

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 without authentication and anonymous APIs with 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 datasets you'll be working with are Catfacts, FCC Filings, and Google Analytics.

Lab Setup

Installations for Module 2

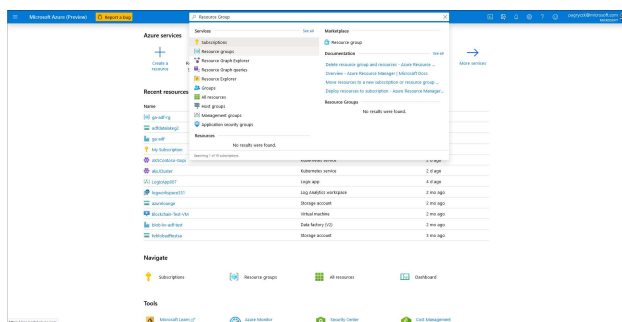
1. Install Visual Studio Code (<https://code.visualstudio.com/download>)
2. Install Node.js (<https://nodejs.org/en/download/>)
3. Install the Azure Functions Core Tools (<https://docs.microsoft.com/en-us/azure/azure-functions/functions-run-local?tabs=windows#v2>)

Log into Azure Portal

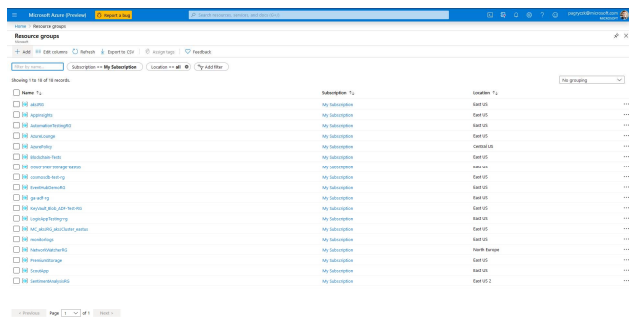
1. Log into the Azure Portal (<https://portal.azure.com>)

Create Resource Group for your Lab Resources

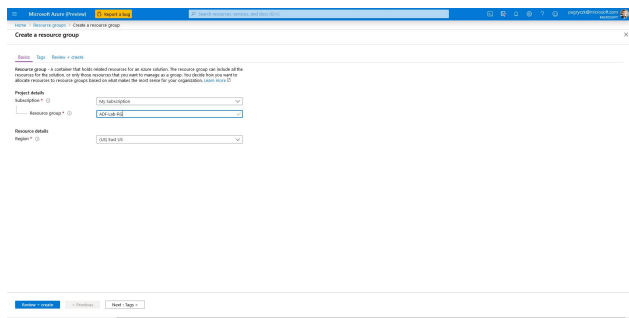
In the Azure Portal navigate to your Resource Groups view



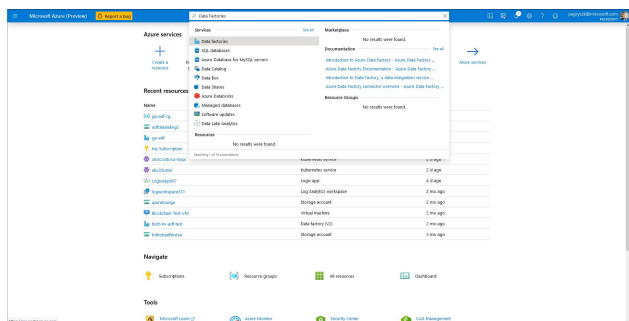
On the Resource Groups page, click “Add”



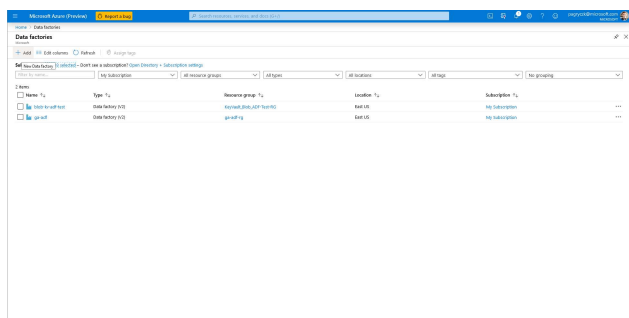
Name your resource group “ADF-Lab-RG” (if you choose a different name for your resource group, keep track of it as you’ll be referencing it as you create additional resources).



