





National Mathematics and Statistics Conference 2022

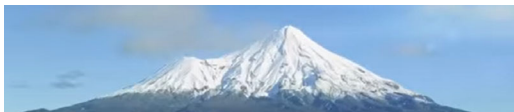
- 2½ years since I was invited to give this talk
- During that time some momentous things happened to me ...
 - I Turned 70 ! *(Arghhhh!)*
 - & retired fully
 - any “teaching” is looking after our 1-year-old granddaughter

“Big talks” have been quite a part of my life ...

KEYNOTES ITINERARY SINCE 2003 (on 5 continents)

ROYAL STATISTICAL SOCIETY (UK)
 INTERAMERICAN STATISTICAL INSTITUTE
 AUSTRALIAN STATISTICAL CONFERENCE
 STATISTICAL SOCIETY OF CANADA
 CANADIAN MATHEMATICAL SOCIETY
 INTERNATIONAL CONFERENCE ON THE TEACHING OF STATISTICS, ICOTS (*2)
 US CONFERENCE ON TEACHING STATISTICS, USCOTS (*3)
 OZCOTS (AUSTRALIA)
 INTERNATIONAL RESEARCH FORUM ON STATISTICAL REASONING, THINKING AND LITERACY (*3)
 IASE ROUNDTABLE CONFERENCE
 PRIESTMAN MEMORIAL LECTURES (NEW BRUNSWICK)
 THE BELZ LECTURE (VICTORIA)
 DELTA
 NZ STATISTICAL ASSOCIATION (*2)
 WATERLOO ANNIVERSARY CONFERENCE (CAN)
 NZAMT (*2)

BUT THIS IS MY LAST EVER “BIG TALK”



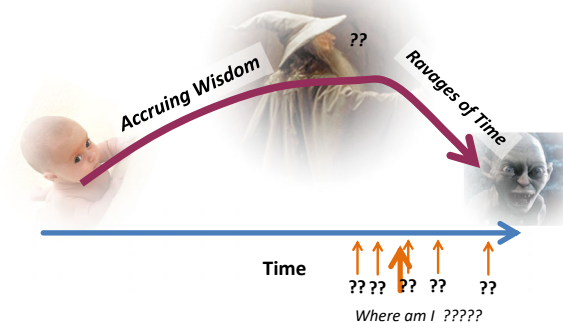
So, what to do in this last-ever talk after a 44-year career?

I must surely have learned something!!

- I should share my accumulated wisdom!
 - (after all, that’s what people do at a time like this)



But there’s a big risk, because as we age ...





So, what to do in this last-ever talk after a 44-year career?
I must surely have learned something!!

Now NZ is in the middle of a **curriculum refresh**

• So I'll include some ...

– guiding **fundamentals**

• & **things I care about** (and hope you will too)

– **forecasting**: from **trends** I see in the discipline, and in the wider society

• **punctuated with ...**

– Music clips (*this is the home of WOMAD!*)

– and a little self-indulgent mining of my “back catalogue”



• So I'll include some ...

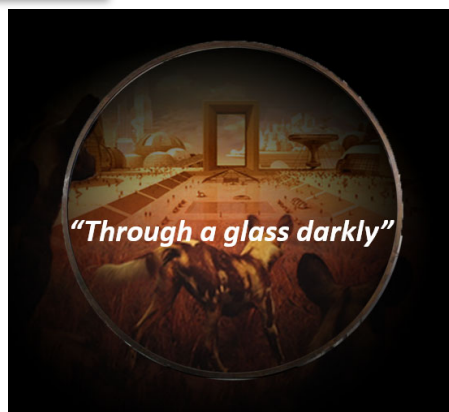
– guiding **fundamentals**

• & **things I care about** (and hope you will too)

– **forecasting**: from **trends** I see in the subject and wider society



Inevitably, this will all be ...



“Through a glass darkly”



Forecasting

“... projects patterns from the past into the future”
from “Statistical Literacy as the Earth Moves”

“Through a glass darkly”





**It's
Back to the Future**

Forecasting ...



*is like driving when all you can see is
in the rear-view mirror*

looking backwards to look forwards

So we shouldn't be surprised by the odd crash !



STARTING POINT 1

WHAT'RE SOME THINGS THAT HAVE BEEN HAPPENING RECENTLY?

- Data
- Dis- & mis-information
- Automation
 - From specialist to generalist
 - Democratization

A BRIEF HISTORY OF DATA

In the beginning....





The Good

Nate Silver *"Skills for a Lifetime"*
Commencement Address, 18 May 2018

... lots and lots of careers now in "data science" — a term that hardly even existed until a few years ago

when I graduated, people using statistics in fields like sports and politics and journalism were really on the outside looking in

Nowadays ... everyone is using data ... the "nerds" are no longer on the outside looking in. Instead, they're probably running the company

Power has shifted toward people and companies with a lot of proficiency in data science

THE UNIVERSITY OF AUCKLAND
DEPARTMENT OF STATISTICS

Slides and links at: <https://bit.ly/nzamt22>

NZAMT 2022

The explosion of data & it's impact on our lives

- brings **opportunities** for the appropriately educated
- But can also **accentuate** societal inequities
 - (between those **with** the resources to exploit the opportunities & those **without**)

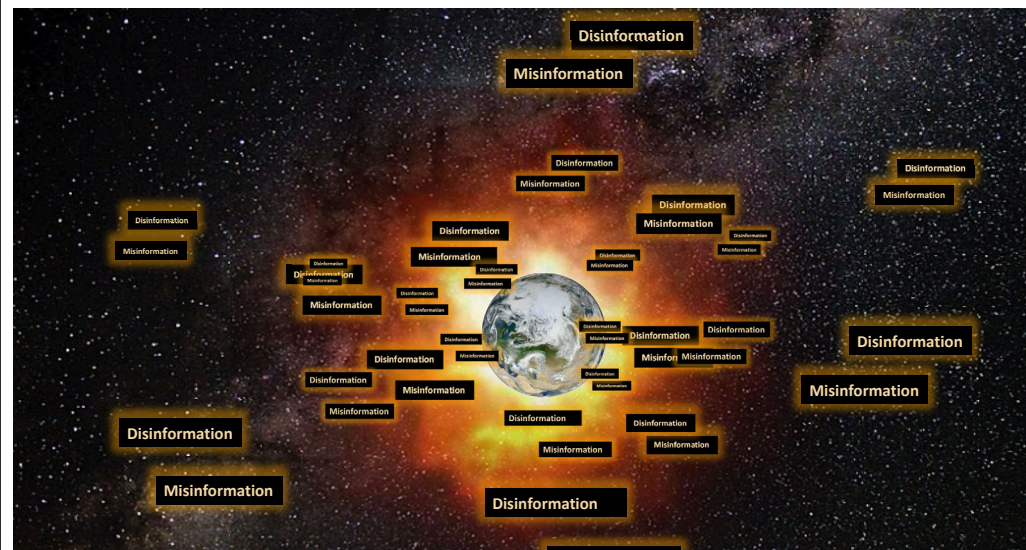
THE UNIVERSITY OF AUCKLAND
DEPARTMENT OF STATISTICS

Slides and links at: <https://bit.ly/nzamt22>

8

NZAMT 2022

But data isn't all that's been exploding



- Explosive spread through social media, accelerated by recommender algorithms, of
- **disinformation** (deliberate/malicious)
 - **misinformation** (in error, ignorance and misplaced trust)

A **hazard** impeding good decision making

What do we most want for our children/students?

Maximise opportunities (for good outcomes)

- We can foster ...
 - Awarenesses
 - Knowledge and skills
 - Imagination

What can we do about it?

Minimise hazards ...

(missing out on essentials that open up opportunities is an **opportunity-cost hazard**)

from false information

- We can foster ...
 - Awareness, recognition and critical thinking skills
 - Knowledge about characteristics of good and bad sources
 - Habitualise critical thinking

But something else has also been exploding ...

