

SQL *days* konferenz

Reicht auch Power BI only
für mein Projekt?

oder brauche ich was von Azure / SQL Server?

Stefan Kirner

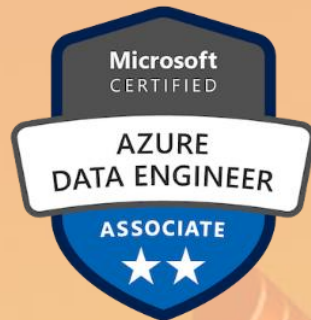
Director Business Intelligence @scieneers

Über mich

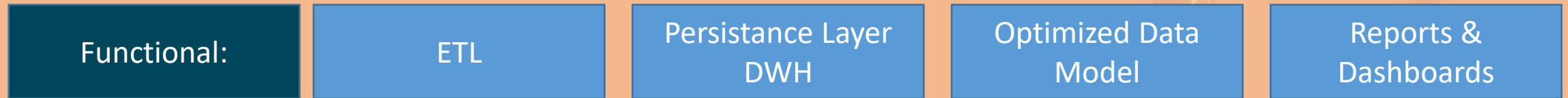
Stefan Kirner



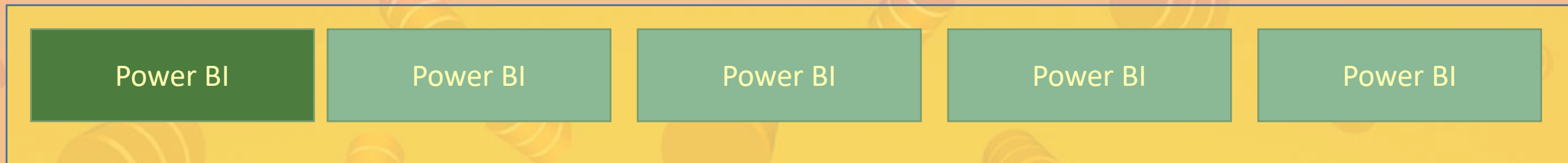
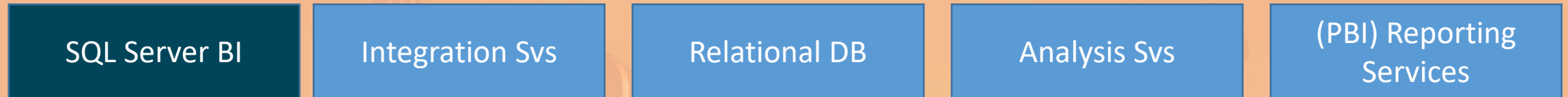
- › PASS Chapter Lead Karlsruhe ski@sqlpass.de
- › Director Business Intelligence scieneers GmbH
- › Microsoft P-TSP Data Platform
- › Twitter: @KirnerKa



