

Azure Data-Bricks:

The screenshot shows the Microsoft Azure Databricks interface. The left sidebar contains navigation options: New, Workspace, Recents, Catalog, Workflows, Compute, SQL (SQL Editor, Queries, Dashboards, Alerts, Query History, SQL Warehouses), Data Engineering (Job Runs, Data Ingestion, Delta Live Tables), Machine Learning, and Experiments. The main content area is titled 'Settings' and 'Developer'. The 'Developer' section includes 'Access tokens', 'SQL query snippets', 'Editor settings' (General), 'Notebook Notifications' (On), and 'Spark tips' (On). A notification at the top right states 'Notebook detached cluster not in usable state'.

Microsoft Azure databricks Search data, notebooks, recents, and more... CTRL + P casestudy1 shellunext_1693422719943@npune...

Settings

- User
- Profile
- Preferences
- Developer**
- Linked accounts
- Notifications

Developer
Manage your development settings

Access tokens
Set up secure authentication to Databricks API using access tokens [Manage](#)

SQL query snippets
Configure SQL query snippets [Manage](#)

Editor settings

General

Notebook Notifications
Controls whether browser notifications are shown for common Notebook events like when a cell is finished running. ☒

Spark tips
Enriches notebook error stack traces by displaying high-level "error hints" which explain otherwise-confusing errors. ☒

Notebook detached
cluster not in usable state

The screenshot shows the Microsoft Azure Databricks interface with a notebook titled '2023-10-05 - DBFS Example'. The notebook code reads a CSV file from DBFS and displays the results. The output is a table with 7 rows and 5 columns: _c0, _c1, _c2, _c3, and _c4. The table contains data for various car models and their specifications.

Microsoft Azure databricks Search data, notebooks, recents, and more... CTRL + P casestudy1 shellunext_1693422719943@npune...

2023-10-05 - DBFS Example Python [☆](#)
File Edit View Run Help Last edit was 8 minutes ago Provide feedback [Run all](#) [Shellunext's Cluster](#) [Schedule](#) [Share](#)

```
6 infer_schema = "false"
7 first_row_is_header = "false"
8 delimiter = ","
9
10 # The applied options are for CSV files. For other file types, these will be ignored.
11 df = spark.read.format(file_type) \
12     .option("inferSchema", infer_schema) \
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 ... 10 more fields]

Table [+](#) New result table: OFF

	_c0	_c1	_c2	_c3	_c4
1	Make	Model	Vehicle Class	Engine Size(L)	Cylinders
2	ACURA	ILX	COMPACT	2	4
3	ACURA	ILX	COMPACT	2.4	4
4	ACURA	ILX HYBRID	COMPACT	1.5	4
5	ACURA	MDX 4WD	SUV - SMALL	3.5	6
6	ACURA	RDX AWD	SUV - SMALL	3.5	6
7	ACURA	RLX	MID-SIZE	3.5	6

Microsoft Azure databricks Search data, notebooks, recents, and more... CTRL + P casestudy1 shellunext_1693422719943@npune...

2023-10-05 - DBFS Example Python ☆

File Edit View Run Help Last edit was 8 minutes ago Provide feedback Run all Shellunext's Cluster Schedule Share

1 %sql
2
3 /* Query the created temp table in a SQL cell */
4
5 select * from 'CO2_Emissions_Canada__1__csv'

▶ (1) Spark Jobs
▶ _sqlidf: pyspark.sql.dataframe.DataFrame = [_c0: string, _c1: string ... 10 more fields]

Table + New result table: OFF

	_c0	_c1	_c2	_c3	_c4
1	Make	Model	Vehicle Class	Engine Size(L)	Cylinders
2	ACURA	ILX	COMPACT	2	4
3	ACURA	ILX	COMPACT	2.4	4
4	ACURA	ILX HYBRID	COMPACT	1.5	4
5	ACURA	MDX 4WD	SUV - SMALL	3.5	6
6	ACURA	RDX AWD	SUV - SMALL	3.5	6
7	ACURA	RLX	MID-SIZE	2.5	6

7,386 rows | 1.08 seconds runtime Refreshed 6 hours ago

This result is stored as PySpark data frame _sqlidf and in the IPython output cache as Out[3]. Learn more

Command took 1.08 seconds -- by shellunext_1693422719943@npune.onmicrosoft.com at 10/5/2023, 12:26:14 PM on Shellunext's Cluster

Microsoft Azure Data Factory casestudyfactory Search factory and documentation Shellunext_1693422719943@npune.onmicrosoft.com

Your insights matter! Participate in our brief survey about our CDC top-level resource, and help us enhance your experience.

main branch Validate all Save all Publish Preview experience Off

Factory Resources

- Pipelines 3
 - pipeline1
 - pipeline2
 - pipeline3
- Change Data Capture (preview) 1
 - adfc1
- Datasets 2
 - dataflow1
- Power Query 0
- Templates 0

Activities

- Move and transform
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
 - Notebook
 - Jar
 - Python
- Data Lake Analytics
 - General
 - HDInsight
 - Iteration & conditionals
 - Machine Learning
 - Power Query

pipeline3

Save Save as template Validate Debug Add trigger

Notebook Notebook1

Properties

General Related

Name * pipeline3

Description

Annotations + New

Parameters Variables Settings Output

Pipeline run ID: cfd2f552-133d-4c33-adaf-6c3e42b09f8a View debug run consumption

Pipeline status Succeeded

All status Monitor in Azure Metrics Export to CSV

Showing 1 - 1 of 1 items

Activity name	Activity status	Activity type	Run start
Notebook1	Succeeded	Notebook	10/5/2023, 4:23:15 PM