

Spark SQL Homework

1. Terraform apply:

```
MINGW64:/c/Users/YuriiiHordiichuk/Desktop/m07_sparksql_python_azure-mast...
azurerm_databricks_workspace.bdcc: Still creating... [10m10s elapsed]
azurerm_databricks_workspace.bdcc: Still creating... [10m20s elapsed]
azurerm_databricks_workspace.bdcc: Still creating... [10m30s elapsed]
azurerm_databricks_workspace.bdcc: Still creating... [10m40s elapsed]
azurerm_databricks_workspace.bdcc: Still creating... [10m50s elapsed]
azurerm_databricks_workspace.bdcc: Still creating... [11m00s elapsed]
azurerm_databricks_workspace.bdcc: Still creating... [11m10s elapsed]
azurerm_databricks_workspace.bdcc: Still creating... [11m20s elapsed]
azurerm_databricks_workspace.bdcc: Still creating... [11m30s elapsed]
azurerm_databricks_workspace.bdcc: Still creating... [11m40s elapsed]
azurerm_databricks_workspace.bdcc: Still creating... [11m50s elapsed]
azurerm_databricks_workspace.bdcc: Still creating... [12m00s elapsed]
azurerm_databricks_workspace.bdcc: Still creating... [12m10s elapsed]
azurerm_databricks_workspace.bdcc: Still creating... [12m20s elapsed]
azurerm_databricks_workspace.bdcc: Still creating... [12m30s elapsed]
azurerm_databricks_workspace.bdcc: Still creating... [12m40s elapsed]
azurerm_databricks_workspace.bdcc: Still creating... [12m50s elapsed]
azurerm_databricks_workspace.bdcc: Still creating... [13m00s elapsed]
azurerm_databricks_workspace.bdcc: Still creating... [13m10s elapsed]
azurerm_databricks_workspace.bdcc: Creation complete after 13m14s [id=/subscriptions/316ab800-22b5-40ab-be41-8595de4fb2d4/resourceGroups/rg-dev-westeu...
providers/Microsoft.Databricks/workspaces/dbw-dev-westeu...
Releasing state lock. This may take a few moments...
```

Apply complete! Resources: 5 added, 0 changed, 0 destroyed.

Outputs:

```
resource_group_name = "rg-dev-westeu...

```

```
AzureAD+YuriiiHordiichuk@EPPLWROW0218 MINGW64 ~/Desktop/m07_sparksql_python_azur...
-master/terraform
$ terraform output resource_group_name
"rg-dev-westeu...

```

2. Resources created:

Name	Type	Resource Group
sf324c0c2c224e3daae70065599bb609	Virtual machine	databricks-rg-rg-dev-westeu...
dbstorage5c3l3fhjhg	Storage account	databricks-rg-rg-dev-westeu...
deepamcourseyh	Storage account	resource_main
stdevwesteuropew0oc	Storage account	rg-dev-westeu...

3. Storage account -> container show:

Home > Storage center | Blob Storage > stdevwesteuropew0oc | Containers >

data Container

Search <>

Add Directory Upload Refresh Delete Copy Paste Rename Acquire lease Break lease

Overview

Diagnose and solve problems

Access Control (IAM)

Settings

data > m07sparksq

Authentication method: Access key (Switch to Microsoft Entra user account)

Search blobs by prefix (case-sensitive)

Showing all 4 items

Name	Last modified	Access
[..]	11/26/2025, 2:09:24 PM	
expedia	11/26/2025, 2:09:32 PM	
hotel-weather	11/26/2025, 2:09:24 PM	Hot (In)
.DS_Store	11/26/2025, 2:09:24 PM	Hot (In)
DataDescription.md	11/26/2025, 2:09:24 PM	Hot (In)

4. Job set up + notification:

Jobs & Pipelines > Databricks Spark SQL Hometask ★ Lakeflow Jobs UI: ON ▾ Send feedback

Runs Tasks

No task selected Choose a task from the graph to edit its properties

+ Add task

Triggered run: 792231827362157 View run

Job details

- Job ID: B11661242669135
- Creator: fhwidw@gmail.com
- Run as: fhwidw@gmail.com
- Description: Add description

Schedule

- None
- Add trigger

Job parameters

No job parameters are defined for this job

Edit parameters

Compute

Job_cluster

Single node: Standard_Dv2s_v5 - DBR: 17.3 LTS (includes Apache Spark 4.0.0, Scala 2.13)

Configure Swap

Tags

Add tag

Job notifications

yura1234px@ On failure

Edit notifications

Duration and streaming backlog thresholds

No thresholds defined

Add metric thresholds

5. Job run show:

Jobs & Pipelines > Databricks Spark SQL Hometask ★ Lakeflow Jobs UI: ON ⚡ Send feedback

Runs Tasks

Runs

Run total duration: 5m 10s
2m 35s

Tasks
run_01_create_metadata
02_load_bronze_data
03_bronze_to_silver
04_silver_to_gold

Go to the latest successful run

Start time	Run ID	Launched	Duration	Status	Error code	Run parameters
Nov 26, 2025, 04:58 PM	315586636618358	Manually	5m 11s	Succeeded		

Started before: < Previous Next >

Job details

Job ID: 811661242669135
Creator: fwsdw@gmail.com
Run as: fwsdw@gmail.com
Description: Add description

Schedule

None
Add trigger

Job parameters

No job parameters are defined for this job
Edit parameters

Compute

Job cluster
Single node: Standard_D4ds_v5 - DBR: 17.3 LTS (includes Apache Spark 4.0.0, Scala 2.13)
Configure Swap

Tags

Add tag

Job notifications

yura7234gor@... On failure

Jobs & Pipelines > Databricks Spark SQL Hometask run Lakeflow Jobs UI: ON ⚡ Send feedback

Graph Timeline List

Delete job run Repair

Job run details

Job ID: 811661242669135
Job run ID: 315586636618358
Launched: Manually
Started: Nov 26, 2025, 04:58 PM
Ended: Nov 26, 2025, 05:03 PM
Duration: 5m 11s
Execution time: 3m 49s
Queue duration: -
Status: Succeeded
View run events

Compute

Job cluster
Single node: Standard_D4ds_v5 - Release: 17.3.1
View details Spark UI Logs Metrics

```
graph LR; A[run_01_create_metadata] --> B[02_load_bronze_data]; B --> C[03_bronze_to_silver]; C --> D[04_silver_to_gold]
```