

Lab 0: Prerequisites and Setup

Saturday, February 2, 2019 12:58 PM

Background

The labs in this Azure Data Factory Deep Dive workshop are intended to give you experience working with a number of key components and data sources that are commonly used in production solutions. Participation in each lab activity is optional, but in order to complete all the labs in this workshop, you will need to configure the following resources.

Prerequisites

Azure Subscription

If you do not have an Azure subscription, create a **free account** on <https://azure.microsoft.com/free/> before you begin.

If you have an existing Azure subscription, ensure that you have **Service Administrator** or **Owner** level of access on the subscription. To view the permissions that you have in the subscription, in the Azure portal, select your username in the upper-right corner, click the ellipsis next to Switch directory, and then select *My Permissions*. If you have access to multiple subscriptions, select the appropriate subscription.

If you do not have Service Administrator or Owner level of access, we recommend that you create a free account. This is because some of the labs may not work if you can't create resources, assign permissions, or make modifications to the Azure Active Directory.

Whitelisted Preview of Azure Data Factory Mapping Data Flows

The Mapping Data Flows feature in Azure Data Factory is still in preview. To participate in these labs, you must sign up for the preview on <http://aka.ms/dataflowpreview>. Make sure you enter "SQLBits" as the Business Name, as this will speed up the approval process.

Local Machine

Currently, the Azure Data Factory UI is only supported in the **Microsoft Edge** and **Google Chrome** browsers. Ensure that you have one of these browsers installed.

We also recommend that you have the following applications installed locally:

- **SQL Server Management Studio (SSMS)**: <https://docs.microsoft.com/en-us/sql/ssms/download-sql-server-management-studio-ssms>
- **SQL Server Data Tools (SSDT)**: <https://docs.microsoft.com/en-us/sql/ssdt/download-sql-server-data-tools-ssdt>
Please note: Make sure you use the **standalone installer** that installs tools for SQL Server Integration Services
- **Azure Storage Explorer**: <https://azure.microsoft.com/en-us/features/storage-explorer/>

Setup: Configure Own Environment

Log into the Azure Portal (<https://portal.azure.com/>). You will create the following resources:

1. Resource Group
2. Azure Data Lake Storage Gen1
3. Azure Storage Account
4. Azure Key Vault
5. Azure SQL Server
6. Azure SQL Database
7. Azure SQL Data Warehouse

8. Azure Analysis Services
9. Azure Databricks
10. Virtual Machine (Free SQL Server License: SQL Server 2017 Developer on Windows Server 2016) with AdventureWorksLT Demo Database

NAME	PUBLISHER
 Resource group	Microsoft
 Data Lake Storage Gen1	Microsoft
 Storage account	Microsoft
 Key Vault	Microsoft
 SQL server (logical server)	Microsoft
 SQL Database	Microsoft
 SQL Data Warehouse	Microsoft
 Analysis Services	Microsoft
 Azure Databricks	Microsoft
 Free SQL Server License: SQL Server 2017 Developer on Windows Server 2016	Microsoft

1. Create Resource Group

Resource Groups enable you to manage all your resources in an application together. All the remaining resources you create will be created in this Resource Group.

Step	Setting	Value	Notes
1	Resource Group	sqlbitsadf<999>	The name of the Resource Group. Use the prefix sqlbitsadf followed by a unique suffix such as your initials and/or a number. Make note of this name as you will refer back to it frequently in later activities.
2	Region	UK South	Choose UK South . It is important to note that not all Azure resources can be created in each region, but we selected UK South because <i>most</i> of the resources exist within that region.

Unless specified above, use the default settings. It is not necessary to add any tags.

Home > New > Marketplace > My Saved List > Resource group > Create a resource group

Create a resource group

Basics Tags Review + Create

Resource group - A container that holds related resources for an Azure solution. The resource group can include all the resources for the solution, or only those resources that you want to manage as a group. You decide how you want to allocate resources to resource groups based on what makes the most sense for your organization. [Learn more](#)

PROJECT DETAILS

* Subscription: Visual Studio Premium with MSDN

* Resource group: sqlbitsadfcw01 1

RESOURCE DETAILS

* Region: UK South 2

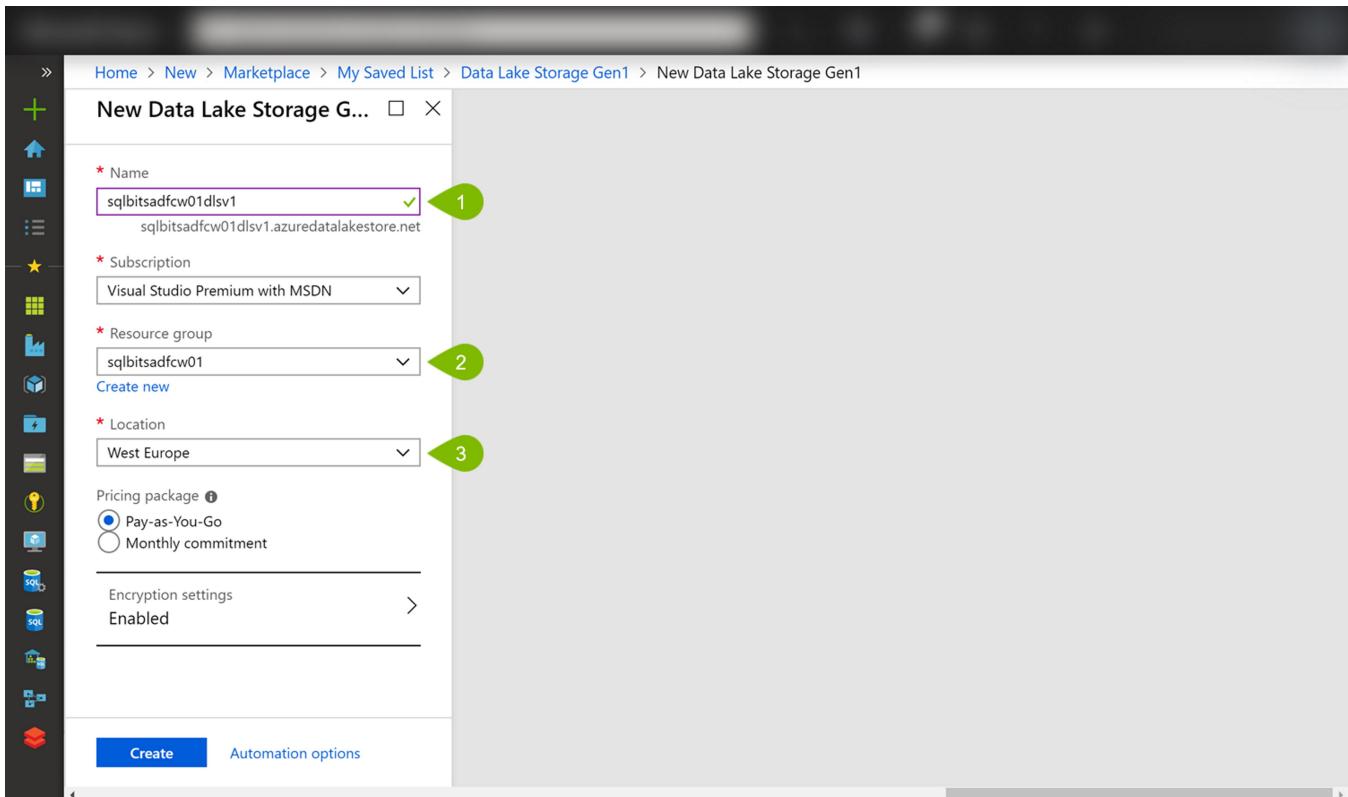
Review + Create **Next : Tags**

2. Create Azure Data Lake Storage Gen1

Please note: Azure Data Lake Storage Gen2 is now generally available. For these labs, you will still use Azure Data Lake Storage Gen1. This is because not all Azure Data Factory activities fully support Gen2 yet.

Step	Setting	Value	Notes
1	Name	sqlbitsadf<999>dlsv1	Use the name of the Resource Group and add dlsv1 as a suffix.
2	Resource Group	sqlbitsadf<999>	Choose the Resource Group you created in the first step.
3	Location	West Europe	This resource cannot be created in UK South, so choose West Europe .

Unless specified above, use the default settings.



Additional information:

<https://docs.microsoft.com/en-us/azure/data-lake-store/data-lake-store-get-started-portal>

3. Create Azure Storage Account

Storage accounts allow you to store objects like blobs, files, and tables.

Step	Setting	Value	Notes
1	Resource Group	sqlbitsadf<999>	Choose the Resource Group you created in the first step.
2	Storage Account Name	sqlbitsadf<999>wasb	Use the name of the Resource Group and add wasb as a suffix.
3	Location	UK South	Choose UK South .

Unless specified above, use the default settings.

