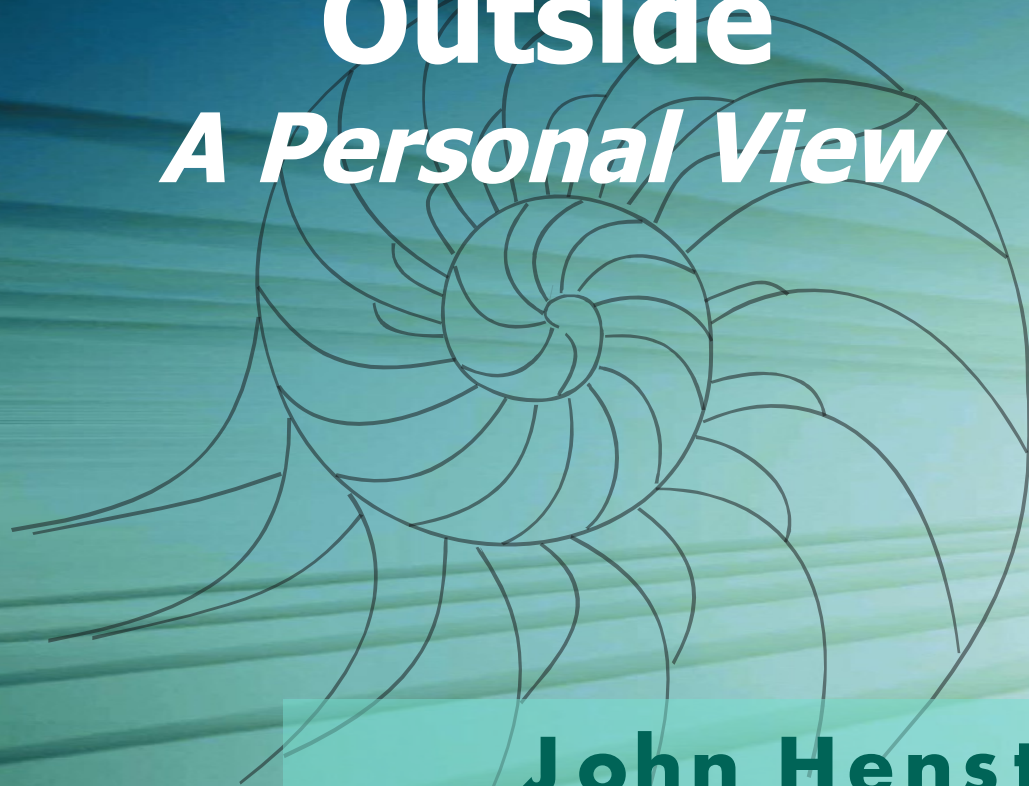


# Mathematics from the Outside *A Personal View*



**John Henstridge**

*Data Analysis Australia*

**September 2010**

# Overview

---

## → Background

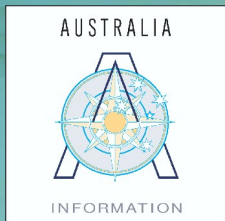
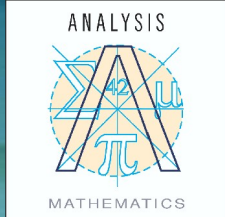
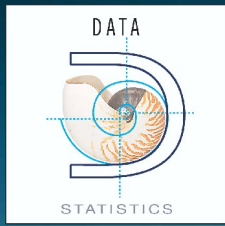
- Education
- Work Experience
- Mentors

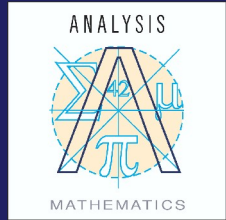
## → Examples

## → How to work

## → What I am doing now

- Data Analysis Australia





# How Did I Get Into It?



# Undergraduate

## → Flinders in 1969

- New radical university, 1400 students
- Intended to study physics
- Took about one week to discover mathematics
- Due to Igor Kluvanek

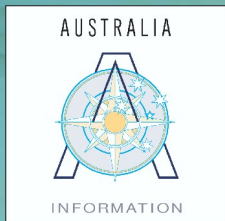
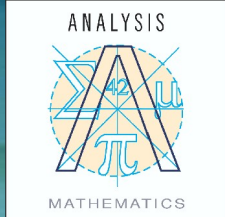
## → In 2<sup>nd</sup> year changed to 100% mathematics