Typische Schichten einer BI Lösung



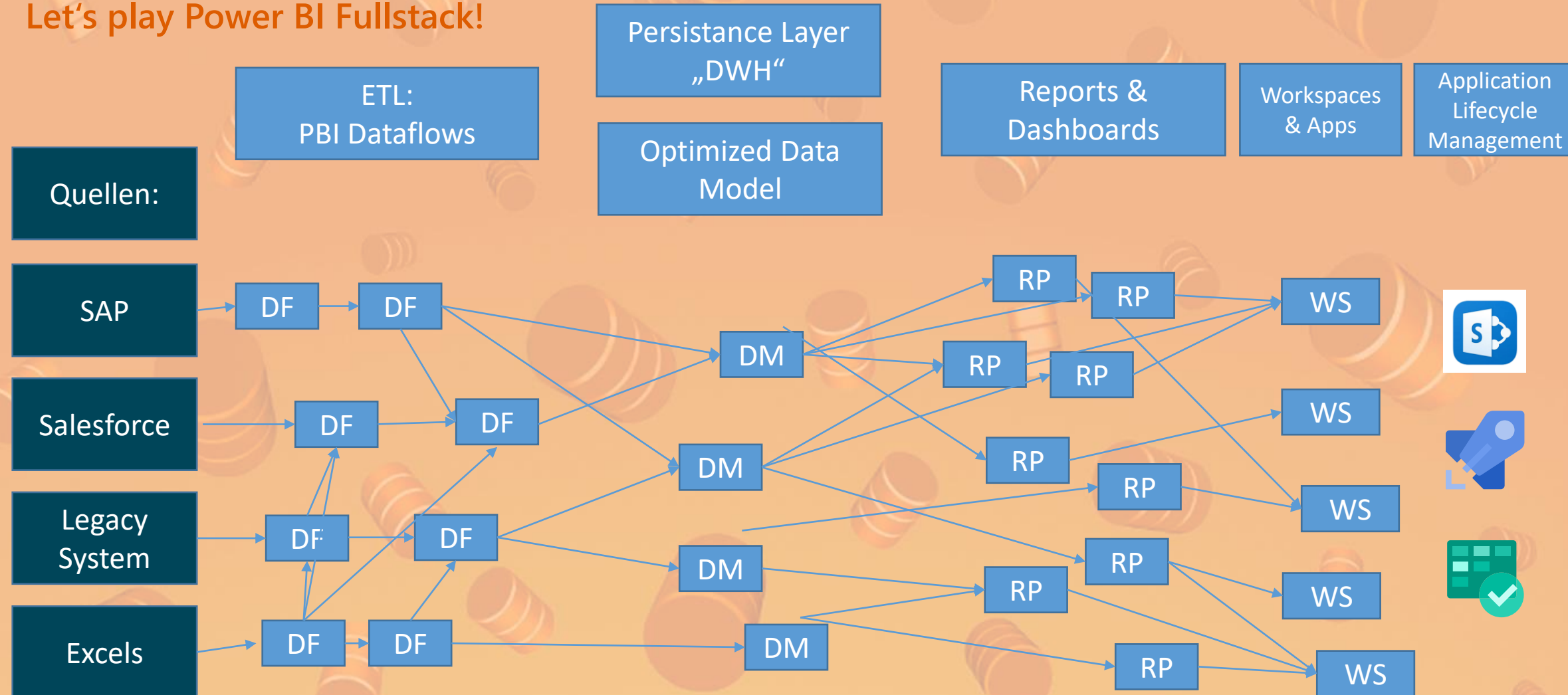
Microsoft Data Platform Stacks:



Ist doch alles da!
Reicht das nicht ?!

Ein Beispielprojekt aus dem wahren Leben

Let's play Power BI Fullstack!



Power BI only für so ein Projekt- geht das denn?!

Prinzipiell schon, aber mit Schmerzen bei

- Entwicklung & Application Lifecycle Management im Team
- Performance bei großen Datenmodellen
- Transparenz über Business Logik
- Komplexität Row-Level Security
- Hohe Komplexität für Fachabteilung bei BI-Betrieb
- U.a.



Woran liegt das?

Einschränkungen von Power BI

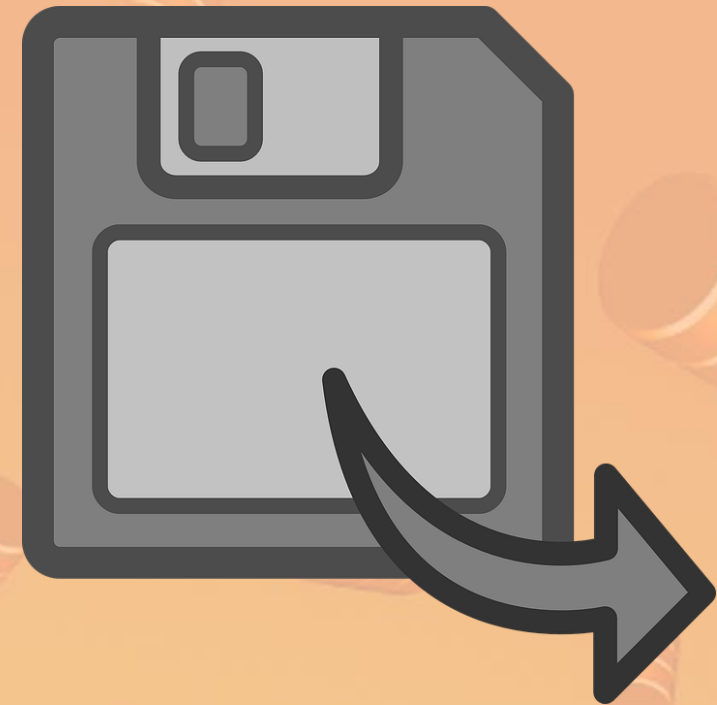
- Keine „echte“ Persistenz
- Keine „echte“ Entwicklungsumgebung
- Kein „echtes“ Error-Handling
- Problematisch für „richtig“ große Datenmengen
- Deployment Pipelines unvollständig
- Kapazitäts-Limitierungen bes. in Pro Version



