Team # 5 Name(s): Muriel Banze; Sahil Mehul Shah; Siddharth Chauhan

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 center[1] 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
O	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
D	Origin Phone Number
D	Origin Phone Type
D	Destination Phone Number
D	Destination Phone Type
O	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
F	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 Data 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
O	Transaction ID

Call Detail Tracking

Attribute Name	Attribute Description	Sample Values
Record Type	Call Detail Record	
Customer Name	Unique identifier for each customer. This will be represented in multiple ways including name and customer identifier. For residential customers, a single individual is designated	Commercial Corp.; Kyle Jones
Call Date	The specific day that an activity took place.	06/04/1998; 06/05/1998
Call Time	The minute that a customer is using telephone service.	11:32 AM; 3:07 PM
Minutes of Use	The total number of minutes for this call.	
Bill to Street Address	The street address for this customer's billing address. Note: the address number, suite/box, street name will be split apart for complete flexibility.	123 Main Street, Suite 1000
Bill to City	The name of the city of this customer's billing address.	Sunny Town
Bill to State	The name of the state of this customer's billing address.	New York
Bill to Zip	The 5 digit zip code for this customer's billing address.	12345
Bill to Country	The name of the country of this customer's billing address.	USA
Customer Street Address	The street address for this customer location. Note: the address number, suite/box, street name will be split apart for complete flexibility.	123 Main Street, Suite 1000
Customer City	The name of the city where this customer location is.	Sunny Town
Customer State	The name of the state where this customer location is.	New York
Customer Zip	The 5 digit zip code for this customer location.	12345
Customer Country	The name of the country where this customer location is.	USA
Product Name	Describes the individual product or service offering that a customer purchased.	Cell Air Time; Local Toll Service; Voice Mail; Call Waiting
Product Type	Indicates whether this is a specific product or a service offering.	Product; Usage Service
Product Class	A collection of products.	Cellular Phone; Cellular Usage; Fixed Line; Line Enhancements
Product Group	A collection of product classes.	Cellular; Stationary
Price Package	Uniquely describes a collection of price information that is sold as a package to customers.	One Rate; Deep Discount Weekends; Call Anytime; Voicemail
Weekend Discount Description	Details the specific terms and conditions associated with weekend usage.	Unlimited Use

Weekend Discount %	Indicates the percentage discount off regular price for all weekend usage.	40%
Weekend Rate	Indicates the actual rate for all weekend usage.	10 cents per minute
Evening Discount Description	Details the specific terms and conditions associated with evening and holiday usage.	Discount only with minimum of $20 usage per month
Evening Discount %	Indicates the percentage discount off regular price for all evening usage.	25%
Evening Rate	Indicates the actual rate for all evening and holiday usage.	12 cents per minute
Daytime Discount Description	Details the specific terms and conditions associated with daytime usage.	Regular residential rates
Daytime Discount %	Indicates the percentage discount off regular price for all daytime usage.	0%
Daytime Rate	Indicates the actual rate for all daytime usage.	25 cents per minute
Origin Phone Number	The specific phone number from where a call was placed.	444-555-1212
Origin Phone Type	The type of phone from where the call was placed.	Cell; Toll
Destination Phone Number	The specific phone number that was called.	444-555-1212
Destination Phone Type	The type of phone that was called.	Cell; Toll
Transaction ID		

Material Inventory

Attribute Name	Attribute Description	Sample Values
Inventory Date	The specific day that an activity took place.	06/04/1998; 06/05/1998
Purchased Item	Identifies the item that was purchased by Telco. This is represented on the invoice.	Executive Cherry Finish Desk; Large Coil Wire; Phone Cords
Purchased Sub Category	A collection of purchased items as defined by the purchasing department.	Desks; Chairs;
Purchased Category	A collection of sub categories that are related.	Telephone Parts; Wiring Parts; Cellular Components; Furniture; Office Supplies
Purchased Product Group	High level group of items that are purchased by Telco.	Customer Maintenance; Finished Products; Internal Use
Buyer	Represents the individual within the purchasing department that is responsible for this category.	Getty A. Deihl; Layne Hatch
Business Unit	A grouping of cost centers.	Indiana/Ohio; Greater Chicago; Illinois/Missouri
Business Region	Business units that are grouped geographically for field organization and home office by functional area.	Northeast; Midwest; Central Marketing; Finance Staff
Business Division	Major operating areas of the company. Reflects the most recent reorganization into vertical business groups.	Small Business Support; Commercial Service; Residential Service
Business Company	Represents the different companies as the separate legal entities.	Telco Services; CellCo Cellular Services; Business Support
Cost Center City	The city where this cost center is based.	Downers Grove; Cleveland
Cost Center State	The state or province where this cost center is based.	Illinois; Ohio
Cost Center Country	The country where this cost center is based.	USA
Quantity on Hand	Indicates the number of units of this product that are available on site at this organization location.	
Quantity on Order	The number of units of this product that have been ordered, but have not yet been shipped.	
Target Inventory Quantity	The optimal inventory level for this product at this location.	

