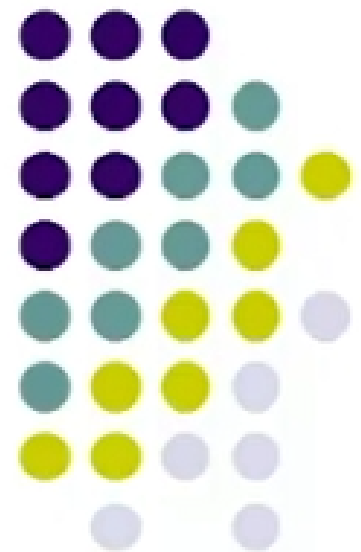


## Лекція 3

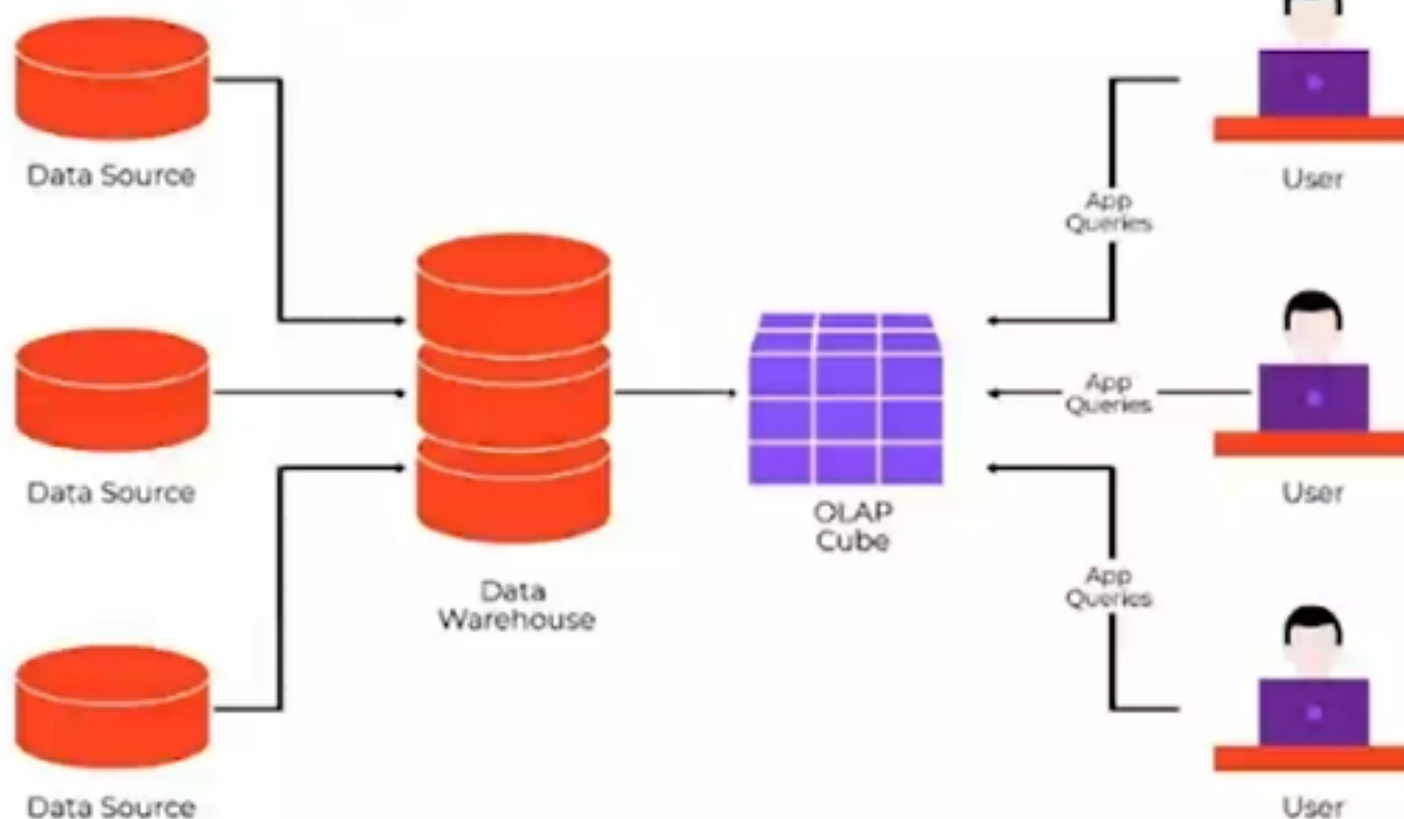
# OLAP

# On-Line Analytical Processing

---



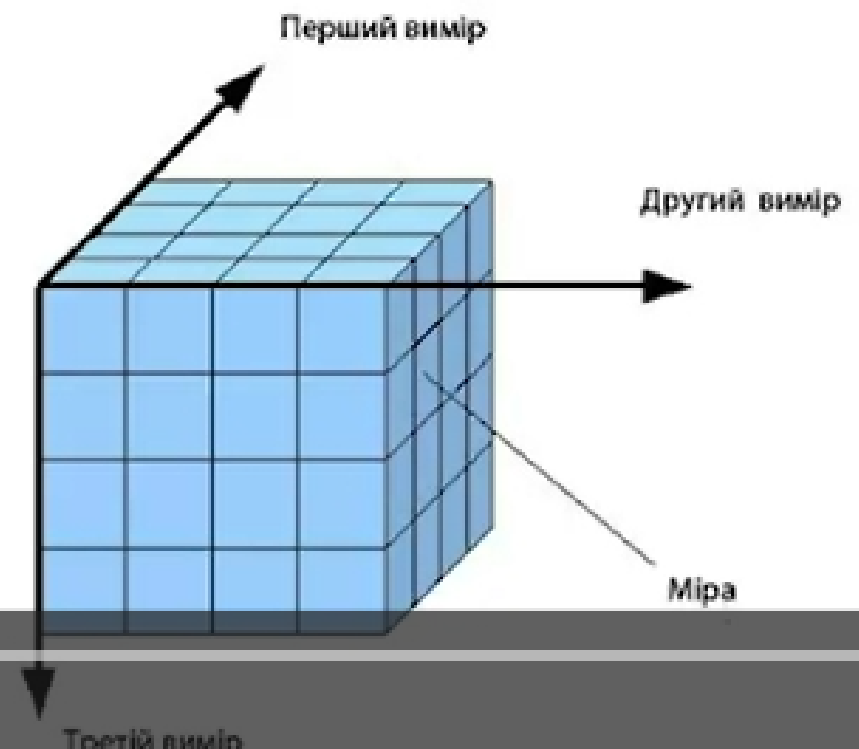
# Як працює OLAP?



