

## Challenge #2

# Moving Data / Azure Data Factory

2022

### IN THIS CHALLENGE:

1. Data Factory
2. Data Pipeline
3. Query Editor
4. Azure Data Studio

Go to Kaggle and download the dataset:

<https://www.kaggle.com/saurabhshahane/twitter-sentiment-dataset>

Load your file in your Azure Storage Account.



## Twitter Sentiment Dataset

Twitter Sentiment Analysis



Saurabh Shahane • updated 7 days ago (Version 1)

^

7

Home - Microsoft Azure x + ▼ - □ X

portal.azure.com/#home ★

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

### Azure services

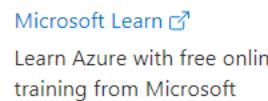
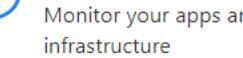
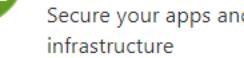
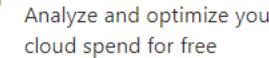
Create a resource  

 SQL databases  Subscriptions  Quickstart Center  Virtual machines  App Services  Storage accounts  Azure Cosmos DB  Kubernetes services  More services

### Navigate

 Subscriptions  Resource groups  All resources  Dashboard

### Tools

 Microsoft Learn   Azure Monitor   Security Center   Cost Management 

### Useful links

[Technical Documentation](#)  [Azure Services](#)  [Recent Azure Updates](#)   
[Azure Migration Tools](#) [Find an Azure expert](#) [Quickstart Center](#)

### Azure mobile app

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

Home > Create a resource > Marketplace >

Data Factory ⚡ ...

Microsoft



Data Factory | Add to Favorite

Microsoft

★★★★☆ 3.4 (472 reviews)

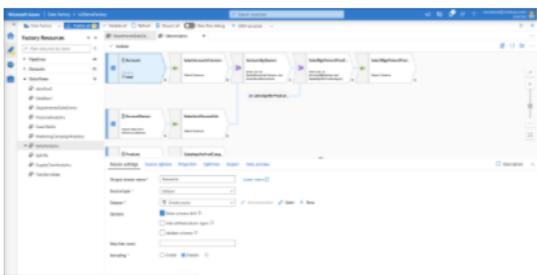
Create

[Overview](#)   [Plans](#)   [Usage Information + Support](#)   [Review](#)

Integrate data silos with Azure Data Factory, a service built for all data integration needs and skill levels. Easily construct ETL and ELT processes code-free within the intuitive visual environment, or write your own code. Visually integrate data sources using more than 90+ natively built and maintenance-free connectors at no added cost. Focus on your data - the serverless integration service does the rest.

- No code or maintenance required to build hybrid ETL and ELT pipelines within the Data Factory visual environment
  - Cost-efficient and fully managed serverless cloud data integration tool that scales on demand
  - Azure security measures to connect to on-premises, cloud-based, and software-as-a-service apps with peace of mind
  - SSIS integration runtime to easily rehost on-premises SSIS packages in the cloud using familiar SSIS tools

Media



Create Data Factory - Microsoft A X +

portal.azure.com/#create/Microsoft.DataFactory

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

Home > Create a resource > Marketplace > Data Factory >

## Create Data Factory

Basics Git configuration Networking Advanced Tags Review + create

### Project details

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

Subscription \* ⓘ Free Trial

Resource group \* ⓘ rg-caio-ai

Create new

Instance details

Region \* ⓘ Brazil South

Name \* ⓘ df-caio-001

Version \* ⓘ V2

Review + create < Previous Next : Git configuration >

Define the name of your instance Azure Data Factory

The screenshot shows the 'Create Data Factory' wizard on the Microsoft Azure portal. The 'Project details' section is active, displaying the 'Subscription' (Free Trial) and 'Resource group' (rg-caio-ai). The 'Instance details' section includes fields for 'Region' (Brazil South), 'Name' (df-caio-001), and 'Version' (V2). A red box highlights the 'Resource group' dropdown, and an orange arrow points from the 'Name' input field towards it, likely indicating a dependency or a common configuration step.

Create Data Factory - Microsoft A X +

portal.azure.com/#create/Microsoft.DataFactory

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

Home > Create a resource > Marketplace > Data Factory >

## Create Data Factory

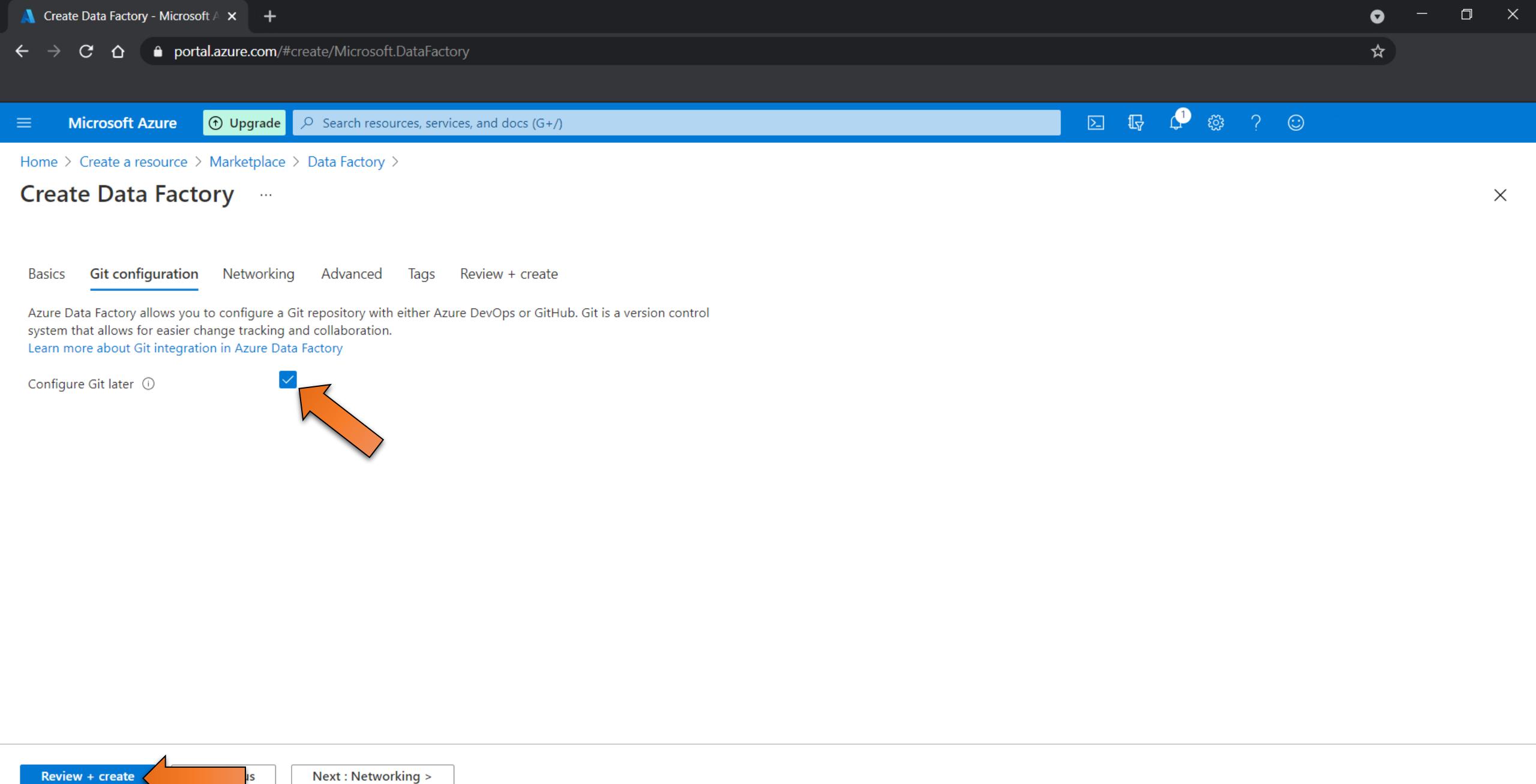
Basics Git configuration Networking Advanced Tags Review + create

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 ⓘ

Review + create  Next : Networking >



Select the option <Configure Git later>

Create Data Factory - Microsoft A X +

portal.azure.com/#create/Microsoft.DataFactory

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

Home > Create a resource > Marketplace > Data Factory >

## Create Data Factory

Validation Passed

Basics Git configuration Networking Advanced Tags Review + create

### 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. See the [Azure Marketplace Terms](#) for additional details.

### Basics

Subscription	Free Trial
Resource group	rg-caio-ai
Region	Brazil South
Name	df-caio-001
Version	V2

**Create**  **Next** Download a template for automation

# Create your Azure Data Factory

Microsoft.DataFactory-20210521122408 x +

portal.azure.com/#blade/HubsExtension/DeploymentDetailsBlade/overview/id/%2Fsubscriptions%2Fbf350b48-f641-49a2-b7e4-598ce72df8ab%2FresourceGroups%2Frg-caio-ai%2Fproviders...

