
UNIT 2
DATA WAREHOUSING
AND ITS APPLICATIONS

Syllabus

- **Data Warehouse Schema-**
- **Introduction-Dimensional Modeling**
- **Star -Snowflake Schema**
- **Aggregate Tables-**
- DBMS Schemas for Decision Support
- Data Extraction
- **Data transformation: Basic tasks-Major transformation types**
- **OLAP –**Hypercube-**OLAP operations**
- **OLAP models-MOLAP**
- Dimensional Analysis

Data Warehouse Schema

- Logical descriptions of database are known as Schema.
- It is the blueprint of the entire database.
- It defines how the data are organized and how the relations among them are associated.
- Data warehouse schema consists of the name and description of records including associated data items and aggregates.

Dimension

- Dimension is a collection of reference information about a measurable event.
- These events are stored in a fact table and are known as facts.
- The dimensions are generally the entities for which an organization wants to preserve records.
- The descriptive attributes are organized as columns in dimension tables by a data warehouse.
- For example, a student's dimension attributes could consist of first and last name, roll number, age, gender, or an address dimension that would include street name, state, and country attributes.

Data Warehouse Schema

Dimension

- A dimension table consists of a primary key column that uniquely identifies each record (row) of dimension.
- It is a framework that consists of one or more hierarchies that classify data.
- Dimensions are de-normalized tables and may have redundant data.
- Normalization is a process of breaking up a larger table into smaller tables free of any possible insertion, updation or deletion anomalies.
- In de-normalization, smaller tables are merged to form larger tables to reduce joining operations.
- De-normalization is performed in those cases where retrieval is a major requirement and insert, update, and delete operations are minimal, as in case of historical data or data warehouse.

location_id is the primary key

item_code as primary key

Location Dimension
Locaton_id
Street_name
City
State_id
Country_code

Item Dimension
Item_code
Item_name
Item_type
Brand_name
Supplier_id

Data Warehouse Schema

Measure

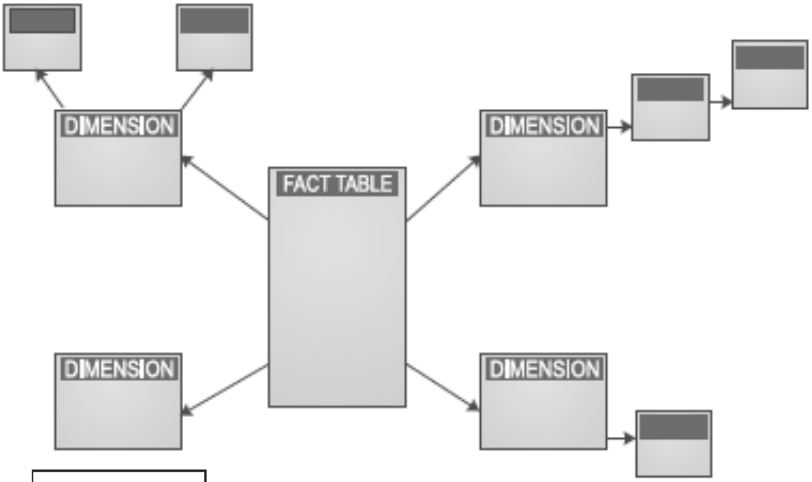
- Measure is used for the values that rely on the dimensions. For example, amount sold, quantity sold, etc.

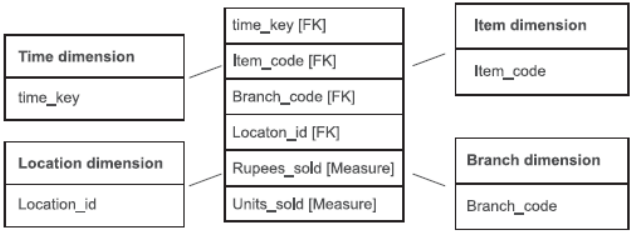
Fact Table

- A ‘fact table’ is a group of associated data items. It consists of values of dimensions and measure.
- This means that a fact table can be defined from the given dimension and measure.
- A fact table consists of two types of columns such as foreign keys and measure.
- Foreign keys are linked to dimension tables and measures consist of numeric facts.
- Fact tables are generally larger in size than dimension tables.
- A fact table can hold a dataset of facts at detailed or aggregated level.

Data Warehouse Schema

Fact Table





Multi-dimensional view of data

- Data consists of multiple dimensions.
- Dimensions often have hierarchies to show parent-child relationships.

DIMENSIONAL MODELING

- Dimensional modeling is a data model design adopted when building a data warehouse.
- A data model is a representation of how data is stored in a database, and it is usually a diagram of the few tables and the relationships that exist between them. This modeling is designed to read, summarize and compute some numeric data from a data warehouse.
- The dimension modeling provide an easy architecture for the end user to write queries and, to reduce the number of relationships between the tables and dimensions hence providing efficient query handling.
- Dimensional modeling populates data in a cube as a logical representation with OLAP data management.
- The concept was developed by **Ralph Kimball**.
- The transaction record is divided into either “facts”, which consists of business numerical transaction data, or “dimensions”, which are the reference information that gives context to the facts.

DIMENSIONAL MODELING

1: Identify the Business Objectives: Selection of the right business process to build a data warehouse. This is very important step otherwise this can lead to repeated process and software defects.

2: Identifying Granularity: The grain literally means each minute detail of the business problem. This is decomposing of the large and complex problem into the lowest level information. This affects the size of the data warehouse.

3: Identifying Dimensions and attributes: The dimensions of the data warehouse can be understood by the entities of the database. like, items, products, date, stocks, time etc. The identification of the primary keys and the foreign keys specifications all are described here.

4: Build the Schema: The database structure or arrangement of columns in a database table, decides the schema. There are various popular schemas like, star, snowflake, fact constellation - summarizing, from the selection of business process to identifying each finest level of detail of the business transactions. Identifying the significant dimensions and attributes would help to build the schema.

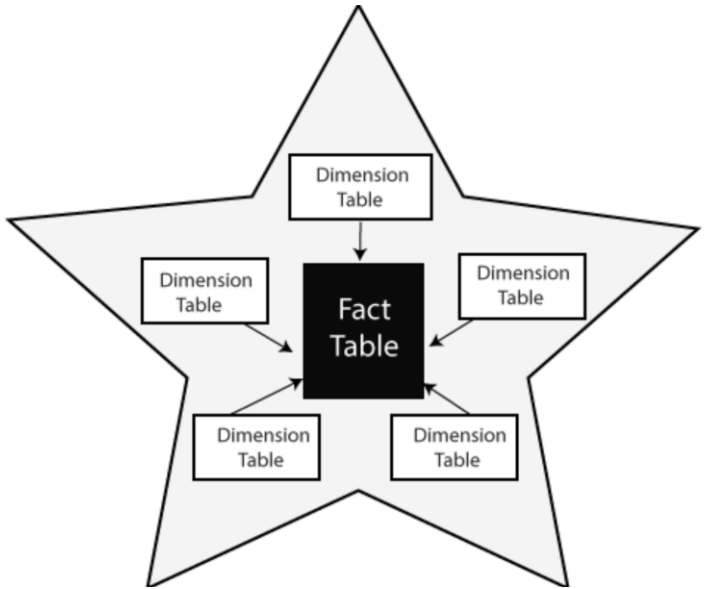
STRENGTHS OF DIMENSIONAL MODELING

- It provides simplicity of architecture/schema to understand and handle various stakeholders from warehouse designers to business clients.
- It reduces the number of relationships between different data elements.
- It promotes data quality by enforcing foreign key constraints as a form of referential integrity check on a data warehouse. The dimensional modeling helps the database administrators to maintain the reliability of the data.
- The aggregate functions used in the schemas optimize the query performance posted by the customers. Since data warehouse size keeps on increasing and with this increased size, the optimization becomes the concern which dimension modeling makes it easy.

Aspect	Facts and Fact Table	Dimensions and Dimension Table
Definition	Fact represents an event or measure, typically numeric, representing business items or transactions with associated context data. Fact Table contains descriptions of primary keys from tables used in business processes, acting as foreign keys, and includes aggregate functions to compute business processes.	Dimensions describe one business dimension, providing contextual background for facts. Dimension tables establish the context of facts and store fields describing them.
Relationship with Data	Facts represent the measurable aspects of business events or transactions. Fact tables store data related to these facts and their context.	Dimensions serve as a framework for facts, providing the context over which OLAP (Online Analytical Processing) is performed. Dimension tables store data describing these dimensions and establish context for facts.
Structure	The number of columns in a fact table is less than that in dimension tables.	Dimension tables often contain many columns compared to fact tables.
Normalization	Fact tables are usually more normalized compared to dimension tables.	Dimension tables are often denormalized to provide faster query performance.
Attributes	Attributes in a fact table represent numeric measures relative to dimensions.	Attributes in a dimension table describe the characteristics of the dimension and are used as row and column headings in query results.
Purpose	Fact tables store quantitative data that can be analyzed to derive insights into business performance.	Dimension tables provide the necessary context for the facts stored in the fact table and enable meaningful analysis of data.

Star Model

- It represents the multidimensional model.
- The data is organized into facts and dimensions.
- Star schemas are easy for end users and applications to understand and navigate.
- It has one broad central table (fact table) and a set of smaller tables (dimensions) arranged in a star design.



Star Model

Features of Star Schema

- The data is in denormalized database.
- It provides quick query response
- Star schema is flexible can be changed or added easily.
- It reduces the complexity of metadata for developers and end users.

Advantages of Star Schema

Easily understood

- A star schema is easy to understand and navigate, with dimensions joined only through the fact table.
- These joins are more significant to the end user, because they represent the fundamental relationship between parts of the underlying business.
- Users can also browse dimension table attributes before constructing a query.

Star Model

Advantages of Star Schema

Built-in referential integrity

- A star schema has referential integrity built in when data is loaded.
- Referential integrity is enforced because each record in a dimension table has a unique primary key, and all keys in the fact tables are legitimate foreign keys drawn from the dimension tables.
- A record in the fact table that is not related correctly to a dimension cannot be given the correct key value to be retrieved.

Load performance and administration

- Structural simplicity also reduces the time required to load large batches of data into a star schema database.
- By defining facts and dimensions and separating them into different tables, the impact of a load operation is reduced.
- Dimension tables can be populated once and occasionally refreshed.
- New facts are added regularly and selectively by appending records to a fact table.

Star Model

Advantages of Star Schema

Query performance

- Star schema database has a small number of tables and clear join paths, hence queries run faster than they do against an OLTP system.
- Small single table queries, usually of dimension tables.
- Large join queries that involve multiple tables take only seconds or minutes to run.
- In a star schema database design, the dimensions are linked only through the central fact table.
- When 2D tables are used in a query, only one join path, intersecting the fact table, exists between those two tables.
- This design feature enforces accurate and consistent query results.

Star Model

Disadvantages of Star Schema

Decreased data integrity:

- Because of the denormalized data structure, star schemas do not enforce data integrity very well.
- Although star schemas use countermeasures to prevent anomalies from developing, a simple insert or update command can still cause data incongruities.

Less capable of handling diverse and complex queries:

- Databases designers build and optimize star schemas for specific analytical needs.
- As denormalized data sets, they work best with a relatively narrow set of simple queries.
- Comparatively, a normalized schema permits a far wider variety of more complex analytical queries.

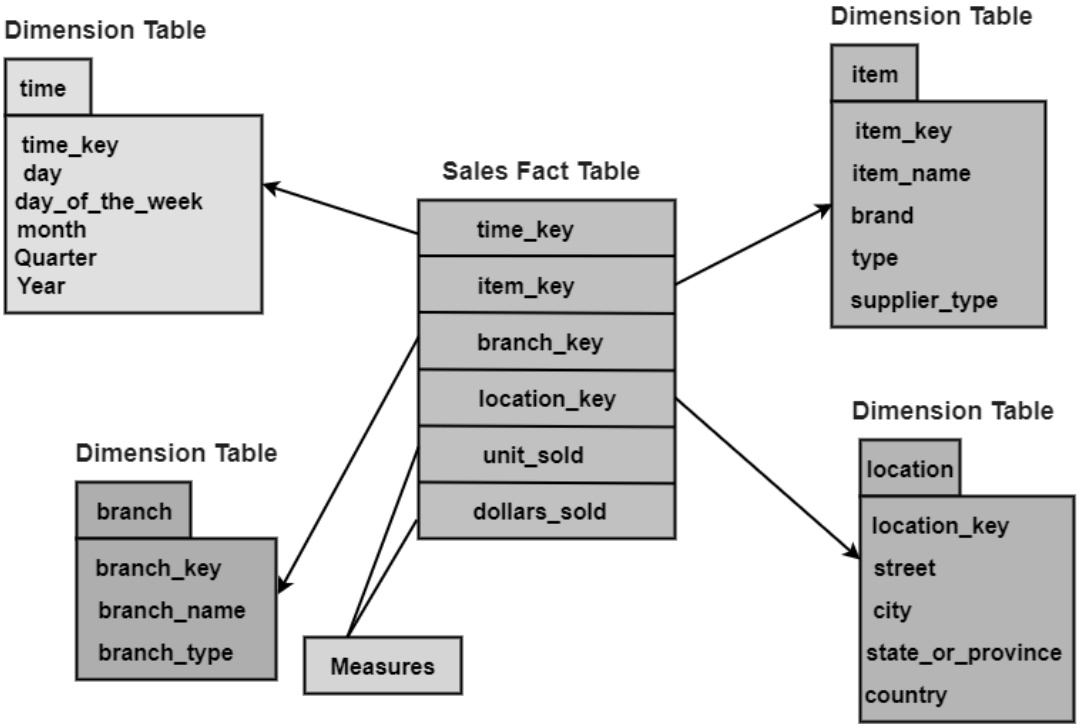
No Many-to-Many Relationships:

- Because they offer a simple dimension schema, star schemas don’t work well for “many-to-many data relationships”

Star Model

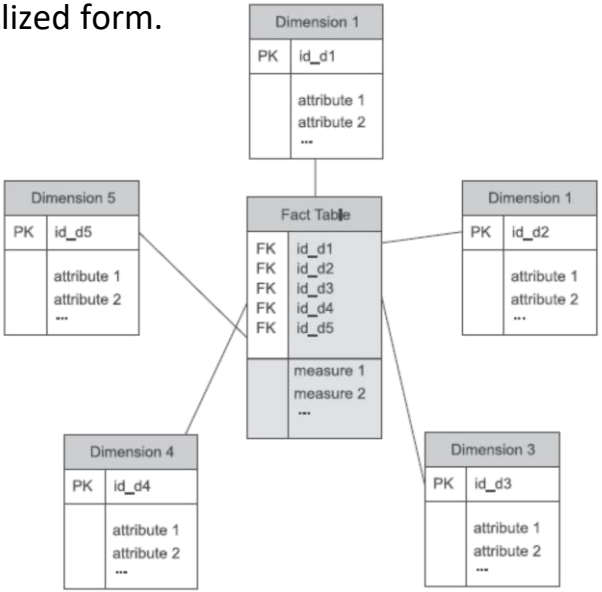
- Star schema is composed of a **Sales** fact and several dimension tables connected to it for Time, Branch, Item and Location.
- **Fact Table- Sales** is the Fact table.
- **Time** table has a column for each day, month, quarter, year etc..
- **Item** table has columns for each item_key, item_name, brand, type and supplier_type.
- The **Branch** table has columns for each branch_key, branch_name and branch_type.
- **Location** table has columns of geographic data, including street, city, state, and country. Unit_Sold and Dollars_Sold are the Measures.

