Common Data Model Folder Setup Instructions

# Background

Data stored in the Common Data Model format provides semantic consistency across apps and deployments. With the evolution of the Common Data Model metadata system, the model brings the same structural consistency and semantic meaning to the data stored in Azure Data Lake Storage Gen2 with hierarchical namespaces and folders that contain schematized data in standard Common Data Model format. The standardized metadata and self-describing data in an Azure data lake facilitates metadata discovery and interoperability between data producers and consumers such as Power BI, Azure Data Factory, Azure Databricks, and Azure Machine Learning service.

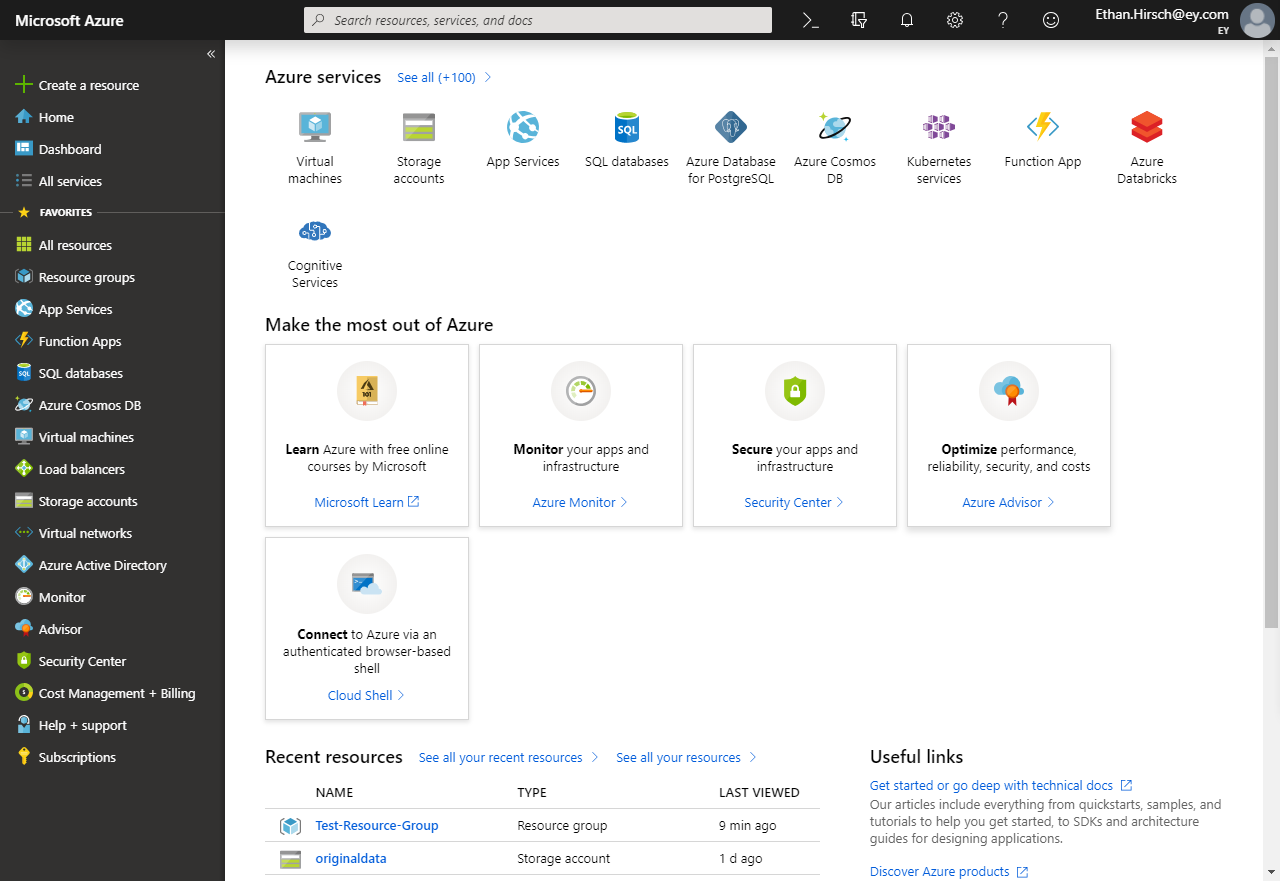
# Prerequisites

1. An Active Azure Subscription.
2. A preconfigured and valid CDM model.json file.
3. A Power BI Subscription.

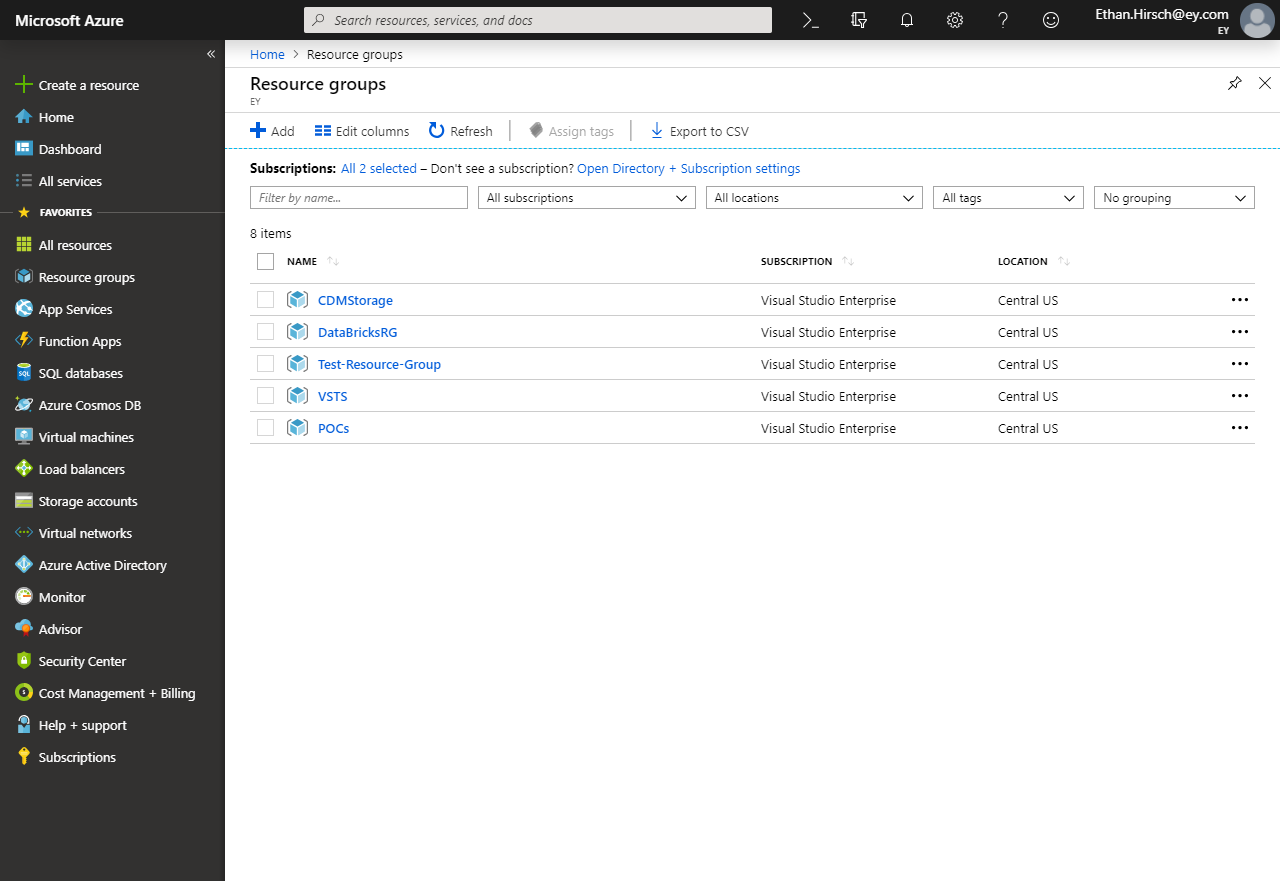
# Create Azure Data Lake Storage Gen2 Storage Account

There are multiple ways to setup an Azure Data Lake Storage Gen2. The instructions below focus on performing these steps manually through the Azure Portal.

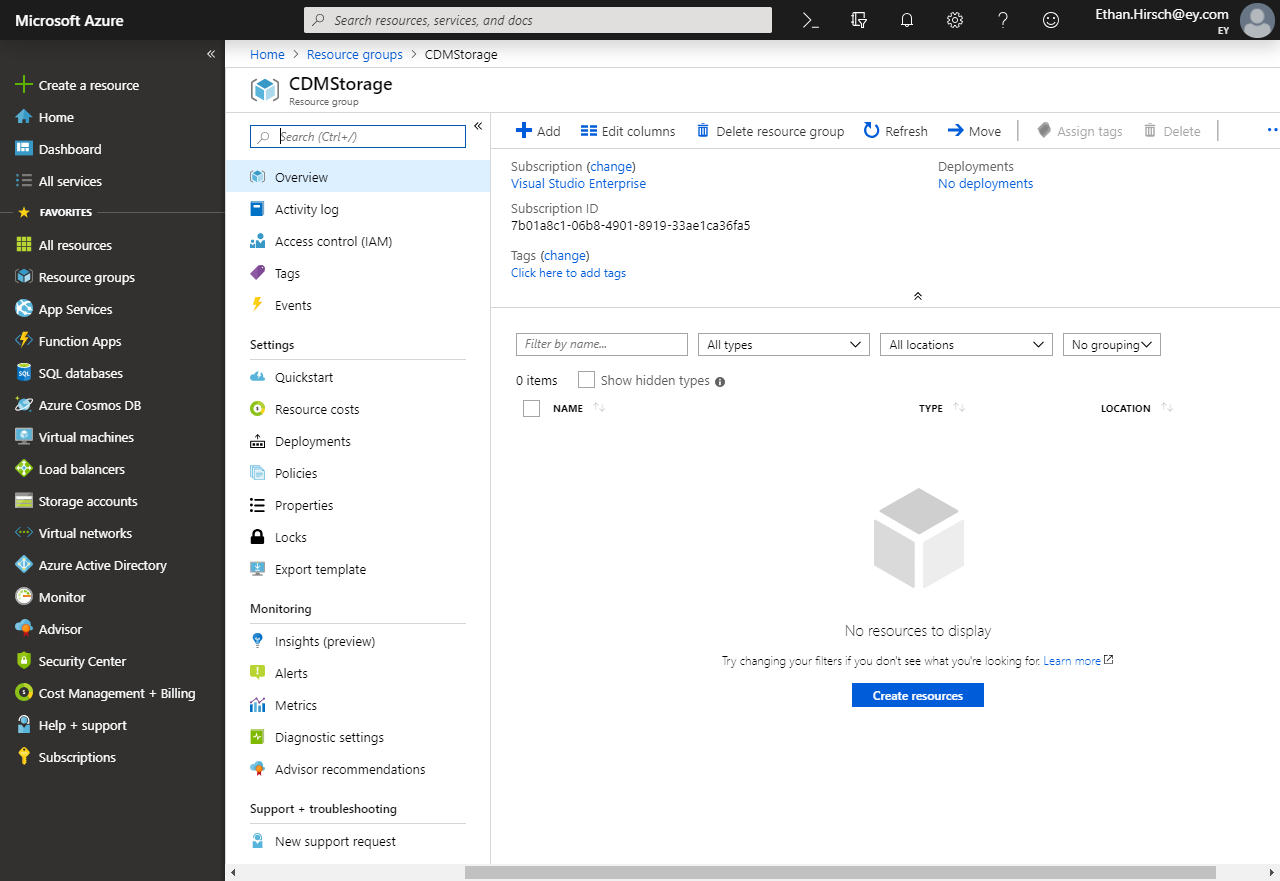
1. Login to your Azure Portal at <https://portal.azure.com>.
2. Before continuing, ensure that you are logged into the same tenant as the Power BI Subscription that will consume this data.
3. Navigate to Resource Groups.



