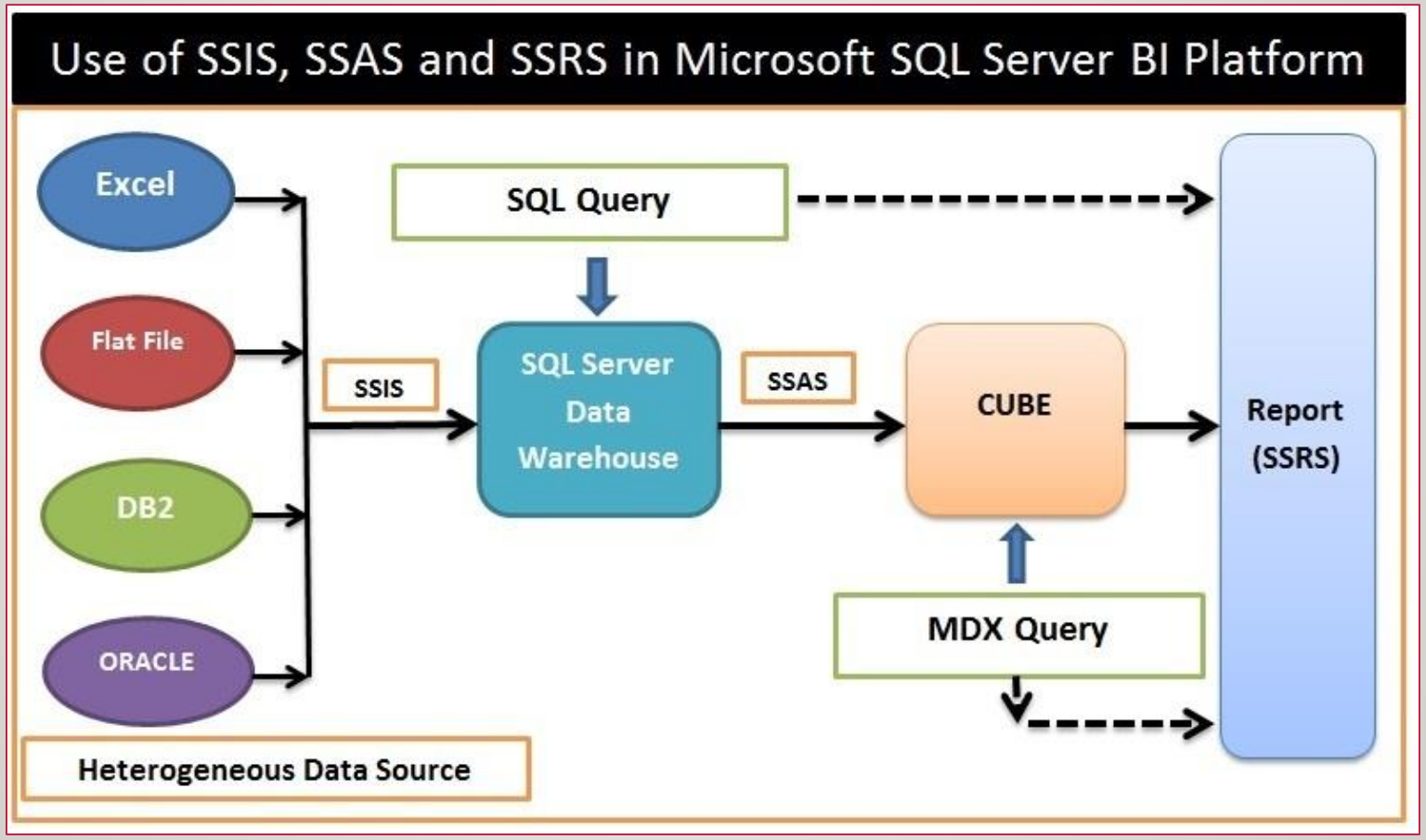
DM LAB 6

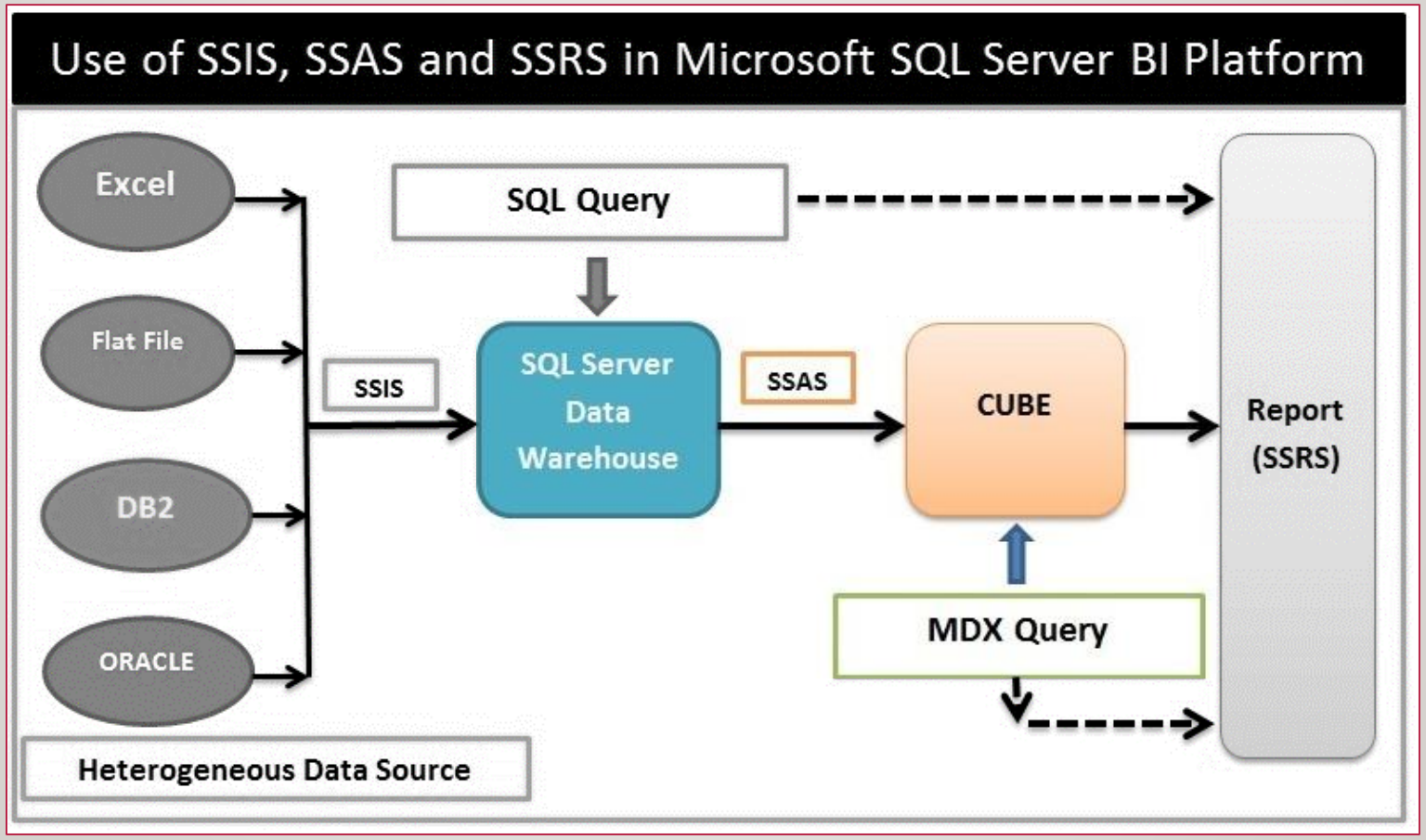
BY

ENG. JOUD KHATTAB

2 VISUAL STUDIO BUSINESS INTELLIGENCE SQL SERVER ANALYSIS SERVICES (SSAS)



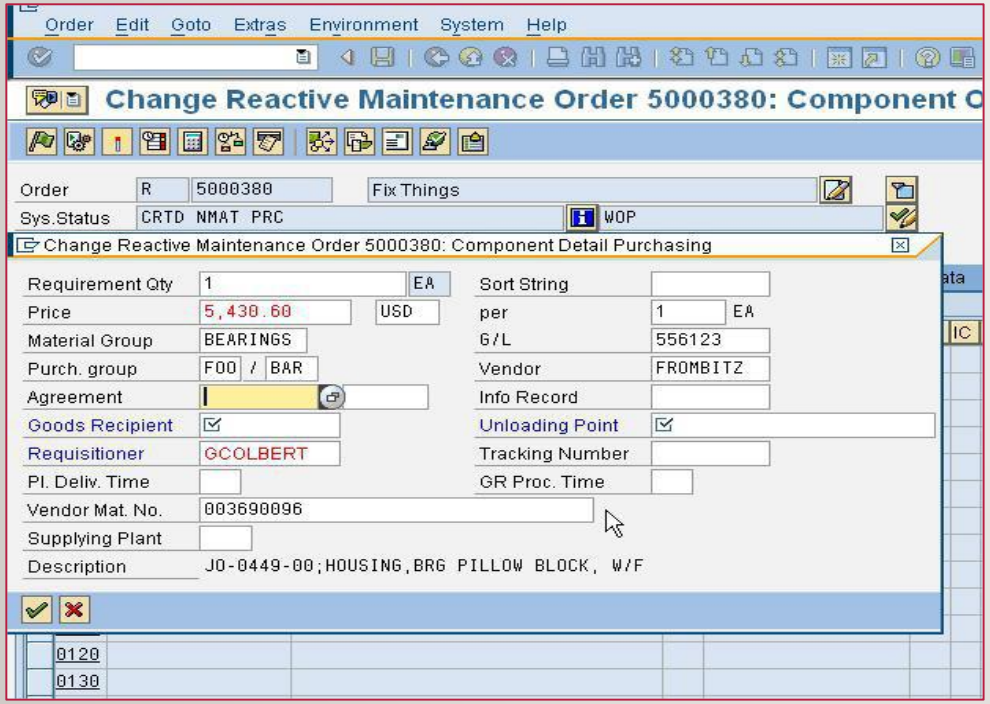
3 VISUAL STUDIO BUSINESS INTELLIGENCE SQL SERVER ANALYSIS SERVICES (SSAS)



