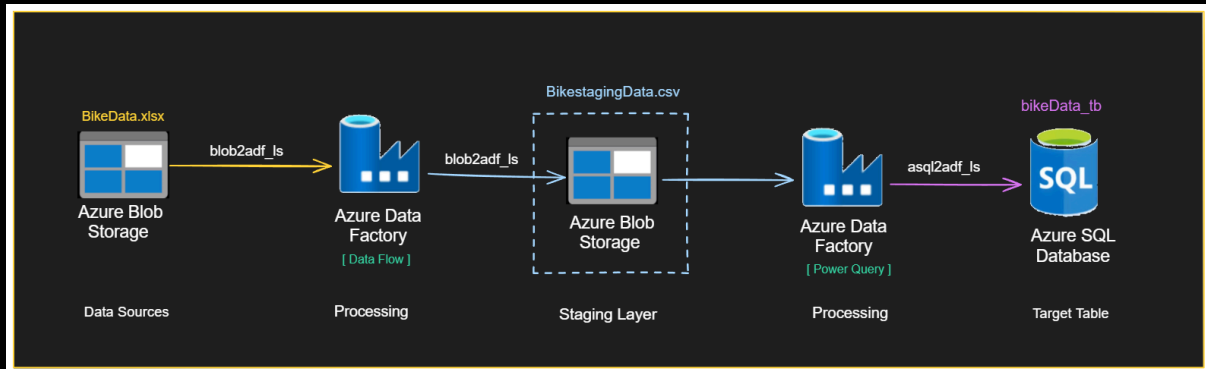


Blob Storage → Data Factory → staging (blob) → powerquery(adf) → SQL Database ( pipeline ):

Pipeline Architecture :

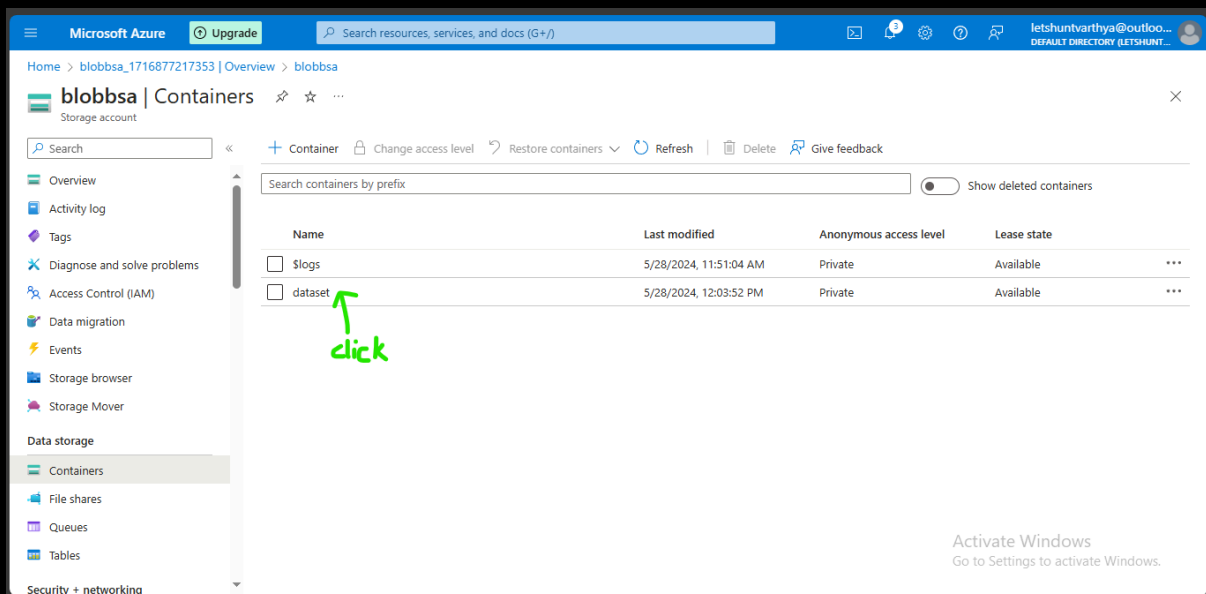
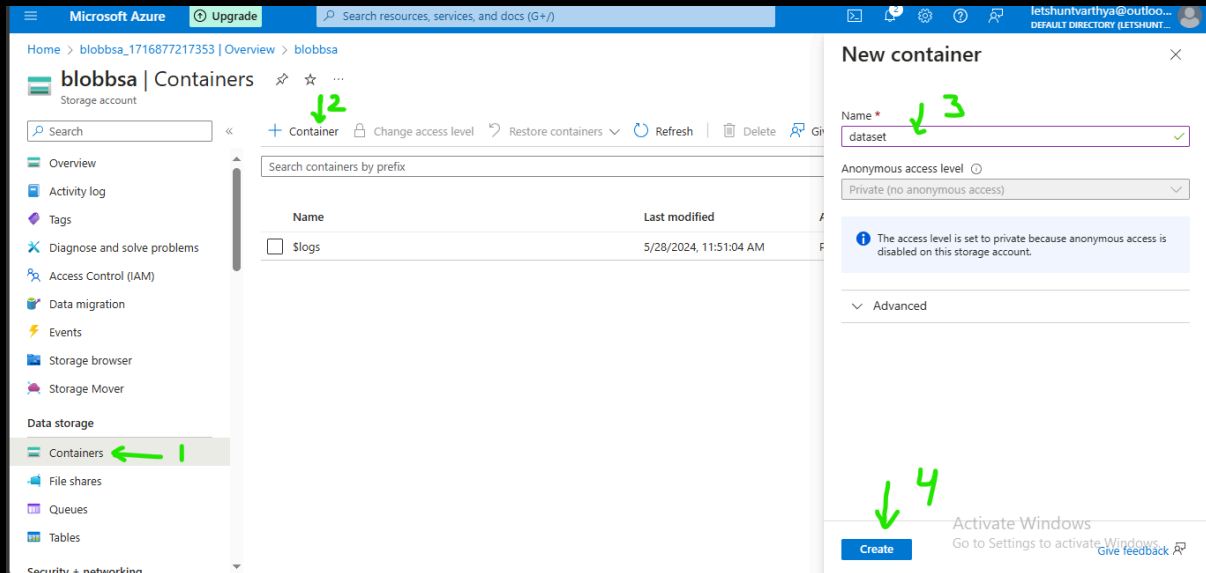


Azure Services Required :

1. Azure blob storage
2. Azure DataFactory
3. Azure SQL Database

## Azure Blob Storage Creation :

- Creation of azure blob storage can be shown in the document below .
- Parameters used while creating azure blob storage.
- Storage account name : blobbsa
- [https://docs.google.com/document/d/1vZCMfM9ieALTQm6Jdw12JzxcMkwnyvkSJHCrry2kVlo/edit?usp=drive\\_link](https://docs.google.com/document/d/1vZCMfM9ieALTQm6Jdw12JzxcMkwnyvkSJHCrry2kVlo/edit?usp=drive_link)



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

Home > blobbsa\_1716877217353 | Overview > blobbsa | Containers >

**dataset** Container

Search

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: dataset

Search blobs by prefix (case-sensitive)

Add filter

Name	Modified	Access tier
No results		

Upload blob

1 file(s) selected: 2666141-Bike\_Data (1).csv  
Drag and drop files here or Browse for files

Overwrite if files already exist

Advanced

Upload

Give feedback

Activate Windows  
Go to Settings to activate Windows.

