Fellowmind

From zero to Kusto hero at 30.000 ft

A unified analytics solution for the era of AI

sqlbits
2024

Brian Bønk Rueløkke

Principal & Enterprise architect, Data & AI

Fellowmind

https://linkedin.com/in/brianbonk
https://brianbonk.dk
https://github.com/brianbonk

Microsoft MVP

Microsoft FastTrack Recognized Solution Architect Power BI 2022 >>

Microsoft Certified Trainer Data Platform 2018 >>

AGENDA

The history of Kusto
Where does Kusto and RTA fit in the Data area
RTA in Fabric – incl. roadmap
Capabilities using Kusto
Get started for free
Introduction to the KQL language

BREAK

Kusto data in Power BI – with ninja tricks 🥷
Kusto Functions
Next level KQL language
Outliers
Visualization
Dash-boarding

Jaques Cousteau
1910-1997

Jaques Cousteau
1910-1997

Image is AI generated

The history of Kusto

Azure Sentinel

Log Analytics

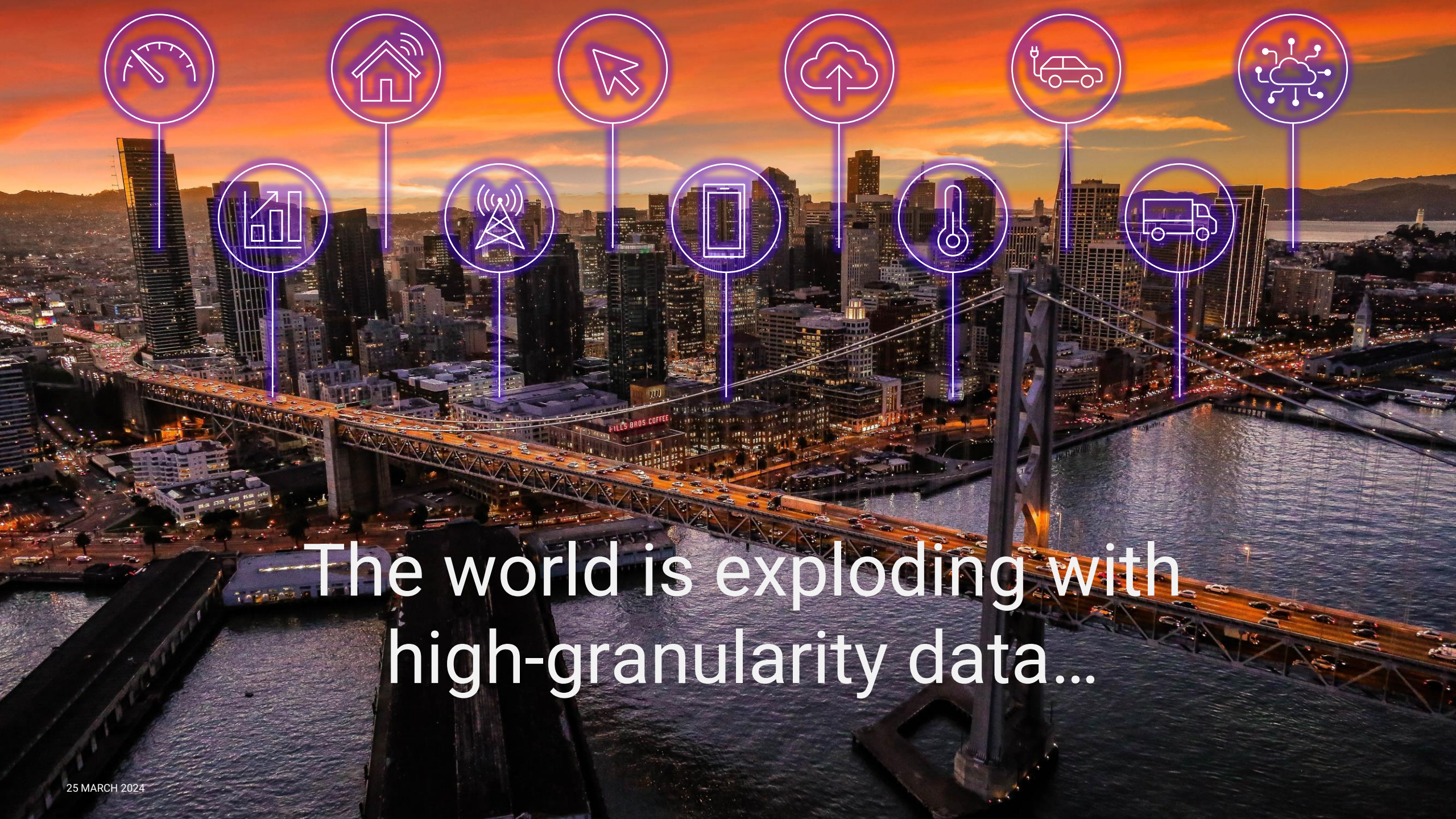
Real-Time Analytics

Azure resource graph

Microsoft 365 Defender

CMPivot

CMPivot



The world is exploding with high-granularity data...

