



GOLD SPONSORS



Volvo Group IT

SILVER SPONSORS



BRONZE SPONSOR



STRATEGIC PARTNER



Transformuj swoje dane w Power BI

Adrian Chodkowski

O mnie

- Adrian Chodkowski
- Data platform enthusiast...
- Business Intelligence Consultant
- MCP, MCSA, MCSE:BI,
MCSE: Data Management and Analytics
- **seequality.net**
- Adrian.Chodkowski@outlook.com
- @Twitter: Adrian_SQL
- LinkedIn: <http://tinyurl.com/adrian-sql>



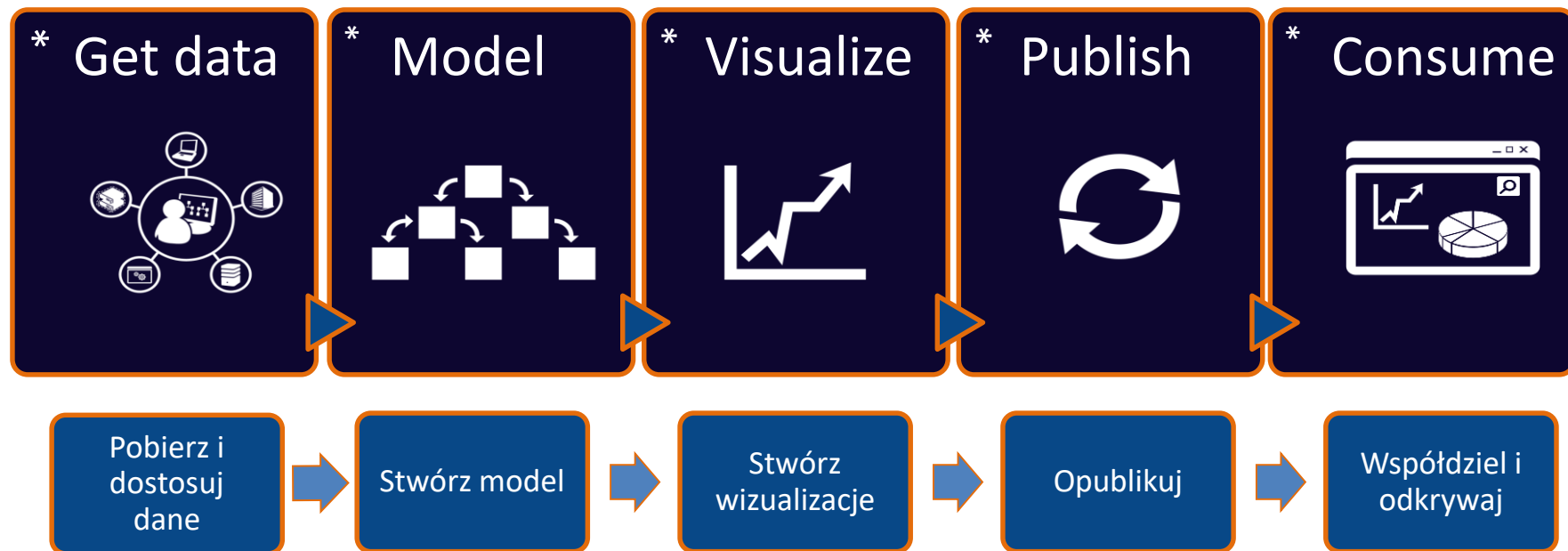
seequality
the way to understand

Agenda

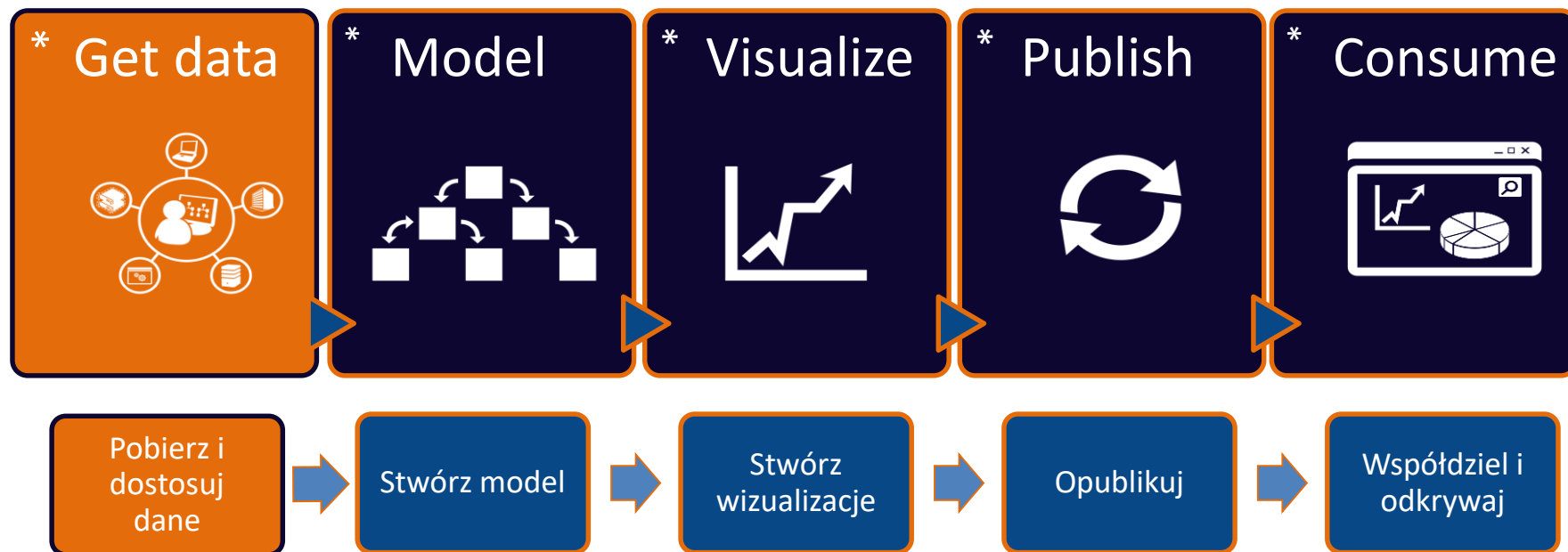
- ✓ WSTĘP DO POWER QUERY W POWER BI
- ✓ POBIERANIE DANYCH I TRANSFORMACJE Z GUI
- ✓ QUERIES I QUERY STEPS
- ✓ LET IN I KONSTRUKCJE JĘZYKA M