AUTOMATION

Hazards

& opportunities

Automation and the Future of Work ...

In the **first decade of the 21st century**, the United States **lost one-third** of all its **manufacturing jobs**, and the **vast majority** of them were **killed by automation**. They didn't 'go' anywhere. They just **vanished**.

Job destruction then slowed down until other new computer-driven technologies matured: self-driving vehicles, online shopping, 'dark' factories and warehouses. **But they are ready now**, and the **carnage in retail jobs, driving jobs and warehouse jobs is just getting underway**.

By 2022 it's now even affecting "creatives", e.g., taking work from actors who do voice work (e.g. ads, news reading, talking books, ...)

Technology will disrupt how we do almost everything
• and recast the *skills* people *will & will not* need

*If it can be done by machines ...
sooner or later it will be*

In the data world ...

Software is making more things easier for more people

- Everywhere software people are trying to do this
 - (free software) The whole R-package culture is about this
 - (commercial software) The driver is saleable products
- **Automation**, esp involving Machine Learning/Artificial Intelligence
 - bigger and more powerful black boxes (has some huge downsides)

Consequences for our students' lives and careers are enormous

"Today's programming job is tomorrow's mouse-click"

Complex programming job

Progression ...

Function/package calls

Mouse click

Voice commands
Full Automation



Your AI skills are worth less than you think" Data Science/Statistics/Maths

Your AI skills are worth less than you think



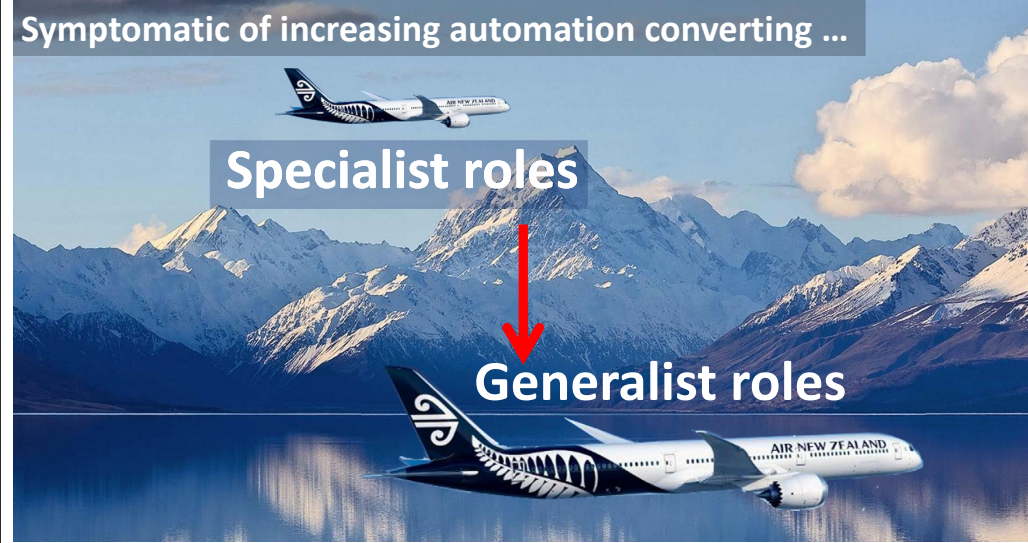
<https://www.kdnuggets.com/2019/01/your-ai-skills-worth-less-than-you-think.html>

As exciting as the progress is, it's **bad news** for both companies and **individuals who have invested heavily in AI skills**.

Today, they give you a solid competitive advantage, as training a competent ML engineer requires plenty of time spent reading papers, and a solid math background to start with.

However, **as the tools get better, this won't be the case anymore**. It'll become more about reading tutorials than scientific papers.

If you don't realize your advantage soon, a band of interns with a library may eat your lunch.



International Data Science in Schools Project

Curriculum Frameworks for

Introductory Data Science

IDSSP Curriculum Team

September 2019

(for the last 2 years of high school)

Endorsements of Frameworks (so far)



AMERICAN STATISTICAL ASSOCIATION
Promoting the Practice and Profession of Statistics

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS



The New Zealand
Statistical Association

Statistical Society of
Australia



International Data Science in Schools Project (IDSSP)

DS Curriculum Topics not in std intro stat

- Data-handling Pipeline
- Dynamic visualization, multiple variables
- Interactive visualization
- Inference via bootstrapping & randomisation
- Time series data
- **Map data** (data relating to geographic locations)
- **Text data** (from pieces of text – tweets, comments, articles, books, ...)
- **Machine Learning: Supervised** ("predictive analytics")
- **Machine Learning: Unsupervised** (clustering "like" with "like")
- **Image Data** (from images)
- **Recommender systems**
- **Network data** (coming)

Including a significant emphasis on code writing



International Data Science in Schools Project (IDSSP)

DS Curriculum Topics not in std intro stat

- Data-handling Pipeline
- Dynamic visualization, multiple variables
- Interactive visualization
- Inference via bootstrapping & randomisation
- Time series data
- Map data
- Text data
- Machine Learning: Supervised
- Machine Learning: Unsupervised
- Image Data
- Recommender systems
- Network data



- Data-handling Pipeline
- Dynamic visualization, multiple variables
- Interactive visualization
- Inference via bootstrapping & randomisation
- Time series data
- Map data
- Text data
- Machine Learning: Supervised
- Machine Learning: Unsupervised
- Image Data
- Recommender systems
- Network data (*coming*)



What makes this achievable is software that automates the details

I've done enough work on these, and software prototypes for them, to see that some **serious practical capabilities** in all of these areas **can be within the grasp of high school students**

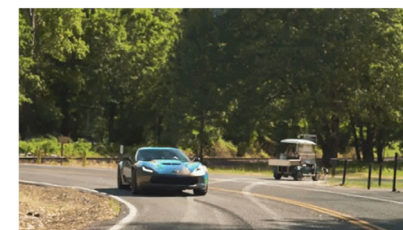


Additionally, because of the tendency in society for ...

Specialist roles

Generalist roles

Increasing numbers of areas
are becoming "just like driving a car"



- And for them ...



... Society needs ...
lots more people ...

- to be good at this
- than this



And you can be
very good at this

Without being
much good at this

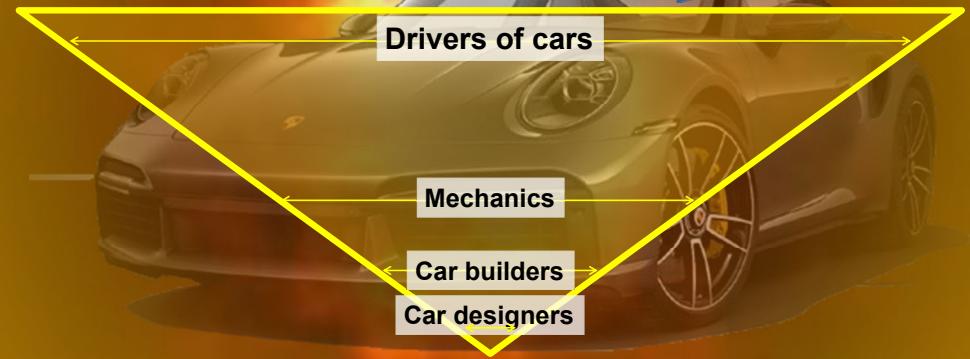
& vice versa!



Increasing numbers of areas
are becoming just like driving a car

There's a pattern ...

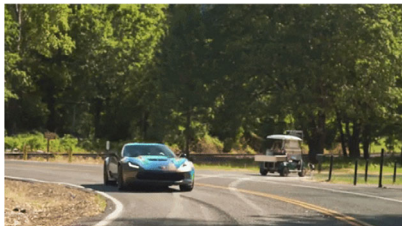
A qualitative impression of
the numbers of people involved



Educationally, we do a lot of ...



And little, if any, of ...



We aren't as bad as others internationally
• but have stumbled on new ways of bogging
things down in quicksand

Us



&

Great!!
But, 1st you'll have to learn
to use this ...

