Microsoft

DATA TOBOGGAN

What's New and Cool in Azure Synapse Analytics

Pawel Potasinski
Senior Program Manager

@pawelpotasinski
/in/pawelpotasinski

Microsoft Azure

Stay up to date with Azure Synapse Analytics:

aka.ms/SynapseMonthlyUpdate

Synapse + Power BI

- Data Integration
- Data Engineering
- Data Warehousing
- Data Science
- Observational Analytics
- Business Intelligence
- Governance

Data
Integration
Data
Engineering
Data
Warehousing
Data
Science
Observational
Analytics
Business
Intelligence

Data
Integration

Data Integration

April-June 2022

- **Azure Synapse Link for SQL [Public Preview]**
- Dataverse connector added to Synapse Data Flows
- Synapse Pipelines Web activity response timeout improvement
- Export pipeline monitoring as a CSV
- Incremental data loading made easy for Synapse and Azure Database for PostgreSQL and MySQL
- Assert Error Handling
- Mapping data flows projection editing
- **Fuzzy Join option in Join Transformation**
- Map Data [Generally Available]
- **Rerun pipeline with new parameters**
- User Defined Functions [Generally Available]

Public Preview

Synapse Link for Microsoft SQL

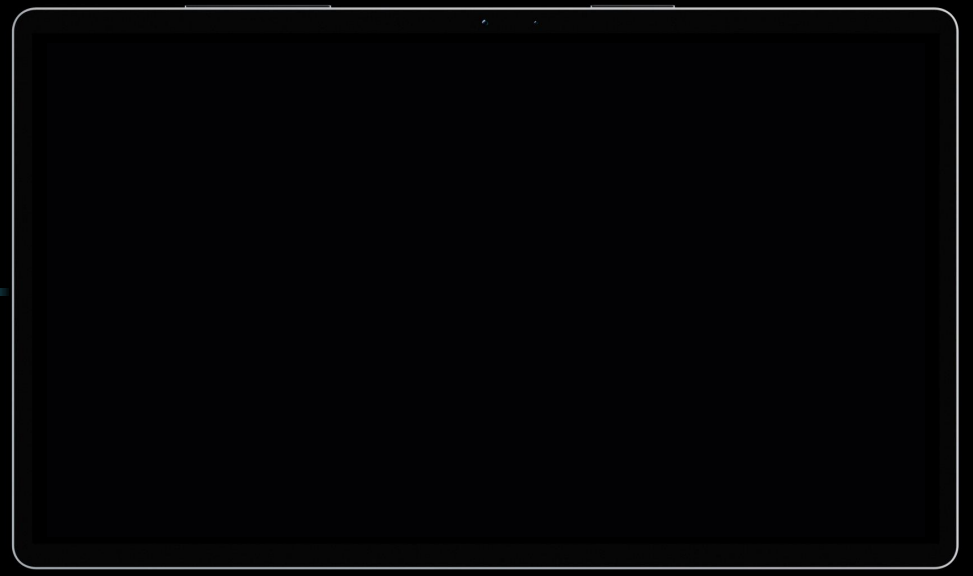
Near real-time operational analytics in Synapse

