SQL days konferenz

SQL days konferenz

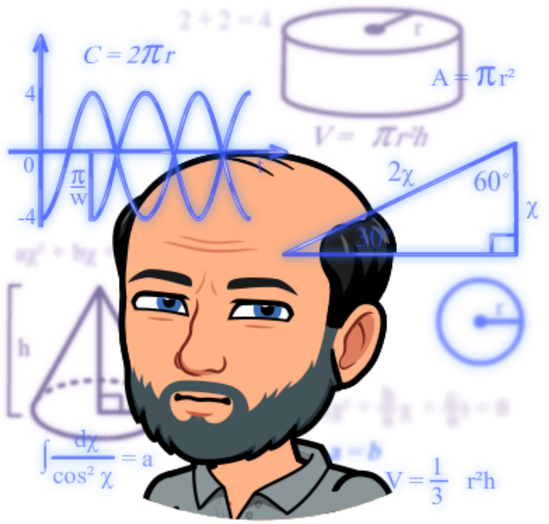
From SQL to KQL via Azure Data Explorer

Thomas Hütter

redgate

ppedv

From SQL to KQL
via Azure Data Explorer



Thomas Hütter

SQL Days Erding
07. - 10.10.2024

From SQL to KQL via Azure Data Explorer

Thomas Hütter, Diplom-Betriebswirt

- Developer for ERP apps, SQL scripts and BI stuff
- Worked at consultancies, ISVs, end user companies
- 1995: SQL Server, 2014: R, 2020: Power*, 2024: Arduino
- Speaker at SQL / data / dev events around Europe

@DerFredo https://twitter.com/DerFredo
de.linkedin.com/in/derfredo
https://techhub.social/@DerFredo

sessionize MOST ACTIVE SPEAKER 2022

SWECON

Developer! DeveloperDeveloper UK Community Events

sqlbits

{ } NDC Conferences

DATA INNOVATION SUMMIT 2023

SQL Konferenz 2023 powered by PASS Deutschland e.V.

Agenda

- Where we come from: SQL - Structured Query Language
- Starring: the `SELECT` statement
- Where we are going: KQL - Kusto Query Language
- Helping us along the way: ADX - Azure Data Explorer
- Hands-on examples: KQL in action
- Next level: visualizations, dashboards, time series
- Round-up, resources, learning

Where we come from: SQL - Structured Query Language

- The Structured Query Language (SQL for short) is a set-based declarative, domain-specific, cross-platform 4GL language that was first released by IBM in 1974.
- Based upon relational algebra and tuple calculus
  (☞ Edgar F. Codd's relational model of data).
- It is mainly used for managing relational data in an RDBMS.
- Standardized by ISO and ANSI, latest version ISO/IEC 9075:2023 or SQL:2023, but all the vendors roll their own set of additions, extensions, omissions to the standard.
- The concept of NULL values enforces 3-valued logic.
- Whitespace is generally ignored, allowing to format for good readability.

Where we come from: SQL - Structured Query Language

Lately, SQL has been divided into 4 sublanguages:

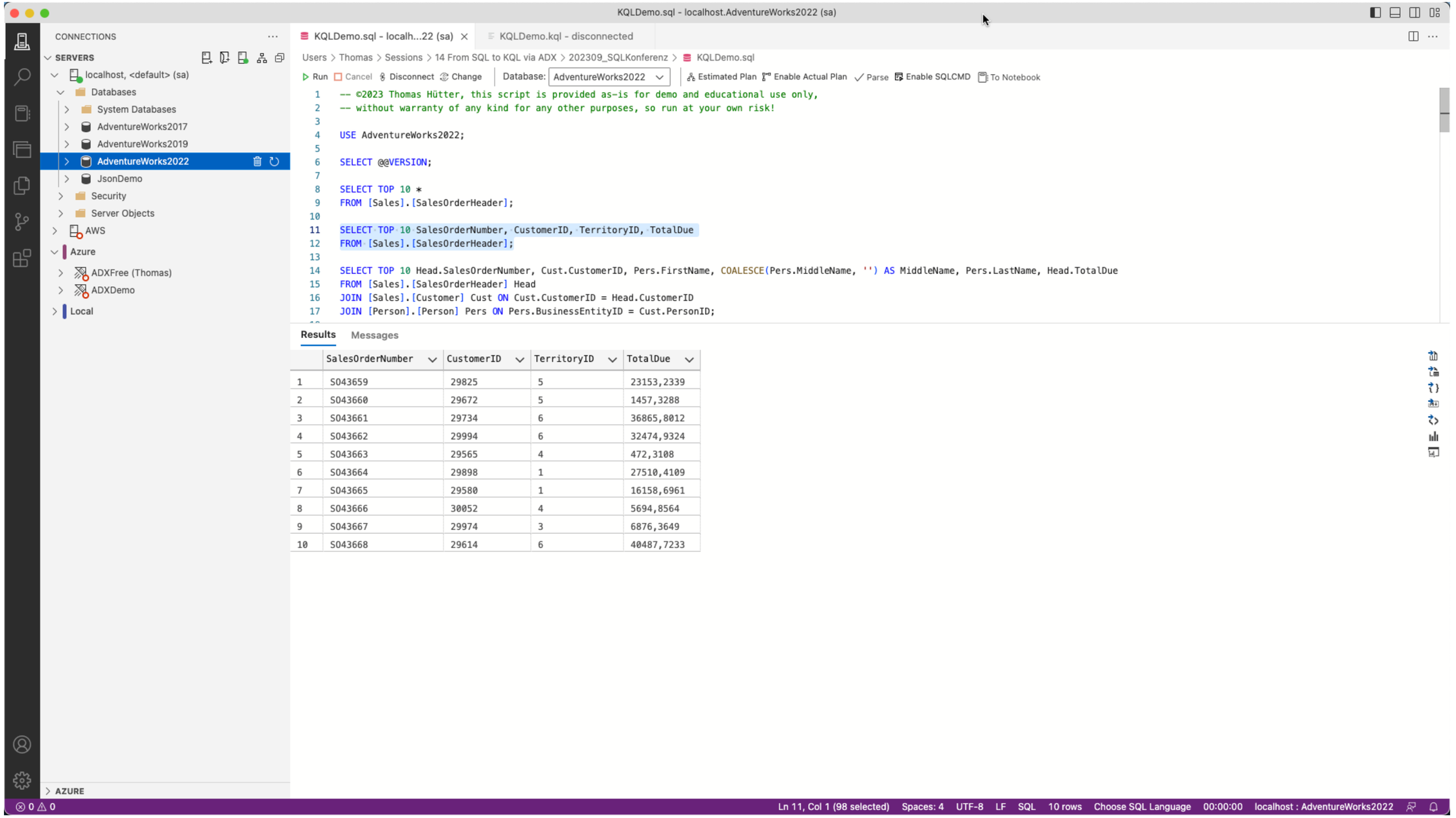
- DCL - Data Control Language
  `GRANT, DENY, REVOKE`
- DDL - Data Definition Language
  `CREATE, DROP, ALTER, TRUNCATE`
- DML - Data Manipulating Language
  `INSERT, UPDATE, DELETE`
- DQL - Data Query Language
  `SELECT`

Starring: the `SELECT` statement

The `SELECT` statement and it's clauses:

- 5. **SELECT** (8. **TOP** <n> | 6. **DISTINCT**) <columns>, <aggregates>
- 1. **FROM** <table1>
  (<jointype> **JOIN** <table2> **ON** <join_condition> …)
- 2. **WHERE** <constraint>
- 3. **GROUP BY** <column_list>
- 4. **HAVING** <constraint_on_group>
- 7. **ORDER BY** <column_list> (**ASC** | **DESC**)
- 8. **LIMIT** <n> / **OFFSET** <m>

