

Azure Data Factory Lab – Introduction and Setup

Overview

In this workshop, you will work on a series of lab modules that teach you how to use Azure Data Factory to import data from REST API data sources.

The first module will guide you through copying data from anonymous APIs and APIs with basic API Key authentication.

The second module will guide you through setting up an Azure Function to handle the OAuth authentication to an API and allow you to call it through ADF.

The third module will guide you through sending the data you are copying from the APIs in the first two modules to Azure Synapse.

The datasets you'll be working with are Catfacts, FCC Filings, and Google Analytics.

Lab Setup

Installations for Module 2

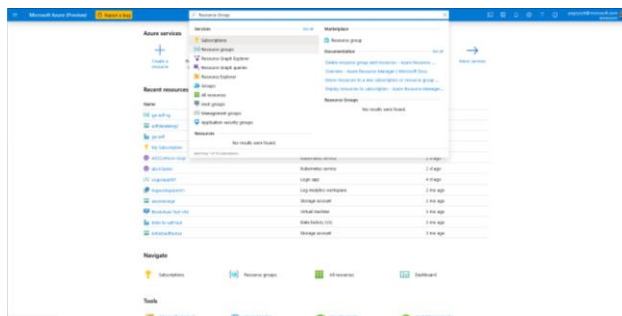
1. Install Visual Studio Code (<https://code.visualstudio.com/download>)
2. Install Node.js (<https://nodejs.org/en/download/>)
3. Install the Azure Functions Core Tools (<https://docs.microsoft.com/en-us/azure/azure-functions/functions-run-local?tabs=windows#v2>)

Log into Azure Portal

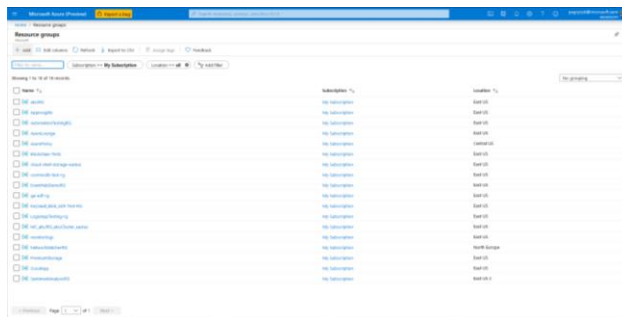
1. Log into the Azure Portal (<https://portal.azure.com>)

Create Resource Group for your Lab Resources

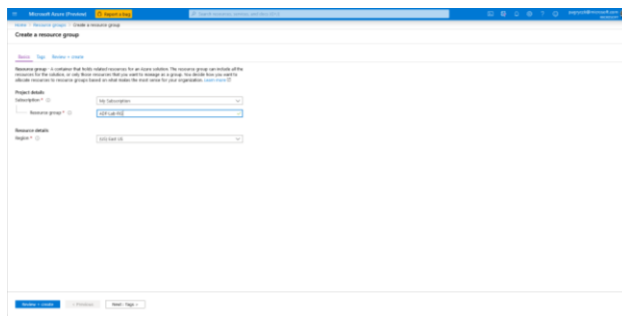
In the Azure Portal navigate to your Resource Groups view



On the Resource Groups page, click “Add”

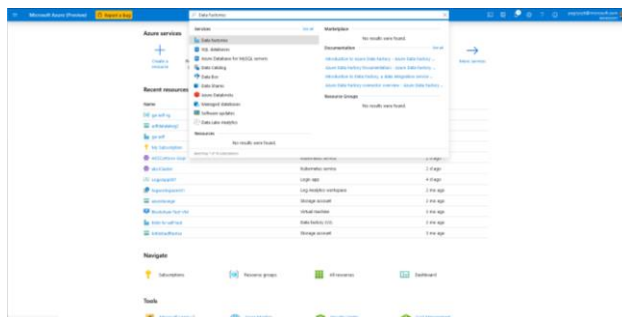


Name your resource group “ADF-Lab-RG” (if you choose a different name for your resource group, keep track of it as you’ll be referencing it as you create additional resources).

