



Create an Azure Synapse Lake Database without writing code

Erwin de Kreuk

Principal Consultant – Lead Data & AI
InSpark



Iberian
Technology
Summit_



Erwin de Kreuk

Principal Consultant – Lead Data & AI
InSpark



@erwindekrek



linkedin.com/in/erwindekrek



erwindekrek.com github.com/edkreuk



Iberian
Technology
Summit_





International
Workplace
Group



Microsoft



PROXIMOX

Iberian
Technology
Summit_

Objectives of today

- Database Templates
- Design Lake Database
- Map data
- Findings & experiences

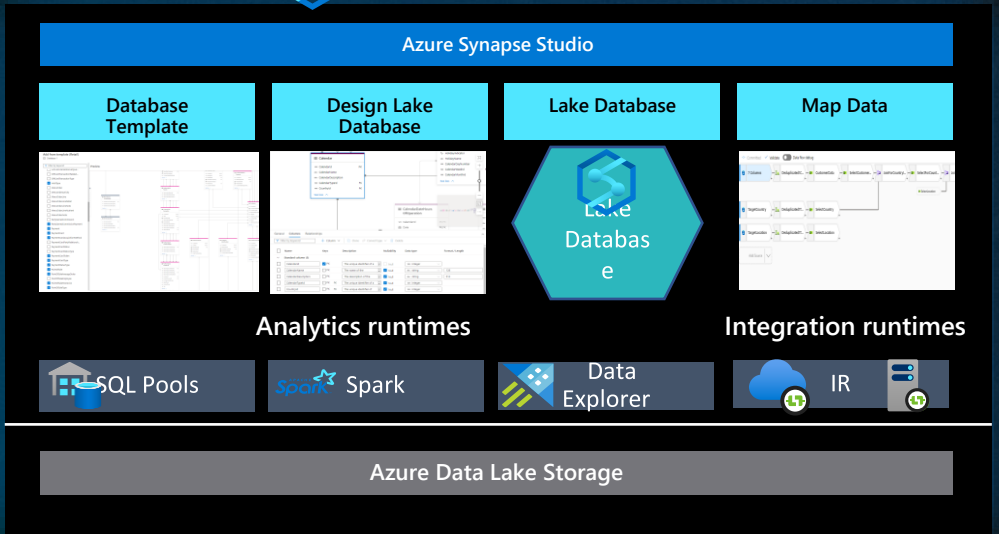




Resource Group Development



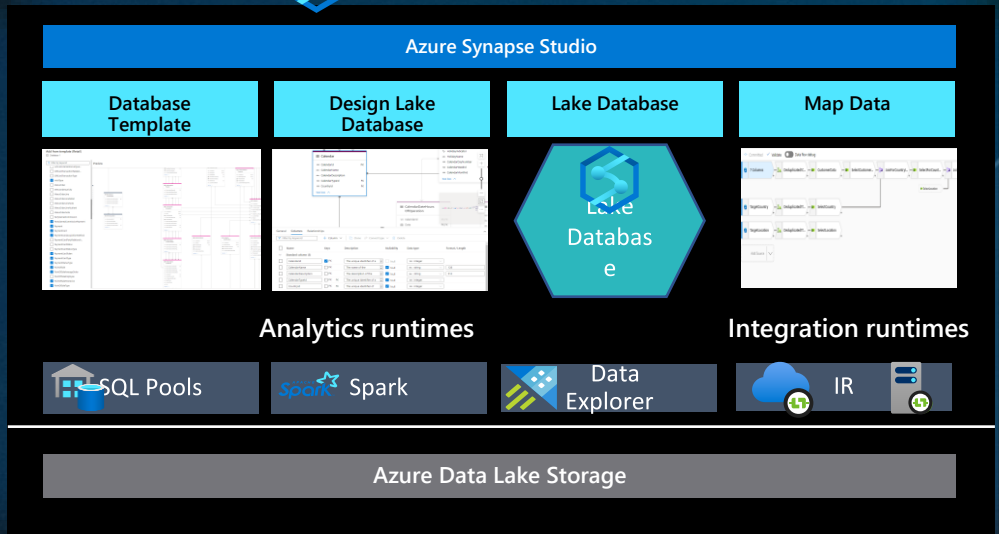
Azure Synapse Analytics



Resource Group Production



Azure Synapse Analytics



Iberian
Technology
Summit

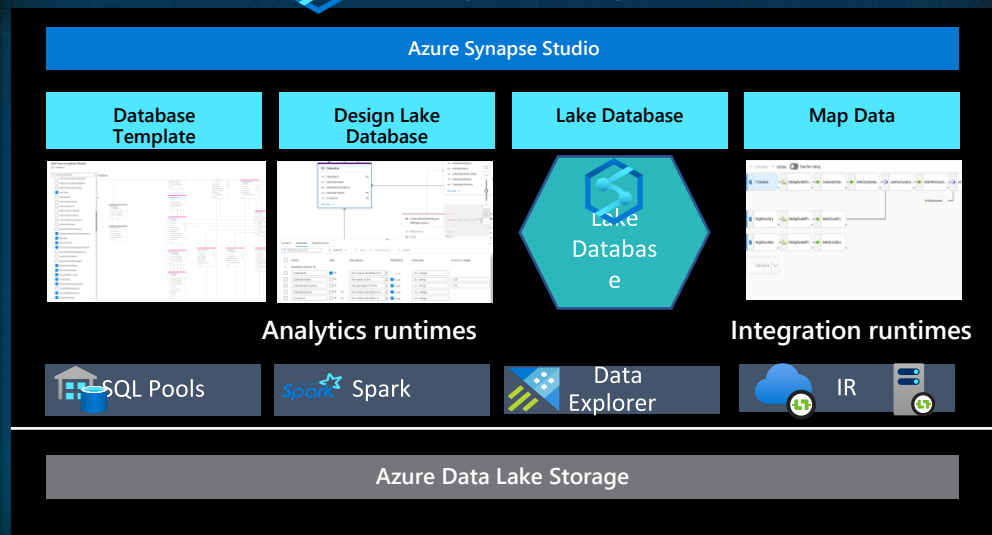




Resource Group Development



Azure Synapse Analytics



Iberian
Technology
Summit

Why

- Accelerate time to insights based on a standardized business area schemas for different industries
- Identify gaps and opportunities in your existing enterprise data model
- Consolidate data silos and query from a single pane of glass
- Create a well-formed data lake ready for the analytics at scale
- Enrich your data with Azure Cognitive Services and Azure Machine Learning
- Develop reports easily using Power BI

