



# DSX Setup Instructions for Pok Cognitive Masters

## Master the art of data science

Solve your toughest data challenges with the best tools and the latest expertise in a social environment built by data scientists.



Watch Now (2:31)

Sign Up For a Free Trial

**Now Available! Watson Machine Learning**

Machine Learning lets you create and train predictive analytics models. [Learn More](#)



# datascience.ibm.com

## Try Watson Data Platform

### Powered by IBM Cloud

Build smarter applications and quickly visualize, share and gain insights from your data with Watson Data Platform.

 IBM Data Science Experience

Solve data challenges in a social environment built by data scientists.

[Remove](#)

### Try our available applications



IBM Data Catalog

Securely discover, catalog, and govern enterprise data.

[Select](#)



IBM Data Refinery

## Register for IBM Cloud

To sign up for IBM Data Science Experience, you must first register for IBM Cloud.

### Create your IBM Cloud Account

Email

### Accept the terms

☐ I accept the [IBM Data Science Experience terms](#), the [IBM Cloud privacy policy](#) and [IBM Cloud terms](#).

[Next](#)

## Already have an IBM Cloud account?

Use it to sign up for IBM Data Science Experience.



Already have an IBM Cloud account? [Log in](#)

Email\*



dustin.van.stee@gmail.com

First Name\*



Dustin

Last Name\*



VanStee

Company

IBM

Country or Region\*

United States



Password\*



••••••••



Keep me informed of products, services, and offerings from IBM companies worldwide.

☒ By email

By clicking Create Account, I accept the [IBM Cloud privacy policy](#) and [IBM Cloud terms](#).

Create Account

Do annoying captcha  
Verify email



Hello Dustin,

Thank you for signing up for IBM Cloud! Confirm your account to get started.

**Confirm Account**



# Try Watson Data Platform

## Powered by IBM Cloud

Build smarter applications and quickly visualize, share and gain insights from your data with Watson Data Platform.

 IBM Data Science Experience

Solve data challenges in a social environment built by data scientists.

[Remove](#)

## Try our available applications



IBM Data Catalog

Securely discover, catalog, and govern enterprise data.

[Select](#)

## Register for IBM Cloud

To sign up for IBM Data Science Experience, you must first register for IBM Cloud.

### Create your IBM Cloud Account

Email

### Accept the terms

☐ I accept the [IBM Data Science Experience terms](#), the [IBM Cloud privacy policy](#) and [IBM Cloud terms](#).

[Next](#)

## Already have an IBM Cloud account?

[Use it to sign up for IBM Data Science Experience.](#)

# Select Organization and Space

Confirm your IBM Cloud organization and space information below.

Or create new organization and space

Select IBM Cloud account

IBM



IBM Cloud Organization

dustin.van.stee@gmail.com



IBM Cloud Space

dev



Continue



[Get started](#) ▾



## Welcome Dustin!

IBM Data Science Experience is part of Watson Data Platform.

[Try out](#) other Watson Data Platform apps.

### Get started with key tasks



New project



New notebook



New SPSS Modeler flow



New model



New streams flow



Hide ▾

## New project

### Define project details

Name

Example Project

84

Description

Project description

3000

### Choose project options

☐ Restrict who can be a collaborator ⓘ

### Define storage

#### ① Select storage type

☐ Object Storage (Swift API) ☒ IBM Cloud Object Storage

#### ② Add

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

#### ③ Refresh

### Define compute engine

#### ① Select Spark service

Spark service

**Add**

Add IBM Analytics for Apache Spark, then return to this page and click Refresh.

#### ② Refresh



# Cloud Object Storage

Existing

New

## Cloud Object Storage

IBM Cloud Object Storage is a highly scalable cloud storage service, designed for high durability, resiliency and security. Store, manage and access your data via our self-service portal and RESTful APIs. Connect applications directly to Cloud Object Storage and integrate IBM Cloud Services

### Features

#### Storage for IBM Cloud native applications such as Watson and developer services

Leverage Cloud Object Storage to build your cloud applications. SDKs are available and include support a common set of S3 API functions for programmatic access so dev ops teams can easily connect applications such as Watson and developer services to scalable cloud storage and integrate with IBM Cloud Services.

#### IAM Policies - Bucket level access management

IBM Identity and Access Management (IAM) integration allows for granular access control at the bucket level using role based policies.

#### Encryption for data at rest (SSE, SSE-C, SSE-Key Protect)

All data is encrypted at-rest and in-flight by default, and with additional option of using customer provided keys. Keys can either be self-managed or by using IBM Key Protect\*. (\*Available in US-south (Dallas) and EU-gb (London) Regional service only)

#### Regional and Cross Region resiliency options

Select the best resiliency option for your data workload. Choose "Cross Region" to store unstructured data across three or more regions, or choose "Regional" resiliency to store your data within a single region.

#### Data storage classes for Active, Less Active, Archive and Dynamic workloads

Choose storage classes for frequently accessed data, occasionally accessed data and long-term data retention with Standard, Vault and Cold Vault. Or, choose Flex class for dynamic data access needs that fluctuate month to month.

