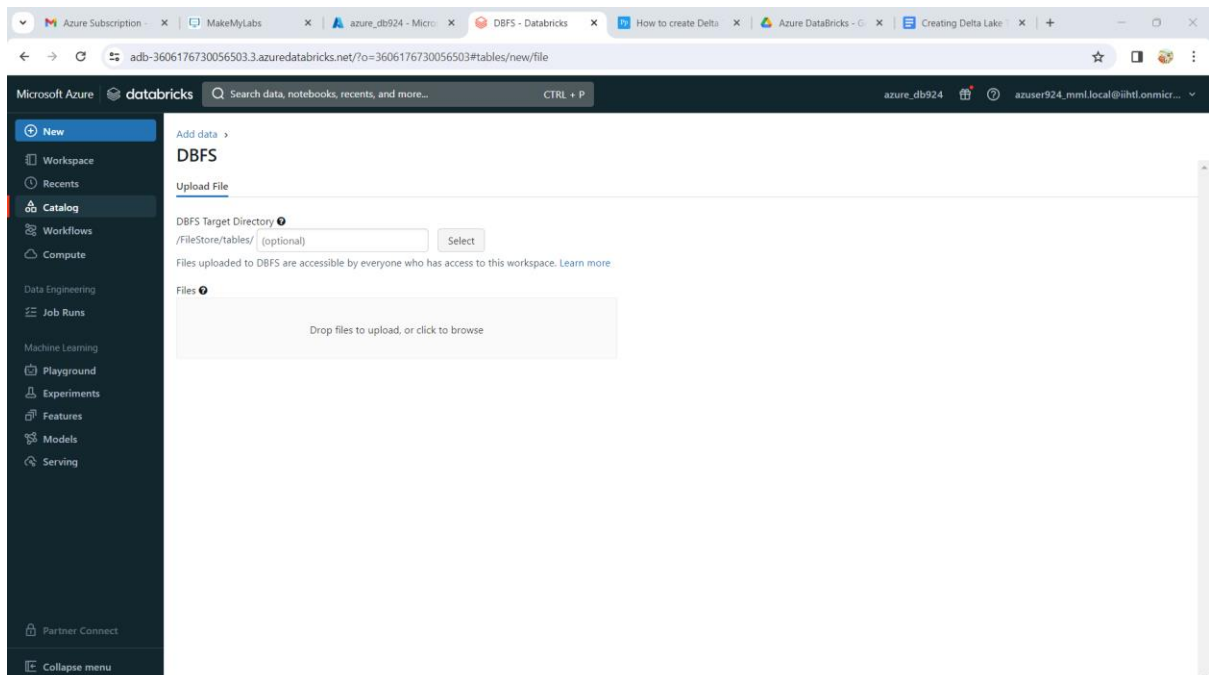
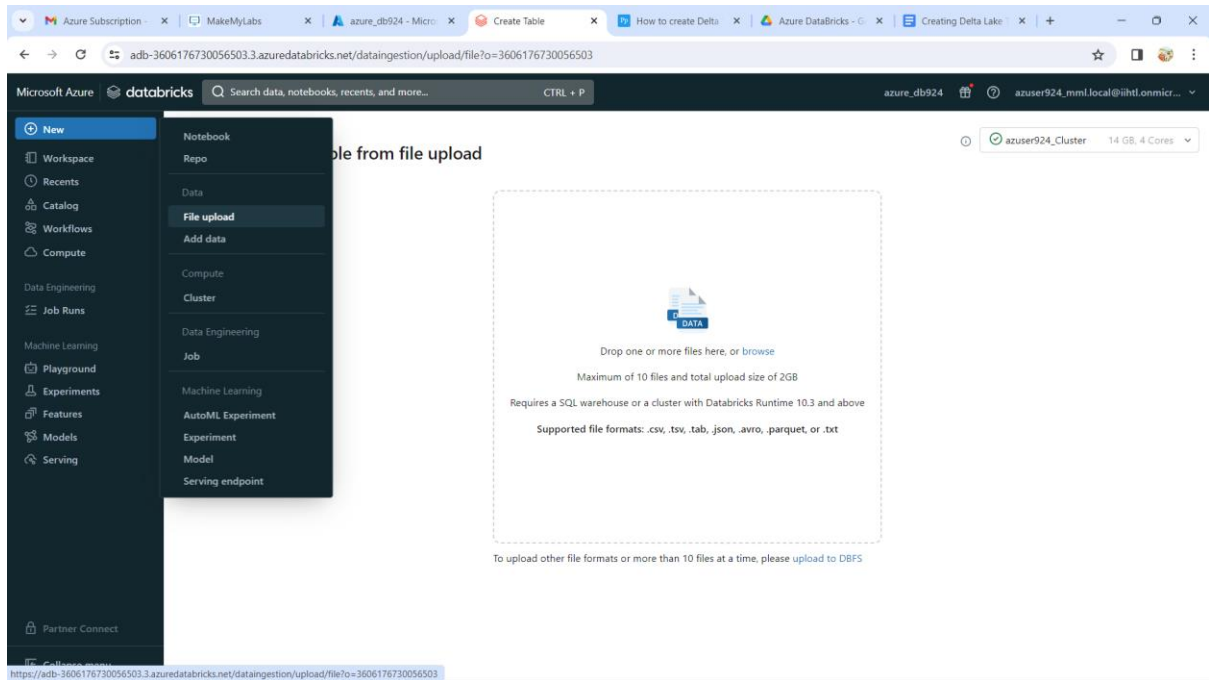