4 OLTP/OLAP COMPARISON

OLTP

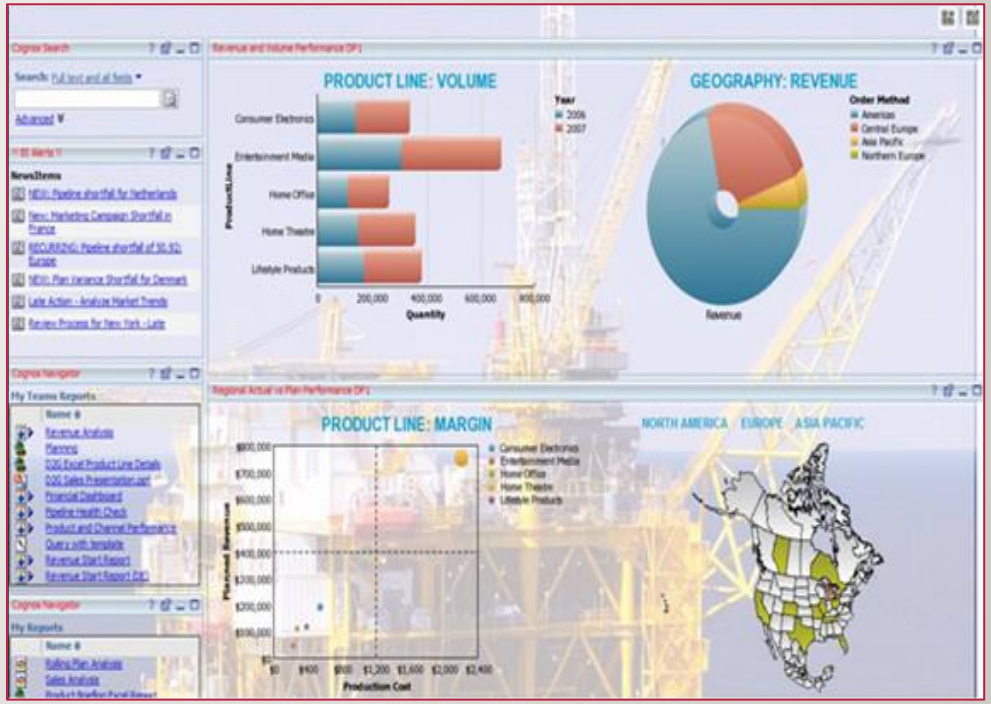
ON-LINE TRANSACTION PROCESSING



OLTP query for a single Sales order

OLAP

ON-LINE ANALYTIC PROCESSING

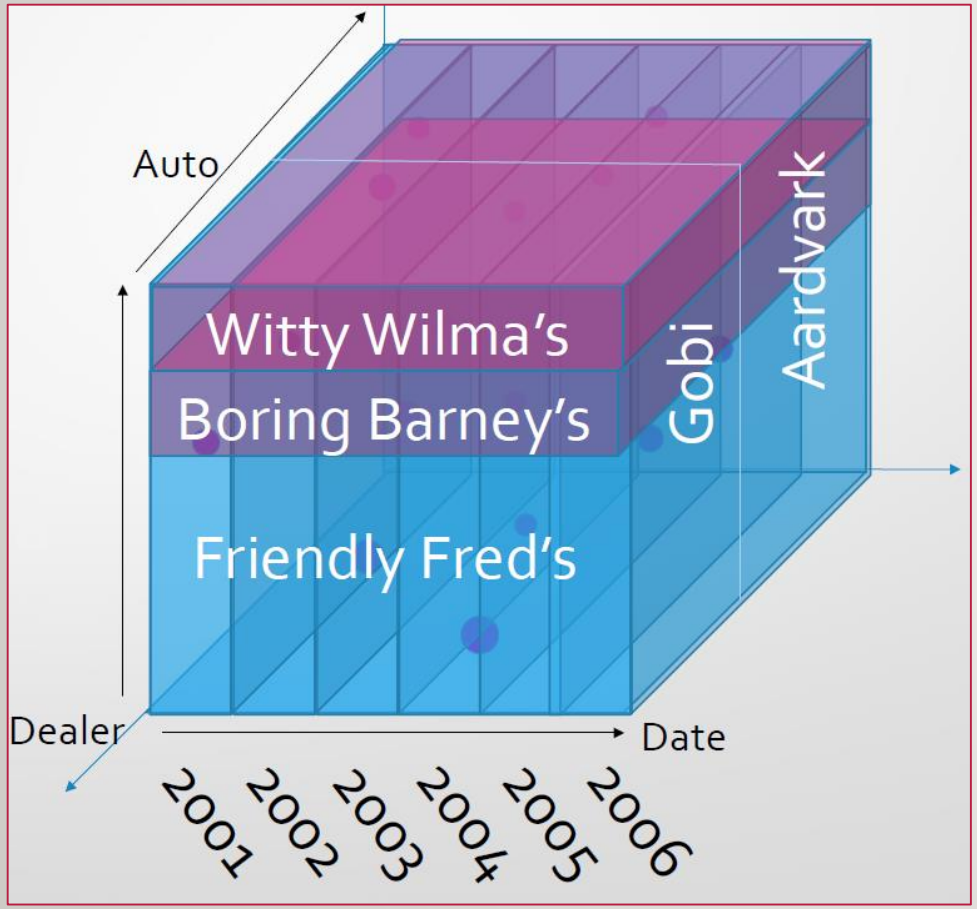


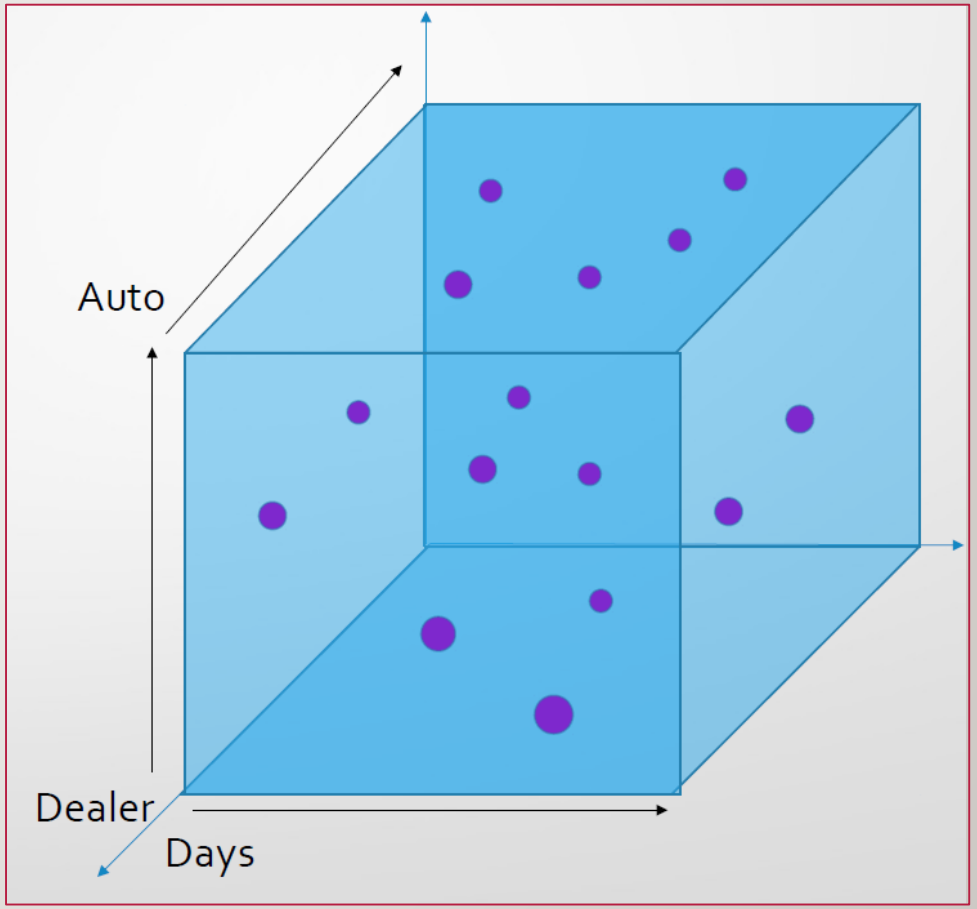
OLAP query for revenue by geography / product line

5 SQL SERVER ANALYSIS SERVICES (SSAS) DEFINITION

- A middle-tier server for OLAP and data mining.
- Manages multi-dimensional cubes of data for analysis and provides rapid client access.
- Allows you to create data mining models from both OLAP and relational data sources.

