

TRANSFORMING

DATA INTO

INFORMATION

USING

VISUALISATION



TODAY

Why Visualise Data?

Elements of Data Visualisation

Encouraging Curiosity through Data Visualisation

WHY VISUALISE DATA?

Exploratory

- Familiarisation with data
- Discover hidden patterns
- Explore large datasets

Explanatory

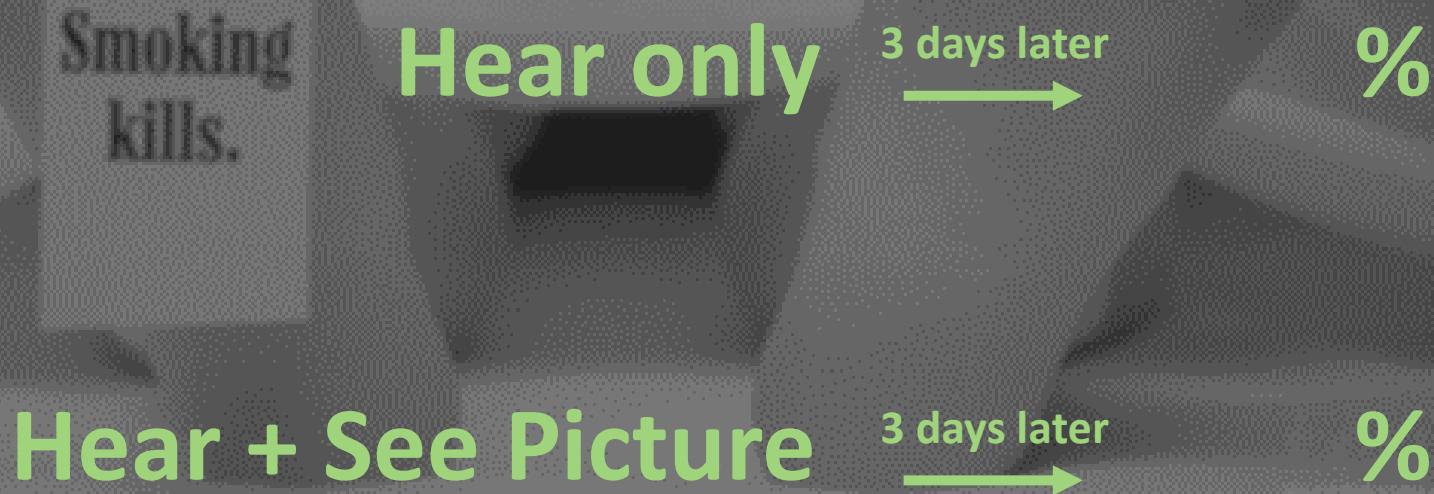
- Communicate insight
- Highlight presence of phenomenon
- Convey meaning

About 440,000 Americans die each year from diseases related to smoking.
90% of them started as teen smokers.



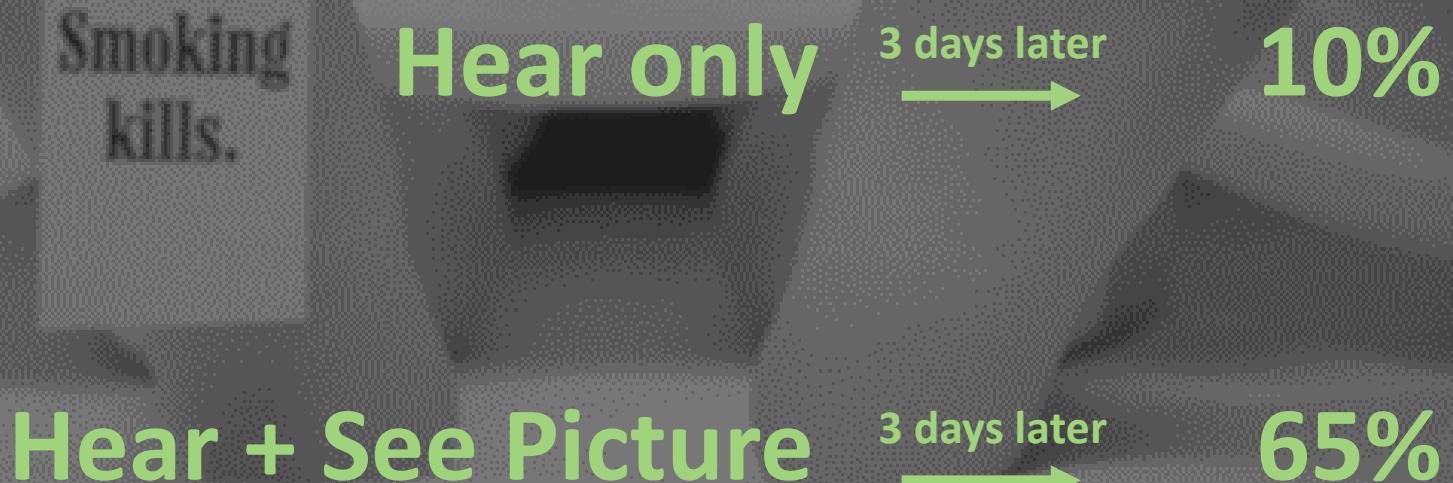
PICTURE SUPERIORITY EFFECT

About 440,000 Americans die each year from diseases related to smoking.
90% of them started as teen smokers.

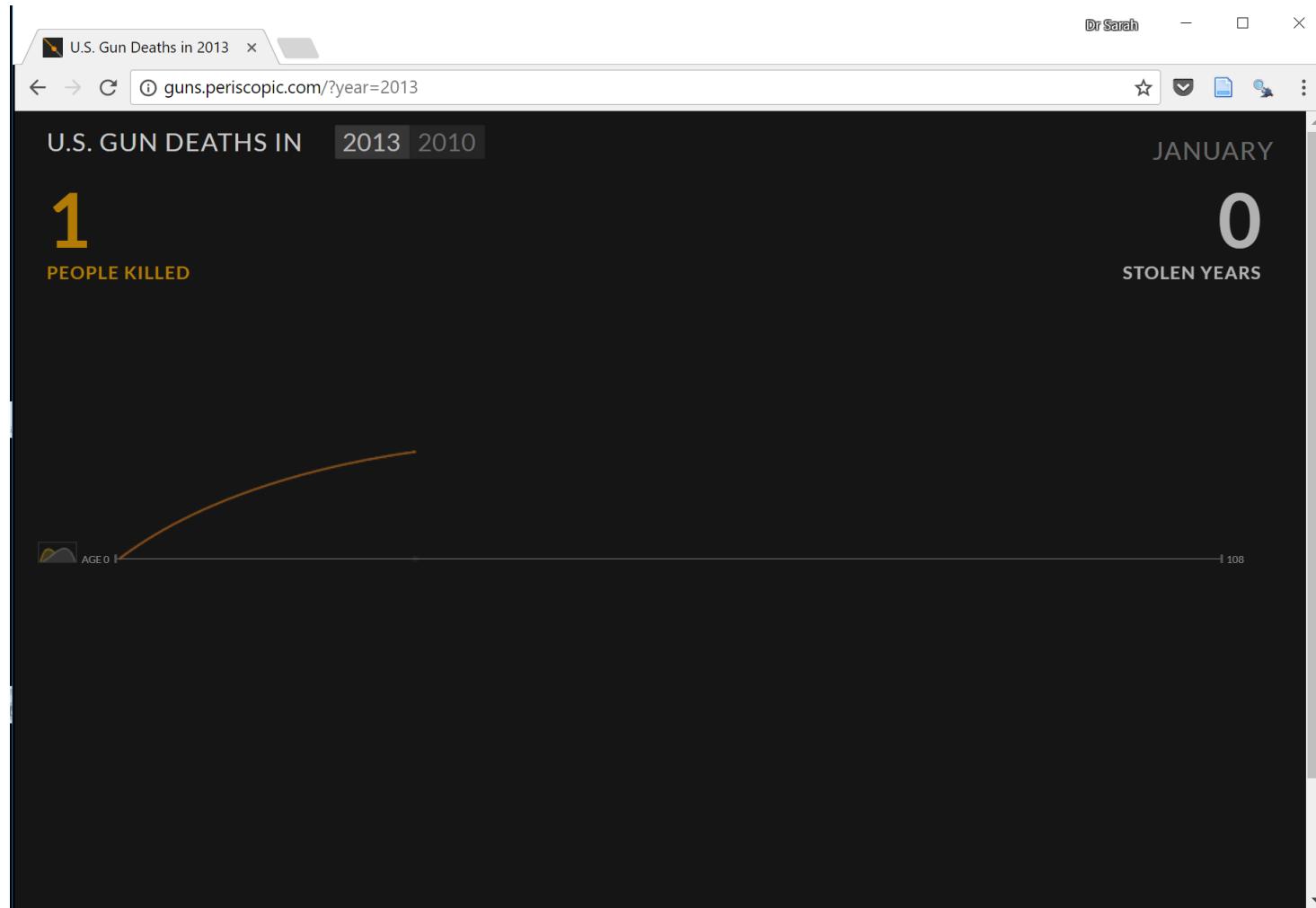


PICTURE SUPERIORITY EFFECT

About 440,000 Americans die each year from diseases related to smoking.
90% of them started as teen smokers.



GUN DEATHS IN THE USA



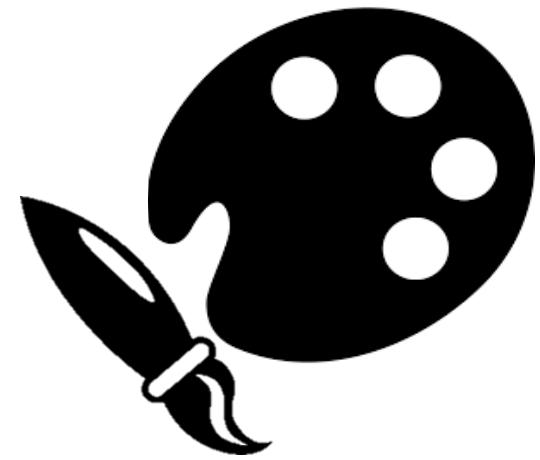
<http://guns.periscopic.com/>

ELEMENTS OF DATA VISUALISATION

1 2 5
3 4 6
7 8 9

Data

Visualisation
Tools



Design
Principles

DATA

1 2 5
3 4 6
7 8 9

DATA TYPES

Numeric

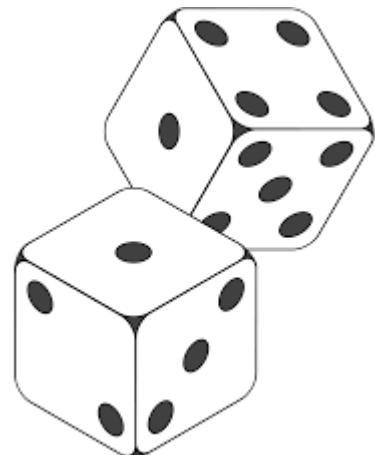
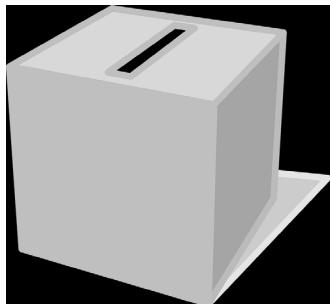
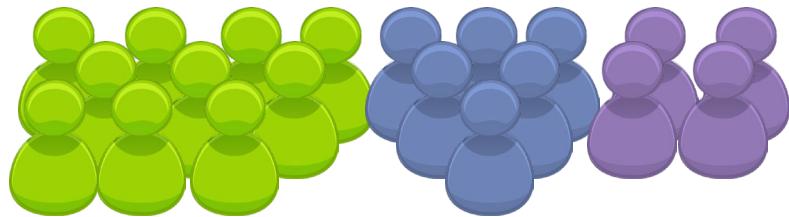
- Values are numbers
- Continuous or discrete

Categorical

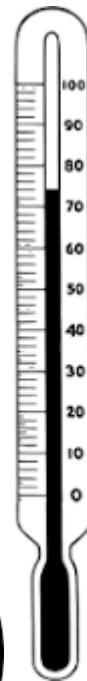
- Values are selected from a number of categories.
- Nominal or ordinal

NUMERICAL DATA

Discrete



Continuous



ACTIVITY: TYPE OF DATA

Discrete or Continuous?

- Number of people in each workshop today
- Result from rolling a dice
- Height of a participant in this workshop
- Time taken to get to AUT this morning
- Number of cats at home
- Age of a participant in this workshop

CATEGORICAL DATA

Nominal

No order

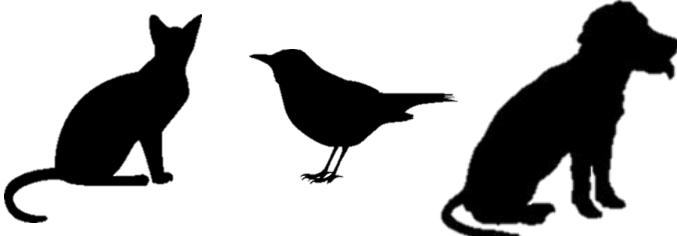
Home town



Hair Colour



Favourite Animal



Ordinal

Meaningful order



A+ A- A
B+ B B- ...

