

Riccardo Perico



EmPower your BI Architecture

Organizers



Sponsors



Riccardo Perico

rperico@solidq.com | @R1k91  | 

Data Platform & BI Specialist @

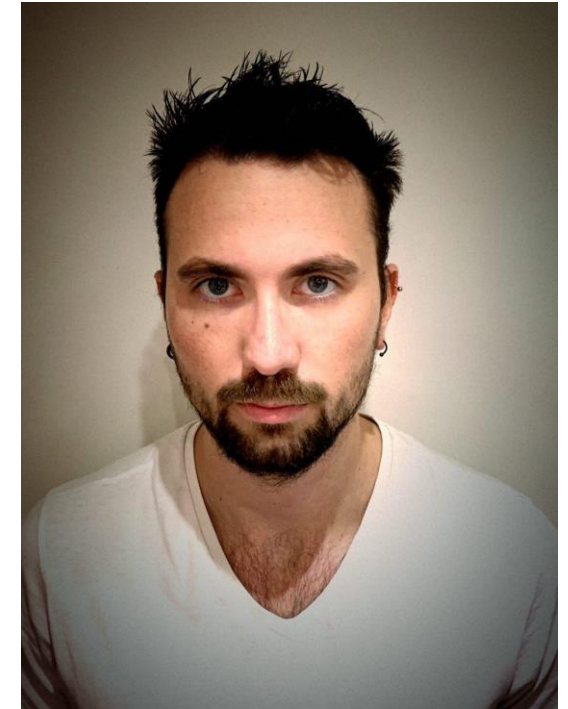


10 years in Microsoft «Data Realm»

MCP: MTA, MCSA



Speaker at Azure Saturday, SQL Saturday and DIAD tutor

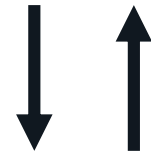
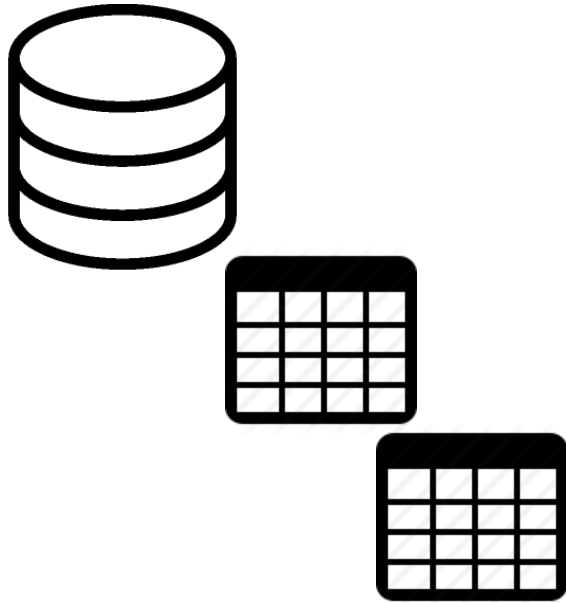


Agenda? The Right Questions

1. What's **Storage Mode** and which one best suits to me?
2. Do I need an On-premises Data **Gateway**?
3. How to **refresh** my data?
4. Is it **just on-cloud** solution?
5. Can I **«encapsulate»** PowerBI?
6. Do I need to care about **dataflows**?
7. Can I have some **advisable scenarios**?

What's Storage Mode and
which one best suits to me?

Storage Mode: Import Mode



Data Model Cached into RAM

Storage Mode: Import Mode

Model **Cached** inside Power BI Service

High **Compression** thanks to VertiPaq (*)

Full **Power BI functionalities and DAX** functions support




Multiple Data Sources mashup (all available)

10 GB model size for PREMIUM / **1 GB** for PRO



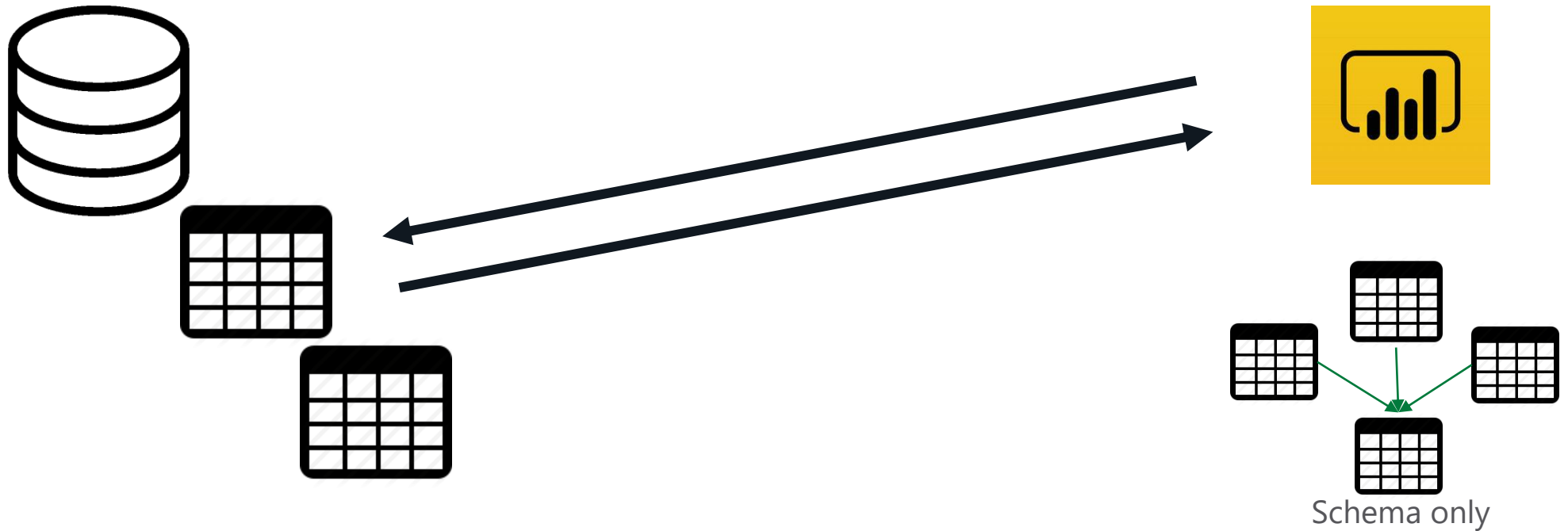
Large models for Power BI Premium



S9V2 Standard ★	
1,280	Query Processing Units
	Up to 400 GB Cache
	Dedicated service
	SSL
20,629.41	
EUR/MONTH (ESTIMATED)	