Microsoft Azure Upgrade Search resources, services, and docs (G+/-) ☰ ⓘ 2 🔍 ⚙️ ? 😊

Home >

## Microsoft.DataFactory-20210521122408 | Overview

Deployment

Search (Ctrl+ /) « Delete Cancel Redeploy Refresh

We'd love your feedback! →

✓ Your deployment is complete

Deployment name: Microsoft.DataFactory-20210521122408  
Subscription: Free Trial  
Resource group: rg-caio-ai

Start time: 5/21/2021, 12:27:06 PM  
Correlation ID: ba37d1d3-1f83-48c3-a9a3-6f3c969c0b04

Deployment details (Download)

Next steps

Go to resource

Security Center

Secure your apps and infrastructure  
[Go to Azure security center >](#)

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 >](#)

Azure Data Factory created!

df-caio-001 - Microsoft Azure +    

portal.azure.com/#@caiogasparinegmail.onmicrosoft.com/resource/subscriptions/bf350b48-f641-49a2-b7e4-598ce72df8ab/resourceGroups/rg-caio-ai/providers/Microsoft.DataFactory/facto... 

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

Home > Microsoft.DataFactory-20210521122408 >

 df-caio-001  

    Delete

 Overview 

 Activity log

 Access control (IAM)

 Tags

 Diagnose and solve problems

**Settings**

 Networking

 Properties

 Locks

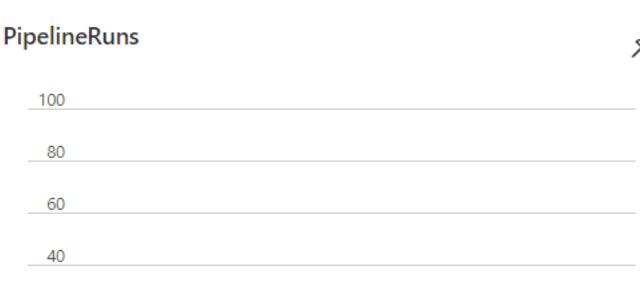
**Getting started**

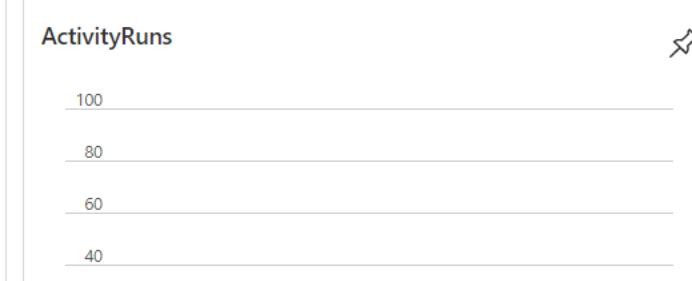
 Quick start

**Monitoring**

 Documentation  Author & Monitor

**Monitoring**

 PipelineRuns  
100  
80  
60  
40

 ActivityRuns  
100  
80  
60  
40

Azure Data Factory up and running...



storageaccountcaio - Microsoft A X +

portal.azure.com/#@caiogasparinegmail.onmicrosoft.com/resource/subscriptions/bf350b48-f641-49a2-b7e4-598ce72df8ab/resourcegroups/rg-caio-ai/providers/Microsoft.Storage/storageA... ☆

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

Home > storageaccountcaio

Storage account

Search (Ctrl+ /) Open in Explorer Delete Move Refresh Feedback

Microsoft recommends upgrading to the new alerts platform to ensure no interruptions in your alerts. Classic alerts will be retired starting in 2021. Upgrade to the new alerts platform. [Learn more](#)

JSON View

Essentials

Resource group (change) : rg-caio-ai	Performance/Access tier : Standard/Hot
Location : East US	Replication : Locally-redundant storage (LRS)
Subscription (change) : Free Trial	Account kind : StorageV2 (general purpose v2)
Subscription ID : bf350b48-f641-49a2-b7e4-598ce72df8ab	Provisioning state : Succeeded
Disk state : Available	Created : 5/20/2021, 11:20:33 AM

Tags (change) :

Properties Monitoring Capabilities (7) Recommendations Tutorials Developer Tools

Storage Explorer (preview)

Containers File shares Queues Tables

Blob service

Hierarchical namespace	Disabled	Secure transfer required	Enabled
Default access tier	Hot	Storage account key access	Enabled
Blob public access	Enabled	Minimum TLS version	Version 1.2
Blob soft delete	Enabled (7 days)	Infrastructure encryption	Enabled
Container soft delete	Disabled		
Versioning	Disabled		
Change feed	Disabled		

Security

Secure transfer required	Enabled
Storage account key access	Enabled
Minimum TLS version	Version 1.2
Infrastructure encryption	Enabled

Networking

Allow access from	All networks
-------------------	--------------

https://portal.azure.com/#@caiogasparinegmail.onmicrosoft.com/resource/subscriptions/bf3...

Select Storage Explorer (preview)...

A storageaccountcaio - Microsoft A X +

portal.azure.com/#@caigasparinegmail.onmicrosoft.com/resource/subscriptions/bf350b48-f641-49a2-b7e4-598ce72df8ab/resourcegroups/rg-caio-ai/providers/Microsoft.Storage/storageA... ☆

Microsoft Azure Upgrade Search resources, services, and docs (G+/-) ☰ 5 ⓘ ? ☺

Home > storageaccountcaio

## storageaccountcaio | Storage Explorer (preview)

Storage account

Search (Ctrl+/) Search

BLOB CONTAINERS my-container

FILE SHARES QUEUES TABLES

Upload Download Open New Folder Copy URL Select All Copy Paste More

Use the UPLOAD option to load your file from your local computer to your BLOB STORAGE

← → Active blobs (default) my-container Search by prefix (case-sensitive) ↗

NAME	ACCESS TIER	ACCESS TIER LAST MODIFIED	LAST MODIFIED	BLOB TYPE	CONTENT TYPE	SIZE	STATUS	REMAINING DAYS	DELETED TIL
NOTICE.txt	Cool	5/21/2021, 12:33:48 PM	5/20/2021, 1:16:25 PM	Block Blob	text/plain	358.9 KB	Active		
Twitter_Data.csv	Cool	5/21/2021, 12:33:58 PM	5/21/2021, 12:33:01 PM	Block Blob	application/vnd.ms-excel	19.9 MB	Active		

← → Active blobs (default) my-container Search by prefix (case-sensitive) ↗

Showing 1 to 2 of 2 cached items

Select your Blob Storage and load your file

The screenshot shows the Azure Storage Explorer (preview) interface. On the left, there's a sidebar with various navigation links like Overview, Activity log, Tags, etc., and a 'Data storage' section with 'Containers', 'File shares', 'Queues', and 'Tables'. The main area is titled 'storageaccountcaio | Storage Explorer (preview)' and shows a 'BLOB CONTAINERS' section with a single container named 'my-container'. An orange arrow points to this container. At the top of the main pane, there are buttons for 'Upload', 'Download', 'Open', 'New Folder', 'Copy URL', 'Select All', 'Copy', 'Paste', and 'More'. A yellow callout box says 'Use the UPLOAD option to load your file from your local computer to your BLOB STORAGE'. Below the container list, there's a table showing two files: 'NOTICE.txt' and 'Twitter\_Data.csv'. The table includes columns for NAME, ACCESS TIER, ACCESS TIER LAST MODIFIED, LAST MODIFIED, BLOB TYPE, CONTENT TYPE, SIZE, STATUS, REMAINING DAYS, and DELETED TIL. Both files are listed as 'Block Blob' type with 'text/plain' and 'application/vnd.ms-excel' content types respectively. The status is 'Active' for both. At the bottom, it says 'Showing 1 to 2 of 2 cached items'.

Select your Blob Storage and load your file

A storageaccountcaio - Microsoft A X +

portal.azure.com/#@caigasparinegmail.onmicrosoft.com/resource/subscriptions/bf350b48-f641-49a2-b7e4-598ce72df8ab/resourcegroups/rg-caio-ai/providers/Microsoft.Storage/storageA... ☆

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

Home > storageaccountcaio

# storageaccountcaio | Storage Explorer (preview)

Storage account

Search (Ctrl+ /) Search

Upload Download Open New Folder Copy URL Select All Copy Paste More

BLOB CONTAINERS my-container

FILE SHARES QUEUES TABLES

NAME ACCESS TIER ACCESS TIER LAST MODIFIED LAST MODIFIED BLOB TYPE CONTENT TYPE SIZE STATUS REMAINING DAYS DELETED TIER

NAME	ACCESS TIER	ACCESS TIER LAST MODIFIED	LAST MODIFIED	BLOB TYPE	CONTENT TYPE	SIZE	STATUS	REMAINING DAYS	DELETED TIER
NOTICE.txt	Cool	5/21/2021, 12:33:48 PM	5/20/2021, 1:16:25 PM	Block Blob	text/plain	358.9 KB	Active		
Twitter_Data.csv	Cool	5/21/2021, 12:33:58 PM	5/21/2021, 12:33:01 PM	Block Blob	application/vnd.ms-excel	19.9 MB	Active		

Search by prefix (case-sensitive)

Change your **ACCESS TIER** to save money.  
We are not concerned about performance.