Service Call Tracking

Attribute Name	Attribute Description	Sample Values
Service Call Date	The specific day that an activity took place.	06/04/1998; 06/05/1998
Customer Name	Unique identifier for each customer. This will be represented in multiple ways including name and customer identifier. For residential customers, a single individual is designated	Commercial Corp.; Kyle Jones
Customer Street Address	The street address for this customer location. Note: the address number, suite/box, street name will be split apart for complete flexibility.	123 Main Street, Suite 1000
Customer City	The name of the city where this customer location is.	Sunny Town
Customer State	The name of the state where this customer location is.	New York
Customer Zip	The 5 digit zip code for this customer location.	12345
Customer Country	The name of the country where this customer location is.	USA
Service Call Type	Represents the type of call that resulted in the need for a work order to be generated.	Line out of order; Static on line; Additional Line Requested;
Service Call Group	Represents general grouping of service call types.	New Request; Problem Report
Service Call Status	Represents the status of a customer service call.	Open; Closed
Cost Center City	The city where this cost center is based.	Downers Grove; Cleveland
Cost Center State	The state or province where this cost center is based.	Illinois; Ohio
Cost Center Country	The country where this cost center is based.	USA
Resolution Date	.	
Employee Name	Represents individuals employed by Telco. Employees will be able to be viewed by employee number, name and social security number.	Lois Alden; Clark Smith
Employee Status	The status of the individuals employment.	Active; Retired; No Longer with Company
Employee Gender	Indicates the gender of this employee.	Male; Female
Employee Data of Birth	Indicates the date of the birth for this employee.	04/24/1974
Employee Education Level	Indicates the highest level of formal education that the employee has completed.	High School; Bachelors; Masters
Employee Original Hire Date	Indicates the first hire date for this employee.	09/15/1986

Employee Current Hire Date	Indicates the most recent date that the employee was hired.	08/01/1991
Employee Zip		
Employee Street Address		
Employee City		
Employee State		
Employee Country		
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	
Number of Service Calls	Indicates the number of service calls that were taken for a given time period.	
Transaction ID		