- Дані спочатку витягуються з різних джерел даних та у різних форматах (текстові файли, електронні таблиці).
- Потім ці дані зберігаються у сховищі даних.
- Далі дані очищуються, трансформуються та зберігаються в кубах OLAP.
- Потрапляючи в куби OLAP, інформація попередньо калькулюється та агрегується для подальшого аналізу.
- Нарешті, користувач отримує дані з кубів OLAP, подаючи запити до них.

# Концепція OLAP

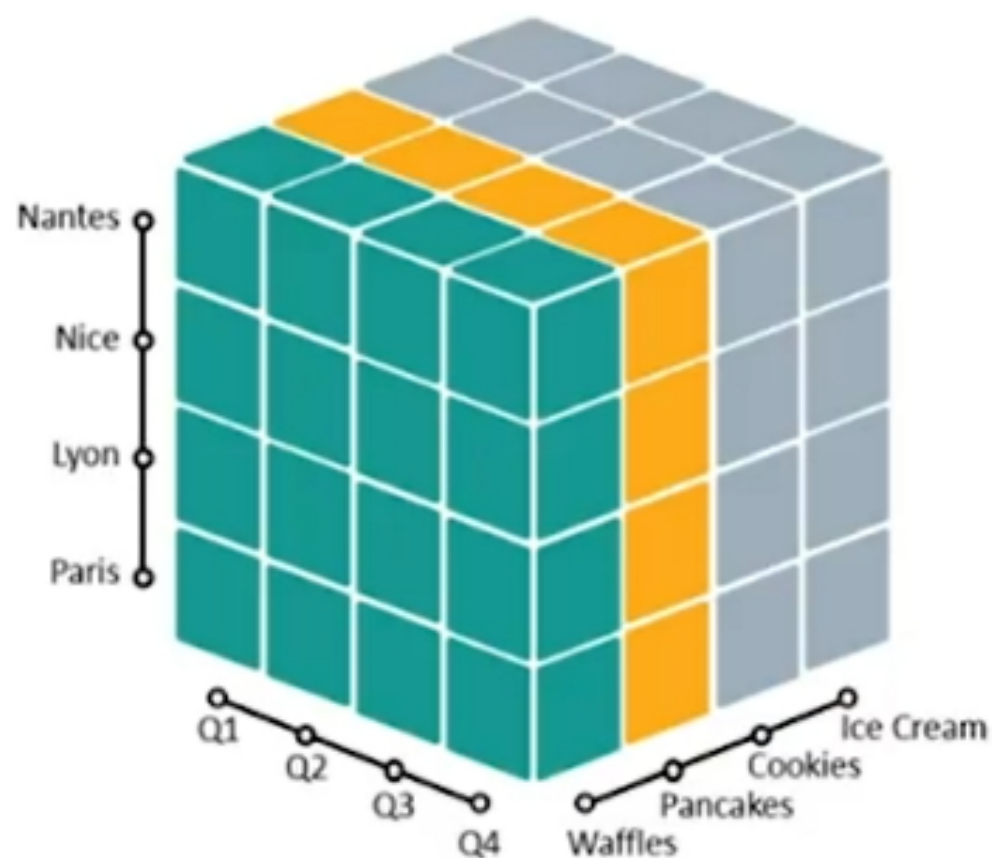
- Olap – on-line analytical processing, аналітична обробка в реальному часі – технологія оперативної аналітичної обробки даних, що використовує методи для збору інформації, збереження та аналізу багатовимірних даних з метою підтримки процесу прийняття рішень.
- Multi-dimensional conceptual view



# Приклад OLAP куба



3-dimensional Cube in Data Warehousing Example  
with dimensions Location, Quarter, Product categories



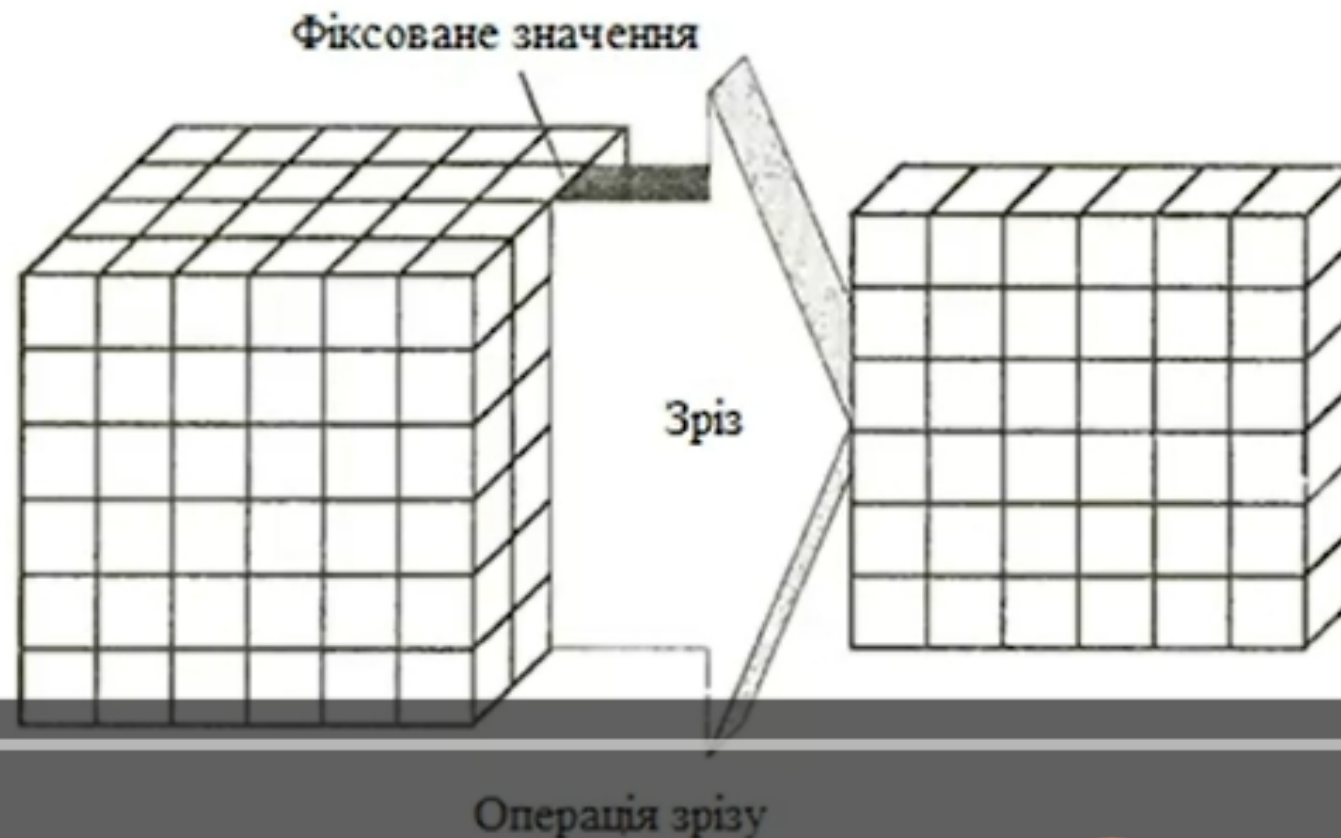
Example: You want to summarize sales by specific product, by the time, and by store location. These are data cube dimensions.



