

Advanced Data Analysis for the *Physical Sciences*

Dr Martin Hendry
School of Physics and Astronomy
University of Glasgow
martin@astro.gla.ac.uk





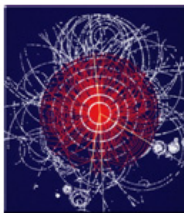
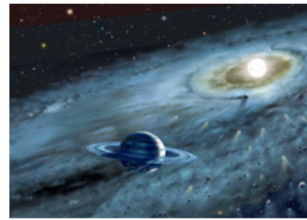
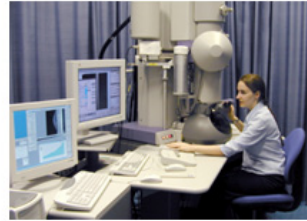
2-day residential course which aims to:

- Review the theoretical foundations and practical methods of modern data analysis in the physical sciences
- Summarise the basic tools of parameter estimation and testing goodness of fit
- Explore and contrast Bayesian and classical (frequentist) methods
- Highlight modern computational advances in analysis of very large datasets.

The SUPA Graduate School is a cornerstone of the SUPA Alliance.

Shared Delivery of Graduate Education:

- New specialist courses are made viable by pooling expertise.
- Delivery by a variety of modes, lectures, distance learning, short courses.
- SUPA-funded Access Grid rooms in each Institution for two-way video delivery.
- Scotland-wide relay of research seminars and colloquia.



Courtesy: Horiba Jobin Yvon IBH Ltd; Coherent Scotland Ltd; CRLO Displays Ltd

Prize Studentship Competition:

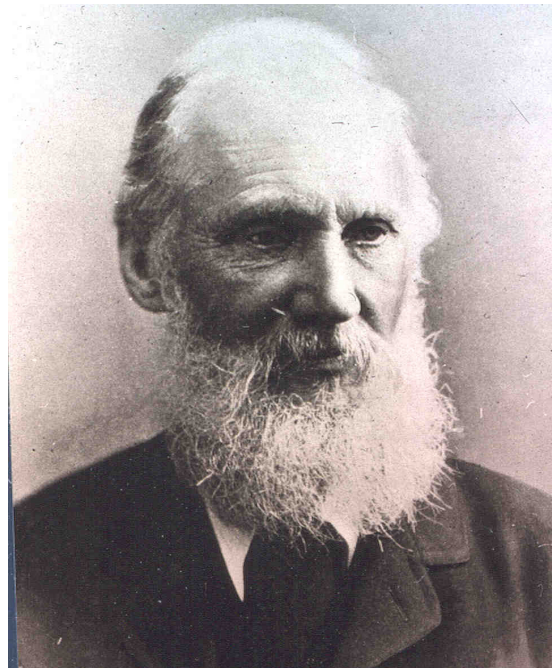
- High profile and prestigious awards.
- Attracting the brightest and best, worldwide.
- 8 awards per year, starting October 2005.
- To be held at any SUPA University and in any theme.

International Summer Schools:

- Building on a 40-year track-record of SUSSP.
- Attracting world-renowned lecturers.

Distinguished Visitor Programme:

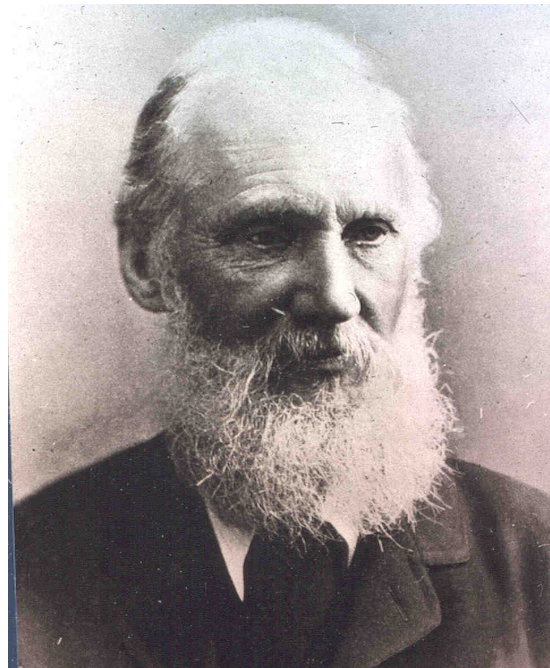
- The best scientists in the world spending time in Scotland.
- Adding variety and weight to the local teaching base.
- Stimulating research in new areas.
- Enhancing the profile of Scottish Physics on the world stage.



William Thomson
(Lord Kelvin)
1824 - 1907



William Thomson
(Lord Kelvin)
1824 - 1907



*"There is nothing new
to be discovered in
physics now. All that
remains is more and
more precise
measurement"*

Course Programme

Wednesday 5th January Room 222 Kelvin Building

| | |
|---------------|---|
| 10.00 - 11.15 | Introduction and theoretical foundations |
| 11.15 - 11.45 | Coffee break |
| 11.45 - 13.00 | Introduction and theoretical foundations contd. |
| 13.00 - 14.15 | Lunch |
| 14.15 - 15.30 | Parameter estimation and goodness-of-fit: Part one |
| 15.30 - 16.00 | Coffee break |
| 16.00 - 17.00 | Parameter estimation and goodness-of-fit: Part two |
| ~19.00 | Dinner at Mother India |

Course Programme

Thursday 6th January Room 222 Kelvin Building

| | |
|---------------|--|
| 09.30 - 11.00 | An advanced toolbox for Bayesian inference |
| 11.00 - 11.30 | Coffee break |
| 11.30 - 12.45 | Bayesian Model Selection |
| 12.45 - 14.00 | Lunch |
| 14.00 - 15.30 | Advanced Numerical Methods and Wrap-up session - future plans |

Lecture notes available on my.SUPA

See also: <http://www.astro.gla.ac.uk/users/martin/supa-da/>

Course Assessment

- Series of 15 'Pop Quiz' MCQs, spread throughout the lectures
- Series of 10 numerical problems, to be posted on my.SUPA after the course.
- Mock data challenge posted on my.SUPA after the course; more details tomorrow.

