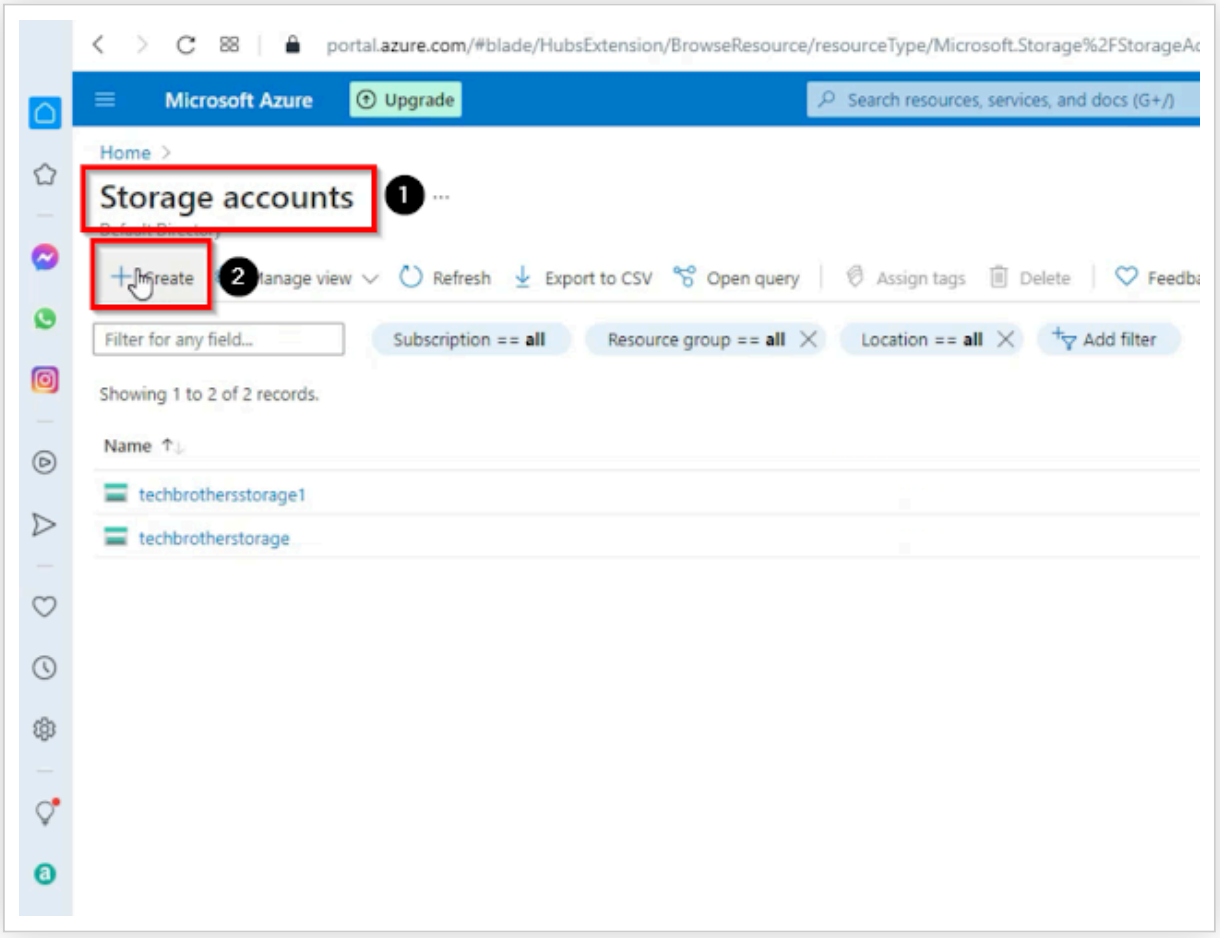
How to Copy files from one folder to another Folder and Delete in Source Folder in Azure Data Factory

Issue: How to Copy files from one folder to another Folder and Delete in Source Folder in Azure Data Factory.

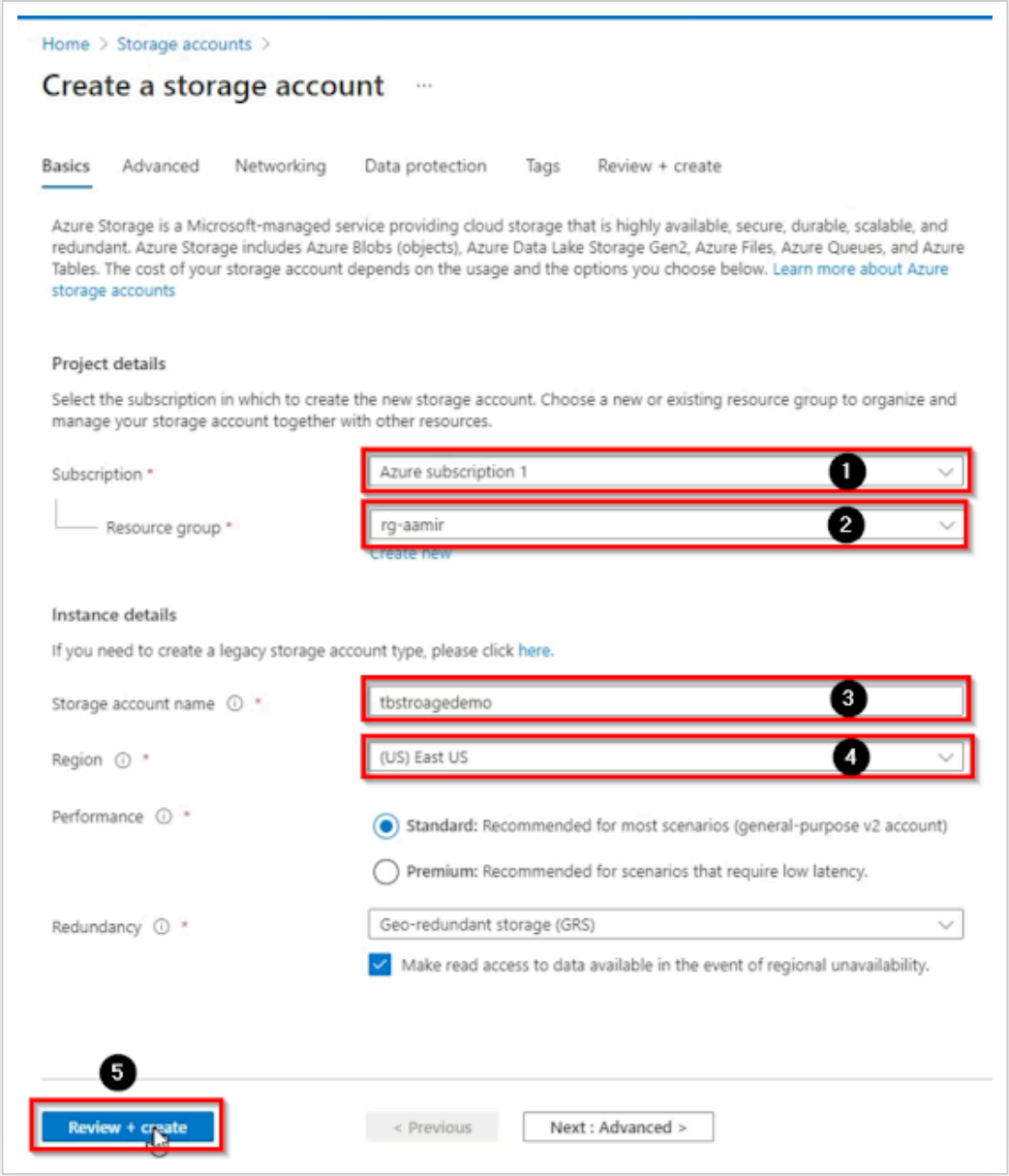
In this article, we are going to learn how to copy files from one folder to another folder and delete them from the source folder in Azure Data Factory, First of all, we will create a storage account and then will learn how to create a folder and copy files from one folder to another folder and delete from the source folder.

How to Create a Storage Account.

Open Azure data factory portal and then click on the Storage Accounts on the dashboard, then click on + create button to create a new Storage account.



