

Lab-1: Setup Environment

Introduction

This lab will set up the Cloud Pak for Data environment for subsequent labs and introduce you to the Project and Gallery features of Cloud Pak for Data. Cloud Pak for Data is an integrated platform of tools, services, data, and meta-data to help companies and agencies accelerate their shift to be data driven organizations. The platform enables data professionals such as data scientists, data engineers, business analysts, and application developers collaboratively work with data to build, train, deploy machine learning and deep learning models at scale to infuse AI into business to drive innovation. Cloud Pak for Data is designed to support the development and deployment of data and analytics assets for the enterprise.

End-to-End Data Science

The general flow of the End to End Data Science PoT will be guided by the activities shown in Figure 1- End to End Flow. This lab will focus on the Create Project and Research Topics activities.

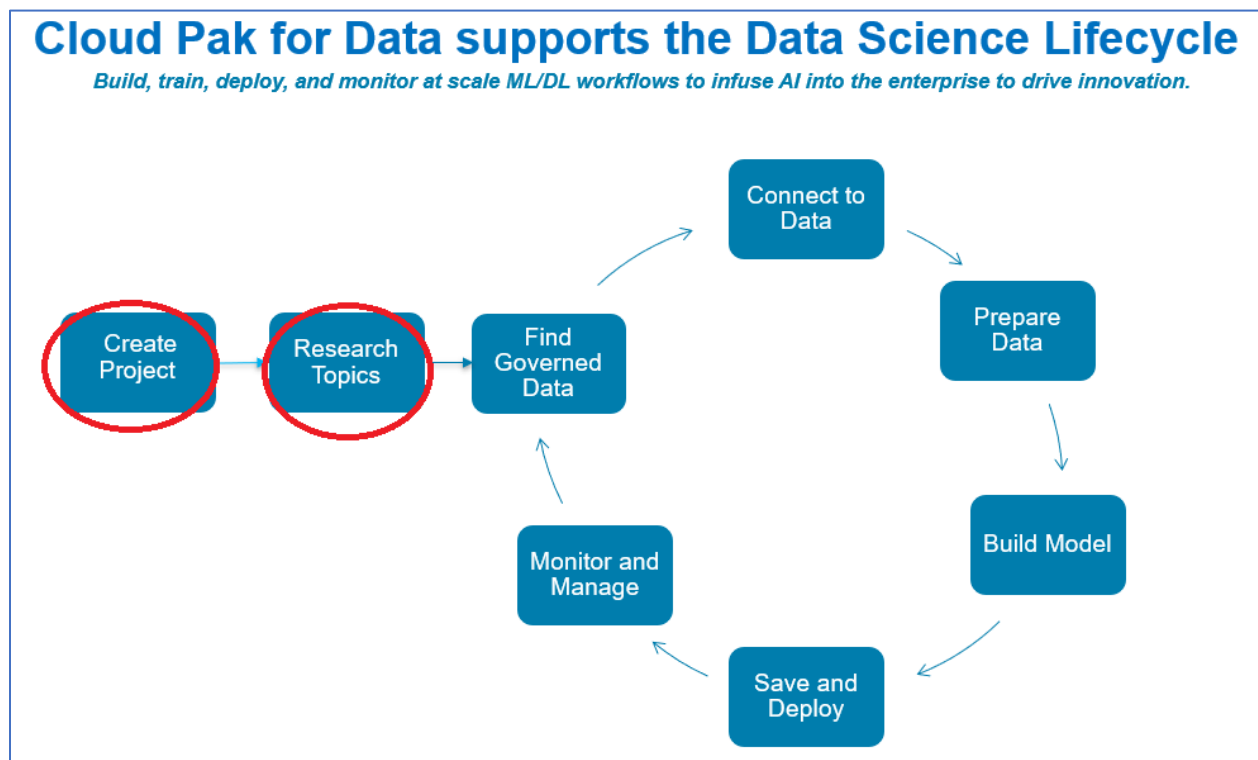


Figure 1- End to End Flow

Objectives

The goal of this lab is to familiarize the user with the Project and Gallery features of Cloud Pak for Data, and to set up the environment for subsequent labs. Projects are a core component of Cloud Pak for Data. Projects enable you to organize your analytic and data assets in one place.

Projects are also the home base for collaboration. Colleagues can be added as collaborators on a project with administrator, editor, or viewer access.

The [Watson Gallery](#) contains samples that you can use in your project:

- Run sample notebooks to learn new techniques or to use as templates for your own notebooks.
- Add sample data sets to your project

The [Watson Community](#) contains resources to help you learn about data science:

- Read articles from many sources to keep current with data science trends.
- Read tutorials for multiple skill levels to learn how to do specific data science tasks.

The Watson Gallery and Watson Community features support the “Research Topics” activity in the end-to-end process shown above.

After completing this lab, you will be familiar with these features of Watson Studio.

1. Create a project
2. Associate an existing Watson Machine Learning service instance with the project
3. Associate an existing Watson OpenScale service instance with the project
4. Add a collaborator to the project
5. Research topics by searching for a notebook in the Gallery

Create a Project

1. Log into your Cloud Pak for Data account by typing in the url **dataplatfrom.cloud.ibm.com** in your Firefox or Chrome browser.
2. Enter the **Username, Password** and click **Log in**.

Log in to IBM Cloud Pak for Data
Starter edition

Need an account? [Sign up and try for free](#)


Username [Forgot ID?](#)

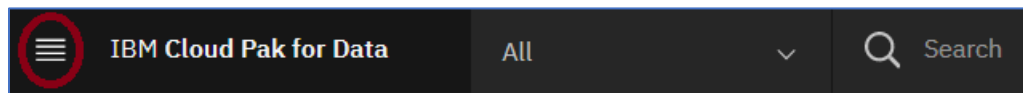
Password [Forgot password?](#)

☐ Remember ID

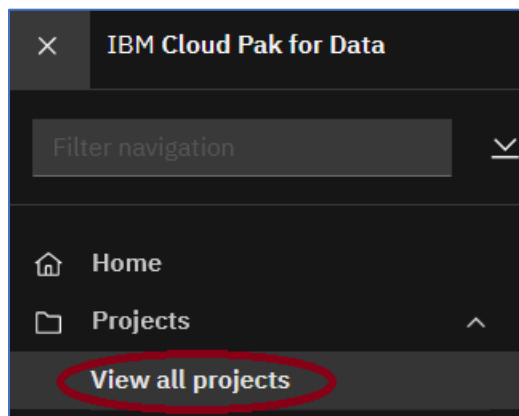
You will log into Dallas ▾

Need help? [Contact the IBM help desk](#)

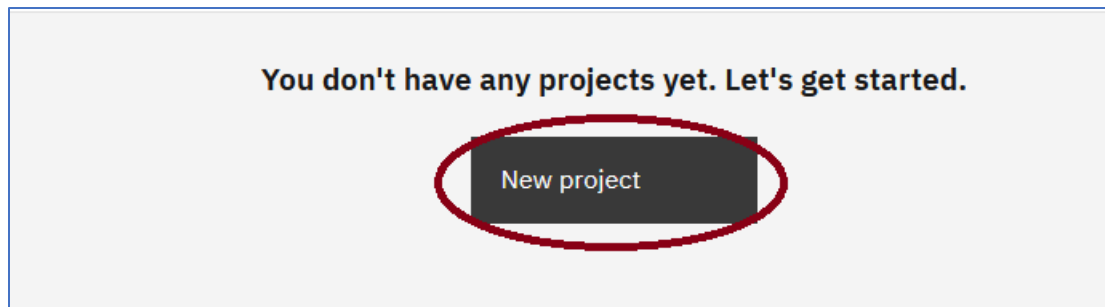
3. Click on the hamburger icon .