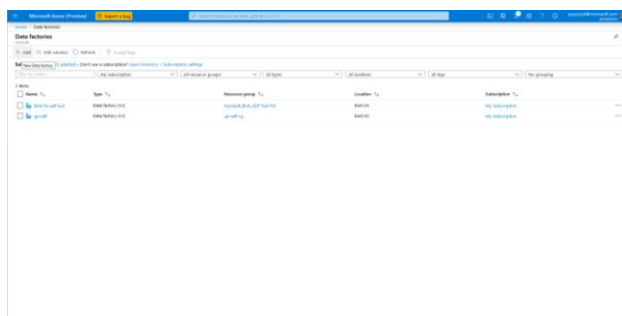


Create an Azure Data Factory

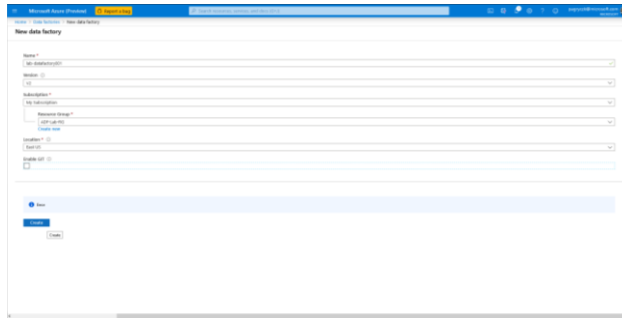
Navigate to your Data Factories page in the Azure Portal



Click “Add”



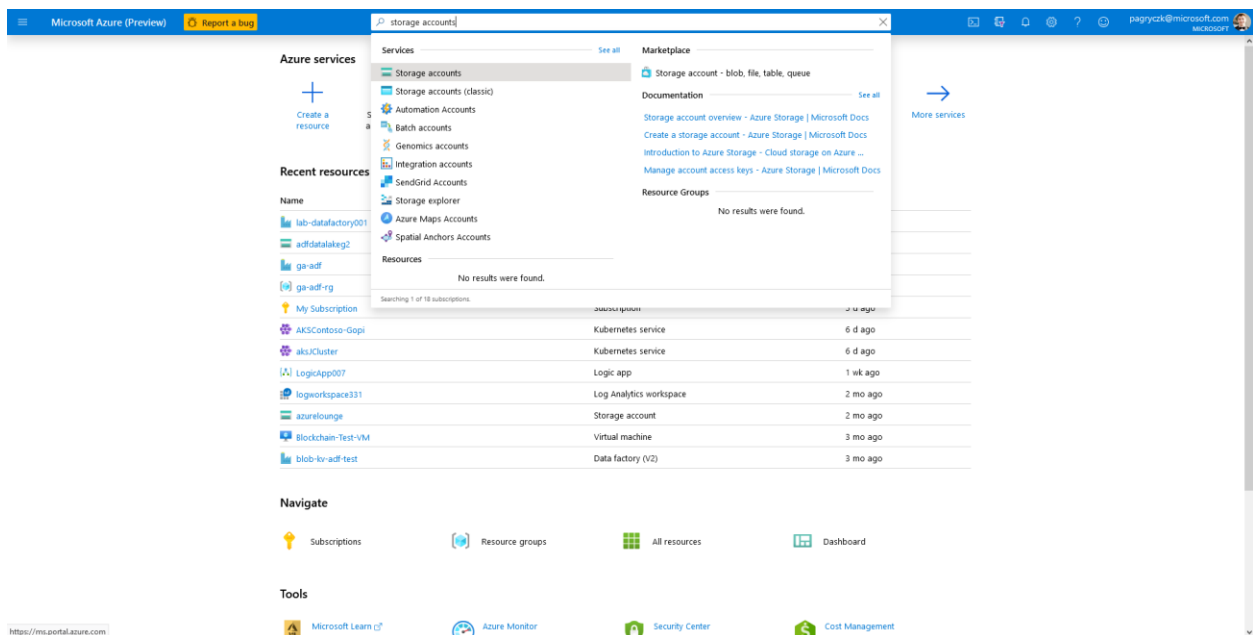
Name your Data Factory, tie it to your ADF-Lab-RG resource group, uncheck “Enable GIT,” and click “Create”. (You will need to pick a globally unique name for your Data Factory.)



