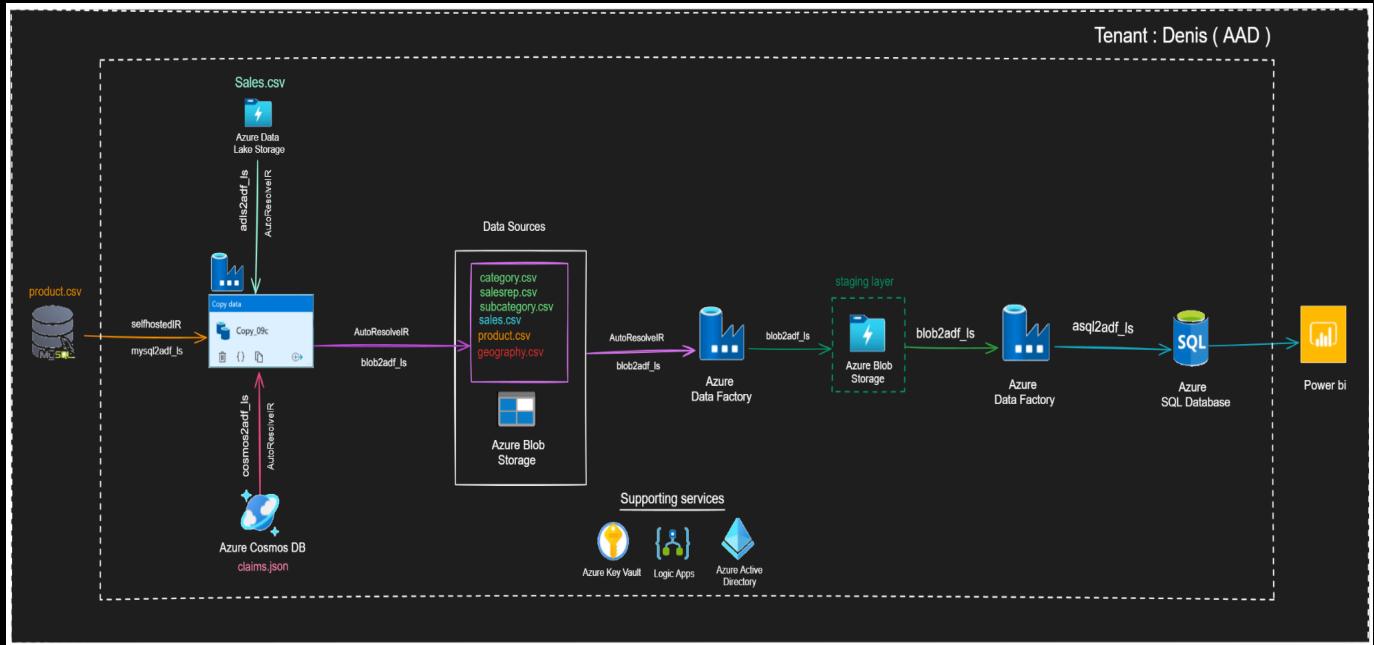


Project Architecture :



Services Required:

1. Azure Blob Storage
2. Azure data lake storage
3. Azure cosmos DB for mongo DB
4. Azure SQL Database
5. Azure Data Factory
6. On prem Mysql

Azure Blob Storage Creation:

The screenshot shows the Microsoft Azure portal interface. At the top, there's a navigation bar with links for 'Create a resource', 'Cost Management', 'Resource groups', 'Subscriptions', 'Quickstart Center', 'Azure AI services', 'Kubernetes services', 'Virtual machines', 'App Services', and 'More services'. Below this is a 'Resources' section with tabs for 'Recent' and 'Favorite'. It lists three resources: 'Free Trial' (Subscription, last viewed 20 hours ago), 'projects_rg' (Resource group, last viewed 3 days ago), and 'projects-datafactory' (Data factory (V2), last viewed 3 days ago). There's also a 'See all' link. Further down, there's a 'Navigate' section with links for 'Subscriptions', 'Resource groups', 'All resources', and 'Dashboard'. The main content area is titled 'Create a resource' and shows a sidebar with categories like Containers, Databases, Developer Tools, DevOps, Identity, Integration, Internet of Things, IT & Management Tools, Media, Migration, Mixed Reality, Monitoring & Diagnostics, Networking, Security, Storage, and Web. To the right, there's a grid of service cards, each with an icon, name, and 'Create | Docs' link. The cards include:

- Function App
- Ubuntu Server 22.04 LTS
- Key Vault
- Red Hat Enterprise Linux 7.4
- Data Factory
- Essentials 50K
- Template deployment (deploy using custom templates)
- MongoDB Atlas (pay-as-you-go)
- Logic App
- Standard
- Automation
- Microsoft Defender for Endpoint
- Public IP address
- Azure Backup - AVS

The URL in the browser bar is <https://portal.azure.com/#create/hub>.

Microsoft Azure

Search resources, services, and docs (G+)

Home > Create a resource

Get Started

Recently created

Categories

- AI + Machine Learning
- Analytics
- Blockchain
- Compute
- Containers
- Databases
- Developer Tools
- DevOps
- Identity
- Integration
- Internet of Things
- IT & Management Tools

Popular Azure services See more in All services

- Azure Cosmos DB
- Storage account
- Azure File Sync
- Data Lake Storage Gen1
- Azure NetApp Files
- Azure Data Box
- Azure Cosmos DB for MongoDB

Getting Started? Try our Quickstart center

Popular Marketplace products See more in Marketplace

- Azure Blob Storage on IoT Edge
- Azure Cost Management plan
- Cloud Manager (by Cap PYGO by Hour, WORM and data services)
- Dionar Managed Azure - CSP
- SFTPGo Standard
- FTP Server Solution for Azure File Share
- SFTP Gateway 3.4

letshunvarthy@outlook...
DEFAULT DIRECTORY (LETSHUN...)

Microsoft Azure

Search resources, services, and docs (G+)

Home > Create a storage account

Basics Advanced Networking Data protection Encryption 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 in which to create the new storage account. Choose a new or existing resource group to organize and manage your storage account together with other resources.

Subscription * Free Trial

Resource group * projects_rg Create new

Instance details

Storage account name * denissblobsa

Region * (US) East US Deploy to an Azure Extended Zone

Performance * Standard: Recommended for most scenarios (general-purpose v2 account)

Redundancy * Locally-redundant storage (LRS)

Previous Next Review + create Give feedback

Microsoft Azure

Search resources, services, and docs (G+)

Home > denissblobsa_1716545281419 | Overview

Your deployment is complete

Deployment name: denissblobsa_1716545281419
Subscription: Free Trial
Resource group: projects_rg

Start time: 5/24/2024 3:38:34 PM
Correlation ID: 8820da7c-61ca-4397-a194-7b5a8ca0346

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 >

1 Microsoft Azure | Overview > denissblobsa | Overview > denissblobsa | Containers

2 + Container

3 Name: denidata

4 Create

5 Advanced

6 The access level is set to private because anonymous access is disabled on this storage account.

7 Give feedback

1 Microsoft Azure | Overview > denissblobsa | Overview > denissblobsa | Containers

2 + Container

3 Show deleted containers

4 denidata

5 Last modified

6 Anonymous access level

7 Lease state

8 ...

9 ...

10 Give feedback

1 Microsoft Azure | Overview > denissblobsa | Overview > denissblobsa | Containers > denidata

2 Upload

3 Authentication method: Access key (Switch to Microsoft Entra user account)

4 Location: denidata

5 Search blobs by prefix (case-sensitive)

6 Add filter

7 Open

8 Categories.xlsx

9 Date modified

10 Modified

11 Access tier

12 Arct

13 Overwrite if files already exist

14 Advanced

15 Upload

16 Give feedback

17 Drag and drop files here or

18 Browse for files

19 File name: "SubCategories.xlsx" "Categories.xlsx"

20 All files

21 Open

22 Cancel

Microsoft Azure Search resources, services, and docs (G+) letshunvarthy@outlook... DEFAULT DIRECTORY (LETSHUN...)

denisdata Container

Overview Diagnose and solve problems Access Control (IAM) Settings Shared access tokens Access policy Properties Metadata

Upload Change access level Refresh Delete Change Authentication method: Access key (Switch to Microsoft Entra user account) Location: denisdata Search blobs by prefix (case-sensitive) Add filter

Name Modified Access tier Archive status Blob type Size Lease state

No results

Upload blob

3 file(s) selected: Categories.xlsx, SalesRep.xlsx, SubCategories.xlsx Drag and drop files here or Browse for files Overwrite if files already exist Advanced Upload Give feedback

Microsoft Azure Search resources, services, and docs (G+) letshunvarthy@outlook... DEFAULT DIRECTORY (LETSHUN...)

denisdata Container

Overview Diagnose and solve problems Access Control (IAM) Settings Shared access tokens Access policy Properties Metadata

Authentication method: Access key (Switch to Microsoft Entra user account) Location: denisdata Search blobs by prefix (case-sensitive) Show deleted blobs

Add filter

Name	Modified	Access tier	Archive status	Blob type	Size	Lease state	...
Categories.xlsx	5/24/2024, 3:55:20 PM	Cool (Inferred)		Block blob	17.03 KiB	Available	...
SalesRep.xlsx	5/24/2024, 3:55:20 PM	Cool (Inferred)		Block blob	17.62 KiB	Available	...
SubCategories.xlsx	5/24/2024, 3:55:20 PM	Cool (Inferred)		Block blob	18.09 KiB	Available	...

Azure Data Lake Storage Creation:

The screenshot shows the Microsoft Azure portal interface. At the top, there's a navigation bar with the Microsoft Azure logo, a search bar, and various icons for account management. Below the navigation bar, the main content area is titled "Azure services". It features a "Create a resource" button with a plus sign icon, followed by icons for Resource groups, Cost Management, Subscriptions, Quickstart Center, Azure AI services, Kubernetes services, Virtual machines, App Services, and a "More services" link. Under the "Resources" section, there's a table listing recent resources: "denissblobsa" (Storage account, 17 minutes ago), "projects_rg" (Resource group, 18 minutes ago), "databricks-rg-optumdatabricks-h4mbsvthosihs" (Resource group, 23 minutes ago), "Free Trial" (Subscription, 24 minutes ago), and "projects-datafactory" (Data factory (V2), 3 days ago). Below the table is a "See all" link. The bottom section is titled "Navigate" and shows the URL "https://portal.azure.com/#create/hub". A modal window titled "Create a resource" is open, showing a sidebar with categories like Containers, Databases, Developer Tools, DevOps, Identity, Integration, Internet of Things, IT & Management Tools, Media, Migration, Mixed Reality, Monitoring & Diagnostics, Networking, Security, Storage, and Web. To the right of the sidebar, there's a grid of service cards, each with an icon, name, "Create" link, and "Learn more" link. The visible services include Function App, Key Vault, Data Factory, Template deployment (deploy using custom templates), Logic App, Automation, Public IP address, Ubuntu Server 22.04 LTS, Red Hat Enterprise Linux 7.4, Essentials 50K, MongoDB Atlas (pay-as-you-go), Standard, Microsoft Defender for Endpoint, and Azure Backup - AVS.

Microsoft Azure

Search resources, services, and docs (G+)

Home > Create a resource

Get Started

Recently created

Categories

- AI + Machine Learning
- Analytics
- Blockchain
- Compute
- Containers
- Databases
- Developer Tools
- DevOps
- Identity
- Integration
- Internet of Things
- IT & Management Tools

Popular Azure services See more in All services

- Azure Cosmos DB Create | Docs | MS Learn
- Storage account Create | Learn more
- Azure File Sync Create | Docs | MS Learn
- Data Lake Storage Gen1 Create | Docs | MS Learn
- Azure NetApp Files Create | Docs | MS Learn
- Azure Data Box Create | Docs | MS Learn
- Azure Cosmos DB for MongoDB Create | Docs | MS Learn

Popular Marketplace products See more in Marketplace

- Azure Blob Storage on IoT Edge Create | Learn more
- Azure Cost Management plan Create | Learn more
- Cloud Manager (by Cap PYGO by Hour, WORM and data services) Set up + subscribe | Learn more
- Dionar Managed Azure - CSP Create | Learn more
- SFTPGo Standard Create | Learn more
- FTP Server Solution for Azure File Share Create | Learn more
- SFTP Gateway 3.4

https://portal.azure.com/#

Microsoft Azure

Search resources, services, and docs (G+)

Home > Create a resource

Create a storage account

Basics Advanced Networking Data protection Encryption 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 in which to create the new storage account. Choose a new or existing resource group to organize and manage your storage account together with other resources.

Subscription: Free Trial

Resource group: projects_rg

Instance details

Storage account name: denissadissa

Region: (US) East US

Performance: Standard: Recommended for most scenarios (general-purpose v2 account)

Redundancy: Locally-redundant storage (LRS)

Previous Next Review + create Give feedback

Microsoft Azure

Search resources, services, and docs (G+)

Home > Create a resource

Create a storage account

Minimum TLS version: Version 1.2

permitted scope for copy operations (preview): From any storage account

Hierarchical Namespace

Enable hierarchical namespace

Access protocols

Enable SFTP

Enable network file system v3

Blob storage

Allow cross-tenant replication

Access tier

Hot Optimized for frequently accessed data and everyday usage scenarios

Cool: Optimized for infrequently accessed data and backup scenarios

Review + create Give feedback

3

Create a storage account

Networking

Network connectivity
You can connect to your storage account either publicly, via public IP addresses or service endpoints, or privately, using a private endpoint.

Network access *

Enable public access from all networks
 Enable public access from selected virtual networks and IP addresses
 Disable public access and use private access
Enabling public access from all networks might make this resource available publicly. Unless public access is required, we recommend using a more restricted access type. [Learn more](#)

Network routing
Determine how to route your traffic as it travels from the source to its Azure endpoint. Microsoft network routing is recommended for most customers.

Routing preference *

Microsoft network routing
 Internet routing



Previous Next Review + create Give feedback

Create a storage account

Data protection

Recovery
Protect your data from accidental or erroneous deletion or modification.

Enable point-in-time restore for containers
Use point-in-time restore to restore one or more containers to an earlier state. If point-in-time restore is enabled, then versioning, change feed, and blob soft delete must also be enabled. [Learn more](#)

Enable soft delete for blobs
Soft delete enables you to recover blobs and directories that were previously marked for deletion. [Learn more](#)
Days to retain deleted blobs

Enable soft delete for containers
Soft delete enables you to recover containers that were previously marked for deletion. [Learn more](#)
Days to retain deleted containers

Enable soft delete for file shares
Soft delete enables you to recover file shares that were previously marked for deletion. [Learn more](#)
Days to retain deleted file shares



Previous Next Review + create Give feedback

Create a storage account

Encryption

Encryption type *

Microsoft-managed keys (MMK)
 Customer-managed keys (CMK)

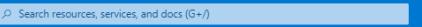
Enable support for customer-managed keys

Blobs and files only
 All service types (blobs, files, tables, and queues)
This option cannot be changed after this storage account is created.

Enable infrastructure encryption



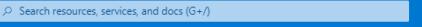
Previous Next Review + create Give feedback

Microsoft Azure  letshunvarthy@outlook...
Home > Create a resource >
Create a storage account ...

Basics Advanced Networking Data protection Encryption **Tags** Review + create

Name	Value	Resource
		All resources selected

Previous Next **Review + create** 

Microsoft Azure  letshunvarthy@outlook...
Home > Create a resource >
Create a storage account ...

Basics Advanced Networking Data protection Encryption **Tags** **Review + create**

[View automation template](#)

Basics

Subscription	Free Trial
Resource group	projects_rg
Location	East US
Storage account name	denissadlissa
Performance	Standard
Replication	Locally-redundant storage (LRS)

Advanced

Enable hierarchical namespace	Enabled
Enable SFTP	Disabled
Enable network file system v3	Disabled
Allow cross-tenant replication	Disabled
Access tier	Cool
Enable large file shares	Enabled

Previous Next **Create** 

Microsoft Azure  letshunvarthy@outlook...
Home >
denissadlissa_1716546997551 | Overview  ...

Deployment

Search     Refresh

Overview  Your deployment is complete

Deployment name: denissadlissa_1716546997551 Start time: 5/24/2024, 4:07:37 PM
Subscription: Free Trial Correlation ID: 6d0e95f4-41cc-4ae2-a393-2fb29d81c4c4 

 Deployment details  Next steps

Go to resource 

Give feedback 

 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 | denissadlssa | Overview | denissadlssa

New container

Name *: denisfactdata

Anonymous access level: Private (no anonymous access)

The access level is set to private because anonymous access is disabled on this storage account.

Create

Microsoft Azure | denissadlssa | Overview | denissadlssa

Containers

Name	Last modified	Anonymous access level	Lease state
\$logs	5/24/2024, 4:08:05 PM	Private	Available
denisfactdata	5/24/2024, 4:23:00 PM	Private	Available

Microsoft Azure | denissadlssa | Overview | denissadlssa | Containers > denisfactdata

Upload blob

Drag and drop files here or **Browse for files**

Upload

File name: "Sales 2014" "sales 2015" "sales 2016" "sales 2017" **Open**

Microsoft Azure

denisadlssa_1716546997551 | Overview > denisadlssa | Containers >

denisfactdata Container

Search

Overview

Diagnose and solve problems

Access Control (IAM)

Settings

Shared access tokens

Manage ACL

Access policy

Properties

Metadata

Name Modified Access tier

No results

Upload blob

4 file(s) selected: Sales 2014.csv, sales 2015.csv, sales 2016.csv...

Drag and drop files here or [Browse for files](#)

Overwrite if files already exist

Advanced

Upload

Give feedback

<https://portal.azure.com/#home>

Microsoft Azure

denisadlssa_1716546997551 | Overview > denisadlssa | Containers >

denisfactdata Container

Search

Overview

Diagnose and solve problems

Access Control (IAM)

Settings

Shared access tokens

Manage ACL

Access policy

Properties

Metadata

Name Modified Access tier Archive status Blob type Size Lease state

Sales 2014.csv 5/24/2024, 4:26:45 PM Cool (Inferred) Block blob 617.38 KiB Available ...

sales 2015.csv 5/24/2024, 4:26:46 PM Cool (Inferred) Block blob 640.35 KiB Available ...

sales 2016.csv 5/24/2024, 4:26:45 PM Cool (Inferred) Block blob 625.03 KiB Available ...

sales 2017.csv 5/24/2024, 4:26:45 PM Cool (Inferred) Block blob 915.34 KiB Available ...

Show deleted objects

Azure Cosmos DB for MongoDB creation:

The screenshot shows the Microsoft Azure portal homepage. At the top, there's a search bar and several service icons. Below the search bar, there's a section titled "Azure services" with a "Create a resource" button highlighted by a green arrow. To the right of this are icons for Resource groups, Cost Management, Subscriptions, Quickstart Center, Azure AI services, Kubernetes services, Virtual machines, App Services, and More services. Below this is a "Resources" section with a "Recent" tab selected, showing a list of resources including Storage accounts, Resource groups, and Data factories. At the bottom, there's a "Navigate" section with links to Subscriptions, Resource groups, All resources, and Dashboard.

The screenshot shows the "Create a resource" page in the Microsoft Azure portal. On the left, there's a sidebar with categories like Blockchain, Compute, Containers, Databases, Developer Tools, DevOps, Identity, Integration, Internet of Things, IT & Management Tools, Media, Migration, Mixed Reality, Monitoring & Diagnostics, Networking, Security, Storage (highlighted with a green arrow), and Web. The main area lists various Azure services with their icons and "Create | Learn more" links. One item, "Azure Cosmos DB for MongoDB", is highlighted with a green arrow. Other items include Azure File Sync, Data Lake Storage Gen1, Azure NetApp Files, Azure Data Box, SFTPGo Standard, FTP Server Solution for Azure File Share, SFTP Gateway 3.4, SFTP Secure Server SSH on Windows Server 2016, Azure Backup as a service, and SUSE Enterprise Linux for SAP 12 SP5 - BYOS.

The screenshot shows the "Create Azure Cosmos DB Account - Choose Architecture" page. It asks "Which type of resource?" and explains that Azure Cosmos DB offers two resource types: Request unit (RU) database accounts and vCore clusters. It says "To start, select the type to create a resource. The resource selection cannot be changed after creation." Two options are shown: "Request unit (RU) database account" and "vCore cluster (Recommended)". Both have a list of benefits and a "Create" button. The "Request unit (RU) database account" option is highlighted with a green arrow pointing to its "Create" button.

Create Azure Cosmos DB Account - Azure Cosmos DB for MongoDB

Project Details
Select the 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
Account Name * Availability zone settings for your account. You cannot change these settings once the account is created.

Availability Zones Enable Disable

Location * Available locations are determined by your subscription's access and availability zone support (if that is enabled). If you don't see or cannot select your desired location, please open a support request for region access.
[Click here for more details on how to create a region access request](#)

Capacity mode Provisioned throughput Serverless
[Learn more about capacity mode](#)

With Azure Cosmos DB free tier, you will get the first 1000 RU/s and 25 GB of storage for free in an account. You can enable free tier on up to one account per subscription. Estimated \$64/month discount per account.

Apply Free Tier Discount Apply Do Not Apply

Review + create **Previous** **Next: Global distribution** 

Create Azure Cosmos DB Account - Azure Cosmos DB for MongoDB

Global distribution Basics Global distribution Networking Backup Policy Encryption Tags Review + create

Configure global distribution and regional settings for your account. You can also change these settings after the account is created.

Geo-Redundancy Enable Disable

Multi-region Writes Enable Disable

Review + create **Previous** **Next: Networking** 

Create Azure Cosmos DB Account - Azure Cosmos DB for MongoDB

Basics Global distribution Networking Backup Policy Encryption Tags Review + create

Network connectivity
You can connect to your Azure Cosmos DB account either publicly, via public IP addresses or service endpoints, or privately, using a private endpoint.

Connectivity method All networks Public endpoint (selected networks) Private endpoint

All networks will be able to access this CosmosDB account. [Learn More](#)

Connection Security Settings
Minimum Transport Layer Security Protocol 
This account only accepts this protocol. [Learn more](#)

Review + create **Previous** **Next: Backup Policy** 

Microsoft Azure Search resources, services, and docs (G+) Home > Create a resource > Create Azure Cosmos DB Account - Choose Architecture > Create Azure Cosmos DB Account - Azure Cosmos DB for MongoDB

Basics Global distribution Networking **Backup Policy** Encryption Tags Review + create

Azure Cosmos DB provides three different backup policies. You will not be able to switch to Periodic mode once you adopt Continuous mode. [Learn more](#) about the differences of the backup policies and pricing details.

Backup policy **Periodic**: Backup is taken at periodic interval based on your configuration
 Continuous (7 days): Provides backup window of 7 days / 168 hours and you can restore to any point of time within the window. This mode is available for free.
 Continuous (30 days): Provides backup window of 30 days / 720 hours and you can restore to any point of time within the window. This mode has cost impact.

Backup interval Minute(s)

Backup retention Hours(s)

Copies of data retained

Review + create Previous Next: Encryption  Feedback

Microsoft Azure Search resources, services, and docs (G+) Home > Create a resource > Create Azure Cosmos DB Account - Choose Architecture > Create Azure Cosmos DB Account - Azure Cosmos DB for MongoDB

Basics Global distribution Networking Backup Policy **Encryption** Tags Review + create

Data Encryption

Azure Cosmos DB encryption protects your data at rest by seamlessly encrypting your data as it's written in our datacenters, and automatically decrypting it for you as you access it.

By default your Azure Cosmos DB account is encrypted at rest using service-managed keys. At the moment, you will not be able to switch back to service-managed key after opting into using custom-managed key while creating your account. [Learn More](#)

Data Encryption * Service-managed key
 Customer-managed key (CMK)

Review + create Previous Next: Tags  Feedback

Microsoft Azure Search resources, services, and docs (G+) Home > Create a resource > Create Azure Cosmos DB Account - Choose Architecture > Create Azure Cosmos DB Account - Azure Cosmos DB for MongoDB

Basics Global distribution Networking Backup Policy Encryption **Tags** Review + create

Tags are name/value pairs that enable you to categorize resources and view consolidated billing by applying the same tag to multiple resources and resource groups. [Learn more](#)

Note that if you create tags and then change resource settings on other tabs, your tags will be automatically updated.

Key	Value	Resource Type
<input type="text"/>	<input type="text"/>	Azure Cosmos DB account

Review + create Previous Next: Review + create  Feedback

Create Azure Cosmos DB Account - Azure Cosmos DB for MongoDB

Validation Success

Basics Global distribution Networking Backup Policy Encryption Tags [Review + create](#)

Creation Time

Estimated Account Creation Time (in minutes) 3

The estimated creation time is calculated based on the location you have selected

Basics

Subscription	Free Trial
Resource Group	projects_rg
Location	East US 2
Account Name	(new) deniscosmossa
API	Azure Cosmos DB for MongoDB
Capacity mode	Provisioned throughput
Geo-Redundancy	Disable
Multi-region Writes	Disable
Availability Zones	Disable

[Create](#) [Previous](#) [Next](#) [Download a template for automation](#) [Feedback](#)

Microsoft.Azure.CosmosDB-20240524163639 | Overview

Your deployment is complete

Deployment name : Microsoft.Azure.CosmosDB-20240524163639
Subscription : Free Trial
Resource group : projects_rg

Start time : 5/24/2024, 4:36:52 PM
Correlation ID : 04687416-5651-43da-bdde-79e861f757e6

[Go to resource](#)

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 >](#)

deniscosmossa | Overview

Welcome to your Azure Cosmos DB Free Tier account! Your first 1000 RU/s and 25 GB of storage will be free for the lifetime of this account. Click here to learn more. →

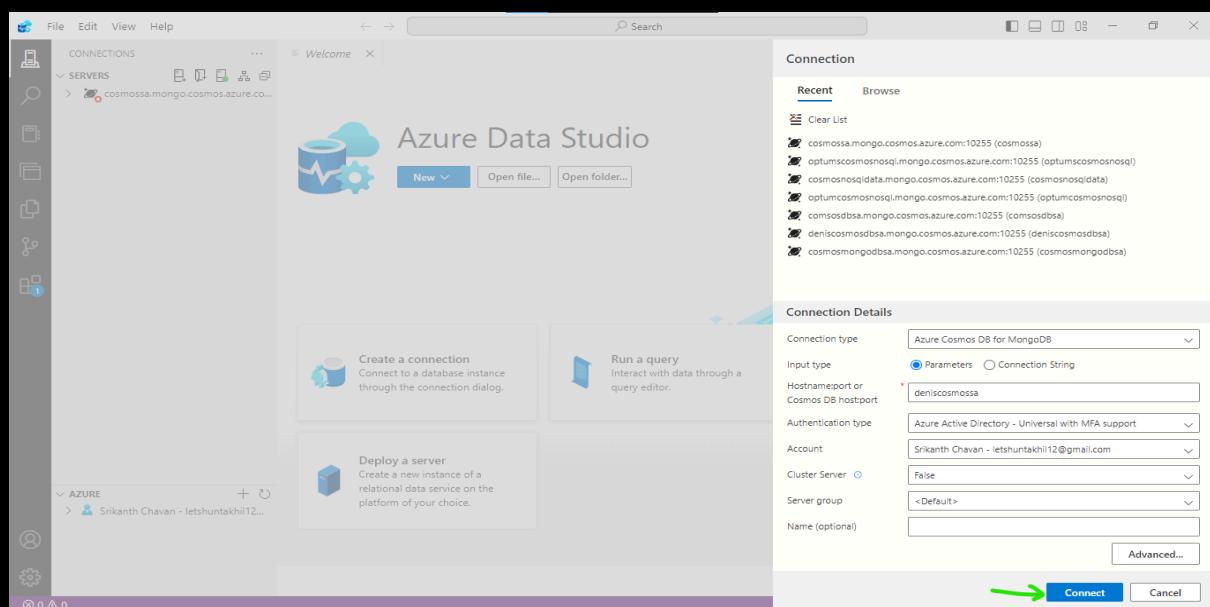
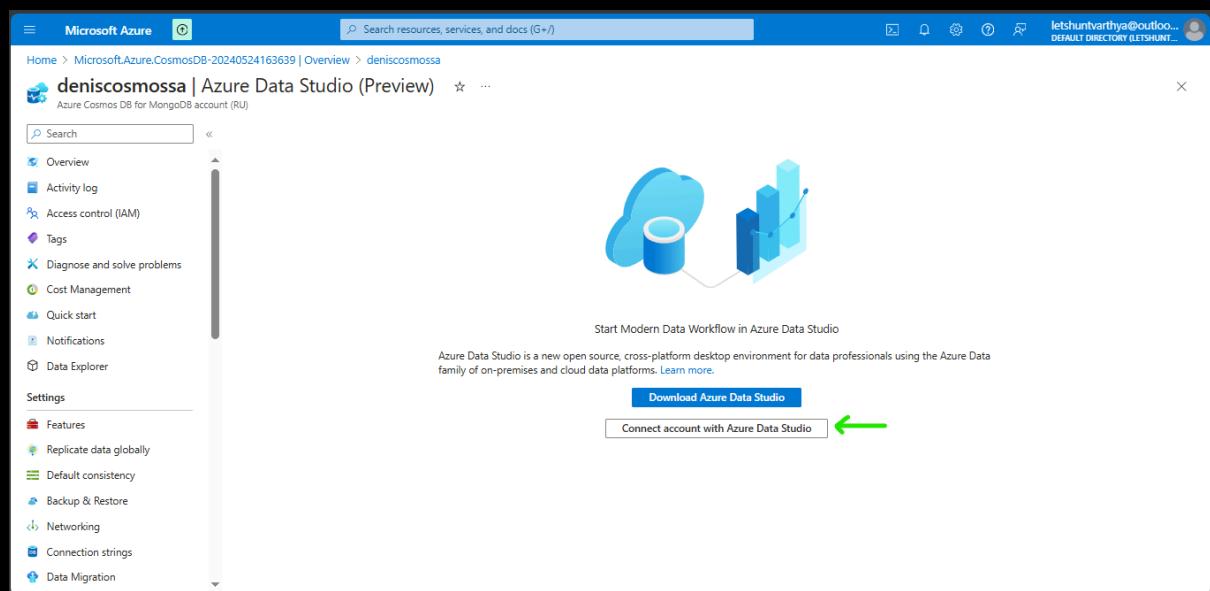
Essentials

Status : Online	Read Locations : East US 2
Resource group (move) : projects_rg	Write Locations : East US 2
Subscription (move) : Free Trial	URI : https://deniscosmossa.mongo.cosmos.azure.com:443/
Subscription ID : e92e5287-784e-4a20-9e6c-85499472ee6	Server Version : 6.0
Total throughput limit : 1000 RU/s	Free Tier Discount : Opted In

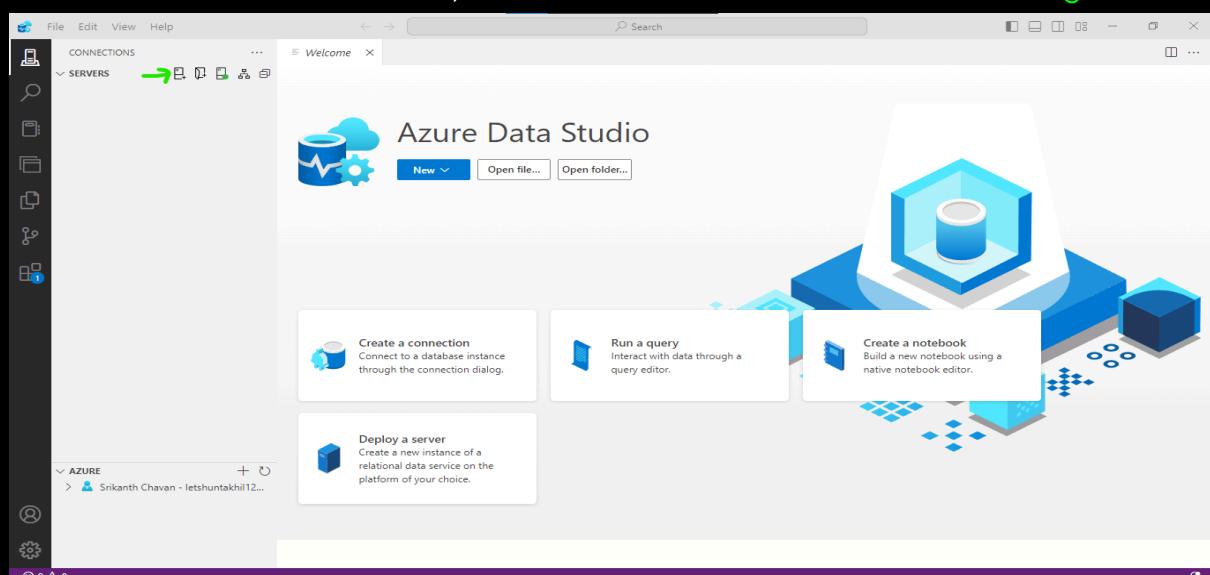
Collections
Looks like you don't have any collections yet. [Data Explorer](#)

Monitoring
Show data for last [1 hour](#) **24 hours** [7 days](#) [30 days](#)

Number of requests per 5 minutes [Request Charge](#)



- If that doesn't connect , we can connect it with **connection strings**



Microsoft Azure | denicosmossa | Connection strings

Overview Activity log Access control (IAM) Tags Diagnose and solve problems Cost Management Quick start Notifications Data Explorer Settings Features Replicate data globally Default consistency Backup & Restore Networking Connection strings Data Migration

Search Refresh Feedback

10255

USERNAME: denicosmossa
SSL: true

Azure Cosmos DB has strict security requirements and standards. Azure Cosmos DB accounts require authentication and secure communication via SSL.

Read-write Keys Read-only Keys

PRIMARY PASSWORD: [REDACTED]

Last regenerated: 24/5/2024 (0 days ago). Learn more

SECONDARY PASSWORD: [REDACTED]

Last regenerated: 24/5/2024 (0 days ago). Learn more

PRIMARY CONNECTION STRING: mongodbs://denicosmossacdA4QFV74jUkFrQq3VNIIWWt3fnK6k1RywWkIdyL7xyS8hqS8330mEKC2BMESLPu5sk1ACDbb0ar1A==@denicosmossa.mongo.cosmos.azure.com:10255

SECONDARY CONNECTION STRING: [REDACTED]

File Edit View Help

CONNECTIONS SERVERS

Welcome

Create a connection Run a query Deploy a server

cosmos.mongo.cosmos.azure.com:10255 (cosmos)

Recent Browse

Clear List

Paste

Connection Details

Connection type: Microsoft SQL Server

Input type: Connection String

Connection string: mongodbs://denicosmossacdA4QFV74jUkFrQq3VNIIWWt3fnK6k1RywWkIdyL7xyS8hqS8330mEKC2BMESLPu5sk1ACDbb0ar1A==@denicosmossa.mongo.cosmos.azure.com:10255/?ssl=true&replicaSet=globaldb&retryWrites=false&maxidleTimeMS=120000&appName=@denicosmossa@

Server group: <Default>

Name (optional):

Connect Cancel

File Edit View Help

CONNECTIONS SERVERS

Welcome Enter database name (Press 'Enter' to confirm or 'Escape' to cancel)

Home > denicosmossa.mongo.cosmos.azure.com:10255

New Database Open Mongo Shell Refresh Learn more

General Databases

Getting started

GeographyDB

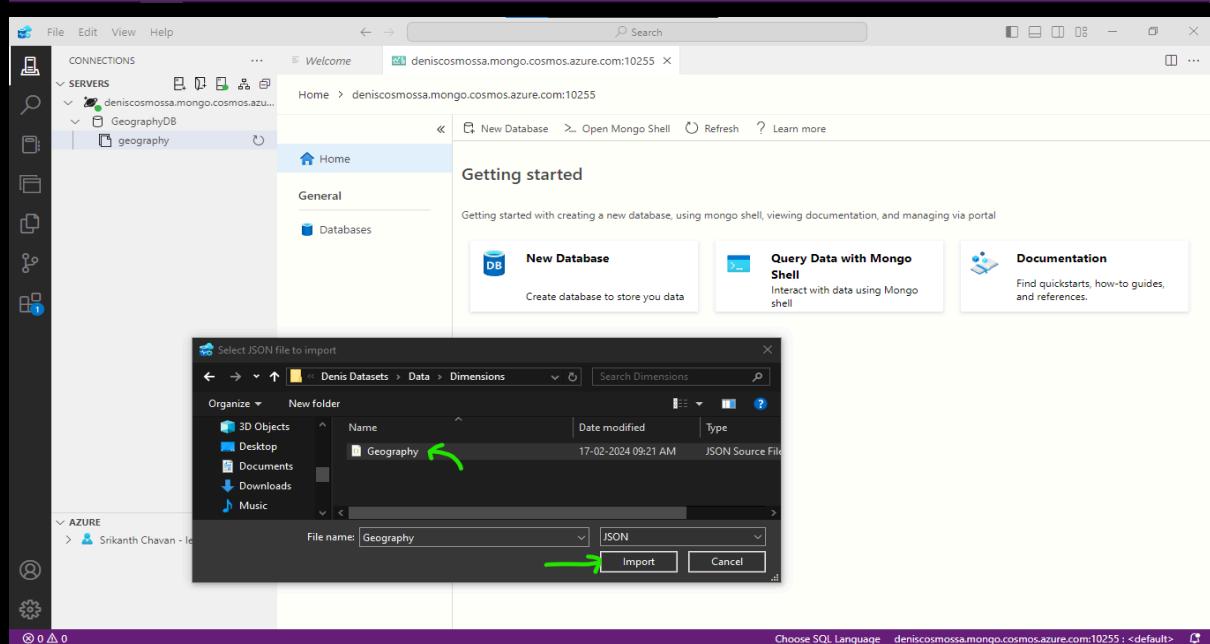
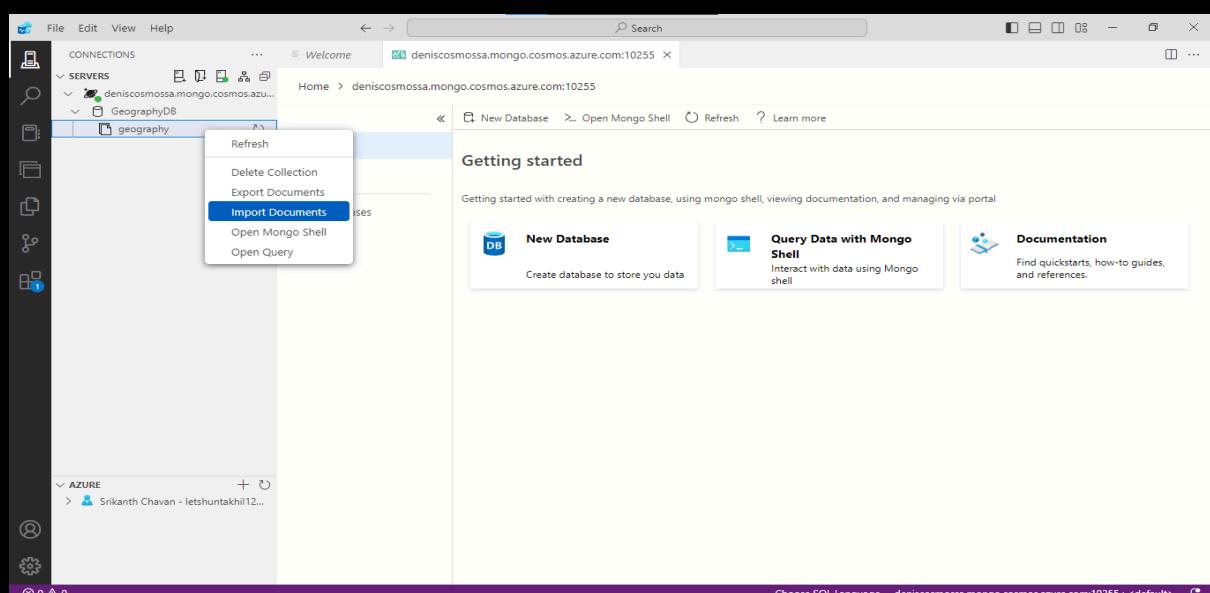
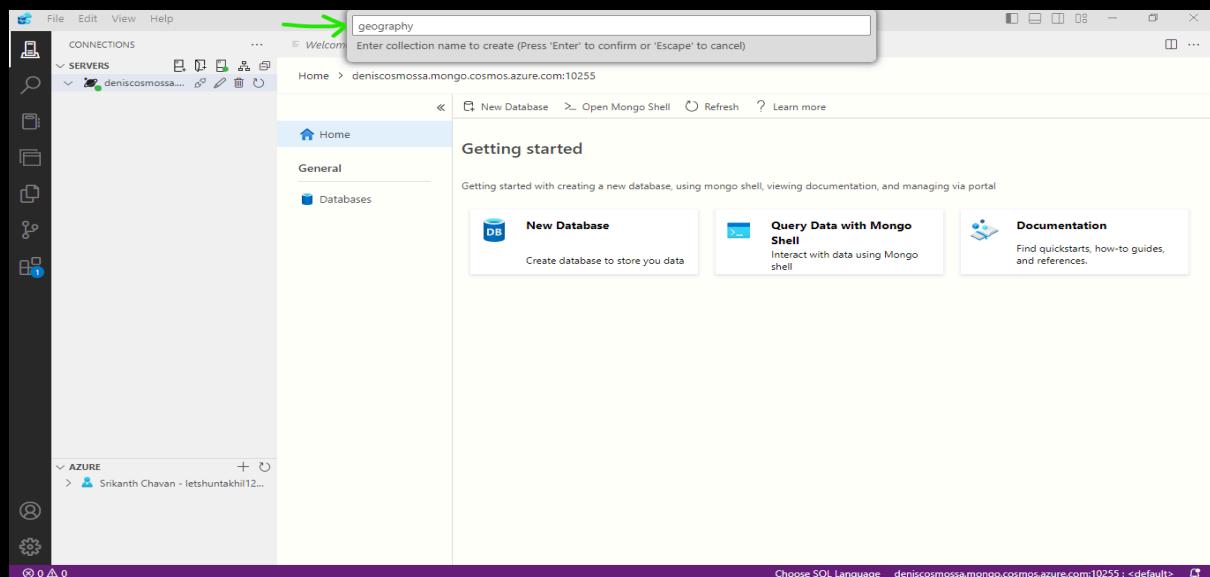
New Database Create database to store your data

Query Data with Mongo Shell Interact with data using Mongo shell

Documentation Find quickstarts, how-to guides, and references.

AZURE Srikanth Chavan - letshuntakhil12...

