



## ***“Through a glass darkly”***

Chris Wild  
University of Auckland

ICOTS  
Looking back, looking forward

Slides at <http://bit.ly/icots10>

## **What to expect**

- Abstract promised *riffing* on a series of themes
  - It also promised music
- “riff” – *a series of notes, chord patterns or musical phrases that are repeated*
- Musical analogy for this talk ...  
**improvisational jazz** with Salvador Dalí
- Expect a kaleidoscope of metaphors, impressions and images rather than structured arguments



## **A game of two halves**

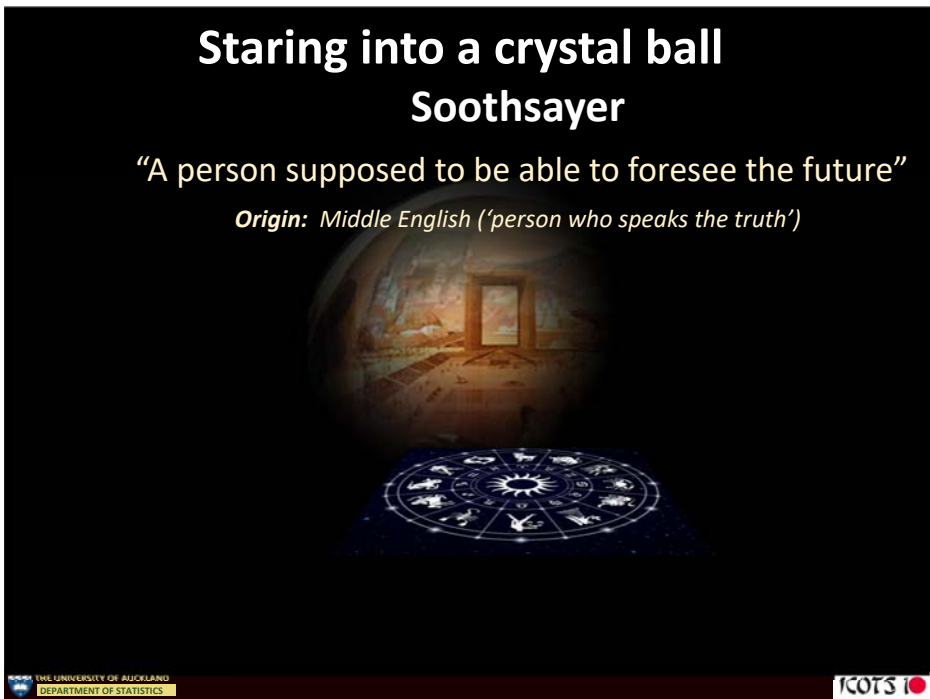
- **Part I : Back to the Future**
- **Part II: “Oh Say, Can we See?”**