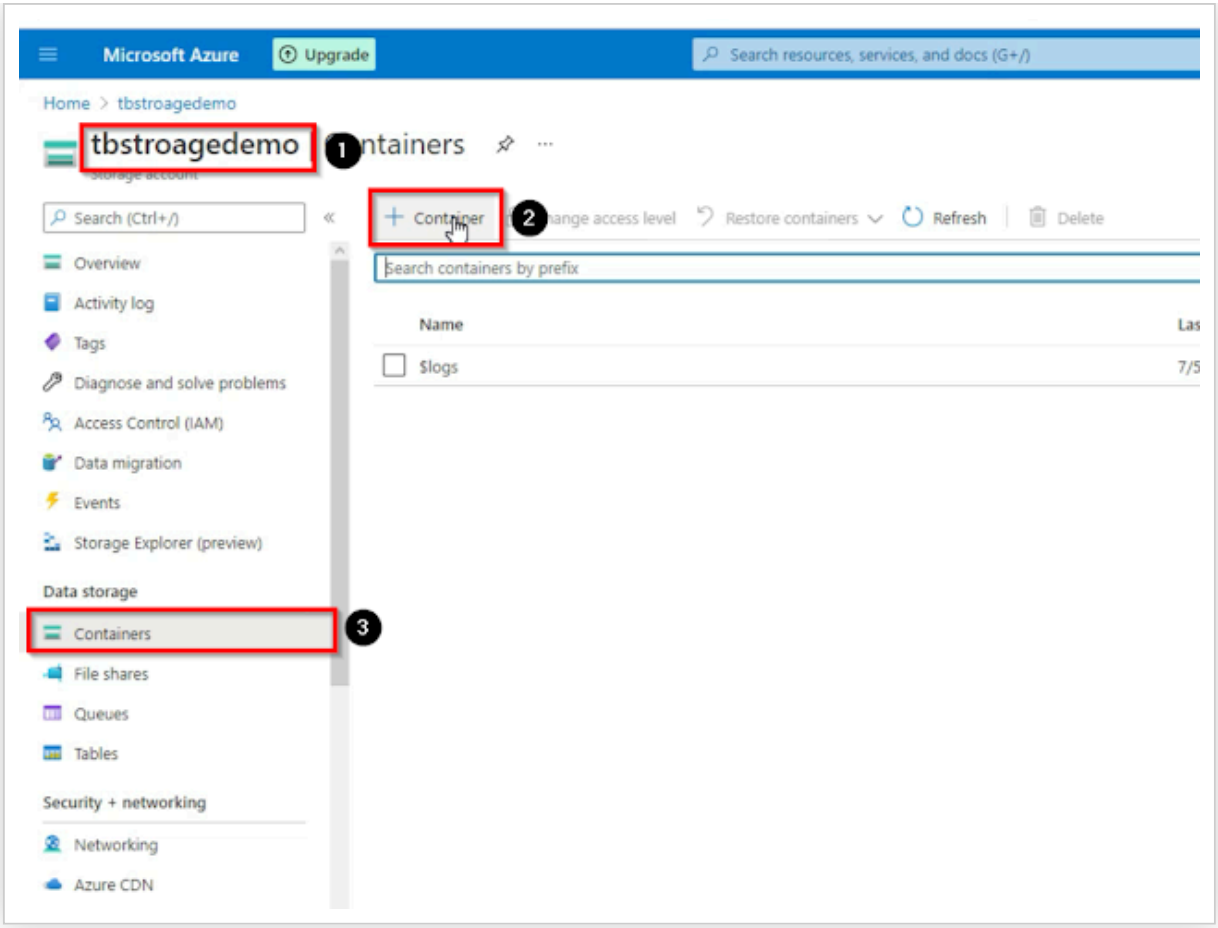
Select Azure subscription, then select Resource group, then provide the storage account name, then select your region, and then click on Review + Create then hit Create.



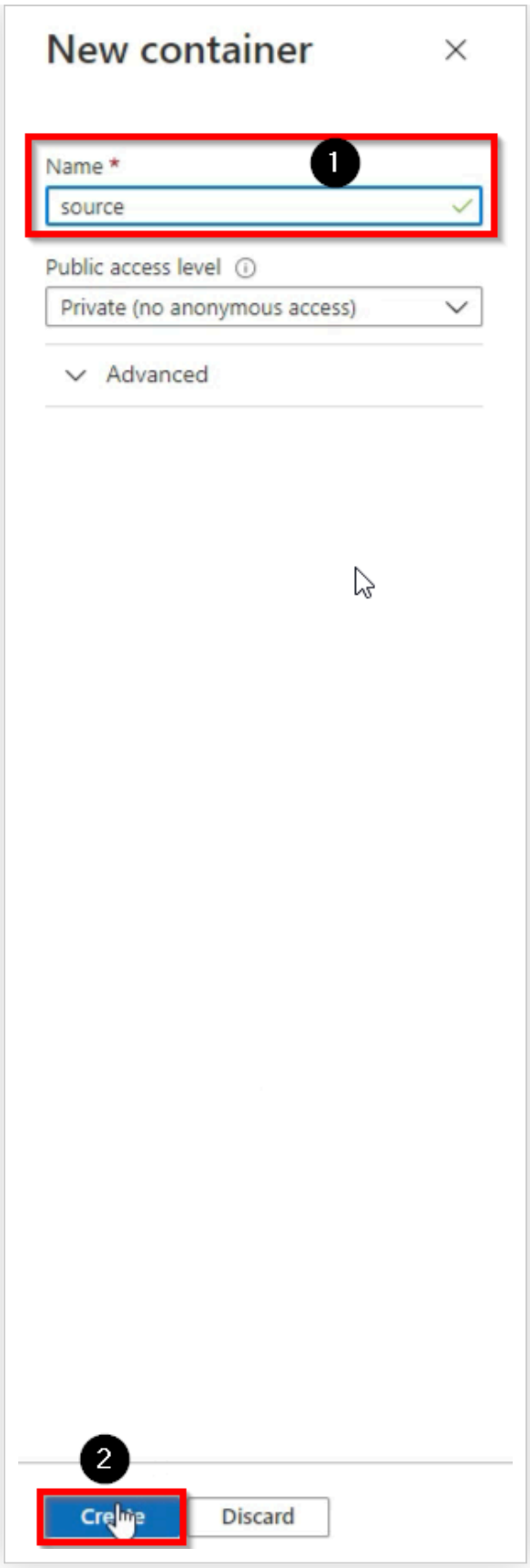
Once our storage is created, now we have to create containers or folders inside our newly created storage.

How to Create Container / Folder in Azure Data Factory.

Click on the Containers under the Data storage tab and then click on the + Container button to create a new Container.



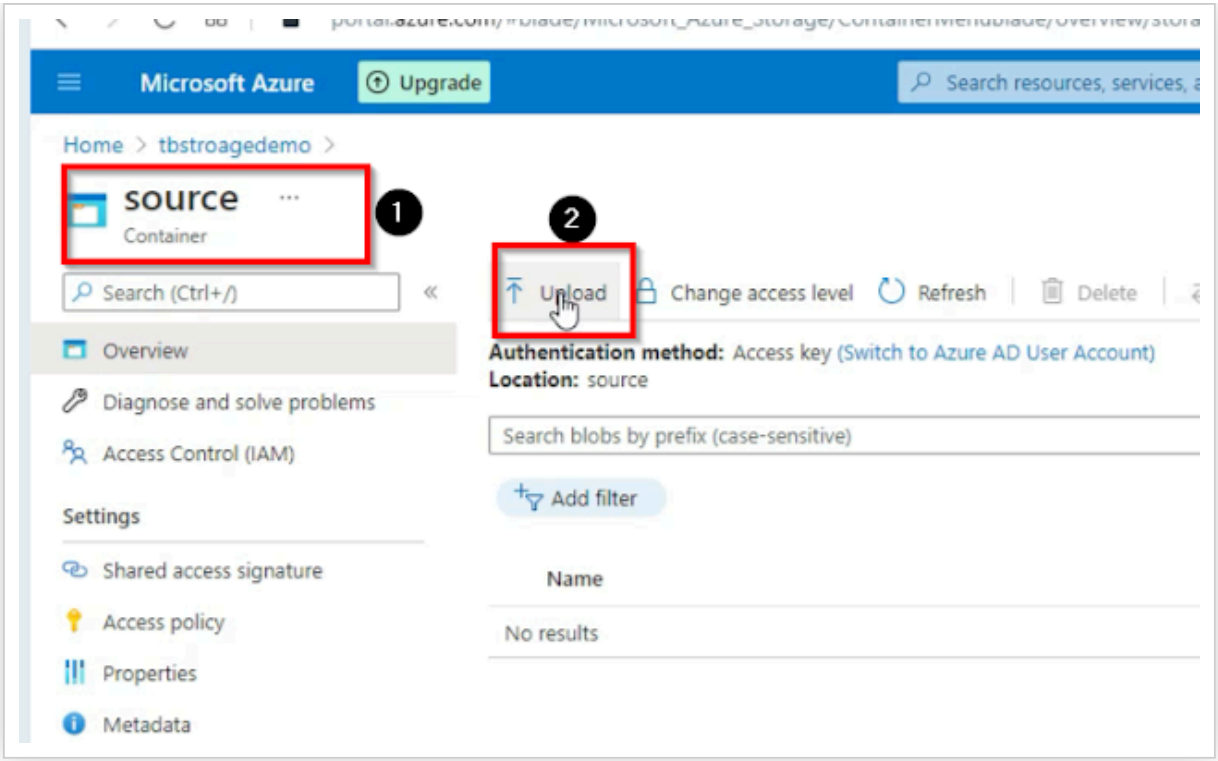
Name the container and then click on Create button, it will be created and appear in the storage, then again click on the + Container button and create another container, you can create tons of containers as per your requirement.



