wortell

Lab 1 - Preparing the Azure environment

Requirements

To start this training, it's important that you have your own Azure environment with sufficient rights to set up Azure Data Factory and its associated resources.

Goal

During the training, you need a working environment to be able to get **hands-on** experience with the Azure Data Factory.

In this series of assignments, the environment is set up, you prepare the data, and set up the database. Follow the assignments step by step.

Assignment 1 - Deploy Azure services

1. Click on the **deploy to azure** button below. If you have received the handout physically, go to the Github repo indicated by the trainer.

   [⯅ **Deploy** to Azure]

2. In the deploy screen, all required fields are already filled with values. You only need to create a resource group by clicking on **create new**. Fill in a name and then click OK.

> A *best practice* is to start the name of a resource group with `rg-` followed by a descriptive name and ending with a `-environment`.
>
> - Practical example:* `rg-dataplatform-dev`
> - Training example:* `rg-adf-training`

3. Click on the **Review + Create** button. The code will then be validated. If the validation is successful, this will be indicated by a green bar and the **Create** button will become available. Click on the button and the deployment of the environment will start. If anything goes wrong during the deployment, let the trainer know.

4. Once the deployment is complete, you will be notified and hooray! Your environment is ready to use!

Assignment 2 - Preparing data part 1

1. In the Github folder of Lab1, you'll find the LabEnvironment folder. Herein:
   - A SQL file for the database
   - Two CSV files named **SalesPersonal.csv** and **ProductCategoryDiscount.csv**
2. Download the Zip file for the entire repository using the following link: Download ZIP
3. Both files should now be on your local machine, probably in the downloads folder.
4. Go back to the Azure Portal, and search in your resource group for the **Virtual Machine** and click on it.
5. In the horizontal navigation bar, you'll see a few options such as:

- Connect
- Start
- Stop
- Capture
- Delete
- etc

6. Start the VM if it has not started yet
7. Click on **Connect** and choose **RDP**.
8. You are now brought to a new screen and the required fields are already filled in for you. Then click on **Download RDP File**.
9. When the download is complete, click on the RDP file and then click on **Connect**. A new screen will appear asking for a username and password, fill in the following details:
   - *Username:* `sqladmin`
   - *Password:* `WortellSmartLearning.nl`
     There is a chance that Windows tries to log in with your domain account (the account of your own organization). In that case, choose *more choices --> Use a different account*:
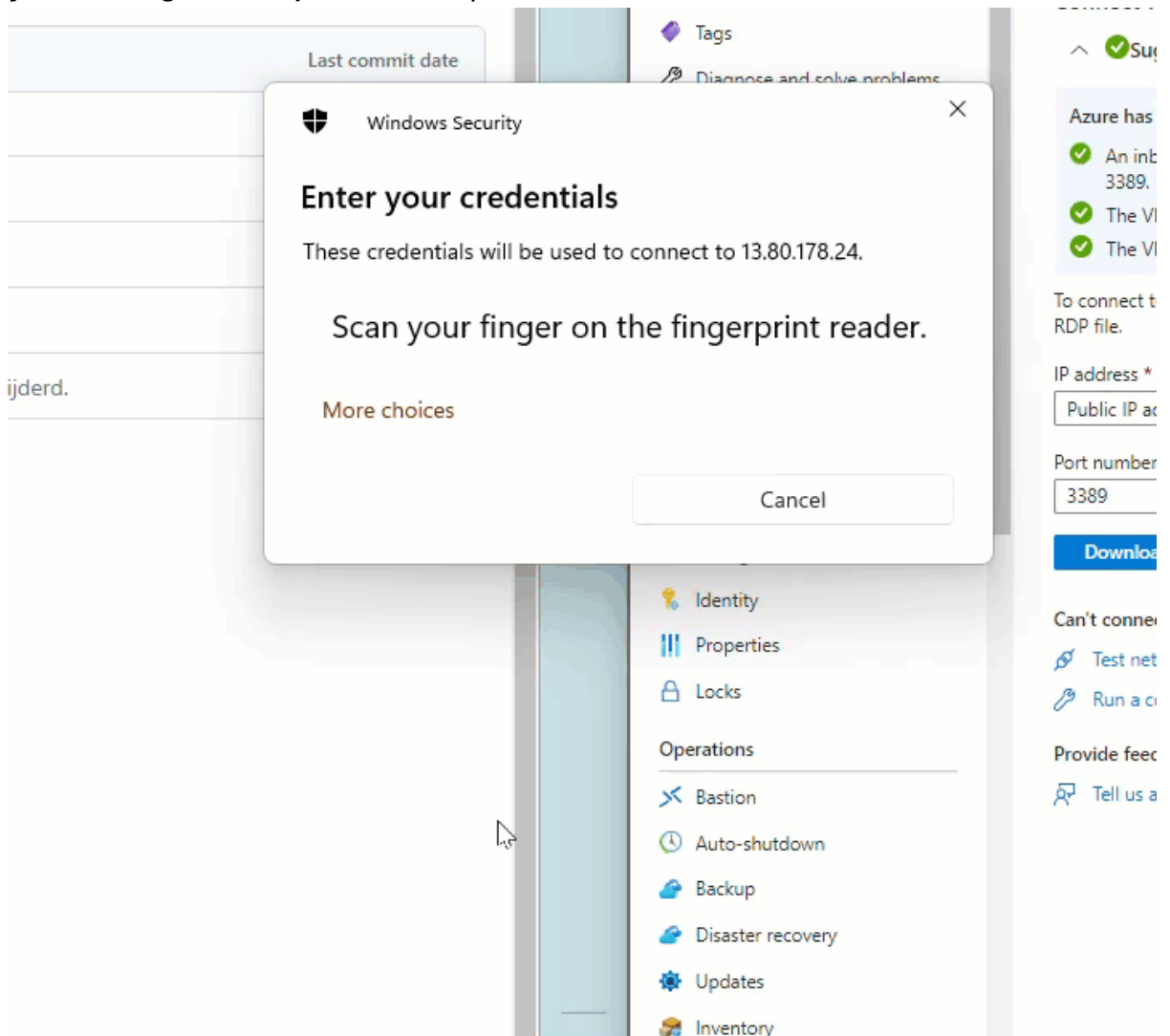
10. After filling in the details, click OK.
11. A new screen may appear to trust the certificate, then click on **Yes**.
12. A connection is now being established with your Virtual Machine. This can take a while and there's a chance that a screen comes up with sliders to accept certain things. Set everything to **No** and click **Accept**.

