#### Power BI "X Files"

# 5 Things They "Forgot" to Tell You...

Nikola Ilic

Data Mozart, Microsoft Data Platform MVP



#### Thank you!



















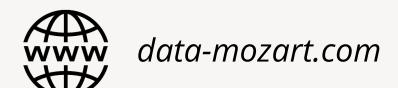
Many thanks to our sponsors, without whom such an event would not be possible.



#### Nikola Ilic

#### Data Platform Consultant & Trainer

- I'm making music from the data!
- Power BI and "all the data things" addict, blogger, speaker...
- Father of 2, Barca & Leo Messi fan...







# WHAT'S IN FOR YOU?

- ✓5 things you won't find in docs
- ✓5 things you won't find in courses





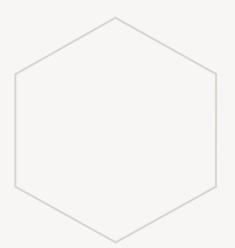
# **DISCLAIMER!**



This is **NOT** a beginner-friendly session!

This is **NOT** a session for data visualization wizards!





# **Agenda**



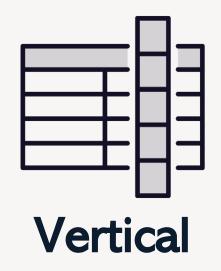




### **DAX Fusion**

#### What is DAX Fusion?

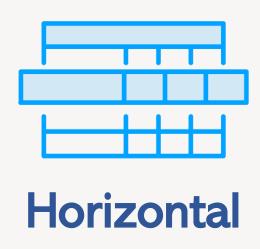
- **✓** DAX optimization technique
- √ Reduces the number of Storage Engine queries







