DAT224x

Developing a Multidimensional Data Model

Lab 02 | Creating the Multidimensional Project

Estimated time to complete this lab is 30 minutes

Overview

In this lab, you will commence the development of a Multidimensional Project based on the SQL Server **AdventureWorksDW2016** database. This will involve preparing data, creating a Multidimensional Project, and then adding a data source and data source view to the project.

Note: The four labs in this course are accumulative. You cannot complete this lab if you did not successfully complete **Lab 01**.

Getting Started

In this exercise, you will start the VM provisioned in **Lab 01**. You will then connect to the VM to complete the exercises in this lab.

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)**.
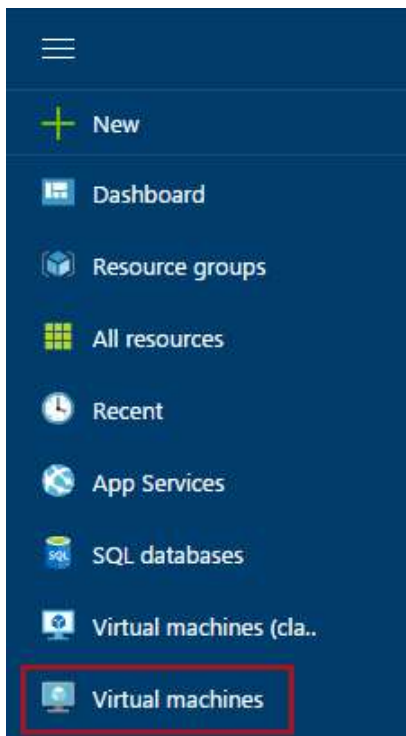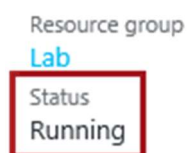

3.   In the **Virtual Machines** blade, select the VM you provisioned in **Lab 01**.
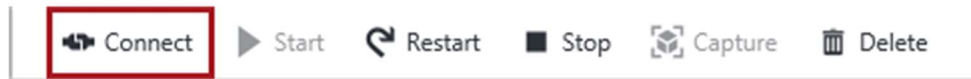
4.   In the VM blade, click **Start**.

5.   Wait for the VM status to update to **Running**.

    It usually takes 1-2 minutes for the VM to start.

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

   Take care not to use the RDP file downloaded in the previous lab. It is likely that a different IP address has be assigned.


7. When prompted to open the Remote Desktop File, click **Open**.

8. If prompted to connect to the unknown publisher, click **Connect**.

9. If prompted, in the **Windows Security** dialog window, click **Use Another Account**.
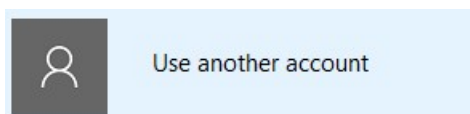
10. Enter the credentials you created for your VM.

11. Check the **Remember My Credentials** checkbox.
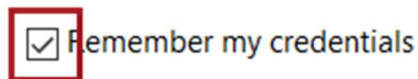
12. Click **OK**.

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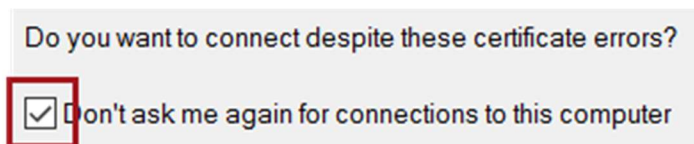
14. Click **Yes**.

Exercise 1: Creating the Multidimensional Project

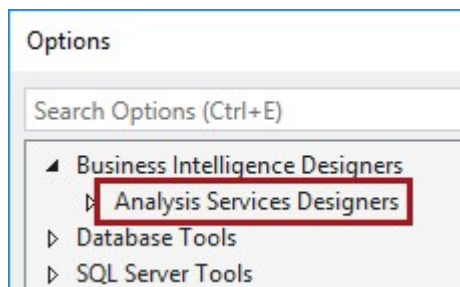In this exercise, you will create the Multidimensional Project.

Configuring the Analysis Services Options

In this task, you will launch SSDT and configure options that will be applied to new Multidimensional Projects.

1. To launch SSDT, on the task bar, click the **Visual Studio 2015** shortcut.

2. To configure the Analysis Services options, on the **Tools** menu, select **Options**.

3. In the **Options** window, expand the **Business Intelligence Designers** group (you will need to scroll down the list to locate this group), and then select the **Analysis Services Designers** page.