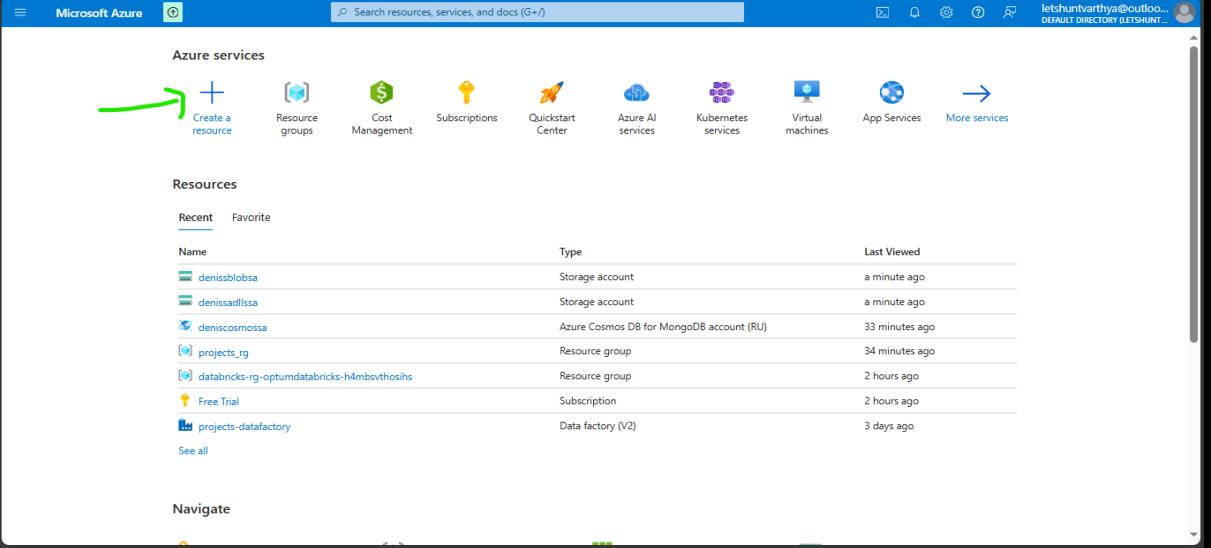
Choose SQL Language denicosmossa.mongo.cosmos.azure.com:10255 : <default>



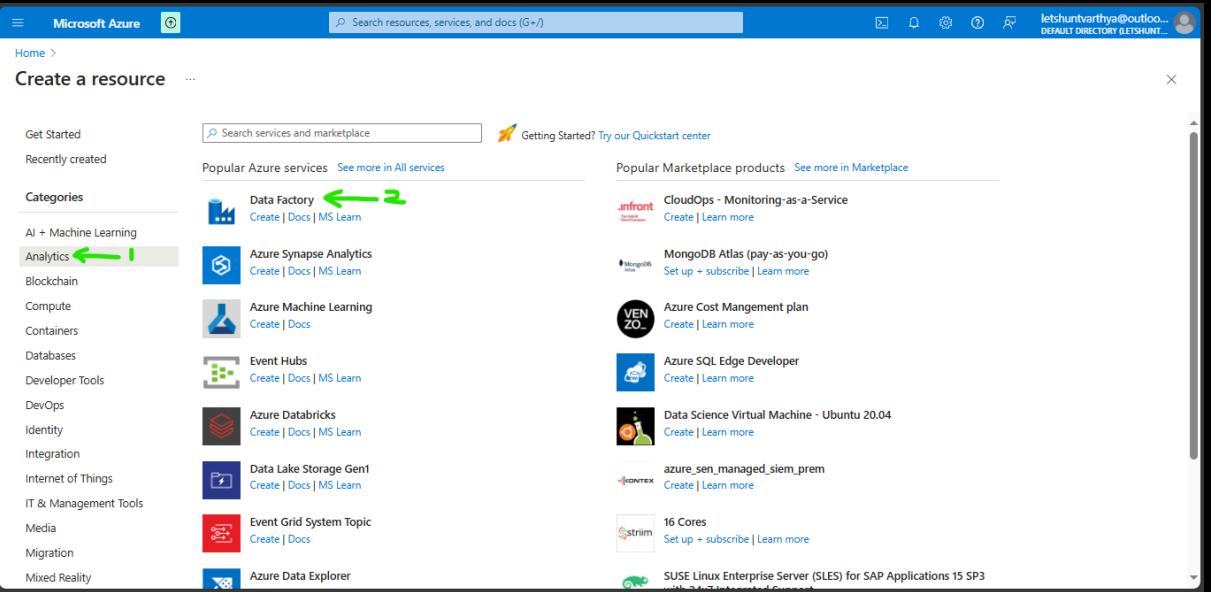
The screenshot shows the Azure portal interface. On the left, the 'Connections' sidebar lists a single server named 'denicosmossa.mongo.cosmos.azure.com:10255'. The main content area is titled 'Home' and displays the 'Getting started' section for MongoDB. It includes three cards: 'New Database' (Create database to store your data), 'Query Data with Mongo Shell' (Interact with data using Mongo shell), and 'Documentation' (Find quickstarts, how-to guides, and references). A green arrow points from the text 'Interact with data using Mongo shell' to the 'Query Data with Mongo Shell' card. A message bar at the bottom right indicates 'Successfully inserted 7 documents (took 2s)'. The status bar at the bottom shows 'Choose SQL Language denicosmossa.mongo.cosmos.azure.com:10255: <default>'.

The screenshot shows the Azure portal interface with the same connection details. The main content area now displays the 'TERMINAL' tab, which shows the MongoDB shell command-line interface. The terminal output includes the connection string, MongoDB version (6.0.0), and Mongosh version (1.6.0). It then shows a command being run: 'use GeographyDB' (marked with a green arrow 1), followed by 'geography.find()' (marked with a green arrow 2). An error message 'ReferenceError: geography is not defined' is displayed. The status bar at the bottom shows 'Choose SQL Language denicosmossa.mongo.cosmos.azure.com:10255: <default>'.

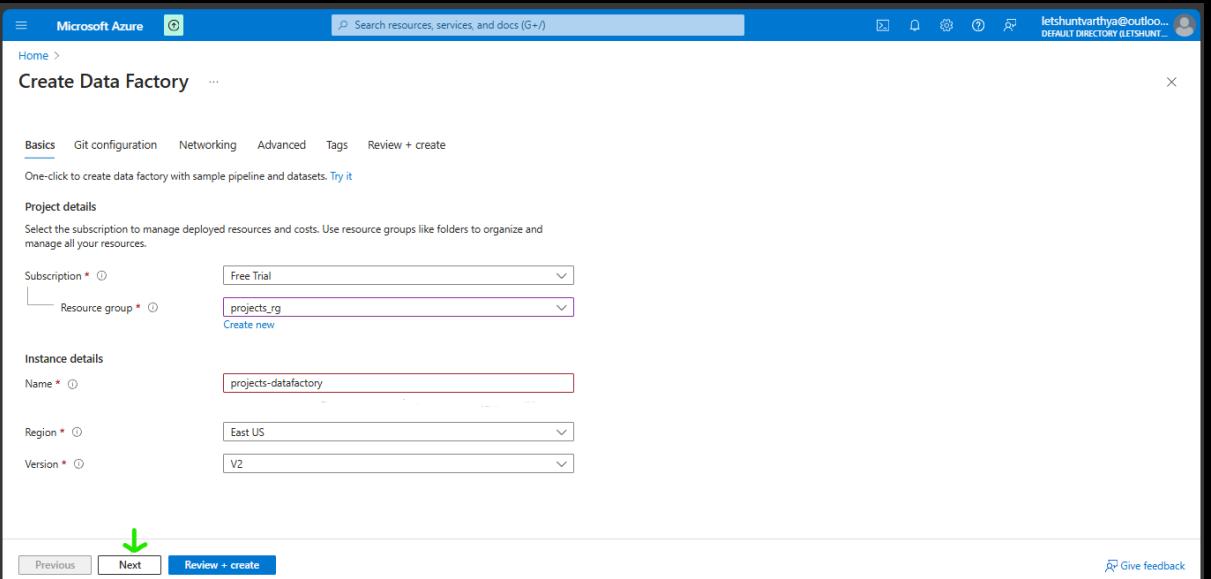
Azure Data Factory Creation:



The screenshot shows the Microsoft Azure portal home page. At the top, there's a search bar and several navigation icons. Below the search bar, there's a section titled "Azure services" with a "Create a resource" button highlighted by a green arrow. Other service icons include Resource groups, Cost Management, Subscriptions, Quickstart Center, Azure AI services, Kubernetes services, Virtual machines, App Services, and More services.



This screenshot shows the "Create a resource" dialog. On the left, there's a sidebar with categories like AI + Machine Learning, Analytics (highlighted with a green arrow), Blockchain, Compute, Containers, Databases, Developer Tools, DevOps, Identity, Integration, Internet of Things, IT & Management Tools, Media, Migration, and Mixed Reality. The main area shows "Popular Azure services" and "Popular Marketplace products". Under "Popular Azure services", the "Data Factory" option is highlighted with a green arrow. Other options include Azure Synapse Analytics, Azure Machine Learning, Event Hubs, Azure Databricks, Data Lake Storage Gen1, Event Grid System Topic, and Azure Data Explorer.



This screenshot shows the "Create Data Factory" wizard. The "Basics" tab is selected. It asks for a subscription (Free Trial) and a resource group (projects_rg). The "Name" field is set to "projects-datafactory". The "Region" is set to "East US" and the "Version" is set to "V2". At the bottom, there are "Previous" and "Next" buttons, and a "Review + create" button highlighted with a green arrow.

Create Data Factory

Git configuration

Azure Data Factory allows you to configure a Git repository with either Azure DevOps or GitHub. Git is a version control system that allows for easier change tracking and collaboration.

Learn more about Git integration in Azure Data Factory

Configure Git later

Previous  Next Review + create Give feedback

Create Data Factory

Networking

Choose whether you want the default AutoResolveIntegrationRuntime to be provisioned on demand inside an ADF-managed virtual network. If this setting is disabled, after the data factory is created, you can still choose whether to provision explicitly created Azure integration runtime inside an ADF-managed virtual network.

Learn more

Enable Managed Virtual Network on the default AutoResolveIntegrationRuntime

Self-hosted integration runtime inbound connectivity to Azure Data Factory service

Choose whether to connect your self-hosted integration runtime to Azure Data Factory via public endpoint or private endpoint. This applies to self-hosted integration runtime running either on premises or inside customer managed Azure virtual network.

Learn more

Connect via * Public endpoint Private endpoint

You can change this or configure another connectivity method after this resource is created. [Learn more](#)

Previous  Next Review + create Give feedback

Create Data Factory

Advanced

Datafactory Encryption

By default, data is encrypted with Microsoft-managed keys. For additional control over encryption keys, you can supply customer-managed keys to use for encryption of blob and file data. Customer-managed keys must be stored in an Azure Key Vault. You can either create your own keys and store them in a key vault, or you can use the Azure Key Vault APIs to generate keys. The storage account and the key vault must be in the same region, but they can be in different subscriptions.

Enable encryption using a Customer Managed Key

Previous  Next Review + create Give feedback

Create Data Factory

Basic Git configuration Networking Advanced Tags Review + create

Tags are name/value pairs that enable you to categorize resources and view consolidated billing by applying the same tag to multiple resources and resource groups. [Learn more about tags](#)

Note that if you create tags and then change resource settings on other tabs, your tags will be automatically updated.

Name	Value	Resource
		Data factory (V2)

Previous Next Review + create Give feedback

Microsoft.DataFactory-20240524172645 | Overview

Deployment

Search Delete Cancel Redeploy Download Refresh

Your deployment is complete

Deployment name : Microsoft.DataFactory-20240524172645
Subscription : Free Trial
Resource group : projects_rg

Start time : 5/24/2024, 5:32:31 PM
Correlation ID : 4d98a809-50e8-4449-bb7a-6cdcbcb26e

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 >](#)

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 Status Succeeded Location East US Subscription (move) Free Trial Subscriptions 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 'denissadilisa_1716546997551' to resource group 'projects_rg' was successful.
an hour ago

Microsoft Azure | Data Factory > projects-datafactory

Linked services

Linked service defines the connection information to a data store or compute. Learn more [\[link\]](#)

+ New **2**

Filter by name Annotations : Any

Showing 1 - 1 of 1 items

Name	Type
[REDACTED]	

New linked service

Data store Compute

Search **3**

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

4

Continue Cancel

Microsoft Azure | Data Factory > projects-datafactory

Linked services

Linked service defines the connection information to a data store or compute **1**

+ New

Filter by name Annotations : Any

Showing 1 - 4 of 4 items

Name	Type
[REDACTED]	
[REDACTED]	
[REDACTED]	
[REDACTED]	

New linked service

Azure Blob Storage [Learn more \[link\]](#)

Name * **blob2adf_ls**

Description

Connect via integration runtime * **AutoResolveIntegrationRuntime**

Authentication type **Account key**

Connection string Azure Key Vault

Account selection method **From Azure subscription**

Azure subscription **Free Trial (c92e5287-784e-4a20-9e6c-8549947f2ee6)**

Storage account name * **blob2adf_ls**

Create Back **5** **4** **Activate Windows**

Microsoft Azure | Data Factory > projects-datafactory

Linked services

Linked service defines the connection information to a data store or compute. Learn more [\[link\]](#)

+ New **1**

Filter by name Annotations : Any

Showing 1 - 2 of 2 items

Name	Type
blob2adf_ls	Azure Blob Storage

New linked service

Data store Compute

Search **2**

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

3

Continue Cancel

Microsoft Azure | Data Factory > projects-datafactory

Linked services

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

+ New

Filter by name Annotations : Any

Show 1 - 2 of 2 items

Name	Type
blob2adf_ls	Azure Blob Storage

New linked service

Azure Data Lake Storage Gen2 [Learn more](#)

Name * adls2adf_ls

Description

Connect via integration runtime * [AutoResolveIntegrationRuntime](#)

Authentication type Account key

Account selection method From Azure subscription Enter manually

Azure subscription * Free Trial (c92e5287-784e-4a20-9e6c-8549947f2ee6)

Storage account name * denissadllsa

Test connection To linked service To file path

Create Back **Test connection** **Cancel**

Microsoft Azure | Data Factory > projects-datafactory

Linked services

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

+ New

Filter by name Annotations : Any

Show 1 - 3 of 3 items

Name	Type
adls2adf_ls	Azure Data Lake Storage Gen2
blob2adf_ls	Azure Blob Storage

New linked service

Data store Compute

Search

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
My	My	My

Continue Cancel

Microsoft Azure | Data Factory > projects-datafactory

Linked services

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

+ New

Filter by name Annotations : Any

Show 1 - 3 of 3 items

Name	Type
adls2adf_ls	Azure Data Lake Storage Gen2
blob2adf_ls	Azure Blob Storage

New linked service

Azure Cosmos DB for MongoDB [Learn more](#)

Name * cosmos2adf_ls

Description

Connect via integration runtime * [AutoResolveIntegrationRuntime](#)

Connection string Azure Key Vault

Account selection method From Azure subscription Enter manually

Azure subscription * Free Trial (c92e5287-784e-4a20-9e6c-8549947f2ee6)

Azure Cosmos DB account name * denisscosmossa

Database name * GeographyDB

Annotations

Create Back **Test connection** **Cancel**

Microsoft Azure | Data Factory > projects-datafactory

Linked services

New linked service

Data store **Compute**

Search **Annotations: Any**

All **Azure** **Database** **File** **Generic protocol** **NoSQL** **Services and apps**

IBM **Informix** **MariaDB** **Microsoft Access**

MySQL **MySQL** **Netezza** **Oracle**

MySQL **Netezza** **Oracle**

MySQL **Netezza** **Oracle**

Continue

Microsoft Azure | Data Factory > projects-datafactory

Linked services

New linked service

Name * **MySQL11**

Description

Connect via integration runtime * **AutoResolveIntegrationRuntime**

+ New

AutoResolveIntegrationRuntime

Port **3306**

Database name *

User name *

Create **Back** **Test connection** **Cancel**

Microsoft Azure | Data Factory > projects-datafactory

Linked services

Integration runtime setup

Network environment:

Choose the network environment of the data source / destination or external compute to which the integration runtime will connect to for data flows, data movement or dispatch activities:

Azure
Use this for running data flows, data movement, external and pipeline activities in a fully managed, serverless compute in Azure.

Self-Hosted
Use this for running activities in an on-premises / private network
View more

External Resources:

You can use an existing self-hosted integration runtime that exists in another resource. This way you can reuse your existing infrastructure where self-hosted integration runtime is setup.

Linked Self-Hosted
Learn more

Continue

Integration runtime setup

Install integration runtime on Windows machine or add further nodes using the Authentication Key.

Name: self-hostedIR

Option 1: Express setup
Click here to launch the express setup for this computer

Option 2: Manual setup
Step 1: Download and install integration runtime

Step 2: Use this key to register your integration runtime
Name: Authentication key

Key1: IR@523aa7c2-86f8-424d-b3f6-d500d5950763@projects-datafactory@
Key2: IR@523aa7c2-86f8-424d-b3f6-d500d5950763@projects-datafactory@

Close

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

Register Integration Runtime (Self-hosted)

Welcome to Microsoft Integration Runtime Configuration Manager. Before you start, register your Integration Runtime (Self-hosted) node using a valid Authentication Key.

Authentication Key: **Show Authentication Key** **Learn how to find the Authentication Key**

HTTP Proxy
Current Proxy: No proxy Change

Register **Cancel**

Register Integration Runtime (Self-hosted)

Welcome to Microsoft Integration Runtime Configuration Manager. Before you start, register your Integration Runtime (Self-hosted) node using a valid Authentication Key.

Authentication Key: **Show Authentication Key** **Learn how to find the Authentication Key**

HTTP Proxy
Current Proxy: No proxy Change

✓ Integration Runtime (Self-hosted) node has been registered successfully.

Note: You can associate up to 4 physical nodes with a Self-hosted Integration Runtime. This enables high availability and scalability for the Self-hosted Integration Runtime.
We recommend you setup at least 2 nodes for higher availability. [See Integration Runtime \(Self-hosted\) article for details.](#)

Launch Configuration Manager **Close**

The screenshot shows the Microsoft Azure Data Factory interface. On the left, the navigation menu includes General, Connections (selected), 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 endpoints, Workflow orchestration manager, and Apache Airflow.

In the center, the 'Linked services' section lists three existing items: adls2adfls (Azure Data Lake Storage Gen2), blob2adfls (Azure Blob Storage), and cosmos2adfls (Azure Cosmos DB for MongoDB). A 'New' button is visible.

The right panel displays the 'New linked service' configuration for MySQL. The steps are numbered 1 through 6:

- Step 1: Set 'Name' to MySQL.
- Step 2: Set 'Server name' to localhost.
- Step 3: Set 'Database name' to denis.
- Step 4: Set 'User name' to root.
- Step 5: Set 'Password' to a value (the field is redacted).
- Step 6: Click 'Create'.

A green arrow points from the 'Server name' field to step 2. Another green arrow points from the 'User name' field to step 3. A third green arrow points from the 'Password' field to step 5. A green arrow points from the 'Create' button to step 6.

- Creating Datasets

The screenshot shows the 'Factory Resources' page. The left sidebar lists Pipelines, Change Data Capture (preview), Datasets (selected), Data flows, and Power Query. A green arrow points to the 'Datasets' item.

The screenshot shows the 'Datasets' page. The left sidebar lists Pipelines, Change Data Capture (preview), Datasets (selected), Data flows, and Power Query. A context menu is open over the 'Datasets' item, with options 'New dataset' and 'New folder'. A green arrow points to the 'New dataset' option.

Microsoft Azure | Data Factory | projects-datafactory | Search factory and documentation | Validate all | Publish all | letshunvarthya@outlook.com | DEFAULT DIRECTORY

Factory Resources | Filter resources by name | +

- Pipelines 0
- Change Data Capture (preview) 0
- Datasets 0
- Data flows 0
- Power Query 0

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

Search | 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

Use this dataset in a pipeline activity | [Activate Windows](#) | Go to Settings to activate Windows | Cancel

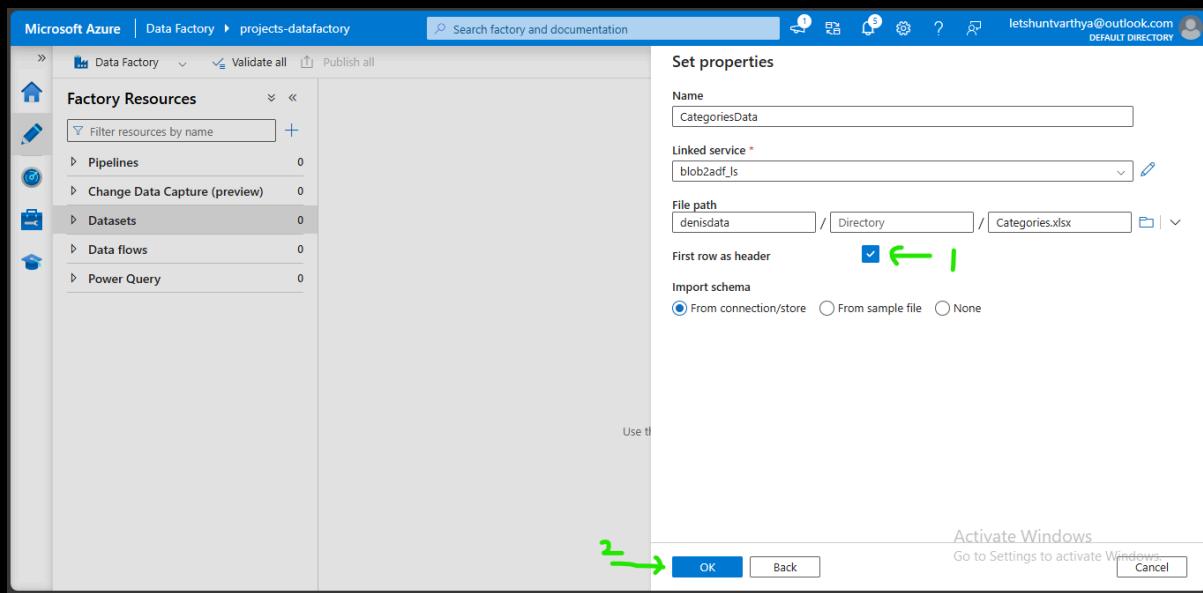
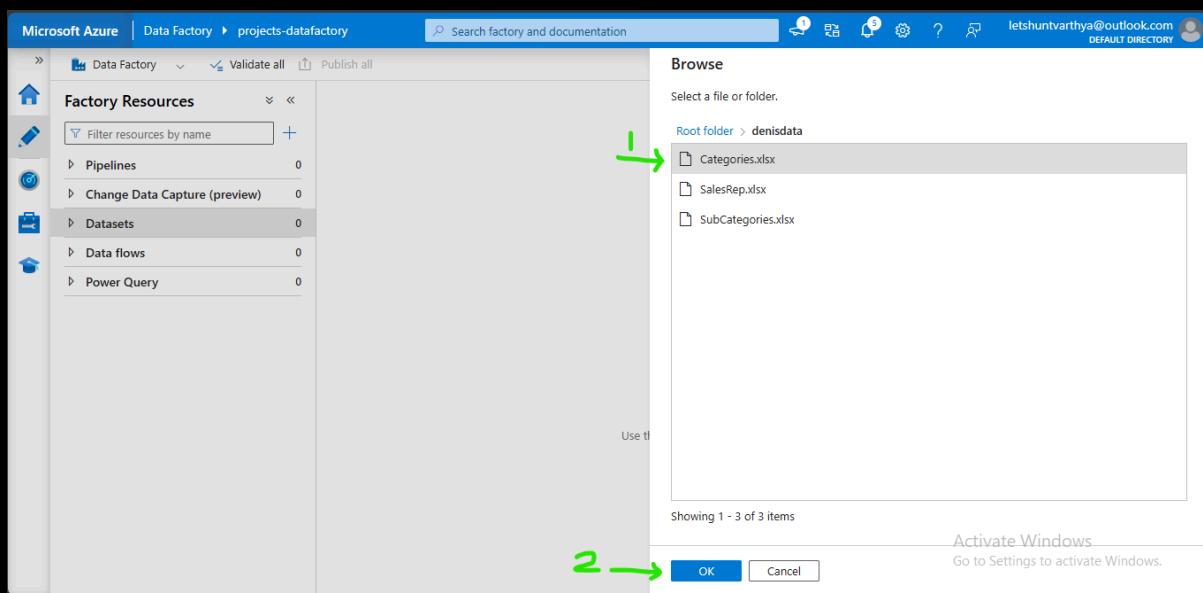
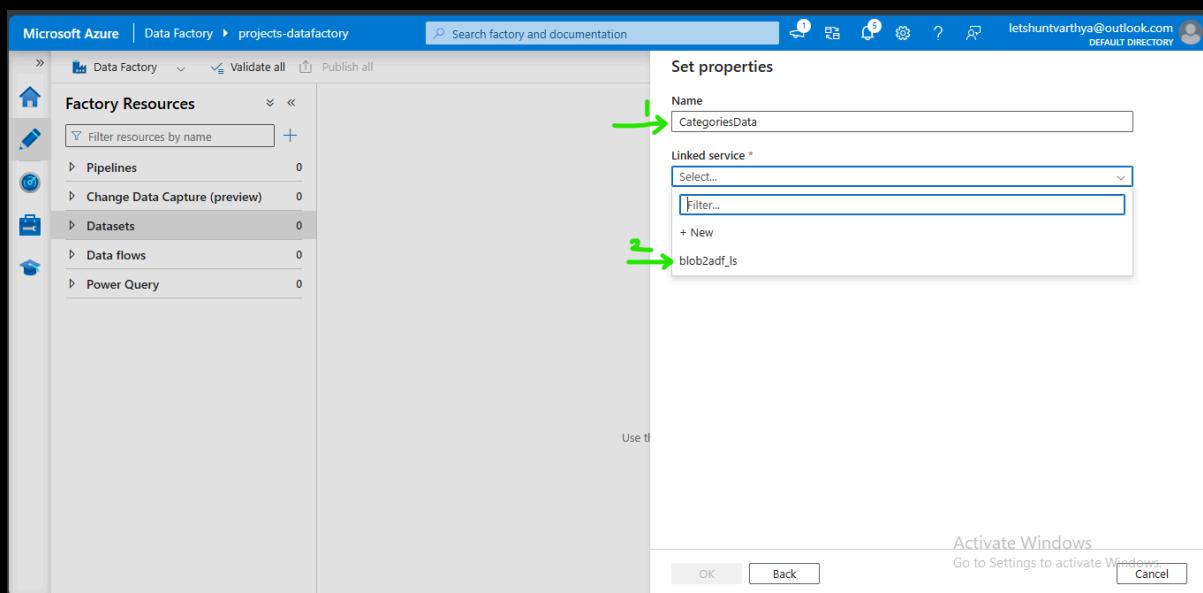
3 → Continue

Select format

Choose the format type of your data

 Avro	 Binary	 DelimitedText
 Excel	 JSON	 ORC
		Activate Windows Go to Settings to activate Windows. Cancel

2 ↓ Continue | Back



Microsoft Azure | Data Factory | projects-datafactory

Search factory and documentation

Preview experience: Off

letshunvarthy@outlook.com DEFAULT DIRECTORY

Factory Resources

- Pipelines: 3
- Change Data Capture (preview): 0
- Datasets:** 8
 - CategoriesData
 - GeographyData
 - ProductData
 - RawGeographyData
 - RawProductData
 - SalesRepData
 - StagingProductData
 - SubCategoriesData
- Data flows: 1
- Power Query: 0

CategoriesData (Excel)

Properties

General Related

Name: CategoriesData

Description:

Annotations: + New

Connection Schema Parameters

Linked service: blob2adf_ls

Test connection: Edit + New Learn more

File path: denisdata / Directory / Categories.xlsx

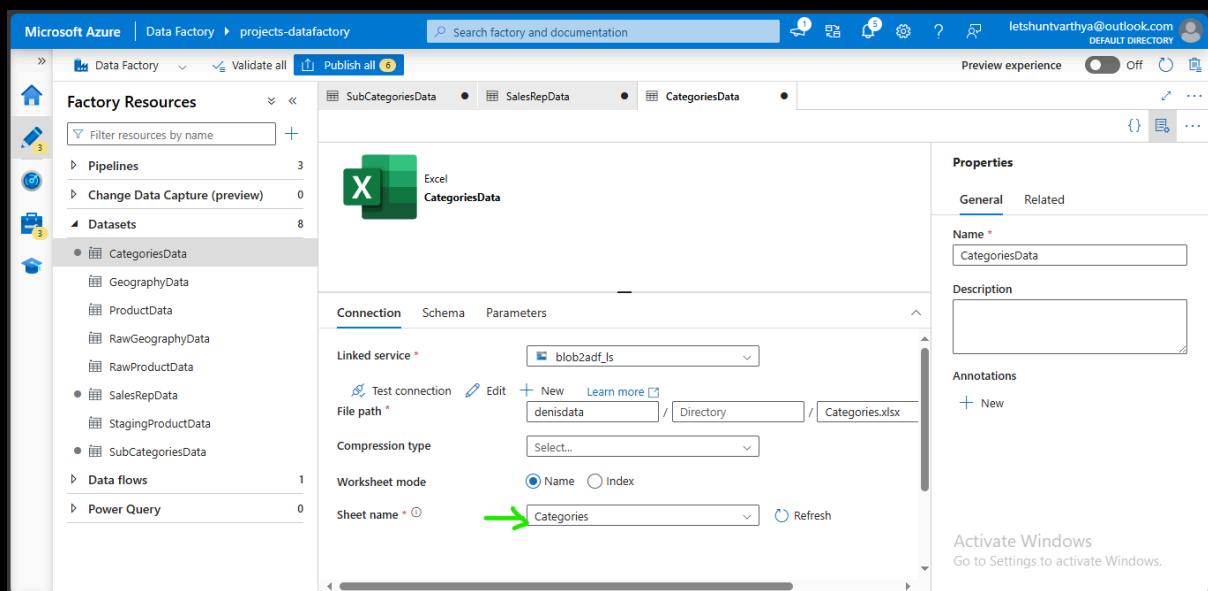
Compression type: Select...

Worksheet mode: Name (radio button selected) Index

Sheet name: Categories (highlighted with a green arrow)

Refresh

Activate Windows: Go to Settings to activate Windows.



Microsoft Azure | Data Factory | projects-datafactory

Search factory and documentation

letshunvarthy@outlook.com DEFAULT DIRECTORY

Factory Resources

- Pipelines: 0
- Change Data Capture (preview): 0
- Datasets:** 1
 - CategoriesData
- Data flows: 0
- Power Query: 0

CategoriesData (DelimitedText)

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

Search: (highlighted with a green arrow)

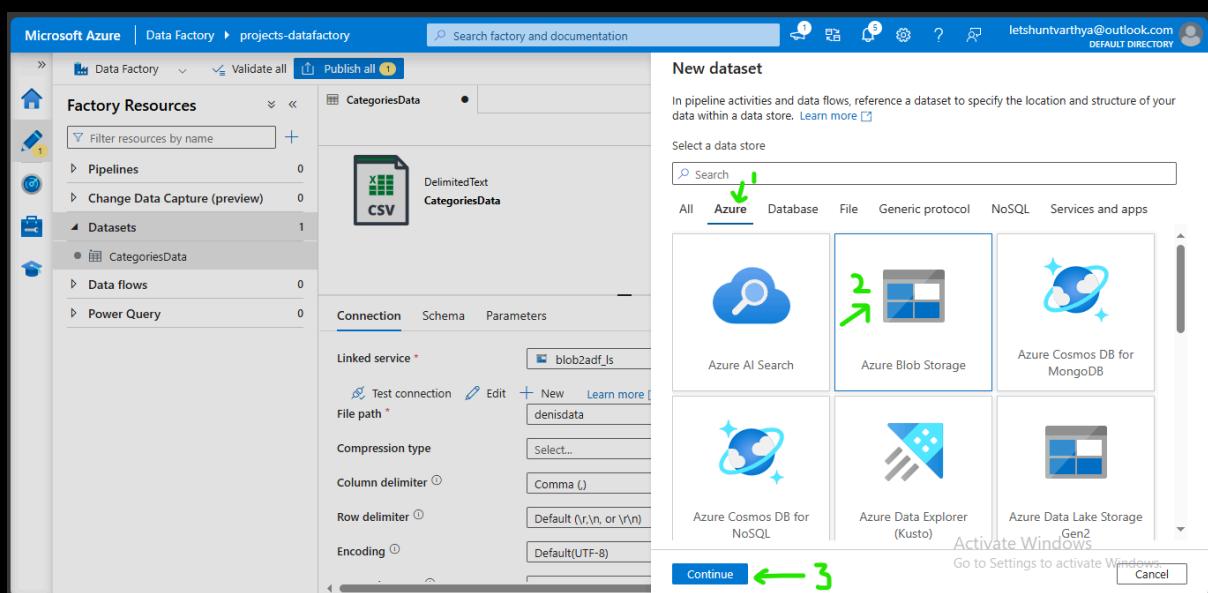
All Azure Database File Generic protocol NoSQL Services and apps

Azure AI Search	Azure Blob Storage (highlighted with a green arrow)	Azure Cosmos DB for MongoDB
Azure Cosmos DB for NoSQL	Azure Data Explorer (Kusto)	Azure Data Lake Storage Gen2

Activate Windows: Go to Settings to activate Windows.

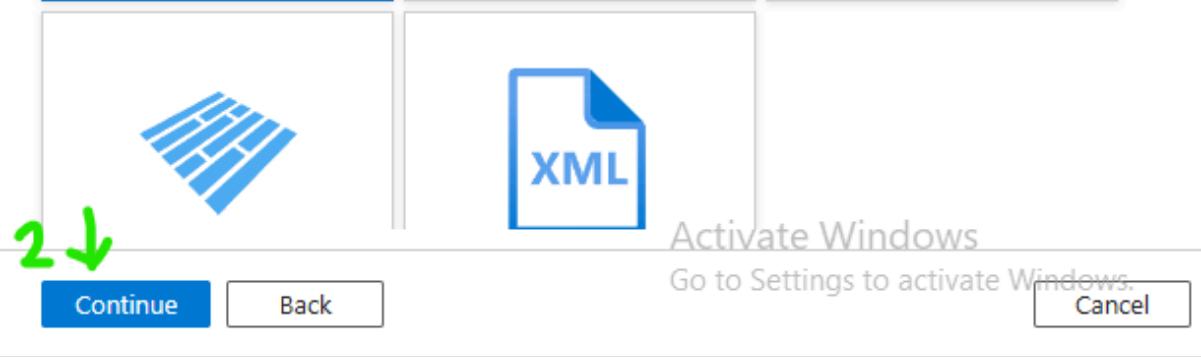
Continue (highlighted with a green arrow)

Cancel



Select format

Choose the format type of your data



Microsoft Azure | Data Factory > projects-datafactory

Set properties

Name: SalesRepData

Linked service: Select... (blob2adf_ls)

File path: denisdata

Compression type: Select...

Column delimiter: Comma (,)

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

Encoding: Default(UTF-8)

OK Back

Activate Windows
Go to Settings to activate Windows
Cancel

Factory Resources

- Pipelines: 0
- Change Data Capture (preview): 0
- Datasets: 1
 - CategoriesData
- Data flows: 0
- Power Query: 0

CategoriesData

DelimitedText CategoriesData

Connection Schema Parameters

Linked service: blob2adf_ls

Test connection Edit New Learn more

File path: denisdata

Compression type: Select...

Column delimiter: Comma (,)

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

Encoding: Default(UTF-8)

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Search factory and documentation

Factory Resources

CategoriesData

CSV

DelimitedText CategoriesData

Set properties

Name: SalesRepData

Linked service: blob2adf_ls

File path: Container / Directory / File name

First row as header:

Import schema: From connection/store, From sample file, None

Connection Schema Parameters

Linked service: blob2adf_ls

Test connection, Edit, New, Learn more

File path: denisdata

Compression type: Select...

Column delimiter: Comma (,)

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

Encoding: Default(UTF-8)

OK Back

Activate Windows

Go to Settings to activate Windows

Cancel

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Search factory and documentation

Factory Resources

CategoriesData

CSV

DelimitedText CategoriesData

Browse

Select a file or folder.

Root folder: denisdata

Showing 1 item

OK Cancel

Activate Windows

Go to Settings to activate Windows.

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Search factory and documentation

Factory Resources

CategoriesData

CSV

DelimitedText CategoriesData

Browse

Select a file or folder.

Root folder: denisdata

Showing 1 - 3 of 3 items

1. Categories.xlsx

2. SalesRep.xlsx

3. SubCategories.xlsx

OK Cancel

Activate Windows

Go to Settings to activate Windows.

Microsoft Azure | Data Factory | projects-datafactory | Search factory and documentation | Validate all | Publish all 1

Factory Resources | CategoriesData | Set properties

Name: SalesRepData
Linked service: blob2adf_ls
File path: denisdata / Directory / SalesRep.xlsx
First row as header: 

Connection Schema Parameters

Linked service: blob2adf_ls
File path: denisdata
Compression type: Select...
Column delimiter: Comma (,)
Row delimiter: Default (\r\n, or \n)
Encoding: Default(UTF-8)

OK Back  

Microsoft Azure | Data Factory | projects-datafactory | Search factory and documentation | Validate all | Publish all 5

Factory Resources | SalesRepData | Properties

Name: SalesRepData
Description:
Annotations: + New

Connection Schema Parameters

Linked service: blob2adf_ls
File path: denisdata / Directory / SalesRep.xlsx
Compression type: Select...
Worksheet mode: Name (radio button selected)  Index
Sheet name: Sales rep  Refresh

Activate Windows Go to Settings to activate Windows. 

Microsoft Azure | Data Factory | projects-datafactory | Search factory and documentation | Validate all | Publish all 1

Factory Resources | CategoriesData | Properties

Name: CategoriesData
Description:
Annotations: + New

Connection Schema Parameters

Linked service: blob2adf_ls
File path: denisdata / Directory / Categories.xlsx
Compression type: Select...
Column delimiter: Comma (,)
Row delimiter: Default (\r\n, or \n)
Encoding: Default(UTF-8)

New dataset   New folder

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all

Search factory and documentation

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

CategoriesData SalesRepData

DelimitedText SalesRepData

Connection Schema Parameters

Linked service * blob2adf_ls

Test connection Edit New Learn more

File path * denisdata

Compression type Select...

Column delimiter Comma (,)

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

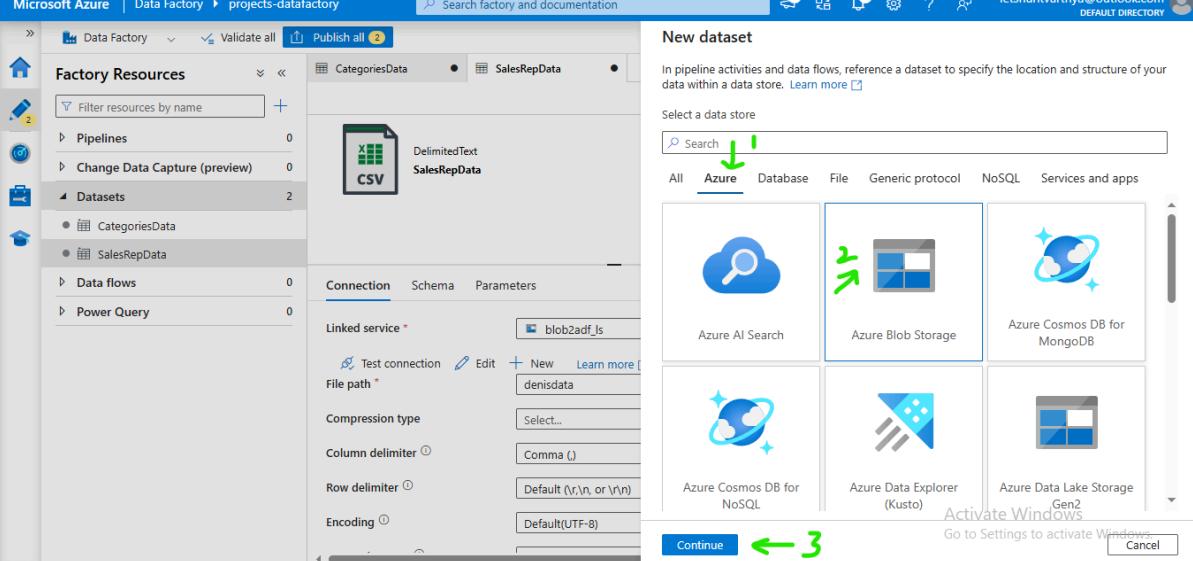
Encoding Default(UTF-8)

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

Activate Windows Go to Settings to activate Windows Cancel

Continue ← 3



Select format

Choose the format type of your data

Avro

Binary

DelimitedText

Excel

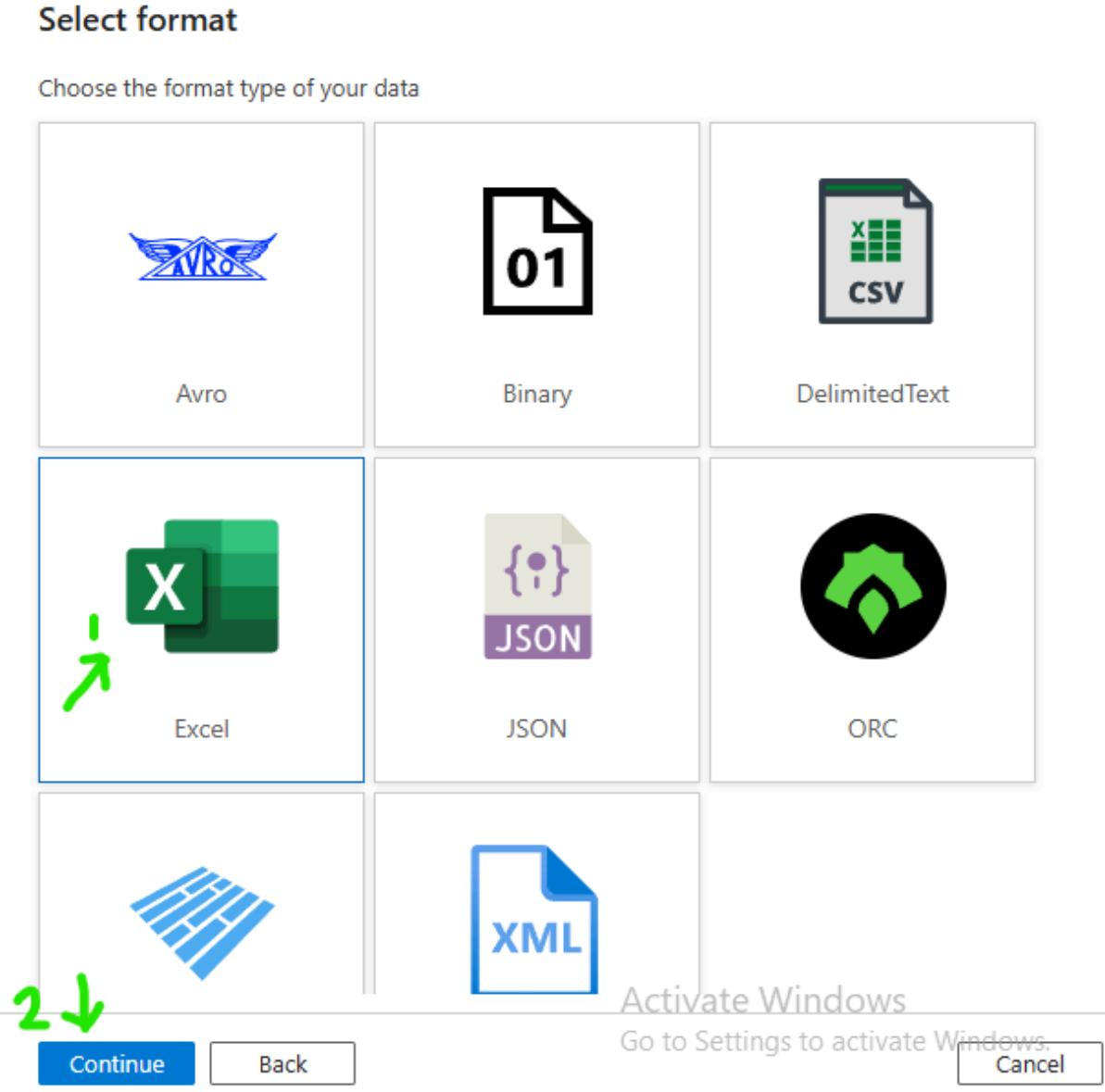
JSON

ORC

XML

Activate Windows
Go to Settings to activate Windows. Cancel

2 ↓ Continue Back



Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Validate all Publish all

Factory Resources

Datasets

- CategoriesData
- SalesRepData

Set properties

Name: SubCategoryData

Linked service: Select... Filter... + New blob2adf_ls

Connection Schema Parameters

Linked service: blob2adf_ls

File path: denisdata

Compression type: Select...

