

6th pipeline onprem mysql → azure data factory → azure blob storage

Pipeline Architecture :

Services Required :

1. On prem mysql
 2. Azure data factory
 3. Azure blob storage
- Assuming product data is already there in onprem mysql.

Azure Blob Storage Creation :

- Creation of an azure blob storage is shown in the document below.
- <https://docs.google.com/document/d/1vZCMfM9ieALTQm6Jdwl2JzxcMkwnyvkSJHCrry2kVlo/edit?usp=sharing>
- Create a container called denisdata.

Azure Data Factory Creation :

- Creation of the data factory is shown in the document below.
- Parameters used while creating azure data factory.
- Datafactory name : **projects-datafactory**
- https://docs.google.com/document/d/1lpvA7XumJjbIP0wPWWUlfD_gdvh6jTn_0a12HWv0jcM/edit?usp=sharing
- Click on **Go to resource**.
- Click on **launch workspace**.

Microsoft Azure | Search resources, services, and docs (G+)

Home > Microsoft.DataFactory-20240524172645 | Overview

Deployment

Search << Delete Cancel Redeploy Download Refresh

Overview

Inputs

Outputs

Template

✓ Your deployment is complete

Deployment name : Microsoft.DataFactory-20240524172645 Start time : 5/24/2024, 5:32:31 PM

Subscription : Free Trial Correlation ID : 4d98a809-50e8-4449-bb7a-6cdcdbcb26e

Resource group : projects_rg

> Deployment details

✓ Next steps

Go to resource

Give feedback

Tell us about your experience with deployment

Cost management

Get notified to stay within your budget and prevent unexpected charges on your bill.

Set up cost alerts >

Microsoft Defender for Cloud

Secure your apps and infrastructure

Go to Microsoft Defender for Cloud >

Free Microsoft tutorials

Start learning today >

Work with an expert

Azure experts are service provider partners who can help manage your assets on Azure and be your first line of support.

Find an Azure expert >

Microsoft Azure | Search resources, services, and docs (G+)

Home > projects_rg > projects-datafactory

Data factory (V2)

Search << Delete

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Settings

Networking

Managed identities

Properties

Locks

Getting started

Quick start

Monitoring

Alerts

Metrics

Diagnostic settings

Logs

Essentials

Resource group (move) projects_rg

Type Data factory (V2)

Status Succeeded

Location East US

Subscription (move) Free Trial

Subscription ID c92e5287-784e-4a20-9e6c-8549947f2ee6

JSON View

Azure Data Factory Studio

Launch studio

Quick Starts

Tutorials

Template Gallery

Training Modules

Notifications

More events in the activity log > Dismiss all

✓ Deployment succeeded

Deployment 'Microsoft.DataFactory-20240524172645' to resource group 'projects_rg' was successful.

Pin to dashboard Go to resource group

3 minutes ago

✓ Deployment succeeded

Deployment 'Microsoft.Azure.CosmosDB-20240524163639' to resource group 'projects_rg' was successful.

Go to resource Go to resource group

56 minutes ago

✓ Successfully uploaded blob(s)

Successfully uploaded 4 blob(s).

an hour ago

✓ Successfully created storage container

Successfully created storage container 'denisfactdata'.

an hour ago

✓ Deployment succeeded

Deployment 'denisadlissa_1716546997551' to resource group 'projects_rg' was successful.

Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Data Factory > Validate all Publish all

General

Factory settings

Connections

Linked services

Integration runtimes

Microsoft Purview

Source control

Git configuration

ARM template

Author

Triggers

Global parameters

Data flow libraries

Security

Credentials

Customer managed key

Outbound rules

Managed private endpoint

Workflow orchestration

Linked services

Linked service defines the connection information to a data store or compute. Learn more

+ New

Filter by name Annotations: Any

Showing 1 - 4 of 4 items

Name	Type
adls2adf_ls	Azure Data Lake Storage Gen2
blob2adf_ls	Azure Blob Storage
cosmos2adf_ls	Azure Cosmos DB for MongoDB

