



Building a User-friendly semantic layer

A Practical Guide
Power BI Feszt 2024



Introduction



Mihaly Kavasi

Group Manager

Delivery Lead and Trainer

Avanade

- ✓ 8+ Years of Power BI Experience
- ✓ 7+ Year of Training Experience
- ✓ Microsoft Certified Solution Expert – Data Management and Analytics
- ✓ Microsoft Certified Trainer
- ✓ Azure Enterprise Data Analyst
- ✓ Fasttrack Recognized Solution Architect

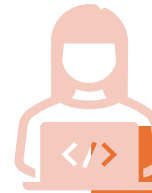


Emergence of Self-Service



Business

- Agility
- Access
- Time to insight
- Ease of use



IT\BI

- Stability
- Security
- Avoid Rework
- Ease of Integration

There are conflicting interests between Business and IT.

This results in the emergence of shadow BI from Business and strict access rules from IT.

Self-Service BI models emerged to try to ease the collaboration.

How to split the responsibilities between Business and IT/BI?



IT/BI is responsible for:

- Data quality
- Data freshness
- Security
- Correctness of insights

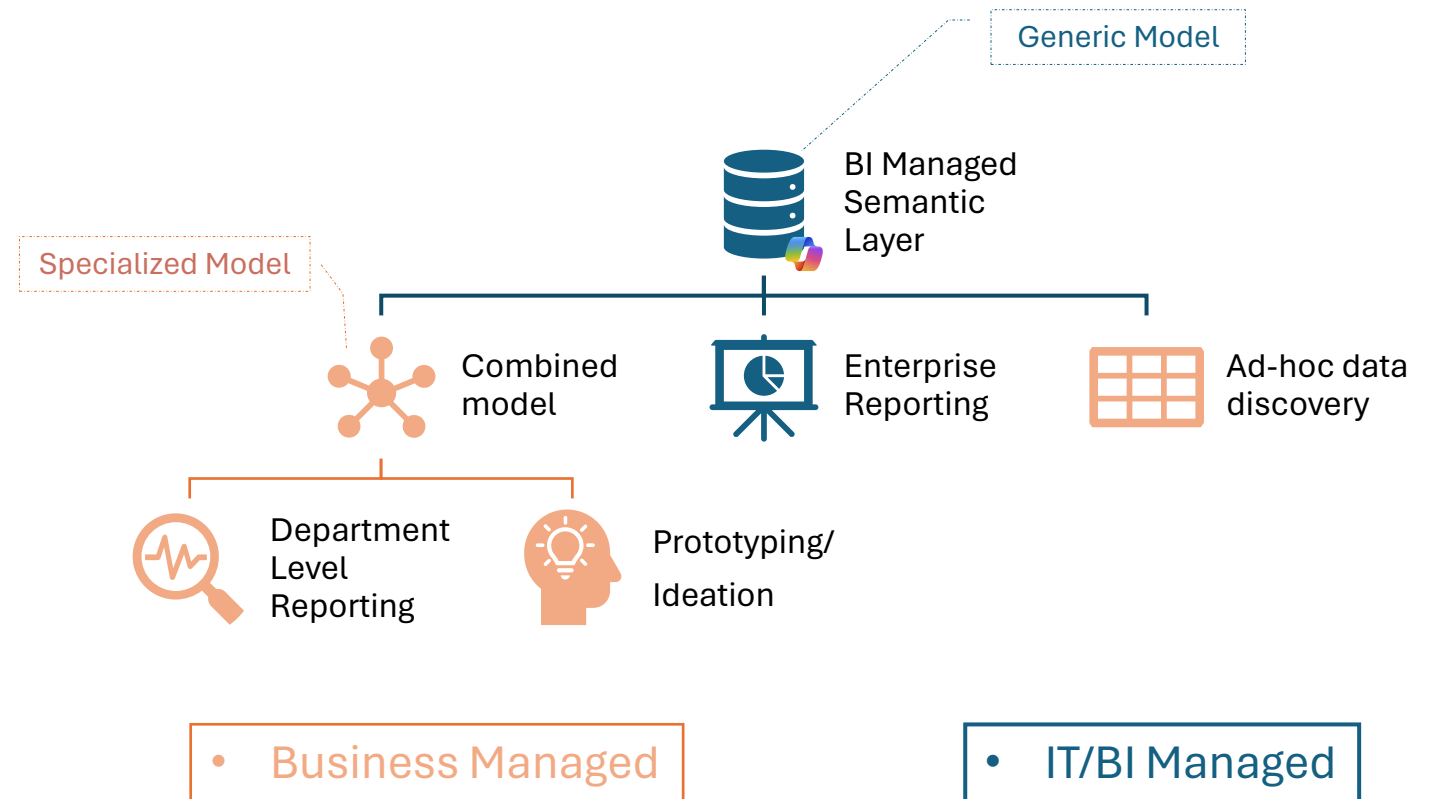



Business is responsible for:

- Validating concepts/prototypes
- Creating reports
- Design
- Adoption support

Why should we create a user-friendly semantic layer?

- Datasets in Power BI contain a most important part of any analytics solution. The data model with the calculations.
- The ability to reuse datasets is one of the keyways to eliminate unnecessary duplications of effort and data. It also helps you get closer to achieve a single source of truth.
- One good dataset can be a source of 10s or even 100s of reports. If you think about maintenance and change management than it is obvious why having a smaller number of good quality, optimized and clean datasets is better.

