

Data Ingestion and Loading in Azure Synapse Pipeline

Design and implement a Synapse Pipeline that ingests data from an Azure SQL Database into Azure Data Lake Storage (ADLS), followed by loading the data into a Synapse Dedicated SQL Pool for analysis and processing.

Step 1: Open Azure Portal

Step 2: Go to search and type “Azure Synapse Analytics” and select “create”

Step 3: **Basics**

- Resource group - Create New - “synapseresource”
- Workspace name: “synapse1ws-rd”
- Region: “Central US”
- Select Data Lake Storage Gen2: “From Subscription”
- Account name - create new - “accountds”
- File system name - create new - “filesystemds”
- Next Security.

Create Synapse workspace

* Basics * Security Networking Tags Review + create

Create a Synapse workspace to develop an enterprise analytics solution in just a few clicks.

Project details

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

Subscription * ⓘ Azure subscription 1 ✓

Resource group * ⓘ (New) synapseresource ✓
[Create new](#)

Managed resource group ⓘ Enter managed resource group name ✓

Workspace details

Name your workspace, select a location, and choose a primary Data Lake Storage Gen2 file system to serve as the default location for logs and job output.

Workspace name * synapse1ws-rd ✓

Region * Central US ✓

Select Data Lake Storage Gen2 * ⓘ ☒ From subscription ☐ Manually via URL

Account name * ⓘ (New) accountds ✓
[Create new](#)

File system name * (New) filesystemds ✓
[Create new](#)

☒ Assign myself the Storage Blob Data Contributor role on the Data Lake Storage Gen2 account to interactively query it in the workspace.

i We will automatically grant the workspace identity data access to the specified Data Lake Storage Gen2 account, using the [Storage Blob Data Contributor](#) role. To enable other users to use this storage account after you create your workspace, perform these tasks:

[Review + create](#) < Previous Next: Security >

Step 4: **Security**

- Authentication method: “Use both local and Microsoft Entra ID authentication”
- SQL Server admin login: “sqladminuser”
- Password: “*****”
- Confirm Password: “*****”

* Basics * Security Networking Tags Review + create

Configure security options for your workspace.

Authentication

Choose the authentication method for access to workspace resources such as SQL pools. The authentication method can be changed later on. [Learn more](#)

Authentication method ⓘ

- ☒ Use both local and Microsoft Entra ID authentication
☐ Use only Microsoft Entra ID authentication

SQL Server admin login * ⓘ

sqladminuser

SQL Password ⓘ

..... ✓

Confirm password

..... ✓

- Next Networking (no changes - keep it default)
- Next Tags (no changes - keep it default)
- Next Review + create
- Select “create”
- Deployment is in process.
- Deployment is successful.

The screenshot shows the 'Overview' page of an Azure Synapse Analytics deployment. The title bar indicates the deployment name 'Microsoft.Azure.SynapseAnalytics-20240916120851'. The left sidebar contains navigation links: Overview (selected), Inputs, Outputs, and Template. The main content area displays a green checkmark icon and the message 'Your deployment is complete'. Below this, deployment details are listed: Deployment name, Subscription, and Resource group. A 'Next steps' section includes a 'Go to resource group' button. At the bottom, there is a 'Give feedback' link. The right sidebar contains several informational cards: 'Cost management', 'Microsoft Defender for Cloud', and 'Free Microsoft tutorials'.

Step 5: Select “Go to Resource Group”

- You will be in “synapseresource”
- Storage account: “accountds” and Synapse Workspace: “synapse1ws-rd” is created and will be seen.

Home > Resource groups >

Resource groups

Default Directory

+ Create Manage view ...

Filter for any field...

Name ↑

- my-rg1
- Projectweek2
- synapserserresource**
- synapseworkspace-managedrg-22a3cd...
- synapseworkspace-managedrg-6bb67...

synapserserresource

Subscription (mouse) > Azure subscription 1

Subscription ID : 43b9891c-2464-4632-80b4-705b996e9216

Tags (edit) : Add tags

Deployments : 2 Succeeded

Location : Central US

Essentials

Activity log

Access control (IAM)

Tags

Resource visualizer

Events

Settings

Cost Management

Monitoring

Automation

Help

Resources Recommendations

Filter for any field... Type equals all X Location equals all X Add filter

Showing 1 to 2 of 2 records. ☐ Show hidden types

No grouping List view

Name ↑	Type ↑	Location ↑
accounts	Storage account	Central US
synapse1ws-rd	Synapse workspace	Central US

< Previous Page 1 of 1 Next >

Give feedback

Step 6: Select - storage account - “accounts” (This step is to validate the file is created in the container or not)

- Select - Data Storage
- Select - Containers
- File is created in the container with the name “filesystemds” (The file name which we have given at step 3).