New linked service

Data store Compute

Search

All Azure Database File Generic protocol NoSQL Services and apps

IBM Informix

MariaDB

Microsoft Access

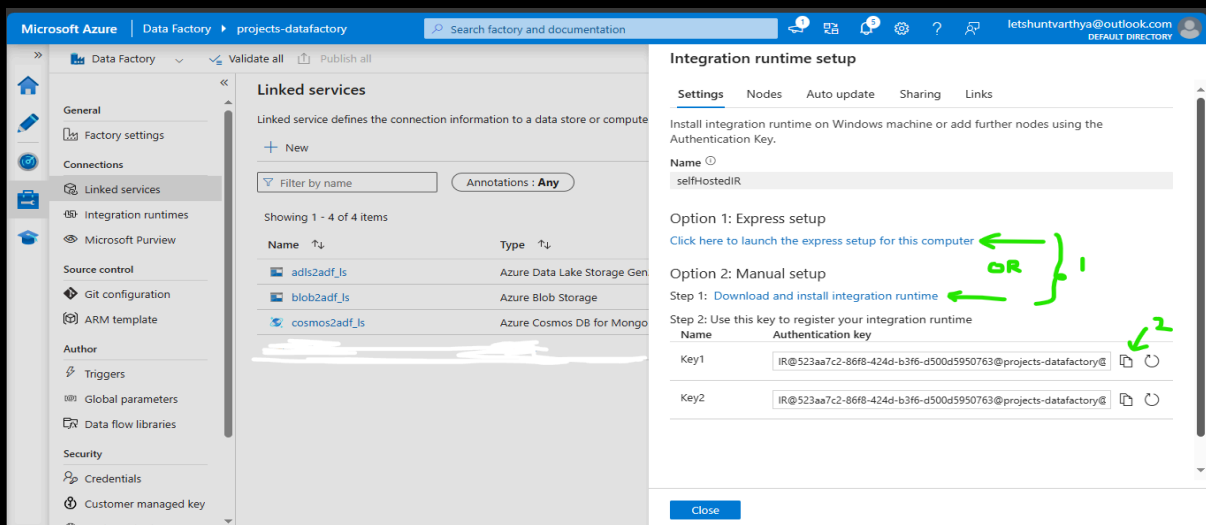
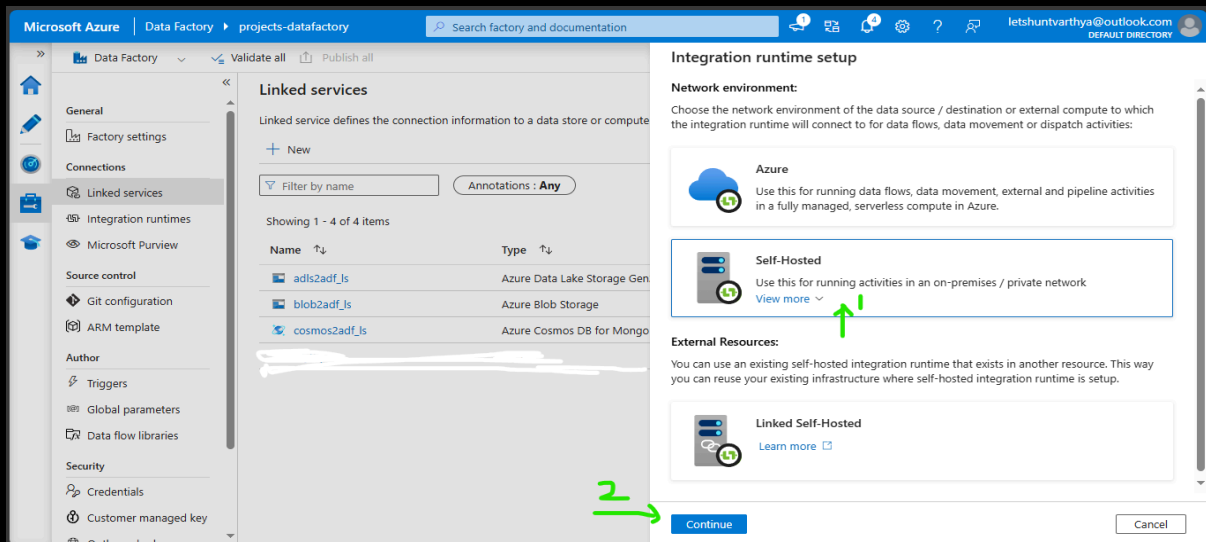
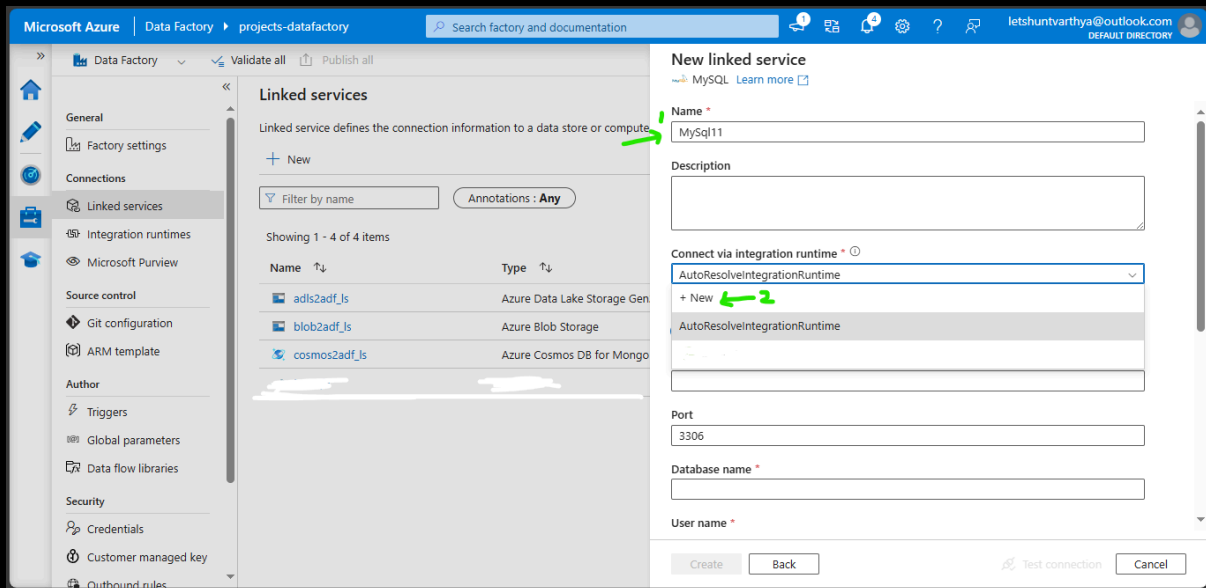
MySQL