Persistenz – Historisierung / SCDs

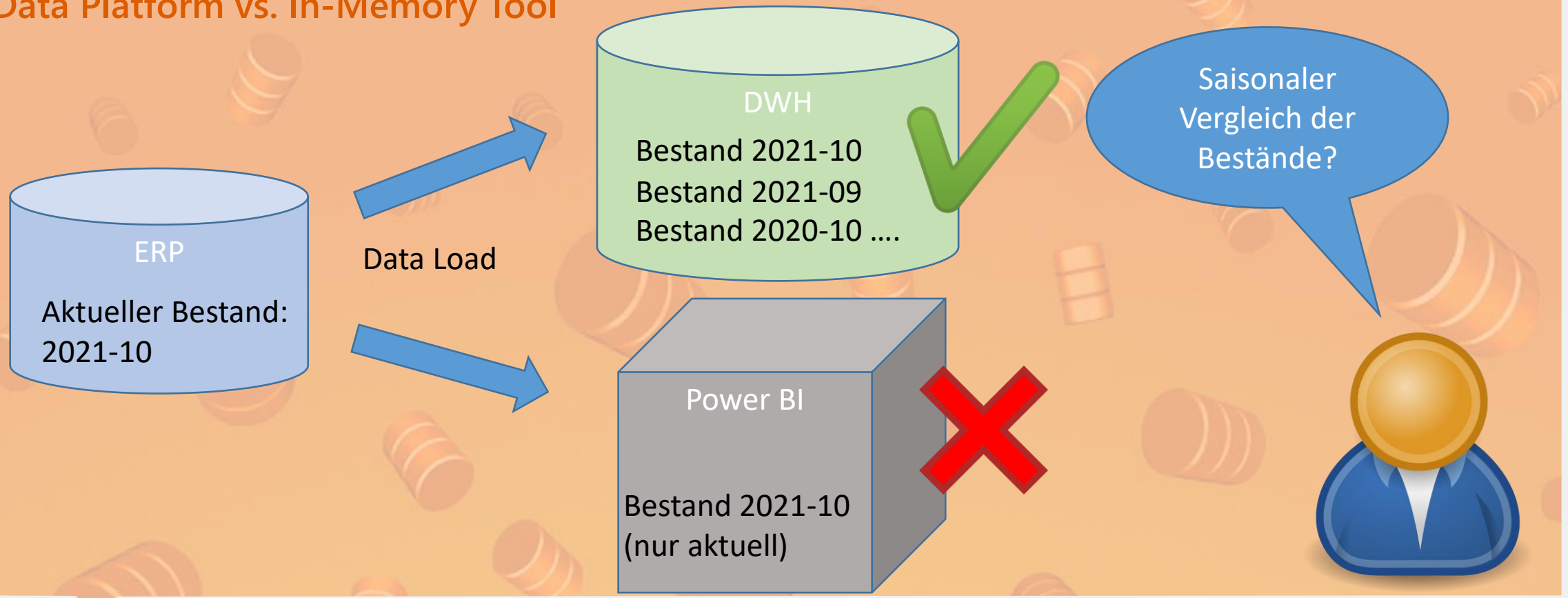
Typische Beispiele – über die Zuordnung

- Beispiele
 - Sich verändernde Zuordnung von Vertriebsgebieten
 - Stichtagesbestände (Lager, Bestellungen, Aufträge...)
 - Daten in Quellen sind nicht auf Dauer verfügbar
- Braucht i.d.R. historisierte Stände von Daten
- In DWH einfach und flexibel lösbar
- In Power BI auch, aber nur kompliziert über Basteleien



