University of Nottingham
UK | CHINA | MALAYSIA

Sharing your science:
Three easy pieces

Denis Schluppeck

ABDSA Limuru
Kenya, 2024

1. **the big picture:** science communication

2. **the nitty-gritty:** some practical tips (formats, basic principles, version control)

3. **an example:** bringing a jupyter notebook to life

University of Nottingham
UK | CHINA | MALAYSIA

Engaging the media / public + presenting your work – a scientist's view

Denis Schluppeck

ABDSA Limuru
Kenya, 2024

Some thoughts about our wider "responsibilities"...

as scientists

- … increase **public awareness** of (your) science / technology

- try to influence your institutions / local administrators (+ government) to support our work

- build communication competence skills (also of students and colleagues around you)

Challenges

– my view

Challenges

– my view

- get **news organisations** interested in covering research and science (sales? incentives?)

Challenges

– my view

- get **news organisations** interested in covering research and science (sales? incentives?)

- get **readers / consumers** to engage

Challenges

– my view

- get **news organisations** interested in covering research and science (sales? incentives?)

- get **readers / consumers** to engage

- get **scientists and researchers** to understand the importance of media / public engagement

Challenges

– my view

- get **news organisations** interested in covering research and science (sales? incentives?)

- get **readers / consumers** to engage

- get **scientists and researchers** to understand the importance of media / public engagement

"The **BSA Media Fellowships** provide a unique opportunity for practising scientists, clinicians and engineers to spend 2 to 6 weeks working at the heart of a media outlet…"

–British Science Association website

Science for Africa Foundation
AAAS: https://www.aaas.org/programs/mass-media-fellowship
BSA website: https://www.britishscienceassociation.org/

4 weeks at the

FINANCIAL TIMES

Denis Schluppeck

Mentor: Clive Cookson

… my aims for this bit:

- How did I get into this & what did I do?
- What have I learnt?
- How can *you* get scientists and researchers around you to engage with “sci-comm”, print and broadcast journalists, online media, …?

How did I get into this?

Fwd: Would you like to experience life as a science journalist? UoN BSA Media Fellowships - entries close midni...

McGraw Paul Inbox - Exchange 11 March 2016 at 16:05

MP

Fwd: Would you like to experience life as a science journalist? UoN BSA Media Fellowships - entries close midnight 16 March 2016

To: LP-Academic

> Please see opportunity below. If you are interested then let me know.
>
> Best wishes,
>
> Paul

> Dear Paul
>
> This year The University of Nottingham is funding three places for academics to participate in the British Science Association (BSA) Media Fellowship scheme. It is a unique opportunity for practising **scientists, clinicians, engineers and social sciences and arts academics** to spend two to six weeks working at the heart of a media outlet such as the Guardian, the BBC or the Times. Fellows are mentored by professional journalists. They learn how the media operates and reports on science, how to communicate with the media and engage the wider public with science through the media. This year there are three funded places – for Arts, Social Sciences and Science.
>
> This placement would involve support from their Head of School.
>
> I wondered if it would be possible to flag this golden opportunity to all School of Psychology academics?
>
> Kind regards
> Lindsay
>
> **Lindsay Brooke**
> **Media Relations Manager (Science)**
> Press Office
> External Relations
> The University of Nottingham

Fwd: Would you like to experience life as a science journalist? UoN BSA Media Fellowships - entries close midni...

McGraw Paul Inbox - Exchange 11 March 2016 at 16:05

MP

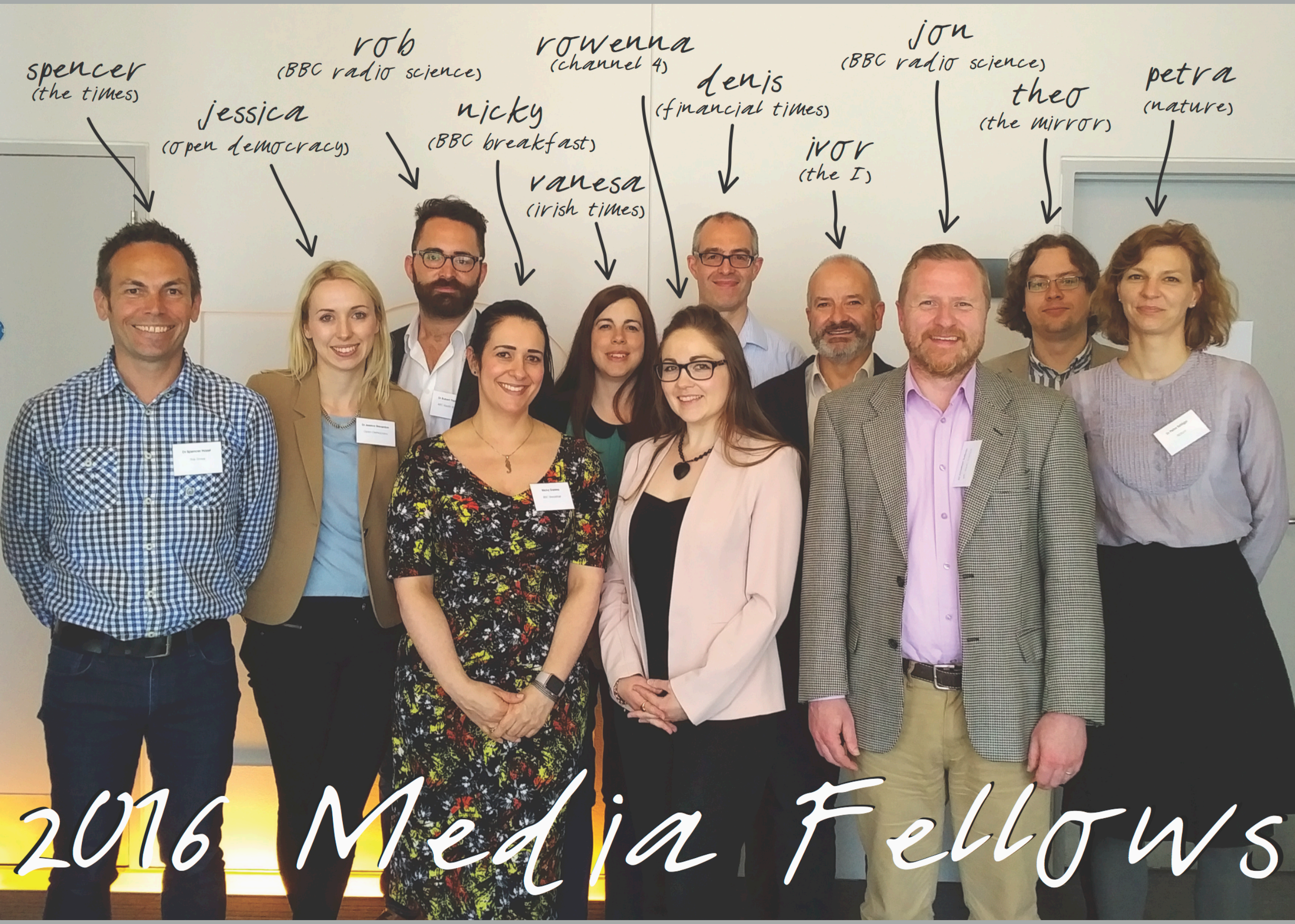
Fwd: Would you like to experience life as a science journalist? UoN BSA Media Fellowships - entries close midnight 16 March 2016

To: LP-Academic

> Please see opportunity below. If you are interested then let me know.
>
> Best wishes,
>
> Paul

> Dear Paul
>
> This year The University of Nottingham is funding three places for academics to participate in the British Science Association (BSA) Media Fellowship scheme. It is a unique opportunity for practising **scientists, clinicians, engineers and social sciences and arts academics** to spend two to six weeks working at the heart of a media outlet such as the Guardian, the BBC or the Times. Fellows are mentored by professional journalists. They learn how the media operates and reports on science, how to communicate with the media and engage the wider public with science through the media. This year there are three funded places – for Arts, Social Sciences and Science.
>
> This placement would involve support from their Head of School.
>
> I wondered if it would be possible to flag this golden opportunity to all School of Psychology academics?
>
> Kind regards
> Lindsay
>
> **Lindsay Brooke**
> **Media Relations Manager (Science)**
> Press Office
> External Relations
> The University of Nottingham

Lindsay Brooke


spencer
(the times)
jessica
(open democracy)
rob
(BBC radio science)
nicky
(BBC breakfast)
vanesa
(irish times)
rowenna
(channel 4)
denis
(financial times)
ivor
(the I)
jon
(BBC radio science)
theo
(the mirror)
petra
(nature)
2016 Media Fellows

Hosts (2016)

The Times

The Guardian

Financial Times

The I

The Daily Mirror

Daily Telegraph

Daily Mail

BBC Radio Science (+online)

Channel 4

BBC Breakfast

Open Democracy

Nature News

Next stop: *FT*

FINANCIAL TIMES

- daily readership: 2.2m
- circulation >234k (in 2014)
- daily, emphasis on business and economic news
- founded 1888, in 2015 bought by Nikkei, Inc. (Tokyo)

source: via wikipedia

my mentor

Clive Cookson

started as journalist
straight out of Oxford (1974)

Times Higher ES (Science)
The Times (Technology)
BBC Radio Science

FT (1987)
Science Editor / FT (1991)

my mentor

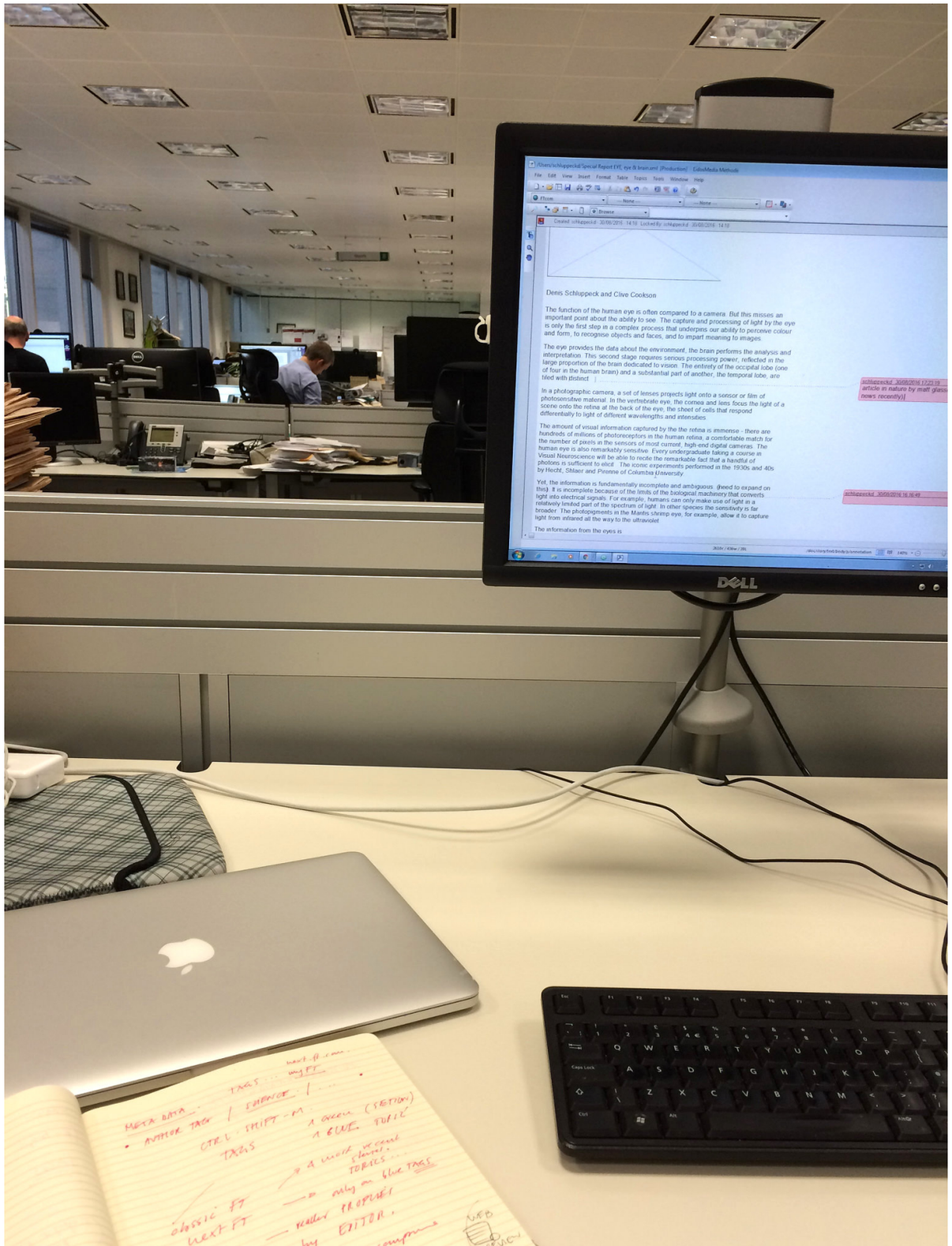
Clive Cookson

started as journalist
straight out of Oxford (1974)

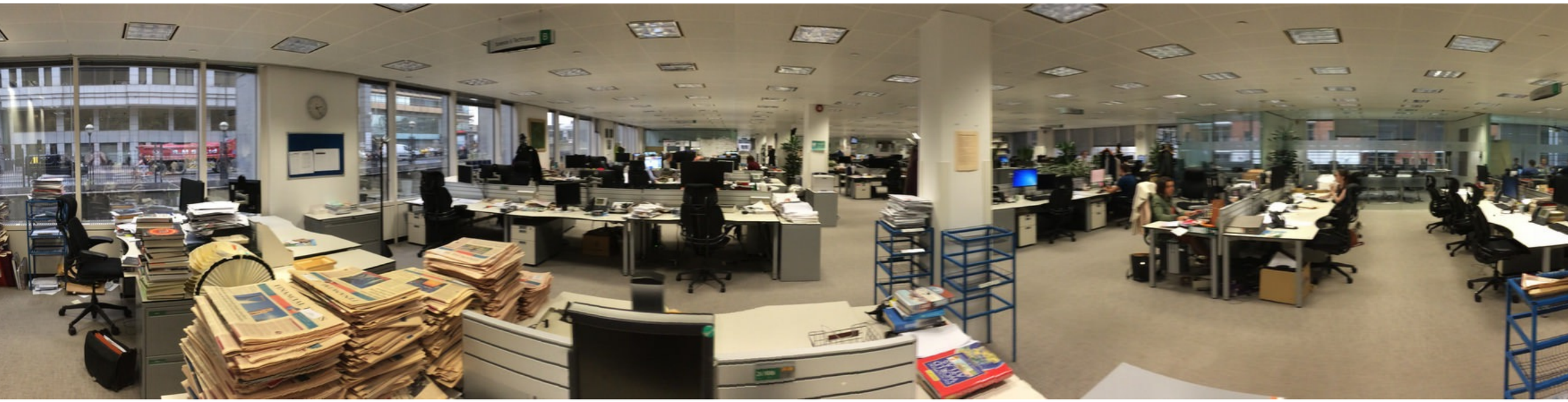
Times Higher ES (Science)
The Times (Technology)
BBC Radio Science

FT (1987)
Science Editor / FT (1991)

at the FT (Aug–Sept)



work !



a typical day

- 730am: wake up and start checking **EurekAlert** website

- 9am. arrive at desk (via Borough Market coffee shop)

- Discuss possible stories with Clive

- 10am Hop on tube to get to **press briefing**
  Today: the Royal Astronomical Society; ESA Gaia mission

- 1245h. Back at FT. Check in with Clive. Start story (lunch *al desko*)

- 1500-1530h, **a coherent 450w piece** is ready. Grab an espresso.

- 1600h. **File story with the newsdesk**. Within 20-30 min, the editor assigned to the story comes back with any questions and an edited version.

- 1700h. The **picture desk gets in touch**. What photo to run with the story.

ted for 10 years. Don't believe

migration. Switzerland, by contrast, has

Over a decade-long period, Switzer-

Santander UK and a former Bus

European single
tinued access is se

A memo on Jap
posted on the mi
website, calls on
negotiate a post
guards almost
rights in the sing

The memo s
lured some Jap
ain on the basi
gateway to Eu
has a moral ob
promises.

"We strong
consider this
in a respons
any harmfu
nesses," sai
Japan's posi
negotiators.

"Japanes
European
may decid
office func
if EU laws
UK after
says.

It adds t
tions migh
ations fro
ments in
their rig
obtained
market.

Almost
ment inter
to the UK
main des
ment sto
of last ye

But Ja
difficult
closely
tionshi
voters
reform

Japa
access
the U
tariff
cedu

Th
nise
betw

M
"co
nati
mir
me

G2
Ha
th
th
m
st

paper 'sting'

z to 'stand le' from ne affairs mittee

MP Keith Vaz is expected airman of the Commons select committee within wspaper report about his

nister, whose role on the given him a high profile immigration and drugs s speaking to his solici- that he was targeted in

ports in the Sunday l paid for the services Vaz said: "It is deeply national newspaper individuals to have

s currently carrying itain's prostitution t issued to the Mail id he would inform rrow of his inten- It was not clear if y or permanent. er solicitor and was first elected East in 1987. He in Tony Blair's r Valerie is MP

Mr Vaz's posi- n and for the Mr Vaz — who tional execu- favour of Jer- getting on to dership elec- cy party has ontest. rted that Mr oppers, the e two male ed attempts

mer culture r Mr Vaz to n the areas esponsible, a sensible

A 3D image of a skull believed to be that of a carpenter on the Mary Rose, left, a facial reconstruction of the carpenter, above, and a painting of Henry VIII's flagship, below

Heads up
Digital focus on crew of Mary Rose

Scientists have produced highly detailed 3D reconstructions of skulls found on the Mary Rose as they try to unravel the mystery of the 460 crew who died when Henry VIII's flagship sank in 1545.

The crew's identities are largely unknown, but researchers from Swansea and Oxford universities hope to glean hints about their age, health and even occupation from the interactive, photorealistic models of each skull, and by sharing the results online.

Capturing museum collections digitally is now commonplace, but a particular focus of this project is to establish whether the quality of the 3D reconstruction is high enough for serious study of skeletal remains.

"This technology, and the appetite of museums and researchers to open their collections to larger global communities, can have huge implications for [the] speed that science is done," said Dr Richard Johnston of Swansea University.

