Introduction to Seminars:
How to Write Papers

University of Freiburg, Germany

Stefan Feuerriegel

Chair for Information Systems Research
October 2014

Agenda

1 Introduction

2 Paper Writing Process

3 Searching & Structuring Literature

4 Organizing the Paper

5 Writing the Paper

These are personal recommendations; there is no single "right" way

Agenda

1 Introduction

2 Paper Writing Process

3 Searching & Structuring Literature

4 Organizing the Paper

5 Writing the Paper

Research Interests

Data Analytics Research Group

- Focus on geo-spatial analysis joining maps with operative data
- Applications to the optimal placement of electric vehicle charge points
- Decision support for car sharing providers and geo-spatial variation in crime incidents

Sustainability Research Group

- Investigates mechanisms & business models to increase economic sustainability
- Applications to demand response, renewable energy sources and electric mobility
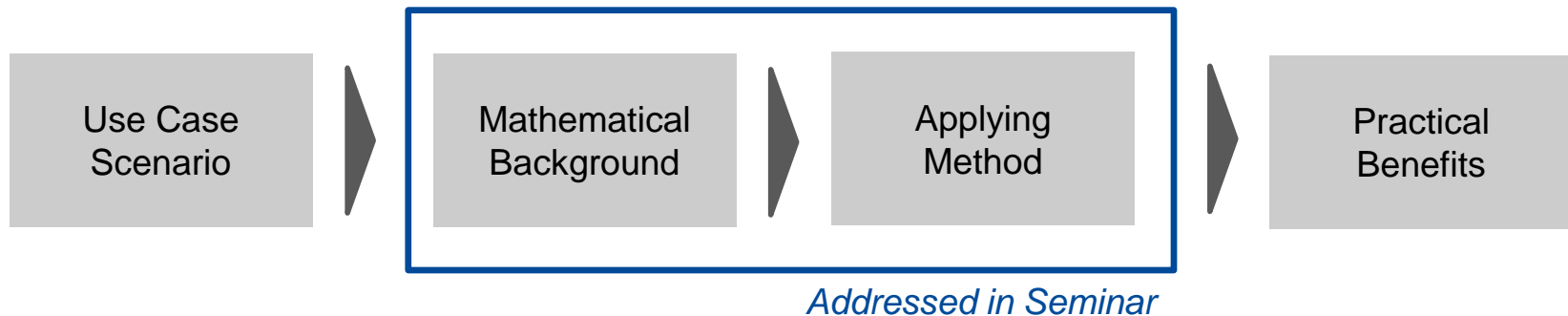
Finance Research Group

- Uses text mining of news to predict and explain stock market changes
- Applications to speculation, financial disclosures, crude oil, real estate and IPOs

Seminar covers data mining as methodological prerequisites → thesis

Motivating Data Mining

- Turning enterprise data into knowledge

- With the aim to increase (monetary) benefits

- Difficult because of the properties of Big Data
  → 4 V's: Volume, Variety, Velocity, Veracity

- Techniques of how data is collected, analyzed and visualized

| Use Case Scenario | → | Mathematical Background | → | Applying Method | → | Practical Benefits |

Addressed in Seminar

Example: potential annual value to US health care is $300 bn

Style of Assignments

Use Case Scenario	• Questions to be addressed in the introduction: When and where can method X be used?

▼

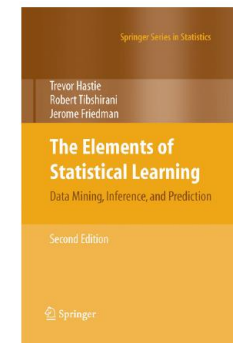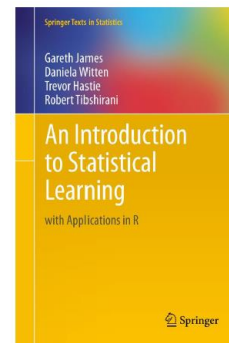|	
Mathematical Background	• Explain the mathematical background of method X as requested

▼

Applying Method	• Show how to use the corresponding routine in R • Provide and present results of the execution

