Contents

[Infraestructura Cloud 2](#_Toc22166473)

[Bucket en Google Cloud (Cloud Storage) 2](#_Toc22166474)

[Máquina virtual en Google Cloud (Compute engine) 2](#_Toc22166475)

[Base de datos MySQL en Google Cloud (Cloud SQL) 4](#_Toc22166476)

[Configurar redes en Google Cloud 8](#_Toc22166477)

[Instalar Software en nuestra VM 11](#_Toc22166478)

[Herramienta de visualización – Qlik Sense 13](#_Toc22166479)

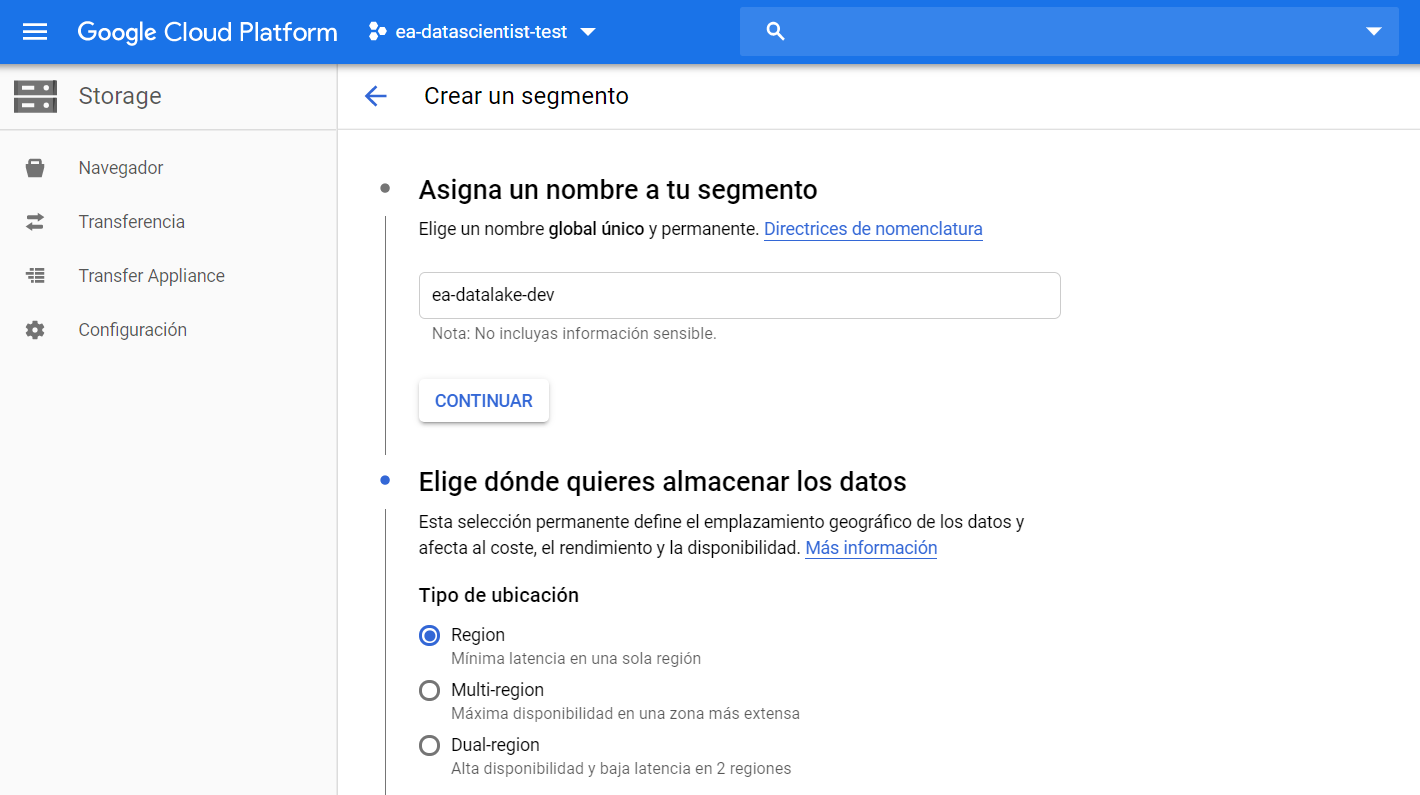
[Conectar Qlik Sense con nuestra base de datos 14](#_Toc22166480)

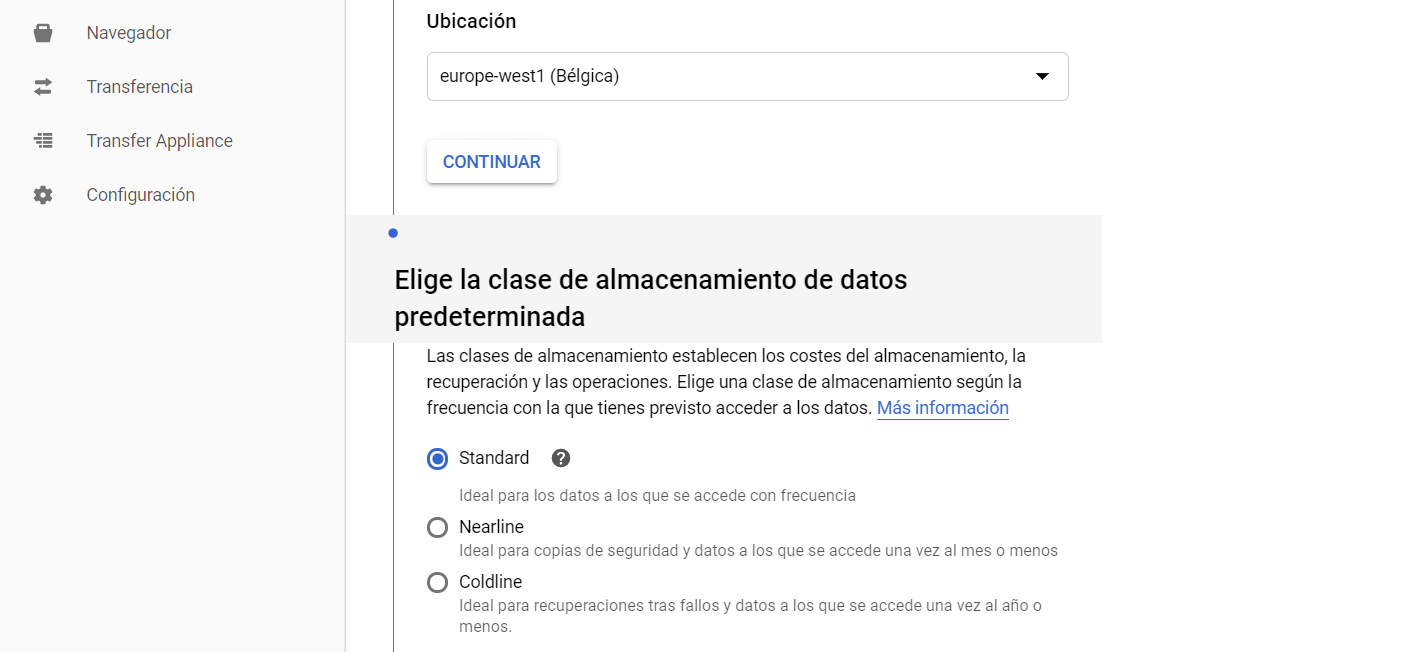
# Infraestructura Cloud

La infraestructura necesaria para resolver el problema planteado se ha levantado sobre la plataforma de Google Cloud. Esta plataforma ofrece diversos componentes que permiten acelerar la creación de recursos y el desarrollo de aplicaciones de ML.

## Bucket en Google Cloud (Cloud Storage)

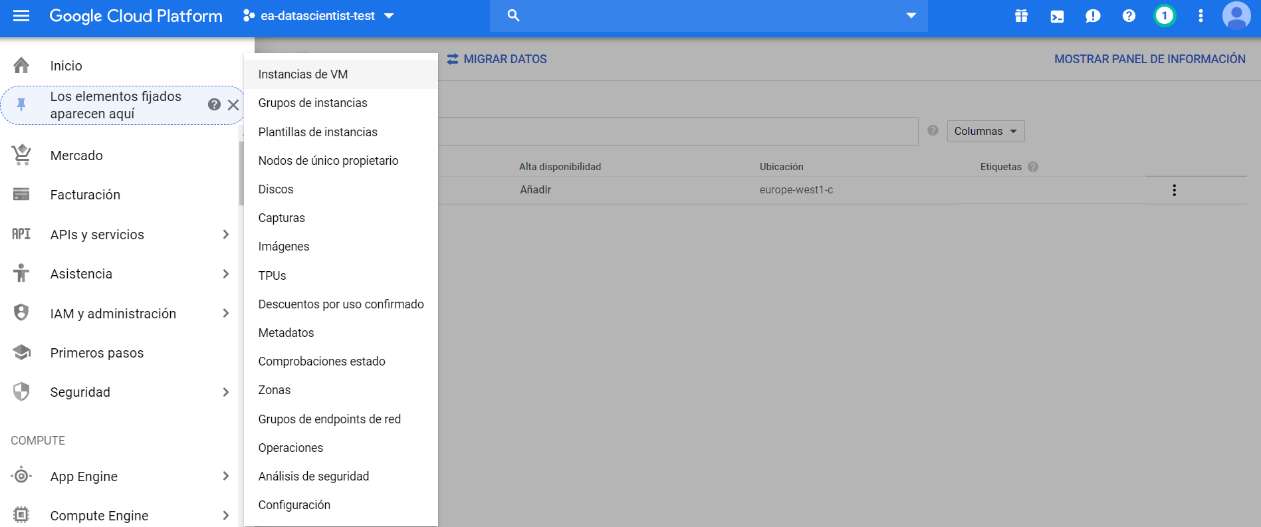
En primer lugar, vamos a crear un bucket en Cloud Storage donde almacenar los datos. Para ello, accedemos a través del menú de navegación a “**Storage**” y configuramos las características del bucket, incluyendo: nombre, tipo de almacenamiento, región,



