Subject Name: DATA WAREHOUSING AND MINING

Unit No:2

Unit Name: Online Analytical Processing(OLAP) & ETL

Faculty Name Mr.Tushar Ghorpade

Ms. Saguna Ingle

Ms. Dhanashri A. Bhosale

Index

OLTP V/s OLAP

Parameters	OLTP	OLAP
Process	It is an online transactional system. It manages database modification.	OLAP is an online analysis and data retrieving process.
Characteristic	It is characterized by large numbers of short online transactions.	It is characterized by a large volume of data.
Functionality	OLTP is an online database modifying system.	OLAP is an online database query management system.
Method	OLTP uses traditional DBMS.	OLAP uses the data warehouse.
Query	Insert, Update, and Delete information from the database.	Mostly select operations
Table	Tables in OLTP database are normalized.	Tables in OLAP database are not normalized.
Source	OLTP and its transactions are the sources of data.	Different OLTP databases become the source of data for OLAP.
Data Integrity	OLTP database must maintain data integrity constraint.	OLAP database does not get frequently modified. Hence, data integrity is not an issue.
Response time	It's response time is in millisecond.	Response time in seconds to minutes.

Lecture no 7: OLTP Vs OLAP ,OLAP Operations

OLTP V/s OLAP

|		
Data quality	The data in the OLTP database is always detailed and organized.	The data in OLAP process might not be organized.
Usefulness	It helps to control and run fundamental business tasks.	It helps with planning, problem-solving, and decision support.
Operation	Allow read/write operations.	Only read and rarely write.
Audience	It is a market orientated process.	It is a customer orientated process.
Query Type	Queries in this process are standardized and simple.	Complex queries involving aggregations.
Back-up	Complete backup of the data combined with incremental backups.	OLAP only need a backup from time to time. Backup is not important compared to OLTP
Design	DB design is application oriented. Example: Database design changes with industry like Retail, Airline, Banking, etc.	DB design is subject oriented. Example: Database design changes with subjects like sales, marketing, purchasing, etc.
User type	It is used by Data critical users like clerk, DBA & Data Base professionals.	Used by Data knowledge users like workers, managers, and CEO.

Lecture no 7: OLTP Vs OLAP ,OLAP Operations

OLTP V/s OLAP

	OLTP	OLAP
Purpose	Designed for real time business operations.	Designed for analysis of business measures by category and attributes.
Performance metric	Transaction throughput is the performance metric	Query throughput is the performance metric.
Number of users	This kind of Database users allows thousands of users.	This kind of Database allows only hundreds of users.
Productivity	It helps to Increase user's self-service and productivity	Help to Increase productivity of the business analysts.
Challenge	Data Warehouses historically have been a development project which may prove costly to build.	An OLAP cube is not an open SQL server data warehouse. Therefore, technical knowledge and experience is essential to manage the OLAP server.
Process	It provides fast result for daily used data.	It ensures that response to the query is quicker consistently.
Characteristic	It is easy to create and maintain.	It lets the user create a view with the help of a spreadsheet.
Style	OLTP is designed to have fast response time, low data redundancy and is normalized.	A data warehouse is created uniquely so that it can integrate different data sources for building a consolidated database

Lecture no 7: OLTP Vs OLAP ,OLAP Operations

OLAP operations

- **Online Analytical Processing (OLAP)** is a category of software that allows users to analyze information from multiple database systems at the same time.

- It is a technology that enables analysts to extract and view business data from different points of view.

- Analysts frequently need to group, aggregate and join data. These operations in relational databases are resource intensive.

- With OLAP data can be pre-calculated and pre-aggregated, making analysis faster.

- OLAP databases are divided into one or more cubes.

- The cubes are designed in such a way that creating and viewing reports become easy.

Lecture no 7: OLTP Vs OLAP
,OLAP Operations

OLAP Cube

- At the core of the OLAP concept, is an OLAP Cube.

- The OLAP cube is a data structure optimized for very quick data analysis.

