

Instructor Notes:

Data Warehousing Concepts

Copyright © 2011 IGATE Corporation. All rights reserved. No part of this publication shall be reproduced in any way, including but not limited to photocopy, photographic, magnetic, or other record, without the prior written permission of IGATE Corporation.
IGATE Corporation considers information included in this document to be Confidential and Proprietary.

Instructor Notes:

Explain the lesson coverage

DWH Concepts

A dimension table is a table containing data that defines the dimension. A time dimension for example could contain dates, names of the days of the week, week numbers, months names, month numbers and year. A fact table contains the measures, that is aggregatable data that can be counted, summed, multiplied, etc.

Fact tables also contain references (foreign keys) to the dimension tables in the cube so the facts can be grouped by the dimensional data.

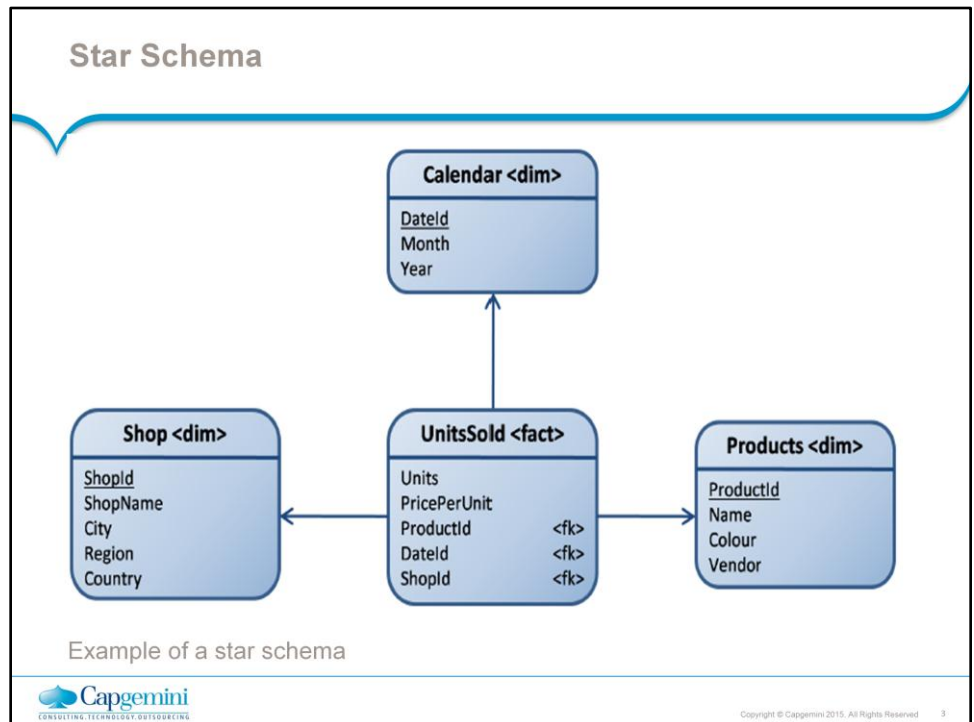
A data warehouse is generally structured as a star schema or a snowflake schema.



Copyright © Capgemini 2019. All Rights Reserved 2

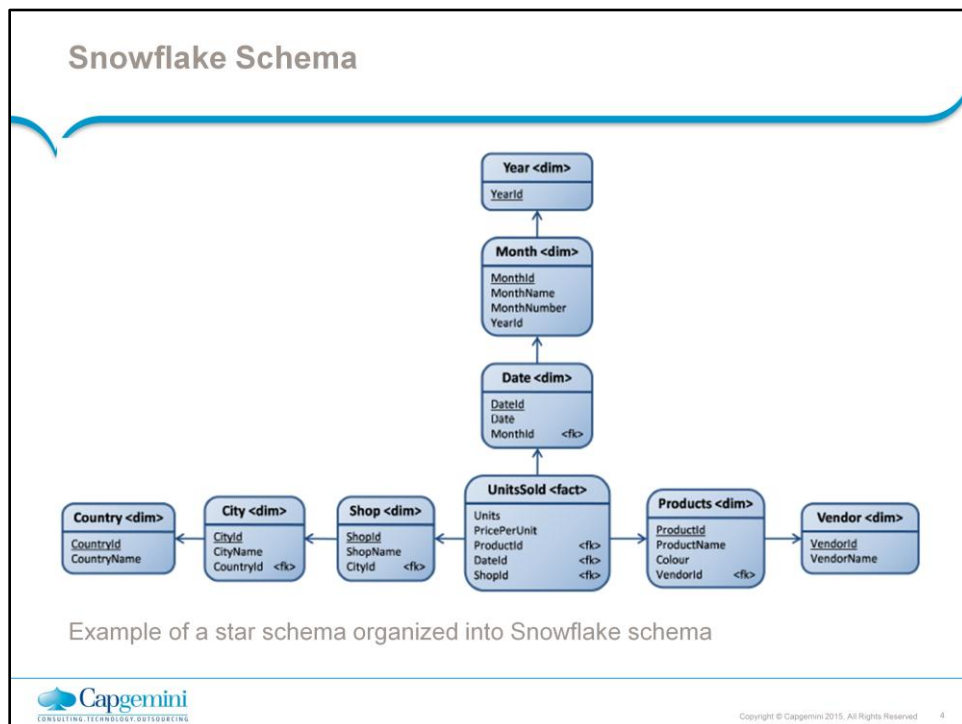
Instructor Notes:

Explain the lesson coverage



Instructor Notes:

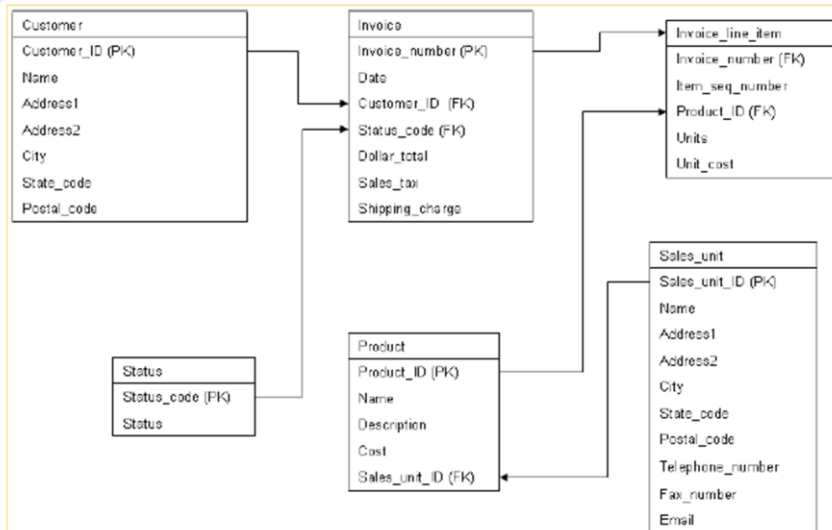
Explain the lesson coverage



Instructor Notes:

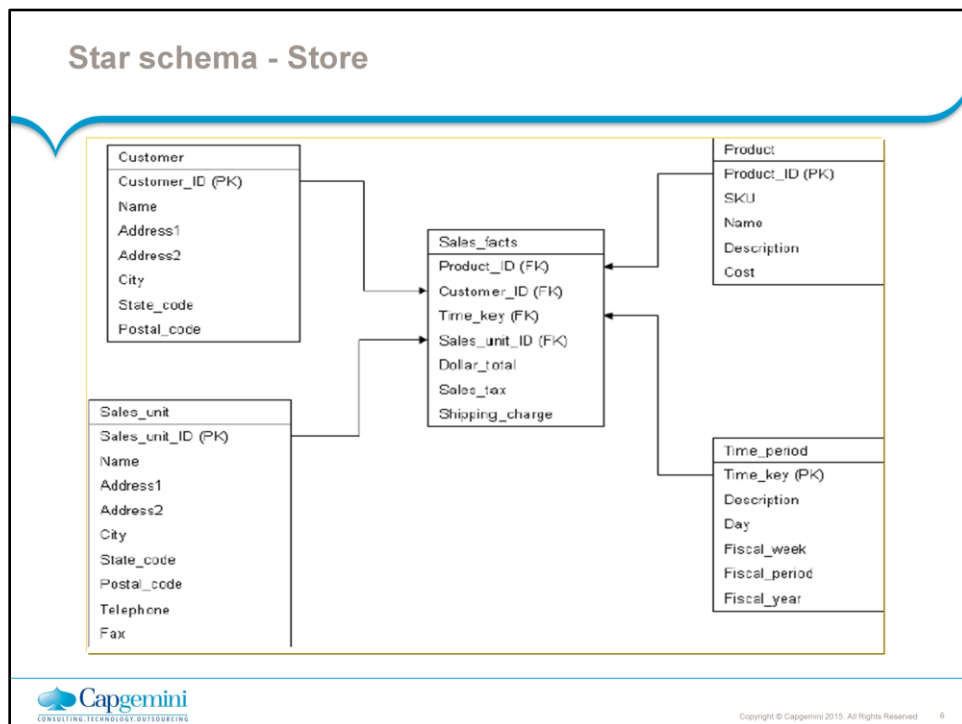
Explain the lesson coverage

Normalized database schema - Store



Instructor Notes:

Explain the lesson coverage



Instructor Notes:

Explain the lesson coverage

Multidimensional Conceptual View

The end user's view of the OLAP model must be multidimensional since the analysis to be done is multidimensional by nature. Note that having a multidimensional view doesn't necessarily mean that the data is actually stored multidimensionally.

