

# Caloudi 教育訓練課程

## Azure Synapse Analytics Hands-On Lab

# Agenda

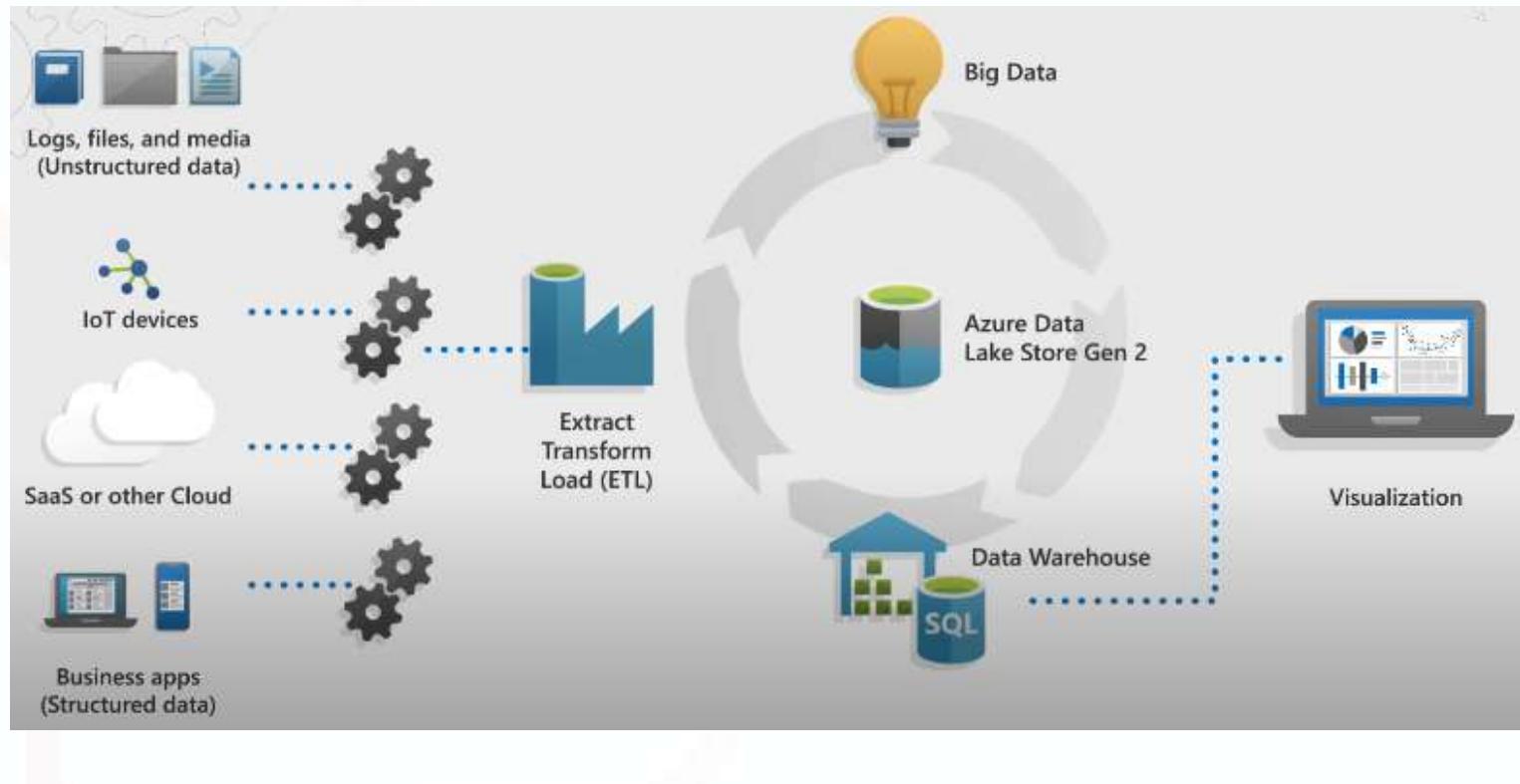
- Hands-On Lab 架構
- 建立 Synapse Workspace
- 使用 Azure Open Datasets
- 建立 Apache Spark Pool
- 建立 Dedicated SQL Pool
- 透過 Power BI 展現資料
- 課後作業



# Hands-On Lab 架構

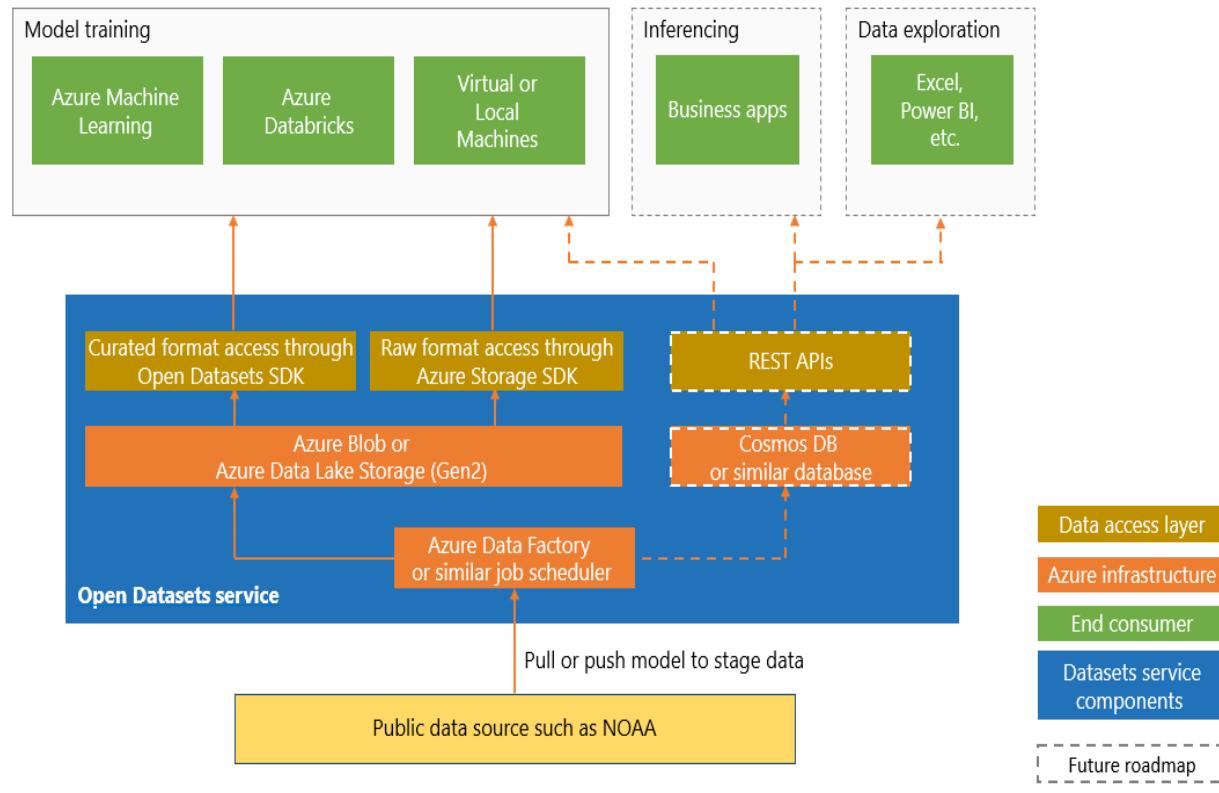
# Scenario

- SQL Database on Azure
- Azure Data Factory
- Azure Data Lake Storage Gen2
- Apache SQL Pool
- Dedicated SQL Pool
- Power BI Service/Desktop



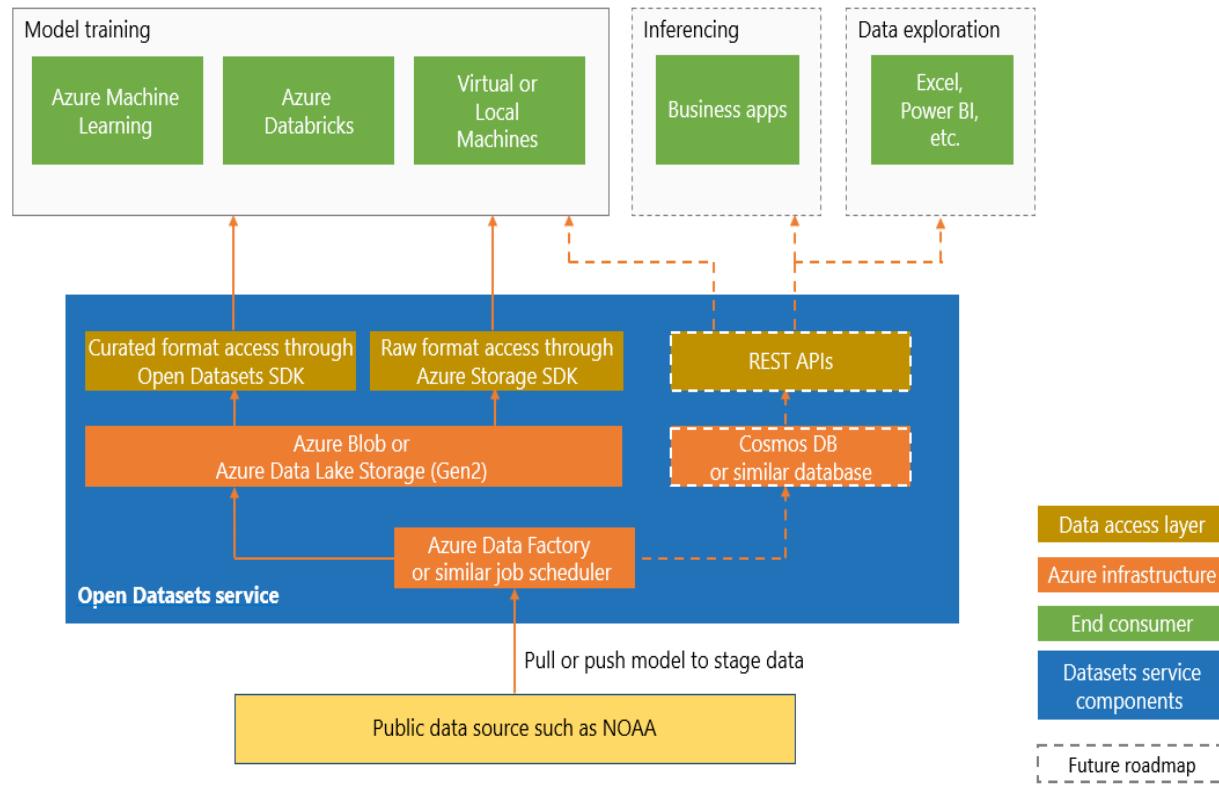
# Azure Open Datasets

## Azure Open Datasets service: Building blocks



# Azure Open Datasets

## Azure Open Datasets service: Building blocks



# New York City Safety Data

All New York City 311 service requests from 2010 to the present.

The screenshot shows the NYC 311 website homepage. At the top, there is a navigation bar with links for Services, News, Government, COVID-19 Vaccine, and a search bar. Below the navigation is a blue header bar with the text "Local and Regional Authorities" and the "NYC 311" logo. On the right side of the blue bar is a yellow "NYC 311" logo. Below the blue bar, there is a "SHARE" button and an envelope icon. A black sidebar on the left contains links for "SECTIONS", "Our Mission", "Core Services", "Who We Serve", and "Contact Information". The main content area features a section titled "Our Mission" with text about 311's mission and how it helps government bodies manage their workload efficiently. The text reads:

311's mission is to provide the public with quick, easy access to all New York City government services and information while maintaining the highest possible level of customer service.

311 allows other government bodies to focus on their core missions and manage their workload efficiently. 311 also helps City agencies improve their service delivery by providing accurate and consistent data tracking and analysis of all service requests.

下載網址：[Azure Open Datasets](#)

參考資料：  
New York City Safety Data  
NYC 311

# Microsoft Contoso BI Demo Dataset

A fictitious retail demo dataset used for presenting Microsoft Business Intelligence products.

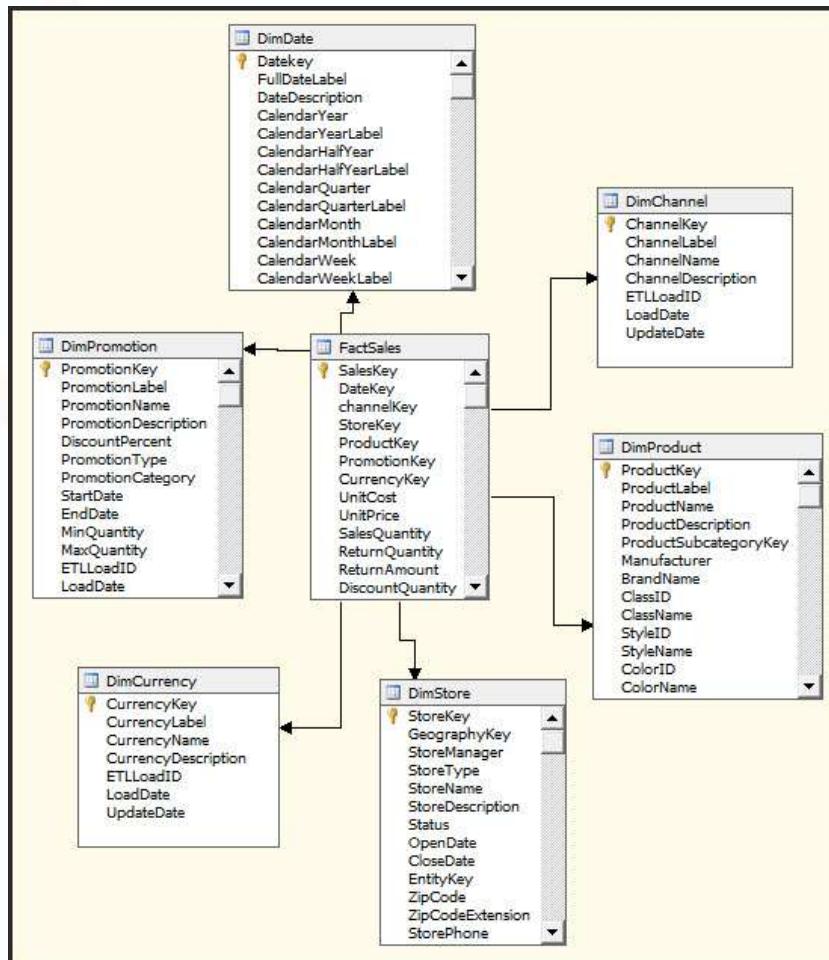
“



*The Contoso BI Demo dataset is used to demonstrate DW/BI functionalities across the entire Microsoft Office product family. This dataset includes C-level, sales/marketing, IT, and common finance scenarios for the retail industry and support map integration. In addition, this dataset offers large volumes of transactions from OLTP and well-structured aggregations from OLAP, along with reference and dimension data.*

下載網址：[Microsoft Contoso BI Demo Dataset for Retail Industry](#)

# Contoso BI Demo Dataset Schema



參考資料：[Adding Dimensions to your Cube](#)

# Dimensional Model

Customer = Dimension

	Customer Identifier
	Frist Name
	Last Name
	Address 1 Text
	Address 2 Text
	City Name
	State Code
	Zip Code
	Phone Number
	Email Address Text

Supplier = Dimension

	Supplier Identifier
	Supplier Name

Time = Dimension

	Time Identifier
	Calendar Day Number
	Calendar Weekday Name
	Calendar Week Number
	Calendar Month Name
	Calendar Quarter Number
	Calendar Year Number

Sales Order = Fact

	Customer Identifier (FK)
	Employee Number (FK)
	Supplier Identifier (FK)
	Product Identifier (FK)
	Time Identifier (FK)
	Location Identifier (FK)

Quantity Count = Measure  
Net Amount = Measure

Employee = Dimension

	Employee Number
	Employee Frist Name
	Employee Last Name
	Sex Code
	Birthday Date
	Employee Phone Number

Product = Dimension

	Product Identifier
	Product Name
	Product Description
	Product Unit Amount

Location = Dimension

	Location Identifier
	Location Name
	Location Description
	City Name
	State Code
	Zip Code
	Country Name

# System Requirement

上課學員必須：

- 有 Subscription 的 Owner 權限
  1. 要在 Subscription 裡面動態產生 Resource Group
  2. 要幫 Storage Account 加入 Storage Blob Data Contributor 權限給 Synapse Workspace 與學員帳號
- 有桌機或是 VM 之類的電腦環境
  1. 要有 Browser
  2. 要有可以下載檔案的權限
  3. 要在電腦裡頭安裝 Power BI Desktop
- 有 Power BI Pro 帳號
  1. Power BI Pro 帳號與 Azure Portal 帳號相同
  2. 可以是正式的 Subscription，也可以是有效的試用帳號

# 建立 Synapse Workspace

# (Workspace) Step 01

- 登入 Azure Portal
- 搜尋 Azure Synapse Analytics

The screenshot shows the Microsoft Azure portal interface. The top navigation bar has 'Microsoft Azure' and a search bar containing 'Synapse'. The user's email 'kc.su@auovide.onmicrosoft.com' is visible in the top right. The left sidebar includes 'Azure services' (with 'Create a resource' and 'More services'), 'Navigate' (with 'Subscriptions' and 'Dashboard'), and a 'Did you mean?' section with 'Try searching in Activity Log' and 'Try searching in Azure Active Directory'. The main content area displays search results under 'Services': 'Azure Synapse Analytics' (selected), 'Azure Synapse Analytics (private link hubs)', and 'Schemas'. Under 'Marketplace', there are links to 'Azure Synapse Analytics (private link hubs)', 'Azure Synapse Analytics', 'Datometry Hyper-Q for Azure Synapse Analytics', and 'Xpert BI with Azure Synapse'. Under 'Documentation', there are links to 'Azure Synapse Analytics - Azure Synapse Analytics ...', 'What is Azure Synapse Analytics? - Azure Synapse Analytics ...', 'Microsoft.Synapse/workspaces - ARM template reference ...', and 'Plan to manage costs for Azure Synapse Analytics - Azure ...'. Under 'Resource Groups', it says 'No results were found.'

# (Workspace) Step 02

- 按下 + Create，建立 Synapse Workspace

Microsoft Azure Search resources, services, and docs (G+/-) kc.su@auovide.onmicro... AUO VIDE (AUOVIDE.ONMICROS...)

Home > Azure Synapse Analytics AUO vIDE (auovide.onmicrosoft.com)

+ Create Manage view Refresh Export to CSV Open query Assign tags Feedback

Filter for any field... Subscription == all Resource group == all Location == all Add filter

Showing 0 to 0 of 0 records. No grouping List view

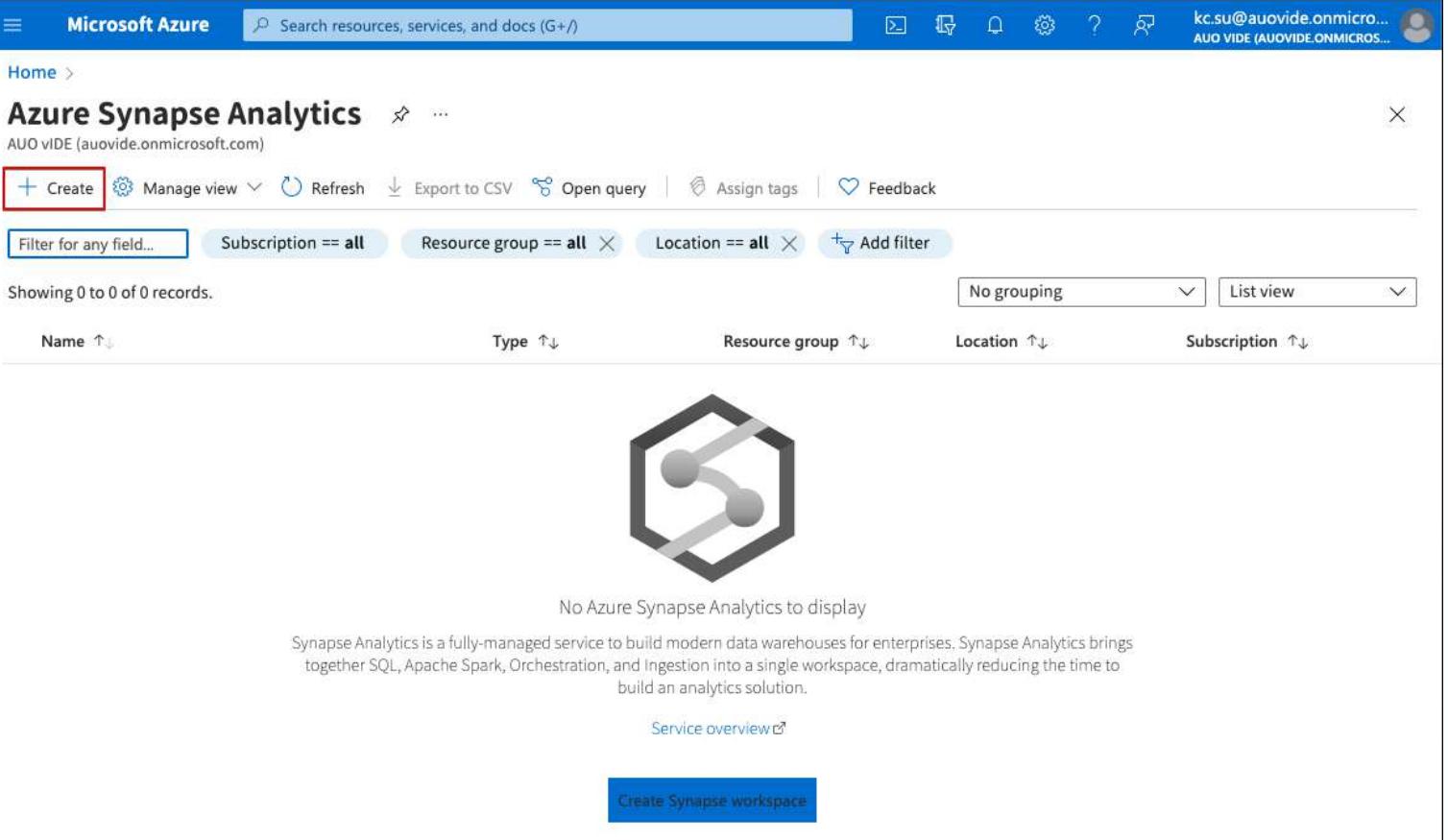
Name ↑↓	Type ↑↓	Resource group ↑↓	Location ↑↓	Subscription ↑↓
---------	---------	-------------------	-------------	-----------------

No Azure Synapse Analytics to display

Synapse Analytics is a fully-managed service to build modern data warehouses for enterprises. Synapse Analytics brings together SQL, Apache Spark, Orchestration, and Ingestion into a single workspace, dramatically reducing the time to build an analytics solution.

Service overview

Create Synapse workspace



# (Workspace) Step 03 (Basics)

- Resource group : 按下 Create new , 建立 M\_員工編號\_RG
- Managed resource group : M\_員工編號\_RG\_DC
- Workspace name 設定 synapseworkspaceM員工編號
- Account name 按下 Create new 設定 adlsg2m員工編號
- File system name 按下 Create new 設定 data

Microsoft Azure Search resources, services, and docs (G+/-) kc.su@auovide.onmicro... AUO VIDE (AUOVIDE.ONMICRO...)

Home > Azure Synapse Analytics >

Azure Synapse Analytics < AUO VIDE (auovide.onmicrosoft.com)

+ Create Manage view ...

Filter for any field... Name ↑

No Azure Synapse Analytics to display

Synapse Analytics is a fully-managed service to build modern data warehouses for enterprises. Synapse Analytics brings together SQL, Apache Spark, Orchestration, and Ingestion into a single workspace, dramatically reducing the time to build an analytics solution.

Service overview ↗

Create Synapse workspace

## 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 \*  ⓘ The Synapse and SQL resource providers are now registered with this subscription.

Resource group \*  ⓘ Create new

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 \*  ⓘ Create new

File system name \*  ⓘ Create new

# (Workspace) Step 04 (Security)

- SQL Server admin login 輸入 sqladminuser
- SQL Password 與 Confirm password 自行設定
- 勾選 Allow pipelines to access SQL pools.

Microsoft Azure Search resources, services, and docs (G+) kc.su@auovide.onmicrosoft.com AUO VIDE (AUOVIDE.ONMICROS...)

Home > Azure Synapse Analytics > Azure Synapse Analytics <

**Create Synapse workspace**

\* Basics \* **Security** Networking Tags Review + create

Configure security options for your workspace.

**SQL administrator credentials**  
Provide credentials that can be used for administrator access to the workspace's SQL pools. If you don't provide a password, one will be automatically generated. You can change the password later.

SQL Server admin login \*  sqladminuser

SQL Password  .....

Confirm password  .....

**System assigned managed identity permission**  
Choose the permissions that you would like to assign to the workspace's system-assigned identity. [Learn more](#)

Allow pipelines (running as workspace's system assigned identity) to access SQL pools. ①

Allow network access to Data Lake Storage Gen2 account. ②

ⓘ The selected Data Lake Storage Gen2 account does not restrict network access using any network access rules, or you selected a storage account manually via URL under Basics tab. [Learn more](#)

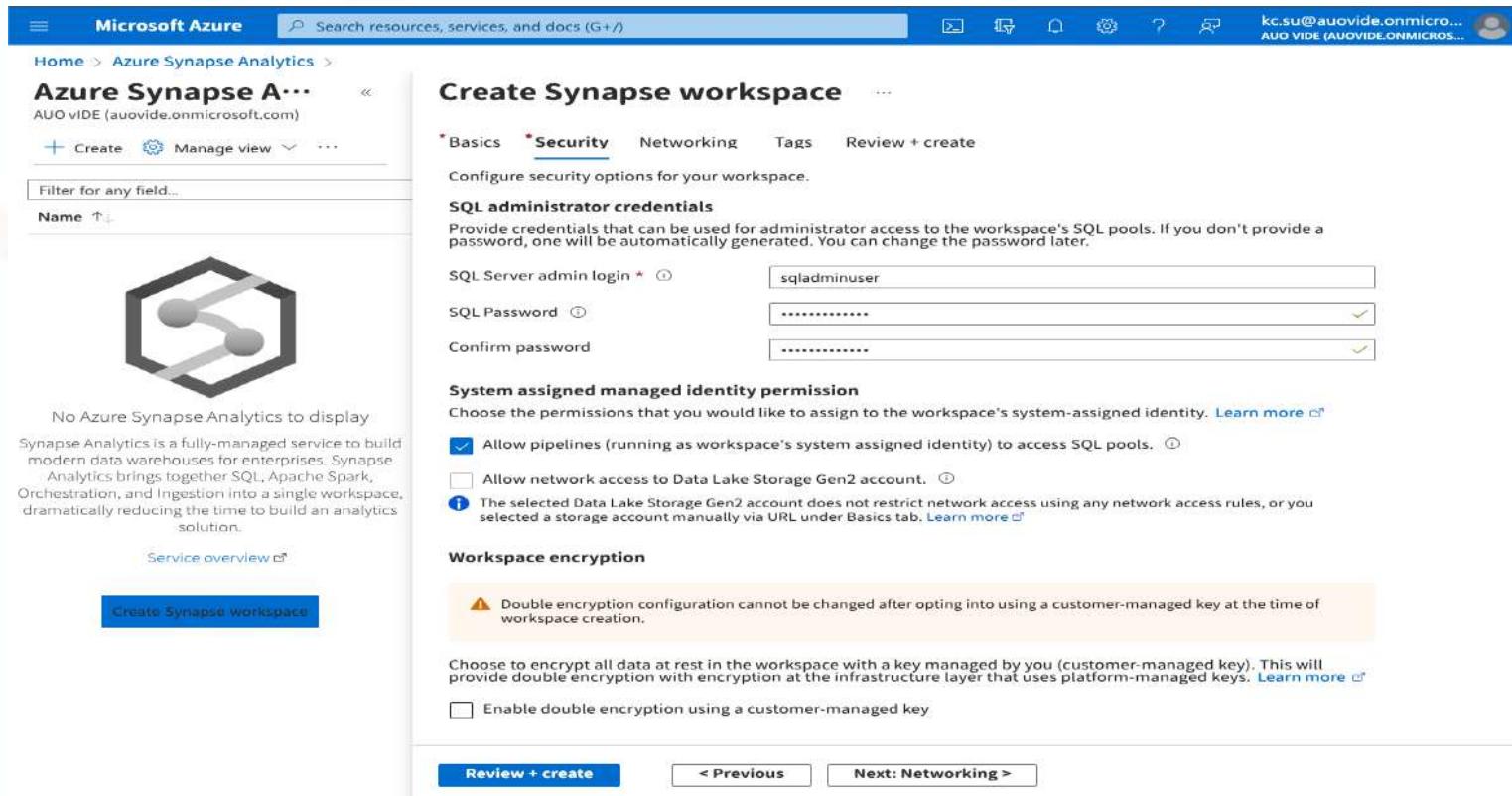
**Workspace encryption**

⚠ Double encryption configuration cannot be changed after opting into using a customer-managed key at the time of workspace creation.