** Up to available memory on your capacity, excluding memory required for query execution and processing, which is often 20-30% of total memory.*

Storage Mode: Direct Query



Storage Mode: Direct Query

Limited data sources

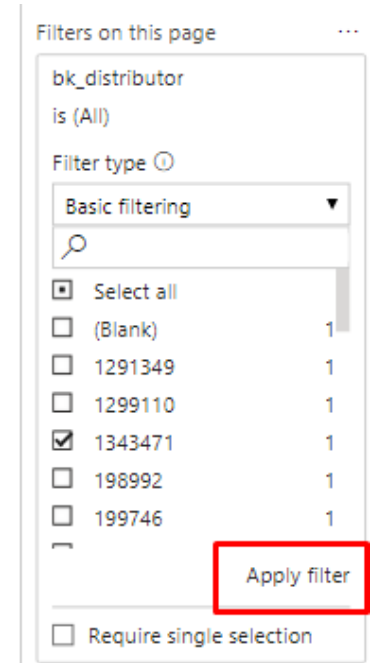
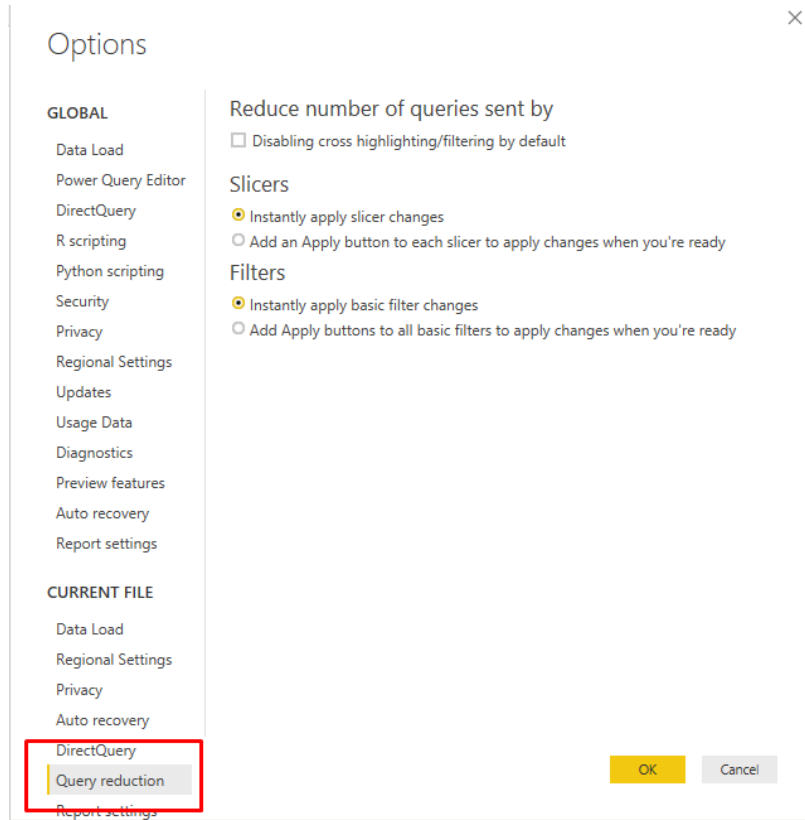
Not full **Power Query** and **DAX** support

Pay attention to **performance** (1 render = 1 query)

Overcome model **size limits**

Real-time / Near Real-time scenarios

Direct Query: Quick Tip



Storage Mode: Live Connection



PowerBI acts just like an interface

Storage Mode: Live Connection

Data Sources: **AAS**, **SSAS** Tabular/Multidimensional and **PBI dataset**
Data **Preparation** and Data **Modeling** done in the source
Interaction with report = **DAX query sent to source**

Overcome model **size limits**

Partitioned Load

Tends to be **faster than DQ**

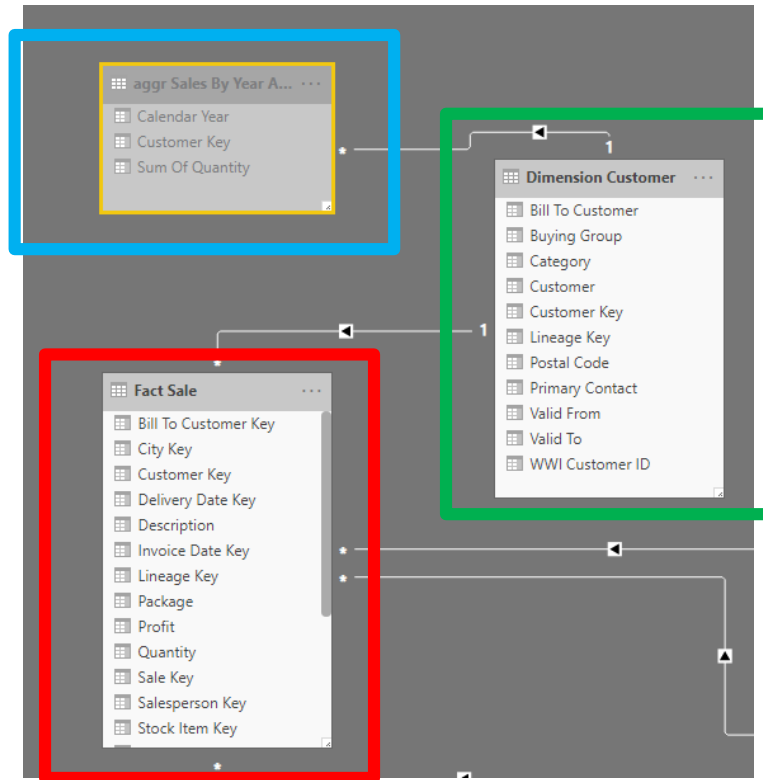
Never prefer «Import» over «Live Connection» for supported sources

