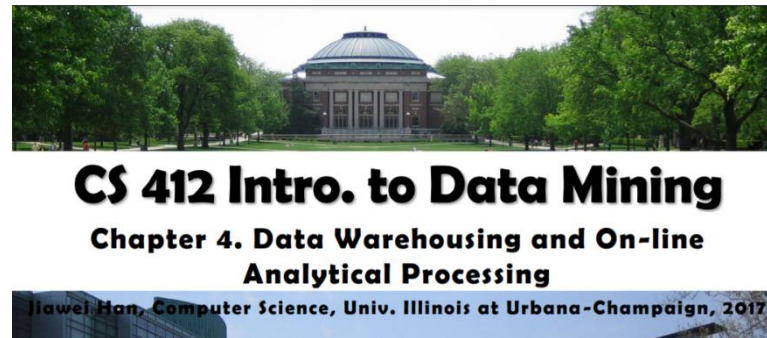


สรุป Chapter 4



What is a Data Warehouse?

- ❑ Defined in many different ways, but not rigorously
 - ❑ A decision support database that is maintained **separately** from the organization's operational database
 - ❑ Support **information processing** by providing a solid platform of consolidated, historical data for analysis
 - ❑ "A data warehouse is a **subject-oriented, integrated, time-variant, and nonvolatile** collection of data in support of management's decision-making process."—W. H. Inmon
 - ❑ Data warehousing:
 - ❑ The process of constructing and using data warehouses
- มีการกำหนดไว้แบบหลาย ๆ วิธี
 - จะต้องมีความหมายว่าจะทำอะไร

From Tables and Spreadsheets to Data Cubes

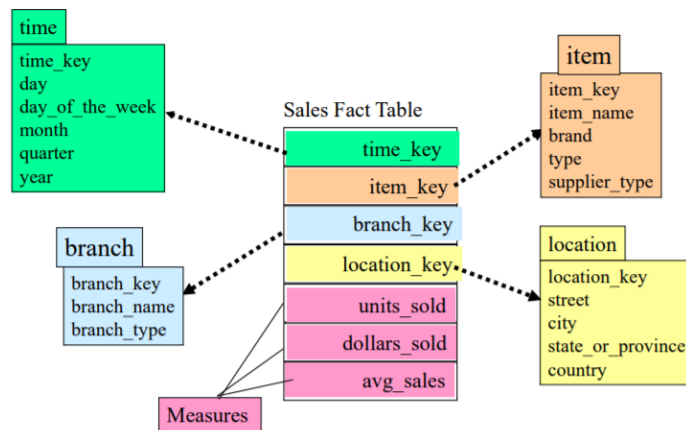
- ❑ A **data warehouse** is based on a multidimensional data model which views data in the form of a data cube
- ❑ A data cube, such as sales, allows data to be modeled and viewed in multiple dimensions
 - ❑ **Dimension tables**, such as item (item_name, brand, type), or time(day, week, month, quarter, year)
 - ❑ **Fact table** contains **measures** (such as dollars_sold) and keys to each of the related dimension tables
- ❑ **Data cube**: A lattice of cuboids
 - ❑ In data warehousing literature, an n-D base cube is called a **base cuboid**
 - ❑ The top most 0-D cuboid, which holds the highest-level of summarization, is called the **apex cuboid**
 - ❑ The lattice of cuboids forms a **data cube**.

