|  |
| --- |
|  |
|  |

Image goes here.

Contents

[1. Business Description 3](#_Toc412572569)

[1.1. Business background 3](#_Toc412572570)

[1.2. Problems because of poor data management 3](#_Toc412572571)

[1.3. Benefits from implementing a Data Warehouse 3](#_Toc412572572)

[2. Dimensions of a Business 3](#_Toc412572573)

[3. Logical Scheme 3](#_Toc412572574)

[4. Data Flow 3](#_Toc412572575)

[5. Fact Table Partitioning Strategy 3](#_Toc412572576)

[6. Strategy of Parallel Load 3](#_Toc412572577)

[7. Report Layouts 3](#_Toc412572578)

# Business Description

## Business background

Internet-shop of sports equipment and sports nutrition

## Problems because of poor data management

Loss of customers because of narrow range of products

## Benefits from implementing a Data Warehouse

Increment of activity, sales growth, understand the needs of some equipment and its enrichment.

# Dimensions of a Business

# Dimensional layer of business process

Шаг 1. Выбор бизнес-процесса:

Для постраения хранилища данных был выбран процесс продажи (sport equipment and sport nutrition ).

Шаг 2. Определение зерна

Зерном будет являться продажа товара.

Шаг 3. Определение измерений:

Измерения могут быть следующими:

Customer:

CREATE TABLE DimCustomer

(

CUST\_ID NUMBER(8) NOT NULL PRIMARY KEY,

CUST\_CODE VARCHAR2(8),

LAST\_NAME VARCHAR2(50),

FIRST\_NAME VARCHAR2(50),

MIDDLE\_NAME VARCHAR2(50),

AGE VARCHAR2(3),

GENDER VARCHAR2(50),

ADDRESS VARCHAR2(50),

CITY VARCHAR2(50),

ATE VARCHAR2(50)

);

Supplier:

CREATE TABLE DimSupplier

(

SUP\_ID NUMBER(8) NOT NULL PRIMARY KEY,

SUP\_CODE VARCHAR2(8),

SUP\_NAME VARCHAR2(50),

PHONE NUMBER(8),

E\_MAIL VARCHAR2(50),

ADDRESS VARCHAR2(50),

CITY VARCHAR2(50),

ATE VARCHAR2(50),

CIUNTRY VARCHAR2(50),

START\_DT DATE NOT NULL,

END\_DT DATE NOT NULL,

IS\_ACTIVE VARCHAR2(4) NOT NULL

);

Product:

CREATE TABLE DimProduct

(

PROD\_ID NUMBER(8) NOT NULL PRIMARY KEY,

PROD\_CODE VARCHAR2(8),

PROD\_NAME VARCHAR2(50),

CATEG\_NAME VARCHAR2(25),

SUBCAT\_NAME VARCHAR2(25),

PROD\_PRICE NUMBER(8),

PROD\_DESCRIPTION VARCHAR2(50),

START\_DT DATE NOT NULL,

END\_DT DATE NOT NULL,

IS\_ACTIVE VARCHAR2(4) NOT NULL

);

Date:

CREATE TABLE DimDate

(

"DATE\_ID" DATE NOT NULL PRIMARY KEY,

"DA\_OF\_WEEK" NUMBER(2),

"DAY\_NAME\_OF\_WEEK" VARCHAR2(25),

"DAY\_OF\_MONTH" NUMBER(8),

"DAY\_OF\_YEAR" NUMBER(8),

"WEEK\_OF\_MONTH" NUMBER(8),

"WEEK\_OF\_YEAR" NUMBER(8),

"MONTH\_NUMBER" NUMBER(8),

"MONTH\_NAME" VARCHAR2(20),

"QUARTER" NUMBER(2),

"FIRST\_DAY\_OF\_MONTH" DATE,

"LAST\_DAY\_OF\_MONTH" DATE,

"YEAR" NUMBER(8)

);

Payment:

CREATE TABLE DimPayment

(

PAYMENT\_ID NUMBER(8) NOT NULL PRIMARY KEY,

PAYMENT\_CODE VARCHAR2(8),

PAYMENT\_TYPE VARCHAR2(4)

);

Шаг 4. Определение фактов:

Фактами будут являться – quantity, price\_sum.

