Gestão de Dados e de Bases de Dados – 2019/2020

Tema:	5.2. Dimensional Modeling
Exercício:	**Record Supermarket Chain**

Adapted from (Hafez & Benchikh, n.d.)

Record is a large supermarket chain having many supermarkets spread over all the country cities. Each supermarket has many products, called stock keeping units (SKUs), on its shelves with a bar code for each of these products. Each supermarket has several departments like: grocery, frozen foods, dairy, meat, bakery, health/beauty, etc. Data is collected at the cash registers as customers purchase products. The supermarket scans the bar codes in the point-of-sale (POS) system. The sales, marketing and logistics departments want to develop data marts (DMs) to support its strategic decisions particularly by understanding how effectively it is running its business. Precisely, the department wants to maximize profit in an extremely competitive market. The profit comes from charging as much as possible for each product, lowering costs for product acquisition and overhead, and at the same time attracting as many customers (who are especially price-sensitive) as possible in a highly competitive pricing environment. The DMs would allow analyzing the trends in the product sales and profits. Besides, some of the most significant management decisions have to do with pricing and promotions. Indeed, the most direct and effective way to create an increase in the volume of product sold is to lower the price. One DM would also allow analyzing promotions, along with sales and profits, in order to develop the appropriate ones. The forms of promotions in the supermarket are:

- Temporary price reductions,
- Ads in TV channels,
- Display in the grocery store (including end-aisle displays), and
- Coupons.

For such, extensive interviews have been conducted with key end users (Directors of Sales, Marketing and Logistics, salespeople, market analysts and shipping and receiving associates) to

determine how they look at the business, and what types of business analysis questions they want to answer. The main first questions were as follows:

- What is the revenue in Record by supermarket in a daily basis? How does this amount of sales in Record differ from the day before?
- Show me the gross profit margin[1] broken out by supermarket, by products and by days?
- What is the COGS amount in Record by supermarkets summarized by product and by day?
- What is the sale quantity for any product in any promotion by day?
- What is the sale quantity for any customer by product by promotion and by day?
- What is the total sales quantity by supermarkets, by product and by day?
- How does sales quantity trending this day versus to prior day?
- What are the most profitable supermarkets for a specific day? What is the percent change in profitability from the prior day?
- What is the average sales amount by transaction, by supermarket and by day?

After re-discussing with the end users, the following issues have been revealed:

a) The above requested statistics should not only have to be on a daily basis but could also be summarized in weeks, months, quarters or years;

b) Business analysts wish to analyze facts at city level to recognize trends, drill-down to supermarket level to identify reasons for these trends and to roll up to the region level to see what these trends have on a larger sector of the business;

c) Users might analyze facts at finer/lower granularity/level, i.e. Individual SKUs. They can then be interested in viewing facts at general levels knowing that individual SKUs make up brands, brands fall into categories, and that categories roll up to departments. The reverse drilling path (i.e. drill-down) could also be considered;

d) Users want to analyze what products were on promotion in a certain supermarket and day but did not sell?

e) Users also want to analyze daily quantity-on-hand inventory levels by supermarket and by product.

f) In addition to the quantity sold, inventory analysts are also interested in value of the inventory at cost, as well as the value at the latest selling price.

Develop a dimensional model to support the answers of the above requirements. You have to precise/identify the following:

1. The Data Warehouse Bus Matrix;

---

[1] Gross Profit Margin= $\frac{Revenue - COGS}{Revenue}$ where COGS is Cost Of Goods Sold

2. For each business process, describe:

    2.1. The grain;

    2.2. The dimensions for facts;

    2.3. The list of dimension attributes;

    2.4. The fact tables and their type;

    2.5. The measures or facts.