Choose to encrypt all data at rest in the workspace with a key managed by you (customer-managed key). This will provide double encryption with encryption at the infrastructure layer that uses platform-managed keys. [Learn more](#)

Enable double encryption using a customer-managed key

**Review + create** < Previous Next: Networking >



# (Workspace) Step 05 (Networking)

- 勾選 Allow connections from all IP addresses (Provision 之後再限制 IP)
- 按下 Review + create 按鈕

Microsoft Azure Search resources, services, and docs (G+/-) kc.su@auovide.onmicrosoft.com AUO VIDE (AUOVIDE.ONMICROS...)

Home > Azure Synapse Analytics >

## Azure Synapse A...

AUO VIDE (auovide.onmicrosoft.com)

+ Create Manage view ...

Filter for any field...

Name ↑



No Azure Synapse Analytics to display

Synapse Analytics is a fully-managed service to build modern data warehouses for enterprises. Synapse Analytics brings together SQL, Apache Spark, Orchestration, and Ingestion into a single workspace, dramatically reducing the time to build an analytics solution.

Service overview ↗

### Create Synapse workspace

\* Basics \* Security Networking Tags Review + create

Configure networking settings for your workspace.

**Allow connections from all IP addresses**

**⚠** Azure Synapse Studio and other client tools will only be able to connect to the workspace endpoints if this setting is allowed. Connections from specific IP addresses or all Azure services can be allowed/disallowed after the workspace is provisioned.

Allow connections from all IP addresses to your workspace's endpoints. You can restrict this to just Azure datacenter IP addresses and/or specific IP address ranges after creating the workspace.

Allow connections from all IP addresses

**Managed virtual network**

Choose whether you want a Synapse-managed virtual network dedicated for your Azure Synapse workspace. [Learn more ↗](#)

Enable managed virtual network ⓘ

**Create Synapse workspace** Review + create < Previous Next: Tags >

# (Workspace) Step 06

- 按下 Create 按鈕

Microsoft Azure Search resources, services, and docs (G+/)

kc.su@auovide.onmicrosoft.com AUO VIDE (AUOVIDE.ONMICRO...)

Home > Azure Synapse Analytics >

## Azure Synapse Analytics

AUO VIDE (auovide.onmicrosoft.com)

+ Create Manage view ...

Filter for any field...

Name ↑↓



No Azure Synapse Analytics to display

Synapse Analytics is a fully-managed service to build modern data warehouses for enterprises. Synapse Analytics brings together SQL, Apache Spark, Orchestration, and Ingestion into a single workspace, dramatically reducing the time to build an analytics solution.

Service overview ↗

Create Synapse workspace

### Create Synapse workspace

Validation succeeded

\* Basics \* Security Networking Tags Review + create

#### Product Details

Azure Synapse Analytics workspace by Microsoft Serverless SQL est. cost/TB ⓘ  
127.73 TWD  
[Terms of use](#) | [Privacy policy](#)

#### Terms

By clicking Create, I (a) agree to the legal terms and privacy statement(s) associated with the Marketplace offering(s) listed above; (b) authorize Microsoft to bill my current payment method for the fees associated with the offering(s), with the same billing frequency as my Azure subscription; and (c) agree that Microsoft may share my contact, usage and transactional information with the provider(s) of the offering(s) for support, billing and other transactional activities. Microsoft does not provide rights for third-party offerings. For additional details see [Azure Marketplace Terms](#).

#### Basics

Subscription	azuretrain01
Resource group	(new) M00_RG
Region	Southeast Asia

**Create** < Previous Next > Download a template for automation

# (Workspace) Step 07

- 檢視 Synapse Workspace

Screenshot of the Microsoft Azure portal showing the Synapse workspace details for 'synapseworkspacem00'.

The page title is 'synapseworkspacem00' - Synapse workspace.

The left sidebar navigation includes:

- 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
- Security
  - Encryption
  - Firewalls
  - Managed identities
  - Private endpoint connections

The main content area displays the following workspace details:

Setting	Value
Resource group (change)	M00_RG
Status	Succeeded
Location	Southeast Asia
Subscription (change)	azuretrain01
Subscription ID	a278626c-1d76-422d-a5b2-4c2cf191a0ac
Managed virtual network	No
Managed Identity object ID	b6a978ff-324d-4417-9e7a-774e8614d44c
Workspace web URL	<a href="https://web.azure.synapse.net?workspace=%2bsubscriptions%2b%2bsynapseworkspacem00">https://web.azure.synapse.net?workspace=%2bsubscriptions%2b%2bsynapseworkspacem00</a>
Tags (change)	<a href="#">Click here to add tags</a>
Firewalls	<a href="#">Show firewall settings</a>
Primary ADLS Gen2 account URL	<a href="https://adlsg2m00.dfs.core.windows.net">https://adlsg2m00.dfs.core.windows.net</a>
Primary ADLS Gen2 file system data	
SQL admin username	sqladminuser
SQL Active Directory admin	<a href="#">kc.su@auovide.onmicrosoft.com</a>
Dedicated SQL endpoint	<a href="#">synapseworkspacem00.sql.azuresynapse.net</a>
Serverless SQL endpoint	<a href="#">synapseworkspacem00-on-demand.sql.azuresynapse.net</a>
Development endpoint	<a href="#">https://synapseworkspacem00.dev.azuresynapse.net</a>

**Getting started**

- Open Synapse Studio**  
Start building your fully-integrated analytics solution and unlock new insights.  
[Open](#)
- Read documentation**  
Learn how to be productive quickly. Explore concepts, tutorials, and samples.  
[Learn more](#)

# (Workspace) Step 07

- Workspace web URL  
`https://web.azure-synapse.net?workspace=...`
- Firewalls 可以 Allow / Disallow 特定 IP
- Primary ADLS Gen2 account URL / file system  
`https://adlsg2m00.dfs.core.windows.net` 與 `data`
- SQL admin username : `sqladminuser`
- Dedicated SQL endpoint  
`synapseworkspace00.sql.azure-synapse.net`
- Serverless SQL endpoint  
`synapseworkspace00-on-demand.sql.azure-synapse.net`
- Development endpoint  
`https://synapseworkspace00.dev.azure-synapse.net`
- 可以 + New dedicated SQL pool / + New Apache Spark pool
- 可以 Open Synapse Studio

# (Workspace) Step 08

- 按下 Open Synapse Studio，或是連上 <https://web.azuresynthesize.net/>

## Getting started



### Open Synapse Studio

Start building your fully-integrated analytics solution and unlock new insights.

[Open ↗](#)



### Read documentation

Learn how to be productive quickly. Explore concepts, tutorials, and samples.

[Learn more ↗](#)

## Analytics pools

Search to filter items...

Name	Type	Size
<strong>SQL pools</strong>		
Built-in	Serverless	Auto
<strong>Apache Spark pools</strong>		
No pools provisioned		

# (Workspace) Step 09

Synapse Studio 畫面：

The screenshot shows the Microsoft Azure Synapse Studio workspace interface for the workspace "synapseworkspacem00". The top navigation bar includes "Microsoft Azure", "Synapse Analytics", and the workspace name "synapseworkspacem00". The top right corner shows the user email "kc.su@auovide.onmicrosoft.com" and a video camera icon. The left sidebar has a "New" button and several icons: Home, Data, Compute, Machine Learning, and Storage. The main area displays three key features: "Ingest" (Perform a one-time or scheduled data load), "Explore and analyze" (Learn how to get insights from your data), and "Visualize" (Build interactive reports with Power BI capabilities). A large, stylized graphic in the background features a globe with network connections and a bar chart. At the bottom, there are links for "Discover more", "Knowledge center", and "Browse partners".

# (Workspace) Step 10

Resource Group 畫面：

The screenshot shows the Microsoft Azure Resource Group (M00\_RG) overview page. The top navigation bar includes the Microsoft Azure logo, a search bar, and user information (kc.su@auovide.onmicrosoft.com). The main content area displays the following details:

- Subscription (change):** azuretrain01
- Deployment:** 1 Succeeded
- Subscription ID:** a278626c-1d76-422d-a5b2-4c2cf191a0ac
- Location:** Southeast Asia
- Tags (change):** Click here to add tags

Below this, there are filter options: Filter for any field..., Type == all, Location == all, and Add filter. The table below shows two records:

Name	Type	Location	Actions
adlsg2m00	Storage account	Southeast Asia	...
synapseworkspaceM00	Synapse workspace	Southeast Asia	...

# (Workspace) Step 11

Storage Account 的 IAM 要設定 Synapse Workspace Storage Blob Data Contributor :

The screenshot shows the Microsoft Azure portal interface for managing IAM roles on a storage account named 'adlsg2m00'. On the left, the 'Access Control (IAM)' section is selected. In the center, the 'Add role assignment' dialog is open, showing the 'Role' dropdown set to 'Storage Blob Data Contributor' and the 'Assign access to' dropdown set to 'User, group, or service principal'. A search bar contains the text 'syna'. Below the search bar, it says 'No users, groups, or service principals found.' On the right, a table lists one item: 'Contributor' for user 'davidhsu3' (email 'davidhsu3@...' and type 'User'). At the bottom of the dialog are 'Save' and 'Discard' buttons.

# (Workspace) Step 12

Storage Account 的 IAM 設定完成畫面：

The screenshot shows the Microsoft Azure portal interface for managing IAM roles in a storage account. The top navigation bar includes the Microsoft Azure logo, a search bar, and user information (kc.su@auovide.onmicrosoft.com). The current page is 'adlsg2m00 | Access Control (IAM)' under the 'Storage account' section.

The left sidebar contains navigation links: Overview, Activity log, Tags, Diagnose and solve problems, Access Control (IAM) (which is selected), Data migration, Events, and Storage Explorer (preview).

The main content area has a header with actions: + Add, Download role assignments, Edit columns, Refresh, Remove, and Got feedback?. Below this is a summary: 'Number of role assignments for this subscription' (16 / 2000).

Filtering options include: Search by name or email, Type : All, Role : Storage Blob Data Contributor (highlighted in blue), Scope : All scopes, and Group by : Role.

A message box states: 'Showing a filtered set of results. Total number of role assignments: 13'.

The table displays 1 item:

Name	Type	Role	Scope	Condition
Storage Blob Data Contributor	App	Storage Blob Data Contributor	This resource	Add

The table row for the 'Storage Blob Data Contributor' role is highlighted in blue.

# (Workspace) Step 13

Storage Account 的 IAM 要設定你的帳號 Storage Blob Data Contributor :

The screenshot shows the Microsoft Azure portal interface for managing IAM roles on a Storage Account named 'adlsg2m00'. The left sidebar lists various services like Overview, Activity log, Tags, and Storage Explorer (preview). The 'Access Control (IAM)' section is selected and highlighted with a red box. The main pane displays the 'Role assignments' tab, showing 16 assignments out of a total capacity of 2000. A modal window titled 'Add role assignment' is open on the right, prompting for a role ('Storage Blob Data Contributor') and an access target ('User, group, or service principal'). The target field contains the email 'kc.su@auovide.onmicrosoft.com', which is also highlighted with a red box. Below the target field, a message states 'No users, groups, or service principals found.' The 'Selected members:' section shows one entry: 'kc.su' with the email 'kc.su@auovide.onmicrosoft.com', accompanied by a 'Remove' link. At the bottom of the modal are 'Save' and 'Discard' buttons.

# (Workspace) Step 14

Storage Account 的 IAM 設定完成畫面：

The screenshot shows the Microsoft Azure portal interface for managing Azure role assignments. The top navigation bar includes the Microsoft Azure logo, a search bar, and various navigation icons. The main title is "kc.su | Azure role assignments". On the left, a sidebar titled "Manage" lists several options: Profile, Assigned roles, Administrative units, Groups, Applications, Licenses, Devices, Azure role assignments (which is selected and highlighted in grey), and Authentication methods. The main content area displays a table of role assignments for the user "kc.su" in the subscription "azuretrain01". The table columns are Role, Resource Name, Resource Type, Assigned To, and Condition. The data shows three entries:

Role	Resource Name	Resource Type	Assigned To	Condition
Owner	azuretrain01	Subscription	kc.su	None
Contributor	azuretrain01	Subscription	kc.su	None
Storage Blob Data Contributor	adlsg2m00	Storage account	kc.su	None

# 使用 Azure Open Datasets

# (Dataset) Step 01

- 按下 Knowledge center
- 按下 Browse gallery

Microsoft Azure | Synapse Analytics > synapseworkspace00

Home > Knowledge center

## Knowledge center

Get started with Azure Open Datasets and sample code. Return to the knowledge center periodically as we provide updated content.



**Experience limitless scale**

Deliver insights from all your data, across data warehouses and big data analytics systems, with blazing speed.

● ○ ○ ○

**Use samples immediately**  
Click once and we'll create everything you need, from scripts and notebooks to pools and data.

**Browse gallery**  
Select from sample code and Azure Open Datasets to quickly get started in your workspace.

**Tour Synapse Studio**  
Familiarize yourself with key features of Synapse Studio. Start by taking a tour of the homepage.

# (Dataset) Step 02

- 在 Datasets 找 New York City Safety Data，然後按下 Continue 按鈕

The screenshot shows the Microsoft Azure Synapse Analytics Datasets gallery. The top navigation bar includes 'Microsoft Azure', 'Synapse Analytics', 'synapseworkspacem00', and a user profile for 'kc.su@auovide.onmicrosoft.com'. The left sidebar has a 'Gallery' section with icons for Home, Datasets, Notebooks, SQL scripts, and Pipelines. The 'Datasets' tab is selected. A search bar says 'Filter by keyword' and 'Tags : All'. Below are eight dataset cards:

- NYC Taxi & Limousine Commission - yellow taxi trip records**: Includes a sample link.
- New York City Safety Data**: This card is highlighted with a blue border. Description: "This dataset contains all New York City 311 service requests from 2010 to the present. It's stored in Parquet format." ID: city\_safety\_newyork. Includes a sample link.
- Oxford COVID-19 Government Response Tracker**: ID: oxford-covid-19-govern...
- Public Holidays**: Worldwide public holiday data sourced from PyPI holidays package and Wikipedia, covering 38 countries. ID: public\_holiday. Includes a sample link.
- Sample: Diabetes**
- San Francisco Safety Data**
- Seattle Safety Data**
- US Consumer Price Index**

At the bottom are 'Continue' and 'Close' buttons.

# (Dataset) Step 03

- 按下 Add dataset，把 Dataset 加到 Synapse Studio 的 Data 標籤頁

The screenshot shows the Microsoft Azure Synapse Analytics dataset details page for 'New York City Safety Data'. The top navigation bar includes 'Microsoft Azure', 'Synapse Analytics', 'synapseworkspacem00', and a user profile. The main content area has a sidebar with icons for Home, Datasets, Pipelines, and Monitoring, followed by sections for 'Description', 'Volume and Retention', 'Storage Location', 'Additional Information', and 'Notices'. The 'Description' section states: 'All New York City 311 service requests from 2010 to the present.' The 'Volume and Retention' section notes: 'This dataset is stored in Parquet format. It is updated daily, and contains about 12M rows (500MB) in total as of 2019.' The 'Storage Location' section indicates: 'This dataset is stored in the East US Azure region. Allocating compute resources in East US is recommended for affinity.' The 'Additional Information' section links to terms of use: 'This dataset is sourced from New York City government. More details can be found from [here](#). Reference [here](#) for the terms of using this dataset.' The 'Notices' section is currently empty. To the right, a 'Preview' table displays sample data with columns: dataType, dataSubtype, dateTime, category, subcategory, and status. The data shows various service requests like Street Conditions, Noise, and Illegal Parks.

dataType	dataSubtype	dateTime	category	subcategory	status
Safety	311_All	7/7/2021 2:...	Street Cond...	Pothole	Open
Safety	311_All	7/7/2021 2:...	Noise - Stre...	Loud Music...	In Progress
Safety	311_All	7/7/2021 2:...	Illegal Parki...	Blocked Hy...	In Progress
Safety	311_All	7/7/2021 2:...	Noise - Veh...	Car/Truck ...	In Progress
Safety	311_All	7/7/2021 2:...	Blocked Dri...	No Access	In Progress
Safety	311_All	7/7/2021 2:...	Homeless P...	Chronic	Assigned
Safety	311_All	7/7/2021 2:...	Homeless P...	Chronic	Assigned
Safety	311_All	7/7/2021 2:...	Noise - Resi...	Banging/Po...	In Progress
Safety	311_All	7/7/2021 2:...	Noise - Stre...	Loud Music...	In Progress
Safety	311_All	7/7/2021 2:...	Illegal Parki...	Blocked Hu...	In Progress

At the bottom, there are buttons for 'Add dataset' (highlighted in blue), 'Back', and 'Close'.

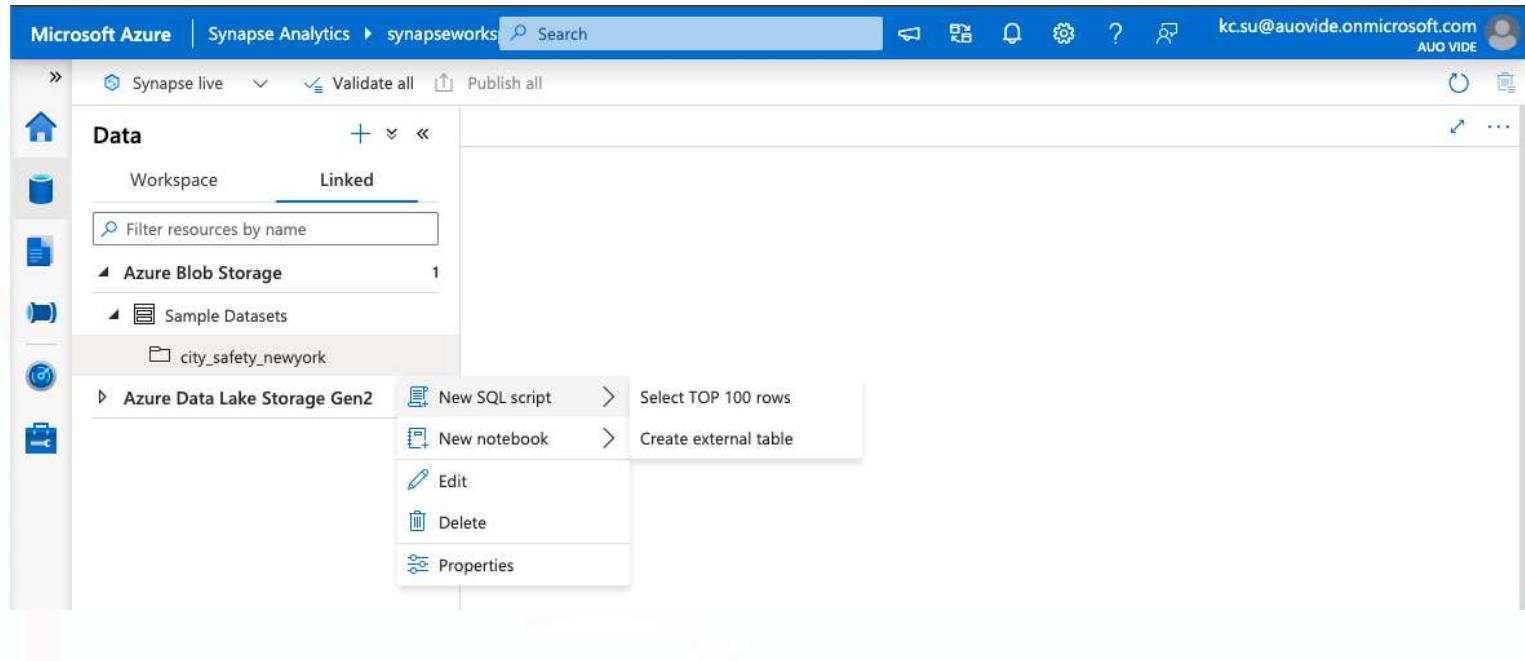
# (Dataset) Step 04

- `city_safety_newyork` 就會以 Parquet 格式出現在 Azure Blob Storage 底下的 Sample Datasets 裡面

The screenshot shows the Microsoft Azure Synapse Analytics Data workspace interface. The top navigation bar includes 'Microsoft Azure', 'Synapse Analytics', 'synapseworks', a search bar, and user information 'kc.su@auovide.onmicrosoft.com AUO VIDE'. The left sidebar has icons for Home, Workspace, Data, Metrics, Pipelines, and Jobs. The main area is titled 'Data' and shows two tabs: 'Workspace' and 'Linked'. Under 'Linked', there is a 'Filter resources by name' search bar. Below it, the 'Azure Blob Storage' section is expanded, showing 'Sample Datasets' which contains the dataset 'city\_safety\_newyork'. Another section, 'Azure Data Lake Storage Gen2', is partially visible.

# (Dataset) Step 05

- 按下旁邊的 ... 圖示，選取 New SQL script、Select TOP 100 rows



# (Dataset) Step 06

- Synapse Studio 就會產生 T-SQL 敘述，按下 Run 就可以進行 Query

The screenshot shows the Microsoft Azure Synapse Studio interface. The top navigation bar includes 'Microsoft Azure', 'Synapse Analytics', 'synapseworks', a search bar, and user information 'kc.su@auovide.onmicrosoft.com AUO VIDE'. The left sidebar has a 'Data' section with 'Workspace' selected, showing 'Azure Blob Storage' (1 item) and 'Sample Datasets' (1 item: 'city\_safety\_newyork'). The main area displays a 'SQL script 1' tab with the following T-SQL code:

```
1 -- This is auto-generated code
2 SELECT
3     TOP 100 *
4 FROM
5     OPENROWSET(
6         BULK      'https://azureopendatastorage.blob.core.windows.net/city_safety_newyork/parquet',
7         FORMAT   = 'parquet'
8     ) AS [result];
```

The 'Results' tab shows the query results in a table format:

dataType	dataSubtype	dateTime	category
Safety	311_All	2010-01-01T19:00:00.000Z	Noise - Residence
Safety	311_All	2017-09-20T10:00:00.000Z	Street Light Condition

A message at the bottom states '00:00:21 Query executed successfully.'

The right side features a 'Properties' panel with tabs for 'General' (selected) and 'Related (0)', showing the name 'SQL script 1' and a description field.

# (Dataset) Step 07

- 可以查詢前 100 筆 Data

```
SELECT TOP 100 * FROM OPENROWSET(
    BULK 'https://azureopendatastorage.blob.core.windows.net/.../city=NewYorkCity/*.parquet',
    FORMAT = 'parquet'
) AS [result]
```

- 可以查詢一共有多少筆 Data

```
SELECT COUNT(*) FROM OPENROWSET(
    BULK 'https://azureopendatastorage.blob.core.windows.net/.../city=NewYorkCity/*.parquet',
    FORMAT = 'parquet'
) AS [result]
```

- 可以查詢 Data 的日期範圍

```
SELECT MIN(dateTime) AS MinDate, MAX(dateTime) AS MaxDate FROM OPENROWSET(
    BULK 'https://azureopendatastorage.blob.core.windows.net/.../city=NewYorkCity/*.parquet',
    FORMAT = 'parquet'
) AS [result]
```

# (Dataset) Step 08

- 可以修改 Properties 的 Name，然後按下 Publish all 按鈕儲存

The screenshot shows the Microsoft Azure Synapse Analytics workspace interface. On the left, there's a sidebar with icons for Home, Data, Workspace, and Linked. The 'Data' section is selected, showing 'Workspace' and 'Linked' tabs, with 'Azure Blob Storage' selected under 'Linked'. A search bar at the top right contains 'synapseworks'. The main area displays a SQL script:

```
8 SELECT COUNT(*) FROM OPENROWSET(
9   BULK 'https://azureopendatastorage.blob.core.windows.net/cit
10  FORMAT = 'parquet'
11 ) AS [result]
12
13 SELECT MIN(dateTime) AS MinDate, MAX(dateTime) AS MaxDate FROM
14   BULK 'https://azureopendatastorage.blob.core.windows.net/cit
15  FORMAT = 'parquet'
16 ) AS [result]
17
```

The 'Properties' panel on the right shows the following details:

- Name:** SQL script - NYCSafetyData (highlighted with a red box)
- Description:** (empty)
- Type:** .sql script
- Size:** 249 bytes
- Results settings per query:**
  - First 5000 rows (default)
  - All rows

At the bottom, a message says "00:01:25 Query executed successfully."

# (Dataset) Step 09

- 再按一次 Publish 按鈕，就可以儲存 SQL Script 檔案

The screenshot shows the Microsoft Azure Synapse Analytics workspace interface. On the left, there's a sidebar with icons for Home, Data, Workspace, and Linked. The Data section is active, showing a list of resources: Azure Blob Storage (1 item), Sample Datasets (1 item), and Azure Data Lake Storage Gen2 (1 item). Below this is a search bar labeled "Filter resources by name". In the center, there's a large code editor window titled "SQL script - NYCSafe... X". It contains the following SQL script:

```
8 SELECT COUNT(*) FROM
9 BULK 'https://az...
10 FORMAT = 'parquet'
11 ) AS [result]
12
13 SELECT MIN(dateTime)
14 BULK 'https://az...
15 FORMAT = 'parquet'
16 ) AS [result]
17
```

Below the code editor, there are tabs for "Results" and "Messages". The "Results" tab shows a table with one row labeled "MinDate" containing the value "2010-01-01T00:00:00.000000". The "Messages" tab is empty. At the bottom of the code editor, a message says "00:01:25 Query executed successfully". To the right of the code editor, there's a "Publish all" dialog box. It says "You are about to publish all pending changes to the live environment. Learn more". Under "Pending changes (1)", it shows a table with three columns: NAME, CHANGE, and EXISTING. There is one entry: "SQL script" with a status of "-". At the bottom of the dialog are "Publish" and "Cancel" buttons.

# (Dataset) Step 10

- 可以按下右上角的 Properties 圖示收納 Properties 視窗

The screenshot shows the Microsoft Azure Synapse Analytics workspace interface. The left sidebar displays 'Data' resources, including 'Workspace' and 'Linked' datasets. Under 'Linked', there are 'Azure Blob Storage' (with 'city\_safety\_newyork' dataset) and 'Azure Data Lake Storage Gen2'. The main area is a 'SQL script - NYCSafe...' editor window. A red box highlights the 'Properties' icon in the top right corner of the editor toolbar. The SQL code retrieves data from a blob storage container using OPENROWSET and BULK statements. The results pane at the bottom shows the minimum and maximum dates from the query.

```
2 SELECT TOP 100 *
3 FROM OPENROWSET(
4      BULK 'https://azureopendatastorage.blob.core.windows.net/citydatacontainer/Safety/Release/city=NewYorkC',
5      FORMAT = 'parquet'
6 ) AS [result];
7
8 SELECT COUNT(*) FROM OPENROWSET(
9      BULK 'https://azureopendatastorage.blob.core.windows.net/citydatacontainer/Safety/Release/city=NewYorkC',
10     FORMAT = 'parquet'
11 ) AS [result];
12
13 SELECT MIN(dateTime) AS MinDate, MAX(dateTime) AS MaxDate FROM OPENROWSET(
14      BULK 'https://azureopendatastorage.blob.core.windows.net/citydatacontainer/Safety/Release/city=NewYorkC',
15      FORMAT = 'parquet'
16 ) AS [result]
17
```

MinDate	MaxDate
2010-01-01T00:00:00.0000000	2021-07-11T02:00:40.0000000

00:01:25 Query executed successfully.

