

Team # 1 Name(s): Ellie Parobek, John Warner

PE04: Dimensional Modeling Exercise

Overview

In this exercise, you have an opportunity to investigate an approach for scoping and designing a set of data marts, which logically form a warehouse, starting from the source operational data systems in a realistic business scenario. This is a composite exercise, composed of six (6) separate individual exercises labeled: Exercise 1, Exercise 2, ..., Exercise 6.

After completing this exercise, you should be able to:

- Review data values from source operational systems to determine their potential use in a data mart/data warehouse.
- Identify possible dimensions and fact groups.
- Design a set of dimensions by identifying dimension attributes and any hierarchies.
- Design a set of fact groups.
- Use a matrix tool, called a Data Mart Matrix, to identify dimensions that need to be conformed across data marts.
- Design a STAR schema for a fact group.

Business Scenario

You are working in the IT department of TelComprehensive Inc., a large, multinational telephone company. The company provides telecommunications service, both land based and cellular, to customers across the Western Hemisphere. Your team has been asked to build a set of conformed data marts for an overall data warehouse development project.

The company keeps track of the following corporate-wide processes:

- Customer service and billing inquiries
- Customer billing
- Equipment installation and service
- Yellow page advertising
- Marketing
- Call detail (billing perspective)
- Call detail (network switching perspective)
- Customer inventory (phone equipment, Centrex systems, features)
- Network equipment inventory (switches, lines, etc)
- Real estate inventory (poles, buildings, etc)
- Employees and payroll
- Computer system job processing &
- chargeback Purchase orders to suppliers
- Deliveries from suppliers

For this phase of the development, data marts will be sourced from three
(3) main operational systems:

Call Detail Tracking This system records the details about the individual phone calls made by the company's customers. Information is maintained in the following categories:

- The customer: home address and the billing address
- The call: when, from where, to whom, how long the call lasted
- Any phone service, i.e. "product," provided by type, class and product group
- Pricing: the pricing package to which the call applies, rates, any special discounts that apply (daytime, evening, weekend)

Service Call Tracking This system keeps track of the details related to the calls for repair services from the company's customers. Information is maintained in the following categories:

- The service call: problem categorization, call request and resolution dates, the number of days required to resolve the problem, the number of customer calls related to this problem, and the status of the service request (open; closed)
- The employee who handled the call: name, address and demographic information, employment information
- The cost center¹ that handled the service request
- The customer who placed the service request: name, address

Materials Inventory Since this is a service business, this inventory system focuses on the supplies and spare parts used in their service calls. They do not re-manufacture or resell the items that they purchase for this purpose.

This system allows the company to evaluate their inventory situation at an individual item basis and to establish policies for the efficient control of inventory levels. It includes:

- Information about items used for service/repair: the categorization of each item, when it was purchased, the quantity currently on hand, the quantity on order, and a target "optimum" inventory level
- The buyer who purchases the items for the company
- The cost center that uses these items and where it fits within the corporate hierarchy

Details about the meaning and data formats of the data values sourced from these systems are shown on the following pages.

¹ For the purposes of this scenario, a cost center is a business entity – i.e. a department or a division – that has a budget, that maintains information pertinent to cost generation, and that the ability to track and control those costs. While they add to the operational costs of the organization they do not directly add to the profits. Examples are Research and Development, Marketing, and Customer Service, etc.

Dimensional Modeling Exercise 1: Identifying Facts and Dimensions

- Suppose you are employed by a large telephone company, and your task is to build the data marts for the overall data warehouse
- You will be extracting data from three operational systems
- Identify each field
 - D - a dimensional attribute
 - F - a fact
 - O - operational only, not to be included

D/F/O	Call Detail Tracking
D	Record Type
D	Customer Name
D	Call Date
D	Call Time
F	Minutes of Use
D	Bill to Street Address
D	Bill to City
D	Bill to State
D	Bill to Zip
D	Bill to Country
D	Customer Street Address
D	Customer City
D	Customer State
D	Customer Zip
D	Customer Country
D	Product Name
D	Product Type
D	Product Class
D	Product Group
D	Price Package
D	Weekend Discount Description
D	Weekend Discount %
D	Weekend Rate
D	Evening Discount Description
D	Evening Discount %
D	Evening Rate
D	Daytime Discount Description
D	Daytime Discount %
D	Daytime Rate
F	Origin Phone Number
D	Origin Phone Type
F	Destination Phone Number
D	Destination Phone Type
D	Transaction ID