ASSESSMENT

NAME-SWARNA LATHA GUVVALA

DATE-02/01/2024

CREATING A DELTA TABLE:



Microsoft Azure databricks Search data, notebooks, recents, and more... CTRL + P azure_db924 azuser924_mml.local@ihlt.onmicr...

New Workspace Recents Catalog Workflows Compute Data Engineering Job Runs Machine Learning Playground Experiments Features Models Serving

Add data > DBFS Upload File

DBFS Target Directory Select

Files uploaded to DBFS are accessible by everyone who has access to this workspace. Learn more.

Files 0.2 KB Remove file

✓ File uploaded to /FileStore/tables/small_zipcode.csv

Create Table with UI Create Table in Notebook

Microsoft Azure databricks Search data, notebooks, recents, and more... CTRL + P azure_db924 azuser924_mml.local@ihlt.onmicr...

New Workspace Recents Catalog Workflows Compute Data Engineering Job Runs Machine Learning Playground Experiments Features Models Serving

2024-01-03 - DBFS Example Python ☆ File Edit View Run Help Last edit was 1 minute ago Provide feedback

Run all azuser924_Cluster Schedule Share

```
13 .option("header", first_row_is_header) \
14 .option("sep", delimiter) \
15 .load(file_location)
16
17 display(df)
```

(2) Spark Jobs

df: pyspark.sql.dataframe.DataFrame = [c0: string, c1: string ... 4 more fields]

	_c0	_c1	_c2	_c3	_c4	_c5
1	id	zipcode	type	city	state	population
2	1	704	STANDARD	null	PR	30100
3	2	704	null	PASEO COSTA DEL SUR	PR	null
4	3	709	null	BDA SAN LUIS	PR	3700
5	4	76166	UNIQUE	CINGULAR WIRELESS	TX	84000
6	5	76177	STANDARD	null	TX	null

6 rows | 1.49 seconds runtime Refreshed now

Command took 1.49 seconds -- by azuser924_mml.local@ihlt.onmicrosoft.com at 1/3/2024, 10:29:39 AM on azuser924_Cluster

```
1 # Create a view or table
2
3 temp_table_name = "small_zipcode_csv"
4
5 df.createOrReplaceTempView(temp_table_name)
```

Microsoft Azure databricks Search data, notebooks, recents, and more... CTRL + P azure_db924 azuser924_mml.local@ih1t1.onmicr...

New Workspace Recents Catalog Workflows Compute Data Engineering Job Runs Machine Learning Playground Experiments Features Models Serving Partner Connect Collapse menu

Add data > Create or modify table from file upload

Salary_Data.csv uploaded 454,008 Create new table

Preview mode Catalog Schema Table name

hive_metastore Select schema salary_data Advanced attributes

Create table Cancel

Microsoft Azure databricks Search data, notebooks, recents, and more... CTRL + P azure_db924 azuser924_mml.local@ih1t1.onmicr...

New Workspace Recents Catalog Workflows Compute Data Engineering Job Runs Machine Learning Playground Experiments Features Models Serving Partner Connect Collapse menu

Explore hive_metastore.default.salary_data Python ☆

File Edit View Run Help Last edit was now Provide feedback Run all azuser924_Cluster Schedule Share

```
1 %sql
2 SELECT * FROM 'hive_metastore`.`default`.`salary_data`;
```

(1) Spark Jobs

_sqldf: pyspark.sql.dataframe.DataFrame = [YearsExperience: double, Salary: double]

	YearsExperience	Salary
1	1.1	39343
2	1.3	46205
3	1.5	37731
4	2	43525
5	2.2	39891
6	2.9	56642
7	3	60150

30 rows | 2.29 seconds runtime Refreshed now

This result is stored as PySpark data frame _sqldf and in the IPython output cache as Out[1]. Learn more

Command took 2.29 seconds -- by azuser924_mml.local@ih1t1.onmicrosoft.com at 1/3/2024, 18:19:17 AM on azuser924_Cluster

Shift+Enter to run
Shift+Ctrl+Enter to run selected text