# (Dataset) Step 11

- 到目前為止都是在 master Database 進行 Query，但是 master Database 不能定義 External Data Source
- 建立 NYCSafetyData Database，並且切換到 NYCSafetyData Database
- 定義 CityData External Data Source
- 查詢前 100 筆 Data
- 查一下 Emergency vs. Non-Emergency

```
CREATE DATABASE [NYCSafetyData] COLLATE Latin1_General_100_BIN2_UTF8

USE NYCSafetyData

CREATE EXTERNAL DATA SOURCE [CityData] WITH (
    LOCATION = 'https://azureopendatastorage.blob.core.windows.net/citydatacontainer'
)

SELECT TOP 100 *
FROM OPENROWSET(
    BULK '/Safety/Release/city=NewYorkCity/*.parquet',
    DATA_SOURCE = 'CityData',
    FORMAT = 'parquet'
) AS [result]

SELECT category, COUNT(*)
FROM OPENROWSET(
    BULK '/Safety/Release/city=NewYorkCity/*.parquet',
    DATA_SOURCE = 'CityData',
    FORMAT = 'parquet'
) AS [result]
WHERE category LIKE '%Emergency%'
GROUP BY category
```

# (Dataset) Step 11

- 注意 USE DATABASE

The screenshot shows the Microsoft Azure Synapse Analytics workspace interface. The top navigation bar includes 'Microsoft Azure', 'Synapse Analytics', 'synapseworks', a search bar, and user information 'kc.su@auovide.onmicrosoft.com AUO VIDE'. The left sidebar has icons for Home, Databases, Tables, Functions, Pipelines, and Jobs. The main area shows a SQL script named 'SQL script - NYCSafe... X' with line numbers 25 to 37. The script uses OPENROWSET and BULK commands to query data from 'NYCSafetyData' database. A red box highlights the 'Use database' dropdown set to 'NYCSafetyData'. Below the script, the 'Results' tab is selected, showing a table with columns: dataType, dataSubtype, dateTime, category, subcategory, status, address, latitude, and longitude. Two rows of data are visible. At the bottom, a message indicates '00:00:12 Query executed successfully.'

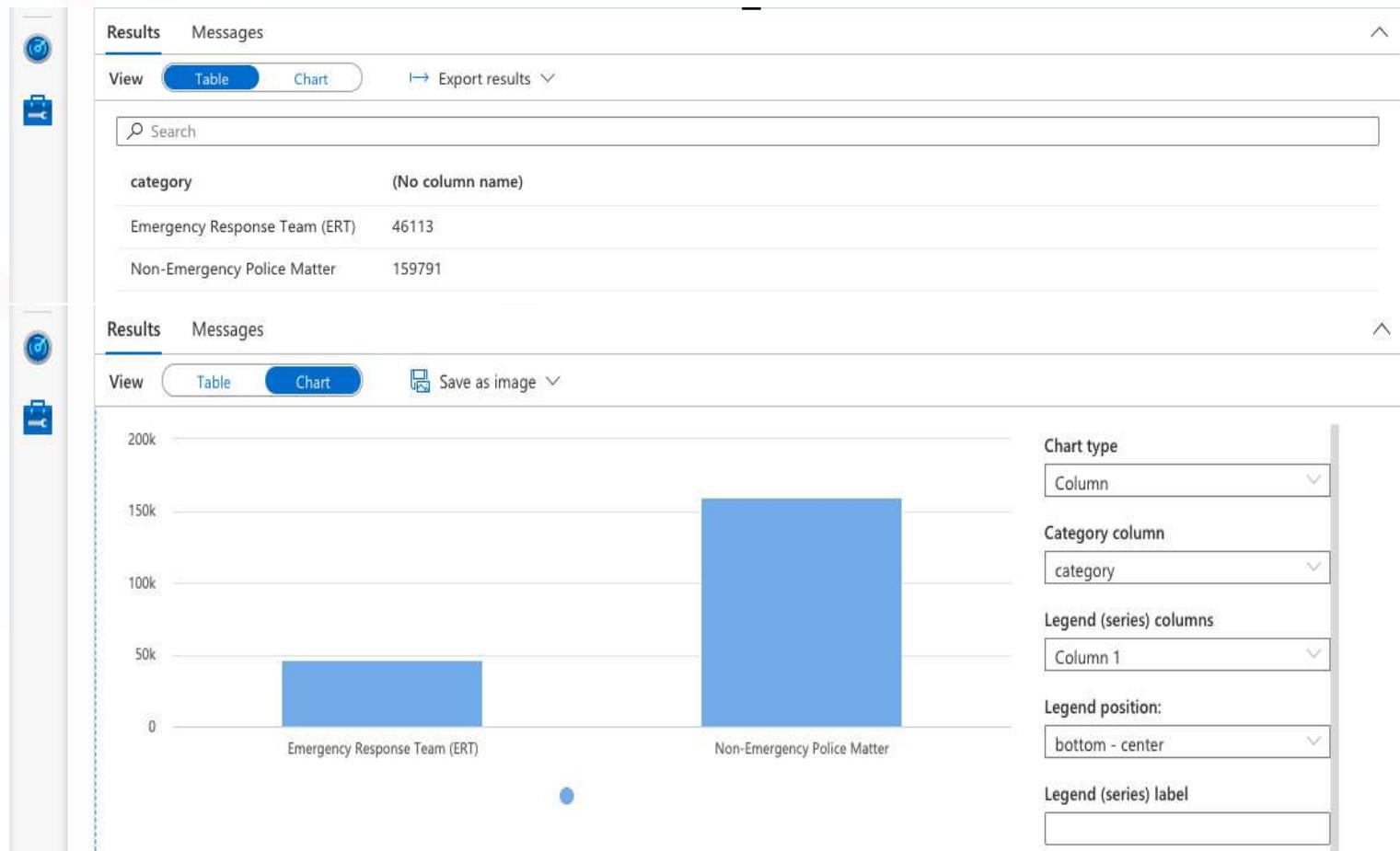
```
25
26
27 SELECT TOP 100 *
28 FROM OPENROWSET(
29   BULK '/Safety/Release/city=NewYorkCity/*.parquet',
30   DATA_SOURCE = 'CityData',
31   FORMAT = 'parquet'
32 ) AS [result]
33
34 SELECT category, COUNT(*)
35 FROM OPENROWSET(
36   BULK '/Safety/Release/city=NewYorkCity/*.parquet',
37   DATA_SOURCE = 'CityData',
```

dataType	dataSubtype	dateTime	category	subcategory	status	address	latitude	longitude
Safety	311_All	2010-01-01T19...	Noise - Residen...	Loud Music/Par...	Closed	335 SUTTER AV...	40.6675789109...	-73.907261891
Safety	311_All	2017-09-20T10...	Street Light Co...	Street Light Cyc...	Closed	(NULL)	40.7700963150...	-73.790515163

00:00:12 Query executed successfully.

# (Dataset) Step 11

- View 可以點選 Chart，以圖表呈現



# (Dataset) Step 12 (Version Check)

```
SELECT @@VERSION
```

```
Microsoft SQL Azure (RTM) - 12.0.2000.8 Jun 10 2021 06:17:24 Copyright (C) 2019 Microsoft Corporation
```

```
SELECT SERVERPROPERTY('ProductVersion')
```

```
12.0.2000.8
```

```
SELECT SERVERPROPERTY('EngineEdition')
```

```
11
```

# 建立 Apache Spark Pool

# (Spark) Step 01

- 按下 Synapse Studio 的 Manage 標籤頁
- 選取 Analytics pools 底下的 Apache Spark pools
- 按下 + New 新增 Apache Spark Pool

The screenshot shows the Microsoft Azure Synapse Studio interface. The left sidebar has a 'Manage' section selected. The main area shows the 'Apache Spark pool' section, which is currently empty. A large 'Create Apache Spark pool' button is at the bottom.

Microsoft Azure | Synapse Analytics > synapseworkspace00

Synapse live | Validate all | Publish all

kc.su@auovide.onmicrosoft.com AUO VIDE

Home | Data | Develop | Integrate | Monitor | Manage

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 | Git configuration

Apache Spark pool

Apache Spark pools can be tuned to run different kinds of Apache Spark workloads using specific configuration libraries, permissions, etc. [Learn more](#)

+ New | Refresh

Filter by name

Showing 0-0 of 0 item

Name	Node size family	Size
------	------------------	------

No items to show

Try changing your filter or create new Apache Spark pool

Create Apache Spark pool

# (Spark) Step 02 (Basics)

- Apache Spark pool name 輸入 sparkpoolm00
- Node size 可以是 Small 到 XXXLarge，選取 Small (4 vCores / 32 GB) 即可
- Autoscale 選取 Enabled
- Number of nodes 上下限可以是 3 到 200，設定 3 到 5 即可

Microsoft Azure | Synapse Analytics > synapseworkspace00

Home Data Develop Integrate Monitor Manage

### Create Apache Spark pool

Basics • Additional settings \* Tags Review + create

Create an Synapse Analytics Apache Spark pool with your preferred configurations. Complete the Basics tab then go to Review + Create to provision with smart defaults, or visit each tab to customize.

**Apache Spark pool details**

Name your Apache Spark pool and choose its initial settings.

Apache Spark pool name \*

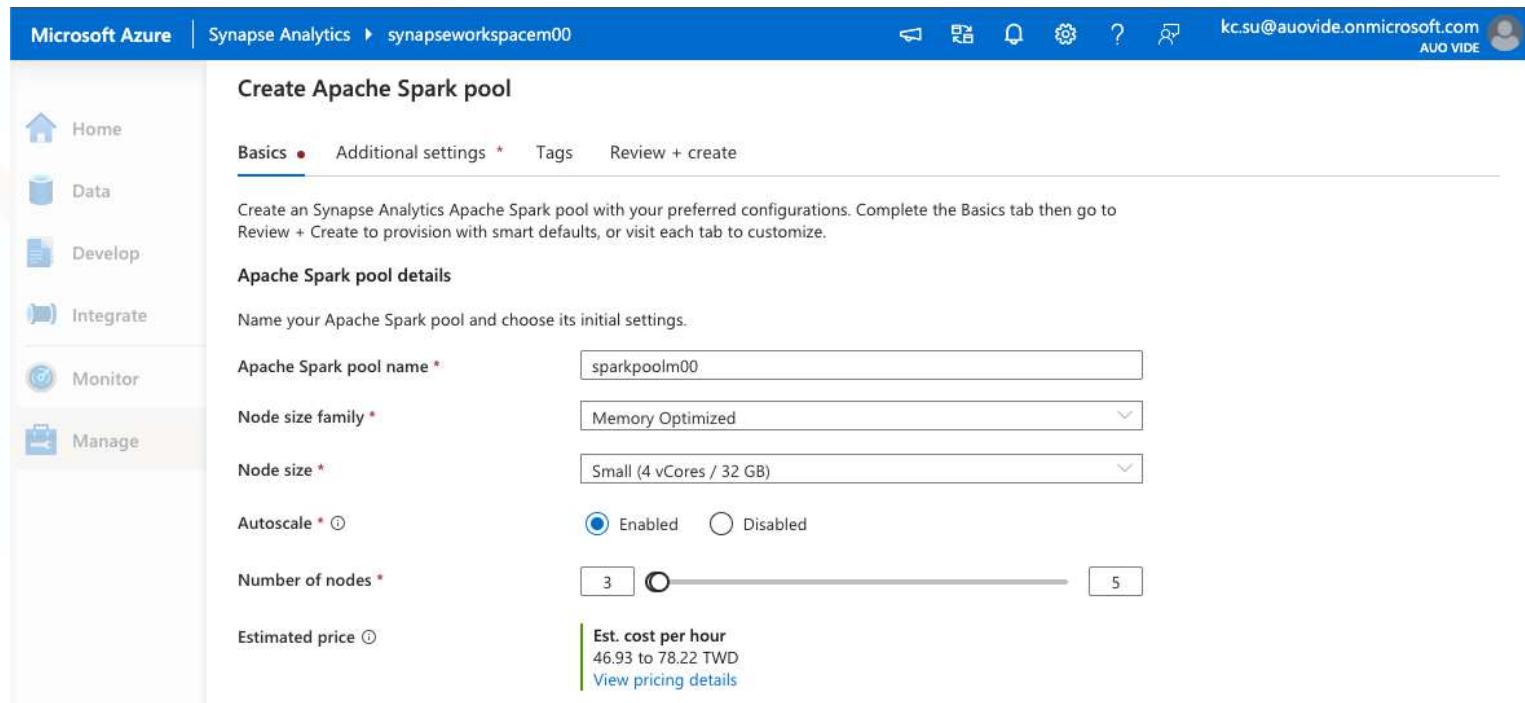
Node size family \*

Node size \*

Autoscale \*  Enabled  Disabled

Number of nodes \*

Estimated price ⓘ   
 Est. cost per hour  
 46.93 to 78.22 TWD  
[View pricing details](#)



# (Spark) Step 03 (Additional settings)

- Automatic pausing 選取 Enabled
- Number of minutes idle 預設是 15 分鐘
- Apache Spark 版本預設是 2.4，3.0 還在 Preview

Microsoft Azure | Synapse Analytics > synapseworkspace00

Create Apache Spark pool

Basics \* Additional settings \* Tags Review + create

Customize additional parameters including pause settings and component versions.

**Automatic pausing**

Configure the pause settings for the Apache Spark pool.

Automatic pausing \*  Enabled  Disabled

Number of minutes idle \*

**Component version**

Select the Spark version for your Apache Spark pool.

Apache Spark *	
Python	3.6
Scala	2.11.12
Java	1.8.0_272

# (Spark) Step 04 (Additional settings)

- 按下 Review + create 按鈕

The screenshot shows the 'Create Apache Spark pool' configuration page in the Microsoft Azure Synapse Analytics portal. The left sidebar includes links for Home, Data, Develop, Integrate, Monitor, and Manage. The main content area displays the following settings:

- Create Apache Spark pool**:
  - Scala: 2.11.12
  - Java: 1.8.0\_272
  - .NET Core: 3.1
  - .NET for Apache Spark: 1.0
  - Delta Lake: 0.6
- Apache Spark configuration**:

Upload a Spark configuration file to specify additional properties on the Apache Spark pool. This will be referenced to configure Spark applications upon job submission.

File Upload:

Upload
- Packages**:

Configure settings related to packages and how they can be installed onto your Spark pool.

Allow session level packages \*  Enabled  Disabled

At the bottom are navigation buttons: **Review + create**, **< Previous**, **Next: Tags >**, and **Cancel**.

# (Spark) Step 05

- 按下 Create 按鈕，建立 Apache Spark Pool

The screenshot shows the 'Create Apache Spark pool' review step in the Microsoft Azure portal. The left sidebar includes links for Home, Data, Develop, Integrate, Monitor, and Manage. The main content area has tabs for Basics, Additional settings, Tags, and Review + create (which is underlined). A green success message box says 'Validation succeeded.' Below it, the 'Product details' section shows 'Azure Synapse Analytics Apache Spark pool by Microsoft' with an 'Est. cost per hour' of '46.93 to 78.22 TWD' and a 'View pricing details' link. The 'Terms' section contains a detailed legal agreement text and a 'Azure Marketplace Terms' link. At the bottom, there are buttons for 'Create' (in blue), '< Previous', 'Download template for automation', and 'Cancel'.

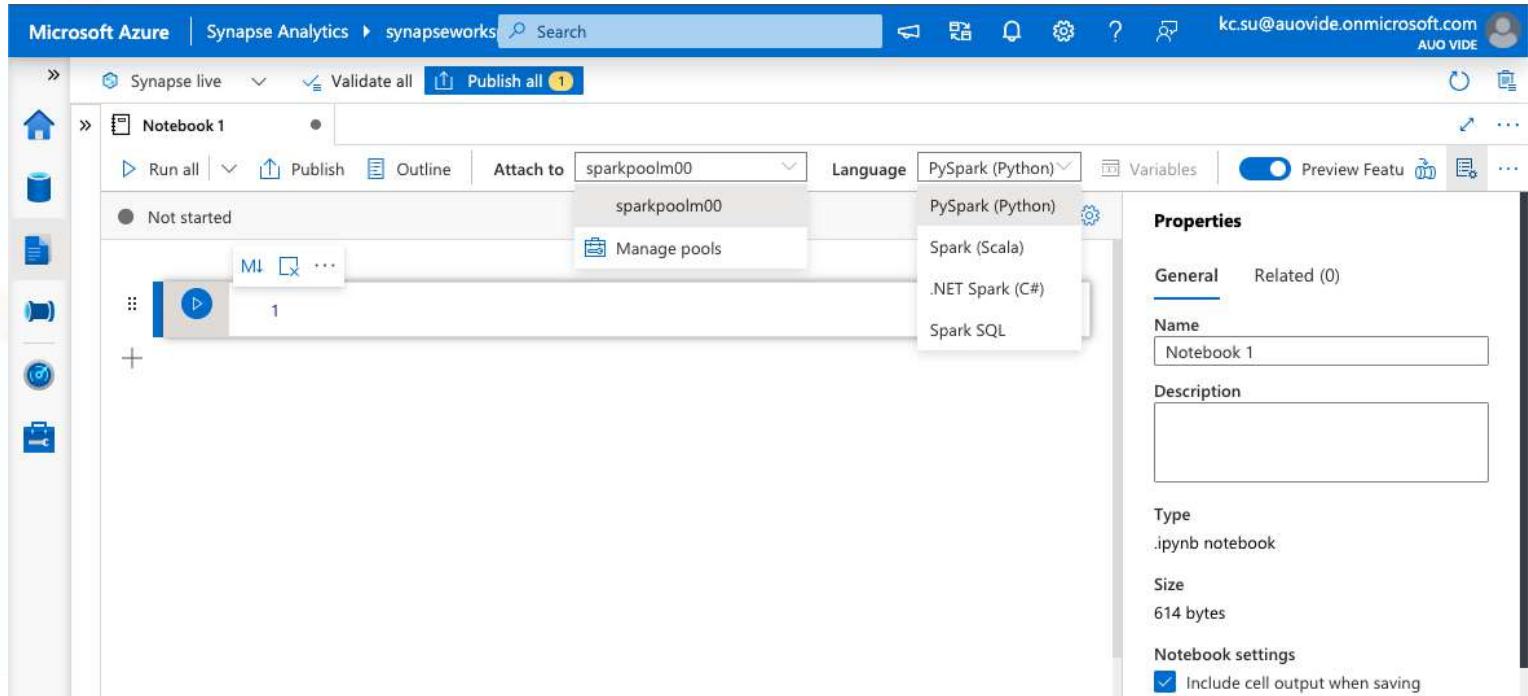
# (Spark) Step 06

- 按下 Synapse Studio 的 Develop 標籤頁
- 雖然 Apache Spark 也支援 SQL，但是我們這次新增 Notebook

The screenshot shows the Microsoft Azure Synapse Studio interface. The top navigation bar includes 'Microsoft Azure', 'Synapse Analytics', 'synapseworks', a search bar, and various icons for notifications and settings. The left sidebar has navigation links: Home, Data, Develop (which is selected and highlighted in blue), Integrate, Monitor, and Manage. The main content area is titled 'Develop' and contains a sub-menu with options: 'SQL script', 'Notebook', 'Data flow', 'Apache Spark job definition', 'Browse gallery', and 'Import'. A sub-item 'SQL script - NYCSafetyData' is also visible under 'SQL scripts'. To the right of the menu, there are two cylindrical icons with code snippets and a callout text 'Select an item' followed by the instruction 'Use the resource explorer to select or create a new item'. The overall theme is light blue and white.

# (Spark) Step 07

- Attach to 選取剛剛建立的 Apache Spark Pool , Language 選取 PySpark



Spark Notebook Primary Language :

- 可以在右上角的 Dropdown List 直接選取 Python / Scala / C# / SparkSQL
- 也可以在每個 Cell 手動加上 %%pyspark / %%spark / %%csharp / %%sql

# (Spark) Step 08

- 使用 Azure Open Datasets 裡面的 New York City Safety Data

```
from azureml.opendatasets import NycSafety
from datetime import datetime
from dateutil import parser

startDate = parser.parse("2021-01-01")
endDate = parser.parse("2021-01-31")
nycData = NycSafety(start_date=startDate, end_date=endDate)
dfData = nycData.to_spark_dataframe()
```

The screenshot shows the Microsoft Azure Synapse Analytics notebook interface. The top navigation bar includes 'Microsoft Azure', 'Synapse Analytics', 'synapseworks', 'Search', and user information 'kc.su@auovide.onmicrosoft.com' with 'AUO VIDE' status. The main area displays a notebook titled 'Notebook - NYCSafe...'. The code cell contains the Python script provided above. Below the code cell, a message indicates 'Command executed in 28 sec 844 ms by kc.su on 11:16:47 PM, 7/15/21'. A 'Job execution' section shows a successful run with 'Spark 2 executors 8 cores'. The table details the execution of two tasks:

ID	Description	Status	Stages	Tasks	Submission Time	Duration
Job 0	parquet at NativeMethodAccessImpl.java:0	Succeeded	1/1	1/1 succeeded	11:16:39 PM, 7/15/21	6 sec
Stage 0	parquet at NativeMethodAccessImpl.java:0	Succeeded	-	1/1 succeeded	11:16:39 PM, 7/15/21	6 sec

# (Spark) Step 09

- 顯示前 100 筆 Data

```
display(dfData.limit(100))
```

The screenshot shows the Microsoft Azure Synapse Analytics interface. The top navigation bar includes 'Microsoft Azure', 'Synapse Analytics', 'synapseworks', a search bar, and user information 'kc.su@auovide.onmicrosoft.com AUO VIDE'. The left sidebar has icons for Home, Databricks, Notebooks, Pipelines, and Data Flow. The main workspace shows a notebook titled 'Notebook - NYCSafe...'. The code cell contains the command: `display(dfData.limit(100))`. Below the code, it says 'Command executed in 41 sec 246 ms by kc.su on 11:19:26 PM, 7/15/21'. The execution status is 'Job execution Succeeded' with 'Spark 2 executors 8 cores'. The results are displayed in a table view, with 'Table' selected. The table has columns: dataType, dataSubtype, dateTime, category, and subcatego. The data shows various safety incidents across different dates and categories.

dataType	dataSubtype	dateTime	category	subcatego
Safety	311_All	2021-01-28T15:51:09.000Z	Elevator	Elevator -
Safety	311_All	2021-01-24T09:43:36.000Z	Illegal Parking	Blocked H
Safety	311_All	2021-01-16T15:25:25.000Z	Illegal Parking	Blocked H
Safety	311_All	2021-01-24T22:54:11.000Z	Noise - Residential	Loud Musi
Safety	311_All	2021-01-07T12:13:16.000Z	Noise - Residential	Loud Musi
Safety	311_All	2021-01-26T00:03:38.000Z	HEAT/HOT WATER	ENTIRE BU

# (Spark) Step 10

- 統計所有 Data 筆數

```
dfData.count()
```

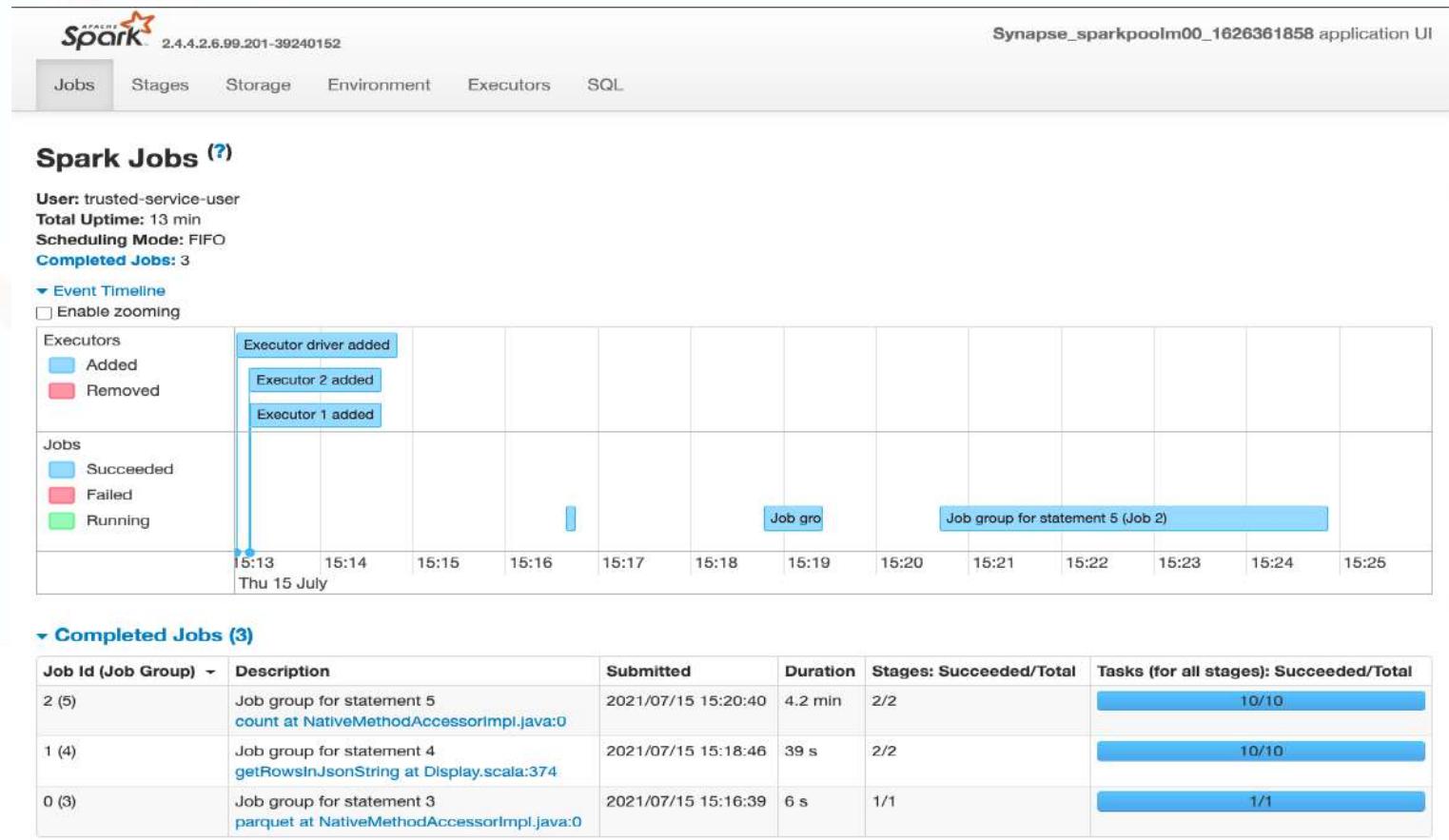
The screenshot shows the Microsoft Azure Synapse Analytics notebook interface. The top navigation bar includes 'Microsoft Azure', 'Synapse Analytics', 'synapseworks', a search bar, and user information 'kc.su@auovide.onmicrosoft.com AUO VIDE'. The main area displays a notebook titled 'Notebook - NYCSafe...'. The code cell contains '1 dfData.count()'. Below the code, a message states 'Command executed in 4 min 13 sec 434 ms by kc.su on 11:24:53 PM, 7/15/21'. A detailed view of the 'Job execution' shows it succeeded with Spark 2 executors and 8 cores. The table below lists the stages and tasks:

ID	Description	Status	Stages	Tasks	Submission Time	Duration
Job 2	count at NativeMethodAccessorImpl.java:0	Succeeded	2/2	10/10 succeeded	11:20:40 PM, 7/15/21	4 min 12 sec
Stage 3	count at NativeMethodAccessorImpl.java:0	Succeeded	-	9/9 succeeded	11:20:40 PM, 7/15/21	4 min 12 sec
Stage 4	count at NativeMethodAccessorImpl.java:0	Succeeded	-	1/1 succeeded	11:24:52 PM, 7/15/21	0 sec

At the bottom of the execution details, the number '193113' is displayed.

# (Spark) Step 11

- 可以透過 Apache Spark UI 檢視執行狀況



# 建立 Dedicated SQL Pool

# (SQL) Step 01

- 按下 Synapse Studio 的 Manage 標籤頁
- 選取 Analytics pools 底下的 SQL pools
- 按下 + New 新增 SQL Pool

The screenshot shows the Microsoft Azure Synapse Studio interface. The left sidebar has 'Manage' selected. The main area shows the 'Analytics pools' section with 'SQL pools' highlighted. A table below lists one item:

Name	Type	Status	Size
Built-in	Serverless	Online	Auto

# (SQL) Step 02 (Basics)

- Dedicated SQL pool name 設定為 sqlpoolm00
- Geo-redundant 視情況需要設定 (選取 No 即可)
- Performance level 預設是 DW1000c (大概 15 USD/Hour)，調整為 DW100c 即可

Microsoft Azure | Synapse Analytics > synapseworkspace00

Create dedicated SQL pool

Basics \* Additional settings \* Tags Review + create

Create a dedicated SQL pool with your preferred configurations. Complete the Basics tab then go to Review + Create to provision with smart defaults. [Learn more](#)

Dedicated SQL pool details

Name your dedicated SQL pool and choose its initial settings.

Dedicated SQL pool name \*

Geo-redundant \*  Yes  No

⚠ Your Azure Synapse Analytics SQL pool data and backups will not be replicated to a paired region for protection against outages. This option ensures your data does not leave your country's geographic boundary. If you'd like to enable replication to your paired region for data protection, select "Yes".

