

Prerequisites

If you want to run all exercises "live" and complete this workshop effectively, there are some pieces of software that you will need to install on your computer. Remember that it is not mandatory at all, but you will be limited to a normal reading if you decide not to install them.

The following steps will guide you through the installation of the necessary modules you need to obtain and install, apart from providing some education links in case that you are not familiar with them. If you don't want or are not allowed to install software on your laptop, you may follow these instructions to provision and customize a minimal *lab client* that you can use during the workshop as your working environment.

1.- bash

You will need to execute commands using your terminal and they use the conventions of the bash shell. All Linux distributions and MacOS include bash or a compatible command-line interface. On Windows, you can try to setup the Windows-Subsystem for Linux (WSL) on your computer. If it is not possible to install bash, you will have to figure out an alternative way of issuing the commands you will see during all the modules. If you are not familiar with bash, it is recommendable to visit online resources to get a basic knowledge. Here you can see one of the many examples you could find.

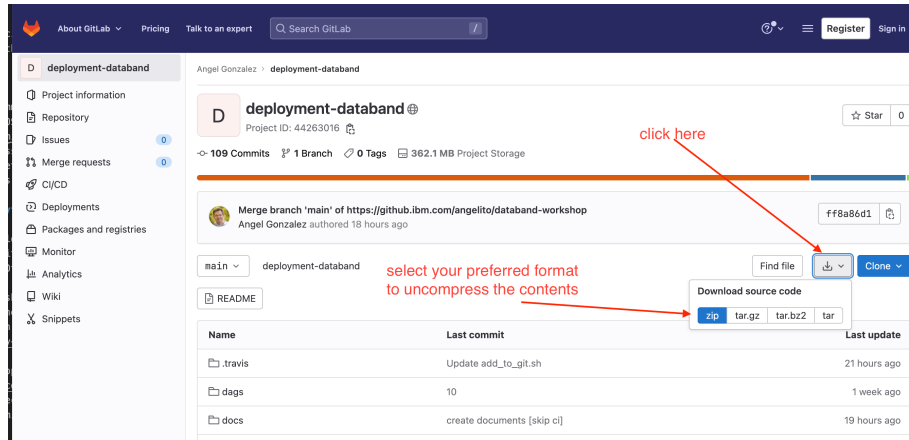
2.- Jupyter Notebooks, Python and git

This workshop has been developed with Jupyter Notebooks technology in order to maximize the user experience. While it is not strictly necessary, you should install it on your local computer. Otherwise, you will have to copy and paste the bash commands and execute them in a terminal.

There are parts of the workshop that have been developed with python which is a requirement of several technologies we will cover. There is no need to be a python programmer but a basic knowledge is highly recommendable. Please review the introduction of any tutorial you may find in internet. The course mentioned above may also be a good start.

The contents of this workshop are placed on a github repository, mirrored in a gitlab repository. Displaying and reading the web pages online is fine but if you want to work comfortably, it is highly recommendable to install git and clone the

repository on your local computer. If you are not familiar with git you should, at least download the repository and open it with jupyter.



Note that all the contents are also available in PDF and Word format, in case that no other option fits for you:

- Documentation in gitlab
- Documentation in github

3.- bash kernel for Jupyter

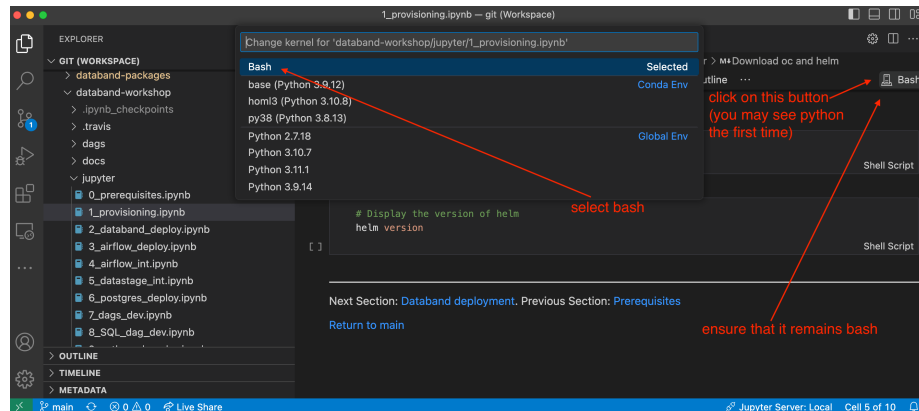
Jupyter is well-know for python development but it is also possible to extend its default kernel to include **bash** capabilities. Once you have installed Jupyter, you will need to run a very simple command in your **bash** terminal:

```
pip install bash_kernel  
python -m bash_kernel.install
```