Column delimiter: Comma (,)

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

Encoding: Default(UTF-8)

OK Back Cancel

Activate Windows Go to Settings to activate Windows

File path

Container / Directory / File name

First row as header

Import schema

From connection/store From sample file None

Up

Browse

Select a file or folder.

Root folder

denisdata

Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Validate all Publish all

Factory Resources

Datasets

- CategoriesData
- SalesRepData

Pipelines Change Data Capture (preview)

Data flows Power Query

Browse

Select a file or folder.

Root folder > denisdata

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

Connection Schema Parameters

Linked service: blob2adf_ls

File path: denisdata

Compression type: Select...

Column delimiter: Comma (,)

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

Encoding: Default(UTF-8)

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

Validate all Publish all

Factory Resources

Datasets

- CategoriesData
- SalesRepData

Pipelines Change Data Capture (preview)

Data flows Power Query

Set properties

Name: SubCategoryData

Linked service: blob2adf_ls

File path: denisdata / Directory / SubCategories.xlsx

First row as header:

Import schema: From connection/store

OK Back

Activate Windows Go to Settings to activate Windows.

Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Validate all Publish all

Factory Resources

Datasets

- GeographyData
- ProductData
- RawGeographyData
- RawProductData
- SalesRepData
- StagingProductData
- SubCategoriesData

Pipelines Change Data Capture (preview)

Data flows Power Query

Properties

General Related

Name: SubCategoriesData

Description:

Annotations + New

Preview experience: Off

SubCategoriesData

Excel SubCategoriesData

Connection Schema Parameters

Linked service: blob2adf_ls

File path: denisdata / Directory / SubCategories.xlsx

Compression type: Select...

Worksheet mode: Name (radio button selected) Index

Sheet name: Sub categories

Refresh

Activate Windows Go to Settings to activate Windows.

The image consists of three vertically stacked screenshots of the Microsoft Azure Data Factory interface, illustrating the publishing process for datasets.

Screenshot 1: Shows the 'Factory Resources' blade with the 'Datasets' section selected. A green arrow points to the 'Publish all' button in the top right corner of the main content area.

Screenshot 2: Shows the 'Publishing' blade. It displays pending changes for three datasets: 'CategoriesData', 'SalesRepData', and 'SubCategoryData'. A green arrow points to the 'Publish' button at the bottom right of the blade.

Screenshot 3: Shows the 'Factory Resources' blade again, but now with the 'Datasets' section expanded. A green arrow points to the 'New dataset' link under the 'SubCategoryData' dataset entry.

Microsoft Azure | Data Factory | projects-datafactory

Search factory and documentation

Validate all Publish all

Factory Resources

Datasets

CategoriesData SalesRepData

DelimitedText SubCategoryData

Connection Schema Parameters

Linked service * blob2adf_ls

Test connection Edit New Learn more

File path * denisdata

Compression type Select...

Column delimiter Comma (,)

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

Encoding Default(UTF-8)

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

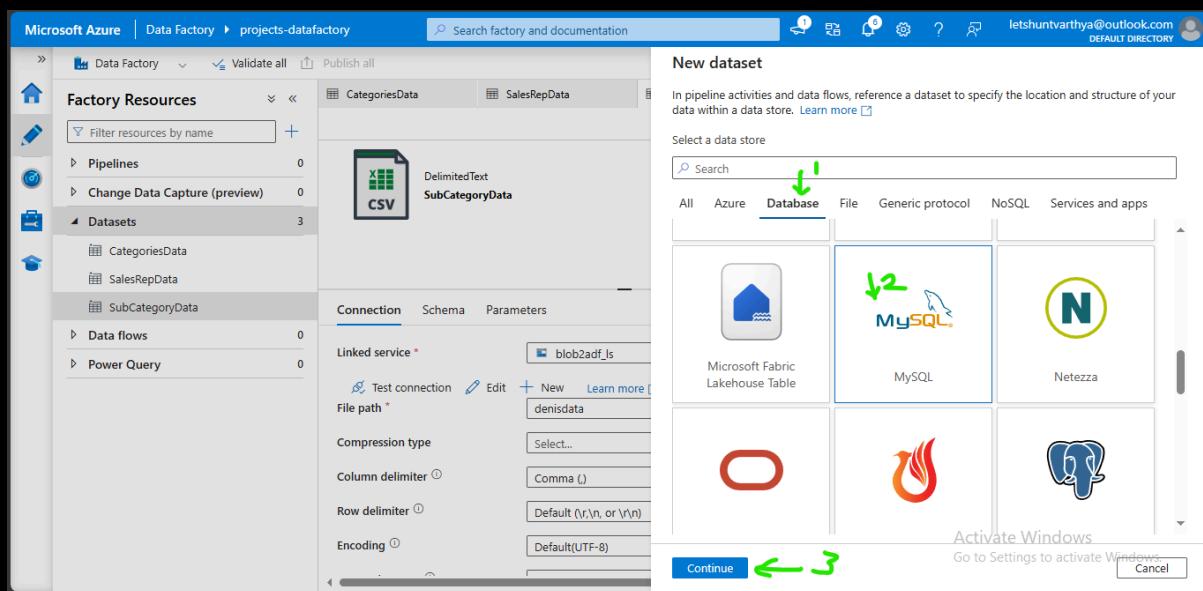
All Azure Database File Generic protocol NoSQL Services and apps

MySQL

Microsoft Fabric Lakehouse Table MySQL Netezza

Activate Windows Go to Settings to activate Windows Cancel

Continue



Microsoft Azure | Data Factory | projects-datafactory

Search factory and documentation

Validate all Publish all

Factory Resources

Datasets

CategoriesData SalesRepData

DelimitedText SubCategoryData

Connection Schema Parameters

Linked service * Select...

+ New Filter... MySQL1

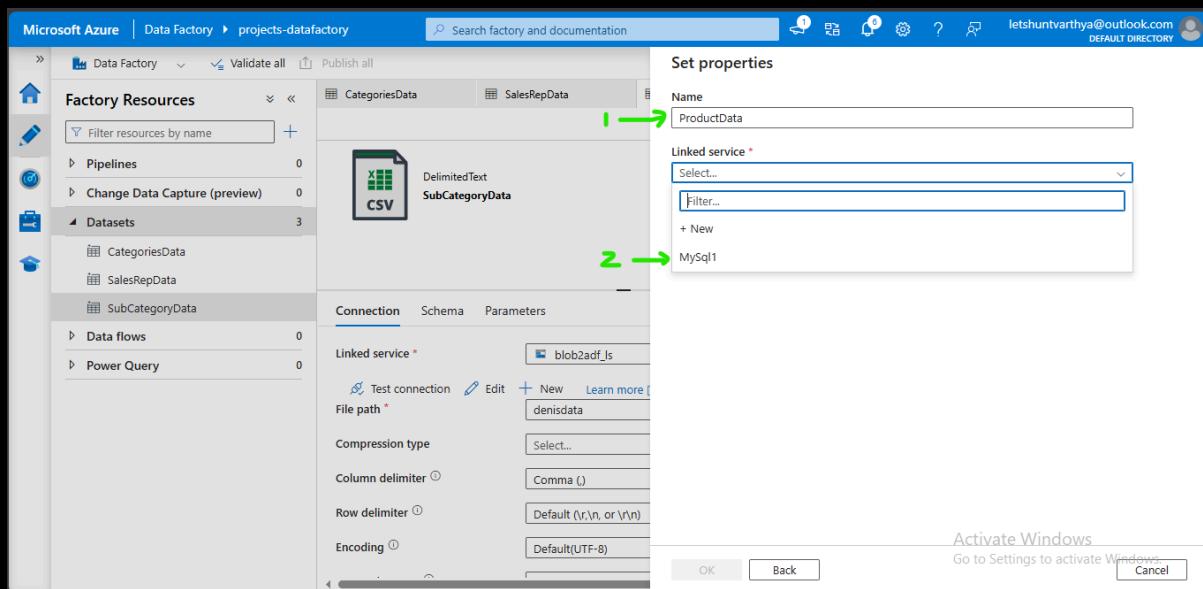
Set properties

Name ProductData

Linked service * Select...

+ New Filter... MySQL1

Activate Windows Go to Settings to activate Windows Cancel



Microsoft Azure | Data Factory | projects-datafactory

Search factory and documentation

Validate all Publish all

Factory Resources

Datasets

CategoriesData SalesRepData

DelimitedText SubCategoryData

Connection Schema Parameters

Linked service * MySQL1

Connect via integration runtime * selfhosted

Table name product (product)

Filter...

product (product)

Set properties

Name ProductData

Linked service * MySQL1

Connect via integration runtime * selfhosted

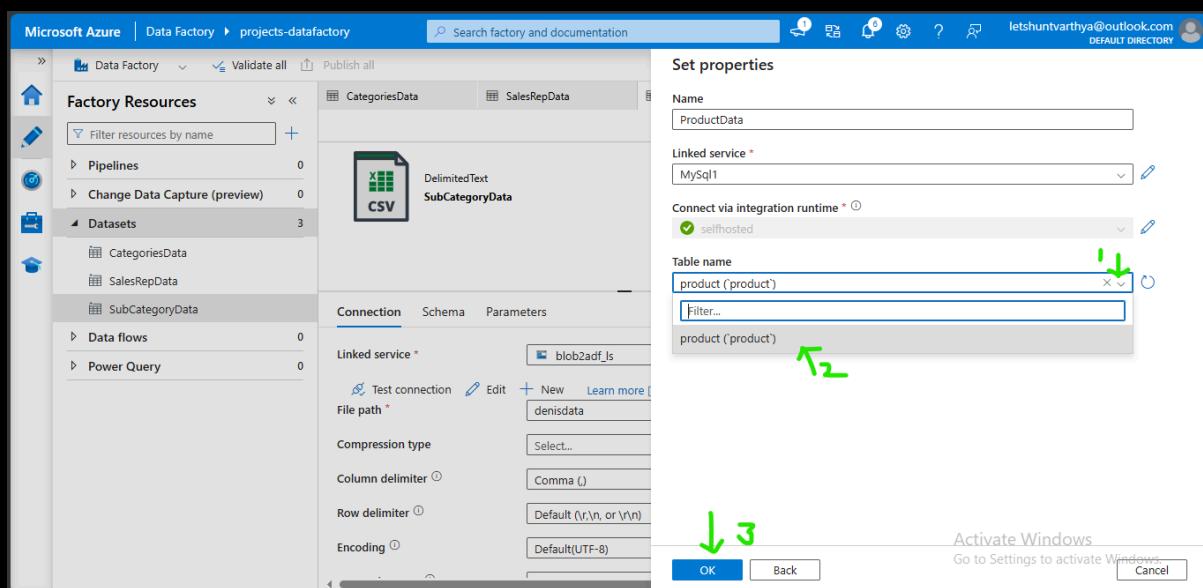
Table name product (product)

Filter...

product (product)

Activate Windows Go to Settings to activate Windows Cancel

OK Back



Microsoft Azure | Data Factory | projects-datafactory

Search factory and documentation

Preview experience: Off

letshunvarthy@outlook.com DEFAULT DIRECTORY

Factory Resources

- Pipelines: 0
- Change Data Capture (preview): 0
- Datasets**: 3
- CategoriesData
- SalesRepData
- New dataset** ← 1
- New folder

Properties

General Related

Name: SubCategoryData

Description:

Annotations:

Microsoft Azure | Data Factory | projects-datafactory

Search factory and documentation

letshunvarthy@outlook.com DEFAULT DIRECTORY

Factory Resources

- Pipelines: 0
- Change Data Capture (preview): 0
- Datasets**: 4
- CategoriesData
- ProductData
- SalesRepData
- SubCategoryData

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

Search: →

All Azure Database File Generic protocol NoSQL Services and apps

Azure AI Search	Azure Blob Storage	Azure Cosmos DB for MongoDB ← 2
Azure Cosmos DB for NoSQL	Azure Data Explorer (Kusto)	Azure Data Lake Storage Gen2

Continue ← 3

Activate Windows: Go to Settings to activate Windows. Cancel

Microsoft Azure | Data Factory | projects-datafactory

Search factory and documentation

letshunvarthy@outlook.com DEFAULT DIRECTORY

Factory Resources

- Pipelines: 0
- Change Data Capture (preview): 0
- Datasets**: 4
- CategoriesData
- ProductData
- SalesRepData
- SubCategoryData

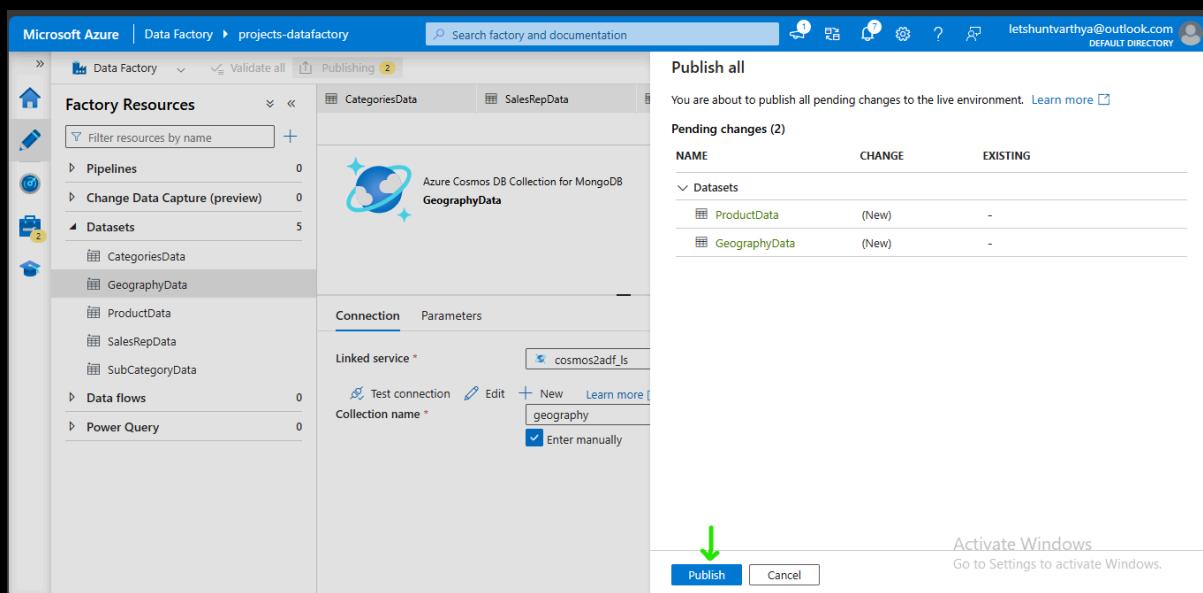
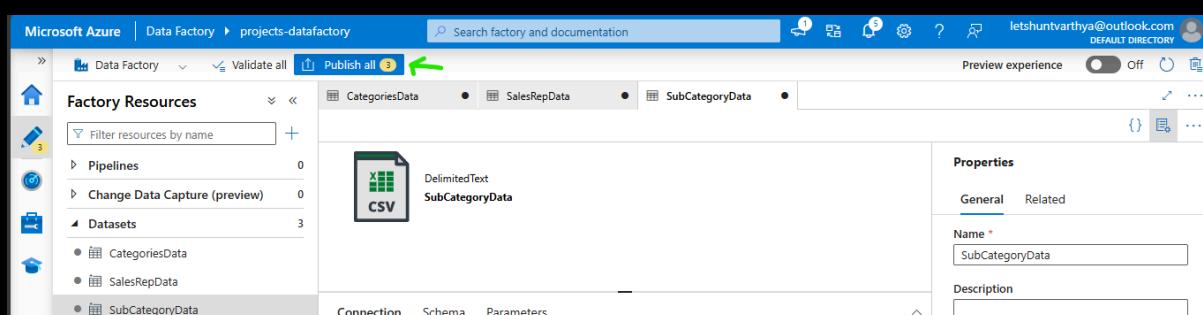
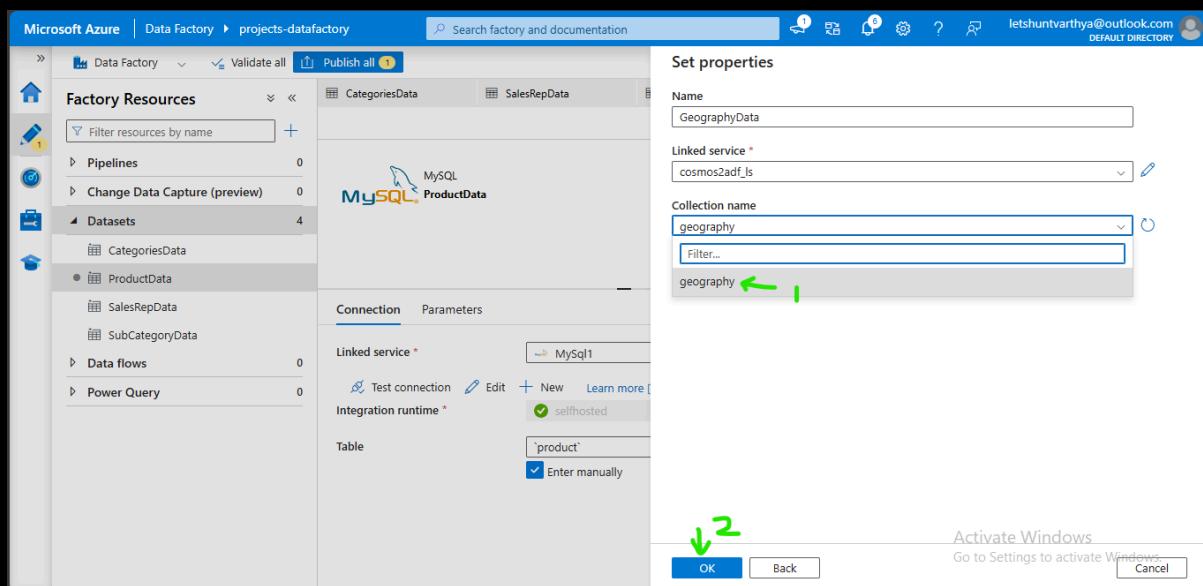
Set properties

Name: GeographyData ← 1

Linked service: Select... ← 2

cosmos2adf_ls ← 3

Activate Windows: Go to Settings to activate Windows. Cancel



- We need to do small **transformations** with this **dataset** like **removing duplicates**, **null values** etc. This data should be **loaded** into a **staging layer** and for that we need to create a **container** called **stagingdenisdata** in **denissblobsa** and **container** called **stagingdenisfactdata** in **denissadllssa**.

Microsoft Azure | denissblobsa | Containers

New container

Name * 2

Anonymous access level

Advanced

Create 2 Go to Settings to activate Windows Give feedback

Search resources, services, and docs (G+/)

Home > denissblobsa

denissblobsa | Containers

Storage account

Overview Activity log Tags Diagnose and solve problems Access Control (IAM) Data migration Events Storage browser Storage Mover Data storage Containers File shares Queues

Containers

Name Last modified

Slogs	5/24/2024, 3:39:05 PM
denisdata	5/24/2024, 3:45:03 PM

Search containers by prefix

Microsoft Azure | denissadllssa | Containers

New container

Name * 2

Anonymous access level

Advanced

Create 2 Go to Settings to activate Windows Give feedback

Search resources, services, and docs (G+/)

Home > denissadllssa

denissadllssa | Containers

Storage account

Overview Activity log Tags Diagnose and solve problems Access Control (IAM) Data migration Events Storage browser Data storage Containers File shares Queues Tables

Containers

Name Last modified

Slogs	5/24/2024, 4:08:05 PM
denisfactdata	5/24/2024, 4:23:00 PM

Search containers by prefix

- Migration of product data to **denisdata**.

Microsoft Azure | Data Factory | projects-datafactory

Factory Resources

Pipelines 2

Change Data Capture (preview)

Datasets

- CategoriesData
- SalesRepData
- SubCategoryData
- ProductData
- GeographyData

Data flows

Power Query

CategoriesData

SalesRepData

SubCategoryData

ProductData

GeographyData

Connection Parameters

Linked service * 2

Test connection 2

Collection name * 2

Enter manually

Properties

General Related

Name *

Description

Annotations

Activate Windows Go to Settings to activate Windows.

Microsoft Azure | Data Factory > projects-datafactory

Activities

Copy data

Properties

Name: onPremsql2blob_PL

Description:

Activity state: Activated

General Source Sink Mapping Settings User properties

Copy data1

Microsoft Azure | Data Factory > projects-datafactory

Activities

Copy data

Properties

Name: onPremsql2blob_PL

Description:

Activity state: Activated

General Source Sink Mapping Settings User properties

Copy data1

Source dataset:

- Select...
- Filter...
- CategoriesData
- GeographyData
- ProductData
- SalesRepData
- SubCategoryData

General Source Sink¹ 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

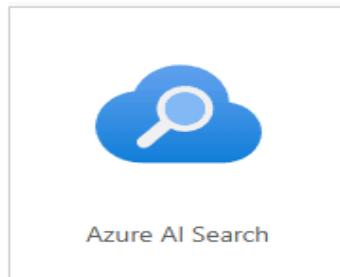
Database

File

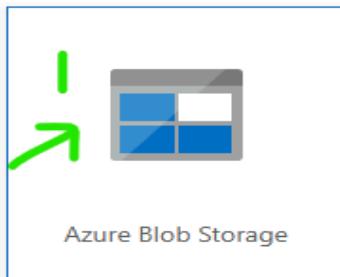
Generic protocol

NoSQL

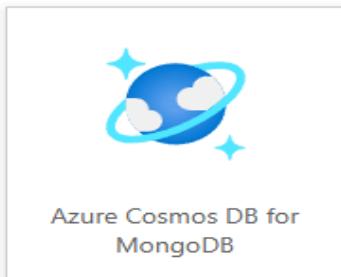
Services and apps



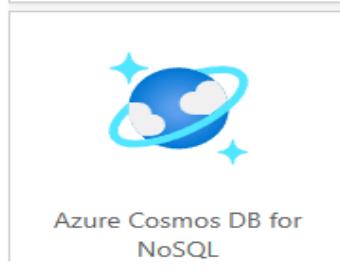
Azure AI Search



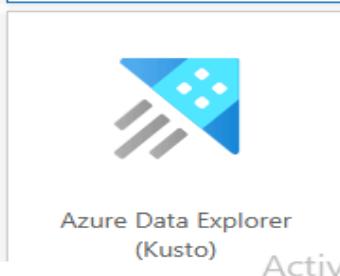
Azure Blob Storage



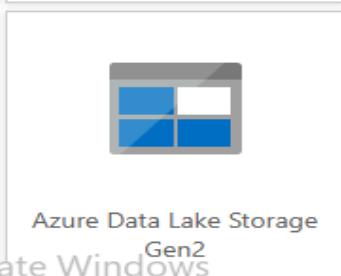
Azure Cosmos DB for MongoDB



Azure Cosmos DB for NoSQL



Azure Data Explorer (Kusto)



Azure Data Lake Storage Gen2

Activate Windows

Go to Settings to activate Windows.

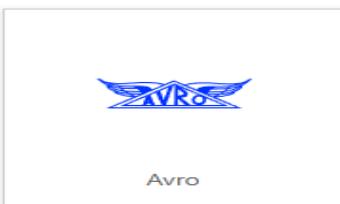
[Cancel](#)

[Continue](#)

2

Select format

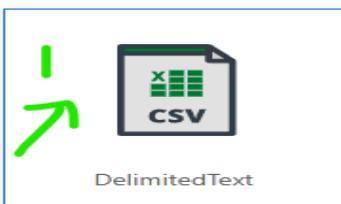
Choose the format type of your data



Avro



Binary



DelimitedText



JSON



ORC



Parquet

[Continue](#)

2

Activate Windows

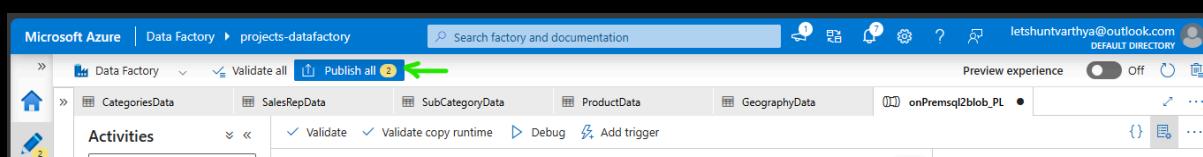
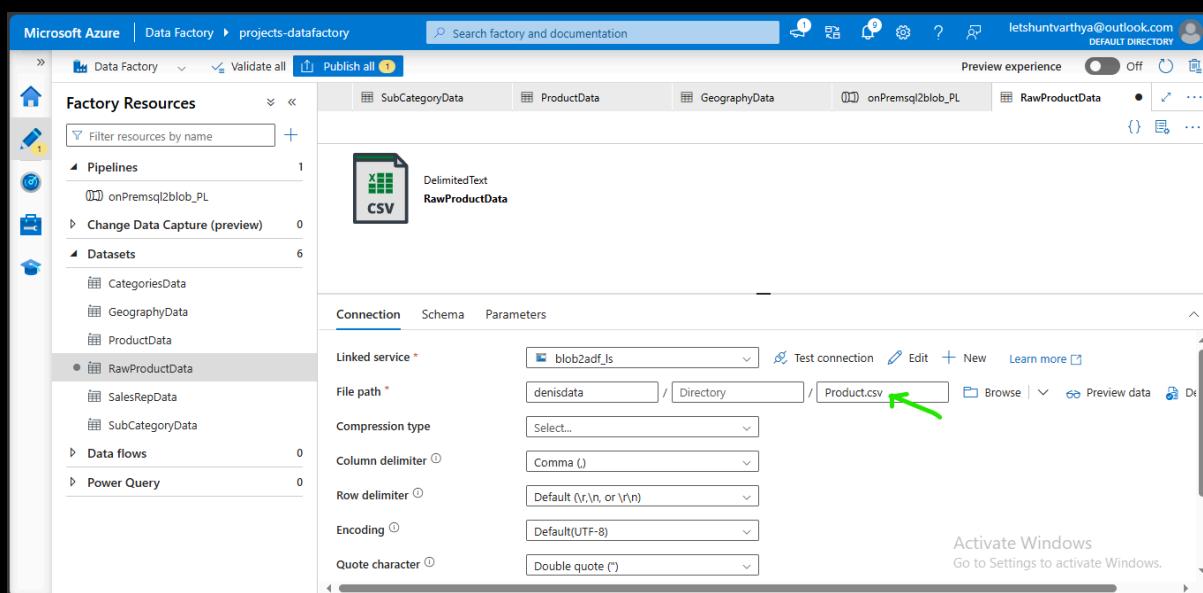
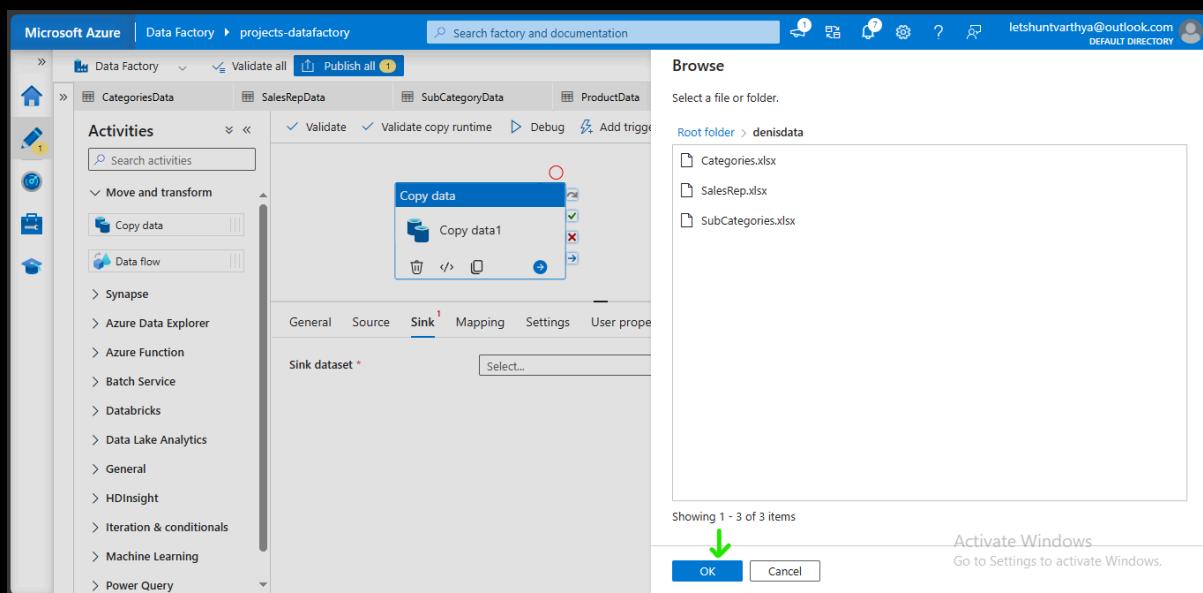
Go to Settings to activate Windows.

[Cancel](#)

The screenshot shows the Microsoft Azure Data Factory interface. A 'Copy data' activity is selected in the center pane. The 'Sink' tab is active. In the 'Set properties' dialog, the 'Name' field is set to 'RawProductData'. The 'Linked service' dropdown is open, showing a list of available services. One item, 'blob2adf_ls', is highlighted with a green arrow labeled '3'. The bottom right corner of the dialog has a watermark: 'Activate Windows. Go to Settings to activate Windows. Cancel'.

This screenshot shows the 'File path' configuration dialog. It includes fields for 'Container', 'Directory', and 'File name'. A 'First row as header' checkbox is checked. Below it, there's an 'Import schema' section with three radio button options: 'From connection/store', 'From sample file', and 'None', where 'None' is selected. On the right side, there are standard file selection icons (Browse, Open, Save).

The screenshot shows a 'Browse' dialog for selecting a file or folder. The 'Root folder' section displays two items: 'denisdata' and 'stagingdenisdata'. A large green arrow points to the 'denisdata' folder. To the right of the folder list, there are standard file selection icons (Browse, Open, Save).



Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Publish all

You are about to publish all pending changes to the live environment. [Learn more](#)

Pending changes (2)

NAME	CHANGE	EXISTING
Pipelines		
onPremsql2blob_PL	(New)	-
Datasets		
RawProductData	(New)	-

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 data1

General Source Sink Mapping Settings User properties

Sink dataset * RawProductData

Copy behavior Select...

Max concurrent connections

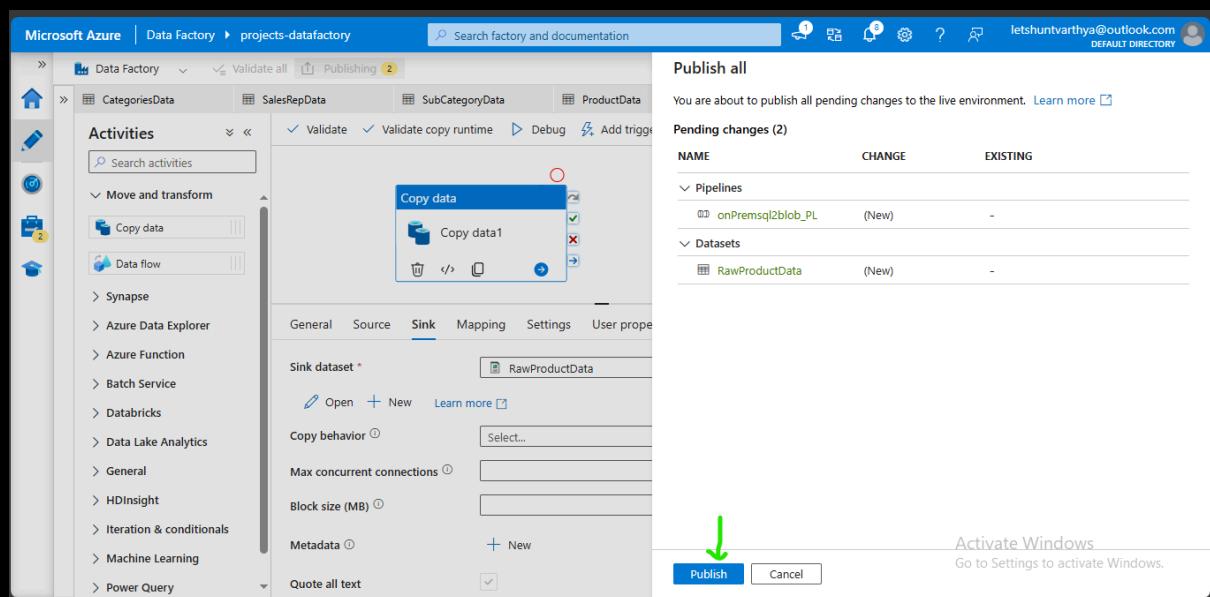
Block size (MB)

Metadata

Quote all text

Activate Windows
Go to Settings to activate Windows.

Publish Cancel



Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Preview experience Off

Add trigger

Trigger now New/Edit

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

onPremsql2blob_PL

Properties

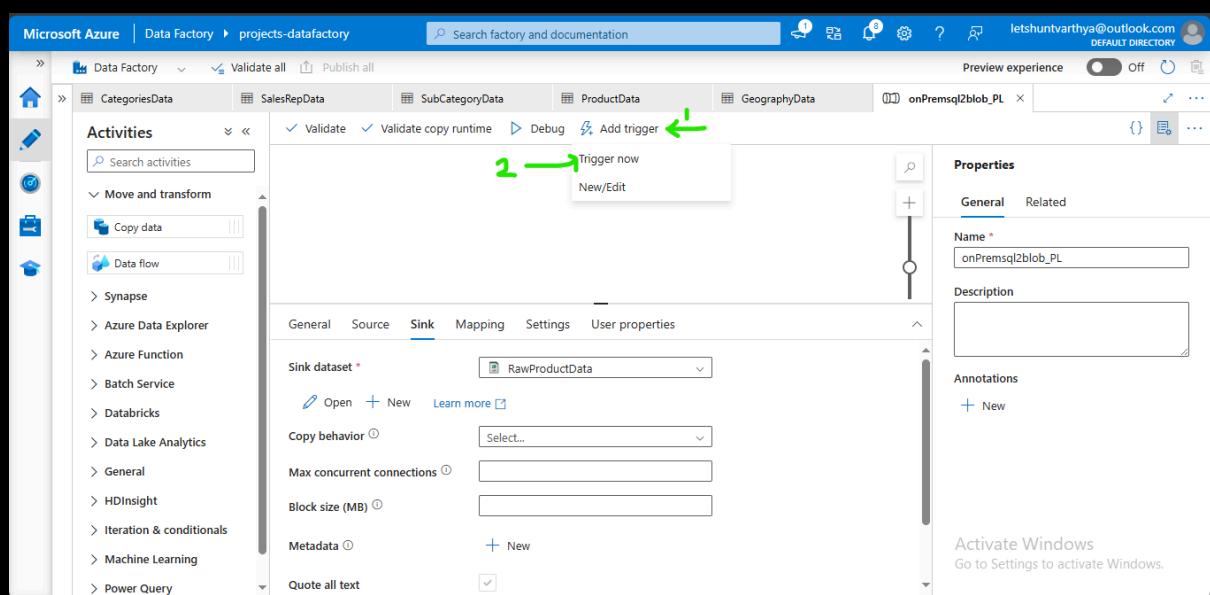
General Related

Name * onPremsql2blob_PL

Description

Annotations

Activate Windows
Go to Settings to activate Windows.



Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Validate all Publish all

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

onPremsql2blob_PL

Pipeline run

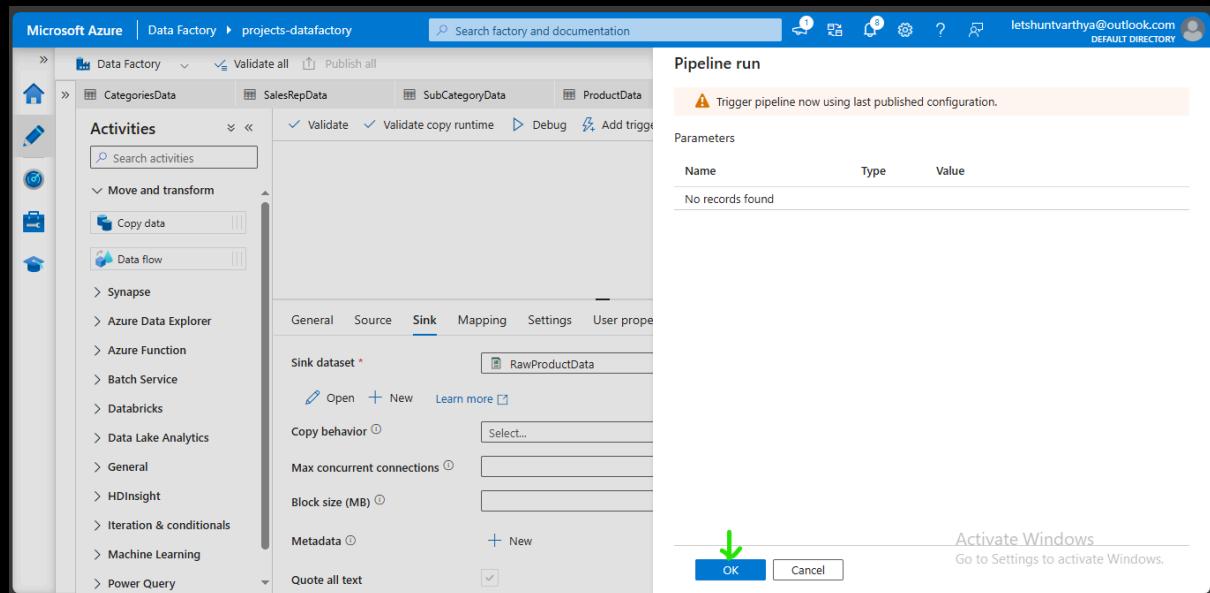
Trigger pipeline now using last published configuration.

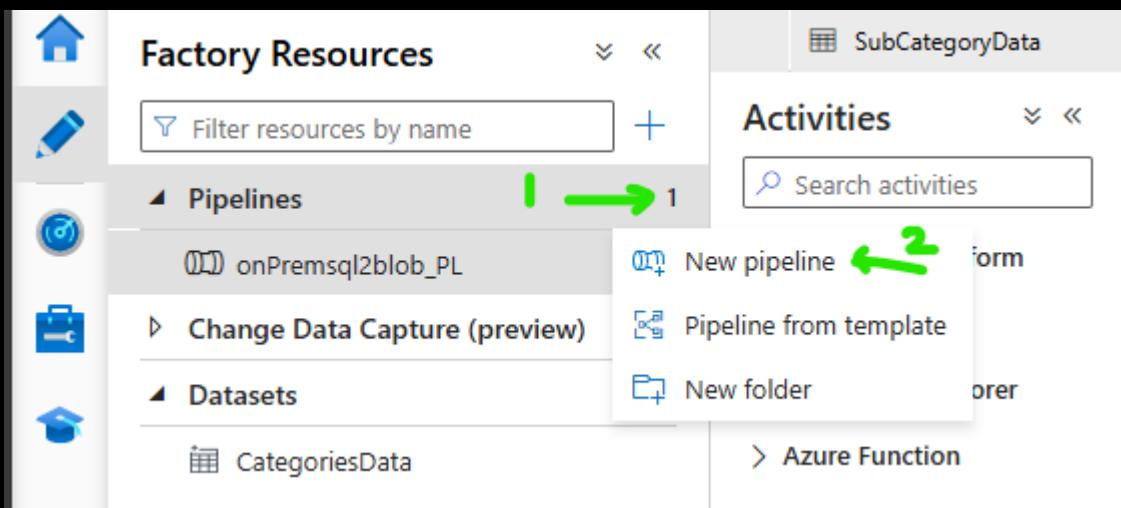
Parameters

Name	Type	Value
No records found		

Activate Windows
Go to Settings to activate Windows.