Them

I want to learn to drive
I want to see the Amalfi Coast



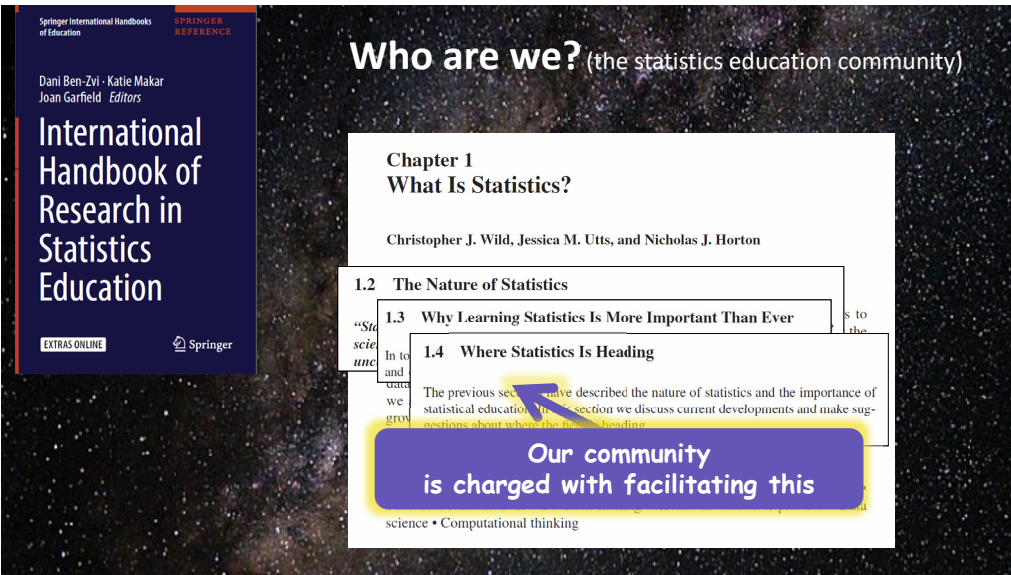
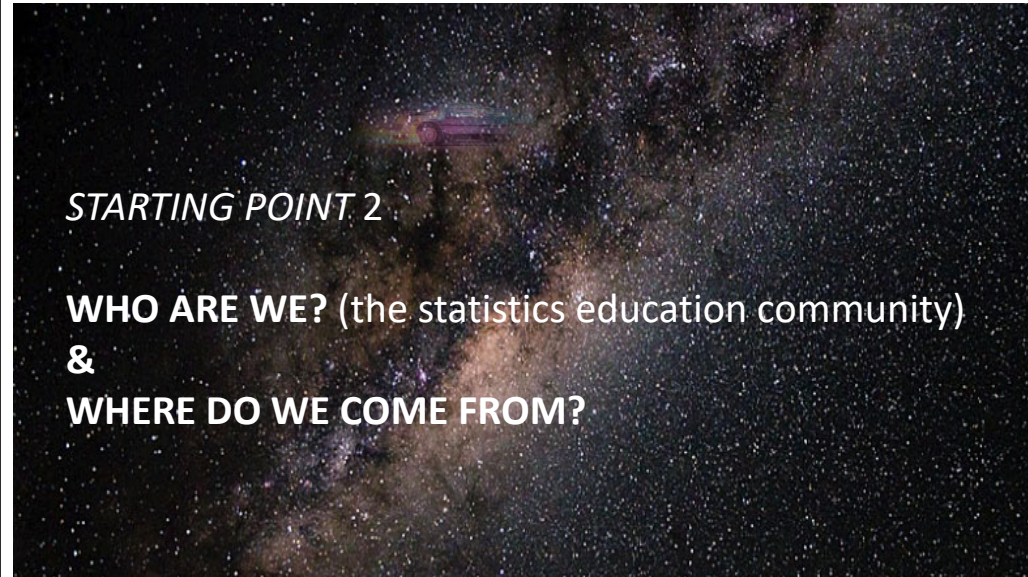


**Big consequences of increasing automation
for students' lives and careers ...**



***What do we need for future-proofing
our children/students?***

The ability to do things machines can't do!



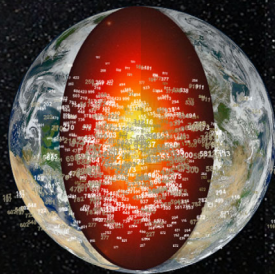
Preparing students for

- Data capabilities for life and work
 - Understanding data-based displays and arguments
 - Investigative skills
 - Capabilities in data harvesting, data wrangling and data analysis
 - Ethical behaviour in collection, storage and use of data
- Protecting themselves from
 - the deluge of disinformation & misinformation
 - Scepticism and critique

Who are we?

"Statistics is the science of learning from data, and of measuring, controlling and communicating uncertainty." — American Statistical Association

- The **focus** of statistics is:
 - understanding the world through data
- The **raw materials** for statistics are:
 - real-world questions and data
- The **tools** of statistics are
 - statistical ways of thinking & investigating
 - & computer software
 - » often based on mathematical models and derivations



Who are we?

Understanding the world through data
Statistics

Computing

Mathematics

motivation

motivation

tools

tools

Main roll in statistics

- Algorithmic ways of thinking
 - Programming new capabilities
- Anything generically useful gets "packaged in software" so you don't need to programme it any more*

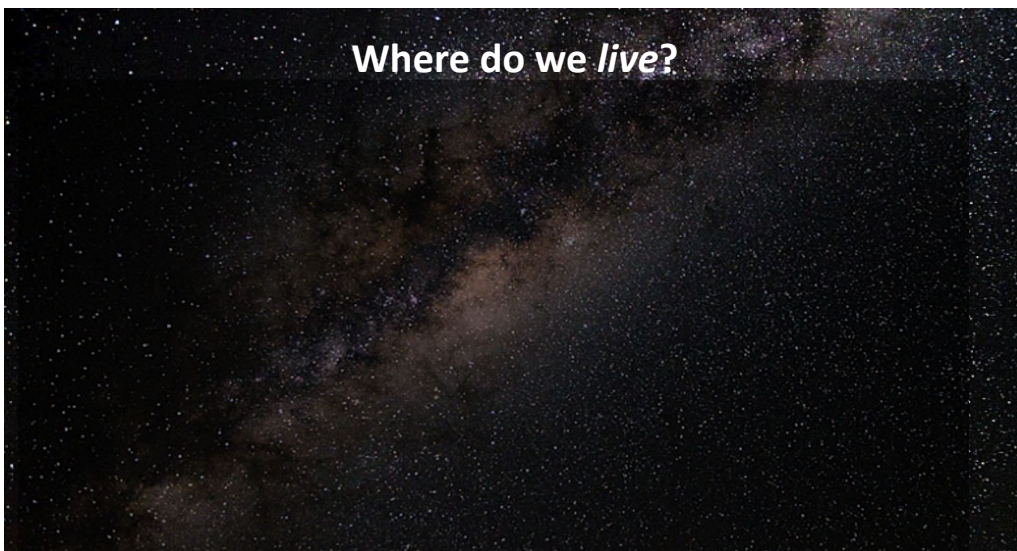
Main role in statistics

- Mathematical ways of thinking
- Algebraic & calculus skills
- Used for developing new models & methodologies

Doing stuff that available software doesn't already do



Where do we live?



Where do we live?

In the galaxy of forming views about how the world around us operates ...



Where do we live?

In the galaxy of forming views about how the world around us operates ...



Where do we live?

In the galaxy of forming views about how the world around us operates ...



statistical inquiry is a tectonic plate on the planet **"Purposefully Finding Out"**

From "On Locating Statistics in the World of Finding Out"



Where do we live?

We share our planet (*Purposefully Finding Out*) with other tectonic plates

Some of our neighbours ...

