

# Power Platform Dataflows

Power Query (and more) in the Cloud

*Jo Servus!*  
*Griass Eich!*

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



25. August 2020

# 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

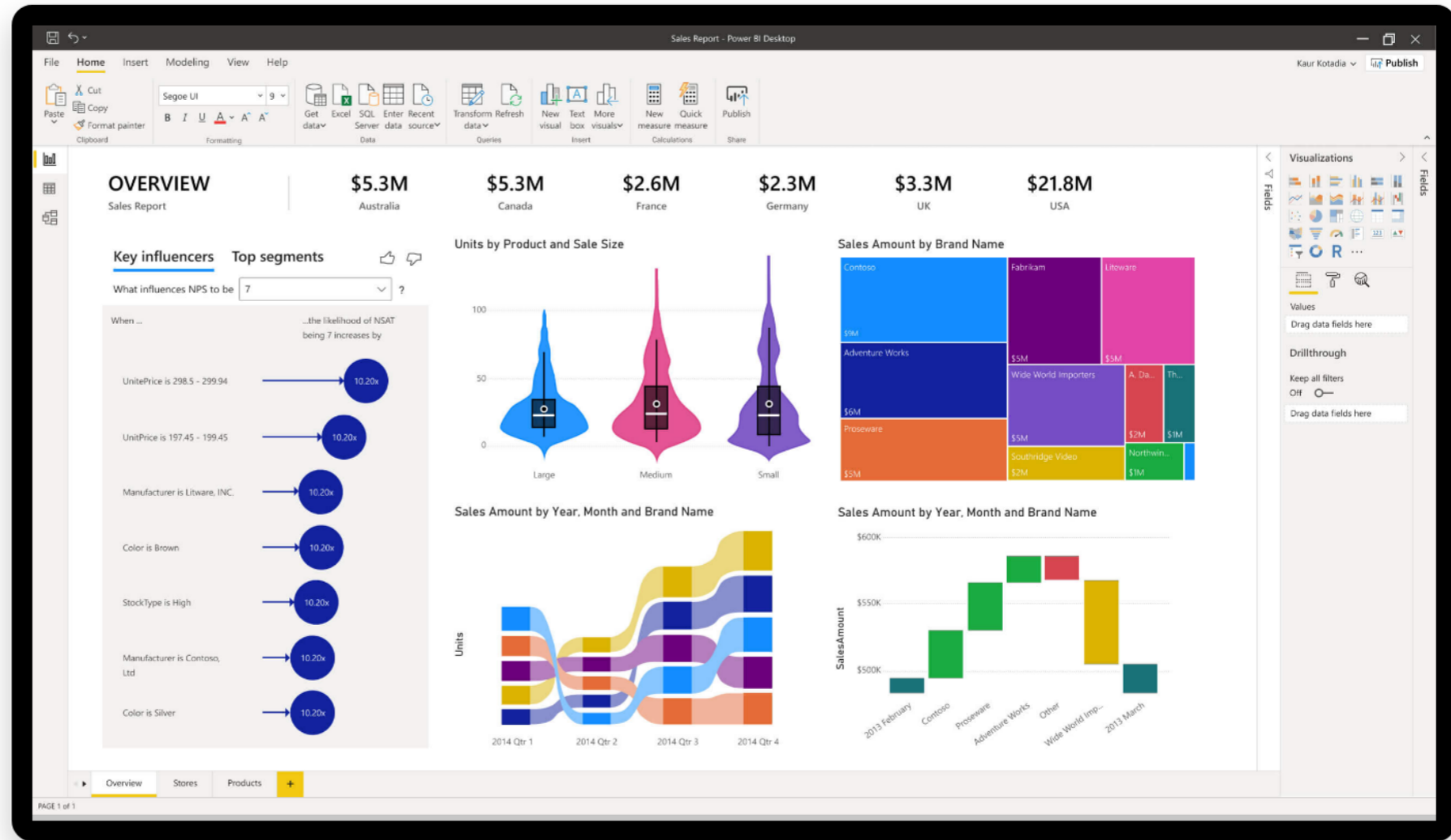
Boost business productivity to get more done by giving everyone the ability to automate organizational processes.

[Learn more >](#)

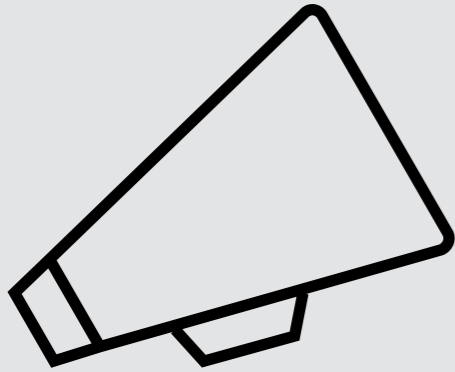
### Power Virtual Agents

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

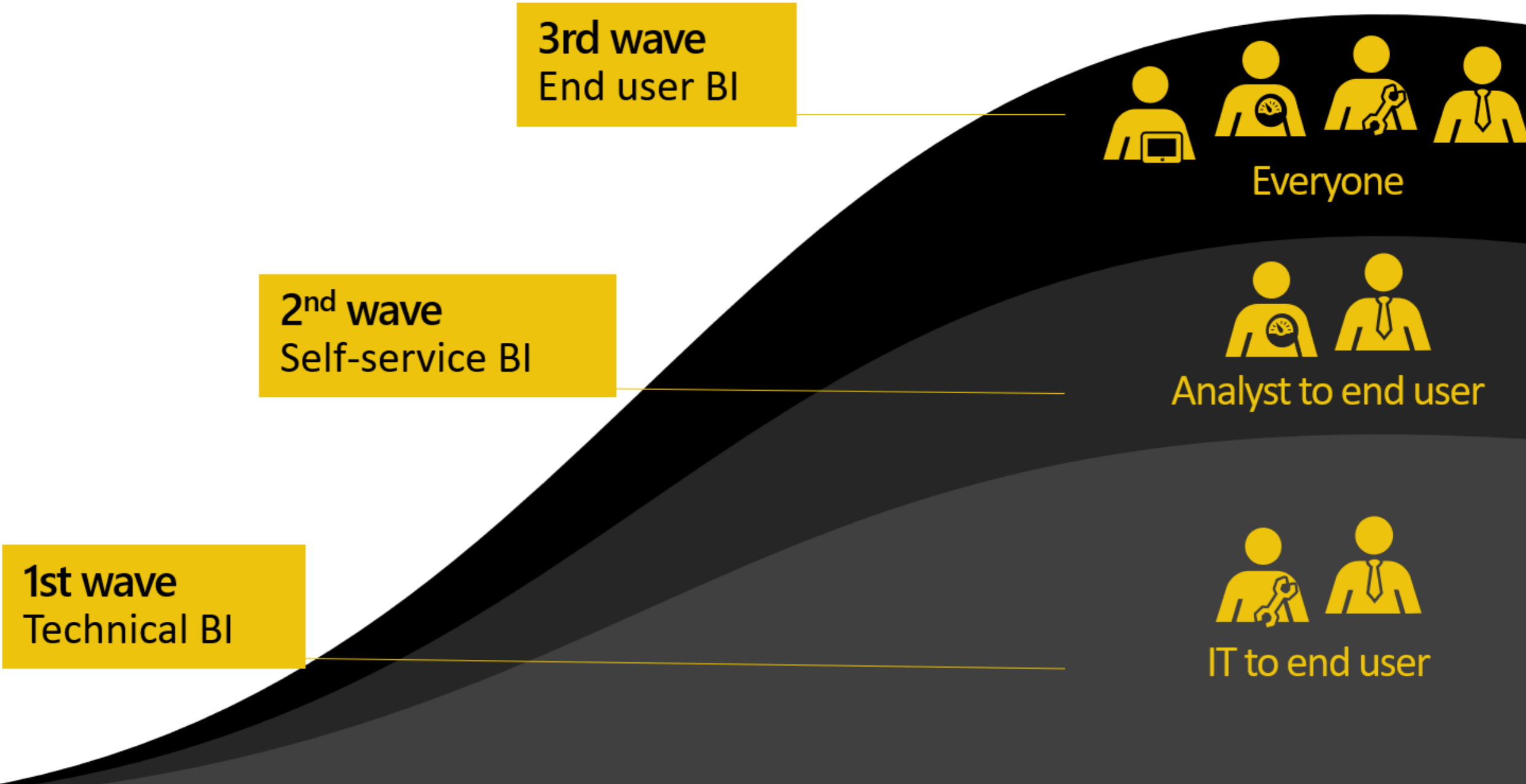
Reusability

Data Lake

Common Data Model (CDM)

Power Query (Online)