OK Cancel





Activities

- Move and transform 2
- Copy data 3
- Data flow

Properties

Name 1: nosql2blob_PL

Description

Activity state: Activated

Timeout: 0:12:00

Activities

- Move and transform
- Copy data
- Data flow

Properties

Name: nosql2blob_PL

Description

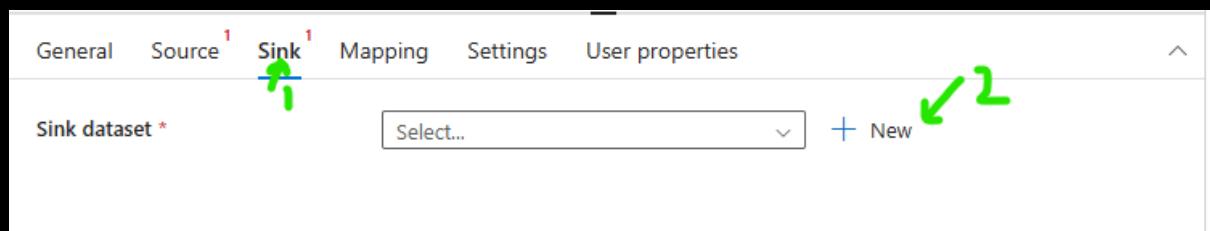
Activity state: Activated

Timeout: 0:12:00

Source 1

Source dataset

Select... 2



Microsoft Azure | Data Factory | projects-datafactory | Search factory and documentation

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

Search

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
Activate Windows		

Continue ← 2

Microsoft Azure | Data Factory | projects-datafactory | Search factory and documentation

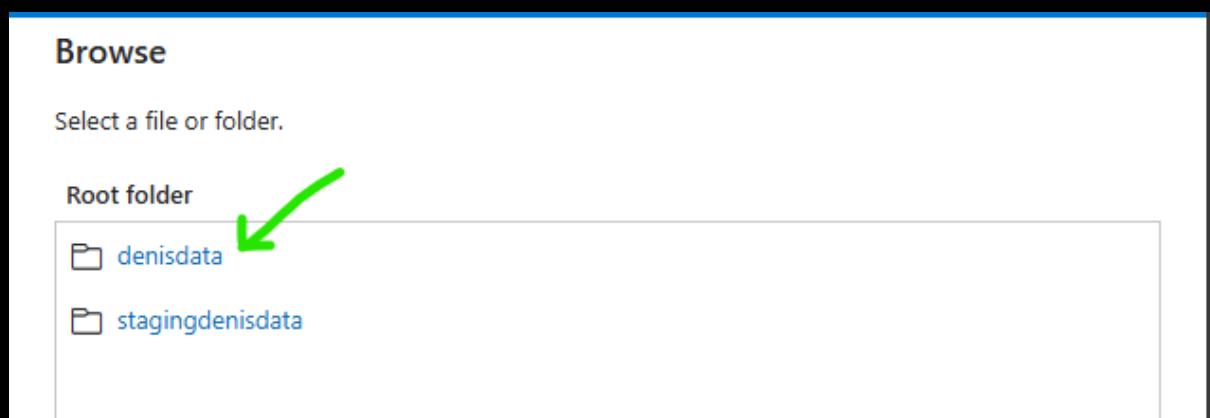
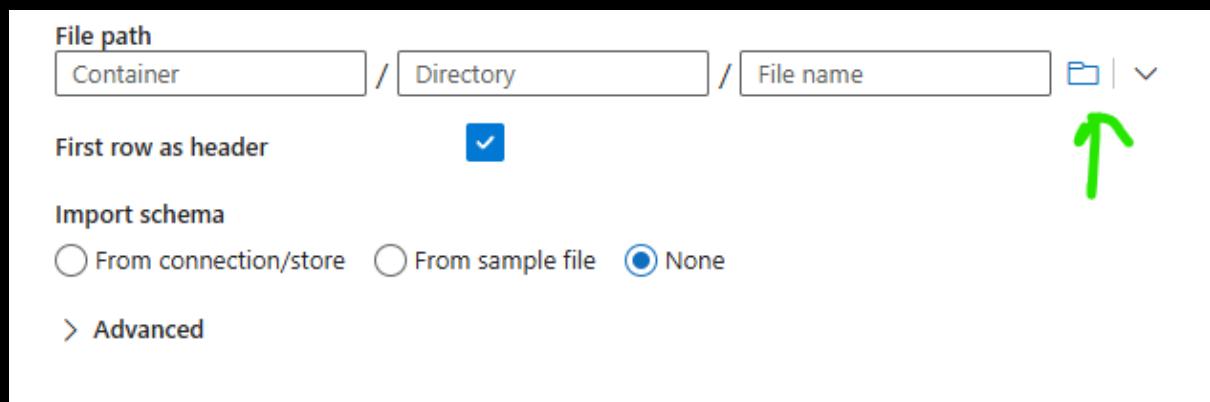
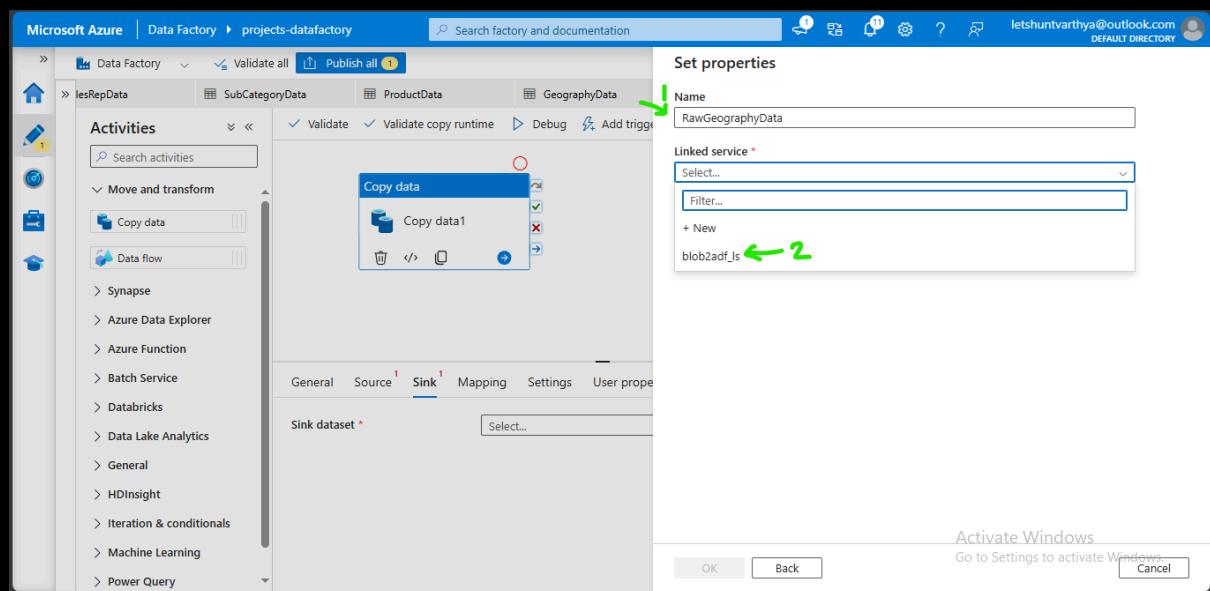
Select format

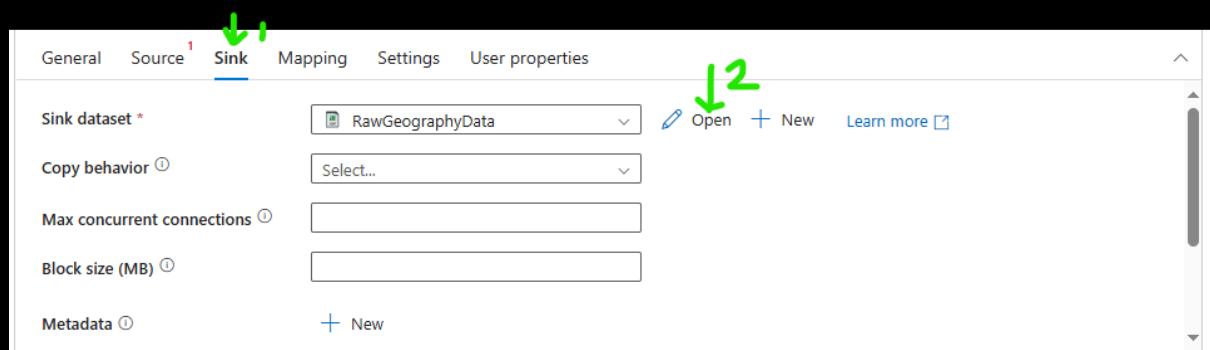
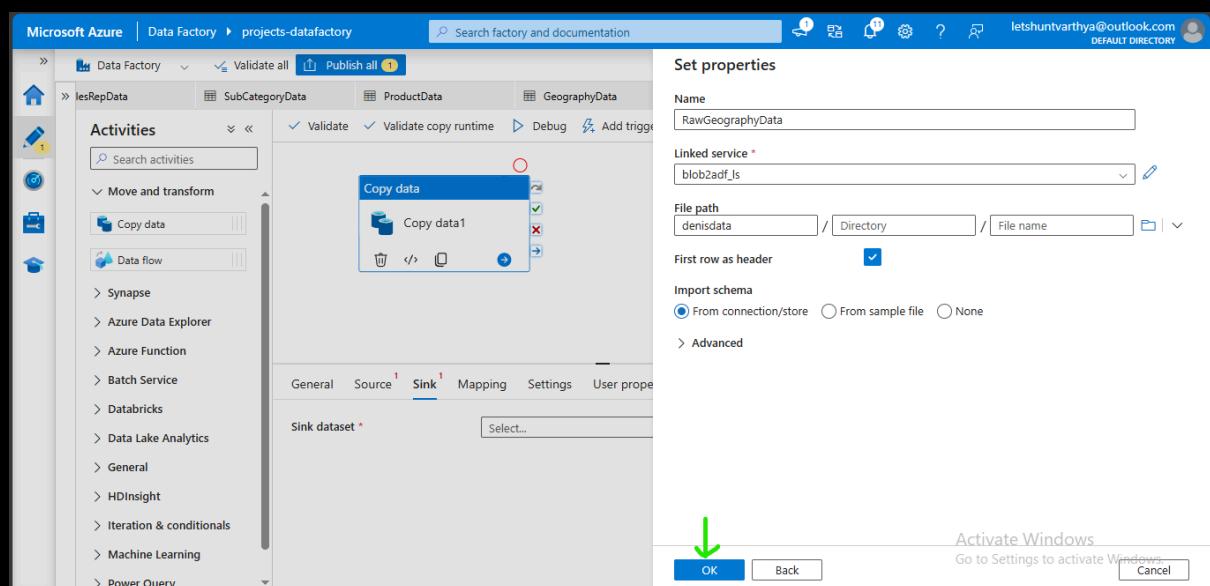
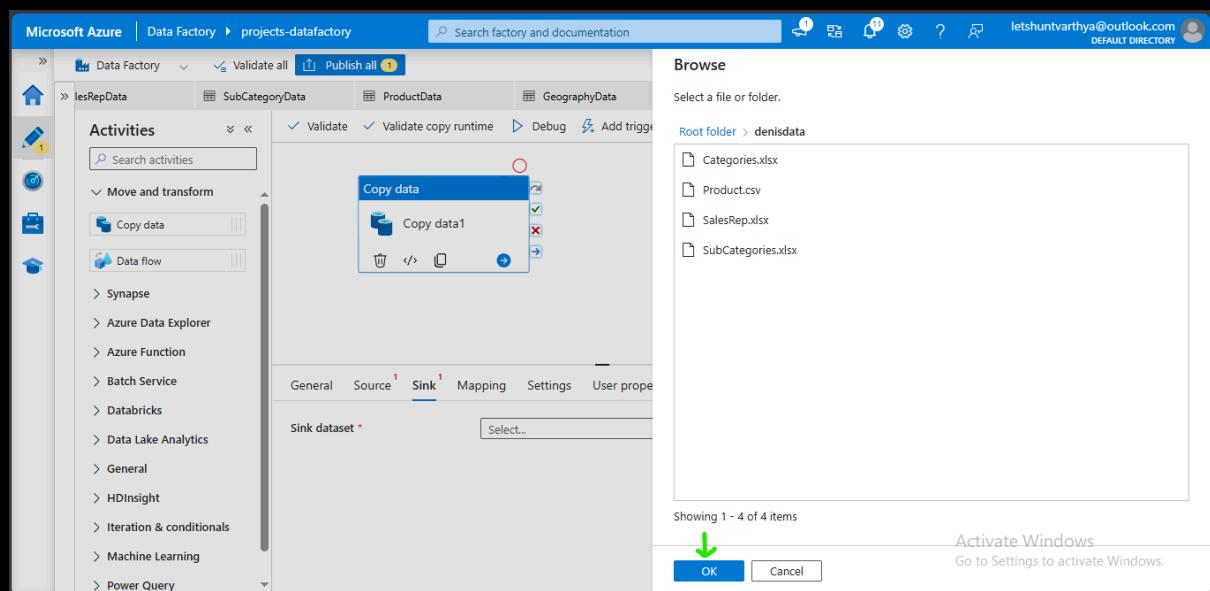
Choose the format type of your data

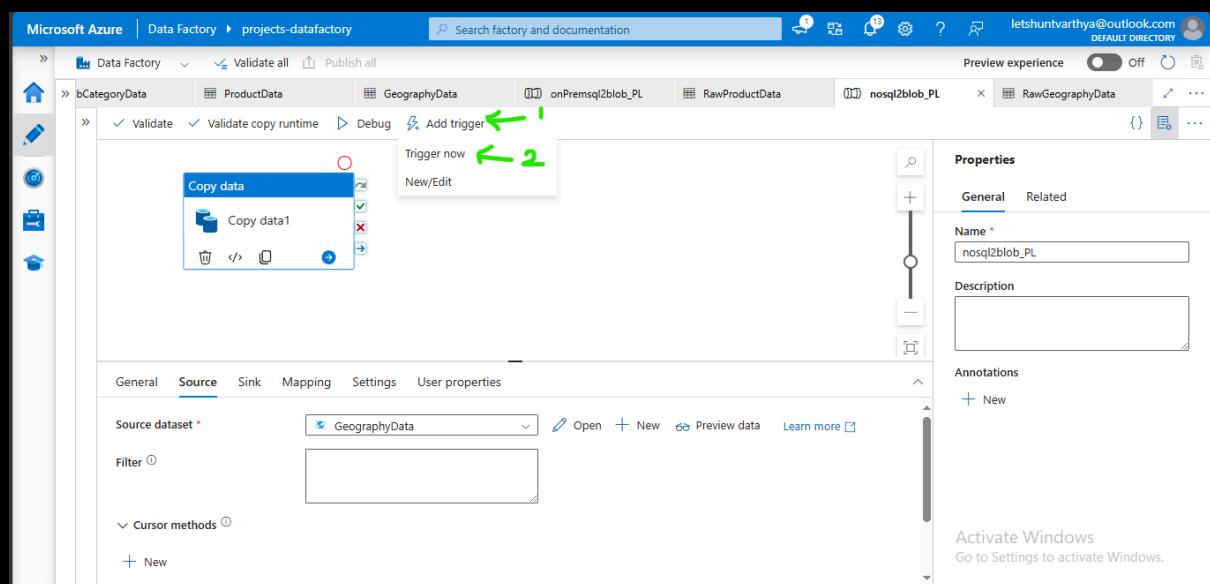
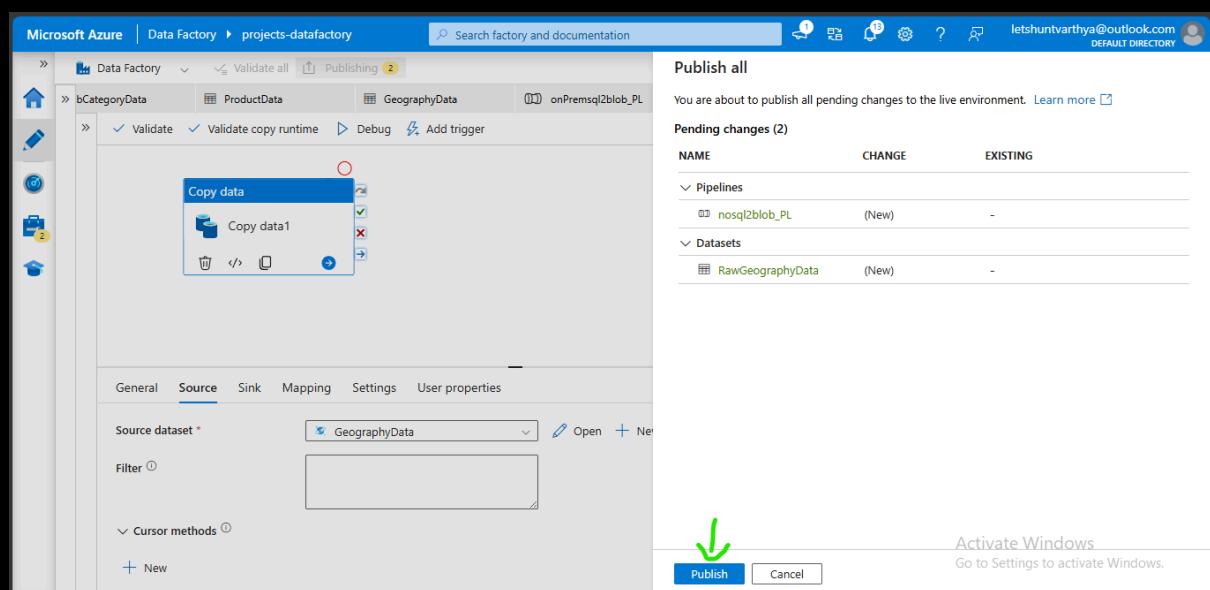
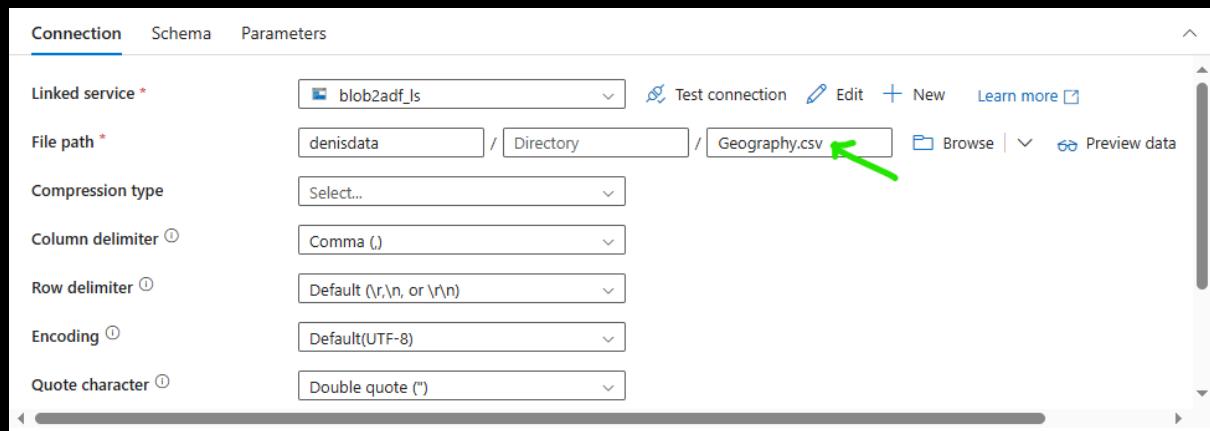
Avro	Binary	CSV
JSON	ORC	Parquet

Activate Windows

Continue ← 2







Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all

Search factory and documentation

bCategoryData ProductData GeographyData onPremsql2blob_PL

Validate Validate copy runtime Debug Add trigger

Copy data

Copy data1

OK Cancel

Activate Windows Go to Settings to activate Windows.

This screenshot shows the Microsoft Azure Data Factory interface. A pipeline named 'Copy data' is selected. The 'Source' tab is active, showing a 'Source dataset' dropdown set to 'GeographyData'. The 'General' tab is also visible. A green arrow points to the 'OK' button at the bottom right of the dialog.

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all

Search factory and documentation

Factory Resources bCategoryData ProductData GeographyData onPremsql2blob_PL RawProductData nosql2blob_PL

Pipelines 2 Change Data Capture (preview) 0 Datasets 7 Data flows 0 Power Query

New data flow

Properties General Related

Name nosql2blob_PL

Description

Annotations

Activate Windows Go to Settings to activate Windows.

This screenshot shows the 'Factory Resources' blade in the Azure Data Factory interface. Under 'Data flows', there is a 'New data flow' button highlighted with a green arrow. The 'Properties' pane on the right shows a data flow named 'nosql2blob_PL'.

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Search factory and documentation

onPremsql2blob_PL RawProductData nosql2blob_PL RawGeographyData ProductData_DF

Validate Data flow debug

Add Source

Properties General Related

Name ProductData_DF

Description

Parameters Settings

+ New

This screenshot shows the 'Factory Resources' blade again. A data flow named 'ProductData_DF' is selected. The 'Properties' pane on the right shows the name 'ProductData_DF' highlighted with a green arrow. The 'Add Source' button is also highlighted with a green arrow.

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all

ProductData GeographyData onPremsql2blob_PL RawProductData nosql2blob_PL RawGeographyData ProductData_DF

Validate Data flow debug

source1
Columns: 0 total

Add Source

Source settings 2 → Select... Options Sampling 3 → RawProductData

Properties General Related
Name * ProductData_DF
Description

Activate Windows Go to Settings to activate Windows.

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all

ProductData GeographyData onPremsql2blob_PL RawProductData nosql2blob_PL RawGeographyData ProductData_DF

Validate Data flow debug

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all

ProductData GeographyData onPremsql2blob_PL RawProductData nosql2blob_PL RawGeographyData ProductData_DF

Validate Data flow debug

source1
Columns: 0 total

Add Source

Source settings Source options Projection Optimize Inspect Data preview

Define default format Detect data type Import projection Reset schema

Turn on data flow debug

Integration runtime AutoResolveIntegrationRuntime

AutoResolveIntegrationRuntime Region AutoResolve Compute size Small

Debug time to live 1 hour

Activate Windows Go to Settings to activate Windows.

OK Cancel

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all

ProductData GeographyData onPremsql2blob_PL RawProductData nosql2blob_PL RawGeographyData ProductData_DF

Validate Data flow debug

Debug Settings

Preview experience Off

Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Data Factory Validate all Publish all

productData GeographyData onPremsql2blob_PL RawProductData nosql2blob_PL RawGeographyData ProductData_DF

Validate Data flow debug Debug Settings

source1

Columns: 6 total

Add Source

Properties

Name: ProductData_DF

Description

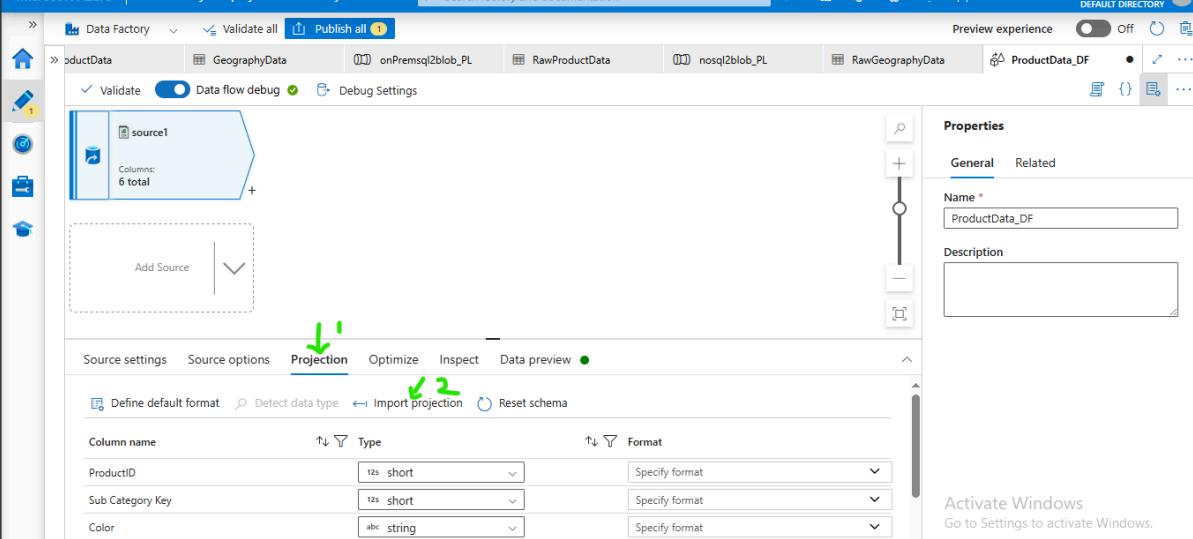
Source settings Source options Projection Optimize Inspect Data preview

Define default format Detect data type Import projection Reset schema

Column name Type Format

ProductID	12s short	Specify format
Sub Category Key	12s short	Specify format
Color	abc string	Specify format
ProductName	abc string	Specify format

Activate Windows Go to Settings to activate Windows.



Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Data Factory Validate all Publish all

productData GeographyData onPremsql2blob_PL RawProductData nosql2blob_PL RawGeographyData ProductData_DF

Validate Data flow debug Debug Settings

source1

Columns: 6 total

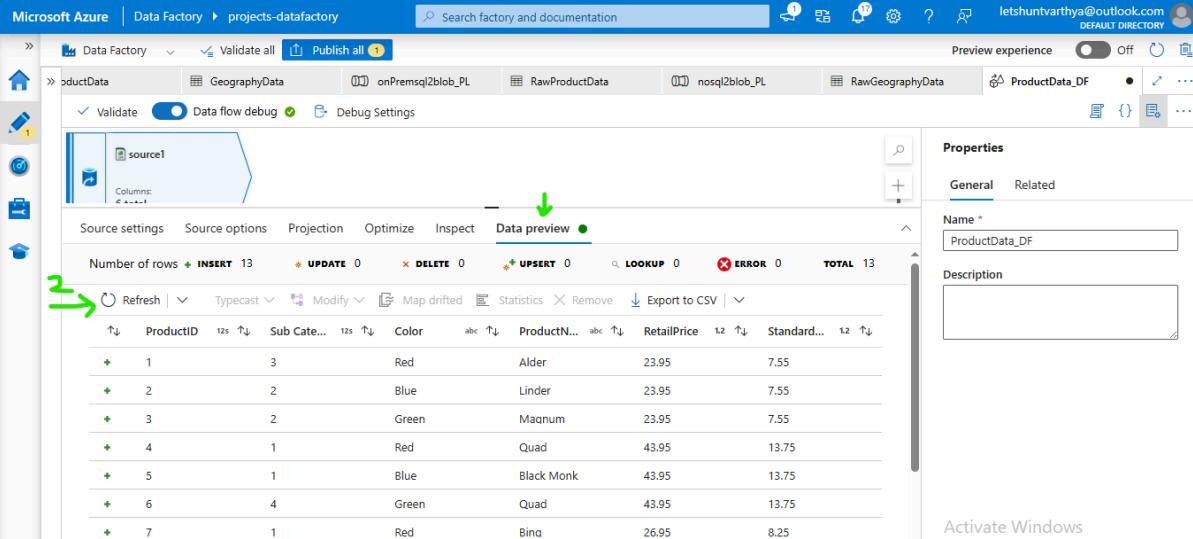
Source settings Source options Projection Optimize Inspect Data preview

Number of rows: 13 * INSERT: 13 * UPDATE: 0 * DELETE: 0 * UPSERT: 0 * LOOKUP: 0 * ERROR: 0 TOTAL: 13

Refresh Typecast Modify Map drifted Statistics Remove Export to CSV

ProductID	Sub Category Key	Color	ProductName	RetailPrice	StandardPrice
1	3	Red	Alder	23.95	7.55
2	2	Blue	Linder	23.95	7.55
3	2	Green	Magnum	23.95	7.55
4	1	Red	Quad	43.95	13.75
5	1	Blue	Black Monk	43.95	13.75
6	4	Green	Quad	43.95	13.75
7	1	Red	Bing	26.95	8.25
8	3	Blue	VanHelen	26.95	8.25

Activate Windows Go to Settings to activate Windows.



Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Data Factory Validate all Publish all

productData GeographyData onPremsql2blob_PL RawProductData nosql2blob_PL RawGeographyData ProductData_DF

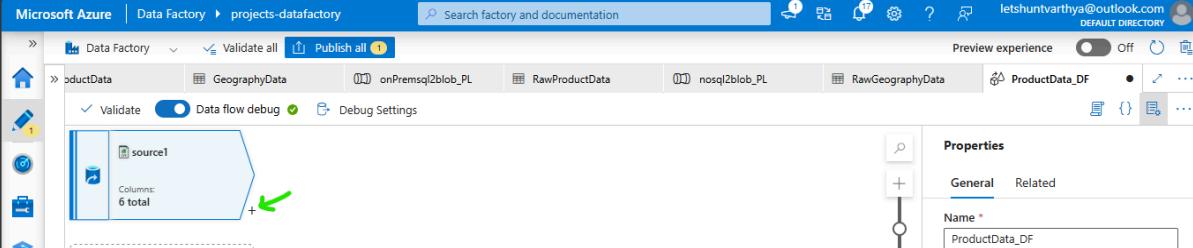
Validate Data flow debug Debug Settings

source1

Columns: 6 total

Properties

Name: ProductData_DF



Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all

ProductData

Validate Data flow debug Debug Settings

source1

Columns: 6 total

Lookup

Schema modifier

Add Source

Derived Column

Select

Aggregate 

Surrogate Key

Source settings Source opts

Number of rows + INSERT 1:

- Pivot
- Unpivot
- Window
- Rank
- External Call
- Cast

Formatters

Data preview

UPSERT 0 LOOKUP 0 ERROR 0 TOTAL N/A

ProductN...	RetailPrice	Standard...
Alder	23.95	7.55
Linder	23.95	7.55
Magnum	23.95	7.55

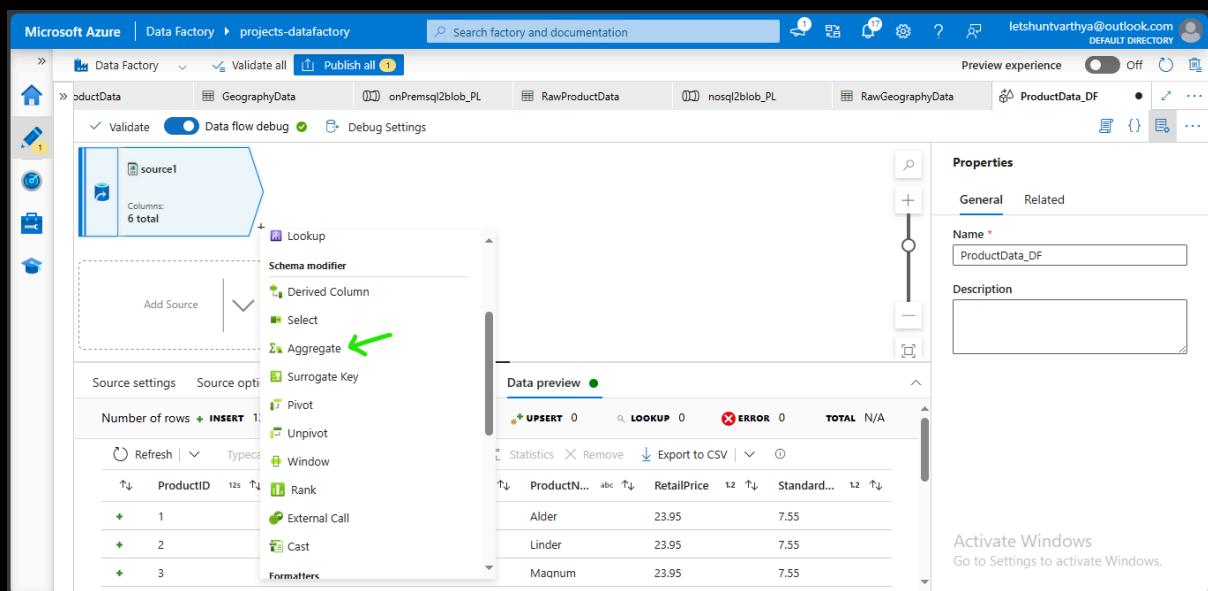
Properties

General Related

Name * ProductData_DF

Description

Activate Windows Go to Settings to activate Windows.



Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all

ProductData

Validate Data flow debug Debug Settings

Aggregate settings Optimize Inspect Data preview

Output stream name * aggregate1

Description Add aggregate columns

Incoming stream * source1

Group by Aggregates 

Columns Name as

12s ProductID	ProductID
Filter...	

12s ProductID

12s Sub Category Key

abc Color

abc ProductName

12s RetailPrice

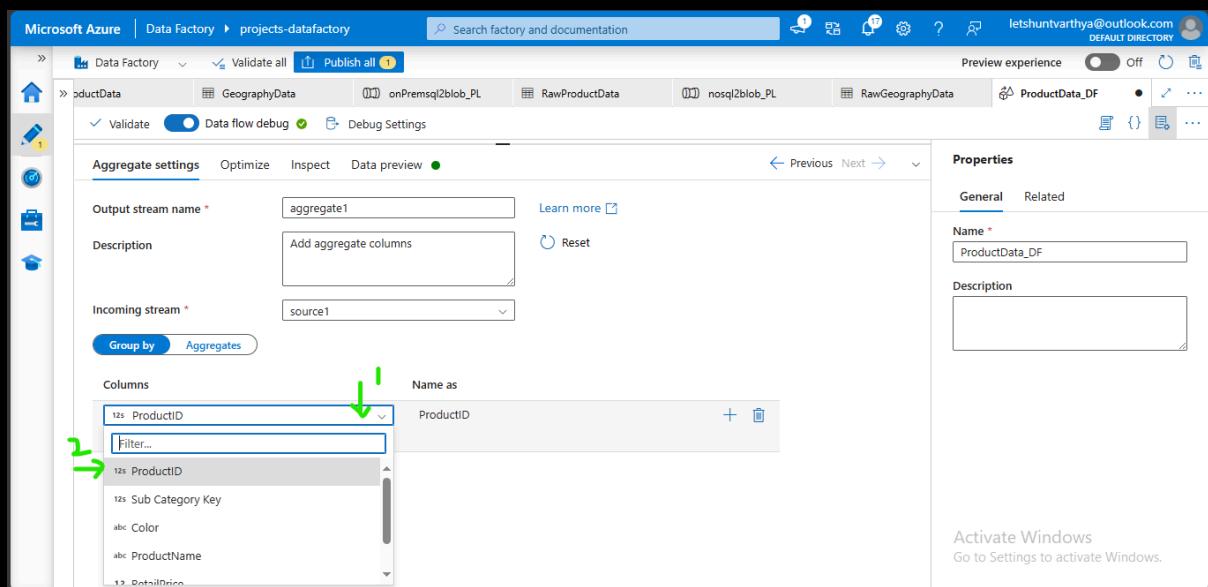
Properties

General Related

Name * ProductData_DF

Description

Activate Windows Go to Settings to activate Windows.



Aggregate settings Optimize Inspect Data preview

Output stream name * aggregate1

Description Add aggregate columns

Incoming stream * source1

Group by Aggregates 

Columns Name as

12s ProductID	ProductID

+

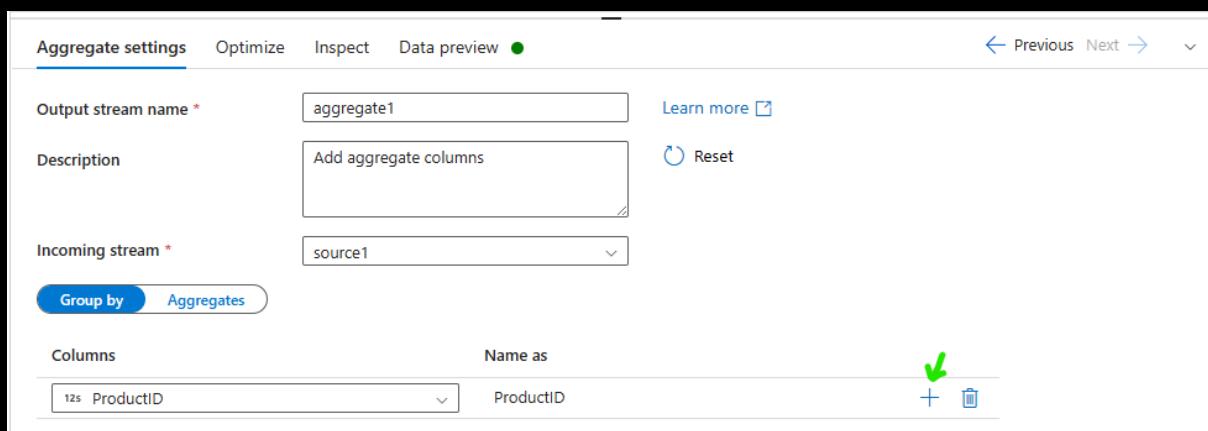
Properties

General Related

Name * ProductData_DF

Description

Activate Windows Go to Settings to activate Windows.



Aggregate settings Optimize Inspect Data preview ●

[← Previous](#) [Next →](#) [...](#)

Output stream name *	<input type="text" value="aggregate1"/>	Learn more
Description	<input type="text" value="Add aggregate columns"/> Reset	
Incoming stream *	<input style="width: 200px;" type="text" value="source1"/> Group by Aggregates	
Columns	Name as	
12s ProductID	ProductID	
abc ProductName	ProductName	
12s Sub Category Key	Sub Category Key	
abc Color	Color	
1.2 RetailPrice	RetailPrice	
1.2 StandardCost	StandardCost	

Aggregate settings Optimize Inspect Data preview ●

← Previous Next →

Output stream name * aggregate1 [Learn more](#)

Description Aggregating data by 'ProductID, ProductName, Sub Category Key, Color, RetailPrice, StandardCost' producing

Reset

Incoming stream * source1

[Group by](#) **Aggregates**

Grouped by: ProductID, ProductName, Sub Category Key, Color, RetailPrice, StandardCost

+ Add [Clone](#) [Delete](#) [Open expression builder](#)

<input type="checkbox"/> Column	Expression
<input type="checkbox"/> ProductCount	<code>count(ProductID)</code>

121 + [Delete](#)

The screenshot shows the Microsoft Azure Data Factory Data Flow blade. A pipeline step named 'aggregate1' is selected, which contains an 'Aggregate' transformation. The 'Schema modifier' pane on the right is open, showing various options like 'Derived Column', 'Select' (which is highlighted with a green arrow labeled '2'), and 'Aggregate'. The 'Properties' pane on the far right shows the name 'ProductData_DF'. The pipeline diagram at the top shows 'source1' connected to 'aggregate1', with a 'Lookup' step also present.

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Search factory and documentation

Preview experience Off

Data Flow Editor

ProductData

GeographyData

onPremsql2blob_PL

RawProductData

nosql2blob_PL

RawGeographyData

ProductData_DF

Validate Data flow debug Debug Settings

Source1: Import data from RawProductData

aggregate1: Aggregating data by ProductID, ProductName, SubCategory Key, Color, RetailPrice, StandardCost

select1: Columns: 7 total

Select settings Optimize Inspect Data preview

Skip duplicate output columns

Auto mapping Reset Add mapping Delete 7 mappings: All inputs mapped

aggregate1's column Name as

aggregate1's column	Name as
12s ProductID	ProductID
abc ProductName	ProductName
12s Sub Category Key	Sub Category Key
abc Color	Color
12 RetailPrice	RetailPrice
12 StandardCost	StandardCost
12l ProductCount	ProductCount

Properties

Name: ProductData_DF

Description:

Activate Windows Go to Settings to activate Windows.

Select settings Optimize Inspect Data preview

Number of rows + INSERT 11 * UPDATE 0 × DELETE 0 * UPSERT 0 ↗ LOOKUP 0 ✘ ERROR 0 TOTAL 11

Refresh Typecast Modify Map drifted Statistics Remove Export to CSV

ProductID	ProductName	Sub Category Key	Color	RetailPrice	StandardCost
1	Alder	3	Red	23.95	7.55
2	Linder	2	Blue	23.95	7.55
3	Magnum	2	Green	23.95	7.55
4	Quad	1	Red	43.95	13.75
5	Black Monk	1	Blue	43.95	13.75
6	Quad	4	Green	43.95	13.75

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Search factory and documentation

Preview experience Off

Data Flow Editor

ProductData

GeographyData

onPremsql2blob_PL

RawProductData

nosql2blob_PL

RawGeographyData

ProductData_DF

Validate Data flow debug Debug Settings

Source1: Import data from RawProductData

aggregate1: Aggregating data by ProductID, ProductName, SubCategory Key, Color, RetailPrice, StandardCost

select1: Columns: 6 total

Select settings Optimize Inspect Data preview

Number of rows + INSERT 11 * UPDATE 0 × DELETE 0 * UPSERT 0 ↗ LOOKUP 0 ✘ ERROR 0 TOTAL 11

Refresh Typecast Modify Map drifted Statistics Remove Export to CSV

ProductID	ProductName	Sub Category Key	Color	RetailPrice	StandardCost
1	Alder	3	Red	23.95	7.55
2	Linder	2	Blue	23.95	7.55
3	Magnum	2	Green	23.95	7.55
4	Quad	1	Red	43.95	13.75
5	Black Monk	1	Blue	43.95	13.75
6	Quad	4	Green	43.95	13.75

Properties

Name: ProductData_DF

Description:

Activate Windows Go to Settings to activate Windows.

Row modifier

- Flatten
- Parse
- Stringify

Row modifier

- Filter
- Sort
- Alter Row
- Assert

Flowlets

- Flowlet

Destination

- Sink

Microsoft Azure | Data Factory | projects-datafactory | Validate all | Publish all

Data flow debug | Debug Settings

Source1: Import data from RawProductData

aggregate1: Aggregating data by ProductID,ProductName,SubCategory Key,Color,RetailPrice,StandardCost

select1: Renaming aggregated1 to select1 with columns ProductID,ProductName,SubCategory Key,Color,...

sink1: Columns: 6 total

Sink | Settings | Errors | Mapping | Optimize | Inspect | Data preview

Output stream name *: sink1 | Learn more

Description: Add sink dataset | Reset

Incoming stream *: select1

Sink type *: Dataset | Inline | Cache

Dataset *: Select... | New (green arrow)

Options: Allow schema drift (checked) | Validate schema

Properties

Name: ProductData_DF

Description:

Activate Windows

Go to Settings to activate Windows.

Microsoft Azure | Data Factory | projects-datafactory | Validate all | Publish all

Data flow debug | Debug Settings

Source1: Import data from RawProductData

aggregate1: Aggregating data by ProductID,ProductName,SubCategory Key,Color,RetailPrice,StandardCost

select1: Renaming aggregated1 to select1 with columns ProductID,ProductName,SubCategory Key,Color,...

sink1: Columns: 6 total

Sink | Settings | Errors | Mapping | Optimize | Inspect | Data preview

Output stream name *: sink1 | Learn more

Description: Add sink dataset | Reset

Incoming stream *: select1

Sink type *: Dataset | Inline | Cache

Dataset *: Select... | New

Options: Allow schema drift (checked) | Validate schema

New dataset

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

Select a data store

Search

All | Azure | Database | File | Generic protocol | NoSQL | Services and apps

Amazon S3 | Azure Blob Storage | Azure Cosmos DB for NoSQL

Azure Data Explorer (Kusto) | Azure Data Lake Storage Gen2 | Azure Database for MySQL

Continue (green arrow) | 2

Activate Windows

Go to Settings to activate Windows.

Cancel

Microsoft Azure | Data Factory | projects-datafactory | Validate all | Publish all

Data flow debug | Debug Settings

Source1: Import data from RawProductData

aggregate1: Aggregating data by ProductID,ProductName,SubCategory Key,Color,RetailPrice,StandardCost

select1: Renaming aggregated1 to select1 with columns ProductID,ProductName,SubCategory Key,Color,...

sink1: Columns: 6 total

Sink | Settings | Errors | Mapping | Optimize | Inspect | Data preview

Output stream name *: sink1 | Learn more

Description: Add sink dataset | Reset

Incoming stream *: select1

Sink type *: Dataset | Inline | Cache

Dataset *: Select... | New

Options: Allow schema drift (checked) | Validate schema

Select format

Choose the format type of your data

Avro | DelimitedText | JSON

ORC | Parquet | Binary

Continue (green arrow) | 2

Activate Windows

Go to Settings to activate Windows.

Cancel

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Search factory and documentation

Data Factory Validate Data flow debug Debug Settings

source1 aggregate1 select1

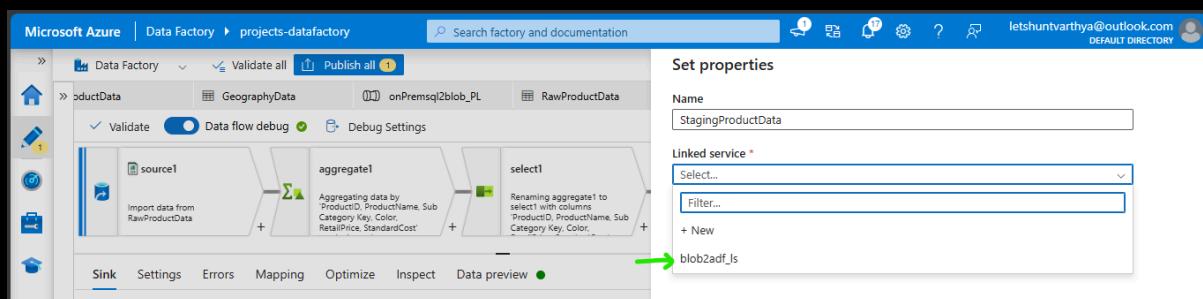
Import data from RawProductData Aggregating data by ProductID, ProductName, SubCategory Key, Color, RetailPrice, StandardCost Renaming aggregate1 to select1 with columns ProductID, ProductName, SubCategory Key, Color, -

Sink Settings Errors Mapping Optimize Inspect Data preview

Set properties

Name: StagingProductData

Linked service: Select... Filter... + New blob2adf_ls



File path

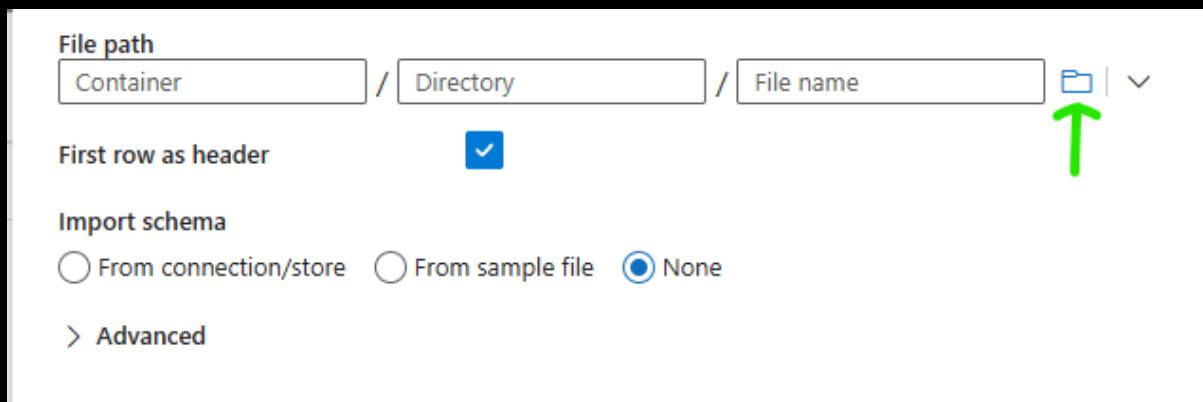
Container / Directory / File name

First row as header

Import schema

From connection/store From sample file None

> Advanced



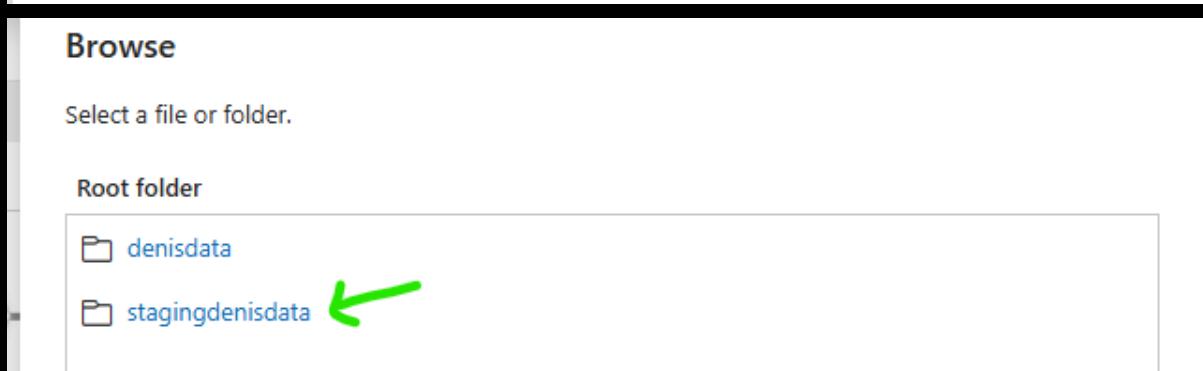
Browse

Select a file or folder.