## → 3<sup>rd</sup> year statistics

- Closest to operations research
- John Darroch, Miloslav Jirina

## → Overloaded 4<sup>th</sup> year

- Joint with Adelaide



# Postgraduate Study

## → ANU with Ted Hannan

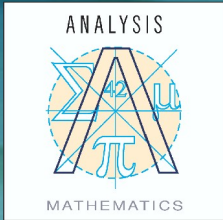
- Time series for signal detection with arrays
- Applied to seismology and geophysics

## → ANU department great place

- Leadership of Pat Moran
- Continual stream of visitors
- Continual seminars, including morning teas

## → And I learnt to ski

- And met Cheryl



# University of Western Australia

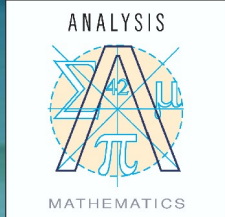
---

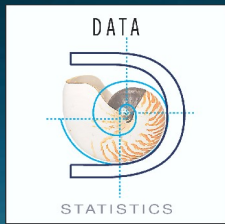
## → Initially mathematics

- Terry Speed + consulting
- Statistical software in time series

## → Then Agriculture

- Teaching + research + consulting
- Sabbatical in UK
- Tenure
- An offer from another university





# Mutton Bird Growth Curves

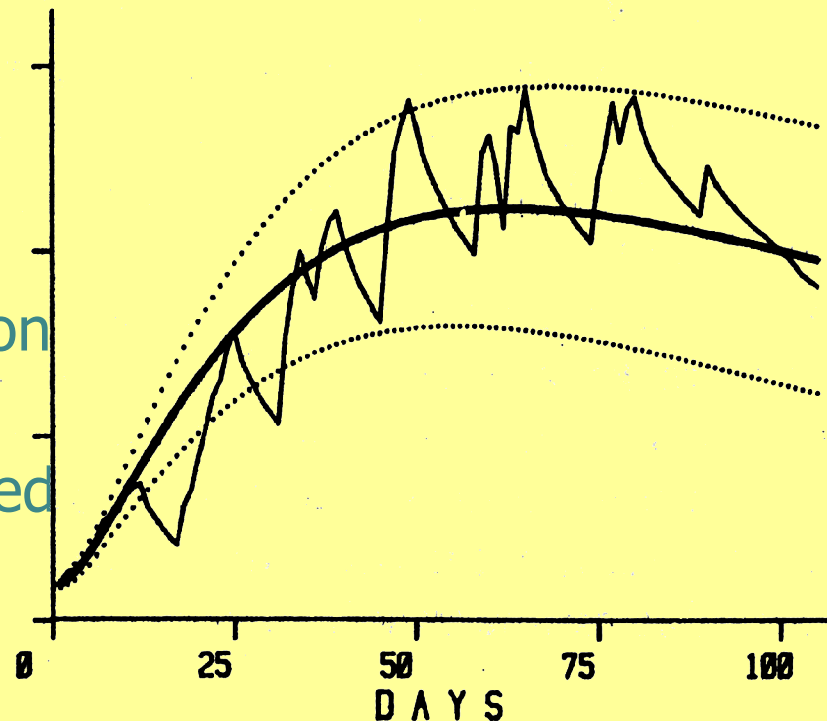
## → Unusual growth

- Chicks fed irregularly
- Weight declines between feeds

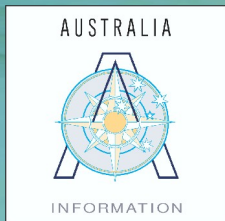
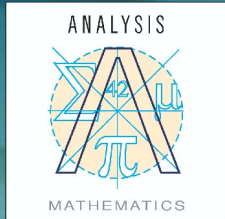
## → Storage model

- Non-stationary Poisson model for feeds
- Gamma model for feed size
- Exponential decline between feeds

BIRD 2



With Richard Tweedie



# Siromath – 1983-87

## → Commercial outgrowth of CSIRO

→ Managed by Richard Tweedie

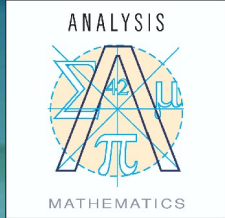
## → First Australian attempt at commercial statistics

→ Made possible by “cheap” computing - about 100 times as expensive as now

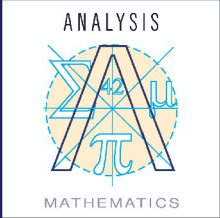
→ “Modern software”