Star Model

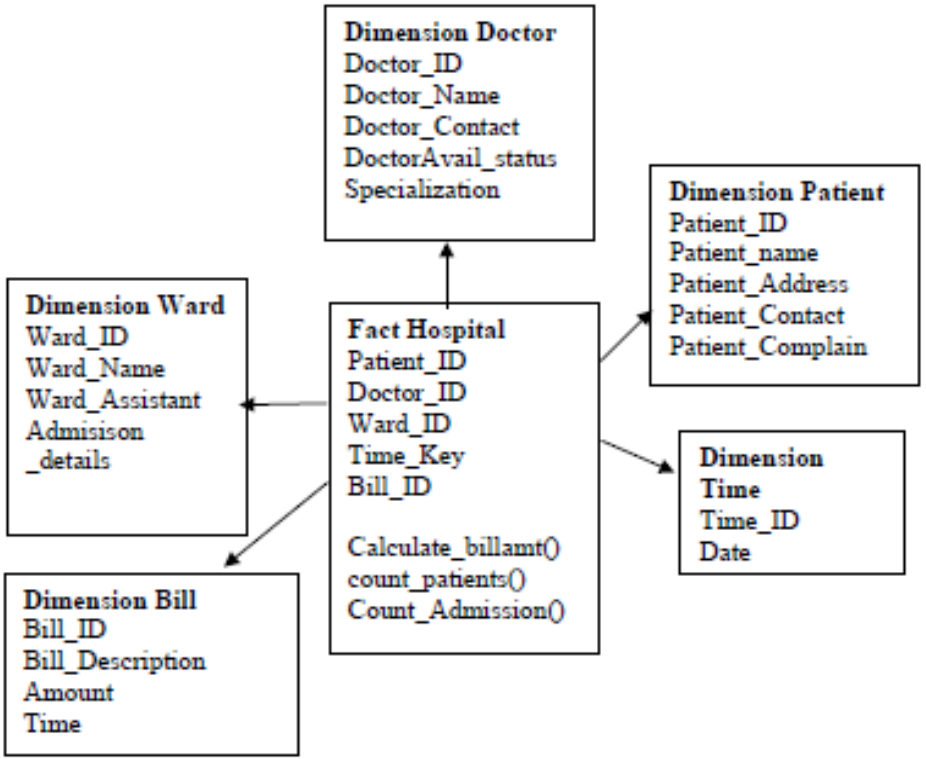

Dimension Table
time
time_key
day
day_of_the_week
month
Quarter
Year
Dimension Table
item
item_key
item_name
brand
type
supplier_type
Sales Fact Table
time_key
item_key
branch_key
location_key
unit_sold
dollars_sold
Dimension Table
branch
branch_key
branch_name
branch_type
Measures
Dimension Table
location
location_key
street
city
state_or_province
country

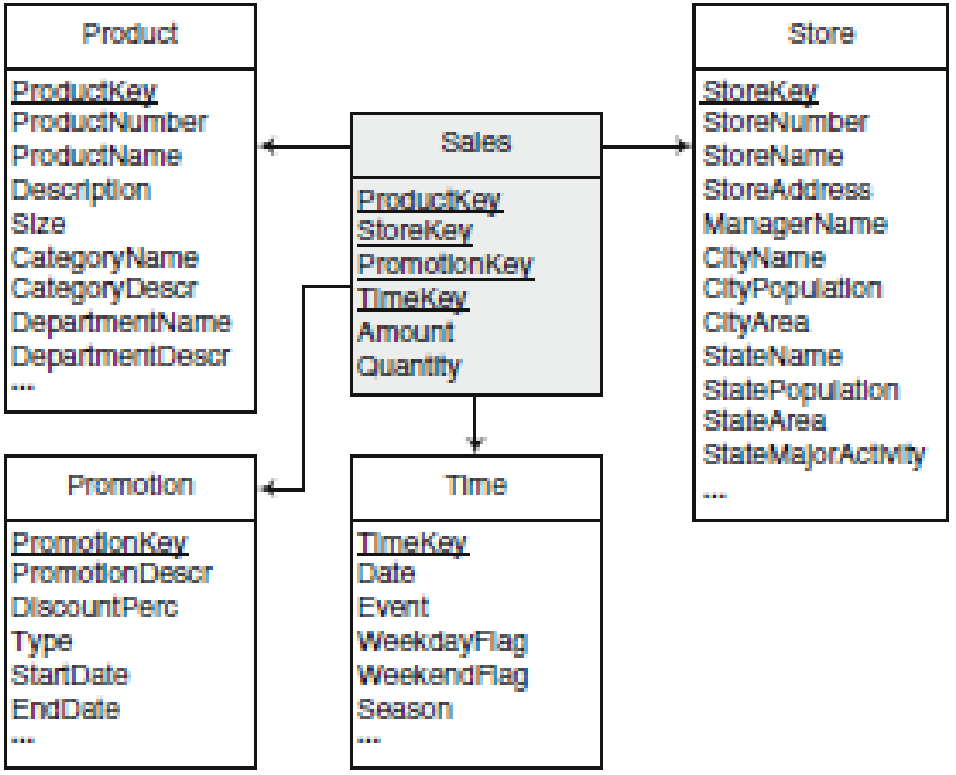
Star Model

- Each dimension in a star schema represents a one-dimensional table only and the dimension table consists of a set of attributes.
- Dimension tables comprise of relatively small numbers of records in comparison to fact tables, but each record may consist of many attributes to describe the fact data.
- Fact tables usually consist of numeric facts and foreign keys to dimensional data.
- Fact tables are in third normal form (3NF) in the case of star schema while dimensional tables are in de-normalized form.



a. Star Schema of Hospital Management



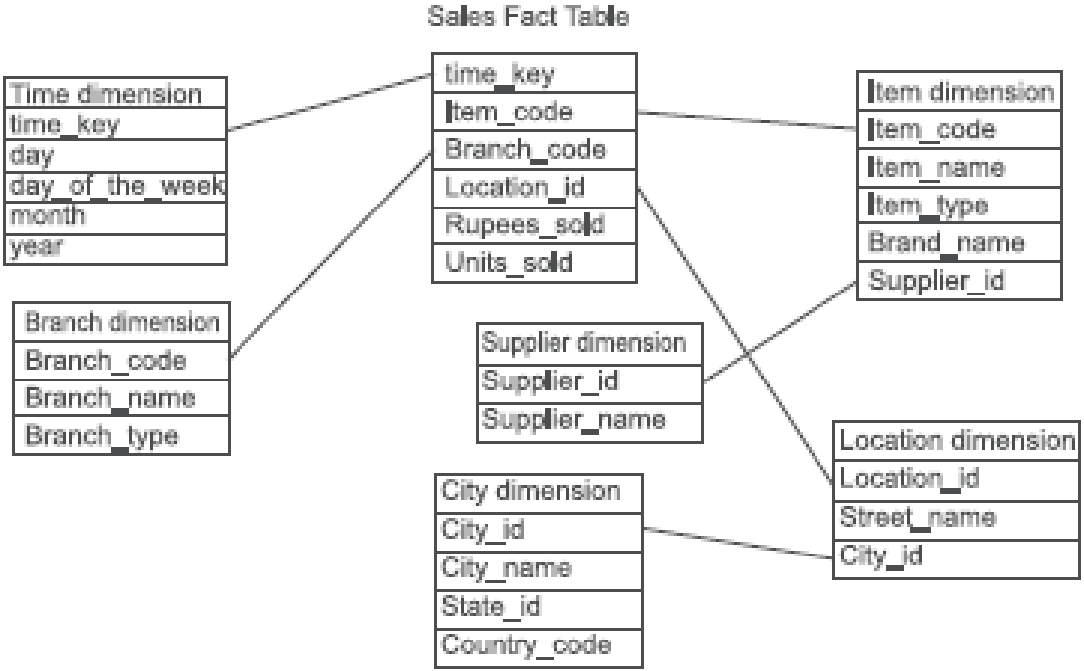


Fact table Sales is normalized, and its key is composed by ProductKey, StoreKey, PromotionKey, and TimeKey.

SNOWFLAKE SCHEMA

- Snowflake schema is the extension of star schema which adds more dimensions to give more meaning to the logical view of the database.
- These additional tables are more normalized than star schema.
- The arrangement of data is like that the centralized fact table relates to multiple related dimensional tables.
- This can become more complex if the dimensions are more detailed and at multiple levels.
- It is a logical arrangement of tables in a multidimensional database such that the ER diagram resembles a snowflake shape.

SNOWFLAKE SCHEMA



SNOWFLAKE SCHEMA

Advantages of Snowflake schema:

- Occupies lesser disk space means more convenience and less hassle.
- Snowflake schema because of how safe if it is.
- Data is easy to maintain and more structured.
- Data quality is better than star schema.

Disadvantages of Snowflake Schema

Complex data schemas:

- In storing data, snowflake result in performance declines while browsing complex joins.
- Still, processing technology advancements have resulted in improved snowflake schema query performance.

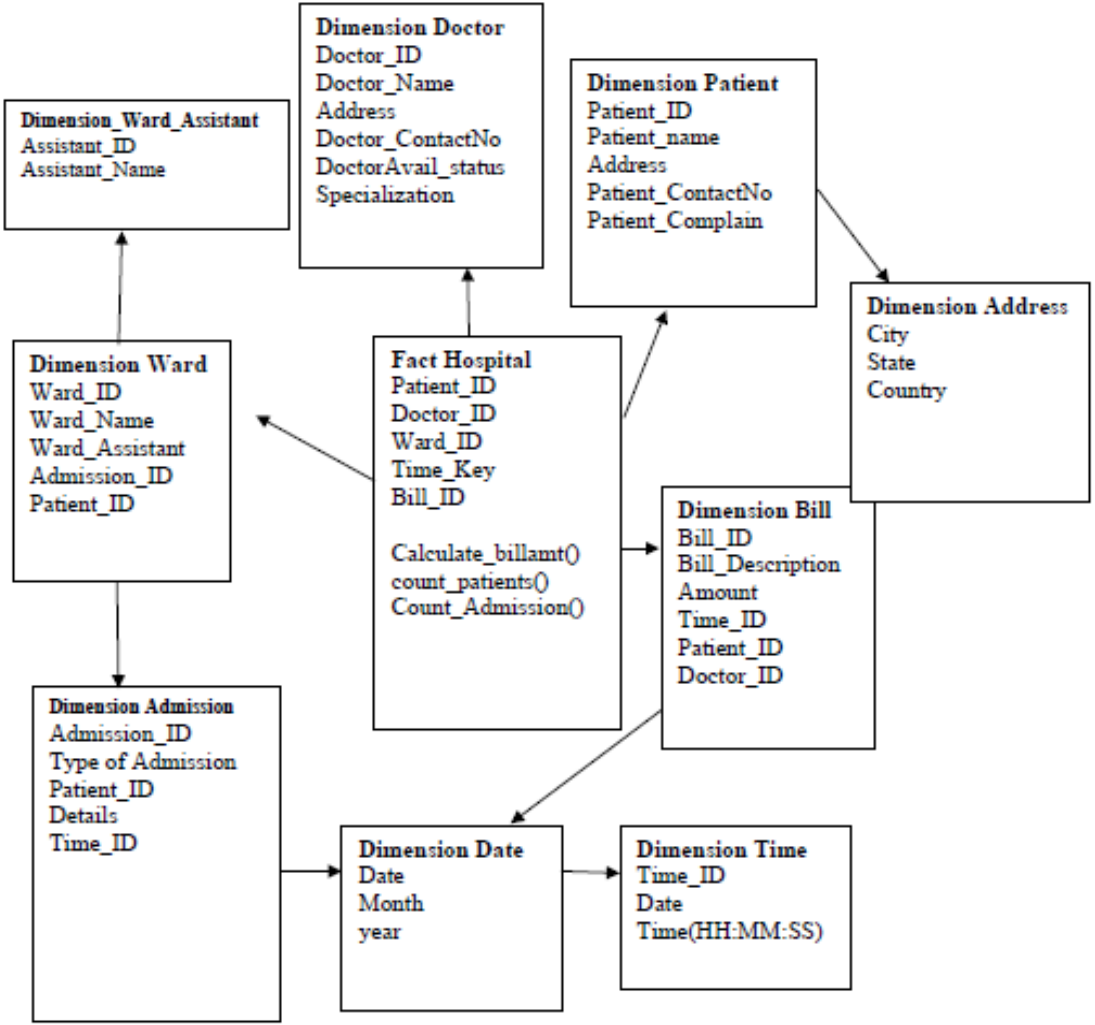
Slower at processing cube data:

