

Analiza sprzedaży w kawiarni

Przygotuj dane i przeanalizuj sprzedaż w kawiarni. Rozwiązania dwóch pierwszych zadań znajdziesz w artykule i plikach SQL. Trzecie wykonasz samodzielnie, korzystając ze wskazówek i szkieletu zapytania.

Wszystkie dane, daty i zdarzenia są fikcyjne i służą wyłącznie do nauki. Ćwiczenie obejmuje jedną tabelę ze sprzedażą, zwrotami i anulowaniami.

Wybierz pliki dla swojego silnika SQL

Silnik SQL	Skrypt przygotowania danych	Rozwiązania i szkielet zadania 3
SQL Server	01_setup_sql_server.sql	02_rozwiazanie_sql_server.sql
PostgreSQL	01_setup_postgresql.sql	02_rozwiazanie_postgresql.sql
MySQL 8.0+	01_setup_mysql.sql	02_rozwiazanie_mysql.sql
SQLite	01_setup_sqlite.sql	02_rozwiazanie_sqlite.sql

Jak zacząć?

Otwórz edytor SQL połączony z bazą do ćwiczeń. Uruchom cały plik 01_setup dla swojego silnika. Utworzy tabelę tymczasową i wstawi dane bez importowania CSV. Kontrola powinna pokazać liczbę 24.

W MySQL wybierz istniejącą bazę, np. poleceniem `USE nazwa_bazy`. Konto musi mieć uprawnienie `CREATE TEMPORARY TABLES`. W SQLite otwórz bazę lub połączenie w pamięci.

W tym samym połączeniu wykonuj zapytania z artykułu lub pliku rozwiązania. SQL Server używa tabeli `#cafe_sales`, a pozostałe silniki `cafe_sales`. Ponowne uruchomienie skryptu zastępuje tylko tę tabelę tymczasową. Skrypt pracuje w istniejącej bazie i zachowuje stałe tabele.

Zachowaj to samo połączenie

Tabela tymczasowa znika po zamknięciu połączenia. Zachowaj otwartą sesję do ostatniego zapytania. Sprawdź, czy edytor używa tego samego połączenia: ta sama zakładka tego nie gwarantuje.

Jeśli korzystasz z CSV

`sprzedaz_kawiarni.csv` jest taki sam dla wszystkich silników. Plik ma kodowanie UTF-8, separator przecinek, kropkę dziesiętną i daty `YYYY-MM-DD`. Pusta wartość `original_event_id` oznacza `NULL`.

W SQLite przelicz `amount_pln` z CSV na grosze, zanim wstawiś kwoty do `amount_grosze`. W skrypcie przygotowania danych kwoty są już zapisane w groszach.

Co właściwie analizujesz?

Jeden wiersz opisuje zdarzenie dotyczące jednej kategorii sprzedaży. Nie oznacza jednego klienta ani całego paragonu.

Kolumna	Znaczenie
event_id	Identyfikator zdarzenia.
event_date	Dzień, w którym nastąpiło zdarzenie.
category	Kategoria: kawa, desery lub kanapki.
channel	Kanał sprzedaży: lokal albo na_wynos.
status	Rodzaj zdarzenia: sprzedaż, zwrot lub anulowanie.
amount_pln	Dodatnia kwota w PLN. W SQLite używamy amount_grosze.
original_event_id	Identyfikator sprzedaży, której dotyczy zwrot. Dla innych zdarzeń jest pusty.

Co oznaczają statusy?

completed to zrealizowana sprzedaż. refunded to osobne zdarzenie zwrotu. cancelled to anulowanie, które nie przyniosło przychodu. Status opisuje rodzaj zdarzenia, a nie aktualny stan zamówienia.

Kwota zwrotu jest dodatnia; odejmujesz ją dopiero w obliczeniu. original_event_id łączy zwrot z wcześniejszą sprzedażą completed o tej samej kwocie, kategorii i kanale. Sprzedaż i zwrot pozostają osobnymi zdarzeniami i mieszczą się w tym samym okresie.

Jak liczymy sprzedaż?

Sprzedaż po zwrotach = suma kwot completed minus suma kwot refunded. Anulowania pomijamy. Nie sumuj wszystkich kwot bez uwzględnienia statusu.

Nie obliczamy sprzedaży bez VAT. Dane nie zawierają informacji o podatkach ani kosztach, więc wynik nie jest zyskiem i nie pozwala ocenić rentowności.

Kwoty w różnych silnikach SQL

SQL Server, PostgreSQL i MySQL przechowują kwoty w amount_pln jako DECIMAL(10,2). SQLite używa amount_grosze, czyli liczb całkowitych: 18,00 zł to 1800 groszy. Dzięki temu sumowanie jest dokładne.

W SQLite najpierw zsumuj grosze. Dopiero wynik podziel przez 100.0 i użyj ROUND(..., 2), żeby otrzymać kwotę w PLN. SQLite może pokazać 178.0 zamiast 178.00 - to ta sama kwota.

Od pytania do wyniku

01 / Porównaj kategorie

Która kategoria przyniosła najwyższą wartość sprzedaży po zwrotach w całym zestawie danych? Zwróć kategorię i kwotę. Posortuj wyniki od najwyższej do najniższej wartości.

Najpierw spróbuj samodzielnie. Następnie przejdź przez wyjaśnienie w artykule lub porównaj swoje zapytanie z plikiem 02_rozwiazanie dla wybranego silnika SQL.

02 / Porównaj kanały

Który kanał przyniósł najwyższą wartość sprzedaży po zwrotach w całym zestawie danych? Zastosuj tę samą zasadę uwzględniania statusów. Pokaż oba kanały i ich kwoty, uporządkowane malejąco.

Najpierw spróbuj zmienić zapytanie z zadania 1. Następnie porównaj je z rozwiązaniem w artykule lub pliku SQL. Sprzedaż na miejscu (lokal) powinna wynieść 277,00 PLN, a sprzedaż na wynos (na_wynos) 213,50 PLN.

03 / Porównaj udział zwrotów

Która kategoria ma najwyższy udział wartości zwrotów w wartości zrealizowanej sprzedaży? Wykonaj tę analizę samodzielnie. W pliku SQL znajdziesz szkielet zapytania do uzupełnienia.

Dla każdej kategorii policz osobno sumę kwot completed i sumę kwot refunded. Podziel wartość zwrotów przez wartość zrealizowanej sprzedaży i pomnóż przez 100. Otrzymasz udział w procentach. Anulowania pomini. Porównujesz wartości w PLN, a nie liczbę transakcji.

Zwróć kolumny category, completed_sales_pln, refunds_pln i refund_share_pct. Zaokrąglaj udział do dwóch miejsc po przecinku i uporządkuj kategorie od najwyższego udziału zwrotów.

Sprawdź wynik

Zestaw zawiera 24 zdarzenia: 18 sprzedaży, 3 zwroty i 3 anulowania. Łączna wartość sprzedaży po zwrotach wynosi 490,50 PLN. Suma wyników dla kategorii oraz suma wyników dla kanałów powinny dawać tę samą kwotę.