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

Home > blobbsa\_1716877217353 | Overview > blobbsa | Containers >

**dataset** Container

Search

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: dataset

Search blobs by prefix (case-sensitive) Show deleted blobs

Add filter

Name	Modified	Access tier	Archive status	Blob type	Size	Lease state
2666141-Bike_Data...	5/28/2024, 12:10:55 ...	Cool (Inferred)		Block blob	6.58 MiB	Available

Activate Windows  
Go to Settings to activate Windows.

- Create a container called stagingdata.

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

Home > blobbsa

**blobbsa** | Containers Storage account

Search

Overview  
Activity log  
Tags  
Diagnose and solve problems  
Access Control (IAM)  
Data migration  
Events  
Storage browser  
Storage Mover

Data storage  
Containers  
File shares  
Queues  
Tables

Security + networking

Container

Search containers by prefix

Name	Last modified
\$logs	5/28/2024, 11:51:04 AM
dataset	5/28/2024, 12:03:52 PM

New container

Name \* stagingdata

Anonymous access level Private (no anonymous access)

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

Advanced

Create

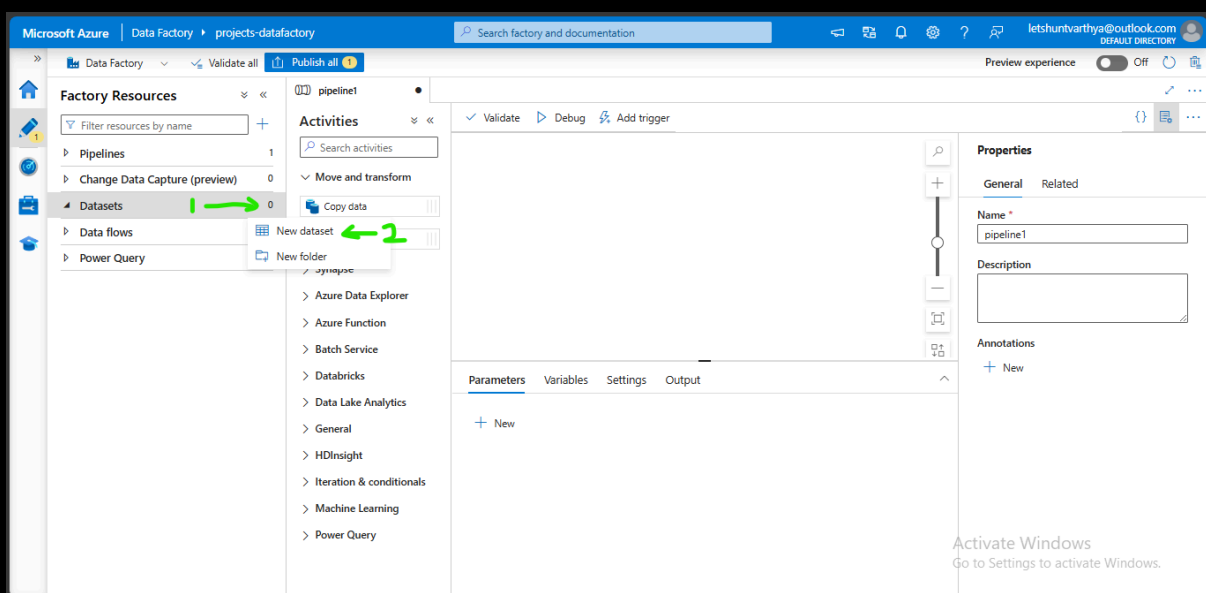
Activate Windows  
Go to Settings to activate Windows.