4. In the **Default Target Server** dropdown list, enter **localhost**.

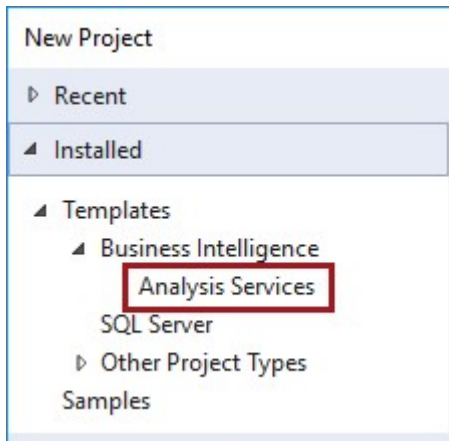   *This is the Analysis Services instance that you installed in **Lab 01**.*
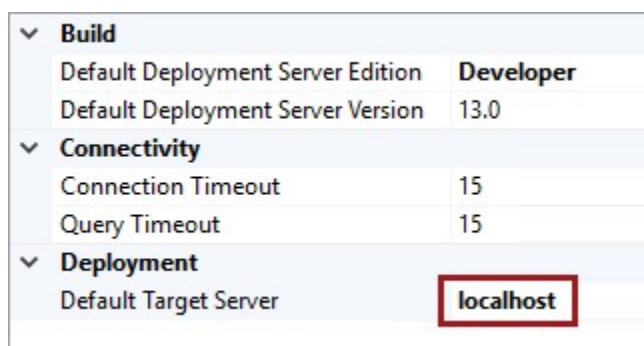
5. Click **OK**.

Creating the Multidimensional Project

In this task, you will create the Multidimensional Project.

1. To create a solution, on the **File** menu, select **New | Project**.

2. In the **New Project** window, in the left pane, from inside the **Business Intelligence** group, select the **Analysis Services** template.
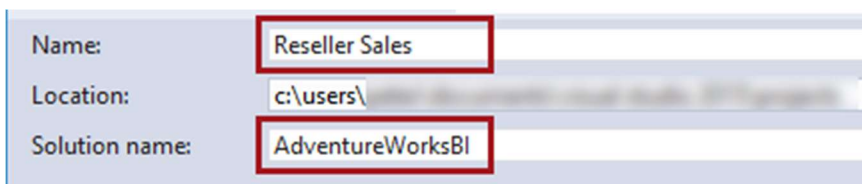

3. Select the **Analysis Services Multidimensional and Data Mining Project** template.
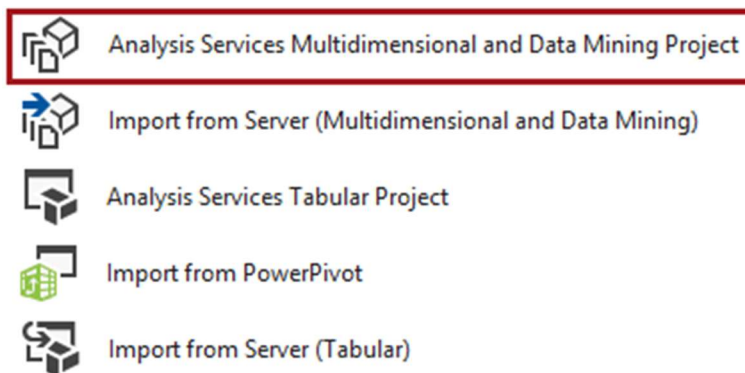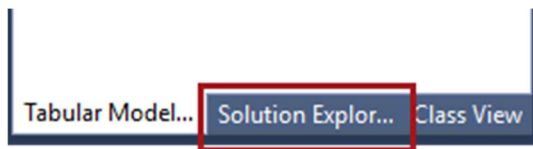

The labs will refer to this simply as a Multidimensional Project. You will not work with the data mining features in this course.

4. In the **Name** box, replace the text with **Reseller Sales**.

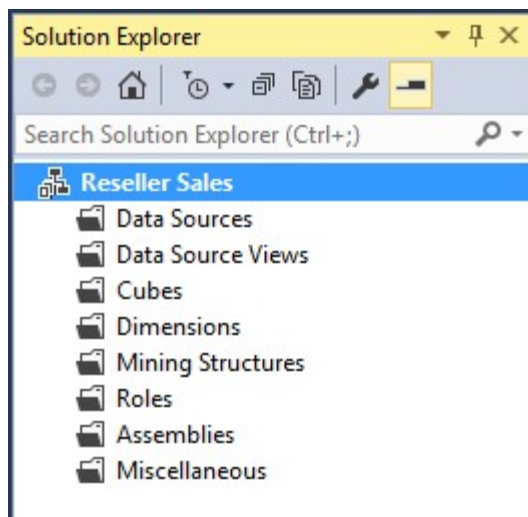5. In the **Solution Name** box, replace the text with **AdventureWorksBI**.