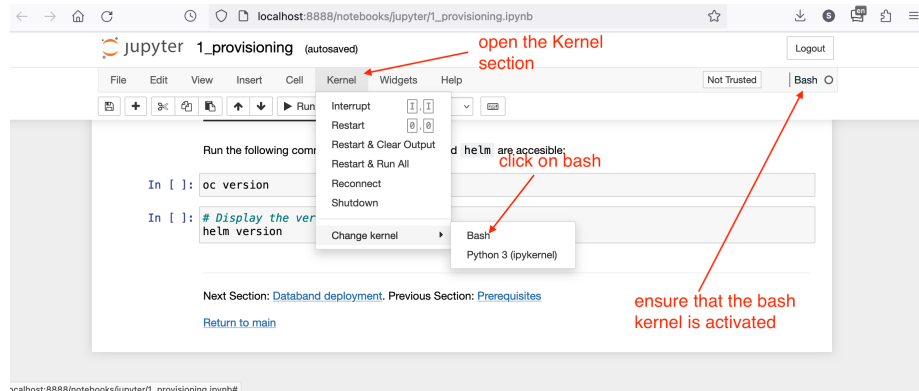
Notice that instaling Jupiter Notebooks implies an underlying python environment, as indicated during the installation instructions mentioned above.

By the way: there are several variations of bash kernels for jupyter. Take a look at the web site where all possible kernels are listed, search for bash and pick one that may fit better for you.

After the installation of the bash kernel, you need to activate it just one time if you want to click on the cells to run the commands. You can do it this way if you are using Microsoft Visual Studio Code:



Alternatively, if you use the native jupyter notebook tool:



4.- Docker

We will run some operations with containers like tagging or transferring them to a different repository. In order to do that we will need docker or podman as an alternative. A very convenient way of getting docker is downloading docker desktop. Alternatively, you may want to install docker with the command-line. In that case, you may want to run the following commands on RHEL8:

First of all, get a clean python environment:

```
sudo yum install curl wget git
sudo yum install gcc openssl-devel bzip2-devel libffi-devel zlib-devel
sudo yum install python3.8
sudo yum install python38-devel
```

Then download, install and setup docker:

```
sudo yum install -y yum-utils
```

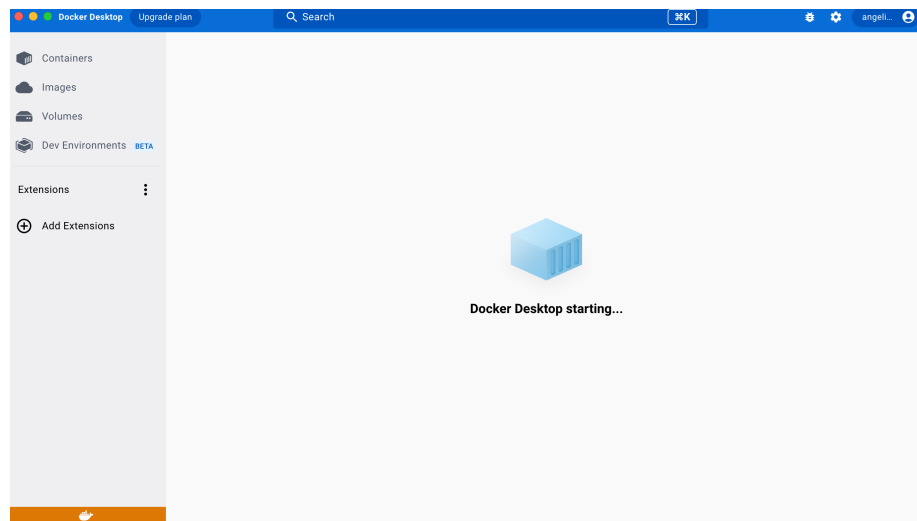
```
sudo yum-config-manager --add-repo https://download.docker.com/linux/rhel/docker-ce.repo
sudo sed -i 's~/rhel/~centos/~g' /etc/yum.repos.d/docker-ce.repo
sudo yum --noplugins install docker-ce docker-ce-cli containerd.io docker-compose-plugin
```

```
sudo systemctl enable --now docker
sudo systemctl start docker
```

```
sudo usermod -aG docker $USER
newgrp docker
```

```
sudo pip3 install docker-compose
```

Don't forget to start docker from the user interface if you installed docker desktop:



Verify that docker is running by displaying the version

```
docker version
```

In case you are not familiar with Docker, here is a good introductory course that will help you to begin your journey into the world of containers and OpenShift.