- ✓ FUNKCJE UŻYTKOWNIKA
- ✓ QUERY FOLDING,
- ✓ PRIVACY LEVELS,
- ✓ PARAMETERS & TEMPLATES
- ✓ INCREMENTAL LOAD
- ✓ PODSUMOWANIE



Cykl tworzenia raportu

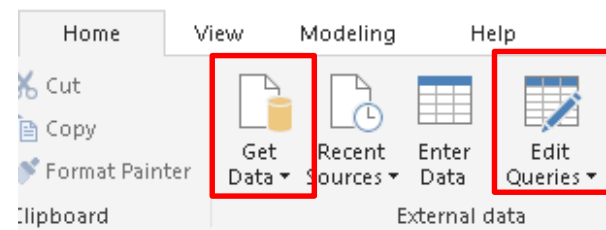


Miejsce ETL w cyklu

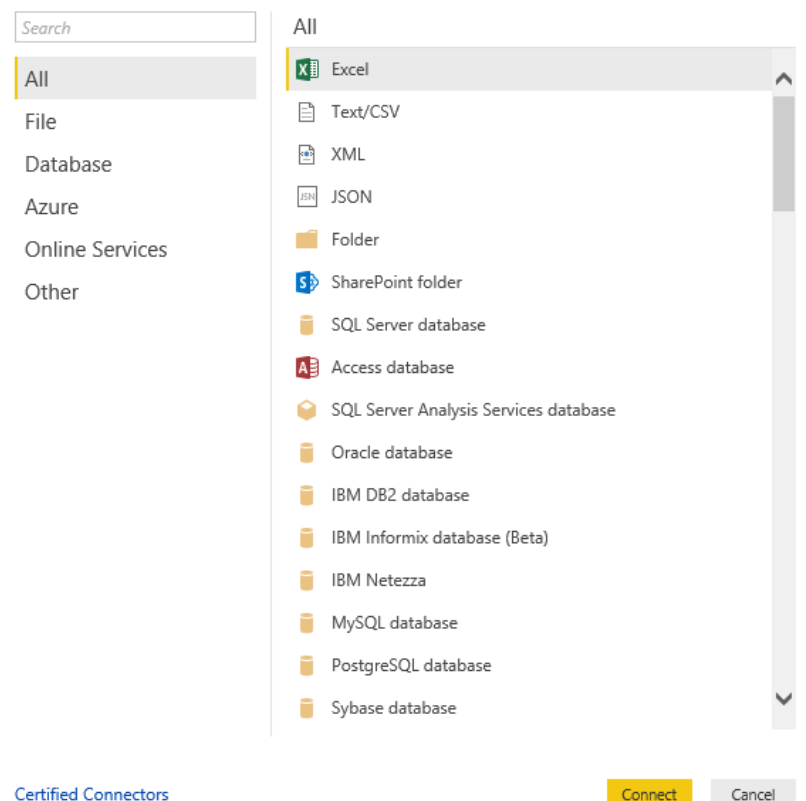


Pobieranie danych do Power BI

- ✓ Aby załadować dane należy użyć opcji dostępnych w sekcji **External data** na zakładce Home
- ✓ Power BI posiada dużą ilość konektorów do różnych systemów i struktur
- ✓ W przypadku gdy dane do których się łączymy nie potrzebują żadnych transformacji po prostu je ładujemy
- ✓ Jeśli jednak dane wymagają dostosowania należy włączyć edytor zapytań i stworzyć odpowiedni proces ETL