Storage Mode: Composite Mode

Import Mode

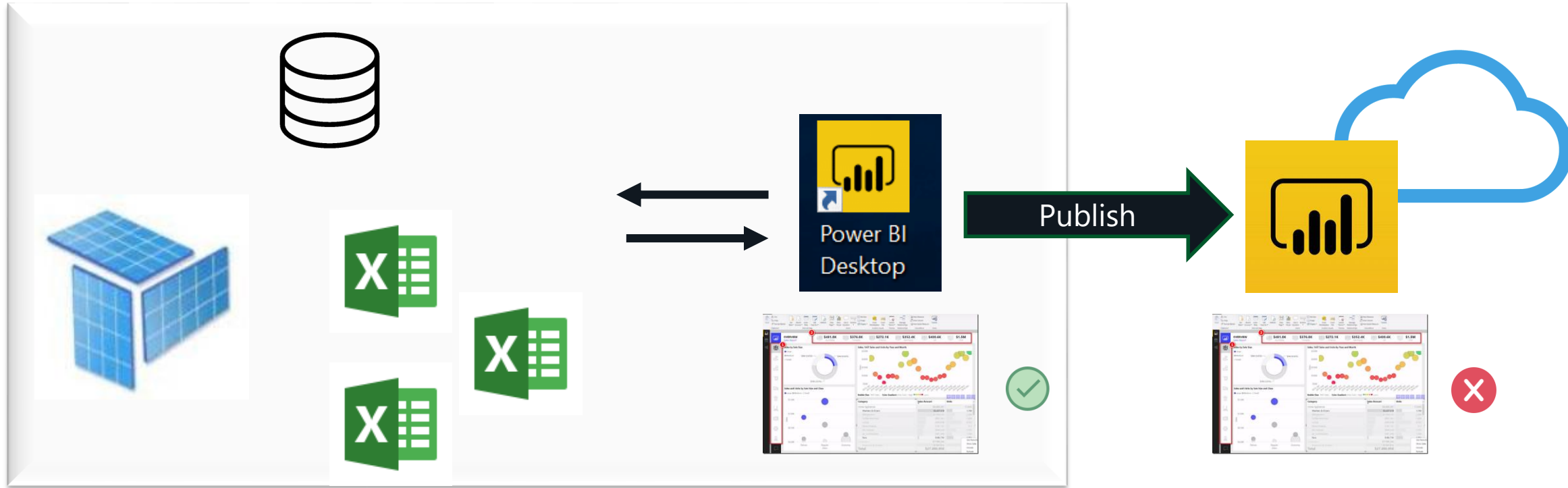
Direct Query

Dual Mode

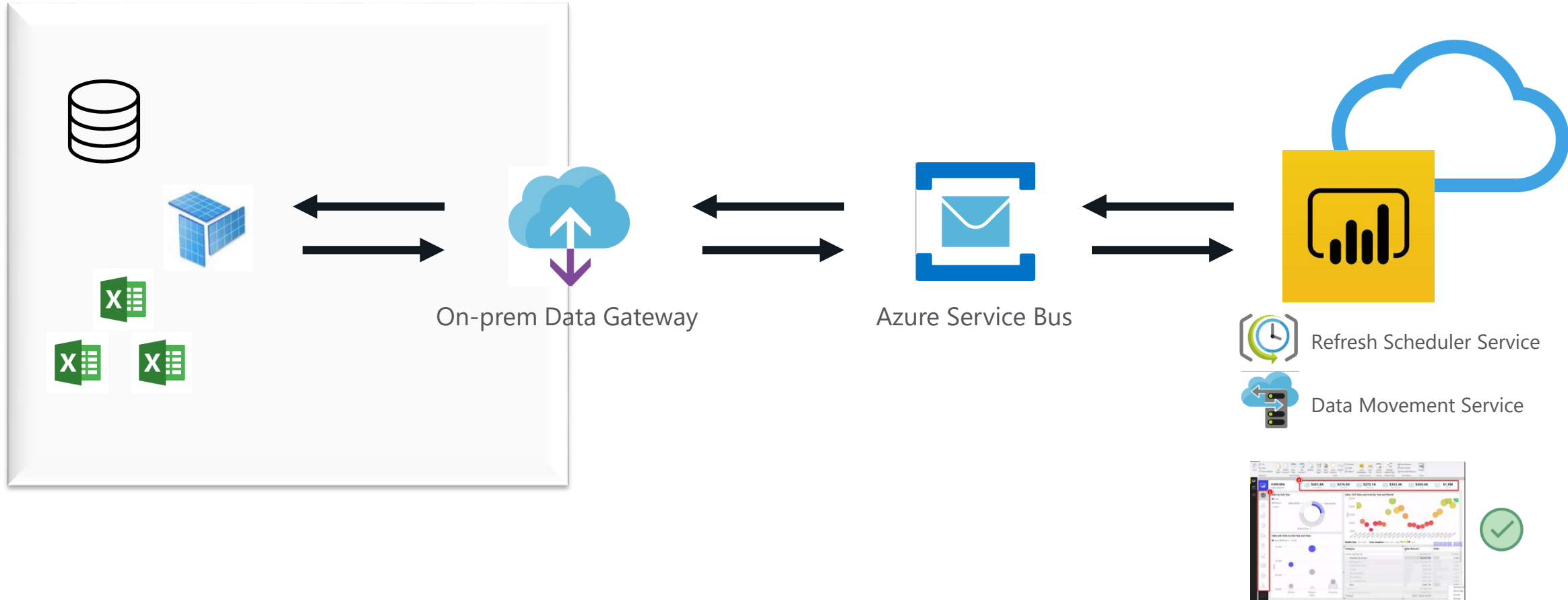


Do I need an On-premises Data Gateway?

