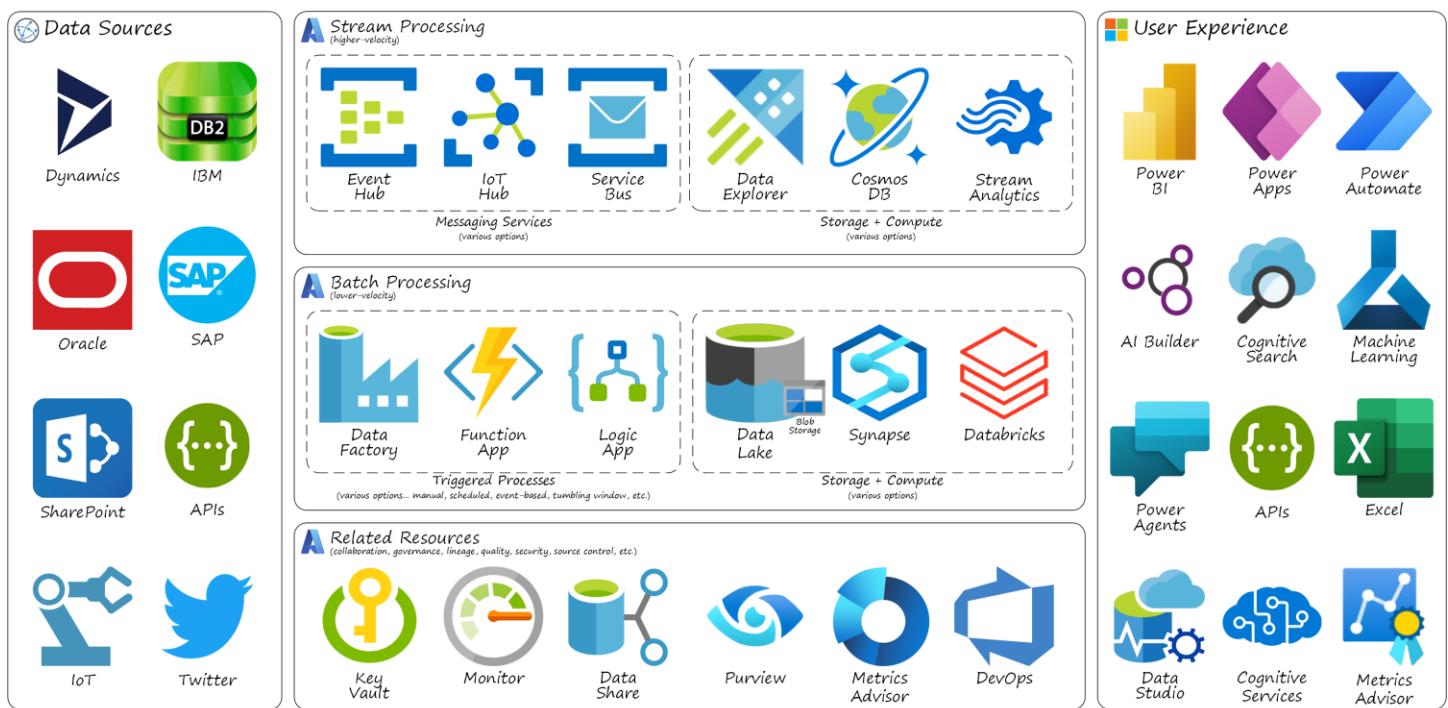


Azure Data Practicum

This document is best viewed in Microsoft Word, "Web Layout" view with Navigation pane turned on

This practicum is an opportunistic capture of things I have learned while working with a variety of amazing solutions.



Sections and Objectives

This document contains many **product | activity-focused objectives** organized into **discipline-specific sections**.

Consider navigating directly to those sections and objectives that align with your focus and interest.

Data Engineering

Product(s)	Activity	Also Applies To...
Synapse (Data and Develop)	Use External Data	
Synapse (Integrate) + Data Lake	Simple Data Ingestion	Data Factory
Synapse (Integration) + SQL + Data Explorer	Simple Data Pipeline	Data Factory
Synapse (Integrate)	Simple Conditional Pipeline	Data Factory
Databricks + Data Lake	Mount Data	
Databricks	Localize Mounted Data	
Databricks	Source from APIs	
Databricks	Batch Upsert Data	
Synapse	Audit Data Usage	
Data Explorer	Query from On-Prem	

Data Analysis

Metrics Advisor + Data Explorer

[Onboard and Investigate](#)

Data Application

SQL + Power Apps + AI Builder	Capture Data	
-------------------------------	------------------------------	--

Data Governance

Purview	Discover Data	
Purview	Classify Data	
Purview	Understand Lineage	

Data DevOps

Synapse + DevOps	Source Control	Data Factory
Synapse	Schema Comparison	

Caveats

- This guidance is designed only to convey foundational knowledge that you can tailor to fit your environment, standards, etc.
- **Keep a watchful eye on incurred costs** ... consider a daily assessment and use of budgets / alerts
- Azure interface and functionality evolve rapidly; **captured snips will become dated over time**

Data Engineering

Objectives in this section describe methods for extracting, transforming, and loading data.

Objective: Synapse | Use External Data

Follow the instructions in this section to **explore methods of using data from assets external to Synapse**.

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Data Lake (with sample data)
- Synapse (with Serverless SQL Database and Apache Spark Pool)

Navigate to your Synapse Analytics workspace and then click the **Open** link on the “Open Synapse Studio” rectangle.

Click the **Data** icon in the left-hand navigation.

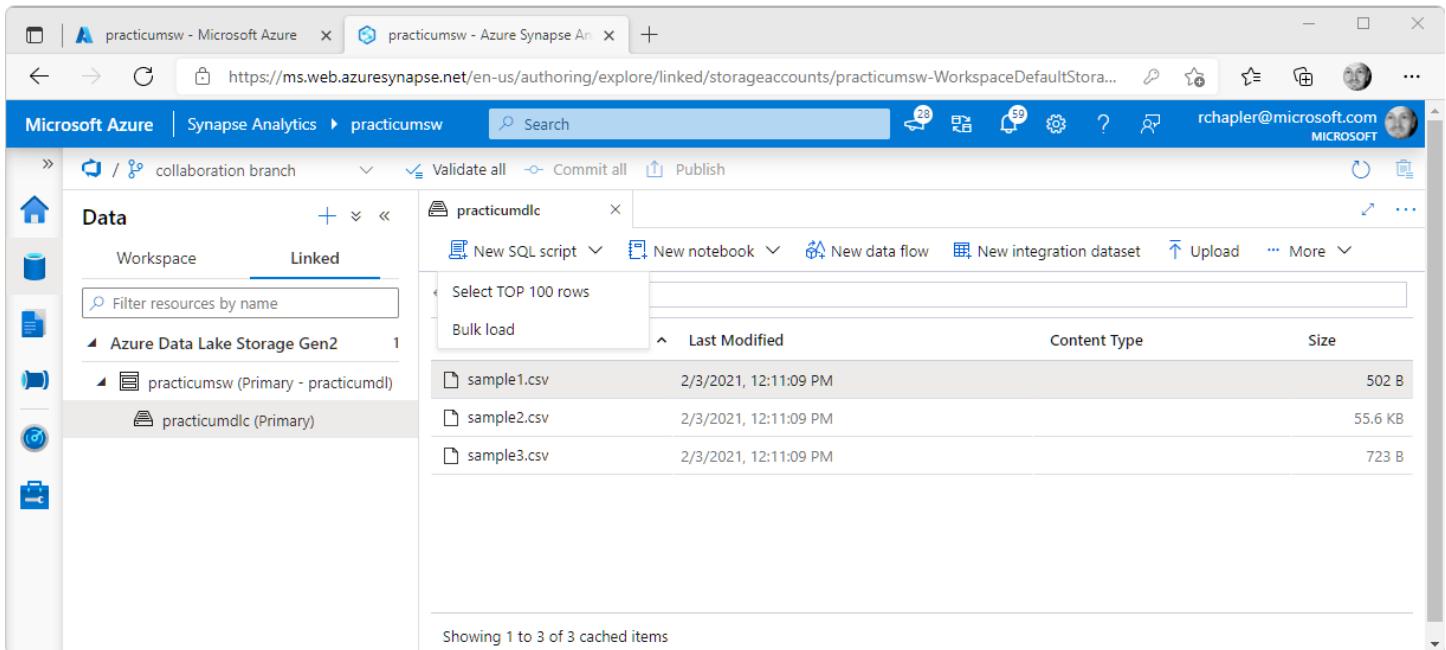
The screenshot shows the Microsoft Azure Synapse Analytics Data Studio interface. The left sidebar has a 'Data' icon selected. The main area shows a 'Linked' tab under 'practicumdlc'. A table lists three CSV files: sample1.csv, sample2.csv, and sample3.csv, all modified on 2/3/2021 at 12:11:09 PM. The table includes columns for Name, Last Modified, Content Type, and Size.

Name	Last Modified	Content Type	Size
sample1.csv	2/3/2021, 12:11:09 PM		502 B
sample2.csv	2/3/2021, 12:11:09 PM		55.6 KB
sample3.csv	2/3/2021, 12:11:09 PM		723 B

Click the **Linked** tab and Expand navigation.

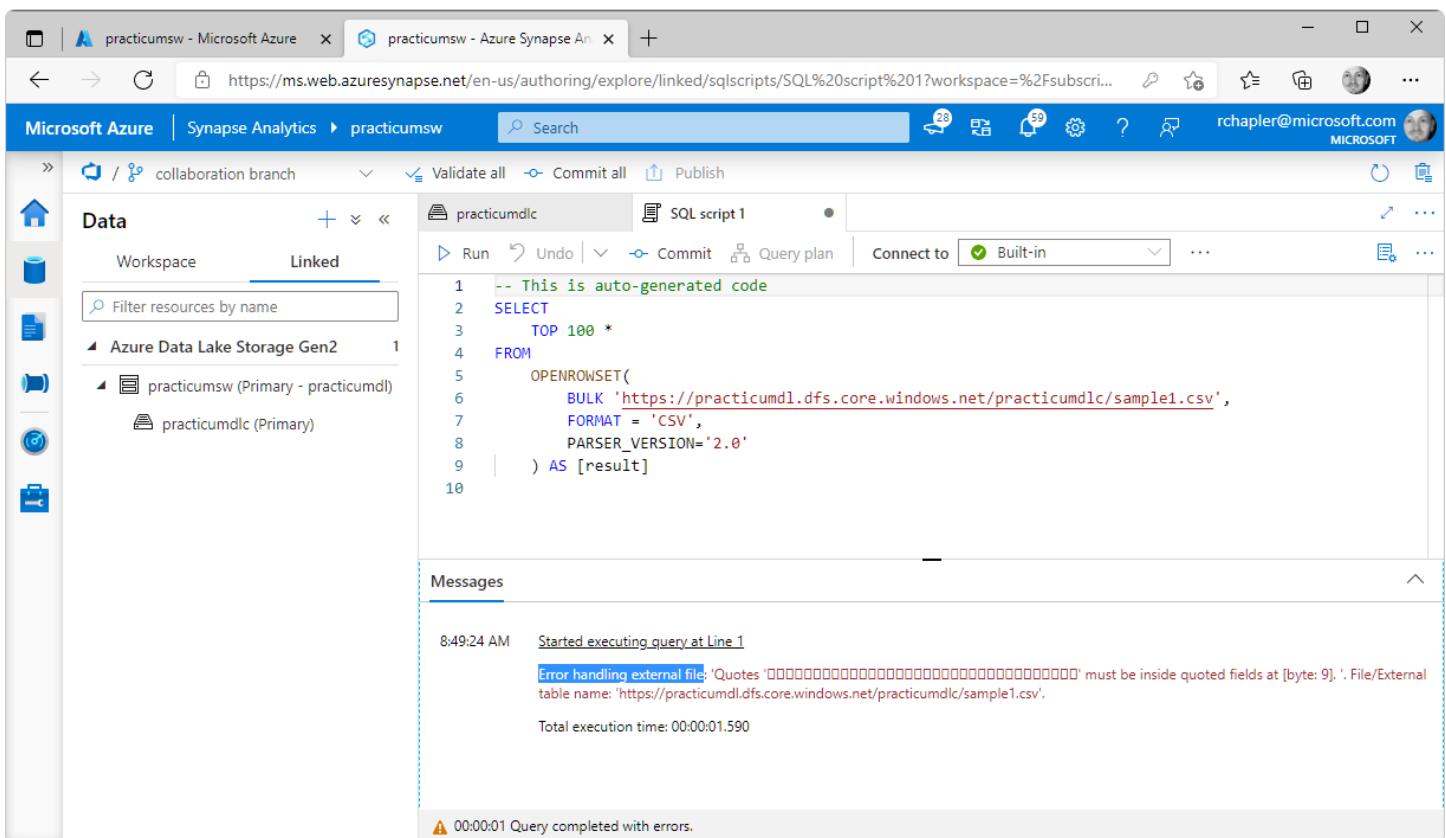
You will see that the Azure Data Lake Storage Gen2 specified during Synapse instantiation {i.e., practicumd1c} is surfaced with no additional configuration.

Method #1: SQL Script (OPENROWSET)



The screenshot shows the Microsoft Azure Synapse Analytics Data blade. On the left, there's a navigation pane with icons for Home, Data, Workspace, and Linked. Under 'Linked', it shows 'practicumdlc' with three CSV files: 'sample1.csv', 'sample2.csv', and 'sample3.csv'. The 'sample1.csv' file is selected. At the top, there are buttons for 'New SQL script', 'New notebook', 'New data flow', 'New integration dataset', 'Upload', and 'More'. The main area displays the contents of 'sample1.csv'.

Click the “New SQL script” button and “SELECT TOP 100 rows” from the resulting drop-down.



The screenshot shows the Microsoft Azure Synapse Analytics SQL script editor. The 'practicumdlc' database is selected. In the center, there's a code editor window titled 'SQL script 1' containing the following auto-generated code:

```
1 -- This is auto-generated code
2 SELECT
3     TOP 100 *
4 FROM
5     OPENROWSET(
6         BULK 'https://practicumdl.dfs.core.windows.net/practicumdlc/sample1.csv',
7         FORMAT = 'CSV',
8         PARSE_DATE='2021-03-02T12:11:09.000Z'
9     ) AS [result]
10
```

Below the code editor is a 'Messages' panel. It shows a log entry for line 1: "Started executing query at Line 1". It also displays an error message: "Error handling external file: 'Quotes '\u0022\u0022 must be inside quoted fields at [byte: 9]'. File/External table name: 'https://practicumdl.dfs.core.windows.net/practicumdlc/sample1.csv'." A note below the error says "Total execution time: 00:00:01.590". At the bottom of the messages panel, a warning icon indicates "00:00:01 Query completed with errors."

The resulting auto-generated code uses OPENROWSET(...) and a built-in, serverless SQL Pool to pull data directly from the CSV file in the ADLS container.

When we run the code (assuming you are using the same sample files, of course), we get an error about quote-handling.

Given the formatting of the sample file, we cannot run this without changes to the code.

Update the code to:

```

SELECT TOP 100 *
FROM OPENROWSET(
    BULK 'https://practicumdl.dfs.core.windows.net/practicumdlc/sample3.csv',
    FORMAT = 'CSV',
    PARSER_VERSION='2.0',
    HEADER_ROW = TRUE,
    FIELDQUOTE = '``'
) AS [result]

```

... and then, re-Run...

The screenshot shows the Microsoft Azure portal interface for a Synapse Analytics workspace named 'practicumsw'. On the left, there's a navigation sidebar with 'Data' selected, showing 'Workspace' and 'Linked' sections. The 'Linked' section lists 'Azure Data Lake Storage Gen2' and 'practicumsw (Primary - practicumdl)' which contains 'practicumdlc (Primary)'. The main area is titled 'practicumdlc' and contains a 'SQL script 1' tab with the following T-SQL code:

```

1 SELECT TOP 100 *
2 FROM OPENROWSET(
3     BULK 'https://practicumdl.dfs.core.windows.net/practicumdlc/sample3.csv',
4     FORMAT = 'CSV',
5     PARSER_VERSION='2.0',
6     HEADER_ROW = TRUE,
7     FIELDQUOTE = '``'
8 ) AS [result]

```

Below the code, the 'Results' tab is active, showing a table with two columns: "Game Number" and "Game Length". The data is as follows:

"Game Number"	"Game Length"
1	30
2	29
3	31

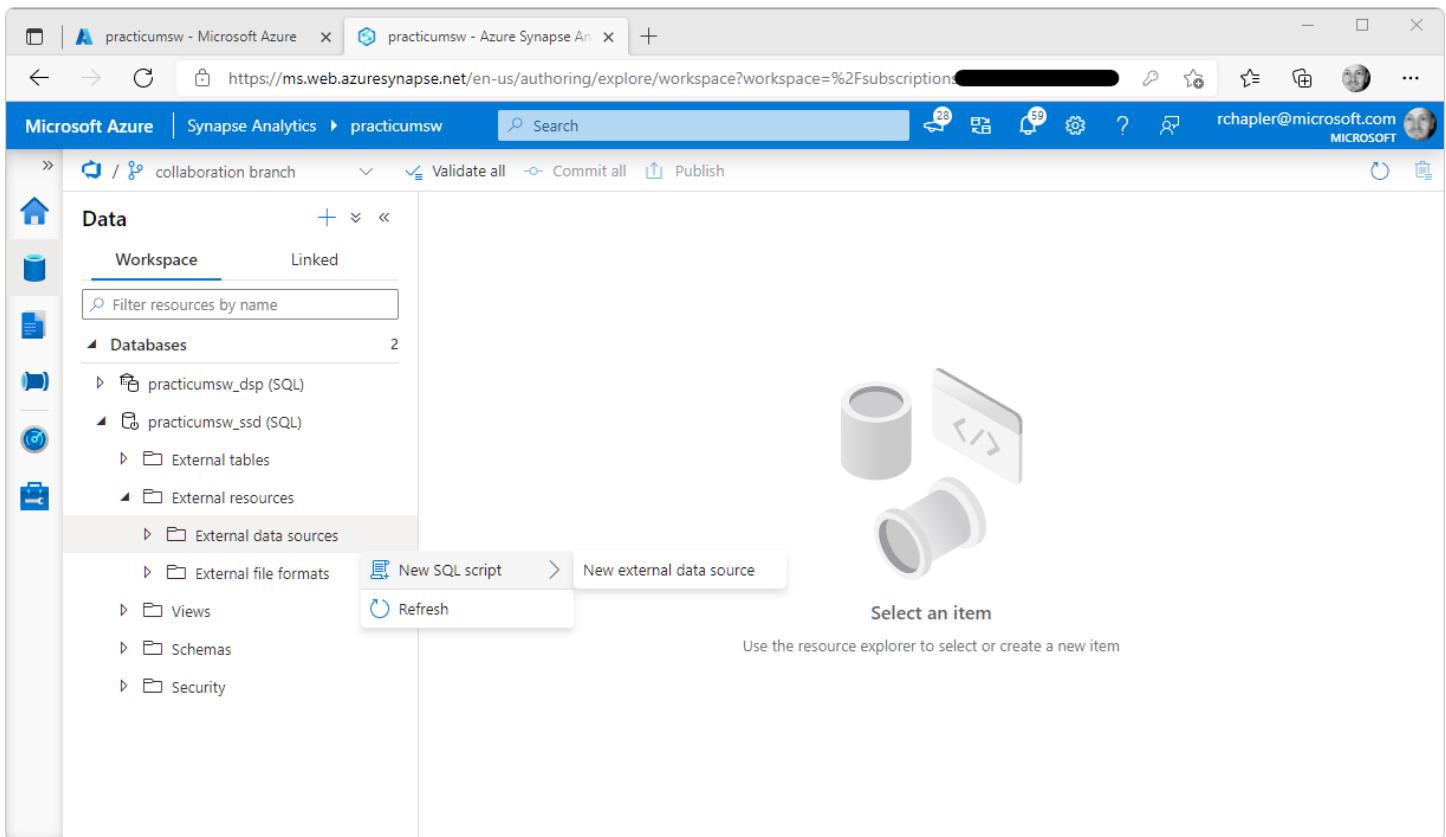
A message at the bottom indicates: **00:00:11 Query executed successfully.**

Method #2: External Table (Serverless SQL Database)

Click the **Workspace** tab.

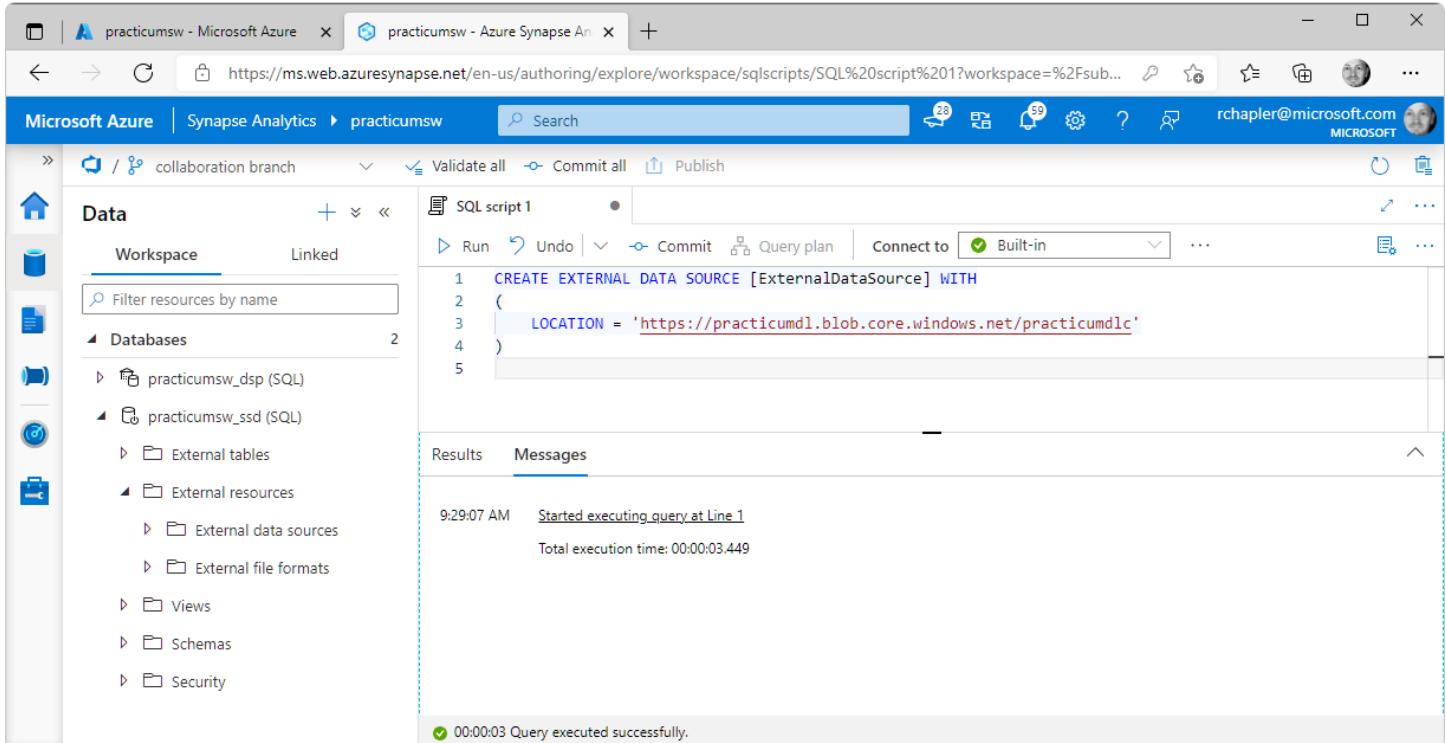
External Data Source

Expand navigation to and then right-click on **“External data sources”**.



The screenshot shows the Microsoft Azure Synapse Analytics workspace Data blade. In the left sidebar, under 'External resources', the 'External data sources' option is selected. A context menu is open at this location, with 'New SQL script' and 'New external data source' being the visible items.

Click “New SQL script”, and then “New external data source” in the drop-down.



The screenshot shows the Microsoft Azure Synapse Analytics workspace Data blade with the 'SQL script 1' editor open. The editor displays the following T-SQL code:

```
1 CREATE EXTERNAL DATA SOURCE [ExternalDataSource] WITH
2 (
3     LOCATION = 'https://practicumdl1.blob.core.windows.net/practicumdlc'
4 )
5
```

The 'Results' tab shows the execution log:

9:29:07 AM Started executing query at Line 1
Total execution time: 00:00:03.449

00:00:03 Query executed successfully.

Replace <STORAGEACCOUNT> and <CONTAINER> in the auto-generated code and then Run.

External File Format

Expand navigation to and then right-click on “External file formats”.

The screenshot shows the Microsoft Azure Synapse Analytics Data Explorer interface. The left sidebar is titled "Data" and contains sections for "Workspace" and "Linked". Under "Workspace", there are two databases: "practicumsw_dsp (SQL)" and "practicumsw_ssd (SQL)". Below these are "External tables", "External resources", "External data sources", and "External file formats". A context menu is open over the "External file formats" folder, with options including "New SQL script" (which is highlighted), "New external file format", and "Refresh". The main pane displays a 3D icon of two pipes connected by a valve, with the text "Select an item" and a placeholder message: "Use the resource explorer to select or create a new item".

Click “New SQL script”, and then “New external file format” in the drop-down.

The screenshot shows the Microsoft Azure Synapse Analytics Data Explorer interface. The left sidebar is titled "Data" and contains sections for "Workspace" and "Linked". Under "Workspace", there are two databases: "practicumsw_dsp (SQL)" and "practicumsw_ssd (SQL)". Below these are "External tables", "External resources", "External data sources", and "External file formats". The "External file formats" section is currently selected. In the center, a SQL script editor window titled "SQL script 1" is open, showing the following code:

```
1 CREATE EXTERNAL FILE FORMAT [ExternalFileFormat] WITH
2 (
3     FORMAT_TYPE = DELIMITEDTEXT, FORMAT_OPTIONS ( FIELD_TERMINATOR = ',', FIRST_ROW = 2 )
4 )
5
```

The "Run" button is highlighted. Below the editor, the "Results" tab shows the execution log:

```
10:24:53 AM Started executing query at Line 1
Total execution time: 00:00:02.705
```

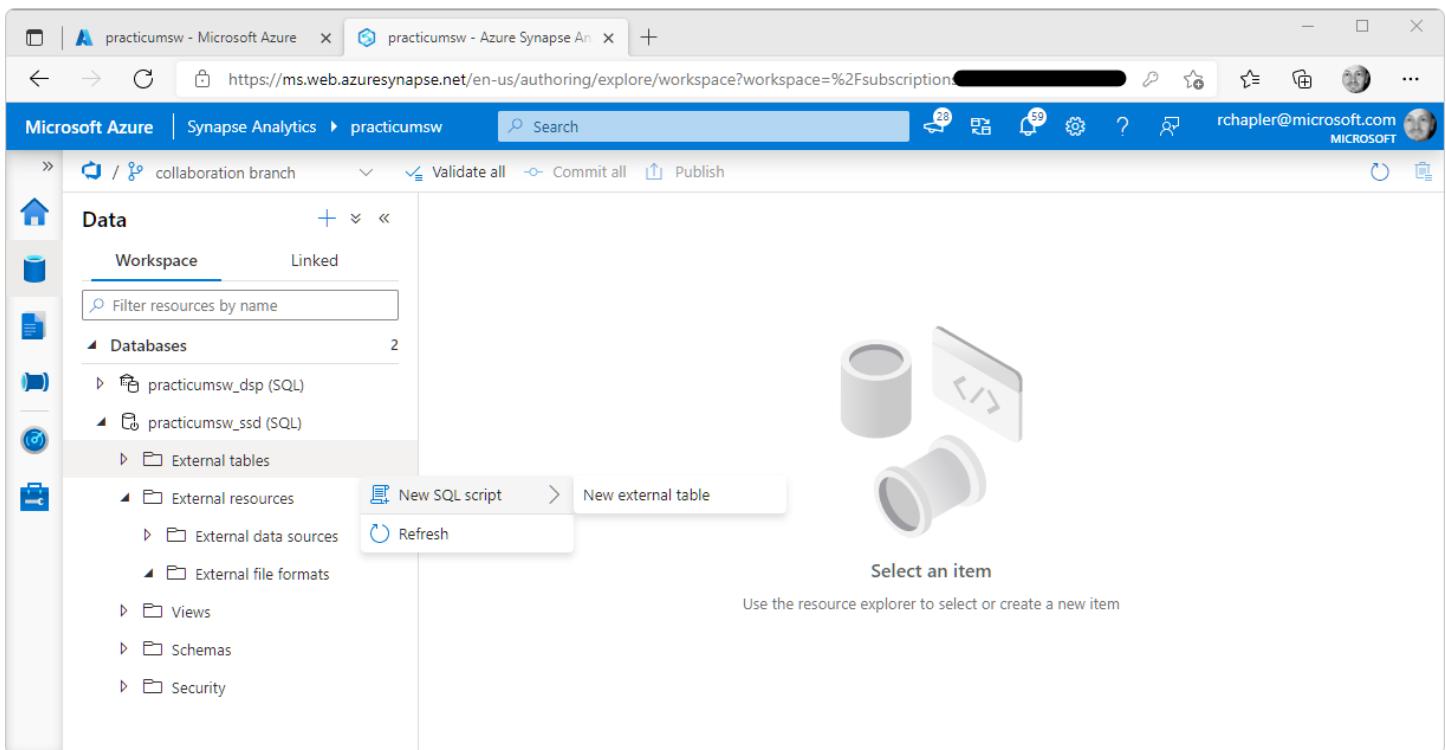
At the bottom, a success message is displayed:

```
00:00:02 Query executed successfully.
```

Append FORMAT_OPTIONS arguments FIELD_TERMINATOR and FIRST_ROW to the default FORMAT_TYPE argument and then Run.

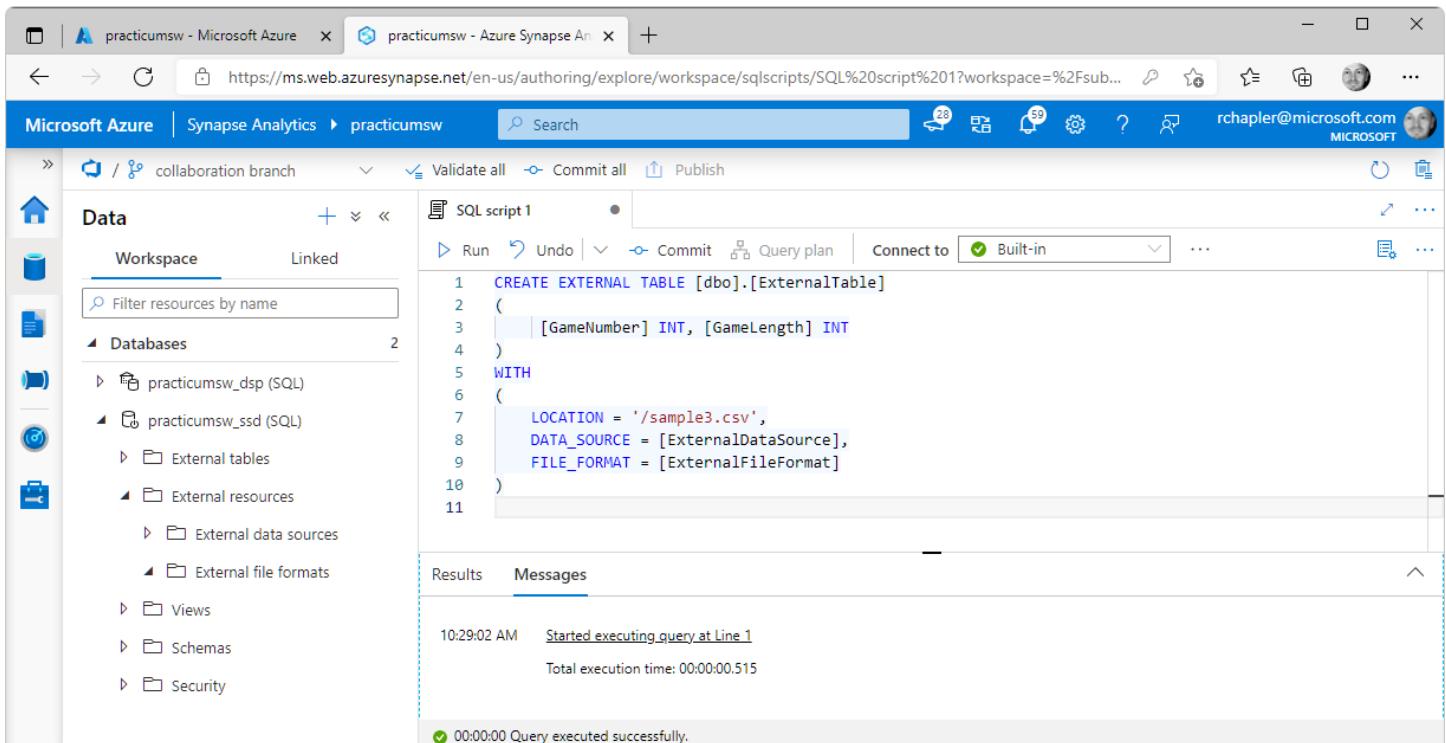
External Table

Expand navigation to and then right-click on “External tables”.



The screenshot shows the Microsoft Azure Synapse Analytics workspace interface. On the left, the 'Data' blade is open, showing a tree view of resources under 'Workspace'. The 'External tables' node is currently selected. A context menu is open at this node, with 'New SQL script' highlighted. Other options in the menu include 'New external table', 'Refresh', and 'Select an item'. The main pane displays a placeholder message: 'Select an item' and 'Use the resource explorer to select or create a new item'. The top navigation bar shows two tabs: 'practicumsw - Microsoft Azure' and 'practicumsw - Azure Synapse Analytics', along with a search bar and various status indicators.

Click “New SQL script”, and then “New external table” in the drop-down.



The screenshot shows the Microsoft Azure Synapse Analytics workspace interface with the 'SQL script' editor open. The editor contains a 'CREATE EXTERNAL TABLE' statement:

```
1 CREATE EXTERNAL TABLE [dbo].[ExternalTable]
2 (
3     [GameNumber] INT, [GameLength] INT
4 )
5 WITH
6 (
7     LOCATION = '/sample3.csv',
8     DATA_SOURCE = [ExternalDataSource],
9     FILE_FORMAT = [ExternalFileFormat]
10 )
11
```

The 'Messages' tab at the bottom of the editor shows the following output:

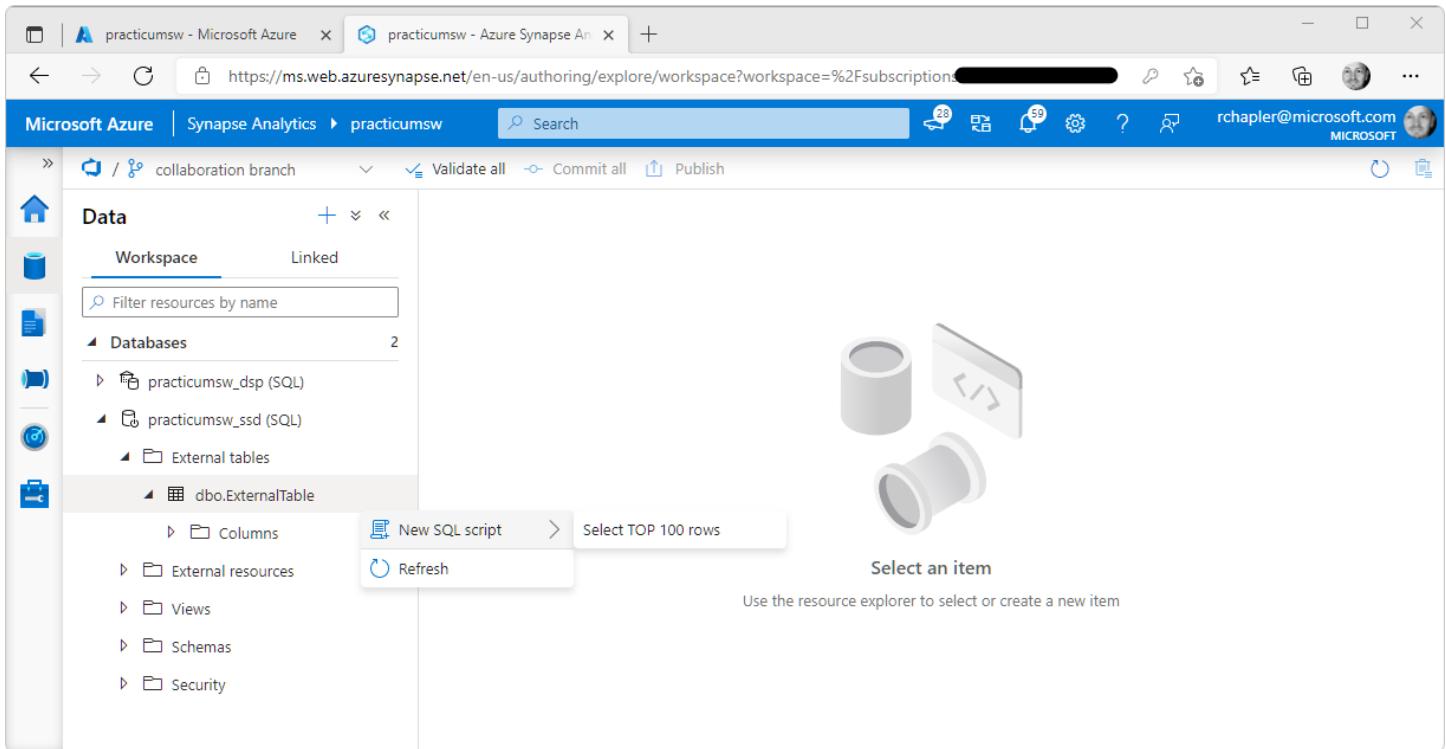
10:29:02 AM Started executing query at Line 1
Total execution time: 00:00:00.515
00:00:00 Query executed successfully.

Update the following items in the default code and then Run.

- **Schema Definition** ... replace `[Id] INT` with columns matching the external data source
- **LOCATION** ... replace `'/folder/file'` with values matching those in your container
- **DATA_SOURCE** ... replace `[DataSource1]` with the name used in [Create External Data Source](#)
- **FILE_FORMAT** ... replace `[FileFormat1]` with the name used in [Create External File Format](#)

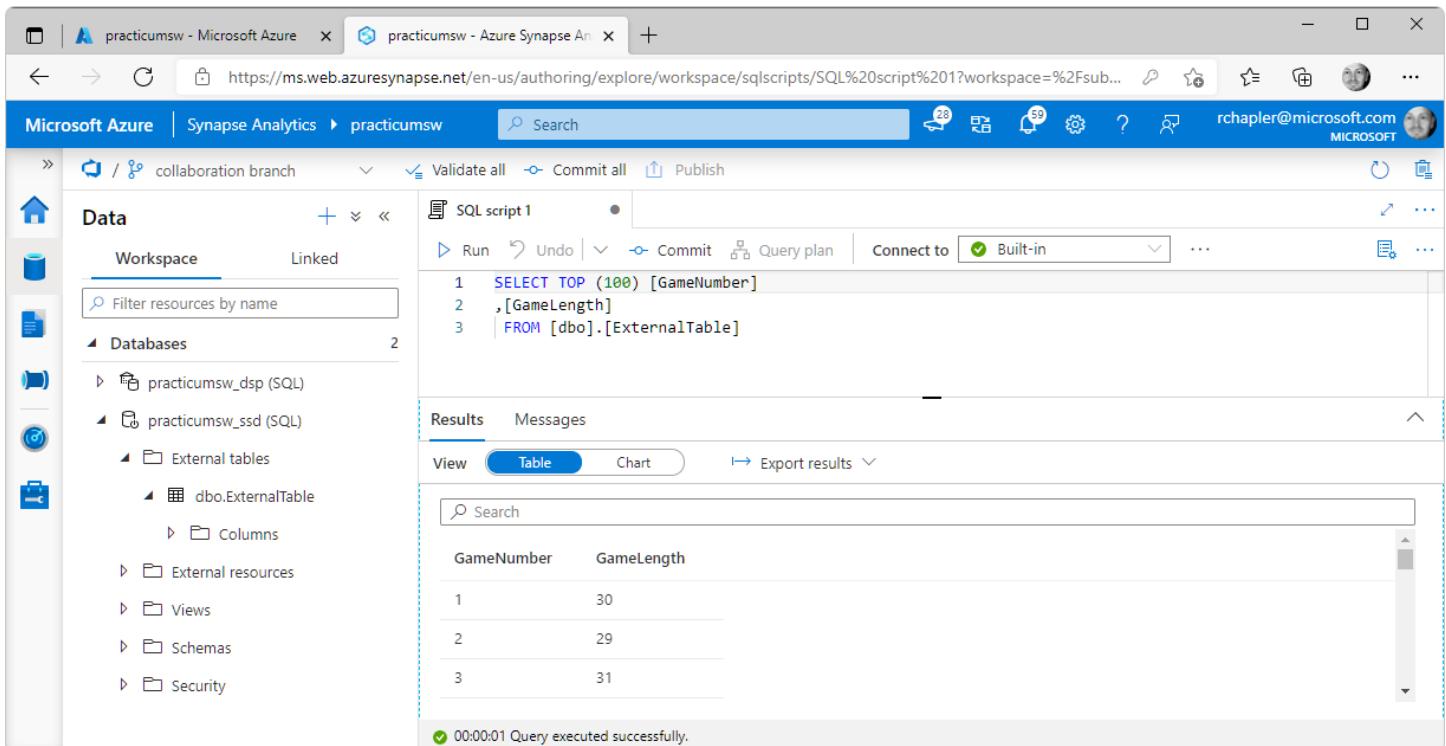
Confirm Success

Right-click on “External tables” and click Refresh.



The screenshot shows the Microsoft Azure Synapse Analytics Data blade. In the left sidebar, under the 'Data' category, the 'Workspace' tab is selected. Under 'External tables', there is a single item named 'dbo.ExternalTable'. A context menu is open over this item, with the 'Refresh' option highlighted. Other options in the menu include 'New SQL script' and 'Select TOP 100 rows'. The main pane displays a 3D icon of two pipes connected by a screen with code symbols.

Right-click on “dbo.ExternalTable”, click on “New SQL script”, and finally click on “Select TOP 100 rows” in the drop-down menu.



The screenshot shows the Microsoft Azure Synapse Analytics Data blade with a SQL script editor open. The script contains the following T-SQL code:

```
1 SELECT TOP (100) [GameNumber]
2 ,[GameLength]
3 FROM [dbo].[ExternalTable]
```

The 'Results' tab is selected, showing the output of the query in a table format. The data consists of two columns: 'GameNumber' and 'GameLength'. The results are:

GameNumber	GameLength
1	30
2	29
3	31

A message at the bottom of the results pane states: '00:00:01 Query executed successfully.'

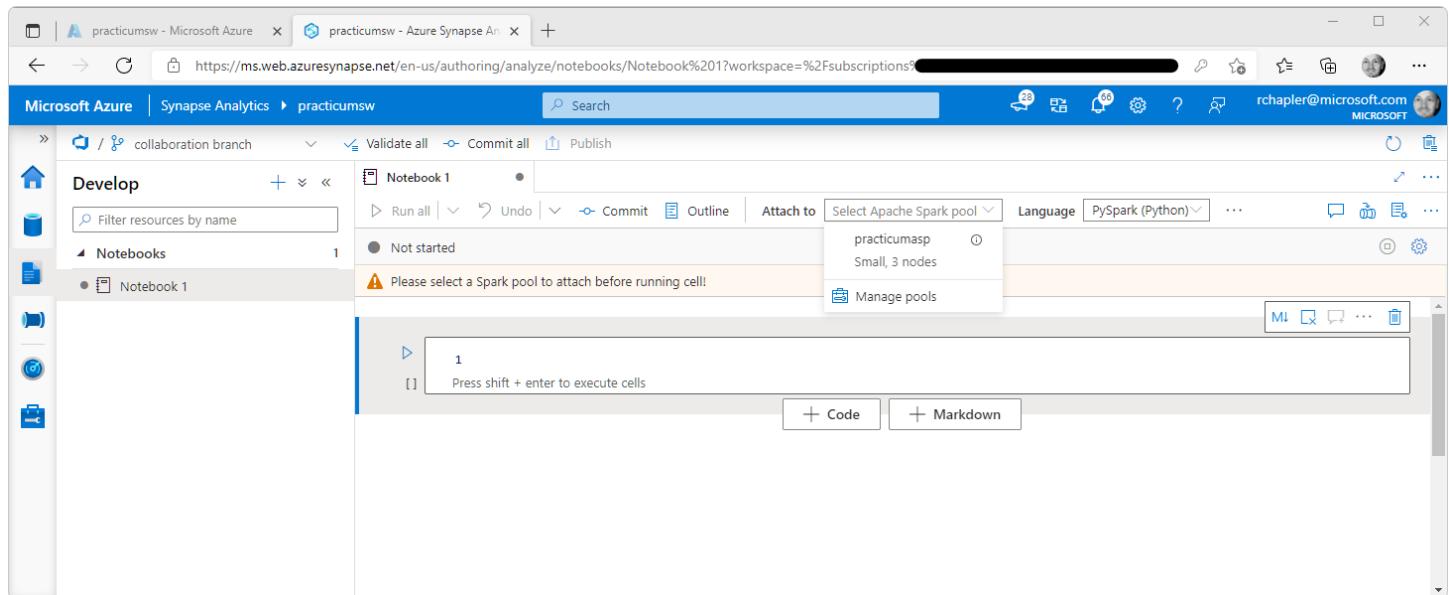
Click Run.

Method #3: Spark Notebook (Python)

Click the **Develop** tab.

Click the + icon just above and to the right of the “**Filter resources by name**” input.

Select **Notebook** from the resulting drop-down.



Select your Apache Spark Pool from the “**Attach to**” drop-down.

Paste the following code to Cell 1 and then click “**Run Cell**”:

```
%pyspark
theData = spark.read.load('abfss://practicumdlc@practicumdl.dfs.core.windows.net/sample1.csv', format='csv',
header=True)
theData.show(10)

print('Converted to Pandas...')
print(theData.toPandas())
```

practicumsw - Microsoft Azure practicumsw - Azure Synapse Analytics

Microsoft Azure | Synapse Analytics > practicumsw

Search

Develop

collaboration branch

Notebook 1

Validate all Commit all Publish

Ready

```
1 %%pyspark
2 theData = spark.read.load('abfss://practicumdlc@practicumdl.dfs.core.windows.net/sample1.csv', format='csv', header=True)
3 theData.show(10)
4
5 print('Converted to Pandas...')
6 print(theData.toPandas())
7
```

[1] ✓ 2 min 29 sec - Apache Spark session started in 2 min 6 sec 348 ms. Command executed in 23 sec 297 ms by rchaper on 1:21:45 PM, 9/23/21

> Job execution Succeeded Spark 2 executors 8 cores

View in monitoring Open Spark UI

Month	Average	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
May	0.1	0	0	1	1	0	0	0	2	0	0	0
Jun	0.5	2	1	1	0	0	1	1	2	2	0	1
Jul	0.7	5	1	1	2	0	1	3	0	2	2	1
Aug	2.3	6	3	2	4	4	4	7	8	2	2	3
Sep	3.5	6	4	7	4	2	8	5	2	5	2	5
Oct	2.0	8	0	1	3	2	5	1	5	2	3	0
Nov	0.5	3	0	0	1	1	0	1	0	1	0	1
Dec	0.0	1	0	1	0	0	0	0	0	0	0	1

Converted to Pandas...

Month	Average	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
0 May	0.1	0	0	...	2	0	0	0	2	0	0	0
1 Jun	0.5	2	1	...	2	2	0	1	2	2	0	1
2 Jul	0.7	5	1	...	0	2	2	1	2	2	1	1
3 Aug	2.3	6	3	...	8	2	2	3	2	2	2	3
4 Sep	3.5	6	4	...	2	5	2	5	2	5	2	5
5 Oct	2.0	8	0	...	5	2	3	0	0	0	0	0
6 Nov	0.5	3	0	...	0	1	0	1	0	1	0	1
7 Dec	0.0	1	0	...	0	0	0	0	0	0	0	1

[8 rows x 13 columns]

+ Code + Markdown



Good Job!

You have successfully completed the objective

Objective: Synapse + Data Lake | Simple Data Ingestion

Note: These instructions also apply (with minor differences) to Azure Data Factory.

Follow the instructions in this section to **load data from Data Lake to Synapse**.

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Data Lake (with container and sample data)
- Synapse (with Dedicated SQL Pool)

Navigate to your Synapse Analytics workspace and then click the **Open** link on the “**Open Synapse Studio**” rectangle.

[practicumsw - Microsoft Azure](#) [practicumsw - Azure Synapse An](#)

<https://ms.web.azuresynthesize.net/en-us/home?workspace=%2Fsubscriptions%2F...>

Microsoft Azure | Synapse Analytics > practicumsw

practicumsw

New

Ingest
Perform a one-time or scheduled data load.

Explore and analyze
Learn how to get insights from your data.

Visualize
Build interactive reports with Power BI capabilities.

Click the **Ingest** button.

Properties

Screenshot of the Microsoft Azure Copy Data tool wizard. The left sidebar shows steps 1 through 5: Properties, Source, Target, Settings, and Review and finish. Step 1 is selected and highlighted with a blue circle.

Properties

Use Copy Data Tool to perform a one-time or scheduled data load from 90+ data sources. Follow the wizard experience to specify your data loading settings, and let the Copy Data Tool generate the artifacts for you, including pipelines, datasets, and linked services. [Learn more](#)

Properties

Select copy data task type and configure task schedule

Task type

Built-in copy task
You will get single pipeline to copy data from 90+ data source easily.

Metadata-driven copy task (Preview)
Metadata is required to be stored in external control tables to load data at large-scale.

You will get single pipeline to quickly copy objects from data source store to destination in a very intuitive manner.

Task cadence or task schedule *

Run once now Schedule Tumbling window

[< Previous](#) [Next >](#) [Cancel](#)

Confirm default select, “Built-in copy task” and “Run once now”, then click the “**Next >**” button.

Source > Dataset

The screenshot shows the 'Copy Data tool' interface in the Microsoft Azure portal. The left sidebar lists steps: Properties (selected), Source, Dataset, Configuration, Target, Settings, and Review and finish. The main panel is titled 'Source data store' and contains the following fields:

- Source type:** Azure Data Lake Storage Gen2
- Connection ***: practicuml (with Edit and New connection buttons)
- Integration runtime ***: AutoResolveIntegrationRuntime (with Edit button)
- File or folder:** practicumlcl/sample1.csv (with Browse button)
- Options:** Binary copy (unchecked), Recursively (checked), Enable partition discovery (unchecked)
- Max concurrent connections:** (empty input field)
- Filter by last modified:** Start time (UTC) and End time (UTC) input fields

At the bottom are 'Previous' and 'Next >' buttons, and a 'Cancel' button.

On the “Source data store” page, enter values for the following items:

Source Type Select “Azure Data Lake Storage Gen2”

Connection Click the “+ New connection” button

The screenshot shows the Microsoft Azure Synapse Analytics Copy Data tool interface. On the left, a navigation pane lists steps: Properties (selected), Source, Dataset, Configuration, Target, Settings, and Review and finish. The main area shows the 'Source' step details. Under 'Source data store', it says 'Specify the source data store'. 'Source type' is set to 'File or folder'. 'Connection *' is 'AutoResolveIntegrationRuntime'. 'Integration runtime *' is 'AutoResolveIntegrationRuntime'. 'Authentication method' is 'Managed Identity'. 'Account selection method' is 'From Azure subscription'. 'Azure subscription' is 'rchapler'. 'Storage account name *' is 'practicumdl'. 'Max concurrent connections' is 100. 'Filter by last modified' and 'Start time (UTC)' fields are empty. Below these are 'Options': 'Binary copy' (unchecked), 'Recursively' (checked), and 'Enable partition discovery' (unchecked). 'Annotations' section has a '+ New' button. 'Parameters' and 'Advanced' sections are collapsed. At the bottom are 'Previous' and 'Next' buttons, a 'Commit' button, and a 'Cancel' button. A success message 'Connection successful' with a green checkmark is shown, along with a 'Test connection' link.

On the “New connection (Azure Data Lake Storage Gen2)” page, enter values for the following items:

Name	Enter a value that is meaningful and aligned with standards
Connect via...	Confirm default selection, AutoResolveIntegrationRuntime
Account Selection Method	Confirm default selection, “ From Azure subscription ”
Azure Subscription	Select your subscription
Storage Account Name	Select your instance of Data Lake

Click “Test connection” to confirm successful connection and then click the **Create** (or **Commit**) button.