Iberian
Technology
Summit



Database Templates



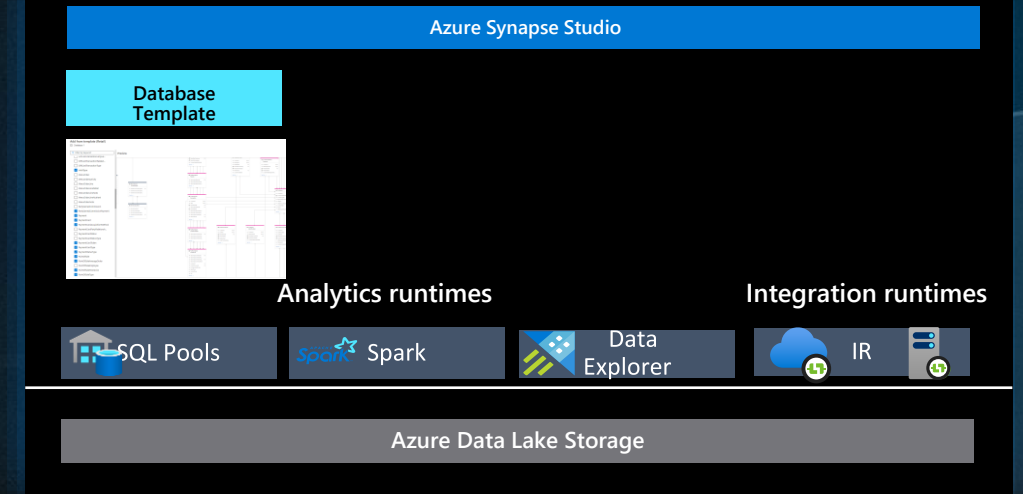
- Industry specific schema definitions



Resource Group Development























Azure Synapse Analytics



Database Templates

- Industry specific schema definitions
- 20 different Templates available

| | | | | |
|--|---|--|---|---|
|  Agriculture For companies engaged in growing crops, raising livestock and dairy production. |  Automotive For companies manufacturing automobiles, heavy vehicles, tires, and other automotive components. |  Banking For companies providing a wide range of banking and related financial services. |  Consumer Goods For manufacturers or producers of goods bought and used by consumers. |  Energy & Commodity Trading For traders of energy, commodities, or carbon credits. |
|  Freight & Logistics For companies providing freight and logistics services. |  Fund Management For companies managing investment funds on behalf of investors. |  Genomics For companies acquiring and analyzing genomic data about human beings or other species. |  Government For organizations controlling, regulating or providing services to a country/region, state or province, or community. |  Healthcare Insurance For organizations providing insurance to cover healthcare needs (sometimes know as Payors). |
|  Healthcare Providers For organizations providing healthcare services. |  Life Insurance & Annuities For companies who provide life insurance, sell annuities, or both. |  Manufacturing For companies engaged in discrete manufacturing of a wide range of products. |  Mining For companies extracting and processing minerals. |  Oil & Gas For companies involved in various phases of the Oil & Gas value chain. |
|  Pharmaceuticals For companies engaged in creating, manufacturing, and marketing pharmaceutical and bio-pharmaceutical products and medical devices. |  Property & Casualty Insurance For companies who provide insurance against risks to property and various forms of liability coverage. |  R&D and Clinical Trials For companies involved in research and development and clinical trials of pharmaceutical products or devices. |  Retail For sellers of consumer goods or services to customers through multiple channels. |  Utilities For gas, electric and water utilities and power generators and water desalinators. |

Database Templates

- Industry specific schema definitions
 - Enterprise Templates

Add from template (Retail)

DemoLake

Filter by keyword

Filter tables: Retail

Filter by keyword

Enterprise template

☒ Retail

Business area template

☐ Accounting & Financial Reporting

☐ Budget & Forecasting

☐ Business Metrics

☐ Channel

☐ Contract

☐ Customer

☐ Customer Account

☐ Customer Credit & Collections

☐ Customer Service & Contact Management

☐ Document Management

☐ Emissions

☐ Geography

☐ Human Resources & Payroll

☐ Inventory

☐ Loyalty

☐ Manufacturing (Continuous)

☐ Marketing & Advertising

Preview

BalanceSheet

| | | |
|-----|-----------------------------------|----|
| 123 | LegalEntityId | PK |
| 123 | FinancialStatementId | FK |
| | PeriodStartDate | FK |
| | PeriodEndDate | FK |
| e* | CurrentAssetsCashAmount | |
| e* | CurrentAssetsNotesReceivable | |
| e* | CurrentAssetsInventories | |
| e* | CurrentAssetsPrepaidExpenses | |
| e* | CurrentAssetsMarketableSecurities | |
| e* | CurrentAssetsAccountsReceivable | |

[See less](#)