6 VISUALIZING MULTIDIMENSIONAL DATA





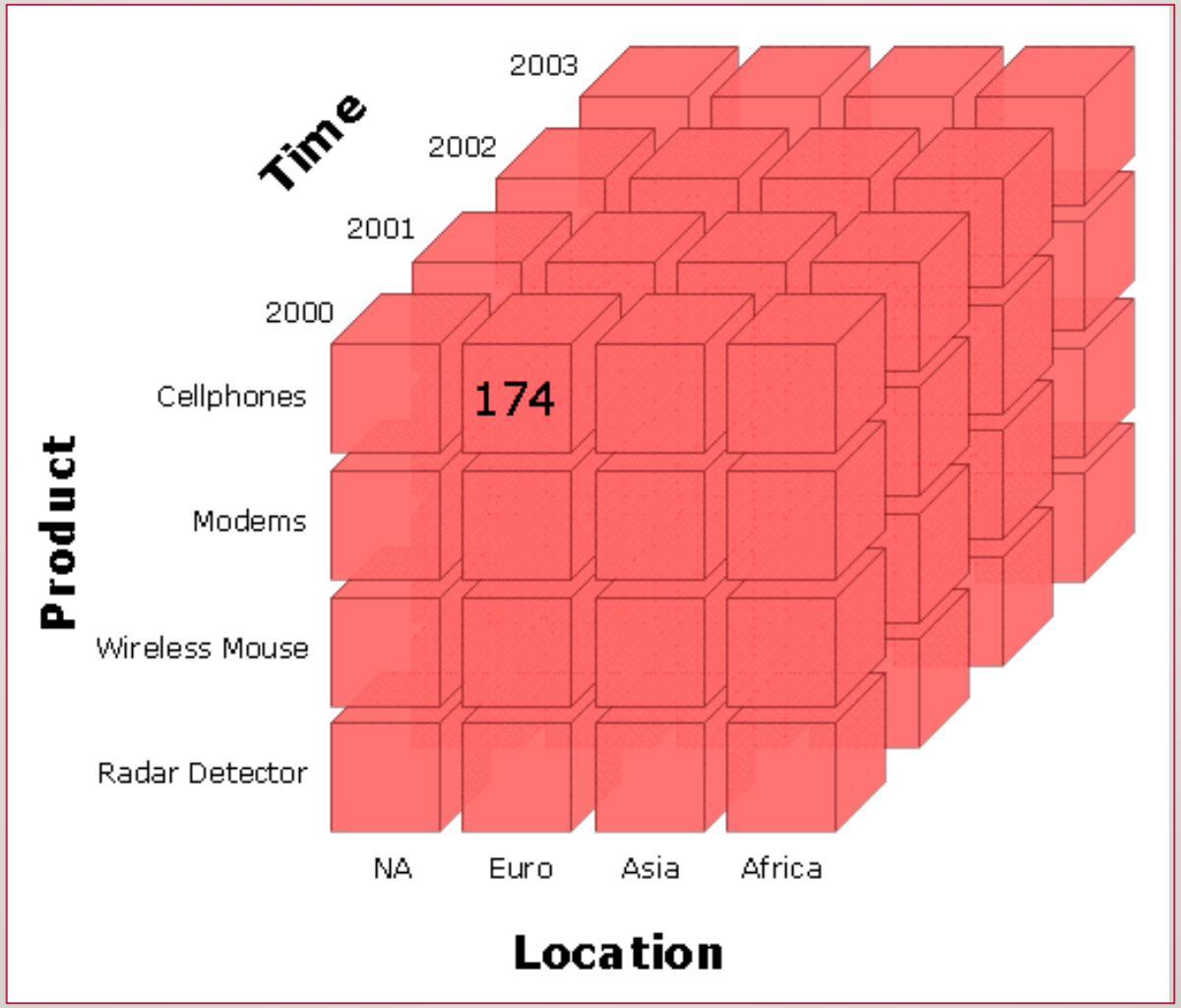
7 VISUALIZING MULTIDIMENSIONAL DATA CUBES

- The data cube is
  - The basic unit of storage and analysis in Analysis Services.
  - A structure that enable OLAP to achieves the multidimensional functionality.
  - Is used to represent data along some measure of interest.
  - An easy way to look at the data (allow us to look at complex data in a simple format).
  - It can be 2-dimensional, 3-dimensional, or higher-dimensional.

VISUALIZING MULTIDIMENSIONAL DATA CUBES TERMINOLOGIES

- Dimension
  - Each cube has one or more dimensions, each based on one or more dimension tables.
  - A dimension represents a category for analyzing business data.
- Fact Table
  - A fact table contains the basic information that you wish to summarize.
  - This might be order detail information, payroll records, or anything else that's amenable to summing and averaging.
  - Measures: represents some fact (or number) such as cost or units of service.

9 VISUALIZING MULTIDIMENSIONAL DATA CUBES TERMINOLOGIES

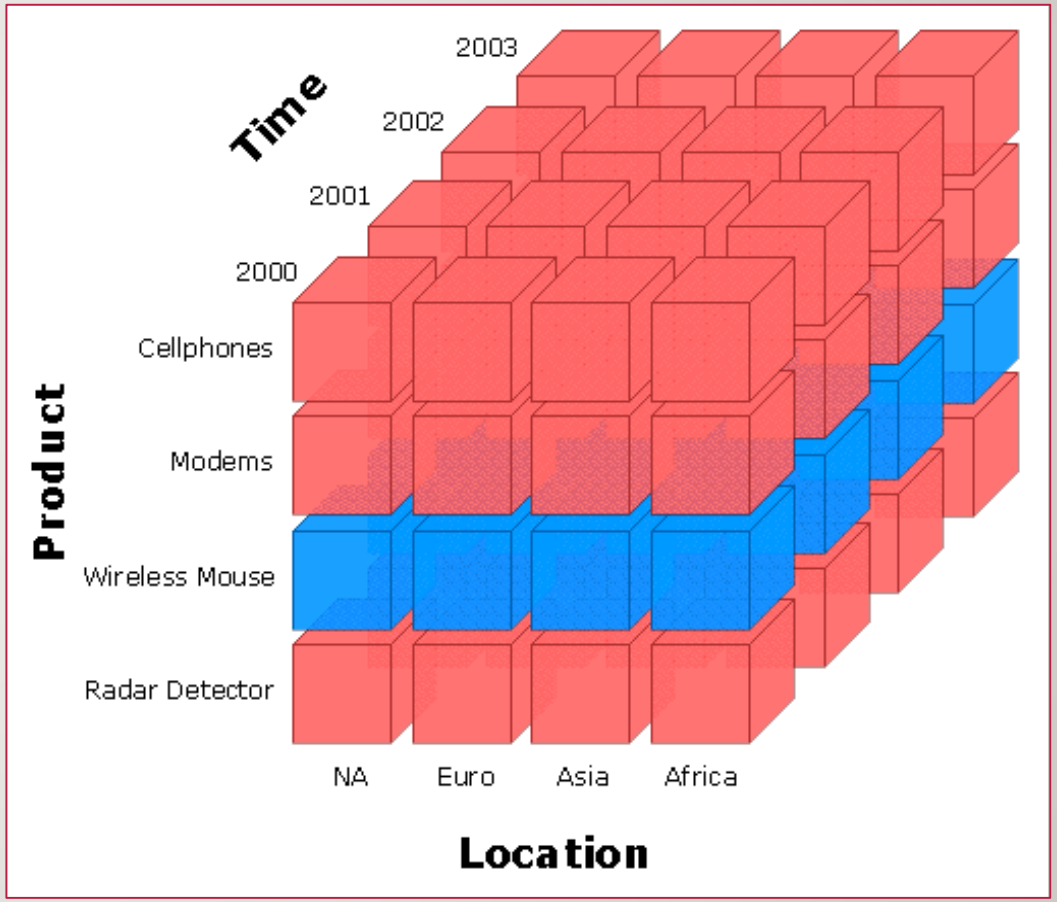


10 DATA CUBES CONCEPTS

- Three important concepts associated with data cubes:
  1. Slicing.
  2. Dicing.
  3. Rotating.

