WIRTSCHAFTS
UNIVERSITÄT
WIEN VIENNA
UNIVERSITY OF
ECONOMICS
AND BUSINESS

Teaching Applied Statistics
A Vision for the 21st Century

Ronald Hochreiter Christoph Waldhauser

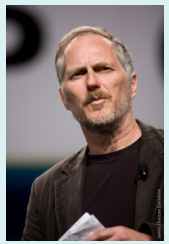
October 4th, 2011

Why Applied Statistics is so Important I

A profession with future – Voices from the industry:

"*For Today's Graduate, Just One Word: Statistics*" – **NYT, 2009**

"*Data is the next Intel Inside.*" – **Tim O'Reilly, 2005**



"*The sexy job in the next ten years will be statisticians.*"
"*The ability to take data – to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it, – that's going to be a hugely important skill in the next decades.*" – **Hal Varian, 2009**

Why Applied Statistics is so Important II

In a data-centric society. . .

. . . the political becomes dependent on statistics.

Problem statement

- Lies,
- Damn lies,
- & Statistics.

Solution

Students of applied statistics are empowered to detect and counteract forged statistics.

In a data-centric society. . .

. . . applied statisticians can live up to their civic responsibility.

What is Applied Statistics

Hal Varian's vision

"The ability to take data – to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it, – that's going to be a hugely important skill in the next decades."

What to Teach

Hal Varian's vision – **a menu**

"The ability to take data – to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it, – that's going to be a hugely important skill in the next decades."

Main course

- Understanding
- Processing
- Extracting value
- Visualizing
- Communicating

Side dishes

- Academic English
- Presentation techniques
- Field specializations
- Mathematical extensions
- Advanced programming

Understanding

Questions

Where do data come from?
What do they mean?

Typical course contents

- applied inferential statistics
- complex survey sampling
- experimental design

Processing

Questions

How to retrieve data?

How to handle data?

Typical course contents

- data structures and data bases
- XML, XPath, JSON
- REST, SOAP, web harvesting

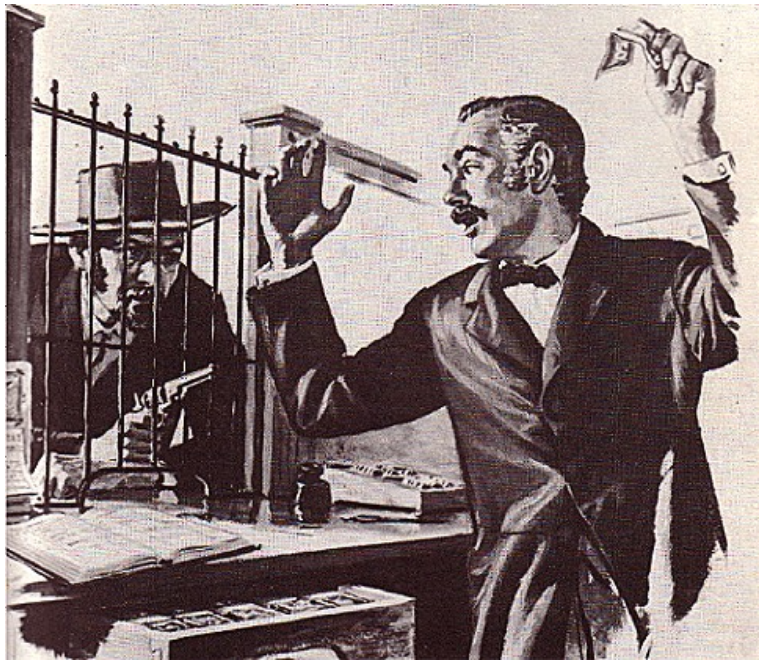
Extracting Value

Questions

Which patterns can be found?
What do they mean to others?

Typical course contents

- machine learning
- collaborative sciences



Visualizing

Questions

Human vision is excellent in finding patterns

- How can we look at data?
- Which techniques can aid in discovering?

Typical course contents

- Data visualization theory
- Processing, GGobi

Communicating

Questions

How to communicate findings

- to academic peers?
- to the masses?

Typical course contents

- Communication psychology
- Creative & academic writing
- Collaborative writing

Cornerstones of Teaching Applied Statistics

Demanding Programs: Challenge Your Students

- Trust your students to learn prerequisites on their own
- One ECTS point indicates at least 30 hours of workload
- In the US, master students sleep in their libraries

Realistic Training

- Enable students to solve real problems
- Use real-world data & examples
- Allow students to discover & explore

Be Smart: Free Software

- Save licensing costs
- Invest in hardware
- permit students lab access 24/7

Free Software

Commercial Software is Expensive

- WU pays EUR 105,000 per year for Windows/SPSS licenses alone
- Windows/SPSS requires top hardware to run (even in labs)
- Administrating Windows is a nightmare

Replacement Options

MS Windows: Ubuntu Linux

MS Office: Open Office, LaTeX, Google Docs

SPSS: R Environment for Statistical Computing

none: GGobi

MATLAB: GNU Octave