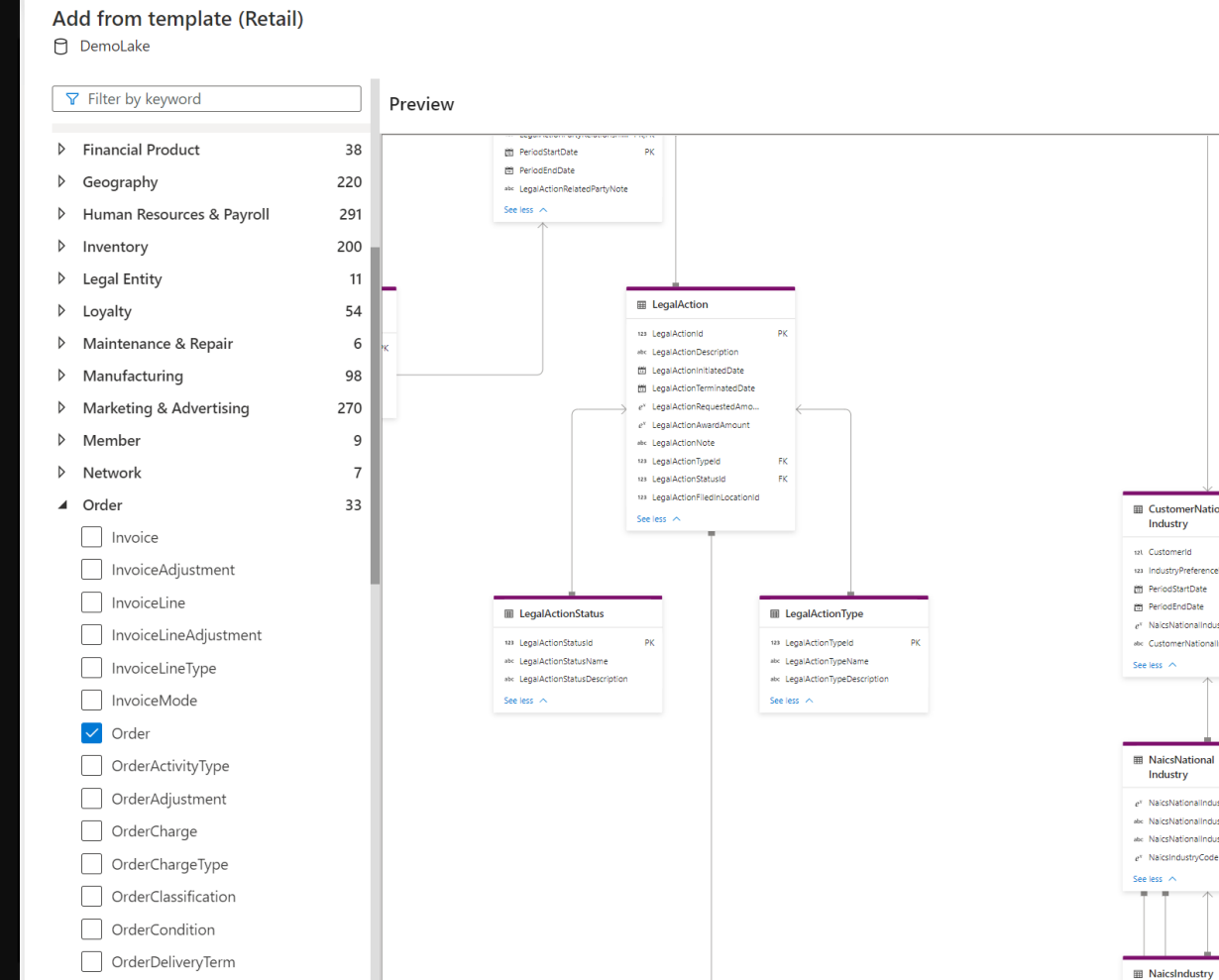
NotesToFinancial Statements

| | | |
|-----|----------------------------|-------|
| 123 | LegalEntityId | PK,FK |
| | PeriodStartDate | PK,FK |
| | PeriodEndDate | PK,FK |
| 123 | FinancialStatementNoteId | PK |
| abc | FinancialStatementNoteText | |
| 123 | AccountingFrameworkId | FK |

Database Templates

- Industry specific schema definitions
 - Enterprise Templates
 - Business Area Templates

Iberian
Technology
Summit



Design Lake Database



- Create database
 - Synapse Permissions
 - Synapse Administrator
 - Synapse Contributor
 - Synapse Artifact
 - Azure Permissions
 - Storage Blob Data Contributor
- Explore Database
 - Synapse Permissions
 - Synapse User

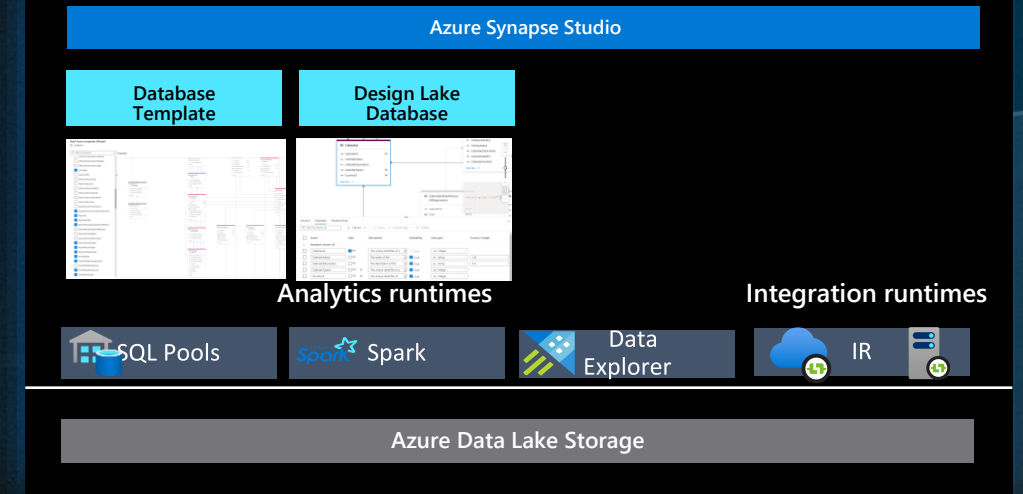
Iberian
Technology
Summit_



Resource Group Development



Azure Synapse Analytics



Design Lake Database

- Create database
 - Empty
 - Name
 - Description
 - Storage Settings
 - Input Folder
 - Data Format
 - Delimited text or Parquet

Iberian
Technology
Summit_

Properties

General Related (0)

Get started
les to build your database.

Choose a name for your Database.
This name can be updated at any time until it is published.

