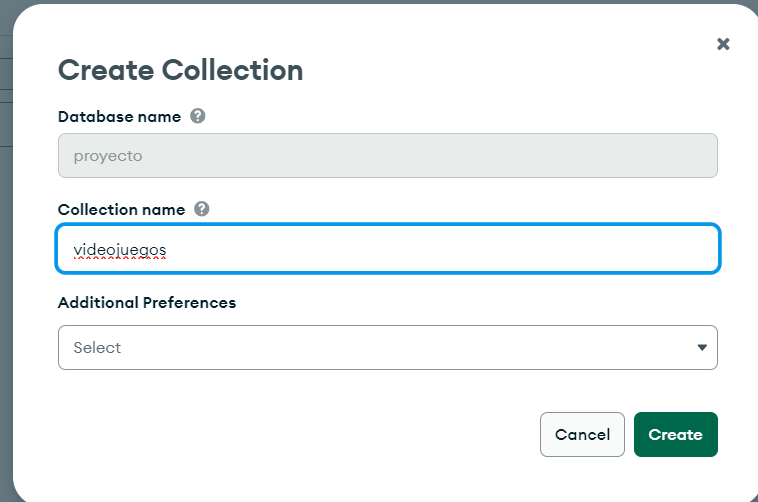
Se nombra la base de datos como se solicita y se crea la primera de las colecciones.

1. **Crear la colección “ventas” dentro de nuestra base de datos “proyecto”.**
2. **Crear la colección “videojuegos” dentro de nuestra base de datos “proyecto”.**

Seleccionamos el botón de más al lado del nombre de nuestra base de datos.

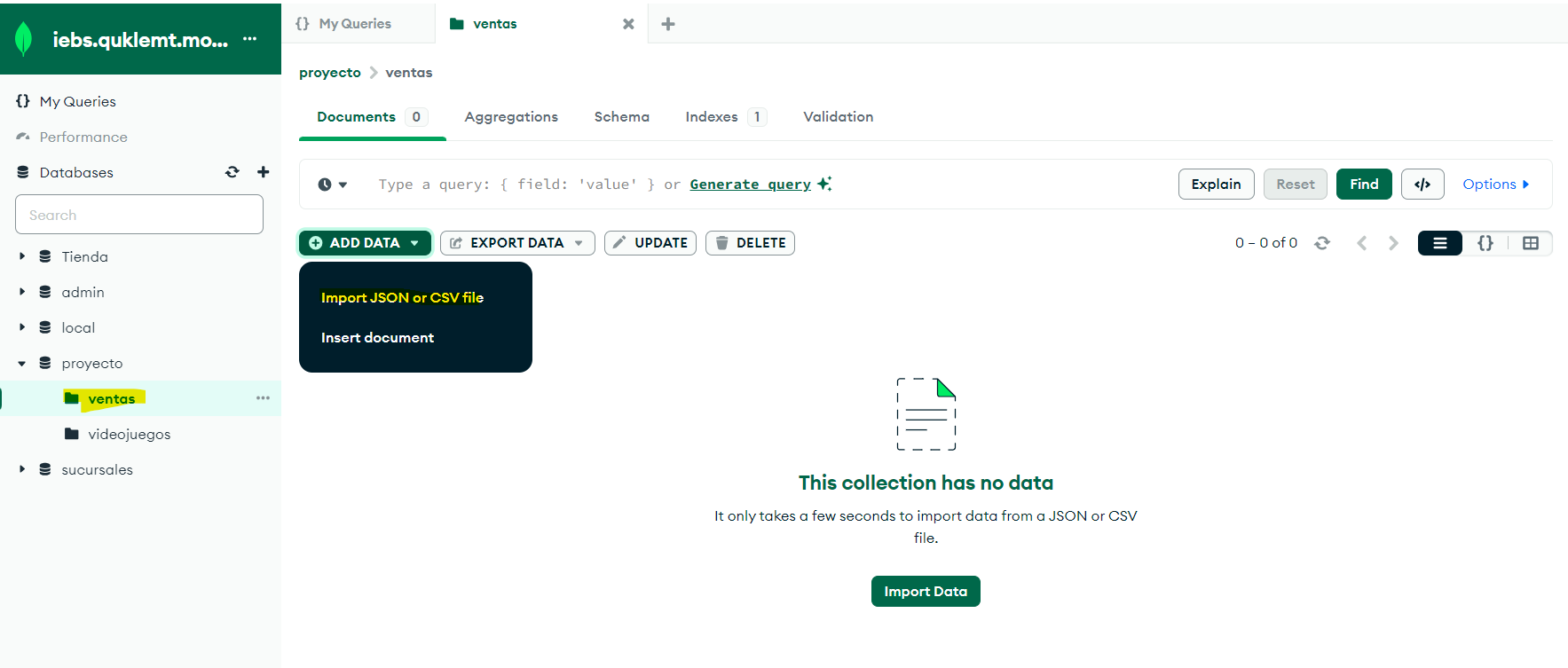
****

Nombramos la colección y le damos a crear

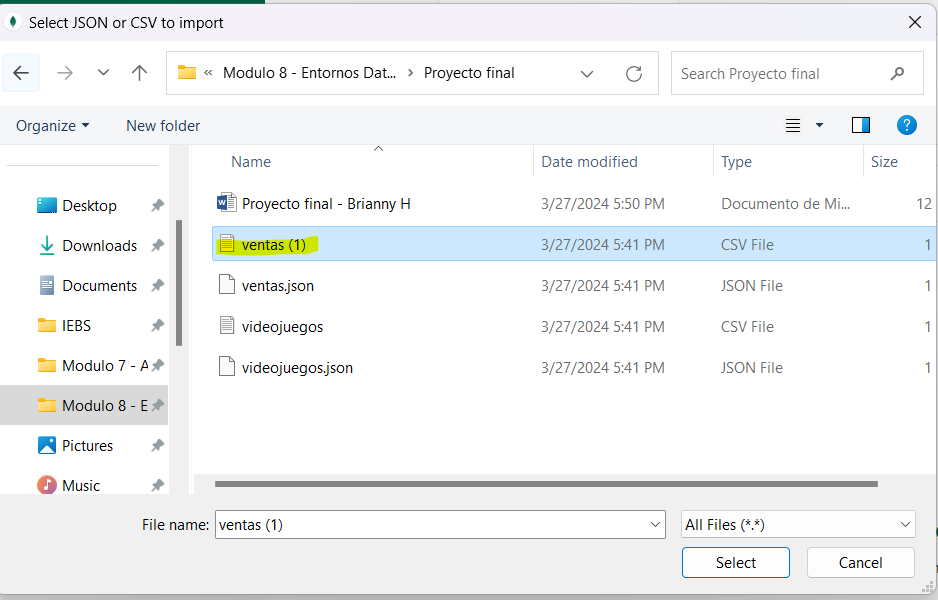


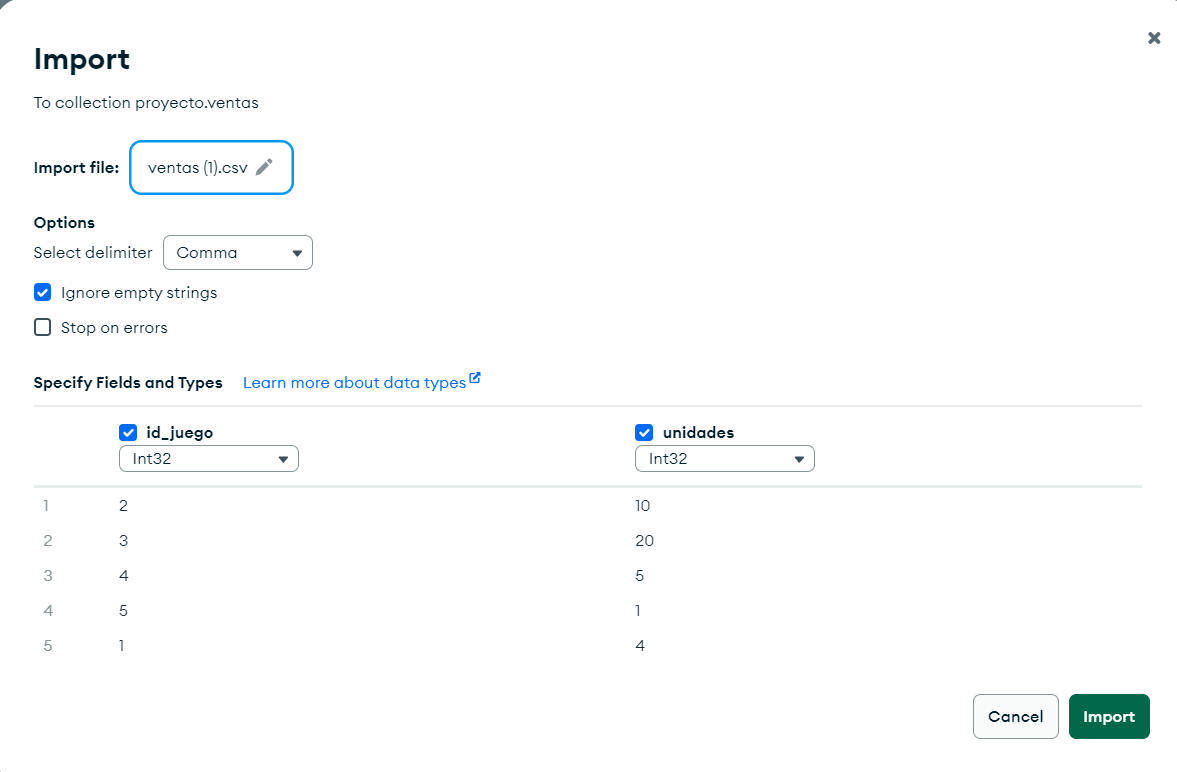
1. **Importar los datos de los ficheros aportados en el proyecto (deberíais de ver 2 de ventas y 2 de videojuegos, uno en formato csv y otro en formato json).**

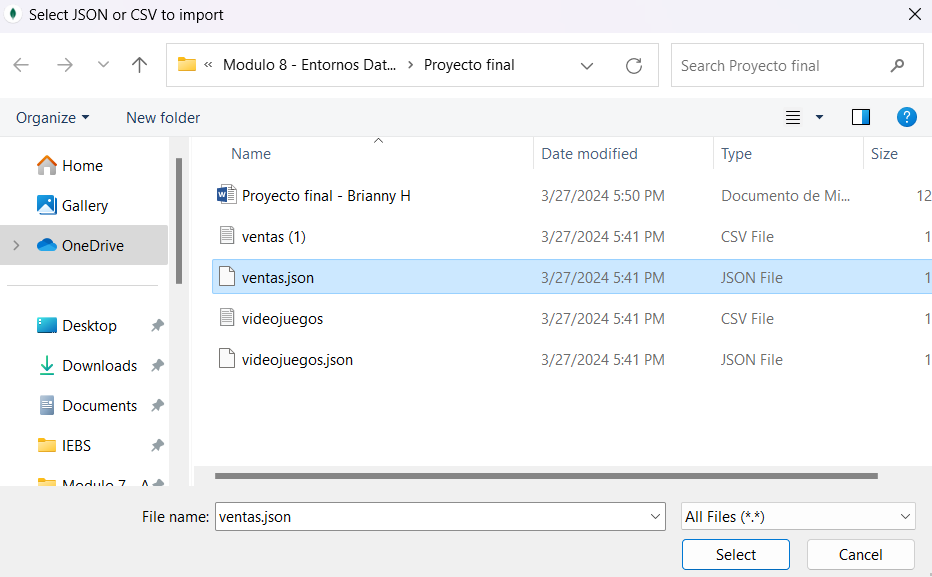
Empezamos por añadir las fuentes de datos a la colección de ventas.

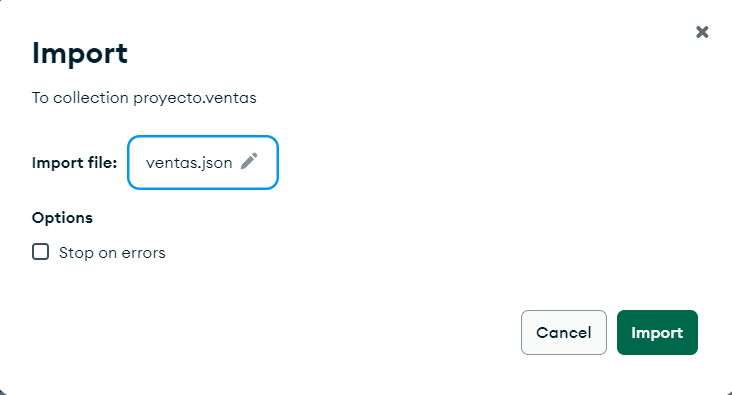


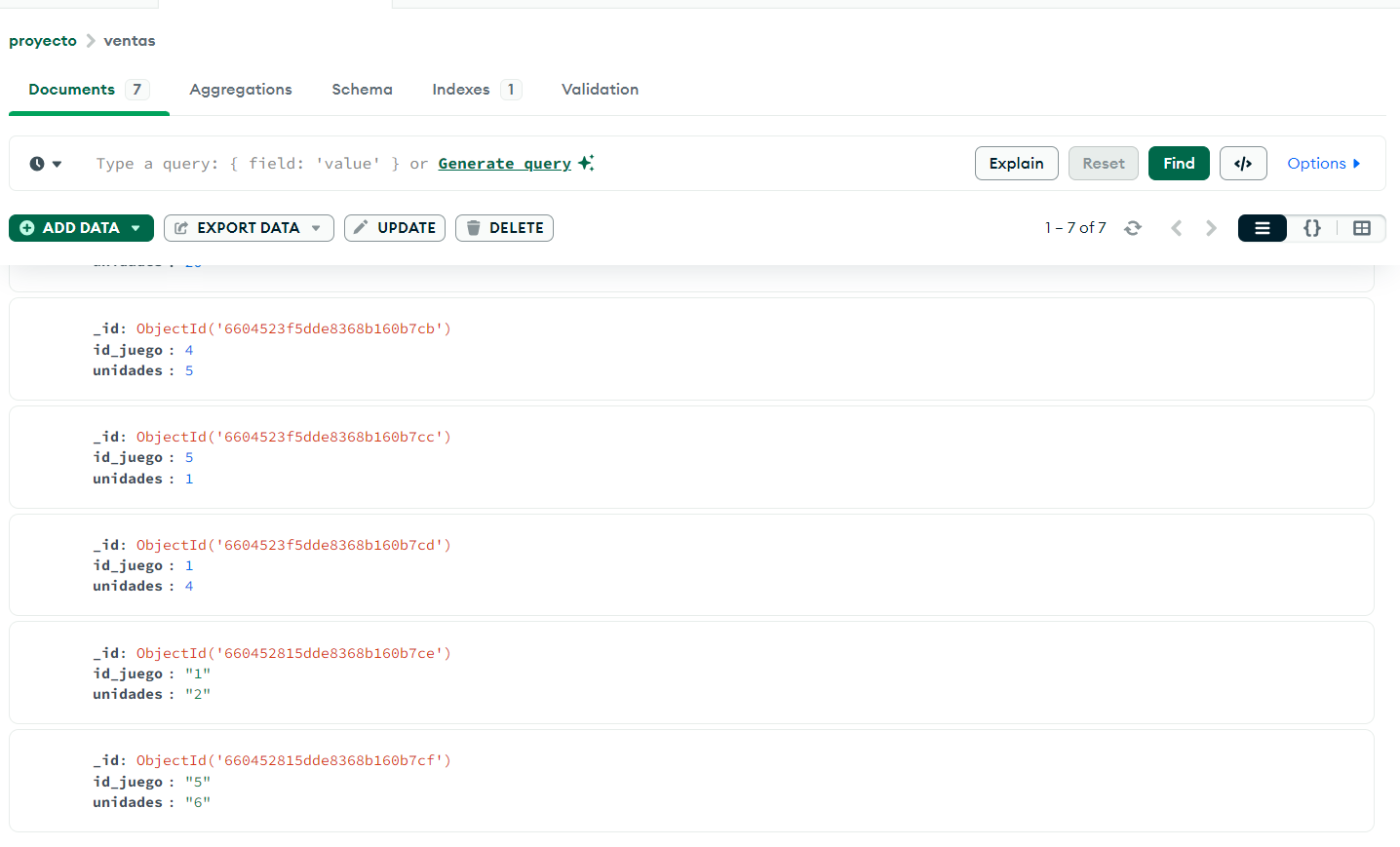
Seleccionamos el .csv

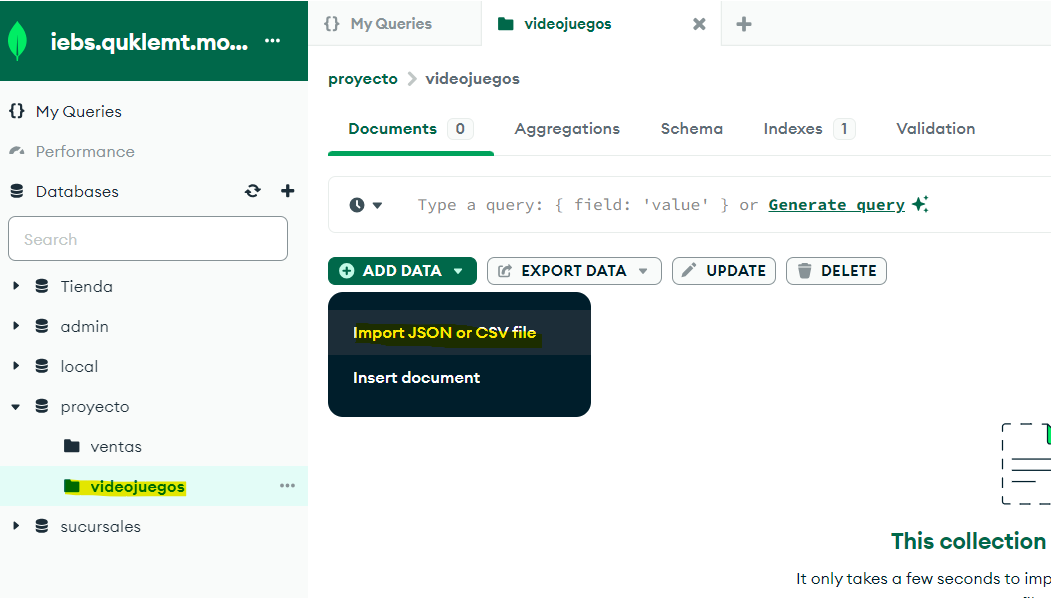


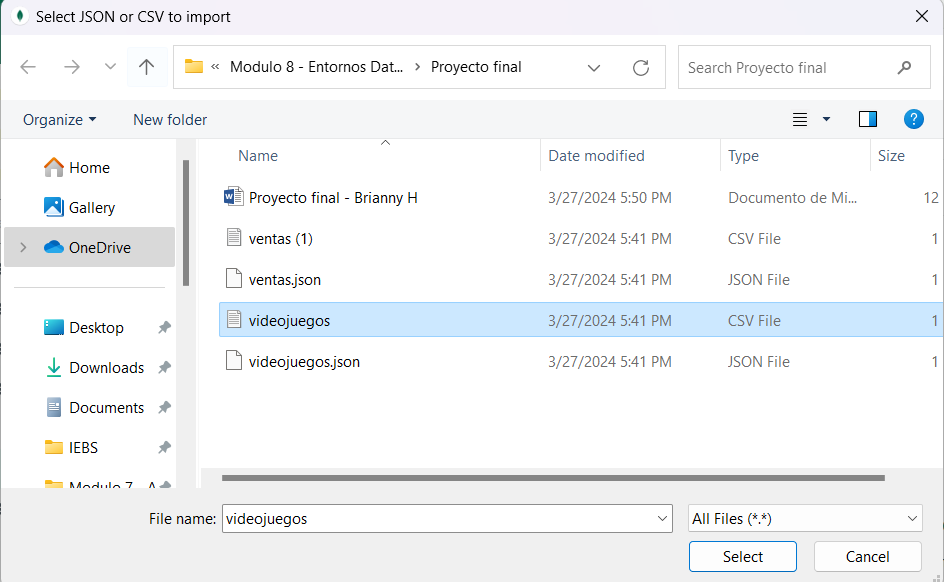
Verificar los datos a importar y presionar el botón de importar.

