**Install Python**

* Go to Python official website download Python 3.11.3

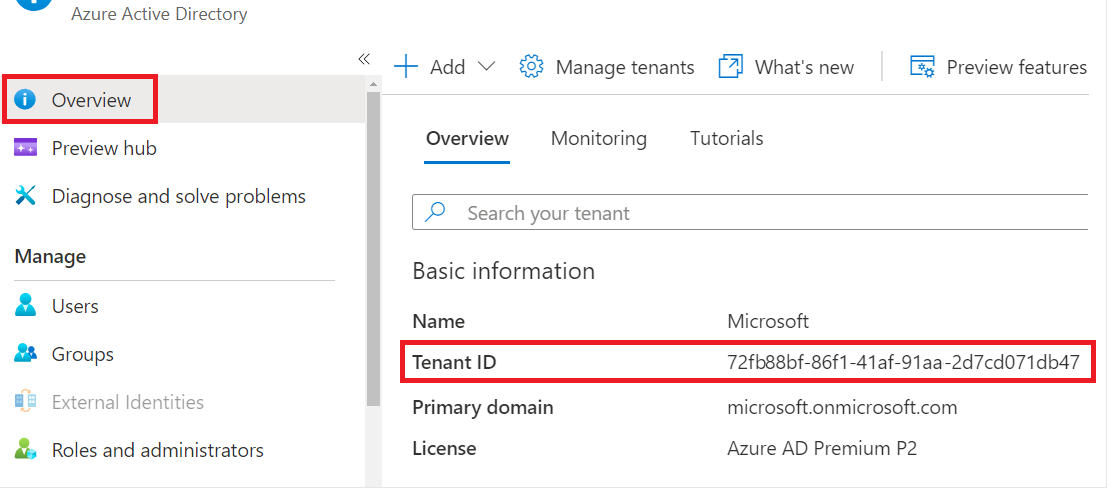
<https://www.python.org/downloads/>

* To install required Python Packages for MS SQL parquet file extraction

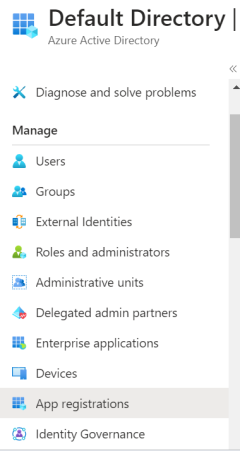
pip install -r requirements.txt.

**Azure Setup**

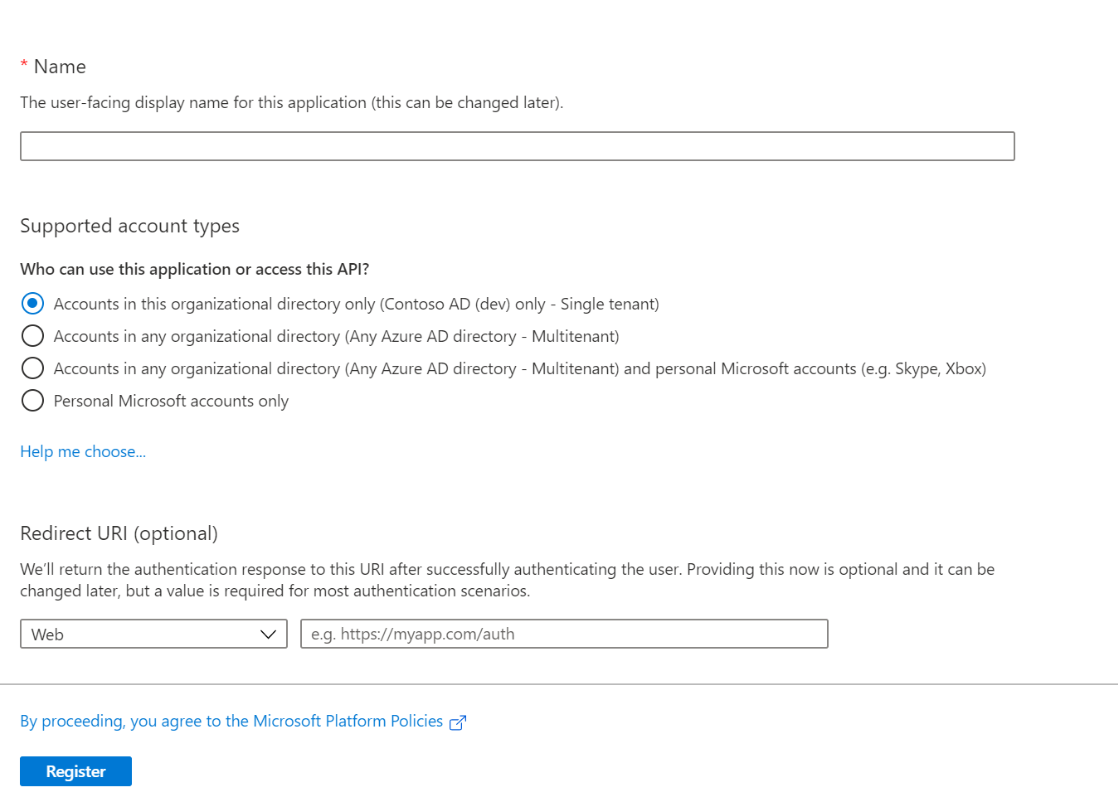
* Create an Azure Active Directory to get TENANT ID
  + Login to your Azure account.
  + In the “Azure Active Directory” pane click on the “Create” button.
  + In the “Basics” tab enter a name for your Azure AD and select the subscription, resource group and location.
  + Click on the “Review + create” tab to preview your settings then click “Create” button to create Azure AD.