Once our required containers are created, let's upload the files from our local machine storage.

How to upload files from local storage to the containers in Azure Data Factory.

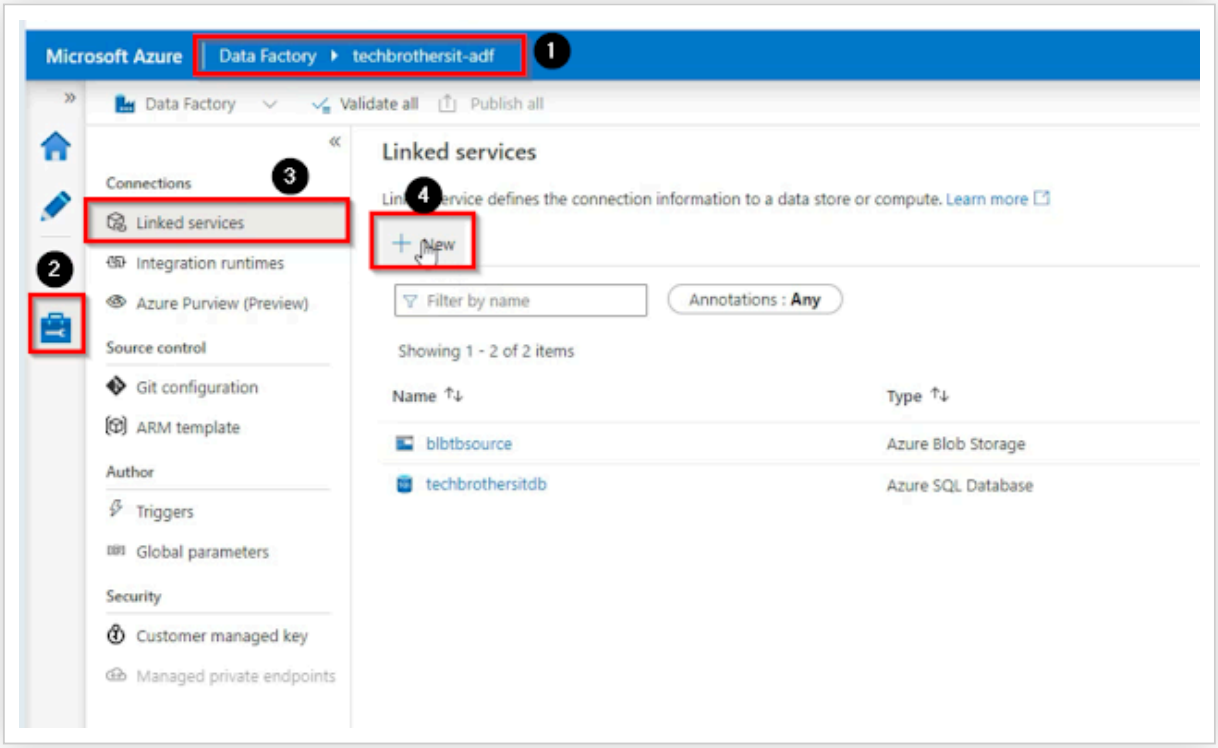
Open the folder in which you want to upload the files, then click on the upload button on the top and then navigate to the files, then select the files and click on upload you can select multiple files at one time.



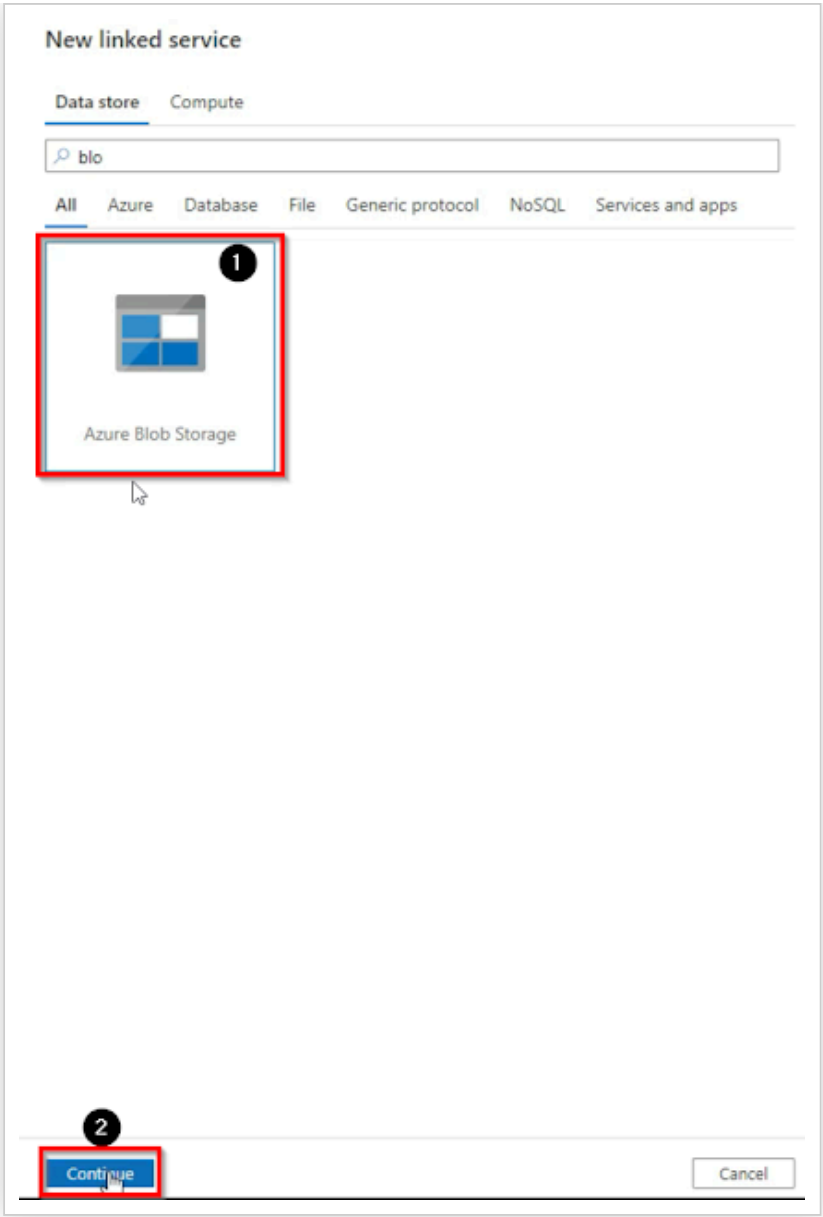
Once our files are uploaded now our next step is to create a linked service.

How to Create a Linked Service in Azure Data Factory.