Back on the “Source data store” page, click **Browse** to select container/file {e.g., “practicumdlc/sample1.csv”}.

No additional settings are required for this exercise.

Click the “**Next >**” button.

Source > Configuration

The screenshot shows the 'Copy Data tool' interface in Microsoft Azure Synapse Analytics. The left sidebar lists five steps: 'Properties' (selected), 'Source', 'Dataset', 'Configuration', 'Target', 'Settings', and 'Review and finish'. The main panel is titled 'File format settings' and contains the following configuration:

- File format:** Text format (dropdown menu)
- Column delimiter:** Comma (,) (dropdown menu)
- Row delimiter:** Line feed (\n) (dropdown menu)
- First row as header:** Checked (checkbox)
- Advanced:** Sub-section with 'Compression type' set to 'None' (dropdown menu) and an 'Additional columns' section with a '+ New' button.

At the bottom are 'Previous' and 'Next >' buttons, and a 'Cancel' button on the right.

On the “**File format settings**” page, confirm default configuration and the resulting data preview.

Make changes if required {e.g., in this example, I check “**First row as header**” to correct handling of incoming data}.

Click the “**Next >**” button.

Target > Dataset

The screenshot shows the 'Copy Data tool' interface in the Microsoft Azure portal. The left sidebar lists steps: Properties (checked), Source (checked), Target (selected), Dataset, Configuration, Settings, and Review and finish. The main panel is titled 'Destination data store' with the sub-instruction: 'Specify the destination data store for the copy task. You can use an existing data store connection or specify a new data store.' It includes a 'Target type' dropdown set to 'Azure Synapse Analytics' and a 'Connection *' dropdown with a 'Select...' option and a '+ New connection' button. Navigation buttons at the bottom include '< Previous', 'Next >', and 'Cancel'.

On the “**Destination data store**” page, enter values for the following items:

Source Type Select “**Azure Synapse Analytics**”

Connection Click the “**+ New connection**” button

practicumsw - Microsoft Azure x practicumsw - Azure Synapse An x + <https://ms.web.azuresynthesize.net/en-us/authoring/orchestrate/pipeline/Pipeline%201?...> 28 4 ⚡ 🌐 ? 🔍 🗃 🏷️ 🎯 🏹 rchapler@microsoft.com MICROSOFT

Microsoft Azure | Synapse Analytics > practicum Search

Copy Data tool

Properties
Source
Target
Dataset
Configuration
Settings
Review and finish

Destination data store
Specify the destination data store.
Target type
Connection *

New connection (Azure Synapse Analytics)

Choose a name for your linked service. This name cannot be updated later.

Name *

Description

Connect via integration runtime * [Edit](#)

Connection string [Azure Key Vault](#)

Account selection method [?](#)
 From Azure subscription Enter manually

Azure subscription

Server name *

Database name *

SQL pool * [Edit](#)

Authentication type *

User name *

Add dynamic content [Alt+Shift+D]

Password [Azure Key Vault](#)

AKV linked service * [?](#) [Edit](#)

Secret name *

Secret version [?](#)

Additional connection properties [New](#)

Annotations [New](#)

Parameters [Advanced](#)

Previous Next Commit Cancel [Test connection](#)

On the “New connection (Azure Synapse Analytics)” pop-out, enter values for the following items:

Name	Enter a value that is meaningful and aligned with standards
Connect via...	Confirm default selection, AutoResolveIntegrationRuntime
Account Selection Method	Confirm default selection, “From Azure subscription”
Azure Subscription	Select your subscription
Server Name	Select your instance of Synapse Workspace
Database Name	Select your Dedicated SQL Pool
Authentication Type	Confirm default value, “SQL authentication”
User Name	Enter the “SQL Server admin login” value used during instantiation of Synapse
AKV Linked Service	Select the name of the Linked Service created for the Key Vault
Secret Name	Enter the Secret Name used to capture the Synapse Workspace administrator password

Click “Test connection” to confirm successful connection and then click the **Create** (or **Commit**) button.

The screenshot shows the Microsoft Azure Copy Data tool interface. On the left, a vertical navigation pane lists steps: Properties (checked), Source (checked), Target (selected), Dataset, Configuration, Settings, and Review and finish. The main panel is titled "Destination data store". It asks to specify the destination data store, using an existing connection or creating a new one. The "Target type" is set to "Azure Synapse Analytics", the "Connection" is "practicumsa", and the "Integration runtime" is "AutoResolveIntegrationRuntime". Below this, a table maps a source dataset ("Azure Data Lake Storage Gen2 file") to a target table ("dbo.Sample1"). A checkbox for "Skip column mapping for all tables" is present. At the bottom are "Previous", "Next >", and "Cancel" buttons.

Back on the “Destination data source” page, enter values for target schema name {e.g., “dbo”} and target table name {e.g., “Sample1”}.

Click the “Next >” button.

Target > Configuration

The screenshot shows the 'Copy Data tool' configuration page in the Microsoft Azure portal. The left sidebar lists steps: Properties, Source, Target (selected), Dataset, Configuration, Settings, and Review and finish. The main area is titled 'Column mapping' and shows 'Table mappings (1)'. It displays a list of columns from a source Azure Data Lake Storage Gen2 file named 'dbo.Sample1' and maps them to a destination table. The 'Source' column contains values like 'Month', '\"Average\"', '\"2005\"', '\"2006\"', '\"2007\"', '\"2008\"', '\"2009\"', '\"2010\"', '\"2011\"', '\"2012\"', '\"2013\"', '\"2014\"', and '\"2015\"'. The 'Type' column for all is 'String'. The 'Destination' column shows the mapped values. The 'Type' column for the destination is also 'String'. Below this, there's a section for 'Azure Synapse Analytics sink properties' with a 'Pre-copy script' input field and an 'Advanced' link. At the bottom are 'Previous' and 'Next >' buttons, and a 'Cancel' button.

Review default column mappings; correct if required.

Click the “**Next >**” button.

Settings

The screenshot shows the 'Copy Data tool' settings page in the Microsoft Azure portal. The left sidebar lists steps: Properties (checked), Source (checked), Target (checked), Settings (checked), and Review and finish (unchecked). The main area is titled 'Settings' and contains the following fields:

- Task name ***: CopyPipeline_8th
- Task description**: (empty text area)
- Data consistency verification**: (radio button)
- Fault tolerance**: (dropdown menu)
- Enable logging**: (checkbox)
- Enable staging**: (checkbox)
- Advanced** section:
 - Copy method**: Bulk insert (selected radio button)
 - Bulk insert table lock**: No (radio button)
- Data integration unit**: Auto (dropdown menu)
- Degree of copy parallelism**: (dropdown menu)

At the bottom are buttons for < Previous, Next >, and Cancel.

On the **Settings** page, enter values for the following items:

Enable Staging	Unchecked
Copy Method	Bulk insert

Click the “**Next >**” button.

Summary

Screenshot of the Microsoft Azure Copy Data tool interface showing a pipeline configuration for copying data from Azure Data Lake Storage Gen2 to Azure Synapse Analytics.

The pipeline steps completed so far are:

- Properties
- Source
- Target
- Settings

The current step is:

- Review and finish

Summary

You are running pipeline to copy data from Azure Data Lake Storage Gen2 to Azure Synapse Analytics.

 Azure Data Lake Storage Gen2 →  Azure Synapse Analytics

Properties

Task name: CopyPipeline_8th

Task description:

Source

Connection name: practicuml
Dataset name: SourceDataset_8th
Column delimiter:
Row delimiter:
Escape character:
Quote char:
First row as header: true
File name: sample1.csv

Target

Connection name: practicumsa
Dataset name: DestinationDataset_8th
Table name: dbo.Sample1

Copy settings

Timeout: 7:00:00:00
Retry: 0
Retry interval: 30
Secure output: false
Secure input: false

Buttons at the bottom:

- < Previous
- Next >**
- Cancel

Review configuration and then click the “**Next >**” button.

Deployment

The screenshot shows the Microsoft Azure Copy Data tool interface. On the left, a vertical navigation bar lists steps: Properties, Source, Target, Settings, Review and finish, Review, and Deployment. The first four steps have green checkmarks. The 'Review and finish' step is currently selected. In the center, there's a diagram showing an arrow pointing from 'Azure Data Lake Storage Gen2' to 'Azure Synapse Analytics'. Below the diagram, the text 'Deployment complete' is displayed. To the right of the diagram, a table titled 'Deployment step' shows two rows: 'Creating datasets' (Status: Succeeded) and 'Creating pipelines' (Status: Succeeded). At the bottom of the screen, there are three buttons: 'Finish' (highlighted in blue), 'Edit pipeline', and 'Monitor'.

Monitor progress.

Click the **Finish** button on success.

Confirm Success

“**CopyPipeline...**” will be added to the list of **Pipelines** in the Synapse, **Integrate** section.

The screenshot shows the Microsoft Azure Synapse Analytics Integrate section. On the left, a sidebar shows a 'collaboration branch' and a list of 'Pipelines' with one item: 'CopyPipeline_8th'. The main area displays the 'CopyPipeline_8th' pipeline details. It shows an 'Activities' list with a single item: 'Copy data' (status: succeeded). Below the activities, there are tabs for 'Parameters', 'Variables', 'Settings', and 'Output'. The 'Output' tab shows a table with one row: 'Copy_8th' (Type: Copy data, Run start: 2021-09-22T17:38:30.320, Duration: 00:00:10, Status: Succeeded, Integration runtime: DefaultIntegrationRuntime). At the bottom right, there are links for 'View debug run consumption' and 'View pipeline run history'.

Click **Debug** and confirm successful execution.

Click the **Data** icon in the left-hand navigation.

The screenshot shows the Microsoft Azure Synapse Analytics workspace interface. On the left, the navigation pane is open with the 'Data' icon selected. Under 'Databases', the 'practicumsw_dsp (SQL)' database is expanded, and under 'Tables', the 'dbo.Sample1' table is selected. A context menu is open over the table, with 'New SQL script' and 'Select TOP 100 rows' highlighted. The main area displays a SQL script window with the following code:

```
1 SELECT TOP (100) [Month]
2 , [ "Average"]
3 , [ "2005"]
```

Below the script, the results are shown in a table format:

	"2006"	"2007"	"2008"	"2009"	"2010"	"2011"
0	1	1	0	0	0	0
1	1	0	0	1	1	1
2	1	1	2	0	1	3
3	2	4	4	4	4	7

A message at the bottom of the results pane says '00:00:00 Query executed successfully.'

Expand **Databases > {Dedicated SQL Pool} > Tables** and confirm that you can see your new table.

Right-click on your table, select “**New SQL script**”, and then “**Select TOP 100 rows**” in the drop-down menu.



Objective: Synapse + SQL + Data Explorer | Simple Data Pipeline

Note: These instructions also apply (with minor differences) to Azure Data Factory.

Follow the instructions in this section to **load data from Data Explorer to Synapse**.

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Data Explorer (with cluster, database, and sample Product table)
- SQL (with sample data)
- Synapse (with linked services)

Navigate to your Synapse Analytics workspace and then click the **Open** link on the “**Open Synapse Studio**” rectangle.

Linked Services

Select the **Manage** icon on the left-hand navigation and then “**Linked Services**” in the resulting navigation menu.

Confirm linked service instantiation for: 1) Key Vault, 2) SQL and 3) Data Explorer.

The screenshot shows the Microsoft Azure portal interface for a workspace named 'practicumsw'. The left sidebar navigation includes 'Analytics pools', 'SQL pools', 'Apache Spark pools', 'External connections' (with 'Linked services' selected), 'Azure Purview (Preview)', 'Integration' (with 'Triggers' and 'Integration runtimes'), 'Security', 'Access control', 'Credentials', 'Managed private endpoints', 'Code libraries', 'Workspace packages', 'Source control', and 'Git configuration'. The main content area is titled 'Linked services' and describes them as connection strings for Azure Synapse Analytics. It features a 'New' button, a 'Filter by name' input field, and an 'Annotations : Any' dropdown. A table lists five items:

Name	Type	Related	Annotations
practicumded	Azure Data Explorer (Kusto)	0	
practicumkv	Azure Key Vault	1	
practicumsd	Azure SQL Database	0	
practicumsw-WorkspaceDefaultSqlServer	Azure Synapse Analytics	0	
practicumsw-WorkspaceDefaultStorage	Azure Data Lake Storage Gen2	0	

Note: you are likely to see system-generated linked services {e.g., "practicumsw-WorkspaceDefault..."} in addition to instantiated items.

Pipeline

Click the **Integrate** icon in the left-hand navigation.

No items to show
Try creating a new item using the + button above. [Learn more](#)

Select an item
Use the resource explorer to select or create a new item

Click the + icon just above and to the right of the “**Filter resources by name**” input.

Select **Pipeline** from the resulting drop-down.

Activities

- Copy data
- Data flow

Source dataset *

Add a “**Copy data**” activity.

Source

On the **Source** tab, click the “**+ New**” button.

The screenshot shows the Microsoft Azure Synapse Analytics pipeline creation interface. On the left, the navigation pane is visible with 'Integrate' selected. Under 'Activities', 'Copy data' is highlighted. A search bar at the top right contains the text 'azure sql'. The main area displays two options under the 'Database' category: 'Azure SQL Database' and 'Azure SQL Database Managed Instance'. A 'Continue' button is at the bottom left, and a 'Cancel' button is at the bottom right.

On the “**New integration dataset**” pop-out, search for and then select “**Azure SQL Database**”.

Click the **Continue** button.

The screenshot shows the Microsoft Azure Synapse Analytics interface. On the left, there's a sidebar with icons for Home, Pipelines, Databricks, Data Lake Analytics, Batch Service, Functions, Data Explorer, Copy Data, Data Flow, Move & Transform, Synapse, and General. The main area shows a pipeline named 'Pipeline 1'. A 'Set properties' dialog box is open, titled 'Set properties'. It contains the following fields:

- Name:** AzureSqlTable1
- Linked service:** practicumsd
- Connect via integration runtime:** AutoResolveIntegrationRuntime
- Table name:** SalesLT.Product
- Import schema:** From connection/store (radio button selected)

At the bottom of the dialog are buttons for OK, Back, and Cancel.

On the “Set properties” pop-out, enter values for the following items:

Name	Enter a value that is meaningful and aligned with standards
Linked Service	Select your Azure SQL linked service
Table Name	Select “SalesLT.Product”
Import Schema	Confirm default selection, “From connection/store”

Leave all other settings with default values.

Click the **OK** button.

The screenshot shows the Microsoft Azure Synapse Analytics pipeline editor interface. On the left, there's a sidebar with icons for Home, Pipelines, Databricks, Data Lake Analytics, General, HDInsight, Iteration & conditionals, and Machine Learning. The main area shows a pipeline named "Pipeline 1" with one activity: "Copy data1". The "Source" tab is selected, showing the source dataset as "AzureSqlTable1". Other tabs include "Sink", "Mapping", "Settings", and "User properties". At the bottom, there's a note: "Please preview data to validate the partition settings are correct before you trigger a run or publish the pipeline." There are also "Additional columns" and a "+ New" button.

Sink

Click on the **Sink** tab and then click the “+ New” button.

The screenshot shows the Microsoft Azure Synapse Analytics pipeline creation interface. On the left, there's a sidebar with icons for Home, Pipelines, Databricks, Data Lake Analytics, and General. The main area shows a pipeline named 'Pipeline 1'. A search bar at the top right has 'Search' and a user profile 'rchapler@microsoft.com MICROSOFT'. Below the search bar, a button labeled 'New integration dataset' is visible. The central part of the screen displays a 'Select a data store' dialog with a search bar containing 'data explorer'. Below it, a list of categories includes 'All', 'Azure', 'Database', 'File', 'Generic protocol', 'NoSQL', and 'Services and apps'. Under 'Database', 'Azure Data Explorer (Kusto)' is highlighted with a blue border and icon. At the bottom of the dialog are 'Continue' and 'Cancel' buttons.

On the “**New integration dataset**” pop-out, search for and then select “**Azure Data Explorer (Kusto)**”.

Click the **Continue** button.

The screenshot shows the 'Set properties' dialog for an integration dataset. The left sidebar is identical to the previous screenshot. The main area is titled 'Set properties' and contains the following fields:

- Name:** AzureDataExplorerTable1
- Linked service ***: practicumed
- Connect via integration runtime ***: AutoResolveIntegrationRuntime
- Table:** None
- Edit
- Advanced** (button)

At the bottom are 'OK', 'Back', and 'Cancel' buttons.

On the “**Set properties**” pop-out, enter values for the following items:

Name	Enter a value that is meaningful and aligned with standards
Linked Service	Select your Azure Data Explorer linked service
Table Name	Select Product

Leave all other settings with default values.

Click the **OK** button.

The screenshot shows the Microsoft Azure Synapse Analytics pipeline editor interface. On the left, the navigation pane is visible with sections like 'Integrate', 'Pipelines', and 'Pipeline 1'. The main workspace displays a pipeline named 'Pipeline 1' containing a single activity: 'Copy data'. This activity is highlighted with a red circle. The 'Sink' tab is selected in the activity configuration pane. Under 'Sink dataset', 'AzureDataExplorerTable1' is chosen from a dropdown. Other fields include 'Database' set to 'practicumded' and 'Ingestion mapping name' left empty. The top bar shows standard browser controls and the URL <https://ms.web.azuresynthesize.net/en-us/authoring/orchestrate/pipeline/Pipeline%201?workspace=%2Fsub...>.

Mapping

Click on the **Mapping** tab and then click the “Import schemas” button.

The screenshot shows the Microsoft Azure Synapse Analytics pipeline editor. On the left, the navigation pane is open with 'Integrate' selected, showing 'Pipelines' and 'Pipeline 1'. The main area displays a 'Copy data' activity named 'Copy data1'. The 'Mapping' tab is active, showing the mapping between source columns (ProductNumber and ListPrice) and destination columns (ProductNumber and ListPrice). The source types are nvarchar and money, and the destination types are String and Decimal respectively. The mapping table includes columns for Source, Type, Destination, and Type.

Confirm Success

Click **Validate** to confirm that there are no errors, and then click **Debug** to confirm that the pipeline runs successfully.

The screenshot shows the Microsoft Azure Synapse Analytics pipeline editor. The 'Output' tab is selected for Pipeline 1. It displays a successful pipeline run with ID 77b27627-6862-4dea-87ba-da3a3e9586a5. Below the run ID, a table provides details about the run, including Name, Type, Run start, Duration, Status, and Integration runtime. The status is marked as 'Succeeded'.



Good Job!

You have successfully completed the objective

Objective: Synapse | Simple Conditional Pipeline

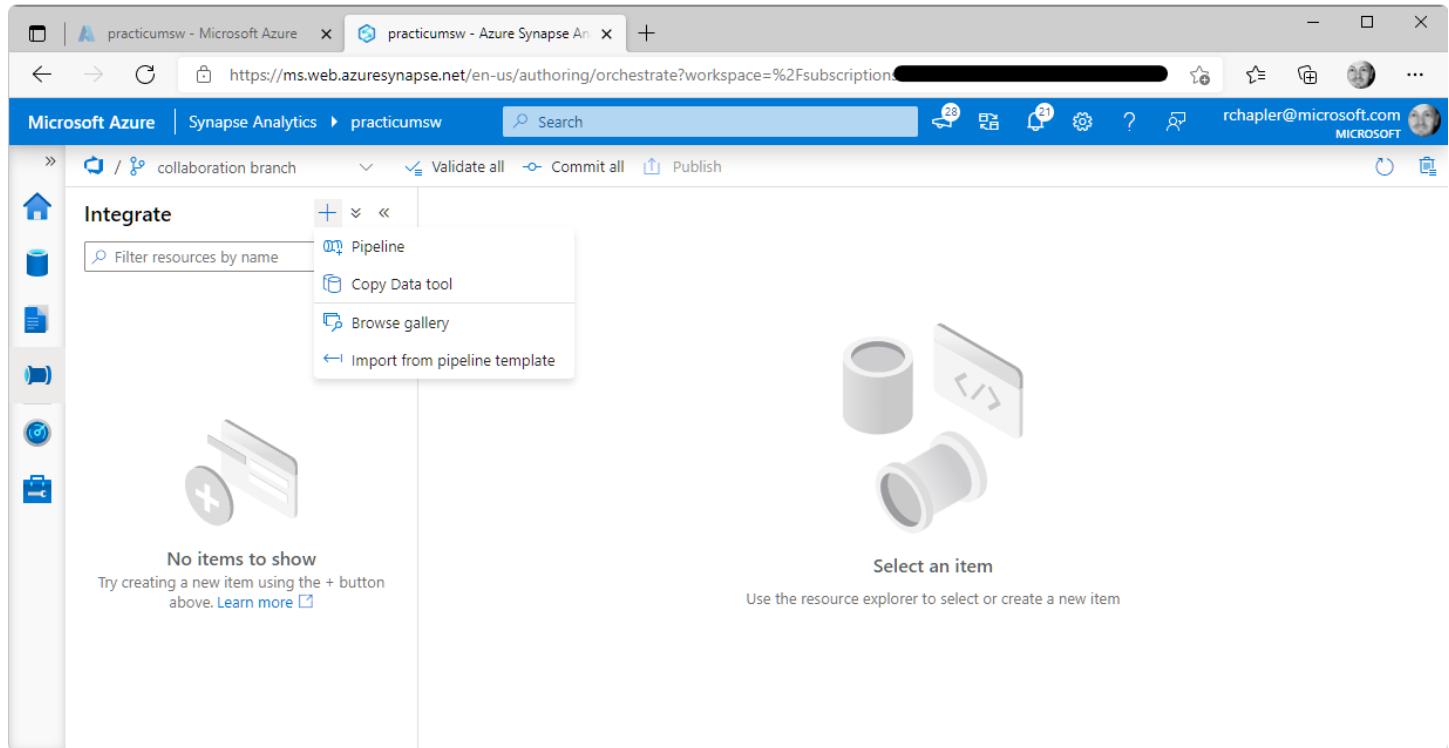
Note: These instructions also apply (with minor differences) to Azure Data Factory.

Follow the instructions in this section to **demonstrate conditional logic in a pipeline**.

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Synapse

Navigate to your Synapse Analytics workspace and then click the **Open** link on the “**Open Synapse Studio**” rectangle.



Pipeline

Click the **Integrate** icon in the left-hand navigation.

Click the **+** icon just above and to the right of the “**Filter resources by name**” input.

Select **Pipeline** from the resulting drop-down.

The screenshot shows the Microsoft Azure Synapse Analytics pipeline editor. On the left, the navigation pane shows 'practicumsw' workspace, 'collaboration branch', and 'Pipeline 1'. The main area displays the 'Activities' bar with 'General' selected. Below it, the 'Variables' tab is active, showing two variables: 'Flag' (Boolean) and 'Result' (String). The 'Flag' variable has a value of 'true'.

Name	Type	Default value
Flag	Boolean	true
Result	String	Value

Click on the **Variables** tab.

Click “+ New” to add each of the following variables:

Flag (Boolean)	Will serve as a TRUE / FALSE trigger for the conditional result
Result (String)	For capturing an anecdotal result {e.g., “Success!”}

Activity #1, Set Variable

Expand **General** in the **Activities** bar, and then drag-and-drop a “Set variable” component into the activity window.

The screenshot shows the Microsoft Azure Synapse Analytics pipeline editor. The 'Activities' bar now includes a 'Set variable' component under the 'General' category. A red circle highlights this component in the activity window. The 'Variables' tab is active, showing the 'Flag' variable with a value of 'true'.

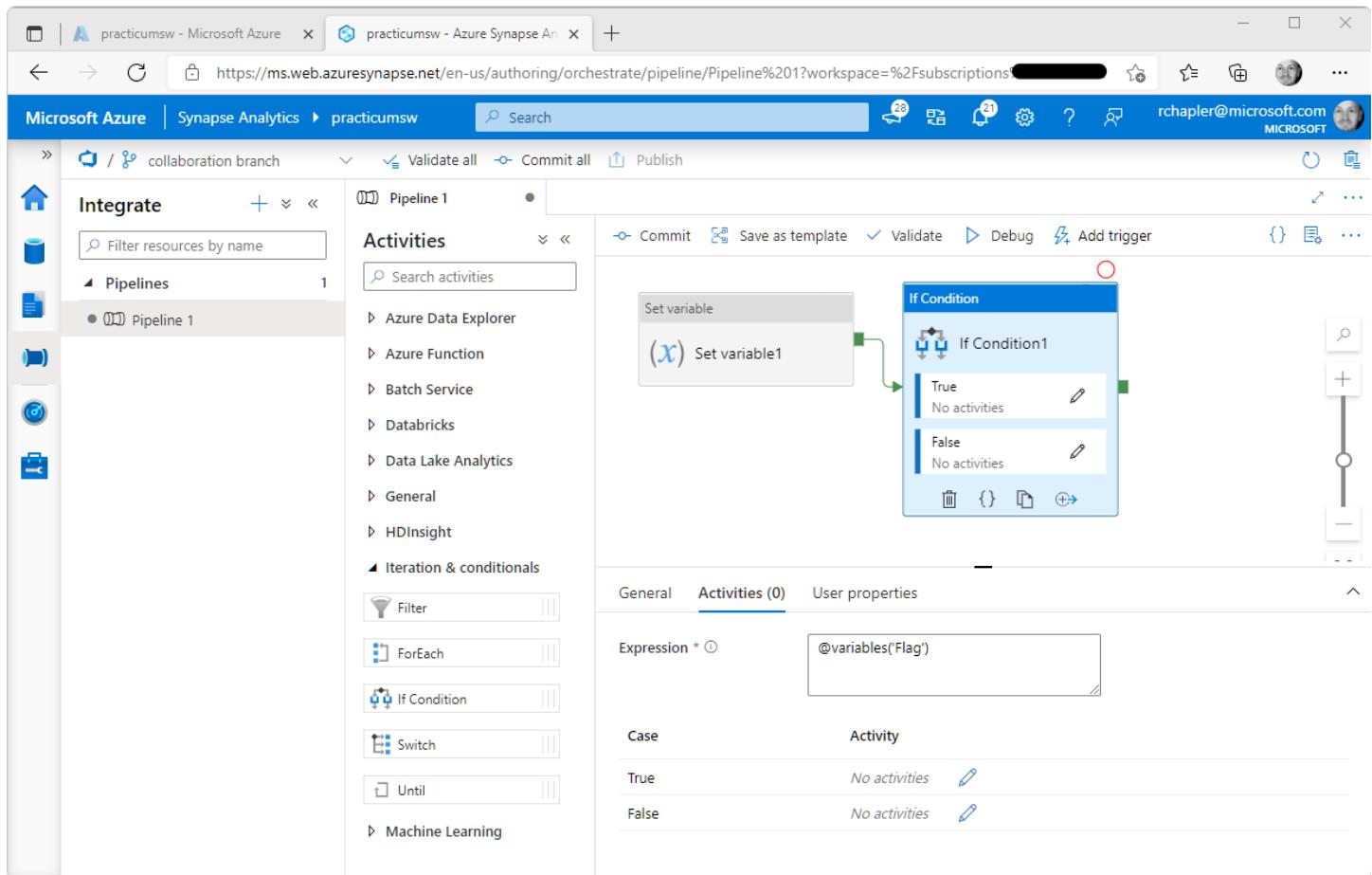
On the **Variables** tab of the “Set Variable...” component, enter values for the following items:

Name	Select Flag from the drop-down
Value	Enter true

Activity #2, If Condition

Expand “Iteration & conditionals” in the **Activities** bar.

Drag-and-drop an “If Condition” component into the pipeline.



Create a dependency between the “Set variable” and “If Condition” components (by drawing an arrow from one to the other).

Click on the “Activities (0)” tab.

Click into the **Expression** textbox, and then enter the following: `@variables('Flag')`

Click on the pencil icon in the “If Condition” component, **True** selection.

The screenshot shows the Microsoft Azure Synapse Analytics pipeline editor. On the left, the navigation pane shows 'Integrate' and 'Pipelines'. Under 'Pipelines', 'Pipeline 1' is selected. The main area displays the pipeline activities. A 'Set variable' activity is currently selected. The 'Variables' tab of the activity configuration pane is open, showing a 'Name' dropdown set to 'Result' and a 'Value' text input set to '"Success!"'. The pipeline structure above shows an 'If Condition' activity followed by a 'True activities' section containing the 'Set variable' activity.

Expand **General** in the **Activities** bar, and then drag-and-drop a “Set variable” component into the activity window.

On the **Variables** tab of the “Set Variable...” component, enter values for the following items:

Name	Select Result from the drop-down
Value	Enter “Success!”

Use cookie crumbs to navigate back to the main window.

Click on the pencil icon in the “If Condition” component, **False** selection.

The screenshot shows the Microsoft Azure Synapse Analytics pipeline editor. The pipeline structure has been modified. The 'If Condition' activity now has two branches: a 'True activities' section containing the previously configured 'Set variable' activity, and a 'False activities' section. In the 'False activities' section, there is a new 'Set variable' activity. The 'Variables' tab of this new activity is open, showing a 'Name' dropdown set to 'Result' and a 'Value' text input set to '@div(1,0)'. The pipeline structure above shows the 'If Condition' activity with both 'True' and 'False' branches visible.

Expand **General** in the **Activities** bar, and then drag-and-drop a “Set variable” component into the activity window.

On the **Variables** tab of the “Set Variable...” component, enter values for the following items:

Name	Select Result from the drop-down
Value	Enter <code>@div(1,0)</code> ... this will force an error result

Confirm Success

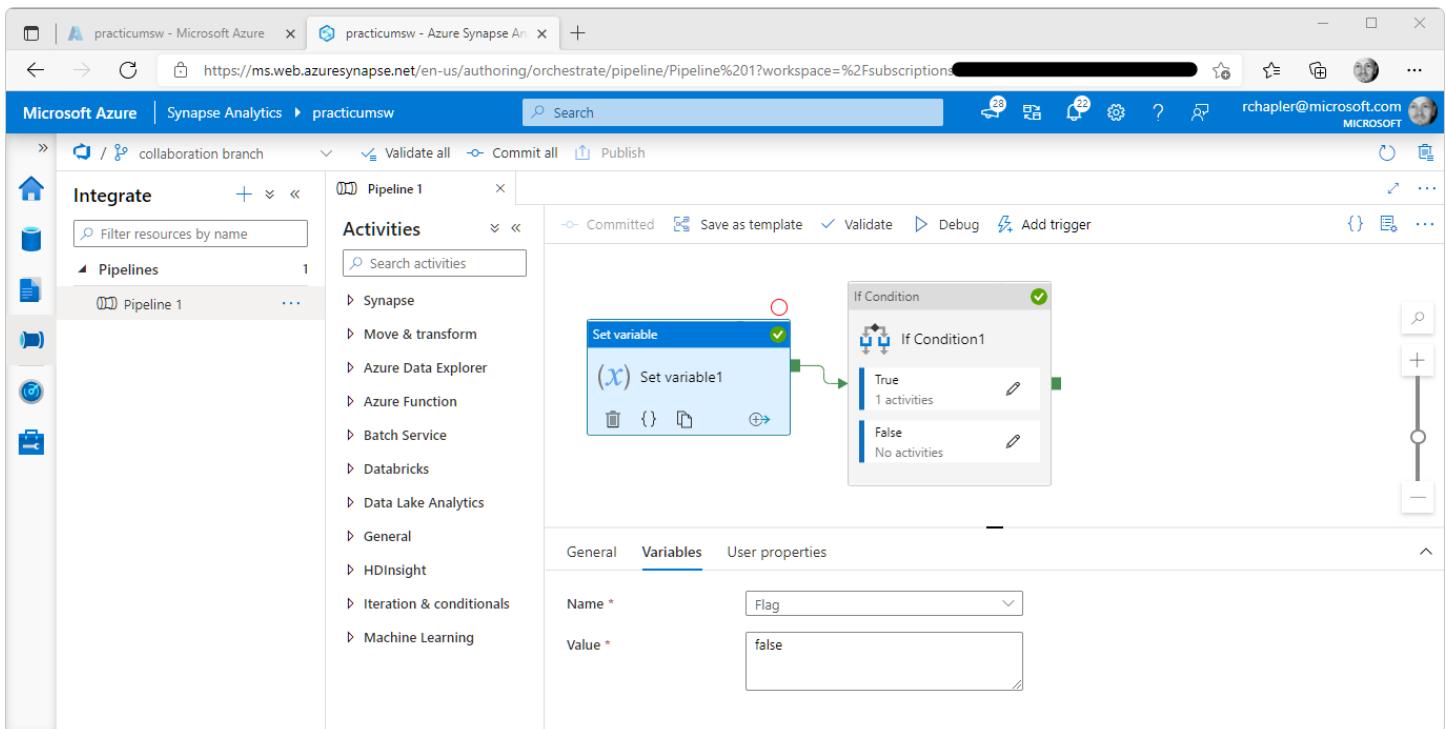
Click **Debug**.

The screenshot shows the Microsoft Azure Synapse Analytics pipeline editor interface. On the left, the sidebar displays 'Integrate' and 'Pipeline 1'. The main area shows a pipeline diagram with a 'Set variable' activity followed by an 'If Condition' activity. The 'If Condition' activity has two branches: 'True' (1 activities) and 'False' (No activities). Below the diagram, the 'Output' tab of the pipeline run details is visible, showing a table of activities and their status. The table includes columns for Name, Type, Run start, Duration, Status, and Integral. All activities listed have a status of 'Succeeded'.

Name	Type	Run start	Duration	Status	Integral
Set variable2	Set variable	2021-09-22T20:13:20.675	00:00:01	Succeeded	
If Condition1	If Condition	2021-09-22T20:13:20.284	00:00:01	Succeeded	
Set variable1	Set variable	2021-09-22T20:13:19.847	00:00:01	Succeeded	

You should see **Succeeded** messages given current variable settings.

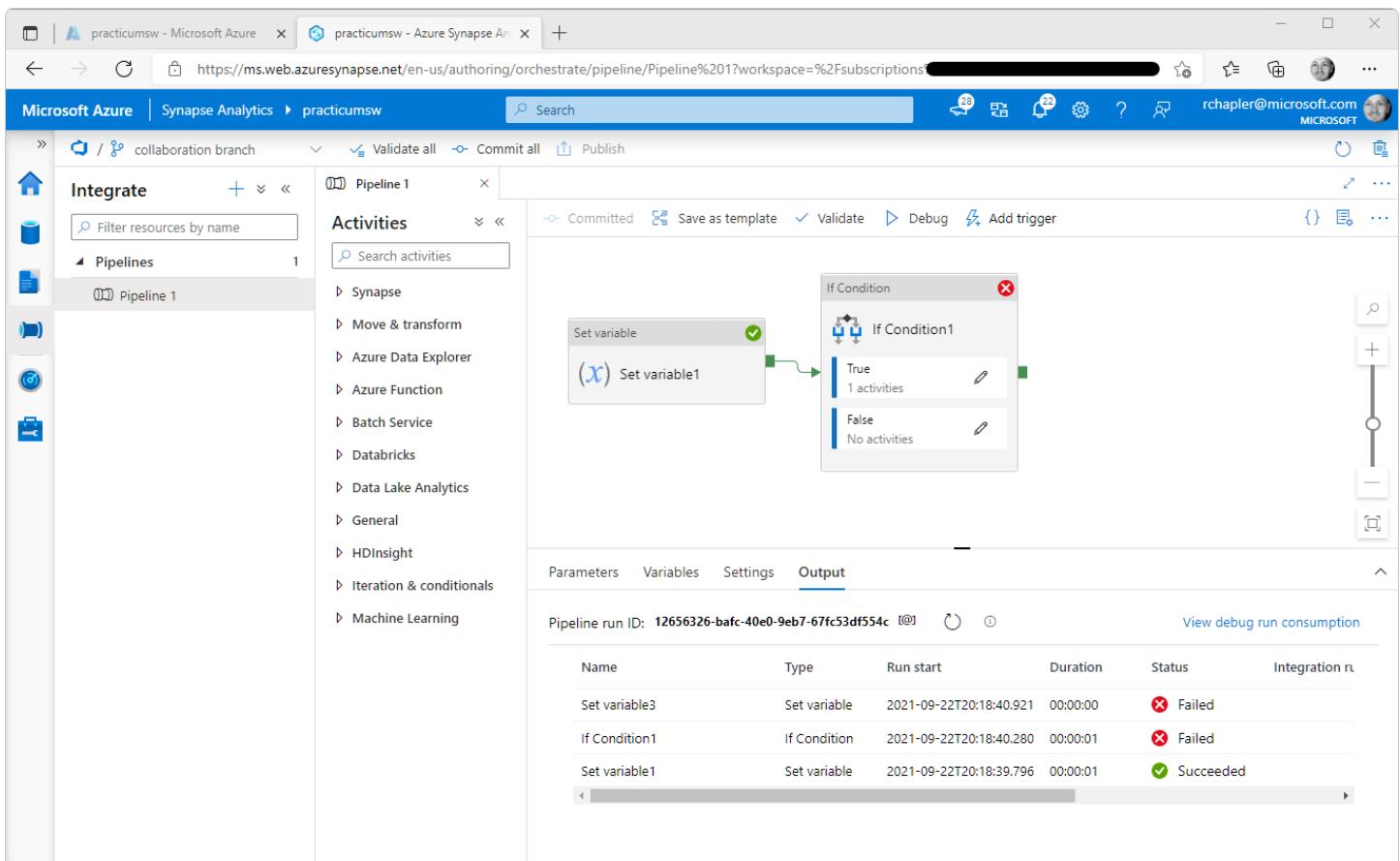
Next, we will make a pipeline change that will force pipeline failure.



Navigate to the “**Set variable1**” component, **Variables** tab.

Replace the **true** value with **false**.

Click **Debug**.



You should see **Failed** messages given current variable settings.



Good Job!

You have successfully completed the objective

Objective: Databricks + Data Lake | Mount Data Lake

Follow the instructions in this section to **establish connection to data in Azure Data Lake Storage**.

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Data Lake (with “Storage Blob Data Reader” permissions set for your Application Registration)
- Databricks (with cluster, notebook, and secret scope)

The Logic

Navigate to Databricks and the **practicumdbn** notebook.

Add a new cell and paste the following code:

```
configs = {"fs.azure.account.auth.type": "OAuth",
           "fs.azure.account.oauth.provider.type":
"org.apache.hadoop.fs.azurebfs.oauth2.ClientCredsTokenProvider",
           "fs.azure.account.oauth2.client.id": dbutils.secrets.get( scope="practicumdbss", key="practicumar-clientid" ),
           "fs.azure.account.oauth2.client.secret": dbutils.secrets.get( scope="practicumdbss", key="practicumar-clientsecret" ),
           "fs.azure.account.oauth2.client.endpoint": "https://login.microsoftonline.com/" +
dbutils.secrets.get( scope="practicumdbss", key="practicumar-tenantid" ) + "/oauth2/token" }

adlsAccount = "practicumdl"
adlsContainer = "practicumdlc"
adlsFolder = ""
mountPoint = "/mnt/practicum"

if not any(mount.mountPoint == mountPoint for mount in dbutils.fs.mounts()):
    dbutils.fs.mount( source = "abfss://" + adlsContainer + "@" + adlsAccount + ".dfs.core.windows.net/" +
adlsFolder, mount_point = mountPoint, extra_configs = configs )
```

Some callouts...

- **practicumdbss** ... refers to the Secret Scope
- **myClientId** ... refers to the Key Vault secret containing the “Application (client) ID”
- **myClientSecret** ... refers to the Key Vault secret containing the “Client Secret”
- **myTenantId** ... refers to the Key Vault secret containing the “Directory (tenant) ID”
- **practicumdl** ... refers to your Data Lake
- **practicumdlc** ... refers to the Data Lake Container
- **adlsFolder** ... placeholder / syntax for inclusion of a folder (null because it is not applicable in this instance)

Run “**Cmd 1**”.

The screenshot shows a Microsoft Azure Databricks notebook interface. The top navigation bar includes tabs for 'practicumdb - Microsoft Azure' and 'practicumdbn - Databricks'. The main area is titled 'practicumdbn (Python)' and contains a single command cell labeled 'Cmd 1'. The code in the cell is as follows:

```
1 configs = {"fs.azure.account.auth.type": "OAuth",
2             "fs.azure.account.oauth.provider.type": "org.apache.hadoop.fs.azurebfs.oauth2.ClientCredsTokenProvider",
3             "fs.azure.account.oauth2.client.id": dbutils.secrets.get(scope="practicumdbss", key="practicumar-clientid"),
4             "fs.azure.account.oauth2.client.secret": dbutils.secrets.get(scope="practicumdbss", key="practicumar-clientsecret"),
5             "fs.azure.account.oauth2.client.endpoint": "https://login.microsoftonline.com/" +
6             dbutils.secrets.get(scope="practicumdbss", key="practicumar-tenantid") + "/oauth2/token"}
7
8 adlsAccount = "practicumdl"
9 adlsContainer = "practicumdlc"
10 adlsFolder = ""
11 mountPoint = "/mnt/practicum"
12
13 if not any(mount.mountPoint == mountPoint for mount in dbutils.fs.mounts()):
14     dbutils.fs.mount( source = "abfss://" + adlsContainer + "@" + adlsAccount + ".dfs.core.windows.net/" + adlsFolder,
15     mount_point = mountPoint, extra_configs = configs )
```

Below the code, there is a section titled '(1) Spark Jobs' and a note indicating the command took 26.98 seconds. A tooltip at the bottom left says 'Shift+Enter to run'.

Quick Idea...

Consider organizing mount locations by associating them with a Databricks database entity; code example:

```
%sql
CREATE DATABASE myDatabase LOCATION "/mnt/practicum/myDatabase"
```

Confirm Success

Proof #1 ... Sample Files at Mount Point

Note: Running this logic requires "Storage Blob Data Reader" permissions on your data lake.

Add a cell below "**Cmd 1**". In cell "**Cmd 2**", paste and run the following code:

```
%fs
ls "/mnt/practicum"
```

The screenshot shows the Microsoft Azure Databricks workspace. On the left, there's a sidebar with icons for Home, Workspace (which is selected), and Recents. The main area has a tab bar with 'practicumdb - Microsoft Azure' and 'practicumdb - Databricks'. Below the tabs is a header with 'Microsoft Azure | Databricks' and a user profile. The main content area is titled 'practicumdbn (Python)' and contains a code editor with two commands:

```
1 %fs
2 ls "mnt/practicum"
```

Below the code editor is a table showing the results of the 'ls' command:

	path	name	size
1	dbfs:/mnt/practicum/sample1.csv	sample1.csv	502
2	dbfs:/mnt/practicum/sample2.csv	sample2.csv	56890
3	dbfs:/mnt/practicum/sample3.csv	sample3.csv	723

Showing all 3 rows.

You should see your sample files at the mount point.

Proof #2 ... Schema and Data

In your Databricks notebook, add a new cell, then paste and run the following code:

```
df = spark.read.csv("dbfs:/mnt/practicum/sample1.csv")
df.printSchema()
df.show()
```

practicumdb - Microsoft Azure prakticumdbn - Databricks

<https://adb-2154823451042175.15.azuredatabricks.net/?o=2154823451042175#notebook/1887258985769069/com...>

Microsoft Azure | Databricks

practicumdbn (Python)

practicumdbc

Cmd 3

```

1 df = spark.read.csv("dbfs:/mnt/practicum/sample1.csv")
2 df.printSchema()
3 df.show()

```

(2) Spark Jobs

df: pyspark.sql.dataframe.DataFrame = [c0: string, c1: string ... 11 more fields]

Month	"Average"	"2005"	"2006"	"2007"	"2008"	"2009"	"2010"	"2011"	"2012"	"2013"	"2014"	"2015"
May	0.1	0	0	1	1	0	0	0	2	0	0	0
Jun	0.5	2	1	1	0	0	1	1	2	2	0	1
Jul	0.7	5	1	1	2	0	1	3	0	2	2	1
Aug	2.3	6	3	2	4	4	4	7	8	2	2	3
Sep	3.5	6	4	7	4	2	8	5	2	5	2	5
Oct	2.0	8	0	1	3	2	5	1	5	2	3	0
Nov	0.5	3	0	0	1	1	0	1	0	1	0	1
Dec	0.0	1	0	1	0	0	0	0	0	0	0	1

Command took 5.57 seconds -- by rchaphler@microsoft.com at 1/26/2021, 9:48:04 AM on practicumdbc

Shift+Enter to run [shortcuts](#)

Some observations:

- Schema ... the schema is not defined, so the resulting interpretation calls out each field as data type string.
- Resultset ... data formatting {e.g., first row as headers, double quotes, etc.} hasn't been applied.



Good Job!

You have successfully completed the objective

Objective: Databricks | Localize Mounted Data

Follow the instructions in this section to **localize data from a Data Lake mount**.

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Data Lake
- Databricks

Why Localize?

Databricks File System (DBFS) is a distributed file system mounted into a Databricks workspace and available on Databricks clusters. DBFS is an abstraction on top of scalable object storage and offers the following **benefits**:

- Allows you to mount storage objects so that you can **seamlessly access data without requiring credentials**.
- Allows you to **interact with object storage using directory and file semantics instead of storage URLs**.
- Persists files to object storage, so you **won't lose data after you terminate a cluster**.

[Databricks File System \(DBFS\) — Databricks Documentation](#) | December 16, 2020

Though it isn't listed here, many customers refer to the following benefits when justifying the decision to localize data from Azure Data Lake Storage to Azure Databricks:

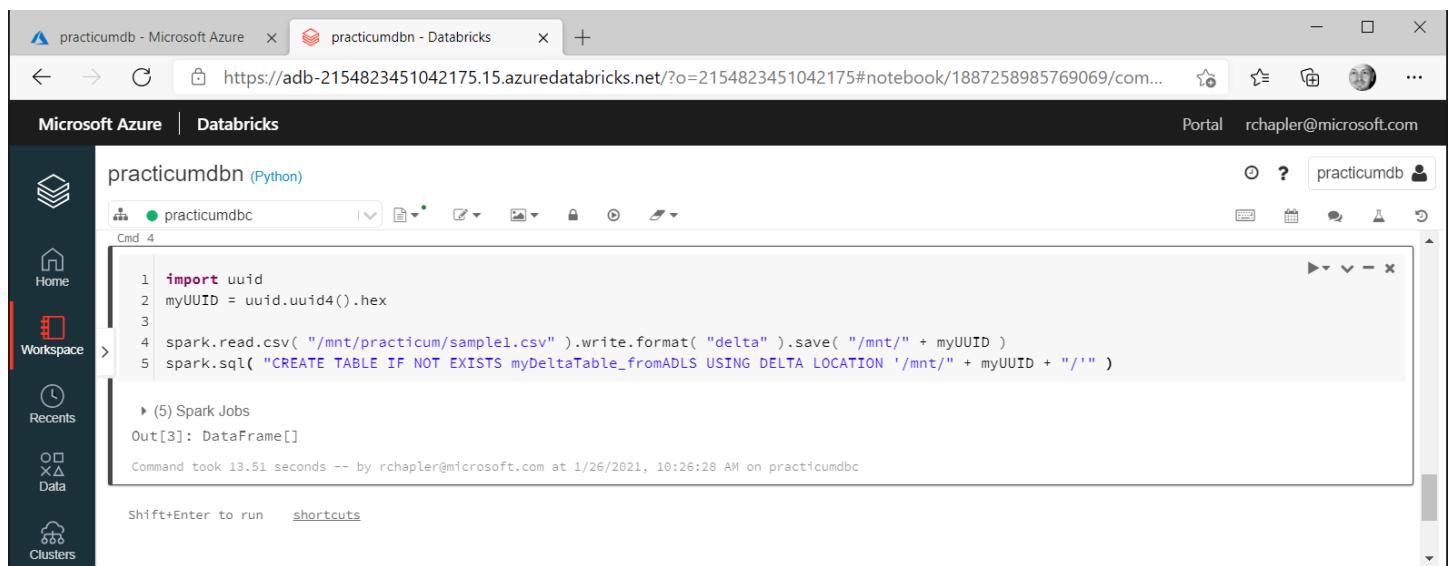
- Performance Gains
- Decreased Compute Cost
- Leverage Benefits of Partitioning

The Logic

Navigate to Databricks and the **practicumdbn** notebook. Add a new cell and paste the following code:

```
import uuid
myUUID = uuid.uuid4().hex

spark.read.csv( "/mnt/practicum/sample1.csv" ).write.format( "delta" ).save( "/mnt/" + myUUID )
spark.sql( "CREATE TABLE IF NOT EXISTS myDeltaTable_fromADLS USING DELTA LOCATION '/mnt/" + myUUID + "/" )
```



The screenshot shows the Microsoft Azure Databricks interface. The top bar displays the workspace name 'practicumdb - Microsoft Azure' and the notebook title 'practicumdbn - Databricks'. The left sidebar has navigation links for Home, Workspace (which is selected), Recents, Data, and Clusters. The main area is titled 'practicumdbn (Python)' and contains a code editor with the following content:

```
import uuid
myUUID = uuid.uuid4().hex

spark.read.csv( "/mnt/practicum/sample1.csv" ).write.format( "delta" ).save( "/mnt/" + myUUID )
spark.sql( "CREATE TABLE IF NOT EXISTS myDeltaTable_fromADLS USING DELTA LOCATION '/mnt/" + myUUID + "/" )
```

Below the code editor, there's a 'Spark Jobs' section with '(5) Spark Jobs' and an 'Out[3]: DataFrame[]' result. A note at the bottom says 'Command took 13.51 seconds -- by rchapler@microsoft.com at 1/26/2021, 10:26:28 AM on practicumdbc'. At the bottom of the code editor, it says 'Shift+Enter to run shortcuts'.

Click on **Data** in the left-hand navigation. Click on the **default** database, then the **mydeltatable_fromadls** table.

practicumdb - Microsoft Azure mydeltatable_fromadls - Databricks

https://adb-2154823451042175.15.azuredatabricks.net/?o=2154823451042175#table/default/mydeltatable_fromadls

Microsoft Azure | Databricks

Portal rchapler@microsoft.com

Data

Databases default

Tables mydeltatable_fromadls

Table: mydeltatable_fromadls

mydeltatable_fromadls

practicumdbc

Details History Refresh

Description:
Created at: 2021-01-26 18:26:29
Last modified: 2021-01-26 18:26:32
Partition columns:
Number of files: 1
Size: 3.54 kB

Schema:

	col_name	data_type	comment
1	_c0	string	
2	_c1	string	
3	_c2	string	
4	_c3	string	
5	_c4	string	
6	_c5	string	
7	_c6	string	
8	_c7	string	

Showing all 16 rows.

Sample Data:

	_c0	_c1	_c2	_c3	_c4	_c5
1	Month	"Average"	"2005"	"2006"	"2007"	"2008"
2	May	0.1	0	0	1	1
3	Jun	0.5	2	1	1	0
4	Jul	0.7	5	1	1	2
5	Aug	2.3	6	3	2	4
6	Sep	3.5	6	4	7	4
7	Oct	2.0	8	0	1	3
8	Nov	0.5	3	0	0	1

Showing all 9 rows.

Observation: Previously noted schema and resultset challenges remain present in the newly created table.

Good Job!



You have successfully completed the objective

Objective: Databricks | Source from APIs

Follow the instructions in this section to **source data from a REST API**.

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Databricks
- Key Vault

Stage Resources

For this exercise, we will use “**Current Weather**” data from <https://openweathermap.org/api>.

The screenshot shows the OpenWeatherMap Weather API page. At the top, there's a navigation bar with links for Guide, API, Pricing, Maps, FAQ, Partners, Blog, Marketplace, Sign in, and Support. Below the bar, the main title is "Weather API". A search bar says "Weather in your city". The page content includes a section about monthly subscriptions and how to start. Three main data collection sections are listed:

- Current Weather Data**: Includes "API doc" and "Subscribe" buttons. A bulleted list details: Access current weather data for any location including over 200,000 cities; We collect and process weather data from different sources such as global and local weather models, satellites, radars and vast network of weather stations; JSON, XML, and HTML formats; Available for both Free and paid subscriptions.
- Hourly Forecast 4 days**: Includes "API doc" and "Subscribe" buttons. A bulleted list details: Hourly forecast is available for 4 days; Forecast weather data for 96 timestamps; Higher geographic accuracy; JSON and XML formats; Available for Developer, Professional and Enterprise accounts.
- One Call API**: Includes "API doc" and "Subscribe" buttons. A bulleted list details: Make one API call and get current, forecast and historical weather data; Minute forecast for 1 hour; Hourly forecast for 48 hours; Daily forecast for 7 days; Historical data for 5 previous days; National weather alerts; JSON format; Available for both Free and paid subscriptions.