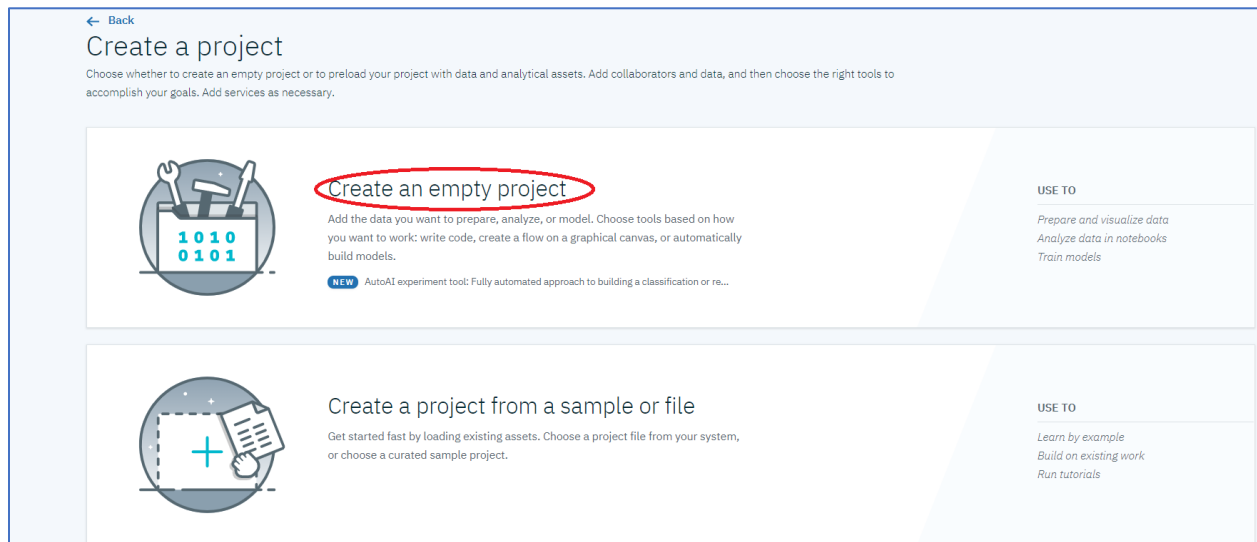
4. Click on **View All Projects**



5. Click on **New Project**.



6. Click on Create an empty project.



7. Enter "Watson Studio Labs" for the **Name**, optionally enter a **Description**, check **Restrict who can be a collaborator** (if it's unchecked) and click **Create**. Note, if you have more than one instance of object storage listed, select an instance. Further note, if you do not have object storage already provisioned, you will see an area marked **Define Storage** and this must be completed before you can click Create. If you need to define object storage, please continue, otherwise go to step 12.

Define details

Name

Description
Project description

Storage

Choose project options

☒ Restrict who can be a collaborator ⓘ

Project includes integration with [Cloud Object Storage](#) for storing project assets.

Cancel

Create

- The next few steps are to be followed if you don't have object storage provisioned. Otherwise skip to step 12. In the **Define Storage** area click on **Add** to add an object storage instance.

New project

Define project details

Name

Description
Project description

Define storage

1 Select storage service

Add

Add an object storage instance, and then return to this page and click Refresh.

2 Refresh

Choose project options

☐ Restrict who can be a collaborator ⓘ

Project includes integration with [Cloud Object Storage](#) for storing project assets.

- Make sure the **Lite** plan is selected, and then click on **Create**

Cloud Object Storage

Author: IBM • Date of last update: Sep 23, 2020 • [Docs](#) • [API Docs](#)

Create

About

Pricing plan

Displayed prices do not include tax. Monthly prices shown are for country or region: United States

Plan	Features	Pricing
Lite	1 COS Service Instance Storage up to 25 GB/month Up to 2,000 Class A (PUT, COPY, POST, and LIST) requests per month Up to 20,000 Class B (GET and all others) requests per month Up to 10 GB/month of Data Retrieval Up to 5GB of egress (Public Outbound) Applies to aggregate total across all storage bucket classes The Lite service plan for Cloud Object Storage includes Regional and Cross Regional resiliency, flexible data classes, and built in security. Lite plan services are deleted after 30 days of inactivity.	Free
Standard	There is no minimum fee, so you pay only for what you use. See pricing details	

Cloud Object Storage
Region: Global
Plan: Lite
Service name: Cloud Object Storage-yp
Resource group: Default

Create

10. Note that it may take a minute for your storage service to show up. Click **Refresh**.

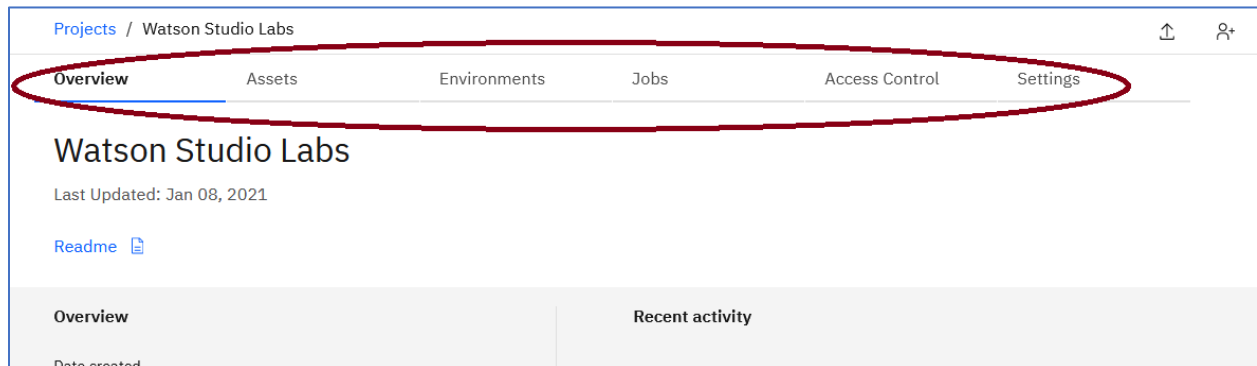
The screenshot shows the 'New project' form. On the left, under 'Define project details', there are fields for 'Name' (containing 'Watson Studio Labs') and 'Description' (containing 'Project description'). Below these is the 'Choose project options' section with a checkbox for 'Restrict who can be a collaborator' and a note: 'Project includes integration with Cloud Object Storage for storing project assets.' On the right, under 'Define storage', there is a step indicator '1 Select storage service' followed by an 'Add' button. Below 'Add' is the text: 'Add an object storage instance, and then return to this page and click Refresh.' A second step indicator '2 Refresh' is circled in red, indicating the next action.