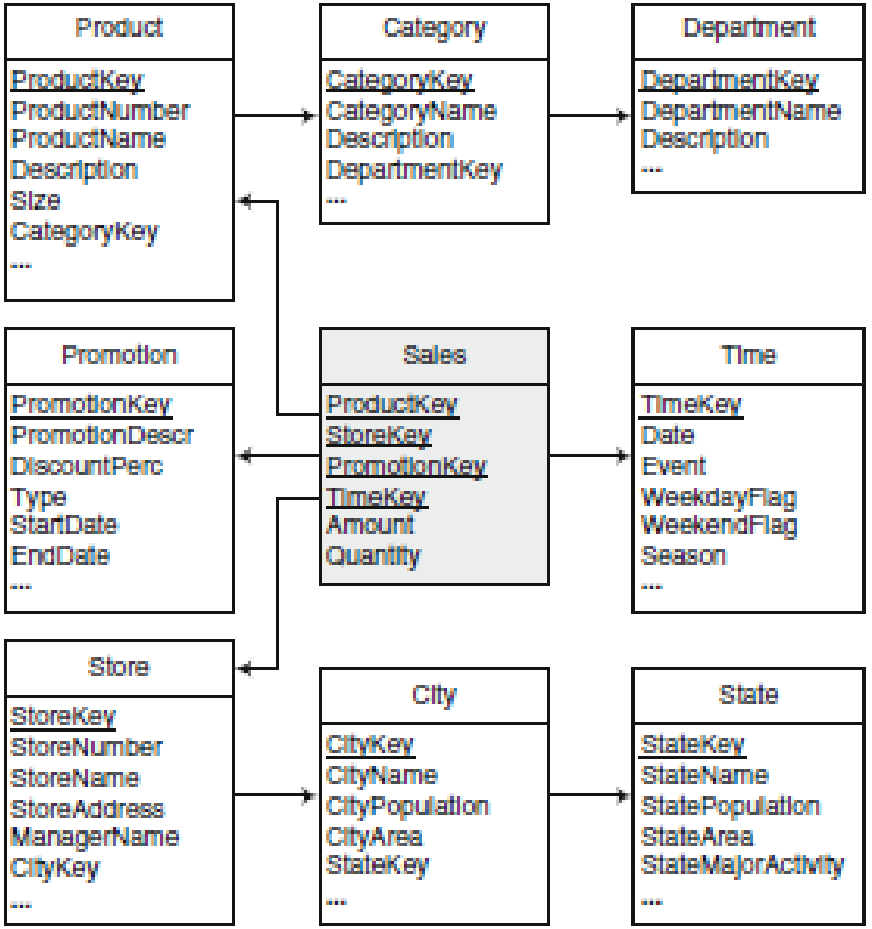
- Complex joins result in slower cube data processing.

Lower data integrity levels:

- While snowflake schemas offer greater normalization and fewer risks of data corruption after performing UPDATE and INSERT commands, they do not provide the level of transnational assurance that comes with a traditional, highly normalized database structure.
- Therefore, when loading data into a snowflake schema, it's vital to be careful and double-check the quality of information post-loading.

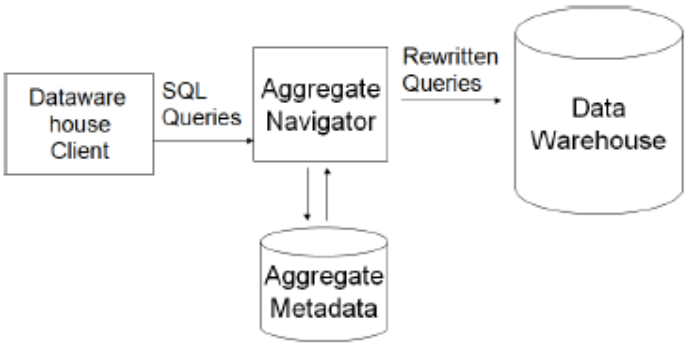
b. Snowflake Schema of Hospital Management



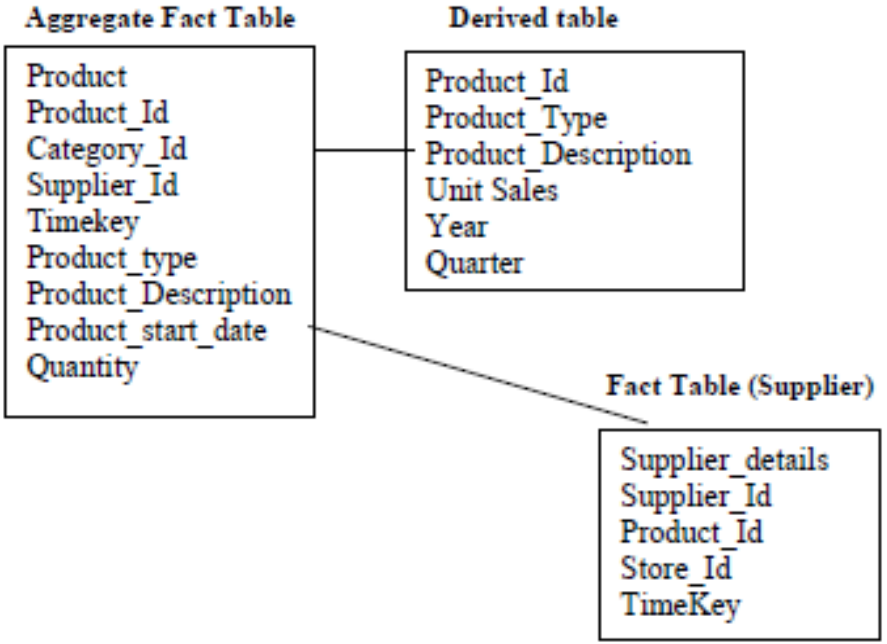

Product
ProductKey
ProductNumber
ProductName
Description
Size
CategoryKey
...
Category
CategoryKey
CategoryName
Description
DepartmentKey
...
Department
DepartmentKey
DepartmentName
Description
...
Promotion
PromotionKey
PromotionDescr
DiscountPerc
Type
StartDate
EndDate
...
Sales
ProductKey
StoreKey
PromotionKey
TimeKey
Amount
Quantity
Time
TimeKey
Date
Event
WeekdayFlag
WeekendFlag
Season
...
Store
StoreKey
StoreNumber
StoreName
StoreAddress
ManagerName
CityKey
...
City
CityKey
CityName
CityPopulation
CityArea
StateKey
...
State
StateKey
StateName
StatePopulation
StateArea
StateMajorActivity
...

AGGREGATE TABLES/Summary tables

- Aggregate is like the aggregation of the database schemas of relational tables.
- Aggregate fact tables roll up the basic fact tables of the schema to improve the query processing.
- The business tools smoothly select the level of aggregation to improve the query performance.
- Aggregate fact tables contain foreign keys referring to dimension tables.
- It contains pre-computed queries of the data warehouse schema.
- It reduces the dimensionality of the base fact tables.
- It can be used to respond to the queries of the dimensions that are saved.
- Aggregate tables also referred to pre-computed tables having partially summarized data.
- It occupies less space than atomic fact tables. It nearly takes the half time of a general query processing.



AGGREGATE TABLES/Summary tables



Data Extraction

- Data extraction, which typically gathers data from multiple, heterogeneous, and external sources.
- Two major factors differentiate the data extraction for a new operational system from the data extraction for a datawarehouse.
- First, for a datawarehouse, you must extract data from many disparate sources.
- Next, for a data warehouse, you must extract data on the changes for ongoing incremental loads as well as for a one-time initial full load.

Data extraction issues:

- Source identification—identify source applications and source structures.
- Method of extraction—for each data source, define whether the extraction process is manual or tool-based.
- Extraction frequency—for each data source, establish how frequently the data extraction must be done: daily, weekly, quarterly, and so on.
- Time window—for each data source, denote the time window for the extraction process.
- Job sequencing—determine whether the beginning of one job in an extraction job stream must wait until the previous job has finished successfully.
- Exception handling—determine how to handle input records that cannot be extracted.

Basic Data Transformation

Selection.

- This is the beginning of the whole process of data transformation.
- Either whole records or parts of several records from the source systems is selected.
- The task of selection usually forms part of the extraction function itself.
- In some cases, the composition of the source structure may not be amenable to selection of the necessary parts during data extraction.
- In these cases, it is prudent to extract the whole record and then do the selection as part of the transformation function.

Splitting/Joining.

- This includes the types of data manipulation needed to be performed on the selected parts of source records.
- Sometimes , during data transformation the selected parts can be splitted even further.
- Joining of parts selected from many source systems is more widespread in the data warehouse environment.

Conversion.

- This is an all-inclusive task.
- It includes a large variety of rudimentary conversions of single fields for two primary reasons—one to standardize among the data extractions from disparate source systems, and the other to make the fields usable and understandable to the users.

Summarization.

- Sometimes it is not feasible to keep data at the lowest level of detail in your data warehouse.
- It may be that none of your users ever need data at the lowest granularity for analysis or querying.

Data Transformation

Enrichment.

- This task is the rearrangement and simplification of individual fields to make them more useful for the data warehouse environment.
- You may use one or more fields from the same input record to create a better view of the data for the data warehouse.
- This principle is extended when one or more fields originate from multiple records, resulting in a single field for the data warehouse.

Major data transformation types

Format Revisions.

- The revisions include changes to the data types and lengths of individual fields.
- In source systems, product package types may be indicated by codes and names in which the fields are numeric and text data types.
- The lengths of the package types may vary among the different source systems.
- It is wise to standardize and change the data type to text to provide values meaningful to the users.

Decoding of Fields.

- When dealing with multiple source systems, the same data items can be described by a plethora of field values.
- The classic example is the coding for gender, 1 and 2 for male and female and another system using M and F.
- Also, many legacy systems are notorious for using cryptic codes to represent business values.
- You need to decode all such cryptic codes and change these into values that make sense to the users.
- Change the codes to Active, Inactive, Regular, and Suspended.

Calculated and Derived Values.

- The extracted data from the sales system contains sales amounts, sales units, and operating cost estimates by product.
- The total cost and the profit margin before data can be stored in the datawarehouse.
- Average daily balances and operating ratios are examples of derived fields.

Splitting of Single Fields.

- Earlier legacy systems stored names and addresses of customers and employees in large text fields.
- The first name, middle initials, and last name were stored as a large text in a single field.
- Similarly, some earlier systems stored city, state, and zip code data together in a single field.
- You need to store individual components of names and addresses in separate fields in your data warehouse for two reasons.
- First, you may improve the operating performance by indexing on individual components.
- Second, your users may need to perform analysis by using ual components such as city, state, and zip code.

Major data transformation types

Merging of Information.

- This type of data transformation does not literally mean the merging of several fields to create a single field of data.
- For example, information about a product may come from different data sources.
- The product code and description may come from one data source.
- The relevant package types may be found in another data source.
- The cost data may be from yet another source.
- In this case, merging of information denotes the combination of the product code, description, package types, and cost into a single entity.

Character set conversion.

- This type of data transformation relates to the conversion of character sets to an agreed standard character set for textual data in the data warehouse.
- If mainframe legacy systems are source systems, the source data from these systems will be in EBCDIC characters.
- If PC-based architecture data warehouse is used, then convert the mainframe EBCDIC format to the ASCII format.
- When source data is on other types of hardware and operating systems, similar character set conversions are used.

Conversion of Units of Measurements.

- Many companies today have global branches.
- Measurements in many European countries are in metric units.
- If company has overseas operations, convert the metrics so that the numbers are all in one standard unit of measurement.

Date/Time Conversion.

- This type relates to representation of date and time in standard formats.
- For example, the American and the British date formats may be standardized to an international format.
- The date of October 11, 2008 is written as 10/11/2008 in the U.S. format and as 11/10/2008 in the British format.
- This date may be standardized to be written as 11 OCT 2008.

Major data transformation types

Summarization.

- This type of transformation is the creating of summaries to be loaded in the data warehouse instead of loading the most granular level of data.
- For example, for a credit card company to analyze sales patterns, it may not be necessary to store in the data warehouse every single transaction on each credit card. Instead, summarize the daily transactions for each credit card and store the summary data instead of storing the most granular data by individual transactions.

Key Restructuring.

- While extracting data from input sources, look at the primary keys of the extracted records.
- You will have to produce keys for the fact and dimension tables based on the keys in the extracted records.
- In the example ,the product code in this organization is structured to have inherent meaning.
- If product code is used as the primary key, there will be problems.
- If the product is moved to another warehouse, the warehouse part of the product key will have to be changed.
- This is a typical problem with legacy systems.
- When choosing keys for your data warehouse database tables, avoid such keys with built-in meanings.
- Transform such keys into generic keys generated by the system itself. This is called key restructuring.

Deduplication.

- In many companies, the customer files have several records for the same customer.
- Mostly, the duplicates are the result of creating additional records by mistake.
- In data warehouse, keep a single record for one customer and link all the duplicates in the source systems to this single record.
- This process is called deduplication of the customer file.
- Employee files and, sometimes, product master files have this kind of duplication problem.

Online Analytical Processing (OLAP)

- OLAP is a design paradigm that provides a method to extract useful information from a physical data store.
- It aggregates information from multiple systems and provides summarized information/ view to the management, while data mining is used to find hidden patterns within the data.
- OLAP summarizes data and makes forecasts.
- The key benefits offered by OLAP are:
    - Increased productivity of end-users.
    - Retention of organizational control over the integrity of corporate data.
    - Reduced backlog of applications development for IT staff.
    - Improved profitability and potential revenue.
    - Reduced query drag and network traffic on the data warehouse or OLTP systems.
- The major strengths of any OLAP are listed below.
    - It is a powerful visualization tool.
    - OLAP tools are good for analyzing time series.
    - It provides fast, interactive response time.
    - OLAP tools are offered by many vendors.
    - It can be of help in identifying outliers and clusters.

Online Analytical Processing (OLAP)

- It is a software which helps the data analysts to collect data from different perspective for developing effective business strategies.
- The query operations like group, join or aggregation can be easily done with OLAP using pre-calculated or preaggregated data hence making it much faster than simple relational databases.
- OLAP is a multi cubic structure, which has many cubes, each cube is pertaining to some database.
- The cubes are designed in such a way that generates reports effectively and efficiently.
- OLAP is the core component of the data warehouse implementation, providing fast and flexible multi-dimensional data analysis for business intelligence (BI) and decision support applications.
- OLAP (for online analytical processing) is a software used to perform high-speed, multivariate analysis of large amounts of data in data warehouses, data markets, or other unified and centralized data warehouses.
- The data is broken down for display, monitoring or analysis.