Netezza

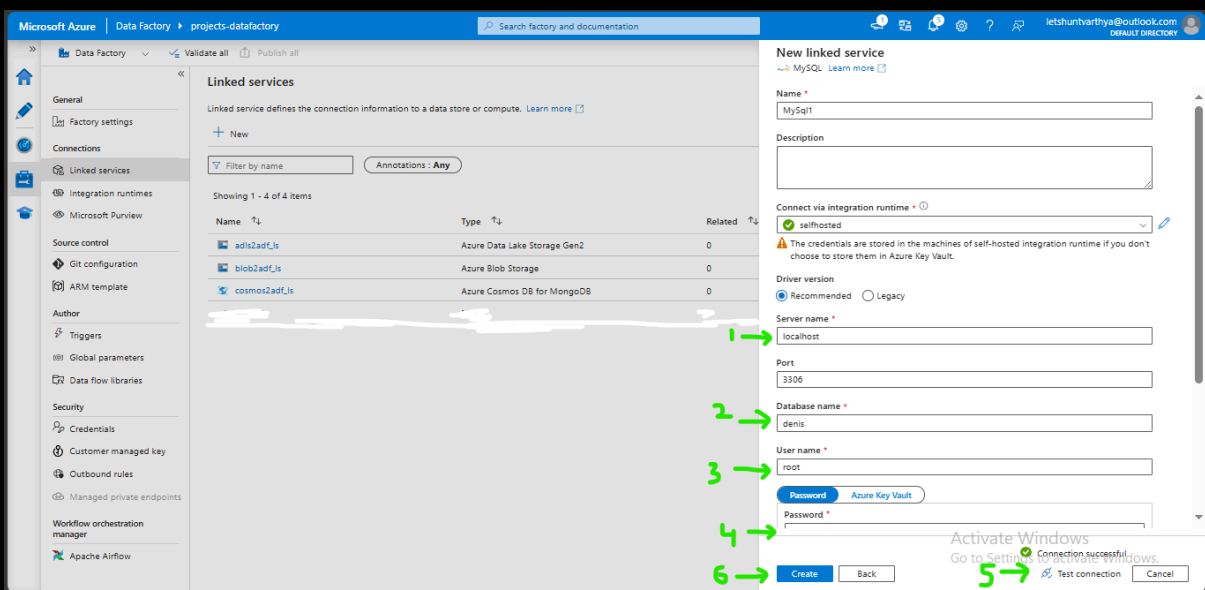
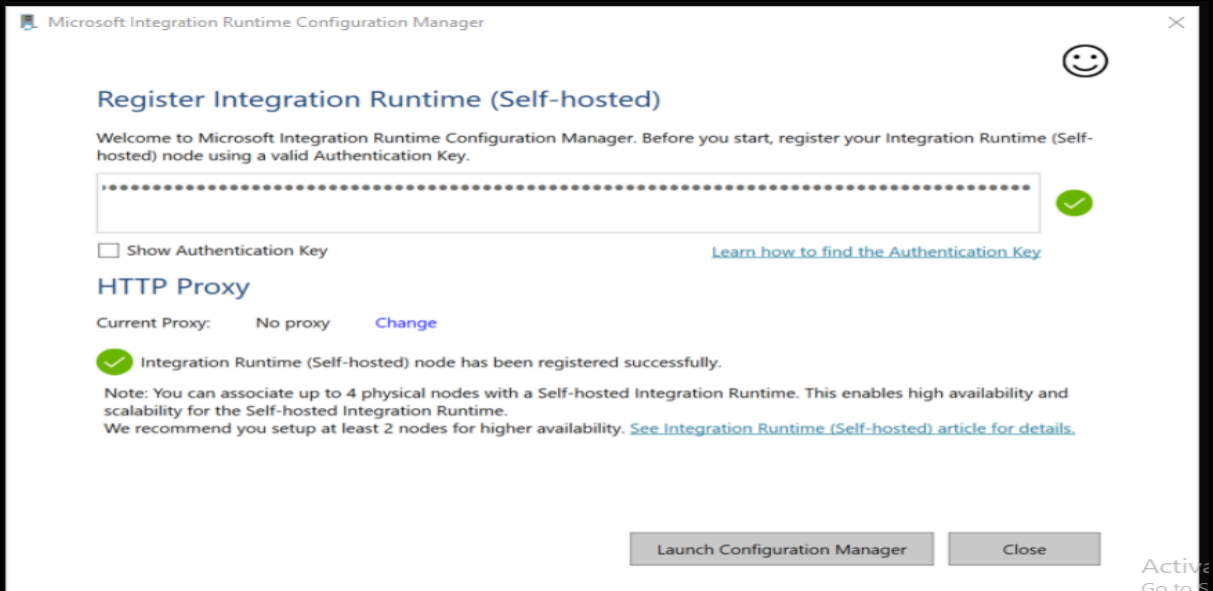
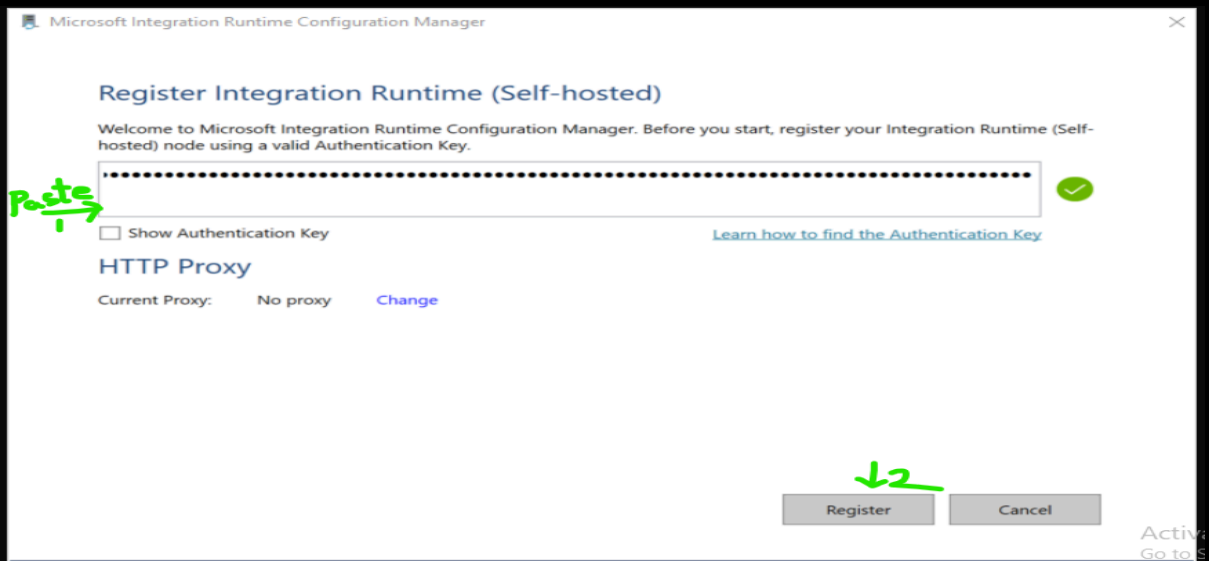
Oracle

