Power Platform Dataflows

Power Query (and more) in the Cloud

Wolfgang Strasser
wolfgang@powerofbi.at
@wstrasser
workingondata.wordpress.com

Microsoft® MVP Most Valuable Professional

UBIDO
AN ACP GROUP COMPANY

Power Platform

Powerful alone. Better together.

The Microsoft Power Platform is more than the sum of its parts. Connect them together—and to Office 365, Dynamics 365, Azure, and hundreds of other apps—and build end-to-end business solutions.

Power BI

Make informed, confident business decisions by putting data-driven insights into everyone's hands.

Learn more ›

Power Apps

Turn ideas into organizational solutions by enabling everyone to build custom apps that solve business challenges.

Learn more ›

Power Automate

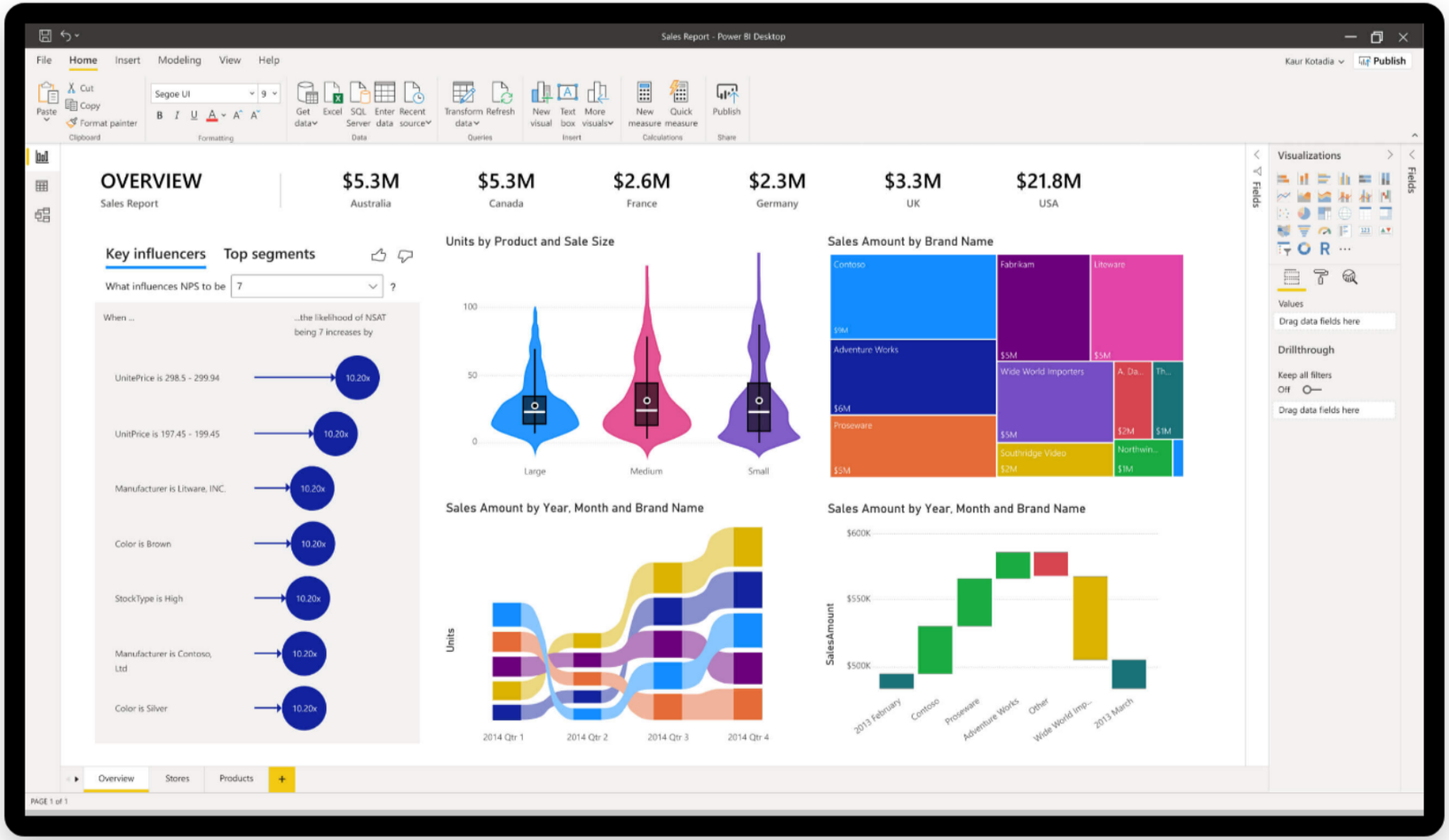
Boost business productivity to get more done by giving everyone the ability to automate organizational processes.

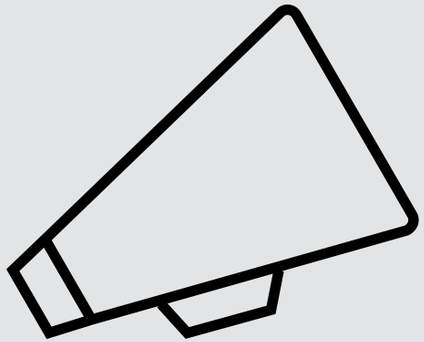
Learn more ›

Power Virtual Agents

Easily build chatbots to engage conversationally with your customers and employees—no coding required.

Learn more ›





Power Platform Dataflows

=

Self-Service

Data Integration in the Cloud

Information Hierarchy

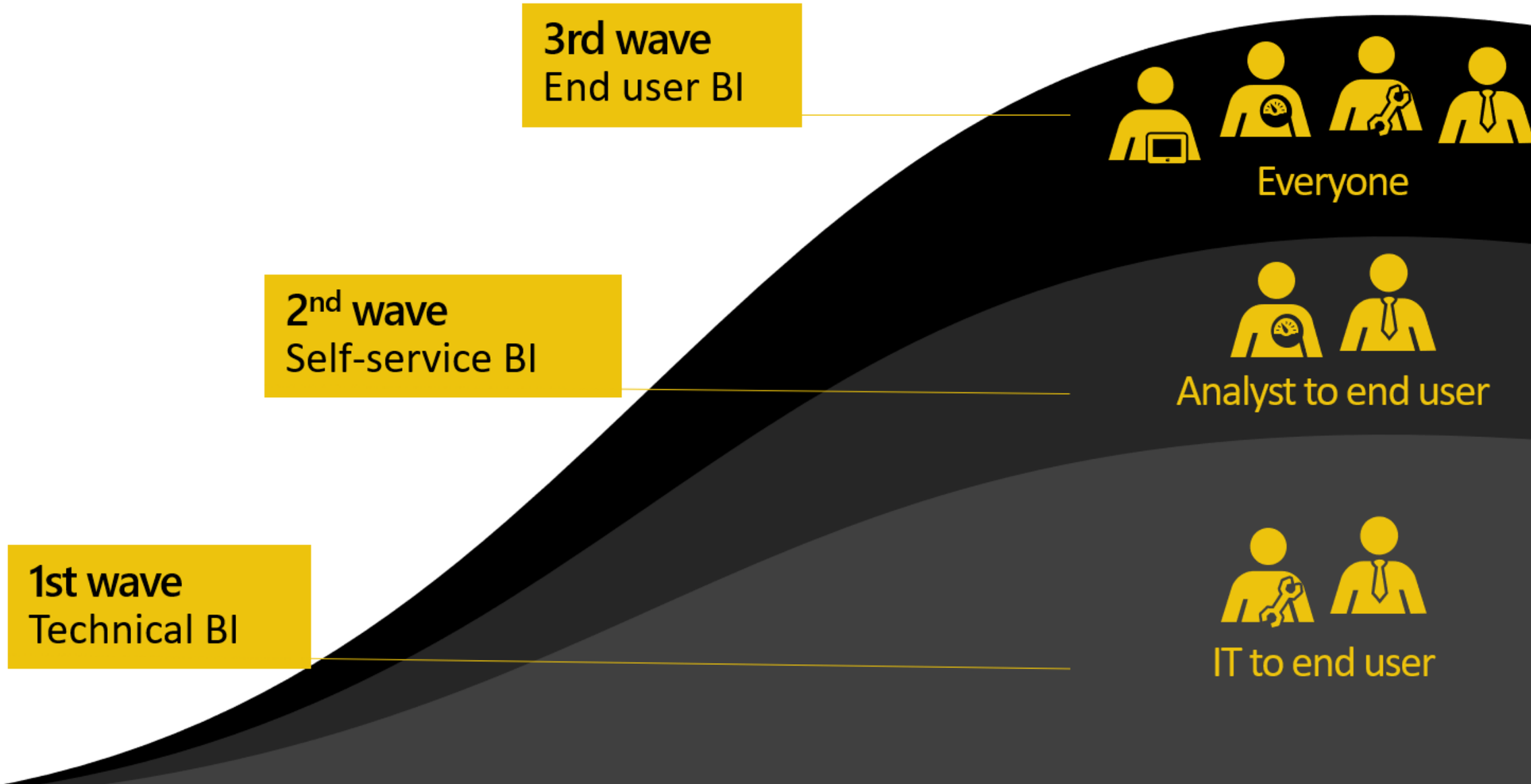
Reusability

Data Lake

Common Data Model (CDM)

Power Query (Online)

Evolution of Business Intelligence

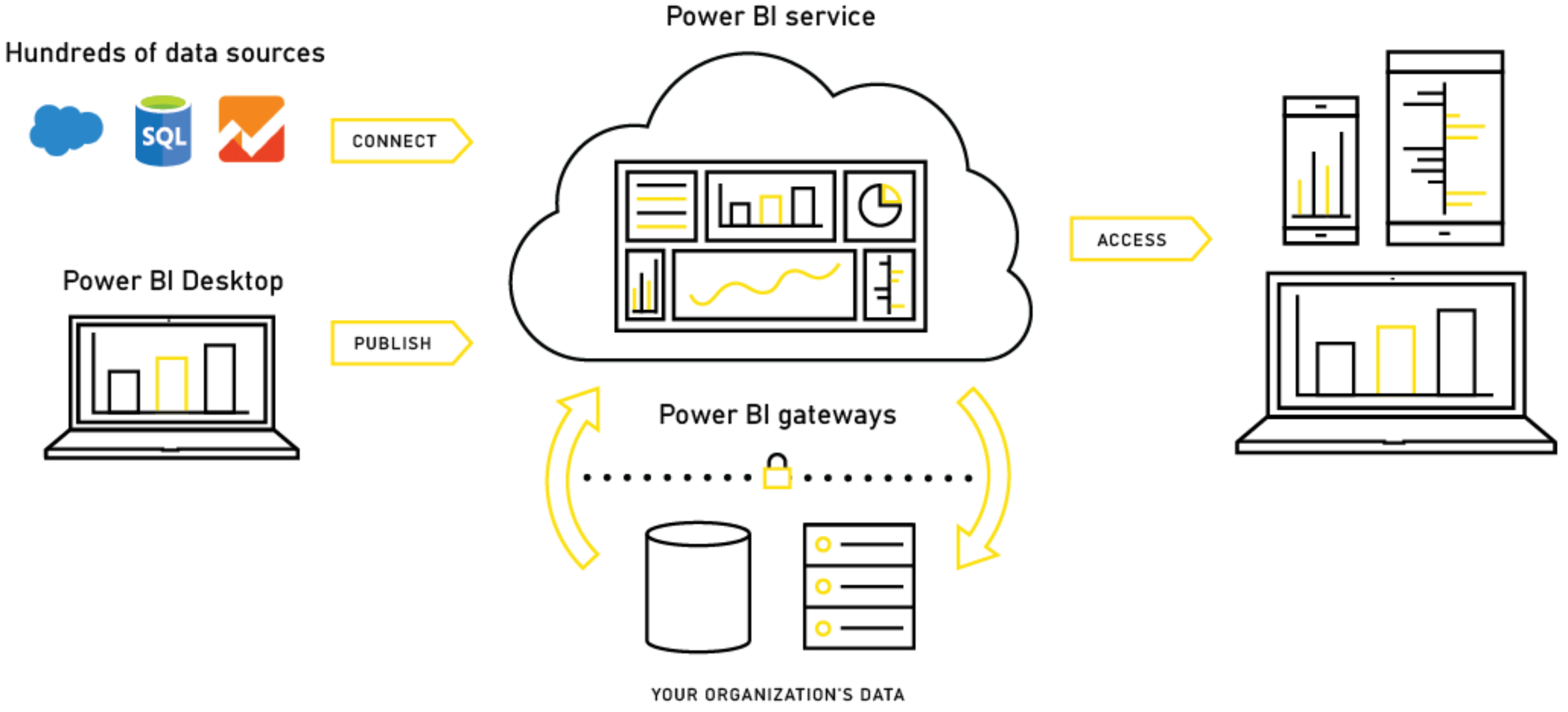


Self Service BI

BI / Reporting ✓

ETL ✗ ✓

IT for innovators.
Analytics & Development


Power BI service
Hundreds of data sources
SQL
CONNECT
Power BI Desktop
PUBLISH
ACCESS
Power BI gateways
YOUR ORGANIZATION'S DATA

