|  |  |  |
| --- | --- | --- |
| Name: \_\_\_Muriel Banze\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | Due: \_\_09\_\_/ 20\_/\_20\_\_\_ |
|  | **Rochester Institute of Technology**  **Golisano College of Computing and Information Sciences**  **School of Information** | |

**PE03: Dimensional Modeling**

# Exercise 1: Identifying Facts and Dimensions

* Suppose a regional dairy products company employs you, and your task is to build the data marts for the overall data warehouse
* The company sells products to grocery stores, convenience stores, gas stations, and mass merchandisers
* You will be extracting data from the **Product Returns** operational system and **Sales**

**Forecast spreadsheet**

* Identify each field

D - a dimensional attribute

F - a fact

|  |  |
| --- | --- |
| **D/F/O** | **Sales Forecast** |
| D | Account Rep |
| D | Month |
| D | Item # |
| D | Item Description |
| F | Forecast Units |
| F | Forecast Amount |
| O | Valid Forecast Flag |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

O - operational only, not to be included

|  |  |
| --- | --- |
| **D/F/O** | **Product Returns** |
| D | Customer Account Number |
| D | Product Category |
| D | Product Brand |
| D | Customer Name |
| D | Product Expiration Date |
| O | Product # |
| O | Product Description |
| D | Package Type |
| D | Plant Number |
| D | Manufacturing Line |
| O | Regular/Low fat |
| D | Customer Ship to Street Address |
| D | Customer Ship to City |
| D | Customer Ship to State |
| D | Customer Ship to Country |
| D | SKU (Stock Keeping Unit) |
| F | Returned Quantity |
| D | Returned Reason |
| F | Expired Quantity |
| F | Damaged Quantity |
| D | Damaged Code |
| D | Returned Date |
| D | Sales Rep |
| D | Sales Region |

# 2: Identifying Dimensions and Fact Groups

Using the same extract files that you used in Exercise #1, identify the following:

* Possible dimensions
* Possible fact groups (facts in each data mart)

## Possible Dimensions

\_Customer\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_Item\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_Time\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Product\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_Date\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Plant\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_Sales Rep\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Returne\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_­Damaged\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## Fact Groups

\_Product returns\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_Sales Forecast\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# 3: Designing Dimensions

Design (i.e., draw a diagram of) each of the dimensions that were identified in Exercises 1 & 2. Follow the dimension representation shown on slide #8 of Week 4 Lecture notes). Specifically:

* Identify dimension attributes
* Identify all hierarchies of the attributes within a dimension

Returned reason

Item Returned

Sales Rep

Region

Sales Representative

Sales Rep ID

Damaged

Damaged Code

Customer

Account Number

Name

Account Rep

Ship to State

Ship to Country

Ship to City

Ship to Street

Day

Year

Week

Quarter

Time

Month

item #

Item description

Item

Category

Brand

Description

Product #

Product

Plant

Package Type

Manufacturing line

Plant No

SKU

Day

Year

Quarter

Week

Month

Date

Returned Date

Expiration Date

# 4: Designing Fact Groups

Design each of the fact groups that were identified in Exercises #1 & #2. Specifically, for each fact group (data mart):

* list the facts that relate to the process that the fact group represents
* write a description for the fact – i.e., define it
* state the default aggregation rule (“sum” if additive; “semi-additive over time” if semi-additive; “N/A” if non-additive)

|  |  |  |
| --- | --- | --- |
| Fact Group: \_\_Item\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | |
| Fact Name | Fact Description | Default Aggregation Rule |
| Returned Quantity | The item returned by the customer | Semi additive over time |
| Damaged Quantity | The item that was damaged | Semi additive over time |
| Expired Quantity | The item that was damaged | Semi additive over time |
|  |  |  |
|  |  |  |

|  |  |  |
| --- | --- | --- |
| Fact Group: \_\_\_\_Sales Forecast\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | |
| Fact Name | Fact Description | Default Aggregation Rule |
| Forecast units | It contains the units for each forecast | Semi additive over time |
| Forecast Amount | Cost of the item | Semi additive over time |
|  |  |  |
|  |  |  |
|  |  |  |

|  |  |  |
| --- | --- | --- |
| Fact Group: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | |
| Fact Name | Fact Description | Default Aggregation Rule |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

# 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 to fit into the Data Warehouse Bus Architecture.

Fill in the data mart matrix using the following table:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Data Mart** | Customer | product | Date | Item returned | Damaged item | Time | Sales rep | Plant |
| Product returns fact | X | X | X | X | X | X |  | X |
| Sales Forecast fact | X |  |  |  |  | X | X |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

# 6: Logical Table Design

Use the dimensional models that you have created so far to:

* Design the actual star schema for each of the fact groups that you defined in Exercise #4.
* Create your Dimensional Models using MySQL Workbench and save it as a pdf file.
* Submit a zip file containing 1) a copy of the answered PE03 and 2) pdf file of EER diagrams to MyCourses PE03 Dropbox by 11:59 PM, Sunday 9/20/2020.
* Bring a hard copy of the answered PE03 & EER pdf to Monday (9/21/20) inperson class.

Name: \_\_\_\_\_Muriel Banze\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Graded By: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# **PE03: Dimensional Modeling Grade sheet**

|  |  |  |
| --- | --- | --- |
| **Requirements** | **Grade** | **Grade Earned** |
| **Exercise 1: Identifying Facts and Dimensions**     * Identify all the attributes correctly       **Exercise 2: Identifying Dimensions and Fact Groups**     * Dimensions * Fact Groups     **Exercise 3: Designing Dimensions**     * Identify all dimensions’ attributes   -   * Identify all hierarchies of the attributes     **Exercise 4: Designing Fact Groups**     * Fact groups * All the facts are included with the fact groups * Fact descriptions and aggregation rules     **Exercise 5: Create the Data Mart Matrix**    **Exercise 6: Logical Table Design**     * Dimensions * Fact tables * Conformed dimension * Correct primary keys and foreign key constraints * Relationships between fact tables and dimensions * Submit star schema to Dropbox & bring a hard copy | 10          6  4        8    5        2  5  5    9        16  8  8  5  9  -20 |  |
| **Total Grade:** | **100** |  |