- But still lots of programming

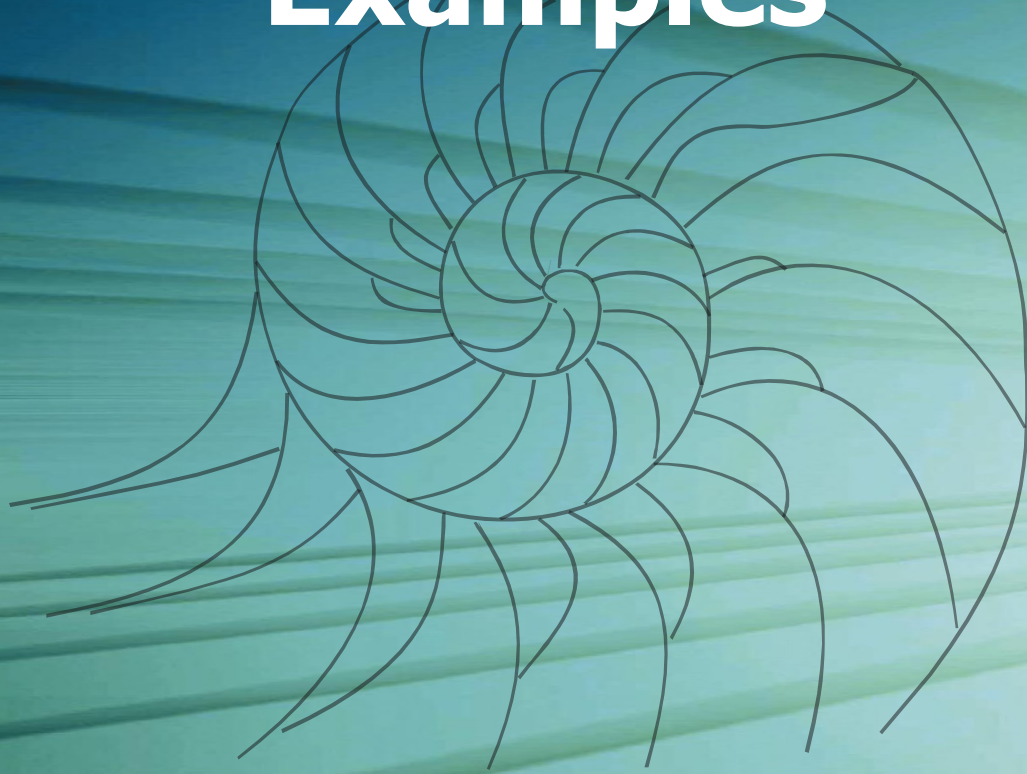
## → Left to set up Data Analysis Australia







# Examples



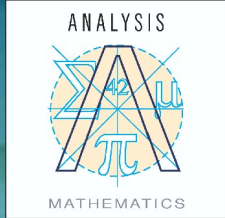
# Checking Out a Casino

## → Check electronic games for new Casino

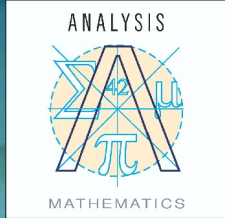
- Return to players
- Security
- Reliability

## → Failed random number generators

- Discovered that they fitted expected distributions *too* well
- Poor algorithm – check against Knuth
- Manufacturer ignored warnings
  - Machines paid out 150% at another casino
  - Went bust



# A Tourism Survey



## → Western Australian Travel Survey WAT S

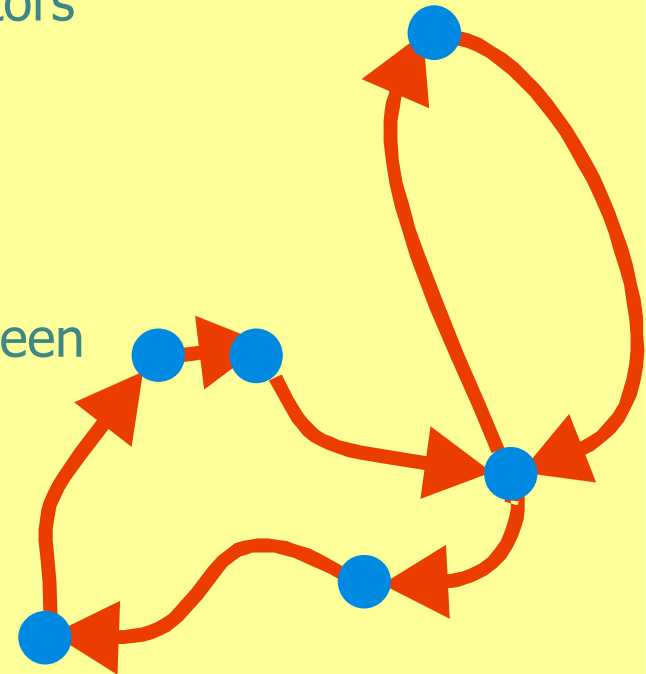
- New model for tourism surveys
- All travellers anywhere in WA