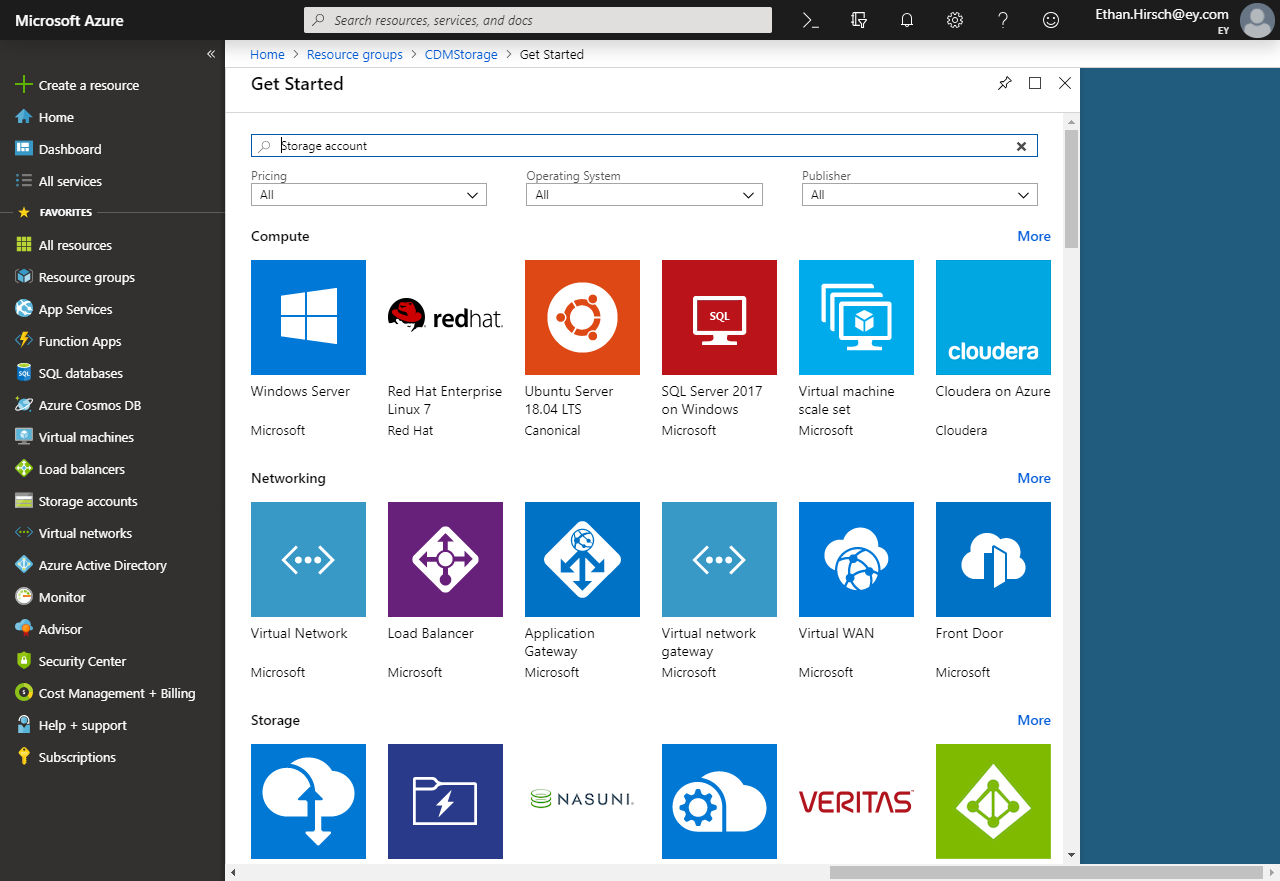
1. Open the Resource Group that the Azure Data Lake Storage Gen2 will be provisioned within. A new Resource Group can also be created for this step if needed or preferred.



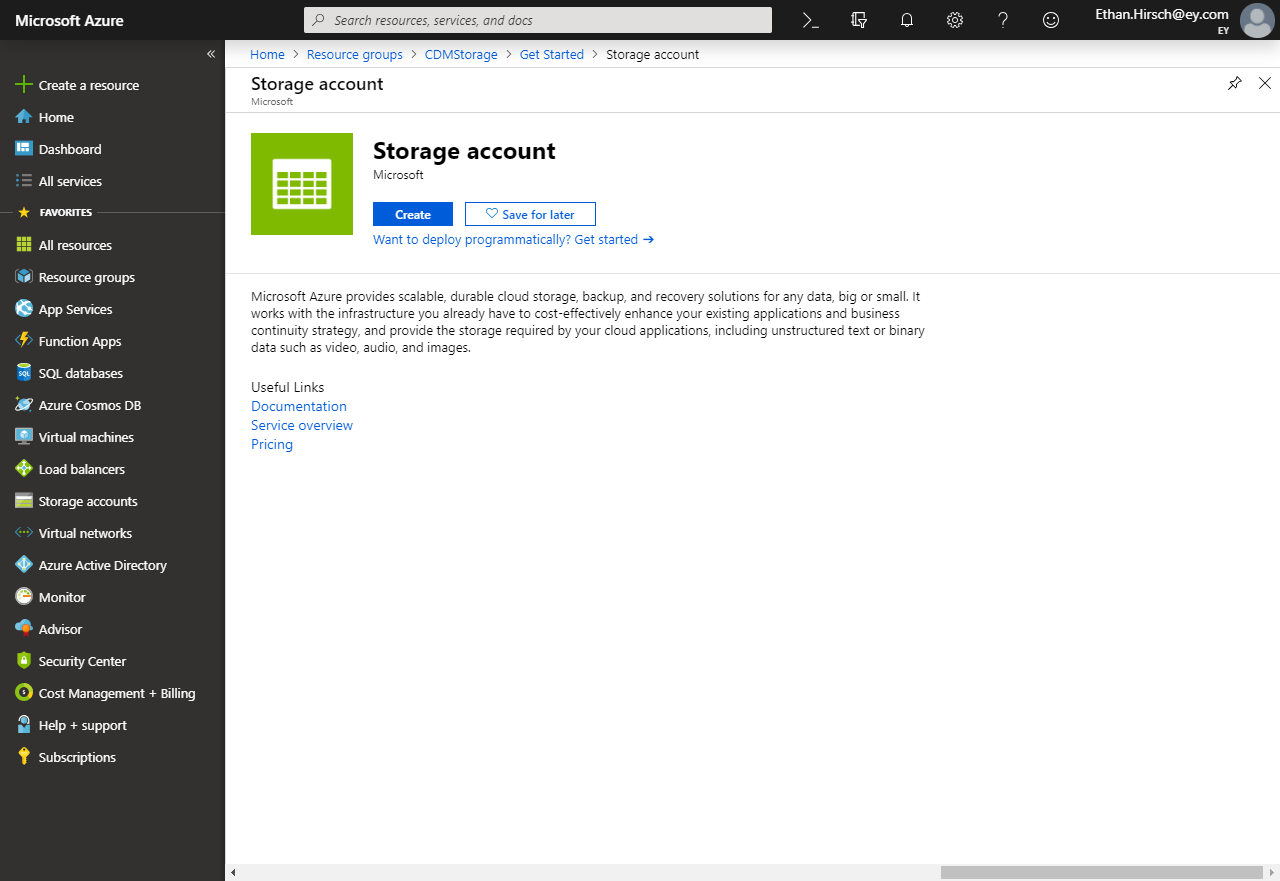
1. Inside the Resource Group, press the Add button to create a new Resource.



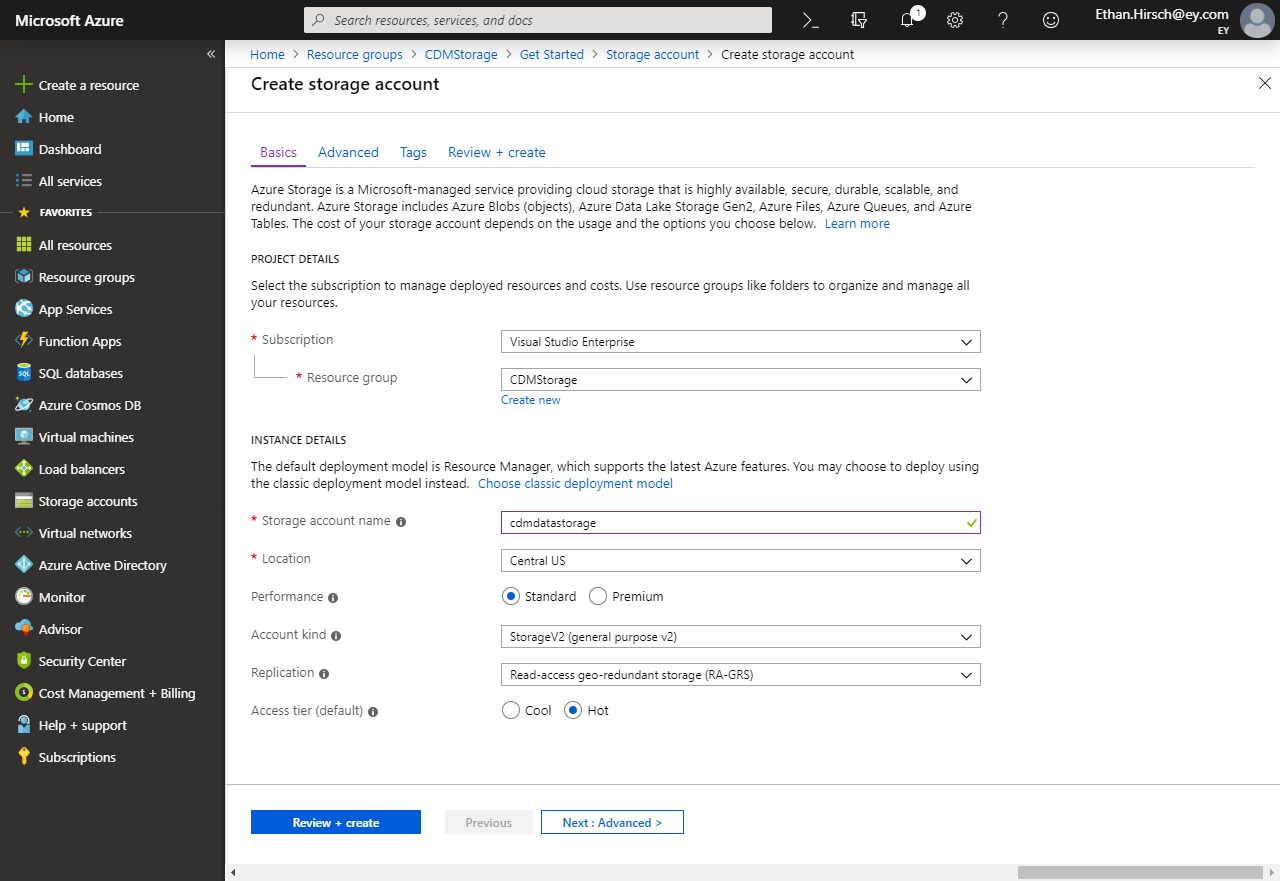
1. In the Get Started window, search for and select Storage Account.