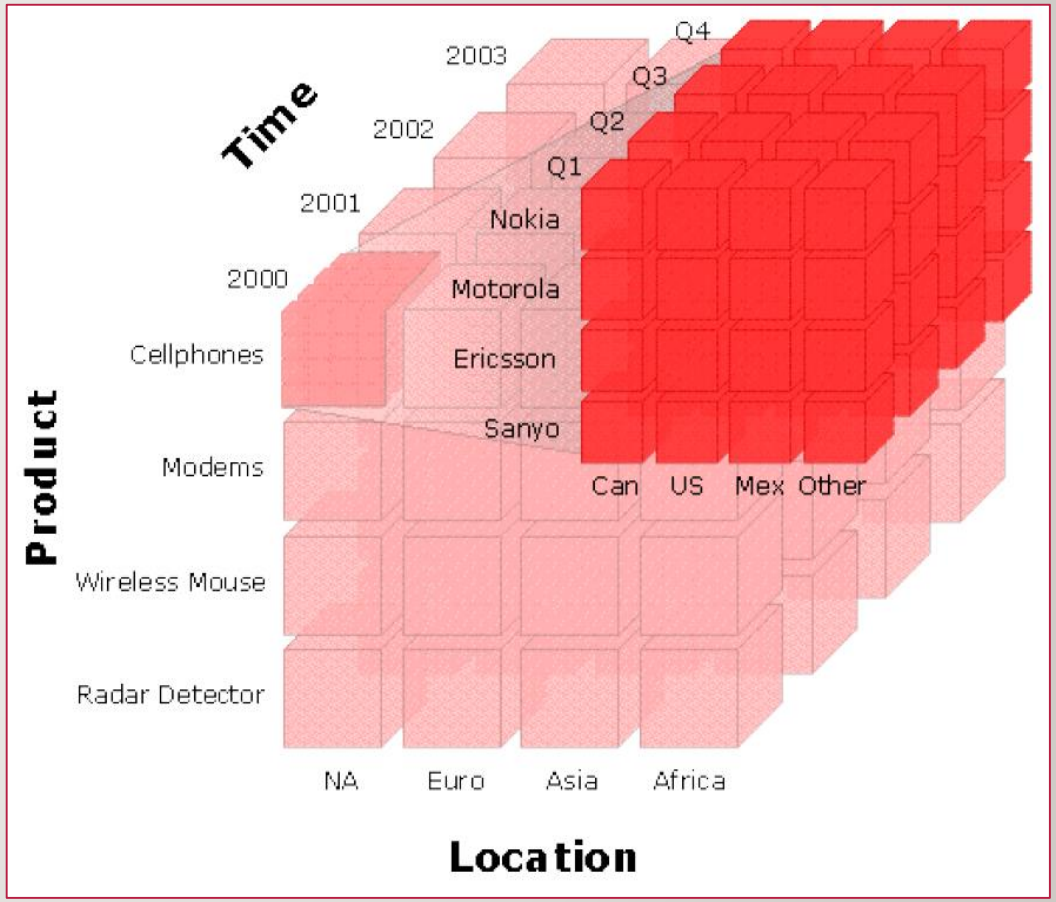
DATA CUBES CONCEPTS SLICING

- The term slice most often refers to a two dimensional page selected from the cube.



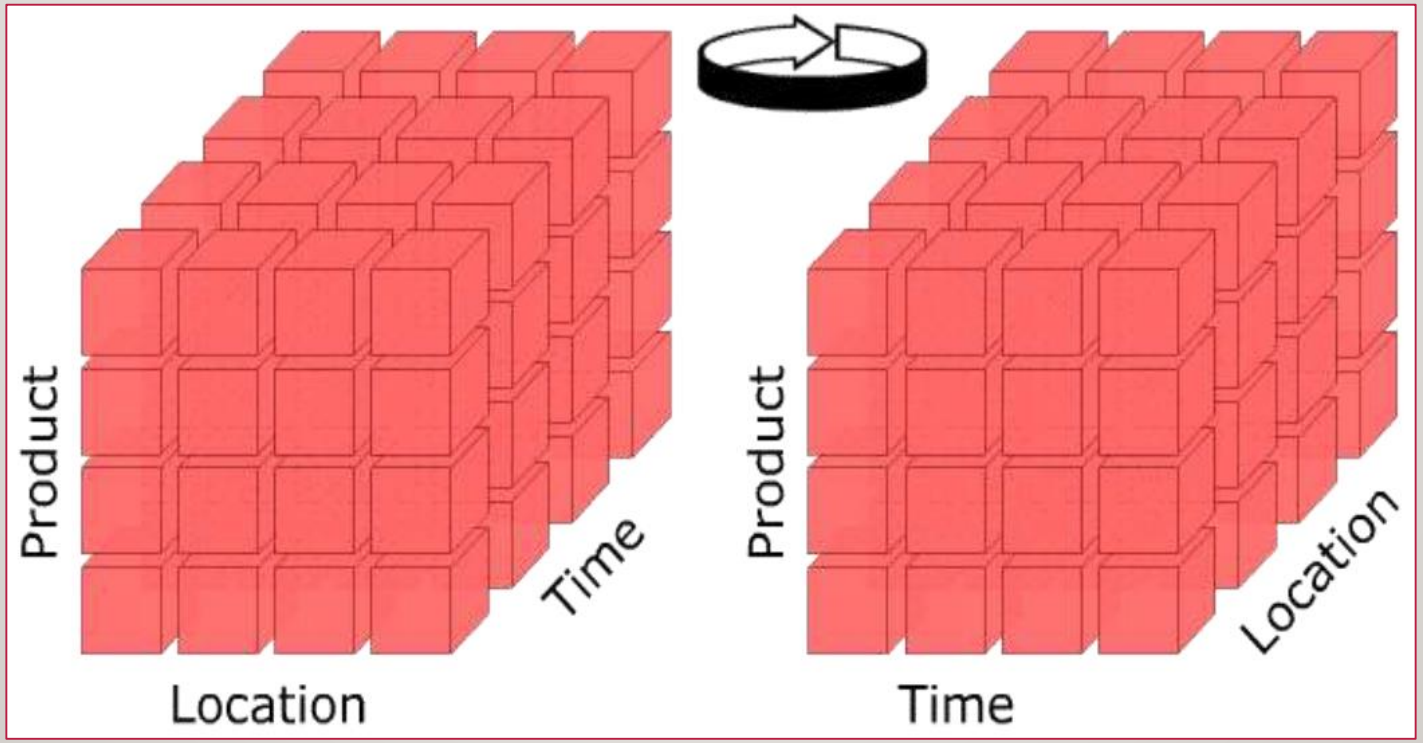
DATA CUBES CONCEPTS DICING

- A related operation to slicing, but in the case of dicing, we define a sub cube of the original space.

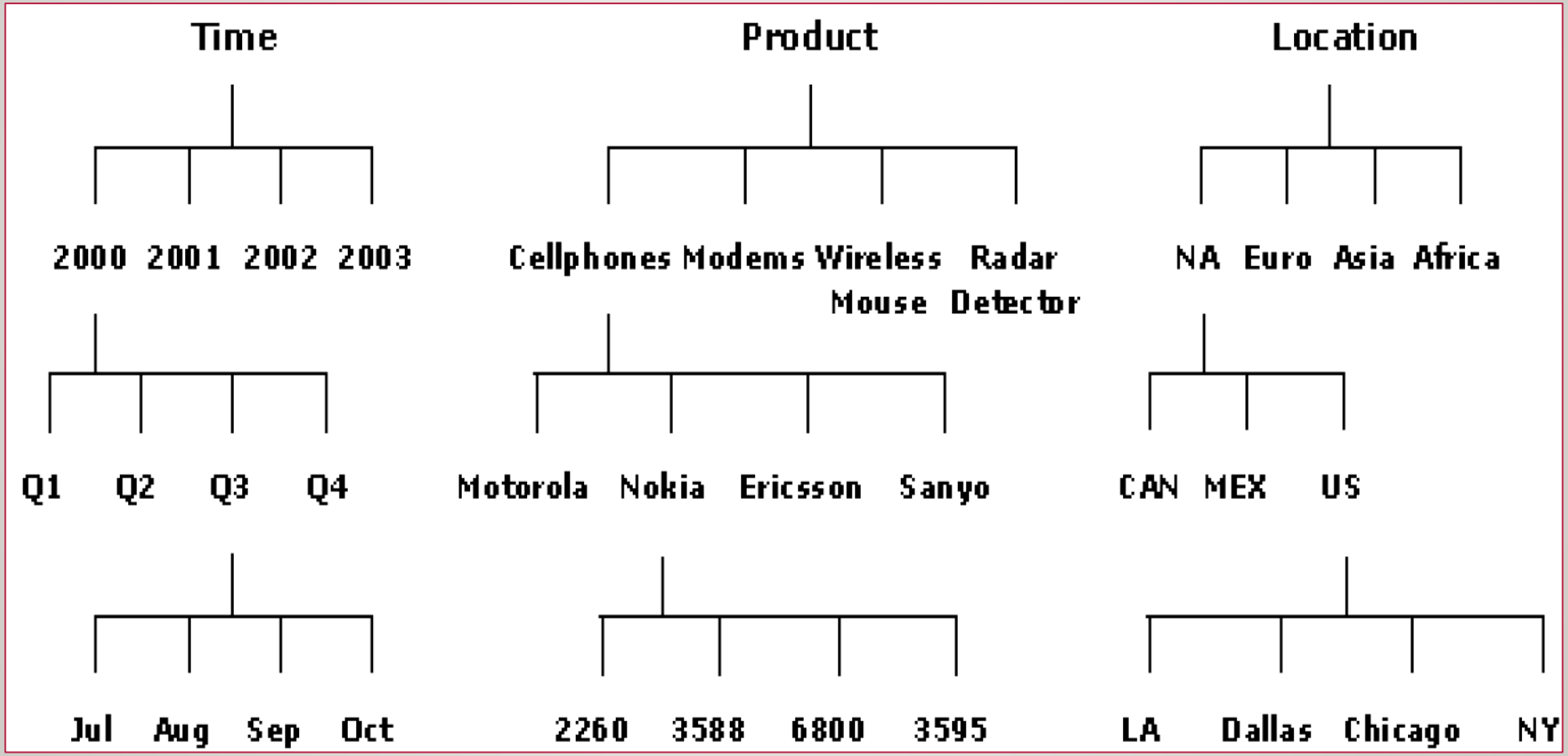


13 DATA CUBES CONCEPTS ROTATING

- Some times called pivoting.
- Rotating changes the dimensional orientation of the report from the cube data.