No data pipelines required

Hybrid integration for SQL Server running on-premise or in other clouds

SQL Server 2022
Public Preview Q2 2022

Azure SQL Database
Public Preview Q2 2022



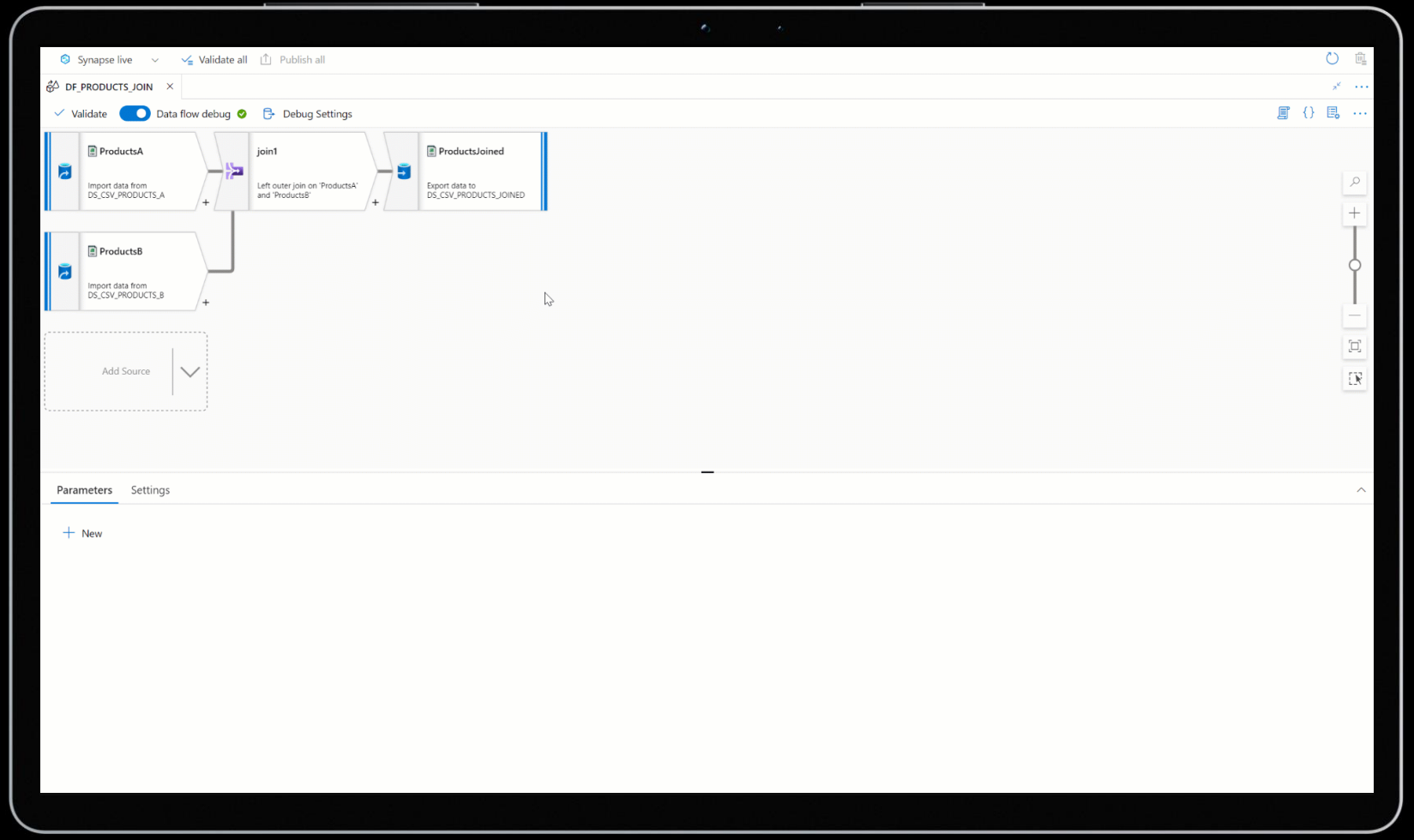
Fuzzy Join

No-code fuzzy logic for joins

Configurable similarity threshold

Similarity score column added

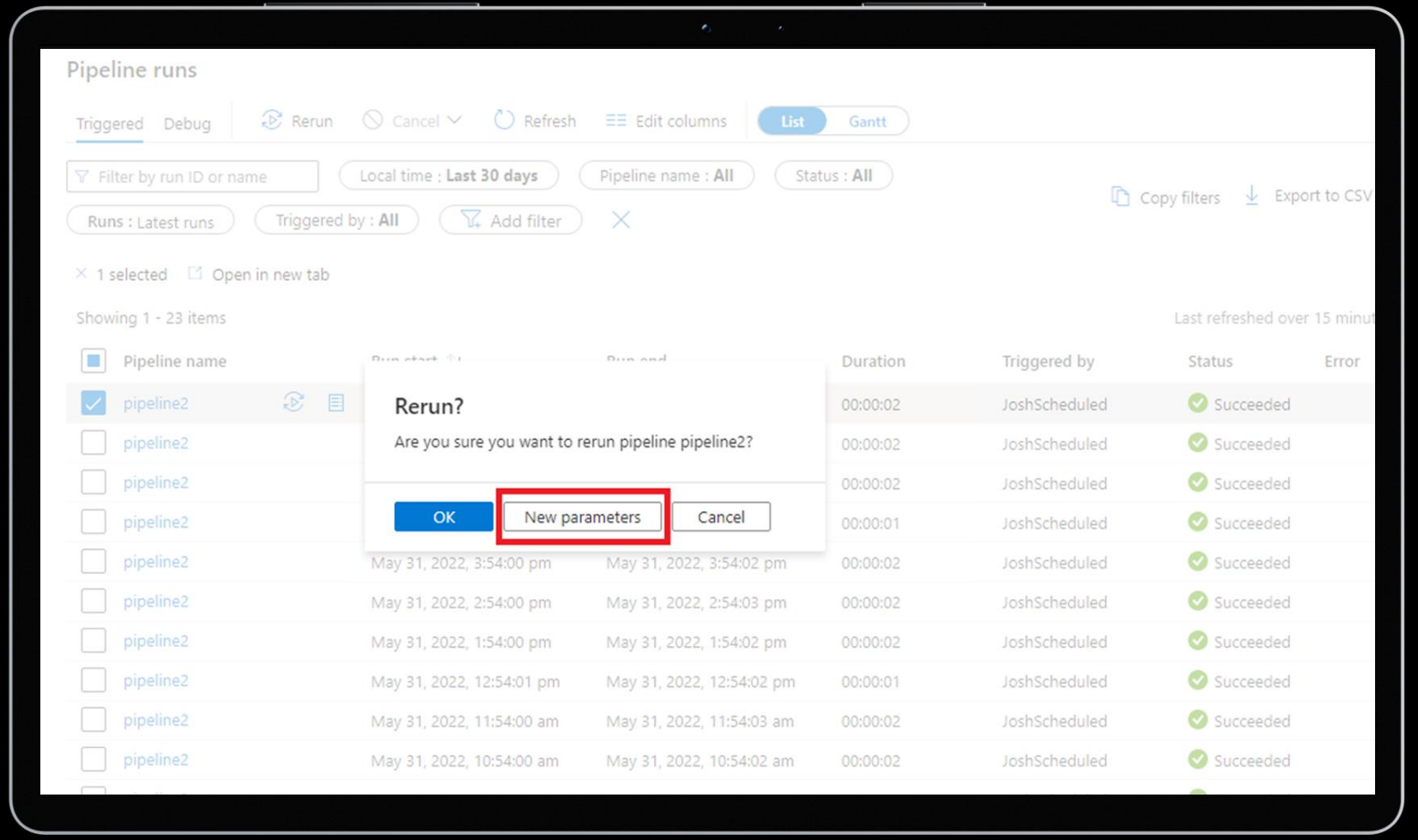
Combine text parts



Rerun pipeline with new parameters

You can change parameter values when re-running a pipeline from the Monitoring page

Considered a new pipeline run



Pre-Order

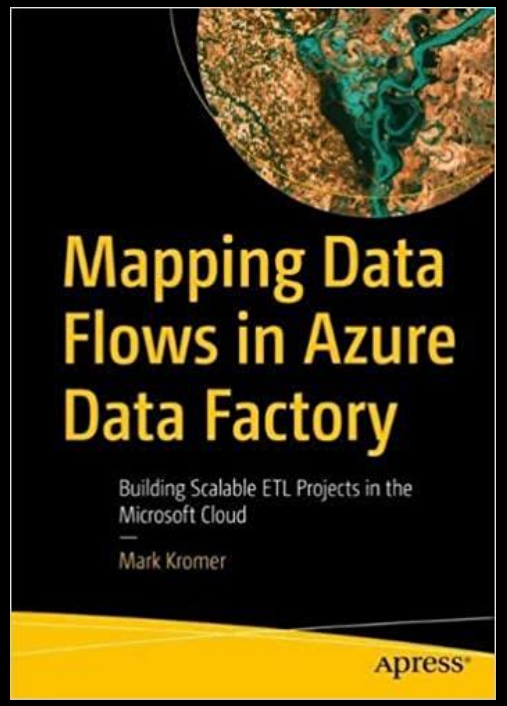
New Book Alert!

Build scalable ETL jobs

Apply best practices for complex data pipelines

Cloud-based ETL patterns

Repeatable ETL design patterns



[Mapping Data Flows in Azure Data Factory: Building Scalable ETL Projects in the Microsoft Cloud by Mark Kromer at Amazon.com]

Data Engineering

Data Engineering

April-June 2022

- **Apache Spark 3.2 [Public Preview]**
- Parameterization for Spark job definition
- Notebook snapshot
- Azure Synapse Dedicated SQL Pool Connector for Apache Spark Now Available in Python
- **Manage Azure Synapse Apache Spark configuration**
- **Reference unpublished notebooks**

