

# OLYMPIC ANALYSIS PROJECT REPORT

## 1. Creating a Storage account

The screenshot shows the 'Create a storage account' wizard on the Microsoft Azure portal. The user is on the first step, 'Project details'. The 'Subscription' dropdown is set to 'Azure subscription 1'. The 'Resource group' dropdown shows '(New) rg\_Project' with a 'Create new' link. The 'Storage account name' field contains 'projectfinalladls'. The 'Region' dropdown is set to '(Asia Pacific) Central India' with a 'Deploy to an Azure Extended Zone' link. At the bottom, there are 'Previous', 'Next', and 'Review + create' buttons.

The screenshot shows the 'Create a storage account' wizard on the second step. The 'the Azure portal' dropdown is selected. The 'Minimum TLS version' dropdown is set to 'Version 1.2'. The 'Permitted scope for copy operations (preview)' dropdown is set to 'From any storage account'. Under 'Hierarchical Namespace', there is a note about Data Lake Storage Gen2 endpoint and access control lists (ACLs). The 'Enable hierarchical namespace' checkbox is checked. Under 'Access protocols', it says 'Blob and Data Lake Gen2 endpoints are provisioned by default'. The 'Enable SFTP' checkbox is unchecked. At the bottom, there are 'Previous', 'Next', and 'Review + create' buttons.

## Storage account Deployment is in Process

The screenshot shows the Azure Storage account overview page for 'projectfinaladls\_1752128943713'. The main heading 'Deployment is in progress' is displayed. Deployment details show the name as 'projectfinaladls\_1752128943713', start time as '7/10/2025, 11:58:37 AM', subscription as 'Azure subscription 1', and resource group as 'rg\_Project'. A correlation ID is also listed. On the right side, there are promotional links for Microsoft Defender for Cloud, Free Microsoft tutorials, and Work with an expert.

The screenshot shows the Azure Storage account overview page for 'projectfinaladls\_1752128943713'. The main heading 'Your deployment is complete' is displayed. Deployment details show the name as 'projectfinaladls\_1752128943713', start time as '7/10/2025, 11:58:37 AM', subscription as 'Azure subscription 1', and resource group as 'rg\_Project'. Below the deployment details, there is a 'Next steps' section with a 'Go to resource' button. On the right side, there are promotional links for Cost Management, Microsoft Defender for Cloud, and Free Microsoft tutorials.

## Azure Data Lake Storage is created

## 2. Created Containers inside the Storage account

Name	Last modified	Anonymous access level	Lease state
\$logs	7/10/2025, 11:59:12 AM	Private	Available
input	7/10/2025, 12:23:38 PM	Private	Available
output	7/10/2025, 12:24:04 PM	Private	Available
raw	7/10/2025, 12:23:52 PM	Private	Available

### Upload the datasets in the input container

Name	Last modified	Access tier	Blob type	Size	Lease state
DIM.Date.Table.csv	7/10/2025, 12:24:59 PM	Hot (Inferred)	Block blob	778.16 KiB	Available
OlympicsAnalysis...	7/10/2025, 12:25:34 PM	Hot (Inferred)	Block blob	5.96 MiB	Available
OlympicsAnalysis...	7/10/2025, 12:25:51 PM	Hot (Inferred)	Block blob	3.22 KiB	Available

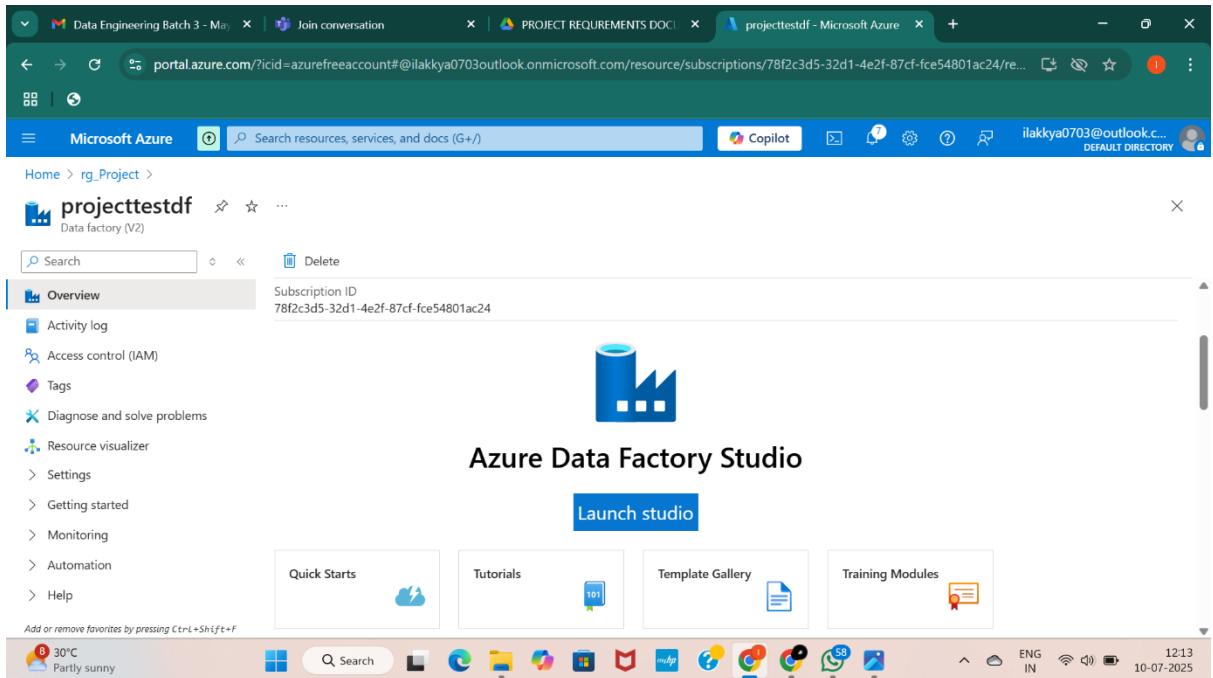
### 3. Ingesting the Input files by using Azure Data Factory (ADF)

The screenshot shows the Microsoft Azure Data factories page. At the top, there is a search bar and a Copilot button. The main heading is "Data factories" with a "Default Directory" link. Below the heading are buttons for "+ Create", "Manage view", "Refresh", "Export to CSV", "Open query", and "Assign tags". A message indicates "You are viewing a new version of Browse experience. Some features may be missing. Click here to access the old experience." There are filter options: "Subscription equals all", "Type equals all", "Resource Group equals all", "Location equals all", and a "Filter for any field..." dropdown. A large icon of a factory building is centered, with the text "No data factories to display" below it. A descriptive paragraph about Azure Data Factory follows. At the bottom, there is a "Create" button and a "Display count" dropdown set to 10.