Persistenz – Historisierung / SCDs

Data Platform vs. In-Memory Tool



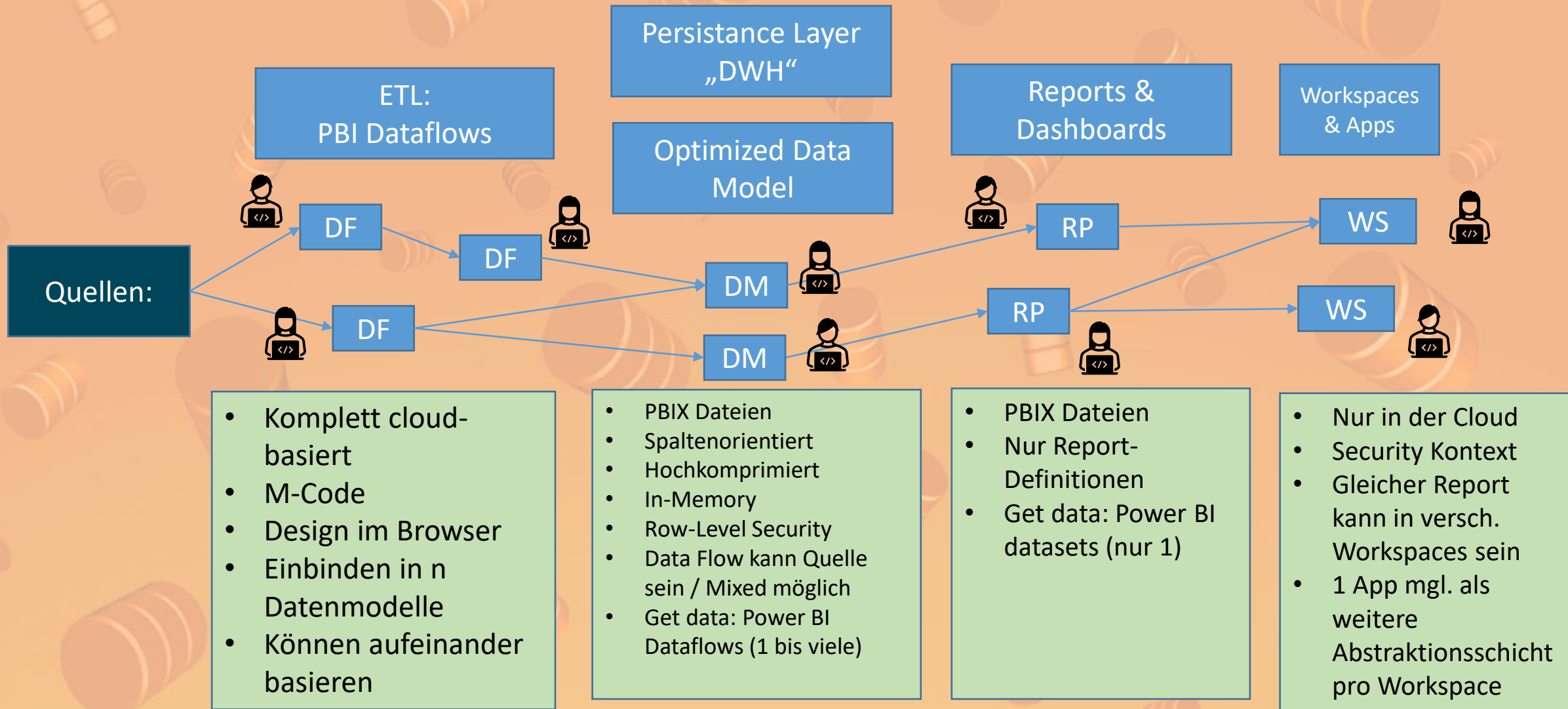
Application Lifecycle Management

Entwicklungsumgebung

- Entwicklung überwiegend in PBIX Dateien:
 - PBI Desktop für Windows only
 - Kein paralleles Bearbeiten
 - Kein integriertes Merge / Diff
- Hohe Anforderungen Hardware bei Datenmenge
- vs. Extra Aufwand Datenset für Entwicklung
- Versionschaos Power BI Desktop App
- Deployment ist in der Regel all-or-nothing
- Workarounds:
 - Auslagern Reports und Modelle
 - Power BI Dataflows
 - Integration Zusatztools PBI Desktop*