Let's draw a typical scenario

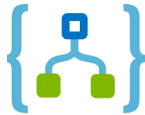
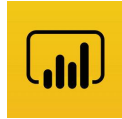
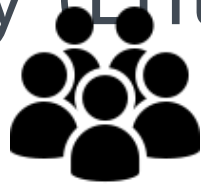


On-Prem Data Gateway is the way



2 Gateway types

On-premises data gateway (Enterprise)



All Storage Mode Support
Scheduled Refresh Support

On-premises data gateway (Personal mode)



Import Mode Support
Scheduled Refresh Support

Gateway: Architecture security

Credentials stored locally and in Power BI Service

Data Source credentials encrypted and stored in the cloud only gateway can decrypt

Gateway polling → No inbound port to open

Outbound port to open TCP: 443, 5671, 5672, 9530...9534

Gateway: Considerations

Keep gateways' name/**recovery key** in a safe place
(migration, take over and recovery)

Do not install Personal and then Enterprise

Sometimes 2 gateways are needed: Import + DQ/Live

Create a cluster for High-Availability: round robin or
balanced load

On-premises data gateway app helps you

New **Powershell 6 cmdlets** to manage Gateway



#SQLSAT895

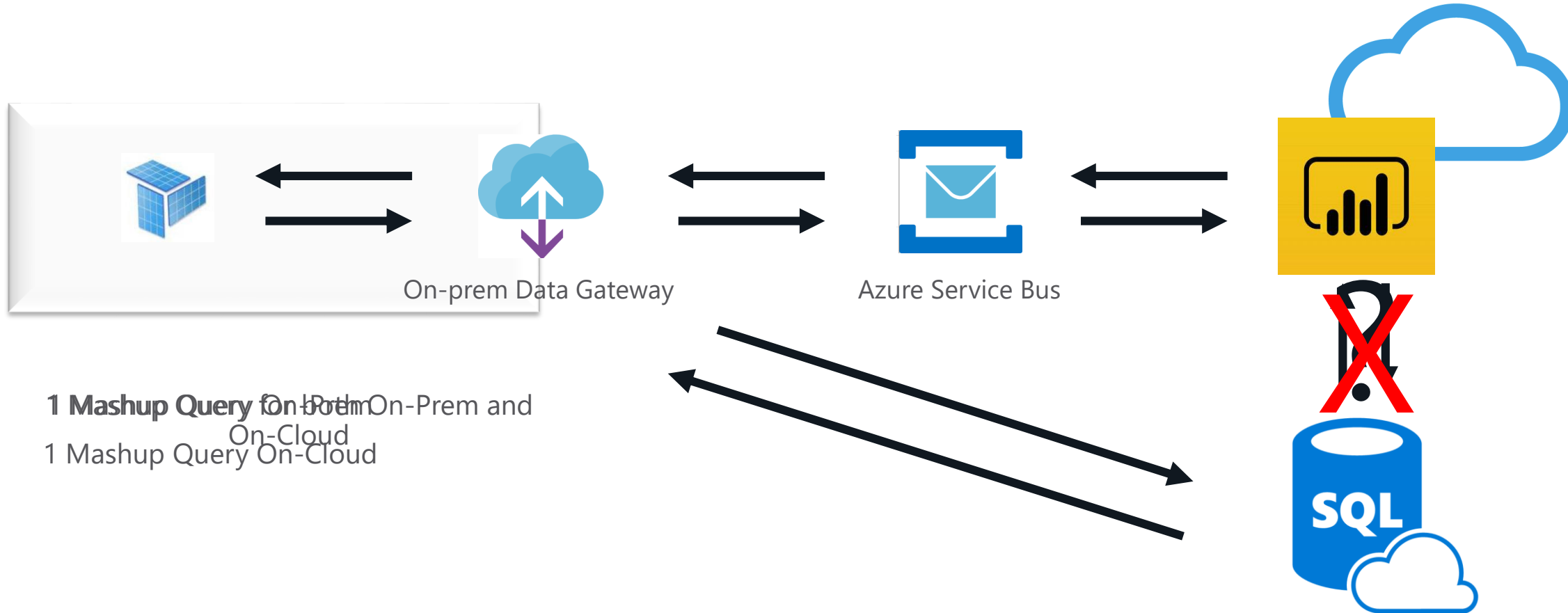




Demo Preview



On-prem & On-Cloud Mashup

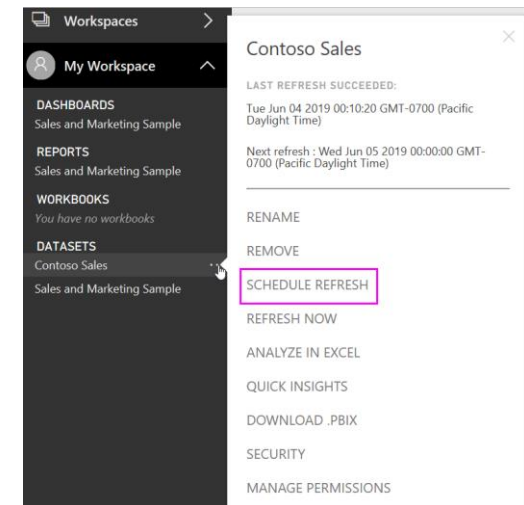
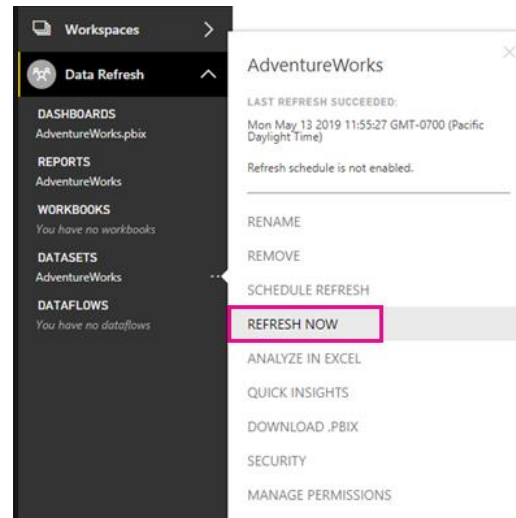