- consulting trusted others (people, articles, books, websites, ...)
- **qualitative research methods**
- data mining
- **machine learning**



From "On Locating Statistics in the World of Finding Out"



Where do we live?



Leverage the *ability of humans to notice things you'd never think to measure*



Where is statistics going?

- “Data science”
- Stochastic modelling of real-world processes
(cf. covid)
- Qualitative methods



Where is the world going?

- Automation



Nate Silver

“Skills for a Lifetime”

Commencement Address, 18 May 2018

Worries

Companies and governments that are capable of using data in powerful ways are also capable of abusing it

What worries me the most ... is the idea that using data science allows one to remove human judgment from the equation

... “machine learning” appeals to people’s notion of a push-button solution ... the computer does all your thinking for you, no human judgment required

a lot of supposedly “objective,” data-driven algorithms to determine criminal sentences can encode and reinforce racial bias based on how variables are chosen



The Ugly?

Nate Silver

“Skills for a Lifetime”

Commencement Address, 18 May 2018

the reality is that working with data requires lots of judgment

it requires moral judgment in deciding what your goals are and in establishing boundaries for your work

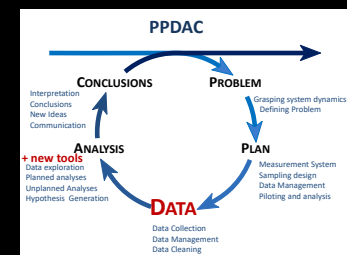
the correct interpretation of the data is rarely obvious, and that the obvious-seeming interpretation isn’t always correct

Sometimes changing a single assumption or a single line of code can radically change your conclusion



Where is “data science” sending us?

- Much more emphasis on ...
 - Harvesting & assembly of data from electronic sources
 - e.g. databases & websites via apis or scraping
 - Data wrangling (whipping data from a variety of sources into shape for analysis)
 - New data types (e.g. “big data”, text, images, sound files)
 - More visualisation, CS paradigms for prediction, ...
- New emphasis on coding (programming)



Who do we serve?

- Intending statistical/data science specialists
- Intending specialists in other areas that need these skills
- **the wider society ...**
 - by spreading valuable data-&information messages & skills as widely as possible, including ...
 - protecting from dis- & misinformation
 - & what it means to be an ethical citizen of the data world



Let's return to this ...

Adding in

& worry about ...

The explosion of data & its impact on our lives

- brings **opportunities** for the appropriately educated
 - **or those who can afford to buy in data and analytics**
- But could also **accentuate societal inequities**
 - (between those that have the resources to exploit opportunities & those who don't)

Harvard Data Science Review • Issue 3.2, Spring 2021

On Democratizing Data Science: Some iNZights Into Empowering the Many

Chris J. Wild¹, Tom Elliott¹, Andrew Sporle¹

¹University of Auckland, New Zealand

Democratizing Data



“Democratizing data” is ...

“empowering the many”



Let's return to this ...

Adding in

& emphasize

The explosion of data & its impact on our lives

- brings **opportunities** for the appropriately educated
 - **or those who can afford to buy in data and analytics**
- But could also **accentuate societal inequities**
 - (between those that have the resources to exploit opportunities & those who don't)

Harvard Data Science Review • Issue 3.2, Spring 2021

On Democratizing Data Science: Some iNZights Into Empowering the Many

Chris J. Wild¹, Tom Elliott¹, Andrew Sporle¹

¹University of Auckland, New Zealand

Introducing Andrew Sporle

- Social, health and official statistics researcher
- Dep. Director of NZ Healthier Lives National Science Challenge
- Board of Int. Sociological Ass. Res. Com. on Racism, Nationalism, Indigeneity and Ethnicity
- founding member of NZ Virtual Health Information Network
- founding member of Te Mana Raraunga (Māori Data Sovereignty Network)
- was inaugural Māori Health Research Manager at the Health Research Council of NZ

Important Keynote at Auckland Statistics Teachers Day 2020
(Here for Video & slides)



Andrew Sporle





Andrew Sporle

Andrew's strong Work Relationships ...

include ...

- Statistical & Health Agencies of small Pacific nations
- Indigenous & community groups,
 - Incl. Māori Data Sovereignty Network & other national movements linked by the Global Indigenous Data Alliance
- Subject-matter researchers (esp. Health, Social Science & Official Statistics)

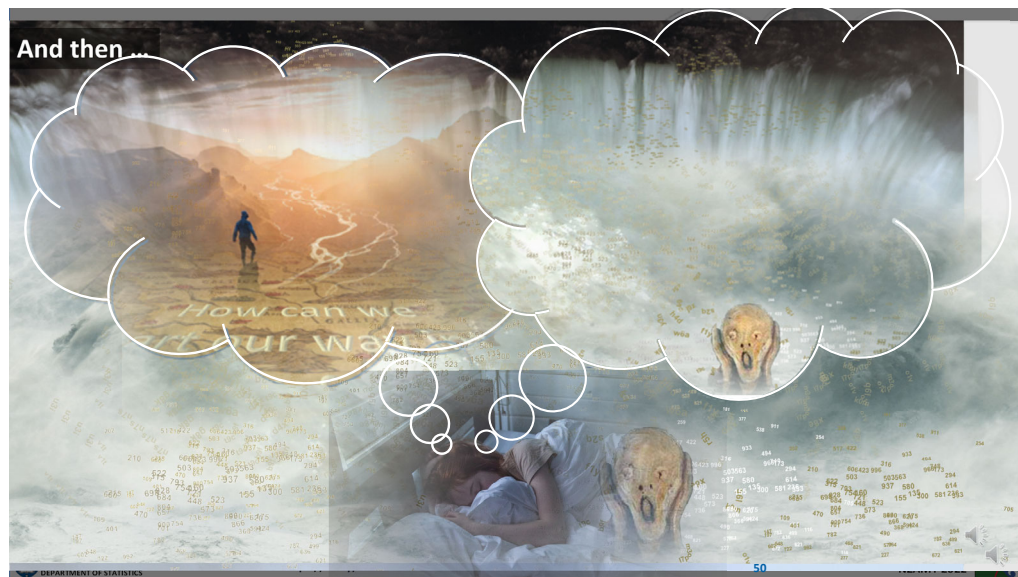
overwhelmed by:

- needs at home
- demands of UN agencies

COMMON REFRAINS

- Big statistical & data science needs, but ...
 - No People
 - No Money
 - No Time
 - Or, at least, very little of these things

49



Greta Thunberg ...

*"While we may all be in the same storm,
we are not all in the same boat"*

Greta Thunberg ... *"While we may all be in the same storm,
we are not all in the same boat"*

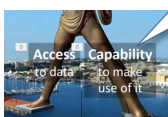
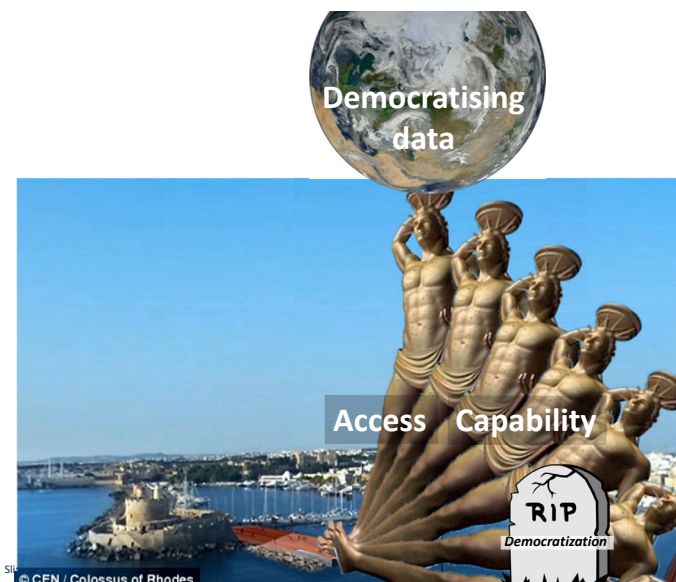
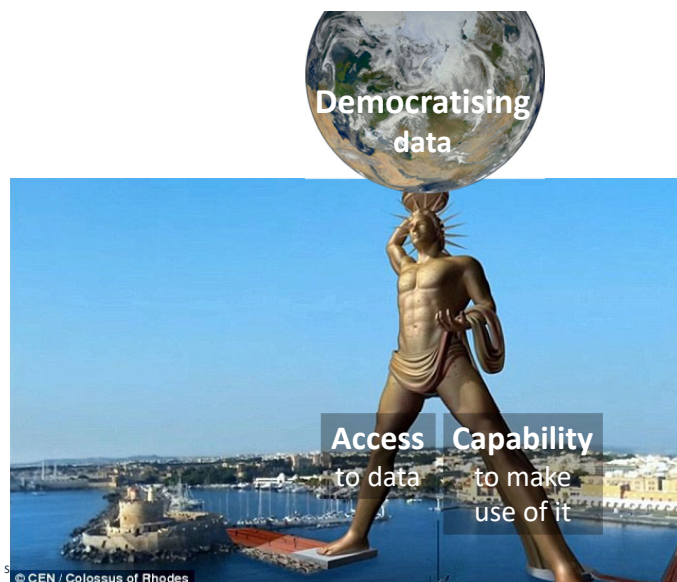


