**Khush Domadiya (fr9739)**

**Lab 2 Part 1: Dimensional Modeling Practice:**

**E1.** UsedCarInc is a used car dealer. When a customer arrive with a car to sell, a dealer agent inspects the car and makes an offer to purchase the car. After purchasing, the car is cleaned and is made available for sale to other customers who come to buy used cars. Any dealer agent can be involved in either the purchase or sale process. Develop a dimensional model to support their main car purchase and sale process. A car purchase by dealer transaction includes information on the seller, dealer agent (employee), car, purchase price, and date. A car sale by dealer includes information on buyer, car, dealer agent (employee), sale price, and date. The company wants to increase profit (sale price minus purchase price) and have higher profit margins, and lower the number of days between car purchase by dealer and car sold by dealer (increase inventory turnover).

1. What is the business process of interest here?

* Car being purchased by dealer
* Car being sold to customer by dealer

1. What are the different levels of granularity present?

* One car purchased by one agent in one day
* One car sold to one customer by an agent on a particular day

1. What are the dimensions for each grain identified in #2?

* Seller dimension, cardimension, date dimension, employee dimension
* Buyer dimension, car, employeer dimension, date dimension, customer dimension, employee dimension

1. What are the facts for each grain in #2 ?

* Purchase price, inspection cost, maintenance cost, clean cost
* Sales price, profit = sales price - purcahse cost - inspection cost - maintenance cost - cleaning cost

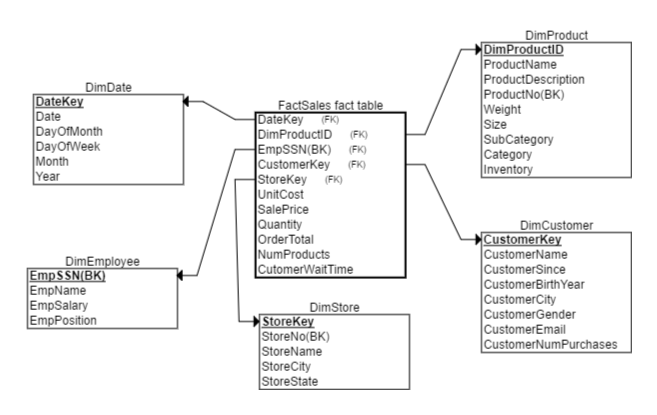
1. Specify if VIN Number and Mileage are included in the model and the most appropriate table where they can be included

* Yes, it can be included in the car dimension

1. Would an accumulating fact table apply in this case?

* Milenstion dates for purchase, inspection, maintenance, sales;
* Some lags can be also be calculated sucha as lags between sale and purchase

**E2.** SunDucks coffee has coffee shops in multiple states. They sell coffee, snacks and related products. You have been provided the following dimensional model for Sales process of SunDucks Coffee, built for the granularity of 1 row per transaction. Identify (1) errors or issues with the model, (2) potential problems due to the errors and (3) propose revisions to correct the errors.



|  |  |
| --- | --- |
| DimProduct | |
| DimProductID | Primary key, sequentially generated |
| ProductName  ProductDescription | Name of product  Description of Product |
| ProductNo | Business Key |
| Weight | Product Weight |
| Size | Product size |
| SubCategory | Product Sub-cateogry |
| Category | Product category |
| Inventory | Number of product in company inventory |
|  |  |
| DimCustomer | |
| CustomerKey | Primary key, sequentially generated |
| CustomerName | Name of Customer |
| CustomerSince | Date of first purchase |
| CustomerBirthYear | Year of Birth of customer |
| CustomerCity | Customer City of residence |
| CustomerGender | Customer gender |
| CustomerEmail | Customer email |
| CustomerNumPurchases | Number of purchases made by customer to date |
|  |  |
| DimStore | |
| StoreKey | Primary key, sequentially generated |
| StoreNo | Business key |
| StoreName | Name of store |
| StoreCity | City in which store is located |
| StoreState | State in which store is located |
|  |  |
| DimEmployee | |
|  | |
| EmpSSN | Employee SSN, Primary Key, Business Key |
| EmpName | Name of Employee |
| EmpSalary | Salary of Employee |
| EmpPosition | Position of Employee |
|  |  |
| FactSales | |
| UnitCost | Unit cost of product sold |
| SalePrice | Price at which product was sold |
| Quantity | Quantity of a product sold ( for e.g. two regular coffees) |
| OrderTotal | Total price of all products sold |
| NumProducts | Total number of products sold in order (for e.g. 3 products coffee, cookies and bottled water) |
| CustomerWaitTime | Total time waited by customer to receive all products in order |

1. Inventory cannot be in product dimension

Issue: where there are any inventory changes, we need to update the product dimension; This type of update will reflect only the latest date, we will lose historical values, as well as dates;

Revison : we can make additional periodical snapshot fact table storing the invenstory balance per day for each product

1. CustomerNumPurchases cannot be in the customer dimension

Revision :

* Transactional fact table contain the quantity purchased for each customer. We can use Bi or analysis tool for aggregating the number of purchases in total till date during the analysis time;
* We can create another peridical snapshot daily for the total amount each customer purchased by them.

1. Employee SSN is not apropriate for the usage as primary key in the table.

Issue : loose sensitive information, difficult to manage or retrieve the employeement information

Revison: ETL system or DW system can automatically generate surrogate keys for employee

1. Fact table has mesurments at different levels of granuality

Issue: Unable to create such a table in practice

Revision:

1. We can keep this table at the transaction level, then we need to remove all the measurements based on the product level
2. We can cover the granularity of the tableto product level.

Order total = total amount paid for all the quantities.

Remove customer wait and number of products measurements