- เอาไว้ใช้ว่าสามารถแยกแยะว่าไอเทมข้อมูลแต่ละชนิดมีอะไรบ้าง

- เอาไว้ใช้อธิบายอะไร จำแนกอย่างไร

- สามารถเก็บเป็นตัวเลข

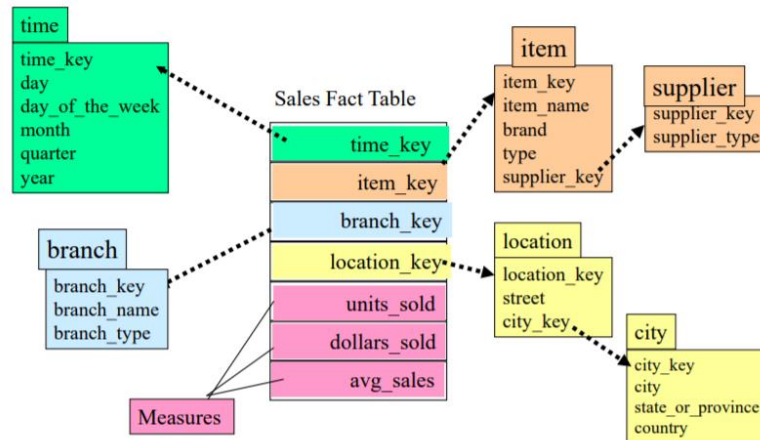
Star Schema: An Example



- ***** สำคัญ *****

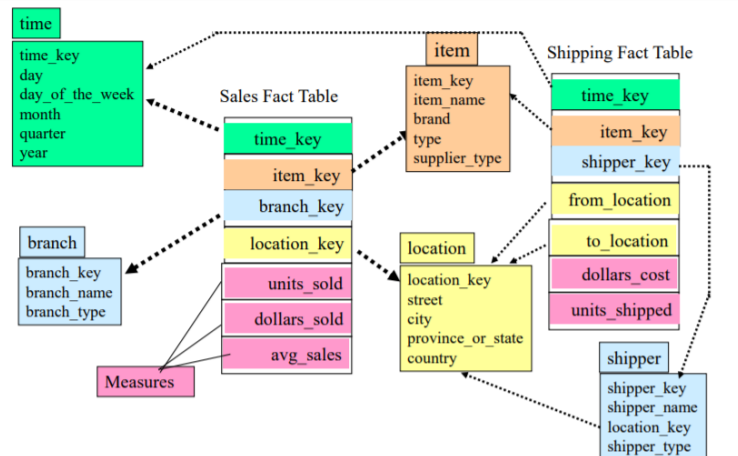
- นำตัวเลขทั้งหมดมารวมได้ตรงกลาง

Snowflake Schema: An Example



- รูปแบบนี้จะเป็นลักษณะเหมือนเกล็ดหิมะ
- สามารถที่จะกระจายแตกแยกออกไปได้

Fact Constellation: An Example

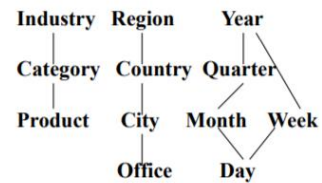
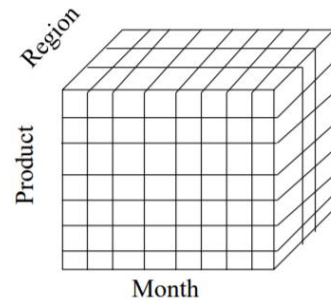


- รูปแบบในลักษณะกลุ่มดาว

Multidimensional Data

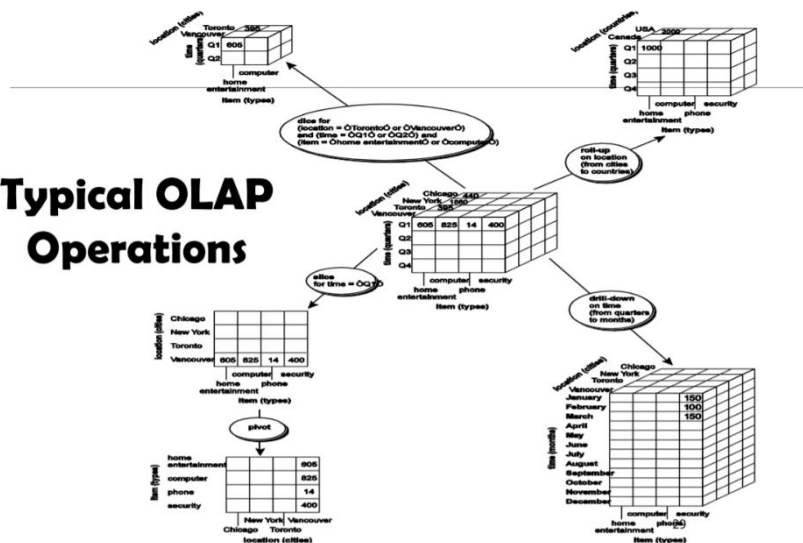
- ❑ Sales volume as a function of product, month, and region

Dimensions: *Product, Location, Time*
Hierarchical summarization paths



- สามารถทำให้มองเห็นได้ละเอียดมากขึ้น

Typical OLAP Operations



- dice เป็นข้อมูลที่มีขนาดใหญ่แล้วตัดออกเป็นข้อมูลขนาดเล็ก
- Slice ตัดมาแค่ในส่วนที่สนใจ
- Roll – up จะเป็นการยุบรวม
- drill down เป็นการแยกออก