## **The (missing?) Link**

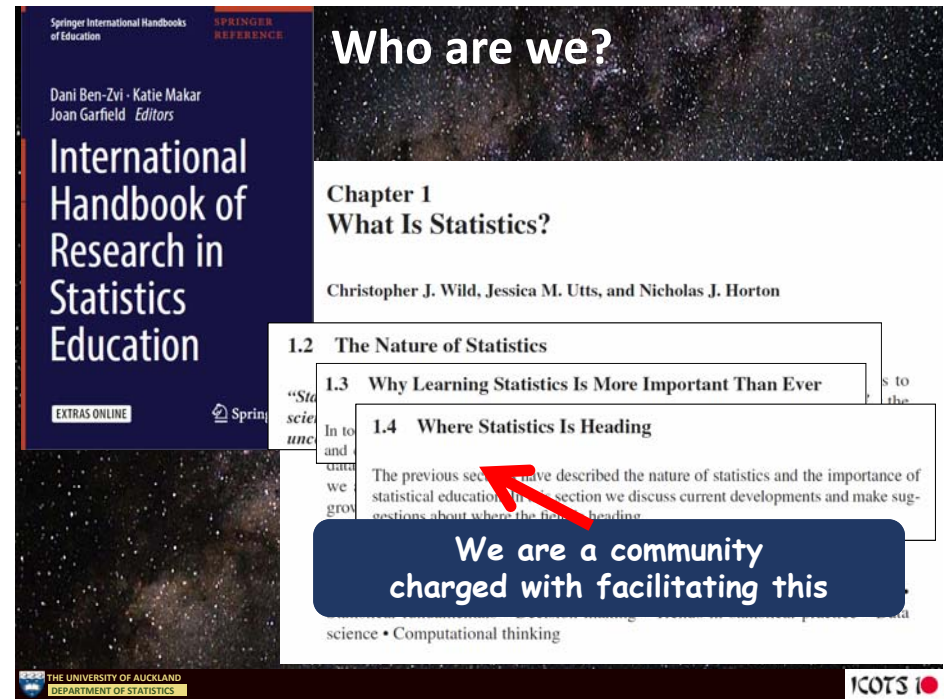
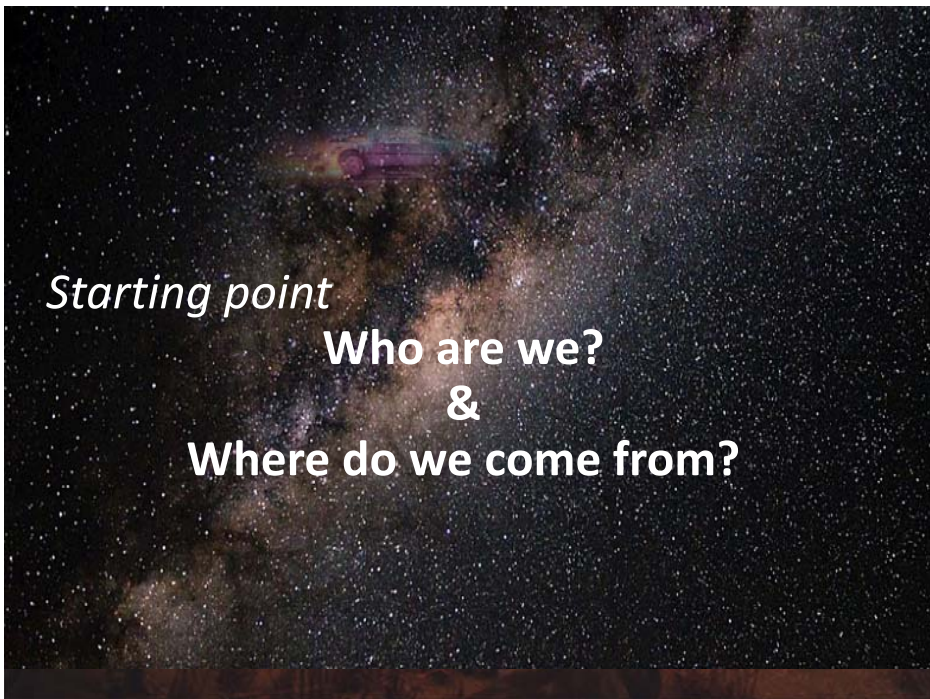


### ***“Through a glass darkly”***

To see “through a glass darkly” is to have an obscure or imperfect vision of reality  
(*dictionary.com*)