11. Click **Create**.

The screenshot shows the 'New project' form after the storage service has been added. The 'Storage' section now displays 'cloud-object-storage-qo'. The 'Define project details' and 'Choose project options' sections remain the same. At the bottom right, there are two buttons: 'Cancel' and 'Create'. The 'Create' button is circled in red, indicating the next action.

- 12.** The Project **Overview** page is shown. This page provides summarized information about the project. In addition to the Overview page, are five other pages described below.
- a. Assets Page** – Analytics and Data assets can be added to the project from this page.
 - b. Environments Page** - Provides information on the current notebook environments that are defined, lists the active notebook environments currently running, and enables users to create custom notebook environments. Similar information for DataStage environments is provided as well. DataStage is a highly performant ETL engine.
 - c. Jobs Page** – Provides the interface to the job subsystem.

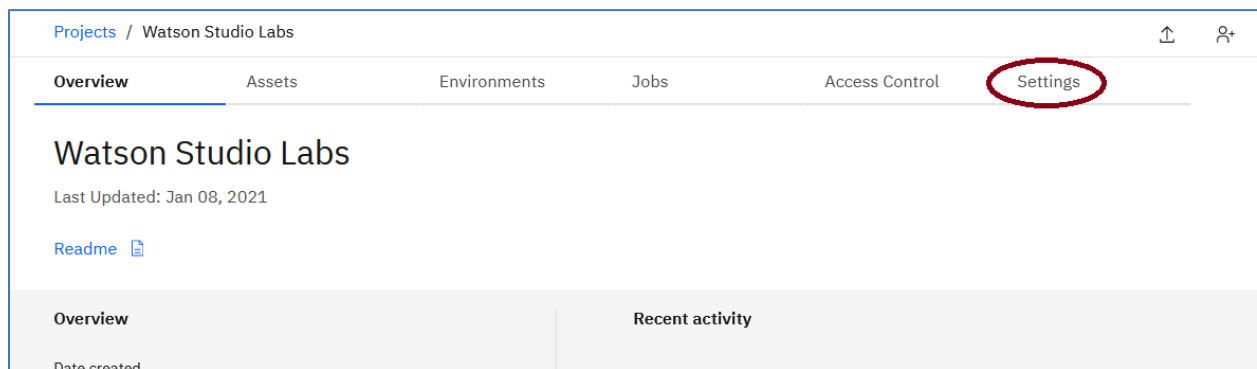
- d. **Access Control** – Lists the project collaborators and enables users to add/remove collaborators.
- e. **Settings** – Enables users to view and set project attributes.



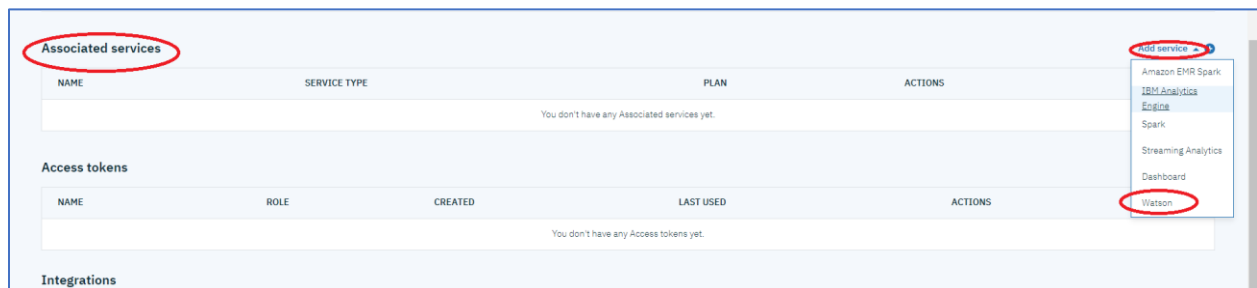
Associate Watson Machine Learning service to the Project

To save and deploy machine learning models, a Watson Machine Learning service must be created (if one doesn't exist) and associated to our project.

1. Click on **Settings** to navigate to the Project **Settings** page.



2. Scroll down to **Associated Services**, click on **Add service**, click on **Watson**.



3. Newer Watson Studio accounts come with the WatsonMachineLearning instance already created. Click on the checkbox adjacent to the WatsonMachineLearning (Type-Machine Learning) service and click **Associate service**. Note, if a service of Type **Machine Learning** does not exist, go to step 6.

Associate service
Choose an existing or add a new service to associate with your project.

Filter by: Resource Groups ▾ 2 x Locations ▾ None ▾

1 item selected Associate service | Cancel

	Name	Type	Plan	Location	Status	Group
<input type="checkbox"/>	WatsonOpenScale	Watson OpenScale	Lite	Dallas	Not associated	Default
<input checked="" type="checkbox"/>	WatsonMachineLearning ⓘ	Machine Learning	Lite	Dallas	Not associated	Default

4. Click on **x** to close the window.

✓ Service WatsonMachineLearning successfully associated with project Watson Studio Labs. x

Choose an existing or add a new service to associate with your project.

Filter by: Resource Groups ▾ 2 x Locations ▾ None ▾

Search New service +

	Name	Type	Plan	Location	Status	Group
<input checked="" type="checkbox"/>	WatsonMachineLearning ⓘ	Machine Learning	Lite	Dallas	Associated	Default

5. The **WatsonMachineLearning** service is associated with the project. Skip to the next section - Add a Project Collaborator

Associated services Add service +










NAME	SERVICE TYPE	PLAN	ACTIONS
WatsonMachineLearning	Watson - Machine Learning		

6. If a service of type Machine Learning does not exist, you will need to create the service. Click on **New service**.


Search New service +

	Name	Type	Plan	Location	Status	Group
<input checked="" type="checkbox"/>	WatsonMachineLearning ⓘ	Machine Learning	Lite	Dallas	Associated	Default