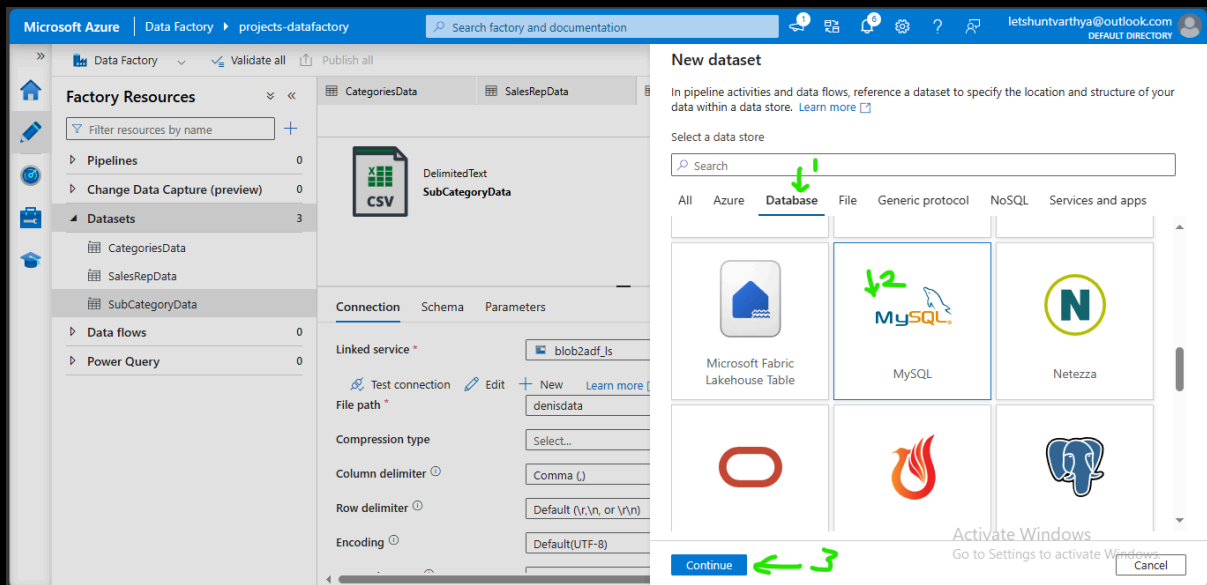
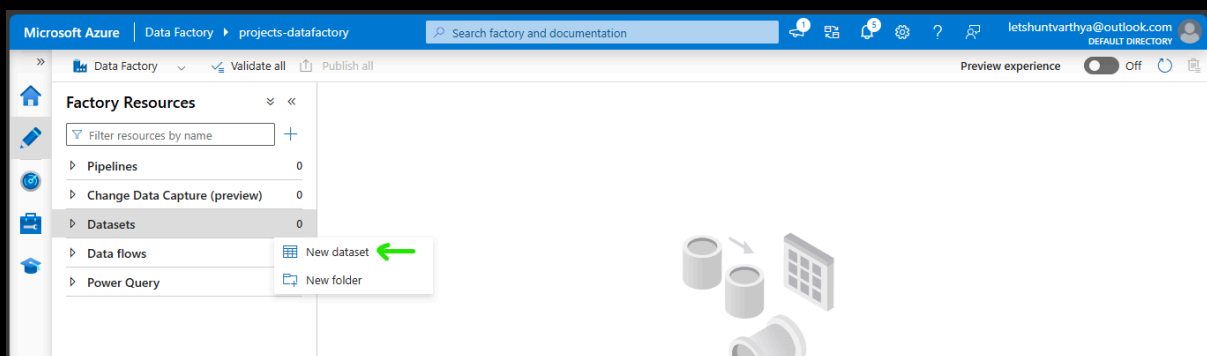
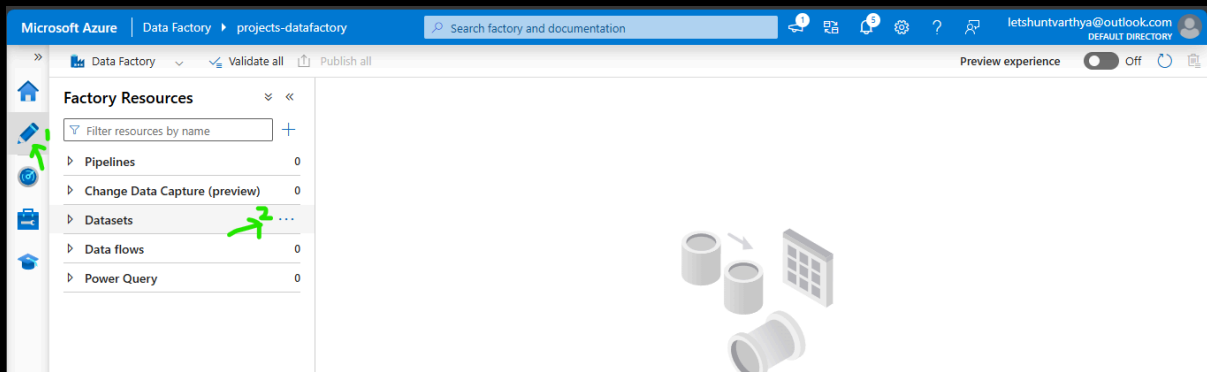
Continue

