

Azure Data Factory

Azure console

The image displays the Azure console interface for managing an Azure Data Factory (ADF) instance. The top navigation bar shows the Microsoft Azure logo, an 'Upgrade' button, a search bar, and user information (malipratik297@gmail.com). The main content area is titled 'quersupportADF' and shows the 'Overview' tab. The 'Essentials' section provides key information about the Data Factory (V2) instance, including its status (Succeeded), location (East US), and subscription details. Below this, the 'Azure Data Factory Studio' logo is prominently displayed, along with a 'Launch studio' button. The bottom section of the console features four tiles: 'Quick Starts', 'Tutorials', 'Template Gallery', and 'Training Modules'. The bottom screenshot shows the 'Azure Data Factory Studio' interface, which is used for designing and monitoring data pipelines. The left sidebar lists 'Factory Resources' such as Pipelines, Datasets, Data flows, and Power Query. The main workspace displays a pipeline named 'get_files' with activities including 'Get Metadata' (List s3 files) and a 'ForEach' loop containing 'IterateFiles' and 'Activities' (getLatestFile, compareFile). The bottom of the studio interface includes tabs for 'Parameters', 'Variables', 'Settings', and 'Output'.

Microsoft Azure | Data Factory | quersupportADF

Search resources, services, and docs (G+)

malipratik297@gmail.com

Home > 295b >

quersupportADF Data factory (V2)

Search

Overview

- Activity log
- Access control (IAM)
- Tags
- Diagnose and solve problems
- Settings
- Getting started
- Monitoring
- Automation
- Help

Essentials

Resource group (move) : 295b

Status : Succeeded

Location : East US

Subscription (move) : Azure subscription 1

Subscription ID : a0a9051b-1d4e-425c-b822-4a6776ad3e69

Type : Data factory (V2)

Getting started : [Quick start](#)

Azure Data Factory Studio

[Launch studio](#)

Quick Starts

Tutorials

Template Gallery

Training Modules

Microsoft Azure | Data Factory | quersupportADF

Search factory and documentation

malipratik297@gmail.com

Microsoft recently announced the public preview of Microsoft Fabric, a brand new and exciting way to build cloud-first data analytics. Click [here](#) to get started with Fabric Data Factory!

Preview experience ☐ Off

Factory Resources

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

Activities

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

get_files

Validate Debug Add trigger

Get Metadata

List s3 files

ForEach

IterateFiles

Activities

getLatestFile

compareFile

Parameters Variables Settings Output

+ New

Microsoft Azure | Data Factory | querysupportADF

Search factory and documentation

malipratik297@gmail.com
DEFAULT DIRECTORY

Microsoft recently announced the public preview of Microsoft Fabric, a brand new and exciting way to build cloud-first data analytics. Click [here](#) to get started with Fabric Data Factory!

» Data Factory | Validate all | Publish all | Preview experience | Off

Factory Resources

Filter resources by name

- Pipelines 1
 - get_files
- Change Data Capture (preview) 0
- Datasets 2
- Data flows 0
- Power Query 0

Activities

Search activities

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

get_files > iterateFiles

Get Metadata
get_latest_file

If Condition
compare_time

True

(X) Get latest modified... → (X) Get latest file name → run_dtb_ntb → +

False

+


Edit linked service

 Azure Databricks [Learn more](#) 

Name *

AzureDatabricks1

Description

Connect via integration runtime * 

AutoResolveIntegrationRuntime

Account selection method *

☐ From Azure subscription ☒ Enter manually

Databrick Workspace URL * 

https://adb-1988582850967560.0.azuredatabricks.net

Authentication type *

Access Token

Access token

Azure Key Vault

Access token * 

.....

Select cluster

Save

Cancel

 Test connection


Edit linked service

 Amazon S3 [Learn more](#) 

Name *

AmazonS31

Description

Connect via integration runtime * 

AutoResolveIntegrationRuntime

Authentication type

Access key

Access key ID

Azure Key Vault

Access key ID *

AKIAX5INISKDOPAPOLPE

Secret access key

Azure Key Vault

Secret access key *

.....

Service URL 

https://s3.amazonaws.com

Save

Cancel



Test connection

Azure Databricks

This screenshot shows the Microsoft Azure portal interface for an Azure Databricks service instance. The top navigation bar includes the Microsoft Azure logo, an 'Upgrade' button, a search bar, and user information for 'malipratik297@gmail.com'. The breadcrumb trail indicates the path: Home > Azure Databricks > 295b. The main content area is divided into a left-hand navigation pane and a central overview section. The navigation pane lists various management options: Overview (selected), Activity log, Access control (IAM), Tags, Diagnose and solve problems, Settings (including Virtual Network, Peerings, Encryption, Networking, Security & compliance, Properties, Locks), Monitoring (Diagnostic settings), and Automation (CLI / PS). The central overview section, titled 'Essentials', displays key information about the instance: Status (Active), Managed Resource Group (databricks-rg-295b-rezjash5mq), Resource group (295b), Location (West US 2), Subscription (Azure subscription 1), and Subscription ID (a0a9051b-1d4e-425c-b822-4a6776ad3e69). It also provides links for Tags (edit) and Add tags. A large red Databricks logo is centered below this information, with a 'Launch Workspace' button underneath. At the bottom, there are four quick-action buttons: Documentation, Getting Started, Import Data from File, and Import Data from Azure Storage.

This screenshot displays the Databricks workspace configuration page for a cluster named 'Pratik Mali's Cluster'. The top navigation bar shows the Microsoft Azure logo, the Databricks logo, a search bar, and user information. The left-hand navigation pane lists various workspace components: New, Workspace, Recents, Catalog, Workflows, Compute (selected), SQL (including SQL Editor, Queries, Dashboards, Alerts, Query History, and SQL Warehouses), Data Engineering (including Job Runs, Data Ingestion, and Delta Live Tables), and a 'More' button. The main content area is titled 'Pratik Mali's Cluster' and includes a 'More' button, a 'Start' button, and an 'Edit' button. The 'Configuration' tab is active, showing settings for Policy (Personal Compute), Access mode (Single user access), and Access user (Pratik Mali). The 'Performance' section shows the Databricks Runtime Version (14.3 LTS ML) and the Node type (Standard_DS3_v2, 14 GB Memory, 4 Cores). A 'Summary' box on the right provides a high-level overview of the cluster's configuration: 1 Driver, 14 GB Memory, 4 Cores, Runtime 14.3.x-cpu-ml-scala2.12, and a performance metric of 0.75 DBU/h. The 'Terminate after' setting is set to 60 minutes of inactivity.

Open in new tab

Jobs	Stages	Storage	Environment	Executors	SQL / DataFrame
Spark Jobs (?)					
User: root					
Total Uptime:					
Scheduling Mode: FAIR					
Completed Jobs: 15					
Event Timeline					
Completed Jobs (15)					
Page: 11 Pages. Jump to 1. Show 100 items in a page. Go					
Job Id (Job Group)	Description	Submitted	Duration	Stages: Succeeded/Total	
14 (7555624075288637712_9222410458427602438_015410834923416381f08da2e8828d22)	df = spark.read.format("csv").opt... wrapper at /root/.ipykernel/6187/command-3077646539820250-3882870502:6	2024/05/10 06:26:56	2 s	1/1	
13 (7555624075288637712_9222410458427602438_015410834923416381f08da2e8828d22)	df = spark.read.format("csv").opt... load at NativeMethodAccessorImpl.java:0	2024/05/10 06:26:55	0.5 s	1/1	
12	df = spark.read.format("csv").opt...	2024/05/10	0.2 s	1/1	