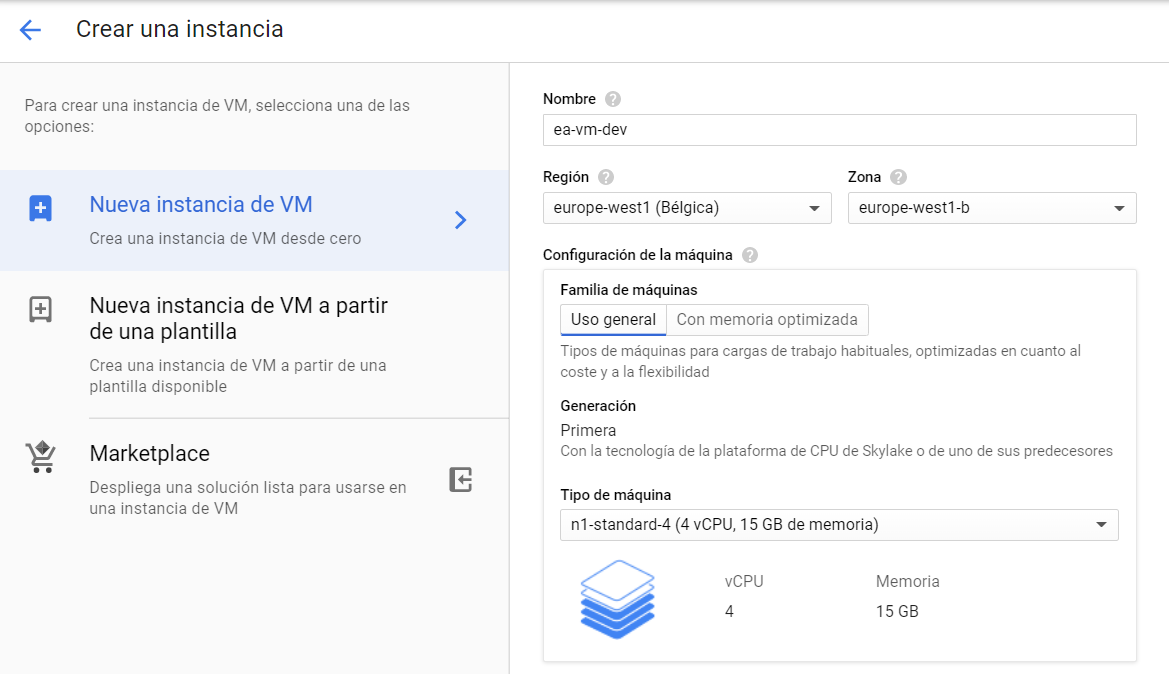


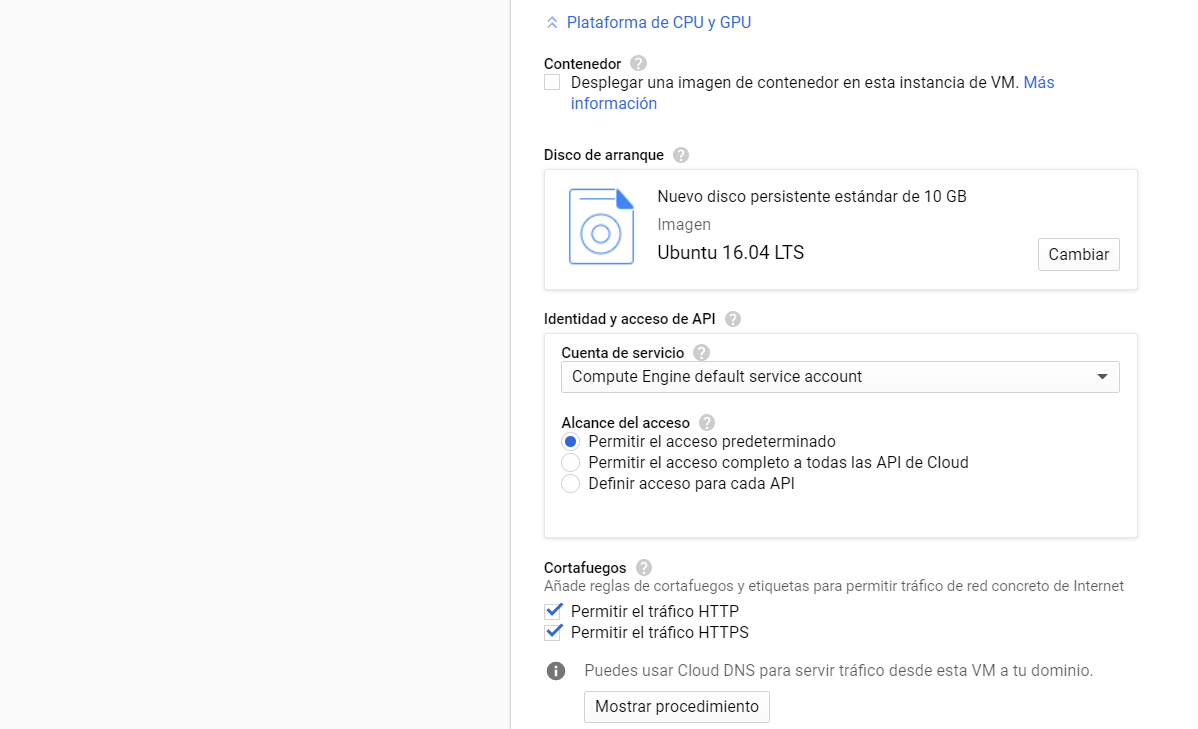
## Máquina virtual en Google Cloud (Compute engine)

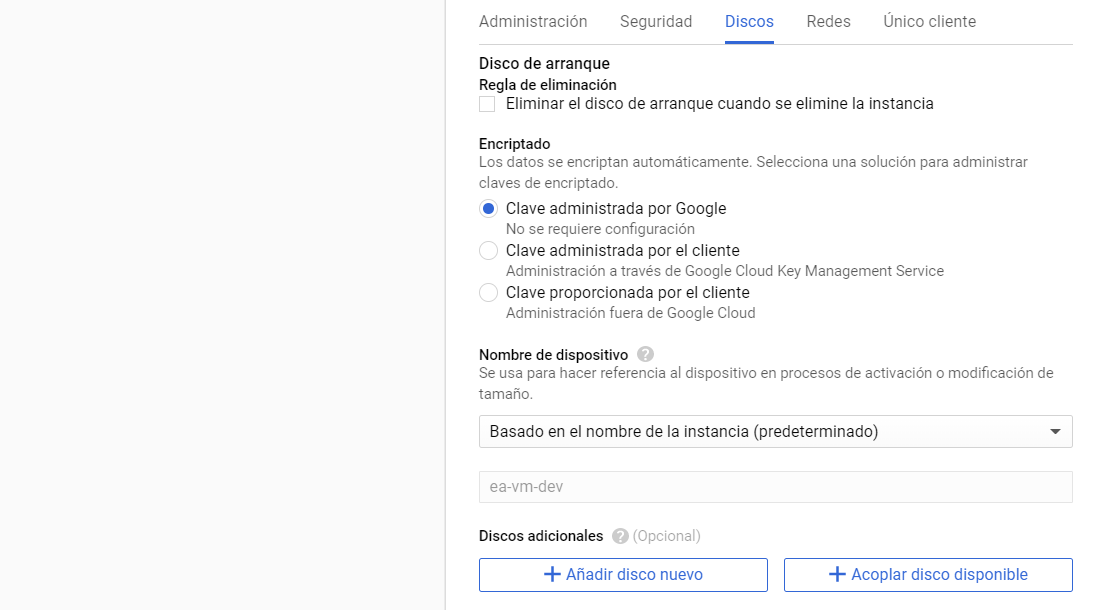
A continuación, vamos a desplegar una máquina virtual en el entorno cloud de Google. Sobre esta máquina virtual ejecutaremos los códigos Python destinados a la carga, tratamiento y análisis de datos. Para ello seleccionamos “**Compute Engine -> Instancias de VM**” en el menú de navegación.



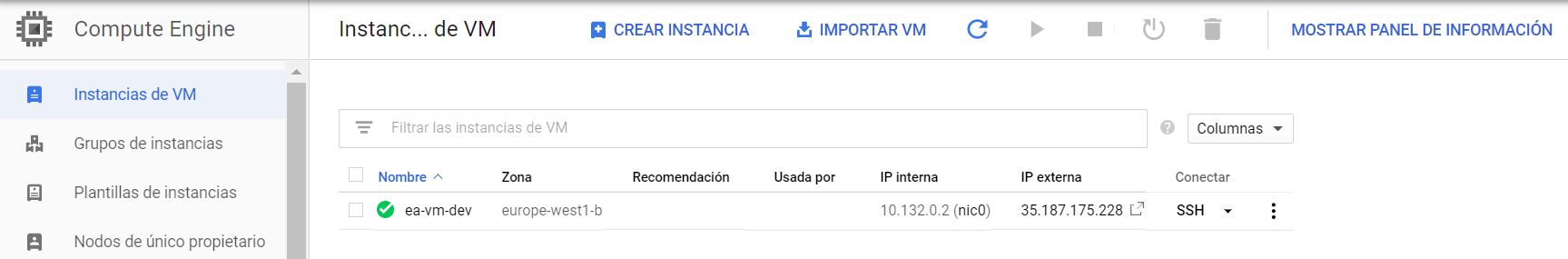
Y configuramos los detalles de la instancia, incluyendo: nombre, región y zona, tipo de máquina, distribución del SO, permisos, reglas del cortafuegos, redes, discos, etc.





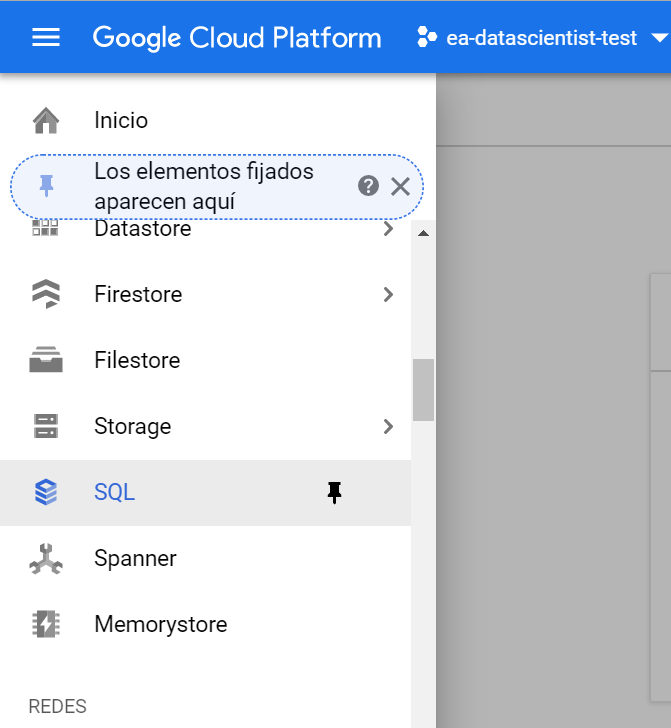


Finalmente, tras unos minutos, tenemos nuestra instancia de VM en funcionamiento.