Performance level  DW100c

Estimated price  46.24 TWD [View pricing details](#)

# (SQL) Step 03 (Additional settings))

- 視情況設定 Use existing data，全新的情況就選取 None
- Collation 預設是 SQL\_Latin1\_General\_CI\_AS
- 按下 Review + create 按鈕

The screenshot shows the 'Create dedicated SQL pool' wizard in the Microsoft Azure portal. The left sidebar includes links for Home, Data, Develop, Integrate, Monitor, and Manage. The main area has tabs for Basics, Additional settings (which is selected), Tags, and Review + create. Under 'Additional settings', there's a note about customizing configuration parameters including collation. The 'Data source' section allows starting with a blank pool or restoring from a backup. The 'Use existing data' dropdown is set to 'None'. The 'SQL pool collation' section notes that collation rules cannot be changed after creation and lists the default as 'SQL\_Latin1\_General\_CI\_AS'. A 'Collation' input field contains 'SQL\_Latin1\_General\_CI\_AS'. At the bottom, there are buttons for 'Review + create', '< Previous', 'Next: Tags >', and 'Cancel'.

# (SQL) Step 04

- 按下 Create 按鈕，建立 Dedicated SQL Pool

The screenshot shows two screenshots of the Microsoft Azure Synapse Analytics portal.

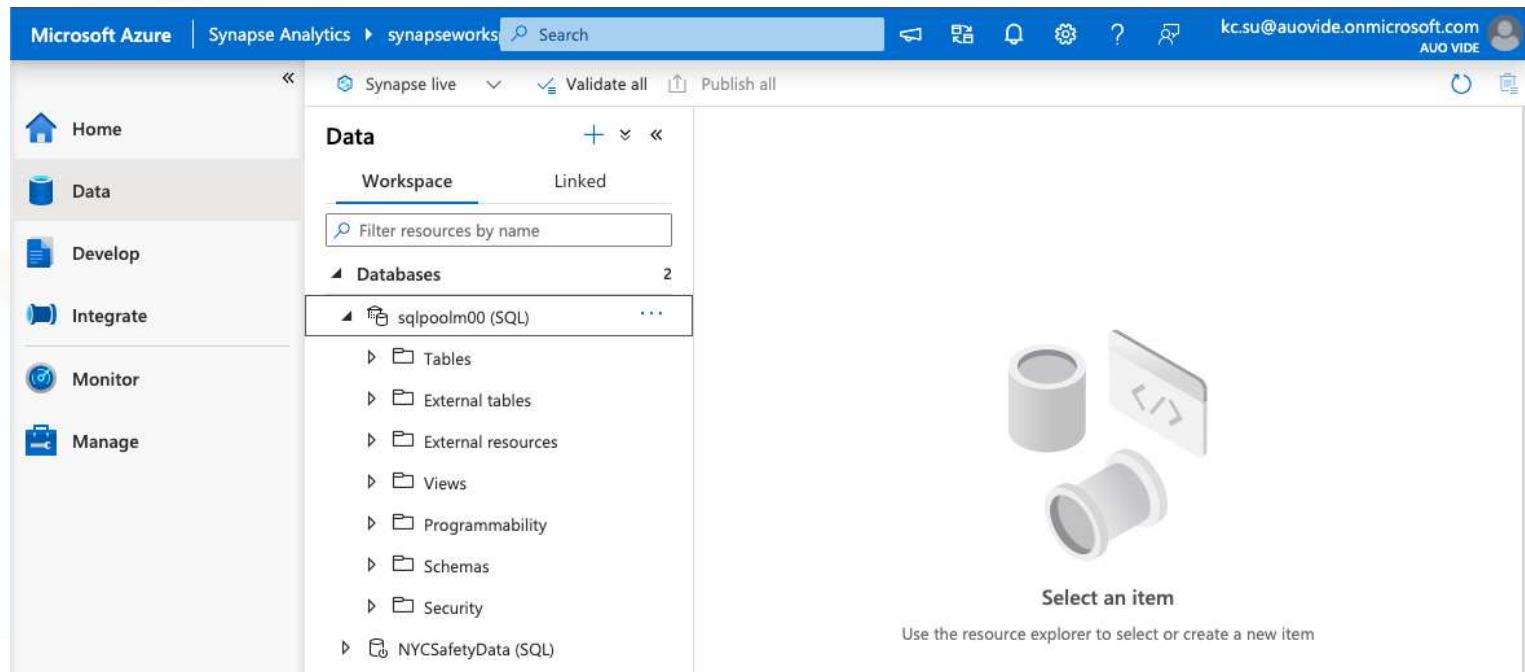
**Top Screenshot:** A 'Create dedicated SQL pool' dialog box. It displays a green validation message 'Validation succeeded.' and tabs for 'Basics', 'Additional settings', 'Tags', and 'Review + create'. Under 'Product details', it shows 'Azure Synapse Analytics dedicated SQL pool by Microsoft' with an 'Est. cost per hour' of '46.24 TWD' and links to 'Terms of use' and 'Privacy policy'. Under 'Terms', there is a detailed legal agreement. At the bottom are 'Create', '< Previous', 'Download template for automation', and 'Cancel' buttons.

**Bottom Screenshot:** A list of SQL pools in the 'Synapse live' workspace. The sidebar shows 'Analytics pools' selected, with 'SQL pools' highlighted. The main area shows a table of SQL pools:

Name	Type	Status	Size
Built-in	Serverless	Online	Auto
sqlpoolm00	Dedicated	Online	DW100c

# (SQL) Step 05

- 按下 Synapse Studio 的 Data 標籤頁
- 按下 Databases 旁邊的 ... 圖示，再按下 Refresh，才會看到剛剛的 SQL Pool



# 透過 Power BI 展現資料

# (PowerBI) Step - Ingest Data 01

- 在 Synapse Studio 按下 Ingest 區塊

The screenshot shows the Microsoft Azure Synapse Analytics workspace interface. The left sidebar contains navigation links: Home, Data, Develop, Integrate, Monitor, and Manage. The main area is titled "Synapse Analytics workspace" and "synapseworkspacem00". It features three prominent buttons: "Ingest" (blue icon), "Explore and analyze" (purple icon), and "Visualize" (yellow icon). A large background graphic displays a network graph with nodes connected by lines and a bar chart overlaid. Below the buttons, there's a "Discover more" section with "Knowledge center" and "Browse partners" links.

# (PowerBI) Step - Ingest Data 02

- Type task 預設選取 Built-in copy task
- Task cadence or task schedule 預設值為 Run once now，也可以視情況設定為 Schedule 或 Tumbling window

Microsoft Azure | Synapse Analytics > synapseworkspace00

Copy Data tool

Properties

Source

Target

Settings

Review and finish

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.

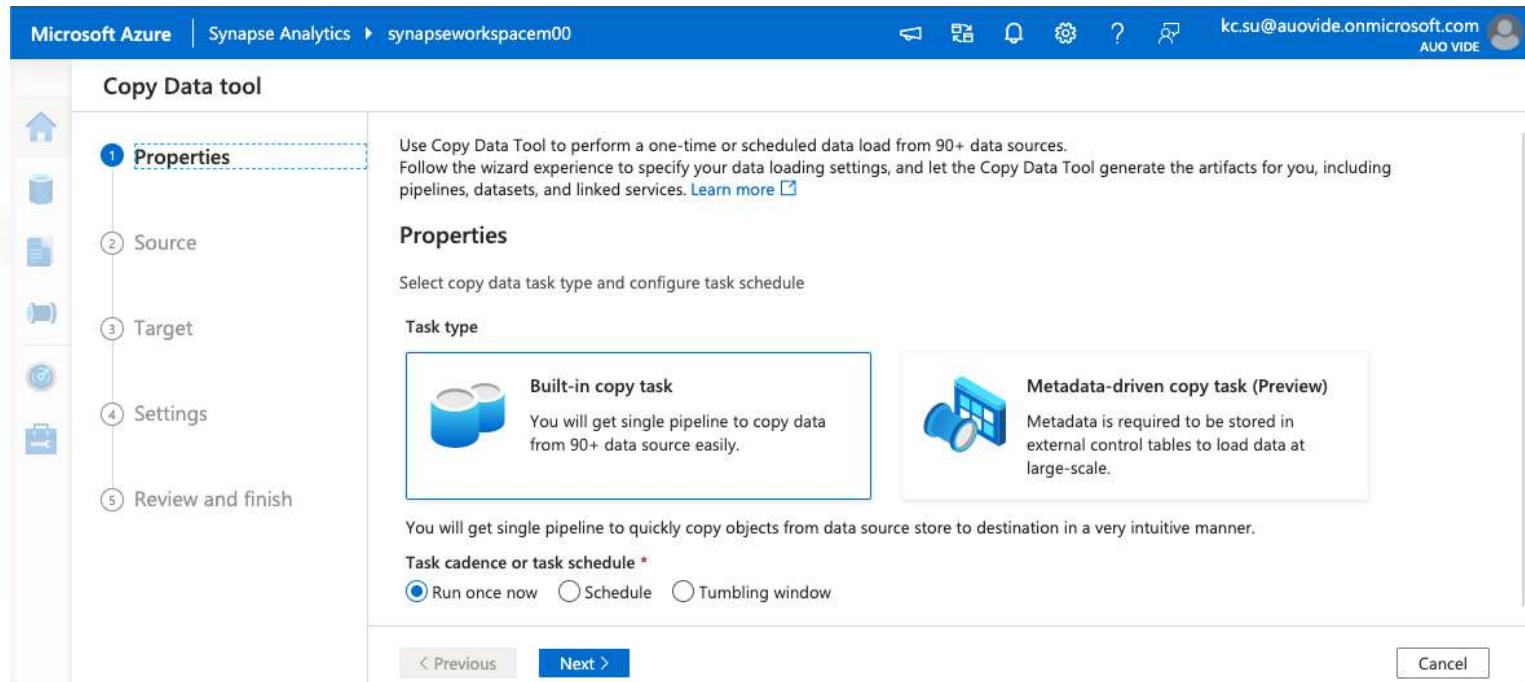
**Metadata-driven copy task (Preview)**  
Metadata is required to be stored in external control tables to load data at large-scale.

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



# (PowerBI) Step - Ingest Data 03

- 接下來其實就是 Copy Data Pipeline
- Source type 選取 SQL server，然後按下 Connection 右邊的 + Create new connection 建立新的 Connection

The screenshot shows the 'Copy Data tool' interface in Microsoft Azure. The top navigation bar indicates the user is in the 'Synapse Analytics' section of 'synapseworkspace00'. The main area is titled 'Copy Data tool' and displays a step-by-step wizard:

- Properties (selected)
- Source (Step 2)
- Dataset
- Configuration
- Target
- Settings
- Review and finish

The 'Source' step is currently being configured. The 'Source data store' section contains the following details:

- Source type: SQL server (selected)
- Connection \*: A dropdown menu is open, showing the following options:
  - Filter...
  - SAP BW via MDX
  - SAP HANA
  - SAP Table
  - SQL server (selected)
  - Spark
  - Sybase
  - Teradata
- Create new connection: A button to the right of the dropdown.

At the bottom of the screen, there are navigation buttons: '< Previous' (disabled), 'Next >', and 'Cancel'.

# (PowerBI) Step - Ingest Data 04

- Name 設定為 ContosoBIDW
- Connect via integration runtime 選取預設的 AutoResolveIntegrationRuntime
- 相關資訊選取透過 Connection string 取得
- Server name 輸入 SQL Server 的 IP Address 23.101.30.33
- Database name 輸入 ContosoRetailDW
- Authentication type 選取 SQL authentication
- User name 輸入 SQL Server 的 Admin 帳號 sqladminuser，然後輸入 Password
- 按下 Test connection 進行測試
- 測試成功就可以按下 Create 按鈕

# (PowerBI) Step - Ingest Data 04

Microsoft Azure | Synapse Analytics > synapseworkspace00

New connection (SQL server)

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

Name \* ContosoBIDW

Description

Connect via integration runtime \* AutoResolveIntegrationRuntime

Connection string Azure Key Vault

Server name \* 23.101.30.33

Database name \* ContosoRetailDW

Authentication type SQL authentication

User name \* sqladminuser

Password Azure Key Vault

\*\*\*\*\*

Connection successful

Test connection Cancel

Copy Data tool

Properties

Source

Dataset

Configuration

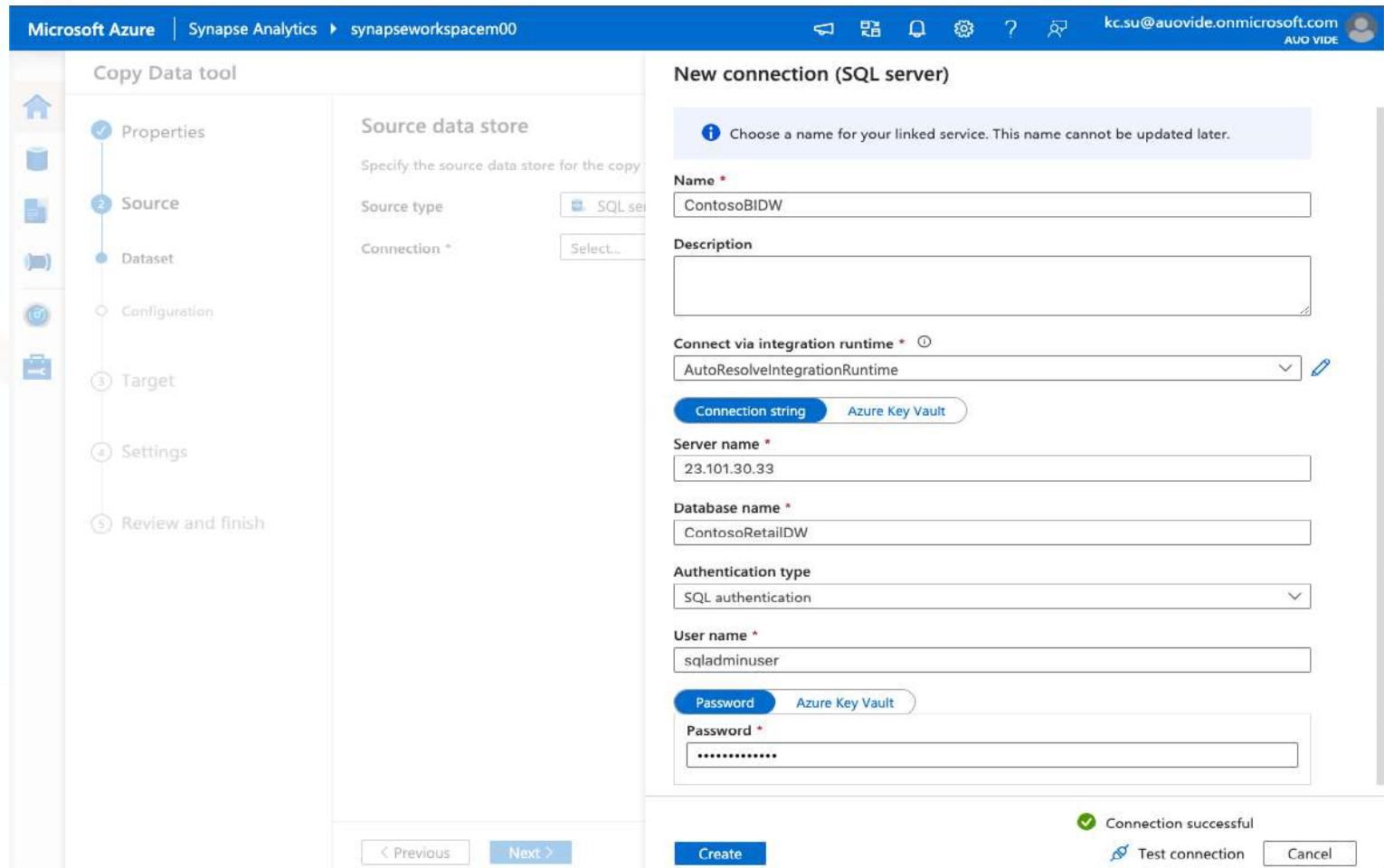
Target

Settings

Review and finish

< Previous Next >

Create



# (PowerBI) Step - Ingest Data 05

- Source tables 預設是 EXISTING TABLES，視情況選取 USE QUERY
- 勾選 DimChannel / DimProduct / DimPromotion / DimStore / FactSales 五個 Table

The screenshot shows the 'Copy Data tool' interface in Microsoft Azure Synapse Analytics. The navigation bar at the top indicates the user is in the 'synapseworkspacem00' workspace. The main area is titled 'Copy Data tool' and shows the 'Source' step selected in the left sidebar. The 'Source data store' section is configured with 'Source type' set to 'SQL server', 'Connection' set to 'ContosoBIDW', and 'Integration runtime' set to 'AutoResolveIntegrationRuntime'. The 'Source tables' section has 'EXISTING TABLES' selected, and it lists 26 tables. Two tables are checked: 'dbo.DimStore' and 'dbo.FactSales'. Other tables listed include 'dbo.FactExchangeRate', 'dbo.FactInventory', 'dbo.FactITMachine', 'dbo.FactITSAL', and 'dbo.FactOnlineSales'. At the bottom of the screen, there are buttons for '< Previous', 'Next >', and 'Cancel'.

# (PowerBI) Step - Ingest Data 06

- 視情況需要個別設定 Isolation level 與 Partition option，也可以 Preview data

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 area is titled 'Apply filter' and shows a list of tables: dbo.DimChannel, dbo.DimProduct, dbo.DimPromotion, dbo.DimStore, and dbo.FactSales. The 'dbo.FactSales' table is highlighted with a gray background. On the right, there are advanced settings for Query timeout (minutes), Isolation level (set to 'None'), and Partition option (set to 'None'). A note at the bottom says: 'Please preview data to validate the partition settings are correct before you trigger a run or publish the pipeline.' At the bottom are 'Previous' and 'Next' buttons, and a 'Cancel' button.

# (PowerBI) Step - Ingest Data 07

- Target type 選取 All
- Connection 選取剛剛建立的 Dedicated SQL Pool

Microsoft Azure | Synapse Analytics > synapseworkspace00

Copy Data tool

Properties

Source

Target

Dataset

Configuration

Settings

Review and finish

Destination data store

Specify the destination data store for the copy task. You can use an existing data store connection or specify a new data store.

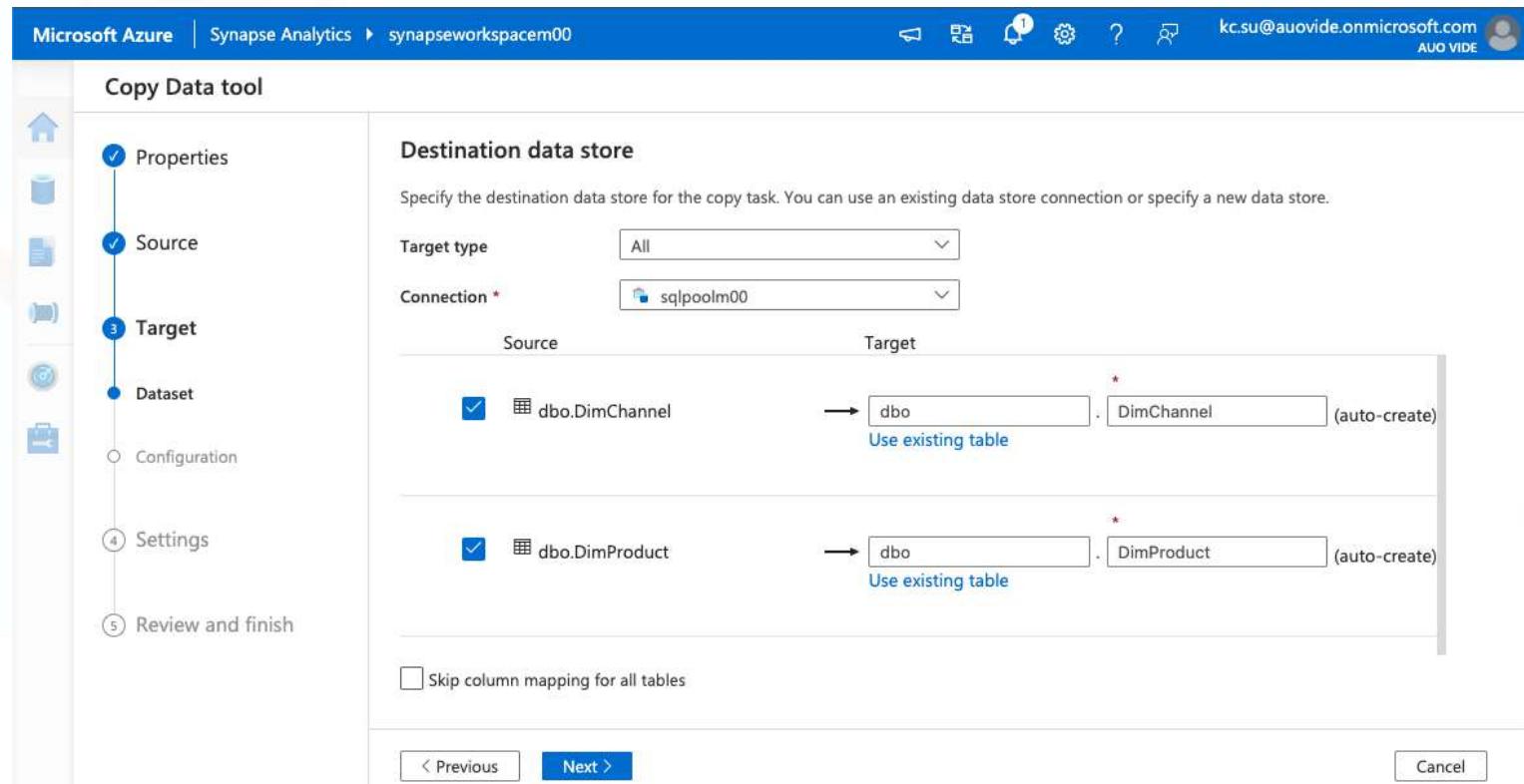
Target type: All

Connection \*: sqlpoolm00

Source	Target
<input checked="" type="checkbox"/> dbo.DimChannel	→ <input type="text"/> dbo . <input type="text"/> DimChannel (auto-create) <a href="#">Use existing table</a>
<input checked="" type="checkbox"/> dbo.DimProduct	→ <input type="text"/> dbo . <input type="text"/> DimProduct (auto-create) <a href="#">Use existing table</a>

Skip column mapping for all tables

< Previous Next > Cancel



# (PowerBI) Step - Ingest Data 08

- Column mapping 視情況調整

The screenshot shows the Microsoft Azure Copy Data tool interface. The left sidebar lists steps: Properties (selected), Source, Target (highlighted in blue), Dataset, Configuration, Settings, and Review and finish. The main area is titled "Column mapping" and shows "Table mappings (5)". It lists five mappings with "Source" and "Target" columns. The first mapping is highlighted with a red border:

Source	Target
dbo.DimChannel	dbo.DimChannel

Below this, there are four more mappings:

Source	Target
dbo.DimProduct	dbo.DimProduct
dbo.DimPromotion	dbo.DimPromotion
dbo.DimStore	dbo.DimStore

To the right, under "Column mappings", there is a table with one row:

Source	Type
ChannelKey	int

Below the table, it says "Azure Synapse dedicated SQL pool sink properties". There is a "Pre-copy script" section with a text input field and an "Advanced" section.

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

# (PowerBI) Step - Ingest Data 09

- Task name 設定為 CopyPipeline\_ContosoBIDW
- 展開 Advanced，視情況需要設定 PolyBase 的 Reject type / Reject value 內容

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

- Task name \***: CopyPipeline\_ContosoBIDW
- Task description**: (empty)
- Fault tolerance**: (dropdown menu)
- Enable logging**: (checkbox)
- Enable staging**: (checkbox)
- Advanced** section:
  - Allow PolyBase**: (checkbox, checked)
  - Reject type**: (dropdown menu)
    - Value (selected)
    - Value
    - Percentage
  - Reject value**: (text input field)
  - Use type default**: (checkbox)
- Data integration unit**: (dropdown menu)
  - Auto (selected)
  - Edit
- Note**: You will be charged # of used DIUs \* copy duration \* \$0.25/DIU-hour. Local currency and separate discounting may apply per subscription type. [Learn more](#)
- Degree of copy parallelism**: (text input field)
- Edit**: (checkbox)

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

# (PowerBI) Step - Ingest Data 10

- Summary 檢視所有設定

Microsoft Azure | Synapse Analytics > synapseworkspace00

Copy Data tool

Properties ✓  
Source ✓  
Target ✓  
Settings ✓  
**Review and finish**  
Review  
Deployment

**Summary**  
You are running pipeline to copy data from SQL server to Azure Synapse dedicated SQL pool.

SQL server → Azure Synapse dedicated SQL pool

**Properties**

Task name: CopyPipeline\_ContosoBIDW  
Task description:

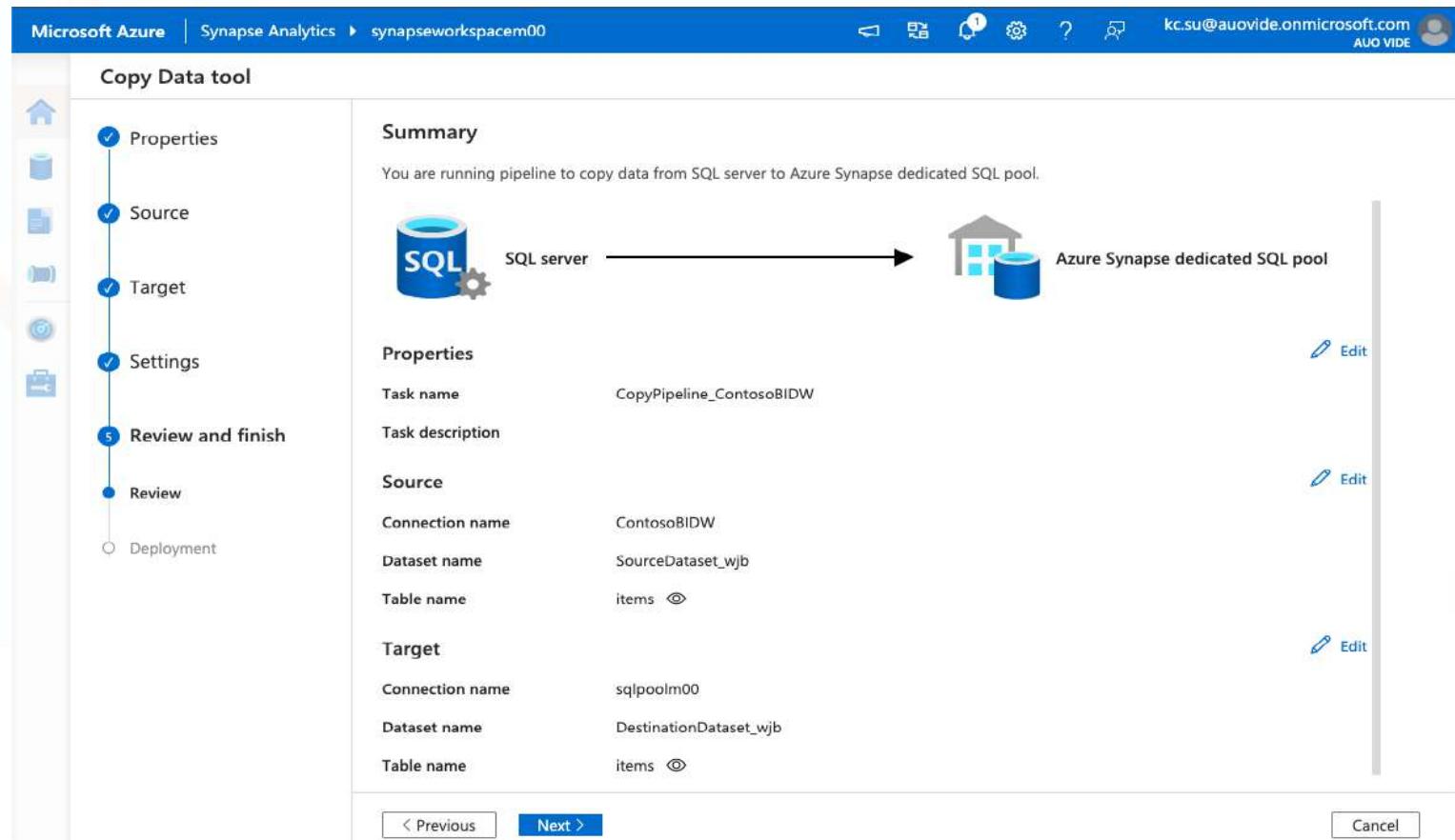
**Source**

Connection name: ContosoBIDW  
Dataset name: SourceDataset\_wjb  
Table name: items