For archaeology, digital imagery on a much larger scale has already brought a revolution. Google Earth has been embraced by scientists as a useful tool for archaeological digs and detailed imagery has been used to reconstruct a replica of an arch from Palmyra in Trafalgar Square.

In other areas of science, from astrophysics to zoology, the push towards sharing data online has made it easier for scientists to collaborate across the globe. Successful projects such as Zooniverse have also allowed researchers to involve the public in their research.

The Mary Rose, which sunk off Portsmouth, continues to capture the imagination. The large number of finds inside the hull have provided an invaluable insight into the Tudor period. Many thousands of artefacts and the carefully conserved hull of the ship are now on display in Portsmouth's Mary Rose Museum.

One of the skulls featured in the project is thought to be that of a carpenter. It was found immediately below the Master Carpenter's cabin near a number of woodworking tools. Ship's carpenters were particularly important during battles and were stationed below decks for repairing damage below the waterline.

The skull reconstruction will be available at www.virtualtudors.org.

Denis Schluppeck

4 weeks at the FT

10 articles in

FINANCIAL TIMES

MENU SEARCH myFT

HOME WORLD US COMPANIES MARKETS OPINION WORK & CAREERS LIFE & ARTS Portfolio My Account

Medical science ✓ Added

Digital research aims to unveil Tudor mystery

Researchers assess skulls of Mary Rose crew from sinking of 1545

Read latest:
If you want to learn, sleep on it
1 SECONDS AGO

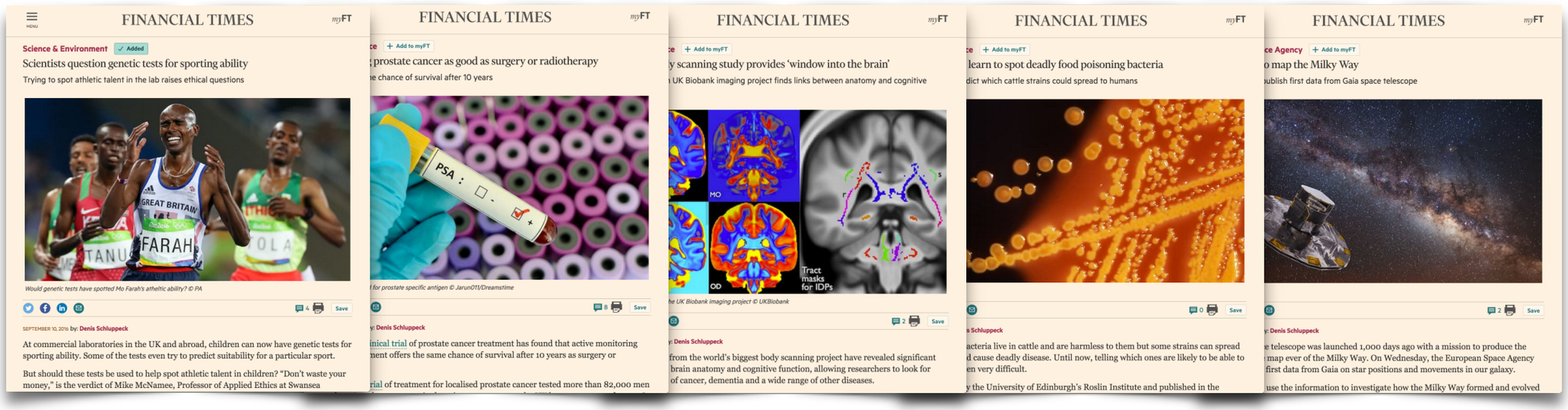
A 3D image of the skull of a carpenter found on the Mary Rose

3 Save

SEPTEMBER 5, 2016 by: Denis Schluppeck

Scientists have produced highly detailed 3D reconstructions of skulls found on the Mary Rose as they try to unravel the mystery of the 460 crew who died when Henry VIII's flagship sank in 1545.

+



FINANCIAL TIMES myFT

Science & Environment ✓ Added

Scientists question genetic tests for sporting ability

Trying to spot athletic talent in the lab raises ethical questions

Would genetic tests have spotted Mo Farah's athletic ability? © PA

SEPTEMBER 10, 2016 by: Denis Schluppeck

At commercial laboratories in the UK and abroad, children can now have genetic tests for sporting ability. Some of the tests even try to predict suitability for a particular sport.

But should these tests be used to help spot athletic talent in children? "Don't waste your money," is the verdict of Mike McNamee, Professor of Applied Ethics at Swansea

FINANCIAL TIMES myFT

+ Add to myFT

prostate cancer as good as surgery or radiotherapy

e chance of survival after 10 years

for prostate specific antigen © JarunOTI/Dreamstime

8 Save

y: Denis Schluppeck

inical trial of prostate cancer treatment has found that active monitoring ment offers the same chance of survival after 10 years as surgery or

rial of treatment for localised prostate cancer tested more than 82,000 men

FINANCIAL TIMES myFT

+ Add to myFT

y scanning study provides 'window into the brain'

UK Biobank imaging project finds links between anatomy and cognitive

he UK Biobank imaging project © UKBiobank

2 Save

y: Denis Schluppeck

from the world's biggest body scanning project have revealed significant brain anatomy and cognitive function, allowing researchers to look for of cancer, dementia and a wide range of other diseases.

FINANCIAL TIMES myFT

+ Add to myFT

learn to spot deadly food poisoning bacteria

dict which cattle strains could spread to humans

0 Save

s Schluppeck

acteria live in cattle and are harmless to them but some strains can spread d cause deadly disease. Until now, telling which ones are likely to be able to en very difficult.

y the University of Edinburgh's Roslin Institute and published in the

FINANCIAL TIMES myFT

ce Agency + Add to myFT

o map the Milky Way

ublish first data from Gaia space telescope

2 Save

y: Denis Schluppeck

e telescope was launched 1,000 days ago with a mission to produce the map ever of the Milky Way. On Wednesday, the European Space Agency first data from Gaia on star positions and movements in our galaxy.

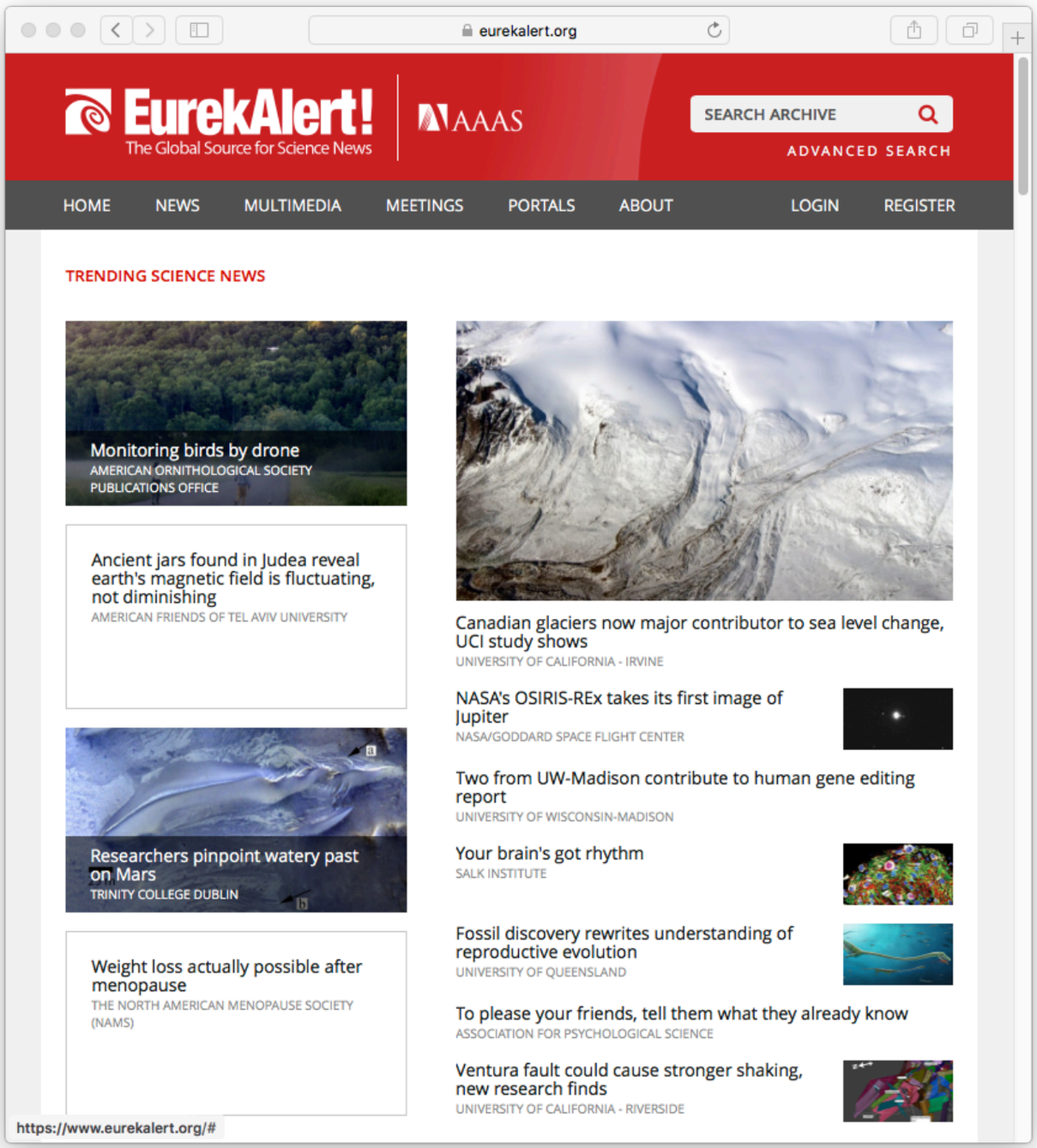
use the information to investigate how the Milky Way formed and evolved

20160905-maryRose/
20160906-bsf-oesophagealCancer/
20160906-supercomputingDrugs/
20160907-cloudyWithAChanceOfPain/
20160908-sleepLab/
20160909-geneticTestingInSports/
20160914-esa-mission/
20160914-prostateCancerTrial/
20160915-eyeVisionBrain/
20160919-biobank-brainImaging/
20160920-machineLearningEColi/
20160921-evolutionOfScience/
20160922-cornWorms/
20160923-hangoverFreeAlcohol/

2

What have I learnt?

For print journalists, *press release* is king!



snappy title

a summary that catches the eye

embargo!

nice imagery / photos (not graphs, tables, …)

