

Data Warehousing Concepts

Lesson 5: Online Analytical
Processing (OLAP)

Lesson Objectives

- In this lesson, you will learn about:
 - The concept of Online Analytical Processing
 - Need for Separate Operational and Informational Systems
 - Nature of OLAP analysis
 - Types Of OLAP
 - OLAP Service Tools
 - OLTP and OLAP
 - Operational versus Informational Systems



5.1: Online Analytical Processing (OLAP)

Concept of OLAP

- OLAP is a functionality available in Data Warehouse applications.
- It enables client applications to efficiently access data in a Data Warehouse or Data Mart.
- It is a multi-dimensional data model.
- It contains a variety of possible views of information.
- It simplifies evaluation of ad hoc complex queries.
- It provides a very fast response time to ad hoc queries.


Online Analytical Processing (OLAP):

- **OLAP** is a category of software technology. It enables the users to gain insight into data through fast, consistent, interactive access to a wide variety of possible views of information, which has been transformed from raw data, to reflect the real dimensionality of the business.
- It provides benefits like **pre-aggregation** of frequently required data, enabling a very fast response time to ad hoc queries. It gives a **multi-dimensional data model** that makes it easy to select, navigate, and explore the data.
- OLAP systems enable managers and analysts to rapidly and easily examine key performance data. OLAP systems allow comparison and trend analysis even on very large volumes. OLAP allows users to view data from various perspectives. It is fast and easy because some aggregations are computed in advance.

5.2: Nature of OLAP Analysis

Use of OLAP

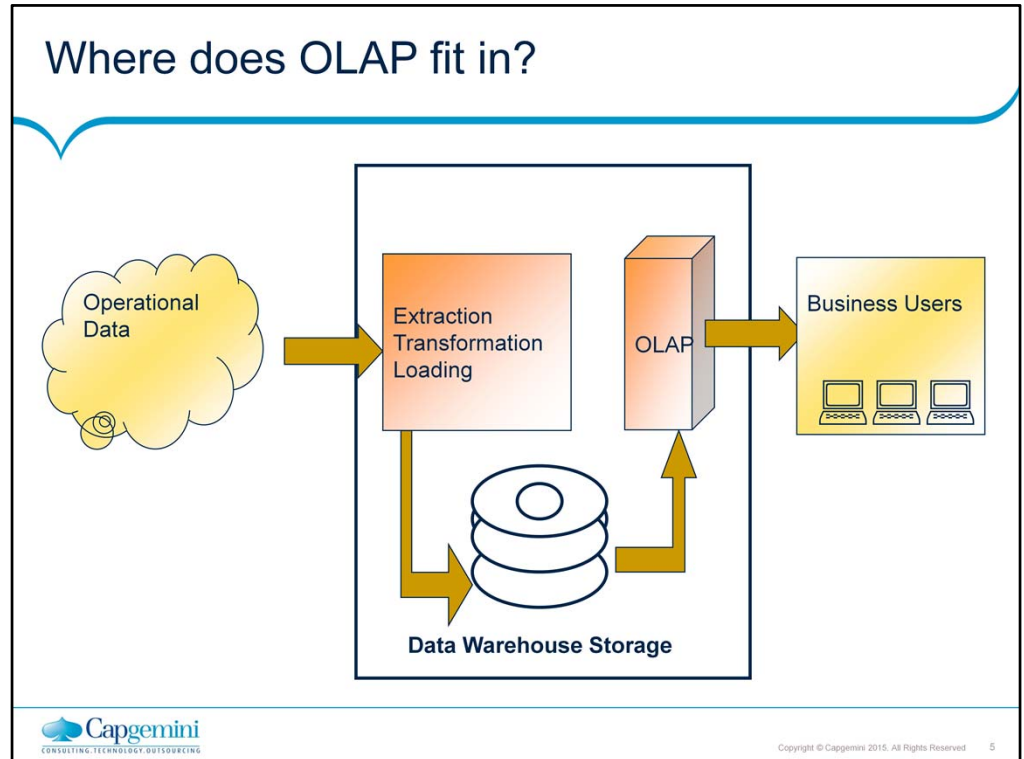
- OLAP analysis is used for:
 - Aggregation
 - Comparison
 - Ranking
 - Access to data
 - Complex criteria specification

