UNIWERSYTET EKONOMICZNY W POZNANIU

UNIWERSYTET EKONOMICZNY W POZNANIU · KATEDRA INFORMATYKI EKONOMICZNEJ

Informatyczne narzędzia wizualizacji danych

Zajęcia 2 · ETL w Power BI: Power Query — pobieranie, czyszczenie i łączenie danych, parametry
rok akademicki 2026/2027 · dr inż. Szczepan Górtowski

Po tych zajęciach

- wczytujesz pliki z folderu jednym zapytaniem i wiesz, co sprawdzić: kodowanie, ogranicznik, kulturę
- czyścisz tekst, typy i duplikaty w kolejności, która nie gubi danych
- łączysz zapytania: scalasz słowniki, dołączasz pliki, szukasz kodów bez pary
- zmieniasz kształt tabeli: anulujesz przestawienie, wypełniasz w dół
- używasz parametru zamiast ścieżki i sprawdzasz wynik sumą kontrolną

Surowe pliki z zajęć: trzy problemy na start

14 → 12

plików w folderze sprzedaży: kopia i opis nie są danymi

35 → 8

zapisów marki w katalogu i marek po oczyszczeniu

48

pustych komórek w lokalizacjach do wypełnienia w dół

ETL: pobierz, przekształć, załaduj

Power Query to „T” w ETL. Ten sam edytor jest w Excelu, w Power BI i w przepływach danych Fabric.

01

Źródła

pliki CSV i Excel, bazy danych, usługi w chmurze, strony WWW

02

Power Query

kroki w języku M: typy, czyszczenie, łączenie, kształt

03

Model

tabele faktów i wymiarów, relacje — zajęcia 3

04

Raport

miary DAX i wizualizacje — zajęcia 4–6

▲ **dziś: zajęcia 2**

Przekształcaj jak najbliżej źródła, jak się da

Źródło

Baza, hurtownia, system firmy. Tu poprawisz błąd raz — dla wszystkich raportów.

Power Query

Kiedy źródła nie zmienisz: typy, czyszczenie, łączenie plików, kształt tabel.

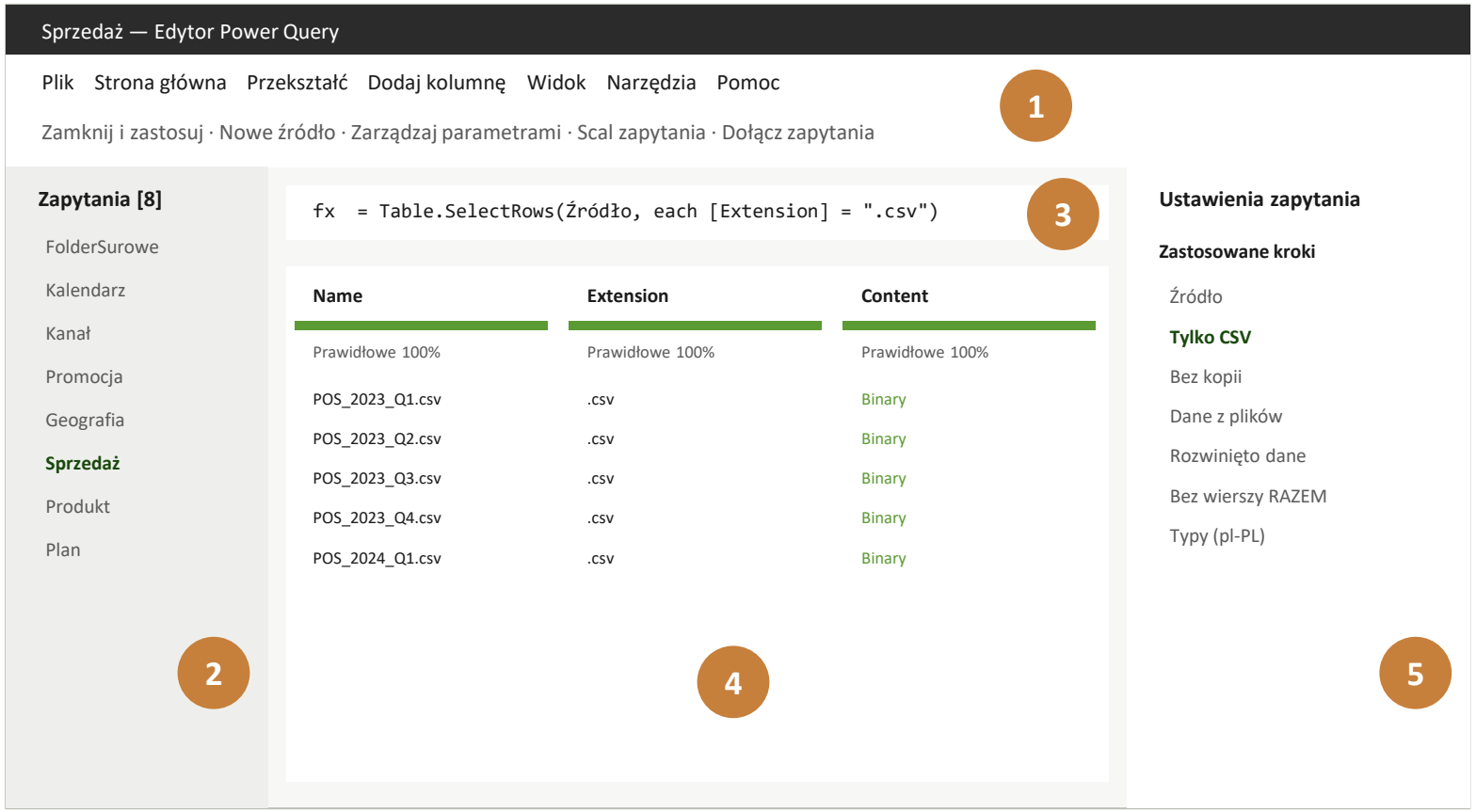
DAX

Tylko to, co zależy od filtrów raportu: sumy, udziały, porównania w czasie.