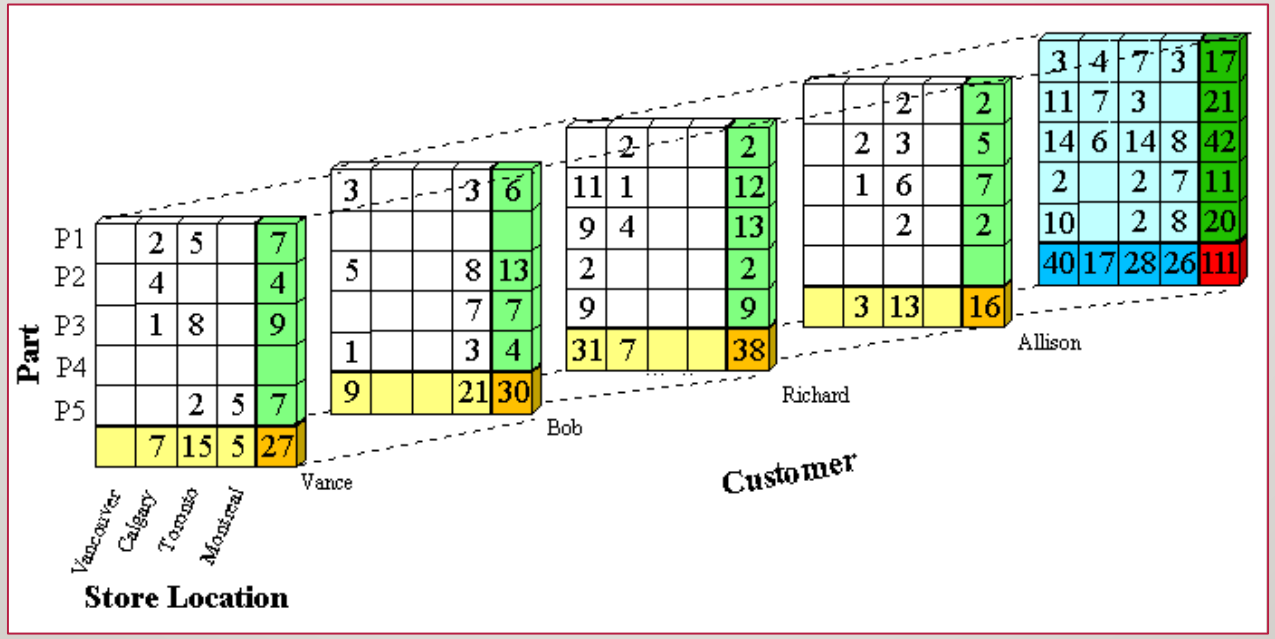


DATA CUBES CONCEPTS DIMENSION HIERARCHY



15 DATA CUBES CONCEPTS REPRESENTATION OF TOTALS

- A simple data cube does not contain totals.
- Storing totals increases the size of the data cube.
- We represent totals by adding an additional layer on n sides of the n-dimensional data cube.



SQL SERVER ANALYSIS SERVICES (SSAS) ADVANTAGES AND FEATURES

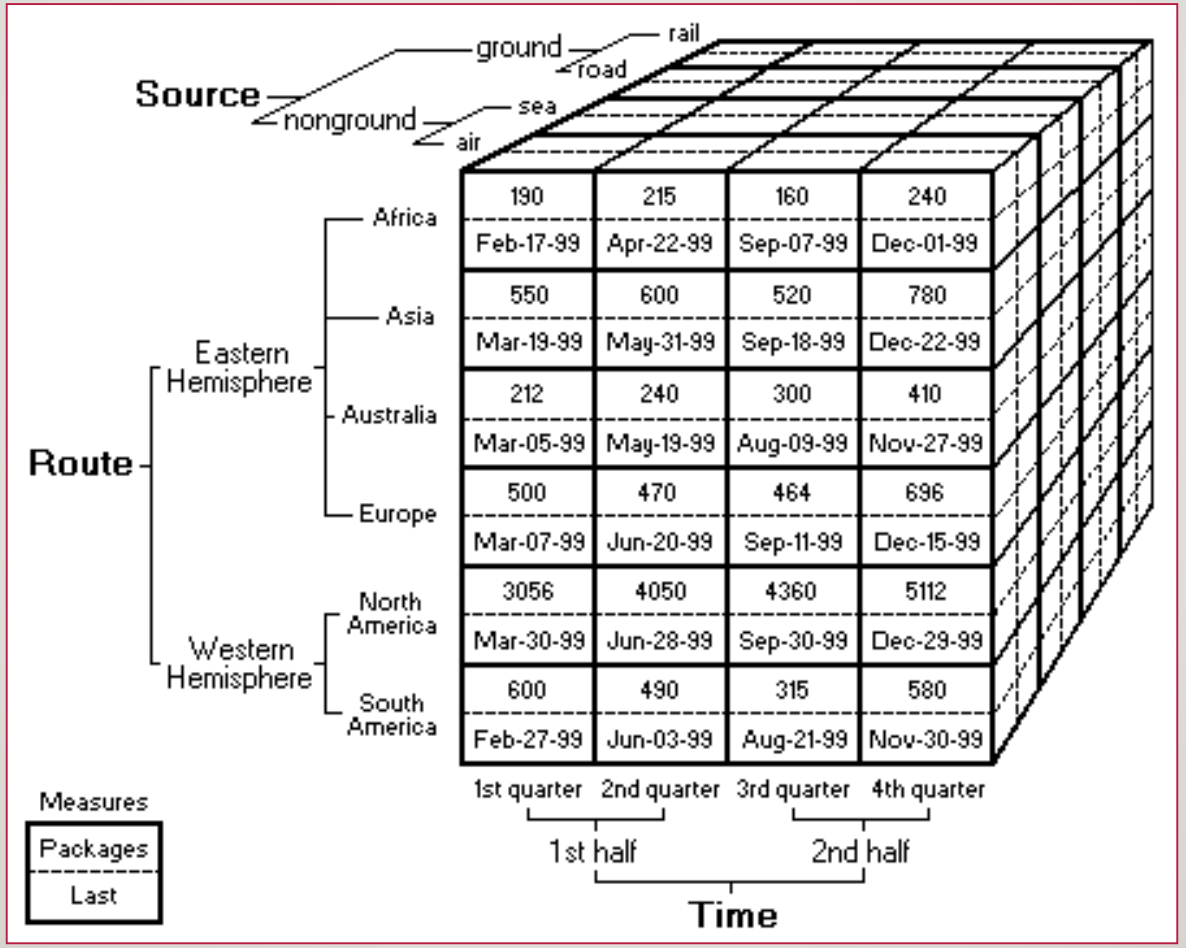
- Ease of use
  - Wizards and editors
  - Data viewers
- Flexible data model
  - Multiple storage options
  - Partitioning
  - Multiple dimension and cube types
  - Write-enabled options
- API's
- Scalability
  - Optimized aggregations
  - Data compression
  - Distributed calculations
  - Partitioning and distributed cubes
- Integration
  - Security
  - Management
  - Other SQL Server tools and features