Assessment designed to be **formative** (for me and you!)

Practical matters: Coffee Breaks and Lunch

- Tea & coffee, soft drinks and snacks available from vending machines in Physics and Astronomy Common Room (450). Follow the signs, and the locals !
- Wide variety of options for lunch.

Byres Road: Little Italy, Subway, Greggs, Peckhams,
Paperinos, Tinderbox, Naked Soup...

On campus: John MacIntyre Cafe, Research Club, QM Union

Ashton Lane: Ketchup, The Loft

Cresswell Lane: Bean Scene, Café Andaluz

Practical matters: Coffee Breaks and Lunch

- Tea & coffee, soft drinks and snacks available from vending machines in Physics and Astronomy Common Room (450). Follow the signs, and the locals !
- Wide variety of options for lunch.

Byres Road: Little Italy, Subway, Greggs, Peckhams, Paperinos, Tinderbox, Naked Soup...

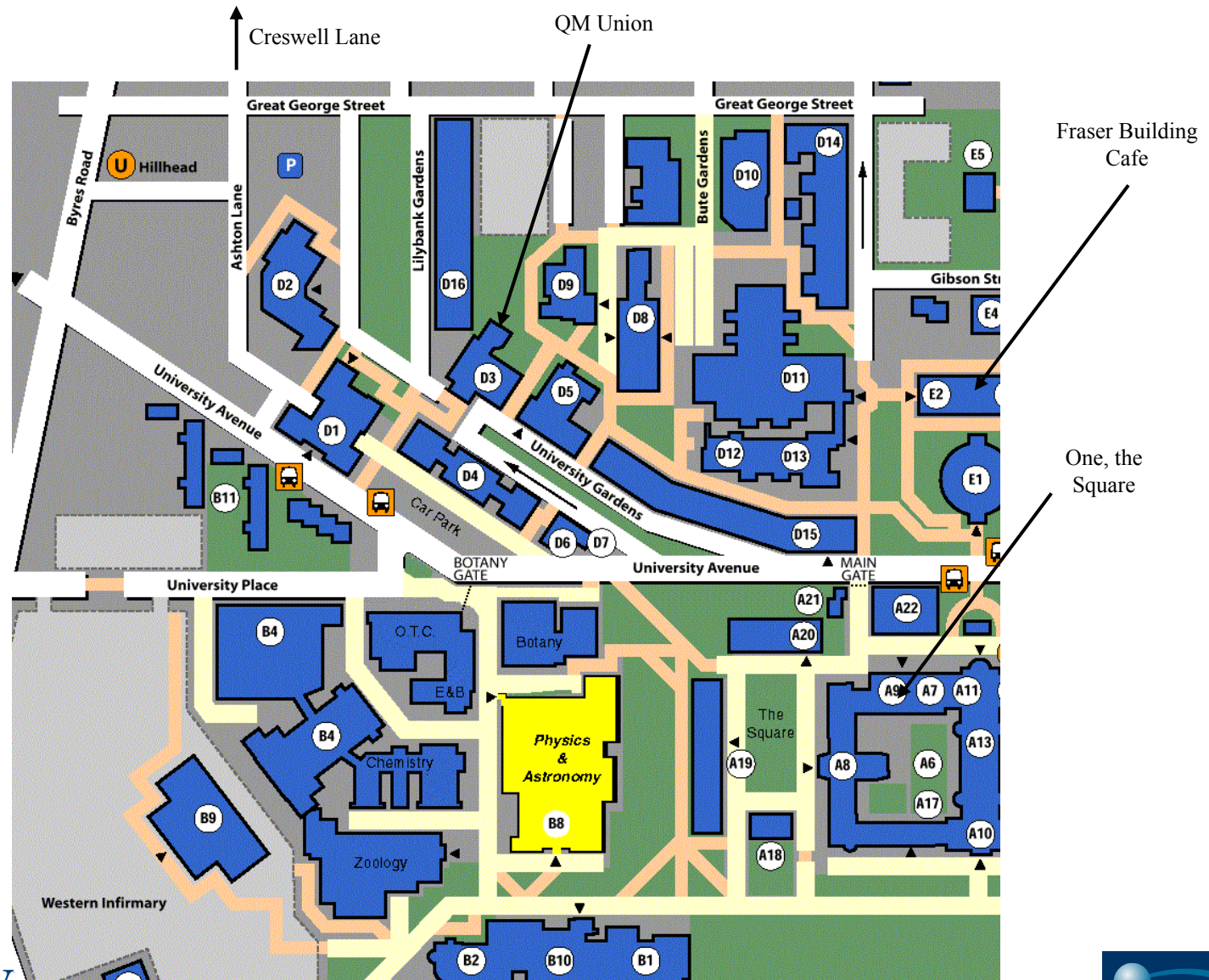
On campus: Fraser Building Cafe, QM Union

