

Proces modelowanie danych

Modelowanie danych to proces analizowania oraz definiowania wszystkich danych zbieranych i tworzonych w firmie, a także relacji między tymi danymi. W procesie modelowania danych jest tworzona graficzna reprezentacja danych używanych w firmie, a sam proces ułatwia zrozumienie i objaśnianie wymagań dotyczących tych danych.

Dzięki modelowaniu danych organizacja:

- Tworzy strukturę współpracy między zespołem IT i zespołami biznesowymi.
- Uwidacznia możliwości udoskonalania procesów biznesowych dzięki zdefiniowaniu potrzeb i zastosowań dotyczących danych.
- Umożliwia zaoszczędzenie czasu i pieniędzy na inwestycje związane z przetwarzaniem i infrastrukturą IT dzięki odpowiedniemu planowaniu z wyprzedzeniem.
- Zmniejsza liczbę błędów (i nadmiarowych, podatnych na błędy wpisów danych), zwiększając równocześnie integralność danych.
- Zwiększa szybkość i wydajność pobierania danych oraz ich analityki dzięki planowaniu pojemności i rozwoju.

Cardinality

Cardinality złączenia między dwiema tabelami to numeryczna relacja między wierszami jednej tabeli i wierszami drugiej.

Rodzaje cardinality to

- jeden-do-jednego
- jeden-do-wielu
- wiele-do-wielu.

Normalizacja i denormalizacja

Normalizacja- to proces tworzenia zestawu schematów do przechowywania niepotrzebnych i spójnych danych. Przeprowadzana w celu zminimalizowania nadmiarowości i niespójności obecnych w danych w relacyjnych bazach danych. Ten proces dzieli głównie duże tabele na mniejsze tabele z mniejszą liczbą redundancji (zwane „formami normalnymi”).

Denormalizacja-jest procesem łączenia danych, dzięki czemu można je szybko sprawdzić. Denormalizacja polega na dodawaniu nadmiarowych danych lub grupowaniu danych w celu optymalizacji wydajności.

Co to jest Datamart?

Datamart to struktura / wzorzec dostępu specyficzny dla środowisk hurtowni danych, używany do pobierania danych skierowanych do klienta. Jest to prosta forma hurtowni danych skoncentrowana na jednym temacie lub branży.

Datamarty oferują następujące korzyści:

- Użytkownicy samoobsługi mogą z łatwością przeprowadzać analizy relacyjnych baz danych, bez konieczności posiadania administratora bazy danych
- Datamarty zapewniają kompleksowe pozyskiwanie, przygotowywanie i eksplorację danych za pomocą SQL, w tym doświadczenia bez kodu
- Umożliwiają budowanie modeli semantycznych i raportów w ramach jednego holistycznego środowiska

Co to jest kostka OLAP (OLAP CUBE, Dax)?

Kostka OLAP – struktura danych, która pozwala na szybką analizę danych. Przechowuje ona dane w sposób bardziej przypominający wielowymiarowe arkusze kalkulacyjne niż tradycyjną, relacyjną bazę danych.

Struktura OLAP nie gromadzi danych w tabelach (jak standardowa, relacyjna baza danych), a w obiektach wielowymiarowych - kostkach. Przy czym trójwymiarowa "kostka" jest tu jedynie terminem umownym - w praktyce wymiarów OLAP może być więcej (mówimy wówczas o tzw. hiperkostce).