# Операції над гіперкубом



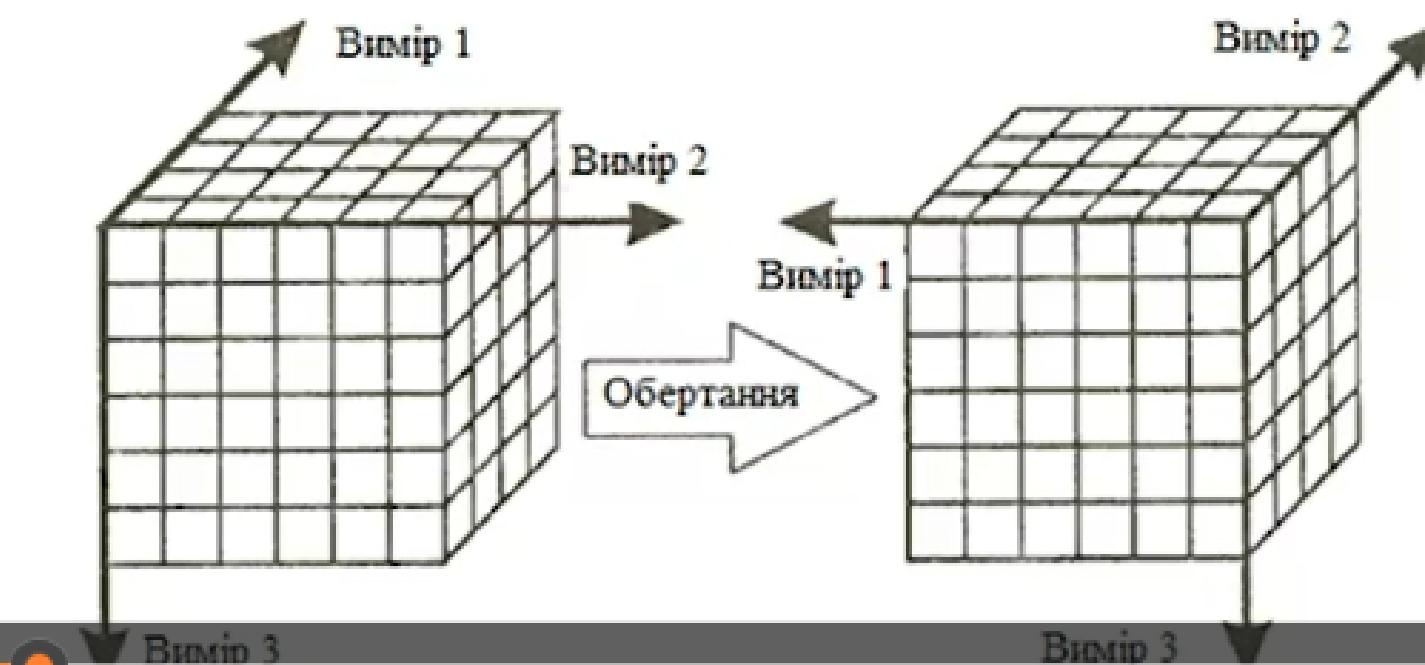
**Зріз** (*Slice*) - формується підмножина багатовимірного масиву даних.



# Операції над гіперкубом



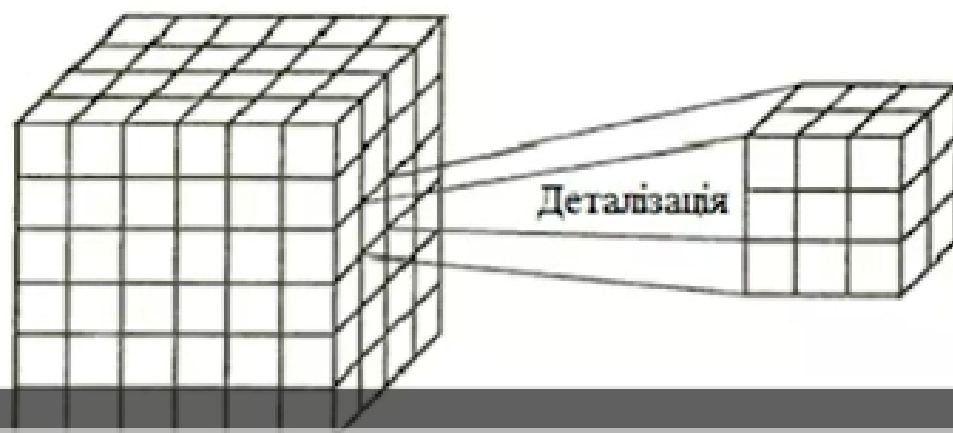
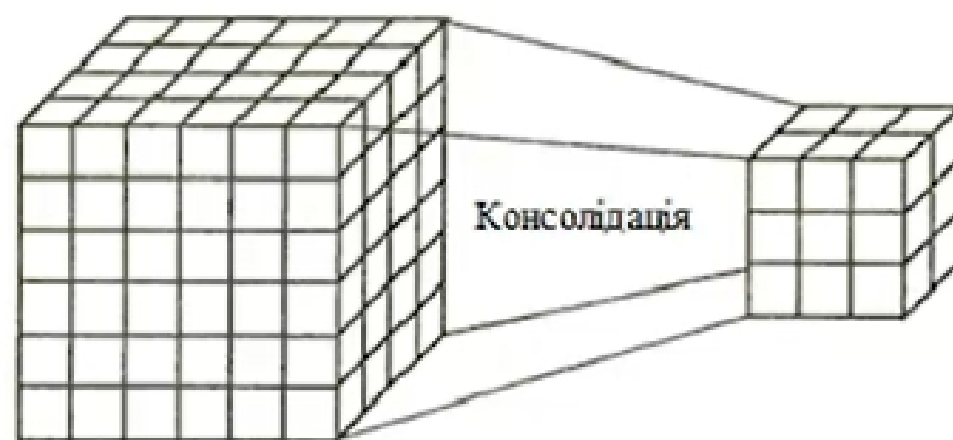
**Обертання** (**Rotate**) - зміна розташування вимірів, представлених у звіті або на сторінці, що відображається.