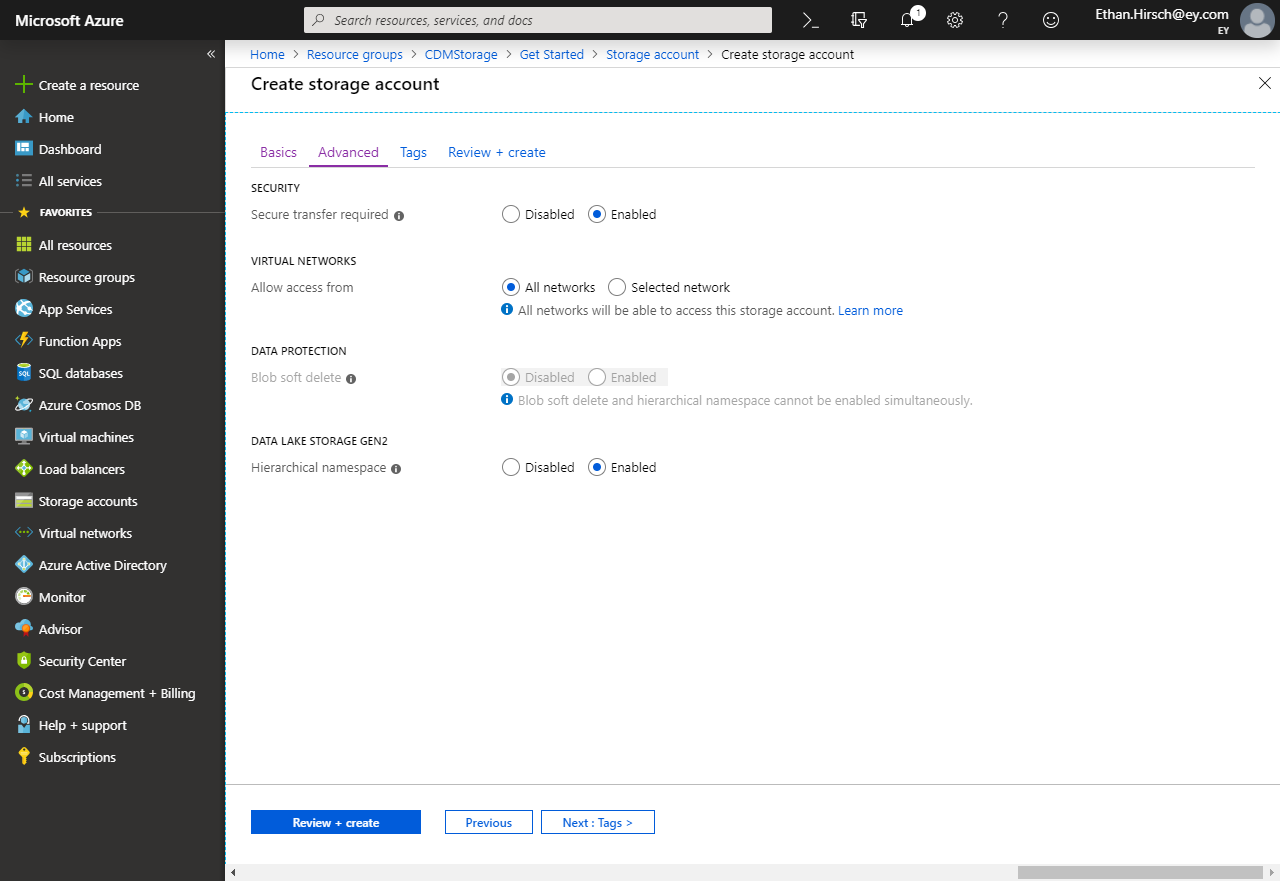
1. Press Create on the Storage account page.



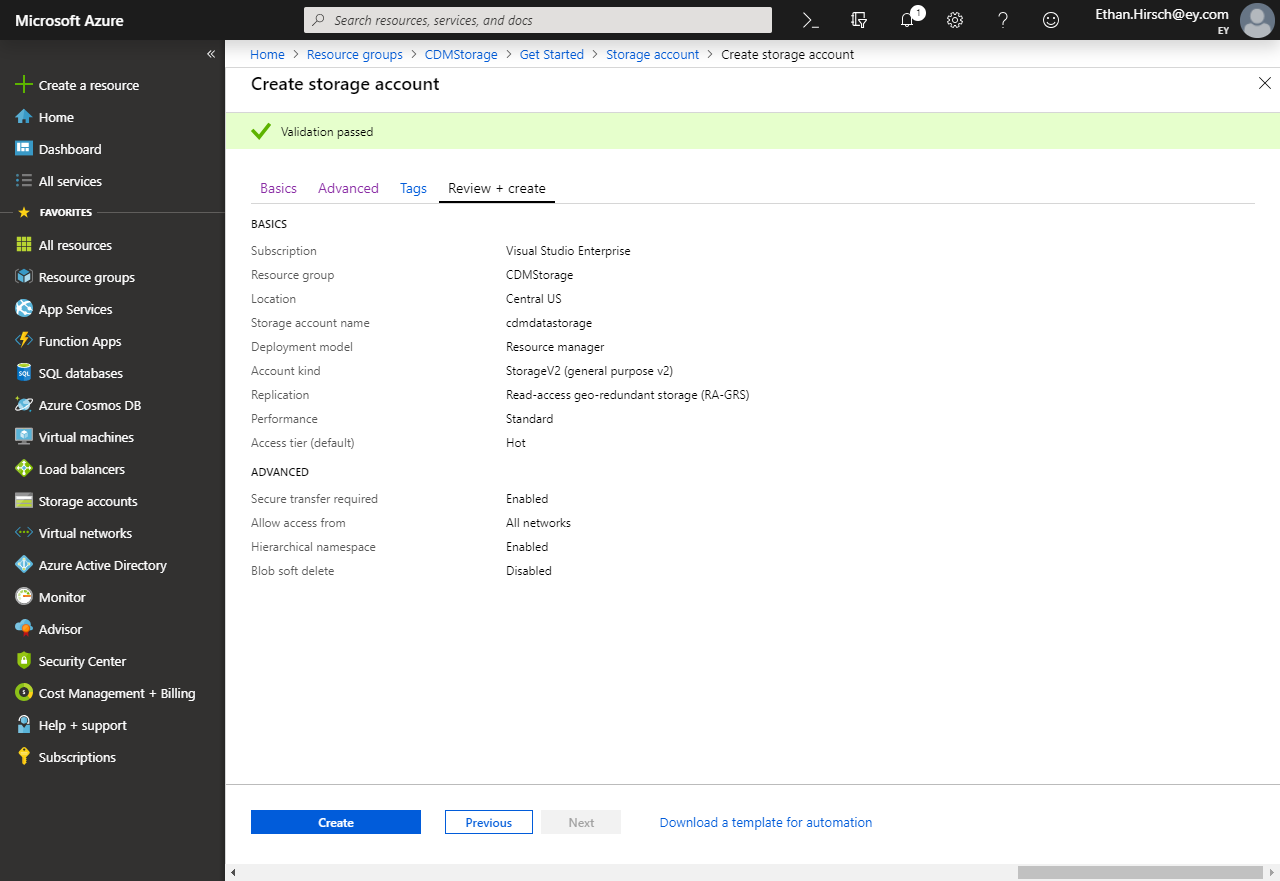
1. Ensure the following information is selected on the first Basics tab in the Create storage account screen.
   1. **Name:** You’re preferred name of the Storage Account.
   2. **Location**: This should match the location of your Power BI tenant. Instructions are provided [below](#_Identify_your_Power) for finding this information.
   3. **Account Kind:** Storage V2 (general purpose v2)
   4. The default values for Performance and Access tier (default) can be used or changed.
   5. It is recommended that Replication be set to Read-access geo-redundant storage (RA-GRS).



1. Press Next : Advanced >to navigate to the next screen.
2. On the Advanced tab, ensure that Hierarchical namespace is enabled.



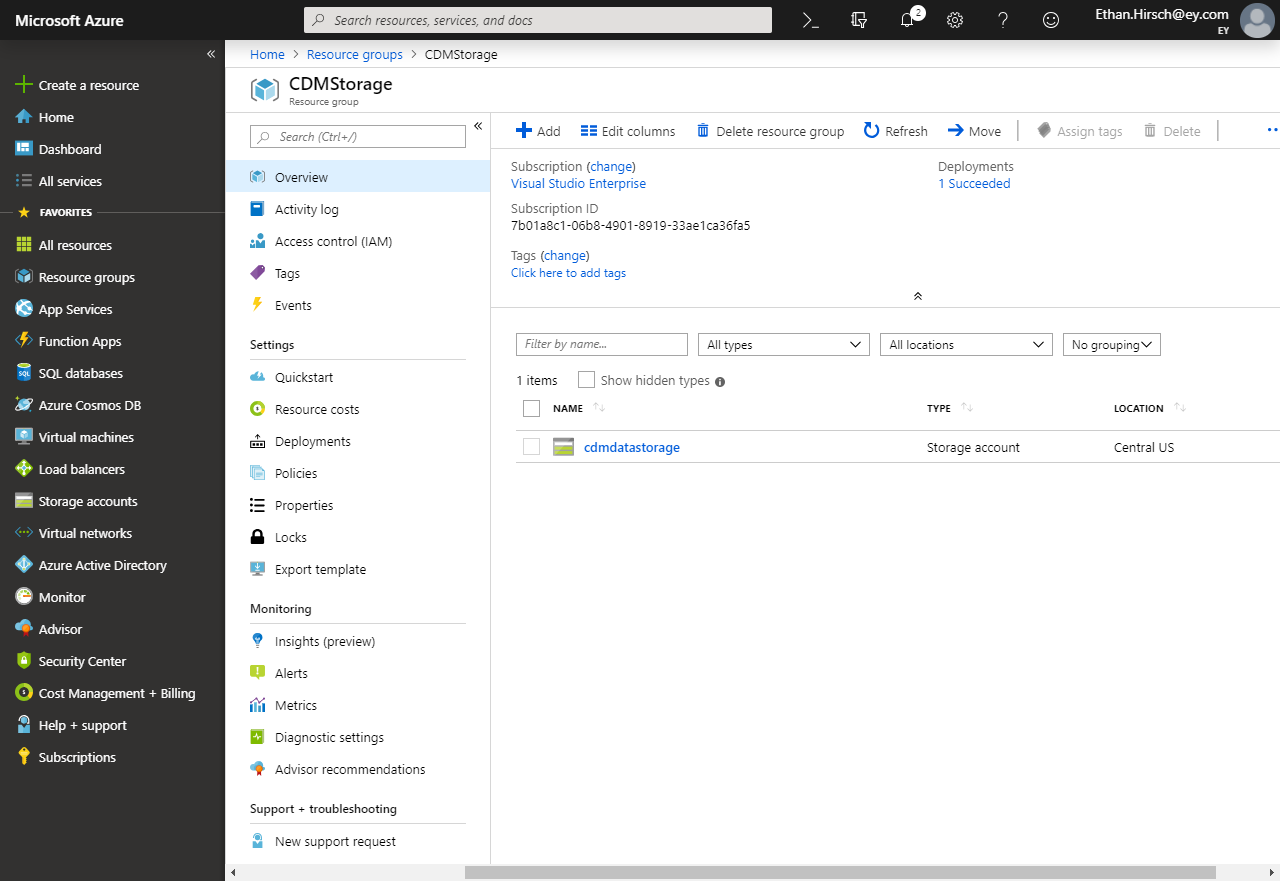
1. Press Review + Create.
2. Press Create to create your Storage account after validation finishes.



