|  |
| --- |
| EPAM Systems, RD Dep. |
| MTN.BI.07 Star Schema Basics |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| REVISION HISTORY | | | | | |
| Ver. | Description of Change | Author | Date | Approved | |
| Name | Effective Date |
| 1.0 | Initial status | [Hanna](mailto:Kiryl_Bucha@epam.com) Klimovich | 8-NOV-2017 |  |  |

Contents

[1. Data Modelling Task 3](#_Toc497924537)

[2. Analytical task 4](#_Toc497924538)

[3. Results 4](#_Toc497924539)

[4. Business Description 5](#_Toc497924540)

[4.1. Business background 5](#_Toc497924541)

[4.2. Problems because of poor data management 5](#_Toc497924542)

[4.3. Benefits from implementing a Data Warehouse 5](#_Toc497924543)

# Data Modelling Task

Here is a typical business model:

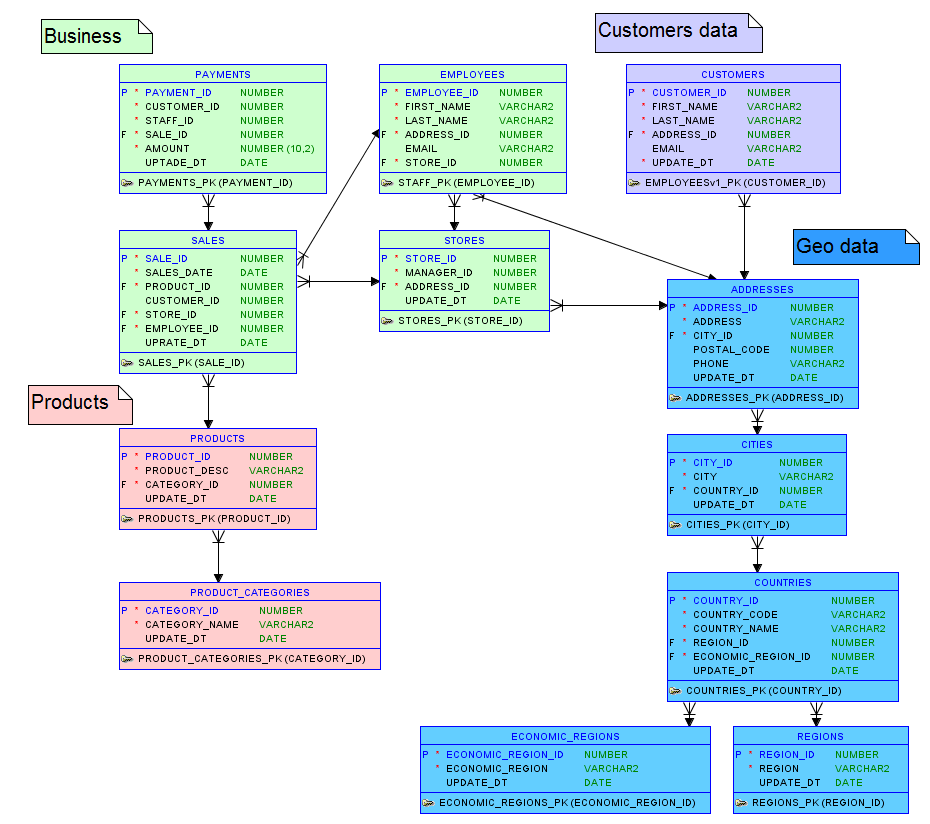


Figure 1 Example

It is in close to 3rd normal form and we can logically divide it in business categories (such as Business, Customer data and Product data).

The task is to build star schema from this model and describe steps of applying 4-step process and denormalizing the model.

The second model that is need to be built is snowflake schema (Hint: to get snowflake apply some normalization rules to Star schema).

**Four-Step Dimensional Design Process**

Step 1: Select the Business Process  
low-level activity: product sales

Business process are supported by sales of different products. Dimensions show detail information on sales. Each store, customer and employee dimension shows geographical data inside.

Step 2: Declare the Grain

One row in a line shows product ordered per customer.

Step 3: Identify the Dimensions

* In StoreDim we can see locations and contacts of the each store.
* In CustomDim each customer location and contact info.
* In EmployeesDim details per each employee is shown (including name and surname, postal code, address and email)
* In DateDim all date information stored
* In ProductionDim shown category and product names.

Step 4: Identify the Facts

Business measures the amount of product sold to customer.

# Analytical task

The main analytical task is to create model of active worldwide business and fill business veins with blood of data after that.

For the next lessons you’ll be creating document describing your chosen business and creating plan of implementing data warehousing for this business.

Your document need to contain next chapters:

1. Overview and some short description of your business. Give some paragraph about purpose and processes in business.
2. The problems that business is facing because of poor data management.
3. What benefit business will get from implementing data warehouse?

# Results

Result of this lab work should be:

* Image of the star schema model.