Online Analytical Processing (OLAP)

The main characteristics of OLAP are as follows:

- **Fast**: OLAP act as bridge between Datawarehouse and front-end. Hence helps in the better accessibility of data yielding faster results.
- **Analysis**: OLAP data analysis and computational measure and their results are stored in separate data files. OLAP distinguishes better zero and missing values. It should ignore missing value and performs the correct aggregate values. OLAP facilitates interactive query handling and complex analysis for the users.
- **Shared**: OLAP operations drill-down or roll-up, it navigates between various dimensions in multidimensional cube making it effective and efficient reporting system.
- **Multidimensional**: OLAP has Multidimensional conceptual view and access of data to different users at different levels. The increasing number of dimensions and report generation performance of the OLAP system does not significantly degrade.
- **Data and Information**: OLAP has calculation power for complex queries anddata. It does data visualization using graphs and charts.

BASIC FEATURES

Multidimensional analysis	Consistent performance	Fast response times for interactive queries
Drill-down and roll-up	Navigation in and out of details	Slice-and-dice or rotation
Multiple view modes	Easy scalability	Time intelligence (year-to-date, fiscal period)

ADVANCED FEATURES

Powerful calculations	Cross-dimensional calculations	Pre-calculation or pre-consolidation
Drill-through across dimensions or details	Sophisticated presentation & displays	Collaborative decision making
Derived data values through formulas	Application of alert technology	Report generation with agent technology

Hypercube

- 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.
- A hypercube is a general metaphor for representing multidimensional data.
- So, we can say that multidimensional Databases can we see hypercube and multi cube.
- Multidimensional cubes have smaller multiple cubes and in hypercube it seems there is one cube as logically all the data seems to be as one unit of cube.
- Hypercube have multiple same dimensions logically.
- A Data Cube is described in terms of dimensions and facts. It represents data in different dimensions.
- Here, the dimensions are generally the entities which an organization preserves as records.
- This intuitive representation is a hypercube, a representation that accommodates more than three dimensions.
- At a lower level of simplification, a hypercube can very well accommodate three dimensions.
- A hypercube is a general metaphor for representing multidimensional data.

Hypercube

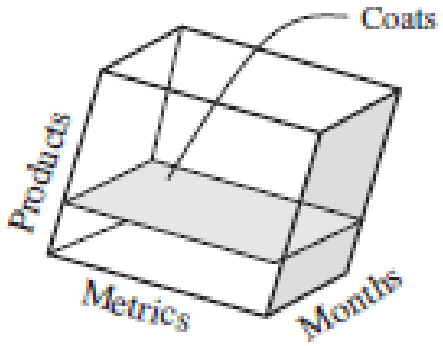
	Multi Cube	Hyper Cube
Metadata	Each dimension can belong to many cubes	Each dimension belongs to one cube only
Dimension	Not necessary all the dimensions should belong to some cube	Every dimension owned by a hypercube
Measure Computation	Complex, data can be retrieved from the all the cubes	Simple, as all the numerical facts are available at one place
Multiple	multicube system, if there are two rows in the DIMENSIONS rowset for which the DIMENSION_NAME value is the same (and the CUBE_NAME value is different), these two rows represent the same dimension. As, sub cubes are built from the same pool of available dimensions.	in a multiple hypercube scenario, it is possible for two hypercubes to have a dimension of the same name, each of which has different characteristics. In this case, the DIMENSION_UNIQUE_NAME value is guaranteed to be different.

PRODUCT: Coats

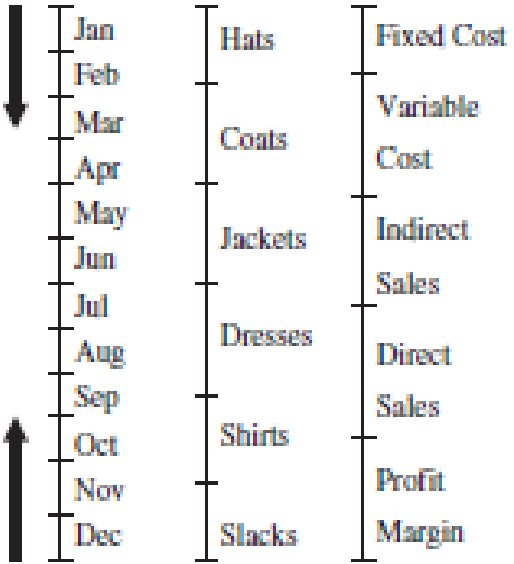
PAGES: PRODUCT dimension COLUMNS: Metrics

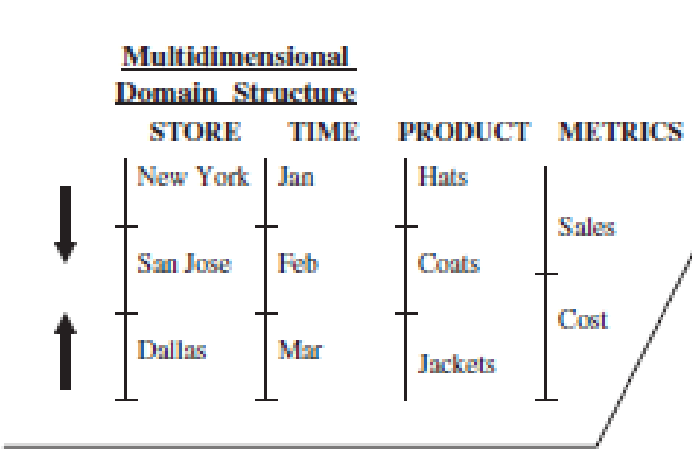
ROWS: TIME dimension

	Fixed Cost	Variable Cost	Indirect Sales	Direct Sales	Profit Margin
Jan	340	110	230	320	100
Feb	270	90	200	260	100
Mar	310	100	210	270	70
Apr	340	110	210	320	80
May	330	110	230	300	90
Jun	260	90	150	300	100
Jul	310	100	180	300	70
Aug	380	130	210	360	60
Sep	300	100	180	290	70
Oct	310	100	170	310	70
Nov	330	110	210	310	80
Dec	350	120	200	360	90



Multidimensional Domain Structure





HOW DISPLAYED ON A PAGE

PAGE: Store Dimension

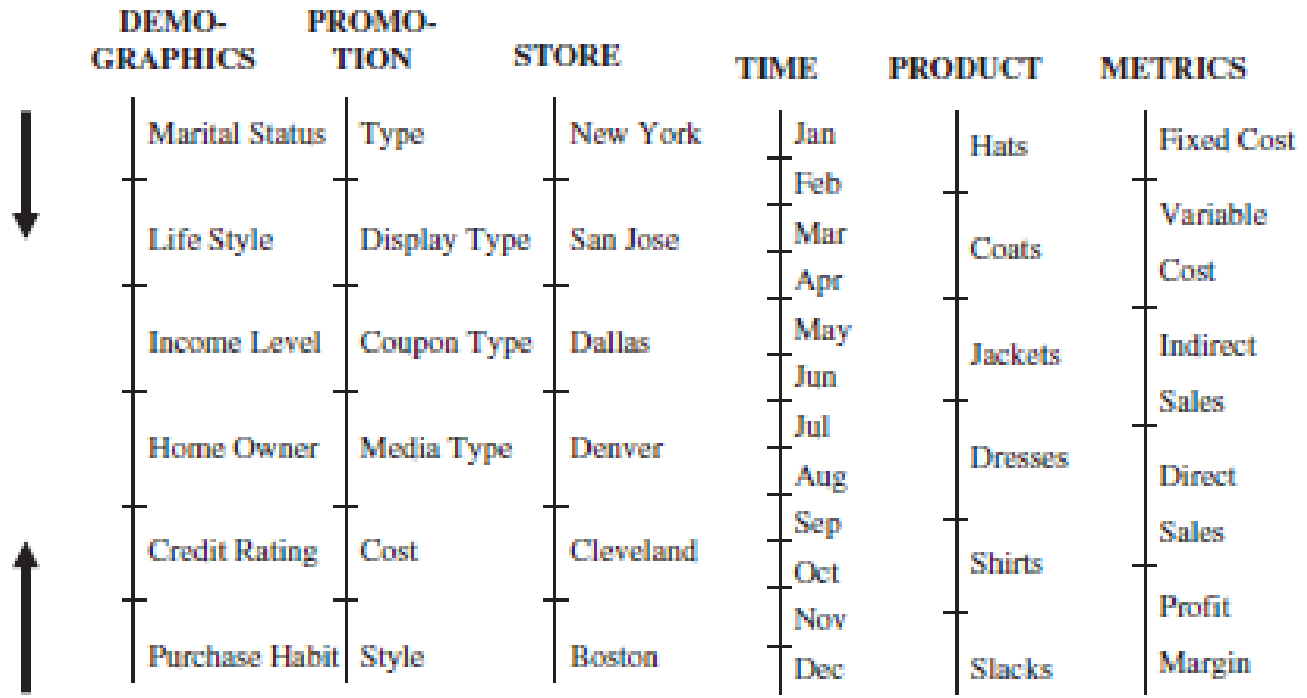
ROWS: Time Dimension

COLUMNS: Product & Metrics combined

New York Store

	Hats:Sales	Hats:Cost	Coats:Sales	Costs:Cost	Jackets:Sales	Jackets:Cost
Jan	450	350	550	450	500	400
Feb	380	280	460	360	400	320
Mar	400	310	480	410	450	400

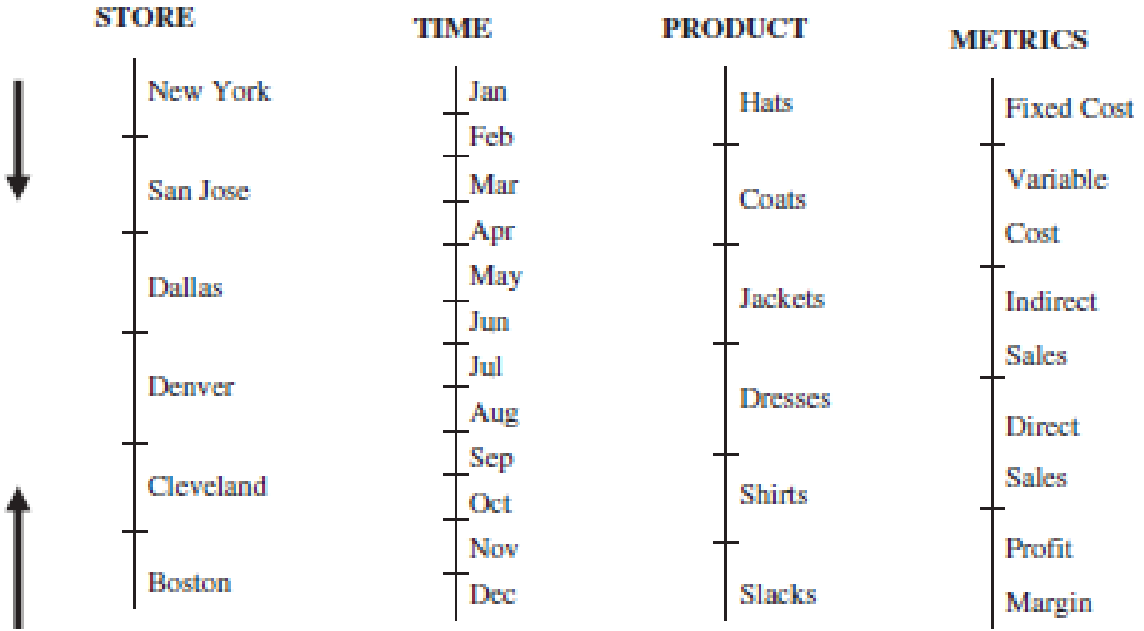
Multidimensional Domain Structure



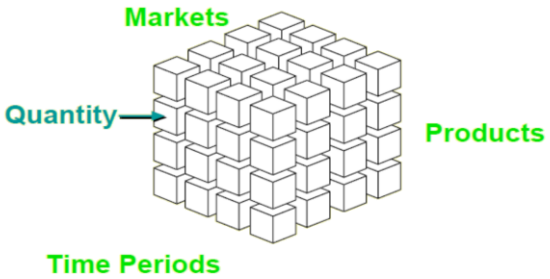
Hypercube

- The figure also shows a cube representing the data points along the edges.
- Relate the three straight lines to the three edges of the physical cube.
- Now the page you see in the figure is a slice passing through a single product and the divisions along the other two straight lines shown on the page as columns and rows.
- With three groups of data—two groups of business dimensions and one group of metrics—we can easily visualize the data as being along the three edges of a cube.
- Now add another business dimension to the model. Let us add the store dimension.
- That results in three business dimensions plus the metrics data.