Additional information:

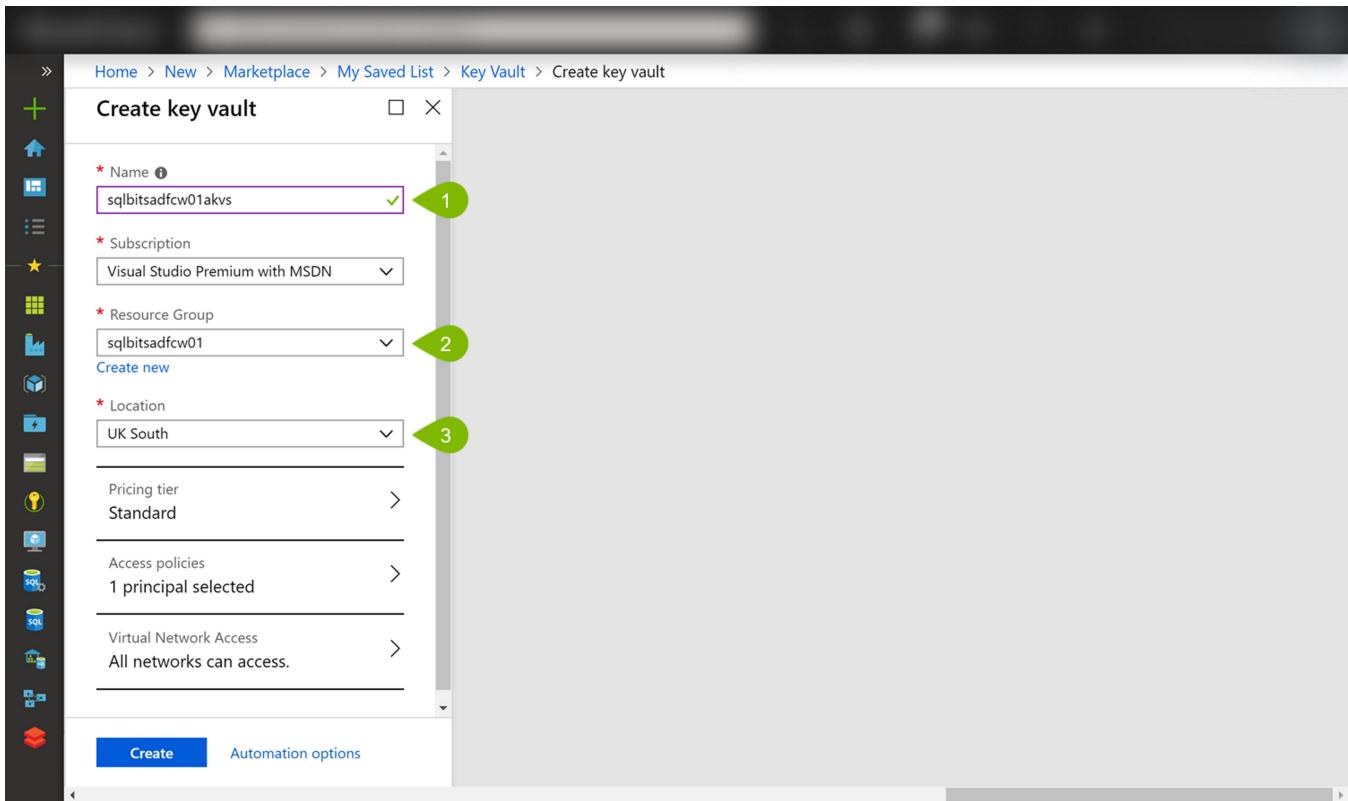
<https://docs.microsoft.com/en-us/azure/storage/common/storage-quickstart-create-account?toc=%2Fazure%2Fstorage%2Fblobs%2Ftoc.json&tabs=azure-portal>

4. Create Azure Key Vault

With Azure Key Vault, you can safely store objects like passwords and encryption keys.

Step	Setting	Value	Notes
1	Name	sqlbitsadf<999>akvs	Use the name of the Resource Group and add akvs as a suffix.
2	Resource Group	sqlbitsadf<999>	Choose the Resource Group you created in the first step.
3	Location	UK South	Choose UK South .

Unless specified above, use the default settings.



Additional information:

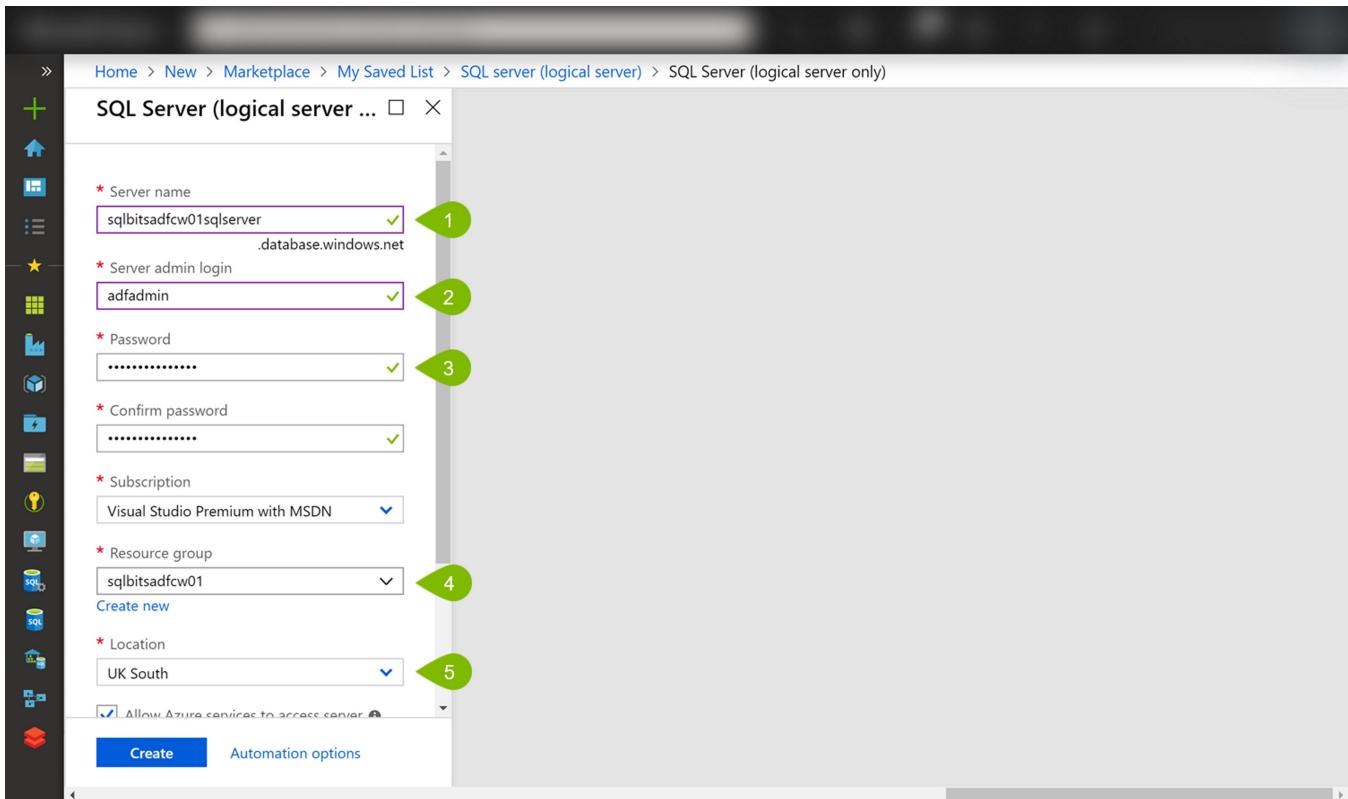
<https://docs.microsoft.com/en-us/azure/key-vault/quick-create-portal>

5. Create Azure SQL Server

The Azure SQL Server will host your SQL Database and SQL Data Warehouse.

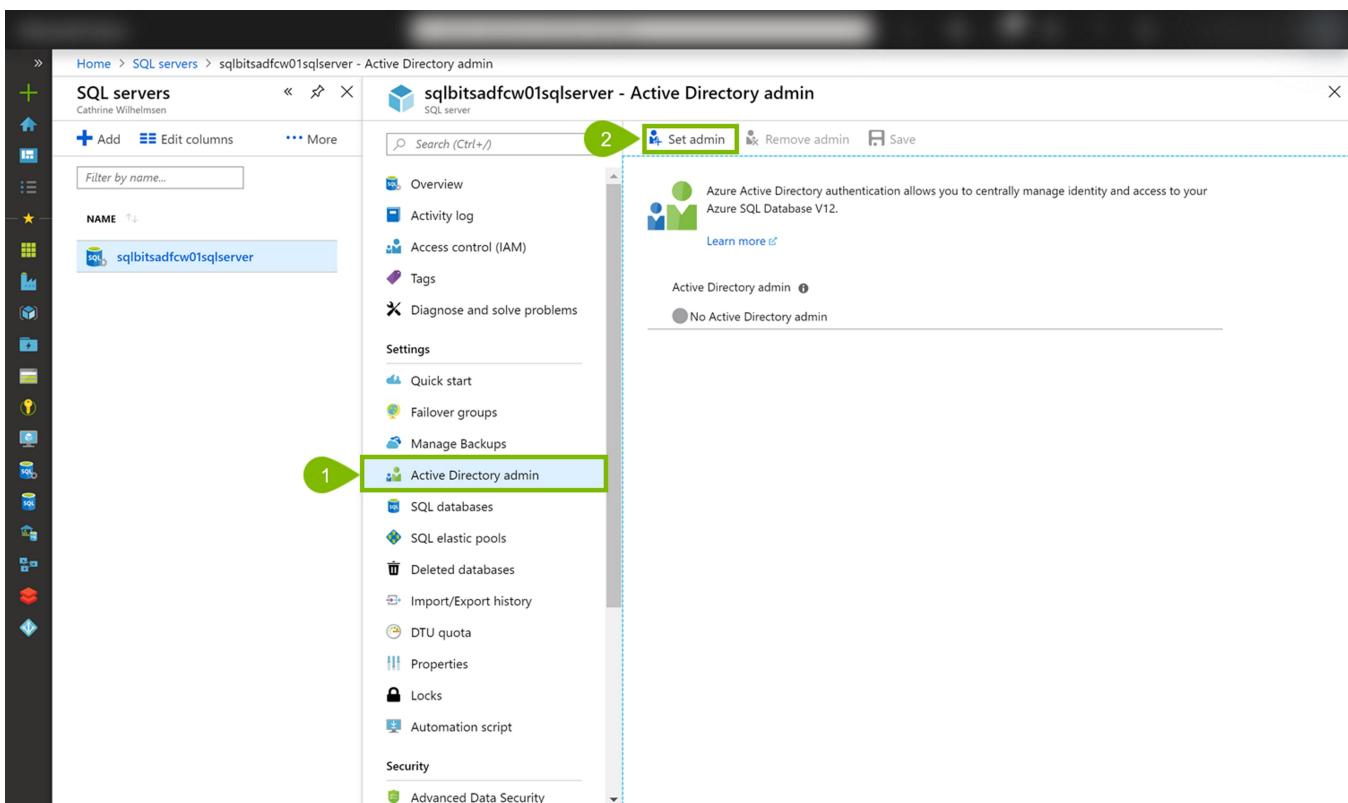
Step	Setting	Value	Notes
1	Server Name	sqlbitsadf<999>sqlserver	Use the name of the Resource Group and add sqlserver as a suffix.
2	SQL Admin Login	adfadmin	Pick a username, for example adfadmin
3	Password	VeryStr0ngPassw0rd!	Pick a password, for example VeryStr0ngPassw0rd!
4	Resource Group	sqlbitsadf<999>	Choose the Resource Group you created in the first step.
5	Location	UK South	Choose UK South .

