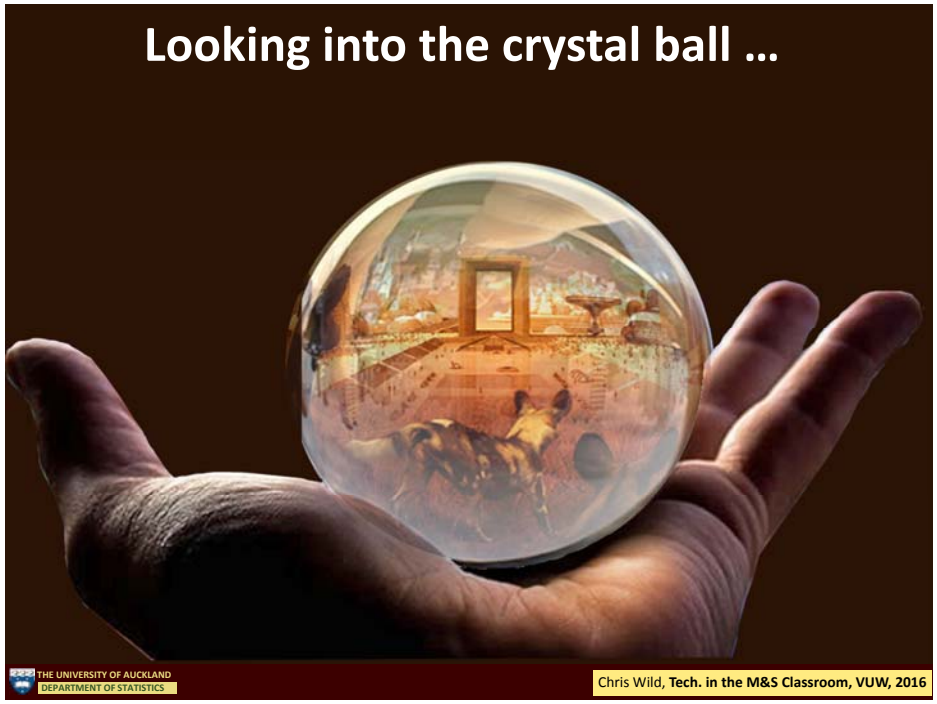
All links at http://bit.ly/2h0DCEA

Gazing into the data Future

Chris Wild

University of Auckland

Looking into the crystal ball …



STEMS: PUTTING STATISTICS INTO STEM IN THE AGE OF DATA

National Colloquium and Workshop | 2-3

Sydney, June 2016

Speakers: *influential decision leaders in Australian government, industry and education*:
https://stems2016.com/program/

(*we have never had anything like this in NZ*)

Basic Focus:

How can the education system start to deliver the volumes of people society needs with highly-developed data skills?

Final Report at:
https://stems2016sydney.files.
stems2016-report.pdf (see

All links at http://bit.ly/2h0DCEA

STEMS Plenary speakers

- Professor Xiao-Li Meng, ***Dean, Graduate School, Harvard University***
- Dr Roslyn Prinsley, ***National Adviser, Science & Mathematics Education and Industry***
- Professor Alan Finkel, ***Chief Scientist, Australia***
- Mr Hamish Treleaven, ***EGM Portfolio & Market Risk Management, Commonwealth Bank of Australia***
- Ms Helen Owens, ***Assistant Secretary, Public Data Branch, Office of Prime Minister & Cabinet***
- Mr Robert Randall, ***CEO, Australian Curriculum, Assessment and Reporting Authority***
- Professor Louise Ryan, UTS and ***Chief Investigator, ARC Centre of Excellence for Mathematical and Statistical Frontiers***
- Professor Nicholas Fisher, a ***Past President of the Statistical Society of Australia***

Trends

- More and more …
  - … **data** being collected and stored
  - … **awareness** of the desirability of **evidence based decision** making
  - … people being given **access** to data
  - … **software tools** empowering **more people** to analyse data
  - … **accelerating demand** for people with good data skills
    - Beyond the capability of present educational systems to deliver
  - … problems that **challenge the ability of computers to cope** ("big data")
    - Basically ***the more you can do***, ***the more*** you can see that ***you'd like to do***
      - *"wouldn't it be great if we could also …"*
- Broad skill sets needed for this
  - Will focus on ***a view from Statistics***

Chapter 1: What is Statistics?

Christopher J. Wild, Jessica M. Utts, and Nicholas J. Horton

Christopher J. Wild
University of Auckland, New Zealand

President, American Statistical Association

Jessica M. Utts
University of California, Irvine, United States

Nicholas J. Horton
Amherst College, United States

Chair, Statistical Education Section, ASA

e-mail: c.wild@auckland.ac.nz, jutts@uci.edu, nhorton@amherst.edu

Abstract

What is Statistics? We attempt to answer this question as it relates to grounding research in statistics education. We discuss the nature of statistics (the science of learning from data), its history and traditions, what characterises statistical thinking and how it differs from mathematics, connections with computing and data science, why learning statistics is essential, and what is most important. Finally, we attempt to gaze into the future, drawing upon what is known about the fast-growing demand for statistical skills and the portents of where the discipline is heading, especially those arising from data science and the promises and problems of big data.

So, what is statistics?

"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**
  - **& computer software**
    - data harvesting/wrangling, storage, organisation, cleaning, visualization and analysis
      - that last often based on **mathematical models** and derivations

"Statistics is **the science of learning from data**, and of measuring, controlling and communicating uncertainty."

-- American Statistical Association

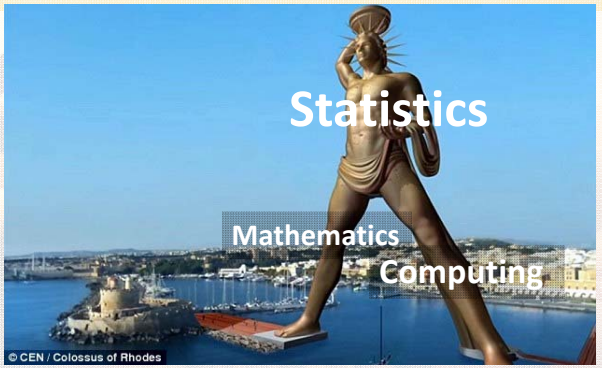
- The **focus** of statistics is:
  - **understanding the world through data**
- The **raw materials** for statistics are:
  - **real-world questions** and **data**
- The **tools** of statistics are:
  - **statistical** ways of **thinking**
  - & **computer software**
    - data wrangling/harvesting, storage, organisation, cleaning, visualization and analysis
      - that last often based on **mathematical models** and derivations

So *for statistics …*

what are mathematics & computer science?

