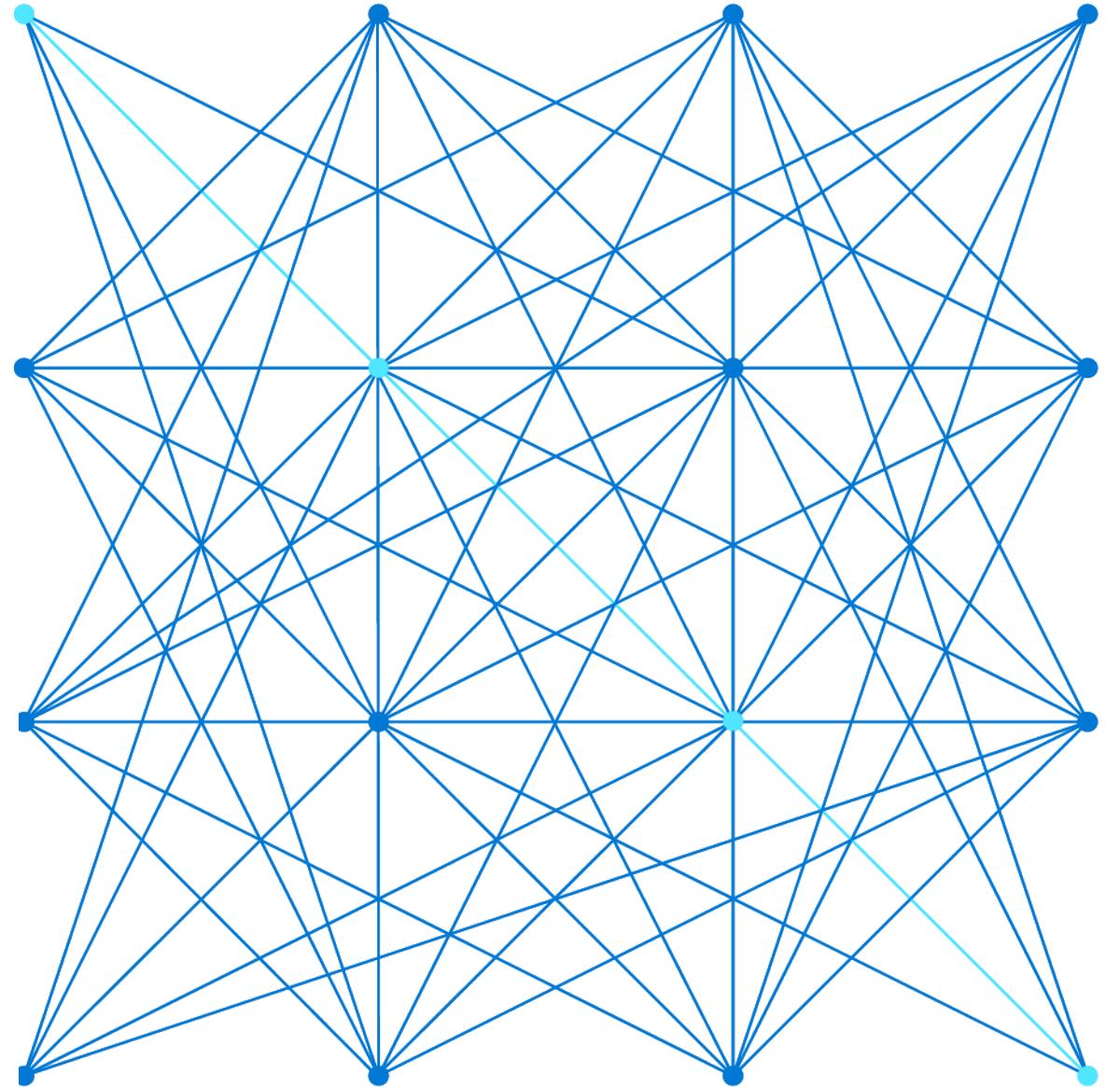


# DA-100 Analyzing Data with Power BI 2

Prof. Ernesto Lee



# **Module 2: Get Data in Power BI**

# Learning Objectives

You will learn the following concepts:

- Getting data from various data sources
- Optimizing Performance
- Resolving data errors