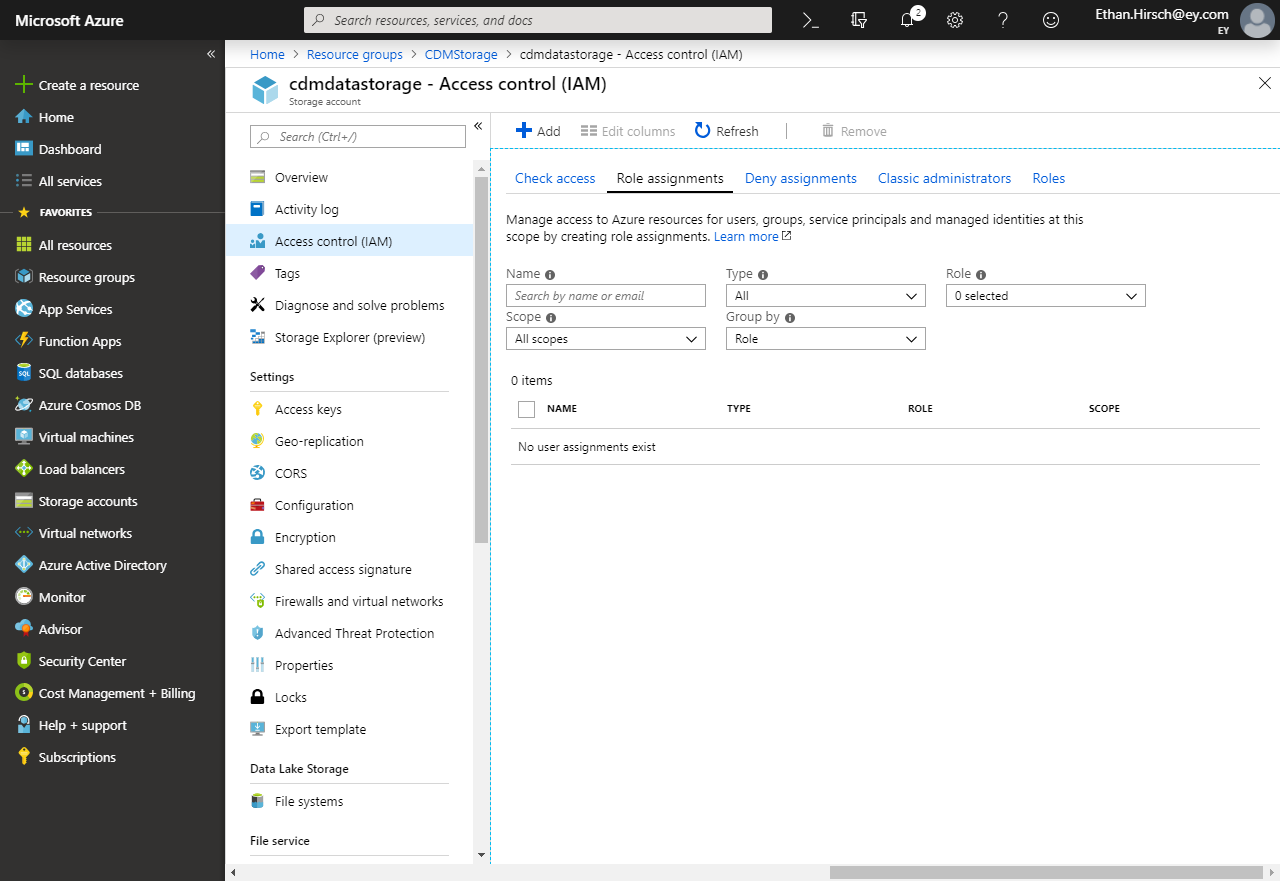
1. After some time, your Azure Storage Account will be created.

# Grant the Power BI Service a Reader Role

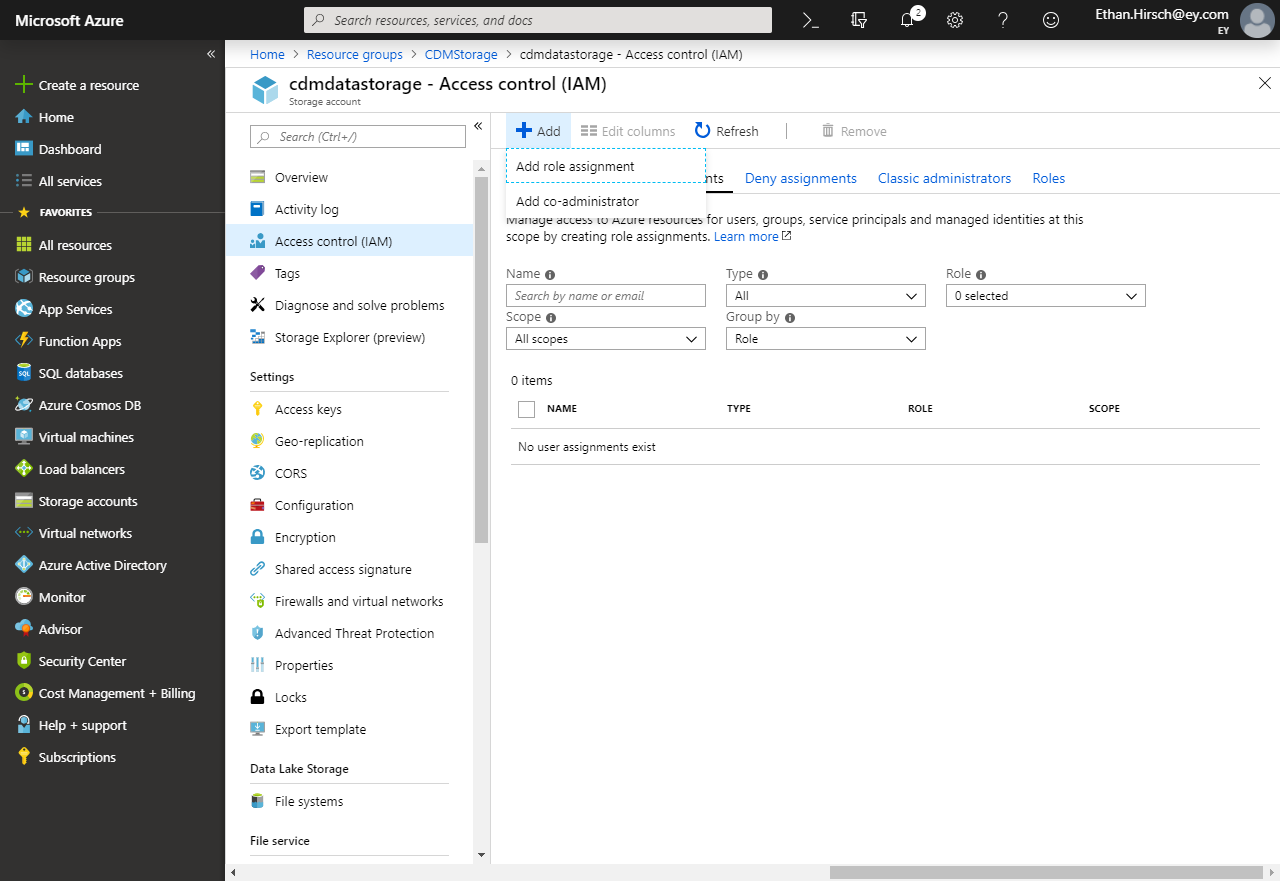
1. Open the Storage account created in the previous step. The simplest location to find it is within its Resource Group.



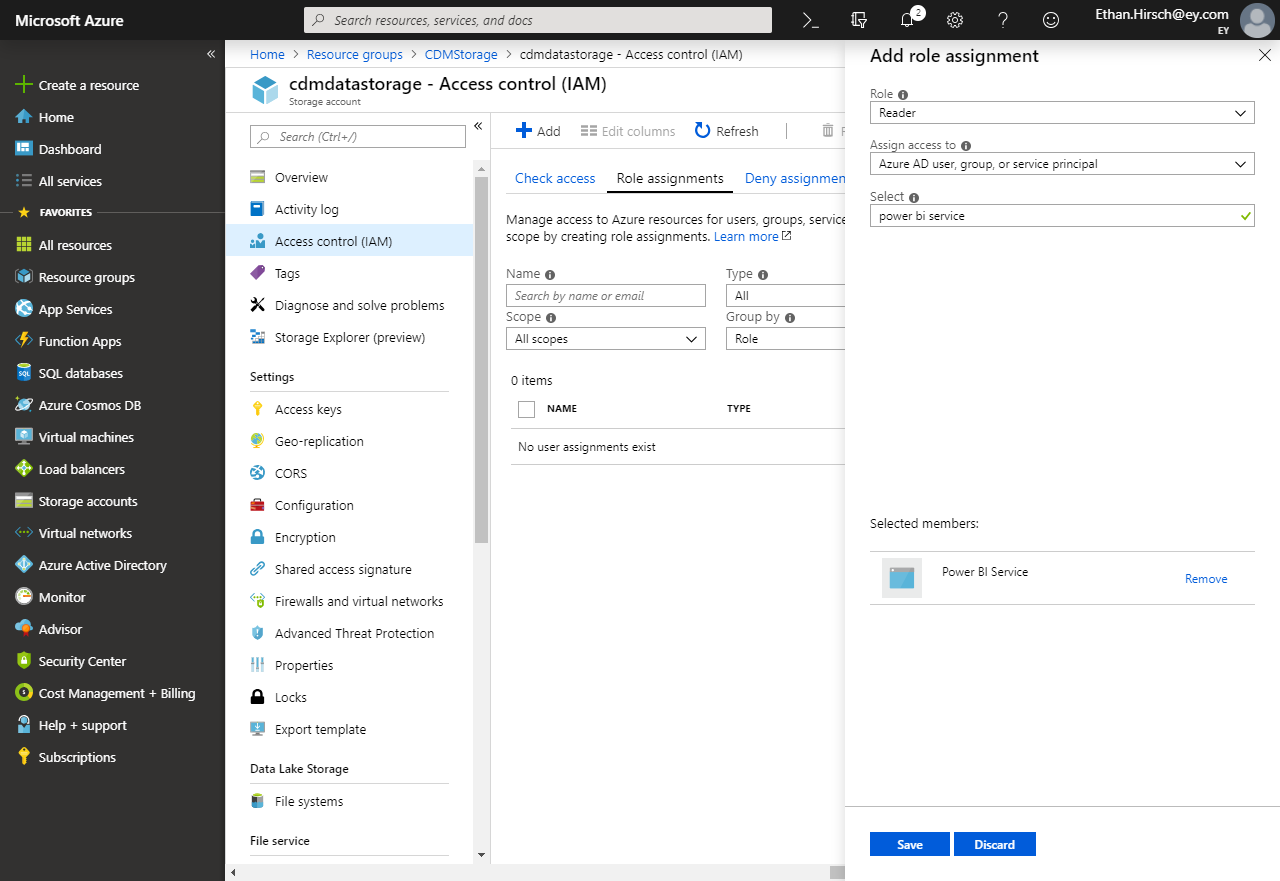
1. Inside the Storage account, click Access control (IAM).
2. Inside the Access control (IAM) pane, click the Role assignments tab.



1. Click Add and then click Add role assignment.



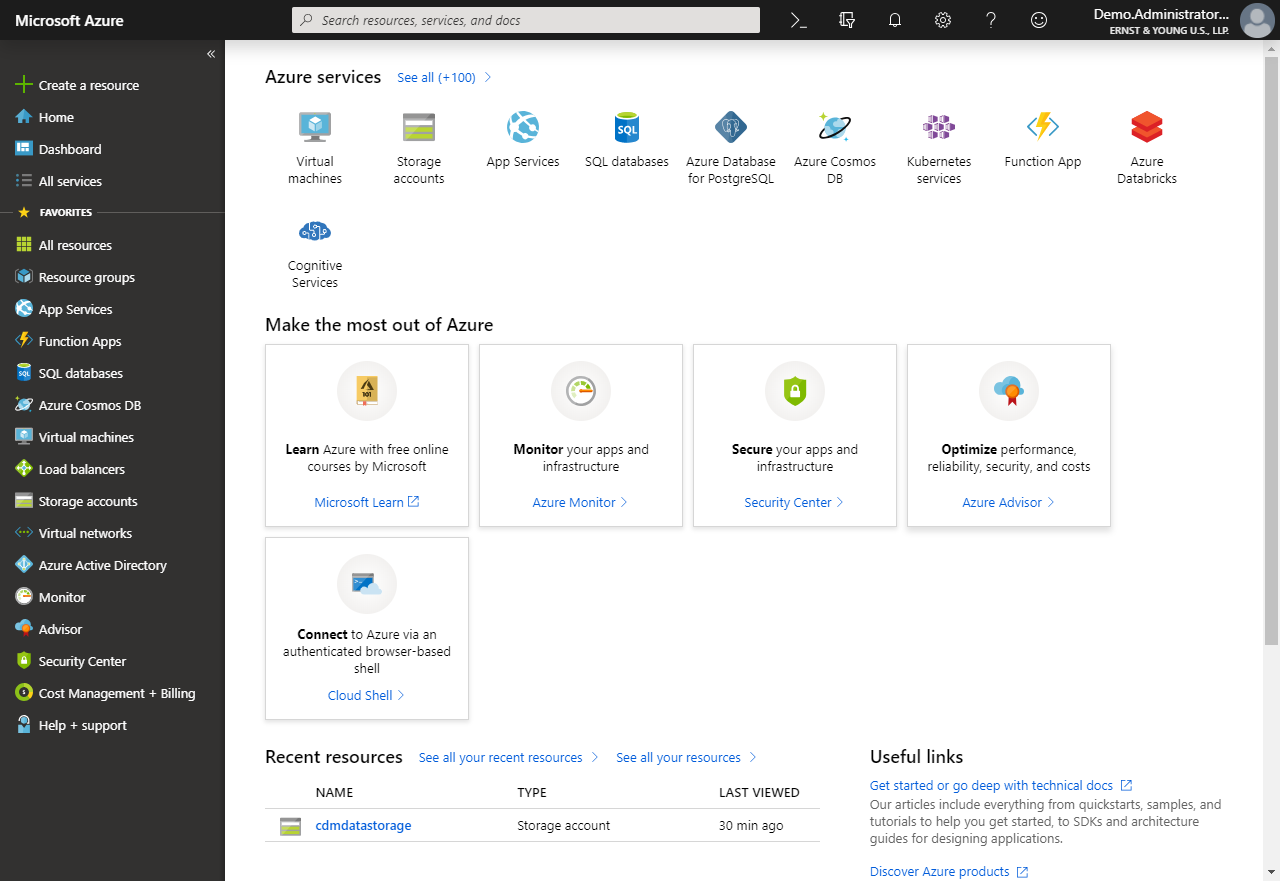
1. In the Add Role assignment pane:
   1. Select Reader within the Role dropdown.
   2. Search for Power BI Service in the Select textbox.
   3. Click on Power BI Service.
   4. Press Save.