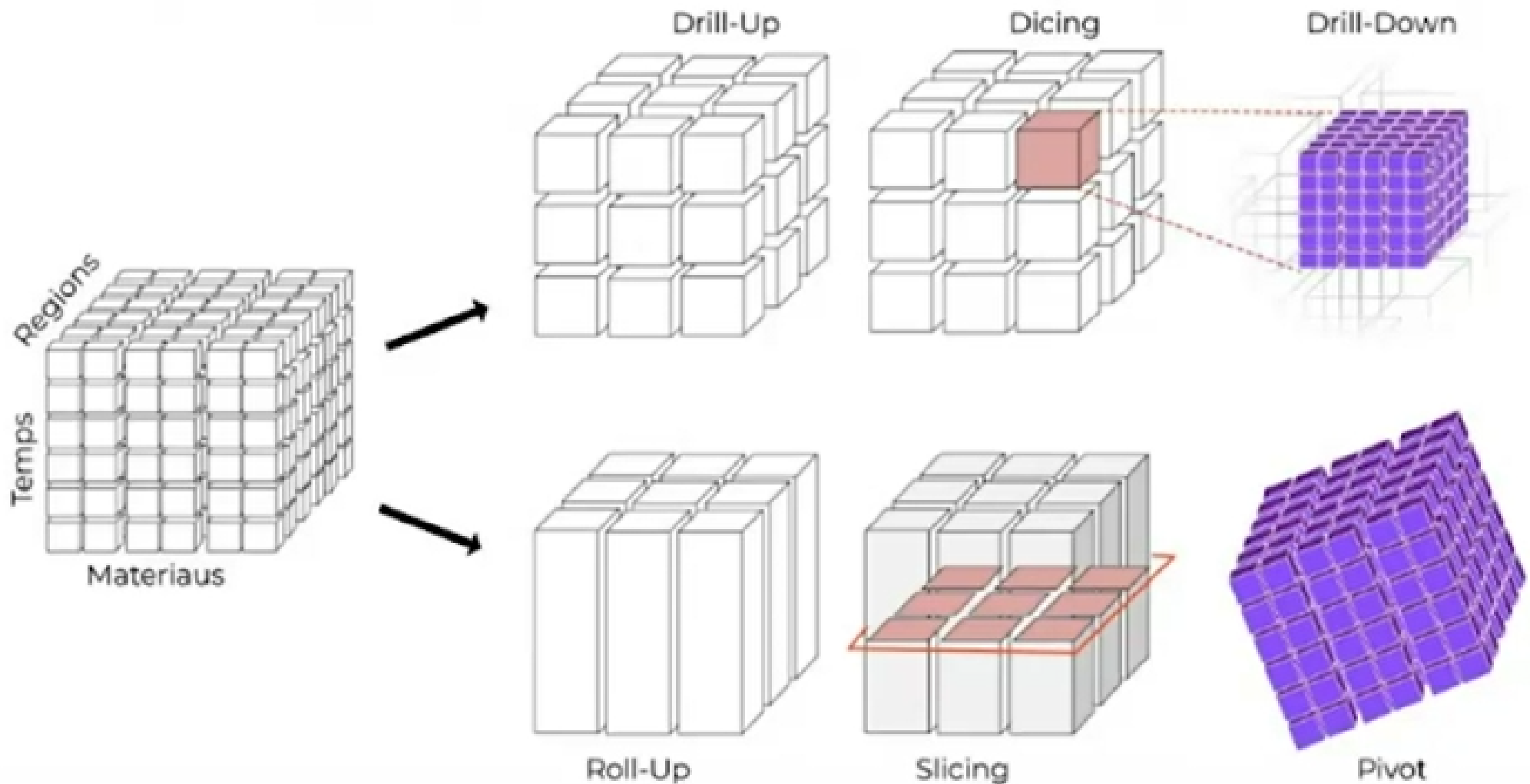
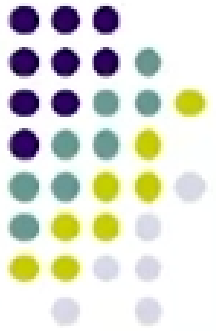
# Операції над гіперкубом

**Консолідація (Drill Up)**  
**і деталізація (Drill Down)** - операції, які визначають перехід вгору за напрямом від детального (down) представлення даних до агрегованого (up) і навпаки, відповідно.





# Операції над гіперкубом





# Визначення OLAP-систем



- OLAP (On-Line Analytical Processing) – це технологія оперативної аналітичної обробки даних, що використовує методи і засоби для збору, зберігання і аналізу багатовимірних даних з метою підтримки процесів прийняття рішень.
- Основне призначення OLAP-систем – підтримка аналітичної діяльності, довільних (часто використовується термін **ad-hoc**) запитів користувачів-аналітиків. Мета OLAP-аналізу – перевірка гіпотез, що виникають при аналізі предметної області.
- Біля витоків технології OLAP стоїть основоположник реляційного підходу Кодд. У 1993 р. він опублікував статтю під назвою "OLAP для користувачів-аналітиків: яким він повинен бути". У даній праці викладено основні концепції оперативної аналітичної обробки та визначено 12 вимог, яким повинні задовольняти продукти для ефективної оперативної аналітичної обробки даних. Згодом він доповнив їх ще 6 вимогами та розділив їх на 4 групи.