I chose this as my sample source because they have a Free subscription, their data is relatable, and it is easy to use.

Navigate to <https://openweathermap.org/price>.

The screenshot shows the OpenWeatherMap Pricing page. At the top, there's a navigation bar with links for Guide, API, Pricing, Maps, FAQ, Partners, Blog, Marketplace, Sign In, and Support. Below the navigation is a search bar labeled "Weather in your city". The main content area has a heading "Pricing" and a sub-heading "Current weather and forecasts collection". A table displays five pricing plans:

Free	Startup 40 USD / month	Developer 180 USD / month	Professional 470 USD / month	Enterprise 2,000 USD / month
Get API key	Subscribe	Subscribe	Subscribe	Subscribe
60 calls/minute 1,000,000 calls/month	600 calls/minute 10,000,000 calls/month	3,000 calls/minute 100,000,000 calls/month	30,000 calls/minute 1,000,000,000 calls/month	200,000 calls/minute 5,000,000,000 calls/month
Current Weather	Current Weather	Current Weather	Current Weather	Current Weather

Click the “Get API key” button in the **Free** section and complete the “[Create a New Account](#)” process.

When you’re successfully created an account, making an API call for “[Current Weather](#)” data is as simple as:

`api.openweathermap.org/data/2.5/weather?q={city name}&appid={API key}`

The screenshot shows a browser window with the URL `api.openweathermap.org/data/2.5/weather?q=redmond&appid=...`. The page title is “Not secure”. The page content displays a JSON response for the weather in Redmond:

```
{"coord":{"lon":-122.1215,"lat":47.674}, "weather":[{"id":804,"main":"Clouds","description":"overcast clouds","icon":"04d"}], "base":"stations", "main": {"temp":279.53,"feels_like":277.4,"temp_min":278.15,"temp_max":280.93,"pressure":1007,"humidity":65,"sea_level":1007,"grnd_level":1006}, "visibility":10000,"wind":{"speed":0.27,"deg":241}, "clouds":{"all":100}, "dt":1611870000, "sys": {"type":3,"id":2010401,"country":"US","sunrise":1611848385,"sunset":1611882185}, "timezone":-28800, "id":5808079, "name": "Redmond", "cod":200}
```

Key Vault

Add your API Key to your instance of Key Vault.

The screenshot shows the Microsoft Azure portal interface for creating a secret. The page title is "Create a secret - Microsoft Azure". The URL in the address bar is <https://ms.portal.azure.com/#@microsoft.onmicrosoft.com/resource/subscriptions/91e9fddc-...>. The user is signed in as rchapler@microsoft.com (MICROSOFT (MICROSOFT.ONM...)).

The main content area is titled "Create a secret". It contains the following fields:

- Upload options: Manual
- Name: myOpenWeatherAPIKey
- Value: (redacted)
- Content type (optional): (empty)
- Set activation date?: (unchecked)
- Set expiration date?: (unchecked)
- Enabled?: Yes (selected)

At the bottom left is a blue "Create" button.

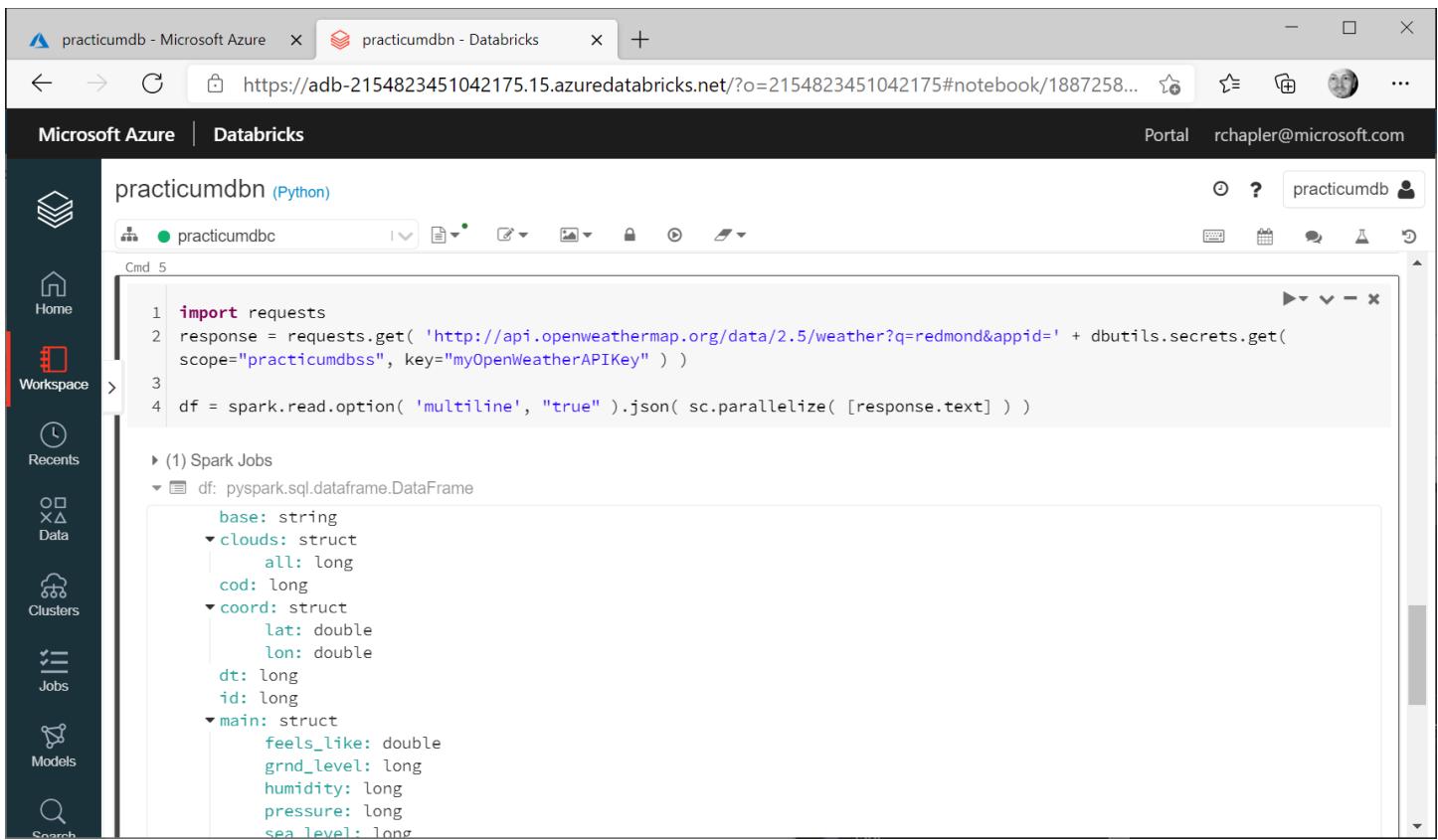
The Logic

Navigate to Databricks and the **practicumdbn** notebook. Add a new cell and paste the following code:

```
import requests
response = requests.get( 'http://api.openweathermap.org/data/2.5/weather?q=redmond&appid=' +
dbutils.secrets.get( scope="practicumdbss", key="myOpenWeatherAPIKey" ) )

df = spark.read.option( 'multiline', "true" ).json( sc.parallelize( [response.text] ) )
```

Run the cell.



```
import requests
response = requests.get( 'http://api.openweathermap.org/data/2.5/weather?q=redmond&appid=' + dbutils.secrets.get(scope="practicumdbss", key="myOpenWeatherAPIKey" ) )
df = spark.read.option( 'multiline', "true" ).json( sc.parallelize( [response.text] ) )

▶ (1) Spark Jobs
  ▶ df: pyspark.sql.dataframe.DataFrame
    base: string
    ▶ clouds: struct
      all: long
      cod: long
    ▶ coord: struct
      lat: double
      lon: double
    dt: long
    id: long
    ▶ main: struct
      feels_like: double
      grnd_level: long
      humidity: long
      pressure: long
      sea_level: long
```

Another Python example (including POST, headers, and body):

```
import json
stringJSON = '{"data": [{"1212":0,"1227":0,..."ZZ9":0}]}'
theJSON = json.loads(stringJSON)
response = requests.post(
    'http://eedfb2cf-deb1-4260-971d-d7a6c308d9b2.eastus.azurecontainer.io/score',
    headers={'CONTENT-TYPE': 'application/json'},
    json = theJSON
)
response.json()
```



Good Job!

You have successfully completed the objective

Objective: Databricks | Batch Upsert Data

Follow the instructions in this section to **provide for recurring upsert of data**.

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Data Lake
- Databricks

CREATE TABLE

Navigate to Databricks and the **practicumdbn** notebook. Add a new cell and paste the following code:

```

import uuid
myUUID = uuid.uuid4().hex

df.write.format("delta").save( "/mnt/" + myUUID )
spark.sql("CREATE TABLE IF NOT EXISTS myOpenWeatherData USING DELTA LOCATION '/mnt/" + myUUID + "'")

```

Notes:

- We are using dataframe **df** created in [Objective #3 ... The Logic](#)
- We are using format **delta** to provide for future UPSERT operations.

The screenshot shows the Microsoft Azure Databricks interface. On the left, there's a sidebar with icons for Home, Workspace (which is selected), Recents, Data, and Clusters. The main area is titled 'practicumdbn (Python)' and shows a code editor with the following Python code:

```

1 import uuid
2 myUUID = uuid.uuid4().hex
3
4 df.write.format("delta").save( "/mnt/" + myUUID )
5 spark.sql("CREATE TABLE IF NOT EXISTS myOpenWeatherData USING DELTA LOCATION '/mnt/" + myUUID + "'")

```

Below the code editor, it says '(4) Spark Jobs' and 'Out[4]: DataFrame[]'. It also notes 'Command took 15.87 seconds -- by rchabler@microsoft.com at 2/1/2021, 12:50:40 PM on practicumdbc'. At the bottom of the code editor, it says 'Shift+Enter to run shortcuts'.

MERGE INTO

Navigate to Databricks and the **practicumdbn** notebook. Add a new cell and paste the following code:

```

import requests
import uuid

response = requests.get( 'http://api.openweathermap.org/data/2.5/weather?q=redmond&appid=' +
dbutils.secrets.get( scope="practicumdbss", key="myOpenWeatherAPIKey" ) )
df = spark.read.option( 'multiline', "true" ).json( sc.parallelize( [response.text] ) )

myUUID = uuid.uuid4().hex
df.write.format("delta").save( "/mnt/" + myUUID )
spark.sql("DROP TABLE myOpenWeatherData_updates")
spark.sql("CREATE TABLE myOpenWeatherData_updates USING DELTA LOCATION '/mnt/" + myUUID + "'")

mySQL = "MERGE INTO myOpenWeatherData existing USING myOpenWeatherData_updates updates ON existing.id = updates.id AND existing.dt = updates.dt"
mySQL += " WHEN MATCHED THEN UPDATE SET existing.main = updates.main"
mySQL += " WHEN NOT MATCHED THEN INSERT (id, dt, main) VALUES (id, dt, main)"

print( mySQL )

```

```
spark.sql( mySQL )
```

Run the cell.

The screenshot shows the Microsoft Azure Databricks workspace interface. On the left, there's a sidebar with icons for Home, Workspace (which is selected), Recents, Data, Clusters, Jobs, Models, and Search. The main area has tabs for 'practicumdb - Microsoft Azure' and 'practicumdbn - Databricks'. The URL in the browser bar is <https://adb-2154823451042175.15.azure.databricks.net/?o=2154823451042175#notebook/1887258...>. The notebook title is 'practicumdbn (Python)'. The code cell contains Python code to import requests, uuid, and spark, then performs a series of operations including reading from a multiline JSON response, creating a Delta table, and performing a MERGE operation into a MySQL table. The output shows the generated SQL command and the resulting DataFrame. A note at the bottom says 'Command took 16.50 seconds -- by rchapler@microsoft.com at 2/1/2021, 1:00:44 PM on practicumdbc'.

Confirm Success

Navigate to Databricks and the **practicumdbn** notebook. Add a new cell and paste the following code:

```
%sql  
SELECT * FROM myOpenWeatherData
```

Run the cell.

The screenshot shows the Microsoft Azure Databricks workspace. On the left is a sidebar with icons for Home, Workspace (selected), Recents, Data, Clusters, Jobs, Models, and Search. The main area has two tabs: "practicumdb - Microsoft Azure" and "practicumdb - Databricks". The "practicumdb - Databricks" tab is active, displaying a Python notebook titled "practicumdbn (Python)". The notebook contains the following code:

```
MERGE INTO myOpenWeatherData existing USING myOpenWeatherData_updates updates ON existing.id = updates.id AND existing.dt = updates.dt WHEN MATCHED THEN UPDATE SET existing.main = updates.main WHEN NOT MATCHED THEN INSERT (id, dt, main) VALUES (id, dt, main)
Out[8]: DataFrame[]

Command took 16.50 seconds -- by rchapler@microsoft.com at 2/1/2021, 1:00:44 PM on practicumdbc
```

Below the notebook is a command editor window titled "Cmd 8" containing:

```
1 %sql
2 SELECT * FROM myOpenWeatherData
```

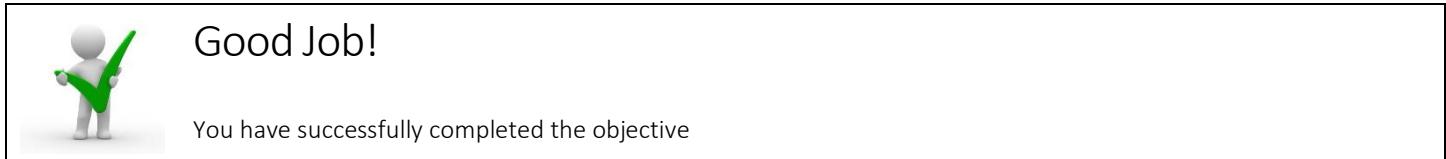
Underneath the command editor is a table titled "(3) Spark Jobs" showing the results of the query:

	base	clouds	cod	coord	dt	id	main
1	stations	►{"all": 100}	200	►{"lat": 47.674, "lon": -122.1215}	1612211690	5808079	►{"feels_like": 280.49, "grnd_level": 1004, "h": 282.59, "temp_min": 281.48}
2	null	null	null	null	1612212650	5808079	►{"feels_like": 279.62, "grnd_level": 1005, "h": 282.59, "temp_min": 281.48}

Showing all 2 rows.

At the bottom of the command editor is the message: "Command took 1.18 seconds -- by rchapler@microsoft.com at 2/1/2021, 1:22:13 PM on practicumdbc".

Note that the row created using MERGE INTO has a limited number of populated cells since only some of the possible cells were included in WHEN NOT MATCHED THEN INSERT (id, dt, main)... we can, of course, expand this to include all fields about which we care.



Objective: Synapse | Audit Data Usage

Follow the instructions in this section to **support regular audit of and alerting on key data assets**.

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Synapse

Synapse | Auditing

Navigate to the **Overview** page in Synapse, Dedicated SQL Pool.

The screenshot shows the Microsoft Azure portal interface for a Dedicated SQL pool named 'practicumdsp'. The left sidebar contains navigation links for Overview, Activity log, Access control (IAM), Tags, Settings (Workload management, Maintenance schedule, Geo-backup policy, Connection strings, Properties, Locks), Security (Auditing, Data Discovery & Classification, Dynamic Data Masking, Security Center), and Notifications (0). The main content area displays the 'Essentials' section with details such as Resource group (practicumrg), Status (Online), Location (West US 2), Subscription (rchapler), Subscription ID (91e9fddc-ef15-416c-9be1-085f8b1b46ed), and Tags (Click here to add tags). It also shows 'Features (4)' (Transparent data encryption, Auditing) and 'Tasks (2)' (All, Security (3), Recovery (1)). The 'Auditing' feature is shown as NOT CONFIGURED.

Click the **Auditing** button.

The screenshot shows the 'Auditing' configuration page for the 'practicumdsp' Dedicated SQL pool. The left sidebar is identical to the previous screenshot. The main content area has a heading 'practicumdsp (practicumsaw/practicumdsp) | Auditing'. It includes a note: 'If Azure SQL Auditing is enabled on the workspace, it will always apply to the SQL Pool, regardless of the SQL Pool settings.' Below this are 'View workspace settings' and 'Workspace-level Auditing: Disabled'. Under 'Azure SQL Auditing', it says 'Azure SQL Auditing tracks SQL Pool events and writes them to an audit log in your Azure storage account.' A 'Learn more about Azure SQL Auditing' link is provided. A switch labeled 'Enable Azure SQL Auditing' is set to 'ON'. Below this is a 'Storage details' section for 'practicumsa'.

Flip “Enable Azure SQL Auditing” switch to ON position. Click to configure “Storage details”.

Screenshot of the Microsoft Azure Storage settings dialog.

The dialog shows the following configuration:

- Subscription: rchapler
- *Storage account: practicumsa
- Retention (Days): 0
- Storage access key: Primary (selected)

At the bottom is an "OK" button.

Save changes.

Click “View audit logs”.

Screenshot of the Microsoft Azure Audit records page.

The page shows the following information:

- Audit source: Workspace audit (selected)
- Showing audit records up to Wed, 03 Feb 2021 20:58:37 UTC.
- Table of audit records:

Event time (UTC)	Principal name	Event type	Action status
2/3/2021 8:58:30 PM	##MS_InstanceCertificate##	BATCH COMPLETED	Succeeded
2/3/2021 8:56:26 PM	##MS_InstanceCertificate##	BATCH COMPLETED	Succeeded
2/3/2021 8:56:26 PM	##MS_InstanceCertificate##	DATABASE AUTHENTICATION SUCCEEDED	Succeeded

Click into one of “BATCH COMPLETED” audit records.

Audit record - Microsoft Azure

Microsoft Azure (Preview)

Search resources, services, and docs (G+)

Dashboard > practicumdsp (practicumsaw/practicumdsp) > Audit records > Audit record

Event time (UTC)
2/8/2021 7:32:50 PM

Event type
BATCH COMPLETED

Server name
practicumsaw

Database name
practicumdsp

Application name
SynapseSqlEditor

Principal name
rchapler@microsoft.com

Client IP
76.121.194.132

Status
Succeeded

STATEMENT

```
SELECT
    s.name AS [schema_name],
    o.name AS [object_name], o.type AS [object_type], o.type_desc AS
    [object_type_desc],
    c.name AS [column_name], TYPE_NAME(c.system_type_id) AS
    [column_type]
FROM sys.schemas s
LEFT OUTER JOIN sys.all_objects o
    ON s.schema_id = o.schema_id
LEFT OUTER JOIN sys.all_columns c
    ON c.object_id = o.object_id
```

Expand statement



Good Job!

You have successfully completed the objective

Objective: Data Explorer | Query from On-Prem

Follow the instructions in this section to **query from an on-prem client**.

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Azure CLI
- Data Explorer
- Postman

Bearer Token

In a Command Prompt window (with [Azure CLI](#) installed), execute the following command: `az login`

Use the resulting browser window to provide credentials.

In a Command Prompt window (with [Azure CLI](#) installed), execute the following command: `az account get-access-token --resource https://practicumdec.westus2.kusto.windows.net --query accessToken --output tsv`

Capture the returned token for use later in this section.

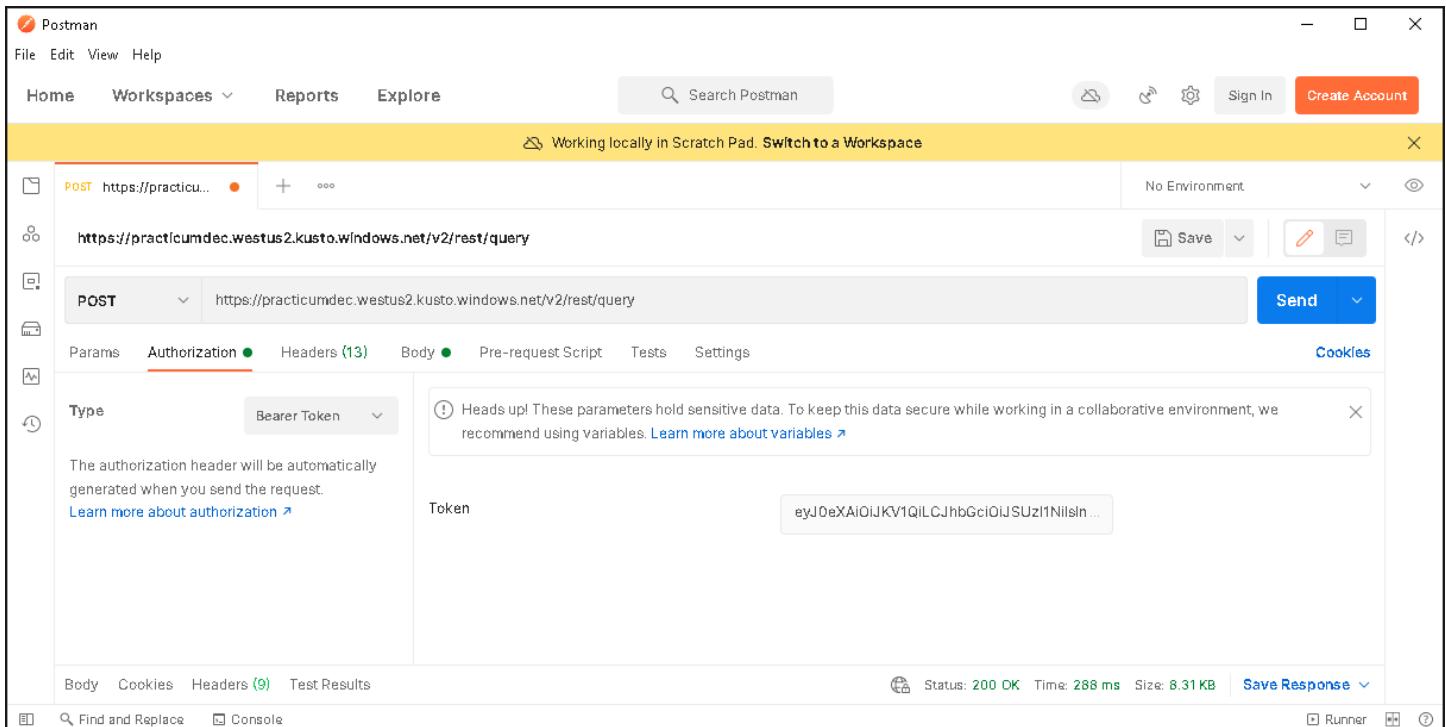


```
C:\Users\rchapler>az account get-access-token --resource https://practicumdec.westus2.kusto.windows.net --query accessToken --output tsv
eyJ0eXAiOiJKV1QiLCJhbGciOiJSUzIiNiIsIng1dCI6ImPbzNaRHJPRFhFsFqS1doWHNsSFJfS1hFZyIsImtpZCI6ImPbzNaRHJPRFhFsFqS1doWHNsSFJfS1hFZyJ9.eyJhdWQ1O1jodHRwczovL3ByYWNoaWN1bWR1Yy53ZXN0dXMyLmt1c3RvLndpbmRvd3MubmVOIiwiAxNzIjoiaHR0cHM6Ly9zdHMu2luZG93cy5uZXQvNzJmOTg4YmYtODZmMS00MWFmLTkxYWItMmQ3Y2QwMTFkYjQ3lyIsImlhdCI6MTYxODk1NjI1NCwibmJmIjoxNjE40TU2MjU0LCJleHAiOjE2MTg5NjAxNTQsImFjciI6IjEiLCJhaW81o1JBVFBDs84VFBUF5Y1NLemdTcmVRNmFrcmh1Ly8xU3lXzmvXVUtyOEUwVh5SlgwUXZ3cGhwSVNOSFhuaXYrc25ROEp6YTNTVewpTUW5QUl3RmdteUpIMFBWWLM3Q1Zydz09IiwiYWIyIjpbInJzYSIsIm1mYSJdLCJhcHBpZCI6Ija0YjaA3NzKLThkZGItNDYxYs11YmVlLTAyZj1lMWJmN2I0NIIsImFwCGlkYWNyIjoIMCIsImRdmljZW1kIjoiMjc3MzgzMWEtMTQ2NS00N2NhLThiODItMDg5NWMI1MGN1MWV1IwiZmfTaWxSX25hbWU1o1JdaGfwbGyIiwiZ2l2ZW5fbmftZS16IljpY2gilCJpcGFkZHii0iI3N14xMjEuMTk0ljezMiIsIm5hbWU1o1JsaNNoIENoXBsZXiLLCJvaQoIjjYmUyOTQ4Yy01NzAxLTQyNjUtODExZC02MzNhMzE0NjczZjAiLCJvbNByZW1fc2lkIjoiUy0xLTUtMjEtMjEyNzUyMTE4NC0xNjA0MDEyOTIwLTE40Dc5Mjc1MjctNDY2NzUyOCiSInB1aWQi0iIxMDAzN0ZGRGwMLFDODYxIiwiicmgi0iIwlkfSb0F2NGo1Y3ZHR3Iwr1JxeTE4ME3JY1I1VjNzQVRialJwR3tUNEMtZUdfZTBZYUFEVS41LCJzY3Ai0i1c2Vx1ltcGVyc29uYXRpb24lCJzdWI10iJsRk9zQ0xkdHb00U1c3VytKtqX3lqam84ax0452zaM0Z4dHdnV3ZNIiwidgk1IjoiNzJn0tG4YmYtODZmMS00MWFmLTkxyWItMmQ3Y2QwMTFkYjQ3IiwidW5pcXV1X25hbWU1o1JyY2hhcGxLckBaWnb3NvZnQuy29tIiwidXBuIjoiemNoYXBsZXJAbwljcm9zb2Z0LmNvbSIsInV0ASI6Iktdqv1hbEdyQ1VhcnndHajNBDB3QUEiLCJ2ZXxi0iXlxjAifq.U6yTE35AdDTQhJgUEjjUPFW6raM90RF07SzFya4azbPzrWkntGBrekyjLzqaR5C-c1zPw9c8dp5Gn-g1Hw0cpXnZTbrA-Mv9-g4Gwsto3I1LxqXkH3bBmhGHqadQv3j1D-N_I1SU1whxj6Fjgsgzfwb1NgCo1gt0nDeEWhrFGv03LPkmwNTFLu3-hUqRQDxm1eu519UF4cxMm00cxSaKegNs2eco0aYAUXxCjgD2cdjv1EgbioL9ZdCNJRb1AgHJcxHVGafPBTyL03XL_Xngc-DpZdyC6MMp1SF4-BEc729N1i14lJHNpGq-hy8DCB2_5bduHwXMGOTnfwbq14PQ
```

C:\Users\rchapler>

API Request

Open a new request in Postman.



Postman

File Edit View Help

Home Workspaces Reports Explore Search Postman Sign In Create Account

Working locally in Scratch Pad. Switch to a Workspace

No Environment

POST https://practicumdec.westus2.kusto.windows.net/v2/rest/query

POST https://practicumdec.westus2.kusto.windows.net/v2/rest/query

Send

Params Authorization Headers (13) Body Pre-request Script Tests Settings Cookies

Type Bearer Token

The authorization header will be automatically generated when you send the request.

Token eyJ0eXAiOiJKV1QiLCJhbGciOiJSUzIiNiIsIng1dCI6ImPbzNaRHJPRFhFsFqS1doWHNsSFJfS1hFZyIsImtpZCI6ImPbzNaRHJPRFhFsFqS1doWHNsSFJfS1hFZyJ9.eyJhdWQ1O1jodHRwczovL3ByYWNoaWN1bWR1Yy53ZXN0dXMyLmt1c3RvLndpbmRvd3MubmVOIiwiAxNzIjoiaHR0cHM6Ly9zdHMu2luZG93cy5uZXQvNzJmOTg4YmYtODZmMS00MWFmLTkxYWItMmQ3Y2QwMTFkYjQ3lyIsImlhdCI6MTYxODk1NjI1NCwibmJmIjoxNjE40TU2MjU0LCJleHAiOjE2MTg5NjAxNTQsImFjciI6IjEiLCJhaW81o1JBVFBDs84VFBUF5Y1NLemdTcmVRNmFrcmh1Ly8xU3lXzmvXVUtyOEUwVh5SlgwUXZ3cGhwSVNOSFhuaXYrc25ROEp6YTNTVewpTUW5QUl3RmdteUpIMFBWWLM3Q1Zydz09IiwiYWIyIjpbInJzYSIsIm1mYSJdLCJhcHBpZCI6Ija0YjaA3NzKLThkZGItNDYxYs11YmVlLTAyZj1lMWJmN2I0NIIsImFwCGlkYWNyIjoIMCIsImRdmljZW1kIjoiMjc3MzgzMWEtMTQ2NS00N2NhLThiODItMDg5NWMI1MGN1MWV1IwiZmfTaWxSX25hbWU1o1JdaGfwbGyIiwiZ2l2ZW5fbmftZS16IljpY2gilCJpcGFkZHii0iI3N14xMjEuMTk0ljezMiIsIm5hbWU1o1JsaNNoIENoXBsZXiLLCJvaQoIjjYmUyOTQ4Yy01NzAxLTQyNjUtODExZC02MzNhMzE0NjczZjAiLCJvbNByZW1fc2lkIjoiUy0xLTUtMjEtMjEyNzUyMTE4NC0xNjA0MDEyOTIwLTE40Dc5Mjc1MjctNDY2NzUyOCiSInB1aWQi0iIxMDAzN0ZGRGwMLFDODYxIiwiicmgi0iIwlkfSb0F2NGo1Y3ZHR3Iwr1JxeTE4ME3JY1I1VjNzQVRialJwR3tUNEMtZUdfZTBZYUFEVS41LCJzY3Ai0i1c2Vx1ltcGVyc29uYXRpb24lCJzdWI10iJsRk9zQ0xkdHb00U1c3VytKtqX3lqam84ax0452zaM0Z4dHdnV3ZNIiwidgk1IjoiNzJn0tG4YmYtODZmMS00MWFmLTkxyWItMmQ3Y2QwMTFkYjQ3IiwidW5pcXV1X25hbWU1o1JyY2hhcGxLckBaWnb3NvZnQuy29tIiwidXBuIjoiemNoYXBsZXJAbwljcm9zb2Z0LmNvbSIsInV0ASI6Iktdqv1hbEdyQ1VhcnndHajNBDB3QUEiLCJ2ZXxi0iXlxjAifq.U6yTE35AdDTQhJgUEjjUPFW6raM90RF07SzFya4azbPzrWkntGBrekyjLzqaR5C-c1zPw9c8dp5Gn-g1Hw0cpXnZTbrA-Mv9-g4Gwsto3I1LxqXkH3bBmhGHqadQv3j1D-N_I1SU1whxj6Fjgsgzfwb1NgCo1gt0nDeEWhrFGv03LPkmwNTFLu3-hUqRQDxm1eu519UF4cxMm00cxSaKegNs2eco0aYAUXxCjgD2cdjv1EgbioL9ZdCNJRb1AgHJcxHVGafPBTyL03XL_Xngc-DpZdyC6MMp1SF4-BEc729N1i14lJHNpGq-hy8DCB2_5bduHwXMGOTnfwbq14PQ

Body Cookies Headers (9) Test Results Status: 200 OK Time: 288 ms Size: 8.31 KB Save Response

Find and Replace Console Runner

Select **POST** from the drop-down. Enter a hostname (with the name you used for your cluster); example:

<https://practicumdec.westus2.kusto.windows.net/v2/rest/query>

On the **Authorization** tab, select “**Bearer Token**” from the drop-down and paste your previously copied token in the resulting interface.

The screenshot shows the Postman application interface. At the top, there's a navigation bar with File, Edit, View, Help, Home, Workspaces, Reports, Explore, and a search bar. On the right of the top bar are icons for Sign In and Create Account. Below the top bar, a yellow banner says "Working locally in Scratch Pad. Switch to a Workspace". The main area shows a list of requests on the left and the details of the current request on the right. The current request is a POST to <https://practicumdec.westus2.kusto.windows.net/v2/rest/query>. The "Headers" tab is selected, showing the following header values:

```
Accept: application/json
Accept-Encoding: deflate
Content-Type: application/json; charset=utf-8
Host: help.kusto.windows.net
```

At the bottom of the interface, there are tabs for Body, Cookies, Headers (9), Test Results, and a status bar showing Status: 200 OK, Time: 288 ms, Size: 8.31 KB, and a Save Response button.

On the **Headers** tab, enter the following items:

Accept application/json

Accept-Encoding deflate

Content-Type application/json; charset=utf-8

Host help.kusto.windows.net

The screenshot shows the Postman application interface. At the top, there's a navigation bar with 'File', 'Edit', 'View', 'Help', 'Home', 'Workspaces', 'Reports', 'Explore', and a search bar. On the right of the top bar are 'Sign In' and 'Create Account' buttons. Below the top bar, a yellow banner says 'Working locally in Scratch Pad. Switch to a Workspace'. The main area shows a request card with 'POST https://practicumdec.westus2.kusto.windows.net/v2/rest/query'. The 'Body' tab is selected, showing a JSON payload:

```
1 {  
2   ... "db": "practicumded",  
3   ... "csl": "StormEvents | take 5"  
4 }
```

Below the body, the status bar shows 'Status: 200 OK Time: 288 ms Size: 8.31 KB' and 'Save Response' button.

On the **Body** tab, paste the following:

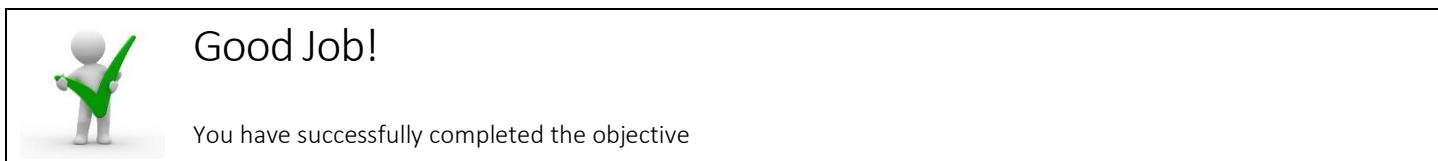
```
{  
  "db": "practicumded",  
  "csl": "StormEvents | take 5"  
}
```

Click the **Send** button.

The screenshot shows the Postman application interface. A POST request is made to <https://practicumdec.westus2.kusto.windows.net/v2/rest/query>. The response status is 200 OK, time 244 ms, and size 8.29 KB. The response body is displayed in Pretty JSON format:

```
1 {
2     "FrameType": "DataSetHeader",
3     "IsProgressive": false,
4     "Version": "v2.0"
5 },
6 {
7     "FrameType": "DataTable",
8     "TableId": 0,
9     "TableKind": "QueryProperties",
10    "TableName": "GetExtendedProperties",
11    "Columns": [
12        {
13            "ColumnName": "TableId",
14            "ColumnType": "int"
15        },
16        {
17            "ColumnName": "Key",
18            "ColumnType": "string"
19        },
20        {
21            "ColumnName": "Value",
22            "ColumnType": "dynamic"
23        }
24    ]
}
```

You should see JSON with both schema and data included.



Data Analysis

Objective(s) in this section describe methods for analyzing data.

Objective: Metrics Advisor + Data Explorer | Onboard and Investigate

Use Case: Company XYZ wants to detect data anomalies for Metric 123. When significant anomalies are detected, the solution should support investigation.

Follow the instructions in this section to **on-board data feeds and investigate findings**.

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Data Explorer
- Metrics Advisor

Navigate to your instance of Metrics Advisor.

The screenshot shows the Microsoft Azure (Preview) portal with the URL <https://ms.portal.azure.com/#@microsoft.onmicrosoft.com/resource/subscriptions/practicumma>. The left sidebar is open, showing the 'Quick start' section under 'Resource Management'. The main content area displays a 'Quick start' guide for Metrics Advisor, divided into three numbered steps:

- 1 Monitor your first metrics using the web-based workspace**

Once you've learnt how the sample data set is configured and how the metrics are monitored by the service, you could go ahead and test with your own data. Pivot to overview blade of your Metrics Advisor resource, open the link displayed in "Web-based workspace". From there you could start onboarding your data. Please read the following document for details.

[Go to your workspace](#)
[Quickstart: Monitor your first metric using the web-based workspace](#)
- 2 Use APIs to customize your own solution**

You are also able to implement your own solution or embed Metrics Advisor into existing solution by using the APIs provisioned with this instance of the service. Please follow the steps in the following documentation to start building your own metrics monitoring solution.

[API reference](#)
[Quickstart: Use API to customize your own solution](#)
- 3 Learn more about the service**

Explore code samples, quickstarts, and the API reference to see what you can do using this product.

[Documentation](#)
[Support options](#)
[Provide feedback](#)

Click the “Go to your workspace” link.

The screenshot shows a Microsoft Edge browser window with the title bar "practicumma - Microsoft Azure" and the address bar "Metrics Advisor" at "https://metricsadvisor.azurewebsites.net". The main content features a large orange sun-like icon and the text "Protect your organization's growth". Below this, a paragraph explains that Metrics Advisor monitors performance metrics like sales revenue and manufacturing operations. To the right, there's a "Welcome to Metrics Advisor!" section with three dropdown menus: "Directory" set to "Microsoft", "Subscription" set to "rchapler", and "Workspace" set to "practicumma". A "Switch account" link is at the top of the dropdowns. At the bottom are "Get started" and "Learn more" buttons.

Welcome to Metrics Advisor!

Select a subscription and a workspace to get started or go to the Azure Portal to subscription and workspace. You can switch subscriptions and workspaces at any time. When switching directory, you will be requested to sign-in again.

Switch account

Directory [?](#) Microsoft

Subscription [?](#) rchapler

Workspace [?](#) practicumma

Create a Metrics Advisor resource

Get started Learn more

On the “Welcome to Metrics Advisor!” page, enter values for the following items:

Directory	Select your directory
Subscription	Select your subscription
Workspace	Select your instance of Metrics Advisor

Click the “Get started” button.

Screenshot of the Microsoft Azure Metrics Advisor Data feeds Onboarding page.

The page title is "Metrics Advisor" and the URL is "https://metricsadvisor.azurewebsites.net/data-feed".

The left sidebar shows the following navigation items:

- Collapse menu
- Onboarding
 - Add data feed
- Monitor & Diagnostic
- Data feeds **(selected)**
- Incident hub
- Metrics graph
- Settings
- API keys
- Hooks
- Credential entity
- Help
 - Documentation
 - Public community

The main content area is titled "Data feeds" and displays three steps:

- Step 1: Onboard time-series data**

Ingest your time-series data from various data sources supported
- Step 2: Tune configuration & subscribe anomaly alerts**

Fine tune Detection configurations to better serve real-world scenarios
Create a hook and subscribe real-time anomaly alerts
- Step 3: Diagnose incidents**

Identify key contributors with dimension tree
Chase down correlations with metrics graph

A blue button at the bottom center says "Onboard my first data feed".

Click the “Onboard my first data feed” button.

The screenshot shows the 'Add data feed' page in the Microsoft Azure Metrics Advisor. The left sidebar has a blue header with the Metrics Advisor logo and a list of navigation items. The 'Add data feed' item is highlighted with a blue background. The main content area has a white background with several input fields and sections. At the top, there's a 'Connection settings' section with a note about granting IP access. Below it are 'Source type' (set to 'Azure Data Explorer (Kusto)'), 'Granularity' (set to 'Daily'), and 'Ingest data since (UTC)' (set to '2007-01-01'). The 'Authentication type' is set to 'Managed Identity'. Under 'Connection string', there's a text input field containing 'Data Source=https://practicumdec.westus2.kusto.windows.net;Initial Catalog=practicumdec'. A note below says 'Tips: Specify the connection string to access your Azure Data Explorer.' and provides a sample connection string. The 'Query' section contains an 'Important' note with two points: 1. Metrics Advisor runs the query continuously to fetch metric data for each interval (equals to granularity). Please use @IntervalStart and @IntervalEnd in your query to limit query result. 2. Please aggregate your metric data by its dimensions, and ensure there's no duplicate values for each dimension combination within each interval. Below this is a code example for 'StormEvents' using Kusto syntax. A 'Load data' button is at the bottom right.

On the “Add data feed” page, enter values for the following items:

Source Type	Select “ Azure Data Explorer (Kusto) ”
Granularity	Confirm default selection, Daily
Ingest Data Since (UTC)	Select a date appropriate for your data set
Authentication Type	Select “ Managed Identity ”
Connection String	Enter in form “Data Source=https://{{ADX Cluster}}.{region}.kusto.windows.net;Initial Catalog={{ADX Database}}
Query	Create and test a query in your database, then parameterize; example below:
	<pre>StormEvents summarize sum(DamageProperty) by startofday(StartTime), State where StartTime >= todatetime("@IntervalStart") and StartTime < todatetime("@IntervalEnd")</pre>

Click the “Load Data” button.

The screenshot shows the Microsoft Azure Metrics Advisor interface for adding a data feed. On the left, a sidebar includes links for 'Credential entity', 'Help', 'Documentation', and 'Public community'. The main area has tabs for 'Query' and 'Schema configuration'. The 'Query' tab displays a DAX-like query:

```
StormEvents  
| summarize sum(DamageProperty) by startofday(StartTime), State  
| where StartTime >= todatetime("@IntervalStart") and StartTime < todatetime("@IntervalEnd")
```

The 'Schema configuration' tab shows sample data and column mappings:

State	StartTime	sum_DamageProperty
NORTH CAROLINA	2007-01-01T00:00:00Z	0
WISCONSIN	2007-01-01T00:00:00Z	0
NEW YORK	2007-01-01T00:00:00Z	20000
ALASKA	2007-01-01T00:00:00Z	0
DELAWARE	2007-01-01T00:00:00Z	0
OKLAHOMA	2007-01-01T00:00:00Z	775000
INDIANA	2007-01-01T00:00:00Z	110000

Below the table, the schema is defined:

Column name	Display name	Column type	Select:
State	State	String	<input type="radio"/> Timestamp <input checked="" type="radio"/> Dimensions <input type="radio"/> Measure
StartTime	Event Date	String	<input checked="" type="radio"/> Timestamp <input type="radio"/> Dimensions <input type="radio"/> Measure
sum_DamageProperty	Count of Events	String	<input type="radio"/> Timestamp <input type="radio"/> Dimensions <input checked="" type="radio"/> Measure

A blue 'Verify schema' button is located at the bottom right of the schema configuration section.

Click the “Verify Schema” button.

practicumma - Microsoft Azure Metrics Advisor https://metricsadvisor.azurewebsites.net/add-data-feed

Advanced settings Help

Automatic roll-up settings Help

My data has already rolled up and the dimension value is represented by NULL or Empty (Default)

I need the service to roll up my data by calculating None and represent it by Set roll-up columns

I do not need to perform root cause analysis into dimensions for my metrics.

Ingestion options Help

Ingestion time offset hours Enter Max concurrency... Stop retrying after hours Minimum retrying interval hours

Data feed not available alert Help

Alert hooks Select Grace period hours Snooze hours consecutive "data feed not available" alerts

By applying hooks, you can receive alerts. You can [create and manage hooks](#).

Misc

Missing points filling for anomaly detection model Help

Smart filling Fill previous Fill custom value: 0 No filling

Action link template Help

Tips: You can use these placeholders in the URL template %datafeed, %metric, %timestamp, %detect_config, %tagset.

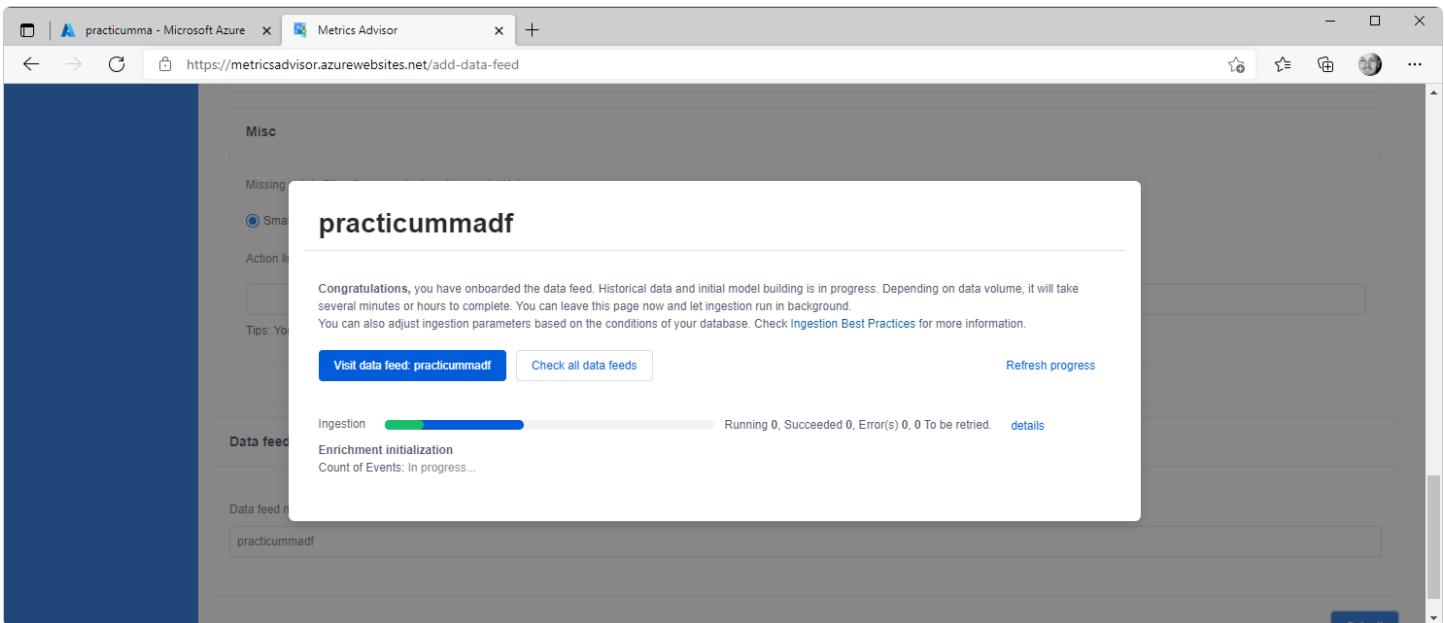
Data feed name

Data feed name practicummadf

Submit

No changes are required to default values in “**Advanced Settings**”.

Enter a “**Data Feed Name**” and then click the Submit button.



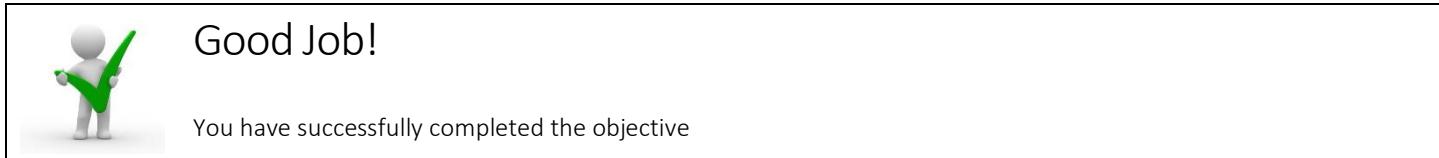
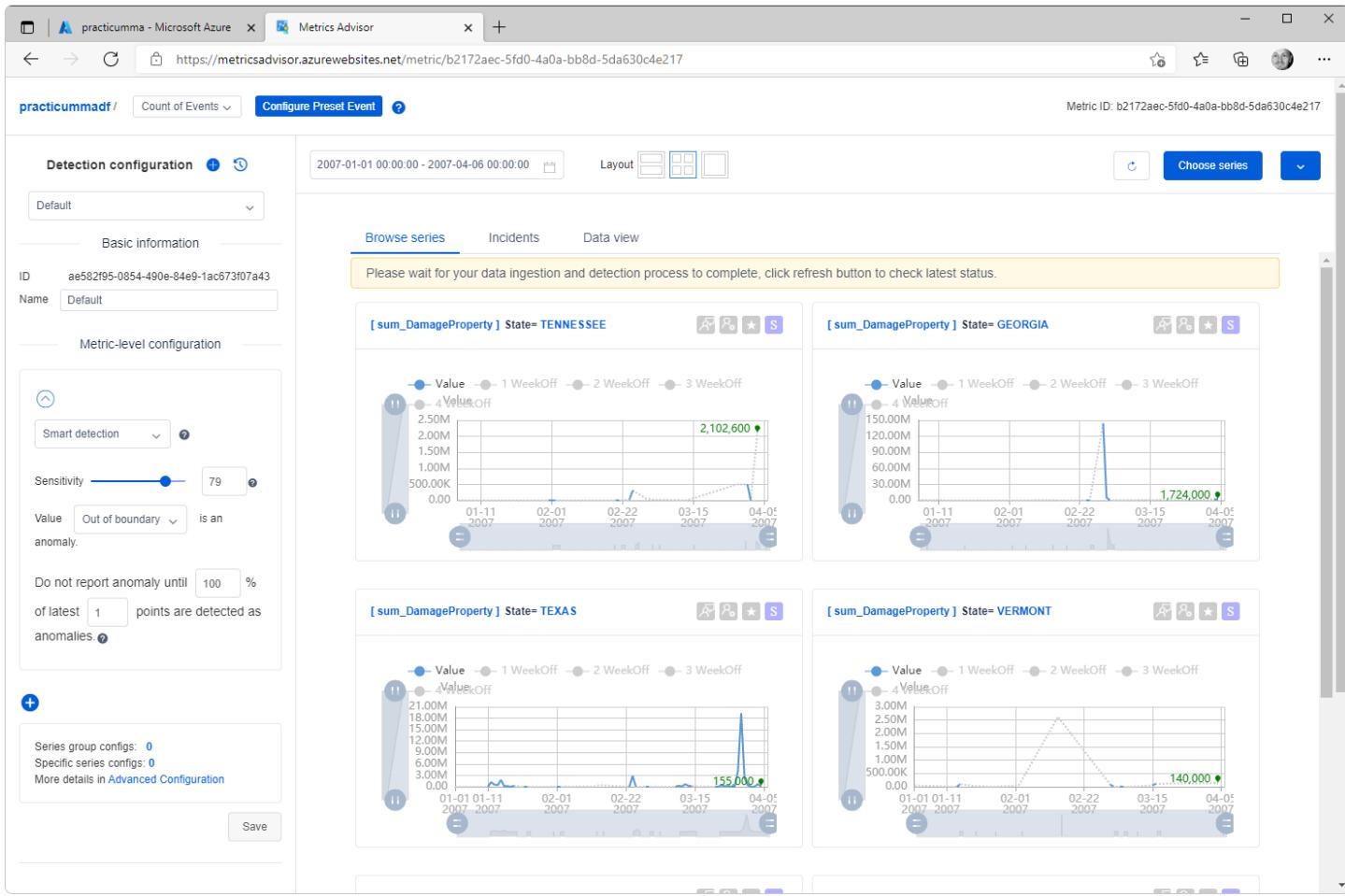
You will receive a “**Congratulations...**” message.

Click the “**Visit data feed...**” button.

A screenshot of a Microsoft Edge browser window showing the 'practicummadf' data feed page. The left sidebar has a blue navigation bar with options like 'Onboarding', 'Add data feed', 'Monitor & Diagnostic', 'Data feeds' (which is selected), 'Incident hub', 'Metrics graph', 'Settings', and 'API keys'. The main content area shows the data feed name 'practicummadf'. It includes an 'Ingestion Progress' bar at 100% completion. Below this, there's a 'Metrics' section with a note: 'It might take 5-8 minutes to see all the metrics here after the data feed is created.' A table lists one metric: 'Count of Events' with ID b2172aec-5fd0-4a0a-bb8d-5da630c4e217, series count 0, and start time 2007-01-01T00:00:00Z. There's also a link to add a new metric. To the right, there's a 'Data feed Information' panel with details: Name: practicummadf, ID: dbf98b41-f6f3-4fcc-8fd3-328b6be1a311, Created time: 2021-09-21T15:06:28Z, Dimensions: State. Buttons for 'Backfill', 'Refresh Progress' (which is blue), 'Delete', 'Export', and 'Start a tour' are at the top right.

Monitor progress on the data feed page.

Click into the Metric Name link to see analysis.



Data Application

Objective(s) in this section describe methods for using data assets in user-facing applications, workflows, and artificial intelligence.

Objective: SQL + Power Apps + AI Builder | Capture Data

Use Case: Company XYZ wants to produce a customer-facing app that can be used to capture images and metadata for an insurance claim {e.g., a vehicle with damage to a headlight}.

Follow the instructions in this section to **build an application that employs artificial intelligence with minimal code**.

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Azure SQL
- PowerApps

Create Target Table

Navigate to your instance of SQL.

