

# Implementing an Azure Databricks Environment with Microsoft Azure

---

## IMPLEMENTING AN AZURE DATABRICKS ENVIRONMENT



**Michael Bender**

AUTHOR EVANGELIST, PLURALSIGHT

@MichaelBender



# Overview



**Introduction to Azure Databricks**

**Fundamental Components of Azure Databricks**

**Working with workspaces, notebooks, tables and jobs**



# Exercise Files

Slides

Code

Links to Resources

The screenshot shows the course page for "PowerShell: Getting Started" by Michael Bender. The page is dark-themed. At the top, there's a search bar and the author's name "Michael Bender" with his email "mbender@bentech.net". The main title "PowerShell: Getting Started" is prominently displayed, followed by "by Michael Bender". Below this, a description states: "This is an introductory course on PowerShell and how to use it for basic IT Operations support." There are four buttons: "Resume Course" (orange), "Bookmark", "Add to Channel", and "Download Course". A "Course Breakdown" section lists topics like "Introduction to PowerShell", "PowerShell Basics", "Gathering information with PowerShell", "Remoting with PowerShell", and "Writing Scripts with PowerShell". The "Exercise files" tab is highlighted with a yellow box. Below the tabs, a paragraph explains: "These exercise files are intended to provide you with the assets you need to create a video-based hands-on experience. With the exercise files, you can follow along with the author and re-create the same solution on your computer. We find this to be even more effective than written lab exercises." A "Download exercise files" button is highlighted with a yellow box. On the right, the "Course author" section shows Michael Bender's profile and a bio. The "Course info" section lists "Level: Beginner", "Rating: ★★★★★ (602)", "My rating: ★★★★★", "Duration: 3h 5m", and "Updated: 22 May 2019". At the bottom right, there are social media icons for Facebook, Twitter, and LinkedIn.

What do you want to learn?

Michael Bender  
mbender@bentech.net

## PowerShell: Getting Started

by Michael Bender

This is an introductory course on PowerShell and how to use it for basic IT Operations support.

[Resume Course](#) [Bookmark](#) [Add to Channel](#) [Download Course](#)

### Course Breakdown

- Introduction to PowerShell
- PowerShell Basics
- Gathering information with PowerShell
- Remoting with PowerShell
- Writing Scripts with PowerShell
- Next Steps

Who Is This Course For?

Table of contents Description **Exercise files** Discussion Learning Check Recommended

These exercise files are intended to provide you with the assets you need to create a video-based hands-on experience. With the exercise files, you can follow along with the author and re-create the same solution on your computer. We find this to be even more effective than written lab exercises.

[Download exercise files](#)

Course author

**Michael Bender**

Michael is a six-time Microsoft Most Valuable Professional, author, technical trainer, and community leader. Having been in the IT industry since the 90's, his experiences covers the gamut of...

Course info

Level **Beginner**

Rating ★★★★★ (602)

My rating ★★★★★

Duration **3h 5m**

Updated **22 May 2019**

Share course

[f](#) [t](#) [in](#)