M. Roche: „dane przekształcaj tak wysoko w górę strumienia, jak to możliwe, i tak nisko, jak to konieczne” (Roche's Maxim of Data Transformation, 2021).

Edytor Power Query: pięć miejsc, które musisz znać

Narzędzia główne → Przekształć dane. Zmiany zapisują się jako kroki; do raportu wracasz przez Zamknij i zastosuj.



1 · wstążka: Przekształć zmienia kolumnę, Dodaj kolumnę tworzy nową

2 · zapytania: każde to przyszła tabela modelu albo pomocnik

3 · pasek formuły: kod M zaznaczonego kroku (Widok → Pasek formuły)

4 · podgląd z profilem kolumn: jakość, rozkład, błędy

5 · zastosowane kroki: przepis wykonywany od nowa przy każdym odświeżeniu

Każdy krok to jedna linia kodu M

Zapytanie Geografia z pliku startowego. Kliknięcia w edytorze piszą ten kod za ciebie — warto umieć go przeczytać.

```
let
    Źródło = Excel.Workbook(File.Contents(FolderSurowe & "lokalizacje.xlsx"), null, true),
    Sklepy = Źródło{[Item = "Sklepy", Kind = "Sheet"]}[Data],
    #"Usunięto górne wiersze" = Table.Skip(Sklepy, 2),
    #"Nagłówki o podwyższonym poziomie" = Table.PromoteHeaders(#"Usunięto górne wiersze", [Prom…
    #"Zmieniono typ" = Table.TransformColumnTypes(#"Nagłówki o podwyższonym poziomie", {{"Makro…
in
    #"Zmieniono typ"
```

- let … in: lista kroków i wynik
- krok = funkcja(poprzedni krok, …)
- #"Nazwa ze spacją" — nazwy kroków w cudzysłowie
- each … [Kolumna] — „dla każdego wiersza”
- zmiana kroku w środku przelicza wszystkie następne

JAK CZYTAĆ M

Od dołu: wynik to ostatni krok. Każdy krok bierze poprzedni i coś w nim zmienia — tak jak historia zmian w Excelu, tylko zapisana i powtarzalna.

Ćw. 0 · Otwórz plik i obejrzyj dane

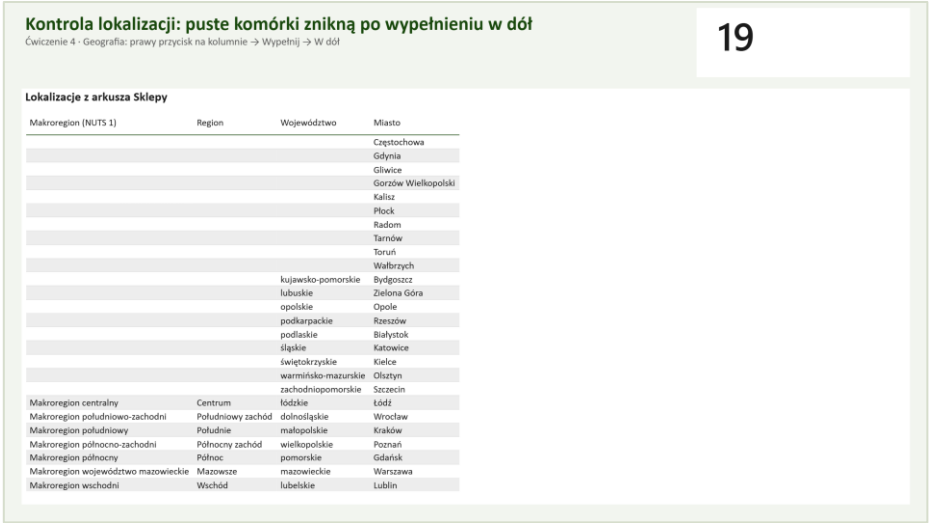
5 min

1. Pobierz INWD_pq_cwiczenia.zip, rozpakuj do C:\ i otwórz INWD_pq_cwiczenia.pbip. Odśwież dane.

2. Narzędzia główne → Przekształć dane (albo panel Dane → … przy tabeli → Edytuj zapytanie). Obejrzyj zapytania Kalendarz, Kanał, Promocja i Geografia.

3. Widok → Jakość kolumn, Rozkład kolumn, Profil kolumny.

4. Na pasku stanu: Profilowanie kolumn w oparciu o cały zestaw danych. Ile pustych komórek ma Region w Geografii?



Wskazówka: Profil domyślnie liczy tylko pierwsze 1000 wierszy — przy większych tabelach przełącz go na cały zestaw danych.

Strona Kontrola w pliku startowym: miasta bez regionu.

01

Pobieranie: folder, kodowanie, kultura

Plik CSV to nie tabela: cztery pytania przed importem

Pierwsze wiersze i ostatni wiersz pliku POS_2023_Q1.csv — dokładnie tak, jak zapisał je system POS.

```
Eksport POS · sieć sklepów z elektroniką (dane fikcyjne)
Okres 01.01.2023–31.03.2023 · wygenerowano 02.04.2023 06:00
Data;Kod produktu;Kod sklepu;Kanał;Kod promocji;Ilość;Cena jedn.;Rabat;Wartość po rabacie;Koszt;Zwroty szt.;Zwroty wartość
01.01.2023;P0001;G11;Aplikacja mobilna;PR5;13;201,47;130,96;2488,15;2088,45;0;0,00
01.01.2023;P0069;G20;Aplikacja mobilna;;21;177,33;0,00;3723,93;1995,00;1;177,33
…
RAZEM;;;;;37308;;952157,48;56865733,05;40647251,40;1129;1781859,22
```

KODOWANIE

Windows-1250 — polskie znaki w nagłówkach („Ilość”, „Wartość”).

OGRANICZNIK

średnik, bo przecinek jest separatorem dziesiętnym.

KULTURA

przecinek dziesiętny i data dd.mm.rrrr — pl-PL, nie en-US.

ŚMIECI

dwa wiersze opisu nad nagłówkiem i wiersz RAZEM na końcu.