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 the following fields and values:

- Name: lab-datafactory001
- Region: US
- Subscription: No subscription
- Resource group: ADF-Lab-RG
- Location: East US
- Enable Git: ☐

At the bottom of the form, there is a 'Create' button.

Create a Storage Account

Navigate to the Storage Accounts page in the Azure Portal

The screenshot shows the Azure Portal interface with the 'Storage accounts' page selected. The left sidebar shows the 'Create a resource' button. The main content area shows a list of storage accounts with columns for Name, Subscription, Service, and Last updated. The 'Add' button is visible at the bottom of the list.

Name	Subscription	Service	Last updated
lab-datafactory001	AKSContoso-Gopi	Kubernetes service	6 d ago
adfdatalake2	AKSContoso-Gopi	Kubernetes service	6 d ago
ga-adf	AKSContoso-Gopi	Logic app	1 wk ago
ga-adf-rg	AKSContoso-Gopi	Log Analytics workspace	2 mo ago
		Storage account	2 mo ago
		Virtual machine	3 mo ago
		Data factory (V2)	3 mo ago

Click 'Add' to create a new Storage Account

Home > Storage accounts

Storage accounts

Microsoft

+ Add Manage view Refresh Export to CSV Assign tags Delete Feedback Leave preview

Filter: Add Subscription == My Subscription Resource group == all Location == all Add filter

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
cs247d99e0a1fa449bea	Storage account	Storage	cloud-shell-storage-eastus	East US	My Subscription
kvblobadtesta	Storage account	Storage	KeyVault_Blob_ADF-Test-RG	East US	My Subscription
mscoutappstorage	Storage account	StorageV2	ScoutApp	East US	My Subscription
premiumstorage-test93	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 (ctrl)

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.

Microsoft Azure (Preview) [Report a bug](#) Search resources, services, and docs (Ctrl+K)

Home > Storage accounts > Create storage account

Create storage account

Basics Networking **Advanced** Tags Review + create

Security

Secure transfer required ☐ Disabled ☒ Enabled

Azure Files

Large file shares ☐ Disabled ☐ Enabled

The current combination of storage account kind, performance, replication and location does not support large file shares.

Data protection

Blob soft delete ☒ Disabled ☐ Enabled

Data protection and hierarchical namespace cannot be enabled simultaneously.

Data Lake Storage Gen2

Hierarchical namespace ☐ Disabled ☒ Enabled

NFS v3

☒ Disabled ☐ Enabled

Signup is currently required to utilize the the NFS v3 feature on a per-subscription basis. [Signup for NFS v3](#)

[Review + create](#) [< Previous](#) [Next : Tags >](#)

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.

Microsoft Azure (Preview) [Report a bug](#) Function App

Azure services

[Create a resource](#)

Recent resources

Name	Created
lab-datafactory001	6 d ago
adfdatabaleg2	6 d ago
ga-adf	1 wk ago
ga-adf-rig	2 mo ago
My Subscription	2 mo ago
AKSCentoso-Gopi	2 mo ago
aksJCluster	2 mo ago
LogicApp007	3 mo ago
logworkspace331	3 mo ago
azurelounge	3 mo ago
Blockchain-Test-VM	3 mo ago
blob-for-adf-test	3 mo ago

Services

[Function App](#)

[Service catalog managed application definitions](#)

[App Dynamics](#)

[App Configuration](#)

[App registrations](#)

[App Services](#)

[Application Gateways](#)

[Application Insights](#)

[Client apps](#)

[Enterprise applications](#)

[Resources](#)

[No results were found.](#)

Marketplace

[Function App](#)

[Function App \(Classic\)](#)

[Documentation](#)

[Create a function app from the Azure Portal | Microsoft Docs](#)

[Create your first serverless function from the Azure ...](#)

[Guidance for developing Azure Functions | Microsoft Docs](#)

[Create your first function in Azure using Visual Studio ...](#)

Resource Groups

[No results were found.](#)

Navigate

[Subscriptions](#) [Resource groups](#) [All resources](#) [Dashboard](#)

Tools

[Microsoft Learn](#) [Azure Monitor](#) [Security Center](#) [Cost Management](#)

<https://ms.portal.azure.com>

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-b53b	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.

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 1 – Copying Data from Rest APIs

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

ADF Lab 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