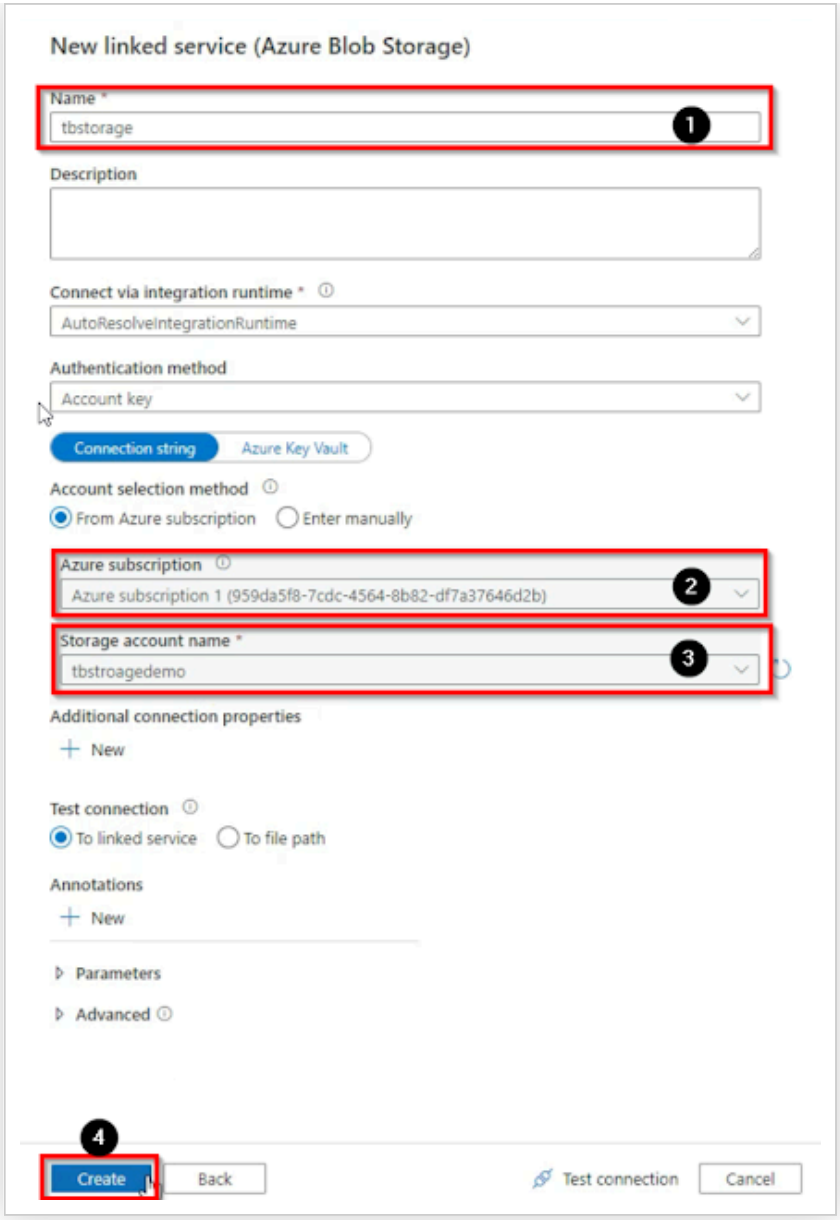
Open the Azure Data Factory studio, then go to the Manage tab and then click on linked services then click on the + New button to create a new linked service.



Select the Azure Blob storage, then click on continue.



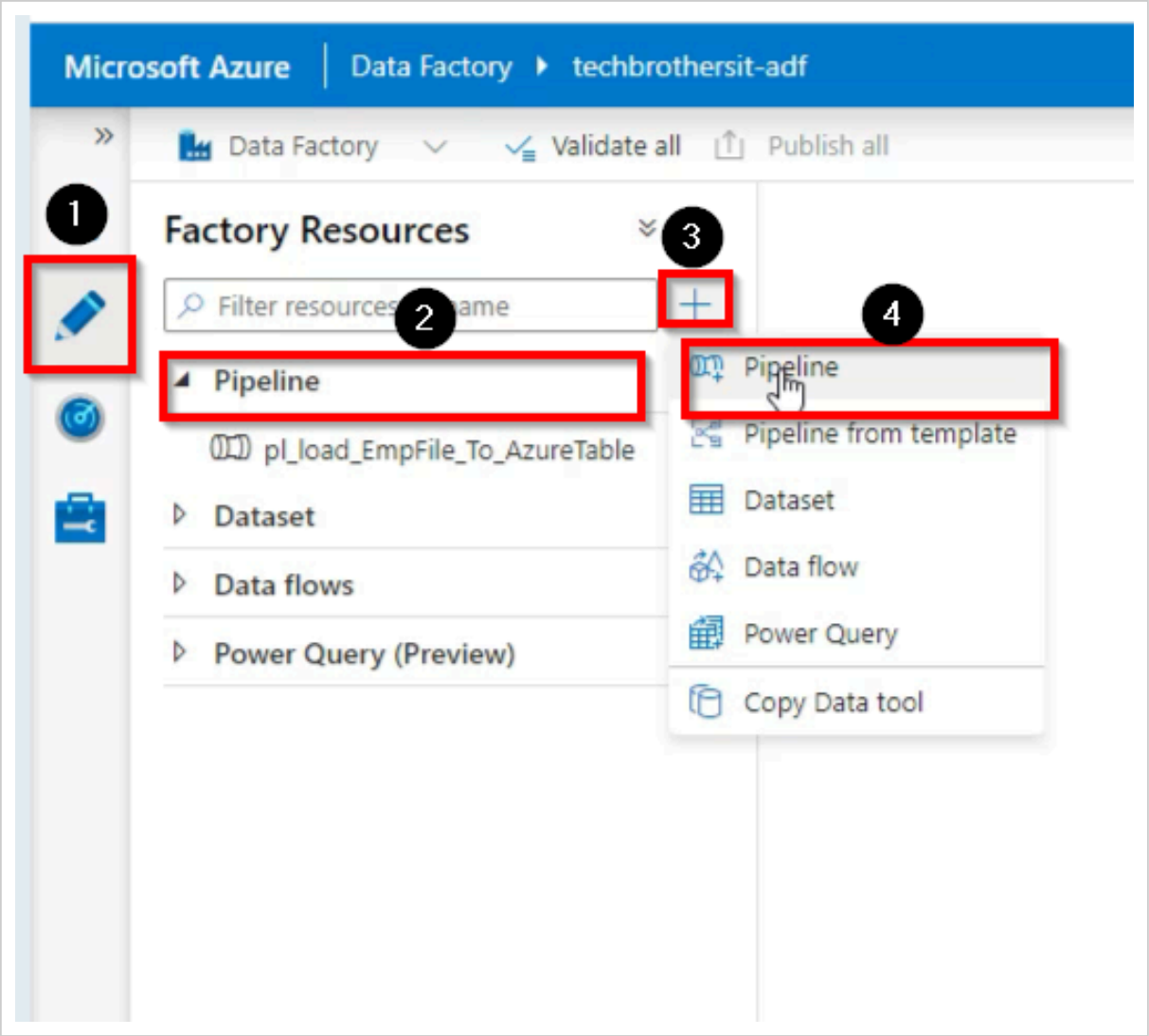
Name the linked service, then select Azure subscription, then select a storage account name, and then click on Create.



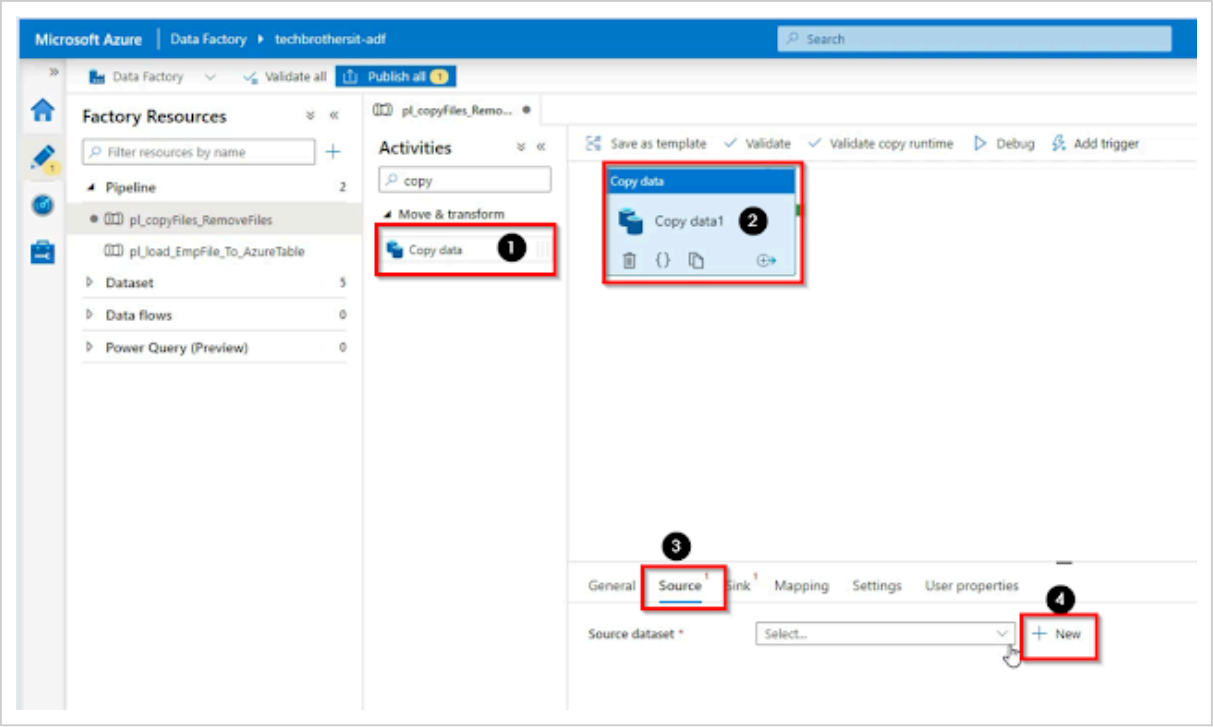
Once our linked service is created, the next step is to create a pipeline.