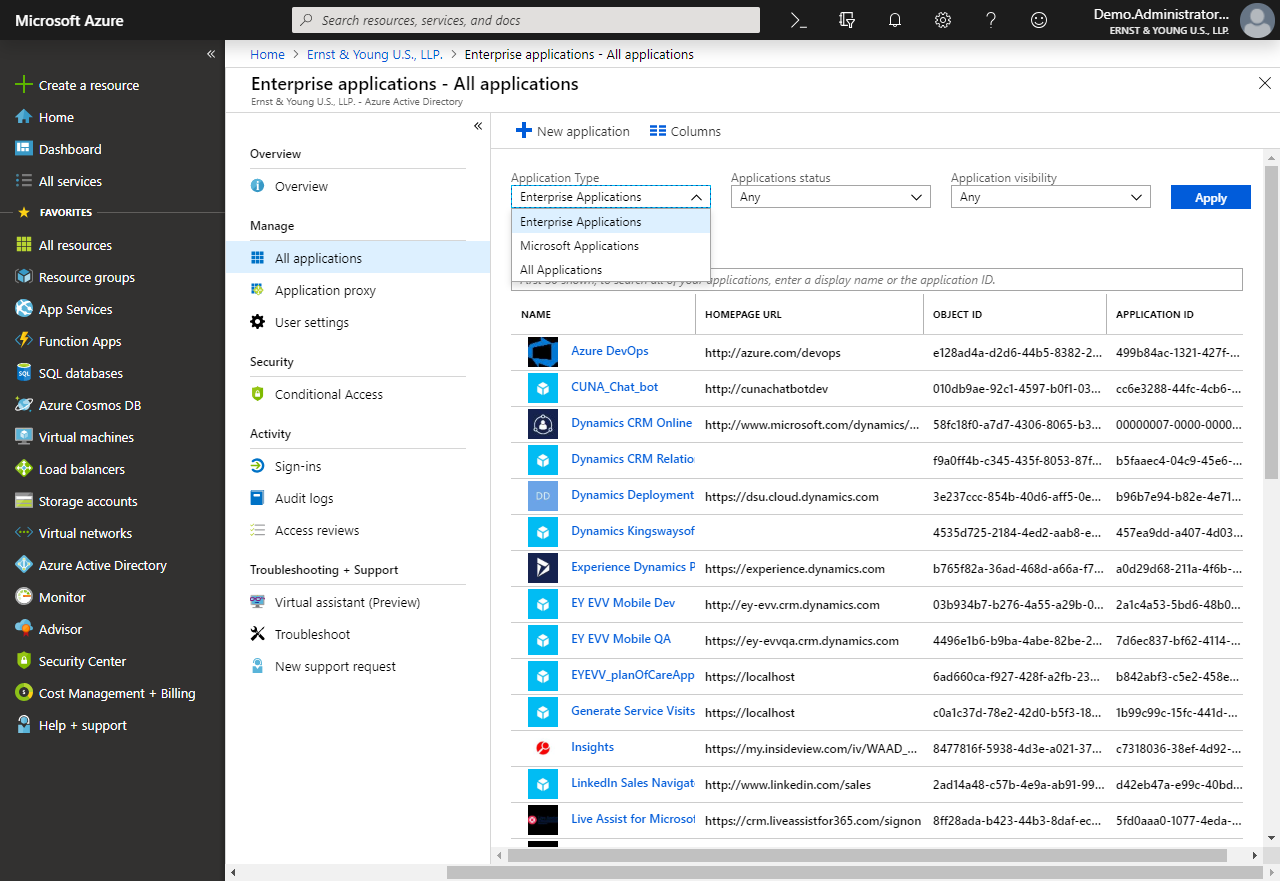
1. The Power BI Service will now have access to this Storage account, but Power BI may take 30 minutes to reflect these changes. This will not affect subsequent sections.

# Identify Power BI Service Ids

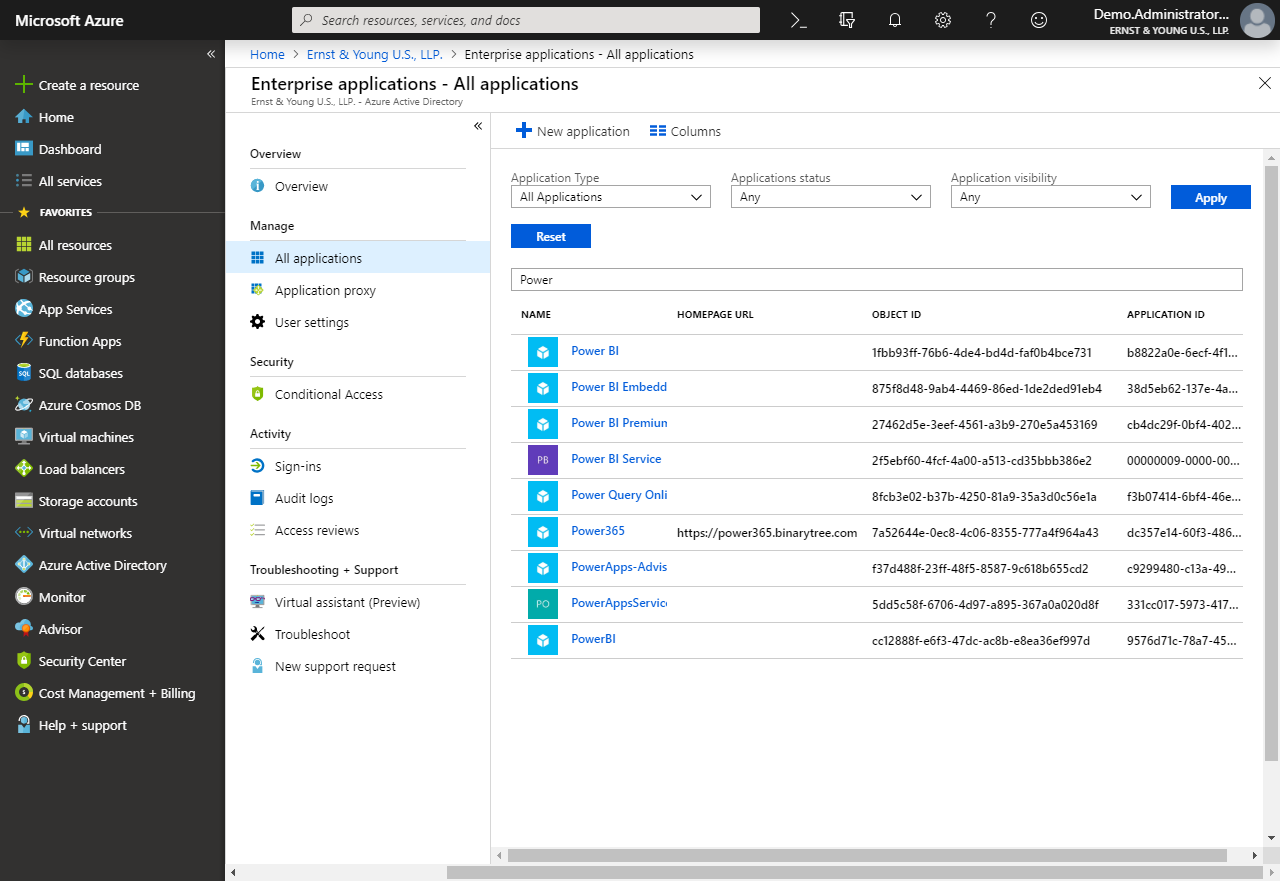
1. Login to <https://portal.azure.com>.
2. Navigate to Azure Active Directory.



1. Within Azure Active Directory, select Enterprise applications.
2. In the All Applications window, change Application Type to All Applications and press Apply.