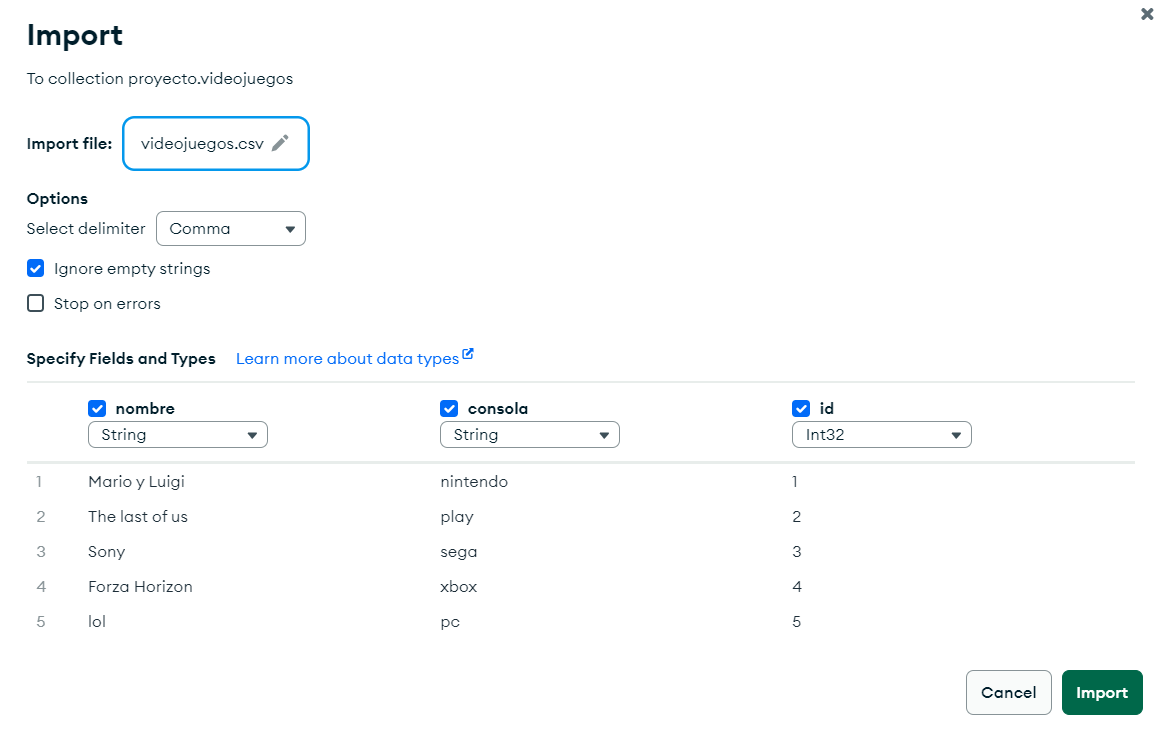
Repetimos el paso anterior pero esta vez con el .json

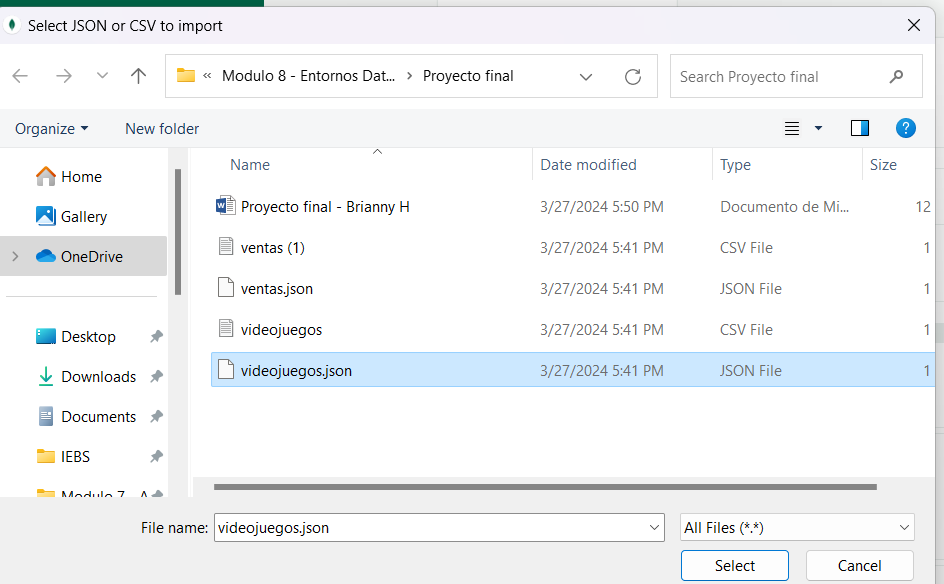


Tendremos un resultado de 7 documentos.

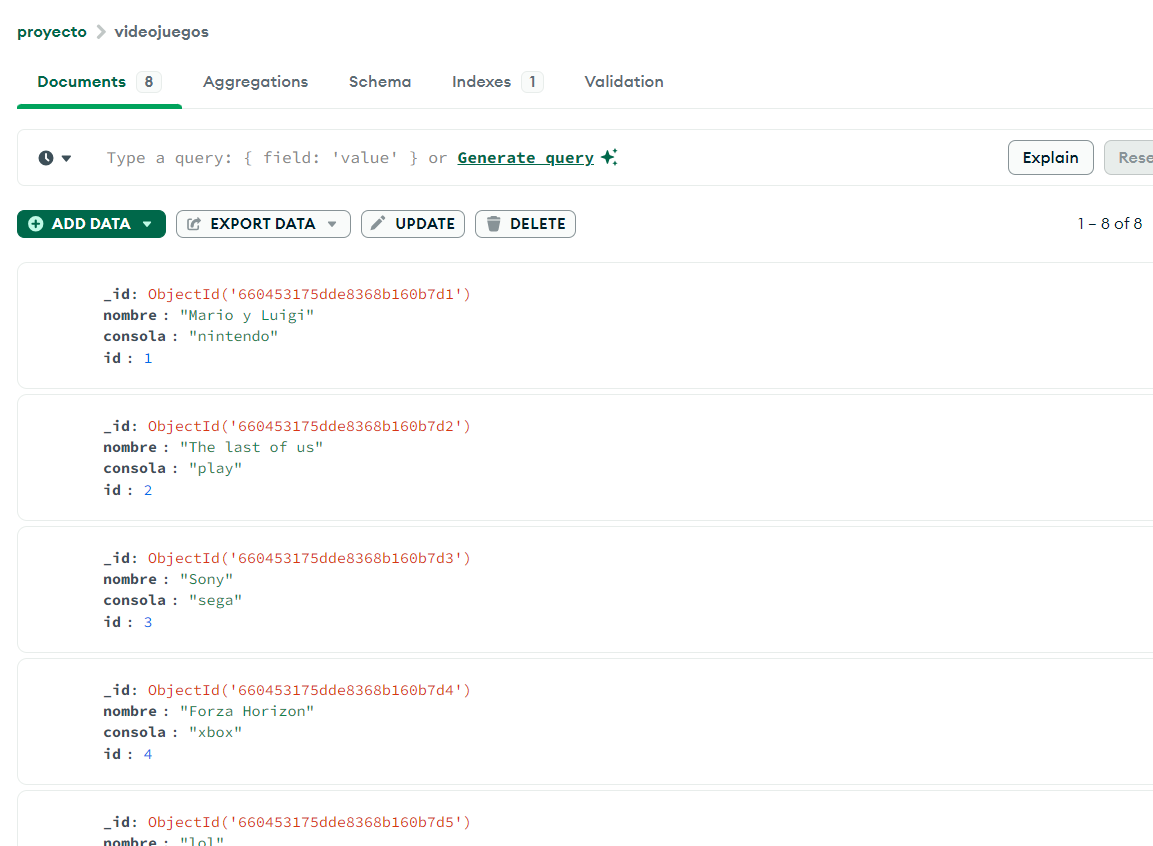
Ahora repetimos los pasos con la colección de videojuegos.

Importamos el .csv

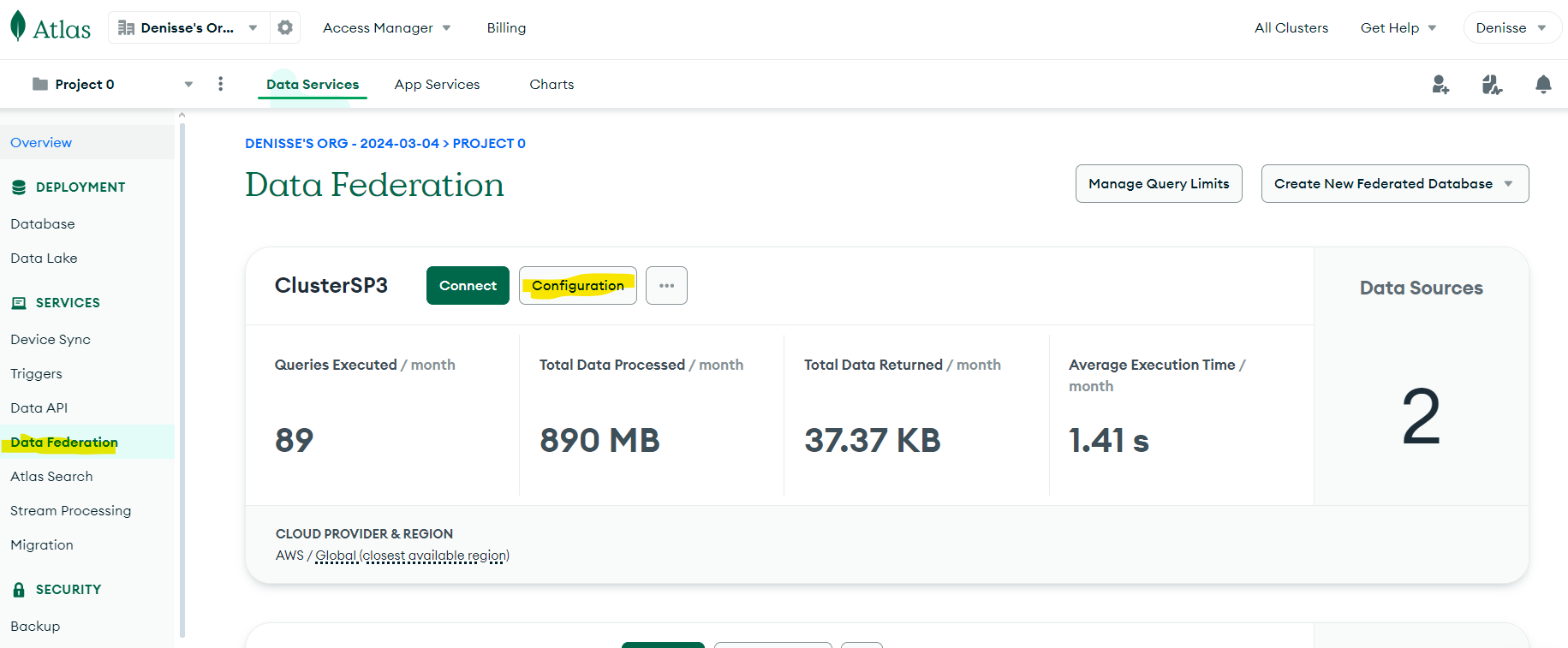


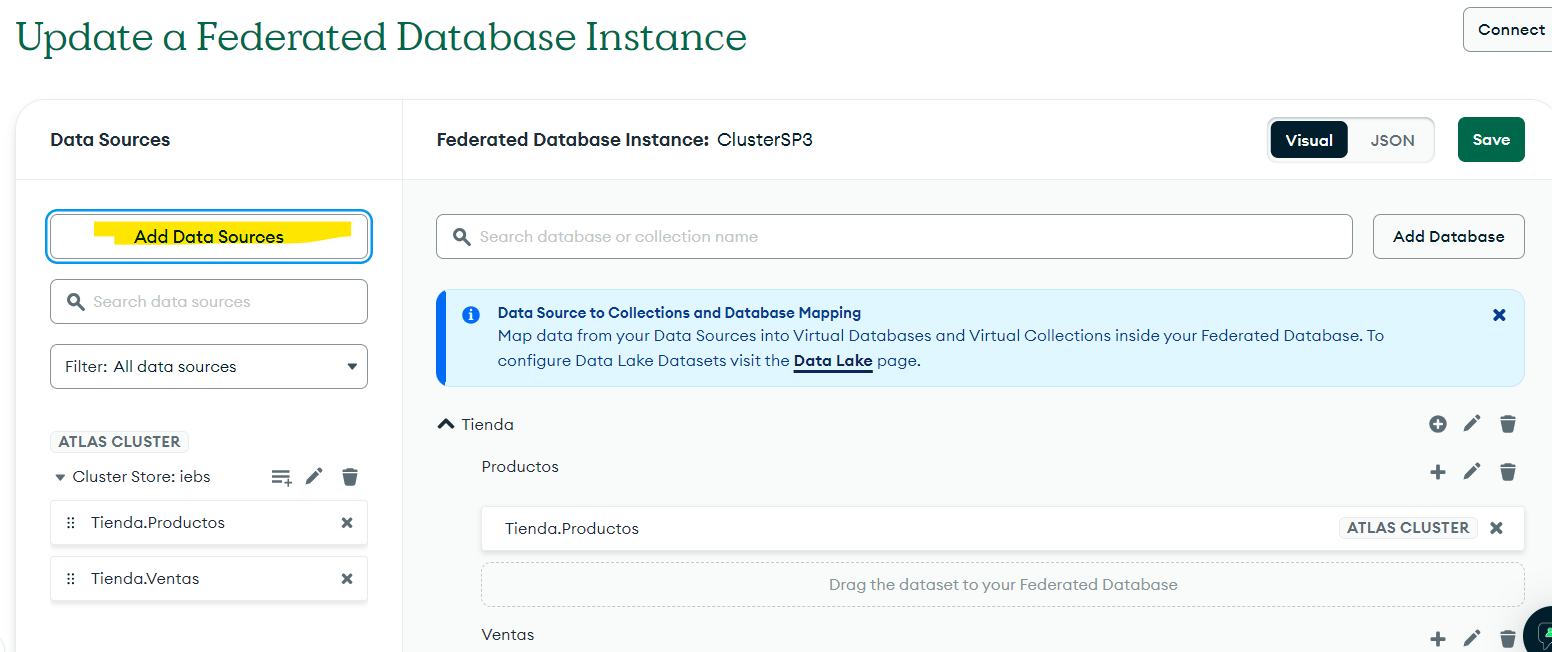
Importamos el .json

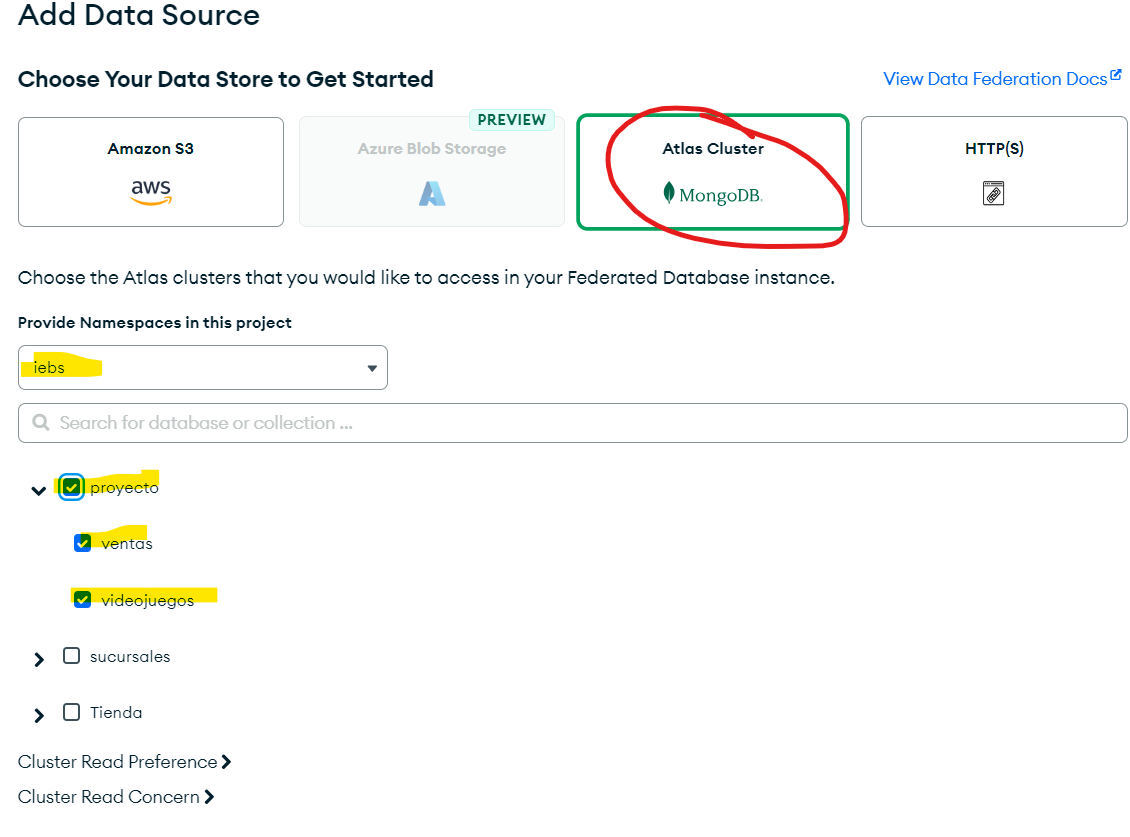


Dejándonos un total de 8 documentos.

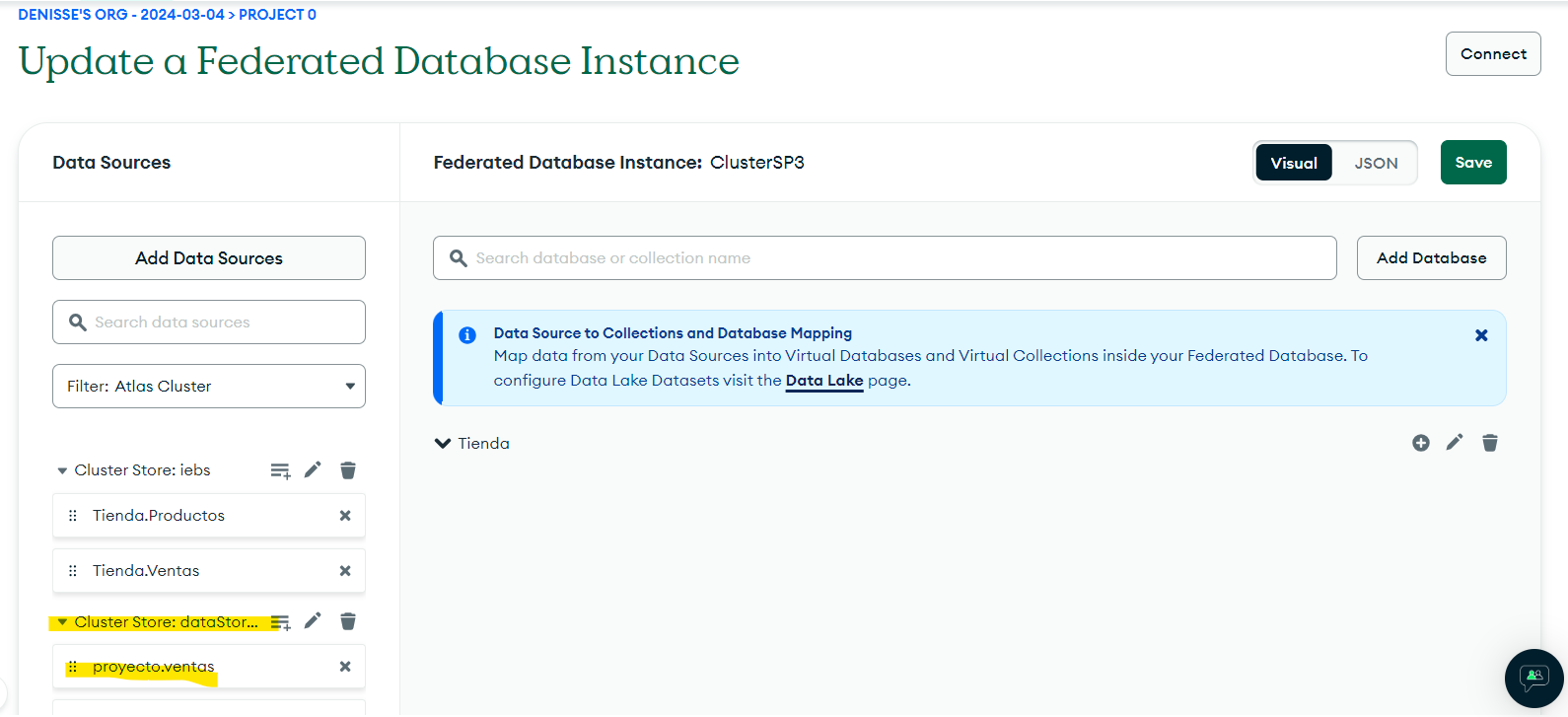
1. **Conectar el Data Federation con nuestro cluster introduciendo la base de datos “proyecto”.**

En este caso usare el mismo del sprint anterior, para conectar la base de datos proyecto. Y seleccionamos la opción “configuration”

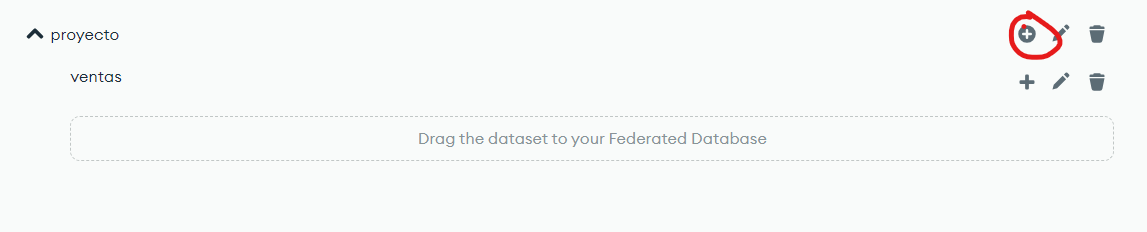
Se escoge la opción de “add data source”.

