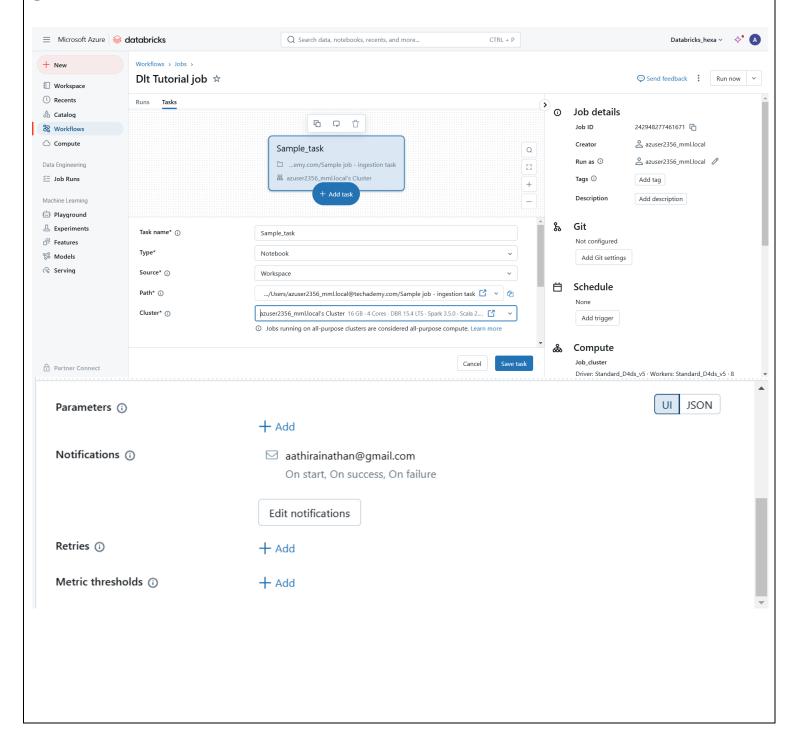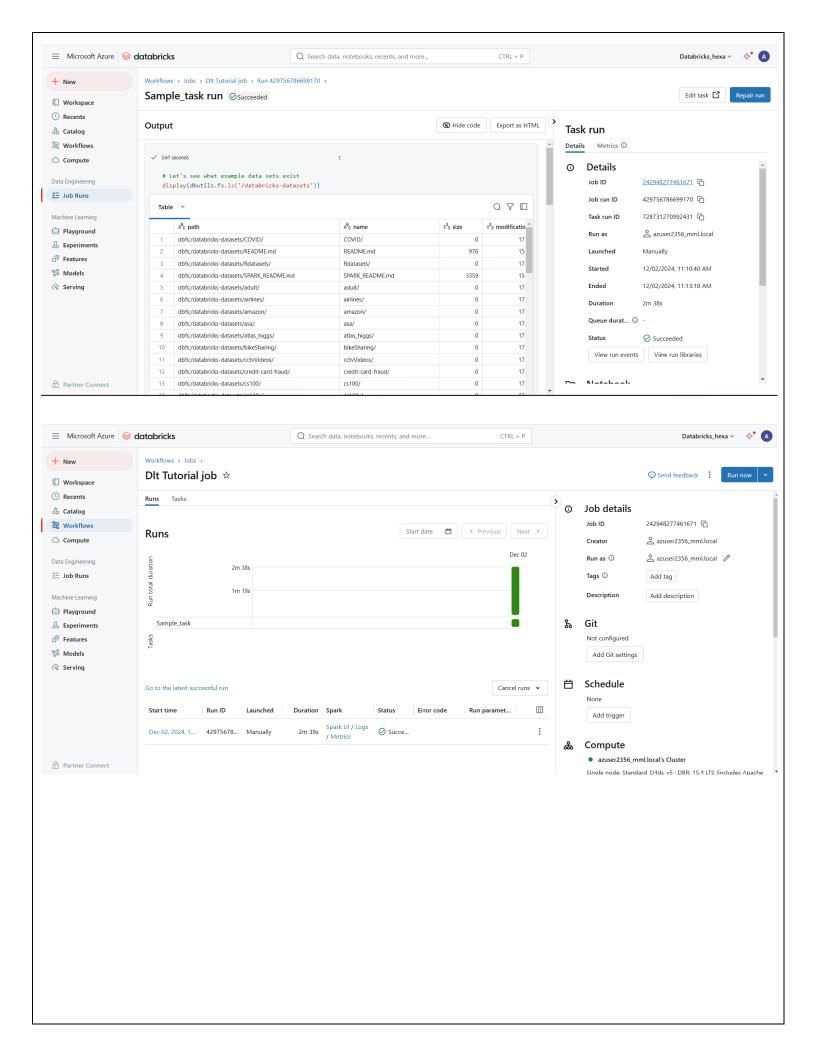# Creating Jobs in Databricks

