

Miniproject 1 - DIS

Lars Andersen, Mathias Winde Pedersen & Søren Skibsted Als

September 15, 2014

B

By looking at the data, different business processes have been considered. A central one is that of sale. Another process to model is the payment process, which is the pre-insertion of money.

We choose to focus on the sale business process.

For the granularity, the data already contains a timestamp for each sale, as of such we maintain that granularity.

That is timestamp without timezone, containing year, month, day, hour, minute and second. However, a weekday could be relevant as well, and when extracting the data, it could be interesting to include this information, same goes for season, holidays, and week number.

The granularity for a member is year, as we cannot go deeper there, without coming up with some fictive data.

The choices are reasonable for paying customers, which is reflected in the questions that can be asked.

For sales, it could be interesting to see if the sales increase during breaks, and the granularity here is fine with down to minutes, for seconds it is still fine as it does not require much effort to record this as well.

For member creation the year of creation is as close as we can get, thus this is also reasonable for the given data.

Questions that can be asked:

Does sales increase during break times (12:00 to 12:30)?

How has the total sales per year evolved?

Which product has been sold the most?

Which member spends the most money?

Which weekday, weeknumber, season, holiday have the largest revenue?

C

For SCDs products could be useful, as the price changes slowly over time, same goes for members, as they can be active or inactive.

For the time dimension, we skipped weekdays etc. as we would have to find a library of some sort to translate this. The dimensions have been specified in Figure 1. The schema can be seen in the code below and in Figure 2.

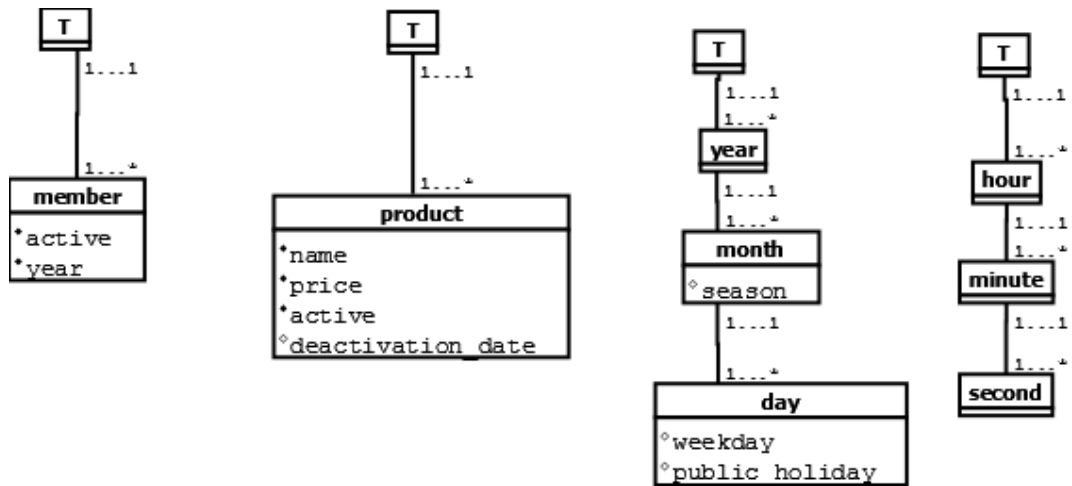


Figure 1: Dimensions, part C.

```

---DROP TABLE sale;
---DROP TABLE product;
---DROP TABLE member;
---DROP TABLE timeofday;
---DROP TABLE date;

```

```

CREATE TABLE "date"
(
  id serial PRIMARY KEY,
  year int NOT NULL,
  month int NOT NULL,
  day int NOT NULL,
  weekday int,
  season int,
  public_holiday int
);

```

```

CREATE TABLE timeofday
(
  id serial PRIMARY KEY,
  hour int NOT NULL,
  minute int NOT NULL,
  second int NOT NULL
);

```

```

CREATE TABLE member
(
  surrogateid serial PRIMARY KEY,
  memberid int NOT NULL,
  active int NOT NULL,

```

```

year int NOT NULL,
version int NOT NULL,
changedate timestamp without time zone — this is for future changes,
—fklub does not keep track of this
);

```

```

CREATE TABLE product
(
surrogateid serial PRIMARY KEY,
id int NOT NULL,
price int NOT NULL,
active int NOT NULL,
deactivation_date timestamp without time zone,
version int NOT NULL
);

```

```

CREATE TABLE sale
(
id serial PRIMARY KEY,
memberid int NOT NULL,
productid int NOT NULL,
dateid int NOT NULL,
timeofdayid int NOT NULL,
FOREIGN KEY(memberid) REFERENCES member(surrogateid),
FOREIGN KEY(productid) REFERENCES product(surrogateid),
FOREIGN KEY(dateid) REFERENCES "date"(id),
FOREIGN KEY(timeofdayid) REFERENCES timeofday(id)
);

```

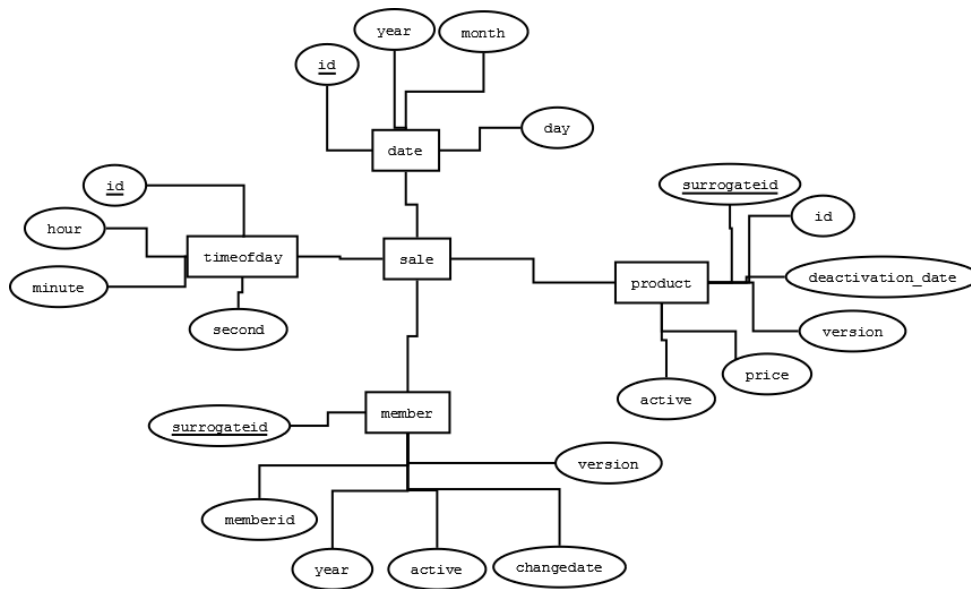


Figure 2: schema, part C.