Get Data



Interfejs

The screenshot shows the Microsoft Power Query interface. The ribbon at the top contains tabs: Plik, Home, Transform, Add Column, View, and Help. The ribbon is divided into groups: Close & Apply, New Source, Recent Sources, Enter Data, Data source settings, Manage Parameters, Refresh Preview, Properties, Advanced Editor, Manage, Choose Columns, Remove Columns, Keep Rows, Remove Rows, Sort, Split Column, Group By, Transform, Merge Queries, Append Queries, and Combine Files. The main area is divided into three panes: a left pane for 'Queries [13]', a central data view, and a right pane for 'QUERY SETTINGS'. The central data view shows a table with columns: CustomerKey, GeographyKey, CustomerAlternateKey, Title, NameStyle, and Suffix. The right pane shows the 'APPLIED STEPS' list, which includes: Source, Promoted Headers, Changed Type, Reordered Columns, Duplicated Column, Split Column by Delimiter, Changed Type1, Removed Columns, Renamed Columns, Reordered Columns1, Inserted Merged Column, and Renamed Columns1. The bottom status bar indicates '31 COLUMNS, 999+ ROWS' and 'PREVIEW DOWNLOADED AT 13:11'.

1. Wstążka i opcje dostępne z jej poziomu

2. Panel podglądu danych

3. Panel zdefiniowanych zapytań

4. Zestaw kroków w wybranym zapytaniu

5. Pasek formuły z kodem M w wybranym kroku zapytania



Wstęp do ETL w Power BI

DEMO

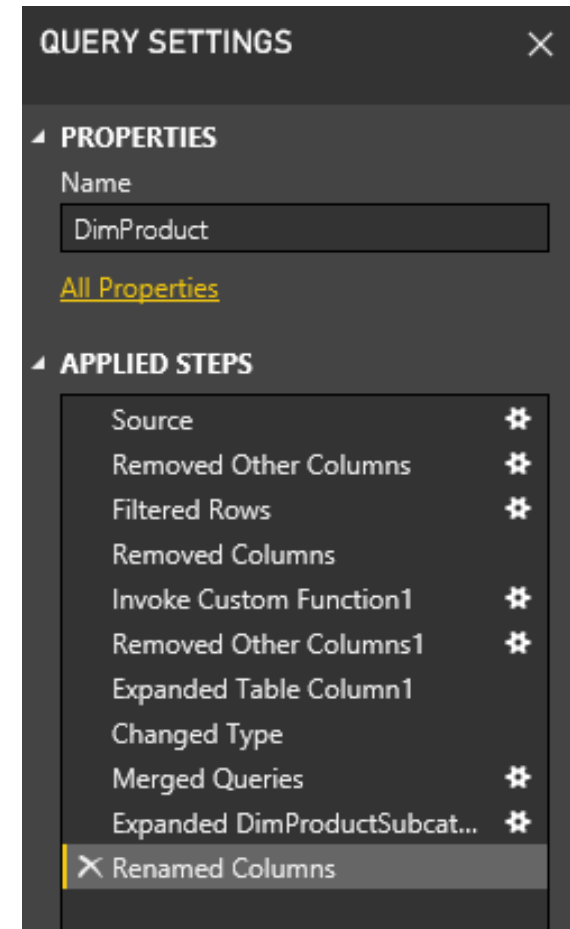
Pierwsze kroki

- ✓ Znakomitą część transformacji można zrobić z GUI
- ✓ Pod spodem generowany jest kod M dostępny po włączeniu **Formula Bar** i/lub **Advanced Editor**
- ✓ Zapytania to nic innego jak zapisany zestaw kroków
- ✓ Nawet gdy nie edytujemy zapytań to pod spodem i tak angażowany jest **kod M**
- ✓ Transformacji można używać zarówno w trybie Import jak i Direct Query (mimo, że w tym drugim nie jest wspierany pełny zakres transformacji)



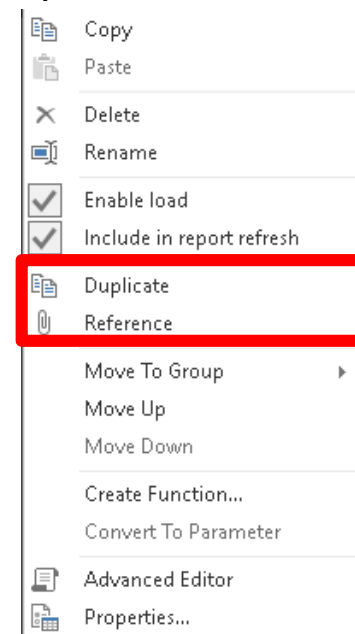
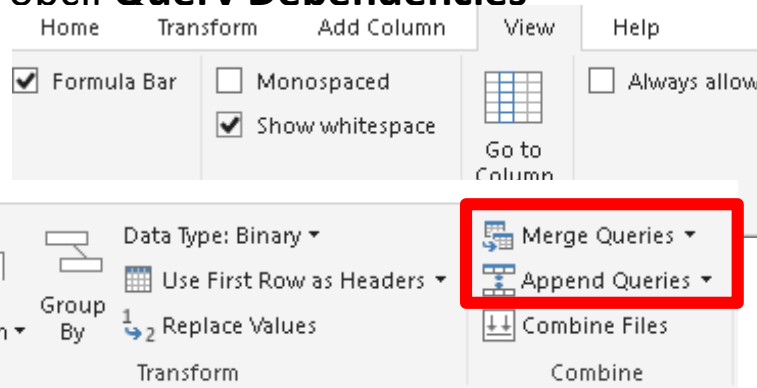
Applied steps