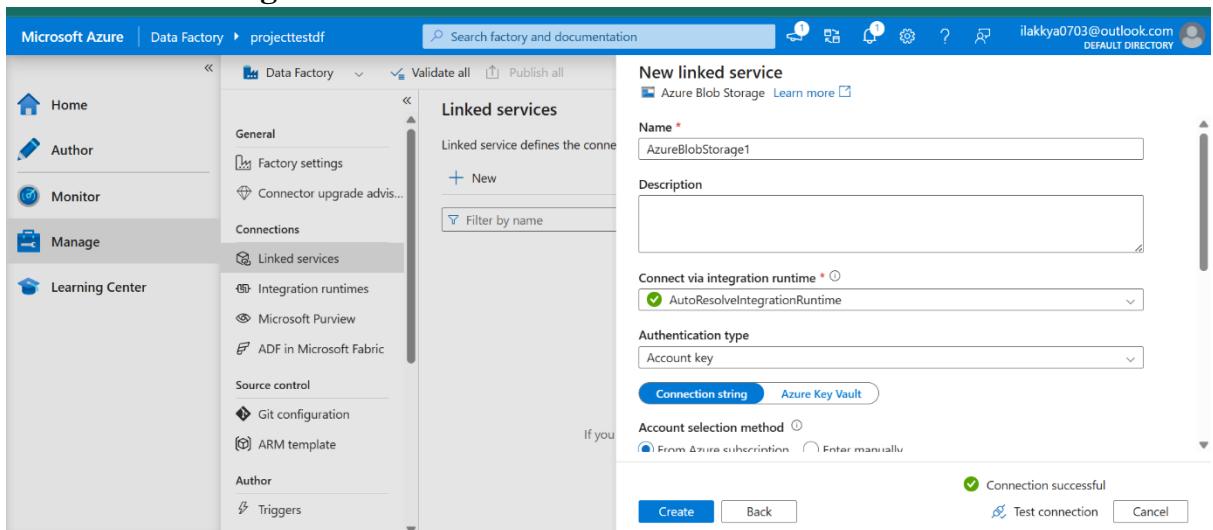
The screenshot shows the "Create Data Factory" wizard. The first step, "Project details", is displayed. It asks to select a subscription and resource group. The subscription dropdown shows "Azure subscription 1" and the resource group dropdown shows "rg\_Project" with a "Create new" option. The "Instance details" section includes fields for "Name" (set to "projectdf"), "Region" (set to "East US"), and "Version" (set to "V2"). At the bottom, there are "Previous" and "Next" buttons, and a "Review + create" button.

The screenshot shows the "Microsoft.DataFactory-20250710121158 | Overview" page. The deployment status is shown as "Your deployment is complete". Deployment details include: Deployment name: Microsoft.DataFactory-2025071..., Start time: 7/10/2025, 12:11:46 PM, Subscription: Azure subscription 1, Correlation ID: 7178e4f2-ffe-47be-a3ee-560a..., and Resource group: rg\_Project. There are sections for "Deployment details" and "Next steps". Buttons for "Go to resource", "Give feedback", and "Tell us about your experience with deployment" are present. On the right side, there are promotional cards for "Cost management", "Microsoft Defender for Cloud", and "Free Microsoft tutorials".

## Now that Data Factory is created, launch the studio



## 4. Create Linked Services →For Blob Storage



## →For ADLS

The screenshot shows the Microsoft Azure Data Factory interface. On the left, the navigation menu includes Home, Author, Monitor, Manage, and Learning Center. The main content area is titled "Linked services" under the "Connections" section. A sub-menu for "Linked services" is open, showing "General", "Factory settings", "Connector upgrade advis...", "Connections", and "Linked services". Under "Linked services", there is a "New" button and a search bar. Below it, a table shows one item: "Name" (AzureBlobStorage1), "Type" (Azure Blob Storage), and "Related" (0). To the right, a form for creating a new linked service for "Azure Data Lake Storage Gen2" is displayed. It includes fields for "Storage account name" (projectfinaladls) and "Test connection" (which is successful). Buttons for "Create", "Back", "Test connection", and "Cancel" are at the bottom.

The screenshot shows the Microsoft Azure Data Factory interface. The left navigation menu is identical to the previous screenshot. The main content area is titled "Linked services" under the "Connections" section. A sub-menu for "Linked services" is open, showing "General", "Factory settings", "Connector upgrade advis...", "Connections", and "Linked services". Under "Linked services", there is a "New" button and a search bar. Below it, a table shows two items: "Name" (ADLS), "Type" (Azure Data Lake Stora...), and "Related" (0); and "Name" (AzureBlobStorage1), "Type" (Azure Blob Storage), and "Related" (0). The table has columns for Name, Type, Related, and Annotations.

## 5. Load the Datasets

The screenshot shows the Microsoft Azure Data Factory interface. The left navigation menu is identical to the previous screenshots. The main content area is titled "Factory Resources" under the "Datasets" section. A sub-menu for "Datasets" is open, showing "Pipelines", "Change Data Capture (preview)", "Datasets", "Data flows", and "Power Query". Under "Datasets", there is a list item "Ds\_Dim\_Table". The details pane on the right shows the properties for "Ds\_Dim\_Table". It includes a preview of a CSV file named "DelimitedText", a "Properties" panel with tabs for "General" and "Related", and a "Linked service" dropdown set to "AzureBlobStorage1". The "File path" is set to "input" and "Compression type" is set to "No compression". A "Test connection" button shows a "Connection successful" message. Buttons for "Edit", "New", and "Learn more" are also present.

Microsoft Azure | Data Factory > projecttestdf

Search factory and documentation

Home Author Monitor Manage Learning Center

Data Factory Validate all Publish all 1

Factory Resources Ds\_Dim\_Table

Filter resources by name +

Pipelines 0 Change Data Capture (preview) 0 Datasets 1 Ds\_Dim\_Table

Data flows 0 Power Query 0

Set properties

Name Olympic\_Analysis\_Source1

Linked service \* AzureBlobStorage1

File path input / Directory / OlympicsAnalysis\_Sour...

First row as header

Import schema From connection/store  From sample file  None

Test connection  Connection successful

File path \*

Compression type

Column delimiter

OK Back Cancel

Microsoft Azure | Data Factory > projecttestdf

Search factory and documentation

Home Author Monitor Manage Learning Center

Data Factory Validate all Publish all 2

Preview experience Off

Factory Resources Ds\_Dim\_Table

Filter resources by name +

Pipelines 0 Change Data Capture (preview) 0 Datasets 2 Ds\_Dim\_Table Olympic\_Analysis\_Source1

Data flows 0 Power Query 0

Ds\_Dim\_Table

Olympic\_Analysis\_Source1

Properties

General Related

Name \* Olympic\_Analysis\_Source1

Description

Annotations + New

Connection Schema Parameters

Linked service \* AzureBlobStorage1

Test connection  Connection successful

File path \* input / Di

Compression type No compression

Microsoft Azure | Data Factory > projecttestdf

Search factory and documentation

Home Author Monitor Manage Learning Center

Data Factory Validate all Publish all 2

Factory Resources Ds\_Dim\_Table

Filter resources by name +

Pipelines 0 Change Data Capture (preview) 0 Datasets 2 Ds\_Dim\_Table Olympic\_Analysis\_Source1

Data flows 0 Power Query 0

Set properties

Name Olympic\_Analysis\_Source2

Linked service \* AzureBlobStorage1

File path input / Directory / OlympicsAnalysis\_Sour...

