

GLOBAL POWER PLATFORM BOOTCAMP

Organized Globally, Held Locally



POWER PLATFORM
BOOTCAMP



KingswaySoft

Power BI dataflows deep dive

Riccardo Perico

Lorenzo Vercellati



#GlobalPowerPlatformBootcamp


Thanks To Our Local Sponsor



#GlobalPowerPlatformBootcamp

Riccardo Perico

rperico@solidq.com | @R1k91  | 

Data Platform & BI Specialist @  **SolidQ**
Think Big. Move Fast.

10 years in Microsoft «Data Realm»

MCP: MTA, MCSA  

Speaker at Azure Saturday, SQL Saturday

DIAD tutor



Lorenzo Vercellati

lvercellati@solidq.com | @supergimi  | 

Power BI Solution Architect @



20 years burned down Databases

MCP: MCSA



Speaker at PowerPlatform World Tour, SQL
Saturday and DIAD tutor

PowerBI User Group Milan Administrator



Session Agenda

- Power BI dataflows: let's do a recap
- ABC for CDM
- What they are and what they're not
- 4 ways to create them
- Power BI dataflows premium feature
- Use your Azure Data Lake Storage gen2
- Best practices, Tips & Tricks
- Q&A



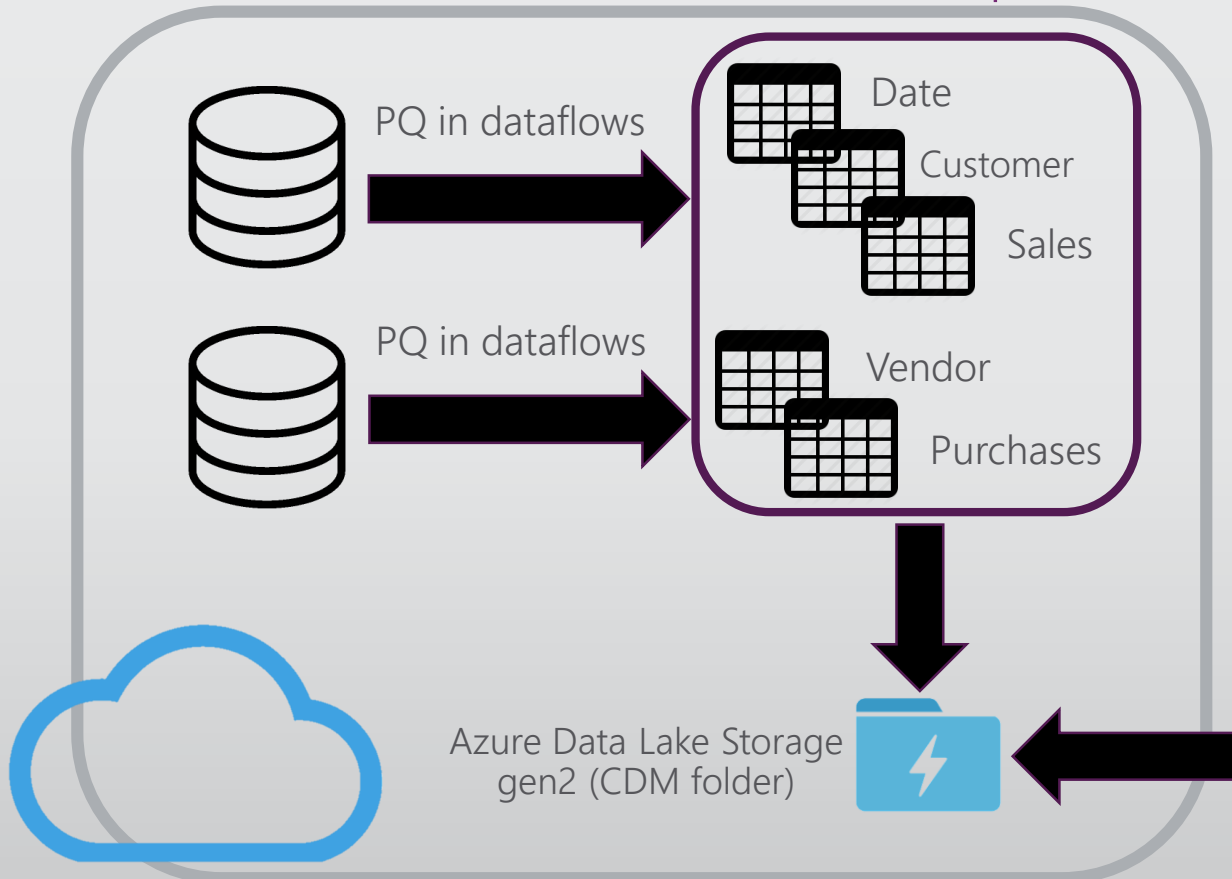
Power BI dataflows: let's do a recap



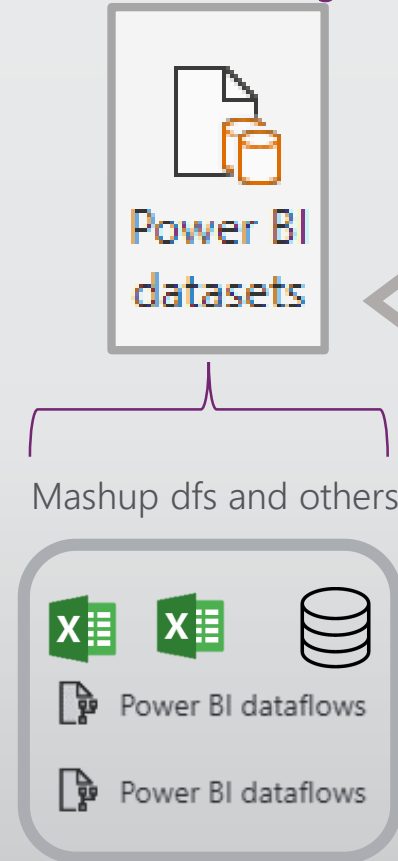
#GlobalPowerPlatformBootcamp

Power BI dataflows: let's do a recap

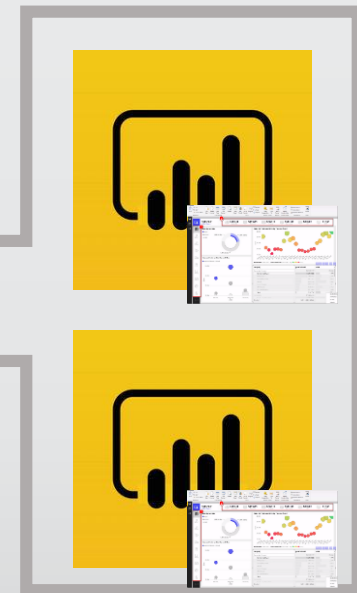
Data Preparation



Data Modeling



Data Presentation



ABC for CDM



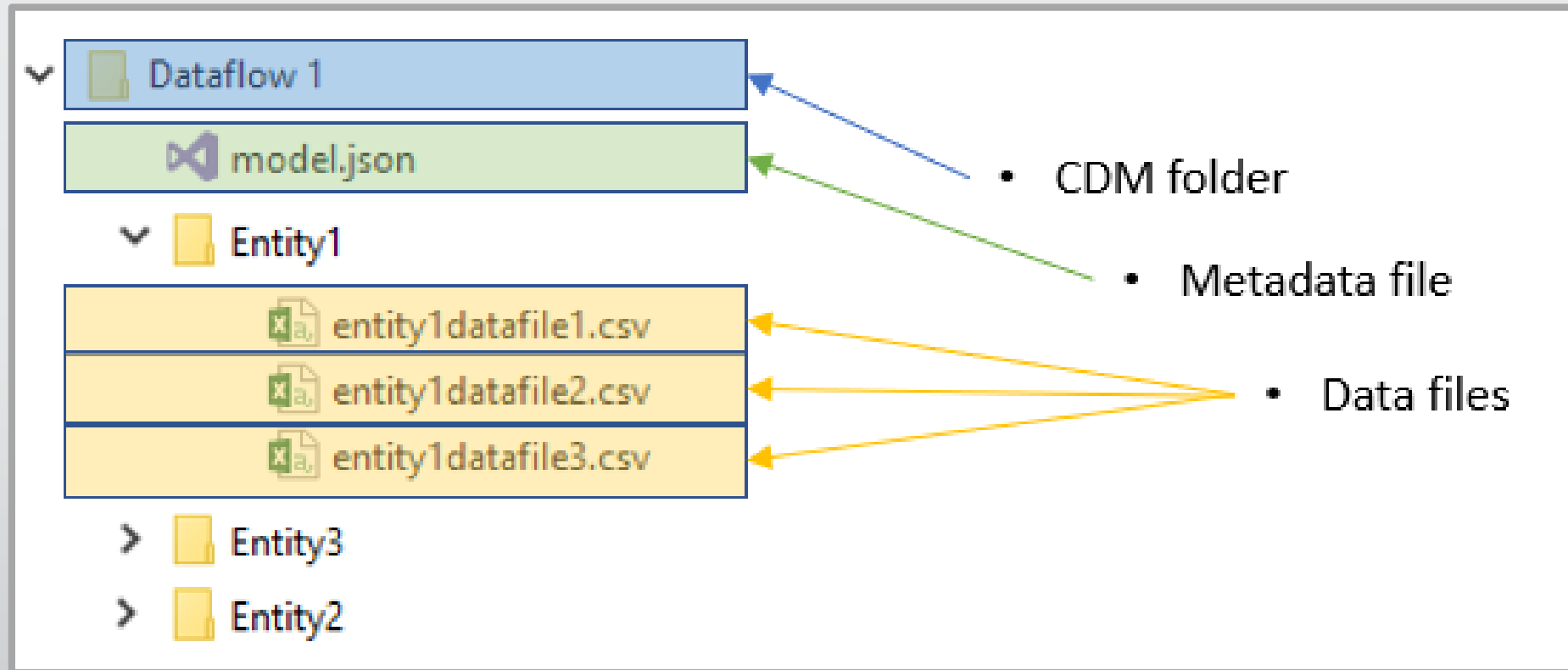
#GlobalPowerPlatformBootcamp

What you should know about CDM

- CDM is a metadata system
- CDM provides a shared language and consistency among different apps (i.e. Power Platform, D365 and Azure)
- Includes a set of **standard entity schemas** (you wouldn't use 9 out of 10 😊)