Ashton Lane: Ketchup, The Loft

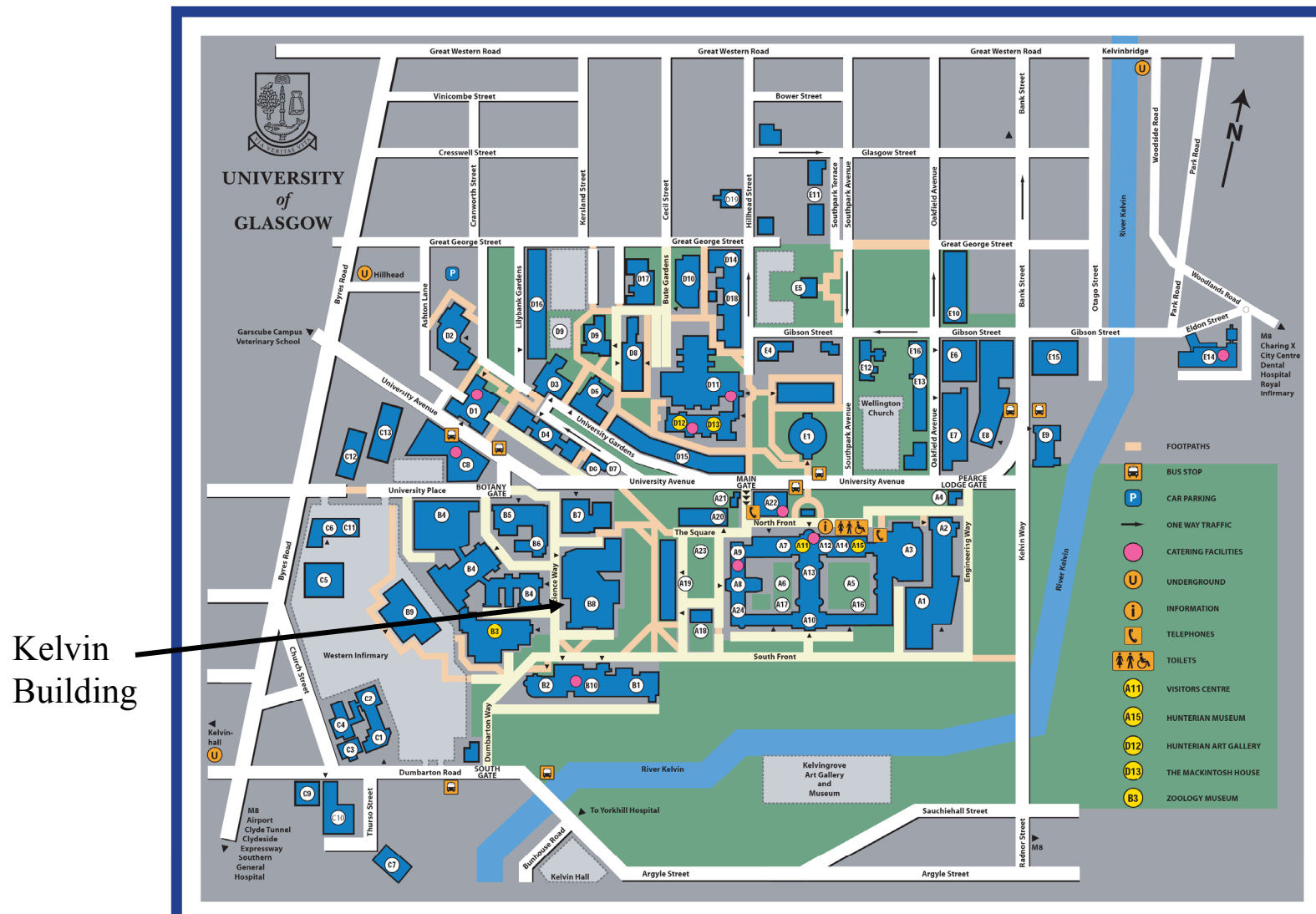
Cresswell Lane: Bean Scene, Café Andaluz

Follow the locals!...

Practical matters: Getting around



Practical matters: Getting around



Practical matters: Dinner - Wednesday evening



[Mother India](#) [Mother India's Café](#) [Ashton Lane](#) [Buccluech Street](#)

MOTHER INDIA

28 WESTMINSTER TERRACE
GLASGOW
G3 7RU
T: 0141 221 1663 F: 0141 221 2217

AUCHENTOSHAN RESTAURANT OF THE YEAR, 2002

- [HOME](#)
- [MENU](#)
- [EVENTS](#)
- [RESERVATIONS](#)
- [RECIPES](#)
- [CONTACT US](#)



Mother India launched its individual style of authentic Indian cuisine on Glasgow's Westminster Terrace in 1995. The stunning, high ceiling property offered something new to Glasgow's discerning diners, a unique take on traditional Indian cuisine. A decade later, owner [Monir Mohammed](#) & manager Asam Rashid are still serving up delights from their innovative menu.