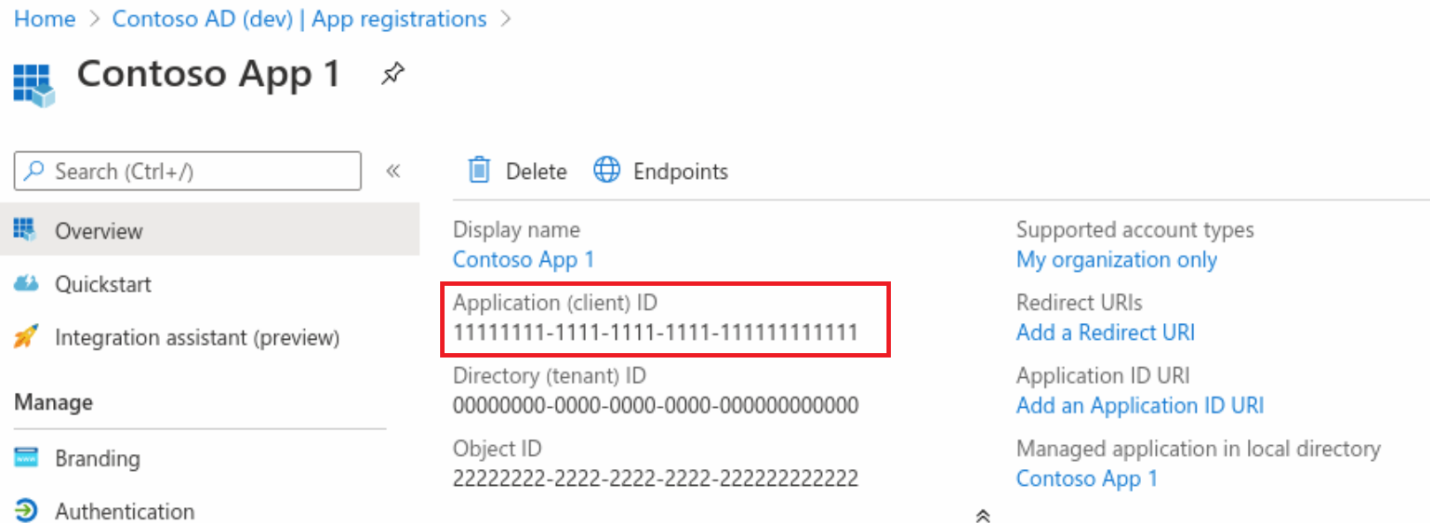
* Register an app to get CLIENT ID and CLIENT SECRET
* Search for and select **Azure Active Directory**.
* Under **Manage**, select **App registrations** > **New registration**.



* Enter a display **Name** for your application.
* Select **Register** to complete the app registration.