The screenshot shows the Microsoft Azure Query editor (preview) interface. On the left, a sidebar lists various database management options like Overview, Activity log, Tags, Diagnose and solve problems, Quick start, and Query editor (preview). The Query editor (preview) option is selected. The main area displays a large "SQL" logo and the text "Welcome to SQL Database Query Editor". It provides two authentication methods: "SQL server authentication" (selected) and "Active Directory authentication". Under SQL server authentication, fields for "Login" (rchapler) and "Password" (redacted) are shown. A red error message states: "Cannot open server 'practicumsds' requested by the login. Client with IP address [REDACTED] is not allowed to access the server. To enable access, use the Azure Portal or run sp_set_firewall_rule on the master database to create a firewall rule for this IP address or address range. It may take up to five minutes for this change to take effect." Below the error message is a link to "Whitelist IP [REDACTED] on server practicumsds". At the bottom right of the authentication section is an "OK" button.

Navigate to “Query editor...” and login. Whitelist your IP address as appropriate.

The screenshot shows the Microsoft Azure Query editor (preview) interface. On the left, there's a sidebar with navigation links like Overview, Activity log, Tags, Diagnose and solve problems, Quick start, and Query editor (preview). The main area has a search bar and tabs for Login, New Query, Open query, and Feedback. A message box says "Showing limited object explorer here. For full capability please open SSDT." Below it is a tree view of objects: Tables (dbo.myTable selected), Views, and Stored Procedures. To the right is a "Query 1" pane with the following T-SQL code:

```
1 CREATE TABLE dbo.myTable( Id INT NOT NULL IDENTITY(1,1) PRIMARY KEY, Name VARCHAR(64), Picture IMAGE )
```

The "Results" tab shows the message "Query succeeded: Affected rows: 0". At the bottom, a yellow bar says "Query succeeded | 0s".

Execute the following T-SQL:

```
CREATE TABLE dbo.myTable( Id INT NOT NULL IDENTITY(1,1) PRIMARY KEY, Name VARCHAR(64), Picture IMAGE )
```

Create Canvas App

Navigate to Power Apps (<https://make.preview.Power Apps.com/>).

Click on **Apps** in the left-hand navigation.

At the top of the “**Apps**” page, click “**+ New app**” and select **Canvas** from the drop-down.

The screenshot shows the "Canvas app from blank" creation dialog in the Power Apps environment. The left sidebar lists Home, Learn, Apps (selected), Create, Data, Tables, Choices, Dataflows, Azure Synapse Link, Connections, Custom Connectors, and Gateways. The main area has a title "Canvas app from blank" with a preview window showing a white canvas with a pencil icon. To the right, there are fields for "App name *": "rchapler", "Format": "Phone" (selected), and a "Type" list showing multiple entries of "Canvas". At the bottom are "Create" and "Cancel" buttons.

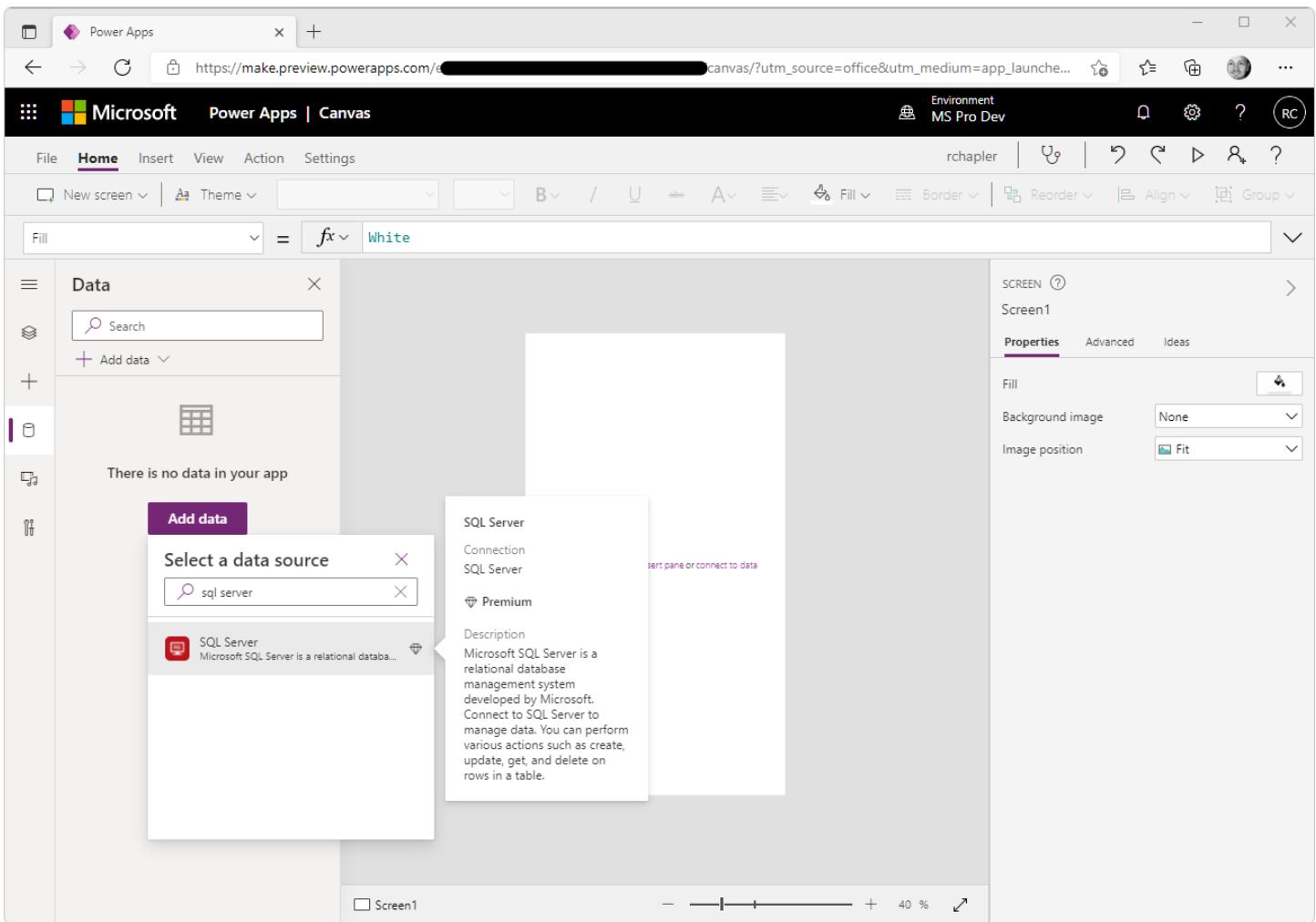
On the “**Canvas app from blank**” pop-up, enter the following items:

App Name	Enter a value that is meaningful and aligned with standards
Format	Select the Phone radio button

Click the **Create** button.

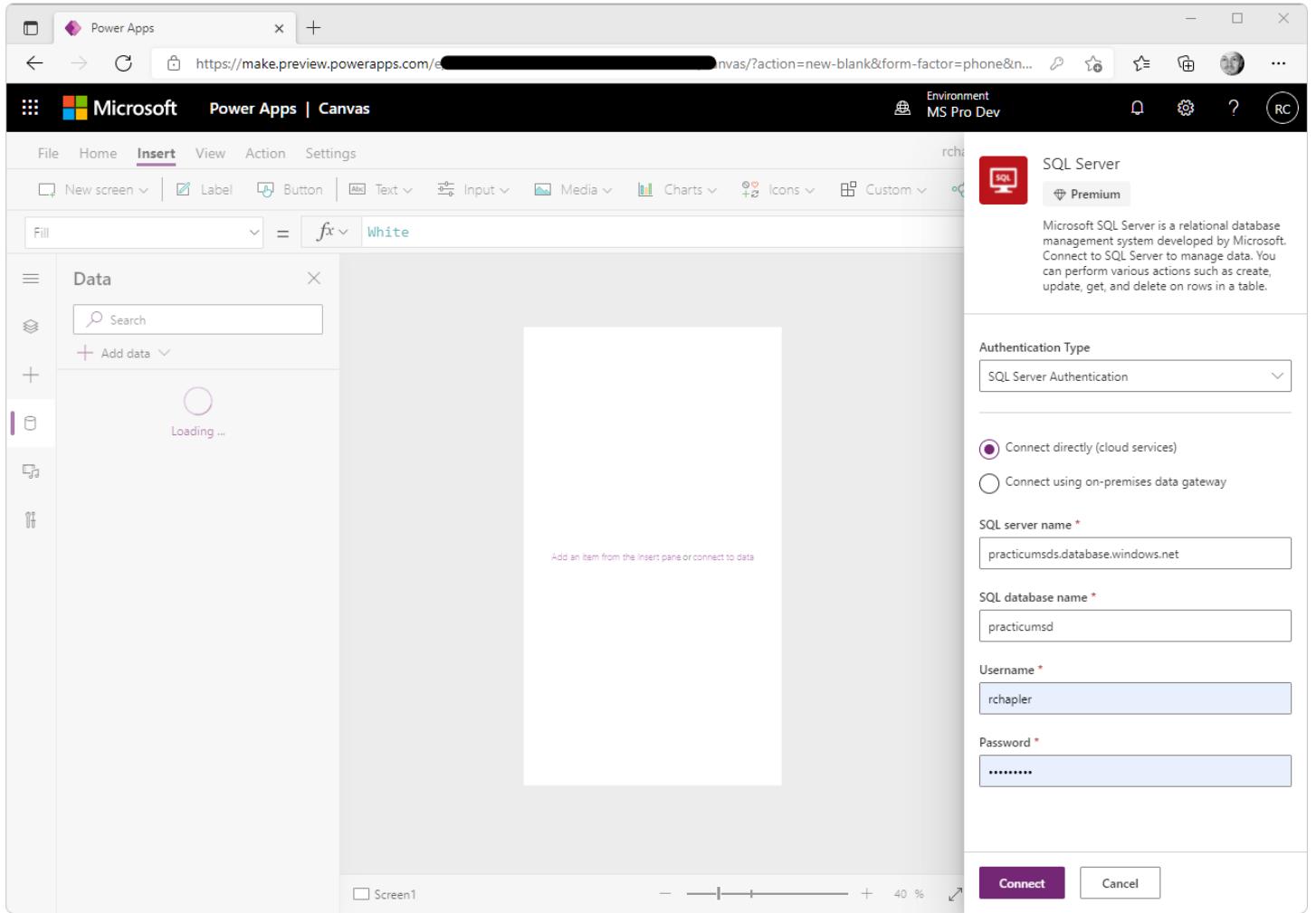
Add Data

Click on **Data** in the left-hand navigation. Click on the “Add data” button.



In “Select a data source”, enter “sql server” in the search input and select “SQL Server” from the results.

Click “+ Add a connection” on the resulting drop-down.



On the “SQL Server” pop-out, enter the following items:

Authentication Type

SQL Server Authentication

SQL Server Name

Enter the values employed during database server creation ([Appendix | SQL](#))

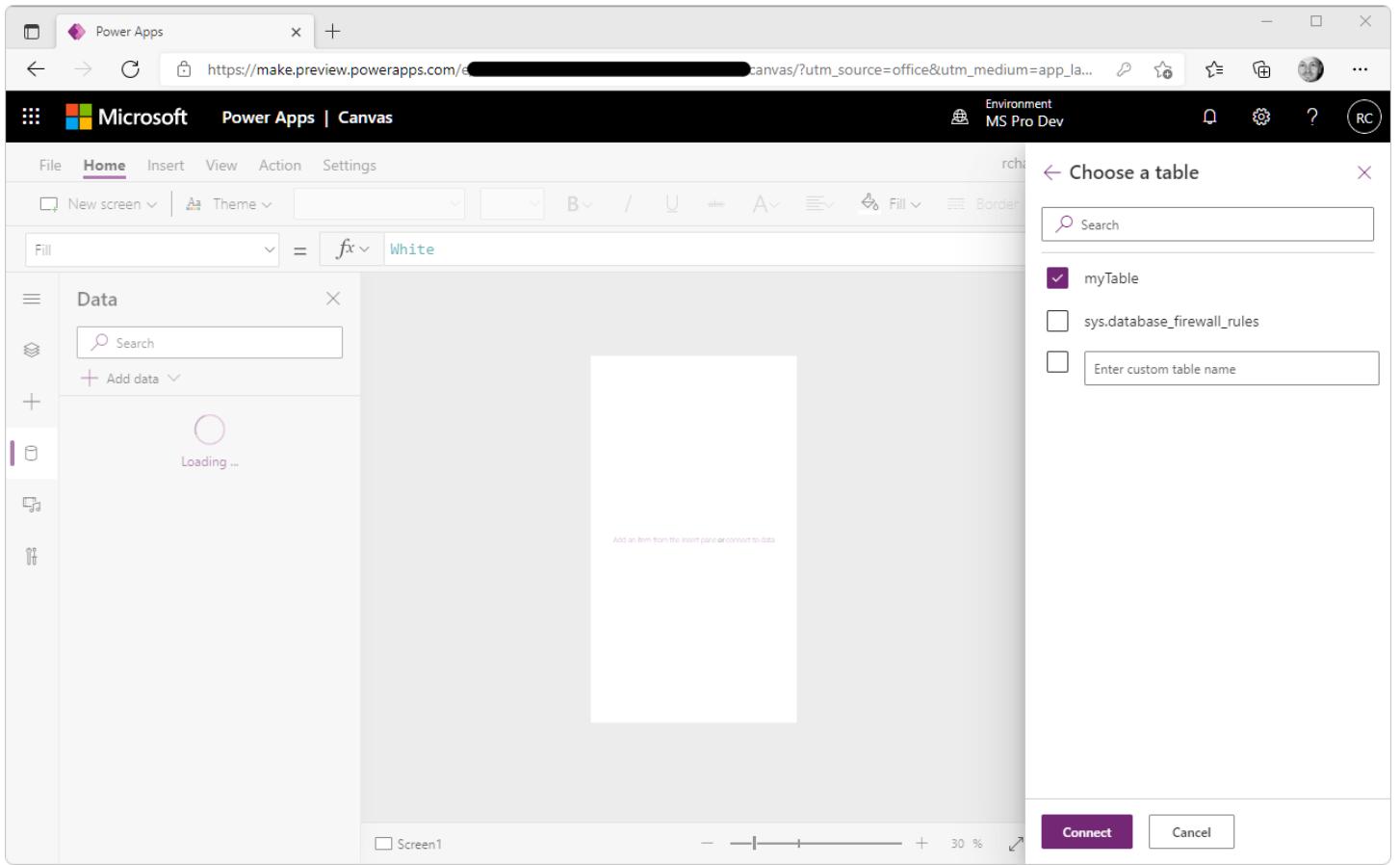
SQL Database Name

Username

Password

Click the **Connect** button.

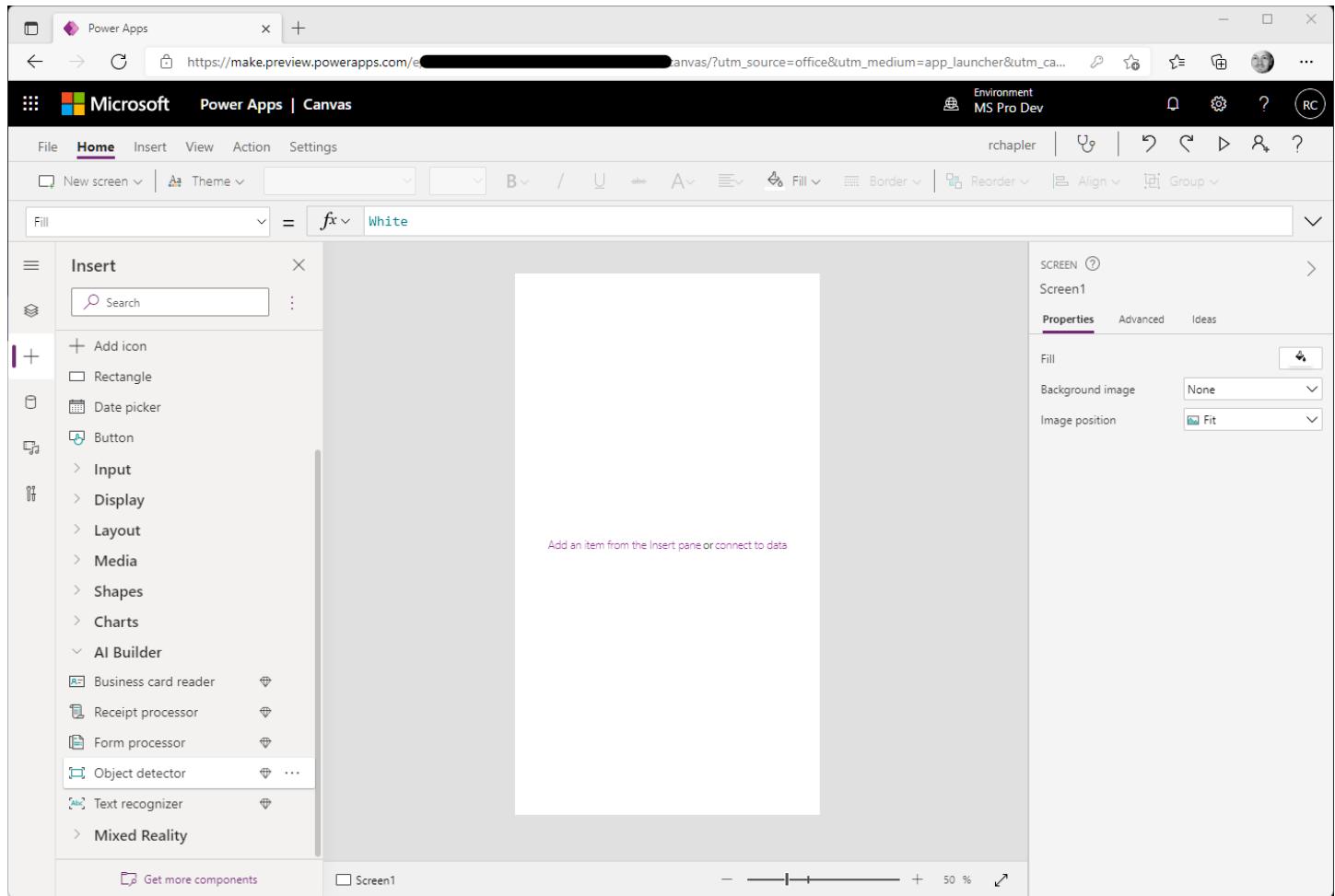
Note: Resolve error messages “We weren’t able to add this connection...” and “Client with IP address #.#.#.# is not allowed...”, by adding your IP address to the Azure SQL Server firewall.



On the “Choose a table” pop-out, check the table created in [Create Target Table](#).

Click the **Connect** button.

Insert Object Detector



Click the “Add an item from the Insert pane” link.

On the resulting, left-hand **Insert** menu, expand “**AI Builder**” and select “**Object detector**”.

On the resulting, right-hand pop-out, click “**+ New model**”.

Create AI Model

The screenshot shows the Microsoft Power Apps | AI Builder interface. On the left, there is a navigation sidebar with options like Home, Learn, Apps, Create, Data, Flows, Chatbots, AI Builder (which is selected), Build, Models, and Solutions. The main content area has a heading "Enhance your business with AI" and a sub-section "Refine a model for your business needs". It displays five cards: "Category Classification", "Entity Extraction", "Form Processing", "Object Detection", and "Prediction". The "Object Detection" card is highlighted with a purple border.

On the new “Power Apps | AI Builder”, “Enhance your business with AI” page, click the “Object Detection” button.

The screenshot shows the "Object Detection" setup page. The left sidebar is identical to the previous one. The main area has a title "Object Detection" and a "Premium" badge. Below it are tabs for "Get started", "Examples", and "Best practices", with "Get started" being active. A large central image shows a document with an upward arrow and a camera icon, with a pink circle highlighting the document. Below this is a section titled "Step 1: Teach the model about your objects" with the sub-instruction "Select or define object names and upload 15+ images of each.". To the right, there is a form field "Name this AI model *" containing "Car Parts", a section "What you'll need" with a note "15+ images of each object", and two buttons at the bottom: "Create" and "Cancel".

Enter a value that is meaningful and aligned with standards and then click the **Create** button.

Select your model's domain

When models focus on specific types of objects, they can be more accurate. If you don't see the right option, select **Common objects**. [Learn more](#)

Common objects

Objects on retail shelves

Brand logos

Combine

Combine

Combine

Next

Car Parts | Save and close

?

Quick tips

Get help or send feedback

Get the answers you need, or tell us about your experiences.

Get help

On the “Select your model’s domain” page, click the “Common objects” and then the “Next” button.

Select from database

Choose objects for your model to detect

You can add them manually or select from your database. [Learn more](#)

Object names

Headlights

+ Add new object

Back

Next

1 object name selected

Car Parts | Save and close

?

Quick tips

Select from database instead

Save time and use objects in Dataverse, rather than creating them manually.

Switch to database

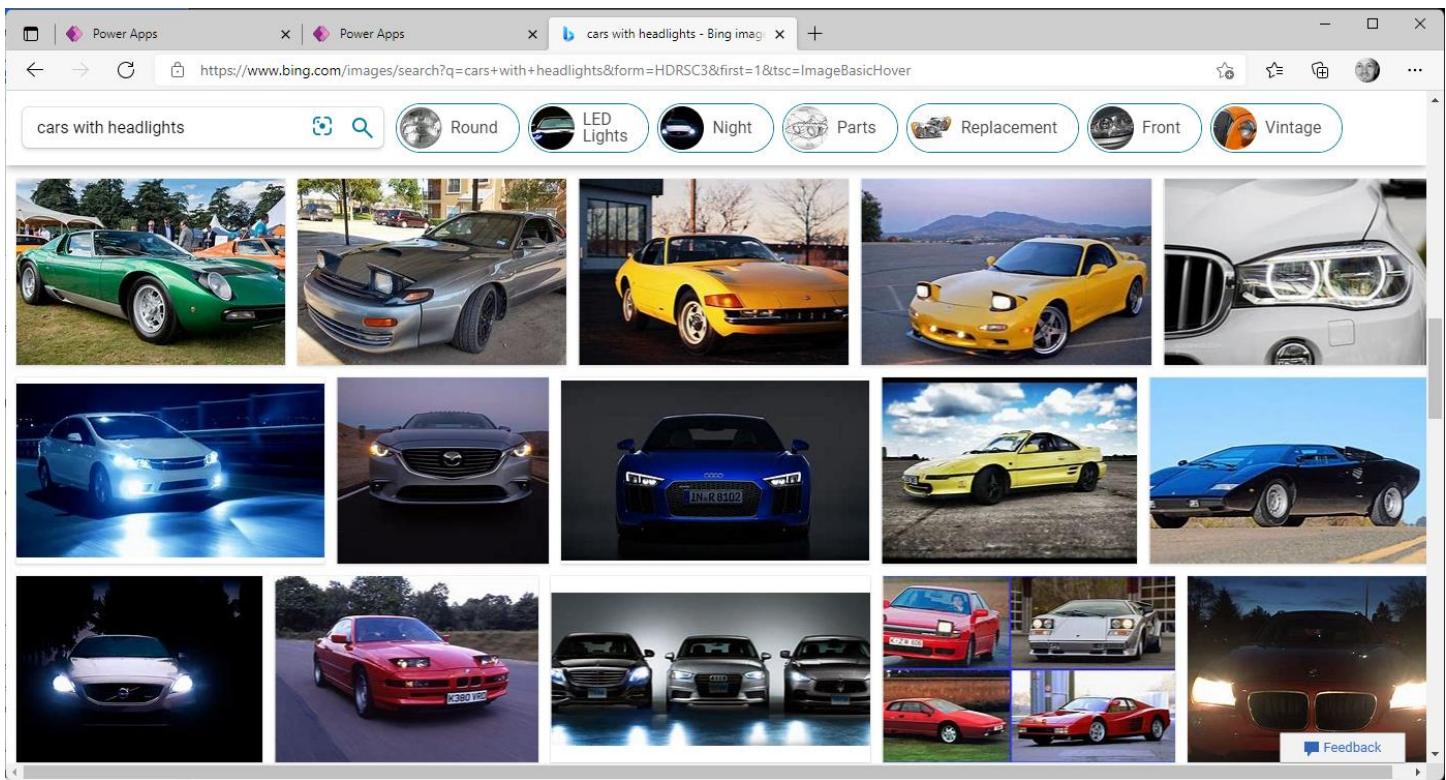
Get help or send feedback

Get the answers you need, or tell us about your experiences.

Get help

On the “Choose objects for your model to detect” page, click “+ Add new object”, enter an object name and then click the “Next” button.

On the “Add example images for your objects” page, we will add image files that will be used to train the AI model. For this exercise, we can pull images of cars (with headlights) from a search engine.



For each of the required 15 images, download an image file to a temporary folder on the Windows desktop.

Power Apps

Power Apps

cars with headlights - Bing images

https://make.preview.powerapps.com/e...ibuilder/models/bbac270f-bec8-491a-b4cd-66663b31a4b2/editor

Microsoft Power Apps | AI Builder

Select domain
Common objects

Choose objects
1 object name selected

Add images

Tag images

Model summary

+ Add images

Add example images for your objects

Find images containing your objects

You need to add at least 15 images for each object you want to detect.

Add images

Back

Next

Environment MS Pro Dev

Car Parts | Save and close

Quick tips

Add lots of images

Start with 15+ images per object, but 50+ is better. It's OK to tag 2+ objects in the same image.

Learn more

Image requirements

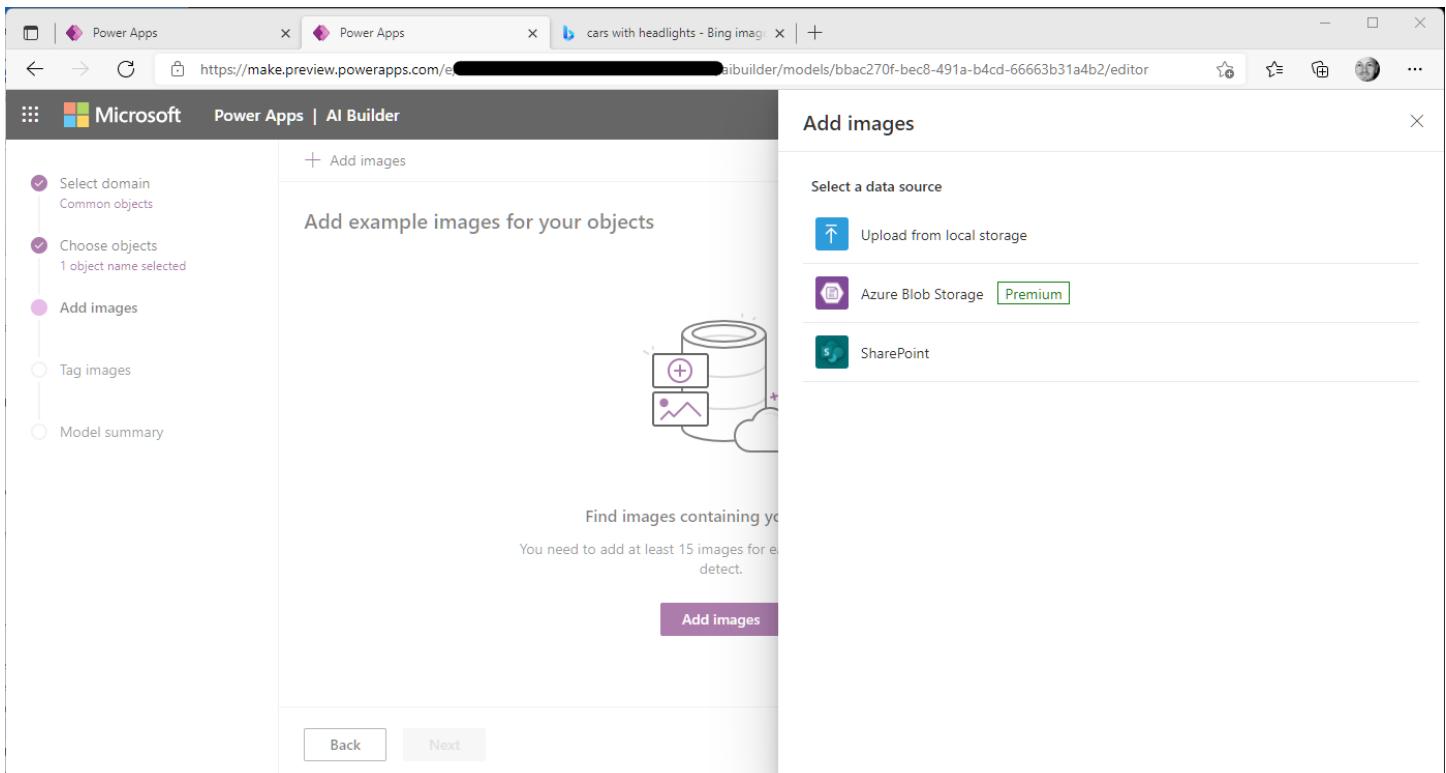
Use .jpg, .png, .bmp image - up to 6 MB each.

Get help or send feedback

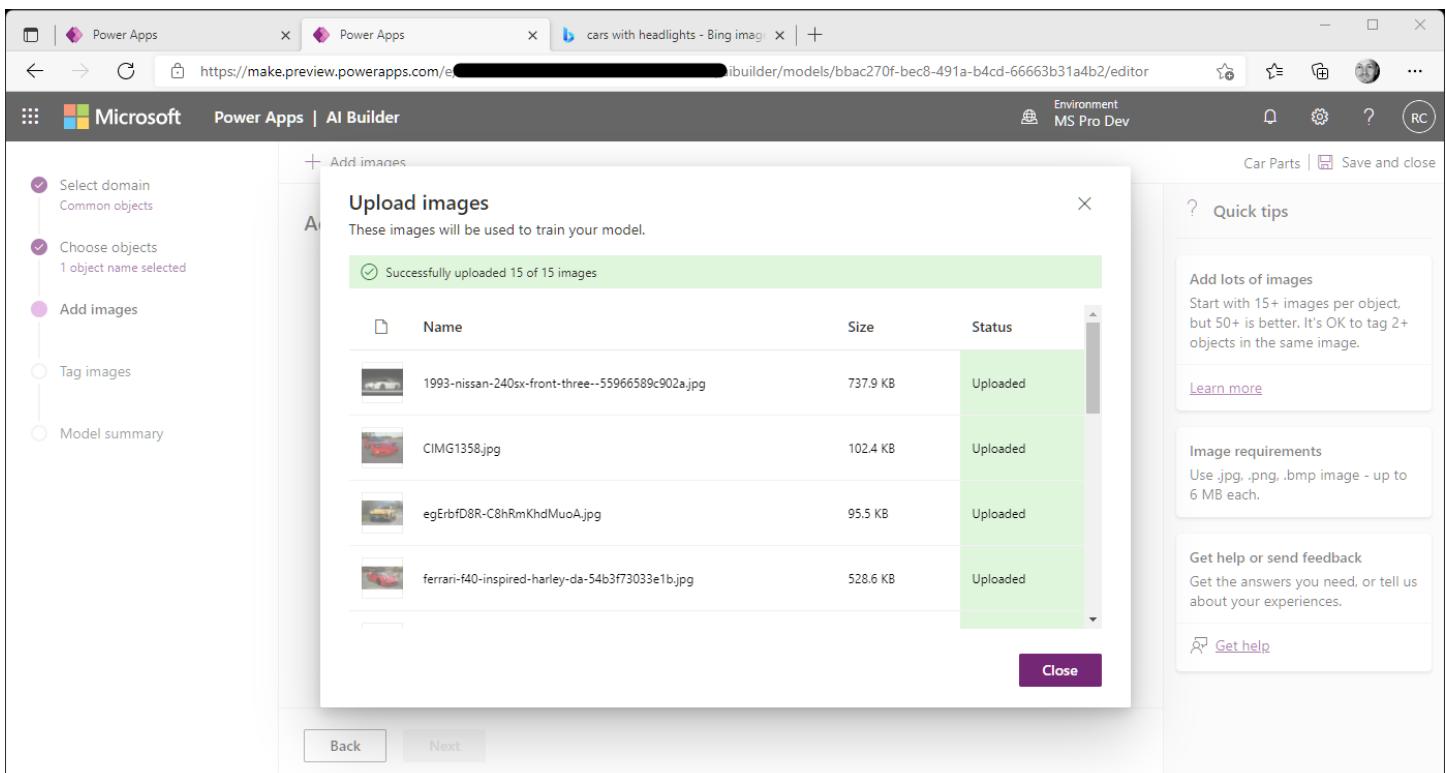
Get the answers you need, or tell us about your experiences.

Get help

On the “Add example images for your objects” page, click the “Add images” button.



On the resulting “Add images” pop-out, click “Upload from local storage”. Select the Example Images downloaded to your desktop.



Review the items in the “Upload images” pop-up, then click the “Upload 15 images” button. Click Close.

Screenshot of the Microsoft Power Apps AI Builder interface showing the "Add images" step. The left sidebar shows completed steps: "Select domain", "Choose objects", and "Add images" (15 images added). The main area displays a table of images from "My device" with 15 images. A sidebar on the right provides tips for adding images and details about image requirements.

Data source	Number of images
My device	15 images

Quick tips:
Add lots of images: Start with 15+ images per object, but 50+ is better. It's OK to tag 2+ objects in the same image.
Learn more
Image requirements: Use .jpg, .png, .bmp image - up to 6 MB each.

Back on the “Add example images for your desktop” page, confirm “Number of images” shows 15 or more images, then click **Next**.

Screenshot of the Microsoft Power Apps AI Builder interface showing the “Tag the objects in your images” step. The left sidebar shows completed steps: “Select domain”, “Choose objects”, “Add images” (15 images added), and “Tag images”. The main area displays a grid of 15 car images with a “NEW” tag. A sidebar on the right provides tagging requirements, a link to add more images, and help options.

You will need to tag at least 15 images per object to continue.

Tag the objects in your images

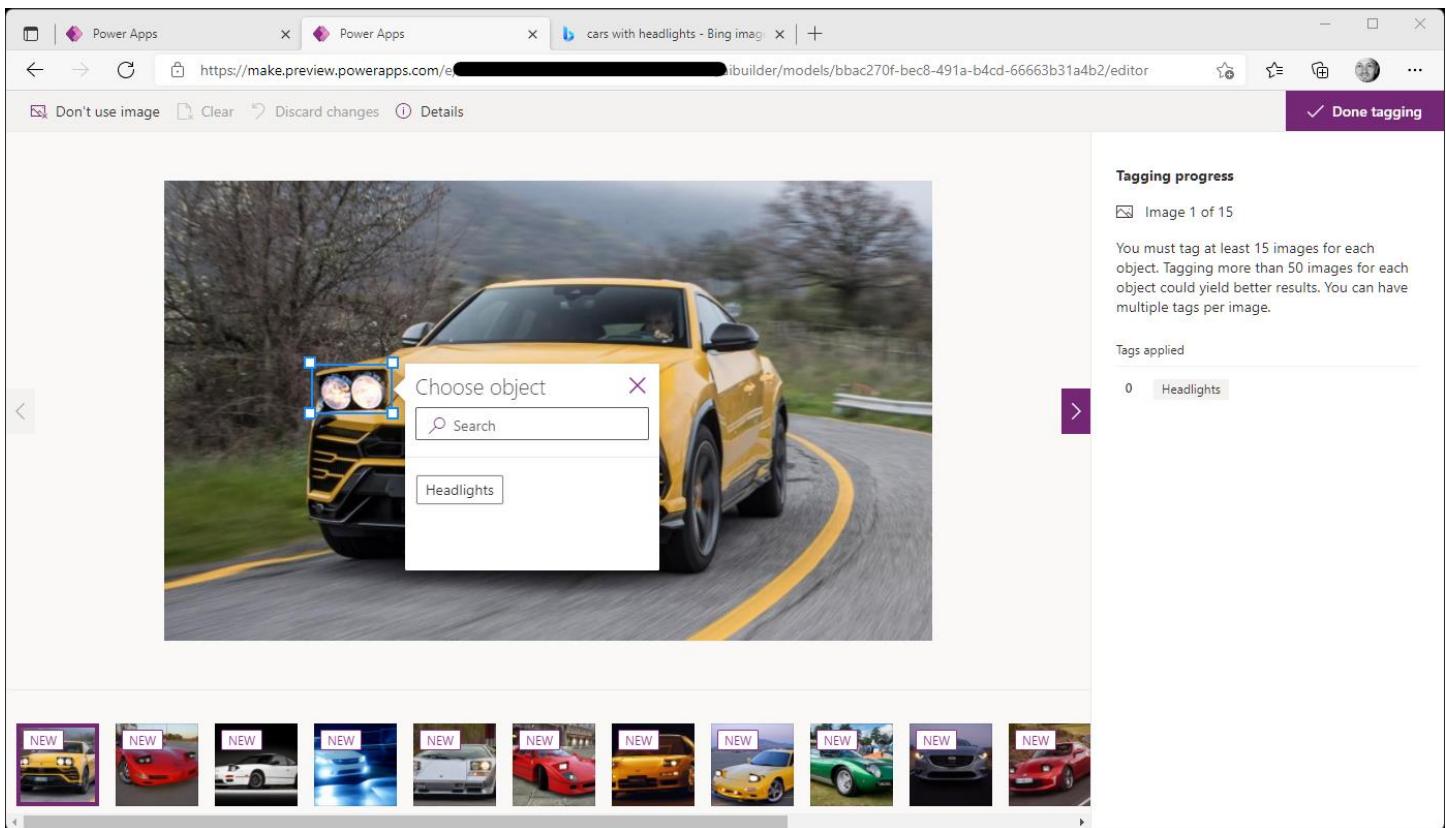
All

Quick tips:
Tagging requirements: You must tag at least 15 images for each object. Tagging more than 50 images for each object could yield better results. You can have multiple tags per image.
0 Headlights

Want to add more images? Adding more images may increase model performance.
+ Add images

Get help or send feedback
Get the answers you need, or tell us about your experiences.
Get help

Select the first image on the “Tag the objects in your images” page.



Drag a rectangle around headlights in the image, then click the Headlights button on the “Choose object” pop-up.
Repeat for all taggable items, click > to move to through images, and then click the “Done tagging” button in the upper-right.

Click the **Next** button.

Screenshot of the Microsoft Power Apps AI Builder interface showing the Model summary step. On the left, a sidebar lists completed steps: Select domain (Common objects), Choose objects (1 object name selected), Add images (15 images added), Tag images (15 images tagged), and Model summary. The main area displays the Model summary details, including Overview (Model type: Object Detection, Owner: Rich Chapler, Object type: Common objects), Image sources (Data source: My device, Number of images: 15 images), and Information to extract (Object: Headlights, Tags: 15). A 'Train' button is visible at the bottom. A 'Quick tips' panel on the right provides information about training, what's next, and help.

Review the “Model Summary” and then click the **Train** button.

Screenshot of the Microsoft Power Apps AI Builder interface showing the Model summary step after clicking the Train button. The status has changed to "Your model is training". A message indicates that training may take a while. A "Go to models" button is visible. The sidebar and right-hand panels remain the same as in the previous screenshot.

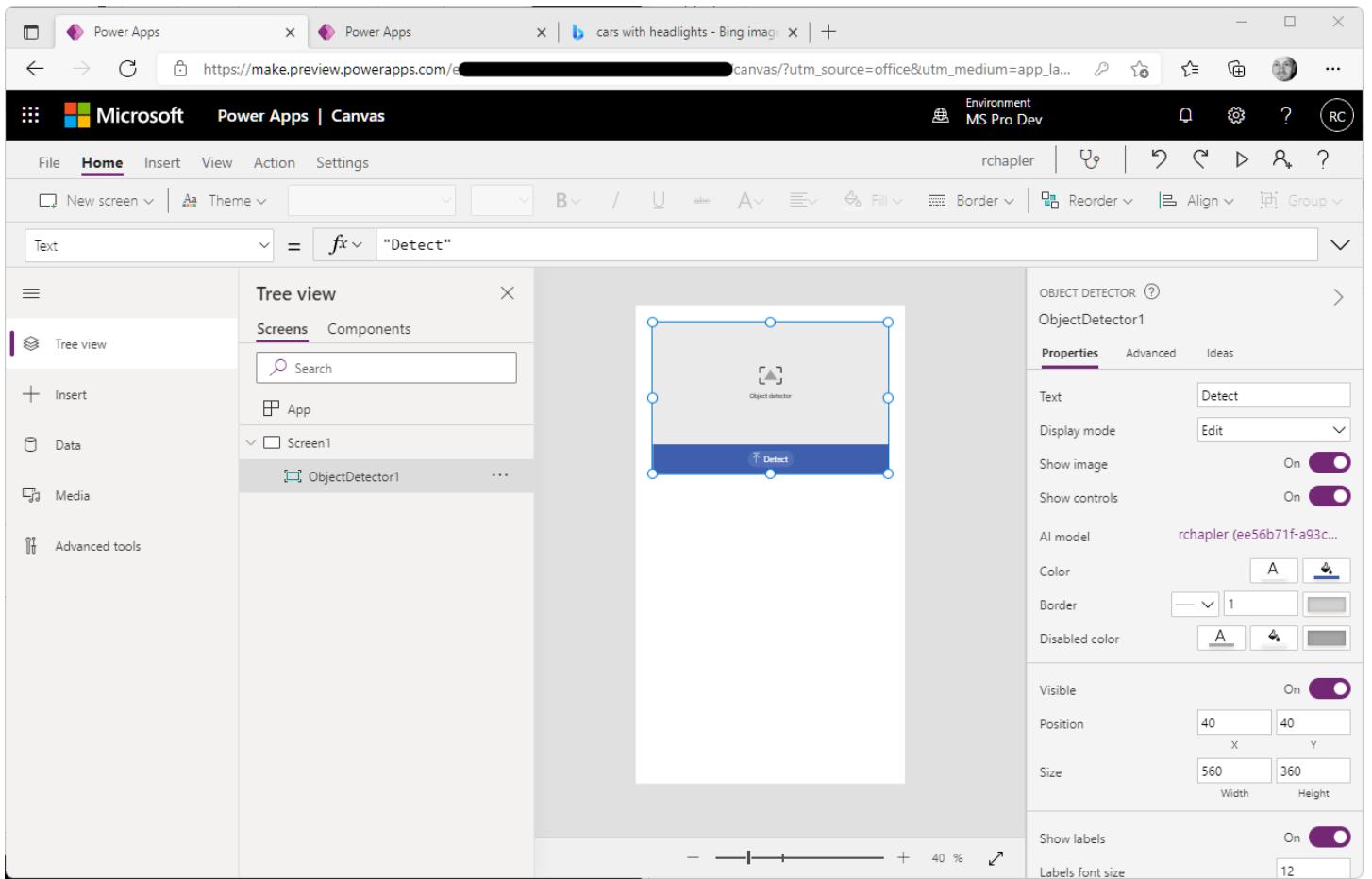
Click the “Go to models” button. After Status changes to **Trained**, navigate to the model.

The screenshot shows the Microsoft Power Apps AI Builder interface. On the left, a sidebar menu includes Home, Learn, Apps, Create, Data, Flows, Chatbots, AI Builder, Build, Models (which is selected), and Solutions. The main content area displays a 'Models > Car Parts' page for 'Object Detection'. A large circular progress bar indicates a 'Performance' of 46%. Below it, a callout box suggests improving the model by providing more training data. Two buttons at the bottom are 'Publish' (in purple) and 'Quick test'. To the right, there's a 'Details' section with 'Training date' (10:29:01 AM), 'Object type' (Common objects), and 'Number of objects' (1). A 'See more' link is also present. On the far right, a panel titled 'How your model is used' lists 'Power Automate' and 'Power Apps', both noting that the model isn't published.

In the snip above, we see that **Performance** {i.e., model performance expectation %} of 46 for the example images included in training.
In the future, you can improve model quality by including more images and re-training the model.

Click the **Publish** button.

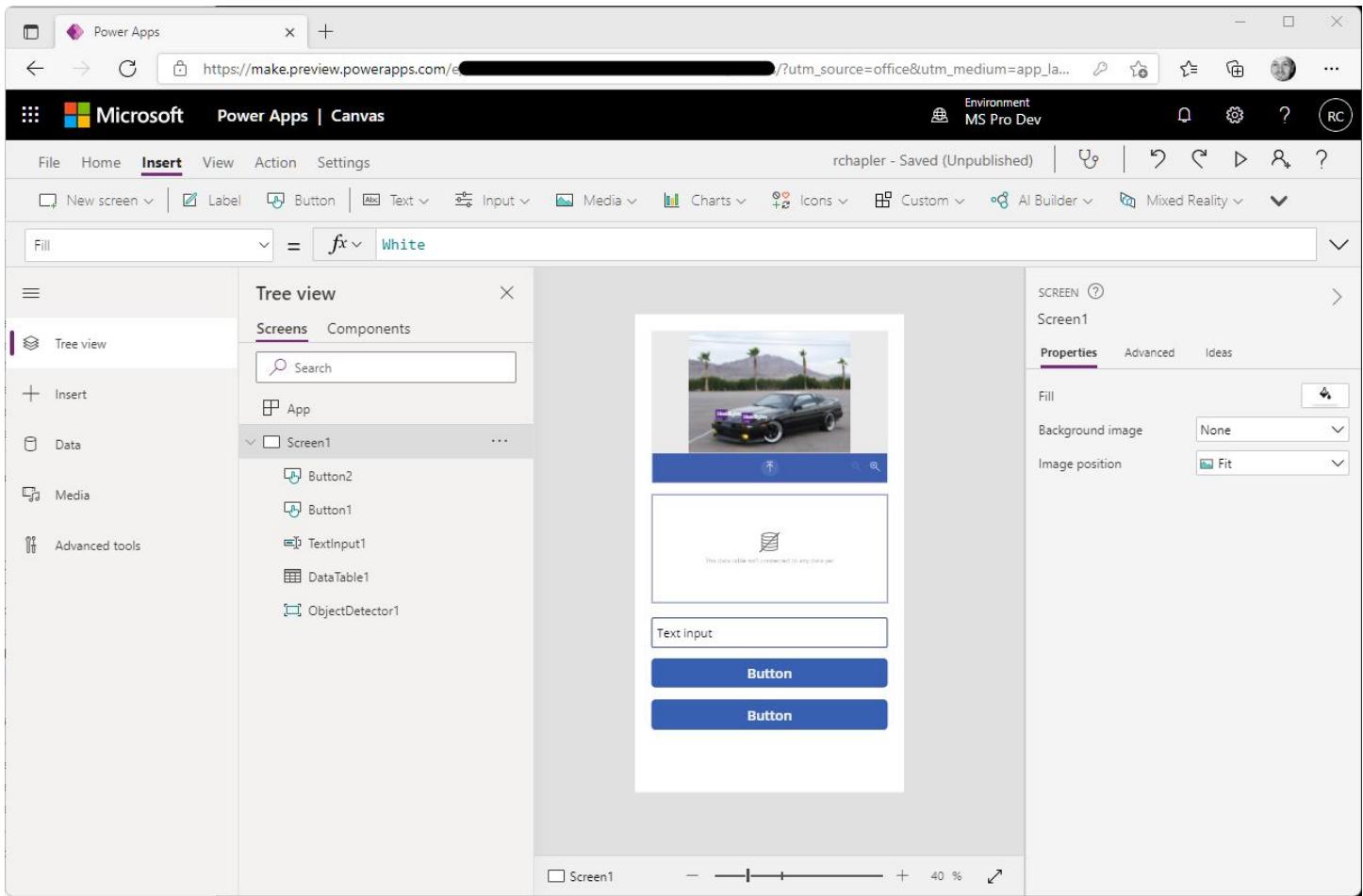
Return to Power Apps and complete the ObjectDetector1 component by selecting your newly created model.



Note: Consider resizing the component to improve visibility.

Insert / Configure Controls

Return by closing Preview Mode.

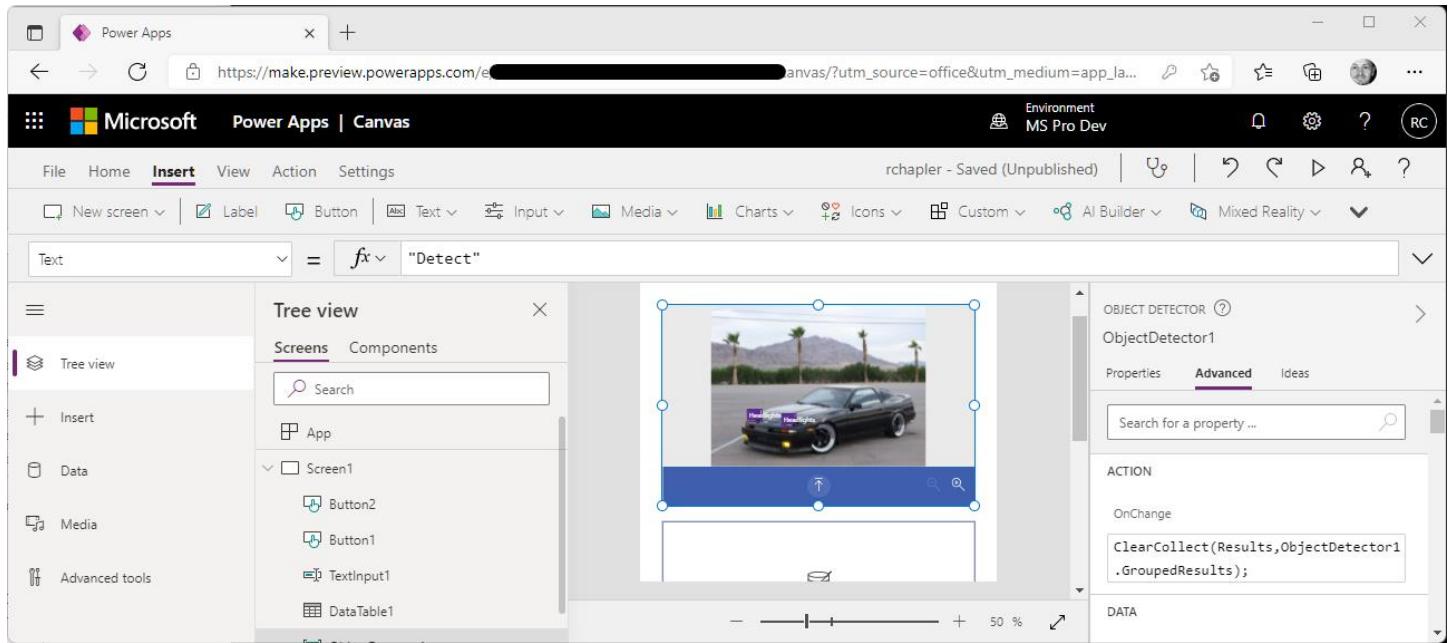


Insert the following controls:

- **Data Table** ... users will use this control to view the results generated by AI Builder
- **Text Input** ... users will use this control to enter comments about the analyzed image
- “Submit” **Button** ... users will use this control to save the image and comment to Azure SQL
- “Reset” **Button** ... users will use this control to clear previously entered values

You might note that I employed formatting {e.g., control width, border size, color, etc.} to enhance usability.

ObjectDetector1

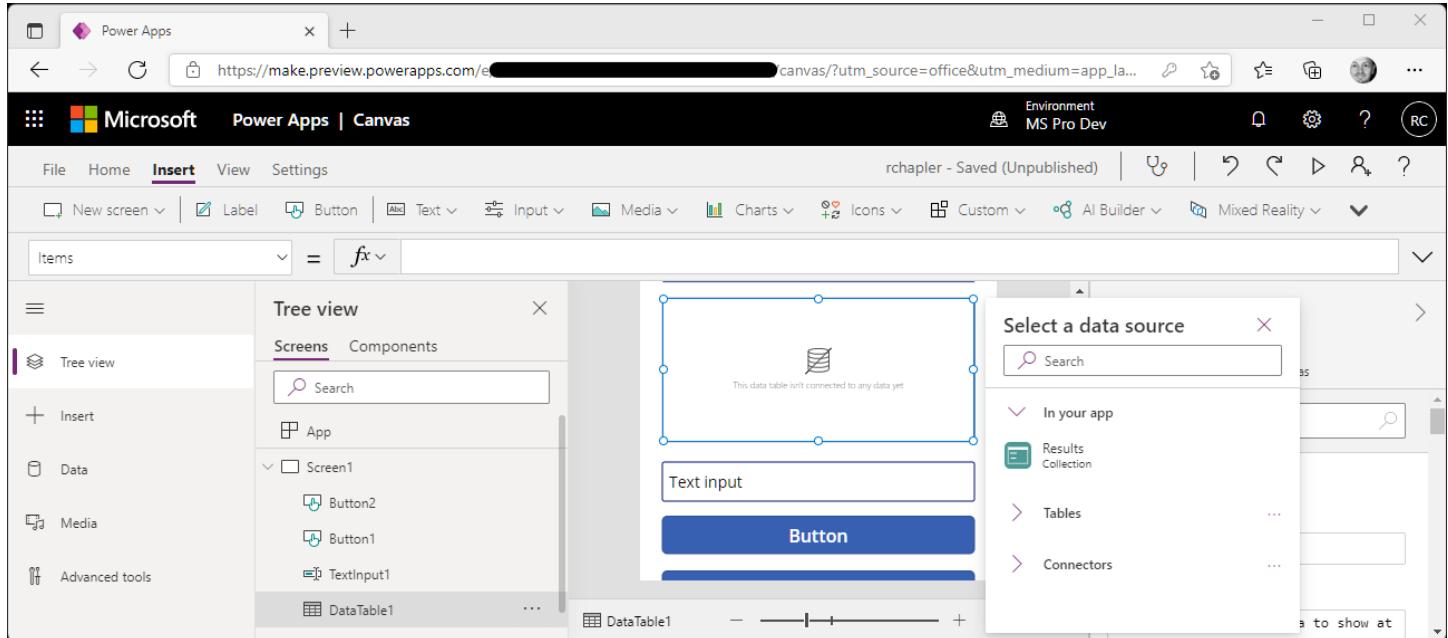


Click on **ObjectDetector1** in the left-hand **Tree View**. On the resulting right-hand pop-out, click the **Advanced** tab.