Multidimensional Domain Structure

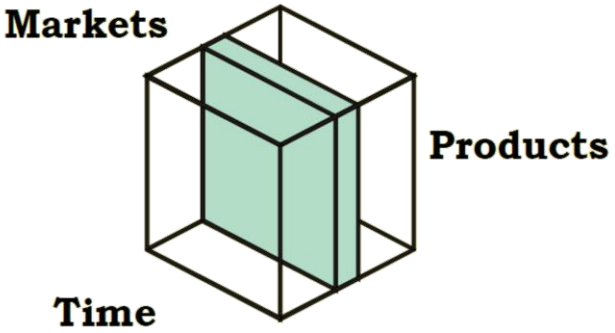


Multidimensional Representation

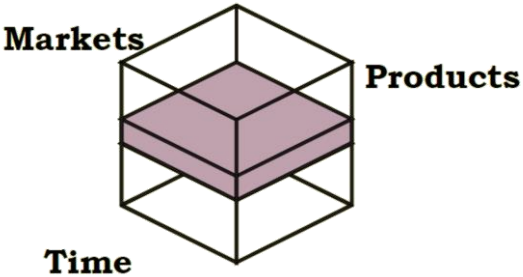


Sales Cube

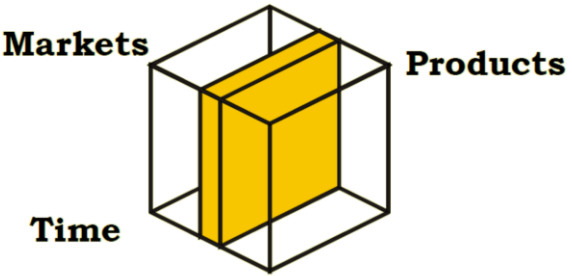
The area manager examines product sales of his/her own markets



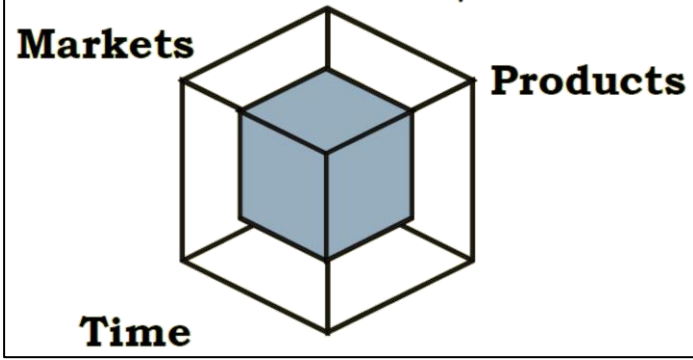
Product manager examines the sales of a specific product in all periods and in all markets



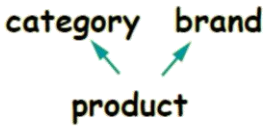
Financial manager examines product sales in all markets, for the current period and the previous one

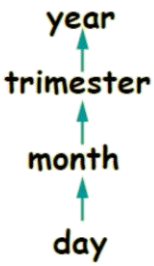


The strategic manager concentrates on a category of products, a specific region and a medium time span



Dimensions and hierarchies

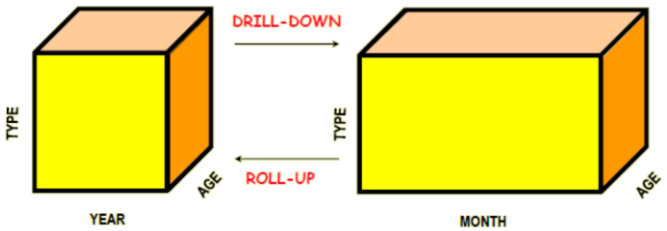


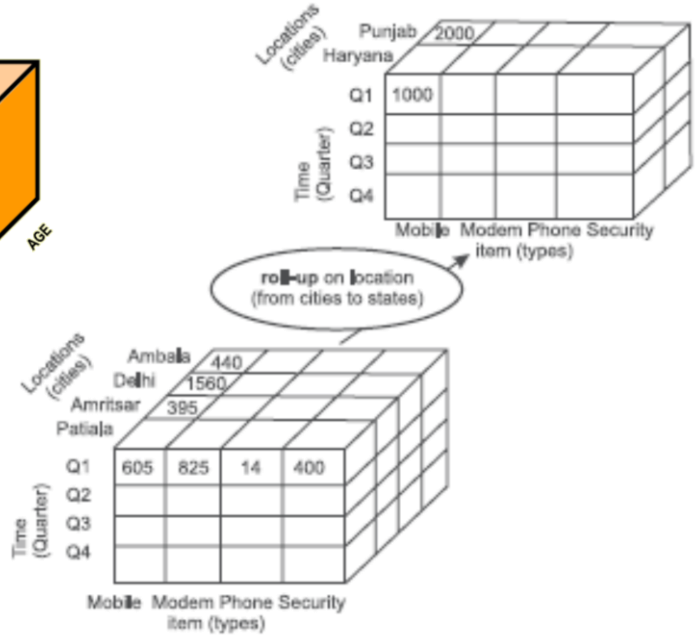


OLAP operations

Roll-up

- Roll-up is like zooming out on the data cube.
- Roll-up is used to provide abstract level details to user.
- It performs further aggregation of data by reduction in dimension or by stepping up a concept hierarchy for a dimension.
- Here, aggregation is performed on cities to climb up to state for dimension reduction.
- The aggregation can also be performed on Time (Quarter) to year, etc. or individual items to group like mobile, modem, etc.

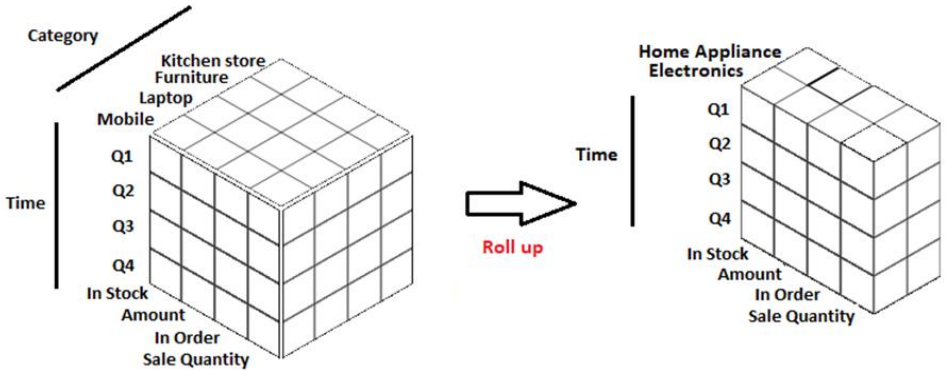




OLAP operations

Roll-up

- The roll-up operation (also called drill-up or aggregation operation) performs aggregation on a data cube, either by climbing up a concept hierarchy for a dimension or by climbing down a concept hierarchy, i.e., dimension reduction.
- In the following example, it is shown a multidimensional cube containing the products of a Home appliances home appliances like laptop, furniture, mobile and kitchen appliances.
- If the manager wants to view the sales of all the products quarterly, the Roll-up operation can be performed on the categories.
- In this aggregation process, data is category hierarchy moves up from mobile to the Kitchen store.
- In the roll-up process at least one or more dimensions get reduced like category here.
- It is also known as consolidation. This operation summarizes the data along the dimension.



Metrics	Dollar Sales									
Customer Region / Month	North-East	Mid-Atlantic	South-East	Central	South	North-West	South-West	England	France	Germany
Jan 97	$ 620	$ 753	$ 30	$ 660	$ 2.405	$ 1.312	$ 440	$ 1.002	$ 1.002	$ 383
Feb 97	$ 258	$ 252	$ 800	$ 975	$ 160	$ 582	$ 744	$ 310	$ 799	$ 118
Mar 97	$ 648	$ 244	$ 148	$ 250	$ 1.085	$ 2.961	$ 650	$ 1.240	$ 119	$ 142
Apr 97	$ 787	$ 588	$ 447	$ 486	$ 226	$ 506	$ 601	$ 119	$ 550	$ 85
May 97	$ 1.350	$ 245	$ 936	$ 159	$ 664	$ 626	$ 107	$ 135	$ 200	$ 177
Jun 97	$ 842	$ 582	$ 1.281	$ 937	$ 240	$ 774	$ 176	$ 1.139	$ 652	$ 254
Jul 97	$ 652	$ 690	$ 486	$ 1.293	$ 605	$ 303	$ 818	$ 103	$ 124	$ 173
Aug 97	$ 1.783	$ 304	$ 1.032	$ 170	$ 398	$ 356	$ 432	$ 190	$ 241	$ 407
Sep 97	$ 581	$ 778	$ 3.558	$ 587	$ 440	$ 1.652	$ 1.071	$ 315	$ 210	$ 202
Oct 97	$ 2.291	$ 1.840	$ 600	$ 656	$ 1.300	$ 718	$ 1.210	$ 427	$ 220	$ 520
Nov 97	$ 39	$ 1.602	$ 1.082	$ 1.187	$ 842	$ 759	$ 745	$ 232	$ 101	$ 1.037
Dec 97	$ 381	$ 1.588	$ 343	$ 118	$ 1.459	$ 635	$ 2.021	$ 259	$ 210	$ 119
Jan 98	$ 311	$ 1.174	$ 2.634	$ 3.130	$ 954	$ 2.083	$ 1.351	$ 747	$ 426	$ 447
Feb 98	$ 2.518	$ 702	$ 1.123	$ 1.336	$ 1.227	$ 3.887	$ 545	$ 268	$ 277	$ 282
Mar 98	$ 2.459	$ 1.523	$ 1.178	$ 4.708	$ 1.420	$ 3.514	$ 1.948	$ 1.705	$ 276	$ 1.168
Apr 98	$ 407	$ 841	$ 524	$ 712	$ 133	$ 2.486	$ 49	$ 390	$ 1.298	$ 221
May 98	$ 667	$ 1.721	$ 440	$ 148	$ 80	$ 1.310	$ 303	$ 104	$ 657	$ 65
Jun 98	$ 699	$ 1.096	$ 898	$ 353	$ 902	$ 839		$ 230	$ 155	$ 105
Jul 98	$ 586	$ 1.897	$ 412	$ 226	$ 406	$ 361	$ 1.628	$ 267	$ 1.011	$ 41
Aug 98	$ 894	$ 326	$ 792	$ 1.832	$ 1.199	$ 295	$ 1.816	$ 277	$ 102	$ 118
Sep 98	$ 338	$ 3.179	$ 505	$ 427	$ 99	$ 2.976	$ 885	$ 135	$ 85	$ 1.110
Oct 98	$ 544	$ 413	$ 1.467	$ 209	$ 679	$ 706	$ 556	$ 480	$ 485	$ 99
Nov 98	$ 671	$ 459	$ 1.471	$ 2.066	$ 701	$ 716	$ 986	$ 1.127	$ 154	$ 440
Dec 98	$ 836	$ 2.096	$ 1.726	$ 3.642	$ 395	$ 1.740	$ 1.943	$ 1.143	$ 366	$ 307

Metrics	Dollar Sales									
Customer Region / Quarter	North-East	Mid-Atlantic	South-East	Central	South	North-West	South-West	England	France	Germany
Q1 1997	$ 1.526	$ 1.249	$ 978	$ 1.885	$ 3.650	$ 4.855	$ 1.834	$ 2.552	$ 1.920	$ 643
Q2 1997	$ 2.979	$ 1.415	$ 2.664	$ 1.582	$ 1.130	$ 1.906	$ 884	$ 1.393	$ 1.402	$ 516
Q3 1997	$ 3.016	$ 1.772	$ 5.076	$ 2.050	$ 1.443	$ 2.311	$ 2.321	$ 608	$ 575	$ 782
Q4 1997	$ 2.711	$ 5.030	$ 2.025	$ 1.961	$ 3.601	$ 2.112	$ 3.976	$ 918	$ 531	$ 1.676
Q1 1998	$ 5.288	$ 3.399	$ 4.935	$ 9.174	$ 3.601	$ 9.484	$ 3.844	$ 2.720	$ 979	$ 1.897
Q2 1998	$ 1.773	$ 3.658	$ 1.862	$ 1.213	$ 1.115	$ 4.635	$ 352	$ 724	$ 2.110	$ 391
Q3 1998	$ 1.818	$ 5.402	$ 1.709	$ 2.485	$ 1.704	$ 3.632	$ 4.329	$ 679	$ 1.198	$ 1.269
Q4 1998	$ 2.051	$ 2.968	$ 4.664	$ 5.917	$ 1.775	$ 3.162	$ 3.485	$ 2.750	$ 1.005	$ 846

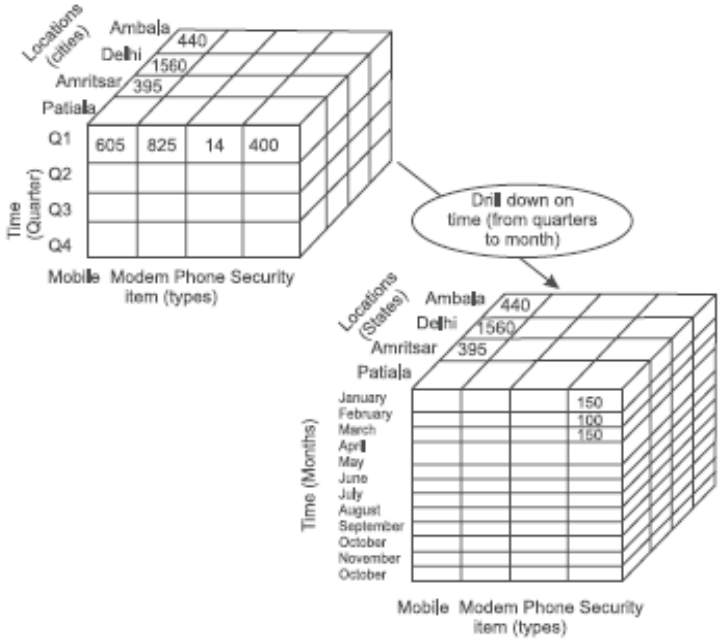
Category	Year	Metrics / Customer Region	Dollar Sales North-East	Mid-Atlantic	South-East	Central	South	North-West	South-West	England	France
Electronics	1997		$ 138	$ 1.774	$ 394	$ 138	$ 2.346	$ 2.554	$ 2.184	$ 566	$ 199
	1998		$ 1184	$ 4.529	$ 1892	$ 7.232	$ 651	$ 9.488	$ 476	$ 2.683	$ 462
Food	1997		$ 759	$ 682	$ 729	$ 252	$ 588	$ 469	$ 807	$ 156	$ 615
	1998		$ 530	$ 925	$ 939	$ 677	$ 213	$ 1.503	$ 201	$ 165	$ 175
Gifts	1997		$ 2532	$ 1.355	$ 1854	$ 1.413	$ 2.535	$ 2.132	$ 1.904	$ 908	$ 375
	1998		$ 1955	$ 2.765	$ 2.800	$ 2.695	$ 1.813	$ 2.844	$ 1.778	$ 1.158	$ 717
Health & Beauty	1997		$ 624	$ 640	$ 1.317	$ 647	$ 588	$ 754	$ 654	$ 143	$ 292
	1998		$ 611	$ 887	$ 556	$ 382	$ 499	$ 1.162	$ 1.044	$ 273	$ 72
Household	1997		$ 5354	$ 4.112	$ 5.410	$ 4.446	$ 3.058	$ 3.974	$ 2.654	$ 3.545	$ 2.875
	1998		$ 5787	$ 5.320	$ 5.416	$ 6.812	$ 4.334	$ 5.008	$ 7.588	$ 2.139	$ 3.649
Kid's Korner	1997		$ 201	$ 398	$ 485	$ 186	$ 409	$ 313	$ 396	$ 105	$ 34
	1998		$ 247	$ 422	$ 441	$ 380	$ 221	$ 592	$ 290	$ 198	$ 19
Travel	1997		$ 624	$ 505	$ 564	$ 386	$ 300	$ 978	$ 416	$ 48	$ 38
	1998		$ 608	$ 559	$ 1.096	$ 611	$ 464	$ 316	$ 573	$ 257	$ 198