# Evolution of Business Intelligence





# Self Service BI

BI / Reporting



ETL

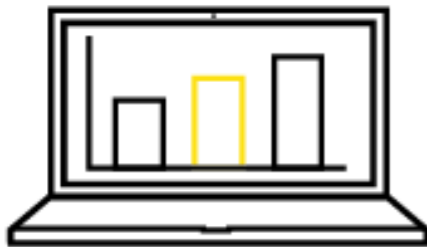


Hundreds of data sources



CONNECT

Power BI Desktop



PUBLISH

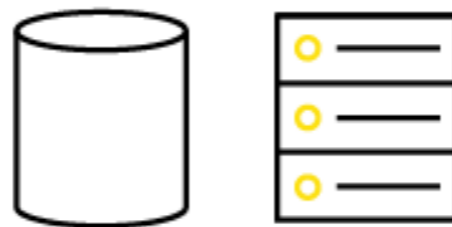
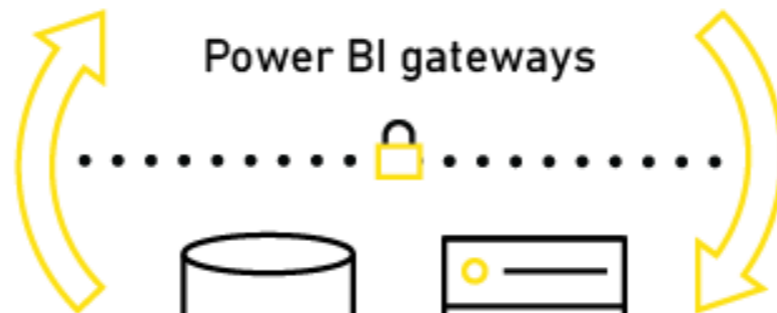
Power BI service