It all starts with data

Telemetry – a key data for digital transformation

Telemetry – a key data for digital transformation

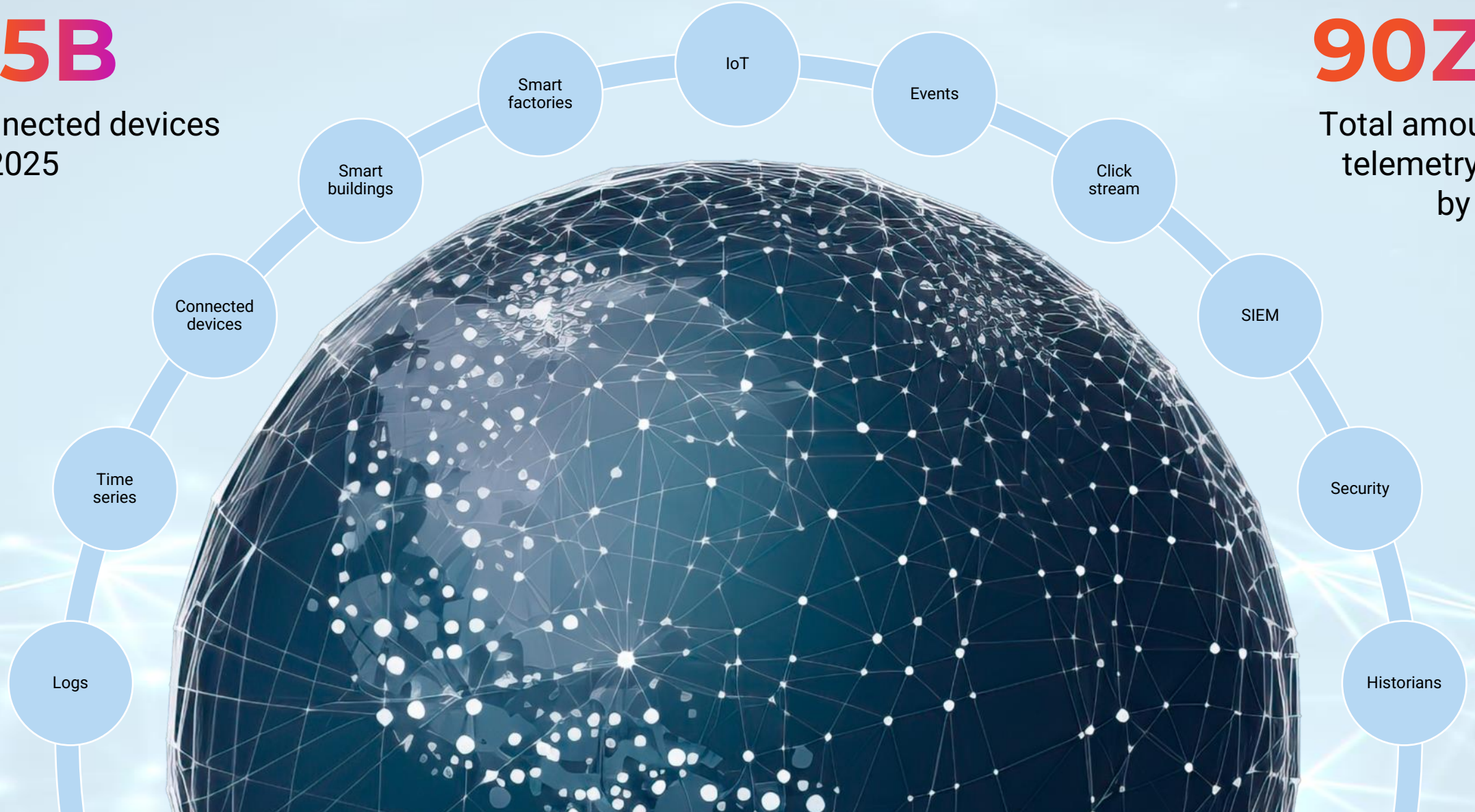
Telemetry – a key data for digital transformation

75B

Connected devices by 2025

90ZB

Total amount of telemetry data by 2025



Digital transformation

Cybersecurity
Asset tracking and management
Predictive maintenance
Supply chain optimization
Customer experience
Energy management
Inventory management
Quality control
Environmental monitoring
Fleet management
Health and safety

Microsoft Fabric

Get data

Data Factory

Prepare data

Synapse Data Engineering

Synapse Data Warehouse

Synapse Data Science

Synapse Real-Time Analytics

Use data

Data Activator

Power BI

Store data

OneLake

Microsoft Fabric

Get data

Data Factory

Prepare data

Synapse Data Engineering

Synapse Data Warehouse

Synapse Data Science

Synapse Real-Time Analytics

Use data

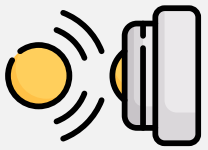
Data Activator

Power BI

Store data

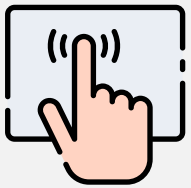
OneLake

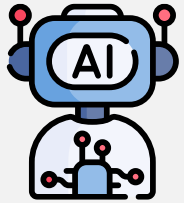
Microsoft Fabric


Event ingestion


Real-Time analytics

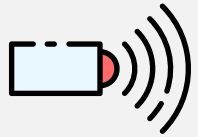

Real-Time dashboards


Real-Time triggers



Real-Time AI


Real-Time applications


Event driven actions

Get data


Data Factory

Prepare data


Synapse Data Engineering


Synapse Data Warehouse


Synapse Data Science


Synapse Real-Time Analytics

Use data


Data Activator

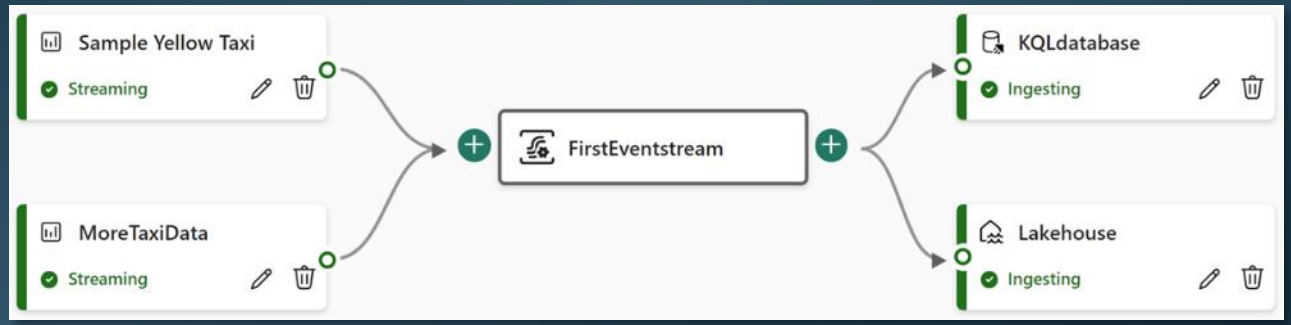

Power BI

Fabric Real-time Analytics solution enables organizations to consume **vast amount of data**, focus and **scale up** their Analytics solution with **data in motion, empower** their business analysts, and **democratize their data** for citizen data scientists and Data Engineers

Streaming data with ease

EVENTSTREAM

The brand-new event stream service, leverages the ability to get data from several sources of streaming data and save it to a wide variety of destinations, including OneLake, KQL databases and Azure services.