Strongly disagree
Disagree
Neither agree nor disagree
Agree
Strongly agree

ACTIVITY: COLLECTING DATA

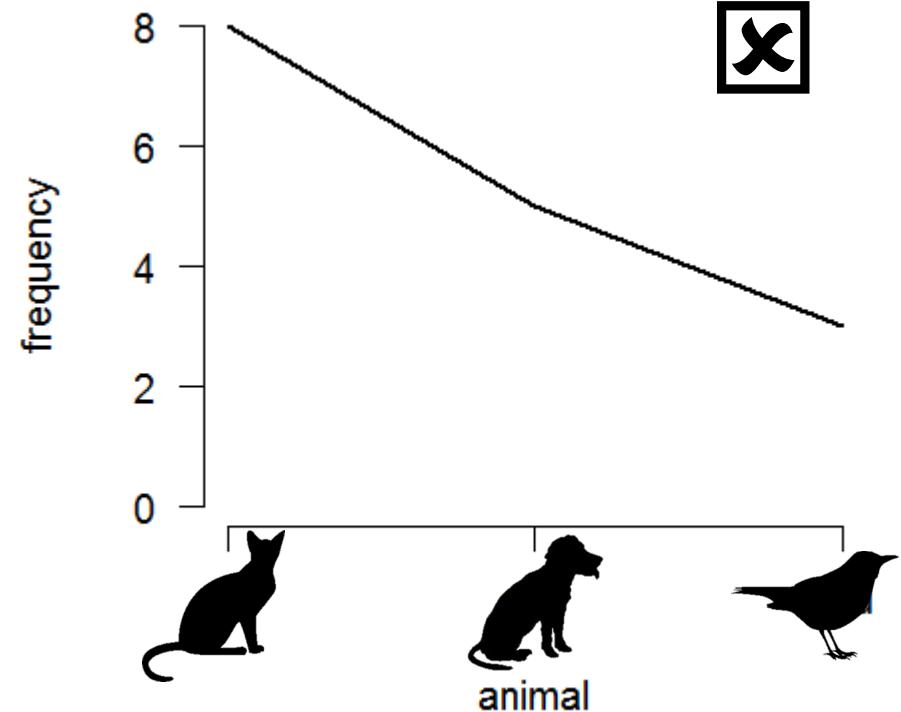
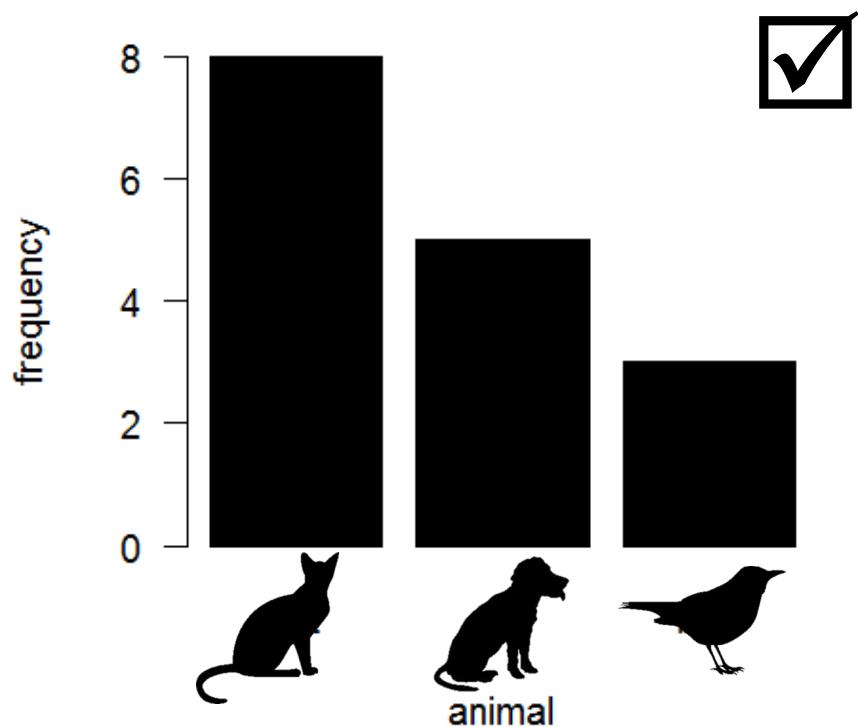
bit.ly/3B1logS



VISUALISATION TOOLS



DISCRETE



CONTINUOUS (TRENDS OVER TIME)

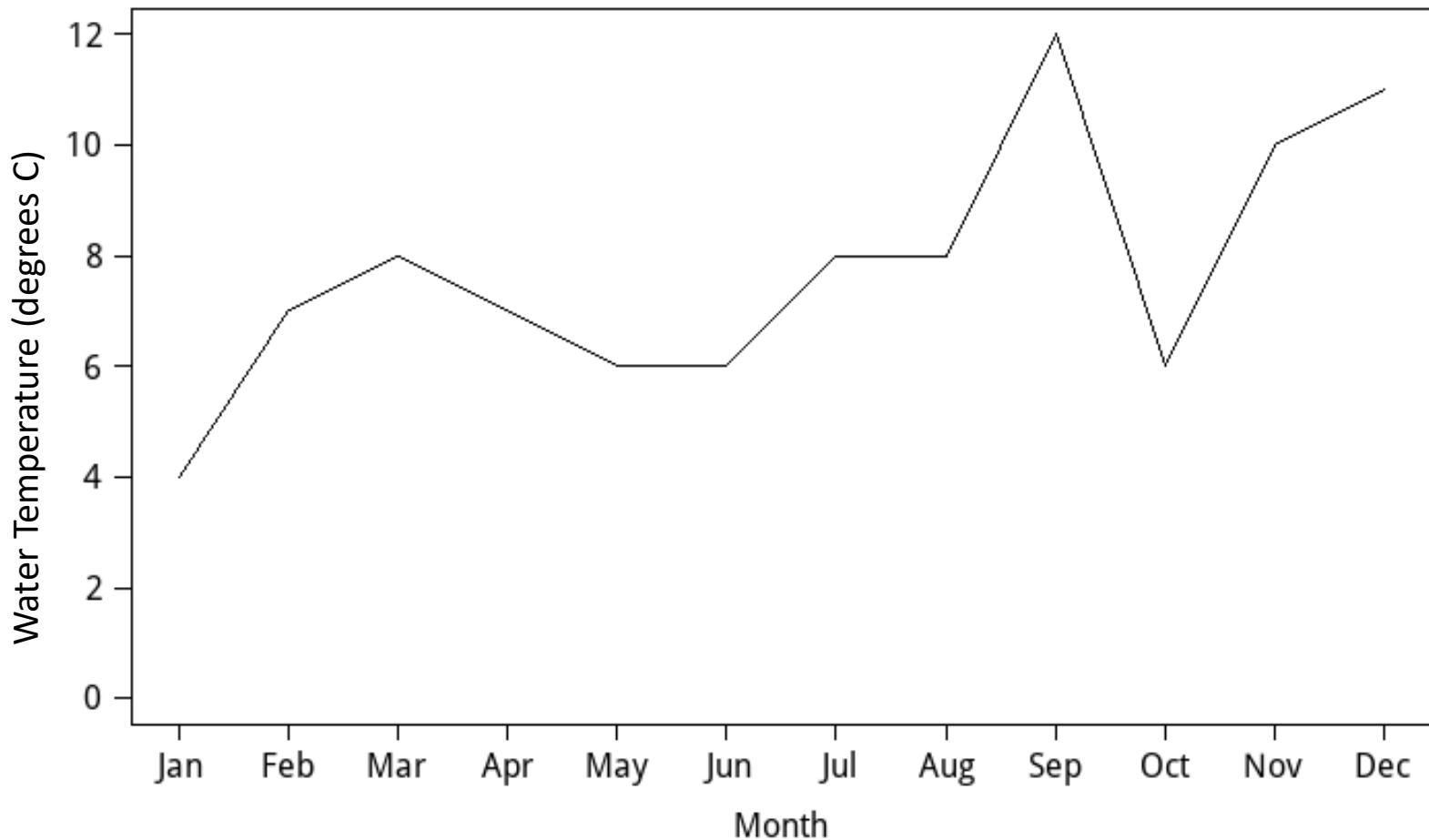
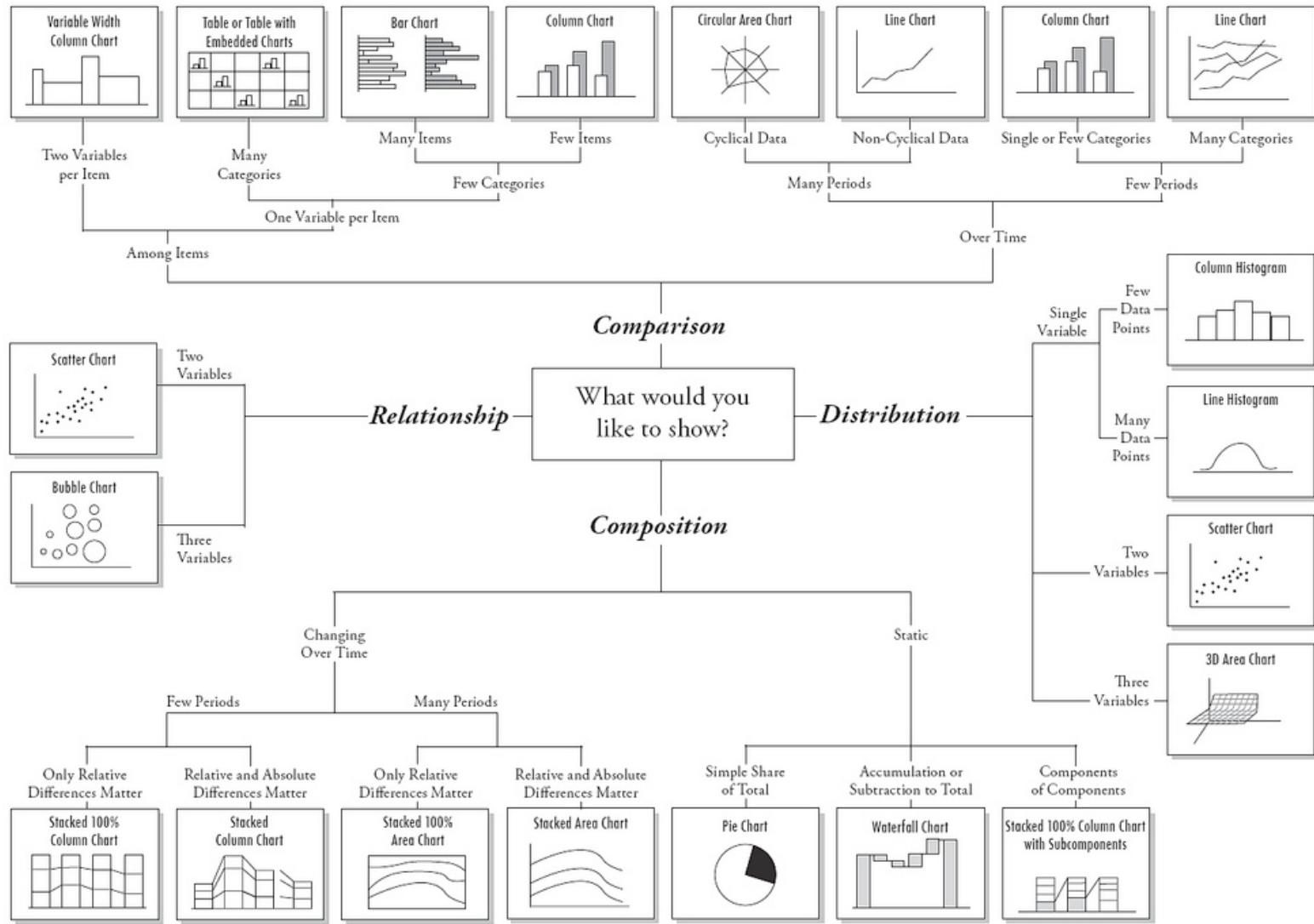


Chart Suggestions—A Thought-Starter



VISUALISATION SOFTWARE

Excel

Google Sheets

R (including Shiny)

