



Power BI on a Diet - 5 TIPS to Slim Down Your Semantic Models

Dhyan Rathore



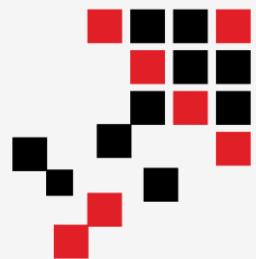


Special thanks to Fabric and Power BI Team at



Microsoft

This Summit presented to you by



RADACAD





Dhyan Rathore

Fabric Enabling Lead @ **Autoliv**

- ✓ **Stockholm, Sweden**
- ✓ **Blogger, Speaker & Hiker**
- ✓ **Automation & Optimization**



Why Data Size Matters?



1 GB

VS.

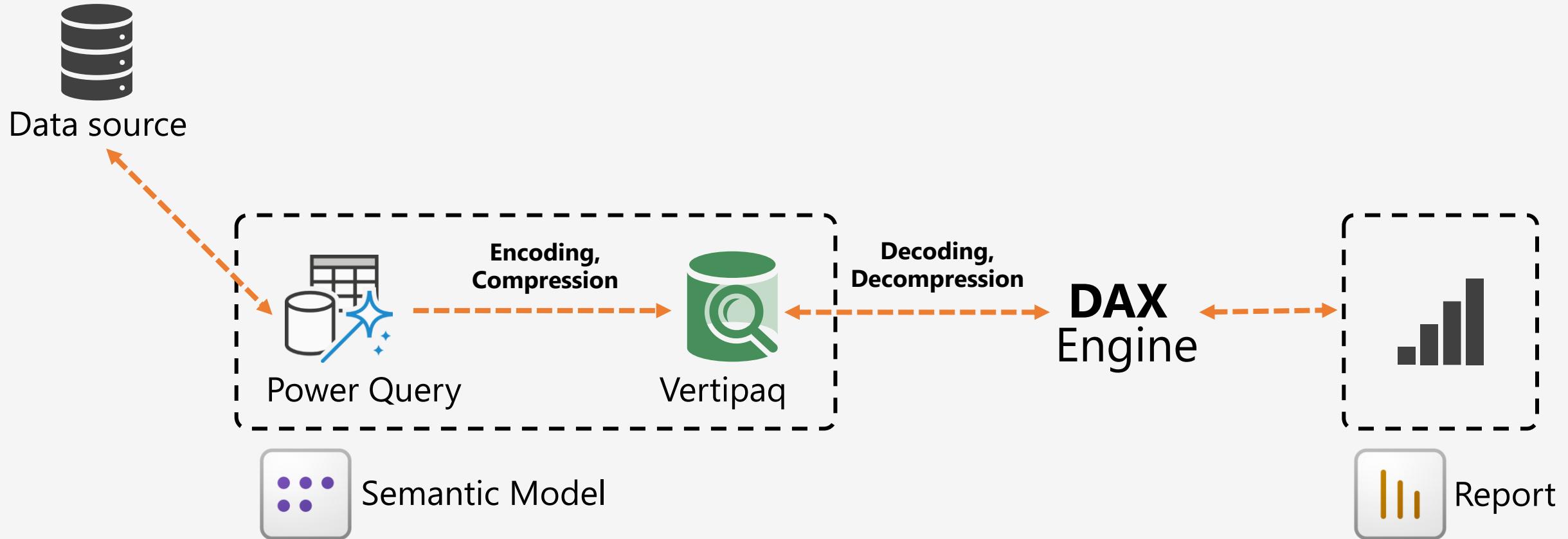


2 GB

- Longer refresh times
- Slower query performance
- Longer waiting time while editing & publishing the model
- Higher memory requirements
- Higher data transfer costs
- Higher data storage cost
- More frequent failures = more retries
- Higher electricity consumption = more Carbon emissions



