



[Course](#) > [1 | Introduction and Overview](#) > [Environment Setup](#) >
Provisioning Azure Data Science VMs

Provisioning Azure Data Science VMs

🔖 Bookmark this page

The Microsoft Data Science Virtual Machine is an Azure virtual machine (VM) image pre-installed and configured with several popular tools that are commonly used for data analytics and machine learning.

Before you can create a Microsoft Data Science Virtual Machine, you must have the following:

- An Azure subscription: To obtain one, see [Get Azure free trial](#).
- An Azure storage account. Alternatively, the storage account can be created as part of the process of creating the VM if you do not want to use an existing storage account.

In this section, we will go through the provisioning of Azure Data Science VM based on Linux (Ubuntu), configuring CNTK and Jupyter notebook, and tunneling from your local machine to the notebook server.

Provisioning Azure Data Science Virtual Machine

1. Go to <https://portal.azure.com> and sign in using your account.
2. Click **New** and type **Data Science** on the Search box.
3. Select **Data Science Virtual Machine for Linux (Ubuntu)**.
4. Provide the requested details and click **OK**.
5. Select a size and click **Select**.
If you have an Azure subscription (not a trial account), you can go for the NC series which has one or more GPU (they need to have HDD as the disk type).
6. Accept the default options and click **OK**.
7. Click **Purchase** to provision your VM.
8. When the VM is Provisioned, the VM dashboard will open.

9. Click **Connect** and note the command to connect to the VM using SSH. It will be something like `ssh username@123.123.123.123`.

Now you will connect to the VM to configure CNTK and run the Jupyter notebook server. You will then tunnel to the VM using the correct port to open the Jupyter notebook.

Connecting to the Azure Data Science VM using SSH

You can use any SSH client. In this example, we will use PuTTY, developed originally by Simon Tatham for the Windows platform.

1. Download PuTTY and run it.
2. Enter the Host Name. This is from the command to connect you noted earlier, minus the `ssh` in the beginning. It will be something like `username@123.123.123.123`.
3. Expand **SSH** in the **Category** and select **Tunnels**.
4. Enter **8000** for Source port, **localhost:8888** for Destination, and click **Add**.
5. Select **Session** in the **Category**, enter a name for **Saved Sessions** and click **Save**.
6. Click **Open** to connect to the VM and click Yes on the dialog box.
7. Enter the password.
8. You are now connected to the Azure Data Science VM you provisioned earlier.

Configuring CNTK on the Azure Data Science VM

1. On the VM command prompt, enter the following command:
`source activate py35`
This will activate the py35 environment.
2. On the py35 environment, enter the following command:
`pip install <url>`
Where `<url>` is the corresponding wheel (.whl) file to install CNTK.
You must select the correct installation from the list here, and substitute `<url>` during the installation.
3. So for a GPU installation with Python 3.5, the command should be as follows:
`pip install https://cntk.ai/PythonWheel/GPU/cntk-2.0-cp35-cp35m-linux_x86_64.whl`
Or if you need super user permission:
`sudo /anaconda/envs/py35/bin/pip install https://cntk.ai/PythonWheel/GPU/cntk-2.0-cp35-cp35m-linux_x86_64.whl`

NOTE: You might need to upgrade pip before installing CNTK.

```
pip install --upgrade pip
```

Or if you need super user permission:

```
sudo /anaconda/envs/py35/bin/pip install --upgrade pip
```

Running Jupyter Notebook Server

1. Navigate to the local directory that you want to be the root folder for the Jupyter notebook.
2. Enter the following command:
`jupyter notebook`
3. Type **A**, and then type **Q**. The notebook server is now running. Note the URL that is displayed. Now, let's connect from your local machine using a browser.
4. On your local machine, open a browser.
5. Enter the URL noted previously, replace the port 8888 with 8000. It will be something like `http://localhost:8000/?token=9cd2f689e8d8e289e63c38b5ac05166d5d906173c2aa98cb`.
6. You can open and run any notebook (.ipynb) file or create a new notebook. The notebook is running on the Azure Data Science VM, but you can view it from your local machine.
7. To create a new notebook, click **New** and select a language.
8. You can now run the newly created notebook.


Discussion

[Show Discussion](#)

Topic: Mod1-3 Environment Setup / Provisioning Azure Data Science VMs

© All Rights Reserved



 English ▼

© 2012–2017 edX Inc. All rights reserved except where noted. EdX, Open edX and the edX and Open edX logos are registered trademarks or trademarks of edX Inc. | 粵ICP备17044299号-2

POWERED BY
OPENedX

