databricks 1.2 - Demo - The Lakehouse Platform

(https://databricks.com)

Demo: The Lakehouse Platform

Learning Objectives

1. Describe the navigational interface to the Databricks Lakehouse Platform, including the organization of services and capabilities on the left-side navigation bar and the availability of settings in the top-right corner.
2. Navigate throughout the Workspace.
3. Describe the Workspace as the solution for organizing assets within Databricks.
4. Perform the actions available in the Workspace.

The Landing Page

In the upper-left corner, right click on the Databricks logo and open the link in a new tab. This will allow you to keep this notebook open while navigating the Databricks workspace. What you see is the Databricks landing page. On the page, you three sections: Get Started, Recents, and Popular.

Get Started

In the Get Started section, there are quick links that will help you with some common tasks like importing data, creating a new notebook, and opening the SQL Editor.

Recents

This section provides a list of the most recently used items. It can include notebooks, queries, dashboards, and other items.

Popular

There are some items you will use more often than others. This section tracks the items that are accessed more often. This will help you quickly find your most commonly used items.

The Sidebar Menu

On the left side of your workspace, you will find the sidebar menu. There are five sections in this menu, and three of the sections can be collapsed if you do not use them often.

1. Click the "Machine Learning" heading in the sidebar menu, and it will collapse.
2. Click again, and it will expand. You can also collapse and expand the entire sidebar menu. At the bottom of the menu, you will find a menu item that provides this feature
3. Click the last menu item in the sidebar menu to either expand or collapse the sidebar menu. Note: you will need to move your mouse off the sidebar menu to see it collapse.

The Workspace Page

Click "Workspace" in the sidebar menu. The workspace is the way you organize your assets in Databricks. There are four sections: Home, Workspace, Repos, and Trash.

Home

The "Home" section is a shortcut to your user folder. This folder can also be found by clicking Workspace -> Users -> *your-username*. You can add sub-folders by clicking Add -> Folder, or you can create new assets like Queries, Alerts, Notebooks, MLFlow experiments, and others.

1. Create a new notebook in your home folder
2. Change the name of the notebook in the upper-left corner of the notebook. "Name it Getting Started."
3. Go back to your home folder by clicking Workspace -> Home. The notebook appears in your Home folder. Let's share the notebook with others.
4. At the far right of the notebook name, click the kebab menu and select "Share (Permissions)." Note that the notebook is shared with you, and the Admins group, by default.
5. Click in the input box, and select "All Users." On the right, you can select the privileges you are granting to this group. They can view, run, edit, or manage the notebook, depending on which option you choose. Note: whichever option

you choose, the group will be granted all privileges below that option (i.e., if they can edit, they can also run and view).

6. For the privilege type, select "Can View."
7. Click "Add."
8. Click the "X" to close the sharing dialog box. Let's delete the notebook.
9. Again, click the kebab menu to the far right of the notebook name. Note: there are many options in this menu (Open in new tab, Clone, Rename, etc). We are not going to discuss these in this course.
10. Select "Move to Trash."
11. Confirm that you wish to move the notebook to the trash by selecting "Confirm and move to Trash." The notebook is moved to the trash. If you wish to undo this action or permantly delete the item, click "Trash" on the left side of the page.
12. Click the kebab menu to the far right of the notebook name. There are two options: "Permanently delete" and "Restore."

Workspace

The Workspace section allows you to organize your assets outside of your home folder and outside of any repositories (discussed next). You can use this section to organize subfolders that are shared with specific users/groups, organize specific projects, or organize based on departments within your organization.

Repos

Databricks Repos allow us to connect our notebooks to repositories, such as Github, to enable CI/CD workflows or to simply backup our work. A personal access token (PAT) is required when setting up a connection. Once the PAT is configured, it is very simple to add new repos. We are going to discuss this in the next lesson.

Trash

As mentioned above, assets that are moved to the trash appear in this section. Any item moved to the trash will be permantently deleted after 30 days, unless you delete it earlier.

"Kebab" and "Add" Menus

Before leaving this lesson, please note that the "kebab" and "Add" menus provide different functionality depending on the section you are in and the type of file involved. For example, the kebab menu for repos will have a "Git" option where you can commit, pull, and otherwise work with Git. Feel free to add new files and experiment with the kebab and Add menus to see how things change.

© 2023 Databricks, Inc. All rights reserved.

Apache, Apache Spark, Spark and the Spark logo are trademarks of the Apache Software Foundation (https://www.apache.org/).