How to create a Pipeline in Azure Data Factory.

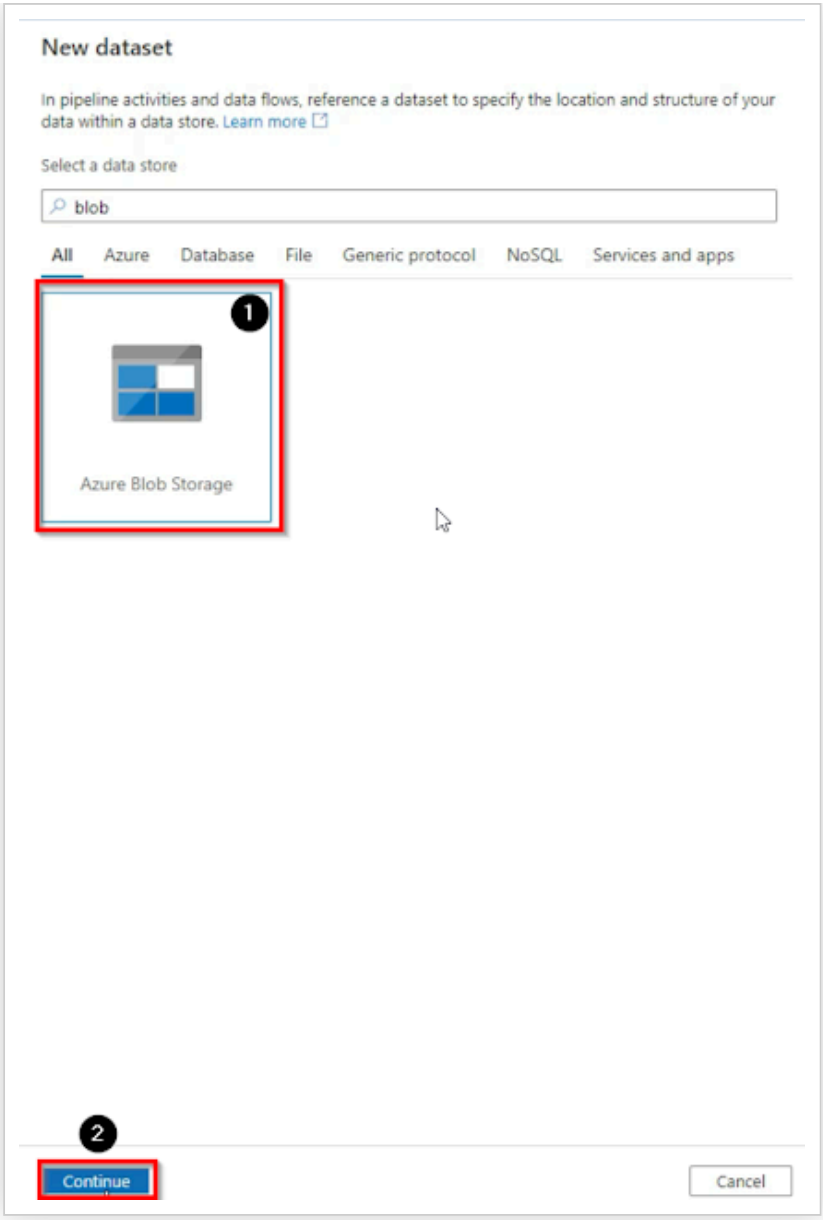
Go to the author tab and then click on pipelines then click on the + button, and then click on the new pipeline to create a new pipeline.



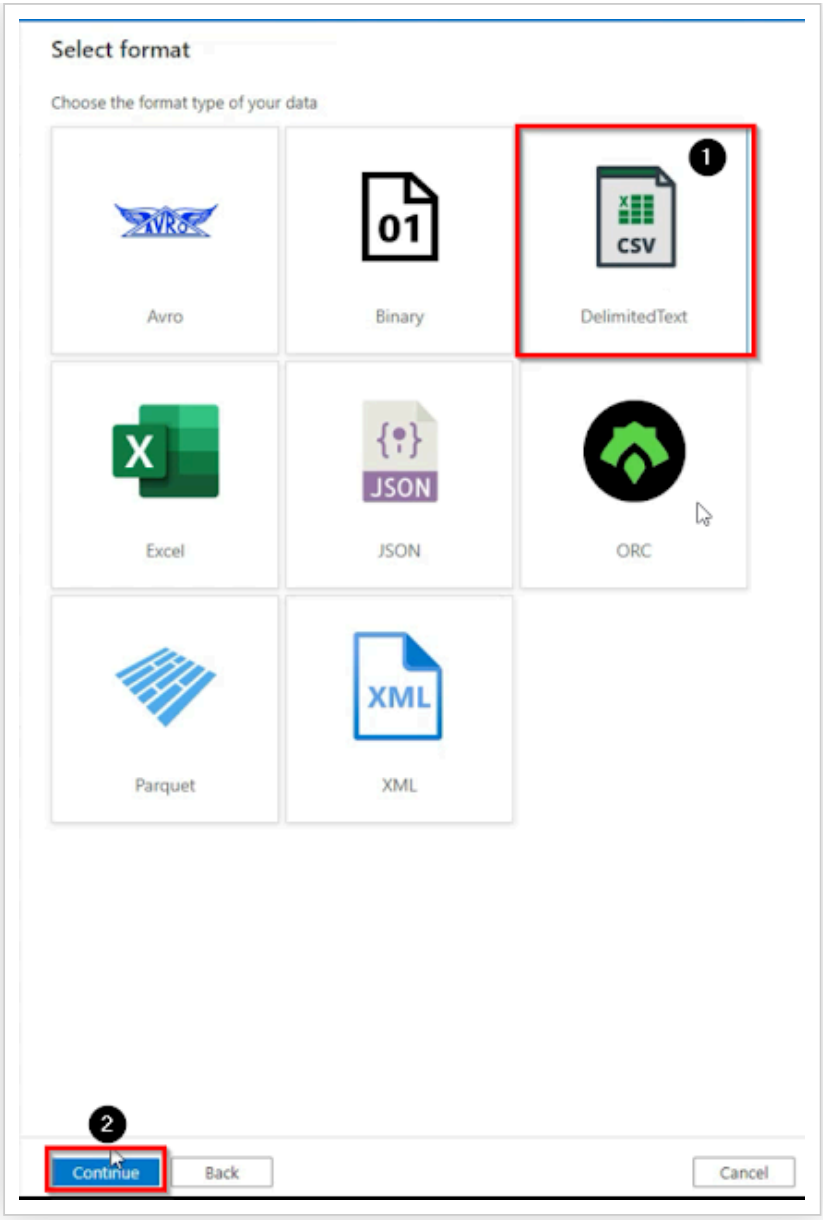
Name the pipeline and then go to the activities tab and find and drag the copy data activity to the working window, and click on the source tab then click on the + New button to create a new source dataset.