- They are ***means*** to be pressed into service
  for reaching statistical ends
  - Just as they are for Physics, Economics, Engineering, ….
- The ***core*** *of statistics* is sets of fundamentally statistical concepts constructed using algorithmic, mathematical & scientific conceptual building blocks

Developing Statistics



- But you can't build it that way educationally
  - *Too slow and shuts out too many people*
- and luckily you don't have to
  - Software is increasingly empowering people to acquire substantial data understanding and capability with minimal programming and maths
    - But they then have to rely on tools (software and models) built by others
      - *they can't develop new methods without specialist help*
      - *Capability limited by the tools they know how to use*

Developing Statistics

Computing

motivation

tools

Mathematics

motivation

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

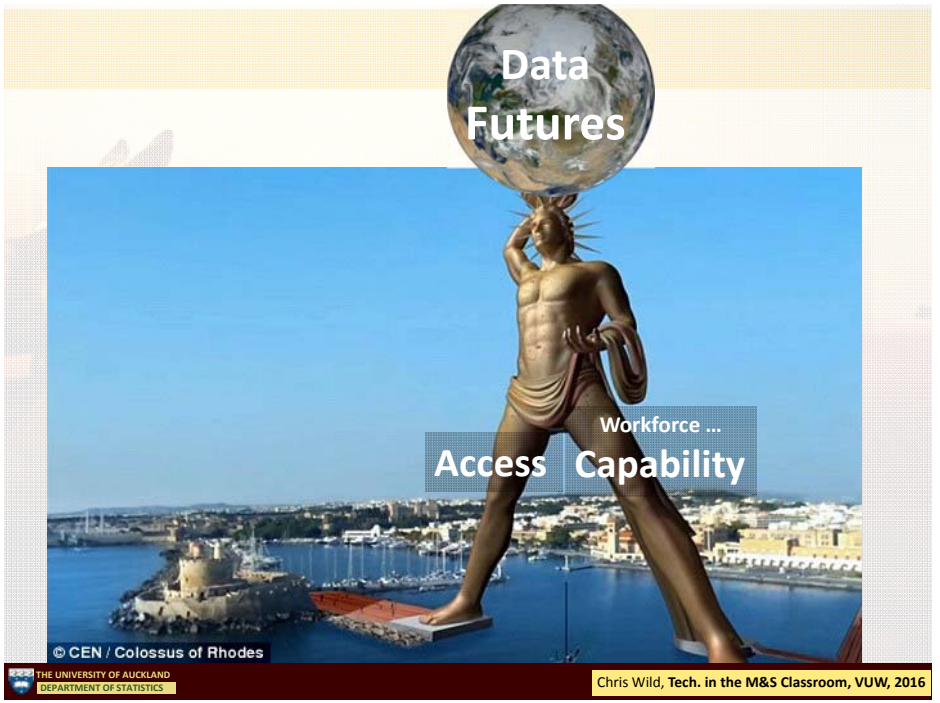
Main roll in statistics
- ***Mathematical*** ways of ***thinking***
- Algebraic & calculus skills
  - Used for developing new models & methodologies

Data Futures for New Zealand

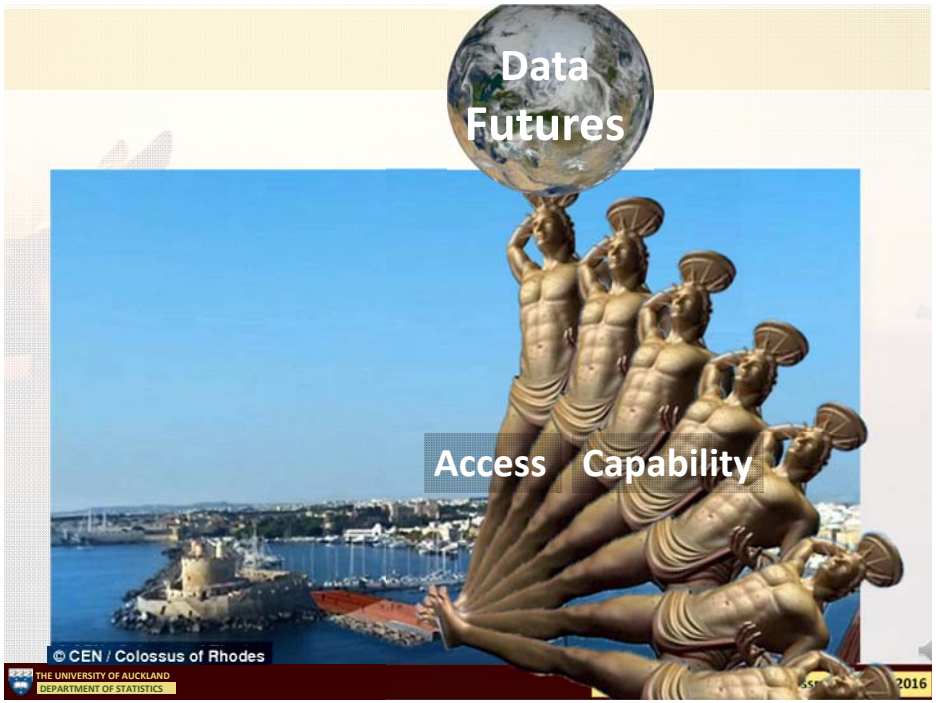
Two key ingredients:

1. **<u>Access</u> to data**
   - Integrated Data Infrastructure etc …
2. **<u>Capability</u> of NZ'ers**
   (Workforce & communities)
   to gain value from it

Social & Economic



THE UNIVERSITY OF AUCKLAND DEPARTMENT OF STATISTICS
Chris Wild, **Tech. in the M&S Classroom, VUW, 2016**



THE UNIVERSITY OF AUCKLAND DEPARTMENT OF STATISTICS
2016

THE UNIVERSITY OF AUCKLAND DEPARTMENT OF STATISTICS
Chris Wild, **Tech. in the M&S Classroom, VUW, 2016**

