A Template for a Scientific DAWA Paper

Markus Goldstein, Max Muster
University of Applied Sciences Ulm
markus.goldstein@thu.de,
max.muster@thu.de

***Abstract*—A short introduction to Latex is given by this example paper. Typically, your abstract should consist of a summary and your key findings. Write your abstract as the last item of your paper and (as a rule of thumb) address every section of your paper with one (sub-)sentence.**

I. Introduction

Here you give an introduction to your Lab Project. Describe what you want to achieve. The structure about the sections in this template is just an example, you may modify everything to your needs.

A. Scenario

Here you can describe your project in more detail, basically referring to the *Business Understanding* phase of the CRISP-DM process [1].

B. Structure of the Paper

The paper is structured as follows: in Section II the dataset about X is described. Then, in Section III the ... Finally, we conclude our work in Section VI with ...

II. Open Data: MyDataSource X

As part of the *Data Understanding* phase for this data science project, the two available data sources are described: The CSV file Y contains ... The material presented in this section is based on downloads/ online information (see [2], [4] and [1]) as well as the text books [5] and [3]...

III. Data Preparation

Describe your data preparation steps following the *Data Preparation* phase of CRISP-DM. You can also mention used tools here. It is also possible that you use paragraphs to structure your work.

a) Issue 1: The first question would be how to deal with the NULL values

b) Issue 2: Another challenge was to ...

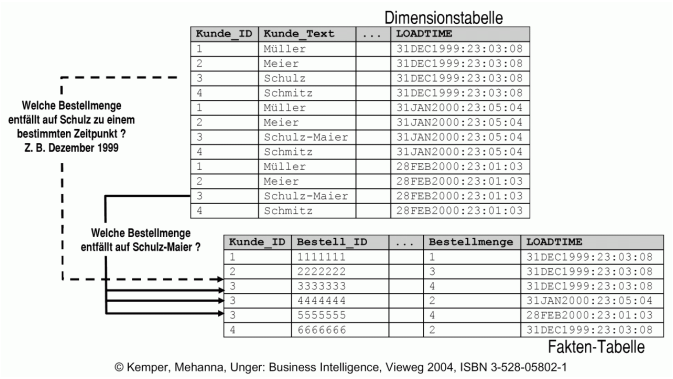

Fig. 1. Overview of solution strategy

A. Overall Solution Strategy

The solution approach is visualized in Figure 1. This means, we can solve the problem in steps as follows:

- step 1
- step 2
- ...
- step n

IV. Concept for CDWH

Of course you also need to describe your CDWH. Typically this is done by using an ERD in this section. In your descriptions, you can also use enumerations, for example what you will do:

1) Create the tables
2) Populate the dimension tables from coarsest to finest granularity
3) Populate the fact table
4) Create indices on ...

You can also describe your DM in this section or in more detail in a separate section. Make sure to mention also VIEWs, if used on top of the DM.

V. Analytics

This section contains your analytical results, including charts and/or pivot tables. Please also mention, how you connected MySQL to LibreOffice or BIRT.

VI. Conclusion

In this paper it was shown that ...

References

[1] *CRISP DM*. Wikipedia. https://en.wikipedia.org/wiki/Cross_Industry_Standard_Process_for_Data_Mining

[2] *Slowly Changing Dimensions*. Wikipedia. http://de.wikipedia.org/wiki/Slowly_Changing_Dimensions

[3] Kemper, Hans-Georg ; Mehanna., Walid ; Unger, Carsten: *Business Intelligence – Grundlagen und praktische Anwendungen*. 2. Aufl. ViewegTeubner, 2006

[4] Kimball, Ralph: *Slowly Changing Dimensions*. DBMS Magazine (Online edition). http://www.dbmsmag.com/9604d05.html. Version: 1996

[5] Kimball, Ralph ; Ross, Mary: *The Data Warehouse Toolkit. The Complete Guide to Dimensional Modeling*. 2nd ed. John Wiley & Sons, 2002