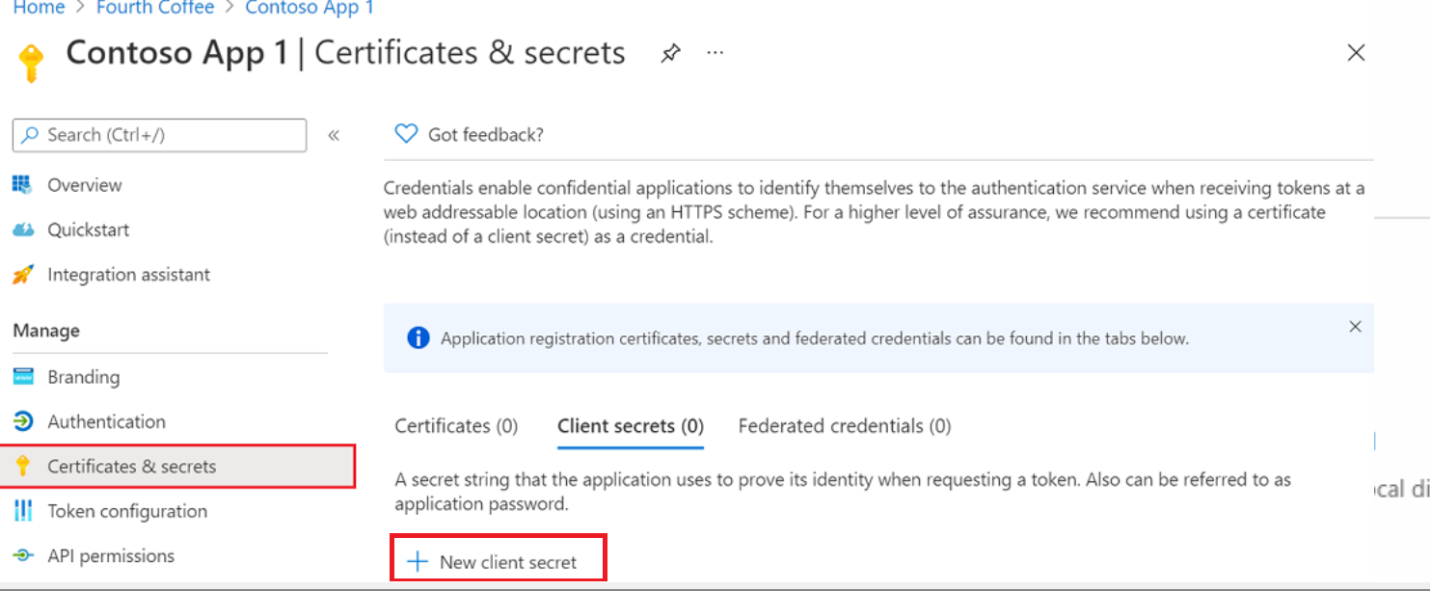
****

The app registration's automatically generated Application (client) ID.