Starring: the `SELECT` statement



Where we are going: KQL - Kusto Query Language

- Kusto* Query Language (KQL) is
  „… the query language used by Azure Data Explorer, Real-Time Analytics in Microsoft Fabric, Log Analytics in Azure Monitor, Microsoft Sentinel, Microsoft 365 Defender, Azure Resource Graph, and Resource Manager, among other Microsoft products. Querying data using KQL lets you gain insights about your IT, business, and security from large data sets you collect, in near real-time.“
- A Kusto query is
  a read-only request to the query engine to process data, stated in plain-text, which will be evaluated top-down and can consist of one or more statements, separated by a semicolon (;).
- A query statement can be
  a tabular expression, a „let“ statement or a „set“ statement; statements are separated by a pipe (|).
- A management / control command
  is used to retrieve/modify metadata or objects, ingest or export data etc…;
  always starts with a period (.); result may be piped as input to a query.

*: Why „Kusto“? Think of Jacques Cousteau diving the oceans to find hidden treasures 😉

Where we are going: KQL - Kusto Query Language

- A tabular expression
  is usually composed of a tabular data source (often a table reference or a table literal), tabular data operators such as filters or projections, optional rendering operators; each step connected by a pipe (|), expecting tabular input and providing tabular output (except graphic renderers).

- A rendering statement
  is used to render the result into a graphical output - generate a visualization
  (usually the last statement in a query).

- A „let“ statement
  defines a named variable that can hold a scalar or tabular value, a function or a virtual table.

- A „set“ statement
  sets an option value, valid for the scope of the query.

Helping us along the way: ADX - Azure Data Explorer

- Azure Data Explorer (ADX) is
  „… a fully managed, high-performance, big data analytics platform that makes it easy to analyze high volumes of data in near real time. The Azure Data Explorer toolbox gives you an end-to-end solution for data ingestion, query, visualization, and management.“

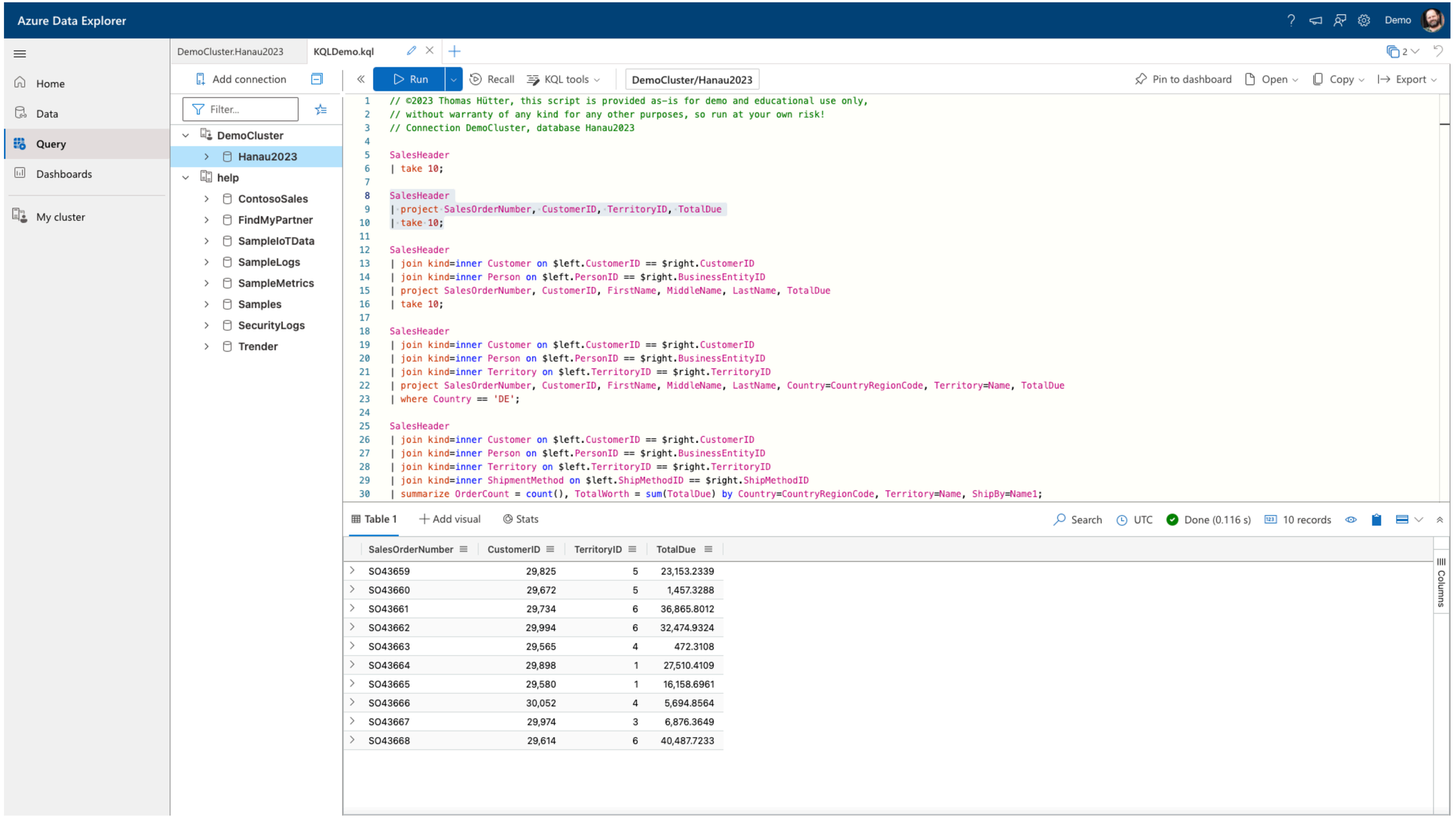
- ADX allows for advanced analytics, including versatile data visualization, using a user-friendly (i.e. easy to read, understand and learn) query language.

