Assignment 1

TDT4300

Data Warehouse and Data Mining

Marte Løge and Linn Vikre

Exercise 1

1. **OLTP:** Stands for 'Online transaction processing' and are a series of short online transactions. OLTP is mainly used when you want to achieve fast query processing.
   OLAP: Stands for 'Online analytical processing' and has a low volume of transactions but more complex queries which means that it requires fewer queries to do the same operation.

2. A **datacube** is an multidimensional array which have three or more dimensions and contains multiple set of data.
   A **cuboid** is a subcube or a cube formed set of data. It might refere to a set of data and a result from a query.

3. Operations you can do on a cube is slice, dice, rollup and drill-down.
   Slice: Here you select a part of the cube, often removing a side of the cube, too analyse it.
   Dice: Here you split the datacube into multiple smaller datacubes.
   Rollup: In this operation you put together slices of data to create a new cube.
   Drill-down: This operation takes out a portion in the cubes middle part, often a dice or just a single value.

Exercise 2a)

Make a star or snowflake schema for the case description:

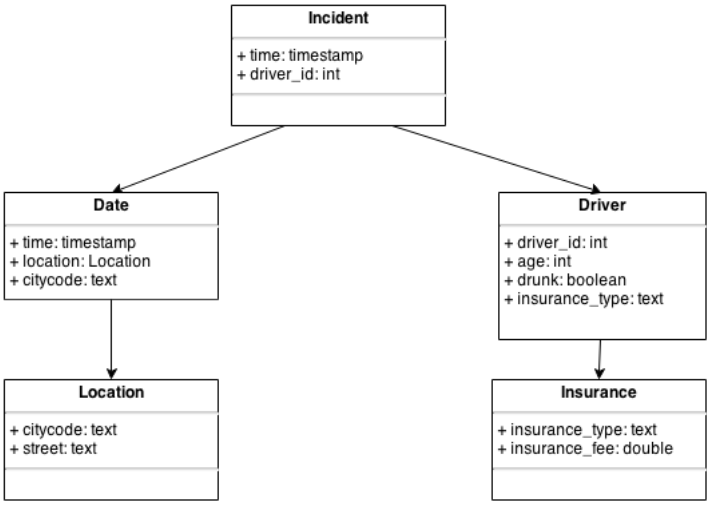

Figure 1: Star diagram

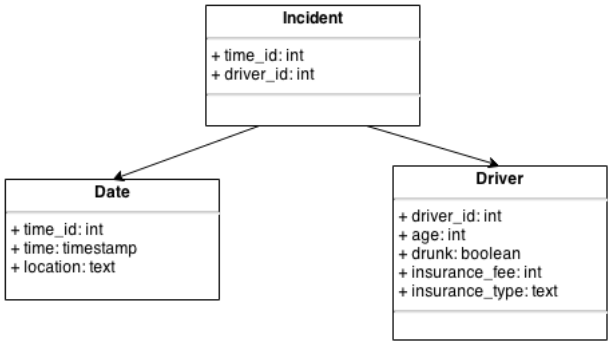

Figure 2: Snowflake diagram

Exercise 2b)

Define two different concept hierarchies:

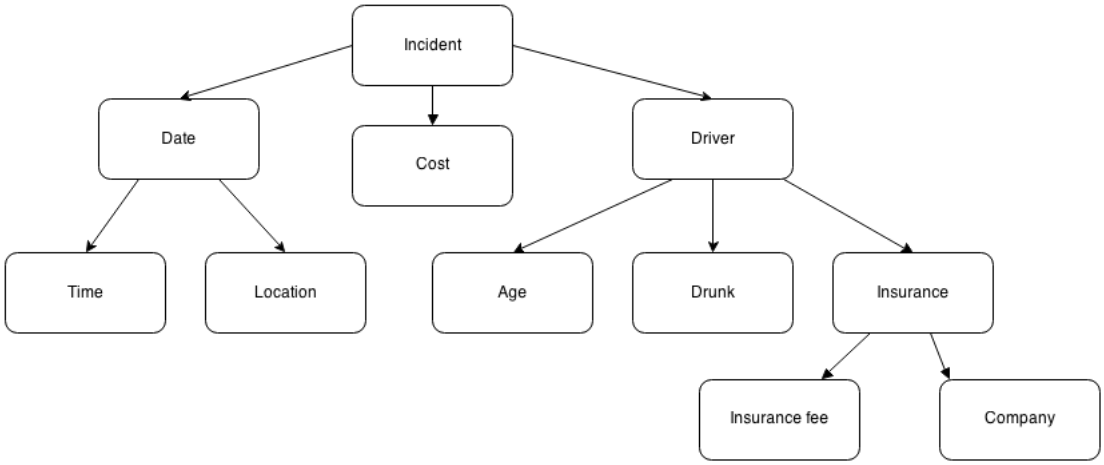

Figure 3: Concept hierarchy 1

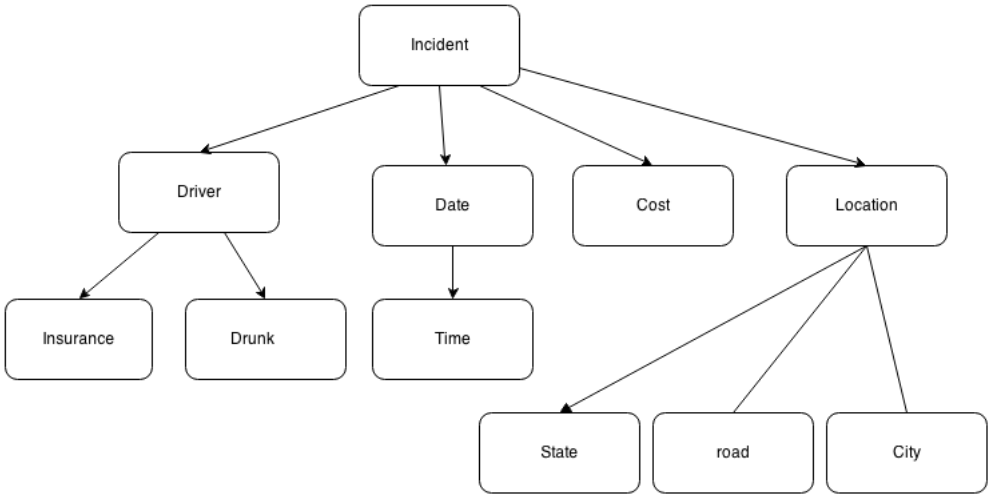

Figure 4: Concept hierarchy 2

Exercise 3

In this exercise we got a lot of problems with the setup of the iccube program. We did manage to get some results, but the queries are quite simple.

Incidents for Location/Driver

```
SELECT [Driver].[Driver] ON 0,
[Location].[Location] ON 1
FROM Car
```

	1.0	3.0	5.0	7.0	10.0	11.0	24.0	25.0	33.0
4845.0	1.0								
8556.0		3.0							
4878.0			5.0						
5587.0				7.0					
1122.0					10.0	11.0			
4588.0							24.0		
4154.0								25.0	
1258.0									33.0
4587.0									
3328.0									
1558.0									

Incidents for each driver with postcode and street

```
select
[Location].[LocationH].[Location] on 0,
non empty [Driver].[DriverH].[Driver].members on 1
from [Car]
```

	1122.0	Gate A	1123.0	Gate B	1127.0	1258.0	1548.0	1558.0
1.0	1							
3.0			1					
5.0					1			
7.0						1		
10.0							1	
11.0								1
24.0								
25.0								
33.0								
64.0								
66.0								
74.0								
78.0								
85.0								

Distinct incidents for each driver

```
select [Measures].[DistinctCountDriver] on 0,
non empty [Driver].[DriverH].[Driver].members on 1
from [Car]
```

	DistinctCountDriver
1.0	1
3.0	1
5.0	1
7.0	1
10.0	1
11.0	1
24.0	1
25.0	1
33.0	1
64.0	1
66.0	1
74.0	1
78.0	1
85.0	1
86.0	1
88.0	1