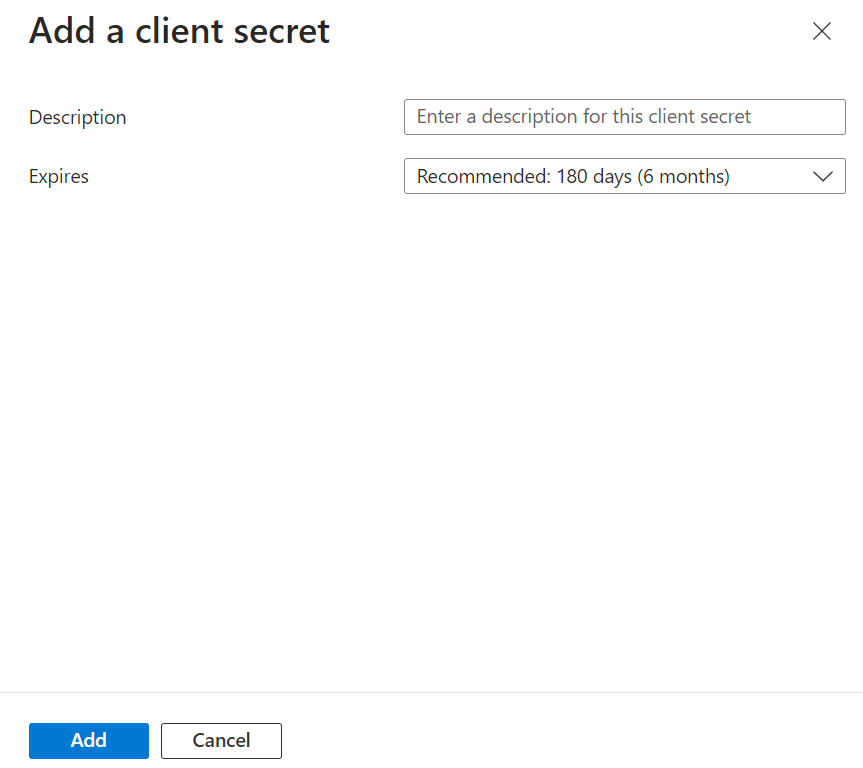


**Add a client secret**

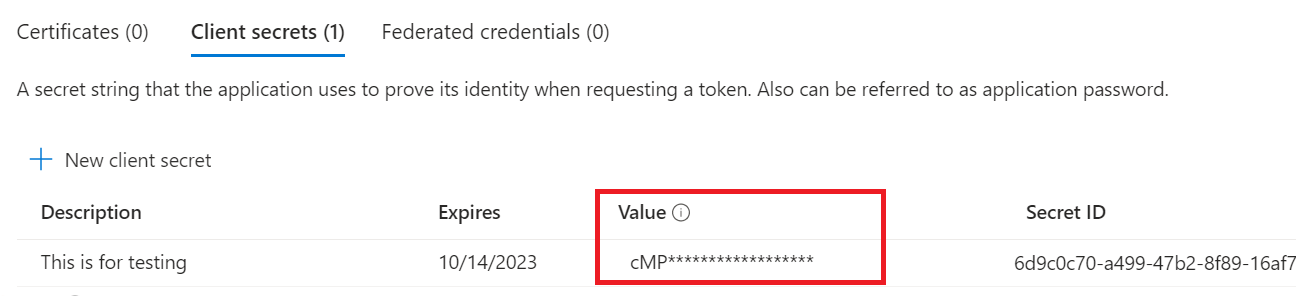
* + In the Azure portal, in **App registrations**, select your application.
  + Select  **Certificates & secrets** > **Client secrets** > **New client secret**.



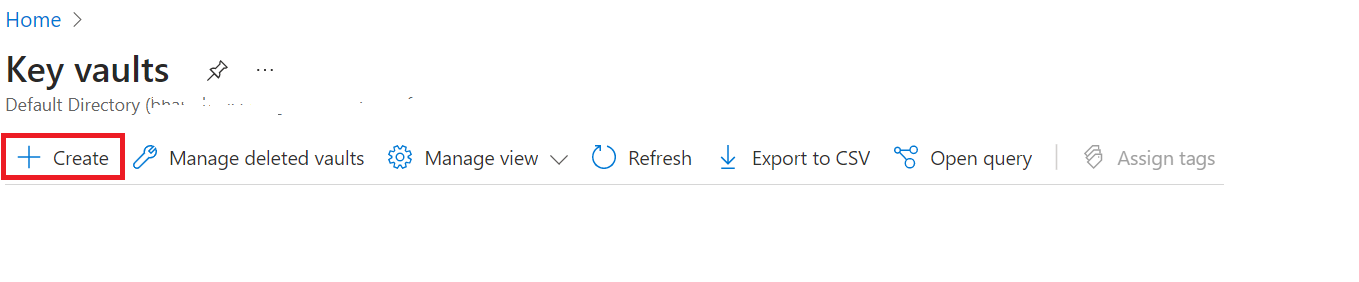
* Add a description and expiration for your client secret.



* + Select **Add**.
  + Record the secret's value for use in your client application code. This secret value is *never displayed again* after you leave this page.