7. Click on **Machine Learning**.

 <p>Watson OpenScale AI / Machine Learning</p> <p>IBM Watson OpenScale is an enterprise-grade environment for AI infused applications that provides...</p> <p>Lite • Free</p>	 <p>Watson Assistant AI / Machine Learning</p> <p>Watson Assistant lets you build conversational interfaces into any application, device, or channel.</p> <p>Lite • Free</p>	 <p>Watson Discovery AI / Machine Learning</p> <p>Add a cognitive search and content analytics engine to applications.</p> <p>Lite • Free</p>
 <p>Language Translator AI / Machine Learning</p> <p>Translate text, documents, and websites from one language to another. Create industry or region-specific...</p> <p>Lite • Free</p>	 <p>Natural Language Classifier AI / Machine Learning</p> <p>Natural Language Classifier uses advanced natural language processing and machine learning techniques to...</p> <p>Free</p>	 <p>Natural Language Understanding AI / Machine Learning</p> <p>Analyze text to extract meta-data from content such as concepts, entities, emotion, relations, sentiment and...</p> <p>Lite • Free</p>
 <p>Personality Insights AI / Machine Learning</p> <p>The Watson Personality Insights derives insights from transactional and social media data to identify...</p> <p>Lite • Free</p>	 <p>Machine Learning AI / Machine Learning</p> <p>IBM Watson Machine Learning - make smarter decisions, solve tough problems, and improve user outcomes.</p> <p>Lite • Free</p>	 <p>Speech to Text AI / Machine Learning</p> <p>Low-latency, streaming transcription</p> <p>Lite • Free</p>

8. Make sure the Lite service is selected and click **Create**.



Machine Learning

Author: IBM SPSS • Date of last update: Dec 11, 2020 • [Docs](#) • [API Docs](#)

Create

About

Select a region

Select a region

Dallas

Pricing plan

Displayed prices do not include tax. Monthly prices shown are for country or region: United States

Plan	Features	Pricing
Lite	Service instance 20 capacity unit-hours (CUH) included: Maximum 1 (one) GPU Capacity Type: • 1 (one) NVIDIA K80 GPU = 3 capacity units required per hour • 1 (one) NVIDIA V100 GPU = 10 capacity units required per hour • 1 vCPU and 4 GB RAM = 0.5 capacity units required per hour • 2 vCPU and 8 GB RAM = 1 capacity units required per hour • 4 vCPU and 16 GB RAM = 2 capacity units required per hour	Free

Summary

Machine Learning

Region: Dallas

Plan: Lite

Service name: Machine Learning-gp

Resource group: Default

Create


View terms

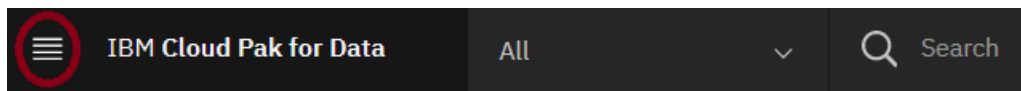
Add a Project Collaborator

Colleagues can gain access to a project's data and analytic assets by being made a collaborator. Permissions are based on the assigned role. The roles are administrator, editor, and viewer.

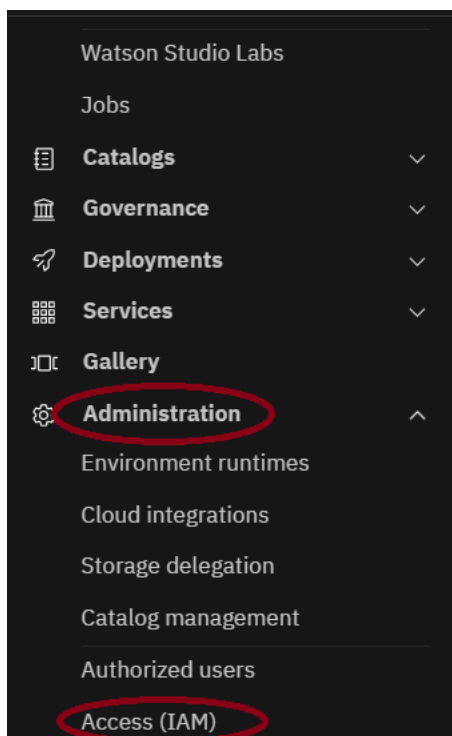
Recall, the project that was created in this lab restricts who can collaborate. This option was required to demonstrate the Watson Knowledge catalog features in lab-2. The restriction limits the collaborators to be members of your company (if your company has federated SAML with IBM Cloud), or a member of the project creator's IBM Cloud account. Given the restriction, to demonstrate adding collaborators to the project, we will need to first add the collaborator to your IBM Cloud account. Instructions to remove the collaborator are provided at the end of lab-2.