Złe kodowanie widać od razu w nagłówkach

Power Query zgaduje kodowanie z początku pliku. Jeśli zgadnie źle, polskie litery zamieniają się w „krzaki”.

Plik zapisany w	Odczytany jako	Tekst w pliku	Co zobaczysz
Windows-1250	UTF-8	Ilość	Ilo□□
UTF-8	Windows-1250	Ilość	IloĹ›Ä‡
Windows-1250	UTF-8	Wartość po rabacie	Warto□□ po rabacie

JAK NAPRAWIĆ

W oknie importu: Pochodzenie pliku → 1250: Środkowoeuropejski (Windows). W kodzie M: Csv.Document(…, [Encoding = 1250]). Pliki „CSV UTF-8” z Excela i eksporty z chmury to zwykle 65001: Unicode (UTF-8).

Bez kultury pl-PL 1 060 z 1 096 dni wyjdzie źle

Ta sama data w tekście i dwie kultury. Zapytanie Kontrola odczytało każdy dzień z eksportów na oba sposoby.

Tekst w pliku	Kultura pl-PL	Kultura en-US
01.02.2023	1 lutego 2023	2 stycznia 2023 — bez błędu, ale źle
13.02.2023	13 lutego 2023	błąd: nie ma 13. miesiąca
2488,15	2 488,15	248 815 — przecinek to dla en-US separator tysięcy

1 096

dni w 12 eksportach POS

664

dni, których en-US nie odczyta (błąd)

396

dni, które en-US odczyta jako inną datę — bez błędu

Poprawnie: Zmień typ → Używając ustawień regionalnych → Polski (Polska). W M: Table.TransformColumnTypes(…, "pl-PL"). Cichy błąd jest groźniejszy niż czerwony: nie widać go w podglądzie.

Folder zamiast dwunastu importów

Jedno zapytanie czyta wszystkie pliki z folderu. Nowy kwartał wrzucasz do folderu i klikasz Odśwież.

```
Źródło = Folder.Files(FolderSurowe & "sprzedaz")
#"Tylko CSV" = Table.SelectRows(Źródło, each [Extension] = ".csv")
#"Bez kopii" = Table.SelectRows(#"Tylko CSV", each not Text.Contains([Name], "kopia"))
#"Dane z plików" = Table.AddColumn(#"Bez kopii", "Dane", each
    Table.PromoteHeaders(Table.Skip(Csv.Document([Content],
        [Delimiter = ";", Columns = 12, Encoding = 1250]), 2)))
```

- Pobierz dane → Więcej… → Folder → Przekształć dane (nie „Połącz”) — najpierw zobaczysz listę plików
- filtrujesz pliki, potem kolumną niestandardową zamieniasz każdą zawartość (Content) w tabelę
- przycisk Połącz pliki robi to samo, ale tworzy cztery zapytania pomocnicze — trudniej je poprawić
- Folder.Files czyta też podfoldery: trzymaj w folderze tylko to, co ma trafić do modelu

Kopia pliku zawyżyłaby sprzedaż o 58,2 mln zł

W folderze sprzedaz leży 14 plików. Do modelu trafia 12: bez kopii i bez opisu formatu.

Plik	Co to jest	Wczytać?
POS_2024_Q2 - kopia.csv	kopia POS_2024_Q2.csv: 4 311 wierszy	nie — dubluje kwartał
opis_eksportu.txt	opis formatu dla ludzi	nie — to nie dane
POS_2023_Q1.csv … POS_2025_Q4.csv	12 eksportów kwartalnych	tak

ZASADA

Najpierw filtruj listę plików (Extension = .csv, nazwa bez „kopia”), dopiero potem łącz zawartość. Kopia nie daje błędu — daje tylko o 4 311 wierszy za dużo.

Suma kontrolna: różnica 0,00 zł

Każdy eksport kończy się wierszem RAZEM. Suma tych wierszy to kontrola dla całego zapytania.

850 573 496,17 zł

suma kolumny „Wartość po rabacie” z 12 wierszy RAZEM

850 573 496,17 zł

suma kolumny Wartość sprzedaży w modelu

0,00 zł

różnica — wiersze RAZEM odrzucone, kopia pominięta, liczby odczytane z kulturą pl-PL

- po każdej zmianie kroków porównaj liczbę wierszy i sumę z wartością kontrolną
- wiersz RAZEM podwoiłby sumę, a przy zmianie typu dałby błąd w kolumnie Data — filtr stoi przed typami
- zapisz kontrolę w dzienniku decyzji projektu: skąd liczba kontrolna, jaki wynik

Parametr zamiast ścieżki wpisanej na sztywno

Ścieżka z twojego komputera nie zadziała u nikogo innego. Parametr zmieniasz w jednym miejscu.

ŹLE

```
Źródło = Folder.Files("C:\Users\Anna\Downloads\eksporty\sprzedaz")
```

DOBRZE

```
Źródło = Folder.Files(FolderSurowe & "sprzedaz")
```

- w edytorze: Strona główna → Zarządzaj parametrami → Nowy parametr
- Nazwa FolderSurowe, typ Tekst, wartość bieżąca C:\INWD\dane_surowe\
- w kroku Źródło zamień ścieżkę na FolderSurowe & "sprzedaz"
- w raporcie: Przekształć dane → Edytuj parametry

PO CO

Projekt otworzysz na pracowni, w domu i w usłudze Power BI. Parametr przyda się też do filtrów: np. pierwszy rok danych albo nazwa pliku.

Ćw. 1 · Sprzedaż z dwunastu plików

20 min

1. Nowe źródło → Więcej… → Folder → C:\INWD\dane_surowe\sprzedaz → Przekształć dane. Nazwa: Sprzedaż.
2. Źródło: ścieżkę zamień na FolderSurowe & "sprzedaz". Filtr: Extension = .csv, Name nie zawiera „kopia”.
3. Kolumna niestandardowa Dane z formułą poniżej. Zostaw tylko Dane i ją rozwiń. Filtr: Data ≠ RAZEM.
4. Typy: Używając ustawień regionalnych → Polski. Nazwy kolumn jak w modelu. Pusty kod promocji → PR0.

```
Table.PromoteHeaders(Table.Skip(Csv.Document([Content],
    [Delimiter = ";", Columns = 12, Encoding = 1250]), 2))
```

PYTANIE

Zamknij i zastosuj. Ile wierszy ma Sprzedaż (Widok tabeli, pasek stanu)? Oczekiwane: 57 222.