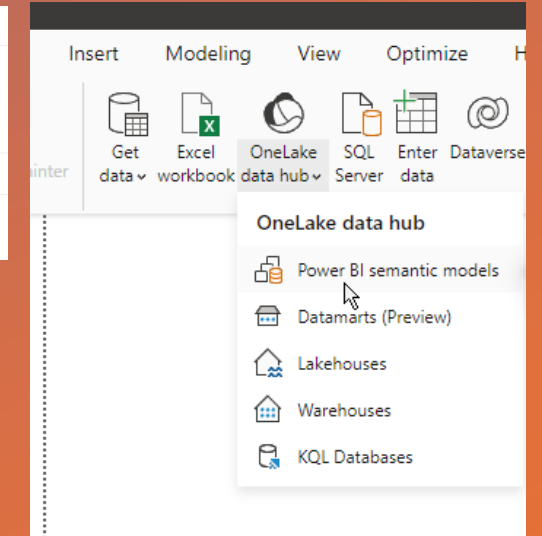
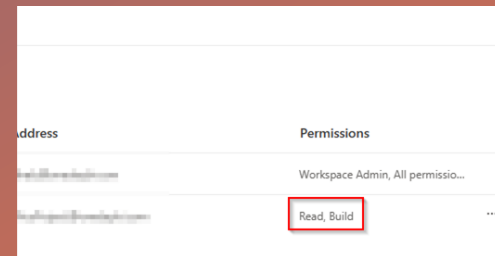
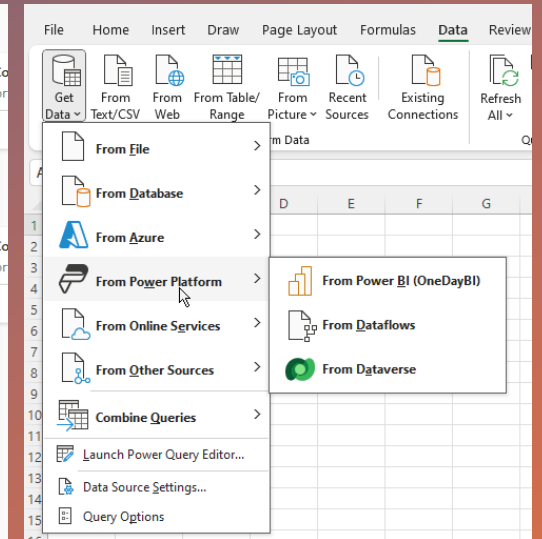
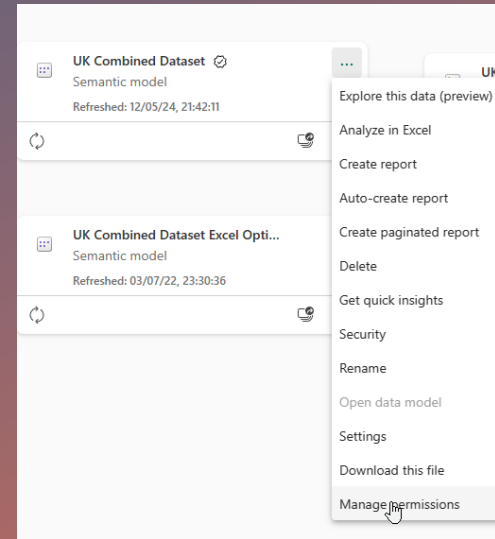




Difference between creating for yourself and others

When you want to create a semantic layer....

Share the dataset but not the workspace to provide access but maintain security



... make sure your audience knows which one is the best for their data needs

- Use the certification system
- Set Discoverability
- Add Semantic Layer

Descriptions

Endorsement and discovery

Help coworkers find your quality content by endorsing this semantic model and making it discoverable. [Learn more](#)

☐ None
This semantic model will appear in search results but isn't endorsed.

☒ Promoted
When you're ready to distribute the semantic model to your coworkers, promote it to let them know.

☐ Certified
Certify your semantic model to show coworkers that it's been reviewed and meets your org's certification criteria. [How do I get my semantic model certified?](#)
This Datamart is a core source of organizational data records.

☒ Make discoverable
Allow users without access to this semantic model to discover it and request permissions to access the data. [Learn more](#)

☐ This dataset will be made discoverable. Others in your org will be able to find it by such details as name, tables, columns, etc. [Learn more](#)

Apply Discard

OneLake data hub

Discover data from across your org and use it to create reports

[All](#) This dataset combines publicly available data from the UK. Population, housing, education, crime, company related information available. Data is refreshed monthly. Row level security is not applied.

Filter by keyword Filter

		Refreshed	Location	Endorsement	Sensitivity
UK Combined Dataset	Kávási Mihály	12/05/24, 14:38:31	UK Combined Dataset	Promoted	—
Power BI Tenant Settings	Kávási Mihály	04/08/22, 21:26:16	SelfService BI Blog	Certified	—
Power BI data sources	Kávási Mihály	04/08/22, 21:27:42	SelfService BI Blog	Certified	—
UK Combined Dataset Excel Optim...	Kávási Mihály	03/07/22, 23:30:36	UK Combined Dataset	—	—
Company Master Data	Kávási Mihály	28/06/22, 23:23:07	UK Combined Dataset	—	—
Management Demo	Kávási Mihály	29/04/18, 23:42:49	Demo	—	—
Management Demo	Kávási Mihály	17/04/18, 00:30:25	Demo	—	—

Connect Cancel

Power BI Datasets

com

Competitor_Analysis

Workspace: Demo
Owner: Kávási Mihály
Refreshed: 17/04/2018, 00:32:05

Tables in this dataset (0)

Reports using this dataset (0)

+ Insert PivotTable + Insert Table

Promoted

UK Combined Dataset

Workspace: UK Combined Dataset
Owner: Kávási Mihály
Refreshed: 28/06/2022, 23:39:25

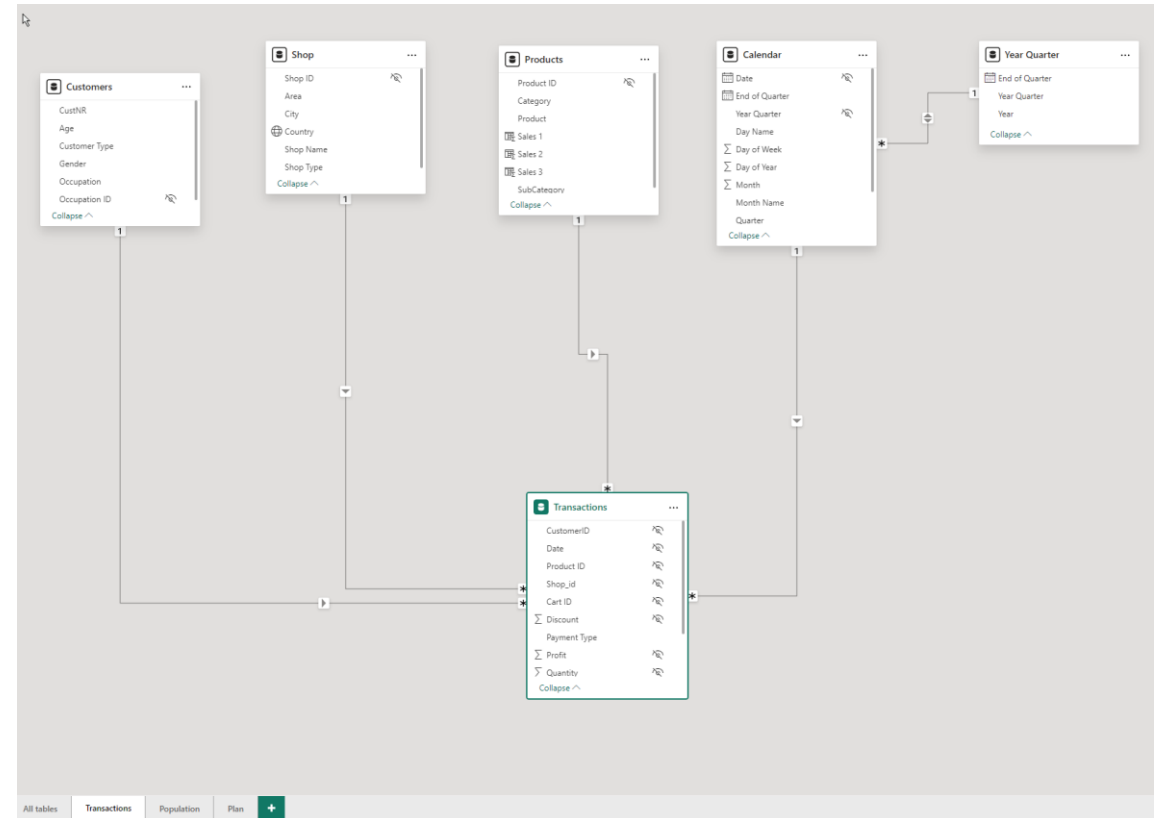
Tables in this dataset (0)

