

DXC REALTIME PROJECTS

AZ-900, DP - 203



JUNE 10, 2022
[DXC TECHNOLOGY]

Name: Bhogadi Naga Iswarya Lakshmi

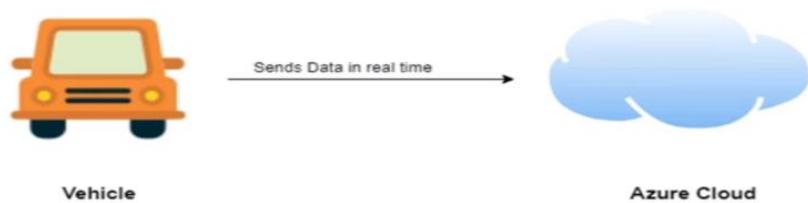
Reg No: DXC262AB12005

Project1 Name: Smart Vehicles

Date:10-06-2022

Project 1 : Connected Vehicles

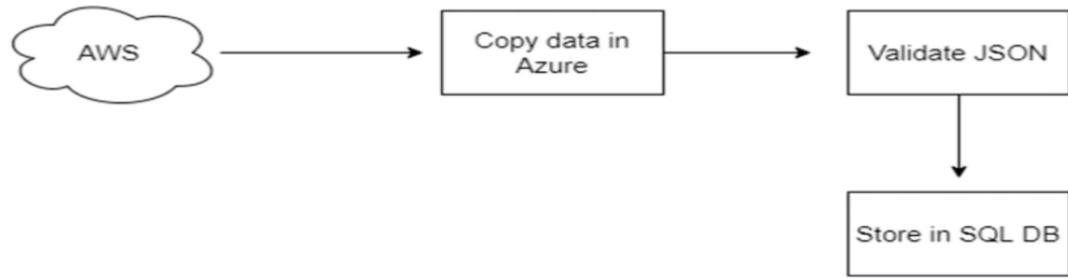
- General Motors is one of the leading heavy vehicle manufacture company. To improve their service they are planning to rollout lot new features based on IoT.



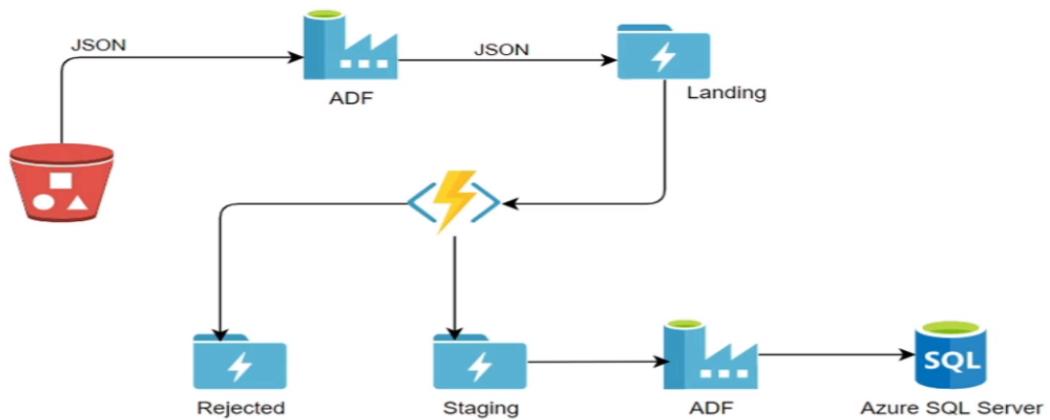
Project 1 : Connected Vehicles

- Vehicle has third party IoT device which will send the telemetry data (in JSON format) over the AWS cloud.
- You need to move data from third party AWS to General Motors Azure cloud.
- You need to validate the JSON sometime it could be incomplete or wrong JSON which need to be rejected.
- Once JSON got validated this data would be stored in the SQL database which will be further utilized by data science team.

Project 1 : Connected Vehicles



Project 1 : Connected Vehicles



Architecture Diagram for Connected Vehicle Project

Practical Lab: Create **Azure Data Factory** Account For Data pipelines

Azure Data Factory is a cloud-based data integration service that orchestrates and automates the movement and transformation of data.

Step-1: First open your Microsoft Azure account and search for Data Factories.

The screenshot shows the Microsoft Azure portal interface. The search bar at the top contains the query "data factories". Below the search bar, there are several tabs: All, Services (32), Documentation (27), Resources (0), Resource Groups (0), and Marketplace (0). The "Services" tab is selected. Under the "Services" heading, "Data factories" is highlighted with a blue selection bar. Other service options listed include Data Catalog, Data Shares, Azure Databricks, and Managed databases. To the right of the main search area, there is a sidebar titled "See all services" with a large arrow icon pointing right.

Step 2: Click on create to create a new data factory account.

The screenshot shows the "Data factories" blade in the Microsoft Azure portal. The top navigation bar includes the Microsoft Azure logo, a search bar, and user information (dx262ab1205_1654530... MANIPAL PRO LEARN (MANIPAL...)). Below the navigation, the breadcrumb path is "Home > Data factories". The main content area displays a table header with columns: Name, Type, Subscription, Resource group, and Location. A message "No data factories to display" is centered, along with a note "Try changing or clearing your filters." and a "Create data factory" button.

Step 3: Complete the Project and Instance details.

The screenshot shows the "Create Data Factory" blade in the Microsoft Azure portal. The top navigation bar includes the Microsoft Azure logo, a search bar, and user information (dx262ab1205_1654530... MANIPAL PRO LEARN (MANIPAL...)). The breadcrumb path is "Home > Data factories > Create Data Factory". The main form is divided into two sections: "Project details" and "Instance details".

Project details:

- Subscription: Azure-DXC262AB12Lab
- Resource group: dxcrg2317 (with a "Create new" option)

Instance details:

- Name: vehiclesdf123
- Region: East US
- Version: V2 (Recommended)

At the bottom of the form are three buttons: "Review + create", "< Previous", and "Next : Git configuration >".

Step 4: After completing those details continue to Git configuration.

The screenshot shows the 'Create Data Factory' wizard on the 'Git configuration' tab. At the top, there's a navigation bar with 'Microsoft Azure' and a search bar. Below it, the breadcrumb path is 'Home > Data factories > Create Data Factory'. The main content area has tabs for 'Basics', 'Git configuration' (which is selected), 'Networking', 'Advanced', 'Tags', and 'Review + create'. A note says '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.' Below this is a link 'Learn more about Git integration in Azure Data Factory'. A checkbox 'Configure Git later' is checked. At the bottom, there are buttons for 'Review + create', '< Previous', and 'Next : Networking >'.

Step 5: After step-4, continue the completion of details of Networking ,Advanced ,Tags and finally goes to Review and Create and also completes validation process.

The screenshot shows the 'Create Data Factory' wizard on the 'Review + create' tab. At the top, there's a navigation bar with 'Microsoft Azure' and a search bar. Below it, the breadcrumb path is 'Home > Data factories > Create Data Factory'. The main content area shows a green banner with a checkmark and the text 'Validation Passed'. Below this, the tabs are 'Basics', 'Git configuration', 'Networking', 'Advanced', 'Tags', and 'Review + create' (which is selected). A 'TERMS' section contains legal text about agreeing to terms and privacy statements. The 'Basics' section shows 'Subscription' as 'Azure-DXC262AB12Lab' and 'Resource group' as 'dxcrg2317'. At the bottom, there are buttons for 'Create', '< Previous', 'Next', and 'Download a template for automation'.

Step 6: After completing validation process, Click on create and Deployment process starts.

The screenshot shows the Microsoft Azure Data Factory Deployment Overview page. At the top, there's a search bar and a navigation bar with icons for Home, Search, Delete, Cancel, Redeploy, Refresh, and Help. The main title is "Microsoft.DataFactory-20220610162942 | Overview". On the left, a sidebar has tabs for Overview, Inputs, Outputs, and Template, with "Overview" selected. A feedback link "We'd love your feedback!" is present. The main content area displays deployment details: "Deployment is in progress", "Deployment name: Microsoft.DataFactory-20220610162942", "Subscription: Azure-DXC262AB12Lab", "Resource group: dxrcg2317", "Start time: 6/10/2022, 4:35:37 PM", and "Correlation ID: d9b03e83-0cbd-4fb8-85c8-35fb43ad8852". Below this, a table titled "Deployment details (Download)" shows "No results." under columns for Resource, Type, Status, and Operation details.