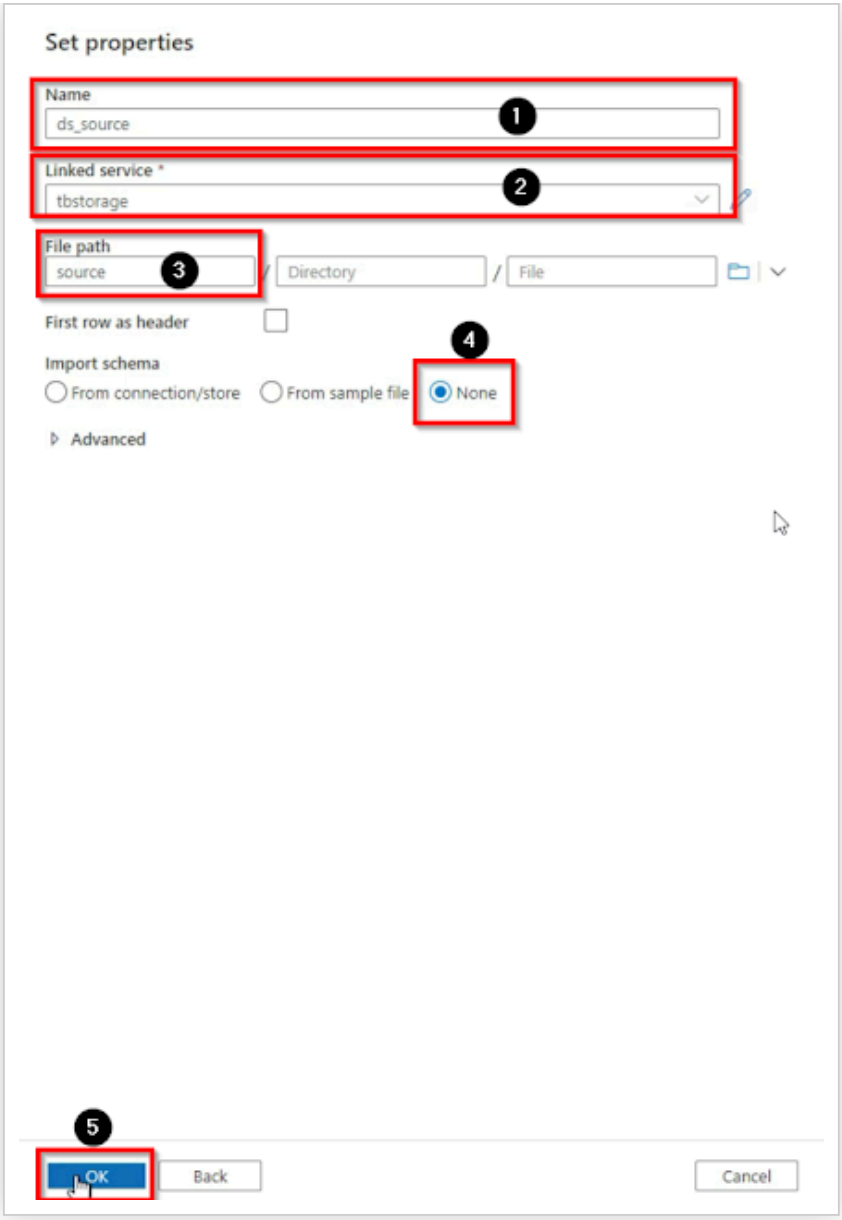
Select the Azure Blob storage and click on continue.



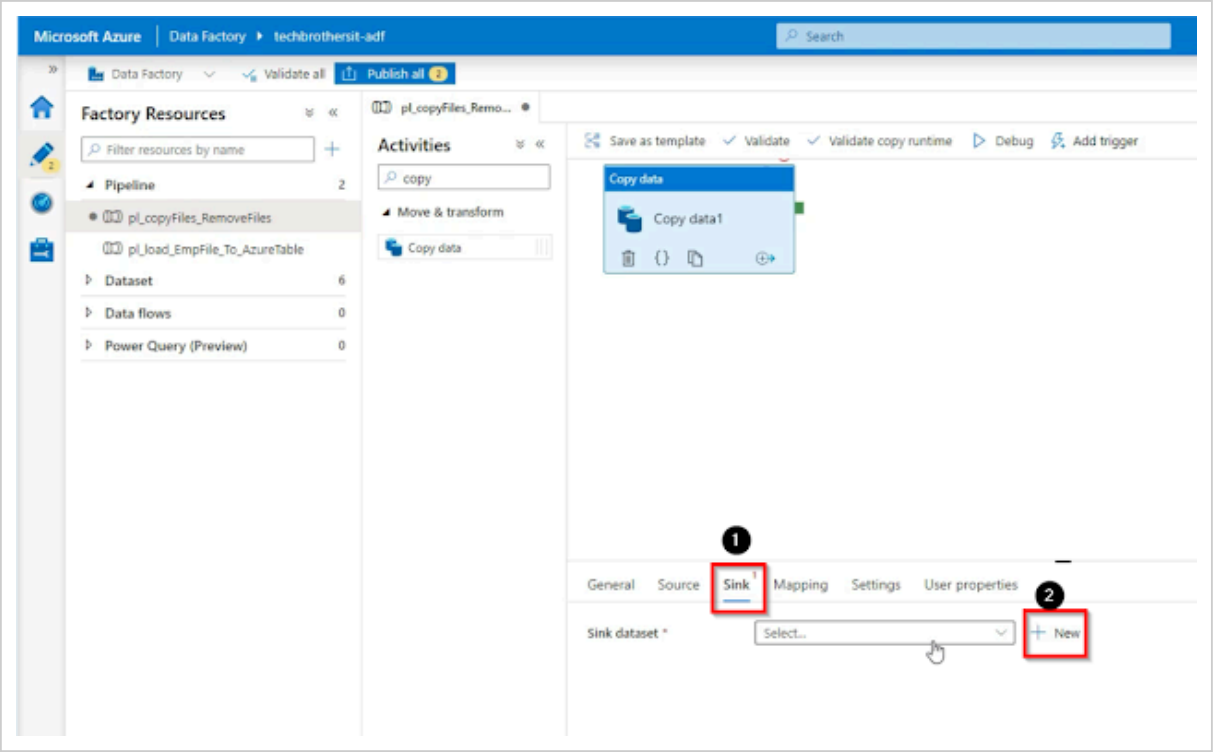
Select the source file format, in my case it is CSV file, then click on continue.



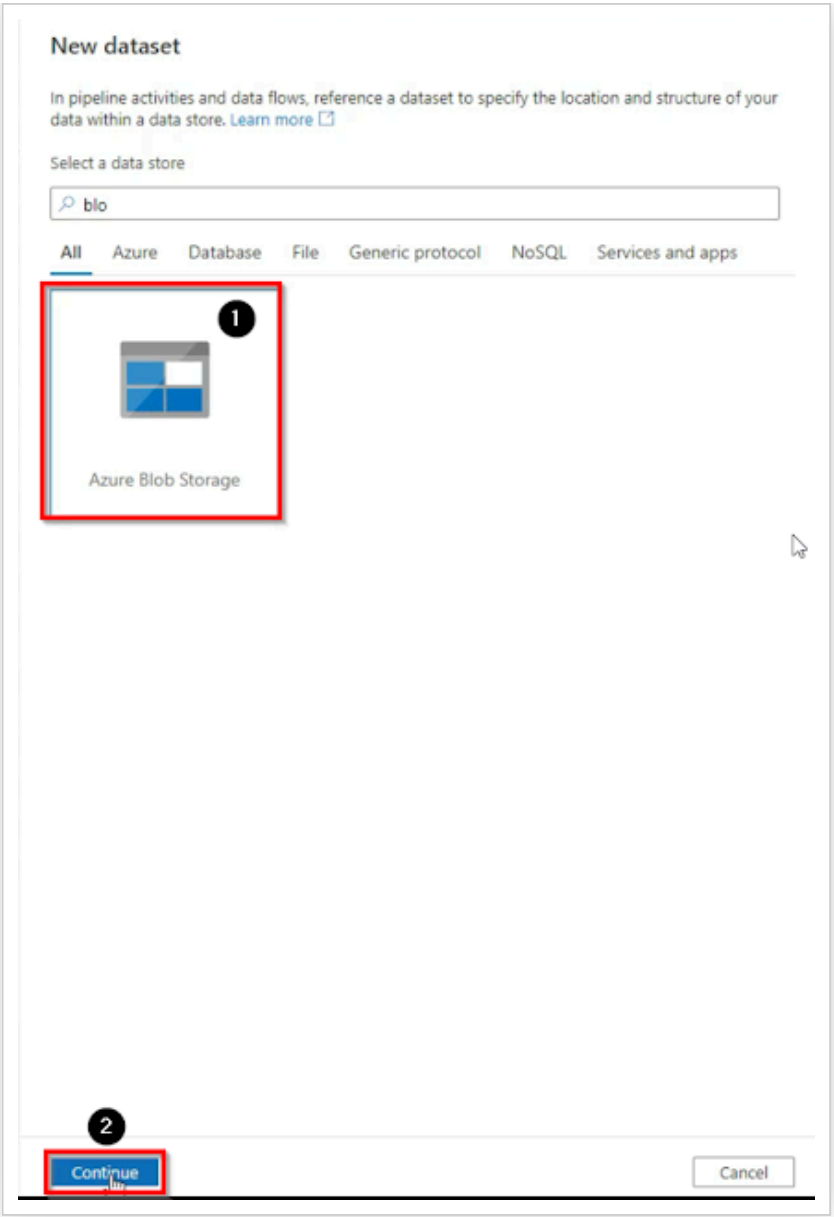
Then give a name to the dataset, then select linked service that we have created before, then select source folder path, then select none for import schema, and then click on ok.