Name: Aathirainathan P

Date: 02-12-2024

---

1. Run job on ingestion of load tutorial cluster and configure parameters as given.

New

Workspace
Recents
Catalog
Workflows
Compute

Data Engineering
Job Runs

Machine Learning
Playground
Experiments
Features
Models
Serving

Partner Connect

Workflows › Jobs › Dlt Tutorial job › Run 429756786699170 ›

Sample_task run ✓ Succeeded

Edit task ↗   Repair run

Output

Hide code   Export as HTML   ›

✓ 3.67 seconds   1

```
# Let's see what example data sets exist
display(dbutils.fs.ls('/databricks-datasets'))
```

Table ∨   🔍 ▽ ▣

	path	name	size	modificatio
1	dbfs:/databricks-datasets/COVID/	COVID/	0	17
2	dbfs:/databricks-datasets/README.md	README.md	976	15
3	dbfs:/databricks-datasets/Rdatasets/	Rdatasets/	0	17
4	dbfs:/databricks-datasets/SPARK_README.md	SPARK_README.md	3359	15
5	dbfs:/databricks-datasets/adult/	adult/	0	17
6	dbfs:/databricks-datasets/airlines/	airlines/	0	17
7	dbfs:/databricks-datasets/amazon/	amazon/	0	17
8	dbfs:/databricks-datasets/asa/	asa/	0	17
9	dbfs:/databricks-datasets/atlas_higgs/	atlas_higgs/	0	17
10	dbfs:/databricks-datasets/bikeSharing/	bikeSharing/	0	17
11	dbfs:/databricks-datasets/cctvVideos/	cctvVideos/	0	17
12	dbfs:/databricks-datasets/credit-card-fraud/	credit-card-fraud/	0	17
13	dbfs:/databricks-datasets/cs100/	cs100/	0	17

Task run

Details   Metrics ⓘ

ⓘ **Details**

Job ID	242948277461671 ⧉
Job run ID	429756786699170 ⧉
Task run ID	728731270992431 ⧉
Run as	👤 azuser2356_mml.local
Launched	Manually
Started	12/02/2024, 11:10:40 AM
Ended	12/02/2024, 11:13:18 AM
Duration	2m 38s
Queue durat... ⓘ	-
Status	✓ Succeeded

View run events   View run libraries

🗂 **Notebook**

Workflows › Jobs ›

Dlt Tutorial job ☆

🗨 Send feedback   ⋮   Run now ∨

Runs   Tasks

Runs

Start date 📅   ‹ Previous   Next ›

Dec 02

2m 38s

1m 19s

Sample_task

Go to the latest successful run   Cancel runs ∨

Start time	Run ID	Launched	Duration	Spark	Status	Error code	Run paramet...	▣
Dec 02, 2024, 1...	42975678...	Manually	2m 39s	Spark UI / Logs / Metrics	✓ Succe...			⋮

ⓘ **Job details**

Job ID	242948277461671 ⧉
Creator	👤 azuser2356_mml.local
Run as ⓘ	👤 azuser2356_mml.local ✎
Tags ⓘ	Add tag
Description	Add description

⅄ **Git**

Not configured

Add Git settings

📅 **Schedule**

None

Add trigger

⛛ **Compute**

● azuser2356_mml.local's Cluster

Single node: Standard_D4ds_v5 · DBR: 15.4 LTS (includes Apache

2. Run job on already done notebooks of anydelta or pyspark programs:

adb-2455489288945509.9.azuredatabricks.net/jobs/553150298936508/new-task?o=2455489288945509

Microsoft Azure · databricks

Search data, notebooks, recents, and more...   CTRL + P

Databricks_hexa

+ New

Workflows › Jobs ›

New Job 2024-12-02 11:19:18 ☆

⊘ Send feedback   ⋮   Run now ▾

Runs   **Tasks**

□ Workspace
⊘ Recents
⊕ Catalog
⊗ **Workflows**
☁ Compute

Custom_task
📁 ... Table with Existing data in Databricks
🖥 azuser2356_mml.local's Cluster