### Lite and pay-as-you-go plans

Pricing Plan: Monthly Process shown above reflect the: **United States**

Plan	Features	Pricing
<div><div></div> Lite</div>	1 COS Service Instance Storage up to 25 GB/mo. Up to 20,000 GET requests/mo. Up to 2,000 PUT requests/mo. Up to Data Retrieval 10 GB/mo. Up to 5GB Public Outbound Applies to aggregate total across all storage bucket classes	Free

The Lite service plan for Cloud Object Storage includes Regional and Cross Regional resiliency, flexible data classes, and built in security.

<div><div></div> Standard</div>	There is no minimum fee, so you pay only for what you use.	-
---------------------------------	--	---

Cancel

Create

# Confirm Creation

Plan

Lite



Resource group

Default



Service name

cloud-object-storage-jd

Cancel

Confirm



## New project

### Define project details

Name

Example Project

84

Description

Project description

3000

### Choose project options

☐ Restrict who can be a collaborator ⓘ

### Define storage

#### ① Select storage type

☐ Object Storage (Swift API) ☒ IBM Cloud Object Storage

#### ② Add

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

#### ③ Refresh

### Define compute engine

#### ① Select Spark service

Spark service

Add

Add IBM Analytics for Apache Spark, then return to this page and click Refresh.

#### ② Refresh

## New project

### Define project details

Name

Example Project

84

Description

Project description

3000

### Choose project options

☐ Restrict who can be a collaborator ⓘ

### Define storage

#### ① Select storage type

☐ Object Storage (Swift API) ☒ IBM Cloud Object Storage

#### ② Add

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

#### ③ Refresh

### Define compute engine

#### ① Select Spark service

Spark service

Add

Add IBM Analytics for Apache Spark, then return to this page and click Refresh.

#### ② Refresh



# Apache Spark

Existing

New

## Apache Spark

Apache Spark is an open source cluster computing framework optimized for extremely fast and large scale data processing, which you can access via the newly integrated notebook interface IBM Analytics for Apache Spark. You can connect to your existing data sources or take advantage of the on-demand big data optimization of Object Storage. Spark plans are based on the maximum number of executors available to process your analytic jobs. Executors exist only as long as they're needed for processing, so you're charged only for processing done.

### Features

#### Incredibly Fast

Apache Spark delivers 100x the performance of Apache Hadoop for certain workloads because of its advanced in-memory computing engine.

#### Easy to Use and Powerful

Apache Spark's Streaming and SQL programming models backed by MLlib and GraphX make it incredibly easy for developers and data scientists to build apps that exploit machine learning and graph analytics. Because the service is 100% compatible with Apache Spark, developers can build their apps and run them against the IBM managed service to benefit from operational, maintenance, and hardware excellence.

#### Convenient Data Storage

Object Storage enables a convenient way to upload your data from a file for immediate use by your Spark instance. You can set up Object Storage directly from the Spark service interface.

**Pricing Plan:** Monthly Process shown above reflect the: **United States**

Plan

Features

Pricing

● Lite

2 Spark Executors

Free

An entry level plan to run programs using up to 2 Spark executors

Terms

Cancel

Create



# Confirm Creation

Plan

Lite



Resource group

Default



Service name

cloud-object-storage-jd

Cancel

Confirm



## New project

### Define project details

Name

Example Project

84

Description

Project description

3000

### Choose project options

☐ Restrict who can be a collaborator ⓘ

### Define storage

#### ① Select storage type

☐ Object Storage (Swift API) ☒ IBM Cloud Object Storage

#### ② Add

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

#### ③ Refresh

### Define compute engine

#### ① Select Spark service

Spark service

**Add**

Add IBM Analytics for Apache Spark, then return to this page and click Refresh.

#### ② Refresh



## New project

## Define project details

Name

Example Project

84

Description

Project description

3000

## Choose project options

☐ Restrict who can be a collaborator ⓘ

## Define storage

## ✓ Select storage type

- ☐ Object Storage (Swift API) ☒ IBM Cloud Object Storage

Target Cloud Object Storage Instance

cloud-object-storage-bh ▼

## Define compute engine

## ✓ Select Spark service

Spark service

Spark-cp ▼

⚠ If you associate the same Spark service with multiple projects, the Spark history server will display job history information for all the projects.

Cancel

Create



## Overview

## Assets

## Bookmarks

## Deployments

## Collaborators

## Settings



What assets are you looking for?

## ▼ Data assets

0 assets selected.

☐

NAME

TYPE

## SERVICE

CREATED BY

LAST MODIFIED

## ACTIONS

you currently have no data assets

## ▼ Notebooks

**(+) New notebook**

NAME

SHARED

**SCHEDULED**

STATUS

LANGUAGE

LAST EDITOR

LAST MODIFIED

## ACTIONS

you currently have no notebooks



## New notebook

Blank

From file

From URL

**Name\****Type Notebook Name here***Description***Type your Description here***Notebook URL\****Remote notebook served by HTTP or HTTPS*<http://ibm.biz/pok-masters-mnist> LAB1<http://ibm.biz/pok-masters-sql> LAB2**Spark service\***

Spark-cp



Associate this notebook with the Spark Service of your choice.