Containers File shares Queues Tables

Networking Azure CDN

Showing 1 to 2 of 2 cached items

Change your ACCESS TIER and select HOME

## Azure services



## Recent resources

Name	Type	Last Viewed
storageaccountcaio	Storage account	7 minutes ago
df-caio-001	Data factory (V2)	11 minutes ago
rg-caio-ai	Resource group	15 minutes ago
sql-caio-ai1	SQL database	a day ago

## Navigate



Subscriptions



Resource groups



All resources



Dashboard

## Tools



Microsoft Learn



Azure Monitor



Security Center



Cost Management

df-caio-002 - Microsoft Azure

portal.azure.com/#@caiogasparinegmail.onmicrosoft.com/resource/subscriptions/5342ccb2-6e74-4d46-ab21-889285fad578/resourceGroups/resource-group-ai1006/providers/Mic...

Microsoft Azure

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

Home > df-caio-002 Data factory (V2)

Create a resource

Home

Dashboard

All services

**FAVORITES**

All resources

Resource groups

App Services

Function App

SQL databases

Azure Cosmos DB

Virtual machines

Load balancers

Storage accounts

Virtual networks

Azure Active Directory

Monitor

Advisor

Microsoft Defender for Cloud

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Settings

Networking

Managed identities

Properties

Locks

Getting started

Quick start

Monitoring

Alerts

Metrics

Delete

Essentials

Resource group (move) : [resource-group-ai1006](#) Type : Data factory (V2)  
Status : Succeeded Getting started : [Quick start](#)  
Location : East US  
Subscription (move) : [MSDN Platforms Subscription](#)  
Subscription ID : 5342ccb2-6e74-4d46-ab21-889285fad578

Getting started

[Open Azure Data Factory Studio](#)  
[Start authoring and monitoring your data pipelines and data flows.](#)  
[Open](#)

Read documentation  
Learn how to be productive quickly. Explore concepts, tutorials, and samples.  
[Learn more](#)

Monitoring

PipelineRuns

ActivityRuns

JSON View

Select < Open Azure Data Factory Studio >

<https://adf.azure.com/en/home?factory=%2Fsubscriptions%2F5342ccb2-6e74-4d46-ab21-889285fad578%2FresourceGroups%2Fresource-group-ai1006%2Fproviders%2FMicrosoft.DataFactory%2Ffactories%2Fdf-caio-002#loginHint=caiogasparine@gmail.com>

df-caio-002 - Microsoft Azure df-caio-002 - Azure Data Factory + [adf.azure.com/en/home?factory=%2Fsubscriptions%2F5342ccb2-6e74-4d46-ab21-889285fad578%2FresourceGroups%2Fresource-group-ai1006%2Fproviders%2FMicrosoft.DataFa...](https://adf.azure.com/en/home?factory=%2Fsubscriptions%2F5342ccb2-6e74-4d46-ab21-889285fad578%2FresourceGroups%2Fresource-group-ai1006%2Fproviders%2FMicrosoft.DataFa...)   

Microsoft Azure | Data Factory ▶ df-caio-002     ? 

»  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

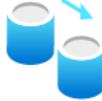
 Set up code repository

Data factory  
**df-caio-002**

New 

 **Ingest**  
Copy data at scale once or on a schedule.

 **Orchestrate**  
Code-free data pipelines.

 **Transform data**  
Transform your data using data flows.

 **Configure SSIS**  
Manage & run your SSIS packages in the cloud.

 Discover more

 [Browse partners \(preview\)](#)

 [Pipeline templates](#)

 [SAP pipeline templates](#)

Select < Ingest >

df-caio-002 - Microsoft Azure df-caio-002 - Azure Data Factory + [adf.azure.com/en/home/copywizard?factory=%2Fsubscriptions%2F5342ccb2-6e74-4d46-ab21-889285fad578%2FresourceGroups%2Fresource-group-ai1006%2Fproviders%2FMic...](https://adf.azure.com/en/home/copywizard?factory=%2Fsubscriptions%2F5342ccb2-6e74-4d46-ab21-889285fad578%2FresourceGroups%2Fresource-group-ai1006%2Fproviders%2FMic...)

**Microsoft Azure** | **Data Factory** ▶ **df-caio-002**

**Copy Data tool**

**Properties**

Use Copy Data Tool to perform a one-time or scheduled data load from 90+ data sources. Follow the wizard experience to specify your data loading settings, and let the Copy Data Tool generate the artifacts for you, including pipelines, datasets, and linked services. [Learn more](#)

**Properties**

Select copy data task type and configure task schedule

**Task type**

**Built-in copy task**  
You will get single pipeline to copy data from 90+ data source easily.

**Metadata-driven copy task**  
You will get parameterized pipelines which can read metadata from an external store to load data at a large scale.

You will get single pipeline to quickly copy objects from data source to destination in a very intuitive manner.

**Task cadence or task schedule \***

Run once now  Schedule  Tumbling window

< Previous **Next >** Cancel

Select < Built-in copy task > and < Next >

df-caio-002 - Microsoft Azure df-caio-002 - Azure Data Factory +

adf.azure.com/en/home/copywizard?factory=%2Fsubscriptions%2F5342ccb2-6e74-4d46-ab21-889285fad578%2FresourceGroups%2Fresource-group-ai1006%2Fproviders%2F...

Microsoft Azure | Data Factory ▶ df-caio-002

Copy Data tool

Properties

Source

Dataset

Configuration

Target

Settings

Review and finish

Source data store

Specify the source data store for the copy task. You can use an existing data store connection or specify a new data store.

Source type: Azure Blob Storage

Connection \*: Select... New connection

< Previous Next > Cancel

Select < Source Type > and < + New Connection >

df-caio-002 - Microsoft Azure df-caio-002 - Azure Data Factory + 

Microsoft Azure | Data Factory ▶ df-caio-002

Copy Data tool

Source data store

Specify the source data store for the copy task. You can use an existing data store connection or specify a new one.

Source type: Azure Blob Storage

Connection \*: Select... New connection

Properties (1) Source (2) Dataset (3) Configuration (4) Target (5) Settings (6) Review and finish

< Previous Next >

New linked service

Azure Blob Storage [Learn more](#)

Name \*: AzureBlobStorage1

Description:

Connect via integration runtime \*: AutoResolveIntegrationRuntime

Authentication type: Account key

Connection string (selected) Azure Key Vault

Account selection method:

From Azure subscription  Enter manually

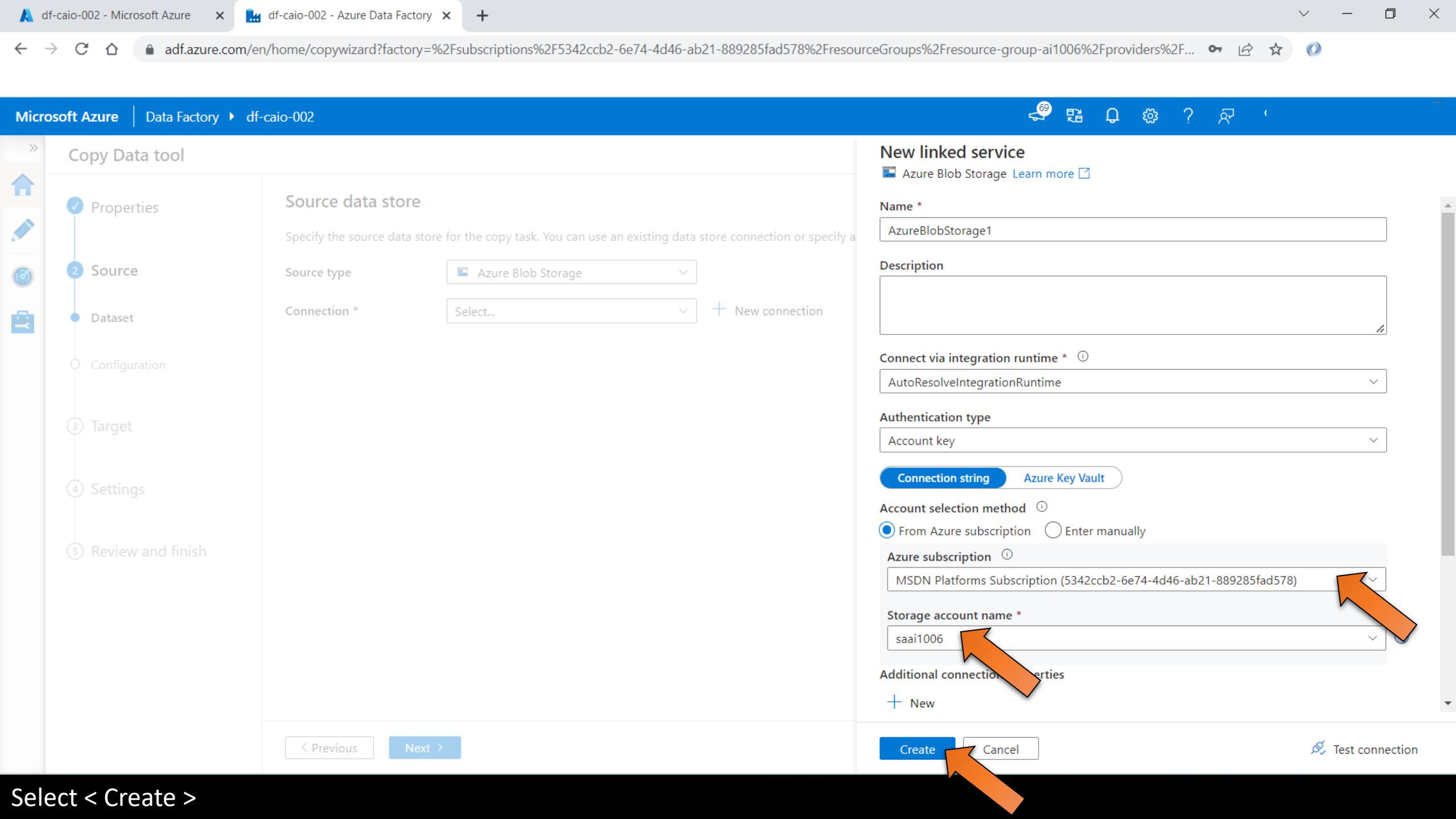
Azure subscription: MSDN Platforms Subscription (5342ccb2-6e74-4d46-ab21-889285fad578)

Storage account name \*: saai1006

Additional connection properties: + New

Create Cancel Test connection

Select < Create >



df-caio-002 - Microsoft Azure df-caio-002 - Azure Data Factory +  adf.azure.com/en/home/copywizard?factory=%2Fsubscriptions%2F5342ccb2-6e74-4d46-ab21-889285fad578%2FresourceGroups%2Fresource-group-ai1006%2Fproviders%2F...

Microsoft Azure | Data Factory ▶ df-caio-002

Copy Data tool

Properties

Source

Dataset

Configuration

Target

Settings

Review and finish

Source data store

Specify the source data store for the copy task. You can use an existing data store connection or specify a new data store.

Source type: Azure Blob Storage

Connection: AzureBlobStorage1

File or folder: my-container/Twitter\_Data.csv

Options:

- Binary copy
- Recursively
- Enable partition discovery

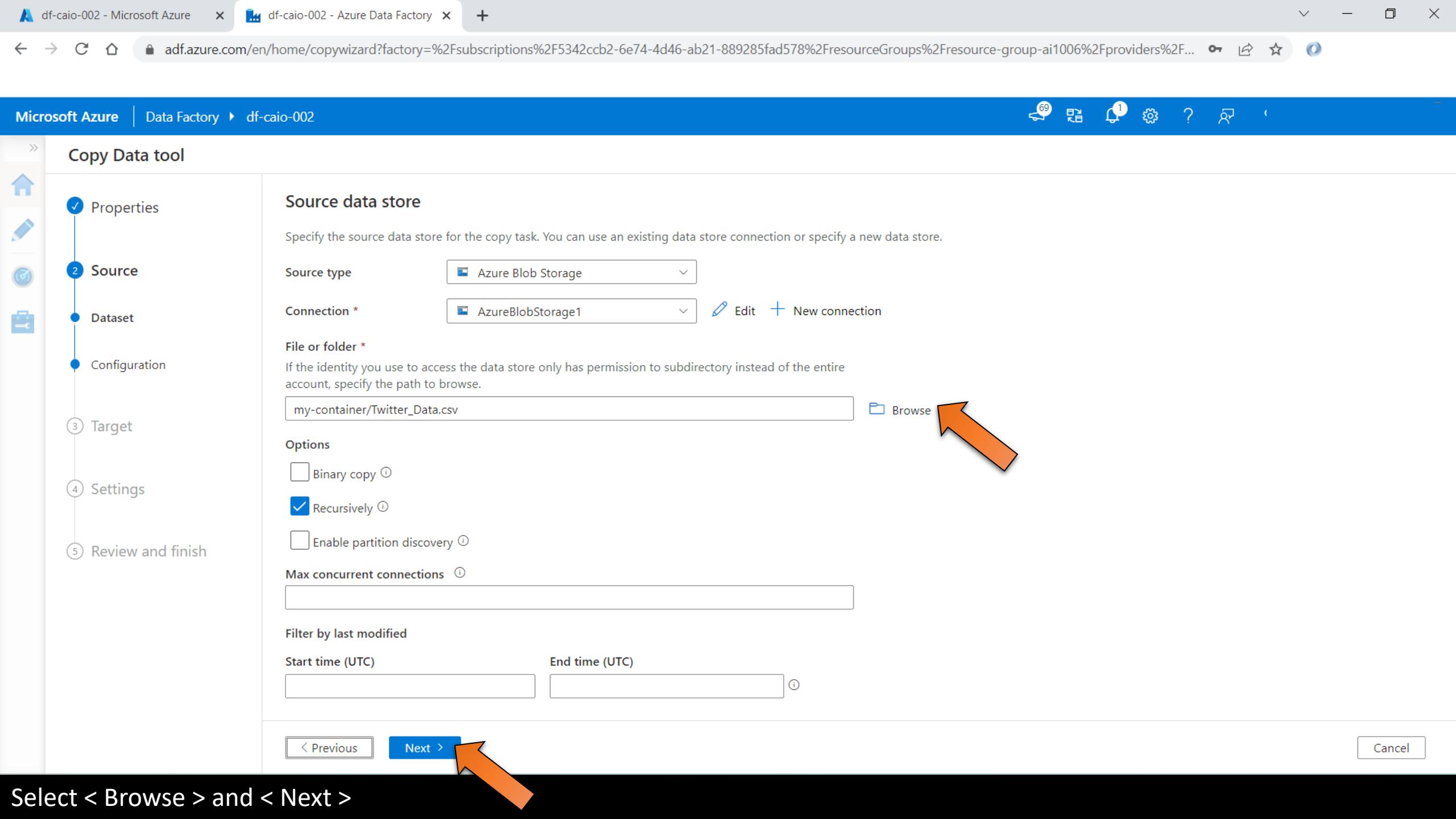
Max concurrent connections:

Filter by last modified

Start time (UTC):

End time (UTC):

< Previous Next > Cancel



Select < Browse > and < Next >

df-caio-002 - Microsoft Azure df-caio-002 - Azure Data Factory +

adf.azure.com/en/home/copywizard?factory=%2Fsubscriptions%2F5342ccb2-6e74-4d46-ab21-889285fad578%2FresourceGroups%2Fresource-group-ai1006%2Fproviders%2F...

Microsoft Azure | Data Factory ▶ df-caio-002

Copy Data tool

Properties

Source

Dataset

Configuration

Target

Settings

Review and finish

File format settings

File format ⓘ

DelimitedText

Detect text format

Preview data

Column delimiter

Comma (,)

Edit

Row delimiter

Line feed (\n)

Edit

First row as header ⓘ

Advanced

Compression type

None

Additional columns ⓘ

New

< Previous

Next >

Cancel

Select < Next >



df-caio-002 - Microsoft Azure df-caio-002 - Azure Data Factory +

adf.azure.com/en/home/copywizard?factory=%2Fsubscriptions%2F5342ccb2-6e74-4d46-ab21-889285fad578%2FresourceGroups%2Fresource-group-ai1006%2Fproviders%2F...

Microsoft Azure | Data Factory ▶ df-caio-002

Copy Data tool

Properties ✓  
Source ✓  
Target 3  
Dataset  
Configuration  
Settings  
Review and finish

Destination data store

Specify the destination data store for the copy task. You can use an existing data store connection or specify a new data store.

Target type: All

Connection \*: Select... + New connection

< Previous Next > Cancel



Select < + New Connection >

df-caio-002 - Microsoft Azure df-caio-002 - Azure Data Factory + [adf.azure.com/en/home/copywizard?factory=%2Fsubscriptions%2F5342ccb2-6e74-4d46-ab21-889285fad578%2FresourceGroups%2Fresource-group-ai1006%2Fproviders%2F...](https://adf.azure.com/en/home/copywizard?factory=%2Fsubscriptions%2F5342ccb2-6e74-4d46-ab21-889285fad578%2FresourceGroups%2Fresource-group-ai1006%2Fproviders%2F...) 69 1 🔍 ⚙️ ? 🔍

Microsoft Azure | Data Factory ▶ df-caio-002

Copy Data tool

Destination data store

Specify the destination data store for the copy task. You can use an existing data store connection or specify a new one.