Things to watch out for:
- Strict formatting rules
- First paper and revised paper with the chance for improvements
- No excess length in your presentation
- Presentation should address other students as audience → make it understandable

R Basics

- What is R? Free software aimed at statistical computing
- The #1 software for data mining[1]

- All practical applications must be implemented in R
  - R Studio recommended as editor
    → http://www.rstudio.com
  - Introductory materials can be found in the library or online, e.g.
    → http://cran.r-project.org/doc/manuals/R-intro.pdf

- Books on data mining with R:

[1] Rexer's Annual Data Miner Survey (2013)

Agenda

http://www.nature.com/scitable/content/ne0000/ne0000/ne0000/ne0000/14239512/ECS_scientific-papers_ksm.jpg

But, then, how do you write papers?

Writing Process

- Structure is important → late changes are expensive

Idea

Structure — Prepare a table with the structure and use this also to note the content of each section with a few bullet points

Bullet Points — Make bullet points for each paragraph with a message
→ 1 paragraph = 1 message, and consistent flow within paragraph

Draft — Write quickly, to get a draft

Writing — Then improve the writing iteratively
→ follow advice in Gopen: The Science of Scientific Writing (1990)

Agenda

Searching Literature

Reasons for Referencing	Sources
▪ Similar research	▪ **Google Scholar**
▪ Proof of relevance	▪ Web of Knowledge
▪ Proof of novelty (by pointing out differences)	▪ ScienceDirect
▪ Same methodology	▪ Journals (Springer, IEEE, …)
▪ Links for background search	→ University's WLAN/VPN necessary to get access

Software for Collecting Literature	▪ **Citavi** is powerful, but Windows only (recommended)
	▪ EndNote, Zotero are common, but less powerful
	▪ MS Word possible, but cannot organize references
	▪ Mendeley is web-based, but citation styles troublesome
	▪ BibDesk can be an alternative on Mac OS

Plagiarism

- **Plagiarism** is the stealing of someone's language, thoughts, ideas or expression, such as
  - Taking passes from someone else's work without citation
  - Rewriting without citation
- Considered as a breach of ethics
  → grade 5.0 as an immediate response

- Instead: give correct source!
  → add necessary references at the corresponding position
- When copying exact words, use quotation
  → this type is mostly avoided in science

Agenda

Sample Outline

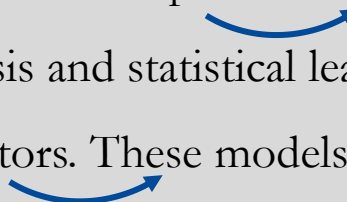| Section	Content
Abstract	What is the content in a nutshell?
Introduction	What is the problem?
Related Work (not necessary)	What are differences/similarities?
Methods	How is the problem solved?
Results	What are the findings?
Discussion (optional)	What does it mean?
Conclusion & Outlook	
Bibliography	What are the works referred to?
Appendices (optional)	Extra information

Introduction

- Helpful tools for structuring
  - Visualizing relationships through **graphics**
  - **Research questions** or hypothesis (as own highlighted paragraph)
- Don't write "… is important", instead give arguments