First row as header

Import schema From connection/store  From sample file  None

Test connection  Connection successful

File path \*

Compression type No compression

OK Back Cancel

**Properties**

**General**

- Name: Olympic\_Analysis\_Source2
- Description:
- Annotations:

**Connection**

- Linked service: AzureBlobStorage1
- Test connection: Connection successful

**File path**: input

**Compression type**: No compression

## For Raw container

**Properties**

**General**

- Name: Ds\_Dim\_Table\_Raw
- Description:
- Annotations:

**Connection**

- Linked service: ADLS
- Test connection: Connection successful

**File path**: raw

**Compression type**: No compression

**Set properties**

**Name**: Olympic\_Analysis\_Source1\_Raw

**Linked service**: ADLS

**File path**: raw / Directory / File name

**First row as header**: checked

**Import schema**

- From connection/store: selected
- From sample file: unselected
- None: unselected

**Compression type**: No compression

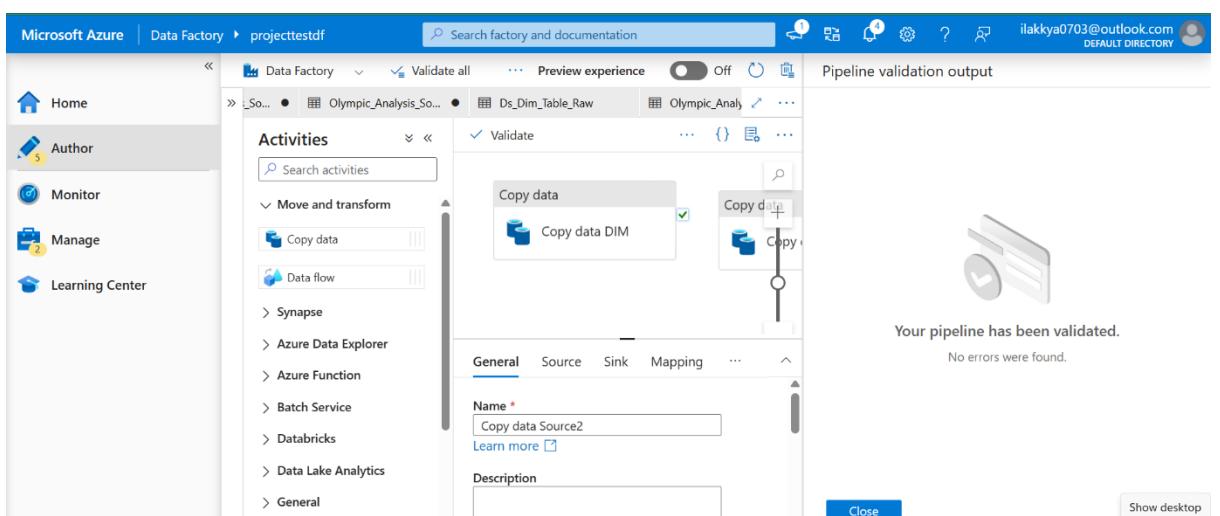
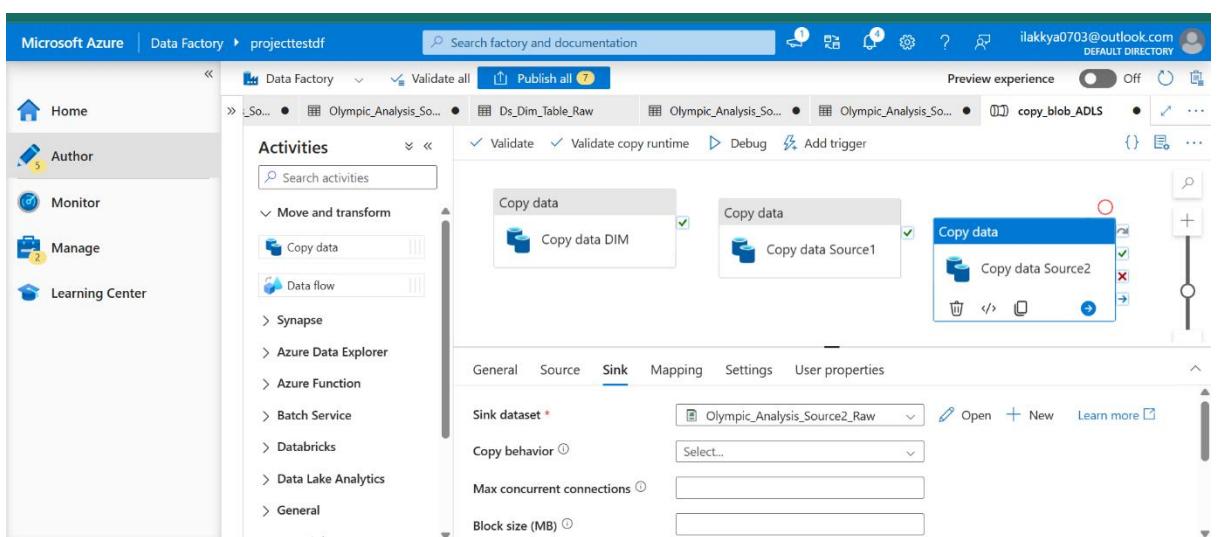
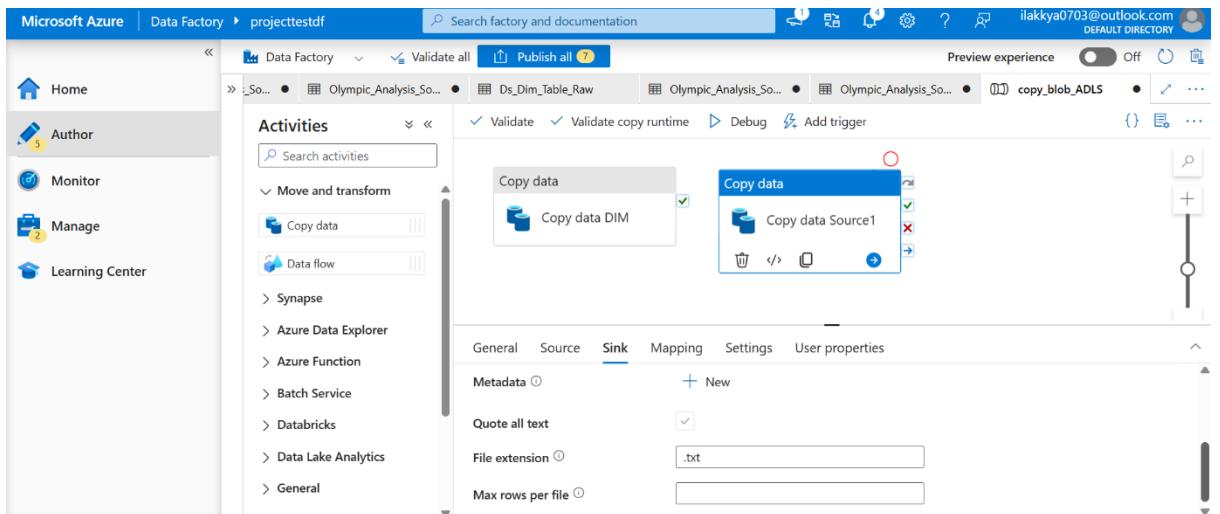
**OK** | **Back** | **Cancel**