The Future of Work

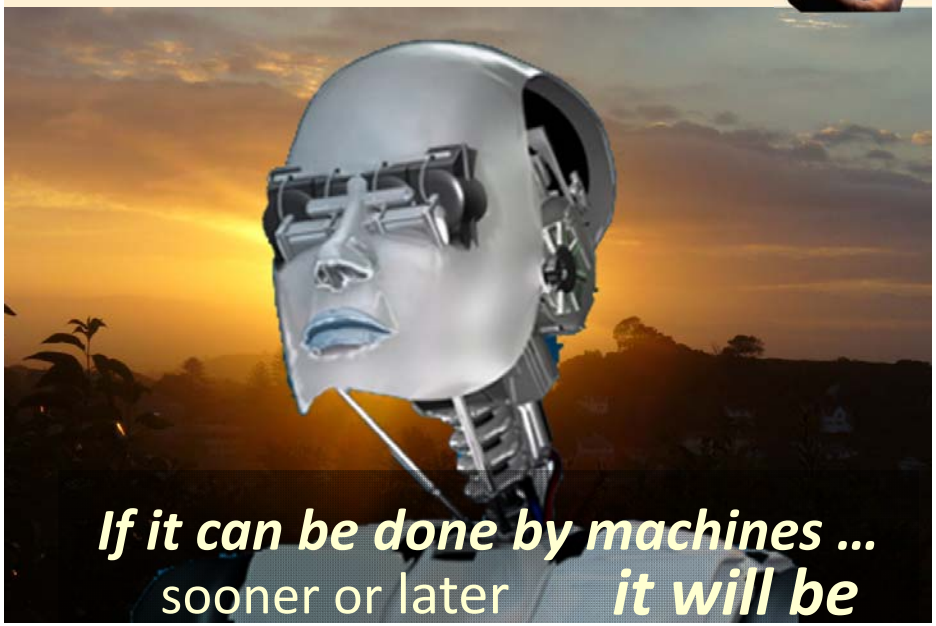
The Future of Work …

- Whatever you think of the Labour Party's

THE FUTURE OF WORK

the nature of work for is changing
and we need coping strategies

The Future of Work …



The Future of Work …

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

The ability to do things machines can't do!

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***

For this …

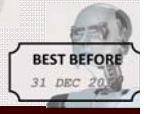
- **Big-picture conceptions …**
  have **long-term value**
- **Details** are **death dated**

The ability to operate …

a ***recipe / procedure / algorithm*** **…**
has **long-term value**

The ability to operate …

any particular ***recipe / procedure / algorithm*** **…**
is **death dated**



The ability to …

code (program) **…**
has **long-term value**

The ability to …

code (program) **…**
anything particular (*in any particular language*)
is **death dated**

Software, technology

- Best thought of as:
  - ***Automators*** of mechanical processes
  - ***Aids to human thinking*** & problem solving

    Echoes of:

    He aha te mea nui o te ao
    What is the most important thing in the world?
    He tangata, he tangata, he tangata
    It is the people, it is the people, it is the people
    -- Māori proverb

    But this time, the most important thing …

    "It is the people's *thinking*, It is the people's *thinking*, It is the people's *thinking*"!
- Operating particular procedures / Driving software is not a worthy end goal
  - it's a temporary (stop-gap) measure

What else helps life-long learning?

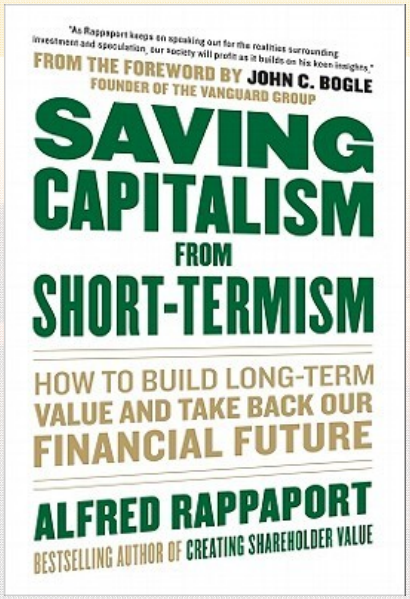
Intimations of what's **around the next bend …**

or just over the horizon …

Intimations of what just might be possible

the seeds of problem solutions and discovery

Change of tack …



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 *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

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

That's terrible, isn't it ??!!

Students and short termism

- Focus on "*What do I need to pass/ get an Excellence*?"
- "***Just give me the answer!***"
  - or at least a simple recipe!
  - ***Anything but*** tell me **"*I actually have to think*"!**

Terrible, isn't it ??!!

Recipes and **The Future of Work**

"If you** can only **think like a robot …

you'll be replaced by one!"

– Andrew Balemi

Qualifications vs capabilities

Short-termism and …

Teachers:

- A focus on
  - just getting it done
  - and getting "good NCEA/CIE results"

University teachers:

- A focus on
  - just getting it done
  - and getting good student evaluations

That's all terrible, isn't it ??!!

Universities:

- When, "***What does the university need*?**" results in a list of topics to cover!

Aside: What do universities need?

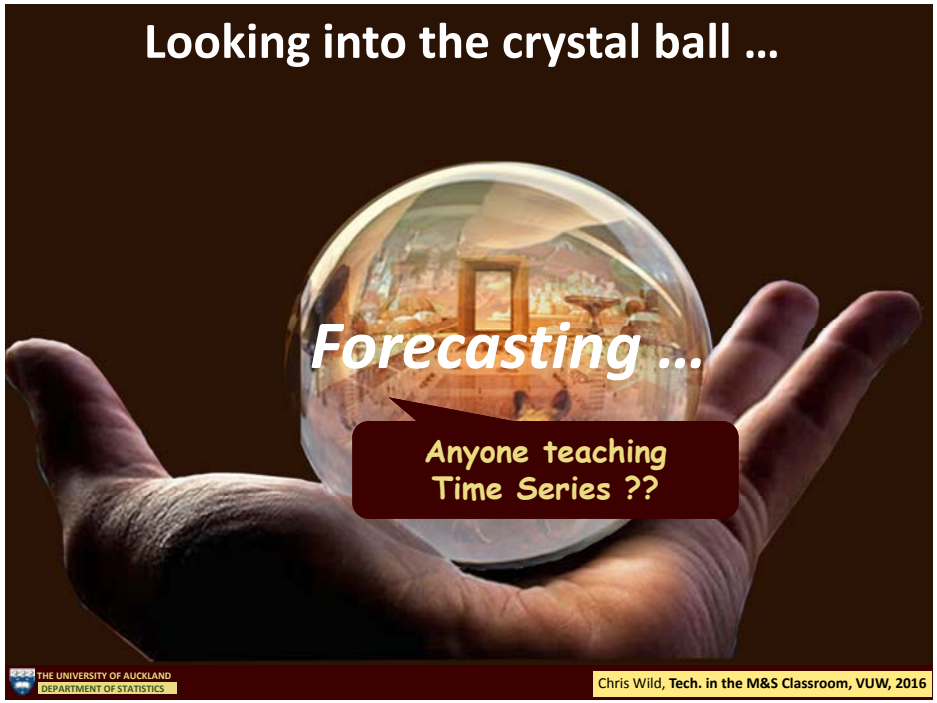
- ***Do they really need what they say they need?***
  - Universities have total control over what they teach, when and how
  - What we most need is a ***good supply*** of bright ***students*** who have been ***turned on to our subject***
    - *When teachers can turn their students on to their subjects we should be abjectly and eternally grateful*
      - Without that we whither and die
      - With just that (if we are half-way smart) we can prosper
        - Bright highly-motivated people learn specifics fast