Step 1 – Add Collaborator to the IBM Account

1. Click on the hamburger  icon

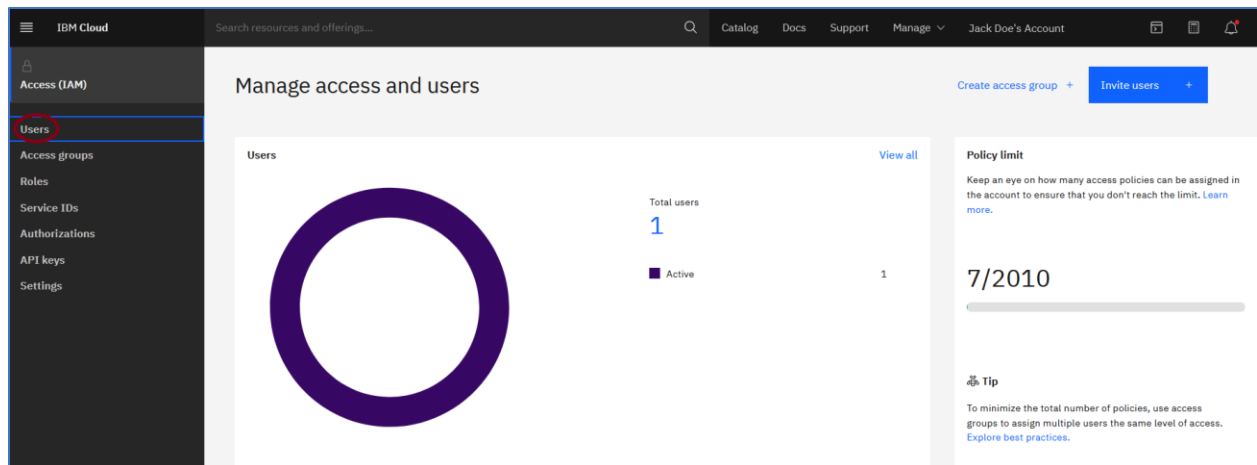


2. Click on **Administration** and then click on **Access (IAM)**. Scroll down if necessary.

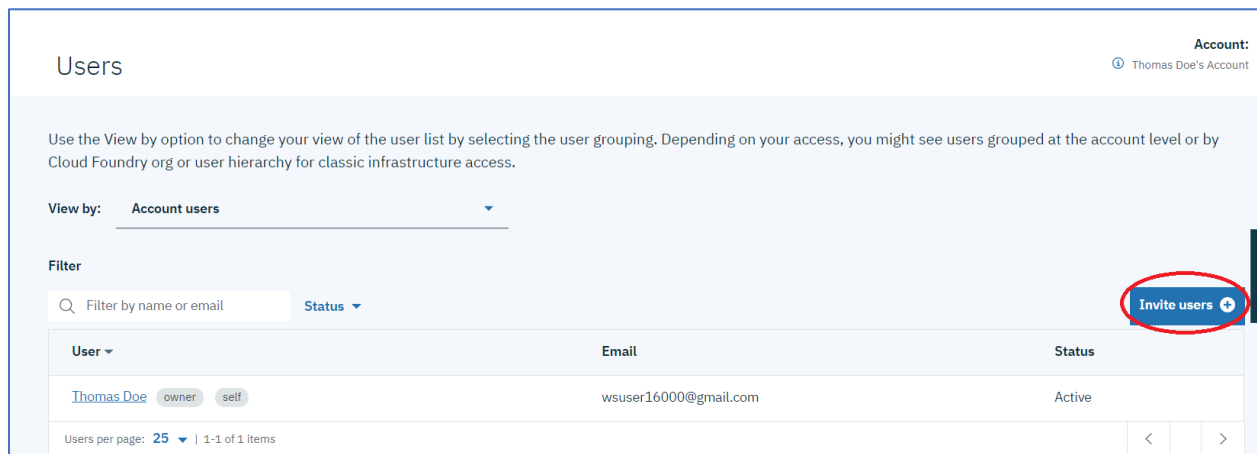


3. An **Identity and Access Management (IAM)** browser tab is created providing the IBM Cloud user interface to the IAM subsystem. Note, you may have to log-in first. Click on **Users** in the menu panel.

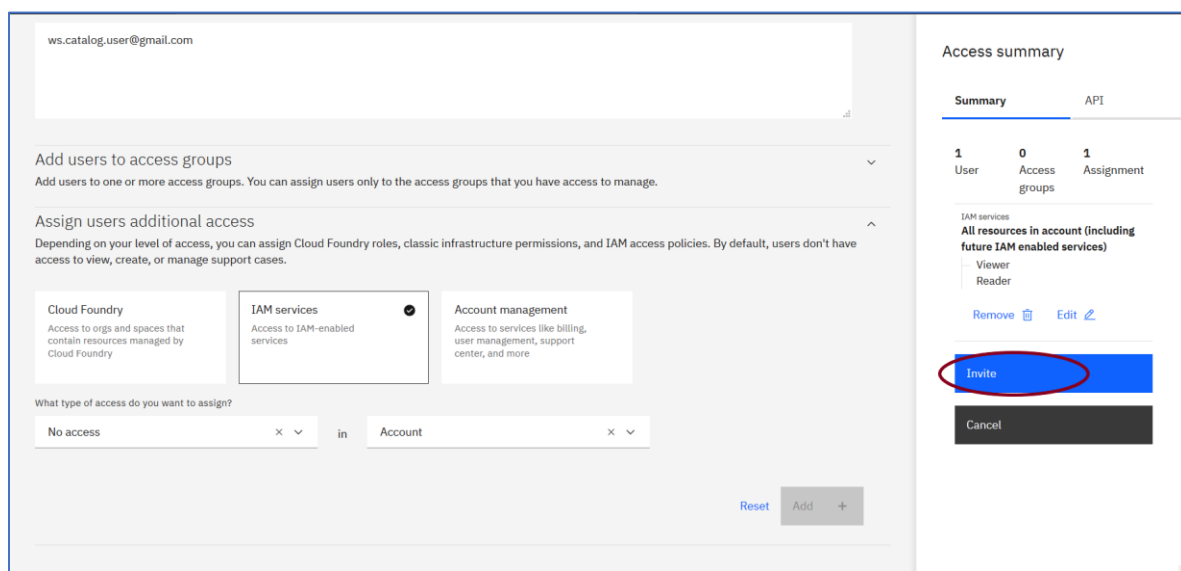




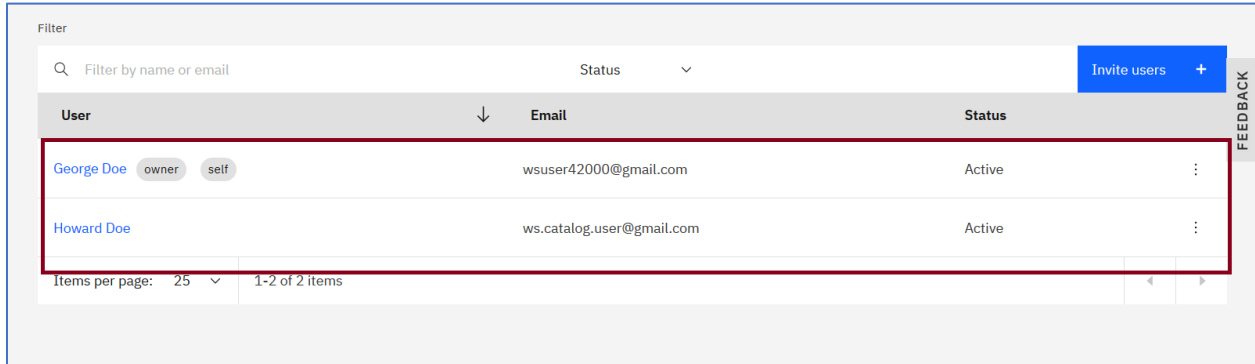
4. Click on **Invite users**.



5. For **E-mail address**, enter **ws.catalog.user@gmail.com** and click **Invite**



- You should have two users in the account. The second user should be “Howard Doe”. Note that “Howard Doe” will need to accept the invite to be put into an **Active Status** (this is a recent change in behavior for Cloud Pak for Data). The status will move from **Processing** to **Pending**. Once “Howard Doe” accepts the invite, the status should move to **Active**. The instructors will be accepting the invite on behalf of Howard Doe. You don’t need to wait until the status changes. Please continue with the lab.



User	Email	Status
George Doe owner self	wsuser42000@gmail.com	Active
Howard Doe	ws.catalog.user@gmail.com	Active

Items per page: 25 1-2 of 2 items

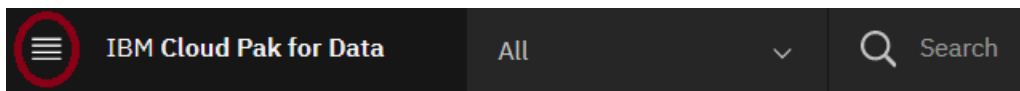
Step 2 – Add Collaborator to the Project

Now that the collaborator has been added to your IBM Cloud Account, you can add the collaborator to the project.