6. Click **OK**.

---

7.   In the right pane, select **Solution Explorer**.

8.   Notice that the **Reseller Sales** project consists of a single project contains multiple folders.

A Multidimensional Project consists of many item types. Relevant to multidimensional model development, these include data sources, data source views, cubes, dimensions, roles and assemblies. When deployed for the first time, the project creates a database on the target Analysis Services instance. You will first deploy the project in the next lab.

9.   To save the project, on the **File** menu, select **Save All**.

It is a good practice to regularly save the solution to protect your development effort in case of an unexpected application crash.

*Tip: The **Save All** function is also available on the standard toolbar.*

Exercise 2: Creating the Data Source

In this exercise, you will create a data source.
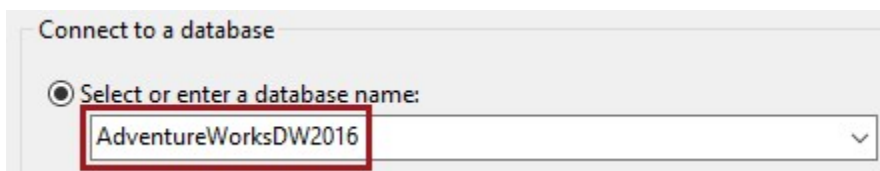
Creating the Data Source

In this task, you will create the **AdventureWorksDW2016** data source.

1. In **Solution Explorer**, right-click the **Data Sources** folder, and then select **New Data Source**.

2. In the **Data Source Wizard**, click **Next**.

3. At the **Select How to Define the Connection** step, in the **Data Connections** list, notice the selected option to create a data source based on an existing or new connection.

4. To create a connection, click **New**.

5. In the **Connection Manager** window, in the **Server Name** box, enter **localhost** (do not open the dropdown list).

6. In the **Select or Enter a Database Name** dropdown list, select **AdventureWorksDW2016**.

7. Click **OK**.

8. Click **Next**.

9. At the **Impersonation Information** step, select the **Use the Service Account** option.

    A preferred practice is to use a dedicated domain account. For simplicity, you will use the service account in this lab.

10. Click **Next**.

11. At the **Completing the Wizard** step, in the **Data Source Name** box, replace the text with **AdventureWorksDW2016**.

   When encountering object names in camel case or Pascal case, the Analysis Services wizards will insert spaces between the words. This is usually helpful to produce friendly names—but this is not useful for the data source name.

12. Click **Finish**.

13. In **Solution Explorer**, notice the addition of the **AdventureWorksDW2016** data source.
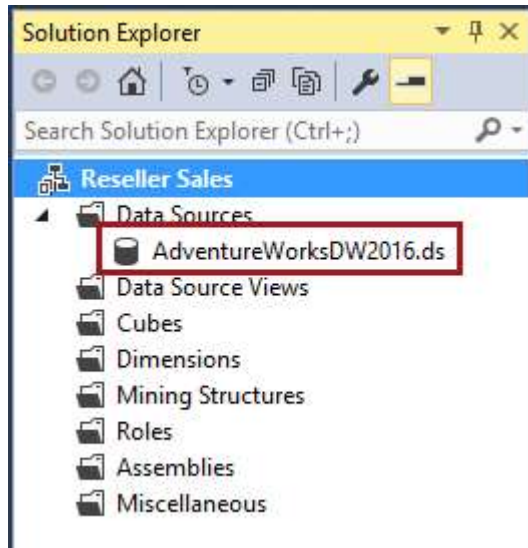
14. To save the solution, on the **File** menu, select **Save All**.

Exercise 3: Developing the Data Source View

In this exercise, you will develop and enhance a data source view. The data source view is a virtual schema upon which the dimensions and cube will be based.

Creating the Data Source View

In this task, you will create the **Reseller Sales** data source view.

1. In **Solution Explorer**, right-click the **Data Source Views** folder, and then select **New Data Source View**.

2. In the **Data Source View Wizard**, click **Next**.

3. At the **Select a Data Source** step, notice that the data source you created in previous exercise is selected, and then click **Next**.