**Target**

Connection name: sqlpoolm00  
Dataset name: DestinationDataset\_wjb  
Table name: items

< Previous Next > Cancel



# (PowerBI) Step - Ingest Data 11

- Pipeline 建立佈署完畢，按下 Finish 按鈕結束

Microsoft Azure | Synapse Analytics > synapseworkspacem00

Copy Data tool

Properties  
Source  
Target  
Settings  
Review and finish  
Review  
Deployment

SQL server → Azure Synapse dedicated SQL pool

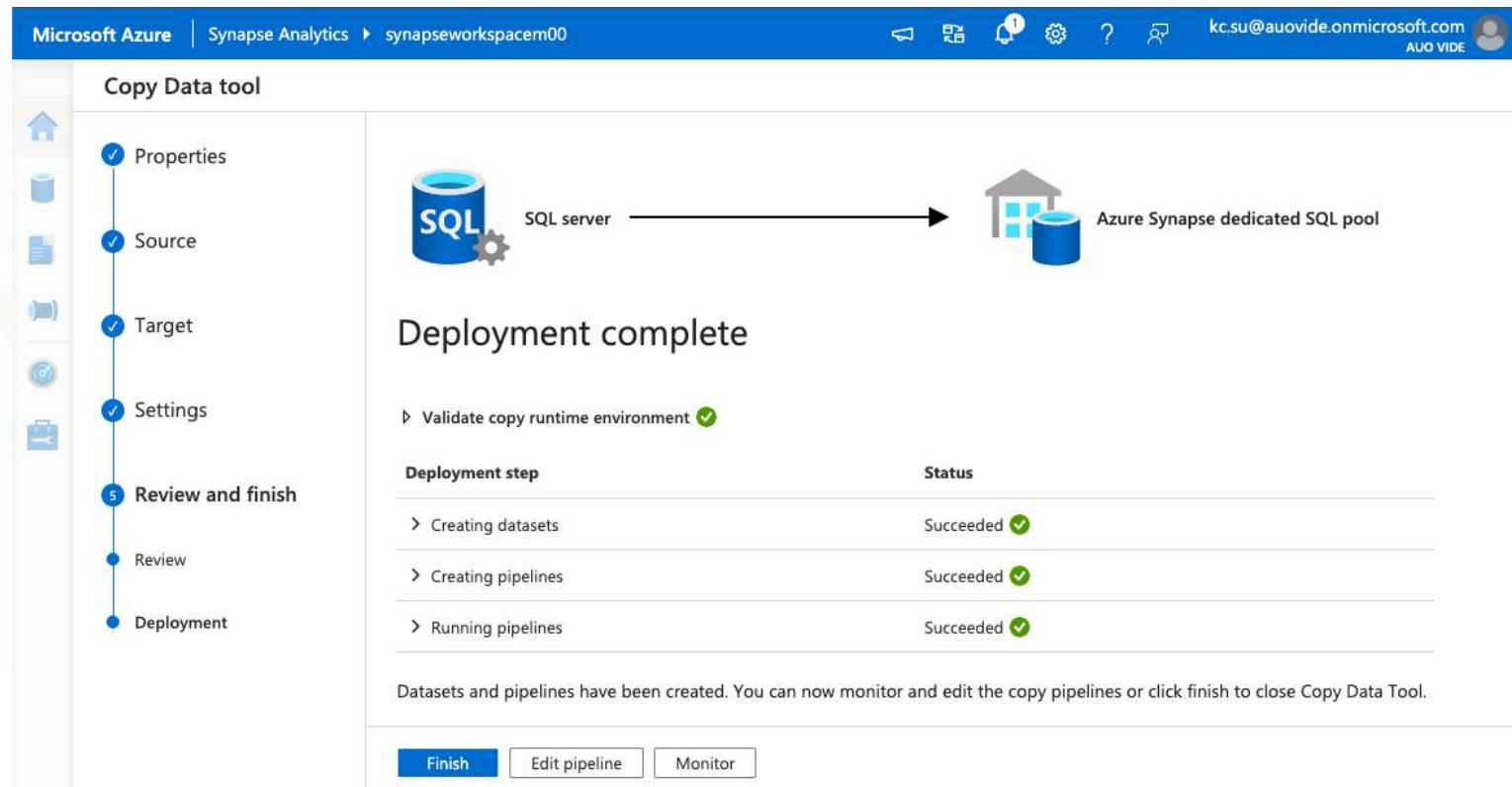
Deployment complete

Validate copy runtime environment ✓

Deployment step	Status
> Creating datasets	Succeeded ✓
> Creating pipelines	Succeeded ✓
> Running pipelines	Succeeded ✓

Datasets and pipelines have been created. You can now monitor and edit the copy pipelines or click finish to close Copy Data Tool.

Finish Edit pipeline Monitor



# (PowerBI) Step - Ingest Data 12

- 過程中可以去 Monitor 標籤頁監看執行狀況

The screenshot shows the Microsoft Azure Synapse Analytics pipeline runs interface. The left sidebar navigation includes: Home, Analytics pools, SQL pools, Apache Spark pools, Activities (selected), SQL requests, Apache Spark applications, Data flow debug, Integration, Pipeline runs (selected), Trigger runs, and Integration runtimes. The main content area displays the 'CopyPipeline\_ContosoBIDW' pipeline. It shows a 'ForEach' activity named 'ForEach\_dmm' which contains a 'Copy\_dmm' activity. Below this, the 'Activity runs' section lists six items, all of which are in progress or succeeded. The table columns are: Activity name, Activity type, Run start, Duration, Status, Error, and Log.

Activity name	Activity type	Run start ↑	Duration	Status	Error	Log
Copy_dmm	Copy data	7/17/21, 10:15:18 AM	00:02:39	In progress		
Copy_dmm	Copy data	7/17/21, 10:15:14 AM	00:00:13	Succeeded		
Copy_dmm	Copy data	7/17/21, 10:15:14 AM	00:00:13	Succeeded		
Copy_dmm	Copy data	7/17/21, 10:15:14 AM	00:00:12	Succeeded		
Copy_dmm	Copy data	7/17/21, 10:15:14 AM	00:00:12	Succeeded		
ForEach_dmm	ForEach	7/17/21, 10:15:14 AM	00:02:43	In progress		

# (PowerBI) Step - Ingest Data 13

- 展開 Synapse Studio 的 Data 標籤頁，就可以看到 Import 進來的 Table

The screenshot shows the Microsoft Azure Synapse Studio interface. The left sidebar has icons for Home, Data (which is selected), Develop, Integrate, Monitor, and Manage. The main area is titled 'Data' and shows a 'Workspace' tab selected. Under 'Databases', there are two entries: 'sqlpoolm00 (SQL)' and 'NYCSafetyData (SQL)'. 'sqlpoolm00 (SQL)' is expanded to show its tables: 'dbo.DimChannel', 'dbo.DimProduct', 'dbo.DimPromotion', 'dbo.DimStore', 'dbo.FactSales', 'External tables', 'External resources', 'Views', 'Programmability', 'Schemas', and 'Security'. A large graphic of two cylinders and a screen with code tags is displayed in the center-right, with the text 'Select an item' and 'Use the resource explorer to select or create a new item'.

# (PowerBI) Step - Explore Data 01

- 在 FactSales Table 按下滑鼠右鍵，選取 SELECT TOP 100 rows，Synapse Studio 就會執行相對應的 SQL 敘述

The screenshot shows the Microsoft Azure Synapse Studio interface. The left sidebar has icons for Home, Data, Develop, Integrate, Monitor, and Manage. The main area is titled 'Data' and shows the 'Workspace' tab selected. Under 'Databases', there is a tree view of databases and tables. The 'dbo.FactSales' table is selected. A context menu is open over the table, with the option 'Select TOP 100 rows' highlighted. The 'SQL script 1' pane contains the following SQL code:

```
1  SELECT TOP (100) [SalesKey]
2  ,[DateKey]
3  ,[channelKey]
4  ,[StoreKey]
5  ,[ProductKey]
6  ,[PromotionKey]
7  ,[CurrencyKey]
8  ,[UnitCost]
9  ,[UnitPrice]
10 , [SalesQuantity]
11 , [ReturnQuantity]
12 , [ReturnAmount]
13 , [DiscountQuantity]
14 , [DiscountAmount]
15 , [TotalCost]
```

The 'Results' pane shows the first 15 rows of the query output:

Key	StoreKey	ProductKey
309	1914	
96	1150	

At the bottom, a message says '00:00:01 Query executed successfully.'

# (PowerBI) Step - Explore Data 02

- 在 FactSales Table 按下滑鼠右鍵，選取 New notebook 裡面的 Load to DataFrame，Synapse Studio 就會建立 Spark Notebook 與相對應的 Scala 程式
- 重點是 Attach to 沒有設定，Language 也不是 PySpark

The screenshot shows the Microsoft Azure Synapse Analytics Studio interface. On the left, the 'Data' sidebar is open, showing a 'Workspace' tab and a 'Databases' section with 'sqlpoolm00 (SQL)' expanded, displaying tables like 'dbo.FactSales'. A context menu is open over 'dbo.FactSales', with 'New notebook' selected, which has a submenu 'Load to DataFrame'. The main workspace shows 'Notebook 1' with the following Scala code:

```
1 %%spark
2 val df = spark.read.sqlAnalytics("sqlpoolm00.dbo.FactSales")
3 df.write.mode("overwrite").saveAsTable("default.t1")
```

# (PowerBI) Step - Explore Data 03

- Attach to 選取剛剛建立的 Apache Spark Pool , Language 選取 PySpark
- 使用 Scala 載入 Data 放到 Data Frame
- 註解掉 `df.write.mode()`

```
%%spark
val df = spark.read.sqlanalytics("SQLPool名稱 dbo.FactSales")
// df.write.mode("overwrite").saveAsTable("default.t1")
```

- 再建立 Temporary View

```
%%spark
df.createOrReplaceTempView("tvFactSales")
```

```
[1] 1 %%spark
2 val df = spark.read.sqlanalytics("sqlpoolm00 dbo.FactSales")
3 // df.write.mode("overwrite").saveAsTable("default.t1")
```

Apache Spark session started in 3 min 8 sec 263 ms. Command executed in 4 sec 139 ms by kc.su on 12:13:50 AM, 7/16/21

df: org.apache.spark.sql.DataFrame = [SalesKey: int, DateKey: timestamp ... 17 more fields]

```
[2] 1 %%spark
2 df.createOrReplaceTempView("tvFactSales")
```

Command executed in 2 sec 134 ms by kc.su on 12:14:42 AM, 7/16/21

# (PowerBI) Step - Explore Data 04

- 就可以使用 SparkSQL 處理 Temporary View 分析 Data

```
%%sql  
SELECT COUNT(*) FROM tvFactSales
```

```
%%sql  
SELECT * FROM tvFactSales  
WHERE DateKey >= "2009-01-01T00:00:00Z" AND DateKey <= "2009-12-31T23:59:59"  
ORDER BY DateKey
```

The screenshot shows a Jupyter Notebook interface with two code cells and their execution results.

**Code Cell 1:**

```
%%sql  
SELECT COUNT(*) FROM tvFactSales
```

**Code Cell 2:**

```
%%sql  
SELECT * FROM tvFactSales  
WHERE DateKey >= "2009-01-01T00:00:00Z" AND DateKey <= "2009-12-31T23:59:59"  
ORDER BY DateKey
```

**Execution Results:**

Command executed in 16 sec 507 ms by kc.su on 12:20:42 AM, 7/16/21

> Job execution Succeeded Spark 2 executors 8 cores

View in monitoring Open Spark UI

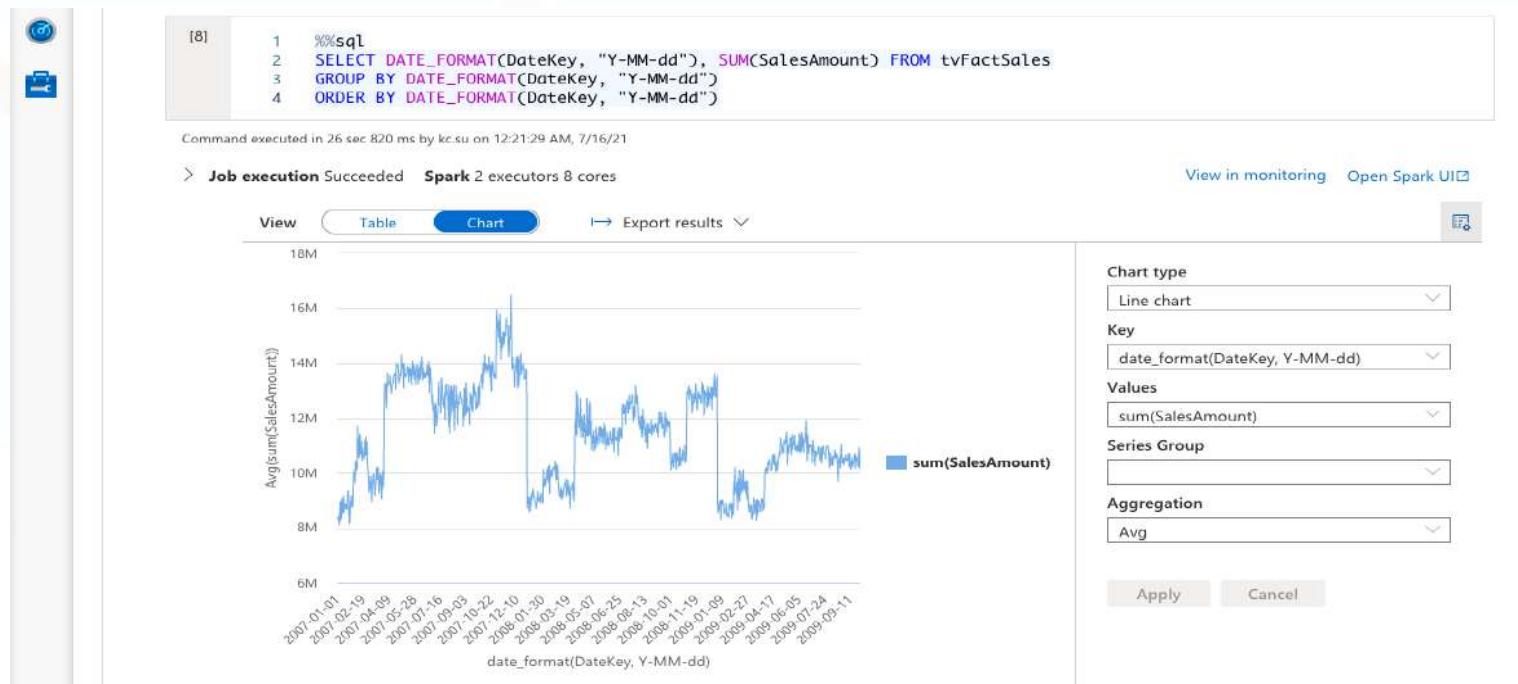
View Table Chart Export results

SalesKey	DateKey	channelKey	StoreKey	ProductKe
500052	2009-01-01T00:00:00Z	1	262	2163
1005445	2009-01-01T00:00:00Z	2	307	2072
246779	2009-01-01T00:00:00Z	1	214	348
70466	2009-01-01T00:00:00Z	1	16	392
337823	2009-01-01T00:00:00Z	2	199	577
55979	2009-01-01T00:00:00Z	3	200	737

# (PowerBI) Step - Explore Data 05

- 把顯示結果的 View 從 Table 切換到 Chart，就可以看到視覺化的圖表
- 如果不滿意還可以調整 Chart type 之類的設定

```
%%sql
SELECT DATE_FORMAT(DateKey, "Y-MM-dd"), SUM(SalesAmount) FROM tvFactSales
GROUP BY DATE_FORMAT(DateKey, "Y-MM-dd")
ORDER BY DATE_FORMAT(DateKey, "Y-MM-dd")
```



# (PowerBI) Step - Power BI Workspace 01

- 在 Power BI 網站按下 Sign in 使用 Azure 帳號登入

The screenshot shows the Microsoft Power BI homepage. At the top, there is a navigation bar with the Microsoft logo, a 'Power BI' button, and links for 'Overview', 'Products', 'Pricing', 'Solutions', and 'More'. On the right side of the nav bar are 'Search', 'Power BI service', 'Sign in' (with a user icon), 'Try free', and a 'Buy now' button. The main content area features a large, blurred background image of a woman in an office environment. Overlaid on this image is the text 'Find clarity when you need it most' in large white letters, followed by a subtitle: 'Empower team members to discover insights hidden in your data with Microsoft Power BI.' Below the subtitle is a yellow 'Start free >' button. At the bottom of the page, there is a black footer bar with the text 'Watch Microsoft Inspire now to hear from global partners and experts about delighting customers and propelling your business forward >'.

# (PowerBI) Step - Power BI Workspace 02

- 登入之後再按下帳號旁邊的 Power BI service 進入 Power BI Portal

The screenshot shows the Power BI Home page. On the left is a navigation sidebar with links like Favorites, Recent, Create, Datasets, Goals, Apps, Shared with me, Discover, Learn, Workspaces, and My workspace. The main area features a "Good morning" greeting and a section titled "Data stories from the Power BI community". It displays three cards: "THE DEFINITIVE 100 MOST USEFUL PRODUCTIVITY TIPS" by Alice\_Drummond, "Cancer statistics in the USA" by mmatey, and "Sports ranked by degree of difficult..." by Patrick Baumgartner. A modal window in the top right corner informs the user that a "Power BI free license assigned" and provides a link to learn more.

# (PowerBI) Step - Power BI Workspace 03

- 展開 Workspaces 標籤頁，按下 Create a workspace 按鈕

The screenshot shows the Power BI Home interface. On the left, there is a navigation sidebar with various options like Favorites, Recent, Create, Datasets, Goals, Apps, Shared with me, Discover, Learn, Workspaces, My workspace, and Get data. The 'Workspaces' option is currently selected, indicated by a yellow arrow icon. In the main content area, there is a search bar and a 'New report' button. Below that, there's a 'My workspace' section with a search bar and a message saying 'No results found'. To the right, there are several data visualizations: a map of the USA with cancer statistics, a bar chart for USA cancer statistics, a scatter plot titled 'Ranking Sports by Degree of Difficulty for Key Skills', and a heatmap titled 'Sports ranked by degree of difficult...'. At the bottom, there are buttons for 'Create a workspace' and 'How to create reports'.

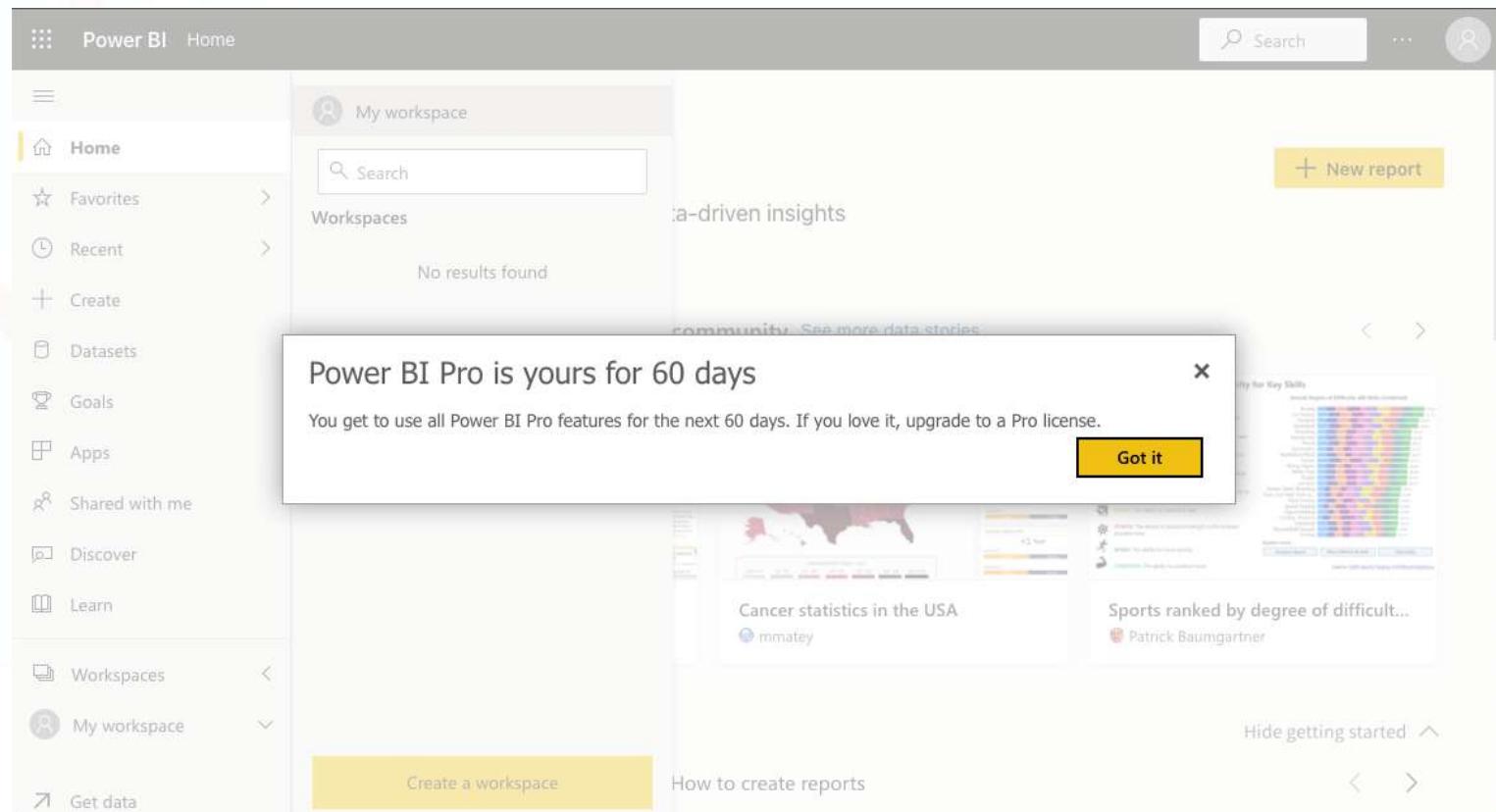
# (PowerBI) Step - Power BI Workspace 04

- Workspace 是 Power BI Pro 的功能，所以請按下 Try free 按鈕，先試用 60 天

The screenshot shows the Power BI Home interface. On the left is a navigation sidebar with options like Home, Favorites, Recent, Create, Datasets, Goals, Apps, Shared with me, Discover, Learn, Workspaces, My workspace, and Get data. The 'My workspace' section is currently selected. In the center, there's a search bar and a 'New report' button. Below the search bar, it says 'My workspace' and 'Search'. A message 'No results found' is displayed under 'Workspaces'. A large modal window titled 'Upgrade to Power BI Pro' is open in the center. It contains the text: 'To use Power BI Pro features, upgrade to a Pro license, or try Pro free for 60 days. [Learn more about Pro](#)'. Below that, it says 'By clicking "Try free" you agree to the [terms and conditions](#)'. There are three buttons at the bottom of the modal: 'Buy now' (yellow), 'Try free' (white), and 'Not now' (gray). In the background, there are preview cards for reports: 'Cancer statistics in the USA' by mmatay and 'Sports ranked by degree of difficult...' by Patrick Baumgartner. At the bottom of the page, there are links for 'Create a workspace' and 'How to create reports'.

# (PowerBI) Step - Power BI Workspace 05

- 按下 Got it 按鈕，試用 60 天 Power BI Pro



# (PowerBI) Step - Power BI Workspace 06

- 再按一次 Got it 按鈕，確定收到 60 天 Power BI Pro 才有的 Premium 功能

The screenshot shows the Power BI Home page. At the top, there's a navigation bar with 'Power BI Home', a search bar ('Search'), and a user profile icon. A message box in the top right corner says 'Trial: 59 days left' and 'You now have all Premium features. We've given you access to all Premium features for 60 days.' It includes a 'Learn more' link, a 'Read the Privacy Statement' link, and a yellow 'Got it' button.

The main area has a 'Good morning' greeting and a section for 'Data stories from the Power BI community'. It features three cards:

- Explore the 100 most useful product...** by Alice\_Drummond. This card shows a thumbnail of a dashboard titled 'THE DEFINITIVE 100 MOST USEFUL PRODUCTIVITY TIPS'.
- Cancer statistics in the USA** by mmately. This card shows a map of the United States where states are colored according to cancer incidence rates.
- Ranking Sports by Degree of Difficulty for Key Skills**. This card displays a complex heatmap and a legend explaining various sports skills like 'ANALYTIC', 'MANUAL', 'PHYSICAL', 'MENTAL', 'SOCIAL', 'TECHNICAL', 'PERFORM', and 'CREATE'.

At the bottom, there's a 'Getting started with Power BI' section with links to 'Power BI basics', 'Sample reports', and 'How to create reports'. A 'Hide getting started' button is also present.

# (PowerBI) Step - Power BI Workspace 07

- 再展開一次 Workspaces 標籤頁，按下 Create a workspace 按鈕

The screenshot shows the Power BI Home interface. On the left, there's a sidebar with various navigation options like Favorites, Recent, Create, Datasets, Goals, Apps, Shared with me, Discover, Learn, Workspaces, My workspace, and Get data. The 'Workspaces' option is currently selected. In the main area, there's a search bar and a 'New report' button. Below that, there's a 'Getting started' section with a 'Cancer Analytics Dashboard' card showing a map of the USA and some statistics. To the right of that is a 'Ranking Sports by Degree of Difficulty for Key Skills' card with a heatmap. At the bottom of the main content area, there are buttons for 'Create a workspace' and 'How to create reports'. A yellow banner at the very bottom of the screen also says 'Create a workspace'.

# (PowerBI) Step - Power BI Workspace 08

- Workspace name 設定為 ContosoBI
- Description 設定為 ContosoBI Workspace
- 然後按下 Save 按鈕

The screenshot shows the Power BI Home interface on the left and a 'Create a workspace' dialog box on the right.

**Power BI Home (Left):**

- Home
- Favorites
- Recent
- Create
- Datasets
- Goals
- Apps
- Shared with me
- Discover
- Learn
- Workspaces
- My workspace
- Get data

**Create a workspace (Right):**

Trial: 59 days left | Search | User icon

### Create a workspace

Workspace image

Workspace name:

Available

Description:

[Learn more about workspace settings](#)

# (PowerBI) Step - Power BI Workspace 09

- 就可以進入 Power BI Pro 準備的 ContosoBI Workspace

The screenshot shows the Power BI workspace interface. At the top, there's a navigation bar with the title "Power BI ContosoBI", a "Trial: 59 days left" message, a search bar, and a user profile icon. Below the navigation bar is a sidebar containing links like Home, Favorites, Recent, Create, Datasets, Goals, Apps, Shared with me, Discover, Learn, Workspaces (which is currently selected), ContosoBI, and Get data. The main content area is titled "ContosoBI" and "ContosoBI Workspace". It features a "Create app" button, a "Welcome to workspaces" message with a "Start tour" button, and a large "Add content to this workspace" section with an "Add content" button. A central graphic shows a dashboard with a chart and a small device screen, both with plus signs indicating they can be added.

# (PowerBI) Step - Linked Service 01

- 展開 Synapse Studio 的 Manage 標籤頁
- 按下 External connections 底下的 Linked services，再按下 + New

The screenshot shows the Microsoft Azure Synapse Studio interface. The left sidebar has a 'Manage' tab selected. Under 'External connections', the 'Linked services' option is highlighted with a red box. The main content area is titled 'Linked services' and contains a list of existing linked services:

Name	Type	Related	Annotations
city_safety_newyork	Azure Blob Storage	0	
city_safety_newyork	Azure Blob Storage	0	
ContosoBIDW	SQL server	1	
synapseworkspace...	Azure Synapse Analytics	0	
synapseworkspace...	Azure Data Lake Storage ...	0	

# (PowerBI) Step - Linked Service 02

- 按下 Connect to Power BI 按鈕，或是搜尋 Power BI 再點擊

The screenshot shows the Microsoft Azure portal interface for Synapse Analytics. The left sidebar menu is visible with options like Home, Data, Develop, Integrate, Monitor, and Manage. Under the Manage section, 'Linked services' is selected. The main content area is titled 'New linked service' and features a yellow banner with the text 'Build interactive reports with integrated Power BI capabilities.' and a 'Connect to Power BI' button. Below the banner is a search bar with the text 'Power'. A filter bar at the top of the list shows categories: All, Azure, Compute, Database, File, Generic protocol, and NoSQL. The list displays several entries, with one entry for 'Power BI' highlighted, showing a small icon and the name 'Power BI'.