Demo: Mashup and Gateway

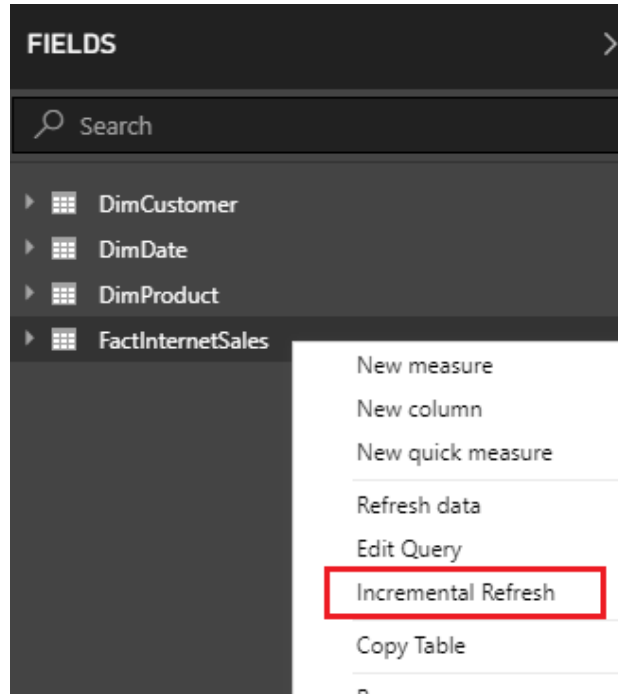
How to refresh my data?

Data are old... let's refresh!

Storage Mode	Refresh Needed	Max Refresh Per Day	Max Refresh Time
Import Mode Pro	✓	8	2 hours
Imp. Premium	✓	48	5 hours
Direct Query	✗		-
Live Connection	✗		-



Too much data?



Incremental Refresh

You can improve the speed of refresh for large tables by using incremental refresh in Premium workspaces. This setting will apply once you've published a report to the Power BI service.

① Once you've deployed this table to the Power BI service, you won't be able to download it back to Power BI Desktop. [Learn more](#)

Table Incremental Refresh
FactInternetSales ☒ On

Store rows in the last:
5 Year

Refresh rows in the last:
10 Day

☒ Detect data changes [Learn more](#)

Only refresh data in the last 10 days if the maximum value of this datetime column changes:

LastUpdateDate

☒ Only refresh complete days [Learn more](#)

Apply all

Cancel

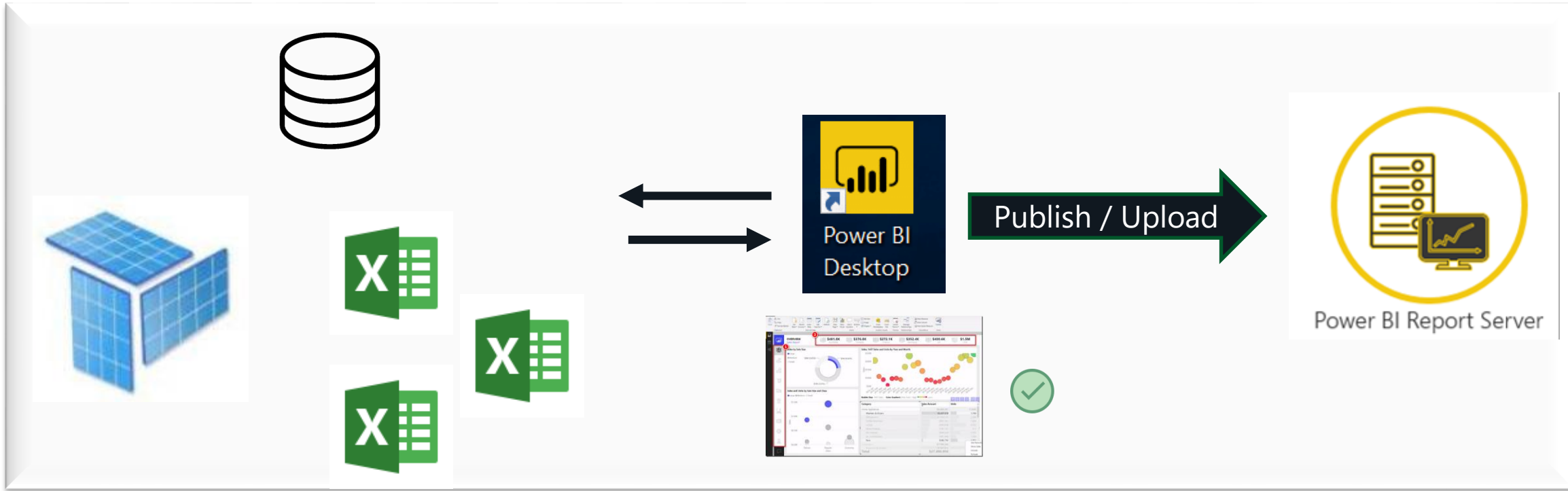
FOR PRO IS COMING



A large, teal-colored abstract graphic on the left side of the slide, consisting of several overlapping, curved, ribbon-like shapes that create a sense of depth and movement.

Is it just on-cloud solution?

Cloud is not the only option



PBI Service Vs PBI Report Server

[illegible]

<https://docs.microsoft.com/it-it/power-bi/report-server/compare-report-server-service>



#SQLSAT895





Can I «encapsulate» PowerBI?