ACCESS



Power BI gateways



YOUR ORGANIZATION'S DATA

# Interactive Reporting

## Sales Report



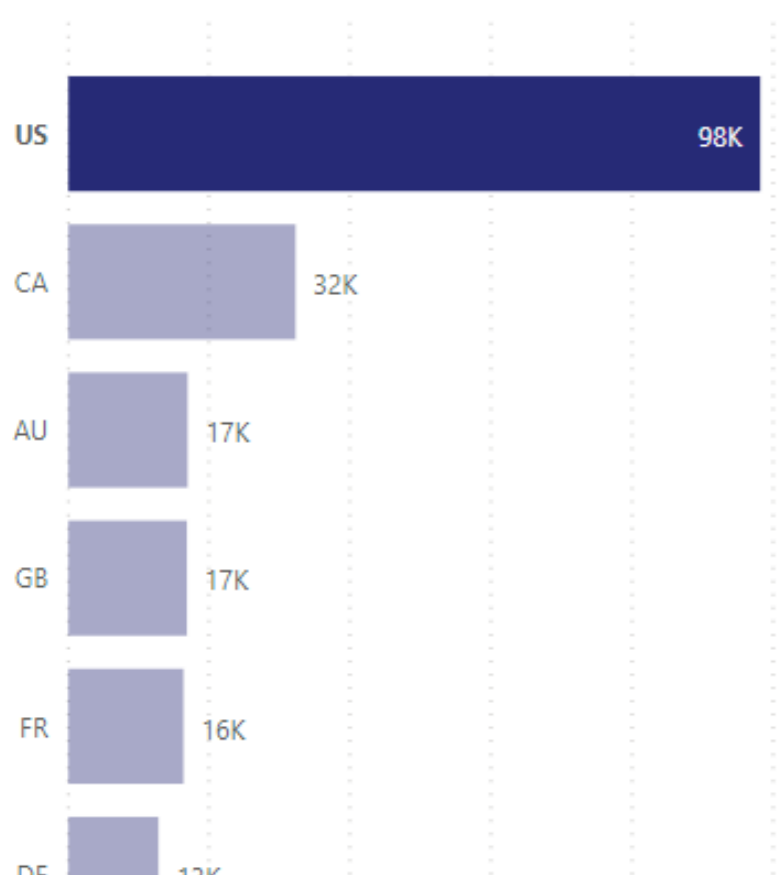
98K

Units Sold

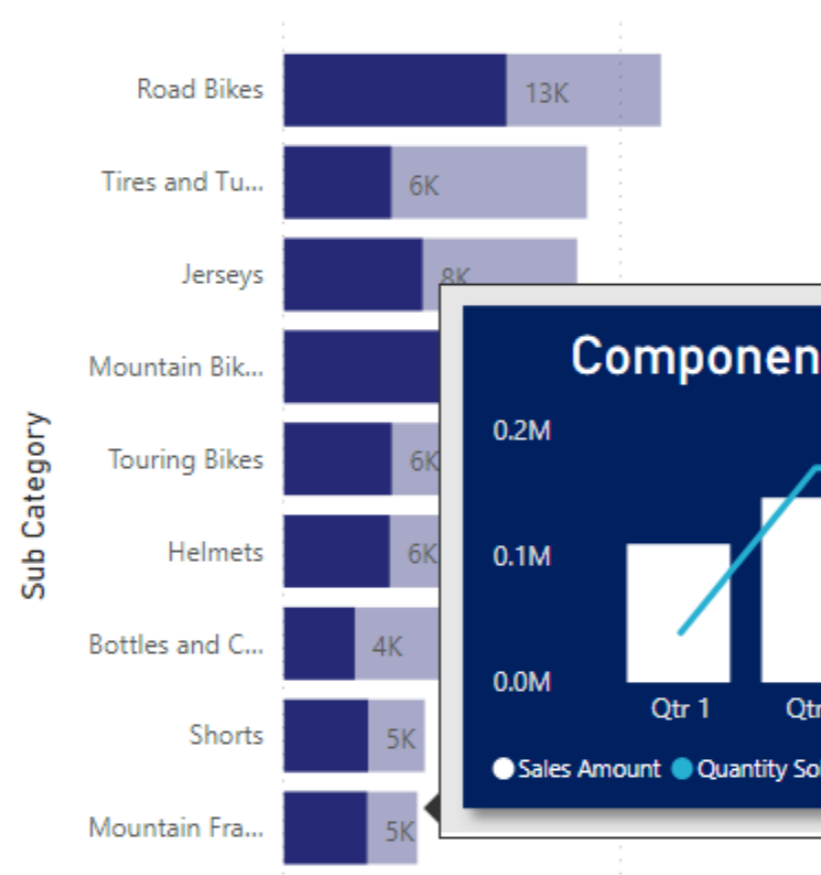
14.79M

Sales Amount

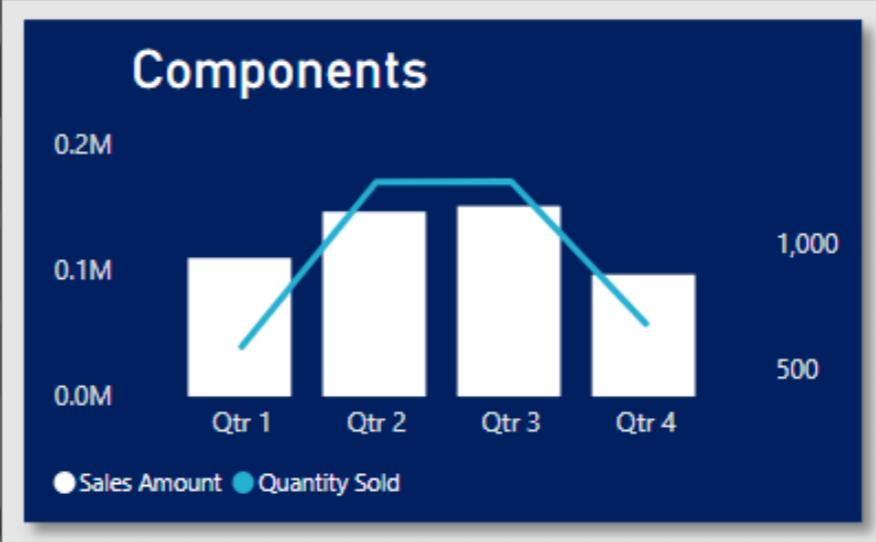
### Country Breakdown



### Product Breakdown

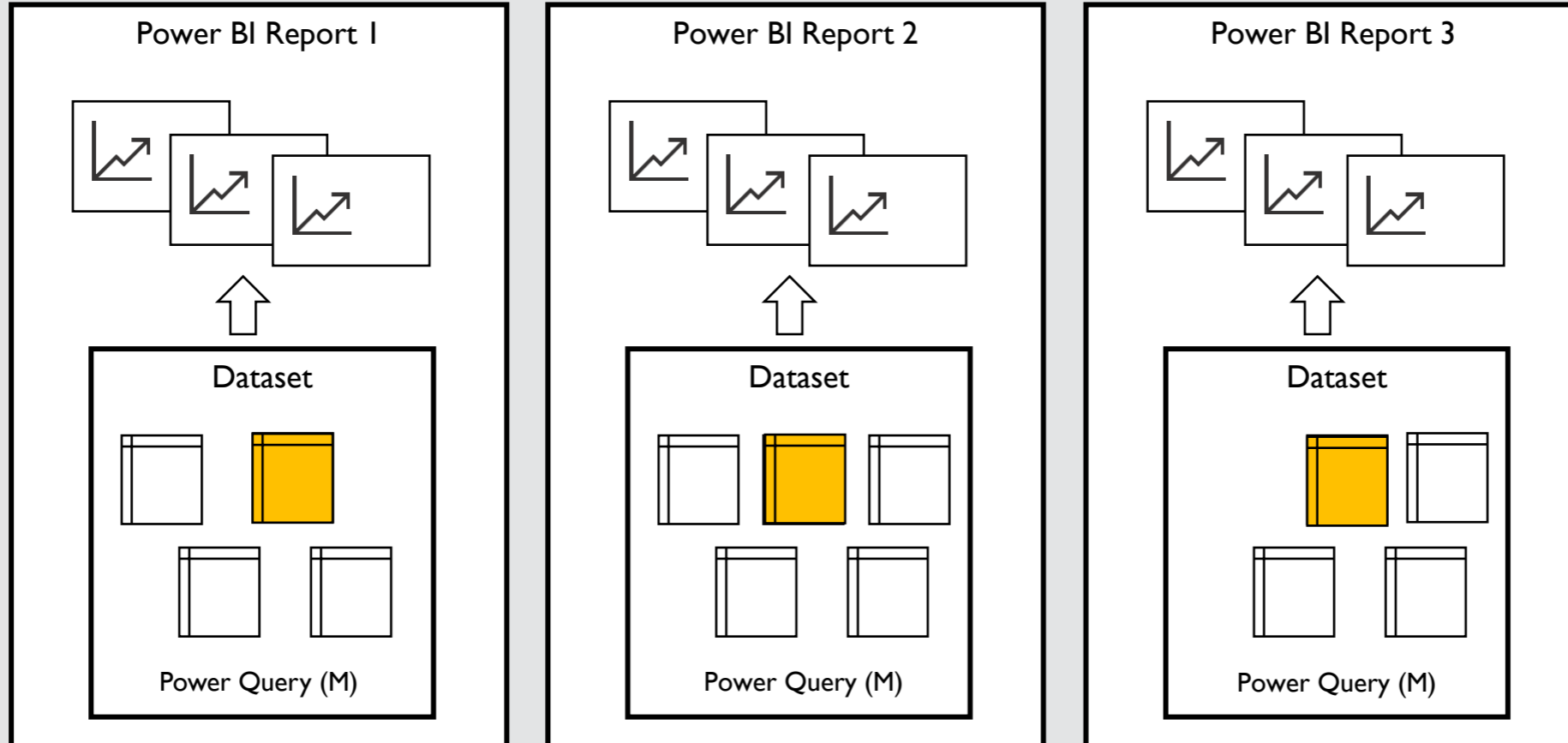


Quantity sold per year as column chart





# (Reusable) ETL in Power BI



# Power Query

DIAD Final Report - Power Query Editor

File Home Transform Add Column View Help

Close & Apply Close New Source Recent Sources Enter Data Data source settings Manage Parameters Refresh Preview Advanced Editor Choose Columns Remove Columns Keep Rows Remove Rows Sort Split Column Group By Data Type: Text Use First Row as Headers Merge Queries Append Queries Combine Files Text Analytics Vision Azure Machine Learning

Queries [9]

- Transform File from International [3]
  - Sample Query [2]
    - Sample File
    - Sample File Parameter1 (Sample Fl...
    - Transform Sample File from Internat...
    - Transform File from International
  - Other Queries [5]
    - Sales
    - Product
    - Geography
    - Manufacturer
    - International

Query Settings

PROPERTIES

Name: Geography

APPLIED STEPS

- Source
- Navigation
- Changed Type
- Removed Top Rows
- Promoted Headers
- Changed Type1

Table: TransformColumnTypes(#"Promoted Headers",{{"Zip", type text}, {"City", type text}, {"State", type text}, {"Region", type

Zip	City	Region	District	Country
22654	Star Tanner	East	District #07	USA
22655	Stephens Ci	East	District #07	USA
22656	Stephenson	East	District #07	USA
22657	Strasburg, V	East	District #07	USA
22660	Toms Brook	East	District #07	USA
22663	White Post,	East	District #07	USA
22664	Woodstock,	East	District #07	USA
22701	Culpeper, V	East	District #07	USA
22709	Aroda, VA, U	East	District #07	USA
22711	Banco, VA, U	East	District #07	USA
22712	Bealeton, VA	East	District #07	USA
22713	Boston, VA,	East	District #07	USA
22714	Brandy Stati	East	District #07	USA
22715	Brightwood,	East	District #07	USA
22716	Castleton, V	East	District #07	USA
22718	Elkwood, VA	East	District #07	USA
22719	Etlan, VA, U	East	District #07	USA
22720	Goldvein, VA, USA	East	District #07	USA
22722	Haywood, VA, USA	East	District #07	USA

# Power BI Dataflows

# Power BI Information Hierarchy (before Dataflows)

**Dashboards**

**Interactive Reports**

**Paginated Reports**

**Datasets**

**Raw Data**

# Power BI Information Hierarchy

**Dashboards**

**Interactive Reports**

**Paginated Reports**

**Datasets**

**Dataflows**

**Raw Data**

# DEMO










































## Power BI Dataflows

# Power Query Online

## Power Query - Choose data source

All categories [File](#) [Database](#) [Power Platform](#) [Azure](#) [Online services](#) [Other](#)

