|  |
| --- |
|  |
| MTN.BI.07 Star Schema Basics Labwork |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| REVISION HISTORY | | | | | |
| Ver. | Description of Change | Author | Date | Approved | |
| Name | Effective Date |
| 1.0 |  | [Vitaliya](mailto:Kiryl_Bucha@epam.com) Adamchuk | 08-NOV-2017 |  |  |

Contents

[1. Data Modelling Task 3](#_Toc497925990)

[2. Analytical task 4](#_Toc497925991)

# Data Modelling Task

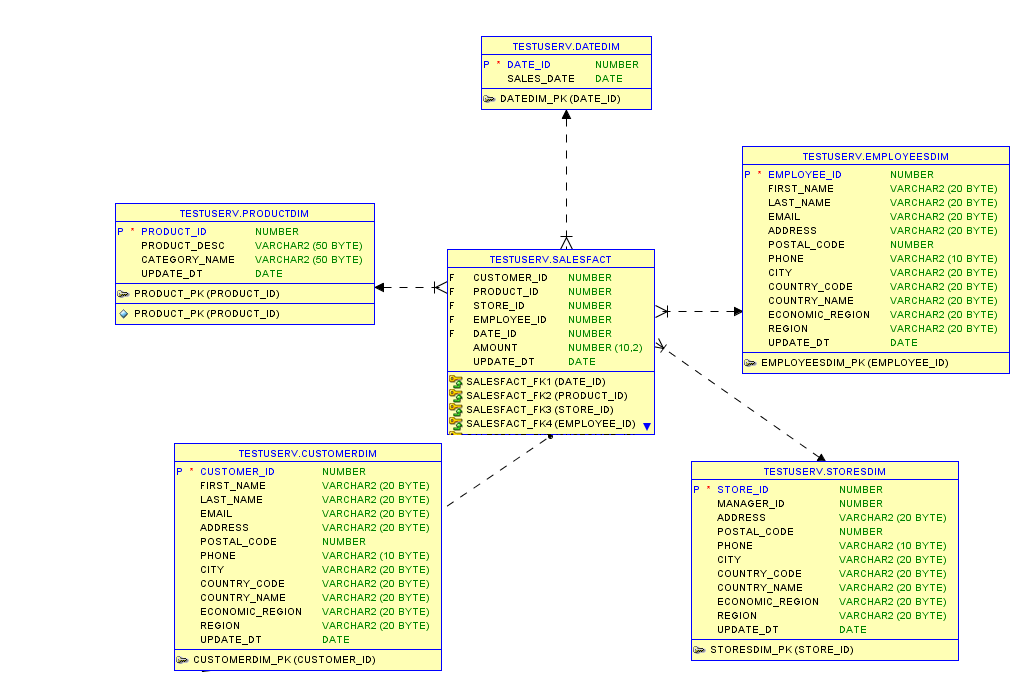


Figure 1 - Star schema

**Steps of denormolazing the model:**

Step 1: Select the Business Process

low level activity : product sales

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 is represented.

- In EmployeesDim details per each employee is shown (including name and surname, postal code, address and email)

- In DateDim information on Sales Date is represented

- In ProductionDim category and product names are representred.

Step 4: Identify the Facts

Fact consists from composition of different ID from dimentions. Fact table represents the fact of bargin.

