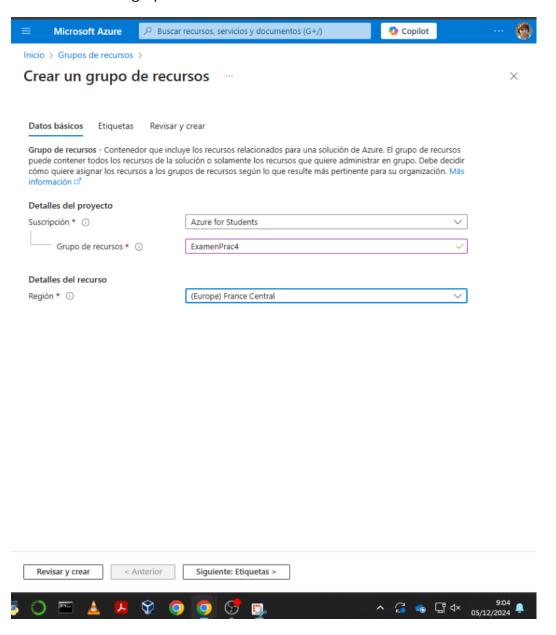
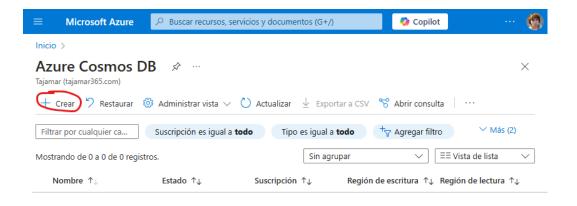
Paso 1: Hacer un grupo de recursos de Azure



Paso 2: Crear en Azure una CosmosDB





No hay Cuentas de Azure Cosmos DB para mostrar

Cree una base de datos completamente administrada, de varios modelos y distribuida globalmente mediante la API que prefiera. O pruebe el servicio de forma gratuita durante 30 días con hasta 20 RU/s y renovación ilimitada.