- ✓ Pojedyncze zapytanie składa się z kroków
- ✓ Każdy krok jest sparametryzowaną operacją
- ✓ Pierwszym krokiem każdego zapytania jest zazwyczaj krok pobierający dane ze źródła, każdy kolejny krok pozwala na zdefiniowanie transformacji
- ✓ Kolejność kroków nie ma znaczenia z punktu widzenia M ale ma znaczenie dla sposobu wyświetlania Applied Steps
- ✓ Całe zapytanie jest zdefiniowany za pomocą języka M
- ✓ Możemy się do niego dostać używając tzw. **Advanced Editor**,
- ✓ Kod pod pojedynczym krokiem możemy podejrzeć używając **Formula Bar**



Zależności pomiędzy zapytaniami

- ✓ Pomiedzy zapytaniami mogą istnieć zależności
- ✓ Zapytanie można duplikować wraz ze wszystkimi krokami (**Duplicate**)
- ✓ Można również z poziomu jednego zapytania odnosić się do rezultatu innego zapytania (**Reference**)
- ✓ Mogą one również być ze sobą łączone wertykalnie (**Append**) lub horyzontalnie (**Merge**)
- ✓ Powiązania pomiędzy zapytaniami można odczytać używając opcji **Query Dependencies**






Zależność pomiędzy zapytaniami

DEMO

Ładowanie zawartości listy plików

- ✓ Odczytać metadane plików jak i ich zawartość możemy za pomocą konektora **Folder**
- ✓ Pomaga on nam pobrać również zawartość plików z określonej lokalizacji



	Content	Name	Extension	Date accessed	Date modified	Date created	Attributes	Folder Path
1	Binary	Products1-100.txt	.txt	29.04.2018 14:11:12	29.04.2018 14:11:13	29.04.2018 14:11:12	Record	C:\Users\adria\Desktop\sqlday\Demos\Old\Product data\
2	Binary	Products101-200.txt	.txt	29.04.2018 14:12:13	29.04.2018 14:12:14	29.04.2018 14:12:13	Record	C:\Users\adria\Desktop\sqlday\Demos\Old\Product data\
3	Binary	Products201-400.txt	.txt	29.04.2018 14:13:10	29.04.2018 14:13:10	29.04.2018 14:13:10	Record	C:\Users\adria\Desktop\sqlday\Demos\Old\Product data\
4	Binary	Products401-1000.txt	.txt	29.04.2018 14:14:11	29.04.2018 14:14:12	29.04.2018 14:14:11	Record	C:\Users\adria\Desktop\sqlday\Demos\Old\Product data\

Combine Files

Specify the settings for each file. [Learn more](#)

Example File:

First file

File Origin

Delimiter

Data Type Detection

1200: Unicode

Tab

Based on first 200 rows

ProductKey	ProductAlternateKey	ProductSubcategoryKey	WeightUnitMeasureCode	SizeUnitMeasureCode	ProductName	Standard
1	AR-5381				Adjustable Race	
2	BA-8327				Bearing Ball	
3	BE-2349				BB Ball Bearing	
4	BE-2908				Headset Ball Bearings	
5	BL-2036				Blade	
6	CA-5965				LL Crankarm	
7	CA-6738				ML Crankarm	
8	CA-7457				HL Crankarm	
9	CB-2903				Chaining Bolts	
10	CN-6137				Chaining Nut	
11	CR-7833				Chaining	
12	CR-9981				Crown Race	
13	CS-2812				Chain Stays	
14	DC-8732				Decal 1	
15	DC-9824				Decal 2	
16	DT-2377				Down Tube	
17	EC-M092				Mountain End Caps	