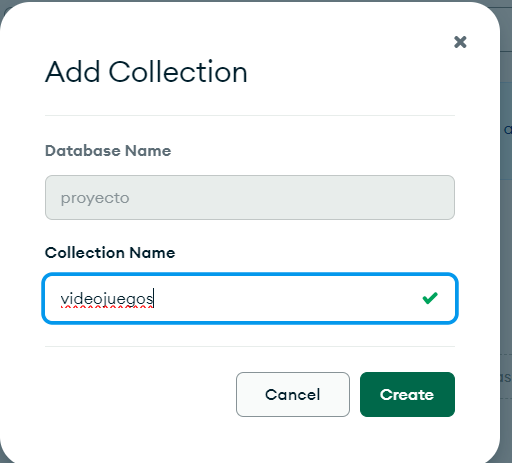
Escoger “atlas cluster” e identificamos el nombre de nuestro cluster y la base de datos requerida junto a sus dos colecciones.

1. **Crear la colección “ventas” y “videojuegos” en la base de datos “proyecto” del Data Lake.**

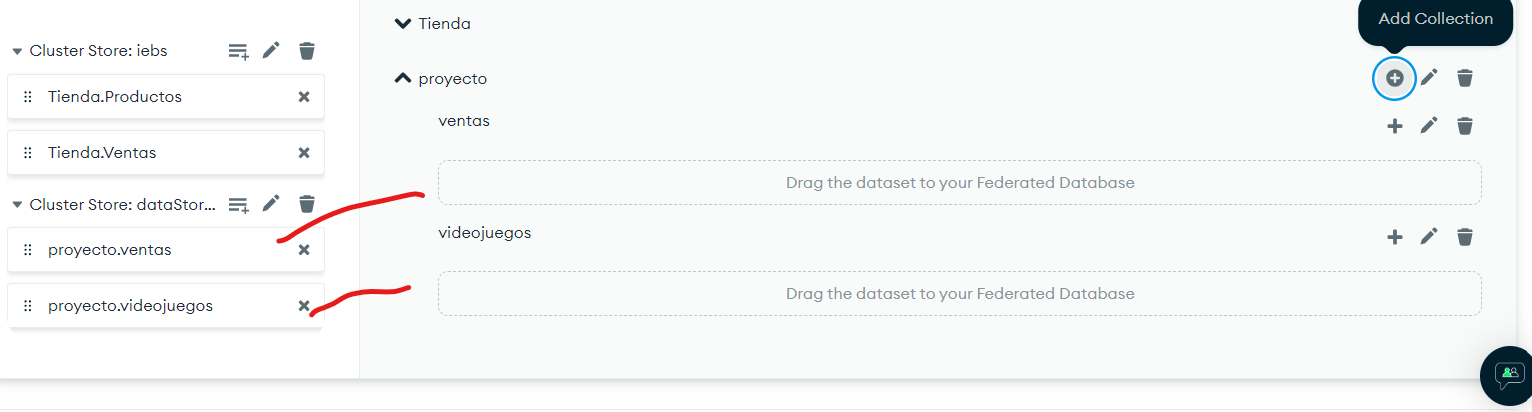
Verificamos que las colecciones seleccionadas están en el data federation. Le damos a añadir base de datos.

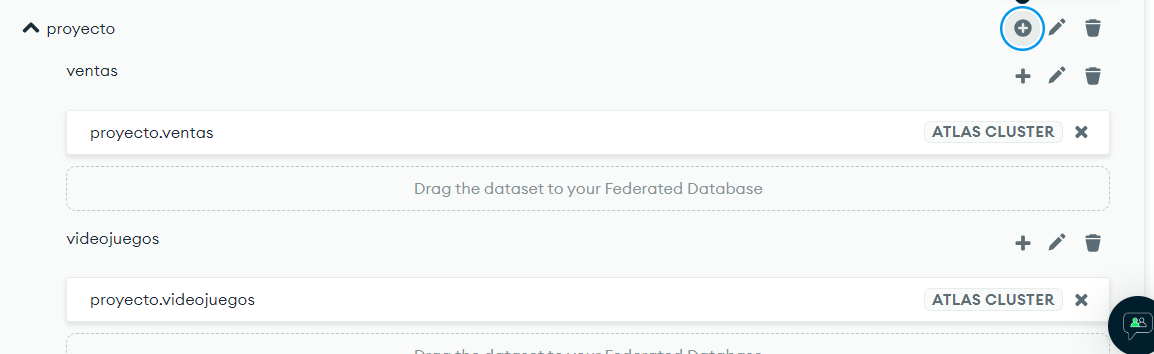
Ponemos el mismo nombre de la base de datos “proyecto” y una de las colecciones “ventas” y le damos a crear.

Añadimos la segunda colección de “proyecto”.

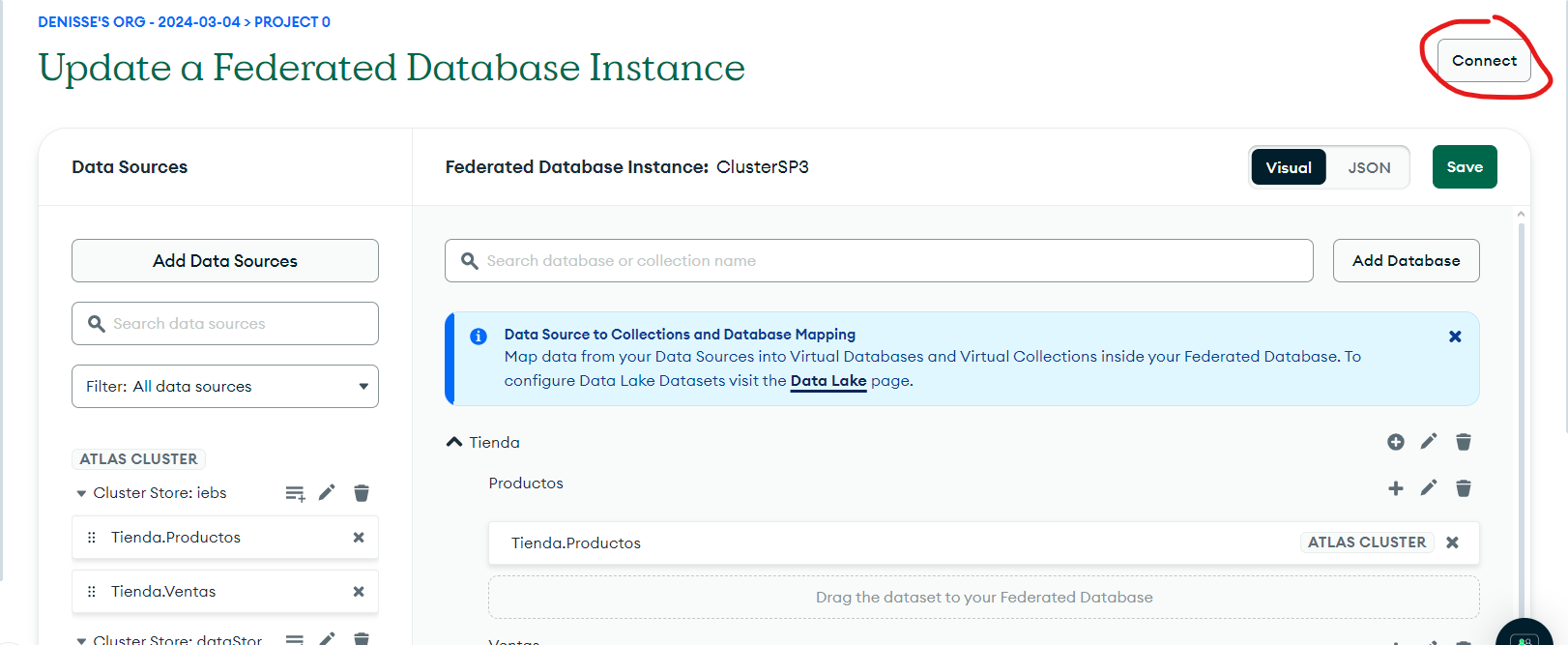


1. **Poner todos los datos de nuestras colecciones “ventas” y “videojuegos” a las del Data Federation.**

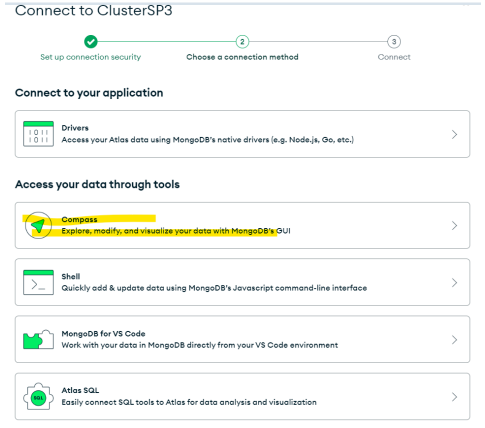
Se arrastra cada una de las colecciones añadidas anteriormente a la correspondiente creada.



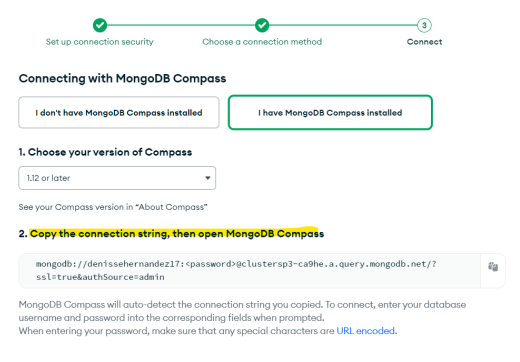
1. **Conectarnos desde MongoDB Compass.**

Buscamos la database federated y seleccionamos el botón de connect.

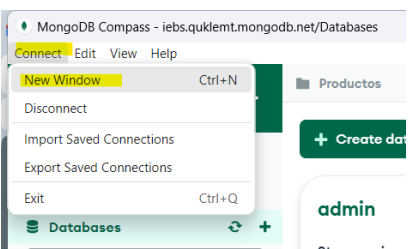
Seleccionamos Compass.



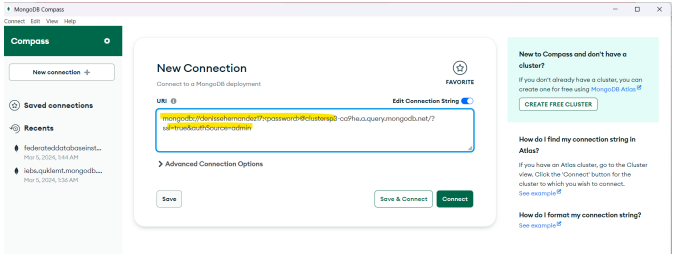
Copiamos el connection string que nos trae este apartado y nos dirigimos a Compass.

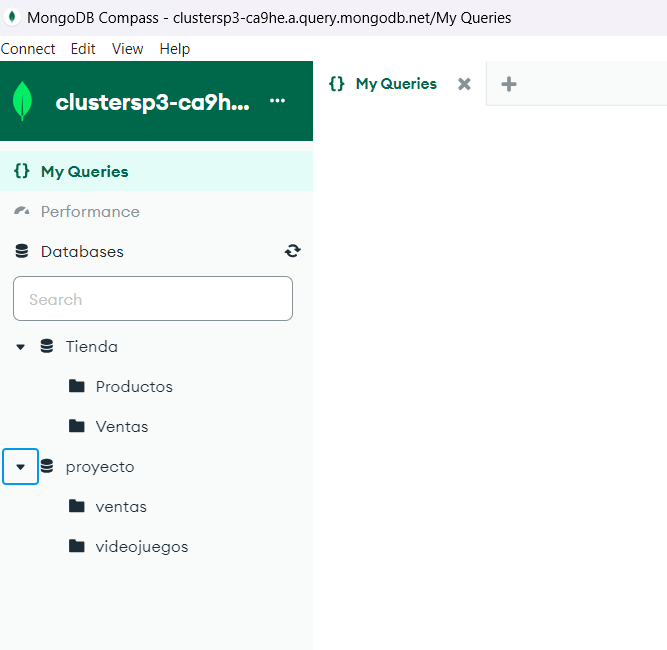


En compass, seleccionamos conectar y luego new Windows.

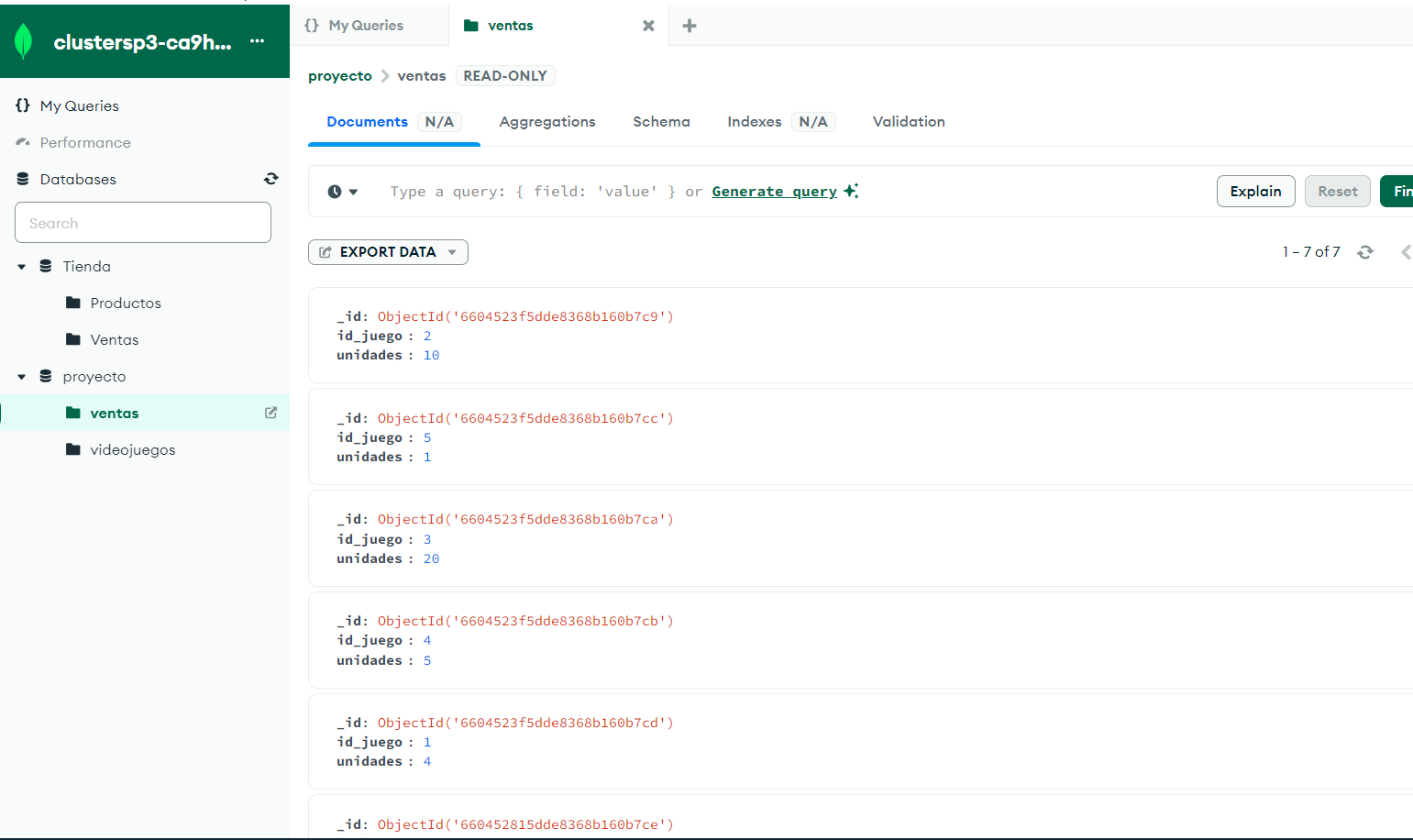


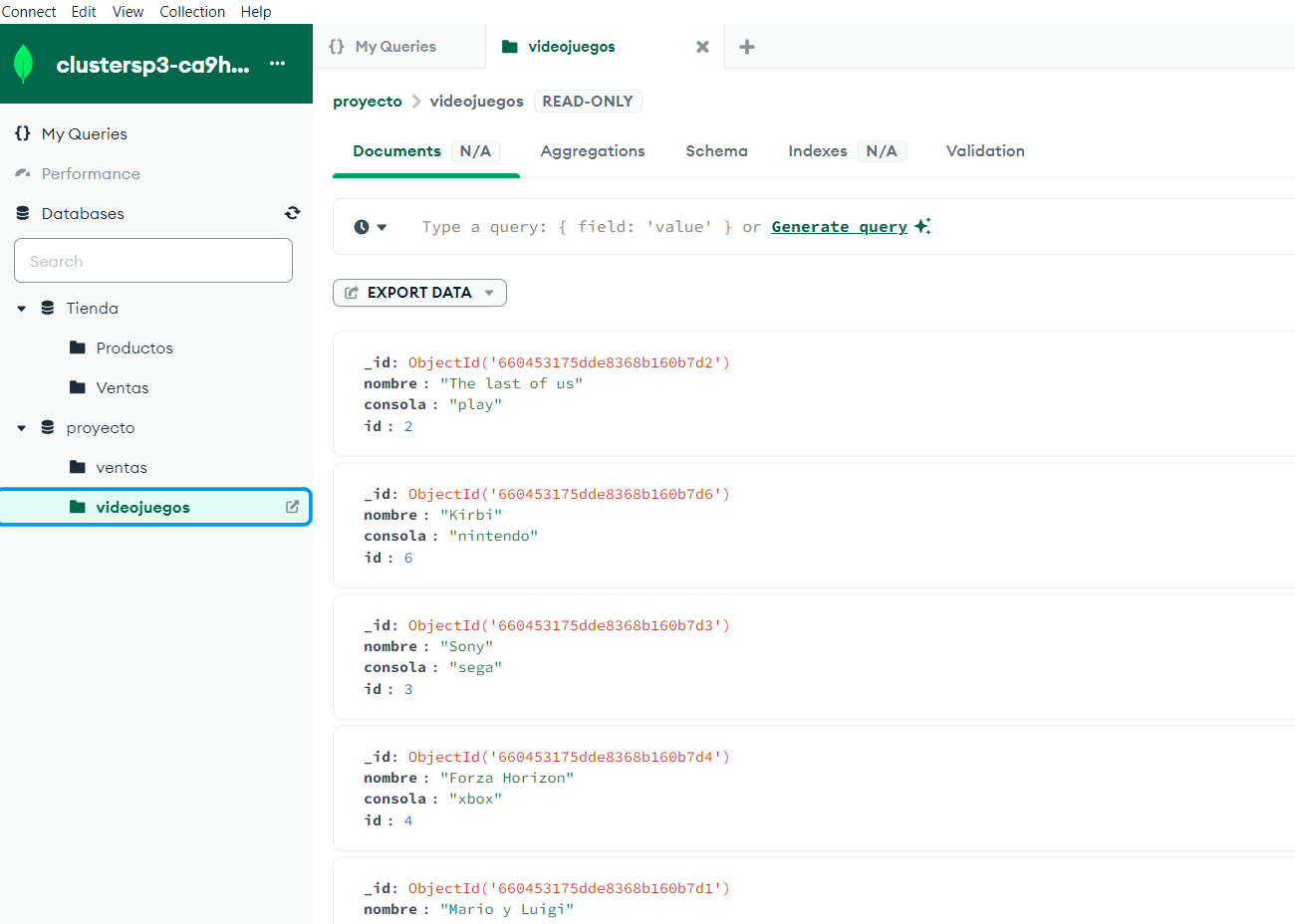
Pegamos la conexión que acabamos de copiar.



Y estaremos conectados.