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

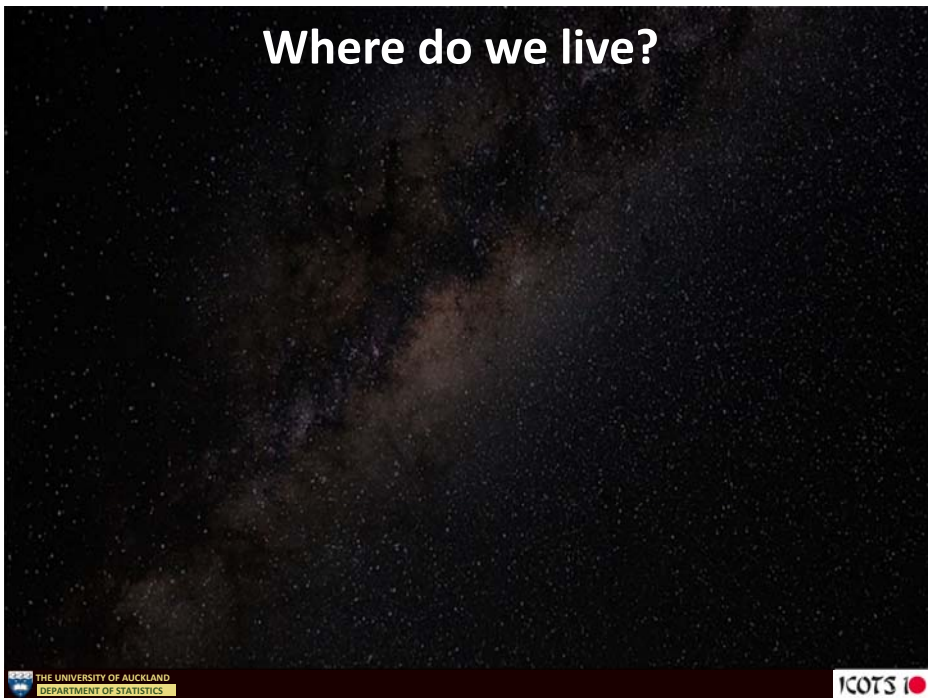
## Who do we serve?

- Intending statistical/data science specialists
- Intending specialists in other areas that know their students need these skills
- **The wider society ...**
  - by spreading valuable messages as widely as possible

Bright  
Data  
Futures

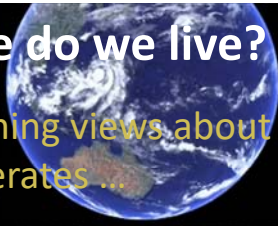
Societal ...  
Access Capability





## 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
- **qualitative research methods**
- data mining
- **machine learning**



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

## Where do we live?



Totally Human

Totally Automated

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



## Where are we going?

- “Data science”
- Qualitative methods
- Stochastic modelling of real-world processes



## Where is the world going?

- Automation

## 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

## Where are we going?

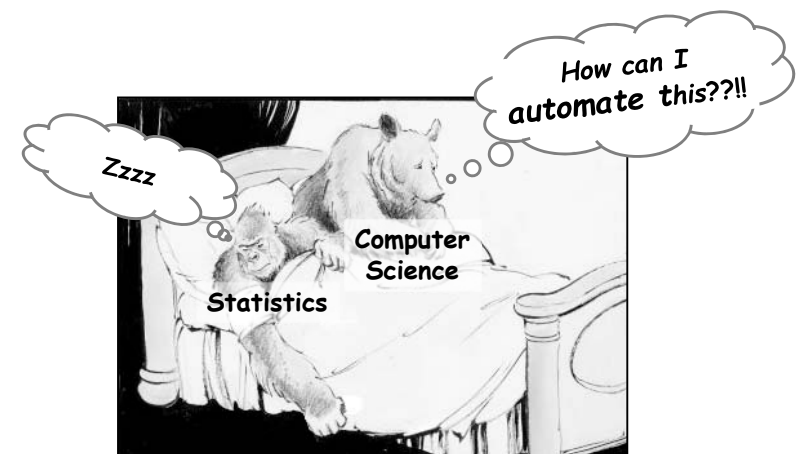
- “Data science”
  - It’s where the **demand** is
  - It’s where the **investment** funding is (e.g. new/replacement positions)
- Qualitative methods
- Stochastic modelling of real-world processes



## Where is the world going?

- Automation

## Bed fellows in data science



**The Bad**

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

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 UNIVERSITY OF AUCKLAND DEPARTMENT OF STATISTICS ICOTS

**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


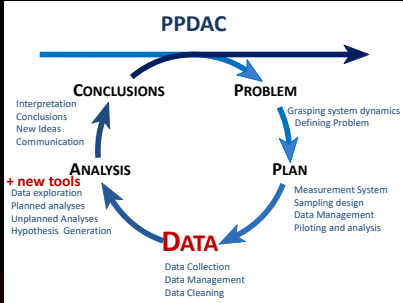
the correct interpretation of the data is rarely obvious, and that the obvious-seeming interpretation isn't always correct