The service computes the data once and can pipe it out to several destinations at once. All configured and maintained from within the Microsoft Fabric portal and "coded" with your mouse.

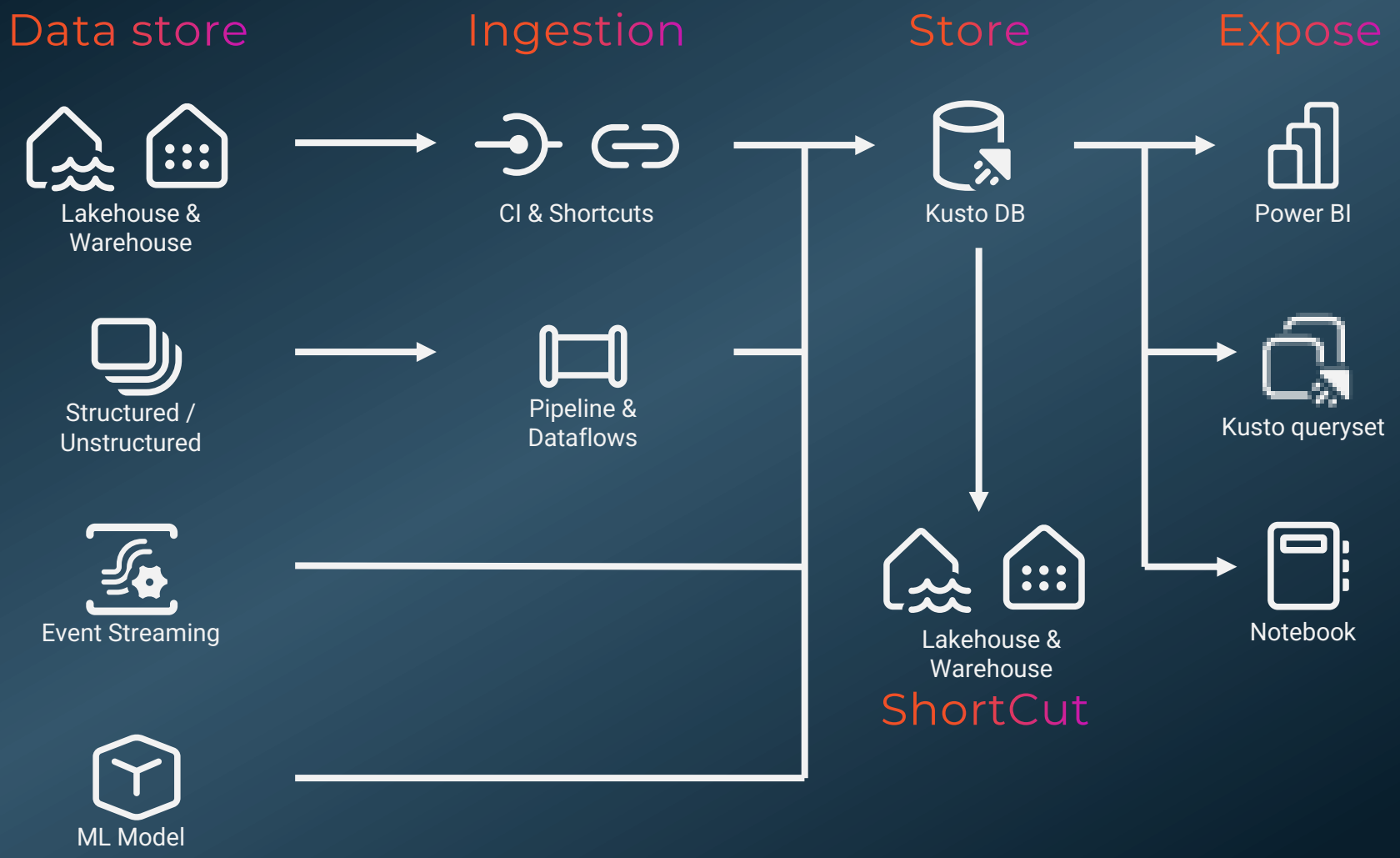
Imagine scenarios of IoT devices loading data to both the data warehouse and other 3-rd party destinations – this can now be done using the low-code approach from Event Stream.

Data Activator

KQL database
Key capabilities

Unlimited Scale (query, ingestion and storage)	Any data source	Any data format
Structured Semi-structured Free-text	Real-time transformation og complicated data strcutures	Streaming analytics in Near-Real-Time
High performance Low latency High freshness	Timeseries database	Everything is indexed and partitioned