02

Czyszczenie: tekst, typy, duplikaty

Katalog produktów: 35 zapisów marki, 8 marek

Ten sam produkt przed i po czyszczeniu. Surowe wiersze z produkty.xlsx, czyste — z tabeli Produkt w modelu.

	Kod	Nazwa	Marka	Kategoria / podkategoria	Cena katalogowa
surowe	„P0002”	Słuchawki E185	„ Fyro”	Audio / Słuchawki	409,00 zł
po	P0002	Fyro Słuchawki E185	Fyro	Audio | Słuchawki	409,00
surowe	„P0O17”	Soundbar E729	„Gaveo”	Audio / Soundbary	1019,00 zł
po	P0017	Gaveo Soundbar E729	Gaveo	Audio | Soundbary	1 019,00
surowe	„P0045”	Kamera sportowa S876	„Delto”	Aparaty i kamery / Kamery sportowe	1189,00 zł
po	P0045	Delto Kamera sportowa S876	Delto	Aparaty i kamery | Kamery sportowe	1 189,00

Spacje i małe litery w kodach, literówka P0O17 zamiast P0017, marka pisana na 35 sposobów, kategoria z podkategorią w jednej kolumnie, cena jako tekst, 3 zdublowane wiersze.

Tekst porządkujesz, zanim zaczniesz porównywać

Problem	Przekształć →	Wynik
„P0014 ” — spacja na końcu	Format → Przytnij	P0014
„p0020” — małe litery w kodzie	Format → Wielkie litery	P0020
„ALVERO”, „alvero”, „ Alvero”	Przytnij, potem Zamień pierwszą literę każdego wyrazu na wielką	Alvero
„Audio / Słuchawki”	Podziel kolumnę → Według ogranicznika „ / ”	Audio | Słuchawki
„209,00 zł”	Zamień wartości „ zł” → nic, potem typ z kulturą pl-PL	209,00
znaki niedrukowalne z systemów	Format → Wyczyść	bez znaków sterujących

Power Query porównuje tekst dokładnie: „Alvero” i „ALVERO” to dla niego dwie marki, a „P0014 ” i „P0014” — dwa kody. Stąd 35 zapisów zamiast 8 marek.

Trzy sposoby na nową kolumnę

Dodaj kolumnę: z podziału, z warunku albo z formuły. Kroki z zapytania Produkt w pliku rozwiązania.

PODZIEL KOLUMNĘ

```
#"Podzielono kategorię" = Table.SplitColumn(#"Poprawiono literówkę w kodzie", "Kategoria / podkategoria",
    Splitter.SplitTextByDelimiter(" / ", QuoteStyle.Csv), {"Kategoria", "Podkategoria"})
```

KOLUMNA WARUNKOWA

```
#"Dodano Nr klasy" = Table.AddColumn(#"Cena bez zł", "Nr klasy",
    each if [Klasa] = "Ekonomiczna" then 1 else if [Klasa] = "Standard" then 2 else 3, Int64.Type)
```

KOLUMNA NIESTANDARDOWA

```
#"Dodano Produkt" = Table.AddColumn(#"Dodano Nr klasy", "Produkt", each [Marka] & " " & [Nazwa], type text)
```

Kolumna z przykładów (Dodaj kolumnę → Kolumna z przykładów): wpisujesz dwa, trzy wyniki, a Power Query sam dobiera formułę. Sprawdź ją potem w pasku formuły.

Typ danych to decyzja, nie formalność

Kolumna	Tekst w pliku	Typ	Uwaga
Data	01.02.2023	Data, kultura pl-PL	bez kultury dzień zamieni się z miesiącem
Wartość po rabacie	2488,15	Liczba dziesiętna, pl-PL	przecinek jako separator dziesiętny
Ilość	13	Liczba całkowita	nie liczba dziesiętna — sztuk nie dzielisz
Kod produktu	P0001	Tekst	kod to nie liczba, nawet gdy ma same cyfry (np. 00123)
Cena katalogowa	209,00 zł	najpierw usuń „ zł”	potem Liczba dziesiętna, pl-PL

Błąd w komórce: Zachowaj błędy pokaże wiersze, w których konwersja się nie udała. Zamień błędy na wartość tylko wtedy, gdy wiesz, skąd się biorą.

Duplikaty usuwasz po czyszczeniu: 205 → 202 produktów

Te same produkty wpisane dwa razy — raz „Delto”, raz „DELTO”. Przed czyszczeniem wyglądają na różne.

Kod	Nazwa	Marka (surowa)	Marka po czyszczeniu
P0045	Kamera sportowa S876	„Delto”	Delto
P0045	Kamera sportowa S876	„DELTO”	Delto
P0110	Konsola S570	„Holmar”	Holmar
P0110	Konsola S570	„HOLMAR”	Holmar
P0163	Projektor E731	„Delto”	Delto
P0163	Projektor E731	„DELTO”	Delto

Kolejność: Przytnij → Zamień pierwszą literę każdego wyrazu na wielką → Usuń duplikaty (kolumna Kod). Usunięcie duplikatów całych wierszy przed czyszczeniem zostawi te pary („Delto” ≠ „DELTO”), kod nie będzie unikatowy i relacja jeden do wielu Produkt → Sprzedaż nie powstanie.