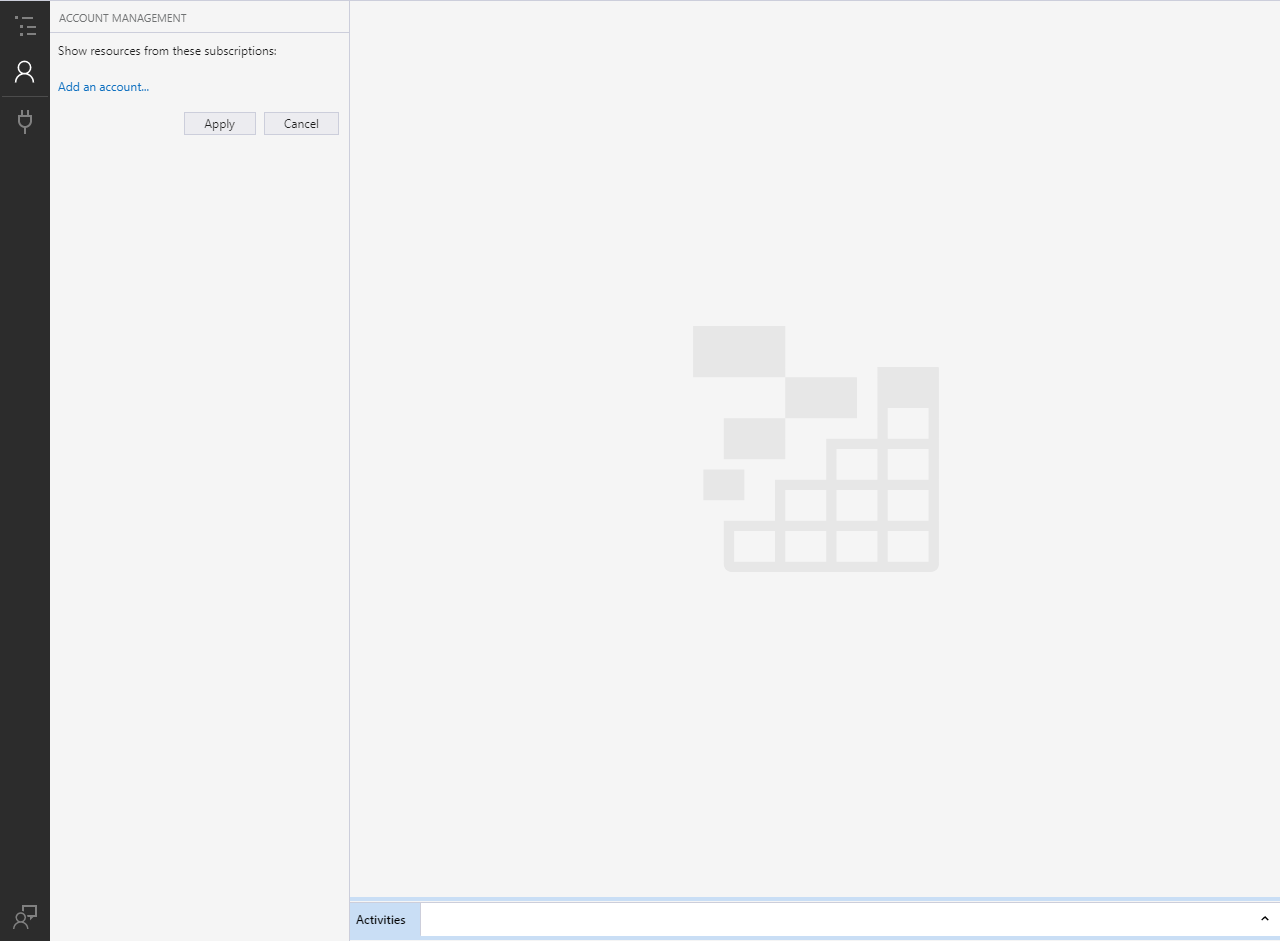


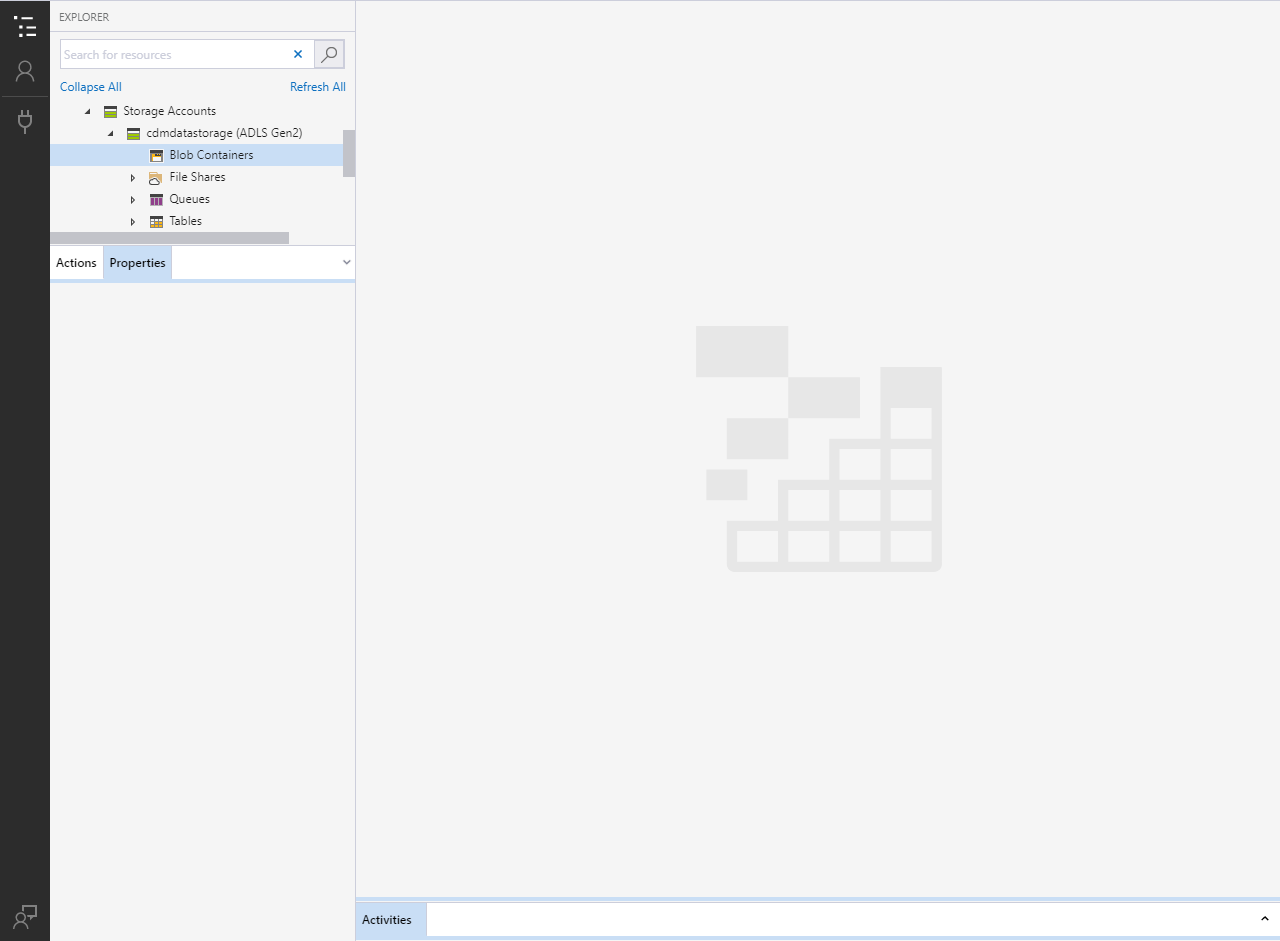
1. In the textbox in the All Applications page, type in Power.
2. Save the Names and Object ids for the following Applications for use in the next section:
   1. Power BI Service
   2. Power BI Premium
   3. Power Query Online



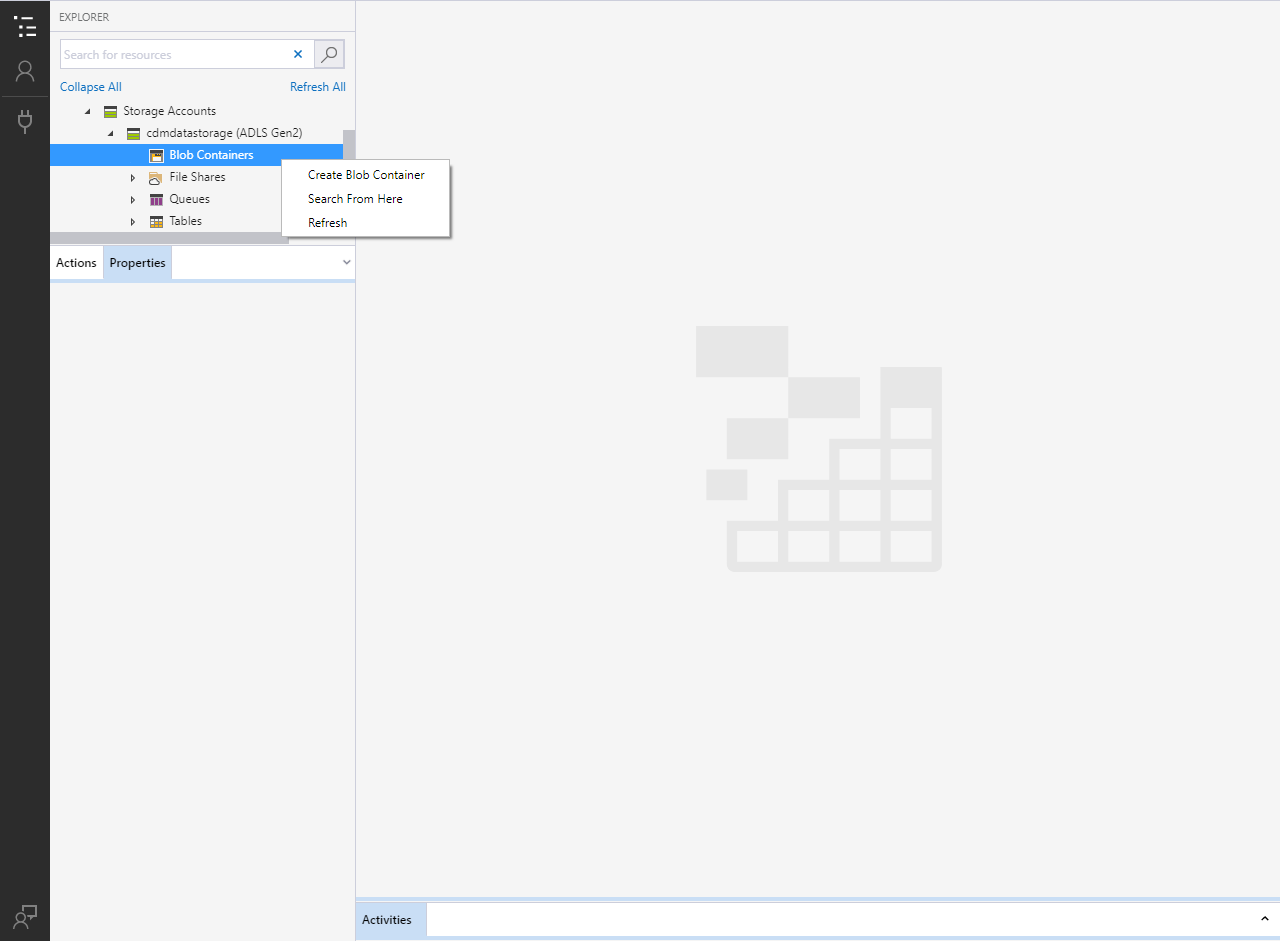
# Create a file system for the CDM data

## Initial Setup

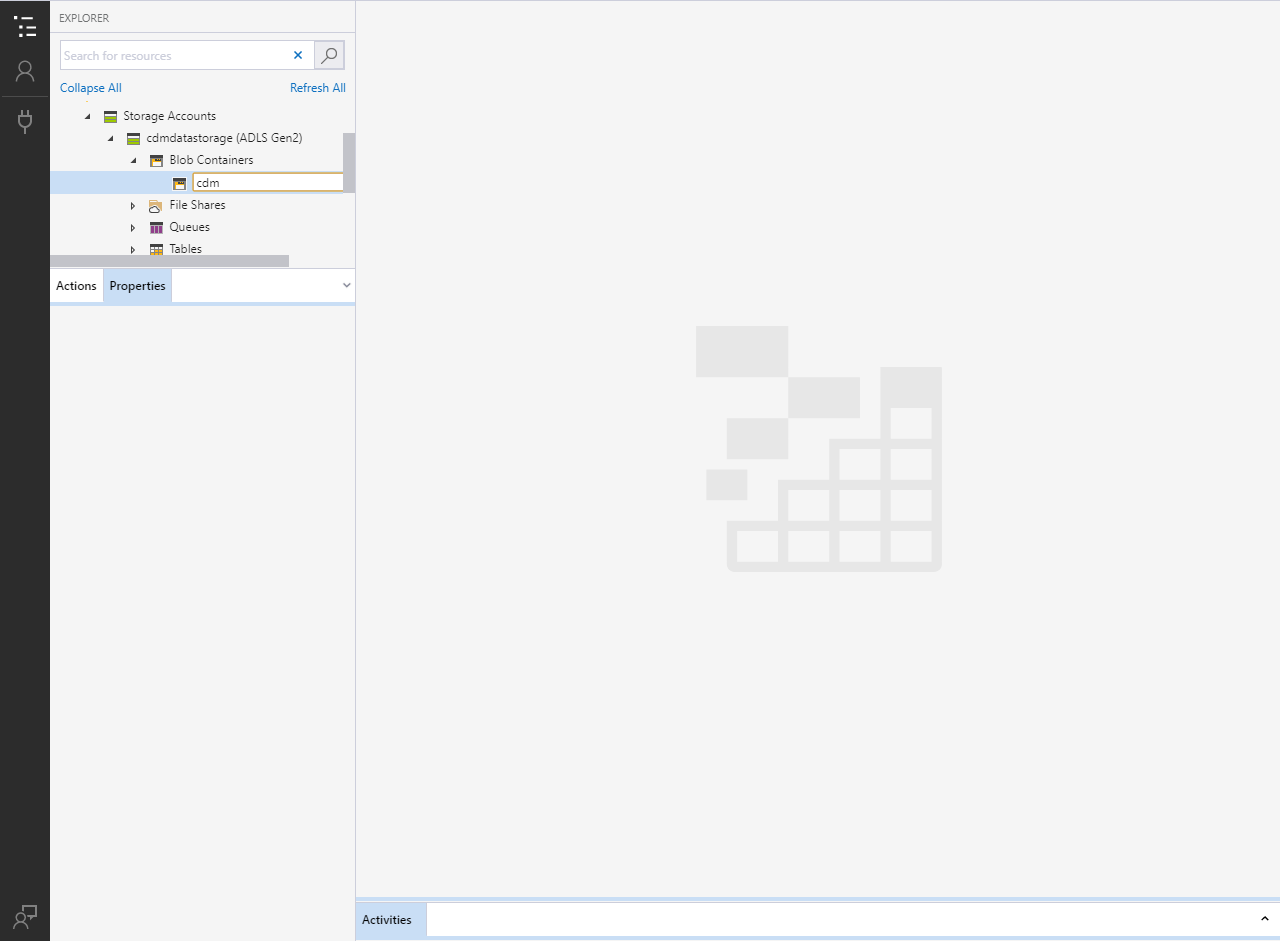
1. Install Azure Storage Explorer from this link: <https://azure.microsoft.com/en-us/features/storage-explorer/>
2. Open Azure Storage Explorer after it installs.
3. Click Add an account… and login using the same account you use to access your Azure tenant. 
4. Expand your Storage Account within Azure Storage Explorer so that Blob Containers is visible.