Lecture no 7: OLTP Vs OLAP
,OLAP Operations
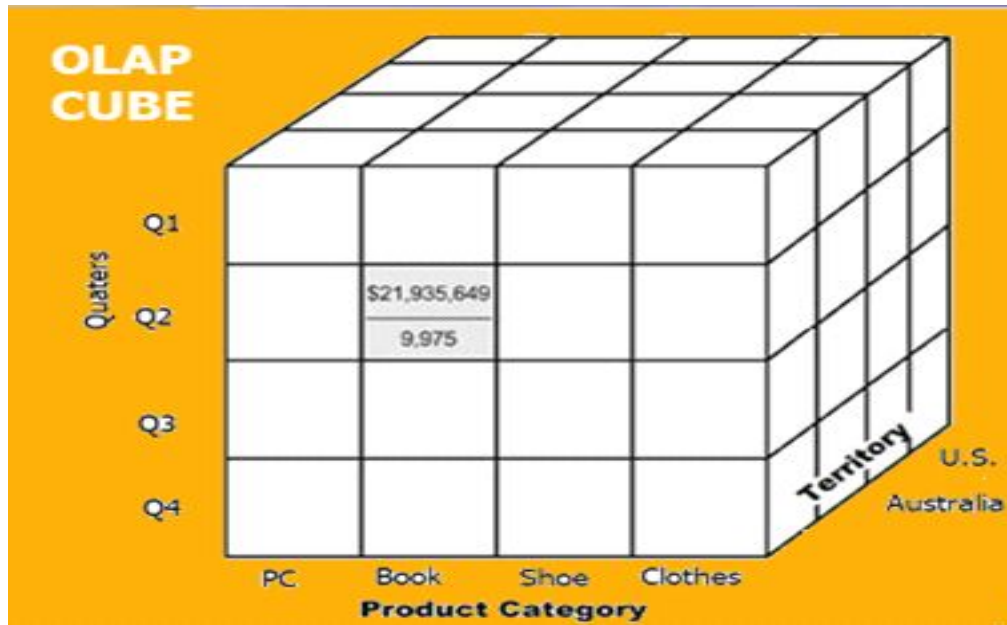
OLAP cube

- The OLAP Cube consists of numeric facts called measures which are categorized by dimensions.

- OLAP Cube is also called the **hypercube**.

Why OLAP Cube?
- Usually, data operations and analysis are performed using the simple spreadsheet, where data values are arranged in row and column format. This is ideal for two-dimensional data. However, OLAP contains multidimensional data, with data usually obtained from a different and unrelated source. Using a spreadsheet is not an optimal option.

- The cube can store and analyze multidimensional data in a logical and orderly manner.

OLAP cube

How does it work?

•A Data warehouse would extract information from multiple data sources and formats like text files, excel sheet, multimedia files, etc.

•The extracted data is cleaned and transformed. Data is loaded into an OLAP server (or OLAP cube) where information is pre-calculated in advance for further analysis.

Lecture no 7: OLTP Vs OLAP ,OLAP Operations

OLAP operations: Drill down and Roll up

**Basic analytical operations of OLAP**

Five types of analytical operations in OLAP are:

- Roll-up(drill-up) : Summarize data

- Drill-down(roll down): reverse of roll-up

- Slice : select

- Dice : project

- Pivot (rotate): reorient the cube

Lecture no 7: OLTP Vs OLAP
,OLAP Operations

Roll-up

- Roll-up is also known as "consolidation" or "aggregation."

- The Roll-up operation can be performed in 2 ways
1. Reducing dimensions
2. Climbing up concept hierarchy.

- Concept hierarchy is a system of grouping things based on their order or level.

- Consider an example:

Roll-up
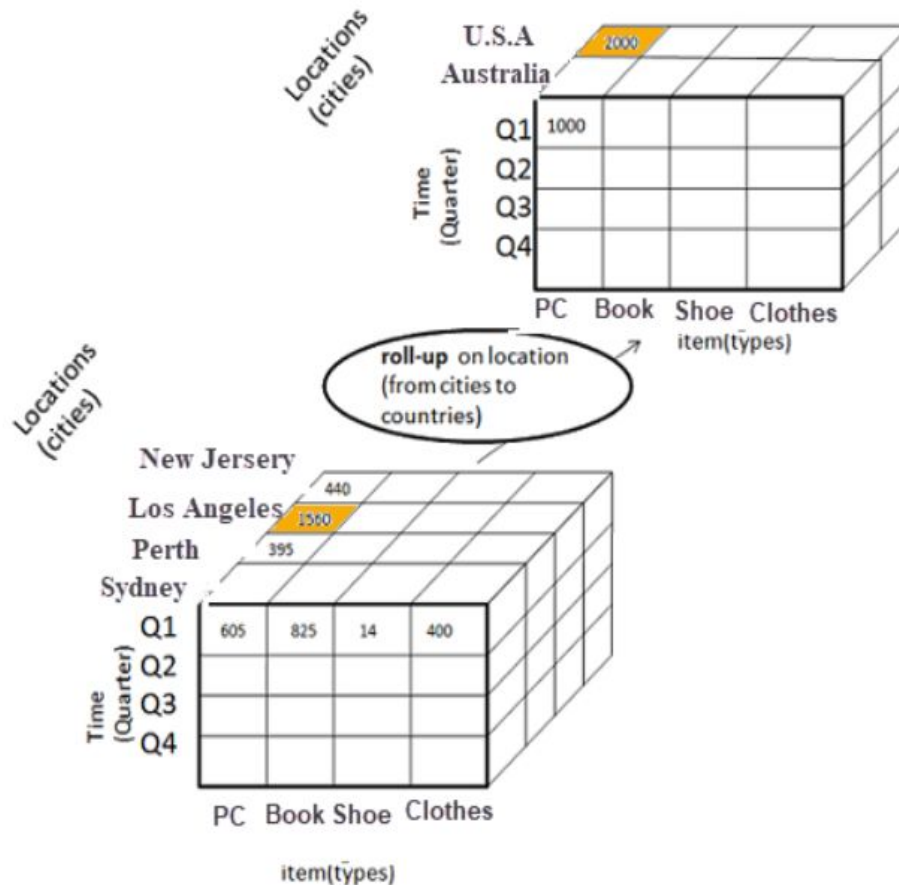
- In this example, cities New jersey and Lost Angles are rolled up into country USA