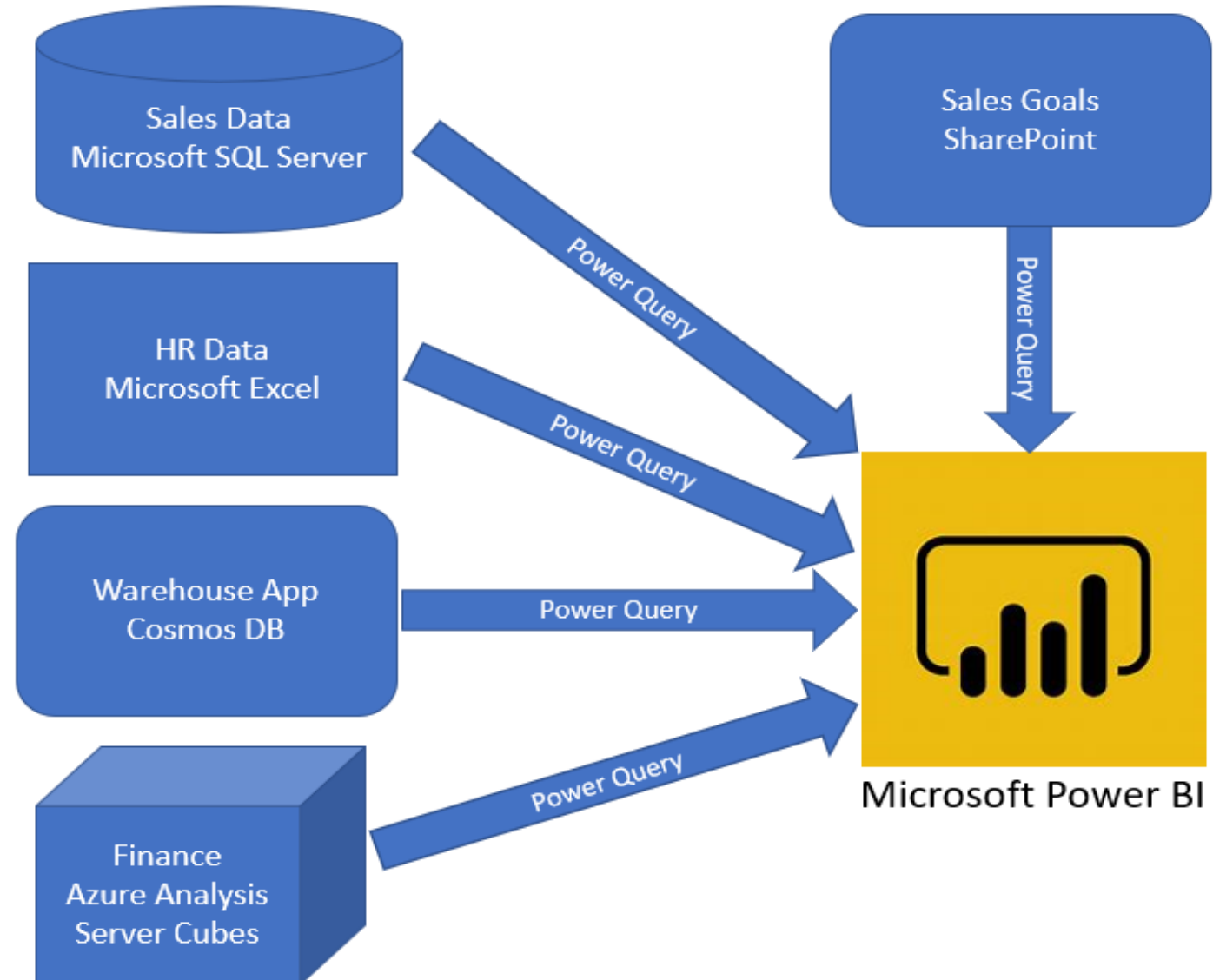
# Lesson 1: Get data from various data sources





# Introduction to getting data

The first step in the data analysis process is identifying and getting data.