The screenshot shows the Microsoft Azure Data Factory interface. On the left, the navigation bar includes Home, Author, Monitor, Manage, and Learning Center. The main area displays 'Factory Resources' with sections for Pipelines, Datasets, and Data flows. A specific dataset, 'Olympic\_Analysis\_Source1\_Raw', is selected. The properties pane on the right shows the following details:

- Name:** Olympic\_Analysis\_Source1\_Raw
- Connection:** ADLS
- Description:** (empty)
- Annotations:** + New
- File path:** raw /
- Compression type:** No compression

## 6. Creating Pipelines

The screenshot shows the Microsoft Azure Data Factory interface with a validation dialog open. The dialog title is 'Validate' and it contains a single activity named 'Copy data DIM'. The validation status is shown as successful with a green checkmark icon. The validation output pane on the right displays the message: 'Your pipeline has been validated. No errors were found.' The validation status is also indicated by a green checkmark icon.



## Debugging the Pipeline

Microsoft Azure | Data Factory > projecttestdf

Search factory and documentation

Validate all Publish all

Preview experience Off

ilakkya0703@outlook.com DEFAULT DIRECTORY

Home Author Monitor Manage Learning Center

Activities

Move and transform

Copy data

Copy data DIM

Copy data Source1

Copy data Source2

Validate Cancel options Add trigger

Parameters Variables Settings Output

Activity name Activity st... Activit... Run start Duration Integration

Activity name	Activity st...	Activit...	Run start	Duration	Integration
Copy data Source1	Queued	Copy data	7/10/2025, 1:00:33 PM	53s	
Copy data DIM	Queued	Copy data	7/10/2025, 1:00:33 PM	53s	
Copy data Source2	Queued	Copy data	7/10/2025, 1:00:33 PM	53s	

Microsoft Azure | Data Factory > projecttestdf

Search factory and documentation

Validate Debug Add trigger

Copy data DIM

Copy data Source1

Copy data Source2

Expand toolbox pane

Parameters Variables Settings Output

Activity name Activity st... Activit... Run start Duration Integration runtime

Activity name	Activity st...	Activit...	Run start	Duration	Integration runtime
Copy data Source1	Succeeded	Copy data	7/10/2025, 1:00:33 PM	16s	AutoResolveIntegrationRuntime (East US)
Copy data DIM	Succeeded	Copy data	7/10/2025, 1:00:33 PM	15s	AutoResolveIntegrationRuntime (East US)
Copy data Source2	Succeeded	Copy data	7/10/2025, 1:00:33 PM	16s	AutoResolveIntegrationRuntime (East US)

## Publish All

Microsoft Azure | Data Factory > projecttestdf

Search factory and documentation

Validate all Publishing

ilakkya0703@outlook.com DEFAULT DIRECTORY

Home Author Monitor Manage Learning Center

Copy data DIM

Copy data Source1

Copy data Source2

Publish all

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

Pending changes (7)

NAME	CHANGE	EXISTING
copy_blob_ADLS	(New)	-
Ds_Dim_Table	(New)	-
Ds_Dim_Table_Raw	(New)	-
Olympic_Analysis_Source1	(New)	-
Olympic_Analysis_Source2	(New)	-
Olympic_Analysis_Source1...	(New)	-
Olympic_Analysis_Source2...	(New)	-

Parameters Variables Settings Output

Activity name Activity st... Activit... Run start Duration

Activity name	Activity st...	Activit...	Run start	Duration
Copy data Source1	Succeeded	Copy data		
Copy data DIM	Succeeded	Copy data		
Copy data Source2	Succeeded	Copy data		

Publish Cancel

## Now, we can see that the Files are copied to our raw Container

The screenshot shows the Microsoft Azure Storage Explorer interface. The left sidebar has a 'raw' container selected under 'Containers'. The main area shows a table of blobs in the 'raw' container:

	Name	Last modified	Access tier	Blob type	Size	Lease state
<input type="checkbox"/>	DIM.Date.Table.csv	7/10/2025, 1:00:45 PM	Hot (Inferred)	Block blob	778.16 KiB	Available
<input type="checkbox"/>	OlympicsAnalysis...	7/10/2025, 1:00:46 PM	Hot (Inferred)	Block blob	5.96 MiB	Available
<input type="checkbox"/>	OlympicsAnalysis...	7/10/2025, 1:00:46 PM	Hot (Inferred)	Block blob	3.22 KiB	Available

## 7. Create a Databricks Workspace

The screenshot shows the 'Create an Azure Databricks workspace' wizard. The 'Basics' tab is selected. The 'Project Details' section shows:

- Subscription: Azure subscription 1
- Resource group: rg\_Project

The 'Instance Details' section shows:

- Workspace name: Olympics\_Analysis

At the bottom, there are 'Review + create' and 'Next : Networking >' buttons.

The screenshot shows the 'rg\_Project\_Olympics\_Analysis | Overview' page. The 'Deployment' section indicates 'Deployment is in progress' with the following details:

- Deployment name: rg\_Project\_Olympics\_Analysis
- Subscription: Azure subscription 1
- Resource group: rg\_Project
- Start time: 7/10/2025, 3:22:40 PM
- Correlation ID: e3621b22-e753-4c99-93de-1d...

The 'Deployment details' section shows a table with columns: Resource, Type, Status, and Operator. A note at the bottom says 'There are no resources to display.'

On the right side, there are promotional links for Microsoft Defender for Cloud, Microsoft tutorials, and Azure experts.

## Launch your Workspace

The screenshot shows the Microsoft Azure Databricks Service Overview page for the 'Olympics\_Analysis' workspace. The top navigation bar includes 'Microsoft Azure', 'Upgrade', 'Search resources, services, and docs (G+)', 'Copilot', and user information 'ilakkya0703@outlook.c... DEFAULT DIRECTORY'. The main area features a large red cube icon and a 'Launch Workspace' button. On the left, a sidebar lists options like 'Overview', 'Activity log', 'Access control (IAM)', 'Tags', 'Diagnose and solve problems', 'Resource visualizer', 'Settings' (with 'Virtual Network Peering', 'Encryption', 'Networking', and 'Security & compliance' sub-options), and a 'Documentation' section with 'Getting Started' and 'Import Data from File' and 'Import Data from Azure Storage' buttons.

