

# TOOLS FOR IMPROVING THE DEVELOPER EXPERIENCE

---

Theresa Hammer  
June 11, 2024



# DEVELOPER ECOSYSTEM

What does the developer ecosystem entail?



IDE Plugins

Databricks CLI (DABs)

Terraform Provider

SDKs (Go, Python, JavaScript, Java)

OpenAPI Models



# SDKs

## Why Databricks SDKs?



- Databricks REST APIs are vast
  - 600 methods across 80 services
  - 10 authentication types across 3 clouds
  - 73 iterators with 7 different paginations
- **Goal:** idiomatic, complete, consistent, unified authentication, minimal dependencies
- **Core SDKs:** Go, Python, Java, JavaScript (in beta)
- 40+ Databricks solutions built on top (e.g. CLI, Terraform, VS Code, UCX)

# SDKs

## Example: List all new objects in workspace

### Without the SDK

- Lookup how to authenticate
- Go to our REST API docs
- Make the request, parse the response from JSON
- Pagination
- Recursively list directories
- To extend, learn an entirely new API

### With the Python SDK

```
from databricks.sdk import WorkspaceClient

w = WorkspaceClient()
w.workspace.list(
    "/",
    notebooks_modified_after=1000)
```

# CLI

## Overview Databricks CLI



- **Use cases**
  - Ad hoc experimentation
  - Scripting
  - Invoking the REST API directly
  - Manage local authentication profiles
  - Sync code (e.g. notebook) from IDE to workspace
- **Unified authentication**
- **Command Groups (e.g. workspace, jobs, clusters)**

```
databricks <command-group> -h
```

```
databricks <command-group> <command-name> -h
```

```
databricks <command-group> <command-name> <subcommand-name> -h
```

# DATABRICKS ASSET BUNDLES (DABs)



## Databricks Asset Bundles in a nutshell

### What are DABs?

- Declarative format for describing resources and code
- Bundle CLI command group
- Premade and custom templates
- Supported assets: notebooks, wheels, files
- Supported resources: DLT, workflows, model serving end points, MLflow models and experiments

### Why use DABs?

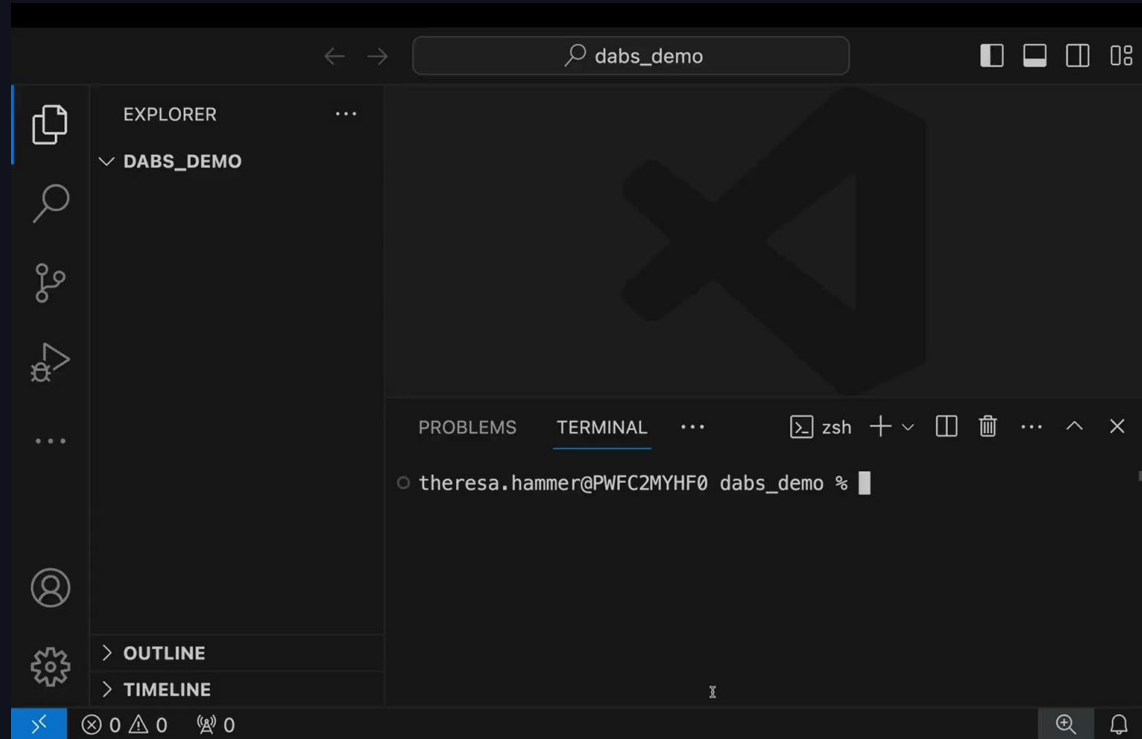
- Version, test and deploy your project as one unit
- Adopt software engineering best practices
- Isolate development copies of the project
- Improve developer productivity
- Repeatable deployments and auditable changes

