|  |
| --- |
| EPAM Systems, RD Dep. |
| Analysis of Amazone.com Sales |



Contents

[1. Business Description 3](#_Toc498071105)

[1.1. Business background 3](#_Toc498071106)

[1.2. Problems because of poor data management 3](#_Toc498071107)

[1.3. Benefits from implementing a Data Warehouse 3](#_Toc498071108)

[2. Dimensions of a Business 3](#_Toc498071109)

[2.1. Select the Business Process 3](#_Toc498071110)

[2.2. Declare the Grain 3](#_Toc498071111)

[2.3. Identify the Dimensions 3](#_Toc498071112)

[2.4. Identify the Facts 3](#_Toc498071113)

[3. Logical Scheme 3](#_Toc498071114)

[4. Data Flow 3](#_Toc498071115)

[5. Fact Table Partitioning Strategy 3](#_Toc498071116)

[6. Strategy of Parallel Load 3](#_Toc498071117)

[7. Report Layouts 3](#_Toc498071118)

# Business Description

## Business background

Amazon (Amazon.com) is the world’s largest online retailer. Amazon offers its services all over the world. It operates its own websites in the UK, Germany, China, Australia, as well as many South American and Asian countries. Amazon.com serves 20 million customers every day. Its annual income is over 34 billion dollars. Amazon is a business empire that sells almost anything cheaper and faster than anyone else.

## Problems because of poor data management

It may be occurred such problems as:

* Technical data not recorded properly;
* Data not stored properly;
* Data Sharing Difficulties.

## Benefits from implementing a Data Warehouse

DWH structure is going to meet the following benefits:

* Quick and easy access to data;
* Improving the performance;
* Consistency of data;
* The opportunity to forecast;
* The opportunity for marketing campaign.

# Dimensions of a Business

## Select the Business Process

Business Process for analysis is Amazon’s sales per different metrics.

## Declare the Grain

Each row in the table shows product that have been ordered by customer.

## Identify the Dimensions

* Dim\_Products
* Dim\_Product\_Type
* Dim\_Product\_Score
* Dim\_Manufacturer
* Dim\_Categories
* Dim\_Subcategories
* Dim\_Customers
* Dim\_Customer\_Location
* Dim\_Sellers
* Dim\_Date
* Dim\_PaymentMethod
* Dim\_DeliveryMethod

## Identify the Facts

Fact\_Orders:

* Product\_id
* Customer\_id
* Seller\_id
* PaymentMethod\_id
* Amount
* Order\_date\_id
* Delivery\_date\_id
* DeliveryMethod\_id
* PaymentMethod\_id
* Flag(Delivered or not)
* Update\_Date

# Logical Scheme

# Data Flow

# Fact Table Partitioning Strategy

# Strategy of Parallel Load

# Report Layouts