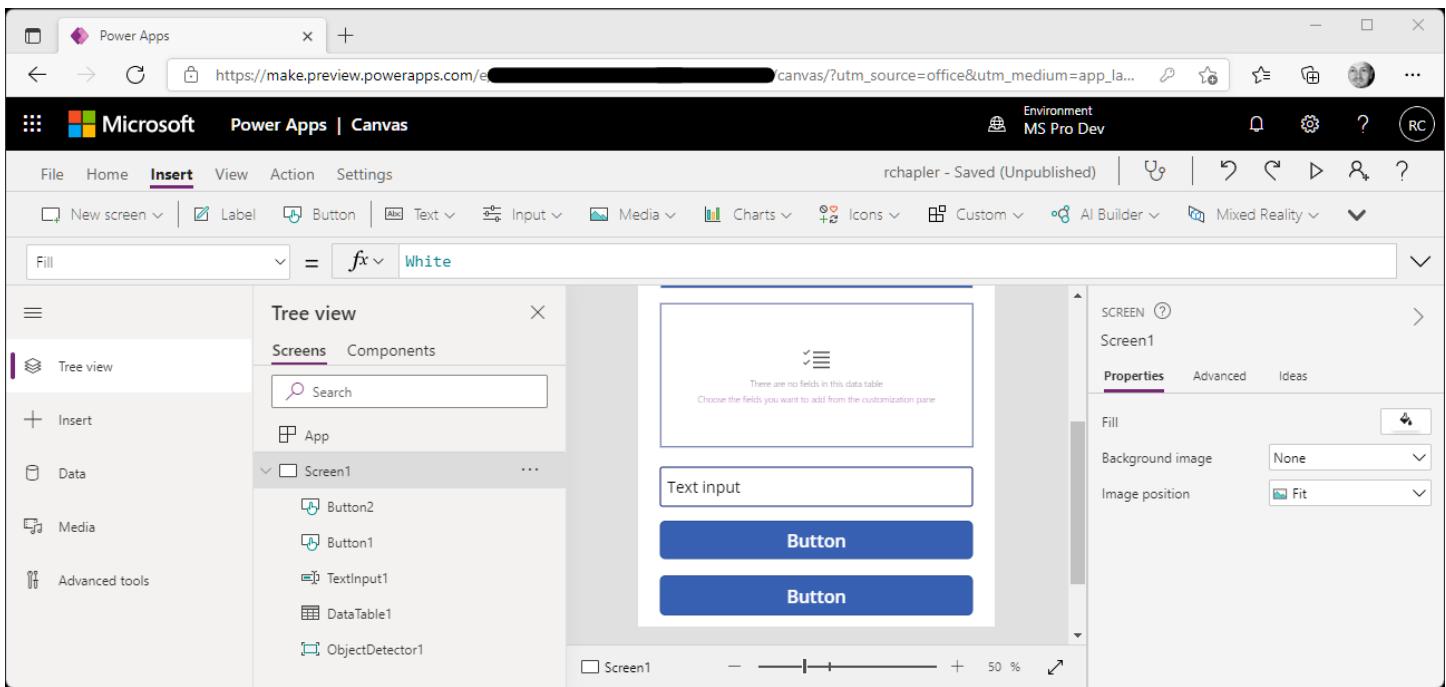
Paste the following logic into **Action | OnChange**:

```
ClearCollect(Results, ObjectDetector1.GroupedResults);
```

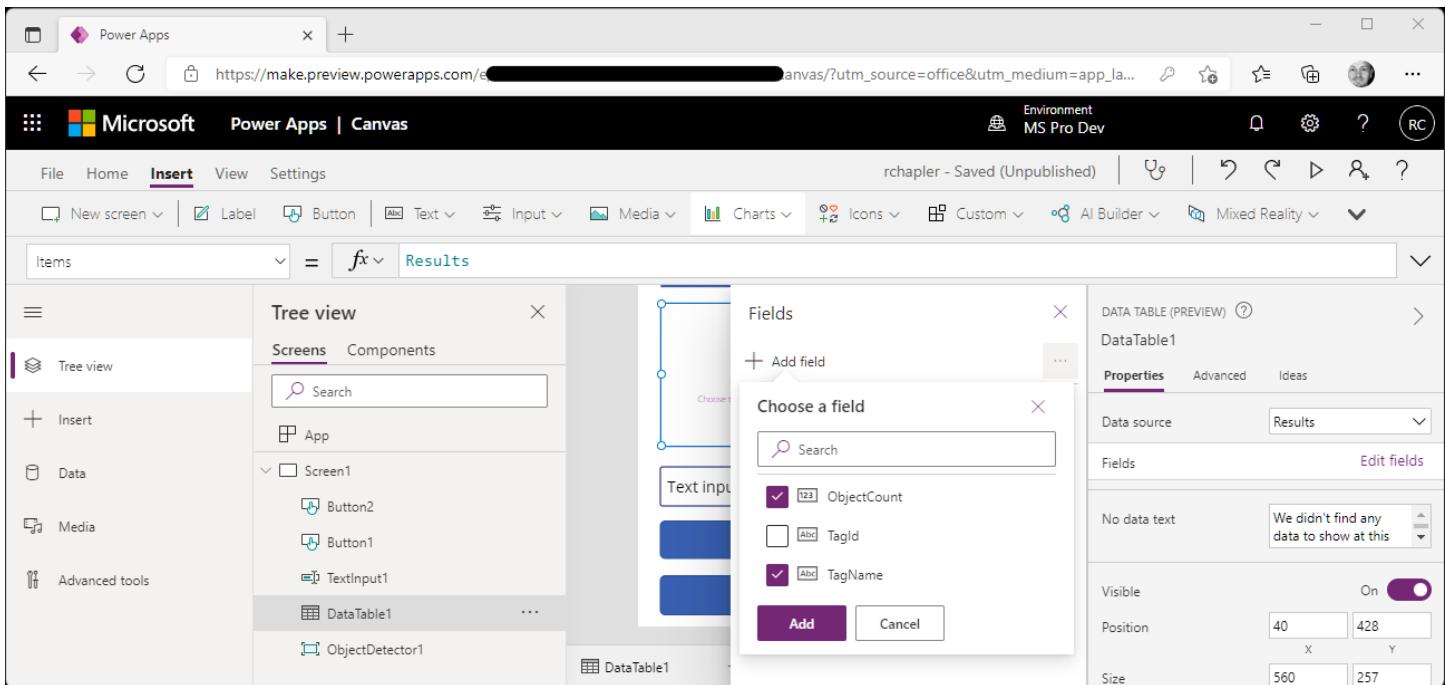
DataTable1



Click on **DataTable1** in the left-hand **Tree View**. On the resulting pop-up, select “**In your app**” | **Results**.



On the updated control, click the “Choose the fields you want to add...” link.



On the resulting **Fields** pop-out, click “+ Add field”.

In the resulting “Choose a field” pop-up, check **TagName** and then **ObjectCount**.

Click the **Add** button.

TextInput1

The screenshot shows the Microsoft Power Apps Canvas editor interface. On the left, there's a tree view of the app structure under 'Screen1'. In the center, a 'Text input' component is selected, displaying a table with one row: 'Headlights' and '2'. To the right, a 'TEXT INPUT' pop-out window is open for 'TextInput1'. The 'Properties' tab is selected, showing the following settings:

Default	Text input
Format	Text
Hint text	"Enter Comments"
Font	Open Sans
Font size	21
Font weight	B Normal
Font style	/ U +
Text alignment	Left Center

Click on **TextInput1** in the left-hand **Tree View**.

On the resulting “**TEXT INPUT**” pop-out, **Advanced** tab, enter the following items:

Default	""
Hint Text	"Enter Comments"

Button1

The screenshot shows the Microsoft Power Apps Canvas editor interface. On the left, there's a tree view of the app structure under 'Screen1'. In the center, a 'Button1' component is selected, displaying a button labeled 'Submit'. To the right, a 'BUTTON' pop-out window is open for 'Button1'. The 'Properties' tab is selected, showing the following settings:

Text	Submit
Display mode	Edit
Visible	On
Position	X: 40 Y: 819
Size	Width: 560 Height: 67
Padding	Top: 5 Bottom: 5

Click on **Button1** in the left-hand **Tree View**.

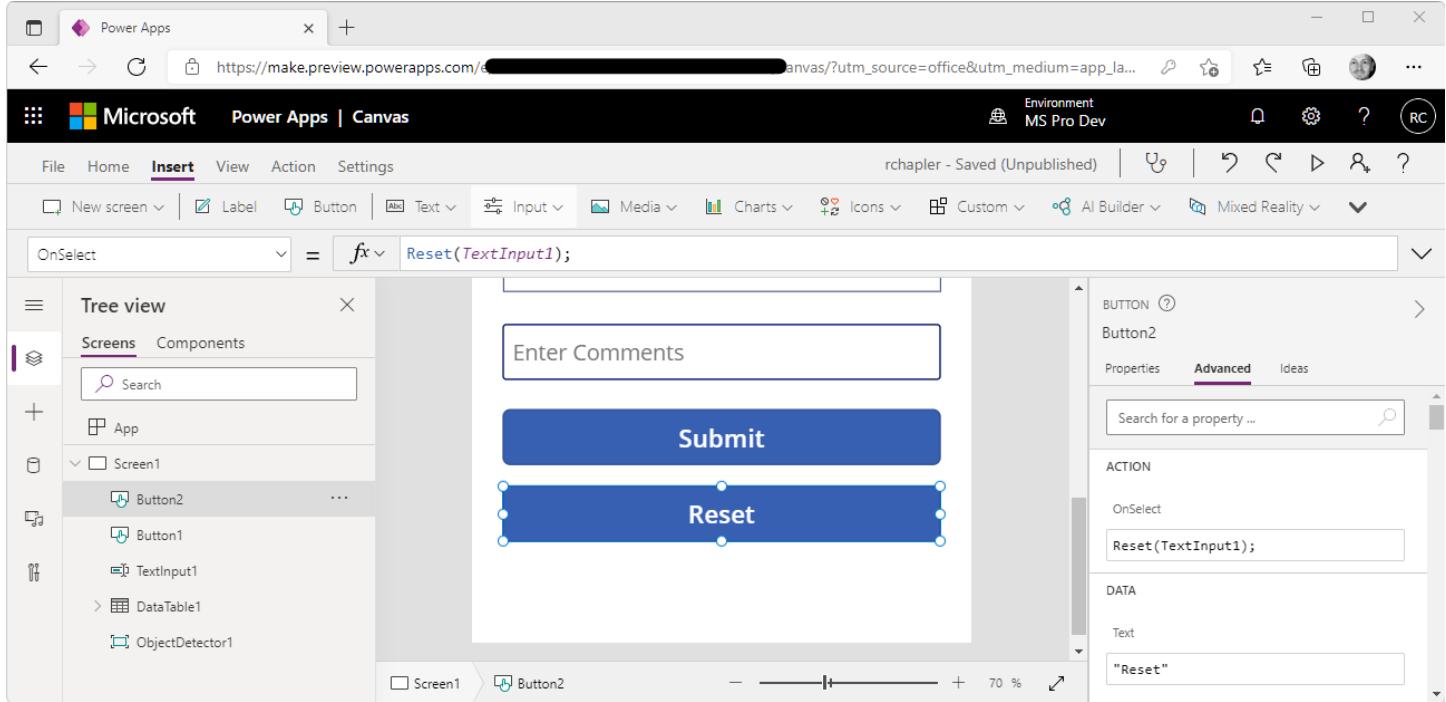
On the resulting right-hand pop-out, click the **Properties** tab, and enter the word “Submit” in the **Text** input.

On the resulting right-hand pop-out, click the **Advanced** tab,

Paste the following logic into **Action | OnSelect**:

```
Patch(myTable,Defaults(myTable),{Name:TextInput1.Text,Picture:ObjectDetector1.OriginalImage})
```

Button2



Click on **Button2** in the left-hand **Tree View**.

On the resulting right-hand pop-out, click the **Properties** tab, and enter the word “Reset” in the **Text** input.

On the resulting right-hand pop-out, click the **Advanced** tab,

Paste the following logic into **Action | OnSelect**:

```
Reset(TextInput1);
```

Confirm Success

Click **File** in the menu bar, then “**Save As**”, confirm the app name, and finally, click the **Save** button in the bottom-right.

Return to the main page, click the “**Preview the app**” button {i.e., Play icon} in the upper-right.

Power Apps

Microsoft Power Apps | Canvas

File Home Insert View Action Settings

rchapter - Saved (Unpublished)

New screen Label Button Text Input Media Charts Icons Custom AI Builder

Fill = fx White

Tree view

Screens Components

Search

App

Screen1

Button2

Button1

TextInput1

DataTable1

ObjectDetector1

Preview the app (F5)

Preview the app to see how it works.

SCREEN

Screen1

Properties Advanced Ideas

Search for a property ...

ACTION

OnVisible

OnHidden

DATA

BackgroundImage

ContentLanguage

DESIGN

Fill

White

ImagePosition

ImagePosition.Fit

Enter Comments

Submit

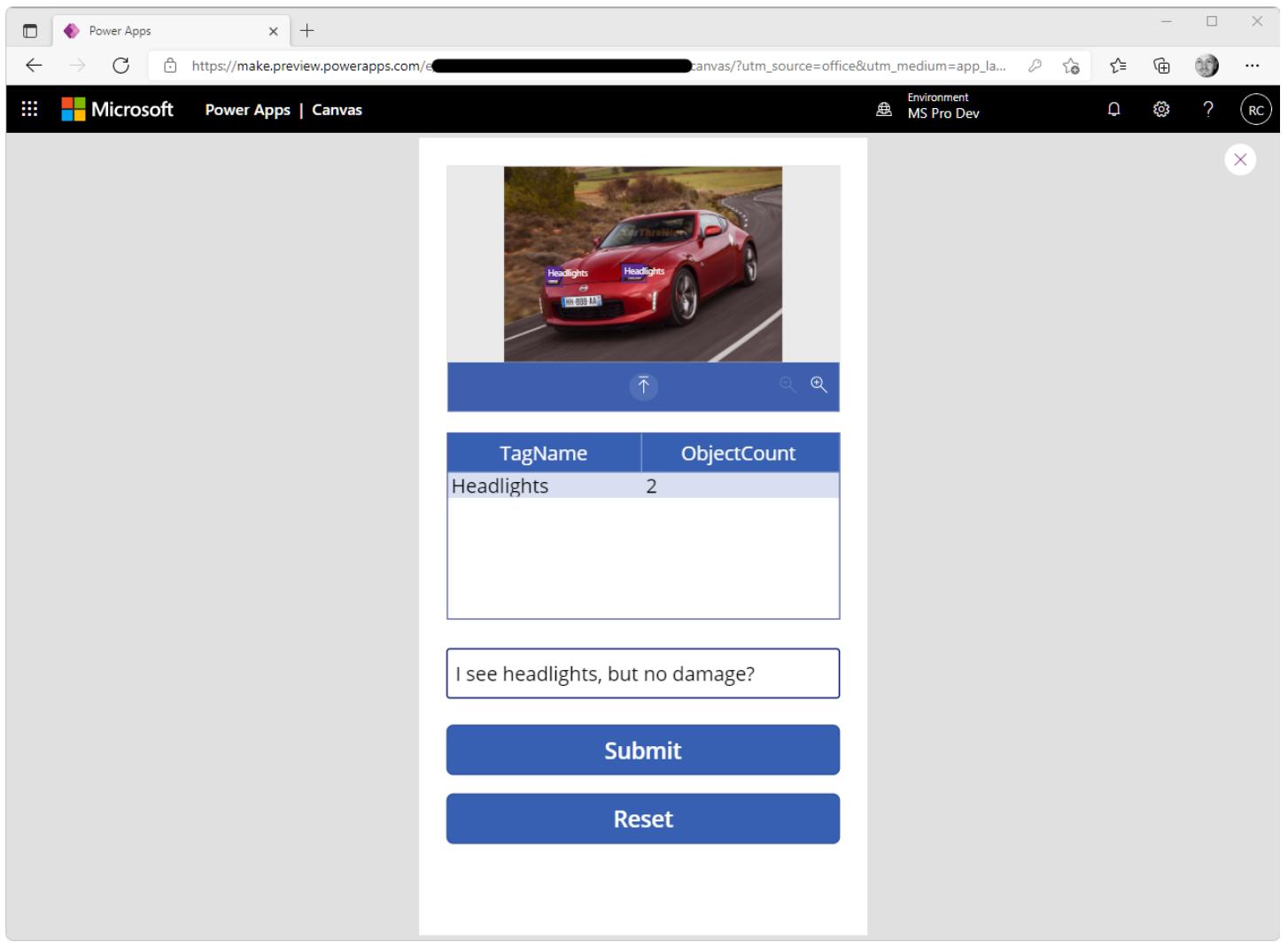
Reset

Screen1

50 %

The screenshot displays the Microsoft Power Apps Canvas editor interface. On the left, a tree view shows the project structure with 'Screens' and 'Components'. Under 'Screens', 'Screen1' is selected, revealing its contents: a red sports car image, a table with one row ('Headlights' and '2'), an input field for comments, and two buttons ('Submit' and 'Reset'). The properties panel on the right is open, specifically the 'Advanced' tab, which contains settings for 'ACTION' (OnVisible, OnHidden), 'DATA' (BackgroundImage, ContentLanguage), and 'DESIGN' (Fill, ImagePosition). A preview window at the top right shows the application running on a mobile device. A tooltip 'Preview the app (F5)' is visible, indicating how to test the application.

Your app will be presented in a phone simulation.



Click the **Detect** button and select an image file to test objection detection.

Review the resulting values in the data table.

Enter a comment in the text input and click the **Submit** button.

The screenshot shows the Microsoft Azure Query editor (preview) interface. On the left, there's a navigation sidebar with links like Overview, Activity log, Tags, Diagnose and solve problems, Quick start, and Query editor (preview). Below that are sections for Power Platform (Power BI, Power Apps, Power Automate), Settings (Compute + storage, Connection strings, Maintenance, Properties, Locks), Data management (Replicas, Sync to other databases), and Integrations (Stream analytics). The main area has tabs for 'Query 1' and 'Query 2'. Under 'Query 2', the status bar says 'Showing limited object explorer here. For full capability please open SSDT.' The object explorer shows 'Tables' with 'dbo.myTable' selected. The results pane displays a table with two rows:

ID	Name	Picture
1	Nice car, but no apparent quality d... /9j/4AAQSkZIRgABAQEAAAAAAA...	
2	I see headlights, but no damage? /9j/2wBDAAYEBQYFBAYGBQYHBwY...	

At the bottom, a message says 'Query succeeded | 0s'.

Navigate to **Query Editor** in the Azure SQL Database and confirm that your comment has been added as a new row.

Consider publishing your app and loading it on your phone as one additional confirmation of success.



Data DevOps

Objective(s) in this section speak to use of Azure DevOps, Data Studio, etc. for source control and release management.

Objective: Synapse + DevOps | Source Control

Note: These instructions also apply (with minor differences) to Azure Data Factory.

Follow the instructions in this section to **connect Synapse to DevOps, create branches, and pull requests**.

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- DevOps
- Synapse (with Dedicated SQL Pool)

Setup Code Repository

Navigate to your Synapse Analytics workspace and then click the **Open** link on the “**Open Synapse Studio**” rectangle.

The screenshot shows the Microsoft Azure portal interface. The left sidebar has a 'Synapse live' dropdown open, showing 'Set up code repository' under the 'Linked' section. The main content area displays a 'Resource Explorer' with a tree view of databases, tables, and other resources. To the right, there's a large icon of two cylinders and a screen with code brackets, labeled 'Select an item' with the sub-instruction 'Use the resource explorer to select or create a new item'.

Click “**Synapse live**” and then select “**Set up code repository**” from the resulting drop-down.

The screenshot shows a 'Configure a repository' dialog box overlaid on the Azure portal. The dialog title is 'Configure a repository' and it says 'Specify the settings that you want to use when connecting to your repository.' It includes fields for 'Repository type' (set to 'Azure DevOps Git') and 'Azure Active Directory' (set to 'Microsoft'). At the bottom are 'Continue' and 'Cancel' buttons.

On the “**Configure a repository**” pop-out, enter values for the following items:

Repository Type	Select “Azure DevOps Git” from the drop-down
Azure Active Directory	Select the value appropriate for your organization

Click the **Continue** button.

On the second “Configure a repository” pop-out, enter values for the following items:

Azure DevOps Organization...	Select your DevOps account
Project Name	Select your DevOps project
Repository Name	Select the repo created with your DevOps project
Collaboration Branch	Dropdown, click “+ Create new”, enter an appropriate name in the resulting pop-up, and click Create
Publish Branch	Confirm default, workspace_publish
Root Folder	Confirm default, /
Import Existing Resources...	Unchecked

Click the **Apply** button.

The screenshot shows the Microsoft Azure Synapse Analytics workspace interface. On the left, there's a sidebar with icons for Home, Data, Databricks, Data Lake, Data Factory, and Data Integration. The main area shows a workspace named 'collaboration_synapse' with one database listed: 'practicumsw_dsp (SQL)'. At the top right, there's a 'Set working branch' dialog box. It has two radio buttons: 'Create new' (unchecked) and 'Use existing' (checked). Below the radio buttons is a dropdown menu showing 'collaboration_synapse'. At the bottom of the dialog is a blue 'Save' button.

Confirm values and then click the **Save** button.

When processing is complete, you will note that we are now working in the new branch.

Confirm Success

In this section, we will confirm success by simulating a deployment workflow {i.e., modification, and then pull request}.

Modify one or more of the following:

- SQL Script
- Spark Notebook
- Data Flow
- Pipeline

Pull Request

Next, we will promote changes from the collaboration branch to the master branch via Pull Request.

Click on the “{collaboration} branch” drop-down control and select “**Create pull request...**” from the resulting drop-down.

A third tab in your browser (pointing to <https://dev.azure.com...>) will open.

The screenshot shows the 'New pull request' interface in Azure DevOps. On the left, a sidebar lists project navigation options like Overview, Boards, Repos, Pull requests, Pipelines, Test Plans, Artifacts, and Compliance. The 'Pull requests' option is currently selected. The main content area displays a 'New pull request' form. At the top, it shows the source branch as 'collaboration_synapse' and the target branch as 'master'. Below this, tabs for 'Overview', 'Files 6', and 'Commits 4' are visible. The 'Overview' tab is active. The main form fields include:

- Title:** A text input field containing 'My Pull Request'.
- Description:** A large text area with placeholder text 'Describe the code that is being reviewed'.
- Markdown supported:** A note indicating that Markdown is supported, with a link to documentation.
- Link work items:** A note with a link to add work items.
- Reviewers:** A search bar labeled 'Search users and groups to add as reviewers'.
- Add required reviewers:** A button to add required reviewers.
- Work items to link:** A search bar labeled 'Search work items by ID or title'.
- Tags:** An empty input field for adding tags.

A blue 'Create' button is located at the bottom right of the form. The browser address bar shows the URL: https://dev.azure.com/practicumado/practicumadop/_git/practicumadop/pullrequestcreate?sourceRef=collaboration_synapse&targetRef=master

On the second “New pull request” page, enter values for the following items:

Branches	{collaboration branch} into {master branch} ... you may have to create a master branch if one does not exist
Title	Enter a title that is meaningful for you (and aligned with your naming standards)

Enter values for items like **Reviewers** and “Work items...” as appropriate.

Review settings on remaining tabs {e.g., **Files** and **Commits**}. No additional settings are required for this exercise.

Click the **Create** button.

The screenshot shows the Azure DevOps interface for a pull request. On the left, there's a sidebar with project navigation (Overview, Boards, Repos, etc.) and a 'Pull requests' section highlighted. The main area is titled 'My Pull Request' and shows an active pull request from 'Rich Chapler' to 'collaboration_synapse' into 'master'. The pull request has a green checkmark indicating 'No merge conflicts' (last checked 'Just now'). Below this are tabs for 'Overview', 'Files', 'Updates', and 'Commits'. To the right, there are sections for 'Reviewers' (with a note 'Required' and 'No required reviewers'), 'Tags' (with a note 'No tags'), and 'Work items' (with a note 'No work items'). A comment input field says 'Add a comment...' and a recent activity log shows 'Rich Chapler created the pull request' 'Just now'. At the top right, there are buttons for 'Approve' and 'Complete'.

At this point (and if specified), designated reviewers would be notified by email, and they will need to review and approve for the Pull Request to move forward.

Additional configuration can provide for validation of other gating criteria {e.g., inclusion of text patterns that might be secret like an account key}.

We have not included gating factors in this example.

Click the **Complete** button.

The screenshot shows the same Azure DevOps interface as before, but with a modal dialog box titled 'Complete pull request' overlaid. The dialog contains fields for 'Merge type' (set to 'Merge (no fast forward)'), a diagram of a merge commit, and 'Post-completion options' with three checked checkboxes: 'Complete associated work items after merging', 'Delete collaboration_synapse after merging', and 'Customize merge commit message'. At the bottom right of the dialog are 'Cancel' and 'Complete merge' buttons.

Confirm selections in the “**Complete pull request**” pop-out.

Click the “**Complete merge**” button.

The screenshot shows the Azure DevOps interface for a pull request titled "My Pull Request". The status bar indicates it is "Completed" by Rich Chapler at just now, merging "collaboration_synapse" into "master". The "Overview" tab is selected. On the right, there are sections for "Reviewers" (Required: No required reviewers), "Optional" (No optional reviewers), "Tags" (No tags), and "Work items" (No work items). A "Description" field is present. At the bottom, a comment from Rich Chapler is shown: "Rich Chapler completed the pull request Just now". A "Cherry-pick" and "Revert" button is visible above the merge message.

Note: Post-completion options such as "**Cherry-pick**" and **Revert** provide for operational control even after a code merge.



Objective: Synapse | Schema Comparison

Follow the instructions in this section to **compare two database schemas** {e.g., old, and new}.

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Data Studio
- Synapse

Although this example focuses on Synapse, the same solution applies to SQL.

Stage Resources

Navigate to your Synapse Workspace.

The screenshot shows the Microsoft Azure portal with the URL <https://ms.portal.azure.com/#@microsoft.onmicrosoft.com/resource/subscriptions/91e9fddc-ef15-416c-9be1-085f8b1b46e...>. The page title is "practicumsaw - Microsoft Azure". The main content area displays the "practicumsaw" Synapse workspace details. Key information includes:

- Resource group: practicumrg
- Status: Succeeded
- Location: West US 2
- Subscription: rchabler
- Subscription ID: 91e9fddc-ef15-416c-9be1-085f8b1b46ed
- Managed virtual network: No
- Managed Identity object ID: 78190823-9f09-4ed0-94af-118b5e80a55b
- Workspace web URL: <https://web.azuresynapse.net?workspace=%2...>
- Tags: Click here to add tags
- Firewalls: Show firewall settings
- Primary ADLS Gen2 account: <https://practicumdl.dfs.core.windows.net>
- Primary ADLS Gen2 file system: practicumdls
- SQL admin username: sqladminuser
- SQL Active Directory admin: rchabler@microsoft.com
- Dedicated SQL endpoint: practicumsaw.sql.azuresynapse.net
- Serverless SQL endpoint: practicumsaw-ondemand.sql.azuresynapse.net
- Development endpoint: <https://practicumsaw.dev.azuresynapse.net>

The sidebar on the left contains the following navigation links:

- Overview
- Activity log
- Access control (IAM)
- Tags
- Diagnose and solve problems
- Settings
- SQL Active Directory admin
- Properties
- Locks
- Analytics pools
 - SQL pools
 - Apache Spark pools
 - Scope pools

Follow the directions at [Instantiate Resources | Synapse | SQL Pool](#) to create two SQL Pools named **practicumdsp_dev** and **practicumdsp_prod**.

We will delete both SQL Pools at the end of this objective.

Open Azure Data Studio.

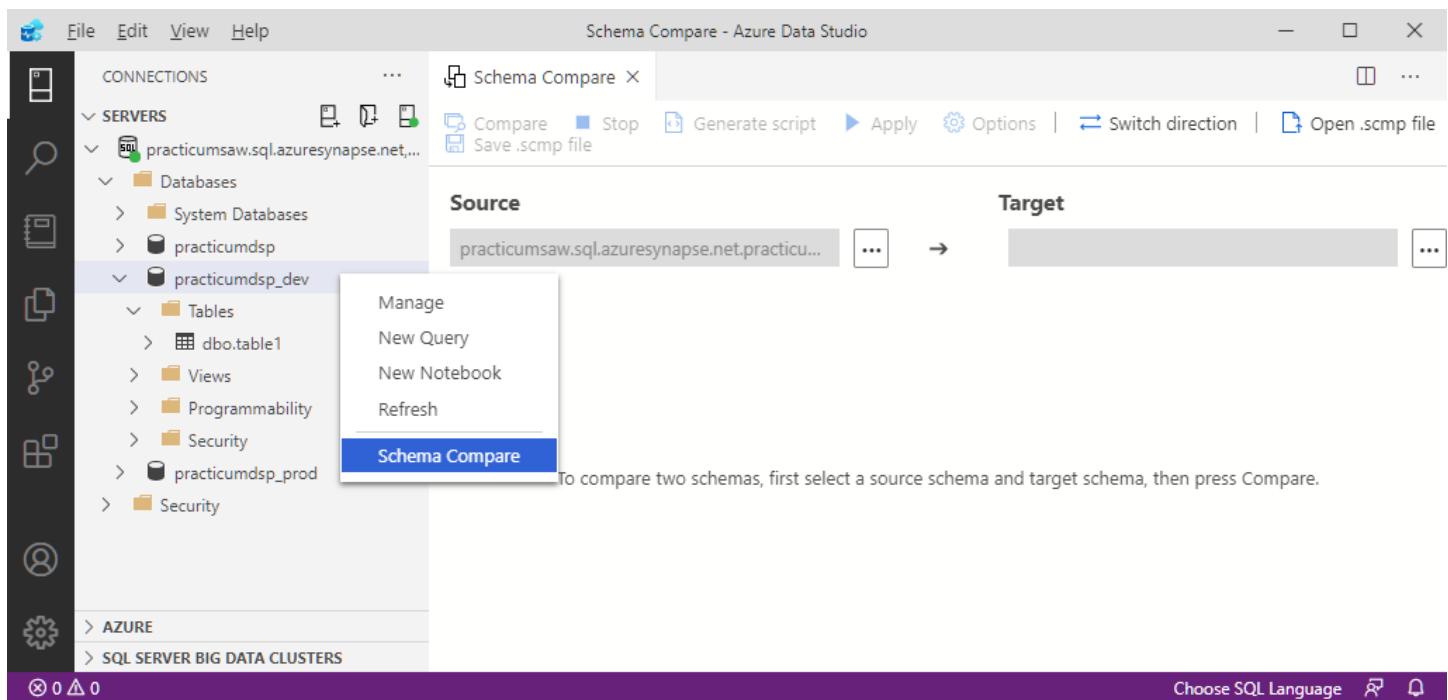
The screenshot shows the Azure Data Studio interface with the following details:

- File Edit View Help
- CONNECTIONS
- SERVERS: practicumsaw.sql.azuresynapse.net,practicumdsp_prod (selected)
- Databases: System Databases, practicumdsp, practicumdsp_dev, practicumdsp_prod (selected)
- Messages:
 - Started executing query at Line 1
 - Commands completed successfully.
 - Total execution time: 00:00:01.257
- Bottom status bar: Ln 6, Col 2 Spaces: 4 CRLF SQL MSSQL 0 rows 00:00:01 practicumsaw.sql.azuresynapse.net : practicumdsp_prod

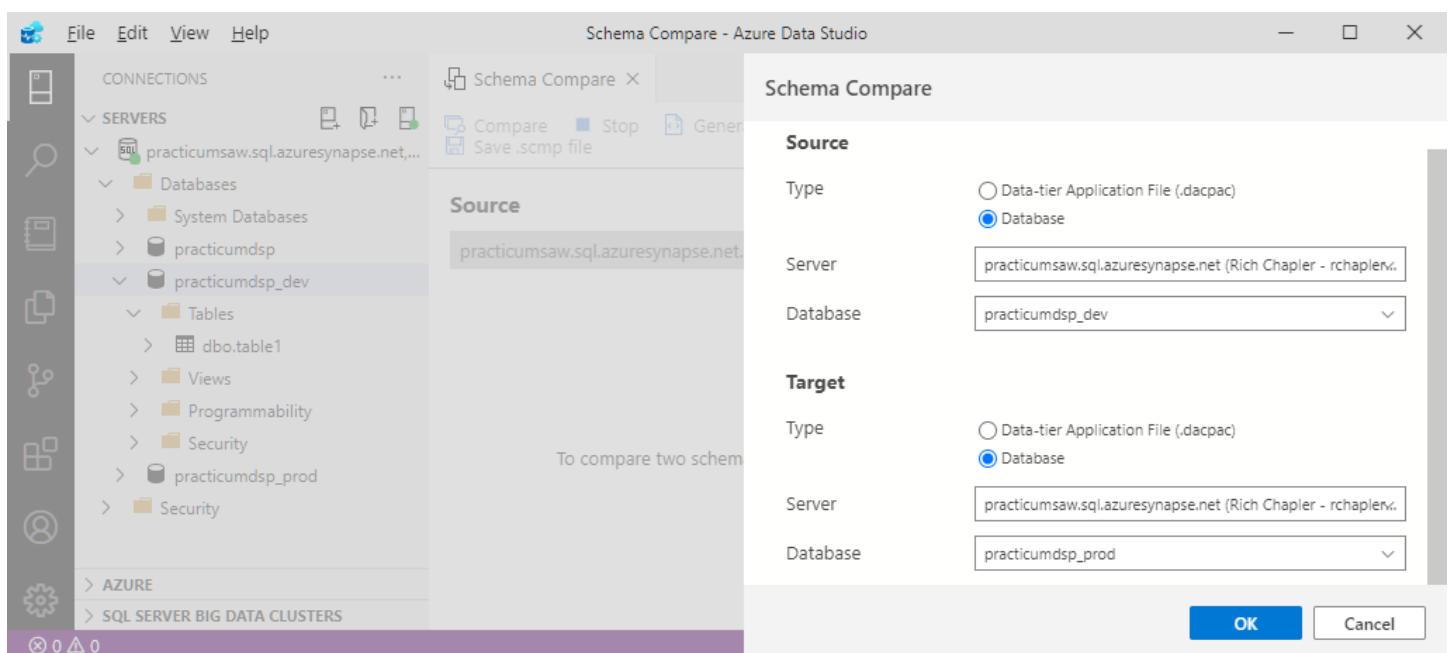
Execute the following T-SQL on practicumdsp_dev:

```
CREATE TABLE [dbo].[table1]
(
    [column1] [nvarchar](128) NULL,
    [column2] [nvarchar](128) NULL,
    [column3] [nvarchar](128) NULL
)
```

Schema Compare



Right-click on the SQL Pool and click “Schema Compare” in the resulting drop-down.



Click the ellipses button.

On the “Schema Compare” pop-out, enter values for the following items:

Source Type	Select the Database radio button
Source Server and Database	Confirm selection of your Synapse server and the practicumdsp_dev database
Target Type	Select the Database radio button
Target Server and Database	Confirm selection of your Synapse server and the practicumdsp_prod database

Click the **OK** button.

The screenshot shows the Schema Compare interface in Azure Data Studio. The left sidebar displays connections and servers, with 'practicumsaw.sql.azuresynapse.net...' selected. The main area shows a comparison between 'Source' (practicumsaw.sql.azuresynapse.net.practicumdsp_dev) and 'Target' (practicumsaw.sql.azuresynapse.net.practicumdsp_prod). A table named 'dbo.table1' is selected in the Source list. The 'COMPARE DETAILS' section shows T-SQL code for creating the table, including columns and constraints. The 'Actions' column for the table row has a checkbox labeled 'Add' checked.

On the “Schema Compare” tab, click the **Compare** button.

We only created a table on the development database, so Schema Compare should surface that table as something missing from the production database. Click on the row to see T-SQL for the identified item.

Click the **Apply** button to update the target.

The screenshot shows the Schema Compare interface after applying changes. The 'TARGET' section now lists 'practicumsaw.sql.azuresynapse.net.practicumdsp_prod'. The 'Tasks' tab shows a success message: 'Apply schema compare changes succeeded' at 9:47:58 AM - 9:48:09 AM (00:00:10).



Good Job!

You have successfully completed the objective

Data Governance

Objective(s) in this section describe methods for cataloging and classifying data.

Objective: Purview | Discover Data

Data discovery is “an organic process based on communal knowledge”, characterized by the following questions:

- **Search** ... “does the data that I need exist?”
- **Location** ... “how do I connect to Data Asset X?”
- **Related** ... “are there related data assets that I might find valuable?”
- **Contacts** ... “who are the owners and experts for Data Asset X?”

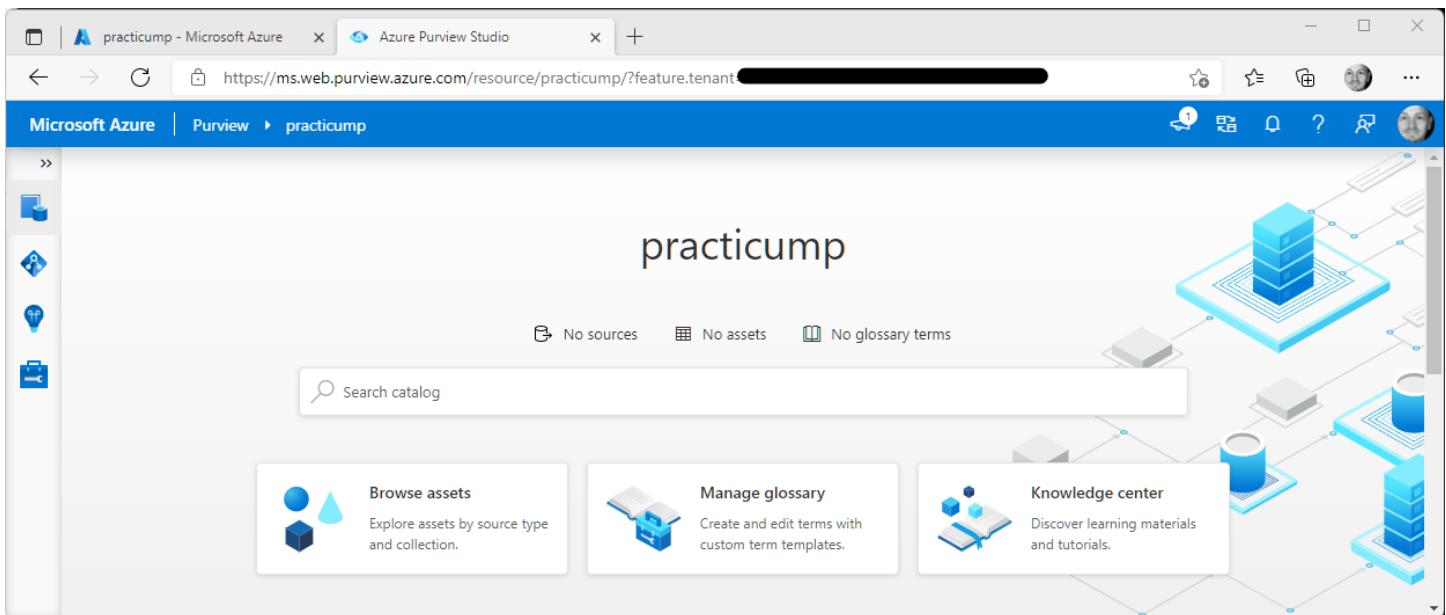
In this section, we will configure Purview and successfully respond to these questions.

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Purview
- SQL

Navigate to your instance of Purview in the Azure portal.

Click the “Open Purview Studio” button.



Register Source

Click the **Data Map** icon on the left-hand navigation.

Click the **Register** button on the resulting page.

practicump - Microsoft Azure Azure Purview Studio

https://ms.web.purview.azure.com/resource/practicump/main/datasource/registeredSources?feature.tenant=██████████

Microsoft Azure | Purview > practicump

Search assets



- > Sources
- Collections
- Source management
- Scan rule sets
- Integration runtimes
- Annotation management
- Classifications
- Classification rules

Sources

[Register](#) [Refresh](#) [Map view](#)

[Filter by keyword](#)

Showing 1 collection, 0 sources

Register sources

[Filter by keyword](#)

All [Azure](#) [Database](#) [Power BI](#) [Services and apps](#)

Azure [MULTIPLE]	AWS account [MULTIPLE]	Azure Synapse Analytics [MULTIPLE]
Amazon S3	Azure Blob Storage	Azure Cosmos DB (SQL API)
Azure Data Explorer (Kusto)	Azure Data Lake Storage Gen1	Azure Data Lake Storage Gen2
Azure Database for MySQL	Azure Database for PostgreSQL	Azure Dedicated SQL Pool (formerly SQL DW)
Azure Files	Azure SQL Database	Azure SQL Database Managed Instance
Cassandra (Preview)	erwin (Preview)	Google BigQuery (Preview)
Hive Metastore	Looker (Preview)	Oracle
SAP ECC	SAP HANA	

Click the **Azure** tab on the “Register sources” pop-out.

Click to check the “**Azure SQL Database**” option.

Click the **Continue** button.

The screenshot shows the Azure Purview Studio interface. On the left, there's a sidebar with various management options like Sources, Collections, Source management, Scan rule sets, Integration runtimes, Annotation management, Classifications, and Classification rules. The main area is titled 'Sources' and shows a table with one row: 'Showing 0 sources'. To the right of this table is a 'Register sources (Azure SQL Database)' dialog box. It contains fields for 'Name' (set to 'practicumsd'), 'Azure subscription' (set to 'rchapler (91e9fddc-ef15-416c-9be1-085f8b1b46ed)'), 'Server name' (set to 'practicumsd'), 'Endpoint' (set to 'practicumsd.database.windows.net'), and 'Select a collection' (set to 'practicump'). A note below the collection field says 'All assets under this source will belong to the collection you select.' At the bottom of the dialog are 'Register', 'Back', and 'Cancel' buttons.

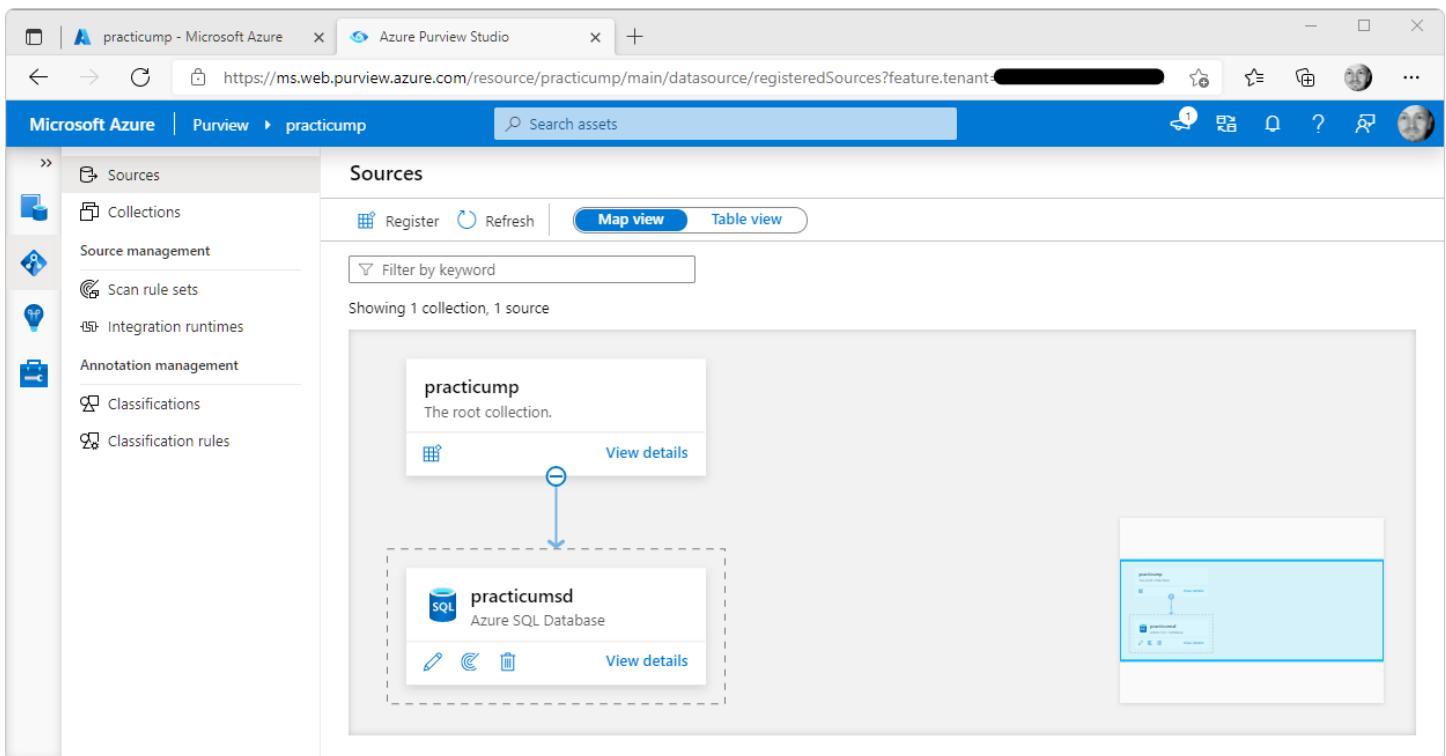
On the “Register sources (Azure SQL Database)” pop-out, enter values for the following items:

Name	Enter a value that is meaningful and aligned with standards
Subscription	Select your subscription
Server Name	Enter the name of your SQL Database Server
Endpoint	Confirm the populated value
Select a Collection	Leave default value

Click the **Register** button.

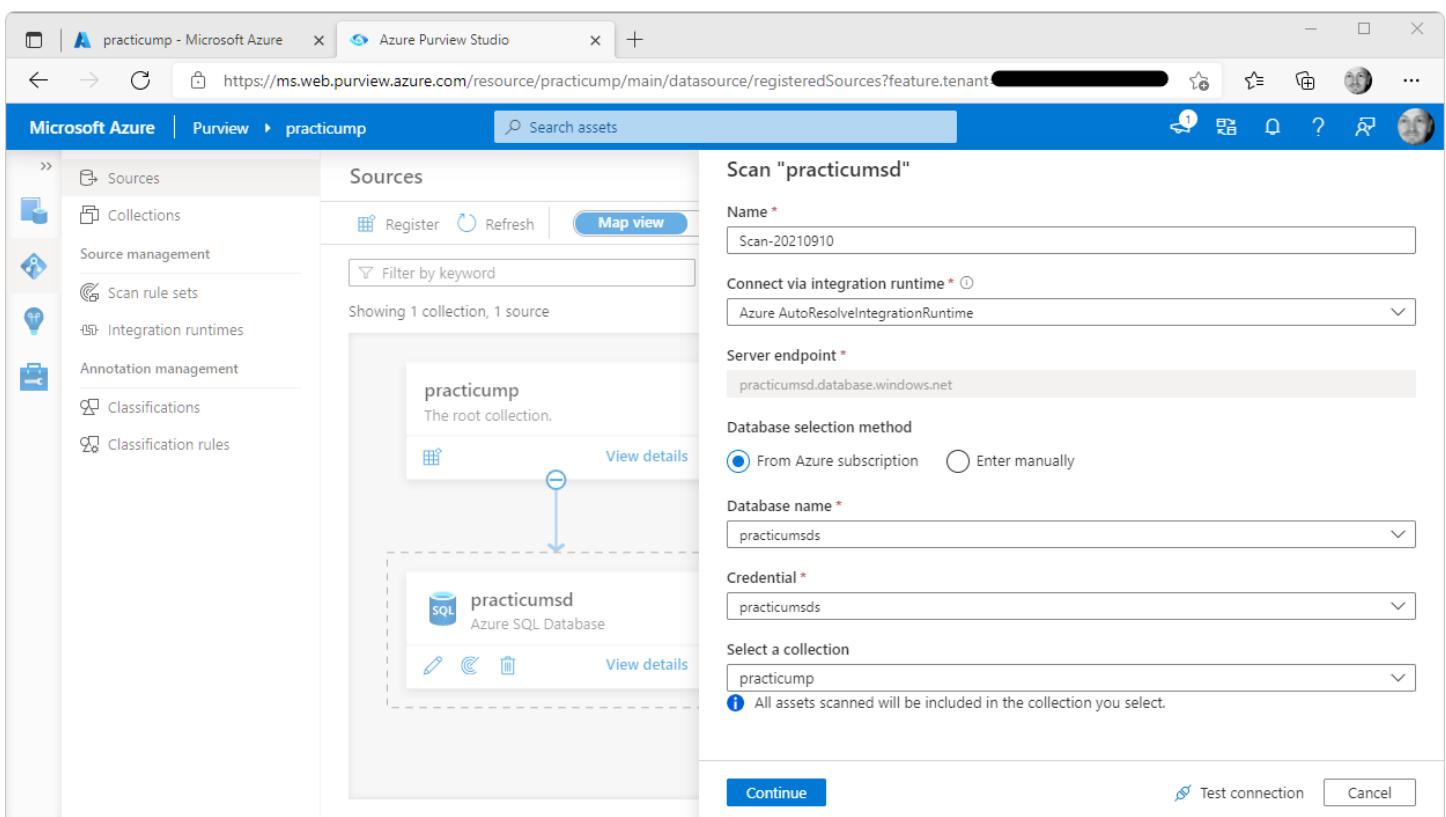
Scan Data

Navigate to Sources, “Map view”.



The screenshot shows the Microsoft Azure Purview Studio interface. The left sidebar has a 'Sources' icon selected. The main area displays a tree view under the 'Sources' tab. A collection named 'practicump' is expanded, showing its root status and a 'View details' link. Below it, a sub-collection 'practicumsd' is shown, also with a 'View details' link. To the right, there's a preview pane showing a simplified version of the collection structure.

Click the “New scan” icon on the newly created item.



The screenshot shows the 'Scan "practicumsd"' configuration dialog. It includes fields for Name (set to 'Scan-20210910'), Connect via integration runtime (set to 'Azure AutoResolveIntegrationRuntime'), Server endpoint (set to 'practicumsd.database.windows.net'), Database selection method (radio button selected for 'From Azure subscription'), Database name (set to 'practicumsds'), Credential (set to 'practicumsds'), and Select a collection (set to 'practicump'). At the bottom are 'Continue', 'Test connection', and 'Cancel' buttons.

On the “Scan...” pop-out, enter values for the following items:

Name	Enter a value that is meaningful and aligned with standards
Database Selection Method	Select the “From Azure subscription” radio button
Database Name	Select your Azure SQL database
Credential	Select your “SQL Authentication” credential

Click “Test connection” to confirm successful configuration and then click the **Continue** button.

The screenshot shows the Azure Purview Studio interface. On the left, a sidebar menu includes 'Sources', 'Collections', 'Source management', 'Scan rule sets' (which is selected), 'Integration runtimes', 'Annotation management', 'Classifications', and 'Classification rules'. The main area is titled 'Sources' and shows a single collection named 'practicump'. Below it is a detailed view of an Azure SQL Database named 'practicumsd'. A 'Scope your scan' pop-up window is open on the right, listing all assets under 'practicump'. These assets include various tables and views from the 'SalesLT' schema, such as 'Customer', 'ProductModel', 'ProductDescription', 'Product', 'ProductCategory', 'BuildVersion', 'ErrorLog', 'Address', 'CustomerAddress', 'SalesOrderDetail', 'SalesOrderHeader', 'ProductModelCatalogDescription', 'ProductAndDescription', and 'GetAllCategories'. Each asset has a checkbox next to it, and most are checked. At the bottom of the pop-up are 'Continue', 'Back', and 'Cancel' buttons.