Root folder

denisdata

stagingdenisdata



Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Search factory and documentation

Data Factory Validate Data flow debug Debug Settings

source1 aggregate1 select1

Import data from RawProductData Aggregating data by ProductID, ProductName, SubCategory Key, Color, RetailPrice, StandardCost Renaming aggregate1 to select1 with columns ProductID, ProductName, SubCategory Key, Color, -

Sink Settings Errors Mapping Optimize Inspect Data preview

Output stream name: sink1

Description: Add sink dataset

Incoming stream: select1

Sink type: Dataset

Dataset: Select... New

Options: Allow schema drift Validate schema

Browse

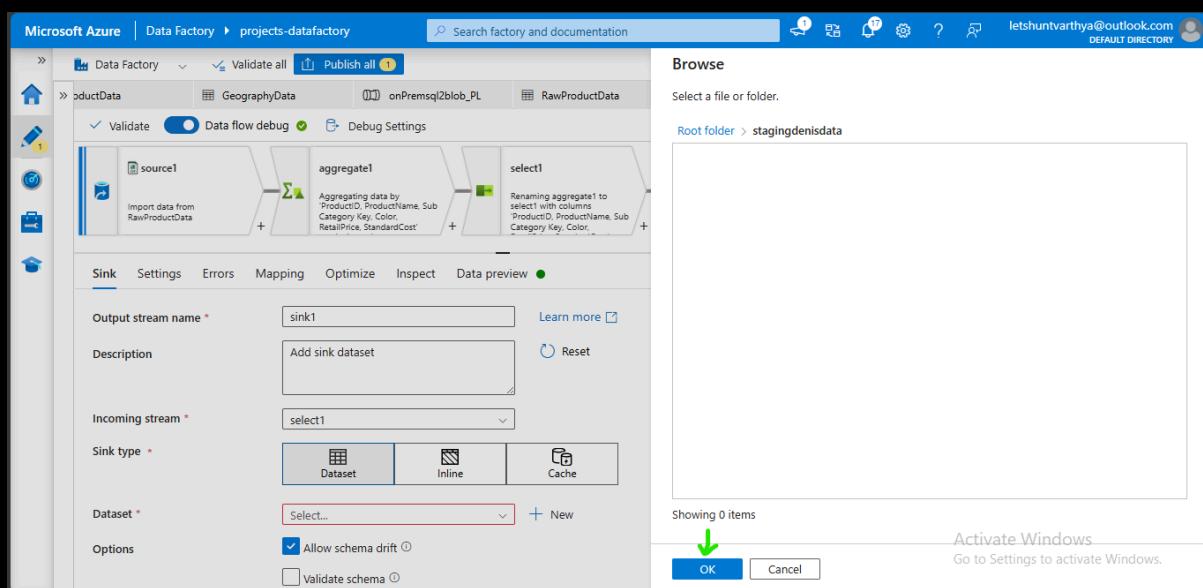
Select a file or folder.

Root folder > stagingdenisdata

Showing 0 items

Activate Windows Go to Settings to activate Windows.

OK Cancel



Microsoft Azure | Data Factory > projects-datafactory | Search factory and documentation

Validate all | Publish all | Set properties

Set properties

Name: StagingProductData
Linked service: blob2adf_ls
File path: stagingdenisdata | / Directory | / File name
First row as header:
Import schema: From connection/store (selected) | From sample file | None
Advanced

Sink

Output stream name: sink1 | Learn more
Description: Add sink dataset | Reset
Incoming stream: select1
Sink type: Dataset | Inline | Cache
Dataset: Select... | + New
Options: Allow schema drift (selected) | Validate schema

OK | Back | Activate Windows | Go to Settings to activate Windows | Cancel

Validate all | Publish all | Set properties

Properties

Name: ProductData_DF
Description:

Sink

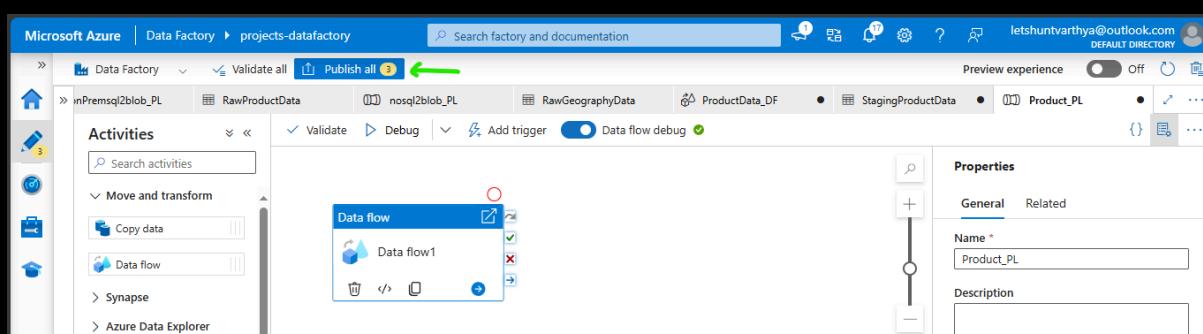
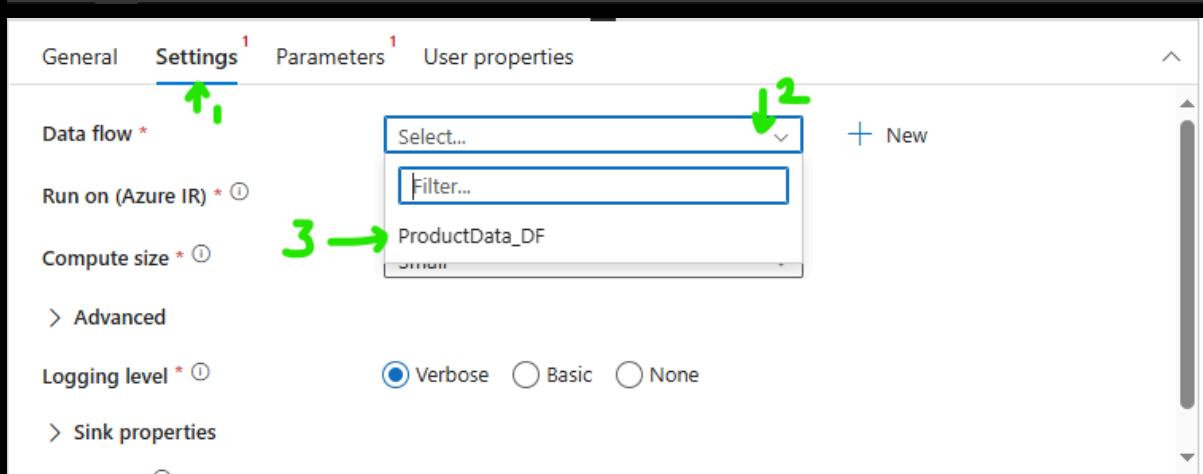
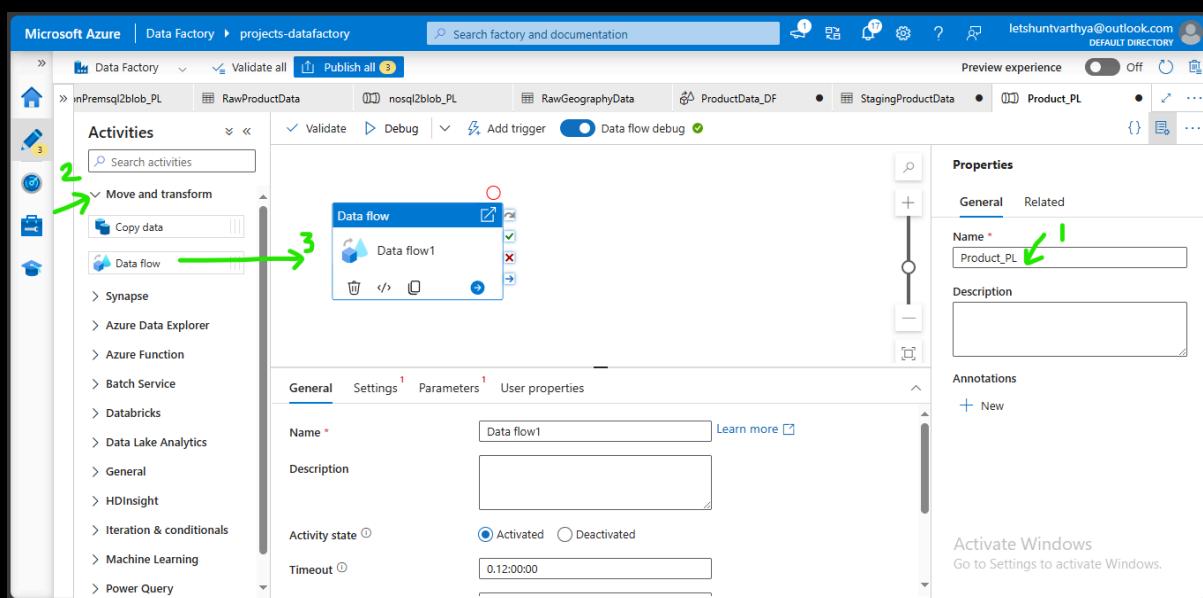
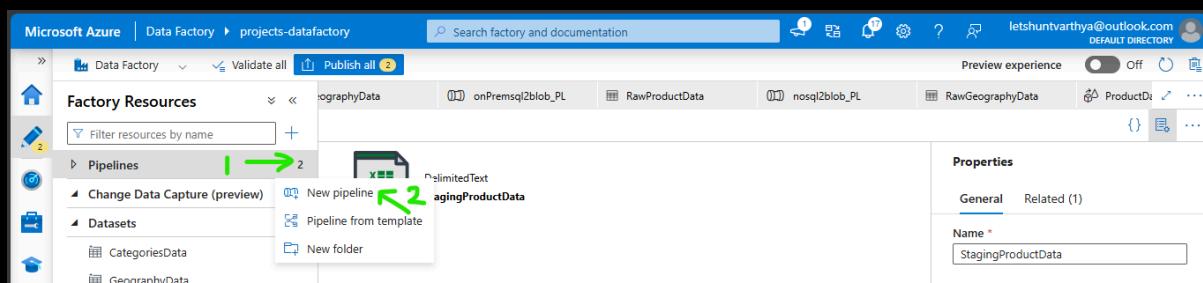
Output stream name: sink1 | Learn more
Description: Export data to StagingProductData | Reset
Incoming stream: select1
Sink type: Dataset | Inline | Cache
Dataset: StagingProductData | Test connection | Open (selected) | + New
Skip line count:
Options: Allow schema drift

Activate Windows | Go to Settings to activate Windows.

Connection Schema Parameters

Connection

Linked service: blob2adf_ls | Test connection | Edit | + New | Learn more
File path: stagingdenisdata | / Directory | / StagingProduct.csv | Browse | Preview data
Compression type: Select...
Column delimiter: Comma (,) | Encoding: Default(UTF-8)
Row delimiter: Default (\r, \n, or \r\n) | Quote character: Double quote (")



Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Data Factory Publishing

Activities

Move and transform

Copy data

Data flow

Synapse

Azure Data Explorer

Azure Function

Batch Service

Databricks

Data Lake Analytics

General

HDIInsight

Iteration & conditionals

Machine Learning

Power Query

Data flow

Data flow1

General Settings Parameters User properties

Data flow * ProductData_DF

Run on (Azure IR) * AutoResolveIntegrationRuntime

Compute size * Small

Advanced

Logging level * Verbose

Sink properties

Publish Cancel

Activate Windows Go to Settings to activate Windows.

This screenshot shows the Microsoft Azure Data Factory publishing interface. It displays a list of pending changes across three categories: Pipelines, Datasets, and Data flows. The Pipelines section includes a new pipeline named 'Product_PL'. The Datasets section includes a new dataset named 'StagingProductData'. The Data flows section includes a new data flow named 'ProductData_DF'. At the bottom, there are 'Publish' and 'Cancel' buttons, and a note about activating Windows.

Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Data Factory Publishing

Factory Resources

Pipelines

nosql2blob_PL

onPremsql2blob_PL

Product_PL

Change Data Capture (preview)

Datasets

Data flows

New data flow

1 → 1

ProductData_DF

Power Query

Settings Parameters User properties

New data flow

New flowlet

New folder

DF

Open New

This screenshot shows the Microsoft Azure Data Factory Factory Resources interface. On the left, the 'Data flows' section is expanded, showing a count of 1. A green arrow points from the number '1' to the 'New data flow' button. In the center, a 'Data flow' panel is open, showing a single data flow named 'Data flow1'. Below it, there are tabs for 'Settings', 'Parameters', and 'User properties'. At the bottom, there are buttons for 'New data flow', 'New flowlet', and 'New folder', along with a dropdown menu set to 'DF'.

Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Data Factory Publishing

Factory Resources

Pipelines

nosql2blob_PL

onPremsql2blob_PL

Product_PL

Change Data Capture (preview)

RawGeographyData

ProductData_DF

StagingProductData

Product_PL

dataflow1

Properties

General Related

Name * dataflow1

This screenshot shows the Microsoft Azure Data Factory Factory Resources interface. On the left, the 'Data flows' section is expanded, showing a count of 0. A green arrow points to the 'Add Source' button. In the center, a 'dataflow1' panel is open, showing a dashed box for adding sources. To the right, there is a 'Properties' pane with tabs for 'General' and 'Related', and a 'Name' field set to 'dataflow1'. At the top, there are buttons for 'Validate', 'Data flow debug' (which is turned on), and 'Debug Settings'.

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all

Factory Resources

- Pipelines 3
- Change Data Capture (preview) 0
- Datasets 8
 - CategoriesData
 - GeographyData
 - ProductData
 - RawGeographyData
 - RawProductData
 - SalesRepData
 - StagingProductData
 - SubCategoriesData
- Data flows 2
 - CategoriesData_DF
 - ProductData_DF
- Power Query 0

Search factory and documentation

Validate Data flow debug

Properties

Name: CategoriesData_DF

Description:

Source settings

Description: Add source dataset

Source type: Dataset

Dataset: Select... (highlighted)

Options: Sampling

Sampling: CategoriesData, RawGeographyData, RawProductData, SalesRepData

```
graph TD; 1[1] --> Name[Name: CategoriesData_DF]; 2[2] --> Dataset[Dataset: Select...]; 3[3] --> Sampling[Sampling: CategoriesData, RawGeographyData, RawProductData, SalesRepData];
```

Validate all Publish all

Validate Data flow debug

Successfully imported Successfully imported the schema for source1 (Source).

Projection

Source settings

Projection

Source options

Projection

Optimize

Inspect

Data preview

Define default format

Column name

CategoryKey	Type: tzs short	Format: Specify format
Category	Type: abc string	Format: Specify format

Properties

Name: CategoriesData_DF

Description:

```
graph TD; 1[1] --> Projection[Projection]; 2[2] --> Type[Type: tzs short]; 3[3] --> Format[Format: Specify format];
```

Source settings

Source options

Projection

Optimize

Inspect

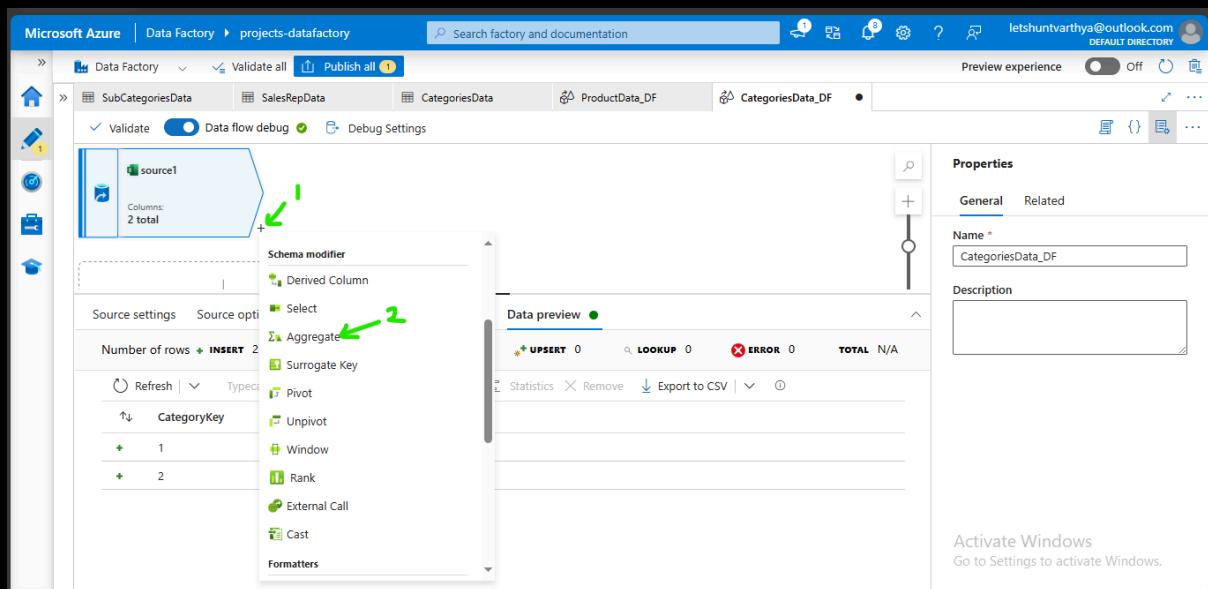
Data preview

Number of rows: + INSERT 2 * UPDATE 0 × DELETE 0 + UPSE 0 LOOKUP 0 ✘ ERROR 0 TOTAL N/A

Refresh | Typecast | Modify | Map drifted | Statistics | Remove | Export to CSV |

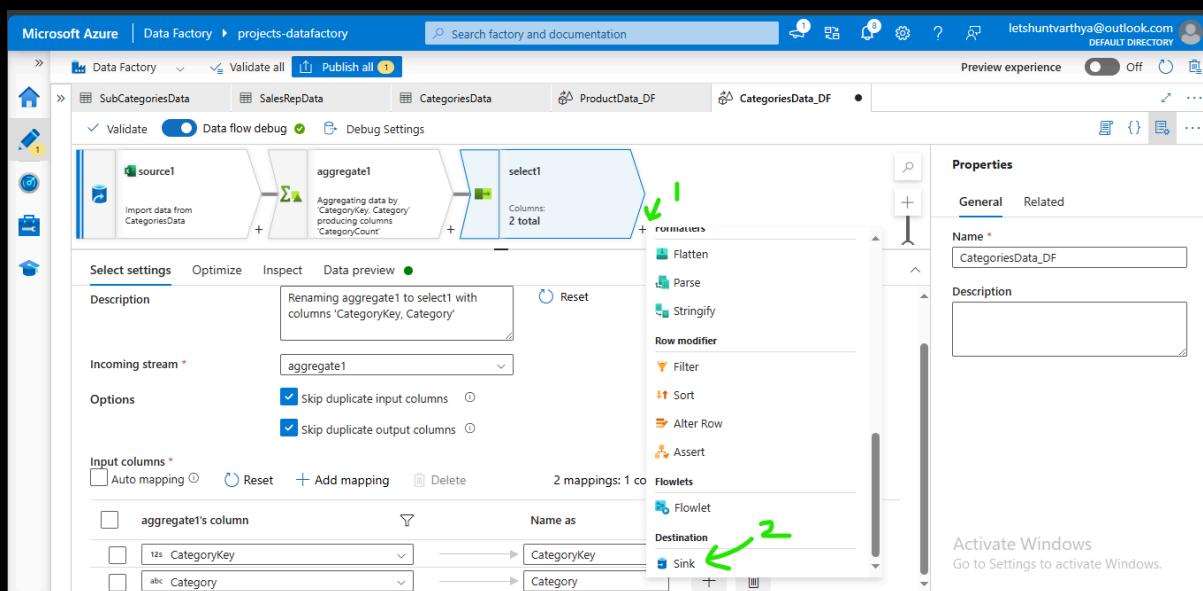
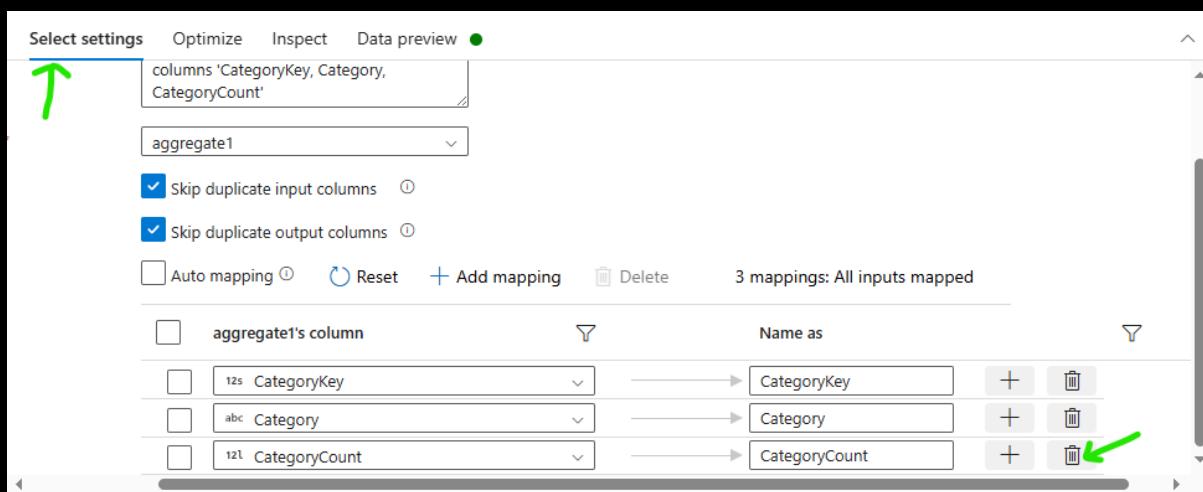
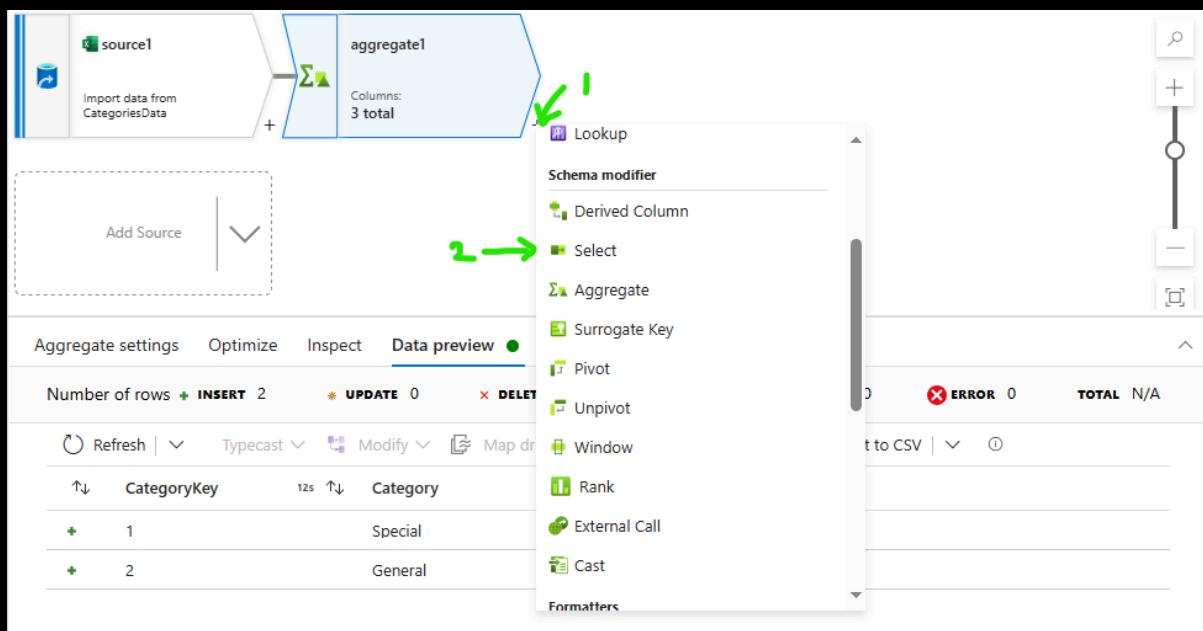
CategoryKey	Category
1	Special
2	General

```
graph TD; 1[1] --> DataPreview[Data preview];
```



This screenshot shows the 'Aggregate settings' blade for the 'aggregate1' stream. It has tabs for 'Aggregate settings', 'Optimize', 'Inspect', and 'Data preview'. Under 'Aggregate settings', there are fields for 'Output stream name' (set to 'aggregate1'), 'Description' (set to 'Add aggregate columns'), and 'Incoming stream' (set to 'source1'). Below these, there are two columns: 'CategoryKey' and 'Category'. Each column has a 'Name as' field (both set to their respective column names) and a '+' button for adding more columns. At the bottom, there are tabs for 'Group by' and 'Aggregates', with 'Aggregates' being the active tab, indicated by a red box and labeled '3'.

This screenshot shows the 'Aggregate settings' blade for the 'aggregate1' stream, with the 'Aggregates' tab selected. It has tabs for 'Aggregate settings', 'Optimize', 'Inspect', and 'Data preview'. Under 'Aggregate settings', there are fields for 'Output stream name' (set to 'aggregate1'), 'Description' (set to 'Aggregating data by 'CategoryKey', 'Category' producing columns 'CategoryCount''), and 'Incoming stream' (set to 'source1'). Below these, it says 'Grouped by: CategoryKey, Category'. There are buttons for '+ Add', 'Clone', 'Delete', and 'Open expression builder'. The 'Open expression builder' button is highlighted with a red box and labeled '3'. In the expression builder, there's a table with a single row for 'CategoryCount'. The 'Column' column contains 'CategoryCount' and the 'Expression' column contains 'count(CategoryKey)'. A green arrow labeled '1' points to the 'Aggregate settings' tab, and a green arrow labeled '2' points to the 'Grouped by' section.



Sink

Output stream name * Learn more [\[link\]](#)

Description [Reset](#)

Incoming stream *

Sink type * Dataset Inline Cache

Dataset * [New](#) 

Options Allow schema drift [\[link\]](#)

Properties

General Related

Name *

Description

Activate Windows Go to Settings to activate Windows.

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

Search

All Azure Database File Generic protocol NoSQL Services and apps

Amazon S3	Azure Blob Storage 	Azure Cosmos DB for NoSQL
Azure Data Explorer (Kusto)	Azure Data Lake Storage Gen2	Azure Database for MySQL

Activate Windows Go to Settings to activate Windows. [Cancel](#)

Select format

Choose the format type of your data

Avro	DelimitedText 	JSON
ORC	Parquet	Binary

Activate Windows Go to Settings to activate Windows. [Cancel](#)

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all

Search factory and documentation

Set properties

Name: StagingCategoriesData

Linked service: Select... Filter... + New blob2adf_ls

Diagram:

```
graph LR; source1[Import data from CategoriesData] --> aggregate1[Aggregating data by CategoryKey producing column 'CategoryCount']; aggregate1 --> target1[Reference: 1 Columns: 2 total]
```

Name: StagingCategoriesData

Linked service *: blob2adf_ls

File path: Container / Directory / File name

First row as header:

Import schema: From connection/store From sample file None

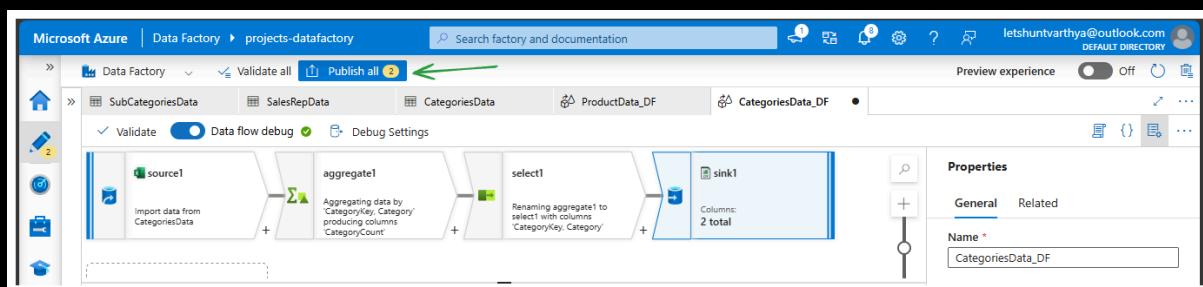
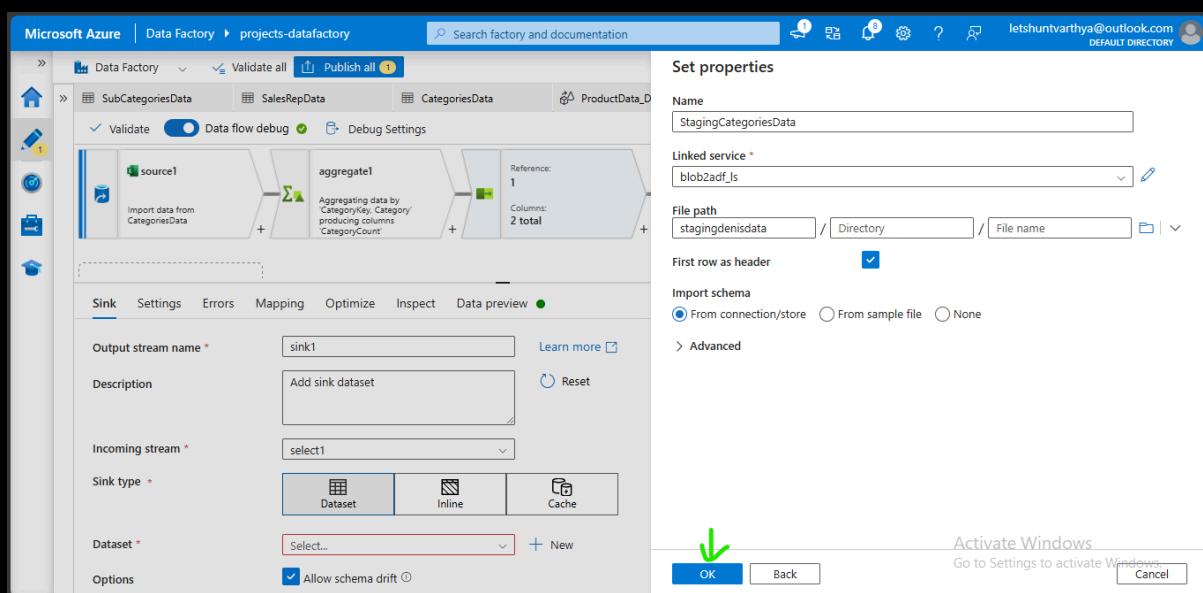
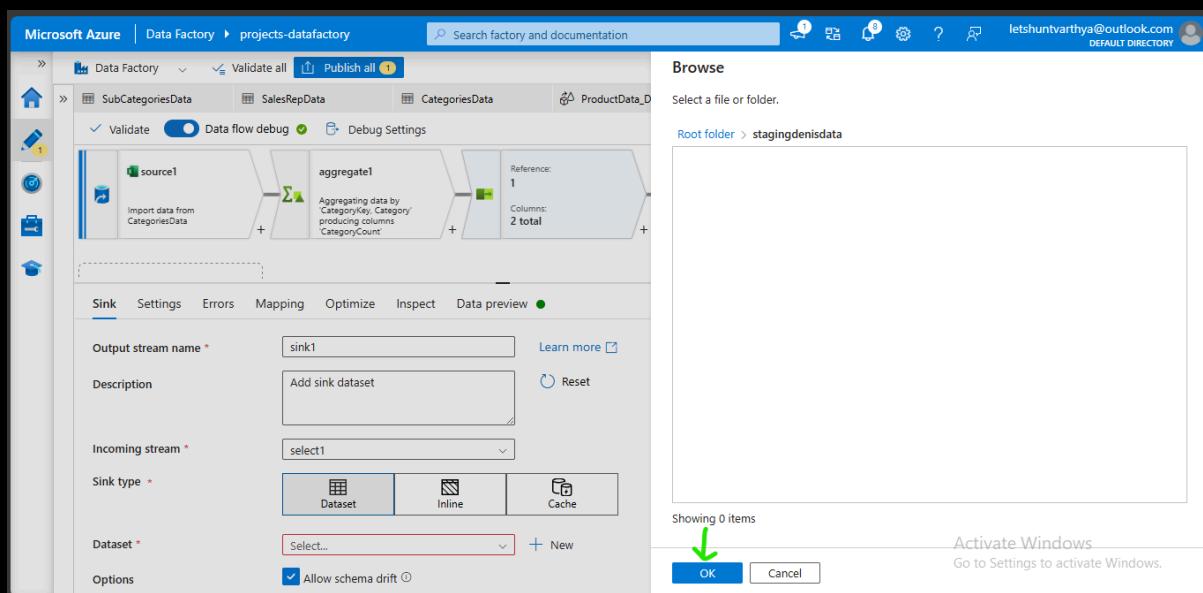
Advanced

Browse

Select a file or folder.

Root folder

- denisdata
- stagingdenisdata



Microsoft Azure | Data Factory projects-datafactory

Search factory and documentation

Publish all

You are about to publish all pending changes to the live environment. [Learn more](#)

Pending changes (2)

NAME	CHANGE	EXISTING
StagingCategoriesData	(New)	-
CategoriesData_DF	(New)	-

Sink Settings Errors Mapping Optimize Inspect Data preview

Output stream name * sink1 Learn more

Description Export data to StagingCategoriesData Reset

Incoming stream * select1

Sink type * Dataset Inline Cache

Dataset * StagingCategoriesData Test connection

Skip line count

Activate Windows
Go to Settings to activate Windows.

Publish **Cancel**

Microsoft Azure | Data Factory projects-datafactory

Search factory and documentation

Factory Resources

Pipelines

New pipeline

Pipeline from template

New folder

Properties

Name *

Preview experience Off

Microsoft Azure | Data Factory projects-datafactory

Search factory and documentation

Activities

Move and transform

Data flow

Copy data

Synapse

Azure Data Explorer

Azure Function

Batch Service

Databricks

Data Lake Analytics

General

HDI insight

Iteration & conditionals

Machine Learning

Power Query

Data flow

Data flow1

General Settings Parameters User properties

Name * Categories_PL

Description

Annotations

Activate Windows
Go to Settings to activate Windows.

Microsoft Azure | Data Factory projects-datafactory

Search factory and documentation

Validate all

Publish all

Preview experience Off

The screenshot shows the Microsoft Azure Data Factory Data Flow blade. On the left, the 'Factory Resources' sidebar lists various datasets and data flows. The 'Data flows' section is currently selected, with a green arrow pointing to it. Below it, a list of existing data flows includes 'CategoriesData_DF' (selected), 'ProductData_DF', and 'Power Query'. A second green arrow points to the 'New data flow' button, which is highlighted with a red box. The main workspace displays a pipeline diagram with three stages: 'source1' (Import data from CategoriesData), 'aggregate1' (Aggregating data by 'CategoryKey, Category' producing column 'CategoryCount'), and 'select1' (Renaming aggregate1 to select1 with columns 'CategoryKey, Category'). The 'Properties' pane on the right shows the details for 'CategoriesData_DF', including its name and description. The 'Sink' tab is selected in the pipeline configuration area.

The screenshot shows the Microsoft Azure Data Factory interface. On the left sidebar, under 'Factory Resources', 'Datasets' is expanded, showing 'CategoriesData', 'GeographyData', 'ProductData', 'RawGeographyData', 'RawProductData', 'SalesRepData', 'StagingCategoriesData', 'StagingProductData', and 'SubCategoriesData'. Under 'Data flows', 'SalesRepData_DF' is selected. The main workspace shows a 'SalesRepData' dataset with a 'Validate' button and a 'Data flow debug' toggle switch. A green arrow points to the 'Add Source' button. The 'Properties' panel on the right shows the data flow is named 'SalesRepData_DF' with no description.

The screenshot shows the Microsoft Azure Data Factory interface. On the left, the 'Factory Resources' sidebar lists various datasets and data flows. The 'Data flows' section is expanded, showing 'SalesRepData_DF' selected. In the main workspace, a data flow named 'source1' is being edited. The 'Source settings' tab is active, showing the 'Source type' as 'Dataset'. A dropdown menu for 'Dataset' is open, with 'Select...' highlighted and a green arrow pointing to it. Below the dropdown, a list of available datasets includes 'CategoriesData', 'RawGeographyData', 'RawProductData', and 'SalesRepData'. A green arrow also points from the bottom of the list towards the 'SalesRepData' entry. The 'Properties' pane on the right shows the dataset's name as 'SalesRepData_DF' and its description as 'Description'. At the top right, there are options for 'Preview experience' (switched off) and other account settings.

The screenshot shows the Azure Data Factory Data Flow editor interface. At the top, there is a source component named "source1" which has two columns: "total". Below the source, there is a green arrow pointing down to the "Projection" tab in the navigation bar. The "Projection" tab is currently selected. The "Source settings" and "Source options" tabs are also visible. Below the tabs, there are several buttons: "Define default format", "Detect data type", "Import projection", and "Reset schema".

Column name	Type	Format
col0	abc string	Specify format
col1	abc string	Specify format

Below the table, there is a "Properties" panel on the right side. It shows the "General" tab selected, with the "Name" field set to "SalesRepData_DF" and the "Description" field empty. There is also a note at the bottom right: "Activate Windows Go to Settings to activate Windows."

Derived column's settings

Incoming stream *: source1

Columns *:

Column: DSalesRepID

Expression: regexReplace(SalesRepID, "[^0-9]", "")

SalesRepID	Sales Rep Name	DSalesRepID
ID - 6	Jan Novotny	6
ID - 7	John White	7
ID - 5	Ellen Woody	5
ID - 3	Mark Spancer	3
ID - 1	Ellie Gill	1
ID - 2	Bill Muray	2
ID - 4	El Bob	4

Properties

Name: SalesRep_DF

Aggregate settings Optimize Inspect Data preview

Output stream name * aggregate1 [Learn more](#)

Description Add aggregate columns [Reset](#)

Incoming stream * derivedColumn1

Group by **Aggregates**

Columns	Name as
SalesRepID	SalesRepID
Sales Rep Name	Sales Rep Name

Aggregate settings Optimize Inspect Data preview

Output stream name * aggregate1 [Learn more](#)

Description Aggregating data by 'SalesRepID, Sales Rep Name' producing columns 'countRecords'

Incoming stream * derivedColumn1

Group by **Aggregates**

Grouped by: SalesRepID, Sales Rep Name

[Add](#) [Clone](#) [Delete](#) [Open expression builder](#)

Column	Expression
countRecords	count(SalesRepID)

Microsoft Azure | Data Factory | projects-datafactory | [Search factory and documentation](#)

Data Factory Validate all [Publish all](#) (1)

productData_DF CategoriesData_DF GeographyData ProductData RawGeographyData RawProductData SalesRep_DF

Validate Data flow debug Debug Settings

source1 derivedColumn1 aggregate1

Properties

Name SalesRep_DF

Description

Aggregate settings Optimize Inspect Data preview

Output stream name * aggregate1 [Learn more](#)

Description Aggregating data by 'SalesRepID, Sales Rep Name' producing columns 'countRecords'

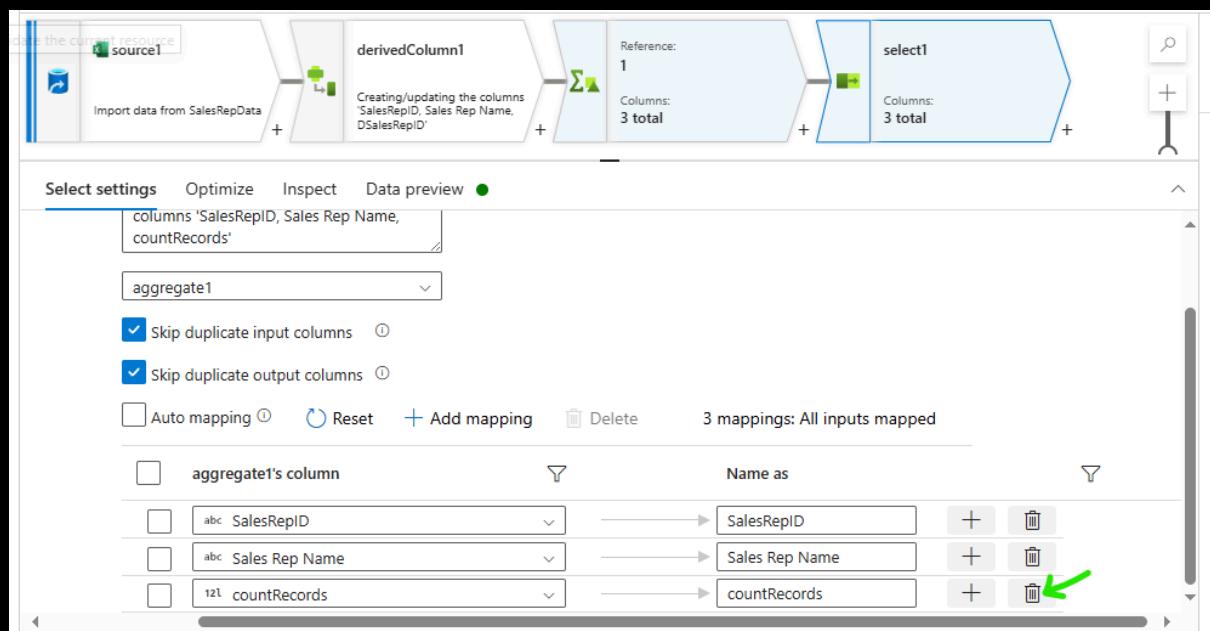
Incoming stream * derivedColumn1

Group by **Aggregates**

Grouped by: SalesRepID, Sales Rep Name

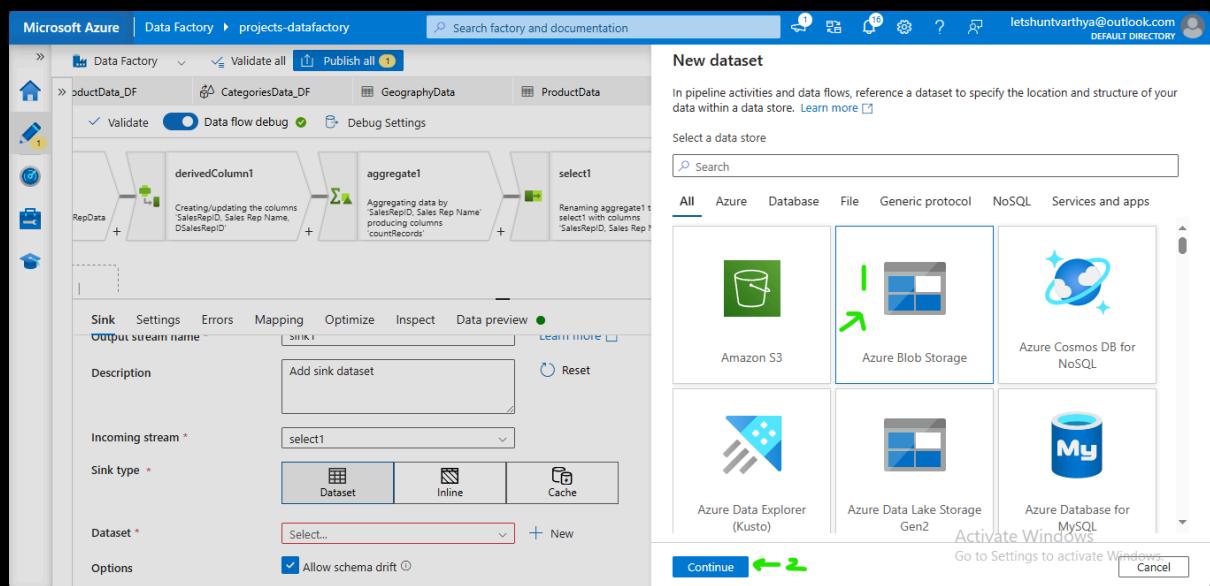
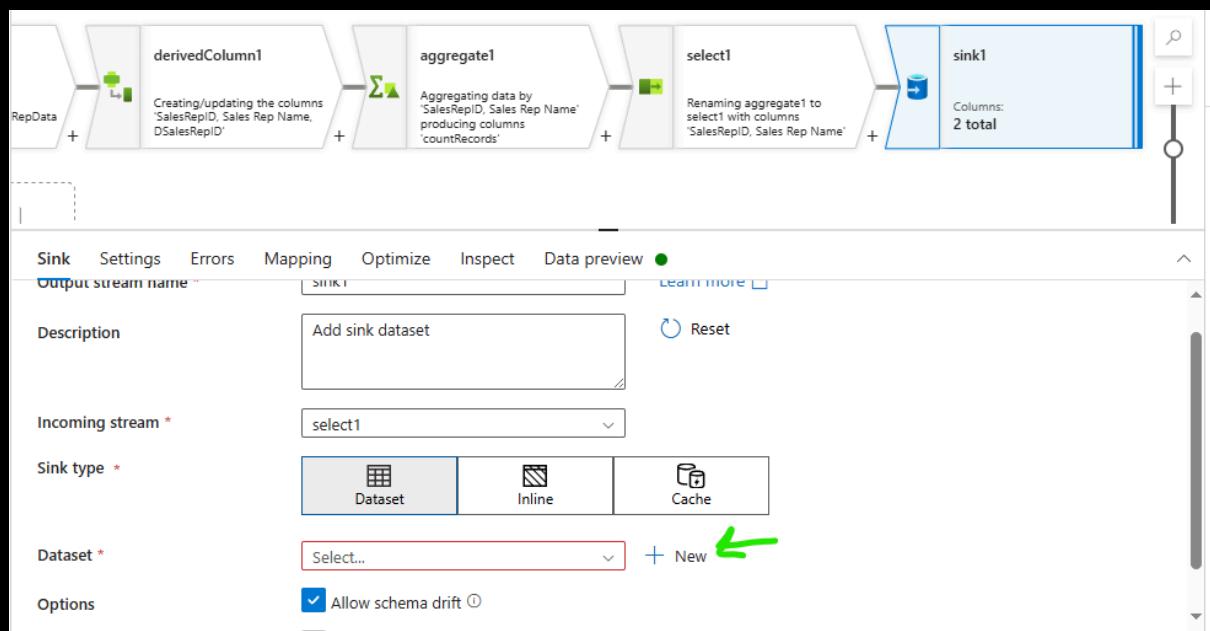
[Add](#) [Clone](#) [Delete](#) [Open expression builder](#)

Activate Windows
Go to Settings to activate Windows.