Public Preview

Spark 3.2

Enables developers can leverage the latest innovations in the Spark ecosystem

Pandas (Koalas) integration

A highly popular and flexible library with broad industry adoption

Adaptive Query Execution (AQE) enabled by default

Significant improvements in query performance out-of-the-box

Small Query execution improvements

Small queries run faster due to reduced initialization overhead

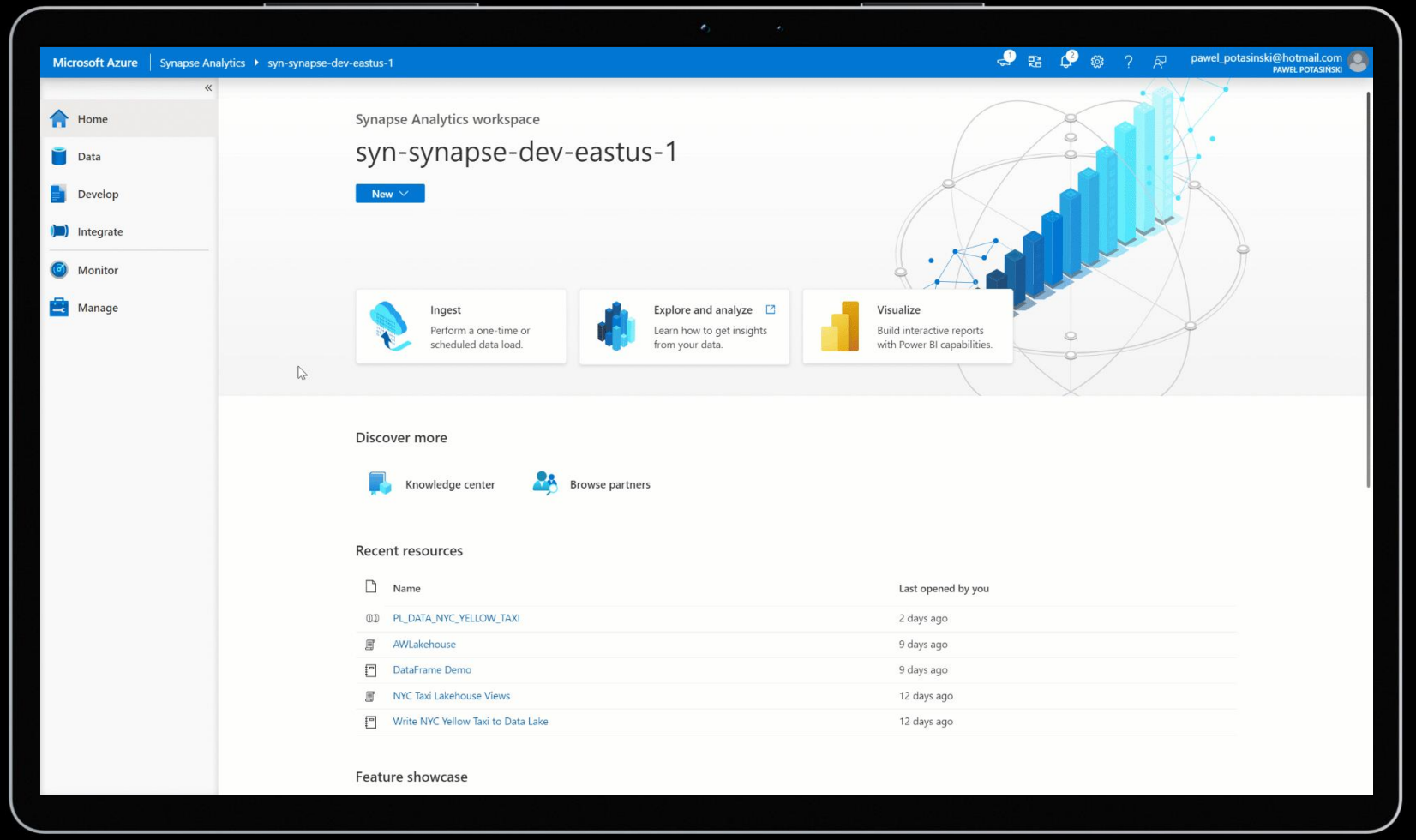
UPCOMING EVENT! Register today for our **Ask the Experts (ATE)** session with **Simon Whiteley on Spark 3.2** in Azure Synapse scheduled for **July 13 (10AM PDT)**. Learn more: https://aka.ms/SynapseInfluencers

Spark Configuration

Spark configuration artifact

Auto-suggestions and built-in validation rules

Can be associated with a Spark pool or a session, re-used and exported

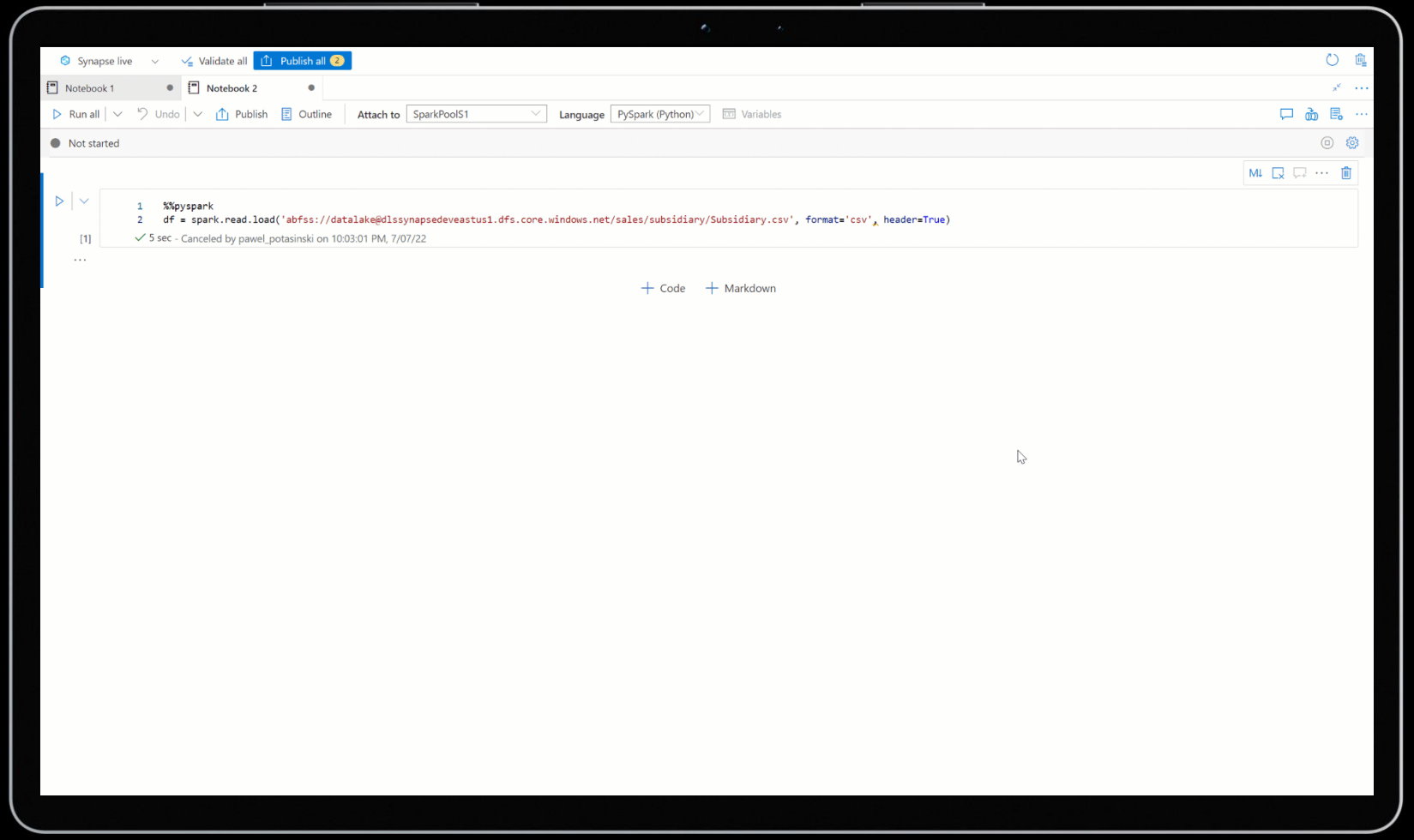


Reference unpublished notebooks

%run magic command can work with unpublished notebook

References notebook content in web cache

Great for local debugging



Data Warehousing

Data Warehousing

April-June 2022

- **Cross-subscription restore for Azure Synapse SQL [Generally Available]**
- Recover SQL pool from dropped server or workspace
- Revamped exploration experience
- **Clone lake database**
- **Use wildcards to specify custom folder hierarchies**
- Data Warehouse Migration guide for Dedicated SQL Pools in Azure Synapse Analytics
- **Specify character column lengths… Not anymore!**
- Result set size limit increase