## Create a Folder and in that create a Notebook

The screenshot shows the Databricks workspace interface. The left sidebar includes 'New' (highlighted), 'Workspace', 'Recents', 'Catalog', 'Jobs & Pipelines', 'Compute', 'Marketplace', 'SQL', 'SQL Editor', 'Queries', 'Dashboards', 'Genie', 'Alerts', 'Query History', and 'SQL Warehouses'. The main area shows a 'Workspace' tree with 'Home', 'Project', 'Workspace', 'Favorites', and 'Trash'. A search bar at the top right says 'Search data, notebooks, recents, and more...' and includes 'CTRL + P'. To the right, a table lists a single project named 'ilakkya0703@outlook.com' with details: Name (ilakkya0703@outlook.com), Type (Folder), Owner (Ilakkya M), and Created at (Jul 10, 2025, 03:27 ...).

## Compute your Cluster

The screenshot shows the Databricks Compute interface. The left sidebar includes 'New' (highlighted), 'Workspace', 'Recents', 'Catalog', 'Jobs & Pipelines', 'Compute' (highlighted), 'Marketplace', 'SQL', 'SQL Editor', 'Queries', 'Dashboards', 'Genie', 'Alerts', 'Query History', and 'SQL Warehouses'. The main area shows a 'Compute' tree with 'New compute' selected. A search bar at the top right says 'Search data, notebooks, recents, and more...' and includes 'CTRL + P'. Below it, a 'Create new compute' form is displayed with tabs for 'Summary' (16 GB Memory, 4 Cores), 'Data access' (Unity Catalog), 'Price' (1 DBU/h), 'General' (Compute name: 'ilakkya M's Cluster 2025-07-10 15:29:14'), 'Performance' (Machine learning: off), and 'Databricks runtime'. Buttons for 'UI' and 'JSON' are at the top right of the form.

## Access your Dataset from the input container through the SAS Key

Accessing through SAS Key

The screenshot shows two Jupyter Notebook cells. The first cell contains Scala code setting up an SAS token for Azure storage:

```
spark.conf.set("fs.azure.account.auth.type.projectfinaladls.dfs.core.windows.net", "SAS")
spark.conf.set("fs.azure.sas.token.provider.type.projectfinaladls.dfs.core.windows.net",
              "org.apache.hadoop.fs.azurebfs.sas.FixedSASTokenProvider")
spark.conf.set("fs.azure.sas.fixed.token.projectfinaladls.dfs.core.windows.net",
              "sv=2024-11-04&ss=bfqt&srt=co&sp=rwdlacupiytx&se=2025-07-09T18:39:15Z&st=2025-07-09T10:39:15Z&
              spr=https&sig=2wP9LyH1PXIIotUgsW7xQuJpvOjsngmJ%2BzRTJ8KWBI%3D")
```

The second cell contains Python code reading a CSV file from an Azure blob storage container:

```
spark.conf.set("fs.azure.account.key.projectfinaladls.blob.core.windows.net", "voxWV/6REP4fvRoo8ldzrWmslRX10oiyV3Z0CBPK3vZx9qnqdNscutPEA/OAsgIkriIi0M4wNqndD+ASTRs+VVQ==")
df = spark.read.option("header", "true").csv(
    "wasbs://input@projectfinaladls.blob.core.windows.net/OlympicsAnalysis_Source1.csv"
)
df.show(5)
```

Below the code, the resulting DataFrame schema and sample data are displayed:

Player_Name	Player_Gender	Player_Age	Player_Height	Player_Weight	Country_Name	Country_id	Olympic_Name	Olympic_Year	Olympic_Season	Olympic_City	Sport_Name	Event_Name	Medal
Aaron	Arthur	Cook	M	25	182	sal	Moldova	2016	Summer	2016	London	Swimming	Gold

After executing the Notebook, we can see the results in Output Container

The screenshot shows the Azure Storage Explorer interface for the 'output' container. The left sidebar includes options like 'Diagnose and solve problems', 'Access Control (IAM)', 'Settings', 'Shared access tokens', 'Access policy', 'Properties', and 'Metadata'. The main area displays a list of blobs:

Name	Last modified	Access tier	Block type	Size	Lease state
DimensionTable...	7/9/2025, 5:49:09 PM	Hot (Inferred)	Block blob	0	Available
FactTables	7/9/2025, 8:10:40 PM	Hot (Inferred)	Block blob	0	Available
Reports					

You can see the Output in Dimension Tables, Fact Tables, and Reports

## 8. Create SQL Database

The screenshot shows the Microsoft Azure portal interface for creating a SQL database server. At the top, there's a navigation bar with 'Microsoft Azure', 'Upgrade', a search bar, 'Copilot', and user information ('ilakky0703@outlook.com'). Below the navigation bar, the main title is 'Create SQL Database Server'. Underneath it, there's a sub-section titled 'Create SQL Database Server' with a 'Microsoft' logo.

The configuration section starts with 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 Microsoft Entra authentication [Learn more](#) using an existing Microsoft Entra user, group, or application as Microsoft Entra admin [Learn more](#), or select both SQL and Microsoft Entra authentication.' There are three radio button options: 'Use Microsoft Entra-only authentication' (selected), 'Use both SQL and Microsoft Entra authentication', and 'Use SQL authentication'.

Below this, there's a section for 'Set Microsoft Entra admin' with a text input field containing 'ilakky0703\_outlook.com#EXT#@ilakky0703outlook.onmicrosoft.com'. It also shows the Admin Object/App ID: '9446dd3f-6eac-4164-ac76-d63d3d264ad3' and a 'Set admin' button.

At the bottom of the configuration area, there's an 'OK' button and a 'Feedback' link.

Below the configuration area, the URL in the browser is 'https://portal.azure.com/#create/Microsoft.SQLServer'. The page title is 'Create SQL Database'. The top navigation bar is identical to the one above.

The main content area shows tabs: 'Basics', 'Networking', 'Security', 'Additional settings', 'Tags', and 'Review + create' (which is underlined, indicating it's the active tab). The 'Product details' section includes 'SQL database by Microsoft' and links for 'Terms of use' and 'Privacy policy'. The 'Estimated cost per month' section is also visible.

The 'Terms' section contains legal text: 'By clicking "Create", I (a) agree to the legal terms and privacy statement(s) associated with the Marketplace offering(s) listed above; (b) authorize Microsoft to bill my current payment method for the fees associated with the offering(s), with the same billing frequency as my Azure subscription; and (c) agree that Microsoft may share my contact, usage and transactional information with the provider(s) of the offering(s) for support, billing and other transactional activities. Microsoft does not provide rights for third-party offerings. For additional details see [Azure Marketplace Terms](#).<sup>1</sup>'

At the bottom, there are buttons for 'Create', '< Previous', and 'Download a template for automation'.