Power BI

Tableau

Google Trends

CODAP

... and others



Caution: default options not necessarily the best options

DESIGN PRINCIPLES



DESIGN PRINCIPLES

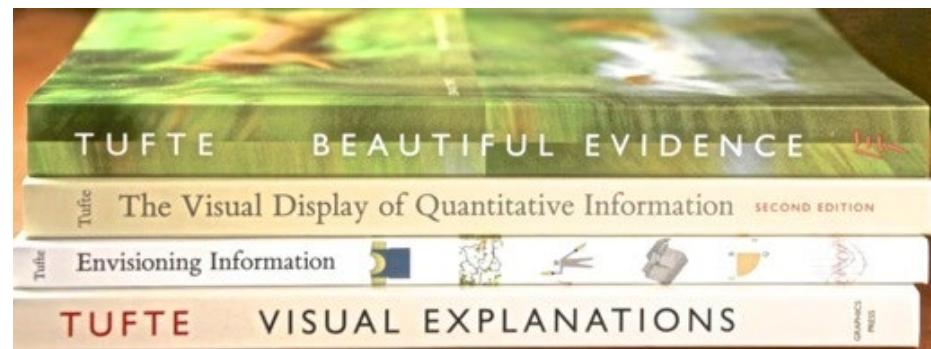
Graphical Integrity

Display data accurately

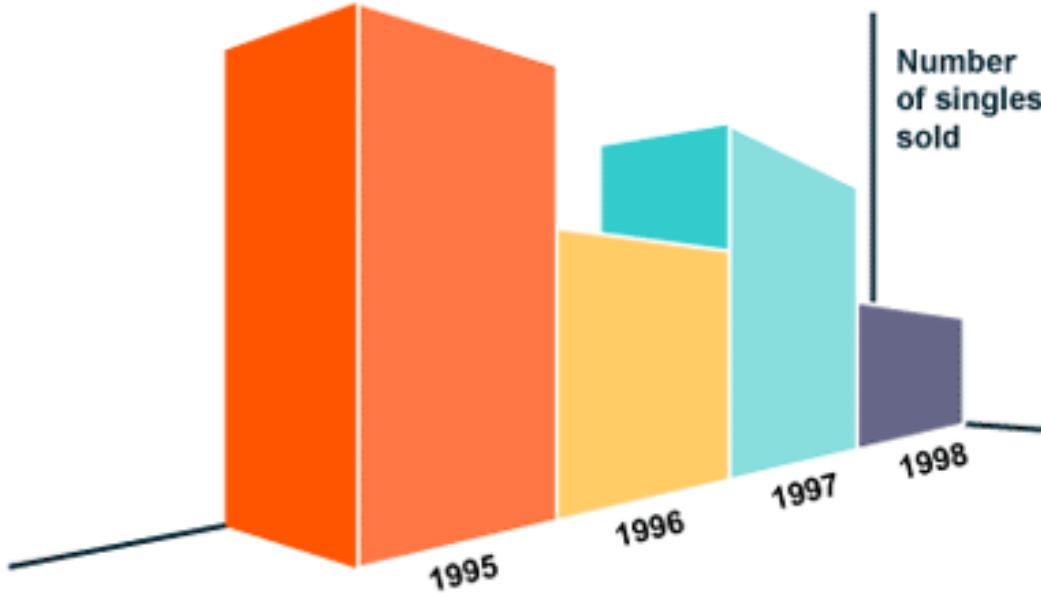
Graphical Excellence

Display data effectively and clearly

Edward Tufte

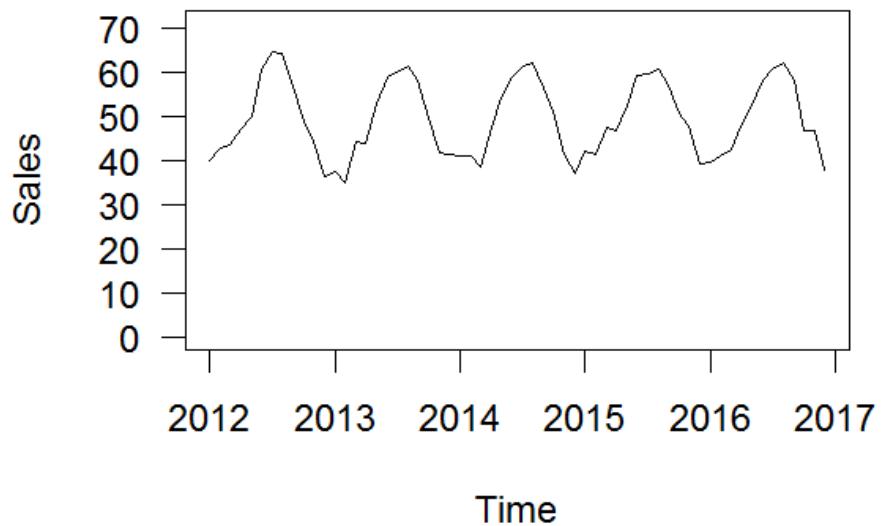
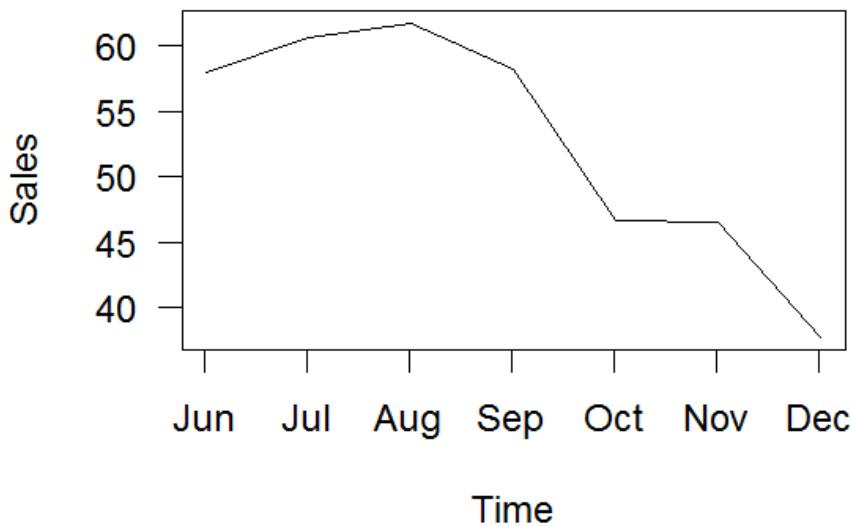


PROPORTION



Sales 1995 > 1997?

CONTEXT & NUMERICAL SCALE



Woollen Blanket Sales

Everything should be made as
SIMPLE
as possible, but not
SIMPLER.

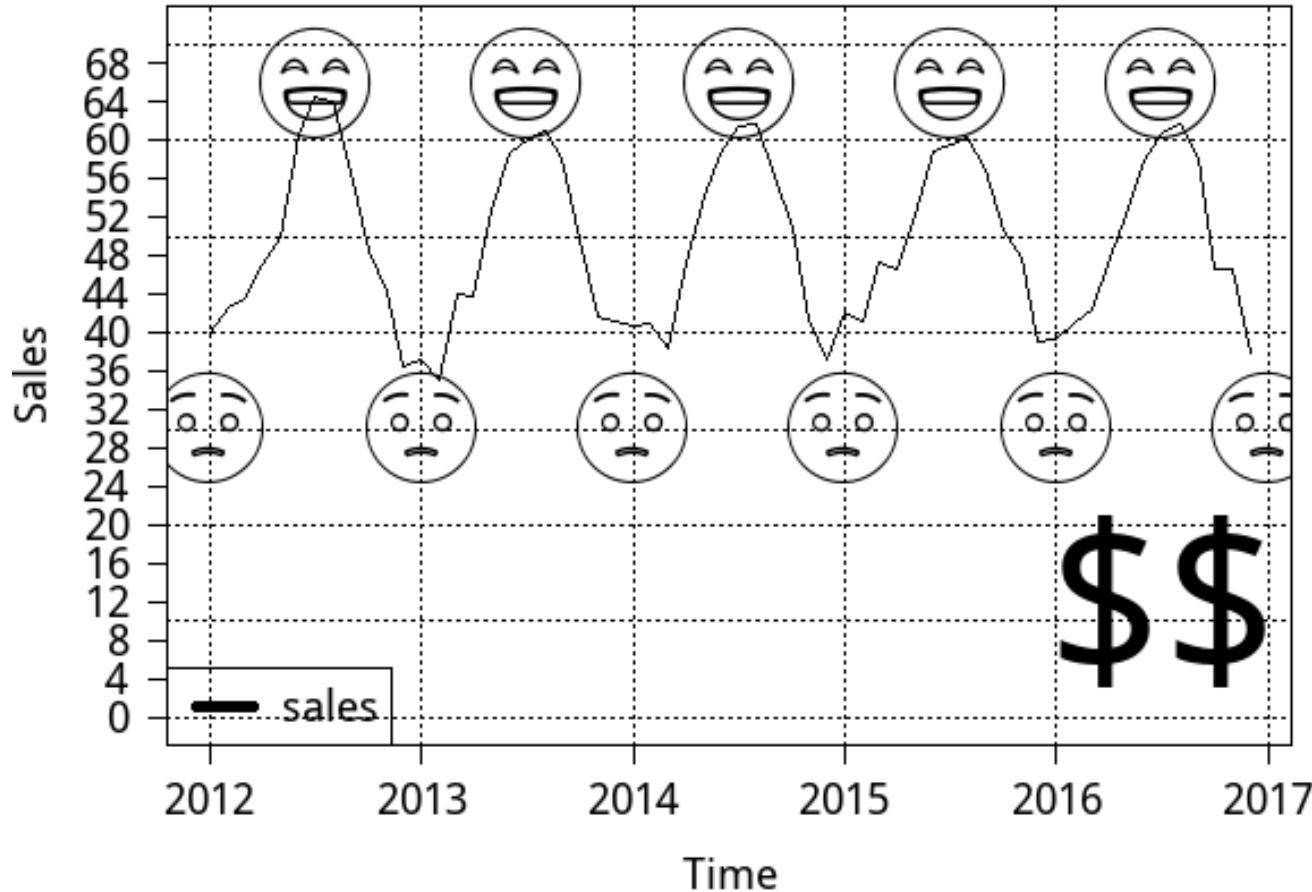
Albert Einstein

MAXIMISE DATA INK RATIO

Data Ink Ratio = $\frac{\text{Data Ink}}{\text{Total Ink Used in Graphic}}$

Minimise “Chart Junk”

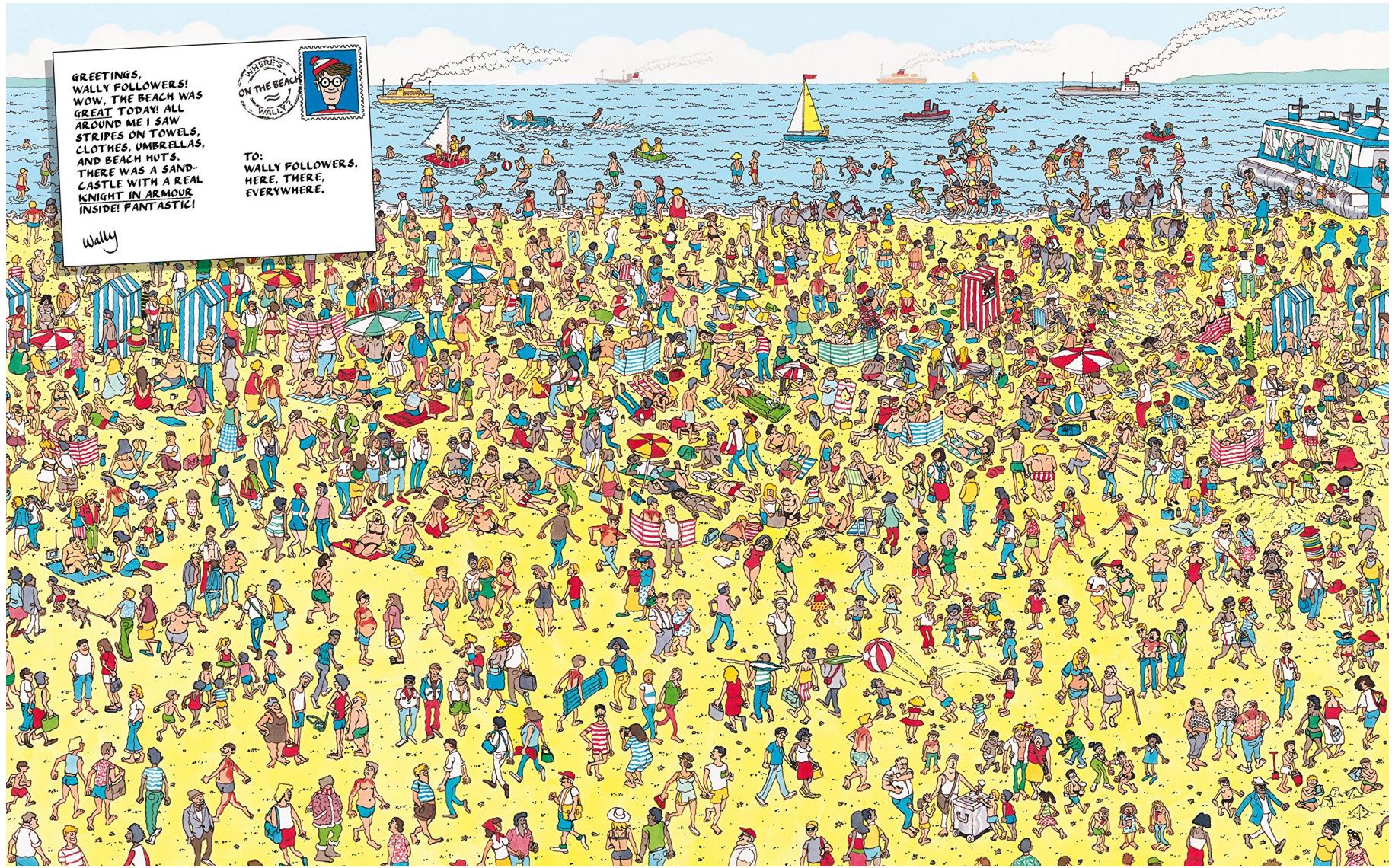
CHART JUNK



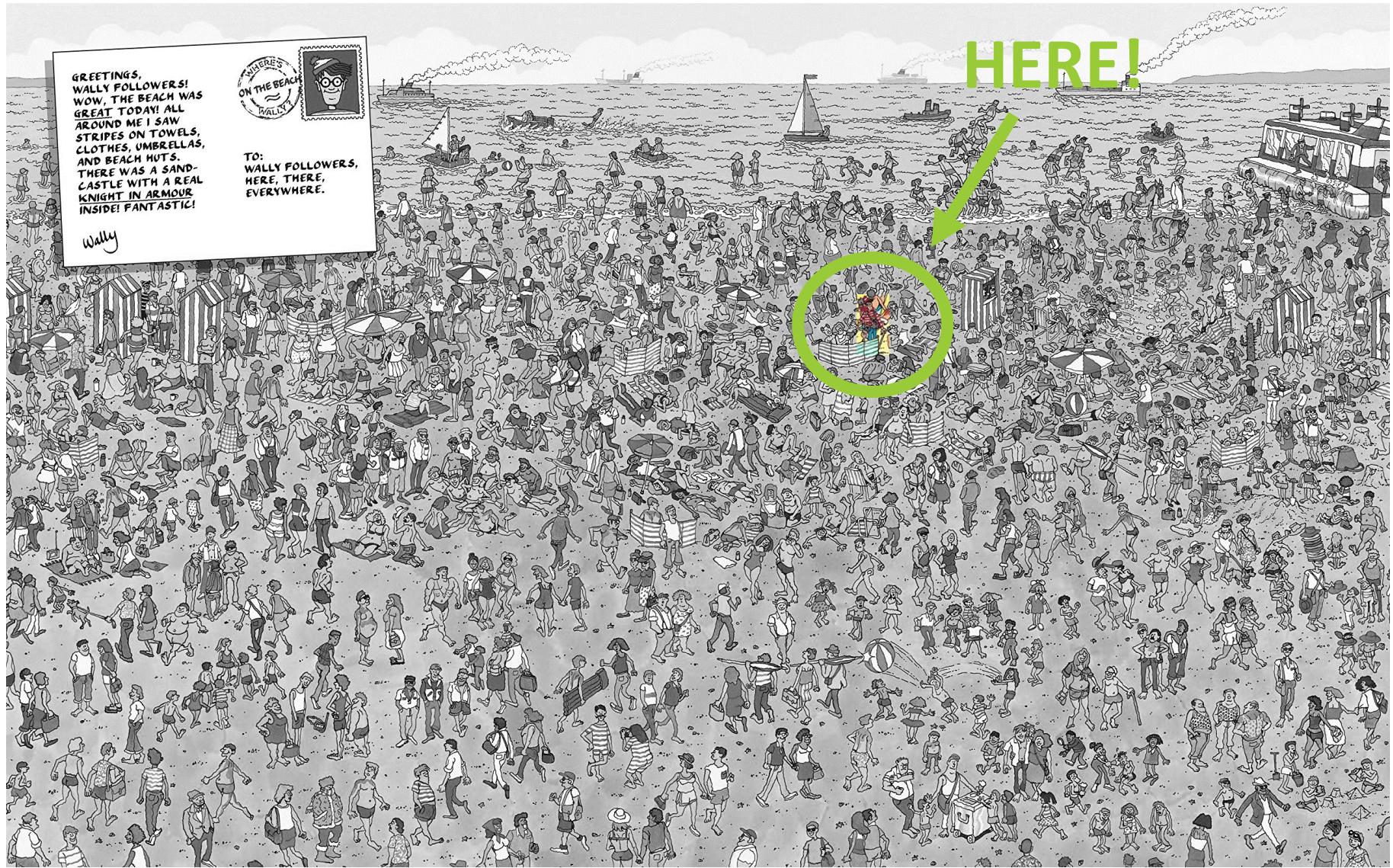
WHERE'S WALLY?