Step 7: After completing deployment process, click on goto resource.

The screenshot shows the Microsoft Azure Data Factory Deployment Overview page after deployment completion. The main title is "Microsoft.DataFactory-20220610162942 | Overview". The deployment status is now "Your deployment is complete". The deployment details remain the same: "Deployment name: Microsoft.DataFactory-20220610162942", "Subscription: Azure-DXC262AB12Lab", "Resource group: dxrcg2317", "Start time: 6/10/2022, 4:35:37 PM", and "Correlation ID: d9b03e83-0cbd-4fb8-85c8-35fb43ad8852". Below the deployment details, there are sections for "Deployment details (Download)", "Next steps", and a prominent blue "Go to resource" button. To the right, there are two side panels: "Cost Management" (with a green dollar sign icon) and "Microsoft Defender for Cloud" (with a shield icon).

Step 8: Then ,click on Open Azure Data Factory Studio and start authoring & monitoring of your data pipelines.

Microsoft Azure | vehiclesdf123

Search resources, services, and docs (G+)

Home > Microsoft.DataFactory-20220610162942 >

vehiclesdf123 Data factory (V2)

Search (Ctrl+F)

Delete

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Settings

Networking

Managed identities

Properties

Locks

Getting started

Quick start

Resource group (move)
dxcrg2317

Status
Succeeded

Location
East US

Subscription (move)
Azure-DXC262AB12Lab

Subscription ID
46b25869-74ec-4ab6-a07d-dc540b41b4f8

Type
Data factory (V2)

Getting started

Open Azure Data Factory Studio

Read documentation Learn how to be productive

Microsoft Azure | vehiclesdf123

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

vehiclesdf123

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.

Practical Lab: Create ADF Pipeline End to end pipeline with triggers enabled

Practical Lab: Create Azure blob trigger logic

Blob source(Input)->ADF->Blob Destination(Output).

Step-1: First open your Microsoft Azure account and search for Storage Accounts.

The screenshot shows the Microsoft Azure search interface. The search bar at the top contains the query "storage accounts". Below the search bar, there are several tabs: All, Services (19), Marketplace (1), Documentation (28), Resources (0), and Resource Groups (0). The "All" tab is selected. Under the "Services" section, "Storage accounts" is highlighted. Other services listed include Azure Active Directory (0), Genomics accounts, Storage accounts (classic), Integration accounts, Automation Accounts, Lab accounts, Batch accounts, Storage Explorer, and Marketplace. There are also sections for Recent, Favorties, and Marketplace. At the bottom, there are links for "Storage account overview - Azure Storage | Microsoft Docs" and "Introduction to Azure Storage - Cloud storage on Azure | Microsoft...". A "See all" link is visible on the right side.

Step 2: Click on create to create a new storage account.

The screenshot shows the Microsoft Azure Storage accounts list page. The top navigation bar includes the Microsoft Azure logo, a search bar, and user information (dx262ab1205_1654530... MANIPAL PRO LEARN (MANIPAL...)). Below the navigation is a breadcrumb trail: Home > Storage accounts. The main area displays a table with columns: Name, Type, Kind, Resource group, Location, and Subscription. A "Create" button is located at the top left of the table. Filter options include "Subscription == all", "Resource group == all", "Location == all", and "Add filter". Below the table, a message states "No storage accounts to display". A descriptive text block explains how to create a storage account, mentioning Blob storage account and access tiers. A "Give feedback" link is at the bottom right.

Step 3: Complete the Project and Instance details.

The screenshot shows the Microsoft Azure "Create a storage account" wizard, specifically the "Basics" step. The top navigation bar includes the Microsoft Azure logo, a search bar, and user information (dx262ab1205_1654530... MANIPAL PRO LEARN (MANIPAL...)). Below the navigation is a breadcrumb trail: Home > Storage accounts > Create a storage account. The main area has tabs for Basics, Advanced, Networking, Data protection, Encryption, Tags, and Review + create. The Basics tab is selected. A descriptive text block explains that Azure Storage is a Microsoft-managed service providing cloud storage. It mentions that the cost of your storage account depends on usage and options chosen. A "Project details" section asks for the subscription to create the new storage account. A dropdown menu shows "Azure-DXC262AB12Lab". Navigation buttons at the bottom include "Review + create" (highlighted in blue), "< Previous", and "Next : Advanced >".

Step 4: After completing those details continue to Advanced.

The screenshot shows the 'Create a storage account' wizard on the 'Advanced' tab. At the top, there's a note: 'Certain options have been disabled by default due to the combination of storage account performance, redundancy, and region.' Below this, the 'Security' section is visible, containing three checked checkboxes: 'Require secure transfer for REST API operations', 'Enable blob public access', and 'Enable storage account key access'. At the bottom of the screen are navigation buttons: 'Review + create' (highlighted in blue), '< Previous', and 'Next : Networking >'.

Step 5: After step-4, continue the completion of details of Networking, Data protection, Encryption, Tags and finally goes to Review and Create and also completes validation process.

The screenshot shows the 'Create a storage account' wizard on the 'Review + create' tab. A green banner at the top indicates 'Validation passed'. Below it, the 'Basics' section displays the following configuration details:

Subscription	Azure-DXC262AB12Lab
Resource Group	dxcrg2317
Location	eastus
Storage account name	vehicles231
Deployment model	Resource manager
Performance	Standard
Replication	Read-access geo-redundant storage (RA-GRS)

At the bottom, there are buttons for 'Create' (highlighted in blue), '< Previous', 'Next >', and 'Download a template for automation'.

Step 6: After completing validation process, Click on create and Deployment process starts.

The screenshot shows the Microsoft Azure Deployment Overview page for a resource named 'vehicles231_1654860223171'. The status bar indicates 'Deployment is in progress'. Deployment details include a name, start time (6/10/2022, 4:53:53 PM), subscription, correlation ID, and resource group. A table shows the resource type as Microsoft.Storage/storageAccounts and status as Accepted. A sidebar on the right provides links to Microsoft Defender for Cloud, free tutorials, and work with experts.

Step 7: After completing deployment process, click on goto resource.

The screenshot shows the Microsoft Azure Deployment Overview page for the same resource. The status bar now indicates 'Your deployment is complete'. Deployment details remain the same. A 'Next steps' section includes a 'Go to resource' button. A sidebar on the right provides links to Cost Management and Microsoft Defender for Cloud.

Step 8: Click on Goto resource and click containers

The screenshot shows the Microsoft Azure Storage account containers page for 'vehicles231'. It lists a single container named '\$logs' with details: Last modified (6/12/2022, 9:42:35 AM), Public access level (Private), and Lease state (Available). The sidebar on the left shows other storage account options like Activity log, Tags, and Data migration.

Click on +container to create a new container to store the data.

The screenshot shows the Microsoft Azure Storage account interface for the 'vehicles231' account. On the left, a sidebar menu includes 'Containers' under 'Data storage'. In the main area, a table lists existing containers: '\$logs' (Last modified: 6/12/2022, 9:42:35 AM, Public access level: Private). A modal window titled 'New container' is open on the right, prompting for a 'Name' (marked with a red asterisk) and setting the 'Public access level' to 'Private (no anonymous access)'. At the bottom of the modal are 'Create' and 'Discard' buttons.