# (PowerBI) Step - Linked Service 03

- Name 設定為 PowerBIWorkspaceContosoBI
- Workspace 選取剛剛建立的 ContosoBI Power BI Workspace
- 然後按下 Create 按鈕

The screenshot shows the Microsoft Azure Synapse Analytics interface. The left sidebar has a 'Manage' section selected, showing options like Home, Data, Develop, Integrate, Monitor, and Manage. Under Manage, 'Linked services' is highlighted. The main area is titled 'New linked service (Power BI)'. It prompts the user to choose a name for the linked service, which is set to 'PowerBIWorkspaceContosoBI'. Below that is a 'Description' field with no input. The 'Tenant' dropdown is set to 'AUO VIDE (e7c24d00-7479-4343-8247-0204699693e8)'. The 'Workspace name' dropdown is set to 'ContosoBI (333979ed-2862-494a-bcca-7a14f92f166f)'. There are sections for 'Annotations' and 'Advanced'. At the bottom are 'Create', 'Back', and 'Cancel' buttons.

# (PowerBI) Step - Linked Service 04

- 記得按下畫面上方的 Publish all，才算真的完成 Linked Service 設定

The screenshot shows the Microsoft Azure Synapse Analytics portal. The left sidebar has navigation links: Home, Data, Develop, Integrate, Monitor, and Manage. The Manage section is currently selected. The main content area is titled "Linked services". It explains that linked services define connection information for Azure Synapse Analytics to connect to external resources. A "New" button is available to create a new linked service. A search bar "Filter by name" and a dropdown "Annotations : Any" are present. Below, a table lists six items:

Name	Type	Related
city_safety_newyork	Azure Blob Storage	0
city_safety_newyork	Azure Blob Storage	0
ContosoBIDW	SQL server	1
PowerBIWorkspaceContosoBI	Power BI	0
synapseworkspace00-WorkspaceDefaultSqlServer	Azure Synapse Analytics	0
synapseworkspace00-WorkspaceDefaultStorage	Azure Data Lake Storage Gen2	0

# (PowerBI) Step - Power BI Dataset 01

- 選取 Synapse Studio 的 Develop 標籤頁，展開剛剛的 Power BI Workspace，選取底下的 Power BI datasets，然後按下右邊上面的 + New Power BI dataset

The screenshot shows the Microsoft Azure Synapse Studio interface. The left sidebar has navigation links: Home, Data, Develop (which is selected), Integrate, Monitor, and Manage. The main area has tabs: Synapse live, Validate all, Publish all, and a search bar. Below these are buttons for New Power BI dataset and Refresh. A list titled 'Power BI datasets (PowerBIWorkspaceContosoBI)' shows a single item: 'SQL script - NYCSafetyData'. A note says it's a read-only view and to go to Power BI to manage datasets. Below this is a section for 'Power BI' datasets, which is currently empty. At the bottom, there's a large plus sign icon and the text 'No items to show'. A list of steps to create a dataset is provided:

1. Select data source
2. Download .pbids file
3. Open and publish with Power BI Desktop

# (PowerBI) Step - Power BI Dataset 02

- 再按下底下的 Start 按鈕

The screenshot shows the Microsoft Azure Synapse Analytics workspace interface. The top navigation bar includes 'Microsoft Azure', 'Synapse Analytics', 'synaps', a search bar, and various icons for notifications and account management. The user's email, 'kc.su@auovide.onmicrosoft.com', and profile picture are visible on the right.

The left sidebar contains navigation links: Home, Data, Develop, Integrate, Monitor, and Manage. The 'Develop' section is currently selected, showing sub-options like 'SQL scripts' (with a file named 'SQL script - NYCSafetyData'), 'Notebooks' (with files 'Notebook - ContosoBIDW' and 'Notebook - NYCSafetyData'), 'Power BI' (with a folder 'PowerBIWorkspaceContosoBI' containing 'Power BI datasets' and 'Power BI reports'), and a 'Start' button at the bottom.

The main content area displays a large circular icon of a laptop with a yellow Power BI logo on its screen. Below the icon, the text reads:

**Let's get started with Microsoft Power BI**

Create a Power BI dataset from a data source and publish it to Power BI to build reports in Azure Synapse Studio.

To begin, you'll need Power BI Desktop installed on your local machine

[Install Power BI Desktop](#)

At the bottom of the main content area, there are 'View documentation' and 'Start' buttons. To the right of the 'Start' button is a 'Cancel' button.

# (PowerBI) Step - Power BI Dataset 03

- Select a data source 選取剛剛的 Dedicated SQL Pool，再按下 Continue 按鈕

The screenshot shows the Microsoft Azure Synapse Analytics interface. The top navigation bar includes 'Microsoft Azure', 'Synapse Analytics', 'synaps', a search bar, and user information 'kc.su@auovide.onmicrosoft.com AUO VIDE'. The left sidebar has navigation links: Home, Data, Develop, Integrate, Monitor, and Manage. The main area is titled 'Develop' and shows a 'Filter resources by name' search bar. Under 'SQL scripts', there is a link to 'SQL script - NYCSafetyData'. Under 'Notebooks', there are links to 'Notebook - ContosoBIDW' and 'Notebook - NYCSafetyData'. Under 'Power BI', there is a folder 'PowerBIWorkspaceContosoBI' containing 'Power BI datasets' and 'Power BI reports'. A large circular icon with two blue cylinders represents a SQL pool. To its right are icons for a data flow and a laptop displaying a chart. Below these icons, the text 'Select a data source' is displayed, followed by the instruction: 'Select a SQL pool below to use as a data source. You'll be able to select tables from this pool when creating your dataset.' A 'Name' input field contains 'sqlpoolm00', which is highlighted with a red border. Below it is a list item 'NYCSafetyData' with a small icon. At the bottom of the screen are three buttons: 'Continue' (highlighted in blue), 'Back', and 'Cancel'.

# (PowerBI) Step - Power BI Dataset 04

- 就可以下載一個以 Dedicated SQL Pool 為主檔名的 .pbids 檔案
- 然後按下 Continue 按鈕

The screenshot shows the Microsoft Azure Synapse Analytics interface. The left sidebar has navigation links: Home, Data, Develop (which is selected), Integrate, Monitor, and Manage. The main area is titled 'Develop' and shows a list of resources under 'Power BI': 'PowerBIWorkspaceContosoBI' (selected), 'Power BI datasets' (underlined), and 'Power BI reports'. To the right of the list are three icons: two blue cylinders (SQL scripts), a blue document with a downward arrow (Notebooks), and a laptop displaying a chart (Power BI datasets). Below the list is a section titled 'Download .pbids file' with the instruction 'Download the .pbids file below and save it to your local drive.' A download link labeled 'sqlpoolm00.pbids' is shown, followed by a 'Download' button. At the bottom are 'Continue', 'Back', and 'Cancel' buttons.

# (PowerBI) Step - Power BI Dataset 05

- Power BI Desktop 可以透過這個檔案存取 Azure Synapse Workspace 的 Data
- 然後按下 Close and refresh 按鈕

Microsoft Azure | Synapse Analytics > synaps Search kc.su@auovide.onmicrosoft.com AUO VIDE

Home Data Develop Integrate Monitor Manage

Synapse live Validate

Develop

Filter resources by name

SQL scripts

SQL script - NYCSafetyData

Notebooks

Notebook - ContosoBIDW

Notebook - NYCSafetyData

Power BI

PowerBIWorkspaceContosoBI

Power BI datasets

Power BI reports

Open in Power BI Desktop

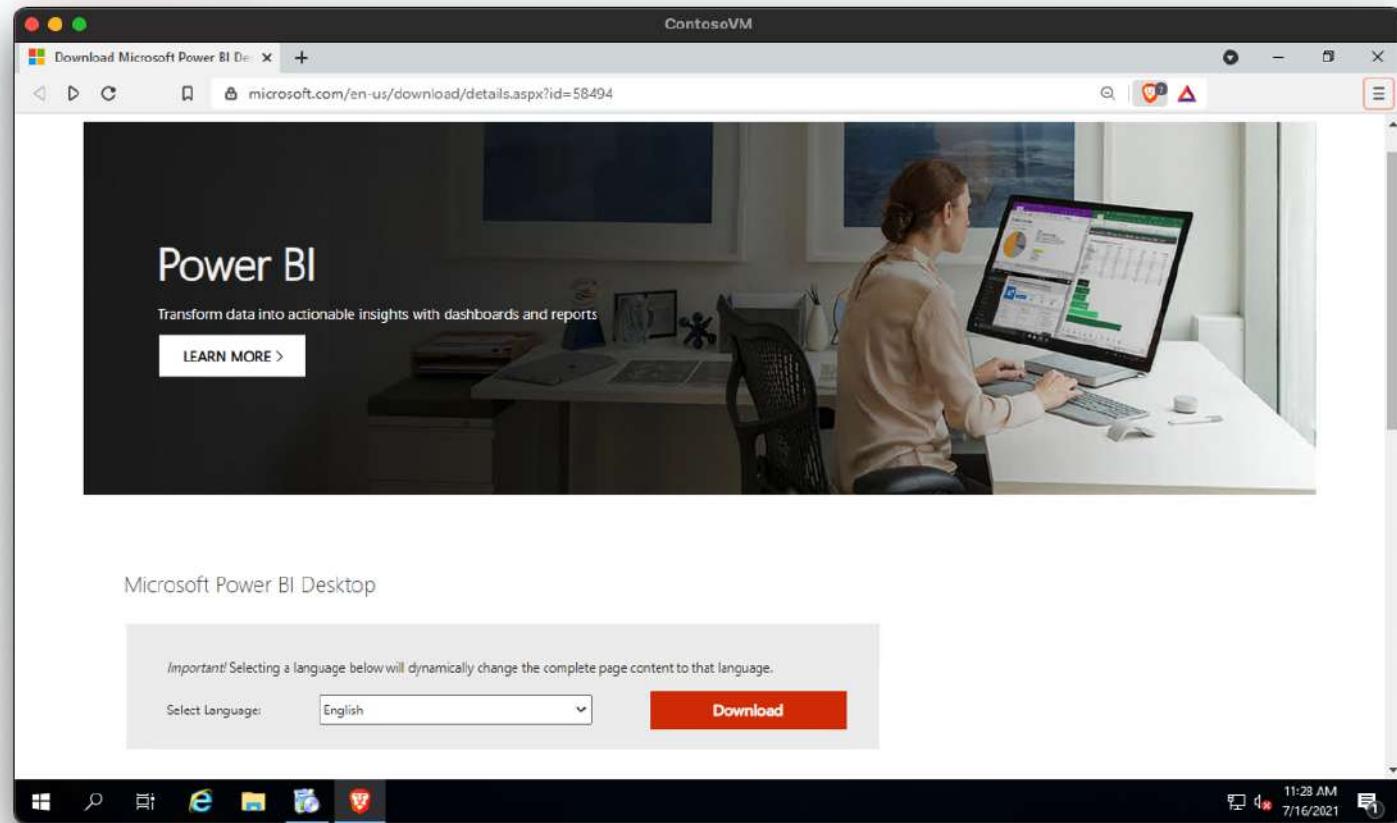
1. Open the downloaded .pbids file in Power BI Desktop.
2. Select the tables you want to include in your dataset and click Load to create the dataset.
3. When you're ready to publish, select File > Publish > Publish to Power BI and select [PowerBIWorkspaceContosoBI] as the destination.
4. Refresh your list of Power BI datasets in Azure Synapse Studio to see the newly published dataset. You can use it to build reports inside Azure Synapse Studio.

View documentation

Close and refresh Back

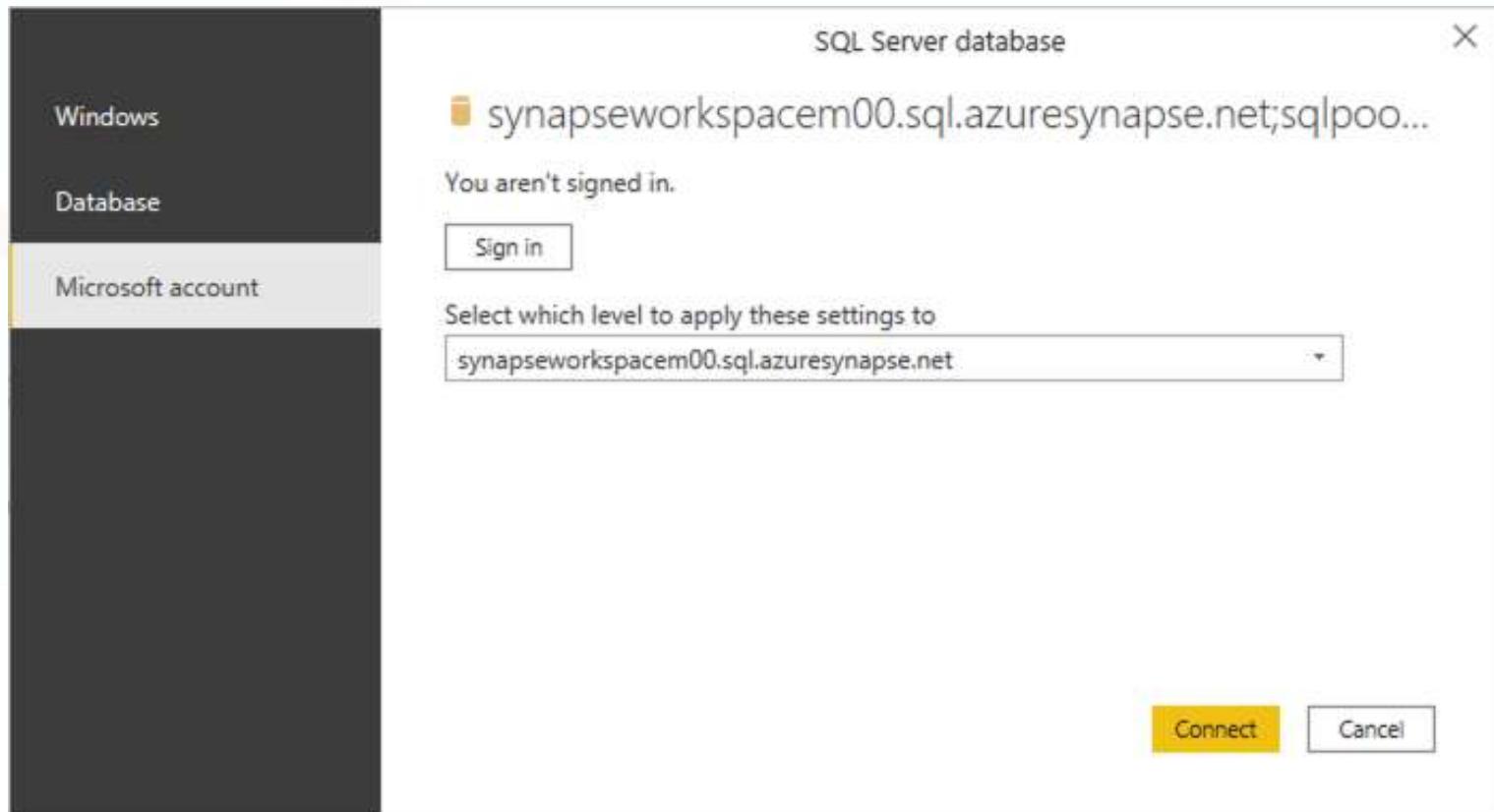
# (PowerBI) Step - Power BI Dataset 06

- 在 Microsoft Download Center 網站下載並安裝 Power BI Desktop



# (PowerBI) Step - Power BI Dataset 06

- 點擊 .pbids 檔案，就會執行 Power BI Desktop
- 點選左邊的 **Microsoft account**，按下 Sign in 按鈕，登入 Azure 帳號
- 再按下 Connect 按鈕，連上指定的 Azure Synapse Workspace



# (PowerBI) Step - Power BI Dataset 07

- 接下來就可以存取 Azure Synapse Workspace 的 Data
- 勾選所有的 Table，然後按下 Load 按鈕

Navigator

Display Options

synapseworkspace00.sql.azuresynapse.net: s...  
 DimChannel  
 DimProduct  
 DimPromotion  
 DimStore  
 FactSales

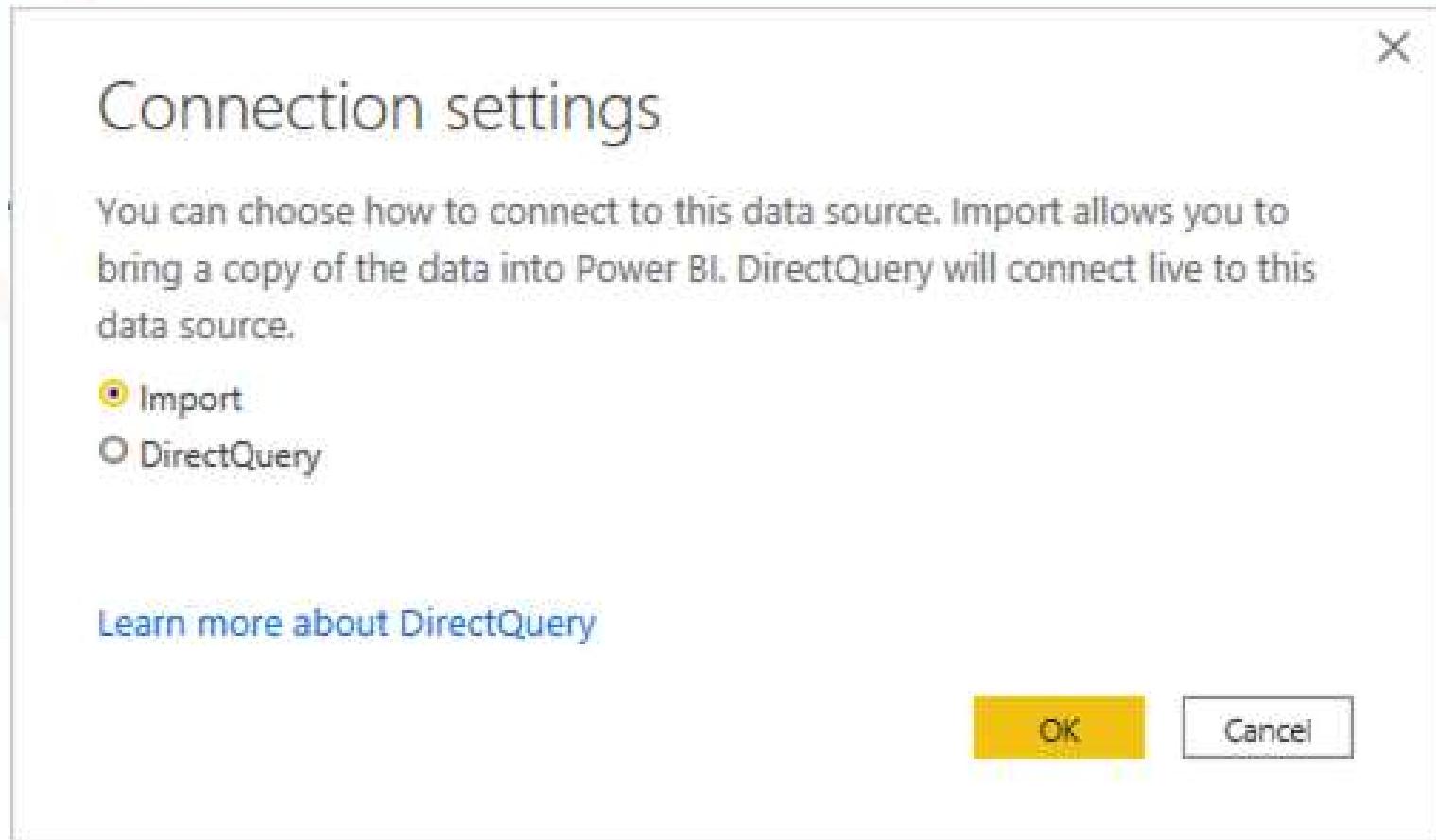
FactSales

SalesKey	DateKey	channelKey	StoreKey	ProductKey
5134	4/25/2009 12:00:00 AM	1	132	1
5378	5/24/2007 12:00:00 AM	1	22	1
5562	10/16/2008 12:00:00 AM	1	167	1
5809	10/23/2008 12:00:00 AM	1	270	2
29195	7/16/2007 12:00:00 AM	1	249	2
45758	6/23/2009 12:00:00 AM	1	165	1
34228	8/25/2009 12:00:00 AM	1	257	1
6703	5/21/2007 12:00:00 AM	1	167	1
34735	8/26/2009 12:00:00 AM	1	269	1
7281	10/11/2007 12:00:00 AM	1	116	1
35235	4/29/2007 12:00:00 AM	1	61	1
7915	5/14/2008 12:00:00 AM	1	232	1
8073	4/30/2009 12:00:00 AM	1	129	1
8251	6/24/2007 12:00:00 AM	1	231	1
8484	6/26/2008 12:00:00 AM	1	168	1
8863	9/5/2009 12:00:00 AM	1	279	2
9034	10/26/2008 12:00:00 AM	1	120	1
9319	10/9/2007 12:00:00 AM	1	129	2
37381	10/28/2007 12:00:00 AM	1	178	2
9772	10/26/2007 12:00:00 AM	1	252	2
10770	6/21/2008 12:00:00 AM	1	246	2
85422	10/26/2008 12:00:00 AM	1	15	1
59865	10/11/2007 12:00:00 AM	1	291	1

Select Related Tables

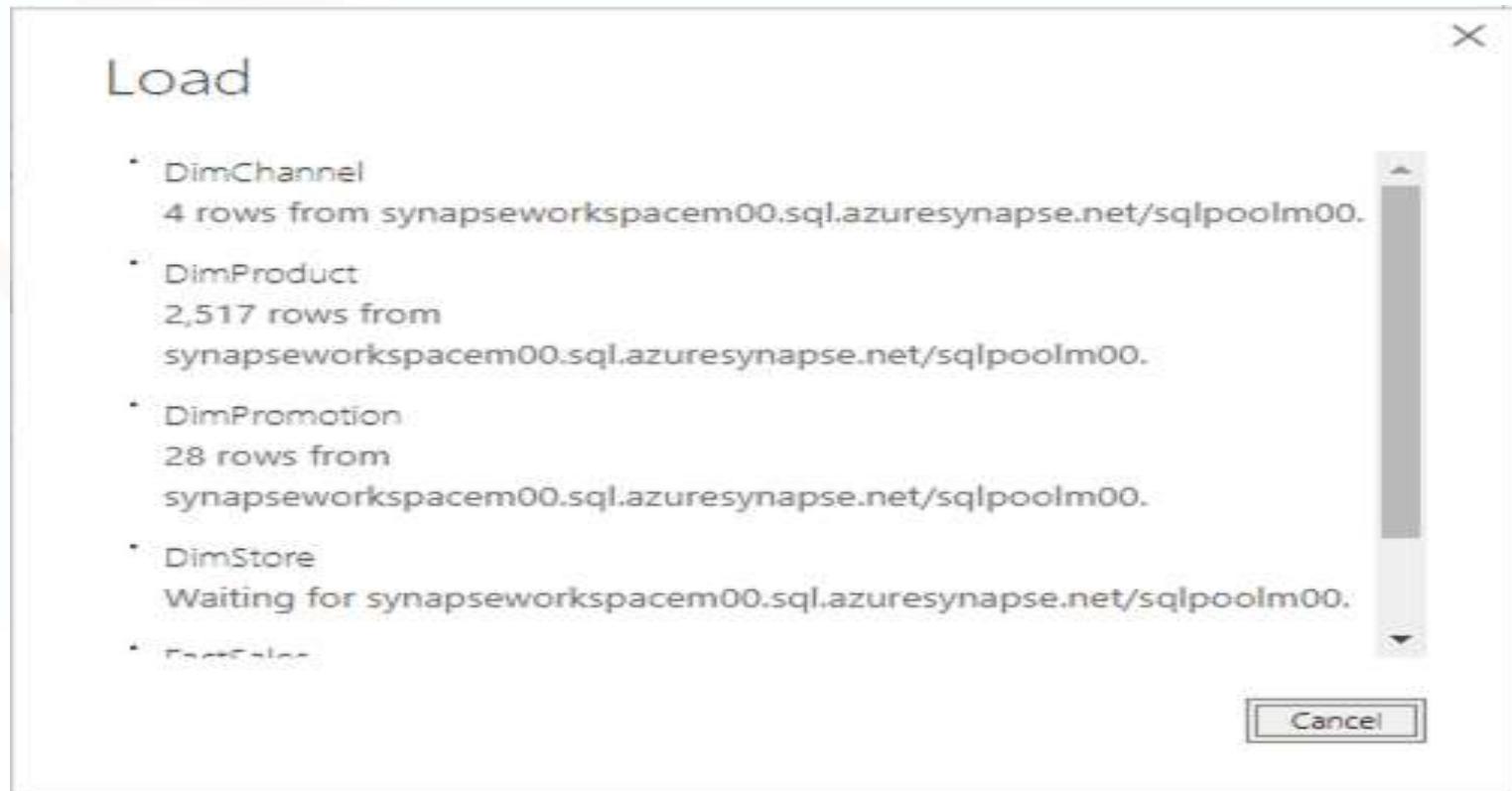
# (PowerBI) Step - Power BI Dataset 08

- Connection settings 視情況需要選取 Import 或 DirectQuery，再按下 OK 按鈕



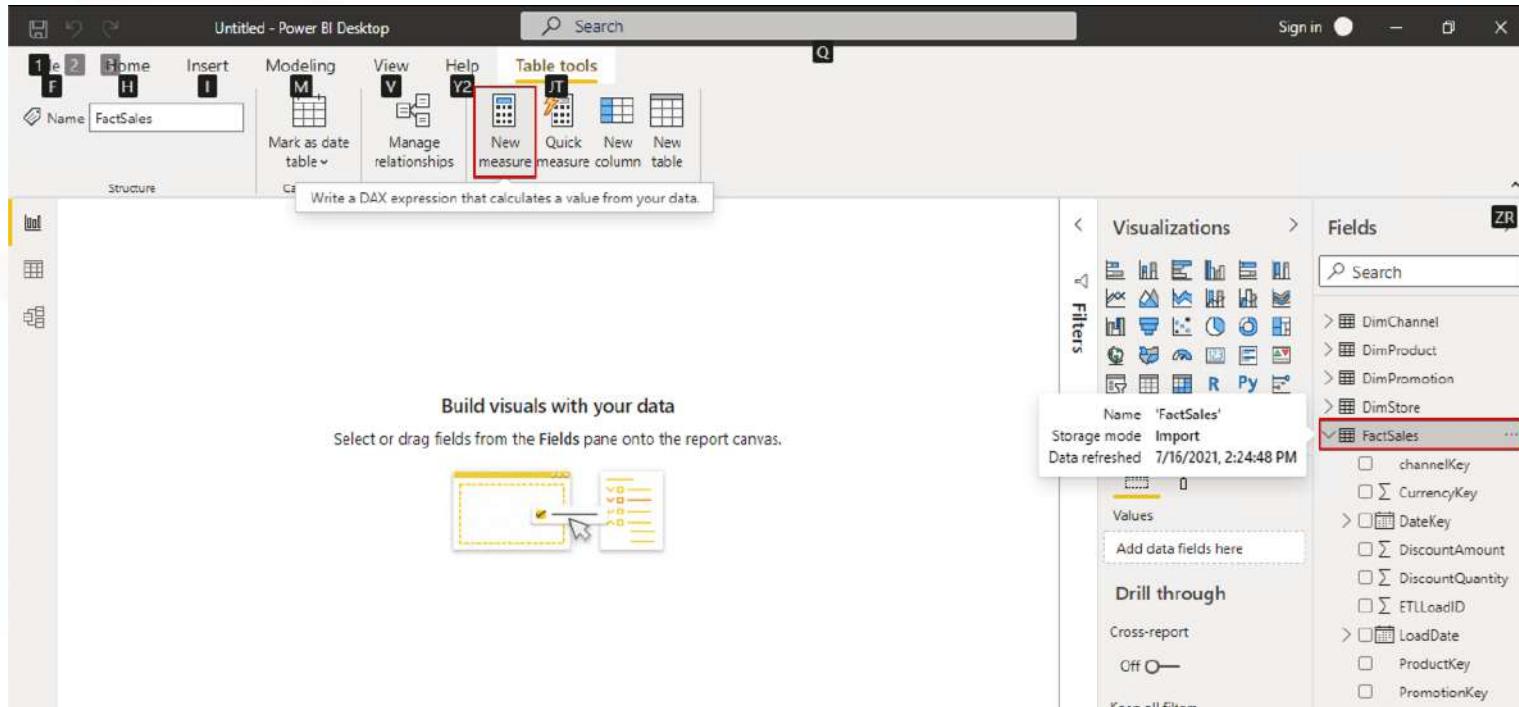
# (PowerBI) Step - Power BI Dataset 09

- Azure Synapse Workspace 的 Data 就會開始載入到 Power BI Desktop
- 記得要將進度儲存到 .pbids 檔案之中