WHERE'S WALLY?



WHERE'S WALLY?



HERE!

HOW MANY THREES?

How MANY THREES?

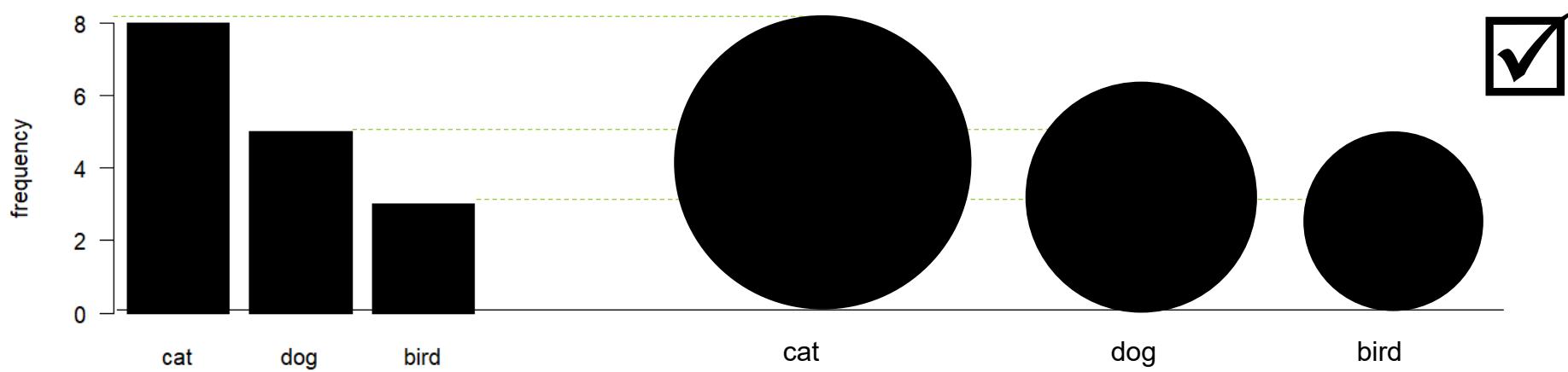
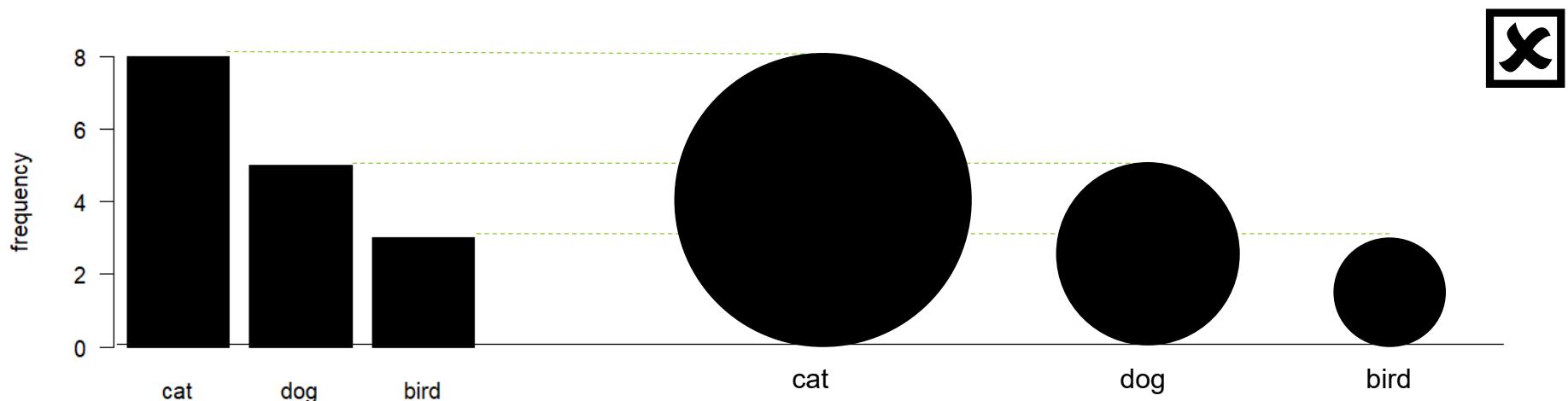
19131862439241485
22327154403504531
68885319313197596
1424215384486

How MANY THREES?

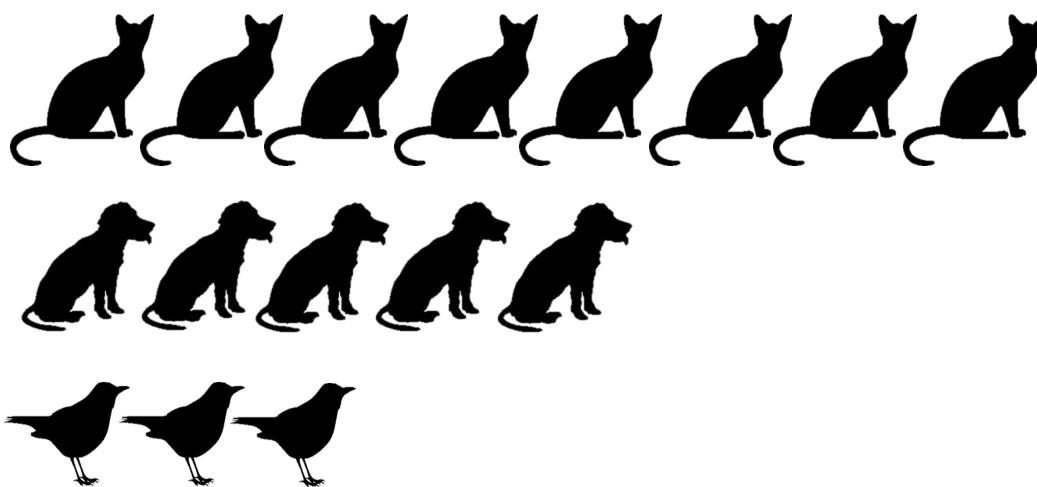
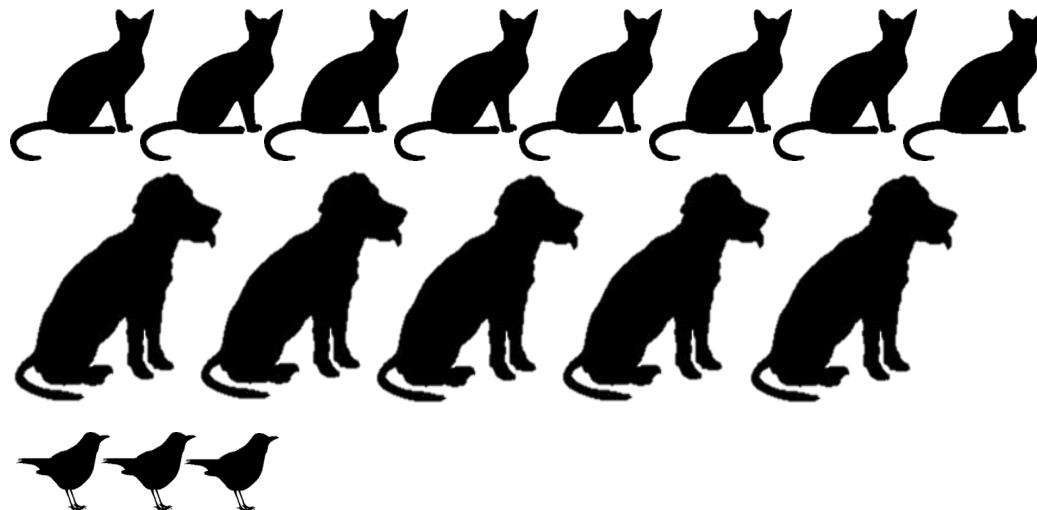
191**3**18624**3**9241485
22**3**27154403504531
68885**3**19313197596
1424215**3**84486

SCALE

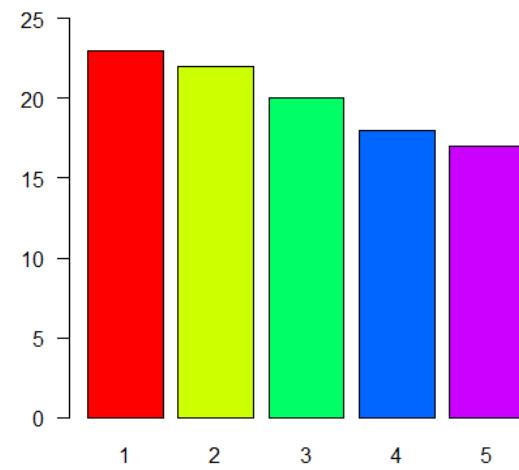
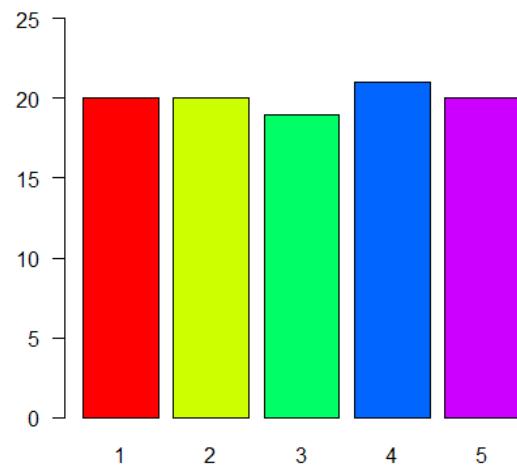
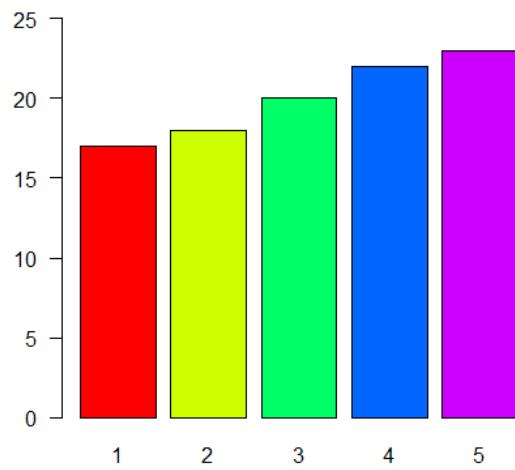
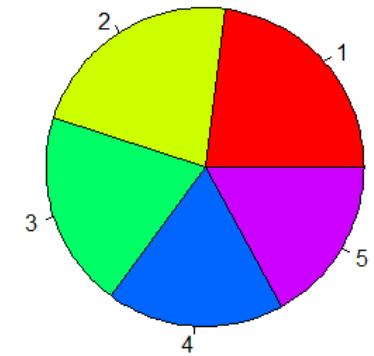
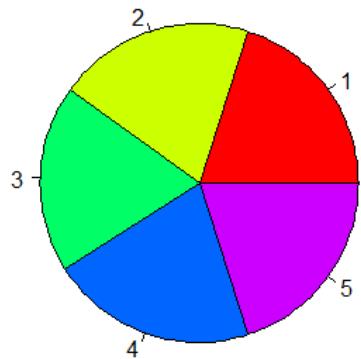
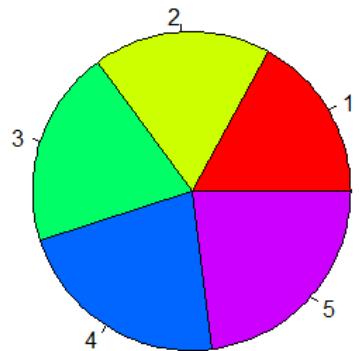
Frequency proportional to area not bar height



SCALE



BARS > PIES



ELEMENTS OF DESIGN

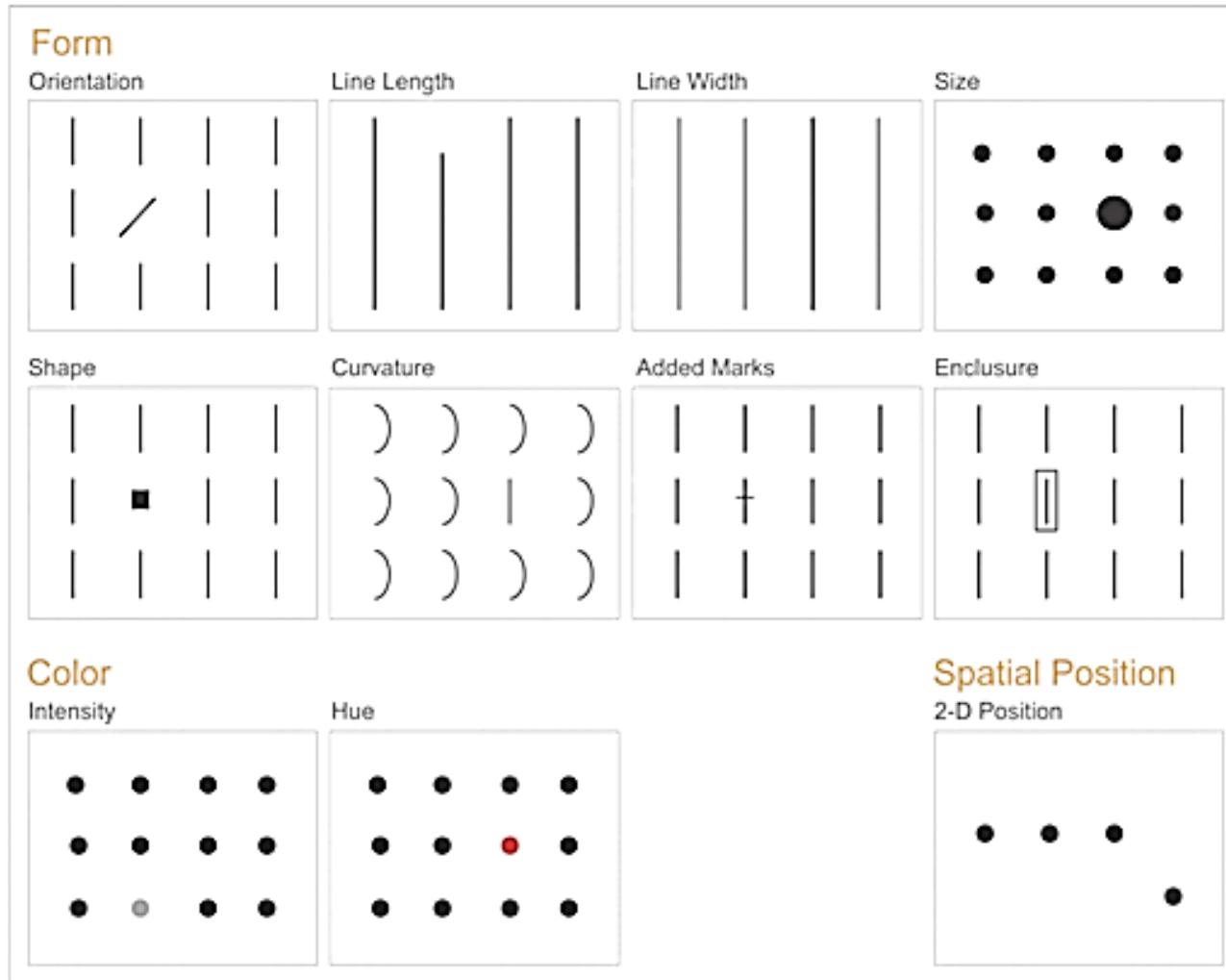
Colour

Contrast

Scale

- Impact effectiveness of visualisations
- Use thoughtfully to convey meaning and highlight key points