Category	Metrics / Year	Dollar Sales
Electronics	1997	$ 10.616
	1998	$ 29.299
Food	1997	$ 5.300
	1998	$ 5.638
Gifts	1997	$ 16.315
	1998	$ 20.047
Health & Beauty	1997	$ 6.042
	1998	$ 5.065
Household	1997	$ 38.383
	1998	$ 50.391
Kid's Korner	1997	$ 2.559
	1998	$ 2.943
Travel	1997	$ 4.497
	1998	$ 4.792

OLAP operations

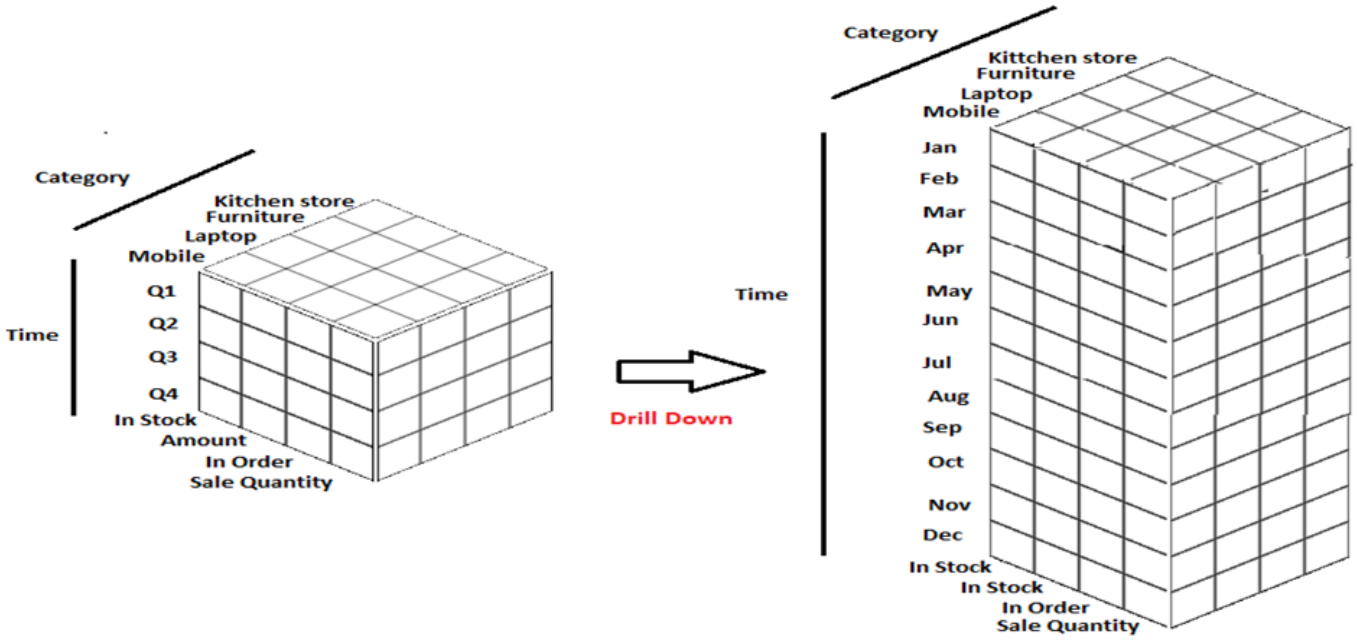
Drill-down

- The drill down operation (also called roll-down) is the reverse of roll up.
- Drill-down is like zooming in on data and is used to provide detailed data to the user.
- It provides detailed information by introducing a new dimension or by moving down a concept hierarchy for a dimension.
- It navigates from less detailed to more detailed data.



OLAP operations

Drill-down



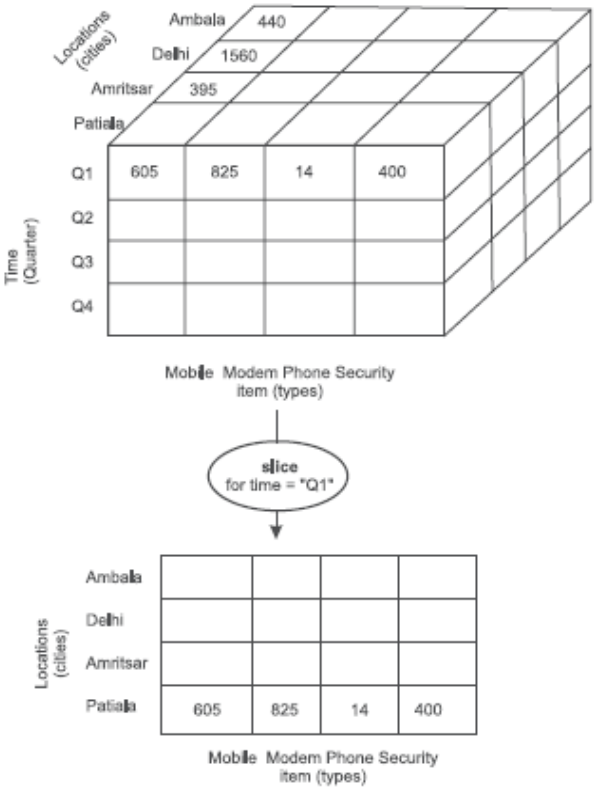
Metrics	Dollar Sales	
Year	1997	1998
Category		
Electronics	$ 10.616	$ 29.299
Food	$ 5.300	$ 5.638
Gifts	$ 16.315	$ 20.047
Health & Beauty	$ 6.042	$ 5.665
Household	$ 38.383	$ 50.391
Kid's Korner	$ 2.559	$ 2.943
Travel	$ 4.497	$ 4.792

Metrics	Dollar Sales											
Customer Region	North-East		Mid-Atlantic		South-East		Central		South		North-West	
Year	1997	1998	1997	1998	1997	1998	1997	1998	1997	1998	1997	1998
Category												
Electronics	$ 138	$ 1.184	$ 1.774	$ 4.529	$ 384	$ 1.092	$ 138	$ 7.232	$ 2.346	$ 651	$ 2.554	$ 9.488
Food	$ 759	$ 538	$ 682	$ 925	$ 729	$ 959	$ 262	$ 677	$ 568	$ 213	$ 469	$ 1.503
Gifts	$ 2.532	$ 1.955	$ 1.355	$ 2.785	$ 1.854	$ 2.800	$ 1.413	$ 2.695	$ 2.515	$ 1.813	$ 2.132	$ 2.844
Health & Beauty	$ 624	$ 611	$ 648	$ 887	$ 1.317	$ 566	$ 647	$ 382	$ 569	$ 499	$ 754	$ 1.162
Household	$ 5.354	$ 5.787	$ 4.112	$ 5.320	$ 5.410	$ 5.416	$ 4.446	$ 6.812	$ 3.059	$ 4.334	$ 3.974	$ 5.008
Kid's Korner	$ 201	$ 247	$ 398	$ 422	$ 485	$ 441	$ 186	$ 380	$ 409	$ 221	$ 323	$ 592
Travel	$ 624	$ 608	$ 505	$ 559	$ 564	$ 1.096	$ 386	$ 611	$ 300	$ 464	$ 978	$ 316

OLAP operations

Slice

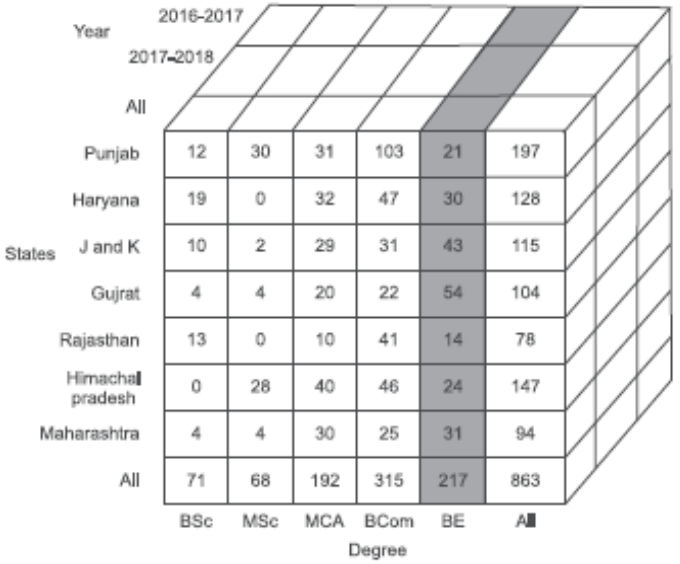
- Slice and dice represent the way to view the information from different perspectives.
- The slice operation gives a new sub-cube by selecting one dimension from a specified cube.
- Thus, a slice is a subset of the cube corresponding to a single value for one or more members of the dimensions.
- It results in reduction in dimension.
- So, a slice operation is performed when the user wants to select one dimension of a three-dimensional cube that results in a two-dimensional slice.
- A slice operation is performed using the criterion time ‘Q1’ for the dimension ‘time’.
- It selects one or more dimensions and gives a new sub-cube.



OLAP operations

Slice

- Let us consider another case of a University database system, where we have a count of the number of students from each state in previous semesters in a particular course.
- Let us suppose management is interested in finding the details of their BE course.
- So, the degree dimension will be fixed as degree = BE and the slice will be taken out. It will not have any information about other degrees.



OLAP operations

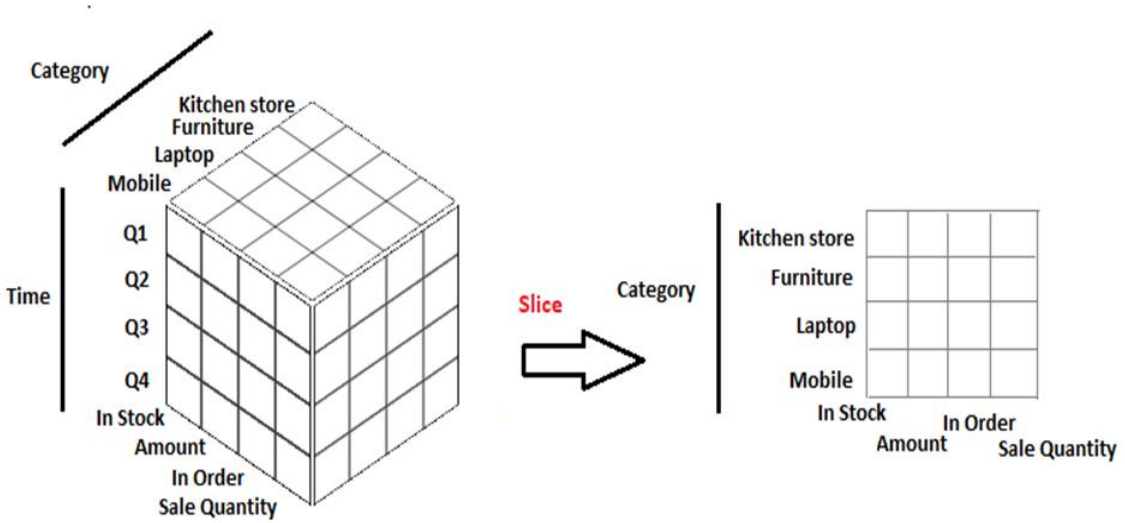
Slice

- The information retrieved therefore is more like a two-dimensional rectangle than a cube for degree = BE.

