

DAT216x

Delivering a Relational Data Warehouse

Lab 01 | Designing a Relational Data Warehouse Schema

Overview

In this lab, you will provision a Microsoft Azure Virtual Machine (VM) which will be used by all three labs in this course. Once the VM is provisioned, you will complete the setup required to support the labs, and then explore the schemas of the **AdventureWorks** sample databases.

Note: The three labs in this course are accumulative. You cannot complete the following two labs if this lab has not been successfully completed.

What You'll Need

To complete this lab, you will need the following:

- High-speed and reliable internet connectivity (for remote connections to the VM)
- A second monitor is preferred (for the Remote Desktop connection)
- A Microsoft account (such as one used for Outlook.com, Hotmail, or other Microsoft services)
- A Microsoft Azure subscription
- The lab files for this course (available for download from GitHub, within the VM)

Creating a Free Trial Azure Subscription

If you already have an Azure subscription, you can skip this section. Otherwise, follow these steps to create a free trial subscription. You will need to provide a valid credit card number for verification, but you will not be charged for Azure services—for more information, refer to <http://azure.microsoft.com/pricing/free-trial-faq/>. Note that the free trial is not available in all regions.

If you already have a Microsoft account that has not already been used to sign up for a free Microsoft Azure trial subscription, you're ready to get started. If not, don't worry—just create a new Microsoft account at <https://signup.live.com>.

After you've created a Microsoft account, browse to <http://azure.microsoft.com> and click the **Free Trial** link. Then follow the instructions to sign up for a free trial subscription to Microsoft

Azure. You'll need to sign in with your Microsoft account if you're not already signed in. Then you'll need to:

- Enter your cellphone number and have Microsoft send you a text message to verify your identity
- Enter the verification code sent to you
- Provide valid payment details—don't worry, your credit card won't be charged for any services you use during the trial period, and the account is automatically deactivated at the end of the trial period, unless you expressly decide to keep it active.

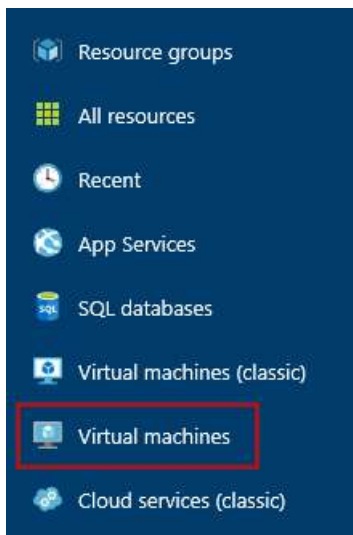
Exercise 1: Provisioning an Azure VM

In this exercise, having signed in to the Azure Portal by using your Azure subscription, you will provision an Azure VM to support all three labs for this course. The Azure VM will should be stopped at the end of the each lab so that your subscription is not charged (for free trial subscriptions, this will ensure you will have sufficient credits left to complete the labs over the duration of the course).

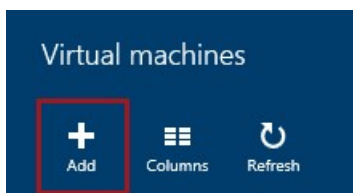
Provisioning an Azure VM

In this task, you will sign in to the Azure Portal, and then provision an Azure VM.

1. Sign in to the **Azure Portal** by using your subscription.
2. In the left pane, select **Virtual Machines**—do not select **Virtual Machines (Classic)**.



3. In the **Virtual Machines** blade, click **Add**.



4. In the **Marketplace** blade, in the search box, enter **SQL Server**, and then press **Enter**.
5. Select the **Free License: SQL Server 2016 SP1 Developer on Windows Server 2016** image.



6. In the image blade, review the text content, and then click **Create**.

A blue rectangular button with the word "Create" in white text.

7. Notice that the **Create Virtual Machine** blade opens, and that also the **Basics** blade (step 1) opens.
8. In the **Name** box, enter a name for the virtual machine (this will become the name of the machine).
9. In the **User Name** and **Password** boxes, enter appropriate values (this will become the machine administrator account).

Be sure to securely record these credentials, as you will be required to use them to sign in to the VM for each lab over the next three weeks.

10. In the **Resource Group** box, enter **Lab**.
11. In the **Location** box, enter a data center that is in close proximity to you.
12. Click **OK**.

A blue rectangular button with the word "OK" in white text.