Name *
DEMO

Description

Storage settings for database

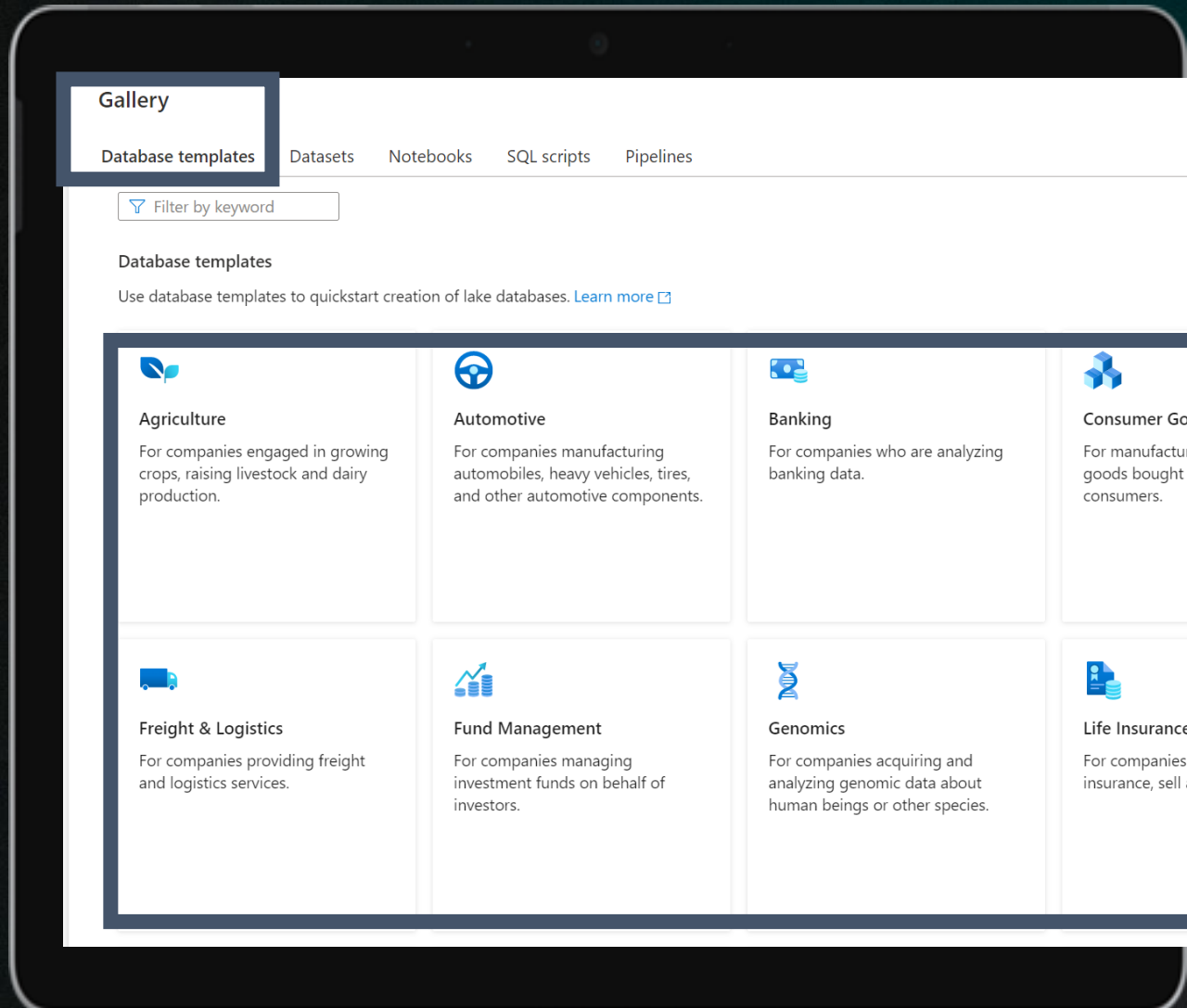
Linked service * ⓘ
demoeuewdvImSyn01-Workspace...

Input folder * ⓘ
temp/DEMO

Data format *
Delimited text

Design Lake Database

- Create database
 - From Database Template
 - Gallery
 - Template





Design Lake Database

- Create database
 - From Database Template
 - Name
 - Description
 - Template version
 - Storage Settings
 - Input Folder
 - Data Format
 - Delimited text or Parquet

Iberian
Technology
Summit

Transaction

- TransactionId PK
- TransactionInitiatedTime...
- TransactionCompletedTime...
- TransactionReceivedTime...
- TransactionEntryTimestamp
- TransactionConfirmationTi...
- CreditCheckTimestamp
- RequestedDeliveryDate
- CommittedDeliveryDate
- ActualDeliveryTimestamp

[See less](#)

Properties

General Related (0)

Choose a name for your Database. This name can be updated at any time until it is published.

Name *

Retail_vvj

Description

Lake database template
Retail 0.1.0

Storage settings for database

Linked service *

demoeuewvImSyn01-Workspace...

Input folder *

temp/Retail_vvj

Data format *

Delimited text

Lake Database

- Database Designer
 - Data Storage
 - CSV Format
 - Parquet
- Database Compute
 - Synapse SQL serverless SQL pool
 - Synapse Apache Spark
- Folder Structure
 - Files

Iberian
Technology
Summit_



Resource Group Development



Azure Synapse Analytics

| Name | Name |
|--|--|
| <input type="checkbox"/> ChannelType | <input type="checkbox"/> _SUCCESS |
| <input type="checkbox"/> Country | <input type="checkbox"/> part-00000-7e1d5778-8c97-4ed3-aaf4-f56a48e1f3fc-c000.snappy.parquet |
| <input type="checkbox"/> Customer | <input type="checkbox"/> part-00000-a4df07be-51b5-4004-93be-a75f8f6ec44b-c000.snappy.parquet |
| <input type="checkbox"/> IndividualCustomer | <input type="checkbox"/> part-00069-7e1d5778-8c97-4ed3-aaf4-f56a48e1f3fc-c000.snappy.parquet |
| <input type="checkbox"/> Item | <input type="checkbox"/> part-00069-a4df07be-51b5-4004-93be-a75f8f6ec44b-c000.snappy.parquet |
| <input type="checkbox"/> Location | <input type="checkbox"/> part-00107-7e1d5778-8c97-4ed3-aaf4-f56a48e1f3fc-c000.snappy.parquet |
| <input type="checkbox"/> Product | <input type="checkbox"/> part-00107-a4df07be-51b5-4004-93be-a75f8f6ec44b-c000.snappy.parquet |
| <input type="checkbox"/> ProductCategoryDim | <input type="checkbox"/> part-00128-7e1d5778-8c97-4ed3-aaf4-f56a48e1f3fc-c000.snappy.parquet |
| <input type="checkbox"/> Territory | <input type="checkbox"/> part-00128-a4df07be-51b5-4004-93be-a75f8f6ec44b-c000.snappy.parquet |
| <input type="checkbox"/> Transaction | <input type="checkbox"/> part-00140-7e1d5778-8c97-4ed3-aaf4-f56a48e1f3fc-c000.snappy.parquet |
| <input type="checkbox"/> TransactionLineItem | <input type="checkbox"/> part-00140-a4df07be-51b5-4004-93be-a75f8f6ec44b-c000.snappy.parquet |

A tablet with a black bezel is centered in the frame. The screen is black and displays the word "DEMO" in a bright yellow, sans-serif font. The background of the entire image is a dark teal color with faint, light-colored circuit board patterns.