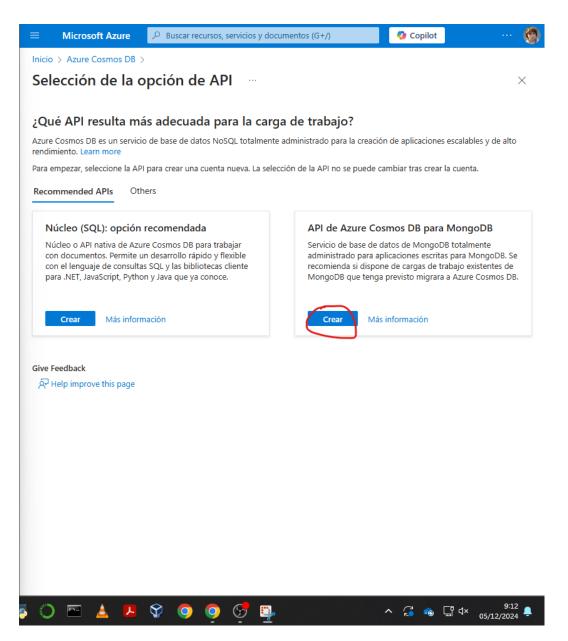
Crear Cuenta de Azure Cosmos DB

Probar ahora 🗗

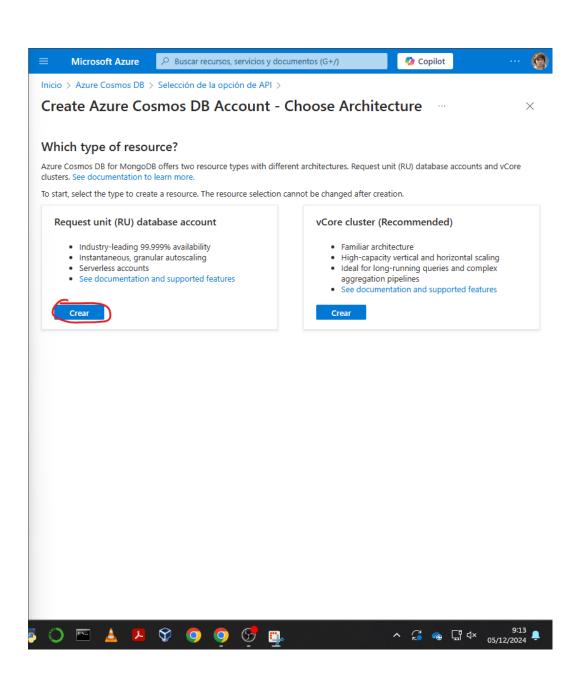
Enviar comentarios

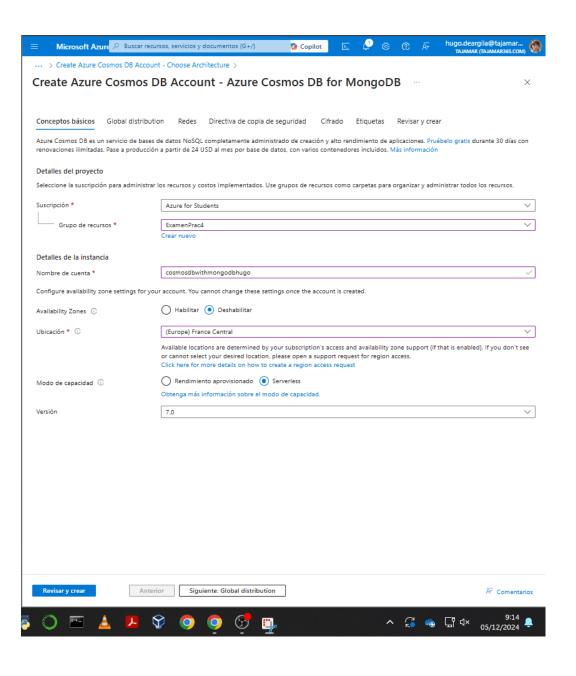


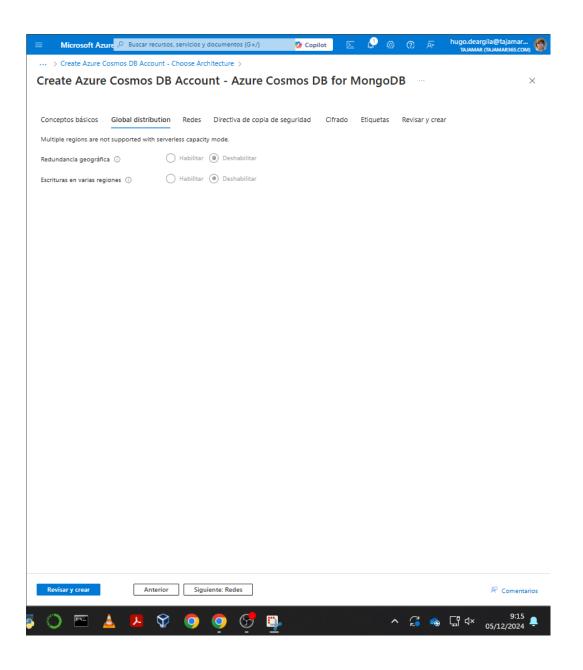
Api:

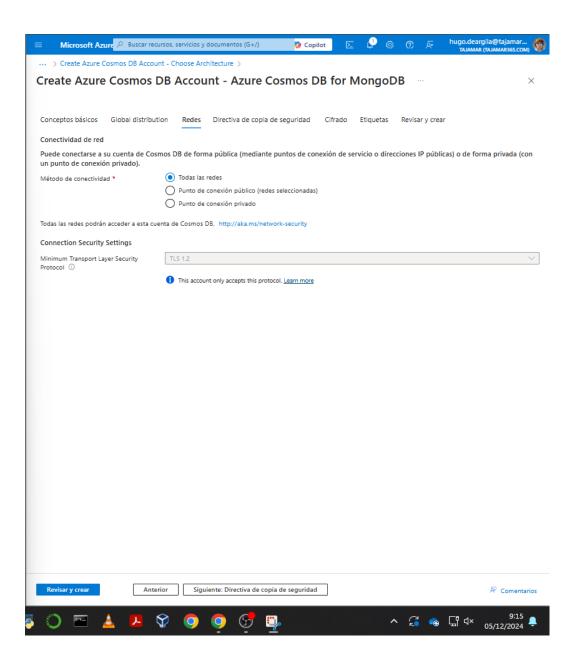


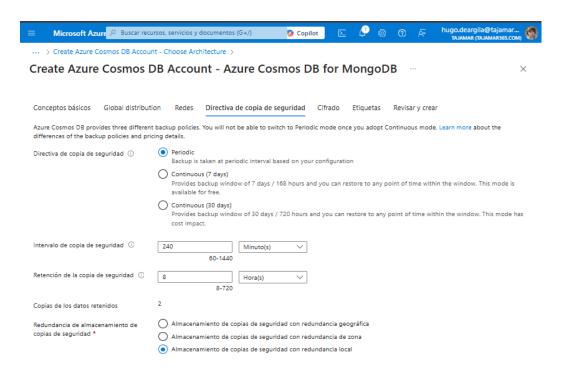
Arquitecture:

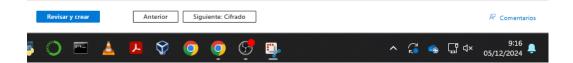


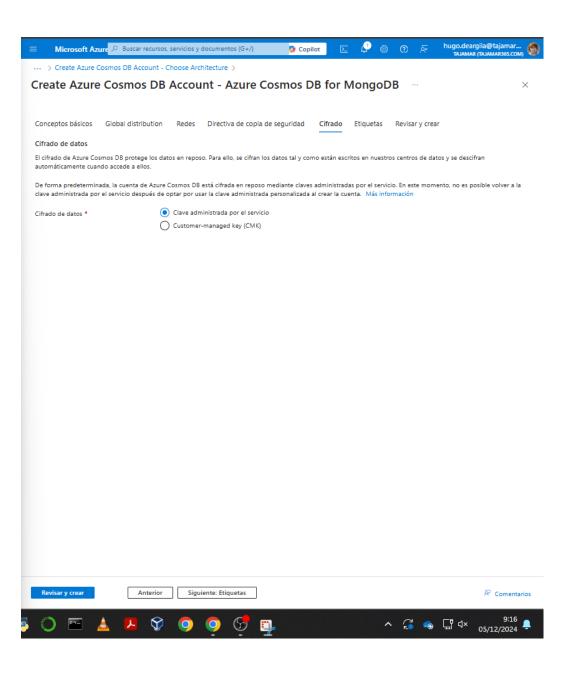


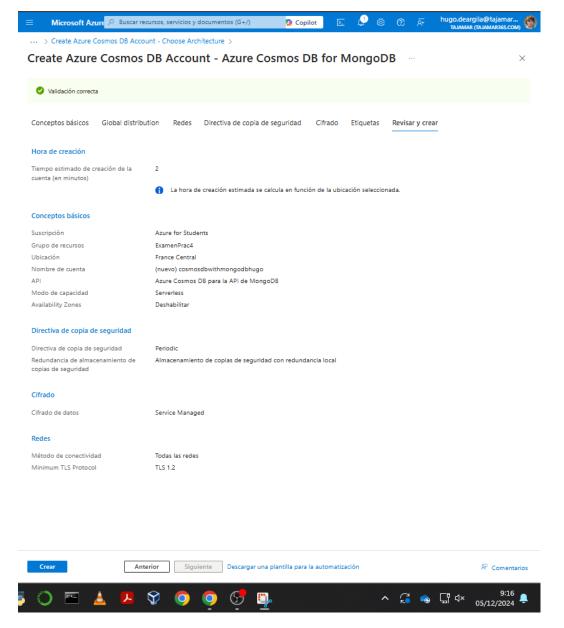












Paso 3: Crear el Entorno de Desarrollo

- 1. Abre la terminal (o PowerShell en Windows) y navega al directorio donde deseas crear tu proyecto.
- 2. Ejecuta el siguiente comando para crear un entorno virtual: python -m venv venv Esto creará una carpeta llamada venv en tu directorio actual.

```
● PS C:\Users\Alumno_AI\Documents\ExamenPrac4> python -m venv venv
```