13. In the **Create Virtual Machine** blade, select **Size** (step 2).
14. In the **Choose a Size** blade, select from the sizes available to your subscription, ensuring that the size contains at least 2 cores and 14GB RAM.
15. Click **Select**.

A blue rectangular button with the word "Select" in white text.

16. In the **Create Virtual Machine** blade, select **Settings** (step 3).
17. In the **Settings** blade, to accept the default settings, click **OK**.

A blue rectangular button with the word "OK" in white text.

18. In the **Create Virtual Machine** blade, select **SQL Server Settings** (step 4).
19. In the **SQL Server Settings** blade, click **Storage Configuration**.

SQL Authentication ⓘ

Disable Enable

Storage configuration ⓘ

General

20. In the **Storage Configuration** blade, in the **Storage Optimization** dropdown list, select **Data Warehousing**.

Storage optimization ⓘ

General

Transactional processing

Data warehousing

21. Click **OK**.

OK

22. In the **SQL Server Settings** blade, click **OK**.

OK

23. In the **Summary** blade, review the configuration, and then click **OK**.

OK

24. On the **Azure Portal** dashboard, notice the tile providing status of the deployment process.

The deployment usually takes 15-20 minutes to deploy, and this time depends largely on the VM size selected. The VM blade will open when the deployment completes.

You cannot proceed to the next task until the deployment completes.

25. Leave the **Azure Portal** dashboard open.

Connecting to the VM

In this task, once the VM has successfully deployed, you will connect to the VM.

26. In the VM blade, notice that the VM status is **Running**.

Resource group

Lab

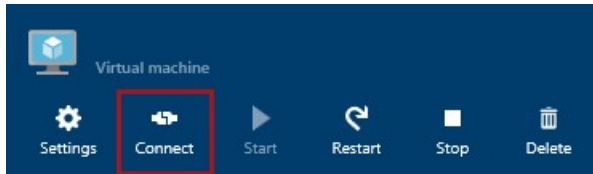
Status

Running

*You are charged when the VM status is **Running**, but you are not charged when the VM status is **Stopped (Deallocated)**.*

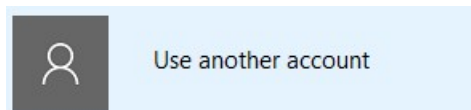
The labs will include steps to remind you to stop the VM.

27. To connect to the VM, click **Connect**.

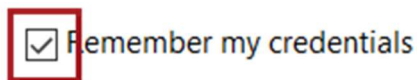


A Remote Desktop File (.rdp) file is downloaded to the desktop.

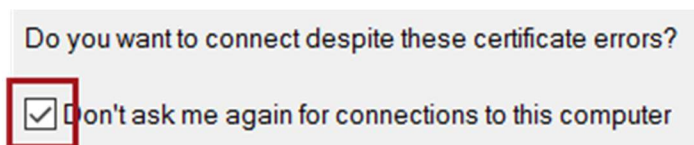
28. When prompted to open the Remote Desktop File, click **Open**.
29. If prompted to connect to the unknown publisher, click **Connect**.
30. In the **Windows Security** dialog window, click **Use Another Account**.



31. Enter the credentials you created for your VM.
32. Check the **Remember My Credentials** checkbox.



33. Click **OK**.
34. In the **Remote Desktop Connection** dialog window, check the **Don't Ask Me Again for Connections to This Computer** checkbox.



35. Click **Yes**.
36. If you have a second monitor, maximize the Remote Desktop window inside a single monitor.

Exercise 2: Setting Up the Azure VM

In this exercise, having connected to the VM, you will complete several setup tasks.

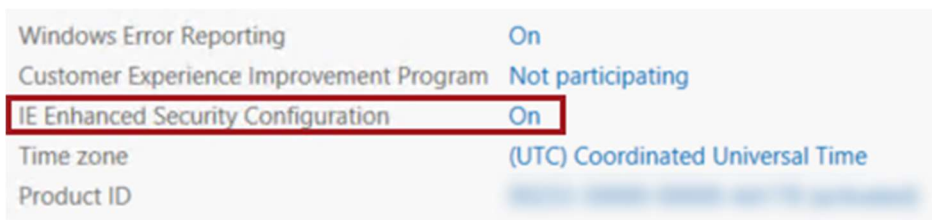
Configuring the Server

In this task, you will configure the server to support the lab experience.