4. At the **Select Tables and Views** step, to fully display the object names, increase the height and width of the window by dragging the bottom right corner of the window.

5. In the **Available Objects** list, select the **FactSalesQuota** table.

6. To add the selected table to the **Included Objects** list, click the right-pointing arrow button.
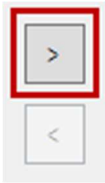
7. To add related tables to the **Included Objects** lists, click **Add Related Tables**.

    Related tables are identified by foreign key relationships in the data source.

8. Notice the addition of the **DimEmployee** and **DimDate** tables to the **Included Objects** list.

9. In the **Available Objects** list, select the **DimGeography** table.

10. While pressing the **Control** key, select the following five additional tables.

Table
DimProduct
DimProductCategory
DimProductSubcategory
DimReseller
FactResellerSales

11. To add the tables to the data source view, click the right-pointing arrow button.

12. Verify that there are nine included objects (order does not matter).

Included objects:

Name	Type
FactSalesQuota (dbo)	Table
DimEmployee (dbo)	Table
DimDate (dbo)	Table
DimGeography (dbo)	Table
DimProduct (dbo)	Table
DimProductCategory (dbo)	Table
DimProductSubcategory (...	Table
DimReseller (dbo)	Table
FactResellerSales (dbo)	Table

13. Click **Next**.

14. At the **Completing the Wizard** step, in the **Name** box, replace the text with **Reseller Sales**.

Data Source View Wizard
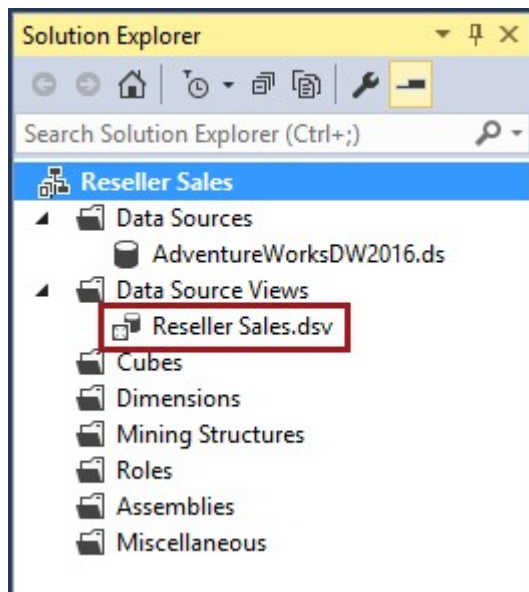
Completing the Wizard
Provide a name, and then click Finish to create the new data source view.

Name:
Reseller Sales

15. Click **Finish**.

16. In **Solution Explorer**, notice the addition of the **Reseller Sales** data source view.

17. Notice also that the data source view designer opens automatically.

18. To save the solution, on the **File** menu, select **Save All**.

Exploring the Data Source View Diagram

In this task, you will explore the data source view diagram.

1. In the data source view designer, at the bottom right-corner of the diagram pane, click the four-headed arrow to open the diagram map, and then drag the frame to navigate about the entire diagram.
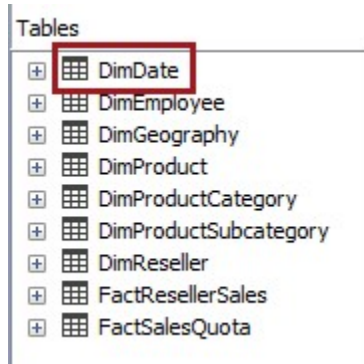

2. Explore the diagram to understand the table and relationship design. In particular note the following:

    - The **FactResellerSales** and **FactSalesQuota** tables are comprised of dimension keys and measures

    - Both fact tables have relationships to the **DimDate** and **DimEmployee** tables

    - There are three relationships between the **FactResellerSales** table and the **DimDate** table – each for a variation of date: **OrderDate**, **DueDate** and **ShipDate**

    - The self-referencing relationship on the **DimEmployee** table to generate a parent-child hierarchy

    - The three snowflake product dimension tables: **DimProduct**, **DimProductSubcategory** and **DimProductCategory**, and that the **DimProduct** table supports Type 2 Slowly Changing Dimensions (it includes **StartDate** and **EndDate** columns)

Enhancing the Data Source View

In this task, you will enhance the data source view by providing friendly names for each of the data source view tables, adding a named calculation, replacing a table with a named query, and removing a relationship that is not required.

