FABRIC HANDBOOK

Focused on DP-700 certification exam

by Eugene Mazarakis

Table of Contents

About The Writer

"Anything worth having takes time."

Meet Eugene Mazarakis, a senior Data Engineer with a wealth of experience across diverse industries, including energy, banking and telecommunications. Over the course of numerous challenging projects, Eugene has honed his expertise in designing, implementing, and optimizing data solutions tailored to complex business needs.

He combines technical prowess with a deep understanding of industry demands, making him a trusted resource for transforming complex concepts into actionable solutions.

He holds a Master of Science degree in Theoretical Computer Science from the National and Kapodistrian University of Athens.

Microsoft CERTIFIED
FABRIC DATA ENGINEER
ASSOCIATE

Microsoft CERTIFIED
FABRIC ANALYTICS ENGINEER
ASSOCIATE

Microsoft CERTIFIED
POWER BI DATA ANALYST
ASSOCIATE

FABRIC EXAMINATION

You can take the examination online, which consists of approximately 60 questions, including multiple-choice questions and use cases.

You will have 1 hour and 40 minutes to complete it. If you are unsure about an answer, you can mark the question and review it later.

The passing score for the examination is 700.

FABRIC

Datastores

What are the datastores,
when do you need them,
and what do they contain.

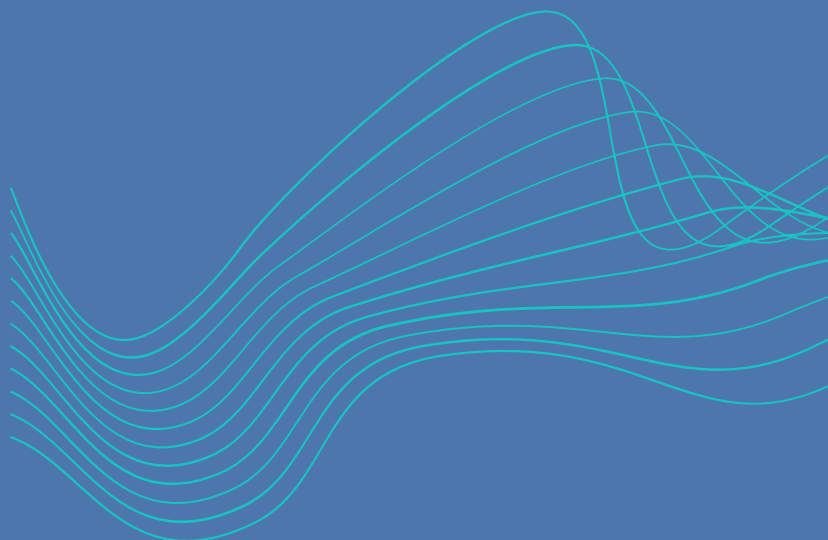# Fabric Data Stores

Lakehouse

It's presented as a database above the Data Lake and uses Delta Format tables.