 Capgemini
CONSULTING TECHNOLOGY ENTREPRENEUR

Copyright © Capgemini 2015. All Rights Reserved 4

Nature of OLAP Analysis:

- The nature of OLAP analysis varies with the multiple ways of using it.
 - It is used as the aggregation for summing up the data.
 - It provides the easy way of comparison.
 - It allows us to rank the data such that you will be able to find the top most and lower most values in the analysis.
 - It also helps in accessing the data in detailed way.
 - It allows to perform complex specification on the criteria.
 - It represents the data in a more simpler way such that it is easily visualized in terms of graphical presentation.
- OLAP Analysis is used for:
 - Aggregation: (total sales, percent-to-total)
 - Comparison: Budget versus Expenses
 - Ranking: Top 10, quartile analysis
 - Access to detailed and aggregate data
 - Complex criteria specification



Nature of OLAP Analysis:

Where does OLAP fit in?

- Data from various sources goes through the ETL process and is integrated into a Data Warehouse. Subsequently, OLAP is used to analyze the data in the Data Warehouse. OLAP focuses on meeting end-user's analytical requirements.
 - **Operational Data:** It is the Customer Database, for example, Sales Database and Product Database.
 - **End User:** High performance is achieved by pre-planning the requirement for joins, summations and periodic reports by end users.
 - **Extract Transform and Load (ETL):**
 - **Extract:** It extracts data from data source and keeps it in staging.
 - **Transform:** It converts data into format required by Data Warehouse.
 - **Load:** It loads data to Data Warehouse.
 - **DWH:** Data Warehouse integrates and aggregates data from various operational and external data bases maintained by different Business Units.
 - **OLAP:** It has been in use to process and record transactions that create new data and update existing information in databases.
 - **Business User:** High performance is achieved by pre-planning the requirement and putting business rules by business users.


OLAP Models

- OLAP models are of different types
- The processing in all these different types is the same:
 - Online Analytical processing
- The storage methods are different in different models
- Different OLAP Models are
 - ROLAP
 - MOLAP
 - HOLAP

5.3: Types of OLAP

ROLAP

- Relational Online Analytical Processing (ROLAP):
 - It stores in a Relational form.
 - It stores Data Mart (Star schema).
- Advantages:
 - It has no data size limitation.
 - It can leverage functions of RDB.
- Disadvantages:
 - Each request must query the RDB.
 - ROLAP itself is limited to RDB functionality.

 Capgemini
CONSULTING TECHNOLOGY SERVICES

Copyright © Capgemini 2015. All Rights Reserved 7

Types of OLAP:

Relational Online Analytical Processing (ROLAP):

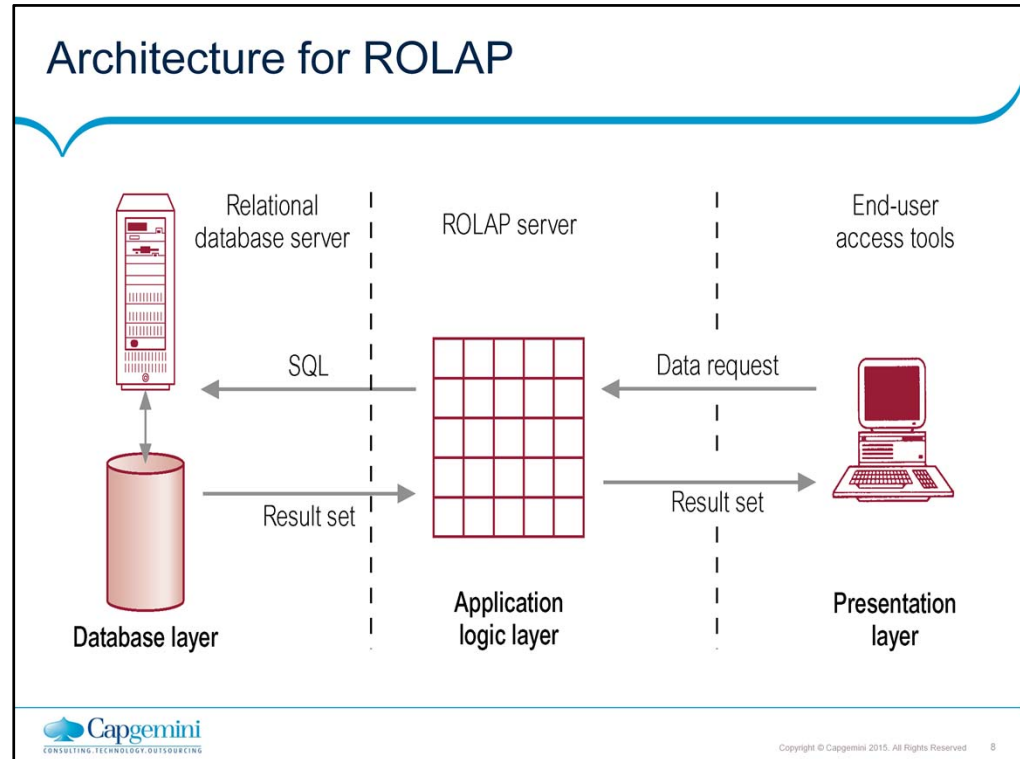
- The data stored in the relational database gives the appearance of traditional OLAP's slicing and dicing functionality.
- In Relational OLAP (ROLAP), Relational DBMS stores Data Mart (Star schema).