### There is more!

- Generate a DAB from an existing job or pipeline
- Target overrides
- Use variables and lookups for modular configuration
- Build and deploy shared code and libraries
- Automate deployments using your CICD system of choice

# DATABRICKS ASSET BUNDLES (DABs)

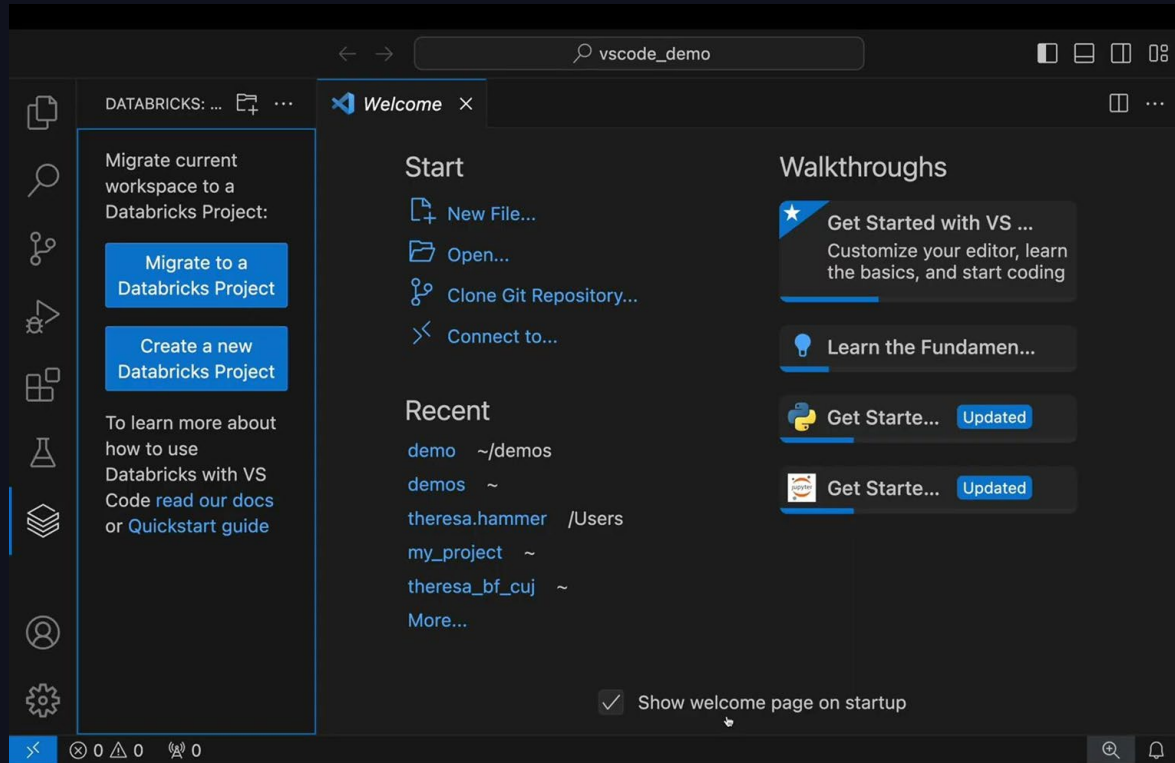
## How to convert an existing job to a DAB?