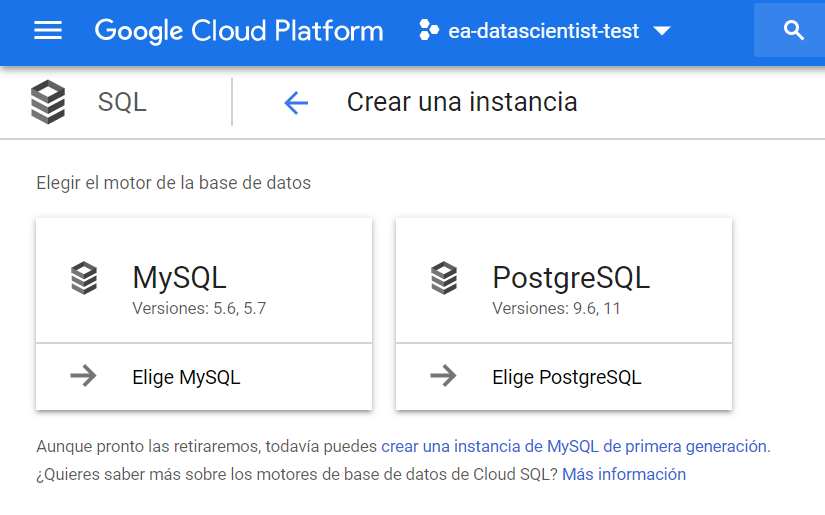


## Base de datos MySQL en Google Cloud (Cloud SQL)

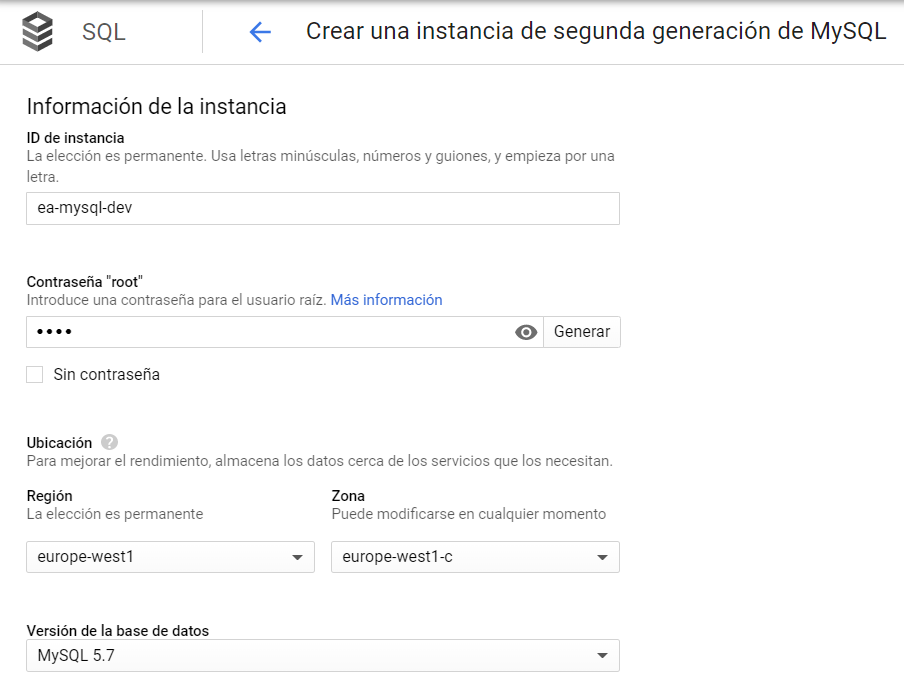
A continuación, desplegaremos una base de datos MySQL en la que se cargarán los datos proporcionados. Para ello es necesario seleccionar Cloud “**SQL**” el menú de navegación.



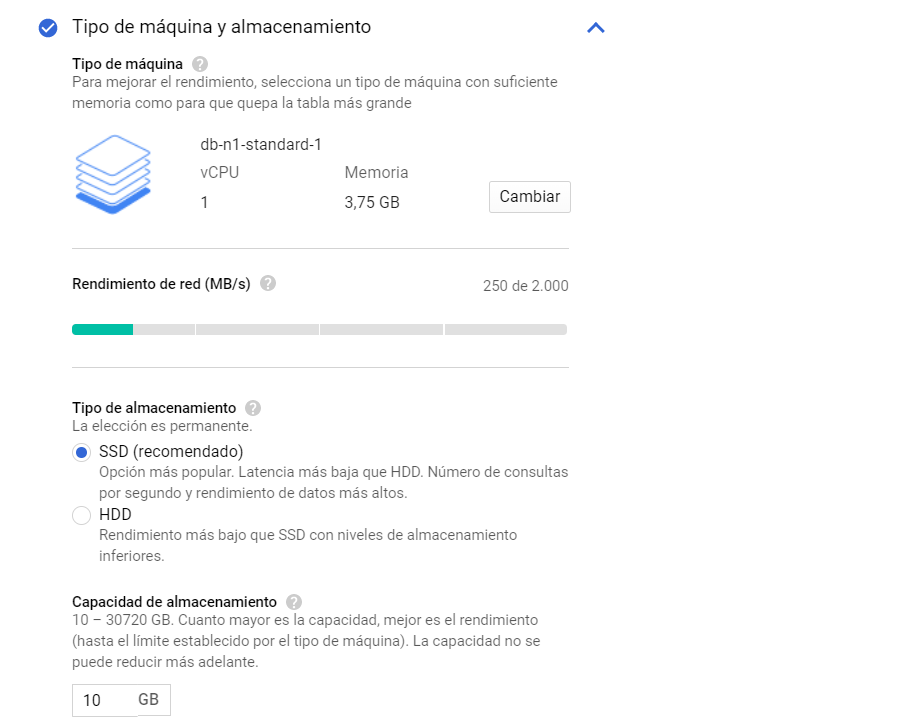
Elegir una instancia MySQL.



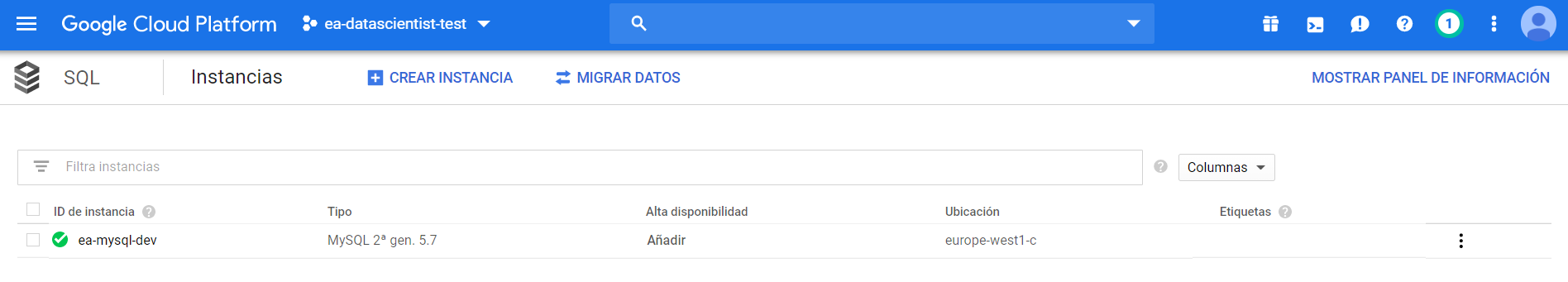
Y configurar los detalles de la instancia, incluyendo: id, password del usuario root, región y zona, versión MySQL, conectividad, tipo de máquina y almacenamiento.