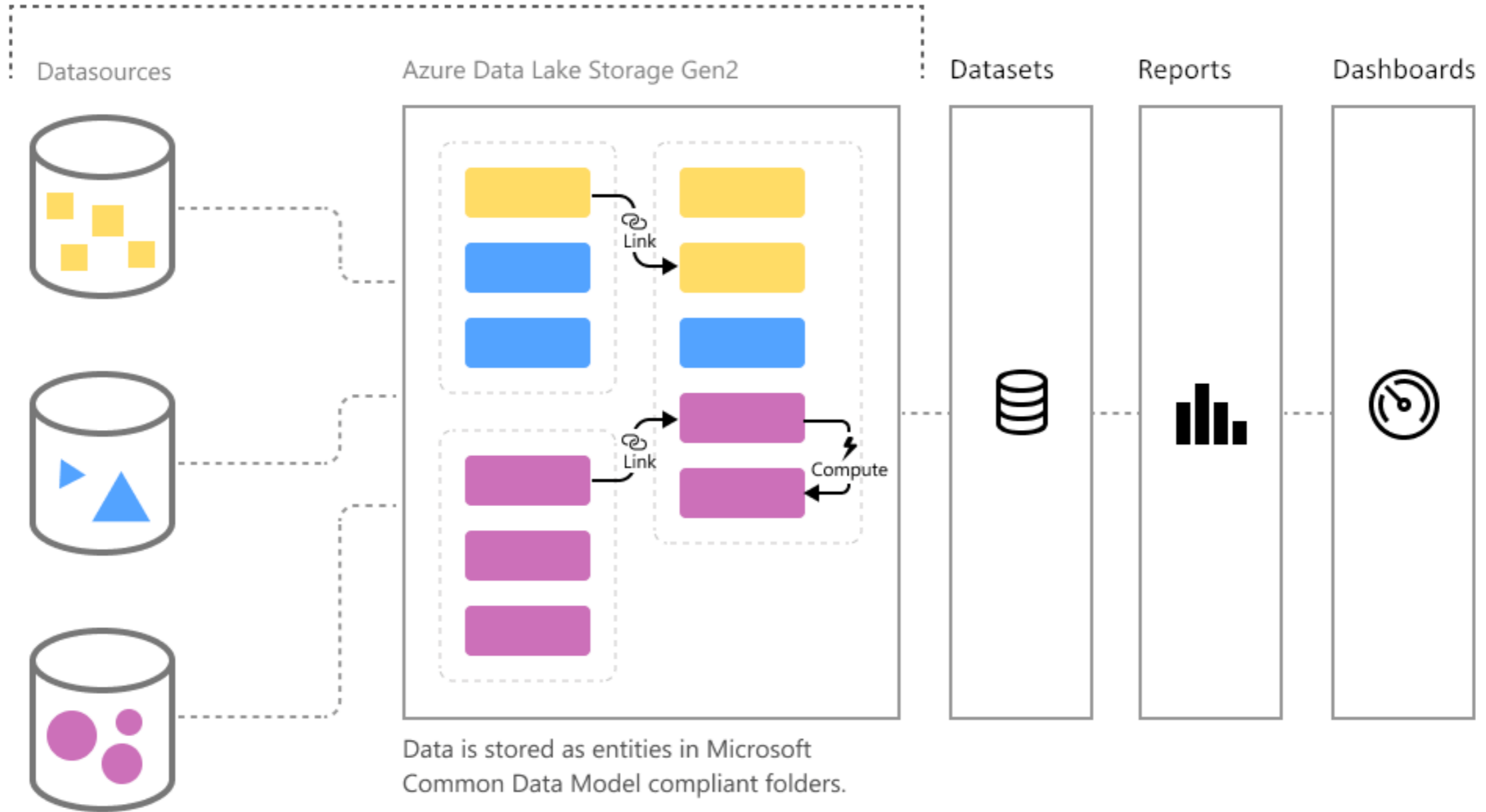
### Data sources

 Access File	 Excel File	 Folder File	 JSON File	 PDF File	 SharePoint folder File
 Text/CSV File	 XML File	 Amazon Redshift Database	 Google BigQuery Database	 IBM Db2 database Database	 MySQL database Database
 Oracle database Database	 PostgreSQL database Database	 SQL Server database Database	 Sybase database Database	 Teradata database Database	 Vertica Database
 Azure Blobs Azure	 Azure Data Explorer (Kusto) Azure	 Azure Data Lake Storage Gen2 Azure	 Azure HDInsight Spark Azure	 Azure SQL Data Warehouse Azure	 Azure SQL database Azure
 Azure Tables Azure	 Microsoft Exchange Online Online services	 Salesforce objects Online services	 Salesforce reports Online services	 SharePoint Online list Online services	 Smartsheet Online services
 Active Directory Other	 Common Data Service Power Platform	 OData Other	 Odbc Other	 Power BI dataflows Power Platform	 SharePoint list Other
 Spark Other	 Web API Other	 Web page Other	 Blank table Other	 Blank query Other	

### Templates

 Accounts, leads, opportunities Salesforce Online services	 Lead to cash Dynamics 365 Sales Online services	 Leads, opportunities Dynamics 365 Sales Online services	 Quotes, orders, invoices Dynamics 365 Sales Online services
---	---	--	---

## Dataflows







# Common Data Model

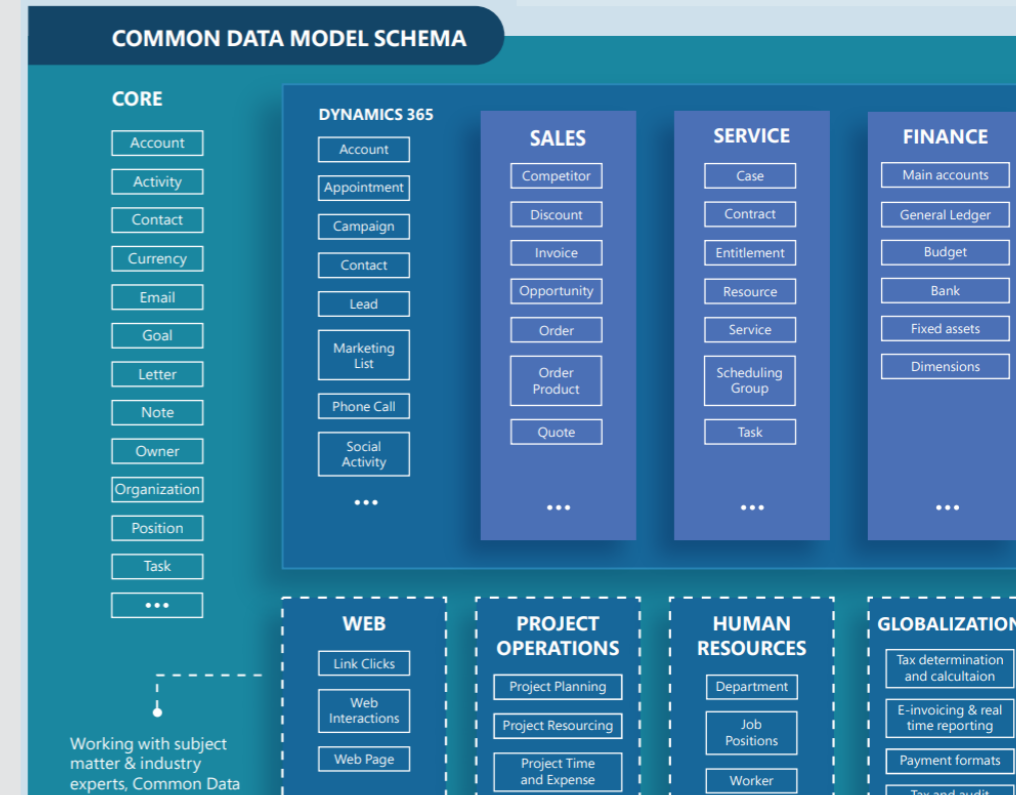
## Set of

- standardized, extensible data schemas

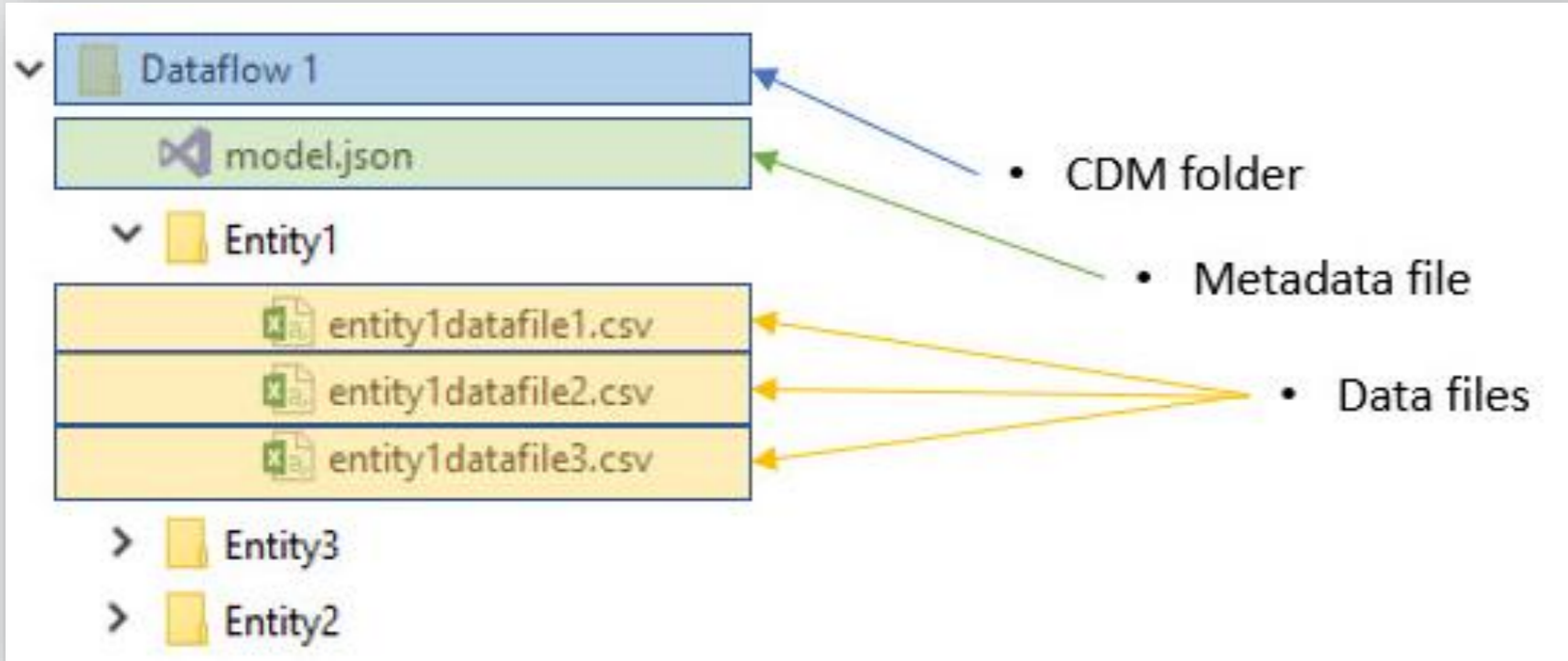
- Entities
- Attributes
- Semantic Metadata & Relationships