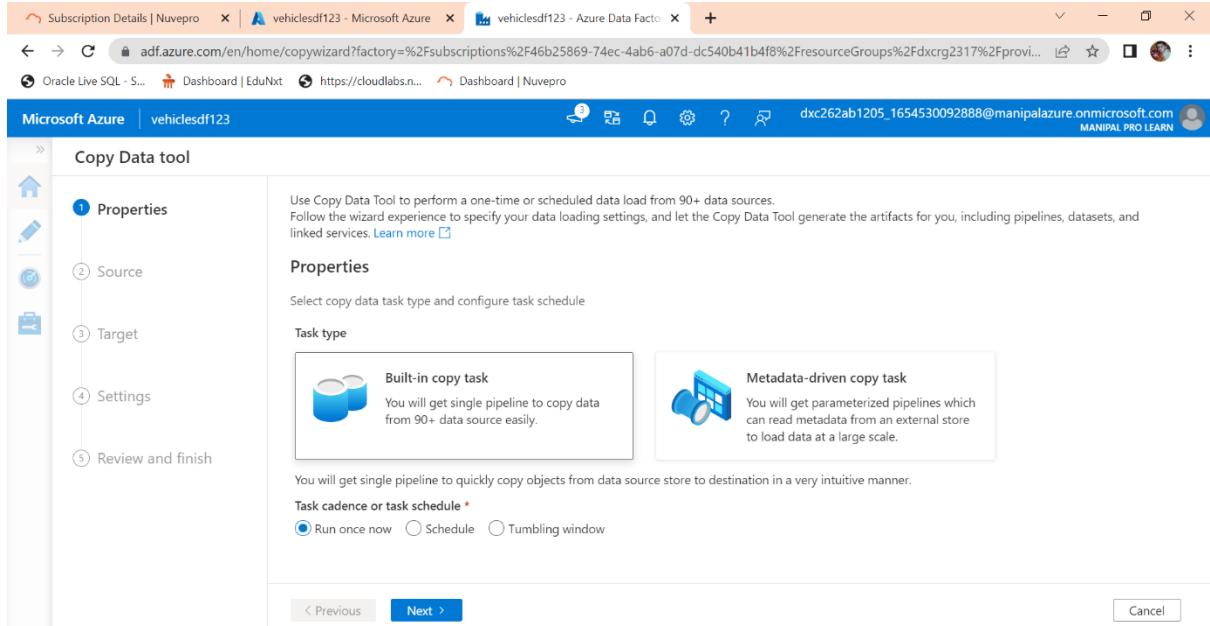
Create a new source container.

The screenshot shows the same Microsoft Azure Storage account interface. The 'Containers' section now includes a new entry: 'source' (Last modified: 6/12/2022, 9:55:18 AM, Public access level: Private, Lease state: Available). The 'Show deleted containers' toggle switch is visible at the top of the table.

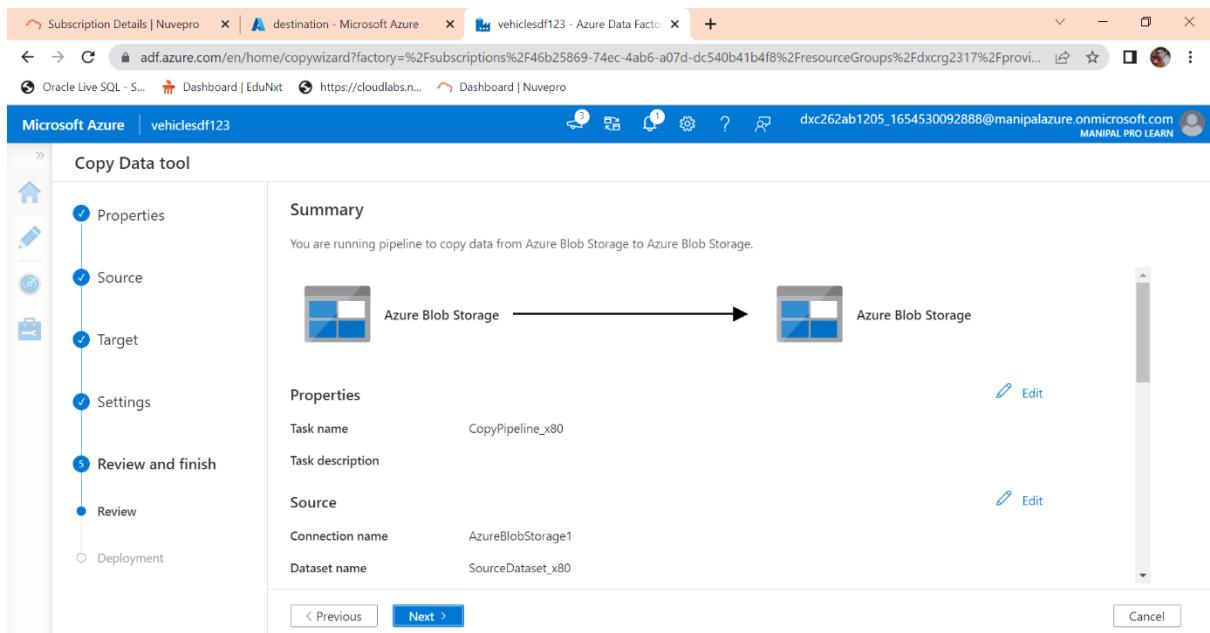
Upload data to source container

The screenshot shows the 'source' container details page. The left sidebar shows 'Overview' selected. The main area displays a table of blobs: 'guests.csv' (Modified: 6/12/2022, 10:12:48 ..., Access tier: Hot (Inferred), Archive status: Not yet archived, Blob type: Block blob, Size: 83 B). Navigation links at the bottom include 'Upload', 'Change access level', 'Refresh', 'Delete', 'Change tier', 'Acquire lease', 'Break lease', 'View snapshots', and 'Add filter'.

Similarly create a destination container ,Open datafactory account and start using copy tool.



Complete the details of properties ,source, target ,settings and come to review+finish.



After completing the copy process, guests.csv file present in destination from source container.

Subscription Details | Nuvepro x destination - Microsoft Azure x vehiclesdf123 - Azure Data Facto x +

← → 🔍 portal.azure.com/#view/Microsoft_Azure_Storage/ContainerMenuBlade/~/overview/storageAccountName%2Fsubscriptions%2F46b25869-74ec-4ab6-a07d-d... 🔍 ☆ 🌐

Oracle Live SQL - ... 📈 Dashboard | EduNxt 🌐 https://cloudlabs.n... 🌐 Dashboard | Nuvepro

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

Home > vehicles231 > destination

destination Container

Search (Ctrl+ /) Upload Change access level Refresh Delete Change tier Acquire lease Break lease View snapshots ...

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

Authentication method: Access key (Switch to Azure AD User Account)
Location: destination

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

Add filter

Name	Modified	Access tier	Archive status	Blob type	Size
guests.txt	6/12/2022, 10:47:11 ...	Hot (Inferred)		Block blob	71 B

At final stage the pipelines are trigger enabled.

All pipeline runs > CopyPipeline_x80

Pipeline runs

Run group ID: be84fbef-626d-4519-b343-5769a9d62427

Refresh Edit columns

Showing 1 - 2 of 2 items

In start	Run end	Duration	Triggered by	Triggered by type	Status	Error
Jun 12, 2022, 10:48:38 am	Jun 12, 2022, 10:48:47 am	00:00:09	Manual trigger	Manual	Succeeded	
Jun 12, 2022, 10:47:01 am	Jun 12, 2022, 10:47:13 am	00:00:12	Manual trigger	Manual	Succeeded	

Practical Lab: Create Azure SQL Server and Database

Step 1: Open Azure account and search for sql databases

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

All Services Marketplace Documentation Resources Resource Groups

Azure Active Directory (0)

Services

- SQL databases
- SQL servers
- SQL elastic pools
- SQL managed instances

Marketplace

- Zumero - Sync SQL Server data with offline SQLite
- SQLstream 6.0.0.1
- Azure Synapse Analytics
- Continue searching in Azure Active Directory

See all

Give feedback

Subscriptions Resource groups All resources Dashboard

https://portal.azure.com/#blade/HubsExtension/BrowseResourceBlade/resourceType/Microsoft.Sql%2fserver%2fdatabases

Step 2: Create a new database by clicking on create and complete the configurations.