Home > synapserserresource > accounts

accounts | Containers

Storage account

+ Container Change access level Restore containers Refresh Delete Give feedback

Search containers by prefix

☐ Show deleted containers

Name	Last modified	Anonymous access level	Lease state
<input type="checkbox"/> filesystemds	9/16/2024, 1:28:00 PM	Private	Available

Containers

File shares

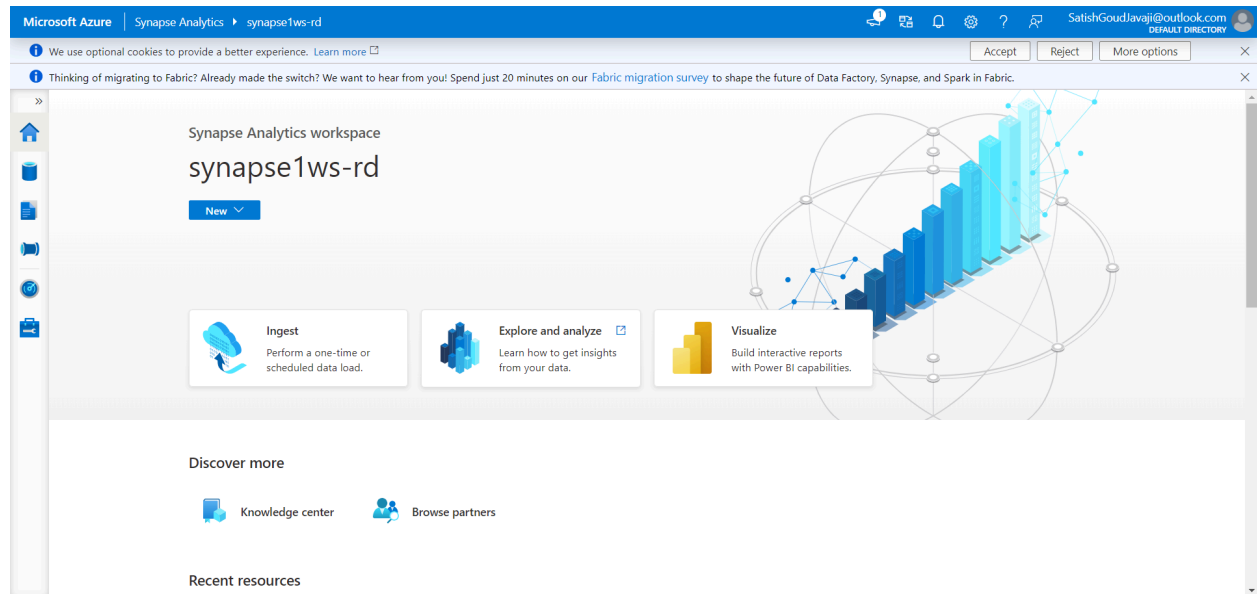
Queues

Tables

Security > networking

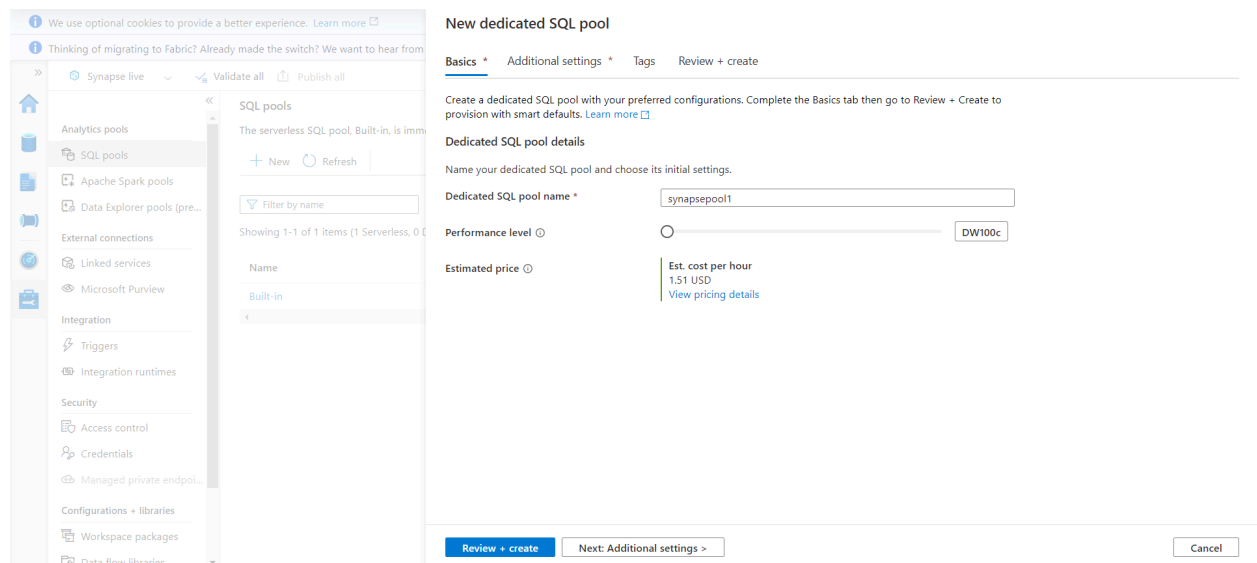
Step 7: Go to Resource group “synapserserresource”

- Select - Synapse workspace - “synapse1ws-rd”