Advantage:

- ROLAP itself places no limitation on data amount.
- Relational database already comes with a host of functionalities. ROLAP technologies, can leverage these functionalities since they sit on top of the relational database.

Disadvantage:


- Each ROLAP report is essentially a SQL query (or multiple SQL queries) in the relational database. The query time can be long if the underlying data size is large.
- ROLAP itself is limited to RDB functionality.



5.3: Types of OLAP

MOLAP

- Multidimensional Online Analytical Processing (MOLAP):
 - Data is stored multi-dimensionally by using Multidimensional Databases (MDDB)
 - MDDB's store data in the form of Multi dimensional cubes
 - MOLAP cubes are built for fast data retrieval and are optimal for slicing and dicing operations.
- Advantages:
 - It has excellent performance.
 - It can return complex calculations.
- Disadvantages:
 - It is limited in scope as definition of cube creates boundaries.
 - Limited volume of data is churned.

 Copyright © Capgemini 2015. All Rights Reserved 9

Types of OLAP:

Multidimensional Online Analytical Processing (MOLAP):

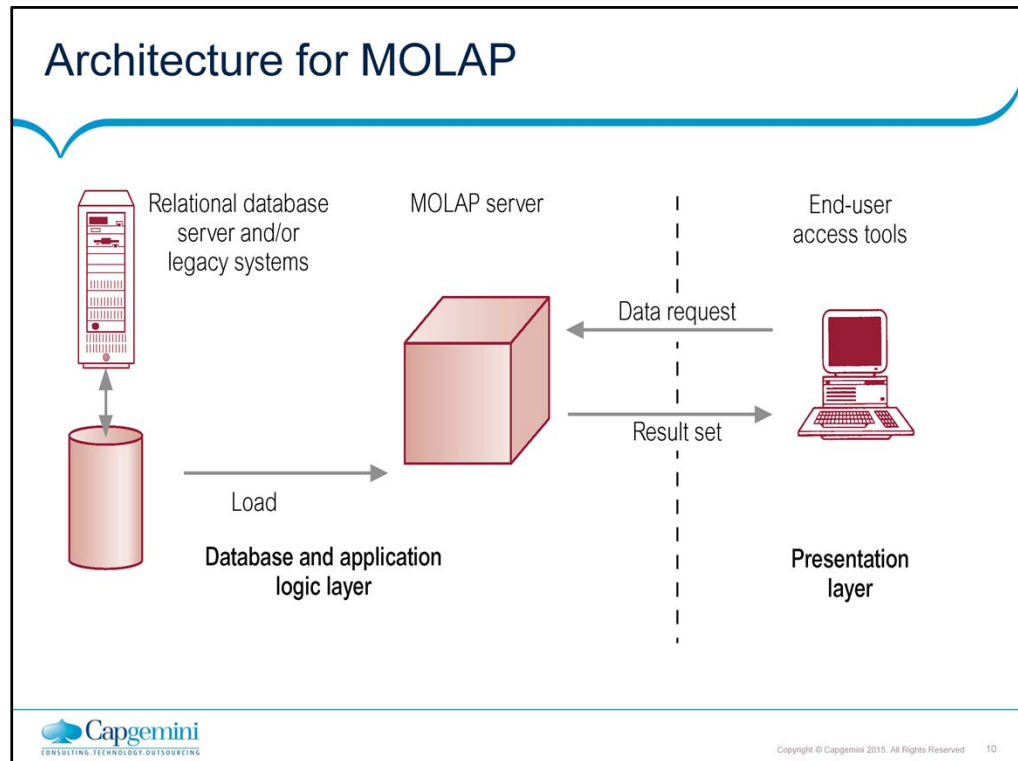
- In MOLAP, the data is stored in a multi-dimensional cube. The storage is not in the relational database, but in proprietary formats.
- MOLAP storage structure is Array-based.