1. `databricks auth login --host <workspace-url>`
1. `databricks bundle init`
1. `databricks bundle generate job --existing-job-id <id>`
1. `databricks bundle deployment bind <name> <id>`
1. `databricks bundle deploy -t <target>`

# IDE - VS Code extension

## Authentication and Project initialization (DABs)



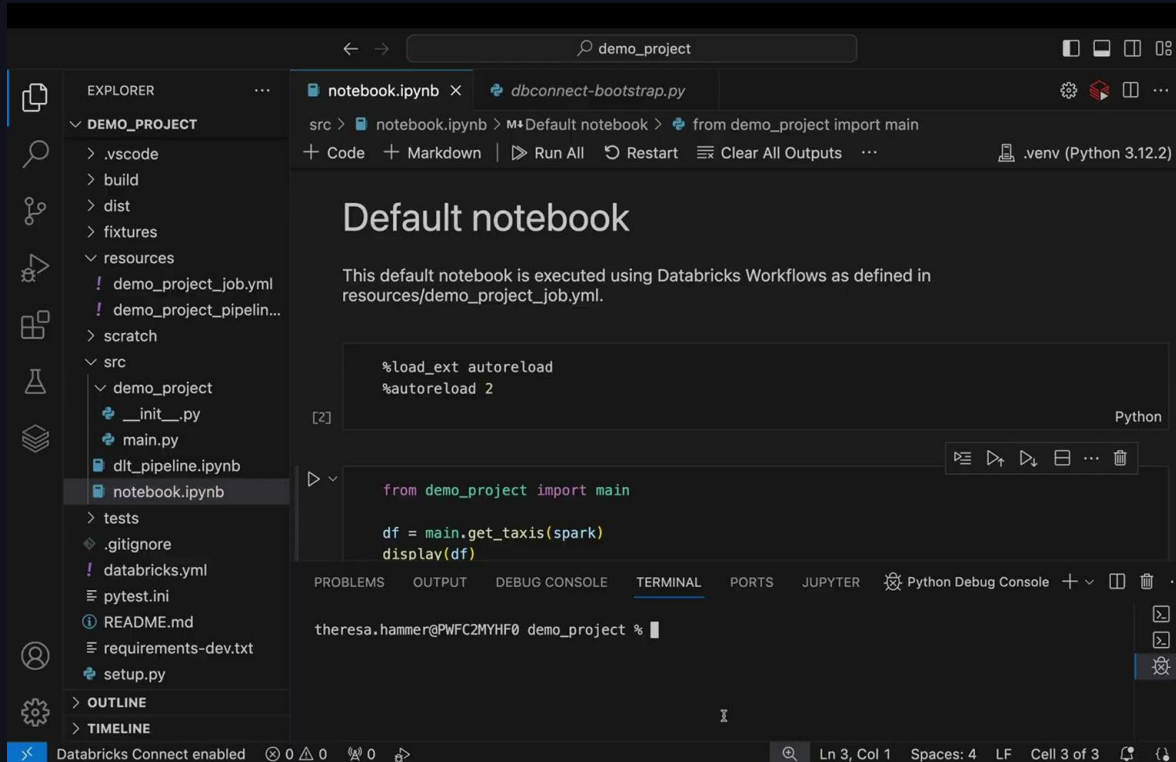
### Topics covered:

Authentication  
Project initialization  
Environment set up  
Deploy and run your bundle



# IDE - VS Code extension

## Debugging



### Topics covered:

- Cell by cell execution
- Run file on Databricks
- Debug a notebook cell
- Debug a file

# TOOLS FOR DIFFERENT USE CASES

## When to use which tool?

User Persona	Tasks	Tool to use
Admins / DevOps Engineers	Create workspaces, catalogs, metastores and enforce permissions	Terraform
	Guarantee environment portability and disaster recovery	Terraform
Data Engineer / Data Analytics Engineer / Data Scientist	Apply best practices and CI/CD to Databricks projects and co-version, co-author and co-deploy your resources and assets as one unit	DABs (CLI)
	Ad-hoc experimentation, shell one liners, sync code from IDE to workspace	CLI
	Integrate with existing deployment systems, create custom Databricks workflows and new web services	SDKs
	Developing from local IDE	IDE plugins

# NEXT STEPS

## What's next on the roadmap and how to learn more?

- Python SDK for DABs “PyDABs” (in private preview)
- DABs in the workspace
- Databricks PyCharm plugin (new private preview)
- Low-latency bindings for model serving endpoints

WHAT	WHEN
<b>Code to Production:</b> Leveraging Your Favorite IDEs with Databricks	<b>Tuesday, Jun 11</b> 10:10 AM - 10:50 AM <b>Wednesday, Jun 12</b> 5:10 PM - 5:50 PM PDT
<b>Databricks Asset Bundles:</b> A unifying tool for deployment of databricks	<b>Thursday, Jun 13</b> 1:40 PM - 2:20 PM 5:10 PM - 5:50 PM
<b>Path to Production:</b> Databricks Project CICD for Seamless Inner to Outer Dev Loops	<b>Thursday, Jun 13</b> 12:30 PM - 1:10 PM