- Select - “Open Synapse Studio”
- It will redirect to Synapse studio page as “synapse1ws-rd”



Step 8: Select - Manage - SQL pools

- In the SQL Pools - Select - "New"
- Dedicated SQL pool name: "synapsepool1"
- Performance level: Drag to min



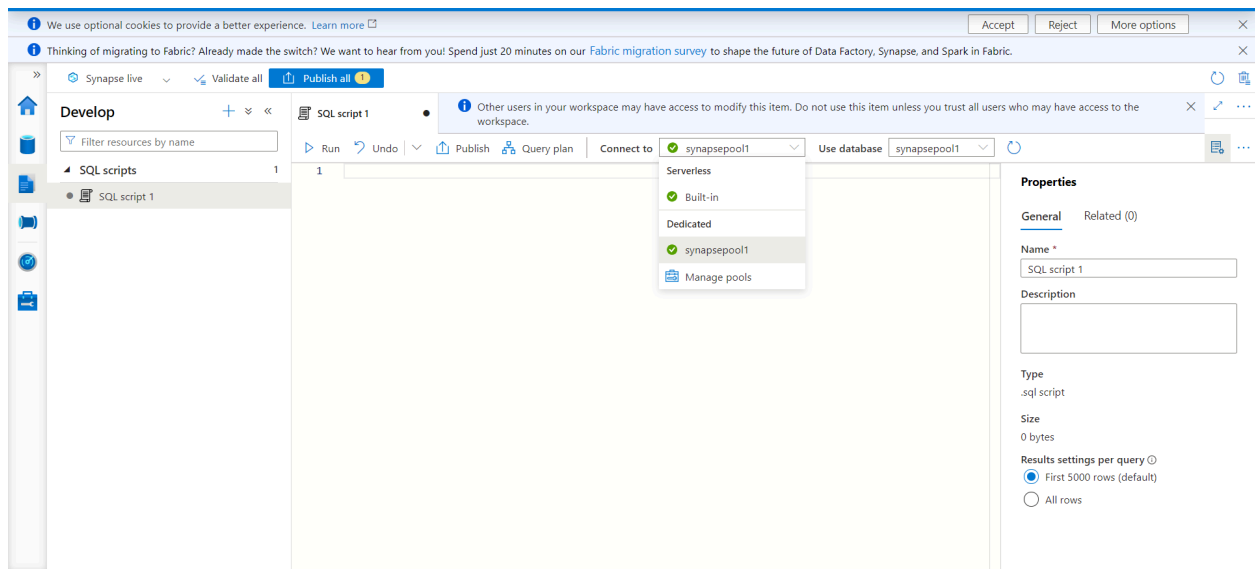
- Select "Review+Create" - Select "Create".
- Deploying in process
- Dedicated SQL pool is created "synapsepool1"

Step 9: Select - Develop

- + select: "SQL script"
- Connect to: "synapsepool1"
- `CREATE TABLE Accounts (`

```
customer_id INT,
first_name VARCHAR(50),
last_name VARCHAR(50),
city VARCHAR(50)
);
```