Ćw. 2 · Katalog produktów z Excela

15 min

1. Nowe źródło → Skoroszyt programu Excel → produkty.xlsx → tabela Katalog (nie arkusz Produkty).
2. Przekształć → Format. Kod: Przytnij i Wielkie litery. Marka: Przytnij i Zamień pierwszą literę każdego wyrazu na wielką.
3. Podziel „Kategoria / podkategoria” według „ / ” na Kategoria i Podkategoria.
4. Cena katalogowa: zamień „ zł” na nic, typ Liczba dziesiętna z kulturą Polski.
5. Kolumna warunkowa Nr klasy (1, 2, 3), niestandardowa Produkt = [Marka] & " " & [Nazwa]. Usuń duplikaty Kod.

PYTANIE

Ile wierszy ma Produkt po usunięciu duplikatów? Oczekiwane: 202. Dlaczego duplikaty usuwasz dopiero na końcu?

03

Łączenie: scal i dołącz

Scal dodaje kolumny, dołącz dodaje wiersze

Scalanie to sprzężenie po kluczu, jak WYSZUKAJ.PIONOWO. Dołączanie to wiersze jeden pod drugim — folder robi to dla wielu plików naraz.

SCAL ZAPYTANIA — NOWE KOLUMNY

Data	Kanał
01.01.2023	Aplikacja mobilna
01.01.2023	Sklep stacjonarny

Kanał	IdKanału
Sklep stacjonarny	1
Aplikacja mobilna	3

Data	Kanał	IdKanału
01.01.2023	Aplikacja mobilna	3
01.01.2023	Sklep stacjonarny	1

DOŁĄCZ ZAPYTANIA — NOWE WIERSZE

Plik	Wiersze
POS_2023_Q1	3 570

Plik	Wiersze
POS_2023_Q2	4 038

Plik	Wiersze
POS_2023_Q1	3 570
POS_2023_Q2	4 038

Sprzedaż ma nazwę kanału, model potrzebuje IdKanału: scalasz Sprzedaż ze słownikiem Kanał po kolumnie Kanał i rozwijasz tylko IdKanału.

Sześć rodzajów sprzężenia: które wiersze zostają

Rodzaj	Co zostaje	Kiedy
Lewe zewnętrzne	wszystkie z pierwszej, pasujące z drugiej	dopisanie kolumny ze słownika
Prawe zewnętrzne	wszystkie z drugiej, pasujące z pierwszej	rzadko — zamień kolejność tabel
Pełne zewnętrzne	wszystkie z obu	porównanie dwóch list
Wewnętrzne	tylko pasujące	część wspólna — uwaga, wiersze znikają po cichu
Lewe anty	z pierwszej, bez pary w drugiej	test jakości: kody bez słownika
Prawe anty	z drugiej, bez pary w pierwszej	produkty, które się nie sprzedały

Scal zapytania → wybierz kolumny kluczy w obu tabelach → Rodzaj sprzężenia. Scal zapytania jako nowe tworzy osobne zapytanie — dobre na testy.

Lewe anty: 1 kod bez pary, 224 wiersze sprzedaży

Sprzedaż (IdProduktu) scalona z katalogiem (Kod) sprzężeniem lewe anty — przed poprawką literówki.

Kod w sprzedaży	Kod w katalogu	Produkt	Wiersze	Sprzedaż
P0017	P0O17 (litera O)	Gaveo Soundbar E729	224	1,6 mln zł

- bez poprawki sprzedaż tego produktu (1 647 733,40 zł) trafia w raporcie do kategorii „(Puste)”
- sprzężenie wewnętrzne ukryłoby problem: te wiersze po prostu by zniknęły
- po poprawce (Zamień wartości P0O17 → P0017) lewe anty zwraca 0 wierszy

TEST W MODELU

Po odświeżeniu: 0 wierszy sprzedaży bez produktu, lokalizacji, kanału albo promocji.

Ćw. 3 · Słownik kanałów i test kodów

12 min

1. Sprzedaż: Scal zapytania → Kanał, kolumna Kanał w obu tabelach, Lewe zewnętrzne. Rozwiń tylko IdKanału.
2. Usuń tekstową kolumnę Kanał — w modelu zostaje identyfikator.
3. Scal zapytania jako nowe: Sprzedaż (IdProduktu) i Produkt (IdProduktu), rodzaj Lewe anty.
4. Znajdź kod bez pary. W Produkt: Zamień wartości P0O17 → P0017. Sprawdź, że lewe anty daje 0.
5. Zapytanie testowe: prawy przycisk → odznacz Włącz ładowanie.

PYTANIE

Ile wierszy sprzedaży zwróciło lewe anty przed poprawką? Co zobaczyłbyś w raporcie bez niej?

04

Kształt danych: anuluj przestawienie, wypełnij w dół

Anuluj przestawienie: 4 wiersze × 12 miesięcy → 48 wierszy

Plan z Excela ma miesiące w kolumnach — wygodne do czytania, złe do modelu. Model potrzebuje jednej kolumny z datą.

Kanał	2025-01	2025-02	…	Razem
Sklep stacjonarny	8 620	8 680	…	131 600
Sklep internetowy	6 860	5 820	…	90 910
Aplikacja mobilna	2 830	2 600	…	41 260
Partnerzy (marketplace)	1 910	1 460	…	24 480
Razem	20 220	18 560	…	288 250



Kanał	Miesiąc	Plan PLN
Sklep stacjonarny	2025-01	8 620 000
Sklep stacjonarny	2025-02	8 680 000
Sklep internetowy	2025-01	6 860 000
Sklep internetowy	2025-02	5 820 000
…	…	…

NAJPIERW USUŃ SUMY

Wiersz i kolumna Razem po anulowaniu przestawienia trafiłyby do danych jak zwykłe liczby: plan wyszedłby 1 153 000 000 zł zamiast 288 250 000 zł — cztery razy za dużo.

Zaznacz kolumnę Kanał → Przekształć → Anuluj przestawienie innych kolumn. Nowe miesiące w Excelu wejdą same — „inne kolumny” to wszystkie poza Kanał.

Wypełnij w dół: 48 pustych komórek w lokalizacjach

Arkusz Sklepy ma układ raportu: makroregion i województwo tylko w pierwszym wierszu grupy.