DEMO

Iberian
Technology
Summit_

Map Data

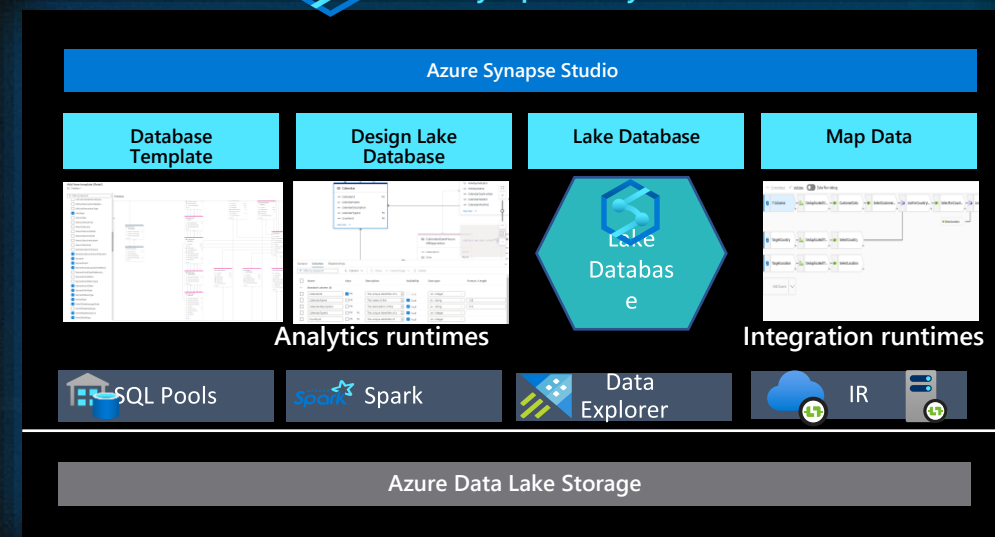
- Create ETL mapping



Resource Group Development

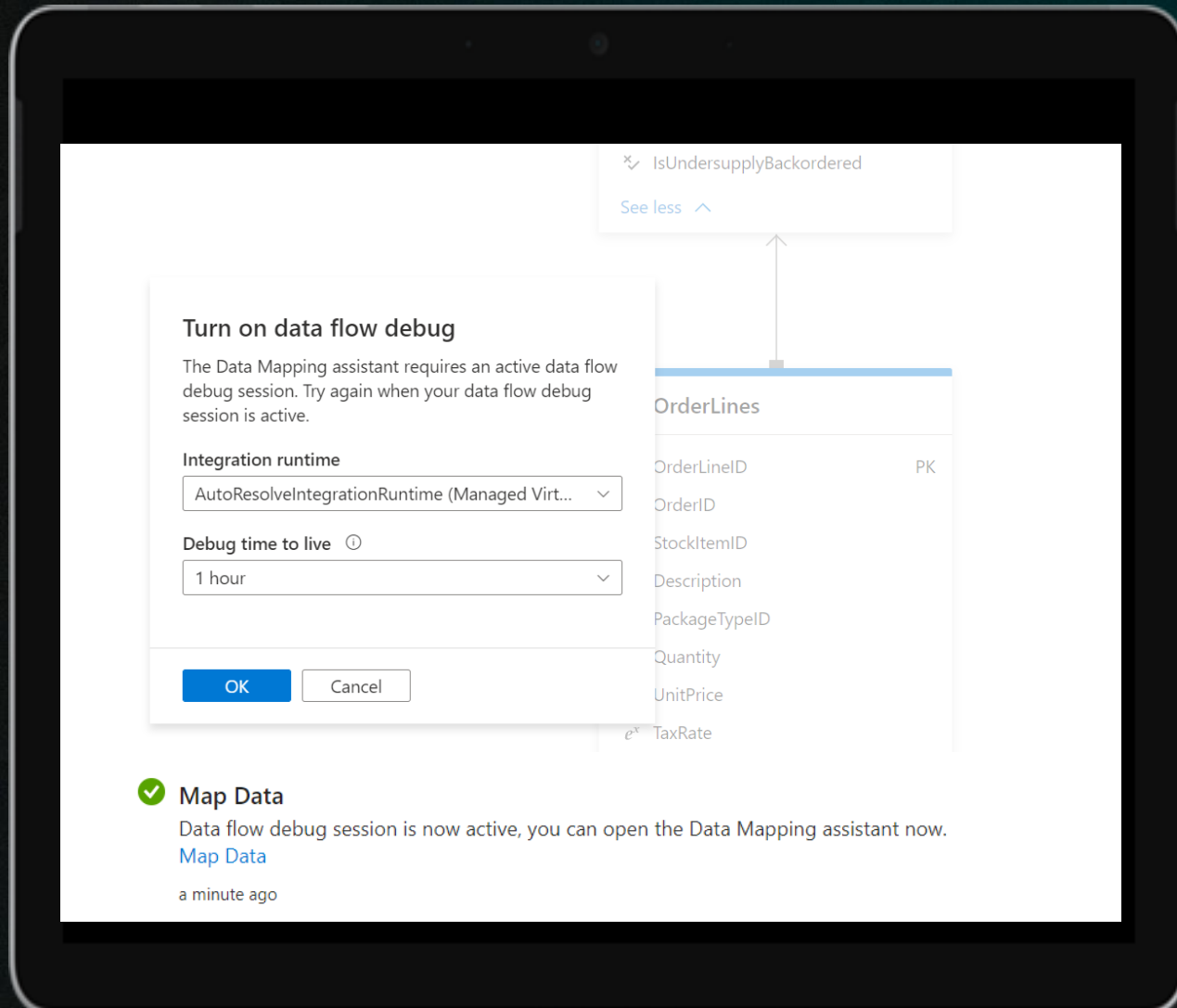


Azure Synapse Analytics



Map Data

- Create ETL mapping
- Data flow debug




Map Data

- Create ETL mapping
- Data flow debug
- Source Type
 - Lake Database

Iberian
Technology
Summit_

New data mapping


The data mapping assistant allows you to easily map data to any of your Lake Databases. [Learn more](#) 

Source type *

 Lake database 

Source database *

Select... 