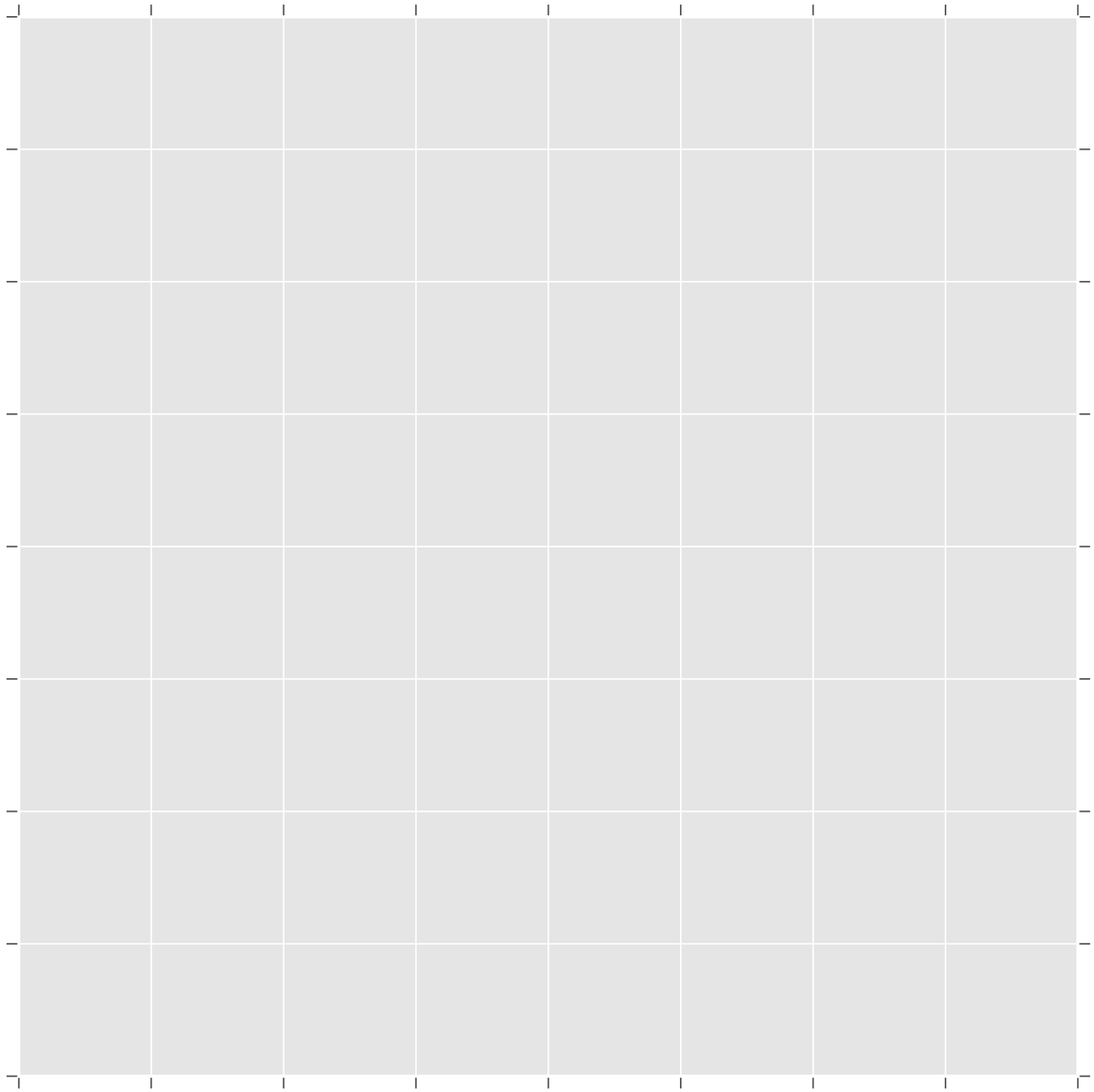
Key sites

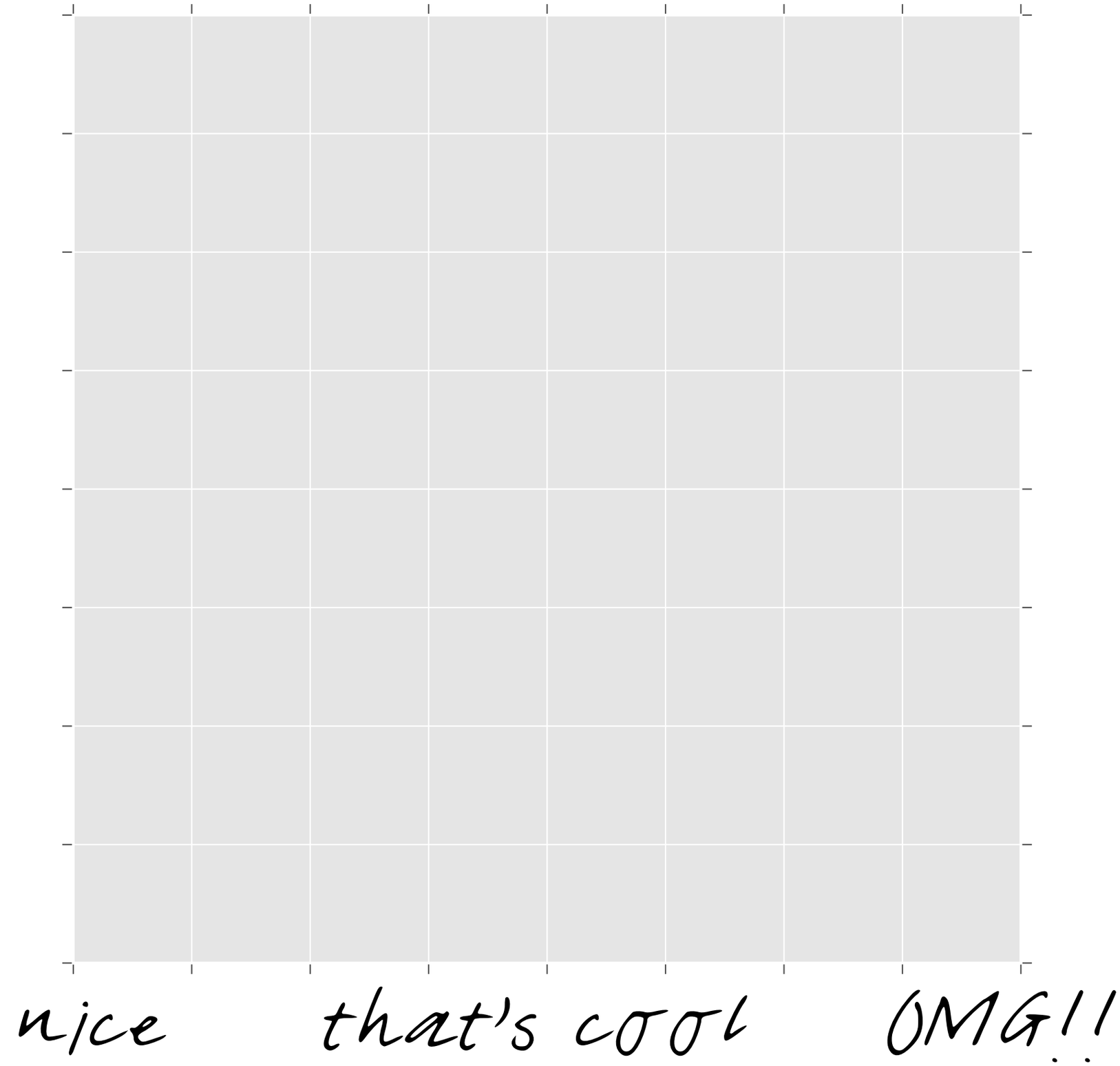
https://www.eurekalert.org/

https://www.alphagalileo.org/

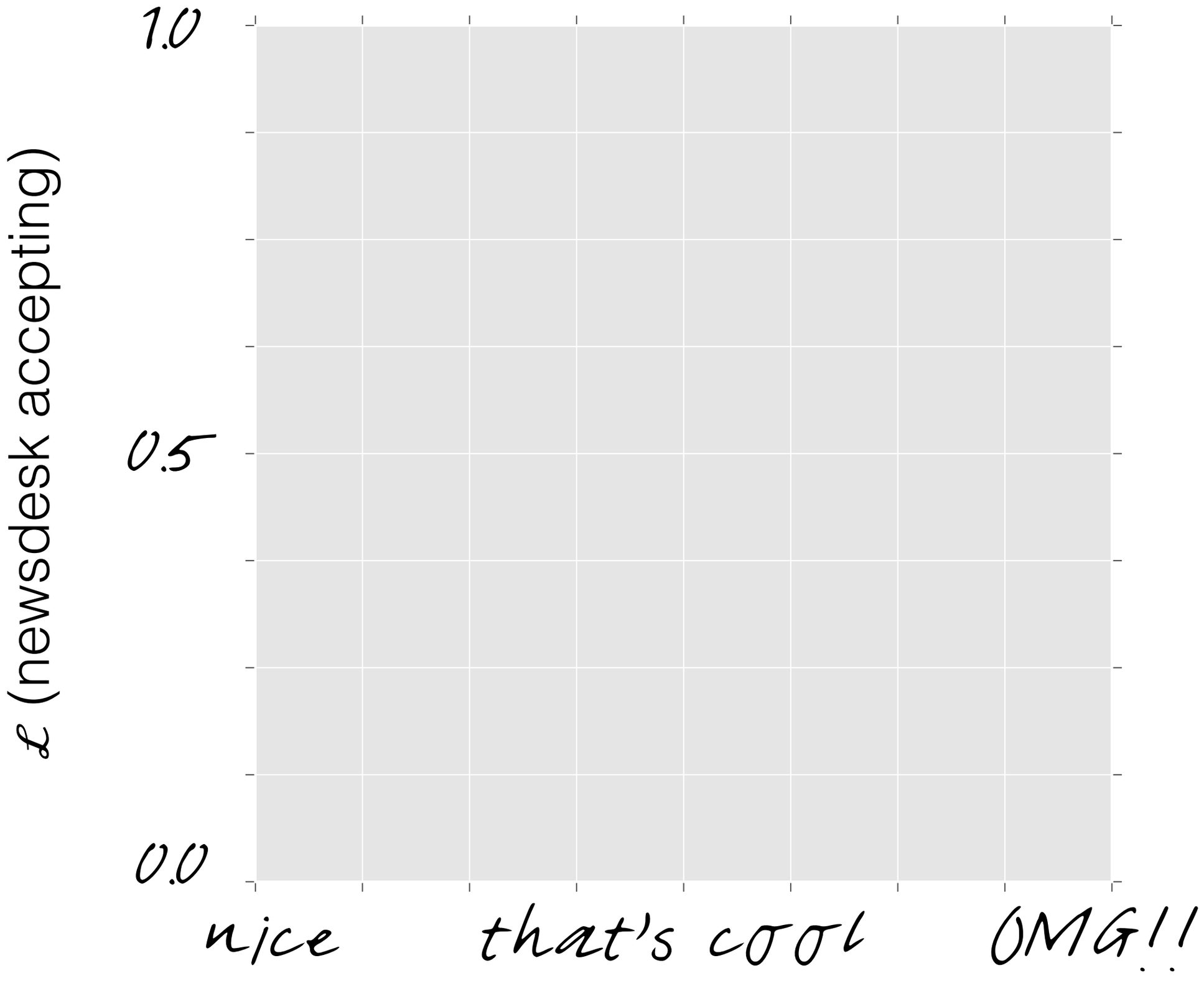
directly or via journals

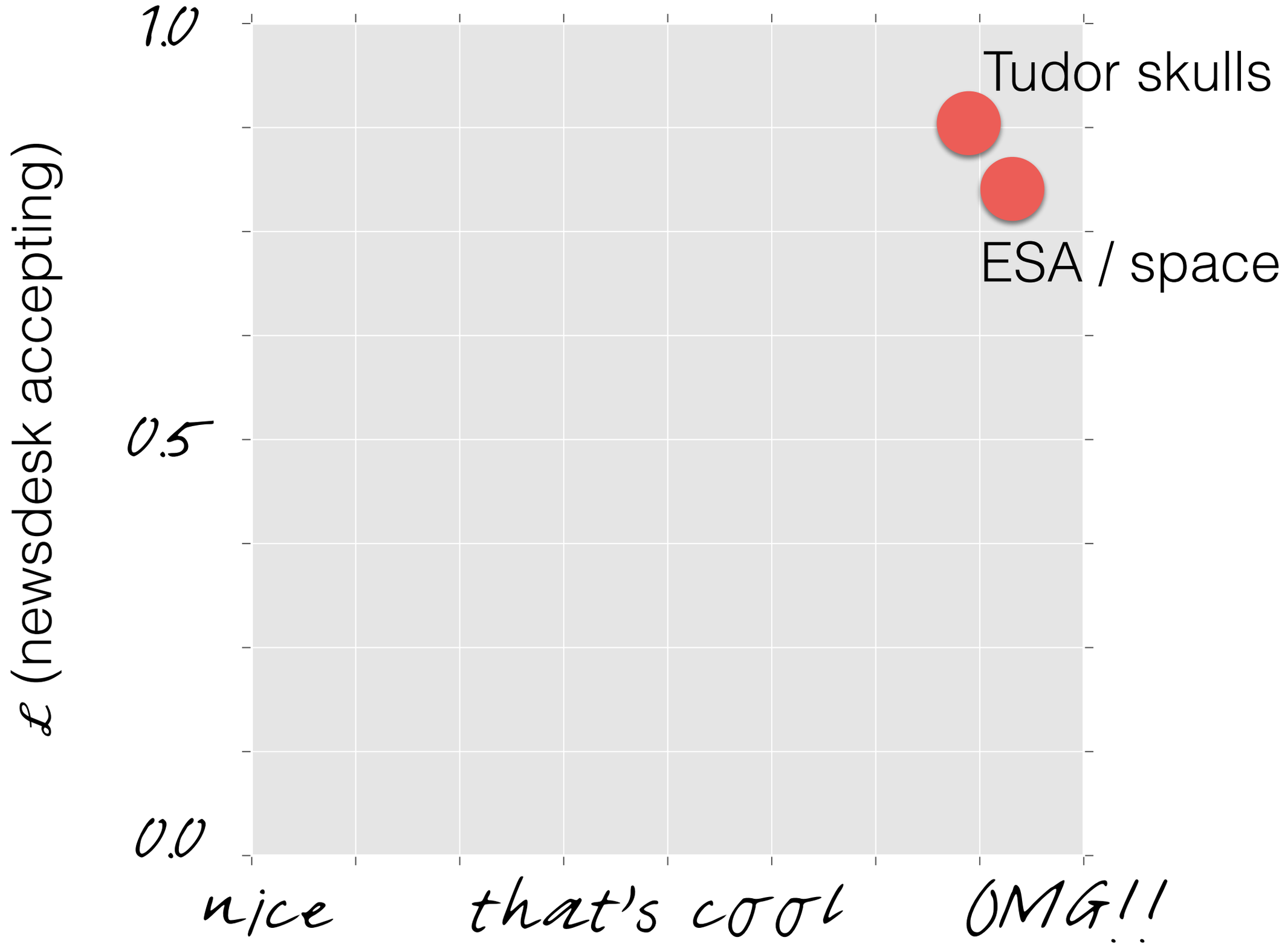
thjngs I learnt about myself

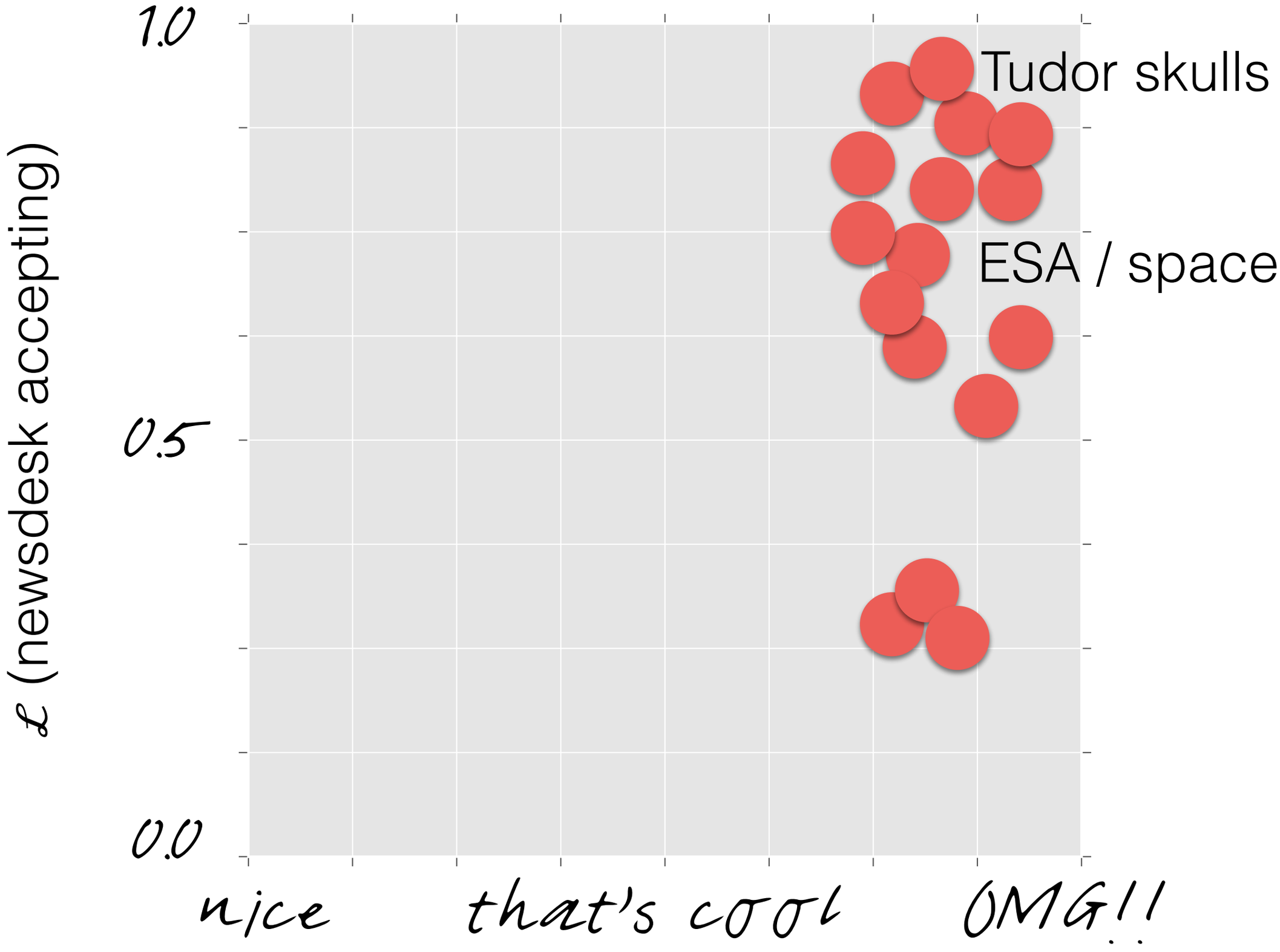


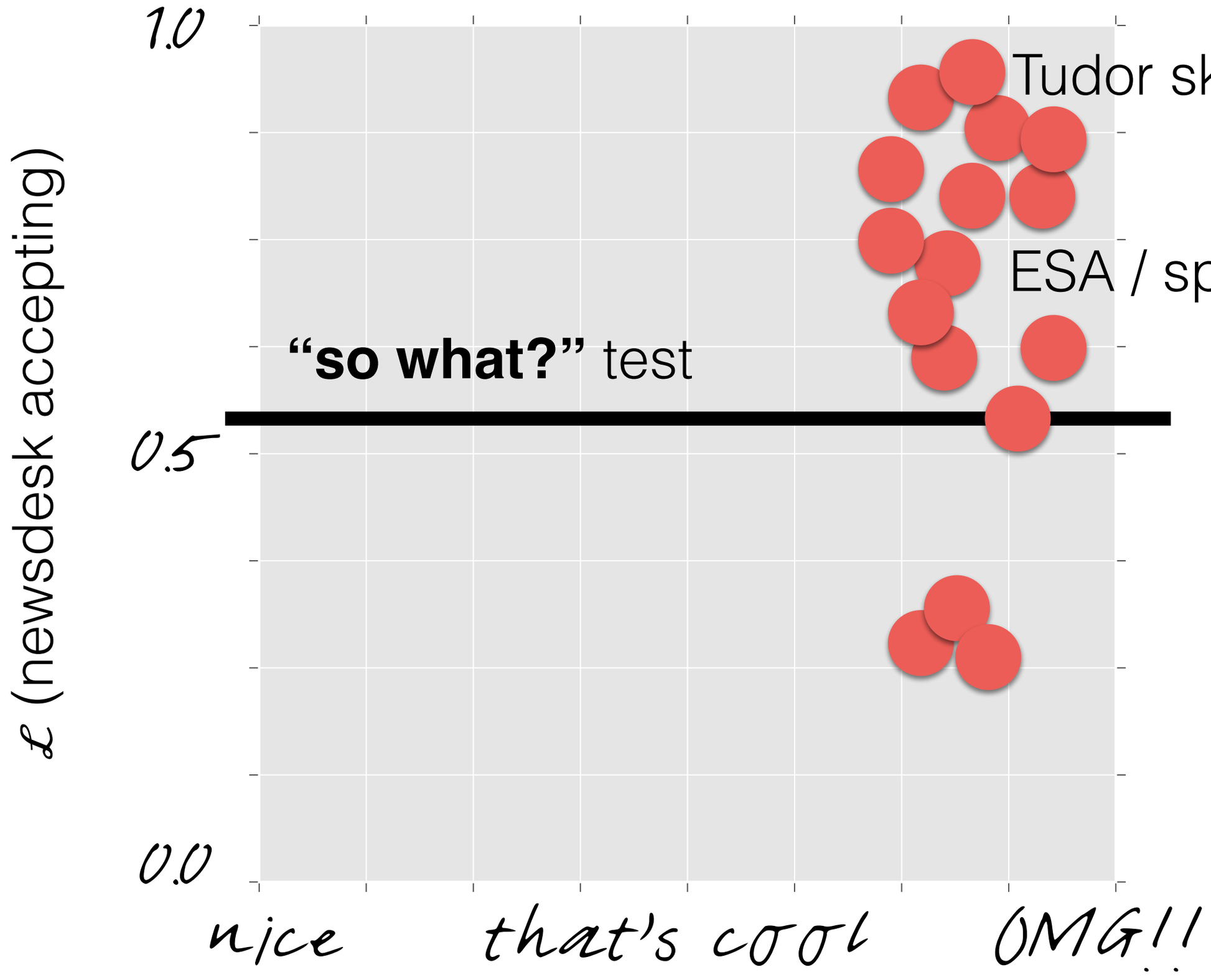


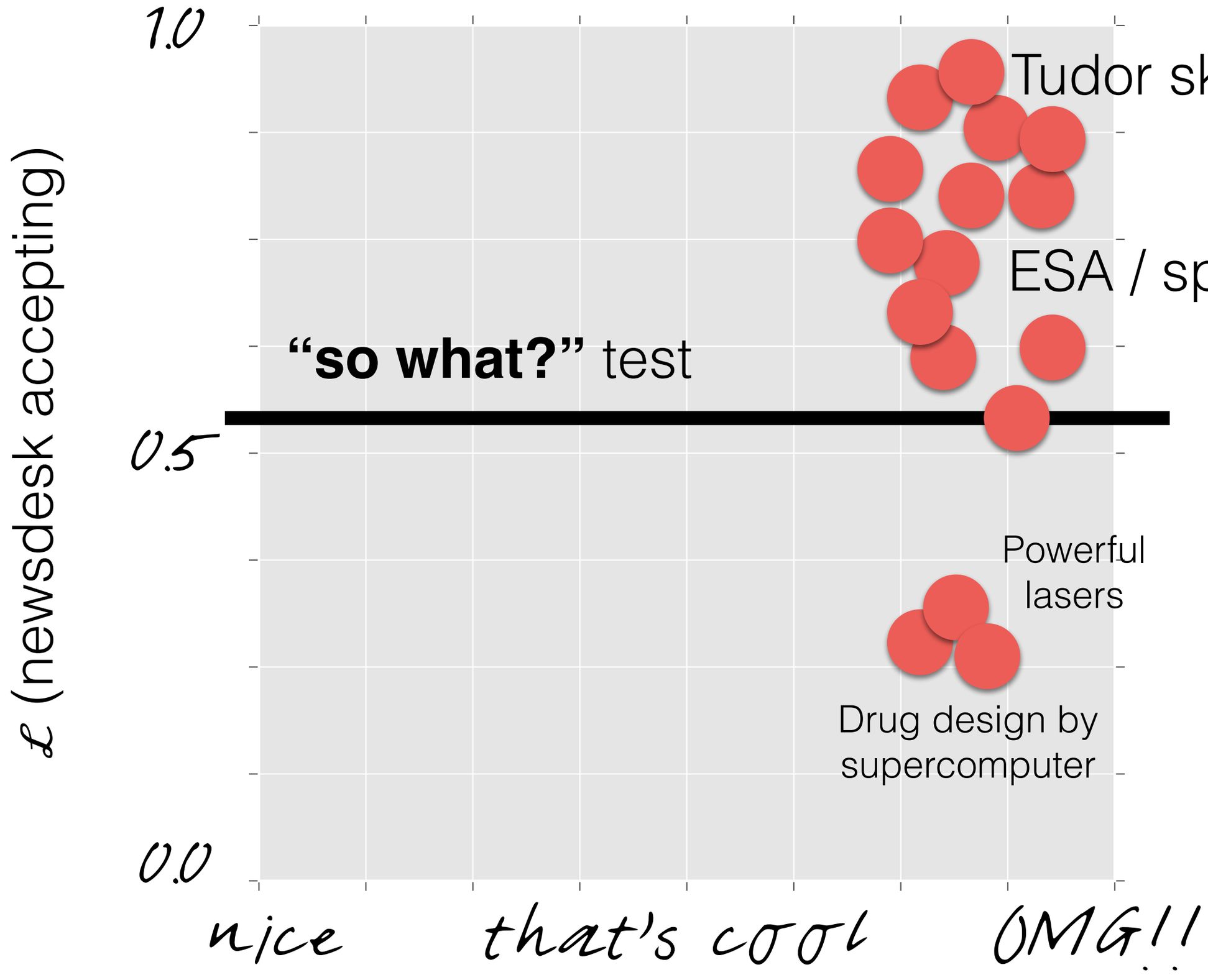
How fascinating do I find this story?