Harvard Data Science Review • Issue 3.3, Spring 2021
**On Democratizing Data
Science: Some iNZights
Into Empowering the Many**
Chris J. Wild*, Tom Elliott*, Andrew Sporle*
*University of Auckland, New Zealand

- **Unmet needs of those lacking in money** and data education are every bit **as important as those who have more**
- Volunteers like **Statistics Without Borders** (pro bono services)
 - as a solution, it **cannot scale** – will never be enough volunteers
- Need to **empower more people to do more for themselves**

cf. "doctors without borders"





Democratising Data/Data Science/Analytics



Enterprise Research Service

Research Brief

The Democratization of Analytics

October 2015

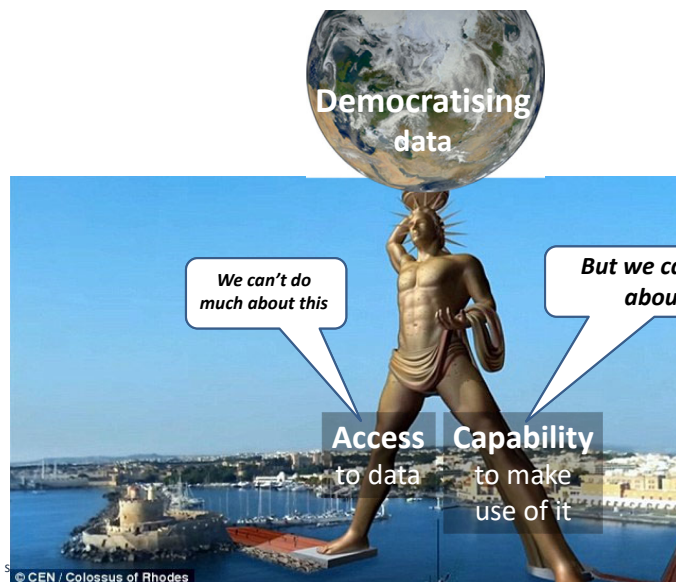
Written by:

Bernard Blais, Senior Manager, Global Technology Practice, SAS

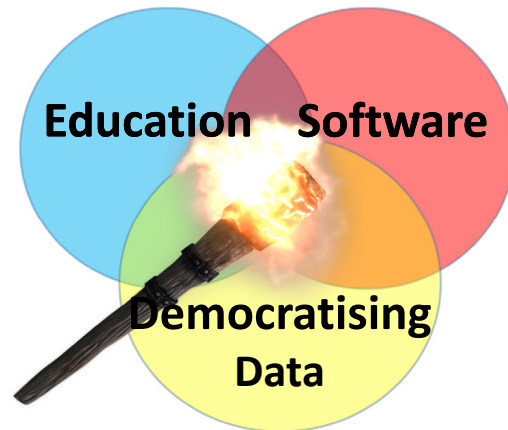
Robert Morrison, BA Faculty

*"Enabling people to access data is one thing.
But empowering them to do something intelligent
with the data, that's democratization of analytics"*

**We need to democratize data skills
to "empower the many" !**



What I care about



So how can we empower the many?



- Education!!!!

- That's us ... *isn't it?*
- We've got this ... *haven't we?*
- *Sorted!!*



– *Why not ????*

Especially for the under-resourced, we need to transfer ...

Specialist skills to



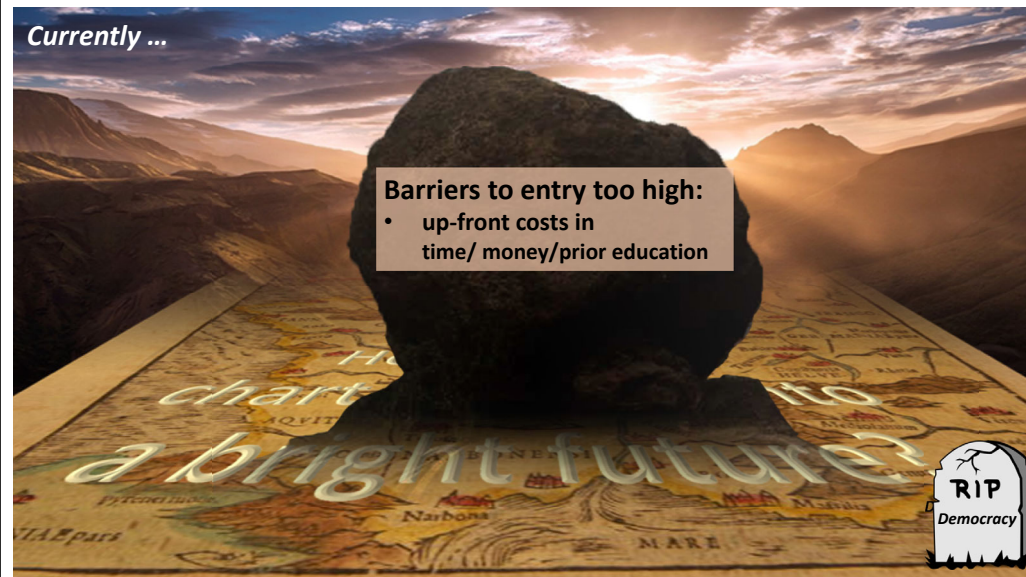
Generalists

empower busy generalists to do things that now *only* specialists can do

Currently ...

Barriers to entry too high:

- up-front costs in time/ money/prior education



Problems for generalists who become “part-time Statisticians”

- little data-related education

“So give them more education!!”

- time-poor

Catch 22!! Gotcha!

– so acquiring knowledge is strangling bottleneck

- And generalists:

– do many different things

– often work in processes infrequently

→ rapidly fading memories of how to do things
and what they mean

Lack of up-front knowledge & fading memories major factors to work around

Educational models

Bottom up

Usual model for statistics education ...

Useful component for ...

- ability to build new things & depth (requires low-level tools)

But ...

way too slow!



Inefficient for getting...

- broad conceptions of possibilities
 - many practical capabilities quickly
- Hopeless for “Motivating the masses”

Our dominant educational model is akin to building something big using Lego bricks – start with learning about small things and then, gradually, about assembling them into something bigger

Educational models

Top Down

Democratization needs top-down (high-level) ...

- awareness of what’s out there
- thinking
- interactions with software

- We glide high above the earth looking down
- We get a rough idea of what’s there - seeing things in relation to one another
- Occasionally we think “that’s interesting” and swoop down for a closer look at detail

For democratization

- We need education & software interfaces that take account of these realities ...