- Select “Run”



Step 10: Select - Manage

- Select - Linked Services

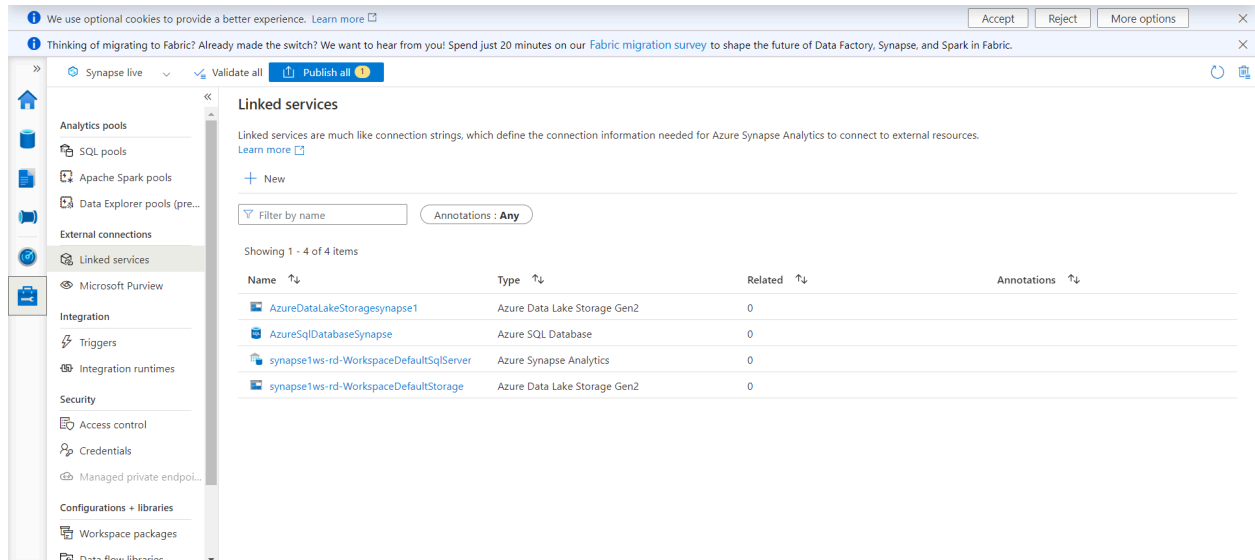
Azure SQL DataBase

- Select - “New”
- In search select “Azure SQL DataBase”
- Name: “AzureSqlDatabaseSynapse”
- Server name: “synapse1ws-rd”
- Database Name: “master”
- Authentication type: “SQL Authentication”
- User Name: “sqladminuser”
- Password: *****
- Select “Test Connection”
- Select - “Create”
- New SQL database is created “AzureSqlDatabaseSynapse”

Azure Data Lake Storage Gen2

- Select - “New”
- In search select “Azure Data Lake Storage Gen2”
- Name: “AzureDataLakeStoragesynapse1”
- Storage account name: “accountds”
- Select “Test Connection”

- Select - “Create”
- New Azure Data Lake Storage Gen2 is created “AzureDataLakeStoragesynapse1”