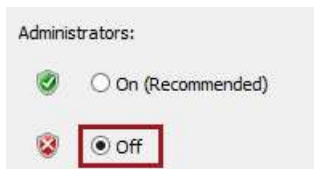
37. Notice that **Server Manager** has automatically opened.
38. In the left pane, select **Local Server**.



39. In the **Properties** pane, notice the **IE Enhanced Security Configuration** is set to **On**.



40. Click the **On** link.
41. In the dialog window, for **Administrators**, select the **Off** option.



42. Click **OK**.
43. Located at the top-right corner, select **Manage**, and then select **Server Manager Properties**.
44. In the dialog window, check the **Do Not Start Server Manager Automatically at Logon**.



45. Click **OK**.
46. Close the Server Manager window.

47. To open the **Start** page, press the **Windows** key.
48. Right-click the **Internet Explorer** tile, and then select **Pin to Taskbar**.
49. To exit the **Start** page, press **Escape**.
50. Notice the **Internet Explorer** shortcut on the taskbar.

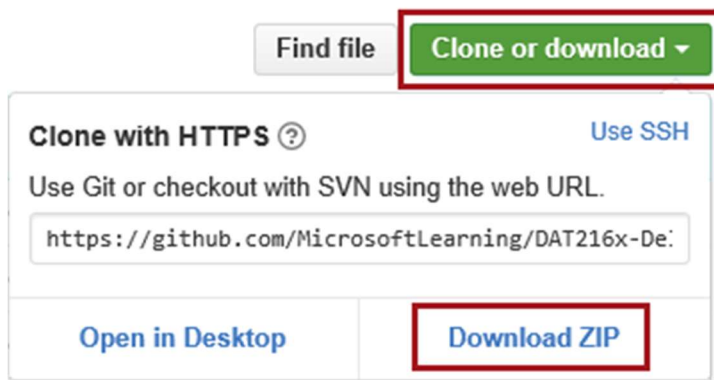


51. Click the **Internet Explorer** shortcut.
52. In the **Internet Explorer 11** dialog window, to accept the recommended settings, click **OK**.

Installing the Lab Resources

In this task, you will download and extract the lab resources to support all labs for this course.

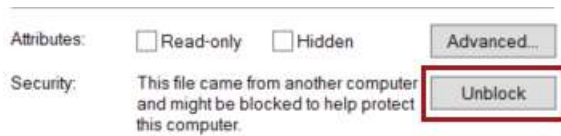
53. In the Internet Explorer **URL** box, enter <https://github.com/MicrosoftLearning/DAT216x-Delivering-EDW>.
54. To download the lab resources, click **Clone or Download**, and then click **Download ZIP**.



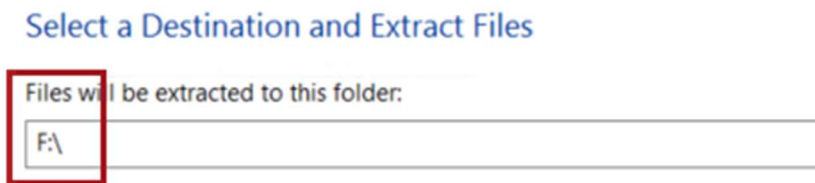
55. Save the downloaded file to **F:**.
56. When downloaded, open Windows Explorer, and navigate to **F:**.



57. Right-click the **DAT216x-Delivering-EDW-master** file, and then select **Properties**.
58. In the dialog window, click **Unblock**.



59. Click **OK**.
60. To extract the file content, right-click the **DAT216x-Delivering-EDW-master** file, and then select **Extract All**.
61. In the dialog window, replace the folder path to **F:**.



62. Click **Extract**.
63. Navigate to the **DAT216x-Delivering-EDW-master** folder.
64. Right-click the **DAT216x-Lab-Resources-v1.part01** file, and then select **Open**.
65. In the dialog window, read the license terms, and if you accept them, click **Accept**.
66. Notice that the destination path is set to **F:**.

As some lab scripts use absolute file paths, it is critical that you extract the lab resources directly to this path.

67. Click **Install**.
68. When the installation has completed, navigate to **F:**.
69. Optionally, delete the **DAT216x-Delivering-EDW-master** file, and also the **DAT216x-Delivering-EDW-master** folder.

Installing the Sample Databases

In this task, you will run a script to install the AdventureWorks sample databases.

70. In File Explorer, navigate to the **F:\Labs\Lab01** folder.
71. Double-click the **Setup** file.

The setup will restore the AdventureWorks2016 and AdventureWorksDW2016 databases.