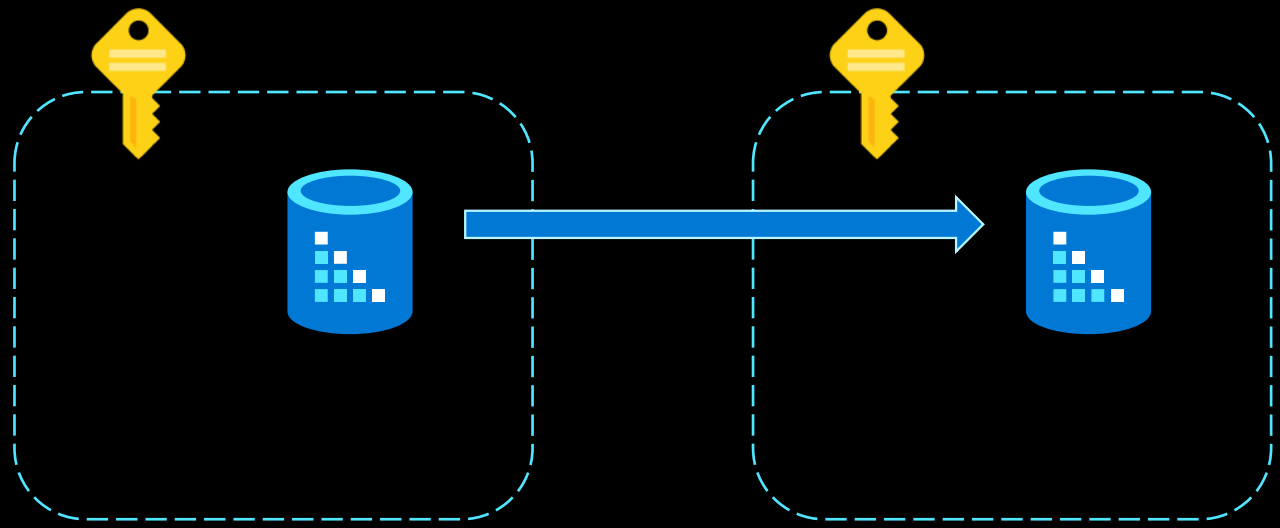
Public Preview

Cross subscription restore

Restore data warehouse across subscription boundary in one command without workarounds