1.0
0.5
0.0
ℓ (newsdesk accepting)
nice
that's cool
OMG!!
How fascinating do I find this story?


1.0
0.5
0.0
£ (newsdesk accepting)
Tudor skulls
ESA / space
nice
that's cool
OMG!!
How fascinating do I find this story?


1.0
0.5
0.0
£ (newsdesk accepting)
Tudor skulls
ESA / space
nice
that's cool
OMG!!
How fascinating do I find this story?


1.0
0.5
0.0
£ (newsdesk accepting)
Tudor skulls
ESA / space
“so what?” test
nice
that's cool
OMG!!
How fascinating do I find this story?


1.0
0.5
0.0
£ (newsdesk accepting)
"so what?" test
Tudor skulls
ESA / space
Powerful lasers
Drug design by supercomputer
nice
that's cool
OMG!!
How fascinating do I find this story?

- **Clarity of writing** and **relevance to the reader** are the most important things.

- **Clarity of writing** and **relevance to the reader** are the most important things.
- If you want to improve, **look for real feedback,** not only what you want to hear

- **Clarity of writing** and **relevance to the reader** are the most important things.

- If you want to improve, **look for real feedback,** not only what you want to hear

- **journalists are busy** and the pace is pretty crazy

“If a good scientific paper is a 5-course menu, a well-written news piece is the perfect serving of sushi.”

–Denis Schluppeck

3

What can *you* (and researchers) do to increase engagment?

Build relationships

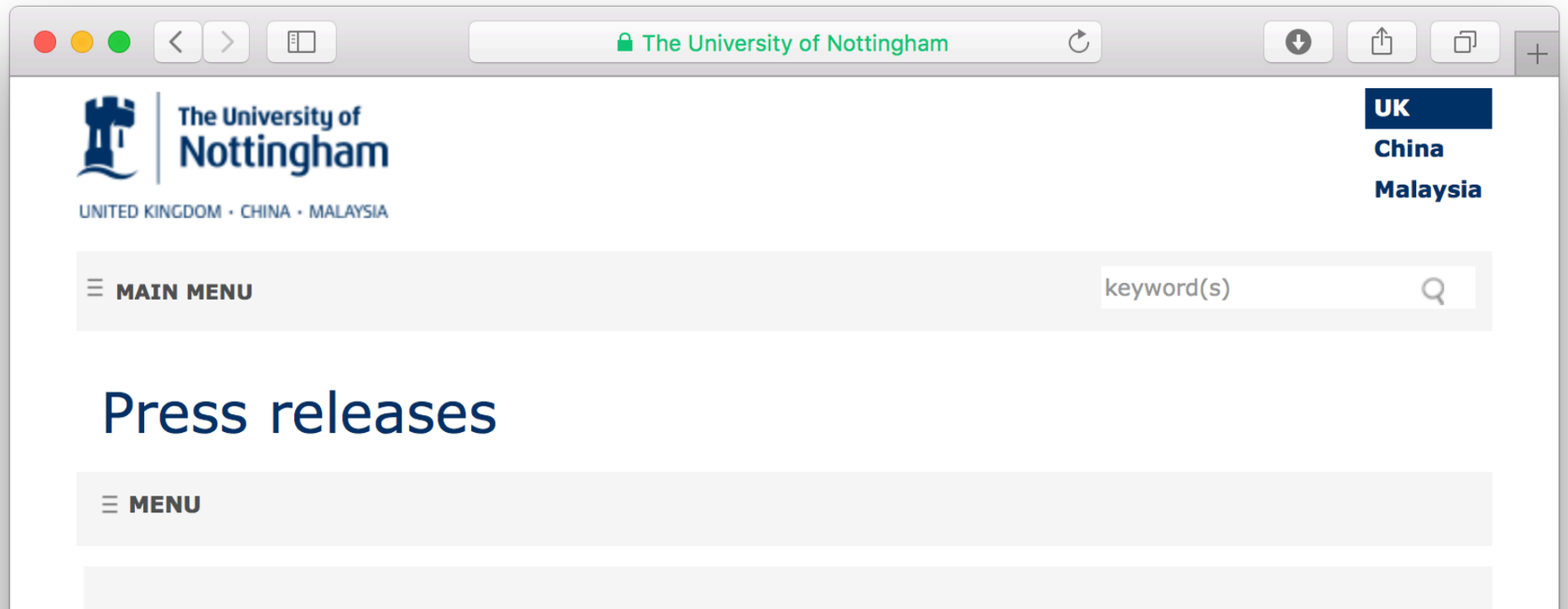
- try to meet science journalists - in person.
  - different outlets (broadsheets – to magazines)
  - local radio (for some stories)
  - … may lead to national coverage
  - … may lead to the next story getting picked up

at ~~Nottingham~~...
your institution

U o Nairobi: https://www.uonbi.ac.ke/news

U o Rwanda: https://ur.ac.rw/spip.php?page=news-and-events

… look up your own media / press office



Timescales

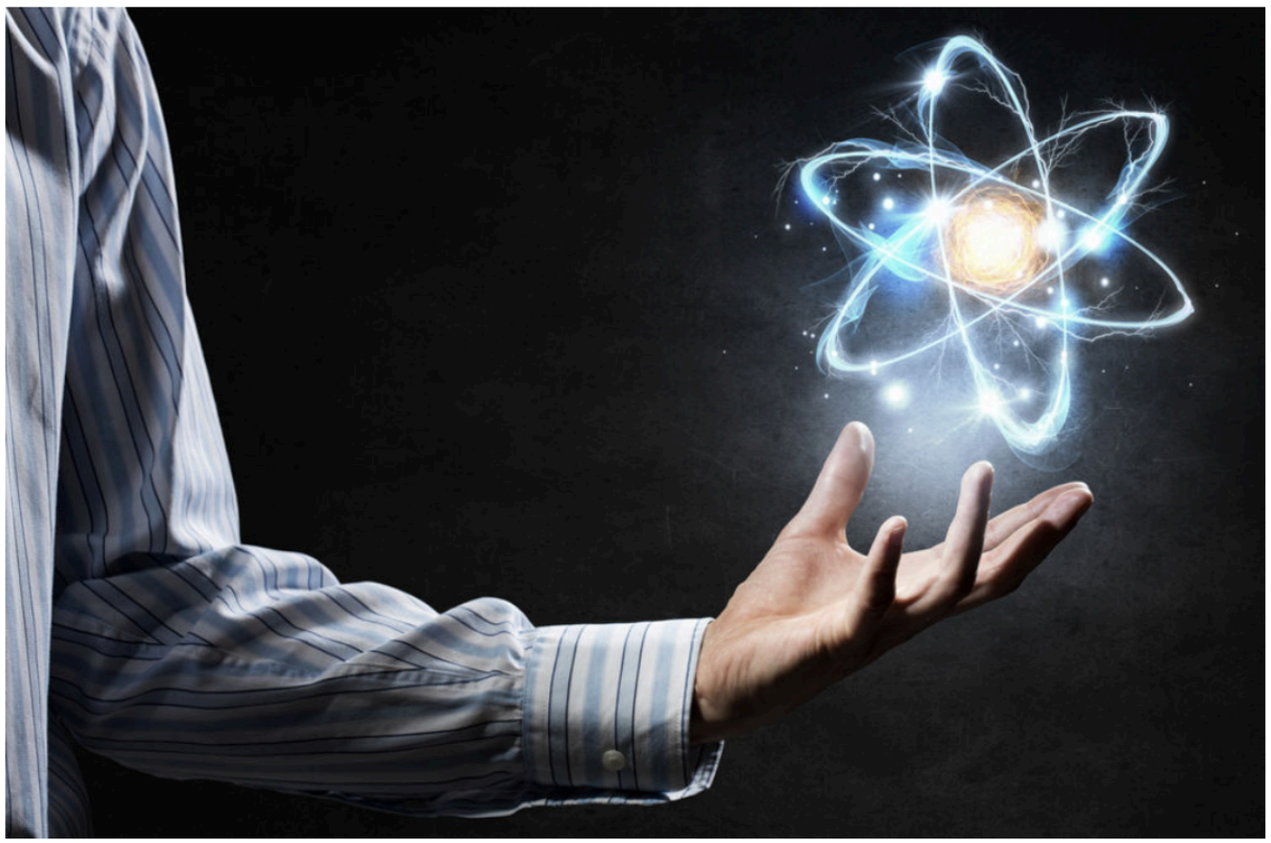
- we / scientists need to learn to plan ahead: talk to media people about work that might be relevant

- media deadlines are much tighter than academic ones – we need to put ourselves in their shoes + be ready

- may require culture shift +/- help from experienced professionals

Different work – different channels

- if a story doesn't end up getting in, don’t take it personally [many reasons why things get “bumped”]
- not everything is front page news
- some work might be more suitable for a blog
- things get re-tweeted, re-discovered, shared on Facebook, go viral…

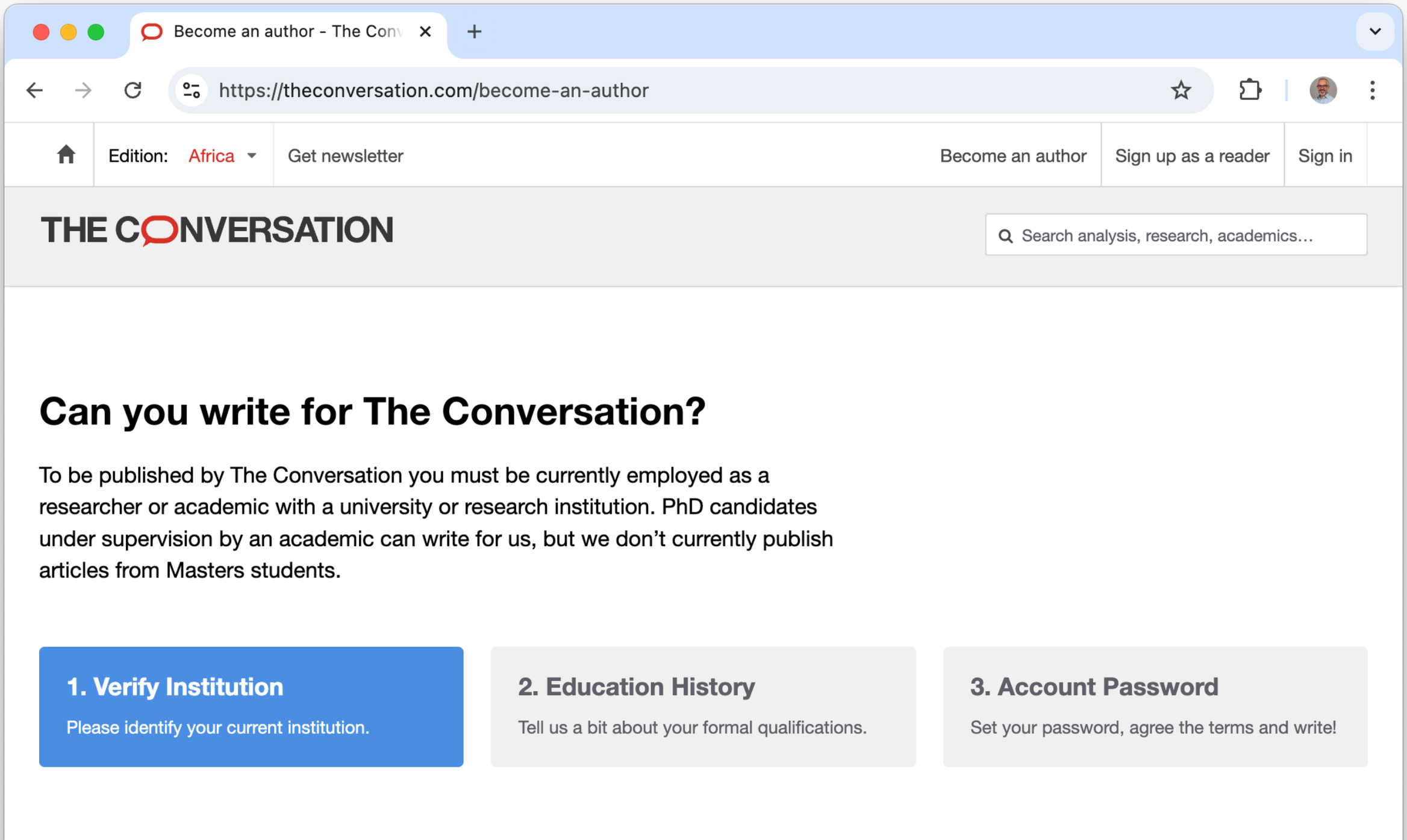


Why can't we see the spaces? Shutterstock

If atoms are mostly empty space, why do objects look and feel solid?

Peter Barker, University of Huddersfield

become an author!



1. ~~**the big picture:** science communication~~

2. **the nitty-gritty:** some practical tips (formats, basic principles, version control)

3. **an example:** bringing a jupyter notebook to life

- **tidy data** principles
- **version control** with git/github
- **markdown** is queen / king

Tidy data

Rethinking your plotting (+ data sharing) habits

Aim

Make us think about how to handle data to be

- reproducible (+ easy to understand later)
- more flexible for exploration
  - in other kinds of plots, visualisations
  - with other tools / programming languages
  - by other people (colleagues, people downloading when you publish)

inspiration for this

Journal of Statistical Software
MMMMMM YYYY, Volume VV, Issue II. *http://www.jstatsoft.org/*

Tidy Data

Hadley Wickham
RStudio

Abstract

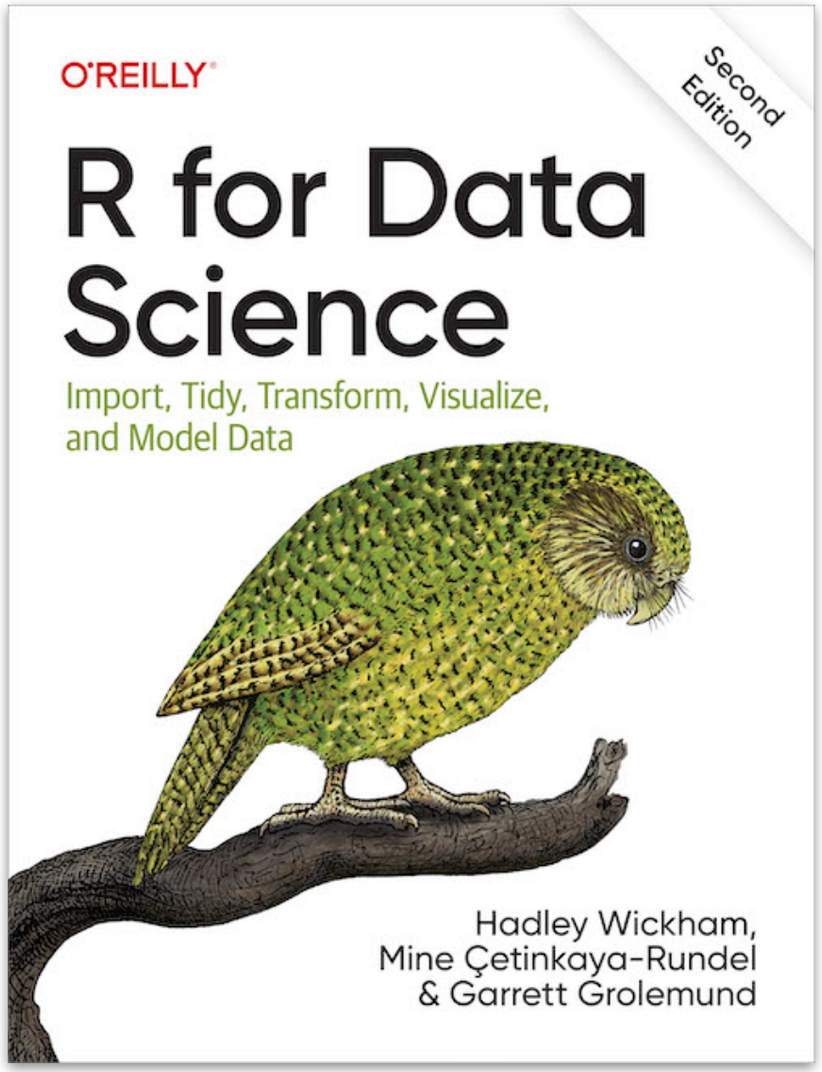
A huge amount of effort is spent cleaning data to get it ready for analysis, but there has been little research on how to make data cleaning as easy and effective as possible. This paper tackles a small, but important, component of data cleaning: data tidying. Tidy datasets are easy to manipulate, model and visualise, and have a specific structure: each variable is a column, each observation is a row, and each type of observational unit is a table. This framework makes it easy to tidy messy datasets because only a small set of tools are needed to deal with a wide range of un-tidy datasets. This structure also makes it easier to develop tidy tools for data analysis, tools that both input and output tidy datasets. The advantages of a consistent data structure and matching tools are demonstrated with a case study free from mundane data manipulation chores.

Keywords: data cleaning, data tidying, relational databases, R.

1. Introduction

It is often said that 80% of data analysis is spent on the process of cleaning and preparing the data (Dasu and Johnson 2003). Data preparation is not just a first step, but must be repeated many over the course of analysis as new problems come to light or new data is collected. Despite the amount of time it takes, there has been surprisingly little research on how to clean data well. Part of the challenge is the breadth of activities it encompasses: from outlier checking, to date parsing, to missing value imputation. To get a handle on the problem, this paper focusses on a small, but important, aspect of data cleaning that I call data **tidying**: structuring datasets to facilitate analysis.