1. **Hacer una query a cada conjunto (ver la colección “ventas” es una concatenación de los datos que tenemos más los públicos).**

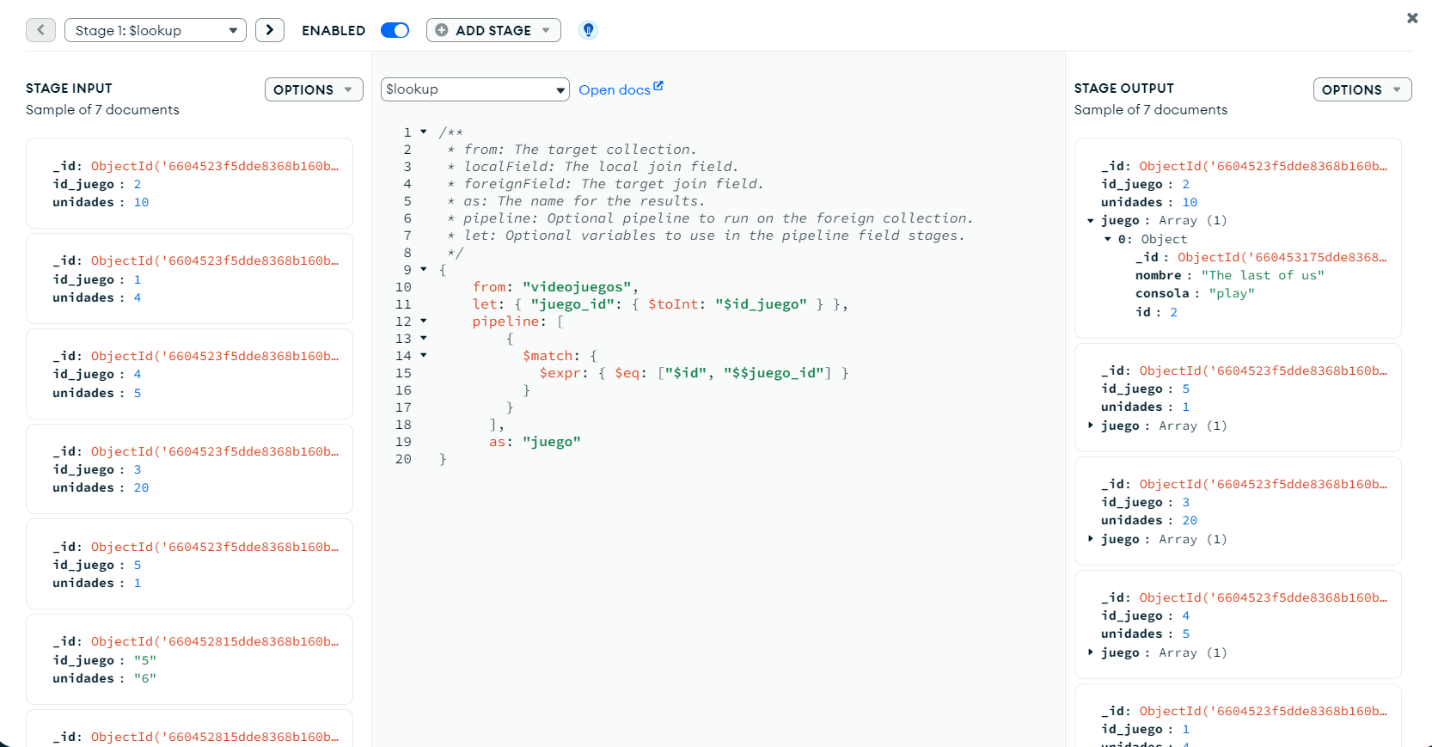


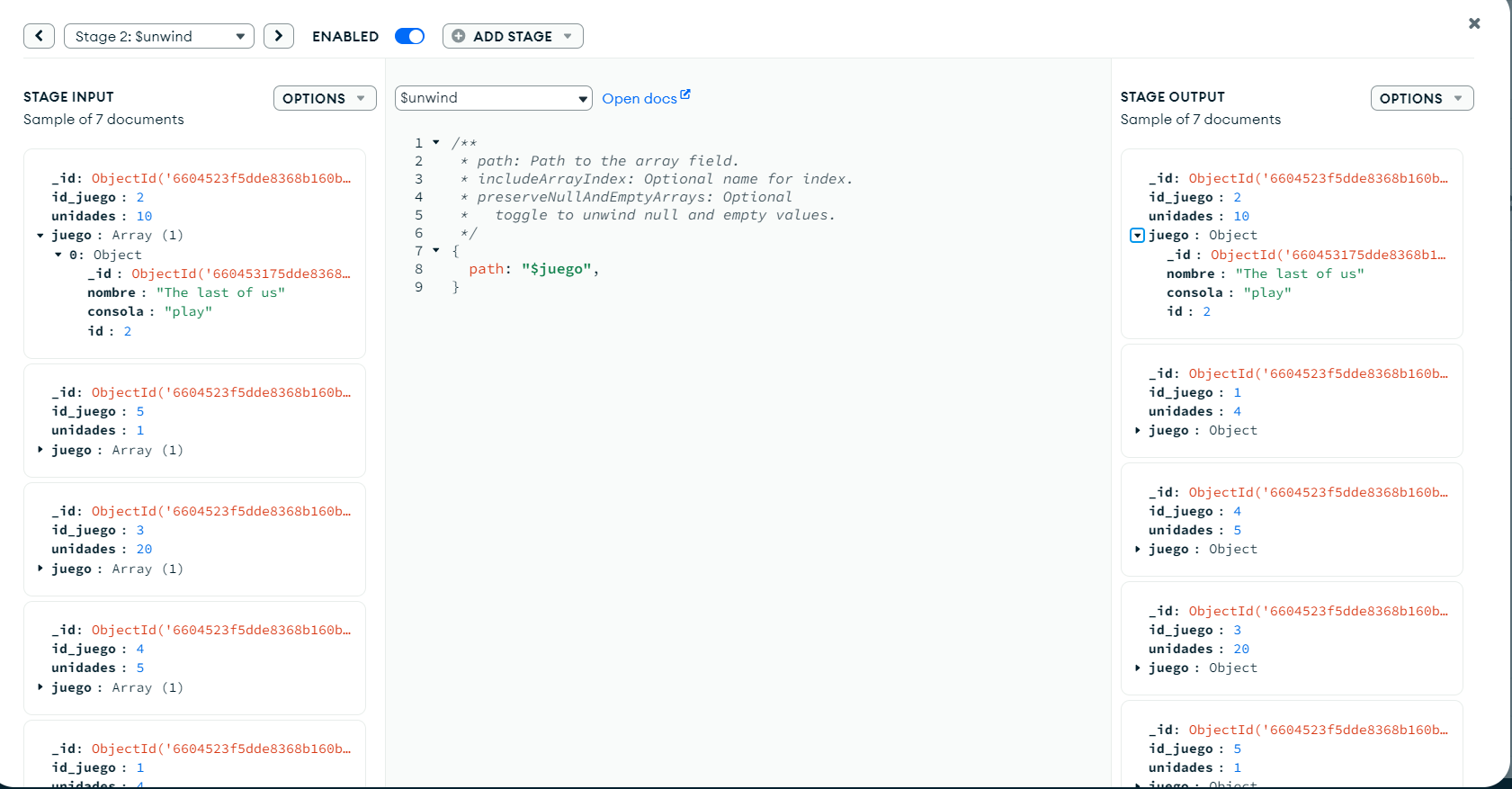


1. **Utilizar las operaciones disponibles en “Aggregation” de MongoDB Compass para conseguir:**

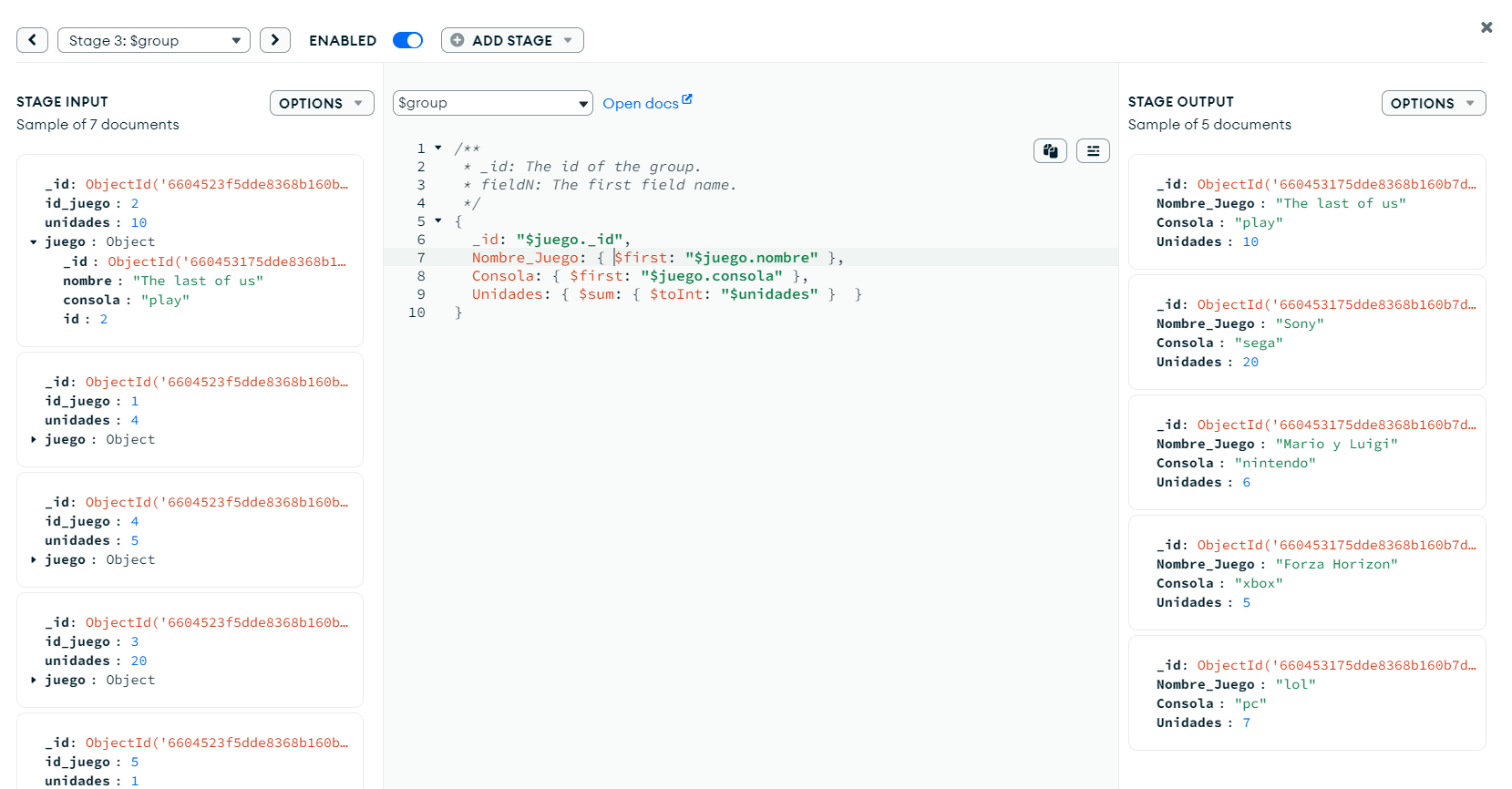
* **Las unidades vendidas totales de cada juego.**

Se utiliza $lookup para realizar una unión entre las colecciones "ventas" y "videojuegos" basándose en el campo "Producto" de "ventas".

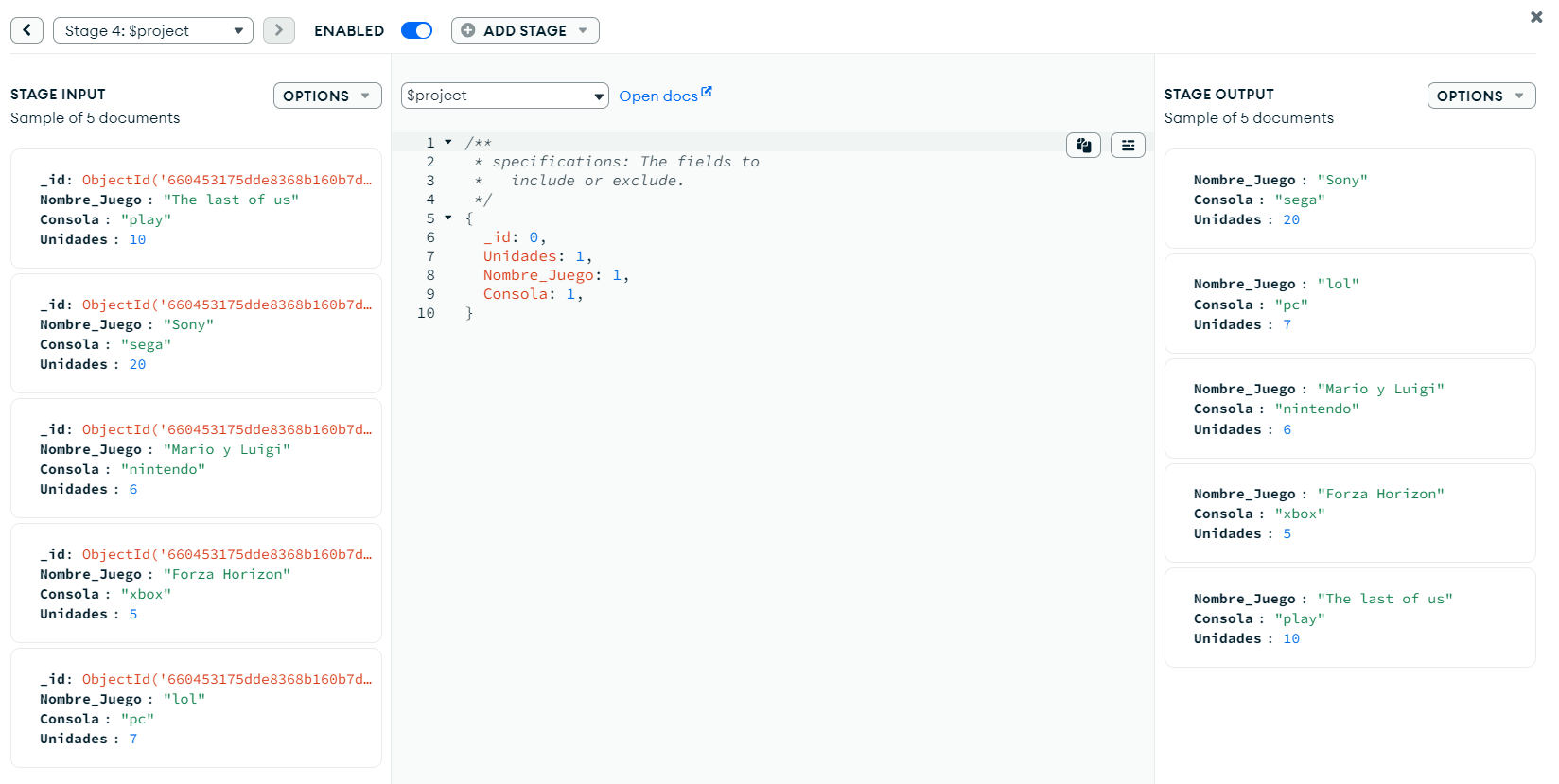
Los campos id\_juego en la colección "ventas" y id en la colección "videojuegos" son de tipos diferentes ("int" en "videojuegos" y "string" en "ventas") y se utiliza $expr se utiliza para convertir el tipo de datos del campo id\_juego en la colección "ventas" a entero, ya que el campo id en la colección "videojuegos" es de tipo entero.

La etapa $unwind se utiliza para descomponer el campo de matriz (array) en múltiples documentos separados.

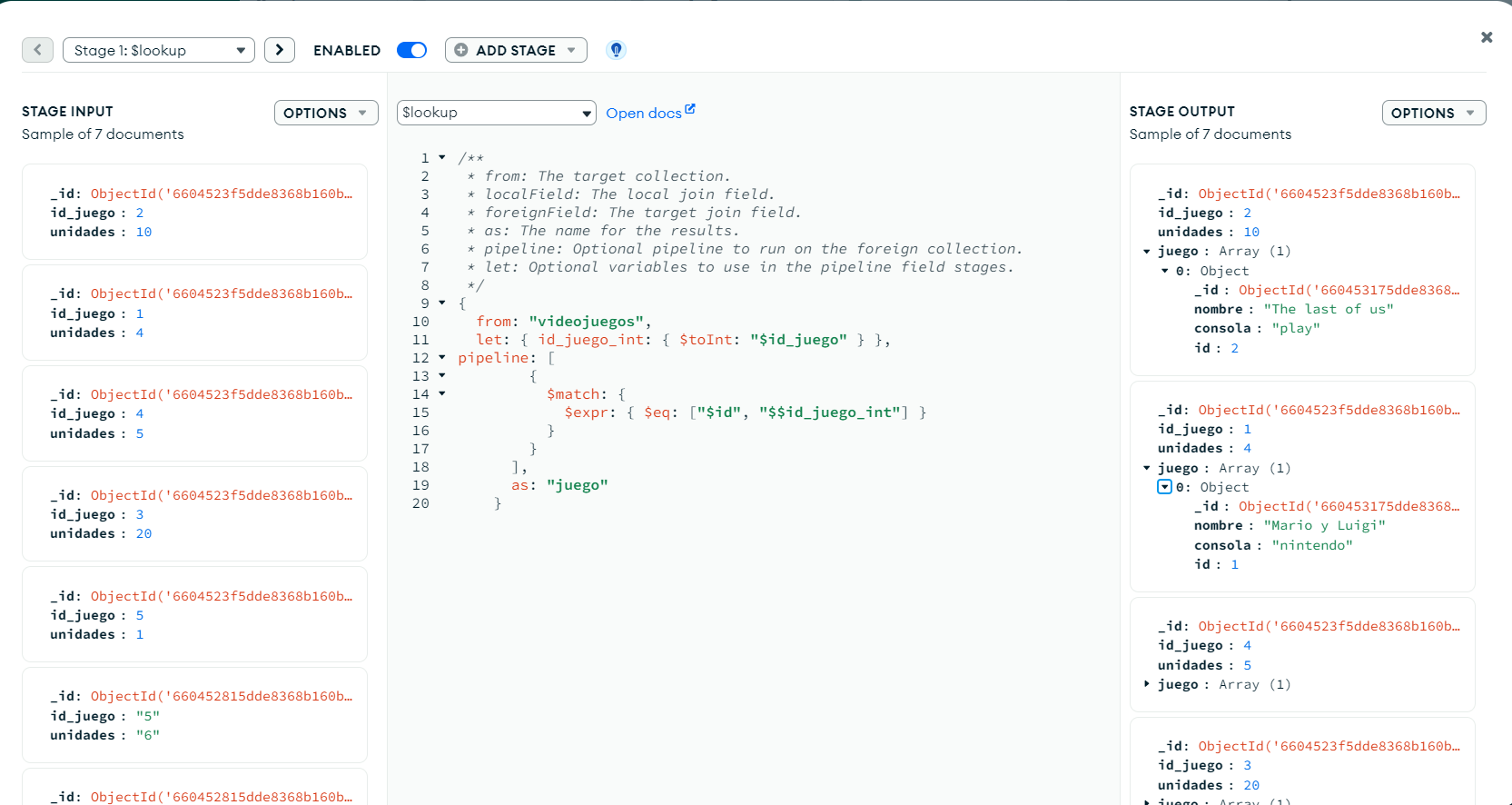
$group se utiliza para agrupar los documentos que provienen de la etapa $unwind, donde cada documento representa información de un juego individual. El campo id\_juego de tipo "string" se convirtió a "int" utilizando $toInt.