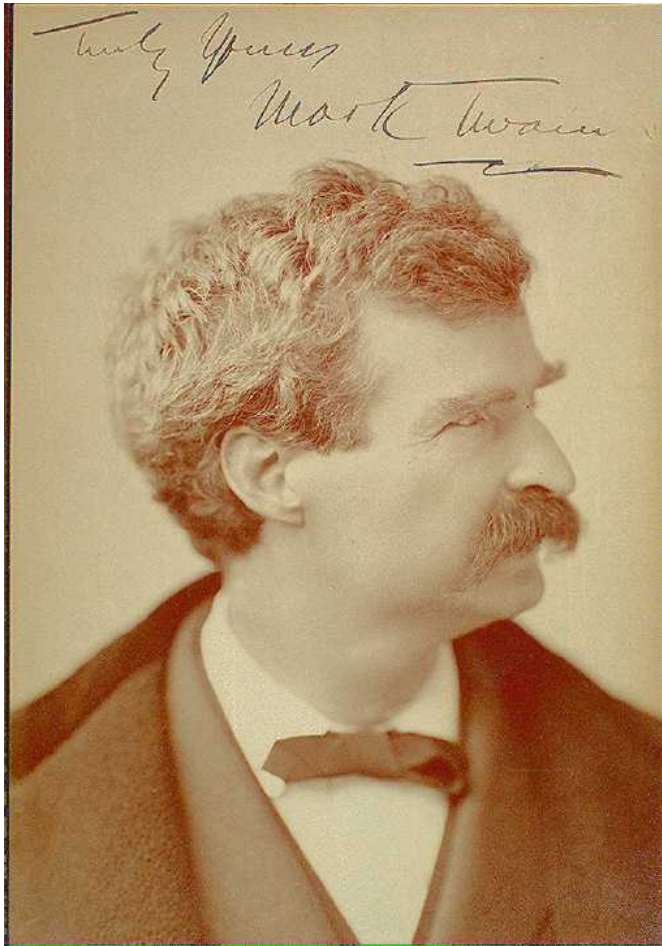
Practical matters: Getting around

Kelvin Building

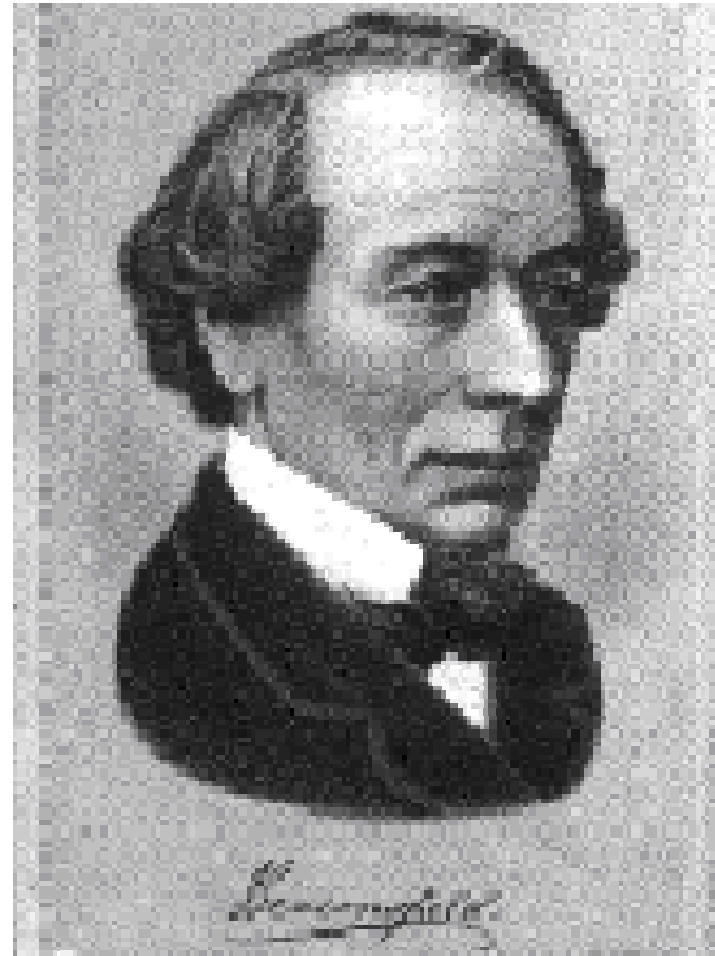


Mother India

Why a course on data analysis?...

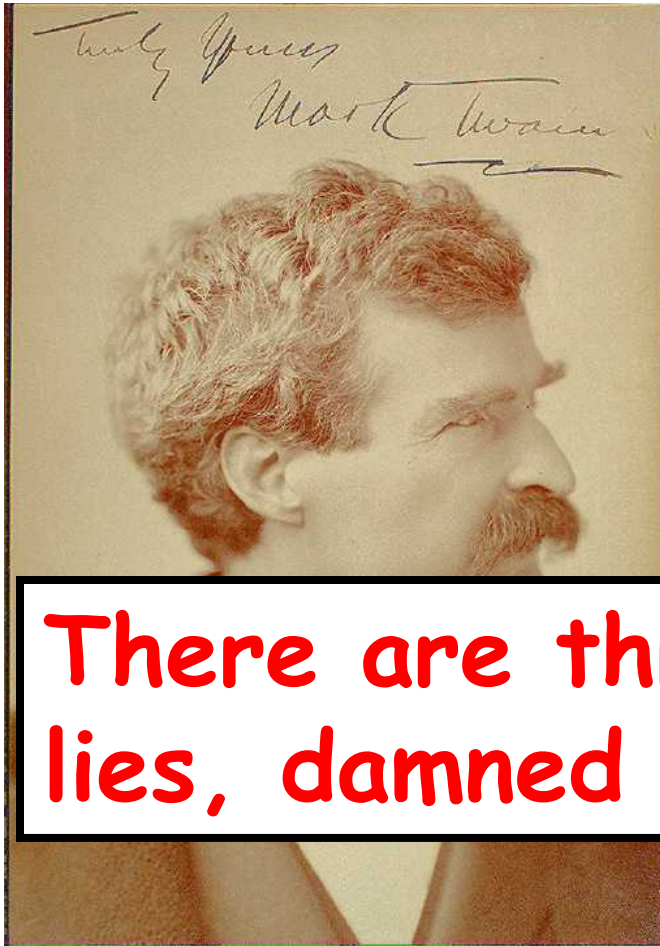


Mark Twain

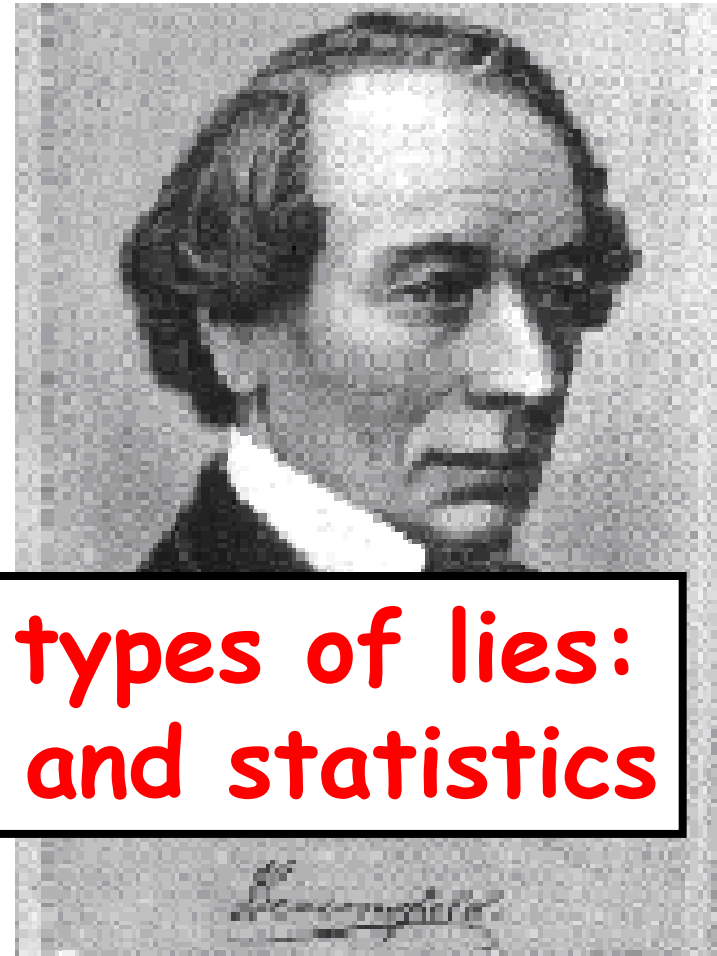


Benjamin Disraeli

Why a course on data analysis?...