# Get data from flat files



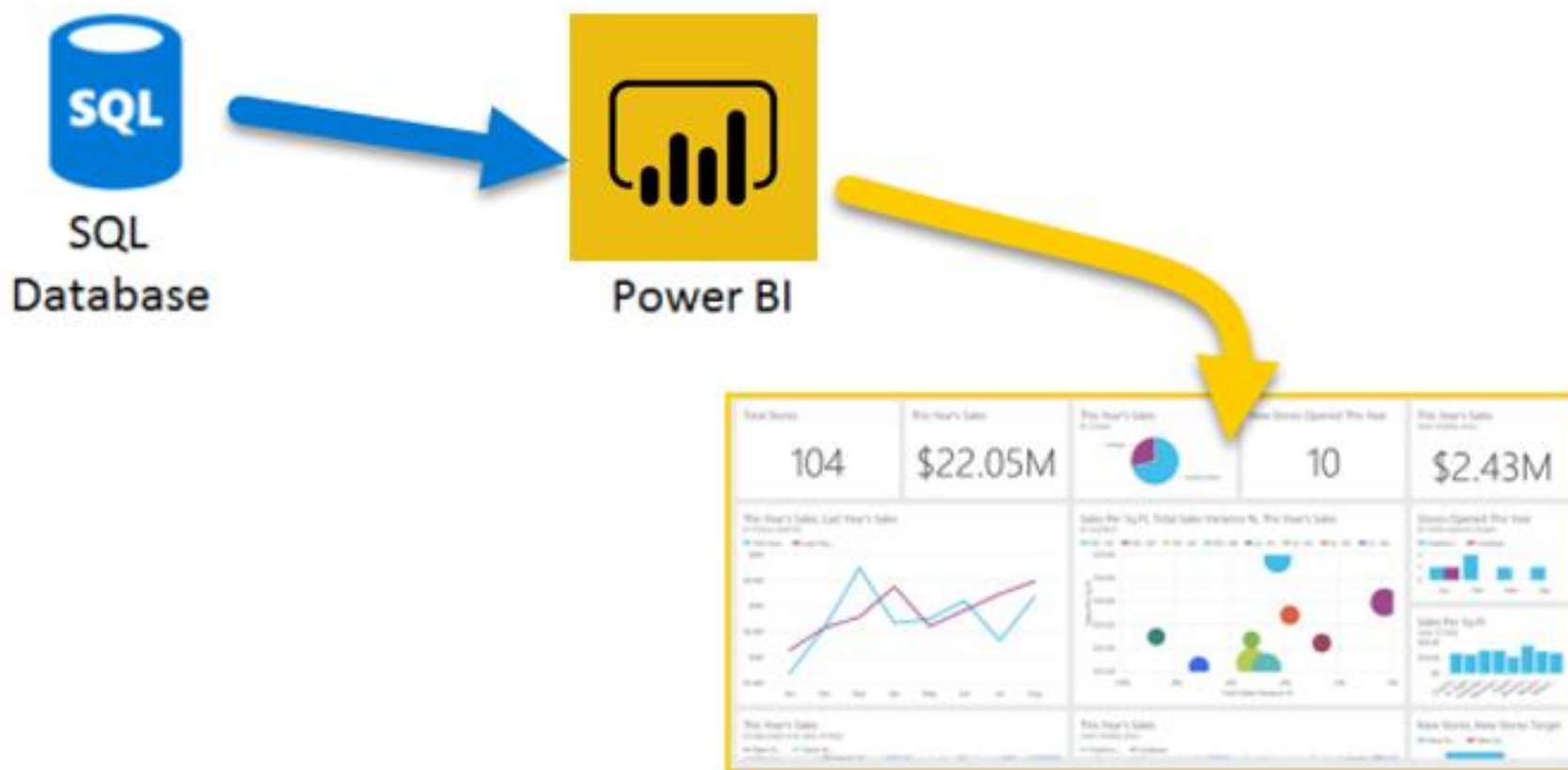
ResellerSales\_202006.csv

File Origin: 1200: Unicode | Delimiter: Comma | Data Type Detection: Based on first 200 rows

