# BI - UNIT III (Part 1)

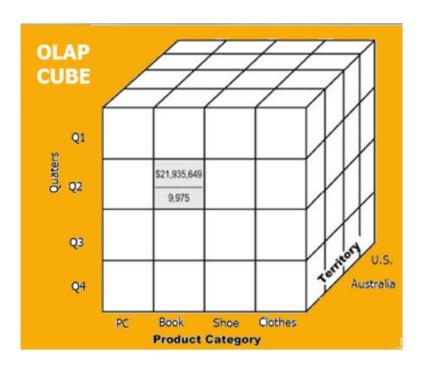
#### 1. EXPLAIN OLAP.

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. OLAP stands for Online Analytical Processing.

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.

# **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.

The OLAP Cube consists of numeric facts called measures which are categorized by dimensions. OLAP Cube is also called the **hypercube**.

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

Notes by Drashti Shrimal

and unrelated source. 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.

## 2. OLAP OPERATIONS.

Four types of analytical operations in OLAP are:

- 1. Roll-up
- 2. Drill-down
- 3. Slice and dice
- 4. Pivot (rotate)

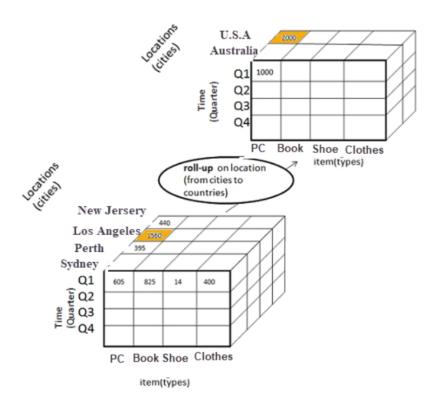
\_\_\_\_\_\_

#### 1) 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 the following diagram:



- In this example, cities New jersey and Lost Angles and 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.

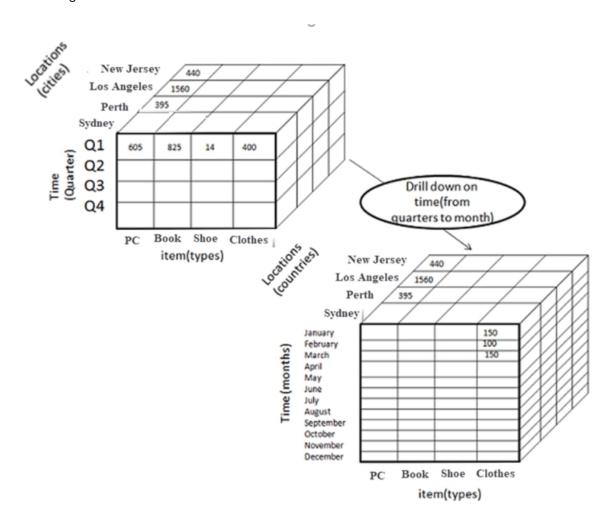
Notes by Drashti Shrimal

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

## 2) Drill-down

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



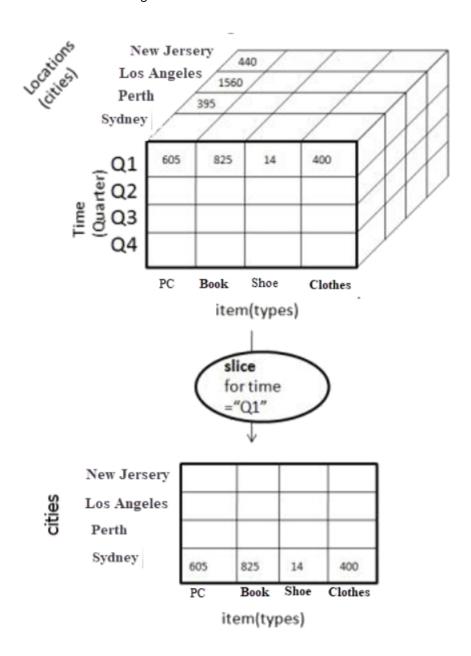
- Quater Q1 is drilled down to months January, February, and March. Corresponding sales are also registers.
- In this example, dimension months are added.