1. In the data source view designer, in the **Tables** pane (located at the left), select the **DimDate** table.

    Tables
    - ⊞ DimDate
    - ⊞ DimEmployee
    - ⊞ DimGeography
    - ⊞ DimProduct
    - ⊞ DimProductCategory
    - ⊞ DimProductSubcategory
    - ⊞ DimReseller
    - ⊞ FactResellerSales
    - ⊞ FactSalesQuota

    *In the diagram pane, notice that the **DimDate** table is also selected. Also, in the **Properties** pane (located at the bottom-right), notice that the table is selected.*

2. In the **Properties** pane, modify the **FriendlyName** property to **Date**.

3. Repeat this process for each of the seven <u>dimension</u> tables to remove the **Dim** table prefix.

    The purpose of this step is to help create a user-friendly data model. It is convenient to configure friendly names in the data source view so that they are consistently propagated to the objects (cubes, dimension and mining structures) based on it.

4. Modify the **FriendlyName** properties of the two fact tables as follows.

Table	Friendly Name
FactResellerSales	Sales
FactResellerSalesQuota	Quota

5. Verify that the table names no longer include the table prefix (Dim or Fact).
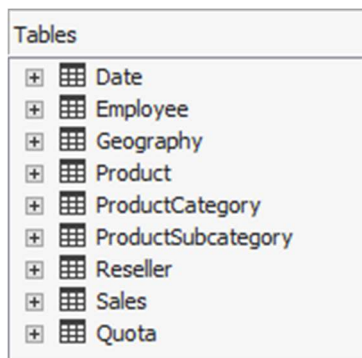
6. To add a named calculation, in the **Tables** pane, right-click the **Employee** table, and then select **New Named Calculation**.

    You can also access the context menu by right-clicking the table in the diagram pane.

7. In the **Create Named Calculation** window, in the **Column Name** box, enter **Employee_Name**.

8. In the **Expression** box, enter the following expression.

    You can copy the expression from the F:\Labs\Lab02\Assets\Snippets.txt file.

    T-SQL
    ```
    [FirstName] + N' ' + [LastName]
    ```

    A named calculation extends the table definition with a new column based on an expression.

9. Click **OK**.

10. In the diagram pane, inside the **Employee** table, notice the addition of the **Employee_Name** column, decorated with the named calculation icon.
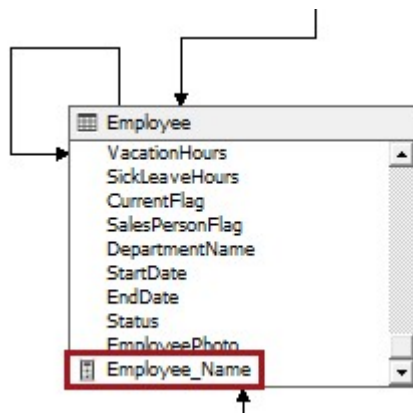
11. To review the new column data, right-click the **Employee** table, and then select **Explore Data**.

12. In the table explorer, use the horizontal scroll bar to scroll to the last column, and then notice the **Employee_Name** column and its values.

13. To close the table explorer, on the **File** menu, select **Close**.

   Tip: To close any project item, you can also click the X inside the item tab.

14. To replace a table with a named query, right-click the **Employee** table, and then select **Replace Table | With New Named Query**.

15. In the **Create Named Query** window, in the **Name** box, replace the text with **Salesperson**.

16. To hide the diagram and grid panes, click the following toolbar buttons.

17. Replace the query in the SQL pane with the following.

   You can copy the query from the F:\Labs\Lab02\Assets\Snippets.txt file.

   T-SQL
```
SELECT
  [e].[EmployeeKey]
  ,[e].[ParentEmployeeKey]
  ,[e].[FirstName] + N' ' + [e].[LastName] AS [Salesperson_Name]
  ,[e].[LoginID]
  ,[e].[EmailAddress]
  ,[e].[Phone]
  ,[st].[SalesTerritoryRegion] AS [Region]
  ,[st].[SalesTerritoryCountry] AS [Country]
  ,[st].[SalesTerritoryGroup] AS [Group]
FROM
  [dbo].[DimEmployee] AS [e]
  INNER JOIN [dbo].[DimSalesTerritory] AS [st]
        ON [st].[SalesTerritoryKey] = [e].[SalesTerritoryKey]
WHERE
  [e].[SalesPersonFlag] = 1;
```