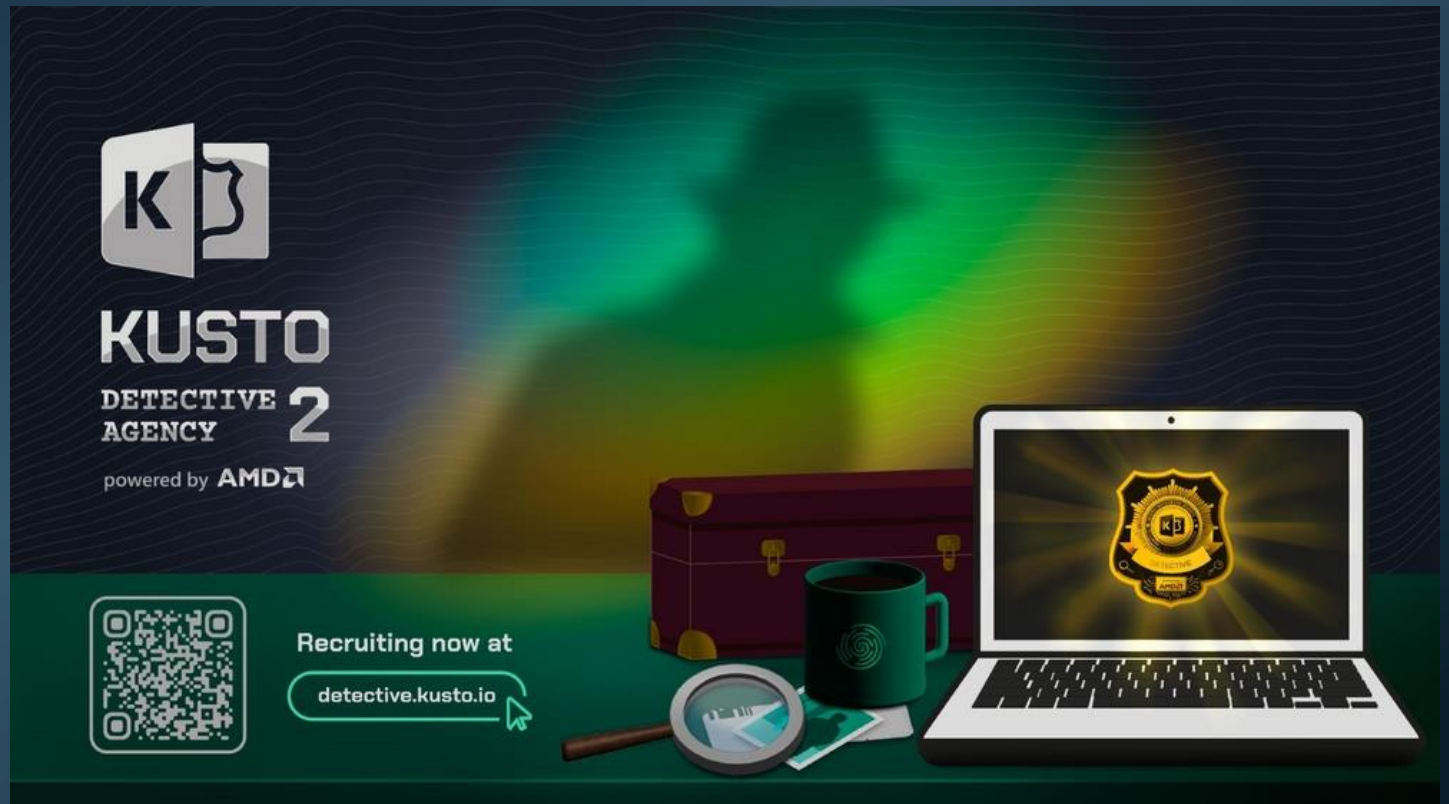
Real-Time Analytics



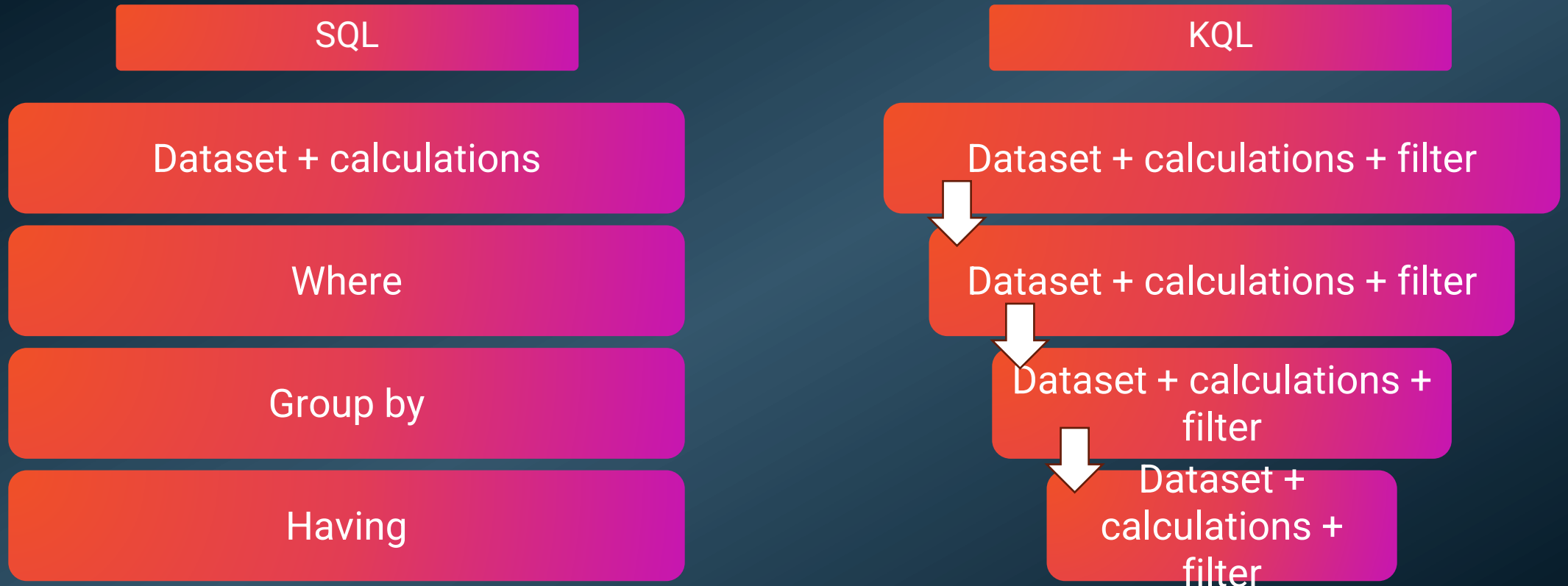
Get started for free

https://dataexplorer.azure.com/freecluster

https://detective.kusto.io



The language and structure



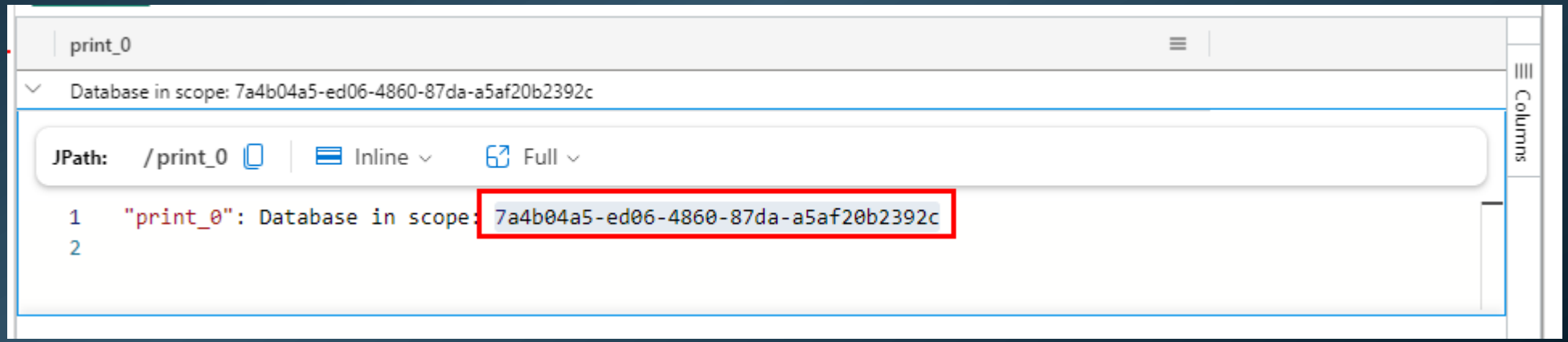
SQL to Kusto query translation - Azure Data Explorer & Real-Time Analytics | Microsoft Learn

The language and structure

Using the Notebook feature in Azure Data Studio to demo

Get the database id from your Fabric Kusto cluster

print strcat("Database in scope: ", current_database())



SQL to Kusto query translation - Azure Data Explorer & Real-Time Analytics | Microsoft Learn

DEMO

Live coding
(hopefully no demo-ghost 👻)

The language and structure

KQL: Kusto Query Language

SQL

```
select * from NYCTaxi
```

KQL

```
NYCTaxi
```

SQL to Kusto query translation - Azure Data Explorer & Real-Time Analytics | Microsoft Learn

The language and structure

SQL

```
select * from NYCTaxi
where VentorID = 2
```

KQL

```
NYCTaxi
| where VendorID == 2
```

SQL to Kusto query translation - Azure Data Explorer & Real-Time Analytics | Microsoft Learn

The language and structure

SQL

```
select * from NYCTaxi
where VentorID = 2
order by passenger_count
```

KQL

```
NYCTaxi
| where VendorID == 2
| order by passenger_count
```

SQL to Kusto query translation - Azure Data Explorer & Real-Time Analytics | Microsoft Learn

The language and structure

SQL

```
select count(*) from
NYCTaxi
```

KQL

```
NYCTaxi
| count
```

SQL to Kusto query translation - Azure Data Explorer & Real-Time Analytics | Microsoft Learn

The language and structure

SQL

```
select
    passenger_count
    ,VendorID
    ,trip_distance
from NYCTaxi
```

KQL

```
NYCTaxi
| project passenger_count, VendorID, trip_distance
```

