Data Warehouse Modeling

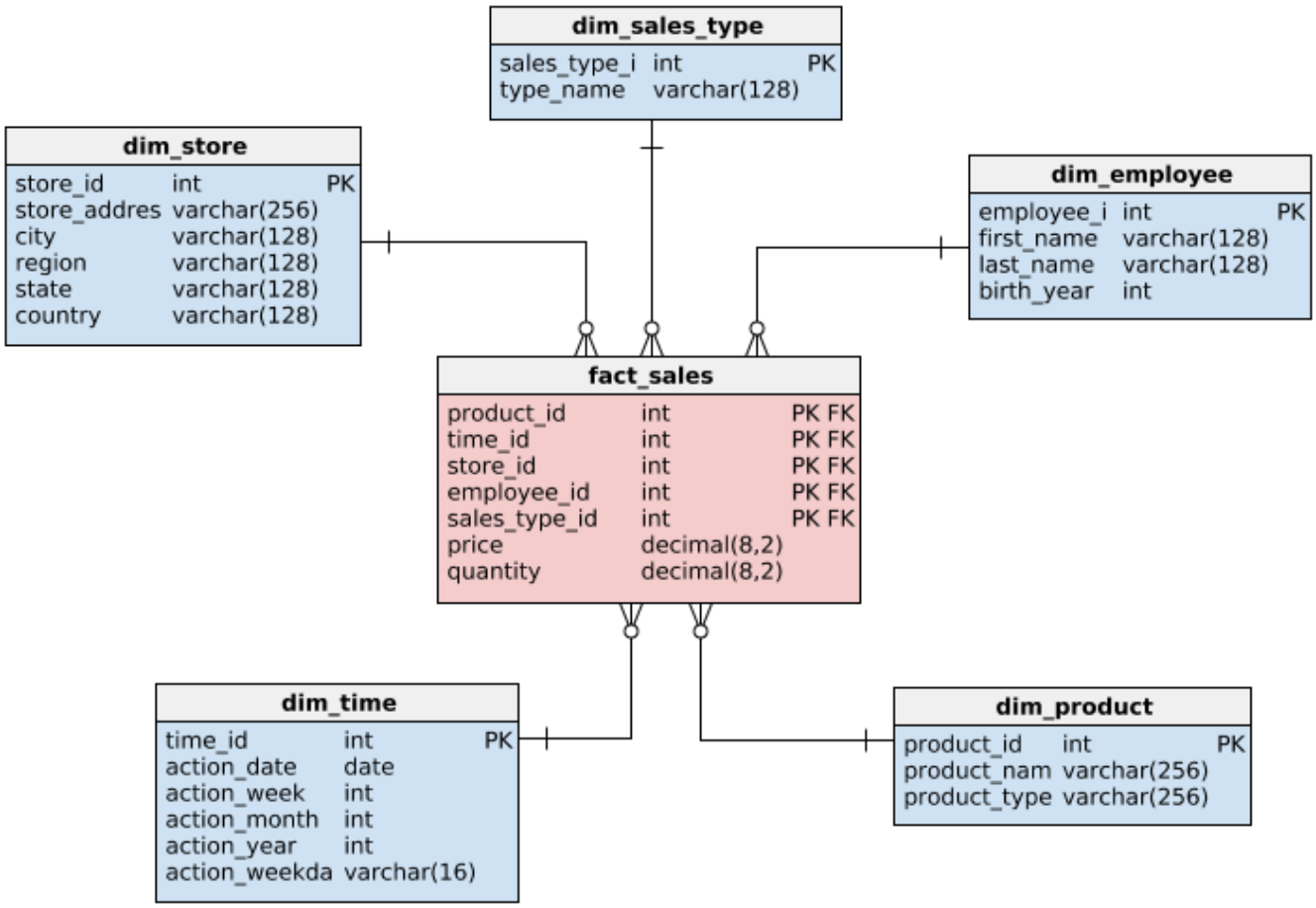
Star Schema vs. Snowflake Schema

The star schema and the snowflake schema are ways to organize data marts or entire data warehouses using relational databases. Both of them use **dimension tables** to describe data aggregated in a **fact table**.

Everyone sells something, be it knowledge, a product, or a service. Storing this information, either in an operational system or in a reporting system, is also a need. So we can expect to find some type of sales model inside the data warehouse of nearly every company.

Let's take one more look at the sales model in both the star and snowflake schemas.