Encapsulate Power BI: 2 options



Opt. 1: Embedded in SharePoint Online

Pros

- Centralized UX
- Easy setup

Cons

- Only reports
- Users won't use service
- Permissions management

Opt. 2: Power BI Embedded

Pros

- Full custom application
- Reports, Dashboards, Datasets and also Q&A
- End user do not need a Power BI Pro License

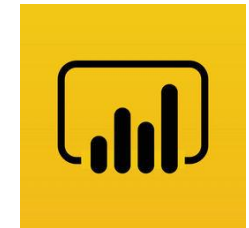
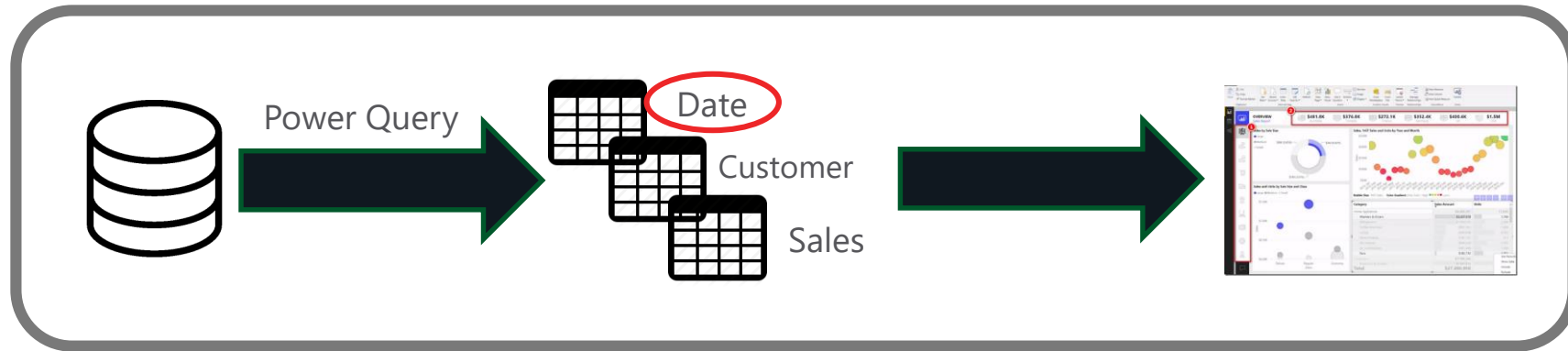
Cons

- Licensing: buckets of range renders
- Registration and management
- Not all service's features
- Web developer needed

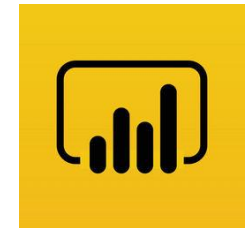
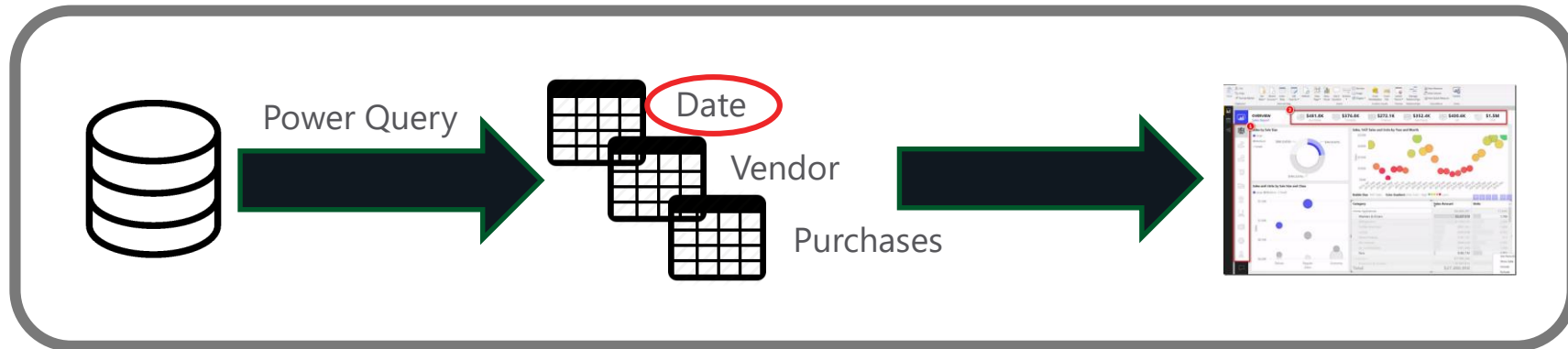
A large, stylized teal graphic on the left side of the slide, consisting of several overlapping, curved, ribbon-like shapes that form a partial 'C' or 'G' shape.

Do I need to care about
dataflows?

dataflows: as-is

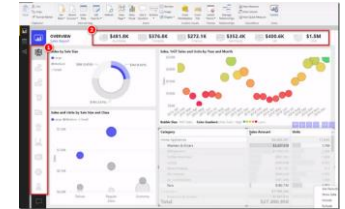
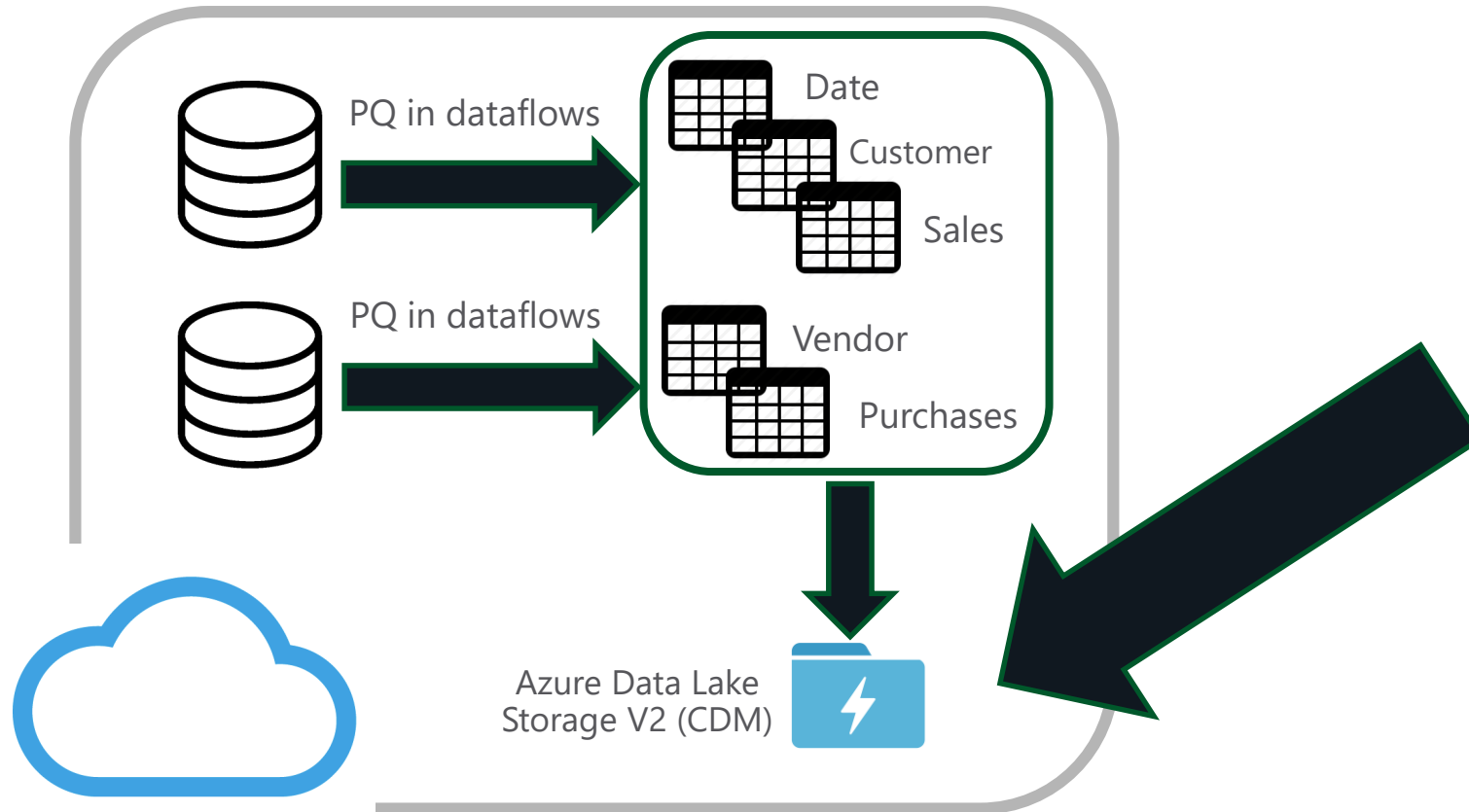