Reports using this dataset (0)

+ Insert PivotTable + Insert Table

Naming Conventions and Standard

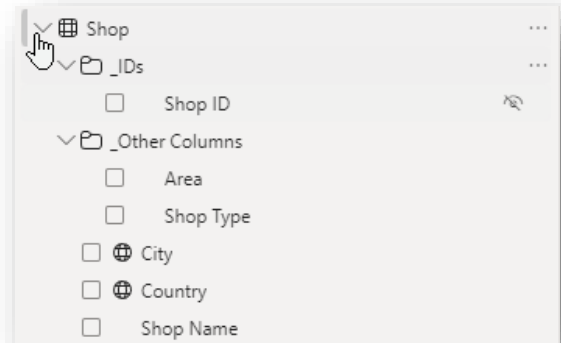
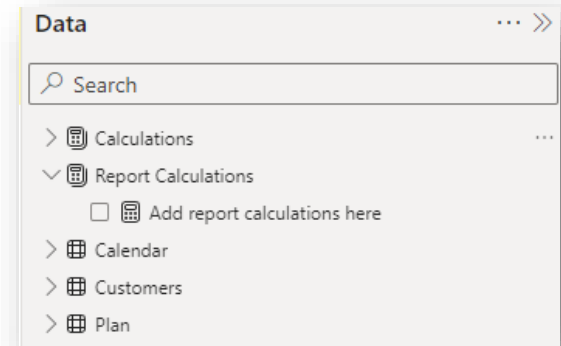
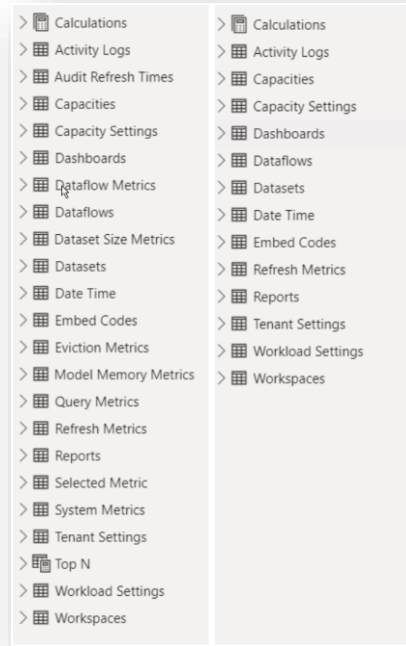
- Unambiguous and business specific naming standards for tables, columns, measures, hierarchies and relationships
- Ensure that each calculation is only can be create by one measure
- Check if measures work with all necessary dimensions
- Remove unnecessary columns
- Make sure that additional filters are not needed to display the right values using the measure. For example, the need to filter out returned sales in the visual filter to give you the proper sales volume, have a measure that already provides that.
- Create Layouts for each Fact table in the Model view



Useful content visible and organized



- Follow Analytical Data Modelling best practices (Star schema)
- Hide all the Fact tables (they should only contain facts and id columns)
- Hide all the id columns
- Do not create measures in data tables
- Create a Calculations table for model metrics
- Create a Report calculations table for report metrics
- Organize your measures into logical folders
- Surface your most important dimension attributes by create folders for the rest



Data should be clean and aligned

Make sure you do not have
referential integrity violations

+
○

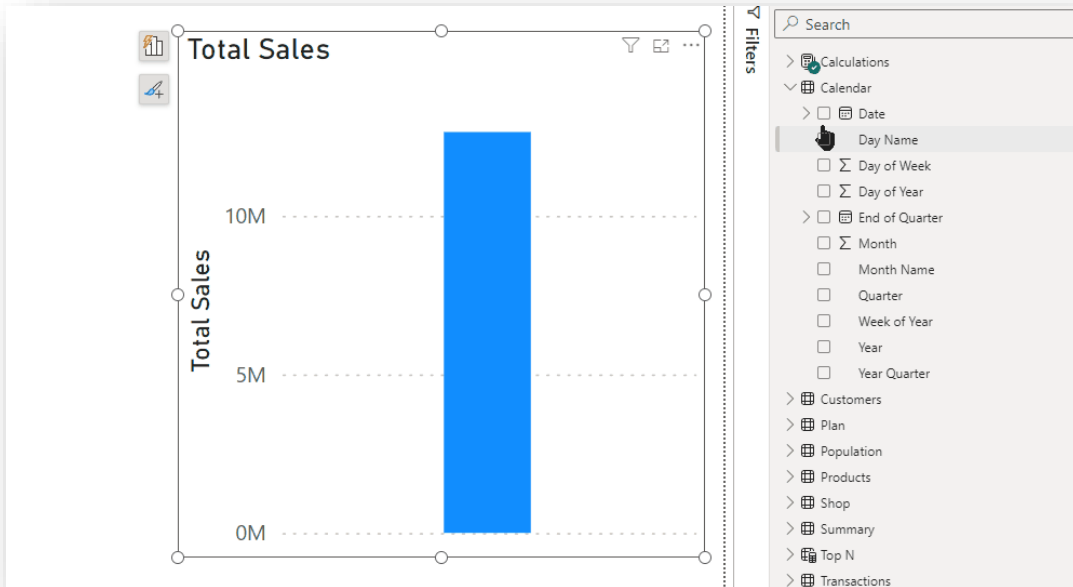
VeriQ Analyzer Metrics										
Tables	Columns	Relationships	Partitions	Summary						
Name	Cardinality	Table Size	Col Size	Data	Dictionary	Hier Size	Encoding	Data Type	RI Violations	User H
Activity	7,438	953,035	953,035	29,760	813,755	109,520	Many	-	-	1
ApprovedValues	9	18,988	18,988	16	18,748	224	Many	-	-	-
Feed Backs	133	186,450	186,450	600	182,682	3,168	Many	-	-	17
Feed Backs (Reviewing)	133	186,450	186,450	600	182,682	3,168	Many	-	-	1
Feed Backs (Reviewed)	133	186,450	186,450	600	182,682	3,168	Many	-	-	3
Feed Backs (Reviewed)	1	17,244	17,244	8	17,172	64	Many	-	-	-
Feed Backs (Reviewed)	6,430	882,975	882,975	56,624	669,871	156,480	Many	-	-	9
Feed	67,936	24,465,431	24,465,431	3,270,256	20,177,535	1,017,640	Many	-	-	-
Feed - Reviewing	67,936	10,078,544	10,078,544	3,275,800	5,513,336	1,289,408	Many	-	-	4
Feed - Reviewed (Feed Backs)	365	357,850	357,850	6,896	339,770	11,184	Many	-	-	-
Feed - Reviewed (Feed Backs)	365	409,630	409,630	6,912	391,230	11,488	Many	-	-	-
Feed - Reviewed (Feed Backs)	83,137	631,136	631,136	310,360	310,360	373,560	Many	-	-	-