- CDM Folder is a storage that use CDM metadata



How does CDM folder look like?



What they are and what they're not



#GlobalPowerPlatformBootcamp

They are... They aren't

- ✓ Self-service data preparation tool
- ✓ It's Power Query
- ✓ Store data in CDM folder in ADLS gen2
- ✓ Encourage data reuse
- ✗ Don't replace datasets
- ✗ Don't replace DWH
- ✗ Don't replace orchestrating tools (ADF v2 or SSIS)
- ✗ Not Premium-only feature
- ✗ Don't have additional costs



dataflows Vs datasets

<https://ssbipolar.com/2019/10/14/power-bite-dataflows-vs-datasets/>

	Power BI dataflows	Power BI datasets
Implementation	CDM folder	Analysis Services Tabular model
Storage	CSV files	Vertipaq
Metadata	CDM – model.json	BISM
Development	Browser	Power BI Desktop
Primary purpose	Data preparation reuse	Data model reuse
Reuse	Acts as data source in multiple datasets	Shared dataset across workspace
Scope of reuse	Entity level reuse	Dataset level reuse
Mashup with other data source	Yes	No
Reuse outside Power BI	Through ADLS gen 2	Through XMLA (only premium)
Data access methods	Import	Import and Direct Query
Storage Mode	Import	Live Connection



#GlobalPowerPlatformBootcamp

Power BI dataflows Pros

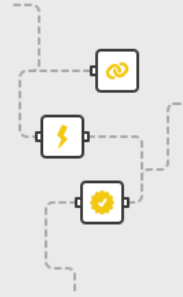
- Centralization of data preparation (reusability and standardization)
- Isolating refresh (1 table a time)
- Attach an existing ADLS gen2 (preview)
- Speed up dataset refresh

4 ways to create them



#GlobalPowerPlatformBootcamp

4 ways to create them



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](#)

