# Zajęcia 14

**Zadanie 1**

Modelowanie danych pod hurtownię danych. Napisz notatkę co to jest:

**Proces modelowanie danych** - jest kluczowym etapem projektowania hurtowni danych i polega na definiowaniu struktury danych, relacji między nimi oraz regułami, które determinują, jak dane są przechowywane i manipulowane w systemie hurtowni danych.

**Cardinality** – odnosi się do relacji między dwoma encjami w modelu danych. Określa, czy relacja między dwiema encjami jest jeden-do-jednego, jeden-do-wielu, wiele-do-jednego czy wiele-do-wielu.

**Normalizacja i denormalizacja** - **normalizacja** jest procesem projektowania struktury bazy danych w taki sposób, aby uniknąć redundancji danych i zachować integralność danych. Celem normalizacji jest podzielenie danych na mniejsze i bardziej znormalizowane tabele, aby zmniejszyć duplikację informacji i zwiększyć spójność danych. Natomiast **denormalizacja** jest procesem przeciwnym do normalizacji. Polega na celowym wprowadzeniu redundancji danych do struktury bazy danych w celu poprawy wydajności odczytu danych. Poprzez denormalizację można zredukować liczbę operacji łączenia tabel w złożonych zapytaniach, co przyspiesza wykonywanie zapytań i zwiększa wydajność systemu.

**Co to jest Datamart** - to specjalnie zaprojektowana i zoptymalizowana struktura danych, która służy do przechowywania i udostępniania określonego zestawu danych biznesowych, które są skoncentrowane na określonym obszarze lub działalności organizacji.

**Co to jest Lakehouse i jak różni się od Hurtowni** - to pojęcie związane z architekturą danych, które łączy cechy hurtowni danych i data lake w jednym rozwiązaniu. Jest to hybrydowa koncepcja, która ma na celu połączenie elastyczności i skalowalności data lake z kontrolą i strukturą hurtowni danych.

Różnice:

-przechowywanie danych: lakehouse przechowuje dane w formie surowej, podczas gdy hurtownia danych przechowuje dane w przetworzonej i uporządkowanej postaci.

-Zarządzanie metadanymi: lakehouse oferuje bardziej scentralizowane i kontrolowane zarządzanie metadanymi, w przypadku hurtowni danych zarządzanie metadanymi może być mniej formalne i elastyczne.

-Elastyczność przetwarzania danych: lakehouse umożliwia elastyczne przetwarzanie danych na różnych etapach, od surowych danych do przetworzonych danych gotowych do analizy. Hurtownia danych ma zazwyczaj bardziej statyczne i zdefiniowane przetwarzanie danych.

-Skalowalność: lakehouse oferuje skalowalność, umożliwiając przechowywanie i przetwarzanie dużych ilości danych. Hurtownia danych może być ograniczona w skalowalności ze względu na jej scentralizowaną strukturę.

**Zadanie 2**

Znajdź informację i napisz krótką notatkę co to jest kostka OLAP (OLAP CUBE, Dax).

kostka OLAP (**OLAP Cube**) to wielowymiarowa struktura danych, która umożliwia efektywną analizę i eksplorację danych z różnych perspektyw.

**DAX** (Data Analysis Expressions) jest językiem zapytań używanym w technologiach Microsoft do manipulowania i analizy danych w kostkach OLAP.