## 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](#)



# CDM Folders

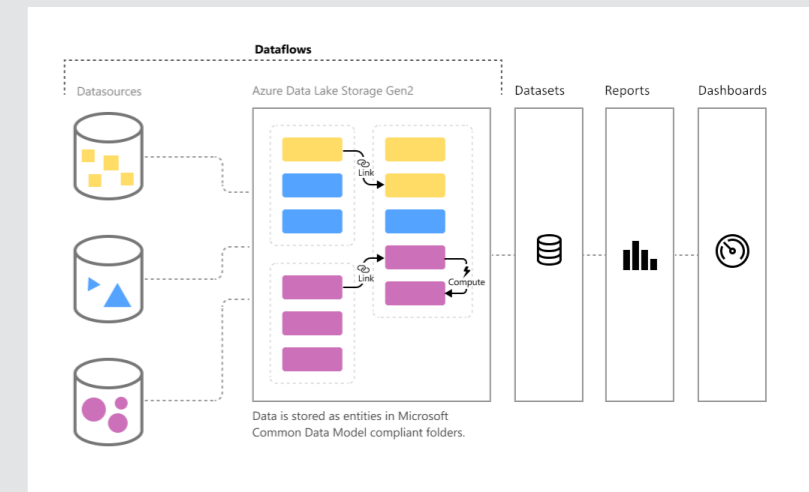


```
{ Sales From Excel.json • First Dataflow.json •  
1 {  
2   "name": "Sales From Excel",  
3   "description": "",  
4   "version": "1.0",  
5   "culture": "en-US",  
6   "pbi:mashup": {  
7     "fastCombine": true,  
8     "allowNativeQueries": false,  
9     "queriesMetadata": {  
10      "Sales": {  
11        "queryId": "40b69972-542b-49ee-ab",  
12        "queryName": "Sales",  
13        "loadEnabled": true  
14      }  
15    },  
16    "document": "section Section1;\r\n\nshared",  
17  },  
18  "entities": [  
19    {  
20      "$type": "LocalEntity",  
21      "name": "Sales",  
22      "description": "",  
23      "pbi:refreshPolicy": {  
24        "$type": "FullRefreshPolicy",  
25        "location": "Sales.csv"  
26      },  
27      "attributes": [  
28        {  
29          "name": "SalesOrderNumber",  
30          "dataType": "string"
```