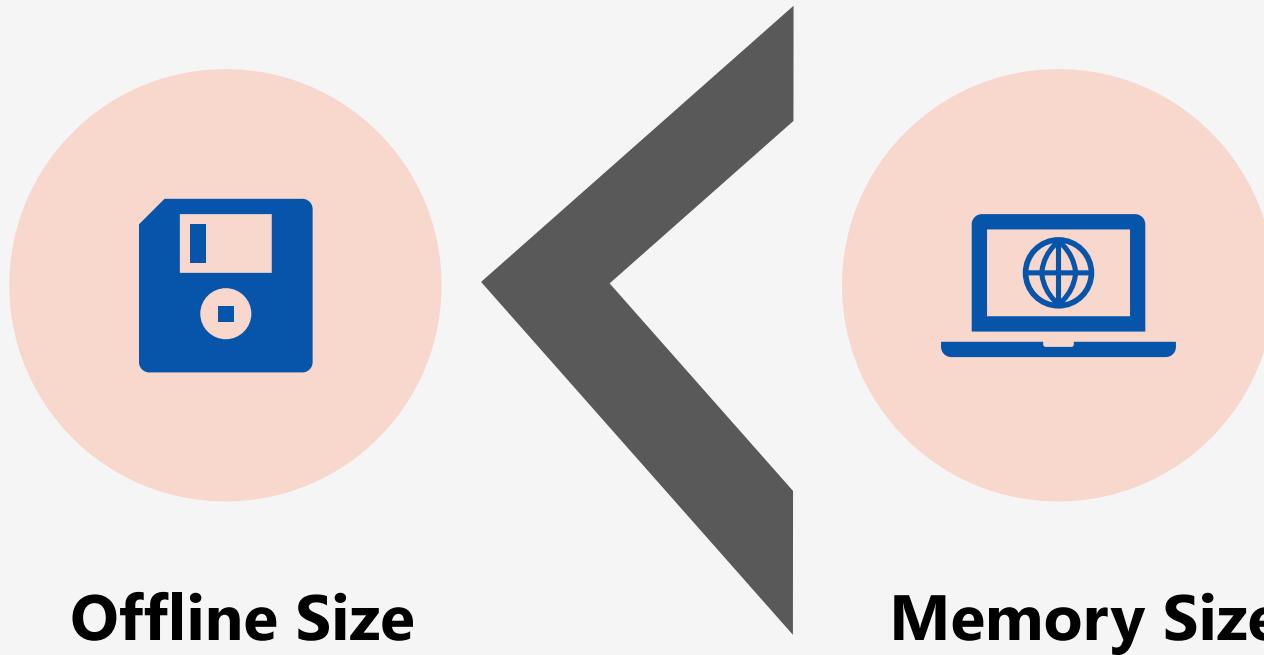
Typical Data Flow



*Oversimplified for clarity



Semantic Model Size



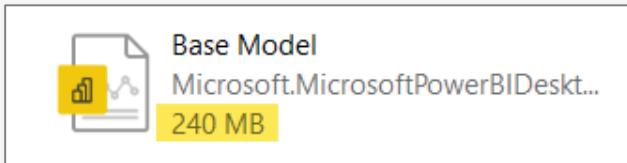
Memory size is always greater than the offline size



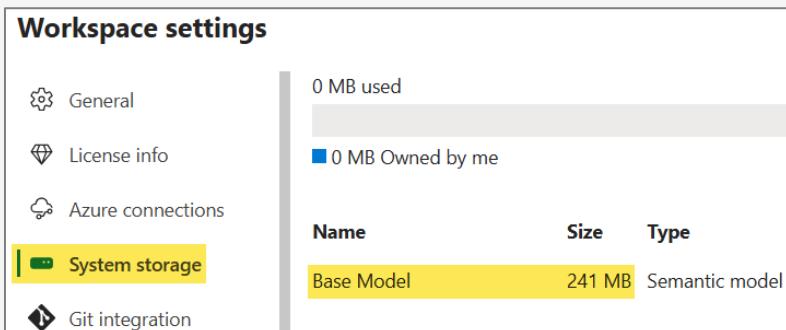
How can I find out the size of my model?