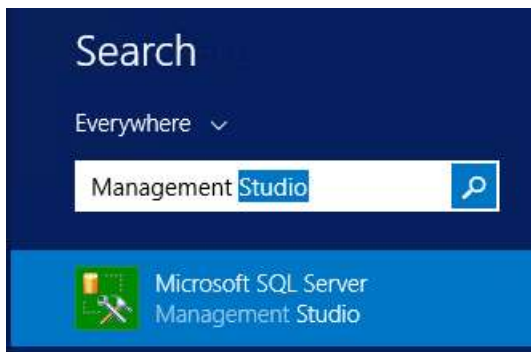
Note that this file can be run at any stage, should you require to reset the databases for the beginning of this lab.

72. When the script execution completes, press any key to close the console window.

Configuring SQL Server Management Studio

In this task, you will configure SQL Server Management Studio (SSMS). This tool is required to explore database, and also to execute scripts.

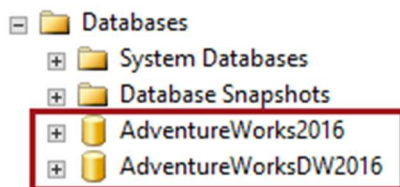
73. To add a shortcut to the taskbar, click the **Windows** key, and then commence typing **Management Studio**.



74. When the **Microsoft SQL Server Management Studio** search result appears, right-click it, and then select **Pin to Taskbar**.
75. Return to the desktop, and then click the **SQL Server Management Studio** shortcut.

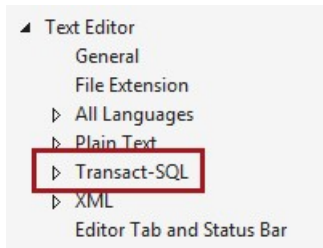


76. In the **Connect to Server** dialog window, to connect **Object Explorer** to the database engine, click **Connect**.
77. To verify that the sample databases are available, in **Object Explorer** (located at the left), expand the **Databases** folder.
78. Verify that the **AdventureWorks2016** and **AdventureWorksDW2016** databases are listed.

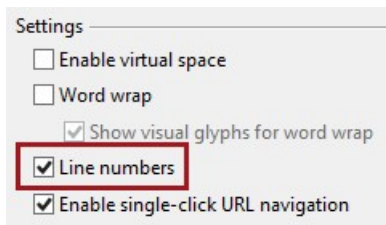


79. On the **Tools** menu, select **Options**.

80. In the Options window, in the left pane, expand **Text Editor**, and then select **Transact-SQL**.



81. Check the **Line Numbers** checkbox.



82. Click **OK**.
83. Leave SQL Server Management Studio open.

Installing SQL Server Tools

In this task, you will install SQL Server Data Tools (SSDT). This tool is required to develop Integration Services (SSIS) packages.

84. In Internet Explorer, navigate to <https://msdn.microsoft.com/en-us/mt186501>.

Tip: You can copy-and-paste the URL into the Remote Desktop window.

85. Scroll down to **Step 3**.

3 Set up an Administrative Install Point (optional)
For locations without internet access, create an **Administrative Install Point** for SQL Server Data Tools by following this procedure:

86. Right-click the link for **English (United States)**, and then select **Save Target As**.

The lab documentation is based on the English version of the SQL Server Tools. If you choose to install a different language, you will need to translate the lab instructions.

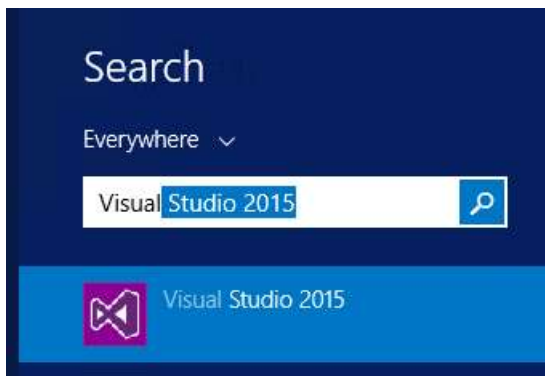
It is recommended that you setup the English version.

87. In the **Save As** dialog window, save the setup file to **F:\Labs\Setup**.
88. When downloaded, click **Open Folder**.

89. To install SSDT, double-click the **SSDTSetup** file.
90. In the installation window, click **Next**.
91. Read the license terms, and if you accept them, check the checkbox.
92. Click **Install**.

The installation usually takes 5-10 minutes to complete.