- too little time to learn “all you’ll need”
- rapidly fading human memories (what’s been “learned” mostly gets forgotten)
- increasing automation

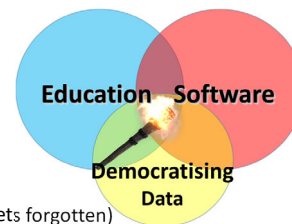
- Leads to trying to do the following ...

– Answer, “What is the *minimal knowledge* someone needs for this?”

– Enable working from that with intelligent GUIs that ...

- Automate decisions whenever good defaults exist
 - But allow users to override defaults (and do things differently)
- Guide them through thinking/decision processes supplemented by just-in-time information

– Move cognitive & human-memory load from non-essentials to understanding what data is saying and the implications of that for action



Graphical User Interfaces

For democratization

- We need education & software interfaces that take account of these realities ...
 - *too little time to learn* “all you’ll need”
 - *rapidly fading* human *memories* (what’s “learned” mostly gets forgotten)
 - *increasing automation*
- Educationally we need internalised big-picture conceptions as scaffoldings to hang on ...
 - what things are
 - how they interrelate
 - what sorts of things can be done / what things can do
 - what things mean
- Else, post formal-education, there will be a dearth of relevant memories to act as starting points for approaching practical problems



For this ...

- Big-picture conceptions ... have long-term value



- Details are death dated



??



Question: “... You must remember this...”
If only 5% of what you “teach” is going to stick long-term, what should that 5% be??

“... the fundamental things apply as time goes by ...”



*For most, I believe,
the future of understanding data ...*

*For most, I believe,
the future of understanding data ... is visual*



"The problem of names"



In most statistical systems ...

- can't do anything 'til you know/remember, its name
 - **Barrier to getting started**
 - Big **time-losses** getting back up to speed **after a period of inactivity**



Slides and links at: <https://bit.ly/nzamt22>

NZAMT 2022



My efforts towards solving these problems ...

has mainly been in the iNZight project <https://inzight.nz>

iNZight Desktop software

iNZight lite Online equivalent

Very useful systems

+ research vehicle for prototyping new ideas

- often using grad students and their Masters and Hons dissertations to create prototypes for new data areas

THE UNIVERSITY OF AUCKLAND
DEPARTMENT OF STATISTICS

Slides and links at: <https://bit.ly/nzamt22>

70

NZAMT 2022



Let's return to this slide

- and then talk about some development that's been done

International Data Science in Schools Project (IDSSP)

DS Curriculum Topics not in std intro stat

- Data-handling Pipeline
- Dynamic visualization, multiple variables
- Interactive visualization
- Inference via bootstrapping & randomisation
- Time series data
- Map data
- Text data
- Machine Learning: Supervised
- Machine Learning: Unsupervised
- Image Data
- Recommender systems
- Network data



THE UNIVERSITY OF AUCKLAND
DEPARTMENT OF STATISTICS

Slides and links at: <https://bit.ly/nzamt22>

68

NZAMT 2022



Specialist modules

iNZight & **iNZight lite**

Topics in Int. Data Science in Schools Project that are not in std intro stat

- Data-handling Pipeline (*heaps of useful tools*)
- Time series data (*"done"*)
- Map data (*"done"*)
- Text data (*ready to go*)
- Machine Learning: Supervised (*prototyped*)
- Machine Learning: Unsupervised (*prototyped*)
- Interactive visualization (*interactive versions of most plots already implemented*)
- Network data (*prototyped*)
- Image Data (*experimenting*)
- Recommender systems (*probably won't do*)

PLUS OTHER "ADVANCED" CAPABILITIES

- Extensive data-wrangling capabilities (*"done" & writing R code*)
- Generalized Linear Models (*"done" & writing R code*) ^{& survival analysis}
- Data from complex surveys (*"done" and integrated*)
- Data from designed experiments (*"done", Lite only*)
- Multivariate- y data (*"done"*)
- Multiple-response data (*"done"*)
- Epidemiological tools (*"done"*)
- Longitudinal data (*prototyped*)
- Hierarchical data (*innovative graphics developed*)

THE UNIVERSITY OF AUCKLAND
DEPARTMENT OF STATISTICS

Slides and links at: <https://bit.ly/nzamt22>

72

NZAMT 2022



Other developments

New capabilities to cater for ...

- **Integrating R coding** and GUI-driven operations, **system-written R-code** (to help with learning to code, providing audit trails, and supplying code blocks for use in programs)
- **Becoming the primary analysis tool for the researchers in ...**
 - The *Growing Up in New Zealand* longitudinal study (2004 – ...)
 - And some other major NZ studies
- **Time series rewritten** to have **professional-level capabilities** (back-end completed)
 - & also work on **small area estimation**
 - esp. to cater for **small Pacific countries** etc
- **Requests by**
 - the Australian Bureau of Statistics
 - Stats NZ in preparation for running in their Data Labs

"Growing Up in New Zealand is this country's largest contemporary longitudinal study of child development. We're tracking the lives of more than 6,000 Kiwi children to discover what life is like growing up in 21st Century New Zealand"

"Give me the R code for what I've just done"



Specialist modules

Topics in IDSSP that are not in std intro stat

- Data-handling Pipeline (many useful tools)
- Time series data ("done")
- Map data ("done")
- Text data (ready to go)
- Machine Learning: Supervised (prototyped)
- Machine Learning: Unsupervised (prototyped)
- Interactive visualization (interactive versions of most plots implemented)
- Network data (prototyped)
- Image Data (experimenting)
- Recommender systems (probably won't do)

PLUS OTHER "ADVANCED" CAPABILITIES

- Extensive data-wrangling capabilities ("done" & writing R code)
- Generalized Linear Models ("done" & writing R code)
- Data from complex surveys ("done" and integrated)
- Data from designed experiments ("done", Lite only)
- Multivariate-y data ("done")
- Multiple-response data ("done")
- Epidemiological tools ("done")
- Hierarchical data (progress on innovative graphics)
- Longitudinal data (prototyped)

Late breaking: Bayesian small-area estimation (aimed at small countries) John Bryant & Tom Elliot

Slides and links at <https://bit.ly/nzamt22>

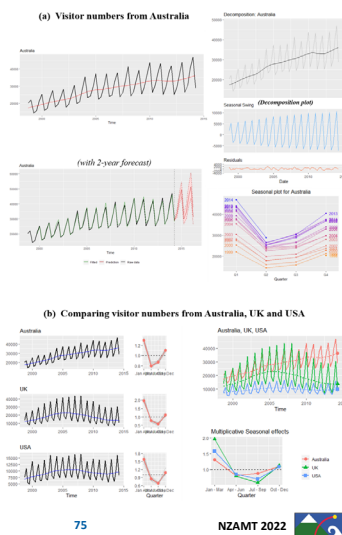
No time to talk about any of these areas in detail

- Instead I'll use what we've *already experienced* in time series as an example of *what can happen* more generally



Lessons from Time Series

- The experience **used to be mainly about ...**
 - **Calculating moving averages**, essentially by hand, in a spreadsheet
 - Low value for revealing what a time series is telling us
- **Starting about 10 years ago ...**
 - a big-picture conceptualising of time series
 - in terms of *trend + seasonal swings + residuals*
 - and graphical tools that can
 - reveal the patterns, and anomalies, in seasonal time series
 - and make predictions
 - providing powerful tools for data exploration and discovery
 - experience is about "what the data is telling us"
 - with graphics that are understandable at senior secondary level



Specialist modules

Topics in IDSSP that are not in std intro stat

- Data-handling Pipeline (many useful tools)
- Time series data ("done")
- Map data ("done")
- Text data (ready to go)
- Machine Learning: Supervised (prototyped)
- Machine Learning: Unsupervised (prototyped)
- Interactive visualization (interactive versions of most plots implemented)
- Network data (prototyped)
- Image Data (experimenting)
- Recommender systems (probably won't do)

PLUS OTHER "ADVANCED" CAPABILITIES

- Extensive data-wrangling capabilities ("done" & writing R code)
- Generalized Linear Models ("done" & writing R code)
- Data from complex surveys ("done" and integrated)
- Data from designed experiments ("done", Lite only)
- Multivariate-y data ("done")
- Multiple-response data ("done")
- Epidemiological tools ("done")
- Hierarchical data (progress on innovative graphics)
- Longitudinal data (prototyped)

Late breaking: Bayesian small-area estimation (aimed at small countries) John Bryant & Tom Elliot

Slides and links at <https://bit.ly/nzamt22>

- This work has revealed that the same sort of thing ...
 - **can be done for almost all** of these areas ...
 - each with **their own big-picture ways of thinking** about the data (cf. *trend + seasonal swings + residuals*) and their **own distinctive types of graphics**
 - conferring **serious real-world capabilities**
 - in ways that are **accessible to senior secondary students**



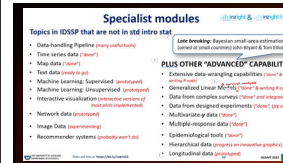
Lessons from Time Series *cont.*

We have a proclivity for

- taking something *potentially exciting* ...
- & making it akin to *wading through a swamp*

Nice if we could also learn from our mistakes ...