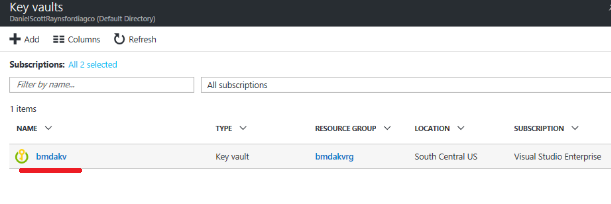


* Create a Azure Key Vault:
* In the Search box, enter **Key Vault** and select.
* On the Key Vault section, choose **Create**.

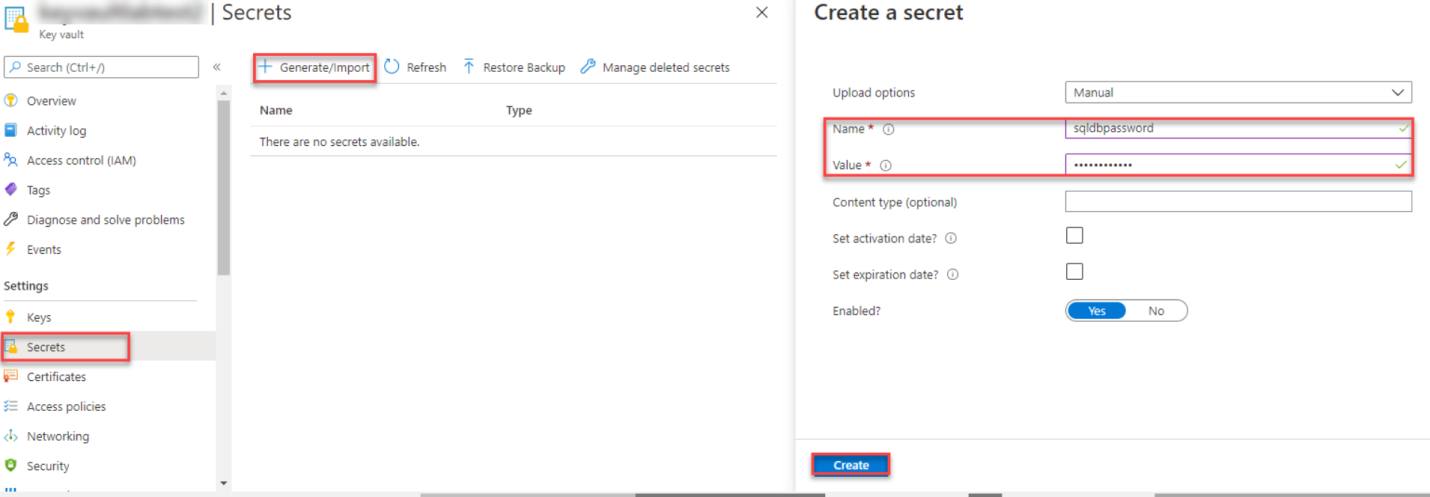


* On the **Create key vault** section provide the following information:
  + **Name**: A unique name is required
  + **Subscription**: Choose a subscription.
  + Under **Resource Group**, select resource .
  + In the Location pull-down menu, choose a location.
  + Leave the other options to their defaults.

Select **Create**.