Warehouse

It's traditional version of a DWH, data is stored in a Delta format and can be queried through SQL.
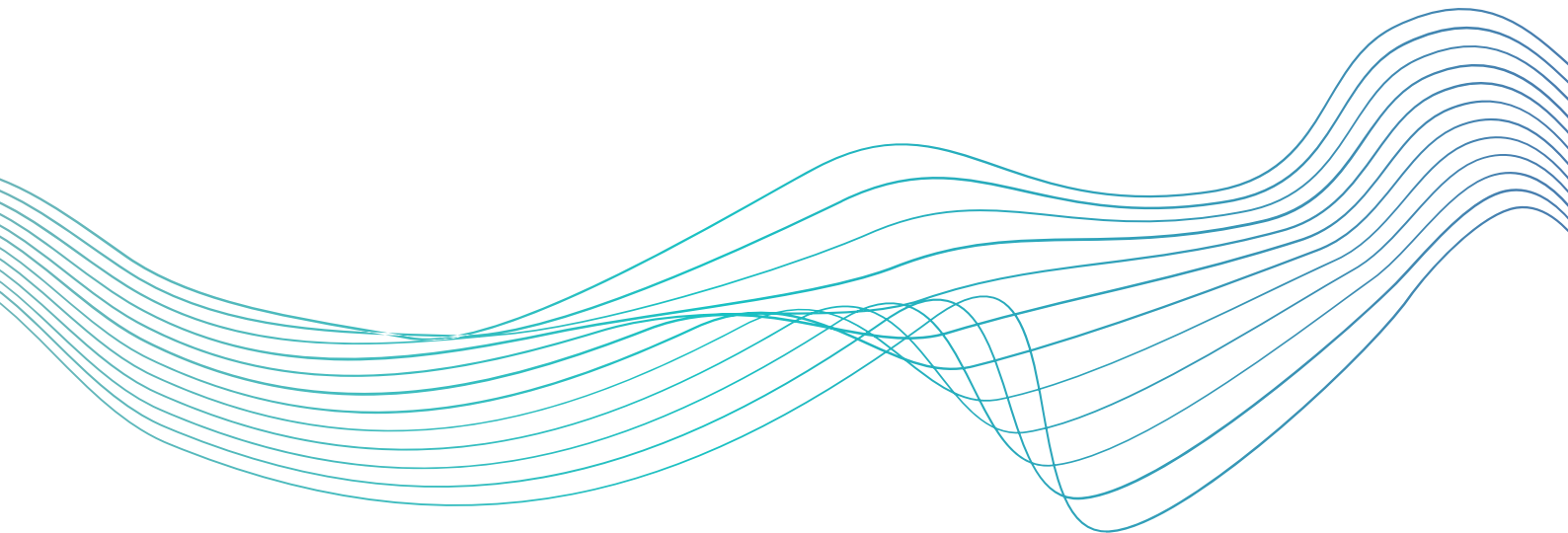
KQL database

It's used to analyze huge amount of data for Real-Time intelligence.

Data Types

	Structured	Semi-Structured	Unstructured	Real-Time
Lakehouse	✓	✓	✓	
DWH	✓			
KQL db				✓

Lakehouse

Shortcuts

Shortcuts are objects in OneLake that point to other storage locations. The location can be internal or external to OneLake.

Some types of sources that you can create shortcut are:
- Amazon S3
- Azure Data Lake Storage (ADLS) Gen 2
- Azure Blob Storage
- Google Cloud Storage

Caching

Shortcut caching can reduce egress costs associated with cross-cloud data access. As files are read through an external shortcut, the files are stored in a cache for the Fabric workspace.
- *Individual files greater than 1 GB in size aren't cached.*
- *Shortcut caching is currently supported for GCS, S3.*

Warehouse

Constraints

SQL analytics endpoint and Warehouse in Microsoft Fabric support these table constraints:

- PRIMARY KEY is only supported when NONCLUSTERED and NOT ENFORCED are both used.
- FOREIGN KEY is only supported when NOT ENFORCED is used.
- UNIQUE constraint is only supported when NONCLUSTERED and NOT ENFORCED are both used.

RLS

Row-level security (RLS) enables you to use group membership or execution context to control access to rows in a database table. To Implement RLS you need a **function** & a **policy.**

Dynamic Data Masking

There are 2 useful functions for this:

- **partial** (prefix, [padding], suffix)
  - exposes the first (prefix) and last (suffix) letters.
- **default** (): For date and time data types, use 1900-01-01 00:00:00.0000000.

Warehouse

Dynamic Management Views (DMVs) Constraints
They are systemic views that they give information about the warehouse.
- **sys.dm_exec_sessions**: Info for every session.