- Assessment focussed *on writing* about *one time series*
 - that then also had to have “all the features” so students could score marks
 - an extended, & regimented, experience of “one thing”
- This is **not** what learning for exploration and discovery looks like ... It needs ...
 - Exploration of many data sets ...
 - spotting different types of features (& wrinkles) in situ
 - thinking about what they might be telling us about the world
- Excitement can build from repeatedly making discoveries
 - in a short timespan (*because of short student attention-spans*)
 - without too much effort (*few have patience for prolonged, concentrated, unrewarded effort*)

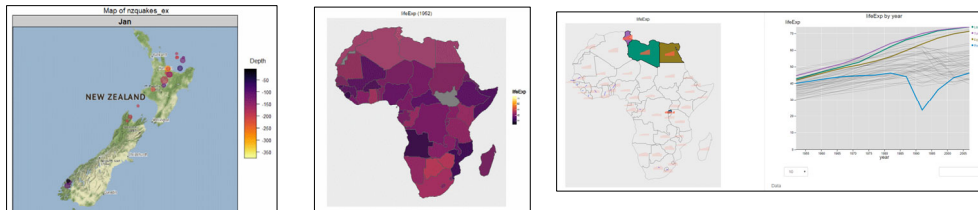


The data world ...
has gotten a whole lot bigger

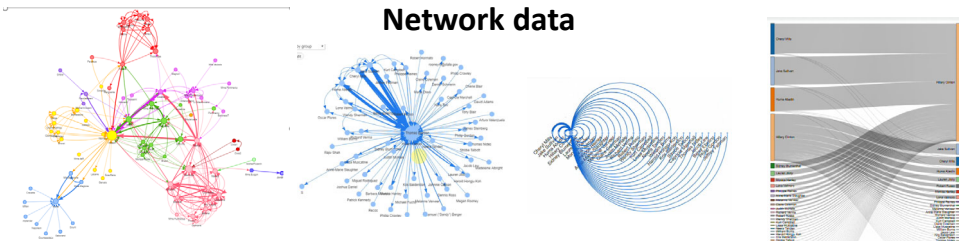
Can't just keep illuminating same small patch

- Would impose huge *opportunity costs* on our students

Data on maps



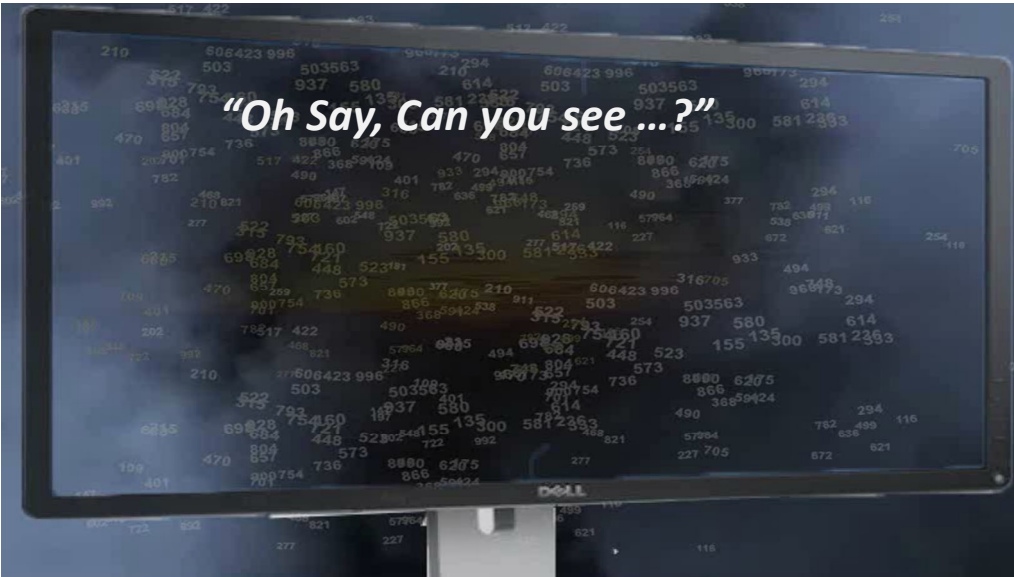
Network data



Seeking the holy grail ...



For most, I believe,
the future of understanding data ... *is visual*

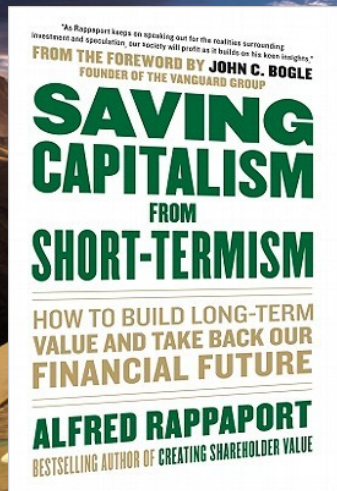


BEST BEFORE
31-12-2025

BEST BEFORE
31-12-2025

BEST BEFORE
31-12-2025





What is short-termism?

The term refers to an excessive focus on short-term results at the expense of long-term interests. When discussing the causes of the financial crisis, a related term was used by columnist Eric Dash of the *New York Times*: the I-B-G-Y-B-G syndrome, "I'll be gone; you'll be gone" before anyone will have to answer for the toxic mortgage securities building up on bank balance sheets.

focus on short-term results at the expense of long-term interests.

Terrible, isn't it ???!!



Students and short termism

Focus on "What do I need to get an Achieved/Excellence?"

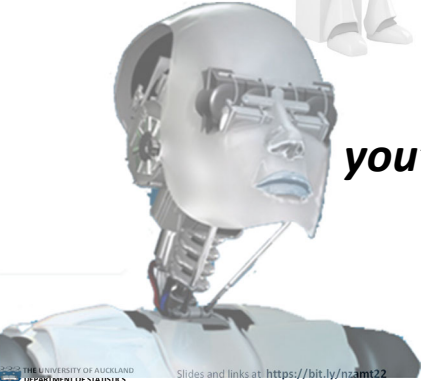
Students and short termism

- Focus on "What do I need to get an Achieved/Excellence?" And when they want help, you know they're thinking ...
- "Just give me the answer!"
— or at least a simple recipe!
- Anything but tell me "I actually have to think"!

Terrible, isn't it ???!!



“If you can only think like a robot ...



you’ll be replaced by one!”

– Andrew Balemi

Partly a misplaced idea of the role of qualifications

Qualifications vs capabilities

Short-term Goal
(foot in the door)

Qualifications

What you can have ...

Capabilities

What you can do ...

Long-term goal

SUCCESS!

THE UNIVERSITY OF AUCKLAND
DEPARTMENT OF STATISTICS
Slides and links at: <https://bit.ly/nzamt22>

NZAMT 2022

But actually, It’s not just students !