The principles of tidy data provide a standard way to organise data values within a dataset. A standard makes initial data cleaning easier because you don't need to start from scratch and reinvent the wheel every time. The tidy data standard has been designed to facilitate initial exploration and analysis of the data, and to simplify the development of data analysis tools that work well together. Current tools often require translation. You have to spend time



links on **references** slide

inspiration for this

Journal of Statistical Software
MMMMMM YYYY, Volume VV, Issue II. *http://www.jstatsoft.org/*

Tidy Data

Hadley Wickham
RStudio

Abstract

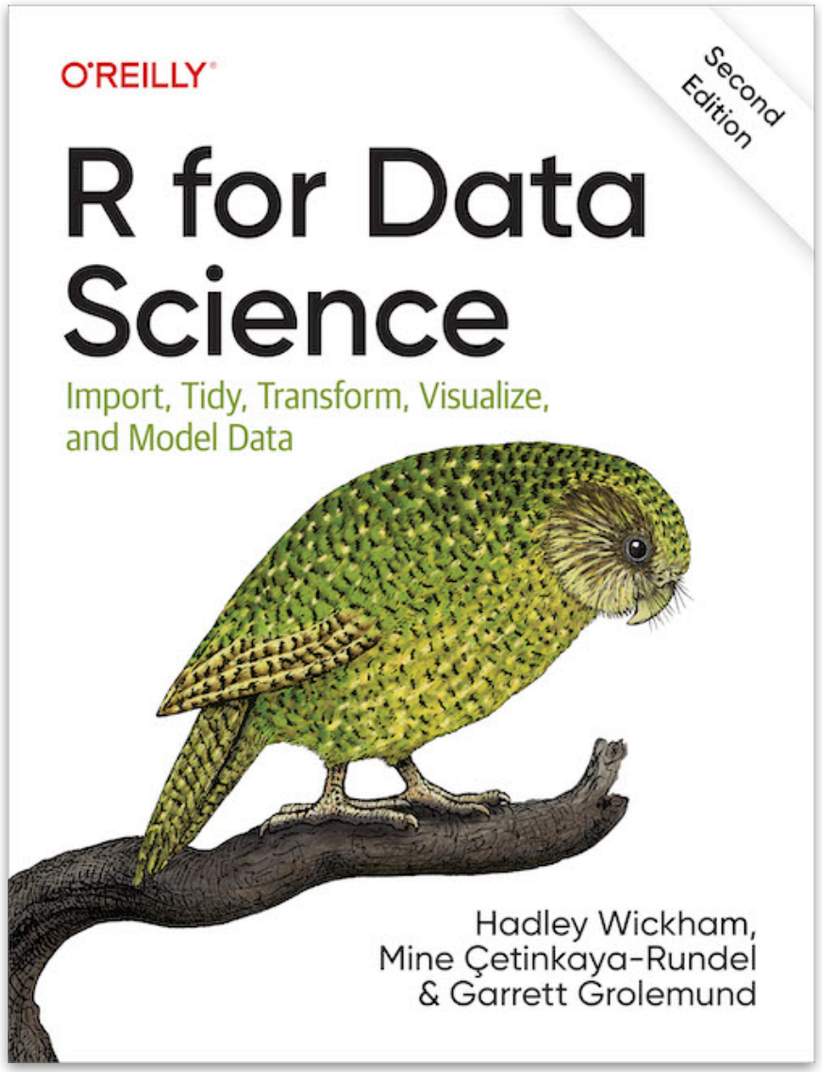
A huge amount of effort is spent cleaning data to get it ready for analysis, but there has been little research on how to make data cleaning as easy and effective as possible. This paper tackles a small, but important, component of data cleaning: data tidying. Tidy datasets are easy to manipulate, model and visualise, and have a specific structure: each variable is a column, each observation is a row, and each type of observational unit is a table. This framework makes it easy to tidy messy datasets because only a small set of tools are needed to deal with a wide range of un-tidy datasets. This structure also makes it easier to develop tidy tools for data analysis, tools that both input and output tidy datasets. The advantages of a consistent data structure and matching tools are demonstrated with a case study free from mundane data manipulation chores.

Keywords: data cleaning, data tidying, relational databases, R.

1. Introduction

It is often said that 80% of data analysis is spent on the process of cleaning and preparing the data (Dasu and Johnson 2003). Data preparation is not just a first step, but must be repeated many over the course of analysis as new problems come to light or new data is collected. Despite the amount of time it takes, there has been surprisingly little research on how to clean data well. Part of the challenge is the breadth of activities it encompasses: from outlier checking, to date parsing, to missing value imputation. To get a handle on the problem, this paper focusses on a small, but important, aspect of data cleaning that I call data **tidying**: structuring datasets to facilitate analysis.

The principles of tidy data provide a standard way to organise data values within a dataset. A standard makes initial data cleaning easier because you don't need to start from scratch and reinvent the wheel every time. The tidy data standard has been designed to facilitate initial exploration and analysis of the data, and to simplify the development of data analysis tools that work well together. Current tools often require translation. You have to spend time

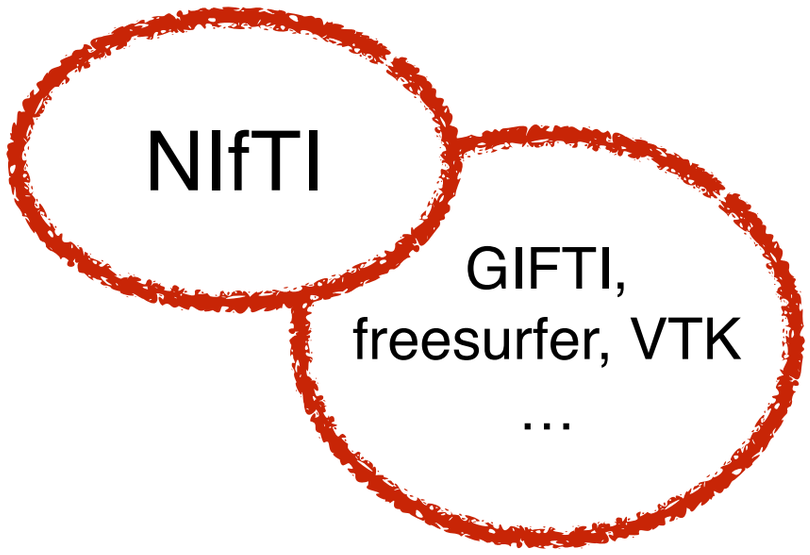


links on **references** slide

Kinds of data?

3d, 4d arrays (+ higher)

imaging data
imaging data + some metadata
+ TXT (e.g. bvecs)



Kinds of data?

3d, 4d arrays (+ higher)

imaging data
imaging data + some metadata
+ TXT (e.g. bvecs)

other stuff (for summary / figures, etc)

tables
key:value
hierarchical

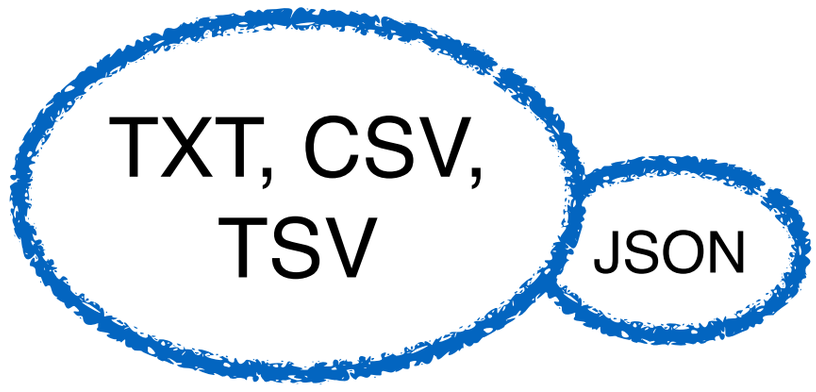
Kinds of data?

3d, 4d arrays (+ higher)

imaging data
imaging data + some metadata
+ TXT (e.g. bvecs)

other stuff (for summary / figures, etc)

tables
key:value
hierarchical



idea: for images

www.nature.com/scientificdata

SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES

» Data publication and archiving

» Research data

Received: 18 December 2015

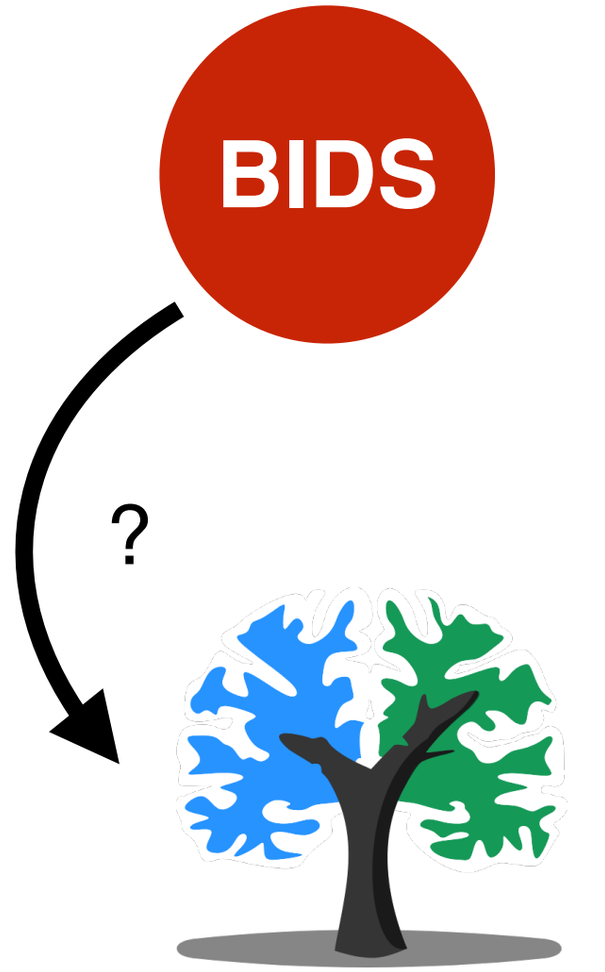
Accepted: 19 May 2016

Published: 21 June 2016

The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments

Krzysztof J. Gorgolewski[1], Tibor Auer[2], Vince D. Calhoun[3,4], R. Cameron Craddock[5,6], Samir Das[7], Eugene P. Duff[8], Guillaume Flandin[9], Satrajit S. Ghosh[10,11], Tristan Glatard[7,12], Yaroslav O. Halchenko[13], Daniel A. Handwerker[14], Michael Hanke[15,16], David Keator[17], Xiangrui Li[18], Zachary Michael[19], Camille Maumet[20], B. Nolan Nichols[21,22], Thomas E. Nichols[20,23], John Pellman[6], Jean-Baptiste Poline[24], Ariel Rokem[25], Gunnar Schaefer[1,26], Vanessa Sochat[27], William Triplett[1], Jessica A. Turner[3,28], Gaël Varoquaux[29] & Russell A. Poldrack[1]

The development of magnetic resonance imaging (MRI) techniques has defined modern neuroimaging. Since its inception, tens of thousands of studies using techniques such as functional MRI and diffusion weighted imaging have allowed for the non-invasive study of the brain. Despite the fact that MRI is routinely used to obtain data for neuroscience research, there has been no widely adopted standard for organizing and describing the data collected in an imaging experiment. This renders sharing and reusing data (within or between labs) difficult if not impossible and unnecessarily complicates the application of automatic pipelines and quality assurance protocols. To solve this problem, we have developed the Brain Imaging Data Structure (BIDS), a standard for organizing and describing MRI datasets. The BIDS standard uses file formats compatible with existing software, unifies the majority of practices already common in the field, and captures the metadata necessary for most common data processing operations.



idea: for tabular data

Journal of Statistical Software

MMMMMM YY Issue II. http://www.jstatsoft.org/

Tidy Data

dley Wickl

Abstract

A huge amount of effort is spent cleaning data to get it ready for analysis, but there has been little research on how to make data cleaning as easy and effective as possible. This paper tackles a small, but important, component of data cleaning: data tidying. Tidy datasets are easy to manipulate, model and visualise, and have a specific structure:

“Happy families are all alike; every unhappy family is unhappy in its own way.” — Leo Tolstoy

“Happy families are all alike; every unhappy family is unhappy in its own way.” — Leo Tolstoy

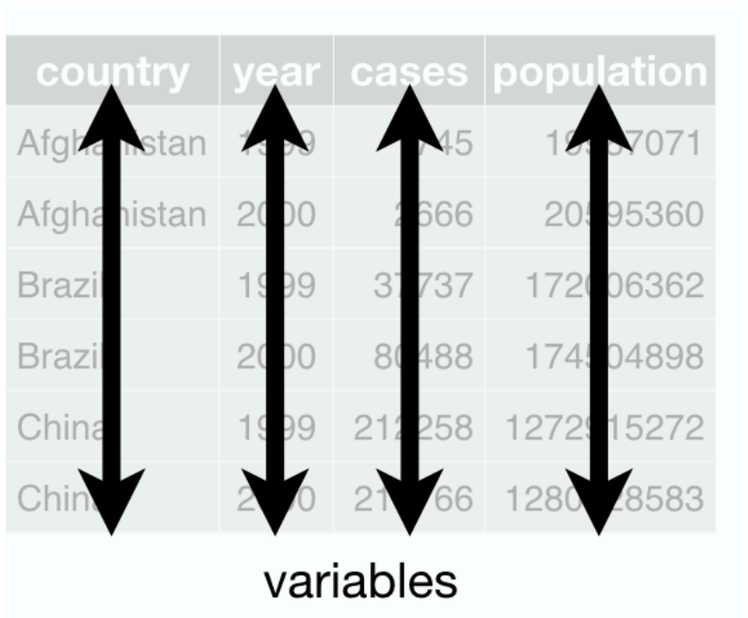
“Tidy datasets are all alike, but every messy dataset is messy in its own way.” — Hadley Wickham

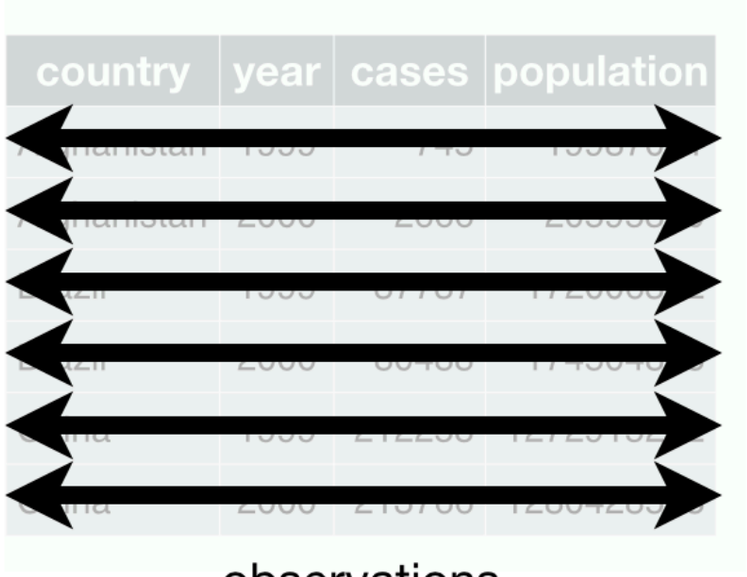
Principles

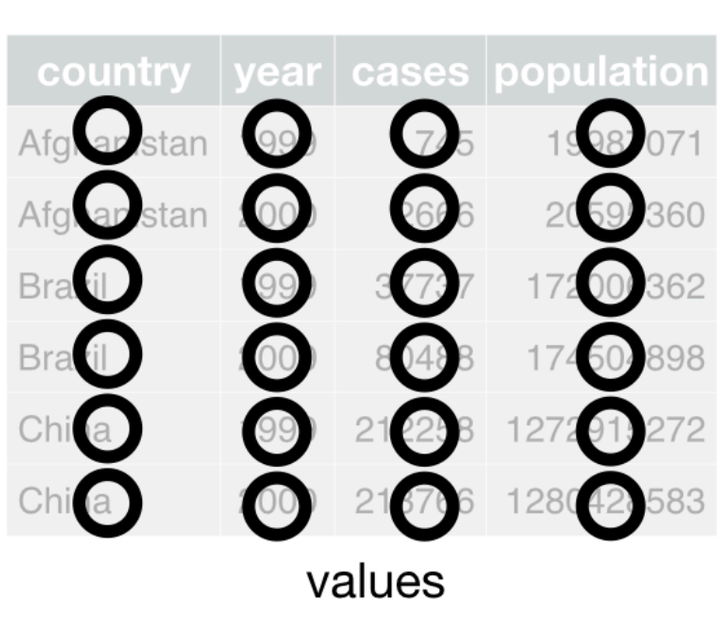
1. Each variable must have its own column.
2. Each observation must have its own row.
3. Each value must have its own cell.