Review and select desired data assets from those listed in the “Scope your scan” pop-out and then click the **Continue** button.

This screenshot shows the 'Select a scan rule set' dialog. It lists a single rule set named 'AzureSqlDatabase' with the status 'SYSTEM DEFAULT'. A description below states: 'Microsoft default scan rule set that includes all supported system classification rules'. There are 'View detail' and 'Edit' buttons next to the rule set name. The background shows the same 'Sources' view as the previous screenshot, with the 'practicump' collection and its 'practicumsd' database detail view. The 'Scan rule sets' option is still selected in the sidebar. At the bottom of the dialog are 'Continue', 'Back', and 'Cancel' buttons.

Confirm selection of the default Scan Rule Set on the “Select a scan rule set” pop-out and then click the **Continue** button.

The screenshot shows the Azure Purview Studio interface. On the left, there's a sidebar with icons for Sources, Collections, Source management, Scan rule sets, Integration runtimes, Annotation management, Classifications, and Classification rules. The main area is titled "Sources" and shows a single item: "practicump The root collection.". To the right, a modal window titled "Set a scan trigger" is open. It contains a description: "Set a scan trigger to run the scan at specific dates and times. If once, the scan will start after setup is completed. If recurring, the scan will start at a date and time you choose. The initial scan is a full scan and every subsequent scan is incremental." Below this, there are two radio buttons: "Recurring" and "Once". The "Once" button is selected. At the bottom of the modal are three buttons: "Continue", "Back", and "Cancel".

Click the **Once** radio button on the “Set a scan trigger” pop-out and then click the **Continue** button.

The screenshot shows the "Review your scan" dialog. On the left, the same sidebar and "Sources" list are visible. The main area now displays the scan configuration. Under "Review your scan", it says "Review your scan before running it." The configuration details are as follows:

Basics	Name: Scan-20210910 Collection: practicump
Credential	Type: SQL authentication Name: practicumsts
Scan Scope	Scope: Full
Scan Rule Set	Name: AzureSqlDatabase Type: System
Scan Trigger	Start at: Immediately Recurrence: Once

At the bottom of the dialog are three buttons: "Save and Run", "Back", and "Cancel".

Review scan settings and then click the **Save and Run** button.

Confirm Scan

Navigate to Sources, “Table view”.

Sources

Showing 1 source

Name	Source type	Collection	Source id	Scans	Registered on
AzureSqlDatabase-P3k	Azure SQL Data...	practicump	practicumsds.d...	1	09/10/21 01:59 PM

Note that the **Scans** column now shows 1.

Click on the **Name** link to drill through to details. Allow time for processing.

practicumsd

Azure SQL Database

New scan Edit source Delete source Refresh

Overview Scans

Source ID: practicumsd.database.windows.net Registered on 09/11/2021 03:18:28 AM

Scans 1 Scanned assets 17

Classified assets 15

Recent scans

Scan name	Last run status	Scan rule set	Last scan time
Scan-20210910	In progress	AzureSqlData...	09/11/21 03:25 AM

→ See all applied scans

Recent failed scans

Scan name	Status	Suggestions
No recent failed scan runs		

→ See all failed scan runs

Collection path
practicump

Source hierarchy
rchapter
Subscription
practicumrg
Resource group

Browse Assets

After scanning, it takes time for **Asset Insights** to reflect new assets.

We do not have to wait to browse assets, however.

Click the **Data Catalog** icon on the left-hand navigation.

The screenshot shows the Azure Purview Studio interface. The top navigation bar includes 'practicump - Microsoft Azure', 'Azure Purview Studio', and a URL 'https://ms.web.purview.azure.com/resource/practicump/?feature.tenant=[REDACTED]'. The left sidebar has icons for Data Catalog, Data Governance, Data Quality, and Data Integration. The main area is titled 'practicump' and displays a network diagram of data assets. Below the title, it says '1 source', 'No assets', and 'No glossary terms'. A search bar says 'Search catalog'. Three buttons are shown: 'Browse assets' (Explore assets by source type and collection), 'Manage glossary' (Create and edit terms with custom term templates), and 'Knowledge center' (Discover learning materials and tutorials).

Click on the “**Browse assets**” button.

The screenshot shows the 'Browse assets' page under the 'Data catalog' section. The top navigation bar is identical to the previous screenshot. The left sidebar shows 'Data catalog >'. The main content area is titled 'Browse assets' with a 'Refresh' button. It offers filtering by 'By collection' or 'By source type'. A 'Filter by keyword' input field is present. Below this, it says 'Showing 1 collection'. A table lists one collection:

Name	Description	Assets	Collection admin
practicump	The root collection.	4	Rich Chapler

Click on the Root Collection link {i.e., name}.

The screenshot shows the Microsoft Azure Purview Studio interface. The left sidebar displays navigation links: Data catalog, Browse assets, Refresh, By collection (selected), By source type, View collection tree, and practicump. Under practicump, Sub collection(s) and Related are listed. A narrow results by section includes Classification (Email Address, Person's Name, U.S. Phone Number, World Cities), Contact, Label, and Glossary term. The main content area shows a list of assets under the 'practicump' collection, filtered by source type (all) and instance (all). The results are sorted by name, showing 1-23 out of 23 results. The assets listed are:

Asset	Type	Location
Address	Azure SQL Table	mssql://practicumsd.database.windows.net/practicumsds/SalesLT/Address
BuildVersion	Azure SQL Table	mssql://practicumsd.database.windows.net/practicumsds/dbo/BuildVersion
Customer	Azure SQL Table	mssql://practicumsd.database.windows.net/practicumsds/SalesLT/Customer
CustomerAddress	Azure SQL Table	mssql://practicumsd.database.windows.net/practicumsds/SalesLT/CustomerAddress
dbo	Azure SQL Schema	mssql://practicumsd.database.windows.net/practicumsds/dbo

Pagination controls at the bottom indicate one page of one result.

Explore results.

Click on the first item, **Address**.

The screenshot shows the Azure Purview Studio interface. At the top, there are two tabs: 'practicump - Microsoft Azure' and 'Azure Purview Studio'. The URL in the address bar is <https://ms.web.purview.azure.com/resource/practicump/main/catalog/entity?guid=81afae99-9619-4559-aa95-b5f6f6f60000§i...>. The main content area displays an asset named 'Address' (SQL Azure SQL Table). The 'Overview' tab is selected. The asset has no description, classifications, or glossary terms. It is part of a collection path 'practicump' and is located in a hierarchy under 'SalesLT' (Azure SQL Schema) which is under 'practicumsd' (Azure SQL Database) which is under 'practicumsd' (Azure SQL Server). The asset was last updated on September 11, 2021, at 10:29 AM UTC by an automated scan.

Notice the automatic “**World Cities**” classification.

Drill-through the tabs {e.g., Properties, Schema, Lineage, etc.}.

Confirm Success

Now that we have implemented the basics in Purview, can we answer the questions previously posed?

Search

Question: “**Does the data that I need exist?**”

Enter a phrase into the search box at the top center of the window {e.g., “address”}.

The screenshot shows the Azure Purview Studio interface with a search query for "address". The results page displays two items: "Address" and "CustomerAddress", both identified as Azure SQL Tables located at `mssql://practicumsd.database.windows.net/practicumsds/SalesLT/Address` and `mssql://practicumsd.database.windows.net/practicumsds/SalesLT/CustomerAddress` respectively. The left sidebar includes filters for Collection, Classification, Contact, Label, and Glossary term.

Asset Type	Name	Location
Azure SQL Table	Address	<code>mssql://practicumsd.database.windows.net/practicumsds/SalesLT/Address</code>
Azure SQL Table	CustomerAddress	<code>mssql://practicumsd.database.windows.net/practicumsds/SalesLT/CustomerAddress</code>

Returned search results will include all assets that include the requested phrase.

Location

Question: “**how do I connect to Data Asset X?**”

Drill into the first search result, Address.

The screenshot shows the Azure Purview Studio interface with the URL <https://ms.web.purview.azure.com/resource/practicump/main/catalog/entity?guid=81afae99-9619-4559-aa95-b5f6f6f60000§i...>. The main content area displays the 'Address' asset, which is an Azure SQL Table. The 'Overview' tab is selected, showing the following details:

- Collection path:** practicump
- Hierarchy:** A tree view showing the database structure:
 - practicumsd.database.windows.net (Azure SQL Server)
 - practicumsds (Azure SQL Database)
 - SalesLT (Azure SQL Schema)
 - Address (Azure SQL Table)

Other tabs visible include Properties, Schema, Lineage, Contacts, and Related. A note at the bottom right indicates the asset was updated on September 11, 2021, 10:29 AM UTC by automated scan.

On the **Overview** tab, you will see **Hierarchy** information which includes the Fully Qualified Domain Name of the Azure SQL Server.

Related

Question: “**are there related data assets that I might find valuable?**”

Continue in the search result for phrase “Address”.

The screenshot shows the Azure Purview Studio interface. The left sidebar has icons for Data Catalog, Purview, and other services. The main area shows a search result for "Address". The "Related" tab is selected, displaying a list of 13 related items under the heading "Showing 1 to 13 of 13 items". The list includes: Address (Azure SQL Table), Customer (Azure SQL Table), CustomerAddress (Azure SQL Table), Product (Azure SQL Table), ProductCategory (Azure SQL Table), ProductDescription (Azure SQL Table), ProductModel (Azure SQL Table), ProductModelProductDescription (Azure SQL Table), SalesOrderDetail (Azure SQL Table), and SalesOrderHeader (Azure SQL Table). The "Overview" tab is also visible on the left.

On the **Related** tab, you will see a list of related assets (as identified by Purview).

Contacts

Question: “who are the owners and experts for Data Asset X?”

Continue in the search result for phrase “Address”.

The screenshot shows the Azure Purview Studio interface. The left sidebar has icons for Data Catalog, Purview, and other services. The main area shows a search result for "Address". The "Contacts" tab is selected, displaying sections for "Experts" and "Owners". The "Experts" section states "No experts for this asset." and the "Owners" section states "No owners for this asset.". The "Overview" tab is also visible on the left.

On the **Contacts** tab, you will see lists of **Experts** and **Owners** you can contact about the asset (as identified by Purview).

If, as in the snip above, there are no Contacts and you can add information not already captured, click the **Edit** button.

The screenshot shows the Azure Purview Studio interface for editing a data asset named "Address". The "Contacts" tab is active. On the left, there are sections for "Experts" and "Owners", each with a search bar and a note about their roles. The "Experts" section shows one result: Rich Chapter, Curator, with the email address Richard.Chapler@microsoft.com. At the bottom are "Save" and "Cancel" buttons.

In either the **Experts** or the **Owners** drop-down, search for the user or group in your directory. Click **Save**.



Objective: Purview | Classify Data

Data classification is a process of annotating data assets to help organizations answer the following questions:

- **Definition** ... “how does data map to a business context?” and “what is the intended use of Data Asset X?”
- **Documentation** ... “where is the documentation of Data Asset X?”
- **Vocabulary** ... “what else {e.g., synonyms, acronyms, etc.} might help me understand Data Asset X?”
- **Standardization** ... “how do I classify all data assets in a standard way?”

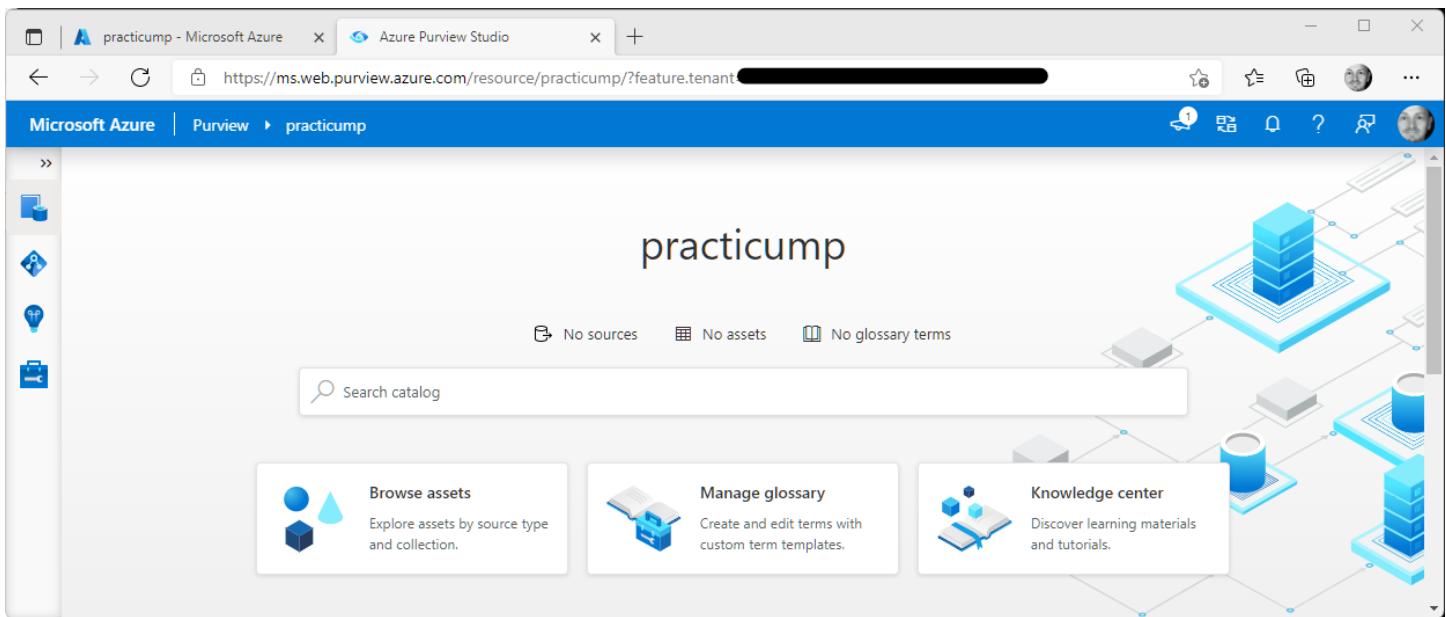
In this section, we will configure Purview and successfully respond to these questions.

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Purview
- SQL

Navigate to your instance of Purview in the Azure portal.

Click the “Open Purview Studio” button.



Click the “Manage Glossary” button.

Create Term

Click the “+ New term” button on the resulting page.

On the resulting pop-out, confirm selection of “System default” and then click the **Continue** button.

The screenshot shows the 'New term' creation page in Azure Purview Studio. The URL in the browser is [https://ms.web.purview.azure.com/resource/practicump/main/catalog/term/new?feature.tenant=\[REDACTED\]](https://ms.web.purview.azure.com/resource/practicump/main/catalog/term/new?feature.tenant=[REDACTED]). The page has a blue header with 'Microsoft Azure' and 'Purview' navigation, and a search bar. On the left, there's a sidebar with icons for Data catalog, Glossary terms, and other Purview features. The main content area is titled 'New term'. It includes fields for 'Term template' (set to 'System default'), 'Status' (set to 'Draft'), and tabs for 'Overview' (selected), 'Related', and 'Contacts'. The 'Overview' tab contains fields for 'Name' (Product Number), 'Definition' (The stock-keeping unit (SKU) that uniquely identifies a product), 'Parent' (Select...), 'Acronym' (SKU), and 'Resources' (Resource name: Wikipedia Definition, Resource link: https://en.wikipedia.org/wiki/Stock_keeping_unit). At the bottom are 'Create' and 'Cancel' buttons.

On the “New term” page, enter values for the following items:

Name	Enter a value that is meaningful and aligned with standards
Definition	Select your subscription
Parent	Select a parent term, if applicable
Acronym	Capture abbreviation, if applicable
Resources	Click “+ Add a resource” to capture hyperlinks with descriptive content

Click the **Create** button.

Use Term

Search for “product” to identify assets that use terms like Product Number.

The screenshot shows the Azure Purview Studio interface. The top navigation bar includes tabs for Microsoft Azure, Purview, and practicum. A search bar at the top right contains the query "product". The main content area is titled "Search results for product". It features a sidebar with filters for "Source types : all" and "Instances : all", and a "Clear all filters" button. A "Filter by keyword" input field is also present. The results list shows seven items, with the first item expanded:

- Product**
Azure SQL Table
mssql://practicumsd.database.windows.net/practicumsds/SalesLT/Product
- ProductModelProductDescription**
Azure SQL Table
mssql://practicumsd.database.windows.net/practicumsds/SalesLT/ProductModelProductDescription
- ProductCategory**
Azure SQL Table
mssql://practicumsd.database.windows.net/practicumsds/SalesLT/ProductCategory
- ProductDescription**
Azure SQL Table
mssql://practicumsd.database.windows.net/practicumsds/SalesLT/ProductDescription

At the bottom, there are navigation buttons for "Page 1 of 1" and "Next >".

Click Product.

The screenshot shows the detailed view for the "Product" entity. The top navigation bar and search bar are identical to the previous screenshot. The main content area shows the "Schema" tab selected. The table structure is as follows:

Column name	Classifications	Sensitivity label	Glossary terms	Data type	Asset description	Description
Color				nvarchar		
DiscontinuedDate				datetime		
ListPrice				money		
ModifiedDate				datetime		
Name				nvarchar		
ProductCategoryID				int		
ProductID				int		
ProductModelID				int		
ProductNumber				nvarchar		
rowguid				uniqueidentifier		

Click the **Schema** tab.

Click on the **ProductNumber** link.

The screenshot shows the Azure Purview Studio interface. The top navigation bar includes 'practicump - Microsoft Azure' and 'Azure Purview Studio'. The address bar shows the URL: https://ms.web.purview.azure.com/resource/practicump/main/catalog/entity?guid=288bec77-3c71-41f7-bf47-09f6f6f60006§ion=entity. The search bar contains 'product'. The main content area displays the 'ProductNumber' asset, which is an 'Azure SQL Column'. Below the asset name are 'Edit', 'Refresh', and 'Delete' buttons. A horizontal menu bar at the bottom of the content area includes 'Overview', 'Properties', 'Lineage', 'Contacts', and 'Related'. To the right of the menu, it says 'Updated on September 11, 2021 10:29 AM UTC'. The 'Overview' tab is selected. The asset details are listed on the left, and their corresponding collection paths are on the right. For example, 'Asset description' is under 'Collection path practicump'. Other sections include 'Description', 'Classification', 'Fully qualified name', and 'Glossary terms'.

Click the **Edit** button.

The screenshot shows the 'Edit "ProductNumber"' screen. The top navigation bar and address bar are identical to the previous screenshot. The main content area has a title 'Edit "ProductNumber"'. It features a horizontal menu bar with 'Overview' and 'Contacts'. Below this are several input fields: 'Name' (set to 'ProductNumber'), 'Asset description' (empty), 'Classifications' (dropdown menu 'Select...'), and 'Glossary terms' (dropdown menu showing 'Product Number' with a checked checkbox). At the bottom are 'Save' and 'Cancel' buttons.

Select the new term, “Product Number” from the “Glossary terms” drop-down.

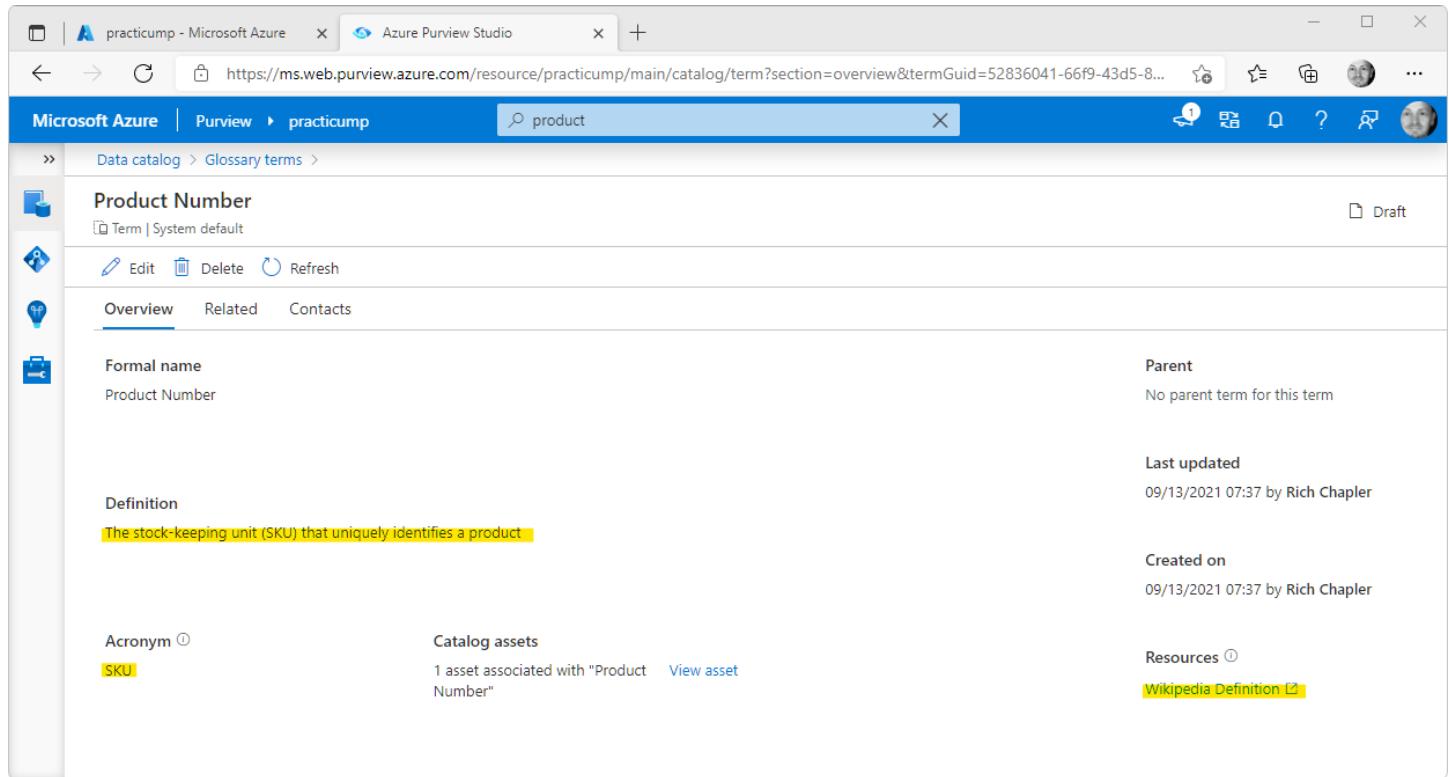
Click the **Save** button.

Confirm Success

Now that we have implemented the basics in Purview, can we answer the questions previously posed?

Definition

Questions: “**how does data map to a business context?**” and “**what is the intended use of Data Asset X?**”



The screenshot shows the Azure Purview Studio interface for managing glossary terms. The URL in the browser is https://ms.web.purview.azure.com/resource/practicump/main/catalog/term?section=overview&termGuid=52836041-66f9-43d5-8... The page displays the 'Product Number' term, which is a system default. The 'Overview' tab is selected. The term has a formal name 'Product Number' and a definition: 'The stock-keeping unit (SKU) that uniquely identifies a product.' It also has an acronym 'SKU' and catalog assets associated with it. The last update was on 09/13/2021 at 07:37 by Rich Chapter. The term is marked as 'Draft'.

This also demonstrates response to questions:

- **Documentation** ... “where is the documentation of Data Asset X?
 - Links to documentation can be included in **Resources**
- **Vocabulary** ... “what else {e.g., synonyms, acronyms, etc.} might help me understand Data Asset X?”

Standardization

Question: “**how do I classify all data assets in a standard way?**”

Edit "Address"

Overview Schema Contacts

Column name

Column name	Glossary terms	Data type	Asset description
AddressID	Select...	int	
AddressLine1	Select...	nvarchar	
AddressLine2	Select...	nvarchar	
City	Select...	nvarchar	
CountryRegion	Select...	nvarchar	
ModifiedDate	Select...	datetime	

Save **Cancel**

Use “**Column level classification**” options to characterize all country- / region-type schema elements as a standard classification.



Good Job!

You have successfully completed the objective

Objective: Purview | Understand Lineage

Data lineage describes provenance {i.e., where data comes from, where it ends up and what happens to it along the way}.

Understanding data lineage enables stakeholders to answer the following question:

- **Accountability** (aka Impact Analysis) ... “who and what would be affected by a change to or problem with Data Asset X?”

In this section, we will configure Purview and successfully respond to this question.

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Data Factory
- Data Share
- Purview
- SQL

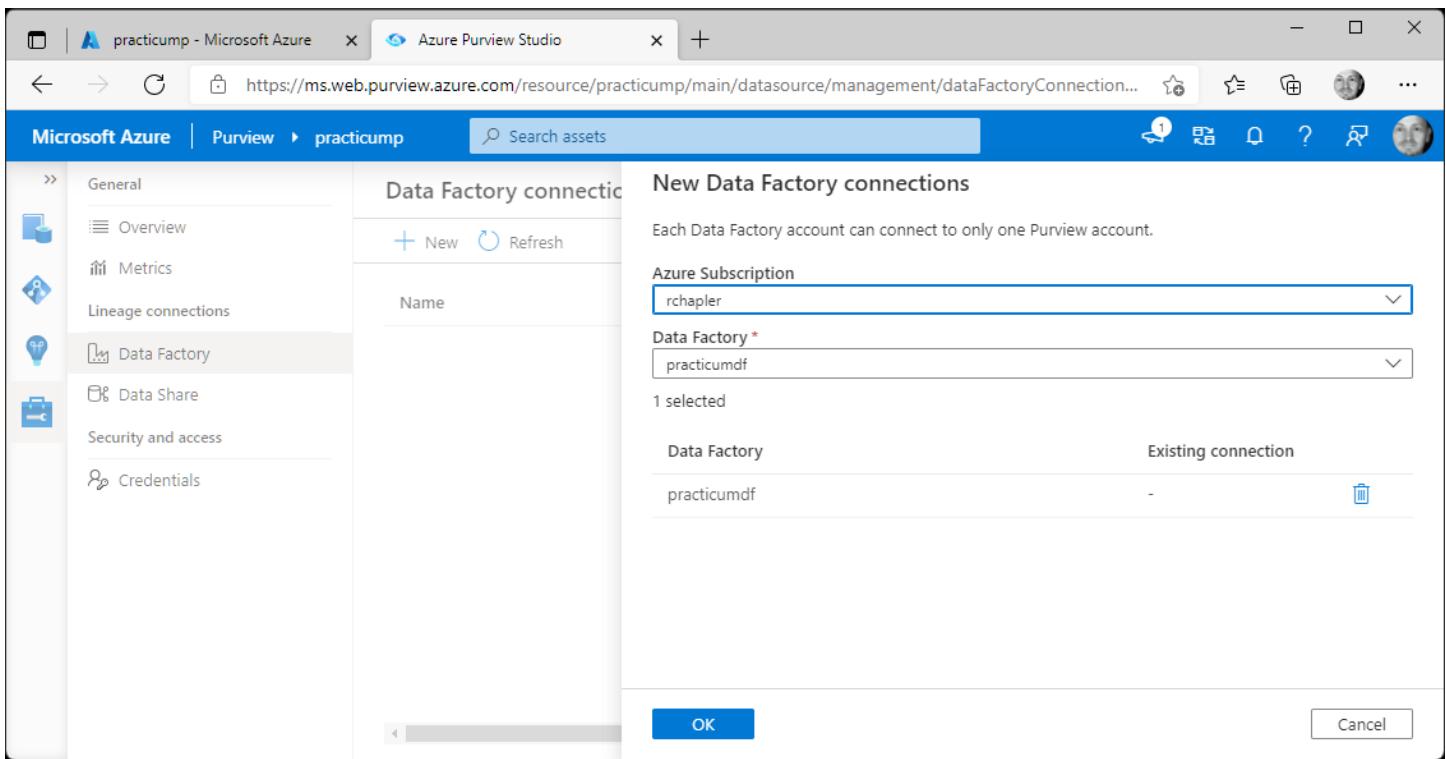
Configure Data Factory

Navigate to your instance of Purview in the Azure portal.

Click the “**Open Purview Studio**” button.

Click the **Management** icon on the left-hand navigation and then “**Data Factory**” from the resulting menu.

Click the “**+ New**” button on the resulting page.



On the “**New Data Factory connections**” pop-out, enter values for the following items:

Azure Subscription	Select your subscription
Data Factory	Select your instance of data factory

Click the **OK** button. Confirm connection.

Confirm Success

Purview will automatically detect data movement and capture data about a pipeline when it is executed in a connected data factory. Navigate to your instance of Data Factory and execute your sample pipeline.

The screenshot shows the Microsoft Azure Data Factory pipeline execution details. The pipeline, named 'pipeline1', contains one activity: 'Copy data1'. The activity is currently running, indicated by a green checkmark. The pipeline run ID is ee386add-a37b-4156-83d0-0e549901c97d. The status of the run is 'Succeeded'.

After successful execution, navigate to your instance of Purview.

The screenshot shows the Microsoft Azure Purview Studio interface. The main search bar displays the text 'pipeline1'. Below the search bar, there are three cards: 'Browse assets', 'Manage glossary', and 'Knowledge center'. The 'Browse assets' card indicates 1 source, 29 assets, and 1 glossary term.

Search for "pipeline1".

The screenshot shows the Azure Purview Studio interface. The top navigation bar includes 'practicump - Microsoft Azure', 'Azure Purview Studio', and a search bar with the query 'pipeline1'. The main content area is titled 'Search results for pipeline1' and displays two results:

- Copy data1**
Azure Data Factory Copy Activity
/subscriptions/[REDACTED]/resourceGroups/practicumrg/providers/Microsoft.Dat...
- pipeline1**
Azure Data Factory Pipeline
/subscriptions/[REDACTED]/resourceGroups/practicumrg/providers/Microsoft.Dat...

On the left, there is a sidebar with filters for 'Collection' (selected 'practicump'), 'Classification', 'Contact', and 'Content type'. The bottom of the page shows pagination controls: '< Previous', 'Page 1 of 1', and 'Next >'.

Click on the “Copy data1” result.

Click on the **Lineage** tab.

The screenshot shows the Microsoft Purview Studio interface, specifically the Data catalog section for the "practicump" resource. A search bar at the top right contains the query "pipeline1". The main content area displays a "Copy data1" activity, which is identified as an "Azure Data Factory Copy Activity". This activity is part of a pipeline named "pipeline1" in the "practicumdf" factory. The activity is connected to two "Product" datasets, one on the left and one on the right. On the left, there is a detailed view of the "Product" dataset, showing its columns: "ProductNumber" and "ListPrice". The "Lineage" tab is selected in the navigation bar. A sidebar on the left provides filtering options for "Input datasets" and "Columns", and a search bar for assets or processes. The status bar at the bottom right indicates the activity was updated on September 15, 2021, at 4:00 PM UTC by the Azure Data Factory pipeline.

Continue to drill-through elements to better understand the capture of Lineage data.

Configure Data Share

Navigate to your instance of Purview in the Azure portal.

Click the “Open Purview Studio” button.

Click the **Management** icon on the left-hand navigation and then “Data Share” from the resulting menu.

Click the “+ New” button on the resulting page.

New Data Share connections

Each Data Share account can connect to only one Purview account.

Azure Subscription
rchapter

Data Share *
practicumds

1 selected

Data Share	Existing connection
practicumds	-

OK Cancel

On the “New Data Factory connections” pop-out, enter values for the following items:

Azure Subscription	Select your subscription	Self-expl
Data Share	Select your instance of data share	Select th

Click the **OK** button.

Article: [Connect to Azure Data Share - Azure Purview | Microsoft Docs](#)

Note from this article... “For Data Share assets to show in Purview, a snapshot job must be run after you connect your Data Share to Purview”



Appendix

Each of the topics in this section describe creation of basic resources necessary to complete other exercises.

Sections are not sequential.

Consider instantiating only those resources required by a specific exercise.

Application Registration (aka Service Principal)

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Data Lake (with container)

Navigate to “Azure Active Directory” > “App Registrations”.

The screenshot shows the Microsoft Azure Active Directory (Preview) interface. The left sidebar is titled "Manage" and includes options like Overview, Preview features, Diagnose and solve problems, Users, Groups, External Identities, Roles and administrators, Administrative units, Enterprise applications, Devices, App registrations (which is selected), and Identity Governance. The main content area is titled "Microsoft | App registrations" and shows a warning message: "If you are building an application for external users that will be distributed by Microsoft, you must register as a first party application to meet all security, privacy, and compliance policies. Read our decision guide." Below this, there are tabs for All applications, Owned applications (which is selected), and Deleted applications (Preview). A search bar says "Start typing a name or Application ID to filter these results". A message states "This account isn't listed as an owner of any applications in this directory." A blue button labeled "View all applications in the directory" is visible.

Click the “+ New Registration” button.

The screenshot shows the 'Register an application' page in the Microsoft Azure portal. The URL is https://ms.portal.azure.com/#blade/Microsoft_AAD_IAM/ActiveDirectoryMenuBlade/RegisterApplication. The page has a dark theme. At the top, there's a navigation bar with 'Microsoft Azure (Preview)', a search bar, and a user profile for rchapler@microsoft.com.

Name: The input field contains 'practicum'. A green checkmark icon is visible to the right of the input field.

Supported account types: The 'Accounts in this organizational directory only (Microsoft only - Single tenant)' option is selected. Other options include 'Any Azure AD directory - Multitenant', 'Any Azure AD directory - Multitenant and personal Microsoft accounts (e.g. Skype, Xbox)', and 'Personal Microsoft accounts only'. A 'Help me choose...' link is available below the list.

Redirect URI (optional): The dropdown menu is set to 'Web' and the input field contains 'e.g. https://example.com/auth'.

Enterprise applications: A note says 'Register an app you're working on here. Integrate gallery apps and other apps from outside your organization by adding from Enterprise applications.'

Agreement: A note states 'By proceeding, you agree to the Microsoft Platform Policies' with a link.

Register button: A blue 'Register' button is located at the bottom left of the main form area.

On the “Register an application” page, enter values for the following items:

Name	Enter a value that is meaningful and aligned with standards
------	---

No additional settings are required for this exercise.

Click the **Register** button.

The screenshot shows the Microsoft Azure portal interface. In the top left, there's a logo for 'practicumar'. The top navigation bar includes links for 'Dashboard', 'Microsoft', and the current application 'practicumar'. A search bar at the top right says 'Search resources, services, and docs (G+)'. On the far right, there's a user profile for 'rchapler@microsoft.com'.

The main content area has a dark header with buttons for 'Delete', 'Endpoints', and 'Preview features'. Below this, a message says 'Got a second? We would love your feedback on Microsoft identity platform (previously Azure AD for developer). →'. The 'Essentials' section contains the following details:

Setting	Value
Display name	practicumar
Application (client) ID	[REDACTED]
Object ID	062d3984-d6b4-4f83-abcd-35d68d49a955
Directory (tenant) ID	[REDACTED]
Supported account types	My organization only

On the left sidebar under 'Manage', the 'Application (client) ID' is highlighted. Other options include 'Branding', 'Authentication', 'Certificates & secrets', 'Token configuration', 'API permissions', and 'Expose an API'.

Make note of the “Application (client) ID” and “Directory (tenant) ID” values for use in the next section.

Client Identifier

Navigate to the App Registration, **Overview** page.

This screenshot is identical to the one above, showing the 'practicumar' application registration in the Microsoft Azure portal. The 'Essentials' section displays the same information: Display name 'practicumar', Application (client) ID '[REDACTED]', Object ID '062d3984-d6b4-4f83-abcd-35d68d49a955', Directory (tenant) ID '[REDACTED]', and Supported account types 'My organization only'. The left sidebar shows the 'Manage' section with 'Application (client) ID' selected.

Copy the “Directory (tenant) ID” and “Application (client) ID” values.

Key Vault Secrets, Tenant and Client Identifiers

Open a new tab and navigate to your instance of Key Vault.

Click **Secrets** in the **Settings** group of the left-hand navigation.

Click the “+ Generate/Import” button.

The screenshot shows the 'Create a secret' page in the Microsoft Azure (Preview) portal. The 'Name' field is set to 'practicumar-clientid'. The 'Value' field contains a redacted value. The 'Enabled' switch is set to 'Yes'. The 'Create' button is visible at the bottom.

On the “Create a secret” page, enter values for the following items:

Name	Enter a value that is meaningful and aligned with standards
Value	Paste the previously copied value

No additional changes are required.

Click the **Create** button.

Repeat for Tenant Id.

Client Secret

In the **Manage** grouping, click “Certificates & secrets”.

Navigate to “Certificates & secrets” in the **Manage** group in the left-hand navigation.

The screenshot shows the Microsoft Azure portal interface. The URL in the address bar is https://ms.portal.azure.com/#blade/Microsoft_AAD_RegisteredApps/ApplicationMenuBlade/Credentials/q.... The top navigation bar includes the Microsoft Azure logo, a search bar, and a user profile for rchapler@microsoft.com.

The main content area is titled "practicumar | Certificates & secrets". On the left, there's a sidebar under "Manage" with the following items:

- Overview
- Quickstart
- Integration assistant
- Branding
- Authentication
- Certificates & secrets** (selected)
- Token configuration
- API permissions
- Expose an API
- App roles
- Owners
- Roles and administrators | Preview
- Manifest

Certificates

Certificates can be used as secrets to prove the application's identity when requesting a token. Also can be referred to as public keys.

[Upload certificate](#)

Thumbprint	Start date	Expires	Certificate ID
No certificates have been added for this application.			

Client secrets

A secret string that the application uses to prove its identity when requesting a token. Also can be referred to as application password.

[New client secret](#)

Description	Expires	Value	Secret ID
No client secrets have been created for this application.			

Click the “+ New client secret” button, and then in the resulting “Add a client secret” popout, click the Add button.

Certificates

Certificates can be used as secrets to prove the application's identity when requesting a token. Also can be referred to as public keys.

Client secrets

A secret string that the application uses to prove its identity when requesting a token. Also can be referred to as application password.

Description	Expires	Value	Secret ID
Password uploaded on Fri Sep ...	3/24/2022	[REDACTED]	[REDACTED]

Copy Value.

Key Vault Secret, Client Secret

Open a new tab and navigate to your instance of Key Vault.
Click **Secrets** in the **Settings** group of the left-hand navigation.
Click the “+ Generate/Import” button.

The screenshot shows the Microsoft Azure portal with the URL <https://ms.portal.azure.com/#@microsoft.onmicrosoft.com/resource/subscriptions>. The user is in the 'Create a secret' section under 'Microsoft Azure (Preview)'. The 'Name' field is set to 'practicumar-clientsecret' and the 'Value' field contains redacted text. The 'Enabled' switch is set to 'Yes'. There are also options for 'Set activation date' and 'Set expiration date', both of which are unchecked. The 'Content type (optional)' field is empty. At the bottom, there is a 'Create' button.

On the “Create a secret” page, enter values for the following items:

Name	Enter a value that is meaningful and aligned with standards
Value	Paste the previously copied value

No additional changes are required.

Click the **Create** button.

Permissions, Data Lake

Navigate to your instance of data lake, your container and then “Access Control (IAM)” in left-hand navigation.

Click the “+ Add” button and then “Add role assignment” in the resulting drop-down.

The screenshot shows the Microsoft Azure (Preview) portal interface. On the left, the navigation pane includes 'Overview', 'Diagnose and solve problems', 'Access Control (IAM)', 'Settings' (with options like 'Shared access tokens', 'Manage ACL', 'Access policy', 'Properties', 'Metadata', and 'Editor (preview)'), and a search bar. The main content area shows 'practicumdlc | Access Control (IAM)' for a 'Container'. It displays 'Number of role assignments for this subscription' (13 total, 2000 limit). A table lists roles: Contributor, Owner, Reader, Storage Blob Data Contributor, Storage Blob Data Reader, and User Access Administrator. On the right, a 'Role assignments' tab is active, showing the 'Add role assignment' dialog. The dialog has fields for 'Role' (set to 'Storage Blob Data Reader'), 'Assign access to' (set to 'User, group, or service principal'), and 'Select' (showing 'practicumar'). Below this, it says 'No users, groups, or service principals found.' At the bottom of the dialog are 'Save' and 'Discard' buttons.

On the “Add role assignment” pop-out, enter values for the following items:

Role	Select “Storage Blob Data Reader”
Assign access to	Confirm default selection, “User, group, or service principal”
Select	Search for and select your application registration

Click the **Save** Button.

CLI

Browse to [Install the Azure CLI for Windows | Microsoft Docs](#)

The screenshot shows a Microsoft Edge browser window displaying the Azure CLI Windows installation documentation. The URL is <https://docs.microsoft.com/en-us/cli/azure/install-azure-cli-windows?tabs=azure-cli>. The page title is "Install Azure CLI on Windows". On the left, there's a navigation sidebar with links like "Overview", "Install, update and run", "Install - Windows", "Install - macOS", "Install - Linux", "Install - beta", "Update", "Sign in", "Tutorials", "Samples", "How-to guides", "Reference Summaries", and "Reference". The main content area has a heading "Install or update" and a sub-section "Azure CLI current version" with a blue button labeled "Current release of the Azure CLI". On the right, there are sections for "Is this page helpful?", "In this article", and "Next Steps".

Download and install the current release of the Azure CLI.

Data Explorer

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Resource Group

The screenshot shows the 'Create an Azure Data Explorer Cluster' page in Microsoft Azure (Preview). The page has a dark theme and includes the following sections:

- PROJECT DETAILS**: A note about managing resources and a dropdown for Subscription (set to 'rchapler') and Resource group (set to 'practicumrg').
- CLUSTER DETAILS**: Fields for Cluster name ('practicumdec'), Region ('West US 2'), and a checked checkbox for 'Enable performance update (EngineV3)'.
- COMPUTE SPECIFICATION**: A nested section with Workload ('Compute optimized'), Size ('Extra Small (2 cores)'), and Compute specifications ('Standard_D11_v2').
- Availability zones**: A dropdown set to '(none)'.
- Buttons**: 'Review + Create' (highlighted in blue) and 'Next : Scale >'.

On the “Create an Azure Data Explorer Cluster” page, enter values for the following form items:

Subscription	Select your subscription
Resource Group	Select your Resource Group
Cluster Name	Enter a value that is meaningful and aligned with standards
Region	Select the value used during Resource Group creation
Enable Performance...	Leave checked
Workload	Select “Compute optimized” from the drop-down
Size	Select “Extra Small (2 cores)” from the drop-down
Compute Specifications	This should be auto populated with “Standard_D11_v2” based on the Workload and Size selections
Availability Zones	Confirm default selection, “(none)”

Review settings on remaining tabs {e.g., Tags}.

No additional settings are required for this exercise.

Click the “Review + create” button, validate, and then click the **Create** button.

Add Database

The screenshot shows the Microsoft Azure (Preview) portal interface. On the left, there is a navigation sidebar for the 'practicumdec' cluster, which includes options like Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Permissions, Query, Settings, Scale up, Scale out, and Configurations. In the center, a message says 'To use Azure Data Explorer, create at least one database.' Below it, under 'Essentials', are details such as Resource group (change) 'practicumrg', Location 'West US 2', Subscription (change) 'rchapler', Subscription ID '91e9fddc-ef15-416c-9be1-085f8b1b46ed', and Engine type 'V3'. On the right, a modal window titled 'Azure Data Explorer Database' is open, prompting the user to 'Create new database'. It shows the Admin email 'rchapler@microsoft.com; Microsoft' and the 'Database name' field containing 'practicumded'. Below that are fields for 'Retention period (in days)' set to '3650' and 'Cache period (in days)' set to '31'. At the bottom of the modal is a blue 'Create' button.

In the newly created Data Explorer Cluster, click the “+ Add database” button.

On the “Create an Azure Data Explorer Database” pop-out, enter values for the following form items:

Database Name	Enter a value that is meaningful and aligned with standards
Retention Period (in days)	Confirm the default value, 3650
Cache Period (in days)	Confirm the default value, 31

Click the **Create** button.

Sample #1, Product Table

Navigate to the newly created Data Explorer database.

Click **Query** in the left-hand navigation.

The screenshot shows the Microsoft Azure (Preview) interface for a Data Explorer cluster named 'practicumdec'. The left sidebar includes options like Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Permissions, Query, Settings, Scale up, Scale out, and Configurations. The main area displays a query editor with the following KQL code:

```
.create table Product ( ProductNumber: string, ListPrice: decimal )
```

The results pane shows a table named 'Table 1' with one record:

TableName	Schema	DatabaseName	Folder	DocString
Product	{"Name": "Product", "OrderedColumns": [{"Name": "Pr..."}]}	practicumdec		

Execute the following KQL query to create a new table named Product:

```
.create table Product ( ProductNumber: string, ListPrice: decimal )
```

Sample #2, StormEvents Data

Follow the instructions in the “Quickstart: Ingest sample data into Azure Data Explorer” article (<https://docs.microsoft.com/en-us/azure/data-explorer/ingest-sample-data>) to populate sample data that we can surface in Power Apps.

The screenshot shows the Azure Data Explorer interface for a cluster named 'practicumdec.westus2'. The left sidebar includes Data, Query, and Dashboards (Preview). The main area displays a query editor with the following KQL code:

```
.create table StormEvents (StartTime: datetime, EndTime: datetime, EpisodeId: int, EventId: int, State: string, EventType: string)
.ingest into table StormEvents 'https://kustosamplefiles.blob.core.windows.net/sampledata/StormEvents.kusto'
StormEvents | sort by StartTime desc | take 10
```

The results pane shows a table named 'Table 1' with 10 records:

StartTime	EndTime	EpisodeId	EventId	State	EventType
2007-12-31 23:53:00.0000	2007-12-31 23:53:00.0000	12,037	65,839	CALIFORNIA	High Wind
2007-12-31 23:53:00.0000	2007-12-31 23:53:00.0000	12,037	65,838	CALIFORNIA	High Wind
2007-12-31 22:30:00.0000	2007-12-31 23:59:00.0000	12,950	71,590	MICHIGAN	Winter Storm
2007-12-31 22:30:00.0000	2007-12-31 23:59:00.0000	12,950	71,588	MICHIGAN	Winter Storm
2007-12-31 22:30:00.0000	2007-12-31 23:59:00.0000	12,950	71,589	MICHIGAN	Winter Storm

Familiarize yourself with the resulting data for use in later sections.

Data Factory

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Key Vault
- Resource Group

The screenshot shows the Microsoft Azure portal with the URL <https://ms.portal.azure.com/#create/Microsoft.DataFactory>. The page title is "Create Data Factory". The navigation bar includes "Microsoft Azure (Preview)" and the user's email "rchapler@microsoft.com". The main content area is titled "Create Data Factory" and has a "Basics" tab selected. The "Project details" section asks to select a subscription and resource group. The "Subscription" dropdown is set to "rchapler" and the "Resource group" dropdown is set to "practicumrg". The "Instance details" section includes fields for "Region" (West US 2), "Name" (practicumdf), and "Version" (V2). At the bottom, there are buttons for "Review + create", "< Previous", and "Next : Git configuration >".

On the “Create Data Factory” page, **Basics** tab, enter values for the following items:

Subscription	Select your subscription
Resource Group	Select your Resource Group
Region	Select the value used during Resource Group creation
Name	Enter a value that is meaningful and aligned with standards
Version	Confirm default selection, “V2”

Click the “Next: Git configuration >” button.

The screenshot shows the Microsoft Azure (Preview) portal with the URL <https://ms.portal.azure.com/#create/Microsoft.DataFactory>. The page title is "Create Data Factory". The "Git configuration" tab is active. A note states: "Azure Data Factory allows you to configure a Git repository with either Azure DevOps or GitHub. Git is a version control system that allows for easier change tracking and collaboration." Below this is a link to "Learn more about Git integration in Azure Data Factory". A checkbox labeled "Configure Git later" is checked. At the bottom, there are buttons for "Review + create", "< Previous", and "Next : Networking >".

On the “Create Data Factory” page, “Git configuration” tab, enter values for the following items:

Configure Git later	Check the box
---------------------	---------------

Review settings on remaining tabs {e.g., Networking, Advanced, Tags}. No additional settings are required for this exercise.

Click the “Review + create” button, review configuration, and then click the **Create** button.

Add Access Policy to Key Vault

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Key Vault

Navigate to your instance of Key Vault.

The screenshot shows the Microsoft Azure Key Vault settings page for the 'practicumkv' vault. The left sidebar shows navigation options like Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Events, and Settings (Keys, Secrets, Certificates, Access policies). The main content area displays the 'Essentials' section with details such as Resource group (practicumrg), Location (West US 2), Subscription (rchapler), and Directory ID (redacted). The 'Access control (IAM)' section is expanded, showing Vault URI (https://practicumkv.vault.azure.net/), Sku (Standard), and other settings like Soft-delete (Enabled) and Purge protection (Disabled). A 'Tags (change)' section is also present.

Click “Access Policies” in the **Settings** group of the left-hand navigation.

The screenshot shows the 'Access policies' page for the 'practicumkv' vault. The left sidebar includes 'Access policies' under the Settings group. The main area shows the 'Enable Access to:' section with checkboxes for Azure Virtual Machines, Azure Resource Manager, and Azure Disk Encryption. It also shows the 'Permission model' section with 'Vault access policy' selected. Below these, there's a '+ Add Access Policy' link and a table for 'Current Access Policies' with one row named 'APPLICATION'.

Click the “+ Add Access Policy” link.

Principal

Select a principal

practicumdf
a383d929-ecd9-47cf-812c-3053a7dbff87
Selected

Selected items

practicumdf
a383d929-ecd9-47cf-812c-3053a7dbff87

Add

Select

On the “Add access policy” page, enter values for the following items:

Secret Permissions Select **Get** and **List** in the drop-down

Select Principal Search for, and then **Select** the managed identity for your Data Factory

Click the **Add** button.

practicumkv | Access policies

Key vault

Search (Ctrl+ /)

Save Discard Refresh

Please click the 'Save' button to commit your changes.

Enable Access to:

Azure Virtual Machines for deployment ⓘ

Azure Resource Manager for template deployment ⓘ

Azure Disk Encryption for volume encryption ⓘ

Permission model

Vault access policy

Azure role-based access control

+ Add Access Policy

Current Access Policies

Name	Email	Key Permissions	Secret Permissions	Certificate Permissions	Action
APPLICATION	practicumdf	0 selected	2 selected	0 selected	Delete

Click the **Save** button.

Add Permissions to Data Explorer

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Data Explorer

Navigate to your instance of Data Explorer.

The screenshot shows the Microsoft Azure portal interface. On the left, there's a navigation menu for an 'Azure Data Explorer Cluster' named 'practicumdec'. The 'Permissions' option is selected. In the main content area, there's a 'New Principals' dialog box. The dialog has a search bar containing 'practicumdf', a 'Role' dropdown set to '2 selected', and a 'Selected' item listed under 'Selected items'. At the bottom of the dialog is a 'Select' button.

Click **Permissions** in the left-hand navigation.

Click “**+ Add**” on the resulting page.

Search for and then click to **Select** your instance of Data Factory.

Data Lake

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Resource Group
- Storage Account

Create storage account

Basics Networking Data protection Advanced Tags Review + create

Azure Storage is a Microsoft-managed service providing cloud storage that is highly available, secure, durable, scalable, and redundant. Azure Storage includes Azure Blobs (objects), Azure Data Lake Storage Gen2, Azure Files, Azure Queues, and Azure Tables. The cost of your storage account depends on the usage and the options you choose below.

[Learn more about Azure storage accounts](#)

Project details

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription * rchapler

Resource group * practicumrg [Create new](#)

Instance details

The default deployment model is Resource Manager, which supports the latest Azure features. You may choose to deploy using the classic deployment model instead. [Choose classic deployment model](#)

Storage account name * practicuumdl

Location * (US) West US 2

Performance Standard

Account kind StorageV2 (general purpose v2)

Replication Read-access geo-redundant storage (RA-GRS)

[Review + create](#) [< Previous](#) [Next : Networking >](#)

On the “Create Storage Account” page, enter values for the following items:

Subscription	Select your subscription
Resource Group	Select your Resource Group
Storage Account Name	Enter a value that is meaningful and aligned with standards
Location	Select the value used during Resource Group creation
Performance	Confirm default radio button selection, “Standard”
Account Kind	Confirm default selection, “StorageV2 (general purpose v2)”
Replication	Confirm default selection, “Read-access geo-redundant storage (RA-GRS)”

Navigate to the **Advanced** tab.