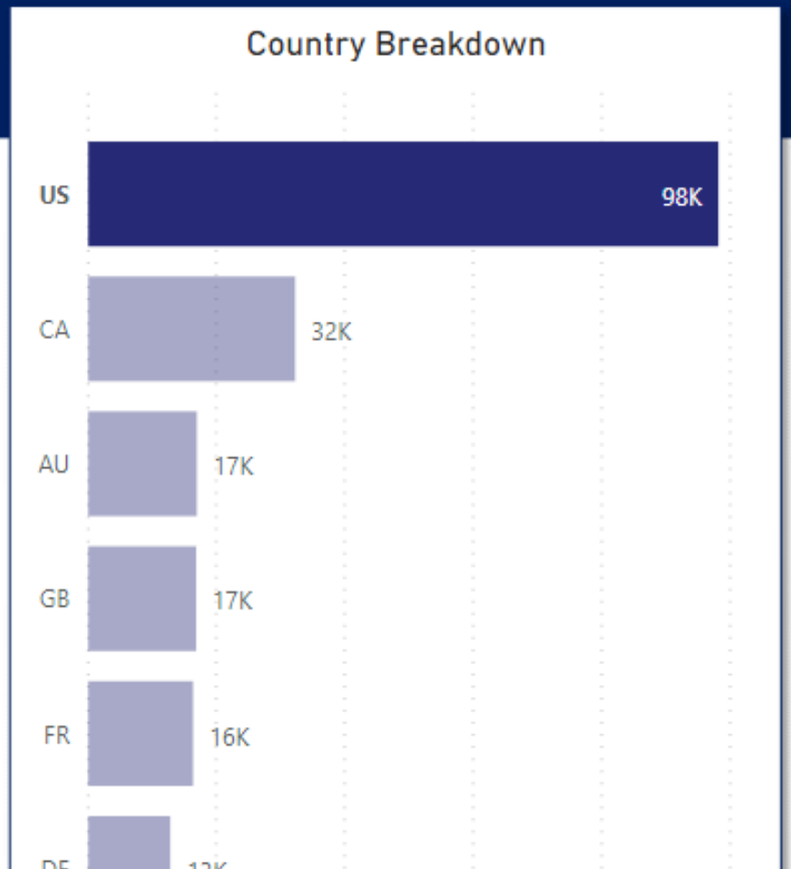
Interactive Reporting

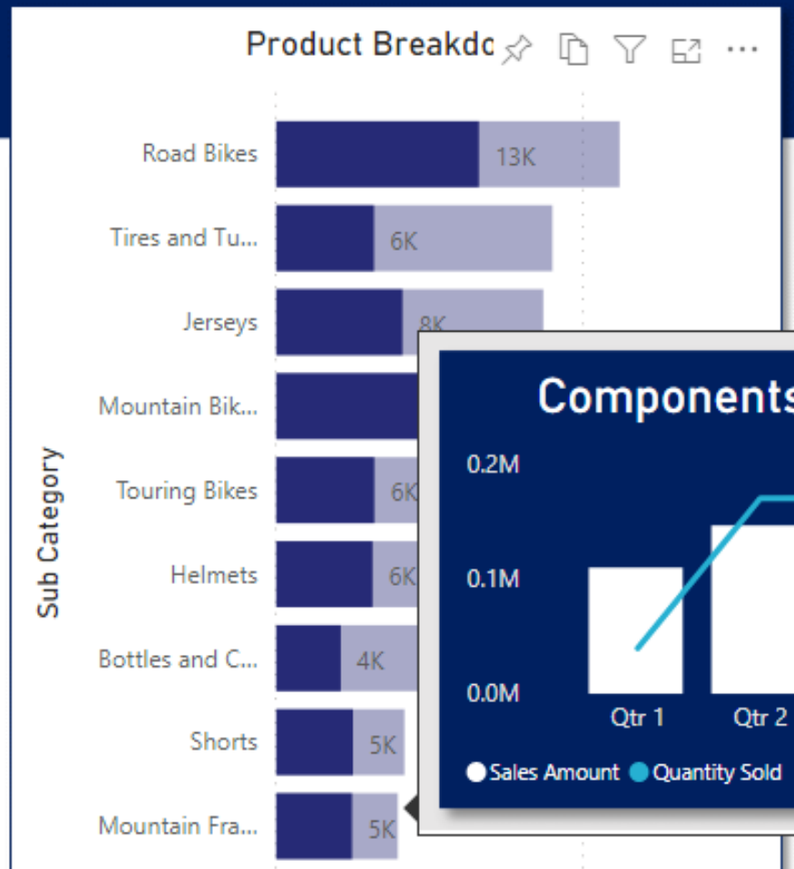
Sales Report

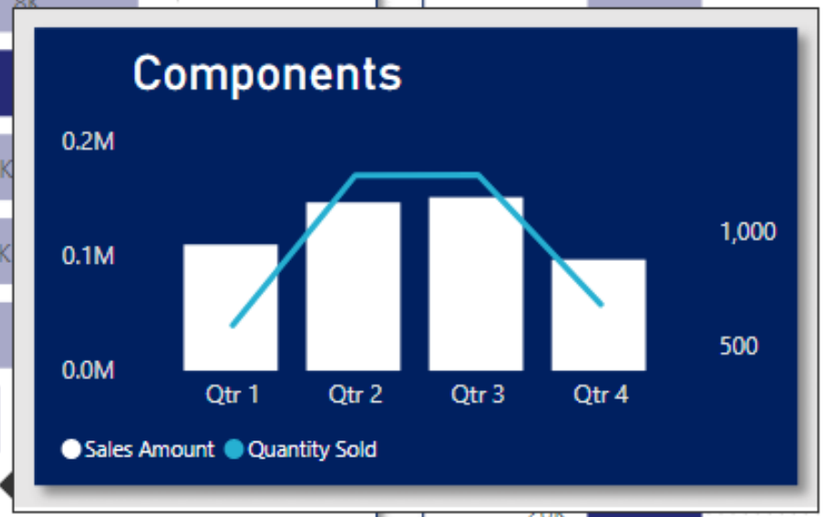
CUBIDO
AN ACP GROUP COMPANY

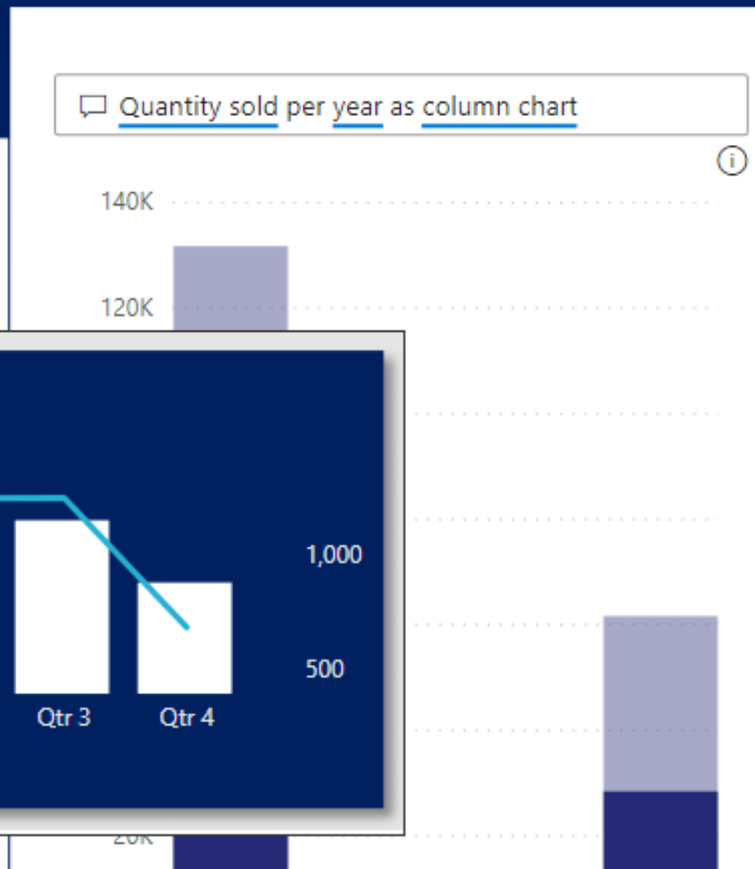
98K
Units Sold

14.79M
Sales Amount

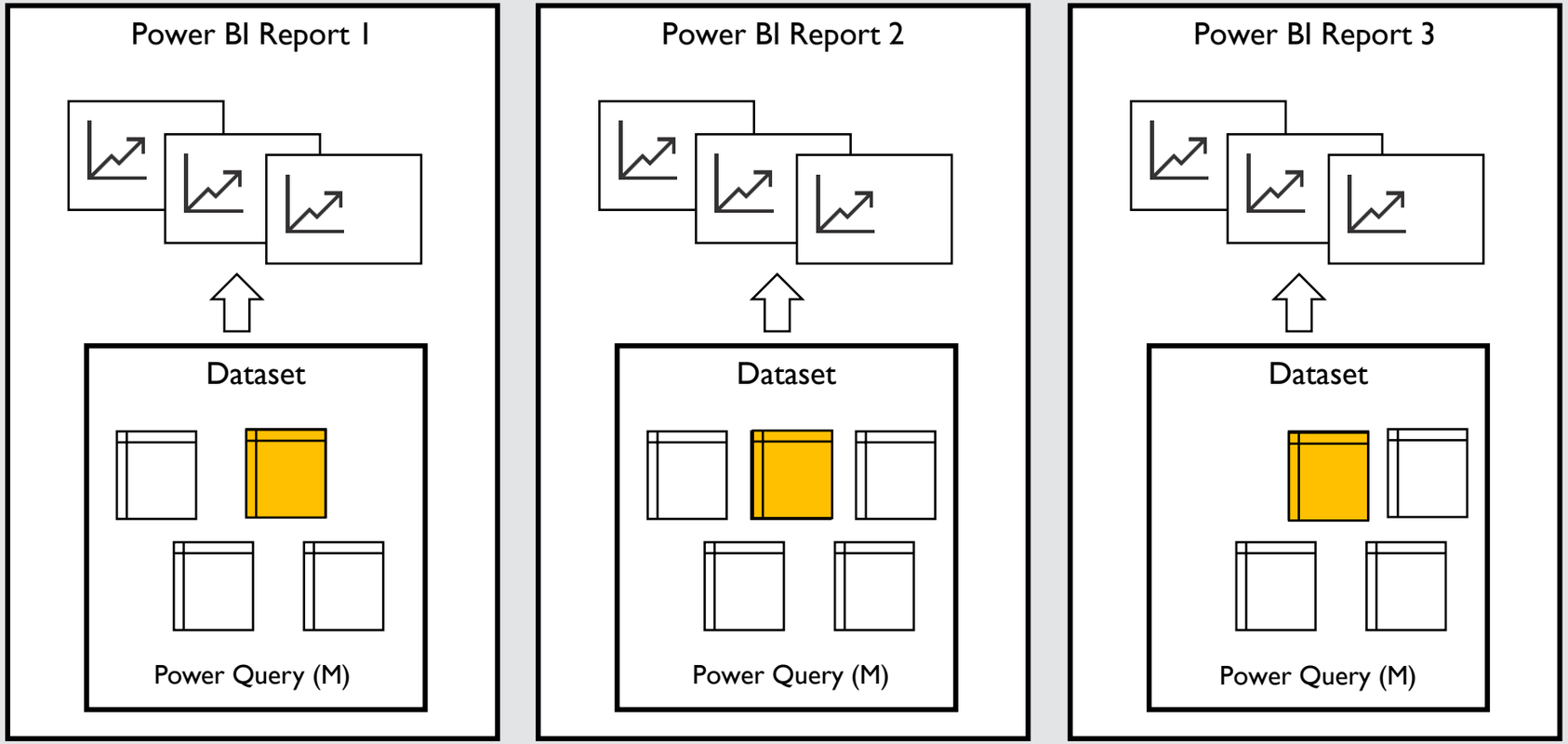




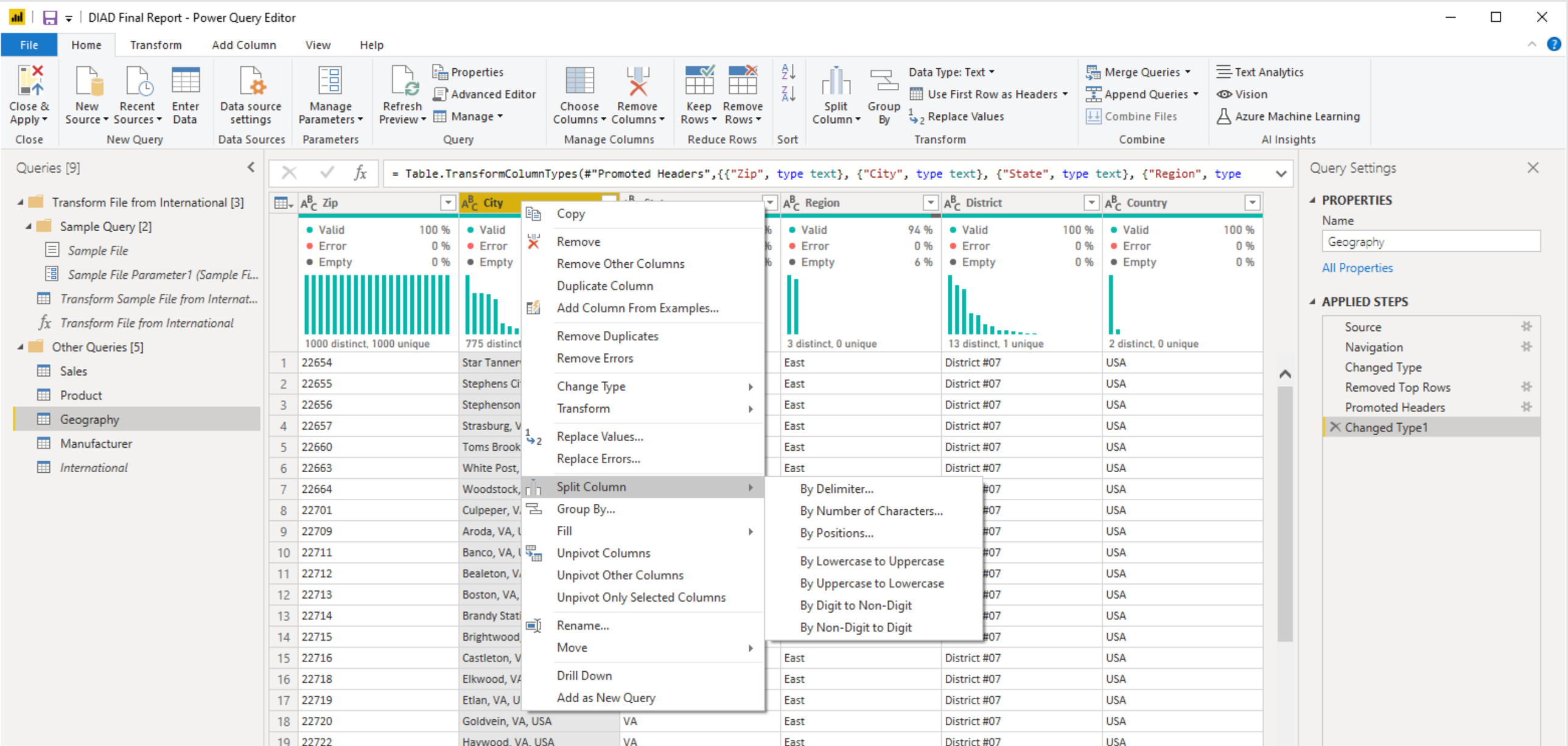




(Reusable) ETL in Power BI

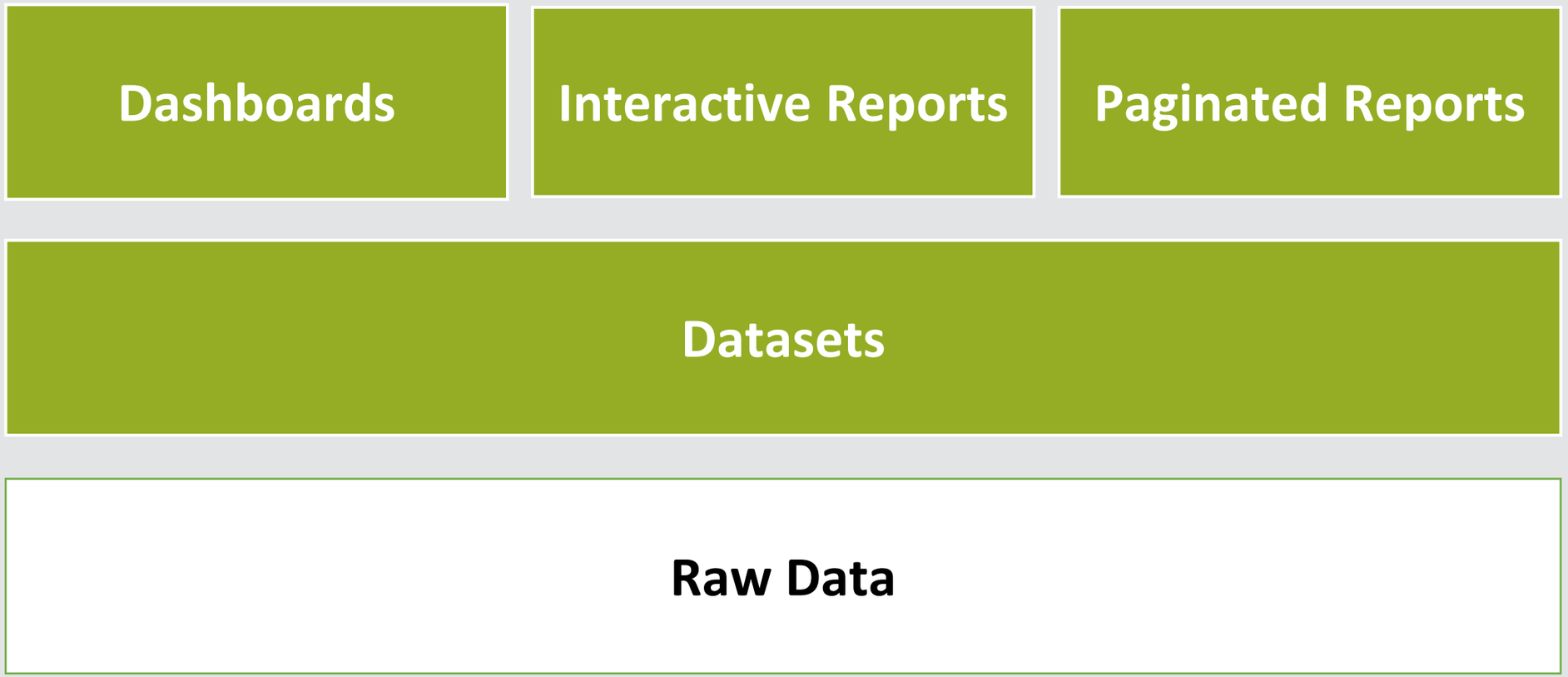


Power Query

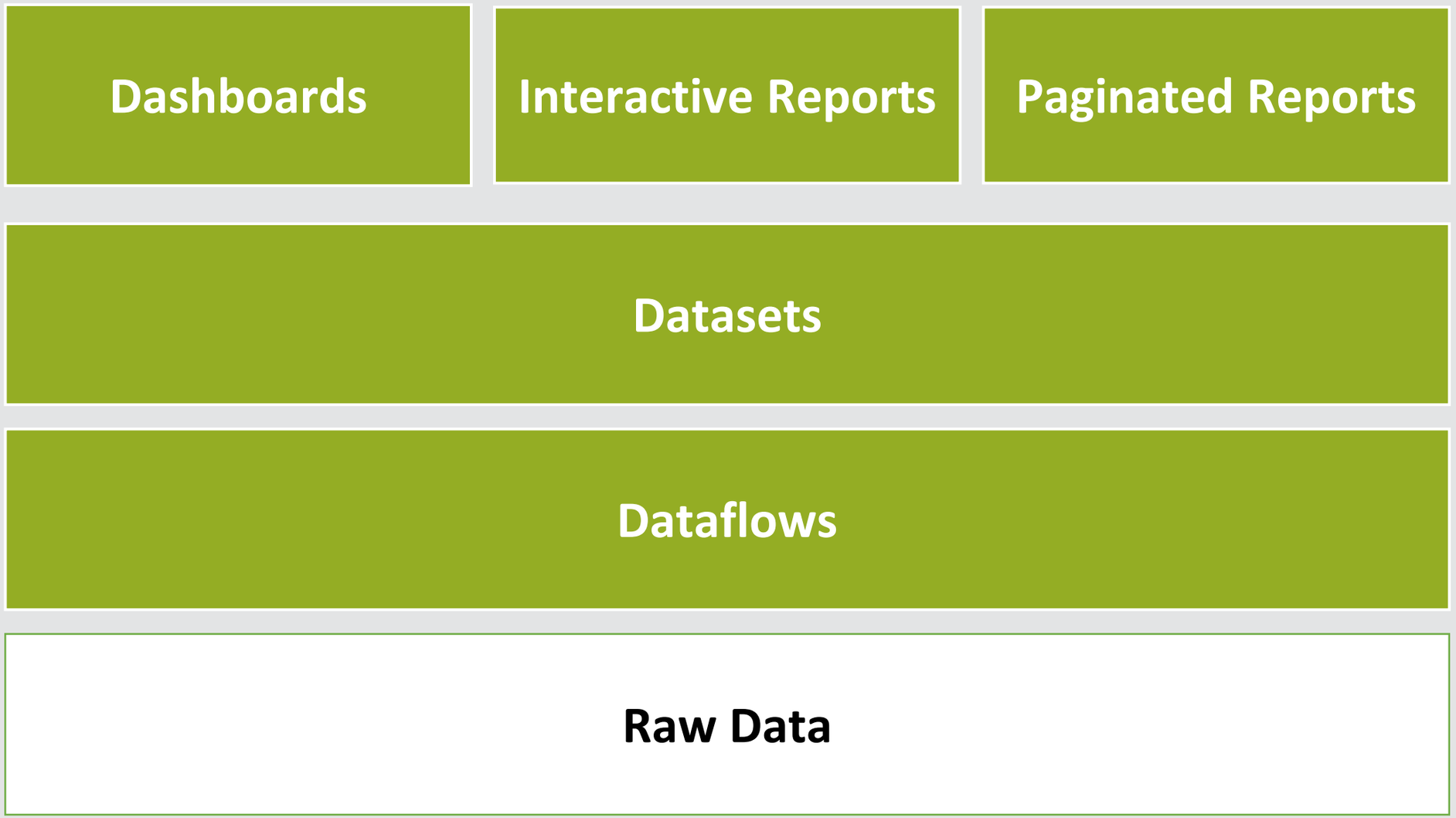

DIAD Final Report - Power Query Editor
File
Home
Transform
Add Column
View
Help
Close & Apply
New Source
Recent Sources
Enter Data
Data source settings
Manage Parameters
Refresh Preview
Properties
Advanced Editor
Manage
Choose Columns
Remove Columns
Keep Rows
Remove Rows
Split Column
Group By
Data Type: Text
Use First Row as Headers
Replace Values
Merge Queries
Append Queries
Combine Files
Text Analytics
Vision
Azure Machine Learning
Close
New Query
Data Sources
Parameters
Query
Manage Columns
Reduce Rows
Sort
Transform
Combine
AI Insights
Queries [9]
Transform File from International [3]
Sample Query [2]
Sample File
Sample File Parameter1 (Sample Fi...
Transform Sample File from Internat...
Transform File from International
Other Queries [5]
Sales
Product
Geography
Manufacturer
International
= Table.TransformColumnTypes(#"Promoted Headers",{{"Zip", type text}, {"City", type text}, {"State", type text}, {"Region", type
Zip
City
Region
District
Country
Valid 100 %
Error 0 %
Empty 0 %
1000 distinct, 1000 unique
775 distinct
Valid 94 %
Error 0 %
Empty 6 %
3 distinct, 0 unique
Valid 100 %
Error 0 %
Empty 0 %
13 distinct, 1 unique
Valid 100 %