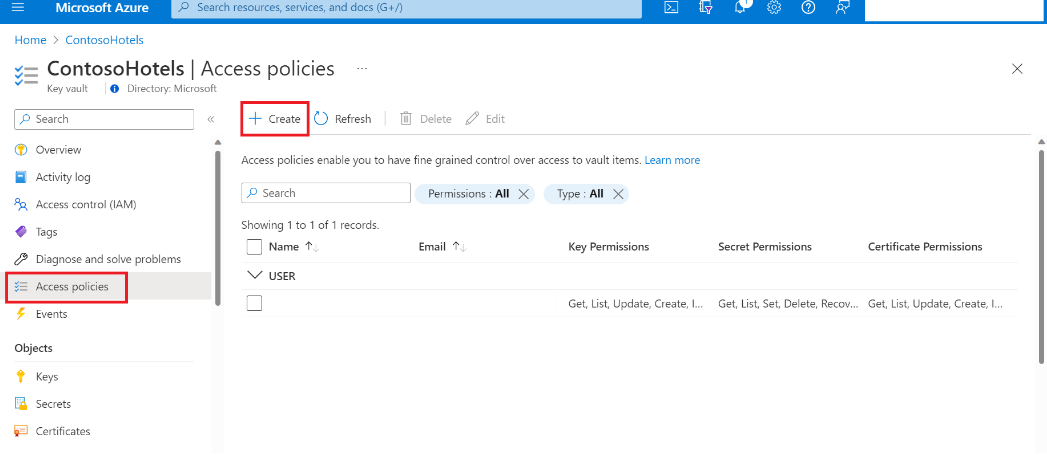


* To add a **SECRET** to the Key vault:
* On the Key Vault settings pages, select **Secrets**.
* Select on **Generate/Import**.
* On the **Create a secret** screen choose the following values:
  + **Upload options**: Manual.
  + **Name**: Type a name for the secret.
  + **Value**: Type a value for the secret.
  + Leave the other values to their defaults. Select **Create**.