SQL to Kusto query translation - Azure Data Explorer & Real-Time Analytics | Microsoft Learn

The language and structure

SQL

```
select
    passenger_count
    ,VendorID
    ,trip_distance
    ,total_amount / passenger_count as AmtPsngr
from NYCTaxi
```

KQL

```
NYCTaxi
| extend AmtPsngr = total_amount / passenger_count
| project passenger_count, VendorID, trip_distance,
AmtPsngr
```

SQL to Kusto query translation - Azure Data Explorer & Real-Time Analytics | Microsoft Learn

The language and structure

SQL

```
select
    sum(passenger_count) as SumPassenger
    ,VendorID
from NYCTaxi
group by VendorID
```

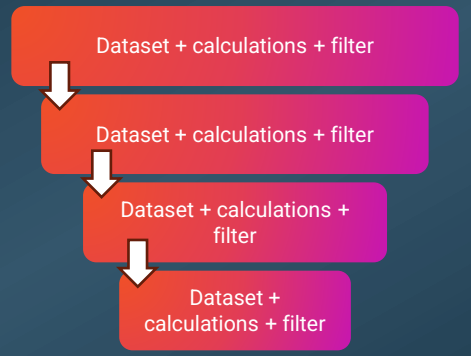
KQL

```
NYCTaxi
| summarize SumPassenger = sum(passenger_count) by
VendorID
```

SQL to Kusto query translation - Azure Data Explorer & Real-Time Analytics | Microsoft Learn

The language and structure

KQL




```
NYCTaxi
| where passenger_count > 1
| project passenger_count, total_amount, VendorID, fare_amount
| extend AmtPsngr = total_amount / passenger_count
| where AmtPsngr > 10
| summarize TotalAmount = sum(total_amount), AvgAmtPsngr = avg(AmtPsngr) by VendorID
| where VendorID  <> 1
```

SQL to Kusto query translation - Azure Data Explorer & Real-Time Analytics | Microsoft Learn

BREAK

Kusto in Power BI

Forget everything you know about

query performance vs data types

&

data modelling best practices

Data modelling Kusto in Power BI

- Single table reporting can be a good option, if you can include all columns from dimensions to the table
- M:M relations are hard to avoid, but not a big deal → all queries will be translated to KQL
- All dimensions must be tagged with “IsDimension=true”
- Dimensions can be imported if they are <1 mio rows.
- INTEGER and DECIMAL er slow joins compared to STRING