5.- oc and helm

The OpenShift command-line interface (oc) and the helm utility are also necessary to run some commands. An alternative to oc is kubectl but this workshop has not been tested with it.

It may be more convenient to defer the installation of these utilities after having provisioned the hardware because the cluster contains direct links to the downloads you need, but you may want to do it now. In that case, you will need

a RedHat account and follow these instructions.

The download that you need to choose will be similar to this:

The screenshot shows the Red Hat Customer Portal interface. At the top, there's a navigation bar with links for Subscriptions, Downloads, Containers, and Support Cases. Below this, the breadcrumb trail reads: Downloads > Red Hat OpenShift Container Platform Packages > openshift-clients-redistributable. The main heading is 'openshift-clients-redistributable'. Underneath, there are dropdown menus for 'Architecture' (set to x86_64) and 'Version' (set to 4.12.0-202301312133.p0.gb05f7d4.assembly.stream.el9 [latest]). Below these are tabs for 'Details' and 'Change Log'. A note indicates the project is OCPBUGS and the component is oc OpenShift Client binaries for Linux, Mac OSX, and Windows. A link 'View more details' is provided. The 'Available From' section shows repos based on active subscriptions, with a 'Show All' link. A table lists the product variant and the repository label. The 'Download' section shows the package name and a 'Download Now' button with a download icon. Below the package name, the SHA-256 hash is displayed, and the file size is 116 MB.

Subscriptions Downloads Containers Support Cases

Red Hat Customer Portal Products & Services Tools Security Community

Search English All Red Hat

Downloads > Red Hat OpenShift Container Platform Packages > openshift-clients-redistributable

openshift-clients-redistributable

Architecture: Version:

x86_64 4.12.0-202301312133.p0.gb05f7d4.assembly.stream.el9 [latest]

Details Change Log

[Maintainer] project: OCPBUGS, component: oc OpenShift Client binaries for Linux, Mac OSX, and Windows

[View more details](#)

Available From

Repos shown are based on your active subscriptions. [Show All](#)

Product (Variant, Version, Architecture)	Repo Label
Red Hat OpenShift Container Platform 4.12 for RHEL 9 x86_64	rhocp-4.12-for-rhel-9-x86_64-rpms

Download

Package - openshift-clients-redistributable-4.12.0-202301312133.p0.gb05f7d4.assembly.stream.el9.x86_64.rpm

SHA-256: a7496ae77de3d6cd94e55445b84dc6f1ed5cd1c043f6e0b905ae8d46285d109e

Download Now

116 MB

Then, just follow the installation procedure described here to get **helm**

At the time of writing these instructions, these were the direct links used during the workshop:

- https://mirror.openshift.com/pub/openshift-v4/x86_64/clients/ocp/stable-4.10/
- https://mirror.openshift.com/pub/openshift-v4/x86_64/clients/helm/latest/

You can try and download the executables for your system from those links but remember that they may have changed now.

6.- Databand images

There are several media packages of databand and we need the helm chart version of the software for this workshop. This section describes the steps that an IBMer would follow to download the right package. Additionally, this is the point where I need you to encourage to adhere to the terms and conditions of a licensed software like databand. In simple words: do not distribute these images illegally or use them for other purposes other than your own education during this workshop.

- IBMers need to be connected to the IBM intranet or use the Cisco Secure Client in order to access the IBM Internal DSW Downloads.
- Read the authorized and not authorized use of the software you will download. If you agree, you can go on
- Search for databand like in the following picture:

The screenshot shows the IBM Internal DSW Downloads search interface. The page has a header with the IBM logo and a search bar. Below the header, there are tabs for 'Search for software', 'Popular downloads', 'Download history', and 'Request assistance'. The 'Search for software' tab is selected. On the left, there is a sidebar with 'Search options', 'Your previous searches', and 'Related Links'. The main content area is titled 'Overview' and contains information about the download search functionality. Below the overview, there is a 'Search options' section with tabs for 'Text', 'Brand', and 'Part number'. The 'Text' tab is selected. Under 'Find by text', there is a search bar with 'Databand' entered. A dropdown menu shows suggestions: 'Databand', 'Data Observability by Databand Self-Hosted V1.0.x English eAssembly', and 'IBM Data Observability by Databand Self-Hosted V1.0.x English eAssembly'. Red arrows point to the search bar and the first suggestion.

type here Databand

select this

Find by text

Product name

Databand

Data Observability by Databand Self-Hosted V1.0.x English eAssembly

IBM Data Observability by Databand Self-Hosted V1.0.x English eAssembly

Type all or part of a product name and click Search or click one of the pop-up suggestions. You can select additional filters below.