# What is Azure Databricks



**Scalable analytics platform in Azure**

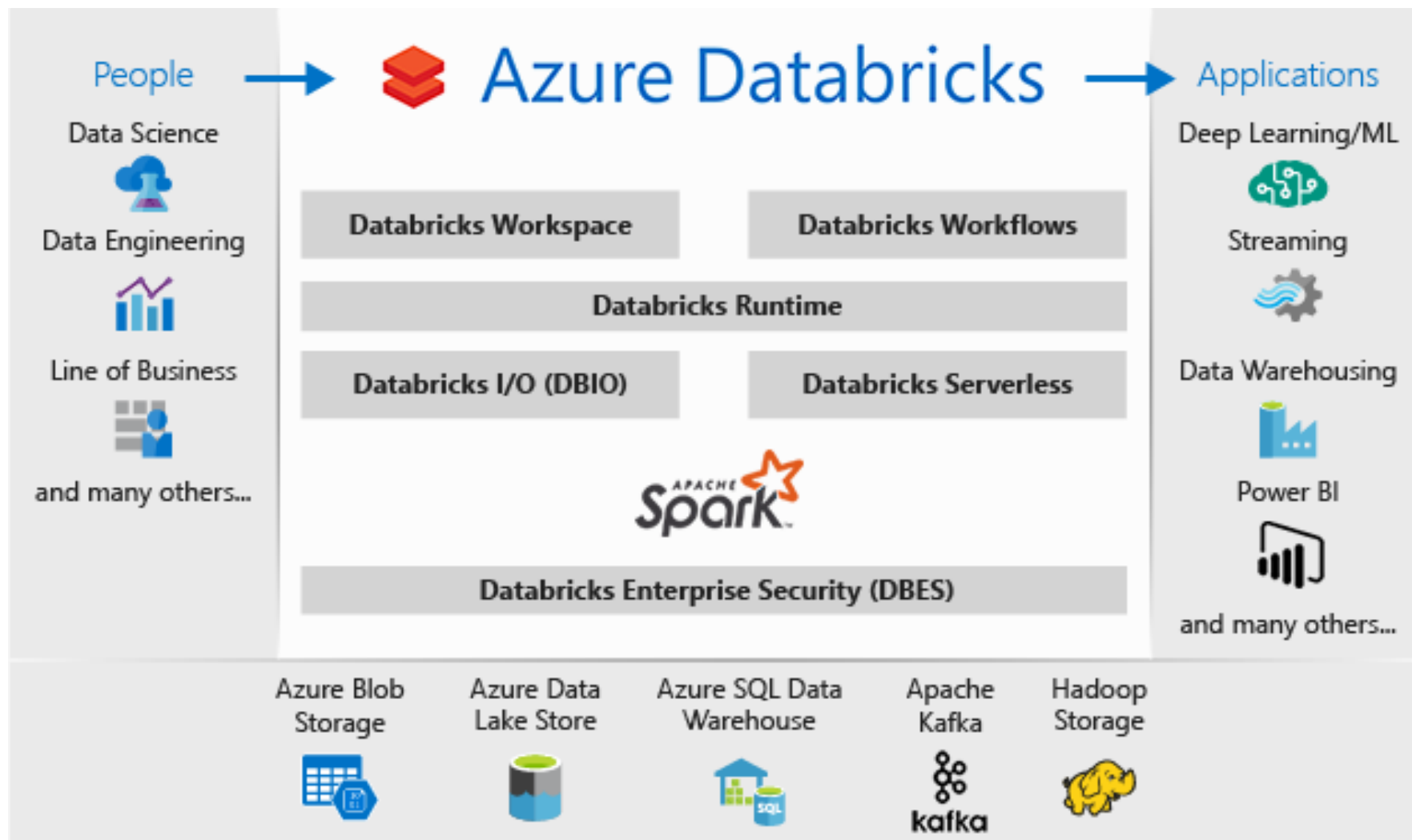
**Based on Apache Spark**

**Workflows and workspaces for data users**

**Native integration with other Azure services**



# What is Azure Databricks



# Users of Azure Databricks