Mark Twain



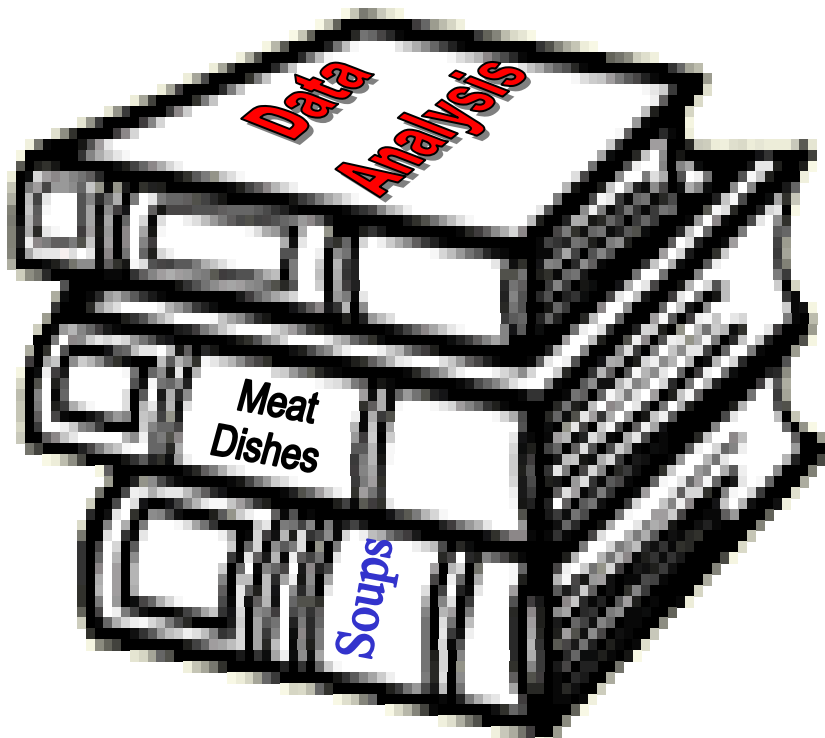
Benjamin Disraeli

**There are three types of lies:
lies, damned lies and statistics**

Why a course on data analysis?...

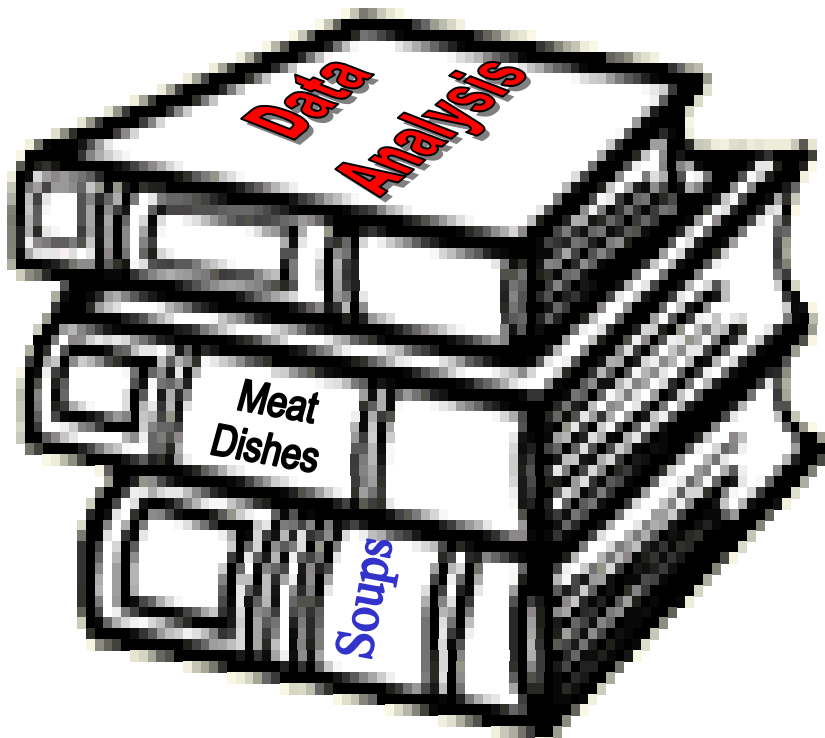
Data analysis methods are often regarded as simple recipes...

...but in physics, sometimes the recipes don't work!!!