13. The desktop of the Virtual Machine is loaded. Navigate to the desktop after loading and click on **File explorer** (The yellow folder). Then click on **This PC** followed by clicking on **Temporary Storage(D:)**.
14. Copy the downloaded file named **SalesPersonal.csv** and paste it on the D-drive of the Virtual Machine.
    - You don't need to do anything with it right now
    - Minimize the screen, you will need the Virtual Machine later today.

Assignment 3 - Preparing data part 2

14. Go back to the Azure portal (portal.azure.com) and search within your resource group for the **Storage Account** and click on it.
15. In the Blade (Navigation bar on the left), under the header **Data storage**, you'll find the **Containers** option, click on it.
16. You will see that there is already a container named **data**, click on it. The container is still empty and we are going to fill it with the remaining file named **ProductCategoryDiscount.csv**.
17. In the horizontal navigation bar, you'll see a few options such as:
    - Upload
    - Change access level
    - Refresh
18. Click on **Upload**, click on **browse for files** and select the file named **ProductCategoryDiscount.csv** or drag and drop the file from explorer to the drag 'n drop area in the portal and then click on the **Upload** button.
19. The file is uploaded and should appear in the container within a few seconds.

Assignment 4 - Setting up the database

20. Go back to your resource group. In the list, you'll see 2 databases (sqldb-source and sqldb-target) and 1 server. The sqldb-source is already set up and filled with data (AdventureWorksLT).
21. Click on the **sqldb-target** database and then in the Blade on **Query editor (preview)**.
    - The login name of the sqladmin account should already be filled in, then enter the password. This is the same password as in Assignment 2.
22. Once logged in, you'll see a query interface in front of you and the tables, views, and stored procedure folders on the left.
23. Go back to the Github folder of Lab1 and in the **LabEnvironment** folder. Then click on the .sql file named **Target-Database.sql**. The code should now be visible, select all code and paste it into the Query editor and then click on the **Run** button.
24. When the query has been executed, your database is ready for use!

Table of Contents

9. Batching and DIUs