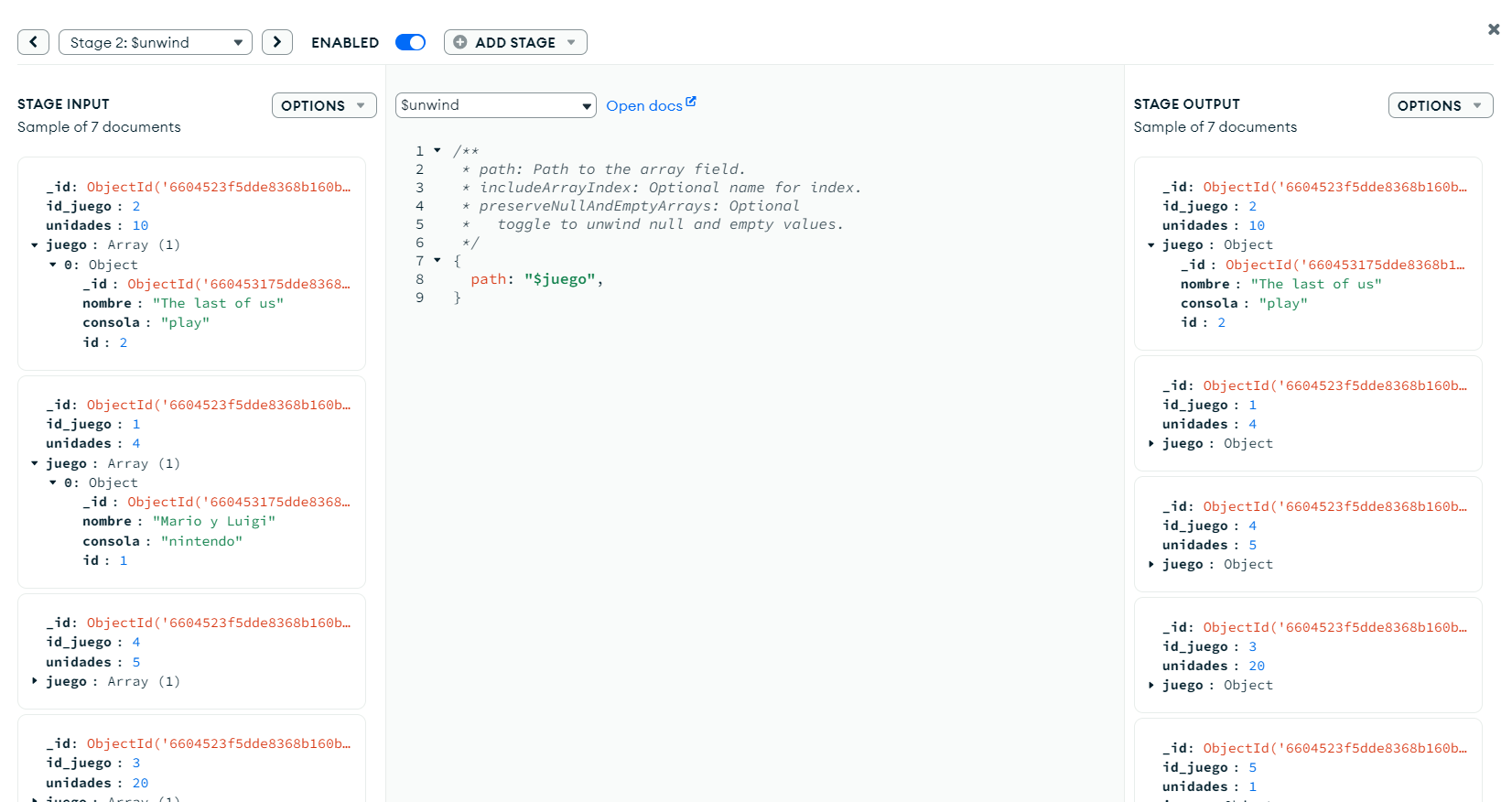


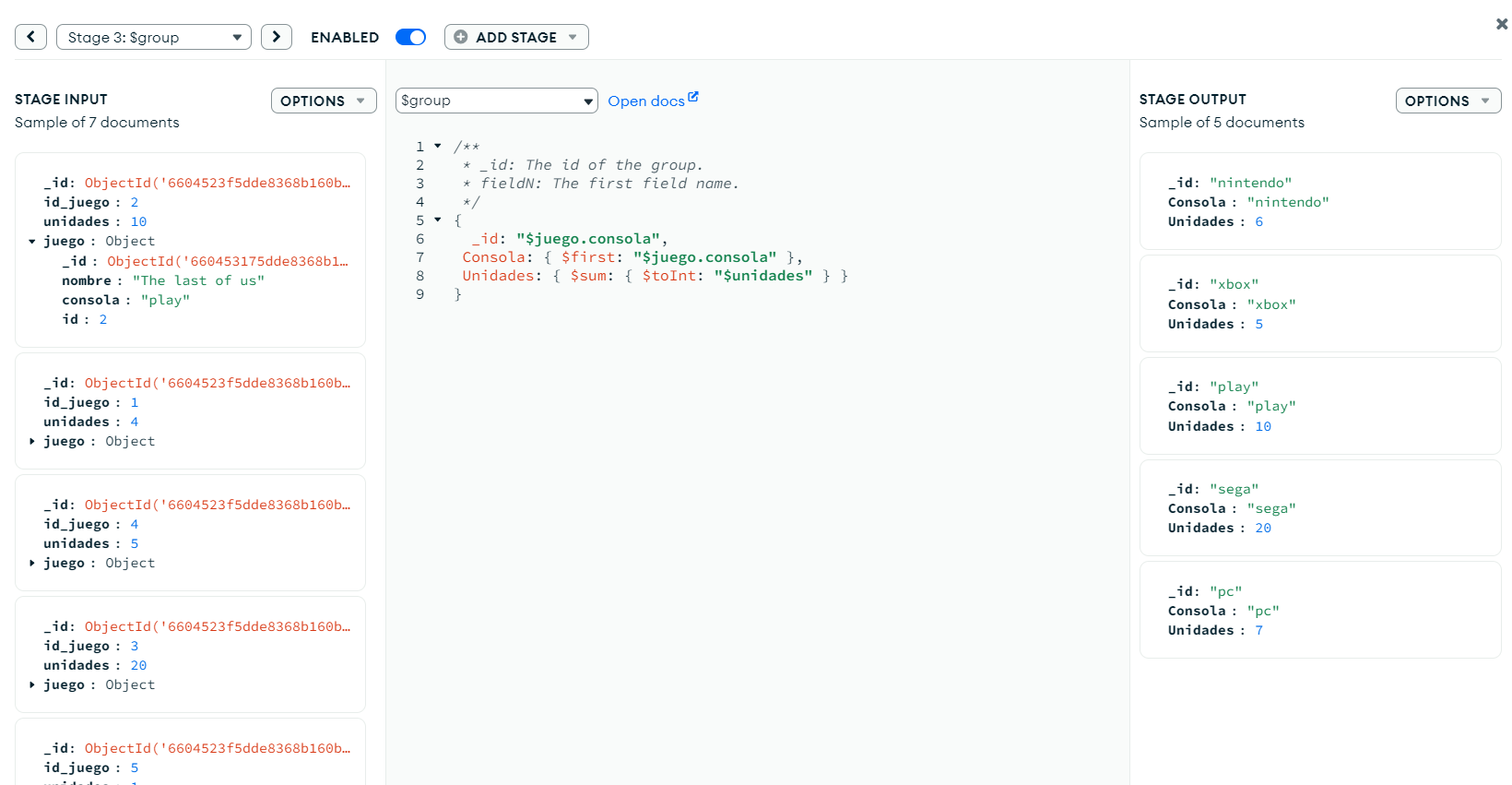
$project se utiliza al final del pipeline de agregación para dar formato a los documentos de salida y seleccionar los campos que deseamos presentar en el resultado final.

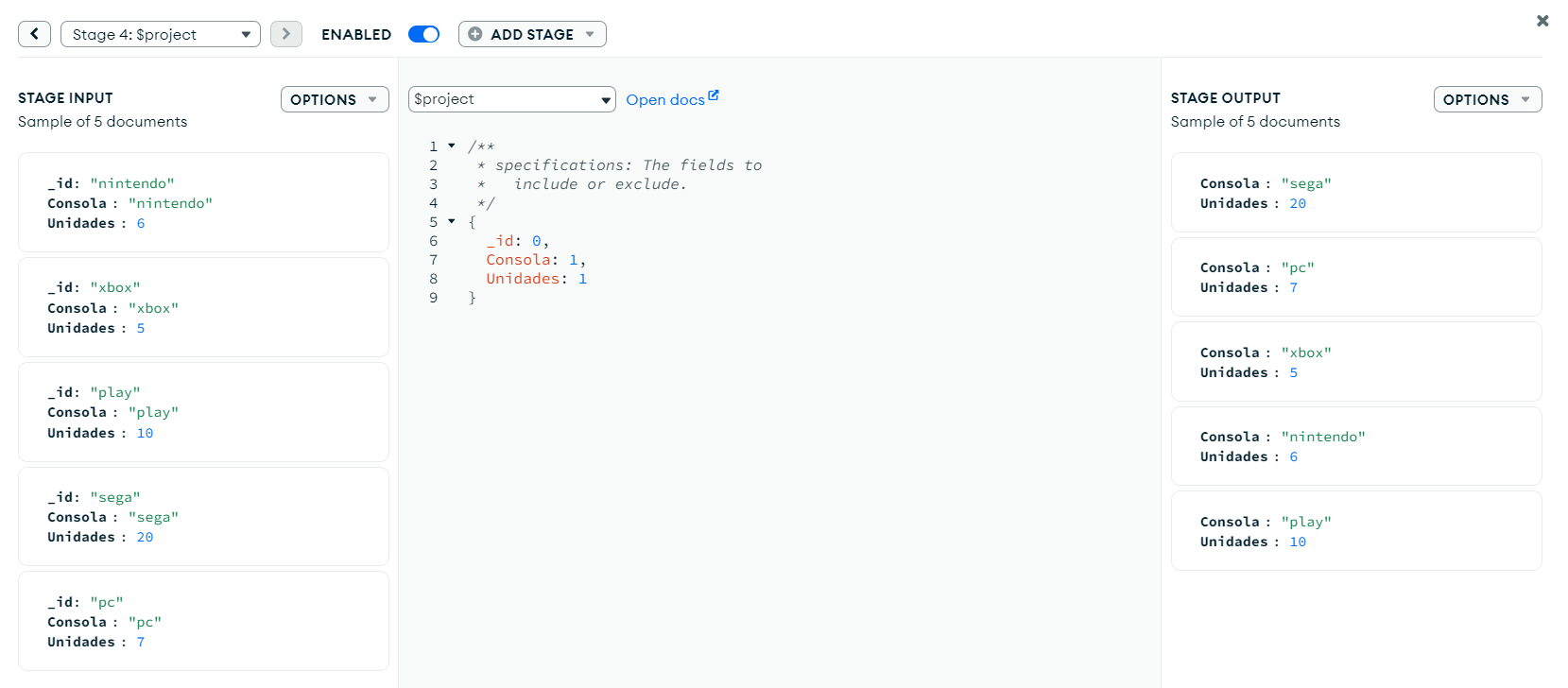


* **Las unidades vendidas totales de cada consola.**

Estas etapas serían las mismas del paso anterior pero para las consolas.

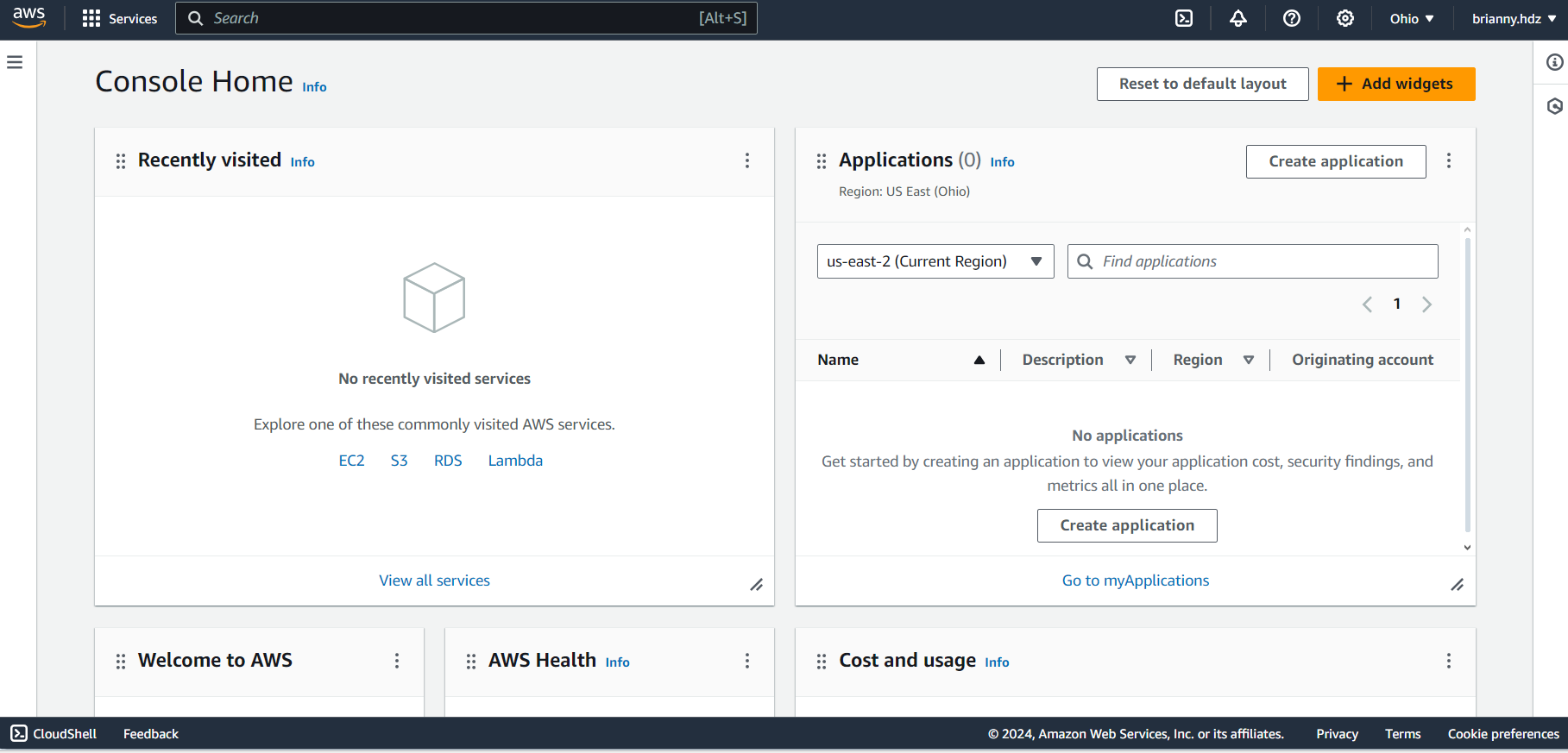


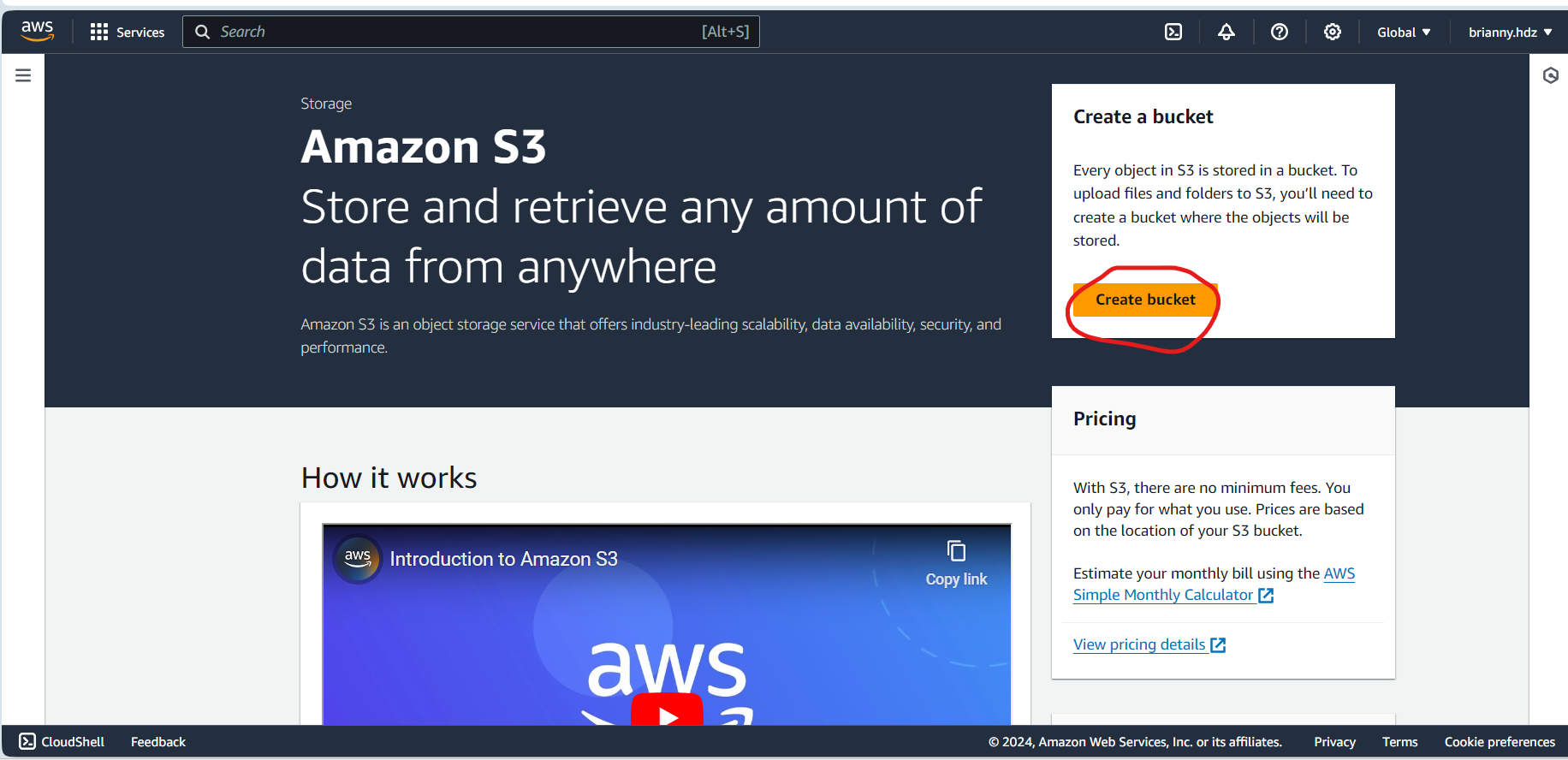


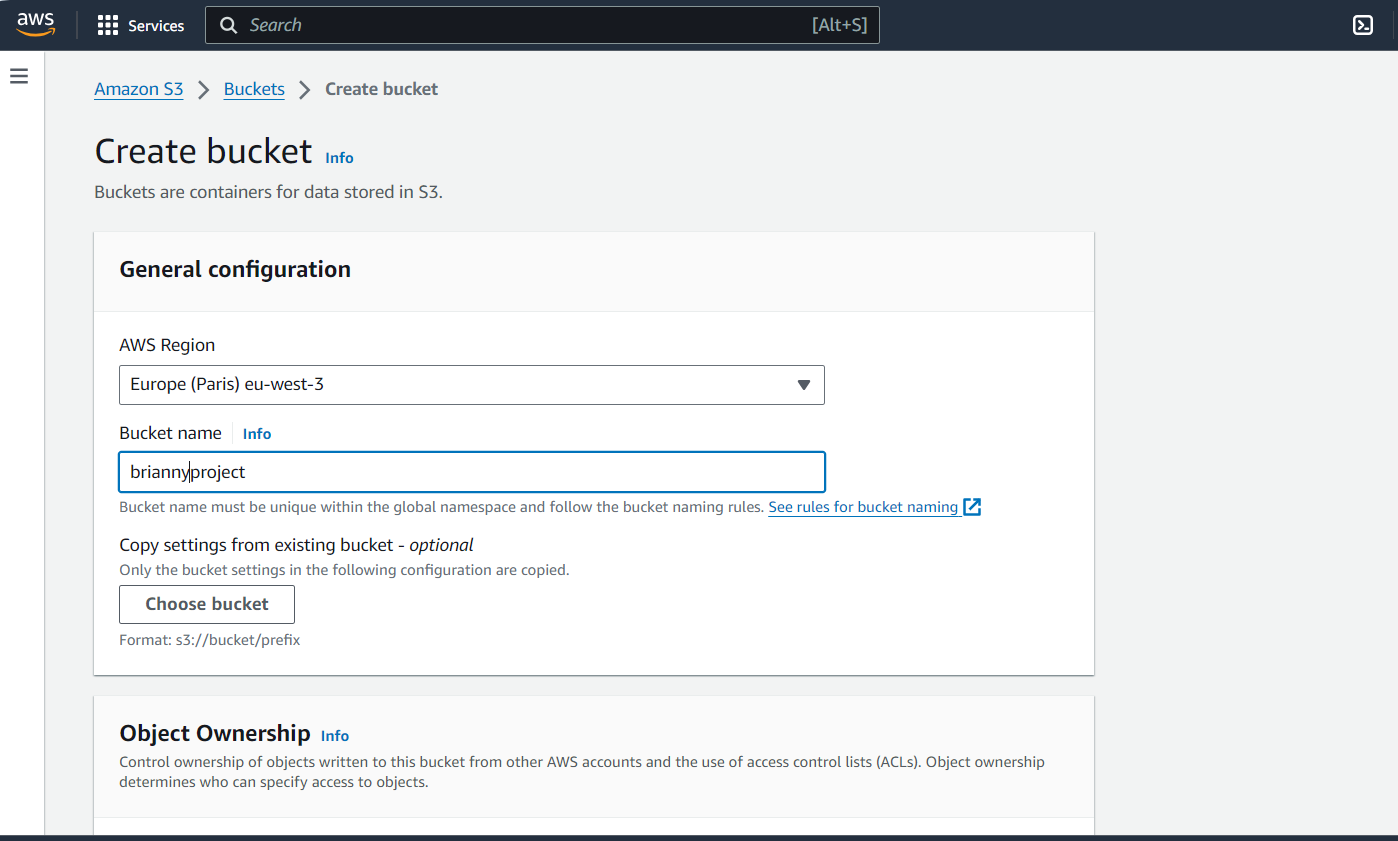


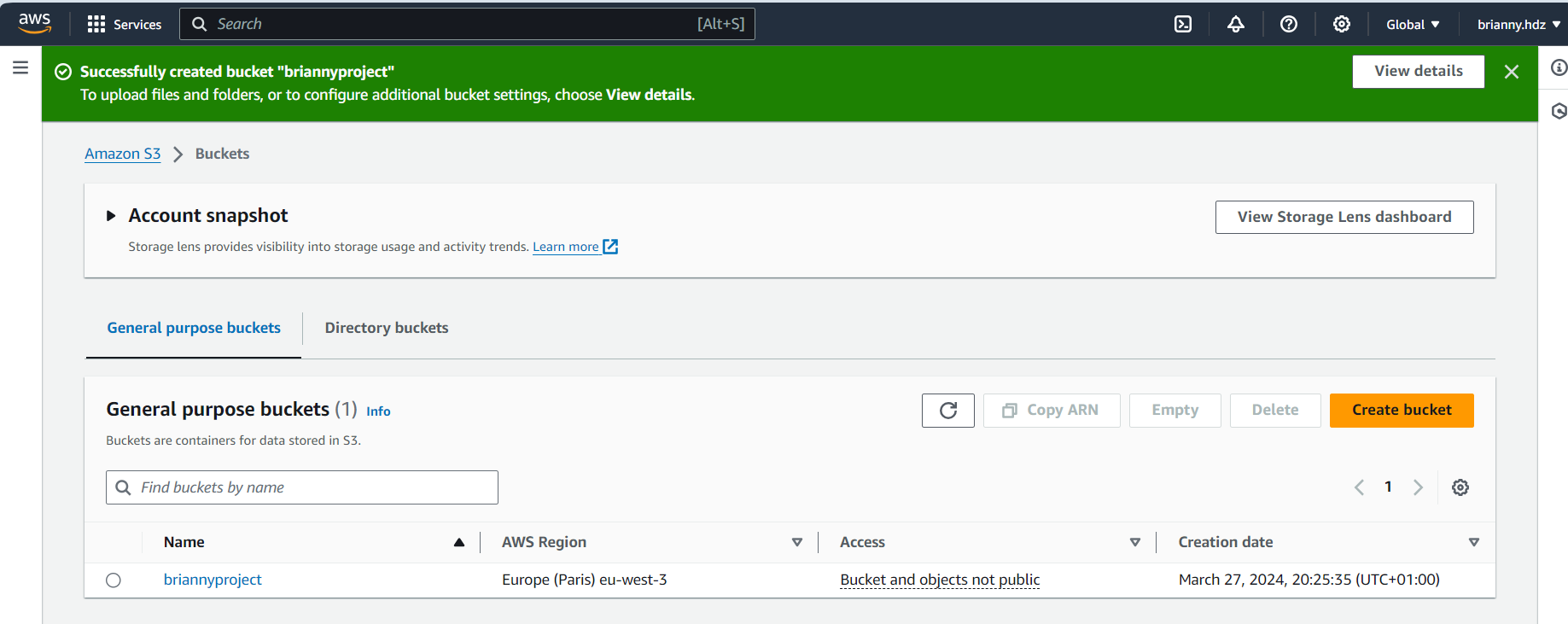
1. **Añadir un fichero de ventas en AWS, ventas.csv en un S3 propietario del alumno, conectarlo con el Data Federation de MongoDB haciendo que cambie el resultado final (o que se muestren datos nuevos).**

Como primer paso, accedemos a nuestra cuenta de AWS, tal como se muestra en la captura de pantalla adjunta a continuación.

