

## Create data visualizations by using Azure Databricks

### Task 1. Clone the Databricks archive

1. From the Azure portal, go to your Databricks workspace and select Launch workspace.
2. In the left pane, select Workspace, select Users, and then select your username (the entry with the house icon).
3. In the pane that appears, select the downward-pointing chevron next to your name, and then select Import.
4. In the Import Notebooks pane, select URL, and paste in the following URL:

<https://github.com/MicrosoftDocs/mslearn-create-data-visualizations-using-azure-databricks-and-power-bi/blob/master/DBC/07-Visualization.dbc?raw=true>

5. Select Import.
6. A folder named after the archive should appear. Select *07-Visualization* folder. The folder contains one or more notebooks that you'll use in completing this lab.

### Task 2. Complete the following notebooks

1. To do the data-visualization exercises, continue working within your Databricks workspace and open the new 07-Visualization folder. Within the folder, you'll find Python, Scala, and Spark subfolders.
2. Choose the language you prefer to use, open the corresponding folder, and then open the 01-Querying-Files notebook.
3. Follow the instructions in the notebook until you've finished the entire notebook.
  - a. 01-Querying-Files: This notebook contains some basic visualizations that use built-in Databricks functions and DataFrames.
  - b. 02-Exploratory-Data-Analysis: In this notebook, you do some basic exploratory analysis on a sample dataset to prepare it for advanced visualization in the next notebooks.
  - c. 03-Power-BI: In this notebook, you connect your Databricks cluster to Power BI and create visualizations by using Power BI tools.
  - d. 04-Matplotlib: In this notebook, you use Matplotlib to create custom visualizations for your data.