

Module 0 - Lab environment setup with a pre-installed virtual machine

The following instructions enables learners to prepare their lab environments for the modules that follow. Please run through these instructions prior to starting Module 1.

Time to complete: It takes around 5 minutes to perform the steps below and initiate the automated setup scripts. The scripts may take an hour or more to complete.

Note: These instructions are designed to be used in the pre-installed virtual machine provided for the course.

Requirements

Before starting setup, you will need an Azure Account with the ability to create an Azure Synapse Workspace.

Important note if using an Azure Pass subscription

If you are using an account for which you have previously redeemed an Azure Pass subscription that has expired, your account may be associated with multiple Azure subscriptions with the same name (*Azure Pass - Sponsorship*). Before starting the setup steps, ensure that only the most recent *active* subscription of this name is enabled by following these steps:

1. Open the Azure portal at <https://portal.azure.com> and sign in using the account associated with your subscription.
2. On the portal toolbar at the top of the page, select the **Directories and Subscriptions** button.
3. In the **Default subscriptions filter** drop-down list, *de-select* any **(Disabled) Azure Pass - Sponsorship** subscriptions, and ensure that only the active **Azure Pass - Sponsorship** subscription that you want to use is selected.

Setup steps

Perform the following tasks to prepare your environment for the labs.

1. Use the Windows **Search** box to search for **Windows PowerShell**, and then run it as an administrator.

Note: Make sure you run **Windows Powershell**, not Windows PowerShell ISE; and be sure to run it as Administrator.

2. In Windows PowerShell, run the following commands to download the required course files. This may take a few minutes.

```
mkdir c:\dp-203

cd c:\dp-203

git clone https://github.com/microsoftlearning/dp-203-data-engineer.git
data-engineering-ilt-deployment
```

3. In Windows PowerShell, run the following command set the execution policy so you can run a local PowerShell script file:

Note: You may need to run Windows PowerShell in Administrator. To complete this, right click on Windows PowerShell and select "Run as Administrator".

```
Set-ExecutionPolicy Unrestricted
```

Note: If you receive a prompt that you are installing the module from an untrusted repository, enter **A** to select the *Yes to All* option.

4. In Windows PowerShell, use the following command to change directories to the folder containing the automation scripts.

```
cd C:\dp-203\data-engineering-ilt-  
deployment\Allfiles\00\artifacts\environment-setup\automation\
```

5. In Windows PowerShell, enter the following command to run the setup script:

```
.\dp-203-setup-Part01.ps1
```

Note: This first script creates and prepares most of the Azure accounts we need (except for Cosmos DB). It takes around 10-15 minutes to complete.

6. When prompted to sign into Azure, and your browser opens; sign in using your credentials. After signing in, you can close the browser and return to Windows PowerShell, which should display the Azure subscriptions to which you have access.
7. When prompted, sign into your Azure account again (this is required so that the script can manage resources in your Azure subscription - be sure you use the same credentials as before).
8. If you have more than one Azure subscription, when prompted, select the one you want to use in the labs by entering its number in the list of subscriptions.
9. When prompted, enter a suitably complex password for the SQL Database (make a note of this password in case you need it later).

While the script runs, your instructor will present the first module of the course. Your environment should be ready for you when it's time to start the first lab.

Note: The first script will take around 10 to 15 minutes to complete. The first script will create the Azure resources with randomly generated names. If the script appears to "stall" (no new information is displayed for 10 minutes) press ENTER and check for any error messages - often the script will continue without any issues. In some rare cases, an identical resource name may already be in use, there may be

capacity constraints for specific resources in the randomly selected region, or transient network issues may occur; causing an error. If this happens, use the Azure portal to delete the **data-engineering-synapse-xxxxxx** resource group created by the script and re-run the script.

If an error indicating that you must supply the **tenant ID** for your Azure pass subscription is displayed, ensure you have followed the instructions in the **Requirements** section above to enable only the Azure Pass subscription that you want to use. You will have **2 additional script** files to run in later labs.