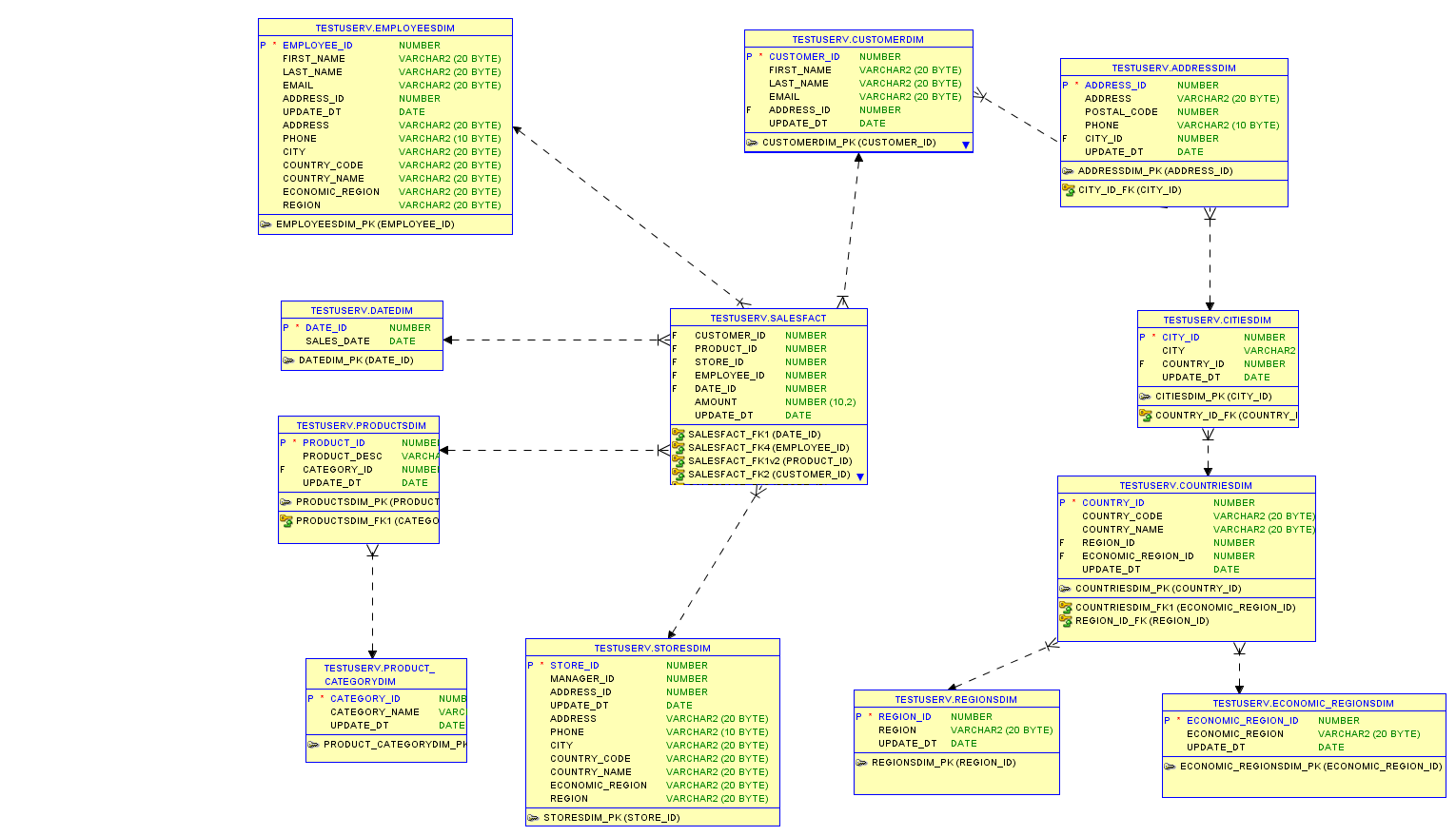


Figure 1 - Snowflake schema

# Analytical task

1. Rossmann operates over 3,000 drug stores in 7 European countries. Rossmann store managers are tasked with predicting their daily sales for up to six weeks in advance.  With a commitment to make health & beauty accessible to all, Rossman maintains the brand’s offering to customers with impeccable retail environments, good assortment of quality health and personal care products and well-trained service staff. The Rossmann name also inspires great confidence from customers as seen in our Own Brands which hold an important position within their product range.
2. Store sales are influenced by many factors, including promotions, competition, school and state holidays, seasonality, and locality. With thousands of individual managers predicting sales based on their unique circumstances, the accuracy of results can be quite varied. So the problem is to predict sales.
3. DWH will be build to enable store managers to create effective staff schedules that increase productivity and motivation. This DWH will help store managers stay focused on what’s most important to them: their customers and their teams.
4. <https://www.kaggle.com/c/rossmann-store-sales/data> - here dataset is represented