Data Engineering
⩴ Job Runs

Machine Learning
▱ Playground
⊿ Experiments
▱ Features
⊗ Models
⊘ Serving

Task name* ⓘ	Custom_task
Type*	Notebook ▾
Source* ⓘ	Workspace ▾
Path* ⓘ	...356_mml.local@techademy.com/Delta Table with Existing data in Databricks ▾ ⧉
Cluster* ⓘ	azuser2356_mml.local's Cluster 16 GB · 4 Cores · DBR 15.4 LTS · Spark 3.5.0 · Scala 2.... ▾

ⓘ Jobs running on all-purpose clusters are considered all-purpose compute. Learn more

🔒 Partner Connect

Cancel   **Create task**

ⓘ **Job details**
Job ID   553150298936508 ⧉
Creator   👤 azuser2356_mml.local
Run as ⓘ   👤 azuser2356_mml.local ✎
Tags ⓘ   Add tag
Description   Add description

📅 **Schedule**
None
Add trigger

⟨⟩ **Job parameters** ⓘ
No job parameters are defined for this job
Edit parameters

🔔 **Job notifications** ⓘ
No notifications
Edit notifications

---

adb-2455489288945509.9.azuredatabricks.net/jobs/553150298936508/runs/1106785774045511?o=2455489288945509

Microsoft Azure · databricks

Search data, notebooks, recents, and more...   CTRL + P

Databricks_hexa

+ New

Workflows › Jobs › New Job 2024-12-02 11:19:18 › Run 1106785774045511 ›

Custom_task run ⊘Succeeded

Edit task ↗   Repair run

□ Workspace
⊘ Recents
⊕ Catalog
⊗ Workflows
☁ Compute

Data Engineering
⩴ **Job Runs**

Output

👁 Hide code   Export as HTML

✓ 2.78 seconds   1

```python
from pyspark.sql import SparkSession

spark = SparkSession.builder.appName("Dlt example").getOrCreate()

# Reading source file using the Delta format and writing to the destination path in Delta format
spark.read.format("delta").load("dbfs:/user/hive/warehouse/export").write.format("delta").mode
("overwrite").save("/FileStore/tables/delta_train/")
```

✓ 0.69 seconds   2

```python
#Below we are listing the data in destination path
display(dbutils.fs.ls("/FileStore/tables/delta_train/"))
```

Table ▾

	ᴬᵇ𝒸 path	ᴬᵇ𝒸 name
1	dbfs:/FileStore/tables/delta_train/_delta_log/	_delta_log/
2	dbfs:/FileStore/tables/delta_train/part-00000-12cd9e41-0c83-476b-bf80-1f0d3036b0fa-c000.snappy.parquet	part-00000-12cd9e41-0c83-

Machine Learning
▱ Playground
⊿ Experiments
▱ Features
⊗ Models
⊘ Serving

🔒 Partner Connect

Task run

Details   Metrics

ⓘ **Details**
Job ID   553150298936508 ⧉
Job run ID   1106785774045511 ⧉
Task run ID   1085450325735405 ⧉
Run as   👤 azuser2356_mml.local
Launched   Manually
Started   12/02/2024, 11:20:50 AM
Ended   12/02/2024, 11:21:02 AM
Duration   12s
Queue durat... ⓘ   -
Status   ⊘ Succeeded

View run events   View run libraries

Notebook

Search data, notebooks, recents, and more... CTRL + P

Databricks_hexa

+ New

Workspace
Recents
Catalog
Workflows
Compute

Data Engineering
Job Runs

Machine Learning
Playground
Experiments
Features
Models
Serving

Partner Connect

Workflows › Jobs ›

New Job 2024-12-02 11:19:18 ☆

Send feedback

Run now

Runs   Tasks

Runs

Start date | ‹ Previous | Next ›

Dec 02

Run total duration

12s

6s

Custom_task

Tasks

Go to the latest successful run

Cancel runs

Start time	Run ID	Launched	Duration	Spark	Status	Error code	Run paramet...	
Dec 02, 2024, 1...	11067857...	Manually	13s	Spark UI / Logs / Metrics	⊘ Succe...			⋮

ⓘ Job details

Job ID   553150298936508

Creator   👤 azuser2356_mml.local

Run as ⓘ   👤 azuser2356_mml.local ✎

Tags ⓘ   Add tag

Description   Add description

Git

Not configured

Add Git settings

Schedule

None

Add trigger

Compute

● azuser2356_mml.local's Cluster

Single node: Standard_D4ds_v5 · DBR: 15.4 LTS (includes Apache

Submitted by:
Aathirainathan P