Target type: All

Connection \*: Select... New connection

Properties ✓

Source ✓

Target 3

Dataset

Configuration

Settings

Review and finish

< Previous Next >

New linked service

Search

All Azure Database File Generic protocol NoSQL Services and apps

Azure Database for MySQL	Azure Database for PostgreSQL	Azure Databricks Delta Lake
Azure File Storage	Azure SQL Database	SQL Database Managed Instance
Azure Search	Azure Synapse Analytics	Azure Table Storage

Select < Azure SQL Database > and < Continue >

Continue

Cancel

The screenshot shows the 'Copy Data tool' wizard in the Azure Data Factory portal. The current step is 'Target'. In the 'Destination data store' section, the 'Target type' dropdown is set to 'All' and the 'Connection' dropdown is set to 'Select...'. To the right, the 'New linked service' pane displays a grid of destination options under the 'Database' category. The 'Azure SQL Database' option is highlighted with a blue border and an orange arrow. Other visible options include 'Azure Database for MySQL', 'Azure Database for PostgreSQL', 'Azure Databricks Delta Lake', 'Azure File Storage', 'Azure Search', 'Azure Synapse Analytics', and 'Azure Table Storage'. At the bottom of the pane, there are 'Continue' and 'Cancel' buttons, with an orange arrow pointing to the 'Continue' button.

Home - Microsoft Azure df-caio-002 - Azure Data Factory

adf.azure.com/en/home/copywizard?factory=%2Fsubscriptions%2F5342ccb2-6e74-4d46-ab21-889285fad578%2FresourceGroups%2Fresource-group-ai1006%2Fproviders%2F...

Microsoft Azure | Data Factory > df-caio-002

Copy Data tool

Destination data store

Specify the destination data store for the copy task. You can use an existing data store connection or specify new connection details.

Target type: All

Connection \*: Select... New connection

Properties ✓  
Source ✓  
Target 3  
Dataset  
Configuration  
Settings  
Review and finish

Next >

New connection

Azure SQL Database Learn more

Name \*: AzureSqlDatabase1

Description:

Connect via integration runtime \* ⓘ: AutoResolveIntegrationRuntime

Connection string Azure Key Vault

Account selection method ⓘ

From Azure subscription  Enter manually

Azure subscription: MSDN Platforms Subscription (5342ccb2-6e74-4d46-ab21-889285fad578)

Server name \*: sql-db-server-ai1006

Database name \*: master

Authentication type \*: SQL authentication

Create Back Test connection Cancel

Select < Create >

Microsoft Azure | Data Factory ▶ df-caio-002

Copy Data tool

Properties  Source  Target  Dataset Configuration Settings Review and finish

Destination data store

Specify the destination data store for the copy task. You can use an existing data store connection or specify a new data store.

Target type: All

Connection \*: AzureSqlDatabase1  

Source: Azure Blob Storage file

Target:  [ ] . Azure Blob Storage file (auto-create)

[Use existing table](#)

Skip column mapping for all tables

< Previous  Next > Cancel

Select < Next >

sql-db-server-ai1006 - Microsoft df-caio-002 - Azure Data Factory

adf.azure.com/en/home/copywizard?factory=%2Fsubscriptions%2F5342ccb2-6e74-4d46-ab21-889285fad578%2FresourceGroups%2Fresource-group-ai1006%2Fproviders%2F...

Microsoft Azure | Data Factory ▶ df-caio-002

Copy Data tool

Properties

Source

**Target**

Dataset

Configuration

Settings

Review and finish

Column mapping

Choose how source and destination columns are mapped

Table mappings (1)

Source  
Azure Blob Storage file  
Target  
new-table.Azure Blob Storage file

Column mappings

Type conversion settings

New mapping Clear Reset Delete

Source	Type	Destination	Type
clean_text	String	clean_text	String
category	String	category	String

Azure SQL Database sink properties

Pre-copy script

Advanced

< Previous Next > Cancel

Select < Next >

## Copy Data tool

 Properties

 Source

 Target

 4 Settings

 5 Review and finish

### Settings

Enter name and description for the copy data task, more options for data movement

Task name \*

CopyPipeline\_ai1006



Task description

Data consistency verification



Fault tolerance



Enable logging



Enable staging



> Advanced

< Previous

Next >

Cancel

Select < Next >

Microsoft Azure | Data Factory ▶ df-caio-002

Copy Data tool

Properties

Source

Target

Settings

Review and finish

Review

Deployment

Summary

You are running pipeline to copy data from Azure Blob Storage to Azure SQL Database.

Azure Blob Storage → Azure SQL Database

Properties

Task name: CopyPipeline\_ai1006

Task description:

Source

Connection name: AzureBlobStorage1

Dataset name: SourceDataset\_u6y

Column delimiter: ,

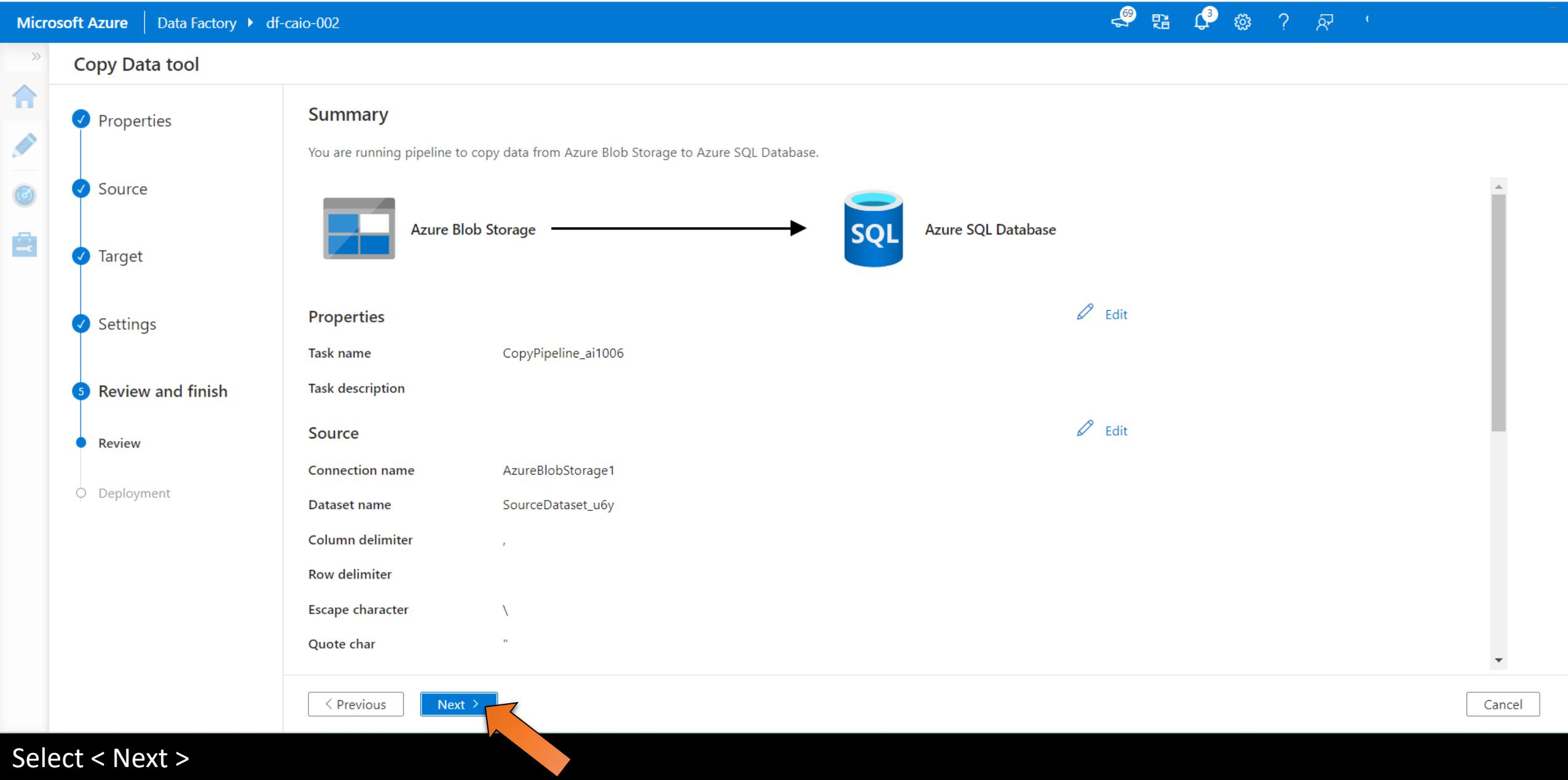
Row delimiter:

Escape character: \

Quote char: "

< Previous Next > Cancel

Select < Next >



Microsoft Azure | Data Factory ▶ df-caio-002

Copy Data tool

Properties  
Source  
Target  
Settings  
**Review and finish**  
Review  
Deployment

Azure Blob Storage → Azure SQL Database

**Deployment complete**

Deployment step	Status
Validating copy runtime environment	Succeeded
Creating datasets	Succeeded
Creating pipelines	Succeeded
Running pipelines	Succeeded

Datasets and pipelines have been created. You can now monitor and edit the copy pipelines or click finish to close Copy Data Tool.