SQL SERVER ANALYSIS SERVICES (SSAS) OLTP VS OLAP

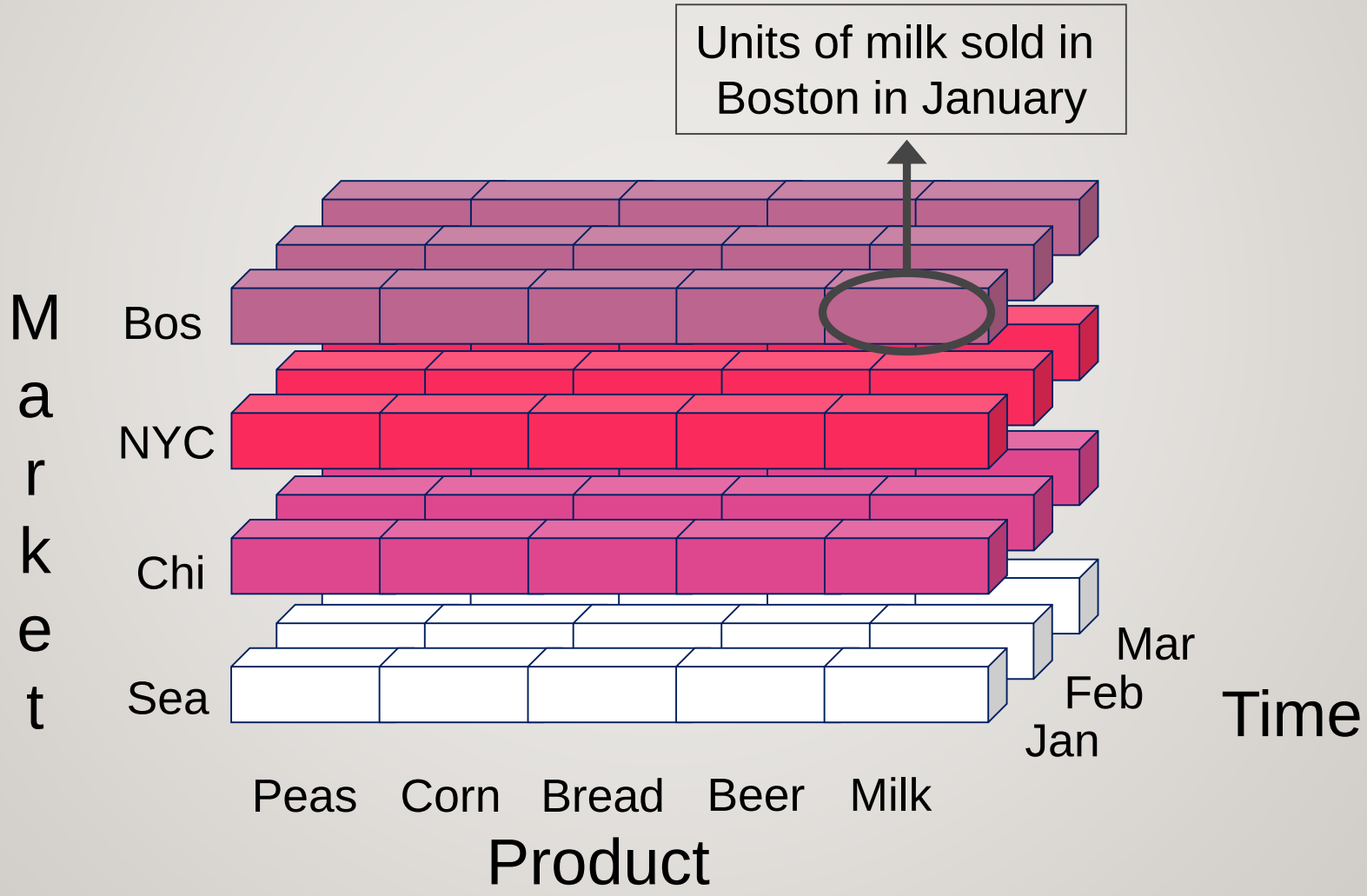
SQL SERVER ANALYSIS SERVICES (SSAS) OLTP VS OLAP

Online Transaction Processing - OLTP (Operational Database)	Online Analytical Processing - OLAP (Data Warehouse)
Support daily operations	Support analyses and forecast business needs
Holds daily latest transactional data	Hold historical data consistent up to the last update occurred in Cube
Data stored in normalized format	Data stored in de-normalized format
Databases size is usually around 100 MB to 100 GB	Databases size is usually around 100 GB to a few TB
Used by normal users	Used by users who are associated with the decision making process
CPU, RAM, HDD space requirement is less	CPU, RAM, HDD space requirement is higher
Query response may be slower if the amount of data is very large, it can impact the reporting performance	Query response is quicker, management can do Trend analysis on their data easily and generate quicker reports
SQL language used for querying on OLTP	MDX is used for querying on OLAP Cube

19 CUBE EXAMPLES



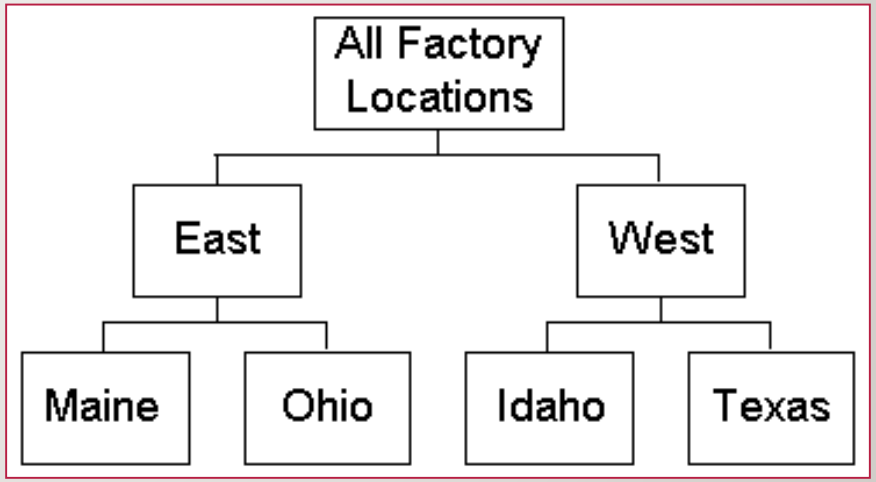
CUBE EXAMPLES



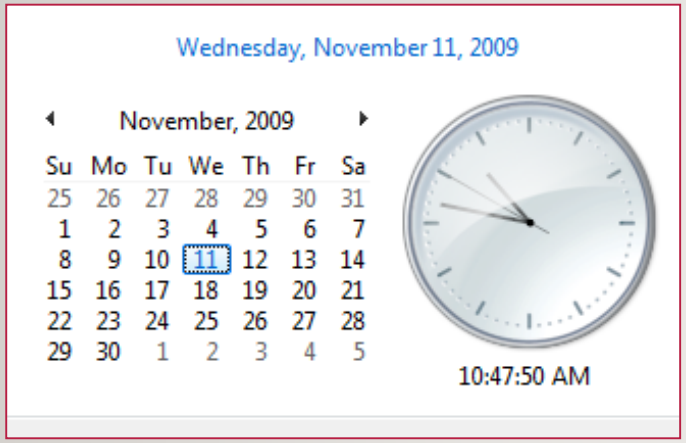
21 DIMENSION EXAMPLES

- Organized hierarchies of categories, levels, and members.
- Used to "slice" and query within a Cube based on an underlying dimension table.

State_ID	Region	State
1	East	Maine
2	East	Ohio
3	West	Idaho
4	West	Texas

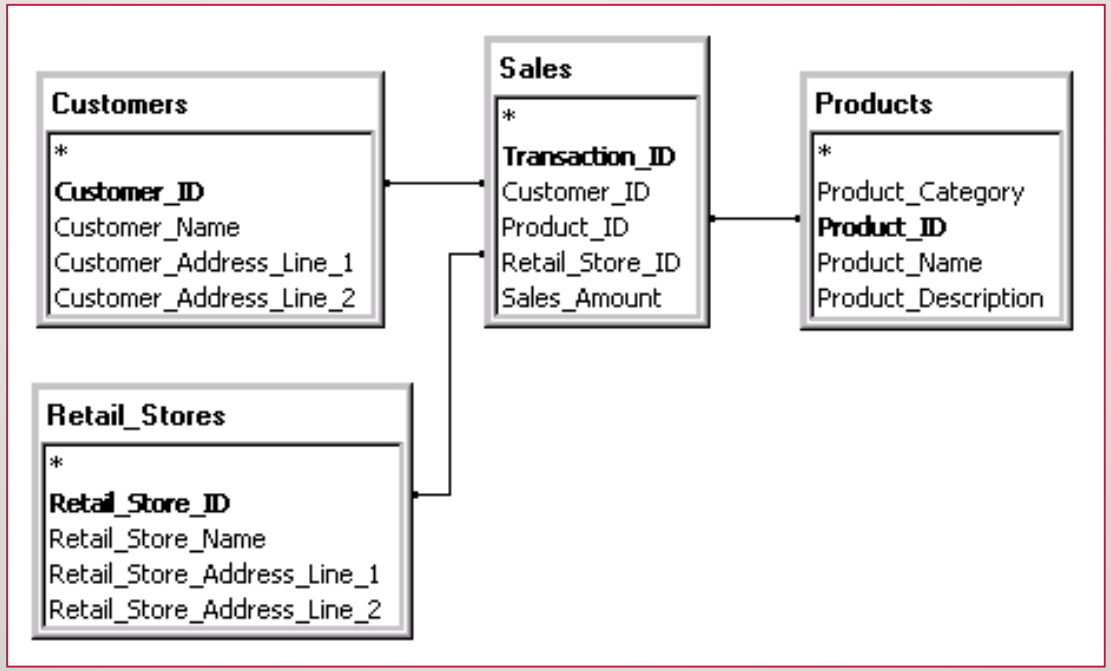


22 DIMENSION EXAMPLES



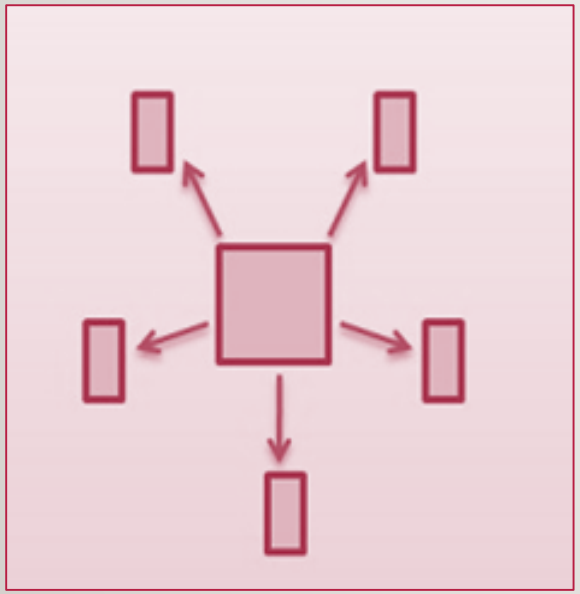
23 MEASURE EXAMPLES

- Contain the data users are interested in
- Created using an aggregation function.
- Examples: Qty Sold, Sales Amount, Cost