Figures / ideas: https://r4ds.had.co.nz/tidy-data.html

Principles



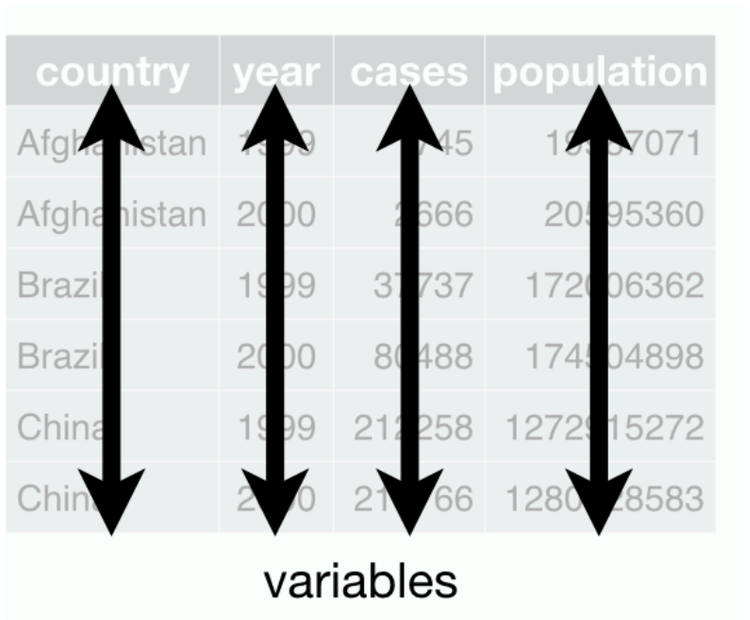


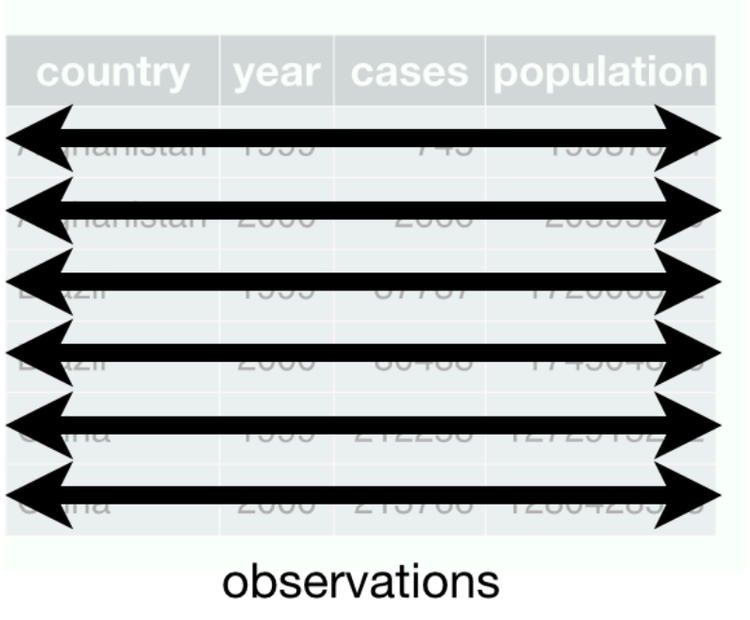


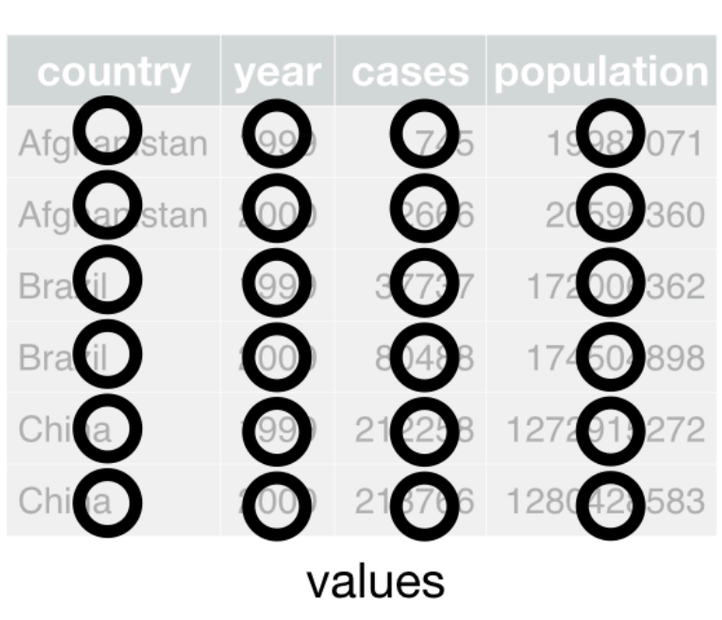
+ even if you have megabytes of data, consider

Figures / ideas: https://r4ds.had.co.nz/tidy-data.html

Principles







+ even if you have megabytes of data, consider

Figures / ideas: https://r4ds.had.co.nz/tidy-data.html

If you stick to this…

… much goodness will result!

If you stick to this…

… much goodness will result!

- predictable (in *your* head + in software)
- many tools handle these kind of data
- secondary analysis + plotting + visualising becomes a breeze

Visualisation tools

- R / ggplot2 (https://ggplot2.tidyverse.org/)
- Matlab / gramm (https://github.com/piermorel/gramm)
- python / pandas (and friends)
  - plotly
  - seaborn
  - plotnine (https://plotnine.readthedocs.io/en/stable/)

“verbs” for dealing with tables

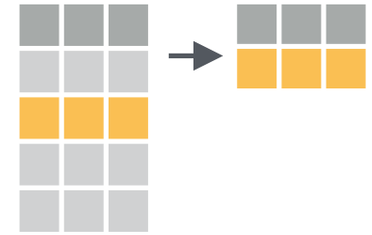
`select()`

variations across APIs, but some version of them usually exists

“verbs” for dealing with tables

`select()`

`filter()`

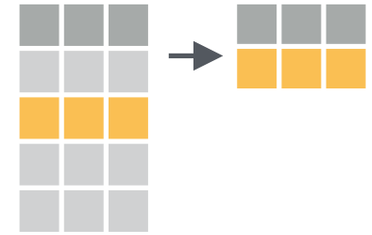


variations across APIs, but some version of them usually exists

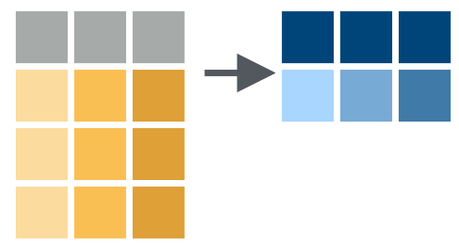
“verbs” for dealing with tables

`select()`

`filter()`



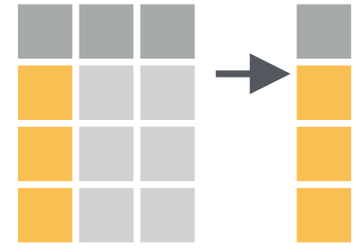
`mutate()`



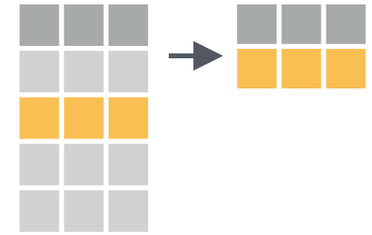
variations across APIs, but some version of them usually exists

“verbs” for dealing with tables

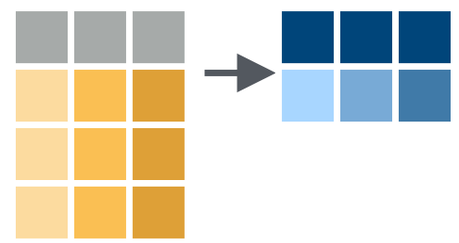
`select()`



`filter()`



`mutate()`



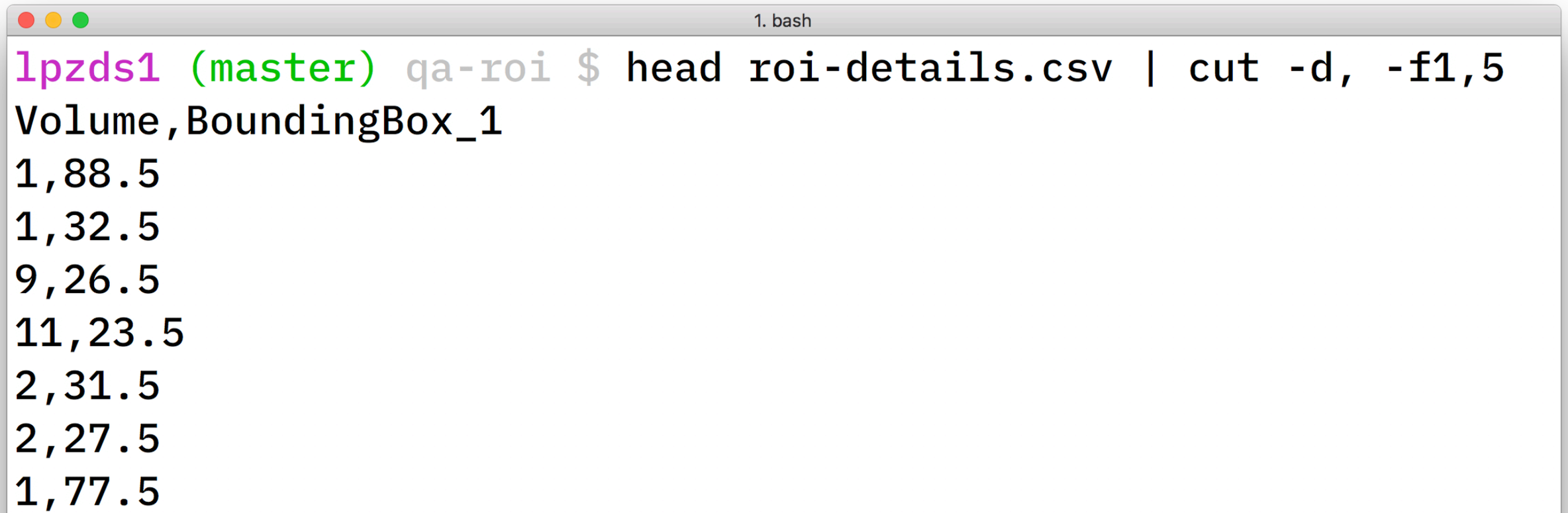
`groupby()/summarize()`

`group_by()/agg()`

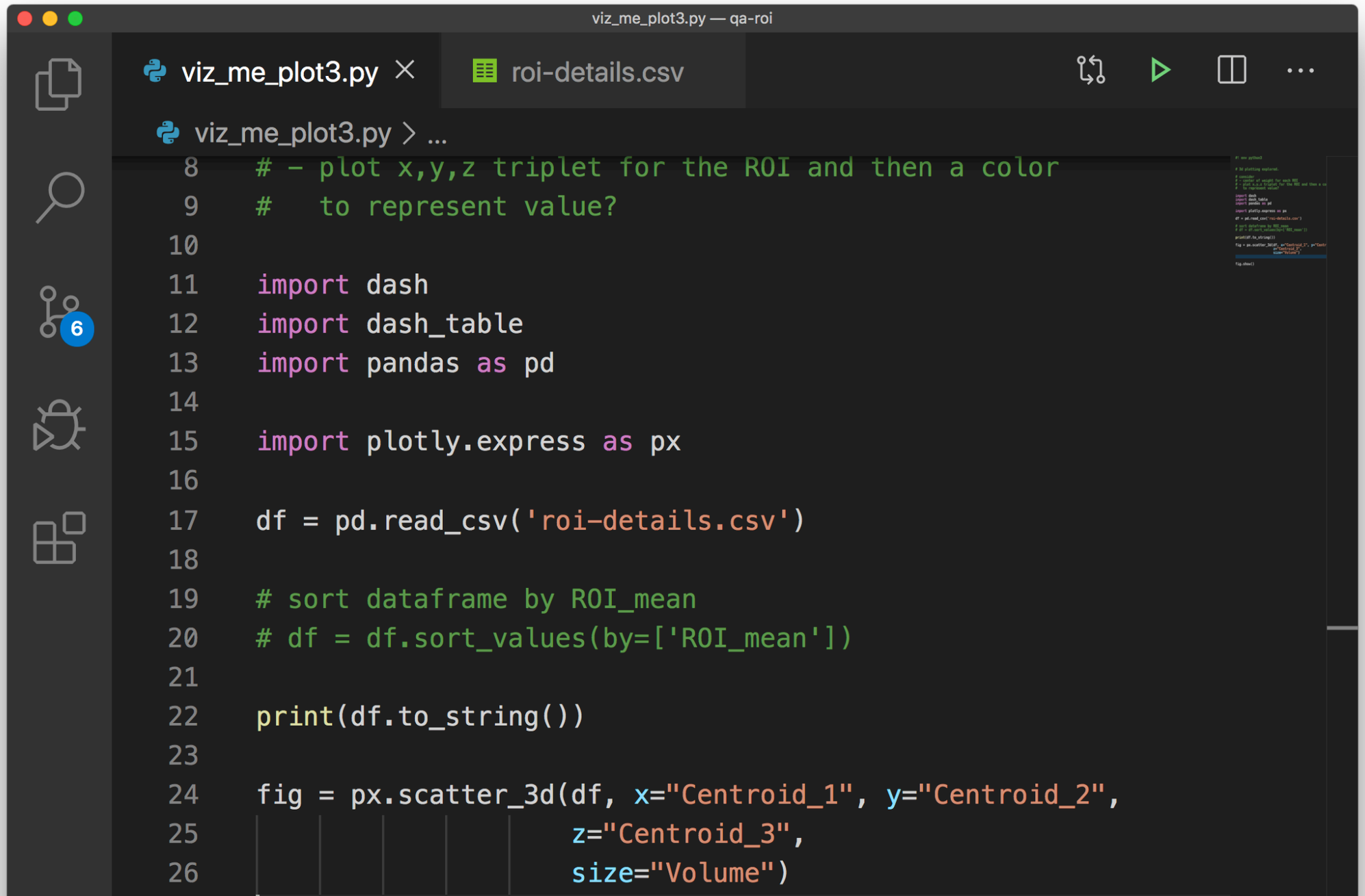
variations across APIs, but some version of them usually exists

unix magic

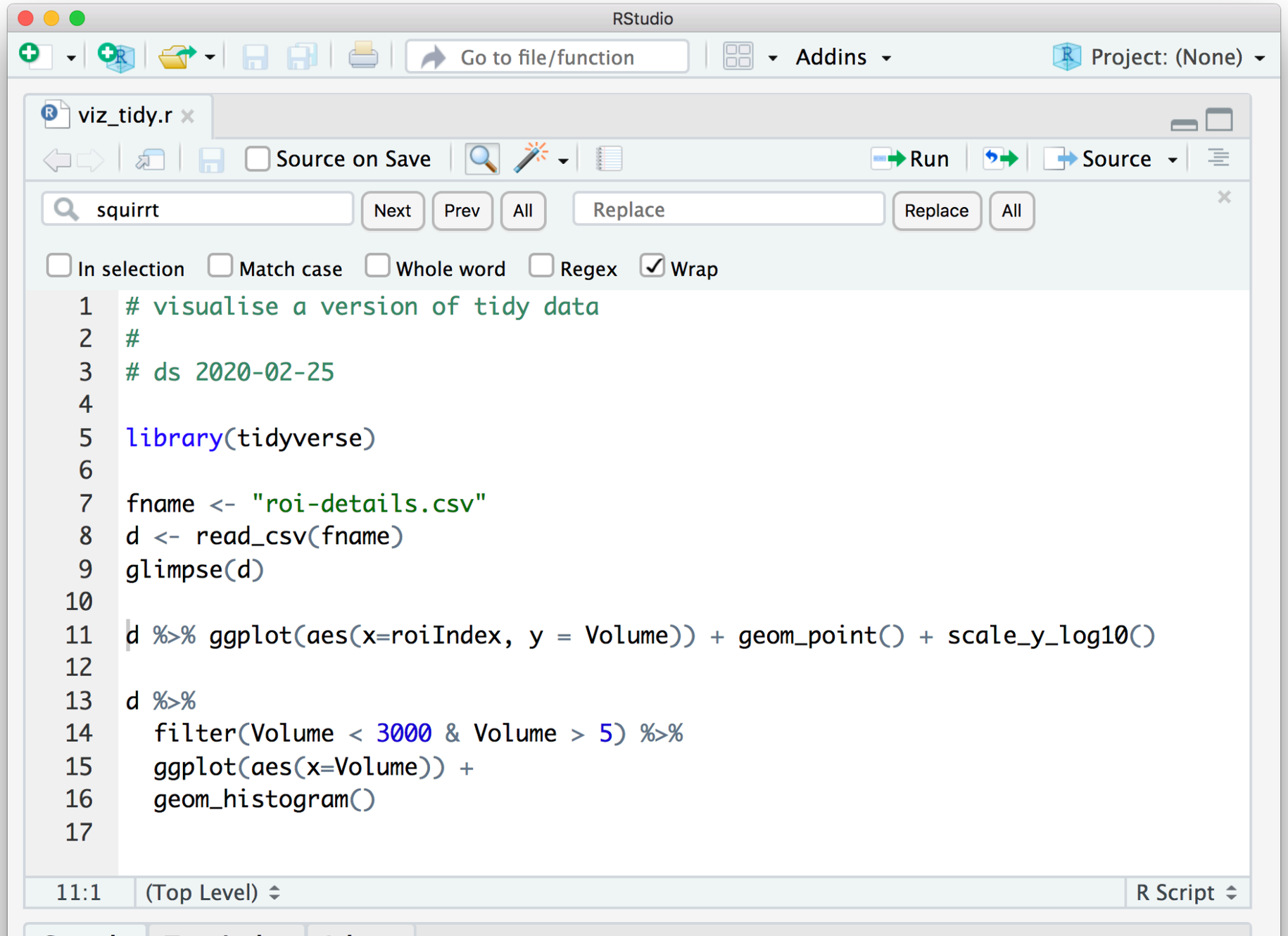
head roi-details.csv
head roi-details.csv l cut -d, -f1,5



python / pandas magic



R / dplyr magic



References

https://r4ds.hadley.nz/

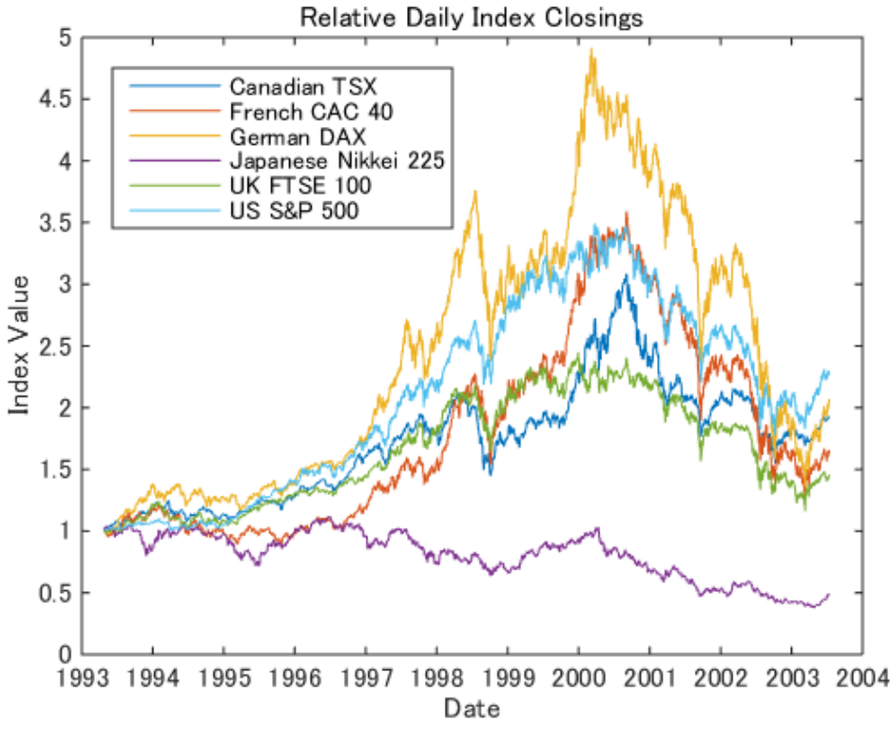
https://vita.had.co.nz/papers/tidy-data.pdf