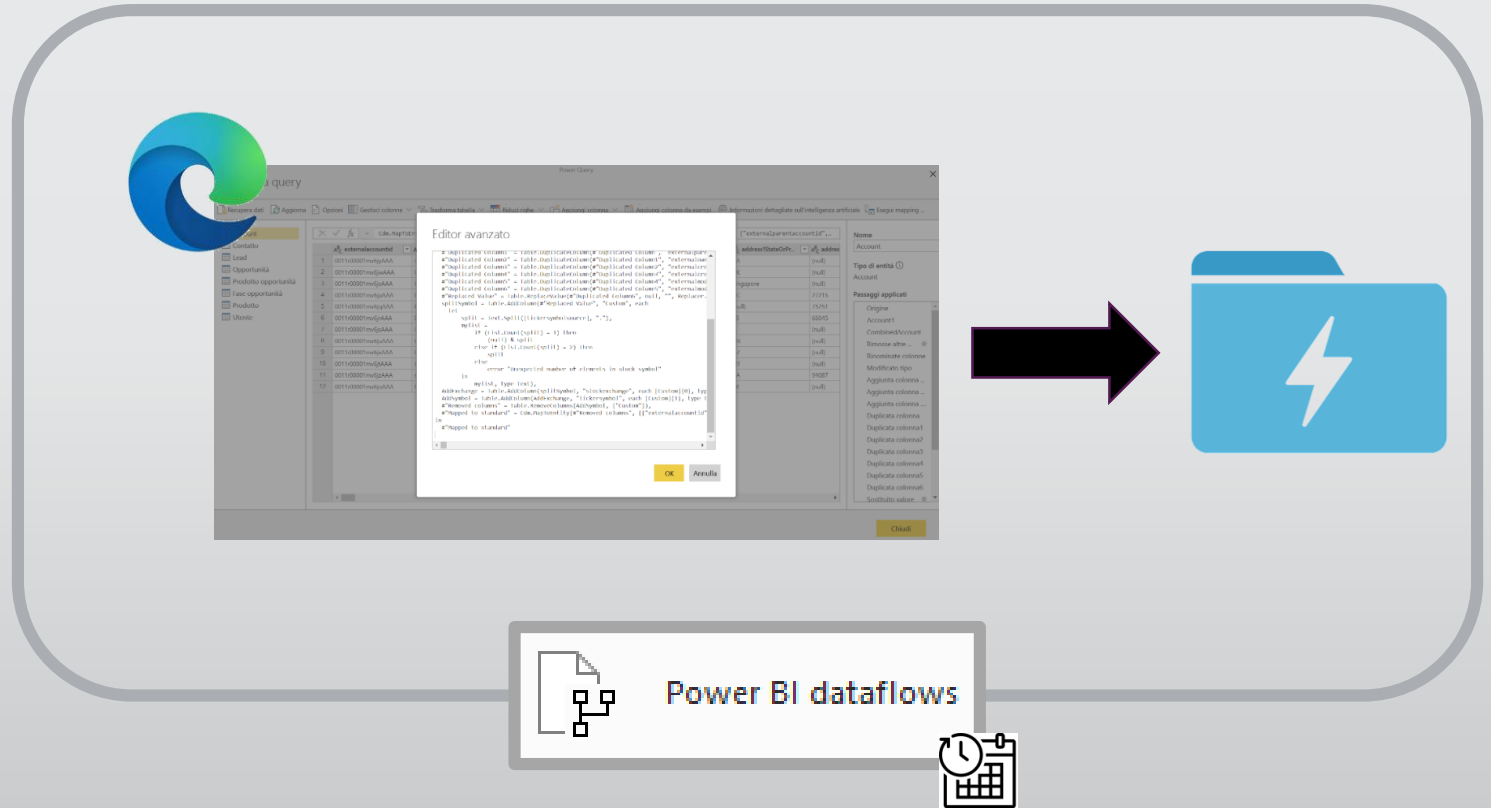
Create and attach

Method 1: Define new entities

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



Demo: Create dataflows method 1



#GlobalPowerPlatformBootcamp

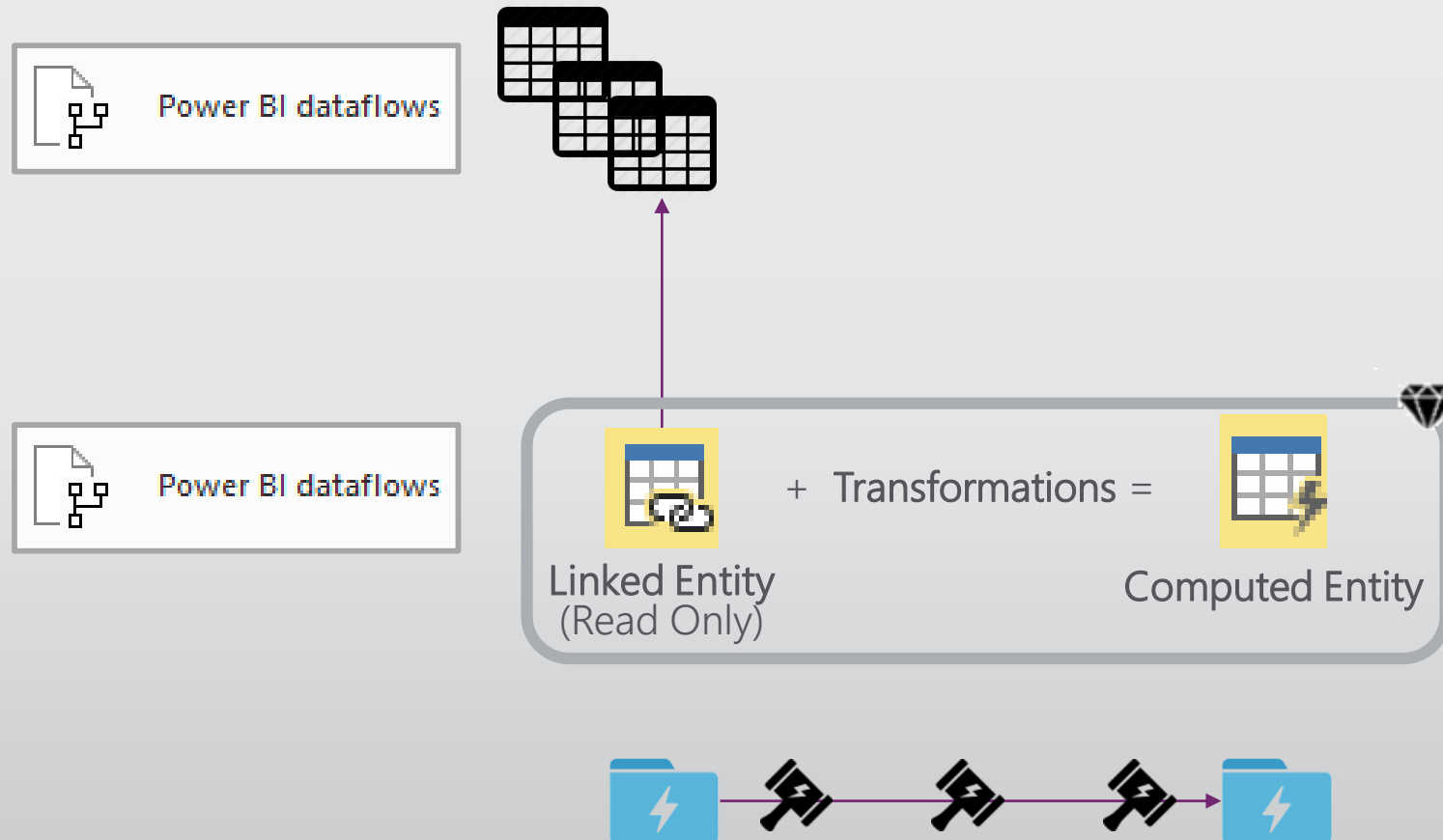
Method 2: Link Entities

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



Method 2: Refresh Matrix

	Same Workspace	Different Workspace
Refresh Source	Source and Destination	Source Only
Refresh Destination	Destination Only	Destination Only

Enable Load

Tasks Done	Resulting Entity	Enable Load on Source(s)	Premium Required	Keep in mind