The screenshot shows the 'New data factory' form in the Azure portal. The form has several fields: 'Name' (with a dropdown for 'lab-datafactory001'), 'Region' (with a dropdown for 'US'), 'Subscription' (with a dropdown for 'My Subscription'), 'Resource Group' (with a dropdown for 'ADF-Lab-RG'), 'Location' (with a dropdown for 'East US'), and 'Enable Git' (with a checkbox). At the bottom, there are 'Add' and 'Cancel' buttons.

Create a Storage Account

Navigate to the Storage Accounts page in the Azure Portal



The screenshot shows the Azure portal interface. The left sidebar has a search bar with 'storage accounts' entered. The main content area shows a list of storage accounts. The 'Storage accounts' section is expanded, showing a list of storage accounts. The 'Marketplace' section is also visible, showing a list of storage accounts. The 'Recent resources' section is also visible, showing a list of recent resources. The 'Navigate' section is at the bottom, showing links to Subscriptions, Resource groups, All resources, and Dashboard. The 'Tools' section is at the bottom, showing links to Microsoft Learn, Azure Monitor, Security Center, and Cost Management.

| Storage accounts | Marketplace |
|----------------------------|---|
| Storage accounts (classic) | Storage account - blob, file, table, queue |
| Automation Accounts | Documentation |
| Batch accounts | Storage account overview - Azure Storage Microsoft Docs |
| Genomics accounts | Create a storage account - Azure Storage Microsoft Docs |
| Integration accounts | Introduction to Azure Storage - Cloud storage on Azure ... |
| SendGrid Accounts | Manage account access keys - Azure Storage Microsoft Docs |
| Storage explorer | Resource Groups |
| Azure Maps Accounts | No results were found. |
| Spatial Anchors Accounts | |
| Resources | |
| No results were found. | |

Click 'Add' to create a new Storage Account

Home > Storage accounts

Storage accounts

[Add](#)
[Manage view](#)
[Refresh](#)
[Export to CSV](#)
[Assign tags](#)
[Delete](#)
[Feedback](#)
[Leave preview](#)

[Filter](#)
[Add filter](#)
 Subscription == My Subscription
 Resource group == all
 Location == all

Showing 1 to 12 of 12 records. No grouping

| Name | Type | Kind | Resource group | Location | Subscription |
|---|-----------------|-----------|----------------------------|----------|-----------------|
| adfdatalake2 | Storage account | StorageV2 | ga-adf-rg | East US | My Subscription |
| adfdemote | Storage account | StorageV2 | ga-adf-rg | East US | My Subscription |
| azurelounge | Storage account | StorageV2 | AzureLounge | East US | My Subscription |
| azurelounge-dev | Storage account | StorageV2 | AzureLounge | East US | My Subscription |
| azurelounge-test | Storage account | StorageV2 | AzureLounge | East US | My Subscription |
| blockchaintestdiag | Storage account | Storage | Blockchain-Tests | East US | My Subscription |
| cs247d99ed0a1fa449bea | Storage account | Storage | cloud-shell-storage-eastus | East US | My Subscription |
| kvblobadftesta | Storage account | Storage | KeyVault_Blob_ADF-Test-RG | East US | My Subscription |
| msscoutappstorage | Storage account | StorageV2 | ScoutApp | East US | My Subscription |
| premiumstorageev93 | Storage account | StorageV2 | PremiumStorage | East US | My Subscription |
| storageaccountgaad95d2 | Storage account | Storage | ga-adf-rg | East US | My Subscription |
| storageaccountgaad9a751 | Storage account | Storage | ga-adf-rg | East US | My Subscription |

[Previous](#)
 Page 1 of 1
 [Next](#)

On the Basics tab of the create view, specify your 'ADF-Lab-RG' resource group and choose a unique name for your storage account. Make note of the name as you will use it later in the lab.