Unlock Dev/Test scenarios and simplify billing at the subscription level

PowerShell and API support available



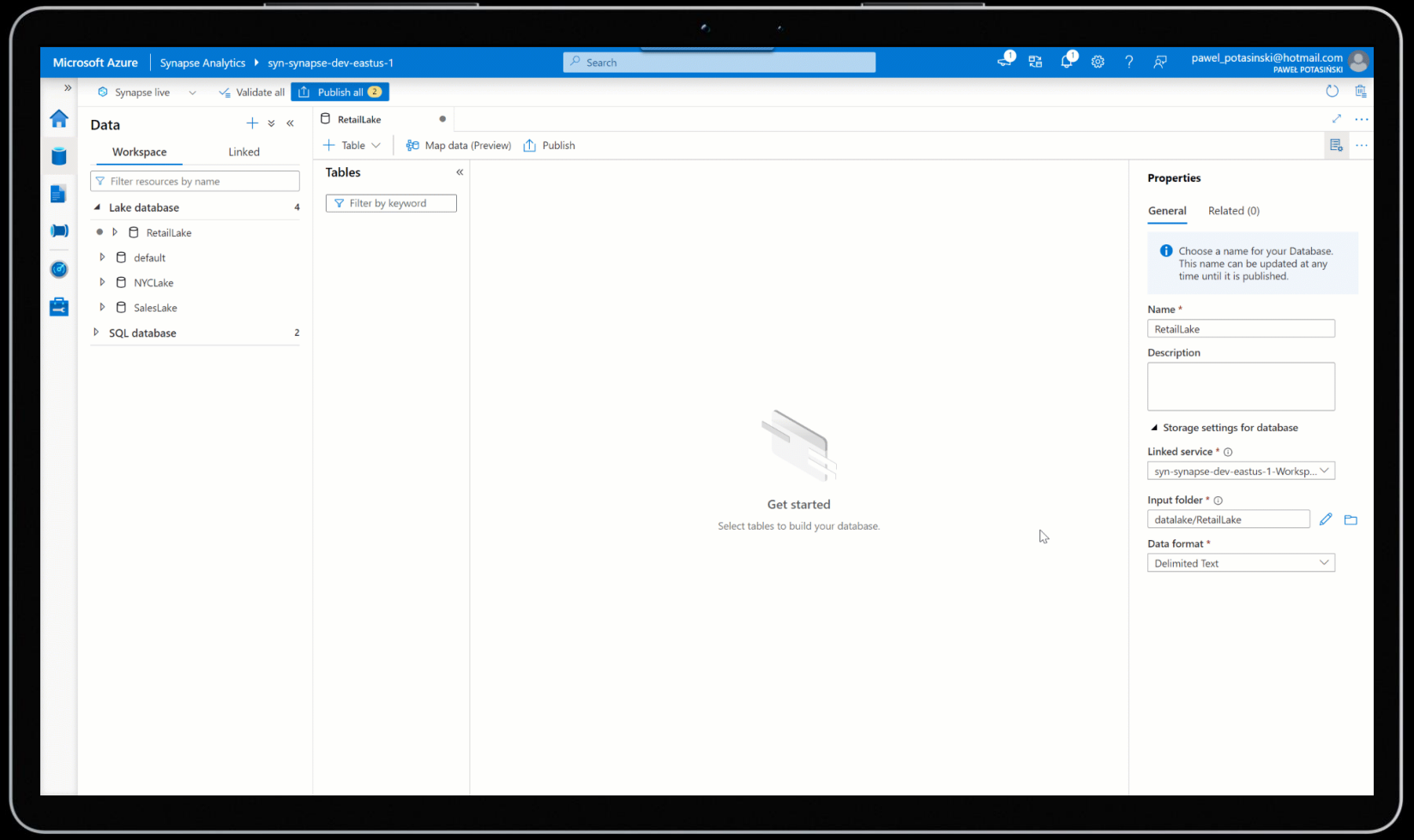
Generally Available

Database Templates & Database Designer

Built-in database templates

Low code database designer

Wildcards for custom folder hierarchies

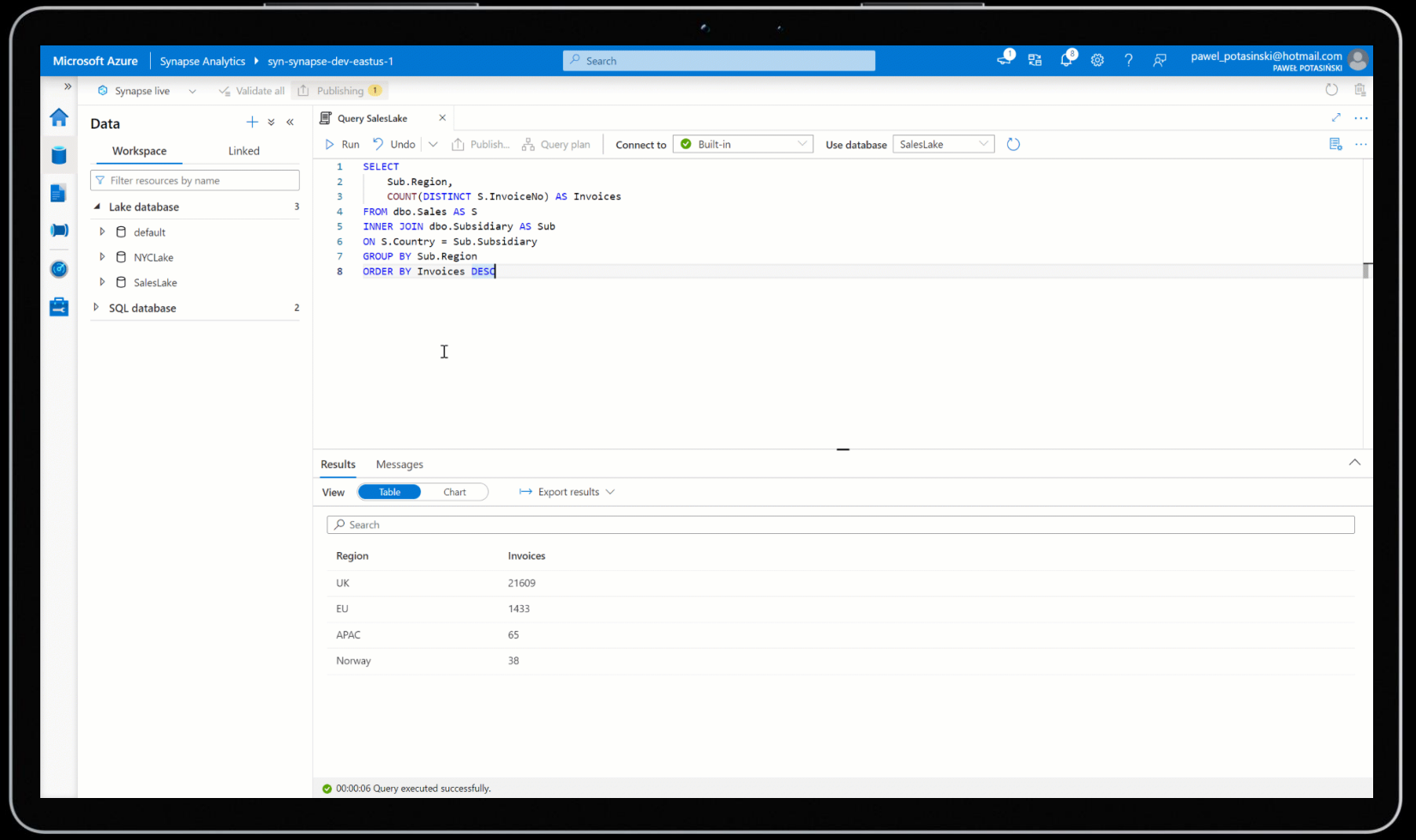


Clone lake database

Quick cloning of lake databases and tables using database designer UI

Editable metadata

No-code deployment



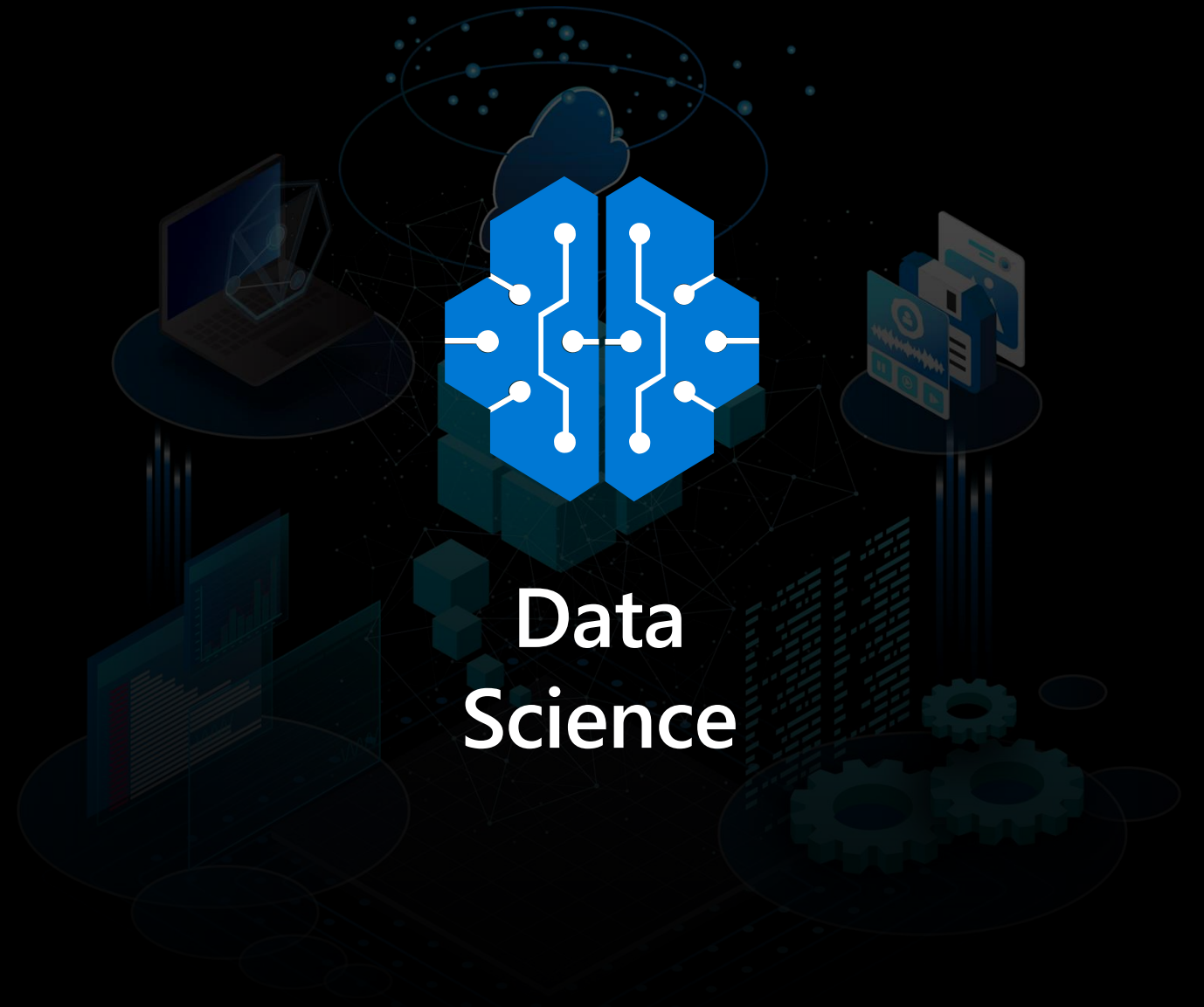
Character column lengths not required

No explicit schema definitione required for performance

Works for OPENROWSET and external tables

Triggers whenever you omit schema or specify character column longer than 100 bytes

```
SELECT *
FROM
OPENROWSET (
  BULK 'https://pandemicdatalake.blob.core.windows.net/public/curated/covid-19/ecdc_cases/latest/ecdc_cases.csv',
     FORMAT = 'CSV',
     PARSER_VERSION = '2.0',
     HEADER_ROW = TRUE
) AS [r]
WITH (
     countries_and_territories VARCHAR(50)
) AS [r]
```


Data
Science

Data Science

April-June 2022

- Distributed Deep Neural Network Training with Horovod and Petastorm [Public Preview]