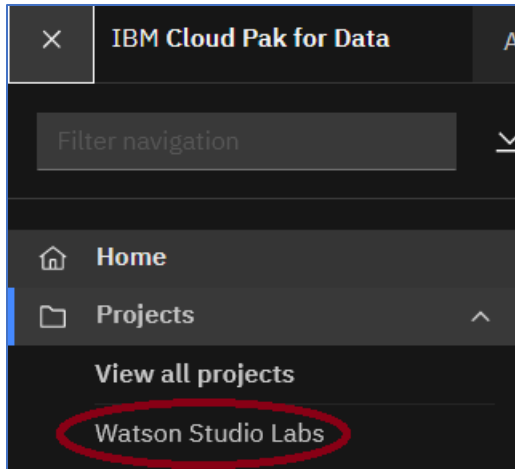
- Close the Identity and Access Management tab.



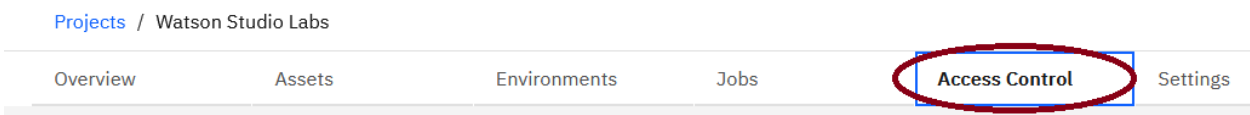
- Click on the **IBM Cloud Pak for Data** tab. Click on the  icon.



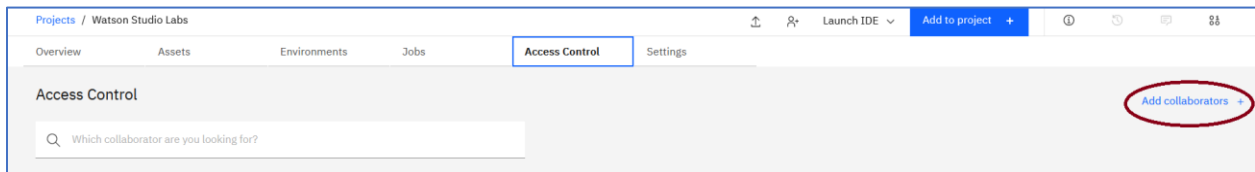
- Click on Watson Studio Labs.



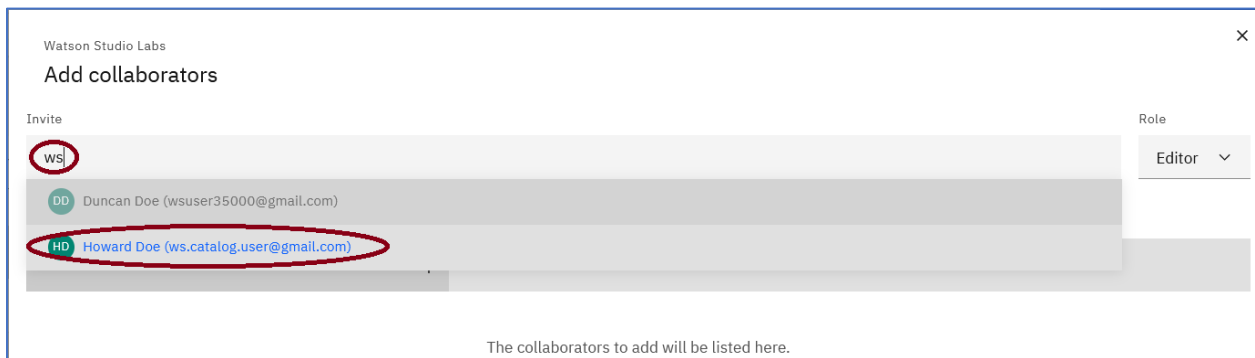
4. Click on the **Access Control** tab.



5. Click on **Add collaborators**.



6. For **Invite**, start entering **ws.catalog.user@gmail.com** and once a match has been made, click on the name in the list.



7. Switch to the role to **Viewer**.

Watson Studio Labs

Add collaborators


Invite

Add more people...

Role

Editor ▾

Add 1

Name	Email	Role	
 Howard Doe	ws.catalog.user@gmail.com	<div> Editor ^ Viewer Editor Admin </div>	✕

8. Click **Invite**.

Watson Studio Labs

Add collaborators


Invite

Add more people...

Role

Editor ▾

Add 1



Name	Email	Role	
 Howard Doe	ws.catalog.user@gmail.com	Viewer ▾	✕

Cancel Invite

9. The collaborator is added.

Access control

Which collaborator are you looking for?

Name	Email	Permission	Status
 Duncan Doe	wsuser35000@gmail.com	Admin ▾	Active
 Howard Doe	ws.catalog.user@gmail.com	Viewer ▾	Active


Add collaborators +

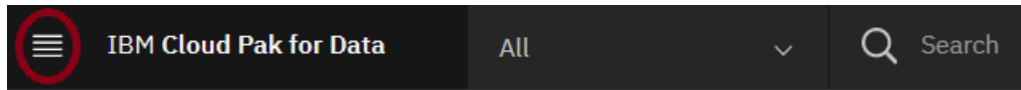
Research Topics

As you work on a data analysis project, you may need to do some research to help find a solution. Cloud Pak for Data provides a built-in capability, accessed via the **Gallery** option, that

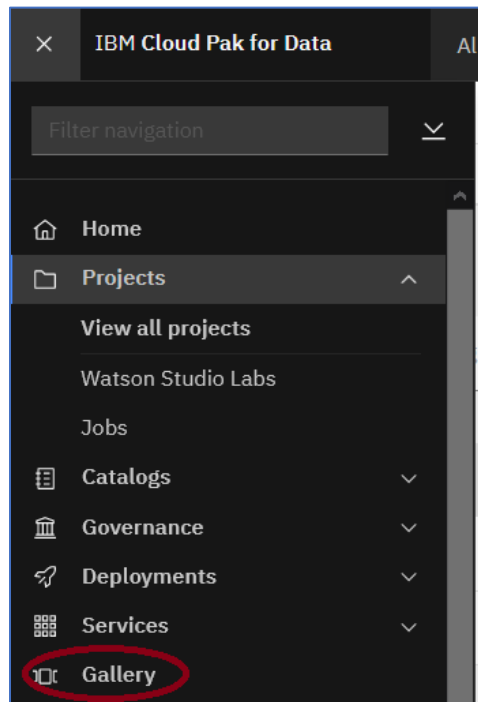
contains sample notebooks, sample datasets, governance content and sample projects. These are curated on a regular basis to provide up-to-date materials.