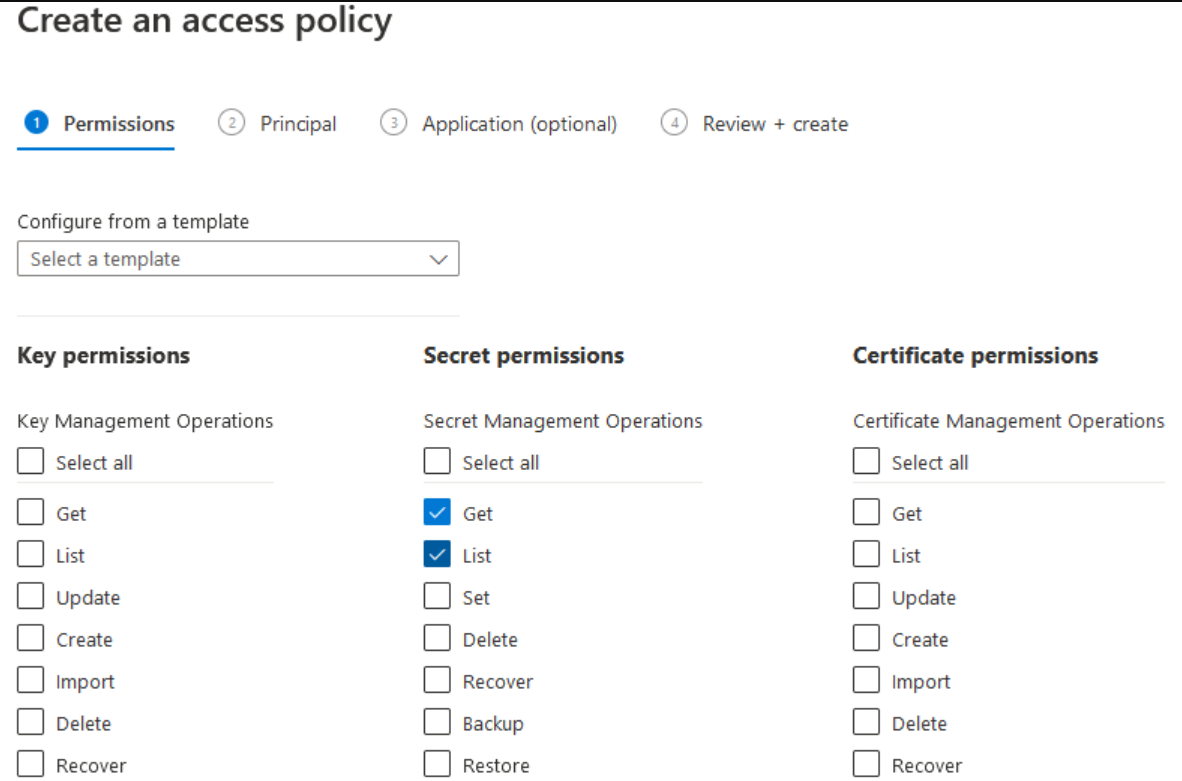


## Assign an access policy

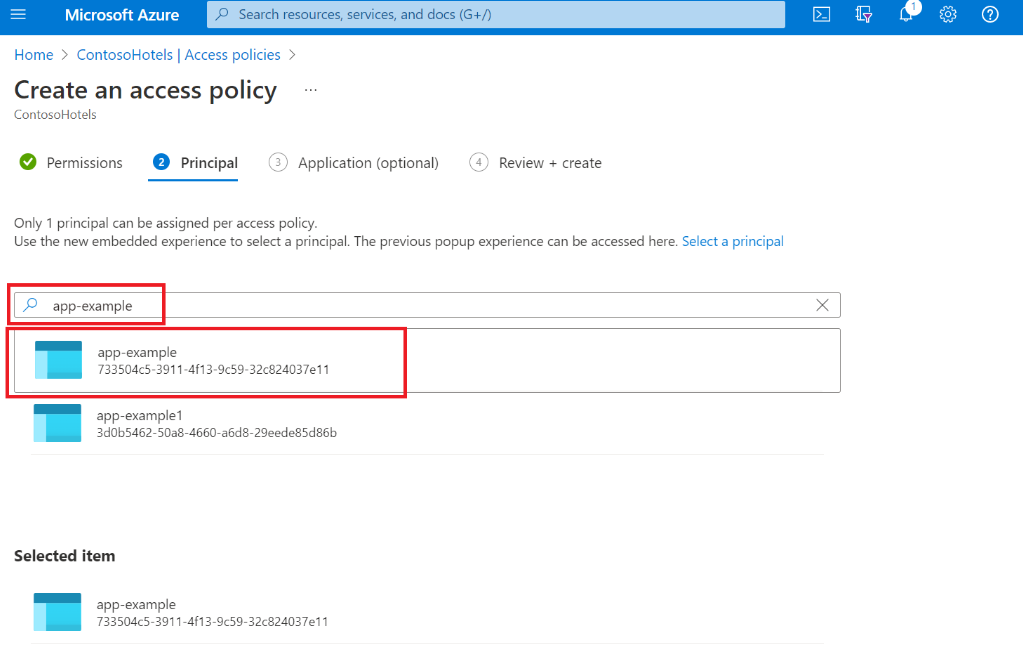
* Navigate to the Key Vault resource.
* Select **Access policies**, then select **Create**:



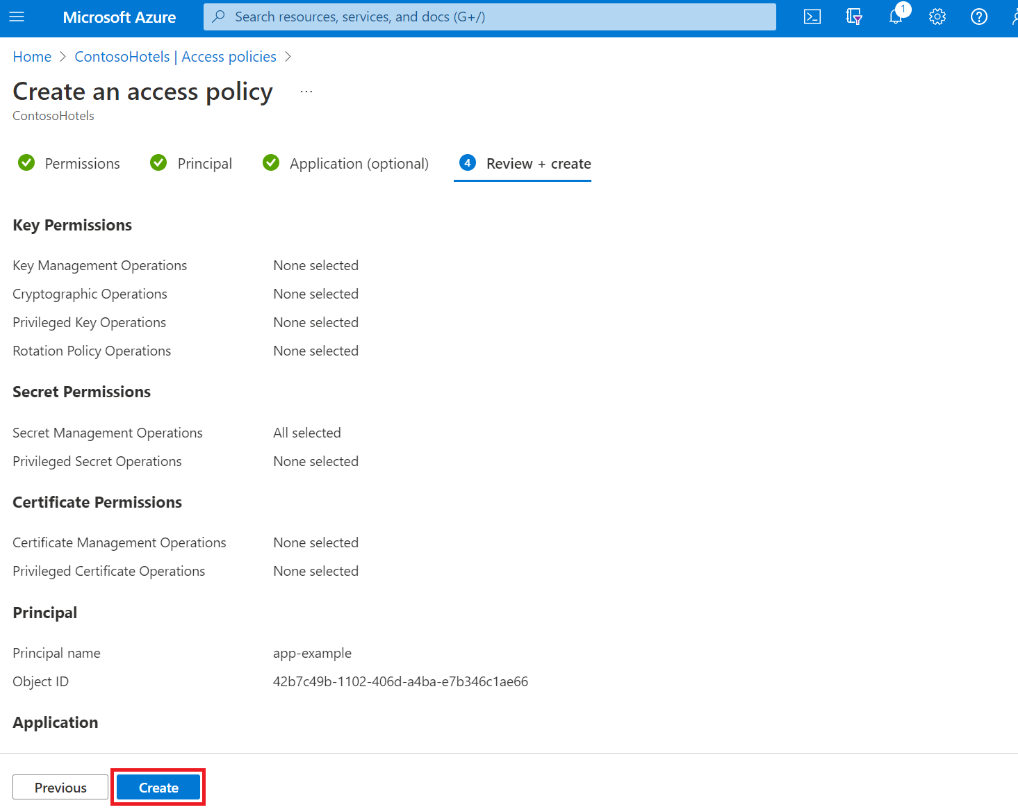
* Select the permissions Get & List under **Secret permissions**.



