DIS Miniproject

Lasse Vang Gravesen
lgrave11@student.aau.dk

Dennis Jakobsen
djakob11@student.aau.dk

Erik Sidelmann Jensen
ejens11@student.aau.dk

1 DIS Miniproject

1.1 Task B

- Sales

- Payments

- Members

- Location

We pick the 'Sales' business process, because it is the most interesting. The granularity is down to seconds in the existing data, for time we keep up to minutes because seconds is not that important and would take up a significant amount of rows for each day. With regards to members, the granularity goes down to years and there is no more information and that is okay. The data is reasonable for paying customers and detailed enough to ask valuable questions.

Examples of questions:

- How much is bought at some point during some day?

- How does the amount sold change over time?

- Which days are the most busy?

- When is it best to restock, given low activity?

- How much revenue is gained each day, week, month, year?

- Which products have changed the most in price from year to year?

- Which department or member have spent the most?

1.2 Task C

Slowly changing dimensions are not that important.

The business process proposed to be modelled in the data warehouse is the 'Sales' business process. With regards to dimensions we pick out product, time, date, and member because those are the most interesting dimensions. Our granularity with regards to time goes down to minutes across two different dimensions to reduce the amount of rows, with regard to product it only has a name attribute, with regard to member it has balance and year.

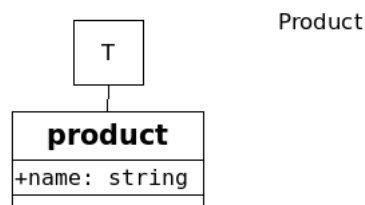The schema for the dimensions can be seen below.
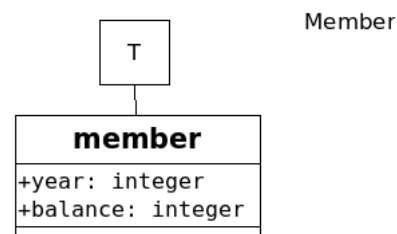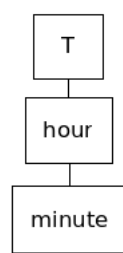

Figure 1.1: The product dimension.

Figure 1.2: The member dimension.

Time

Figure 1.3: The time dimension.

Date

Figure 1.4: The date dimension.

The star schema for it can be seen below. With the schema it is important to note that the dimensions have a surrogate key with a serial integer(it auto increments).
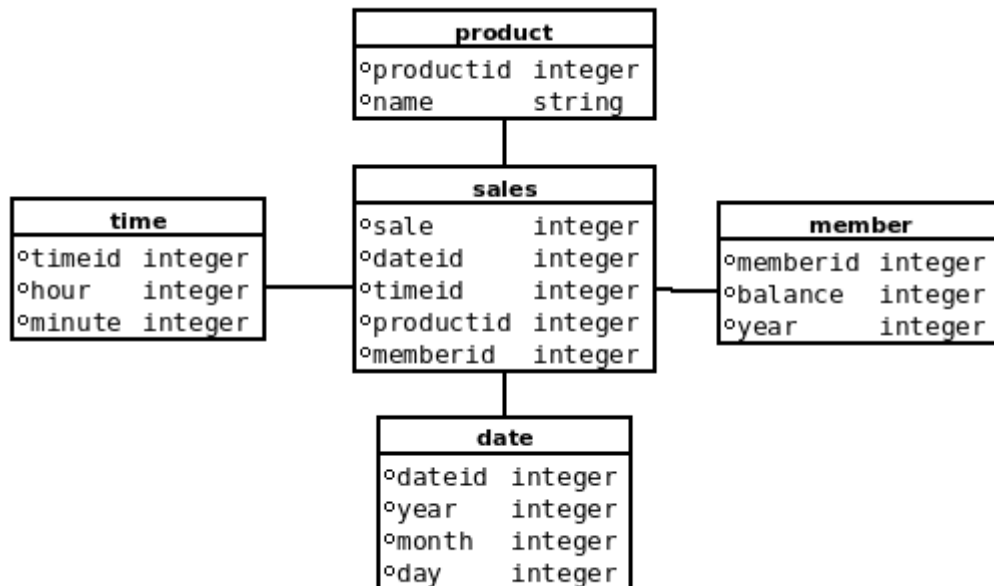

Figure 1.5: The star schema for the data warehouse.