For the lab exercise, assume that you are interested in learning how to develop a Spark model. We will look for a sample notebook that demonstrates this capability and add this notebook to our project. **Note, we are doing this exercise for illustrative purposes on using the Gallery, and not for use in any subsequent lab.**

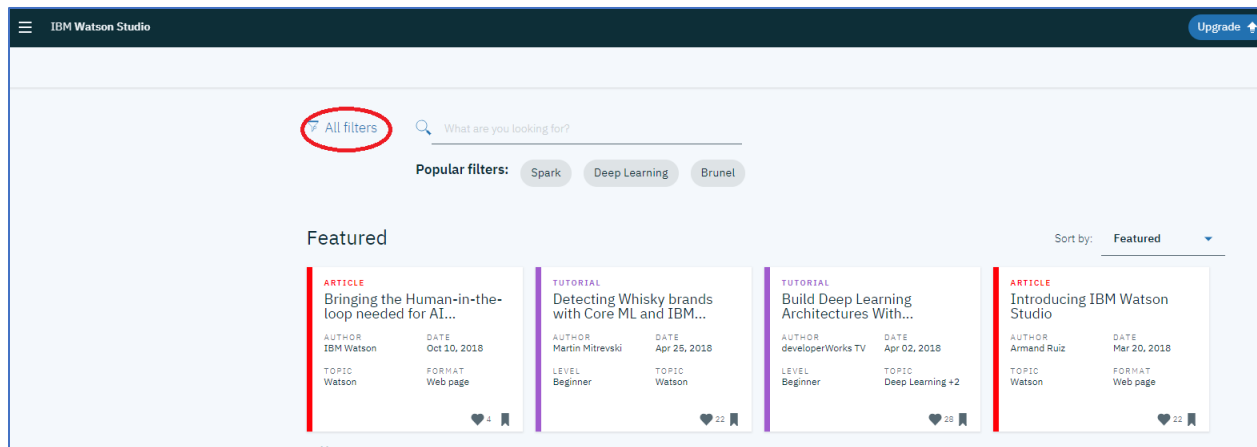
1. Click on the  icon.



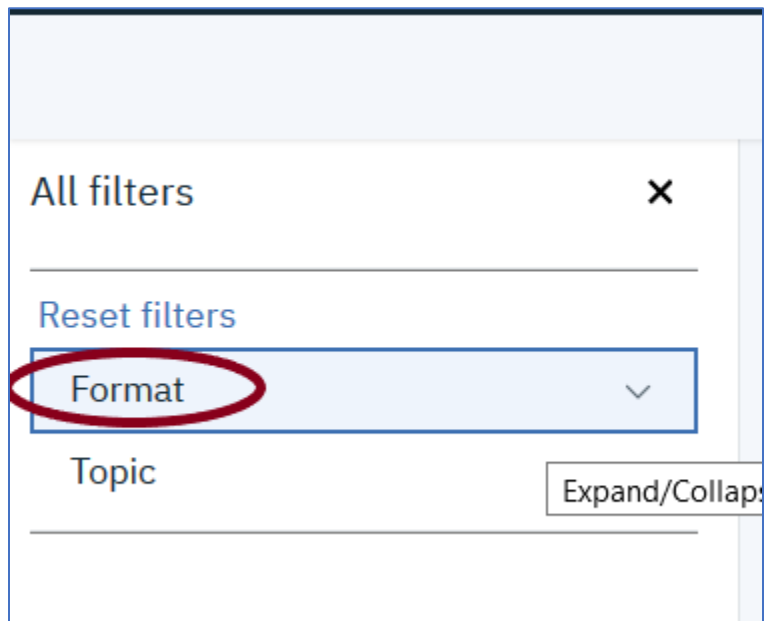
2. Click on **Gallery**.



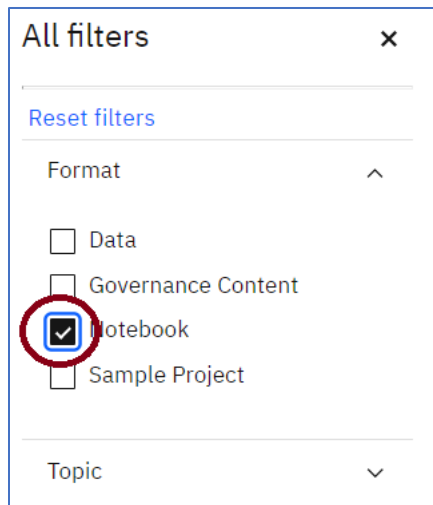
3. The Gallery is displayed. Click on **All filters**.



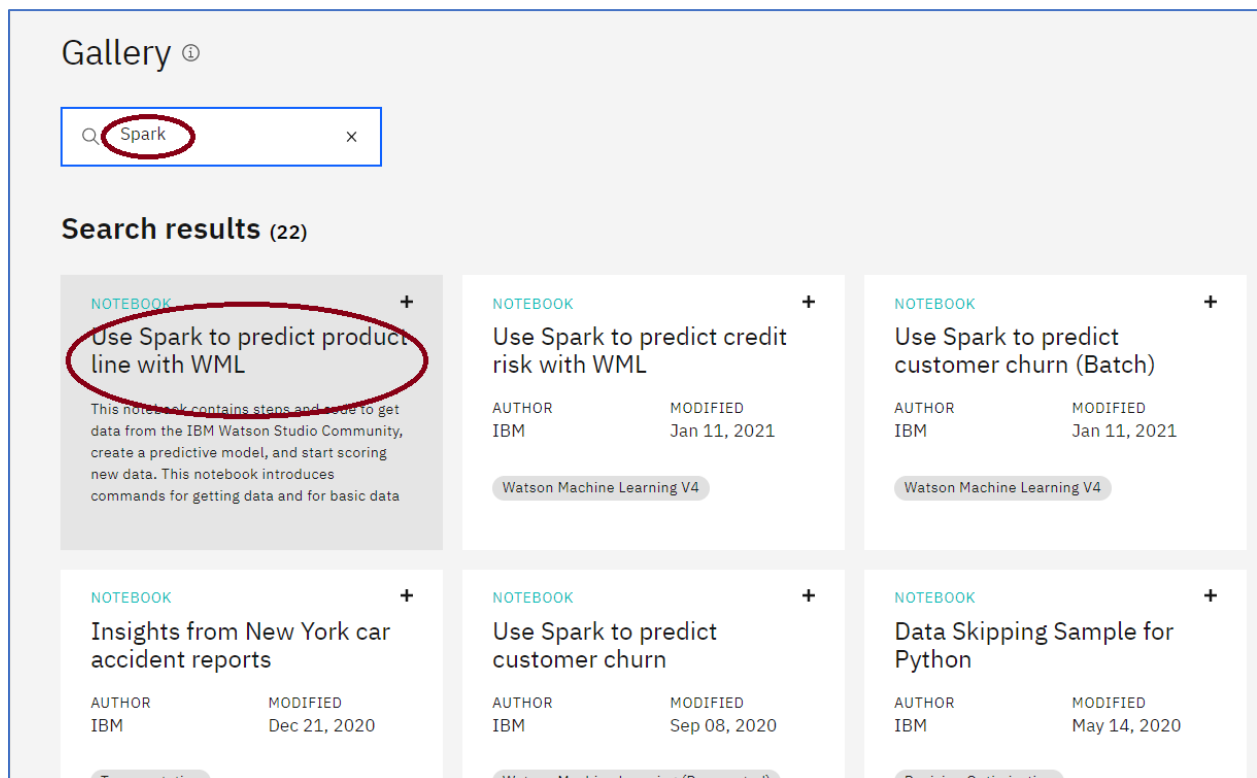
4. Click on **Format**.



