Homework2 Report

1000 Sales Records

All the fiels we used are:

1000\_Sale\_Tecords.csv : It save all the data we need.

Homework2.ipynb : It is all the code we write.

Hoemework2Report.docx : It is this work's report.

uml\_model.png : It is uml model we build for the project.

conceptual\_model.png : It is conceptual model we build for the project.

pyhsical\_model.png : It is the pyhsical model we build for the project.

Abstract

We use 1000 sales records data, after audit, building Conceptual Model, UML Model and Physical Model. Then answer all the questions that need to be answered.

Data

Region

Country

Item Type

Sales Channel

Order Priority

Order Date

Order ID

Ship Date

Units Sold

Unit Price

Unit Cost

Total Revenue

Total Cost

Total Profit

Questions

Questions you must answer about your conceptual model:

1. What are the ranges, data types and format of all of the attributes in your entities?

Region VARCHAR(45)

Country VARCHAR(45)

Item Type VARCHAR(45)

Sales Channel VARCHAR(45)

Order Priority VARCHAR(1)

Order Date DATE

Order ID VARCHAR(20)

Ship Date DATE

Units Sold VARCHAR(45)

Unit Price DECIMAL(20,2)

Unit Cost DECIMAL(20,2)

Total Revenue DECIMAL(20,2)

Total Cost DECIMAL(20,2)

Total Profit DECIMAL(20,2)

2. When should you use an entity versus attribute? (Example: address of a person could be modeled as either)

When a object has some attributes and we can use these attributes to describe this object, we use entity to represent it. For exmaple, in our project, Item is an entity, which has Item Type and Sales Channel two attributes.

3. When should you use an entity or relationship, and placement of attributes? (Example: amanager could be modeled as either)

It depends on the purpose of database.

For example, manager could be an entity or attribute, but there do have some differences.

In a manger table, manager is an entity, with its own attributes like name, gender and age. But in an employee table, manager may be a attribute of managers, along with recruiter, engineers. Also, manager can be regarded as the relationship between two entities like managers and departments, that usually means someone is responsible of this department.

4. How did you choose your keys? Which are unique?

For primary key, usually in the same table, we choose the attribute that can identify the item uniquely in a table. For foreign key, we need to find the relationship between two tables and which attribute build the relationship between them and usually, the foreign key in one table is another primary key in another table. For example, in our order table, order ID is unique and in the item table, the attribute item type is unique.

5. Did you model hierarchies using the “ISA” design element? Why or why not?

No, because ISA represent the relationship of inheritance. But in our model, there is no relations that one entity belongs to anther.

6. Were there design alternatives? What are their tradeoffs: entity vs. attribute, entity vs. relationship, binary vs. ternary relationships?

Actually the design is alternative. When we first design the costTable, revenueTable and profiTable, we wanted to combine them together. But after discussion, we think although it is fine to combine them together as entities, but if we separate the whole table into three table, that can better represent the different entity of each table. And each of them use item type as the foreign key, that can better display relationship between them. Thus, finally them forms a ternary relationship.

7. Where are you going find real-world data populate your model?

This model is related to sales orders. So, we can find real-world data Online-shopping websites like Amazon and Taobao.

Questions you must answer about your physical model:

1. Are all the tables in 1NF?

Yes, each table has a primary key(Order ID or Item type). And the values in each column of a table are atomic and no repeating content in these columns.

2. Are all the tables in 2NF?

Yes, first it met all the requirements of 1 NF. Second, all the columns are depend on the primary key.

3. Are all the tables in 3NF?

Yes, first it met all the requirements of 2 NF. Second, there is no transitive dependencies. All columns are directly depend on the primary key.