Entity + Transformation	Computed	Yes	Yes	
Entity + Transformation	Computed	No	No	Source Entity won't be materialized in ADLS gen2
Referencing another dataflows' entity	Linked	Yes	Yes	
Referencing another dataflows' entity	Linked	No	No	Source Entity won't be available in the destination dataflow

Demo: Create dataflows method 2



#GlobalPowerPlatformBootcamp

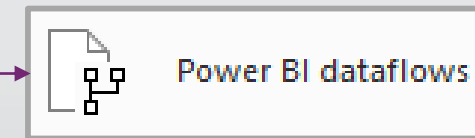
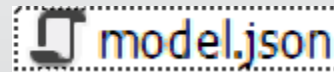
Method 3: Import Model

Import Model

Choose a dataflow model to import into your workspace.

[Learn more](#)

Import model



Demo: Create dataflows method 3



#GlobalPowerPlatformBootcamp

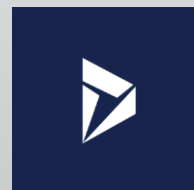
Method 4: Attach a CDM folder (preview)

Attach a Common Data Model folder (preview)

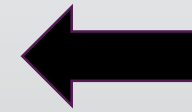
Attach a Common Data Model folder from your Azure Data Lake Storage Gen2 account to a new dataflow, so you can use it in Power BI.

[Learn more](#)

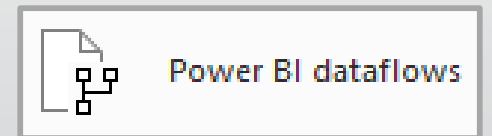
Create and attach



Create



Use



Power BI dataflows Premium Features (\$\$\$\$)



#GlobalPowerPlatformBootcamp

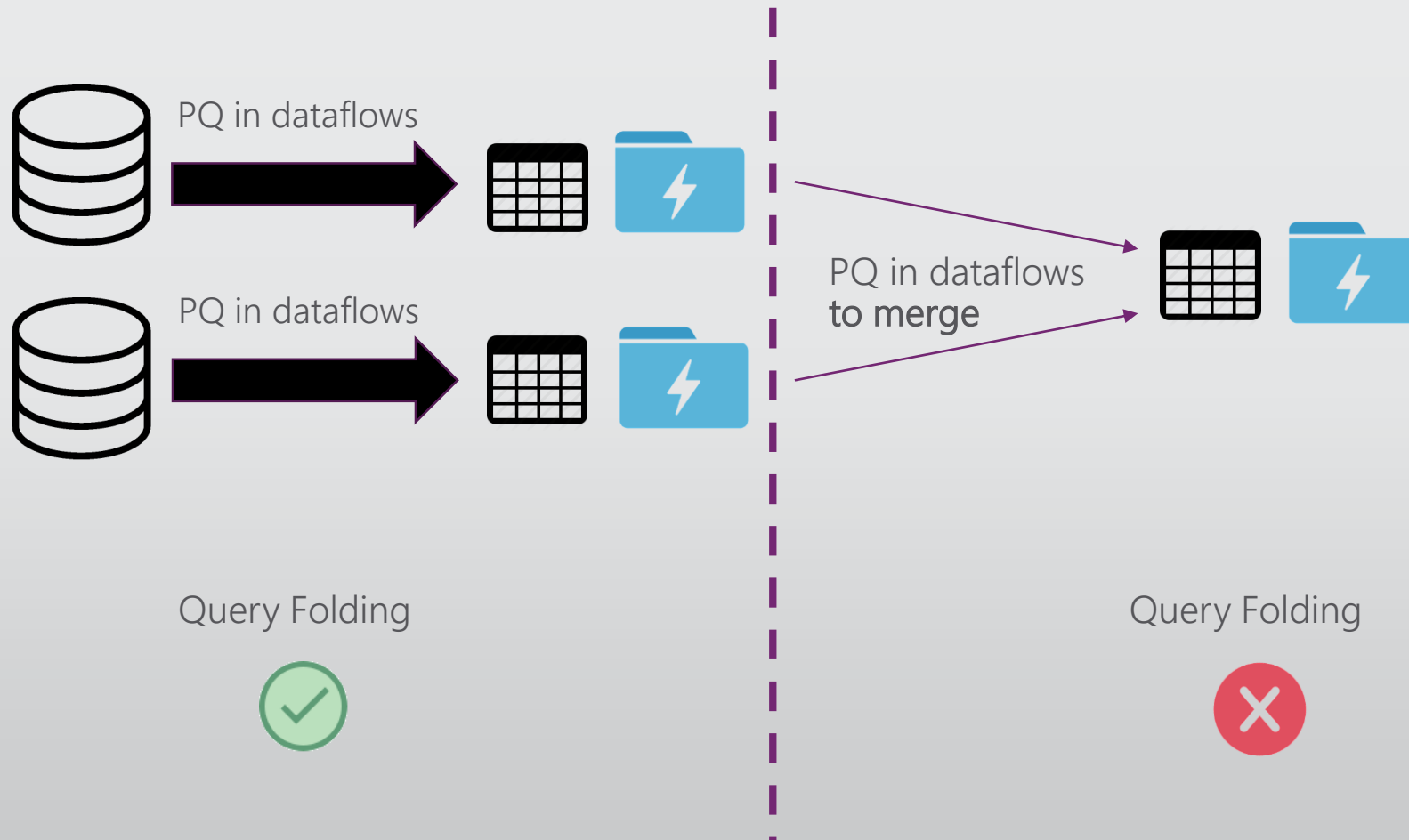
\$\$\$\$: Computed Entities & Linked Entities