Error 0 %
Empty 0 %
2 distinct, 0 unique
1 22654 East District #07 USA
2 22655 East District #07 USA
3 22656 East District #07 USA
4 22657 East District #07 USA
5 22660 East District #07 USA
6 22663 East District #07 USA
7 22664 #07 USA
8 22701 #07 USA
9 22709 #07 USA
10 22711 #07 USA
11 22712 #07 USA
12 22713 #07 USA
13 22714 #07 USA
14 22715 #07 USA
15 22716 East District #07 USA
16 22718 East District #07 USA
17 22719 East District #07 USA
18 22720 Goldvein, VA, USA VA East District #07 USA
19 22722 Haywood, VA, USA VA East District #07 USA
Copy
Remove
Remove Other Columns
Duplicate Column
Add Column From Examples...
Remove Duplicates
Remove Errors
Change Type
Transform
Replace Values...
Replace Errors...
Split Column
Group By...
Fill
Unpivot Columns
Unpivot Other Columns
Unpivot Only Selected Columns
Rename...
Move
Drill Down
Add as New Query
By Delimiter...
By Number of Characters...
By Positions...
By Lowercase to Uppercase
By Uppercase to Lowercase
By Digit to Non-Digit
By Non-Digit to Digit
Query Settings
PROPERTIES
Name
Geography
All Properties
APPLIED STEPS
Source
Navigation
Changed Type
Removed Top Rows
Promoted Headers
Changed Type1

Power BI Dataflows

Power BI Information Hierarchy (before Dataflows)



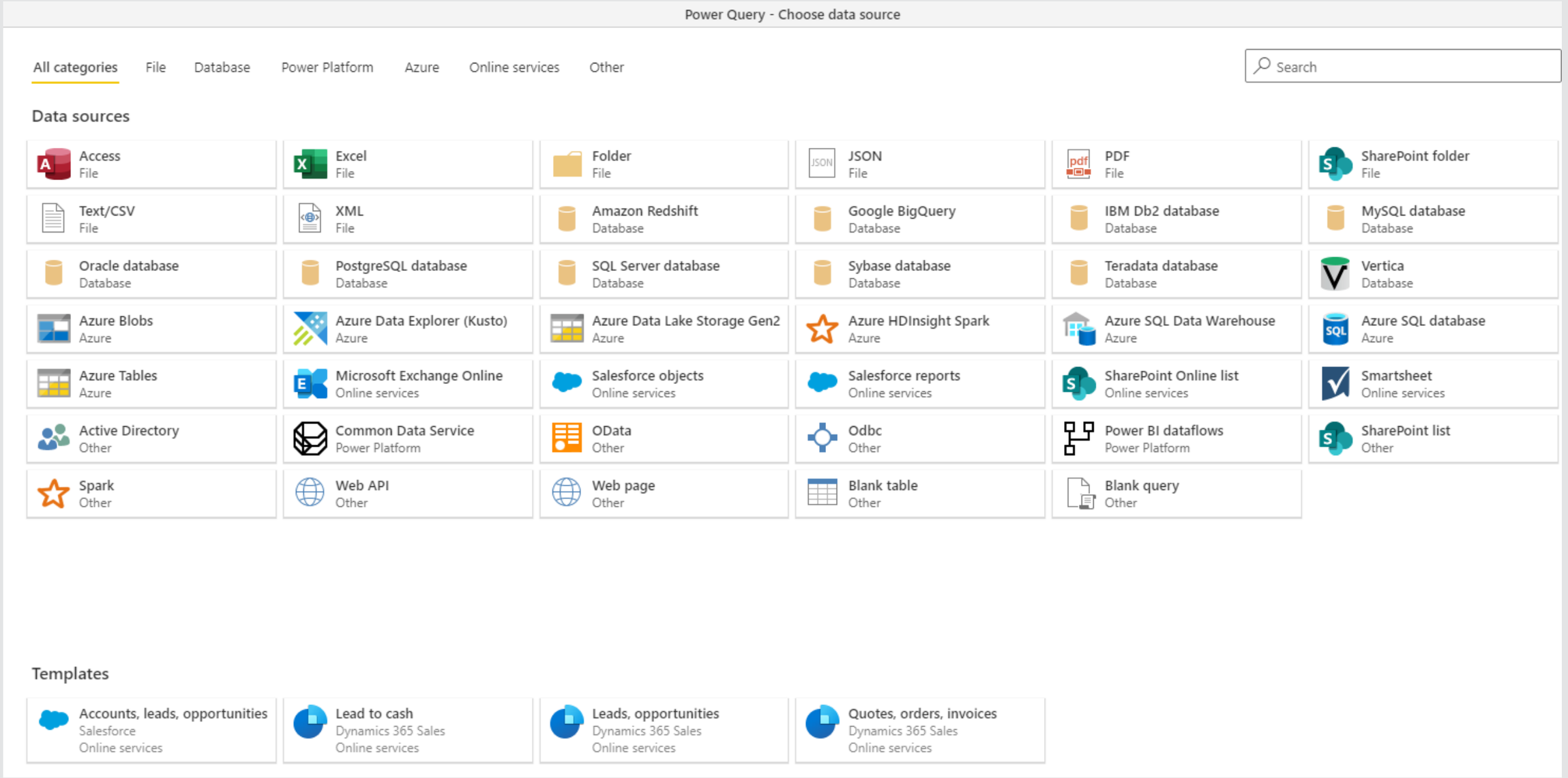
Power BI Information Hierarchy

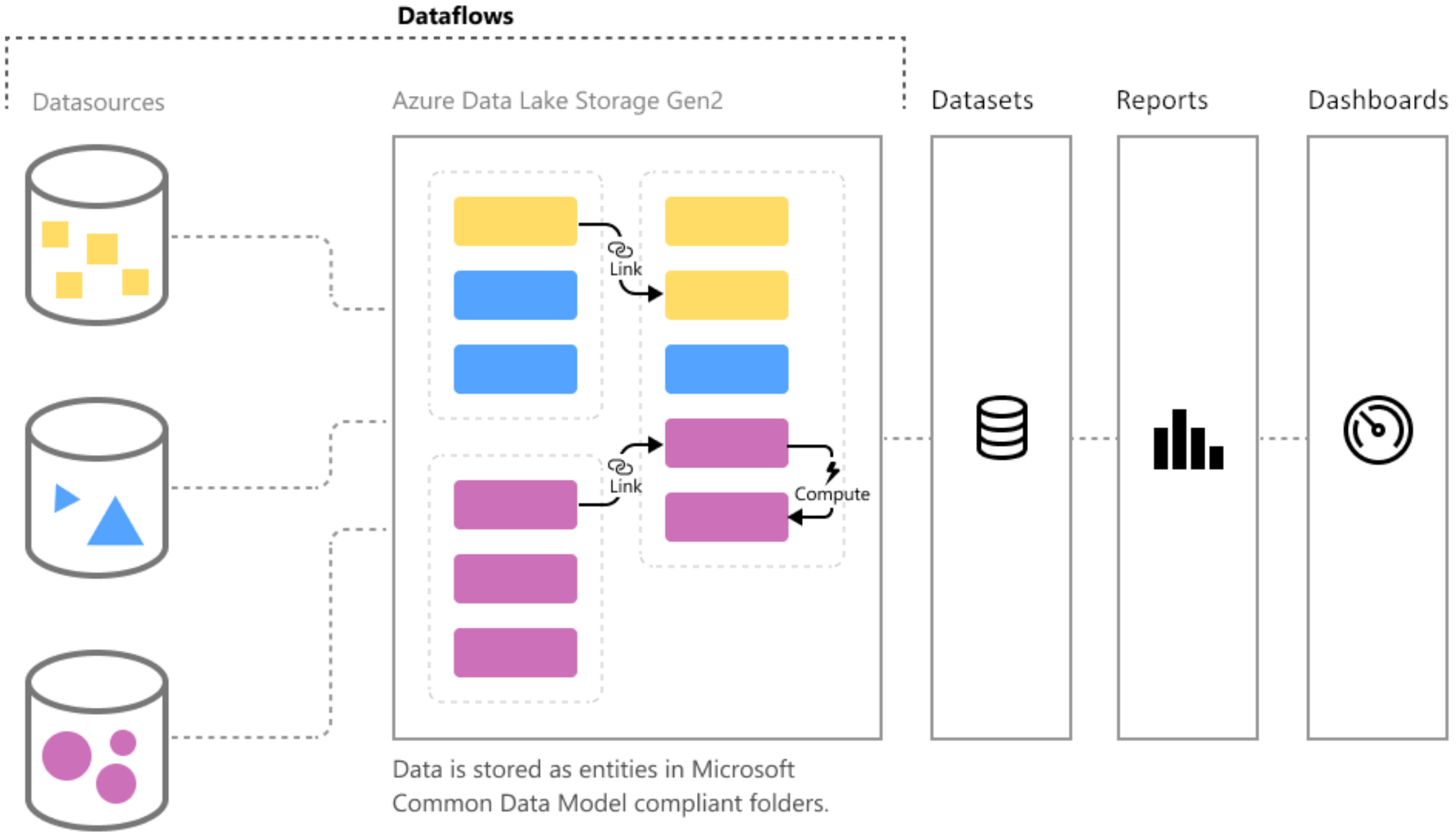


DEMO

Power BI Dataflows

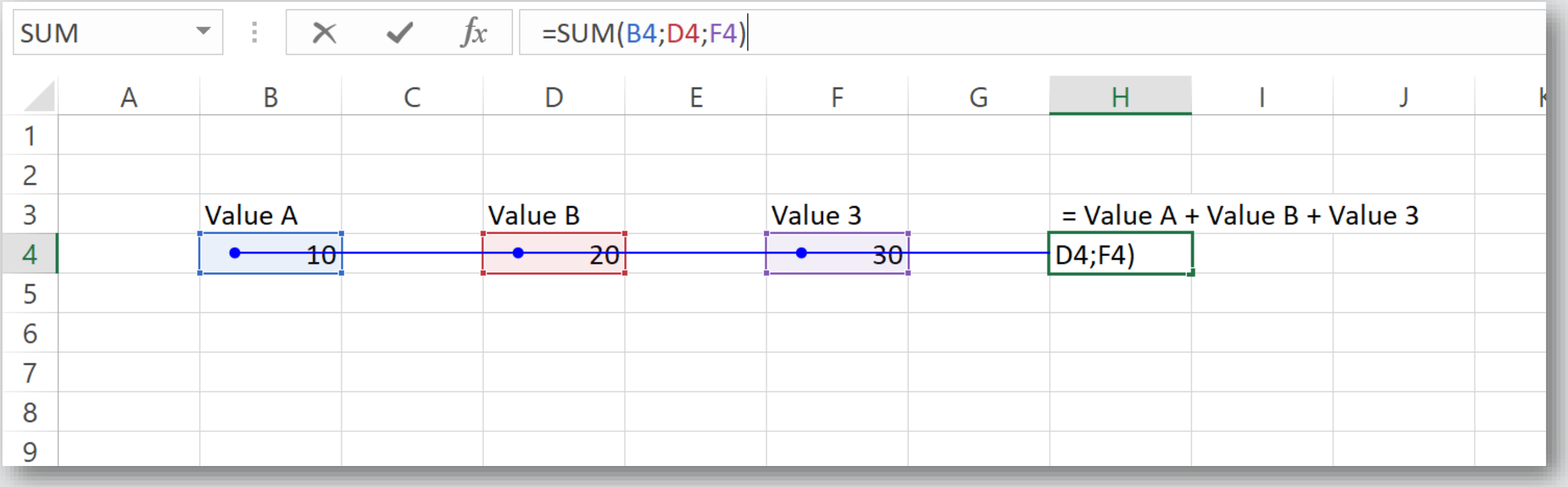
Power Query Online





https://docs.microsoft.com/en-us/power-bi/service-dataflows-overview

Dataflows are like Excel



Common Data Model

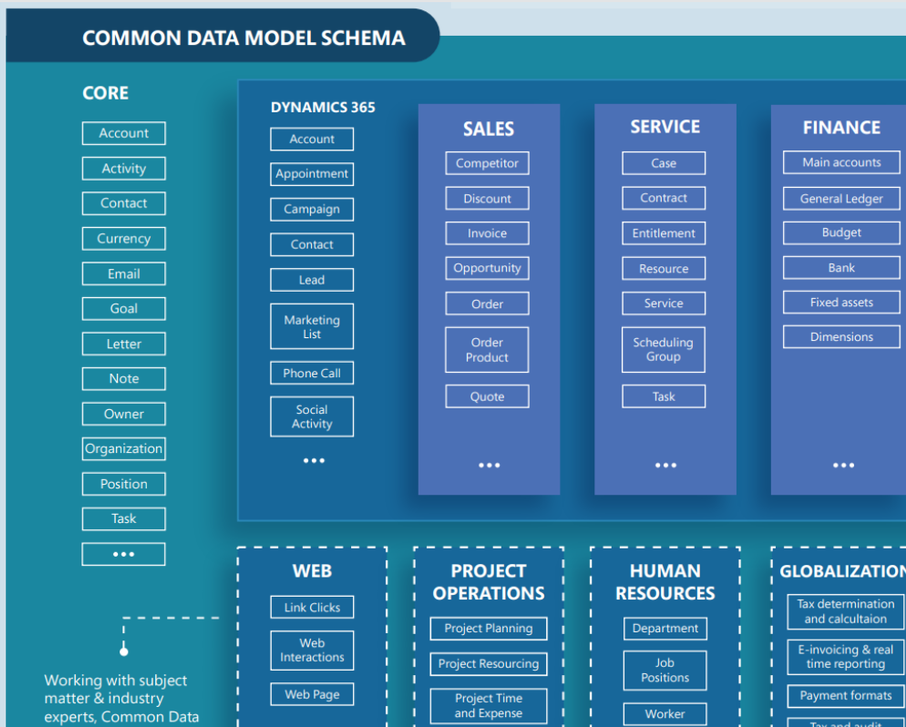
What's in Common Data Model?