☐ Skip files with errors

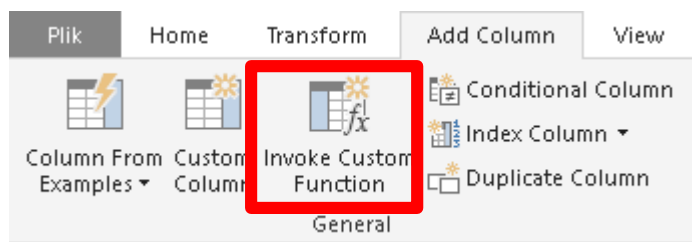
OK

Cancel

ProductKey	ProductAlternateKey	ProductSubcategoryKey	WeightUnitMeasureCode	SizeUnitMeasureCode	ProductName	Standar
1	AR-5381				Adjustable Race	FALSE NA
2	BA-8327				Bearing Ball	FALSE NA
3	BE-2349				BB Ball Bearing	FALSE NA
4	BE-2908				Headset Ball Bearings	FALSE NA
5	BL-2036				Blade	FALSE NA
6	CA-5965				LL Crankarm	FALSE NA
7	CA-6738				ML Crankarm	FALSE NA
8	CA-7457				HL Crankarm	FALSE NA
9	CB-2903				Chaining Bolts	FALSE NA
10	CN-6137				Chaining Nut	FALSE NA
11	CR-7833				Chaining	FALSE NA
12	CR-9981				Crown Race	FALSE NA
13	CS-2812				Chain Stays	FALSE NA
14	DC-8732				Decal 1	FALSE NA
15	DC-9824				Decal 2	FALSE NA
16	DT-2377				Down Tube	FALSE NA
17	EC-M092				Mountain End Caps	FALSE NA
18	EC-8098				Road End Caps	FALSE NA
19	EC-7209				Touring End Caps	FALSE NA
20	FI-6760				Rack End	FALSE NA
21	FI-2961				Frontwheel	FALSE NA
22	FI-1080				Flat Washer 1	FALSE NA
23	FI-1200				Flat Washer 6	FALSE NA
24	FI-1400				Flat Washer 2	FALSE NA
25	FI-1400				Flat Washer 9	FALSE NA
26	FI-1800				Flat Washer 4	FALSE NA
27	FI-1800				Flat Washer 8	FALSE NA
28	FI-1800				Flat Washer 3	FALSE NA
29	FI-1800				Flat Washer 5	FALSE NA
30	FI-1900				Flat Washer 7	FALSE NA
31	FI-1900				Front Crown	FALSE NA
32	FI-1900				Front Derailleur Cage	FALSE NA
33	FI-1900				Front Derailleur Linkage	FALSE NA
34	FI-1900				Grille Pulley	FALSE NA
35	FI-1900				LL Grip Tape	FALSE NA
36	FI-1209				ML Grip Tape	FALSE NA
37	FI-1209				HL Grip Tape	FALSE NA
38	FI-1213				Thin Jam Hex Nut 9	FALSE NA
39	FI-1213				Thin Jam Hex Nut 10	FALSE NA
40	FI-1420				Thin Jam Hex Nut 1	FALSE NA
41	FI-1420				Thin Jam Hex Nut 2	FALSE NA
42	FI-1420				Thin Jam Hex Nut 15	FALSE NA

Funkcje użytkownika

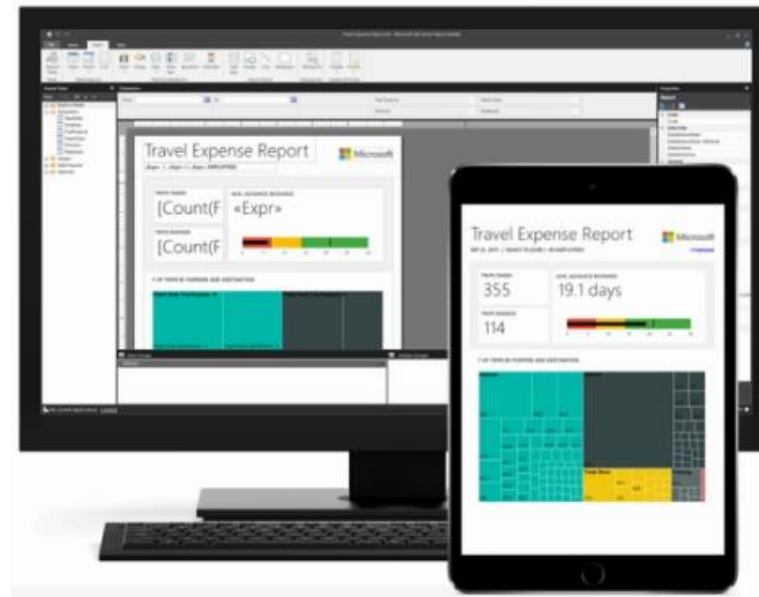
- ✓ Język M umożliwia tworzenie własnych funkcji
- ✓ Funkcja to nic innego jak zapisana pod określoną nazwą sekwencja transformacji
- ✓ Standardowe zapytanie może stać się funkcją przez dodanie () =>
- ✓ Najczęściej jest parametryzowana i na podstawie podanej wartości wykonuje określone czynności
- ✓ Może być wywołana z GUI poprzez opcję **Invoke Custom Function**



```
let
Source = paramFile => let
    Source = Excel.Workbook(File.Contents(paramFile), null, true),
    Dane_Sheet = Source[[Item="Dane",Kind="Sheet"]][Data],
    #"Changed Type" = Table.TransformColumnTypes(Dane_Sheet,{{"Column1",
    #"Removed Top Rows" = Table.Skip(#"Changed Type",2),
    #"Promoted Headers" = Table.PromoteHeaders(#"Removed Top Rows", [Pro
    #"Changed Type1" = Table.TransformColumnTypes(#"Promoted Headers",{{
    #"Unpivoted Other Columns" = Table.UnpivotOtherColumns(#"Changed Typ
    #"Renamed Columns" = Table.RenameColumns(#"Unpivoted Other Columns",

in
    #"Renamed Columns"

in
    Source
```