CREATE TABLE FactOrderSales

( ORDERSALES\_ID NUMBER(8) PRIMARY KEY,

DATE\_ID DATE,

PROD\_ID NUMBER(8) NOT NULL,

SUP\_ID NUMBER(8) NOT NULL,

PAYMENT\_ID NUMBER(8) NOT NULL,

CUST\_ID NUMBER(8) NOT NULL,

QUANTITY NUMBER,

PRICE\_SUM NUMBER

);

# Logical Scheme

By adding following constraints we connect our existing tables in order to have star-scheme:

ALTER TABLE FactOrderSales

ADD CONSTRAINT "FK\_DATE\_ID\_Fact"

FOREIGN KEY ("DATE\_ID")

REFERENCES DimDate("DATE\_ID");

ALTER TABLE FactOrderSales

ADD CONSTRAINT "FK\_PROD\_ID\_Fact"

FOREIGN KEY ("PROD\_ID")

REFERENCES DimProduct("PROD\_ID");

ALTER TABLE FactOrderSales

ADD CONSTRAINT "FK\_PAYMENT\_ID\_Fact"

FOREIGN KEY ("PAYMENT\_ID")

REFERENCES DimPayment("PAYMENT\_ID");

ALTER TABLE FactOrderSales

ADD CONSTRAINT "FK\_SUP\_ID\_Fact"

FOREIGN KEY ("SUP\_ID")

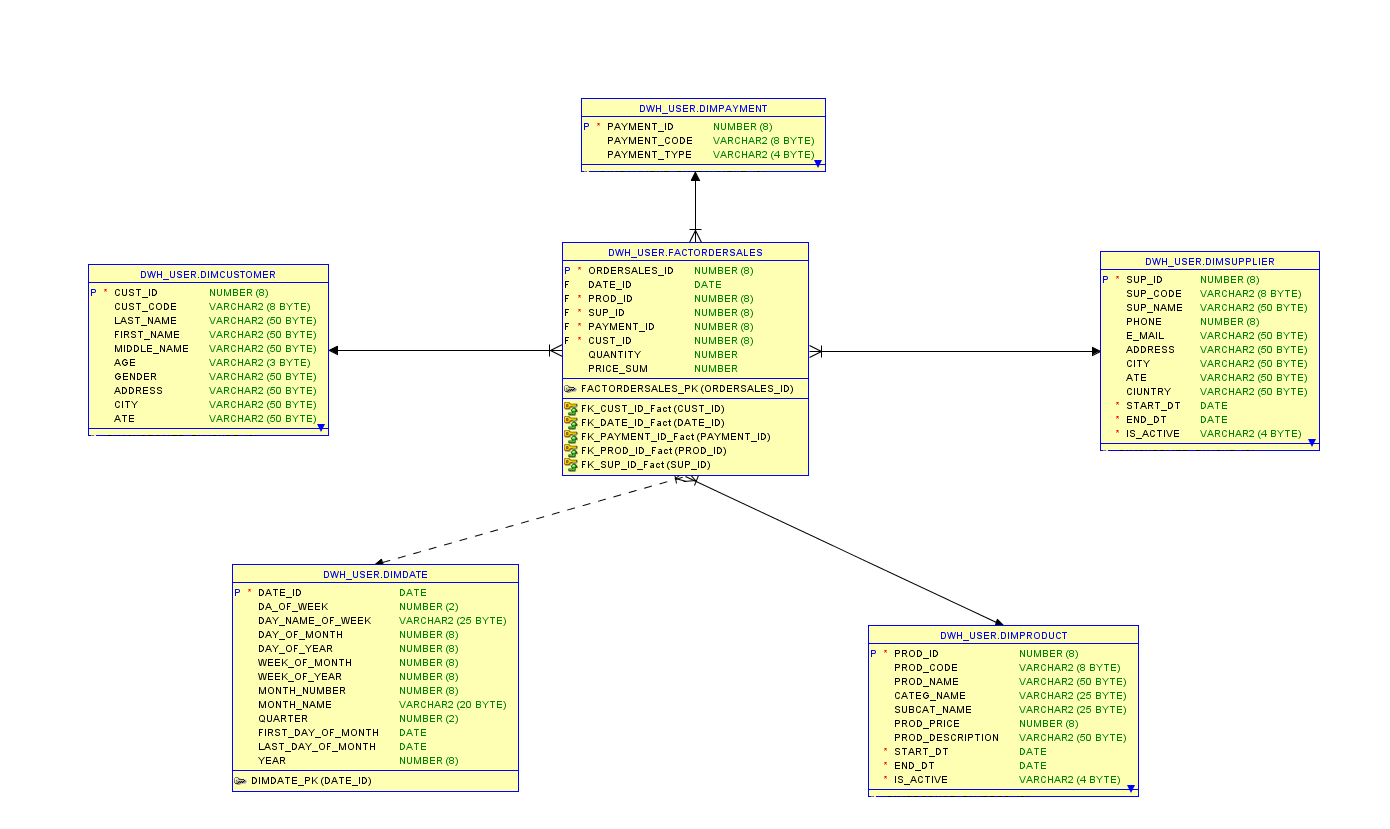
REFERENCES DimSupplier("SUP\_ID");