In addition to the metadata system, Common Data Model includes a set of standardized, extensible data schemas that Microsoft and its partners have published. This collection of predefined schemas includes entities, attributes, semantic metadata, and relationships. The schemas represent commonly used concepts and activities, such as **Account** and **Campaign**, to simplify the creation, aggregation, and analysis of data. This graphic shows some elements of the standard entities. More information: [Common Data Model repository on GitHub]

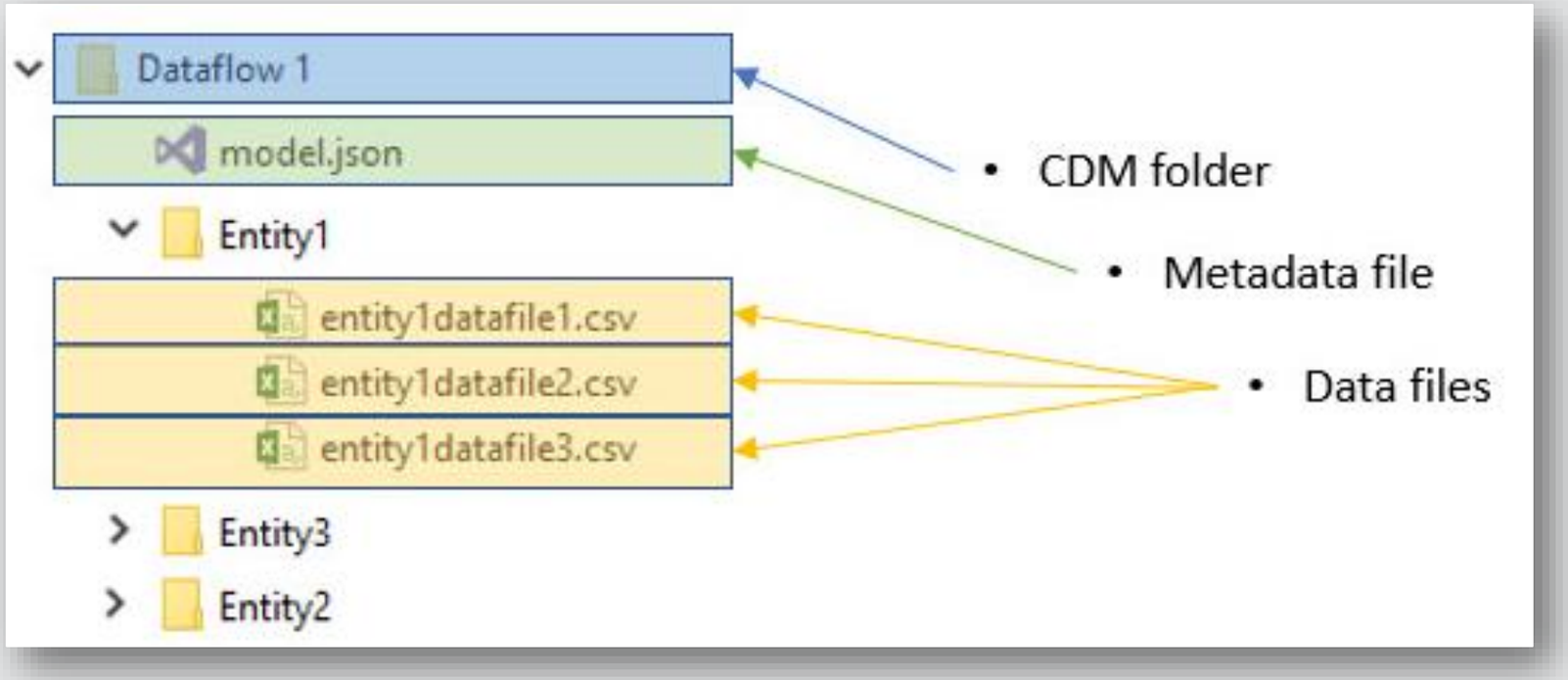
Set of

- standardized, extensible data schemas

- Entities
- Attributes
- Semantic Metadata & Relationships



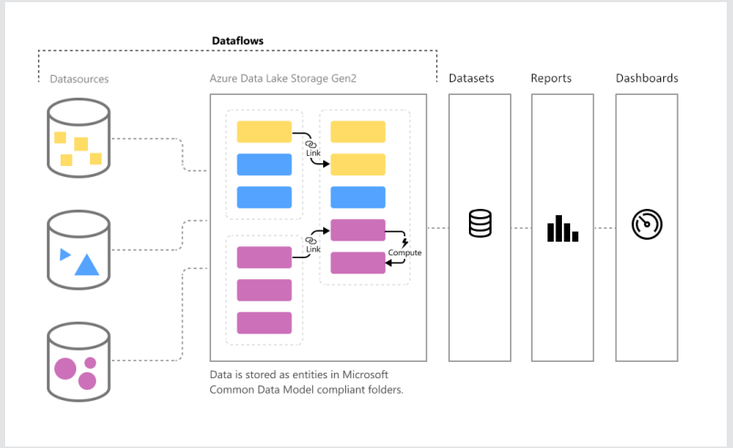
CDM Folders



Sales From Excel.json ● First Dataflow.json ●

```
{
    "name": "Sales From Excel",
    "description": "",
    "version": "1.0",
    "culture": "en-US",
    "pbi:mashup": {
        "fastCombine": true,
        "allowNativeQueries": false,
        "queriesMetadata": {
            "Sales": {
                "queryId": "40b69972-542b-49ee-ab
                "queryName": "Sales",
                "loadEnabled": true
            }
        },
        "document": "section Section1;\r\nshared
    },
    "entities": [
        {
            "$type": "LocalEntity",
            "name": "Sales",
            "description": "",
            "pbi:refreshPolicy": {
                "$type": "FullRefreshPolicy",
                "location": "Sales.csv"
            },
            "attributes": [
                {
                    "name": "SalesOrderNumber",
                    "dataType": "string"
```

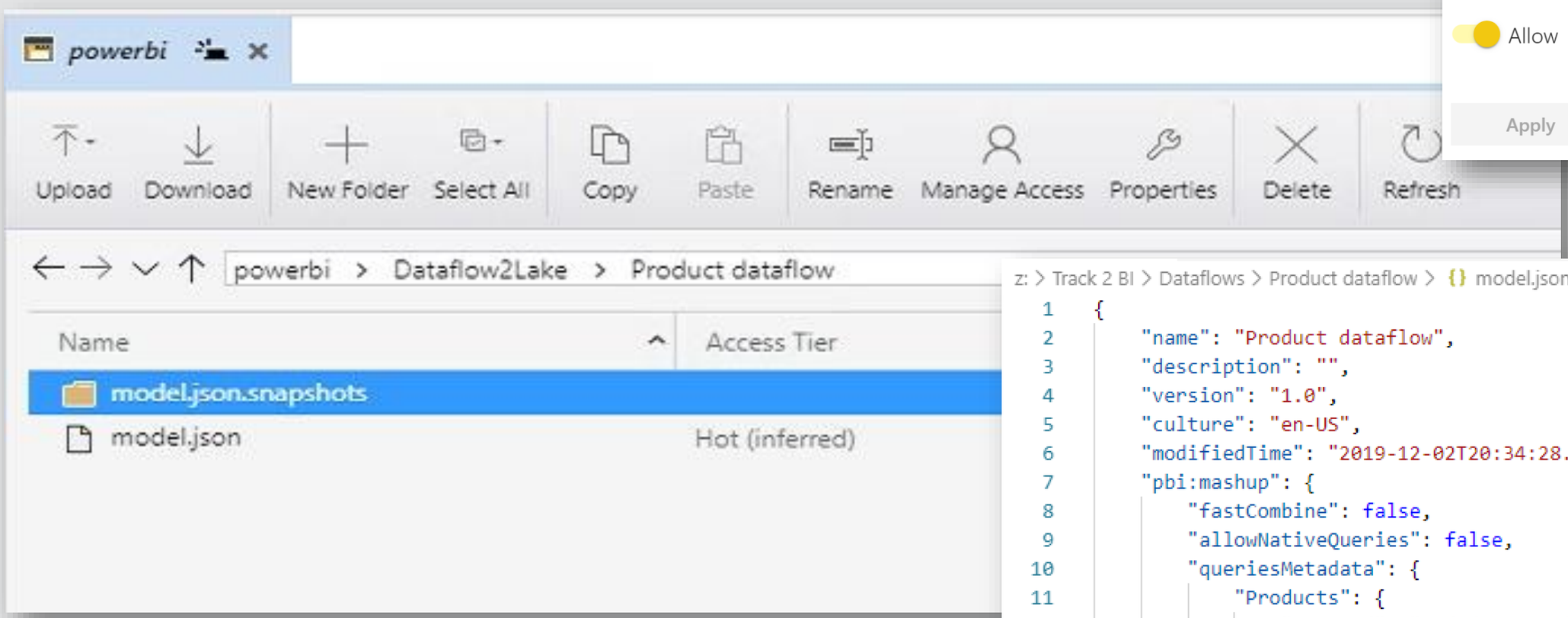
Fact Sheet – PBI Dataflows

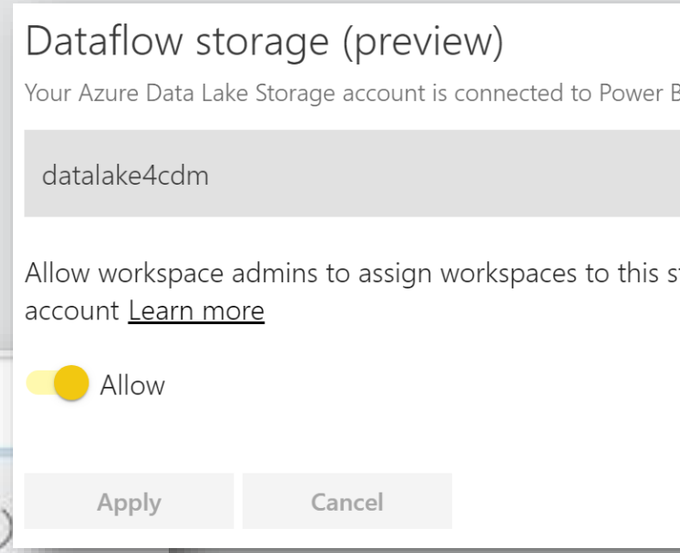


Power Query processes **in the cloud**

- .. run **independently of a Power BI report**
- .. store **output into a data lake** (ADLS Gen2)
- .. **entities** (result tables)
- .. use **Common Data Model** for entity definition
- .. not in “my workspace”

Bring your Own Storage Account (BYOSA)

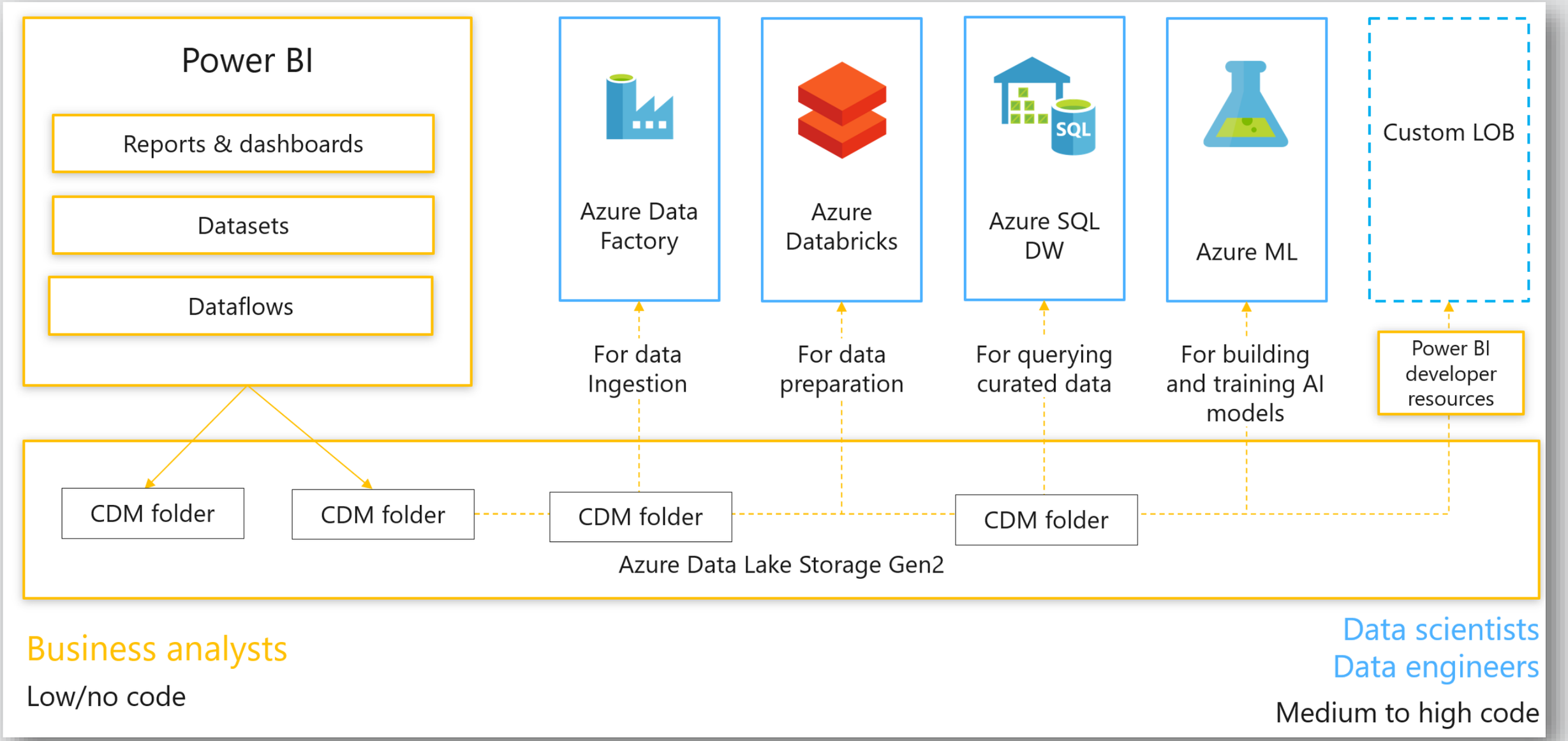




z: > Track 2 BI > Dataflows > Product dataflow > {} model.json > [ ] entities > {} 0 > [ ] attributes > {}

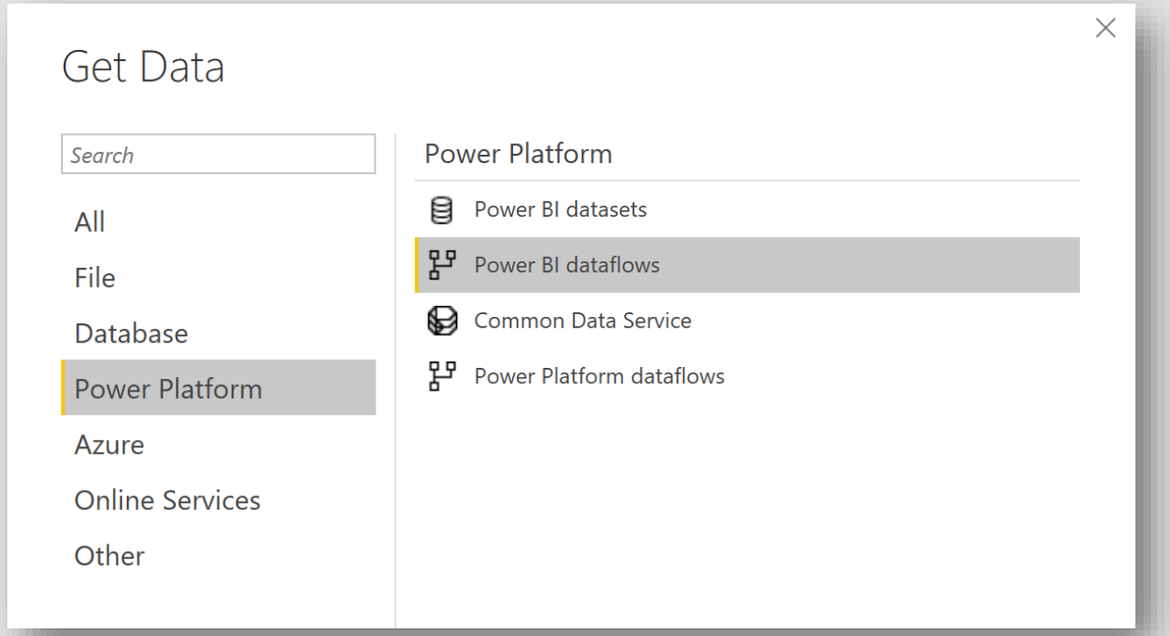
```
{
    "name": "Product dataflow",
    "description": "",
    "version": "1.0",
    "culture": "en-US",
    "modifiedTime": "2019-12-02T20:34:28.5677753+00:00",
    "pbi:mashup": {
        "fastCombine": false,
        "allowNativeQueries": false,
        "queriesMetadata": {
            "Products": {
                "queryId": "eed1d3a0-543f-48fd-bd2a-e5938d922b08",
                "queryName": "Products",
                "loadEnabled": true
            }
        },
        "document": "section Section1;\r\nshared Products = let\r\n  Sour
    },
    "entities": [
```