1. Right click on Blob Containers and select Create Blob Container.

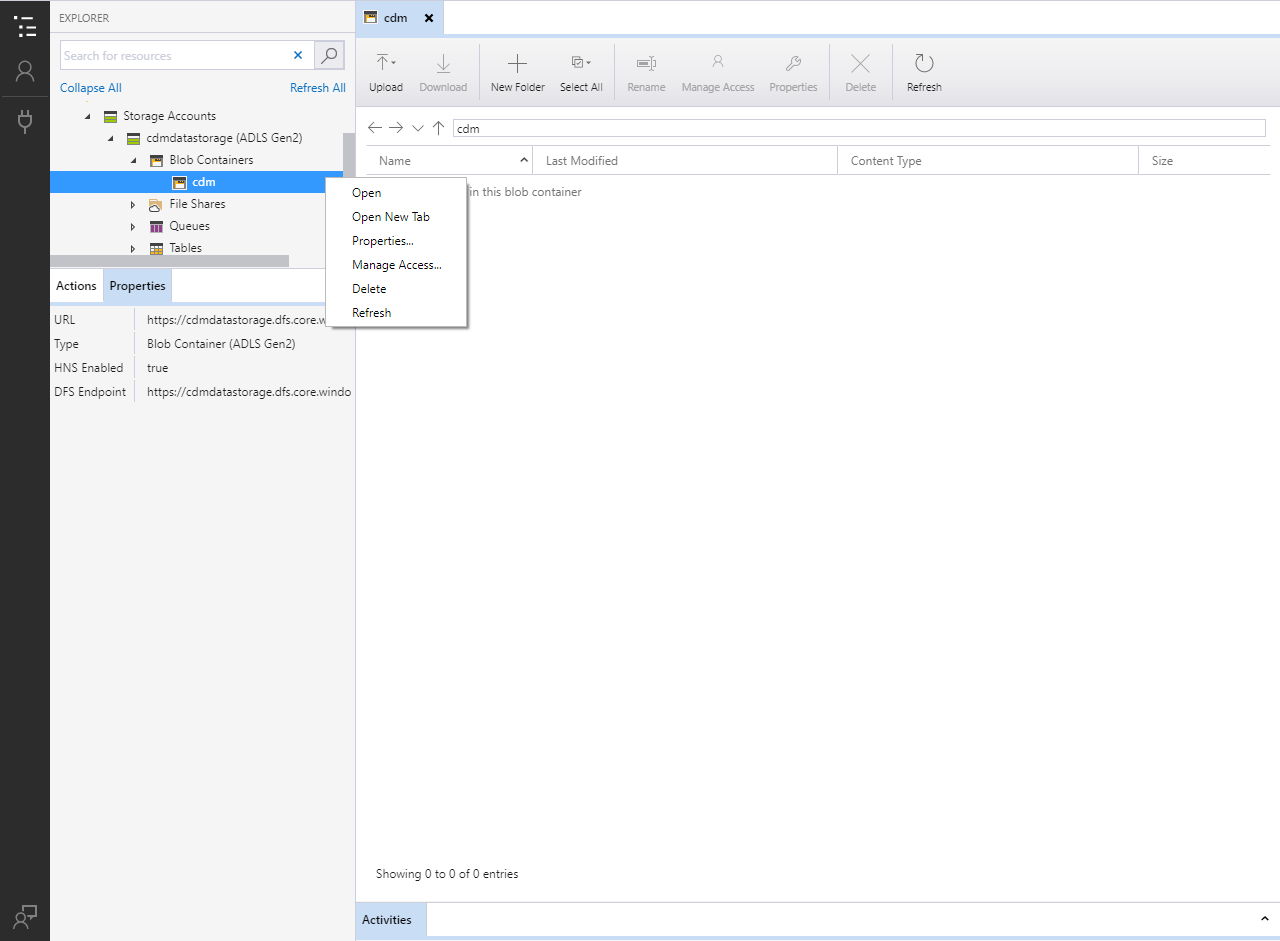


1. In the textbox that appears, enter your preferred name for the blob.
   1. The blob must be named powerbi if it will be written to by a Power BI Dataflow.
   2. If the blob is just being consumed by Power BI Dataflow its name can be any valid name.

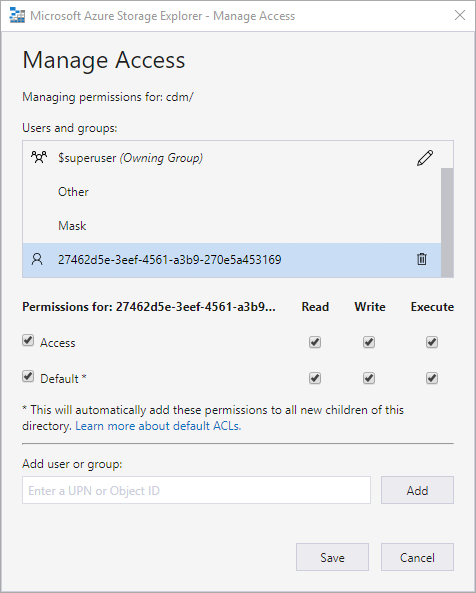


## Blob Security Setup

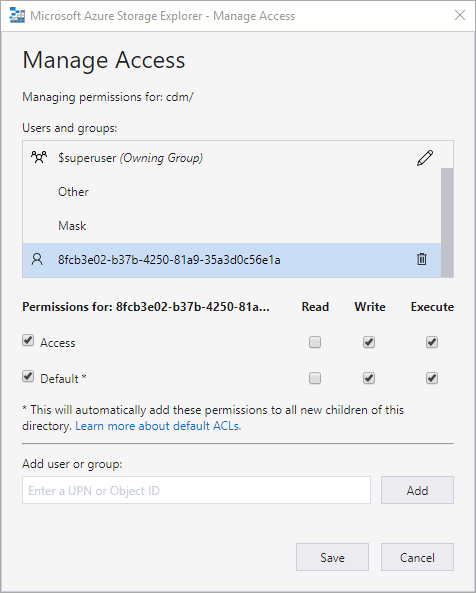
1. Right click on the newly created blob and select Manage access…



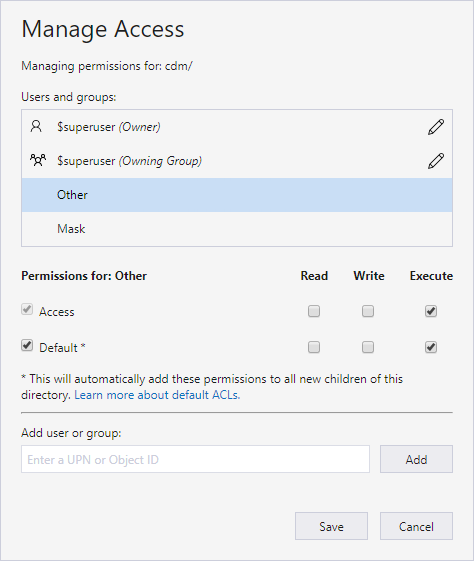
1. Using the GUIDs collected earlier, provide the following access to the following services:
   1. Power BI Premium and Power BI Service (Both require the same access)