Generally Available

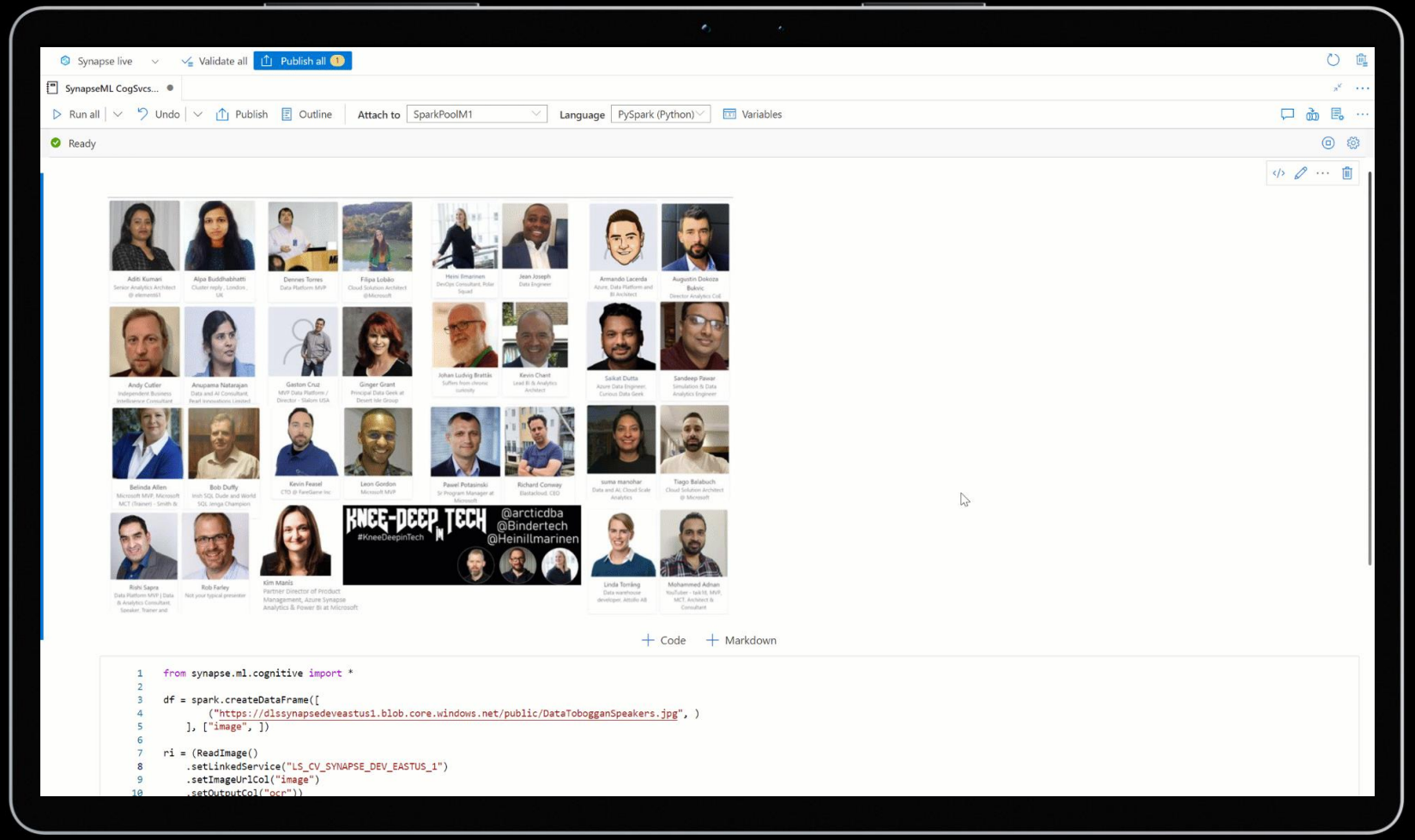
SynapseML

Open-Source Spark Library

Distributed ML algorithms for Apache Spark

Cognitive Services Integration for Spark

Language Binding Generators: Python, R, Java (, .NET)



https://aka.ms/spark

Observational Analytics

Observational Analytics

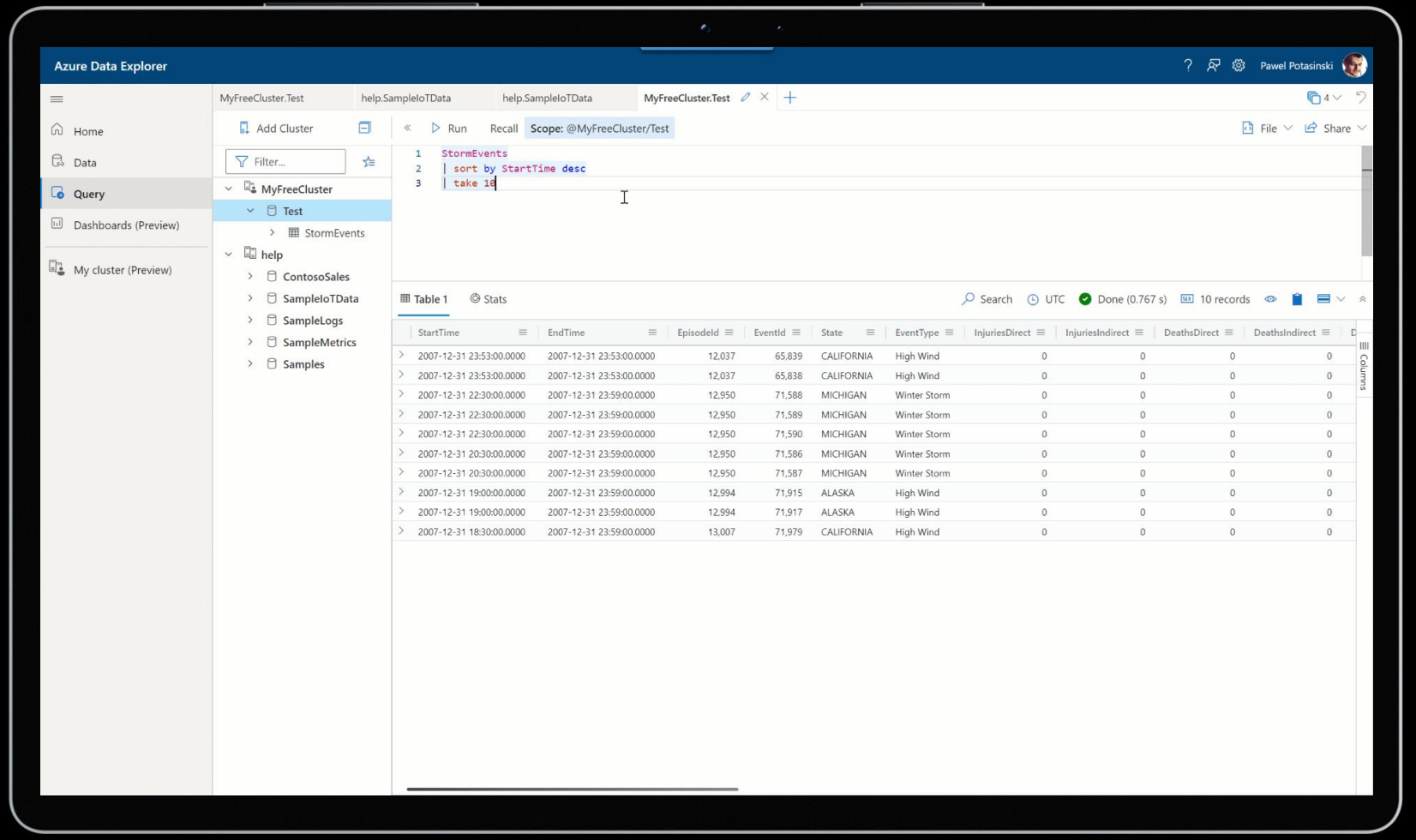
April-June 2022

- **Synapse Data Explorer live query in Excel**
- Use Managed Identities for External SQL Server Tables
- New KQL Learn module (2 out of 3) is live!
- **Azure Synapse Data Explorer connector for Microsoft Power Automate, Logic Apps, and Power Apps**
- Dynamic events routing from event hub to multiple databases
- Configure a database using a KQL inline script as part of JSON ARM deployment template
- **Web Explorer new homepage**
- **Web Explorer sample gallery**
- Web Explorer dashboards drill through capabilities
- Time Zone settings for Web Explorer