## 3) Slice:

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

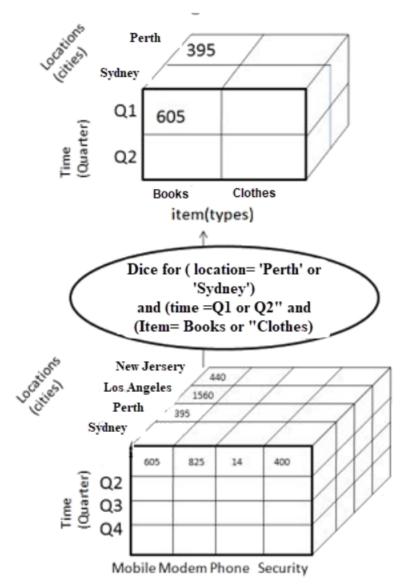
Following diagram explain how slice operation performed:

- Dimension Time is Sliced with Q1 as the filter.
- A new cube is created altogether.



#### Dice:

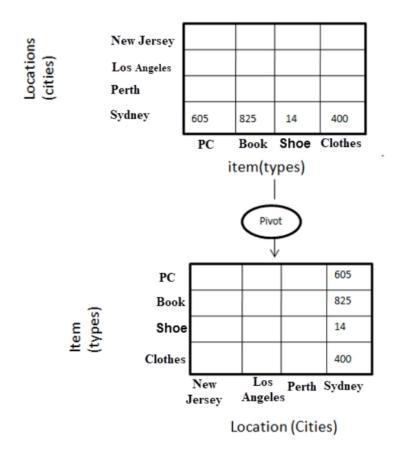
This operation is similar to a slice. The difference in dice is you select 2 or more dimensions that result in the creation of a sub-cube.



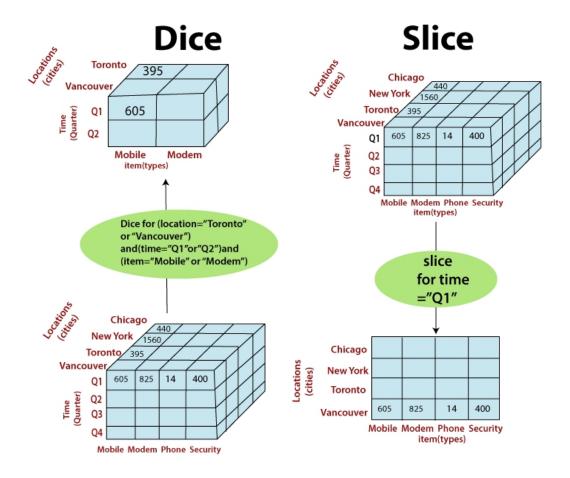
# 4) Pivot

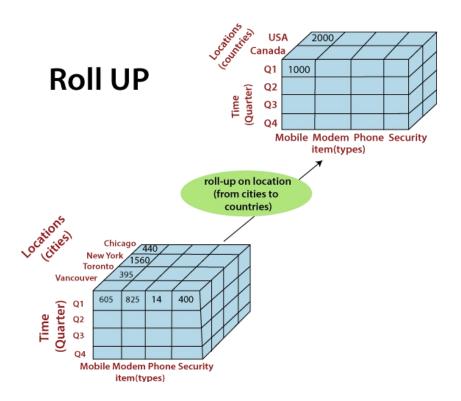
In Pivot, you rotate the data axes to provide a substitute presentation of data.