* 1. Power Query Online

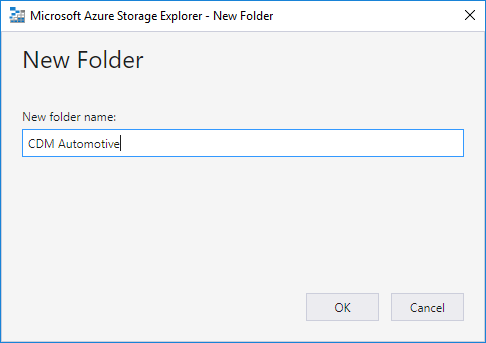


1. Grant the following access to the Other Group:

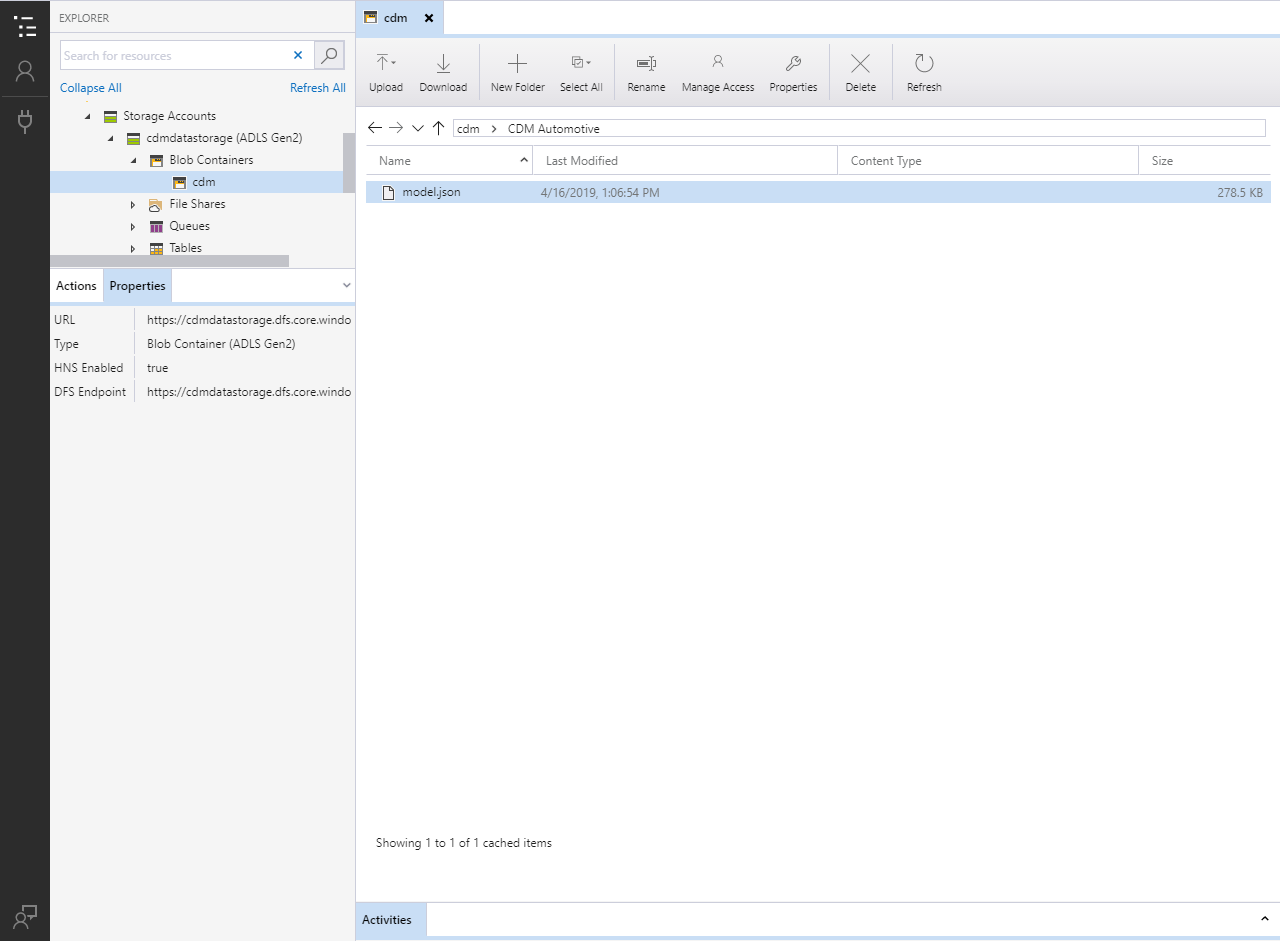


## Folder Setup

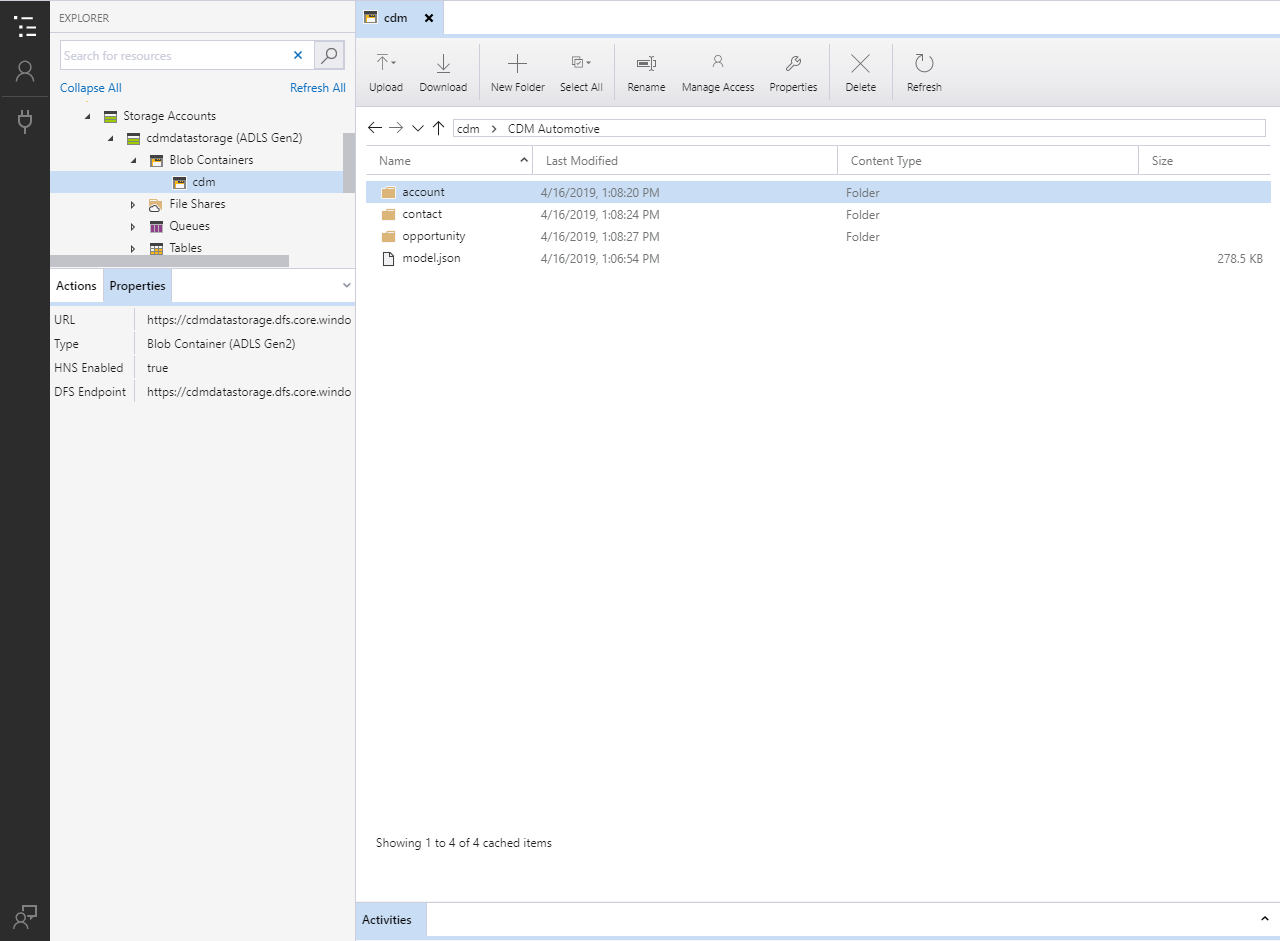
1. Inside the blob within Azure Storage Explorer press the New Folder button.
2. In the New Folder dialog, enter your preferred name for the folder. Then press OK.



1. Refresh the current view to see the newly created folder.
2. Open the newly created folder.
3. Upload the model.json file. Refresh the view if asked or the file does not appear and no errors have occurred.



1. Press the New Folder button and create folders for all entities in the model.json file. Refresh as needed to verify that the folders created successfully.

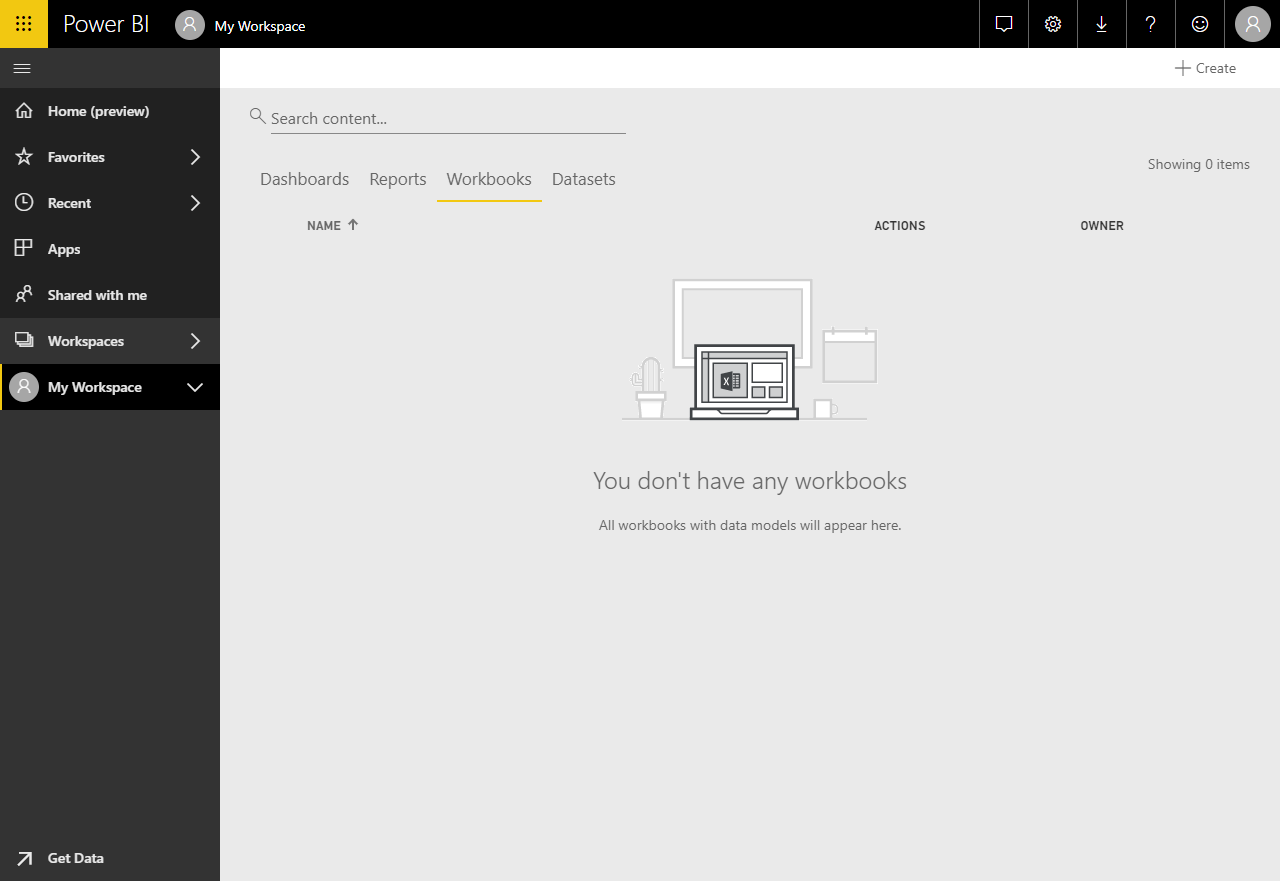


1. The basic structure for the CDM folder is complete and the model.json file can be referenced in a Power BI Dataflow. CSV files can be uploaded within each folder to add data to the Dataflow and the partition property within the model.json file can assist the Dataflow with identifying the latest data file.

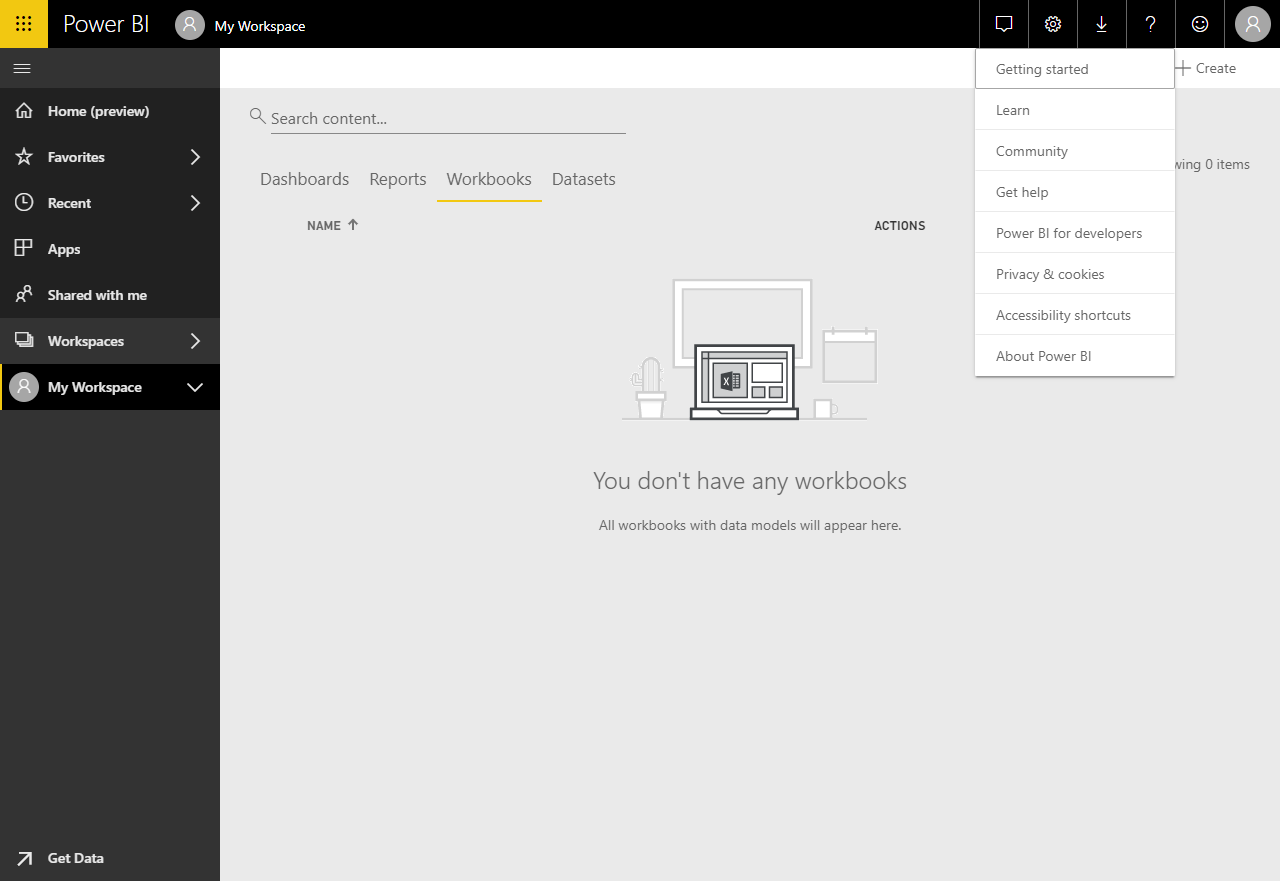
# Additional Instructions

## Identify your Power BI tenant

1. Login to powerbi.com.
2. Click the ? in the upper right-hand corner of the screen.



1. Click About Power Bi in the dropdown menu that appears.



1. The dialog that appears will display the location in which your Power BI data is stored.