- Probably is „The most powerful Azure service you’ve never heard of“ (Patrick LeBlanc)

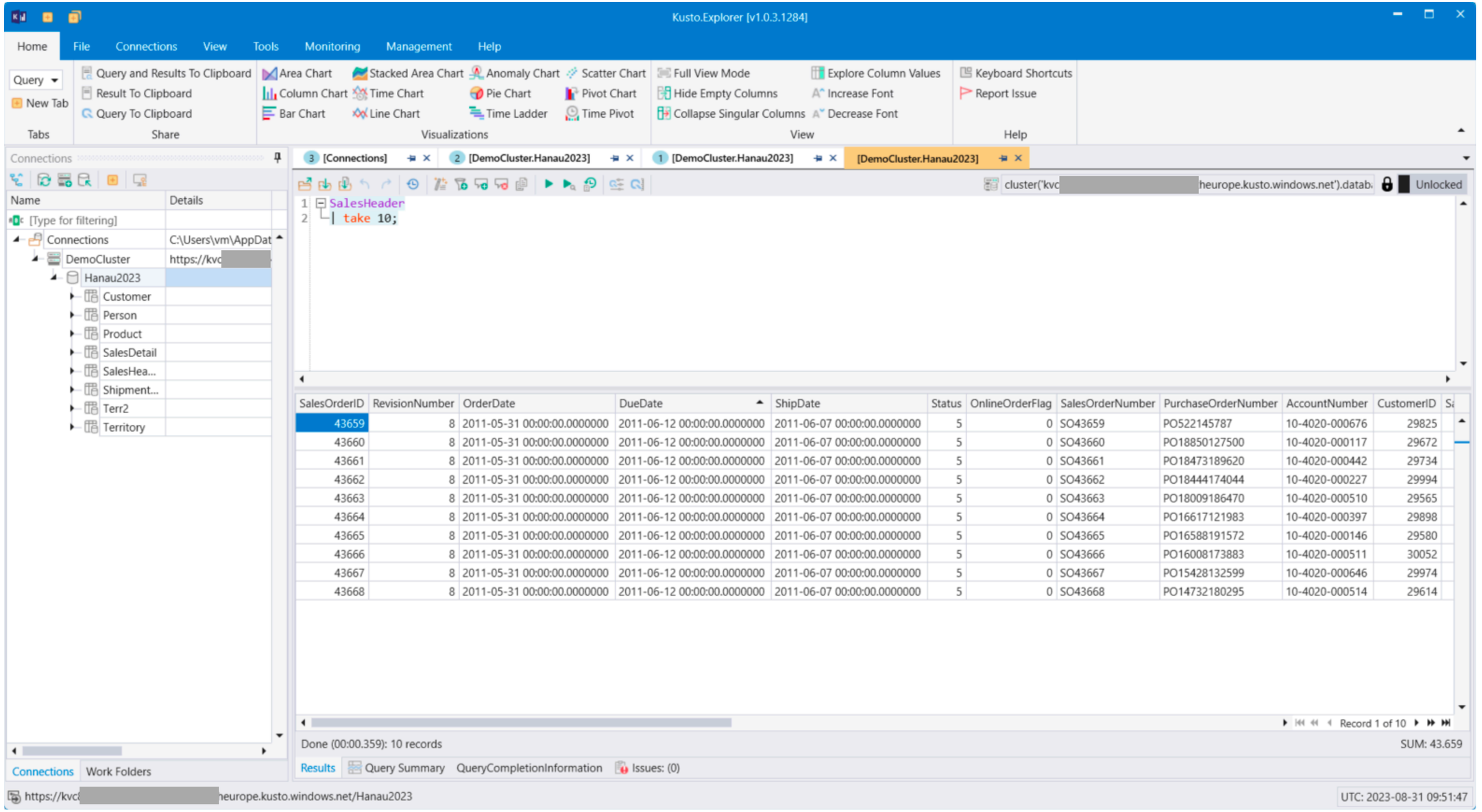
- Can be tried without costs by using the ADX „free cluster“ option.