# Особливості OLAP-систем



## Основні

- Багатовимірність
- Інтуїтивне маніпулювання даними
- Доступність
- Пакетне видобування замість інтерпретації
- Підтримка всіх моделей OLAP аналізу
- Архітектура клієнт-сервер
- Прозорість
- Підтримка багатокористувацького режиму

## Спеціальні

- Обробка ненормалізованих даних
- Збереження результатів окрема від вхідних даних
- Виключення відсутніх значень
- Обробка відсутніх даних

## Подання звітів

- Гнучкість формування звітів
- Постійна продуктивність звітів
- Автоматичне налаштування фізичного рівня

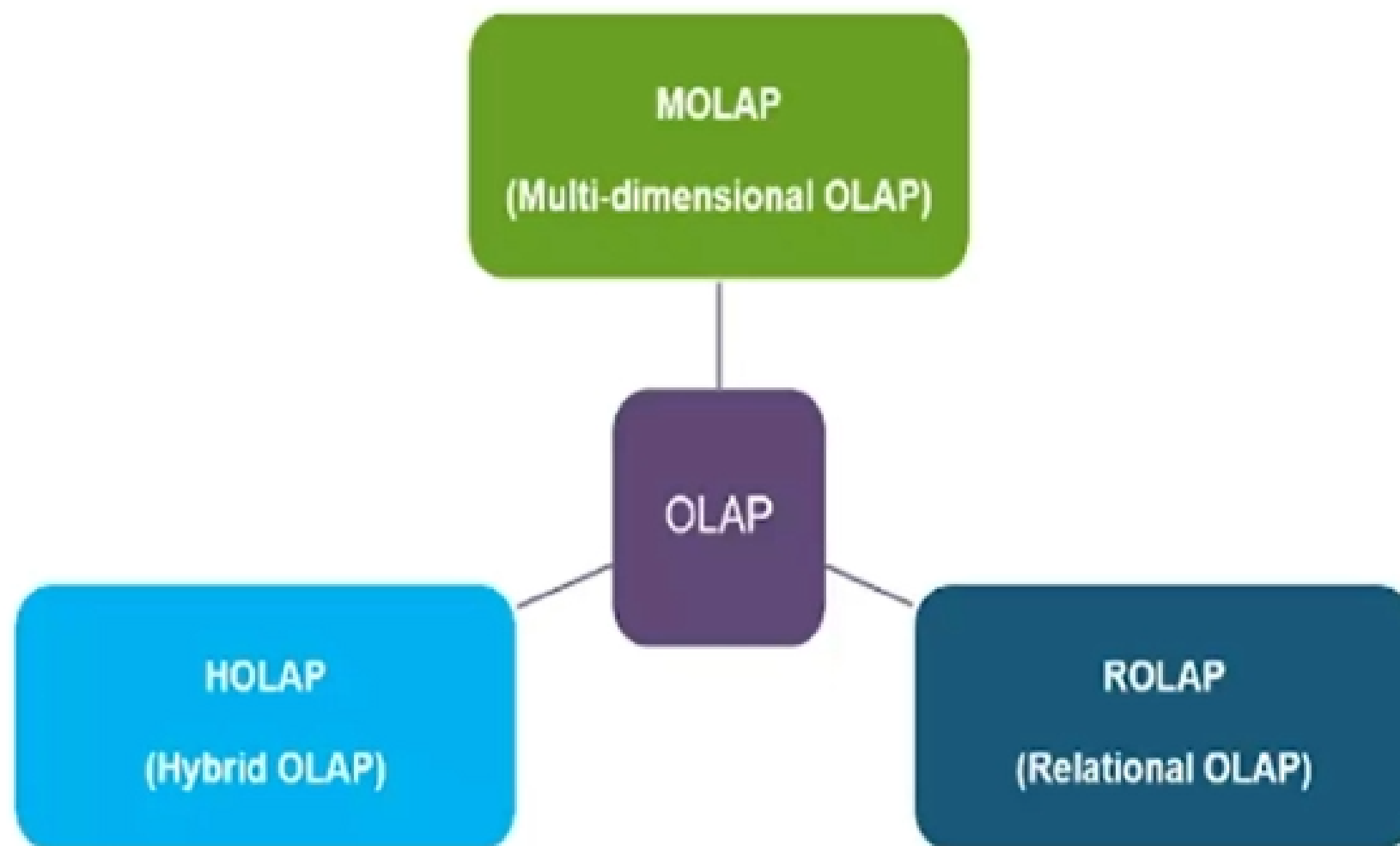
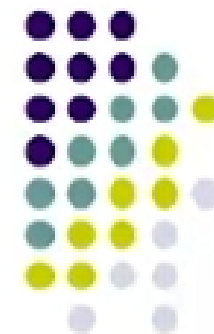
## Управління вимірами

- Універсальність вимірів
- Необмежене число вимірів і рівнів агрегації
- Необмежені операції між розмірностями

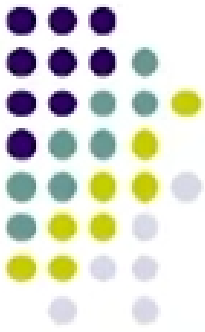
# Тест FASMI



# Архітектура OLAP - систем



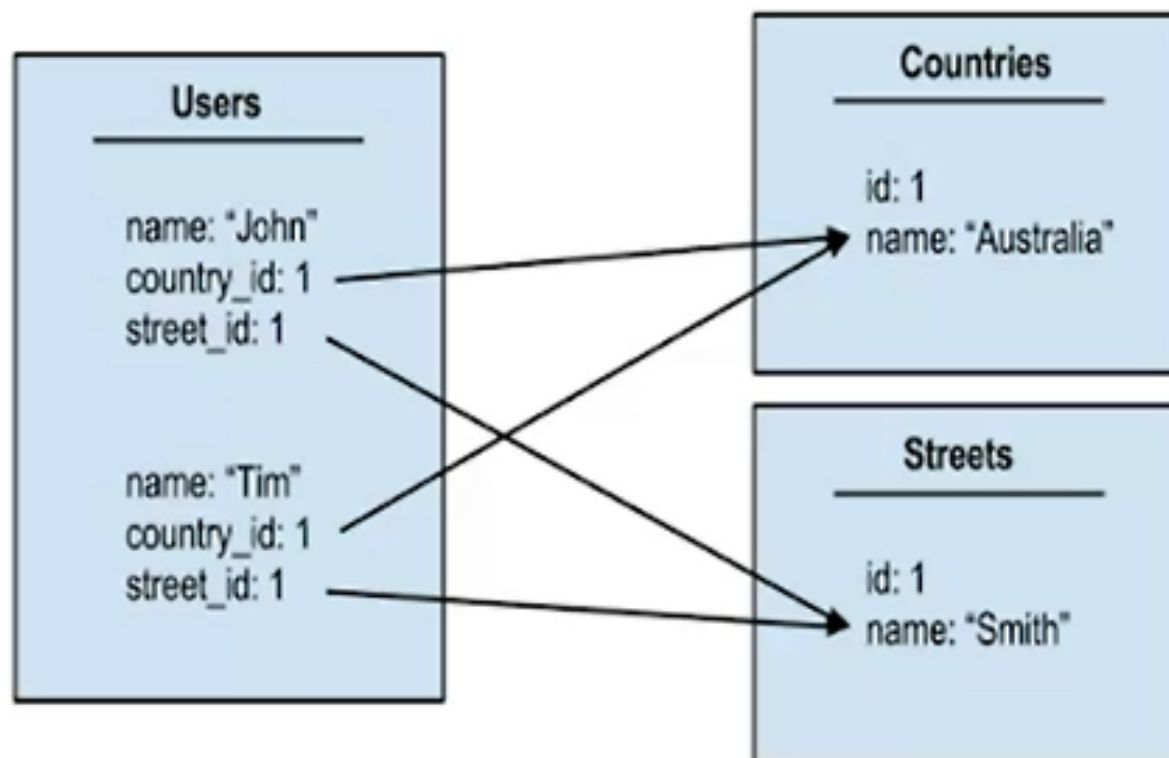
# Денормалізація даних



## MOLAP (Multidimensional OLAP)

- Об'єднання таблиць;
- Розрахунок даних;
- Поля великої довжини виносяться у інші таблиці;