Query Insights
An important view that contains is:
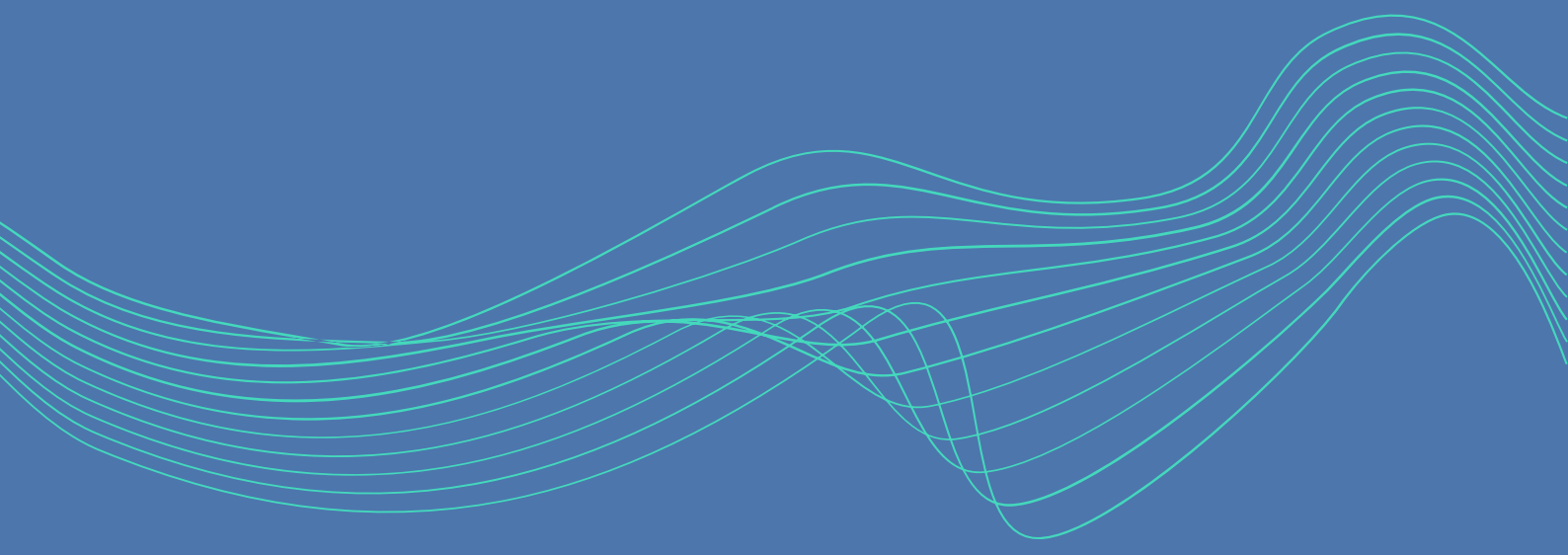- **long_running_queries**: Info for the execution time of the queries.

KQL

KQL is a simple yet powerful language to query structured, semi-structured, and unstructured data

Queries operators
- The SORT and ORDER operators are equivalent. Sorts the rows of the input table into order by one or more columns.
  - Default is **desc**, high to low.

- The WHERE and FILTER operators are equivalent. Filters a table to the subset of rows that satisfy a predicate.