- Look for the helm charts version and download it:

Search results

eAssemblies (1) Images (33)

Brand	# eAssemblies	# Images
Analytics - Platform	1	3

eAssembly	Description	Date posted	# Images	Size (MB)
G06XGEN	IBM Data Observability by Databand Self-Hosted V1.0.x English eAssembly	9/15/2022	3	1,637.5

[Multi-product package terms](#)

Important! This software is licensed as an eAssembly. You may not transfer or remarket individual products from it as this would violate the license terms applicable to the eAssembly. Your acceptance of the license terms applicable to the eAssembly is a precondition to your downloading and using the eAssembly.

<input type="checkbox"/>	Image	Description	Date posted	Size (MB)
<input type="checkbox"/>	MOB8REN	IBM Data Observability by Databand Self-Hosted v1.0.19.0 - Readme English	2/20/2023	0.5
<input type="checkbox"/>	MOB8SEN	IBM Data Observability by Databand Self-Hosted v1.0.19.0 - Docker Compose Release English	2/20/2023	722
<input checked="" type="checkbox"/>	MOB8TEN	IBM Data Observability by Databand Self-Hosted v1.0.19.0 - Helm Charts Release English	2/20/2023	915

check this (arrow pointing to MOB8TEN)

tick here (arrow pointing to checkbox of MOB8TEN)

[View details](#) [License agreement](#) [Download estimate](#) [eAssembly](#)

[View details](#) [License agreement](#) [Download estimate](#) [eAssembly](#)

[View details](#) [License agreement](#) [Download estimate](#) [eAssembly](#)

☒ I agree ☐ I do not agree

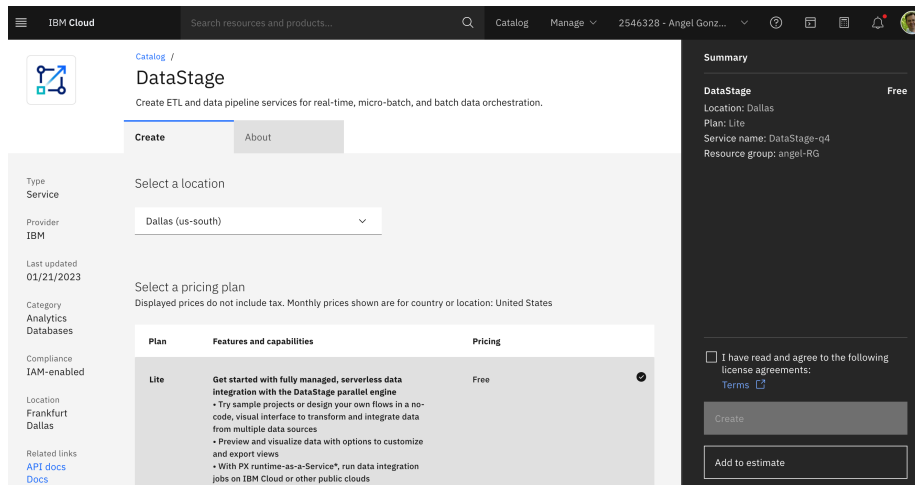
By clicking the "I agree" button, you agree that (1) you have had the opportunity to read and understand the above license agreement(s) and multi-product package terms, if any, and (2) terms of the license agreement(s) govern this transaction. If you do not agree with the terms of the agreement(s), you will be unable to download the software.

[Download now](#)

After pressing the blue button, the download will start and you will get a file similar to `databand-1.0.19-helm-chart.tar.gz` (960MB). Needless to say, names and sizes will change with upcoming versions.

7.- DataStage

If you want to exercise with the integration of DataStage, you will need to create an instance in the IBM Cloud. The lite (free) plan is perfectly fine for this workshop.



8.- Airflow

It is highly recommendable, but not strictly necessary, to master the basic concepts of Apache Airflow. There are plenty of tutorials in internet and visiting any introduction would be enough to follow the workshop. Here is a good example.

9. OpenShift

You need to provision an OpenShift cluster as a part of the hands-on workshop. If you are not familiar with OpenShift, this may be your starting point but, please, review the next chapter Hardware Provisioning for a full description of this pre-requisite.

Next Section: Hardware Provisioning

[Return to main](#)