Advantage:

- MOLAP cubes are built for fast data retrieval, and is optimal for slicing and dicing operations.
- All calculations have been pre-generated when the cube is created. Hence, complex calculations are not only doable, they return quickly, as well.

Disadvantage:

- In case of MOLAP, scope is limited as definition of cube creates boundaries.
- It is not possible to include a large amount of data in the cube itself. This is because all calculations are performed when the cube is built. Only summary-level information will be included in the cube itself.



HOLAP

- HOLAP is the product of the attempt to incorporate the best features of MOLAP and ROLAP into a single architecture.
- HOLAP systems stores larger quantities of detailed data in the relational tables while the aggregations are stored in the pre-calculated cubes.
- HOLAP also has the capacity to “drill through” from the cube down to the relational tables for delineated data.



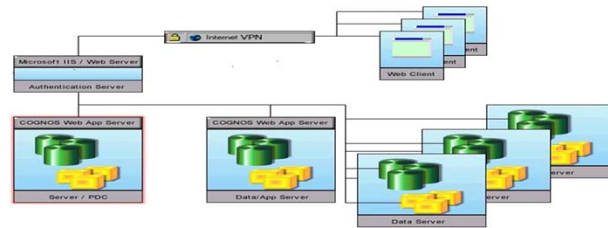
Copyright © Capgemini 2015. All Rights Reserved 11

This tool tried to bridge the technology gap of both products by enabling access or use to both multidimensional database (MDDb) and Relational Database Management System (RDBMS) data stores.

WOLAP Architecture

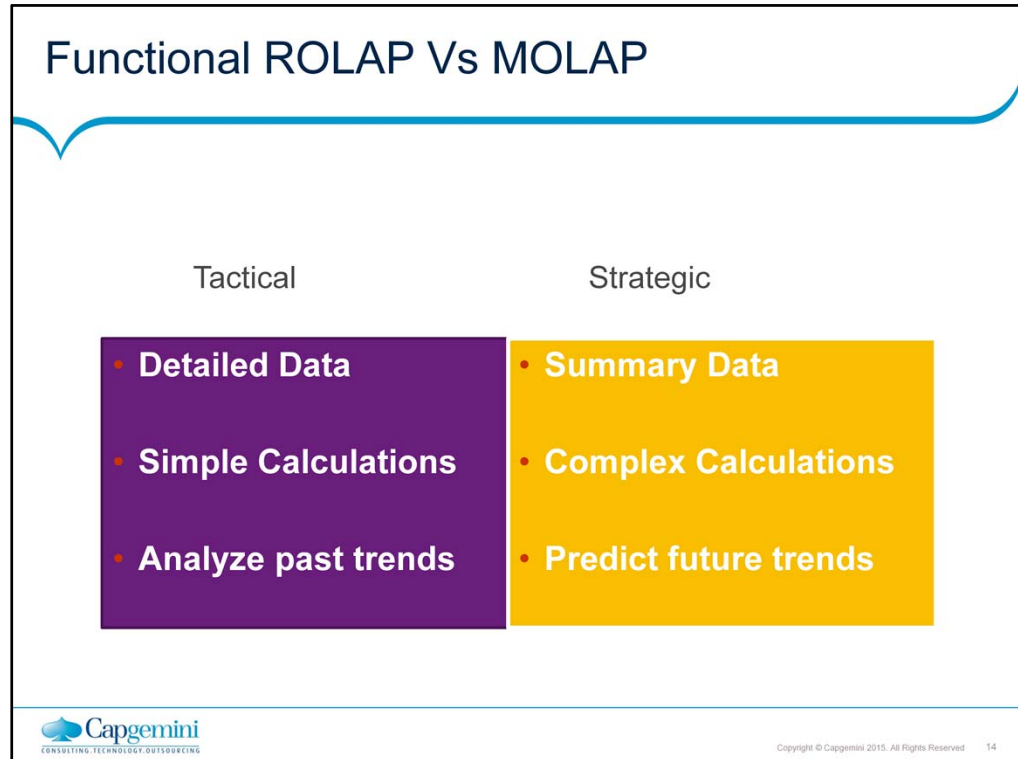
- Web Based Online Analytical Processing: WOLAP, or Web-enabled OLAP, which uses a browser to deliver OLAP. The combination is powerful, say many business managers. The delivery capability of the Web, coupled with the business intelligence tool of OLAP, will allow a broader number of business analysts to benefit from the software.