PRE-ATTENTIVE ATTRIBUTES



KNOW YOUR WHY, WHO & WHAT

Why? Know your reason

Who? Know your audience

(user stories can be helpful)

What? Know your message

(write your story before creating the visuals)

Source: Kat Greenbrook, Data Storyteller and Founder of Rogue Penguin

https://www.linkedin.com/posts/katgreenbrook_an-introduction-to-data-communication-activity-6990492611547475968-W0IH?utm_source=share&utm_medium=member_desktop

REASONS TO VISUALISE

Discover

Audience – usually only you

Inform

Audience – subject expert

Example: Dashboard

Educate

Audience – knowledge seeker

Example: Infographic

Key message should be very obvious

Source: Kat Greenbrook, Data Storyteller and Founder of Rogue Penguin

https://www.linkedin.com/posts/katgreenbrook_an-introduction-to-data-communication-activity-6990492611547475968-W0IH?utm_source=share&utm_medium=member_desktop

AUT BSc REGULATIONS

Bachelor of Science (BSc) Regulations

AK1041

which incorporates the following exit qualification:

Diploma in Applied Science (Level 6) (DipAppSc)

AK3383

These regulations are to be read in conjunction with all other relevant statutes and regulations, including the General Academic Statute and General Academic Regulations.

Completion Requirements

- 1 To qualify for the Bachelor of Science, a student must achieve 360 points, by successfully completing the following:
 - (a) core courses: 120 points with 60 points at level 5, 30 points at level 6 and 30 points at level 7, and
 - (b) a major: 120 points that satisfies the requirements of one of the majors identified in these regulations, and
 - (c) a further 120 points from any one of the following:
 - (i) a second major that satisfies the requirements of one of the majors identified in these regulations; or
 - (ii) two minors of at least 60 points each that satisfy the requirements of each minor as identified in these regulations; or
 - (iii) one minor of at least 60 points that satisfies the requirements of a minor identified in these regulations, and 60 points from courses listed in the Bachelor of Science Table, or any other bachelor's degree offered by the University.
- 2 A student taking a second major may hold up to 30 points in common between the two majors.
- 3 A course completed as part of the core or major requirements shall not be credited towards meeting the requirements of a student's minor.
- 4 A student would normally be expected to achieve the required points specified at each level of study in these regulations before being permitted to enrol in courses at subsequent levels.
- 5 To qualify for the Diploma in Applied Science (level 6), a student must achieve 240 points with at least 75 points at level 6 or above.
- Additional Majors and Minors
- 6 A student may choose an additional major to meet the requirements in 1(c)(i) above or a minor(s) to meet the requirements of 1(c)(ii) or (iii) above. The completion requirements for an additional major or minor are set out in the Additional Majors regulations and Minors regulations on page 561.

Conjoint Programmes of Study

7 A student who completes this degree as part of an approved conjoint programme of study will be exempted from the requirement in 1(c) above. The completion requirements are set out in the Conjoint Programmes of Study section of this Calendar.

Structure

Core Courses

15 points from	Level 5: SCIE504, DIGD507
and 15 points from	Level 5: ASTR500, BIOL500, BMED500, CHEM500, CHEM501, CHEM502, CONSS500, EASC500, ENV5502, ENV5503, FOOD502, FOOD503, GESC500, HEAL505, HEAL507, MAOH501, MASC500, MACS501, MEL5501, MICR501, PHYS501, SOSC585
and 15 points from	Level 5: COMP500, COMP503, MATH502, MATH504, MATH505, STAT500, STAT502
and 15 points from	Level 5: any level 5 course in the BSc Table of Courses.
and 15 points	Level 6: SCIE606
and 15 points from	Level 6: CHEM604, SCIE600
and 30 points	Level 7: SCIE707, SCIE708

Majors

Analytics

30 points	Level 5: MATH505, STAT502
and 45 points	Level 6: STAT603, STAT604, STAT605
and 45 points	Level 7: STAT704, STAT705, STAT706

Mathematical Modelling and Computation

30 points	Level 5: MATH505, STAT502
and 45 points	Level 6: MATH605, MATH606, MATH607
and 30 points	Level 7: MATH707, MATH708
and 15 points from	Level 7: MATH700, MATH709

**BSc Core
Papers**

**Requirements
for majors in
Analytics and
Mathematical
Modelling and
Computation**

AUT BSc ANALYTICS MAJOR

Level 7	SCIE707 Project part 1	SCIE708 Project part 2	STAT704 Industrial & Business Analytics	STAT705 Multivariate Data Analysis	STAT706 Stochastic Modelling	2 nd Major / Minor / Elective	2 nd Major / Minor / Elective	2 nd Major / Minor / Elective
Level 6	SCIE606 Vision Mātauranga: Science Practice in Aotearoa	SCIE600 Scientific Inquiry	STAT603 Forecasting	STAT604 Statistical Inference	STAT605 Statistical Data Analysis	2 nd Major / Minor / Elective	2 nd Major / Minor / Elective	2 nd Major / Minor / Elective
Level 5	DIGD507 Mahitahi Collaborative Practices Or SCIE504 Science and Society	Natural Sciences Elective	COMP500 Programming Concepts and Techniques*	MATH502 Algebra and Discrete Mathematics*	MATH505 Algebra and Calculus I	STAT502 Intro to Probability and Statistics	2 nd Major / Minor / Elective	2 nd Major / Minor / Elective

Core

Major

2nd Major / Minor / Elective

* Recommended paper

AUT BSc DOUBLE MAJOR:

ANALYTICS + MATHEMATICAL MODELLING AND COMPUTATION

Level 7	SCIE707 Project part 1	SCIE708 Project part 2	STAT704 Industrial & Business Analytics	STAT705 Multivariate Data Analysis	STAT706 Stochastic Modelling	MATH707 Mathematical Computation	MATH708 Modelling and Differential Equations II	MATH700 Financial Modelling or MATH709 Modelling Health, Biology
Level 6	SCIE606 Vision Mātauranga: Science Practice in Aotearoa	SCIE600 Scientific Inquiry	STAT603 Forecasting	STAT604 Statistical Inference	STAT605 Statistical Data Analysis	MATH605 Algebra and Calculus II	MATH606 Modelling and Differential Equations I	MATH607 Quantitative Decision Analysis
Level 5	DIGD507 Mahitahi Collaborative Practices or SCIE504 Science and Society	Natural Sciences Elective	COMP500 Programming Concepts and Techniques*	MATH502 Algebra and Discrete Mathematics*	MATH505 Algebra and Calculus I	STAT502 Intro to Probability and Statistics	Elective	Elective

Core

Major

2nd Major / Minor / Elective

* Recommended paper

KNOW YOUR AUDIENCE

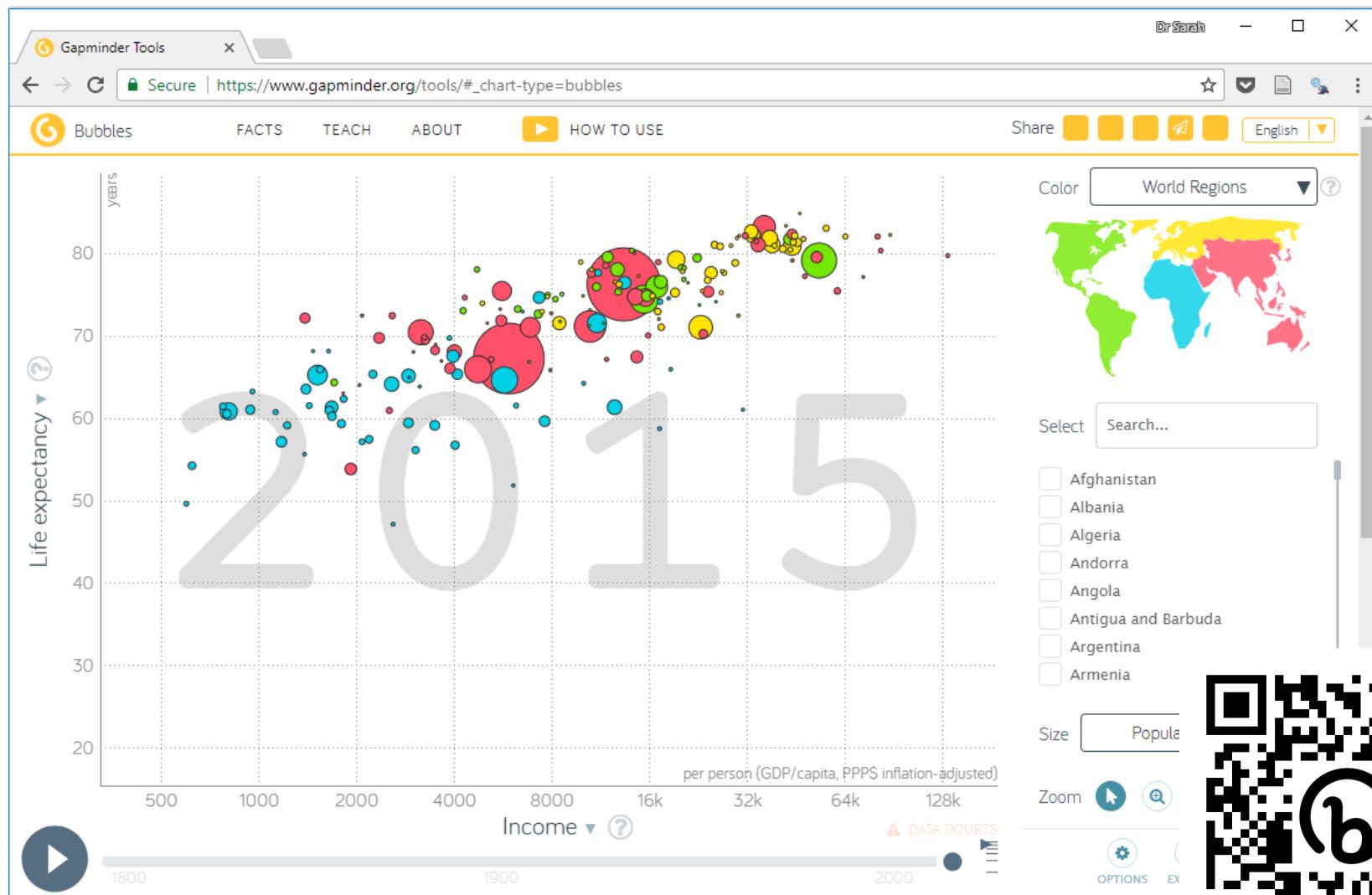
- Who is the audience for information about the BSc degree?
- What type of information/visualisation would be suitable for each audience?