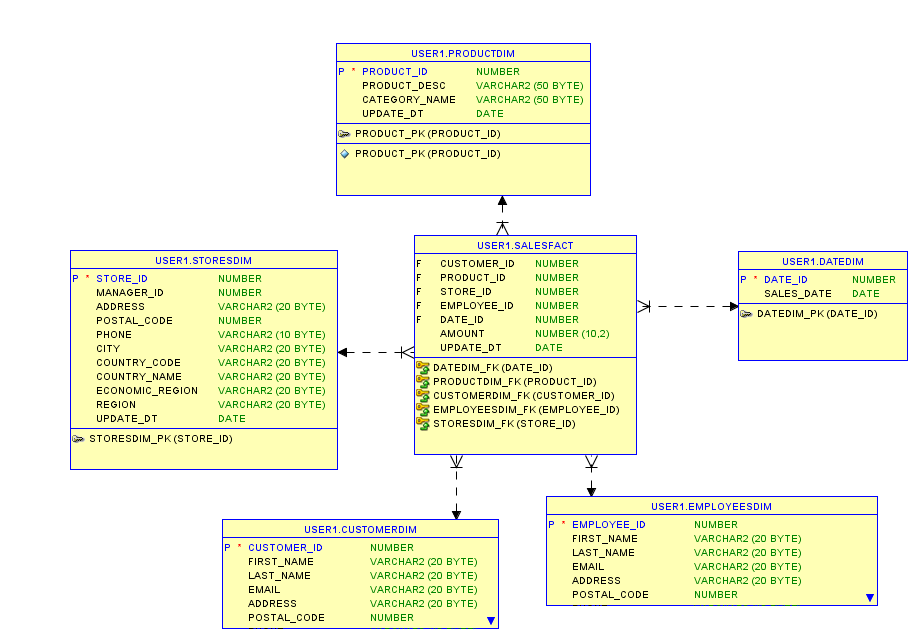


Figure 2 Star schema

* Image of the snowflake schema model.

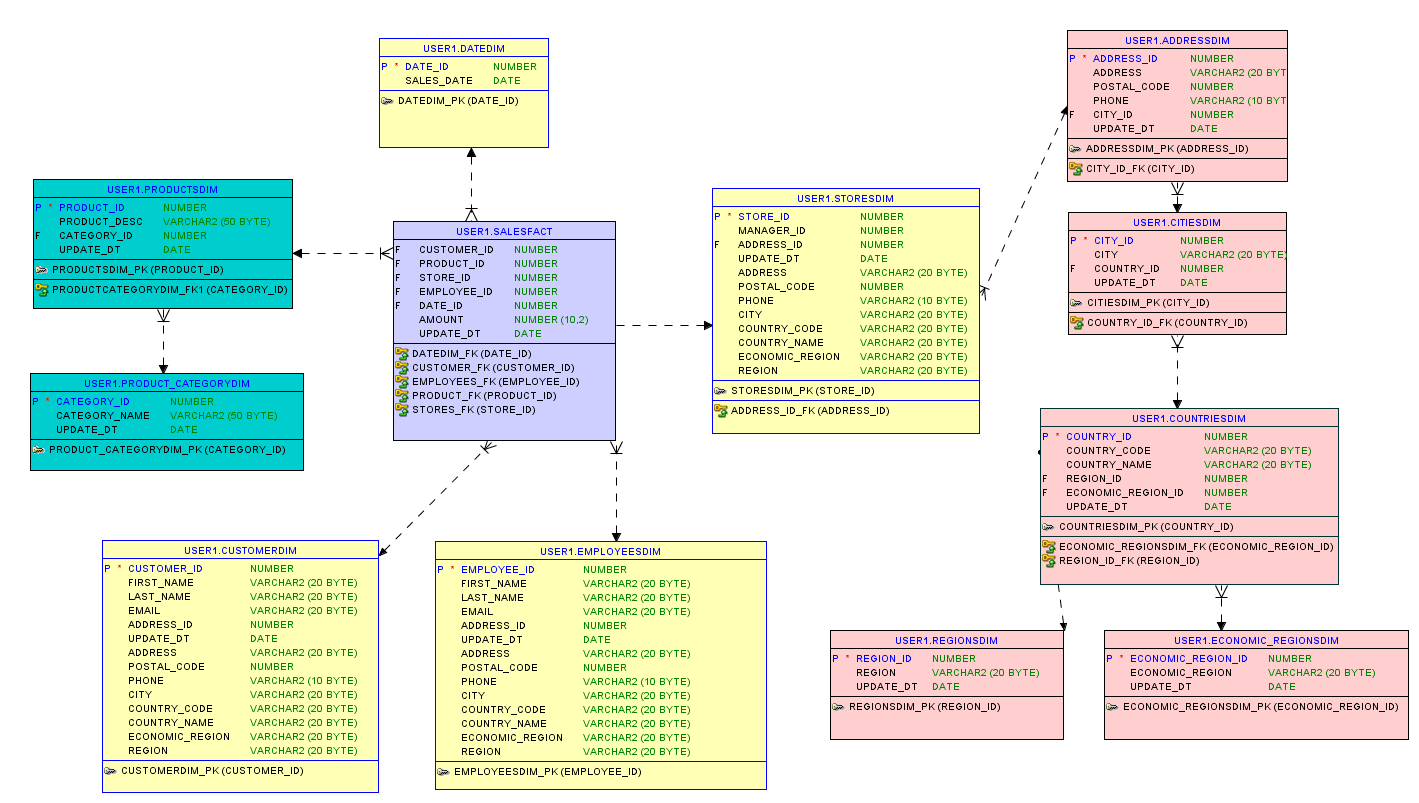


Figure 3 Snowflake

* Document about chosen business.

# Business Description

## Business background

DWH will provide research department of the all airports and companies schedule with flight performance information (departures, arrivals, etc.).

## Problems because of poor data management

Today it’s hard to know/analyze data about flight performance.

## Benefits from implementing a Data Warehouse

As a result the data will have performance statistics that can be analyzed to improve companies and airports schedule and gain more passengers by static time flights.