POWER BI

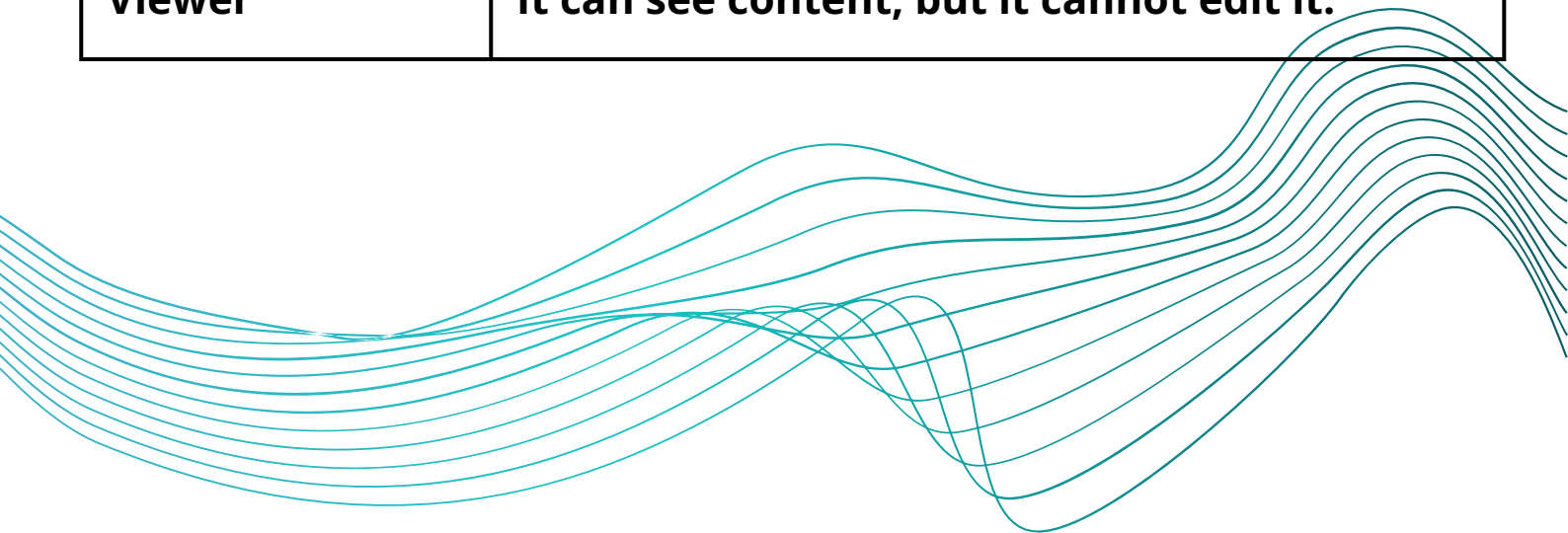## 01

Workspace Level Sharing

Information about roles assigned to a workspace.

Workspace Level Sharing

Access can be granted to individuals or groups, through one of the following roles. The roles are applied to the workspace. A user, depending on the role, can also perform the corresponding actions.

Role	Actions
Admin	Update, delete workspace. Only this role can define another admin.
Member	Add new user to a workspace, with <= permission.
Contributor	It can see content and can edit it.
Viewer	It can see content, but it cannot edit it.

05

Fabric Pipelines

01

Activities

Some useful type of activities

02

Expressions
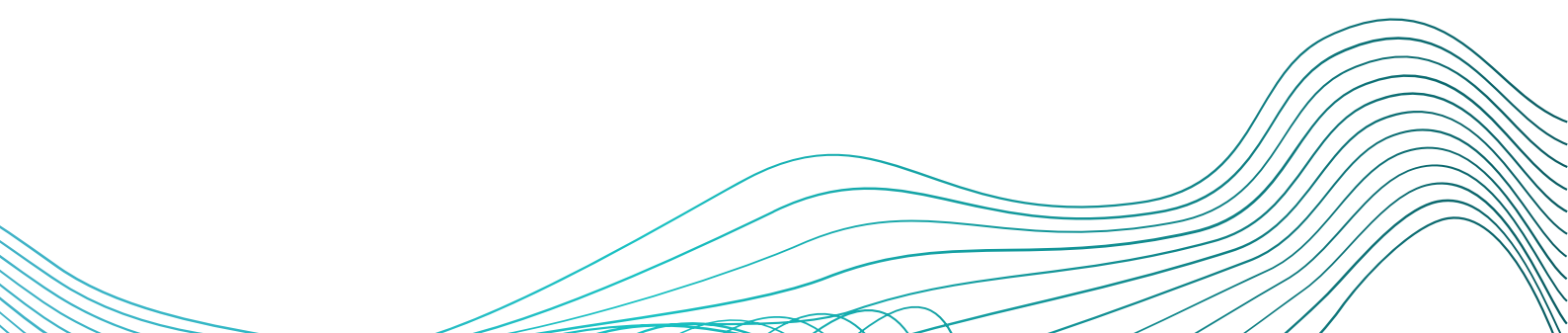
Some useful expressions for the expression builder.