Microsoft Azure    [Upgrade](#)    [Search resources, services, and docs \(G+\)](#)

ilakky0703@outlook.c...  
DEFAULT DIRECTORY

# Microsoft.SQLDatabase.newDatabaseNewServer\_c2fd2a4797f441cfb43f1 | Overview

Deployment

Search    [Delete](#)    [Cancel](#)    [Redeploy](#)    [Download](#)    [Refresh](#)

**Overview**

**Deployment is in progress**

Deployment name : Microsoft.SQLDatabase.newDa... Start time : 7/10/2025, 1:20:51 PM  
Subscription : Azure subscription 1 Correlation ID : 939d10fe-6110-46f4-969e-72f5...  
Resource group : rg\_Project

**Deployment details**

Resource	Type	Status	Operator
ilakky	Microsoft.Sql/servers	Accepted	Operation

Add or remove favorites by pressing **Ctrl+Shift+F**

Microsoft Azure    [Upgrade](#)    [Search resources, services, and docs \(G+\)](#)

ilakkya0703@outlook.c...  
DEFAULT DIRECTORY

Home >

# Microsoft.SQLDatabase.newDatabaseNewServer\_c2fd2a4797f441cfb43f1 | Overview

Deployment

Search    [Delete](#)    [Cancel](#)    [Redeploy](#)    [Download](#)    [Refresh](#)

**Overview**

**Your deployment is complete**

Deployment name : Microsoft.SQLDatabase.newDa... Start time : 7/10/2025, 1:21:03 PM  
Subscription : Azure subscription 1 Correlation ID : 939d10fe-6110-46f4-969e-72f5...  
Resource group : rg\_Project

**Deployment details**

**Next steps**

[Go to resource](#)

**Deployment succeeded**  
Deployment 'Microsoft.SQLDatabase.newDatabaseNewServer\_c2fd2a4797f441cfb43f1' to resource group 'rg\_Project' was successful.

[Go to resource](#)    [Go to resource group](#)

**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**  
[Cloud migration](#)

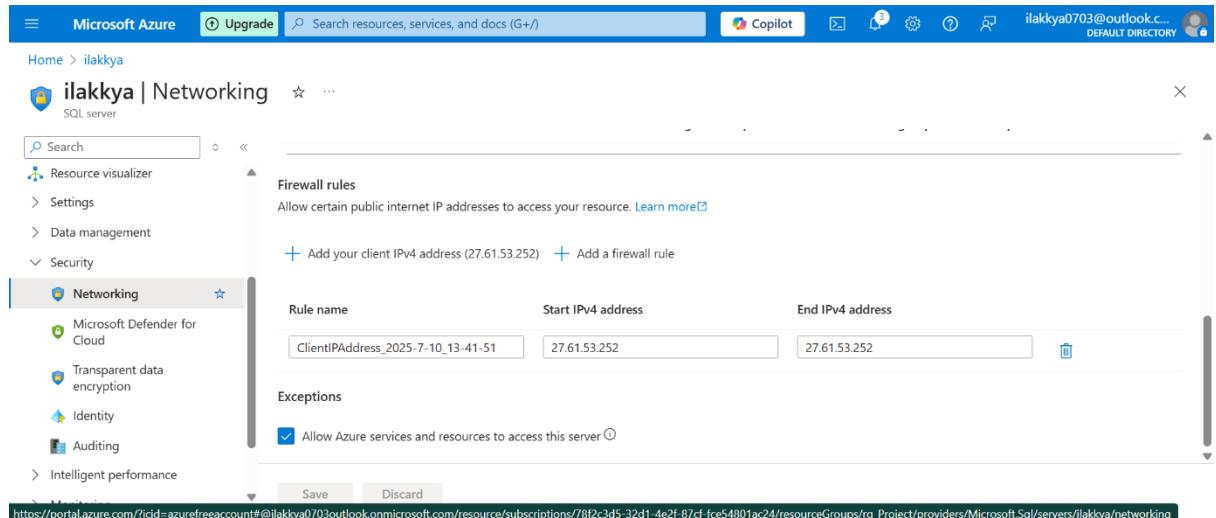
The screenshot shows the Microsoft Azure portal with the URL <https://portal.azure.com/?cid=azurerefreaccount#%2fresource%2fsubscriptions%2f78fc2d5-32d1-4e2f-87cf-fce54801ac...>. The page title is "ProjectFinal (ilakky/ProjectFinal) | Query editor (preview)". The left sidebar has a "Query editor (preview)" link under the "Azure SQL Database" section. The main content area features a large "SQL" logo and the text "Welcome to SQL Database Query Editor". It also includes sections for "SQL server authentication" (Login: CloudSA8639f45f, Password: [redacted]) and "Microsoft Entra authentication" (Logged in as ilakky0703@outlook.com, Continue as ilakky0703@outlook.com).

**Before starting with the Database,**

**--->Make sure that the Firewalls are configured**

**---> Add your client IP address**

**---> Save the Changes**



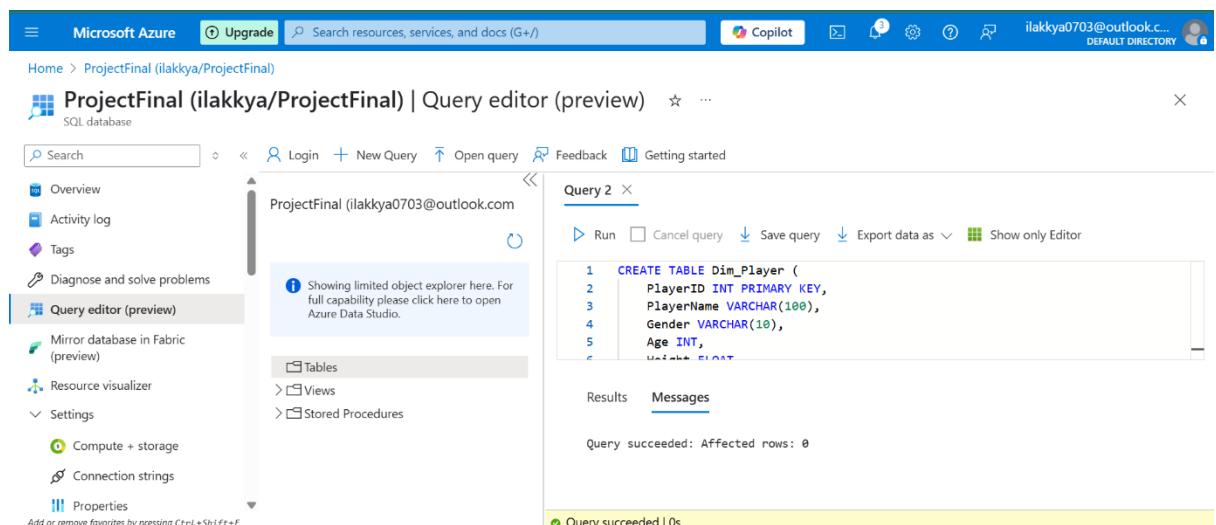
The screenshot shows the Microsoft Azure portal interface for managing a SQL server named 'ilakkya'. In the left sidebar, under the 'Networking' section, there is a 'Firewall rules' subsection. A new rule is being created with the following details:

Rule name	Start IPv4 address	End IPv4 address
ClientIPAddress_2025-7-10_13-41-51	27.61.53.252	27.61.53.252

Below the main table, there is an 'Exceptions' section with a checked checkbox for 'Allow Azure services and resources to access this server'.