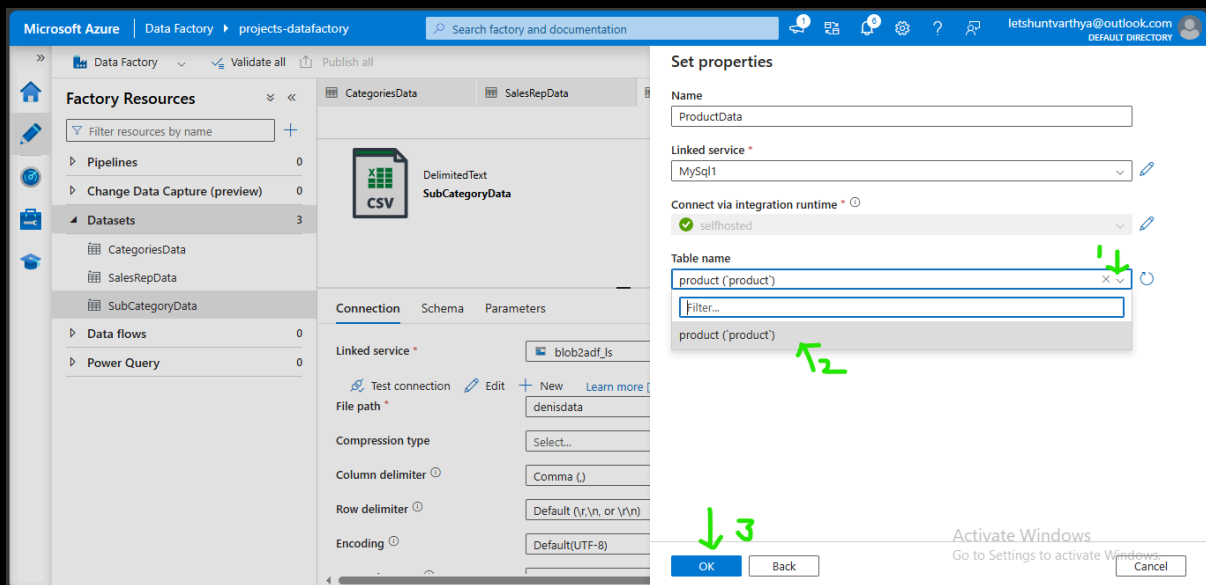
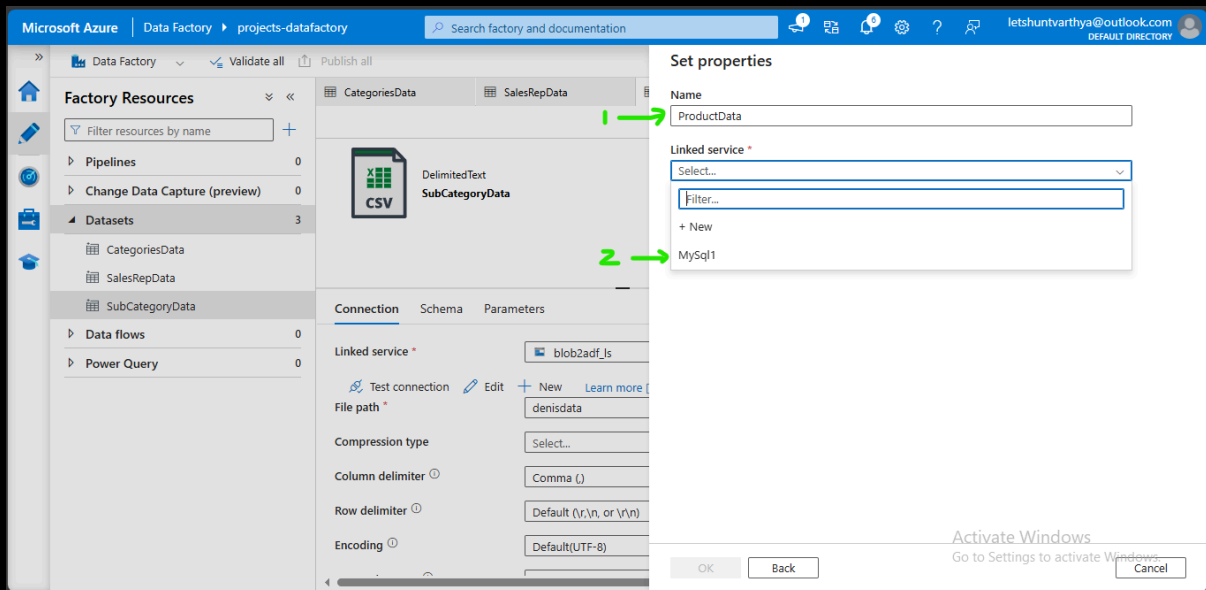
Cancel