The screenshot shows the 'Create storage account' page in Microsoft Azure. The 'Advanced' tab is selected. The configuration is as follows:

- Security**:
 - Secure transfer required: Enabled
 - Allow shared key access: Enabled
 - Minimum TLS version: Version 1.2
 - Infrastructure encryption: Disabled (with a note: "Sign up is currently required to enable infrastructure encryption on a per-subscription basis. [Sign up for infrastructure encryption](#)")
- Blob storage**:
 - Allow Blob public access: Enabled
 - Blob access tier (default): Hot
 - NFS v3: Disabled (with a note: "Sign up is currently required to utilize the NFS v3 feature on a per-subscription basis. [Sign up for NFS v3](#)")
- Data Lake Storage Gen2**:
 - Hierarchical namespace: Enabled

At the bottom, there are buttons for 'Review + create', '< Previous', and 'Next : Tags >'.

Click on the **Enabled** radio button in the “Data Lake Storage Gen2” grouping.

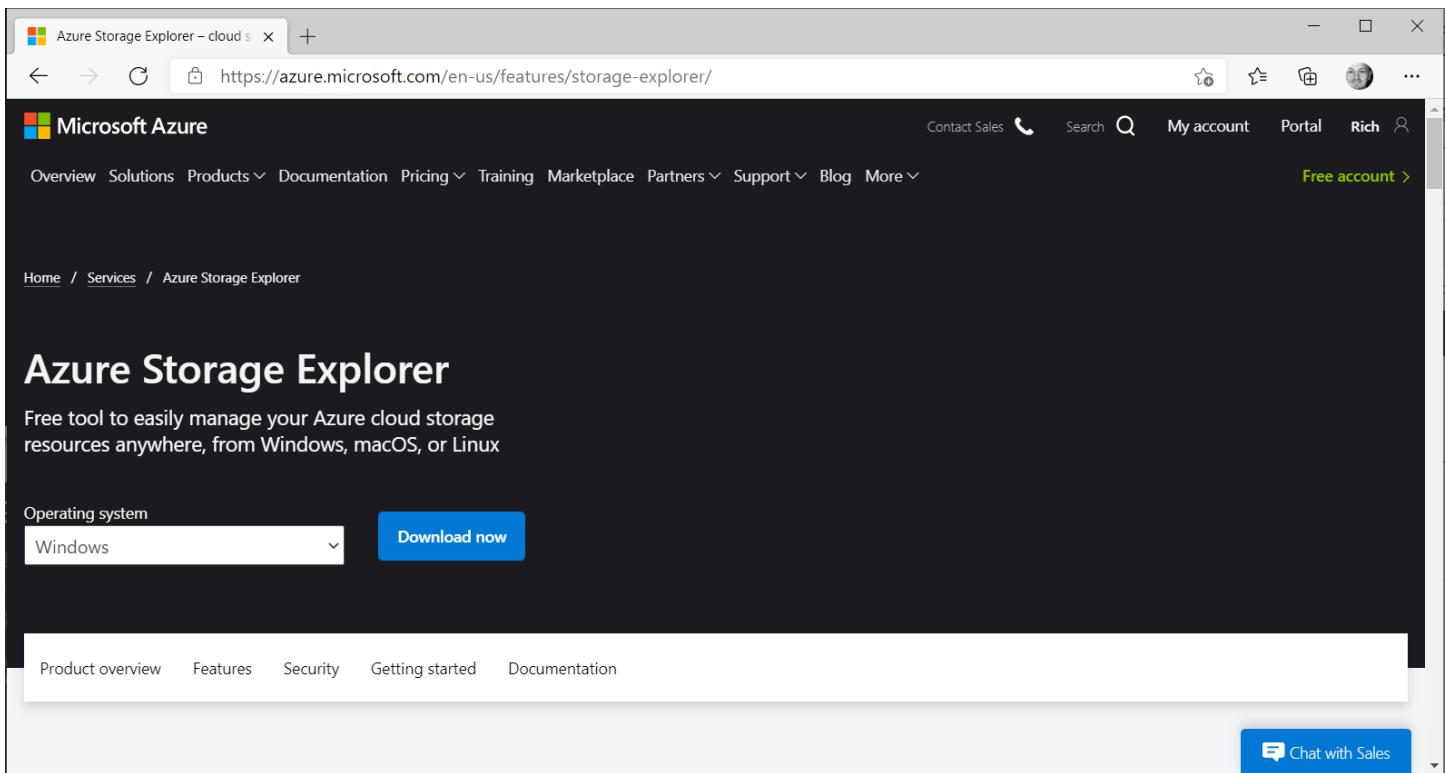
Review settings on remaining tabs {e.g., Networking, Data Protection, Tags}.

No additional settings are required for this exercise.

Click the “Review + create” button, review configuration, and then click the **Create** button.

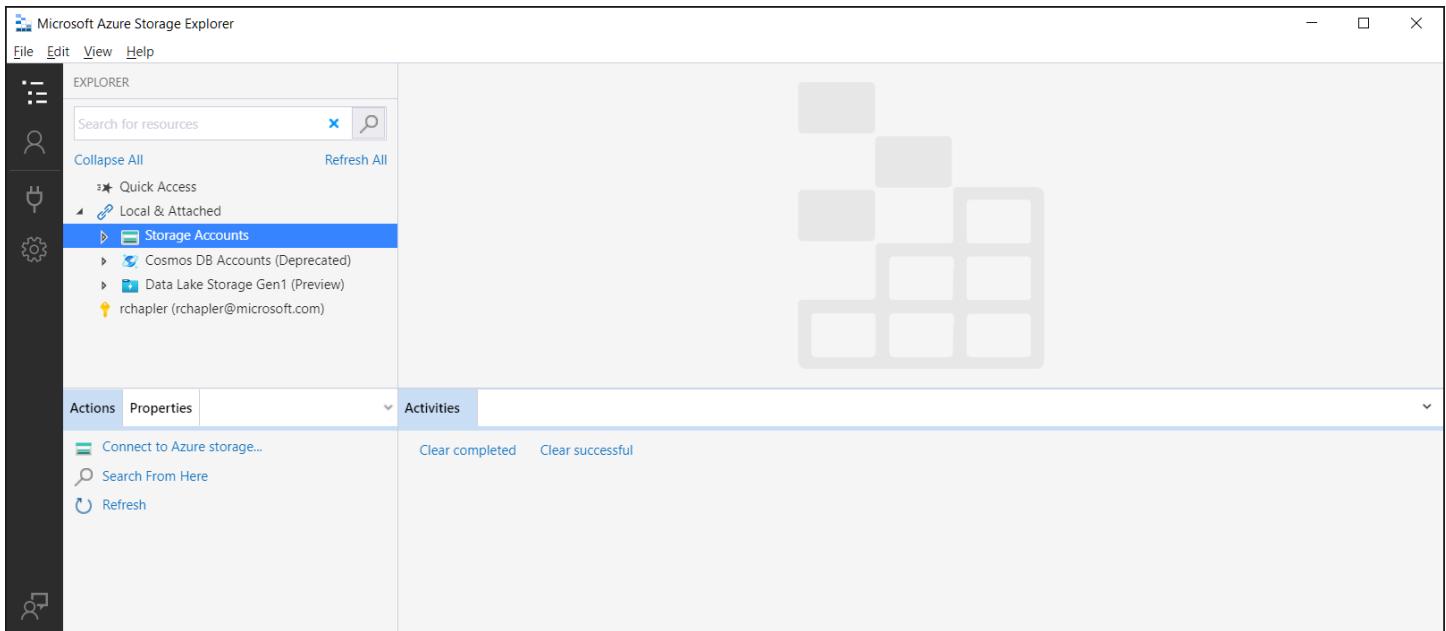
Storage Explorer

Download and install the Microsoft Azure Storage Explorer app ([Azure Storage Explorer – cloud storage management | Microsoft Azure](#)).



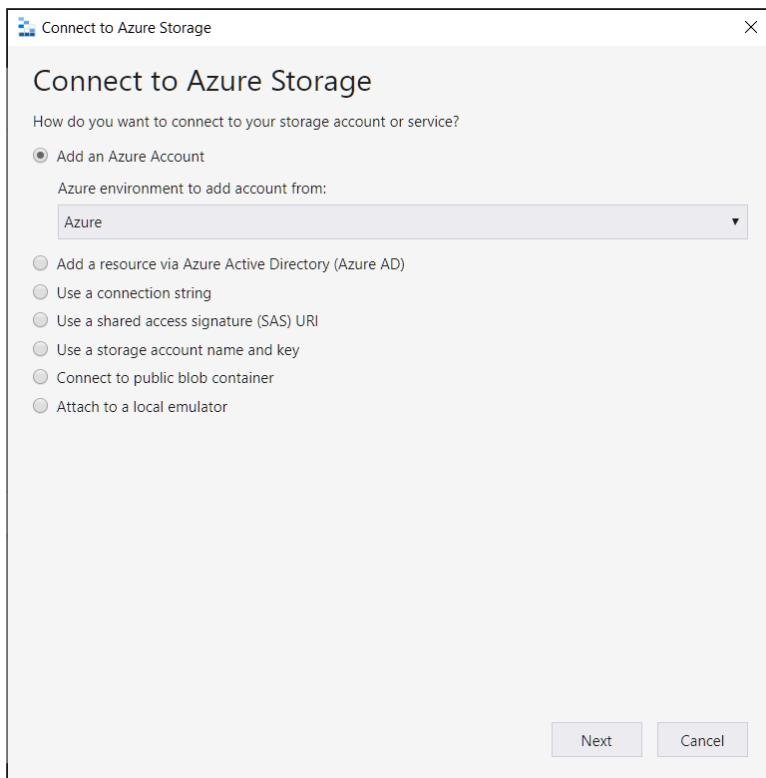
The screenshot shows the Microsoft Azure website with the URL <https://azure.microsoft.com/en-us/features/storage-explorer/>. The page title is "Azure Storage Explorer – cloud storage management". The top navigation bar includes links for Overview, Solutions, Products, Documentation, Pricing, Training, Marketplace, Partners, Support, Blog, More, Contact Sales, Search, My account, Portal, Rich, and Free account. Below the navigation is a breadcrumb trail: Home / Services / Azure Storage Explorer. The main heading is "Azure Storage Explorer" with the subtext "Free tool to easily manage your Azure cloud storage resources anywhere, from Windows, macOS, or Linux". A dropdown menu for "Operating system" is set to "Windows", with a "Download now" button next to it. At the bottom of the page are links for Product overview, Features, Security, Getting started, Documentation, and a "Chat with Sales" button.

Launch the app and provide Azure credentials.



The screenshot shows the Microsoft Azure Storage Explorer application window. The menu bar includes File, Edit, View, and Help. The left sidebar has icons for Quick Access, Local & Attached, Storage Accounts (which is selected and highlighted in blue), Cosmos DB Accounts (Deprecated), Data Lake Storage Gen1 (Preview), and a user account (rchapler). The main pane displays a grid of placeholder icons for storage accounts. The bottom navigation bar has tabs for Actions, Properties, Activities (which is selected and highlighted in blue), and a status message: "Clear completed" and "Clear successful".

Click the “Connect to Azure storage...” link.



On the resulting pop-up, click the “**Add an Azure Account**” radio button. Click the **Next** button. Authenticate with Azure credentials.

Container

Continue with the “**Microsoft Azure Storage Explorer**” app. Use the search bar to focus on your Storage Account.

Right-click on “**Blob Containers**” and click “**Create Blob Container**” in the resulting pop-up menu. Enter a meaningful name.

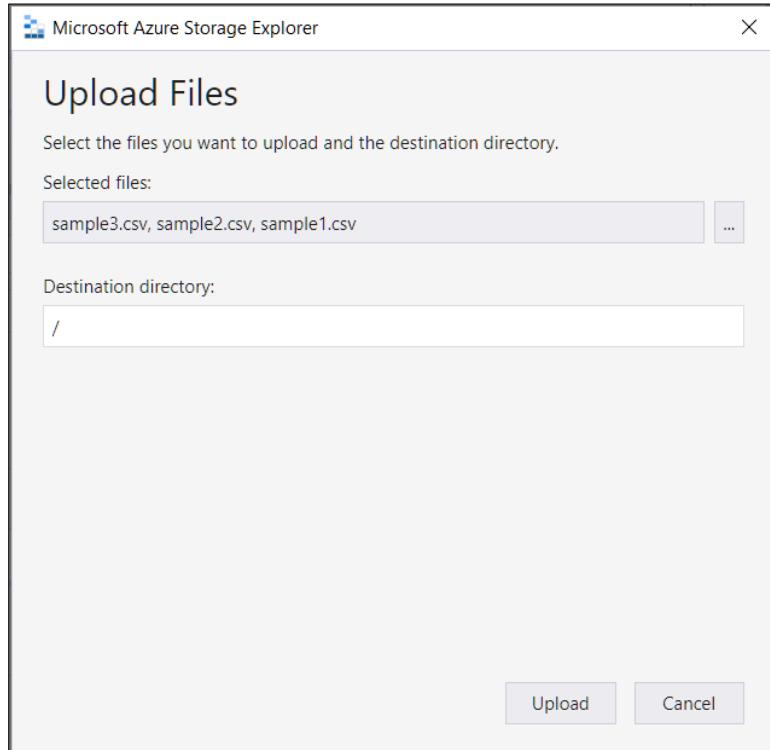
Sample Data

Download sample files from a site like [Sample CSV Files Download - Get Examples Instantly \(filesamples.com\)](http://filesamples.com)

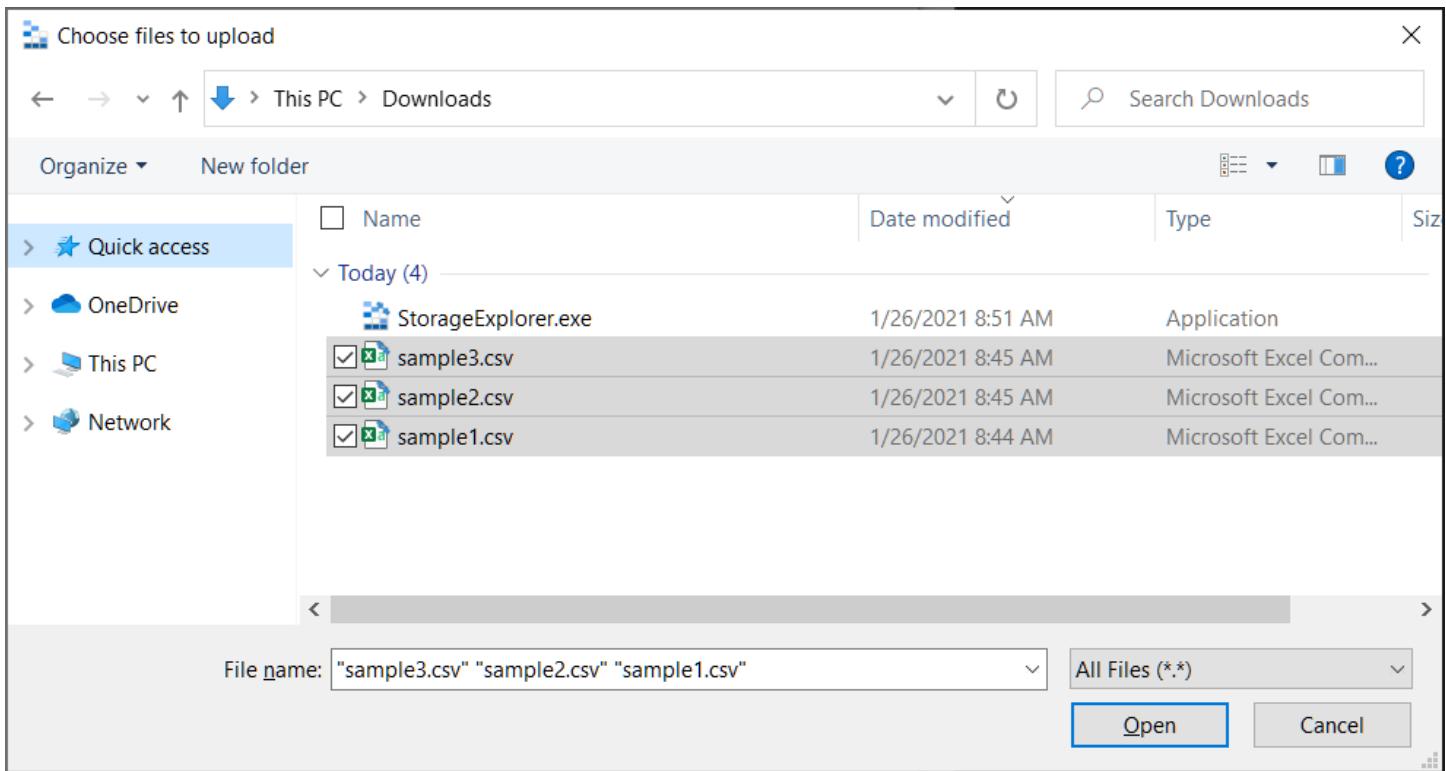
The screenshot shows a Microsoft Edge browser window with two tabs open: "practicumsa - Microsoft Azure" and "Sample CSV Files Download - Ge...". The current page is "https://filesamples.com/formats/csv". The main content is titled "Sample CSV Files Download" and features a "CSV" icon and the text "Comma Separated Values". It describes the purpose of the page: "Below you will find a selection of sample .csv document files for you to download. On the right there are some details about the file such as its size so you can best decide which one will fit your needs." Three CSV files are listed: "sample1.csv" (CSV / 502.00 B), "sample2.csv" (CSV / 55.56 KB), and "sample3.csv" (CSV / 723.00 B). Each entry has a "Download" button. A sidebar on the right contains an advertisement for "Ten-X Commercial" with the text "Extend Your Reach. Transact With Certainty. The End-To-End Solution For CRE Investment." and a "Learn More" link.

Return to the “Microsoft Azure Storage Explorer” app and navigate to your container. Click the **Upload** button and select “Upload Files” in the resulting pop-up menu.

Click the ellipses button to the right of the “Selected Files” box.



Navigate to **Downloads** and select your sample data files.



Click the **Open** button.

On the “Upload Files” pop-up, click the **Upload** button. Allow time for file transfer.

A screenshot of the Microsoft Azure Storage Explorer application. The left sidebar shows a file structure under "practicumdl": "rchapler (rchapler@microsoft.com)" and "practicumdl (ADLS Gen2)" which contains "Blob Containers" and "practicumdlc". The main pane is titled "practicumdlc" and shows a list of files: "sample1.csv", "sample2.csv", and "sample3.csv". The table has columns: Name, Access Tier, Access Tier Last Modified, Last Modified, Blob Type, Content Type, Size, and Lease State. All files are listed as "Hot (inferred)". The "Actions" tab is selected in the bottom left. At the bottom, there is an "Activities" section showing a completed transfer task: "Transfer from 'C:\Users\rchapler\Downloads\' to 'practicumdlc/' complete: 3 items transferred (used SAS, discovery completed)" with a timestamp of "Started at 2/3/2021, 12:11:00 PM Duration: 2 seconds". There is also a "Copy AzCopy Command to Clipboard" link.

Data Share

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Data Explorer (2 instances)
- Resource Group

The screenshot shows the Microsoft Azure portal's "Create Data Share" wizard. The "Basics" tab is active. In the "Project details" section, the subscription dropdown is set to "rchapler" and the resource group dropdown is set to "practicumrg". In the "Instance details" section, the location dropdown is set to "West US 2" and the name input field contains "practicumds". At the bottom, there are navigation buttons: "Review + create" (highlighted in blue), "< Previous", and "Next : Tags >".

On the “Create Data Share” page, enter values for the following items:

Subscription	Select your subscription
Resource Group	Select your Resource Group
Location	Select the value used during Resource Group creation
Name	Enter a value that is meaningful and aligned with standards

Review settings on remaining tabs.

No additional settings are required for this exercise.

Click the “Review + create” button, validate, and then click the **Create** button.

After deployment is complete, click the “Go to resource” button.

The screenshot shows the Microsoft Azure Data Share blade for the 'practicumds' resource. The left sidebar contains navigation links for Overview, Activity log, Access control (IAM), Tags, and Diagnose and solve problems. Under Settings, there are links for Properties and Locks. The main content area is titled 'Essentials' and displays the following information: Resource group (practicumrg), Location (West US 2), Subscription (rchapler), and Subscription ID. Below this, there is a section for Tags with a link to 'Click here to add tags'. At the bottom, there are two large buttons: 'Start sharing your data' and 'View received shares'.

Sample Share

Note: To demonstrate sharing data to a target, you must instantiate a second ADX cluster.

Click the “**Start sharing your data**” button.

On the resulting “**Sent Shares**” page, click the “**+ Create**” button.

The screenshot shows the Microsoft Azure (Preview) portal with the URL <https://ms.portal.azure.com/#@microsoft.onmicrosoft.com/resource/subscriptions>. The user is in the 'practicumds' subscription under the 'Microsoft.DataShare-20210915092919' resource group. The left sidebar shows the 'Data Share' section with 'Sent Shares' selected. The main pane displays the 'practicumds | Sent Shares' page, specifically the '1. Details' tab of the 'Create Data Share' wizard. The 'Share name' field is populated with 'practicumdec'. The 'Share type' dropdown is set to 'In-place'. The 'Description' and 'Terms of use' fields both have placeholder text: 'Enter description for the share' and 'Enter terms of use for the share' respectively. At the bottom of the form are 'Cancel' and 'Continue' buttons.

On the “Sent Shares” page, “1. Details” tab, enter values for the following items:

Share Name	Enter a value that is meaningful and aligned with standards
Share Type	Select “In-place” from the drop-down menu

Click the **Continue** button.

practicumds | Sent Shares

Dashboard > Microsoft.DataShare-20210915092919 > practicumds

1. Details 2. Datasets 3. Recipients 4. Review + Create

Search (Ctrl+ /)

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

Settings
Properties
Locks

Data Share
Sent Shares

Add datasets

Select datasets to be shared. You must have permission to add role assignment to the data store. This permission exists in the Owner role. See Share Your Data tutorial for details

Previous Continue

On the “Sent Shares” page, “2. Datasets” tab, click the “Add datasets” button.

practicumds | Sent Shares

1. Details 2. Datasets 3. Recipients

Select dataset type

Azure Data Explorer

Cancel Next

On the resulting “Select dataset type” pop-out, click the “Azure Data Explorer” button, then click the **Next** button.

The screenshot shows the Microsoft Azure (Preview) portal interface. On the left, a sidebar for the 'practicumds' Data Share is visible, containing sections for Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Settings, Properties, Locks, Data Share (Sent Shares, Received Shares, Feedback), Monitoring (Alerts, Metrics, Diagnostic settings, Logs), and a search bar. The main area shows a step-by-step process: 1. Details, 2. Datasets (selected), and 3. Recipients. A 'Azure Data Explorer' pop-up window is open, prompting the user to 'Select datasets'. It includes fields for Subscriptions (set to 'rchapler'), Resource groups (set to 'practicumrg'), and Azure data explorer clusters (set to 'practicumdec2'). Navigation buttons 'Previous' and 'Next' are at the bottom of the pop-up.

On the resulting “**Azure Data Explorer**” pop-out, enter values for the following items:

Subscriptions	Select your subscription
Resource Groups	Select your Resource Group
Azure Data Explorer Clusters	Select your Data Explorer cluster
Name	Enter a value that is meaningful and aligned with standards

Click the **Next** button.

The screenshot shows the Microsoft Azure portal interface for creating a Data Share named "practicumds". The main page has a left sidebar with options like Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Properties, Locks, and Sent Shares. The "Sent Shares" option is currently selected. The main content area is divided into three steps: 1. Details, 2. Datasets (which is active), and 3. Recipients. A modal window titled "Azure Data Explorer" is open, showing the datasets available in the target cluster "practicumdec2". One dataset, "practicumdec2", is selected. At the bottom of the modal are "Previous" and "Next" buttons.

On the second "...Select datasets" pop-out, check the box next to your cluster and then click the **Next** button.

The screenshot shows the "Azure Data Explorer" modal window from the previous step, now titled "Rename datasets". It contains instructions: "Provide names for your datasets. This is the name that your data consumer will see when they accept the share. Dataset names must be unique." Below this, there is a table with two columns: "Dataset name" and "Path". A single row is present in the table, showing "practicumdec2" in both columns. At the bottom of the modal are "Previous" and "Add datasets" buttons.

On the "...Rename datasets" pop-out, confirm the automatically generated "Dataset name" value.
Click the "Add datasets" button.

practicumds | Sent Shares

2. Datasets

Datasets ↑	Type	Path
practicumdec2	Azure Data Explorer Cluster	practicumdec2/

Previous Continue

Back on the “Sent Shares” page, “2. Datasets” tab, click the **Continue** button.

practicumds | Sent Shares

3. Recipients

Enter email address of the recipients for the share. Please ensure you are using recipient's Azure login email.

Add recipient Delete All Update expiration for all

Email	Share expiration
rchapler@microsoft.com	Fri Oct 15 2021

Previous Continue

On the “Sent Shares” page, “3. Recipients” tab, click the “Add recipient” button.

In the resulting interface, enter an **Email** value.

Click the **Continue** button.

practicumds | Sent Shares

Dashboard > Microsoft.DataShare-20210915092919 > practicumds

1. Details 2. Datasets 3. Recipients 4. Review + Create

Share Contents

Number of datasets 1

Settings

Name of data share practicumdec

Description -

Terms of use -

Recipients

Number of recipients 1

Previous Create

On the “Sent Shares” page, “4. Review + Create” tab, click the **Create** button.

The designated email recipient can expect to receive an email like the one below:

Azure Data Share invitation from Rich Chapler - Message (HTML) Search

File **Message** Help

Delete Respond Share to Teams Quick Steps Move Tags Editing Immersive Translate Zoom Dynamics 365 Insights Report Message Protection Reply with Meeting Poll FindTime

Teams Quick Steps Language Add-in Add-in Protection FindTime

Previous Item Next Item Quick Print

Azure Data Share invitation from Rich Chapler

 Microsoft Azure
To: Rich Chapler

If there are problems with how this message is displayed, click here to view it in a web browser.

 Microsoft Azure

You're invited to access data from Microsoft

You're receiving this email because Rich Chapler from Microsoft wants to share the following data with you.

Share name: practicumdec

Description: NA

[View invitation >](#)

See detailed instructions on how to accept and configure your data share.
If you're new to Azure, [create a free Azure subscription](#).

[f](#) [t](#) [y](#) [in](#)
[Privacy Statement](#)
Microsoft Corporation, One Microsoft Way, Redmond, WA 98052
 Microsoft

Click the “View invitation >” link.

A Data Share Invitations - Microsoft

https://ms.portal.azure.com/#blade/Microsoft_Azure_DataShare/InvitationsBrowseBlade

Microsoft Azure (Preview)

Search resources, services, and docs (G+/)

rchapler@microsoft.com MICROSOFT

Dashboard >

Data Share Invitations

Refresh

Pending invitations sent to your Azure login email are listed. See [accept and receive data tutorial](#) for details.

Invitation	Sender	Company	Status	Received On
practicumdec	Rich Chapler	Microsoft	Pending	9/15/2021 11:02:56 AM

Click the **Invitation** link.

practicumdec - Microsoft Azure

https://ms.portal.azure.com/#blade/Microsoft_Azure_DataShare/InvitationsBrowseBlade

Microsoft Azure (Preview)

Search resources, services, and docs (G+/)

rchapler@microsoft.com MICROSOFT

Dashboard > Data Share Invitations >

practicumdec

Invitation

From	Number of datasets
Rich Chapter	1
Company	Expires on
Microsoft	-
Description	-
Terms of use	-

TARGET DATA SHARE ACCOUNT

Subscription * rchapler

Resource group * practicumrg

Data share account * practicumsd

Received share name * practicumdec

Accept and configure Reject

On the resulting “...Invitation” page, enter values for the following items:

Subscriptions	Select your subscription
Resource Groups	Select your Resource Group
Data Share Account	Select your Data Share account
Received Share Name	Confirm default value

Click the “Accept and configure” button.

The screenshot shows the Microsoft Azure portal interface. The left sidebar has a dark theme with various navigation options. The main content area is titled "practicumds | Received Shares". It displays the following details:

Source share	Provider
practicumdec	Rich Chapler

Below this, it shows:

Provider company	Shared on
Microsoft	11:14:31 AM, 9/15/2021

Other details include:

Number of source datasets	Accepted by
1	Rich Chapler

Received share status: Active (indicated by a green checkmark).

Click the **Datasets** link.

On the resulting “Received Shares...” page, **Datasets** tab, check the box next to your Data Explorer cluster.

Click the “+ Map to target” button.

The screenshot shows the Microsoft Azure portal interface. In the top left, there's a navigation bar with 'practicums - Microsoft Azure' and a search bar. The main area shows 'practicums | Received Shares' under 'Data Share Invitations'. On the left, a sidebar has sections like 'Overview', 'Activity log', 'Access control (IAM)', 'Tags', 'Diagnose and solve problems', 'Properties', 'Locks', 'Sent Shares', 'Received Shares' (which is selected), 'Feedback', 'Monitoring', 'Alerts', 'Metrics', and 'Diagnostic settings'. The central part of the screen displays a 'Map datasets to target' dialog. This dialog has tabs for 'Details' and 'Datasets', with 'Datasets' being active. It shows a list of datasets: 'practicumdec' and 'practicumdec2'. Below this is a table with columns 'Datasets', 'Source Type', and 'Source Path'. The first row shows 'practicumdec' as an 'Azure Data Explorer Cluster' with the path 'practicumdec'. At the bottom of the dialog are 'Cancel' and 'Map to target' buttons.

On the resulting “Map datasets to target” pop-out, enter values for the following items:

Subscriptions	Select your subscription
Resource Groups	Select your Resource Group
Select Kusto clusters	Select your Data Explorer cluster

Click the “Map to target” button.

The screenshot shows the Microsoft Azure (Preview) interface. On the left, there's a navigation sidebar with options like Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Properties, Locks, Data Share (selected), and Sent Shares. The main area is titled 'practicumds | Received Shares' and shows a list of datasets. One dataset, 'practicumdec', is listed with its details: Source Type is 'Azure Data Explorer Cluster', Source Path is 'practicumdec', and Status is 'Mapped' (indicated by a green checkmark). There are buttons for Refresh, Unmap, Map to target, and Delete.

Data Studio

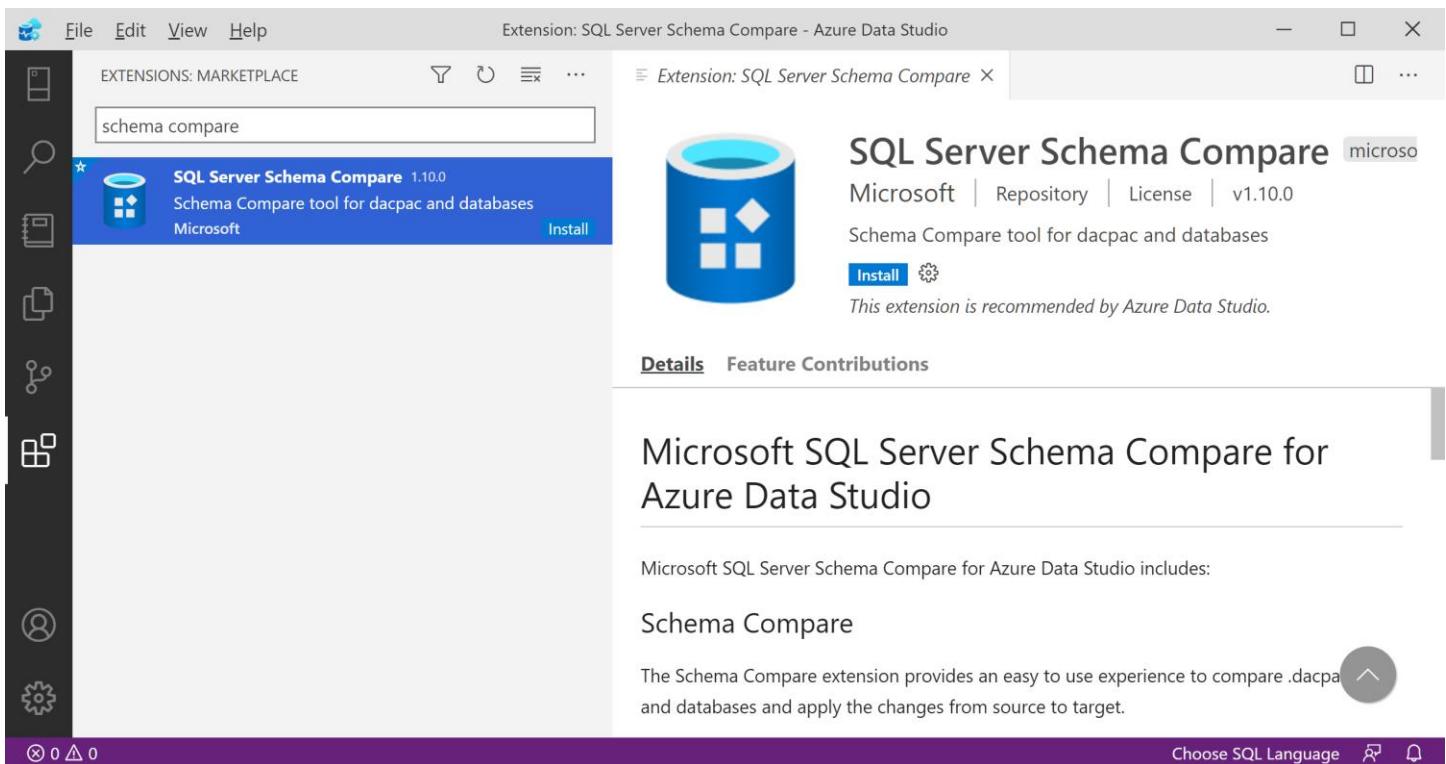
Browse to [Download and install Azure Data Studio - Azure Data Studio | Microsoft Docs](https://docs.microsoft.com/en-us/sql/azure-data-studio/download-azure-data-studio?view=sql-server-ver15)

The screenshot shows the Microsoft Docs page for 'Download and install Azure Data Studio'. The page has a sidebar with sections for Version (SQL Server 2019), Filter by title, and a list of documentation items. The main content area is titled 'Download and install Azure Data Studio' and includes a summary of the tool's features and a link to 'What is Azure Data Studio?'. Below this, there's a section titled 'Download Azure Data Studio' with a table showing download links for Windows, User Installer (recommended), System Installer, and zip file. The right side of the page has a sidebar with links for 'Is this page helpful?' (Yes/No), 'In this article', and various documentation links for Azure Data Studio.

Download, open, and install the “User Installer” for your platform. Once installed, launch Data Studio.

Add Extensions

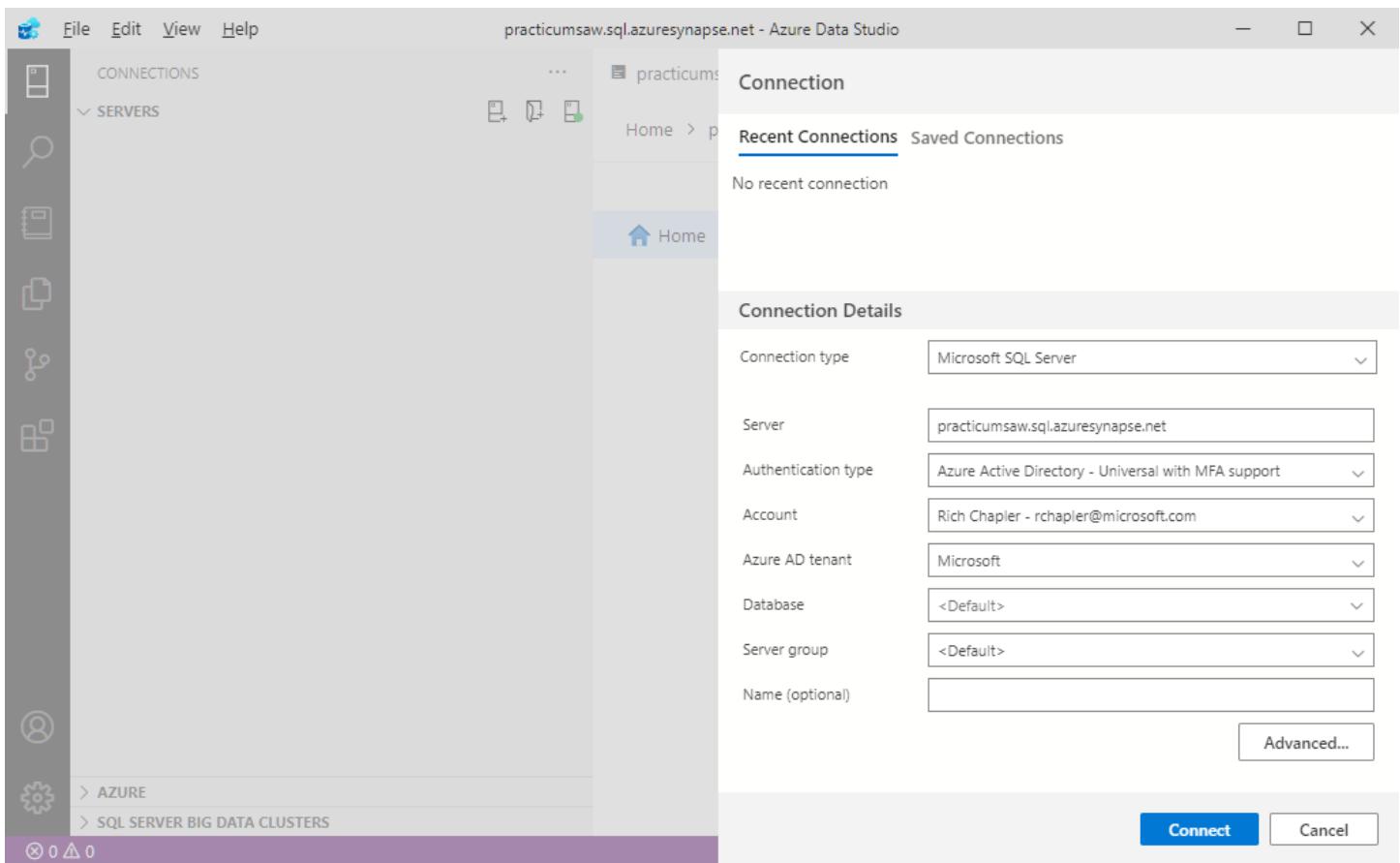
Navigate to **Extensions** in the left-hand navigation.



Search for “schema compare” and install the “**SQL Server Schema Compare**” extension.

Create Connection

Navigate to **Connections** in the left-hand navigation.



Click the “Create Connection” icon (first of three) to the right of the SERVERS section header.

On the “Connection Details” pop-out, enter values for the following items:

Connection Type	Confirm default value, “Microsoft SQL Server”
Server	Paste the “Dedicated SQL endpoint” value copied in Instantiate Resources Synapse
Authentication Type	Select “Azure Active Directory – Universal with MFA support”
Account	Confirm credentials
Azure AD Tenant	Confirm tenant

Click the **Connect** button.

Databricks

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Resource Group

The screenshot shows the 'Create an Azure Databricks workspace' page in the Microsoft Azure portal. The 'Basics' tab is active. In the 'Project Details' section, the subscription is set to 'rchapler' and the resource group is 'practicumrg'. In the 'Instance Details' section, the workspace name is 'practicumdb', the region is 'West US 2', and the pricing tier is 'Premium (+ Role-based access controls)'. At the bottom, there are navigation buttons for 'Review + create', '< Previous', and 'Next : Networking >'.

On the “Create an Azure Databricks workspace” page, enter values for the following items:

Subscription	Select your subscription
Resource Group	Select your Resource Group
Workspace Name	Enter a value that is meaningful and aligned with standards
Region	Select the value used during Resource Group creation
Pricing Tier	Select “ Premium... ” to enable features that we will need for this exercise

Review settings on remaining tabs {e.g., Networking, Tags}.

No additional settings are required for this exercise.

Click the “**Review + create**” button, review configuration, and then click the **Create** button.

New Cluster

Navigate to the **Overview** page in Databricks, click the “**Launch Workspace**” button, and login if required.

The screenshot shows the Azure Databricks landing page. On the left, there is a vertical sidebar with icons for Home, Workspace, Recents, Data, Clusters, Jobs, Models, and Search. The main content area features the Azure Databricks logo at the top. Below it are three main sections:

- Explore the Quickstart Tutorial**: An icon of a document with curly braces and a lightbulb. Description: Spin up a cluster, run queries on preloaded data, and display results in 5 minutes.
- Import & Explore Data**: An icon of a dashed box with a cloud and a plus sign. Description: Quickly import data, preview its schema, create a table, and query it in a notebook.
- Create a Blank Notebook**: An icon of a document with curly braces and a plus sign. Description: Create a notebook to start querying, visualizing, and modeling your data.

Below these sections are three tabs: **Common Tasks**, **Recents**, and **Documentation**. The **Common Tasks** tab is active, showing the following options:

- New Notebook
- Create Table
- New Cluster
- New Job
- New MLflow Experiment
- Import Library
- Read Documentation

The **Recents** tab shows the message: "Recent files appear here as you work." The **Documentation** tab lists three items:

- Documentation
- Release Notes
- Getting Started

In the “Common Tasks” grouping, click “New Cluster”.

Create Cluster

New Cluster

Cancel Create Cluster

Cluster Name: practicumdbc

Cluster Mode: Standard

Pool: None

Databricks Runtime Version: Runtime: 7.4 (Scala 2.12, Spark 3.0.1)

Autopilot Options:

- Enable autoscaling
- Terminate after 120 minutes of inactivity

Worker Type: Standard_DS3_v2 (14.0 GB Memory, 4 Cores, 0.75 DBU)

Driver Type: Same as worker (14.0 GB Memory, 4 Cores, 0.75 DBU)

Min Workers: 2

Max Workers: 8

This Runtime version supports only Python 3.

UI | JSON

On the “Create Cluster” page, enter values for the following items:

Cluster Name Enter a value that is meaningful and aligned with standards

No additional settings are required for this exercise.

Click the “Create Cluster” button.

New Notebook

Return to the start page.

In the “Common Tasks” grouping, click “New Notebook”.

The screenshot shows the 'Create Notebook' dialog box in the Azure Databricks interface. The dialog has fields for 'Name' (set to 'practicumdbn'), 'Default Language' (set to 'Python'), and 'Cluster' (set to 'practicumdbc'). Below the dialog, there's a message about quickly importing data. The background shows the Databricks workspace with various navigation links like Home, Workspace, Recents, Data, Clusters, Jobs, Models, and Search.

Enter a meaningful name and click the **Create** button.

The screenshot shows the Azure Databricks notebook interface with a new notebook titled 'practicumdbn (Python)'. The notebook contains a single cell labeled 'Cmd 1' with the number '1' in it. The left sidebar shows the 'Workspace' tab is selected. The top navigation bar shows the URL as https://adb-2154823451042175.15.azuredatabricks.net/?o=2154823451042175#notebook/1887258985769069/com...

Make note of the URL {i.e., <http://adb-21458...>} for use in the next section.

Secret Scope

Navigate to your instance of Key Vault.

Click **Properties** in the **Settings** group of the left-hand navigation.

practicumkv | Properties

Key vault

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Events

Settings

Keys

Secrets

Certificates

Access policies

Networking

Security

Properties

Locks

Monitoring

Alerts

Metrics

Name: practicumkv

Sku (Pricing tier): Standard

Location: westus2

Vault URI: <https://practicumkv.vault.azure.net/>

Resource ID: /subscriptions/.../resourceGroups/practicumrg...

Subscription ID: [REDACTED]

Subscription Name: rchapler

Directory ID: [REDACTED]

Directory Name: Microsoft

Soft-delete: Soft delete has been enabled on this key vault

Days to retain deleted vaults: 90

Purge protection:

Disable purge protection (allow key vault and objects to be purged during retention period)

Enable purge protection (enforce a mandatory retention period for deleted vaults and vault objects)

Make note of the values in “Vault URI” and “Resource ID”.

On a new tab, navigate to [https://\[databricksInstance\]#secrets/createScope](https://[databricksInstance]#secrets/createScope).

You will replace {databricksInstance} with the start of the URL in your workspace, from my exercise, example:

<https://adb-2154823451042175.15.azure.databricks.net/#secrets/createScope>

The screenshot shows the Microsoft Azure Databricks portal with the URL <https://adb-2154823451042175.15.azuredatabricks.net/?o=2154823451042175#secrets/createScope>. The left sidebar includes icons for Home, Workspace, Recents, Data, Clusters, Jobs, and Models. The main content area is titled "Create Secret Scope" with a "Create" button. It contains fields for "Scope Name" (practicumdbss), "Manage Principal" (Creator), and "Azure Key Vault" (DNS Name: https://practicumkv.vault.azure.net/ and Resource ID: /46ed/resourceGroups/practicumrg/providers/Microsoft.KeyVault/vaults/practicumkv).

On the “Create Secret Scope” page, enter values for the following items:

Scope Name	Enter a value that is meaningful and aligned with standards
DNS Name	Paste the copied “Vault UI” value
Resource ID	Paste the copied “Resource ID” value

Click the **Create** button.

DevOps

Browse to <https://azure.microsoft.com/en-us/services/devops/>

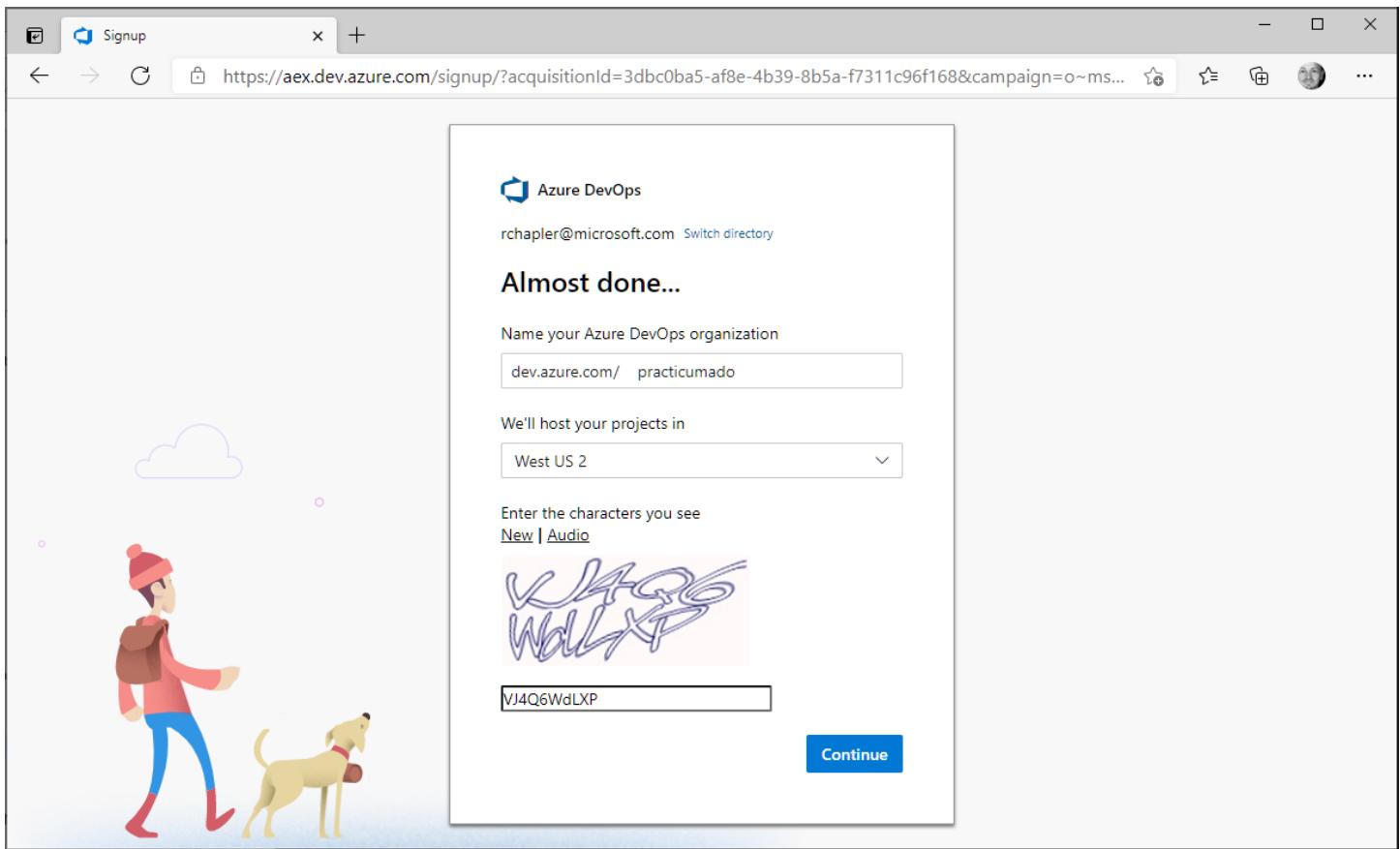
The screenshot shows the Microsoft Azure DevOps Services landing page. At the top, there's a navigation bar with links for Overview, Solutions, Products (which is currently selected), Documentation, Pricing, Training, Marketplace, Partners, Support, Blog, and More. Below the navigation is a breadcrumb trail: Home / Services / Azure DevOps. The main heading is "Azure DevOps" with the subtext "Plan smarter, collaborate better, and ship faster with a set of modern dev services." Two buttons are visible: "Start free" and "Start free with GitHub". To the right is a colorful illustration of people working on a rocket launching from a stack of servers. Below the illustration, there's a link for users who already have an account and a "Sign in to Azure DevOps" button.

Click the “Start free with GitHub” button.

On the resulting Azure DevOps page, click the “New organization” link on the left-hand navigation.

The screenshot shows the Azure DevOps Signup page. On the left, there's a cartoon illustration of a person walking a dog. The main content area has a "Get started with Azure DevOps" heading. It includes a note about agreeing to Terms of Service, Privacy Statement, and Code of Conduct. There's also a checkbox for opting in to information, tips, and offers from Azure DevOps and other Microsoft products. A "Continue" button is at the bottom of the form.

Click the **Continue** button.



On the “Almost done...” page, enter values for the following form items:

Name your Azure DevOps organization Enter a value that is meaningful and aligned with standards

We will host your projects in Select the value used during Resource Group creation

Click the **Continue** button.

Create Project

The screenshot shows the Azure DevOps 'Create a project to get started' page. On the left, there's a sidebar with organization names (practicumado, [REDACTED], [REDACTED], [REDACTED]), a '46 more organizations' link, and a 'New organization' button. A 'What's new' section mentions 'Sprint 184 release notes' about Azure Pipelines grants. Below that is an 'Organization settings' link. The main area has a title 'Create a project to get started'. It includes a 'Project name *' field with 'practicumadop' entered, a 'Description' field (empty), and a 'Visibility' section. The 'Private' option is selected, highlighted with a blue border. Other options shown are 'Public' and 'Enterprise'. At the bottom is a blue '+ Create project' button.

On the “Create a project to get started” page, enter values for the following form items:

Project Name Enter a value that is meaningful and aligned with standards

Visibility Select the value that best aligns with your requirements

Click the “+ Create project” button.

When processing is complete, you will be navigated to a screen like the one snipped below.

The screenshot shows the Azure DevOps interface for the 'practicumadop' project. The left sidebar contains links for Overview, Summary (which is selected), Dashboards, Wiki, Boards, Repos, Pipelines, Test Plans, and Artifacts. The main content area features a green header with the project name 'practicumadop'. Below the header is a cartoon illustration of a person sitting at a desk with a laptop, accompanied by a dog. The text 'Welcome to the project!' is displayed, followed by a question 'What service would you like to start with?'. Below this are buttons for Boards, Repos, Pipelines, Test Plans, and Artifacts, with 'Artifacts' being the active tab. At the bottom, there is a link to 'or manage your services'.

Key Vault

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Resource Group

The screenshot shows the 'Create key vault' wizard in the Microsoft Azure portal. The current step is 'Basics'. The page includes a brief description of Azure Key Vault, project details (subscription and resource group selection), instance details (key vault name, region, and pricing tier), and navigation buttons for 'Review + create' and 'Next : Access policy >'.

Azure Key Vault is a cloud service used to manage keys, secrets, and certificates. Key Vault eliminates the need for developers to store security information in their code. It allows you to centralize the storage of your application secrets which greatly reduces the chances that secrets may be leaked. Key Vault also allows you to securely store secrets and keys backed by Hardware Security Modules or HSMs. The HSMs used are Federal Information Processing Standards (FIPS) 140-2 Level 2 validated. In addition, key vault provides logs of all access and usage attempts of your secrets so you have a complete audit trail for compliance.