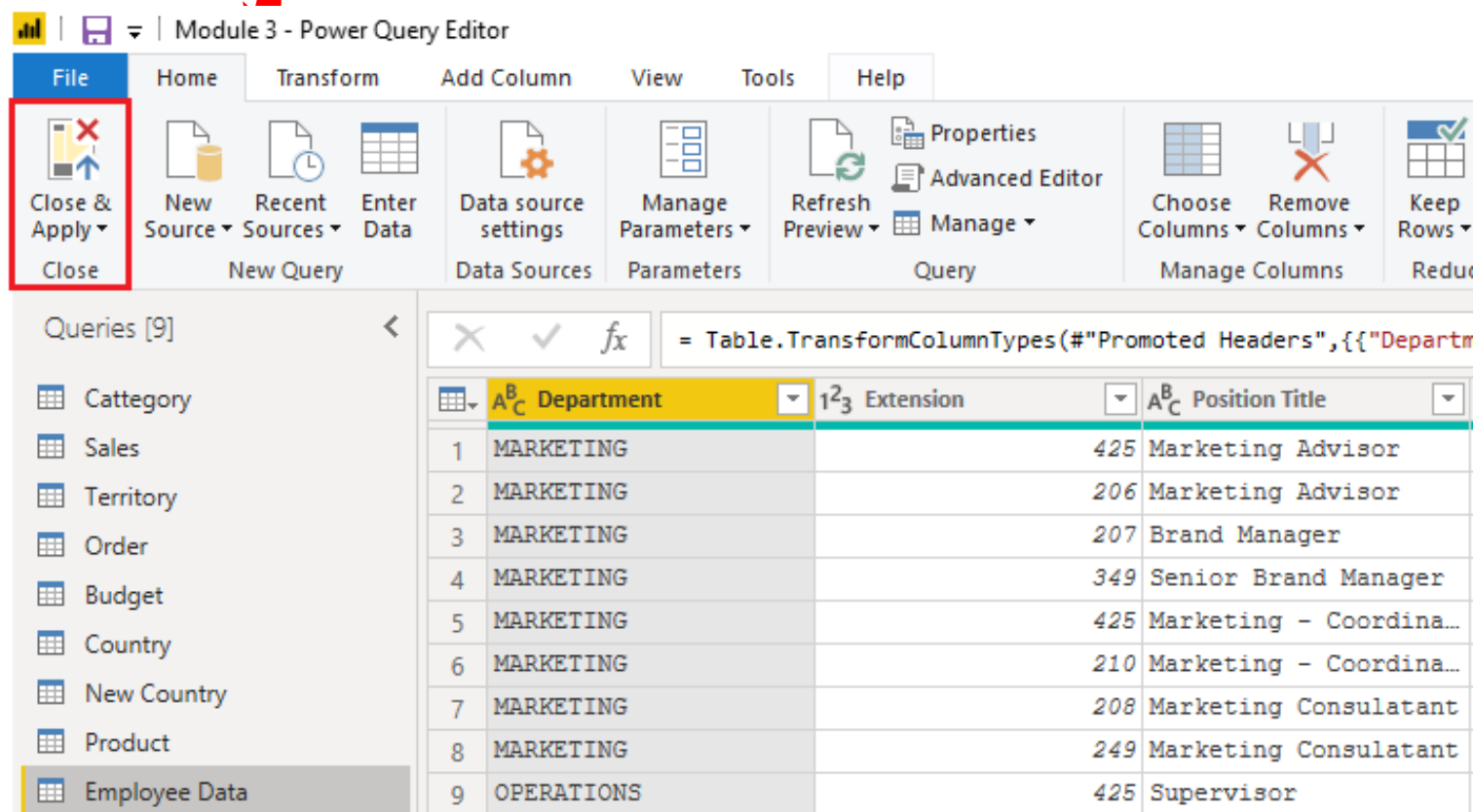
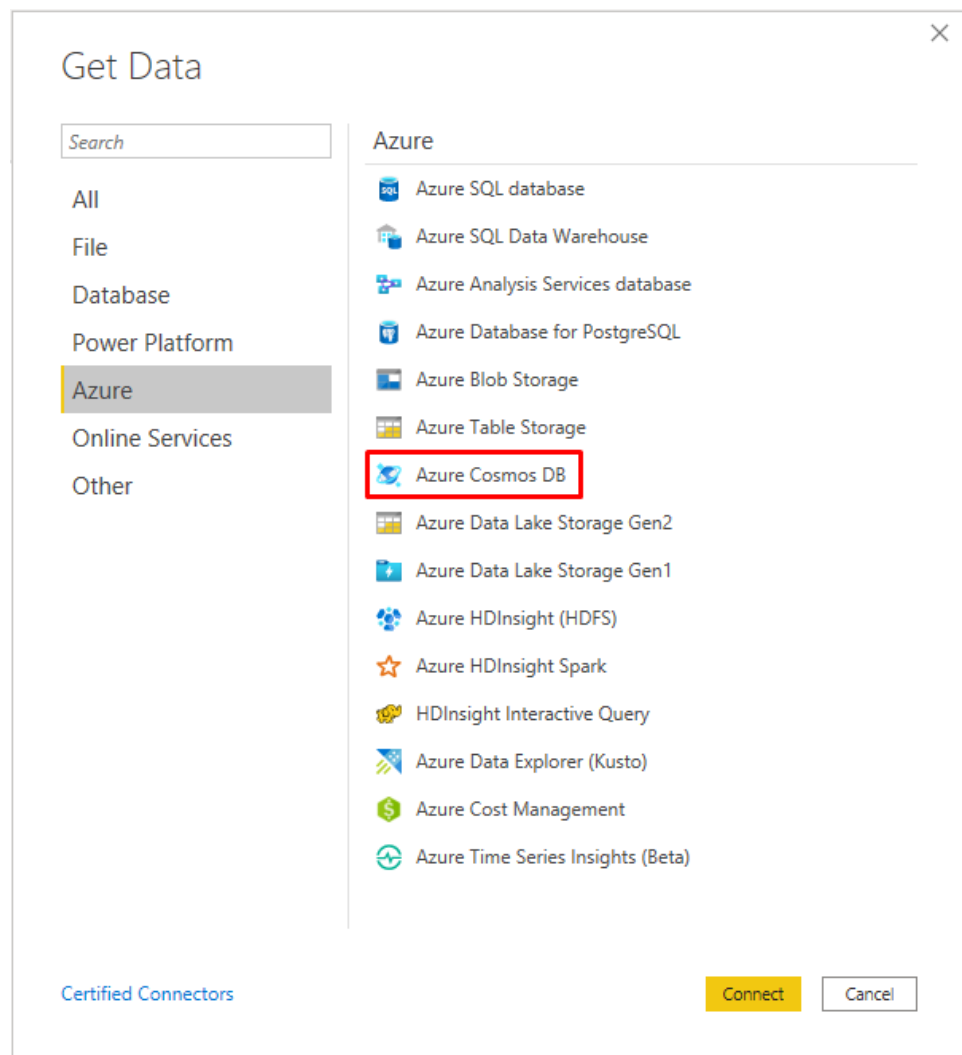
SalesOrderNumber	SalesOrderLineNumber	OrderDate	DueDate	ShipDate	ProductKey	ResellerKey	PromotionKey
SO71691	2	6/1/2020	6/11/2020	6/8/2020	434	104	
SO71691	4	6/1/2020	6/11/2020	6/8/2020	222	104	
SO71774	1	6/1/2020	6/11/2020	6/8/2020	436	609	
SO71774	2	6/1/2020	6/11/2020	6/8/2020	418	609	
SO71775	1	6/1/2020	6/11/2020	6/8/2020	573	595	
SO71775	2	6/1/2020	6/11/2020	6/8/2020	555	595	
SO71775	3	6/1/2020	6/11/2020	6/8/2020	490	595	
SO71776	1	6/2/2020	6/12/2020	6/9/2020	514	106	
SO71777	1	6/2/2020	6/12/2020	6/9/2020	408	128	
SO71777	2	6/2/2020	6/12/2020	6/9/2020	436	128	
SO71778	1	6/2/2020	6/12/2020	6/9/2020	467	557	
SO71778	2	6/2/2020	6/12/2020	6/9/2020	566	557	