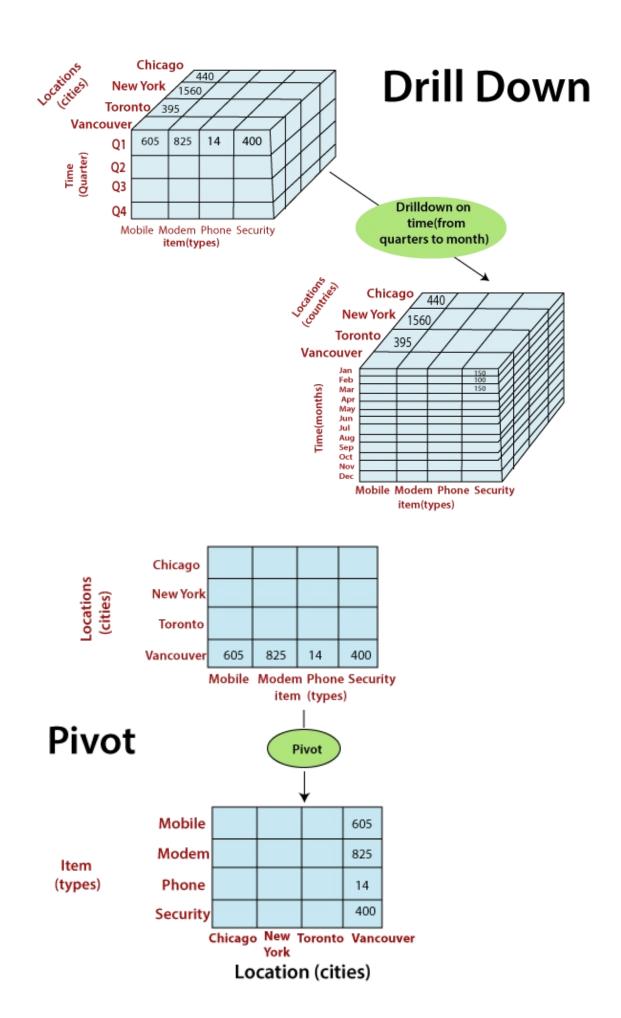
In the following example, the pivot is based on item types.



# Example 2:







### OLTP:

**OLTP** is an operational system that supports transaction-oriented applications in a 3-tier architecture. It administers the day to day transaction of an organization. OLTP is basically focused on query processing, maintaining data integrity in multi-access environments as well as effectiveness that is measured by the total number of transactions per second. The full form of OLTP is Online Transaction Processing.

# **Characteristics:**

- OLTP uses transactions that include small amounts of data.
- Indexed data in the database can be accessed easily.
- OLTP has a large number of users.
- It has fast response times
- Databases are directly accessible to end-users
- OLTP uses a fully normalized schema for database consistency.
- The response time of OLTP system is short.
- It strictly performs only the predefined operations on a small number of records.
- OLTP stores the records of the last few days or a week.
- It supports complex data models and tables.

### **Example:**

An example of the OLTP system is the ATM center. Assume that a couple has a joint account with a bank. One day both simultaneously reach different ATM centers at precisely the same time and want to withdraw the total amount present in their bank account.

However, the person that completes the authentication process first will be able to get money. In this case, the OLTP system makes sure that the withdrawn amount will be never more than the amount present in the bank. The key to note here is that OLTP systems are optimized for transactional superiority instead of data analysis.

# 3. OLTP AND OLAP DIFFERENCE.

| BASIS FOR<br>COMPARISON | OLTP   | OLAP   |
|-------------------------|--|--|
| Basic                   | It is an online<br>transactional system and<br>manages database<br>modification. | It is an online data retrieving and data analysis system.                                    |
| Focus                   | Insert, Update, Delete information from the database.                            | Extract data for analyzing that helps in decision making.                                    |
| Data                    | OLTP and its transactions are the original source of data.                       | Different OLTPs database becomes the source of data for OLAP.                                |
| Transaction             | OLTP has short transactions.   | OLAP has long transactions.  |
| Time                    | The processing time of a transaction is comparatively less in OLTP.              | The processing time of a transaction is comparatively more in OLAP.                          |
| Queries                 | Simpler queries.   | Complex queries.   |
| Normalization           | Tables in OLTP database are normalized (3NF).                                    | Tables in OLAP database are not normalized.  |
| Integrity               | OLTP database must maintain data integrity constraint.                           | OLAP database does not get<br>frequently modified. Hence,<br>data integrity is not affected. |