## → Complex sampling

- Households – residents and visitors
- Establishments

## → How do you weight?

- Model sampling probability
- Each observed trip could have been sampled in many ways
- Complex computation
  - Database organisation



# Public Housing Queues

## → Part of large project

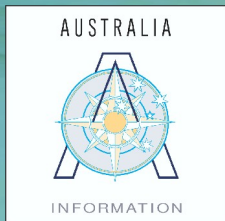
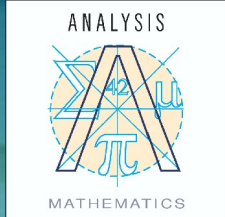
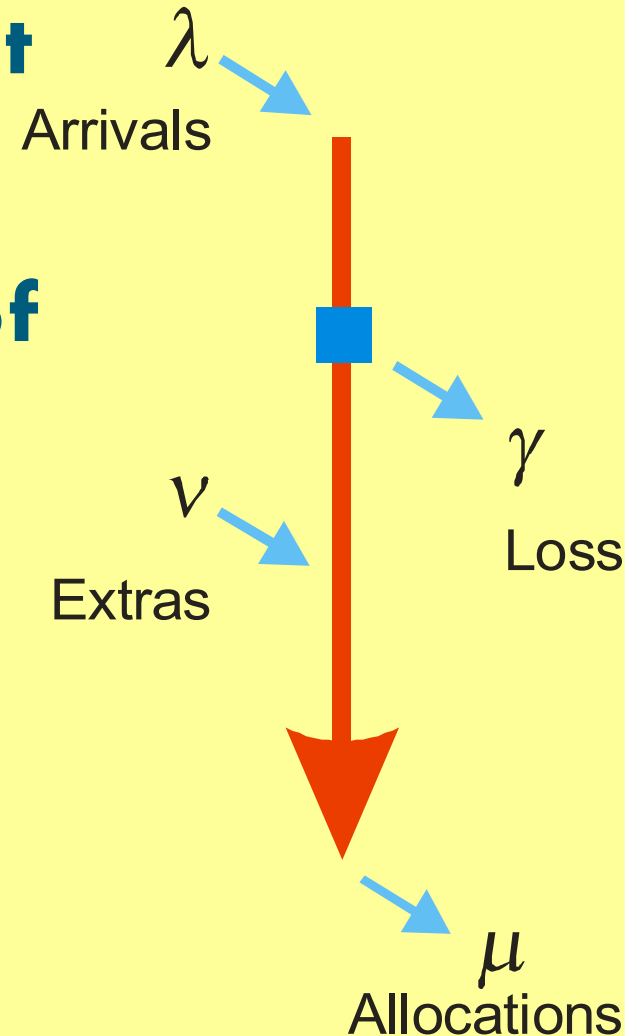
→ The client asked for queuing theoretic approach

## → Needed new form of queue

→ Standard arrivals, servers  
→ But **drop out** before service  
→ And extra queue **jumpers**

## → Model predicted waiting times

→ But understanding more important



# A Fair Trial for Laurie Connell ?

→ **Will a jury have someone with prior certainty that he is guilty?**

→ Probability calculation

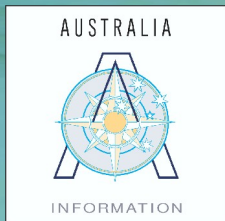
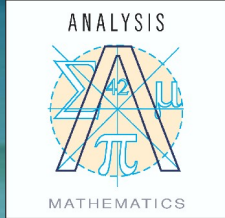
→ Gold standard telephone survey of people who could be jurors

- Strict protocols
- Call-backs

→ **Expert evidence**

→ Explaining it to a Judge

→ Cross examined by prosecutor



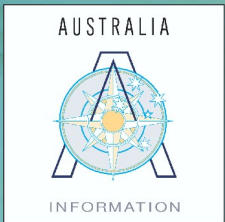
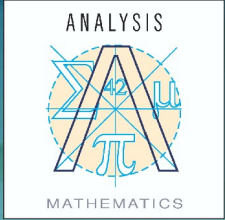
# Christmas Island Airport