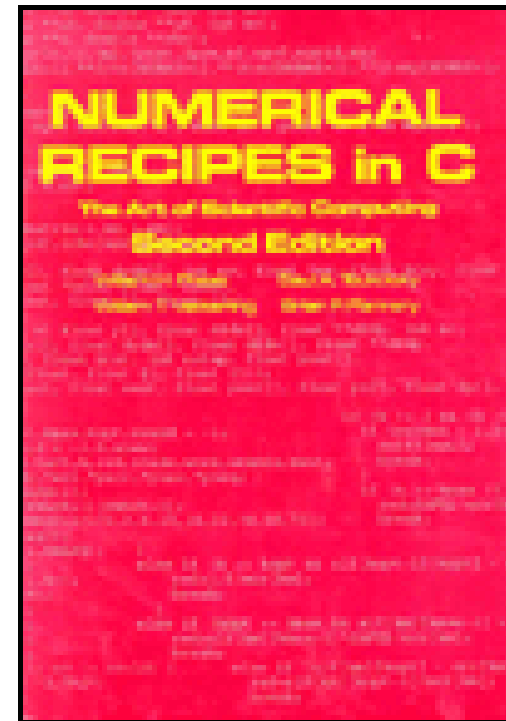
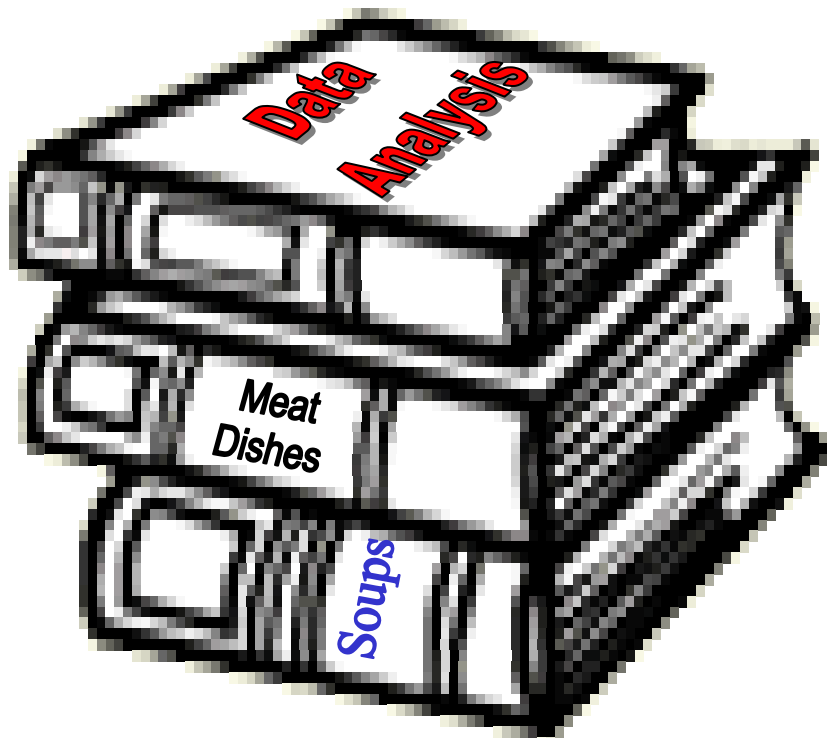
Why a course on data analysis?...

Data analysis methods are often regarded as simple recipes...



Why a course on data analysis?...

Data analysis methods are often regarded as simple recipes...



<http://www.numerical-recipes.com/>

<http://www.nr.com/olderswitcher.html>

Why a course on data analysis?...

Data analysis methods are often regarded as simple recipes...

...but in physics, sometimes the recipes don't work!!!

- Very weak signals
- Correlated 'residuals'
- Incorrect assumptions



SYSTEMATIC ERRORS



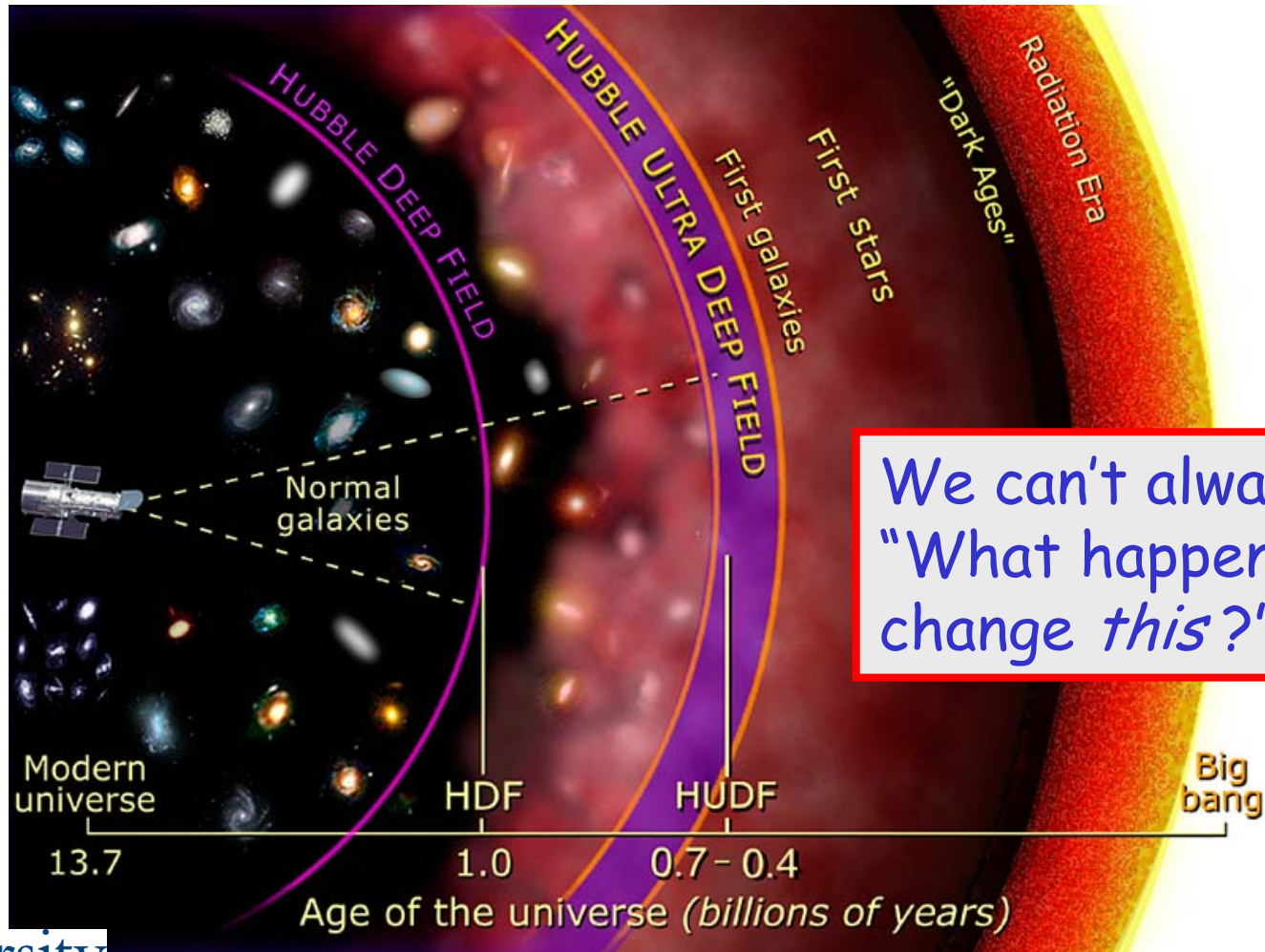
Why a course on data analysis?...