The screenshot shows the 'Create SQL Database' wizard in the Microsoft Azure portal. The title bar says 'Microsoft Azure' and 'Create SQL Database'. The top navigation bar includes 'Search resources, services, and docs (G+)', 'Home > SQL databases >', and user information 'dxc262ab1205_1654530... MANIPAL PRO LEARN (MANIPAL...)'. The main content area is titled 'Create SQL Database ...' and has tabs for 'Basics', 'Networking', 'Security', 'Additional settings', 'Tags', and 'Review + create'. A note at the top says '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.' Below this is a 'Did you know?' box: 'Did you know that new users in Azure can create a free Azure SQL Database and use it for 12 months using Azure free account? [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 * (dropdown) Azure-DXC26AB12Lab
Resource group * (dropdown) Select a resource group

Review + create **Next : Networking >**

Step 3: now create a new database server and fil the details

The screenshot shows the 'Create SQL Database Server' wizard in the Microsoft Azure portal. The title bar says 'Microsoft Azure' and 'Create SQL Database Server'. The top navigation bar includes 'Search resources, services, and docs (G+)', 'Home > SQL databases > Create SQL Database >', and user information 'dxc262ab1205_1654530... MANIPAL PRO LEARN (MANIPAL...)'. The main content area is titled 'Create SQL Database Server ...' and has tabs for 'Server details', 'Authentication', and 'Advanced settings'. Under 'Server details', there are fields for 'Server name *' (input field 'Enter server name .database.windows.net') and 'Location *' (dropdown '(US) East US'). Under 'Authentication', there is a note: '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 Azure AD authentication [Learn more](#) using an existing Azure AD user, group, or application as Azure AD admin [Learn more](#), or select both SQL and Azure AD authentication.' Below this is a section for 'Authentication method' with a radio button for 'Use SQL authentication' (selected) and an 'OK' button.

Step 4:It validates and create a database

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

Home > SQL databases > Create SQL Database

Networking

Allow Azure services and resources to access this server	No
Private endpoint	None
Minimum TLS version	1.2
Connection Policy	Default

Security

Identity	Not enabled
Service principal (preview)	Off
Transparent data encryption	Service-managed key selected
Advanced data security	Not now
Sql Ledger(Database)	Disabled
Digest Storage	Disabled

Buttons: Create, < Previous, Download a template for automation

Step 5: Deployment starts and completes.

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

Home > Microsoft.SQLDatabase.newDatabaseNewServer_6f8aa8d03b7049039adf1 | Overview

Deployment

Search (Ctrl+)/

Delete Cancel Redeploy Refresh

We'd love your feedback! →

Your deployment is complete

Deployment name: Microsoft.SQLDatabase.newDatabaseNewServer_6f8aa8d03b7049039adf1 Start time: 6/12/2022, 11:22:54 AM

Subscription: Azure-DXC262AB12Lab Correlation ID: a0c8e7cb-3ed4-40b5-9225-415e67c7d36d

Resource group: dxcrg2317

Deployment details (Download) Next steps Go to resource

Cost Management Get notified to stay within budget prevent unexpected charges Set up cost alerts >

Step 6: new sql231 server is created

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

Home > Microsoft.SQLDatabase.newDatabaseNewServer_6f8aa8d03b7049039adf1 > sql231 (server232/sql231)

SQL database

Search (Ctrl+)/

Copy Restore Export Set server firewall Delete Connect with... Feedback

This database was just created. Do you need any help [getting started?](#)

Overview

Activity log Tags Diagnose and solve problems Getting started Query editor (preview)

Power Platform

Power BI Power Apps Power Automate

Settings

Compute + storage

Essentials

Resource group (move) dxcrg2317	Server name server232.database.windows.net
Status Online	Connection strings Show database connection strings
Location East US	Pricing tier General Purpose - Serverless: Gen5_1 vCore
Subscription (move) Azure-DXC262AB12Lab	Auto-pause delay 1 hour
Subscription ID 46b25869-74ec-4ab6-a07d-dc540b41b4f8	Earliest restore point No restore point available

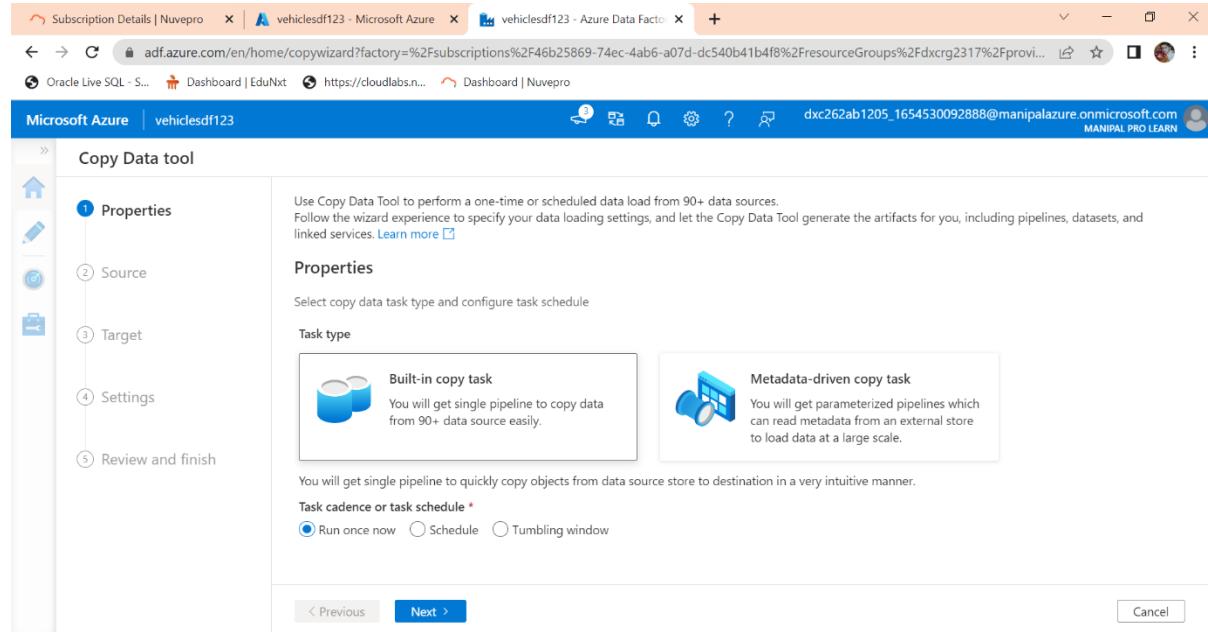
Tags ([edit](#)) Click here to add tags

Show data for last: 1 hour 24 hours 7 days

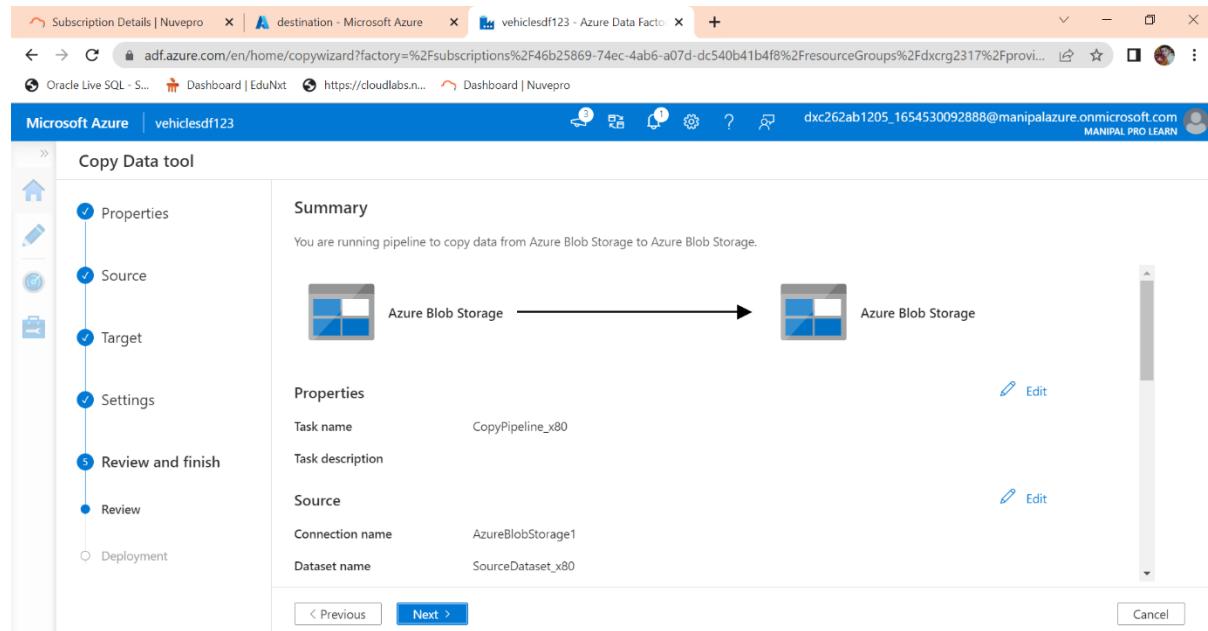
Aggregation type: Max

Practical Lab: Add another pipelines for moving data from Staging to SQL DB

Open datafactory account and start using copy tool.