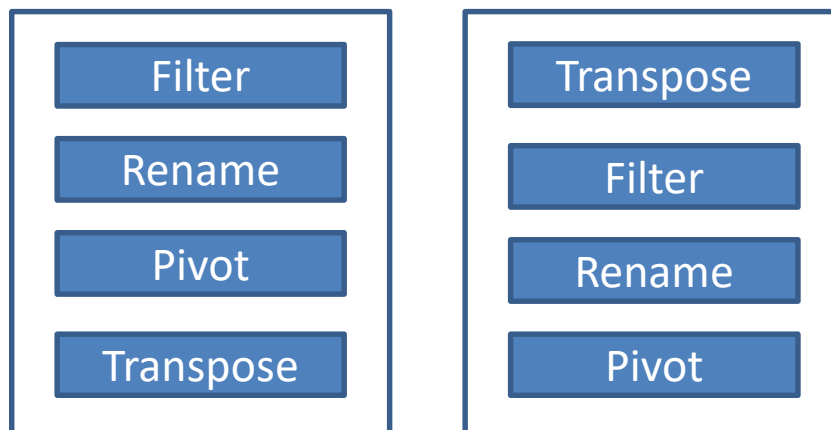
Ładowanie zawartości listy plików i funkcje

DEMO

Query Folding

- ✓ Mechanizm transformujący zapytania M na zapytanie źródła
- ✓ Jeśli coś nie może zostać przerzucone to jest wykonywane lokalnie przez Power BI
- ✓ Zapytanie może zostać przerzucone w całości lub częściowo
- ✓ Zależy od:
 - ✓ Typu źródła
 - ✓ Zastosowanej logiki transformacji
 - ✓ Privacy Level

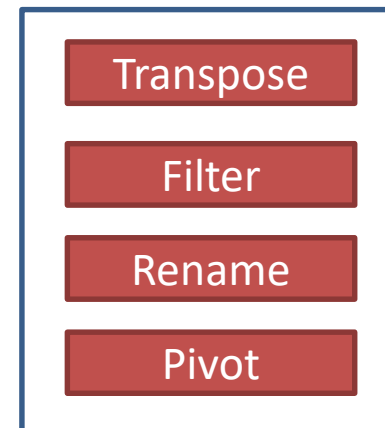
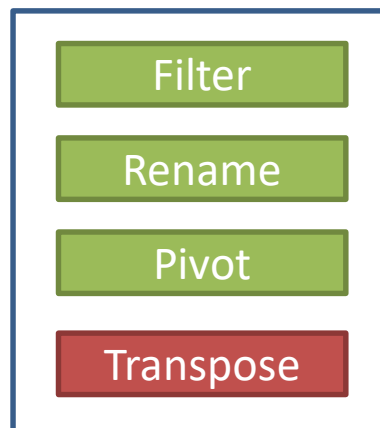
**Które transformacje
zostaną przerzucone
na źródło?**



Query Folding

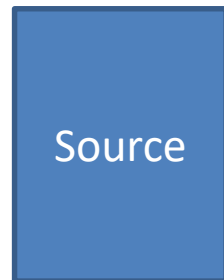
- ✓ Mechanizm wydajnościowy transformujący zapytania M na zapytanie źródła
- ✓ Jeśli coś nie może zostać przerzucone to jest wykonywane lokalnie przez Power BI
- ✓ Zależy od:
 - ✓ Typu źródła
 - ✓ Zastosowanej logiki transformacji
 - ✓ Privacy Level

**Transformacja która nie może być
przerzucona powoduje, że
wszystkie następujące
po niej transformacje też nie
zostaną przerzucone!**



Query Folding

1



```
SELECT  
ProductKey,  
TransactionID,  
Value  
From Tabela  
WHERE CompanyID=1
```



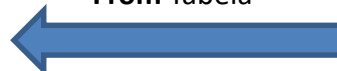
10 tys wierszy



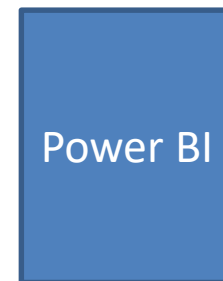
2



```
SELECT  
ProductKey,  
TransactionID,  
Value  
From Tabela
```