**ENCOURAGING
CURIOSITY
AND
CRITIQUE**

**curiosity
is the most
powerful
thing you own**

James Cameron

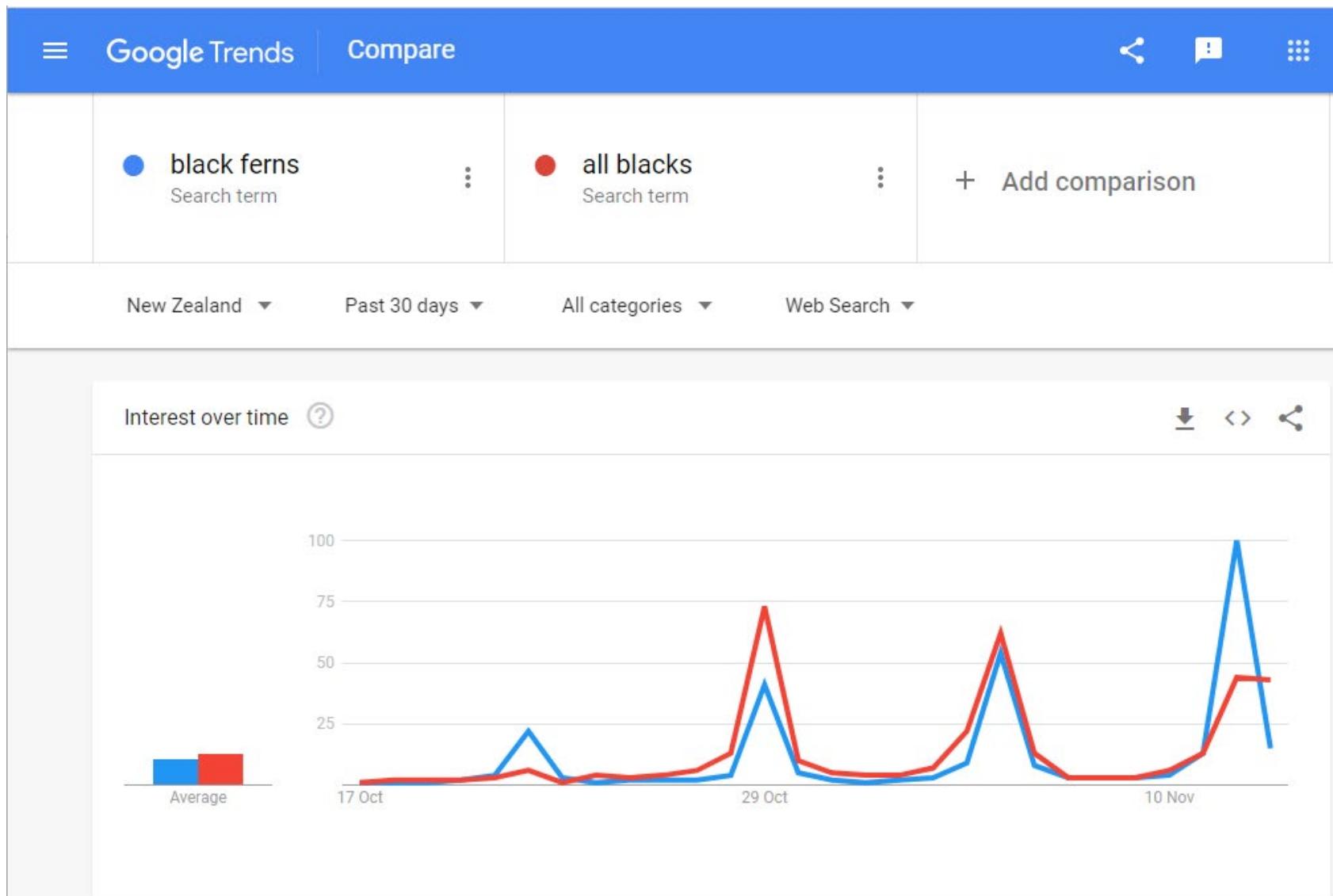
GAP MINDER



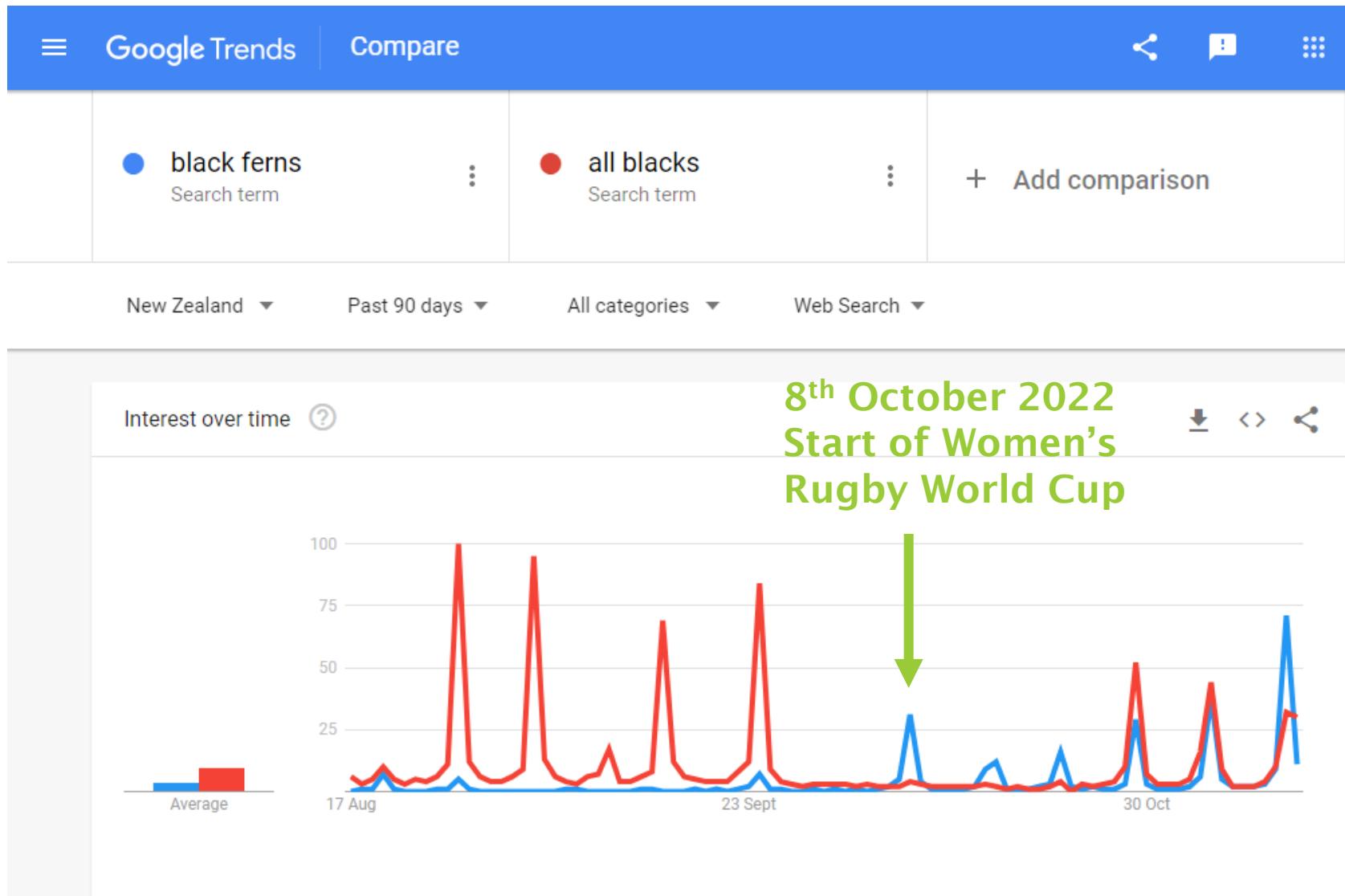
POPULARITY CONTEST: BLACK FERNS VS ALL BLACKS



POPULARITY CONTEST: BLACK FERNS VS ALL BLACKS



POPULARITY CONTEST: BLACK FERNS VS ALL BLACKS



ACTIVITY: GOOGLE TRENDS

<https://trends.google.com/trends/>

Use google trends to explore search terms of your choice

Explore any peaks or dips

Share your findings with your table

STATSNZ: MAPS



<https://statsnz.maps.arcgis.com/apps/MapSeries/index.html?appid=3a61db9ab16d435b91b3997ed3a4b634>

STATSNZ: VISUALISATIONS

Population counts, by age and sex

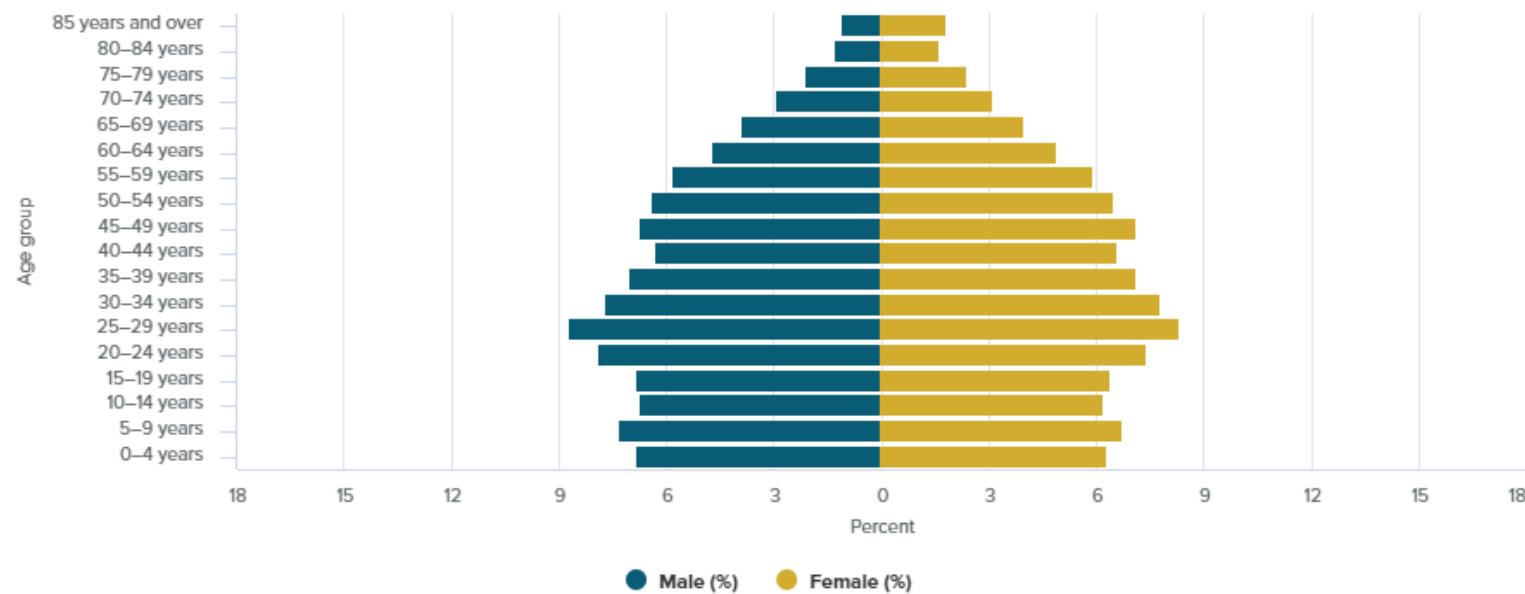
GRAPH

TABLE

METADATA

Age and sex of people in Auckland Region, 2018 Census

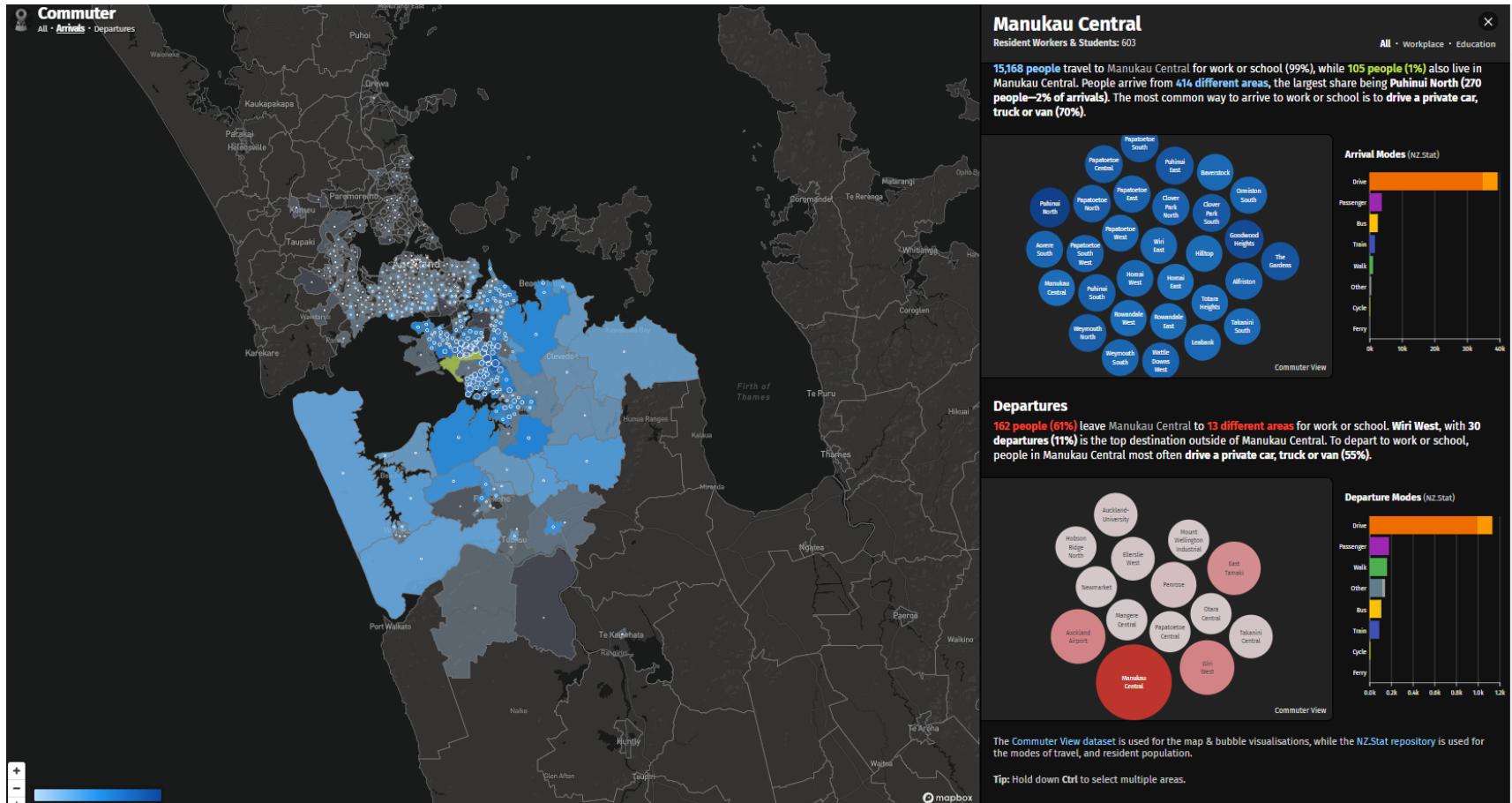
⋮



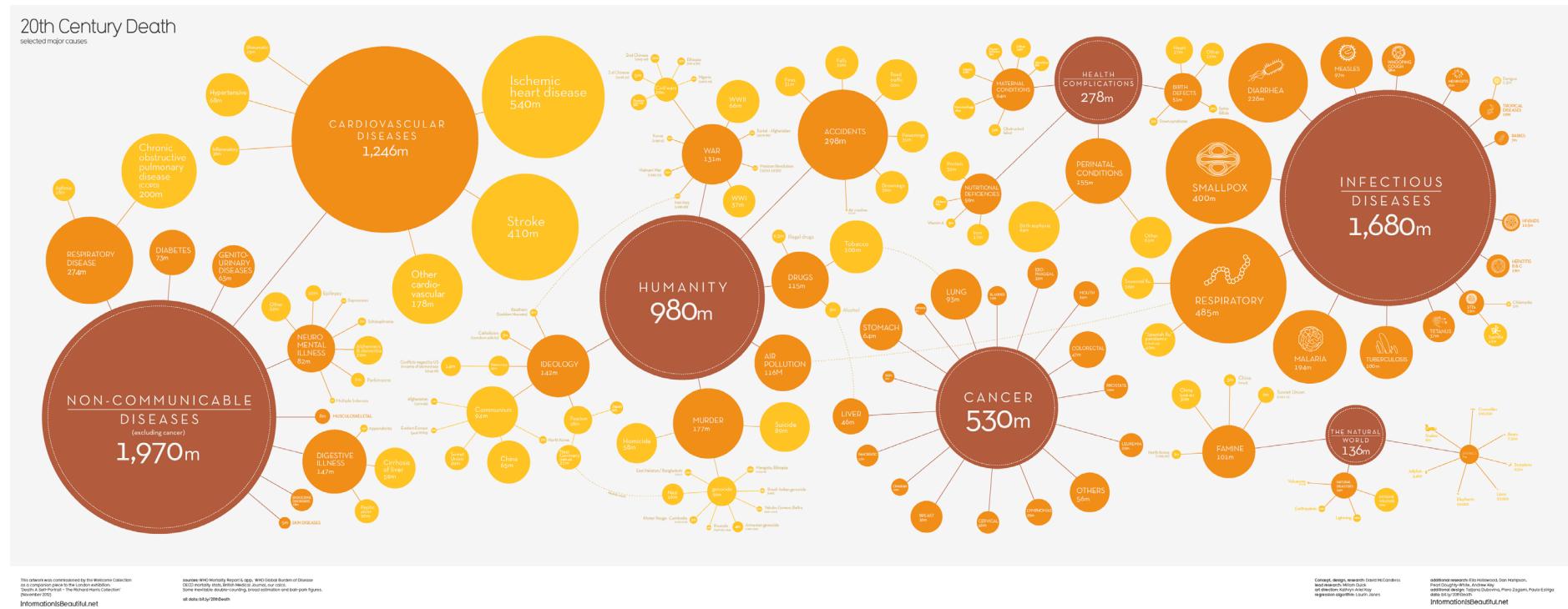
See metadata tab for information about variables and quality.