Unless specified above, use the default settings.

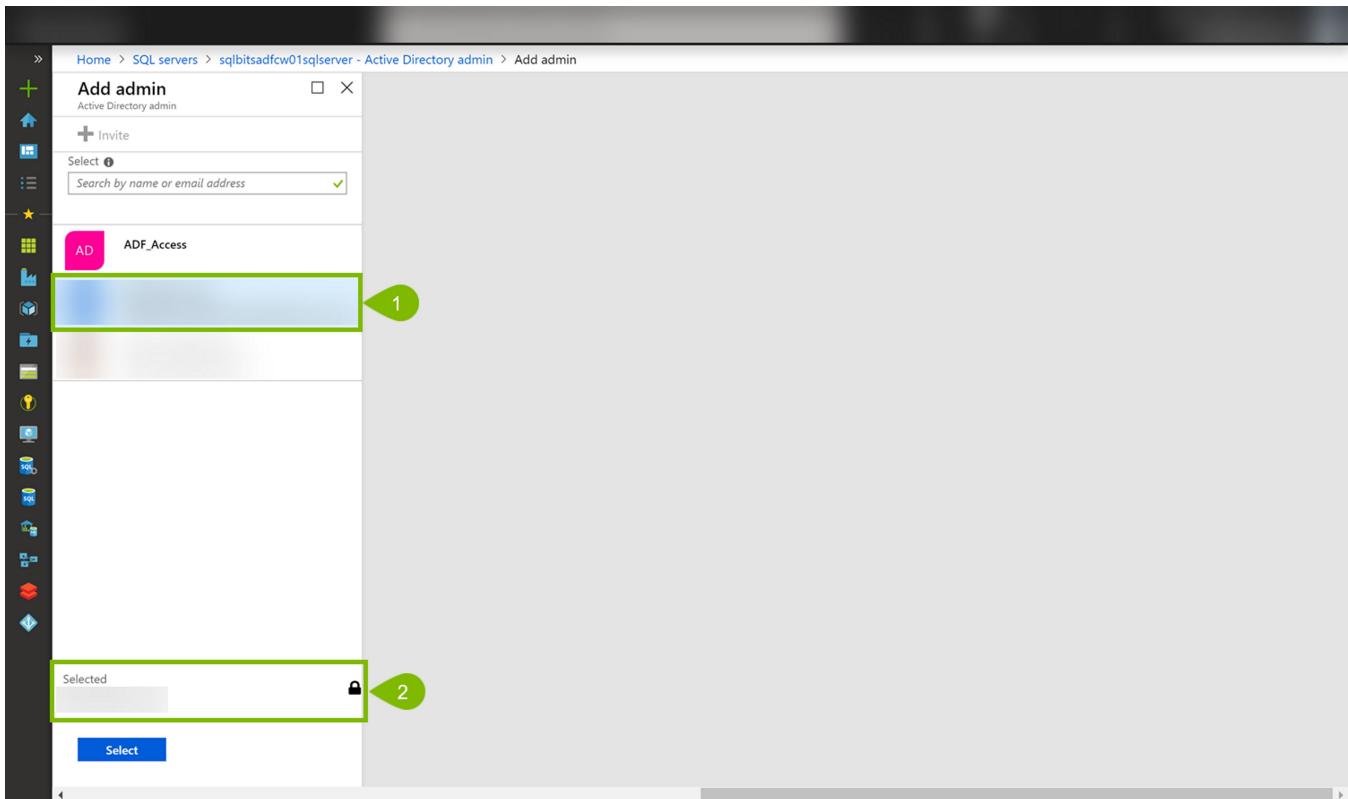


Add yourself as an Active Directory Admin of the Azure SQL Server

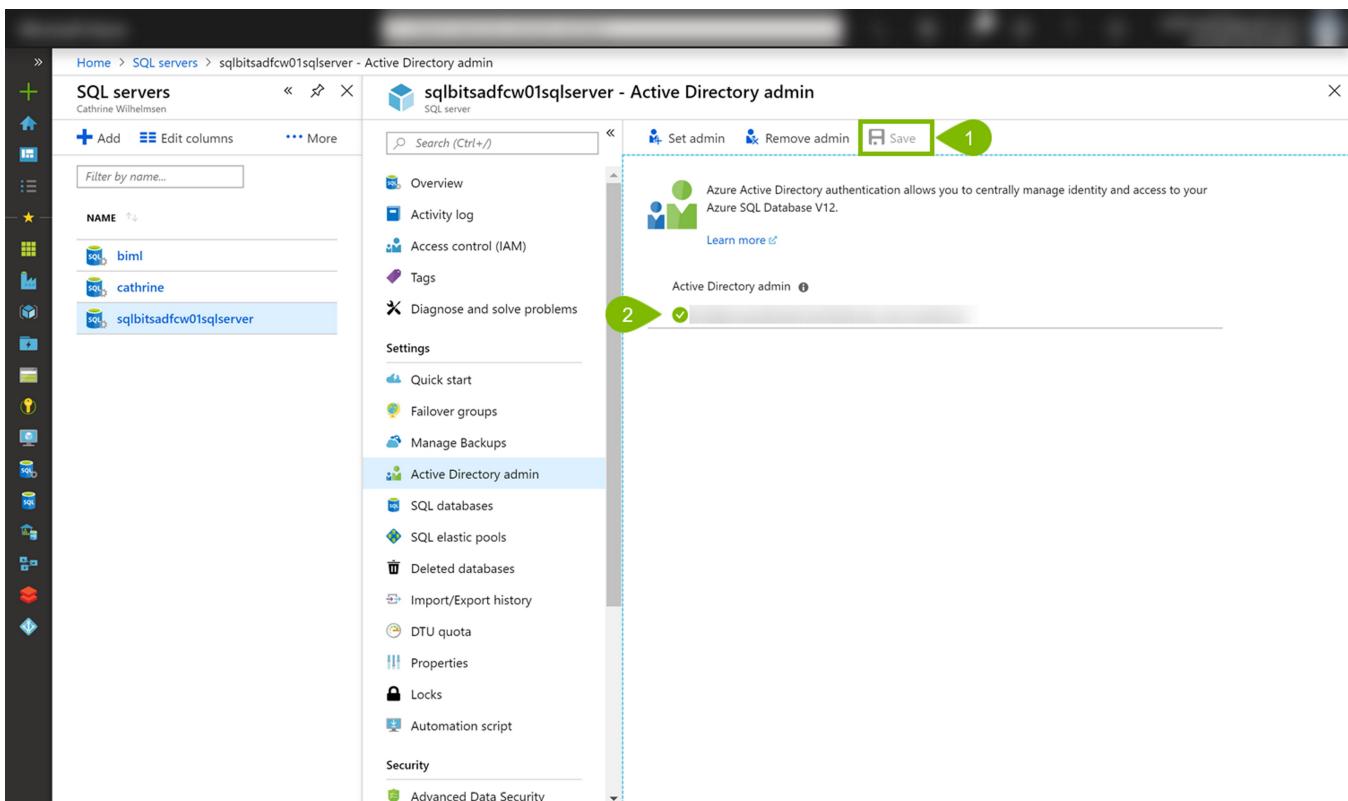
Click on **Active Directory Admin**, then **Set Admin**.



Select **your user**, then click **Select**.



Ensure that you **Save** the new Active Directory Admin and see the **green checkmark**.



6. Create Azure SQL Database

Create the Azure SQL Database in the Azure SQL Server created in the previous step, then restore

Create Azure SQL Database

Step	Setting	Value	Notes
1	Database Name	Staging	
2	Resource Group	sqlbitsadf<999>	Choose the Resource Group you created in the first step.
3	Select Source	Blank database	Keep the default Blank database
4	Server	sqlbitsadf<999>sqlserver	Choose the SQL Server you created in the previous step.
5	Pricing tier	Basic, 2 GB	Change the default Standard to the cheaper Basic.

Unless specified above, use the default settings.

The screenshot shows the 'SQL Database' creation wizard in the Azure portal. The form has five numbered callouts pointing to specific fields:

1. Database name: Staging
2. Resource group: sqlbitsadf01
3. Select source: Blank database
4. Server: sqlbitsadf01sqlserver (UK South)
5. Pricing tier: Basic, 2 GB

Additional information:

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-single-database-get-started>

Dacpac

Download the Staging.dacpac file from <https://github.com/jasonhorner/adfdeepdive/tree/master/Labs/01%20-%20ADF%20Overview>.

