Takaisin välilehdelle

Tehty: Käy oppitunti läpi loppuun asti

Power BI Desktop and the self-service agile approach greatly improve the success of BI projects and this is due to the flexibility of **Power BI**. Power BI allows you to easily and quickly create meaningful relationships with the different tables that have been imported into your data model.

Click link to watch a video about modeling the data.

Organizing data with a star schema

Self-service BI would not be possible without a **functional data model.** Historically, BI projects focused on building data models could take months and even years to develop when working within the rigid structure and constraints of a corporate business intelligence environment. Unfortunately, studies show that about 50 percent of all enterprise BI projects fail.

Another reason for the low success rate is that the business and end users often won't see any results for many months and can grow frustrated with the lack of visibility in key business areas. These longer project timelines are a result of the time it takes to work through the gathering of requirements, architecting a complex data model, and cleaning up the original data sources.

Power BI Desktop and the self-service agile approach greatly improve the success of BI projects and this is due to the flexibility of **Power BI**. Power BI allows you to easily and quickly create meaningful relationships with the different tables that have been imported into your data model.

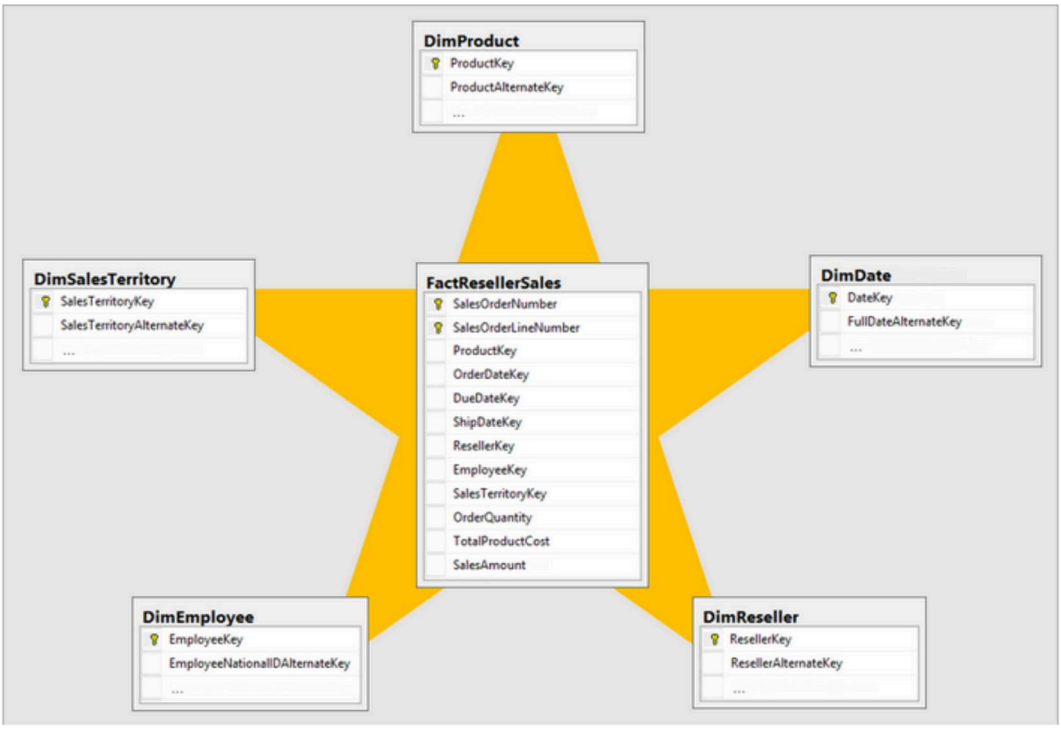
The **star schema** is a way of modeling data that is designed specifically for making it easier to build reports and retrieve analytical value from your data. The star schema method of modeling is part of a much broader topic called **dimensional modeling**. Dimensional modeling has two main types of tables; those tables are **facts** and **dimensions**.

Dimension tables describe business entities—the *things* you model. Entities can include products, people, places, and concepts including time itself. The most consistent table you'll find in a star schema is a date dimension table. A dimension table contains a key column (or columns) that acts as a unique identifier, and descriptive columns.