Tables	Columns	Relationships	Partitions	Summary							
Table / Relationship					Size	Max From Cardinality	Max To Cardinality	1:M Ratio %	Missing Keys	Invalid Rows	Sample Violations
1. Activity					188,480	45,950	3,683,302	2.97%	18,536	544,253	
	Activity				12,872	6,430	6,430	0.97%	0	0	
	Feedback				9,464	4,726	4,761	0.74%	13	85	
	FeedbackReview				15,744	3,510	67,936	1.93%	22	62,455	
2. FeedbackReview	FeedbackReview	Feedback			7,032	3,510	6,440	0.18%	1	19	GB07PGB0700002437
	FeedbackReview	FeedbackReview	Feedback		4,624	1,729	66,046	1.88%	0	0	
	FeedbackReview	FeedbackReview	FeedbackReview		1,992	989	6,430	0.18%	1	1	GB07
	FeedbackReview	FeedbackReview	FeedbackReview		1,472	549	67,936	1.93%	1	1	(blank)
	FeedbackReview	FeedbackReview	FeedbackReview		200	117	3,048	0.09%	1	29,227	0
	FeedbackReview	FeedbackReview	FeedbackReview		200	93	7,640	0.22%	2	11	GB06, GB07
	FeedbackReview	FeedbackReview	FeedbackReview		80	71	133	0.00%	1	204	(blank)
	FeedbackReview	FeedbackReview	FeedbackReview		32	25	120	0.00%	13	3,038	GB04, GB13, GB28
	FeedbackReview	FeedbackReview	FeedbackReview		32	32	38	0.00%	0	0	
	FeedbackReview	FeedbackReview	FeedbackReview		24	16	115	0.00%	1	204	(blank)
	FeedbackReview	FeedbackReview	FeedbackReview		24	20	133	0.00%	1	29,750	(blank)
	FeedbackReview	FeedbackReview	FeedbackReview		16	13	100	0.00%	0	0	
	FeedbackReview	FeedbackReview	FeedbackReview		8	4	8	0.00%	0	0	
	FeedbackReview	FeedbackReview	FeedbackReview		8	1	4	0.00%	0	0	
	FeedbackReview				9,968	3,346	67,936	3.15%	1	1	

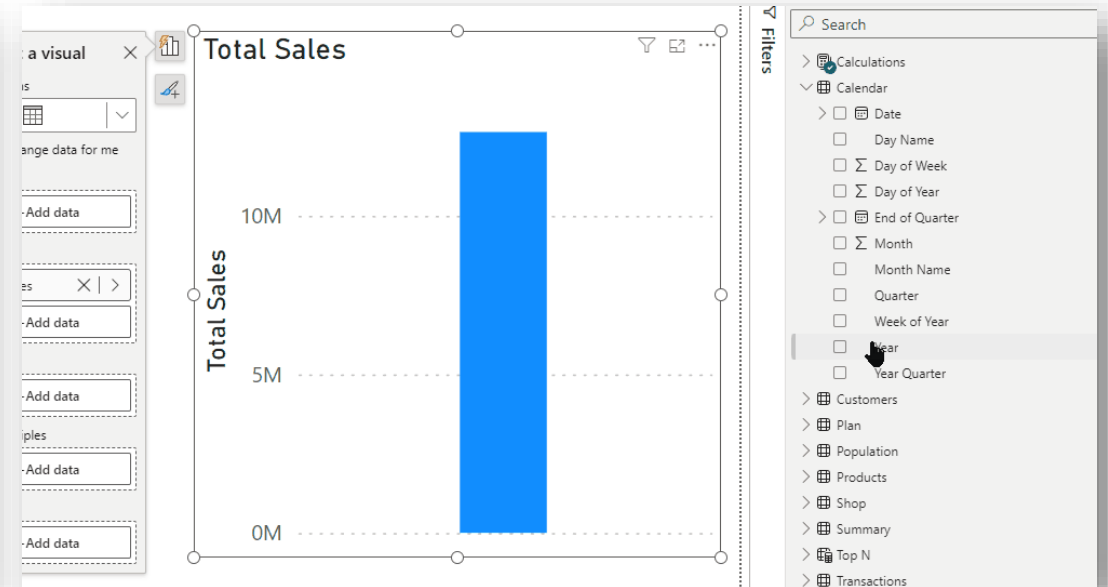
Make sure the data behaves as expected - Sorting



Sorting not set:



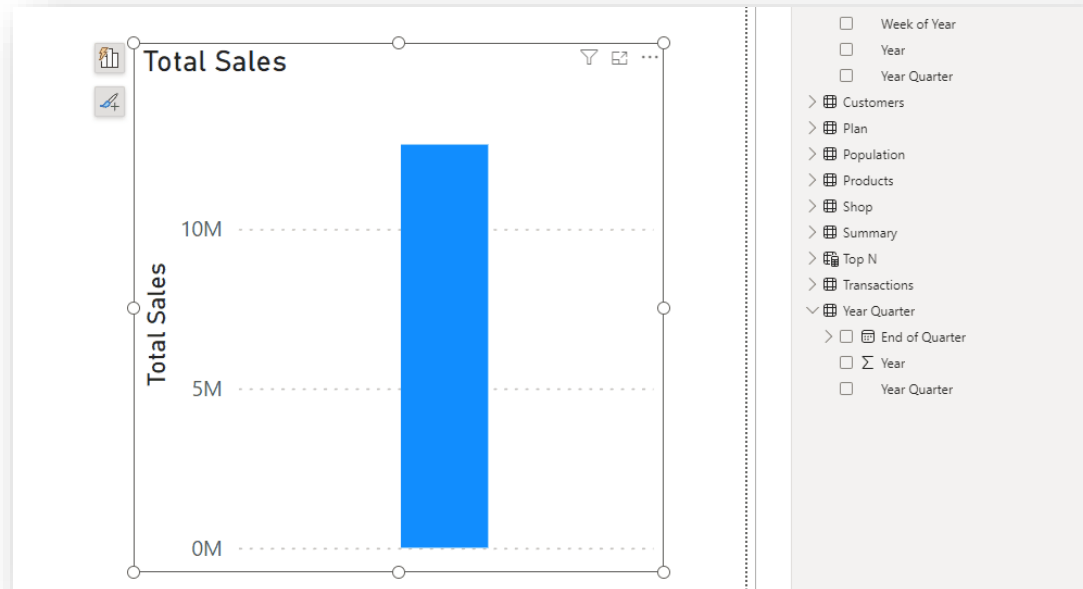
Sorting set:



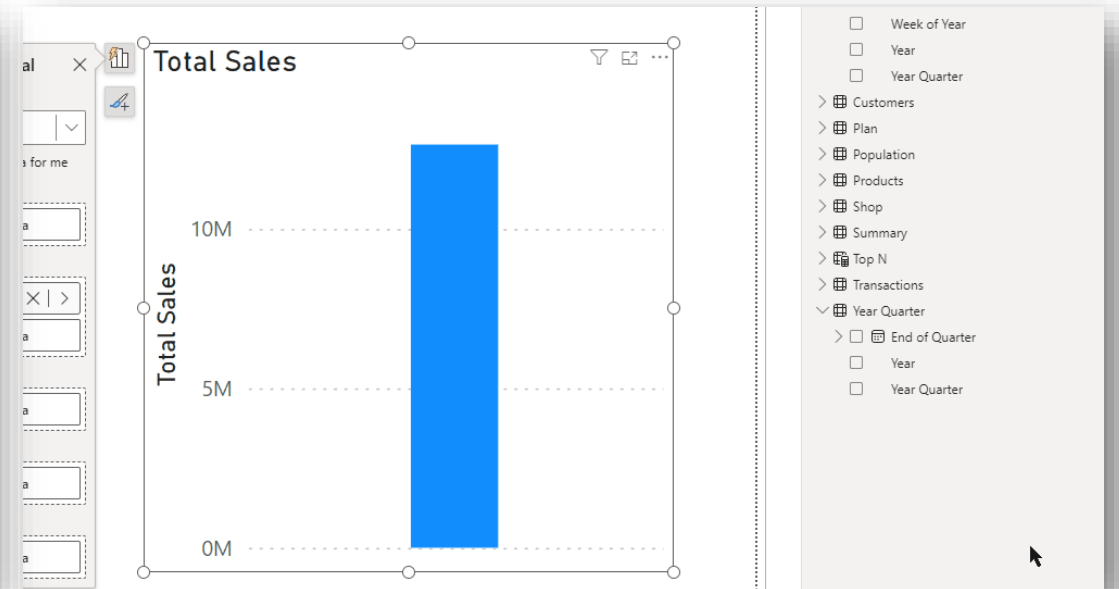
Make sure the data behaves as expected - Summarization



Summarization not set correctly:



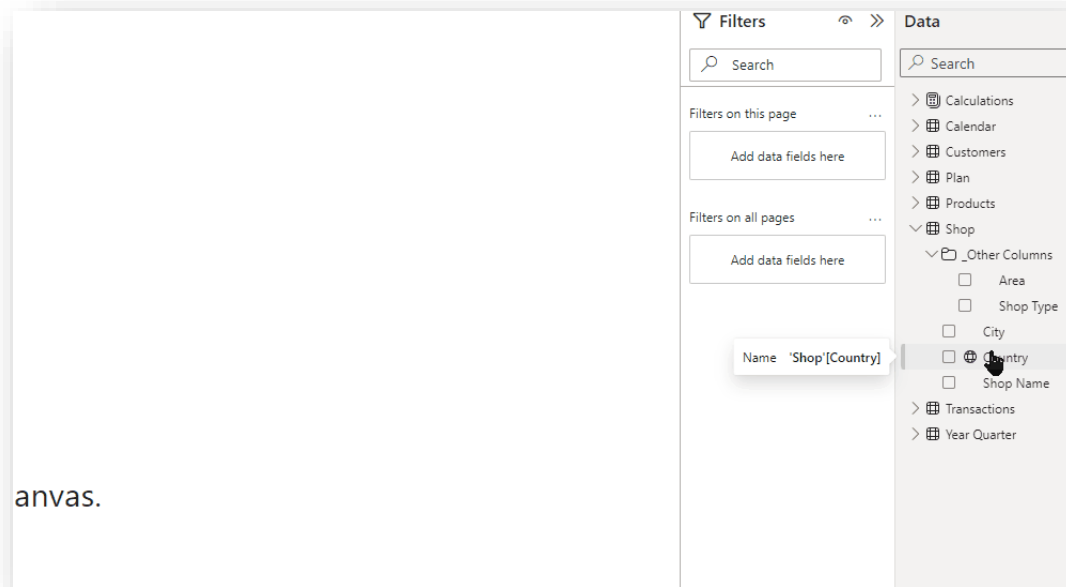
Summarization set correctly:



Make sure the data behaves as expected – Data Category

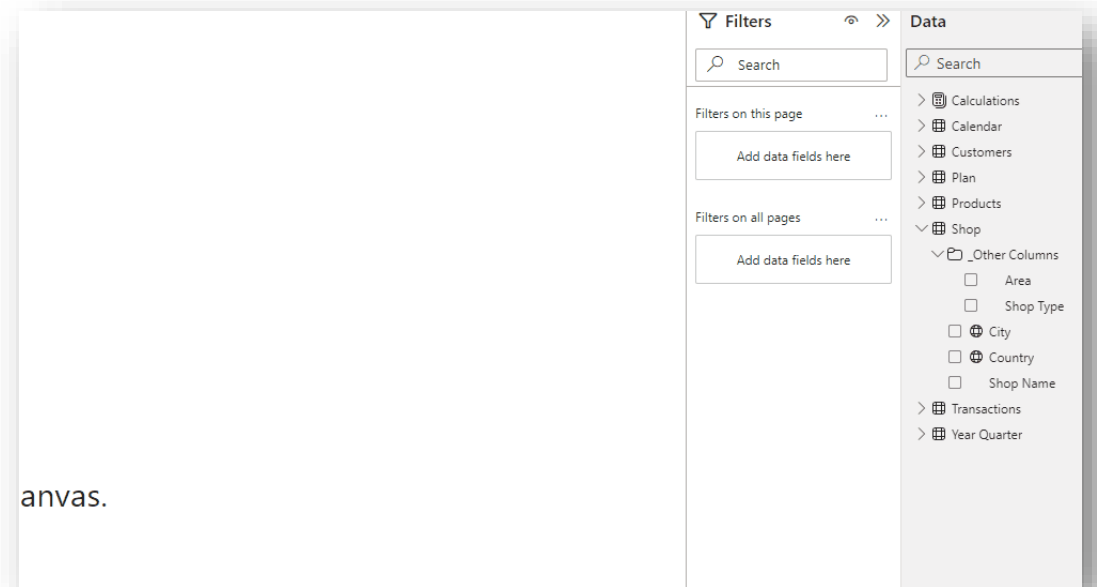


Data category not set correctly:



anvas.