7. Create Azure SQL Data Warehouse

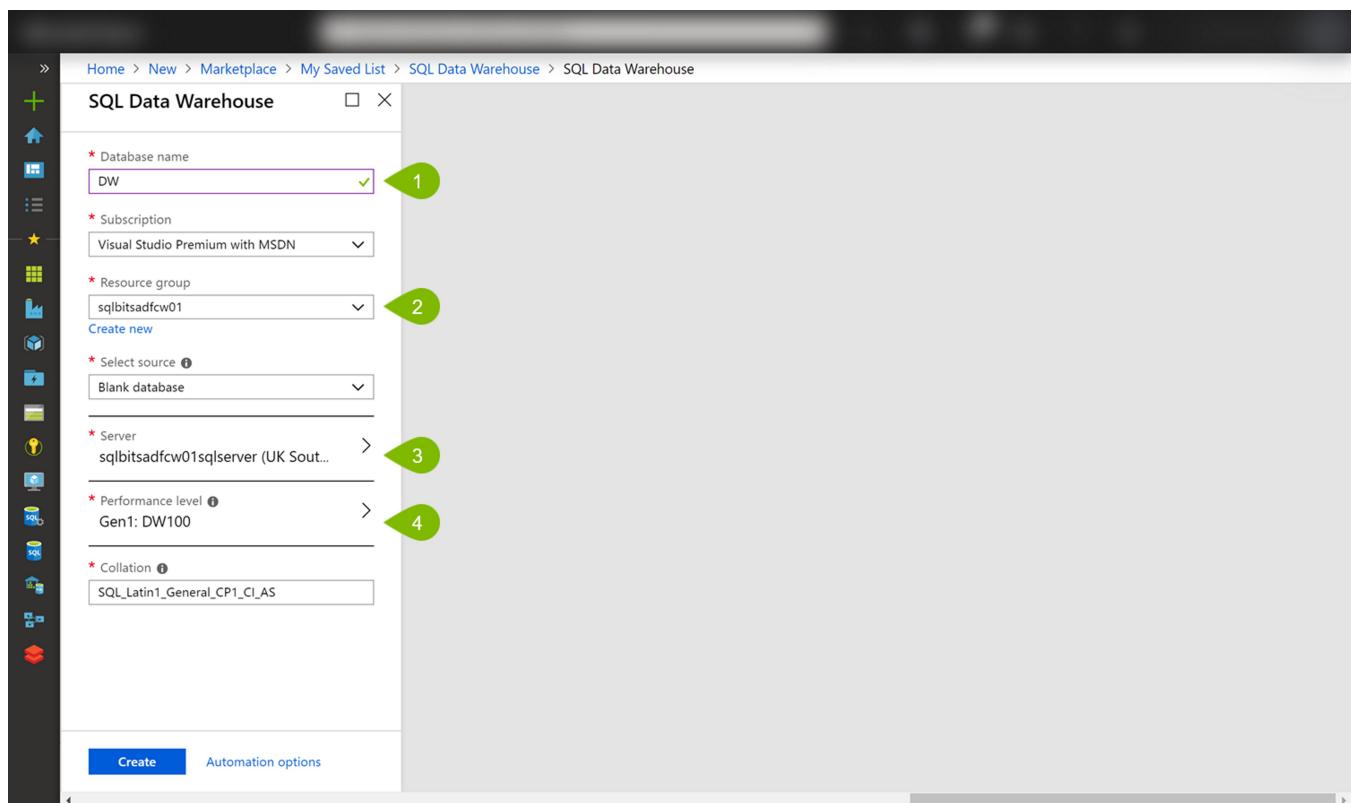
Create the Azure SQL Data Warehouse in the Azure SQL Server created in the previous step.

Please note: It is very important to ensure this resource is paused after creation, as it will incur costs if left running.

Step	Setting	Value	Notes
1	Database Name	DW	
2	Resource Group	sqlbitsadf<999>	Choose the Resource Group you created in the first step.

3	Server	sqlbitsadf<999>sqlserver	Choose the SQL Server you created in the previous step.
4	Performance Level	Gen1: DW100	Change the default Gen2 DW1000c to the cheaper Gen1 DW100. (The cheapest Gen2 is not available by default in the UK South region.)

Unless specified above, use the default settings.



Additional information:

<https://docs.microsoft.com/en-us/azure/sql-data-warehouse/create-data-warehouse-portal>

Once the Azure SQL Data Warehouse has been deployed, pause it immediately.

The screenshot shows the Azure portal interface for a SQL data warehouse named 'DW'. The left sidebar contains navigation links for Overview, Activity log, Tags, Diagnose and solve problems, Settings, Properties, Locks, Automation script, Security, and Common Tasks. The main content area displays the 'Essentials' tab with a 'DWU Usage' chart. The chart shows 'DWU limit (Max)' at 100 and 'DWU used (Max)' at 0. A legend indicates 'DWU' and 'DW'. Below the chart are three tabs: Notifications (0), Features (4), and Tasks (7). Under the Features tab, there are two items: 'Transparent data encryption' (NOT CONFIGURED) and 'Auditing' (NOT CONFIGURED).

Please note: It is very important to ensure this resource is paused after creation, as it will incur costs if left running.

8. Create Azure Analysis Services

Azure Analysis Services integrates with many Azure services enabling you publish Tabular models to Azure.

Please note: It is very important to ensure this resource is paused after creation, as it will incur costs if left running.

Step	Setting	Value	Notes
1	Server Name	sqlbitsadf<999>aas	Use the name of the Resource Group and add aas as a suffix.
2	Resource Group	sqlbitsadf<999>	Choose the Resource Group you created in the first step.
3	Location	UK South	Choose UK South .
4	Pricing Tier	D1	Choose D1 .
5	Administrator	<User>	Choose your own user.

Unless specified above, use the default settings.

Analysis Services

* Server name 1

* Subscription

* Resource group 2

* Location 3

* Pricing tier 4

* Administrator 5

Backup Storage Settings
Backup Storage: Not configured

Storage key expiration
Never

Create [Automation options](#)

Additional information:

<https://docs.microsoft.com/en-us/azure/analysis-services/analysis-services-create-server>

Once the Azure Analysis Services has been deployed, pause it immediately.

sqlbitsadfcw01aas

New model **Pause** Move Delete

Overview

Activity log Access control (IAM) Tags Diagnose and solve problems

Scale

Pricing Tier (Scale QPUs) Replicas

Models

Manage

Settings

Quick Start Analysis Services Admins On-Premises Data Gateway Backups Connection Strings Firewall

Essentials

Resource group [\(change\)](#)
sqlbitsadfcw01

Status
Active

Location
UK South

Subscription name [\(change\)](#)
Visual Studio Premium with MSDN

Subscription ID

Server name
asazure://uksouth.asazure.windows.net/sqlbitsadfcw01aas

Management Server Name
asazure://uksouth.asazure.windows.net/sqlbitsadfcw01aas:rw

Pricing tier
D1

Connection strings
[Show server connection strings](#)

Web designer — preview
[Open](#)

Models on Analysis Services Server

NAME	COMPATIBILITY	DATE MODIFIED	LAST SYNCED TIME	SYNC STATE
No results				

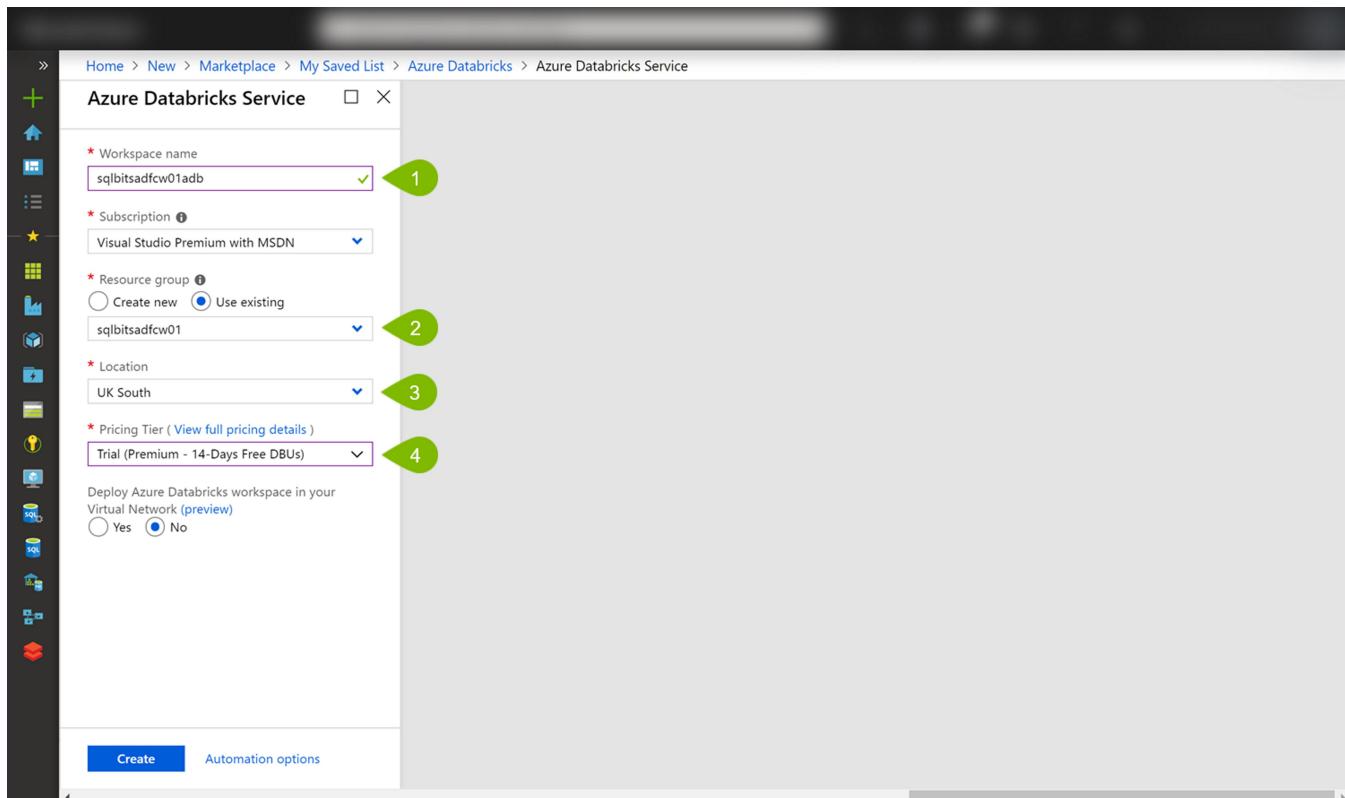
Please note: It is very important to ensure this resource is paused after creation, as it will incur costs if left running.

9. Create Azure Databricks Service

Azure Databricks is an Apache Spark-based analytics platform optimized for the Microsoft Azure cloud services platform.

Step	Setting	Value	Notes
1	Workspace Name	sqlbitsadf<999>adb	Use the name of the Resource Group and add adb as a suffix.
2	Resource Group	sqlbitsadf<999>	Choose the Resource Group you created in the first step.
3	Location	UK South	Choose UK South .
4	Pricing Tier	Trial (Premium - 14 Days Free)	Choose Trial (Premium - 14 Days Free DBUs) . If this is no longer an option for you, choose Standard .

Unless specified above, use the default settings.



Additional information:

<https://docs.azuredatabricks.net/getting-started/try-databricks.html>

10. Create Virtual Machine (Free SQL Server License: SQL Server 2017 Developer on Windows Server 2016) with AdventureWorksLT Demo Database

Create a virtual machine with SQL Server 2017 and the AdventureWorksLT database that will act as your on-premises source.