Sales.pbix

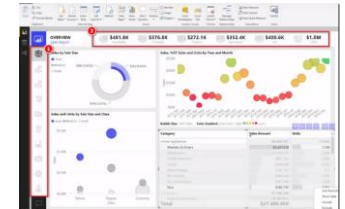


Purchases.pbix

dataflows: a new paradigm



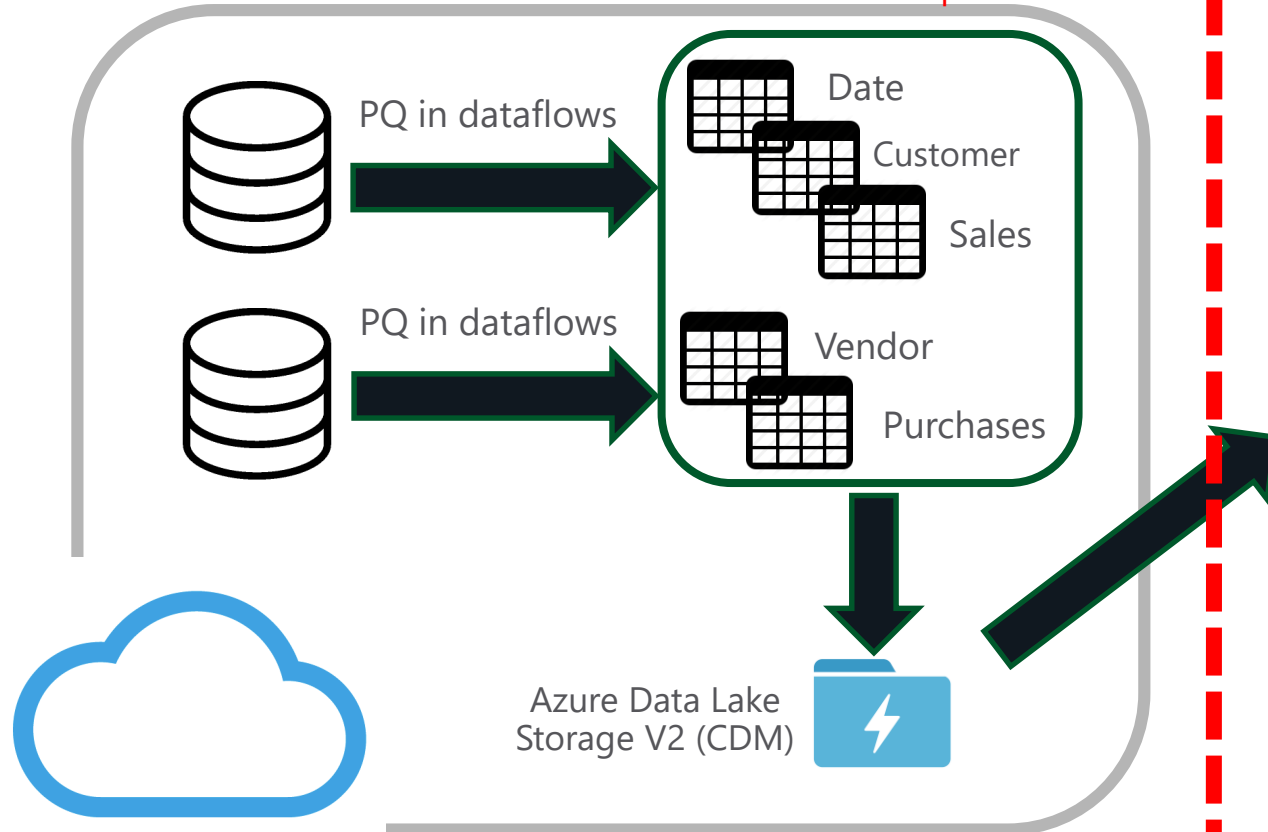
Sales.pbix



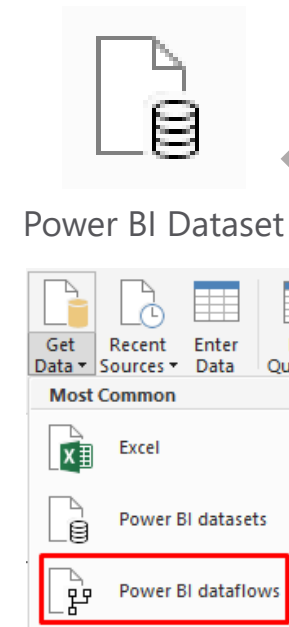
Purchases.pbix

dataflows: 2 steps forward

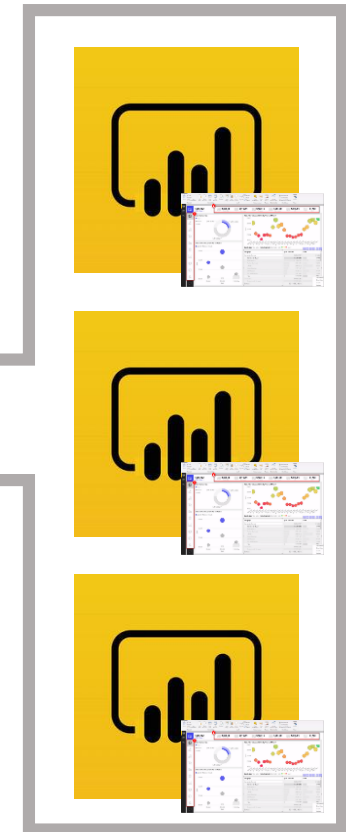
Data Preparation



Data Modeling



Data Presentation



dataflows: pros

Centralization of data preparation (reusability and standardization)

Isolating refresh (1 table a time)

Attach an existing ADLS v2 (preview)

Speed up dataset refresh

dataflows: summary pt. 1

Power BI dataflows FAQ

- ✓ A new capability for self-service data preparation in Power BI
- ✓ Delivered in a familiar Power Query experience
- ✓ Built on the foundation of Azure Data Lake Storage gen2
- ✓ Utilize the CDM folder format for data storage
- ✓ A tool for business users to drive data reuse without requiring IT involvement
- ✓ Enable Excel-like data lineage and orchestration
- ✗ NOT a replacement for datasets
- ✗ NOT a replacement for a data warehouse
- ✗ NOT a replacement for Azure Data Factory or SSIS
- ✗ NOT a Premium-only feature
- ✗ NOT an additional cost or fee
- ✗ NOT spelled with a space or any capital letters



🐦 @DAVERUIJTER | AMSTERDAM | JULY 24th



Storage Mode: Live Connection

	Power BI dataflows	Power BI datasets
Implementation	CDM folder	Analysis Services tabular model
Storage ^[2]	CSV files	Vertipaq
Metadata ^[3]	Common Data Model – model.json	BISM
Development	Power Query Online	Power Query in Power BI Desktop
Primary purpose ^[4]	Data reuse	Data analysis
Reuse ^[5]	Acts as data source in multiple datasets	Shared datasets across workspaces
Scope of reuse ^[6]	Entity level reuse	Dataset level reuse