Offline Size

- **Local machine:** Windows explorer



- **Workspace:** Workspace settings > System storage

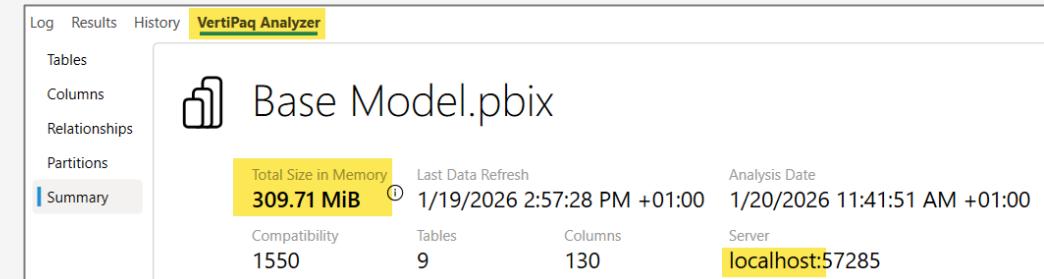


Workspace settings

General 0 MB used
License info 0 MB Owned by me
Azure connections
System storage Base Model 241 MB Semantic model
Git integration

Memory Size

- **Local machine:** Vertipaq Analyzer

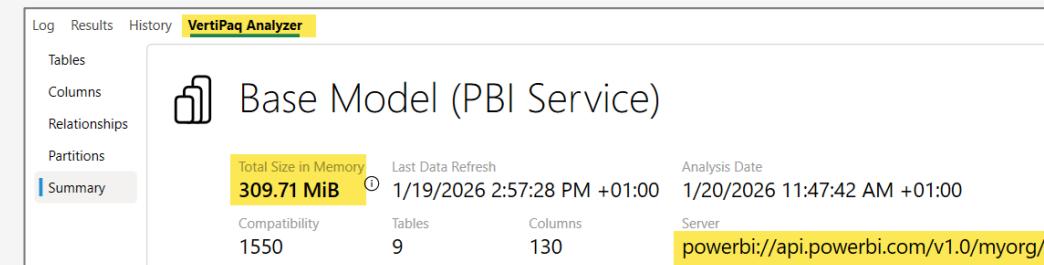


Vertipaq Analyzer

Tables Base Model.pbix
Columns
Relationships
Partitions
Summary

Total Size in Memory: **309.71 MiB** Last Data Refresh: 1/19/2026 2:57:28 PM +01:00 Analysis Date: 1/20/2026 11:41:51 AM +01:00
Compatibility: 1550 Tables: 9 Columns: 130 Server: localhost:57285

- **Workspace:** Vertipaq Analyzer



Vertipaq Analyzer

Tables Base Model (PBI Service)
Columns
Relationships
Partitions
Summary

Total Size in Memory: **309.71 MiB** Last Data Refresh: 1/19/2026 2:57:28 PM +01:00 Analysis Date: 1/20/2026 11:47:42 AM +01:00
Compatibility: 1550 Tables: 9 Columns: 130 Server: powerbi://api.powerbi.com/v1.0/myorg/





TIP 1: Disable auto date/time





Auto date/time

- When the auto date/time is enabled, Power BI Desktop creates a hidden auto date/time table for **each date column** in the model.
- These tables are **permanently hidden**. They don't appear in the Fields pane or the Model view diagram, or in the Table view.
- The table and its column **can't** be directly referenced by DAX expressions.
- They don't work when using **Analyze in Excel**.
- What's the point of auto date/time tables?
 - To support convenient time intelligence reporting based on date columns loaded into a model

Best practice: **Disable auto date/time and add a date table in your semantic models**





Disable auto date/time

- Auto date/time can be configured *globally* or for the *current file*
 - Global* option applies to new Power BI Desktop files
 - Current file* option applies to the current file
- They can be turned on or off at any time.
- Recommendation:** Disable auto date/time for both global and current file
 - File > Options and settings > Options > GLOBAL > Time intelligence
 - File > Options and settings > Options > CURRENT FILE > Time intelligence

Options

GLOBAL

- Data Load** (highlighted with red box)
- Power Query Editor
- DirectQuery
- R scripting
- Python scripting
- Security
- Privacy
- Regional Settings
- Updates

When you load data into Power BI (via import or DirectQuery) are evaluated simultaneously instead of one-by-one, which can adjust the default number of simultaneous query evaluations

Maximum number of simultaneous evaluations

Maximum memory used per simultaneous evaluation (MB)

Time intelligence (highlighted with red box)