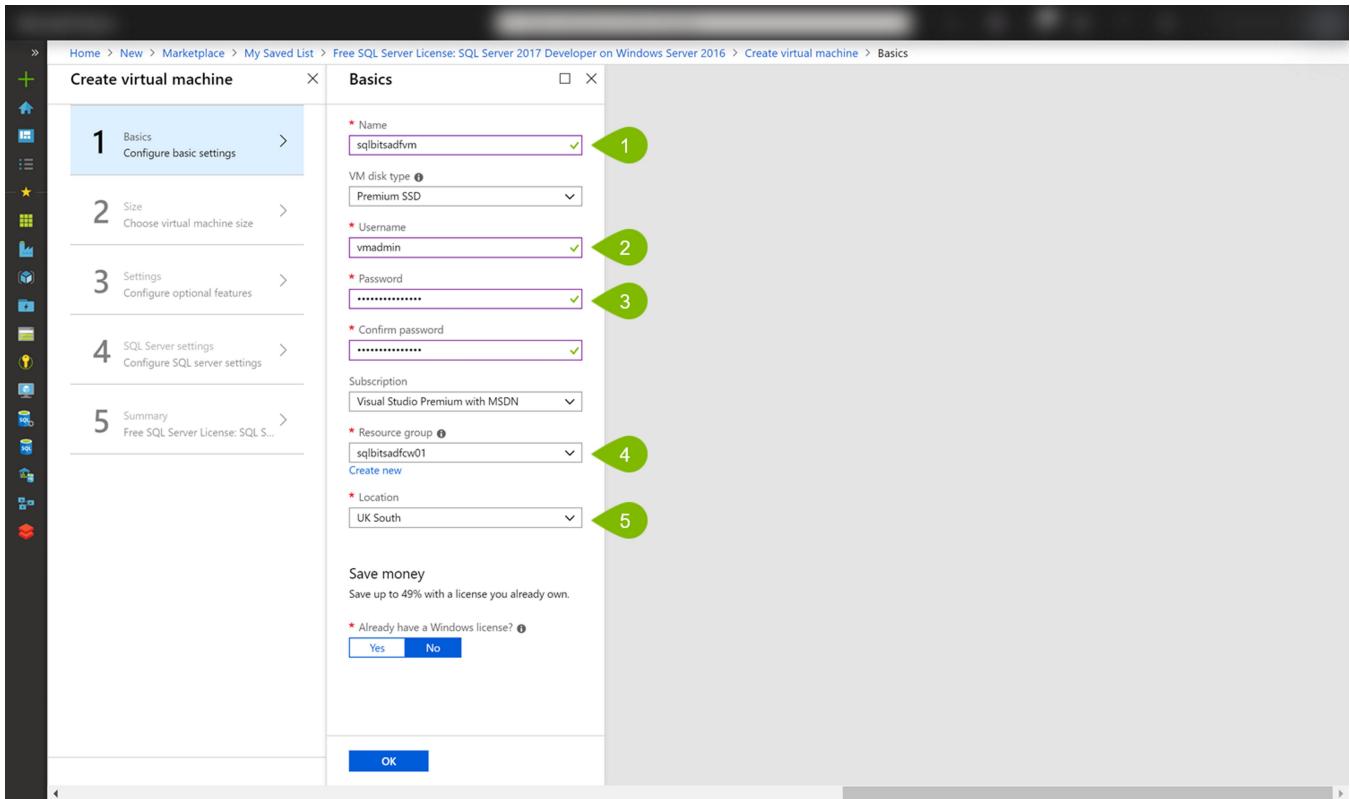
Please note: It is very important to ensure this resource is stopped after creation, as it will incur costs if left running.

This process consists of multiple parts.

Part 1: Basics

Step	Setting	Value	Notes
1	Name	sqlbitsadfvm	Please note that the name has to be between 1-15 characters, so you may not be able to use your chosen suffix.
2	Username	vmadmin	Pick a username, for example vmadmin
3	Password	VeryStr0ngPassw0rd!	Pick a password, for example VeryStr0ngPassw0rd!
4	Resource Group	sqlbitsadf<999>	Choose the Resource Group you created in the first step.
5	Location	UK South	Choose UK South .

Unless specified above, use the default settings.



Part 2: Size

Choose **D2s_v3**.

Create virtual machine

Choose a size

Browse the available sizes and their features

RECOMMENDATION	SKU	TYPE	COMPUTE TYPE	VCPUS	GB RAM	DATA DISK CAPACITY	MAX IOPS	LOCAL SSD CAPACITY	Premium	vCPUs	NOK/MONTH
B2ms	Standard	General purpose	2	8	4	2400	16 GB	Yes		1,2,3	NOK 569.88
B2s	Standard	General purpose	2	4	4	1600	8 GB	Yes		1,2,3	NOK 284.94
B4ms	Standard	General purpose	4	16	8	3600	32 GB	Yes		1,2,3	NOK 1,140.97
B8ms	Standard	General purpose	8	32	16	4320	64 GB	Yes		1,2,3	NOK 2,281.94
D16s_v3	Standard	General purpose	16	64	32	25600	128 GB	Yes		1,2,3	NOK 5,602.23
D2s_v3	Standard	General purpose	2	8	4	3200	16 GB	Yes		1,2,3	NOK 700.28
D4s_v3	Standard	General purpose	4	16	8	6400	32 GB	Yes		1,2,3	NOK 1,400.56
D8s_v3	Standard	General purpose	8	32	16	12800	64 GB	Yes		1,2,3	NOK 2,801.12
DS1_v2	Standard	General purpose	1	3.5	4	3200	7 GB	Yes		1,2,3	NOK 531.25
DS11_v2	Standard	Memory optimized	2	14	8	6400	28 GB	Yes		1,2,3	NOK 1,412.63
DS11_v2	Promo	Memory optimized	2	14	8	6400	28 GB	Yes		1,2,3	NOK 1,412.63
DS11-1_v2	Standard	Memory optimized	1	14	8	6400	28 GB	Yes		1,2,3	NOK 1,412.63
DS12_v2	Promo	Memory optimized	4	28	16	12800	56 GB	Yes		1,2,3	NOK 2,831.30
DS12_v2	Standard	Memory optimized	4	28	16	12800	56 GB	Yes		1,2,3	NOK 2,837.34

Prices presented are estimates in your local currency that include only Azure infrastructure costs and any discounts for the subscription and location. The prices don't include any applicable software costs. Final charges will appear in your local currency in cost analysis and billing views. Recommended sizes are determined by the publisher of the selected image based on hardware and software requirements.

Select

Part 3: Settings

Select the RDP public inbound port. Keep all other defaults.

Create virtual machine

Settings

High availability

Availability zone (Optional)

None

* Availability set (Optional)

None

Storage

Use managed disks (Optional)

No Yes

Network

* Virtual network (Optional)

(new) squbitsadfcw01-vnet

* Subnet (Optional)

default (10.0.0.0/24)

* Public IP address (Optional)

(new) squbitsadfvml-p

Network Security Group (Optional)

Basic Advanced

* Select public inbound ports (Optional)

RDP (3389) ▲

No public inbound ports □

HTTP (80) □

HTTPS (443) □

SSH (22) ▲

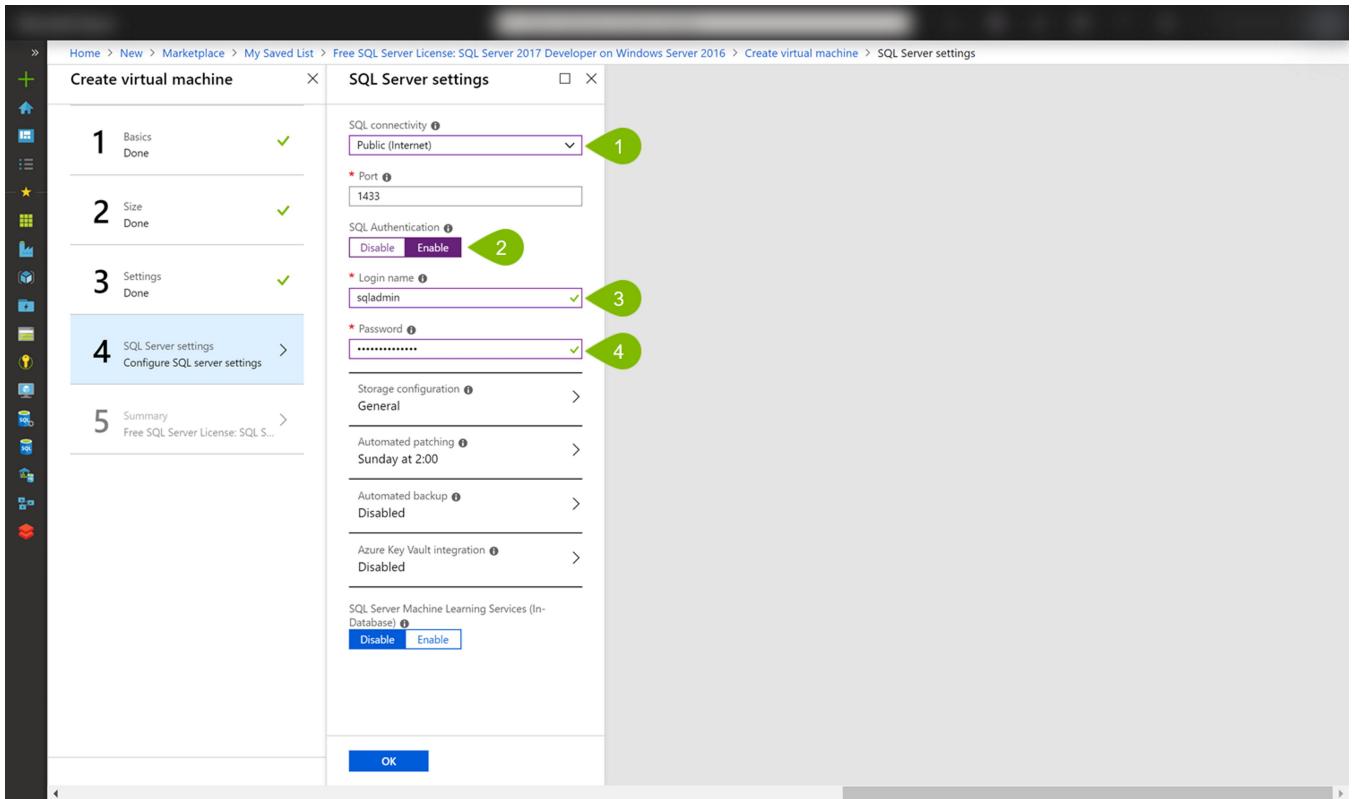
OK

Part 4: SQL Server Settings

Enable public access using SQL authentication.