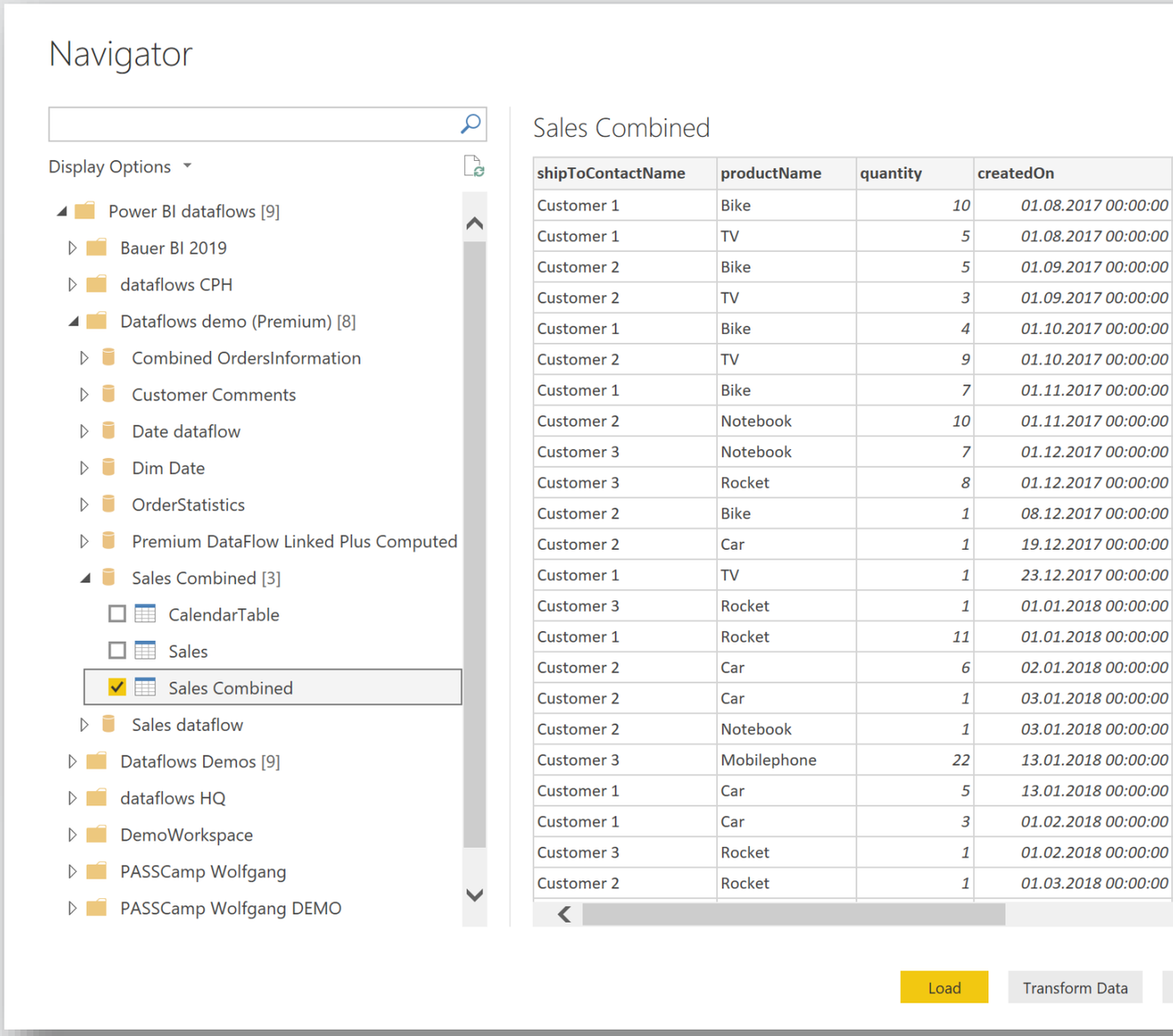
Power BI Dataflows and the Data Lake



DEMO

Consume Dataflow Entities

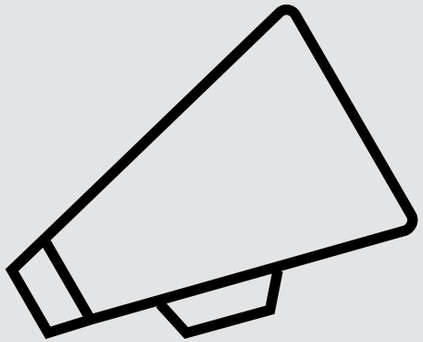




shipToContactName	productName	quantity	createdOn
Customer 1	Bike	10	01.08.2017 00:00:00
Customer 1	TV	5	01.08.2017 00:00:00
Customer 2	Bike	5	01.09.2017 00:00:00
Customer 2	TV	3	01.09.2017 00:00:00
Customer 1	Bike	4	01.10.2017 00:00:00
Customer 2	TV	9	01.10.2017 00:00:00
Customer 1	Bike	7	01.11.2017 00:00:00
Customer 2	Notebook	10	01.11.2017 00:00:00
Customer 3	Notebook	7	01.12.2017 00:00:00
Customer 3	Rocket	8	01.12.2017 00:00:00
Customer 2	Bike	1	08.12.2017 00:00:00
Customer 2	Car	1	19.12.2017 00:00:00
Customer 1	TV	1	23.12.2017 00:00:00
Customer 3	Rocket	1	01.01.2018 00:00:00
Customer 1	Rocket	11	01.01.2018 00:00:00
Customer 2	Car	6	02.01.2018 00:00:00
Customer 2	Car	1	03.01.2018 00:00:00
Customer 2	Notebook	1	03.01.2018 00:00:00
Customer 3	Mobilephone	22	13.01.2018 00:00:00
Customer 1	Car	5	13.01.2018 00:00:00
Customer 1	Car	3	01.02.2018 00:00:00
Customer 3	Rocket	1	01.02.2018 00:00:00
Customer 2	Rocket	1	01.03.2018 00:00:00

Power Platform Dataflows

Power Platform Dataflows



=

Data Integration in the Cloud

Self-Service Data Preparation

75%

“Three-quarters of companies are ***not able to act on the majority of data*** they collect…”

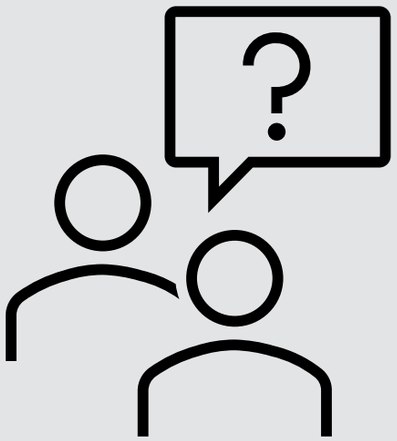
Harvard Business Review, Closing the Customer Experience Gap

80%

“Data Analysts spend ***up to 80% of their time on data preparation*** delaying the time to analysis and decision making.”

- Gartner

How can Dataflows help?



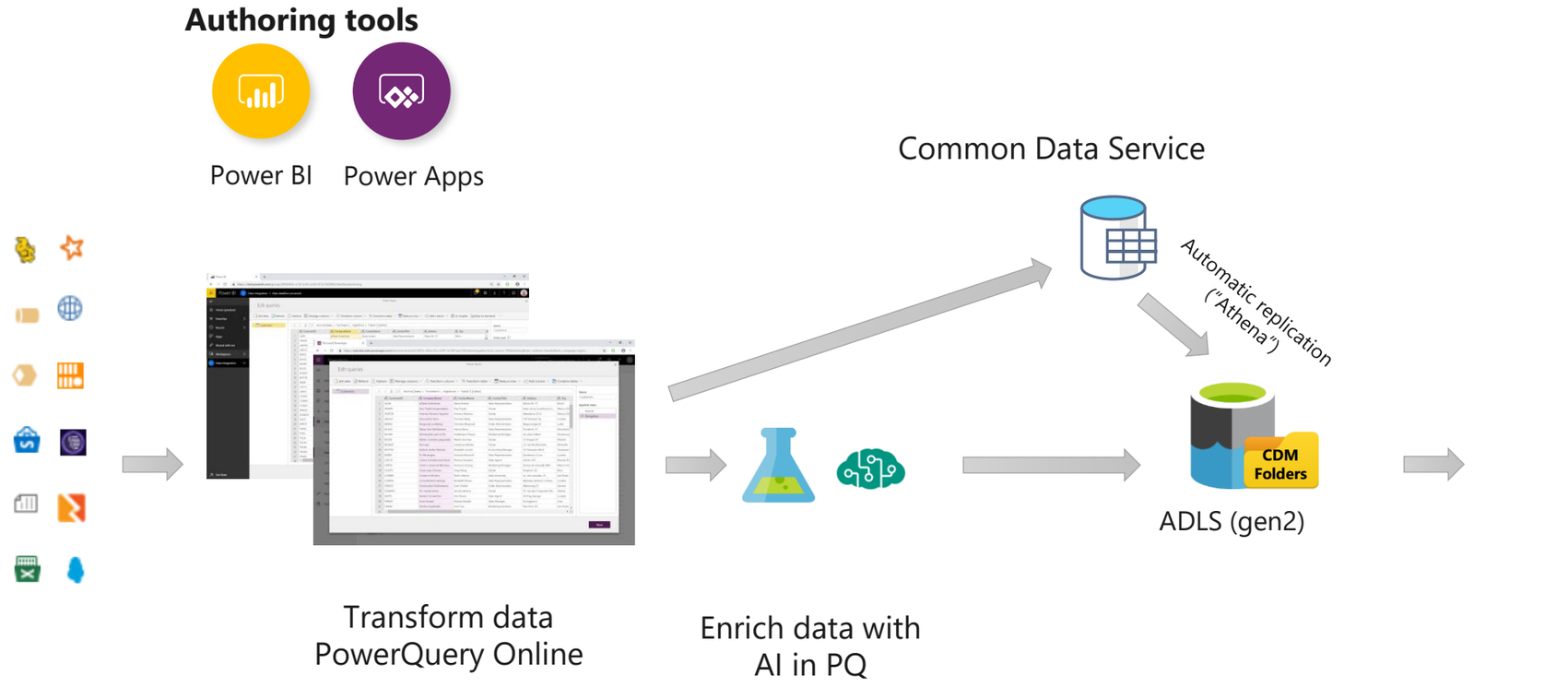
Connectors

Interactive & intuitive

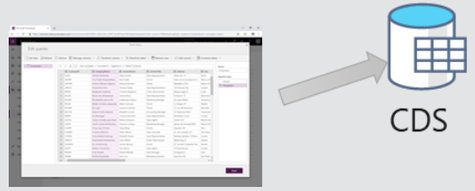
Consistency of experience

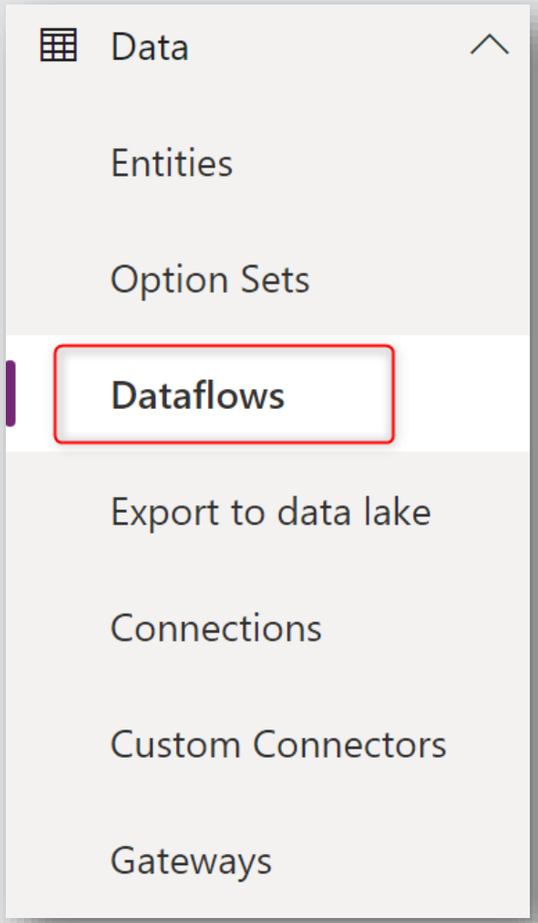
Semantic understanding and data consistency

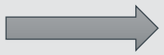
Power Platform Dataflows overview

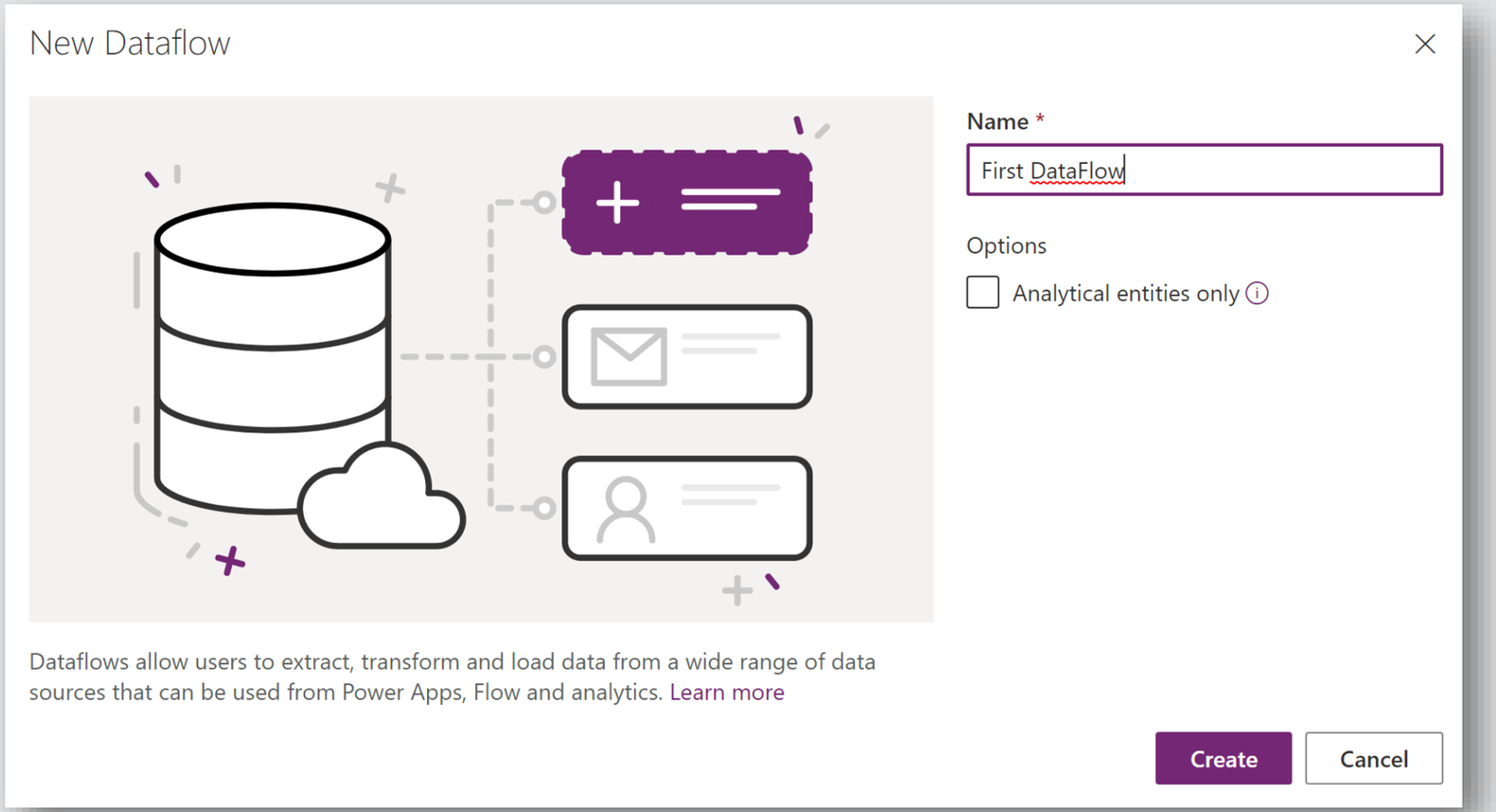


Load Data into CDS

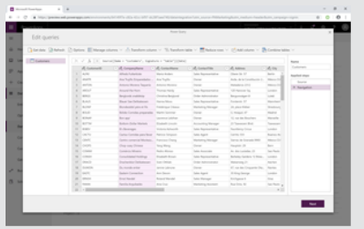






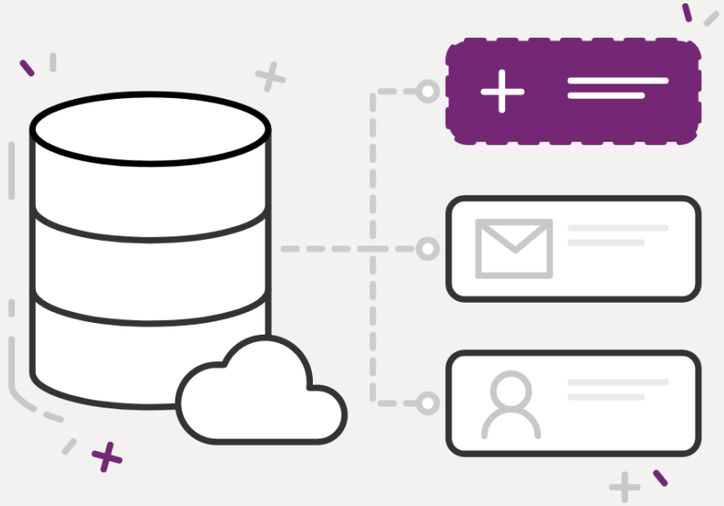


“Standard” versus Analytic Entities



ADLS (gen2)

New Dataflow



Dataflows allow users to extract, transform and load data from a wide range of dat
sources that can be used from Power Apps, Flow and analytics. Learn more

Name *

First DataFlow

Options

Analytical entities only

Queries

Products

Load settings

Load to new entity

Load to existing entity

Do not load

Destination entity

cdmtmp_Product

Entity display name

Product

Entity description

An item that is available for sale.

