

## Gedächtnisprotokoll 1. Teilklausur DW & BI (04.12.18)

35 Points in total

### Aufgabe 1 (5 Points)

10 MC question à 0,5 points

### Aufgabe 2 (7 Points)

#### Aufgabe 2.1 (3 Points)

Assign the followings terms to the different parts of a DW architecture (Data Sources, Backend Tier, DW Tier, OLAP Tier, Frontend Tier)

- Internal sources
- Reporting and Statistical Tools
- OLAP server
- External sources
- Operational Databases
- Data Marts
- ETL process
- Enterprise DW
- Staging Area
- Data Mining Tools
- Metadata

#### Aufgabe 2.2 (4 Points)

Describe the following design approaches, mention one problem and explain it.

- a) Bottom-Up
- b) Top-Down

## Aufgabe 3 (7 Points)

### Aufgabe 3.1 (4 Points)

The following table had empty cells which we had to fill out. 0,25 points for each right cell.

Criteria	OLTP Database	OLAP Data warehouse
Transaction operations	Insert, select, update, delete	Select
Transaction style	Predefined: predictable, stable	Ad-hoc: unpredictable, volatile
Optimized for	Update efficiency and write consistency	Query performance and usability
Update frequency	Real time when business event occur	Periodic (daily) via scheduled ETL, moving to near real time
Update concurrency	High	Low
Historical data access	Current and recent periods	Current + several years of history
Selection criteria	Precise, narrow	Fuzzy, broad
Comparisons	Infrequent	Frequent
Query complexity	Low	High
Tables/joins per transaction	Few (1-3)	Many (10+)
Rows per transaction	Tens	Millions
Transactions per day	Millions	Thousands
Data volumes	Gigabytes-Terabytes	Terabytes-Petabytes (many sources, history)
Data	Mainly raw detailed data	Detailed data, summarized data, derived data
Design technique	Entity-Relationship modeling (normalization)	Dimensional modeling
Data model diagram	ER diagram	Star schema

### Aufgabe 3.2 (3 points)

Explain the abbreviation ETL and the tasks included in every step.

E:

T:

L:

## Aufgabe 4 (6 Points)

### Aufgabe 4.1 (2 Points)

What are virtual views compared to materialized views? When can it be make sense to use a virtual view instead of a materialized view?

### Aufgabe 4.2

Given three tables. Find the solution by using bitmaps. (where gender = 'female' AND month = 'december')

## Aufgabe 5 (10 Points)

Case Study "Linkbook": Create a snowflake schema to answer the questions mentioned in the descriptions by following this guide:

5.1: List every dimension with their hierarchie. Mark Primary and Foreign Keys

5.2: Choose the measures which are needed to answer the questions. Write down the measures with the measure type.

5.3: Write down example dimension values for 6 dimension levels (separate tables)

## Aufgabe 6 (Bonus: 3 Points)

Given four attribute changes. Handle the changes of gender and birth date using type 1 updates, handle changes for credit rating and residence using type 2 updates. Write down the table after each of the four changes.