- We've already talked about it

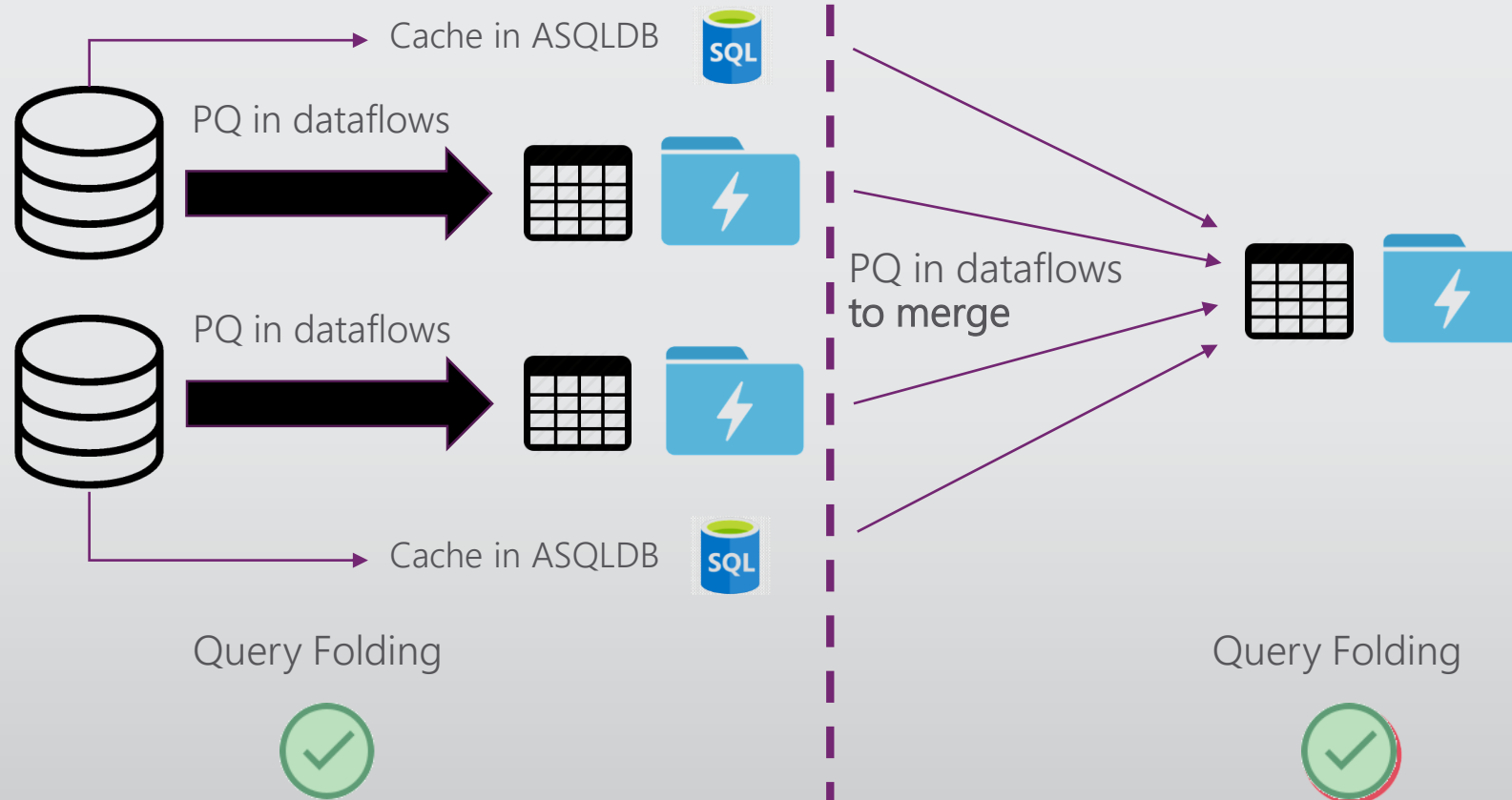
Do You remember?