- The sales figure of New Jersey and Los Angeles are 440 and 1560 respectively. They become 2000 after roll-up

- In this aggregation process, data is location hierarchy moves up from city to the country.

- In the roll-up process at least one or more dimensions need to be removed. In this example, Quater dimension is removed.

Lecture no 7: OLTP Vs OLAP ,OLAP Operations

Drill down

- Give detail information
- In drill-down data is fragmented into smaller parts.
- It is the opposite of the rollup process. It can be done via

- Moving down the concept hierarchy
- Increasing a dimension

Consider an example:

- Quater Q1 is drilled down to months January, February, and March. Corresponding

  sales are also registers.
- In this example, dimension months are added.

Drill down Example
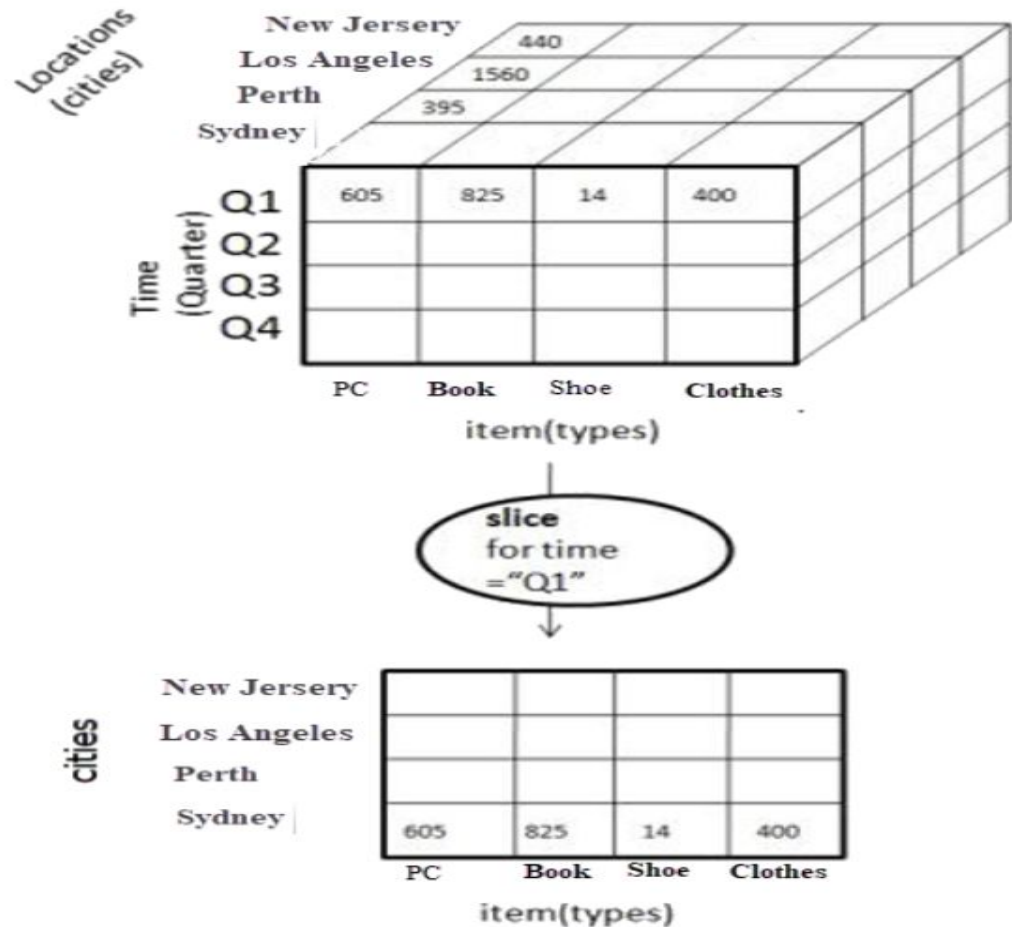
Lecture no 7: OLTP Vs OLAP
,OLAP Operations
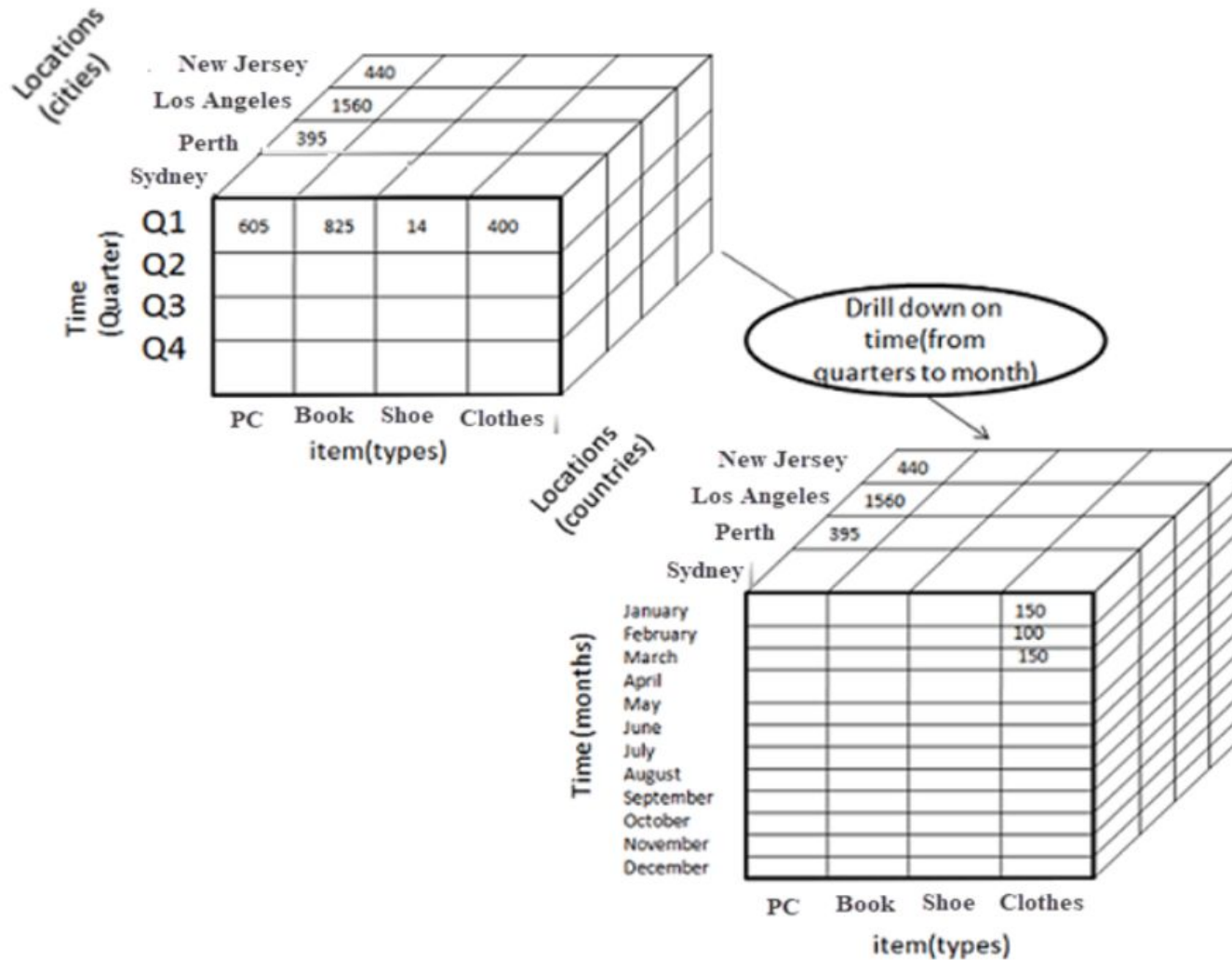
OLAP operations: Slice

Here, one dimension is selected, and a new sub-cube is created.