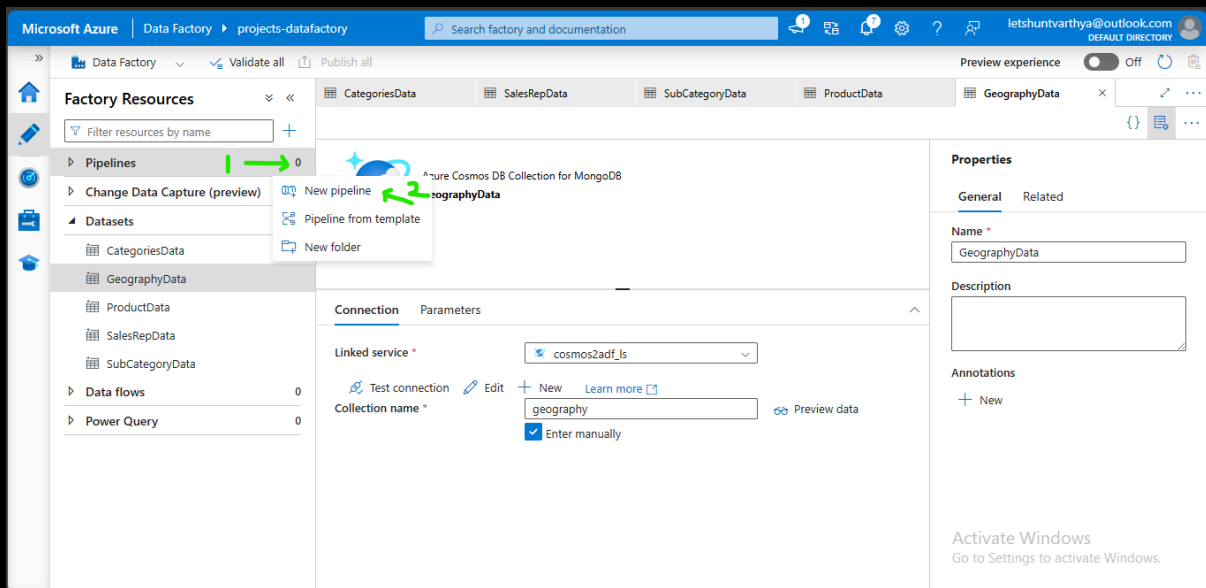
- After downloading the Self Integration Runtime, copy the key1 or key2 and paste in the self Integration Runtime.







Migration of product data from mysql to blob storage.



Microsoft Azure | Data Factory | projects-datafactory

Search factory and documentation

lets huntvarthya@outlook.com

Validate all Publish all

CategoriesData SalesRepData SubCategoryData ProductData GeographyData onPremsql2blob_PL

Activities

Search activities

Move and transform

Copy data

Copy data1

Properties

General Related

Name * onPremsql2blob_PL

Description

Annotations

+ New

Activate Windows

Go to Settings to activate Windows.

General Source Sink Mapping Settings User properties

Name * Copy data1

Description

Activity state ☒ Activated ☐ Deactivated

Microsoft Azure | Data Factory | projects-datafactory

Search factory and documentation

lets huntvarthya@outlook.com

Validate all Publish all

CategoriesData SalesRepData SubCategoryData ProductData GeographyData onPremsql2blob_PL

Activities

Search activities

Move and transform

Copy data

Copy data1

Properties

General Related

Name * onPremsql2blob_PL

Description

Annotations

+ New

Activate Windows

Go to Settings to activate Windows.

General Source Sink Mapping Settings User properties

Source dataset *

Select...

Filter...

CategoriesData

GeographyData

ProductData

SalesRepData

SubCategoryData


General Source Sink Mapping Settings User properties

Sink dataset *

Select...

+ New

New dataset

In pipeline activities and data flows, reference a dataset to specify the location and structure of your data within a data store. [Learn more](#) 

Select a data store

All

Azure

Database

File

Generic protocol

NoSQL

Services and apps



Azure AI Search



Azure Blob Storage



Azure Cosmos DB for
MongoDB



Azure Cosmos DB for
NoSQL



Azure Data Explorer
(Kusto)