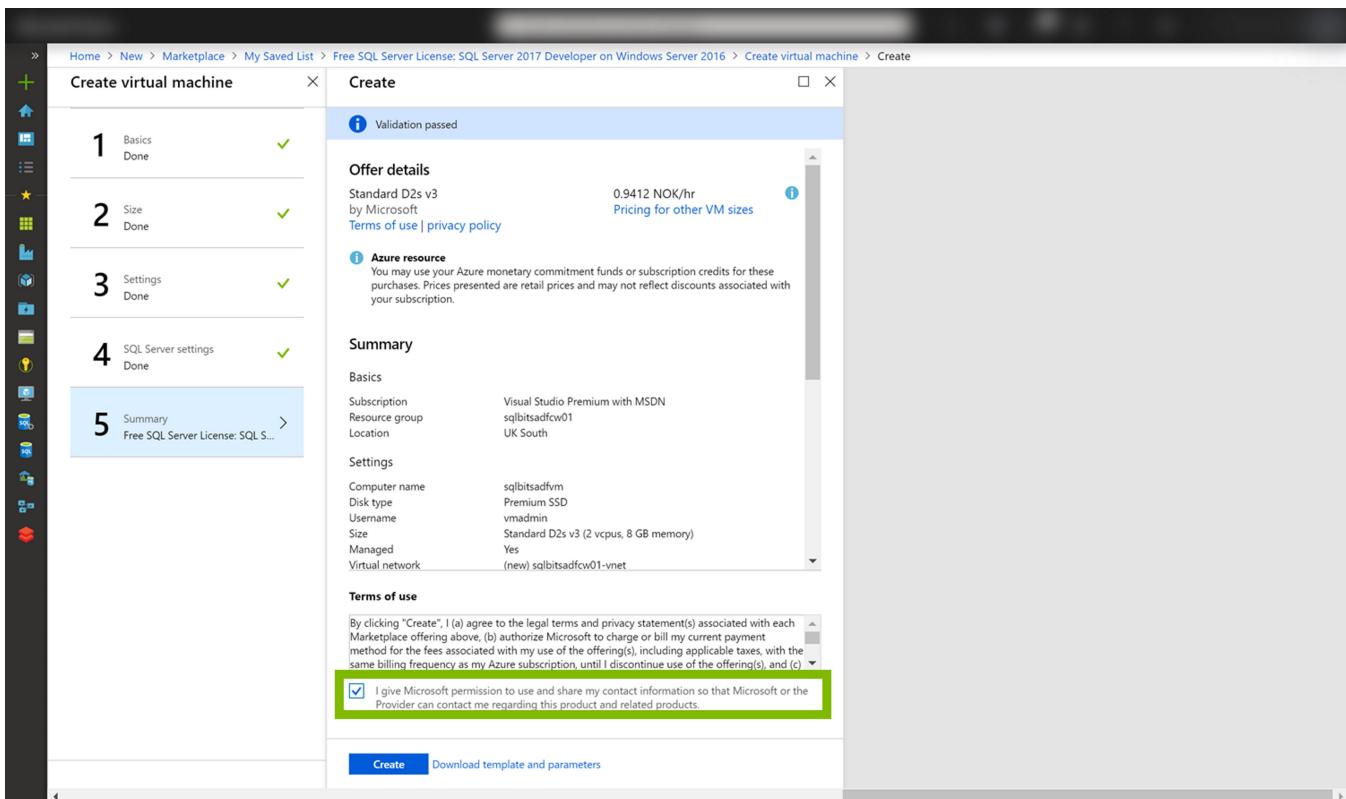
Step	Setting	Value	Notes
------	---------	-------	-------

1	SQL Connectivity	Public (Internet)	We will be using this public setting to simplify demos.
2	SQL Authentication	Enabled	Choose the Resource Group you created in the first step.
3	Username	sqladmin	Pick a username, for example sqladmin
4	Password	VeryStr0ngPassw0rd!	Pick a password, for example VeryStr0ngPassw0rd!



Part 5: Summary

Accept the terms of use and create the VM.



Change the Public IP Address from Dynamic to Static

In this workshop, we don't want to keep our Virtual Machine running all the time as that will incur costs. To be able to start and stop the virtual machine and reconnect to it without having to change the connection settings every time, we need to change the IP address settings. From the Virtual Machine page, click on the **Public IP Address**.

Setting	Value
Computer name	: sqlbitsadfvvm
Operating system	: Windows
Size	: Standard D2s v3 (2 vcpus, 8 GB memory)
Public IP address	: 51.140.105.31
Virtual network/subnet	: sqlbitsadfcw01-vnet/default
DNS name	: Configure

Change the Assignment from **Dynamic** to **Static**. Click **Save**.

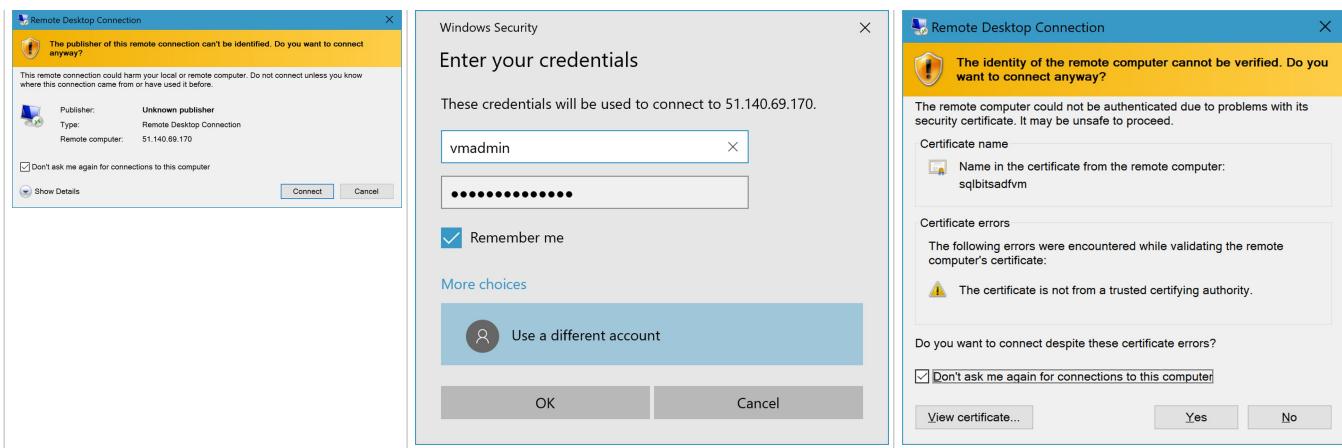
The screenshot shows the 'Configuration' page for a public IP address associated with a virtual machine. The 'Assignment' section is highlighted with a green box and a number '1'. It shows that 'Static' is selected for the IP address assignment. A green circle labeled '2' points to the 'Save' button at the top right. A warning message at the top states: 'The associated virtual machine 'sqlbitsadfvm' may be rebooted. Click here to learn more.'

Connect to the Virtual Machine

Go back to the Virtual Machine page. Click **Connect** and download the **RDP** file.

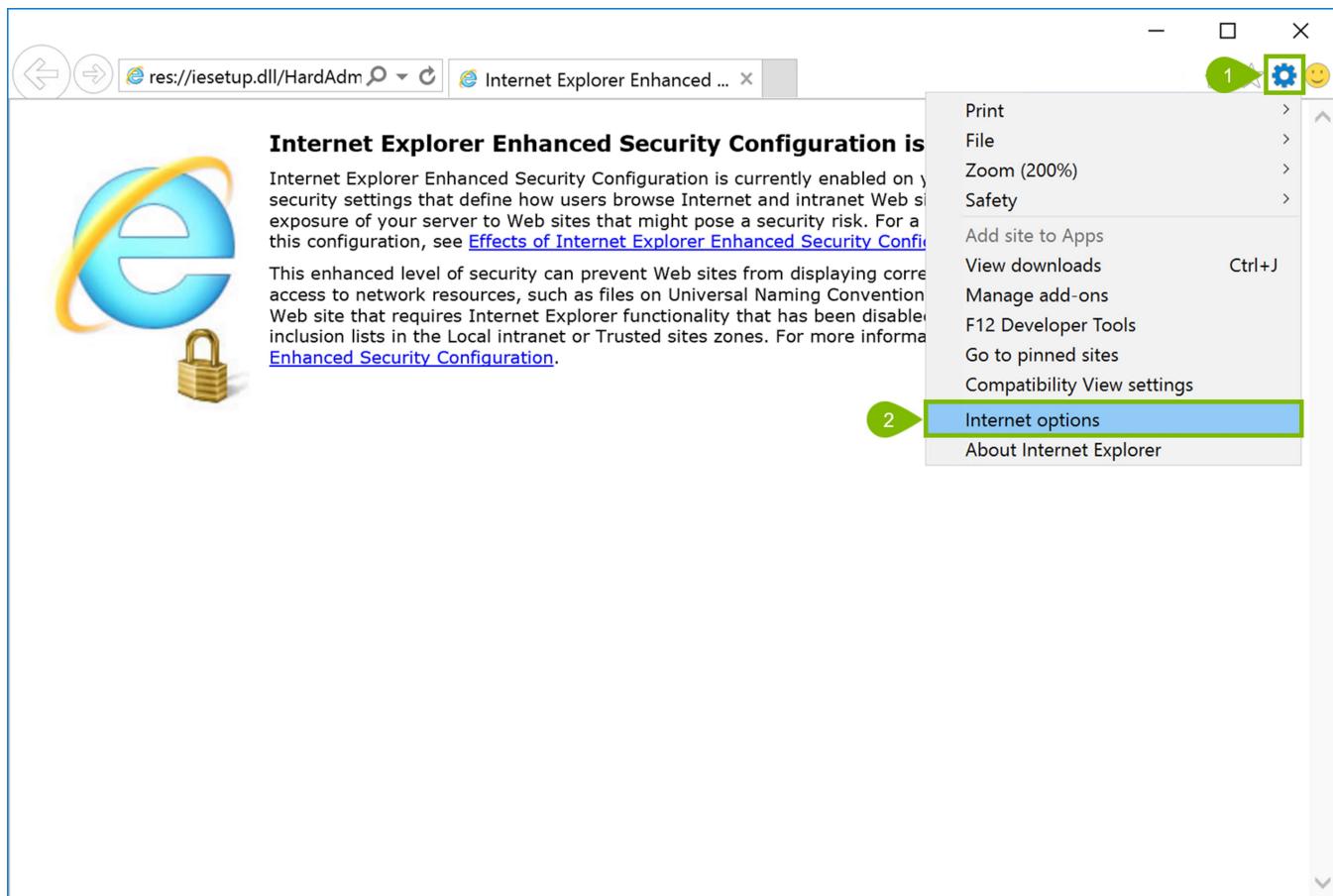
The screenshot shows the 'Virtual machines' page for the 'sqlbitsadfvm' VM. The 'Connect' button in the top navigation bar is highlighted with a green box and a number '1'. An 'RDP' tab is selected in the 'Connect to virtual machine' overlay. The 'Download RDP File' button is highlighted with a green box and a number '2'. A warning message in the overlay says: 'To improve security, enable just-in-time access on this VM.' Below the overlay, there is a note about inbound traffic being blocked and a link to troubleshoot connection issues.