Harness the Power (BI) of Kusto

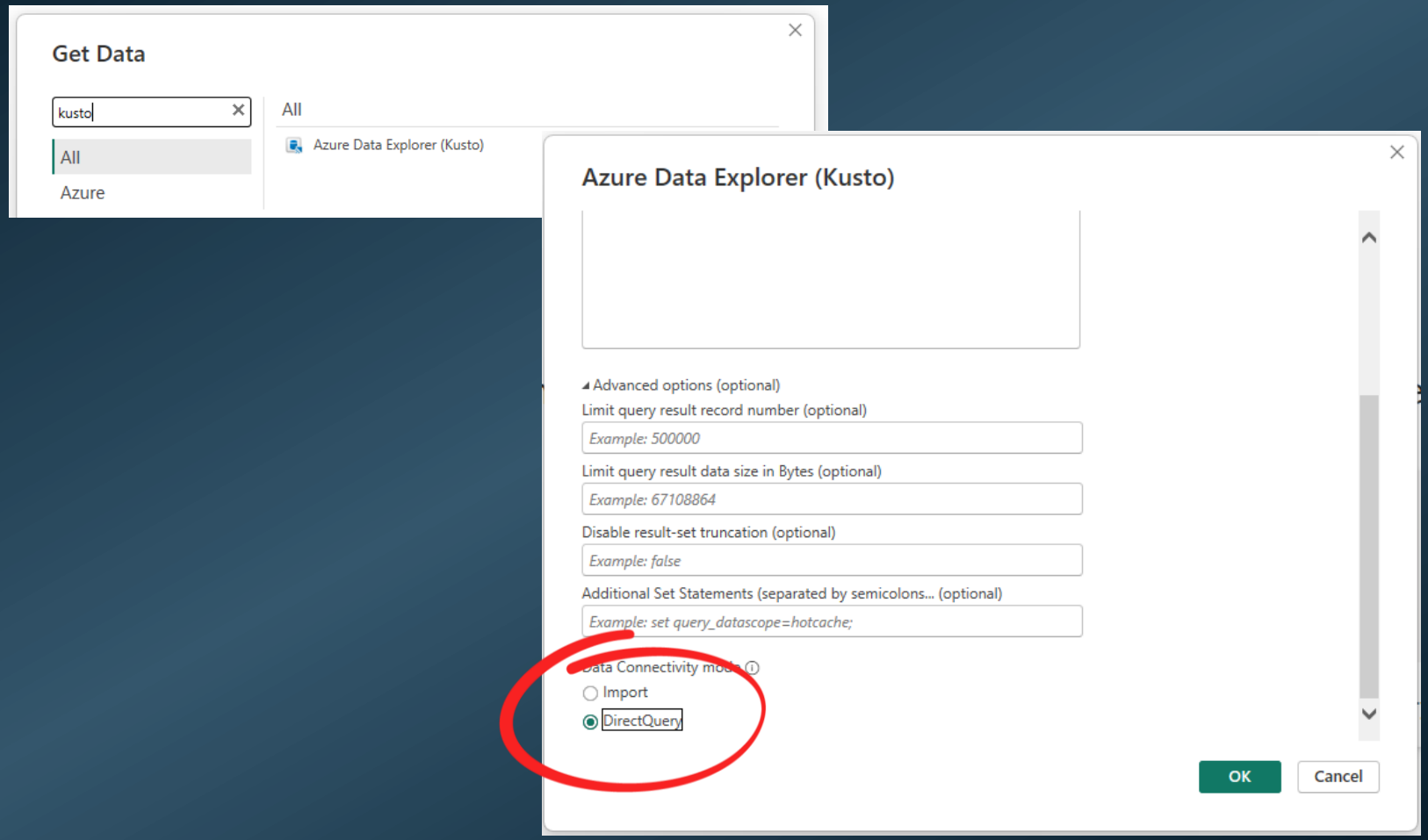
Let Power BI build the KQL

- In Power Query
- Using DAX

Or build a Kusto function

Analysis and reporting

Power BI



Functions

Functions in Kusto is equivalent to a stored procedure in the SQL world.

With additional functionality to be able to go outside of the cluster and service and ask for data from a different place in the world.

```
.create-or-alter function GetSysLogs(TimeWindow:string , Bucket:string )
{
cluster('help').database('SampleLogs').RawSysLogs
| where timestamp > ago(totimespan(TimeWindow))
| summarize LogCount=count() by name, bin(timestamp, totimespan(Bucket))
| order by timestamp asc
}

// to execute the function
GetSysLogs('5d','1h')
```

SQL to Kusto query translation - Azure Data Explorer & Real-Time Analytics | Microsoft Learn

DEMO

Live coding
(hopefully no demo-ghost 👻)

Data discovery and outlier detection

Data discovery is what we've just been through – use select statements and filter your data to find and explore the data given to you.

RENDERING!!

```
NYCTaxi
| where tpep_pickup_datetime between (datetime(2009-01-01)..datetime(2015-01-01))
| extend PickUpdate = startofday(tpep_pickup_datetime)
| summarize SumPsngrCount = sum(passenger_count) by PickUpdate
| project PickUpdate, SumPsngrCount
| render  timechart
    with(
        title = "timechart"
        ,xtitle = "Time"
        ,ytitle = "Fares"
    )
```

SQL to Kusto query translation - Azure Data Explorer & Real-Time Analytics | Microsoft Learn

Data discovery and outlier detection

Outliers

series_outliers() - LINK
series_decompose() - LINK
series_decompose_anomalies() - LINK
series_decompose_forecast() - LINK

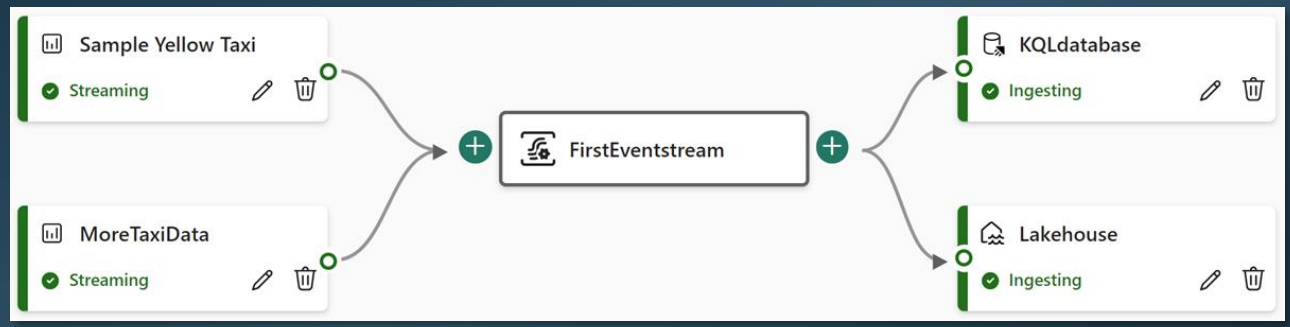
```
range x from 0 to 364 step 1
| extend t = datetime(2023-01-01) + 1d*x
| extend y = rand() * 10
// generate a sample series with outliers at first day of each month
| extend y = iff(monthofyear(t) != monthofyear(prev(t)), y+20, y)
| summarize t = make_list(t), series = make_list(y)
| extend outliers=series_outliers(series)
| extend pos_anomalies = array_iff(series_greater_equals(outliers, 1.5), 1, 0)
| render anomalychart with(xcolumn=t, ycolumns=series, anomalycolumns=pos_anomalies)
```

SQL to Kusto query translation - Azure Data Explorer & Real-Time Analytics | Microsoft Learn

Eventstream and Data Activator

Eventstream

The brand-new event stream service, leverages the ability to get data from several sources of streaming data and save it to a wide variety of destinations, including OneLake, KQL databases and Azure services.

