

Azure Databricks Notebooks and Jobs

Template Version: 2.0

Introduction

During this lab, you will learn how to manage notebooks

Estimated Time

45 minutes

Objectives

At the end of this lab, you will be able to:

- Create/Delete notebooks
- Export/Import notebooks
- Attach/Detach notebooks
- Create dashboards using notebooks
- Create jobs and run notebooks via jobs on scheduled basis

Table of Contents

Lab: Manage Notebooks and Jobs

Exercise 1: Create/Delete, Export/Import, Attach/Detach notebooks

Exercise 2: Create dashboards using notebooks

Exercise 3: Create jobs to run notebooks on scheduled basis

Lab: Manage Notebooks and Jobs

During this lab, you will learn how do manage notebooks in Azure Databricks.

Exercise 1: Create/Delete, Export/Import, Attach/Detach notebooks

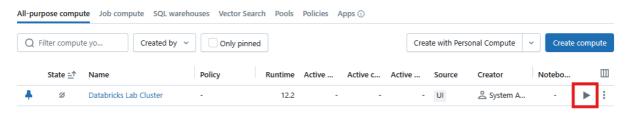
This exercise shows how to work with notebooks.

Tasks

1. Start a cluster

You will use the same cluster that was created in the previous lab. If your cluster is in Terminated state, hover over the cluster and click on the play icon to start the cluster

Compute

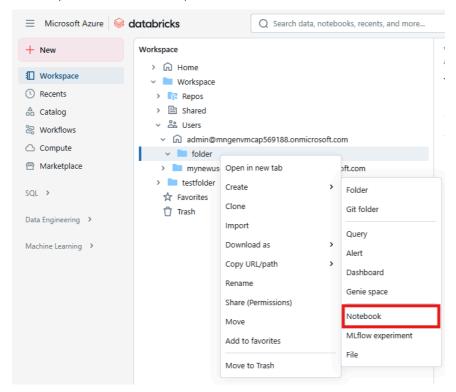


2. Create a Notebook

Once the cluster is started, click on Workspaces or on the Home button on the sidebar pane.

You can use an existing folder created in the previous lab to perform the next step.

In the Workspace or a user folder, click the dropdown and select Create -> Notebook.



The notebook will be created and opened.

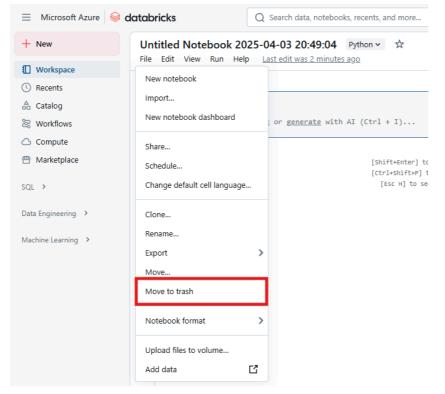


3. How to Delete a Notebook

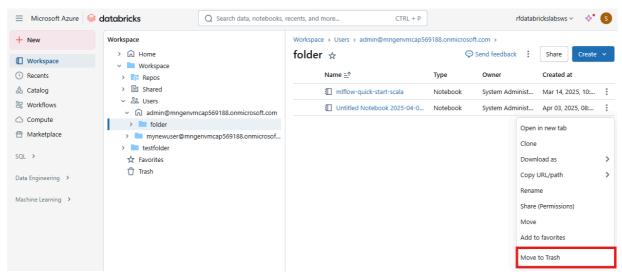
To delete a notebook, do one of the following:

 $\textbf{Note:} \ \mathsf{Do} \ \mathsf{not} \ \mathsf{perform} \ \mathsf{the} \ \mathsf{step} \ \mathsf{of} \ \mathsf{deleting} \ \mathsf{the} \ \mathsf{notebook} \ \mathsf{yet}. \ \mathsf{We} \ \mathsf{will} \ \mathsf{need} \ \mathsf{the} \ \mathsf{notebook} \ \mathsf{for} \ \mathsf{next} \ \mathsf{step}$

Click on the **File** menu while the notebook is opened and click on "**Move to Trash**"

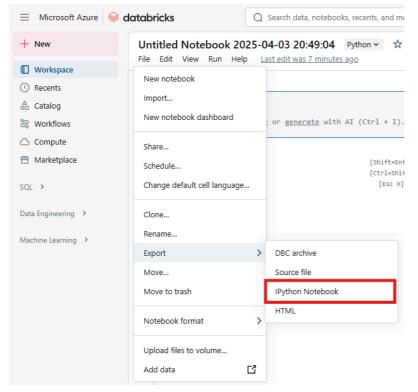


Or, select the notebook from the folder location and select Move to Trash

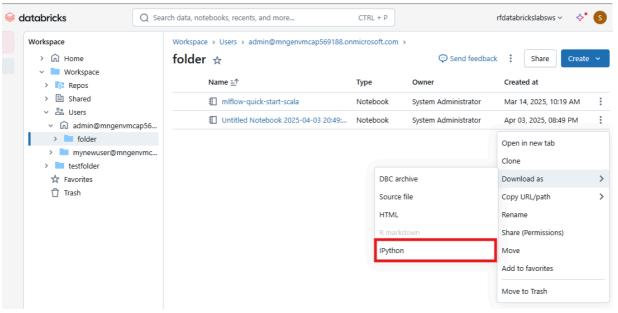


4. Export a Notebook

In the notebook toolbar, select File > Export to Ipython Notebook format and save it locally on the machine.