**MOLAP**-сервери використовують для збереження і управління даними *багатовимірні бази даних*. При цьому дані зберігаються у вигляді впорядкованих багатомірних масивів.



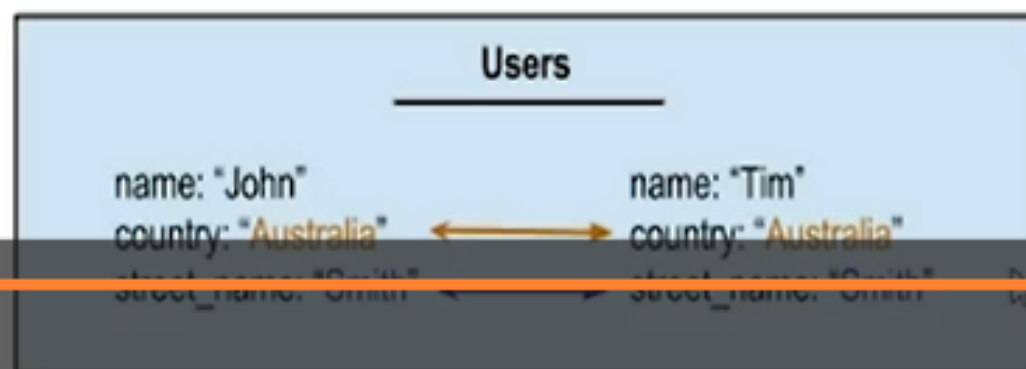
### Normalization

Less duplication

VS

### Denormalization

More duplication



# MOLAP



## Переваги

- Швидкий пошук та вибірка
- Простота використання вбудованих функцій

## Недоліки

- Велика надлишковість
- Гіперкуб зазвичай сильно розріджений
- Чутливість до змін в моделі



# ROLAP (Relational OLAP) - вихідні дані залишаються в тій же реляційній базі даних, де вони завжди і знаходилися. Агрегатні ж дані поміщають у спеціально створені для їх зберігання службові таблиці в тій же базі даних.



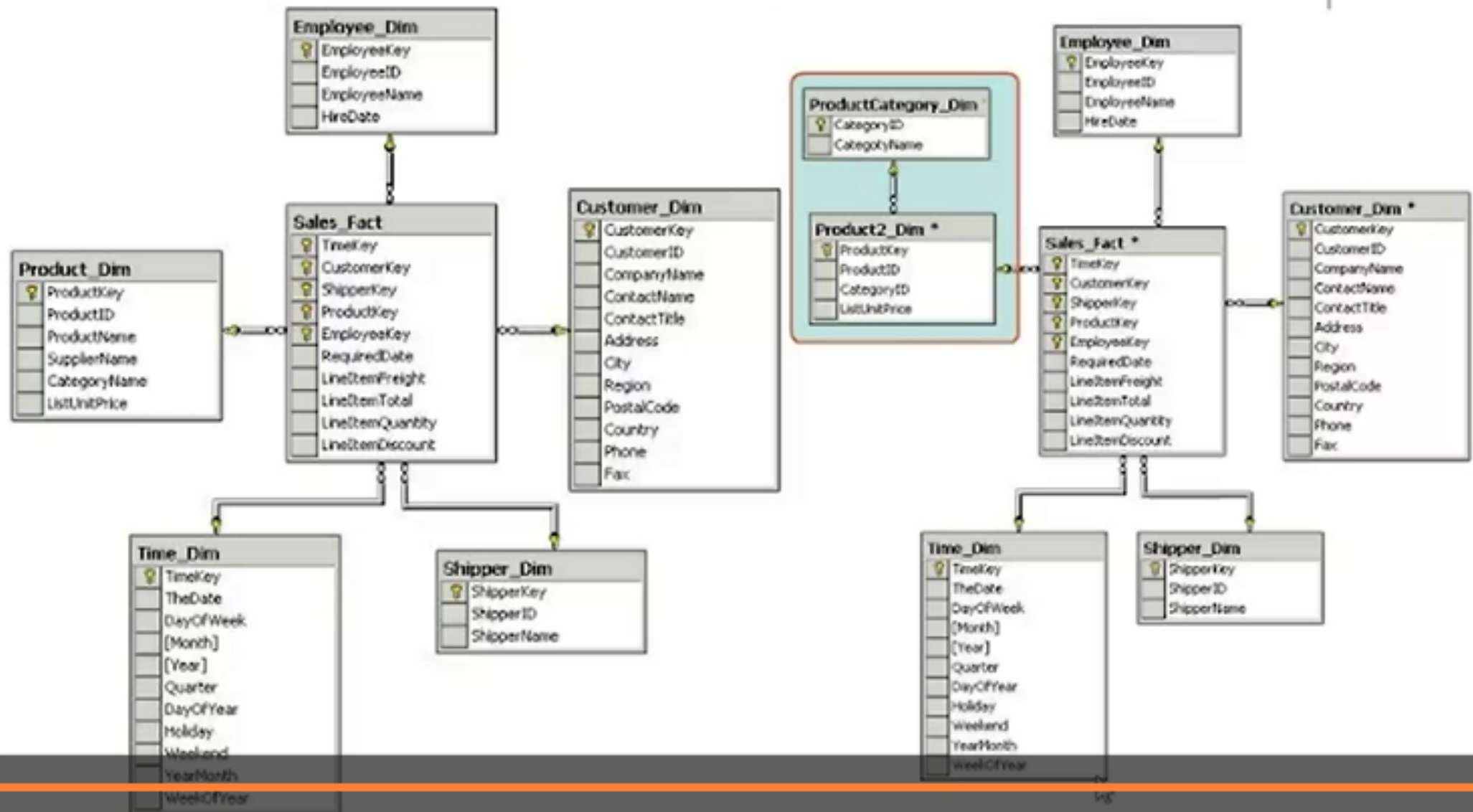
## Переваги

- Менший розмір Сховища Даних;
- Можливість безпосередньо працювати над корпоративними БД;
- Високий рівень захисту даних