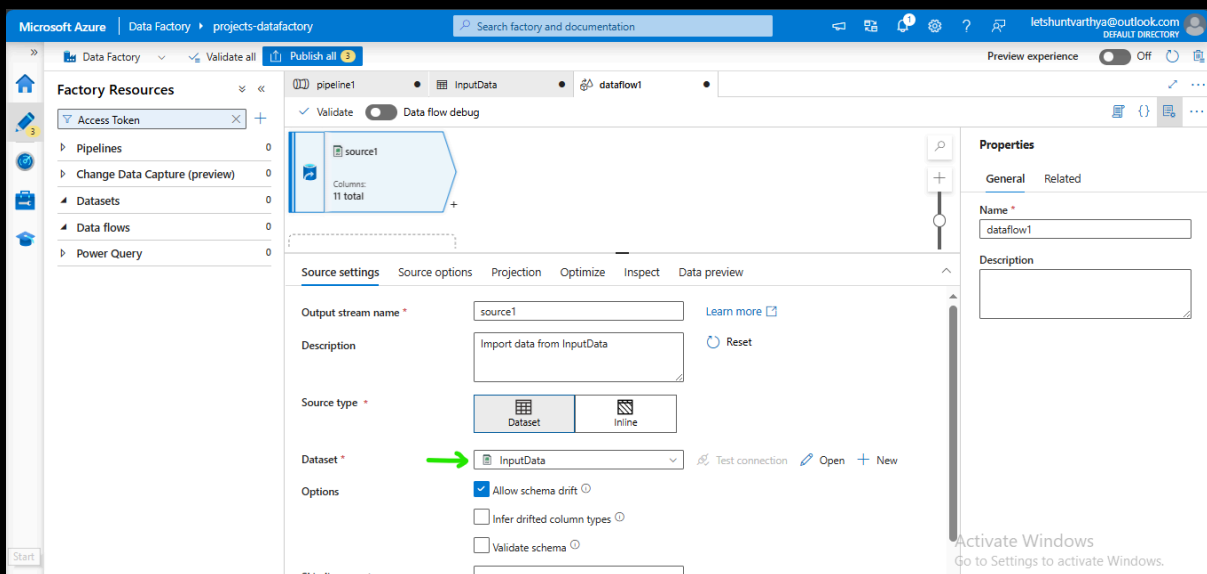
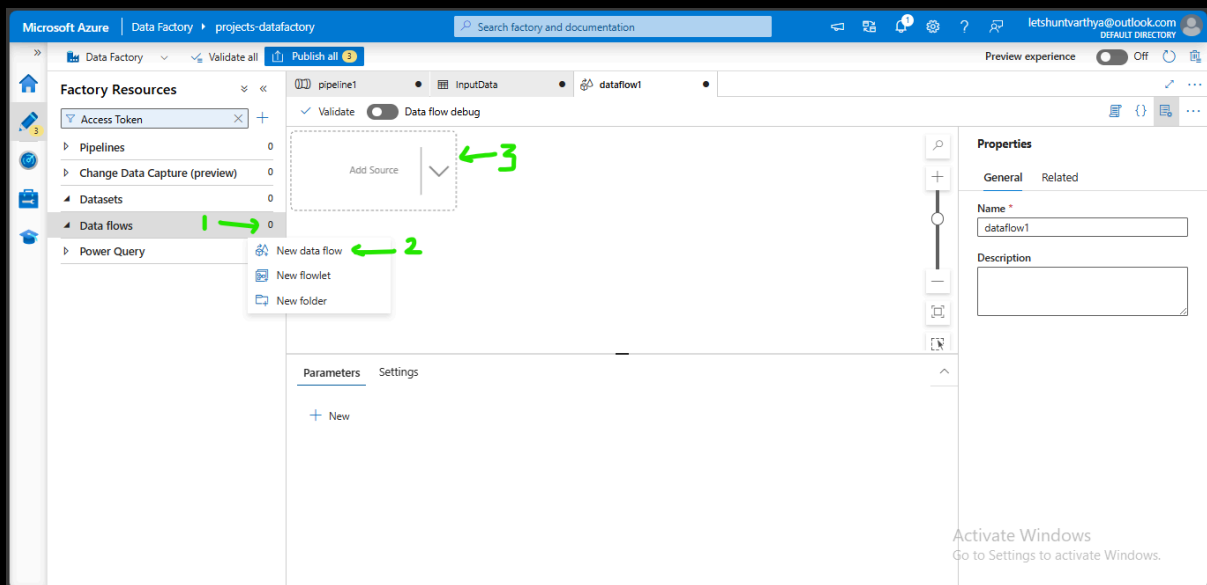
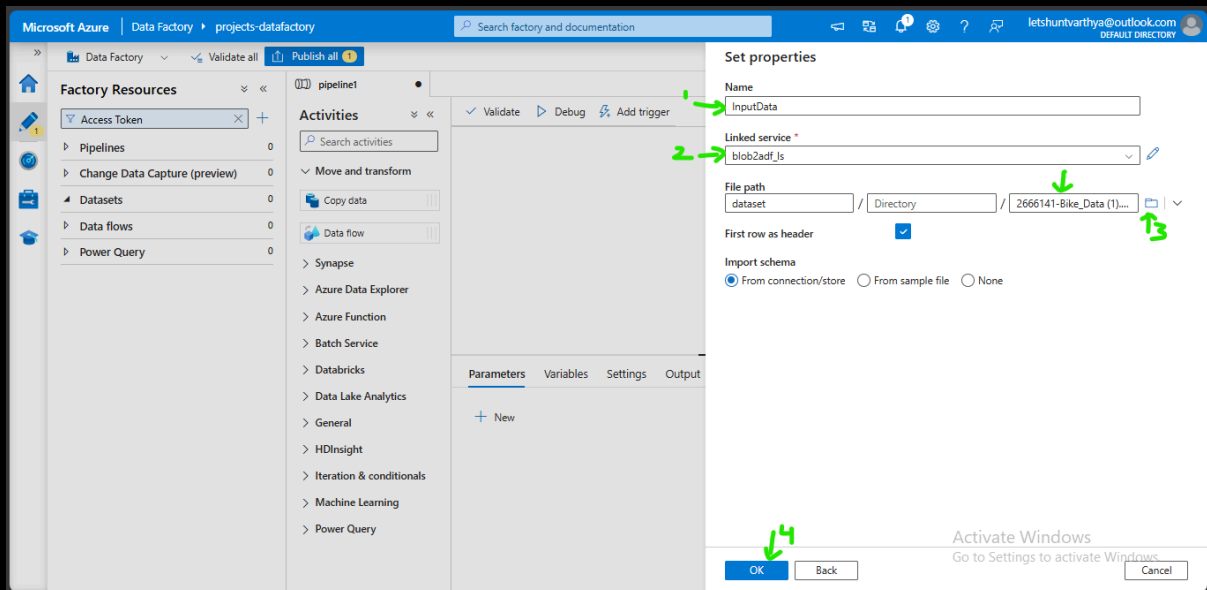
## Azure SQL Database Creation :

- Creation of an azure sql database is shown in the document below.
- Parameters used while creating azure sql database.
- Database name : **sqldb**
- Server name : **projectsserver**
- Server admin login : **project\_admin**
- [https://docs.google.com/document/d/16iB1EsGKHc6-bcgTPSfqkK6BVf3n8\\_fpbat42uNOvXc/edit?usp=drive\\_link](https://docs.google.com/document/d/16iB1EsGKHc6-bcgTPSfqkK6BVf3n8_fpbat42uNOvXc/edit?usp=drive_link)

## Azure Data Factory Creation :

- Creation of the data factory is shown in the document below.
- Parameters used while creating azure data factory.
- Datafactory name : projects-datafactory
- [https://docs.google.com/document/d/1lpvA7XumJjbIP0wPWWUlf\\_d\\_gdvh6jTn\\_0a12HWv0jcM/edit?usp=sharing](https://docs.google.com/document/d/1lpvA7XumJjbIP0wPWWUlf_d_gdvh6jTn_0a12HWv0jcM/edit?usp=sharing)
- Click on Go to resource.
- Click on launch workspace.
- Now let's do the transformations on data





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

Number of rows **INSERT** 100 **UPDATE** 0 **DELETE** 0 **UPSERT** 0

Refresh Typecast Modify Map drifted Statistics Remove Export to CSV

	Region	Country	Custo...	Busine...	Categ...	Model
+	North ...	United...	Advan...	Comp...	Road F...	LL Roa...
+	North ...	United...	Central...	Bikes	Mount...	Mount...
+	North ...	United...	Leadin...	Clothing	Jerseys	Long-S...
+	North ...	United...	Paint S...	Comp...	Mount...	HL Mo...
+	North ...	United...	Scoote...	Bikes	Road B...	Road-...
+	North ...	United...	Scoote...	Bikes	Road B...	Road-...
+	North ...	United...	Scoote...	Bikes	Road B...	Road-...
+	North ...	United...	Moder...	Bikes	Road B...	Road-...
+	North ...	United...	Moder...	Bikes	Road B...	Road-...

source1 Columns: 11 total

derivedColumn1 Creating/updating the columns Region, Country, Customer, Business Segment, Category, Model, Color, SalesDate, ListPrice, UnitPrice, OrderQty