Entwicklung skalieren



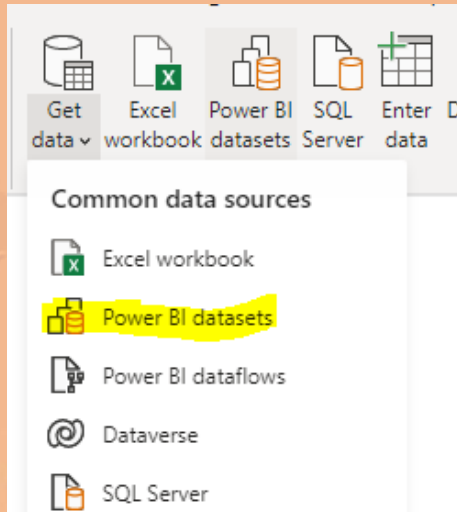
Nutzen von Power BI dataflows

The image shows the 'Get data' menu in Power BI Desktop and the 'Navigator' window. The 'Get data' menu lists 'Common data sources' including Excel workbook, Power BI datasets, **Power BI dataflows** (highlighted), Dataverse, and SQL Server. The 'Navigator' window displays a tree view of dataflows under 'Display Options'. The tree includes 'Power BI dataflows [2]', 'DemoDataflows [2]', 'test myTurn [1]', 'WideWorldImporters-Sales [1]', 'Fifty Shades [1]', and 'ShadesOfRopes [1]'. The 'Ropes' dataset is selected, and its preview is shown on the right. The preview table has two columns: 'Material' and 'Popularity (0 - lowest, 4 - highest)'. The table contains five rows of data.

Material	Popularity (0 - lowest, 4 - highest)
Nylon	
MFP	
Filament Polyester	
Hempex	
Spun Polyester	

At the bottom of the Navigator window, there are buttons for 'Load', 'Transform Data', and 'Cancel'.

Nutzen von Power BI datasets



Select a dataset to create a report

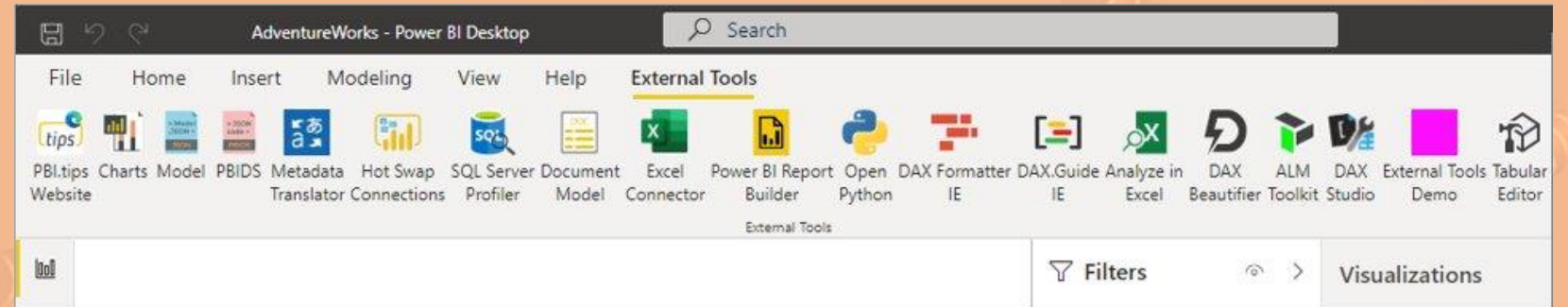
All Recent

	Name	Endorsement	Owner	Workspace	Refreshed
	Azure Consumption scieneers ...	–	Stefan Kirner	all scieneers	9/7/21, 3:59:39 PM
	PBI_Solutions_Showcase_v1	–	Stefan Kirner	Power BI Solution Sh...	10/5/21, 4:34:17 PM
	PBI_Solutions_Showcase_v1	–	Milo Sikora	all scieneers	10/4/21, 3:47:46 PM
	Sales and Marketing Sample P...	–	Stefan Kirner	Sales and Marketing ...	4/10/20, 6:01:06 PM

Zusatztools

- DAX Studio
- Tabular Editor
- ALM Toolkit

...