Delete rows that no longer exist in the query output

Field mapping

Key fields

cdmtmp_ProductNumber

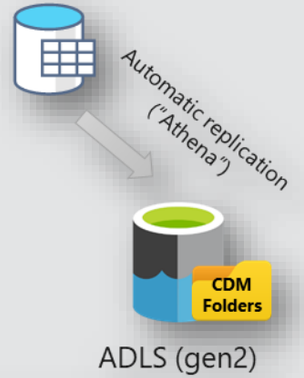
Auto map

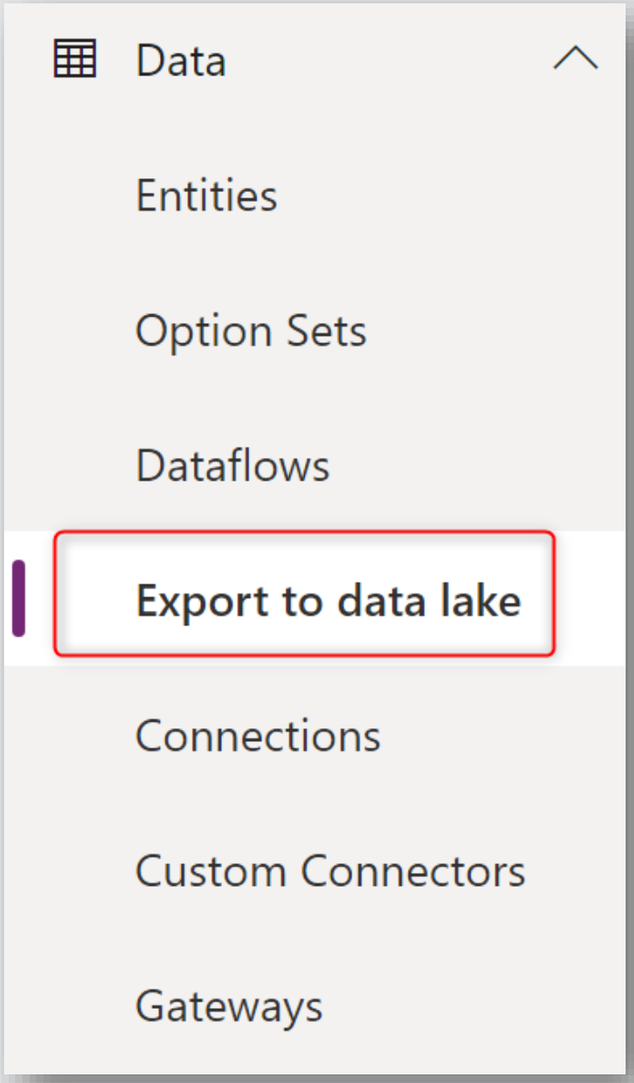
(none)	cdmtmp_DefaultBuyingUnitOfMeasure
(none)	cdmtmp_DefaultSellingQuantity
(none)	cdmtmp_DefaultSellingQuantity_UOM
(none)	cdmtmp_DefaultStockingUnitOfMeasure
(none)	cdmtmp_Description
(none)	cdmtmp_IsStocked
Name	cdmtmp_Name
(none)	cdmtmp_OrganizationId.cdmtmp_OrganizationNumber
(none)	cdmtmp_ProductCategoryId.cdmtmp_CategoryId
ArticleID	cdmtmp_ProductNumber
Category	cdmtmp_ProductType
(none)	cdmtmp_SellingUnitPrice
(none)	cdmtmp_SellingUnitPrice_CurrencyCode
(none)	cdmtmp_StandardCostAmount
(none)	cdmtmp_StandardCostAmount_CurrencyCode
(none)	cdmtmp_Status
(none)	cdmtmp_UnitOfMeasureScale
(none)	EntityImage

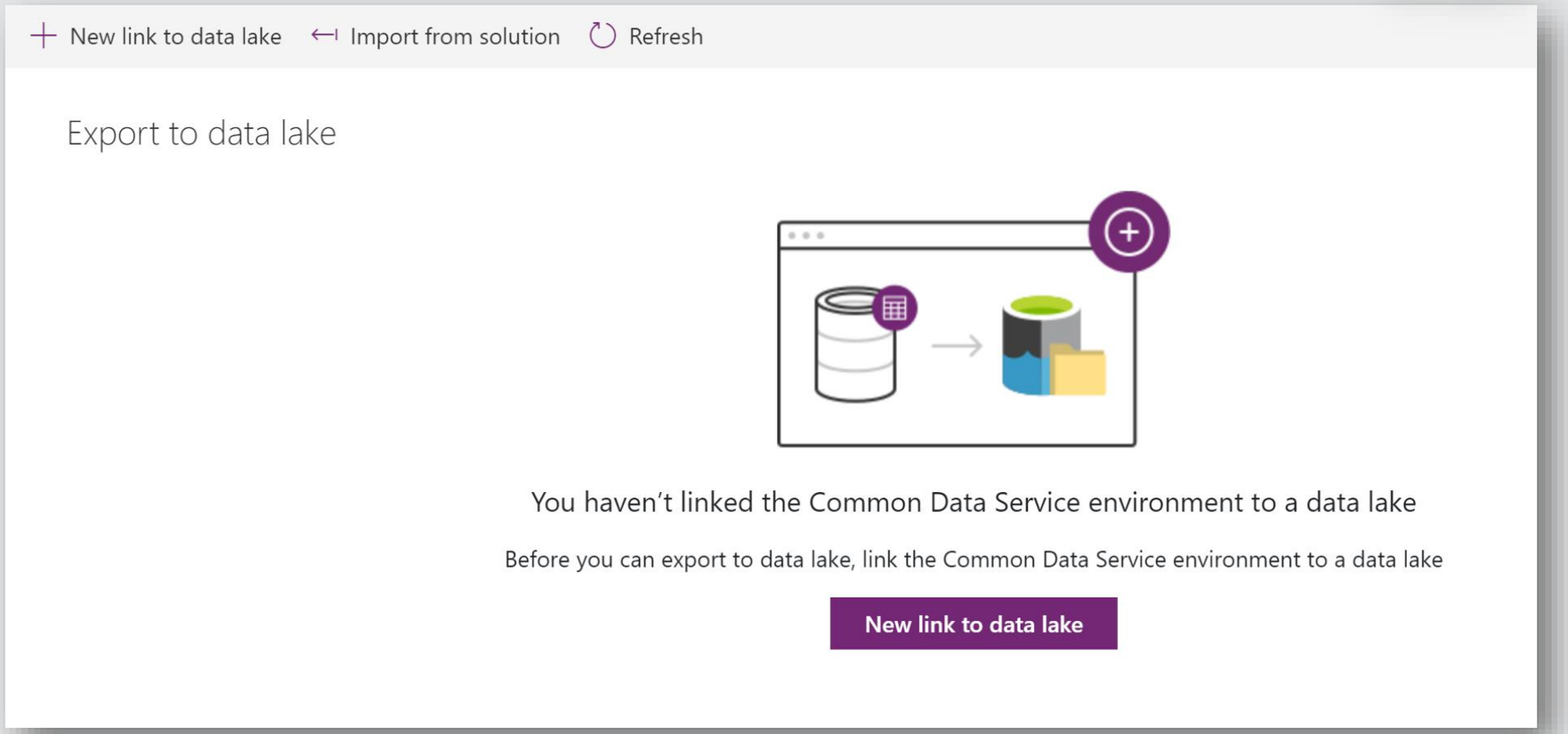
DEMO

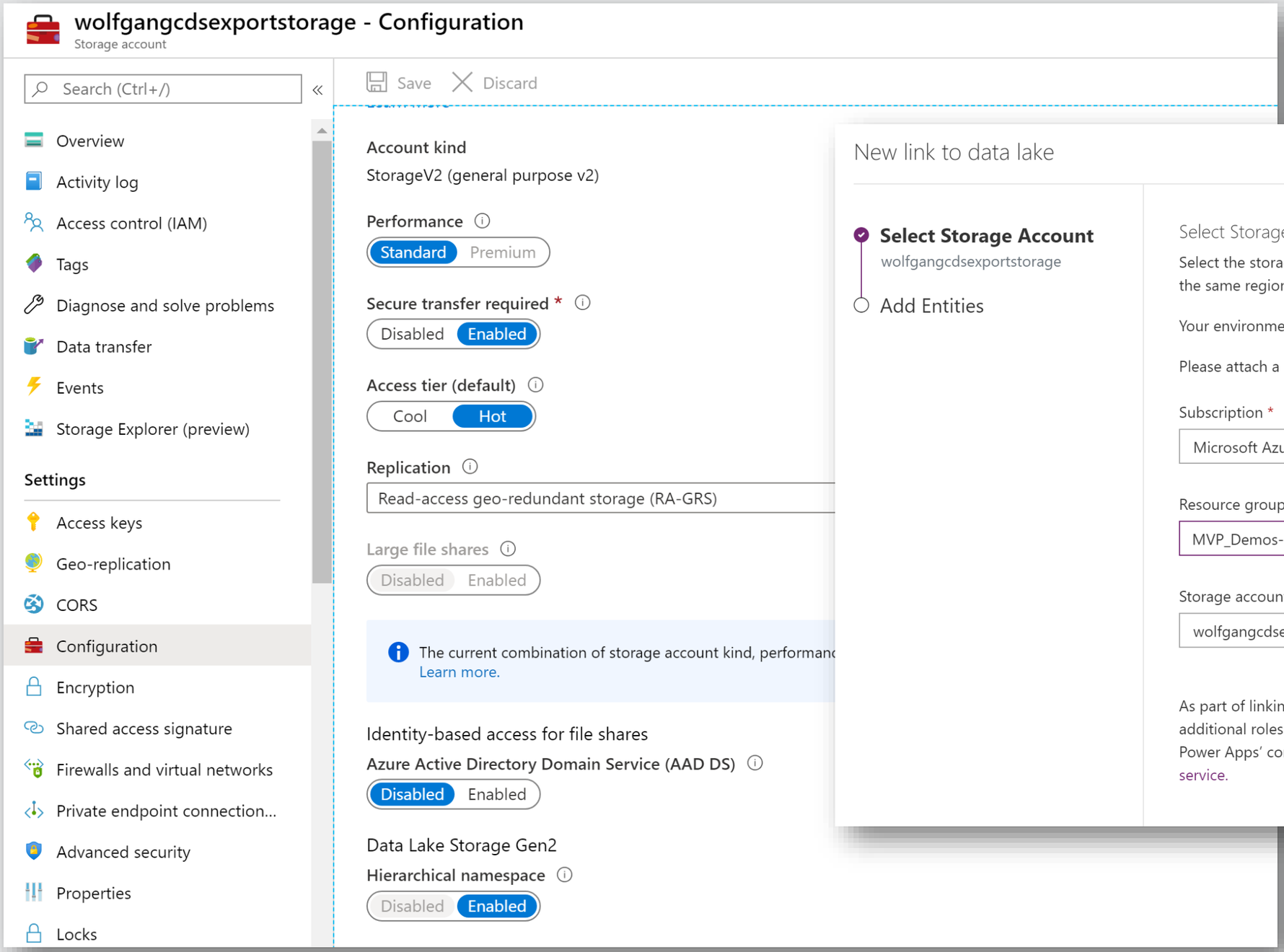
Power Platform Dataflows

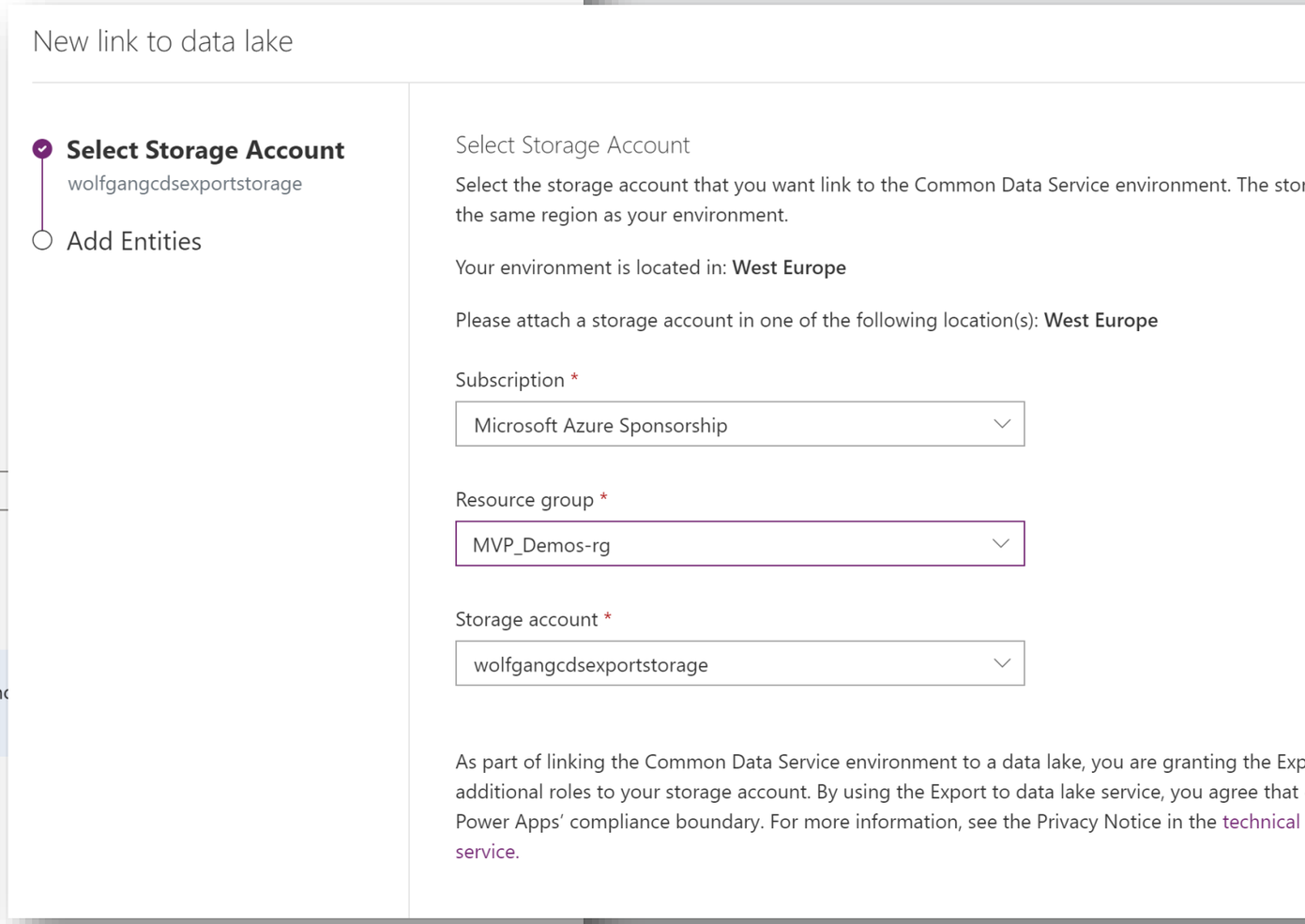
Export to Data Lake – Project “Athena”

