**Modelowanie danych** - proces tworzenia schematów, definiowania zawartości danych zebranych w hurtowni danych. Tworzony jest model koncepcyjny, potem logiczny i na końcu model fizyczny.

**Cardinality** to częstotliwość występowania danych.

**High cardinality** - wartości w kolumnie danych są bardzo rzadkie

**Normal cardinality** - wartości w kolumnie danych są rzadkie, ale nie niepowtarzalne

**Low cardinality** - wartości w kolumnie danych są bardzo typowe

**Normalizacja** - uporządkowanie danych, usunięcie nadmiarowości i zwiększenie integralności bazy danych. Zmniejsza ilość pamięci potrzebnej do przechowywania danych.

Jest to wieloetapowy proces, przekształca dane do formy tabelarycznej i usuwa zduplikowane rekordy z tabel relacyjnych.

**Denormalizacja** - wprowadzenie kontrolowanej nadmierności do bazy w celu przyspieszenia wykonywania na niej operacji, dzięki której unika się kosztownych operacji połączeń tabel, jednak dane zajmują więcej miejsca na dysku i są bardziej podatne na błędy przy insertach.

**Datamart** - jest to tematyczna hurtownia danych, która jest używana przez jeden dział firmy i dane w niej zawarte dotyczą jednego konkretnego tematu biznesowego. Tworzone w celu zapewnienia wsparcia i przyspieszenia procesu podejmowania decyzji osobom odpowiedzialnym za konkretny obszar biznesowy.

Online Analytical Processing (**OLAP**) struktur danych, który pozwala pokonać ograniczenia relacyjnych baz danych, umożliwiając szybką analizę danych. Nie jest zoptymalizowany pod kątem przeprowadzania transakcji, a pod wykonywanie większych zapytań i analiz.

Dane gromadzone w ten sposób są dzielone w zależności od potrzeb w kostkach. Podejście to zauważalnie ułatwia tworzenie raportów.