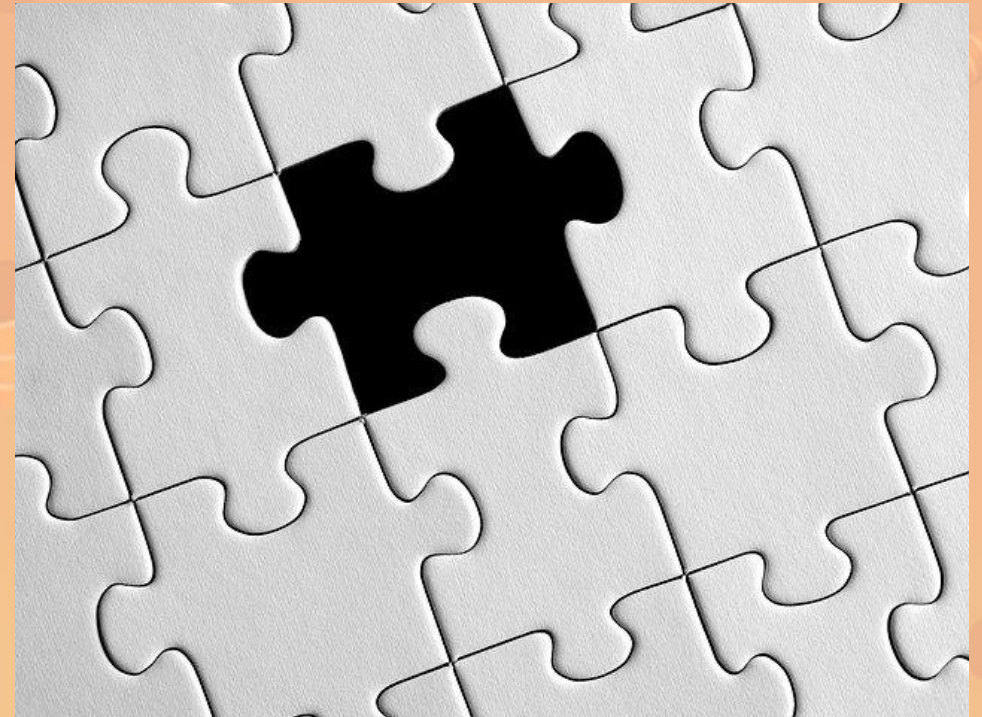
<https://docs.microsoft.com/en-us/power-bi/transform-model/desktop-external-tools>



Application Lifecycle Management

Weitere Aspekte

- Debugging von technischen / fachlichen Fehlern in Dataflows zeitaufwändig und umständlich
- Deployment nur Premium feature, hat Einschränkungen *
 - keine linked entities bei Dataflows
 - Teile von Power BI wie gar nicht unterstützt z.B. Streaming datasets
- Power BI Rest-APIs für eigene Erweiterungen mit Powershell / Azure DevOps unvollständig



* <https://docs.microsoft.com/en-us/power-platform-release-plan/2021wave1/power-bi/deployment-pipelines--dataflows-support>

Deployment pipelines in Power BI

The screenshot displays the Power BI Deployment Pipelines interface for a workspace named "Documentation demo". The interface is organized into three main stages: Development, Test, and Production, each with its own description and a list of deployed content.

Development Stage:

- Description: Design, review, and revise your content in a development workspace. When it's ready to test and preview, deploy the content to the test stage.
- Content Summary: 1 Dataflow (preview), 2 Datasets (Refresh all), 5 Reports, 2 Dashboards.
- Actions: Publish app, Deploy to test.

Test Stage:

- Description: Test and verify your content in a preproduction workspace. When it's ready to distribute, deploy the content to the production stage.
- Content Summary: 1 Dataflow (preview), 2 Datasets (Refresh all), 5 Reports, 2 Dashboards. Deployed: 6/7/2021, 10:03:13 AM.
- Actions: Publish app, Deploy to production.

Production Stage:

- Description: Your content has been tested and is ready to distribute to your consumers as an app or by access to the production workspace.
- Content Summary: 1 Dataflow (preview), 2 Datasets (Refresh all), 3 Reports, 1 Dashboard. Deployed: 6/14/2021, 10:49:23 AM.
- Actions: Publish app, Update app.

The interface also includes a left-hand navigation pane with options like Home, Favorites, Recent, Create, Datasets, Goals, Apps, Shared with me, Deployment pipelines (selected), Discover, Learn, Workspaces, and My workspace. At the bottom of the interface, there is a list of content items: CDM, MyFoods income, Sales analysis, MyFoods income, Sales analysis, CountrySalesPerformance, and OrganizationExpenditures.

Deployment pipelines Azure DevOps

The screenshot displays the Azure DevOps web interface for configuring a deployment pipeline. The left sidebar shows a list of tasks under the 'DEV-PowerBI' pipeline. The main panel is titled 'Power BI Actions' and shows the configuration for a specific task.