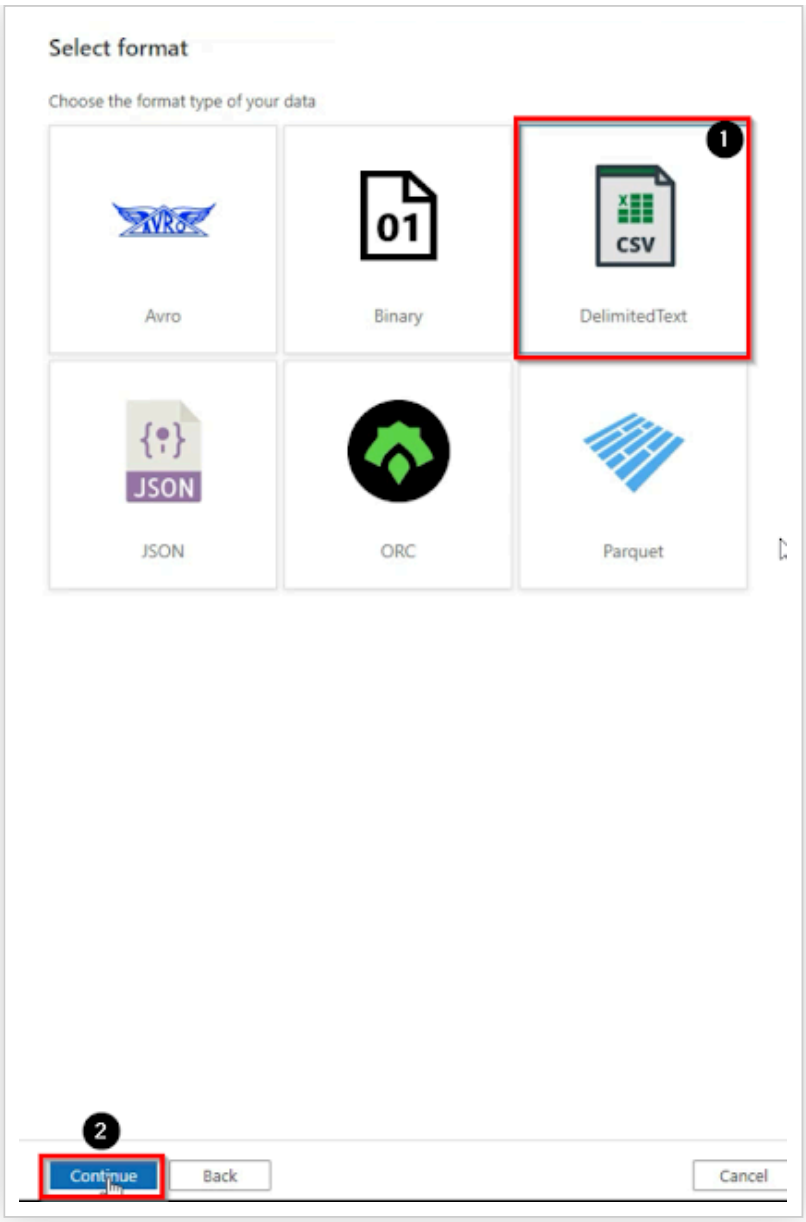
As our source dataset is completed now we have to create a sink dataset, to create a sink dataset, go to the sink tan and then click on the + New button.



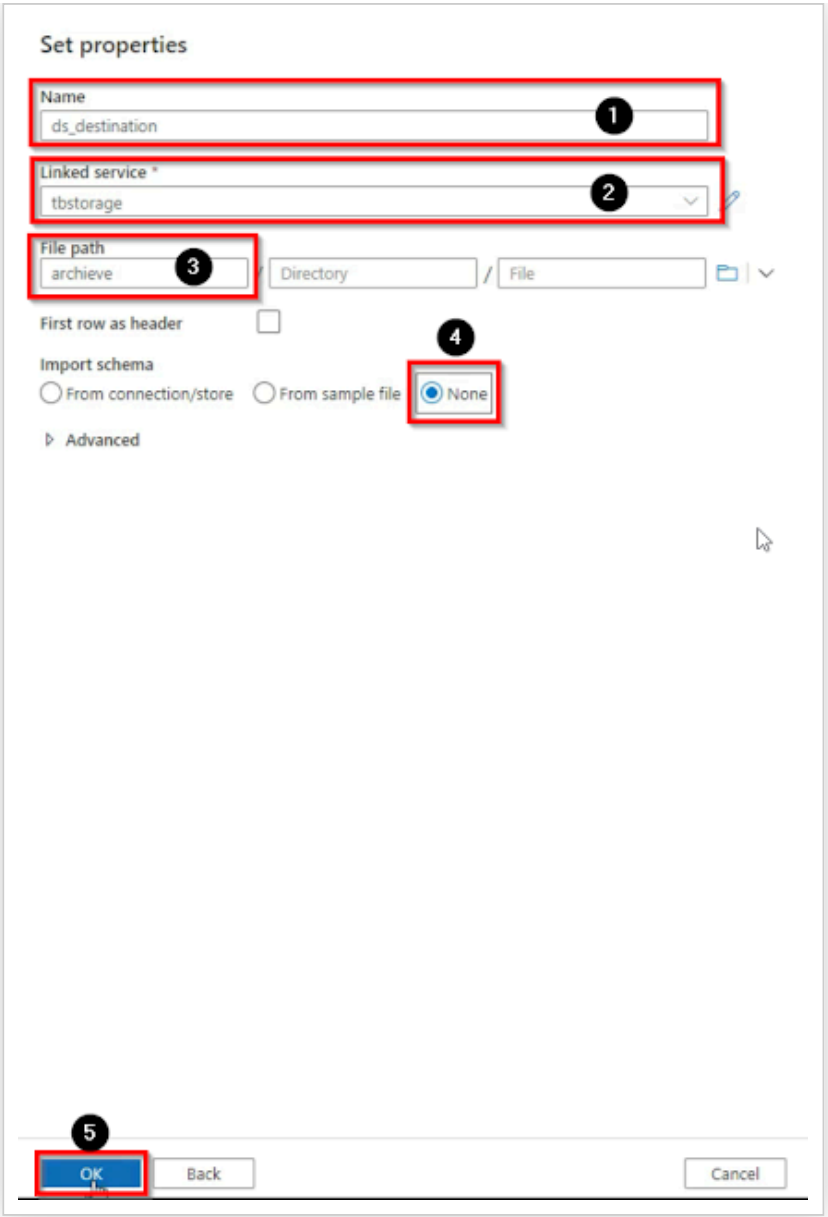
Then select Azure Blob storage and click on continue.