   *A named query is a query that behaves as a table. It has the benefits of defining calculated columns (like the **Salesperson_Name** column), filtering rows with a WHERE clause and joining tables together, as is the case in this query.*

18. To run the query, on the toolbar, click **Run**.

19. Review the query result in the results pane.

20. Click **OK**.

21. In the **Tables** pane, notice the named query icon decorating the **Salesperson** table, and that it is different from the other icons used to decorate direct tables.
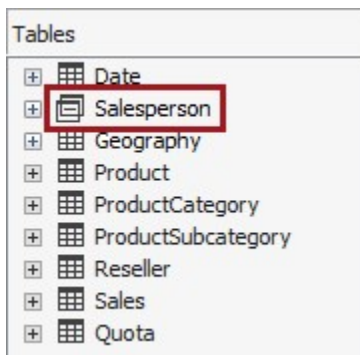
22. To remove an existing relationship, in the **Tables** pane, expand the **Date** table, and then expand the **Relationships** folder (located after the last column).

23. Right-click the **Sales (Using DueDateKey)** relationship, and then select **Delete Relationship**.

24. In the **Delete Objects** window, to confirm the deletion, click **OK**.

   The relationship was deleted as the cube will not be used to analyze sales by due date.

25. In the **Tables** pane, collapse the **Date** table.

26. To save the solution, on the **File** menu, select **Save All**.

27. To close the data view designer, on the **File** menu, select **Close**.

You have now completed the lab. In the next lab, you will add dimensions to the project.

*If you are not immediately continuing with the next lab, you should complete the **Finishing Up** exercise to shut down and stop the VM.*
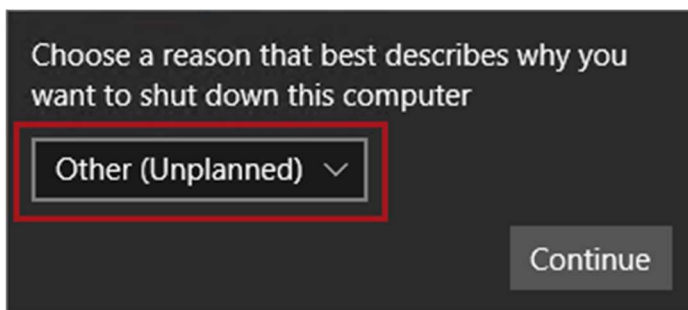
Finishing Up

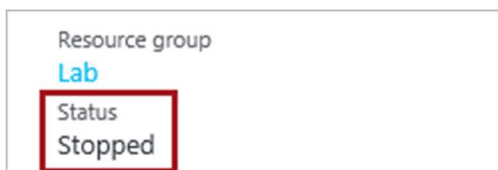In this exercise, you will shut down and stop the VM.

1. Close all open applications.

2. Press the **Windows** key, and then in the **Start** page, located at the bottom-left, click the **Power** button, and then select **Shut Down**.
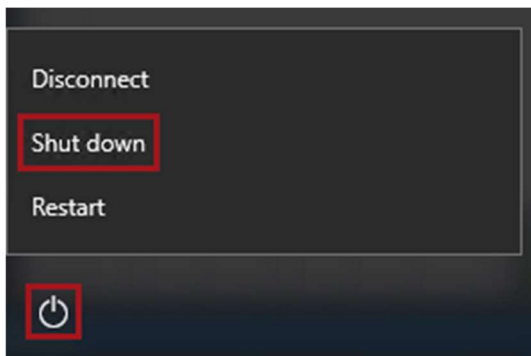
3. When prompted to choose a reason, to accept the default.

4. Click **Continue**.

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

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

6. Optionally, to deallocate the VM, click **Stop**.

   Deallocation will take some minutes to complete, and also extends the time required to restart the VM. Consider deallocating the VM if you want to reduce costs, or if you choose to complete the next lab after an extended period.

   The deallocation can take several minutes to complete.

7. Verify that the VM status updates to **Stopped (Deallocated)**.
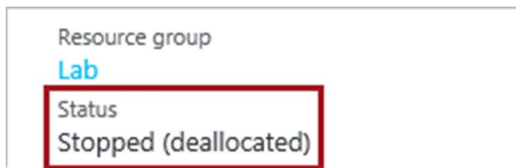
   In this state, the VM is now not billable—except for a relatively smaller storage cost.

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

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