Back to "short-termism"

Actually,

- **we are all short-termists**
  - The reward systems we work under reinforce that
- ***But that shouldn't entirely define us …***
  - "greater good" often comes from considering the longer term

Looking into the crystal ball …

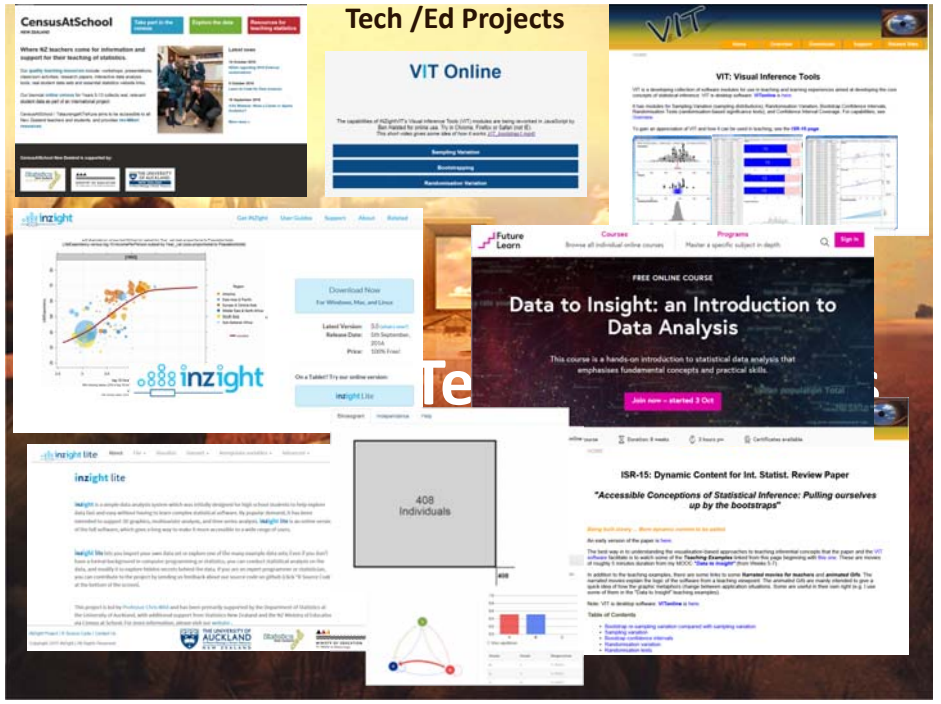


Forecasting …

is like driving when you can only ***see what's in the rear-view mirror***

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

Tech /Ed Projects



My Ed/Tech Projects

Developing in concert …

Software ⟷ **Educational Experiences**

Research ***(Maxine)***

Desired Capabilities

My Ed/Tech Projects

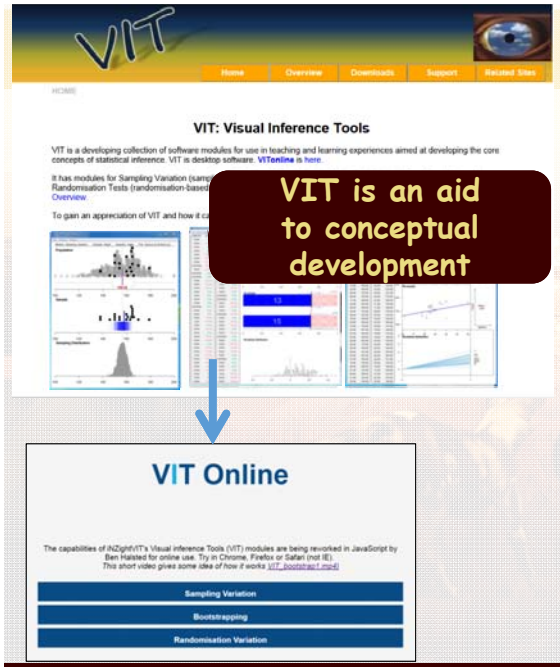
(*Visual Inference Tools*)

inzight VIT

Data Analysis ⟷ **Conceptual development**

Teaching

DATA *to insight*



Why Animate?

- To convey nature of problems
- To show how processes work
- To convey "randomness"
- Critical to connect all the conceptual dots as a process develops
  - *If there is a crack … they will fall down itl*

Confidence Intervals: *What game are we playing??*

Problem …

- ***"my estimate is wrong"*** ≡ *"I can never trust any estimate to be quite right!!"*
  - **Sampling error** is an **important source** of error

A solution …

- **Put margin around *estimate*** to **allow for** likely extent of **sampling error**
  ("*Confidence interval*")
  - *New problem*: Can't see "likely extent of sampling error"
  - What can we do?! *Approximate*

- **How do such intervals behave?**
- **Why should I trust the results?**
- **Now I have one, how do I use it?**

How ??

Postscript: Which error(s) have I accounted for and which not?