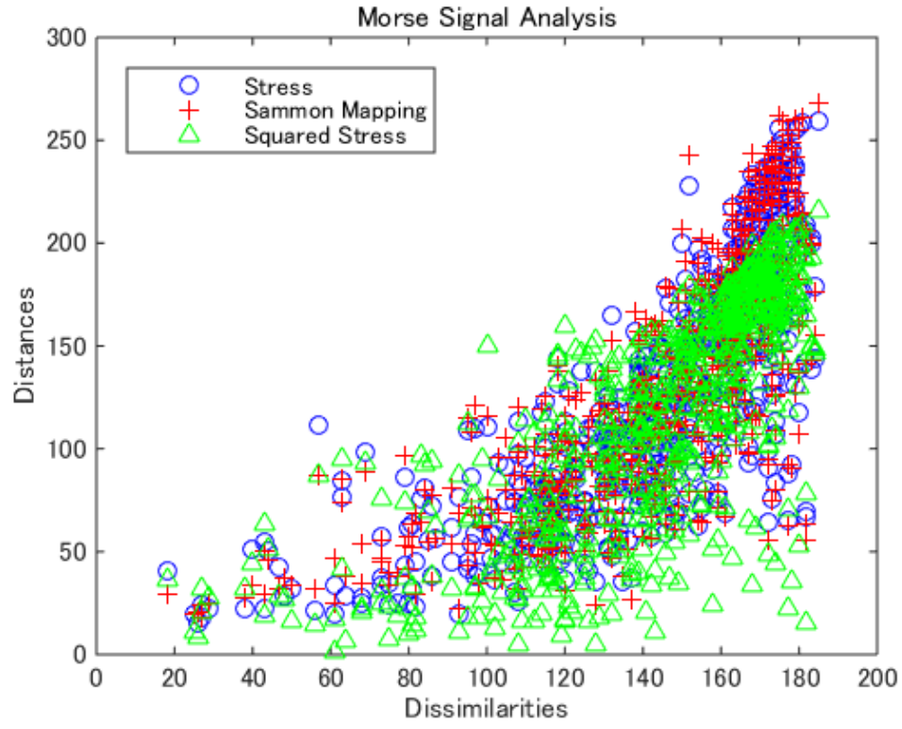
https://pandas.pydata.org/docs/

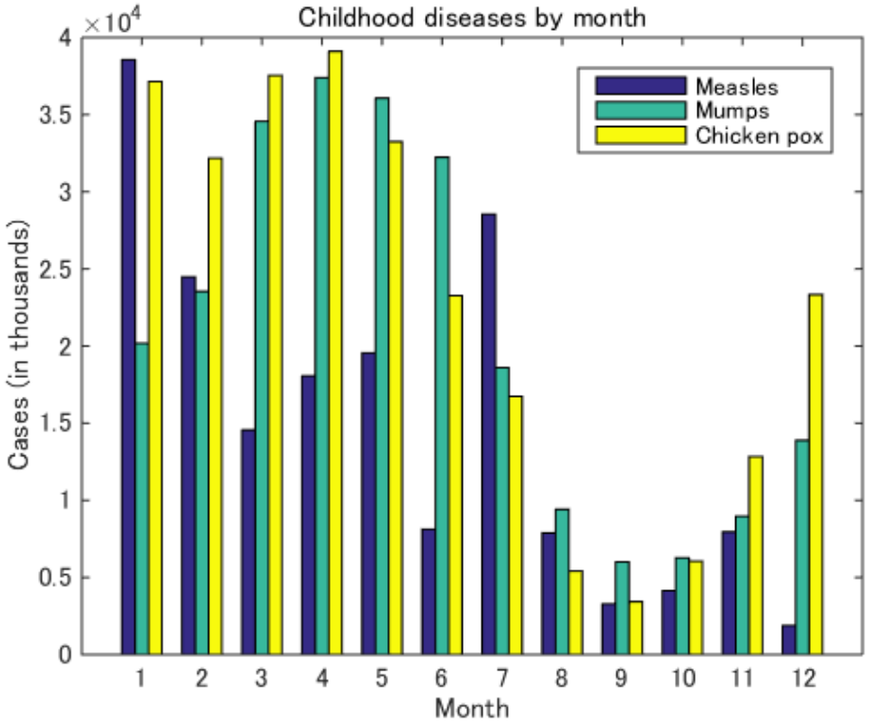
/visualise

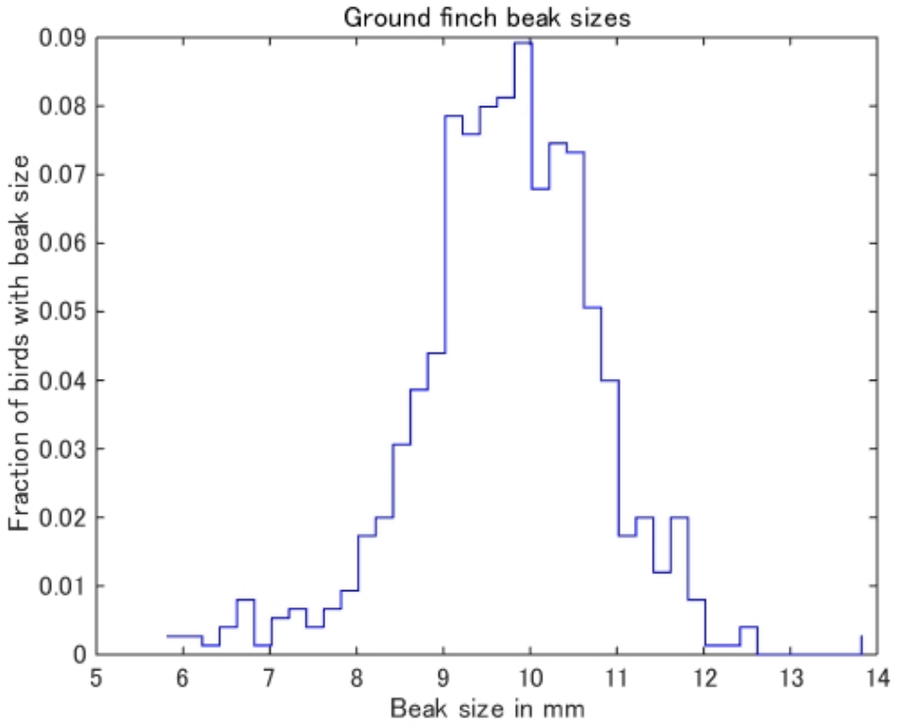
/visualize

imperative



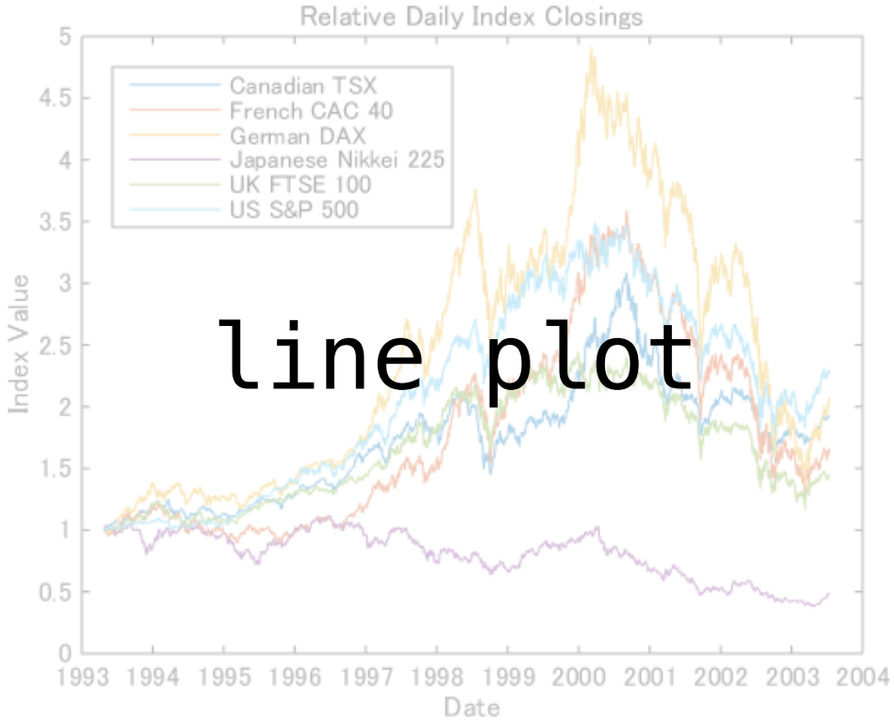


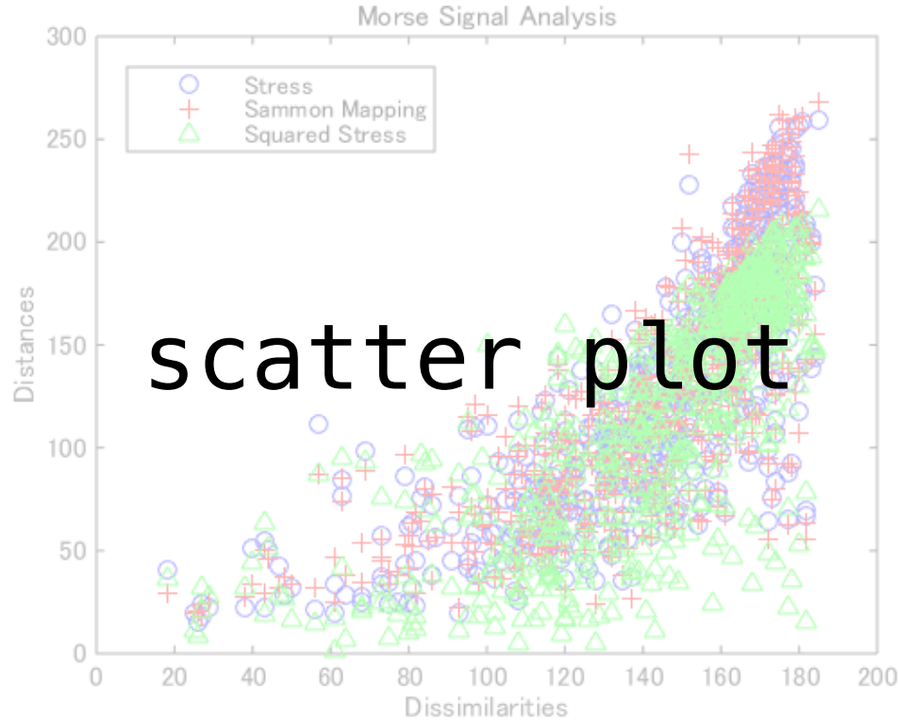


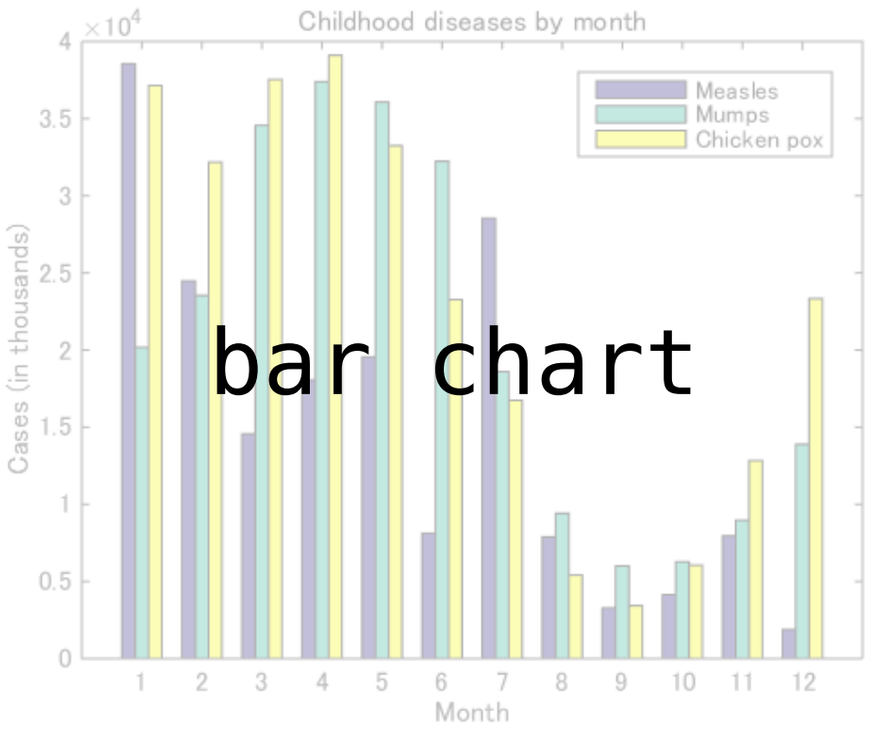


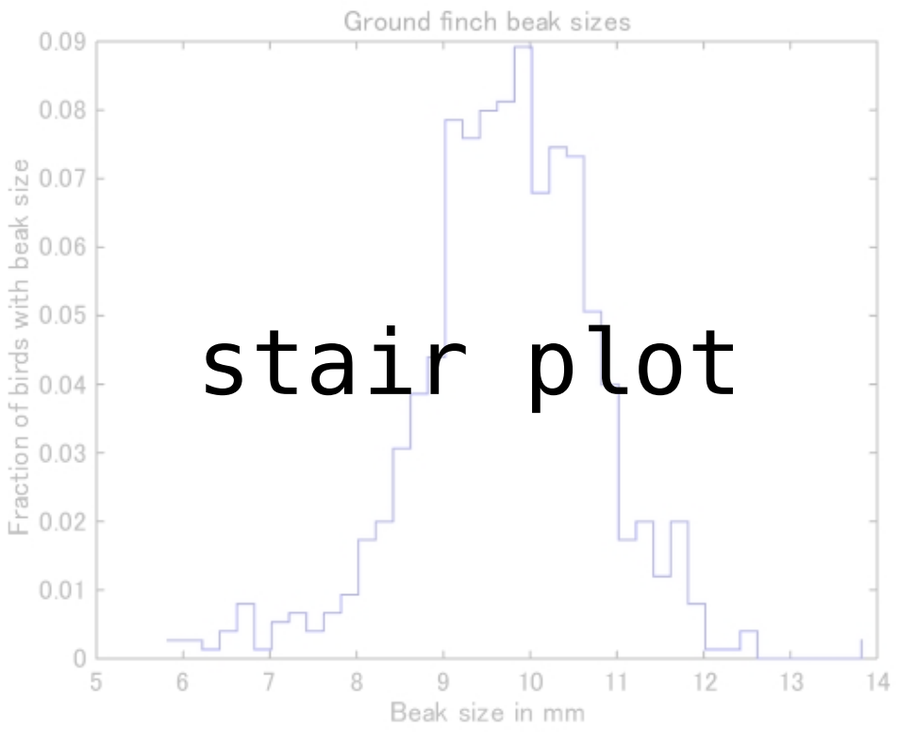
https://uk.mathworks.com/products/matlab/plot-gallery.html

imperative









https://uk.mathworks.com/products/matlab/plot-gallery.html

imperative

declarative

decide on plot type,
build plot step by step, …

–ve: leads to repetitive code / work

declarative

/dɪˈklarətɪv/

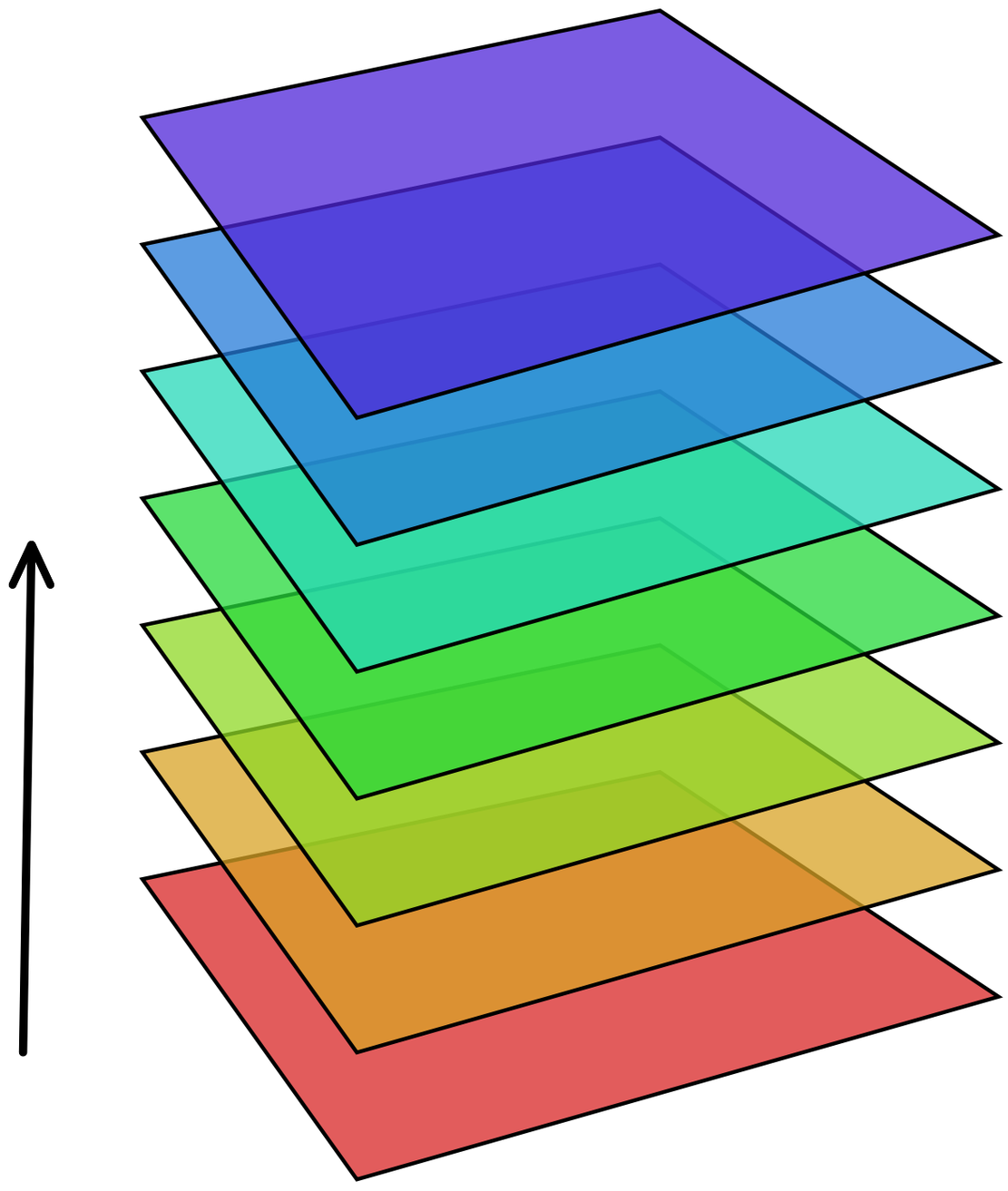
adjective

1. of the nature of or making a declaration.
   "declarative statements"
2. COMPUTING
   denoting high-level programming languages which can be used to solve problems without requiring the programmer to specify an exact procedure to be followed.