Complete the details of properties ,source, target ,settings and come to review+finish.



After completing the copy process, guests.csv file present in destination from source container.

The screenshot displays two windows from the Microsoft Azure portal. The top window shows the 'destination' container in the 'vehiclesdf123' storage account. It lists a single blob named 'guests.txt' with a size of 71 B. The bottom window shows the 'Pipeline runs' page for the 'CopyPipeline_x80' run group, displaying two successful runs on June 12, 2022.

Run start	Run end	Duration	Triggered by	Triggered by type	Status	Error
Jun 12, 2022, 10:48:38 am	Jun 12, 2022, 10:48:47 am	00:00:09	Manual trigger	Manual	Succeeded	
Jun 12, 2022, 10:47:01 am	Jun 12, 2022, 10:47:13 am	00:00:12	Manual trigger	Manual	Succeeded	

Result: In this project a new pipeline is created and transferred data from source to destination by using data tool by using data Azure Data Factory.

Conclusion: Data is successfully validated and stored using Microsoft Azure.

Name:Bhogadi Naga Iswarya Lakshmi

Reg No: DXC262AB12005

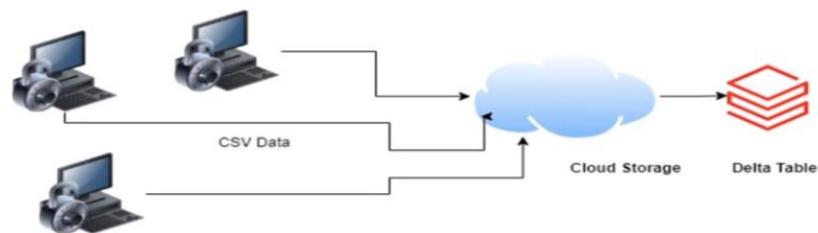
Project2 Name: AP Morgan Data Platform

Date: 10-06-2022

Project 2: AP Morgan Data Platform

Project 2 : AP Morgan

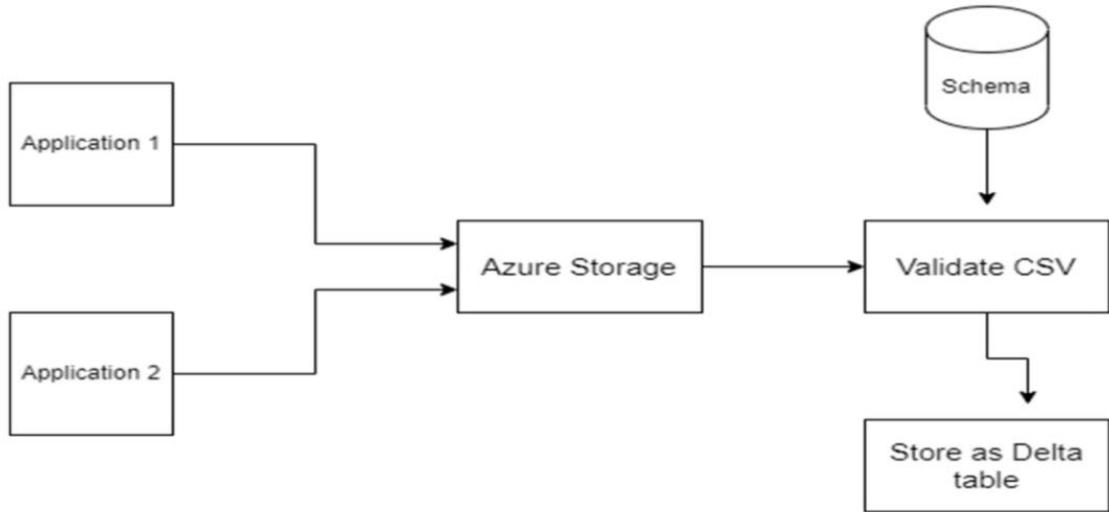
- Multiple Internal applications sends the data(huge size) in CSV format on daily basis in the cloud storage location. There are couple of Data/schema validation needed to be performed on this incoming data. Once everything is passed data to be persisted as Delta table in Databricks for downstream system.



Project 2 : AP Morgan- High Level Detail

- Internal Application sends CSV file in Azure data lake storage.
- Validation needed to apply on this follows:
 - Check for duplicate rows. If it contains duplicate rows, file need to be rejected.
 - Need to validate the date format for all the date fields. Date column names and desired date format is stored in a Azure SQL server. If validation fails file will be rejected.
- Move all the rejected files to Reject folder.
- Move all the passed files to Staging folder.
- Write the passed files as the Delta table in the Azure Databricks

Project 2 : AP Morgan



Practical Lab: Create a Databricks

Azure Databricks is a Scalable analytics in Azure based on Apache Spark Workflows and workspace for data users Native integration with other Azure services.

Step 1: First open your Microsoft Azure account and search for Data Bricks.

The screenshot shows the Microsoft Azure portal interface. The search bar at the top contains the text "data bricks". Below the search bar, there is a navigation bar with tabs: All, Services (33), Marketplace (9), Documentation (28), Resources (0), and Resource Groups (0). The "All" tab is selected. Under the "Services" section, "Azure Databricks" is highlighted with a blue background. Other listed services include Datadog, Data Catalog, Data factories, Data Shares, Unravel for Azure Databricks, Trifactor for Azure, and Lakehouse Monitor. On the left side, there is a sidebar with sections for "Create a resource", "Recent", "Resources", and "Marketplace". The "Recent" section lists "vehiclesdf12", "vehicles231", and "dxcrg2317". The "Resources" section has a "See all" link. The "Marketplace" section also has a "See all" link. At the bottom of the page, there is a "Give feedback" button.

Step 2: Click on create to create a new Databricks account.

Microsoft Azure Search resources, services, and docs (G+/-) dx262ab1205_1654530... MANIPAL PRO LEARN (MANIPAL)

Home > Azure Databricks ...

Manipal Pro Learn (manipalazure.onmicrosoft.com)

+ Create Manage view Refresh Export to CSV Open query Assign tags

Filter for any field... Subscription == all Resource group == all Location == all Add filter

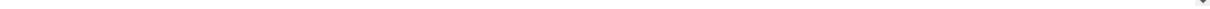
No grouping List view

Name ↑ Type ↑ Resource group ↑ Location ↑ Subscription ↑

No azure databricks services to display

Unlock insights from all your data and build artificial intelligence (AI) solutions with Azure Databricks, set up your Apache Spark environment in minutes, autoscale, and collaborate on shared projects in an interactive workspace.

Give feedback



Step 3: Complete the Project and Instance details.

Microsoft Azure Search resources, services, and docs (G+/-) dx262ab1205_1654530... MANIPAL PRO LEARN (MANIPAL)

Home > Azure Databricks > Create an Azure Databricks workspace ...

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

Subscription * Azure-DXC262AB12Lab

Resource group * dxcrg2317 Create new

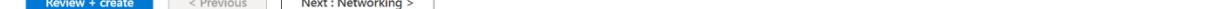
Instance Details

Workspace name * APMorgan23

Region * East US

Pricing Tier * Standard (Apache Spark, Secure with Azure AD)

Review + create < Previous Next : Networking >



Step 4: After step-3, continue the completion of details of Networking, Advanced ,Tags and finally goes to Review and Create and also completes validation process.

Microsoft Azure Search resources, services, and docs (G+/-) dx262ab1205_1654530... MANIPAL PRO LEARN (MANIPAL)

Home > Azure Databricks > Create an Azure Databricks workspace ...