**Finish** Edit pipeline Monitor

Select < Finish >

adf.azure.com/en/home/copywizard?factory=%2Fsubscriptions%2F5342ccb2-6e74-4d46-ab21-889285fad578%2FresourceGroups%2Fresource-group-ai1006%2Fproviders%2F...

Microsoft Azure | Data Factory > df-caio-002

Copy Data tool

Properties  
Source  
Target  
Settings  
Review and finish  
Review  
Deployment

Azure Blob Storage → Azure SQL Database

Deployment complete

Deployment step

- Validating copy runtime environment
- > Creating datasets
- > Creating pipelines
- > Running pipelines

Status

- Succeeded
- Succeeded
- Succeeded
- Succeeded

Datasets and pipelines have been created. You can now monitor and edit the copy pipelines or click finish to close Copy Data Tool.

Finish Edit pipeline Monitor

Select < Finish >

# What is Azure Data Factory?

<https://youtu.be/EpDkxTHAhOs>

<https://docs.microsoft.com/en-us/azure/data-factory/concepts-data-flow-performance>

;-)

## Challenge #2.2

Using Query Editor to visualize your data

## Azure services



## Recent resources

Name	Type	Last Viewed
Free Trial	Subscription	4 days ago
storageaccountcaio	Storage account	4 days ago
sa-caio-ai1	Synapse workspace	7 days ago
rg-caio-ai	Resource group	7 days ago
df-caio-001	Data factory (V2)	2 weeks ago
sql-caio-ai1	SQL database	2 weeks ago

## Navigate



To

<https://portal.azure.com/#@caioquamarine@mail.onmicrosoft.com/asset/SolAzureExtension/Database/subscriptions/bf350b48-f641-49a2-b7e4-598ce72df8ab/resourceGroups/ra-caio-ai/providers/Microsoft.Sql/servers/srv-caio-001/databases/sql-caio-ai1>

Select your SQL instance

sql-caio-ai1 (srv-caio-001/sql-caio-ai1) + More

portal.azure.com/#@caiogasparinegmail.onmicrosoft.com/resource/subscriptions/bf350b48-f641-49a2-b7e4-598ce72df8ab/resourceGroups/rg-caio-ai/providers/Microsoft.Sql/servers/srv-caio-ai1/databases/sql-caio-ai1/queryEditor

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

Home > sql-caio-ai1 (srv-caio-001/sql-caio-ai1) ⚡ ...

SQL database

Search (Ctrl+ /) Copy Restore Export Set server firewall Delete Connect with... Feedback

Overview Activity log Tags Diagnose and solve problems Quick start Query editor (preview)

Resource group (change) : rg-caio-ai Status : Online Location : Brazil South Subscription (change) : Free Trial Subscription ID : bf350b48-f641-49a2-b7e4-598ce72df8ab Tags (change) : Click here to add tags Server name : srv-caio-001.database.windows.net Elastic pool : No elastic pool Connection strings : Show database connection strings Pricing tier : Basic Earliest restore point : 2021-05-25 00:00 UTC

JSON View

Show data for last: 1 hour 24 hours 7 days Aggregation type: Max

Compute utilization

Database data storage 3.22% USED SPACE

Power Platform Power BI (preview) Power Apps (preview) Power Automate (preview)

Settings Compute + storage Connection strings Properties Locks

Data management

Used space 66 MB Allocated space 99 MB

Select Query Editor (preview)

Home > sql-caio-ai1 (srv-caio-001/sql-caio-ai1)

## sql-caio-ai1 (srv-caio-001/sql-caio-ai1) | Query editor (preview) ...

🔍 Search (Ctrl+ /)

>Login + New Query ↑ Open query 💡 Feedback

📊 Overview

💻 Activity log

🏷️ Tags

🔧 Diagnose and solve problems

☁️ Quick start

💻 Query editor (preview)

**Power Platform**

📊 Power BI (preview)

❖ Power Apps (preview)

➤ Power Automate (preview)

**Settings**

⌚ Compute + storage

🔗 Connection strings

☰ Properties

🔒 Locks

**Data management**

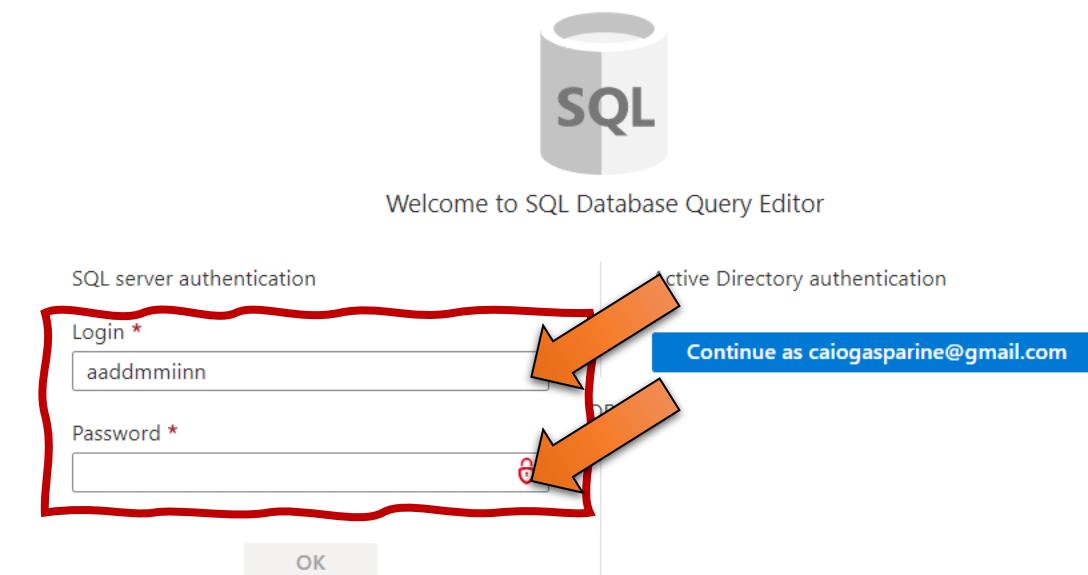
Welcome to SQL Database Query Editor

SQL server authentication Active Directory authentication

Login \* aaddmmiinn Continue as caiogasparine@gmail.com

Password \* ██████████

OK



Enter your user and password to connect with your SQL database

sql-caio-ai1 (srv-caio-001/sql-caio-ai1) + More

portal.azure.com/#@caiogasparinegmail.onmicrosoft.com/resource/subscriptions/bf350b48-f641-49a2-b7e4-598ce72df8ab/resourceGroups/rg-caio-ai/providers/Microsoft.Sql/servers/...

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

Home > sql-caio-ai1 (srv-caio-001/sql-caio-ai1)

## sql-caio-ai1 (srv-caio-001/sql-caio-ai1) | Query editor (preview) ...

SQL database

Search (Ctrl+/) Login New Query Open query Feedback

Overview Activity log Tags Diagnose and solve problems Quick start Query editor (preview)

sql-caio-ai1 (aaddmmiinn) ↻

Showing limited object explorer here. For full capability please open SSDT.

Tables Views Stored Procedures

Query 1 X

Run Cancel query Save query Export data as Show only Editor

1

Results Messages

Search to filter items...

Ready

Select your table to visualize the data...