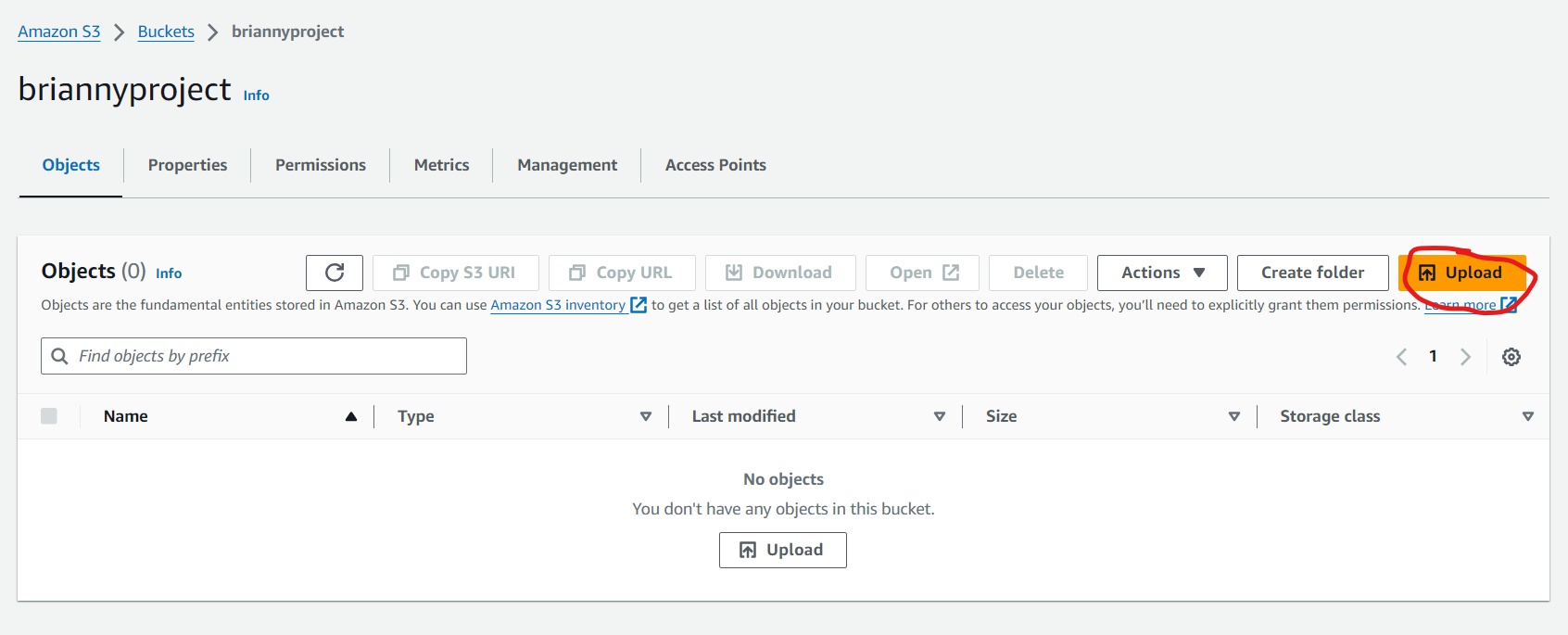


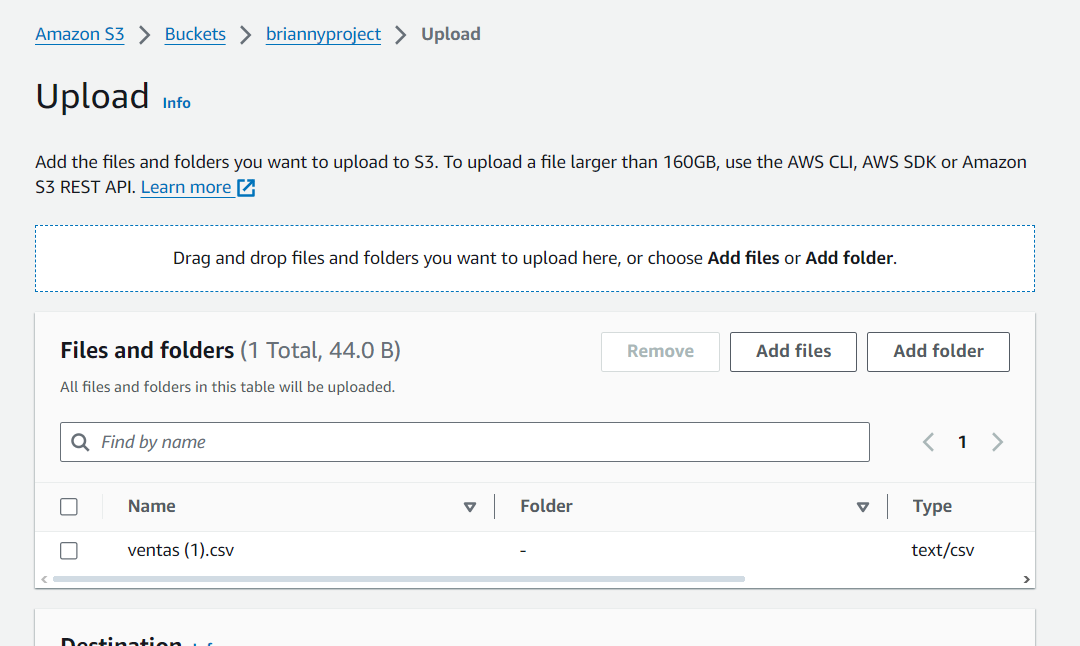
Buscamos en el buscador S3. Clicar en el botón Create bucket que aparece en la parte superior derecha.

Se nos abrirá una pantalla en la que indicaremos la región en la que crearemos el bucket (en mi caso he seleccionado París) y dejar los definidos por defecto.

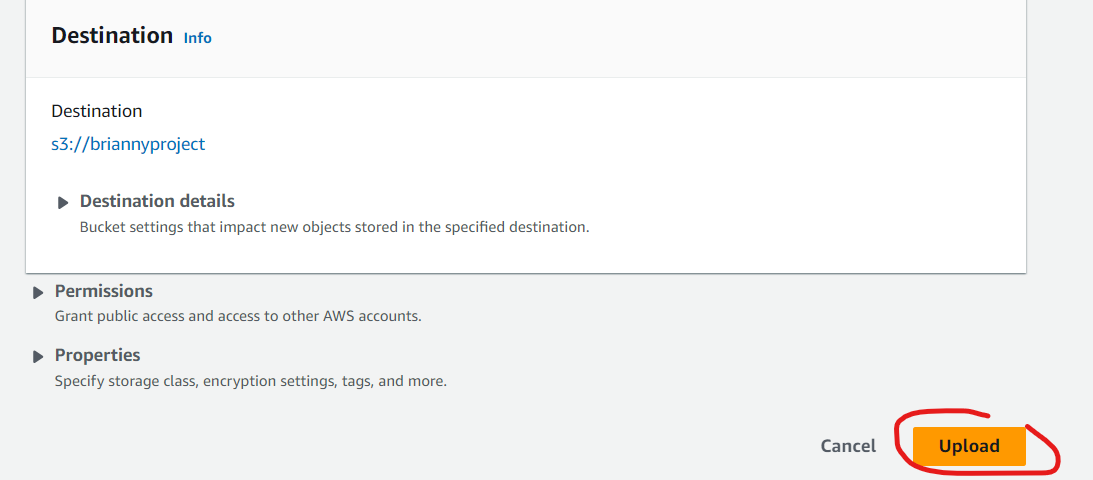
Asimismo, deberemos verlo en el listado de buckets.

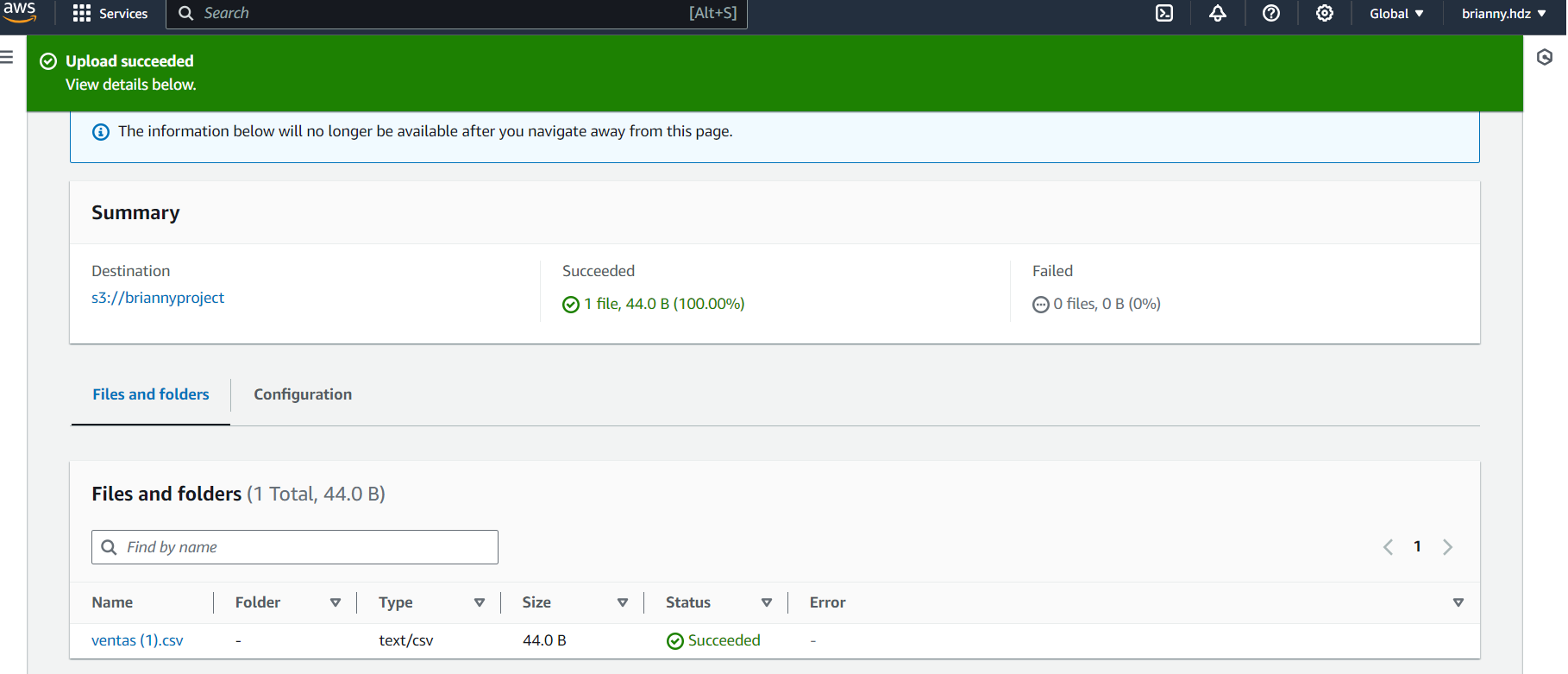
Una vez tenemos creado el bucket, podemos acceder a él clicando en el nombre. Una vez dentro del bucket, podemos subir archivos clicando en el botón Upload.



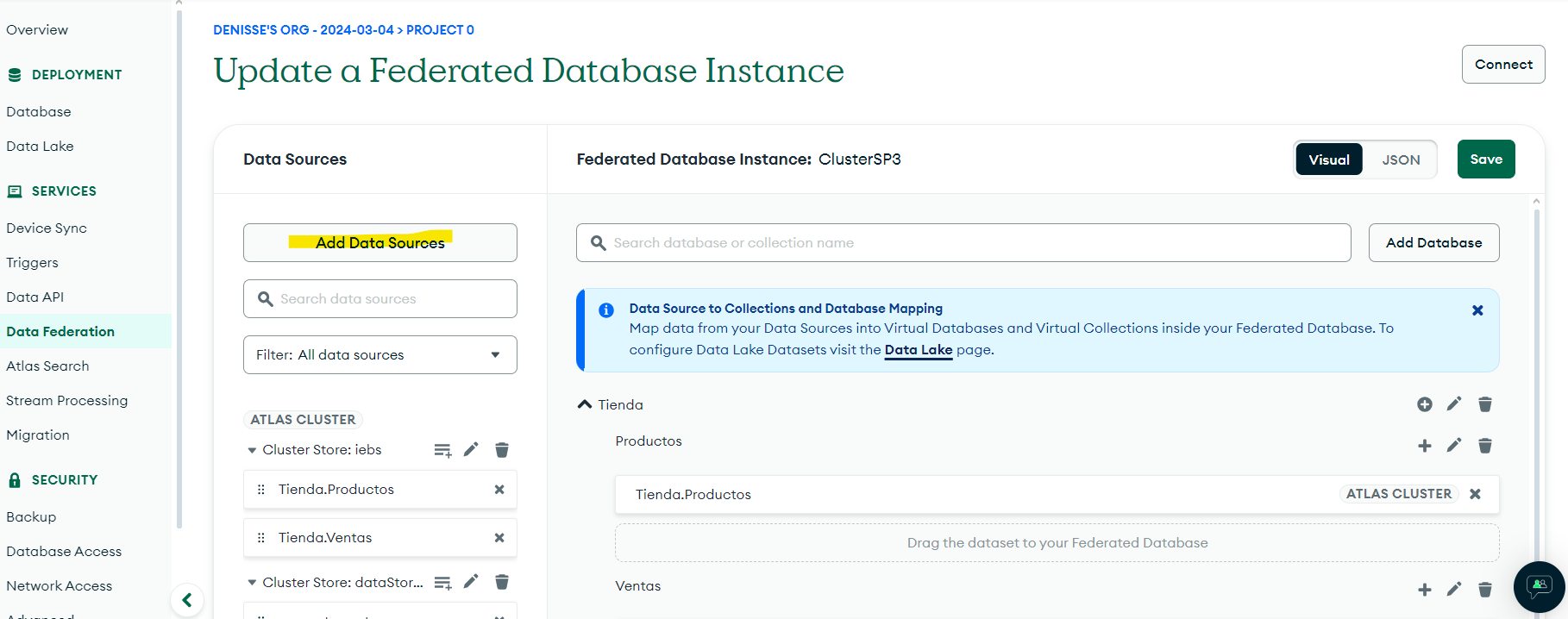


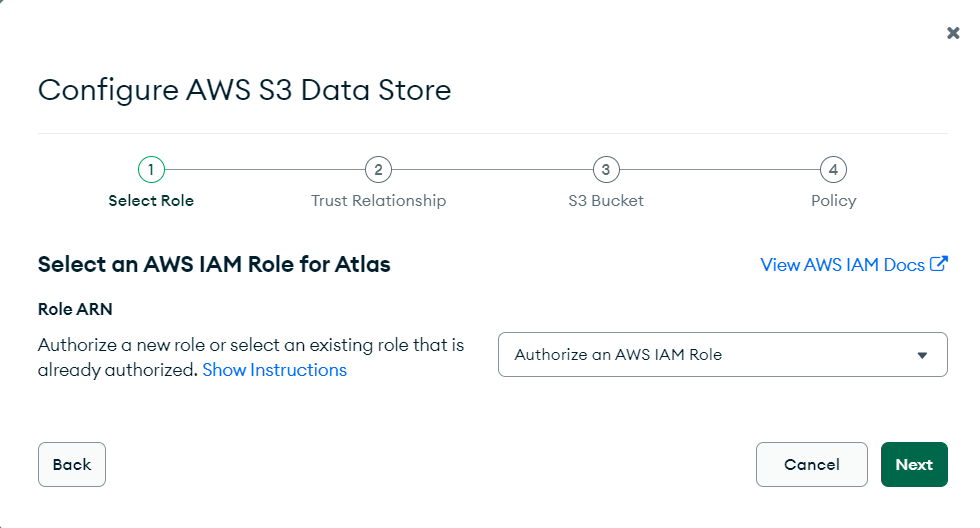
Una vez seleccionado el archivo que queremos subir, habrá que clicar en Upload y cuando se haya completado la subida nos aparecerá la siguiente pantalla de confirmación.



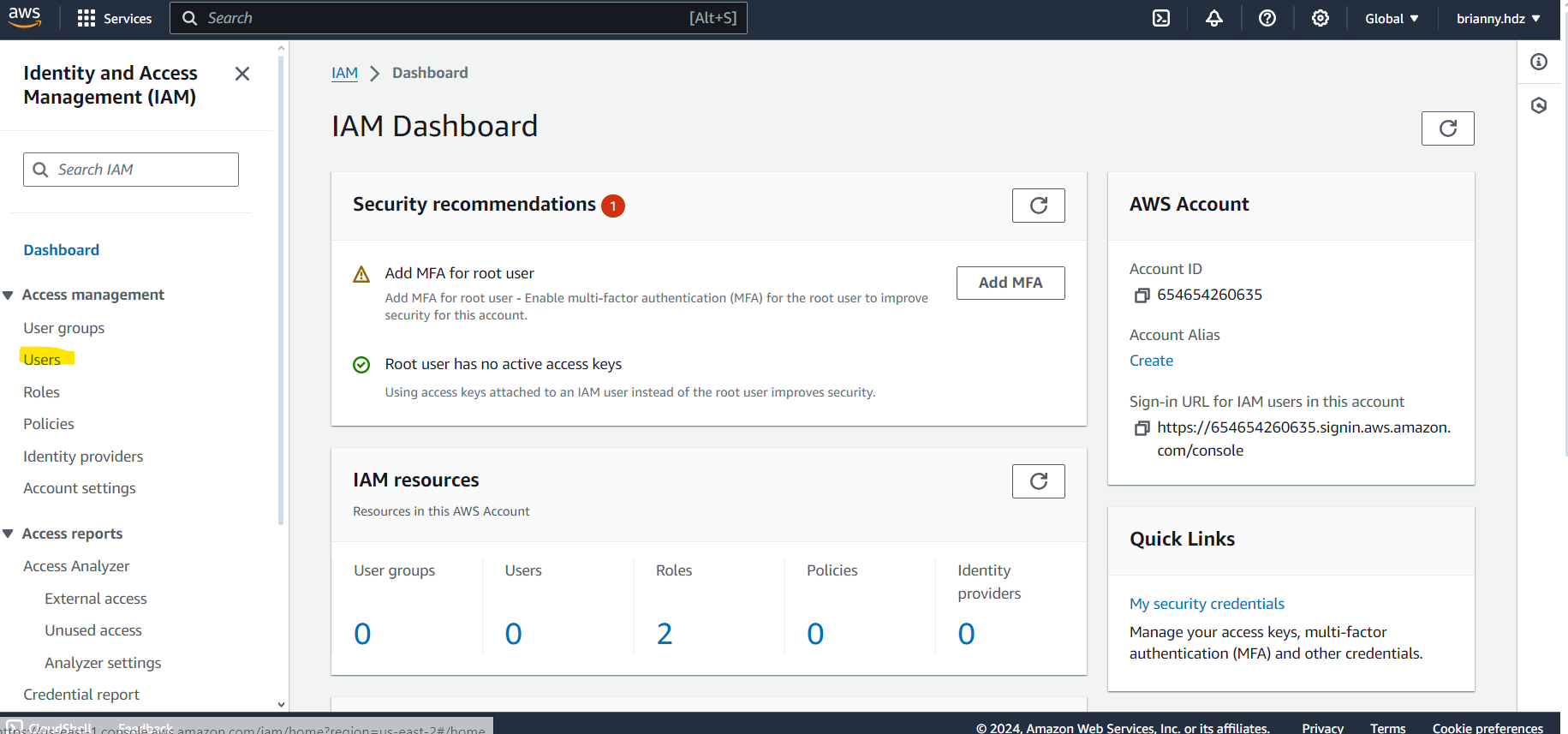


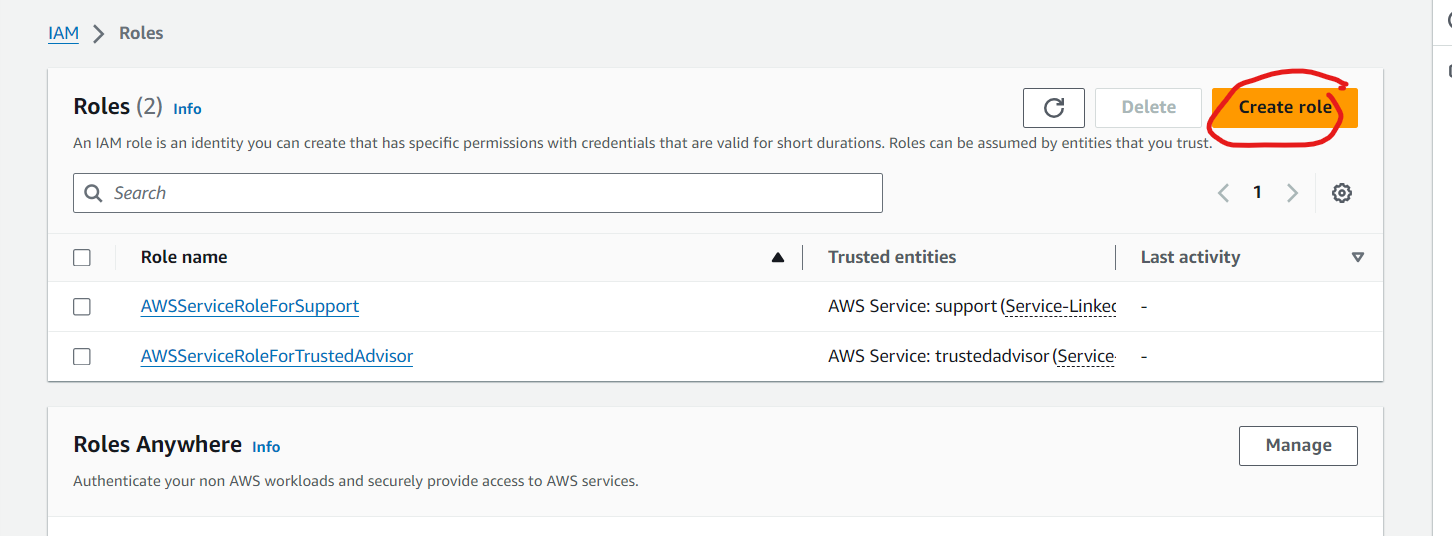
Conectar desde el Data Lake de MongoDB al S3

Una vez tenemos creado el bucket en S3 y el archivo subido, procedemos a conectar el data lake creado en Mongo Atlas con AWS S3.