Confidence Intervals: *What game are we playing??*

Problem …

- ***"my estimate is wrong"***
  - **Sampling error** is an **important source** of error

A solution …

- **Put margin around** estimate to **allow for** likely extent of **sampling error**
  ("*Confidence interval*")
  - ***New problem***: **Can't see "**likely extent of **sampling error"**
  - **What can we do?!** ***Approximate***

Math:
- Make assumptions &
- deduce the behaviour

- Hard to understand
- New situation, new theory

Confidence Intervals: *What game are we playing??*

Problem …

- ***"my estimate is wrong"***
  - **Sampling error** is an **important source** of error

A solution …

- **Put margin around** estimate to **allow for** likely extent of **sampling error**
  ("*Confidence interval*")
  - ***New problem***: **Can't see "**likely extent of **sampling error"**
  - **What can we do?!** ***Approximate***

Bootstrap:
Approx. by resampling error

- Only 1 big idea
- General

Confidence Intervals: What do we need to make this story work?

Problem …

- *"my estimate is wrong"*
  - **Sampling error** is an i…

2. Idea of "putting a margin around" …

1. Appreciation of "sampling error"

A solution …

- **Put margin around *estimate*** to **allow for** likely extent of **sampling error**

  ("*Confidence interval*")

  - ***Approximate*** **"**likely extent of **sampling error"** by **re-sampling error**

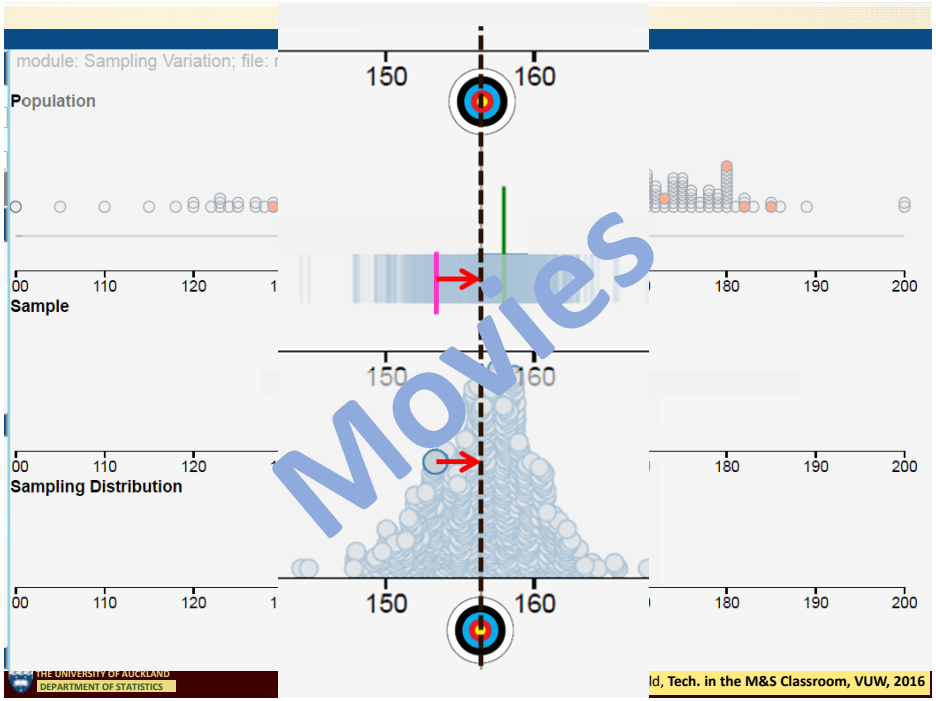
4. Needed for this idea …

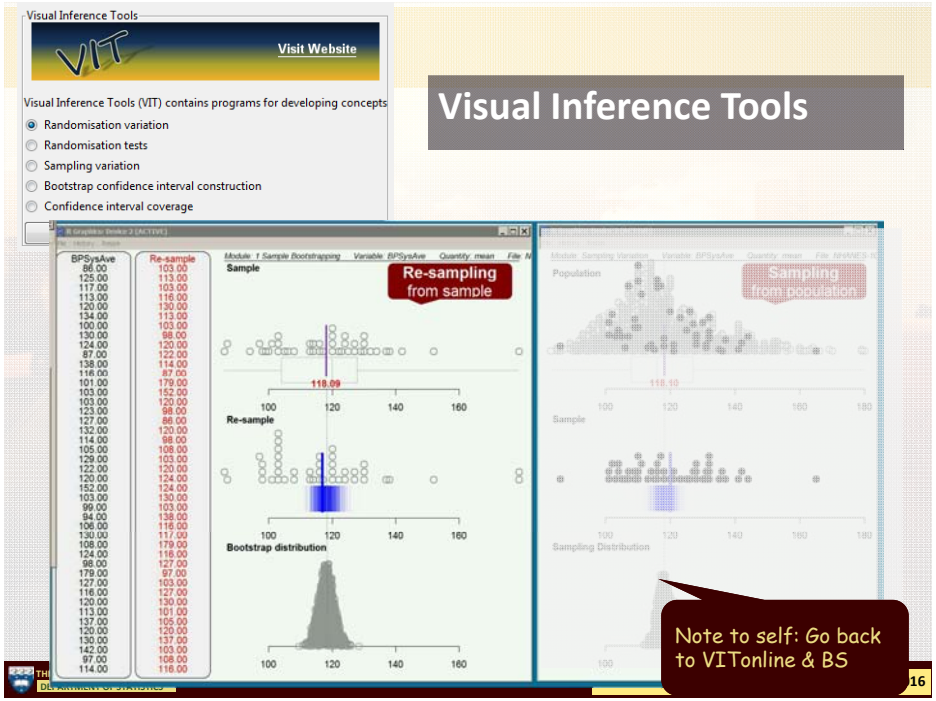
3. This idea …

Boot… by resampling error

5. An appreciation of a construction process …

How can visualisation help?





Confidence Intervals: What do we need to make this story work? ...laying??

Problem ...

- *"my estimate is wrong"*
  - **Sampling error** is an i...

2. Idea of "putting a margin around" ...

1. Appreciation of "sampling error"

A solution ...