5. Click on **Notebook**.




6. Enter **Spark** in the **Search** area. The Gallery view is updated. Locate the Gallery Card “Use Spark to predict product line with WML”. Hover the mouse over the card. The descriptive text provides a notebook summary. This notebook appears to be a good candidate for having code demonstrating the use of the Spark. Click on the **Gallery Card**.

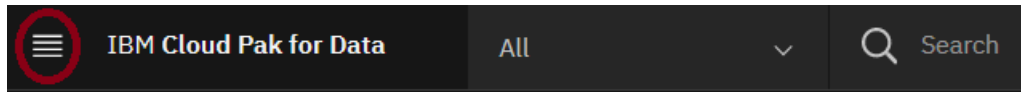


7. Here you can review the notebook documentation and then add it to the project if this notebook was a good starting point. We are not using this notebook for our labs so don't add it to the project. Our purpose was just to demonstrate the **Gallery** feature.

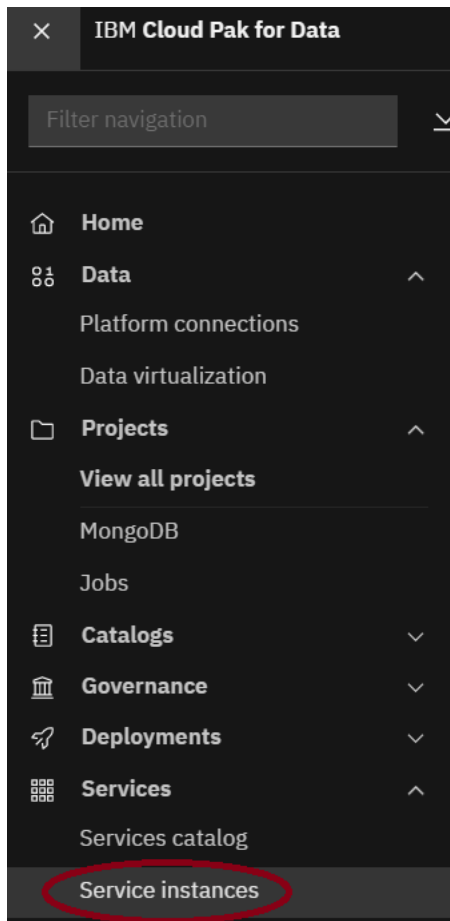
Provision Watson OpenScale

In this section, we will provision a Watson OpenScale service for use in a later lab, if one has not already been provisioned. Note, newer Cloud Pak for Data accounts are provisioned with a Watson OpenScale service.

1. Click on the  icon.



2. Click on **Services** and then **Service instances**



3. Check to see if **Watson OpenScale** is in the list of services. In the screenshot below, Watson OpenScale is an existing service. You have completed Lab-1! In the case where Watson OpenScale is not listed, go to the next step.

Service instances

To upgrade a service plan, first [upgrade](#) your IBM Cloud account. Then choose **Upgrade service** or **Manage in IBM Cloud** from the menu in the service's row.

Filter by: Resource Groups 2 x Locations ▼ None ▼ Product ▼ Service plan ▼

Find service instances

Name	Group	Location	Product	Plan	Status
CloudObjectStorage	Default	Global	Cloud Object Storage	Lite	Active
WatsonMachineLearning	Default	Dallas	Machine Learning	Lite	Active
KnowledgeCatalog	Default	Dallas	Watson Knowledge Catalog	Lite	Active
WatsonOpenScale	Default	Dallas	Watson OpenScale	Lite	Active
WatsonStudio	Default	Dallas	Watson Studio	Lite	Active

[Add service](#) +

4. Click on **Add service** if WatsonOpenScale is not listed.

Service instances

To upgrade a service plan, first [upgrade](#) your IBM Cloud account. Then choose **Upgrade service** or **Manage in IBM Cloud** from the menu in the service's row.

Filter by: Resource Groups 2 x Locations ▼ None ▼ Product ▼ Service plan ▼

Find service instances

Name	Group	Location	Product	Plan	Status
CloudObjectStorage	Default	Global	Cloud Object Storage	Lite	Active
WatsonMachineLearning	Default	Dallas	Machine Learning	Lite	Active
KnowledgeCatalog	Default	Dallas	Watson Knowledge Catalog	Lite	Active
WatsonOpenScale	Default	Dallas	Watson OpenScale	Lite	Active
WatsonStudio	Default	Dallas	Watson Studio	Lite	Active

[Add service](#) +

5. Click on **Watson OpenScale**

Services catalog

Find services

Category

- ☐ AI / Machine Learning
- ☐ Databases
- ☐ Mobile
- ☐ Analytics
- ☐ Integration
- ☐ Storage

AI / Machine Learning

Watson OpenScale
AI / Machine Learning

IBM Watson OpenScale is an enterprise-grade environment for AI infused applications that provides...

Lite • Free

Watson Assistant
AI / Machine Learning

Watson Assistant lets you build conversational interfaces into any application, device, or channel.

Lite • Free

Watson Studio
AI / Machine Learning

Embed AI and machine learning into your business. Create custom models using your own data.

Lite • Free

Watson Knowledge Catalog
AI / Machine Learning

Discover, catalog, and securely share enterprise data.

Watson Discovery
AI / Machine Learning

Add a cognitive search and content analytics engine to applications.

Language Translator
AI / Machine Learning

Translate text, documents, and websites from one language to another.

6. Make sure the Lite plan is selected and click **Create**.

Watson OpenScale

Author: IBM • Date of last update: Jan 7, 2021 • [Docs](#) • [API Docs](#)

Create About

Select a region

Select a region

Dallas

Pricing plan

Displayed prices do not include tax. Monthly prices shown are for country or region: United States

Plan	Features	Pricing
Lite	Lite - Maximum 5 deployed models to be monitored Maximum 2 protected features per model and 50,000 payload rows (cumulative) across 5 deployed models, per month for Fairness monitoring Maximum 50,000 rows of input (cumulative) across 5 deployed models, per month for active De-Biasing Maximum 20 transactions explained per month through the Explainability feature Maximum 50,000 feedback records (cumulative) across 5 models, per month for Accuracy monitoring Internal Postgres db instance that supports the above capacity; Data cannot be migrated from this instance; If upgrading to Standard plan and continuity is required from Lite plan, please provision your own instance of paid Postgres db from IBM Cloud catalog	Free

Create

You have completed Lab-1!

- ✓ Created a project
- ✓ Associated an existing Watson Machine Learning service instance with the project
- ✓ Added a collaborator to the project
- ✓ Demonstrated researching topics by searching the Gallery.
- ✓ Created an instance of Watson OpenScale if required