Mashup with other data sources ^[7]	Yes	No
Used for reporting ^[8]	Not directly	Yes
Reuse outside Power BI ^[9]	Yes, through ADLSg2	Yes, through XMLA
Data access methods ^[10]	Import	Import, DirectQuery
Connection methods ^[11]	Import	Live Connection
What else am I missing?	Please let me know!	Seriously, you should let me know.

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

Demo: dataflows, datasets,
refresh... all in!

Can I have some advisable
scenarios?

Some guidelines

Relational DWH	AAS/SSAS	Model Size	Advice
Yes	No	Under 1 GB	Storage Mode: Import Data License: Pro Data Preparation: Power Query or dataflows
No	No	Under 1 GB	Storage Mode: Import Data (also shared dataset) License: Pro Data Preparation: dataflows
Yes	Yes	-	Storage Mode: Live Query Connection License: Pro
Yes/No	No	Over 1 GB	Storage Mode: Import Data License: Premium Data Preparation: Power Query or dataflows
Yes/No	No	Over 1 GB	Storage Mode: Direct Query or Dual Mode with Aggregations License: Pro Data Preparation: Power Query

Useful Links

Live or Import? <http://tiny.cc/fa9uaz>

Gateway <http://tiny.cc/lb9uaz>

PowerShell Cmdlets for data gateway: <http://tiny.cc/ebia1>

Incremental Refresh: <http://tiny.cc/4c9uaz>

dataflows & CDM: <http://tiny.cc/5d9uaz>

Dataflows best practices: <http://tiny.cc/dmgebi>

Shared Datasets replace SSAS: <http://tiny.cc/8imtbz>

Q&A



Riccardo Perico

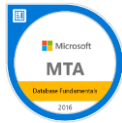
rperico@solidq.com | @R1k91  | 

Data Platform & BI Specialist @

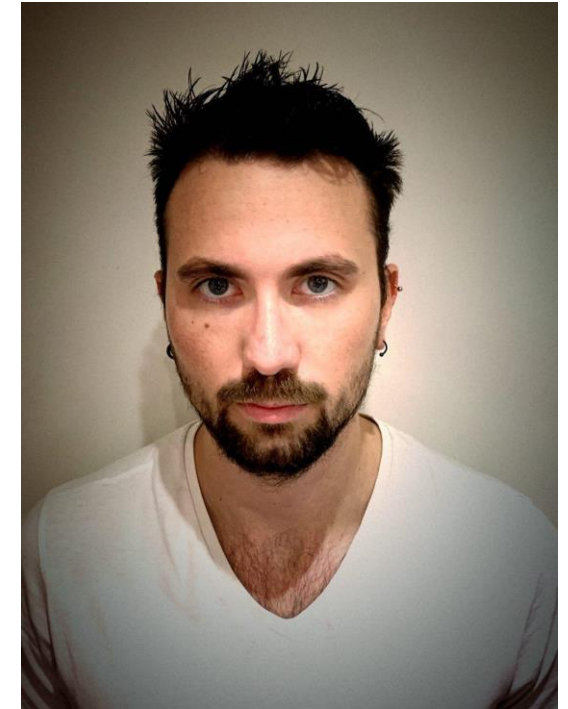


10 years in Microsoft «Data Realm»

MCP: MTA, MCSA



Speaker at Azure Saturday, SQL Saturday and DIAD tutor





Thanks!