PRZED: ARKUSZ SKLEPY

Region	Województwo	Miasto
Południe	małopolskie	Kraków
		Tarnów
	śląskie	Katowice
		Gliwice
		Częstochowa
Północny zachód	wielkopolskie	Poznań

PO: PRAWY PRZYCISK → WYPEŁNIJ → W DÓŁ

Region	Województwo	Miasto
Południe	małopolskie	Kraków
Południe	małopolskie	Tarnów
Południe	śląskie	Katowice
Południe	śląskie	Gliwice
Południe	śląskie	Częstochowa
Północny zachód	wielkopolskie	Poznań

```
#"Wypełniono w dół" = Table.FillDown(#"Zmieniono typ", {"Makroregion (NUTS 1)", "Region", "Województwo"})
#"Dodano Kraj" = Table.AddColumn(#"Wypełniono w dół", "Kraj", each "Polska", type text)
```

Wypełnij w dół działa tylko na pustych komórkach (null). Pusty tekst z CSV najpierw zamień na null: Zamień wartości „” → null.

Ćw. 4 · Plan 2025 i lokalizacje

13 min

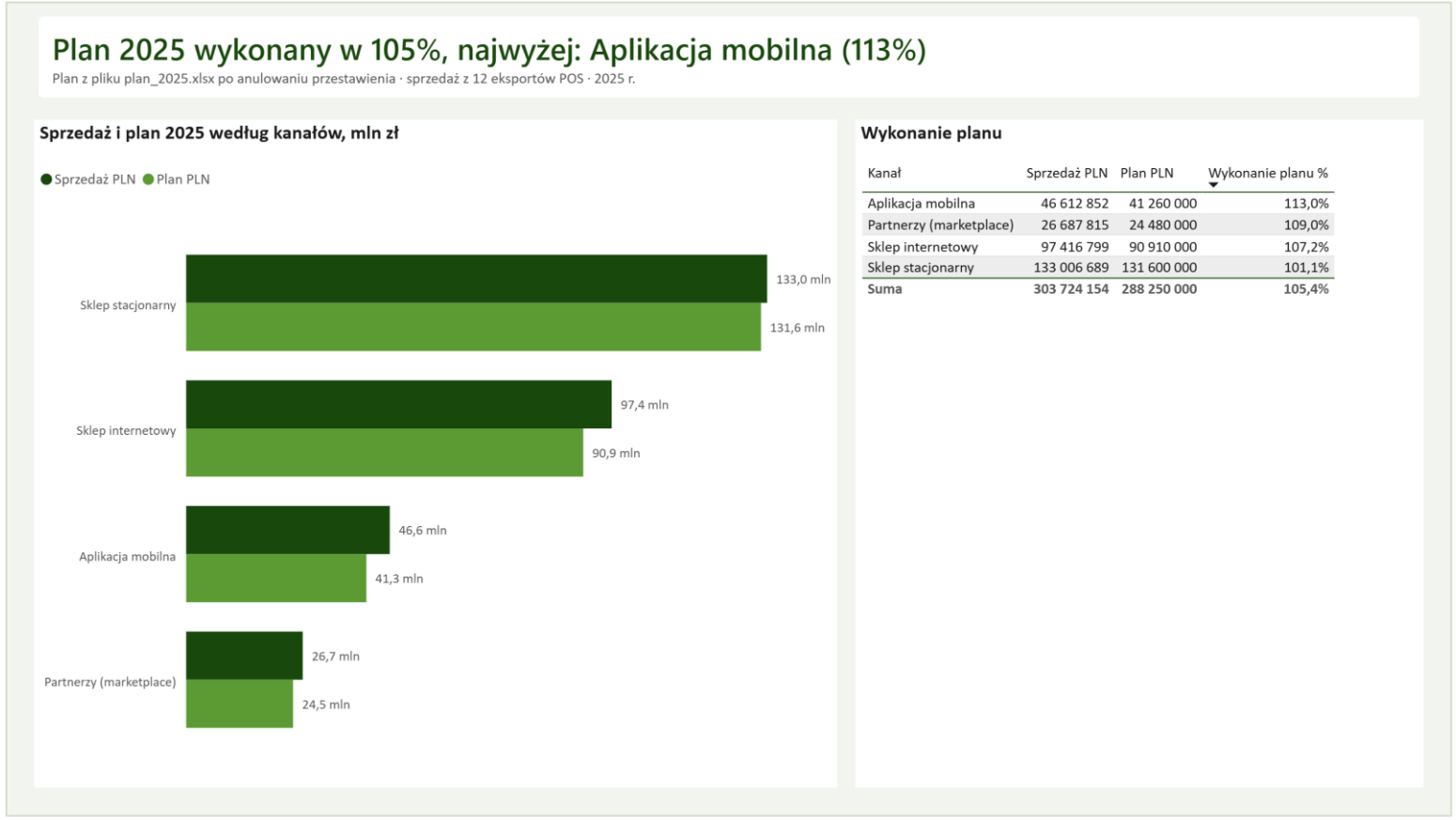
1. Geografia: zaznacz Makroregion, Region i Województwo → prawy przycisk → Wypełnij → W dół.
2. Dodaj kolumnę niestandardową Kraj = "Polska". Zmień nazwę Kod → IdLokalizacji.
3. Plan: Excel plan_2025.xlsx → tabela Plan2025. Usuń kolumnę Razem i wiersz Razem.
4. Zaznacz Kanał → Anuluj przestawienie innych kolumn. Nazwy: Miesiąc, Plan PLN. Pomnóż przez 1000.
5. Kolumna niestandardowa Data = Date.FromText([Miesiąc] & "-01"). Zamknij i zastosuj, połącz Plan z Kalendarz i Kanał.

PYTANIE

Ile wierszy ma Plan? Oczekiwane: 48. Na stronie Kontrola: ile miast bez regionu?

Rezultat: plan 2025 wykonany w 105%

Najwyżej Aplikacja mobilna (113%), najniżej Sklep stacjonarny (101%). Strona Plan 2025 z pliku rozwiązania, tytuł z miary.