#### **Horizontal Fusion**

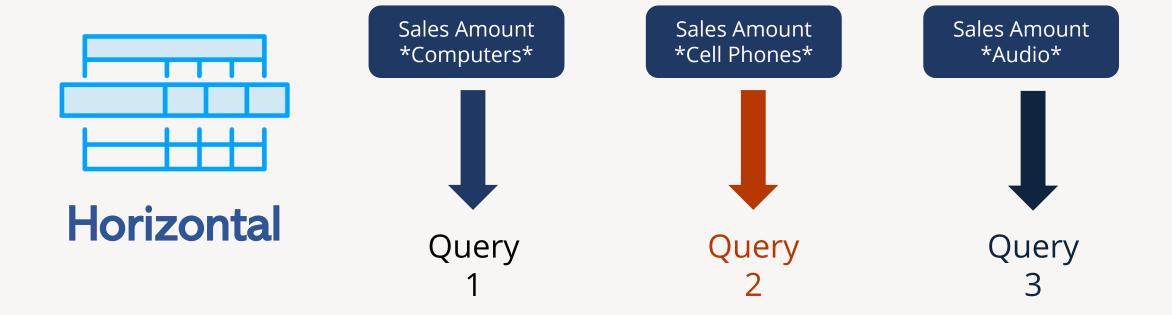


- ✓ Same calculation different filter(s)
- √ Usually much better performance

Available out-of-the-box since September 2022

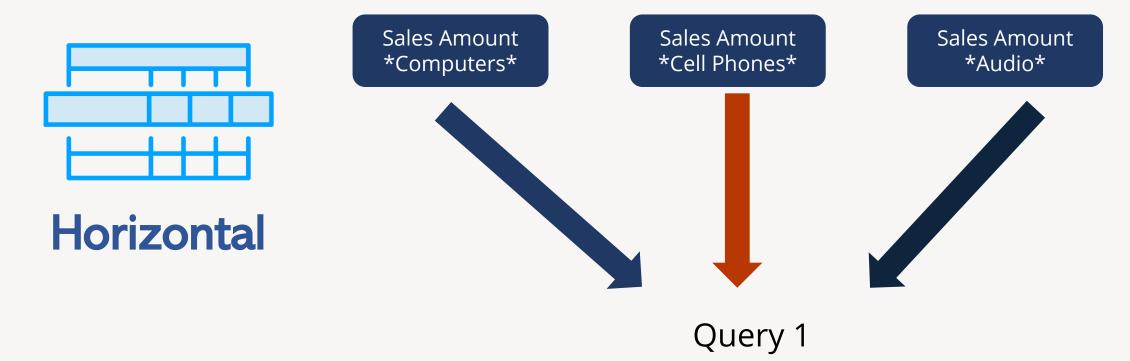


# Horizontal Fusion – Before September 2022





# Horizontal Fusion – After September 2022







Sales Amount Computers	Sales Amount Cell Phones	Sales Amount Audio
250 M	150 M	100 M



### **Horizontal Fusion**

**DEMO** 

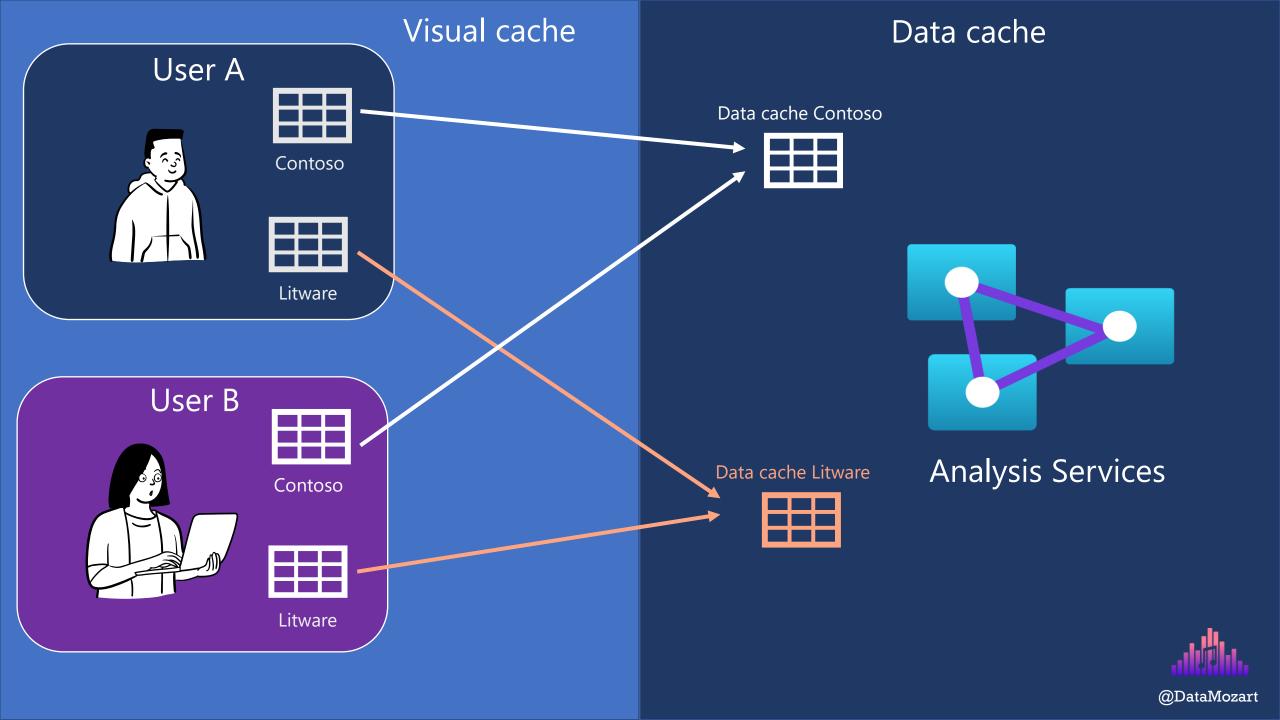


# Cache me if you can



# Cache me if you can

**DEMO** 





# Two cache types in Power BI/Tabular

#### Visual cache

Data cached in the session scope

#### Data cache

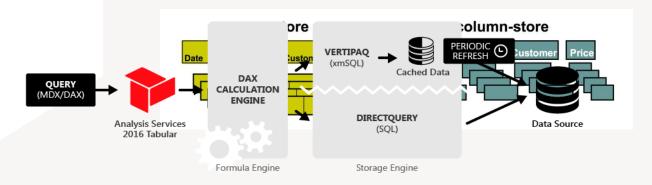
Data cached in the scope of the AS instance



