

**Data Modeling**

**2018/2019**

**Adventure Works Cycles  
OLTP to OLAP model**https://docs.google.com/drawings/d/sc2cNfYB47Oppg6M5fTXd8w/image?w=557&h=1&rev=1&ac=1&parent=1SWU-hxAxzsNmfPMsh6xgnfdnNGvATs9nRW_aLsqCa2Q

Subject proposal

**Group members: Group ID: 12**

50932, João Gonçalo Vaz Oliveira

54842, Ondřej Zemánek

54888, Hubert Kamil Olkiewicz

November 19, 2018

1. Abstract

This project is focused on creating a multidimensional model from existing relational database for analytical purposes. It describes business model of a fictional company called Adventure Works Cycles. This company is a multinational bicycle manufacturer and it stores the sales data, customers data, and much more. Next part is about finding the best analytical tasks, which suits the business model most. Based on this analytical task is built model (schema) from existing relational model. In last part project is focused on implementation of this model, migrating data from relational database and comparing speed of defined analytical tasks in OLTP and OLAP.

1. Detailed description
2. Definition of business model

In our project we figure out what is this company doing through the prism of the database: connections between sales from more general and distinct product perspective, discounts, people and territories or countries.

1. Finding the most suitable analytical tasks

In this part we will focus on how to correctly define analytical tasks for our business model. We will use the business model definition for definition of analytical tasks, that will be really interesting and usable for real user of this system. We will define (at least) five analytical tasks:

1. Relation between time (example: day of week) and most sold products/subcategory products.
2. What kind of products are customers buying when purchasing a bike.
3. How does the season influence sales of each product\_category and what is the best “special\_offer” for each season.
4. How much can “special\_offer” and its discount amount improve sales of certain product, which were rarely sold before.
5. Do more experienced salespersons selling more products in winter season.
6. Building multidimensional model

In this step we will create OLAP model from OLTP E/R model with set of rules:

* 1. In memory model relational model
* Representation of our relational database model as ER diagram
  1. Classification of entities
* Classification entities of OLTP database into three categories:

1. Transaction Entities
2. Component Entities
3. Classification Entities
   1. Identifying hierarchies

* Identifying Hierarchy of sequences of entities joined by one-to-many relationships.
  1. Building the model

1. Collapse Hierarchy
2. Aggregation
3. Choosing the most suitable data model
4. Evaluation and model tuning
5. Implementation of the schema

Implementation of the schema from the built OLAP model.

1. Populating database

Migrating data from original relational database into OLAP database.

1. Data

We are using a dataset containing the sales data for fictitious, multinational bicycle manufacturer called Adventure Works Cycles created by Microsoft. This dataset is for use with Microsoft SQL Server database.

Data model consists of 71 tables with 486 columns in total. The dataset contains more than 750 000 rows. It contains information about product sales, customers, shops, employees, product and much more. [1]

1. Scope

Business model definition will be based only on information obtained from database model. Also comparison is focused on performance of the models in certain analytical tasks, that will be described in second part of our project. We will also try to find dependencies between transaction, component and classification entities.

1. Teamwork

Hubert Kamil Olkiewicz - Detailed description and project analysis. Finding out dependencies between data. Building OLAP model.

João Gonçalo Vaz Oliveira - Taking part on gathering analytical tasks (studying the business) in order to choose five interesting analytical tasks of value for the final user of the system. Building OLAP model and task performance comparison both in OLTP and OLAP.

Ondřej Zemánek - I will took a part in analytical tasks, building OLAP model, migrating data and performance comparison.

1. References

[1] Release AdventureWorks sample databases · Microsoft/sql-server-samples · GitHub. The world’s leading software development platform · GitHub [online]. Copyright © 2018 [cit. 19.11.2018]. Dostupné z: <https://github.com/Microsoft/sql-server-samples/releases/tag/adventureworks>