Auto date/time for new files ⓘ Learn more

Data Cache Management Options ⓘ

Options

GLOBAL

- Data Load
- Power Query Editor
- DirectQuery
- R scripting
- Python scripting
- Security
- Privacy
- Regional Settings
- Updates
- Usage Data
- Diagnostics
- Preview features
- Auto recovery
- Report settings
- Copilot (preview)

Type Detection

Detect column types and headers for unstructured sources

Relationships

Import relationships from data sources on first load ⓘ

Update or delete relationships when refreshing data ⓘ

Autodetect new relationships after data is loaded ⓘ

Learn more

Time intelligence (highlighted with red box)

Auto date/time ⓘ Learn more

Background Data

Allow data previews to download in the background

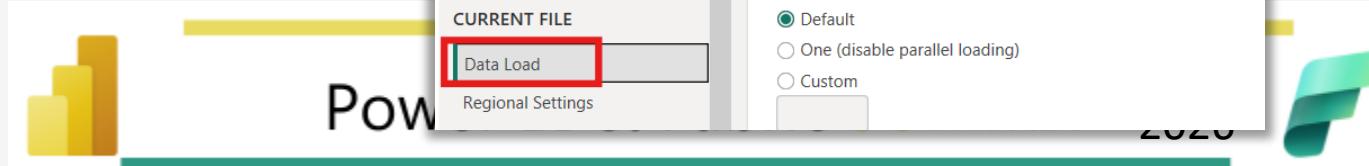
Parallel loading of tables ⓘ

Maximum number of concurrent jobs Learn more

Default

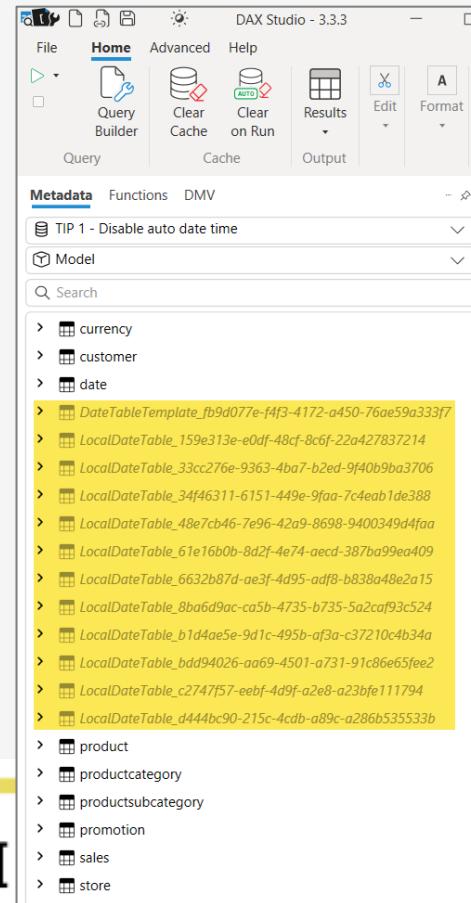
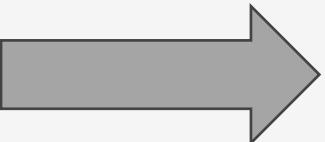
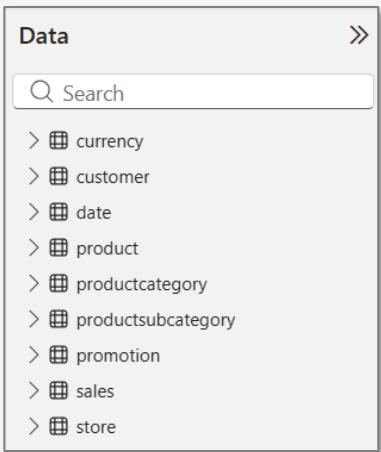
One (disable parallel loading)

Custom



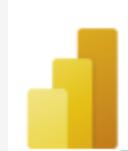
Tool: DAX Studio

- Is there a way to view the auto date/time tables when they are hidden from the model?
- How can we find out how much space the auto date/time tables use?