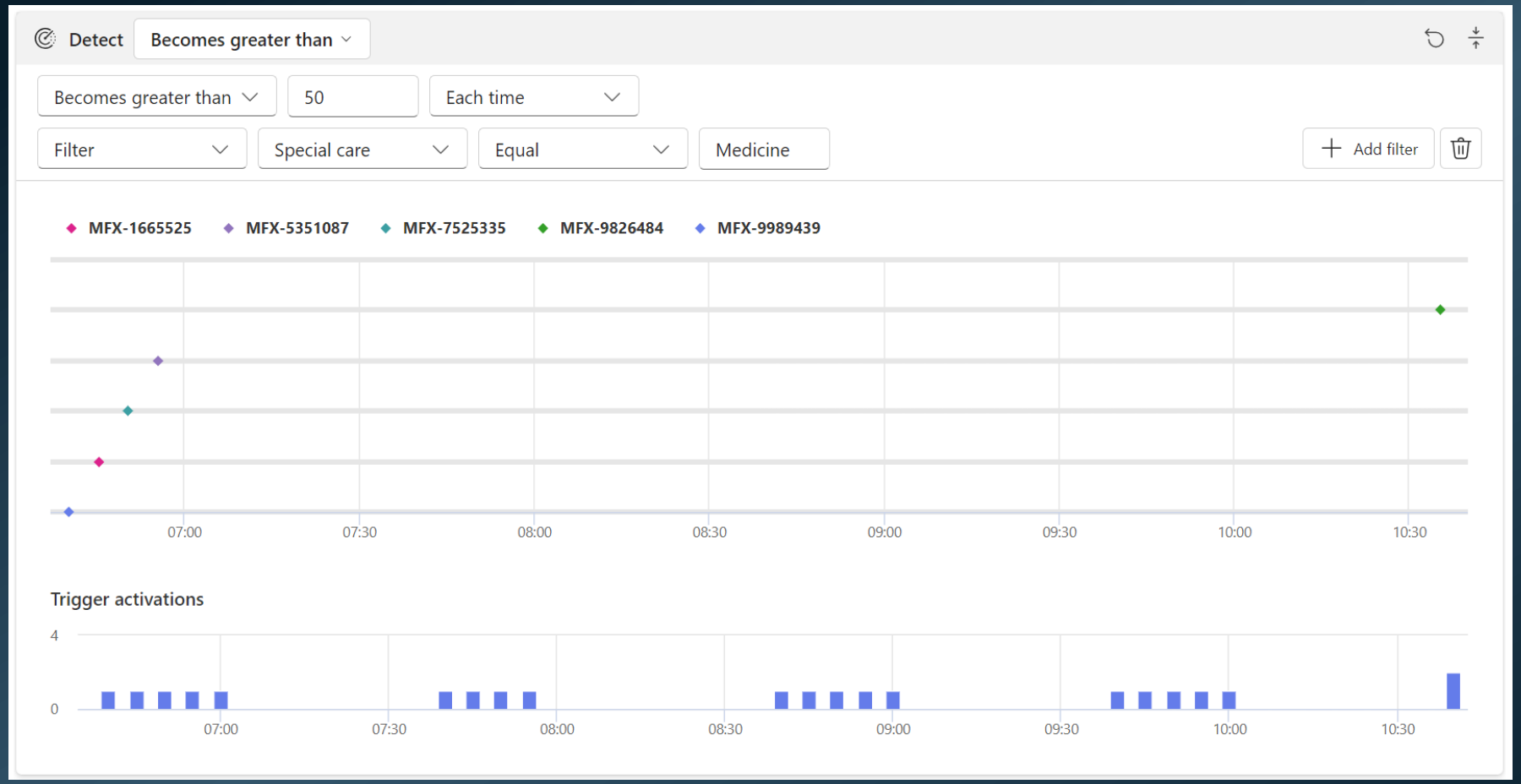


Data Activator

Activeli listens to your data from either the Eventstream service or a Power BI dataset.
Can react to values outside of defined boundaries and, for now, send an e-mail for a Teams message.

