



Demo: Working with Notebooks

Notebooks are the primary means of developing and executing code interactively on Databricks. This lesson provides a basic introduction to working with Databricks notebooks.

Learning Objectives

1. Describe Databricks Notebooks as the most common interface for data engineers when working with Databricks.
2. Recognize common use cases for data engineers when working with Notebooks.
3. Create a new Notebook.
4. Write code in Notebook cells.
5. Run code in Notebook cells.
6. Write markdown-based notes in a Notebook.
7. Run code using multiple languages within the same Notebook.
8. Change the name of a Notebook.
9. Explore basic visualization capabilities in Databricks Notebooks.

Attach to a Cluster

In the previous lesson, you should have either deployed a cluster or identified a cluster that an admin has configured for you to use.

At the top right corner of your screen, click the cluster selector ("Connect" button) and choose a cluster from the dropdown menu. When the notebook is connected to a cluster, this button shows the name of the cluster.

NOTE: If your cluster is not running at the time you select it from the list, it will begin the startup process. Deploying a cluster can take several minutes. A solid green circle will appear to the left of the cluster name once resources have been deployed.

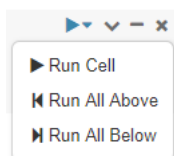
Notebooks Basics

Notebooks provide cell-by-cell execution of code. Multiple languages can be mixed in a notebook. Users can add plots, images, and markdown text to enhance their code.

Throughout this course, our notebooks are designed as learning instruments. Notebooks can be easily deployed as production code with Databricks, as well as providing a robust toolset for data exploration, reporting, and dashboarding.

Running a Cell

- Run the cell below using one of the following options:
 - Place your cursor in the cell, and type **CTRL+ENTER** or **CTRL+RETURN**
 - Place your cursor in the cell, and type **SHIFT+ENTER** or **SHIFT+RETURN** to run the cell and move to the next one
 - Click on the right-pointing triangle image in the upper-right corner of the cell and using **Run Cell**, **Run All Above** or **Run All Below** as seen here



Lesson Setup

Run the following cell. It will install a data file, called `customers.csv`, that we will use throughout this demo.

```
%run ../Includes/Classroom-Setup-05
```

NOTE: Cell-by-cell code execution means that cells can be executed multiple times or out of order. Unless explicitly instructed, you should always assume that the notebooks in this course are intended to be run one cell at a time from top to bottom. If you encounter an error, make sure you read the text before and after a cell to ensure that the error wasn't an intentional learning moment before you try to troubleshoot. Most errors can be resolved by either running earlier cells in a notebook that were missed or re-executing the entire notebook from the top.

View customers.csv

Run the following cell to view the contents of `customers.csv`

```
display((spark.read
          .csv(f"{DA.paths.datasets}/customers.csv")
          ))
```

Run a SQL Cell

As shown immediately to the right of this notebook's name, the default language for the notebook is Python. If we want to run a SQL command in this notebook, we can start the cell with `%sql`. This is called a magic command.

Run the cell below. It contains a SQL command to view the contents of the same file as above.

```
%sql
SELECT * FROM csv.`dbfs:/mnt/dbacademy-datasets/get-started-with-data-engineering-on-databricks/v01/customers.csv`
```

Markdown

With the magic command `%md` we can render Markdown in a cell:

- Double click this cell to begin editing it
- Then hit `Esc` to stop editing

Title One

Title Two

Title Three

This is a test of the emergency broadcast system. This is only a test.

This is text with a **bold** word in it.

This is text with an *italicized* word in it.

This is an ordered list

1. one
2. two
3. three

This is an unordered list

- apples

- peaches
- bananas

Links/Embedded HTML: Markdown - Wikipedia (<https://en.wikipedia.org/wiki/Markdown>)



And of course, tables:

name	value
Yi	1
Ali	2
Selina	3

Create a New Notebook

There are many ways to create a new notebook. If you are already in a notebook, perform the following:

1. Click the **File** menu and select **New notebook**

The notebook opens in a new tab.

2. Change the name of the new notebook to "My Notebook" by clicking the current name ("Untitled Notebook ...") in the upper-left corner

© 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/>).

Privacy Policy (<https://databricks.com/privacy-policy>) | Terms of Use (<https://databricks.com/terms-of-use>) | Support