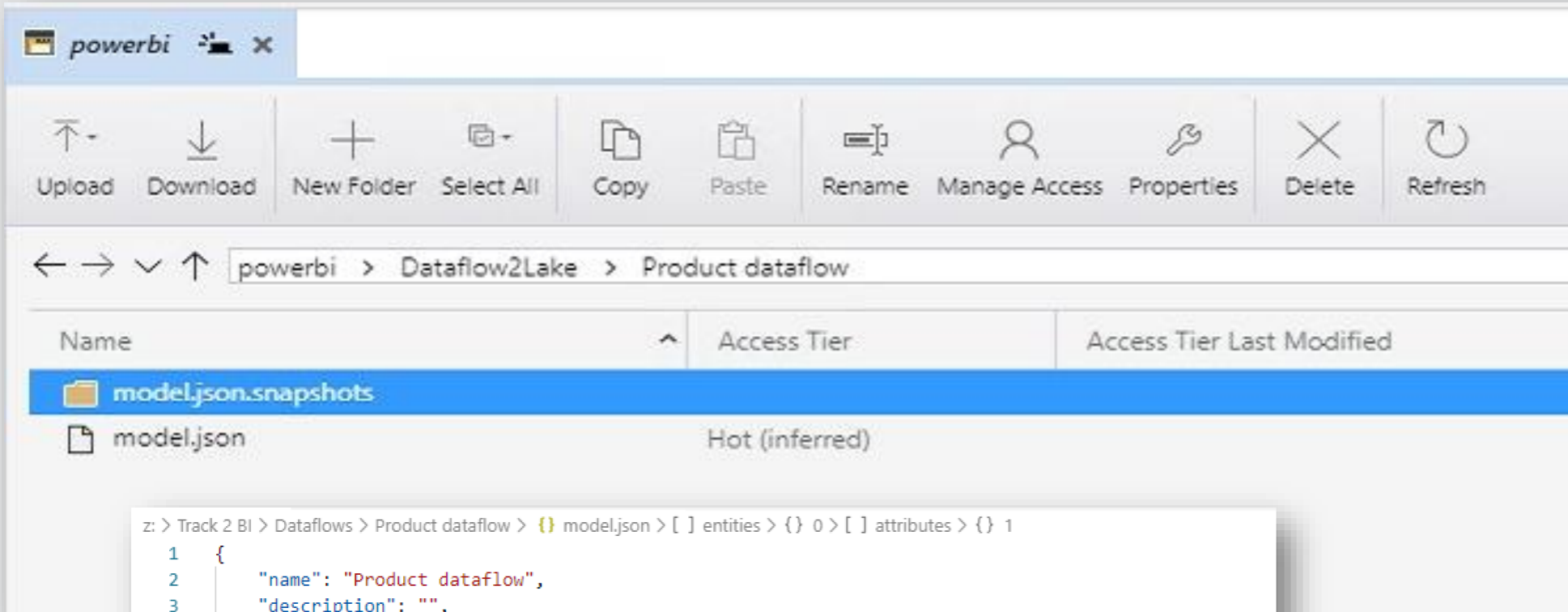
# Fact Sheet – PBI Dataflows

## 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 > {} 1
1 {
2   "name": "Product dataflow",
3   "description": "",
4   "version": "1.0",
5   "culture": "en-US",
6   "modifiedTime": "2019-12-02T20:34:28.5677753+00:00",
7   "pbi:mashup": {
8     "fastCombine": false,
9     "allowNativeQueries": false,
10    "queriesMetadata": {
11      "Products": {
12        "queryId": "eed1d3a0-543f-48fd-bd2a-e5938d922b08",
13        "queryName": "Products",
14        "loadEnabled": true
15      }
16    },
17    "document": "section Section1;\r\n\nshared Products = let\r\n\n Source = Sql.Database(\"passcan
18  },
19  "entities": [
```

### Dataflow storage (preview)

Your Azure Data Lake Storage account is connected to Power BI

datalake4cdm

Allow workspace admins to assign workspaces to this storage account [Learn more](#)

Allow

Apply Cancel

### Settings

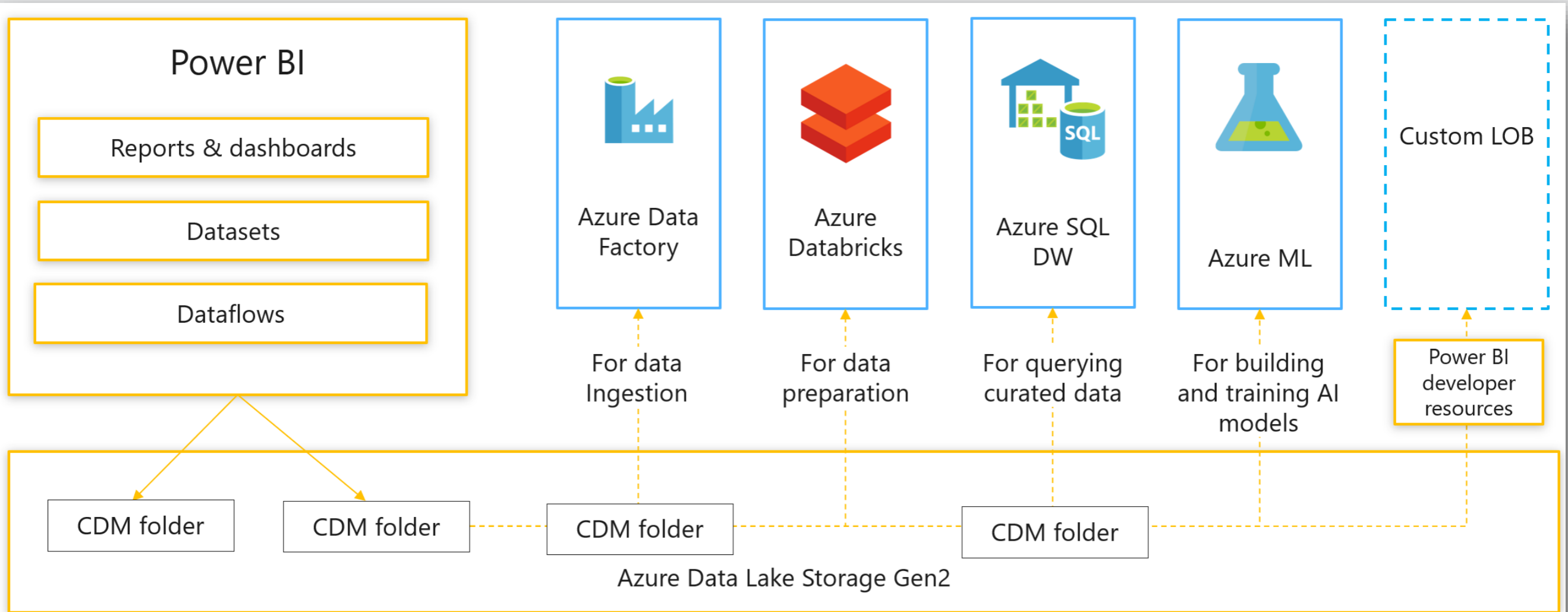
#### DIAD

About Premium Azure connections

Dataflow storage (preview) ⓘ

On

# Power BI Dataflows and the Data Lake



Business analysts

Low/no code

Data scientists  
Data engineers

Medium to high code

# DEMO

## Consume Dataflow Entities

# Get Data

Search

- All
- File
- Database
- Power Platform**
- Azure
- Online Services
- Other

## Power Platform

- Power BI datasets
- Power BI dataflows**
- Common Data Service
- Power Platform dataflows

# Navigator

Search

Display Options

- Power BI dataflows [9]
  - Bauer BI 2019
  - dataflows CPH
  - Dataflows demo (Premium) [8]
    - Combined OrdersInformation
    - Customer Comments
    - Date dataflow
    - Dim Date
    - OrderStatistics
    - Premium DataFlow Linked Plus Computed
    - Sales Combined [3]
      - CalendarTable
      - Sales
      - Sales Combined**
    - Sales dataflow
  - Dataflows Demos [9]
  - dataflows HQ
  - DemoWorkspace
  - PASSCamp Wolfgang
  - PASSCamp Wolfgang DEMO

## Sales Combined

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

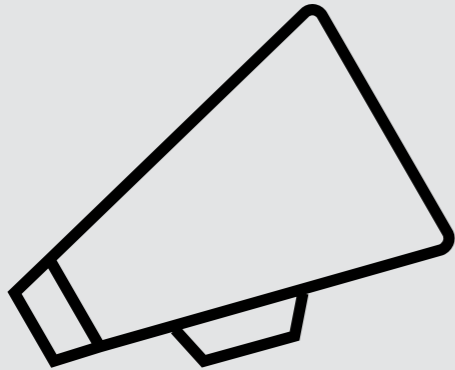
Load

Transform Data



# 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*

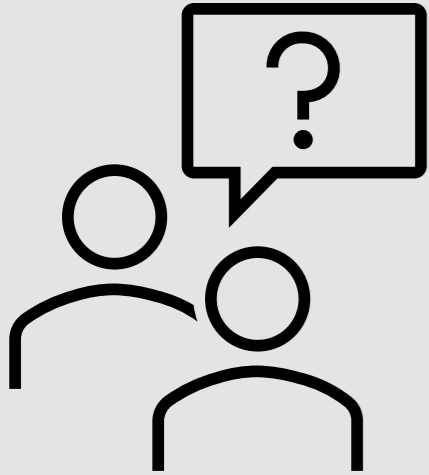
# 80%

“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**

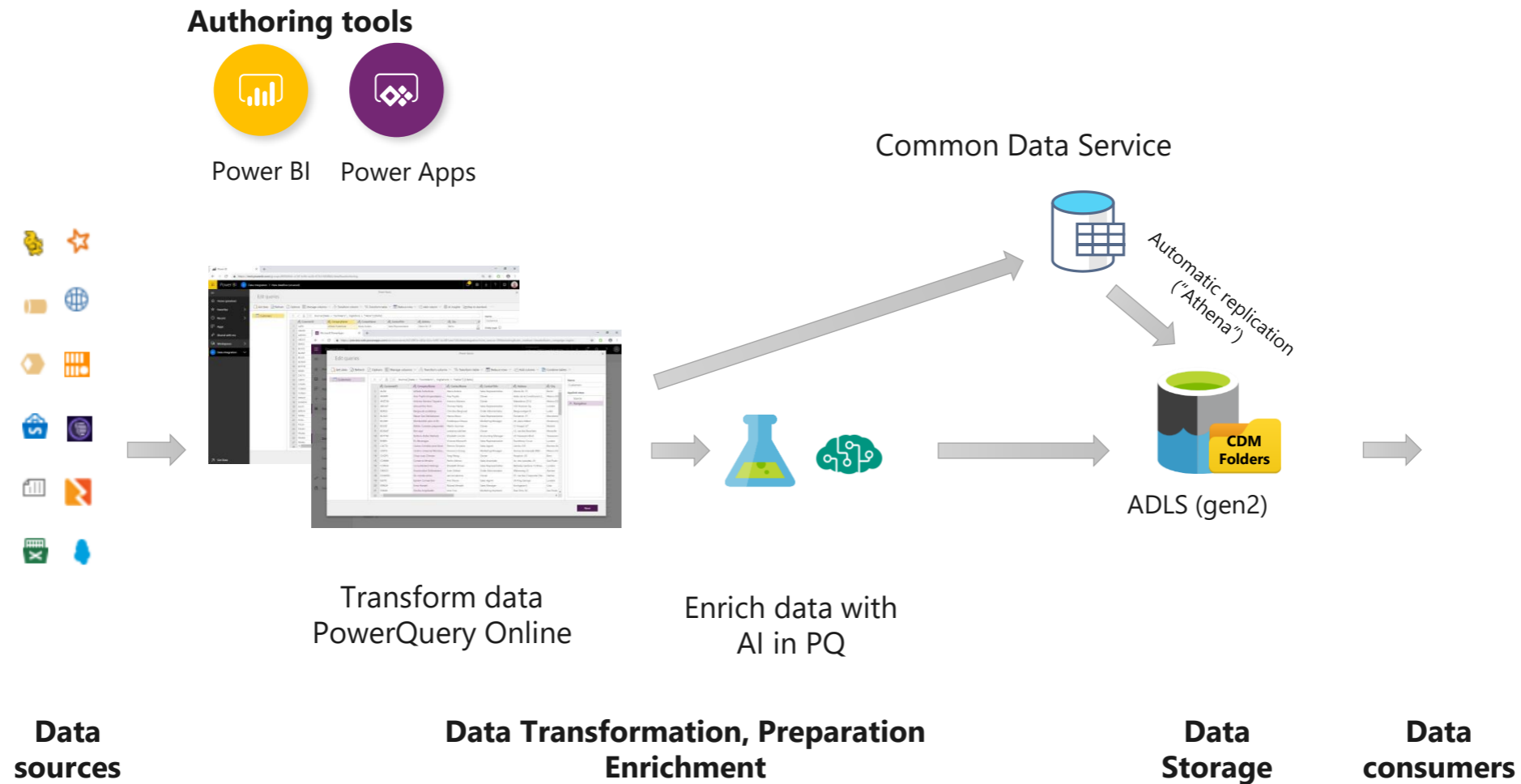


**Interactive & intuitive**

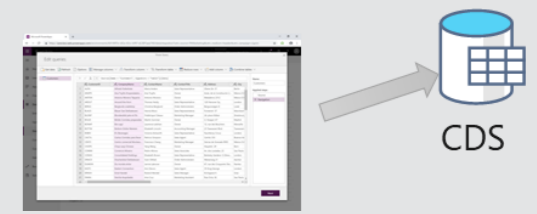
**Consistency of experience**

**Semantic understanding and data consistency**

# Power Platform Dataflows overview



# Load Data into CDS



Data

Entities

Option Sets

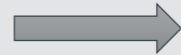
**Dataflows**

Export to data lake

Connections

Custom Connectors

Gateways