\$\$\$\$: Enhanced compute engine, why?



\$\$\$\$: Enhanced compute engine, what?



\$\$\$\$: Enhanced compute engine, how?

DATAFLOWS - Active
Your workload is ready to use.

☒ On

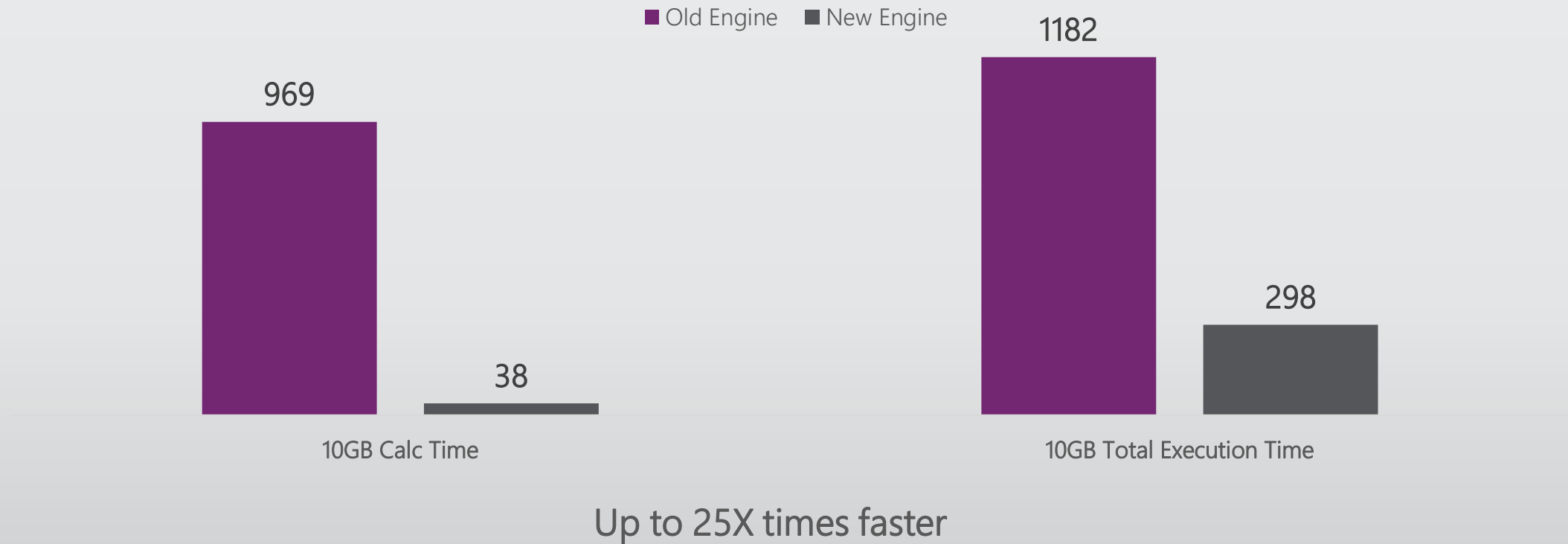
Max Memory (%)

Enhanced Dataflows Compute Engine (Preview)

☒ On

Container Size (Mb)

\$\$\$\$: Enhanced compute engine, so?



**Note: loading data is slower as the data is being written twice*

\$\$\$\$: AI Insights



Sentiment Analysis
Key Phrase Extraction
Language Detection
Image Tagging



Call custom **Azure ML** models
adding insights to your data

\$\$\$\$: Incremental Refresh

Incremental refresh settings

Appointment

Incremental refresh updates only data that's changed, to speed refresh, reduce capacity usage, and store historic data. [Learn more](#)

☒ On

Choose a DateTime field to filter by

createdon ▼

Store rows from the past

5 Years ▼

Refresh rows from the past

10 Days ▼

☐ Detect data changes [Learn more](#)

Only refresh data if the maximum value in this field changes

Choose a field ▼

☐ Only refresh complete [Learn more](#)

When you save these settings, data from the past [storage period] will be loaded to your dataflow storage the next time this dataflow is refreshed. Subsequent refreshes will update only data that's changed in the past [refresh period].