Open the RDP file, click Connect, then choose **Use a different account** under More choices, enter your username and password (for example **vmadmin / VeryStr0ngPassw0rd!**), and click Yes.

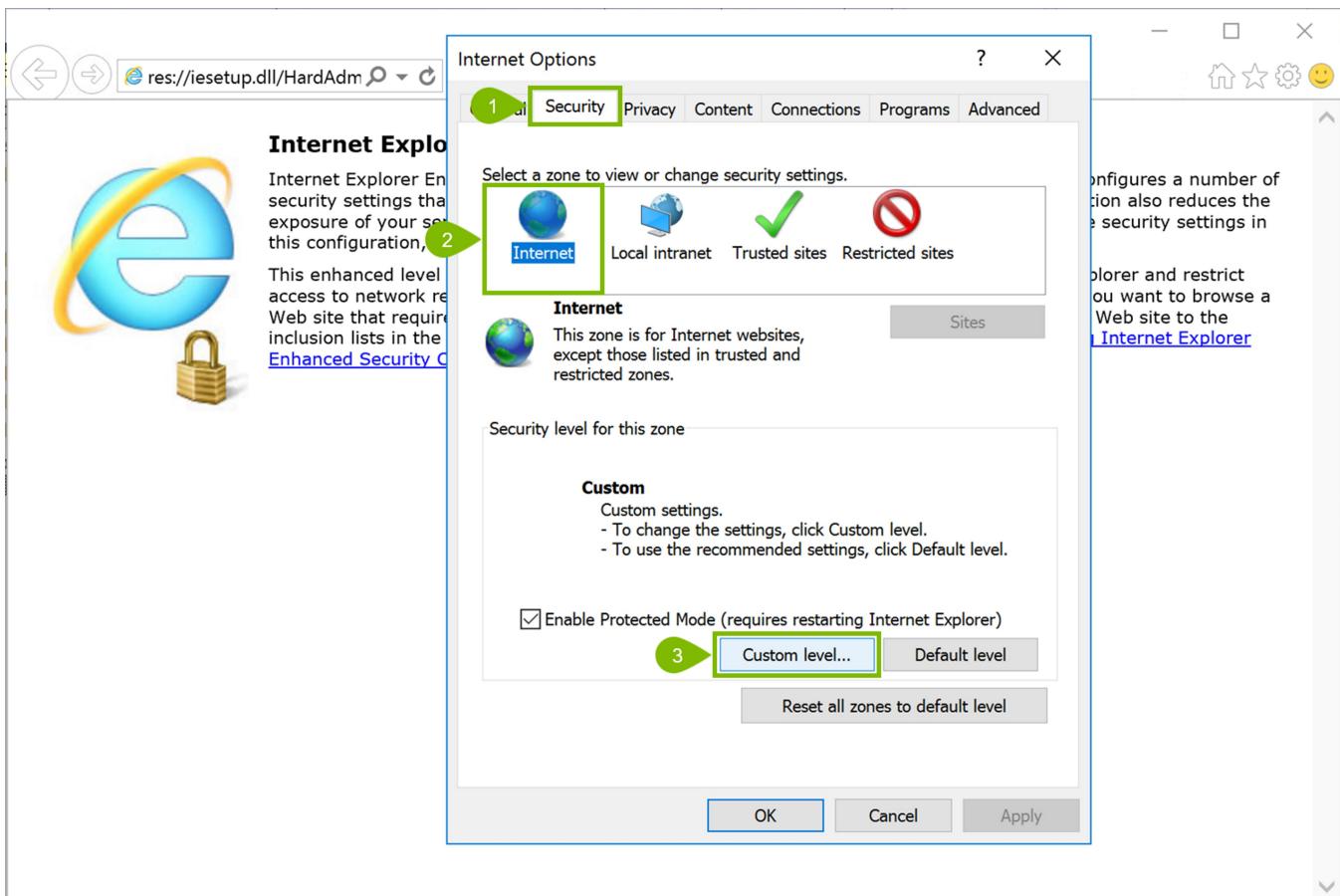


Download the AdventureWorksLT Demo Database

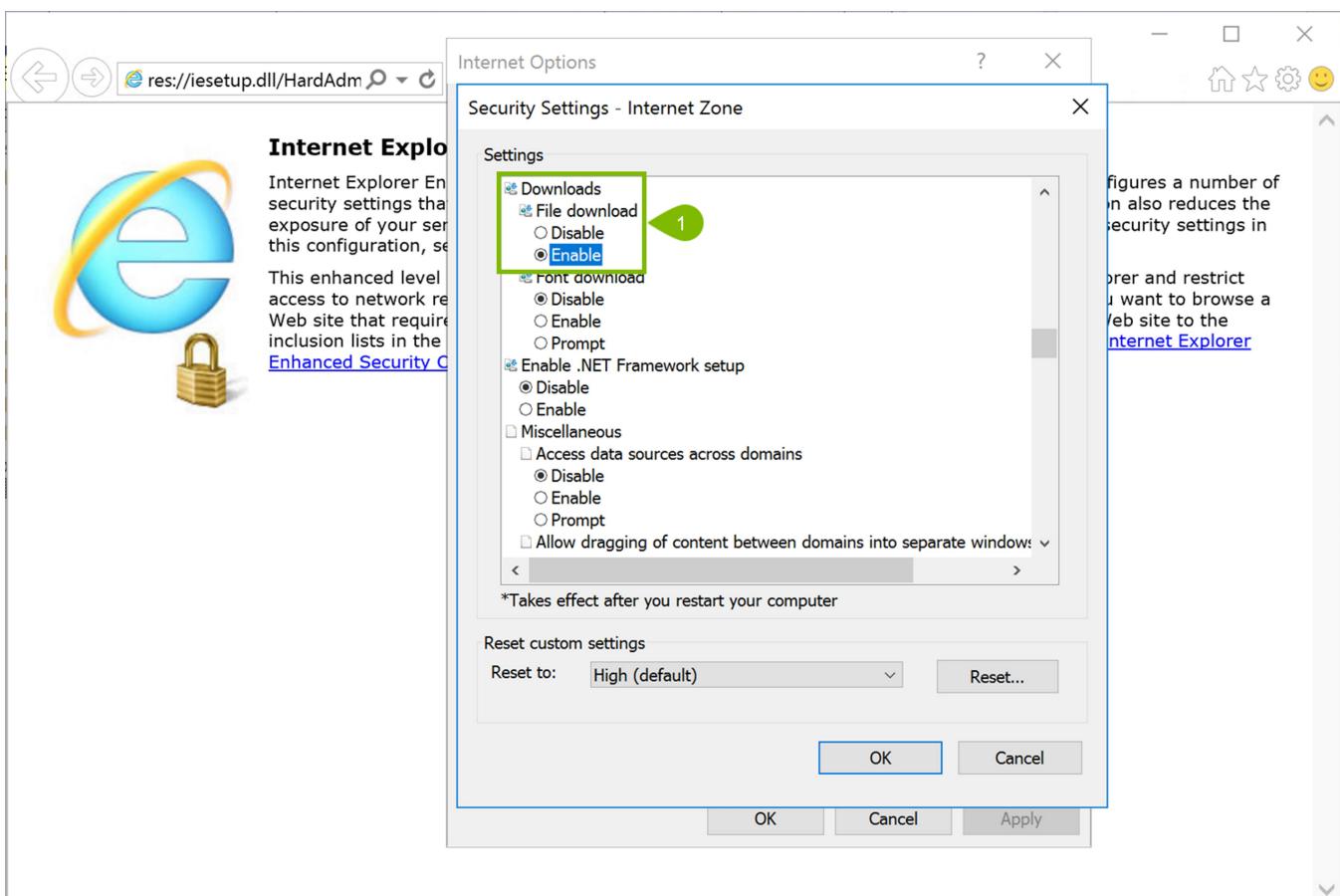
Once inside the Virtual Machine, open Internet Explorer. To be able to download the AdventureWorksLT demo database, we first need to change the security settings. Click **Settings**, then **Internet Options**.



Click the **Security** tab, select **Internet**, then click **Custom level**.