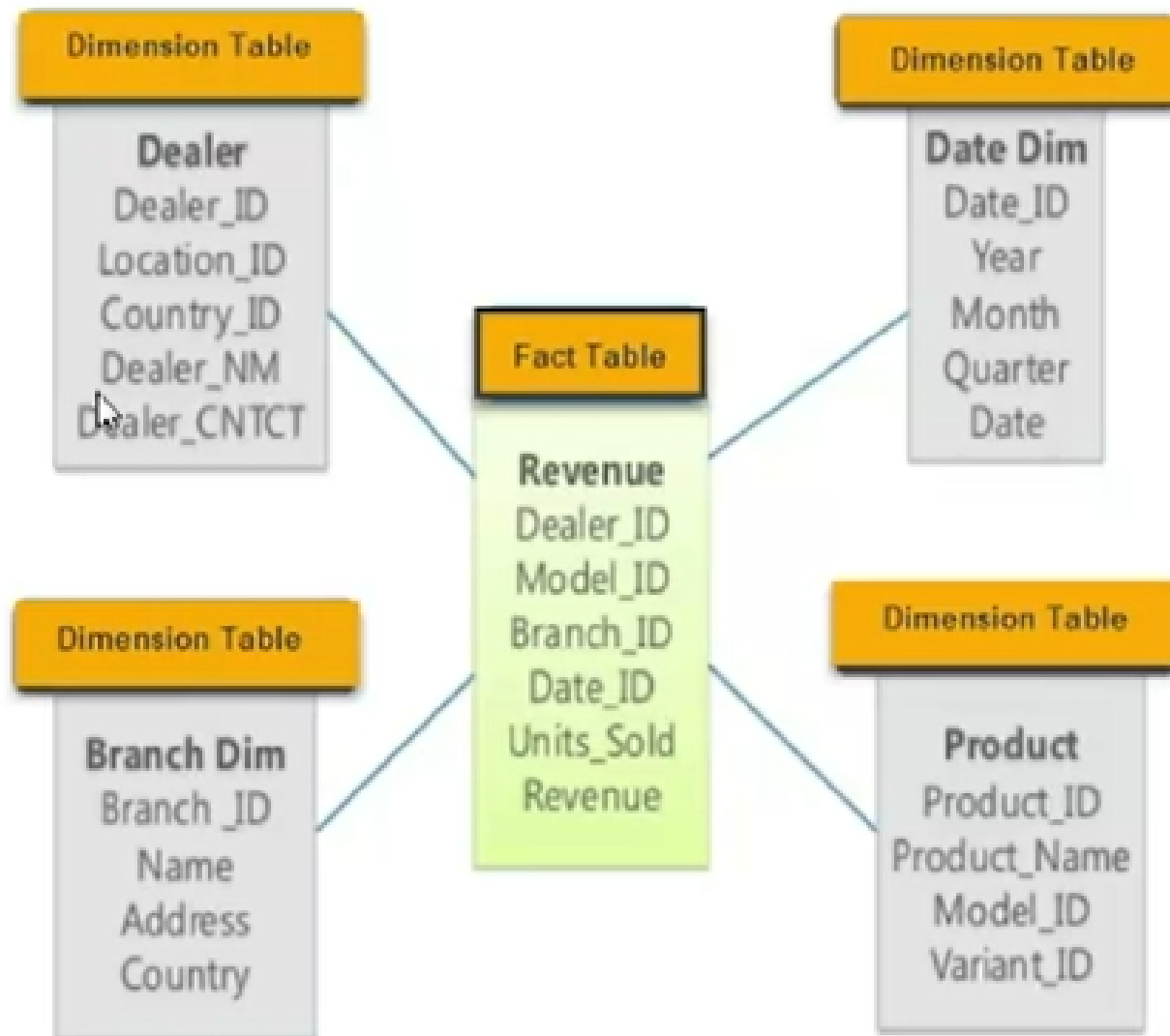
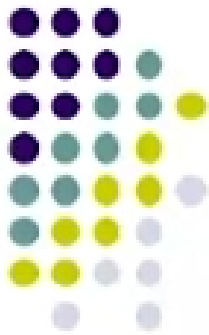
## Недоліки

- Менша продуктивність

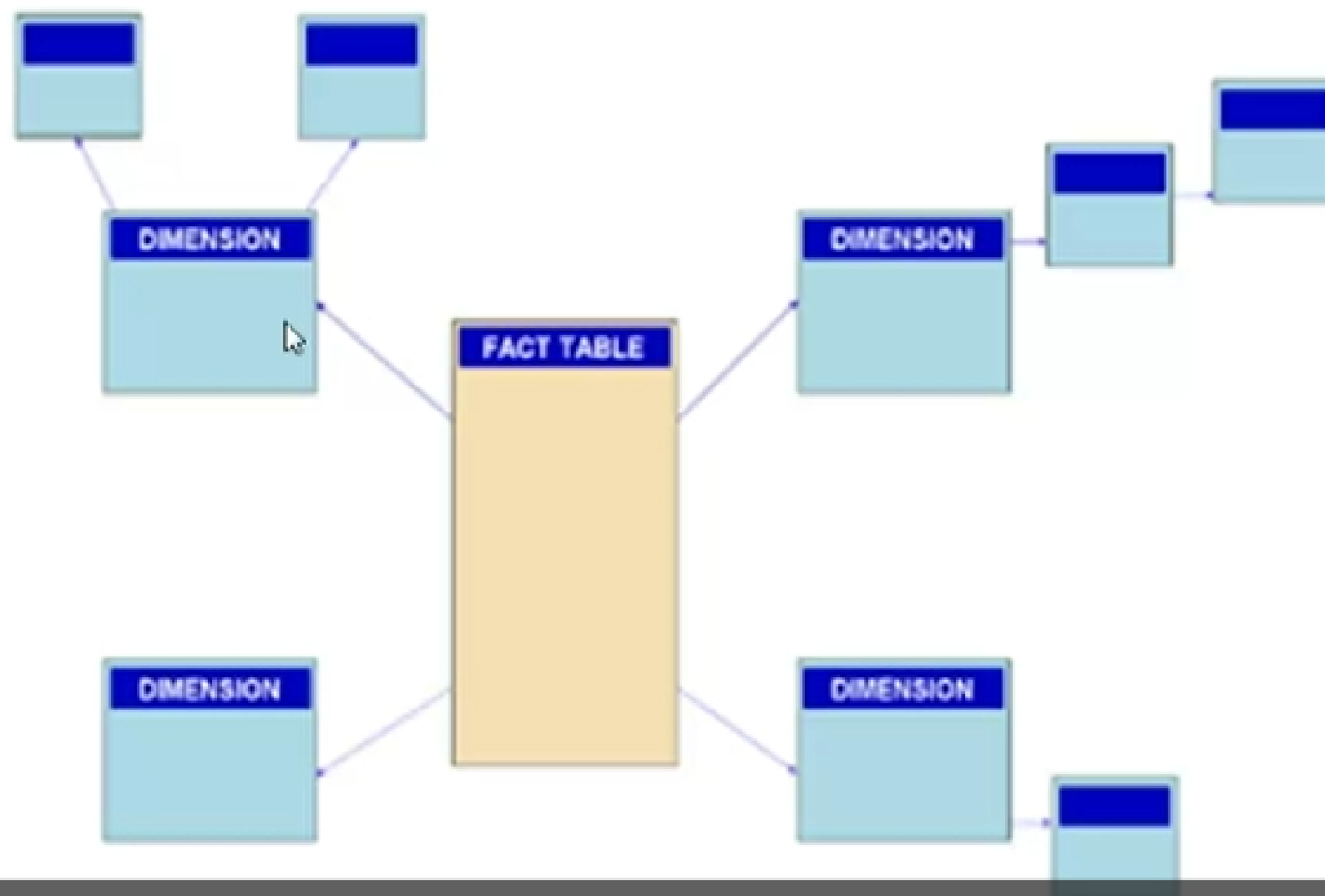
# Схеми «Зірка», «Сніжинка»