Validation Succeeded

Basics Networking Advanced Tags Review + create

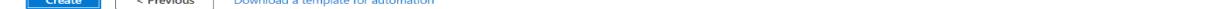
Summary

Basics

Workspace name	APMorgan23
Subscription	Azure-DXC262AB12Lab
Resource group	dxcrg2317
Region	East US
Pricing Tier	standard

Networking

Create < Previous Download a template for automation



Step 5: After completing validation process, Click on create and Deployment process starts.

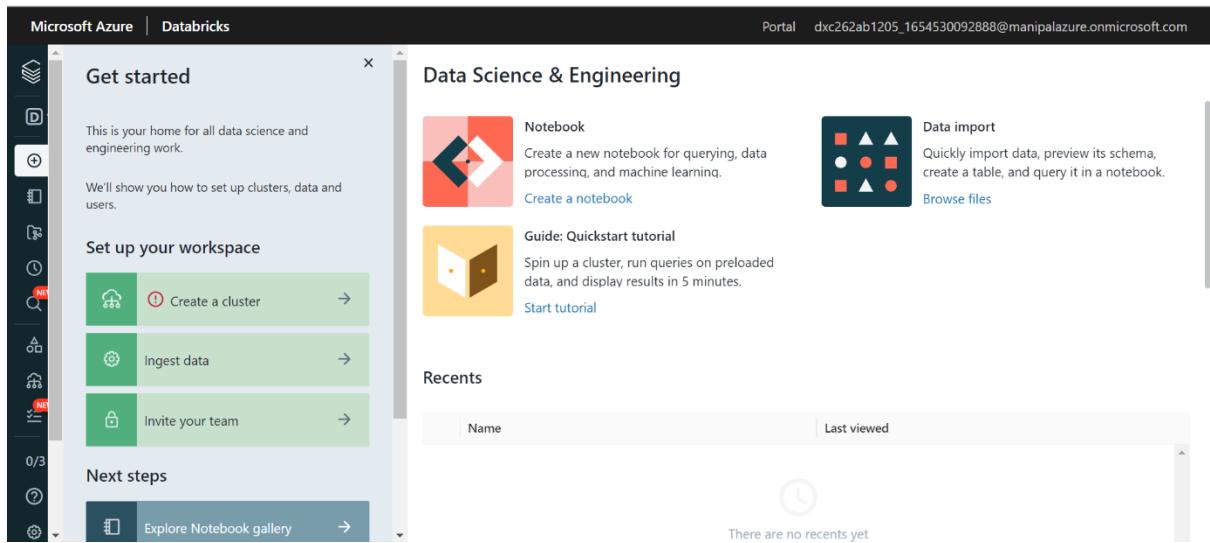
The screenshot shows the Microsoft Azure Overview page for a resource group named 'dxcrg2317_APMargin23'. The main header bar includes the Microsoft Azure logo, a search bar, and various navigation icons. Below the header, the title 'dxcrg2317_APMargin23 | Overview' is displayed, along with a 'Deployment' icon. A prominent message box in the top right corner states 'Deployment is in progress...' with the subtext 'Deployment to resource group 'dxcrg2317' is in progress.' Below this, there are tabs for 'Overview', 'Inputs', 'Outputs', and 'Template'. The 'Overview' tab is selected. A large central area displays deployment details, including the deployment name 'dxcrg2317_APMargin23', subscription information 'Azure-DXC262AB12Lab', and resource group 'dxcrg2317'. It also shows the start time '6/10/2022, 5:46:54 PM' and correlation ID 'e8f2e6e8-37e5-4d3c-b065-c5c520b45188'. A section titled 'Deployment details (Download)' is present, showing a table with columns 'Resource', 'Type', 'Status', and 'Operation details'. The table body contains the message 'No results.'

Step 6: After completing deployment process, click on goto resource.

The screenshot shows the Microsoft Azure Overview page for the same resource group 'dxcrg2317_APMargin23'. The main header bar and title are identical to the previous screenshot. The message box now indicates 'Your deployment is complete'. The deployment details remain the same. A new section titled 'Next steps' is visible, containing a blue 'Go to resource' button. To the right of the main content area, there are two promotional cards: 'Cost Management' (with a green dollar sign icon) and 'Microsoft Defender for Cloud' (with a blue shield icon). Both cards include descriptive text and a 'Go to [service]' link.

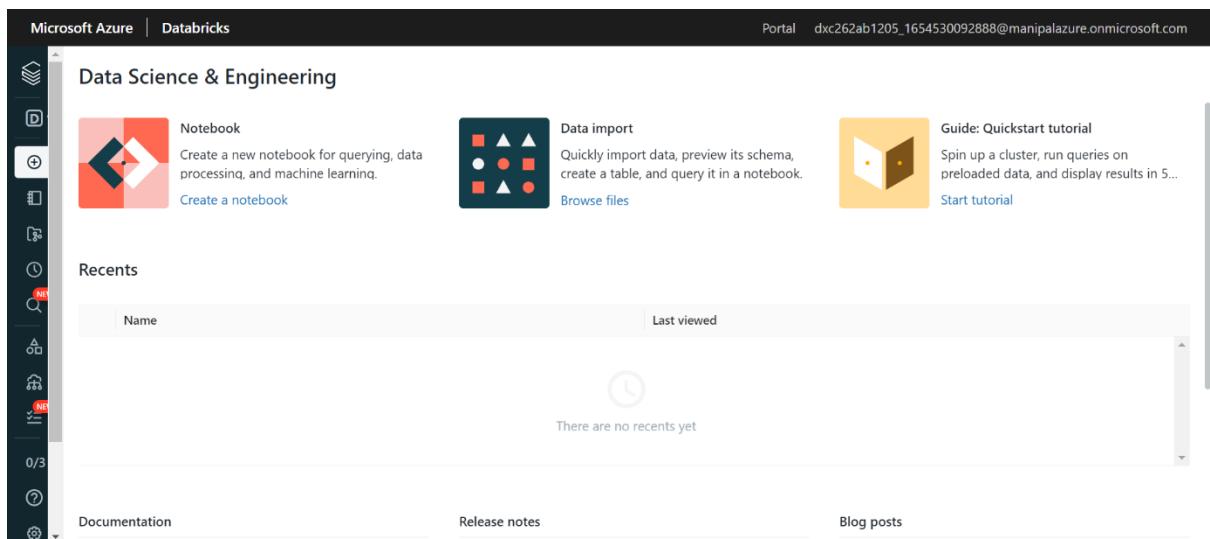
The screenshot shows the Microsoft Azure Overview page for an Azure Databricks service named 'APMargin23'. The header bar includes the Microsoft Azure logo, a search bar, and various navigation icons. The title 'APMargin23 | Overview' is displayed, along with an 'Azure Databricks Service' icon. A 'Delete' button is visible. The left sidebar contains a navigation menu with sections like 'Overview', 'Activity log', 'Access control (IAM)', 'Tags', 'Settings', 'Virtual Network Peerings', 'Encryption', 'Properties', 'Locks', 'Automation', 'Tasks (preview)', and 'Export template'. The 'Overview' section is selected. The main content area displays 'Essentials' information, including status 'Active', resource group 'dxcrg2317', location 'East US', subscription 'Azure-DXC262AB12Lab', and a URL 'https://adb-7716123982395834.14.azuredatabricks.net'. There is also a 'Managed Resource Group' entry for 'databricks-rg_APMargin23-3j6jagen7lq4e'. A 'JSON View' button is located at the top right of the essentials table. At the bottom center of the page is the red Databricks logo.

Step 7: Click on overview and launch workspace.



Practical Lab: Create Cluster in Azure Databricks

Step 1: Open Azure Databricks



Step 2: Click on create, and create a cluster.

The screenshot shows the Microsoft Azure Databricks home page. On the left, there's a sidebar with various navigation options like Data Science & Machine Learning, Workspace, Repos, Recents, Search, Data, Compute, Workflows, and Help. A red box highlights the 'Cluster' option under the Compute section. The main content area has sections for Notebook, Table, Cluster, Job, Repo, and AutoML Experiment. There are also links for Data import, Guide: Quickstart tutorial, and Start tutorial. At the bottom, there are Release notes and Blog posts.

