Zajęcia 14

**Zadanie 1**

Modelowanie danych pod hurtownię danych.

**Proces modelowanie danych** polega na projektowaniu struktury danych i relacji między nimi, uwzględniając wymagania biznesowe. Obejmuje analizę potrzeb, projektowanie modeli danych oraz implementację i ładowanie danych do hurtowni. Celem tego procesu jest stworzenie optymalnej struktury danych umożliwiającej efektywną analizę i raportowanie.

**Cardinality (moc zbioru)** oznacza liczbę unikalnych wartości w danym zbiorze danych lub relacji między nimi. Jest to miara ilościowa, która informuje o rozmiarze i złożoności danych.

**Normalizacja** to proces organizacji danych w celu minimalizacji redundancji i zależności funkcyjnych poprzez podział danych na różne tabele i ustanowienie relacji między nimi. Celem jest zapewnienie spójności i integralności danych, kosztem zwiększonej złożoności zapytań. **Denormalizacja** to odwrotne podejście, polegające na łączeniu danych z różnych tabel w celu zwiększenia wydajności zapytań kosztem redundancji danych.

**Data mart** to zoptymalizowany pod kątem konkretnej dziedziny lub zastosowania podzbiór danych z hurtowni danych, dostarczający gotowych zestawień i analiz danych dla użytkowników. Ułatwia szybki dostęp i zrozumienie informacji biznesowych.

**Lakehouse** to hybrydowa architektura danych łącząca cechy hurtowni danych i data lake. Zapewnia przechowywanie nieprzetworzonych danych w oryginalnym formacie, jednocześnie umożliwiając elastyczne analizy i skalowalność. Jest bardziej wszechstronne i łatwiejsze w zarządzaniu niż tradycyjna **hurtownia danych**.

**Zadanie 2**

**Kostka OLAP** (ang. OLAP cube) – struktura danych, która pozwala na szybką analizę danych. Przechowuje ona dane w sposób bardziej przypominający wielowymiarowe arkusze kalkulacyjne niż tradycyjną, relacyjną bazę danych. Można ją również zdefiniować jako zdolność manipulowania i analizowania danych z różnych punktów widzenia. Rozmieszczenie danych w kostkach pokonuje ograniczenia relacyjnych baz danych.

*źródło:* [*https://pl.wikipedia.org/wiki/Kostka\_OLAP*](https://pl.wikipedia.org/wiki/Kostka_OLAP)

Kostka OLAP pozwala na zapisanie wielowymiarowych danych np. o produkcie, czas, miasta, budżet. Można dokonywać na niej różnych operacji takich jak: selekcja, wycinanie, obracanie, zwijanie itd.