Nos pedirá que se autorize un AWS IAM Role para que Mongo Atlas pueda acceder a nuestro S3. Clicamos en Next.

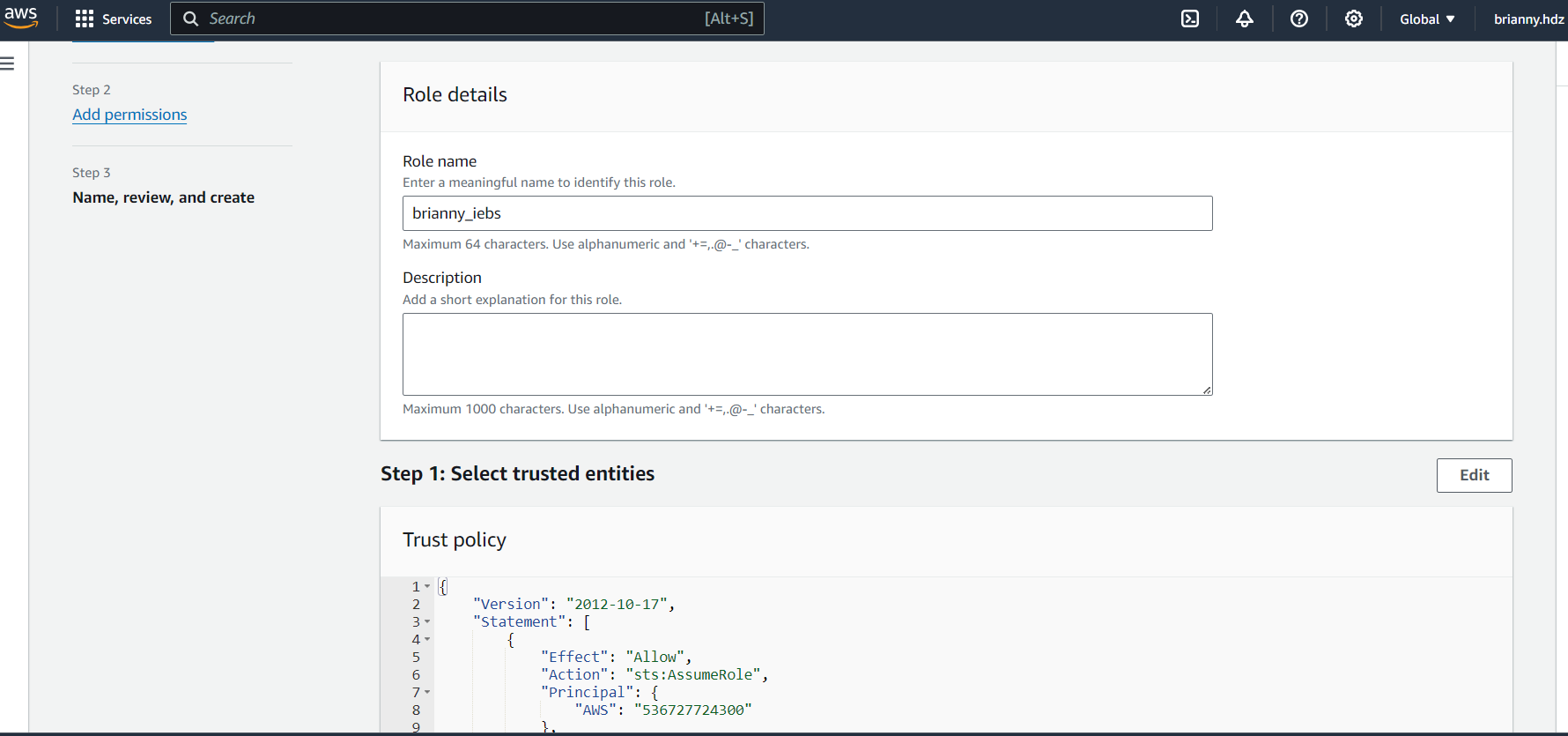
En AWS, accedemos a IAM y clicamos en Roles. Deberemos crear un nuevo rol con la información que nos proporcionará Mongo Atlas para crear la conexión entre ambos servicios.

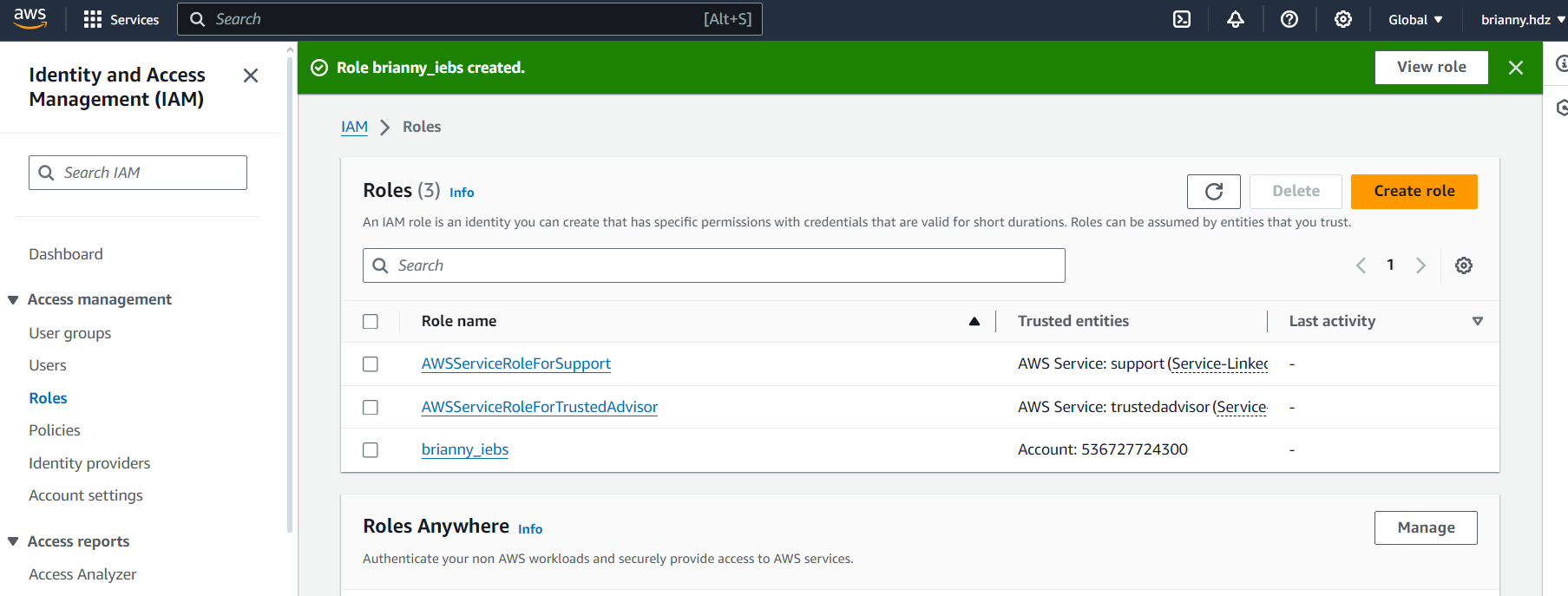




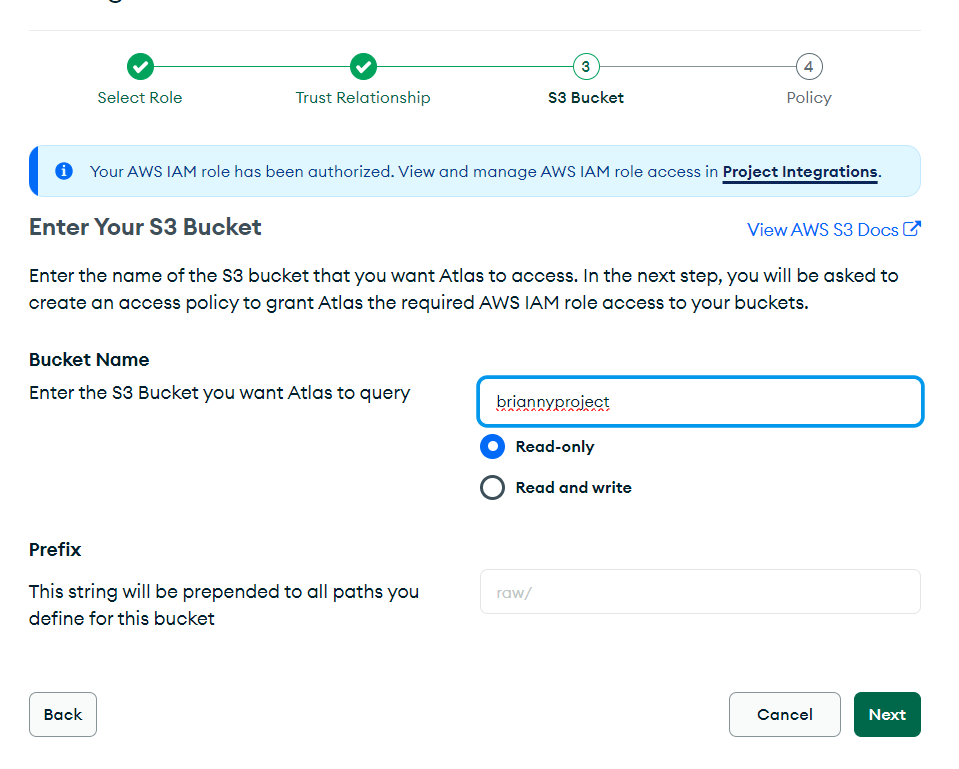
En Mongo Atlas, por otro lado, al clicar Next nos aparecerá la siguiente información que será la que utilizaremos para crear la conexión.

Después de configurar el role, le asignamos un nombre y guardamos.

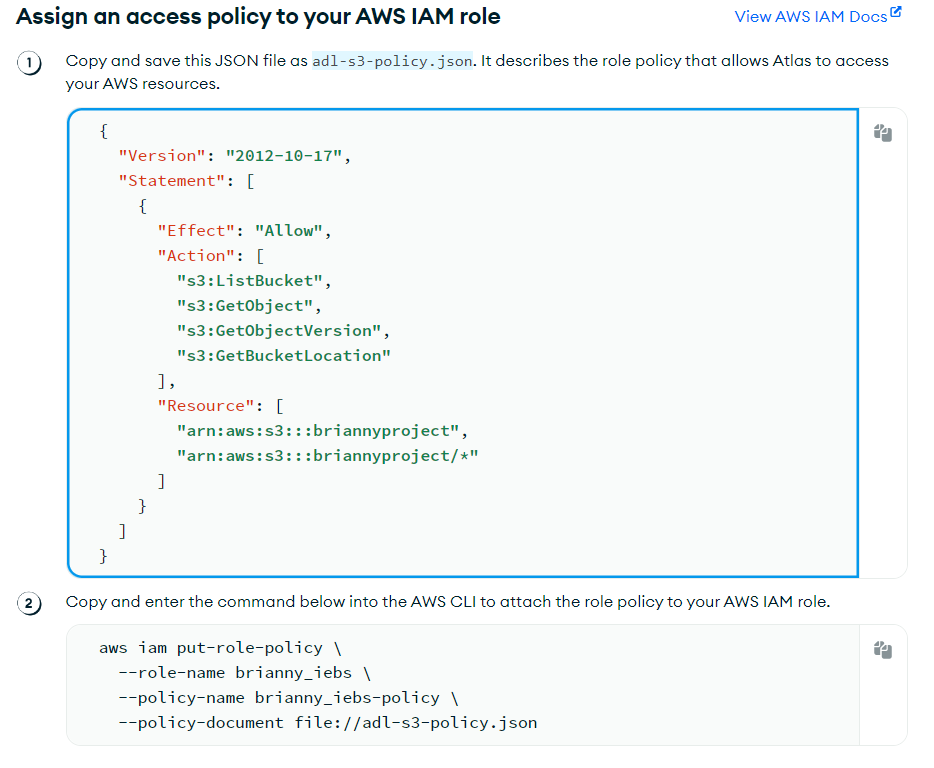


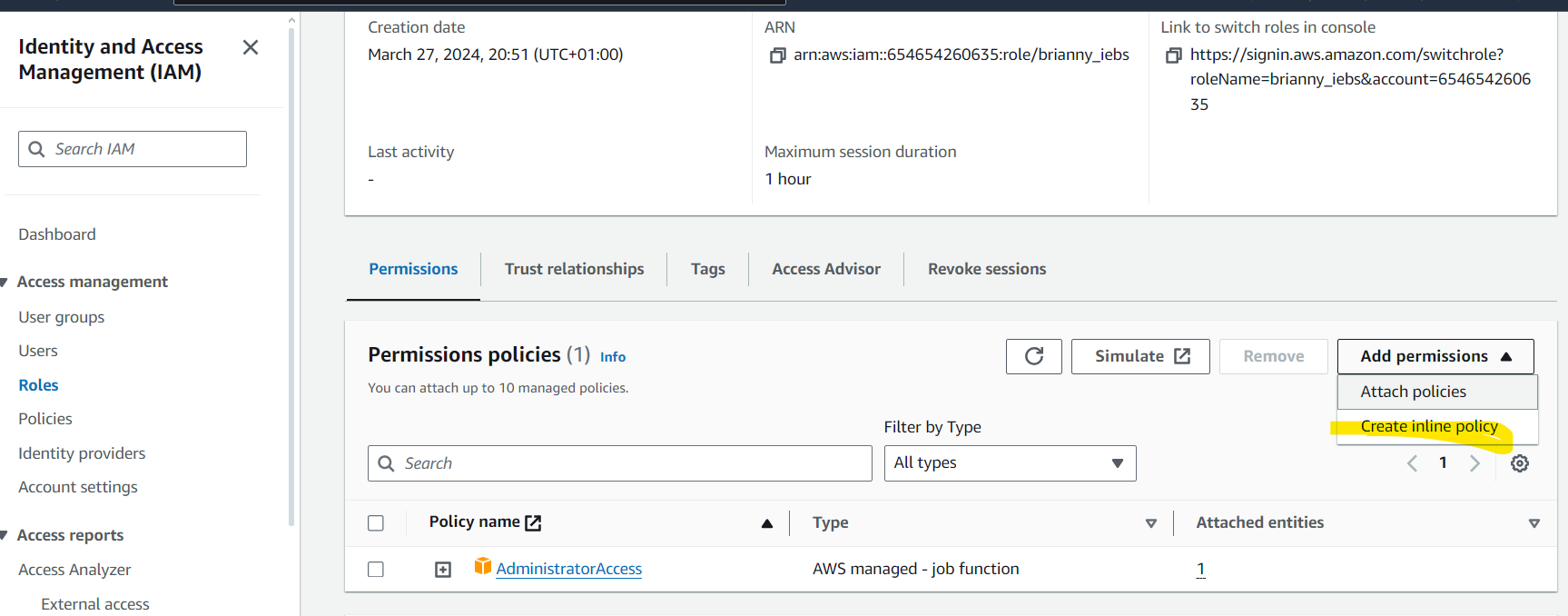


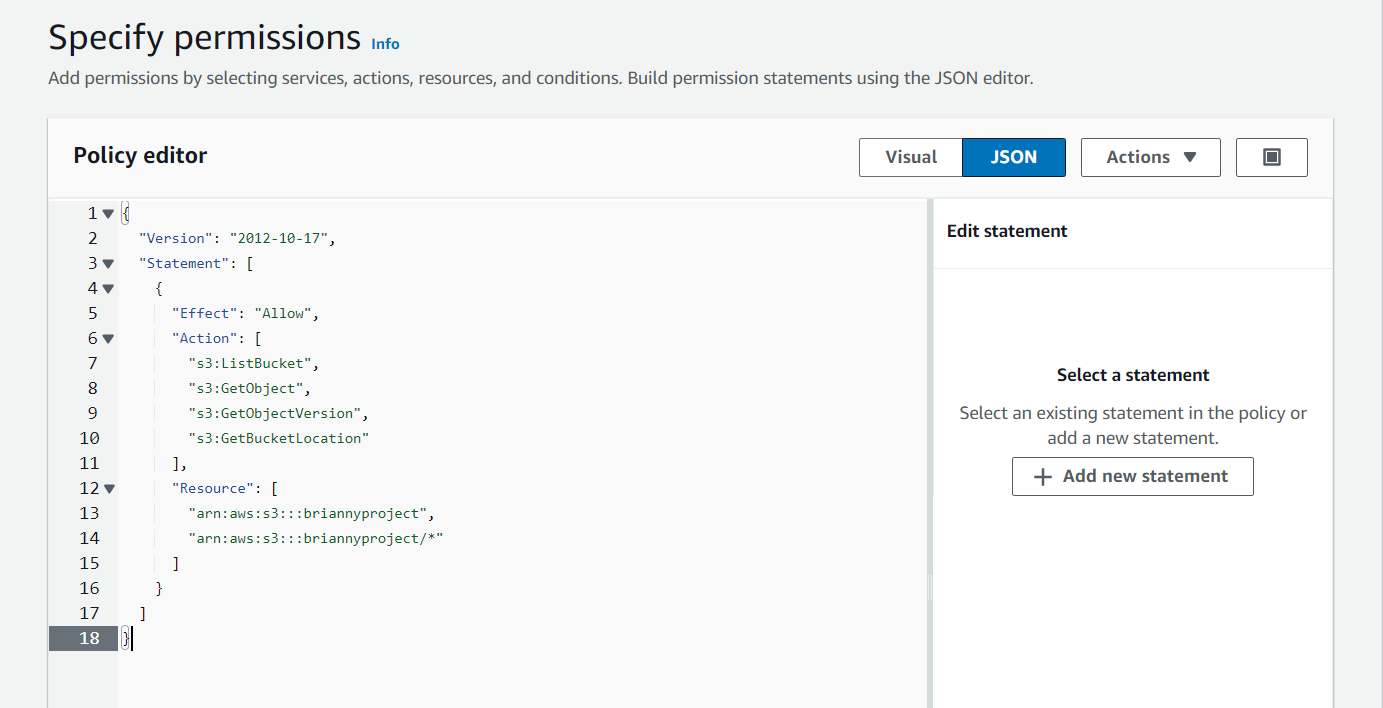
Indicamos el nombre del bucket al que queremos que acceda y clicamos en Next.

