

DATA WAREHOUSING

tdt4300-undass@idi.ntnu.no

Spring 2023

Problem

Download and unzip the file *data_bike.zip*. Inside are files *stations.csv* and *trips.csv* that show data about Austin's bike share trips in the years 2013-2017. Make yourself familiar with the dataset. Your task is to provide following reports:

Make yourself familiar with the datasets. Your task is to provide the following reports:

Report 1 Trip counts in each year.

Report 2 The total duration each bike was ridden in 2014.

Report 3 The trip count, revenue and average trip duration throughout the months from 2013 to 2015. Can you observe some patterns?

Report 4 Subscriber types that generated revenue greater than 400 (USD) in the year 2015.

Report 5 The share of trips per neighborhood and per year. Consider only years 2014 and 2015.

1 Modeling

Given the dataset create a star schema compatible with the requirements above. Also define the concept hierarchies for each dimension. Briefly explain any assumptions you have made. We are primarily looking for you to show modeling principles for data warehousing.

2 OLAP Operations

On top of your schema specify sequences of OLAP operations generating Report 1, Report 2 and Report 3. Keep in mind following assumption: For all concept hierarchies assume the "All"-level (level with zero granularity) as default.

3 Implementation of the Cube

You are going to use the icCube (an OLAP server with a web-based user interface) to implement the schema/cube you have created in the section 1. Download¹ and install icCube on your computer. The provided dataset does not require any changes and only the *flights.csv* is needed step 3 and 4. All you need for creating a new schema/cube is the Builder tab. There you need to create a new schema, set the data source, define the dimensions and concept hierarchies, and define the cube and the measures.

4 Multi-Dimensional Expressions (MDX)

Deploy your schema and switch to the MDX tab. Finally, define and execute MDX queries generating the requested reports.

¹<http://www.iccube.com/download/>

Submission Requirements

In this assignment we expect you to submit following artifacts:

- A PDF file reporting the mdx code and the results of the sections 1, 2 and 4.
 - **Text must not be handwritten.** (Hand drawn figures are acceptable, but not wanted.)
 - Make sure that the document follows the usual conventions (**names**, assignment/task number, etc.).
- A file (**.icc-schema*) with the exported schema.

All assignment artifacts are to be delivered using **BlackBoard**. You are allowed to **work in maximum group of size 2**, and , one artifact must be delivered within the group submission.