“You must unlearn what you have learned”

Why a course on data analysis?...

Many areas of physics are *Remote sensing*



Why a course on data analysis?...

Even if you (will) do little data analysis yourself, you will need to assess critically results in the literature of your field.

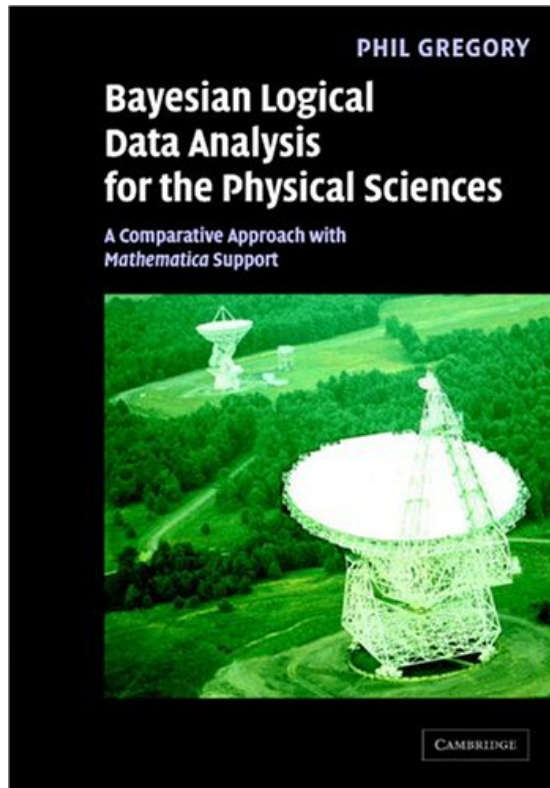
—————> Determine **significance** of old and new theories and models.

Many (all?) current analyses are carried out within a particular framework - the 'frequentist' approach to probability:

- 'classical' data analysis methodology
- slowly losing ground to **Bayesian inference**.

Important to understand the differences between these two approaches, and the strengths of the new Bayesian paradigm.

Why a course on data analysis?...

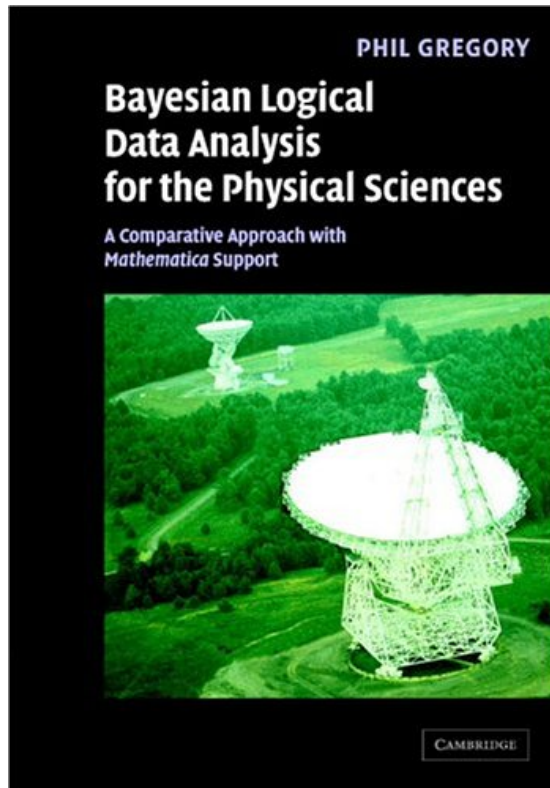


Cambridge Univ. Press

ISBN: 052184150X

Why a course on data analysis?...

PREFACE



Cambridge Univ. Press

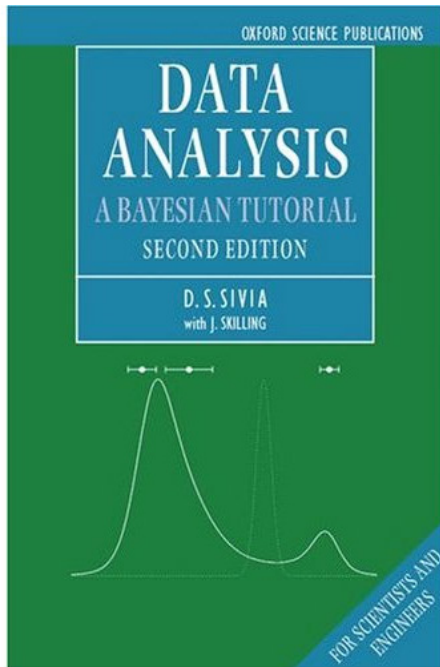
ISBN: 052184150X

The goal of science is to unlock nature's secrets... Our understanding comes through the development of theoretical models capable of explaining the existing observations as well as making testable predictions... Statistical inference provides a means for assessing the plausibility of one or more competing models, and estimating the model parameters and their uncertainties. These topics are commonly referred to as "data analysis".

We are currently in the throes of a major paradigm shift in our understanding of statistical inference based on... Bayesian Probability Theory... The Bayesian paradigm is becoming very visible at international meetings of physicists and astronomers. However, the majority of scientists are still not at home with the topic and much of the current scientific literature still employs the conventional "frequentist" statistical paradigm.

This book is an attempt to help new students to make the transition while at the same time exposing them to some of the essential ideas of the frequentist paradigm that will allow them to comprehend much of the current and earlier literature and interface with his or her research supervisor.

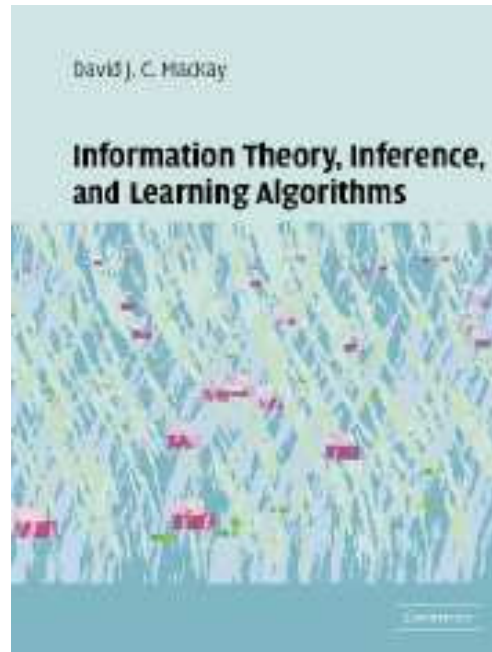
Why a course on data analysis?...