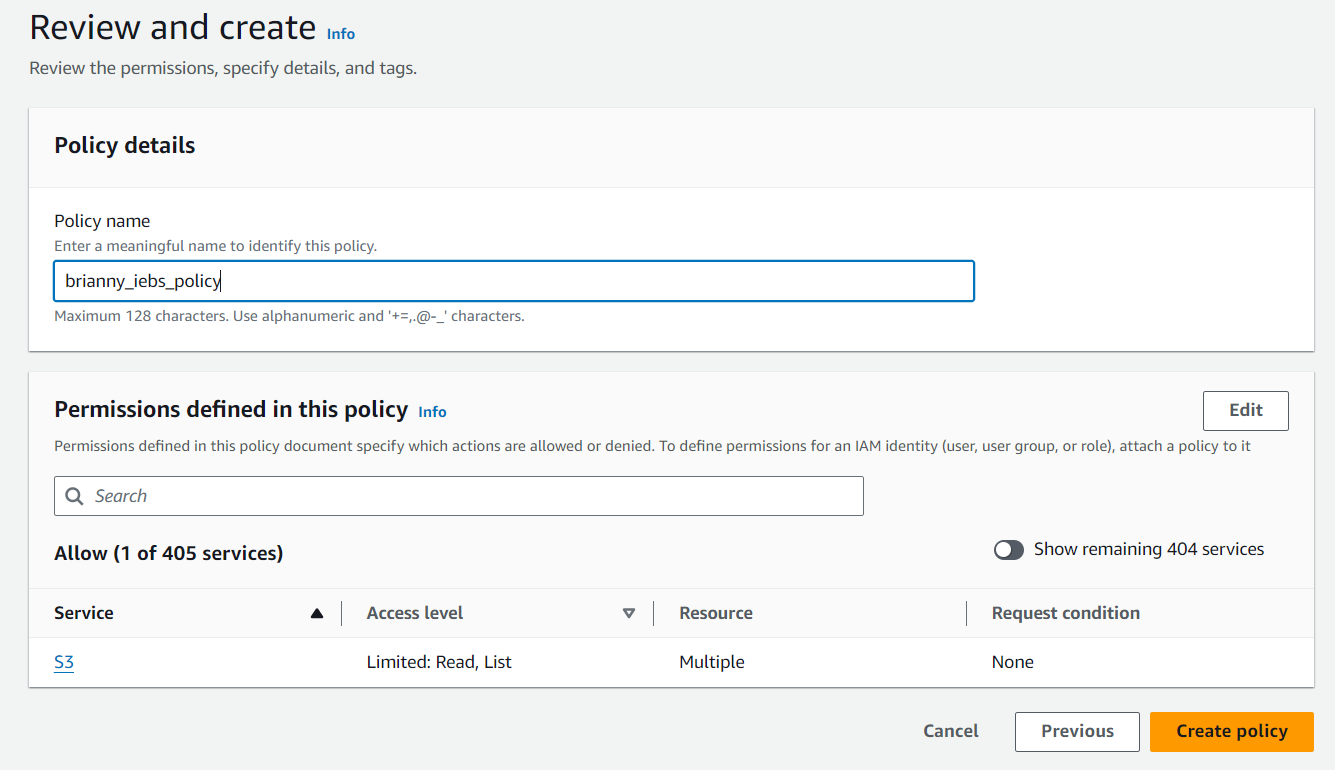
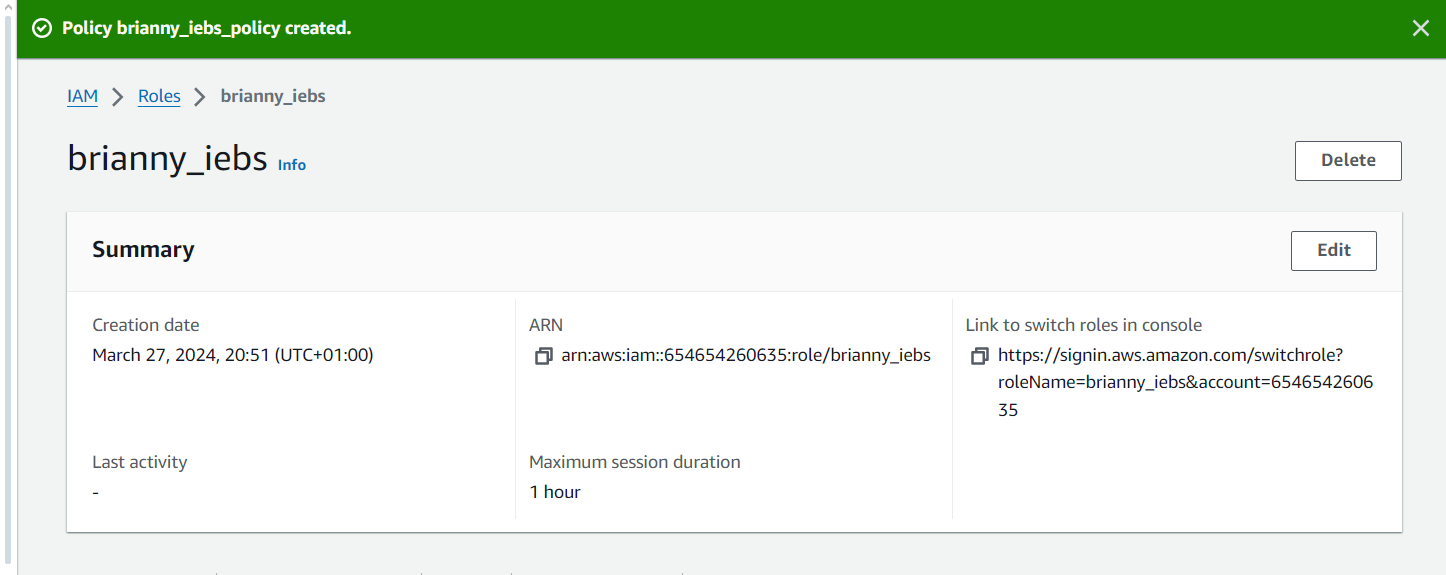


En la siguiente pantalla, nos aparecerá la S3-policy en formato JSON que deberemos copiar y crearla y añadirla al rol que acabamos de crear.

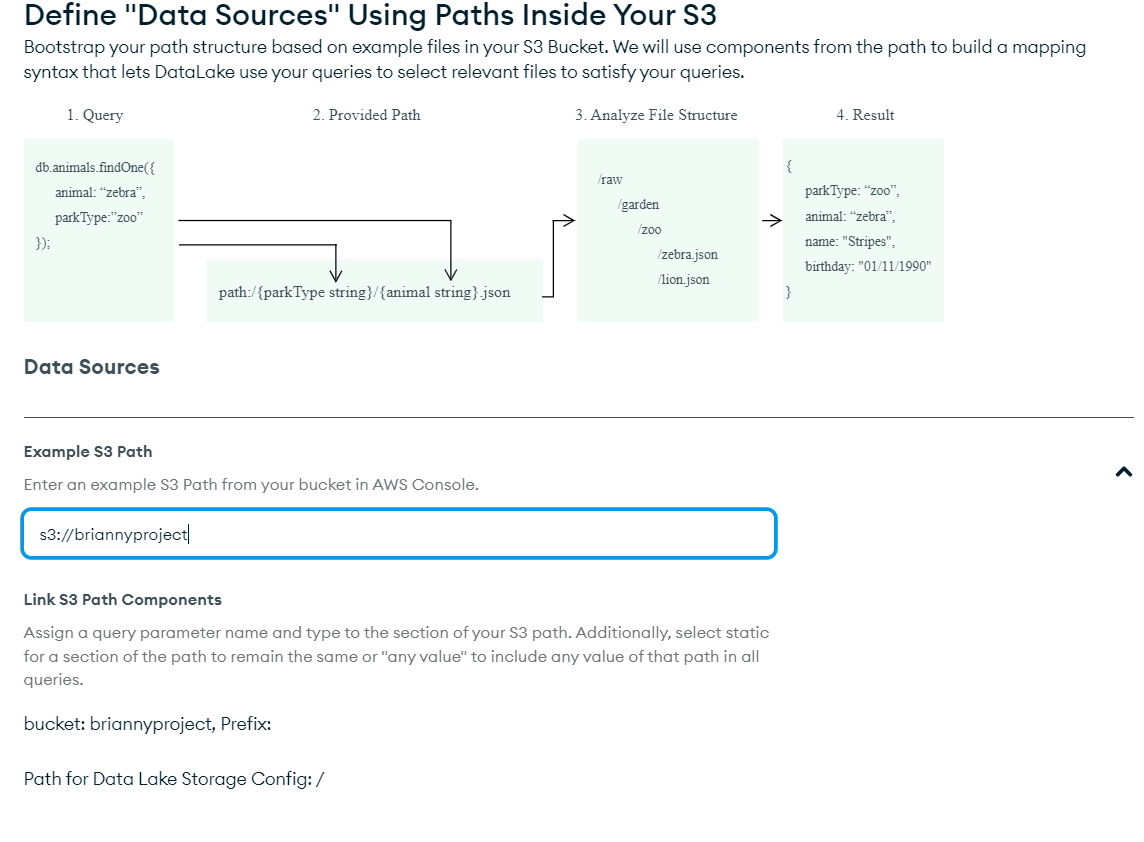


En AWS IAM, en la pantalla de Roles y en concreto dentro del rol que acabamos de crear, clicaremos en Create inline policy. Se nos abrirá la pantalla que se muestra a continuación y simplemente deberemos pegar el código JSON que acabamos de copiar en Mongo Atlas.

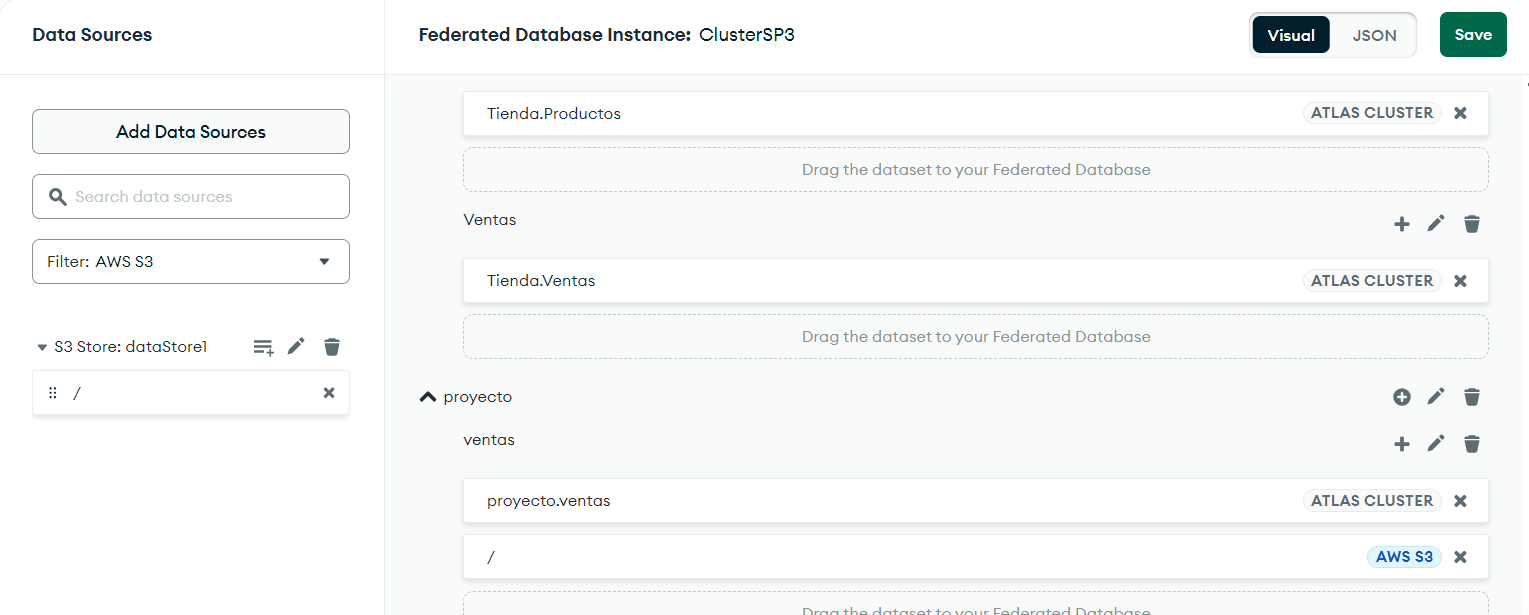
Indicamos un nombre para la policy que acabamos de crear y clicamos en Create policy. Automáticamente se nos añadirá a nuestro rol, tal como se puede observar en la segunda pantalla.



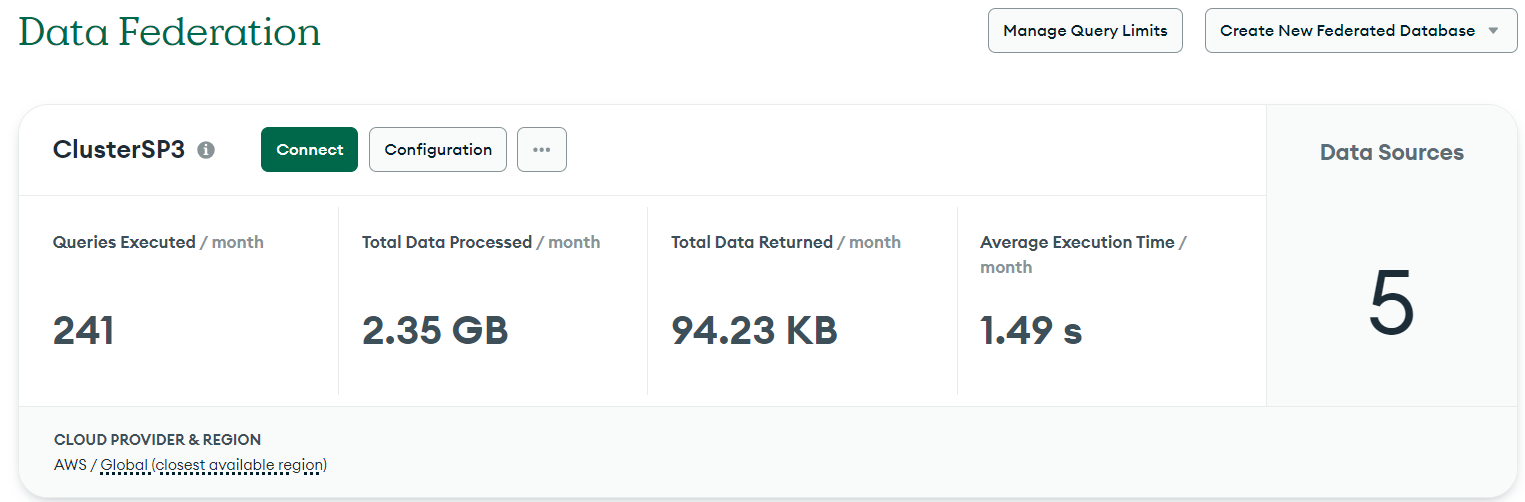
De vuelta a Mongo Atlas, se nos pedirá el path del bucket en S3.



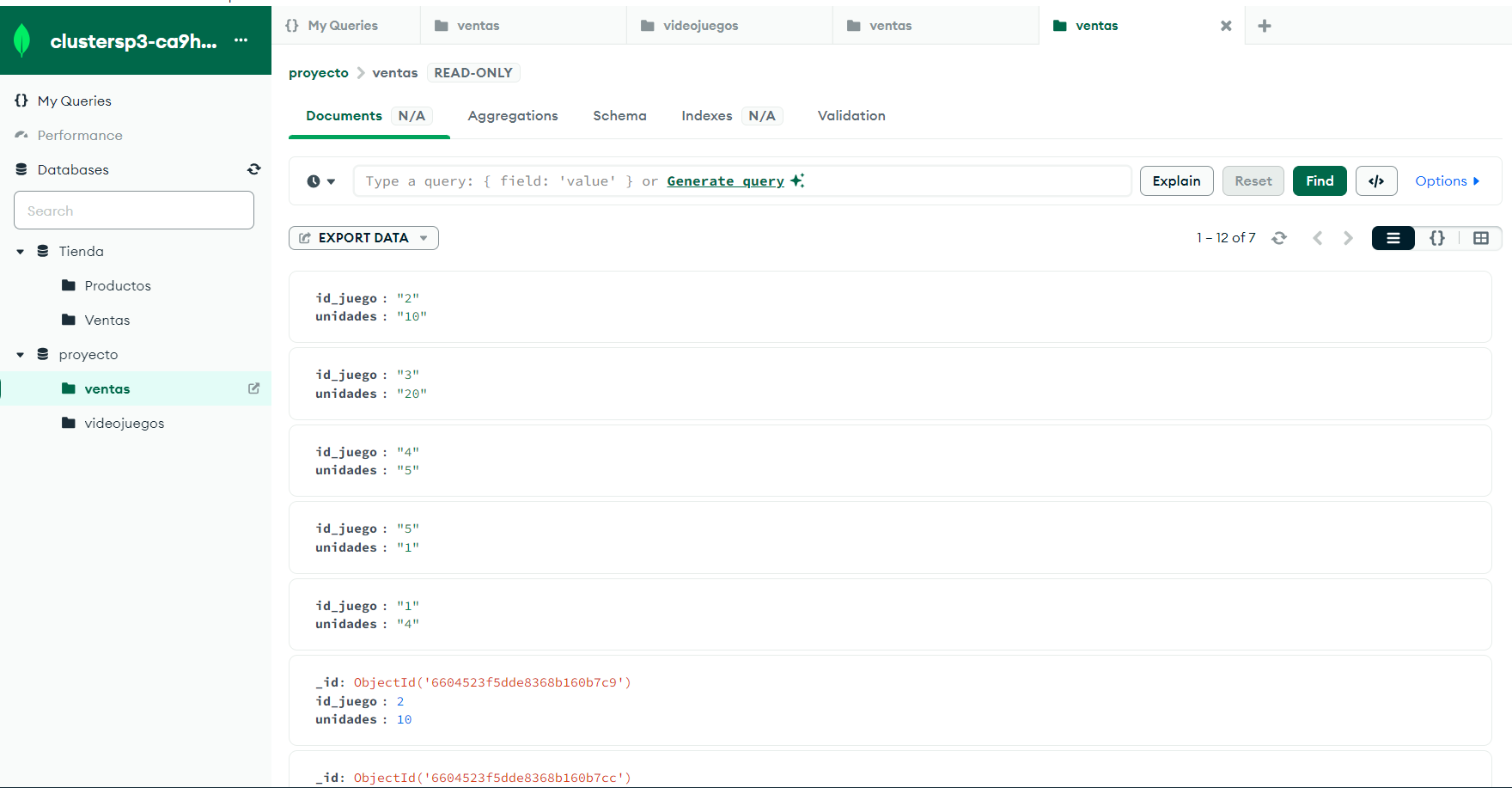
Una vez que hemos seguido los pasos anteriores, veremos que nos aparece AWS S3 como fuente de datos. Ahora podemos añadirla a la colección que siga el mismo esquema, en este caso Ventas.

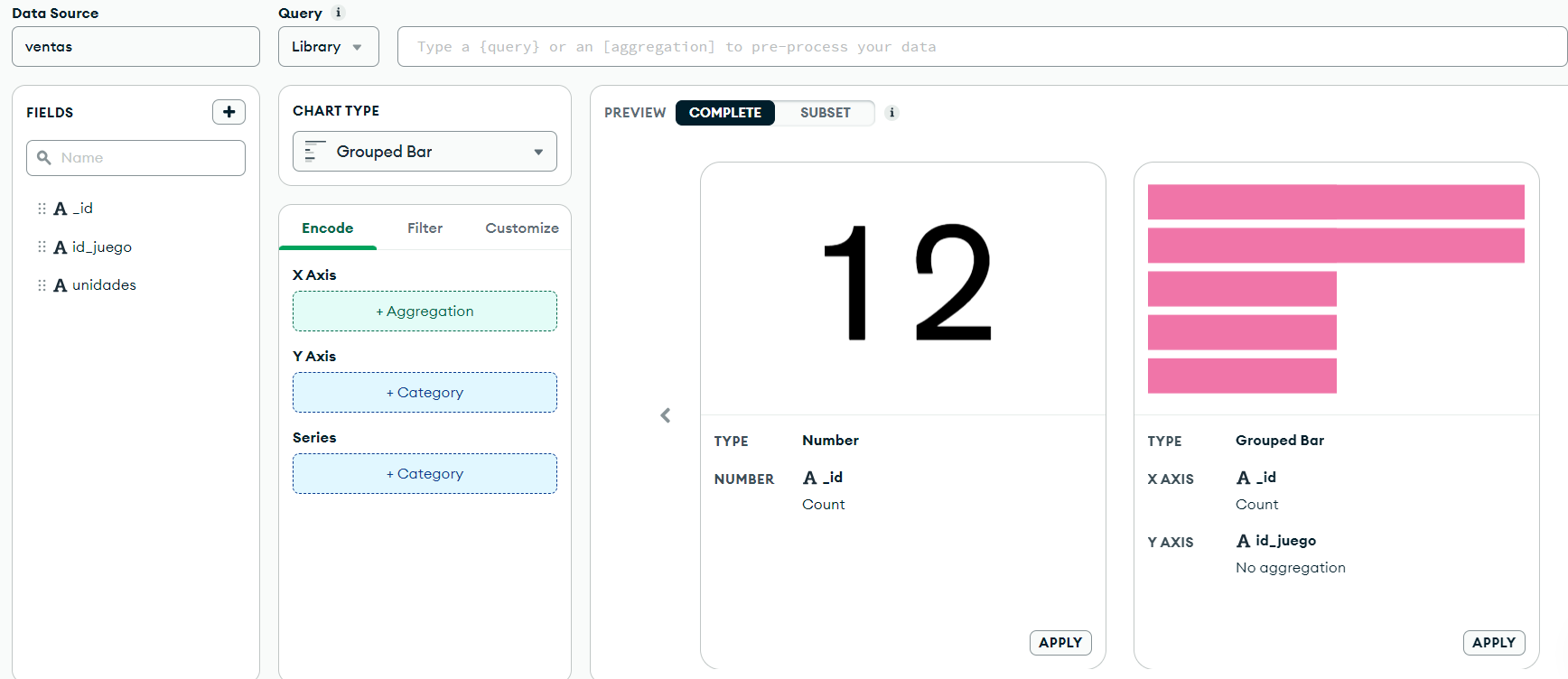


Guardamos y visualizamos los cambios.



Query con nuevos datos.



1. **Fase adicional: Crear gráficos a partir de los datos creados mejorando la comprensión de las conclusiones.**