Data Category set correctly:



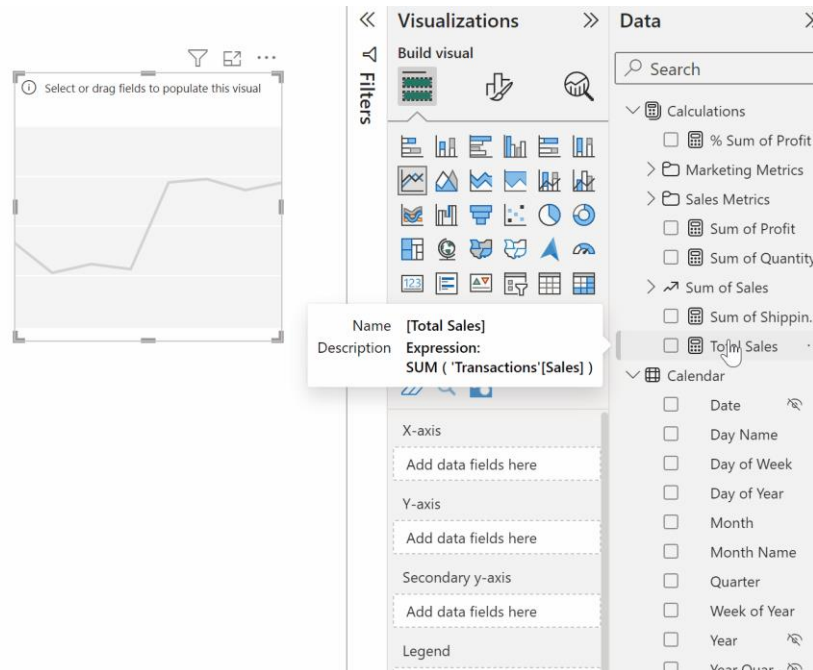
anvas.

Make sure the data behaves as expected – Auto Date Time

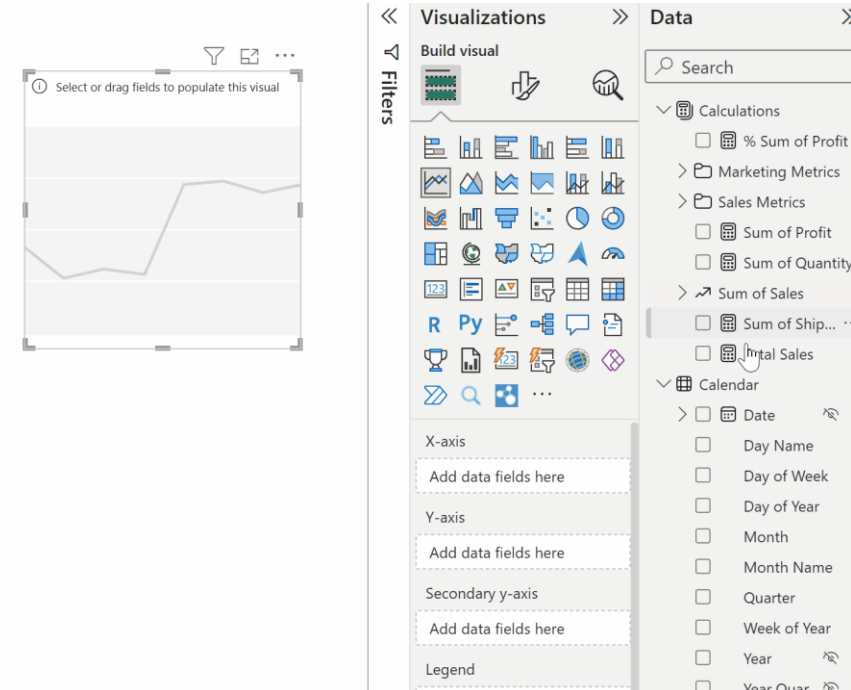
Time intelligence

☒ Auto date/time ⓘ [Learn more](#)

Auto date time disabled:



Auto date time enabled:

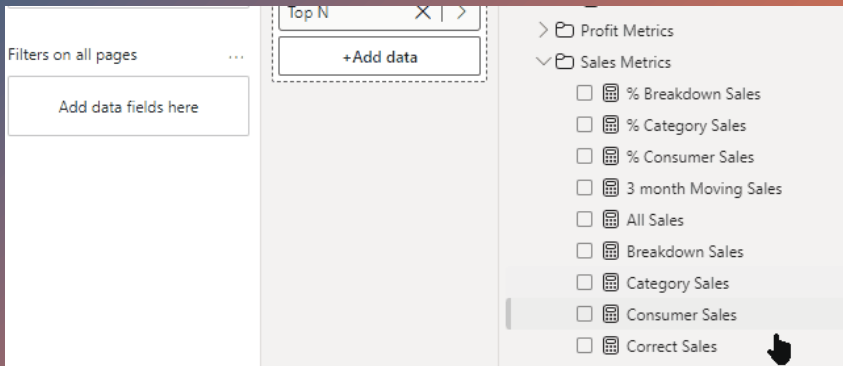


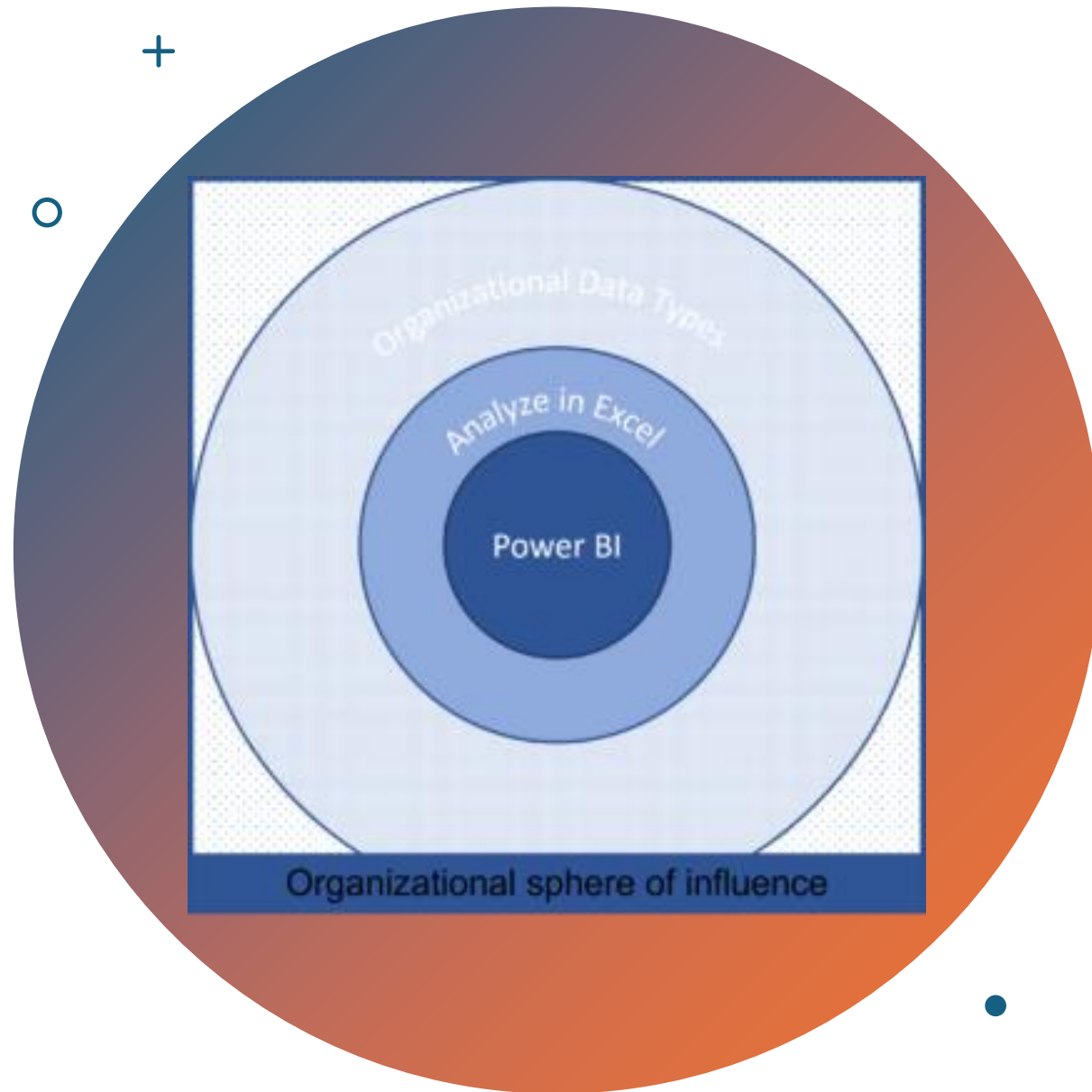
Warning this is controversial suggestion!

Make your measures understandable

- Give meaningful names
- Only keep visible the reusable measures (hide any intermediary or report specific measures)
- Format your measures
- Add the measure definition to the description
- Semantic Layer measures should not contain design related logic (like +0)

```
1 3 month Moving Sales =  
2 IF([Total Sales],  
3   CALCULATE (  
4     [Total Sales],  
5     DATESBETWEEN (  
6       'Calendar'[Date],  
7       CALCULATE ( MAX ( 'Calendar'[Date] ), DATEADD ( 'Calendar'[Date], -3, MONTH ) ) + 1,  
8       MAX ( 'Calendar'[Date] )  
9     )  
10  )  
11 )
```





Think about all your potential users

- Do not only think about Power BI report creators
- Analyze in Excel in many cases are more likely used for quick data discovery
- Also, you can improve your reach even further by setting up Organizational Data Types

Accommodate Excel Features – Show Details



Detailed Rows Expression not set:

Row Labels	Total Sales
Computer Parts	
FIC	64269.573
FOXCONN	53927.517
FUJITSU	1467385.07
HITACHI	57187.095
HP	643651.7761
HTC	34254.855
HYUNDAI	47753.706
IBM	40395.687
MESH COMPUTERS	20340.903
MICROSOFT	36548.412
RAZER	29772.585
SAMSUNG	1325572.643
SAPPHIRE TECHNOLOGY	32481.3
SGI	39990.756
SHUTTLE	40501.623
SIRAGON	25545.933
SUPERMICRO	50914.848
TOSHIBA	658595.5255

Detailed Rows expression set correctly:

Row Labels	Total Sales
Computer Parts	
FIC	64269.573
FOXCONN	53927.517
FUJITSU	1467385.07
HITACHI	57187.095
HP	643651.7761
HTC	34254.855
HYUNDAI	47753.706
IBM	40395.687
MESH COMPUTERS	20340.903
MICROSOFT	36548.412
RAZER	29772.585
SAMSUNG	1325572.643
SAPPHIRE TECHNOLOGY	32481.3
SGI	39990.756
SHUTTLE	40501.623
SIRAGON	25545.933
SUPERMICRO	50914.848
TOSHIBA	658595.5255

Accommodate Excel Features – Data Types



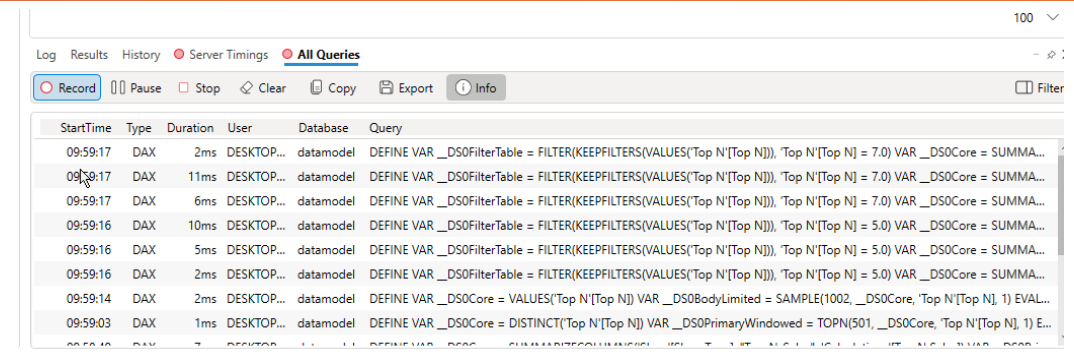
The screenshot shows the Microsoft Excel interface with the 'Data' ribbon selected. The ribbon includes sections for 'Get & Transform Data', 'Queries & Connections', 'Data Types', 'Sort & Filter', 'Data Tools', and 'Forecast'. The 'Data Types' section is expanded, showing 'Schools' and 'GP Practices' as available data types. The worksheet contains a table with the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
4	School Name	Ofsted Rating	Gender mix											
5	Woodhill Preparatory School	No Information	Mixed											
6	Wellstead Primary School	Good	Mixed											
7	Berrywood Primary School	Good	Mixed											
8	Kings Copse Primary School	Good	Mixed											
9	Freegrounds Junior School	Good	Mixed											
10	Shamblehurst Primary School	Good	Mixed											
11	Botley Church of England Controlled Primary School	Outstanding	Mixed											
12	Saint James' Church of England Primary School	Good	Mixed											
13	Freegrounds Infant School	Good	Mixed											
14	Wildern School	Outstanding	Mixed											
15														
16														
17														
18														
19														
20														
21														
22														
23														

Continuously Optimize Performance

During Development:

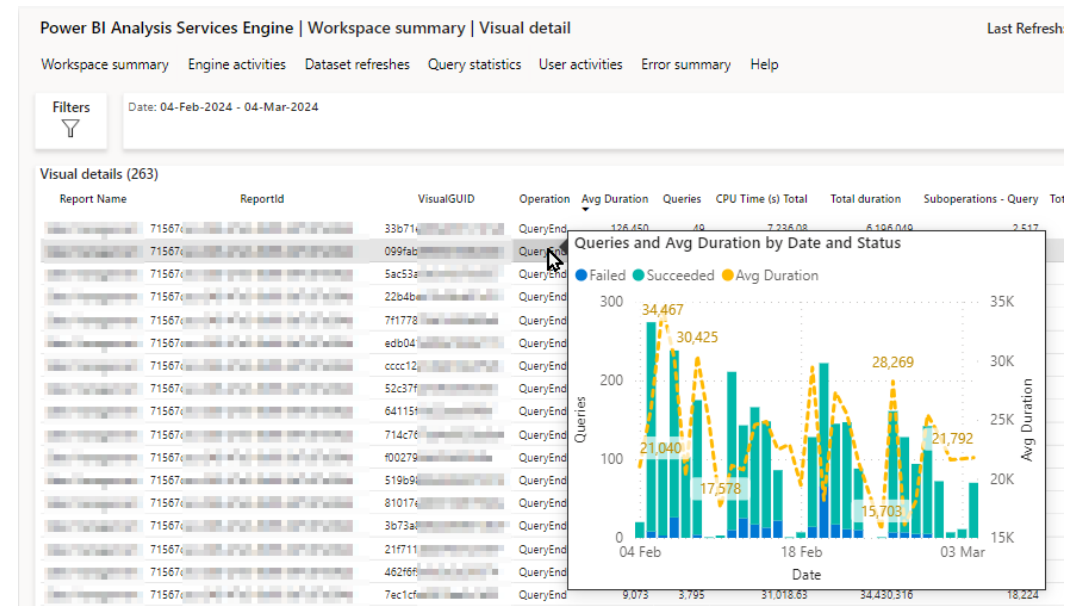
- Use tools like Performance Analyzer and DAX Studio to test and optimize the performance.

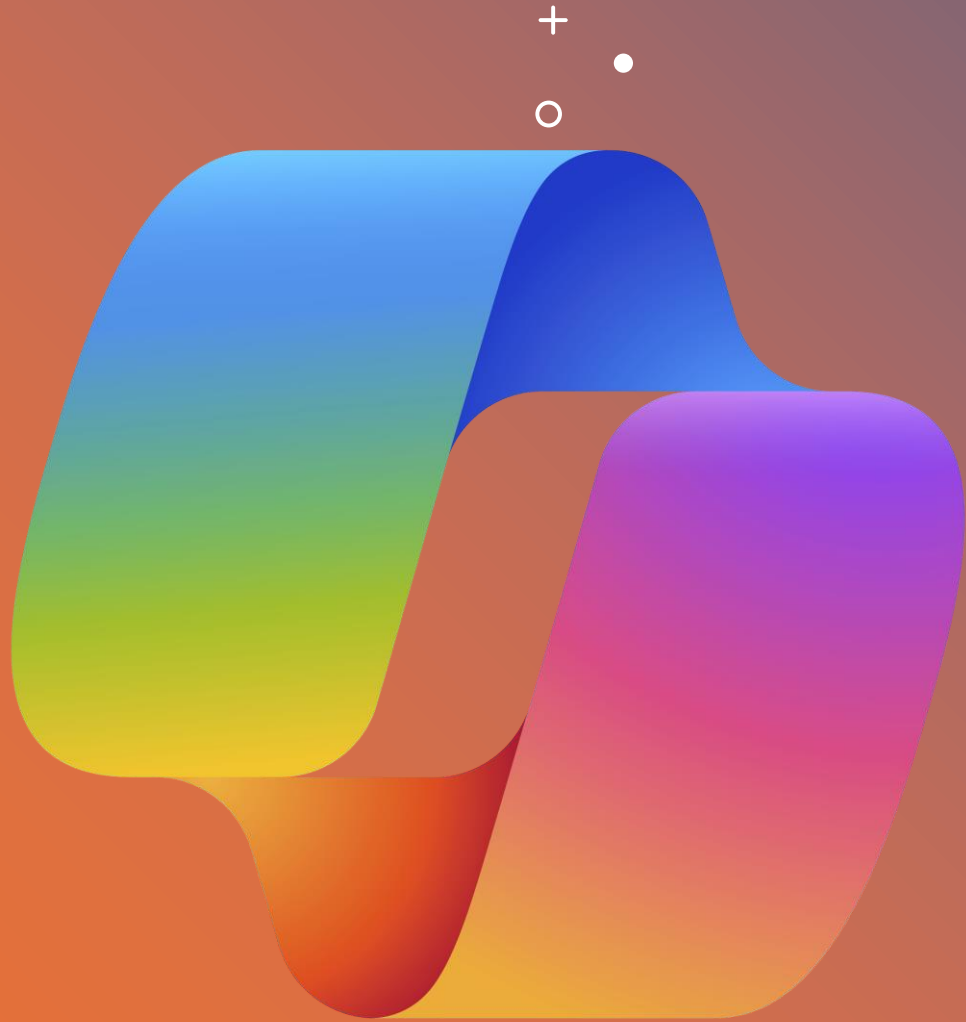


StartTime	Type	Duration	User	Database	Query
09:59:17	DAX	2ms	DESKTOP...	datamodel	DEFINE VAR __DSOFilterTable = FILTER(KEEPFILTERS(VALUE...
09:59:17	DAX	11ms	DESKTOP...	datamodel	DEFINE VAR __DSOFilterTable = FILTER(KEEPFILTERS(VALUE...
09:59:17	DAX	6ms	DESKTOP...	datamodel	DEFINE VAR __DSOFilterTable = FILTER(KEEPFILTERS(VALUE...
09:59:16	DAX	10ms	DESKTOP...	datamodel	DEFINE VAR __DSOFilterTable = FILTER(KEEPFILTERS(VALUE...
09:59:16	DAX	5ms	DESKTOP...	datamodel	DEFINE VAR __DSOFilterTable = FILTER(KEEPFILTERS(VALUE...
09:59:16	DAX	2ms	DESKTOP...	datamodel	DEFINE VAR __DSOFilterTable = FILTER(KEEPFILTERS(VALUE...
09:59:14	DAX	2ms	DESKTOP...	datamodel	DEFINE VAR __DSOCore = VALUES('Top N'[Top N]) VAR __DSO...
09:59:03	DAX	1ms	DESKTOP...	datamodel	DEFINE VAR __DSOCore = DISTINCT('Top N'[Top N]) VAR __DSO...

During Production:

- Use Log Analytics to continuously monitor performance and proactively optimize based on utilization patterns





These same
principles will
make our
semantic layer
Copilot Ready!