- **Put margin around *estimate*** to **allow for** likely extent of **sampling error**

  ("*Confidence interval*")

  - ***Approximate*** "likely extent of **sampling error**" by **re-sampling error**

4. Needed for this idea ...

3. This idea ...

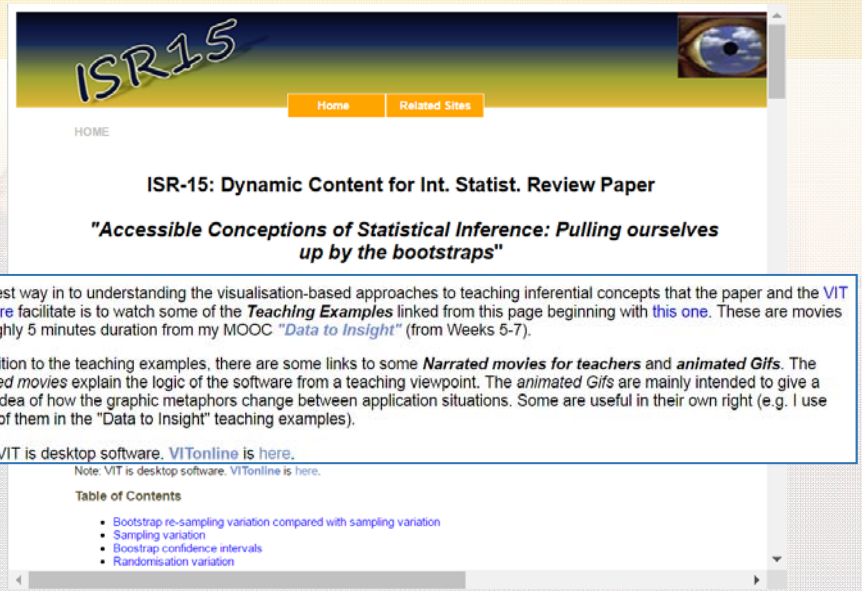
5. An appreciation of a construction process ...

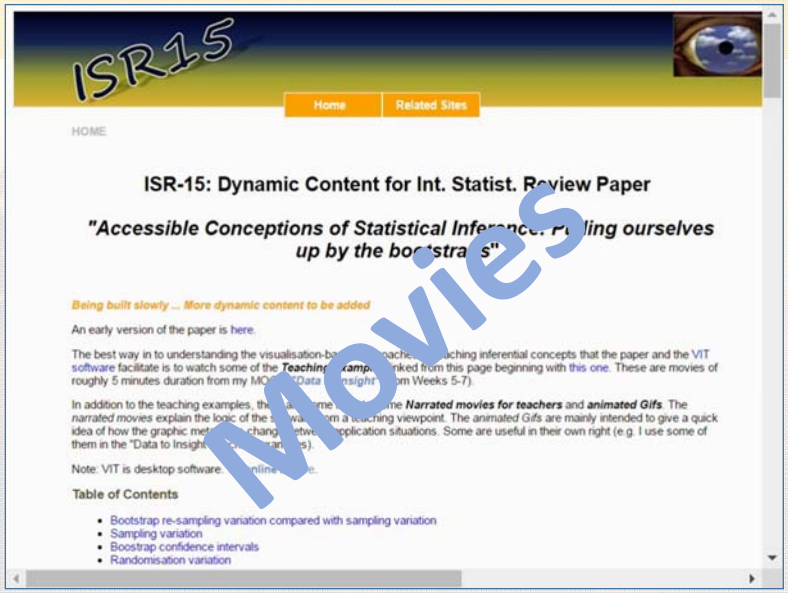
Job done, right?

The job's not done when you've explained it to your students ...

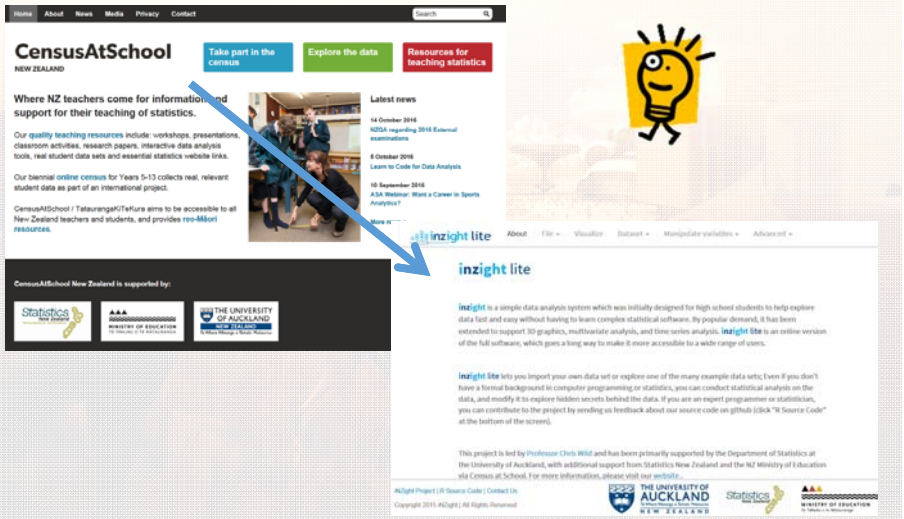
The job's done when your students can explain it to you