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

THE UNIVERSITY OF AUCKLAND DEPARTMENT OF STATISTICS ICOTS

**Where is "data science" sending us?**

- Into new-for-us areas ...
  - Harvesting & assembly of data from electronic sources
    - e.g. databases & websites via apis or scraping
  - Data wrangling (getting data 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)

THE UNIVERSITY OF AUCKLAND DEPARTMENT OF STATISTICS ICOTS

Back to the Future ...

Roadblocks: **Short termism**

a bright future!

THE UNIVERSITY OF AUCKLAND DEPARTMENT OF STATISTICS ICOTS



*"As Rappaport keeps on speaking out for the realities surrounding investment and speculation, our society will profit as it builds on his keen insights."*  
**FROM THE FOREWORD BY JOHN C. BOGLE**  
 FOUNDER OF THE VANGUARD GROUP

# SAVING CAPITALISM FROM SHORT-TERMISM

HOW TO BUILD LONG-TERM VALUE AND TAKE BACK OUR FINANCIAL FUTURE

**ALFRED RAPPAPORT**  
 BESTSELLING AUTHOR OF CREATING SHAREHOLDER VALUE

**What is short-termism?**  
 The term refers to an excessive focus on short-term results at the expense of long-term interests.

*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.*

**Terrible, isn't it ???!**

THE UNIVERSITY OF AUCKLAND ICOTS

**BILL PROUD**

*"Our long-term plan is to find another short-term plan"*

## Students and short termism

Focus on *"What do I need to pass/ get an A?"*

*chart our way into a bright future?*

THE UNIVERSITY OF AUCKLAND ICOTS

## Qualifications vs capabilities

**Qualifications**  
 What you (can) have ...

**Capabilities**  
 What you can do ...

**SUCCESS!**

THE UNIVERSITY OF AUCKLAND ICOTS





## Students and short termism

- Focus on *“What do I need to pass/ get an A?”*

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 ???!**

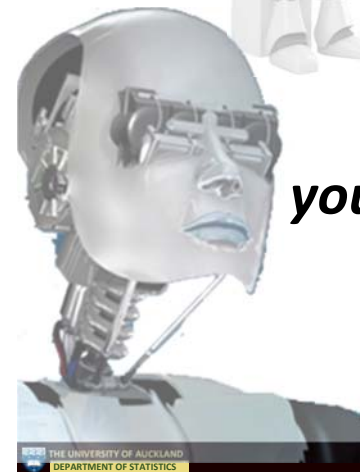
## Qualifications vs capabilities



**Qualifications**  
What you  
(can) **have** ...



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



*you’ll be replaced by one!”*

– Andrew Balemi



## But actually, *It's not just them !*

- **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



## The Future of Work



## The Future of Work ...



nzherald.com

BUSINESS

Comment: A global job cull is coming and NZ must prepare

12 July 2018 9:10pm

Automation, not trade, is what Tru  
be concerned about

Temper tantrum notwithstanding, we need to understand what Tru  
coming back, writes Gwynne Dyer

OPINION Jun 11, 2018 by Gwynne Dyer 25 Hamilton Spectator

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.

Technology will disrupt how we do almost everything

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

# 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 ...

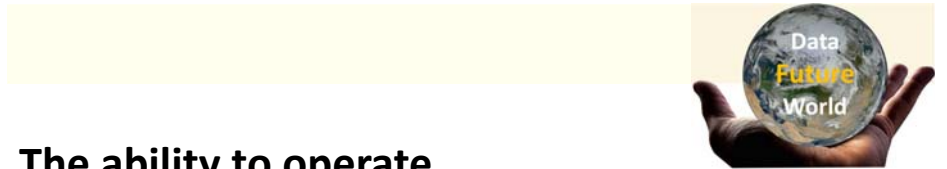
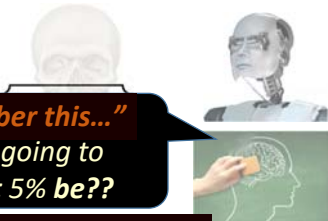


- Big-picture conceptions ...  
have long-term value



- Details are death dated

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



# 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

"Today's complex programming task is tomorrow's mouse click"  
-- Lucid Dreams about the Future

Anything generically useful gets wrapped up in a software module

BEST BEFORE  
31 DEC 2018



However, the ability to devise ...