D/F/O	Material Inventory
D	Inventory Date
D	Purchased Item
D	Purchased Sub Category
D	Purchased Category
D	Purchased Product Group
D	Buyer
D	Business Unit
D	Business Region
D	Business Division
D	Business Company
D	Cost Center City
D	Cost Center State
D	Cost Center Country
F	Quantity on Hand
F	Quantity on Order
O	Target Inventory Quantity

D/F/O	Service Call Tracking
D	Service Call Date
D	Customer Name
D	Customer Street Address
D	Customer City
D	Customer State
D	Customer Zip
D	Customer Country
D	Service Call Type
D	Service Call Group
D	Service Call Status
D	Cost Center City
D	Cost Center State
D	Cost Center Country
D	Resolution Date
D	Employee Name
D	Employee Status
D	Employee Gender
D	Employee Date of Birth
D	Employee Education Level
D	Employee Original Hire Date
D	Employee Current Hire Date
D	Employee Zip
D	Employee Street Address
D	Employee City
D	Employee State
D	Employee Country
F	Days to Resolution
F	Number of Service Calls
D	Transaction ID

Dimensional Modeling Exercise 2: Identifying Dimensions and Fact Groups

Using the same extract files, identify

- Possible dimensions
- Possible fact groups

Possible Dimensions

Call Detail Date_____

Customer_____

Product_____

Price Package_____

Service Call Date_____

Service Call Resolution Date_____

Employee_____

Cost Center_____

Material Inventory Date_____

Purchase_____

Fact Groups

Call Details_____

Service Call_____

Material Inventory_____

Dimensional Modeling Exercise 3: Designing Dimensions

Design each of the dimensions that were identified in Exercises 1 & 2.

- Identify dimension attributes
- Identify all hierarchies of attributes within a dimension

Call Detail Date -> Year, Month, Day, Hour, Minute, Second

Customer -> Country, State, City, Zip, Street Address
-> Billing Country, Billing State, Billing City, Billing Zip, Billing Street Address

Product -> Name
-> Type
-> Group, Class

Price Package -> Evening Discount Description
-> Evening Discount Percent
-> Evening Rate
-> Weekend Discount Description
-> Weekend Discount Percent
-> Weekend Rate
-> Daytime Discount Description
-> Daytime Discount Percent
-> Daytime Rate

Service Call Date -> Year, Month, Day

Service Call Resolution Date -> Year, Month, Day

Employee -> Name
-> Status
-> Gender
-> Date of Birth
-> Education Level
-> Original Hire Date
-> Current Hire Date
-> Country, State, City, Zip, Street Address

Cost Center -> Business Unit, Business Region, Business Division, Business Company
-> Country, State, City

Material Inventory Date -> Year, Month, Day

Purchase -> Group, Category, Sub-Category, Purchased Item
-> Buyer

Dimensional Modeling Exercise 4: Designing Fact Groups

Design each of the fact groups that were identified in Exercises 1 & 2.

Fact Group: Call Details

Fact Name	Fact Description	Default Aggregation Rule
Minutes of Use	The total number of minutes for this call	Semi-additive over time
Origin Phone Number	The specific phone number from where a call was placed	Sum
Destination Phone Number	The specific phone number that was called	Sum

Fact Group: Service Call

Fact Name	Fact Description	Default Aggregation Rule
Days to Resolution	The total number of days from the date the service call was opened to the date that the service call was closed	Semi-additive over time
Number of Service Calls	Indicates the number of service calls that were taken for a given time period	Semi-additive over time

Fact Group: Material Inventory

Fact Name	Fact Description	Default Aggregation Rule
Quantity on Hand	Indicates the number of units of this product that are available on site at this organization location	Semi-additive over time
Quantity on Order	The number of units of this product that have been ordered, but have not yet been shipped	Semi-additive over time

Dimensional Modeling Exercise 5: Create the Data Mart Matrix

The data mart matrix shows the relationship between the possible data marts and dimensions. Any dimension (column) with more than one X implies that this dimension must be conformed across multiple data marts in order to fit into the Data Warehouse Bus Architecture. Fill in the data mart matrix using the following format:

Fact Groups	Call Detail Date	Customer	Product	Price Package	Service Call Date	Service Call Resolution Date	Employee	Cost Center	Material Inventory Date	Purchase
Call Details	X	X	X	X						
Service Call		X			X	X	X	X		
Material Inventory								X	X	X

Dimensional Modeling Exercise 6: Logical Table Design

- Use the dimensional model that you have created so far
- Design the actual star schema for the fact groups

Create your Dimensional Model using MySQL Workbench, save it as a pdf file, and one member from each team to submit it to MyCourses PE04 Drop Box. And bring a hard copy of the pdf file to next class.

Each team member should submit a peer evaluation form to the dropbox. No PE04 grade without the Peer Evaluation form submitted. The form is available on MyCourses.