## 9. Write your Query in the Query Editor

**--->Create Tables using SQL Query according to the Dimensional Tables and Fact Tables**



The screenshot shows the Microsoft Azure portal interface for a database named 'ProjectFinal'. In the left sidebar, under the 'Query editor (preview)' section, a query is being run in 'Query 2':

```
1 CREATE TABLE Dim_Player (
2     PlayerID INT PRIMARY KEY,
3     PlayerName VARCHAR(100),
4     Gender VARCHAR(10),
5     Age INT,
```

The results show that the query succeeded with 0 affected rows.

The screenshot shows the Microsoft Azure Query editor (preview) interface. The top navigation bar includes 'Microsoft Azure', 'Upgrade', 'Search resources, services, and docs (G+)', 'Copilot', and user information 'ilakkya0703@outlook.c... DEFAULT DIRECTORY'. The main title is 'ProjectFinal (ilakkya/ProjectFinal) | Query editor (preview)'. The left sidebar lists various database management options like Overview, Activity log, Tags, Diagnose and solve problems, and a prominent 'Query editor (preview)' section which is selected. Below this are options for Mirror database in Fabric (preview), Resource visualizer, Settings (Compute + storage, Connection strings), and Properties. The central pane displays a query titled 'Query 9' containing the following SQL code:

```

1 CREATE TABLE Fact_CountryDetails (
2     CountryFactID INT PRIMARY KEY,
3     CountryCode VARCHAR(10),
4     OlympicID INT,
5     OlympicYear INT,
6     CountryName VARCHAR(100)

```

The right pane shows tabs for 'Results' and 'Messages', with a search bar below.

**To connect with SQL using Azure Data Factory, remember your admin ID and Password**

The screenshot shows the Microsoft Azure SQL server blade for the 'ilakkya' server. The top navigation bar is identical to the previous screenshot. The main title is 'ilakkya | SQL server'. The left sidebar lists Overview, Activity log, Access control (IAM), Tags, Quick start, Diagnose and solve problems, Resource visualizer, and Settings (Microsoft Entra ID, SQL databases, SQL elastic pools). The central pane displays server details under the 'Essentials' section:

- Resource group: rg\_Project
- Status: Available
- Location: Central India
- Subscription: Azure subscription 1
- Subscription ID: 78f2c3d5-32d1-4e2f-87cf-fce54801ac24
- Tags: (edit), Add tags

A success message box is visible in the top right corner: 'Successfully reset administrator password' and 'Successfully reset administrator password for server: ilakkya'. At the bottom, there are tabs for Notifications (1), Features (6), and buttons for All, Info (1), and Recommendations (0).

## 10. Loading to SQL using Azure Data Factory (ADF)

---> Create Linked Service for SQL

The screenshot shows the 'Edit linked service' dialog in the Microsoft Azure Data Factory interface. The 'Name' field is set to 'AzureSqlDatabase1'. The 'Connect via integration runtime' dropdown is set to 'AutoResolveIntegrationRuntime'. The 'Version' is set to '2.0 (Recommended)'. The 'Account selection method' is set to 'Enter manually'. A 'Connection successful' message is displayed at the bottom right. The left sidebar shows navigation options like Home, Author, Monitor, Manage, and Learning Center.

## Loading Input dataset from ADLS

The screenshot shows the 'Set properties' dialog for a dataset named 'Ds\_Dim\_Table\_input'. The 'Linked service' is set to 'Ls\_ADLS Output'. The 'File path' is set to 'output / Dimension.Tables / File name'. The 'Import schema' section shows 'From sample file' selected. A file named 'Dim\_Player.csv' is selected in the 'Select file' dropdown. The dialog includes 'OK', 'Back', and 'Cancel' buttons.

The screenshot shows the 'Set properties' dialog for a dataset named 'Ds\_Dim\_Sport\_input'. The 'Linked service' is set to 'Ls\_ADLS Output'. The 'File path' is set to 'output / Dimension.Tables / File name'. The 'Import schema' section shows 'From sample file' selected. A file named 'Dim\_Sport.csv' is selected in the 'Select file' dropdown. The dialog includes 'OK', 'Back', and 'Cancel' buttons.

Microsoft Azure | Data Factory > projecttestdf

Search factory and documentation

Home Author Monitor Manage Learning Center

Data Factory Validate all Publish all

Factory Resources

- Pipelines 1
- Change Data Capture (preview) 0
- Datasets 8
  - Ds\_Dim\_Player\_input
  - Ds\_Dim\_Sport\_input
    - Ds\_Dim\_Table
    - Ds\_Dim\_Table\_Raw
    - Olympic\_Analysis\_Source1
    - Olympic\_Analysis\_Source1\_Raw
    - Olympic\_Analysis\_Source2
    - Olympic\_Analysis\_Source2\_Raw
- Data flows 0

Set properties

Name: Ds\_Dim\_Olympic\_input

Linked service: Ls\_ADLS Output

File path: output / Dimension\_Tables / File name

First row as header:

Import schema:  From sample file  None

Select file: Dim\_Olympic.csv

OK Back Cancel

This screenshot shows the 'Set properties' dialog for a dataset named 'Ds\_Dim\_Olympic\_input'. The 'Name' field is filled with the dataset's name. The 'Linked service' dropdown is set to 'Ls\_ADLS Output'. The 'File path' field contains 'output / Dimension\_Tables / File name'. The 'First row as header' checkbox is checked. Under 'Import schema', the 'From sample file' radio button is selected. A 'Select file' input field shows 'Dim\_Olympic.csv' with a 'Browse' button next to it. At the bottom are 'OK', 'Back', and 'Cancel' buttons.