There are two main architectures that traditionally have been discussed in the field of database research; Multidimensional OLAP (MOLAP) and Relational OLAP (ROLAP).

MOLAP: In this type of OLAP, a cube is aggregated from the relational data source (data warehouse). When user generates a report request, the MOLAP tool can generate the create quickly because all data is already pre-aggregated within the cube.

Multidimensional Online Analytical Processing is based upon the philosophy that since the cube is multidimensional in its nature, the data should be stored multidimensionally.

Instructor Notes:

Explain the lesson coverage

Multidimensional Conceptual View

Thus, the data is copied from the data warehouse to the cube storage and aggregations of different combinations of dimensions are pre-calculated and stored in the cube in an array based data structure.

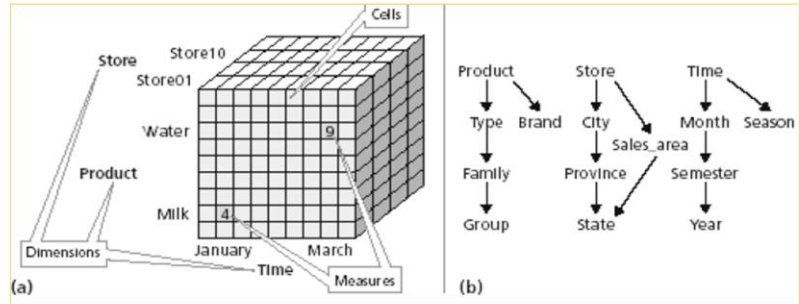
This means that the query response time is very short since no calculations have to be done at the time a query is executed.

Instructor Notes:

Explain about OLAP.
Describe what all it involves.

Multidimensional Conceptual View

Consider the database of sales information maintained by a company. This particular data cube has three feature attributes - *store*, *product*, and *time* - and a single *measure attribute* - product sales for a large chain of stores (sales is computed with the sum function).

Practical data cube example

Instructor Notes:

Describe the nature of OLAP analysis.

ROLAP

Relational Online Analytical Processing is, just as the name suggests, based on the relational model.

The main idea here is that it is better to read data from the data warehouse directly, than to use another kind of storage for the cube.

As is the case with MOLAP, data can be aggregated and pre-calculated in ROLAP too, using materialized views, i.e. storing the aggregations physically in database tables.

A ROLAP architecture is more flexible since it can pre-calculate some of the aggregations, but leave others to be calculated on request.

Instructor Notes:

Describe the nature of OLAP analysis.

Relational Cubes

A retailer company has sold jeans and gloves in two colors; black and blue.

The total amount of items sold during 2007 and 2008 is summed separately for each year and color in the Sales column.

Now, a manager would probably want to look at the figures in the form of sums of the total.

Product	Year	Colour	Sales
Jeans	2007	Black	231
Jeans	2007	Blue	193
Jeans	2008	Black	205
Jeans	2008	Blue	236
Gloves	2007	Black	198
Gloves	2007	Blue	262
Gloves	2008	Black	168
Gloves	2008	Blue	154

Instructor Notes:

Describe the nature of OLAP analysis.

Relational Cubes

Sales during one year

- the total amount of jeans sold during 2007,
- or maybe the amount of blue gloves sold during 2008.

You can arrange the data for this purpose to use a pivot table. In this kind of table, data is structured with both row and column labels and is also summed along its dimensions in order to get a clear view of the data set.

Example of a Pivot Table is given below

	2007		2007	2008		2008	Grand Total
	Black	Blue		Black	Blue		
Jeans	231	193	424	205	236	441	865
Gloves	198	262	460	168	154	322	782
Grand Total	429	455	884	373	390	763	1647



Instructor Notes:

Describe the nature of OLAP analysis.

Relational Cubes

The cube operator expresses the pivot table data set in a relational database context.

This operator is basically a set of group by clauses put together with a union operation.

It groups the data by the given dimensions (attributes), in this case *Product*, *Year* and *Colour*, and aggregations are made for all possible combinations of dimensions.