- Alternatives: Kusto.Explorer (Windows app), Kusto extension for Azure Data Studio

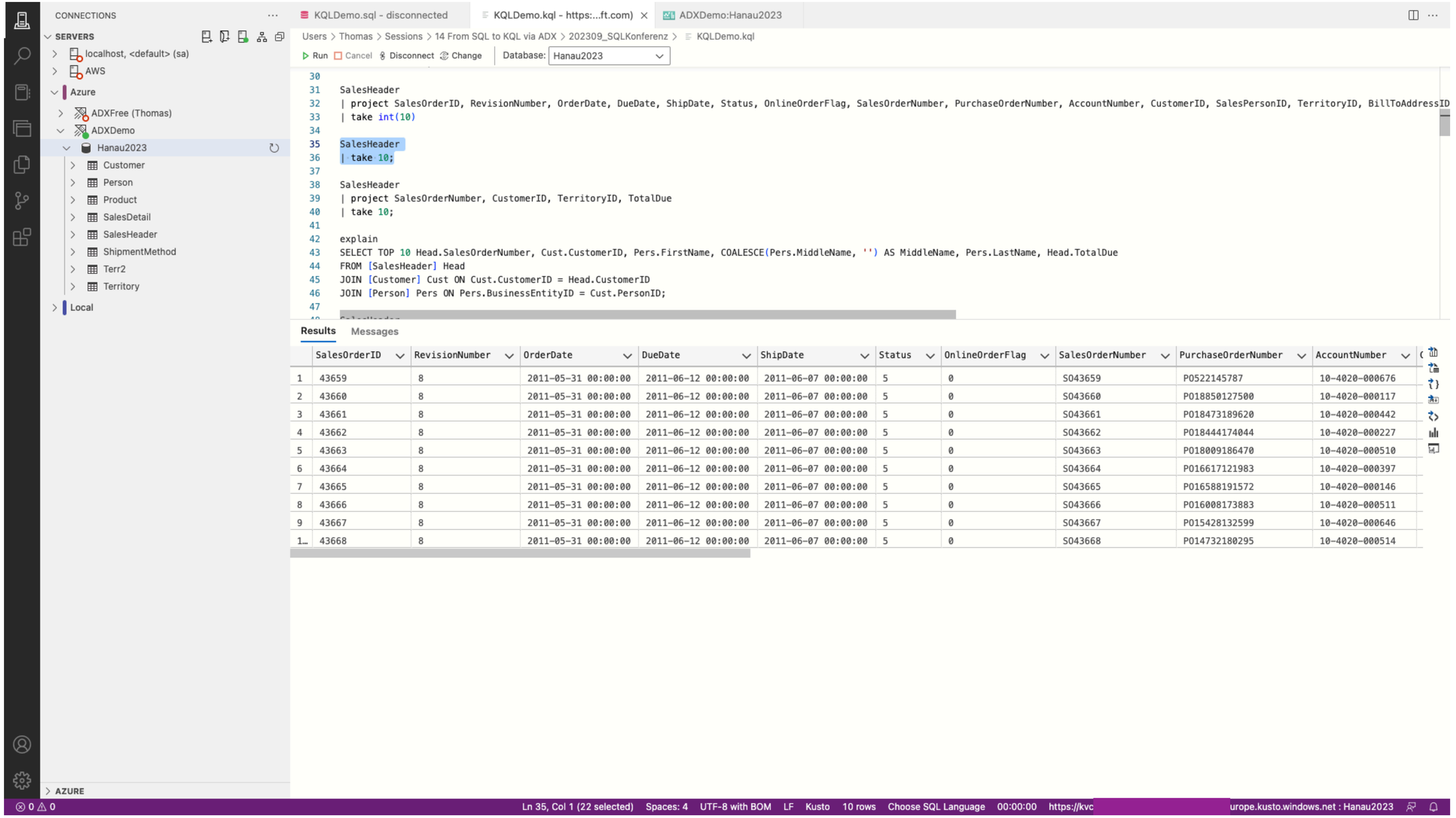
Helping us along the way: ADX - Azure Data Explorer