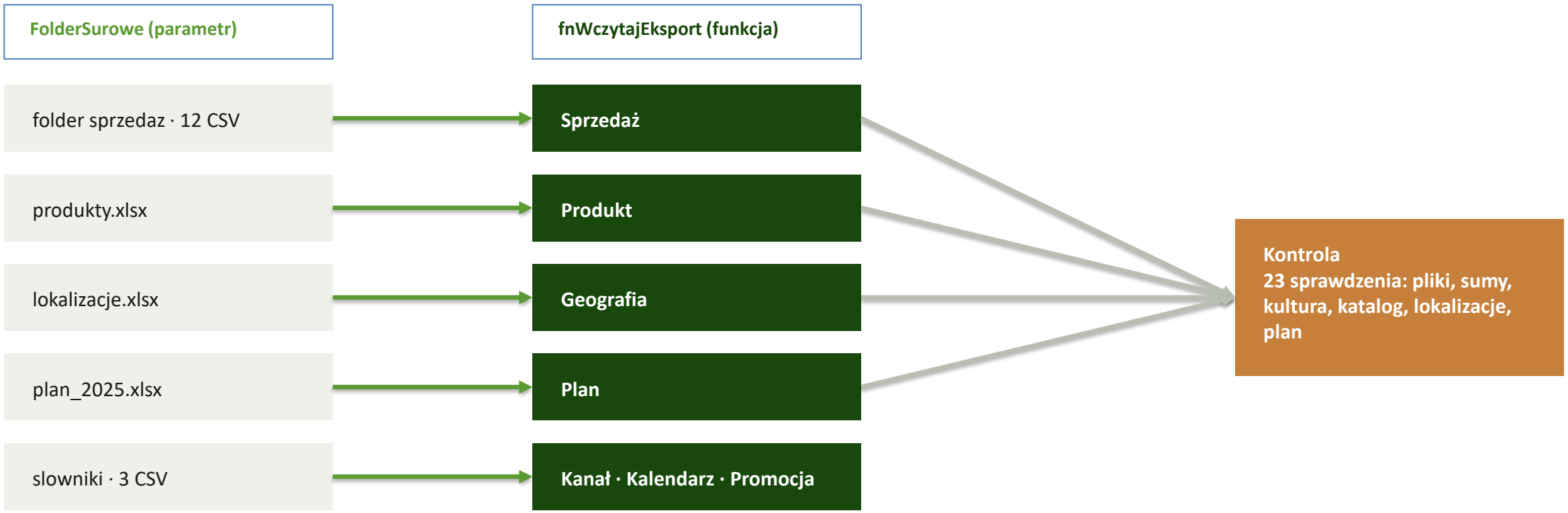
Plan 2025: 288,2 mln zł, sprzedaż 303,7 mln zł. Relacje: Plan[Data] → Kalendarz[Data], Plan[Kanał] → Kanał[Kanał].

05

Porządek, wydajność i AI

Zależności zapytań: skąd płyną dane rozwiązania

Widok → Zależności zapytań. Parametr i funkcja są wspólne; Kontrola sprawdza surowe pliki i cztery tabele modelu.



Łączysz dane z kilku źródeł? Desktop zapyta o poziomy prywatności albo zgłosi błąd Formula.Firewall. Pliki zajęć je ignorują (Opcje → Bieżący plik → Prywatność) — w usłudze Power BI to ustawienie nie działa.

Porządek w zapytaniach: nazwy, pomocnicy, funkcja

Za miesiąc otworzysz projekt i musisz zrozumieć, co zrobiłeś. Recenzent też.

- kroki nazywaj opisowo (F2): „Bez wierszy RAZEM”, a nie „Przefiltrowano wiersze1”
- zapytania pomocnicze: prawy przycisk → odznacz Włącz ładowanie — nie trafią do modelu
- grupy zapytań porządkują listę (prawy przycisk → Przenieś do grupy)
- powtarzalny kod zamień w funkcję: jedna definicja, wiele użyć
- przy trudnym kroku: Właściwości → Opis. Zobaczysz go po najechaniu na krok

FUNKCJA Z FORMUŁY ĆWICZENIA 1

```
fnWczytajEksport = (plik as binary) as table =>
    Table.PromoteHeaders(
        Table.Skip(
            Csv.Document(plik,
                [Delimiter = ";", Columns = 12,
                 Encoding = 1250]), 2))
```

W rozwiązaniu funkcję fnWczytajEksport wywołują dwa zapytania: Sprzedaż i Kontrola. Poprawka w funkcji działa w obu.

Wydajność: mniej danych wcześniej, bufor tam, gdzie trzeba

Tabela w M nie jest zapamiętana: każde użycie zapytania czyta źródło od nowa.

- usuwaj kolumny i wiersze jak najwcześniej — każdy następny krok ma mniej pracy
- pliki CSV i Excel czyta się w całości; bazy SQL wykonują część kroków same (składanie zapytań)
- przy bazie trzymaj kroki składane na początku; prawy przycisk na kroku → Wyświetl zapytanie natywne
- nie ładuj do modelu zapytań pomocniczych, których raport nie używa

LISTA SPRAWDZANA W KAŻDYM WIERSZU

```
KodyBezPary = List.Buffer(List.Difference(
    List.Distinct(Model[IdProduktu]), KodyKatalogu)),
…Table.SelectRows(Model,
    each List.Contains(KodyBezPary, [IdProduktu]))
```

ZMIERZONE NA PLIKU ROZWIĄZANIA

Bez List.Buffer lista kodów liczyła się od nowa dla każdego z 57 222 wierszy — odświeżenie nie skończyło się w 10 minut. Z buforem zapytanie Kontrola liczy się ok. 25 s.

Składanie zapytań opisuje Microsoft Learn: „Understanding query evaluation and query folding in Power Query”. Table.Buffer i List.Buffer używaj świadomie — bufor zajmuje pamięć.