Load Transform Data Cancel

# Get data from relational data sources

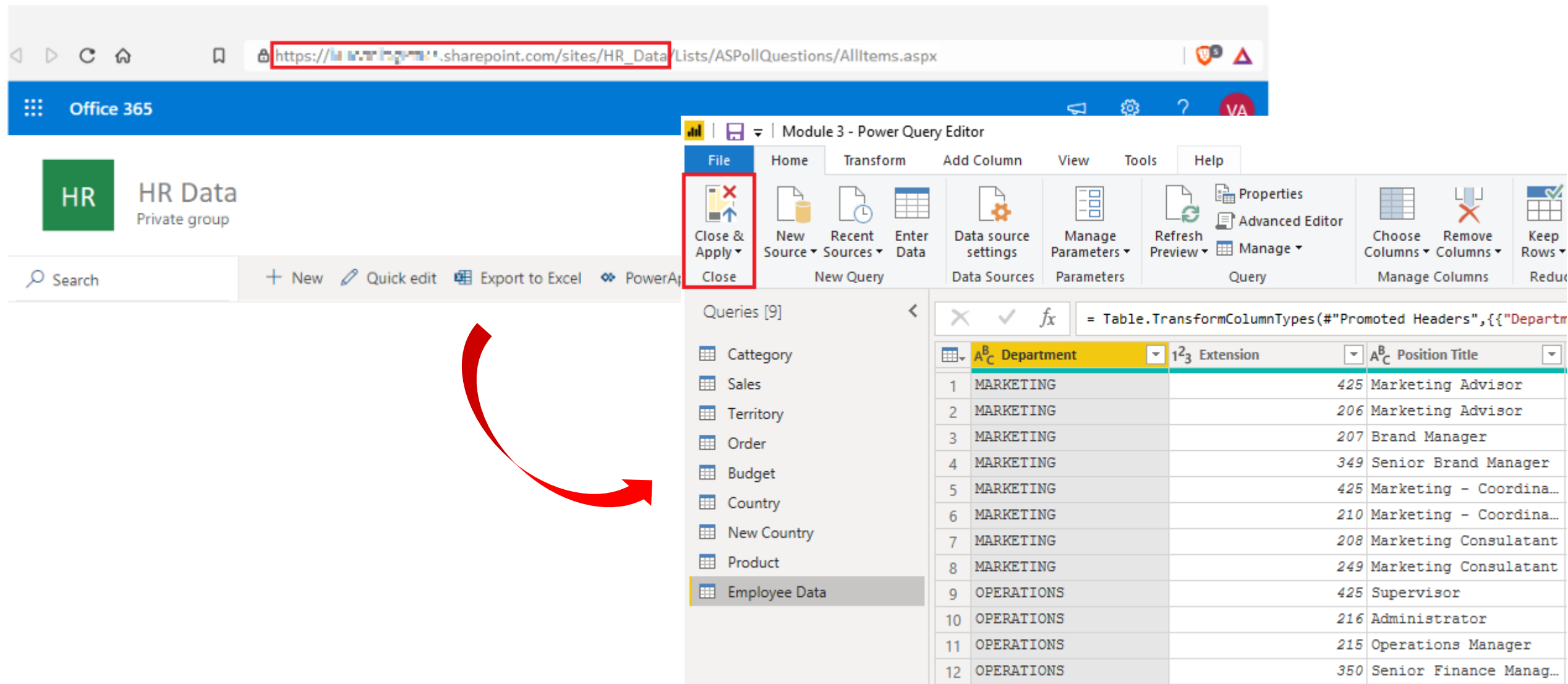


# Get data from NoSQL





# Get data from applications



The screenshot shows a SharePoint list named 'ASPollQuestions' with the following URL highlighted in a red box: [https://\[redacted\].sharepoint.com/sites/HR\\_Data/Lists/ASPollQuestions/AllItems.aspx](https://[redacted].sharepoint.com/sites/HR_Data/Lists/ASPollQuestions/AllItems.aspx). The list is part of the 'HR Data' private group in an Office 365 environment.

Overlaid on the list is the 'Module 3 - Power Query Editor'. The 'File' tab is active, and the 'Close & Apply' button is highlighted with a red box. A red arrow points from this button to the 'Employee Data' query in the 'Queries [9]' pane on the left.

The 'Employee Data' query is selected, and the data table is displayed. The table has columns for 'Department', 'Extension', and 'Position Title'. The data is as follows:

	Department	Extension	Position Title
1	MARKETING	425	Marketing Advisor
2	MARKETING	206	Marketing Advisor
3	MARKETING	207	Brand Manager
4	MARKETING	349	Senior Brand Manager
5	MARKETING	425	Marketing - Coordina...
6	MARKETING	210	Marketing - Coordina...
7	MARKETING	208	Marketing Consulatant
8	MARKETING	249	Marketing Consulatant
9	OPERATIONS	425	Supervisor
10	OPERATIONS	216	Administrator
11	OPERATIONS	215	Operations Manager
12	OPERATIONS	350	Senior Finance Manag...

# Get data from Analysis Services

An analytical data engine that lets you digest data from multiple data sources and create calculations on the fly.