*recipes*  
*procedures* ...  
*algorithms*



has **HUGE** long-term value

## Software, technologies



- Best approached as:
  - **Automaters** of mechanical processes
  - **Aids to human thinking** & problem solving
- We need to ...
  - **Identify and emphasize the key decision points** where **human insights add value**
  - Think in terms of the rest being automated
    - Except for those who will be tool makers rather than tool users
- The most important **human intervention points**
  - are around **goal setting** and **critique**
    - e.g. of assumptions

And what do we know about any assumption? ...

"It ain't necessarily so. Oh yeah it ain't necessarily so"



## What else helps life-long learning?

**Peripheral** **awarenesses**



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



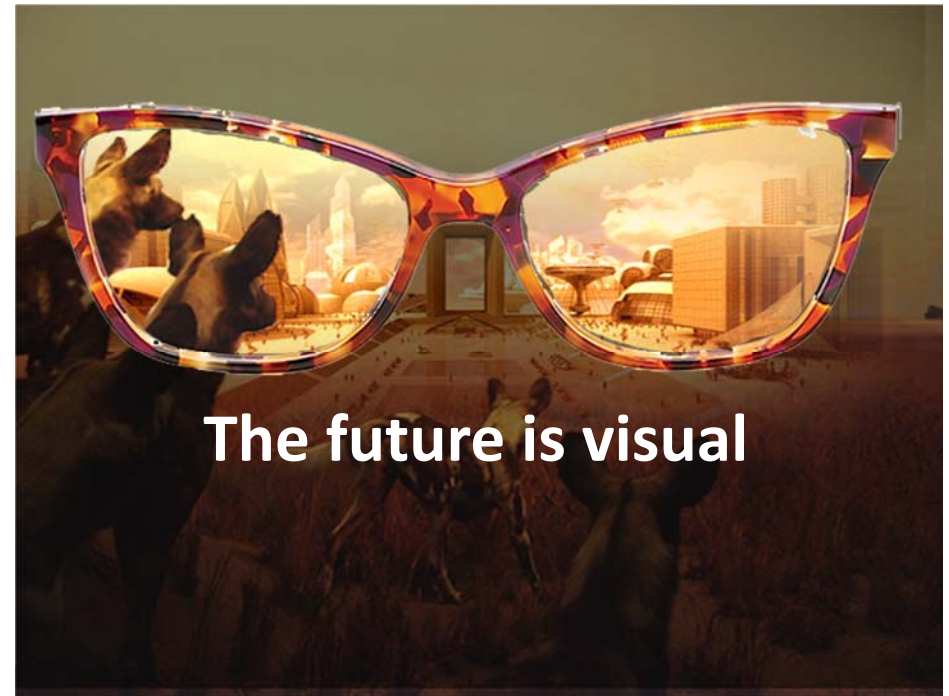
or just over the horizon ...



Intimations of *what just might be possible*



the seeds of problem solutions and discovery



The future is visual



# The future is visual

The biggest hope for increasing the capabilities of a broad cross-section of society

- to better understand data
  - To better understand data-related concepts
- Involves graphics/visualisations
- especially interactive and dynamic graphics online

Part II of this talk involves a lot of graphics for data  
The next ? Slides are teasers for conceptual graphics

# Conceptual graphics – Teasers

scampy

DATA MODE

Enter/paste your data (actual times) as comma separated values:

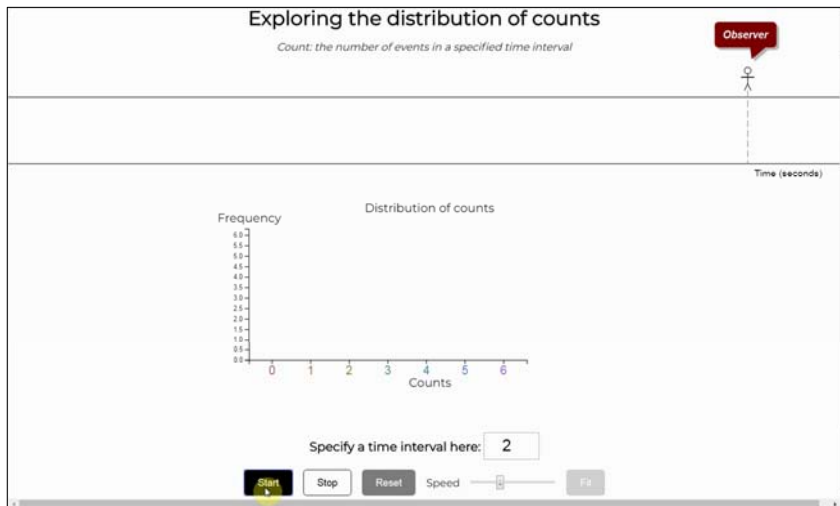
```
6,16,21,42,45,45,54,62,73,77,92,96,96,105,116,120,123,130,140,155,165,173,181,187,194,203,220,229,230,242,261,263,267,292,296,30
6,314,328,329,345,355,360,367,382,385,393,404,410,422,423,432,446,460,463,478,501,511,524,527,549,549,561,561,561,572,592,592,599
,623,630,654,654,665,679,680,692,696,705,715,738,739,746,777,782,783,783,792,809,810,812,827,837,838,843,855,864,881,887,908,911,
916,919,936,939,948,961,976,979,993,1000,1011,1016,1019,1030,1032,1045,1049,1051,1076,1101,1124,1124,1146,1146,1162,1173,1179,119
7,1204,1215,1228,1233,1249,1255,1268,1280,1281,1305,1307,1310,1313,1324,1327,1342,1344,1355,1365,1376,1378,1393,1405,1405,1440,14
43,1453,1466,1487,1493,1505,1506,1513,1520,1530,1542,1551,1561,1573,1579,1590,1594,1601,1606,1609,1633,1646,1646,1659,1666,1696,1
703,1708,1743,1748,1755,1757,1757,1772,1793,1793,1802,1808,1808,1821,1824
```

Units: seconds



NEXT >

# Conceptual graphics – Teasers



# Conceptual graphics – Teasers

# Why Animate?

## VIT: Visual Inference Tools

VIT is a developing collection of software modules for use in teaching and learning experiences aimed at developing the core concepts of statistical inference. VIT is desktop software. [VITonline is here](#)



### VIT Online

The capabilities of ANZStats VIT's Visual Inference Tools (VIT) modules are being rewritten in JavaScript by Ben Haines for online use. Try in Chrome, Firefox or Safari (not IE). This short video gives some idea of how it works. [VITonline](#)

- Sampling Variation
- Bootstrapping
- Randomisation Variation

- To convey nature of problems
- To show how processes work
- To convey “randomness”
- The key: *Carefully connecting all the conceptual dots as a process develops*

# “Data science” given us something new

- a growing cadre of undergrad students so allured by a career in “data science” that we can get close to ...

**“Jump!” “How high?”**

**But for the rest ...**

- the broad cross-section we need to reach to impact the wider society ...

this adage still applies ...



# We need the “3 E’s” of stats-ed

- Entice
- Excite
- Empower





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

Populate

their imaginations

with possibilities

"... and the vision that was planted in my brain still remains ..."

## Two Challenges

How can we trigger ...

*"This data world could light up my life" ?*

Or how can we make ...

intro stat like Speed Dating ?

*"Give me a call sometime. We'll go speed dating ..."*

(Hint: For most I don't think it's grinding out code)

Enough "hard thinking" for one  
talk!!

AND NOW  
FOR SOMETHING  
COMPLETELY  
DIFFERENT