Live query in Excel

Web explorer can create an Excel workbook with connection to the cluster

KQL executed by PowerQuery

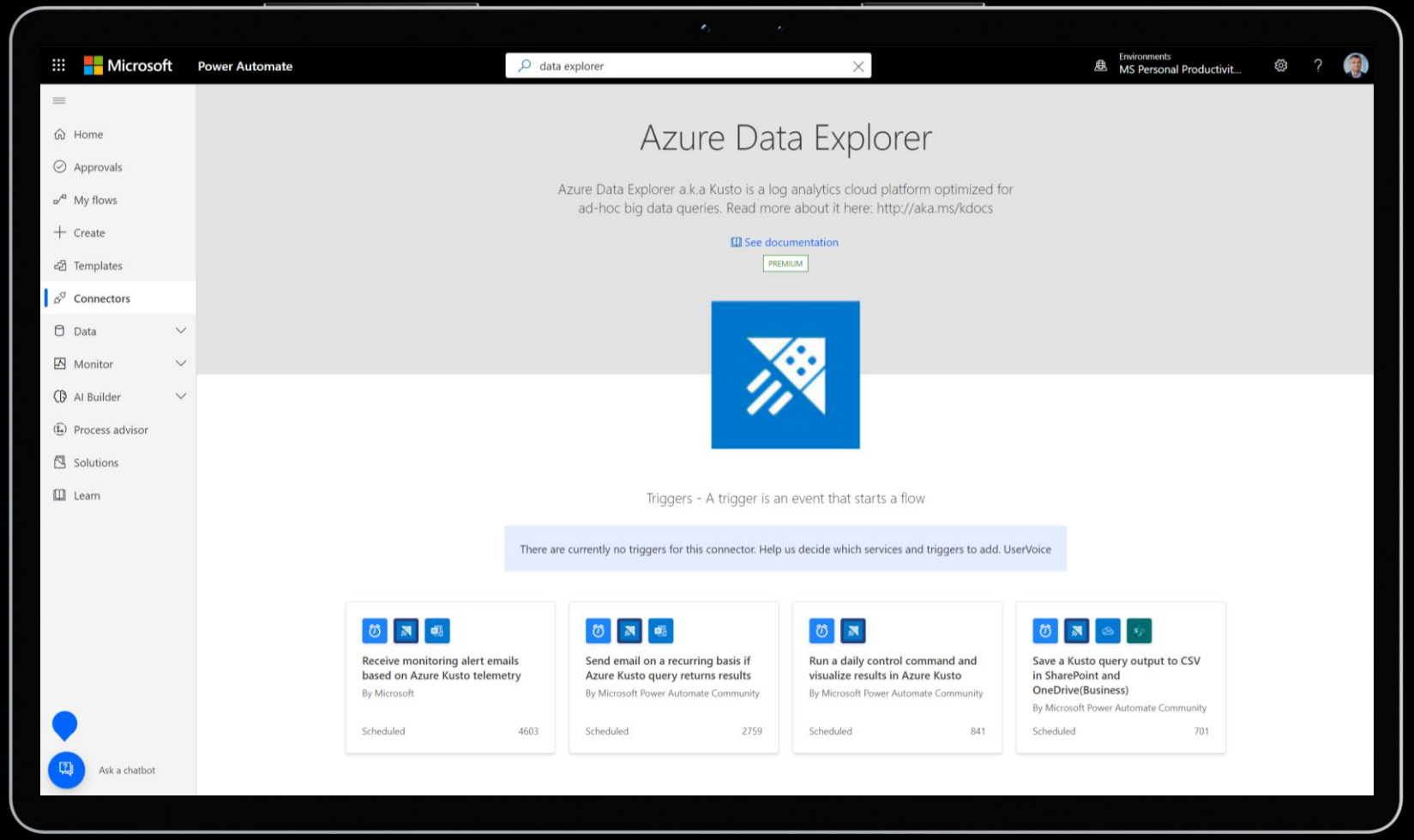


Connector for Power Automate, Logic Apps and Power Apps

Run KQL queries and render charts

Send notifications, alerts, reports

Schedule jobs using control commands on clusters

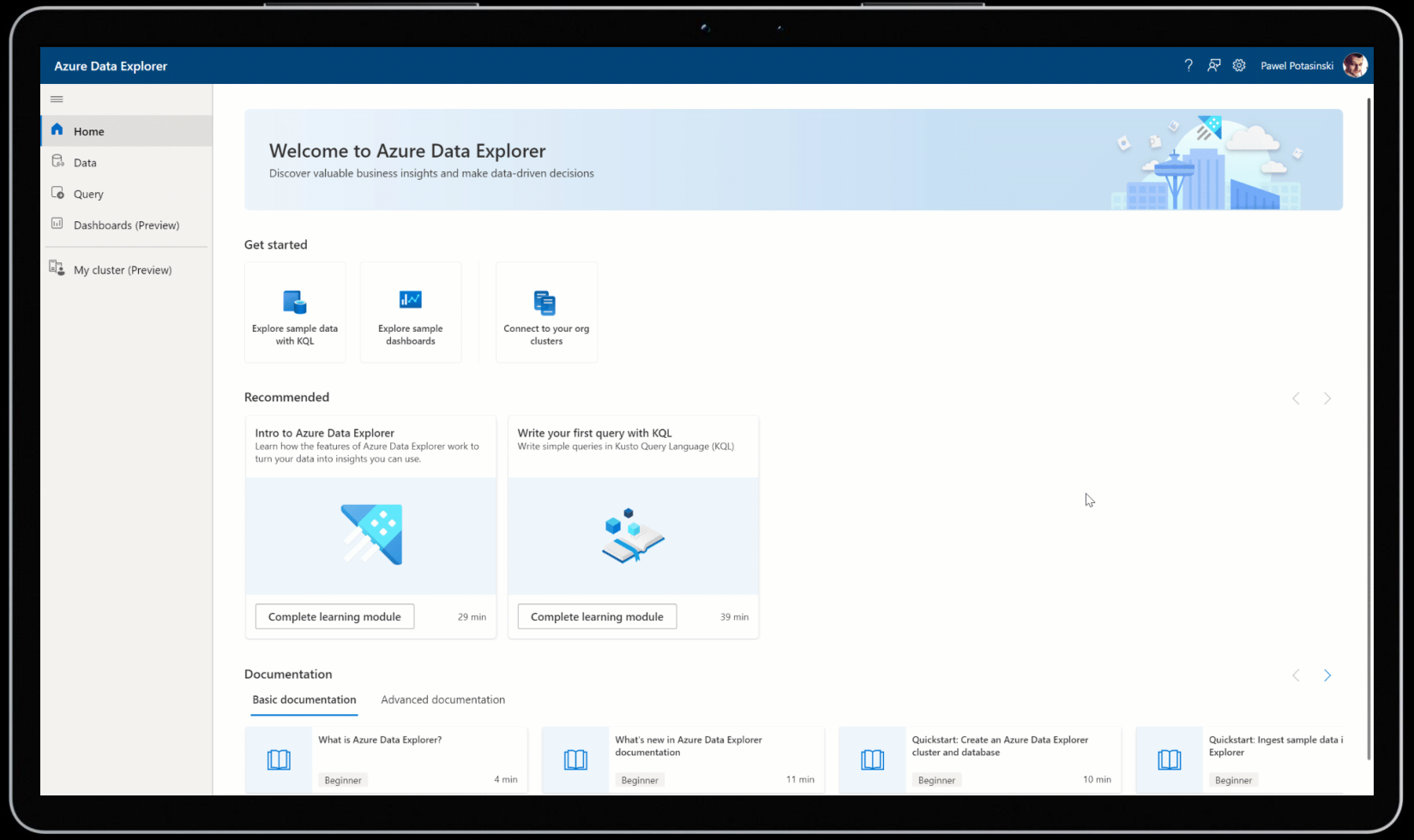


Web explorer new homepage and sample gallery

Get started with samples

Learn ADX and KQL using recommended MS Learn modules

Get quick access to documentation



Start & learn for free: [Azure Data Explorer in 60 minutes](#) | https://aka.ms/kustofree