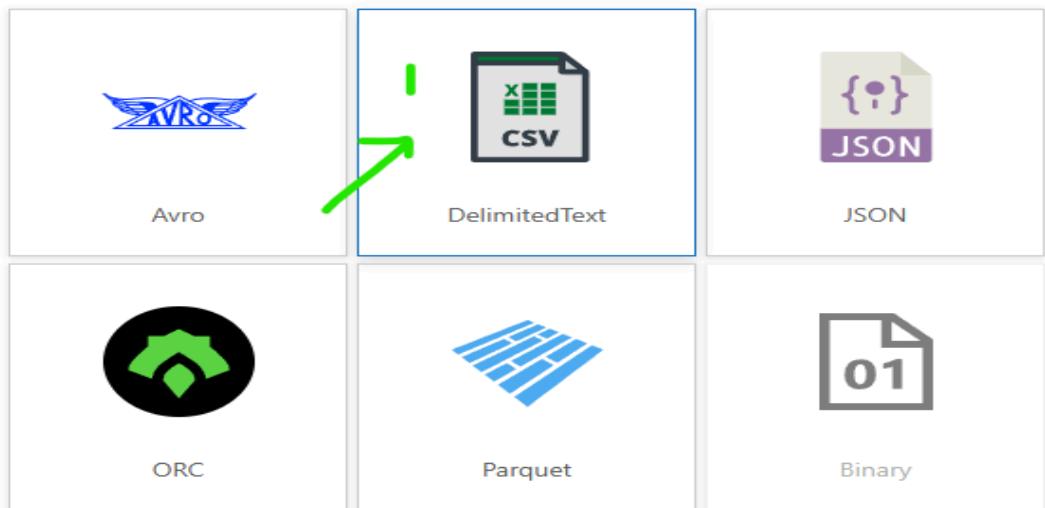
This screenshot shows the Data preview tab in the Microsoft Data Factory Data Flow designer. It displays 7 rows of data with columns **SalesRepID**, **Sales Rep Name**, and **countRecords**. A green arrow points from the source table icon to the **SalesRepID** column header. The total number of rows is listed as **TOTAL 7**.

This screenshot shows the Microsoft Azure Data Factory Data Flow blade. It displays the same data preview as the previous screenshots, with 7 rows of data and columns **SalesRepID**, **Sales Rep Name**, and **countRecords**. A green arrow points from the source table icon to the **SalesRepID** column header. On the right side, there is a sidebar with various properties and options, including a 'Properties' section and a 'Sink' button.



Select format

Choose the format type of your data



Continue

Back

Activate Windows

Go to Settings to activate Windows.

Cancel

Microsoft Azure | Data Factory | projects-datafactory | Search factory and documentation | DEFAULT DIRECTORY

ProductData_DF | Validate all | Publish all | Set properties

Name: StagingSalesRepData

Linked service: Select... (blob2adf_ls)

Set properties dialog box:

- Name: StagingSalesRepData
- Linked service: Select... (blob2adf_ls)
- + New
- blob2adf_ls

ProductData_DF Data Flow:

```
graph LR; RepData --> derivedColumn1[derivedColumn1]; derivedColumn1 --> aggregate1[aggregate1]; aggregate1 --> select1[select1];
```

Sink settings:

- Output stream name: SINK
- Description: Add sink dataset
- Incoming stream: select1
- Sink type: Dataset
- Dataset: Select... (blob2adf_ls)
- Options: Allow schema drift

Activate Windows

Go to Settings to activate Windows.

Cancel

Set properties

Name
StagingSalesRepData

Linked service *
blob2adf_ls

File path
Container / Directory / File name 

First row as header 

Import schema
 From connection/store From sample file None

[Advanced](#)

Browse

Select a file or folder.

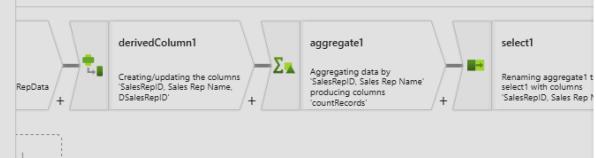
Root folder

- [denisdata](#)
- [stagingdenisdata](#) 

Microsoft Azure | Data Factory > projects-datafactory   DEFAULT DIRECTORY

ProductData_DF CategoriesData_DF GeographyData ProductData

Validate  



Sink [Settings](#) [Errors](#) [Mapping](#) [Optimize](#) [Inspect](#) [Data preview](#)

Output stream name [Sink 1](#) 

Description Add sink dataset 

Incoming stream * select1

Sink type *   

Dataset * [Select...](#) 

Options Allow schema drift  

Browse
Select a file or folder.
Root folder > stagingdenisdata

Showing 0 items  

Activate Windows
Go to Settings to activate Windows.

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Set properties

Name: StagingSalesRepData

Linked service: blob2adf_ls

File path: stagingdenisdata / Directory / File name

First row as header: ✓

Import schema: From connection/store (selected)

Advanced:

Description: Add sink dataset

Incoming stream: select1

Sink type: Dataset

Dataset: Select...

Options: Allow schema drift (selected)

OK **Back**

Activate Windows
Go to Settings to activate Windows
Cancel

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 2

ProductData DF CategoriesData DF GeographyData ProductData RawGeographyData RawProductData SalesRep DF

Properties

Name: SalesRep_DF

Description:

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all

Factory Resources

Datasets

- CategoriesData
- GeographyData
- ProductData
- RawGeographyData
- RawProductData
- SalesRepData
- StagingCategoriesData
- StagingProductData
- StagingSalesRepData
- SubCategoriesData

Data flows

- New data flow 2
- New flowlet
- New folder

Power Query

ProductData RawGeographyData RawProductData SalesRep DF

Properties

Name: SalesRep_DF

Description:

Preview experience: Off

Validate Data flow debug Debug Settings

Sink Settings Errors Mapping Optimize Inspect ...

Output stream name: sink1

Description: Export data to StagingSalesRepData

Incoming stream: select1

Sink type: Dataset

Dataset: StagingSalesRepData

Options:

Activate Windows
Go to Settings to activate Windows.

Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Preview experience Off

Factory Resources

- Datasets (10)
 - CategoriesData
 - GeographyData
 - ProductData
 - RawGeographyData
 - RawProductData
 - SalesRepData
 - StagingCategoriesData
 - StagingProductData
 - StagingSalesRepData
 - SubCategoriesData
- Data flows (4)
 - CategoriesData_DF
 - SubCategoriesData_DF
 - ProductData_DF
 - SalesRep_DF

source1

Source settings

Output stream name: source1

Description: Add source dataset

Source type: Dataset

Dataset: Select... (highlighted with green arrow)

Options

Sampling: 2

Properties

Name: SubCategoriesData_DF

Description:

Activate Windows
Go to Settings to activate Windows.

Projection

Define default format

Detected data type

Import projection

Reset schema

Column name	Type	Format
SubCategoryKey	12s short	Specify format
CategoryKey	abc string	Specify format
SubCategory Name	abc string	Specify format

Data preview

Number of rows: 4

SubCategoryKey	CategoryKey	SubCategory Name
1	ID -1	Extra
2	ID - 2	Regular
3	ID - 1	Micro
4	ID - 2	Super

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Search factory and documentation

Preview experience Off

Data Factory SubCategorieData_DF

Validate Data flow debug Debug Settings

Properties

General Related

Name: SubCategorieData_DF

Description:

Activate Windows Go to Settings to activate Windows.

source1 Import data from SubCategoriesData

derivedColumn1 Columns: 4 total

Derived column's settings Optimize Inspect Data preview

Output stream name: derivedColumn1

Description: Creating/updating the columns 'SubCategoryKey, CategoryKey, SubCategory Name, DcategoryKey'

Incoming stream: source1

Columns:

Column	Expression
DcategoryKey	regexReplace(CategoryKey, "[^0-9]", "")

Number of rows: INSERT 4 UPDATE 0 DELETE 0 UPSERT 0 LOOKUP 0 ERROR 0 TOTAL 4

Refresh Typecast Modify Map drifted Statistics Remove Export to CSV

SubCategoryKey	CategoryKey	SubCategory Name	DcategoryKey
1	ID -1	Extra	1
2	ID - 2	Regular	2
3	ID - 1	Micro	1
4	ID - 2	Super	2

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Search factory and documentation

Data Factory Validate Data flow debug Debug Settings

SubCategoriesData_DF GeographyData ProductData RawGeographyData RawProductData SalesRep_DF SubCategorieData...

Preview experience off

Properties General Related

Name: SubCategorieData_DF

Description

Derived column's settings Optimize Inspect Data preview

Number of rows: + INSERT 4 * UPDATE 0 × DELETE

Derived Column Select (green arrow)

Aggregate Surrogate Key Pivot Unpivot Window Rank External Call Cast

Formatters

DerivedColumn1: Columns: 4 total

SubCategoryKey CategoryKey

1 ID -1
2 ID -2
3 ID -1
4 ID -2

Schema modifier

0 ERROR 0 TOTAL 4

Activate Windows Go to Settings to activate Windows.

source1 Import data from SubCategoriesData

derivedColumn1 Creating/updating the columns 'SubCategoryKey', 'CategoryKey', 'SubCategory Name', 'DcategoryKey'

select1 Columns: 3 total

Select settings Optimize Inspect Data preview

Incoming stream: derivedColumn1

Options: Skip duplicate input columns (checked), Skip duplicate output columns (checked)

Input columns:

- derivedColumn1's column
- SubCategoryKey → SubCategoryKey
- CategoryKey → CategoryKey (green arrow)
- SubCategory Name → SubCategory Name
- DcategoryKey → CategoryKey (green arrow)

4 mappings: 1 column(s) from the inputs left unmapped

Reference: 1

derivedColumn1 Creating/updating the columns 'SubCategoryKey', 'CategoryKey', 'SubCategory Name', 'DcategoryKey'

select1 Columns: 3 total

Select settings Optimize Inspect Data preview

Number of rows: + INSERT 4 * UPDATE 0 × DELETE 0 * UPINSERT 0 LOOKUP 0 × ERROR 0 TOTAL 4

Refresh Typecast Modify Map drifted Statistics Remove Export to CSV

SubCategoryKey	SubCategory Name	CategoryKey
1	Extra	1
2	Regular	2
3	Micro	1
4	Super	2

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Search factory and documentation

Preview experience off

letshunvarthy@outlook.com DEFAULT DIRECTORY

Data Factory SubCategorieData_DF GeographyData ProductData RawGeographyData RawProductData SalesRep_DF SubCategorieData... Validate Data flow debug Debug Settings

Properties General Related Name SubCategorieData_DF Description

Select settings Optimize Inspect Data preview

Incoming stream * derivedColumn1 Options Skip duplicate input columns Skip duplicate output columns

Input columns + Add mapping Delete 3 mappings: 1 co

derivedColumn1's column Name as SubCategoryKey SubCategory Name CategoryKey

Schema modifier Derived Column Select Aggregate Surrogate Key Pivot Unpivot Window Rank External Call Cast Formatters

Aggregate 2

Activate Windows Go to Settings to activate Windows.

source1 Import data from SubCategoriesData

derivedColumn1 Creating/updating the columns SubCategoryKey, CategoryKey, SubCategory Name, DcategoryKey

select1 Renaming derivedColumn1 to select1 with columns SubCategoryKey, SubCategory Name,

aggregate1 Columns: 3 total

Aggregate settings Optimize Inspect Data preview

Output stream name * aggregate1 Learn more

Description Add aggregate columns Reset

Incoming stream * select1

Group by Aggregates

Columns Name as

- SubCategoryKey SubCategoryKey
- CategoryKey CategoryKey
- SubCategory Name SubCategory Name

Aggregate settings Optimize Inspect Data preview

Description Aggregating data by 'SubCategoryKey', 'CategoryKey', 'SubCategory Name' producing columns 'countRecords'

Incoming stream * select1

Group by Aggregates

Grouped by: SubCategoryKey, CategoryKey, SubCategory Name

+ Add Clone Delete Open expression builder

Column	Expression
countRecords	count(SubCategoryKey)

1 2 3

Aggregate settings Optimize Inspect **Data preview**

Number of rows + **INSERT 4** * **UPDATE 0** x **DELETE 0** * **UPSERT 0** **LOOKUP 0** ✘ **ERROR 0** **TOTAL 4**

Refresh | Typecast | Modify | Map drifted | Statistics | Remove | Export to CSV |

↑↓	SubCategoryKey	12s ↑↓	CategoryKey	abc ↑↓	SubCategory Name	abc ↑↓	countRecords	12l ↑↓
+	1		1		Extra		1	
+	2		2		Regular		1	
+	3		1		Micro		1	
+	4		2		Super		1	

Microsoft Azure | Data Factory | projects-datafactory | Search factory and documentation

Data Factory Validate all Publish all Preview experience Off

SubCategroiesData_DF GeographyData ProductData RawGeographyData SalesRep_DF SubCategorieData...

Validate Data flow debug Debug Settings

Properties

General Related

Name: SubCategorieData...

Description:

Activate Windows Go to Settings to activate Windows.

Select settings Optimize Inspect Data preview

Skip duplicate input columns

Skip duplicate output columns

Auto mapping Reset Add mapping Delete 4 mappings: All inputs mapped

aggregate1's column	Name as
12s SubCategoryKey	SubCategoryKey
abc CategoryKey	CategoryKey
abc SubCategory Name	SubCategory Name
12l countRecords	countRecords

Select settings Optimize Inspect **Data preview**

Number of rows + **INSERT 4** * **UPDATE 0** x **DELETE 0** * **UPSERT 0** **LOOKUP 0** ✘ **ERROR 0** **TOTAL 4**

Refresh | Typecast | Modify | Map drifted | Statistics | Remove | Export to CSV |

↑↓	SubCategoryKey	12s ↑↓	CategoryKey	abc ↑↓	SubCategory Name	abc ↑↓
+	1		1		Extra	
+	2		2		Regular	
+	3		1		Micro	
+	4		2		Super	

Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Data Factory Validate all Publish all

Preview experience Off

SubCategroiesData_DF

Properties General Related

Name SubCategroiesData_DF

Description

Activate Windows Go to Settings to activate Windows.

The screenshot shows the Microsoft Azure Data Factory Data Flow blade. It displays a pipeline with three stages: 'select1', 'aggregate1', and 'select2'. The 'select2' stage is currently selected. In the properties pane on the right, the 'Sink' operation is highlighted with a green arrow labeled '2'. The pipeline data preview shows four rows of data with columns: SubCategoryKey, CategoryKey, and SubCategory Name. The 'sink1' stage is also visible at the end of the pipeline.

Sink Settings Errors Mapping Optimize Inspect Data preview

Description Add sink dataset

Incoming stream * select2

Sink type * Dataset

Dataset * Select... + New

Options Allow schema drift Validate schema

This screenshot shows the 'Sink' configuration pane within the Data Flow blade. It includes tabs for Sink, Settings, Errors, Mapping, Optimize, and Inspect. Under the Sink tab, there is a 'Description' field and an 'Incoming stream' dropdown set to 'select2'. A 'Sink type' section shows 'Dataset' selected. Below it, a 'Dataset' dropdown is highlighted with a red box, and a green arrow labeled '1' points to the '+ New' button. There are also 'Options' settings for 'Allow schema drift' (checked) and 'Validate schema'.

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

tegoriesData_DF GeographyData ProductData RawGeographyData

Validate Data flow debug Debug Settings

select1 aggregate1 select2

Renameing derivedColumn1 to select1 with columns SubCategoryKey, SubCategory Name.

Aggregating data by SubCategoryKey, CategoryKey, SubCategory Name producing columns

Renameing aggregate1 to select2 with columns SubCategoryKey, CategoryKey, SubCategory

Sink Settings Errors Mapping Optimize Inspect Data preview

Description Add sink dataset

Incoming stream * select2

Sink type * Dataset Inline Cache

Dataset * Select... New

Options Allow schema drift Validate schema

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

Amazon S3 Azure Blob Storage Azure Cosmos DB for NoSQL

Azure Data Explorer (Kusto) Azure Data Lake Storage Gen2 Azure Database for MySQL

Activate Windows Go to Settings to activate Windows Cancel

Continue ← 2

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

tegoriesData_DF GeographyData ProductData RawGeographyData

Validate Data flow debug Debug Settings

select1 aggregate1 select2

Renameing derivedColumn1 to select1 with columns SubCategoryKey, SubCategory Name.

Aggregating data by SubCategoryKey, CategoryKey, SubCategory Name producing columns

Renameing aggregate1 to select2 with columns SubCategoryKey, CategoryKey, SubCategory

Sink Settings Errors Mapping Optimize Inspect Data preview

Description Add sink dataset

Incoming stream * select2

Sink type * Dataset Inline Cache

Dataset * Select... New

Options Allow schema drift Validate schema

Select format

Choose the format type of your data

Avro DelimitedText JSON

ORC Parquet Binary

Activate Windows Go to Settings to activate Windows Cancel

Continue ← 2 Back

Set properties

Name

StagingSubCategoriesData

Linked service *

Select... Filter... + New blob2adf_ls ← 2

Set properties

Name
StagingSubCategoriesData

Linked service *
blob2adf_ls

File path
Container / Directory / File name

First row as header

Import schema
 From connection/store From sample file None

> Advanced

Browse

Select a file or folder.

Root folder

Microsoft Azure | Data Factory > projects-datafactory

tegoriesData_DF

select1

Sink Settings Errors Mapping Optimize Inspect Data preview

Incoming stream * select2

Sink type *

Dataset * Select...

Options

Browse
Select a file or folder.
Root folder > stagingdenisdata
Showing 0 items

Activate Windows
Go to Settings to activate Windows.

Step 1: Publish Data Flow

The screenshot shows the Azure Data Factory Data Flow blade. A complex pipeline is displayed with four main stages: 'select1', 'aggregate1', 'select2', and 'sink1'. The 'sink1' stage is configured to export data to a dataset named 'StagingSubCategoriesData'. The 'Publish all' button at the top is highlighted with a green arrow. To the right, the 'Properties' panel shows the dataset name 'SubCategorieisData_DF'.

Step 2: Publish Pending Changes

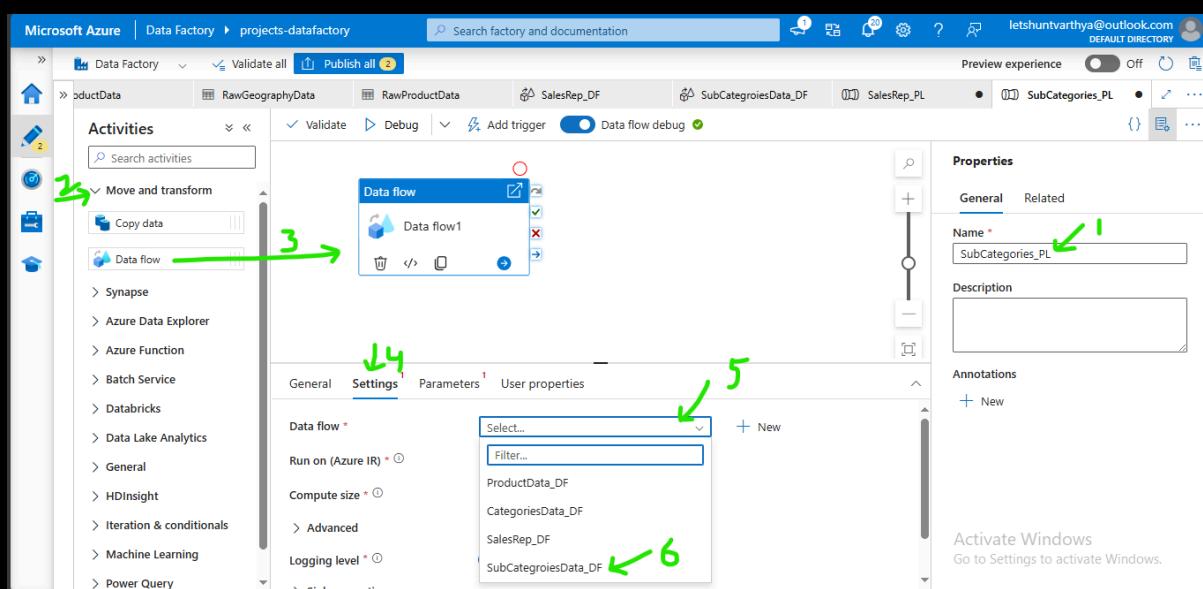
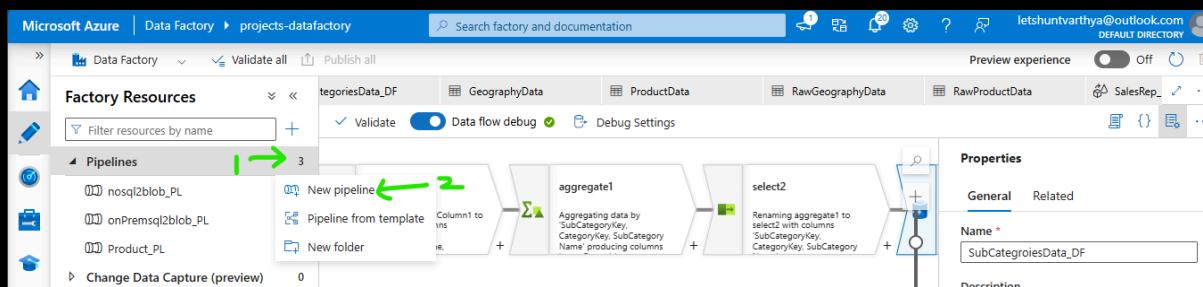
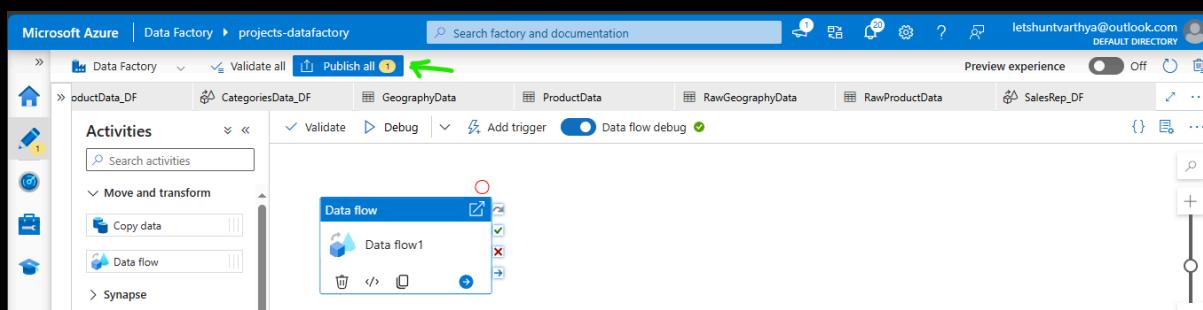
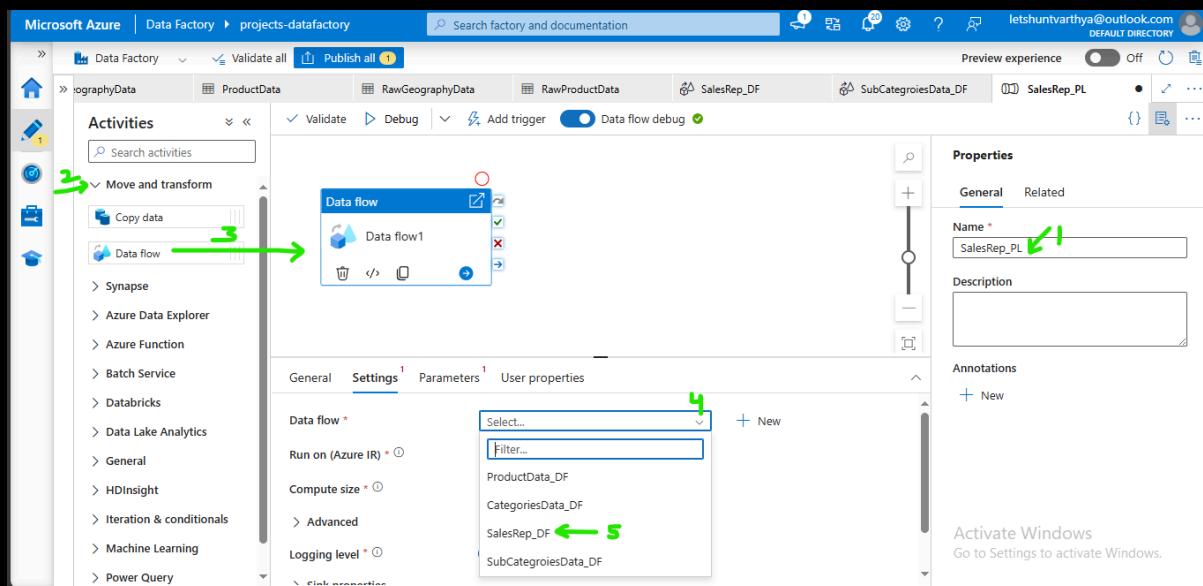
The screenshot shows the 'Publishing' blade. It lists pending changes for datasets and data flows. Under 'Pending changes (2)', there are entries for 'StagingSubCategoriesData' (New) and 'SubCategorieisData_DF' (New). The 'Publish' button at the bottom is highlighted with a green arrow.

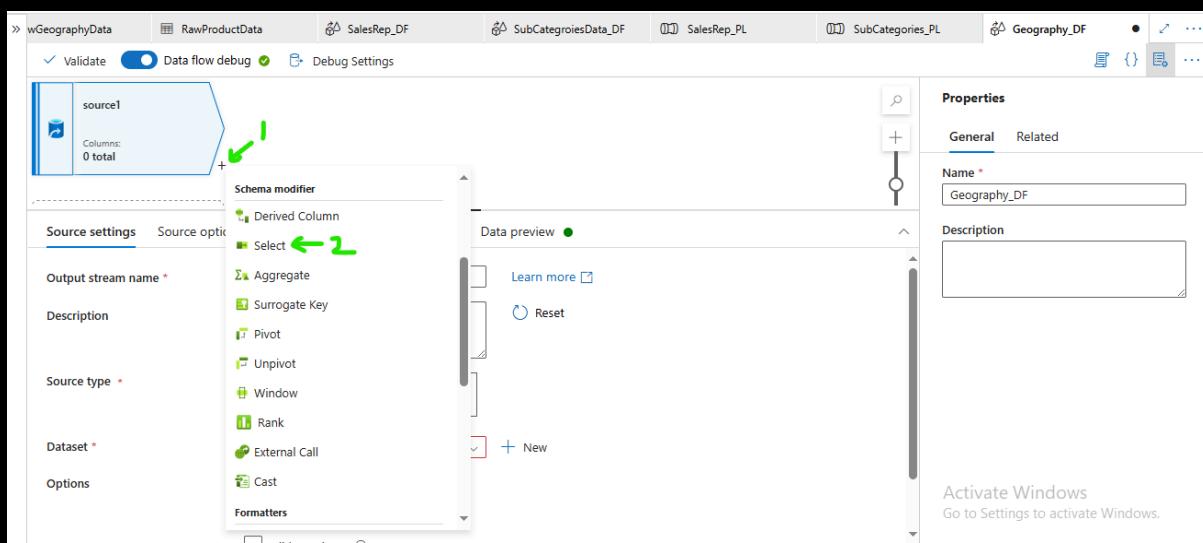
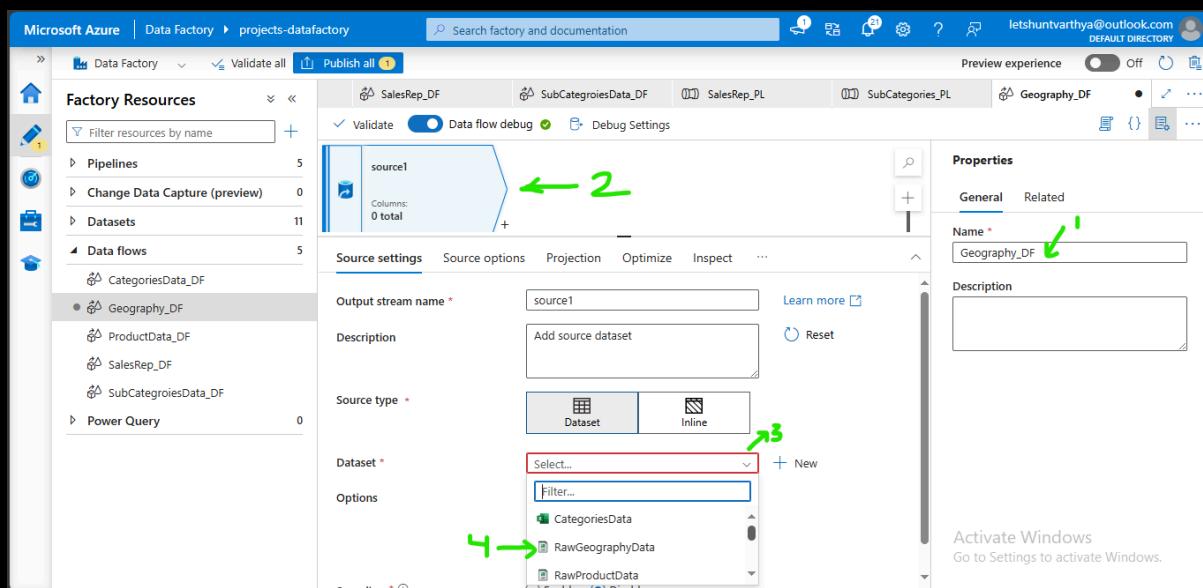
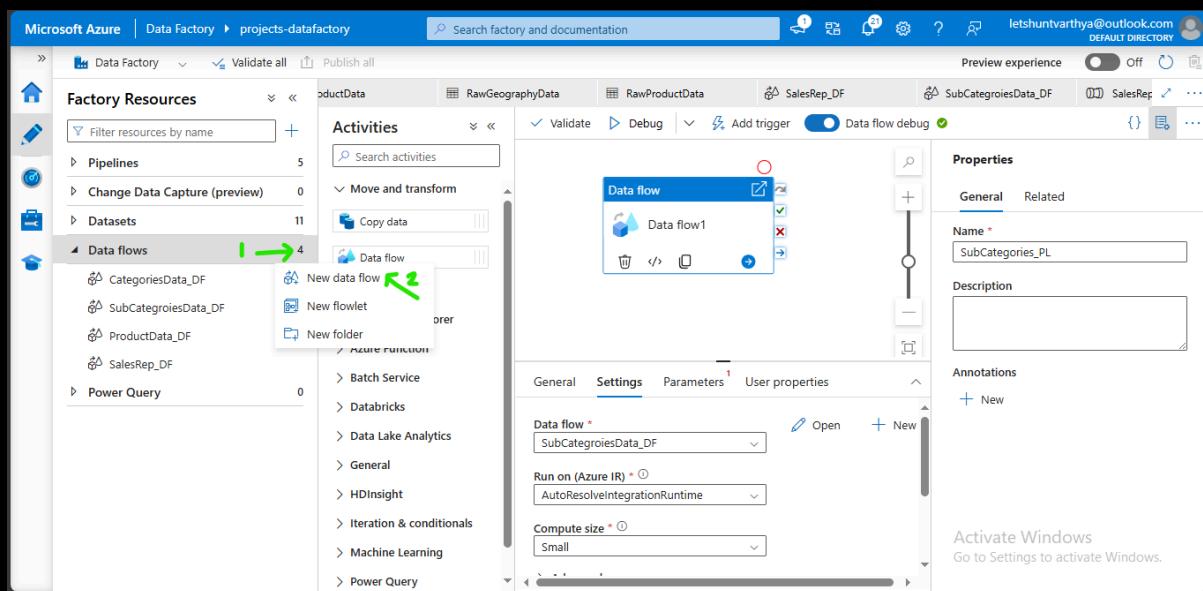
Step 3: Create a New Pipeline

The screenshot shows the 'Factory Resources' blade with the 'Pipelines' section selected. A green arrow points to the 'New pipeline' button, which is highlighted with a green arrow.

Step 4: Configure Pipeline Settings

The screenshot shows the 'Activities' blade. The 'Data flow' activity is selected and highlighted with a green arrow. The properties panel on the right shows the pipeline name 'SalesRep_PL'.





Source settings Source options **Projection** Optimize Inspect Data preview

Define default format Detect data type Import projection Reset schema

Column name	Type	Format
Country	abc string	Specify format
Town	abc string	Specify format
Wikipedia	abc string	Specify format

Select settings Optimize Inspect Data preview

am * source1

Skip duplicate input columns

Skip duplicate output columns

Auto mapping Reset Add mapping Delete 3 mappings: All inputs mapped

source1's column	Name as
Country	Country
Town	Town
Wikipedia	Wikipedia

Select settings Optimize Inspect **Data preview**

Number of rows + INSERT 7 * UPDATE 0 × DELETE 0 + UPSERT 0 ? LOOKUP 0 ✎ ERROR 0 TOTAL 7

Refresh | Typecast | Modify | Map drifted | Statistics | Remove | Export to CSV |

Country	Town
Czech republic	Prague
Denmark	Copenhagen
Germany	Frankfurt
Germany	Berlin
Germany	Dresden
Czech republic	Brno

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Search factory and documentation

Preview experience Off

Data Factory wGeographyData RawProductData SalesRep_DF SubCategoriessData_DF SalesRep_PL SubCategories_PL Geography_DF

Validate Data flow debug Debug Settings

source1 Import data from RawGeographyData select1 Columns: 2 total

Lookup Schema modifier

- Derived Column
- Select
- Aggregate
- Surrogate Key
- Pivot
- Unpivot
- Window
- Rank
- External Call
- Cast

Properties General Related

Name * Geography_DF

Description

Number of rows + INSERT 7 * UPDATE 0 × DELETE 0

Refresh Typecast Modify Map drifted Statistics Remove Export to CSV

to CSV

Activate Windows Go to Settings to activate Windows.

Number of rows + INSERT 7 * UPDATE 0 × DELETE 0

Refresh Typecast Modify Map drifted Statistics Remove Export to CSV

to CSV

Activate Windows Go to Settings to activate Windows.

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Search factory and documentation

Preview experience Off

Data Factory wGeographyData RawProductData SalesRep_DF SubCategoriessData_DF SalesRep_PL SubCategories_PL Geography_DF

Validate Data flow debug Debug Settings

source1 Import data from RawGeographyData select1 Renaming source1 to select1 with column: 'Country, Town'

surrogateKey1 Columns: 3 total

Surrogate key settings Optimize Inspect Data preview

Output stream name * surrogateKey1 Learn more

Description Adding new key GeographyID starting from 1 with step 1

Incoming stream * select1

Key column * GeographyID

Start value * 1

Step value 1

Properties General Related

Name * Geography_DF

Description

Number of rows + INSERT 7 * UPDATE 0 × DELETE 0

Refresh Typecast Modify Map drifted Statistics Remove Export to CSV

to CSV

Activate Windows Go to Settings to activate Windows.

Number of rows + INSERT 7 * UPDATE 0 × DELETE 0

Refresh Typecast Modify Map drifted Statistics Remove Export to CSV

to CSV

Activate Windows Go to Settings to activate Windows.

Surrogate key settings Optimize Inspect Data preview

Number of rows + INSERT 7 * UPDATE 0 × DELETE 0

Number of rows + UPSERT 0

Number of rows + LOOKUP 0

Number of rows + ERROR 0

Number of rows + TOTAL 7

Refresh Typecast Modify Map drifted Statistics Remove Export to CSV

to CSV

Activate Windows Go to Settings to activate Windows.

Country	Town	GeographyID
Czech republic	Prague	1
Denmark	Copenhagen	2
Germany	Frankfurt	3
Germany	Berlin	4
Germany	Dresden	5
Czech republic	Brno	6
Czech republic	Ostrava	7

Number of rows + TOTAL 7

Refresh Typecast Modify Map drifted Statistics Remove Export to CSV

to CSV

Activate Windows Go to Settings to activate Windows.

