

Spark SQL Homework

1. Terraform apply:

```
MINGW64:/c:/Users/YuriiHordiichuk/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-westeurop...
providers/Microsoft.Databricks/workspaces/dbw-dev-westeurop...w0oc]
Releasing state lock. This may take a few moments...

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

Outputs:

resource_group_name = "rg-dev-westeurop...w0oc"

AzureAD+YuriiHordiichuk@EPPLWROW0218 MINGW64 ~/Desktop/m07_sparksql_python_azure-master/terraform
$ terraform output resource_group_name

"rg-dev-westeurop...w0oc"
```

2. Resources created:

Resources			
Default Directory (ffwsdgm@gmail.onmicrosoft.com)			
Manage view Refresh Export to CSV Open query Assign tags Add to service group			
You are viewing a new version of Browse experience. Click here to access the old experience.			
Filter for any field... Subscription equals all Type equals all Resource Group equals all Location equals all Add filter			
Name	Type	Resource Group	
5f324c0c2c224e3daae70065599bb609	Virtual machine	databricks-rg-dev-westeurop...w0oc	
dbstorgen5c3l3fhyjqe	Storage account	databricks-rg-dev-westeurop...w0oc	
deepamcourseyh	Storage account	resource_main	
stdevwesteuropew0oc	Storage account	rg-dev-westeurop...w0oc	

3. Storage account -> container show:

The screenshot shows the Azure Storage Explorer interface. The breadcrumb path is Home > Storage center | Blob Storage > stdevwesteuropew0oc | Containers >. The selected container is named 'data'. The left sidebar shows the 'Overview' tab selected, with options for 'Diagnose and solve problems', 'Access Control (IAM)', and 'Settings'. The top toolbar includes actions like 'Add Directory', 'Upload', 'Refresh', 'Delete', 'Copy', 'Paste', 'Rename', 'Acquire lease', and 'Break lease'. Below the toolbar, the authentication method is set to 'Access key' with a link to 'Switch to Microsoft Entra user account'. A search bar is present with the placeholder text 'Search blobs by prefix (case-sensitive)'. The main area displays 'Showing all 4 items' in a table:

<input type="checkbox"/>	Name	Last modified	Access
<input type="checkbox"/>	..		
<input type="checkbox"/>	expedia	11/26/2025, 2:09:24 PM	
<input type="checkbox"/>	hotel-weather	11/26/2025, 2:09:32 PM	
<input type="checkbox"/>	.DS_Store	11/26/2025, 2:09:24 PM	Hot (l...
<input type="checkbox"/>	DataDescription.md	11/26/2025, 2:09:24 PM	Hot (l...

4. Job set up + notification:

The screenshot shows the Databricks Jobs UI for a job named 'Databricks Spark SQL Hometask'. The job is currently in a 'Triggered run' state with ID 792231827362137. The main area displays a task graph with four tasks: 'run_01_create_metadata', '02_load_bronze_data', '03_bronze_to_silver', and '04_silver_to_gold'. Each task is associated with a specific job cluster. A '+ Add task' button is visible below the graph. The right sidebar contains configuration options for the job:

- Job details:** Job ID 811661242669135, Creator fhwswd@gmail.com, Run as fhwswd@gmail.com, and an 'Add description' button.
- Schedule:** Set to 'None' with an 'Add trigger' button.
- Job parameters:** A message stating 'No job parameters are defined for this job' with an 'Edit parameters' button.
- Compute:** The 'Job cluster' is set to 'Single node: Standard D4ds_v5 - DBR: 17.3 LTS (includes Apache Spark 4.0.0, Scala 2.13)'. 'Configure' and 'Swap' buttons are available.
- Tags:** An 'Add tag' button.
- Job notifications:** A notification is configured for 'yura1234gor@...' on 'On failure', with an 'Edit notifications' button.

5. Job run show:

Jobs & Pipelines >

Databricks Spark SQL Hometask

Lakeflow Jobs UI: ON

Send feedback

Runs

Tasks

Runs

Started before

Previous

Next

Run total duration

Nov 26

5m 10s

2m 35s

run_01_create_metadata

02_load_bronze_data

03_bronze_to_silver

04_silver_to_gold

Go to the latest successful run

Cancel runs

Start time	Run ID	Launched	Duration	Status	Error code	Run parameters
Nov 26, 2025, 04:58 PM	315586636618350	Manually	5m 11s	Succeeded		

Job details

Job ID: 811661242669135

Creator: ffrwdw@gmail.com

Run as: ffrwdw@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

yura1234gor@... On failure

Jobs & Pipelines >

Databricks Spark SQL Hometask run

Lakeflow Jobs UI: ON

Send feedback

Graph

Timeline

List

run_01_create_metadata

02_load_bronze_data

03_bronze_to_silver

04_silver_to_gold

Job run details

Job ID: 811661242669135

Job run ID: 315586636618350

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