Task Configuration:

- Task version:** 4.*
- Display name:** Power BI Action: Publish Duplicate Address
- Authentication type:** User
- Power BI service connection:** Power BI
- Action:** Upload PowerBI Report

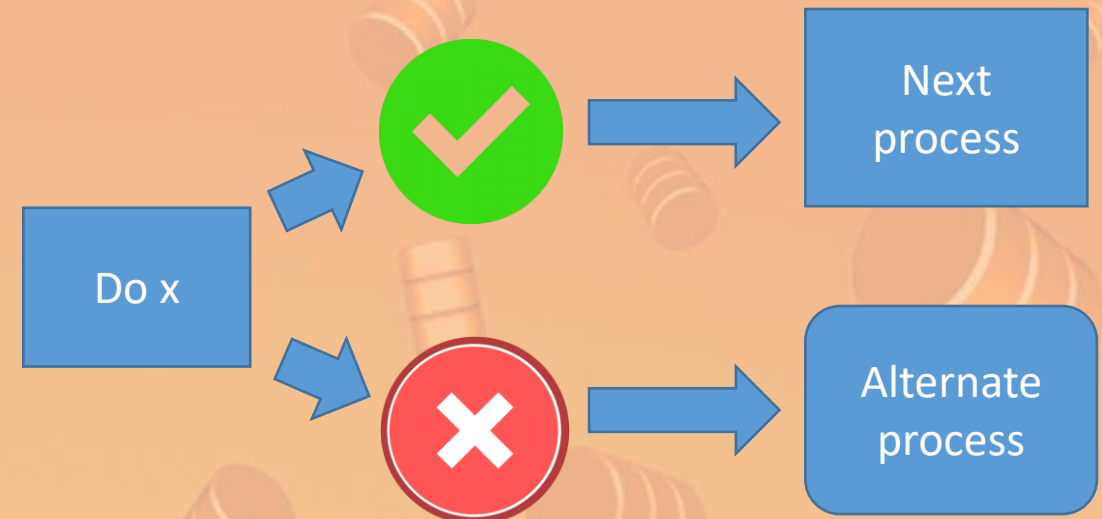
The task list on the left includes:

- Power BI Action: Publish Duplicate Address
- Power BI Action: Publish ...
- Power BI Action: Publish ... Template
- Power BI Action: Publish ... Workshop Teilnehmera...
- Power BI Action: Publish ...
- Power BI Action: Publish ... Qualifizierungsreport
- Power BI Action: Publish ... langfristige Quali Wunsch
- Power BI Action: Publish: ... Leistungs- und Erfolgsre...
- Power BI Action: Publish Marktcheck
- Power BI Action: Publish IT-Auswertungen

Error- Handling

Abfangen von Fehlern in M eher schwierig

- Keine einfache Möglichkeit Fehler generisch abzufangen wie error constraint in ADF oder SSIS
- Try catch in M gibt's über Umwege, aber nicht unbedingt intuitiv und intransparent*



*<https://4pbi.com/m-language-try-catch-expression/>

Grenzen bei Datenmodellen

Einschränkungen vor allem beim Service*

- Import Model - max Anzahl Refresh / Tag (8x/48x)
- Max Speicher pro Workspace:
 - 10 GByte bei Shared per Pro Lizenz im App Workspace
 - 100 GByte bei Premium
- Pro Dataset max 1 Gbyte nutzbar
- Zeit für full load über Netz benötigt, incremental load nur über Premium
 - Nur begrenzt Einfluss auf Ladeperformance (ETL) bei Managed Service
- Komprimierung hilft, aber wie stark hängt von Daten / Modell ab
 - Nur begrenzt Einfluss auf Abfrage-Performance der Modelle vor allem in Shared Workspaces



Vgl. <https://docs.microsoft.com/en-us/power-bi/admin/service-admin-manage-your-data-storage-in-power-bi>

Transparenz Business Logik

In M oder DAX oder beides?

- Vorberechnete Werte besser für Abfrage-Performance (M)
- Aber: Manche Berechnungen Kontext-Abhängig (DAX)
- Also oft beides notwendig
- Komplexität steigt bei Verteilung auf verschiedene Artefakte (dataflows & model & reports)
- Schwierige Kapselung von Funktionen und klar definierte Schichten
- Führt schnell zu Intransparenz und aufwändiger Fehlersuche