# (PowerBI) Step - Analyze Data (Date) 01

- 右邊的 Fields 視窗裡面選取 FactSales Table
- 然後按下畫面上方的 New measure 按鈕



# (PowerBI) Step - Analyze Data (Date) 02

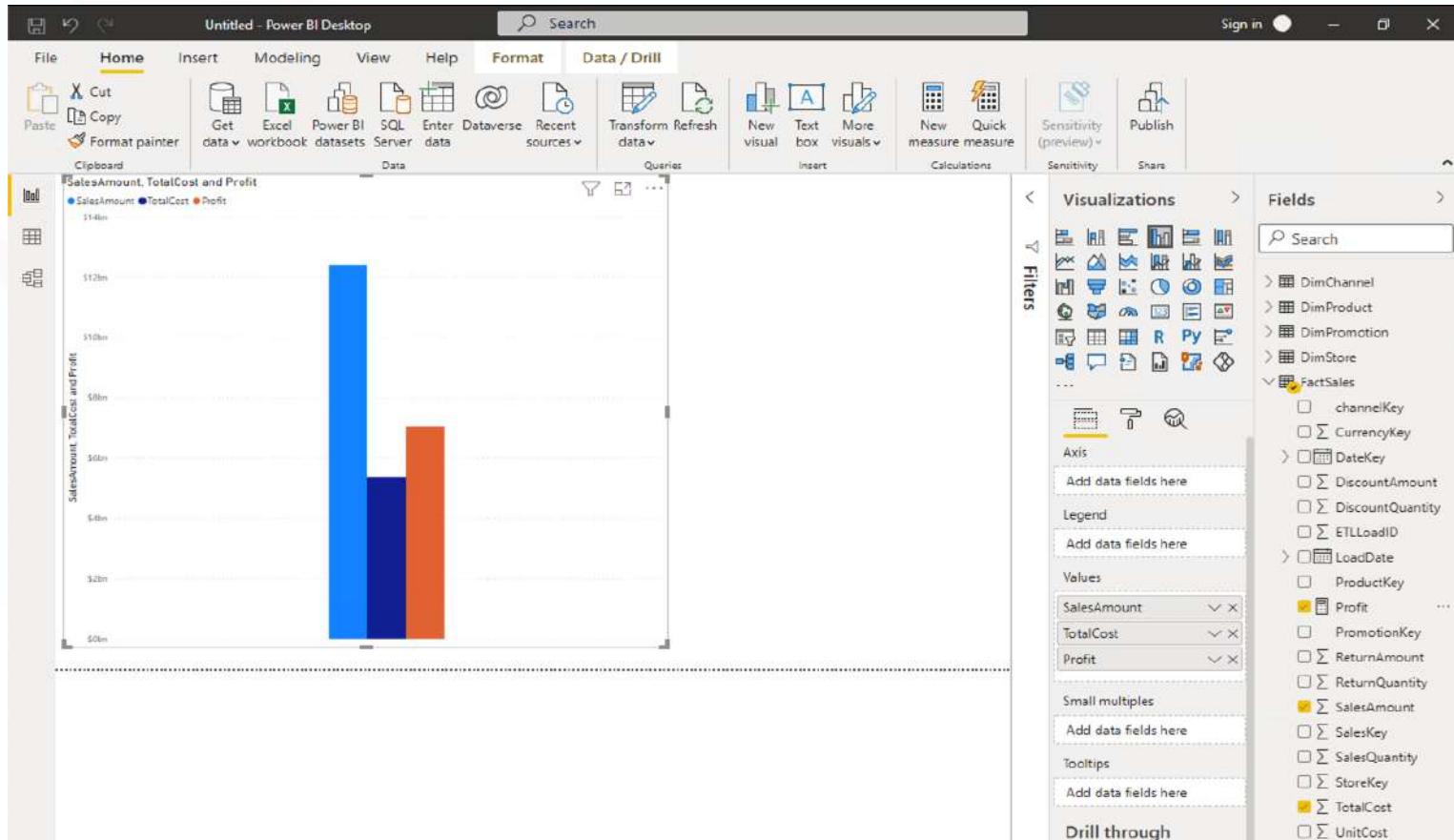
- 輸入底下的 Formula，再按下左邊的勾勾

Profit = SUM(FactSales[SalesAmount]) - SUM(FactSales[TotalCost])

The screenshot shows the Power BI Desktop interface with the title bar "Untitled - Power BI Desktop". The ribbon is visible with the "Table tools" tab selected. In the "Measure tools" section, a new measure named "Profit" is being defined. The formula is entered as "Profit = SUM(FactSales[SalesAmount]) - SUM(FactSales[TotalCost])". A red box highlights this formula entry. The "Formatting" and "Properties" tabs are also visible. On the right side, the "Fields" pane is open, showing the FactSales table with various columns listed. The "Profit" column is highlighted with a gray background.

# (PowerBI) Step - Analyze Data (Date) 03

- 右邊的 Fields 視窗裡面選取 FactSales Table 底下的 SalesAmount、TotalCost、與 Profit 三個 Field，就可以建立一個 Chart Visualization



# (PowerBI) Step - Analyze Data (Date) 04

- 不要點選任何 Chart
- 右邊的 Fields 視窗裡面選取 FactSales Table 底下的 DateKey、SalesAmount、TotalCost、與 Profit 四個 Field，就可以建立一個 Table View Visualization

The screenshot shows the Power BI Desktop interface with the following details:

- Home Tab:** Selected.
- Data Source:** Untitled - Power BI Desktop.
- Visualizations:** A bar chart titled "SalesAmount, TotalCost and Profit" showing data for November 2007. The Y-axis ranges from \$0m to \$12m. The chart has three bars: SalesAmount (blue), TotalCost (dark blue), and Profit (orange).
- Fields pane:** Shows the "FactSales" table with the following selected fields:
  - DateKey
  - SalesAmount
  - TotalCost
  - Profit
- Filters:** A section showing filters applied to the DateKey, Year, Quarter, Month, and Day fields.
- Drill through:** Set to "Off".
- Cross-report:** Set to "Keep all filters".
- Page Number:** 107 / 131.

# (PowerBI) Step - Analyze Data (Date) 05

- 右邊的 Visualizations 視窗中間，按下 Day 與 Month 的 X 圖示，就會從原來細到 Day 的內容，變成細到 Quarter，將這些 Data 以季度的方式呈現

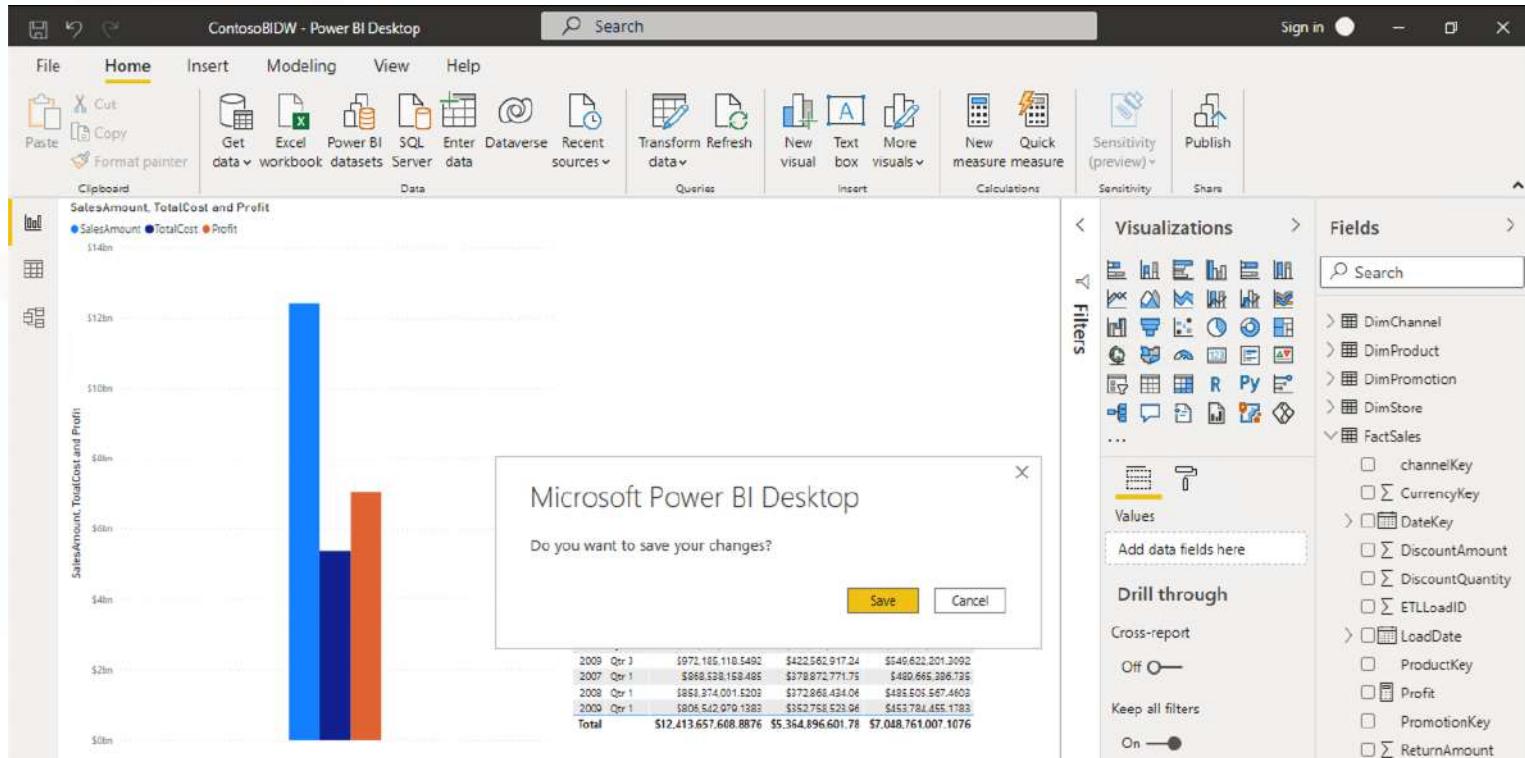
The screenshot shows the Power BI Desktop interface with the following details:

- Home Tab:** Selected.
- Data:** SalesAmount, TotalCost and Profit.
- Visualizations:** A bar chart showing SalesAmount, TotalCost, and Profit over time. The Y-axis ranges from \$0Bn to \$12Bn. The X-axis shows years and quarters (e.g., 2007 Qtr 4, 2008 Qtr 1).
- Fields pane:** Shows the FactSales table with columns: Year, Quarter, SalesAmount, TotalCost, Profit.
- Values section:** Contains DateKey, Year, and Quarter, all highlighted with a red box.
- Drill through:** Options include SalesAmount, TotalCost, and Profit.
- Visualizations section:** Options include DimChannel, DimProduct, DimPromotion, DimStore, and FactSales.
- Fields section:** Options include DimChannel, DimProduct, DimPromotion, DimStore, FactSales, and various measures like CurrencyKey, DiscountAmount, etc.

Year	Quarter	SalesAmount	TotalCost	Profit
2007	Qtr 4	\$1,302,823,992.0526	\$500,793,150.75	\$742,030,841,9020
2007	Qtr 2	\$1,233,387,004.9306	\$531,410,121.81	\$701,976,883,1286
2007	Qtr 3	\$1,157,191,798,945	\$495,630,924.5	\$661,360,874,445
2008	Qtr 4	\$1,108,106,254,3249	\$479,614,390,45	\$628,491,963,0749
2008	Qtr 3	\$1,009,384,328,059	\$472,746,072,19	\$528,631,255,869
2008	Qtr 2	\$1,045,368,950,7709	\$448,120,042,02	\$597,248,907,8599
2009	Qtr 2	\$982,018,016,3377	\$418,011,183,03	\$564,006,833,3077
2009	Qtr 4	\$979,737,005,1571	\$431,308,166,12	\$548,428,836,0371
2009	Qtr 3	\$972,185,118,5492	\$422,562,317,24	\$549,622,201,3092
2007	Qtr 1	\$668,538,158,485	\$378,872,771,75	\$499,665,306,735
2008	Qtr 1	\$952,374,001,5203	\$372,888,434,06	\$485,505,567,4603
2009	Qtr 1	\$805,542,979,1383	\$352,758,523,96	\$453,784,455,1783
Total		\$12,413,657,608,8876	\$5,364,896,601.78	\$7,048,761,007,1076

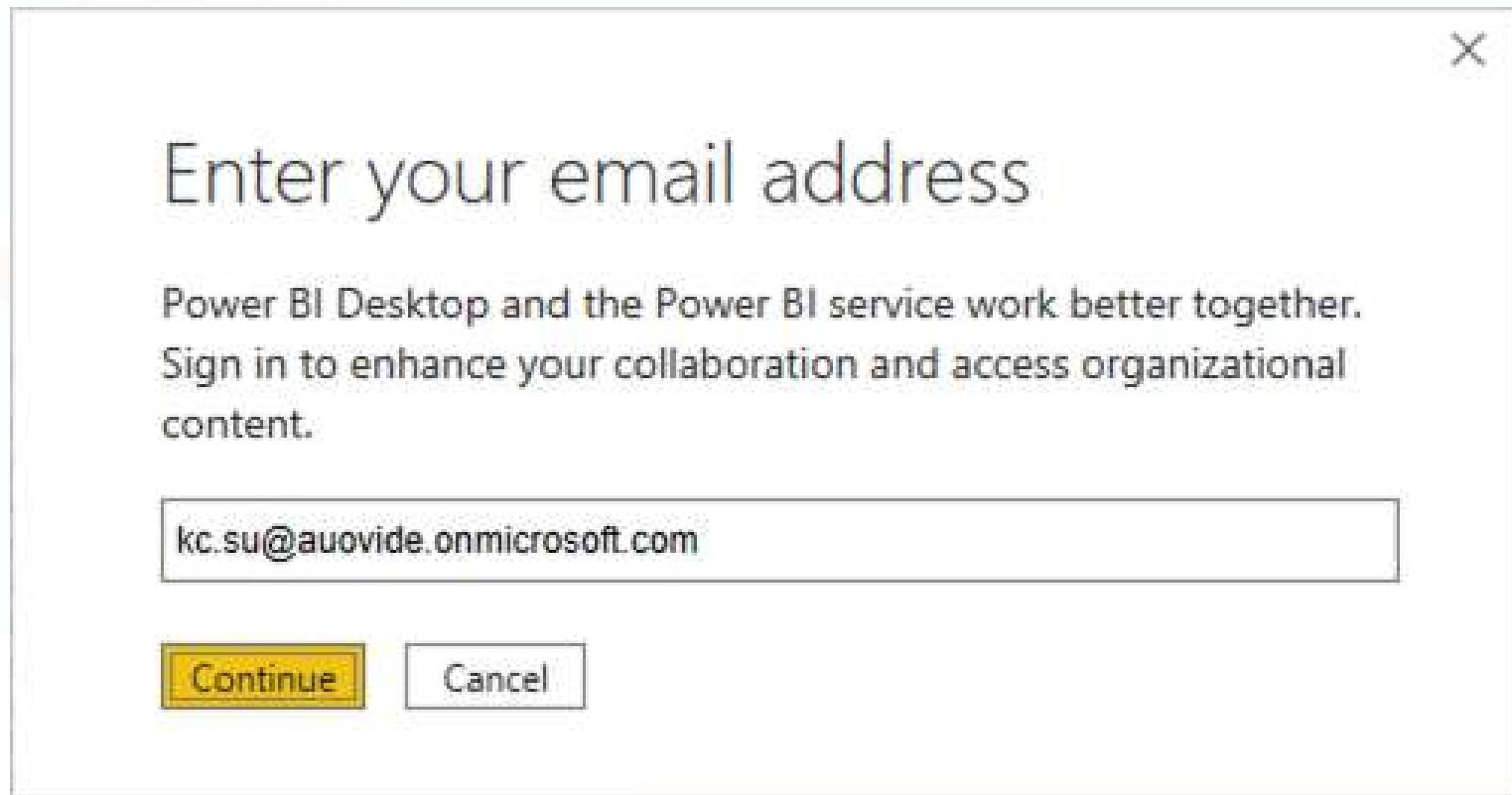
# (PowerBI) Step - Publish Report 01

- 按下畫面右上方的 Publish 按鈕
- 按下對話方塊中的 Save 按鈕，儲存所做的異動到 ContosoBIDW.pbix 檔案



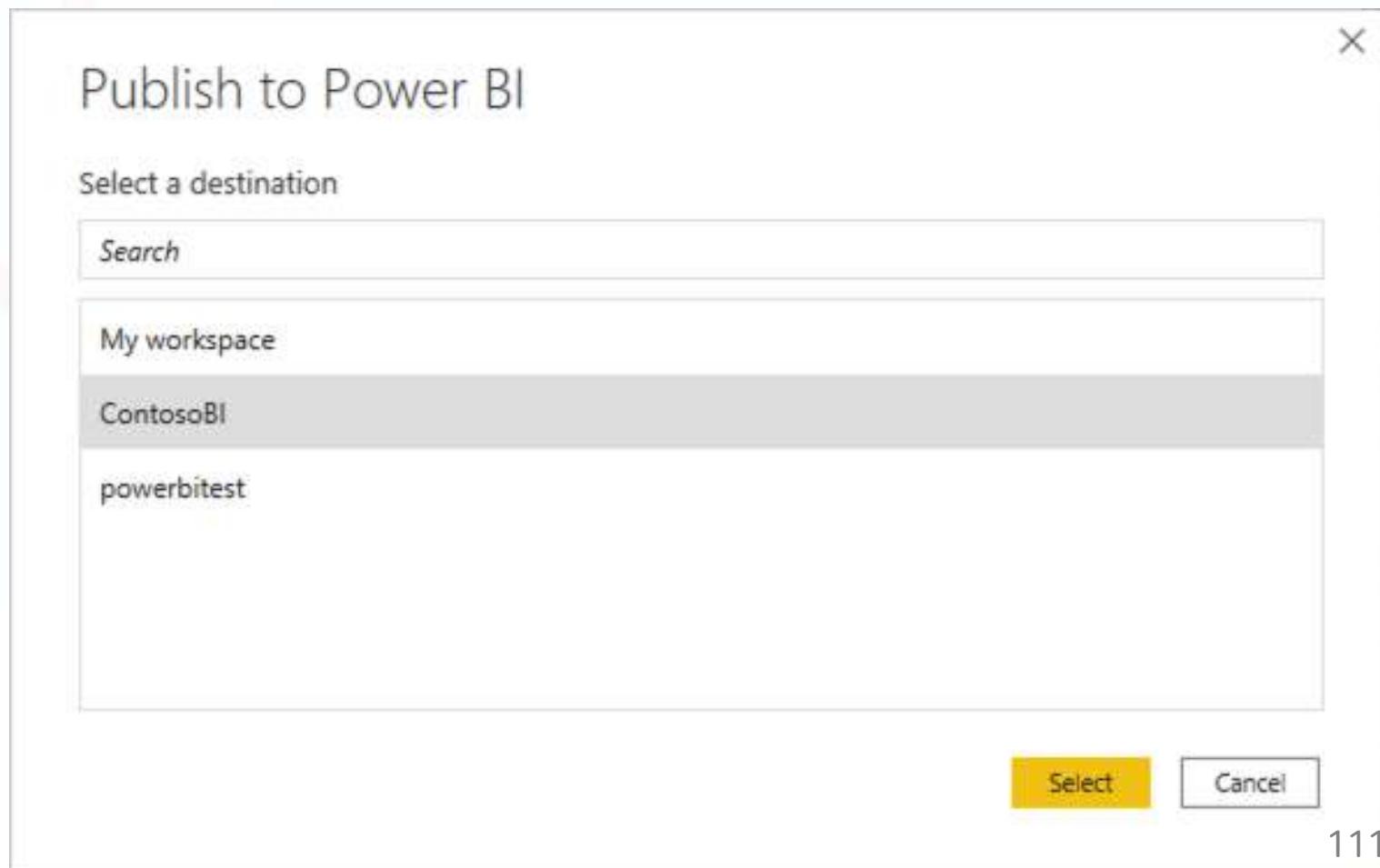
# (PowerBI) Step - Publish Report 02

- 輸入 Power BI Portal 帳號 (就是 Azure 帳號)
- 然後按下 Continue 按鈕，輸入密碼登入帳號



# (PowerBI) Step - Publish Report 03

- 選取 ContosoBI Workspace
- 然後按下 Select 按鈕



# (PowerBI) Step - Publish Report 04

- 就會將分析結果發佈到 Power BI Service

Publishing to Power BI

✓ Success!

[Open 'ContosoBIDW.pbix' in Power BI](#)

[Get Quick Insights](#)

---

 **Did you know?**

You can create a portrait view of your report, tailored for mobile phones.  
On the View tab, select Mobile Layout. [Learn more](#)

[Got it](#)

# (PowerBI) Step - Publish Report 05

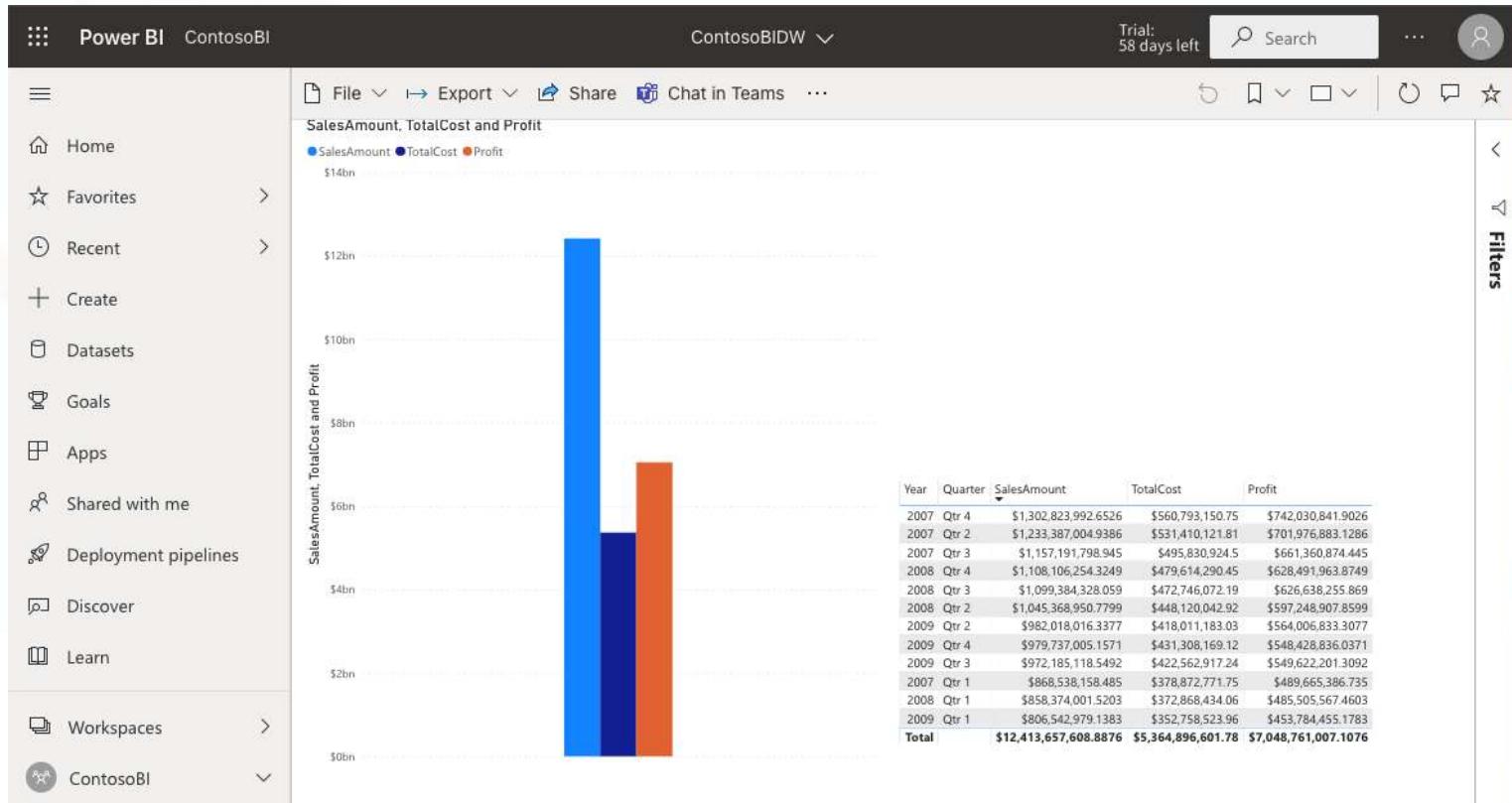
- 之後可以在 Synapse Workspace 開啟這份分析結果 Report，也可以進一步將這份 Report 以 PowerPoint / PDF / Excel 等方式匯出

The screenshot shows the Microsoft Azure Synapse Analytics workspace interface. The left sidebar displays the 'Develop' section with a list of resources: SQL scripts, Notebooks, and Power BI datasets. The 'Power BI' section is expanded, showing a folder for 'PowerBIWorkspaceContosoBI' which contains 'Power BI datasets' and 'Power BI reports'. A specific report named 'ContosoBIDW' is selected and highlighted in grey. The main workspace area shows a bar chart titled 'SalesAmount, TotalCost and Profit' with two bars: one blue and one red. Below the chart is a table with data for each quarter from 2008 to 2010. The right sidebar contains sections for 'Visualizations', 'Fields', 'Filters', and 'Values', along with a search bar and a 'Drill through' button.