Microsoft Azure | Data Factory | projects-datafactory | Validate all | Publish all 1 | Search factory and documentation | Preview experience Off

wGeographyData | RawProductData | SalesRep_DF | SubCategoriessData_DF | SalesRep_PL | SubCategories_PL | Geography_DF

Validate | Data flow debug | Debug Settings

source1 | select1 | surrogateKey1 | select2

Renaming source1 to select1 with columns 'Country, Town'

surrogate key settings | Optimize | Inspect | Data preview

Number of rows: INSERT 7, UPDATE 0, DELETE 0, UPSERT 0

	Country	Town	GeographyID
+	Czech republic	Praque	1
+	Denmark	Copenhagen	2
+	Germany	Frankfurt	3
+	Germany	Berlin	4
+	Germany	Dresden	5
+	Czech republic	Brno	6
+	Czech republic	Ostrava	7

Properties

Name: Geography_DF

Description:

Activate Windows

source1 | select1 | surrogateKey1 | select2

Renaming source1 to select1 with columns 'Country, Town'

Adding new key GeographyID starting from 1 with step 1

surrogateKey1 | select2

Columns: 3 total

Select settings | Optimize | Inspect | Data preview

columns: 'GeographyID, Country, Town'

surrogateKey1

Skip duplicate input columns

Skip duplicate output columns

Auto mapping

Reset | Add mapping | Delete | 3 mappings: All inputs mapped

surrogateKey1's column	Name as
121 GeographyID	GeographyID
abc Country	Country
abc Town	Town

select1 | surrogateKey1 | select2

Renaming source1 to select1 with columns 'Country, Town'

Adding new key GeographyID starting from 1 with step 1

select2 | flattener

Columns: 3 total

Properties

Name: Geography_DF

Description:

Activate Windows

Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Data Factory Validate all Publish all 1

wGeographyData RawProductData SalesRep_DF SubCategroiesData_DF

Validate Data flow debug Debug Settings

select1 Renaming source1 to select1 with columns 'Country, Town' + surrogateKey1 Adding new key GeographyID starting from 1 with step 1 + select2 Renaming surrogateKey1 to select2 with columns GeographyID, Country, Town

Sink Settings Errors Mapping Optimize Inspect Data preview

Description Add sink dataset

Incoming stream * select2

Sink type * Dataset

Dataset * Select... New

Options Allow schema drift Validate schema

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

Amazon S3 Azure Blob Storage Azure Cosmos DB for NoSQL

Azure Data Explorer (Kusto) Azure Data Lake Storage Gen2 Azure Database for MySQL

Activate Windows Go to Settings to activate Windows Cancel

2 3

Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Data Factory Validate all Publish all 1

wGeographyData RawProductData SalesRep_DF SubCategroiesData_DF

Validate Data flow debug Debug Settings

select1 Renaming source1 to select1 with columns 'Country, Town' + surrogateKey1 Adding new key GeographyID starting from 1 with step 1 + select2 Renaming surrogateKey1 to select2 with columns GeographyID, Country, Town

Sink Settings Errors Mapping Optimize Inspect Data preview

Description Add sink dataset

Incoming stream * select2

Sink type * Dataset

Dataset * Select... New

Options Allow schema drift Validate schema

Select format

Choose the format type of your data

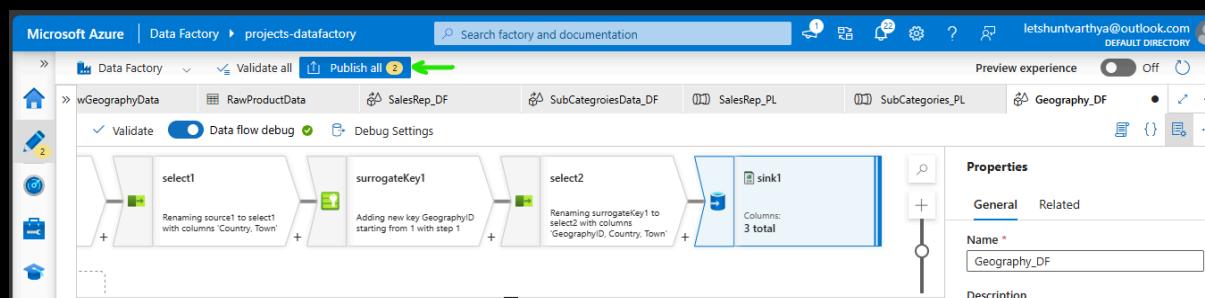
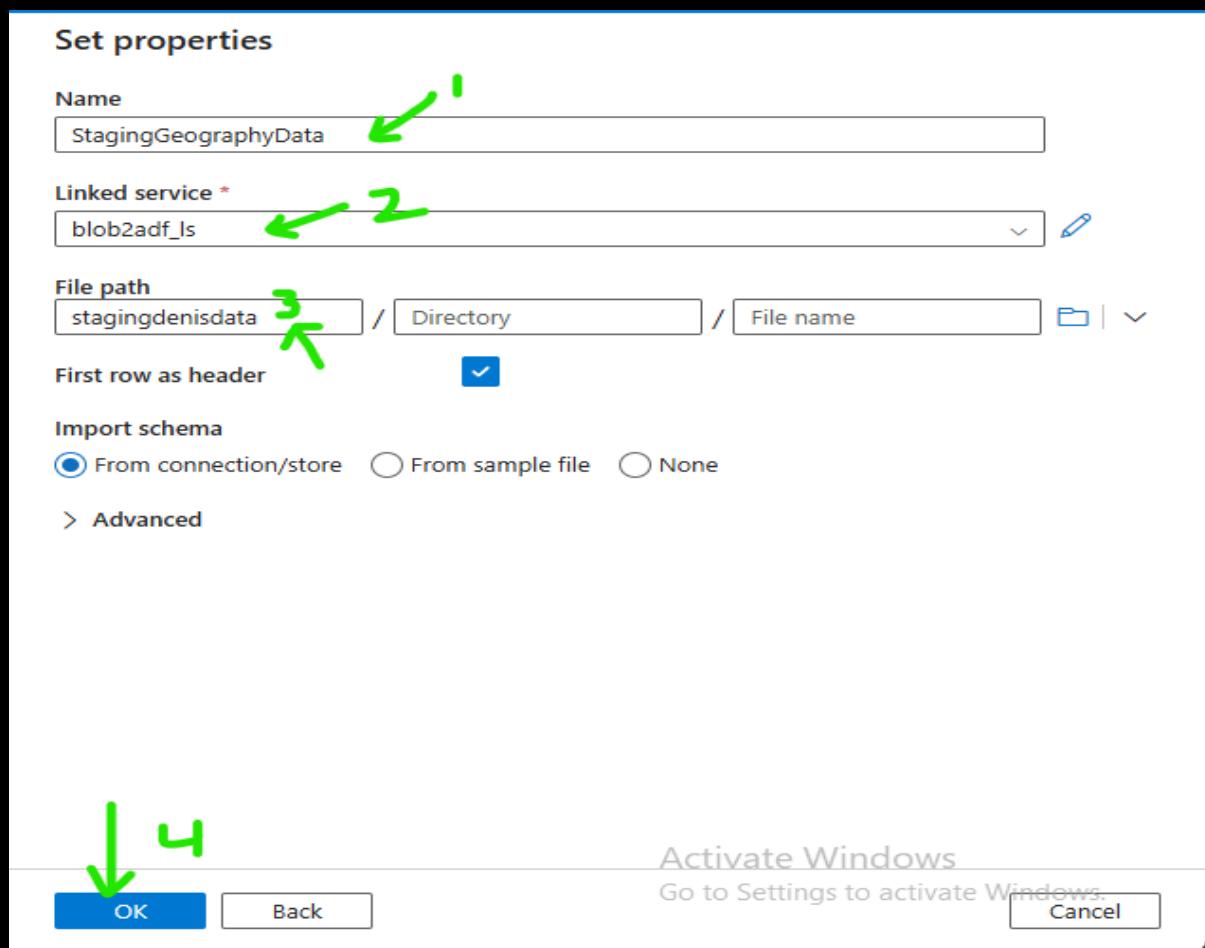
Avro CSV JSON

DelimitedText

ORC Parquet Binary

Activate Windows Go to Settings to activate Windows Cancel

2 1



Microsoft Azure | Data Factory > projects-datafactory

Publishing 2

Validate all Publishing

Data Factory

wGeographyData RawProductData SalesRep_DF SubCategorieData_DF

✓ Validate Data flow debug Debug Settings

Publish all

You are about to publish all pending changes to the live environment. [Learn more](#)

Pending changes (2)

NAME	CHANGE	EXISTING
Datasets		
StagingGeographyData	(New)	-
Data flows		
Geography_DF	(New)	-

Description: Export data to StagingGeographyData

Incoming stream: select2

Sink type: Dataset

Dataset: StagingGeographyData

Skip line count:

Options: Allow schema drift

Activate Windows Go to Settings to activate Windows.

Publish **Cancel**

Microsoft Azure | Data Factory > projects-datafactory

Publish all

Factory Resources

wGeographyData RawProductData SalesRep_DF SubCategorieData_DF SalesRep_PL SubCate ...

Filter resources by name +

✓ Validate Data flow debug Debug Settings

Pipelines 5

New pipeline 2

Change Data Capture (preview)

Datasets

Data flows

CategoriesData_DF

Properties

General Related

Name: Geography_DF

Description:

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Data flow

Activities

Move and transform

Copy data

Data flow 3

Synapse

Azure Data Explorer

Azure Function

Batch Service

Databricks

Data Lake Analytics

General

HDInsight

Iteration & conditionals

Machine Learning

Power Query

Properties

General Related

Name: Geography_PL

Description:

Annotations

Activate Windows Go to Settings to activate Windows.

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Data flow

Activities

Move and transform

Copy data

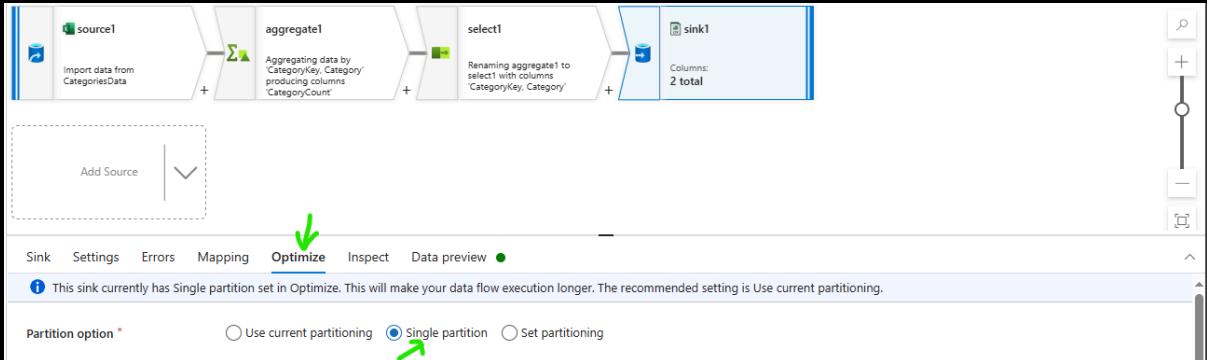
Data flow

Properties

General Related

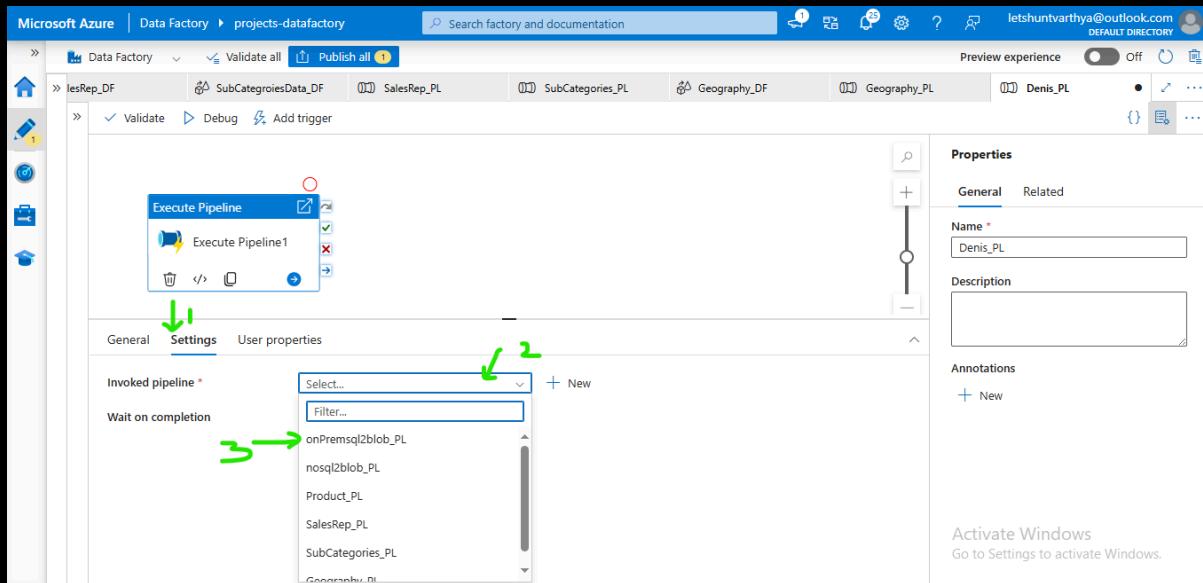
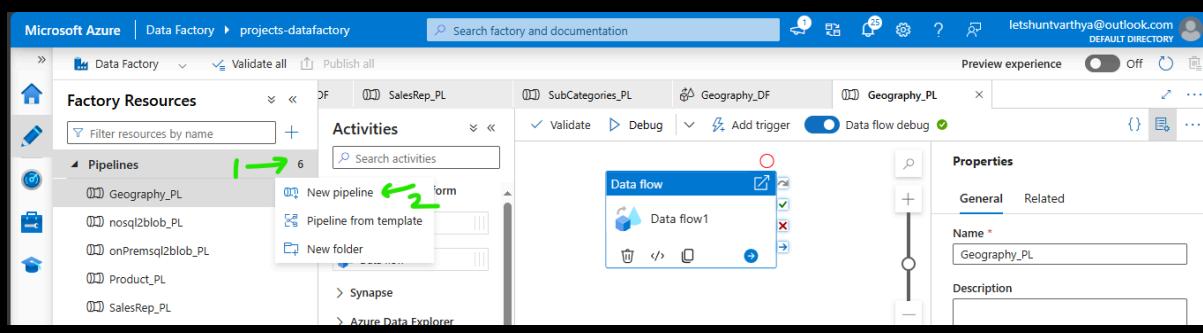
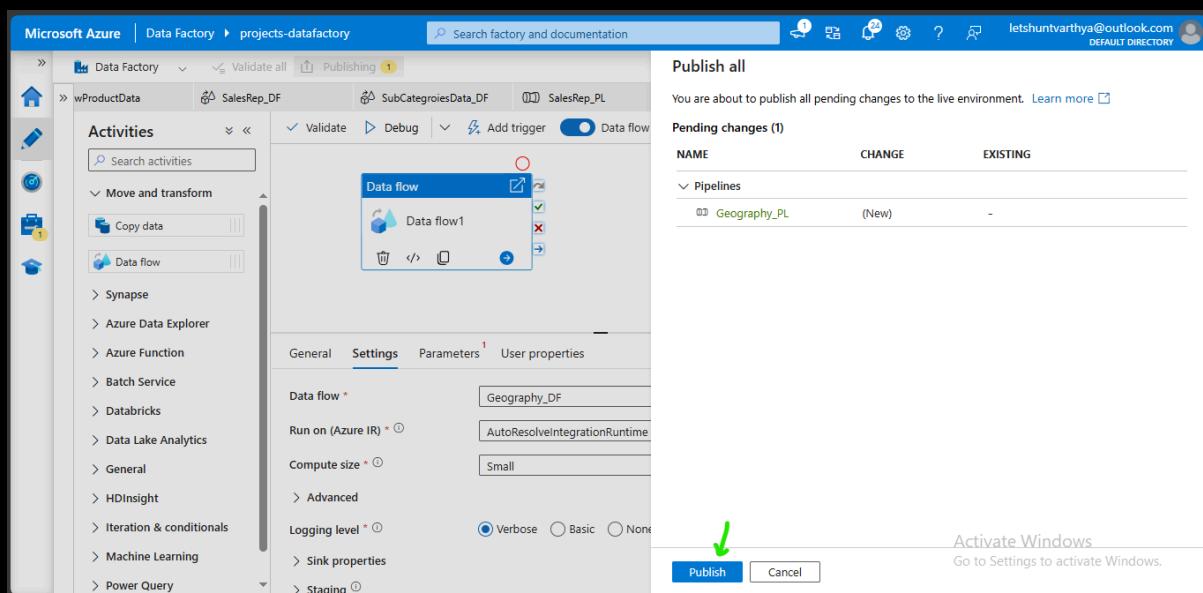
Name: Geography_PL

- Set the **partition option** to **single partition** in **optimise** which is in the **sink** of all DataFlow (ProductData_DF , SalesRep_DF, CategoriesData_DF, SubCategoriesData_DF, Geography_DF) .

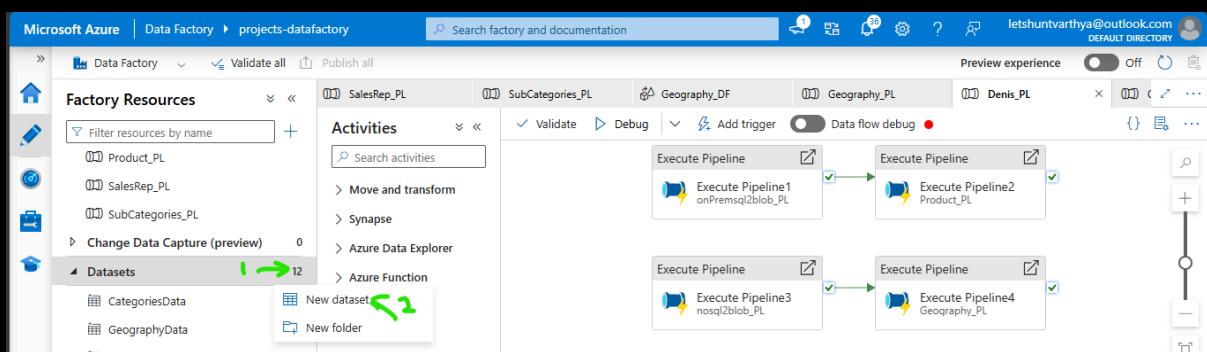
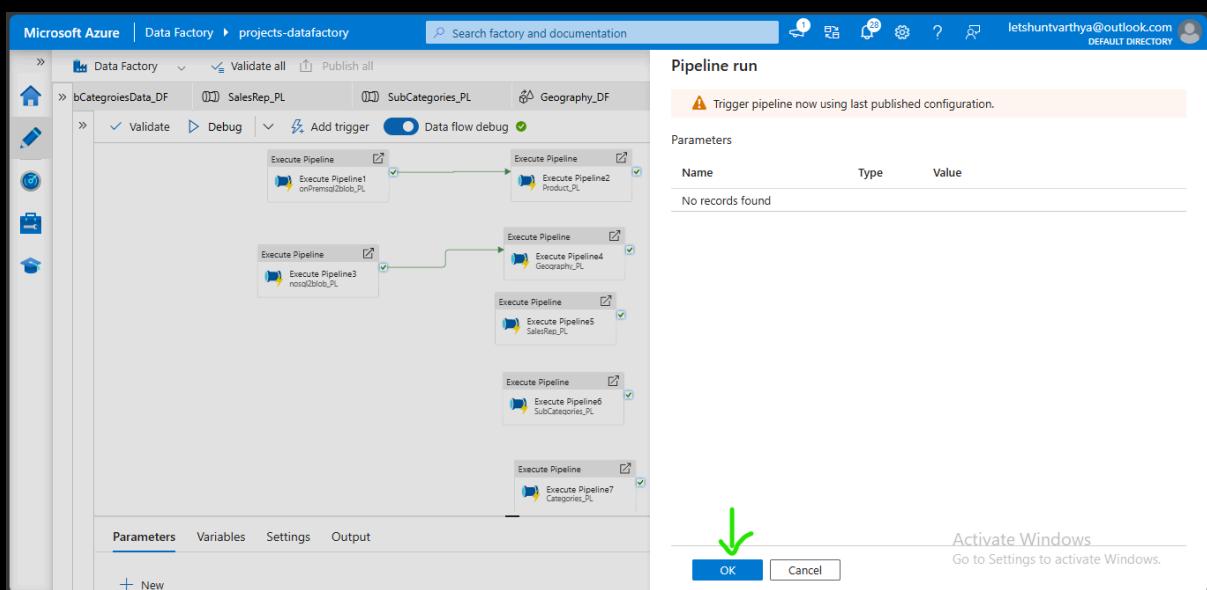
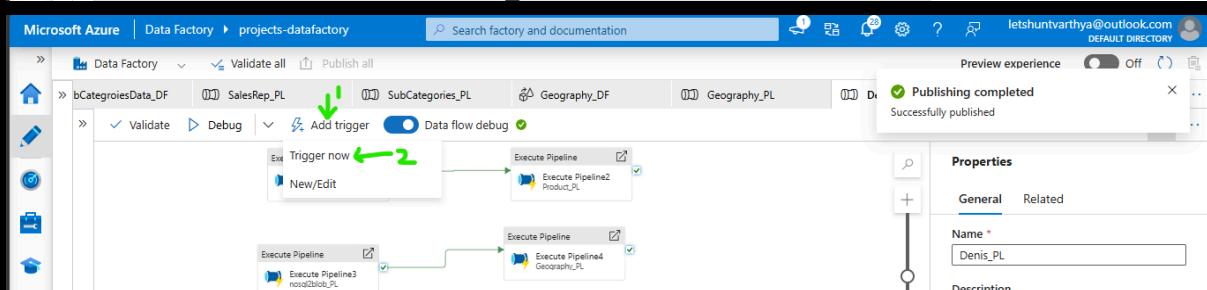
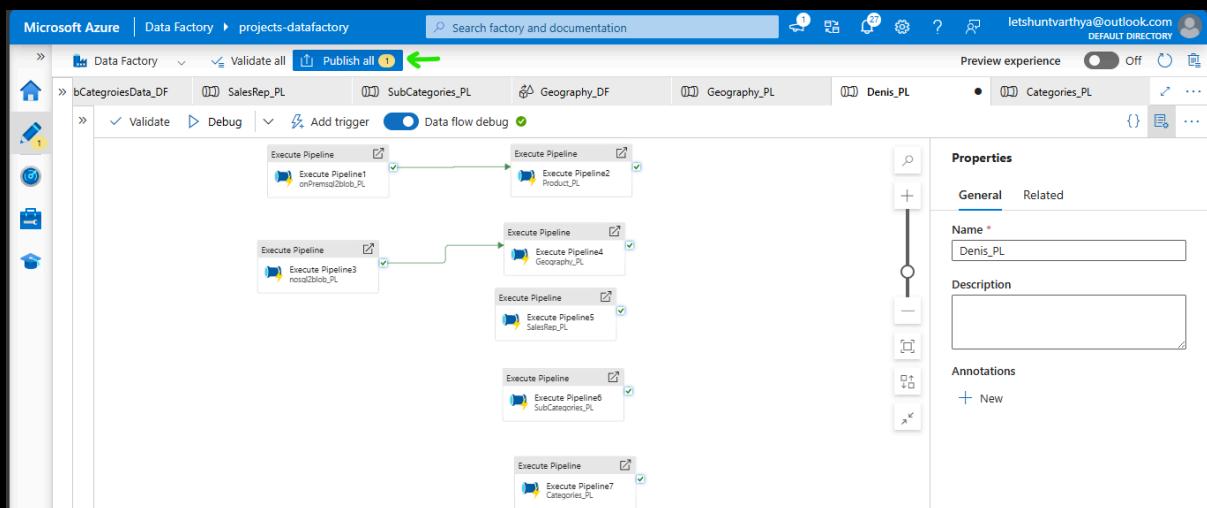


- Select **File name Option** as **Pattern** and Provide the name as (**StagingSubCategoriesData.csv,StagingCategories.csv,StagingProduct.csv,StagingSalesRep.csv** and **StagingGeographyData.csv**) for filename in pattern which is in the **DataFlows(SubCategories_DF, Categories_DF, Product_DF, SalesRep_DF and Geography_DF)**.

The screenshot shows the Power BI Data Factory Resources pane. On the left, under the "Data flows" section, several data flows are listed: CategoriesData_DF, Geography_DF, ProductData_DF, SalesRep_DF, and SubCategoriesData_DF. Each of these data flows is highlighted with a green arrow. On the right, the "Settings" tab for a specific data flow (Geography_DF) is selected. It shows a data flow diagram with steps: source1, select1, surrogateKey1, select2, and sink1. The sink step (sink1) is highlighted with a blue border. Below the diagram, the "File name option" section is selected, indicated by a green arrow pointing to it. The "Pattern" field contains the value "StagingGeographyData.csv", highlighted with a green arrow. Other settings shown include "Clear the folder" (unchecked), "Quote All" (unchecked), and "Headers" (set to "Enter expression... ANY"). A tooltip at the bottom right says: "Activate Windows Go to Settings to activate Windows."



- Repeat this process for the rest of the pipelines as in the image below and **publish** it.



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	CSV
Excel	JSON	ORC
XML		

[Continue](#) [Back](#)

Set properties

Name 1

Linked service 2

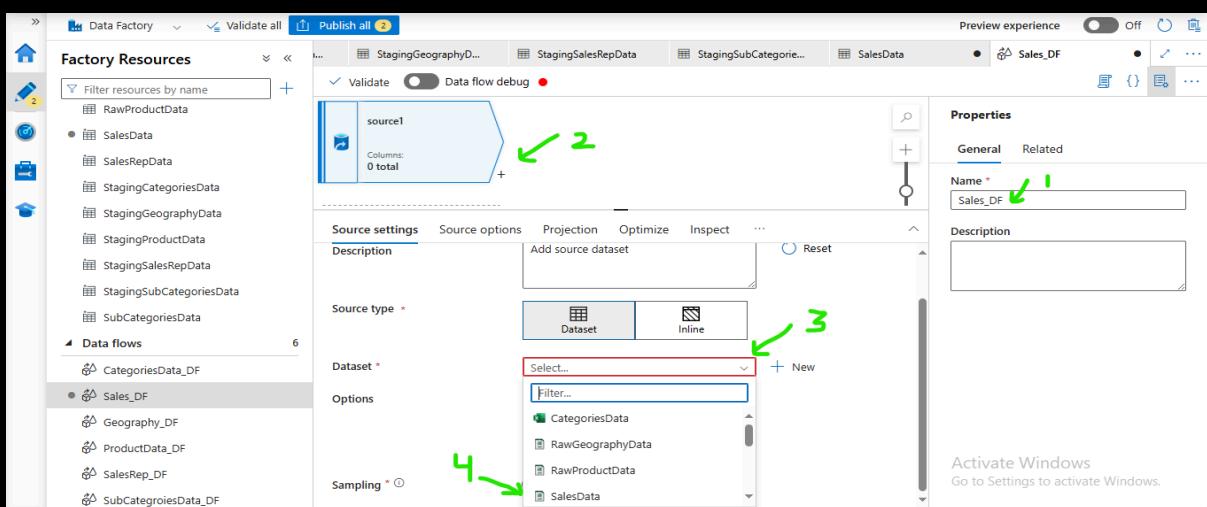
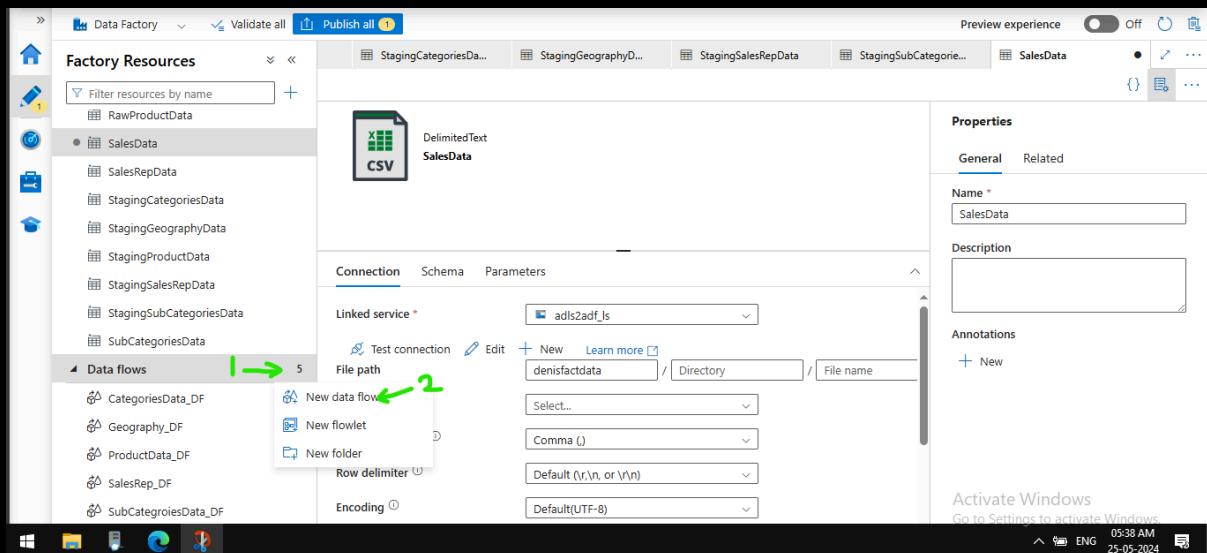
File path / / 3

First row as header

Import schema

From connection/store From sample file None 4

[OK](#) [Back](#)



Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all

Preview experience Off

Properties

- General Related
- Name * Sales_DF
- Description

Activate Windows Go to Settings to activate Windows.

Data preview

Number of rows: 100

fSalesPri...	ProductID	SalesRepID	Location	Date	Units
36943	10	7	Germany...	1.10.2017	60
36944	8	2	Denmark...	26.4.2017	71
36945	9	6	Germany...	28.9.2017	68
36946	2	5	Germany...	16.4.2017	157
36947	8	2	Denmark...	13.2.2017	103

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all

Preview experience Off

Properties

- General Related
- Name * Sales_DF
- Description

Activate Windows Go to Settings to activate Windows.

Data preview

Number of rows: 100

Location	Date	Units
Germany...	1.10.2017	60
Denmark...	26.4.2017	71
Germany...	28.9.2017	68
Germany...	16.4.2017	157
Denmark...	13.2.2017	103

Multiple inputs/outputs

- Join
- Conditional Split
- Exists
- Union
- Lookup
- Derived Column
- Select
- Aggregate

Microsoft Azure | Data Factory > projects-datafactory

Validate all Data flow debug

Properties

- General Related
- Name * Sales_DF
- Description

Activate Windows Go to Settings to activate Windows.

Derived column'settings

Creating/updating the columns 'fSalesPrimaryKey, ProductID, SalesRepID, Location, Date, Units,'

Incoming stream: source1

Columns:

Column	Expression
country	regexSplit(Location,": \\n") [1]
cities	regexSplit(Location,": \\n") [2]
correctDate	toDate(toTimestamp(Date, 'dd.M.yyyy'))

The screenshot shows three views of a Data Flow pipeline in Azure Data Factory:

- Data preview view:** Shows the raw data from the SalesData source. A green arrow points up from the preview area towards the derived column settings.
- Derived column's settings view:** Shows the configuration for the "derivedColumn1". It includes a "Select" step (indicated by a green arrow) and a "Lookup" step. The "Properties" pane on the right shows the name is "Sales_DF".
- Select settings view:** Shows the mapping of columns from the source to the destination. The "Name as" column contains new names: "fSalesPrimaryKey", "ProductID", "SalesRepID", "Location", "Date", "Units", "PercentOfStandardCost", "RevenueDiscount", "country", "cities", and "correctDate". Green arrows point from the original column names to their new mappings.

Data Preview View (Top):

ID	abc	Date	abc	Units	abc	Percent...	abc	Revenue...	abc	country	abc	abc	correctDate
IV...		1.1.2017		31		0.954		0.25		Germany			2017-01-01
IV...		1.1.2017		35		0.954		0.25		Germany			2017-01-01
E...		2.1.2017		3		0.96		0		Czech re...			2017-01-02
K...		9.1.2017		29		0.956		0.25		Denmark			2017-01-09
E...		11.1.2017		68		0.967		0.35		Czech re...			2017-01-11
IV...		12.1.2017		77		0.962		0.4		Germany			2017-01-12

Derived column's settings view (Middle):

Select Step:

```

graph LR
    source1[Import data from SalesData] --> derivedColumn1[derivedColumn1  
Columns: 11 total]
    derivedColumn1 --> select1[select1  
Creating/updating the columns  
fSalesPrimaryKey, ProductID, SalesRepID, Location, Date, Units, PercentOfStandardCost, RevenueDiscount, country, cities, correctDate]
  
```

Properties pane (Middle):

Select settings view (Bottom):

Original Column	Name as
abc fSalesPrimaryKey	fSalesPrimaryKey
abc ProductID	ProductID
abc SalesRepID	SalesRepID
abc Location	Location
abc Date	Date
abc Units	Units
abc PercentOfStandardCost	PercentOfStandardCost
abc RevenueDiscount	RevenueDiscount
abc country	country
abc cities	cities
correctDate	Date

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all

Validate Data flow debug Debug Settings

Properties

General Related

Name * Sales_DF

Description

Activate Windows Go to Settings to activate Windows.

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all

Validate Data flow debug Debug Settings

Sink Settings Errors Mapping Optimize Inspect Data preview

Output stream name * sink1 Learn more

Description Add sink dataset

Incoming stream * select1

Sink type * Dataset

Dataset * Select... New

Options Allow schema drift

Continue ← 3 Activate Windows Go to Settings to activate Windows Cancel

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all

Validate Data flow debug Debug Settings

Sink Settings Errors Mapping Optimize Inspect Data preview

Output stream name * sink1 Learn more

Description Add sink dataset

Incoming stream * select1

Sink type * Dataset

Dataset * Select... New

Options Allow schema drift

Select format

Choose the format type of your data

Avro	CSV	JSON
DelimitedText	ORC	Parquet
JSON	Binary	

Continue ← 2 Activate Windows Go to Settings to activate Windows Cancel

Data Factory Pipeline Editor

Set properties

- Name:** StagingSalesData (marked with green arrow 1)
- Linked service:** adls2adf_ls (marked with green arrow 2)
- File path:** stagingdenisfactdata / Directory / File name (marked with green arrow 3)
- First row as header:**
- Import schema:**
 - From connection/store
 - From sample file
 - None
- Advanced**

OK **Back**

Sink Settings

- Clear the folder:**
- File name option:** Pattern (marked with green arrow 2)
- Pattern:** SatgingSalesData.csv (marked with green arrow 3)
- Quote All:**
- Headers:** Enter expression... (marked with green arrow 4)
- Umask:**

Owner	<input type="checkbox"/> R	<input type="checkbox"/> W	<input type="checkbox"/> X
Group	<input type="checkbox"/> R	<input checked="" type="checkbox"/> W	<input type="checkbox"/> X
Others	<input type="checkbox"/> R	<input checked="" type="checkbox"/> W	<input type="checkbox"/> X
Octal	1777		

Properties

- General**
- Name:** Sales_DF
- Description:** (empty)

Activate Windows
Go to Settings to activate Windows.

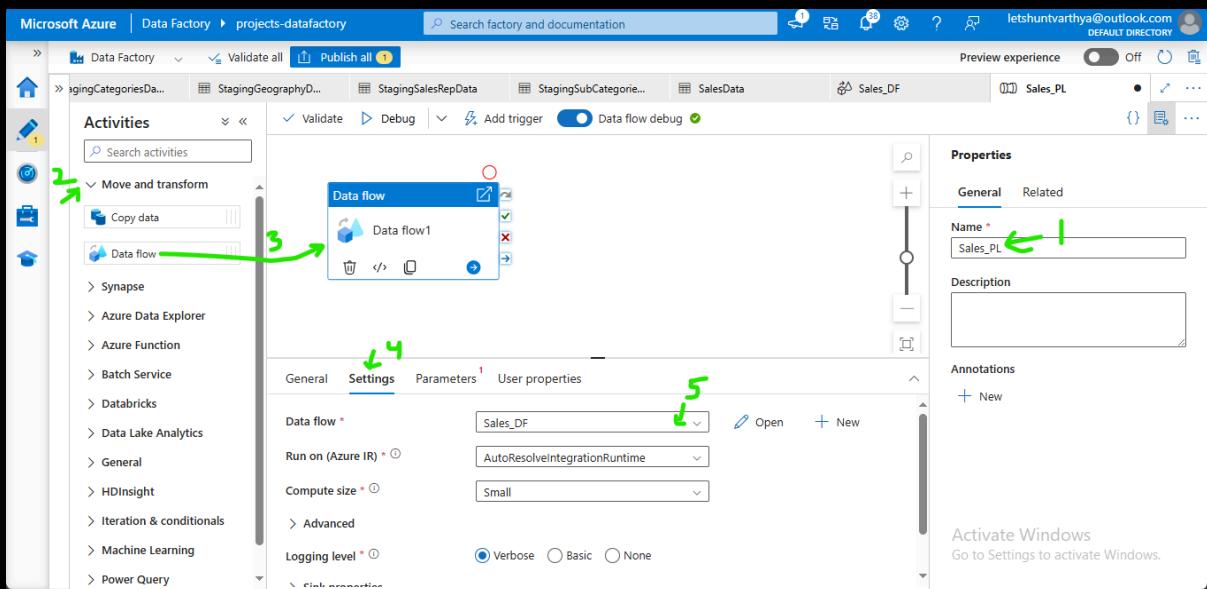
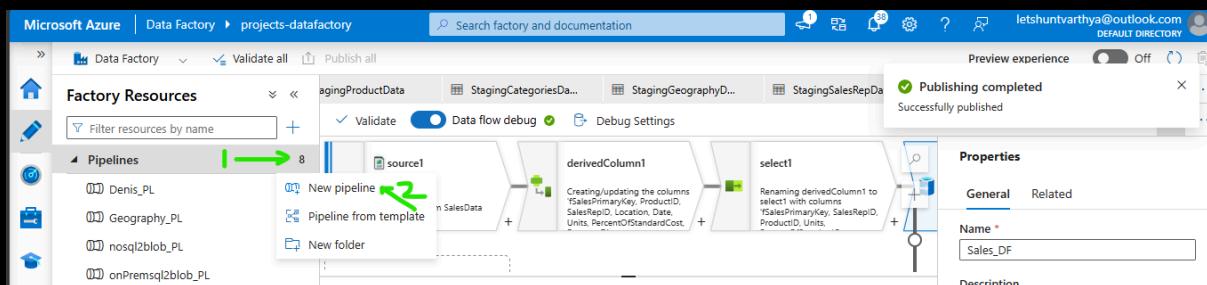
Publish all

You are about to publish all pending changes to the live environment. [Learn more](#)

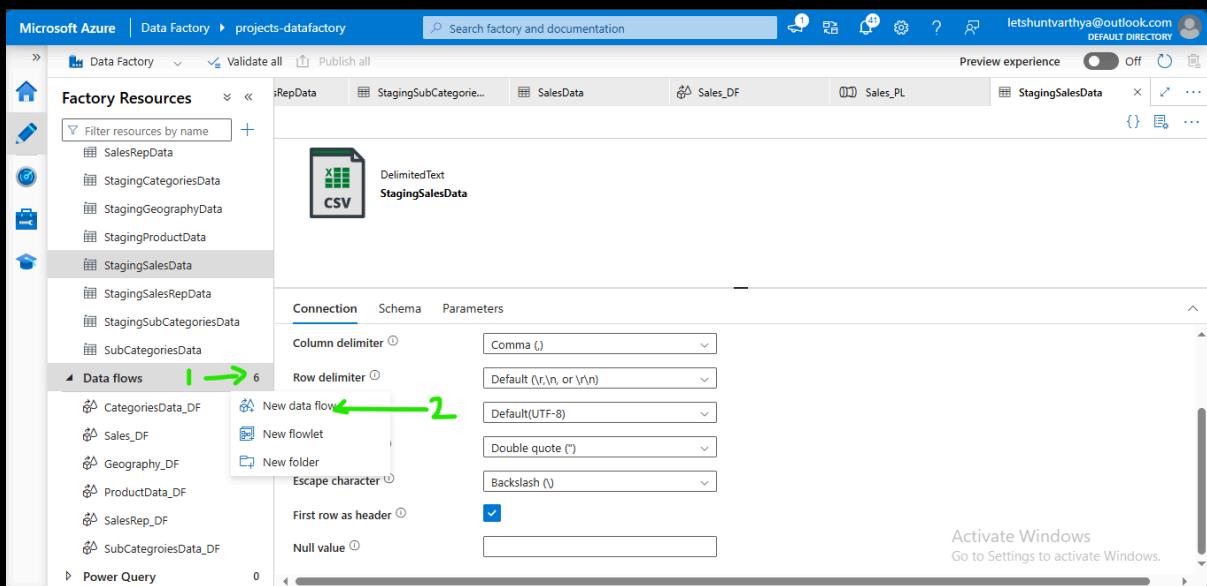
Pending changes (3)

NAME	CHANGE	EXISTING
Datasets		
SalesData	(New)	-
StagingSalesData	(New)	-
Data flows		
Sales_DF	(New)	-

Activate Windows
Go to Settings to activate Windows.



Publish



Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Search factory and documentation

Preview experience Off

Data Factory StagingSalesRepData SalesData Sales_DF Sales_PL StagingSalesData Denis_DF

Validate Data flow debug Debug Settings

Properties General Related

Name * Denis_DF

Description

Source settings Source options Projection Optimize Inspect Data preview

Output stream name * 3 getCategoryData Learn more

Description Import data from StagingSubCategoriesData

Reset

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Search factory and documentation

Preview experience Off

Data Factory StagingProductData StagingCategoriesDa... StagingGeographyD... StagingSalesRepData StagingSubCategorie... SalesData Sales_DF

Validate Data flow debug Debug Settings

Properties General Related

Name * Denis_DF

Description

Source settings Source options Projection Optimize Inspect Data preview

Output stream name * getCategoryData Learn more

Description Import data from StagingCategoriesData

Source type * Dataset Inline

Dataset * 3 StagingCategoriesData Test connection Open New

Options

Allow schema drift

Infer drifted column types

Validate schema

Skip line count

Sampling * Enable Disable

Activate Windows Go to Settings to activate Windows.

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Search factory and documentation

Preview experience Off

Data Factory StagingSalesRepData StagingSubCategorie... SalesData Sales_DF Sales_PL StagingSalesData Denis_DF

Validate Data flow debug Debug Settings

Properties General Related

Name * Denis_DF

Description

Source settings Source options Projection Optimize Inspect Data preview

Output stream name * 3 getSubCategoriesData Learn more

Description Import data from StagingSubCategoriesData

Source type * Dataset Inline

Dataset * 3 StagingSubCategoriesData Test connection Open New

Activate Windows Go to Settings to activate Windows.

- Projections → Import projection

Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Validate all Publish all 1

Preview experience Off

letshunvarthy@outlook.com DEFAULT DIRECTORY

Data Factory Validate Data flow debug Debug Settings

getCategoryData (Column 0 total) + getSubCategoriesData (Import data from StagingSubCategoriesData)

Multiple inputs/outputs

Join (highlighted with green arrow 1) Conditional Split Exists Union Lookup Schema modifier Derived Column Select Aggregate Surrogate Key Pivot

Properties General Related

Name * Denis_DF

Description

Activate Windows Go to Settings to activate Windows.

Microsoft Azure | Data Factory > projects-datafactory

Search factory and documentation

Validate all Publish all 1

Preview experience Off

letshunvarthy@outlook.com DEFAULT DIRECTORY

Data Factory Validate Data flow debug Debug Settings

Join settings Optimize Inspect Data preview

Output stream name * join1

Description Inner join on 'getCategoryData' and 'getSubCategoriesData'

Left stream * getCategoryData

Right stream * getSubCategoriesData

Join type * Inner (highlighted with green arrow 2)

Full outer Left outer Right outer Custom (cross)

Use fuzzy matching

Properties General Related

Name * Denis_DF

Description

Use fuzzy matching

Join conditions *

Left: getCategoryData's column Right: getSubCategoriesData's column

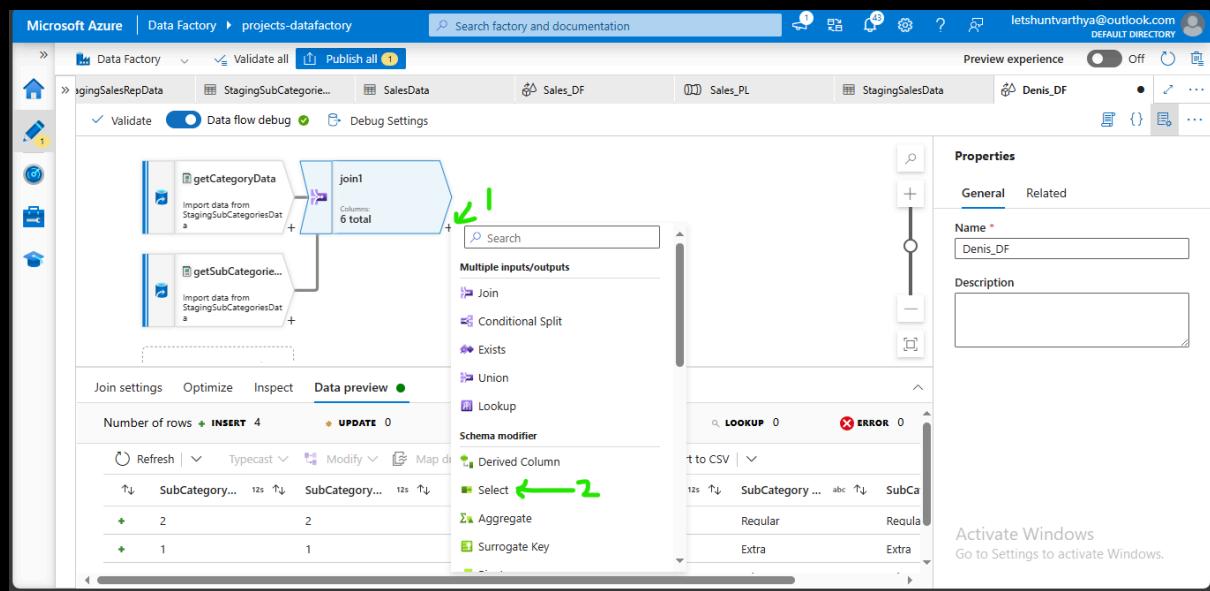
12s CategoryKey == 12s CategoryKey + -

Join settings Optimize Inspect Data preview (highlighted with green arrow 3)

Number of rows + INSERT 4 * UPDATE 0 × DELETE 0 * UPSERT 0 ↗ LOOKUP 0 ✘ ERROR 0 TOTAL 4

Refresh Typecast Modify Map drifted Statistics Remove Export to CSV

↑↓	CategoryKey	12s ↑↓	CategoryKey	12s ↑↓	Category	abc ↑↓	SubCategoryKey	12s ↑↓	SubCategory ...	abc ↑↓
+	1		1		Special		3		Micro	
+	1		1		Special		1		Extra	
+	2		2		General		4		Super	
+	2		2		General		2		Regular	



The screenshot shows the 'join1' component settings. Under 'Input columns', there are five columns from the source: 'getCategoryData@CategoryKey', 'Category', 'SubCategoryKey', 'getSubCategoriesData@CategoryKey', and 'SubCategory Name'. These are mapped to the output stream 'CategorySubCategoryData' under 'Name as'. Two green arrows point to the delete icons for the 'CategoryKey' and 'SubCategory Name' mappings.