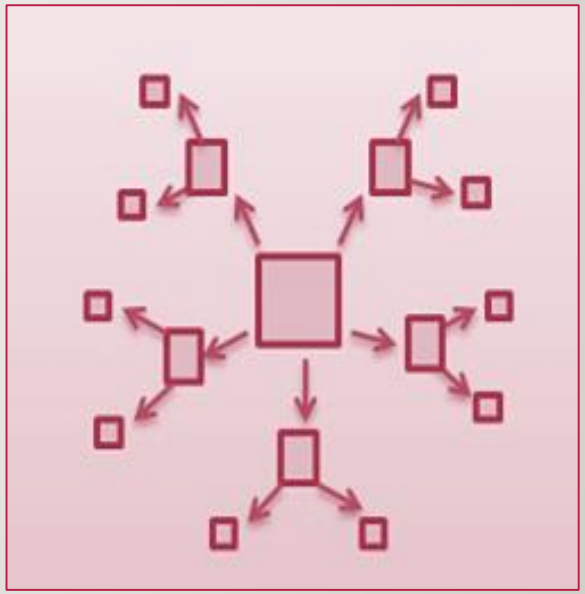


24 OLAP SCHEMA

STAR SCHEMA



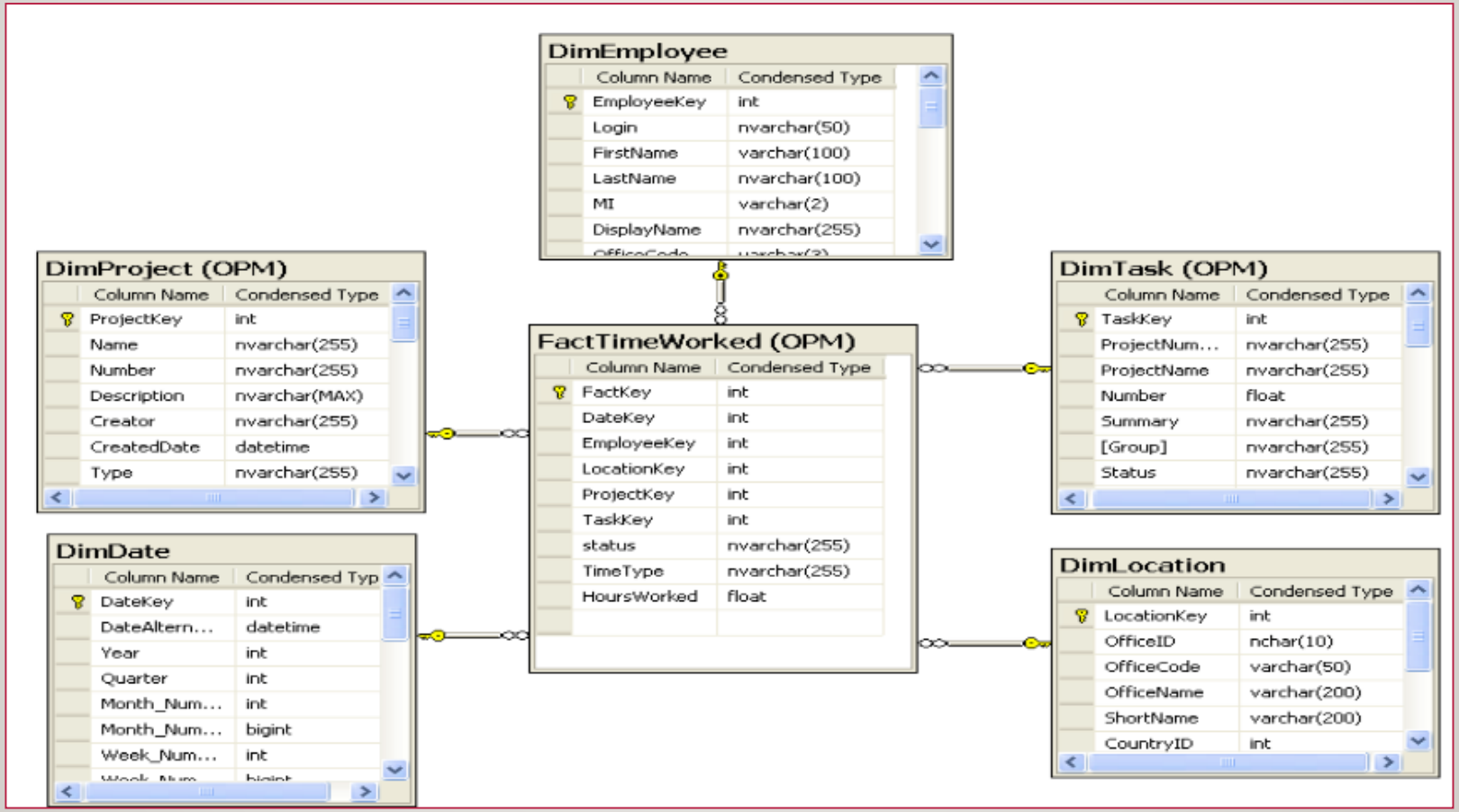
SNOWFLAKE SCHEMA



25 OLAP STAR SCHEMA

- A star schema consists of at least one fact table and it points towards a number of dimension tables.
- Star Schema is highly recommended schema for SSAS cubes.

OLAP STAR SCHEMA EXAMPLE



27 OLAP SNOWFLAKE SCHEMA

- A star schema structure normalized through the use of outrigger tables.
  - Dimension table hierarchies are broken into simpler tables.
- In OLAP, this snow flake schema approach increases the number of joins and poor performance in retrieval of data.

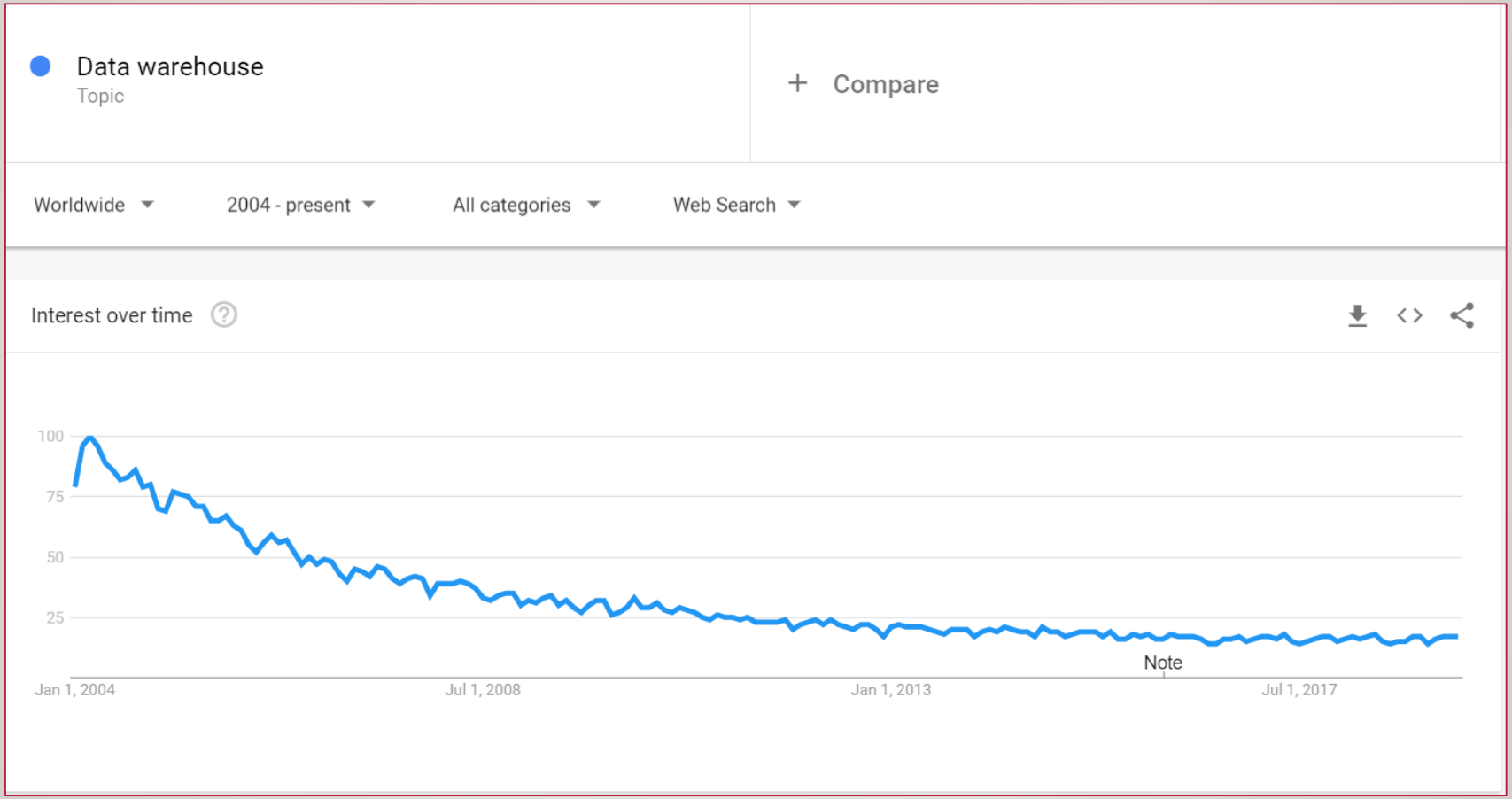
28 STAR SCHEMA V.S. SNOW FLAKE SCHEMA

- In a star schema every dimension will have a primary key.
- In a star schema, a dimension table will not have any parent table.
- Whereas in a snow flake schema, a dimension table will have one or more parent tables.
- Hierarchies for the dimensions are stored in the dimensional table itself in star schema.
- Whereas hierarchies are broken into separate tables in snow flake schema. These hierarchies helps to drill down the data from topmost hierarchies to the lowermost hierarchies.