Dimension Time is Sliced with Q1 as the filter.

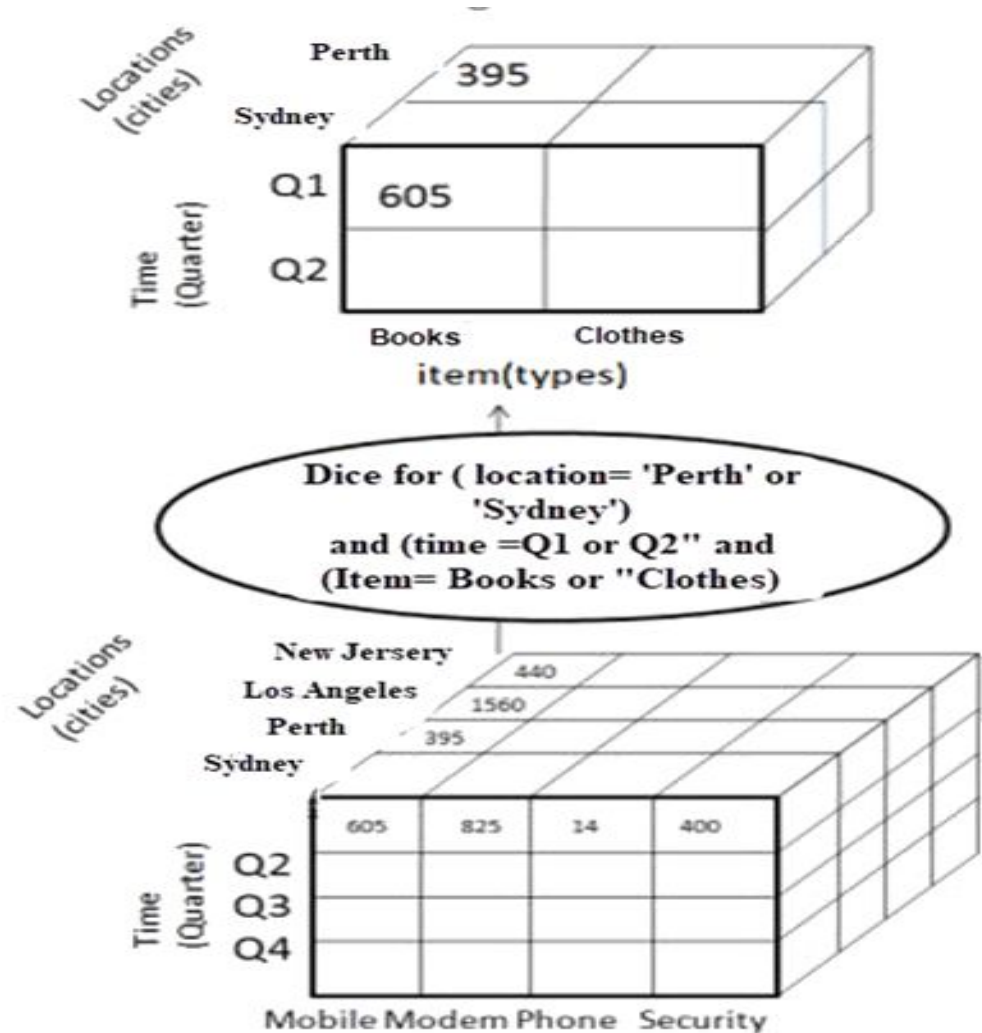A new cube is created altogether.

Lecture no 7: OLTP Vs OLAP ,OLAP Operations

OLAP operations: Dice

This operation is similar to a slice.

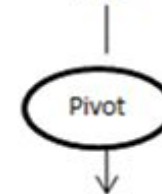The difference in dice is you select 2 or more dimensions that result in the creation of a sub-cube.

Lecture no 7: OLTP Vs OLAP ,OLAP Operations

OLAP operations: Pivot(Rotate)

In Pivot, you rotate the data axes to provide a substitute presentation of data.

In the following example, the pivot is based on item types.

Locations (cities)

	PC	Book	Shoe	Clothes
New Jersey				
Los Angeles				
Perth				
Sydney	605	825	14	400

item(types)

↓ Pivot ↓

Item (types)

	New Jersey	Los Angeles	Perth	Sydney
PC				605
Book				825
Shoe				14
Clothes				400

Location (Cities)

Lecture no 7: OLTP Vs OLAP ,OLAP Operations

OLAP Models

Lecture 9 - Data Mining: Task Primitives, Architecture

Thank You