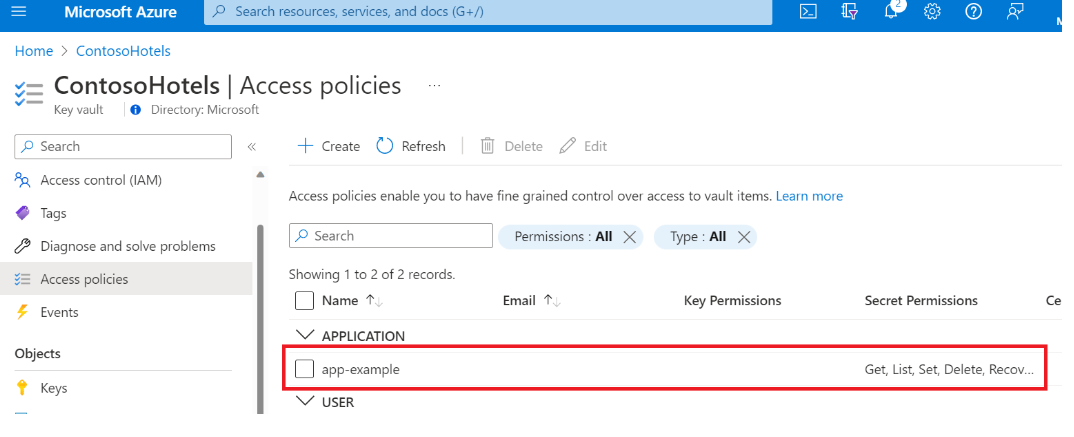
* Under the **Principal** selection pane, enter the name of the user, app or service principal in the search field and select the appropriate result.



* Review the access policy changes and select **Create** to save the access policy.



* Back on the **Access policies** page, verify that your access policy is listed.



* Create a Storage Account
* Select **Storage accounts** and select Create**.**
* Create a Container
* Click the Storage Account.
* Click on the "Containers" and select **“+Container"** tab to create container.
* You need to assign a name and select the access level then click ok.

**Parquet Generator Config**

|  |  |
| --- | --- |
| **EDW\_DB\_SERVER\_HOST** | EDW database server name |
| **EDW\_DB\_NAME** | EDW database name |
| **EDW\_DB\_USER\_NAME** | EDW database username |
| **GP\_DB\_SERVER\_HOST** | GP database server name |
| **GP\_DB\_USER\_NAME** | GP database username |
| **KEY\_VAULT\_NAME** | Name of the Azure key vault. |
| **EDW\_KEY\_VAULT\_SECRET\_NAME\_DB\_PASSWORD** | Secret name for EDW database password |
| **GP\_KEY\_VAULT\_SECRET\_NAME\_DB\_PASSWORD** | Secret name for GP database password |
| **MAX\_JOB\_COUNT** | Enter the Max job count to make the parallel execution |
| **QUERY\_GET\_ALL\_TABLES\_FOR\_DB** | SQL query to get all tables from GP database |
| **QUERY\_GET\_ALL\_RECORDS\_FROM\_TABLE** | SQL query to fetch all records from the table GP database |
| **AZURE\_STORAGE\_ACCOUNT\_NAME** | Name of the azure storage account |
| **AZURE\_STORAGE\_ACCOUNTKEY\_KEY\_VAULT\_SECRET\_NAME** | Secret name for the azure storage account key |
| **AZURE\_STORAGE\_ACCOUNT\_CONTAINER\_NAME** | Container name in the azure storage account |

**How to Run Python**

Follow this command to run python file in command prompt.

* Python <File Path\Filename>

Example command:

python gp\_data\_extractor.py

**Screen Blob Storage**