Stats NZ

STATSNZ: INTERACTIVE TOOLS



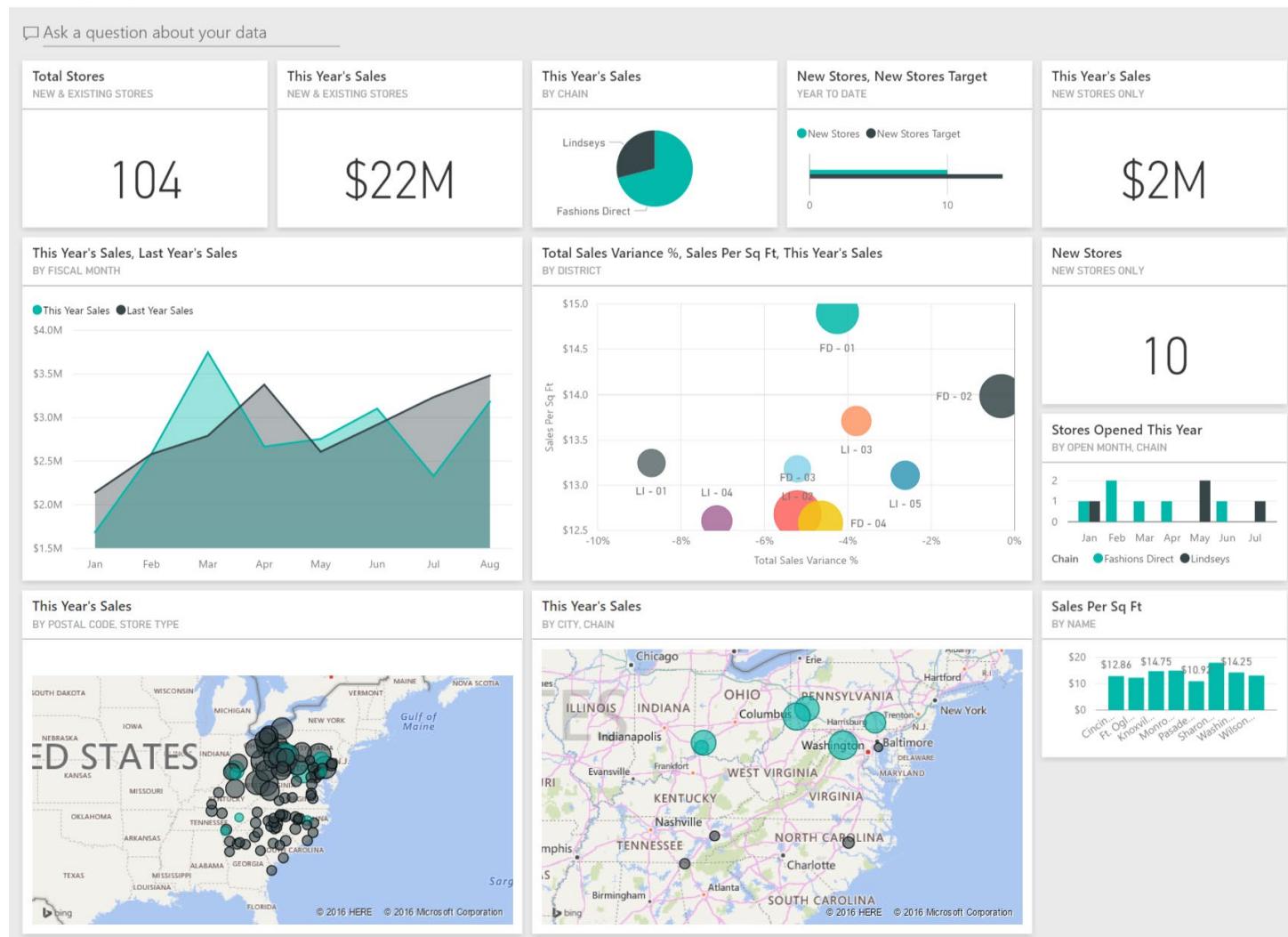
20TH CENTURY DEATH



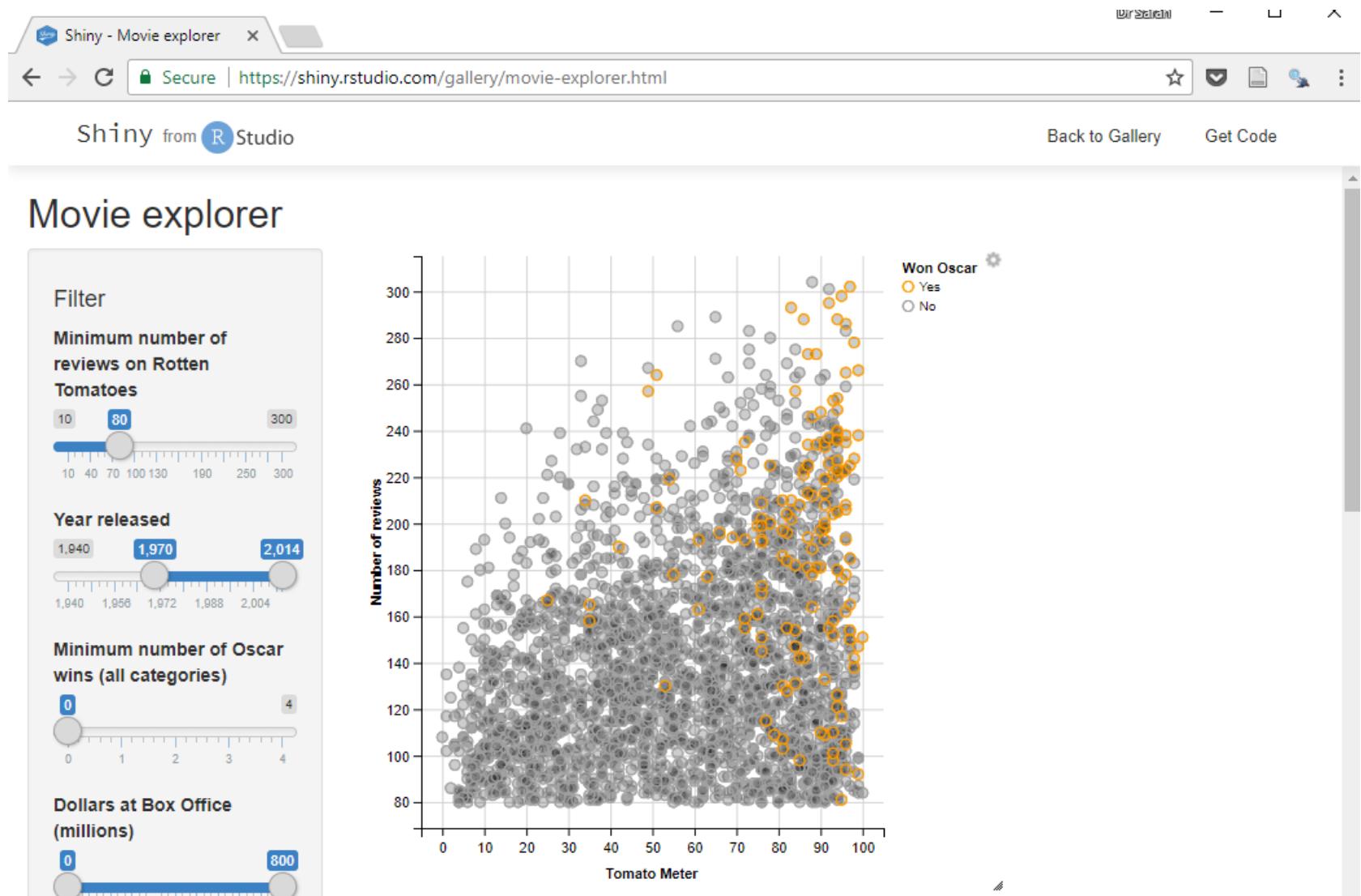
<http://www.informationisbeautiful.net/visualizations/20th-century-death/>

DASHBOARDS

Retail Analysis Sample



R SHINY WEB APP



<https://shiny.rstudio.com/gallery/movie-explorer.html>

ACTIVITY IDEA: CRITIQUE

Find a data visualisation online

What are the key messages?

What are the strengths/weaknesses of the data visualisation?

DATA STORIES

rogue

DATA VIZ

SCROLLING STORIES

A scrolling story is designed for digital reading but can easily be redesigned to fit a poster format. Because of their scrolling nature, they are best designed around a narrative.