# Dark art of table processing

### What is a VertiPaq?

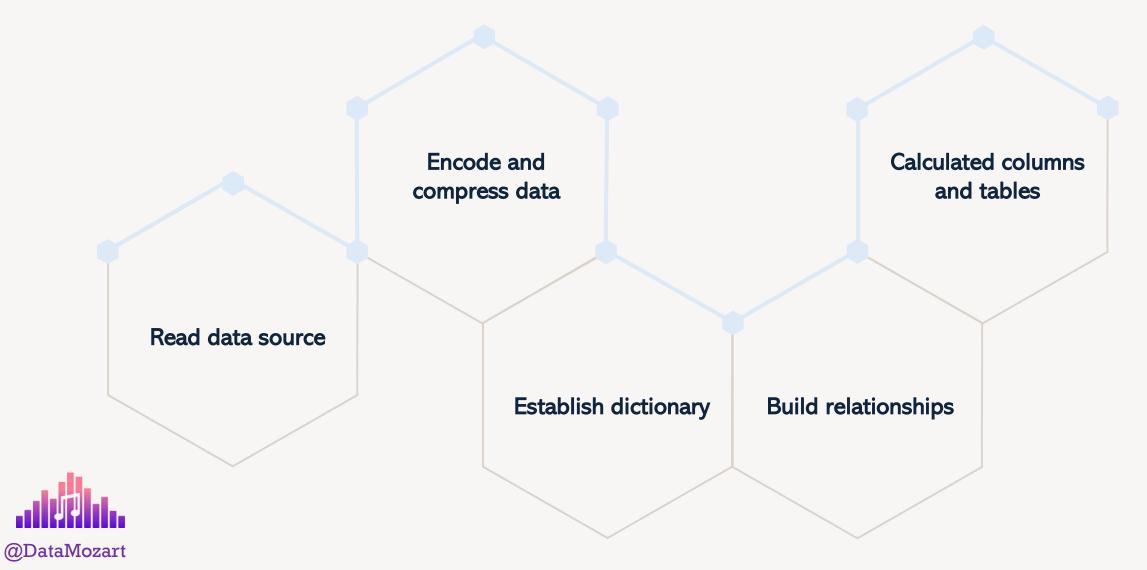
- ✓ Row-store vs columnar databases
- ✓ Formula engine vs storage engine



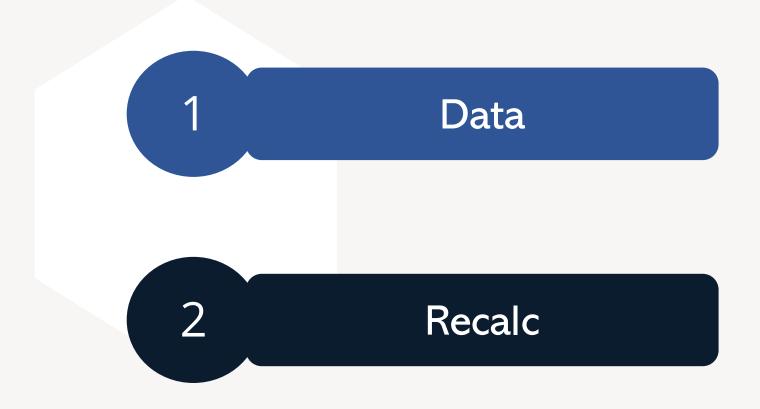
Source: The Definitive Guide to DAX, 2nd Edition (M.Russo & A.Ferrari)