### New Dataflow

A diagram showing a data source (cylinder) connected to a cloud, which is then connected to three processing blocks (purple, envelope, and person icons).

Name \*

First DataFlow

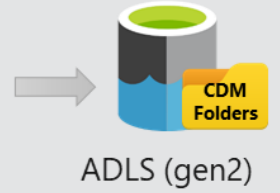
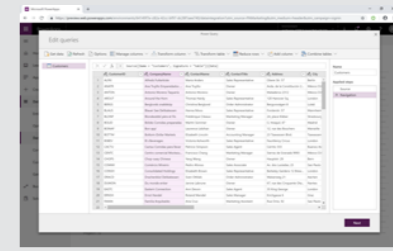
Options

Analytical entities only ⓘ

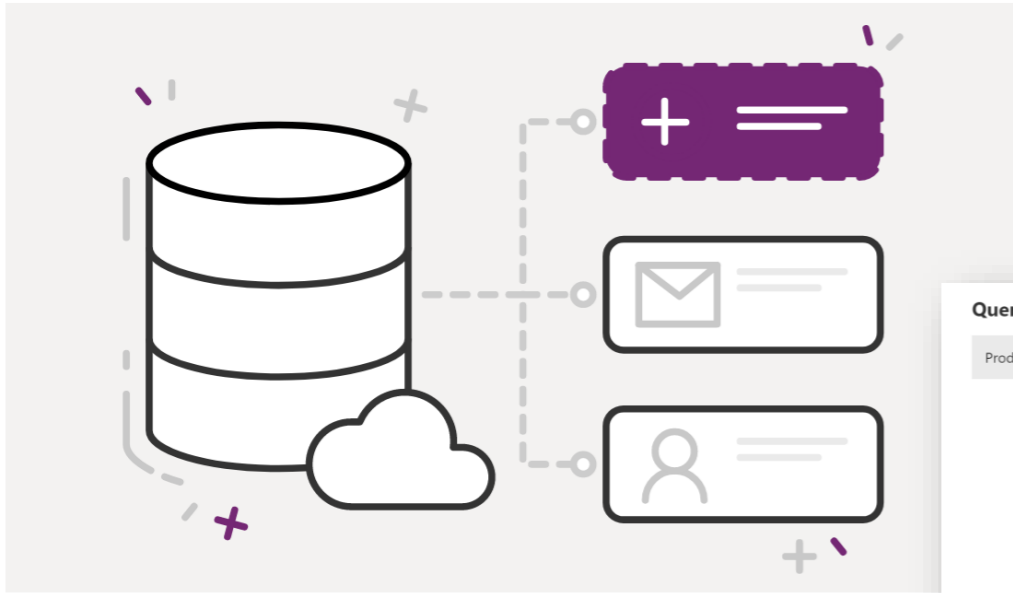
**Create** **Cancel**

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

# “Standard” versus Analytic Entities



New Dataflow



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

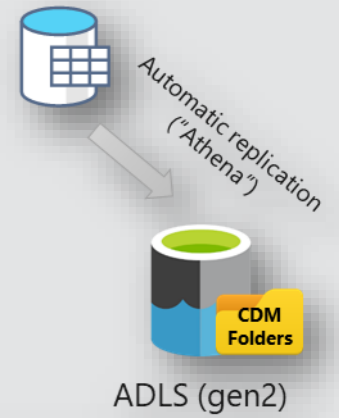
Dataflows allow users to extract, transform and load data from a wide range of data sources that can be used from Power Apps, Flow and analytics. [Learn more](#)



# DEMO

## Power Platform Dataflows

# Export to Data Lake – Project “Athena”



Data

Entities

Option Sets

Dataflows

Export to data lake

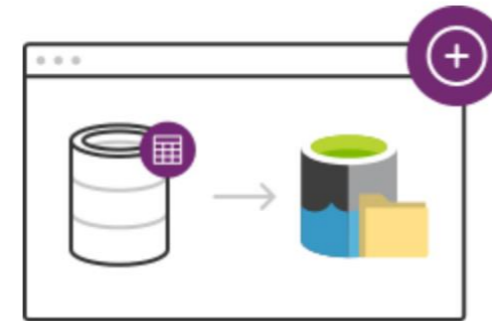
Connections

Custom Connectors

Gateways

+ New link to data lake ← Import from solution ↻ Refresh

Export to data lake



You haven't linked the Common Data Service environment to a data lake  
Before you can export to data lake, link the Common Data Service environment to a data lake

New link to data lake

# 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

**i** The current combination of storage account kind, performance tier, and replication type is not supported. [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 storage account must be in 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 Export to Data Lake service additional roles to your storage account. By using the Export to data lake service, you agree that your storage account will be subject to the 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.



## Manage entities for wolfgangcdsexportstorage

### ✓ Manage entities

2 of 107 selected



<input type="radio"/>	Approval Request	msdyn_flow_approvalrequest
<input type="radio"/>	Approval Response	msdyn_flow_approvalresponse
<input checked="" type="radio"/>	Article	kbarticle
<input type="radio"/>	Attachment	attachment
<input type="radio"/>	Attachment	activitymimeattachment
<input type="radio"/>	Attribute Map	attributemap
<input type="radio"/>	AttributeImageConfig	attributeimageconfig
<input type="radio"/>	Await All Approval Model	msdyn_flow_awaitallapprovalmodel
<input type="radio"/>	Basic Approval Model Data	msdyn_flow_basicapprovalmodel
<input checked="" type="radio"/>	Business Unit	businessunit
<input type="radio"/>	Canvas App	canvasapp
<input type="radio"/>	Category	category

# What's in my Data Lake?



commondataservice-powerofbiatdefaultupgrade-orgbfa74bd0

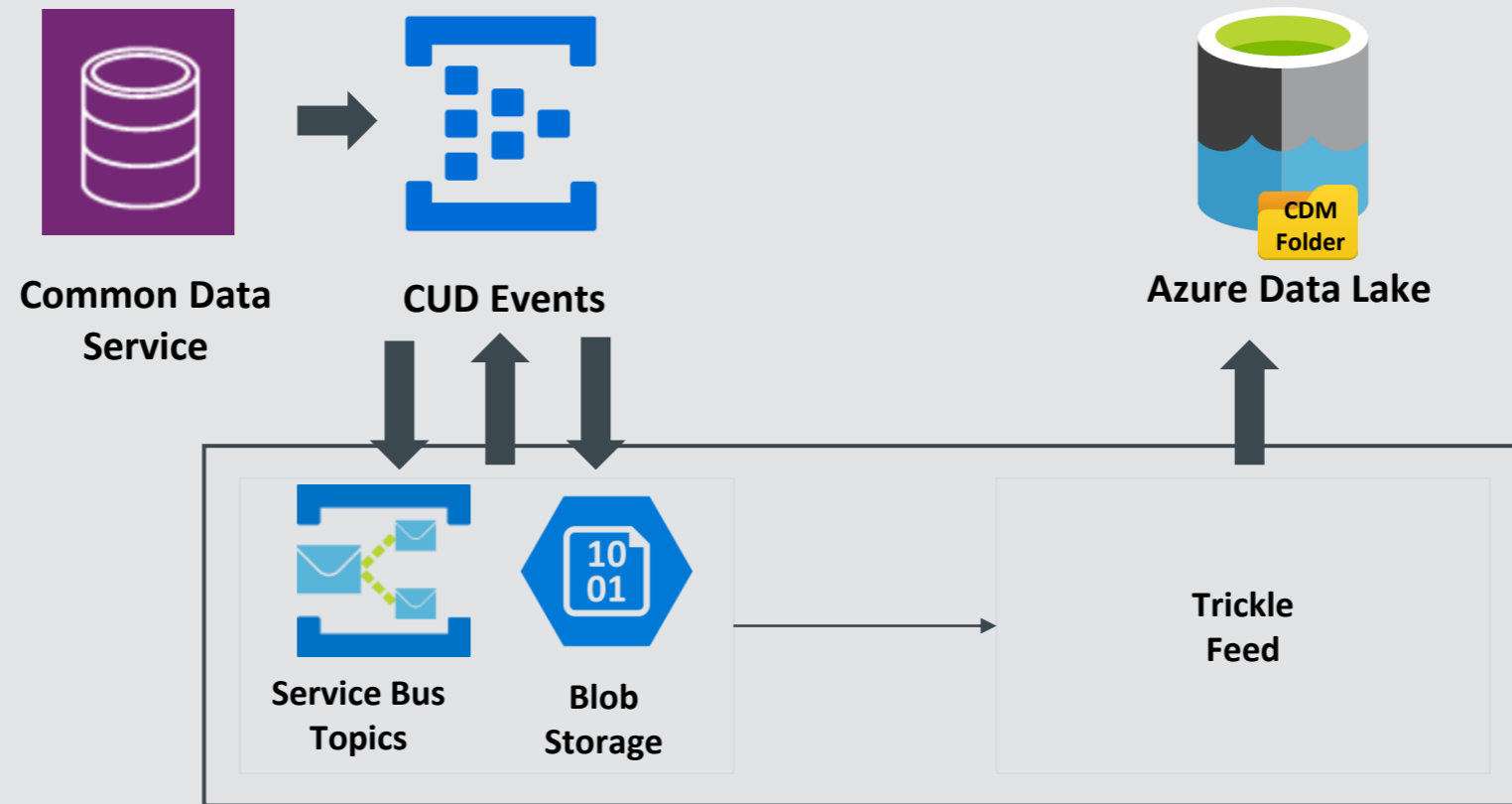
Name

- businessunit
- Microsoft.Athena.TrickleFeedService
- model.json