**Data Engineers**

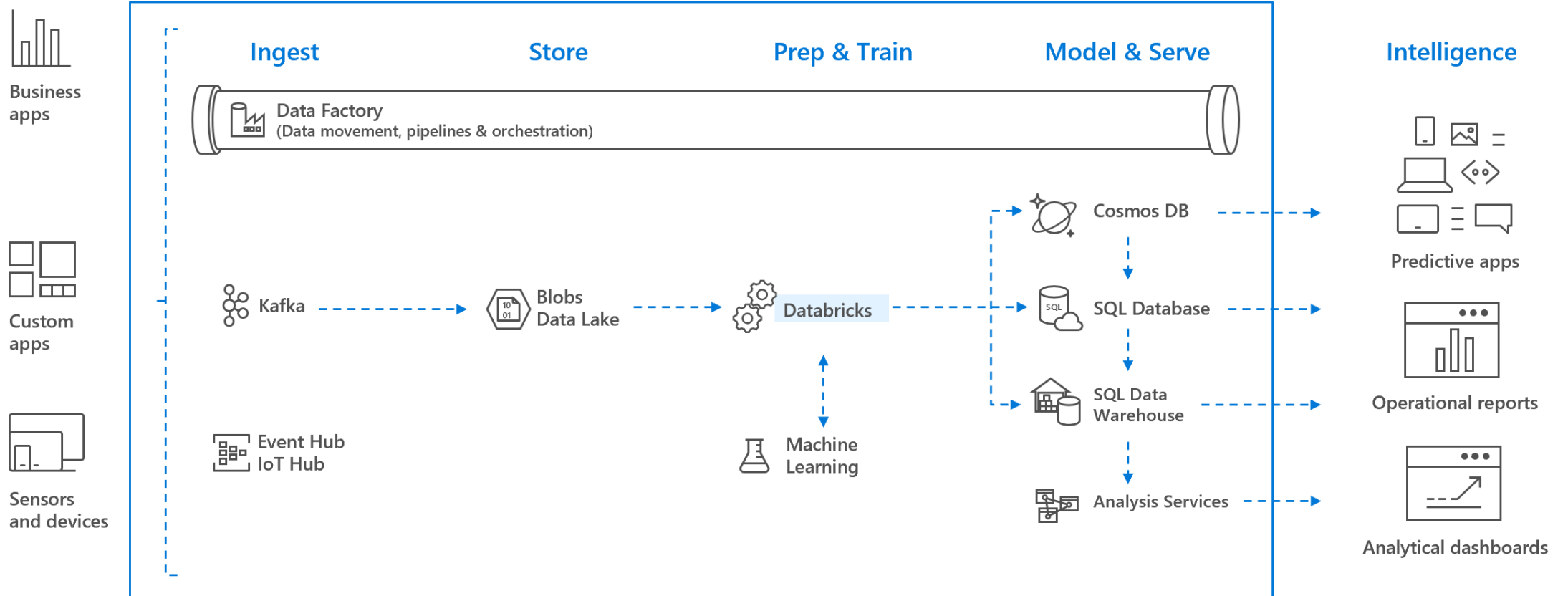


**Data Scientists**

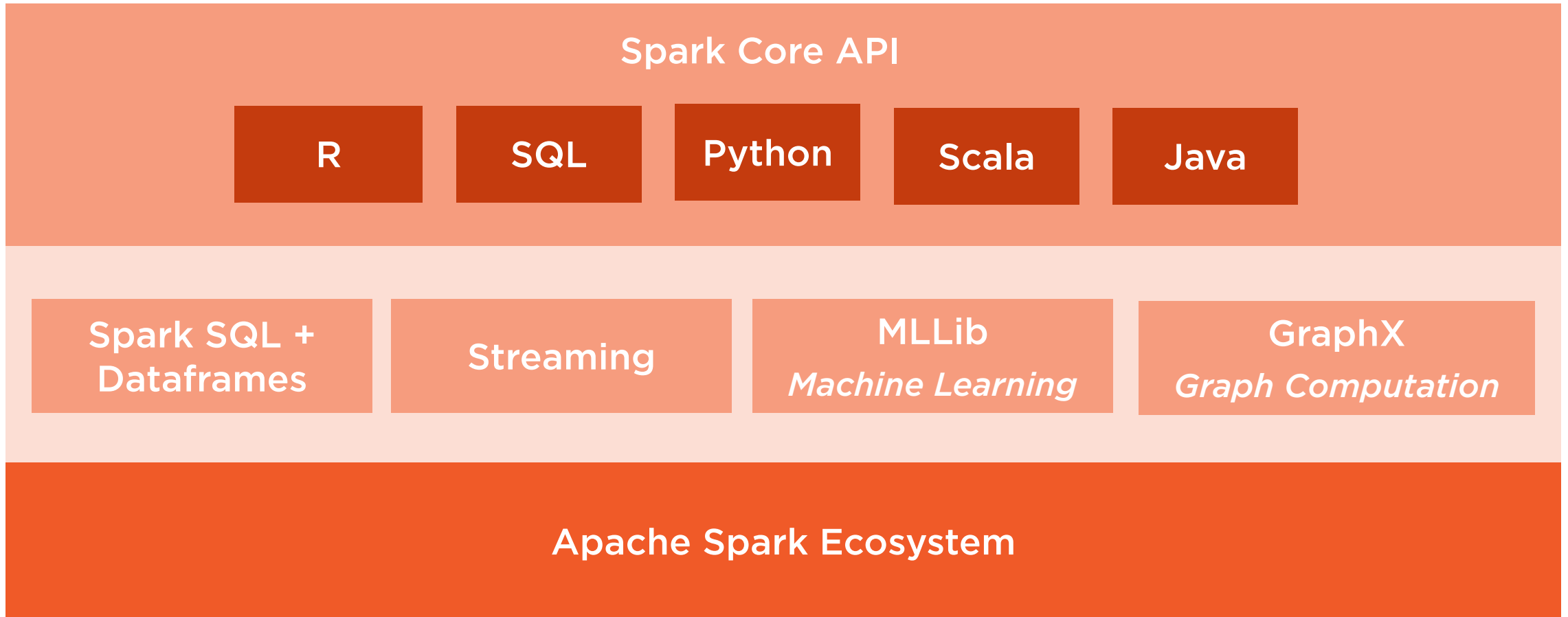


**Business Users**

# Data Pipeline and Azure Databricks



# Apache Spark-based analytics platform





# Typical use cases for Azure Databricks

**Interactive