State \ Year	2014–2015	2015–2016	2016–2017	2017–2018
Punjab	11	5	10	2
Haryana	17	0	13	5
J and K	23	2	20	1
Gujarat	31	4	23	4
Rajasthan	7	0	7	4
Himachal Pradesh	13	8	11	6
Maharashtra	19	4	12	5

OLAP operations

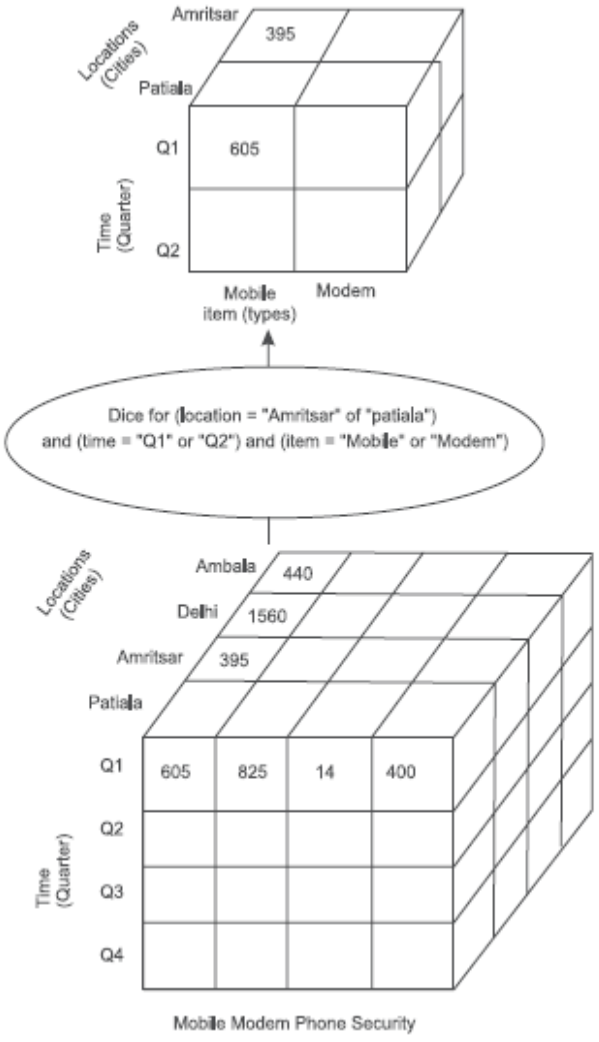
Slice



OLAP operations

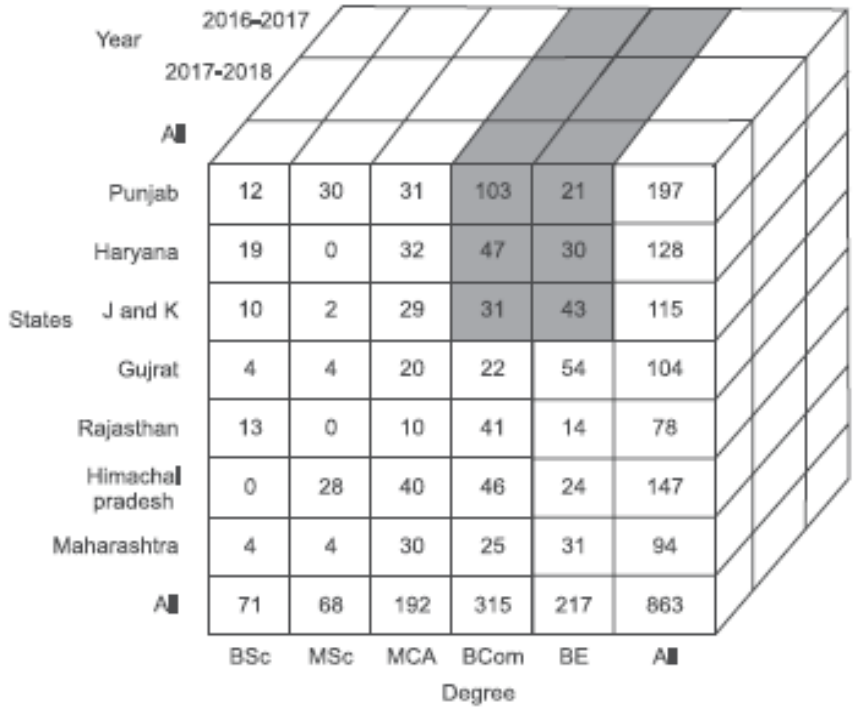
Dice

- The dice operation is analogous to slice without reduction in the number of dimensions.
- The Dice operation gives a new sub-cube by selecting two or more dimensions from a specified cube.
- Here, the dice operation is performed on the cube based on the selection criteria of the following three dimensions.
- • (location = ‘Amritsar’ or ‘Patiala’)
- • (time = ‘Q1’ or ‘Q2’)
- • (item = ‘Mobile’ or ‘Modem’)
- Dice is attained by performing selection on two or more dimensions.
- Similarly, in the case of the university database system, a dice operation can be performed to know the number of students enrolled in both BE and BCom degrees from states Punjab, Haryana and Maharashtra as shown in Figure 14.24.



OLAP operations

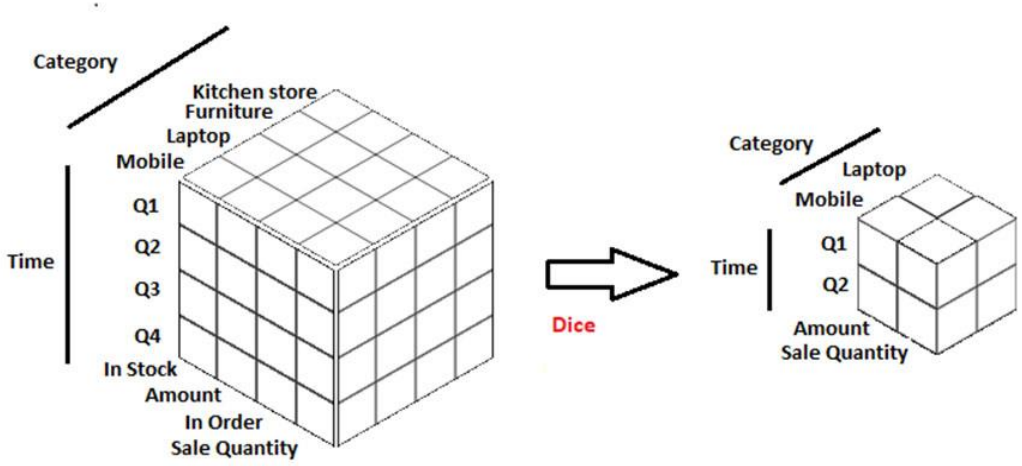
Dice

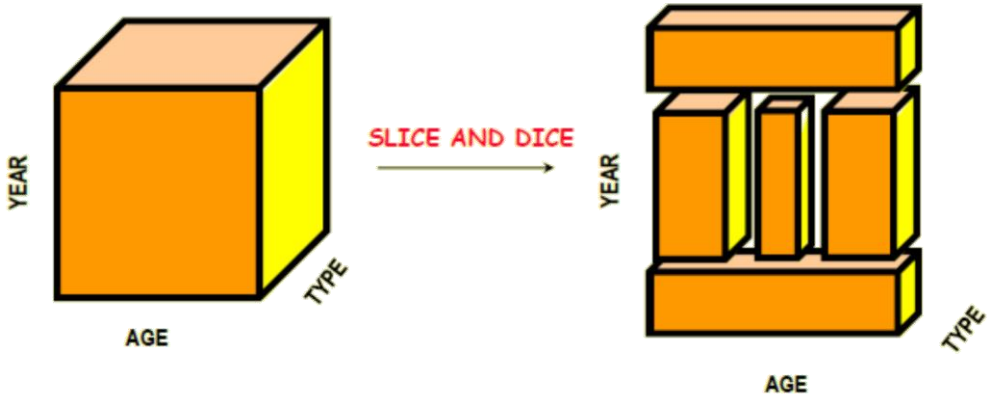


OLAP operations

Dice

- This allows an analyst to select data from multiple dimensions to analyze. This OLAP
- operation is just like the Projection relational query you have read in RDBMS. In this
- technique you select two or more dimensions that results in the creation of a sub cube




YEAR
AGE
TYPE
SLICE AND DICE
YEAR
AGE
TYPE

Metrics / Customer City / Subcategory	Dollar Sales										
	Afton	Akron	Albon	Alcameda	Alka	Allagash	Alta	Altoola	Amestra	Amsterdam	Andersonville
Audio						$ 85					
Automotive								$ 30			
Chocolate	$ 42	$ 42		$ 50		$ 20		$ 22	$ 44		
Christmas	$ 30					$ 25	$ 30	$ 15			
Classic Toys						$ 7	$ 26				$ 36
Coffee			$ 9								
Comfort				$ 59		$ 59					
Furniture							$ 485				
Gadgets							$ 199	$ 79	$ 79		
Games & Puzzles							$ 17		$ 45		$ 45
Gift Baskets			$ 55	$ 43							
Golf	$ 25							$ 25	$ 14		$ 25
Hearth									$ 15		
Jewelry	$ 75			$ 189		$ 24	$ 77	$ 189	$ 24		
Kitchen						$ 55	$ 21		$ 76		
Lawn & Garden	$ 75		$ 100		$ 15	$ 63	$ 100		$ 180	$ 67	$ 40
Learning	$ 16							$ 37			
Meat & Cheese		$ 40		$ 20			$ 20				$ 25
Miscellaneous		$ 200	$ 1,320		$ 200	$ 139			$ 993		
Natural Remedies	$ 13								$ 13		
Pets	$ 215		$ 26			$ 30	$ 68	$ 115	$ 25		$ 34
Plants & Flowers	$ 65	$ 65	$ 65				$ 50	$ 60			
Safety & Security								$ 30	$ 22	$ 22	
Skin Care											
Sleeping			$ 18								
Toys & Accessories							$ 29	$ 185	$ 744		

Filter Details:
Category = Electronics
AND
Dollar Sales > 60
AND
Customer Region = North-West
AND
Year = 1997

Metrics / Customer City / Subcategory	Dollar Sales					
	Alta	Armstrong	Avery Heights	Lane	Mt. Everest	San Fransisco
Audio		$ 98		$ 123	$ 85	
Comfort			$ 118		$ 1,495	
Gadgets	$ 199					$ 199

Category	Year	Metrics / Customer Region	Dollar Sales North-East	Mid-Atlantic	South-East	Central	South	North-West	South-West	England
Electronics	1997		$ 138	$ 1,774	$ 384	$ 138	$ 2,346	$ 2,554	$ 2,184	$ 560
	1998		$ 1,184	$ 4,529	$ 1,892	$ 7,232	$ 651	$ 9,488	$ 476	$ 2,683
Food	1997		$ 759	$ 682	$ 729	$ 262	$ 588	$ 469	$ 807	$ 156
	1998		$ 538	$ 925	$ 959	$ 677	$ 213	$ 1,503	$ 261	$ 165
Gifts	1997		$ 2,532	$ 1,355	$ 1,854	$ 1,413	$ 2,535	$ 2,132	$ 1,904	$ 908
	1998		$ 1,955	$ 2,785	$ 2,800	$ 2,695	$ 1,813	$ 2,844	$ 1,778	$ 1,158
Health & Beauty	1997		$ 624	$ 640	$ 1,317	$ 647	$ 568	$ 754	$ 654	$ 140
	1998		$ 611	$ 887	$ 566	$ 382	$ 499	$ 1,162	$ 1,044	$ 273
Household	1997		$ 5,354	$ 4,112	$ 5,410	$ 4,446	$ 3,658	$ 3,974	$ 2,654	$ 3,545
	1998		$ 5,787	$ 5,320	$ 5,416	$ 6,812	$ 4,334	$ 5,008	$ 7,588	$ 2,139
Kid's Korner	1997		$ 201	$ 399	$ 485	$ 196	$ 409	$ 323	$ 396	$ 105
	1998		$ 247	$ 422	$ 441	$ 380	$ 221	$ 592	$ 290	$ 198
Travel	1997		$ 624	$ 505	$ 564	$ 386	$ 300	$ 978	$ 416	$ 48
	1998		$ 608	$ 559	$ 1,096	$ 611	$ 464	$ 316	$ 573	$ 257

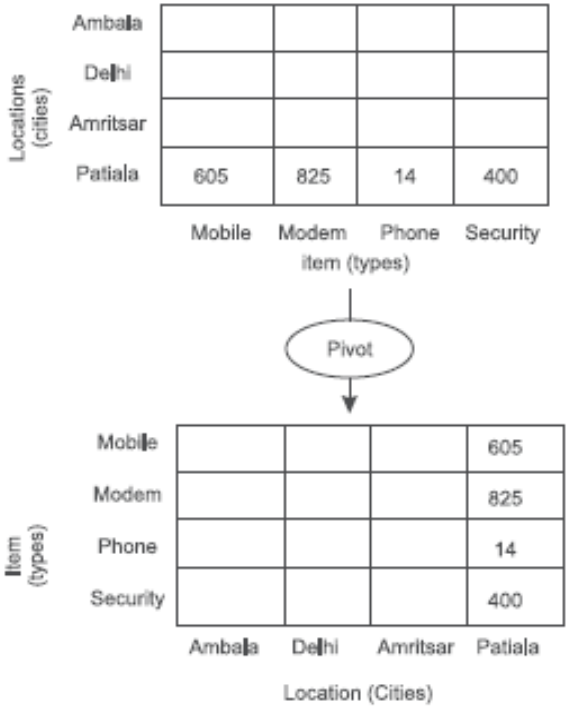
Filter Details:
Year = 1998

Category	Metrics / Customer Region	Dollar Sales North-East	Mid-Atlantic	South-East	Central	South	North-West	South-West	England
Electronics		$ 1,184	$ 4,529	$ 1,892	$ 7,232	$ 651	$ 9,488	$ 476	$ 2,683
Food		$ 538	$ 925	$ 959	$ 677	$ 213	$ 1,503	$ 261	$ 165
Gifts		$ 1,955	$ 2,785	$ 2,800	$ 2,695	$ 1,813	$ 2,844	$ 1,778	$ 1,158
Health & Beauty		$ 611	$ 887	$ 566	$ 382	$ 499	$ 1,162	$ 1,044	$ 273
Household		$ 5,787	$ 5,320	$ 5,416	$ 6,812	$ 4,334	$ 5,008	$ 7,588	$ 2,139
Kid's Korner		$ 247	$ 422	$ 441	$ 380	$ 221	$ 592	$ 290	$ 198
Travel		$ 608	$ 559	$ 1,096	$ 611	$ 464	$ 316	$ 573	$ 257

OLAP operations

Pivot

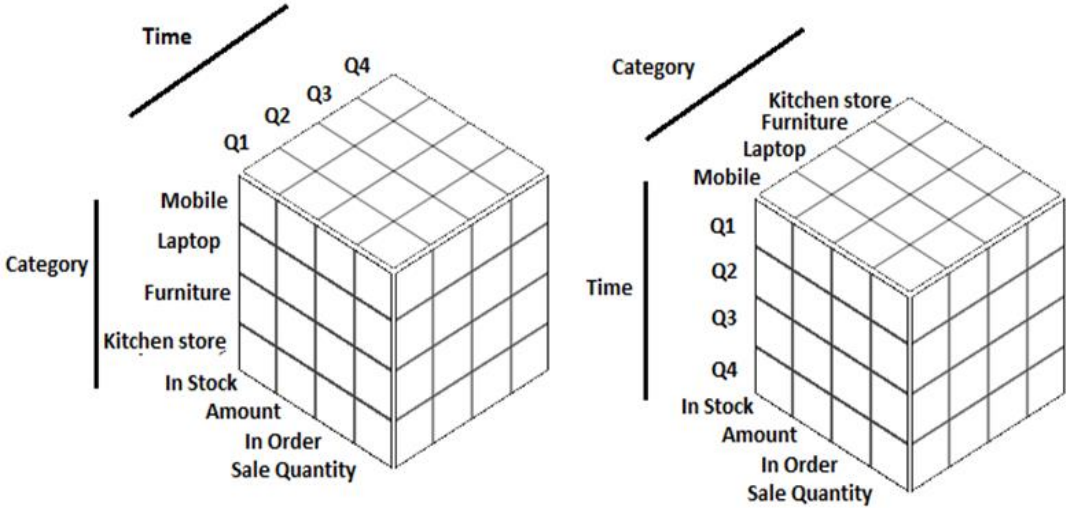
- The pivot operation is also called rotation.
- This operation rotates the data axes in view to get another presentation of data. It may include swapping of the columns and rows.