Step 3: Complete all the data & click on create cluster.

This screenshot shows the 'New Cluster' configuration page. It includes fields for Databricks runtime version (set to 10.4 LTS), Autopilot options (terminating after 30 minutes of inactivity), Node type (Standard_DS3_v2), and DBU / hour (0.75). A promotional discount message is displayed: '50% promotional discount applied to Photon during preview'. The 'Create Cluster' button is at the top right.

Step 4: it configures all the data and creates cluster.

This screenshot shows the configuration page for the 'morgon123' cluster. It displays the same settings as the previous screen: Cluster mode (Single Node), Databricks Runtime Version (10.4 LTS), Autopilot options (terminating after 30 minutes of inactivity), Node type (Standard_DS3_v2), and DBU / hour (0.75). The 'Edit' and 'Terminate' buttons are visible at the top right of the cluster card.

Practical Lab: Add notebook in Databricks and Implement the Business Logic

The screenshot shows the Microsoft Azure Databricks workspace interface. On the left, there's a sidebar with various options like 'Data Science & E...', 'Create', 'Workspace', 'Repos', 'Recents', 'Search', 'Data', 'Compute', 'Workflows', and 'Help'. The 'Create' button is highlighted. In the main area, there's a 'Workspace' section with 'Shared' and 'Users' options. A context menu is open over the 'Create' button, with 'Notebook' being the selected option. The URL in the browser bar is <https://adb-5512239340581971.11.azuredatabricks.net/?o=5512239340581971#setting/clusters/0612-062933-zsgg3vt7/configuration>. The status bar at the bottom shows the date as 12-06-2022 and the time as 12:15.

This screenshot shows the 'Create Notebook' dialog box overlaid on the Databricks workspace. The dialog has fields for 'Name' (set to 'notebook1'), 'Default Language' (set to 'Python'), and 'Cluster' (set to 'morgon123'). There are 'Cancel' and 'Create' buttons at the bottom. The background workspace is visible behind the dialog.

The screenshot shows a Microsoft Azure Databricks notebook titled "notebook1" running Python code. The code in Cmd 1 is:

```
x=100  
y=200  
print(x+y)
```

The output is 300. The command took 0.04 seconds. In Cmd 2, there is a single digit '1' entered, with a note "Shift+Enter to run".

Practical Lab: Azure Data Factory For AP Morgan

Practical Lab: Create Azure Databricks Linked Service in ADF

Open azure home page and search for Data Factory

The screenshot shows the Azure portal homepage. The "Data factories" service is selected in the main navigation bar. Below it, there are sections for "Resources" (Recent and Favorite), "Useful links" (Overview, Get started, Documentation), and "Last Viewed" (Azure Databricks Service, Resource group). At the bottom, there are links for "Subscriptions", "Resource groups", "All resources", "Dashboard", and various tools like Microsoft Learn, Azure Monitor, Microsoft Defender for Cloud, and Cost Management.

Create a new account on Azure Data Factory

Home >

Data factories ⚙️ ...

Manipal Pro Learn (manipalazure.onmicrosoft.com)

+ Create Manage view Refresh Export to CSV Open query Assign tags

Filter for any field... Subscription == all Type == all Resource group == all Location == all Add filter

No grouping List view

Name ↑ Type ↑ Subscription ↑ Resource group ↑ Location ↑

 No data factories to display

Try changing or clearing your filters.

Create data factory Learn more ⓘ

Enter the details for the factory account

Home > Data factories >

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 * Resource group * Create new

Instance details

Name * Region * Version *

Review + create < Previous Next : Git configuration >

Validate and deploy the Data Factory

Home > Data factories >

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; and (b) 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	Azure-DXC262AB12Lab
Resource group	resourcegroup
Name	AP-morgan
Region	East US
Version	V2 (Recommended)

Networking

Connect via	Public endpoint
-------------	-----------------

Actions

- [Create](#)
- [< Previous](#)
- [Next >](#)
- [Download a template for automation](#)

After the deployment is done , click on “Go to resource”

Home >

Microsoft.DataFactory-20220611221708 | Overview

Deployment

Overview

We'd love your feedback!

Your deployment is complete

Deployment name: Microsoft.DataFactory-20220611221708
 Subscription: Azure-DXC262AB12Lab
 Resource group: resourcegroup

Start time: 6/11/2022, 10:18:22 PM
 Correlation ID: aee1bbae-0991-4da2-be42-e577d473e7dd

Deployment details (Download)

Next steps

[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 >](#)

Now open the Data Factory

Home > Microsoft.DataFactory-20220611221708 >

AP-morgan Data factory (V2)

Essentials

- Resource group (move) : **resourcegroup**
- Status : Succeeded
- Location : East US
- Subscription (move) : Azure-DXC-262AB12Lab
- Subscription ID : 4236c42a-d131-4bd6-b609-aec3a598f2d3

Type : Data factory (V2)

Getting started : [Quick start](#)

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

We need to create a pipeline which connects both Data Bricks notebook and Data Factory.

Data Factory Validate all Publish all

Factory Resources

Filter resources by name

- Pipelines
- Datasets
- Data flows
- Power Query

Pipeline

Dataset

Data flow

Power Query

Template gallery

Import from pipeline template

Select an item

Use the resource explorer to select or create a new item

Create a new Pipeline

Data Factory Validate all Publish all

Factory Resources

Filter resources by name

- Pipelines
- Datasets
- Data flows
- Power Query

pipeline1

Activities

Search activities

- Move & transform
- Azure Data Explorer
- Azure Function
- Batch Service
- Notebook**
- Jar
- Python
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

Notebook

Notebook1

General Azure Databricks Settings User properties

Data bricks linked service * Select... New

Properties

General Related

Name * pipeline1

Description

Annotations

+ New

Create a new Linked service for databricks

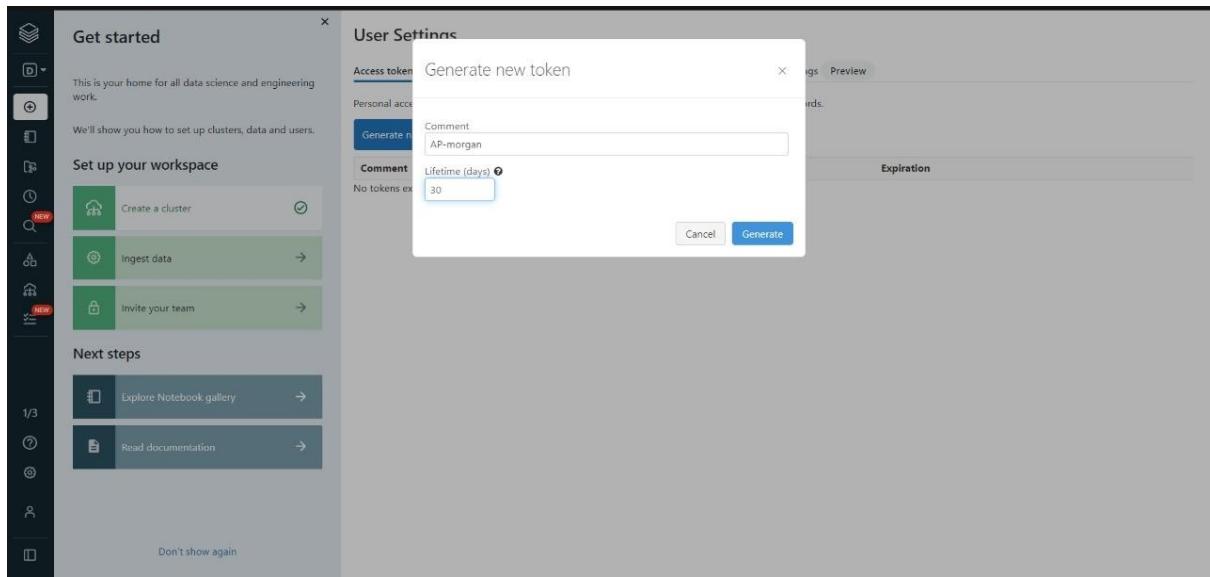
The screenshot shows the Azure Data Factory interface. On the left, there's a sidebar with 'Factory Resources' like Pipelines, Datasets, Data flows, and Power Query. In the center, there's a 'Activities' section with options like Move & transform, Azure Data Explorer, Azure Function, Batch Service, and Databricks. A 'Notebook' activity is selected, showing a sub-menu for Notebook1. On the right, a 'New linked service' dialog is open for 'Azure Databricks'. It asks for 'Account selection method' (From Azure subscription or Enter manually), 'Azure subscription' (selected as 'Azure-DXC262AB12Lab'), 'Databricks workspace' (selected as 'brick101'), 'Select cluster' (Existing job cluster selected), 'Databrick Workspace URL' (https://adb-2843779180221864.azuredatabricks.net), 'Authentication type' (Access Token selected), 'Access token' (a text input field containing a placeholder 'I'), 'Annotations' (a plus sign icon), and 'Create' and 'Cancel' buttons. There's also a 'Test connection' link at the bottom.