Business Intelligence

Business Intelligence

April-June 2022

- **Public preview of datamart in Power BI**
- General availability of automatic aggregations in Power BI
- Power BI integration for PowerPoint (storytelling)
- Goals in Power BI is becoming ‘metrics’ and is moving to GA

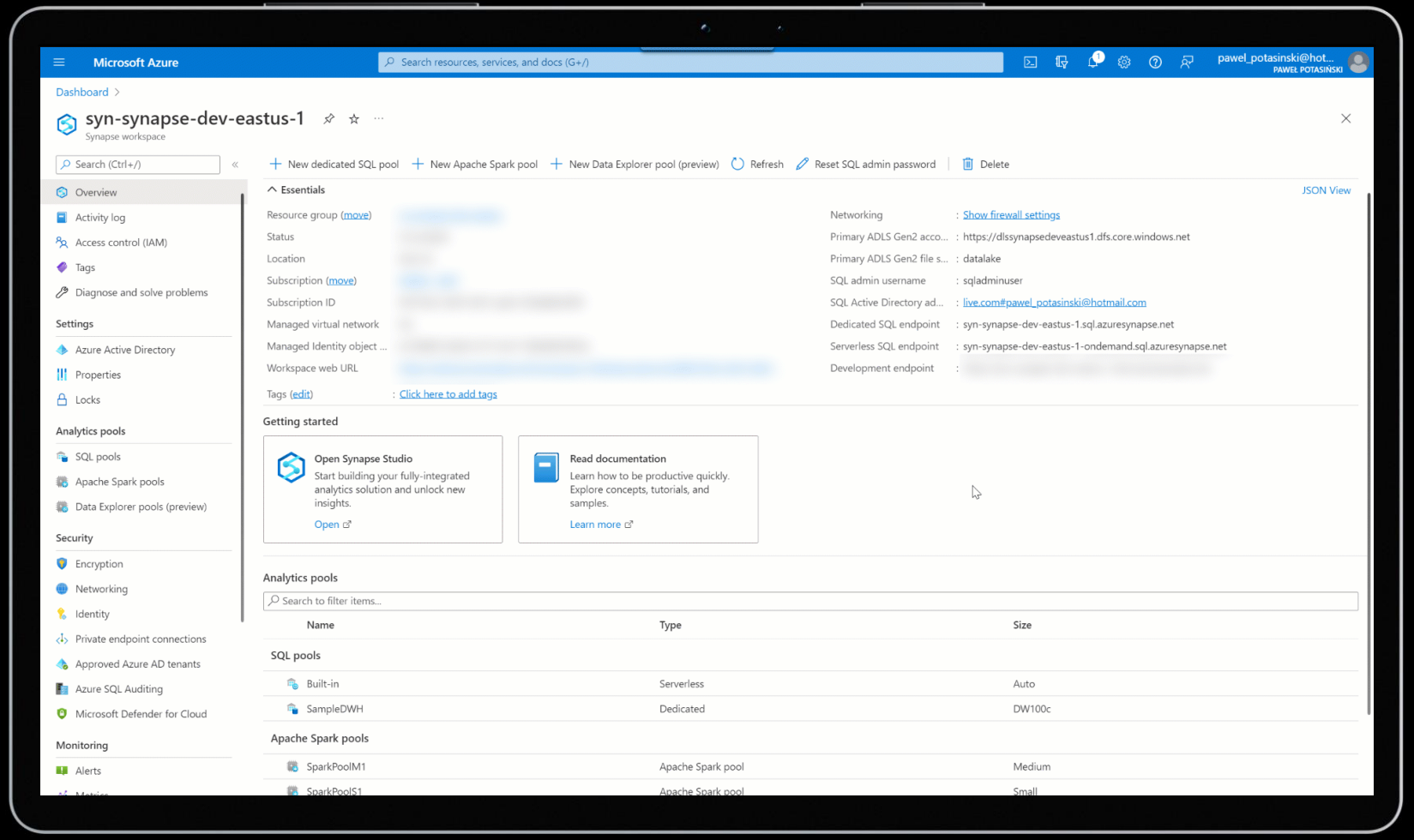
- Power BI April 2022 Feature Summary
- Power BI May 2022 Feature Summary
- Power BI June 2022 Feature Summary

Public Preview

Power BI datamart

Combination of Power BI dataflow + Azure SQL DB + Power BI dataset

Allows model development (relationships, RLS)



Synapse + Power BI

- Data Integration
- Data Engineering
- Data Warehousing
- Data Science
- Observational Analytics
- Business Intelligence
- Governance

Microsoft

Azure Synapse ♡ Community

Azure Synapse Community Resources

Community Call to Action

- ☑ Follow us on Twitter: @Azure_Synapse
- ☑ Read and comment on our blog: https://aka.ms/SynapseBlog
- ☑ Check monthly updates blog: https://aka.ms/SynapseMonthlyUpdate
- ☑ Subscribe to YouTube channel: https://aka.ms/SynapseYouTube
- ☑ Share ideas and vote for them: https://aka.ms/SynapseIdeas
- ☑ Join Azure Synapse Influencers program: https://aka.ms/SynapseInfluencers

Product

- Product page: https://aka.ms/Synapse
- Documentation: https://aka.ms/SynapseDocs
- Reference architectures: https://aka.ms/SynapseArchitectures
- Security whitepaper: http://aka.ms/SynapseSecurity

Technical Discussions and Q&A

- Microsoft Q&A: https://aka.ms/SynapseQuestions
- StackOverflow: https://aka.ms/SynapseStackOverflow

Quickstart

- Get started in 60 minutes: https://aka.ms/SynapseGetStarted
- Azure Synapse Analytics Toolkit: https://aka.ms/SynapseToolkit

Free Learning

- MS Learn – learning paths: https://aka.ms/SynapseLearningPaths
- Synapse Practitioner: https://aka.ms/SynapsePractitioner
- Microsoft Virtual Training Days: https://aka.ms/SynapseMVTD
- 30-Day Cloud Skills Challenge: https://aka.ms/SynapseSkillsChallenge

Samples & Accelerators

- Samples on GitHub: https://aka.ms/SynapseSamples
- Accelerator – End-to-End Analytics: https://aka.ms/azsynapsee2e-git

Microsoft

DATA TOBOGGAN

What's New and Cool in Azure Synapse Analytics

Pawel Potasinski
Senior Program Manager

@pawelpotasinski
/in/pawelpotasinski

Get this slide deck from:
https://github.com/pawelpo/presentations