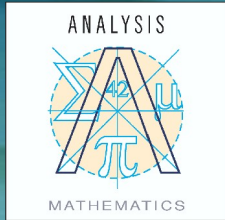
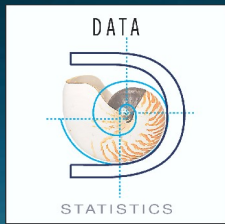
## → Forecasting passengers

- In a very uncertain environment
- Many possible developments

## → Risk analysis approach

- Probabilities of what might happen, when and how big
- Every number a guess
- Simulations to give bounds on what might happen





# Random Breath Testing

## → Simple question – does RBT work?

- Does it reduce road trauma?
- Any other evaluation is secondary
- Typical road safety problem

## → Not easy to answer

- Poor data, not enough data
  - Not enough crashes
  - Non-crashes not recorded
- Not just one road safety measure
  - Confounding

## → Good statistical methods essential

- Done in **WA**, NSW, Tas, Qld and SA

# How Accurate is the Defence Inventory ?

→ **Over 500,000 different stock codes**

→ Paperclips to replacement periscopes

→ Values <1c to >\$1M

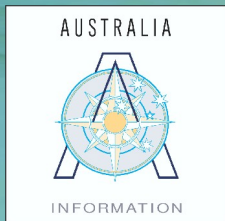
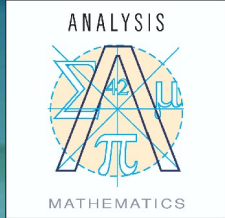
→ **Hundreds of locations**

→ Warehouses across Australia and overseas

→ **Dozens of organisations supplying, managing, consuming**

→ **Have to meet Accounting Standards**

→ Work with accountants





# Odds that Two Horses will Collide?

## → Horses on a training track

- Some galloping down the track
- Others trying to cross it

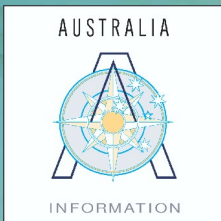
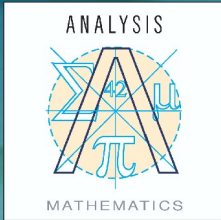
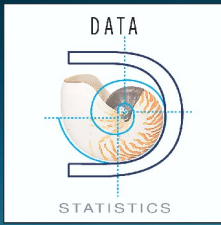
## → Legal question

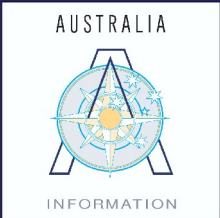
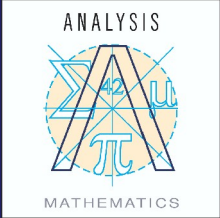
- Were the risks of collision so high that action should have been taken?