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

Define default format Detect data type Import projection Reset schema

Column name	Type	Format
Region	abc string	Specify format
Country	abc string	Specify format
Customer	abc string	Specify format
Business Segment	abc string	Specify format
Category	abc string	Specify format
Model	abc string	Specify format
Color	abc string	Specify format
SalesDate	abc string	Specify format
ListPrice	12 double	Specify format
UnitPrice	12 double	Specify format
OrderQty	123 integer	Specify format

source1 Import data from InputData + derivedColumn1 Columns: 13 total

Derived column's settings Optimize Inspect Data preview ●

Output stream name \* derivedColumn1 [Learn more](#)

Description Creating/updating the columns 'Region, Country, Customer, Business Segment, Category, Model, Color, SalesDate, [Reset](#)

Incoming stream \* source1

+ Add Clone Delete Open expression builder

Columns ①

Column	Expression
Cost	UnitPrice*OrderQty 1.2 + -
sales	ListPrice*OrderQty 1.2 + -

Handwritten green annotations: 1 points to 'Column', 2 points to 'Expression', 3 points to the '+' icon, 4 points to 'Cost', 5 points to 'ListPrice\*OrderQty'.

source1 Import data from InputData + derivedColumn1 Columns: 13 total

Derived column's settings Optimize Inspect Data preview ● ← 1

Number of rows + INSERT 100 UPDATE 0 DELETE 0 UPSERT 0

Refresh Typecast Modify Map drifted Statistics Remove Export to CSV

	Region	Count...	Custo...	Busin...	Categ...	Model	Color	Sales...	ListPri...
+	North ...	Unite...	Advan...	Comp...	Road ...	LL Roa...	Red	4/1/2...	337.22
+	North ...	Unite...	Centr...	Bikes	Moun...	Moun...	Silver	4/1/2...	NULL
+	North ...	Unite...	Leadi...	Clothi...	Jerseys	Long...	Multi	4/1/2...	49.99
+	North ...	Unite...	Paint ...	Comp...	Moun...	HL M...	Black	4/1/2...	NULL
+	North ...	Unite...	Scoot...	Bikes	Road ...	Road-...	Red	4/1/2...	NULL
+	North ...	Unite...	Scoot...	Bikes	Road ...	Road-...	Black	4/1/2...	782.99
+	North ...	Unite...	Scoot...	Bikes	Road ...	Road-...	Red	4/1/2...	782.99
+	North ...	Unite...	Mode...	Bikes	Road ...	Road-...	Red	4/1/2...	NULL
+	North ...	Unite...	Mode...	Bikes	Road ...	Road-...	Black	4/1/2...	782.99

Handwritten green annotations: 1 points to 'Data preview', 2 points to the 'Refresh' button.

- Lets load these data into **stagingdata (blobbsa)**.

Microsoft Azure | Data Factory | projects-datafactory

Search factory and documentation

lets huntvartha@outlook.com  
DEFAULT DIRECTORY

Validate all Publish all

pipeline1 InputData dataflow1

Validate Data flow debug Debug Settings

source1 Import data from InputData

derivedColumn1 Creating/updating the columns Region, Country, Customer, Business Segment, Category, Model, Color, SalesDate...

sink1 Columns: 13 total

Sink Settings Errors Mapping Optimize Inspect Data preview

Output stream name \* sink1 Learn more

Description Add sink dataset Reset

Incoming stream \* derivedColumn1

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

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

- Select csv for file format

Microsoft Azure | Data Factory | projects-datafactory

Search factory and documentation

lets huntvartha@outlook.com  
DEFAULT DIRECTORY

Validate all Publish all

pipeline1 InputData dataflow1

Validate Data flow debug Debug Settings

source1 Import data from InputData

derivedColumn1 Creating/updating the columns Region, Country, Customer, Business Segment, Category, Model, Color, SalesDate...

sink1 Columns: 13 total

Sink Settings Errors Mapping Optimize Inspect Data preview

Output stream name \* sink1 Learn more

Description Add sink dataset Reset

Incoming stream \* derivedColumn1

Sink type \* Dataset Inline Cache

Dataset \* Select... + New

Options ☒ Allow schema drift ☐ Validate schema

Set properties

Name stagingbikedata

Linked service \* blob2adf\_ls

File path stagingdata / Directory / File name

First row as header ☒

Import schema ☒ From connection/store ☐ From sample file ☐ None

> Advanced

OK Back

Microsoft Azure | Data Factory | projects-datafactory

Search factory and documentation

lets huntvartha@outlook.com  
DEFAULT DIRECTORY

Validate all Publish all

pipeline1 InputData dataflow1

Validate Data flow debug Debug Settings

source1 Import data from InputData

derivedColumn1 Creating/updating the columns Region, Country, Customer, Business Segment, Category, Model, Color, SalesDate...