- Example :



Other OLAP tools

- DOLAP – Desktop OLAP
- MOLAP – Mobile OLAP
- SOLAP – Spatial OLAP



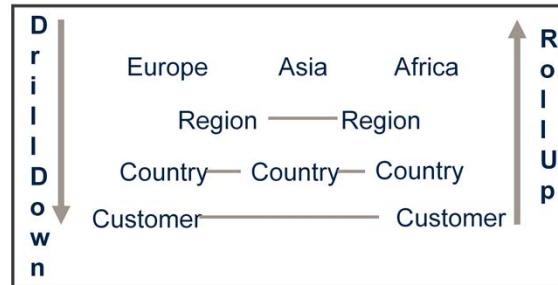
Types of OLAP Operations

- Different OLAP operations are:

- Roll up (drill-up)
- Drill down (roll down)
- Slice and dice
- Pivot (rotate)

- Other operations:

- Drill across
- Drill through



Types Of OLAP Operations:

- **Aggregation / Consolidation / Roll up (drill-up):** Roll Up operation is used to summarize data by climbing up hierarchy or by dimension reduction. This allows you to aggregate data from lower level details to the parent level. For example, the total revenue generated by a particular product type will be the rolled up value of the revenue generated by descendants, that is, products that belong to that particular product type.
- **Drill down (roll down):** Roll down operation is the reverse of roll-up, that is, a drill down from higher level summary to lower level summary or detailed data, or introducing new dimensions. It allows you to view data from a top-level to a detailed view by going down the hierarchy. For example, we can view the Sales data for the Year and drill down from the year level to a quarterly view and further down to a monthly view.
- **Slice and dice:** It is a general term for viewing data from any angle.
- **Pivot (rotate):** Rotate operation is used for reorienting the cube, visualization, 3D to series of 2D planes.

Other operations:

- **Drill Across:** It involves (across) more than one fact table.
- **Drill Through:** It involves drilling through the bottom level of the cube to its back-end relational.

5.5: OLTP and OLAP

Concepts of OLTP and OLAP

- OLTP: Online Transaction Processing:
 - OLTP is used to process and record transactions that create new data.
 - It updates existing information in databases.
- OLAP: Online Analytical Processing:
 - Data is aggregated, warehoused, and then analyzed.
 - Users query and generate reports without modifying any data.



Copyright © Capgemini 2015. All Rights Reserved 16

OLTP and OLAP:

- OLTP is basically operational data, wherein data is frequently changing.
- For example, you can consider an online Railway Reservation system. A passenger books a ticket for two people. This becomes an operational data. S/He can also change the number of passengers travelling online since OLTP data is frequently updated.
- On the other hand, OLAP data is non-operational data, wherein data is read-only data. It used for analytical purpose.
- For example, suppose one wants to see, the number of trains running on the previous day. This becomes analysis of data that is not updatable.