Below table illustrates the resulting data set when the cube operator is used combined with the SUM operator.

Instructor Notes:

Explain the scenario about OLAP application.

Relational Cubes

Product	Year	Colour	Sales	Product	Year	Colour	Sales
Jeans	2007	Black	231	Gloves	2008	ALL	322
Jeans	2007	Blue	193	Gloves	ALL	Black	366
Jeans	2007	ALL	424	Gloves	ALL	Blue	416
Jeans	2008	Black	205	Gloves	ALL	ALL	782
Jeans	2008	Blue	236	ALL	2007	Black	429
Jeans	2008	ALL	441	ALL	2007	Blue	455
Jeans	ALL	Black	436	ALL	2007	ALL	884
Jeans	ALL	Blue	429	ALL	2008	Black	373
Jeans	ALL	ALL	865	ALL	2008	Blue	390
Gloves	2007	Black	198	ALL	2008	ALL	763
Gloves	2007	Blue	262	ALL	ALL	Black	802
Gloves	2007	ALL	460	ALL	ALL	Blue	845
Gloves	2008	Black	168	ALL	ALL	ALL	1647
Gloves	2008	Blue	154				

Instructor Notes:

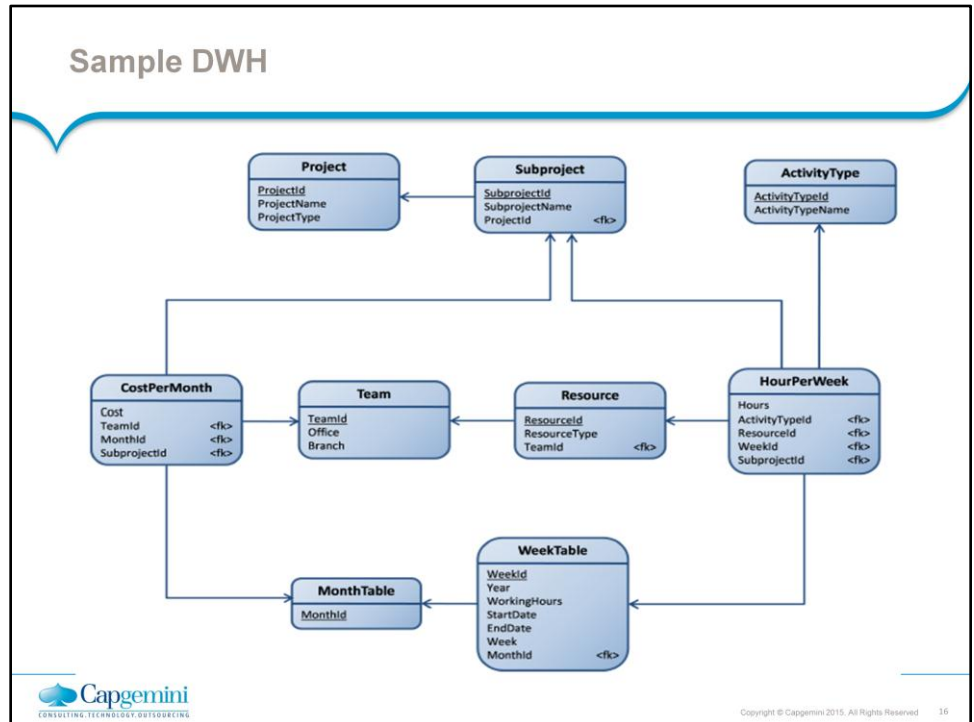
Relational Cubes

Rollup operator works the same way as the cube operator, but with the difference that new aggregations are calculated only from already calculated aggregations.

Product	Year	Colour	Sales	Product	Year	Colour	Sales
Jeans	2007	Black	231	Gloves	2007	Black	198
Jeans	2007	Blue	193	Gloves	2007	Blue	262
Jeans	2007	ALL	424	Gloves	2007	ALL	460
Jeans	2008	Black	205	Gloves	2008	Black	168
Jeans	2008	Blue	236	Gloves	2008	Blue	154
Jeans	2008	ALL	441	Gloves	2008	ALL	322
Jeans	ALL	ALL	865	Gloves	ALL	ALL	782
ALL	ALL	ALL	1647				

Instructor Notes:

Describe in detail ROLAP type. Give advantages & disadvantages of ROLAP.



Instructor Notes:**Answers for the Review Questions:**

- 1.1,3
- 2.True
- 3.Slice & dice

Popular Tools

- [Business Objects](#)
- [IBM Cognos](#)
- [SQL Server Analysis Services](#)
- [MicroStrategy](#)
- Palo OLAP Server