## → Data on speeds and spacing of gallopers, horses crossing

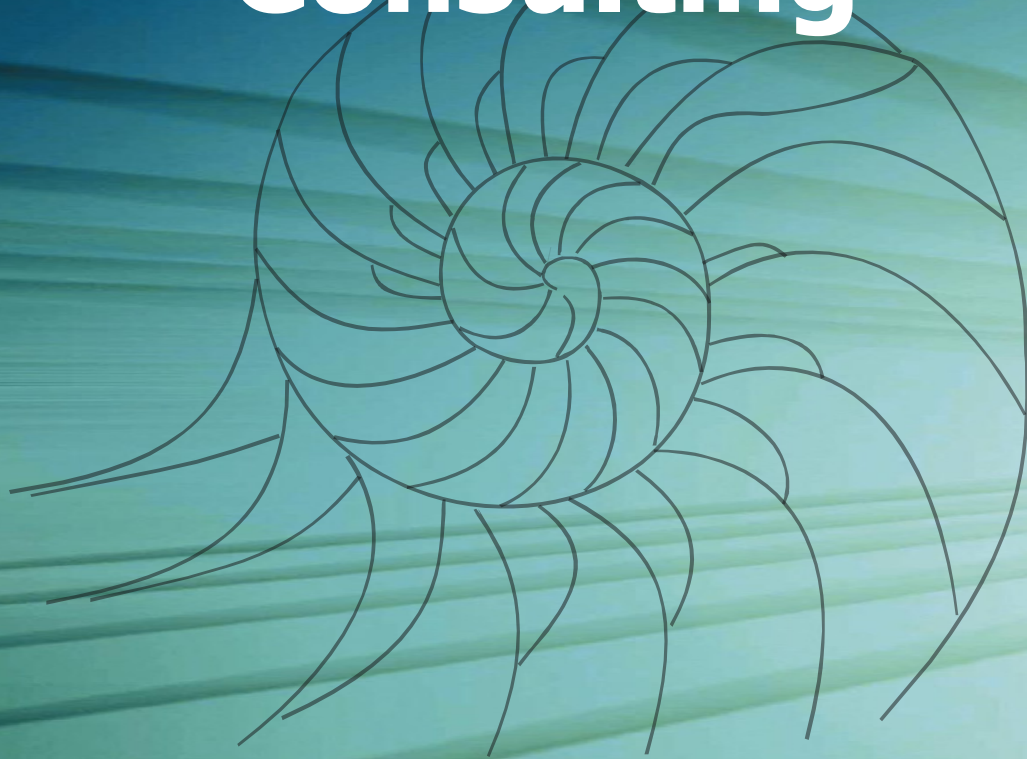
- Video records

## → Australian Standard for Risk Management





# Consulting



# Consulting

→ **Often considered equivalent**

**to**

→ Applied work for other professionals

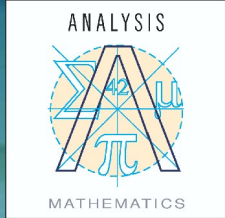
→ Scientific context

→ **But can be quite different**

→ Sometimes the client is not a traditional professional

→ In a group of professionals, the statistician is sometimes the appropriate leader

→ Often the scientific context is less formal



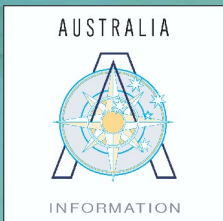
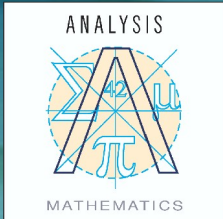
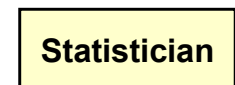
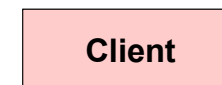
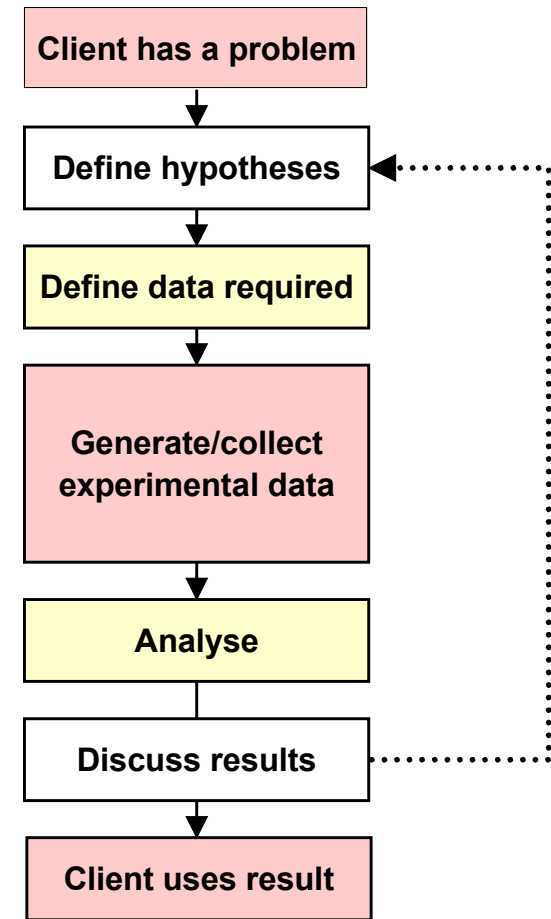
# Standard Consulting Model

## → Statistician's role limited

- The beginning and the end
- Formalising problem into hypotheses
- Defining data requirement
- Analysing data (in terms of hypotheses)
- Discuss results