Data Analysis:
A Bayesian Tutorial

(Oxford Univ Press)
D.S. Sivia & J. Skilling

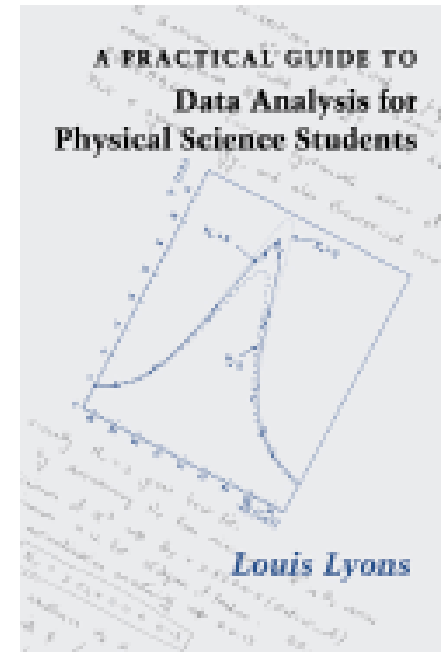
ISBN: 0198568312



Information Theory,
Inference and Learning
Algorithms

(Cambridge Univ Press)
D.J.C. Mackay

ISBN: 0521642981



A Practical Guide to
Data Analysis for the
Physical Sciences

(Oxford Univ Press)
Louis Lyons

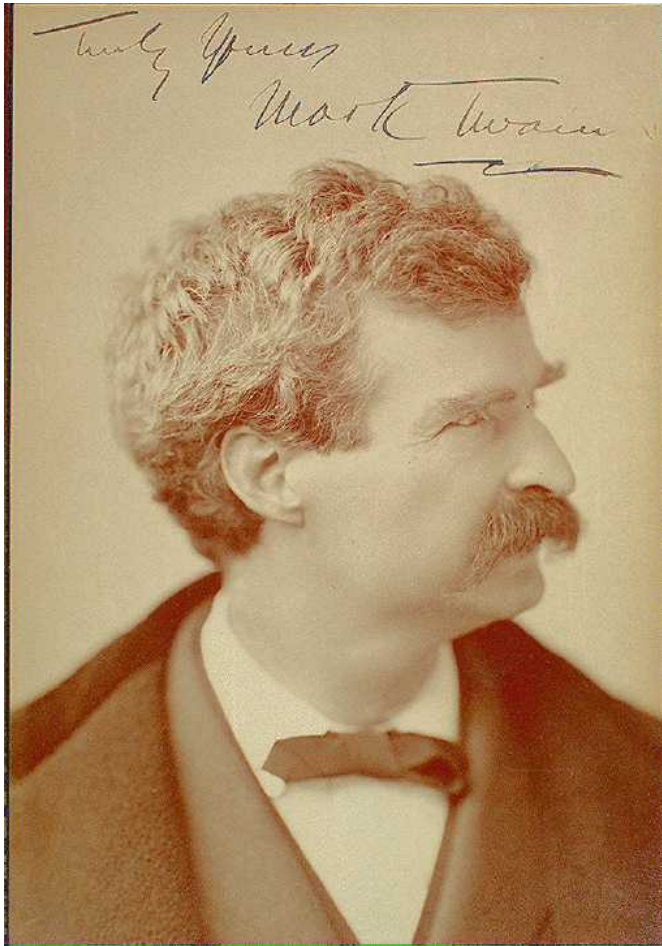
ISBN: 0521424631



University
of Glasgow

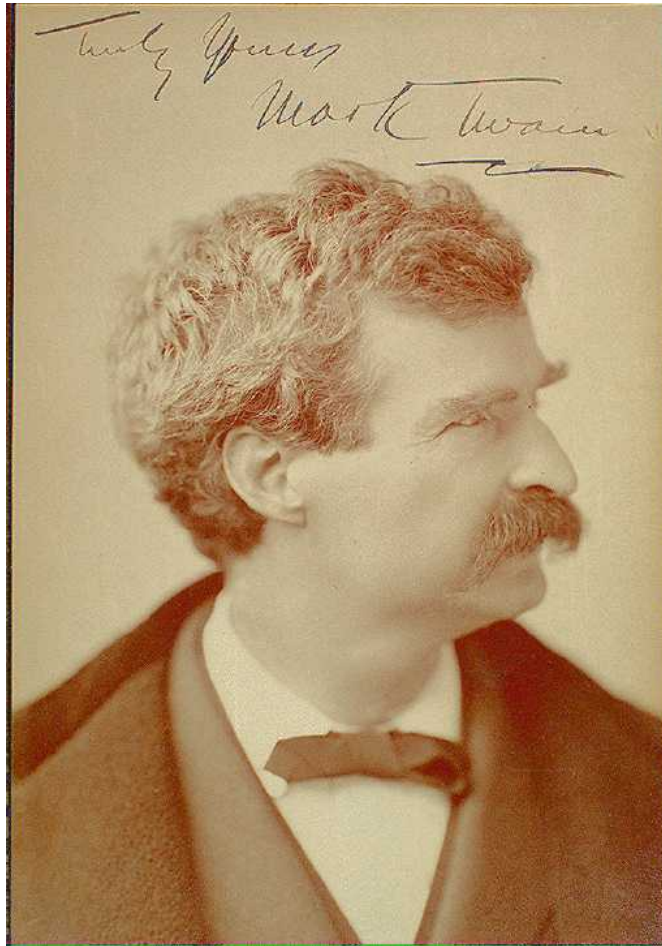
SUPA Advanced Data Analysis Course, Jan 5th – 6th 2011





Mark Twain





Mark Twain

“It is better to keep your mouth closed and let people think you are a fool than to open it and remove all doubt.”