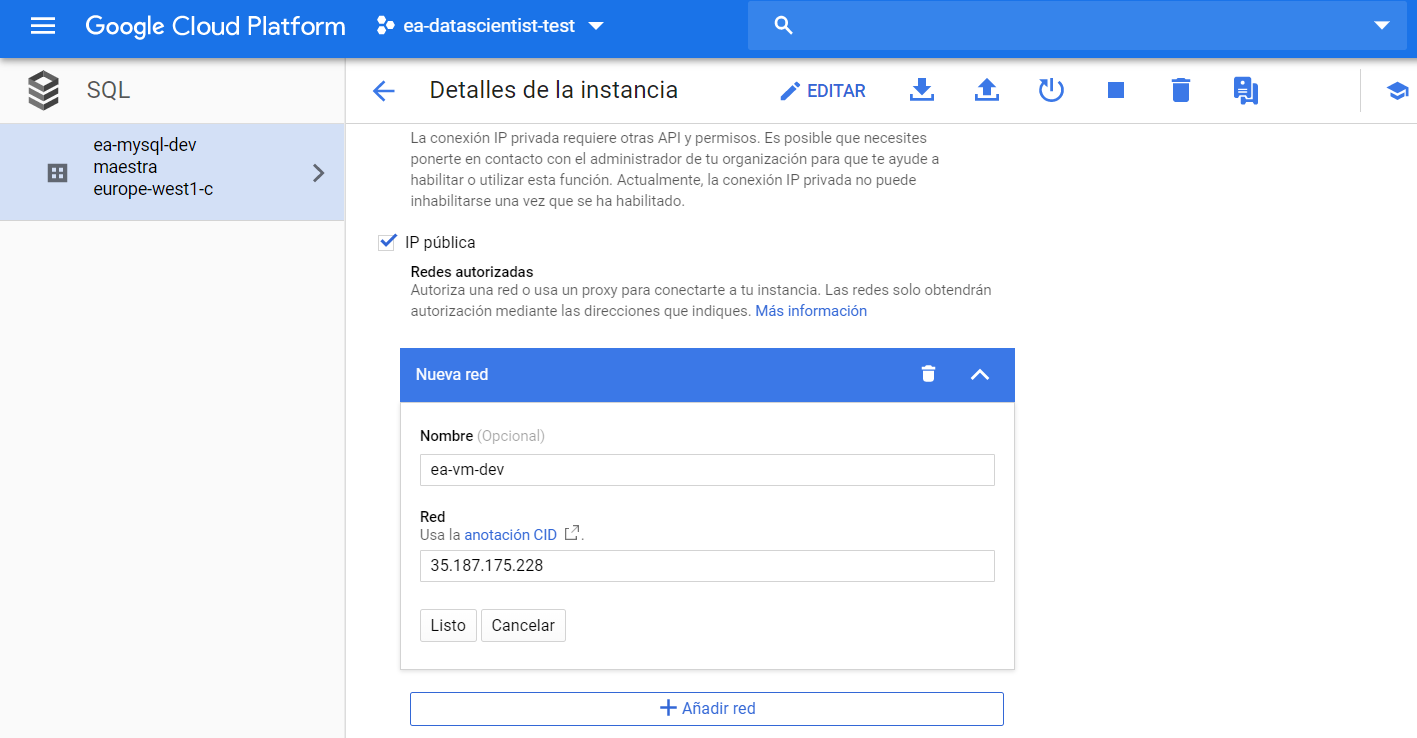




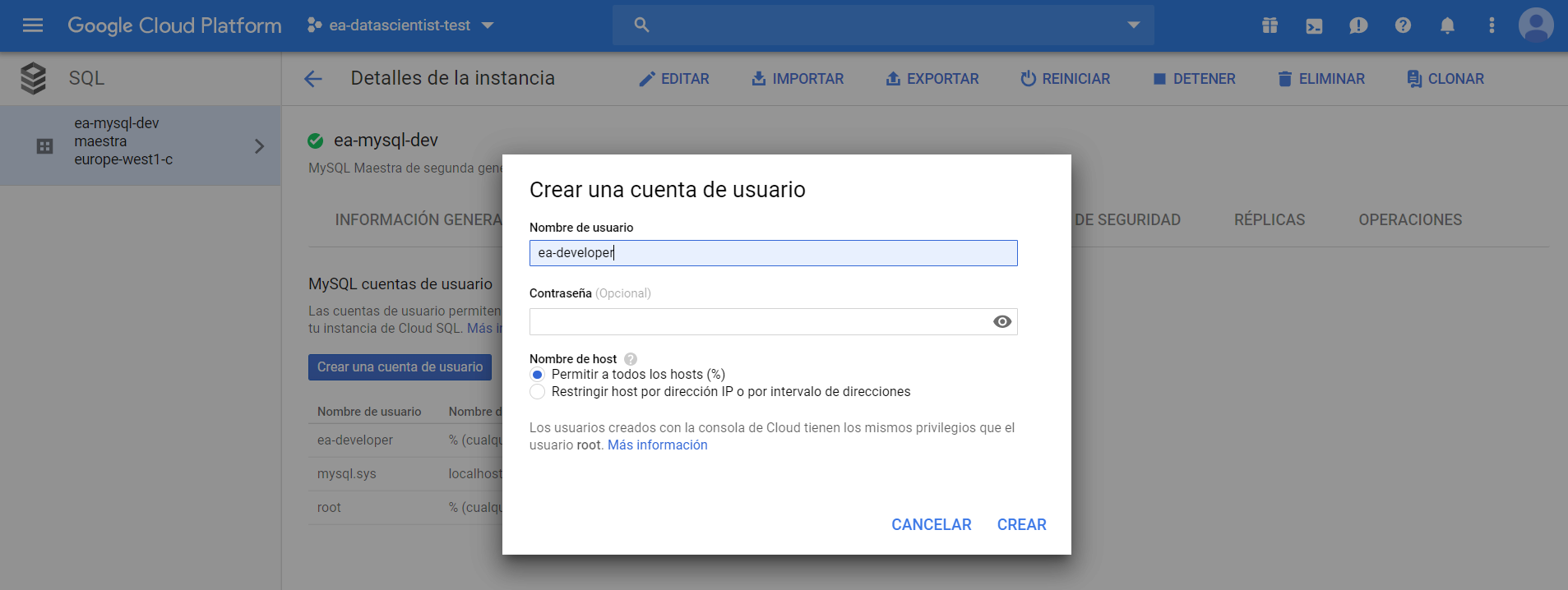
Finalmente, en unos minutos tenemos nuestra base de datos MySQL en funcionamiento:

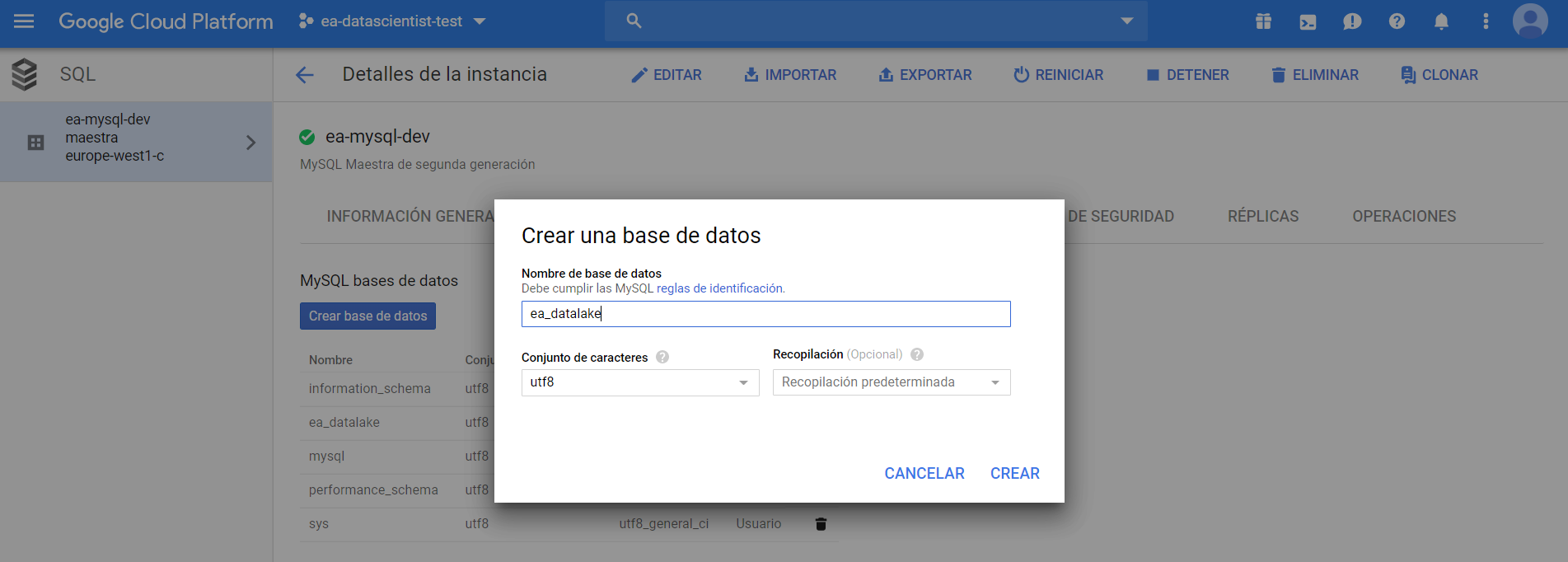


A continuación, vamos a autorizar las conexiones desde la instancia VM de Google Compute Engine que hemos creado.



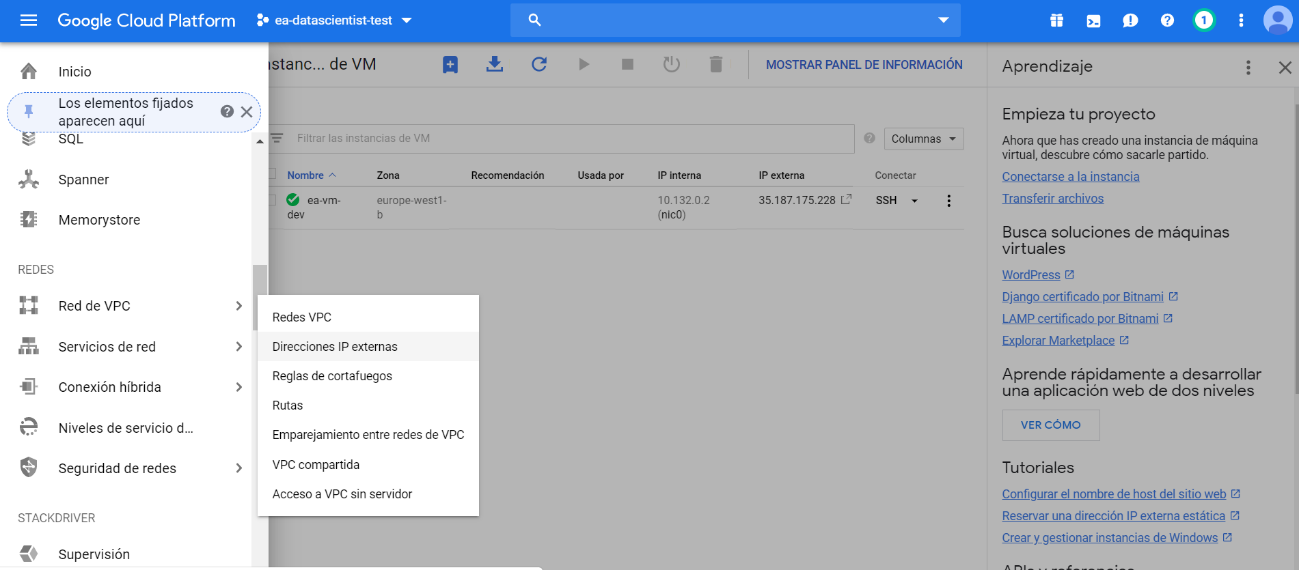
Por último, vamos a crear un usuario y una base de datos dentro de nuestra instancia Cloud SQL.

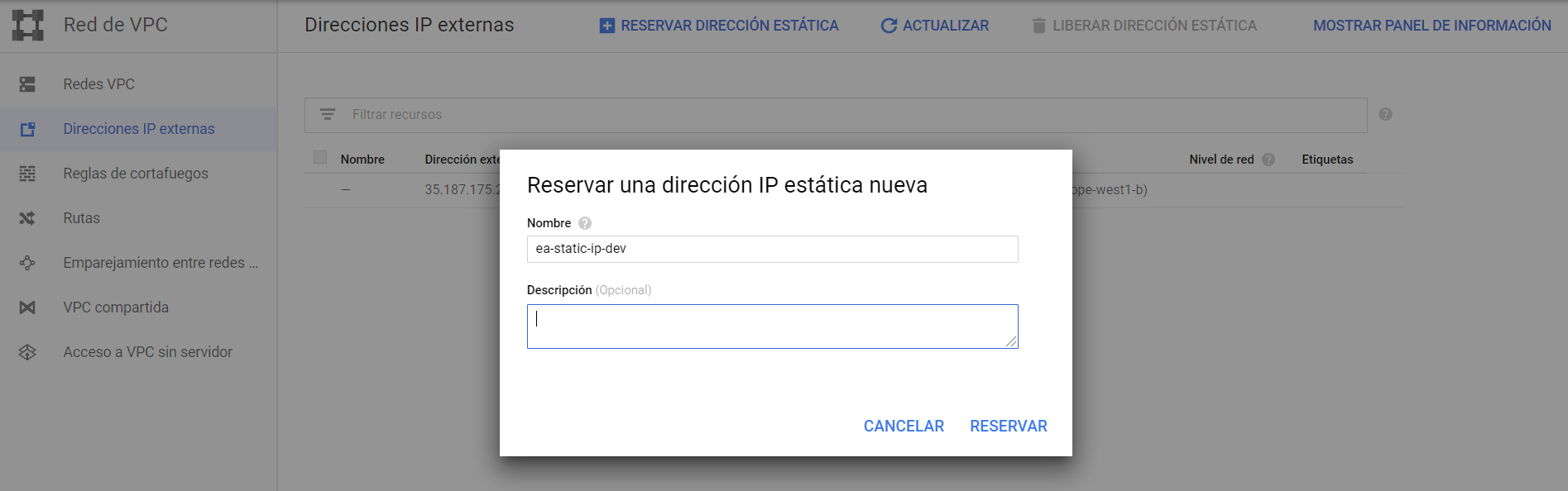




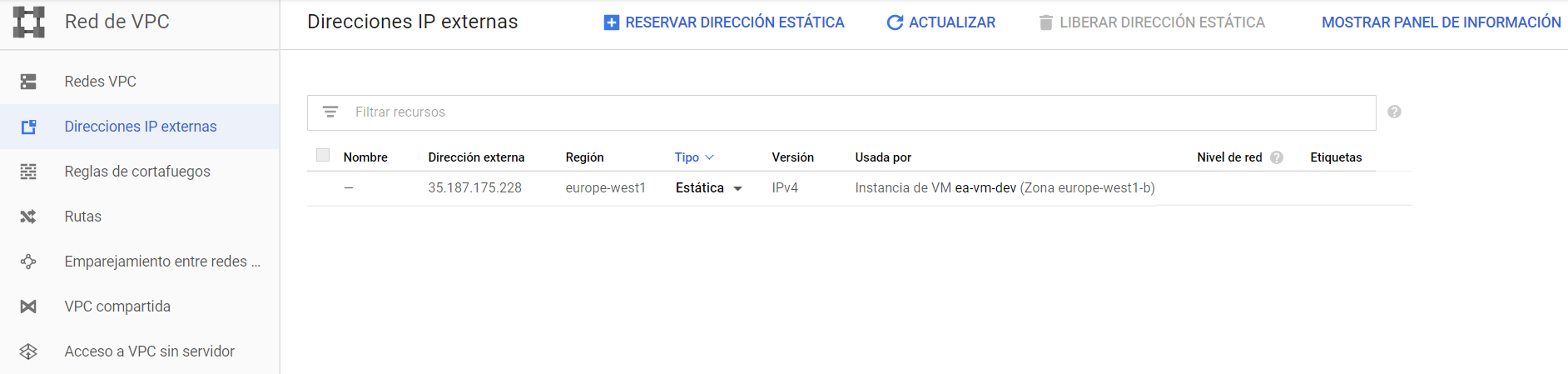
## Configurar redes en Google Cloud

A continuación, vamos a hacer que la ip externa de la vm sea estática de cara a acceder posteriormente a nuestro Jupyter Notebook. Para ello, debemos acceder a “**Red de VPC -> Direcciones IP externas**”.