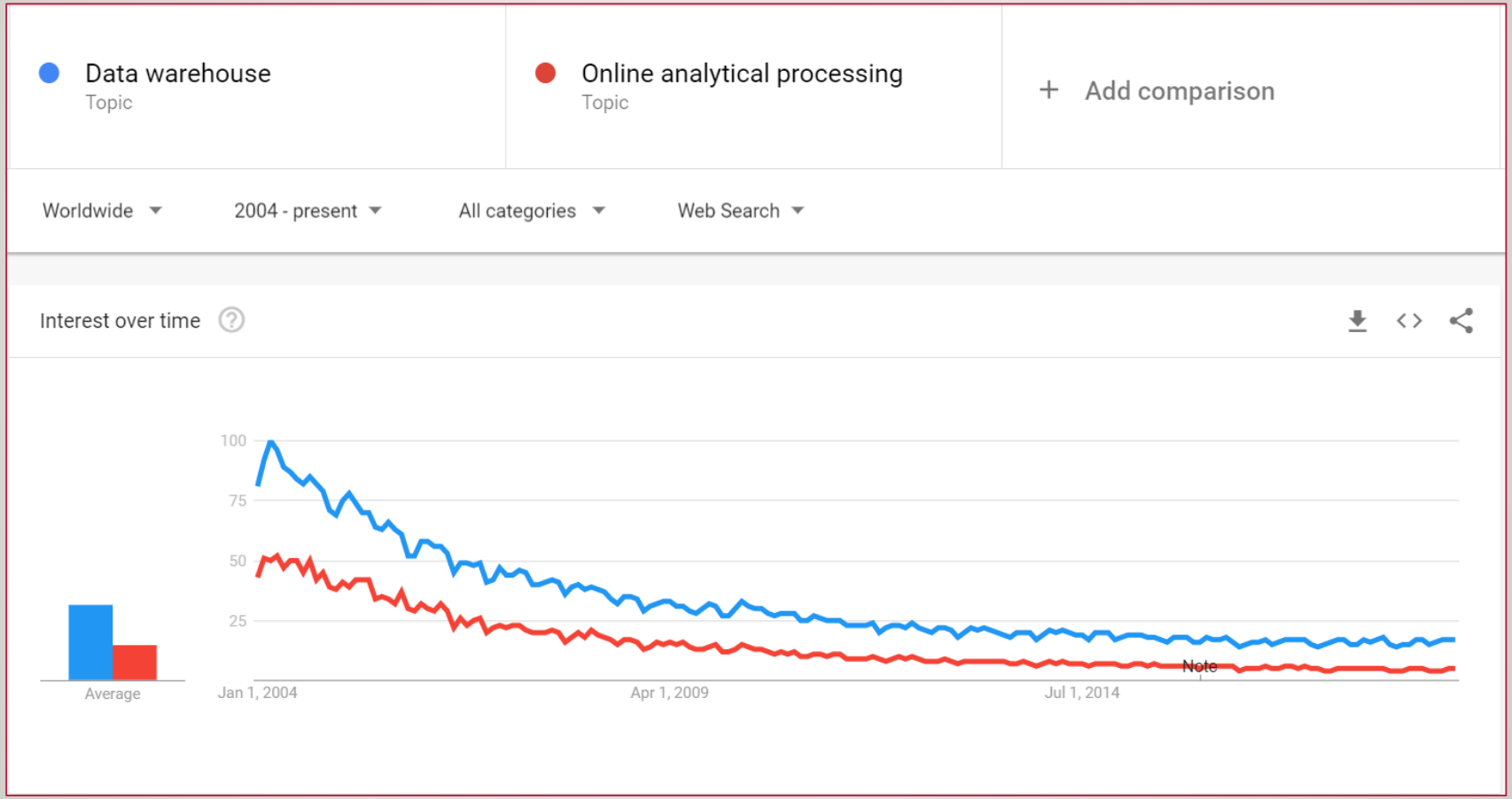
BUILDING THE CUBE IN SSAS

- Preconditions
  - Connecting data sources
  - Defining views
  - Selecting dimensions
- Define fact & dimension tables & time dimension
- Select measures
- Deploy & query the cube
- Demo

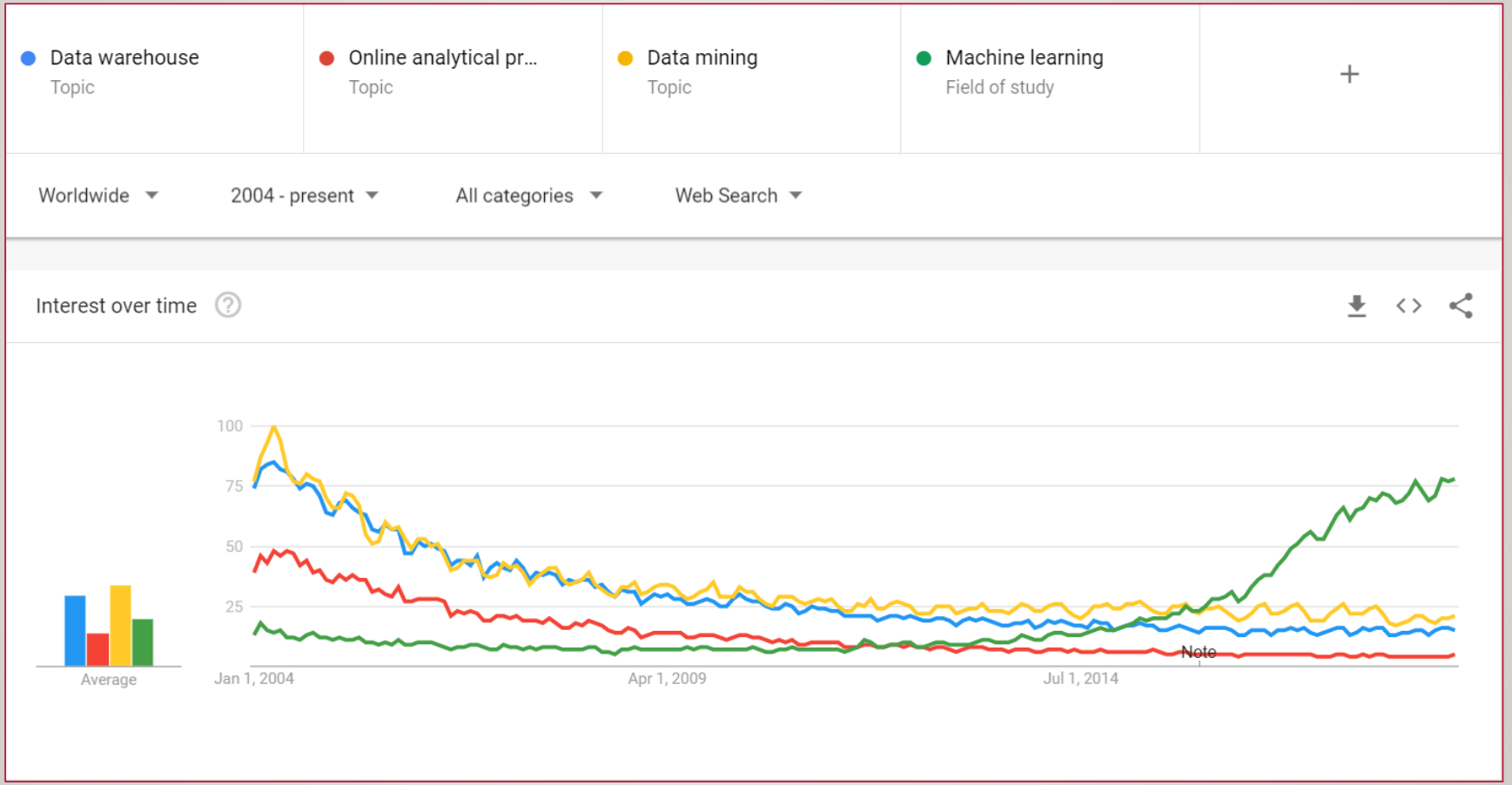
30 SOME STATISTICS



31 SOME STATISTICS



32 SOME STATISTICS



THE END