Dimensional Modeling Exercise 2: Identifying Dimensions and Fact Groups

Using the same extract files, identify

- Possible dimensions
- Possible fact groups

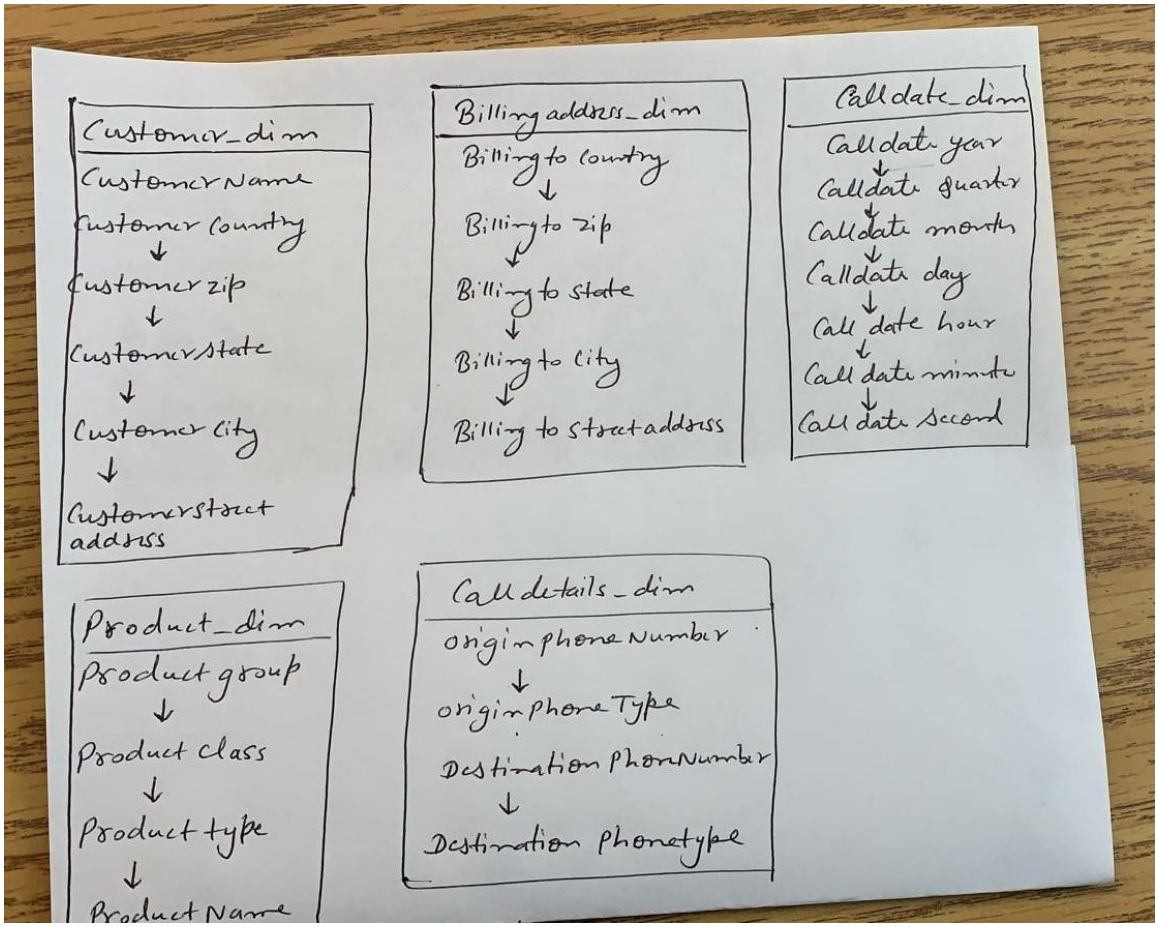
Possible Dimensional Tables:
Customer
Billing Address
Call Date
Product
Call Details (Origin, Destination)
Package (Weekend, Evening, Daytime)
Purchase
Inventory Date
Business
Cost Center
Service Call
Service Call Date
Resolution Date
Employee