TIP 2: Only load the rows and columns that are necessary





Only load the rows that are necessary

- Do you really need all that data in your fact tables?
- If you're making a report only for actuals, then why load data for forecasts?
- Filter out the data whenever you can and load only the rows necessary for your report
- You can always load the additional data later when required
- Following dimensions can be a good starting point:
 - Years
 - Scenarios
 - Geography





Remove the columns that are not required

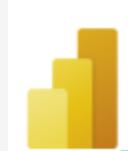
Types of columns in a table

- **Primary or alternate key:** column contains unique value for each row - *keep only if used for relationships*
- **Qualitative attributes:** column can be text or number, used to group or filter rows - *keep*
- **Quantitative attributes:** column is a number and used for calculations or filtering - *keep*
- **Descriptive attributes:** column contains text providing additional information about a row but never used in filters or calculations - *keep only if required for drill-through operations*
- **Technical attributes:** Information recorded for technical reason without a business value - *keep only if required for drill-through operations*





TIP 3: Set the correct data type for columns





Set the correct data type for columns

- Different data types consume different amounts of storage space
- Data type affects the encoding method used by the VertiPaq Engine
- The encoding method affects the final storage acquired by the column
- **Look out for data type impostors:**
 - integers stored as decimals or text
 - decimals stored as text





TIP 4: Optimize datetime columns





Concept: Column Cardinality

- The number of unique values in a column.
- This number is important for reducing column size and directly affects VertiPaq performance.
- DAX operations, including iterations and filters, depend directly on this number.
- **Column cardinality is more significant than the total number of rows in the table.**

What is the minimum level of granularity at which the data should be presented in the report?



Tool: Vertipaq Analyzer

- How can you check the cardinality of a column?
- Tool to analyze VertiPaq storage structures and data model details in Power BI.
- Available with DAX Studio: **Advanced > View Metrics**



**VertiPaq
Analyzer**



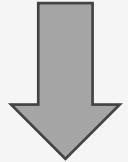
	Name	Cardinality	Total Size	Data	Dictionary	Hier Size	Encoding	Data Type
Tables	temp_activity_logs	670,783	51,672,195	4,056,440	38,996,491	8,177,408	Many	-
Columns	semantic_model_size_metrics	502,447	6,758,038	3,494,776	2,616,750	645,792	Many	-
Relationships	MaximumMemoryInGB	27,127	1,852,417	818,912	816,481	217,024	HASH	String
Partitions	MedianMemoryInGB	24,020	1,648,924	713,728	743,020	192,176	HASH	Double
Summary	MinimumMemoryInGB	22,327	1,494,664	592,120	723,920	178,624	HASH	Double
	SemanticModelID	6,716	1,275,709	910,016	311,949	53,744	HASH	String
	Date	268	257,780	244,992	10,628	2,160	HASH	DateTime
	ExtractionDate	256	227,560	214,880	10,616	2,064	HASH	DateTime





Cardinality of datetime columns

- Do our users prefer the report in milliseconds or in seconds?
- Do you **really** need the datetime to the lowest detail?
 - **Date:** 10 years of date = 3,700 unique values
 - **Time:** 1 day = unique value based on the precision



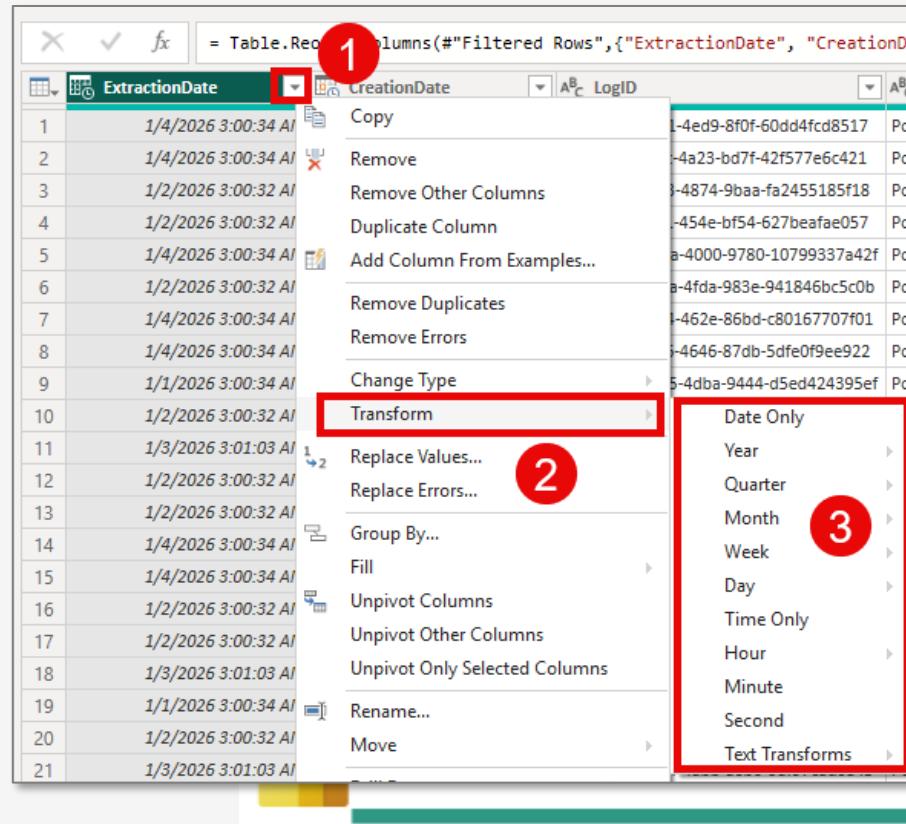
Unit	Cardinality
Hour	24
15 Minutes	96
5 Minutes	288
Minute	1,440
Second	86,400
Milisecond	86,400,000