100 mln wierszy



WHERE
CompanyID=1



10 tys. wierszy





Query Folding

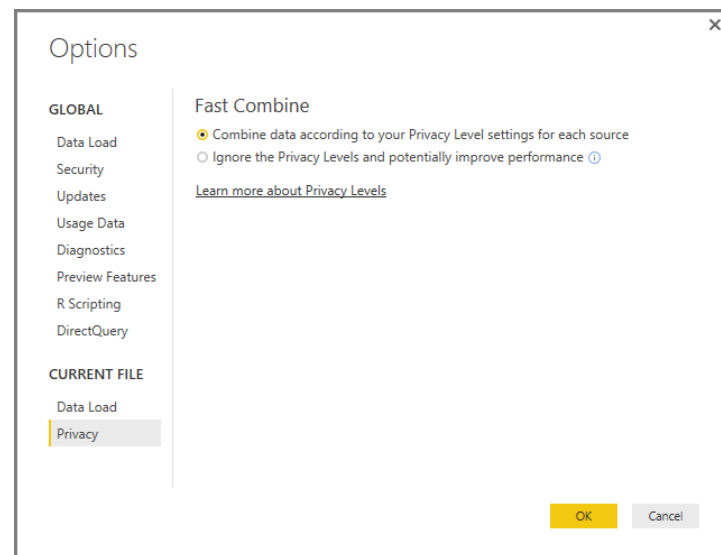
DEMO

Privacy Levels

Definiują poziom izolacji jednego źródła danych od innego, a co za tym idzie interakcję między nimi w przypadku **Query Folding**.

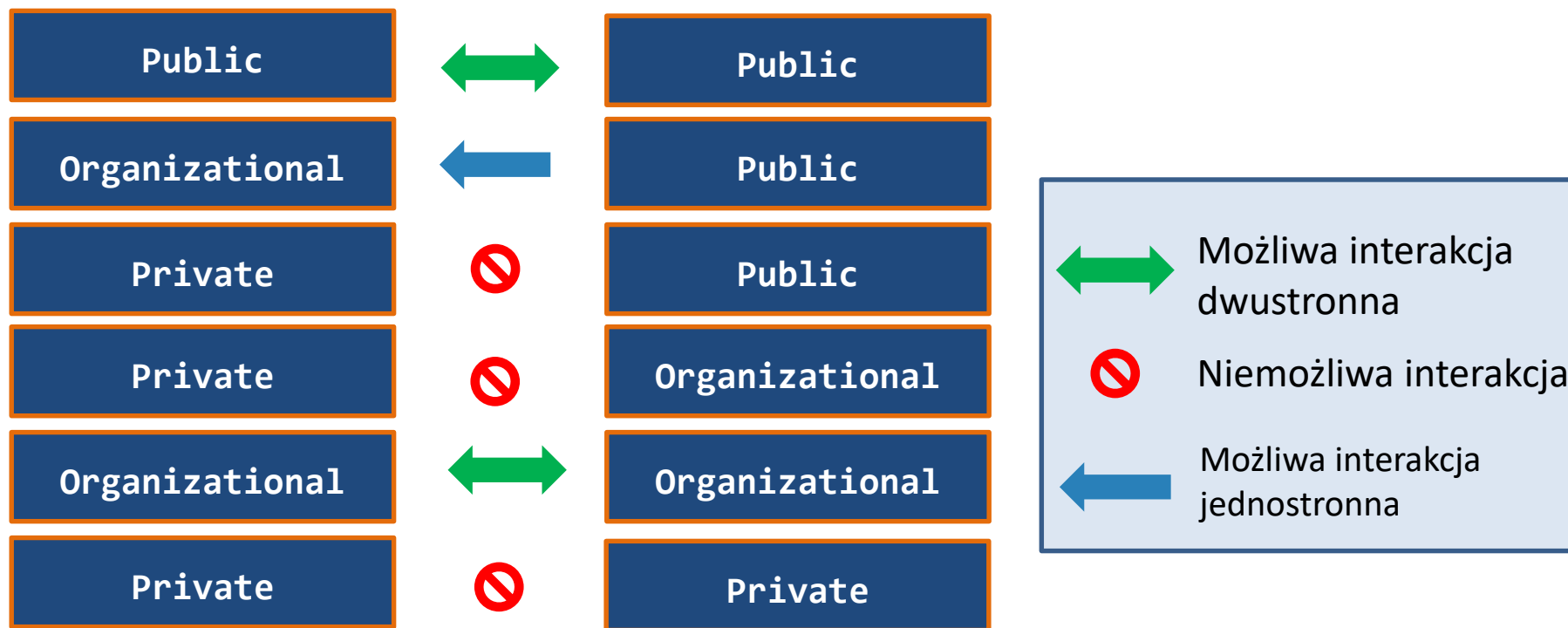
Ustawienie to przyjmuje 4 wartości:

- ✓ **Private** – dane z tego źródła nie mogą być przesłane do innego źródła
- ✓ **Organizational** – dane z tego źródła mogą być przesłane do innych źródeł tego samego typu
- ✓ **Public** – dane z tego źródła mogą być przesłane do innych źródeł Public oraz do źródeł Organizational
- ✓ **None** – ustawienie dziedziczone po źródle rodzicu, jeśli rodzic nie ma zdefiniowanego PL. Jeśli rodzic nie ma zdefiniowanego PL to poziom izolacji jest niezdefiniowany.