 Filter by keyword

LakeDemo

DemoLake

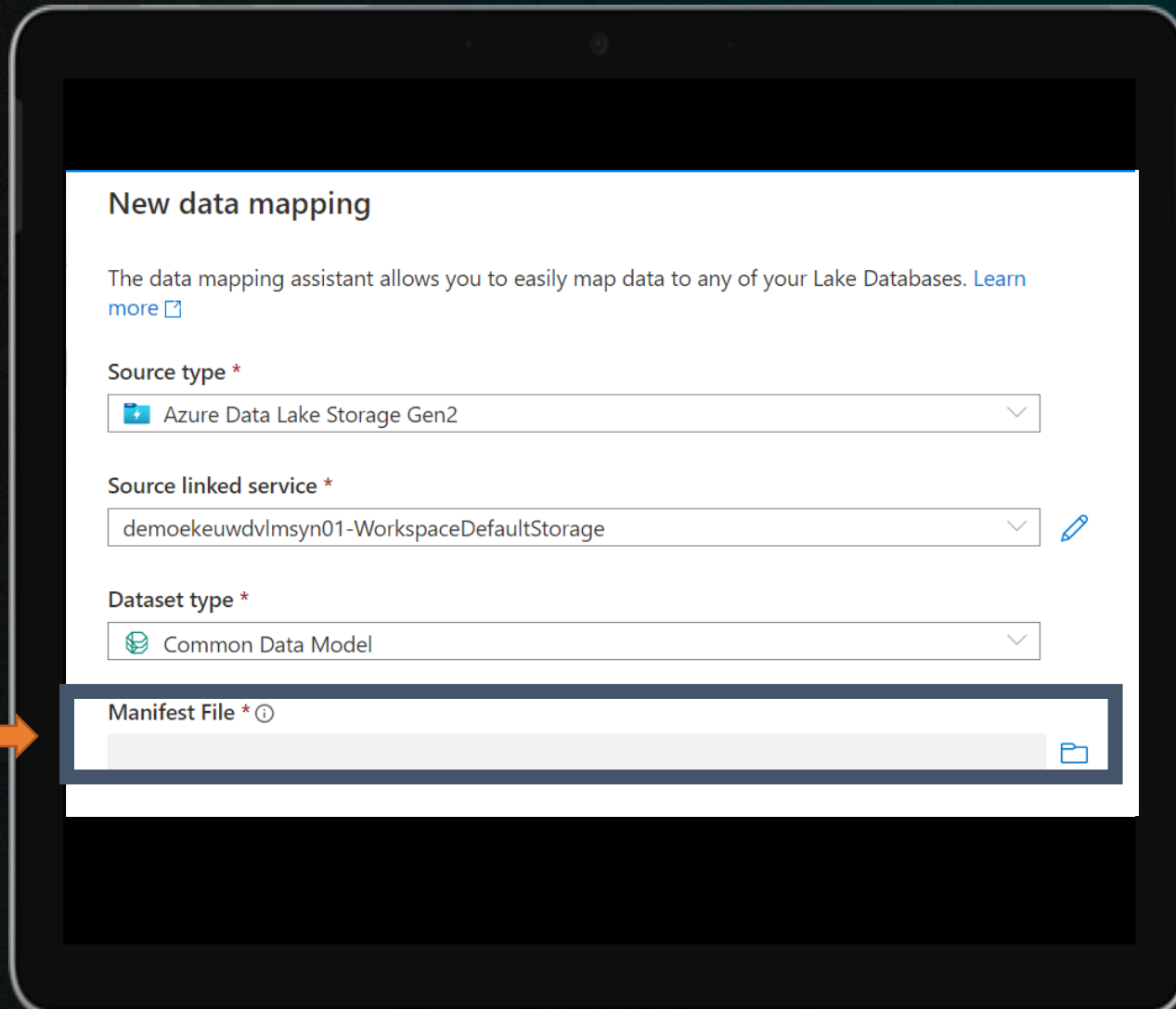
DEMO

Retail_vvj

Map Data

- Create ETL mapping
- Data flow debug
- Source Type
 - Lake Database
 - **Data Lake Storage**
 - **Common Data Model**

Iberian
Technology
Summit_



The screenshot shows a tablet displaying the 'New data mapping' interface. The interface includes a title 'New data mapping', a descriptive paragraph, and four dropdown menus: 'Source type' (set to 'Azure Data Lake Storage Gen2'), 'Source linked service' (set to 'demoeuewvmsyn01-WorkspaceDefaultStorage'), 'Dataset type' (set to 'Common Data Model'), and 'Manifest File' (which is highlighted with a blue border and an orange arrow pointing to it from the left). The 'Manifest File' field has a folder icon on the right. Below the 'Manifest File' field is a light gray input area.

New data mapping

The data mapping assistant allows you to easily map data to any of your Lake Databases. [Learn more](#)

Source type *

Azure Data Lake Storage Gen2

Source linked service *

demoeuewvmsyn01-WorkspaceDefaultStorage

Dataset type *

Common Data Model

Manifest File * ⓘ

Map Data

- Create ETL mapping
- Data flow debug
- Source Type
 - Lake Database
 - **Data Lake Storage**
 - Common Data Model
 - **Delimited text**

Iberian
Technology
Summit_

The screenshot shows the 'New data mapping' configuration window. It includes a title bar, a description, and several input fields. Five orange arrows point from the left towards the following fields: 'Source type', 'Source linked service', 'Source type' (DelimitedText), 'First row as header', and 'Column delimiter'.

New data mapping

The data mapping assistant allows you to easily map data to any of your Lake Databases.
[Learn more](#)

Source type *

Azure Data Lake Storage Gen2

Source linked service *

demoeukeuwdvlsyn01-WorkspaceDefaultStorage

Source type *

DelimitedText

Folder path * ⓘ

☒ First row as header

Column delimiter ⓘ

Comma (,)

☐ Edit

Map Data

- Create ETL mapping
- Data flow debug
- Source Type
 - Lake Database
 - **Data Lake Storage**
 - Common Data Model
 - Delimited text
 - Parquet

Iberian
Technology
Summit_

New data mapping

The data mapping assistant allows you to easily map data to any of your Lake Databases. [Learn more](#)

Source type *

Azure Data Lake Storage Gen2

Source linked service *

demoeuewdvImSyn01-WorkspaceDefaultStorage

Dataset type *

Parquet

Folder path * ⓘ

Compression ⓘ

None

None

GZip

Snappy

A tablet with a black bezel is centered in the frame. The screen is black and displays the word "DEMO" in a bright yellow, sans-serif font. The background of the entire image is a dark teal color with faint, light-colored circuit board patterns.