analytics**

**Data integration**

**Machine learning**

**Stream processing**



# Fundamental Azure Databricks Components

**Workspaces**

**Clusters**

**Notebooks**

**Tables**

**Jobs**



# Workspaces



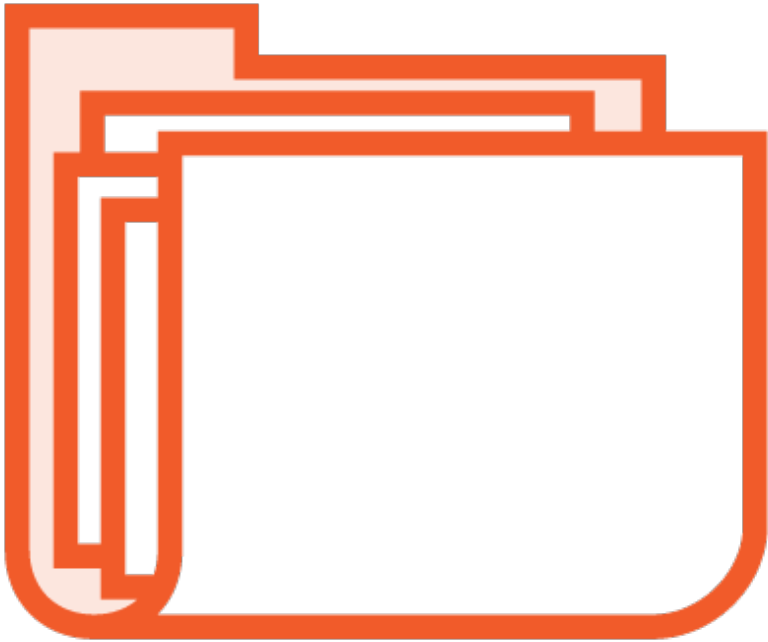
Organizes all of your Azure Databricks assets