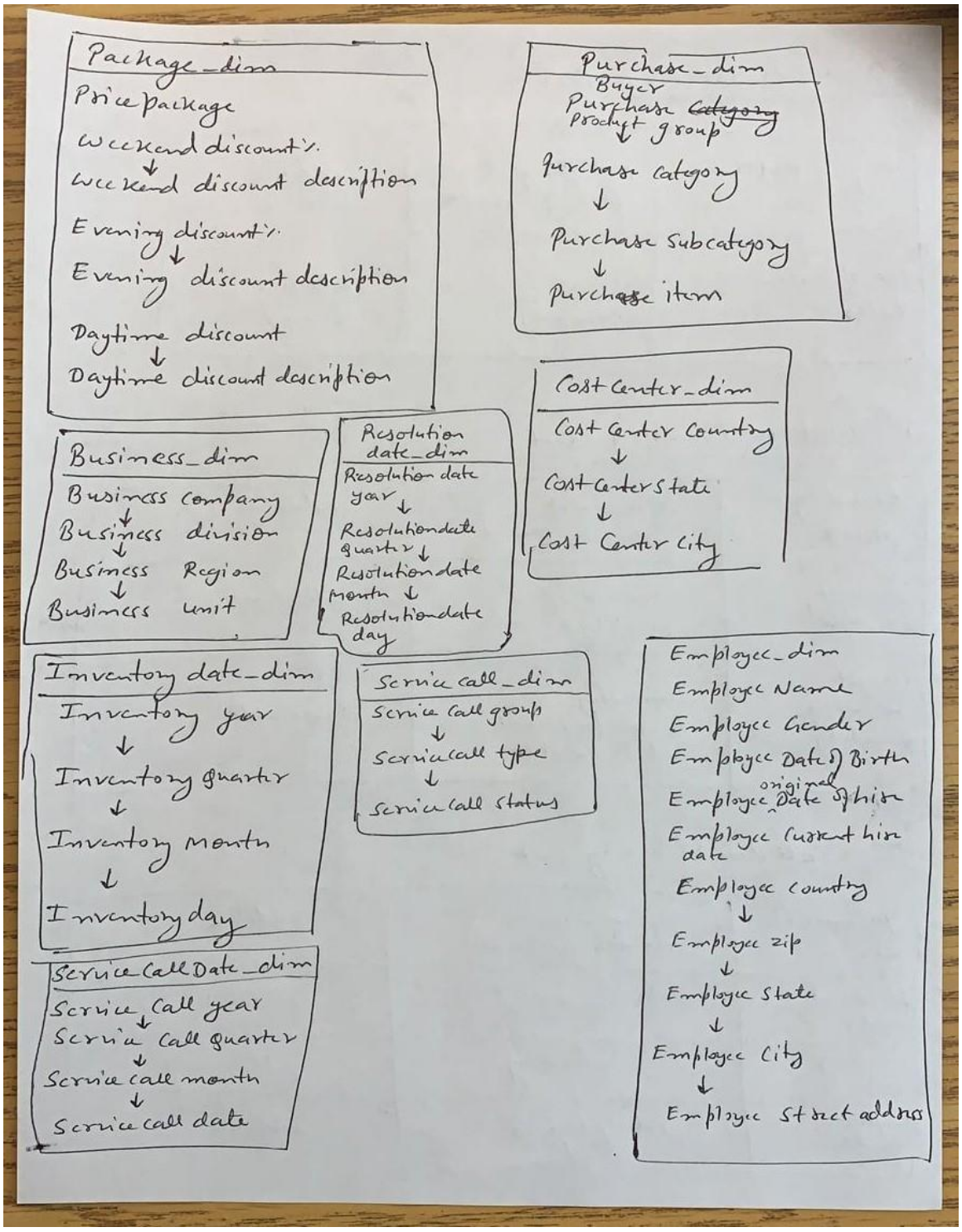
Possible Fact Tables:
Call Detail Tracking
Material Inventory
Service Call Tracking