Azure Data Lake Storage
Gen2

Continue

Cancel

Select format

Choose the format type of your data



Avro



Binary



DelimitedText



JSON



ORC

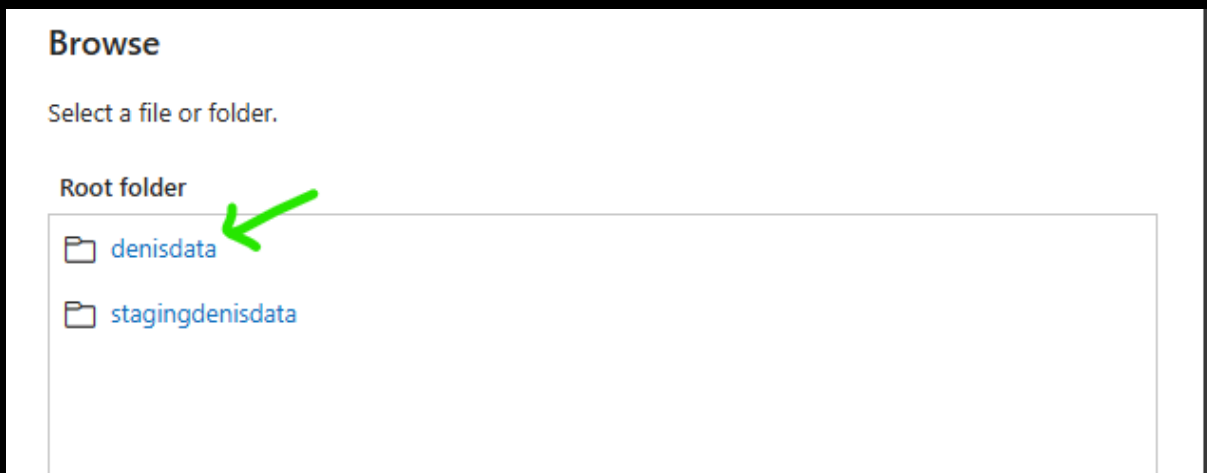
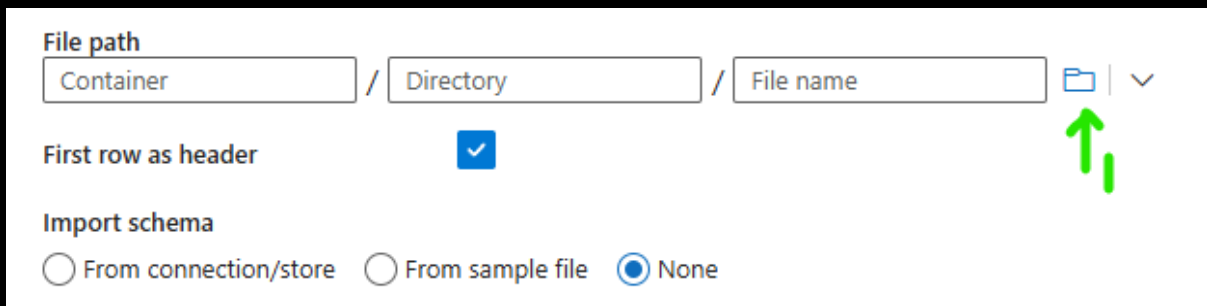
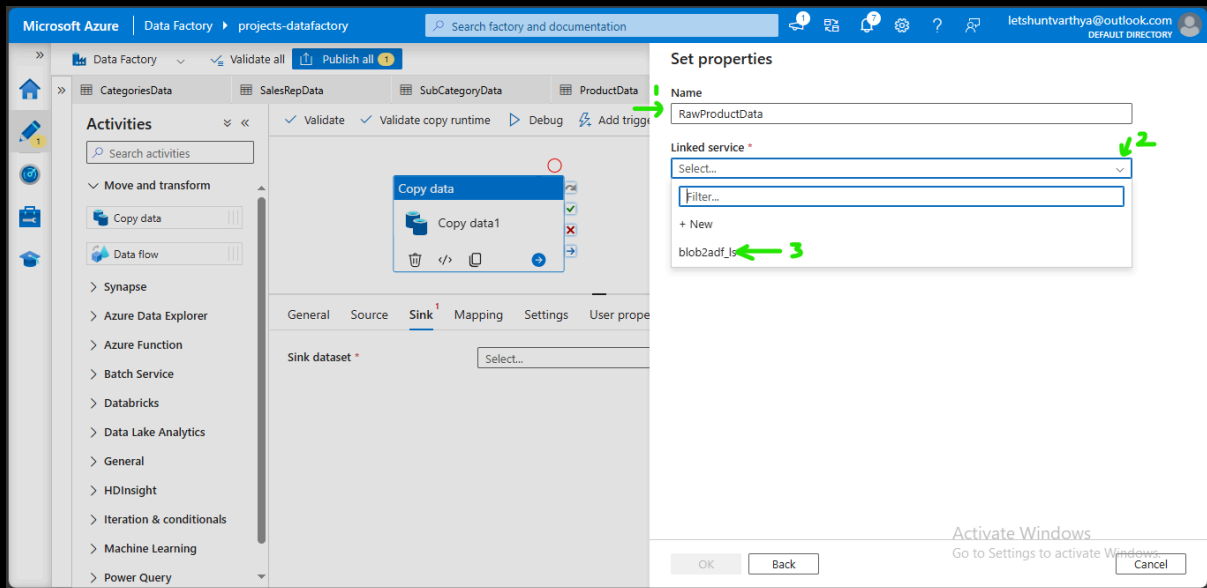


Parquet

Continue

Back

Cancel



Microsoft Azure | Data Factory | projects-datafactory

Search factory and documentation

letshuntvarthya@outlook.com
DEFAULT DIRECTORY

» Data Factory » Validate all » Publish all 1

Activities

- Move and transform
 - Copy data
 - Data flow
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

Copy data

Copy data1

General Source Sink Mapping Settings User properties

Sink dataset *

Select...

Browse

Select a file or folder.

Root folder > denisdata

- Categories.xlsx
- SalesRep.xlsx
- SubCategories.xlsx

Showing 1 - 3 of 3 items

OK Cancel

Activate Windows
Go to Settings to activate Windows.

Microsoft Azure | Data Factory | projects-datafactory

Search factory and documentation

letshuntvarthya@outlook.com
DEFAULT DIRECTORY

» Data Factory » Validate all » Publish all 1

Factory Resources

- Pipelines
 - onPreqsql2blob_PL
- Change Data Capture (preview) 0
- Datasets 6
 - CategoriesData
 - GeographyData
 - ProductData
 - RawProductData
 - SalesRepData
 - SubCategoryData
- Data flows 0
- Power Query 0

DelimitedText
RawProductData

CSV

Connection Schema Parameters

Linked service * blob2adf.js Test connection Edit + New Learn more

File path * denisdata / Directory / Product.csv

Compression type Select...

Column delimiter Comma (,)

Row delimiter Default (\r\n, or \n\r)

Encoding Default(UTF-8)

Quote character Double quote (")

Activate Windows
Go to Settings to activate Windows.