The export option is also available at the workspace level while selecting the notebook as shown below



(Optional)

You may also experiment with saving it to other formats

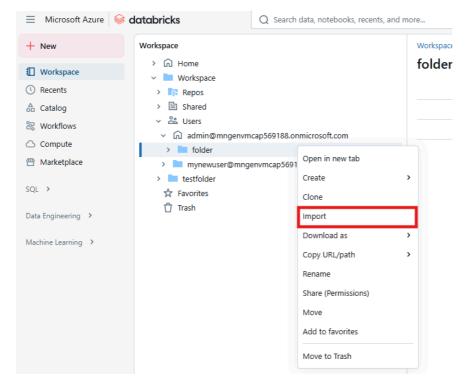
Use the steps in the **Delete** notebook task above to delete the notebook that was exported in this step

5. Import a Notebook

Click the **Workspace** button or the Home button in the sidebar. Do one of the following:

Next to any folder, click the on the right side of the text and select $\mbox{\bf Import}.$

In the Workspace or a user folder, click and select Import.



Browse to the file that was exported in the earlier step. Click Import.

(Optional)

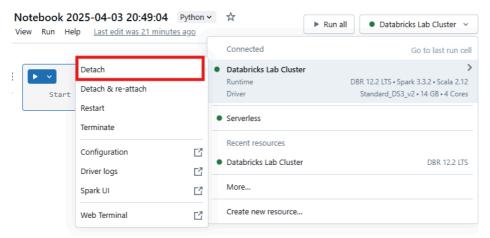
Note that you can import notebooks using a URL too. You can experiment by importing example notebook from the URL link below:

 $https://docs.databricks.com/_extras/notebooks/source/mlflow/mlflow-quick-start-scala.html \\$

1. Detach a notebook from a cluster

To detach a notebook from a cluster:

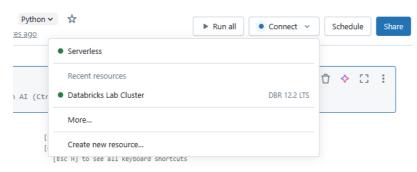
In the notebook toolbar, click the cluster name on the top-right and click $\ensuremath{\mathsf{Detach}}$



2. Attach a notebook to a cluster

To attach a notebook to a cluster:

In the notebook toolbar, click Connect and then select the name of a cluster you want to attach to



Exercise 1 has been completed

Exercise 2: Create dashboards using notebooks

This exercise shows how to create dashboards using notebooks and how to run commands within notebooks

Tasks

1. Use a notebook

You will use the DemoDashboard.ipynb for completing the steps in this lab

Follow the steps in Exercise 1, Task 4 to import the DemoDashboard.ipynb file and attach it to your cluster

Follow the documentation in the notebook to finish the rest of the lab. As you read through the lab, start running the cells to produce and view the output

Exercise 2 has been completed

Exercise 3: Create jobs to run notebooks on scheduled basis

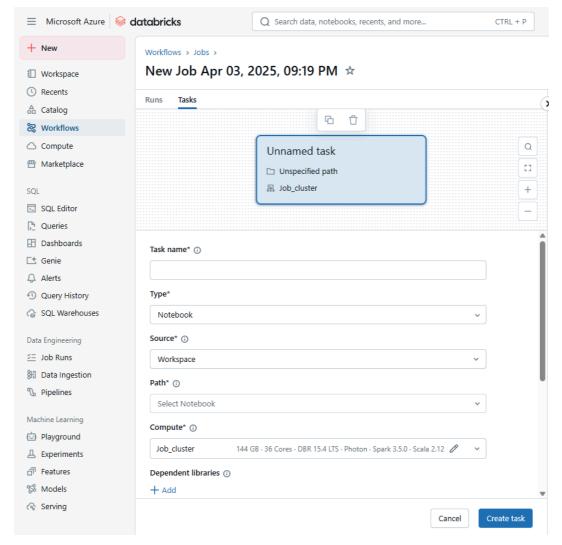
This exercise shows how to create jobs to run notebooks on scheduled basis

Tasks

1. Create a Job Click on the **Workflows** button 📚 **Workflows** on the sidebar panel. Workflows Send feedback Job runs Q Filter jobs Owned by me Accessible by me Favorites Tags v Load tutorial Name <u>=</u>↑ Tags Created by Trigger Recent runs

No jobs found.

Click + Create Job. The tasks detail page displays.