- **We’re all short-termists too**
 - The reward systems we work under reinforce that
 - *Get it done*
 - *Not-too-bad evaluations*
 - *Go on to something that counts more*
- **But that shouldn’t entirely define us ...**
 - “the greater good” often comes from considering the longer term



But this talk is **trying to apply some long-term antidote to our habitual short-termism**



*Not either/or
It's a balancing act*

The Future of Work



Remembering that ...
*"if it can be done by machines ...
sooner or later it will be"*

The Future of Work ...



Theorem:

- Just about the only constant will be change
 - and the *pace of change is accelerating*
 - *"the only constant is change"* – Heraclitus, 500 BC

Corollary:

- **Everyone** will have to be **a life-long learner**

The Future of Work ...



We've seen this ...



- **What do we want for our students?**

The ability to do things machines can't do!



For this ...



and this ...

- Big-picture conceptions ...

have long-term value



- Details are death dated



NZAMT 2022

Continuing in this vein ...

The Future of Work



The ability to do ...

Mathematical derivations ...

has long-term value



The ability to do ...

any particular mathematical derivations

...

is death dated



The skills' value comes from the ability to
produce *new theory* and *new methods*

The Future of Work



The ability to operate ...

a *recipe*
procedure ... has long-term value
algorithm



The ability to operate ...

any particular recipe
procedure ... is death dated
algorithm



NZAMT 2022

The Future of Work



The ability to ...

code (program) ...

has long-term value



The ability to ...

code (program) ...

anything particular (in any particular language)

is death dated



The skills' value comes from the ability to
code things *not yet catered for*
by existing and accessible *software*

The Future of Work



But, the ability to devise new ...

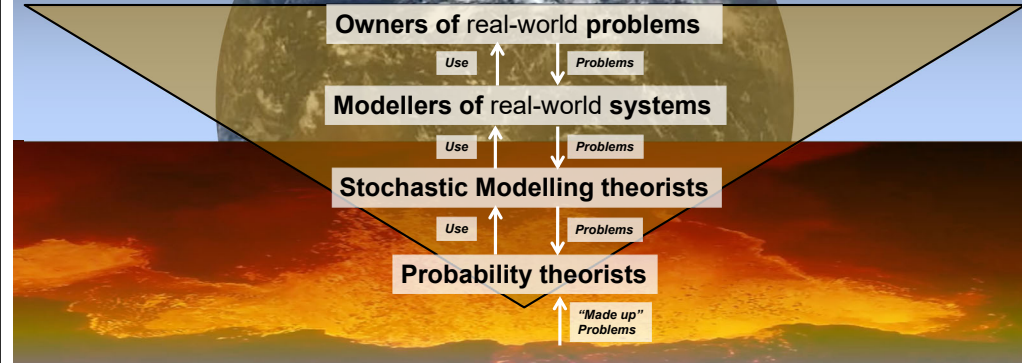
recipes
procedures ...
algorithms
theory
models
....



has **HUGE** long-term value

But a small pool of people capable of doing it

Roles in the probability World



Roles in the probability World

The Future of Work



Now, while ...

But, the ability to devise new ...

recipes
procedures ...
algorithms
theory
models
....



has **HUGE** long-term value

But a small pool of people
able to do it

From "Probability modeling and thinking: What can we learn from practice?"
-- Statistics Education Research Journal, 2016

The Future of Work



The ability to uncover and navigate ...

Human dimensions ...



will always have **HUGE** long-term value

And need everyone's
involvement



What else helps life-long learning and creativity?



And with **problem solving**,
because the more you're aware of, the more possibilities you can entertain and connect




Intimations of what's around the next bend ...



or just over the horizon ...

fuelling a *desire* to
go there





Intimations of *what just might be possible*


Curiosity & awarenesses

seed imagination, problem solutions and discovery

that arise from making novel connections
between elements stored in your brain

Beyond the 3 R's

Need a 4 E's of stats-ed

- Entice
 - Engage
 - Excite
 - Empower
- 

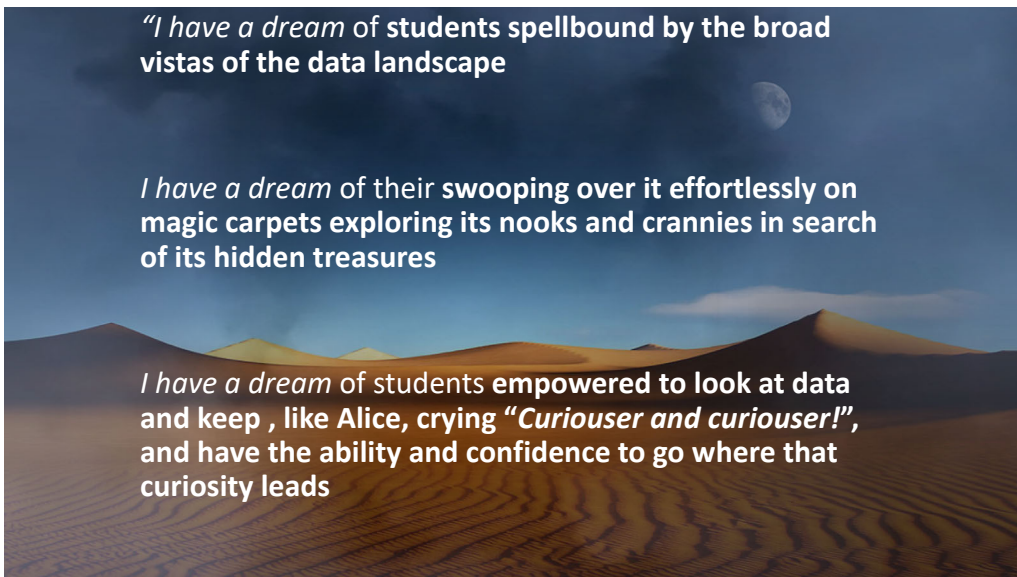
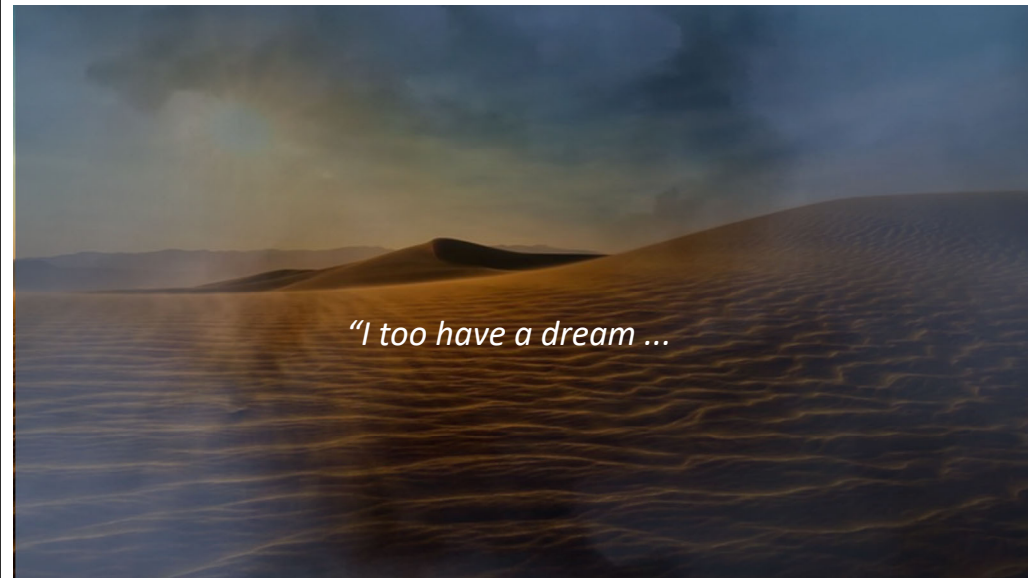
What is *the single best thing*
we can do for our students?

To populate

their imaginations

with possibilities









We Will Plot You!



We Will Rock You!



We Will Plot You!