93. When the installation completes, click **Close**.
94. To add a shortcut to the taskbar, click the **Windows** key, and then commence typing **Visual Studio**.



95. When the **Visual Studio 2015** search result appears, right-click it, and then select **Pin to Taskbar**.
96. Return to the desktop, and then click the **Visual Studio 2015** shortcut.



97. In the getting started dialog window, in the **Development Settings** dropdown list, select **Business Intelligence**.



98. Click **Start Visual Studio**.
99. To close Visual Studio, on the **File** menu, select **Exit**.

*You will work with Visual Studio 2015 in **Lab 03**.*

Exercise 3: Exploring the AdventureWorks Databases

In this exercise, you will explore the two sample AdventureWorks databases, to help you understand subsets of their schemas.

Exploring the AdventureWorks2016 Database

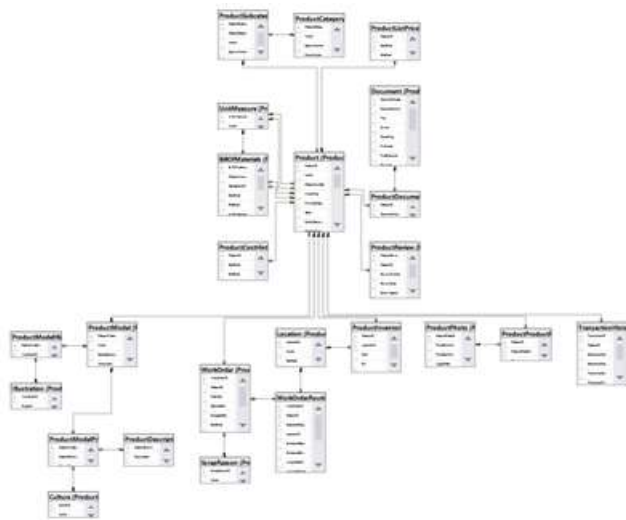
In this task, you will explore the **AdventureWorks2016** database.

100. In **Object Explorer**, expand the **AdventureWorks2016** database.
101. Expand also the **Tables** folder.
102. Notice the large list of tables, typical of a highly normalized database schema.
103. To filter the tables, right-click the **Tables** folder, and then select **Filter > Filter Settings**.
104. In the **Filter Settings** dialog window, configure the following.

Property	Operator	Value
Name	Contains	
Schema	Equals	Production
Owner	Equals	

105. Click **OK**.
106. Review the list of 25 tables which belong to the **Production** schema, noting that this subset of tables represents the Enterprise Resource Planning (ERP) system.
107. To create a database diagram, in **Object Explorer**, right-click the **Database Diagrams** folder, and then select **New Database Diagram**.
108. When the **Add Table** dialog window appears, click **Close**.
109. From **Object Explorer**, drag each of the 25 tables into the diagram (you can drop them randomly, and even on top of each other).
110. Right-click inside a blank area of the diagram, and then select **Arrange Tables**.
111. Right-click again inside a blank area of the diagram, and then select **Zoom > To Fit**.
112. Notice that 24 of the tables are related to one another.

It is not important to understand the details of tables and relationships.



Highly normalized database schemas are optimized for scalable and intensive write operations, and also ensure integrity by defining foreign keys (shown as relationships between tables). However, this design does not usually support efficient scalable read operations, as are required by data warehouse analytics queries.

113. To close the diagram, on the **File** menu, select **Close** (or close the diagram tab).

114. When prompted to save changes, click **No**.

115. In **Object Explorer**, collapse the **AdventureWorks2016** database.

Exploring the AdventureWorksDW2016 Database

In this task, you will explore the **AdventureWorksDW2016** database.

116. In **Object Explorer**, expand the **AdventureWorksDW2016** database.

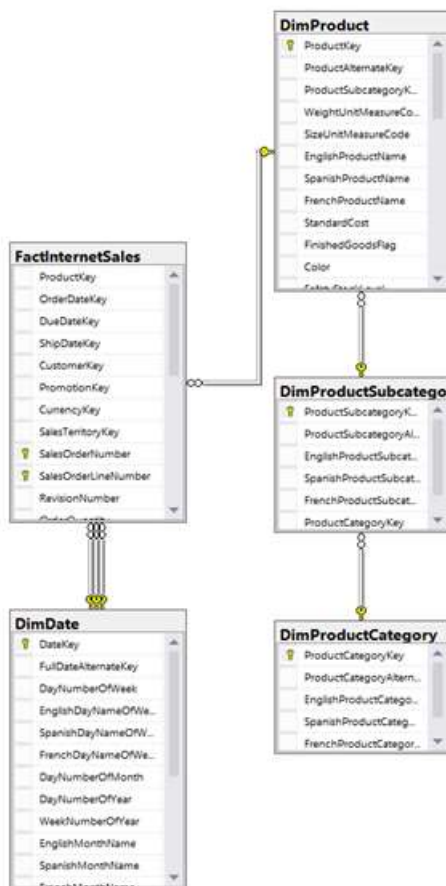
117. Expand also the **Tables** folder.