The screenshot shows the 'CategorySubCategoryData' output stream settings. The 'Output stream name' is set to 'CategorySubCategoryData'. Under 'Input columns', there are three columns: 'SubCategoryKey', 'SubCategory Name', and 'Category', which are mapped to 'SubCategoryKey', 'SubCategory Name', and 'Category' respectively. A green arrow points to the 'Output stream name' field.

Select settings Optimize Inspect Data preview

Number of rows + INSERT 4 * UPDATE 0 X DELETE 0 + UPSERT 0 LOOKUP 0 X ERROR 0 TOTAL 4

Refresh | Typecast | Modify | Map drifted | Statistics | Remove | Export to CSV |

	SubCategoryKey	SubCategory Name	Category
+	2	Regular	General
+	1	Extra	Special
+	3	Micro	Special
+	4	Super	General

Microsoft Azure | Data Factory | projects-datafactory | Search factory and documentation | Preview experience Off

Validate | Data flow debug | Debug Settings

Properties

General Related

Name *

Description

Source settings

Output stream name *

Description

Source type *

Dataset *

Microsoft Azure | Data Factory | projects-datafactory | Search factory and documentation | Preview experience Off

Validate | Data flow debug | Debug Settings

Properties

General Related

Name *

Description

Source settings

Output stream name *

Description

Source type *

Dataset *

Options Allow schema drift

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Preview experience Off

Properties

Name: Denis_DF

Description:

Activate Windows Go to Settings to activate Windows.

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Preview experience Off

Join settings

Left stream: getSalesData

Right stream: getSalesRepData

Join type: Inner (highlighted)

Use fuzzy matching:

Join conditions:

- Left: getSalesData's column 123 SalesRepID
- Right: getSalesRepData's column 123 SalesRepID

Activate Windows Go to Settings to activate Windows.

Microsoft Azure | Data Factory > projects-datafactory

Validate all Publish all 1

Preview experience Off

Multiple inputs/outputs

Join

Conditional Split

Exists

Union

Lookup

Schema modifier

Derived Column

Select (highlighted)

Aggregate

Join settings Optimize Inspect Data preview

+ INSERT 100 * UPDATE 0 X DELETE 0 + UPSERT 0 Q LOOKUP 0

Typecast Modify Map drifted Statistics Remove Export to CSV

ProductID	Units	PercentOrder	Revenue	country	cities
7	10	60	0.99	Germany	Frankfurt

Activate Windows Go to Settings to activate Windows.

Select settings Optimize Inspect Data preview ●

Input columns * Auto mapping Reset Add mapping Delete 11 mappings: 1 column(s) from the inputs left unmapped

join2's column	Name as
abc Sales Rep Name	Sales Rep Name
123 fSalesPrimaryKey	fSalesPrimaryKey
abc country	country
abc cities	cities
Date	Date
12s ProductID	ProductID
12s Units	Units
1..2 PercentOfStandardCost	PercentOfStandardCost
1..2 RevenueDiscount	RevenueDiscount
12s getSalesData@SalesRepID	SalesRepID
12s getSalesRepData@SalesRepID	SalesRepID

Activate Windows
Go to Settings to activate Windows.

Microsoft Azure | Data Factory | projects-datafactory | Search factory and documentation | Preview experience Off

Data Factory | Validate all | Publish all

Validate | Data flow debug | Debug Settings

getSalesRepData

0 Columns

Source settings | Source options | Projection | Optimize | Inspect | Data preview ●

Output stream name * 2 → getProductData | Learn more

Description | Import data from StagingProductData

Source type * | Dataset | Inline

Dataset * 3 → StagingProductData | Test connection | Open | New

Options | Allow schema drift

Activate Windows
Go to Settings to activate Windows.

Microsoft Azure | Data Factory | projects-datafactory | Search factory and documentation | Preview experience Off

Data Factory | Validate all | Publish all

Factory Resources | Filter resources by name

StagingProductData | StagingSalesData | StagingSalesRepData | StagingSubCategoriesData | SubCategoriesData

Data flows | CategoriesData_DF | Denis_DF | Geography_DF | ProductData_DF | Sales_DF | SalesRep_DF | SubCategroiesData_DF | Power Query

getSalesRepData

6 Columns

Add Source | Multiple inputs/outputs | Join 2 → | Conditional Split | Exists | Union | Lookup

Source settings | Define default | Column name | Schema modifier

ProductID | Derived Column | ProductName | Select | Sub Category Key | Aggregate

Inspect | Data preview ●

projection | Reset schema | Format | Specify format | Activate Windows
Go to Settings to activate Windows.

The screenshot shows three views of a Data Flow job in Microsoft Azure Data Factory:

- Top View:** The "Join settings" tab is selected. A green arrow labeled "1" points to the "Right stream" dropdown, which is set to "CategorySubCategoryData". A green arrow labeled "2" points to the "Join type" section, where "Inner" is selected. A green arrow labeled "3" points to the "Join conditions" section, showing a condition where "Left: getProductData's column" equals "Right: CategorySubCategoryData's column" (Sub Category Key). A green arrow labeled "T" points to the "Join type" dropdown.
- Middle View:** The "Join settings" tab is selected. A green arrow labeled "1" points to the "Right stream" dropdown, which is set to "CategorySubCategoryData". A green arrow labeled "2" points to the "Join type" dropdown, which is set to "Inner". A green arrow labeled "3" points to the "Join conditions" section, showing a condition where "Left: getProductData's column" equals "Right: CategorySubCategoryData's column" (SubCategoryKey). A green arrow labeled "T" points to the "Join type" dropdown.
- Bottom View:** The "Input columns" tab is selected. A green arrow labeled "1" points to the "Name as" column for the "SubCategoryKey" mapping, which is set to "SubCategoryKey". A green arrow labeled "2" points to the "Name as" column for the "Sub Category Key" mapping, which is set to "Sub Category Key". A green arrow labeled "3" points to the "Delete" icon for the "SubCategoryKey" mapping. A green arrow labeled "4" points to the "Delete" icon for the "Sub Category Key" mapping. A green arrow labeled "5" points to the "Activate Windows" message at the bottom right.

Top Panel:

- Source settings:**
 - Output stream name:** getGeographyData
 - Description:** Import data from StagingGeographyData
 - Source type:** Dataset
 - Dataset:** StagingGeographyData
 - Options:** Allow schema drift

Middle Panel:

- Stream Flow:** getSalesData → join2 → 10 Columns
- join2:**
 - Output stream name:** SalesRepSales
 - Description:** Renaming join2 to SalesRepSales columns 'Sales Rep Name, fSalesPrimarykey, country, city'
 - Incoming stream:** join2
 - Options:** Skip duplicate input columns, Skip duplicate output columns

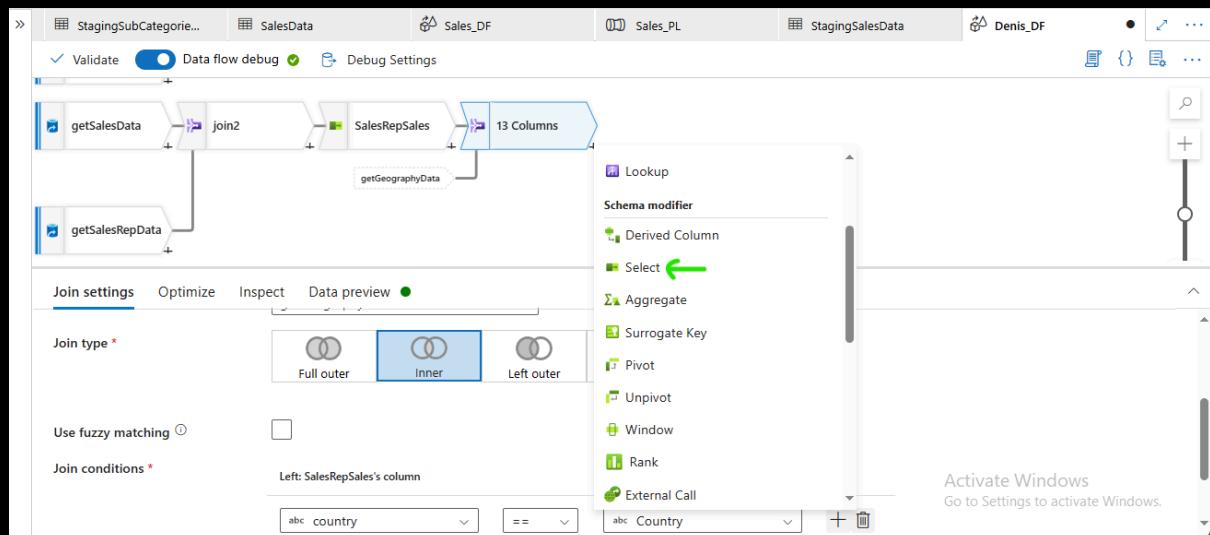
Right Panel (Toolbox):

- Multiple inputs/outputs: Join (highlighted), Conditional Split, Exists, Union, Lookup.
- Schema modifier: Derived Column, Select, Aggregate.

Bottom Panel:

- join2 (Join settings):**
 - Right stream:** getGeographyData
 - Join type:** Inner (highlighted)
 - Join conditions:**
 - Left: SalesRepSales's column: abc_country == abc_Country (highlighted with number 2)
 - Right: getGeographyData's column: abc_cities == abc_Town (highlighted with number 4)
 - Left: SalesRepSales's column: abc_cities == abc_Town (highlighted with number 5)
 - Right: getGeographyData's column: abc_Country == abc_Town (highlighted with number 6)

Activate Windows: Go to Settings to activate Windows.



Select settings **Optimize** **Inspect** **Data preview**

Output stream name * → SalesRepSalesGeo

Description: Renaming join4 to SalesRepSalesGeo with columns: 'Sales Rep Name', fSalesPrimaryKey, country, cities, Date,

Incoming stream *: join4

Options:

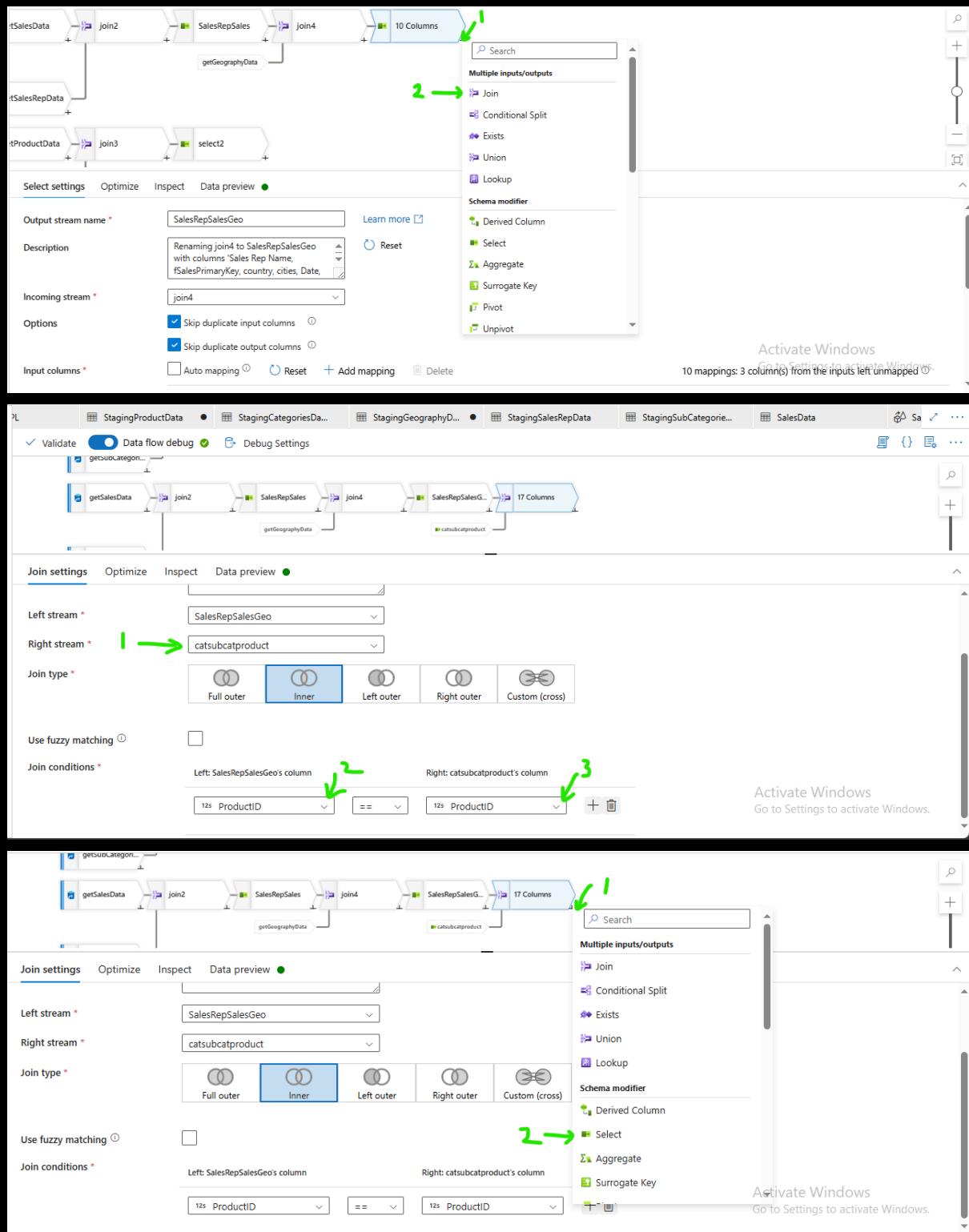
- Skip duplicate input columns
- Skip duplicate output columns

Input columns *

Input columns *	Auto mapping	Reset	Add mapping	Delete
join4's column				
abc Sales Rep Name				
123 fSalesPrimaryKey				
abc SalesRepSales@country				
abc cities				
123 Date				
123 ProductID				
123 Units				
12 PercentOfStandardCost				
12 RevenueDiscount				
123 SalesRepID				
123 GeographyID				
abc getGeographyData@Country				
abc Town				

13 mappings: All inputs mapped

Activate Windows
Go to Settings to activate Windows.



Select settings Optimize Inspect Data preview

Input columns * Auto mapping Reset Add mapping Delete

17 mappings: 1 column(s) from the inputs left unmapped

Join's column	Name as
121 fSalesPrimarykey	fSalesPrimarykey
121 SalesRepID	SalesRepID
abc Sales Rep Name	Sales Rep Name
abc ProductName	ProductName
abc SubCategory Name	SubCategory Name
abc Category	Category
abc Color	Color
abc country	country
abc cities	cities
Date	Date
121 Units	Units
121 PercentOfStandardCost	PercentOfStandardCost
121 RevenueDiscount	RevenueDiscount
121 RetailPrice	RetailPrice
121 StandardCost	StandardCost
121 catSubCatProduct@ProductID	ProductID
121 SalesRepSalesGeo@ProductID	ProductID

Activate Windows
Go to Settings to activate Windows.

Select settings Optimize Inspect Data preview

Number of rows: 100 INSERT 100 UPDATE 0 DELETE 0 UPsert 0 LOOKUP 0

Refresh | Typecast | Modify | Map drifted | Statistics | Remove | Export to CSV |

fSalesPri...	123	SalesRepID	123	Sales Rep...	abc	ProductN...	abc	SubCate...	abc	Category	abc	Color	abc	country	abc	cities	abc	Date	121	Units
36943	7	John White	123	Carlota	Extra	Special	abc	Florescen...	Germany	Frankfurt	2017-10-01	60								
36944	2	Bill Muray	123	VanHelen	Micro	Special	abc	Blue	Denmark	Copenha...	2017-04-26	71								
36945	6	Jan Novo...	123	Magnum	Extra	Special	abc	Green	Germany	Dresden	2017-09-28	68								
36946	5	Ellen Wo...	123	Linder	Regular	General	abc	Blue	Germany	Berlin	2017-04-16	157								
36947	2	Bill Muray	123	VanHelen	Micro	Special	abc	Blue	Denmark	Copenha...	2017-02-13	103								
36948	1	Ellie Gill	123	VanHelen	Micro	Special	abc	Blue	Czech rep...	Prague	2017-12-03	48								
36949	2	Bill Muray	123	Quad	Extra	Special	abc	Red	Denmark	Copenha...	2017-08-26	30								
36950	7	John White	123	Quad	Extra	Special	abc	Red	Germany	Frankfurt	2017-04-16	57								
36951	4	Ei Bob	123	Carlota	Super	General	abc	Florescen...	Germany	Frankfurt	2017-11-03	145								
36952	5	Ellen Wo...	123	Black Monk	Extra	Special	abc	Blue	Germany	Berlin	2017-01-01	31								
36953	6	Jan Novo...	123	Magnum	Extra	Special	abc	Green	Germany	Dresden	2017-03-04	87								
36954	3	Mark Spa...	123	Magnum	Extra	Special	abc	Green	Czech rep...	Ostrava	2017-04-23	50								
36955	1	Ellie Gill	123	Magnum	Regular	General	abc	Green	Czech rep...	Prague	2017-01-21	72								
36956	5	Ellen Wo...	123	Alder	Micro	Special	abc	Red	Germany	Berlin	2017-06-07	76								
36957	5	Ellen Wo...	123	Linder	Regular	General	abc	Blue	Germany	Berlin	2017-11-29	158								

Microsoft Azure | Data Factory | projects-datafactory | Search factory and documentation

Data Factory Validate all Publish all

PL StagingProductData StagingCategoriesDa... StagingGeographyD... StagingSalesRepData StagingSubCategorie... SalesData

Validate Data flow debug Debug Settings

getSalesData -> join2 -> SalesRepSales -> join4 -> SalesRepSalesGeo -> join5 -> 16 Columns

getSalesRepData -> join3 -> catSubCatproduct

getProductData -> join3 -> catSubCatproduct

CategorySubCate...

Select settings Optimize Inspect Data preview

Number of rows: 100 INSERT 100 UPDATE 0 DELETE 0 UPsert 0

Refresh | Typecast | Modify | Map drifted | Statistics | Remove | Export to CSV |

fSalesPri...	123	SalesRepID	123	Sales Rep...	abc	ProductN...	abc	SubCate...	abc	Category	abc	Color	abc	country	abc	cities	abc	Date	121	Units
36943	7	John White	123	Carlota	Extra	Special	abc	Florescen...	Germany	Frankfurt	2017-10-01	60								
36944	2	Bill Muray	123	VanHelen	Micro	Special	abc	Blue	Denmark	Copenha...	2017-04-26	71								
36945	6	Jan Novo...	123	Magnum	Extra	Special	abc	Green	Germany	Dresden	2017-09-28	68								

Activate Windows
Go to Settings to activate Windows.

Search

Multiple inputs/outputs

- Join
- Conditional Split
- Exists
- Union
- Lookup

Schema modifier

- Derived Column
- Select
- Aggregate
- Surrogate Key

Derived column's settings

Output stream name * derivedColumn1

Description Creating/updating the columns 'SalesPrimaryKey, SalesRepID, Sales Rep Name, ProductName, SubCategory'

Incoming stream * select1

Columns * ①

Column	Expression
Total_Revenue	RetailPrice*Units
Total_Cost	StandardCost*Units

Activate Windows
Go to Settings to activate Windows.

Derived column's settings

Optimize

Inspect

Data preview ●

UPSEND 0

LOOKUP 0

ERROR 0

TOTAL 1000

City	Date	Units	PercentOff	Revenue	RetailPrice	StandardCost	Total_Revenue	Total_Cost
Berlin	2017-06-07	76	0.977	0.4	23.95	7.55	1820.2	573.8
Prague	2017-09-27	134	0.989	0.5	23.95	7.55	3209.299...	1011.699...
Berlin	2017-10-09	75	0.992	0.4	23.95	7.55	1796.25	566.25
Frankfurt	2017-10-23	85	0.993	0.4	23.95	7.55	2035.75	641.75
Copenhagen	2017-09-05	35	0.988	0.25	23.95	7.55	838.25	264.25
Frankfurt	2017-09-12	172	0.987	0.55	23.95	7.55	4119.4	1298.6
Prague	2017-11-21	8	0.995	0.0	23.95	7.55	191.6	60.4
Frankfurt	2017-11-26	7	0.996	0.0	23.95	7.55	167.65	52.85

Derived column's settings

Optimize

Inspect

Data preview ●

UPSEND 0

LOOKUP 0

ERROR 0

City	Date	Units	PercentOff	Revenue	RetailPrice	StandardCost	Total_Revenue	Total_Cost
Berlin	2017-06-07	76	0.977	0.4	23.95	7.55	182	182
Prague	2017-09-27	134	0.989	0.5	23.95	7.55	320	320
Berlin	2017-10-09	75	0.992	0.4	23.95	7.55	179	179
Frankfurt	2017-10-23	85	0.993	0.4	23.95	7.55	203	203
Copenhagen	2017-09-05	35	0.988	0.25	23.95	7.55	838	838

Derived column's settings

Optimize

Inspect

Data preview ●

Output stream name * derivedColumn2

Description Creating/updating the columns 'SalesPrimaryKey, SalesRepID, Sales Rep Name, ProductName, SubCategory'

Incoming stream * derivedColumn1

Columns * ①

Column	Expression
Gross_Profit	Total_Revenue - Total_Cost

Date	Units	PercentOff	Revenue	RetailPrice	Standard	Total_Rev	Total_Cost	Gross_Profit
2017-06-07	76	0.977	0.4	23.95	7.55	1820.2	573.8	1246.4
2017-09-27	134	0.989	0.5	23.95	7.55	3209.299...	1011.699...	2197.6
2017-10-09	75	0.992	0.4	23.95	7.55	1796.25	566.25	1230.0
2017-10-23	85	0.993	0.4	23.95	7.55	2035.75	641.75	1394.0
2017-09-05	35	0.988	0.25	23.95	7.55	838.25	264.25	574.0
2017-09-12	172	0.987	0.55	23.95	7.55	4119.4	1298.6	2820.799...

Azure SQL Database creation:

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

letshunvarthy@outlook.com DEFAULT DIRECTORY (LETSHUN...)

Azure services

- Create a resource
- Resource groups
- Cost Management
- Subscriptions
- Quickstart Center
- Azure AI services
- Kubernetes services
- Virtual machines
- App Services
- More services

Resources

Recent

Name	Type	Last Viewed
denissblobs	Storage account	an hour ago
denissadilssa	Storage account	2 hours ago
projects-datafactory	Data factory (V2)	17 hours ago
projects_rg	Resource group	17 hours ago
deniscomossa	Azure Cosmos DB for MongoDB account (RU)	17 hours ago
Free Trial	Subscription	19 hours ago

See all <https://portal.azure.com/#create/hub>

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

letshunvarthy@outlook.com DEFAULT DIRECTORY (LETSHUN...)

Home >

Create a resource ...

Get Started

Recently created

Popular Azure services See more in All services

Categories

- AI + Machine Learning
- Analytics
- Blockchain
- Compute
- Containers
- Databases
- Developer Tools
- DevOps
- Identity
- Integration
- Internet of Things

Getting Started? Try our Quickstart center

Popular Marketplace products See more in Marketplace

Virtual machine	Windows Server 2019 Datacenter
Web App	Windows 11 Pro, version 21H2
SQL Database	Ubuntu Server 20.04 LTS
Function App	Ubuntu Server 22.04 LTS
Key Vault	Red Hat Enterprise Linux 7.4
Data Factory	Essentials 50K

Activate Windows Go to Settings to activate Windows.

Create SQL Database Microsoft

Basics Networking Security Additional settings Tags Review + create

Create a SQL database with your preferred configurations. Complete the Basics tab then go to Review + Create to provision with smart defaults, or visit each tab to customize. [Learn more](#)

Want to try Azure SQL Database for free? Create a free serverless database with the first 100,000 vCore seconds, 32GB of data, and 32GB of backup storage free per month for the lifetime of the subscription. [Learn more](#)

Apply offer (Preview)

SQL Database Hyperscale: Low price, high scalability, and best feature set. [Learn more](#)

Project details

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

Subscription * | [Review + create](#) [Next : Networking >](#)

Activate Windows Go to Settings to activate Windows.

Project details

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

Subscription * | [Review + create](#) [Next : Networking >](#)

Resource group * Create new

Database details

Enter required settings for this database, including picking a logical server and configuring the compute and storage resources

1 → Database name *

2 → Server * Create new

3 → The value must not be empty.

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

Activate Windows Go to Settings to activate Windows.

Server details

Select your preferred authentication methods for accessing this server. Create a server admin login and password to access your server with SQL authentication, select only Microsoft Entra authentication [Learn more](#) using an existing Microsoft Entra user, group, or application as Microsoft Entra admin [Learn more](#), or select both SQL and Microsoft Entra authentication.

Authentication method

1 → Use Microsoft Entra-only authentication

2 → Use both SQL and Microsoft Entra authentication

3 → Use SQL authentication

4 → Server admin login *

5 → Password *

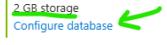
6 → Confirm password *

[OK](#)

Activate Windows Go to Settings to activate Windows.

Workload environment  Development  Production

1 Default settings provided for Development workloads. Configurations can be modified as needed.

Compute + storage *  2 GB storage  **2**

Backup storage redundancy

Choose how your PITR and LTR backups are replicated. Geo restore or ability to recover from regional outage is only available when geo-redundant storage is selected.

Backup storage redundancy  Locally-redundant backup storage  Zone-redundant backup storage  Geo-redundant backup storage

Review + create **Next : Networking >** **←**

Activate Windows
Go to Settings to activate Windows.

Basics **Networking** Security Additional settings Tags Review + create

Configure network access and connectivity for your server. The configuration selected below will apply to the selected server 'denisserver' and all databases it manages. [Learn more](#)

Network connectivity

Choose an option for configuring connectivity to your server via public endpoint or private endpoint. Choosing no access creates with defaults and you can configure connection method after server creation. [Learn more](#)

Connectivity method *  No access  Public endpoint  Private endpoint **1**

SQL

Cost summary

Basic (Basic)	81.49
Cost per DTU (in INR)	x 5
DTUs selected	
ESTIMATED COST / MONTH	407.46 INR

Firewall rules

Setting 'Allow Azure services and resources to access this server' to Yes allows communications from all resources inside the Azure boundary, that may or may not be part of your subscription. [Learn more](#)

Setting 'Add current client IP address' to Yes will add an entry for your client IP address to the server firewall.

Allow Azure services and resources to access this server *  No  Yes **2**

Add current client IP address *  No  Yes **3**

Activate Windows

Connection policy

Configure how clients communicate with your SQL database server. [Learn more](#)

Connection policy  Default - Uses Redirect policy for all client connections originating inside of Azure (except Private Endpoint connections) and Proxy for all client connections originating outside Azure  Proxy - All connections are proxied via the Azure SQL Database gateways  Redirect - Clients establish connections directly to the node hosting the database

Encrypted connections

This server supports encrypted connections using Transport Layer Security (TLS). For information on TLS version and certificates, refer to connecting with TLS/SSL. [Learn more](#)

Minimum TLS version  TLS 1.2 **4**

Review + create **< Previous** **Next : Security >** **←**

Activate Windows
Go to Settings to activate Windows.

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

letshuntvarthy@outlook.com DEFAULT DIRECTORY (LETSHUNT...)

Home > Create a resource >

Create SQL Database

Microsoft

Basics Networking Security Additional settings Tags Review + create

Microsoft Defender for SQL

Protect your data using Microsoft Defender for SQL, a unified security package including vulnerability assessment and advanced threat protection for your server. [Learn more](#)

Get started with a 30 day free trial period, and then 1247.9202 INR/server/month.

Enable Microsoft Defender for SQL * Start free trial Not now

Ledger

Ledger cryptographically verifies the integrity of your data and detects any tampering that might have occurred. [Learn more](#)

Ledger **Not configured** [Configure ledger](#)

Server identity

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

Activate Windows Go to Settings to activate Windows.

Basics Networking Security **Additional settings** Tags Review + create

Customize additional configuration parameters including collation & sample data.

Data source

Start with a blank database, restore from a backup or select sample data to populate your new database.

Use existing data * None Backup Sample

Database collation

Database collation defines the rules that sort and compare data, and cannot be changed after database creation. The default database collation is SQL_Latin1_General_CI_AS. [Learn more](#)

Collation * SQL_Latin1_General_CI_AS [Find a collation](#)

Maintenance window

Select a preferred maintenance window from the drop-down. During maintenance, databases remain available, but some updates may require a failover. The system default maintenance window (5pm to 8am) limits most activities to this time, but urgent updates may occur outside of it. To ensure all updates occur only during the maintenance window, select a

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

Activate Windows Go to Settings to activate Windows.

Basics Networking Security Additional settings **Tags** Review + create

Tags are name/value pairs that enable you to categorize and view consolidated billing by applying the same tag to multiple resources and resource groups. [Learn more](#)

Note that if you create tags and then change resource settings on other tabs, your tags will be automatically updated.

Name	Value	Resource
<input type="text"/>	<input type="text"/> :	<input type="text"/> 2 selected 

Cost summary

Basic (Basic)	
Cost per DTU (in INR)	81.49
DTUs selected	x 5
ESTIMATED COST / MONTH	407.46 INR

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

Activate Windows Go to Settings to activate Windows.

Screenshots illustrating the creation of an Azure SQL Database and its cost summary.

Azure Marketplace Creation:

- Product details:** SQL database by Microsoft. Estimated cost per month: 407.46 INR.
- Terms:** By clicking "Create", I (a) agree to the legal terms and privacy statement(s) associated with the Marketplace offering(s) listed above; (b) authorize Microsoft to bill my current payment method for the fees associated with the offering(s), with the same billing frequency as my Azure subscription; and (c) agree that Microsoft may share my contact, usage and transactional information with the provider(s) of the offering(s) for support, billing and other transactional activities. Microsoft does not provide rights for third-party offerings. For additional details see [Azure Marketplace Terms](#).
- Basics:** Subscription: Free Trial, Resource group: projects_rg.

Cost Summary:

Basic (Basic)	
Cost per DTU (in INR)	81.49
DTUs selected	x 5
ESTIMATED COST / MONTH	
407.46 INR	

Create button highlighted with a green arrow.

Azure SQL Database Overview:

Deployment Status: Your deployment is complete.

Deployment Details: Deployment name: Microsoft.SQLDatabase.newDatabase... Start time: 5/25/2024, 10:24:56 AM. Subscription: Free Trial. Resource group: projects_rg.

Next steps: Go to resource (button highlighted with a green arrow).

Flowchart: Data flow diagram showing stages: SalesRepSalesG... → join5 → select1 → derivedColumn1 → 18 Columns. A green arrow points to the "formatters" section of the context menu.

Data Preview:

Date	Units	PercentO...	Revenue...	RetailPrice
2017-06-07	76	0.977	0.4	23.95
2017-09-27	134	0.989	0.5	23.95
2017-10-09	75	0.992	0.4	23.95
2017-10-23	85	0.993	0.4	23.95
2017-09-05	35	0.988	0.25	23.95
2017-09-12	172	0.987	0.55	23.95

Context Menu:

- formatters
- Flatten
- Parse
- Stringify
- Row modifier
- Filter
- Sort
- Alter Row
- Assert
- Flowlets
- Destination
- Sink** (highlighted with a green arrow, labeled with a red number 2)

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.

The screenshot shows the Microsoft Azure Data Factory interface. On the left, a data flow pipeline is visible with various stages: SalesRepSalesG..., joinS, select1, derivedColumn1, derivedColumn2, and 18 Columns. A green arrow points from the 'Dataset' dropdown in the Sink configuration to the 'New' button in the 'Dataset' section of the configuration pane.

Sink configuration details:

- Output stream name: sink1
- Description: Add sink dataset
- Incoming stream: derivedColumn2
- Sink type: Dataset (selected)
- Dataset: Select... (highlighted with a red box)
- Options: Allow schema drift (checked)

A modal dialog titled "New dataset" is open on the right, showing a grid of data store options under the "Azure" tab. A green arrow points from the "Search" bar to the "Azure" tab. Another green arrow points from the "Continue" button at the bottom of the dialog to the "Continue" button in the main configuration pane.

New dataset dialog:

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

Search	All	Azure	Database	File	Generic protocol	NoSQL	Services and apps
			Azure SQL Database	Azure Synapse Analytics			
			Microsoft Fabric	Azure AI Search	Azure Cosmos DB for	Activate Windows	Cancel

New linked service

Azure SQL Database [Learn more](#)

Name *

Description

Connect via integration runtime * ⓘ

Version

Recommended Legacy

[Import from connection string](#)

Account selection method ⓘ

From Azure subscription Enter manually

Azure subscription

Server name *

Database name *

Authentication type *

User name *

>Password Azure Key Vault

Activate Windows
Go to Settings to activate Windows

Connection successful 

[Create](#) [Cancel](#) [Test connection](#)

Microsoft Azure | Data Factory > projects-datafactory

PL StagingProductData StagingCategoriesData StagingGeographyD... StagingSalesRepData

Validate all Publish all 3

Validate Data flow debug

Sink Settings Errors Mapping Optimize Inspect Data preview

Output stream name * sink1 Learn more ↗

Description Add sink dataset ↗ Reset

Incoming stream * derivedColumn2 ↗

Sink type * Dataset Inline Cache

Dataset * Select... + New

Options Allow schema drift

Set properties

Name DenisFinalData

Linked service * adf2asql_ls ↗

Select from existing table New table 1

Schema and table name Schema name . Denis_Tb ↗

Advanced

OK Back

Activate Windows Go to Settings to activate Windows Cancel

Microsoft Azure | Data Factory > projects-datafactory

PL StagingProductData StagingCategoriesData StagingGeographyD... StagingSalesRepData SalesData

Validate all Publish all 4 ↗

Validate Data flow debug

Pending changes (4)

NAME	CHANGE	EXISTING
StagingProductData	(Edited)	StagingProductData
StagingGeographyData	(Edited)	StagingGeographyData
DenisFinalData	(New)	-
Denis_DF	(Edited)	Denis_DF

Activate Windows Go to Settings to activate Windows.

Microsoft Azure | Data Factory > projects-datafactory

PL StagingProductData StagingCategoriesData StagingGeographyD... StagingSalesRepData

Validate all Publishing 4 ↗

Validate Data flow debug

Publish all

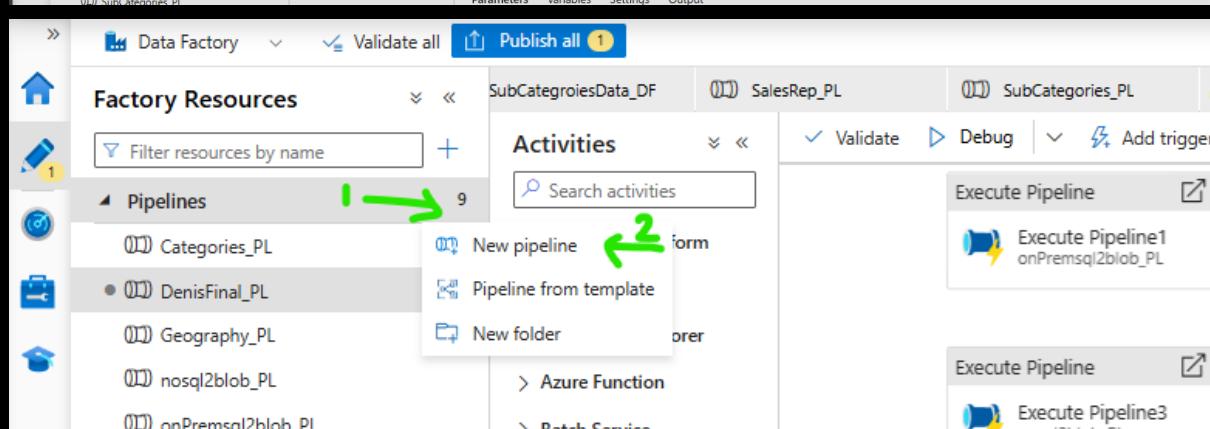
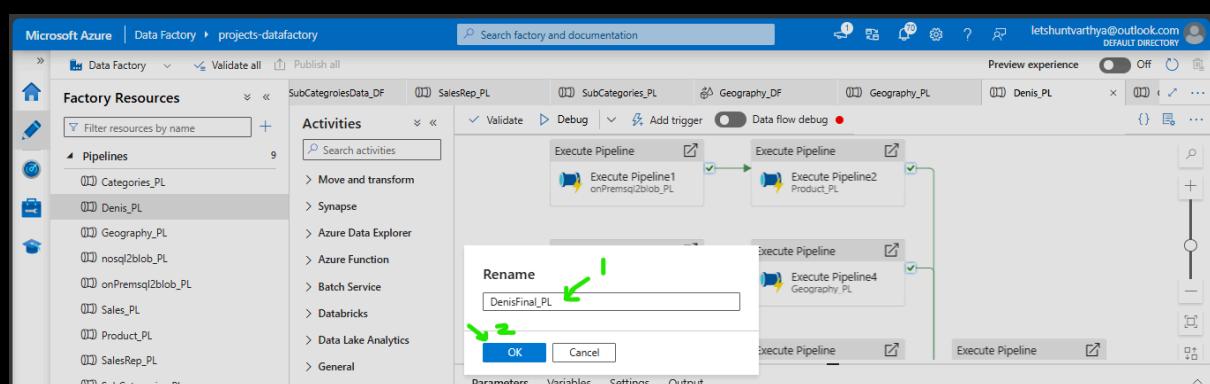
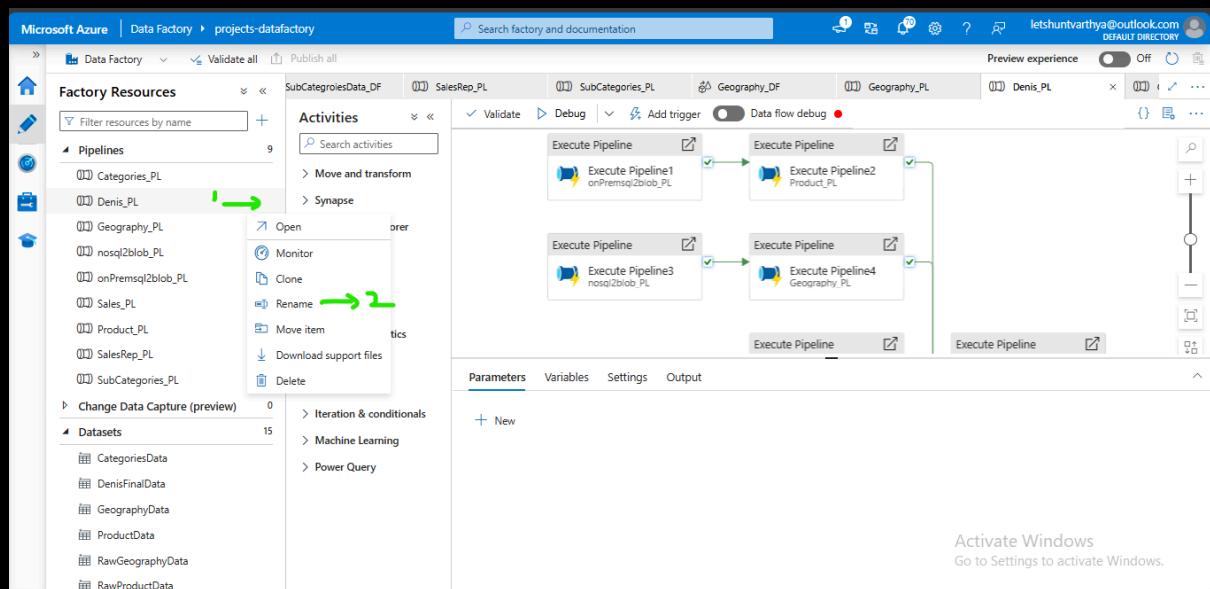
You are about to publish all pending changes to the live environment. Learn more ↗

Pending changes (4)

NAME	CHANGE	EXISTING
StagingProductData	(Edited)	StagingProductData
StagingGeographyData	(Edited)	StagingGeographyData
DenisFinalData	(New)	-
Denis_DF	(Edited)	Denis_DF

Activate Windows Go to Settings to activate Windows.

OK Cancel



This screenshot shows the Microsoft Azure Data Factory interface. On the left, the 'Activities' sidebar is open, showing various options like 'Copy data', 'Data flow', and 'Move and transform'. A green arrow labeled '3' points from the 'Data flow' option to the main canvas area. In the center, a 'Data flow' activity named 'denis_PL' is selected. The 'Properties' panel on the right shows the name 'denis_PL' highlighted with a green arrow labeled '1'. The 'General' tab is selected, showing the run settings: 'Run on (Azure IR)' set to 'AutoResolveIntegrationRuntime', 'Compute size' set to 'Small', and 'Logging level' set to 'Verbose'. A green arrow labeled '4' points to the 'Run on' dropdown.

This screenshot shows the Microsoft Azure Data Factory interface. On the left, the 'Pipelines' section of the 'Factory Resources' sidebar is selected, showing several pipelines listed. A green arrow labeled '1' points to the 'DenisFinal_PL' pipeline. The main canvas shows a pipeline structure with multiple 'Execute Pipeline' activities. The bottom pane shows the 'General' tab of the 'Settings' panel for the 'DenisFinal_PL' pipeline, with the 'Invoked pipeline' dropdown set to 'denis_PL' and the 'Wait on completion' checkbox checked. A green arrow labeled '2' points to the 'General' tab, and another green arrow labeled '3' points to the 'Invoked pipeline' dropdown.

This screenshot shows the Microsoft Azure Data Factory interface. The top navigation bar has a 'Publish all' button highlighted with a green arrow labeled '4'. The main canvas shows a pipeline structure with multiple 'Execute Pipeline' activities. The bottom pane shows the 'General' tab of the 'Settings' panel for the 'DenisFinal_PL' pipeline, with the 'Invoked pipeline' dropdown set to 'denis_PL' and the 'Wait on completion' checkbox checked.

This screenshot shows the Microsoft Azure Data Factory interface. The top navigation bar has a 'Publish all' button highlighted with a green arrow labeled '4'. The main canvas shows a pipeline structure with multiple 'Execute Pipeline' activities. The bottom pane shows the 'General' tab of the 'Settings' panel for the 'DenisFinal_PL' pipeline, with the 'Trigger now' button highlighted with a green arrow labeled '2'. This indicates that the pipeline has been triggered for immediate execution.

The screenshot shows the Microsoft Azure Data Factory interface. On the left, the 'Factory Resources' sidebar lists various pipelines and datasets. In the center, the 'Activities' pane displays a sequence of pipeline steps: 'SubCategorieData_DF', 'SalesRep_PL', and 'SubCategories_PL'. A tooltip above the first step says 'Trigger pipeline now using last published configuration.' Below the activities, there's a 'Parameters' section which is currently empty. At the bottom right of the central area, there's an 'OK' button with a green arrow pointing to it, indicating the next step to trigger the pipeline.

The screenshot shows the Microsoft Azure SQL Database Query editor. The left sidebar shows the database schema with tables like 'dbo.denistb'. The main area contains two tabs: 'Query 1' and 'Query 2'. 'Query 1' contains the following SQL code:

```
1 SELECT TOP (1000) * FROM [dbo].[denistb]
```

The 'Results' tab shows the output of the query:

FSalesPrimarykey	SalesRepID	Sales Rep Name	ProductName	SubCategory Name	Category	Color	country
36943	7	John White	Carlota	Extra	Special	Florescent Pink	Germany
36944	2	Bill Muray	VanHelen	Micro	Special	Blue	Denmark

A message at the bottom of the results table says 'Query succeeded | 2s'.

Here, we can see that data has been loaded in azure sql database.