Fact tables store observations or events, and can be sales orders, stock balances, exchange rates, temperatures, etc. A fact table contains dimension key columns that relate to dimension tables, and numeric measure columns. The dimension key columns determine the *dimensionality* of a fact table, while the dimension key values determine the *granularity* of a fact table. For example, consider a fact table designed to store sale targets that has two dimension key columns **Date** and **ProductKey**. It's easy to understand that the table has two dimensions. The granularity, however, can't be determined without considering the dimension key values. In this example, consider that the values stored in the **Date** column are the first day of each month. In this case, the granularity is at month-product level.

Generally, **dimension tables** contain a relatively small number of rows. **Fact tables**, on the other hand, can contain a very large number of rows and continue to grow over time.

The model that is used throughout this course represents a dimensional model with a fact and multiple dimension tables. When you look at the model view, the tables can appear to be in the shape of a star, hence where the term "star schema" comes from.



Normalization vs. denormalization

To understand some star schema concepts described in this article, it's important to know two terms: **normalization** and **denormalization**.

Normalization is the term used to describe data that's stored in a way that reduces repetitious data. Consider a table of products that has a unique key value column, like the product key, and additional columns describing product characteristics, including product name, category, color, and size. A sales table is considered normalized when it stores only keys, like the product key. In the following image, notice that only the **ProductKey** column records the product.

	SalesOrderNumber	OrderDate	ProductKey	ResellerKey	SalesAmount
1	SO69561	2020-05-31	594	546	226.00
2	SO69560	2020-05-30	513	100	218.45
3	SO69560	2020-05-30	594	100	113.00
4	SO69539	2020-05-28	243	529	858.90
5	SO69539	2020-05-28	378	529	1466.01
6	SO69541	2020-05-28	594	661	113.00
7	SO69542	2020-05-28	243	317	1717.80
8	SO69544	2020-05-28	243	666	3435.60
9	SO69545	2020-05-28	378	436	5864.04
10	SO69532	2020-05-27	594	312	113.00
11	SO69532	2020-05-27	513	312	436.90
12	SO69533	2020-05-27	594	476	226.00

If, however, the sales table stores product details beyond the key, it's considered **denormalized**. In the following image, notice that the **ProductKey** and other product-related columns record the product.

	SalesOrderNumber	OrderDate	ProductKey	Product	Category	Color	Size	ResellerKey	SalesAmount
1	SO69561	2020-05-31	594	Mountain-500 Silver, 48	Bikes	Silver	48	546	226.00
2	SO69560	2020-05-30	513	ML Mountain Frame-W - Silver, 46	Components	Silver	46	100	218.45
3	SO69560	2020-05-30	594	Mountain-500 Silver, 48	Bikes	Silver	48	100	113.00
4	SO69539	2020-05-28	243	HL Road Frame - Red, 44	Components	Red	44	529	858.90
5	SO69539	2020-05-28	378	Road-250 Black, 52	Bikes	Black	52	529	1466.01
6	SO69541	2020-05-28	594	Mountain-500 Silver, 48	Bikes	Silver	48	661	113.00
7	SO69542	2020-05-28	243	HL Road Frame - Red, 44	Components	Red	44	317	1717.80
8	SO69544	2020-05-28	243	HL Road Frame - Red, 44	Components	Red	44	666	3435.60
9	SO69545	2020-05-28	378	Road-250 Black, 52	Bikes	Black	52	436	5864.04
10	SO69532	2020-05-27	594	Mountain-500 Silver, 48	Bikes	Silver	48	312	113.00
11	SO69532	2020-05-27	513	ML Mountain Frame-W - Silver, 46	Components	Silver	46	312	436.90
12	SO69533	2020-05-27	594	Mountain-500 Silver, 48	Bikes	Silver	48	476	226.00

When you source data from an export file or data extract, it's likely that it represents a denormalized set of data. In this case, use **Power Query** to transform and shape the source data into multiple normalized tables.

Building relationships

Olet suorittanut 100 % oppitunnista

100%

◄ Lesson 3 Quiz

Siirry...

Exercise 8 - Building relationships ►