118. Note the naming conventions used for the tables: Either a table is prefixed with **Dim** (for dimension) or **Fact**.

119. Create a new database diagram, and add only the following tables:

- DimDate
- DimProduct
- DimProductCategory
- DimProductSubcategory
- FactInternetSales

120. Arrange the tables, and if necessary, zoom to fit.



In contrast to the AdventureWorks2016 database design, this schema supports efficient scalable read operations, as are required by data warehouse analytics queries.

121. Notice that the product dimension was developed as a snowflake dimension.

*In **Lab 02**, you will improve upon the design of specific tables to optimize storage, and enhance performance.*

*In **Lab 03**, you will develop and deploy an ETL solution to load these tables.*

122. Close the database diagram, and do not save changes.

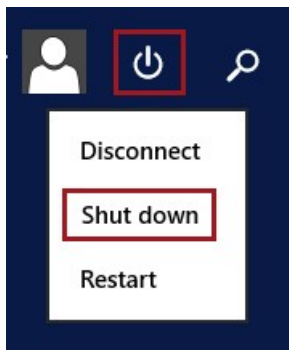
*You have now completed the lab. Be sure to complete the **Finishing Up** exercise to shut down, and stop the VM.*

Finishing Up

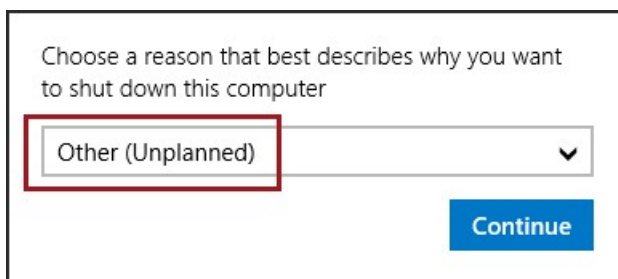
In this exercise, you will shut down and stop the VM.

123. Close all open applications.

124. Press the **Windows** key, and then in the **Start** page, locate at the top-right, select **Shut Down**.

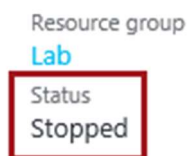


125. When prompted to choose a reason, to accept the default.



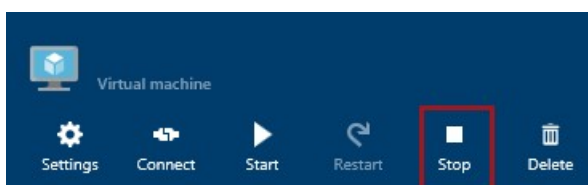
126. Click **Continue**.

127. In the **Azure Portal** Web browser page, wait until the status of the VM updates.

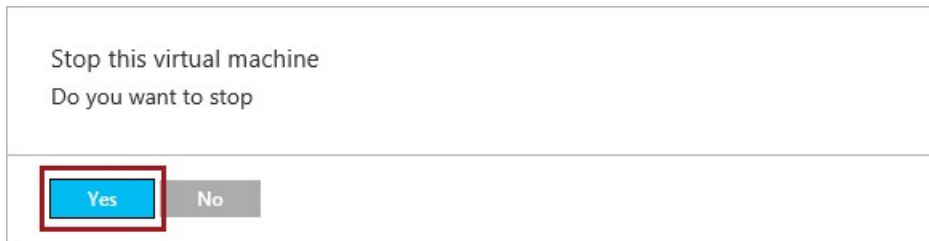


In this state, however, the VM is still billable.

128. To deallocate the VM, click **Stop**.



129. When prompted to stop the VM, click **Yes**.



The deallocation can take several minutes to complete.

130. Verify that the VM status updates to **Stopped (Deallocated)**.



In this state, the VM is now not billable.

Note that a deallocated VM will likely acquire a different IP address the next time it is started.

131. Sign out of the **Azure Portal**.