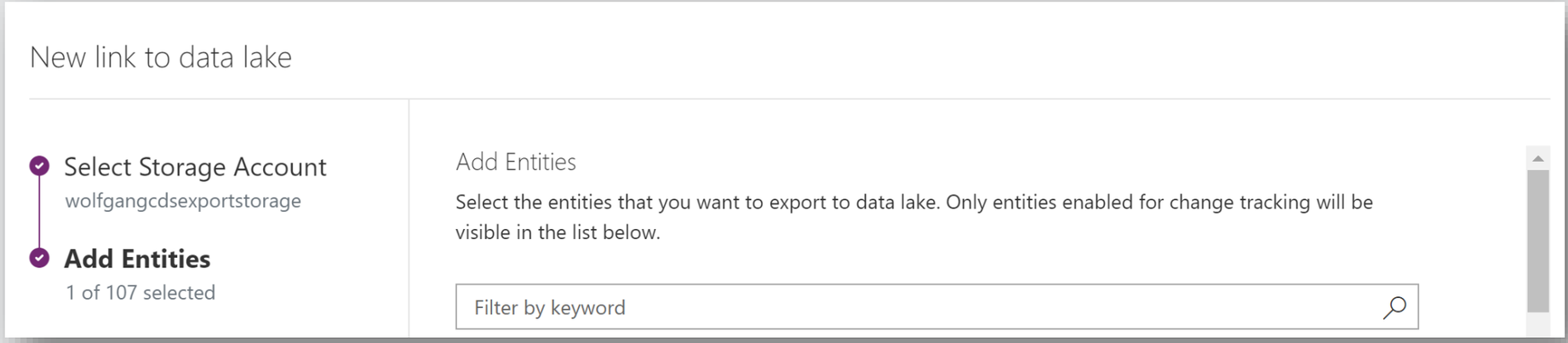


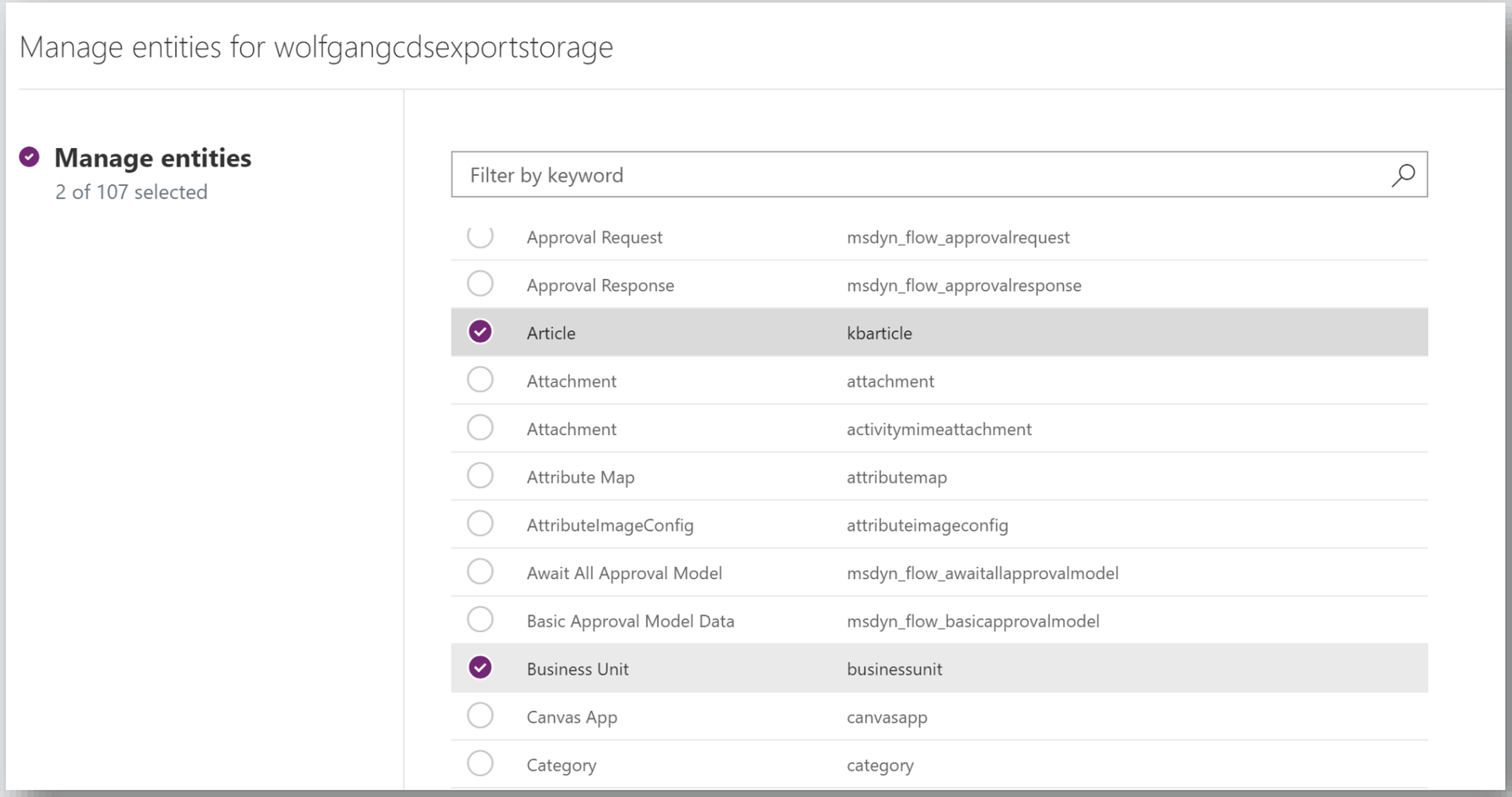




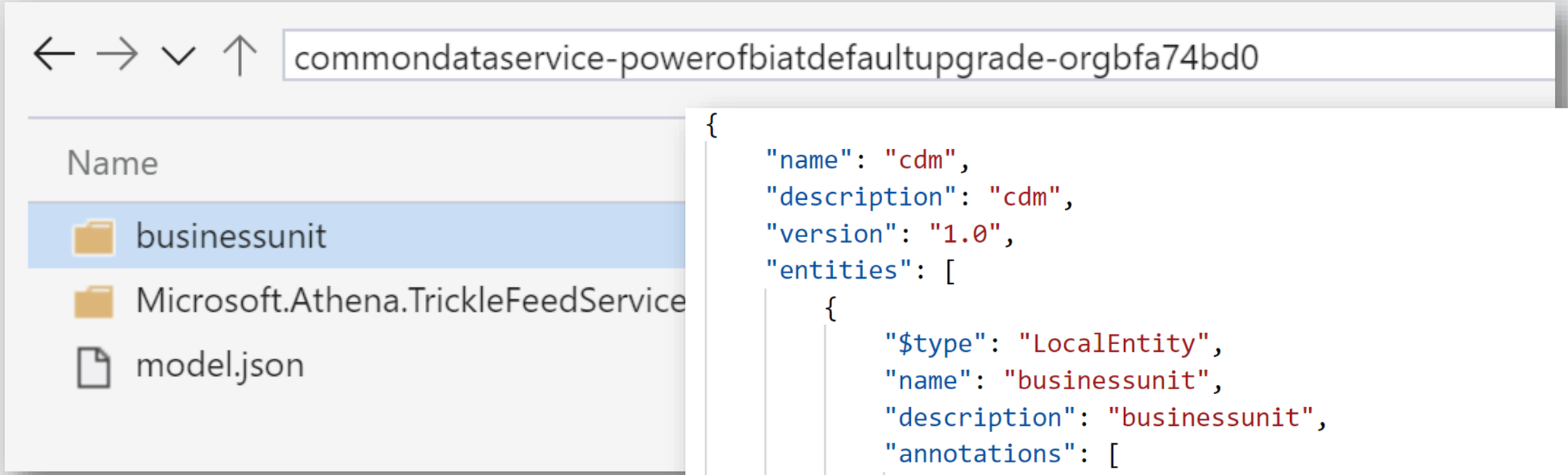

wolfgangcdsexportstorage - Configuration
Storage account
Search (Ctrl+/)
Save
Discard
Overview
Activity log
Access control (IAM)
Tags
Diagnose and solve problems
Data transfer
Events
Storage Explorer (preview)
Settings
Access keys
Geo-replication
CORS
Configuration
Encryption
Shared access signature
Firewalls and virtual networks
Private endpoint connection...
Advanced security
Properties
Locks
Account kind
StorageV2 (general purpose v2)
Performance
Standard
Premium
Secure transfer required *
Disabled
Enabled
Access tier (default)
Cool
Hot
Replication
Read-access geo-redundant storage (RA-GRS)
Large file shares
Disabled
Enabled
The current combination of storage account kind, performanc
Learn more.
Identity-based access for file shares
Azure Active Directory Domain Service (AAD DS)
Disabled
Enabled
Data Lake Storage Gen2
Hierarchical namespace
Disabled
Enabled


New link to data lake
Select Storage Account
wolfgangcdsexportstorage
Add Entities
Select Storage Account
Select the storage account that you want link to the Common Data Service environment. The stor
the same region as your environment.
Your environment is located in: West Europe
Please attach a storage account in one of the following location(s): West Europe
Subscription *
Microsoft Azure Sponsorship
Resource group *
MVP_Demos-rg
Storage account *
wolfgangcdsexportstorage
As part of linking the Common Data Service environment to a data lake, you are granting the Exp
additional roles to your storage account. By using the Export to data lake service, you agree that
Power Apps' compliance boundary. For more information, see the Privacy Notice in the technical
service.


New link to data lake
Select Storage Account
wolfgangcdsexportstorage
Add Entities
1 of 107 selected
Add Entities
Select the entities that you want to export to data lake. Only entities enabled for change tracking will be visible in the list below.
Filter by keyword


Manage entities for wolfgangcdsexportstorage
Manage entities
2 of 107 selected
Filter by keyword
Approval Request msdyn_flow_approvalrequest
Approval Response msdyn_flow_approvalresponse
Article kbarticle
Attachment attachment
Attachment activitymimeattachment
Attribute Map attributemap
AttributeImageConfig attributeimageconfig
Await All Approval Model msdyn_flow_awaitallapprovalmodel
Basic Approval Model Data msdyn_flow_basicapprovalmodel
Business Unit businessunit
Canvas App canvasapp
Category category

What’s in my Data Lake?

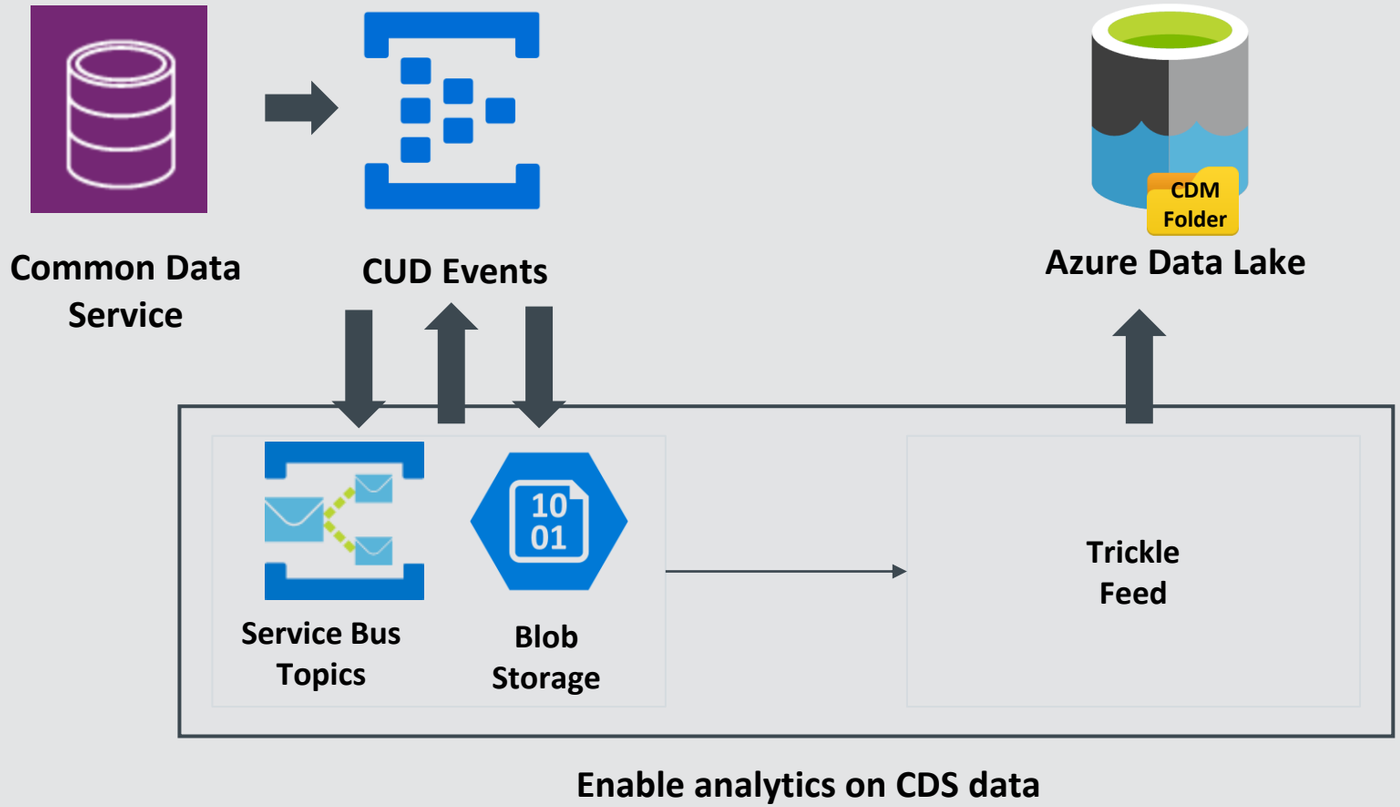



```
{
    "name": "cdm",
    "description": "cdm",
    "version": "1.0",
    "entities": [
        {
            "$type": "LocalEntity",
            "name": "businessunit",
            "description": "businessunit",
            "annotations": [
                {
                    "name": "Athena:PartitionGranularity",
                    "value": "Year"
                },
                {
                    "name": "Athena:InitialSyncState",
                    "value": "Completed"
                },
                {
                    "name": "Athena:InitialSyncDataCompletedTime",
                    "value": "3/1/2020 5:58:33 PM"
```