sink1 Columns: 13 total

Sink Settings Errors Mapping Optimize Inspect Data preview

Output stream name \* sink1 Learn more

Description Export data to stagingbikedata Reset

Incoming stream \* derivedColumn1

Sink type \* Dataset Inline Cache

Dataset \* stagingbikedata Test connection Open + New

Skip line count

Options ☒ Allow schema drift ☐ Validate schema

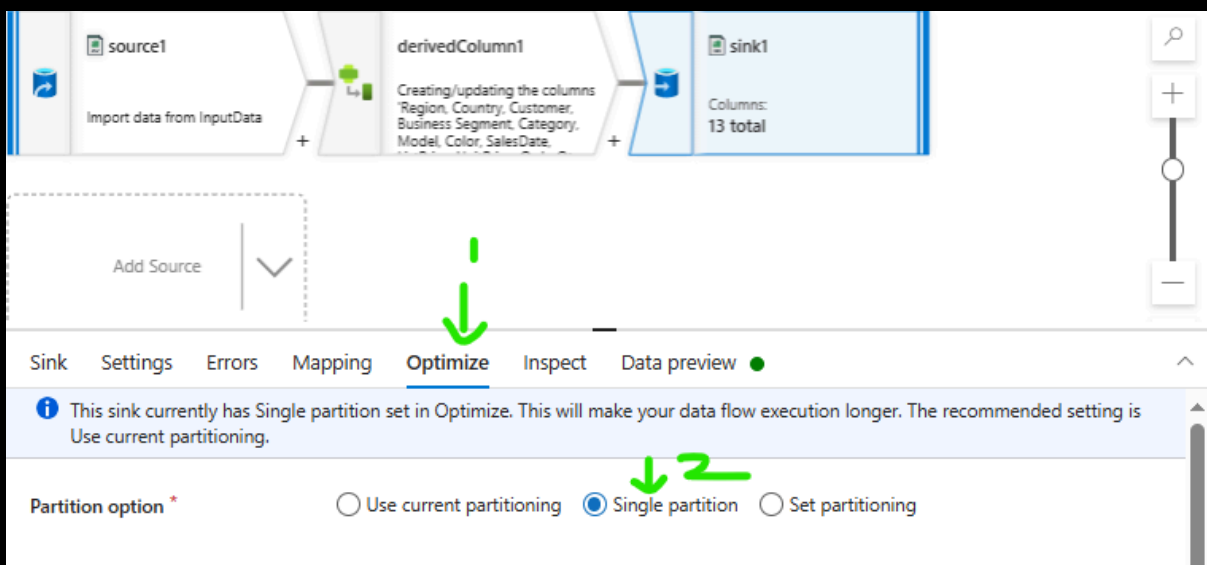
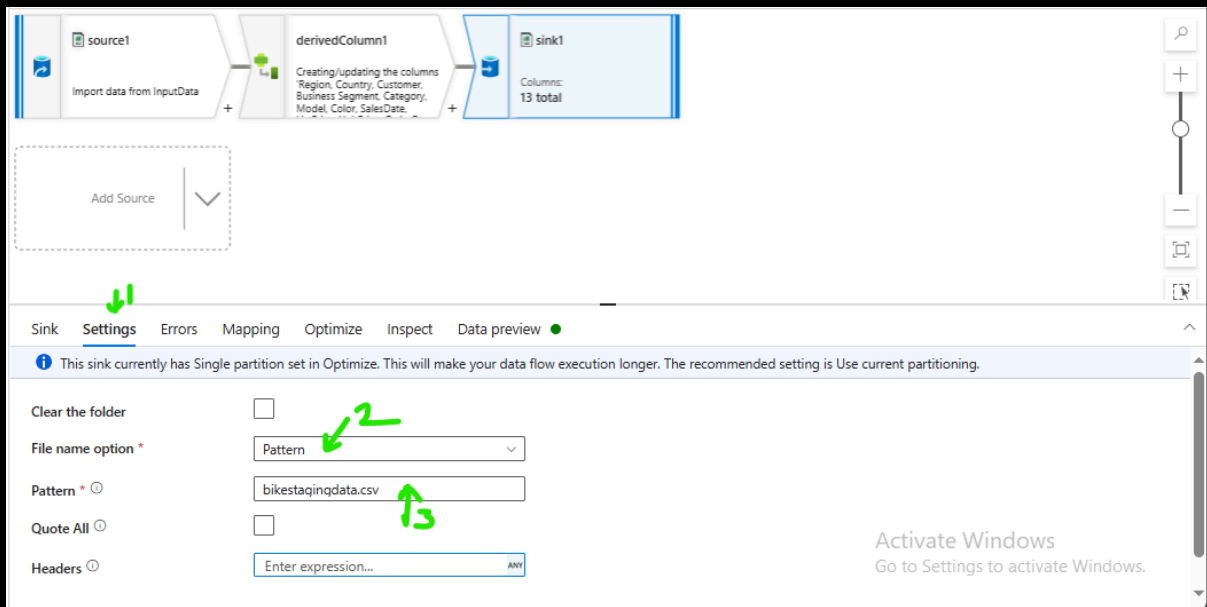
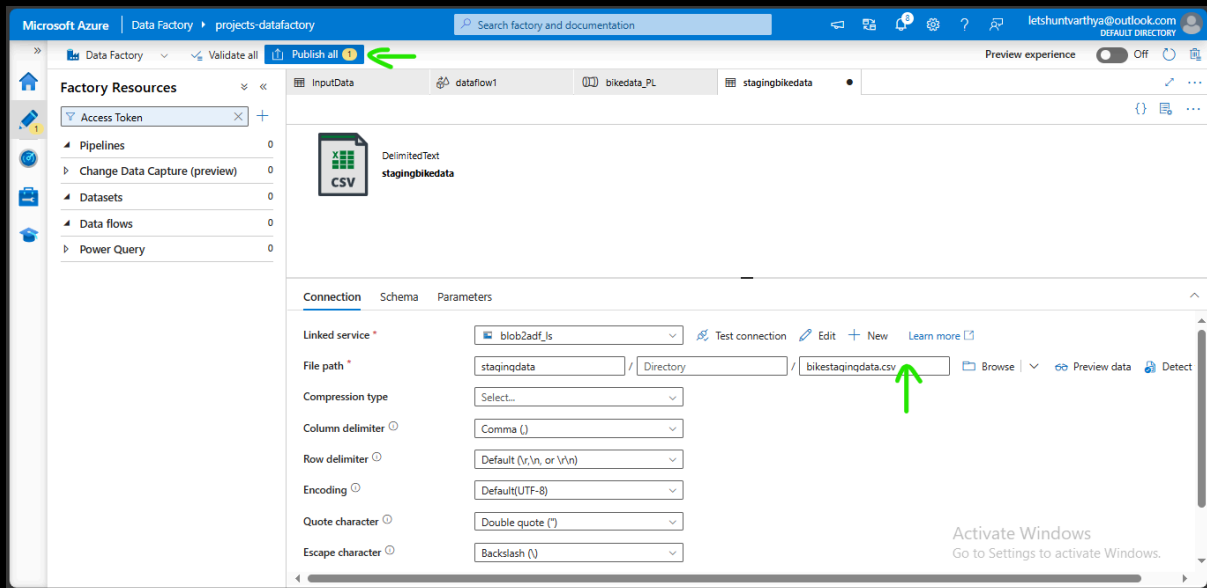
Properties