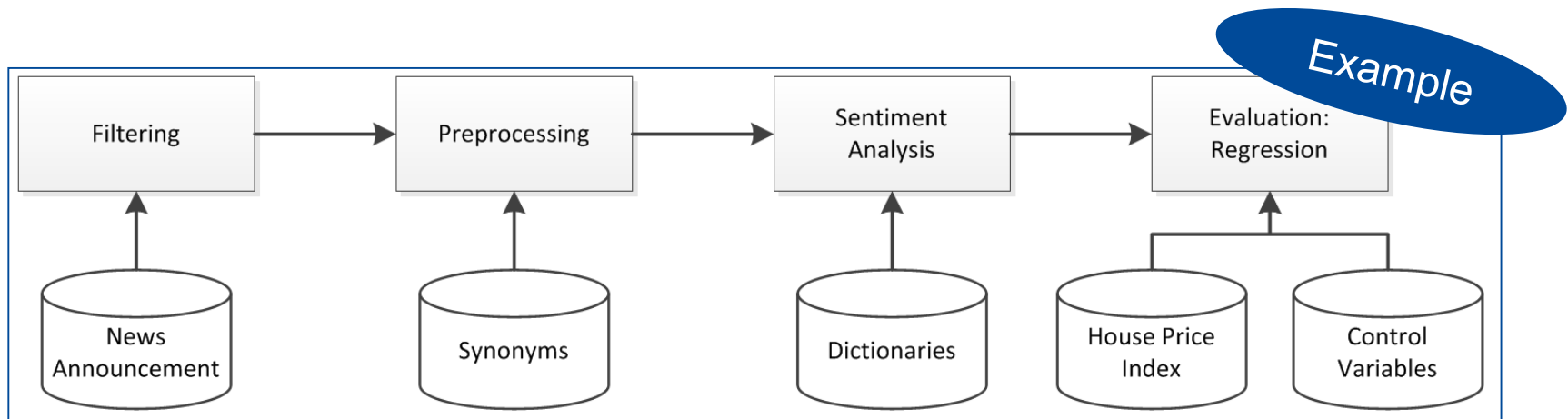| **Structure: 4 paragraphs**	
1	Motivate relevance of problem (probably give context)
2	State the problem briefly
3	Repeat contribution, e.g. as research questions (not in all journals, but recommended)
4	Outline (structure of section as story line)

Outline

Example

The remainder of this paper is structured as follows. Section 2 provides a literature overview of publications forecasting electricity prices, where the majority of models ignores external impacts. To close this research gap, Section 3 utilizes both time series analysis and statistical learning to present models that incorporate exogenous predictors. These models are evaluated in Section 4, which, finally, reveals that external inputs improve forecasting accuracy at a statistically significant level.

- Not just name "Section 2 analyzes …", but link sections
  → this explains and motivates the need for each section

Methodology

- Pay attention to formulae
  → even little errors seem unprofessional

- Visualize via flow diagram if helpful
  → MS Visio as a skillful tool

Example

Conclusion & Outlook

**Structure:
3 paragraphs**

1 Repeat problem and its relevance

2 Repeat contribution (plus quantitative results)

3 Give outlook on further research steps

Agenda

Figures

- Name axes with units, as well as legend

- Always state the message of what you see in the text

- Recommended also to explain plot structure in the text
  → including which data rows, possible trend or outliers

- Reference figure always in the text ("… in Figure 3")

Detailed Captions	- Aim: Most scan papers, thus, plot must be understandable without surrounding text - Name axes and data rows - Give details on set-up of e.g. experiments - Always end with punctuation

Captions

Examples

- Related literature on forecasting electricity prices using exogenous predictors.

- Comparison of forecasting accuracy in terms of root mean squared error (RMSE), measured from Feb 01, 2010 to March 21, 2012 across various models.

- Scatterplot of hourly day-ahead electricity prices and expected generation of wind (left) and solar (right) power with LOWESS trend line.

- Pooled regression comparing news reception in bullish and bearish markets.

- Demand Response potential through load shifting in Germany (Klobasa, 2007); scaled according to retailer's electricity demand.

Beginning a Section

- Always separate two headlines by introductory sentences
- State briefly what you are doing at beginning of each section

Examples	
	- In this section, we present related literature grouped into two categories.
	- To fully capture the concepts of …, we start by presenting different angles of the existing … approaches.
	- This section introduces the theoretical background on modeling …
	- In this section, we evaluate our proposed models in terms of ….

Hints

- Use understandable/clean English as native speakers are rare
- Ask native speaker for proof-reading
- Use highlighting where necessary (italics only)
- Use high quality graphics, not pixelated
- Numbers: use 2–4 significant digits (except amount of money)

Dictionaries	- http://dict.leo.org for Germans to find alternatives - http://ldoceonline.com to check the **meaning** and proper usage - http://dictionary.com is a good alternative - http://thesaurus.com to find **synonyms**