Control Access via Workspace Access Control

Create and manage using UI, CLI or Workspace API



# Folders



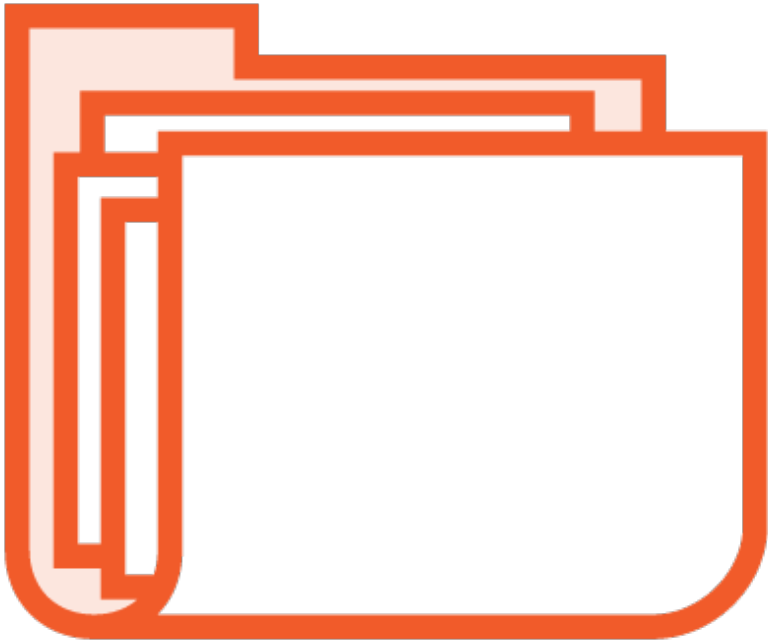
**Storage for all workspaces assets**

**Check the icon for object type contained**

**Use access control to manage collaboration**

**Special folders**

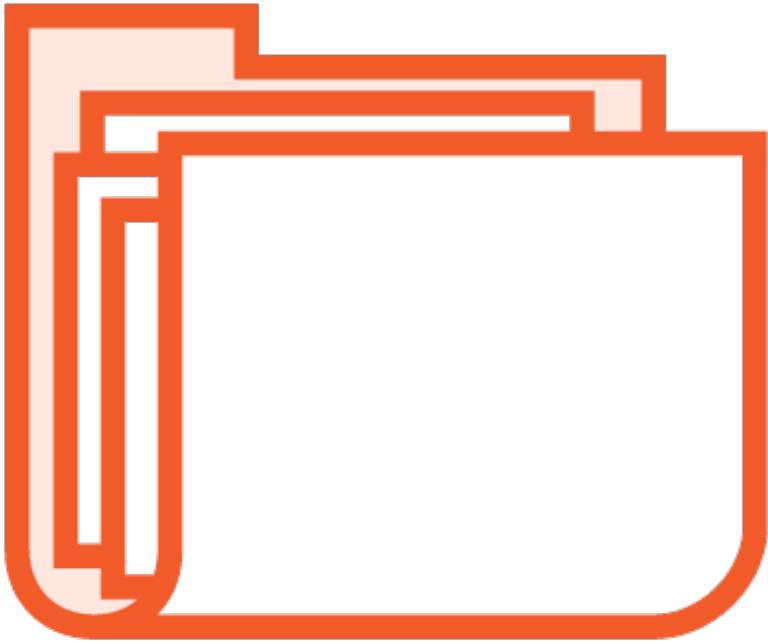
# Folders



**Storage for all workspaces assets**



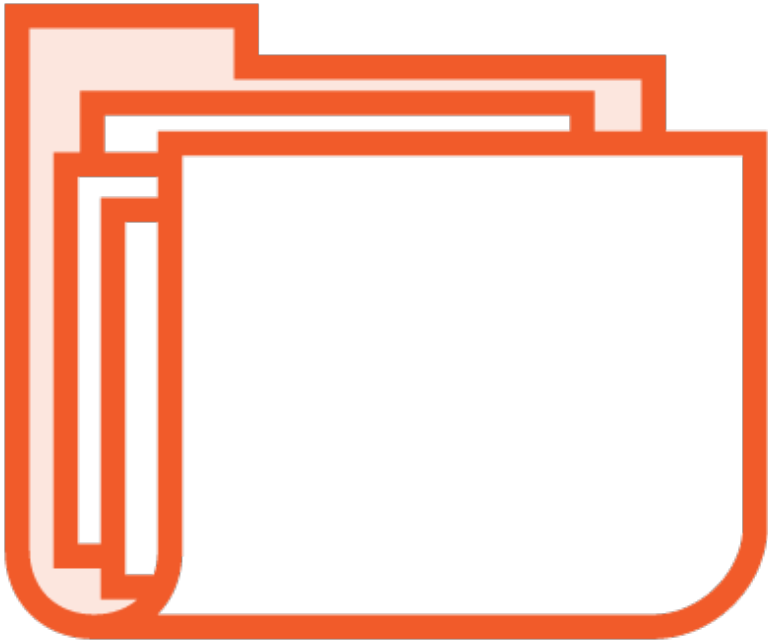
# Folders



**Storage for all workspaces assets**

**Check the icon for object type contained**

# Folders



**Storage for all workspaces assets**

**Check the icon for object type contained**

**Use access control to manage collaboration**



# Folders



**Storage for all workspaces assets**

**Check the icon for object type contained**

**Use access control to manage collaboration**

**Special folders**



# Demo



Creating an Azure Databricks workspace