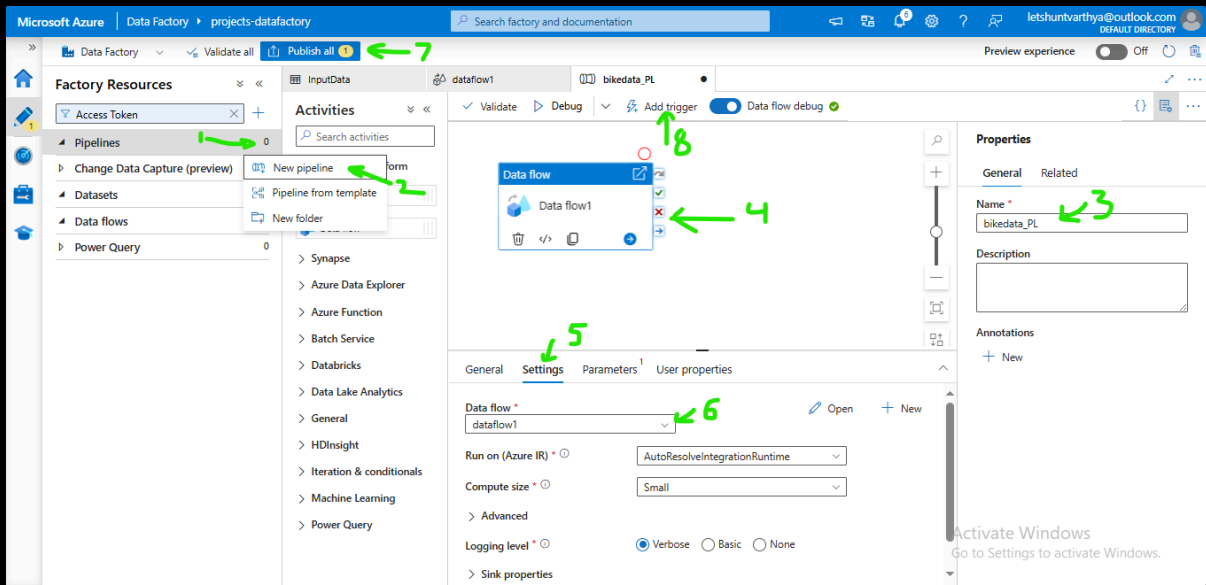
General Related

Name \* dataflow1

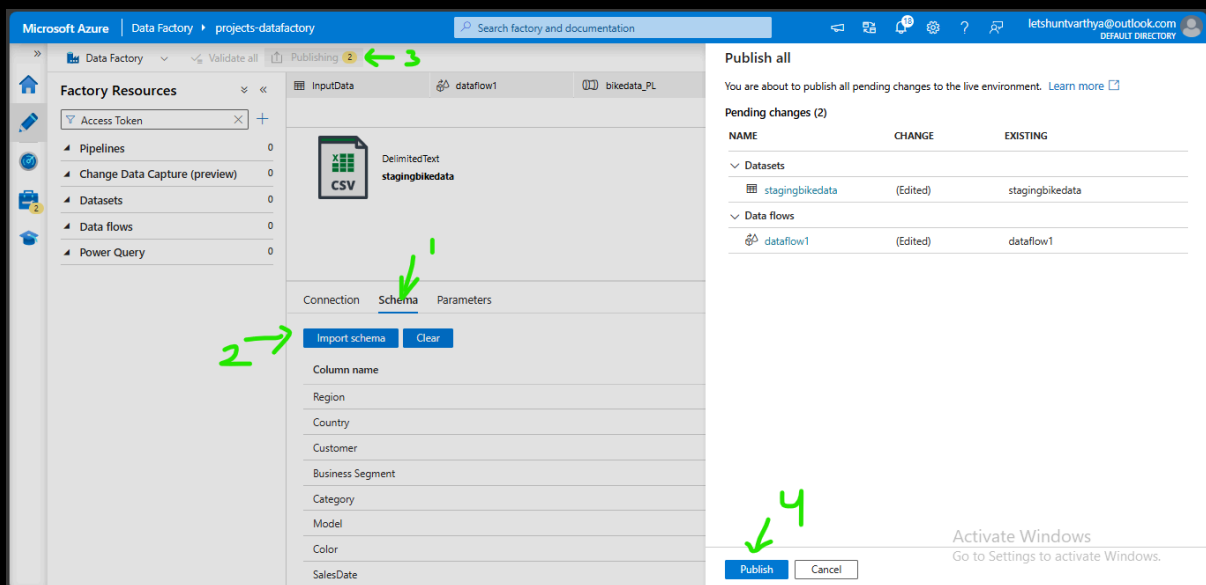
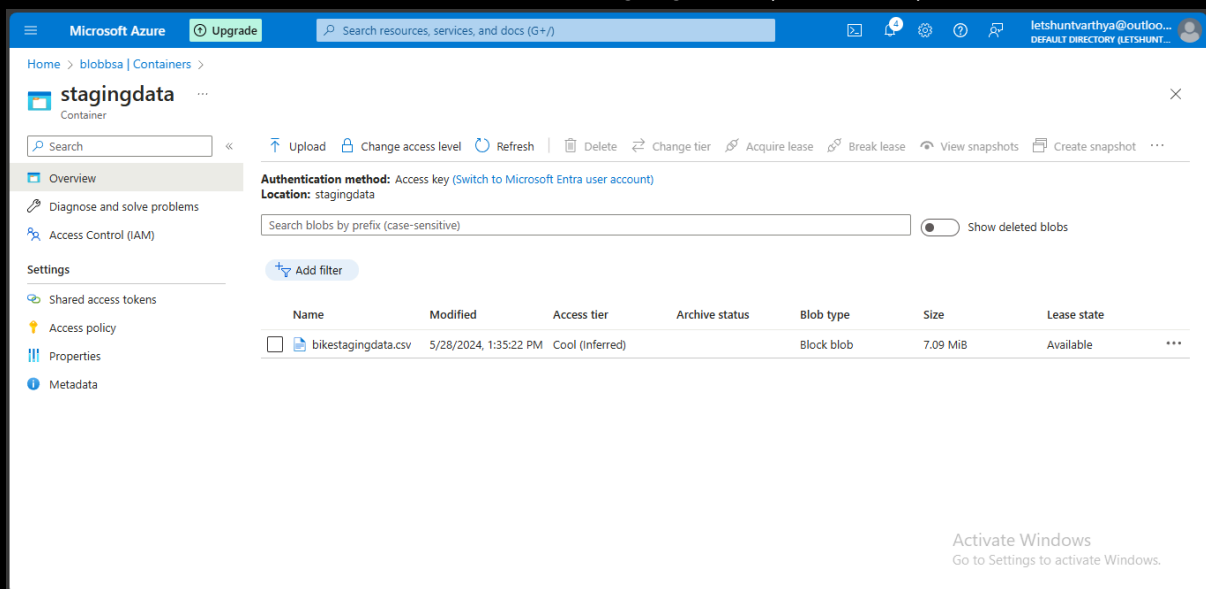
Description







- Data has been loaded into the stagingdata ( blobbsa ).



- Last transformations using power query.

The screenshot shows the Microsoft Power Query Editor interface. The 'Queries' pane on the left lists several queries, with 'Table.TransformColumnTypes' selected. The main area displays a table with columns: Business Segment, Category, and a list of columns (ListPrice, UnitPrice, OrderQty, Cost, sales). The 'Columns' dropdown is open, showing various data types like Decimal number, Currency, Whole number, Percentage, Date/Time, Date, Time, Date/Time/Zone, Duration, Text, True/False, and Binary. The 'Query settings' pane on the right shows the 'General' tab with the query name 'powerquery!' and a description field.

[illegible]

- Deleting the color column.