The screenshot shows the Microsoft Azure SQL Database Query editor (preview) interface. On the left, there's a sidebar with various navigation options like Overview, Activity log, Tags, etc., and a section for Power Platform and Settings. The 'Query editor (preview)' option is selected. In the main content area, it displays a database named 'sql-caio-ai1 (aaddmmiinn)'. A message box says 'Showing limited object explorer here. For full capability please open SSDT.' Below this, there are links for 'Tables', 'Views', and 'Stored Procedures'. An orange arrow points from the 'Tables' link in the sidebar to the 'Tables' link in the main content area. The top navigation bar includes 'Search (Ctrl/)', 'Login', 'New Query', 'Open query', and 'Feedback'. The top right has icons for settings, notifications (1), help, and a smiley face. The bottom right shows a 'Ready' status. At the bottom, it says 'Select your table to visualize the data...'.

sql-caio-ai1 (srv-caio-001/sql-caio-ai1) + More

portal.azure.com/#@caiogasparinegmail.onmicrosoft.com/resource/subscriptions/bf350b48-f641-49a2-b7e4-598ce72df8ab/resourceGroups/rg-caio-ai/providers/Microsoft.Sql/servers/... Key 

Microsoft Azure Upgrade Search resources, services, and docs (G+) More Feedback Help Smile

Home > sql-caio-ai1 (srv-caio-001/sql-caio-ai1)

## sql-caio-ai1 (srv-caio-001/sql-caio-ai1) | Query editor (preview) X

SQL database

Search (Ctrl+/) Login New Query Open query Feedback

Overview Activity log Tags Diagnose and solve problems Quick start Query editor (preview)

sql-caio-ai1 (aaddmmiinn) Overview Activity log Tags Diagnose and solve problems Quick start Query editor (preview)

Showing limited object explorer here. For full capability please open SSDT.

Tables dbo.Azure Blob Storage file ... Select Top 1000 Rows Edit Data (Preview) Rename

Views sys.database\_firewall\_rules ... Edit Data (Preview) Rename

Stored Procedures Edit Data (Preview) Rename

Query 1 Query 2 Run  Cancel query Save query Export data as Show only Editor

1 SELECT TOP (1000) \* FROM [dbo].[Azure Blob Storage file]

Results Messages

Search to filter items...

clean_text	category
when modi promised "minimum government maximum governance" expec...	-1
talk all the nonsense and continue all the drama will vote for modi	0
what did just say vote for modi welcome bjp told you rahul the main camp...	1
asking his supporters prefix chowkidar their names modi did great service ...	1

Query succeeded! 10s

Use the command SELECT TOP 1000 ROWS to visualize your data

sql-caio-ai1 (srv-caio-001/sql-caio-ai1) + More

portal.azure.com/#@caiogasparinegmail.onmicrosoft.com/resource/subscriptions/bf350b48-f641-49a2-b7e4-598ce72df8ab/resourceGroups/rg-caio-ai/providers/Microsoft.Sql/servers/...

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

Home > sql-caio-ai1 (srv-caio-001/sql-caio-ai1)

## sql-caio-ai1 (srv-caio-001/sql-caio-ai1) | Query editor (preview) ...

SQL database

Search (Ctrl+/) Login New Query Open query Feedback

Overview Activity log Tags Diagnose and solve problems Quick start Query editor (preview)

sql-caio-ai1 (aaddmmiinn)

Showing limited object explorer here. For full capability please open SSDT.

Tables > dbo.Azure Blob Storage file ...

Views > sys.database\_firewall\_rules ...

Stored Procedures

Query 1 Query 2

Run Cancel query Save query Export data as Show only Editor

```
1 SELECT TOP (1000) * FROM [dbo].[blob Storage file]
2 WHERE clean_text LIKE '%high%'
3
4
```

Write a SQL statement and add a clause WHERE filtering the field **clean\_text** with the word **high**

Results Messages

Search to filter items...

clean_text	category
yes good job highly insensitivearrogant incompetent ploar needs defeated ...	1
higher voting turnout directly proportional bjp victory wonder modi has la...	1
modi came power because five yrs ago the youth india were takenin his pr...	1
malnutrition india and constituencies among worst affected mallikarjun kha...	-1

Query succeeded | 7s

Improve your SQL statement...

## Challenge #2.3

**Using Azure Data Studio to visualize your data**



CONNECTIONS

... SERVERS

srv-caio-001.database.windows.net, <default> (aaddmmiinn)  
> Databases  
> Security



Would you like to take a quick tour of Azure Data Studio?

Start Tour



## Azure Data Studio

New ▾

Open file...

Open folder...



## Create a connection

Connect to a database instance through the connection dialog.



## Run a query

Interact with data through a query editor.



## Create a notebook

Build a new notebook using a native notebook editor.



## Deploy a server

Create a new instance of a relational data service on the platform of your choice.

## Resources

## History



Name

Location

No recent folders

Show more

Change SQL language provider: srv-caio-001.database.windows.net : &lt;default&gt;



Using Azure Data Studio, locate your previous connection...

CONNECTIONS

...

SERVERS

srv-caio-001.database.windows.net, <default> (aaddmmiinn)

- Databases
  - System Databases
- sql-caio-ai1
  - Tables
    - dbo.Azure Blob Storage file
    - Columns
    - Keys
    - Constraints
    - Triggers
    - Indexes
    - Statistics
  - Views
  - Synonyms
  - Programmability
  - External Resources
  - Storage
  - Security
- Security
- Logins

AZURE

SQL SERVER BIG DATA CLUSTERS

1

Welcome x

Would you like to take a quick tour of Azure Data Studio?

Start Tour

# Azure Data Studio

New Open file... Open folder...

Create a connection  
Connect to a database instance through the connection dialog.

Run a query  
Interact with data through a query editor.

Create a notebook  
Build a new notebook using a native notebook editor.

Deploy a server  
Create a new instance of a relational data service on the platform of your choice.

## Resources

History

Name	Location
No recent folders	

Show more

Change SQL language provider: srv-caio-001.database.windows.net : sql-caio-ai1

Select your database and your imported table

CONNECTIONS

...

SERVERS

srv-caio-001.database.windows.net, <default> (aaddmmiinn)

- Databases
  - System Databases
- sql-caio-ai1
  - Tables
    - dbo.Azure Blob Storage file
      - Columns
      - Keys
      - Constraints
      - Triggers
      - Indexes
      - Statistics
    - Views
    - Synonyms
    - Programmability
    - External Resources
    - Storage
    - Security
  - Logins

New Open file...

AZURE  
SQL SERVER BIG DATA CLUSTERS

Change SOL language provider

Welcome x

Would you like to take a quick tour of Azure Data Studio?

# Azure Data Studio

New Open file...

Create a connection  
Connect to a database instance through the connection dialog.

Run a query  
Interact with data through a query editor.

Create a notebook  
Build a new notebook using a native notebook editor.

Deploy a server  
Create a new instance of a relational data service on the platform of your choice.

Resources

History

Name Location

No recent folders  
Show more

Right click and choose the command SELECT TOP 1000

File Edit View Help

SQLQuery\_2 - srv-caio-001.database.windows.net.sql-caio-ai1 (aaddmmiinn) - Azure Data Studio

CONNECTIONS

SERVERS

srv-caio-001.database.windows.net, <default> (aaddmmiinn)

Databases

> System Databases

sql-caio-ai1

Tables

dbo.Azure Blob Storage file

Columns

clean\_text (nvarchar(max), null)  
category (nvarchar(max), null)

Keys  
Constraints  
Triggers  
Indexes  
Statistics  
Views  
Synonyms  
Programmability  
External Resources  
Storage  
Security  
Logins

AZURE

SQL SERVER BIG DATA CLUSTERS

Welcome SQLQuery\_2 - srv-caio-001.database.windows.net.sql-caio-ai1 (aaddmmiinn) X

Run Cancel Disconnect Change Connection sql-caio-ai1 Explain Enable SQLCMD Export as Notebook

```
1 SELECT TOP (1000) [clean_text]
2 , [category]
3 FROM [dbo].[Azure Blob Storage file]
```

Results Messages

	clean_text	category
1	when modi promised “minimum ...	-1
2	talk all the nonsense and co...	0
3	what did just say vote for m...	1
4	asking his supporters prefix...	1
5	answer who among these the m...	1
6	kiya tho refresh maarkefir c...	0
7	surat women perform yagna se...	0
8	this comes from cabinet whic...	0
9	with upcoming election india...	1
1...	gandhi was gay does modi	1
1...	things like demonetisation g...	1
1...	hope tuthukudi people would ...	1
1...	calm waters wheres the modi ...	1
1...	one vote can make all the di...	0
1...	one vote can make all the di...	0
1...	vote such party and leadersh...	-1
1...	vote modi who has not create...	0

Ln 1. Col 1 Spaces: 4 UTF-8 CRLF SOL MSSQL 1.000 rows 00:00:03 srv-caio-001.database.windows.net : sql-caio-ai1

You are connect to your database!