## → Improvement cycle around experiment

- Can delay decisions till evidence clear



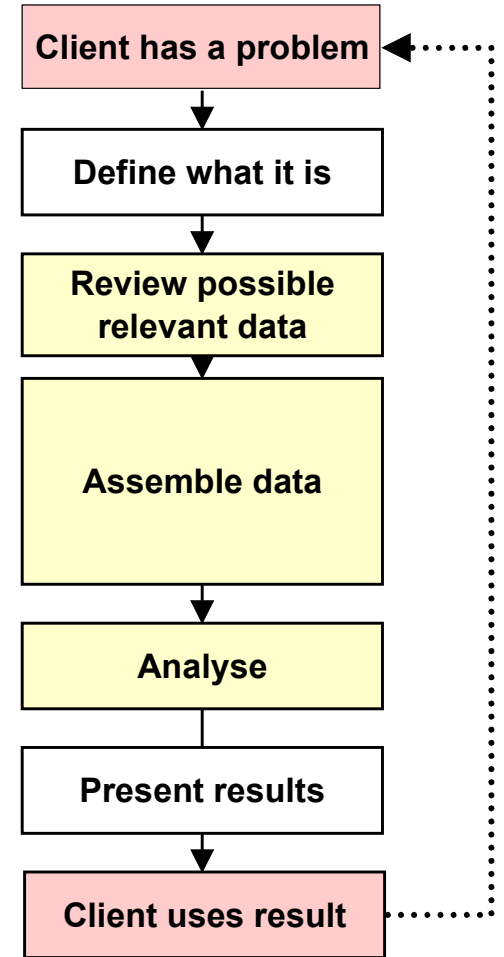
# "Real" Consulting Model

## → Statistician's role much greater

- Helping to define problem
- Using existing and new data
- Managing data
- Analysis and presentation

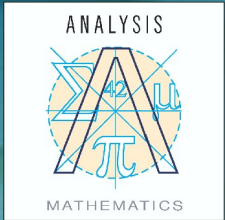
## → Improvement cycle includes client actions

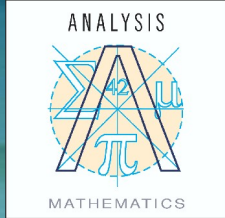
- Intermediate results must be usable



Client

Statistician



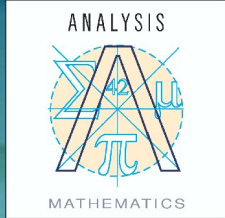


# Consulting is about Problems

---

---

- **Solving *other* people's problems**
  - They define the problem
    - Often not very well
- **Problem defines the solution set**
  - Some choice in solution, but not complete
  - Rarely obvious at the beginning
- **When accepting the problem, don't know what the solution will be like**
  - Cannot have pet solutions



# Quality is "Fit for Purpose"

---

---

→ **Best solution is simple**

→ But not too simple

→ "Everything should be made as simple as is possible but not more simple" – *Einstein*

→ **No value in sophistication for itself**

→ Value for a solution, not the method

→ **Solution must still be correct**

→ To a sensible accuracy

# Solution Might not be Where You Start

---

→ **Client does not know the boundaries we use**

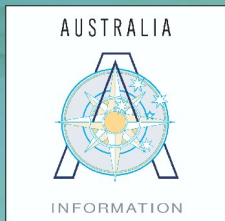
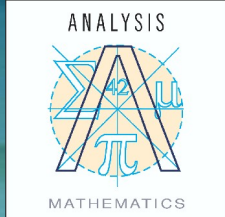
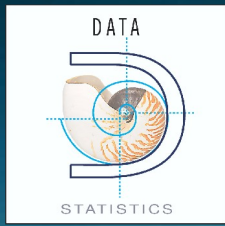
→ But usually it requires a mathematical solution

- Operations research
- Numerical methods
- Data management

→ Sometimes not

→ **Pressure to still give solution**

→ A mathematical/statistical training provides special skills for this





# A range of models

## → Scientific

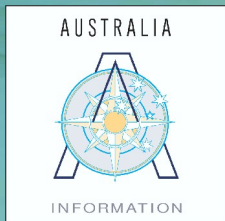
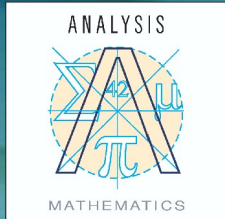
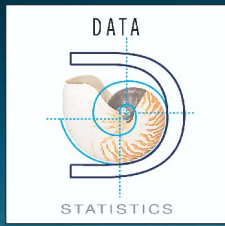
- Conservative
- Existing hypothesis
- Rarely considers costs
- Time less critical
- Statistician's role understood

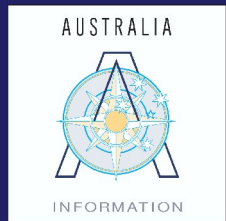
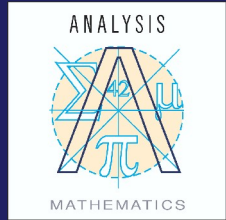
## → Commercial

- Optimistic environment
- Innovation favoured
- Costs critical
- Time critical
- Statistician has larger role

## → Neither model ever perfect

- Elements are applicable to most cases
- Some problems closer to one than the other
  - Medical trials closer to scientific