We need the access token of Data bricks account.

go to Users in settings

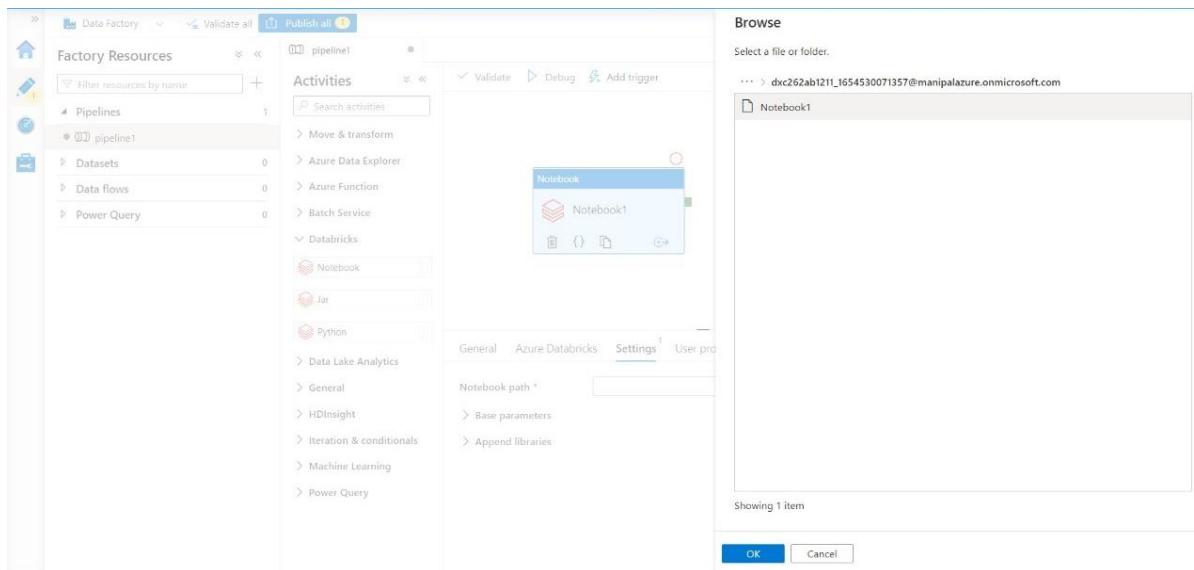
The screenshot shows the Databricks workspace interface. On the left, there's a sidebar with 'Data Science & Engineering', 'Create', 'Workspace', 'Repos', 'Recents', 'Search', 'Data', 'Compute', and 'Workflows'. A 'User Settings' dropdown is open, showing 'Admin Console' and 'Manage Account'. The main area is a 'Notebook1' window in Python mode. It has two code cells: 'Cmd 1' with the code 'print("Do some computation on data")' and output 'Do some computation on data' and 'Command took 0.84 seconds --- by dxc262ab1211_1654530071357@manipalazure.onmicrosoft.com at 6/11/2022, 10:37:09 PM on newCluster'; and 'Cmd 2' with the code '1'. Below the notebook, there's a message 'Shift+Enter to run'. At the bottom, there's a URL: https://adb-2843779180221864.azuredatabricks.net/?o=2843779180221864#setting/account.

Click on “Generate token”



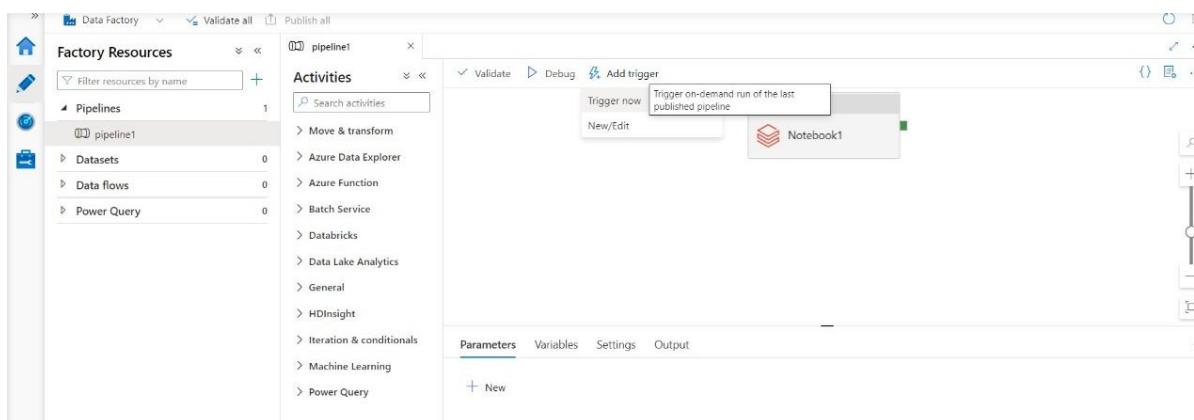
The screenshot shows the Azure Data Factory pipeline editor. On the left, the 'Factory Resources' pane lists Pipelines, Datasets, Data flows, and Power Query. The 'Activities' pane shows a pipeline named 'pipeline1' with several activities listed under 'Move & transform', 'Azure Data Explorer', 'Azure Function', 'Batch Service', and 'Data bricks'. A specific activity for 'Notebook' is selected, showing 'Notebook1' with a preview icon. To the right, a 'New linked service' dialog is open. It's configured for 'Azure Databricks' with the URL 'https://adb-2843779100221864.4.azuredatabricks.net'. The 'Authentication type' is set to 'Access Token', and the 'Access token' field contains a redacted value. Other options include selecting a cluster ('Existing interactive cluster' is chosen) and choosing from existing clusters ('newcluster'). Buttons at the bottom include 'Create', 'Cancel', and 'Test connection'.

Choose the right cluster and click on create

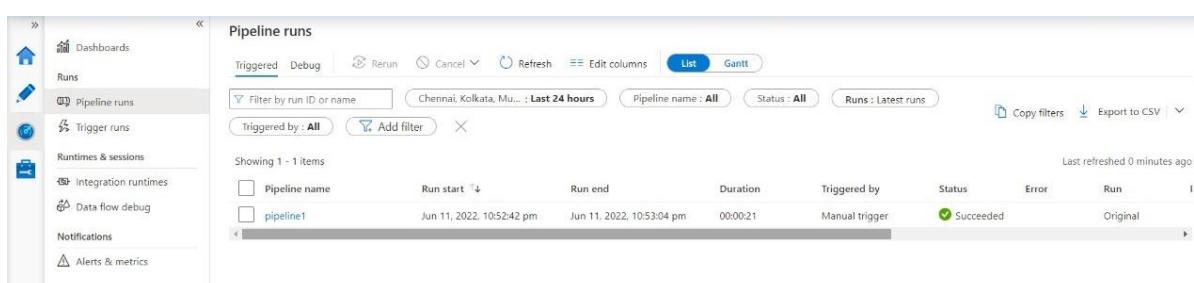


Now choose the notebook in the pipeline and Publish the pipeline in order to trigger it .

After publication is done click on “Trigger now”



Go to Monitor window to check execution



Here we can see we have successfully triggered.

Result: : In this project a data is able to link and trigger azure DataBricks notebook.

Conclusion: Azure Data Factory linked with Azure Data bricks.