File Edit View Help

• SQLQuery\_2 - srv-caio-001.database.windows.net.sql-caio-ai1 (aaddmmiinn) - Azure Data Studio

CONNECTIONS

SERVERS

- srv-caio-001.database.windows.net, <default> (aaddmmiinn)
  - Databases
    - System Databases
  - sql-caio-ai1
    - Tables
      - dbo.Azure Blob Storage file
        - Columns
          - clean\_text (nvarchar(max), null)
          - category (nvarchar(max), null)
        - Keys
        - Constraints
        - Triggers
        - Indexes
        - Statistics
      - Views
      - Synonyms
      - Programmability
      - External Resources
      - Storage
      - Security
    - Security
      - Logins

Welcome SQLQuery\_2 - srv-caio-ai1 (aaddmmiinn)

Run Cancel Disconnect Change Connection sql-caio-ai1 Explain Enable SQLCMD Export as Notebook

```
1 SELECT TOP (1000) [clean_text]
2 ,[category]
3 FROM [dbo].[Azure Blob Storage file]
4 WHERE [clean_text] LIKE '%high%'
```

Write a SQL statement and add a clause WHERE filtering the field **clean\_text** with the word **high**

Results Messages

	clean_text	category
1	yes good job highly insensit...	1
2	higher voting turnout direct...	1
3	modi came power because five...	1
4	malnutrition india and const...	-1
5	growth story indian economy ...	-1
6	chowkidar hee chor hain baap...	1
7	unemployment rate years high...	1
8	raghuram rajan sent list hig...	-1
9	modi while preparing manifes...	1
1...	would like play the devils a...	1
1...	modi govt controls the price...	1
1...	1st all highest court our la...	1
1...	why vote modi dynasty vision...	-1
1...	today dollar highest ever mo...	1
1...	raghuram rajan sent list hig...	-1
1...	jawans killed the border cri...	-1
1...	the last years the economist...	-1

In 4. Col 2 (31 selected) Spaces: 4 CRLF SOL MSSQL 1.000 rows 00:00:07 srv-caio-001.database.windows.net : sql-caio-ai1

Improve your SQL statement...

## Challenge #2.4

Check your subscription cost

Home - Microsoft Azure X

portal.azure.com/#home

**Left Menu:**

- Create a resource
- Home
- Dashboard
- All services
- FAVORITES**
- All resources
- Resource groups
- App Services
- Function App
- SQL databases
- Azure Cosmos DB
- Virtual machines
- Load balancers
- Storage accounts
- Virtual networks
- Azure Active Directory
- Monitor
- Advisor
- Security Center
- Cost Management + Billing** (highlighted with a red box and orange arrow)
- Help + support

**Header:**

Upgrade  X ? ! ⚙️ ? 😊

**Azure services:**

- Create a resource
- Cost Management ...
- Subscriptions
- All resources
- Data factories
- SQL databases
- Quickstart Center
- Virtual machines
- App Services
- More services

**Recent resources:**

Type	Last Viewed
Subscription	a few seconds ago
SQL database	39 minutes ago
Storage account	4 days ago
Synapse workspace	7 days ago
Resource group	7 days ago
Data factory (V2)	2 weeks ago

**Useful links:**

- View
- Get started
- Documentation
- Training Videos

[https://portal.azure.com/#blade/Microsoft\\_Azure\\_GTM/ModernBillingMenuBlade](https://portal.azure.com/#blade/Microsoft_Azure_GTM/ModernBillingMenuBlade)

Left Menu, select Cost Management + Billing...

A Cost Management + Billing - Mic x df-caio-002 - Azure Data Factory x | +

portal.azure.com/#view/Microsoft\_Azure\_GTM/ModernBillingMenuBlade/~/Overview

## Microsoft Azure

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

Home > Cost Management + Billing | Overview

Default Directory

Search (Ctrl+/)

Overview

Cost Management

Diagnose and solve problems

Billing scopes

**Billing**

Subscriptions

Invoices

Payment methods

Properties

Settings

Support + troubleshooting

New support request

Cost Management + Billing

Help + support

Essentials

Account admin : caiogasparine@gmail.com

Currency : BRL

Notification email : [cailogasparine@gmail.com](mailto:cailogasparine@gmail.com)

Billing country/region : BR

Billing account type : Microsoft Online Services Program

Latest billed amount

R\$162.53

Invoice E0800HIHBH generated on 2/2/2022

Download View invoices

Invoices over time

300  
200  
100  
0

01/02 02/02

Total amount R\$162.53

Shortcuts

Update sold to

Edit preferred communication email

Give feedback

Your subscriptions

Current period cost excludes tax. Available credit would be applied when invoice is generated.

Subscription name	Subscription ID	Status	Last billed amount	Due date	Current Cost
MSDN Platforms Subscri...	5342ccb2-6e74-4d46-ab21-889285fad578	Active	Not available	Not available	0.00

Visualize your subscription costs and forecast.

Free Trial - Microsoft Azure + ↘

portal.azure.com/#@caiogasparinegmail.onmicrosoft.com/resource/subscriptions/bf350b48-f641-49a2-b7e4-598ce72df8ab/overview

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

Home > Cost Management + Billing >

**Free Trial** Subscription

Search (Ctrl+/) Up Upgrade Cancel subscription Rename Change directory Transfer billing ownership Feedback

Your free credit expires in 17 days. Upgrade to keep going with your account.

**Costs by resource** View details >

Resource	Cost
sql-caio-air	9.94 BRL
sql-caio-air	0.91 BRL
df-caio-001	0.28 BRL
Others	0.02 BRL

**Spending rate and forecast** View details >

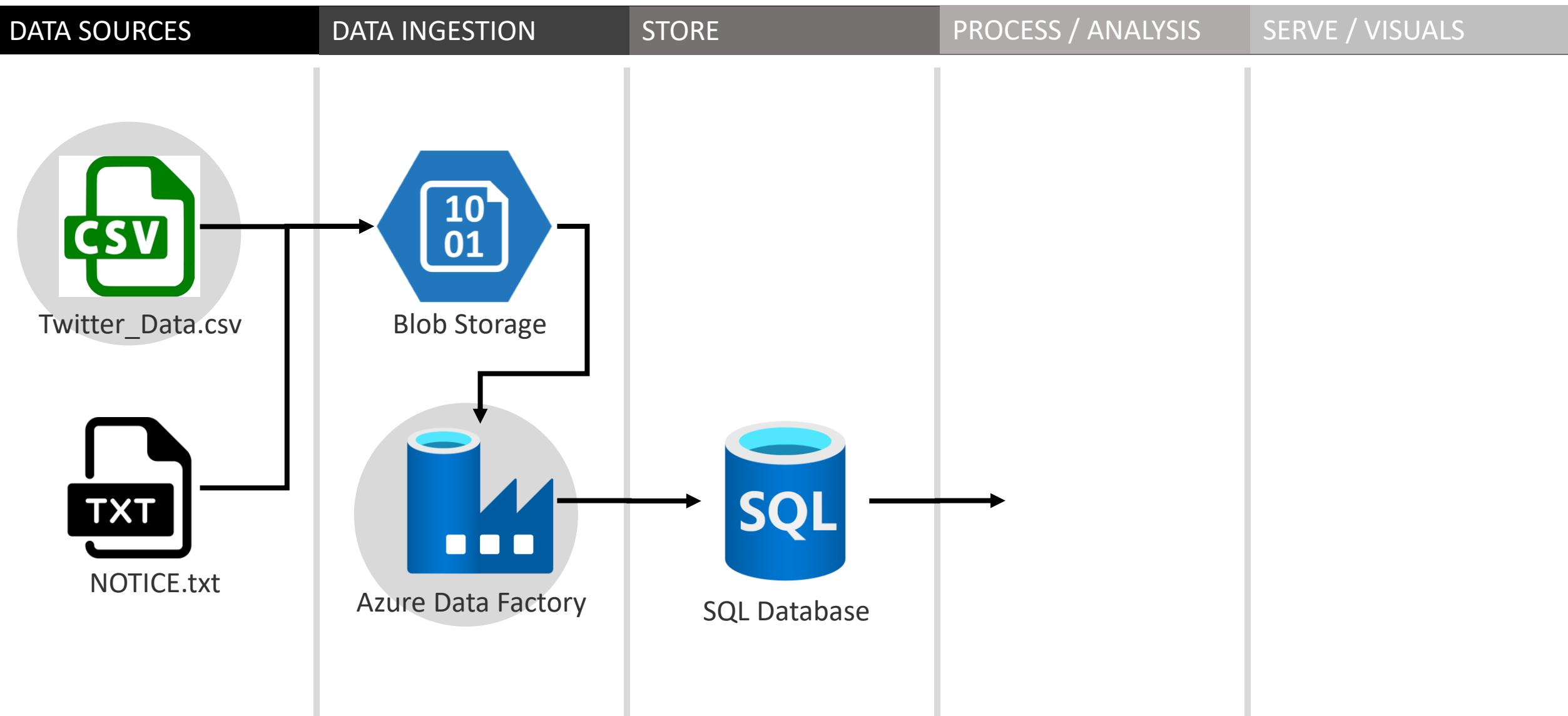
Starting credit: 980 BRL Credit remaining: 968.46 BRL Forecast: 31.53 BRL

**Free services for 12 months** i

Usage period: 5/20/2021 - 6/19/2021

Visualize your subscription costs and forecast.

# Data Architecture so far...



# Thank you! ;-)

Please e-mail me the screenshots of your final steps for each component / service to validate your bonus points.