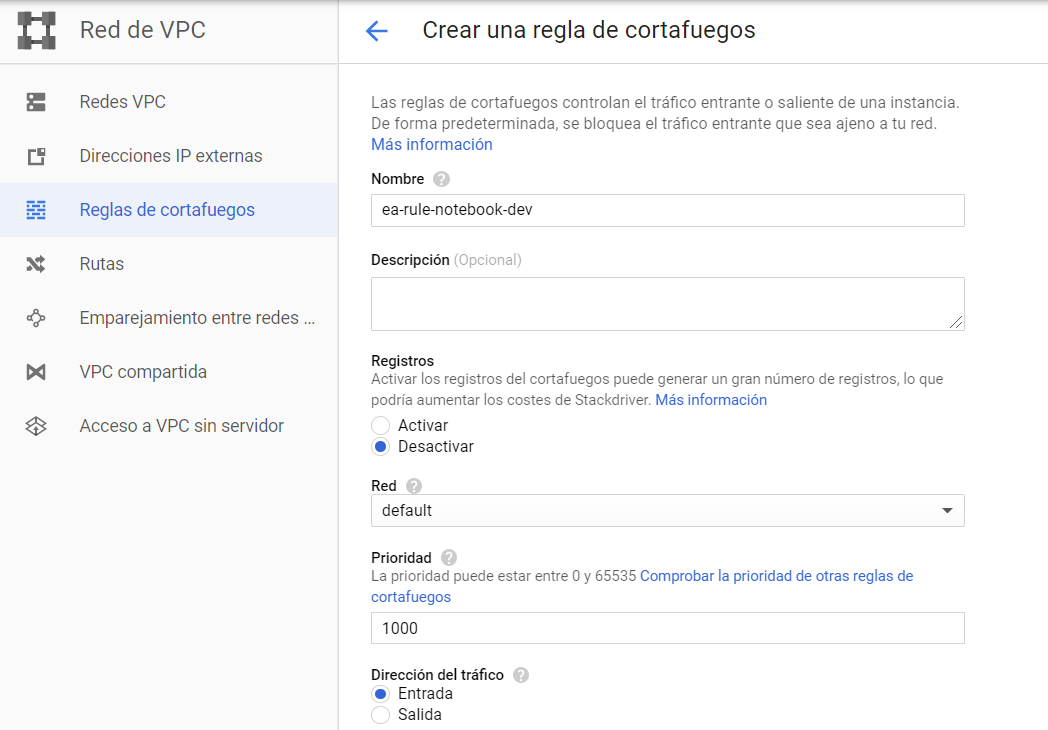


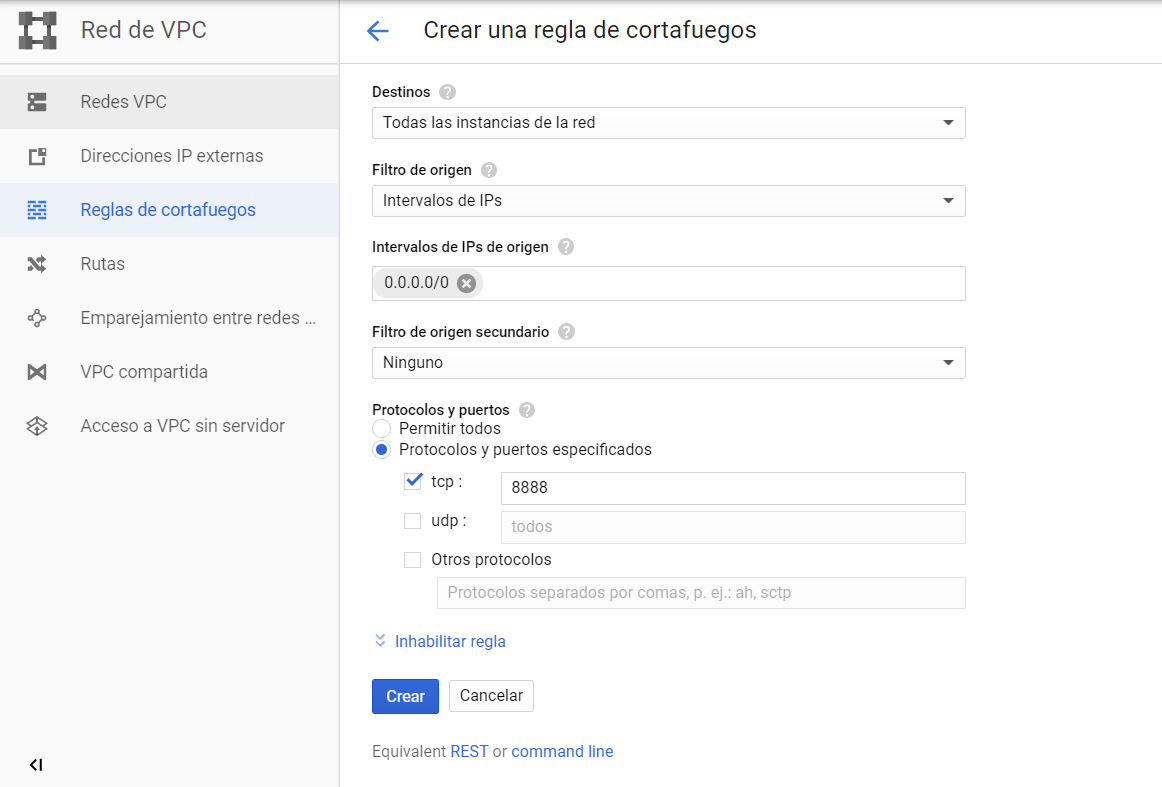


Finalmente, tenemos nuestra dirección IP externa estática: **35.187.175.228**



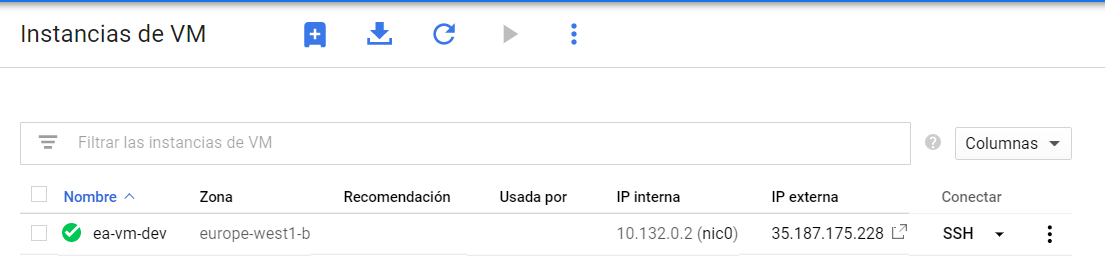
Por último, vamos a crear una regla que permita el acceso a través del puerto 8888 (para poder acceder posteriormente a nuestro Jupyter notebook). Por simplicidad, aunque no es lo recomendable, vamos a abrirlo a todos los rangos IP (0.0.0.0/0).





## Instalar Software en nuestra VM

Nos conectamos a nuestra VM en Goocle Cloud mediante SSH.

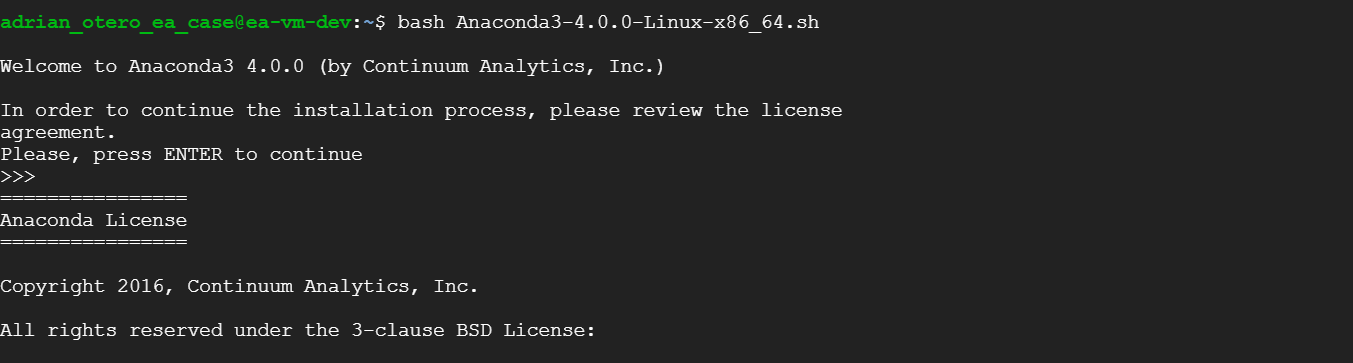


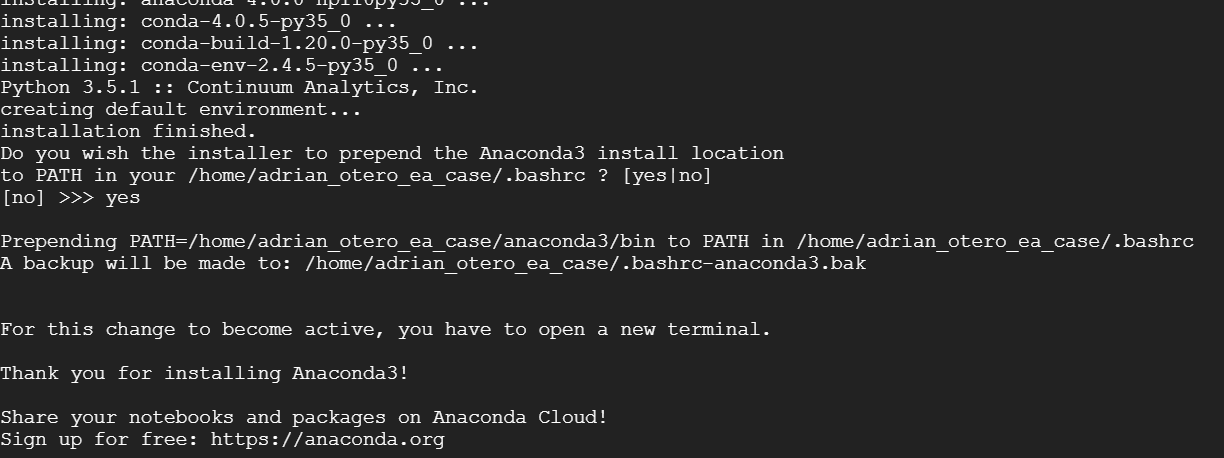
Ejecutamos los siguientes comandos para instalar y configurar, en primer lugar, Anaconda:

**wget** [**http://repo.continuum.io/archive/Anaconda3-4.0.0-Linux-x86\_64.sh**](http://repo.continuum.io/archive/Anaconda3-4.0.0-Linux-x86_64.sh)



**bash Anaconda3-4.0.0-Linux-x86\_64.sh**

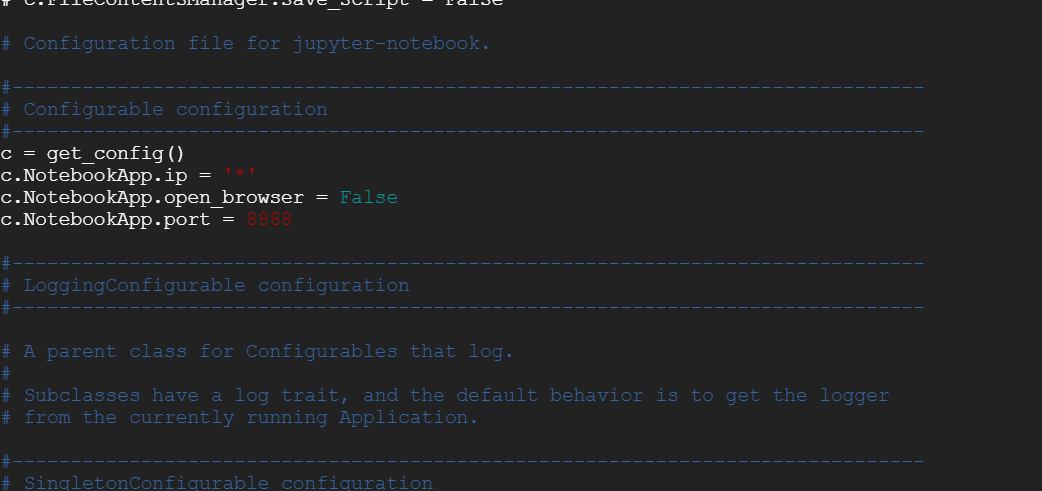




**source ~/.bashrc**

**jupyter notebook --generate-config**

**vi ~/.jupyter/jupyter\_notebook\_config.py**



Añadir al fichero de configuración las 4 líneas de la imagen anterior, incluyendo el puerto configurado previamente en la regla del firewall (8888).

A continuación, vamos a instalar algunos paquetes y librerías de Python usando los siguientes comandos:

**sudo apt update**

**sudo apt upgrade**

**sudo apt install zip**

**sudo apt-get install mysql-client**

**sudo apt-get install libsm6 libxext6 libxrender-dev**

**sudo apt-get install texlive texlive-latex-extra pandoc**

**conda update pandas**

**pip install --upgrade pip**

**pip install mysql-connector**

**pip install mysql-connector-python-rf**

**pip install pandas\_profiling**

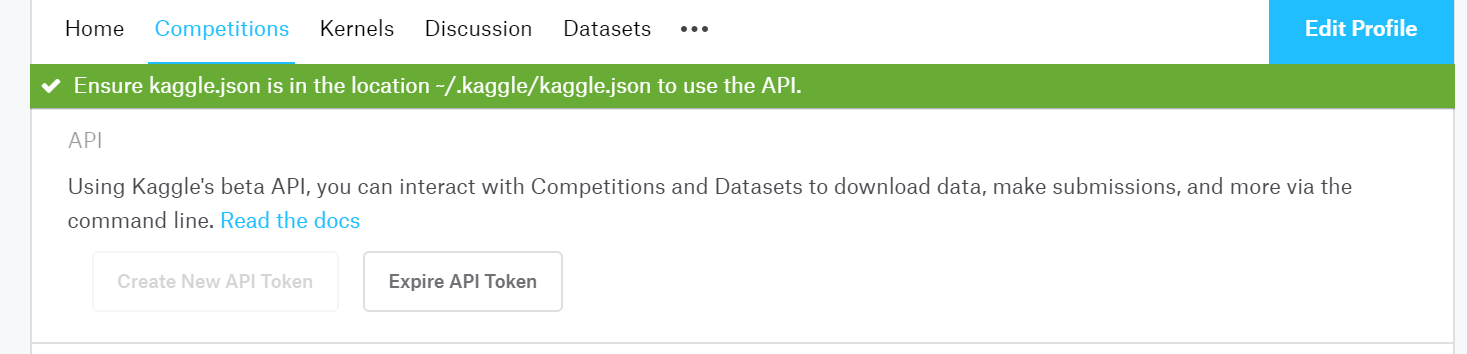
**pip install kaggle**

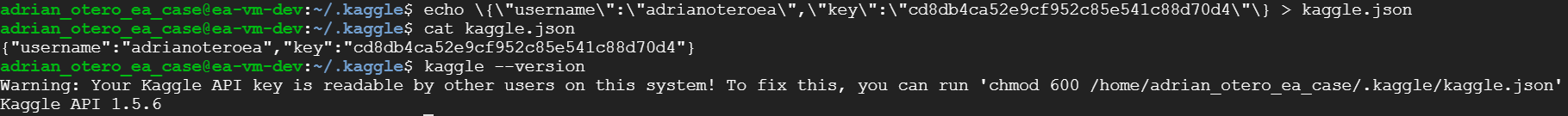
**pip install google-cloud-storage**

**pip install google-cloud-bigquery**

**pip install google-cloud-language**

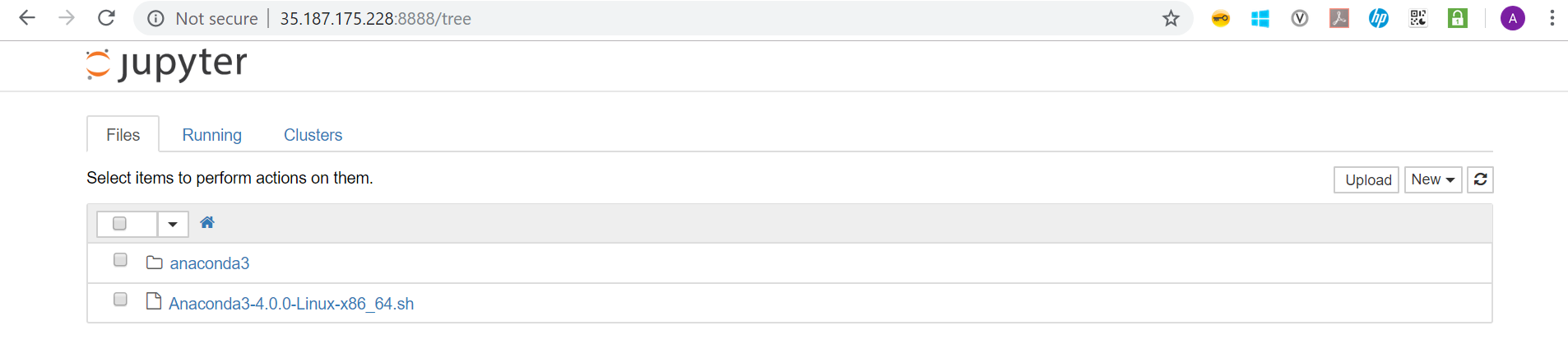
**pip install pyarrow**





Por último, ejecutar el notebook de Python: **jupyter notebook**

Jupyter Notebook será accesible a través de la ip estática y puerto configurados en los pasos previos, en este caso: [http://35.187.175.228:8888](http://35.187.175.228:8888/)

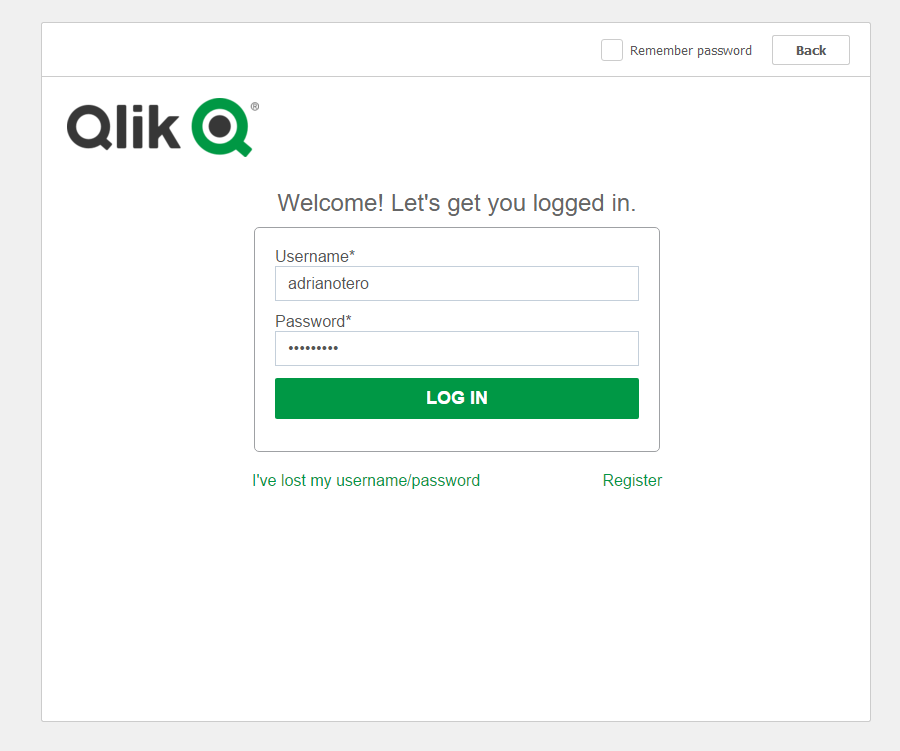


# Herramienta de visualización – Qlik Sense

Como herramienta de visualización se utilizará Qlik Sense. En primer lugar, es necesario descargar e instalar la aplicación. Es posible descargarla de forma gratuita a través de su web oficial:

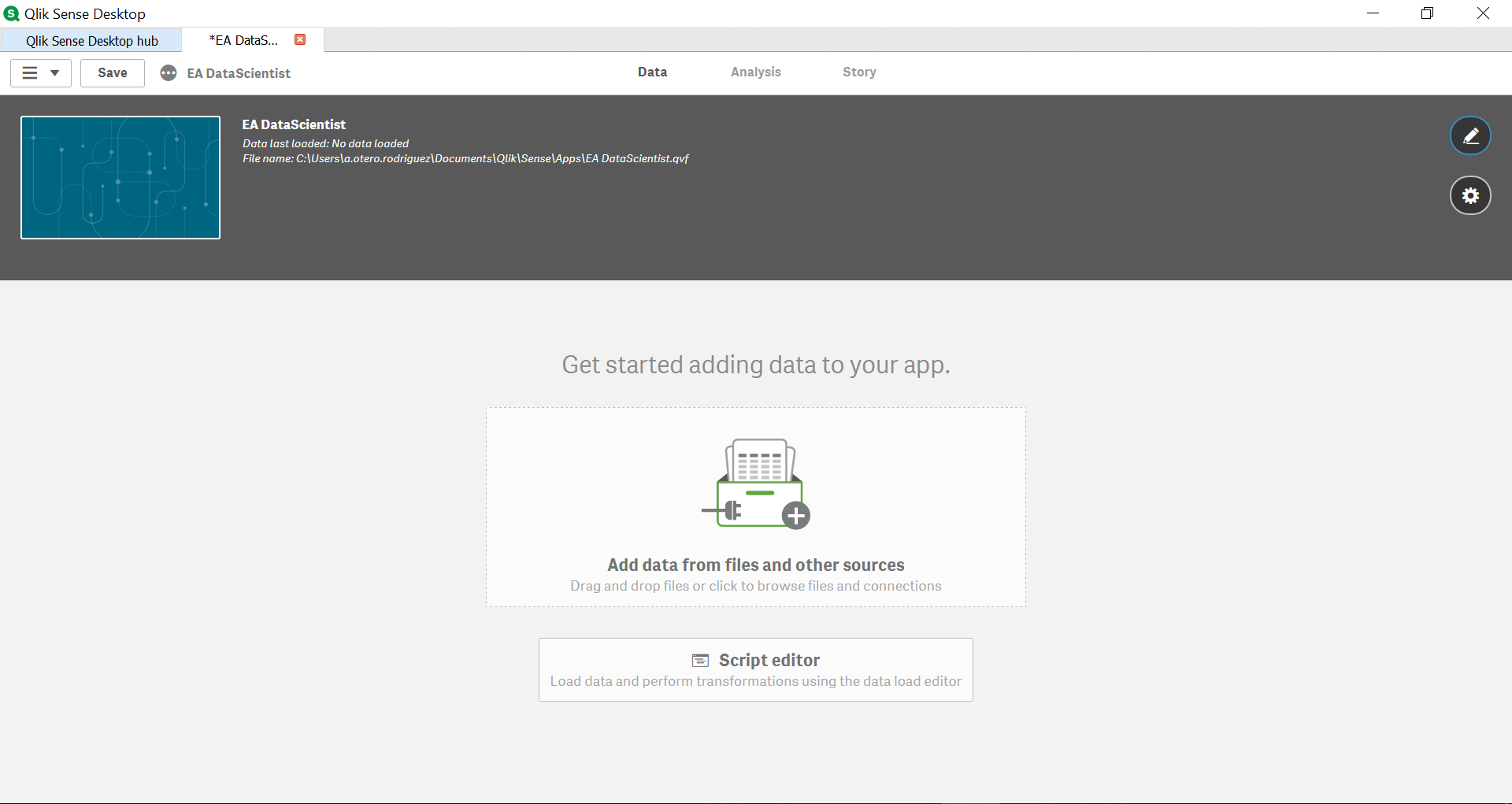
<https://www.qlik.com/es-es/try-or-buy/download-qlik-sense>

Para poder obtener una copia, es necesario registrarse en su web y aceptar el correo de activación. Por otro lado, para instalar Qlik se aceptarán todas las opciones configuradas por defecto.

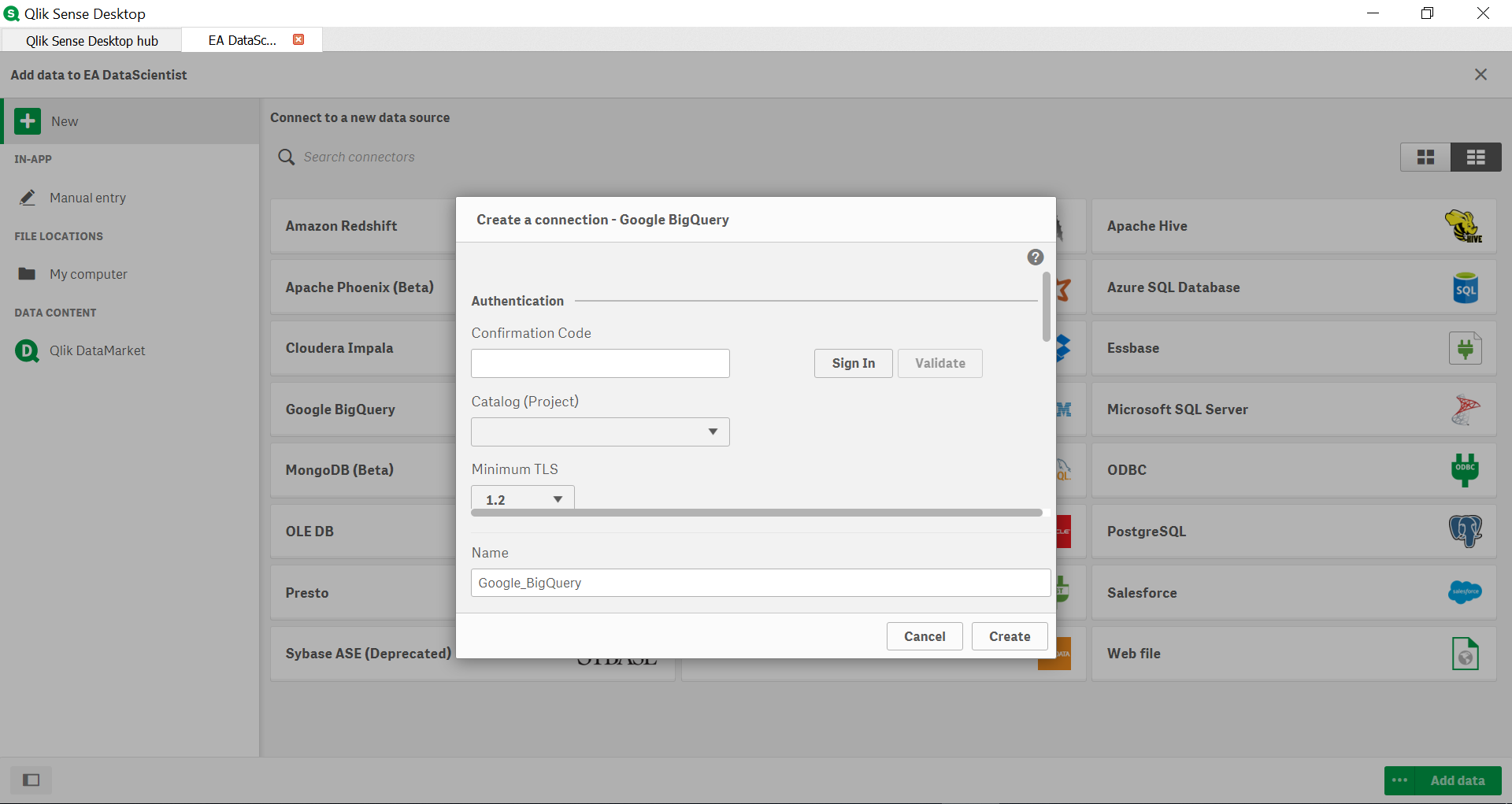


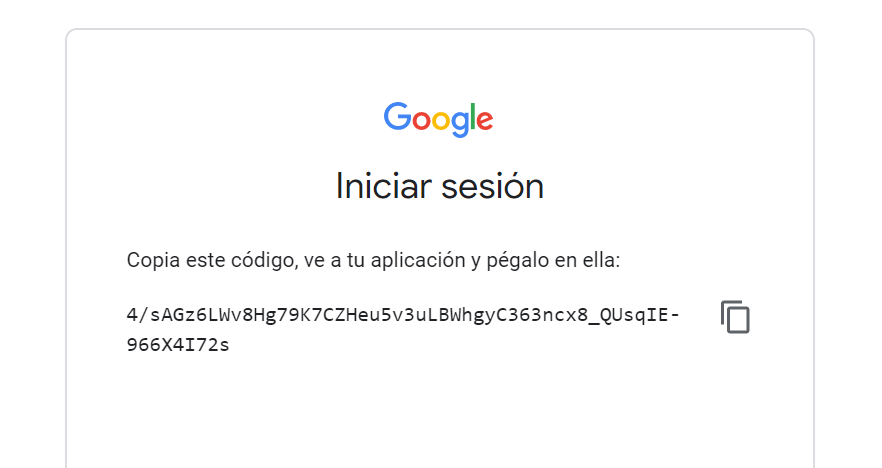
## Conectar Qlik Sense con nuestra base de datos

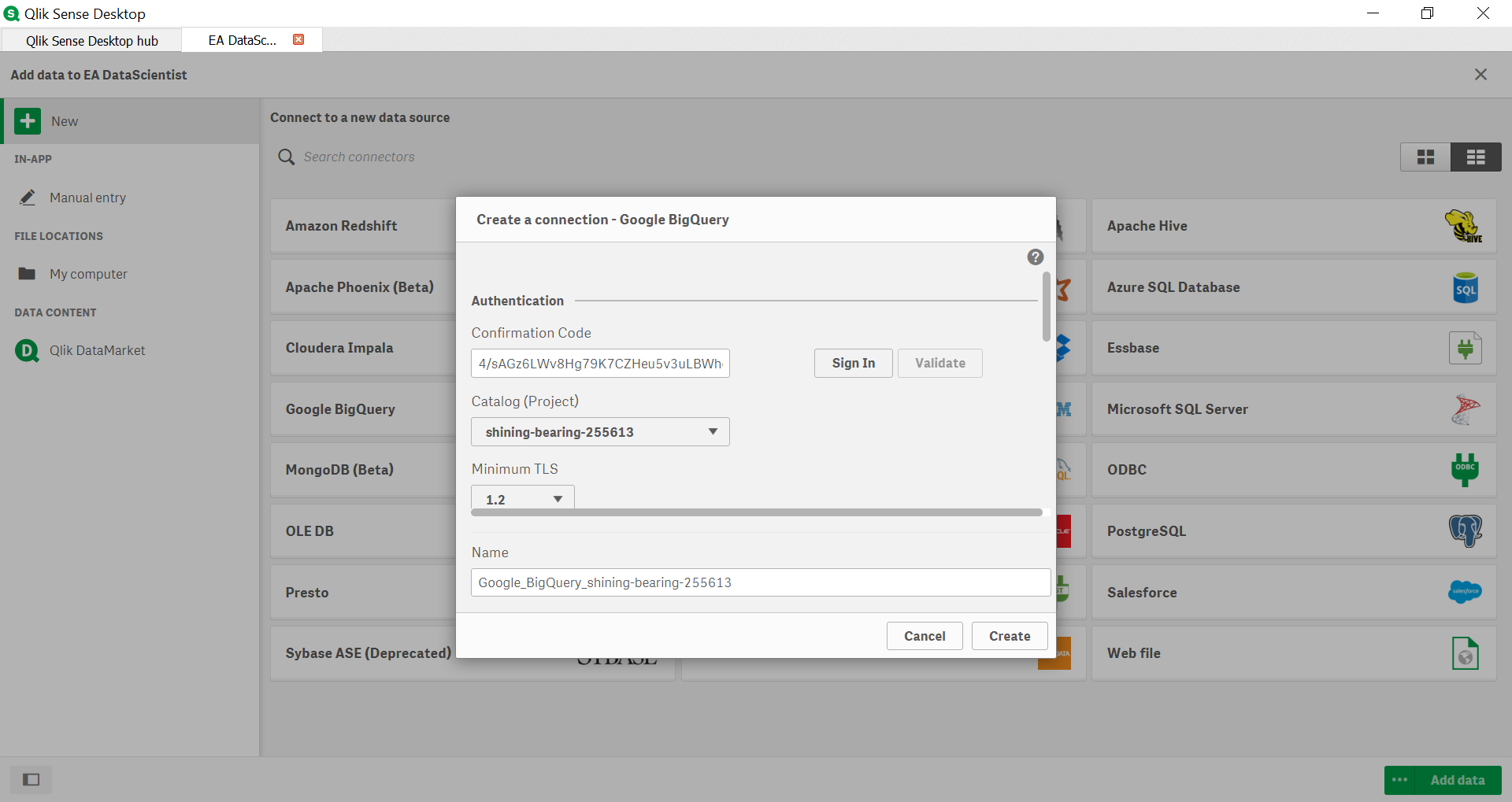
En primer lugar, debemos crear una nueva APP y seleccionar la opción de añadir una nueva fuente:

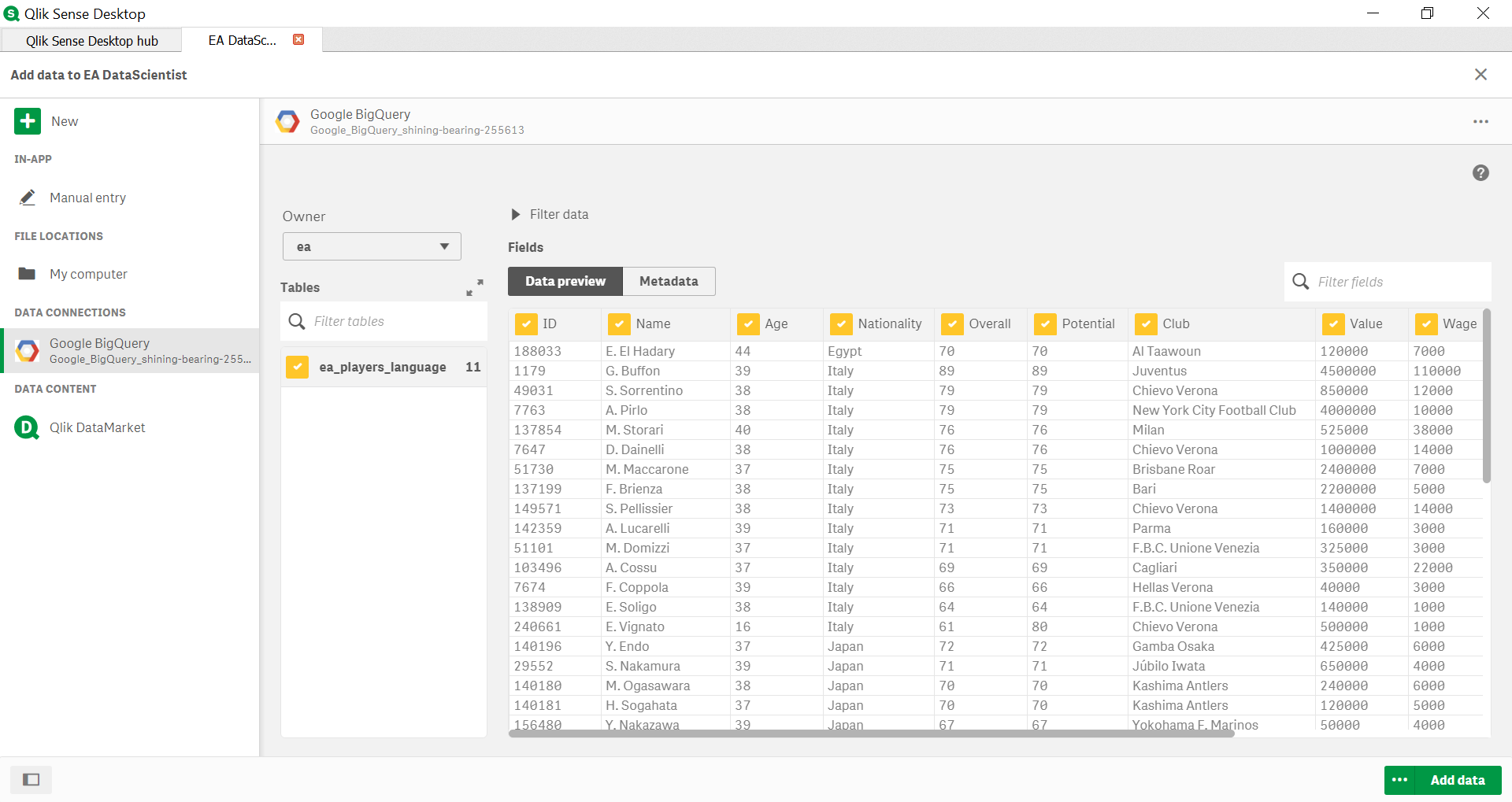


Qlik cuenta con conectores con Google Cloud que permiten acceder a la información almacenada en nuestra base de datos Cloud SQL de forma sencilla.



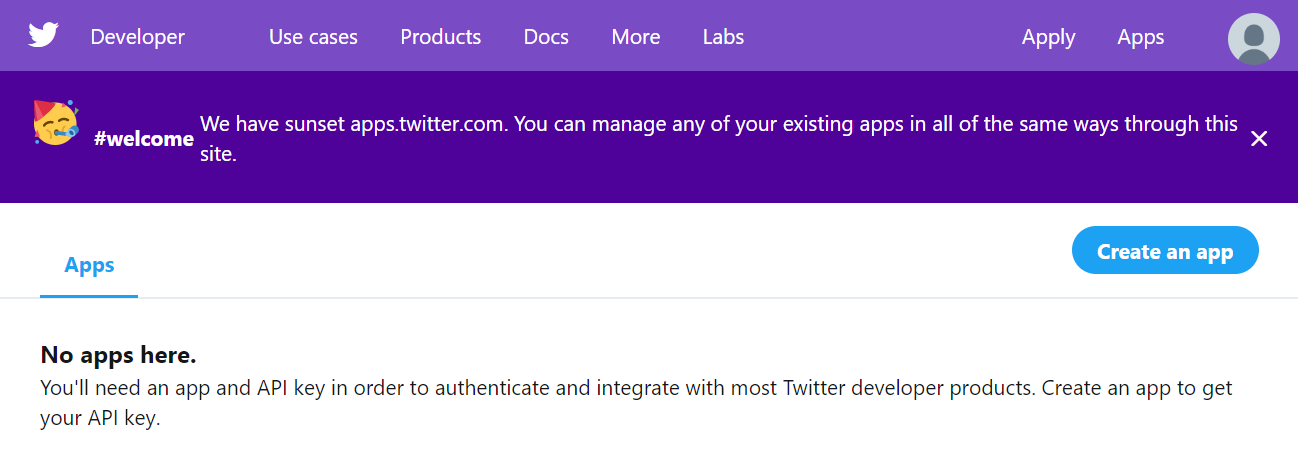




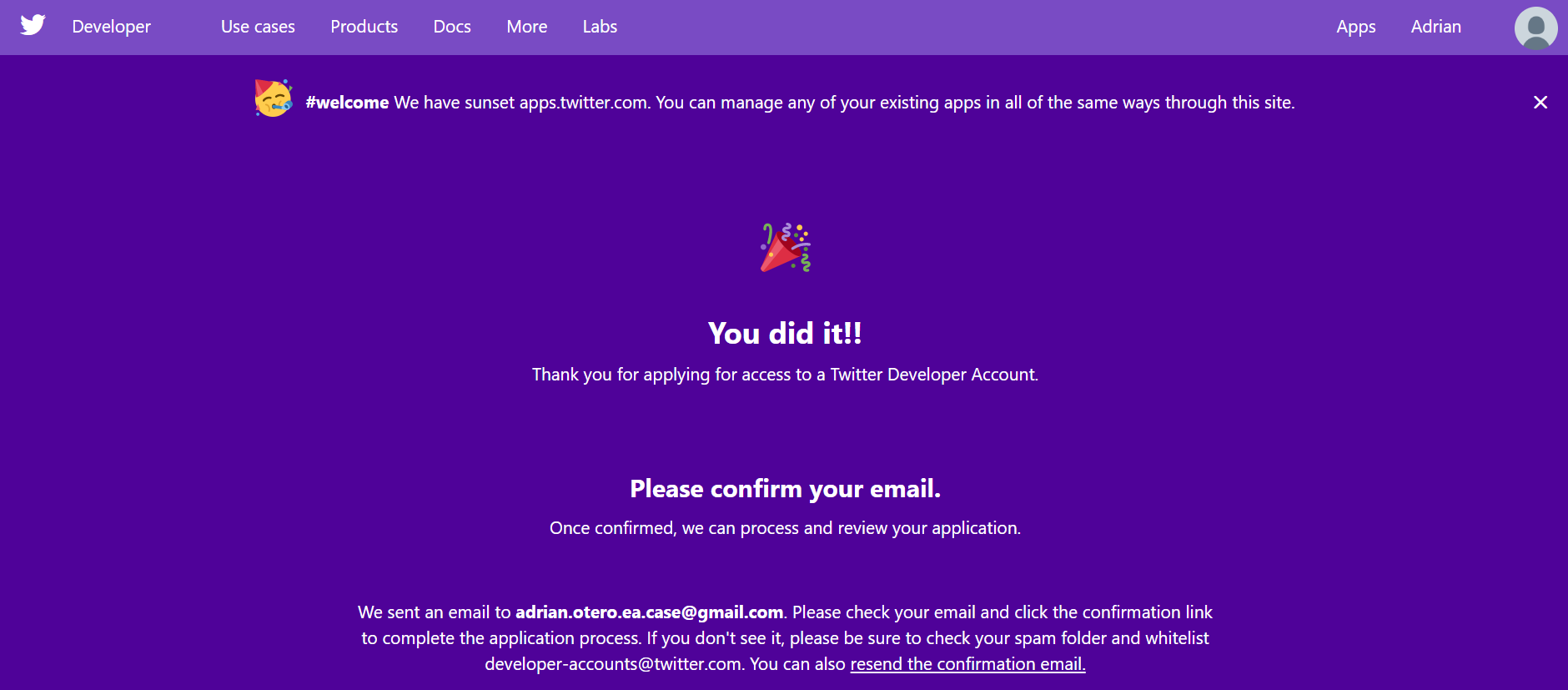


# Redes Sociales – Twitter API

Registrarse en Twitter y acceder a la plataforma de desarrolladores (<https://apps.twitter.com/>) para generar una nueva APP y conseguir las credenciales que nos permitirán acceder a la API desde nuestro sistema.



Presionar en “**Create an app**” y registrarse como desarrollador, rellenando los campos indicados en los diferentes formularios de la plataforma.



# Repositorio GIT

Acceder a Github a través del siguiente enlace: <https://github.com/> y crear una nueva cuenta.