Microsoft Azure (Preview) [Report a bug](#) [Search resources, services, and docs \(v2\)](#)

Home > Storage accounts > Create storage account

Create storage account

[Basics](#)
[Networking](#)
[Advanced](#)
[Tags](#)
[Review + create](#)

Azure Storage is a Microsoft-managed service providing cloud storage that is highly available, secure, durable, scalable, and redundant. Azure Storage includes Azure Blobs (objects), Azure Data Lake Storage Gen2, Azure Files, Azure Queues, and Azure Tables. The cost of your storage account depends on the usage and the options you choose below. [Learn more about Azure storage accounts](#)

Project details

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription * [My Subscription](#)
 Resource group * [ADF-Lab-RG](#) [Create new](#)

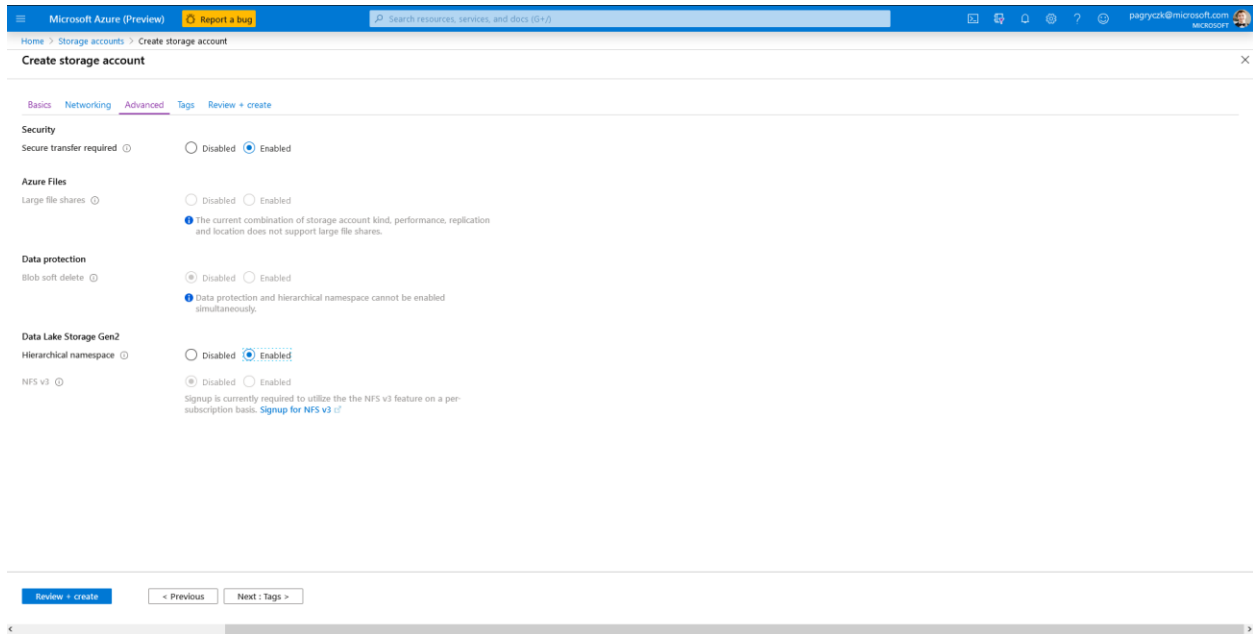
Instance details

The default deployment model is Resource Manager, which supports the latest Azure features. You may choose to deploy using the classic deployment model instead. [Choose classic deployment model](#)

Storage account name * [datalake007](#)
 Location * [\(US\) East US](#)
 Performance ☒ Standard ☐ Premium
 Account kind [StorageV2 \(general purpose v2\)](#)
 Replication [Read-access geo-redundant storage \(RA-GRS\)](#)
 Access tier (default) ☐ Cool ☒ Hot

[Review + create](#)
[Previous](#)
[Next: Networking](#)