Home Transform Add column View Help

Enter data Manage connections Options Manage parameters Refresh Advanced editor Choose columns

Table.RenameColumns(#"Inserted subtraction 1", {"Subtract 1", "Color"})

Queries [2]

Business Segment	Category	Model	Color
Components	Road Frames	LL Road Frame	Red
Components	Mountain Bikes	Mountain-100	Silver
Components	Jerseys	Long-Sleeve Logo Jersey	Multi
Components	Mountain Frames	HL Mountain Frame	Black
Bikes	Road Bikes	Road-450	Red
Bikes	Road Bikes	Road-650	Black
Bikes	Road Bikes	Road-650	Red
Bikes	Road Bikes	Road-450	Red
Bikes	Road Bikes	Road-650	Black
Accessories	Helmets	Sport-100	Black
Bikes	Road Bikes	Road-650	Red
Bikes	Road Bikes	Road-650	Red

Completed (43.67 s) Columns: 14 Rows: 99+

Settings Parameters

Copy preview data  
Remove columns  
Remove other columns  
Duplicate column  
Add column from examples...  
Remove duplicates  
Remove errors  
Split column  
Replace values...  
Replace errors...  
Change type  
Transform column  
Group by...  
Fill  
Unpivot columns  
Unpivot other columns  
Unpivot only selected columns  
Rename...  
Move  
Drill down

Sort Transform Combine Export template

Query settings

1.2 UnitPrice 1.2 OrderQty 1.2 Cost 1.2 sales

Activate Windows  
Go to Settings to activate Windows.

Home Transform Add column View Help

Enter data Manage connections Options Manage parameters Refresh Advanced editor Choose columns Remove columns Keep rows Remove rows Filter rows Sort Split column Group by Data type: Text Use first row as headers Replace values Merge queries Append queries Combine files Export template

Table.RemoveColumns(#"Renamed columns 1", {"Color"})

Queries [2]

Region	Country	Customer	Business Segment	Category	Model	SalesDate	UnitPrice	OrderQty	Cost	sales	Profit
North America	United States	Advanced Bike Components	Components	Road Frames	LL Road Frame	2020-04-01	337.22	183.94	1	183.94	337.22
North America	United States	Central Discount Store	Bikes	Mountain Bikes	Mountain-100	2020-04-01	3399.99	2039.99	1	2039.99	3399.99
North America	United States	Leading Sales & Repair	Clothing	Jerseys	Long-Sleeve Logo Jersey	2020-04-01	49.99	28.84	6	173.04	299.94
North America	United States	Paint Supply	Components	Mountain Frames	HL Mountain Frame	2020-04-01	1349.6	714.7	2	1429.4	2699.2
North America	United States	Scooters and Bikes Store	Bikes	Road Bikes	Road-450	2020-04-01	1457.99	874.79	2	1749.58	2915.9
North America	United States	Scooters and Bikes Store	Bikes	Road Bikes	Road-650	2020-04-01	782.99	419.46	2	838.92	1565.9
North America	United States	Scooters and Bikes Store	Bikes	Road Bikes	Road-650	2020-04-01	782.99	419.46	4	1677.84	3131.96
North America	United States	Modern Bike Store	Bikes	Road Bikes	Road-450	2020-04-01	1457.99	874.79	1	874.79	1457.9
North America	United States	Modern Bike Store	Bikes	Road Bikes	Road-650	2020-04-01	782.99	419.46	2	838.92	1565.9
North America	Canada	Corner Bicycle Supply	Accessories	Helmets	Sport-100	2020-04-01	34.99	20.19	5	100.95	174.95
North America	Canada	Metal Processing Company	Bikes	Road Bikes	Road-650	2020-04-01	782.99	419.46	1	419.46	782.9
North America	United States	Excellent Riding Supplies	Bikes	Road Bikes	Road-650	2020-04-01	782.99	419.46	1	419.46	782.9

Query settings

Microsoft Azure Data Factory projects-datafactory

Search factory and documentation

Validate all Publish all

Factory Resources

- Pipelines
  - bikedata\_PL
- Change Data Capture (preview) 0
- Datasets 2
  - sourcedata
  - stagingbikedata
- Data flows 1
  - dataflow1
- Power Query 1
  - bikeData\_PQ