ADX - Alternative 1: Kusto.Explorer



ADX - Alternative 2: Kusto extension for Azure Data Studio



Hands-on examples: KQL in action

- Understands a limited subset of T-SQL
- „Explain“ function translates SQL to KQL (also limited 😕)
- Observe special behavior:
  - default sort order: descending
  - default equality check: case-sensitive (case-insensitive with =~)
  - `union` allows for differing table schemas
  - `search` works on *all* columns in a table (default case-*insensitive*)
  - `ago` allows for dynamic time intervals
  - …

Next level: visualizations, dashboards, time series

- Visualizations? Yes, easy, just add a `render` clause at the end of your query 😉.
- Several types of visualization possible, including maps (geospatial clustering).
- Then „Pin to dashboard“ does exactly that.
- Dashboards can be filtered, even dynamically.
- Dashboards can be exported and shared with others in your org (tenant).
- Built-in time series analysis capabilities
- Graphical analysis, anomaly detection, decomposition, forecasting

Resources:

- KQL overview: https://learn.microsoft.com/en-us/azure/data-explorer/kusto/query/
- SQL to Kusto cheat sheet: https://learn.microsoft.com/en-us/azure/data-explorer/kusto/query/sqlcheatsheet
- Azure Data Explorer: https://learn.microsoft.com/en-us/azure/data-explorer/data-explorer-overview
- ADX free cluster: https://dataexplorer.azure.com/freecluster
- Visualization overview: https://learn.microsoft.com/en-us/azure/data-explorer/viz-overview, dashboards: https://learn.microsoft.com/en-us/azure/data-explorer/azure-data-explorer-dashboards
- Microsoft learning paths: https://learn.microsoft.com/en-us/training/paths/kusto-query-language/, https://learn.microsoft.com/en-us/training/paths/data-analysis-data-explorer-kusto-query-language/, https://learn.microsoft.com/en-us/training/paths/analyze-monitoring-data-with-kql/
- ADX-in-a-day: https://github.com/Azure/ADX-in-a-Day

From SQL to KQL via Azure Data Explorer

- 𝕏 @DerFredo https://twitter.com/DerFredo
- de.linkedin.com/in/derfredo
- https://techhub.social/@DerFredo

This file and the demo scripts can be found at:
https://bit.ly/DerFredoErding2024

SQL days konferenz

Wir freuen uns auf Feedback!

Scan mich!

redgate

ppedv

SQL days konferenz

Vielen Dank!

… und gerne bis zum nächsten Mal!

redgate

ppedv