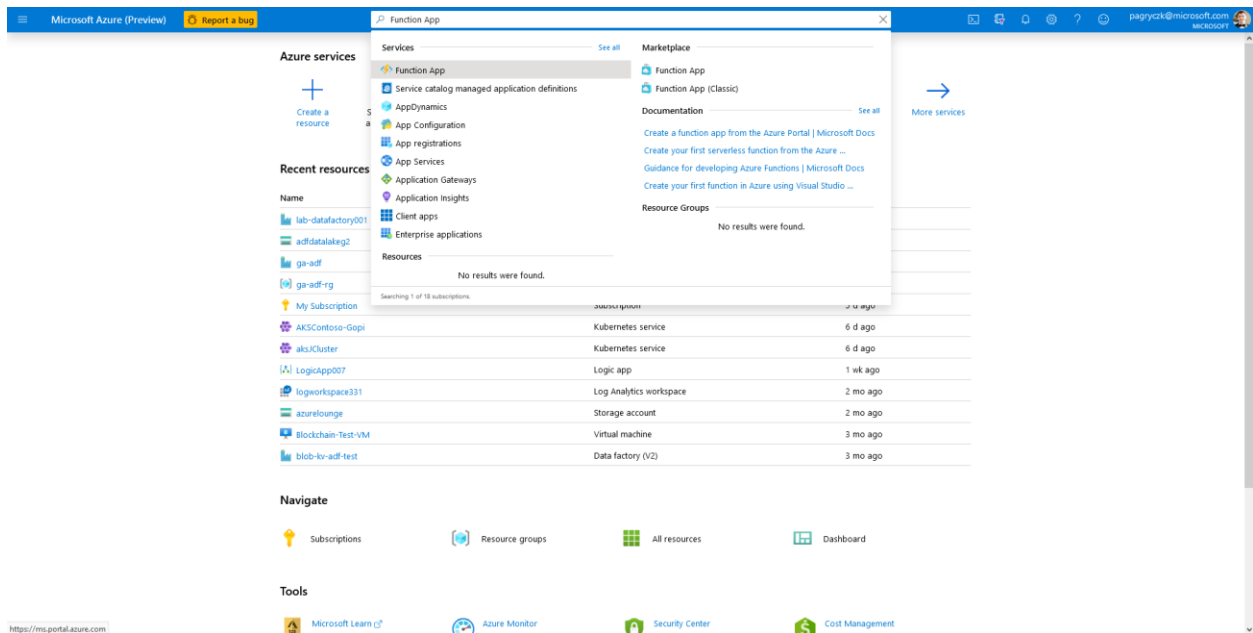
Under the 'Advanced' tab, select 'Enabled' for the option to turn on Hierarchical namespace.



Then click Review and Create and proceed to create your storage account.

Create an Azure Function

Navigate to the Function App page in the Azure Portal.



Click 'Add' to create a new Function App.

| Name | Status | Location | Pricing Tier | App Service Plan | Subscription | App Type |
|--------------------|---------|------------|--------------|-----------------------------|-----------------|--------------|
| funcapp-ga | Running | East US | Dynamic | ASP-gaadfrg-b93b | My Subscription | Function App |
| ga-adf-functionapp | Running | East US | Dynamic | ASP-gaadfrg-9867 | My Subscription | Function App |
| MSScoutApp | Running | Central US | Dynamic | MSScoutAppConsumptionHostin | My Subscription | Function App |

Select your 'ADF-Lab-RG' as the resource group for the Function App, give the function app a unique name and make note of it, and select Python as the Runtime stack. Then proceed to 'Review + Create' and 'Create' the Function App.

Function App

Basics | Hosting | Monitoring | Tags | Review + create

Create a function app, which lets you group functions as a logical unit for easier management, deployment and sharing of resources. Functions lets you execute your code in a serverless environment without having to first create a VM or publish a web application.

Project Details

Select a subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription *

Resource Group * [Create new](#)

Instance Details

Function App name * .azurewebsites.net

Publish * ☐ Code ☐ Docker Container

Runtime stack *

Region *

[Review + create](#) [Previous](#) [Next: Hosting](#)

Lab Overview

ADF Lab: Module 1 – Copying Data from Rest APIs

- Create a Data Factory Pipeline
- Create a Dataset
- Add an Activity to your Pipeline

ADF Lab: Module 2 – Using Azure Functions to Copy Data from Rest API

- Create an Azure Function
- Update and upload Code to Azure Function
- Call Azure Function from ADF

ADF Lab: Module 3 – Sending data to Azure Synapse

- Transform JSON to Delimited Text Format
- Copy Data to Azure Synapse