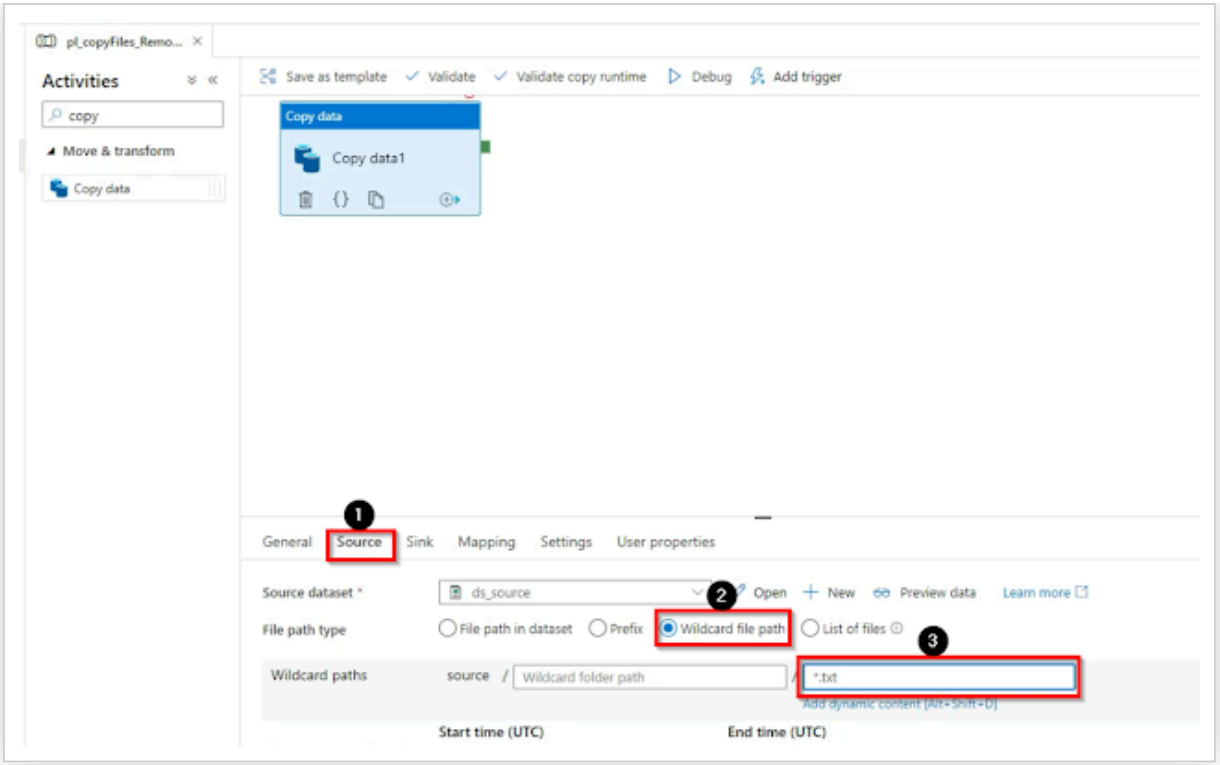
Then select the file format CSV and then click on Continue.



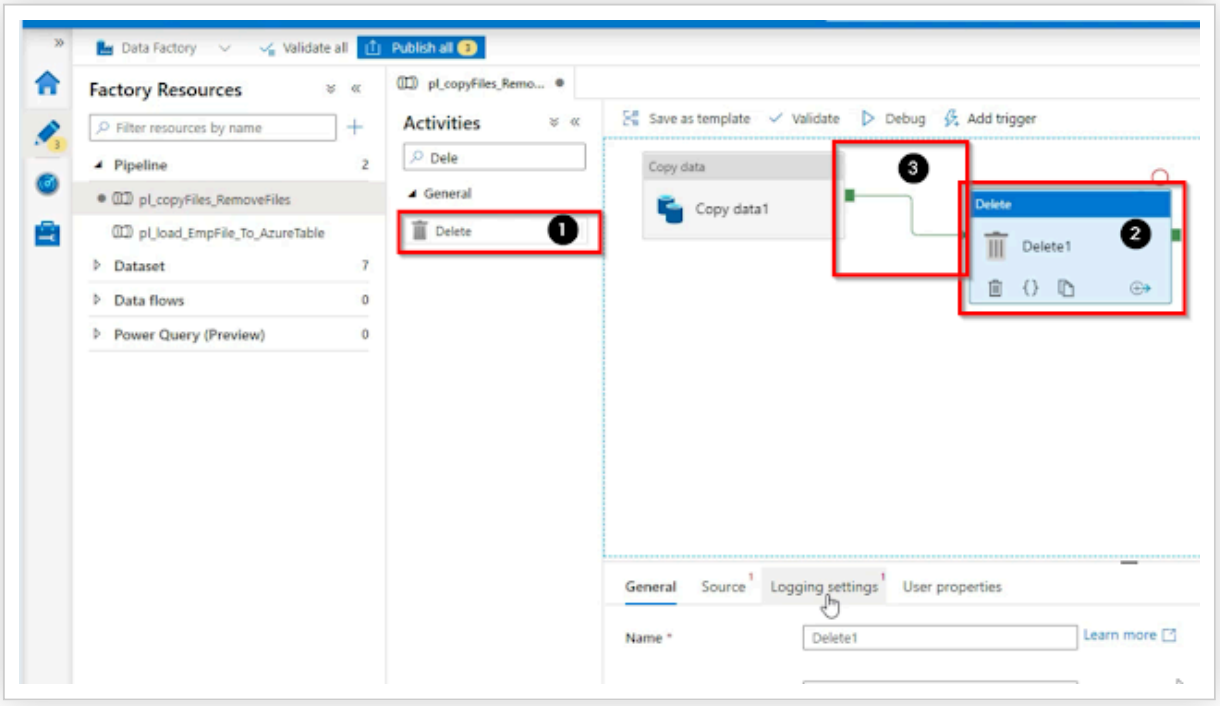
Name the dataset, then select the linked service, in this case, we will select the same linked service that we have selected in the source dataset because we are pointing to the same blob storage, then select the path folder where we need to copy the files, then select none for import schemas and then click on ok.



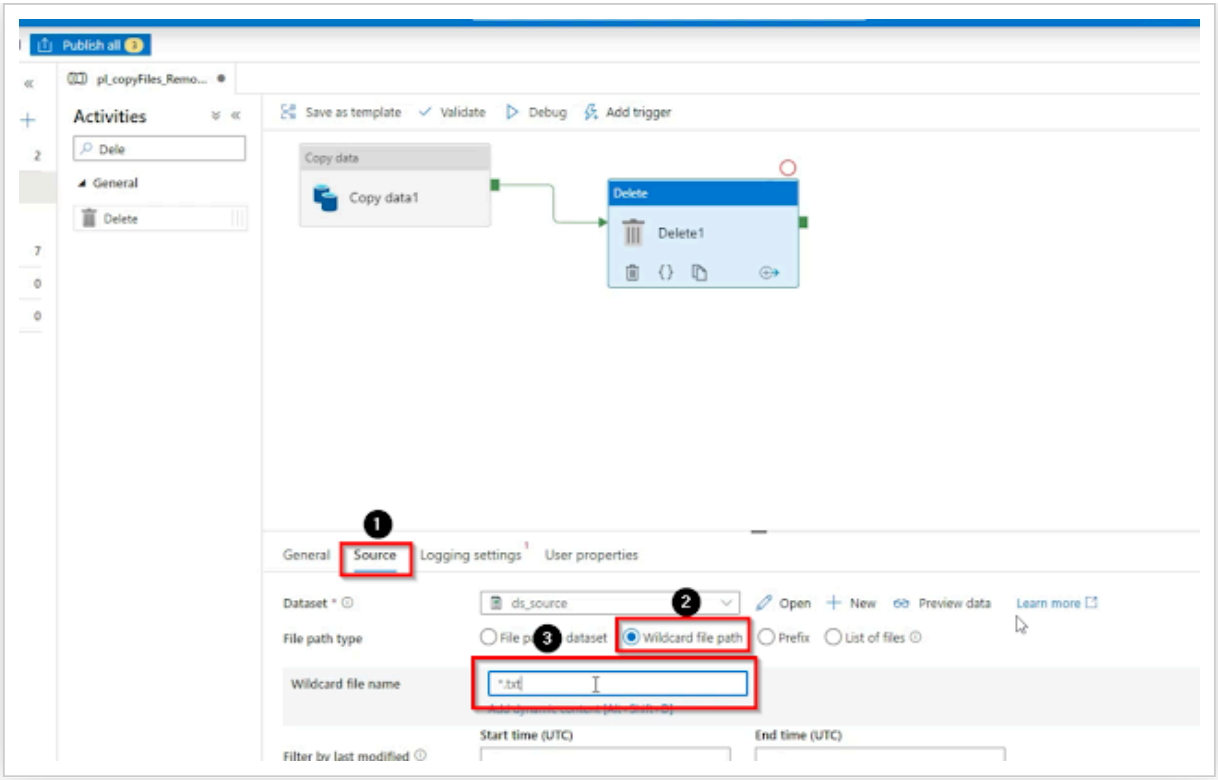
Next, go to the source tab and click on Wildcard file path, as we are going to copy multiple files so we have to give the values, which is (*.txt) which means all the files.



Then go to the activities tab and find and drag the delete activity and connect with the copy activity, as we have to delete files after copying them.



Then click on the source tab where we need to specify which files we want to delete, so click on the wildcard file path then put the values (*.txt) and then go to the pipeline and click on debug.



Once you will click on Debug it will copy all the files from the source path to the sink path and then delete all the files from the source path.