New Zealand Bees



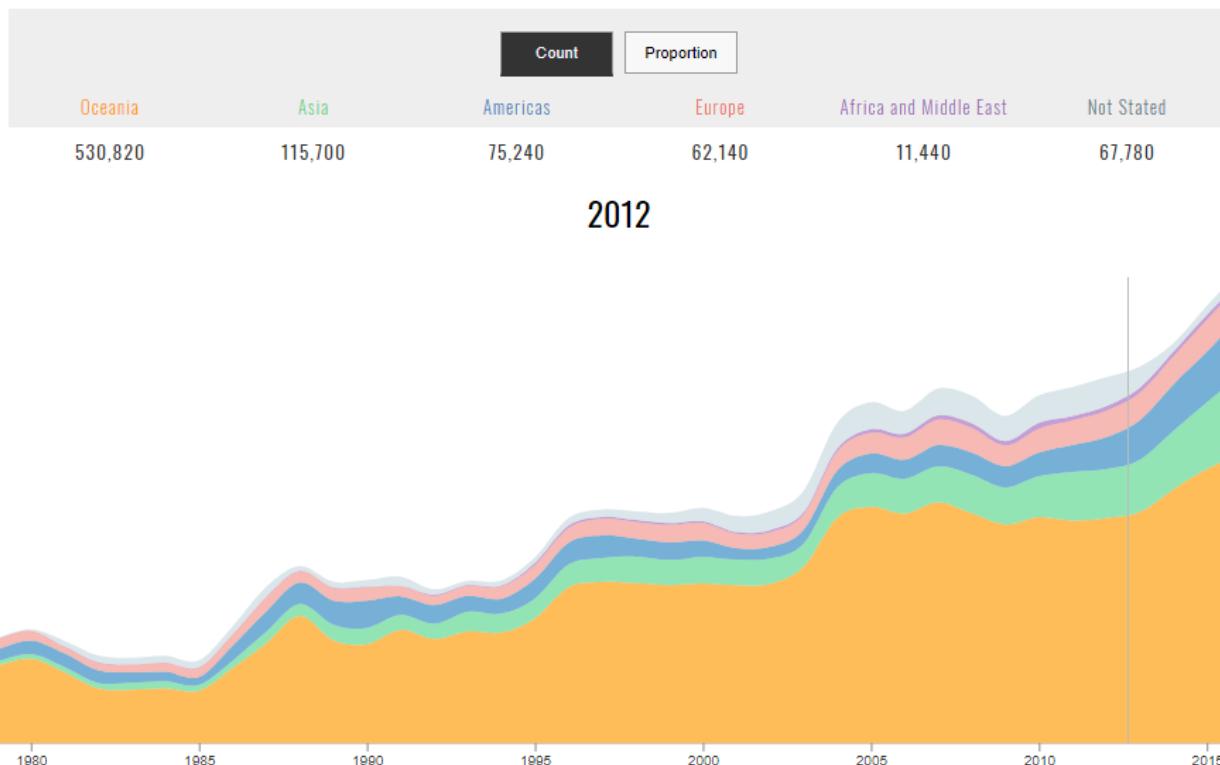
Measles in Aotearoa



Lost Boys

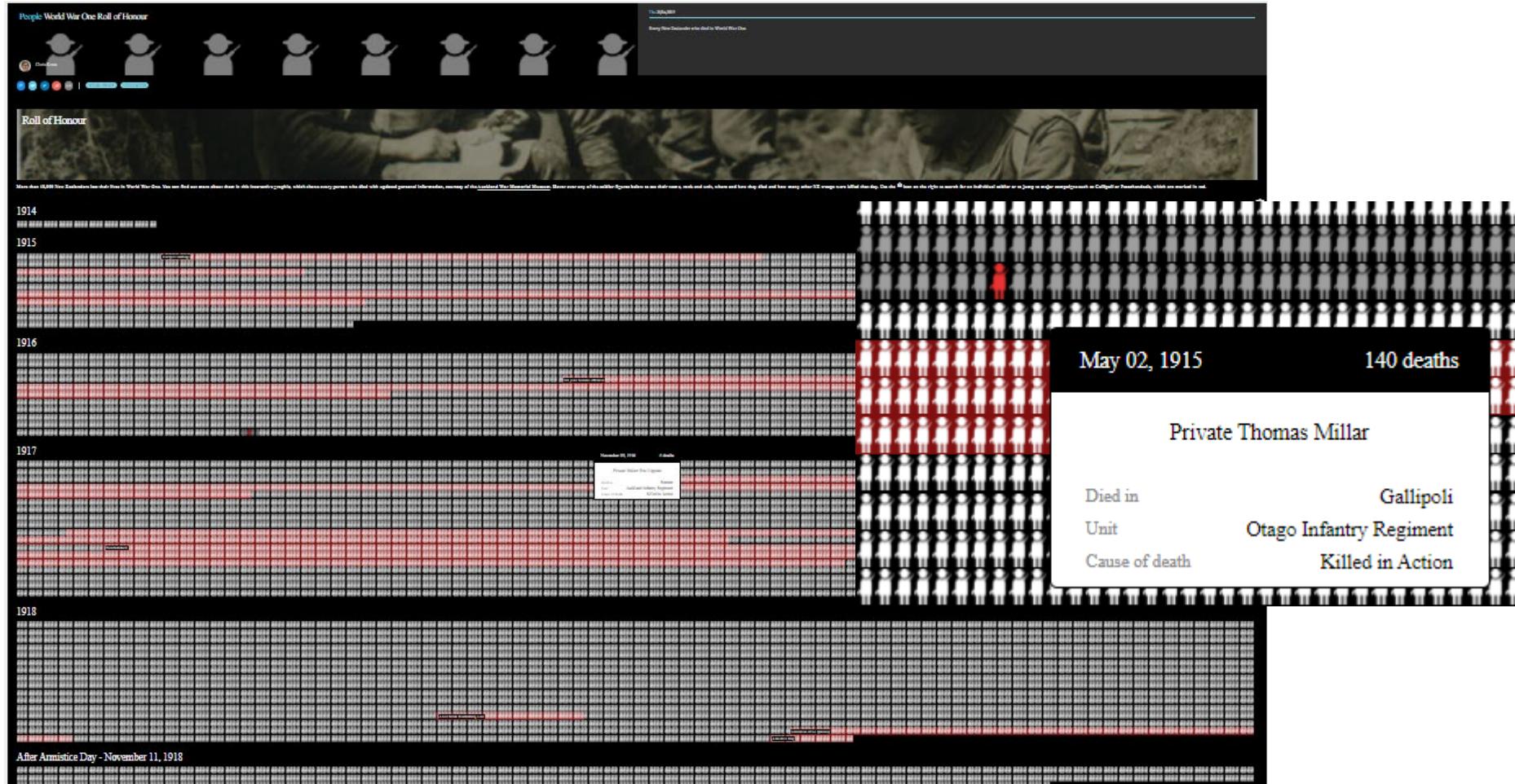
NZ HERALD: INSIGHTS

The proportion of Kiwi holidaymakers visiting Asia has increased significantly from **4.1% in 1979** to **16.8% in 2016**. Travel within the Oceania region still remains dominant.



Since 2000, Indonesia, Thailand and China has seen a massive boom in Kiwi visitors, rising from a combined proportional total of **4.6%** to **8.7% in 2016**.

NZ HERALD: INSIGHTS



ACTIVITY IDEA: INFOGRAPHICS

Choose a social issue of interest to you

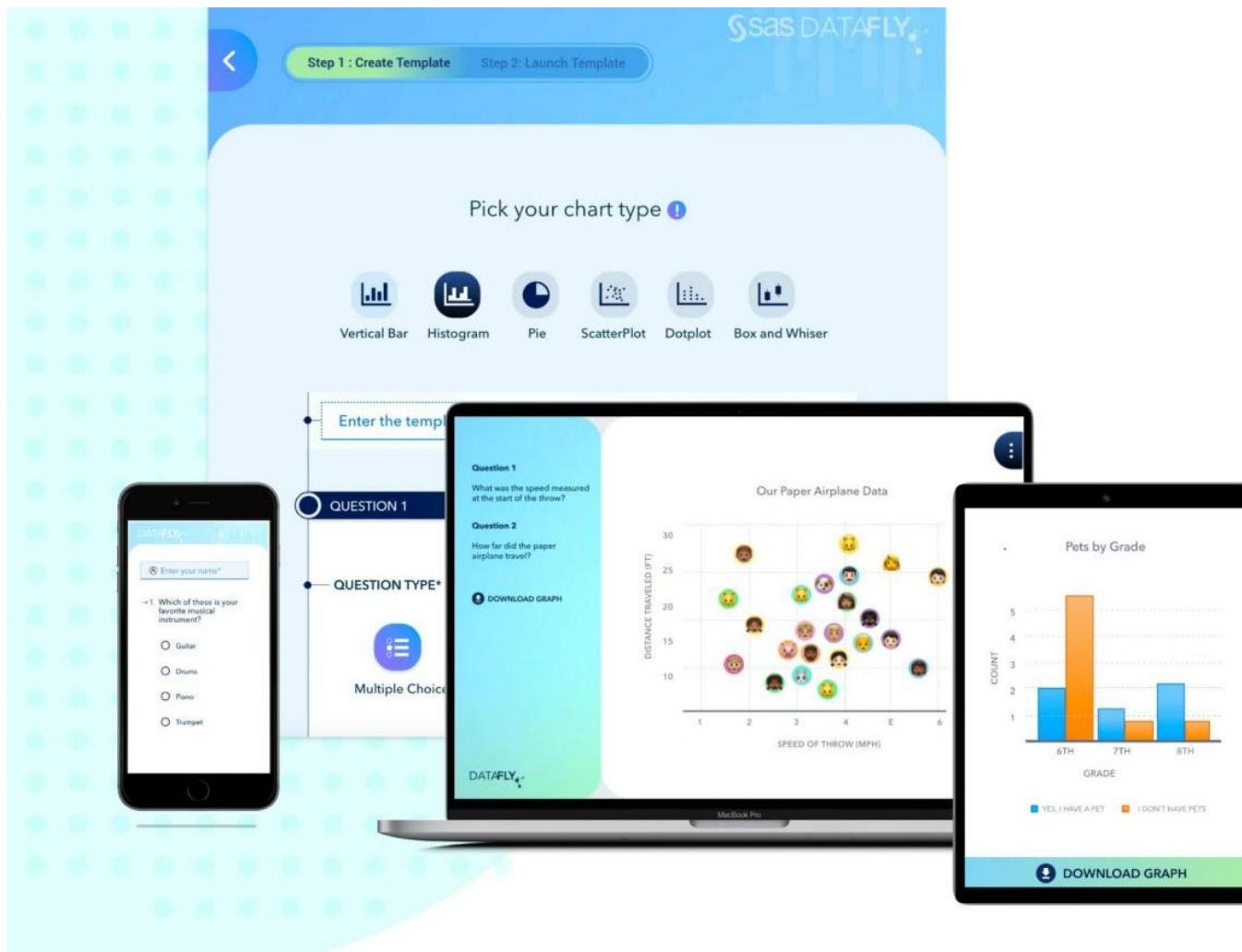
Create an infographic to explain the key issues surrounding this issue

Use a mixture of graphs, text and images

Tell me and I forget.
Teach me and I remember.
Involve me and I learn.

Benjamin Franklin

SAS DATAFLY



<https://curiosity.sas.com/app/datafly> + live session code

ACTIVITY IDEA: DATA COLLECTION & VISUALISATION

- Create a template at
<https://curiosity.sas.com/app/datafly>
- Ask students to add their responses
- Discuss the results

SAS CURIOSITY



GEN Z

GROWING UP IN THE GAP

One woman's look at data visualizations to see the lifetime impact of gender pay inequality



SOCIAL GOOD

THE UNCOUNTED COMMUNITY

Why missing LGBTQ data means missing out on the power to create change.



NATURE

DATA DEEP DIVE

For the Healthy Reefs for Healthy People Initiative, collecting data underwater is challenging but rewarding – and necessary to protect our oceans and the communities that depend on them.



GEN Z

THE ZOOMER SKILLS THAT (WILL) PAY THE BILLS

We asked SAS interns, in roles ranging from R&D to marketing and more, what skills they found



SOCIAL GOOD

FOOD FOR THOUGHT

Hunger Free America turned to data visualizations to show the impact of the social safety net on food insecurity.



SOCIAL GOOD

WHEN YOU WISH UPON YOUR DATA

See how Make-A-Wish Eastern North Carolina took a deeper look at a variety of factors to deliver more life-changing experiences for children.

ACTIVITY: DATA STORYTELLING

- Groups of 3 – 4
- Choose one of the following tools and find or create a visualisation:
 - Google Trends
 - StatsNZ Maps
 - FigureNZ
 - GapMinder
- Add your visualisation to a shared document (e.g. google slides)
- Each group 2 min to explain key messages

DATA SOURCES

- Figure NZ
- OzDASL - Australasian Data and Story Library
<http://www.statsci.org/data/>
- The Data and Story Library <http://lib.stat.cmu.edu/DASL/>
- Google Trends
- StatsNZ and the IDI <https://www.stats.govt.nz/integrated-data/integrated-data-infrastructure/>
- Census at Schools
<https://new.censusatschool.org.nz/explore/>
- NZ Govt <https://www.data.govt.nz/>
- <http://opendata.nz/>
- Kaggle
- ... and many more

INTERESTING VISUALISATIONS I

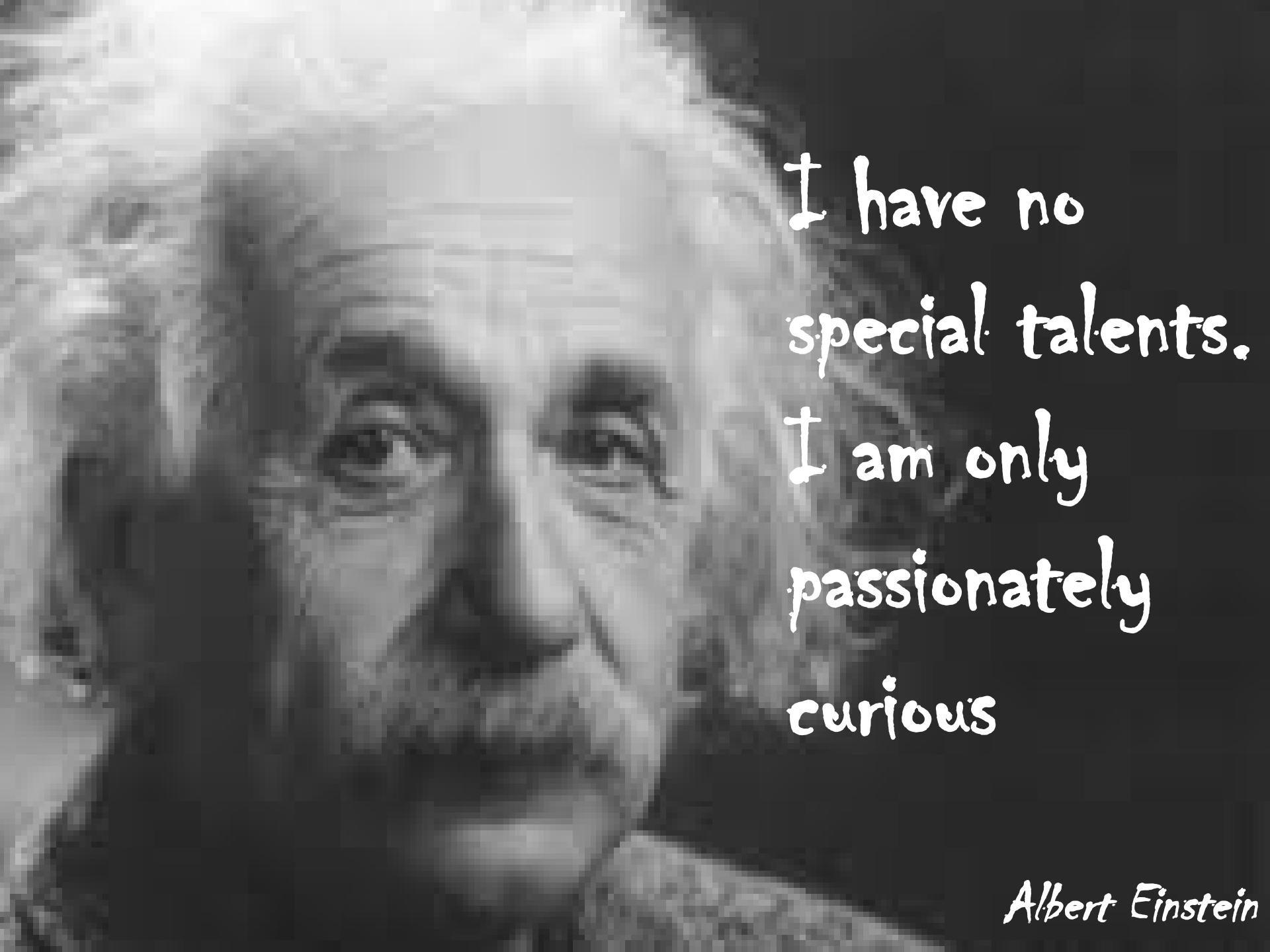
- FigureNZ <https://figure.nz/>
- NZ Herald Insights <http://insights.nzherald.co.nz/>
- Information is Beautiful
<http://www.informationisbeautiful.net/visualizations>
- Rogue Penguin <https://www.roguepenguin.co.nz/>
- The Guardian
<https://www.theguardian.com/news/datablog+technology/data-visualisation>
- The Economist
<https://www.economist.com/blogs/graphicdetail>
- Power BI <https://powerbi.microsoft.com/en-us/documentation/powerbi-sample-retail-analysis-take-a-tour/>

INTERESTING VISUALISATIONS II

- Gap Minder e.g. <http://bit.ly/2qo6rOT>
- Google Trends <https://trends.google.com/trends>
- StatsNZ Visualisations
 - <https://statsnz.maps.arcgis.com/apps/MapSeries/index.html?appid=3a61db9ab16d435b91b3997ed3a4b634>
 - <https://www.stats.govt.nz/tools/2018-census-place-summaries/auckland-region>
 - <https://commuter.waka.app/>
- R Shiny <https://shiny.rstudio.com/gallery/movie-explorer.html>
- SAS Curiosity
 - <https://curiosity.sas.com/app/datafly>
 - <https://curiosity.sas.com/en/home.html>

MORE READING

- Edward Tufte
<https://www.edwardtufte.com/tufte/>
- Stephen Few - *Show Me the Numbers: Designing Tables and Graphs to Enlighten*
- David McCandless – *Information is Beautiful*



I have no
special talents.
I am only
passionately
curious

Albert Einstein

TRANSFORMING DATA INTO INFORMATION USING VISUALISATION



[https://github.com/sarahemmarshall/
dataviz/tree/main/stats-day-2022](https://github.com/sarahemmarshall/dataviz/tree/main/stats-day-2022)

<https://bit.ly/3ET4guQ>