Privacy Levels

Interakcja pomiędzy źródłami jest określona poziomami izolacji. Dane nie mogą przepływać ze źródła bardziej do mniej restrykcyjnego. Jeśli korzystamy z jednego źródła można zignorować poziomy izolacji lub ustawić je jako Organizational.



Query Parameters i Templates

- ✓ **Query Parameters** to obiekty pozwalające wpłynąć na zapytania w języku M
- ✓ Mają określony typ, a ich wartości mogą być wpisywane bez żadnych restrykcji lub na podstawie wartości zwracanych z innych zapytań
- ✓ Standardowym zastosowaniem tych obiektów jest parametryzacja szczegółów połączeń do źródła czy chociażby filtracja danych przychodzących do aplikacji
- ✓ Query Parameters mogą być wykorzystane do tworzenia **Templates** czyli plików o rozszerzeniu **pbit** raportu bez danych, który danymi się wypełni po podaniu wartości parametrów
- ✓ Parametry wykorzystywane są również przy scenariuszu odświeżania przyrostowego (**Incremental Refresh**)



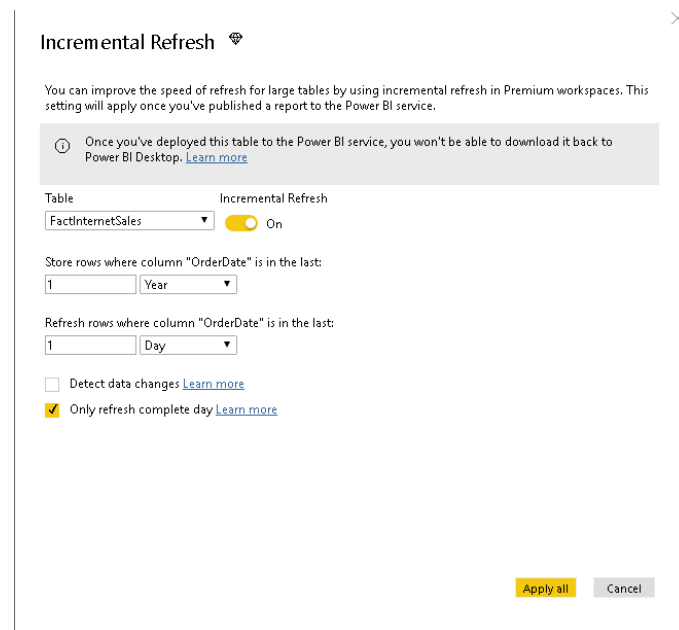


Query Parameters & Templates

DEMO

Incremental refresh*

- ✓ W majowej aktualizacji usługi możemy definiować ładowanie inkrementalne
- ✓ Dotyczy to jedynie **Power BI Premium**
- ✓ Definiowanie tego typu ładowania opiera się na użyciu parametrów i zdefiniowaniu:
 - ✓ Przez jaki okres dane będą przechowywane
 - ✓ Za jaki okres dane mają zostać odświeżone
 - ✓ Czy ma zostać użyta kolumna audytowa służąca do detekcji czy coś się zmieniło
 - ✓ Czy mają ładowanie ma się opierać tylko na pełnych dniach
- ✓ Po implementacji nie można ściągnąć pbix
- ✓ Na ten moment funkcjonalność jest w fazie **preview**



Incremental Refresh

You can improve the speed of refresh for large tables by using incremental refresh in Premium workspaces. This setting will apply once you've published a report to the Power BI service.

① Once you've deployed this table to the Power BI service, you won't be able to download it back to Power BI Desktop. [Learn more](#)

Table: FactInternetSales Incremental Refresh: ☒ On

Store rows where column "OrderDate" is in the last: 1 Year

Refresh rows where column "OrderDate" is in the last: 1 Day

☐ Detect data changes [Learn more](#)

☒ Only refresh complete day [Learn more](#)

[Apply all](#) [Cancel](#)

Materialy

- ✓ Seequality.net ([link](#))
- ✓ Power Query M Reference ([link](#))
- ✓ Power Query formula language specification ([link](#))
- ✓ Chris Webb (<https://blog.crossjoin.co.uk>)
- ✓ Matt Mason (<https://www.mattmasson.com/>)



Dziękuję!

Adrian Chodkowski

adrian.chodkowski@outlook.com

seequality.net



GOLD SPONSORS



Volvo Group IT

SILVER SPONSORS



BRONZE SPONSOR



STRATEGIC PARTNER