DEMO

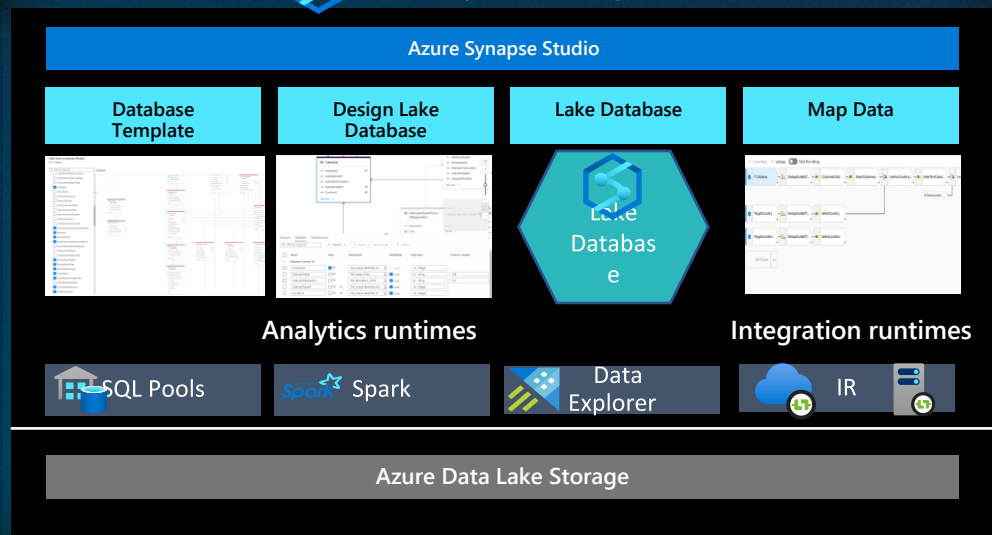
Iberian
Technology
Summit_



Resource Group Development



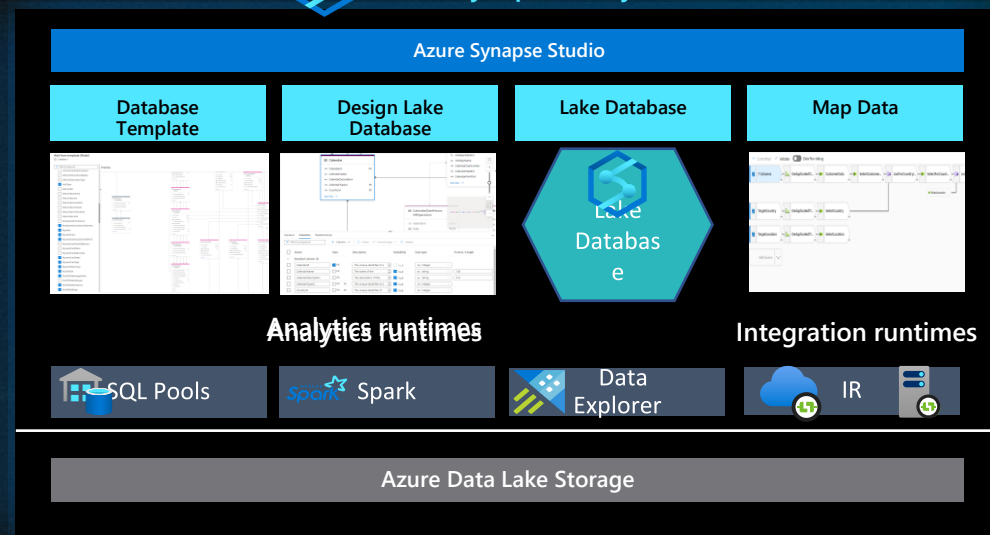
Azure Synapse Analytics



Resource Group Production



Azure Synapse Analytics

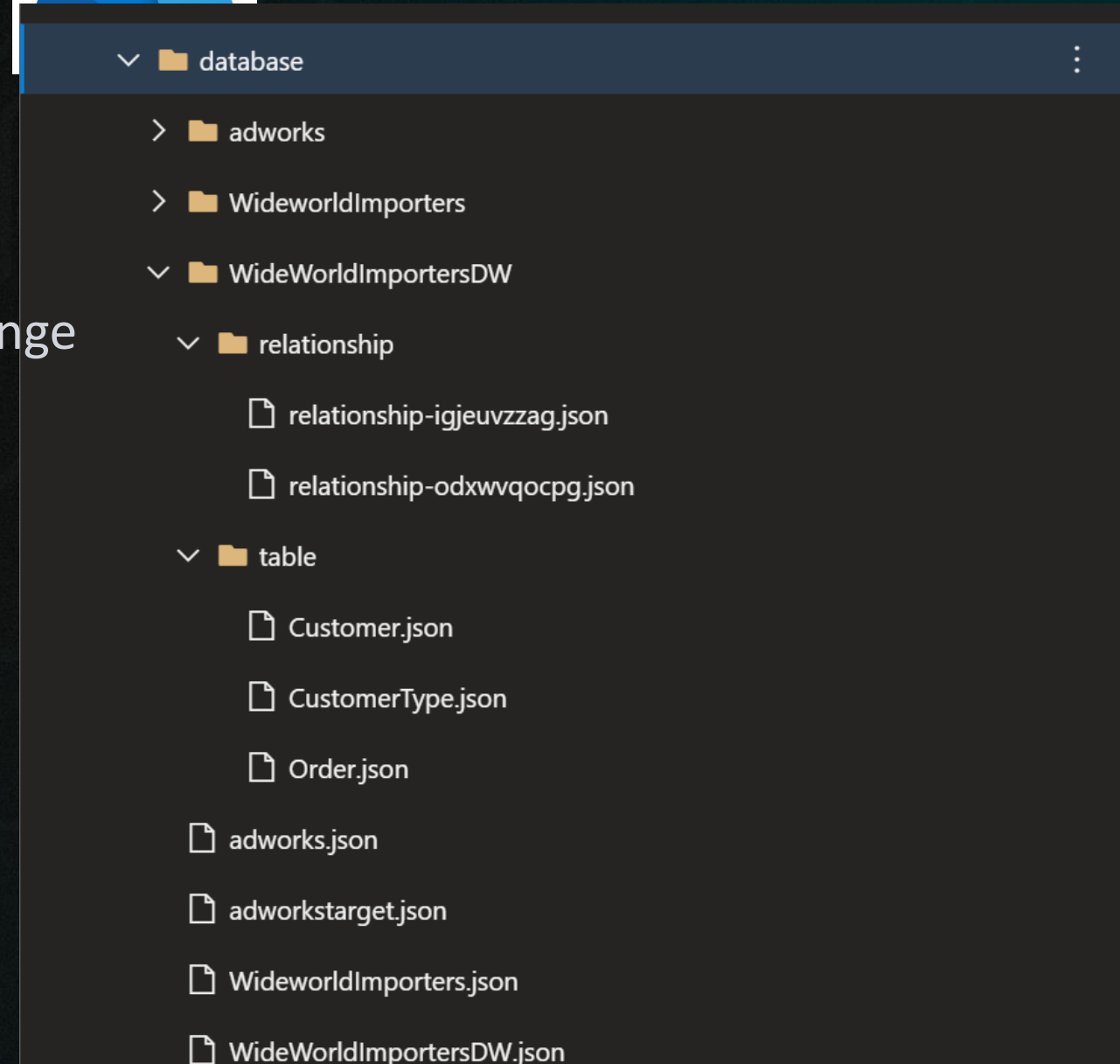


Iberian
Technology
Summit

Azure Dev Ops

- Connect to Azure Dev Ops
 - Database Entity's
 - Lake Linked Service will automatically change

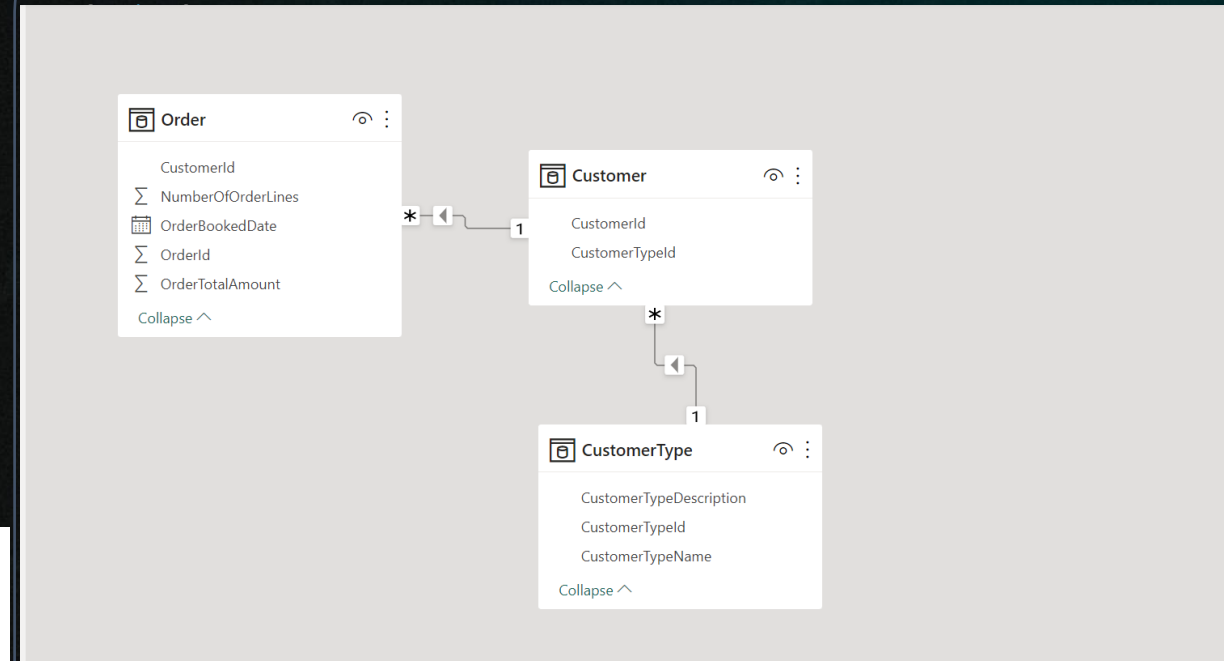
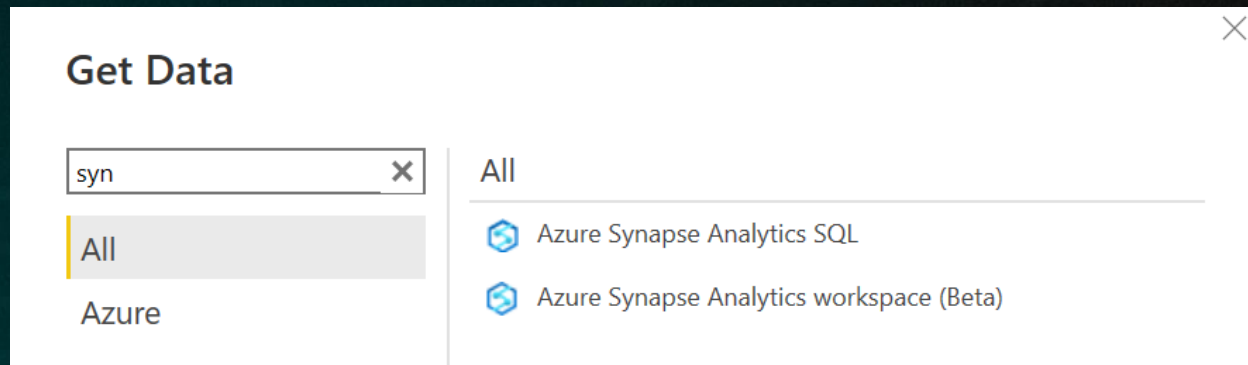
Iberian
Technology
Summit_



Power BI



- Connect to Power BI
 - New connector
 - Azure Synapse Analytics Workspace (beta)
 - Relationship from template will be taken over



Take Aways

- Microsoft Purview will scan your Lake Database
- Publish before you connect
- Lake Database as source is not working (Yet)
- No Deltalake support



International
Workplace
Group



Iberian
Technology
Summit_

Questions ??

Obrigado_
Thank You_
Gracias_



Erwin de Kreuk

Principal Consultant – Lead Data & AI
InSpark



@erwindekrek



linkedin.com/in/erwindekrek



erwindekrek.com github.com/edkrek