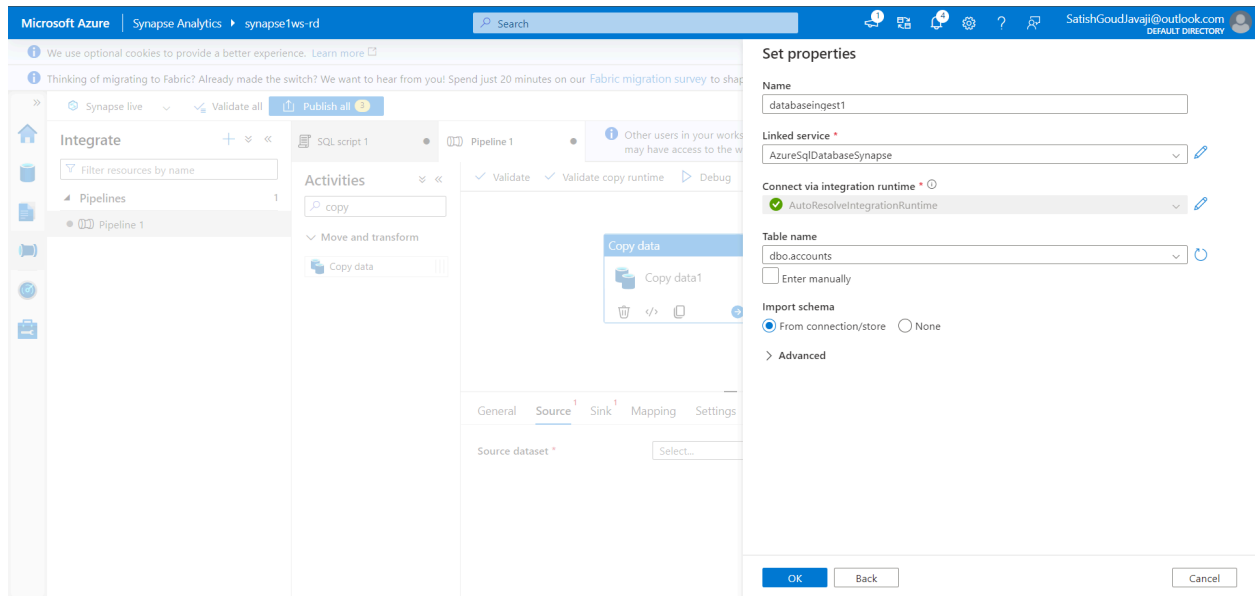
Step additional:

- Manage - Linked services
- Select - “AzureSqlDatabaseSynapse”
- Edit linked services.
- Account selection method: “From Azure subscription”
- Server name: “shreeserverproject2”
- Database name: “CustomerAccountLoanDB”
- Authentication type: “SQL Authentication”
- User name: “shreeserverproject2”
- Password: “*****”
- Select - “Test connection”
- Select - Manage - linked services
- Select - Azure Synapse analytics
- Name: “AzureSynapseAnalyticsingest”
- Account selection method: “From Azure subscription”
- Server name: “synapse1ws-rd (Synapse workspace)”
- Database name: “synapsepool1”
- User name: “sqladminuser”
- Password: “*****”
- Select - “Test connection”
- Select - Create.

Step 11: Select - Integrate

- Select + Select “Pipeline”
- Search Activity: “copy data” (type)
- Drag copy data to pipeline page.

- Select - Source
- Select - New
- Select - "Azure SQL database"
- Name: "databaseingest1"
- Linked service: "AzureSqlDatabaseSynapse"
- Table name: "dbo.Accounts"
- Select - ok



- Select - Sink
- Select - New
- Search "Azure Data Lake Gen 2"
- Select - Parquet
- Name: "datalakeingest1"
- Linked service: "AzureDataLakeStoragesynapse1"
- File path - Select file icon
- Select - "filesystemds" Select "OK"

Step 12: Drag copy data to pipeline page.

- Select - New
- Search "Azure Data Lake Gen 2"
- Select - Parquet
- Name: "datalakeingest3"
- Linked service: "AzureDataLakeStoragesynapse1"
- File path - Select file icon
- Select - "filesystemds" Select
- Select - "dbo.accounts.parquet" and select "OK"
- Select - Sink
- Select - New

- Search “Azure Synapse Analytics dedicated pool”
- Name: “Synapseanalyticsing”
- SQL Pool: “synapsepool1”
- Table name: “dbo.accounts.”
- Select - apply
- Drag copy data 1 tick mark to copy data 2
- Select copy data 2 - Select mapping - Select import schema.
- Publish all
- Validate
- Copy data 1 tick mark - Drag it to copy data 2.
- Publish all
- Debug
- Pipeline status Success

We use optional cookies to provide a better experience. [Learn more](#)

Accept Reject More options

Synapse live Validate all Publish all

Integrate

Filter resources by name

Pipelines 1

Pipeline 1

SQL script 1 Pipeline 1 datalakeingest3

Activities

copy

Move and transform

Copy data

Copy data1

Copy data2

Parameters Variables Settings Output

Pipeline run ID: 2c56695f-3d84-41f4-8af6-8d72c80c3fc1 Pipeline status Succeeded View debug run consumption

All status

Showing 1 - 2 of 2 items

Activity name	Activity status	Activity type	Run start	Duration	Integration runtime
Copy data2	Succeeded	Copy data	9/16/2024, 3:51:13 PM	23s	AutoResolveIntegrator
Copy data1	Succeeded	Copy data	9/16/2024, 3:50:57 PM	16s	AutoResolveIntegrator