Activities

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

Data flow

Power Query1

General Settings Sink User properties

Power Query \*

Run on (Azure IR) \*

Compute type \*

Core count \*

Staging

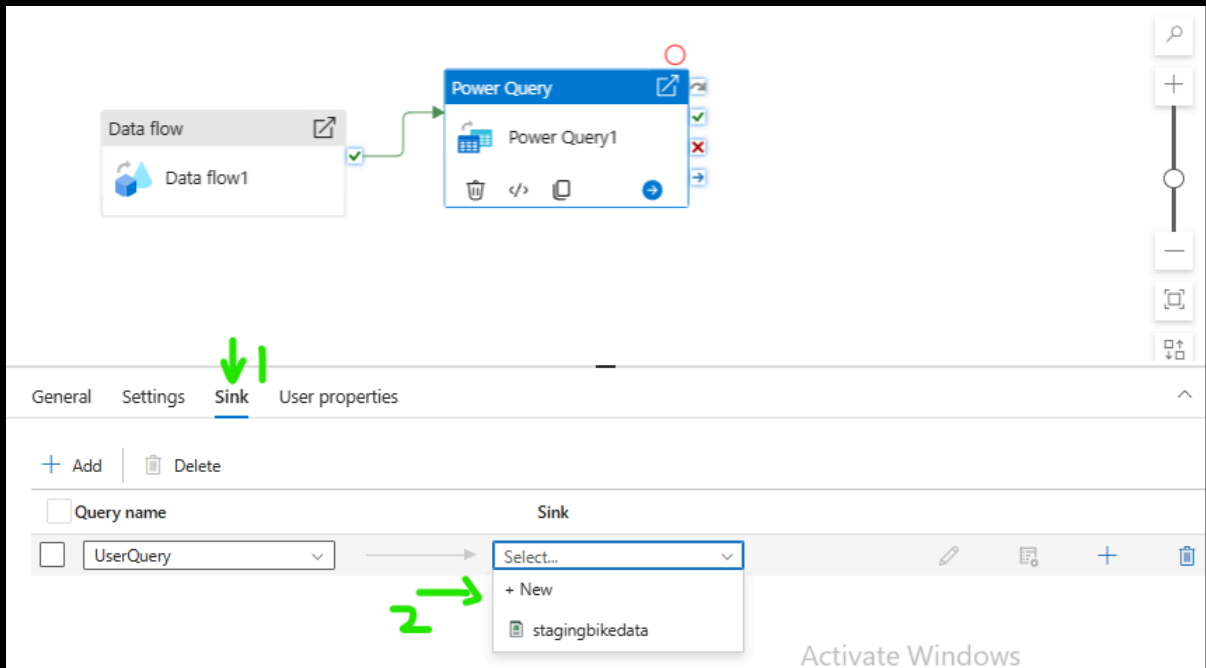
bikeData\_PQ

AutoResolveIntegrationRuntime

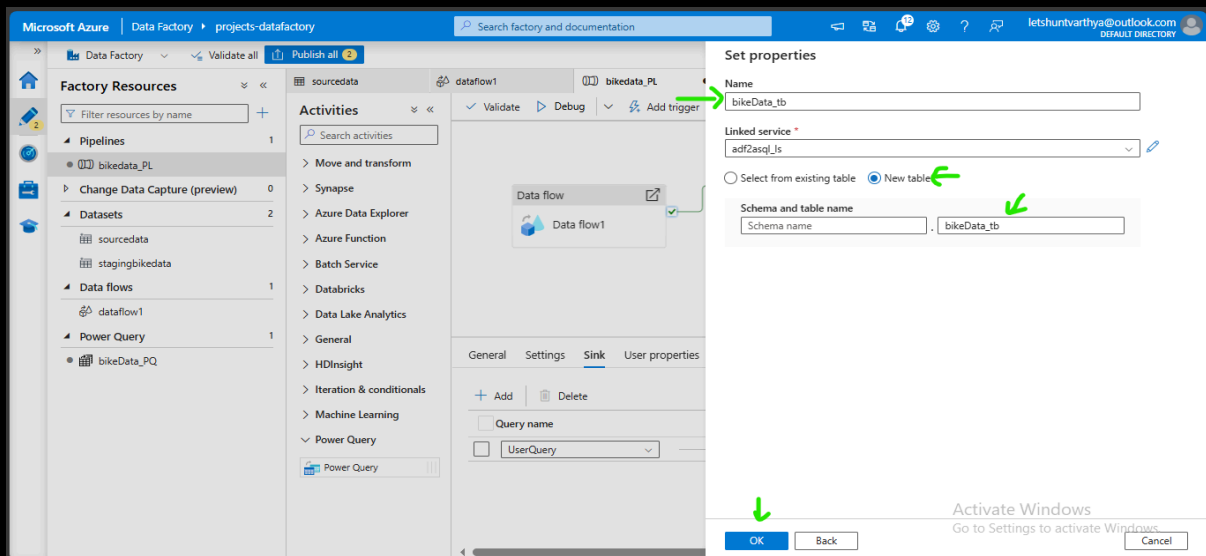
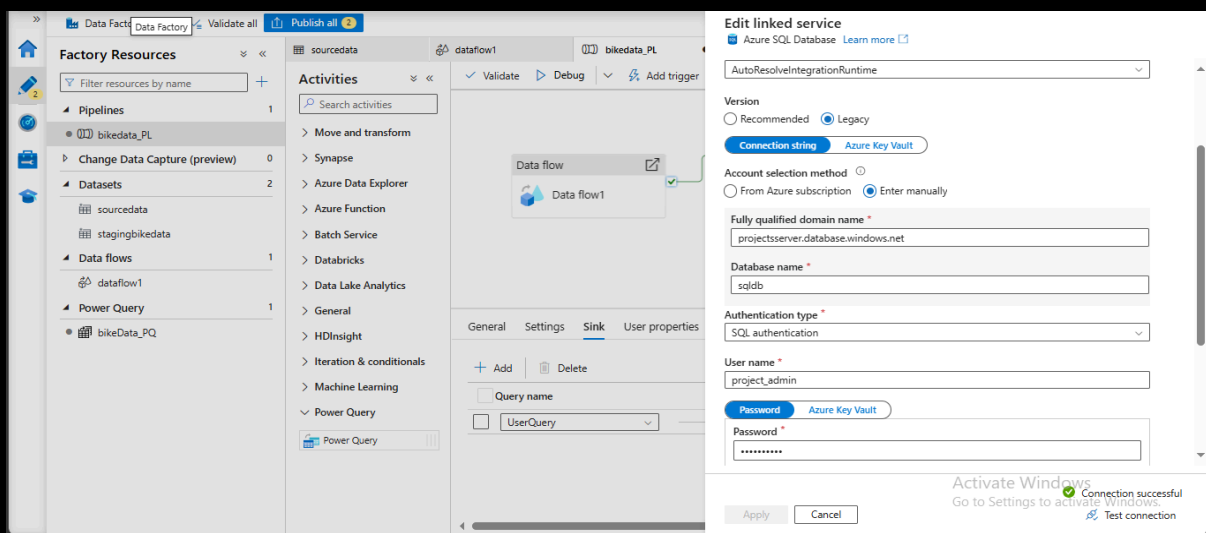
General purpose

4 (+ 4 Driver cores)

Activate Windows  
Go to Settings to activate Windows.



## New Dataset → Azure SQL Database



- Publish all → publish
- Add Trigger → Trigger now → ok

**Activity runs**

Pipeline run ID 1472be41-817c-4b55-a4f8-611312a064cf

All status ▾ Monitor in Azure Metrics Export to CSV

Showing 1 - 2 items

Activity name	Activity status	Activity type	Run start	Duration	Integration runtime	User properties	Activity run ID
Power Query1	Succeeded	Power Query	5/28/2024, 3:16:54 PM	1m 23s	AutoResolveIntegration		22fbae22-8580-4e82-b
Data flow1	Succeeded	Data flow	5/28/2024, 3:14:06 PM	2m 48s	AutoResolveIntegration		427ea7c0-0c9a-4b93-8

- Lets see the data in Azure sql Database

**SQLdb (projectsserver/sqlldb) | Query editor (preview)**

Search resources, services, and docs (G+)

Query editor (preview) is a tool to run SQL queries against Azure SQL Database in the Azure portal. It is designed for lightweight querying and object exploration in your database. For more information and troubleshooting, [Learn more](#)

Welcome to SQL Database Query Editor

SQL server authentication

Login \*

Password \*

Microsoft Entra authentication

[Continue as letshuntvarthya@outlook.c...](#)

OR

[OK](#)

Microsoft Azure

Upgrade

Search resources, services, and docs (G+)

letshuntvarthya@outloo...  
DEFAULT DIRECTORY (LETSHUNT...

Home > sqlldb (projectsserver/sqlldb) | Query editor (preview) ☆ ...

SQL database

> Login + New Query ↑ Open query Feedback Getting started

sqlldb (project\_a...  
Query 3  
Run Cancel query Save query Export data as Show only Editor  
1 SELECT TOP (1000) \* FROM [dbo].[bikeData\_tb]  
Showing limited object explorer here. For full capability please click here to open Azure Data Studio.  
Tables  
dbo.bikeData\_th  
Views  
Stored Procedures  
Select Top 1000 Rows  
Edit Data (Preview)  
Rename  
Country Customer Business Segment Category Model SalesDate  
North America United States Thorough Parts and Repair Services Components Road Frames LL Road Frame 2018-11-0  
North America United States Thorough Parts and Repair Services Accessories Cleaners Bike Wash 2018-11-0  
Query succeeded | 1s