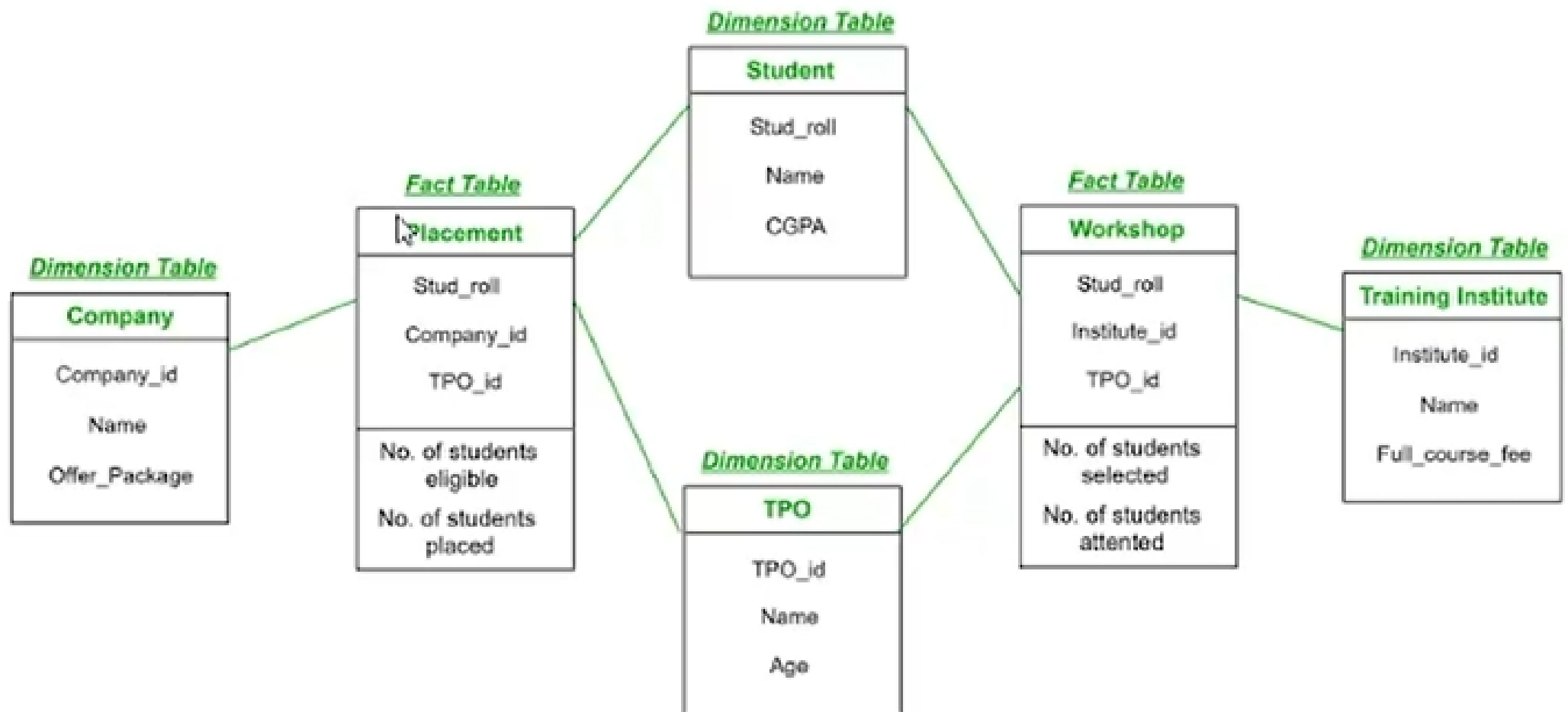
# Схема зірки



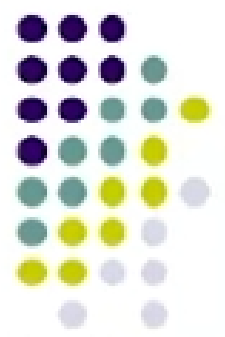
# Схема сніжинки



# Схема галактики



# Підсумок



Отже, OLAP (англ. On-line analytical processing – аналітична обробка в реальному часі) – це технологія обробки інформації, що дозволяє швидко отримувати відповіді на багатовимірні аналітичні запити.

Означення OLAP системи дав Кодд на початку 90-х років XX століття навівши 18 правил, які повинні виконуватись у хорошій системі OLAP аналізу. Згодом ще одне означення дали Н. Пендс і Р. Кріт, яке ще називають **тестом FASMI**.

Ядром будь-якої OLAP-системи є ідея OLAP-куба (багатовимірний куб, або гіперкуб). OLAP-структура, створена з робочих даних, називається OLAP-кубом. Він складається з чисельних фактів, розподілених за вимірами. Зазвичай куб створюється за допомогою з'єднання таблиць із застосуванням схеми «зірка» або схеми «сніжинка».