Pipeline
Activities

For Each

The ForEach Activity defines a repeating control flow in a Microsoft Fabric pipeline. This activity is used to iterate over a collection and executes specified activities in a loop

Copy Data

The Copy activity connects to your data sources and destinations, then moves data efficiently between them
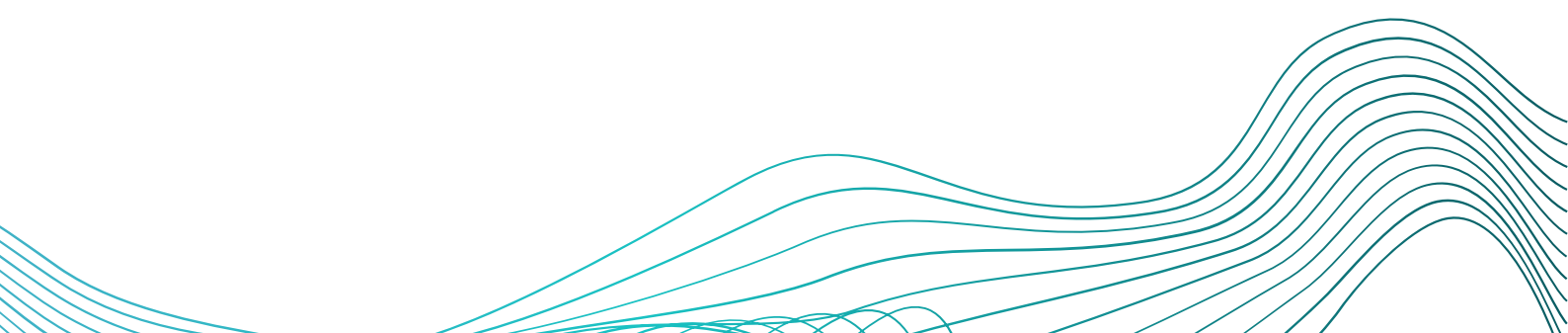
Pipeline
Expressions

@activity()

@activity('activity name').output.value
You can use @activity('activity name') to capture output of activity and make decisions.

@pipeline()

@pipeline ().parameters.paramName
A pipeline may have parameters. To retrieve the value of a parameter, you must use the above syntax.The result is the parameter value, returned in the data type that was defined.

SQL

Functions

Some useful functions, that
you need to know.

SQL
Functions

YEAR

YEAR (date)
Returns an integer that represents the year of the specified date.

CAST

CAST (expression AS data_type)
Convert an expression of one data type to another.

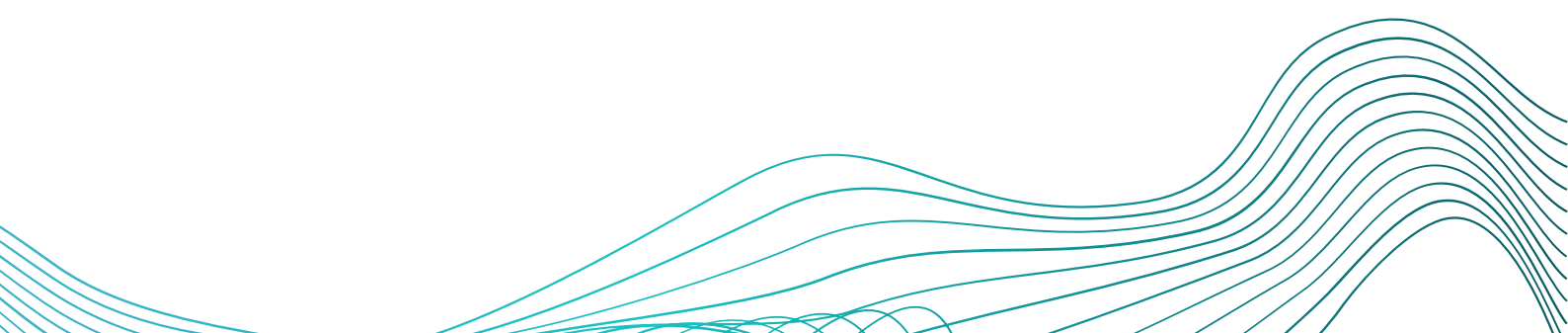**CONVERT**

CONVERT (data_type , expression [, style])
Convert an expression of one data type to another.
Style =7 → Mon dd, yy

TRY_CAST

TRY_CAST (expression AS data_type)
Returns a value cast to the specified data type if the cast succeeds; otherwise, returns null.

STAY UPDATED

YouTube   The 1nf0rmant

Blog   emazarakis.github.io

Linked In   eugene-mazarakis