Microsoft Azure | Data Factory > projecttestdf

Search factory and documentation

Home Author Monitor Manage Learning Center

Data Factory Validate all Publish all

Factory Resources

- Pipelines 1
- Change Data Capture (preview) 0
- Datasets 9
  - Ds\_Dim\_Olympic\_input
  - Ds\_Dim\_Player\_input
  - Ds\_Dim\_Sport\_input
    - Ds\_Dim\_Table
    - Ds\_Dim\_Table\_Raw
    - Olympic\_Analysis\_Source1
    - Olympic\_Analysis\_Source1\_Raw
    - Olympic\_Analysis\_Source2
    - Olympic\_Analysis\_Source2\_Raw
- Data flows 0

Set properties

Name: Ds\_Dim\_Medal\_input

Linked service: Ls\_ADLS Output

File path: output / Dimension\_Tables / File name

First row as header:

Import schema:  From sample file  None

Select file: Dim\_Medal.csv

OK Back Cancel

This screenshot shows the 'Set properties' dialog for a dataset named 'Ds\_Dim\_Medal\_input'. The 'Name' field is filled with the dataset's name. The 'Linked service' dropdown is set to 'Ls\_ADLS Output'. The 'File path' field contains 'output / Dimension\_Tables / File name'. The 'First row as header' checkbox is checked. Under 'Import schema', the 'From sample file' radio button is selected. A 'Select file' input field shows 'Dim\_Medal.csv' with a 'Browse' button next to it. At the bottom are 'OK', 'Back', and 'Cancel' buttons.

Microsoft Azure | Data Factory > projecttestdf

Search factory and documentation

Home Author Monitor Manage Learning Center

Data Factory Validate all Publish all

Factory Resources

- Pipelines 1
- Change Data Capture (preview) 0
- Datasets 11
  - Ds\_Dim\_Event\_input
  - Ds\_Dim\_Medal\_input
  - Ds\_Dim\_Olympic\_input
  - Ds\_Dim\_Player\_input
  - Ds\_Dim\_Sport\_input
    - Ds\_Dim\_Table
    - Ds\_Dim\_Table\_Raw
    - Olympic\_Analysis\_Source1
    - Olympic\_Analysis\_Source1\_Raw
- Data flows 0

Set properties

Name: Ds\_Dim\_Country\_input

Linked service: Ls\_ADLS Output

File path: output / Dimension\_Tables / File name

First row as header:

Import schema:  From sample file  None

Select file: Dim\_Country.csv

OK Back Cancel

This screenshot shows the 'Set properties' dialog for a dataset named 'Ds\_Dim\_Country\_input'. The 'Name' field is filled with the dataset's name. The 'Linked service' dropdown is set to 'Ls\_ADLS Output'. The 'File path' field contains 'output / Dimension\_Tables / File name'. The 'First row as header' checkbox is checked. Under 'Import schema', the 'From sample file' radio button is selected. A 'Select file' input field shows 'Dim\_Country.csv' with a 'Browse' button next to it. At the bottom are 'OK', 'Back', and 'Cancel' buttons.

## Creating Output Datasets

The screenshot shows the 'Set properties' dialog for creating a new dataset named 'Ds\_Dim\_Player\_Output'. The 'Name' field is filled with 'Ds\_Dim\_Player\_Output'. The 'Linked service' dropdown is set to 'AzureSqlDatabase1'. The 'Table name' dropdown is set to 'dbo.Dim\_Player'. The 'Import schema' section has 'From connection/store' selected. The 'File path' section shows a preview of a CSV file with columns: ID, Name, and BirthDate. The 'Compression type' dropdown is set to 'No compression'. At the bottom right are 'OK', 'Back', and 'Cancel' buttons.

The screenshot shows the 'Set properties' dialog for creating a new dataset named 'Ds\_Dim\_Sport\_output'. The 'Name' field is filled with 'Ds\_Dim\_Sport\_output'. The 'Linked service' dropdown is set to 'AzureSqlDatabase1'. The 'Table name' dropdown is set to 'dbo.Dim\_Sport'. The 'Import schema' section has 'From connection/store' selected. The 'Table' dropdown is set to 'dbo.Dim\_Player'. At the bottom right are 'OK', 'Back', and 'Cancel' buttons.

## Create Pipelines

Microsoft Azure | Data Factory > projecttestdf

Search factory and documentation

Preview experience Off

Activities

- Copy data Dim Player
- Copy data Dim Sport
- Copy data Dim Event
- Copy data Dim Medal

Copy data

General Source Sink Mapping Settings User properties

Write batch timeout: e.g. 00:30:00

Write batch size

Max concurrent connections

Disable performance metrics analytics

Microsoft Azure | Data Factory > projecttestdf

Search factory and documentation

Preview experience Off

Factory validation output

Your factory has been validated.

No errors were found.

Close

## Publish all

Microsoft Azure | Data Factory > projecttestdf

Search factory and documentation

Publishing 17

Publish all

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

Pending changes (17)

NAME	CHANGE	EXISTING
<input type="checkbox"/> Load_data_sql	(New)	-
<input type="checkbox"/> Ds_Dim_Player_input	(New)	-
<input type="checkbox"/> Ds_Dim_Sport_input	(New)	-
<input type="checkbox"/> Ds_Dim_Olympic_input	(New)	-
<input type="checkbox"/> Ds_Dim_Medal_input	(New)	-
<input type="checkbox"/> Ds_Dim_Event_input	(New)	-
<input type="checkbox"/> Ds_dim_Country_input	(New)	-
<input type="checkbox"/> Ds_Fact_Country_input	(New)	-

[Publish](#) [Cancel](#)

The screenshot shows the Microsoft Azure Data Factory interface. On the left, the 'Factory Resources' sidebar lists 'Pipelines' (copy\_blob\_ADLS, load\_data\_sql) and 'Datasets' (Ds\_dim\_Country\_input, Ds\_dim\_Country\_output, Ds\_Dim\_Event\_input, Ds\_Dim\_Event\_output, Ds\_Dim\_Medal\_input, Ds\_Dim\_Medal\_output, Ds\_Dim\_Olympic\_input). In the center, a pipeline named 'Ds\_dim\_Country\_out...' is selected for validation. The 'Sink' tab is active, showing settings for 'Sink dataset' (selected), 'Write behavior' (Insert), 'Bulk insert table lock' (Yes), and 'Table option' (User). Notifications on the right show a completed publishing task and two pending deletions for datasets. A 'Close' button is at the bottom right.

## Verify it in SQL Code Editor

The screenshot shows the Microsoft Azure SQL Database Query editor. The left sidebar includes 'Overview', 'Activity log', 'Tags', 'Diagnose and solve problems', 'Query editor (preview)' (which is selected), 'Mirror database in Fabric (preview)', 'Resource visualizer', 'Settings' (Compute + storage, Connection strings), and 'Properties'. The main area displays a query window with tabs for 'Query 10' through 'Query 16'. The 'Query 16' tab is active, containing the query: 'SELECT TOP (1000) \* FROM [dbo].[Dim\_Country]'. The results pane shows a single row: Country\_ID 0, Country\_Code AUT, and Country\_Name Austria. A message at the bottom says 'Query succeeded | 0s'.

## Download the CSV Files by performing the Report Generation Queries