Exploring the Azure Databricks portal

Working with Folders



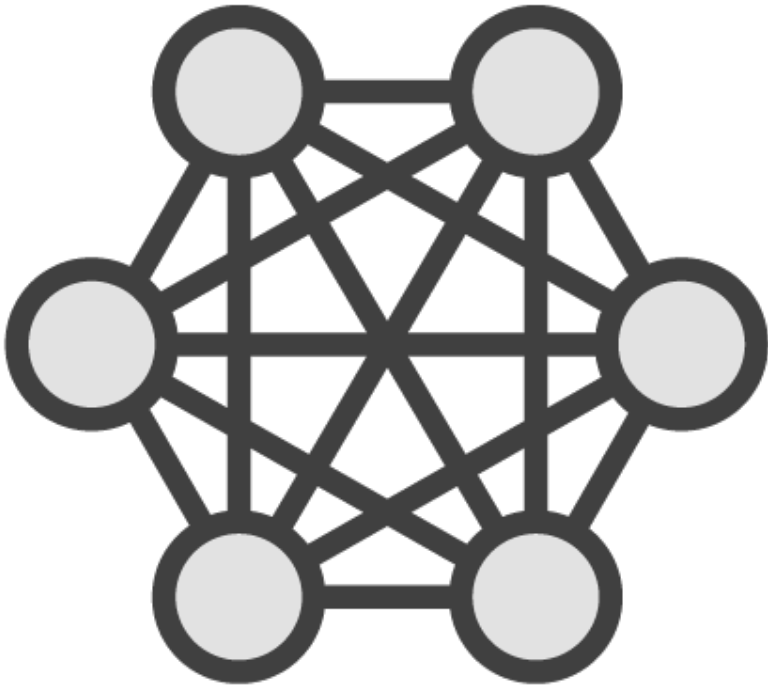
# Demo



## Getting started with the Databricks CLI



# Azure Spark Clusters



**Unified cluster computing platform**

**Interactive or job**

**Manage using UI, CLI or Clusters API**

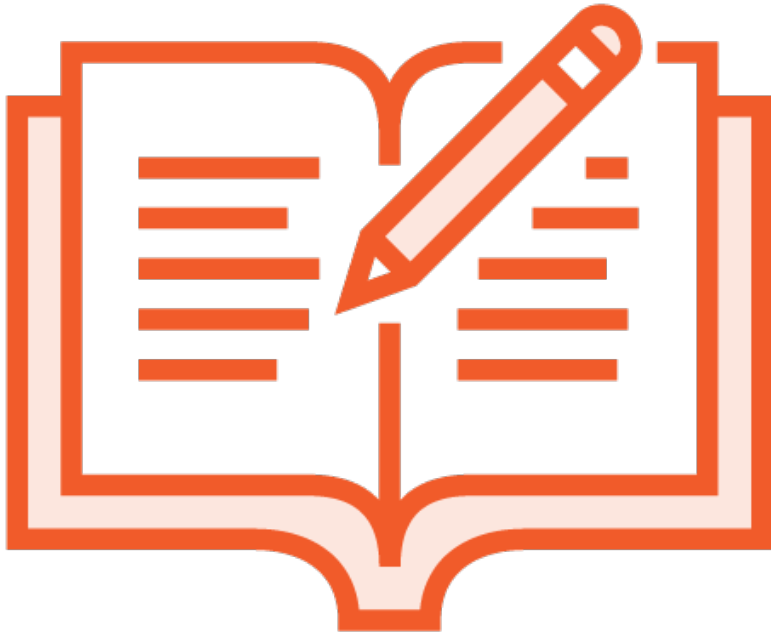
# Demo



## Working with Spark clusters



# Notebooks



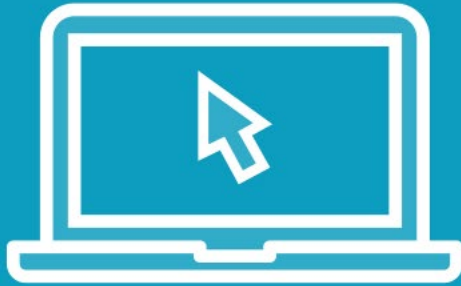
**Web-based interface**

**Combine code, visualizations and text**

**Used for building pipelines**

**Support formats of Python, Scala,  
Markdown and more**

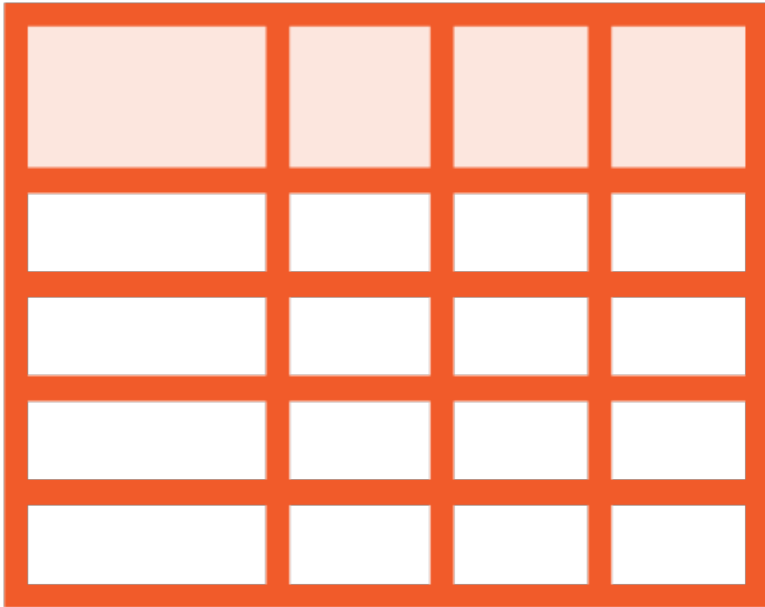
# Demo



## Working with Notebooks in Azure Databricks



# Azure Databricks Tables




**Collection of structured data**

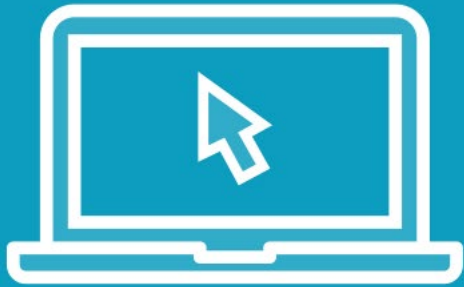
**Equivalent to Apache Spark DataFrame**

**Can be cached, filtered, queried, and more**

**Temporary storage during transformation**



# Demo



## Working with Tables





# Apache Spark Jobs



Run workflows interactively or scheduled

Accessible through UI, CLI or API

Limits



# Demo



## Creating Spark Jobs in Azure Databricks



## Summary



Azure Databricks is quick and easy to set up

Don't forget about auto-scaling for your clusters

Use your favorite Language

Notebooks are key to working in Azure Databricks



# For Further Learning

Azure Databricks documentation at **docs.microsoft.com**  
<https://docs.microsoft.com/en-us/azure/azure-databricks/>

Azure Databricks documentation at  
**docs.azuredatabricks.net**  
<https://docs.azuredatabricks.net/user-guide/index.html>

Remember the module exercise files

Questions? Join on the conversation at [pluralsight.com](https://pluralsight.com)



Next Up:  
Performing ETL (Extract, Transform, Load)  
Operations with Azure Databricks

---