Dimensional Modeling Exercise 3: Designing Dimensions

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

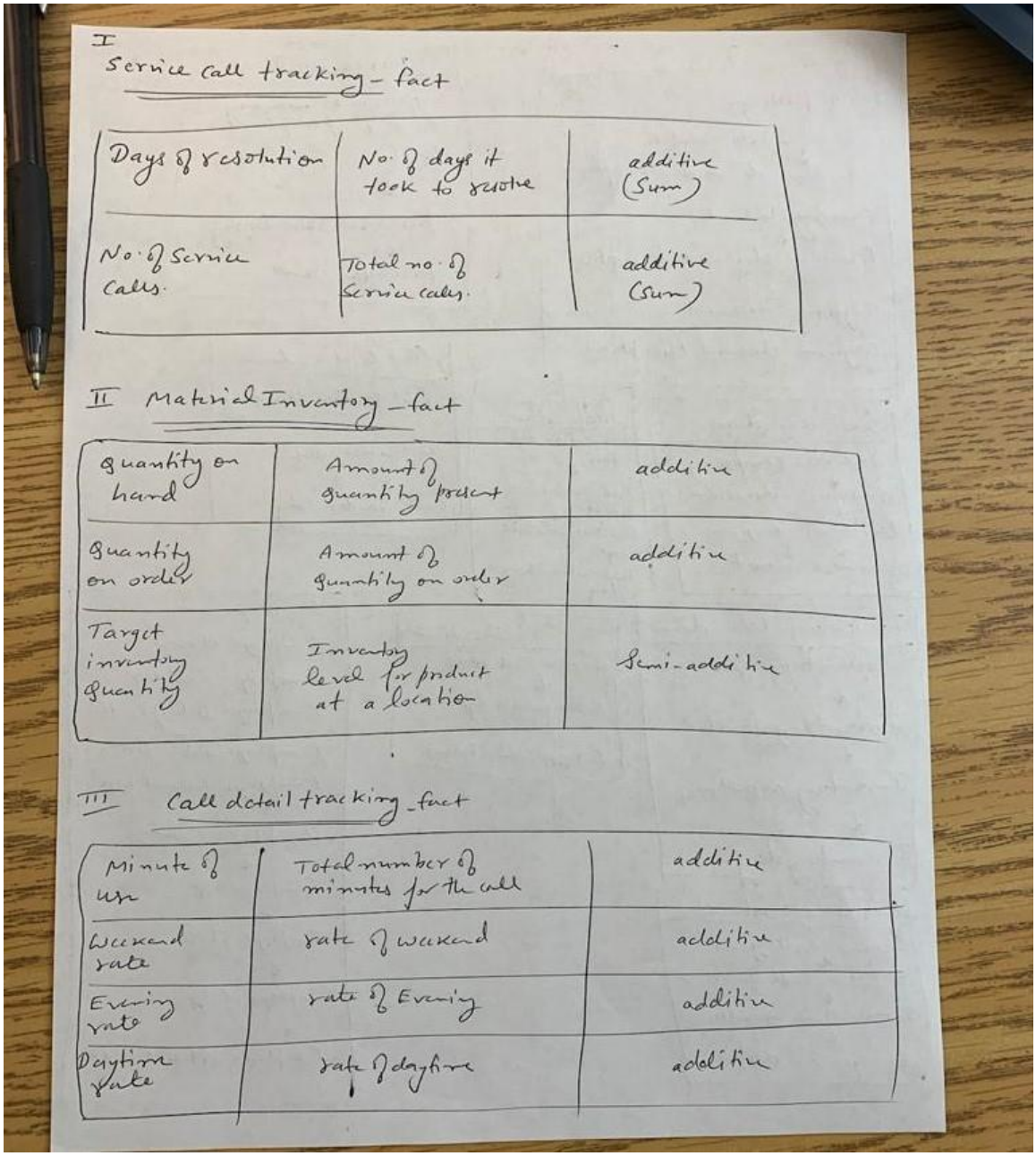
- Identify dimension attributes
- Identify all hierarchies of attributes within a dimension




Package_dim
Price package
Weekend discount %
Weekend discount description
Evening discount %
Evening discount description
Daytime discount
Daytime discount description
Purchase_dim
Buyer
Purchase ~~Category~~
Product group
Purchase category
Purchase subcategory
Purchase item
Cost Center_dim
Cost Center Country
Cost Center State
Cost Center City
Business_dim
Business company
Business division
Business Region
Business unit
Resolution date_dim
Resolution date year
Resolution date quarter
Resolution date month
Resolution date day
Inventory date_dim
Inventory year
Inventory quarter
Inventory month
Inventory day
Service call_dim
Service call group
Service call type
Service call status
Employee_dim
Employee Name
Employee Gender
Employee Date of Birth
Employee original Date of hire
Employee current hire date
Employee country
Employee zip
Employee State
Employee City
Employee Street address
Service Call Date_dim
Service Call year
Service Call quarter
Service call month
Service call date

Dimensional Modeling Exercise 4: Designing Fact Groups

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



I

Service call tracking - fact

Days of resolution	No. of days it took to resolve	additive (Sum)
No. of Service Calls.	Total no. of Service calls.	additive (Sum)

II Material Inventory - fact

quantity on hand	Amount of quantity present	additive
Quantity on order	Amount of quantity on order	additive
Target inventory Quantity	Inventory level for product at a location	Semi-additive

III Call detail tracking - fact

Minute of use	Total number of minutes for the call	additive
Weekend rate	rate of weekend	additive
Evening rate	rate of Evening	additive
Daytime rate	rate of daytime	additive

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	Dimension Names	Customer	Billing Address	Call Date	Product	Call Details	Package	Purchase	Business	Resolution Date	Cost Center	Inventory Date	Service Call Date	Service Call	Employee
Call Detail Tracking		X		X	X	X	X								
Material Inventory								X	X		X	X			
Service Call Tracking		X								X	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.