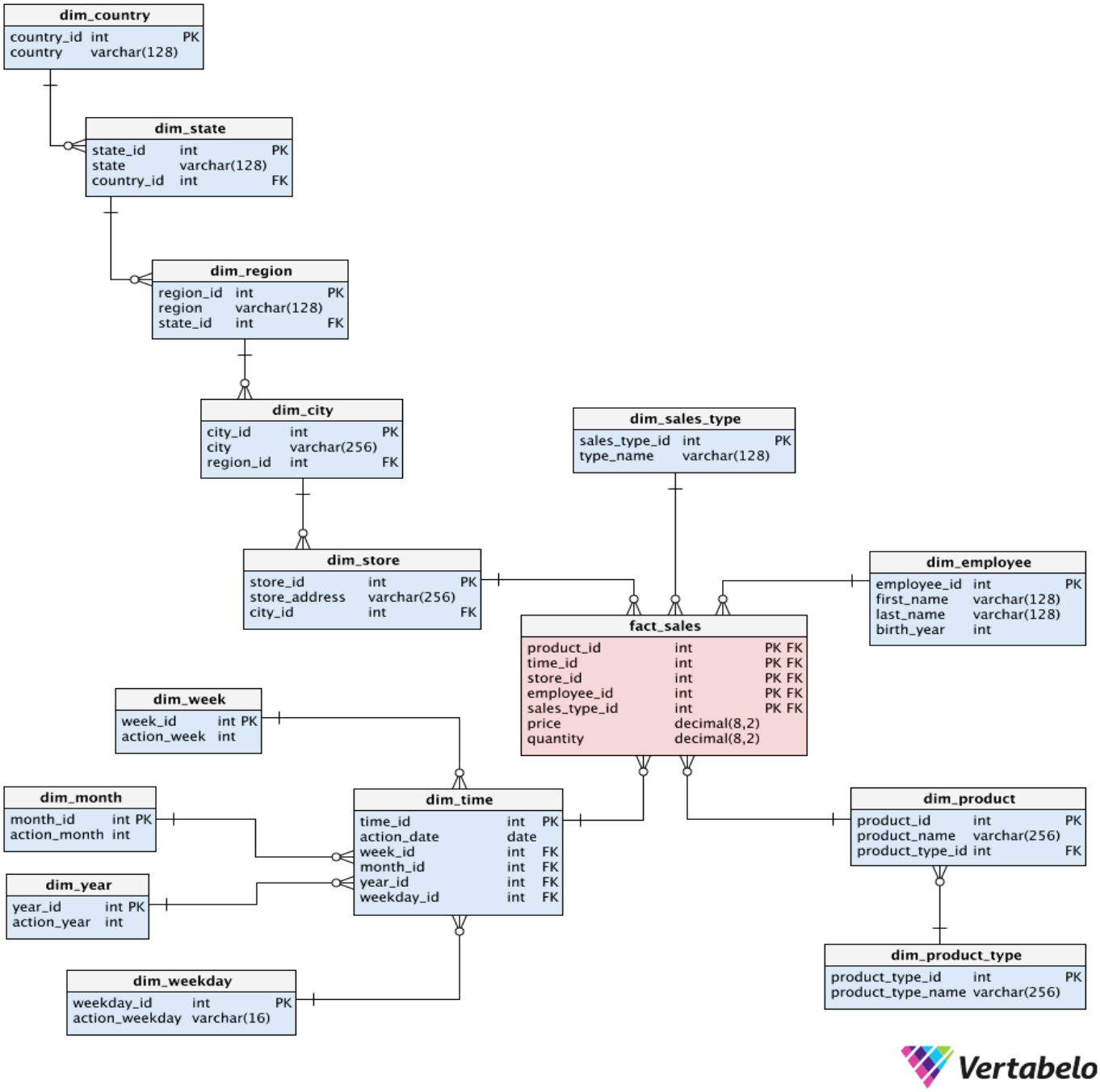
The Star Schema



Vertabelo

The most obvious characteristic of the star schema is that dimension tables are not normalized. In the model above, the pink **fact_sales** table stores aggregated data created from our operational database(s). The light blue tables are dimension tables. We decided to use these five dimensions because we need to create reports using them as parameters. The granulation inside each dimension is also determined by our reporting needs. From this model, we can easily see why this schema is called the 'star schema': it looks like a star, with the dimension tables surrounding the central fact table.

The Snowflake Schema



Vertabelo

This snowflake schema stores exactly the same data as the star schema. The fact table has the same dimensions as it does in the star schema example. The most important difference is that the dimension tables in the snowflake schema are normalized. Interestingly, the process of normalizing dimension tables is called snowflaking.

Once again, visually the snowflake schema reminds us of its namesake, with several layers of dimension tables creating an irregular snowflake-like shape.

Reference: my.vertabelo.com