# Import Mode In-Depth

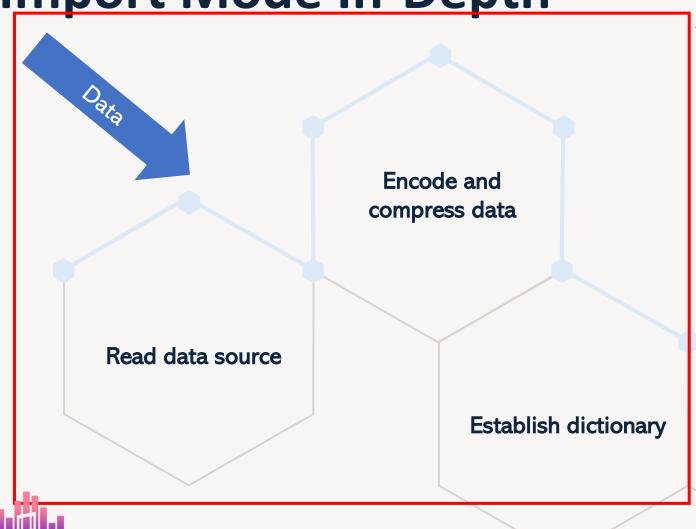


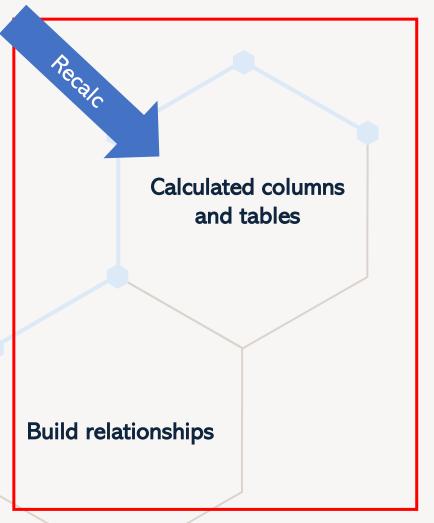
## **Two Phases of Data Processing**





Import Mode In-Depth







# **Tabular Processing Options**

Mode	Scope	Description
Process Data	Table/Partition	Load data into a table/partition <u>without rebuilding</u> hierarchies or relationships or recalculating calculated columns/measures
Process Recalc	Database	Updates and recalculates hierarchies, relationships, and calculated columns
Process Full	Database/Table/Partition	SSAS drops all data in the object and then processes the object. Requires the most resources
Process Default	Database/Table/Partition	Detects the state of an object and performs processing necessary to deliver unprocessed or partially processed objects to a fully processed state. Hierarchies, calculated columns, and relationships are recalculated
Process Add	Partition	Incrementally update partition with new data



## **Choosing the Right Strategy...**

Process Data on each table/partition to create updated dictionaries for each column

2

Process Recalc on the database

Calculated columns/Relationships/Hierarchies unavailable between the start of 1 and the end of 2!!!



## Choosing the Right Strategy...

Phil Seamark's blog post on visualizing the Power BI refresh process







Michael Kovalsky's blog on automating the processing







# Fold on tight

## What is Query Folding?

- √ Translate M to a query language
- √ Pushing heavy lifting on the data source

Ability to generate <u>a single SQL query</u> that's going to be executed on the <u>data source side!</u>



### **Query Folding and Data Sources**

- √ Relational database engines
- ✓ Odata, SSAS, Sharepoint lists, Exchange, AD...
- ✓ Some sources don't support folding at all!



@DataMozart





## **Query Folding and Data Transformations**

- ✓ Not all transformations will fold!
- √ Subtle differences are sometimes decisive



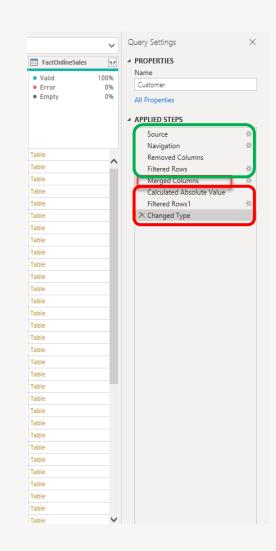
- ✓ Removing/Renaming columns
- √ Filtering rows
- √ Grouping/Summarizing
- ✓ Merging/Appending queries from the same source
- ✓ Simple Custom columns
- ✓ Pivot/Unpivot

- ✓ Merging queries from different sources
- ✓ Appending queries from different sources
- √ Complex Custom columns
- ✓ Index columns
- ✓ Changing column data type ("it depends")!