Project details

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription * rchabler

Resource group * practicumrg

Create new

Instance details

Key vault name * practicumkv

Region * West US 2

Pricing tier * Standard

Review + create < Previous Next : Access policy >

On the “Create Key Vault” page, enter values for the following items:

Subscription	Select your subscription
Resource Group	Select your Resource Group
Key Vault Name	Enter a value that is meaningful and aligned with standards
Region	Select the value used during Resource Group creation
Pricing Tier	Confirm default selection, “Standard”

Review settings on remaining tabs {e.g., Access Policy, Networking, Tags}. No additional settings are required for this exercise.

Click the “Review + create” button, review configuration, and then click the **Create** button.

Log Analytics

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Resource Group

A Log Analytics workspace is the basic management unit of Azure Monitor Logs. There are specific considerations you should take when creating a new Log Analytics workspace. [Learn more](#)

With Azure Monitor Logs you can easily store, retain, and query data collected from your monitored resources in Azure and other environments for valuable insights. A Log Analytics workspace is the logical storage unit where your log data is collected and stored.

Project details
Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription * rchapler

Resource group * practicumrg
[Create new](#)

Instance details
Name * praticumlaw

Region * West US 2

[Review + Create](#) [« Previous](#) [Next : Pricing tier >](#)

On the “Create Log Analytics workspace” page, enter values for the following items:

Subscription	Select your subscription
Resource Group	Select your Resource Group
Name	Enter a value that is meaningful and aligned with standards
Region	Select the value used during Resource Group creation

Review settings on remaining tabs {e.g., Pricing Tier, Tags}; no additional settings are required for this exercise.

Click the “Review + create” button, review configuration, and then click the **Create** button.

Metrics Advisor

The screenshot shows the Microsoft Azure portal in dark mode, specifically the 'Create Metrics Advisor' wizard. The top navigation bar includes the Microsoft Azure logo, a search bar, and a user profile for rchapler@microsoft.com.

The breadcrumb navigation shows: Dashboard > Resource groups > practicumrg > Create a resource > Marketplace > Metrics Advisor > Create Metrics Advisor.

The main content area has tabs at the top: Basics (selected), Virtual network, Tags, and Review + create.

Project details: A descriptive text explains that Metrics Advisor embeds AI-powered monitoring features to stay one step ahead of incidents, requiring no machine-learning expertise. It monitors performance across various sectors like manufacturing and retail.

Subscription: Set to rchapler.

Resource group: Set to practicumrg.

Instance details:

- Region:** Set to West US 2.
- Name:** Set to practicumma.
- Pricing tier:** Set to S0.

[View full pricing details](#)

Storage:

Bring your own storage: Yes No

The deployment could take up to 60 minutes to complete, although it normally finishes in less than 10 minutes.

I confirm I have read and understood the notice below.

[Service Agreement & Terms](#)

At the bottom are buttons for [Review + create](#), [< Previous](#), and [Next : Virtual network >](#).

On the “Create Metrics Advisor” page, enter values for the following items:

Subscription	Select your subscription
Resource Group	Select your Resource Group
Region	Select the value used during Resource Group creation
Name	Enter a value that is meaningful and aligned with standards
Pricing Tier	Select S0
Bring Your Own Storage	Confirm default selection, “No”
I confirm...	Check the box

Review settings on remaining tabs {e.g., Virtual Network, Tags}.

No additional settings are required for this exercise.

Click the “**Review + create**” button, review configuration, and then click the **Create** button.

Add Permissions to Data Explorer

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Data Explorer

Navigate to your instance of Data Explorer.

The screenshot shows the Azure portal interface for managing permissions. On the left, the navigation menu includes 'Overview', 'Activity log', 'Access control (IAM)', 'Tags', 'Diagnose and solve problems', 'Permissions' (which is selected), 'Query', 'Settings', 'Scale up', 'Scale out', and 'Configurations'. The main content area is titled 'practicumdec | Permissions' and shows a table with one item: 'practicumdf' (Type: App). To the right, a modal window titled 'New Principals' is open, showing a search bar with 'practicumma' and a list of results with one item selected: 'practicumma' (GUID: 1b12f297-9025-4ac9-bfc3-98832a38dc38). A 'Selected' button is shown below the list.

Click **Permissions** in the left-hand navigation.

Click “**+ Add**” on the resulting page.

Select **AllDatabasesViewer** from the resulting drop-down.

Search for and then click to **Select** your instance of Metrics Advisor.

Postman

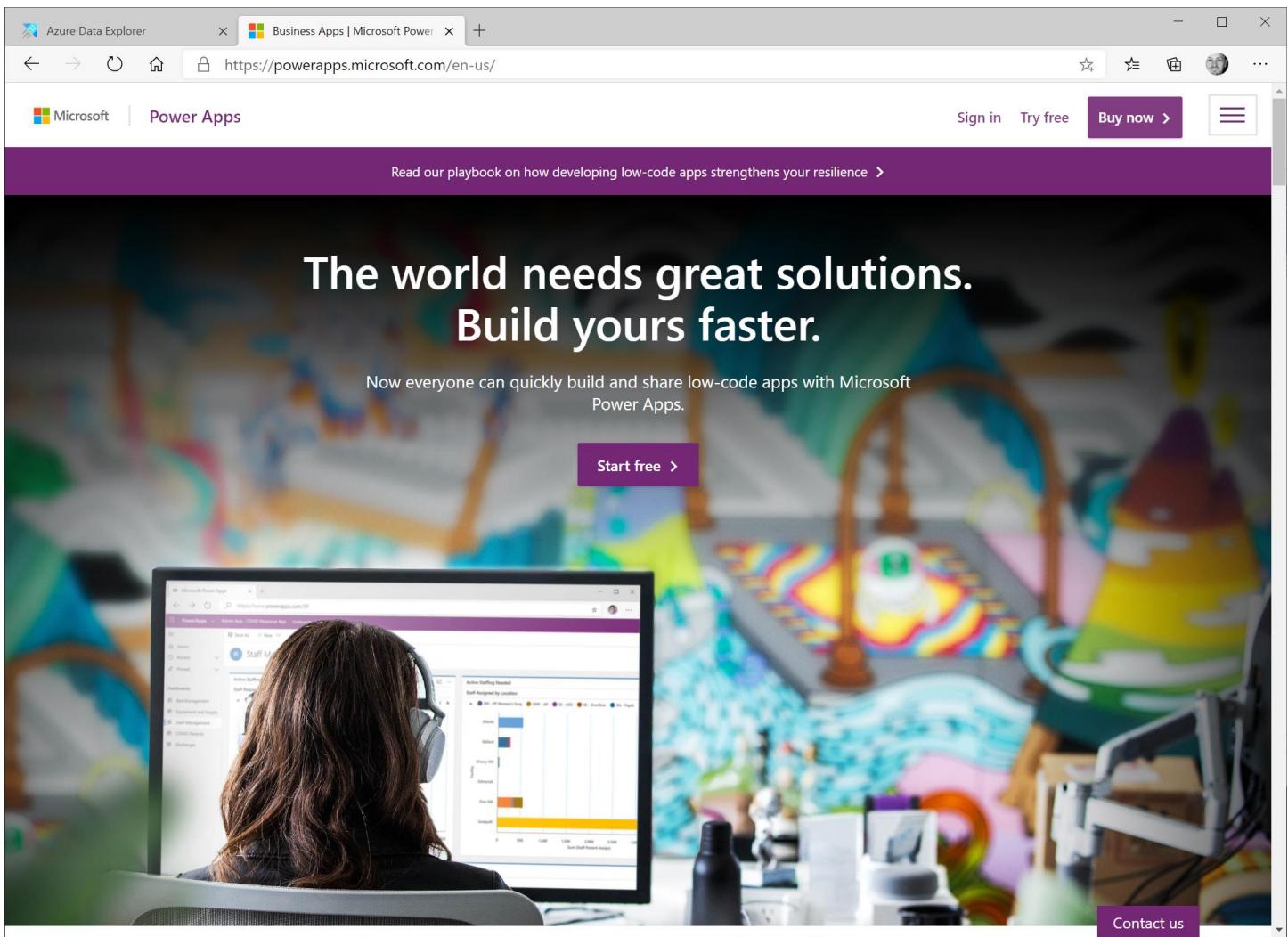
Browse to [Download Postman | Try Postman for Free](#)

The screenshot shows the official Postman download page at <https://www.postman.com/downloads/>. The main heading is "Download Postman". Below it, a sub-section titled "The Postman app" describes the app as ever-improving and full-featured, with a "Download the App" button. To the right, a large screenshot of the Postman application interface is displayed, featuring a toolbar with "New", "Import", "Runner", and other options. The main area shows a "History" tab with a recent entry for "postman-echo.com/get". Arrows point from the "History" tab to the "Collections" and "APIs" tabs. A "Params" section is visible on the right.

Select, download, and install the version of the Postman app appropriate for your system.

Power Apps

If you don't have a working instance of Power Apps, you can get started at <https://Power Apps.microsoft.com/en-us/>



If you are already set up, click “[Sign In](#)”.

Purview

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Key Vault
- Resource Group

Create a Purview account to develop a data governance solution in just a few clicks. A storage account and eventhub will be created in a managed resource group in your subscription for catalog ingestion scenarios. [Learn more](#)

Project details

Subscription * rchapler

Resource group * practicumrg [Create new](#)

Instance details

Purview account name * practicump

Location * West US 2

Managed resources

A resource group, a storage account, and an Eventhub will be created in the selected subscription for catalog ingestion scenarios. The Microsoft.Storage and Microsoft.EventHub resource providers will get registered. [Learn more](#)

Managed resource group name * practicumpmrg

Storage account name *Name will be auto-generated during account creation.*

Event Hubs namespace name *Name will be auto-generated during account creation.*

Review + Create **Previous** **Next: Networking >**

On the “Create Purview account” page, enter values for the following items:

Subscription	Select your subscription
Resource Group	Select your Resource Group
Purview Account Name	Enter a value that is meaningful and aligned with standards
Location	Select the value used during Resource Group creation
Managed Resource Group...	Enter a value that is meaningful and aligned with standards

Review settings on remaining tabs {e.g., Networking, Tags}. No additional settings are required for this exercise.

Click the “Review + Create” button, review configuration, and then click the **Create** button.

Add Access Policy to Key Vault

Navigate to the Key Vault.

The screenshot shows the Microsoft Azure Key Vault settings page for the 'practicumkv' vault. The left sidebar shows navigation options like Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Events, and Settings (Keys, Secrets, Certificates, Access policies). The main content area is titled 'Essentials' and displays resource group ('practicumrg'), location ('West US 2'), subscription ('rchapler'), and directory information ('Vault URI: https://practicumkv.vault.azure.net/'). The 'Access control (IAM)' section is currently selected.

Click “Access Policies” in the **Settings** group of the left-hand navigation.

The screenshot shows the Microsoft Azure Key Vault Access policies page for the 'practicumkv' vault. The left sidebar shows the same navigation options as the previous screen. The main content area is titled 'Access policies' and includes sections for enabling access to Azure VM deployment, Resource Manager deployment, and Disk Encryption, as well as selecting a permission model (Vault access policy or Azure role-based access control). A prominent blue link '+ Add Access Policy' is visible. The 'Access policies' option in the sidebar is also highlighted.

Click the “+ Add Access Policy” link.

The screenshot shows the 'Add access policy' dialog in the Microsoft Azure portal. On the left, there are sections for 'Configure from template (optional)', 'Key permissions' (0 selected), 'Secret permissions' (2 selected), and 'Certificate permissions'. Below these are fields for 'Select principal *' and 'Authorized application'. A blue 'Add' button is visible. On the right, a modal window titled 'Principal' shows a search bar with 'practicump' and a list of results. One result, 'practicump d0206ad4-fa5d-4923-982a-4cf48bcf3559', is selected. This selection is also listed under 'Selected items' with a 'Remove' button. At the bottom of the modal is a 'Select' button.

On the “Add access policy” page, enter values for the following items:

Secret Permissions Select **Get** and **List** in the drop-down

Select Principal Search for, and then **Select** the managed identity for your instance of Purview

Click the **Add** button.

The screenshot shows the 'Access policies' page for the 'practicumkv' key vault. The left sidebar includes 'Overview', 'Activity log', 'Access control (IAM)', 'Tags', 'Diagnose and solve problems', 'Events', 'Settings' (with 'Keys', 'Secrets', 'Certificates', 'Access policies' selected), 'Networking', 'Security', and 'Properties'. The main area displays a message 'Please click the 'Save' button to commit your changes.' Below it, 'Enable Access to:' checkboxes are available for 'Azure Virtual Machines for deployment', 'Azure Resource Manager for template deployment', and 'Azure Disk Encryption for volume encryption'. A 'Permission model' section shows 'Vault access policy' (selected) and 'Azure role-based access control' as options. A '+ Add Access Policy' button is present. A table titled 'Current Access Policies' lists one entry: 'APPLICATION' with principal 'practicump', 'Key Permissions' (0 selected), 'Secret Permissions' (2 selected), and 'Certificate Permissions' (0 selected). A 'Delete' button is located at the bottom right of the table row.

Click the **Save** button.

Connect Key Vault

Navigate to Purview and click on the “**Open Purview Studio**” button.

Select the **Management** icon on the left-hand navigation.

Click **Credentials** in the “**Security and access**” group of the resulting menu.

Click the “**Manage Key Vault connections**” button on the resulting page.

Click the “**+ New**” button the resulting pop-out.

The screenshot shows the 'New Key Vault' creation dialog in the Azure Purview Studio. The dialog has the following fields:

- Name ***: practicumkv
- Description**: Enter description
- Azure subscription**: rchapler [dropdown]
- Key Vault name ***: practicumkv
- Note**: You must grant the Purview managed identity access to your Azure Key Vault. [See more](#)

At the bottom are 'Create', 'Back', and 'Cancel' buttons.

On the “**New Key Vault**” pop-out, enter values for the following items:

Name	Enter a value that is meaningful and aligned with standards
Subscription	Select your subscription
Key Vault Name	Select your instance of Key Vault

Click the **Create** button.

On the resulting “**Confirm granting access**” pop-up, review and then click the **Confirm** button.

Confirm the addition of your instance of Key Vault on the “**Manage Key Vault connections**” pop-out.

Click the **Close** button.

Add New Credential

On the **Credentials** page, click the “**+ New**” button.

The screenshot shows the Azure Purview Studio interface. In the top navigation bar, there are tabs for 'practicump - Microsoft Azure' and 'Azure Purview Studio'. The URL in the address bar is [https://ms.web.purview.azure.com/resource/practicump/main/datasource/management/credentials?feature.tenant=\[REDACTED\]](https://ms.web.purview.azure.com/resource/practicump/main/datasource/management/credentials?feature.tenant=[REDACTED]). The main content area has a blue header bar with icons for notifications, search, and user profile. On the left, a sidebar menu includes 'General', 'Overview', 'Metrics', 'Lineage connections', 'Data Factory', 'Data Share', 'Security and access', and 'Credentials', with 'Credentials' currently selected. The central panel is titled 'Credentials' and shows a sub-section for 'New credential'. The 'Name' field is populated with 'practicumsds'. The 'Description' field is empty. Under 'Authentication method', 'SQL authentication' is selected. The 'User name' field contains 'rchapler'. The 'Password' field is empty. Under 'Key Vault connection', 'practicumkv' is selected. The 'Secret name' field contains 'practicumsds-adminpassword'. The 'Secret version' field contains the placeholder 'Use the latest version if left blank'. At the bottom right of the dialog are 'Create' and 'Cancel' buttons.

On the resulting “**New Credential**” pop-out, enter values for the following items:

Name	Enter a value that is meaningful and aligned with standards
Authentication	Select “SQL authentication”
User Name	Enter the “server admin login” value used during creation of the SQL Database Server
Key Vault Connection	Select the Key Vault created in Appendix Key Vault
Secret Name	Enter the value used in Appendix SQL Add Key Vault Secret

No additional changes are required. Click the **Create** button.

Resource Group

The screenshot shows the Microsoft Azure (Preview) portal with the URL <https://ms.portal.azure.com/#create/Microsoft.ResourceGroup>. The page is titled "Create a resource group". The "Basics" tab is selected. The "Project details" section contains a "Subscription" dropdown set to "rchapler" and a "Resource group" dropdown set to "practicumrg". The "Resource details" section contains a "Region" dropdown set to "(US) West US 2". At the bottom, there are navigation buttons: "Review + create" (highlighted in blue), "< Previous", and "Next : Tags >".

On the “Create a Resource Group” page, enter values for the following items:

Subscription	Select your subscription
Resource Group	Enter a value that is meaningful and aligned with standards
Region	Select a region appropriate for your situation; take into consideration that: <ul style="list-style-type: none">• Some regions {e.g., West US and East US} see higher demand than others• Creation of resources in the same region offers best performance and lowest cost

Review settings on remaining tabs {e.g., Tags}.

No additional settings are required for this exercise.

Click the “Review + create” button, review configuration, and then click the **Create** button.

SQL

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Resource Group

Select SQL deployment option

How do you plan to use the service?

SQL databases
Best for modern cloud applications. Hyperscale and serverless options are available.
Resource type: Single database
Create **Show details**

SQL managed instances
Best for most migrations to the cloud. Lift-and-shift ready.
Resource type: Single instance
Create **Show details**

SQL virtual machines
Best for migrations and applications requiring OS-level access. Lift-and-shift ready.
Image
Create **Show details**

On the “Select SQL deployment option” page, confirm default selection “Single database” in the “SQL databases” > “Resource type” drop-down.

Click the **Create** button.

New server - Microsoft Azure

Create SQL Database

Basics • Networking Security Additional settings Tags Review + create

Create a SQL database with your preferred configurations. Complete the Basics tab then go to Review + Create to provision with smart defaults, or visit each tab to customize. [Learn more](#)

Project details

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription * rchapler

Resource group * praticumrg [Create new](#)

Database details

Enter required settings for this database, including picking a logical server and configuring the compute and storage resources

Database name * praticumsd

Server * Select a server [Create new](#)

New server

Microsoft

Server name * praticumsds .database.windows.net

Server admin login * rchapler

Password *

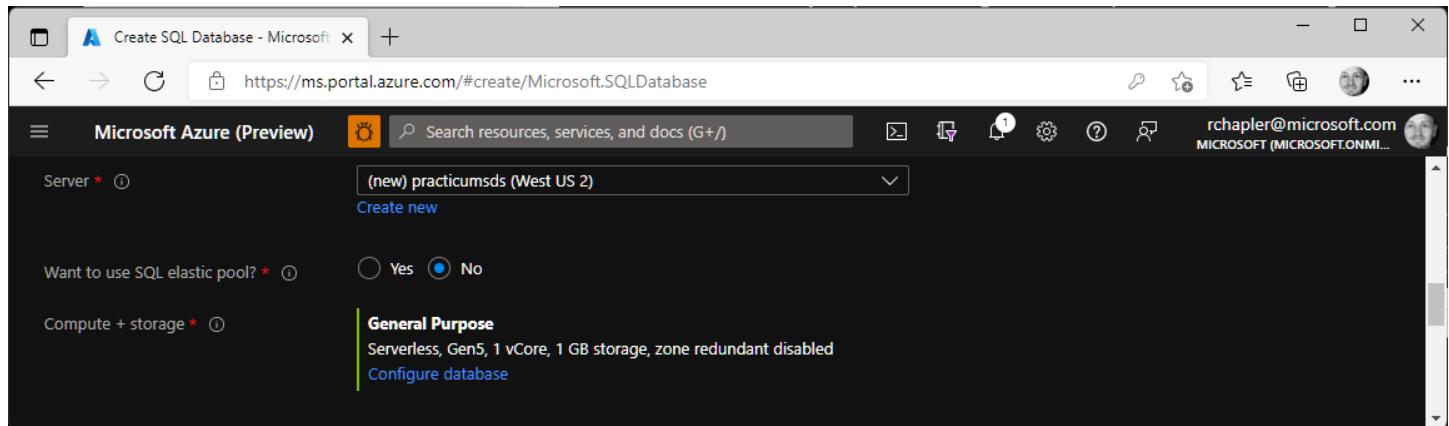
Confirm password *

Location * (US) West US 2

Review + create **Next : Networking >** **OK**

On the “Create SQL Database” page, **Basics** tab, enter values for the following items:

Subscription	Select your subscription
Resource Group	Select your Resource Group
Database Name	Enter a value that is meaningful and aligned with standards
Server	Click the “Create new” link, enter values in the resulting “New server” pop-out and then click the OK button



Further down on the “Create SQL Database” page, **Basics** tab, enter values for the following items:

Want to use SQL elastic pool?	Confirm default value, No
Compute + Storage	Click the “Configure database” link

Service and compute tier

Select from the available tiers based on the needs of your workload. The vCore model provides a wide range of configuration controls and offers Hyperscale and Serverless to automatically scale your database based on your workload needs. Alternately, the DTU model provides set price/performance packages to choose from for easy configuration. [Learn more](#)

Service tier: General Purpose (Scalable compute and storage options)

Compute tier:

- Provisioned** - Compute resources are pre-allocated. Billed per hour based on vCores configured.
- Serverless** - Compute resources are auto-scaled. Billed per second based on vCores used.

Compute Hardware

Select the hardware configuration based on your workload requirements. Availability of compute optimized, memory optimized, and confidential computing hardware depends on the region, service tier, and compute tier.

Hardware Configuration: Gen5
up to 40 vCores, up to 120 GB memory
[Change configuration](#)

Max vCores: 1 vCore

Min vCores: 0.5 vCore

Apply

On the resulting **Configure** page, enter values for the following items:

Service Tier	Confirm default value, “General Purpose...”
Compute Tier	Click the Serverless radio button
Max vCores	Confirm minimum value
Min vCores	Confirm minimum value

The screenshot shows the Microsoft Azure (Preview) portal with the URL <https://ms.portal.azure.com/#create/Microsoft.SQLDatabase>. The user is on the 'Create SQL Database' configuration page. Key settings shown include:

- Auto-pause delay:** Enabled, set to 0 Days, 1 Hour, 0 Minutes.
- Data max size (GB):** Set to 1 GB, resulting in 307.2 MB LOG SPACE ALLOCATED.
- Cost summary:** Details the cost for a Gen5 - General Purpose (GP_S_Gen5_1) database:
 - Cost per GB (in USD): 0.12
 - Max storage selected (in GB): x 1.3
 - ESTIMATED STORAGE COST / MONTH: 0.15 USD
 - COMPUTE COST / VCORE / SECOND¹: 0.000145 USD

At the bottom, there is an 'Apply' button.

Further down on the resulting **Configure** page, enter values for the following items:

Enable Auto-Pause	Confirm default value, checked
Days Hours Minutes	Confirm default values, 0 1 0
Data Max Size (GB)	Set minimum appropriate value
“...database zone redundant”	Confirm default value, No

Review the “Cost Summary” and then click the **Apply** button.

Backup storage redundancy

Choose how your PITR and LTR backups are replicated. Geo restore or ability to recover from regional outage is only available when geo-redundant storage is selected.

Backup storage redundancy (i)

- Locally-redundant backup storage - Preview
- Zone-redundant backup storage - Preview
- Geo-redundant backup storage

i Your use of either of the Preview backup storage redundancy options (ZRS and LRS) is governed by the agreement under which you obtained Microsoft Azure Services. By selecting a Preview redundancy option, you confirm that you agree to the preview terms in such agreement.
Microsoft Azure Legal Information: [Learn more](#)

Review + create Next : Networking >

Back on and further down on the “Create SQL Database” page, **Basics** tab, enter values for the following items:

Backup Storage Redundancy Confirm default value, “Locally-redundant backup storage”

Key Vault Secret, Admin Password

Navigate to your instance of Key Vault.

Click **Secrets** in the **Settings** group of the left-hand navigation.

Click the “+ Generate/Import” button.

Create a secret

Upload options Manual

Name (i) practicumsds-adminpassword

Value (i)

Content type (optional)

Set activation date

Set expiration date

Enabled Yes No

Tags 0 tags

Create

On the “Create a secret” page, enter values for the following items:

Name	Enter a value that is meaningful and aligned with standards
Value	Enter the value used when creating the SQL Database Server

No additional changes are required. Click the **Create** button.

Sample Database

Navigate to the “Additional settings” tab.

The screenshot shows the Microsoft Azure portal's "Create SQL Database" wizard. The URL is https://ms.portal.azure.com/#create/Microsoft.SQLDatabase. The user is signed in as rchaper@microsoft.com. The breadcrumb navigation shows: Dashboard > practicumrg > Create a resource > Marketplace > Azure SQL > Select SQL deployment option > Create SQL Database.

The "Additional settings" tab is selected. The sub-section "Data source" is visible, with the "Sample" option selected under "Use existing data". The note "AdventureWorksLT will be created as the sample database." is displayed. The "Database collation" section shows "SQL_Latin1_General_CI_AS" selected. At the bottom, there are buttons for "Review + create", "< Previous", and "Next : Tags >".

On the “Create SQL Database” page, “Additional Settings” tab, enter values for the following items:

Use Existing Data	Select the “Sample” option
-------------------	----------------------------

No additional settings are required for this exercise.

Click the **Review + create** button, review configuration, and then click the **Create** button.

Configure Firewall

Navigate to the new SQL Server and click the **Show firewall settings** link on the **Overview** page.

The screenshot shows the Azure portal interface for managing a SQL server. On the left, there's a sidebar with various navigation options like Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Quick start, and Settings. Under Settings, there are links for Azure Active Directory, SQL databases, and SQL elastic pools. The main content area is titled 'practicumsds | Firewalls and virtual networks'. It contains several configuration settings: 'Deny public network access' is set to 'Yes'; 'Minimum TLS Version' is set to '1.2'; 'Connection Policy' is set to 'Default'; and 'Allow Azure services and resources to access this server' is set to 'Yes'. There are also buttons for 'Save', 'Discard', and 'Add client IP'. A search bar at the top is empty.

Set “Allow Azure services and resources to access this server” to “Yes”.

Click the **Save** button.

Storage Account

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Resource Group

Azure Storage is a Microsoft-managed service providing cloud storage that is highly available, secure, durable, scalable, and redundant. Azure Storage includes Azure Blobs (objects), Azure Data Lake Storage Gen2, Azure Files, Azure Queues, and Azure Tables. The cost of your storage account depends on the usage and the options you choose below.

[Learn more about Azure storage accounts](#)

Project details

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription * rchapler

Resource group * praticumrg

Create new

Instance details

The default deployment model is Resource Manager, which supports the latest Azure features. You may choose to deploy using the classic deployment model instead. [Choose classic deployment model](#)

Storage account name * praticumsa

Location * (US) West US 2

Performance Standard

Account kind StorageV2 (general purpose v2)

Replication Read-access geo-redundant storage (RA-GRS)

Review + create < Previous Next : Networking >

On the “Create Storage Account” page, enter values for the following items:

Subscription	Select your subscription
Resource Group	Select your Resource Group
Storage Account Name	Enter a value that is meaningful and aligned with standards
Location	Select the value used during Resource Group creation
Performance	Confirm default radio button selection, “Standard”
Account Kind	Confirm default selection, “StorageV2 (general purpose v2)”
Replication	Confirm default selection, “Read-access geo-redundant storage (RA-GRS)”

Review settings on remaining tabs {e.g., Networking, Data Protection, Tags}; no additional settings are required for this exercise.

Click the “Review + create” button, review configuration, and then click the **Create** button.

Synapse

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Data Lake

- Key Vault
- Resource Group
- SQL

Create Synapse workspace

*** Basics** **Security** **Networking** **Tags** **Review + create**

Create a Synapse workspace to develop an enterprise analytics solution in just a few clicks.

Project details

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all of your resources.

Subscription * Resource group * Managed resource group

Workspace details

Name your workspace, select a location, and choose a primary Data Lake Storage Gen2 file system to serve as the default location for logs and job output.

Workspace name * Region * Select Data Lake Storage Gen2 * From subscription Manually via URL

Account name * File system name *

Info We will automatically grant the workspace identity data access to the specified Data Lake Storage Gen2 account, using the [Storage Blob Data Contributor role](#). To enable other users to use this storage account after you create your workspace, perform these tasks:

- Assign other users to the **Contributor** role on workspace
- Assign other users the appropriate [Synapse RBAC roles](#) using Synapse Studio
- Assign yourself and other users to the **Storage Blob Data Contributor** role on the storage account

[Learn more](#)

Review + create **< Previous** **Next: Security >**

On the “Create Synapse workspace” page, enter values for the following items:

Subscription	Select your subscription
Resource Group	Select your Resource Group
Managed Resource Group	Enter a value that is meaningful and aligned with standards This additional Resource Group will hold ancillary resources created specifically for Synapse
Workspace Name	Enter a value that is meaningful and aligned with standards
Region	Select the value used during Resource Group creation

Select Data Lake Storage Gen2	Confirm default, “From subscription”
Account Name	Select the Storage Account created previously
File System Name	Select the Storage Account container created previously
Assign myself...	Check to assign necessary permissions

Click the “Next: Security >” button.

Click the “Next: Security >” button.

Enter values for “SQL Server Admin Login” and “SQL Password”.

No other setting modifications are required.

Click the “Review + create” button, review configuration, and then click the **Create** button.

Key Vault Secret, Admin Password

Navigate to your instance of Key Vault.

Click **Secrets** in the **Settings** group of the left-hand navigation.

Click the “+ Generate/Import” button.

The screenshot shows the 'Create a secret' page in the Microsoft Azure portal. The 'Name' field is set to 'practicumsw-adminpassword' and the 'Value' field contains a redacted password. The 'Enabled' switch is set to 'Yes'. The 'Create' button is visible at the bottom left.

On the “Create a secret” page, enter values for the following items:

Name	Enter a value that is meaningful and aligned with standards
Value	Enter the value used when creating the Synapse Workspace

No additional changes are required. Click the **Create** button.

Add Access Policy to Key Vault

Navigate to your instance of Key Vault.

Click “Access Policies” in the **Settings** group of the left-hand navigation.

practicumkv | Access policies

Key vault

Search (Ctrl+ /)

Save Discard Refresh

Enable Access to:

- Azure Virtual Machines for deployment
- Azure Resource Manager for template deployment
- Azure Disk Encryption for volume encryption

Permission model

- Vault access policy (selected)
- Azure role-based access control

+ Add Access Policy

Name	Email	Key Permissions	Secret Permissions	Certificate Permissions	Action
APPLICATION					

Click the “+ Add Access Policy” link.

Add access policy

Add access policy

Configure from template (optional)

Key permissions

0 selected

Secret permissions

2 selected

Certificate permissions

0 selected

Select principal *

None selected

Authorized application

None selected

Add

Select

practicumsw

e5ddc6b3-a858-4bcf-a9f0-26b5df8a27d7

Selected

Selected items

practicumsw

e5ddc6b3-a858-4bcf-a9f0-26b5df8a27d7

Remove

On the “Add access policy” page, enter values for the following items:

Secret Permissions Select **Get** and **List** in the drop-down

Select Principal Search for, and then **Select** the managed identity for your Synapse Workspace

Click the **Add** button.

The screenshot shows the Microsoft Azure Key Vault Access Policies page for a vault named 'practicumkv'. The left sidebar includes links for Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Events, Keys, Secrets, Certificates, and Access policies (which is selected). The main pane displays a message to save changes, enable access to various services, and choose a permission model (Vault access policy is selected). It also shows a table of current access policies for an application named 'practicumsw'.

Name	Email	Key Permissions	Secret Permissions	Certificate Permissions	Action
practicumsw		0 selected	2 selected	0 selected	<button>De</button>

Click the **Save** button.

Add Permissions to Data Explorer

To complete this objective, use the instructions in the Appendix to instantiate the following resources:

- Data Explorer

Navigate to your instance of Data Explorer.

New Principals
New AllDatabasesViewer Admin Principals

Search resources, services, and docs (G+/)

practicumsw

Role	Name	Type
2 selected	practicumdf	App
	practicumma	App

Selected items

practicumsw
7b669ff9-ec12-4a48-810a-257b7e2120b1

Select Remove

Click **Permissions** in the left-hand navigation.

Click “+ Add” on the resulting page.

Select **AllDatabasesViewer** for read permissions and **AllDatabasesAdmin** for write, etc. permissions from the resulting drop-down.

Search for and then click to **Select** your instance of Synapse.

Databases

Dedicated SQL Pool

Navigate to your instance of Synapse.

The screenshot shows the Microsoft Azure portal interface. The top navigation bar includes the URL <https://ms.portal.azure.com/#@microsoft.onmicrosoft.com/resource/subscriptions>. The main content area displays the 'practicumsw' Synapse workspace. On the left, a sidebar menu lists various options like Overview, Activity log, Access control (IAM), Tags, and Diagnose and solve problems. The central panel shows the workspace's configuration under the 'Essentials' tab, including details such as Resource group (practicumrg), Status (Succeeded), Location (West US 2), Subscription (rchapler), and various endpoint URLs. Buttons for 'New dedicated SQL pool' and 'New Apache Spark pool' are visible at the top of the central panel.

Click the “+ New dedicated SQL pool” button.

The screenshot shows the 'New dedicated SQL pool' creation wizard in the Microsoft Azure portal. The top navigation bar includes the URL <https://ms.portal.azure.com/#create/Microsoft.Synapse>. The main content area is titled 'New dedicated SQL pool' and has tabs for 'Basics', 'Additional settings', 'Tags', and 'Review + create'. The 'Basics' tab is selected. It contains fields for 'Dedicated SQL pool name' (set to 'practicumsw_dsp'), 'Performance level' (set to 'DW100c'), and 'Estimated price' (1.20 USD). At the bottom, there are buttons for 'Review + create' and 'Next: Additional settings >'

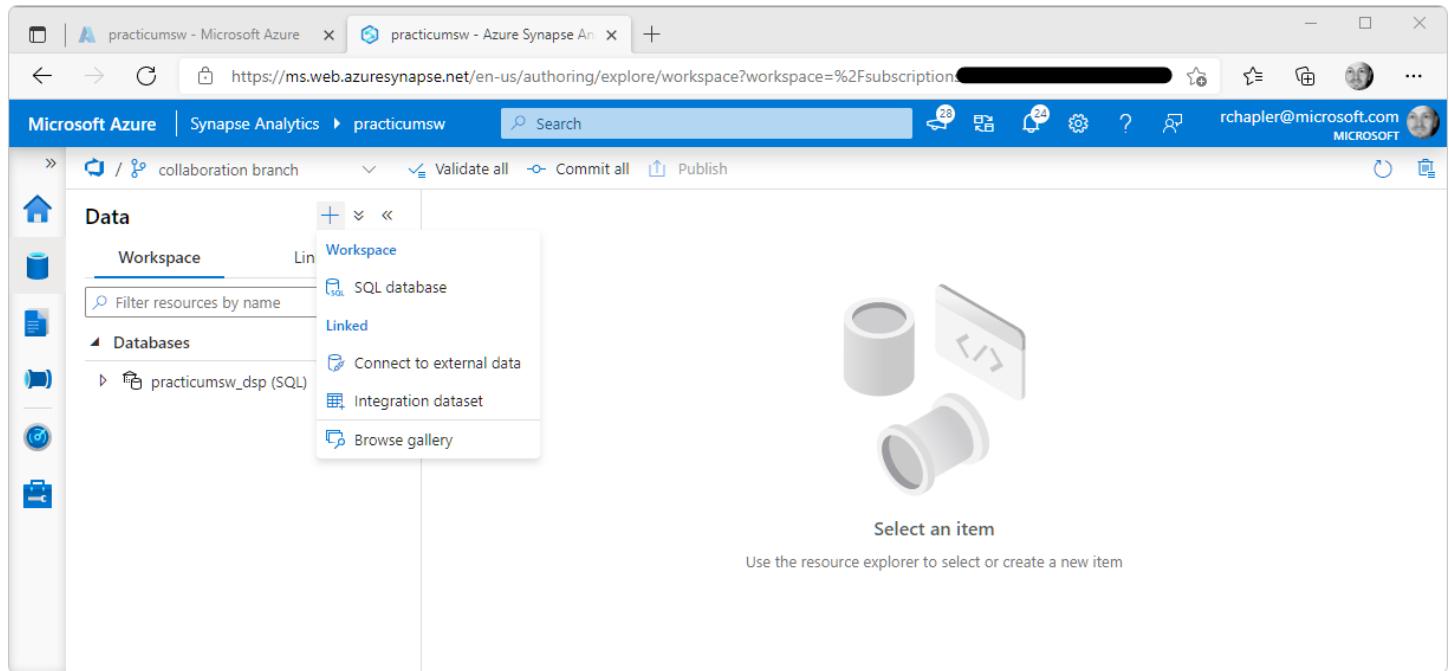
Enter a meaningful name in “**Dedicated SQL pool name**” and choose a pricing tier.

Click the “**Review + create**” button, review configuration, and then click the **Create** button.

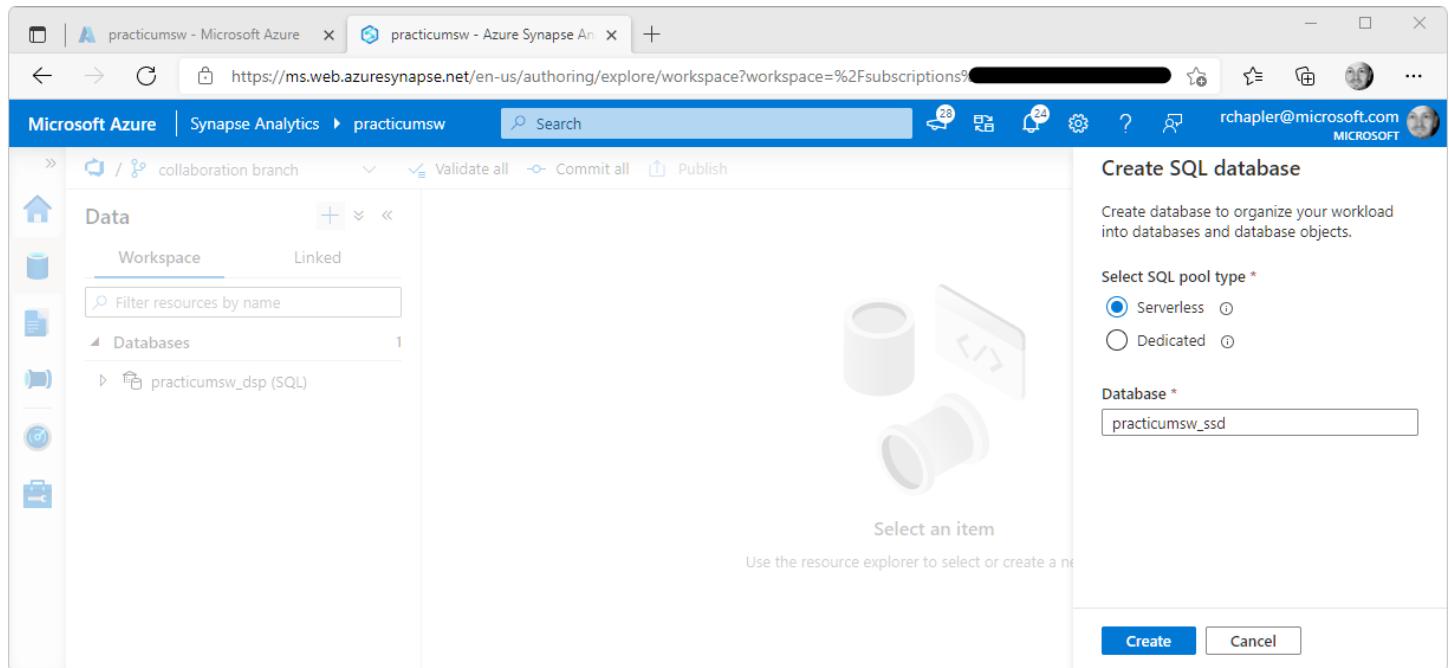
Serverless SQL Database

Navigate to your Synapse Analytics workspace and then click the **Open** link on the “Open Synapse Studio” rectangle.

Click the **Data** icon in the left-hand navigation.



Click the **Workspace** tab. Click the **+** button and “**SQL database**” in the resulting drop-down.



On the “**Create SQL database**” pop-out, enter values for the following items:

Select SQL Pool Type Confirm default, **Serverless**

Database Enter a value that is meaningful and aligned with standards

Click the **Create** button.

Linked Services

Key Vault

Navigate to your Synapse Analytics workspace and then click the **Open** link on the “Open Synapse Studio” rectangle.

Click the **Manage** icon and then “**Linked Services**” in the left-hand navigation.

The screenshot shows the Microsoft Azure portal interface for a Synapse Analytics workspace named "practicumsw". The left sidebar navigation menu is visible, showing options like "Analytics pools", "External connections", and "Linked services", which is currently selected. The main content area is titled "Linked services" and contains a message stating "Linked services are much like connection strings, which define the connection information needed for Azure Synapse Analytics to connect to external resources." Below this, there is a "New" button, a "Filter by name" input field, and an "Annotations : Any" dropdown. A table header is shown with columns for "Name", "Type", "Related", and "Annotations". In the center, there is a large "No linked service to show" message with a "Create linked service" button below it. The browser address bar shows the URL <https://ms.web.azuresynthesize.net/en-us/management/datalinkedservices?workspace=%2Fsubscriptions%2F...>.

Click the “**Create linked service**” button.

New linked service

key vault

All Azure Compute Database File Generic protocol NoSQL

Azure Key Vault

Search for and then select “**Azure Key Vault**”.

Click the **Continue** button.

New linked service (Azure Key Vault)

Name *

Description

Azure key vault selection method From Azure subscription Enter manually

Azure subscription

Azure key vault name *

Annotations

Parameters

Advanced

Commit Back Test connection Cancel

On the “New linked service (Azure Key Vault)” pop-out, enter values for the following items:

Name	Enter a value that is meaningful and aligned with standards
...Selection Method	Confirm default selection, “From Azure subscription”
Subscription	Select your subscription
Azure Key Vault Name	Select your instance of Key Vault

Click “Test connection” to confirm successful connection and then click the **Create** (or **Commit**) button.

Data Explorer

Navigate to your Synapse Analytics workspace and then click the **Open** link on the “Open Synapse Studio” rectangle.

Click the **Manage** icon and then “Linked Services” in the left-hand navigation.

Click the “Create linked service” button.

The screenshot shows the Microsoft Azure Synapse Analytics portal interface. The left sidebar navigation includes options like Analytics pools, SQL pools, Apache Spark pools, External connections, Integration, Security, Access control, Credentials, and Managed private endpoints. The main content area is titled "New linked service" and shows a search bar for "data explorer". Below the search bar, there are tabs for All, Azure, Compute, Database, File, Generic protocol, and NoSQL. A list of existing linked services is displayed, with one item highlighted: "Azure Data Explorer (Kusto)". At the bottom of the dialog are "Continue" and "Cancel" buttons.

Search for and then select “**Azure Data Explorer (Kusto)**”.

Click the **Continue** button.

The screenshot shows the Microsoft Azure portal interface for a Synapse Analytics workspace named "practicumsw". The left sidebar contains navigation links for Analytics pools, SQL pools, Apache Spark pools, External connections, Linked services, Azure Purview (Preview), Integration (Triggers, Integration runtimes), Security, Access control, Credentials, Managed private endpoints, Code libraries, Workspace packages, Source control, and Git configuration. The "Linked services" section is selected.

The main content area displays the "New linked service (Azure Data Explorer (Kusto))" configuration dialog. It includes fields for Name (set to "practicumded"), Description, Connect via integration runtime (set to "AutoResolveIntegrationRuntime"), Authentication method (set to "Managed Identity"), Account selection method (radio button selected for "From Azure subscription"), Azure subscription (set to "rchapler [REDACTED]"), Cluster (set to "practicumdec"), Database (set to "practicumded"), and Annotations. A note at the bottom states: "Managed identity name: **practicumsw**
Managed identity object ID: **88a20c05-f533-4d5b-8163-55f5f814f915**
Grant workspace service managed identity access to your Azure Data Explorer (Kusto). [Learn more](#)". At the bottom right, there are "Commit", "Back", "Test connection" (with a checkmark indicating success), and "Cancel" buttons.

On the “New linked service (Azure Data Explorer (Kusto))” pop-out, enter values for the following items:

Name	Enter a value that is meaningful and aligned with standards
Authentication Method	Select “ Managed Identity ”
Account Selection Method	Confirm default selection, “From Azure subscription”
Azure Subscription	Select your subscription
Cluster	Select your Data Explorer Cluster
Database	Select your Data Explorer Database

Click “Test connection” to confirm successful connection and then click the **Create** (or **Commit**) button.

Data Lake

Navigate to your Synapse Analytics workspace and then click the **Open** link on the “Open Synapse Studio” rectangle.

Click the **Manage** icon and then “**Linked Services**” in the left-hand navigation.

Click the “**Create linked service**” button.

The screenshot shows the Microsoft Azure portal interface for creating a new linked service. The left sidebar navigation includes options like 'Analytics pools', 'SQL pools', 'Apache Spark pools', 'External connections', 'Linked services' (which is selected and highlighted in blue), 'Azure Purview (Preview)', 'Integration', 'Triggers', 'Integration runtimes', 'Security', 'Access control', 'Credentials', 'Managed private endpoints', 'Code libraries', 'Workspace packages', 'Source control', and 'Git configuration'. The main content area is titled 'New linked service' and has a search bar with 'data lake'. Below the search bar is a filter tab bar with 'All' selected, followed by 'Azure', 'Compute', 'Database', 'File', 'Generic protocol', and 'NoSQL'. A table displays two items: 'Azure Data Lake Storage Gen1' and 'Azure Data Lake Storage Gen1 for Cosmos Structured Stream'. The third item, 'Azure Data Lake Storage Gen2', is highlighted with a blue border. At the bottom right are 'Continue' and 'Cancel' buttons.

Search for and then select “**Azure Data Lake Storage Gen2**”.

Click the **Continue** button.

The screenshot shows the Microsoft Azure portal interface for a Synapse Analytics workspace named "practicumsw". The left sidebar navigation includes options like "Analytics pools", "SQL pools", "Apache Spark pools", "External connections", "Linked services" (which is currently selected), "Azure Purview (Preview)", "Integration", "Triggers", "Integration runtimes", "Security", "Access control", "Credentials", "Managed private endpoints", "Code libraries", "Workspace packages", "Source control", and "Git configuration".

The main content area is titled "New linked service (Azure Data Lake Storage Gen2)". It provides instructions to avoid publishing immediately to workspace and to use Azure Key Vault for retrieving secrets securely. It prompts the user to choose a name for the linked service, which is set to "practicumdl".

The "Description" field is empty. Under "Connect via integration runtime", the dropdown is set to "AutoResolveIntegrationRuntime". The "Authentication method" is set to "Account key". The "Account selection method" is set to "From Azure subscription", and the selected subscription is "rchapler". The "Storage account name" is set to "practicumdl".

At the bottom, there are "Commit" and "Back" buttons, and a "Test connection" button which is highlighted with a green checkmark indicating a successful connection.

On the “New linked service (Azure Data Lake Storage Gen2)” pop-out, enter values for the following items:

Name	Enter a value that is meaningful and aligned with standards
Connect via...	Confirm default selection, “AutoResolveIntegrationRuntime”
Authentication Method	Confirm default selection, “Account key”
Azure Subscription	Select your subscription
Storage Account Name	Select your Data Lake

Click “Test connection” to confirm successful connection and then click the **Create** (or **Commit**) button.

SQL

Navigate to your Synapse Analytics workspace and then click the **Open** link on the “Open Synapse Studio” rectangle.

Click the **Manage** icon and then “Linked Services” in the left-hand navigation.

Click the “Create linked service” button.

The screenshot shows the Microsoft Azure Synapse Analytics portal. The left sidebar navigation includes: Home, collaboration branch, Analytics pools, SQL pools, Apache Spark pools, External connections, **Linked services** (which is selected), Azure Purview (Preview), Integration, Triggers, Integration runtimes, Security, Access control, Credentials, Managed private endpoints, Code libraries, Workspace packages, Source control, and Git configuration.

The main content area is titled "New linked service". A search bar at the top right contains "azure sql". Below it, a filter bar has "All" selected. Two items are listed:

- Azure SQL Database (represented by a blue cylinder icon)
- Azure SQL Database Managed Instance (represented by a blue cylinder icon with a cloud and a gear)

At the bottom right are "Continue" and "Cancel" buttons.

Search for and then select "**Azure SQL Database**".

Click the **Continue** button.

Screenshot of the Microsoft Azure Synapse Analytics portal showing the creation of a new linked service (Azure SQL Database). The left sidebar shows navigation options like Analytics pools, SQL pools, External connections, and Linked services. The main area displays the 'New linked service (Azure SQL Database)' configuration page.

New linked service (Azure SQL Database)

Name *: practicumsd

Description: (empty)

Connect via integration runtime *: AutoResolveIntegrationRuntime

Account selection method: From Azure subscription

Azure subscription: rchaper [dropdown]

Server name *: practicumsds

Database name *: practicumsd

Authentication type *: SQL authentication

User name *: rchapler

AKV linked service *: practicumkv

Secret name *: practicumsds-adminpassword

Secret version: Use the latest version if left blank

Always encrypted: (checkbox)

Additional connection properties: + New

Annotations: + New

Parameters: + Advanced

Commit | **Back** | **Connection successful** | **Test connection** | **Cancel**

On the “New linked service (Azure SQL Database)” pop-out, enter values for the following items:

Name	Enter a value that is meaningful and aligned with standards
Connect via...	Confirm default selection, “AutoResolveIntegrationRuntime”
Account Selection Method	Confirm default selection, “From Azure subscription”
Azure Subscription	Select your subscription

Server Name	Select your Azure SQL Server
Database Name	Select your Azure SQL Database
Authentication Type	Select “SQL authentication”
User Name	Enter the “server admin login” value used during instantiation of the SQL Database Server
AKV Linked Service	Select the name of the Linked Service created for the Key Vault
Secret Name	Enter the Secret Name used to capture the Azure SQL Server administrator password

Click “Test connection” to confirm successful connection and then click the **Create** (or **Commit**) button.

Apache Spark Pool

Navigate to your Synapse Analytics workspace and then click the **Open** link on the “Open Synapse Studio” rectangle.

Click the **Manage** icon and then “Apache Spark pools” in the left-hand navigation.

The screenshot shows the Microsoft Azure Synapse Analytics portal. The URL in the browser is <https://ms.web.azuresynthesize.net/en-us/management/apachesparkpools?workspace=%2Fsubscription%2Fpracticumsw>. The page title is "practicumsw - Microsoft Azure". The left sidebar has a tree view with "collaboration branch" expanded, showing "Analytics pools", "SQL pools", "Apache Spark pools" (which is selected), "External connections", "Linked services", "Azure Purview (Preview)", "Integration", "Triggers", "Integration runtimes", "Security", "Access control", "Credentials", "Managed private endpoints", "Code libraries", "Workspace packages", "Source control", and "Git configuration". The main content area has a heading "Apache Spark pool" with a sub-instruction: "Apache Spark pools can be tuned to run different kinds of Apache Spark workloads using specific configuration libraries, permissions, etc. [Learn more](#)". Below this are buttons for "+ New" and "Refresh". A search bar says "Filter by name" and shows "Showing 0-0 of 0 item". A table header includes columns "Name", "Node size family", and "Size". In the center, there's a large "New" button with a plus sign icon and the text "No items to show. Try changing your filter or create new Apache Spark pool".

Click the “New Apache Spark pool” button.

The screenshot shows the Microsoft Azure Synapse Analytics portal. The left sidebar has a tree view with 'Analytics pools' selected. The main area is titled 'New Apache Spark pool' under the 'Basics' tab. It includes fields for 'Apache Spark pool name' (practicumasp), 'Isolated compute' (Disabled), 'Node size family' (Memory Optimized), 'Node size' (Small (4 vCores / 32 GB)), 'Autoscale' (Disabled), and 'Number of nodes' (set to 3). An estimated price of 1.66 USD per hour is shown. At the bottom are 'Review + create', 'Next: Additional settings >', and 'Cancel' buttons.

On the “New Apache Spark pool” pop-out, enter values for the following items:

Apache Spark Pool Name	Enter a value that is meaningful and aligned with standards
Isolated Compute	Confirm default, Disabled
Node Size Family	Confirm default, “ Memory Optimized ”
Node Size	Select “ Small (4 vCores / 32 GB) ”
Autoscale	Select Disabled
Number of Nodes	Slide to lowest possible value (to minimize demonstration cost)
Estimated Price	Review final “Est. cost per hour” and view pricing details, as desired

Click the “Review + create” button, validate, and then click the **Create** button.