Luego activamos el entorno de desarrollo:

```
PS C:\Users\Alumno_AI\Documents\ExamenPrac4> .\venv\Scripts\activate(venv) PS C:\Users\Alumno_AI\Documents\ExamenPrac4> []
```

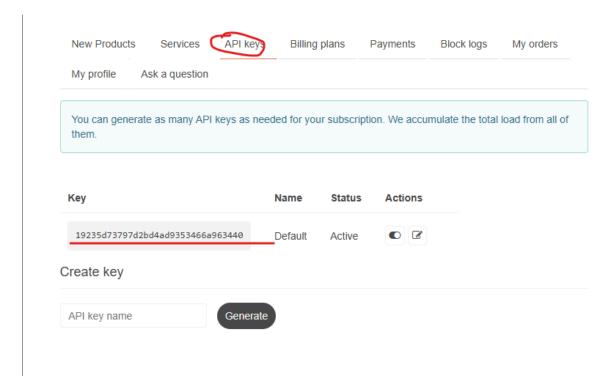
Instalamos dependencias:

```
(venv) PS C:\Users\Alumno_AI\Documents\ExamenPrac4> pip install pymongo requests
```

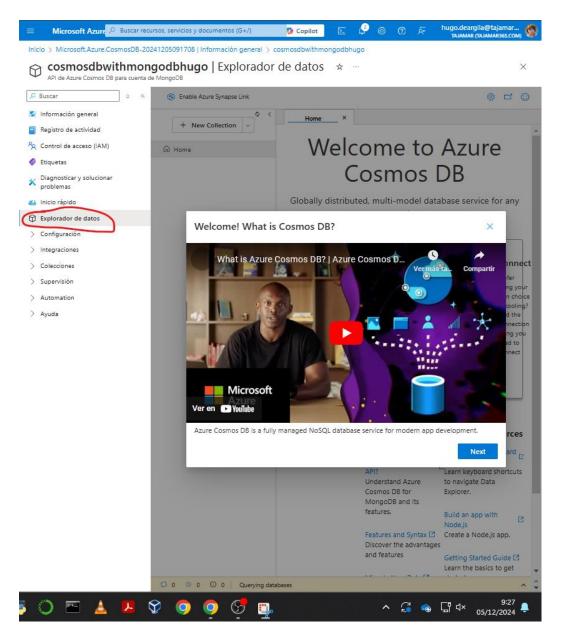
Comprobamos la instalación:

```
(venv) PS C:\Users\Alumno AI\Documents\ExamenPrac4> pip list
 Package
                    Version
 certifi
                    2024.8.30
 charset-normalizer 3.4.0
 dnspython
                    2.7.0
 idna
                    3.10
 pip
                    24.0
                    4.10.1
 pymongo
 requests
                    2.32.3
 urllib3
 [notice] A new release of pip is available: 24.0 -> 24.3.1
 [notice] To update, run: python.exe -m pip install --upgrade pip
○ (venv) PS C:\Users\Alumno_AI\Documents\ExamenPrac4> |
```

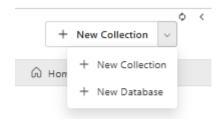
Paso 4: Vamos a https://openweathermap.org/api y obtenemos nuestra key



Paso 5: Ir a Explorador de datos de Cosmos DB

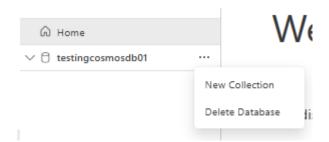


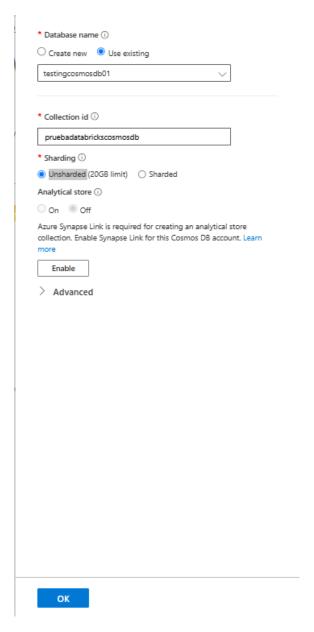
Creamos una New Database





Creamos una New Collection:





Paso 6: Conseguir datos

Primero debemos añadir este script a nuestro entorno virtual, tendremos que cambiar la API_KEY por nuestra key de Openweather y los datos de la BBDD por los de nuestra BBDD:

from pymongo import MongoClient
import requests
import json
import time

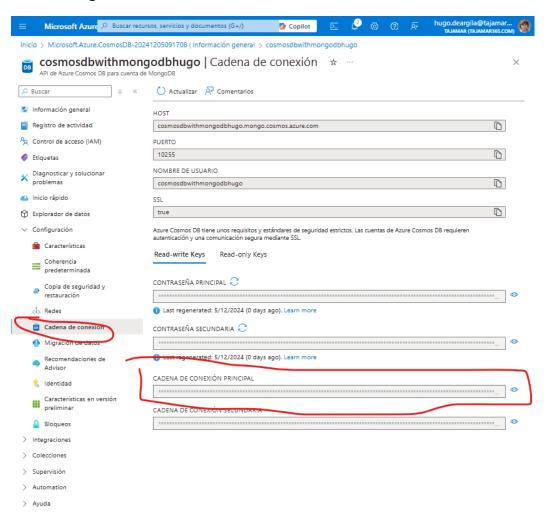
Configuración de OpenWeather API
API_KEY = "8b73515bcec2ae54c33e6d422f504f2a" # Reemplaza con tu API Key
CITY = "Toronto" # Cambia por la ciudad deseada

```
WEATHER_URL =
f"http://api.openweathermap.org/data/2.5/weather?q={CITY}&appid={API_KEY}&unit_
s=metric"
# Configuración de Cosmos DB
DB_NAME = "testingcosmosdb01" # Nombre de tu base de datos
COLLECTION NAME = "pruebadatabrickcosmosdb" # Nombre de tu colección
CONNECTION =
"mongodb://cosmosdbwithmongodb:fUTd69EFQvoWifX0yn4sBTJ3e0yJoMaRAC3hUri
FSZnAKlheVJwezrgr01Jw5GVnFjLWnho3YkRcACDbPhcU3g==@cosmosdbwithmongo
db.mongo.cosmos.azure.com:10255/?ssl=true&replicaSet=globaldb&retrywrites=fal
se&maxIdleTimeMS=120000&appName=@cosmosdbwithmongodb@" # Conexión a
Cosmos DB
# Función para insertar datos en Cosmos DB
def insert_weather_data(client, data):
 db = client[DB_NAME]
 collection = db[COLLECTION NAME]
 document_id = collection.insert_one(data).inserted_id
 print(f"Documento insertado con id: {document id}")
# Función para obtener datos meteorológicos
def get_weather_data():
 try:
   response = requests.get(WEATHER_URL)
   if response.status code == 200:
     weather_data = response.json()
     # Formatear los datos
     formatted_data = {
       "city": weather_data["name"],
       "temperature": weather_data["main"]["temp"],
       "weather": weather_data["weather"][0]["description"],
       "humidity": weather_data["main"]["humidity"],
       "pressure": weather_data["main"]["pressure"],
       "wind_speed": weather_data["wind"]["speed"],
       "timestamp": time.strftime('%Y-%m-%d %H:%M:%S')
     }
     return formatted data
   else:
     print(f"Error al obtener los datos: {response.status code}")
     return None
 except Exception as e:
   print(f"Error al conectar con la API: {e}")
   return None
# Bucle infinito para recopilar e insertar datos
while True:
```

```
client = MongoClient(CONNECTION)
weather_data = get_weather_data()
if weather_data:
    print("Datos obtenidos del clima:")
    print(weather_data)
    insert_weather_data(client, weather_data)
    print("Datos insertados exitosamente en Cosmos DB")
else:
    print("No se pudieron obtener los datos del clima.")

# Pausa entre iteraciones (5 minutos)
time.sleep(20)
```

Para conseguir la WEATHER_URL:

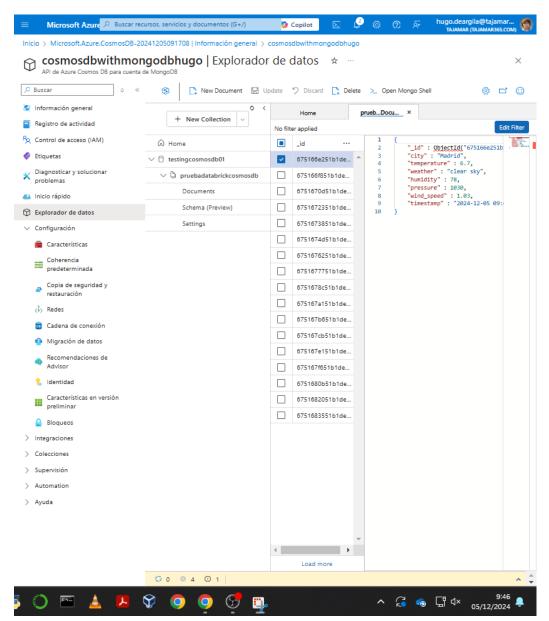




Ahora ejecutaremos el script con el comando:

(venv) PS C:\Users\Alumno_AI\Documents\ExamenPrac4> python weather_to_cosmosdb.py
>> []

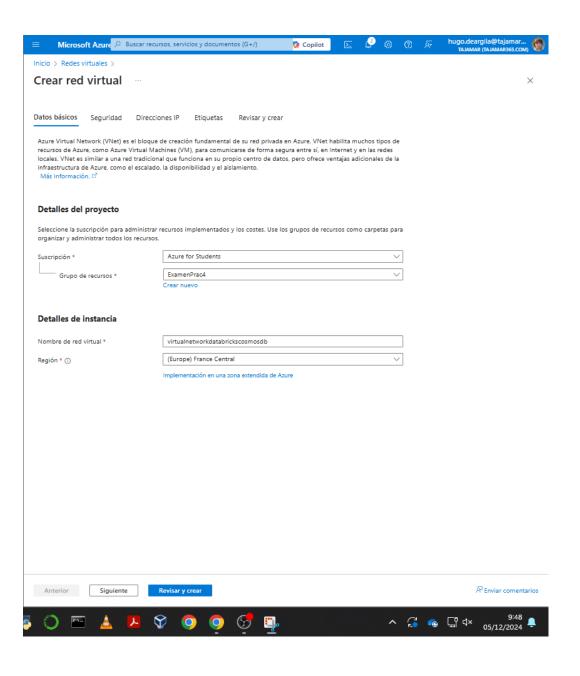
Se nos cargaran los datos en nuestro container:

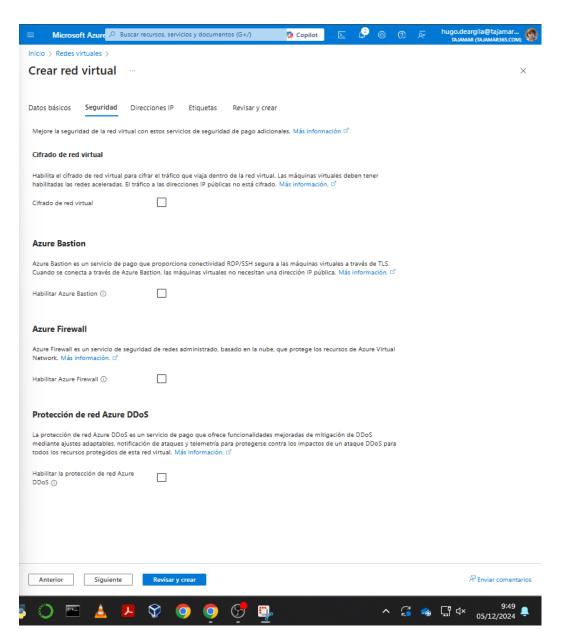


Para parar el indexado de datos usaremos Ctrl +C:

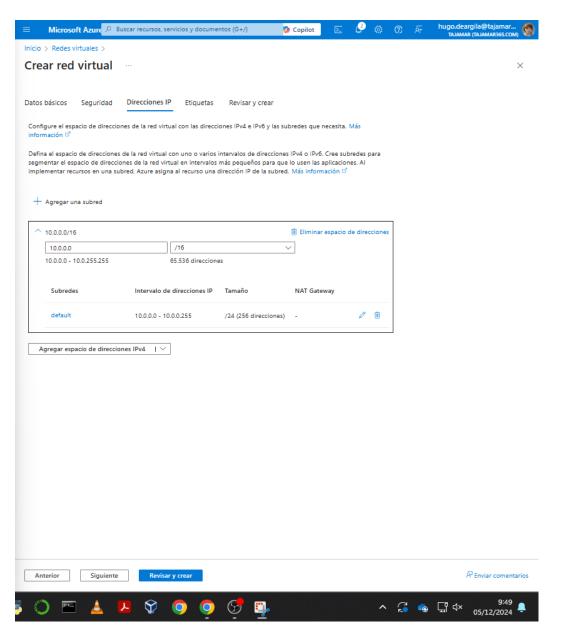
```
time.sleep(20)
KeyboardInterrupt
```

Paso 7: Crear una Red Virtual de Azure

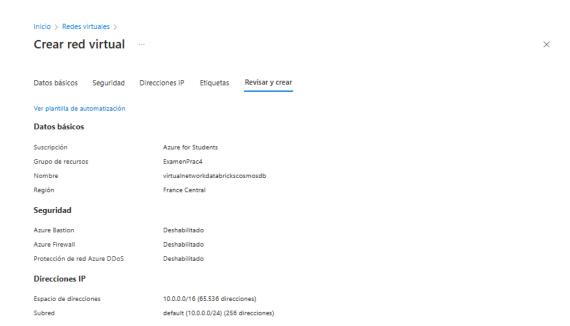




No marcar nada.



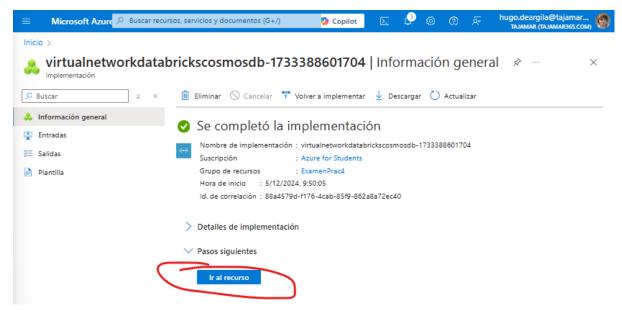
Dejar lo que viene por defecto



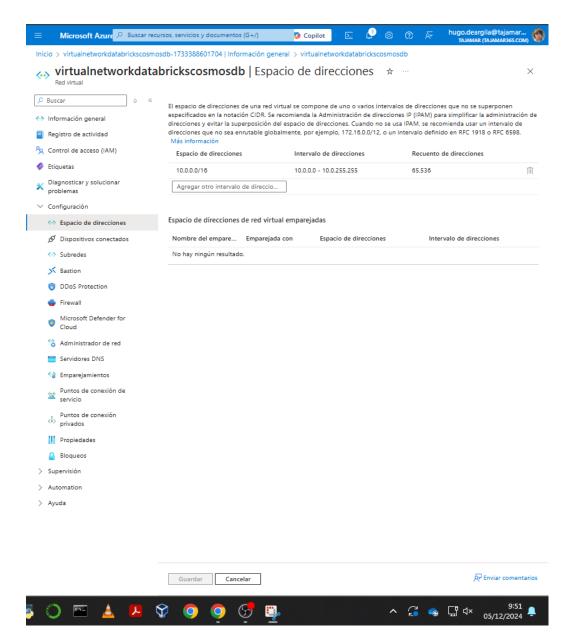


Una vez creado, le damos a go to resource:

Etiquetas



Vamos al Espacio de direcciones:



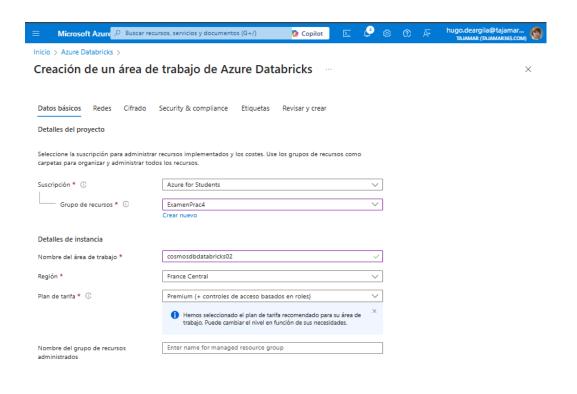
Añadimos una subred:

El espacio de direcciones de una red virtual se compone de uno o varios intervalos de direcciones que no se superponen especificados en la notación CIDR. Se recomienda la Administración de direcciones IP (IPAM) para simplificar la administración de direcciones y evitar la superposición del espacio de direcciones. Cuando no se usa IPAM, se recomienda usar un intervalo de direcciones que no sea enrutable globalmente, por ejemplo, 172.16.0.0/12, o un intervalo definido en RFC 1918 o RFC 6598.