# Data Analysis Australia



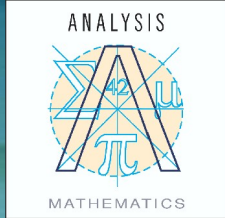
# Data Analysis Australia

## → My effort to do it right

- Planned while in England
- Started in September 1988
- Small at first – just me
- And fewer clients

## → First years getting the technical structure

- Corporate set up – forming a company
- Computing set up
- Accounting
- Image



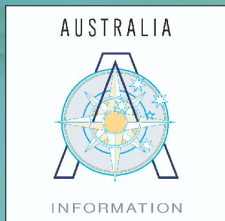
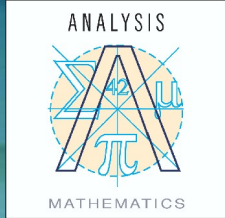
# Employing Staff

## → Usually recent graduates

- Usually honours +
- Look for mix of skills

- Statistics
- Computing Common sense
- Innovation
- Outgoing
- Writing

## → Learn to manage others



# Marketing - Broad

---

→ **Market not used to statisticians and mathematicians**

→ Who they are

→ How they can help

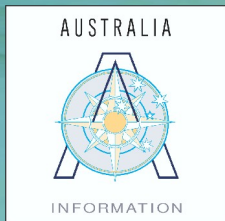
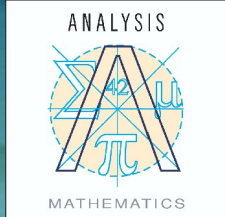
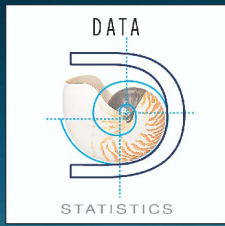
→ Not like marketing a commodity

→ **Constantly raising awareness**

→ Newsletter

→ Website

→ Seminars



# Marketing - Specific

## → Interactive

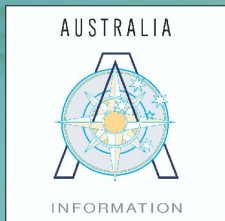
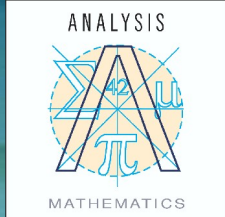
- Understanding client's problem
- Thinking up a solution that might work
- Costing
- Presenting it to the client

## → Problem is the starting point

- Not your favourite solution
- Solution might not even be mathematical
  - But a mathematician might be the best at finding it

## → Work with other consultants

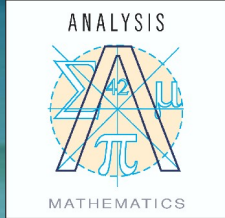
- Economists, social researchers, ...



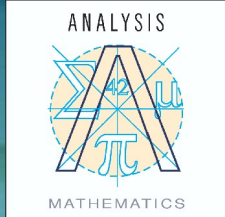
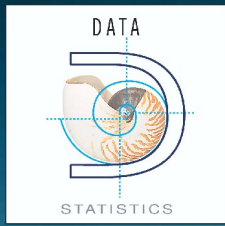
# Rewards

---

- Seeing important decisions better made
- Interesting clients / problems
  - And they like what you can do
- Learning to do more



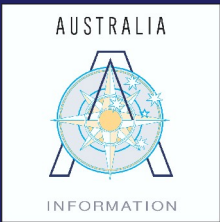
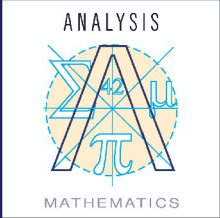
# Publications – Not the Focus



## → Unusual list of journals etc

- *Geophysics*
- *Statist. Soc. Australia News*
- *Proc. 7th Inter. Symp. Gastrointestinal Motility*
- *Discrete Mathematics*
- *COMPSTAT-80, Proceedings in computational statistics*
- *GLIM Newsletter*
- *Gastroenterology*
- *J. Endocrinology*
- *Australian Physical and Engineering Sciences in Medicine*
- *The American Statistician*
- *Proc. STATCOMP 83*
- *Computer Journal*
- *Biometrics*
- *Current Medical Research Opinion*
- *Essays in Time Series and Allied Processes*
- *European Journal of Appl. Physiol.*
- *Proc. Interface Symposium, 1988*
- *Economic and Labour Relations Review*
- *Australasian Transport Research Forum*
- *Gazette Australian Math. Soc*
- *J. Biogeography*
- *Sri Lankan Journal of Applied Statistics*





# Thank you