The higher the cardinality of the column, the higher the storage required for the column – reduce cardinality to reduce size



Optimize datetime columns

- Remove the time part if not required
- Round the time to acceptable granularity, e.g., 5 second interval, 30 minutes interval, etc.
- Split the date and time into two different columns





TIP 5: Optimize high precision (decimal) columns





Cardinality of high precision columns

- Do our users like to see a value like 123.45678 in the report?
- Do we really need to keep the floating columns to 5 digits after the decimal?
- In most cases, it's acceptable to show 123.5 or 123.46

Each digit increases the cardinality of the column





Optimize high precision columns

- Round the column to acceptable granularity, e.g., 2 digits after the decimal
- Split the integer and decimal parts into two different columns

The screenshot shows the Power BI Data Editor interface. A context menu is open over a column named "1.2 MinimumMemoryInGB". The menu items are:

- Copy (highlighted with a red circle)
- Remove
- Remove Other Columns
- Duplicate Column
- Add Column From Examples...
- Remove Duplicates
- Remove Errors
- Change Type (highlighted with a red rectangle)
- Transform (highlighted with a red rectangle)
- Replace Values...
- Replace Errors... (highlighted with a red circle)
- Group By...
- Fill
- Unpivot Columns
- Unpivot Other Columns
- Unpivot Only Selected Columns

The "Transform" option is expanded, showing:

- Round (highlighted with a red rectangle)
- Absolute Value
- Factorial
- Base-10 Logarithm (highlighted with a red circle)
- Natural Logarithm
- Power
- Square Root
- Text Transforms

Column 1	Column 2
0.0021	10/19/2025
0.0632	10/19/2025
6.6753	10/19/2025
32.1051	10/19/2025
0.0909	10/19/2025
28.9994	10/19/2025
0.7374	10/19/2025
11.3409	10/19/2025
0.0078	10/19/2025





Concept: Roche's Maxim of Data Transformation

Data should be transformed as far “upstream” as possible, and as far “downstream” as necessary

“upstream” means closer to where the data is originally produced

“downstream” means closer to where the data is consumed



Have the data already available in the format you need it to be in



Ask the data source team to do the transformations in the source if possible





Summary

- **TIP 1:** Disable auto date/time
- **TIP 2:** Only load the rows and columns that are necessary
- **TIP 3:** Set the correct data type for columns
- **Column Cardinality:** Reduce the number of unique values in a column
 - **TIP 4:** Datetime columns: Only keep the lowest level of time detail required for your report
 - **TIP 5:** Decimal columns: Round to a decimal digit sufficient for your report
- Use Vertipaq Analyzer to analyze the storage distribution of your model
- **Roche's Maxim:** Move the data transformations closer to the data source





Get the slides...



https://github.com/DhyanRathore/Presentations/tree/main/2026/20260224_PowerBI&FabricSummit2026

Stay online for my live Q&A session

Thank You!!

Dhyan Rathore

 **dhyans**

bits2BI



Power BI & Fabric **SUMMIT**

2026