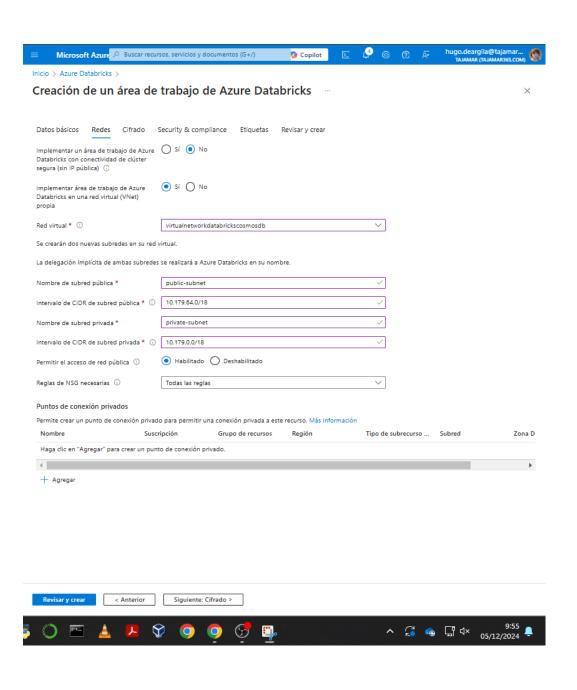
Más información

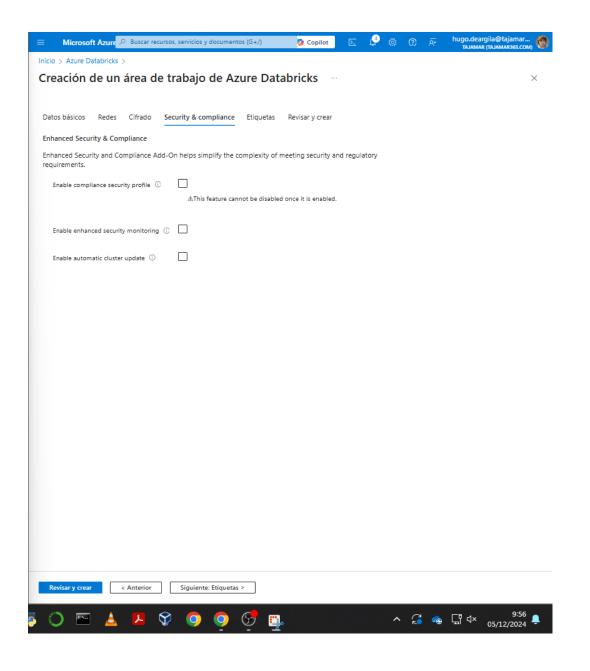


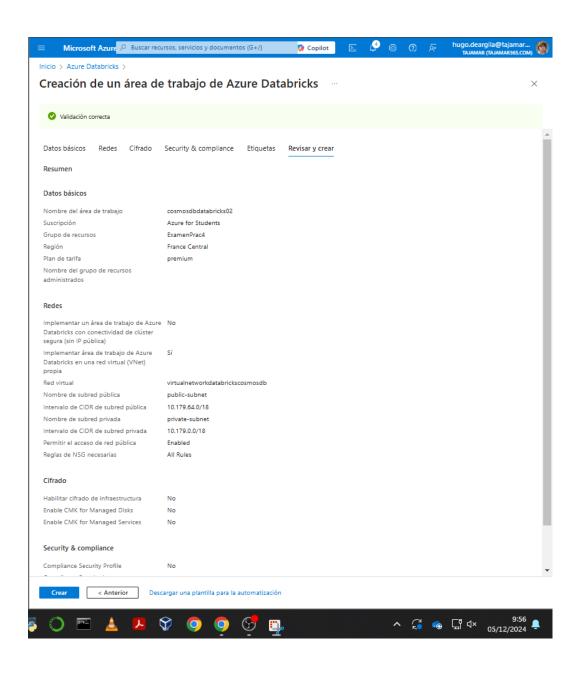
Paso 8: Crear un recurso de Azure Databricks:

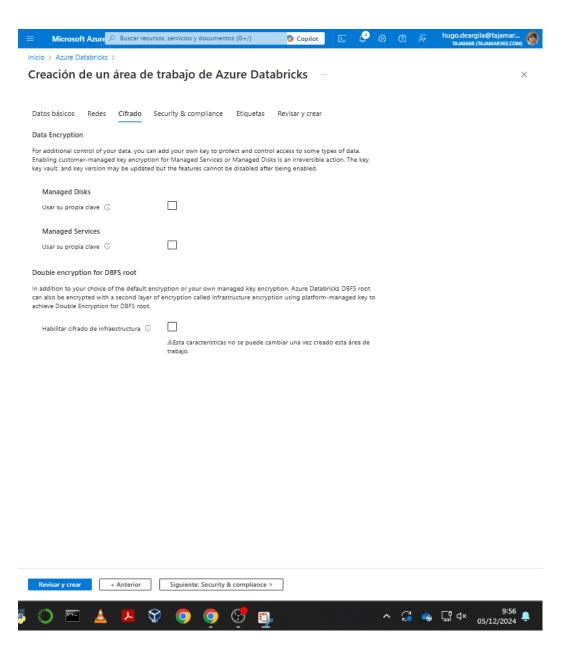




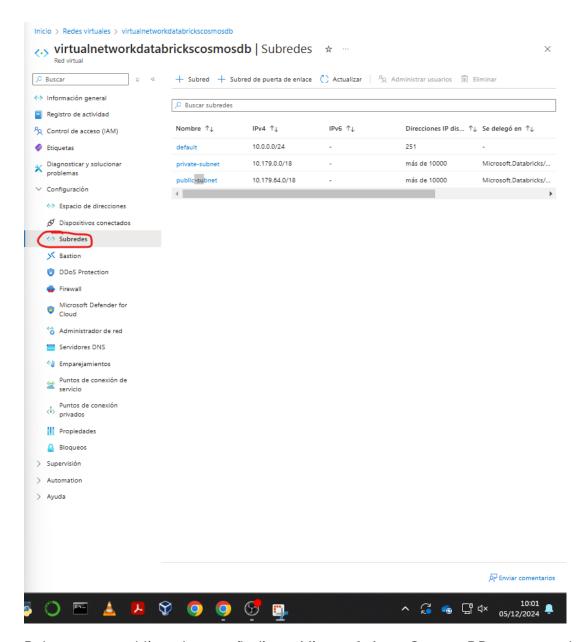








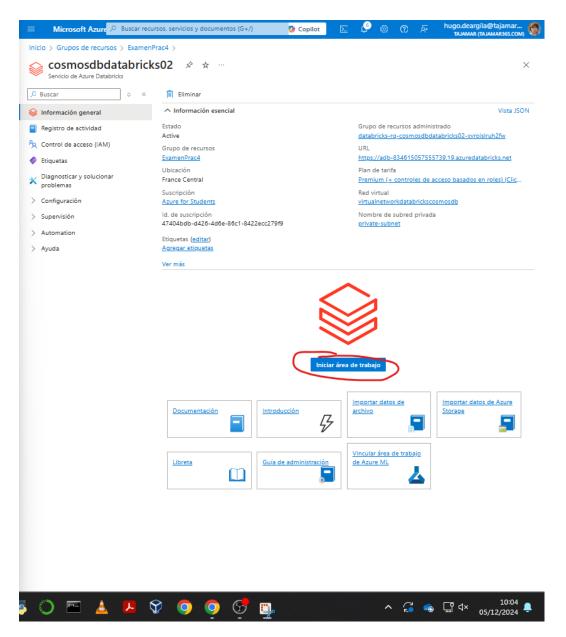
Paso 9: Crear un Service Endpoint para CosmosDB



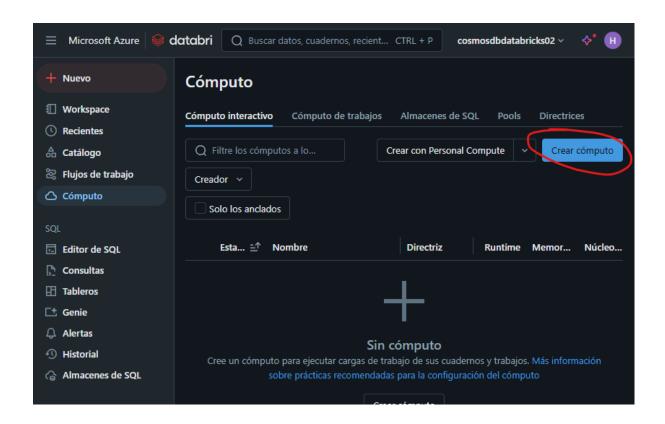
Pulsamos en public-subnet y añadimos Microsoft. Azure Cosmos DB en puntos de conexion del servicio:

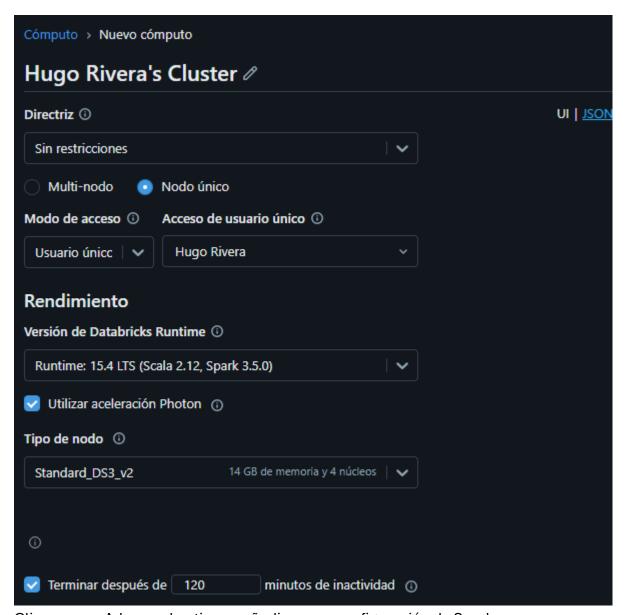


Paso 10: Iniciar Databricks



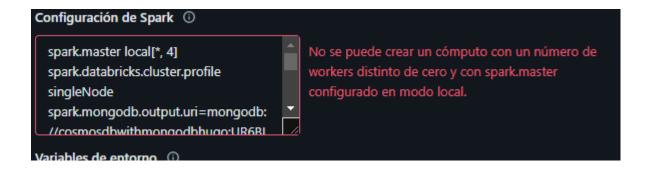
Vamos a Compute para crear un cluster:



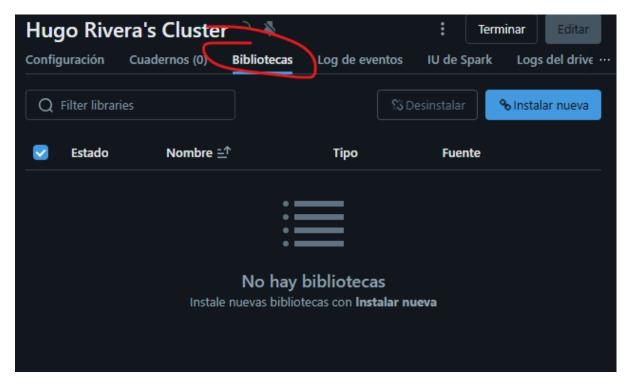


Clicamos en Advanced options y añadimos en configuración de Spark:

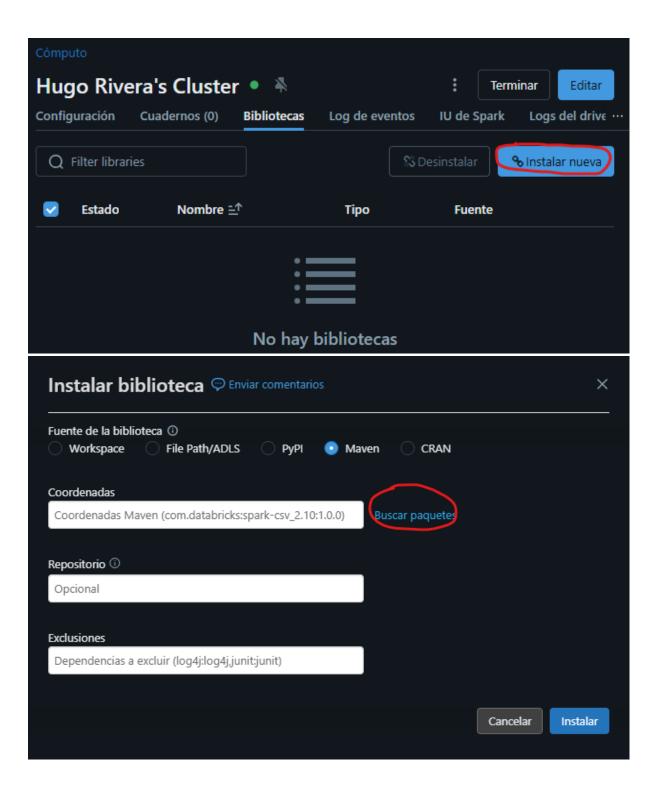
spark.master local[*, 4]
spark.databricks.cluster.profile singleNode
spark.mongodb.output.uri
mongodb://cosmosdbwithmongodbhugo:UR6BI4WFSmq4FsPiRQ60Seo0eVYWNLqA
9POkvT8f1wLGgQzvmotaaTjtwzC2l8kYEtl444mbmVC5ACDbRcBbHA==@cosmosdb
withmongodbhugo.mongo.cosmos.azure.com:10255/?ssl=true&replicaSet=globaldb
&retrywrites=false&maxldleTimeMS=120000&appName=@cosmosdbwithmongodbh
ugo@
spark.mongodb.input.uri
mongodb://cosmosdbwithmongodbhugo:UR6BI4WFSmq4FsPiRQ60Seo0eVYWNLqA
9POkvT8f1wLGgQzvmotaaTjtwzC2l8kYEtl444mbmVC5ACDbRcBbHA==@cosmosdb
withmongodbhugo.mongo.cosmos.azure.com:10255/?ssl=true&replicaSet=globaldb
&retrywrites=false&maxldleTimeMS=120000&appName=@cosmosdbwithmongodbh
ugo@



Una vez creado vamos a Libraries:

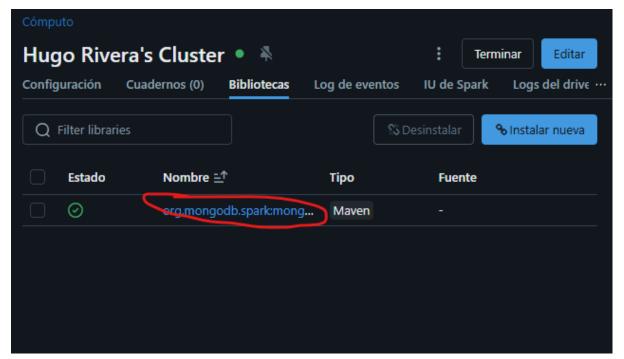


Le damos a instalar nueva:





Una vez este instalado:



Copiamos la coordenada:



Paso 12: Creamos un Notebook

Añadimos el siguiente cuadro de codigo:

```
from pyspark.sql import SparkSession

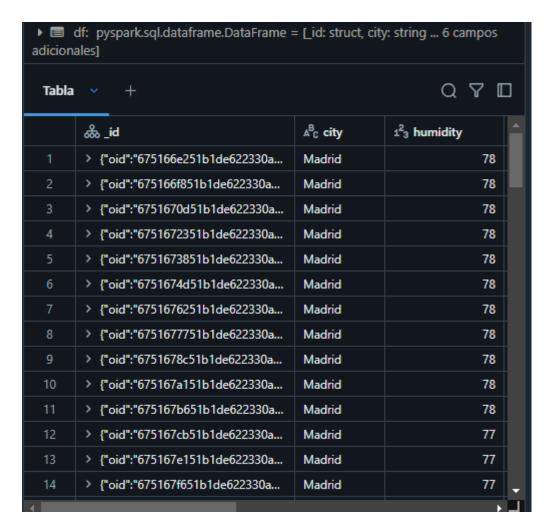
databasename = 'testingcosmosdb01'
collection = 'pruebadatabrickcosmosdb'
uri = "mongodb://<TU_CONEXION_A_COSMOS_DB>"

spark = SparkSession.builder.appName("TestingApp").config('spark.jars.packages', 'org.mongodb.spark:mongo-spark-connector_2.12:3.0.1').getOrCreate()

df = spark.read.format("com.mongodb.spark.sql.DefaultSource") \
.option("uri", uri) \
.option("database", databasename) \
.option("collection", collection) \
.load()

df.show()
```

Con esto obtendremos los datos de nuestra database:



Despues guardamos los datos en una tabla de PowerBI:

df.write.mode("overwrite").saveAsTable("PowerBITable")

Paso 13: Conectar a PowerBI

1. Configurar la conexión Abre Power BI Desktop.

Selecciona Obtener Datos > Azure > Azure Databricks.

Introduce la URL del clúster y las credenciales de acceso.

2. Cargar la tabla Navega hasta la tabla PowerBITable que creaste en Databricks. Carga los datos en Power BI para comenzar a crear las visualizaciones.