Year	Quarter	SalesAmount	TotalCost	Profit
2008	Q1	\$1,234,567.89	\$890,123.45	\$344,444.44
2008	Q2	\$1,234,567.89	\$890,123.45	\$344,444.44
2008	Q3	\$1,234,567.89	\$890,123.45	\$344,444.44
2008	Q4	\$1,234,567.89	\$890,123.45	\$344,444.44
2009	Q1	\$1,234,567.89	\$890,123.45	\$344,444.44
2009	Q2	\$1,234,567.89	\$890,123.45	\$344,444.44
2009	Q3	\$1,234,567.89	\$890,123.45	\$344,444.44
2009	Q4	\$1,234,567.89	\$890,123.45	\$344,444.44
2010	Q1	\$1,234,567.89	\$890,123.45	\$344,444.44
2010	Q2	\$1,234,567.89	\$890,123.45	\$344,444.44
2010	Q3	\$1,234,567.89	\$890,123.45	\$344,444.44
2010	Q4	\$1,234,567.89	\$890,123.45	\$344,444.44
Total		\$11,473,634.76	\$8,901,234.75	\$2,572,400.01

# (PowerBI) Step - Publish Report 06

- 一樣可以在 Power BI Workspace 開啟這份分析結果 Report，也可以進一步將這份 Report 以 PowerPoint / PDF / Excel 等方式匯出

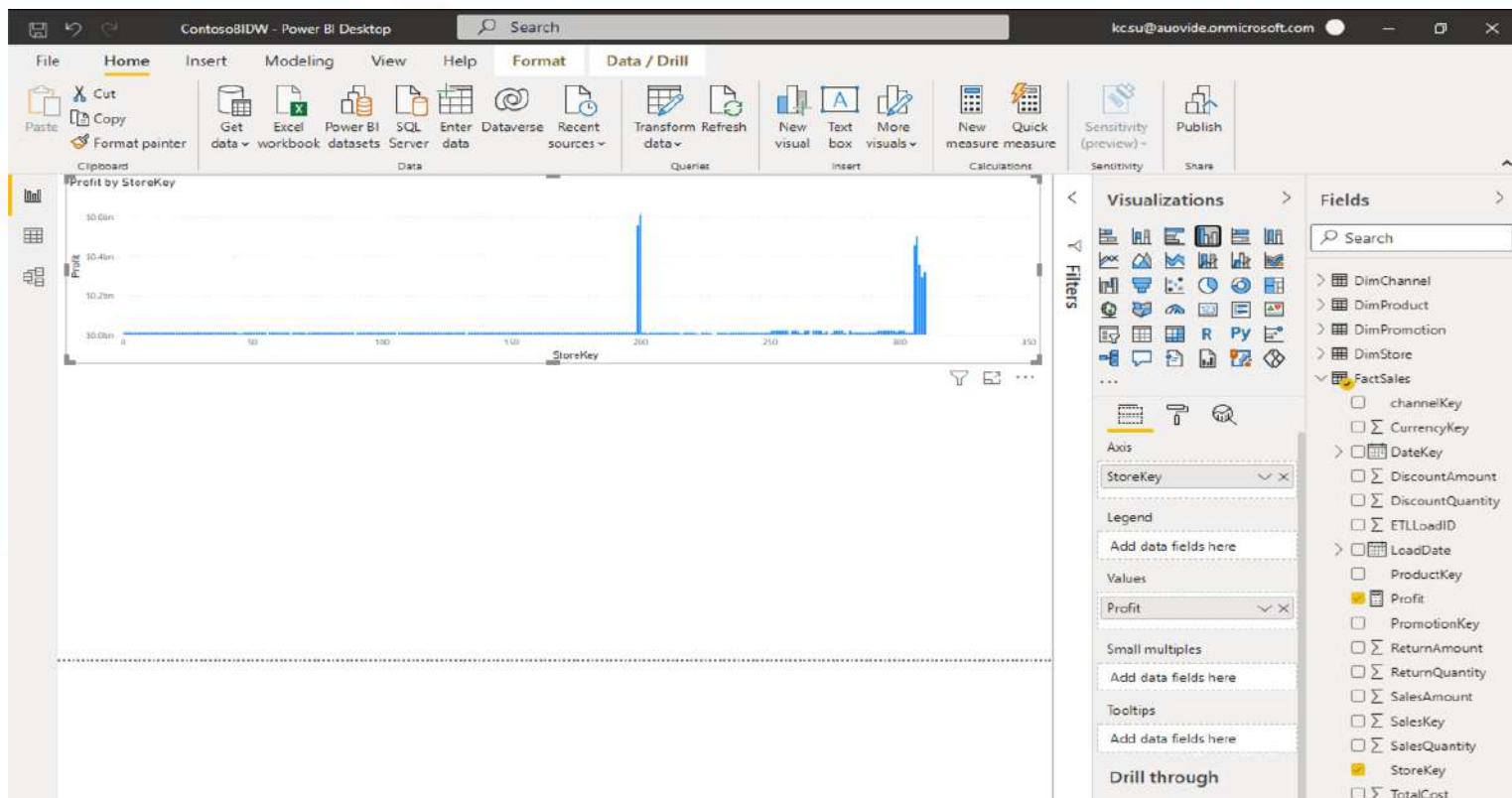


# 課後作業

# 1. Store 獲利分析

# Step - Analyze Data (Store) 01

- 按下畫面底下 Page 1 標籤頁右邊的 + 圖示，建立一個新的 Page
- Fields 視窗選取 FactSales Table 底下的 StoreKey 與 Profit 兩個 Field
- Visualization 視窗選取 Clustered Column Chart Visualization，再把 StoreKey 從 Values 拖曳到 Axis



# Step - Analyze Data (Store) 02

- 不要點選任何 Chart
- Fields 視窗選取 FactSales Table，然後按下畫面上方的 New column 按鈕

The screenshot shows the Power BI Desktop interface. The ribbon is selected, and the 'Table tools' tab is active. A red box highlights the 'New column' button in the ribbon. A tooltip above the button says 'Write a DAX expression that creates a new column in the selected table and calculates values for each row.' The main area displays a bar chart titled 'Profit by StoreKey'. The Y-axis is labeled 'Profit' and ranges from '\$0.0m to \$0.6m. The X-axis is labeled 'StoreKey' and ranges from 0 to 350. The chart shows two distinct peaks: one around StoreKey 200 reaching approximately \$0.55m, and another around StoreKey 300 reaching approximately \$0.5m. The Fields pane on the right shows the FactSales table with columns channelKey, CurrencyKey, and DateKey.

# Step - Analyze Data (Store) 03

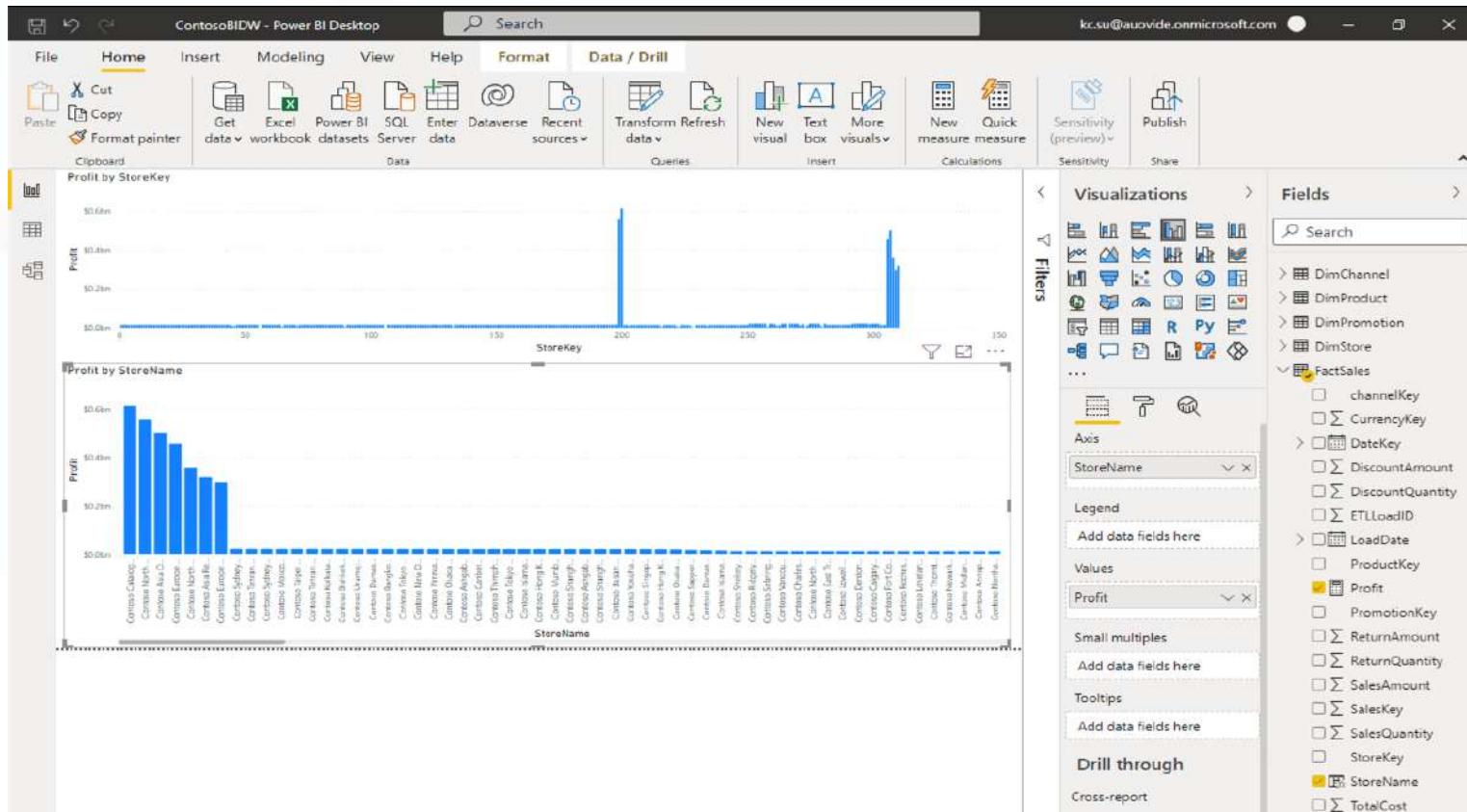
- 輸入底下的 Formula，再按下左邊的勾勾

StoreName = LOOKUPVALUE(DimStore[StoreName], DimStore[StoreKey], FactSales[StoreKey], "Unknown")

The screenshot shows the Power BI Desktop interface with the 'Column tools' ribbon selected. In the center, there is a histogram visual titled 'StoreName' with the formula `1 StoreName = LOOKUPVALUE(DimStore[StoreName], DimStore[StoreKey], FactSales[StoreKey], "Unknown")`. The x-axis is labeled 'StoreKey' and ranges from 0 to 350. The y-axis is labeled 'Profit' and ranges from 0.00m to 10.00m. The histogram shows two distinct peaks: one very tall peak at StoreKey ~200 and another smaller peak at StoreKey ~350. To the right of the visual is the 'Fields' pane, which lists various dimensions and fact tables. The 'FactSales' table is expanded, showing fields like channelKey, DateKey, DiscountAmount, etc. The 'StoreName' field is highlighted with a gray background.

# Step - Analyze Data (Store) 04

- 不要點選任何 Chart
- 選取 StoreName 與 Profit 兩個 Field 來建立 Clustered Column Chart Visualization



## 2. Product 獲利分析

# Step - Analyze Data (Product) 01

- 按下畫面底下 Page 2 標籤頁右邊的 + 圖示，建立一個新的 Page
- Fields 視窗裡面選取 FactSales Table，然後按下畫面上方的 New column 按鈕
- 輸入底下的 Formula，再按下左邊的勾勾

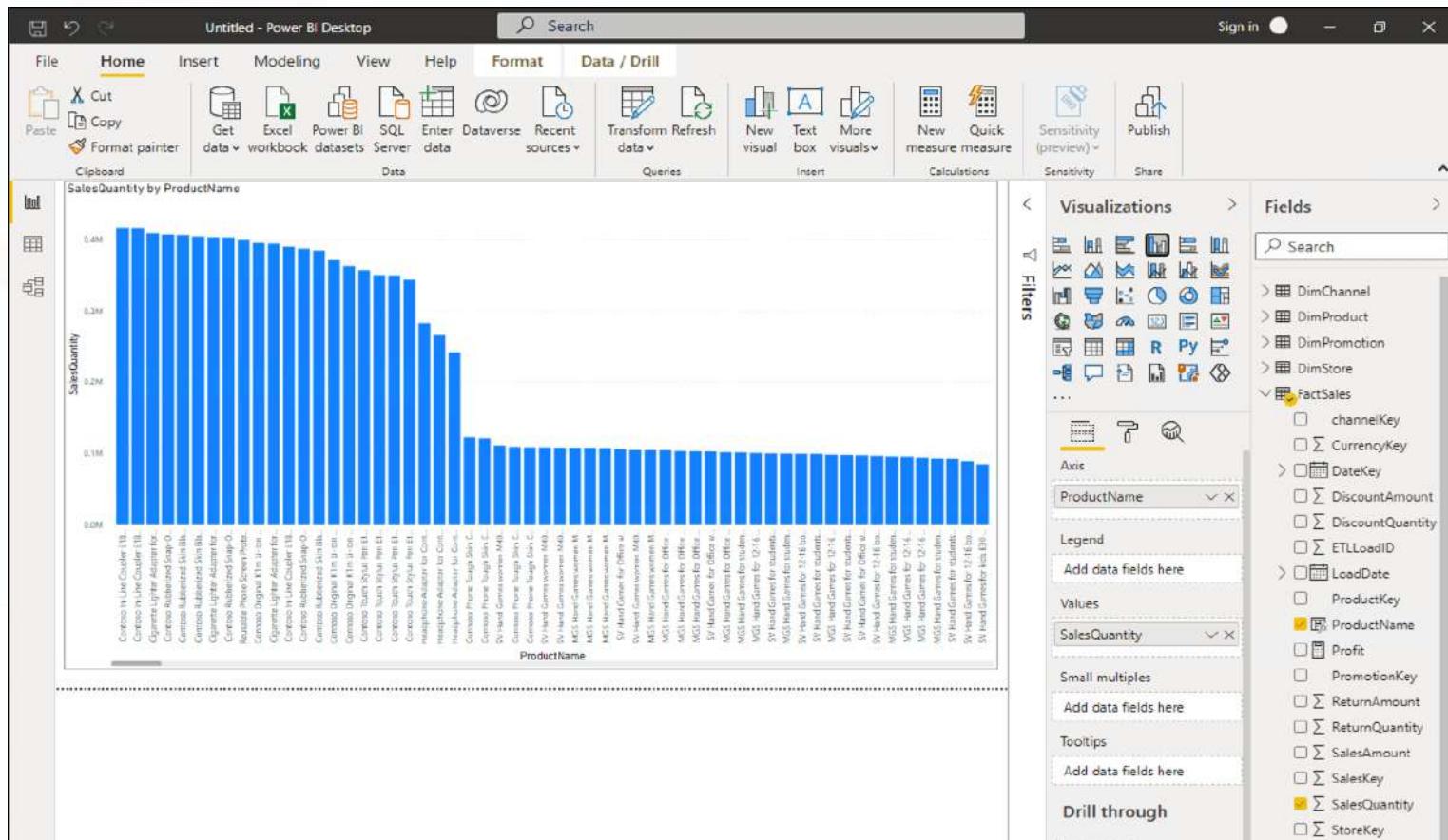
ProductName = LOOKUPVALUE(DimProduct[ProductName], DimProduct[ProductKey], FactSales[ProductKey], "Unknown")

The screenshot shows the Power BI Desktop interface with the following details:

- Top Bar:** ContosoBIDW - Power BI Desktop, Search, kc.su@auovide.onmicrosoft.com.
- Ribbon:** File, Home, Insert, Modeling, View, Help, Table tools (selected), Column tools.
- Column Tools Panel:** Structure, Format, Summarization, Data category, Sort by column, Data groups, Manage relationships, New column, Calculators.
- Formula Bar:** 1 ProductName = LOOKUPVALUE(DimProduct[ProductName], DimProduct[ProductKey], FactSales[ProductKey], "Unknown")
- Fields Pane:** Visualizations, Fields, Search. It lists tables: DimChannel, DimProduct, DimPromotion, DimStore, FactSales. Under FactSales, the ProductName field is selected.
- Bottom Area:** Build visuals with your data, Select or drag fields from the Fields pane onto the report canvas.

# Step - Analyze Data (Product) 02

- 選取 ProductName 與 SalesQuantity 兩個 Field 來建立 Clustered Column Chart Visualization



### 3. Channel 獲利分析

# Step - Analyze Data (Channel) 01

- 按下畫面底下 Page 3 標籤頁右邊的 + 圖示，建立一個新的 Page
- Fields 視窗裡面選取 FactSales Table，然後按下畫面上方的 New column 按鈕
- 輸入底下的 Formula，再按下左邊的勾勾

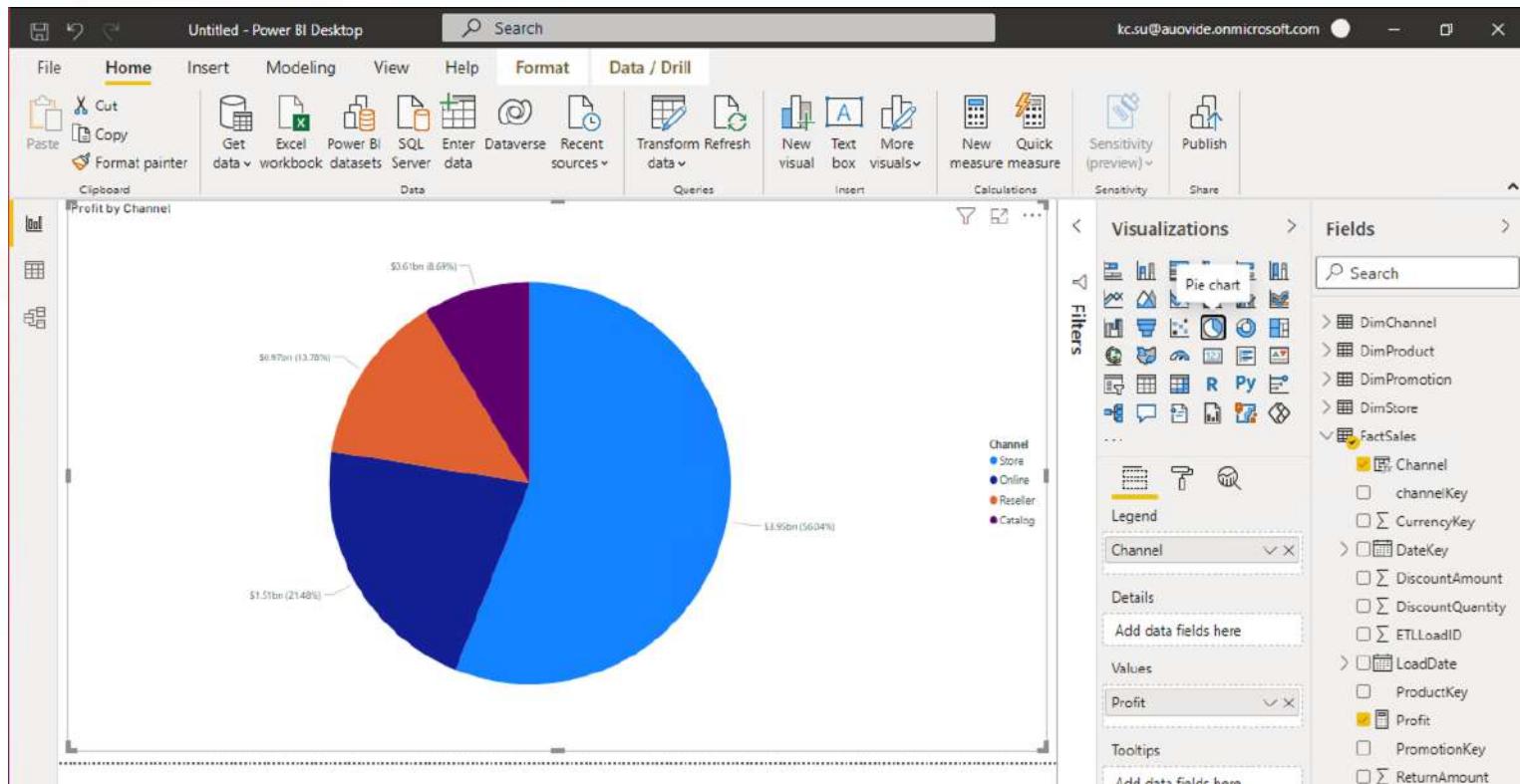
Channel = LOOKUPVALUE(DimChannel[ChannelName], DimChannel[ChannelKey], FactSales[ChannelKey], "Unknown")

The screenshot shows the Power BI Desktop interface with the following details:

- Top Bar:** Untitled - Power BI Desktop, Search, File, Home, Insert, Modeling, View, Help, Table tools (selected), Column tools.
- Column Tools Tab:** Structure, Format, Summarization, Data category, Sort by column, Data groups, Manage relationships, New column, Calculations.
- Formula Bar:** Channel = LOOKUPVALUE(DimChannel[ChannelName], DimChannel[ChannelKey], FactSales[ChannelKey], "Unknown")
- Fields Pane:** Visualizations, Fields, Search, DimChannel, DimProduct, DimPromotion, DimStore, FactSales (expanded), Channel (selected).
- Bottom Area:** Build visuals with your data, Select or drag fields from the Fields pane onto the report canvas.

# Step - Analyze Data (Channel) 02

- 選取 Channel 與 Profit 兩個 Field 來建立 Chart Visualization
- 預設會是 Table View，手動在 Visualizations 視窗改為 Pie Chart 即可



## 4. Promotion 獲利分析

# Step - Analyze Data (Promotion) 01

- 按下畫面底下 Page 4 標籤頁右邊的 + 圖示，建立一個新的 Page
- Fields 視窗裡面選取 FactSales Table，然後按下畫面上方的 New column 按鈕
- 輸入底下的 Formula，再按下左邊的勾勾

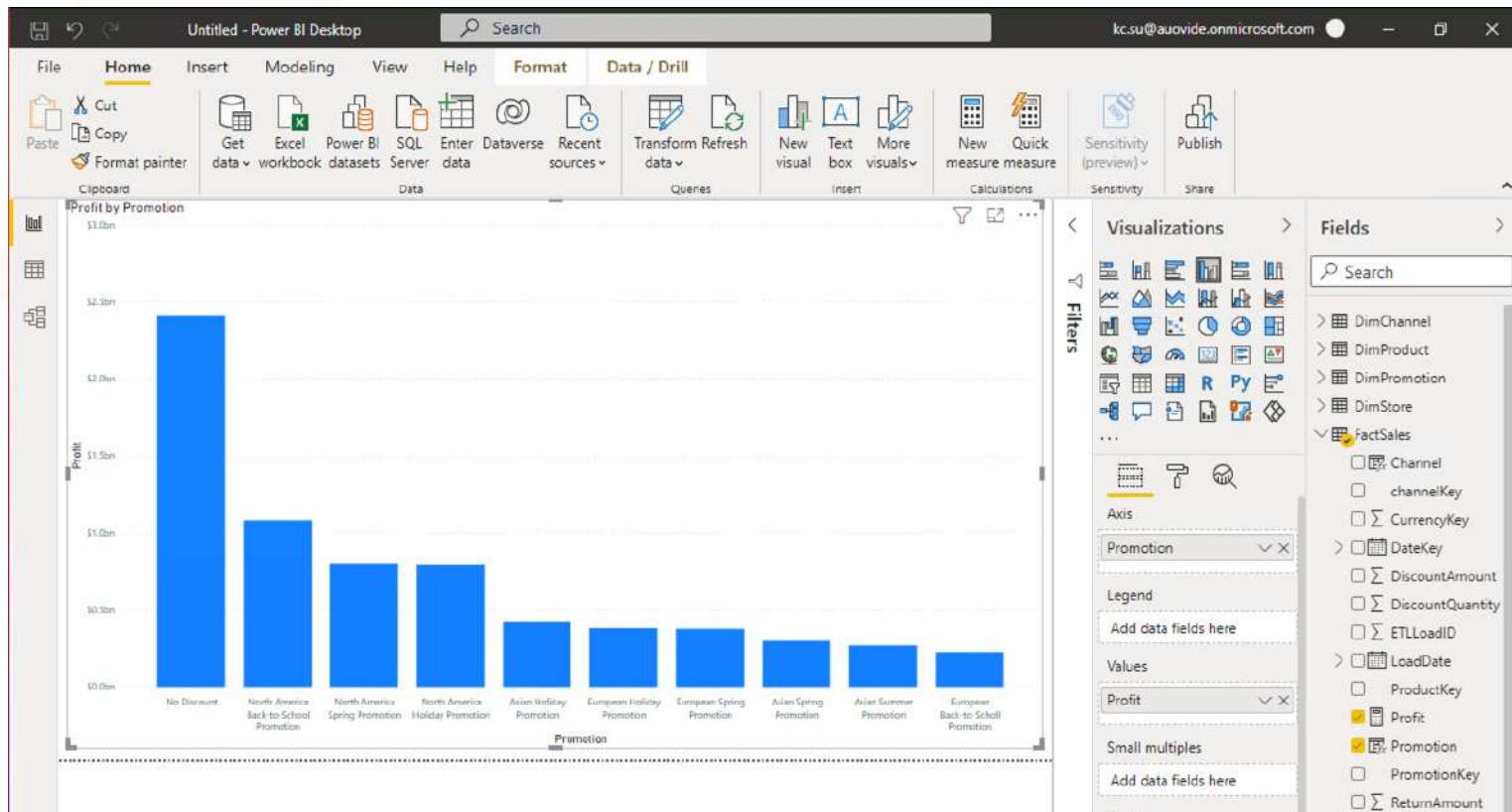
Promotion=LOOKUPVALUE(DimPromotion[PromotionName], DimPromotion[PromotionKey], FactSales[PromotionKey],

The screenshot shows the Power BI Desktop interface with the following details:

- Top Bar:** Untitled - Power BI Desktop, Search, kc.su@auvide.onmicrosoft.com.
- Home Tab:** Selected, showing options like Paste, Cut, Copy, Format painter, Get data, Excel, Power BI, SQL, Enter data, Dataverse, Recent sources, Transform, Refresh data, New visual, Text box, More visuals, New measure, Quick measure, Sensitivity (preview), Publish, and Share.
- Left Sidebar:** Displays icons for List, Grid, and Canvas.
- Middle Area:** Text "Build visuals with your data" and "Select or drag fields from the Fields pane onto the report canvas." Below this is a small icon illustrating data flow between two boxes.
- Right Side:**
  - Visualizations:** A grid of visualization icons.
  - Fields:** A list of fields categorized by table:
    - DimChannel
    - DimProduct
    - DimPromotion
    - DimStore
    - FactSales** (selected):
      - Channel
      - channelKey
      - CurrencyKey
      - DateKey
      - DiscountAmount
      - DiscountQuantity
      - ETLLoadID
      - LoadDate
      - ProductKey
      - Profit
      - Promotion
      - PromotionKey
      - ReturnAmount
      - ReturnQuantity
  - Filters:** Options for Drill through, Cross-report, Off, Keep all filters, and On.

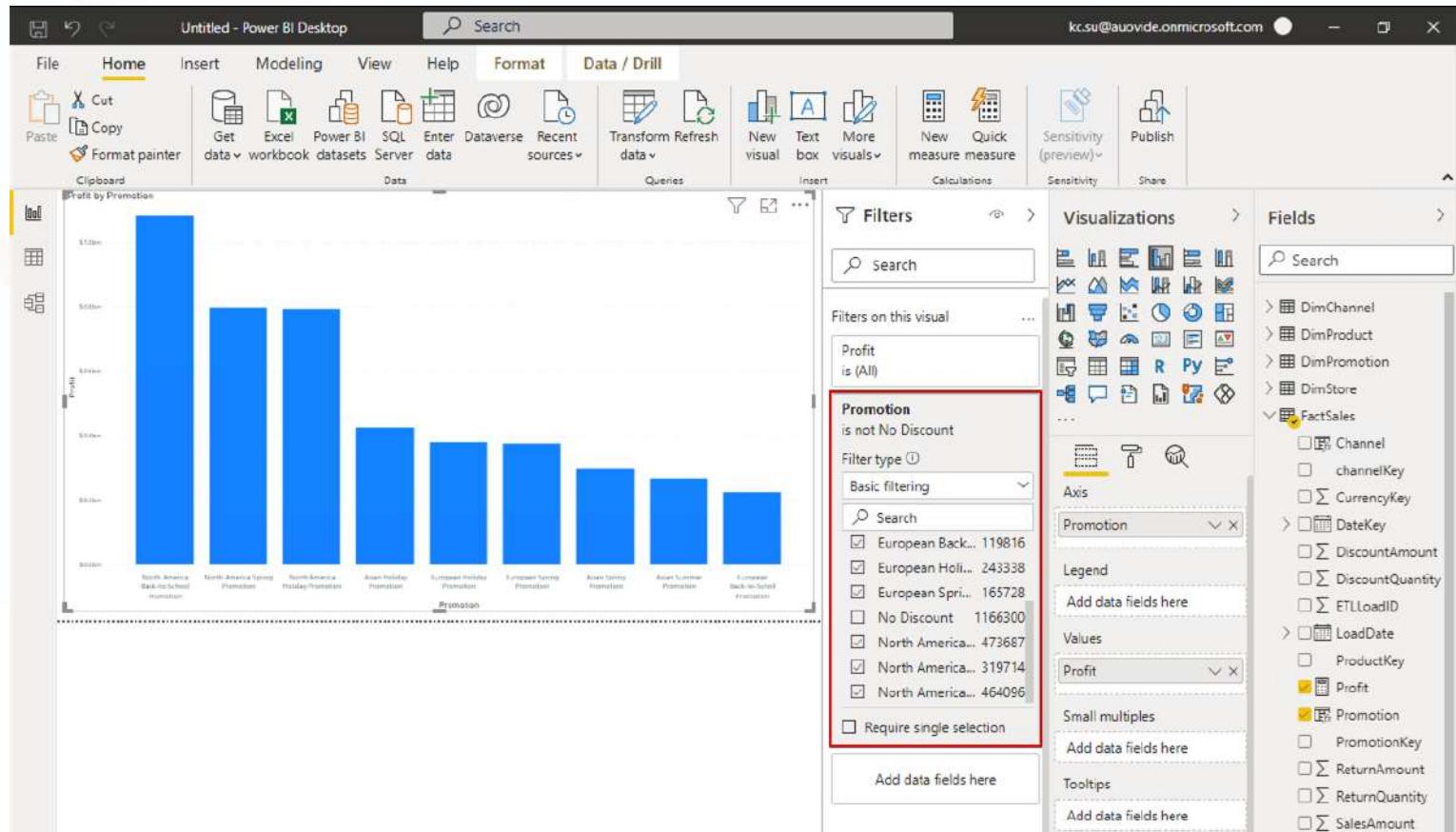
# Step - Analyze Data (Promotion) 02

- 選取 Promotion 與 Profit 兩個 Field 來建立 Clustered Column Chart Visualization



# Step - Analyze Data (Promotion) 03

- Filters 視窗裡頭，選取 Promotion is (All)，Filter type 選取 Basic filtering，然後不要勾選 No Discount



# Step - Analyze Data (Promotion) 04

- Clustered Column Chart 馬上就反應出 Filter 的效果