Kod M z czatu AI: pięć rzeczy do sprawdzenia

Tak mógłby wyglądać kod na polecenie „połącz pliki CSV z folderu”. Działa — na komputerze autora i z błędnymi liczbami.

```
let
    Source = Folder.Files("C:\Users\Anna\Downloads\sprzedaz"),
    Combined = Table.Combine(List.Transform(Source[Content],
        each Table.PromoteHeaders(Csv.Document(_, [Delimiter = ";"])))),
    Typed = Table.TransformColumnTypes(Combined, {{"Data", type date}, {"Wartość po rabacie", type number}})
in
    Typed
```

ŚCIEŻKA

na sztywno z Downloads — zamiast parametru

PLIKI

bez filtra: kopia i plik txt wchodzą do danych

NAGŁÓWKI

brak Table.Skip: nagłówkiem zostaje opis eksportu

KODOWANIE

bez Encoding = 1250: krzaki w polskich nazwach

KULTURA

typy bez pl-PL: daty zamienione, liczby błędne

Po kodzie z AI zawsze: liczba wierszy, suma kontrolna i lewe anty ze słownikami. AI nie widzi twoich plików — zgaduje ich format.

Ćw. 5 · Porządek i kontrola (dla szybszych)

10 min

1. Widok → Zależności zapytań. Które zapytanie zależy od Kanał?
2. Zmień nazwy trzech kroków w Sprzedaży na opisowe (F2).
3. Formułę z ćw. 1 zamień w funkcję fnWczytajEksport (ściąga) i użyj jej w kolumnie Dane.
4. Odwołanie do Sprzedaż → Grupuj według Rok (z Data): liczba wierszy i suma wartości.
5. Porównaj sumy z wierszami RAZEM w plikach. Zapisz wynik w dzienniku decyzji.

PYTANIE

Które kroki trzeba zmienić, gdy dział IT doda do eksportu nową kolumnę na końcu?

Rezultat: 57 222 wiersze, te same liczby co na zajęciach 1

Strona Kontrola jakości z pliku rozwiązania. Model zbudowany z surowych plików daje co do grosza sprzedaż z pliku zajęć 1.



Dane z 12 eksportów zgadzają się z sumami kontrolnymi

Zapytanie Kontrola liczy w Power Query to, co w ćwiczeniach sprawdzasz ręcznie

57 222
Wiersze sprzedaży

850 573 496
Sprzedaż PLN 2023–2025, zł

Kontrole po każdym odświeżeniu

Obszar	Kontrola	Wartość
Pliki	Pliki w folderze sprzedaz	14
Pliki	Pliki CSV wczytane (bez kopii)	12
Pliki	Wiersze w kopii pliku (odrzucone)	4 311
Pliki	Sprzedaż w kopii pliku (odrzucona), zł	58 161 559,27
Sumy kontrolne	Wiersze sprzedaży w modelu	57 222
Sumy kontrolne	Suma z wierszy RAZEM, zł	850 573 496,17
Sumy kontrolne	Suma w modelu, zł	850 573 496,17
Sumy kontrolne	Różnica, zł	0
Kultura	Dni w eksportach	1 096
Kultura	Dni, których en-US nie odczyta	664
Kultura	Dni, które en-US odczyta jako inną datę	396
Produkty	Wiersze katalogu (surowe)	205
Produkty	Produkty po usunięciu duplikatów	202
Produkty	Zapisy marki w katalogu (surowe)	35
Produkty	Marki po oczyszczeniu	8
Produkty	Kody ze sprzedaży bez pary w surowym katalogu	1
Produkty	Wiersze sprzedaży z tymi kodami	224
Lokalizacje	Puste komórki przed wypełnieniem w dół	48
Lokalizacje	Miasta w modelu	26
Plan	Wiersze planu po anulowaniu przestawienia	48
Kultura	Tekst 2488,15 odczytany w en-US	248 815
Kultura	Miesiąc z tekstu 01.02.2023 w en-US	1
Kultura	Miesiąc z tekstu 01.02.2023 w pl-PL	2

Rok	Wiersze	Sprzedaż
2023	18 095	284,8 mln zł
2024	19 222	262,0 mln zł
2025	19 905	303,7 mln zł

- 284,8 · 262,0 · 303,7 mln zł — jak w pliku z zajęć 1
- 202 produktów, 26 miast, 48 wierszy planu

Power Query w projekcie zaliczeniowym

- wczytaj dane projektu przez Power Query: parametr ścieżki, typy z kulturą, opisowe nazwy kroków
- w dzienniku decyzji zapisz, co czyściłeś i jak to sprawdziłeś — sumą kontrolną albo lewym anty
- CP1 przed zajęciami 3: karta projektu z opisem źródeł danych i ich problemów
- na obronie mogę wskazać krok w Power Query i zapytać, co by się stało bez niego

Kryterium „Model danych” (15 pkt) obejmuje czytelne nazwy i poprawne relacje — obie rzeczy zaczynają się w Power Query.

Na następne zajęcia

- Wczytaj dane swojego projektu przez Power Query i zapisz trzy problemy, które naprawiłeś
- Sprawdź typy i kulturę każdej kolumny z datą i kwotą
- Zajęcia 3: modelowanie — schemat gwiazdy, relacje, tabela dat. Kartę projektu (CP1) oddajesz przed nimi
- Przeczytaj: G. Raviv, Power Query w Excelu i Power BI — o łączeniu danych z wielu źródeł i anulowaniu przestawienia

Ćwiczenie 5 możesz dokończyć w domu na pliku ćwiczeniowym.

Trzy rzeczy do zapamiętania

pl-PL

typy dat i liczb zawsze z ustawieniami regionalnymi

filtr plików

najpierw lista plików, potem łączenie zawartości

kontrola

suma kontrolna i lewe anty po każdej zmianie

DZIĘKUJĘ

dr inż. Szczepan Górtowski

Katedra Informatyki Ekonomicznej · Uniwersytet Ekonomiczny w Poznaniu

szczepan.gortowski@ue.poznan.pl · dyżury: kie.ue.poznan.pl/pl/dyzury/