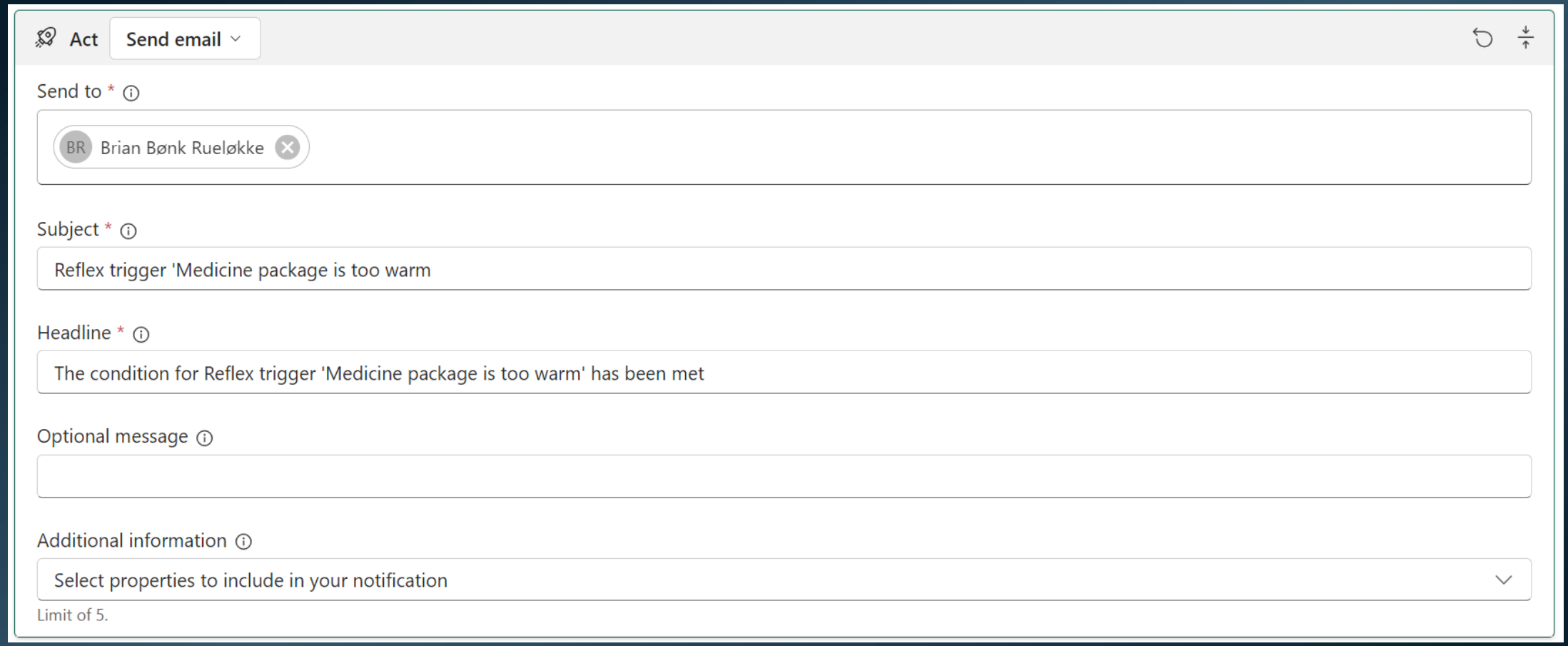
Eventstream and Data Activator

Data Activator



Eventstream and Data Activator

Data Activator



Analysis and reporting

Real-Time Analytics

Dashboards in RTA - planned - to come...

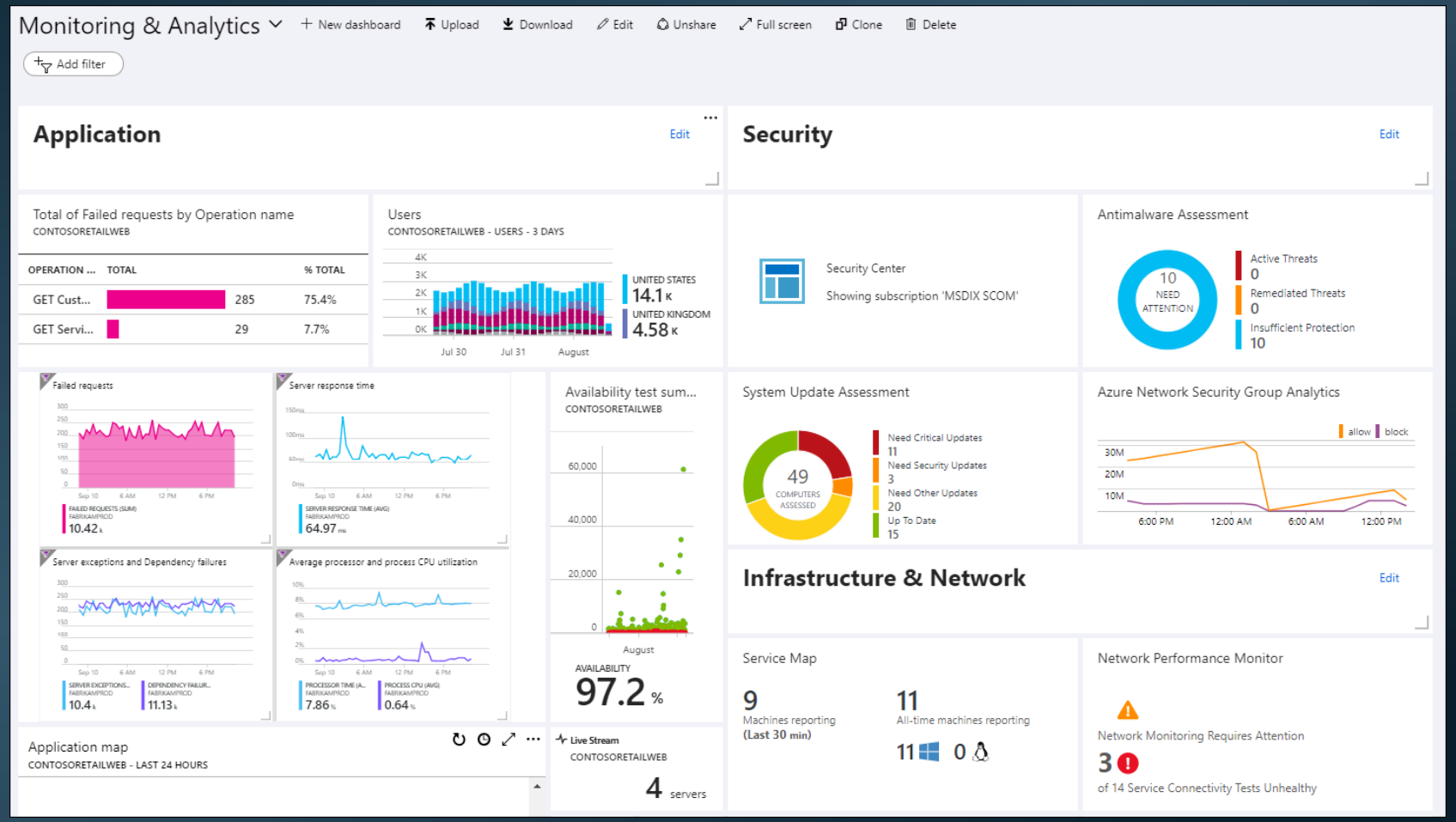


Image from James Westall

Thank you

Connect with me at:

https://linkedin.com/in/brianbonk
https://brianbonk.dk
https://github.com/brianbonk

Connect

Feedback