## Query Folding – why should I care?

- ✓ Data refresh more efficient
- √ With DirectQuery/Dual the query must fold!
- ✓Incremental refresh
- VALL OF NOTHING?
- **≻Slow report**
- >Slow visuals
- ➤ Large data model

  IT'S NOT BECAUSE OF QUERY (NON)FOLDING!!!

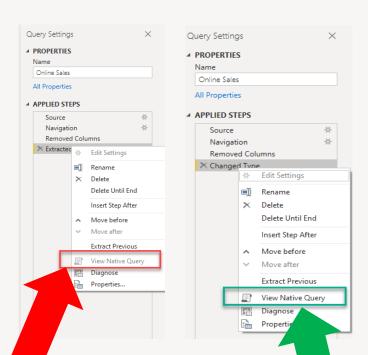


## How do you know if the query folds?

- √ View Native Query
- √lt's not 100% correct!!!
- ✓ Use Query Diagnostics in PQ Editor or SQL Server Profiler
- √ Feature in Power Query Online!







PROBABLY!!!



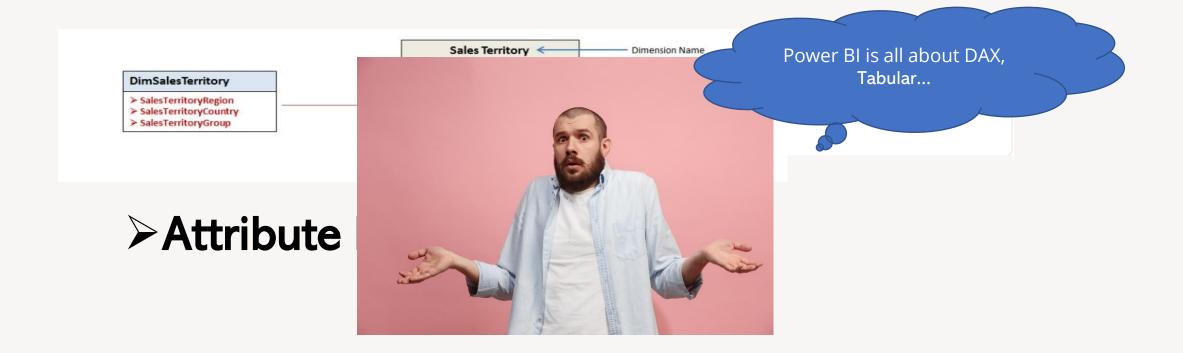
# Fold on tight

**DEMO** 



# Secret sauce of professionals

#### Scenario



@DataMozart

Which language does *Excel* use when connecting to *Tabular*?

#### **ISAVAILABLEINMDX**

- >Set to TRUE by default
- **>Why not FALSE?!**



MANY POWER BI DEVELOPERS ARE NOT AWARE OF THE PROPERTY!



# Secret sauce of professionals

**DEMO** 



#### **Lesson Learned...**

- ➢ Hidden gems in Tabular
- ➢Increase efficiency + keep functionalities

#### WIN-WIN SITUATION!



## Some things to keep in mind...



>DISTINCTCOUNT, MIN, MAX...





## **Summary**

**DAX Fusion** 

Horizontal and vertical

Cache me if you can

Visual vs data cache

Dark art of table processing

Process Data vs Process Recalc Fold on tight

Don't break the fold!

Secret sauce of professionals

Hidden little gem that can make you a Power BI hero





## Feedback please!





https://feedback.dsmuc.de/