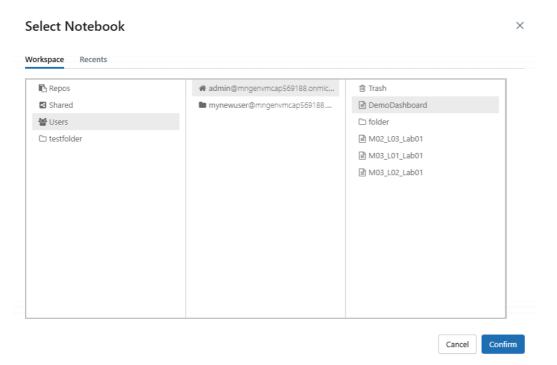


Enter a name in the text field with the placeholder text **Add a name for your job**.

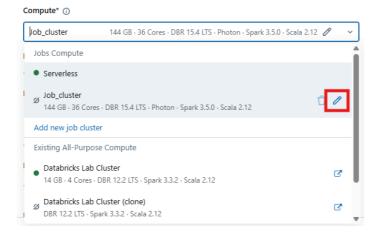
Specify the Task properties:

a. For the Type, select Notebook and then Click **Select notebook**.

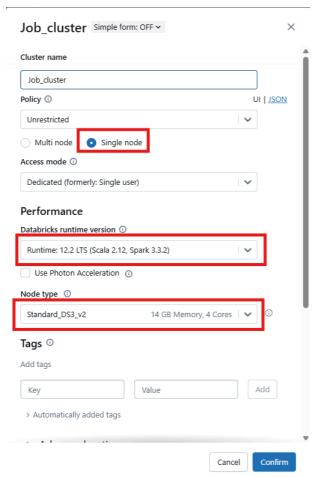
In the Select Notebook, click on the folder you imported the DemoDashboard.ipynb notebook from the exercise 2



b. For the cluster, in the dropdown, click on the Edit button in New Job Cluster



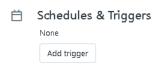
c. In Job Cluster creation form select **Single Node**, choose the same databricks runtime and Node type. Click on **Confirm** button



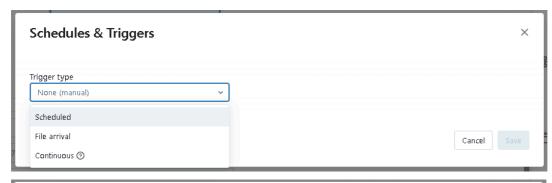
Note: There is a tradeoff between running on a new cluster and on an existing cluster. We recommend that you run on a new cluster for production-level jobs or jobs that are important to complete.

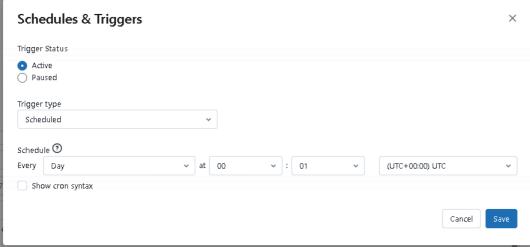
1. Schedule a Job

a. In the Job details click on **Add trigger** under the *Schedules & Triggers* section



b. Select **Scheduled** as Trigger Type and set a schedule as shown below or a scheduled as desired

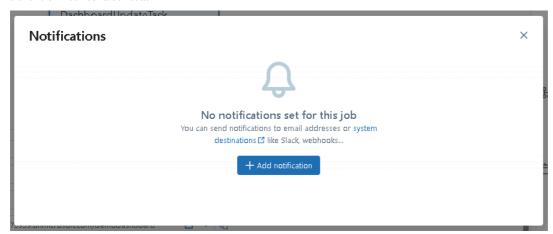




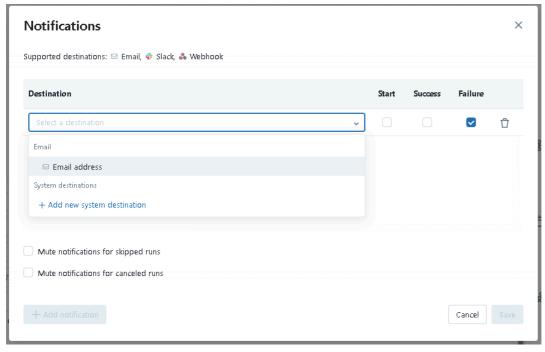
c. Click on **Edit Notifications** under the **Notifications** section



click on the + Add notification button



Select Email as destination:



You can either wait for the job to run on the selected schedule or use the Run Now to start the job immediately and monitor its progress



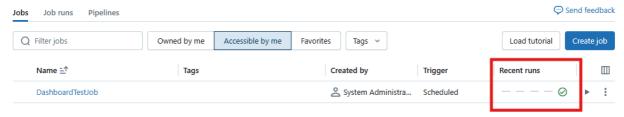
and setup a valid email to receive alerts for the job

2. View a scheduled job

Click on the **Workflows** icon on the sidebar panel

In the Workflows panel, review the job Last Run and ensure that it was successful

Workflows



Click to Succeeded in the Last Run column to view the run details

Exercise 3 has been completed