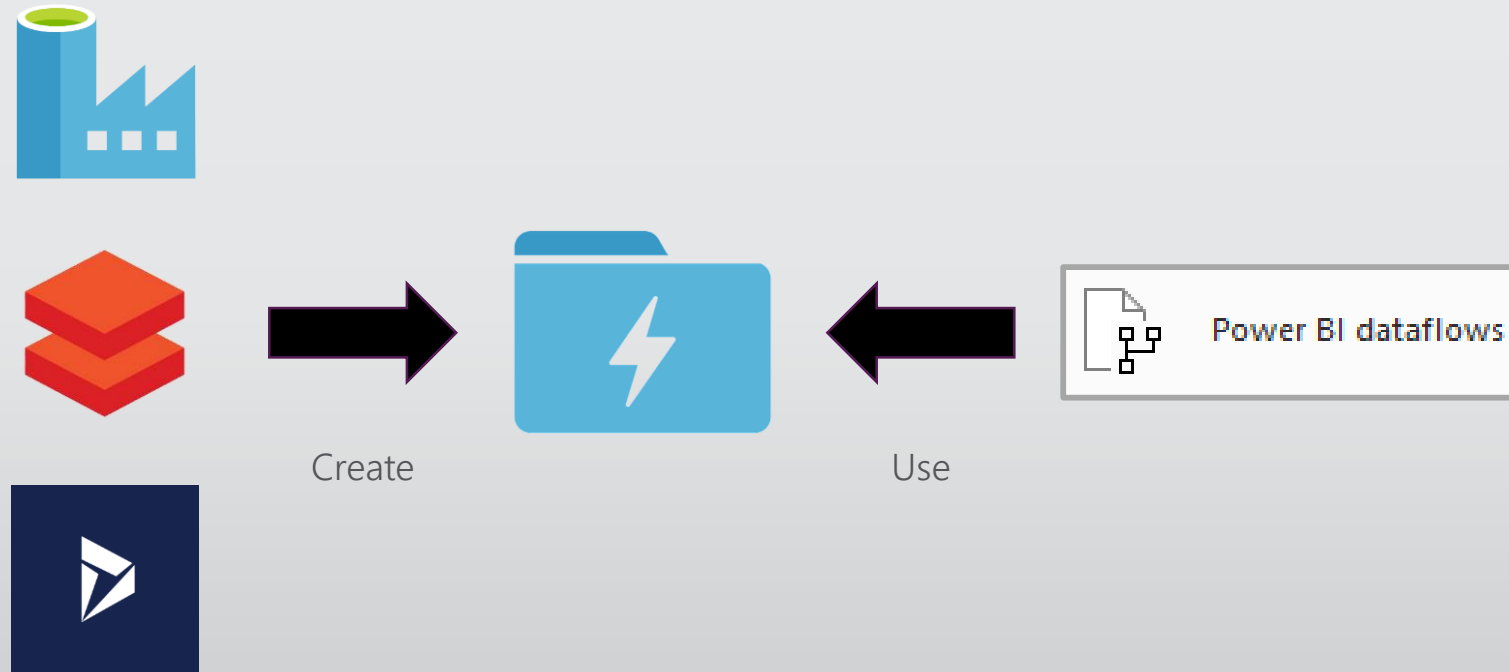
Save Cancel

Use your Azure Data Lake Storage gen2



#GlobalPowerPlatformBootcamp

Use your Azure Data Lake Storage gen2 (preview)



Best Practices, Tips & Tricks



#GlobalPowerPlatformBootcamp

Best Practices

- Split ingestion and transformation (\$\$\$\$)
- Create a separate workspace for ingestion and transformation (\$\$\$\$)
- Do not break query folding (\$\$\$\$)
- Avoid scheduling refresh for linked entities inside the same workspace (\$\$\$\$)
- Separate on-premises sources from cloud sources into different dataflows



Tips & Tricks

- Power BI Gateway is supported for On-Prem sources
- All queries with the **same Gateway**
- Use “**Blank Query**” source for missing transformation UIs



Useful Links

- dataflows & CDM: <http://tiny.cc/5d9uaz>
- Dataflows best practices: <http://tiny.cc/dmgebi>
- Shared Datasets replace SSAS: <http://tiny.cc/8imtbz>
- Datasets & dataflows, my thoughts: <http://tiny.cc/pbidfdd1>

Any Questions?



#GlobalPowerPlatformBootcamp

Help us make better!



#GlobalPowerPlatformBootcamp

**Please fill
out the
survey!
& Win
Swags!!**



<https://www.powerplatformbootcamp.com/survey>



#GlobalPowerPlatformBootcamp

Thank You For Attending

rperico@solidq.com

@R1k91

lvercellati@solidq.com

@supergimi



#GlobalPowerPlatformBootcamp