## SQL Server Analysis Services database

Server ⓘ

asazure://westus.asazure.windows.net/azureanalysiservicesiketest

Database (optional)

adventureworks

☐ Import

☒ Connect live

▸ MDX or DAX query (optional)

OK

Cancel

# Review Questions

- Q01 – Which query language do you use to extract data from Microsoft SQL Server?
- A01 – T-SQL
- Q02 – You're creating a Power BI report with data from an Azure Analysis Services Cube. When the data refreshes in the cube, you would like to see it immediately in the Power BI report. How should you connect?
- A02 – Connect Live
- Q03 – What can you do to improve performance when you are getting data in Power BI?
- A03 – Do some calculations in the original data source.



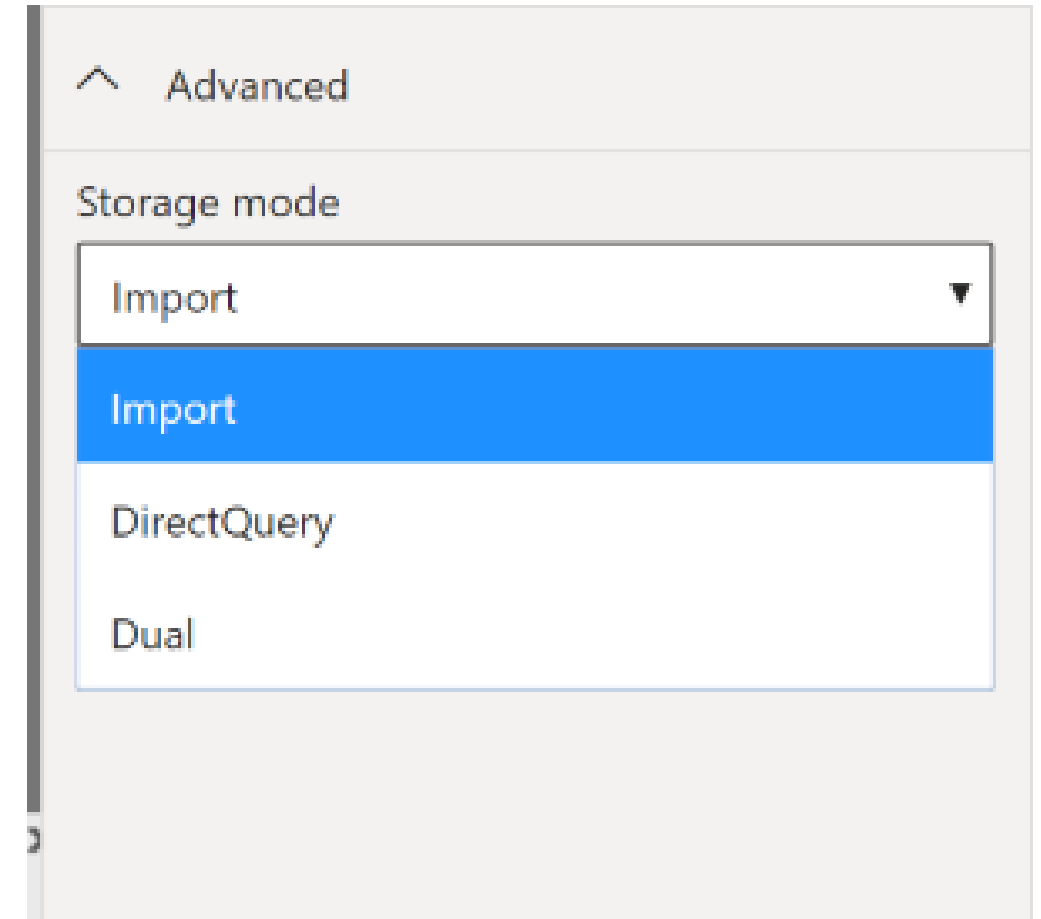
# Lesson 2: Optimize Performance



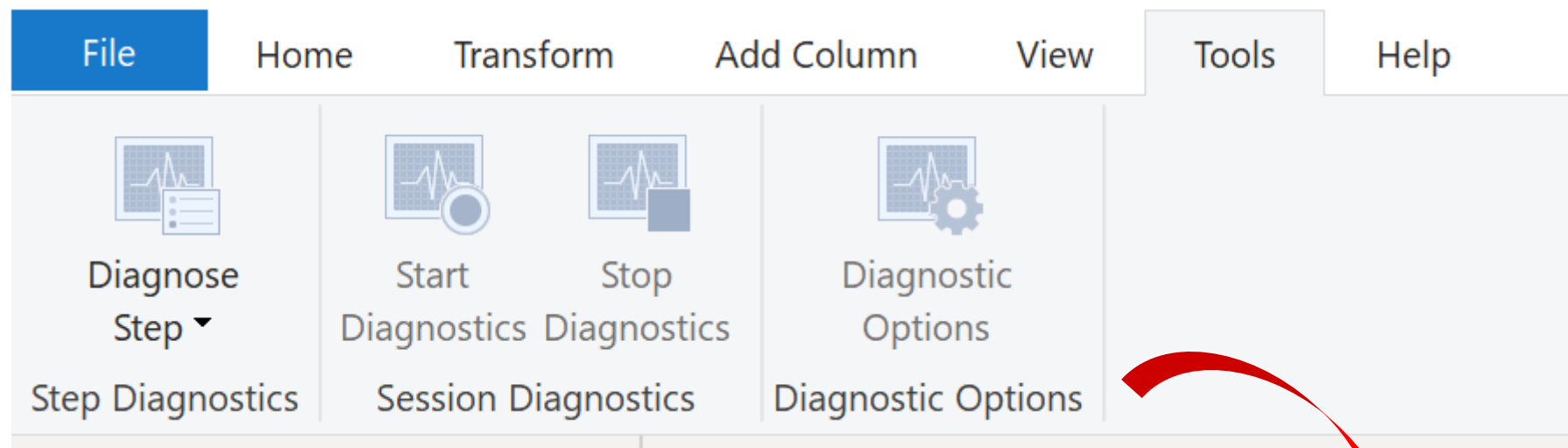