Or better still, to Great Aunt Myrtle
or Great Uncle Horace

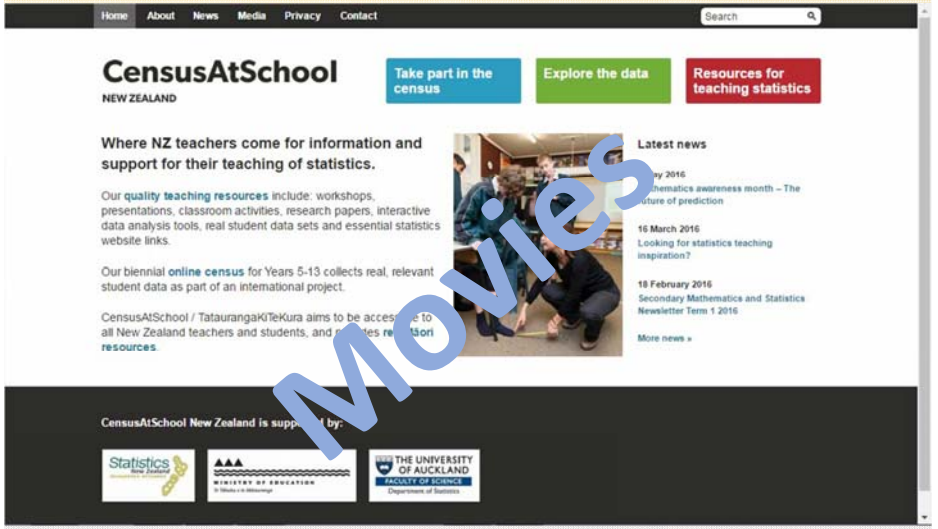




Linking CensusAtSchool & iNZight Lite

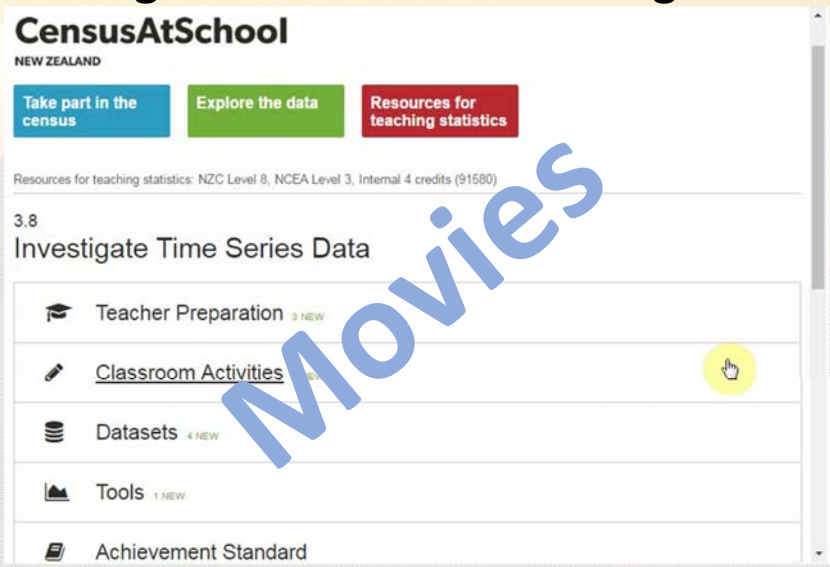


Linking CensusAtSchool & iNZight Lite

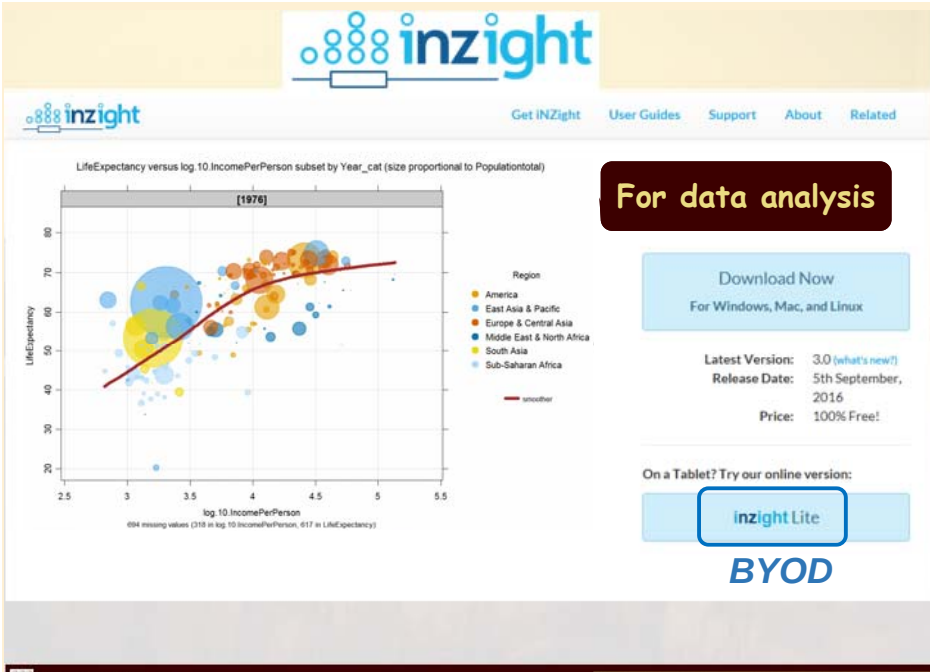


Movies

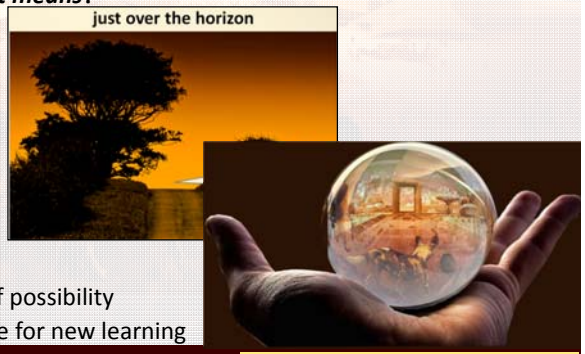
Linking CensusAtSchool & iNZight Lite



Movies

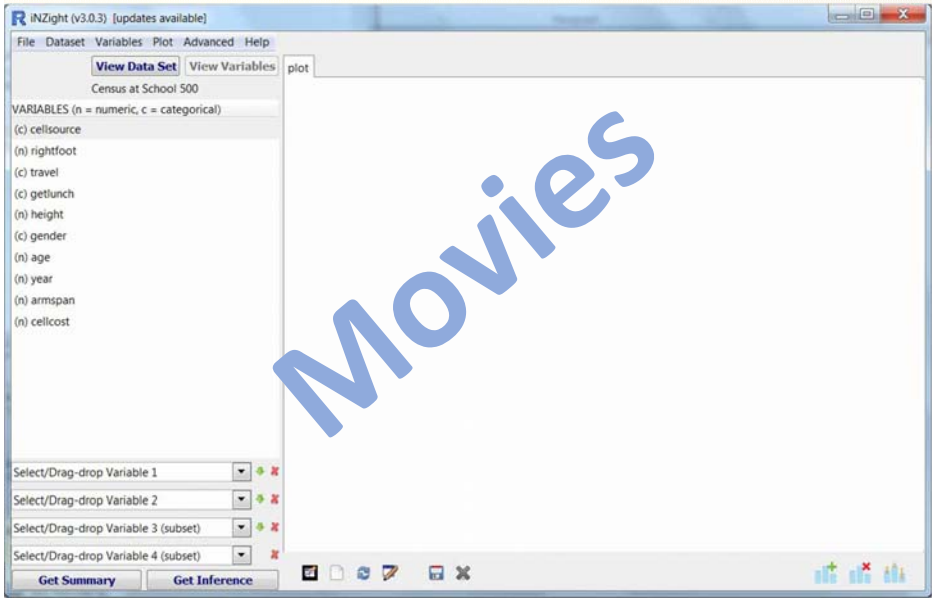


- Bigger needed for NCEA
- Why a plus?
  - Many young people are curious
    - They just play with things, ***"I wonder what this does?"*** and ***"I wonder what would happen if I ..?"*** which can also lead to "***I wonder what that means?***"
    - Intimations of ….



