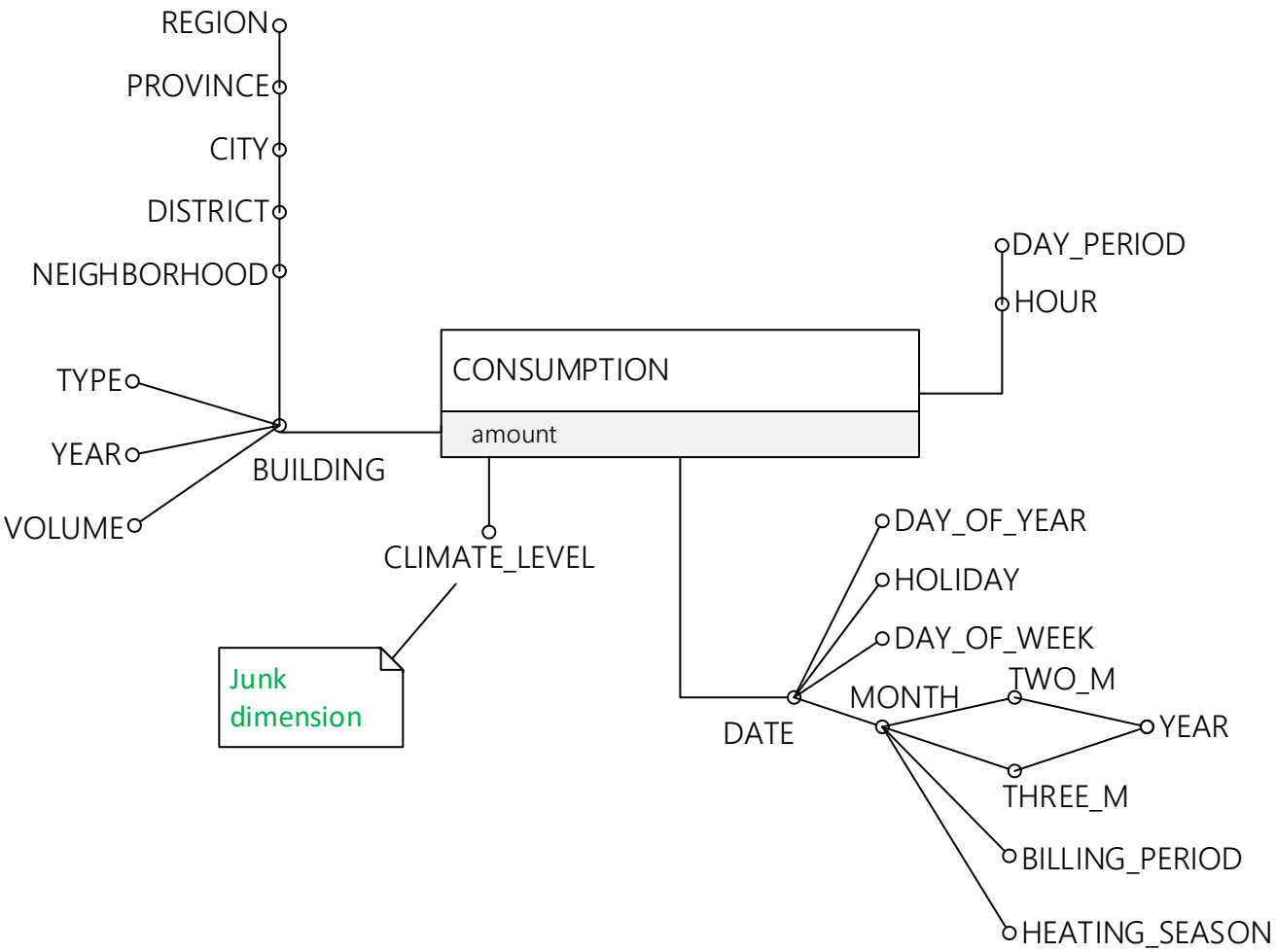

The time dimension is split in two: hour and date. The month determines also the billing period and the heating season (each MONTH is assigned to a single value of BILLING_PERIOD and HEATING_SEASON).
The MONTH, TWO_M and THREE_M attributes also include the year (e.g.: MONTH='September 2015').
Instead the BILLING_PERIOD is one of the following: 'September-October-November', 'December-January', 'February-March', and 'April-May'. The HEATING_SEASON can't be seen as a parent attribute of BILLING_PERIOD because having those values for the attributes, BILLING_PERIOD will have the same values in different HEATING_SEASON. To have a hierarchy between those two attributes, BILLING_PERIOD should include information relative to the HEATING_SEASON, as does the MONTH attribute relatively to the YEAR.
CLIMATE_LEVEL is a junk dimension and is collapsed into the fact table.

Dimension tables:

HOUR(HOUR_ID, HOUR, DAY_PERIOD)
DATE(DATE_ID, DATE, DAY_OF_WEEK, HOLIDAY, DAY_OF_YEAR, MONTH, TWO_M, THREE_M, YEAR, BILLING_PERIOD, HEATING_SEASON)
BUILDING(BUILDING_ID, YEAR, TYPE, VOLUME, NEIGHBORHOOD, DISTRICT, CITY, PROVINCE, REGION)

Fact table:

CONSUMPTION(HOUR_ID, BUILDING_ID, DATE_ID, CLIMATE_LEVEL, AMOUNT)

Queries

Query 1:

```
SELECT BUILDING, CLIMATE_LEVEL, DAY_PERIOD,
  SUM(AMOUNT)/COUNT(*) AS AVG_HOURLY,
  SUM(AMOUNT)/COUNT(*) - (SUM(SUM(AMOUNT)) OVER (PARTITION BY C.BUILDING_ID))/(SUM(COUNT(*)) OVER
(PARTITION BY C.BUILDING_ID)) AS COMPARED

FROM CONSUMPTION C, HOUR H, BUILDING B
WHERE C.HOUR_ID = H.HOUR_ID AND C.BUILDING_ID = B.BUILDING_ID
GROUP BY C.BUILDING_ID, BUILDING, CLIMATE_LEVEL, DAY_PERIOD;
```

The GROUP BY is done including the BUILDING_ID to display correctly the result also in case of duplicate names.
Tuples in the same group have different values for DATE and HOUR. The COUNT counts the number of hour, considering the different dates.
The PARTITION is done according the BUILDING_ID: in the same partition there will be groups with the same BUILDING_ID and different CLIMATE_LEVEL and DAY_PERIOD. The AMOUNT of each group is summed up, and also the number of hours of each group is summed. In order to have the right result, in a HOUR in a DATE there should exist only one DAY_PERIOD (always true) and one CLIMATE_LEVEL, or the HOUR of the DATE will be counted repeated times. This assumption is correct because the CLIMATE_LEVEL is computed every hour.

Query 2:

```
SELECT TYPE, HEATING_SEASON, DAY_OF_WEEK,
  SUM(AMOUNT) / (SUM(SUM(AMOUNT)) OVER (PARTITION BY TYPE, HEATING_SEASON)) * 100 AS PERC_WEEKDAY

FROM CONSUMPTION C, DATES D, BUILDING B
WHERE C.DATE_ID = D.DATE_ID AND C.BUILDING_ID = B.BUILDING_ID
GROUP BY TYPE, HEATING_SEASON, DAY_OF_WEEK;
```

Tuples in the same group have different values for HOUR, DATE and CLIMATE_LEVEL.
The PARTITION puts together groups having the same TYPE and HEATING_SEASON and different DAY_OF_WEEK. The total consumption of each group is summed in the partition to obtain the number in respect to which the consumption of each WEEK_DAY has to be normalized.