OLAP Functional Requirements

- Fast Access and Calculations

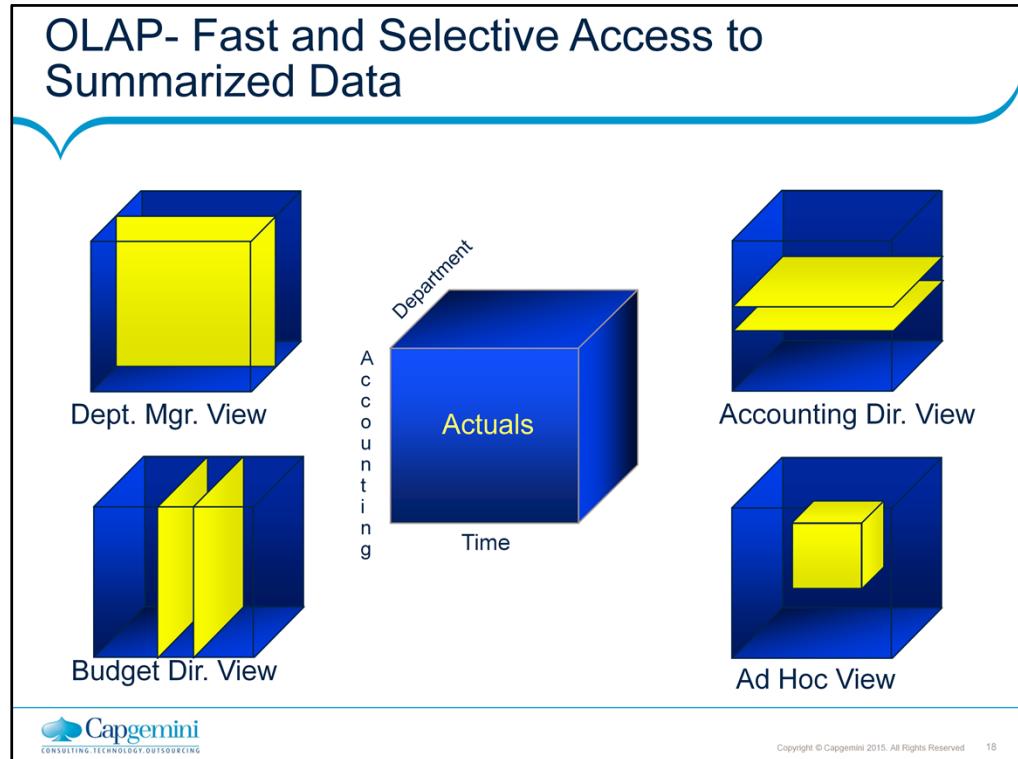
- Speed is critical to maintain an analyst's train of thought.
- An analyst needs to navigate throughout the data which requires aggregations, or roll-ups.

- Powerful Analytical Capabilities

- There is more complicated calculations to OLAP than simple aggregations, or roll-ups.

- Flexibility

- Viewing: graphs, charts, row or columns
- Definitions: format of numbers, name changes
- Analysis: Sales analyze data differently than marketing
- Interfaces: section wise, report looks



5.6: Operational versus Informational system

Points of Difference

	Operational (OLTP)	Informational (DW)
Typical User	Clerical	Management
System usage	Regular business	Analysis
Workload	Read/Write	Read only
Types of queries	Predefined	Ad-hoc
Unit of interaction	Transaction	Query
Level of isolation required	High	Low
No of records accessed	<100	>1,000
No of concurrent users	Thousands	Hundred
Focus	Data in and out	Information out

Note: The above table shows the comparison between **OLTP** and **Data Warehouse**.

Points of Difference

	Operational (OLTP)	Analytical Systems
User	Clerk, IT Professional	Knowledge Worker
Function	Day to day operations	Decision support
DB Design	Application-oriented (E-R based)	Subject Oriented (Star, Snow flake)
Data	Current, Isolated	Historical, Consolidated
View	Detailed, Flat relational	Summarized, Multidimensional
Usage	Structured, Repetitive	Ad hoc
Unit of Work	Short, Simple transaction	Complex Query
Access	Read/Write	Read Mostly
Operation	Index/hash on prim. Key	Lots of scan
# Records accessed	Tens	Millions
#Users	Thousands	Hundreds
Db size	100 MB-GB	100GB-TB
Metric	Transaction throughput	Query throughput response

Note: Above table depicts the comparison between **Operational** and **Analytical systems**.

Summary

- In this lesson, you have learnt:
 - OLAP allows users to view data from various perspectives.
 - The nature of OLAP analysis varies in multiple ways of using it.
 - Aggregation
 - Comparison
 - Ranking
 - Access to data
 - Complex criteria specification



Summary

- OLAP is used to analyze the data in the Data Warehouse.
- Different types of OLAP are:
 - MOALP
 - HOLAP
 - ROLAP



Review Question

- Question 1: OLAP analysis is used for:
 - Option 1: Retrieving data
 - Option 2: Updating data
 - Option 3: Summarizing data
- Question 2: OLAP makes use of multidimensional data model.
 - True/ False
- Question 3: ____ OLAP operation helps for viewing data from any angle.