- Enlarging awareness
- Populating their sense of possibility
- Planting seeds of a desire for new learning

& simple things are easy …



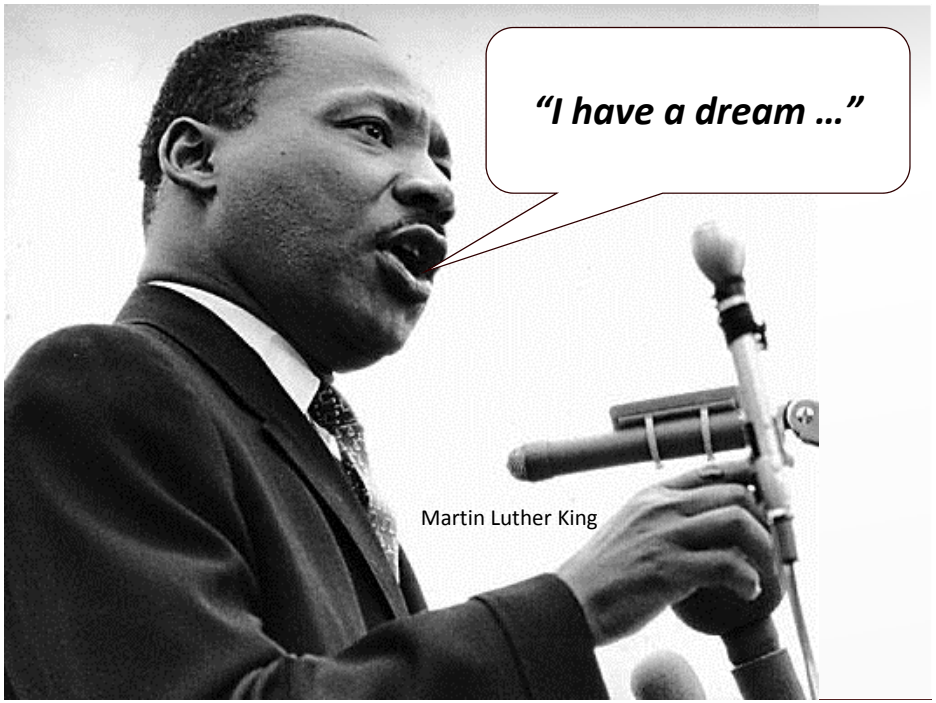
Running again March & October next year

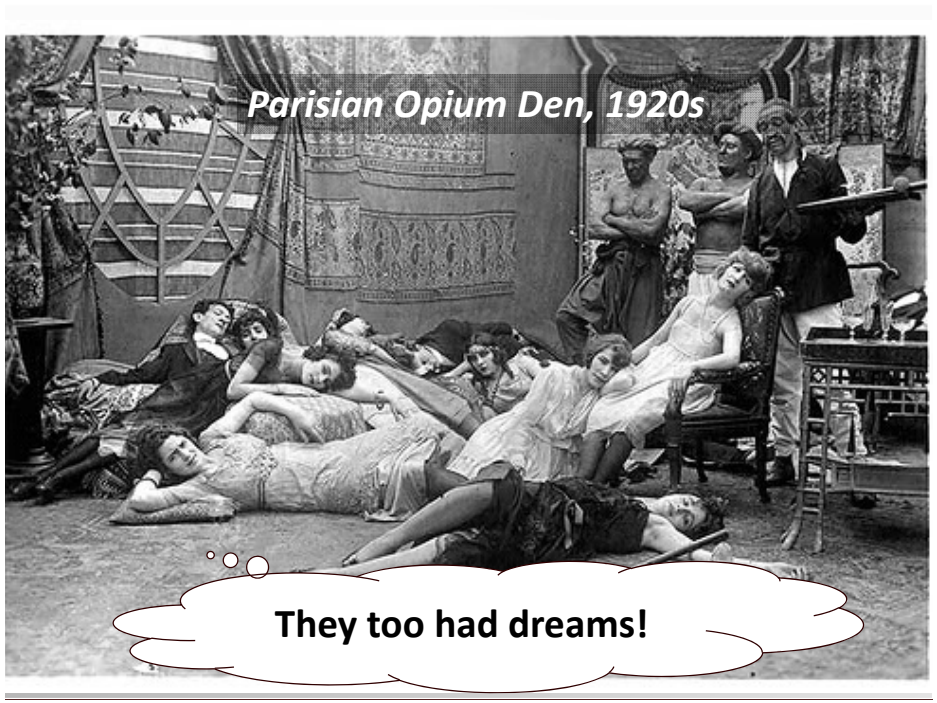
Why did I do it?

1. Prototype a getting-further-faster intro-stat
   - for ourselves and others
     - *learn by building a model*
2. Build PD for NZ high school teachers
3. Gain experience with largish-scale online education
4. Gain some high-quality resources

What's "teaching online" like?

- Very stimulating
  - Questioning and discussion much deeper and broader than a university class
  - Why?
    - All the students are there simply because they wanted to learn
    - a greater level of personal confidence and life's experience in mature (older) learners
    - the presence of many professionals and researchers who wanted to apply the ideas in their work
  - Good for slowly improving your messaging
- Also quite time consuming!





"I have a dream of **students spellbound by the broad vistas of the data landscape**

I have a dream of their **flying on magic carpets that enable them to swoop effortlessly over this landscape exploring its nooks and crannies in search of its hidden treasures**

I have a dream of students **empowered to look at data, explore analysis systems and educational environments designed so that, like Alice in Wonderland, they keep crying "*Curiouser and curiouser!*" and have the ability and confidence to go where that curiosity leads**

I have a dream of **educational and analysis environments designed to leverage the power of "I wonder …?" to draw students in to discovering more and more –**

the power of **"I wonder why that is?",** *the power of* **"I wonder what happens if …?",** *the power of* **"I wonder what that does?",** *the power of* **"I wonder what's around the next bend or just over the horizon?"**

I have a dream of **software that finesses away the mundane, the mind-numbing and the soul-destroying difficulties.**

I have a dream ... "

Thank you

All links at http://bit.ly/2h0DCEA