# Select a Storage Mode

- Specifies the storage mode of a table and lets Power BI determine how to cache data for reports.
- Set the storage mode for each table individually.



# Fix Performance Issues



Queries [17] <

- ▲ Diagnostics [2]
  - 📄 Diagnostics\_Detailed...
  - 📄 Diagnostics\_2020-0...

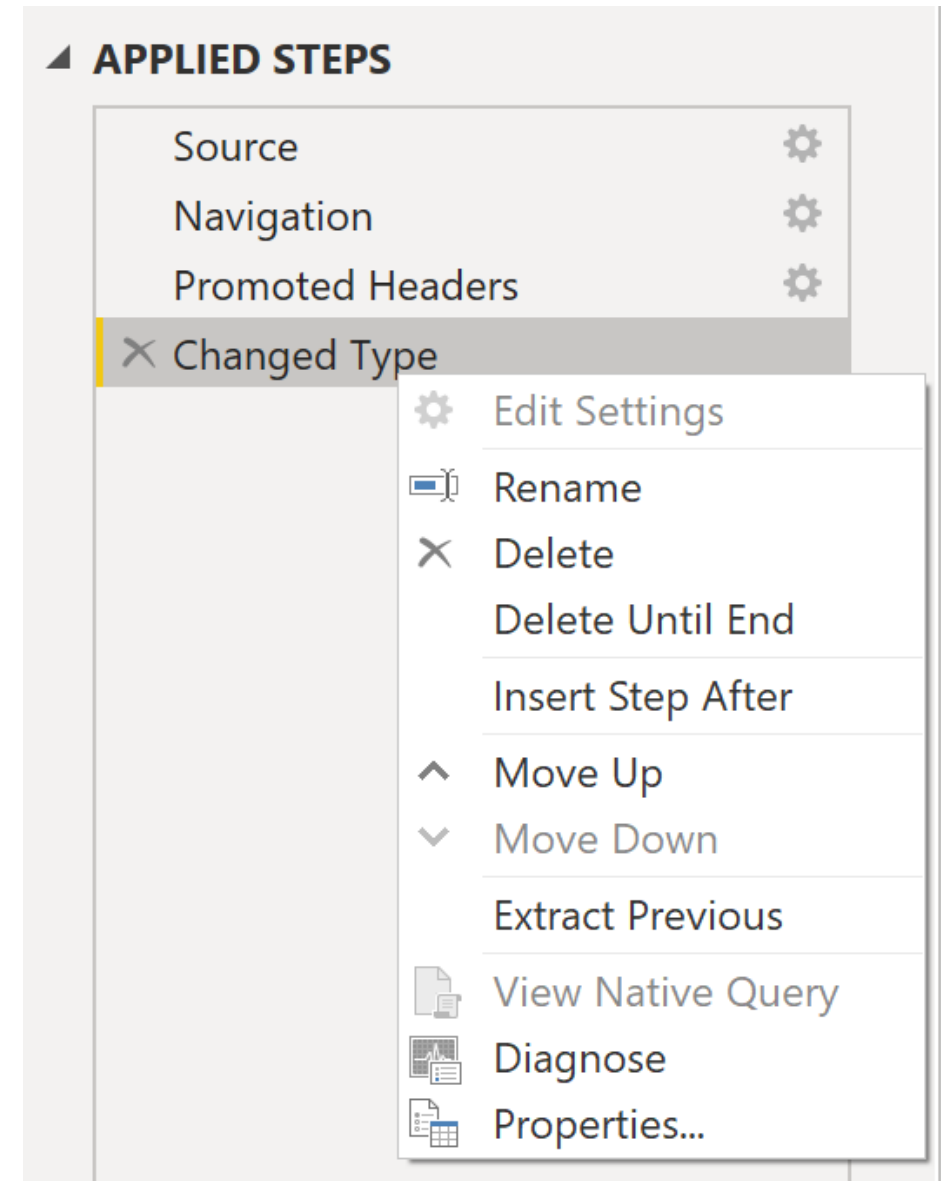
	AB_C Id	AB_C Query	AB_C Step	🕒 Exclusive Duration	AB_C Category	AB_C Data Source Kind
1	1.1	Product	Changed Type	0.00:00:00.0457545	Evaluator	null
2	1.2	Product	Source	0.00:00:01.9567741	Evaluator	null

# Optimize Query Performance

- Performance in Power Query depends heavily on the performance at the data source.
- Follow performance tuning guidelines of the source product.
- Some performance tuning can be done in Power BI.