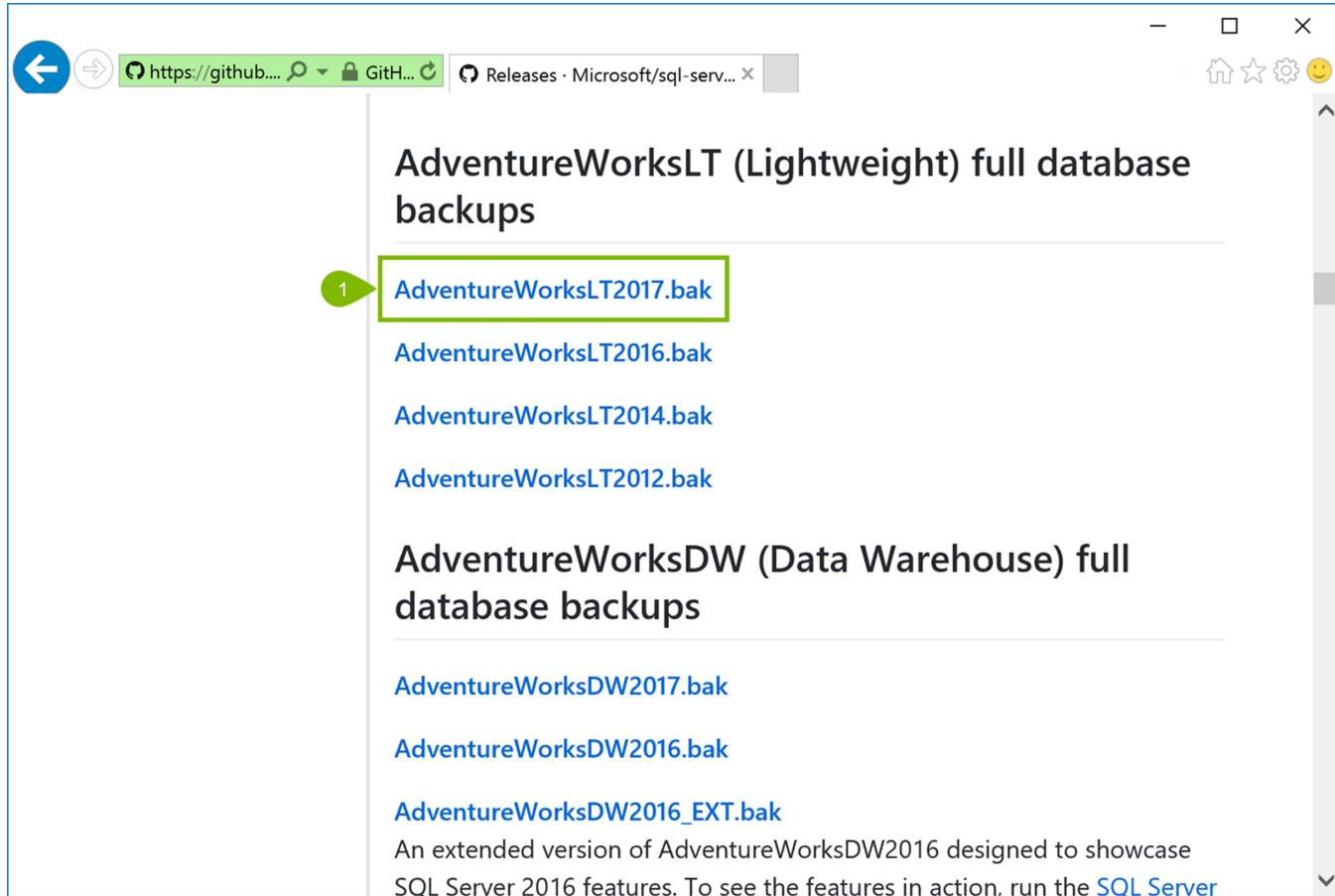


Scroll down to **Downloads**, choose **Enable**.



Scroll further down to **Scripting**, choose **Enable**. Click Yes to confirm you want to change the settings for the zone, click OK to close the Settings windows, and OK again to close the Internet Options window.

Go to <https://github.com/Microsoft/sql-server-samples/releases>, scroll down, and download **AdventureWorksLT2017.bak** to D:\.



The screenshot shows a web browser window with the GitHub URL in the address bar. The page displays two sections of database backups:

- AdventureWorksLT (Lightweight) full database backups**
 - [AdventureWorksLT2017.bak](#) (highlighted with a green border and arrow)
 - [AdventureWorksLT2016.bak](#)
 - [AdventureWorksLT2014.bak](#)
 - [AdventureWorksLT2012.bak](#)
- AdventureWorksDW (Data Warehouse) full database backups**
 - [AdventureWorksDW2017.bak](#)
 - [AdventureWorksDW2016.bak](#)
 - [AdventureWorksDW2016_EXT.bak](#)

An extended version of AdventureWorksDW2016 designed to showcase SQL Server 2016 features. To see the features in action, run the [SQL Server](#)

Open **SQL Server Management Studio (SSMS)**, connect to **localhost**, and restore the **AdventureWorksLT** database.

```
USE [master]
RESTORE DATABASE [AdventureWorksLT]
FROM DISK = N'D:\AdventureWorksLT2017.bak'
WITH FILE = 1,
MOVE N'AdventureWorksLT2012_Data' TO N'F:\Data\AdventureWorksLT2012.mdf',
MOVE N'AdventureWorksLT2012_Log' TO N'F:\Log\AdventureWorksLT2012_log.ldf',
NOUNLOAD,  STATS = 5
GO
```

Open **SQL Server Management Studio (SSMS)**, connect to **localhost**, and restore the following query in the **master** database.

```
USE MASTER
GO

-- Azure ADF Loader Account
CREATE LOGIN DataLoadUser WITH PASSWORD = 'DataFactoryD3m0!';
GO

USE AdventureWorksLT
GO

CREATE USER DataLoadUser FOR LOGIN DataLoadUser;
```

```
-- Create Security

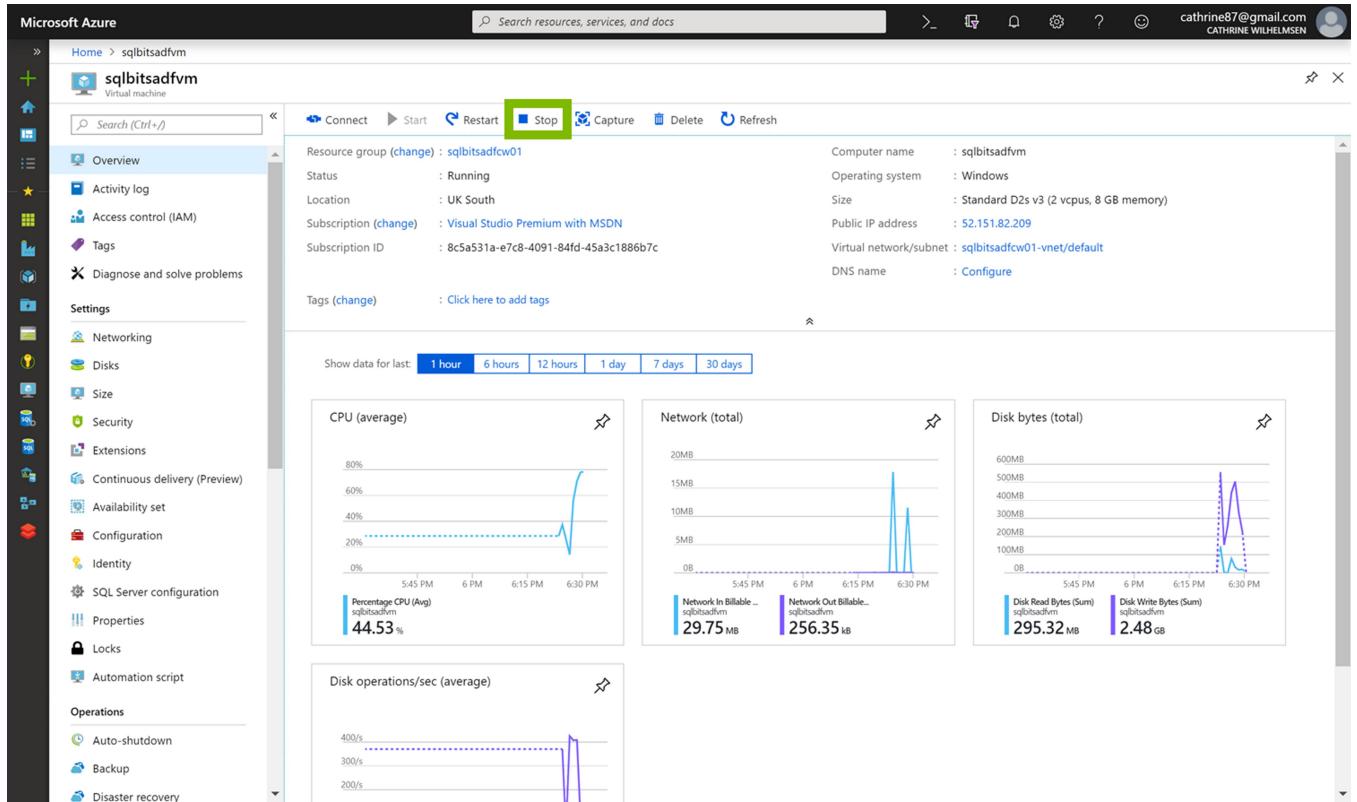
CREATE ROLE Executor;
GRANT EXECUTE TO [Executor]

CREATE ROLE Reader;
GRANT SELECT TO [Reader];
EXEC sp_addrolemember 'Executor', 'Reader';

GRANT VIEW CHANGE TRACKING ON Schema::SalesLT TO [Reader];
GRANT VIEW DEFINITION ON Database::AdventureWorksLT to [Reader];

EXEC sp_addrolemember 'Reader', 'DataLoadUser';
```

Once the AdventureWorksLT database has been restored, exit the Virtual Machine and stop it.



Additional information:

<https://docs.microsoft.com/en-us/azure/virtual-machines/windows/sql/quickstart-sql-vm-create-portal>

Please note: It is very important to ensure this resource is stopped after creation, as it will incur costs if left running.

Post-Setup: Ensure all Resources are Paused or Stopped

	<ul style="list-style-type: none"> Azure SQL Data Warehouse (<i>Paused</i>) Azure Analysis Service (<i>Paused</i>) Virtual Machine (<i>Stopped</i>)
--	--