Row-Level Security in verteilten Datenfiles

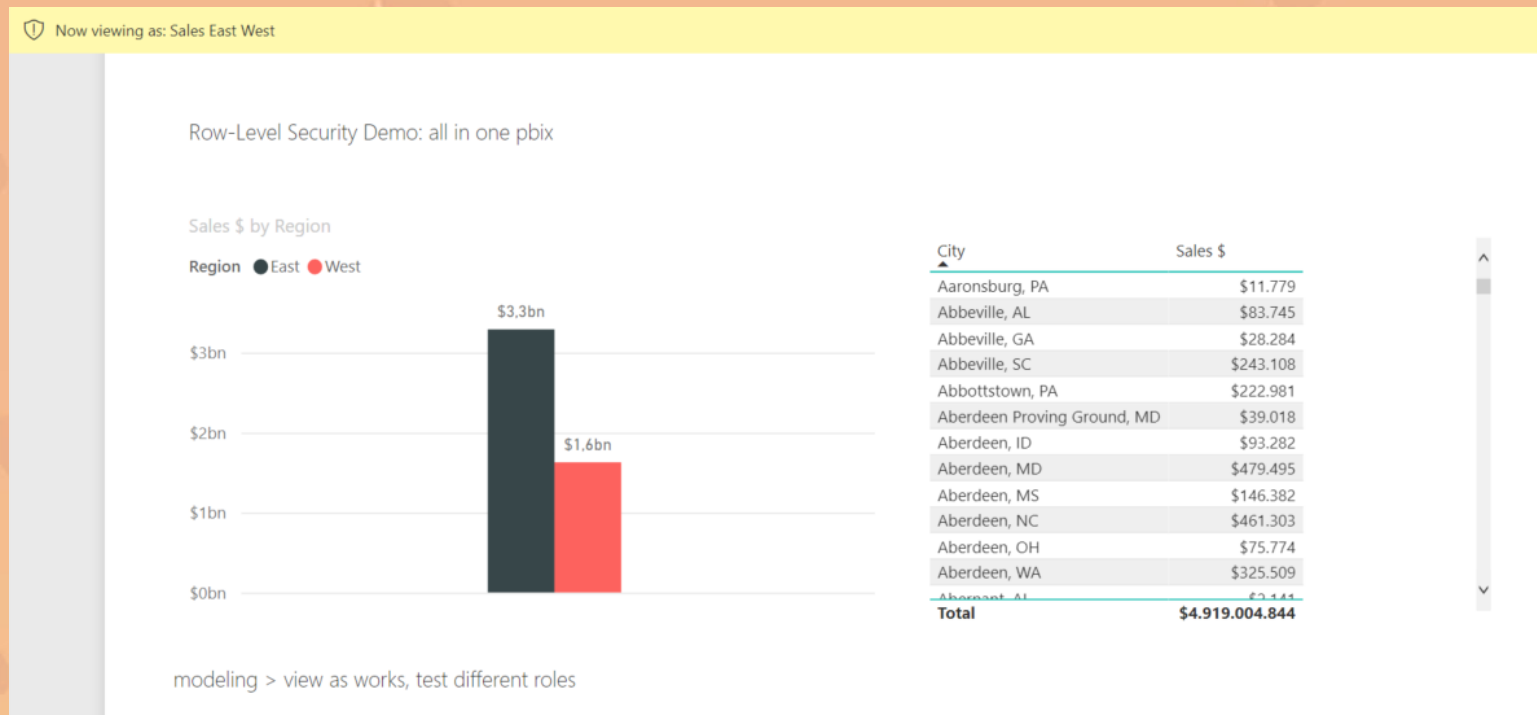
„Anschauen als“ geht nur im Kontext des Modells

- Skalierung der Entwicklung über Aufteilung von Modellen / Reports in unterschiedliche PBIX Dateien möglich
- Testen der Reports über den Security Kontext einer Rolle ist dann aber nur schwierig möglich
- Kopieren des Reports in Modelldatei führt schnell zu unerwünschten Effekten



Row-Level Security in verteilten Datenfiles

„Anschauen als“ geht nur im Kontext des Modells



Die besten Einsatzszenarien für „Power BI only?“

Komfortzonen Power BI

- Fachabteilungslösungen mit wenigen Datenquellen
- Quellsysteme mit eigener Historisierung
- Viele kleine unabhängige Problemstellungen im Unternehmen



Fazit: Reicht Power BI für alles?

Kann man machen – ist dann aber auch kompliziert

- Scope der Projekts zu Beginn abgrenzen für Entscheidungen
 - Weitere Reporting Tools?
 - Weiterverarbeitung der Daten?
- Früh BI-Betriebs-Orga klären
- Premium in Erwägung ziehen
- Weitere Komponenten und Zusatz-Tools nutzen um Schwächen punktuell ausgleichen



Fragen?

Vielen Dank!

Ich freue mich auf Euer Feedback!