ALTER TABLE FactOrderSales

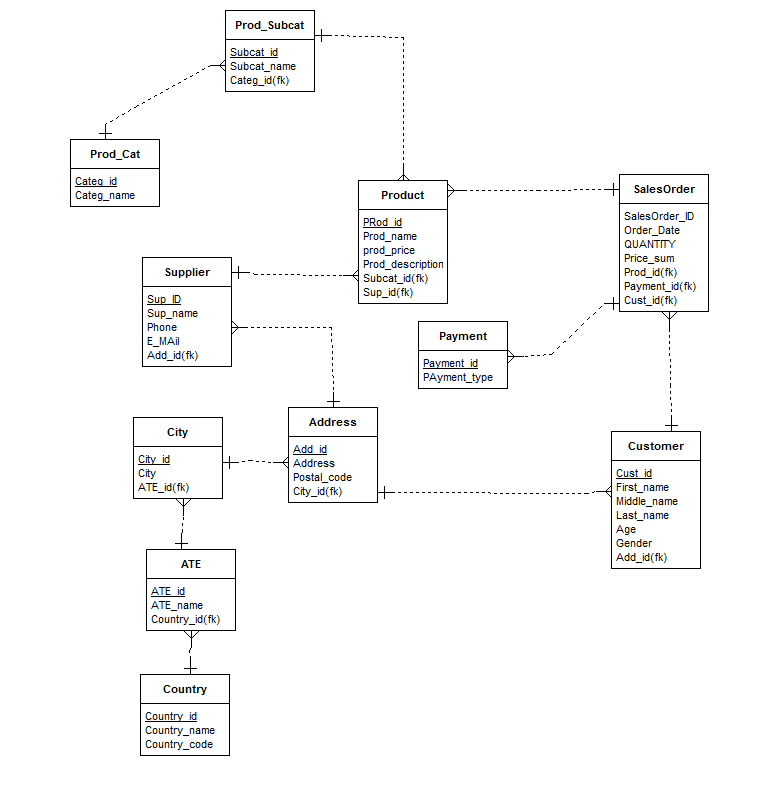
ADD CONSTRAINT "FK\_CUST\_ID\_Fact"

FOREIGN KEY ("CUST\_ID")

REFERENCES DimCustomer("CUST\_ID");



3NF LAYER:



# Data Flow

# Fact Table Partitioning Strategy

It would be helpful to create partitions on table fact\_SalesOrder in order to analyze data about sales from internet-shop. At first partitioning would be organized by date of sale (year). After that creating a subpartition by product category and subcategory can be done.

# Strategy of Parallel Load

# Report Layouts

I would analyze month and year sales of product (also it’s category and subcategory) in slice of customer and geodate(city) .

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| year | month | category | amount\_sold | sales |
| 2017 | jan | powerlifting equipment | 1651 | 1263100 |
| 2017 | jan | nutrition | 2653 | 4196100 |
| 2017 | jan | sport outfit | 1501 | 49564060 |
| 2017 | feb | powerlifting equipment | 156 | 46018400 |
| 2017 | feb | nutrition | 5161 | 465040 |
| 2017 | feb | sport outfit | 15610 | 46546497 |
| 2017 | march | powerlifting equipment | 268 | 4645640 |
| 2017 | march | nutrition | 516 | 46064 |
| 2017 | march | sport outfit | 1623 | 465498 |