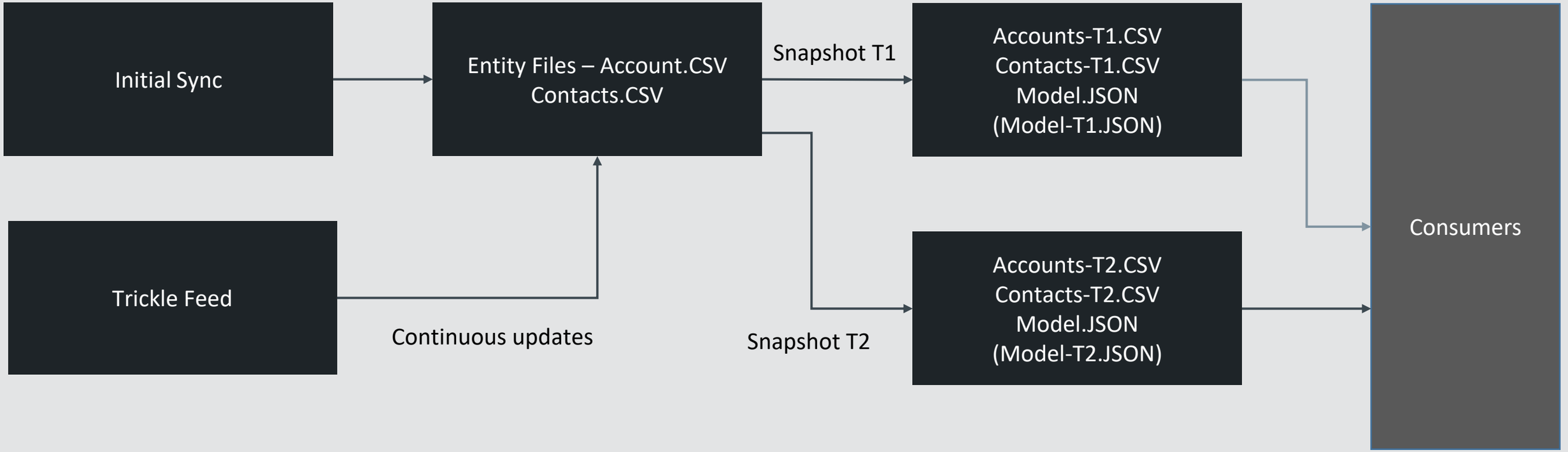
Behind the Scenes

Support for **full** and **incremental**

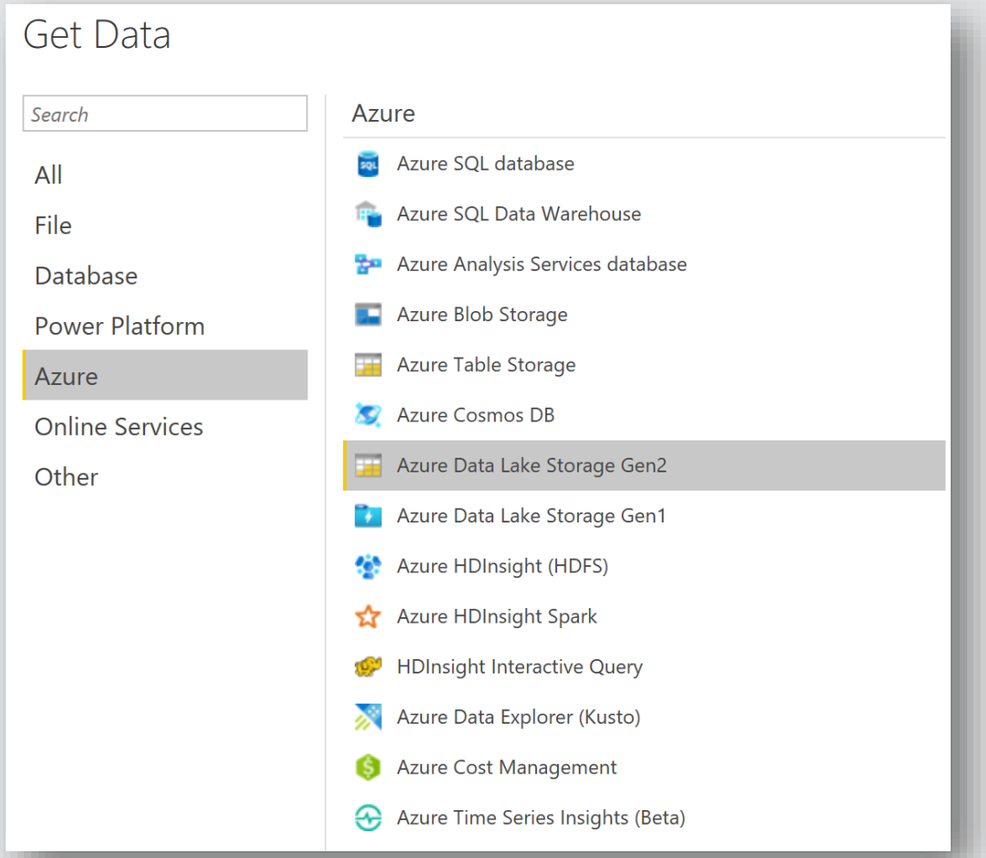
Trickle feed to write data to ADLS Gen2

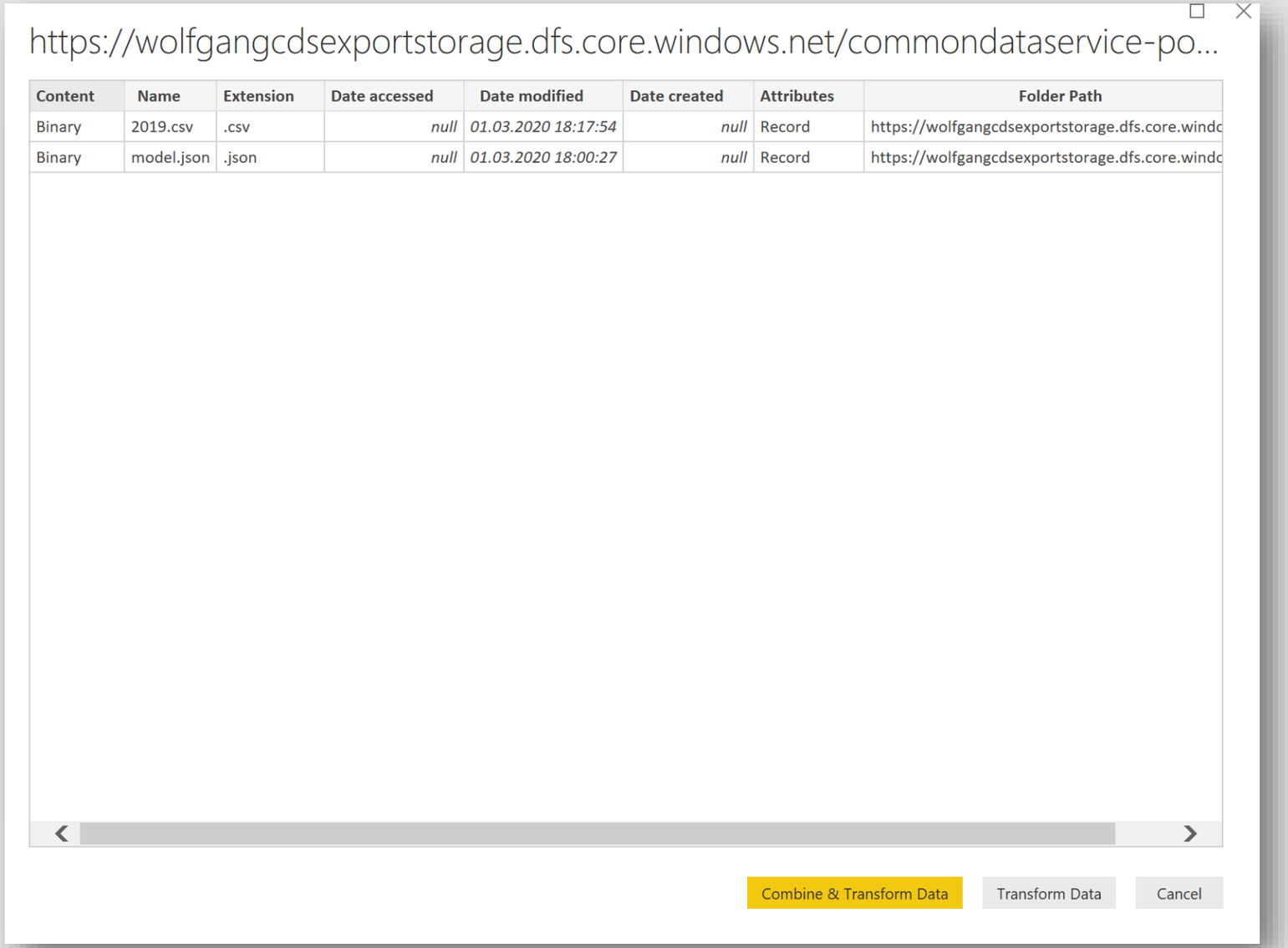


Consistent Data & Metadata Updates

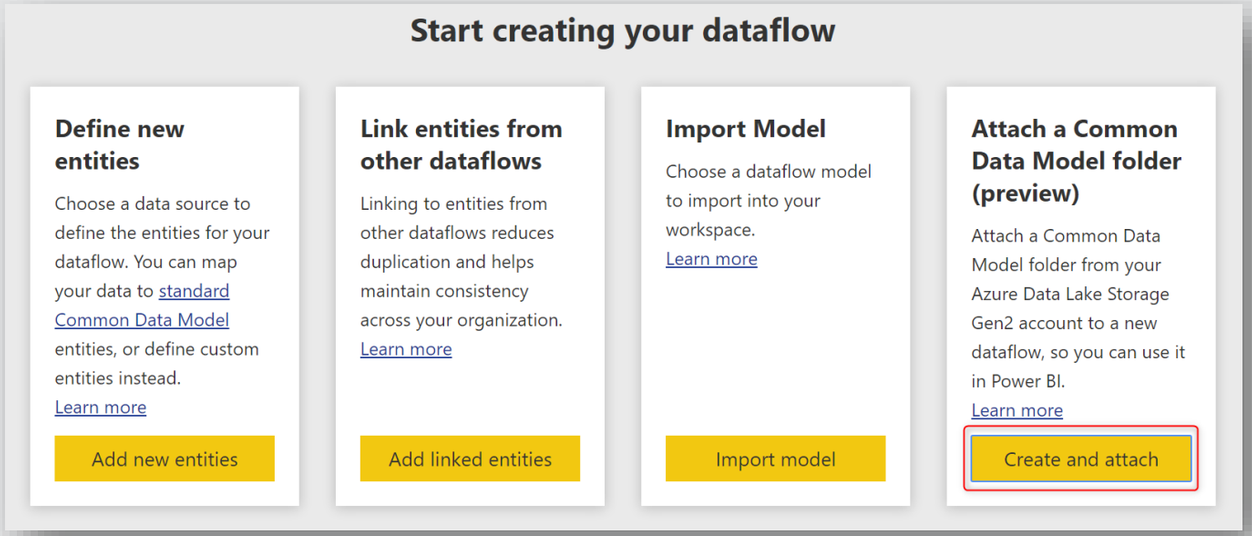


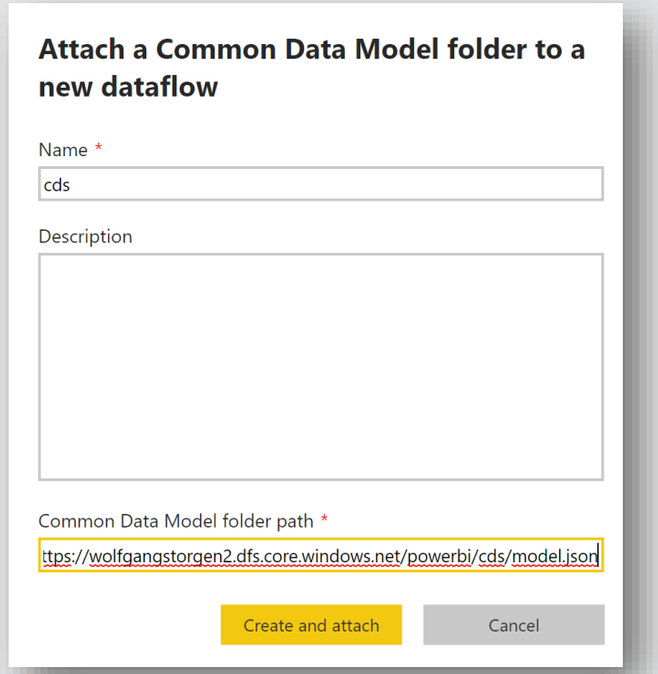
Consume in Power BI (Option 1)

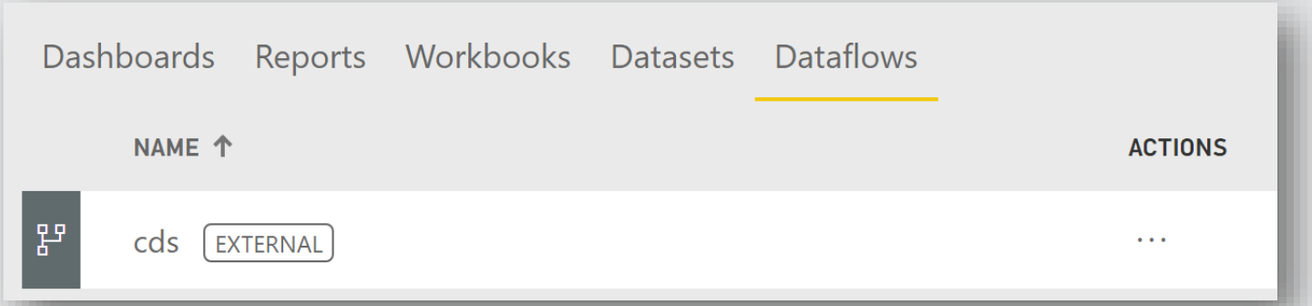




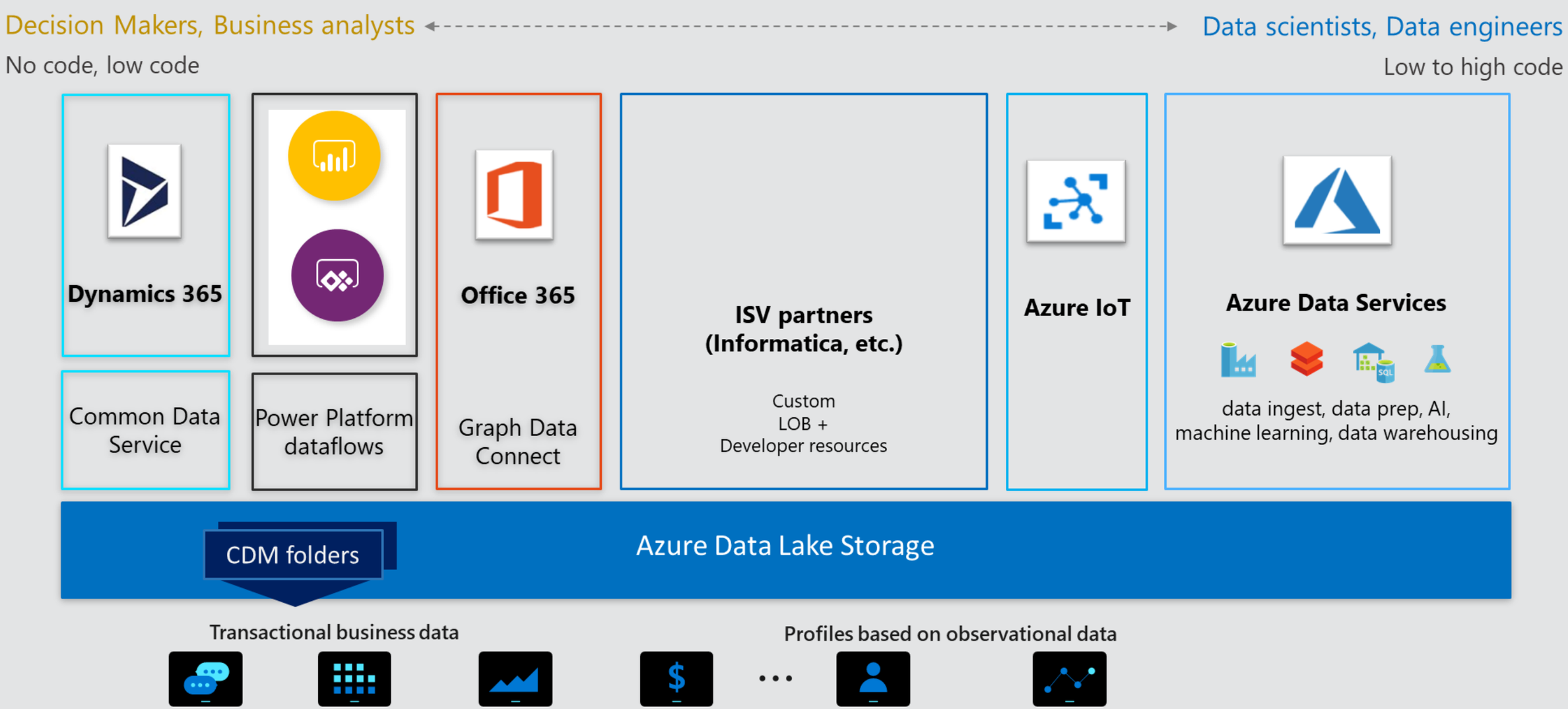
Consume in Power BI (Option 2)



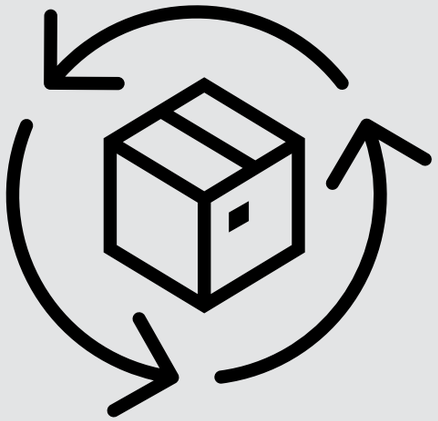




Power Platform Dataflows and the Data Lake



Power Platform Dataflows…



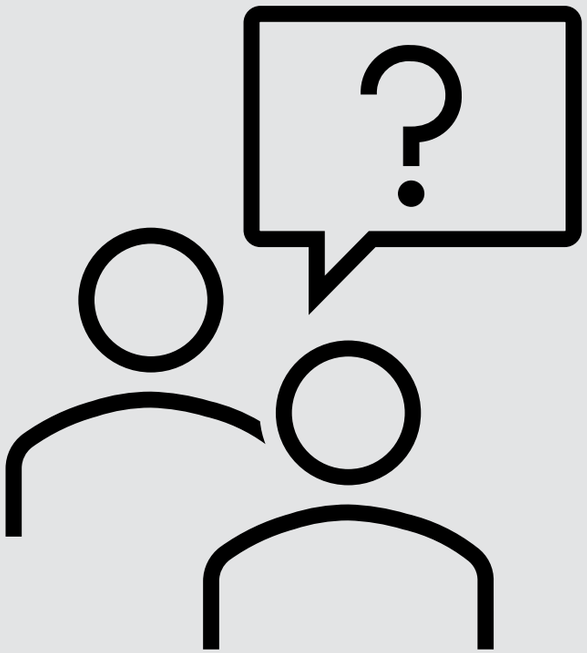
Data Integration & Preparation in the Cloud

Data Lake

Common Set of Technology

- Connectors
- Power Query
- CDM Folders

Data Exchange made easy



Wolfgang Strasser

wolfgang_AT_powerofbi.at

@wstrasser

workingondata.wordpress.com

Microsoft® MVP Most Valuable Professional

UBIDO AN ACP GROUP COMPANY

#DataWeekender

Additional Links

- Common Data Model

  https://docs.microsoft.com/en-us/common-data-model/

- Common Data Model (github)

  https://github.com/Microsoft/CDM

- Additional Datasources for Power BI Dataflows (no UI)

  https://docs.microsoft.com/en-us/power-bi/service-dataflows-data-sources#connecting-to-additional-data-sources

- Power Apps Dataflows

  https://docs.microsoft.com/en-us/powerapps/maker/common-data-service/create-and-use-dataflows