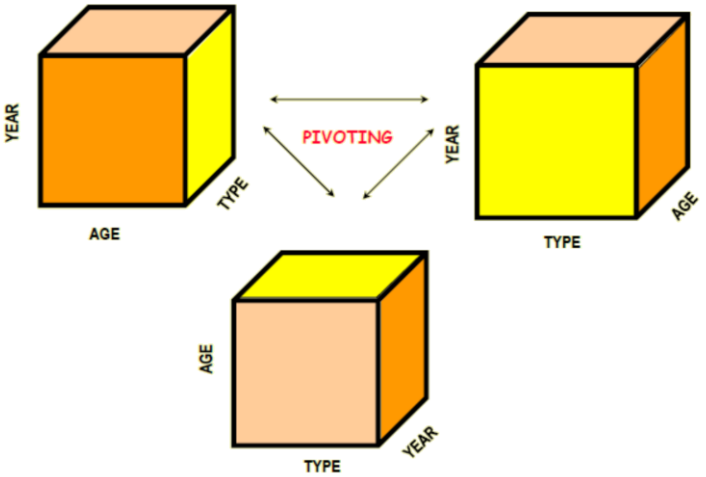


OLAP operations

Pivot

- Analysts can gain a new view of data by rotating the data axes of the cube.
- This OLAP operation fixes one attribute as a Pivot and rotate the cube to fetch the results.
- Like inverting the spreadsheet it gives a different perspective.
- You can observe in the above figure that the presentation of the dimensions has been changed to impart a different perspective of the data cube for data analysis.





Category	Year	Metrics: Dollar Sales
Electronics	1997	$ 10.616
	1998	$ 29.299
Food	1997	$ 5.300
	1998	$ 5.638
Gifts	1997	$ 16.315
	1998	$ 20.047
Health & Beauty	1997	$ 6.042
	1998	$ 5.665
Household	1997	$ 38.383
	1998	$ 50.391
Kid's Korner	1997	$ 2.559
	1998	$ 2.943
Travel	1997	$ 4.497
	1998	$ 4.792

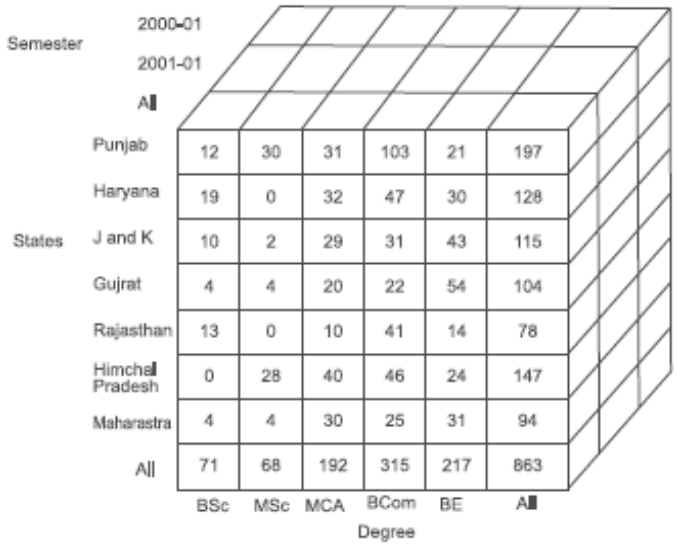
	Metrics	Dollar Sales	
	Year	1997	1998
Category			
Electronics		$ 10.616	$ 29.299
Food		$ 5.300	$ 5.638
Gifts		$ 16.315	$ 20.047
Health & Beauty		$ 6.042	$ 5.665
Household		$ 38.383	$ 50.391
Kid's Korner		$ 2.559	$ 2.943
Travel		$ 4.497	$ 4.792

OLAP models

- **ROLAP** - relational online analytical processing. In this case, the OLAP system is built on top of a relational database.
- **MOLAP** Refers to multidimensional online analytical processing. In this case, the OLAP system is implemented through a specialized multidimensional database.
- **HOLAP** Refers to hybrid online analytical processing. This models attempts to combine the strengths and features of ROLAP and MOLAP.
- **DOLAP** Refers for desktop online analytical processing. DOLAP is meant to provide portability to users of online analytical processing. In the DOLAP methodology, multidimensional datasets are created and transferred to the desktop machine, requiring only the DOLAP software to exist on that machine. DOLAP is a variation of ROLAP.
- **Database OLAP** Refers to a relational database management system (RDBMS) designated to support OLAP structures and to perform OLAP calculations.
- **Web OLAP** Refers to online analytical processing where OLAP data is accessible from a Web browser.

Multi-dimensional OLAP (MOLAP)

- Multi-dimensional OLAP is based on multi-dimensional DBMS for storage and access of data.
- It is known as a top-down approach to OLAP.
- It uses a special purpose file system to store precomputed values in the cube.



Multi-dimensional OLAP (MOLAP)

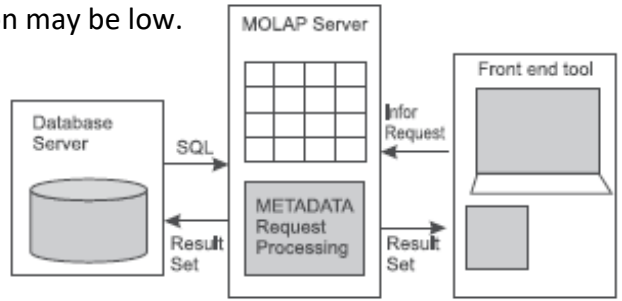
- In this case, cell locations can be inferred from the known dimensions while dimensions can be inferred from cell locations.
- There are some issues in implementations of MOLAP as arrays are likely to be very large for large multi-dimensional databases.
- Since, zero facts are stored in MOLAP it produces the case of sparseness of data.
- To overcome the problem of large sizes, arrays are normally split into chunks so that these can be accommodated in the main memory.
- To overcome the issue of sparseness, the chunks may be compressed.

Advantages

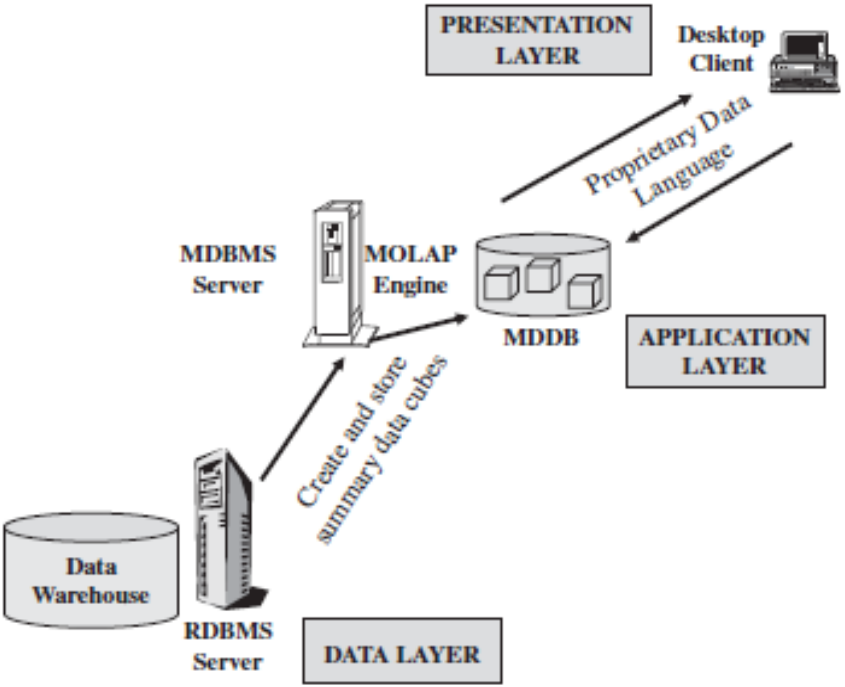
- MOLAP servers allow quickest indexing of pre-computed data.
- It allows users to analyze both large and less-detailed data.
- It is suitable for inexperienced users as MOLAP is easier to use.

Disadvantages

- MOLAP is not able to contain detailed data.
- If the dataset is sparse then the storage utilization may be low.



Multi-dimensional OLAP (MOLAP)



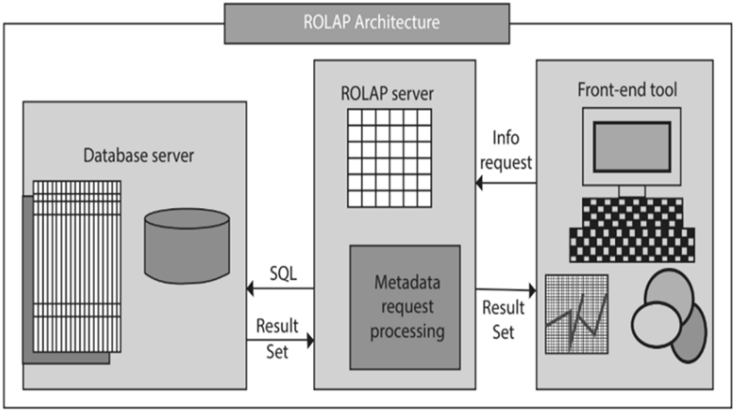
Relational OLAP(ROLAP)

- ROLAP is based on the relational model of DBMS.
- The result of GROUP BY queries can be pre-computed and stored in another relational table or view for implementation by the ROLAP server.
- It is also known as a bottom up approach which uses the star schema of data warehousing.
- The major advantages of ROLAP, is the support of existing RDBMS and efficient storage because no zero facts are returned by the SQL query.
- It means that if there is no record for Professor in department number 20, the SQL query will not return any record for it.
- Records will be returned for it, if at least one record is found for the given job in a particular dept number.
- It shows that GROUP BY clause doe not return or store zero value fact.
- While these queries run on the fly and are not pre computed, low efficiency is the issue with ROLAP servers.

The characteristics of ROLAP are:

- ROLAP utilizes the more processing time and disk space.
- ROLAP enables and supports larger user group in the distributed environment.
- ROLAP processes complex queries utilizing the greater amounts of data.

Relational OLAP



Autocommit Display 10

```
select dno, job, count(*)
from emp group by dno,job order by dno
```

Results Explain Describe Saved SQL History

DNO	JOB	COUNT(*)
10	AP	3
10	ASSOC PROF	1
10	PROF	4
20	AP	2
20	ASSOC PROF	2
30	ASSOC PROF	1
30	PROF	2

7 rows returned in 0.02 seconds CSV Export

Relational OLAP

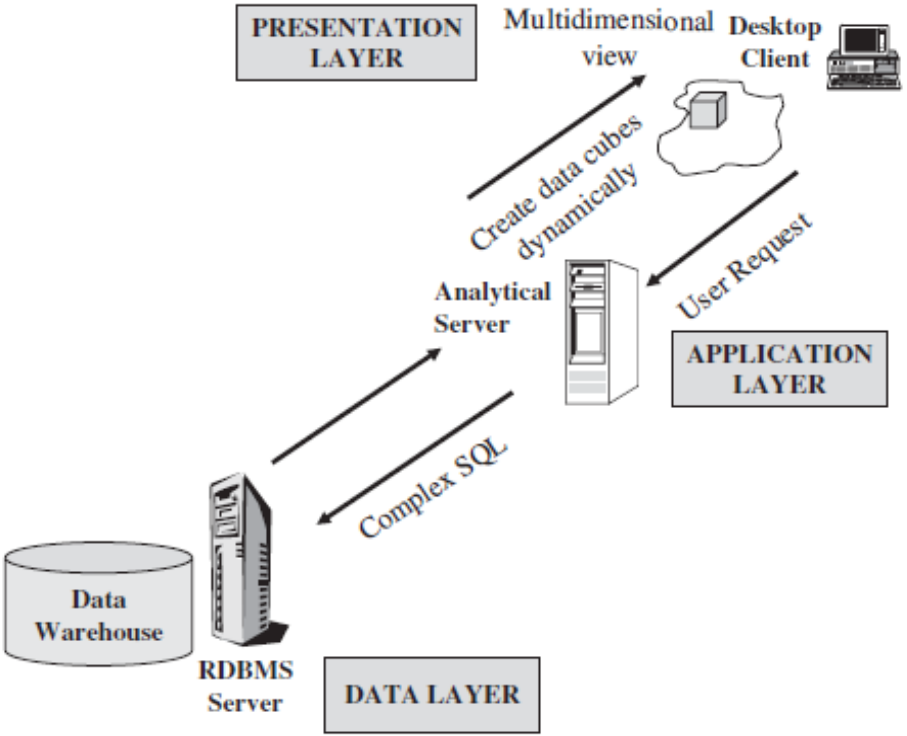
Advantages

- These can be used with existing RDBMS without any problem.
- As no zero facts can be stored, therefore data is stored efficiently.
- ROLAP servers do not use pre-computed data cubes.

Disadvantages

- Poor query performance.

Relational OLAP



Property	*ROLAP*	*MOLAP*
Information retrieval	Information retrieval is comparatively slow.	Information retrieval is fast.
Data Model	Uses relational table.	Uses sparse array to store datasets.
Type of Users	ROLAP is best suited for experienced users.	MOLAP is applicable for inexperienced users as it is easy to use.
Disk Space	It may not require space other than available in data warehouse.	MOLAP maintains an isolated database for data cubes.
Scalability	It is scalable as it easy to add dimension and fact in SQL query.	It has limited scalability as any addition of new dimension requires a complete remapping of values with dimensions in an array.
DBMS Facility	It has strong DBMS facility as it is based on relational model.	It has poor DBMS facility due to its different data structure.