Microsoft Azure | Data Factory | projects-datafactory

Search factory and documentation

letshuntvarthya@outlook.com
DEFAULT DIRECTORY

» Data Factory » Validate all » Publishing 2

Activities

- Move and transform
 - Copy data
 - Data flow
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

Copy data

Copy data1

General Source Sink Mapping Settings User properties

Sink dataset * RawProductData

Open + New Learn more

Copy behavior Select...

Max concurrent connections

Block size (MB)

Metadata + New

Quote all text ☒

Publish all

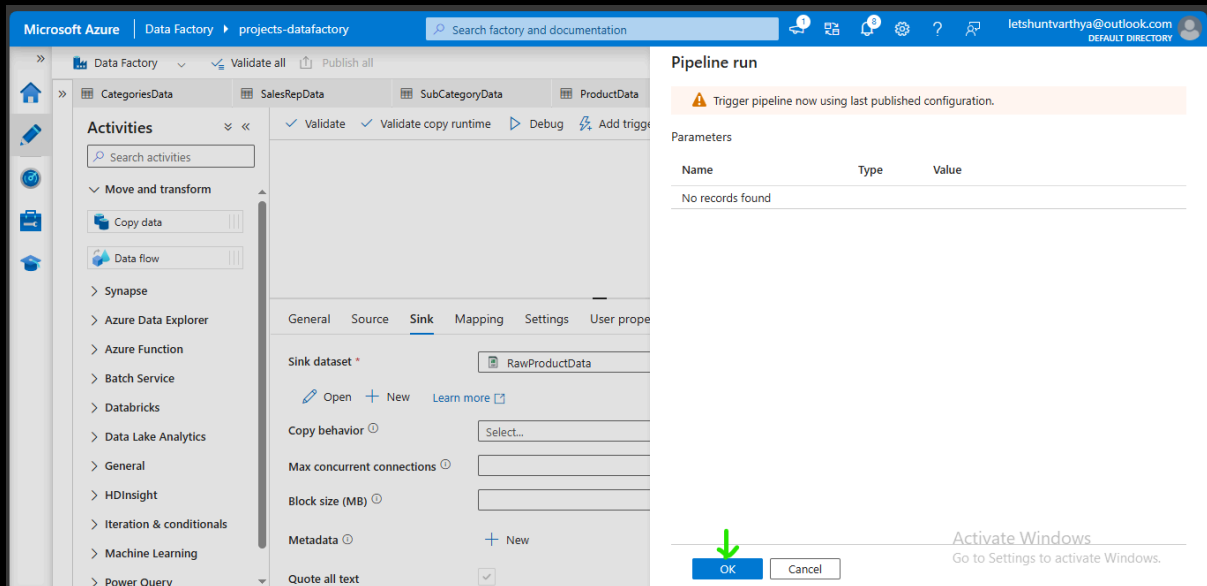
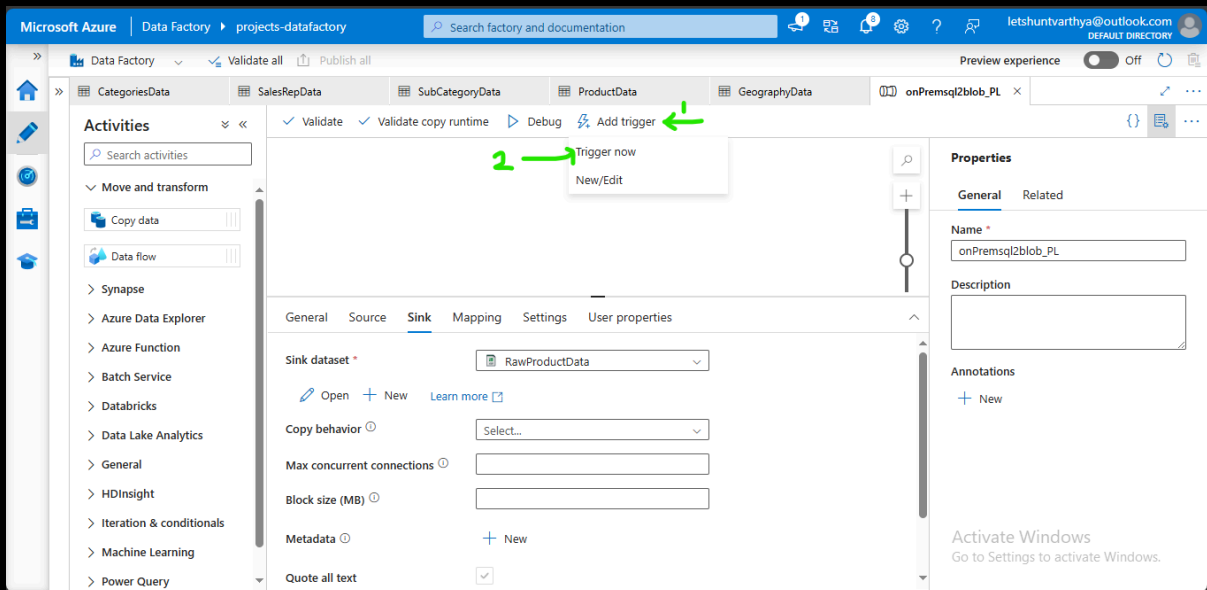
You are about to publish all pending changes to the live environment. Learn more

Pending changes (2)

NAME	CHANGE	EXISTING
Pipelines		
onPreqsql2blob_PL	(New)	-
Datasets		
RawProductData	(New)	-

Publish Cancel

Activate Windows
Go to Settings to activate Windows.



After successful execution of the pipeline we can see the data in blob storage.