```
{
  "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"
        }
      ]
    }
  ]
}
```

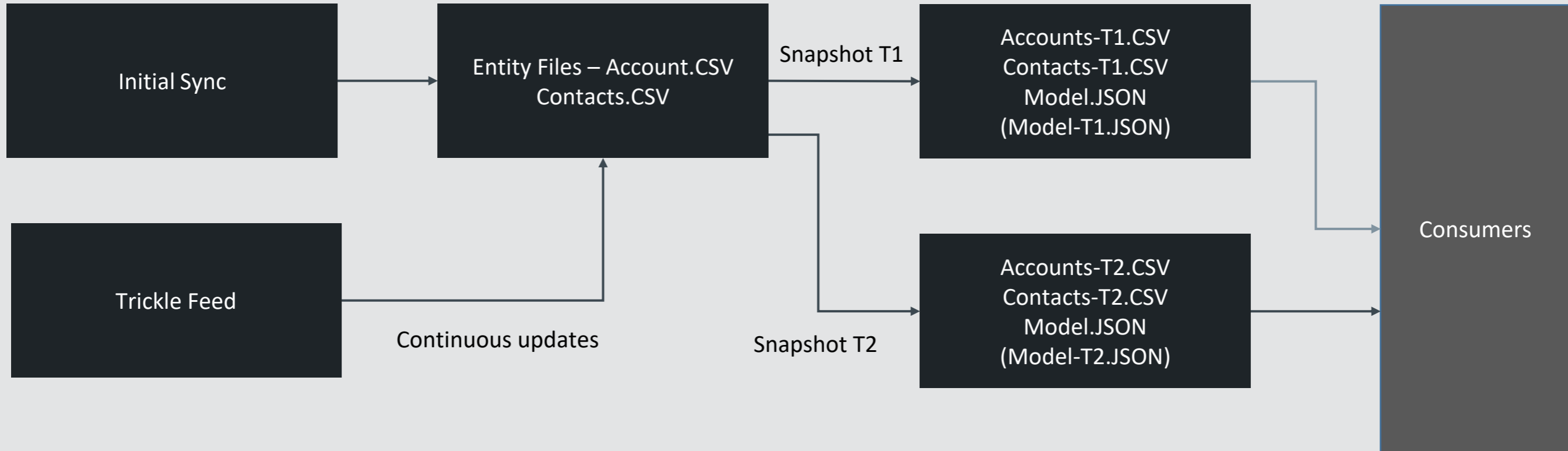
# Behind the Scenes

# Support for **full** and **incremental** **Trickle feed** to write data to ADLS Gen2



Enable analytics on CDS data

# Consistent Data & Metadata Updates





# Consume in Power BI (Option 1)

## Get Data

All

File

Database















Power Platform

Azure

Online Services

Other

### Azure

-  Azure SQL database
-  Azure SQL Data Warehouse
-  Azure Analysis Services database
-  Azure Blob Storage
-  Azure Table Storage
-  Azure Cosmos DB
-  Azure Data Lake Storage Gen2
-  Azure Data Lake Storage Gen1
-  Azure HDInsight (HDFS)
-  Azure HDInsight Spark
-  HDInsight Interactive Query
-  Azure Data Explorer (Kusto)
-  Azure Cost Management
-  Azure Time Series Insights (Beta)

https://wolfgangcdsexportstorage.dfs.core.windows.net/commondataservice-po...

Content	Name	Extension	Date accessed	Date modified	Date created	Attributes	Folder Path
Binary	2019.csv	.csv	null	01.03.2020 18:17:54	null	Record	https://wolfgangcdsexportstorage.dfs.core.windc
Binary	model.json	.json	null	01.03.2020 18:00:27	null	Record	https://wolfgangcdsexportstorage.dfs.core.windc



Combine & Transform Data

Transform Data

Cancel

# Consume in Power BI (Option 2)

## Start creating your dataflow

### Define new entities

Choose a data source to define the entities for your dataflow. You can map your data to [standard Common Data Model](#) entities, or define custom entities instead. [Learn more](#)

Add new entities

### Link entities from other dataflows

Linking to entities from other dataflows reduces duplication and helps maintain consistency across your organization. [Learn more](#)

Add linked entities

### Import Model

Choose a dataflow model to import into your workspace. [Learn more](#)

Import model

### Attach a Common Data Model folder (preview)

Attach a Common Data Model folder from your Azure Data Lake Storage Gen2 account to a new dataflow, so you can use it in Power BI. [Learn more](#)

Create and attach



## Attach a Common Data Model folder to a new dataflow

Name \*

cds

Description

Common Data Model folder path \*

<https://wolfgangstorgen2.dfs.core.windows.net/powerbi/cds/model.json>

Create and attach

Cancel



Dashboards Reports Workbooks Datasets Dataflows

NAME ↑

ACTIONS



cds EXTERNAL

...

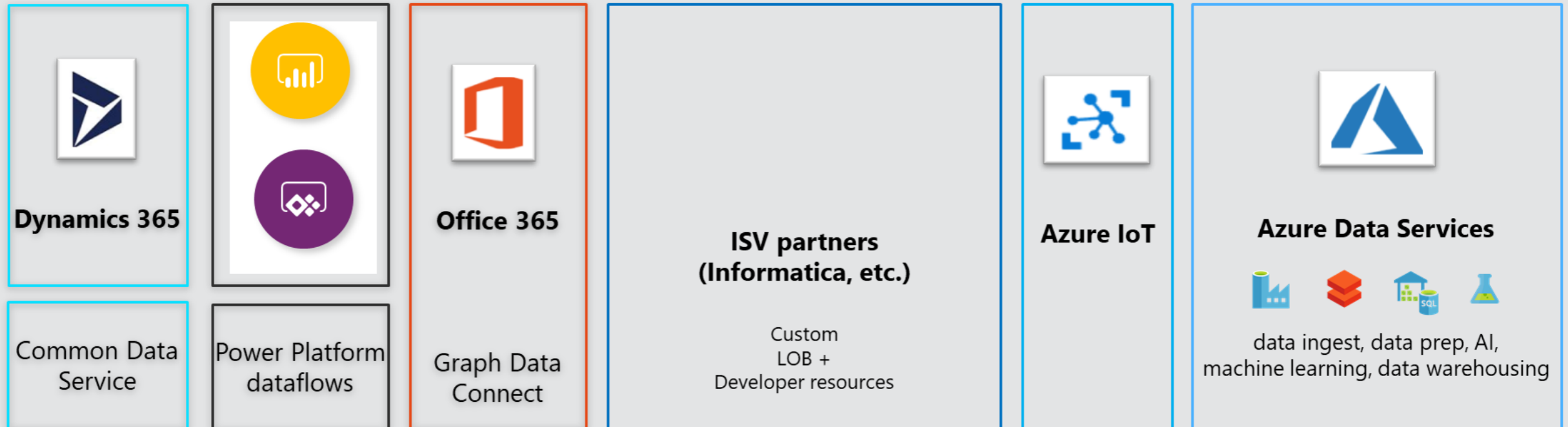
# Power Platform Dataflows and the Data Lake

Decision Makers, Business analysts

Data scientists, Data engineers

No code, low code

Low to high code



Transactional business data



...

Profiles based on observational data



# Power Platform Dataflows...

Data Integration & Preparation in the Cloud

Data Lake

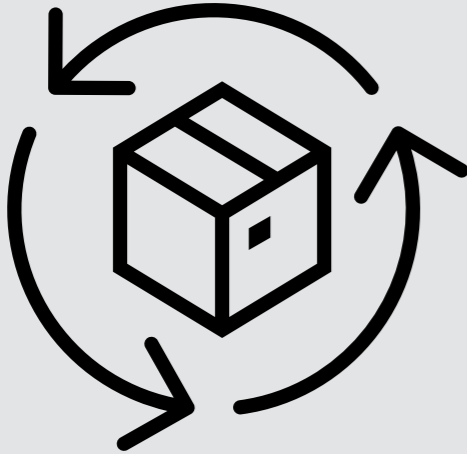
Common Set of Technology

Connectors

Power Query

CDM Folders

Data Exchange made easy

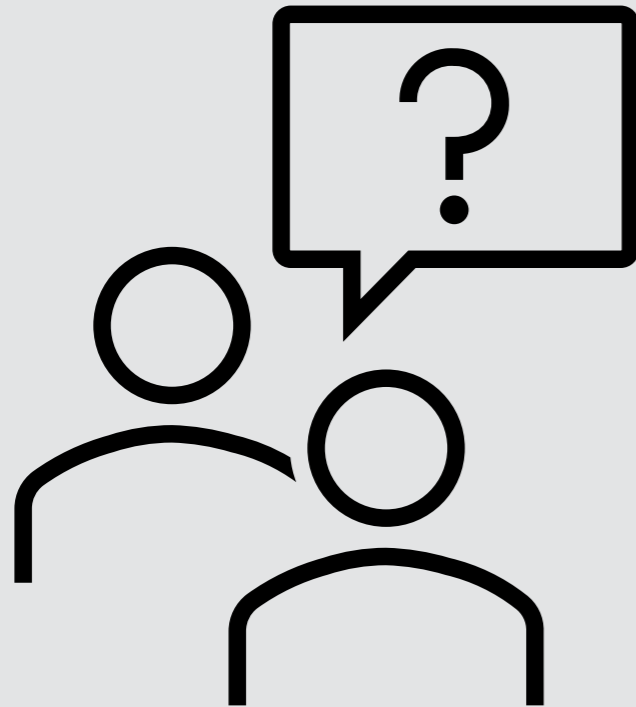


**Wolfgang Strasser**

wolfgang@powerofbi.at

@wstrasser

workingondata.wordpress.com



# 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>

- Power Query / m reference

<https://docs.microsoft.com/en-us/powerquery-m/power-query-m-function-reference>