# Query Folding

The process that lets Power Query generate a single query statement to retrieve and transform source data.





# Review Questions

- Q01 – Which storage mode leaves the data at the data source?
- A01 – DirectQuery
- Q02 – Which technology improves performance by generating a single query statement to retrieve and transform source data?
- A02 – Query Folding



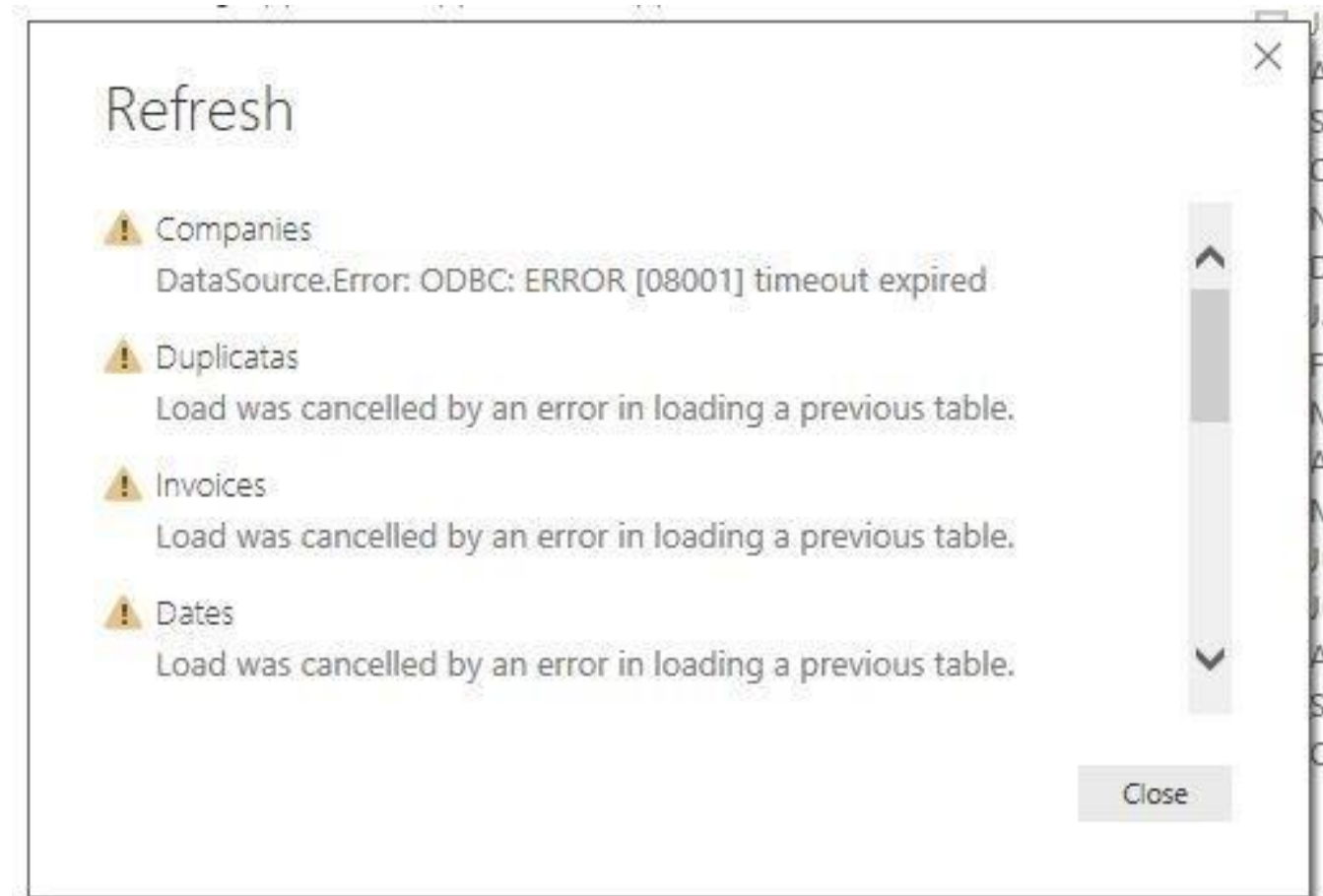
# Lesson 3: Resolve Data Errors





# Identify and Resolve Data Import Errors

- You may encounter the following errors:
  - Query Timeout.
  - Couldn't find data formatted as a table.
  - Could not find file.
  - Data type errors.



# Review Questions

- Q01 – What type of import error might leave a column blank?
- A01 – Data type error



# Module Overview

We covered the following concepts:

- Getting data from various data sources
- Optimizing Performance
- Resolving data errors

# Lab: CONNECT TO DATA

# References

- DA-100 Prepare data for analysis  
<https://docs.microsoft.com/en-us/learn/modules/get-data/>
- DA-100 Clean, Transform, and load data in Power BI  
<https://docs.microsoft.com/en-us/learn/modules/clean-data-power-bi/>