noun

1. a statement in the form of a declaration.

instead: common framework

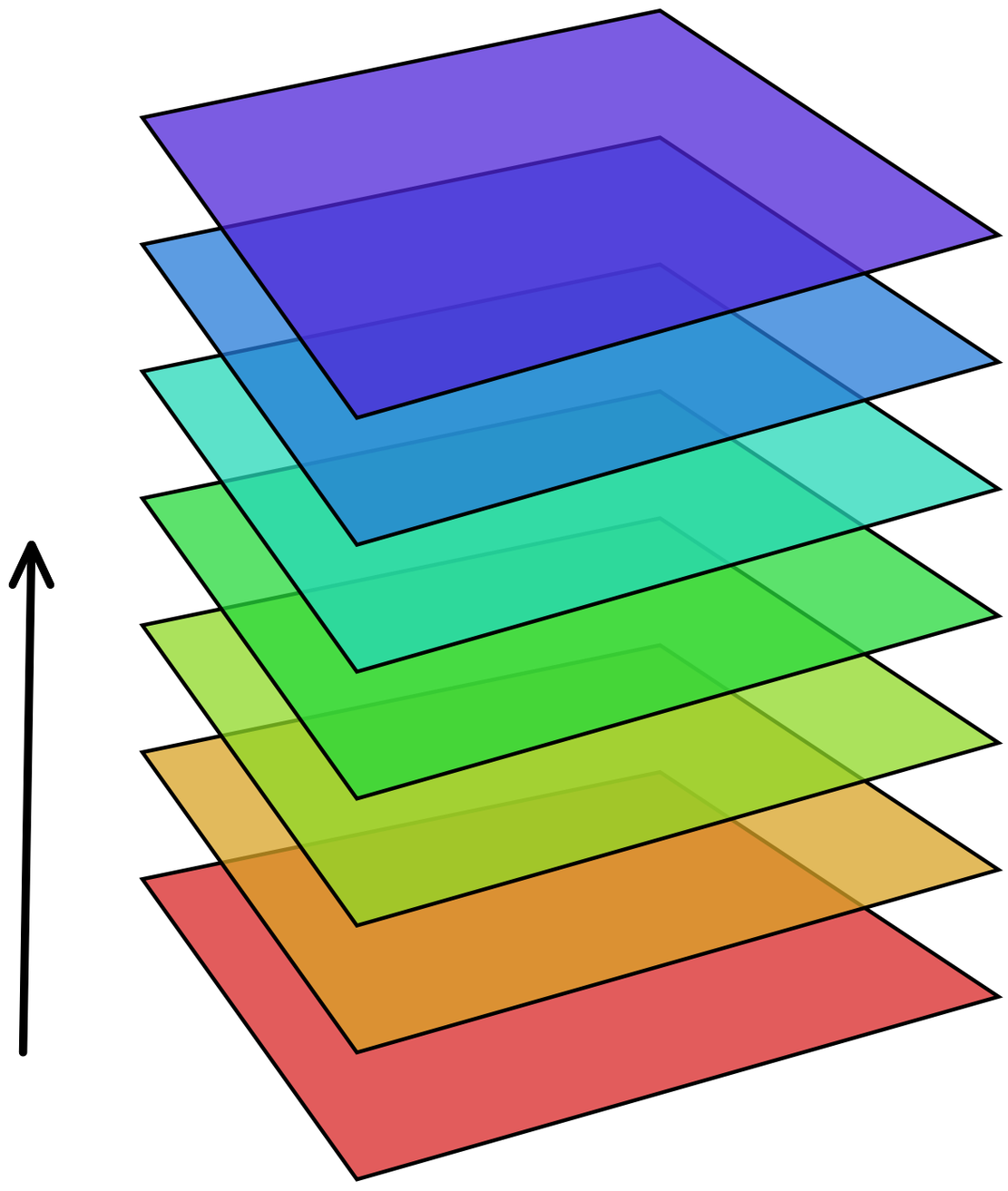


geometry

aesthetics

data

instead: common framework



coordinates

stats

facets

scales

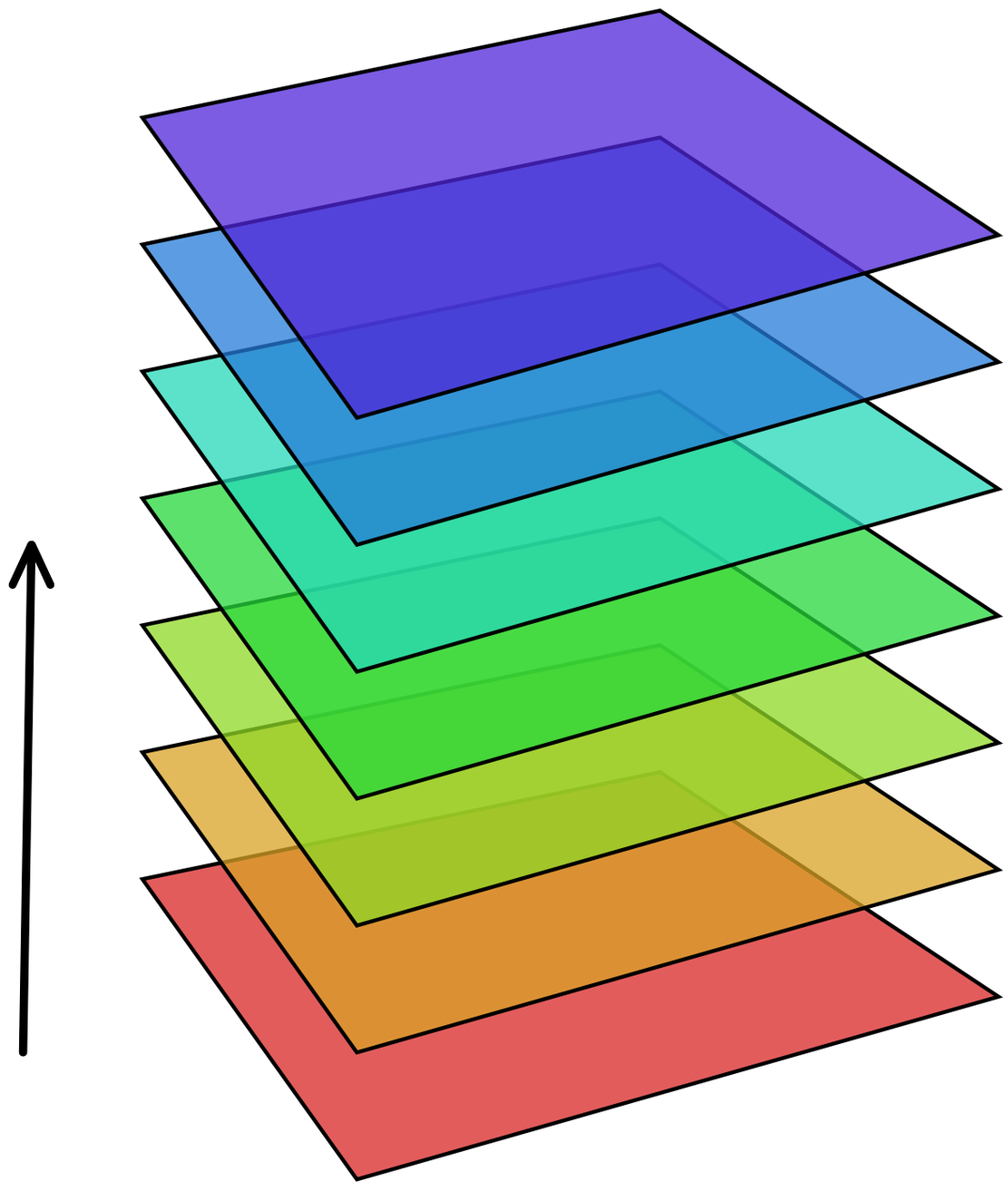
geometry

aesthetics

data

a grammar

instead: ~~common framework~~



coordinates

stats

facets

scales

geometry

aesthetics

data

an example

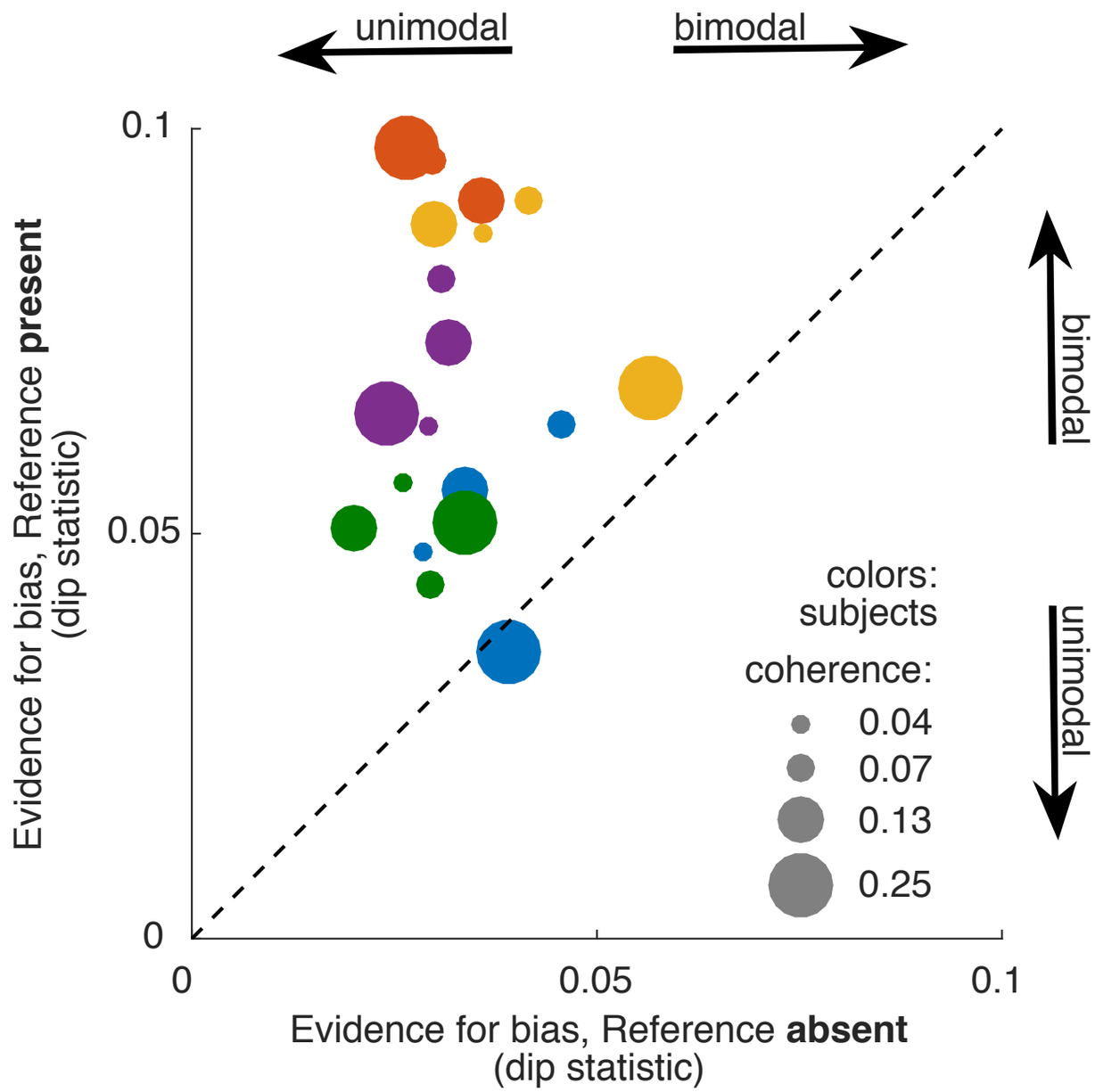


Figure 3b

Zamboni et al (2016) Proc Biol Sci.

recreating this graphic in R/ggplot2

code: https://gist.github.com/schluppeck/9a54b9b7a37793d8959779629b4cd2fc

data, d

```
head(d)
```

```
  X subject coherence   absent  present
1 1       1         4 0.035996 0.087079
2 2       2         7 0.026042 0.056272
3 3       3        13 0.028604 0.047791
4 4       4        25 0.029221 0.063149
5 5       5         4 0.025157 0.096832
6 6       1         7 0.041558 0.091160
```

4 variables that we want to **map** into a plot

aesthetics

- `x, y (position)`
- `alpha, color, fill`
- `size`
- `shape`
- `linetype`

http://docs.ggplot2.org/current/vignettes/ggplot2-specs.html

data

```
head(d)
```

```
  X subject coherence   absent  present
1 1       1         4 0.035996 0.087079
2 2       2         7 0.026042 0.056272
3 3       3        13 0.028604 0.047791
4 4       4        25 0.029221 0.063149
5 5       5         4 0.025157 0.096832
6 6       1         7 0.041558 0.091160
```

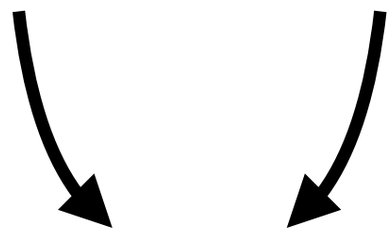
4 variables that we want to **map** into a plot

data

```
head(d)
```

```
  X subject coherence   absent  present
1 1       1         4 0.035996 0.087079
2 2       2         7 0.026042 0.056272
3 3       3        13 0.028604 0.047791
4 4       4        25 0.029221 0.063149
5 5       5         4 0.025157 0.096832
6 6       1         7 0.041558 0.091160
```

4 variables that we want to **map** into a plot



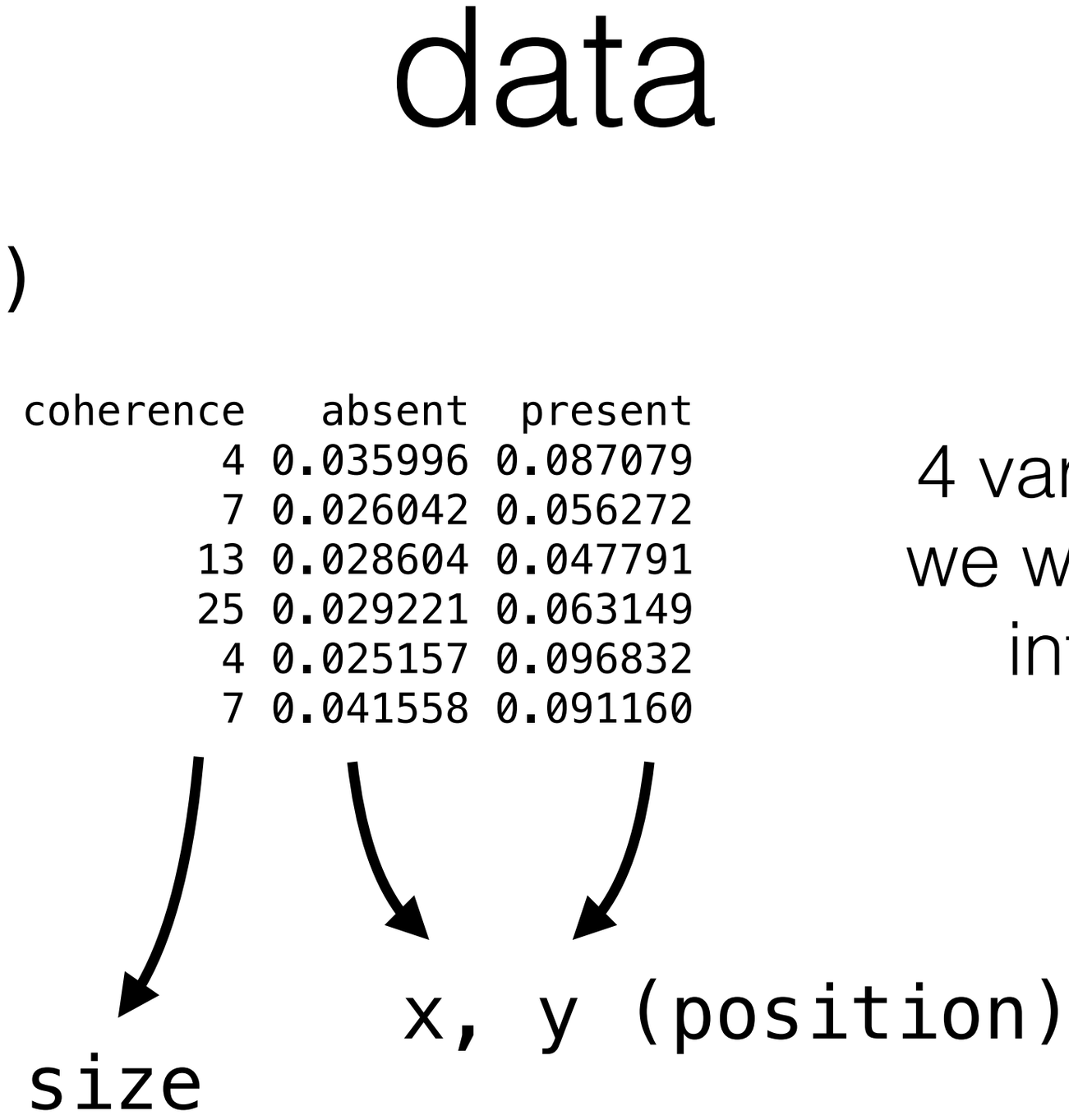
x, y (position)

data

```
head(d)
```

	X	subject	coherence	absent	present
1	1	1	4	0.035996	0.087079
2	2	2	7	0.026042	0.056272
3	3	3	13	0.028604	0.047791
4	4	4	25	0.029221	0.063149
5	5	5	4	0.025157	0.096832
6	6	1	7	0.041558	0.091160

4 variables that we want to **map** into a plot

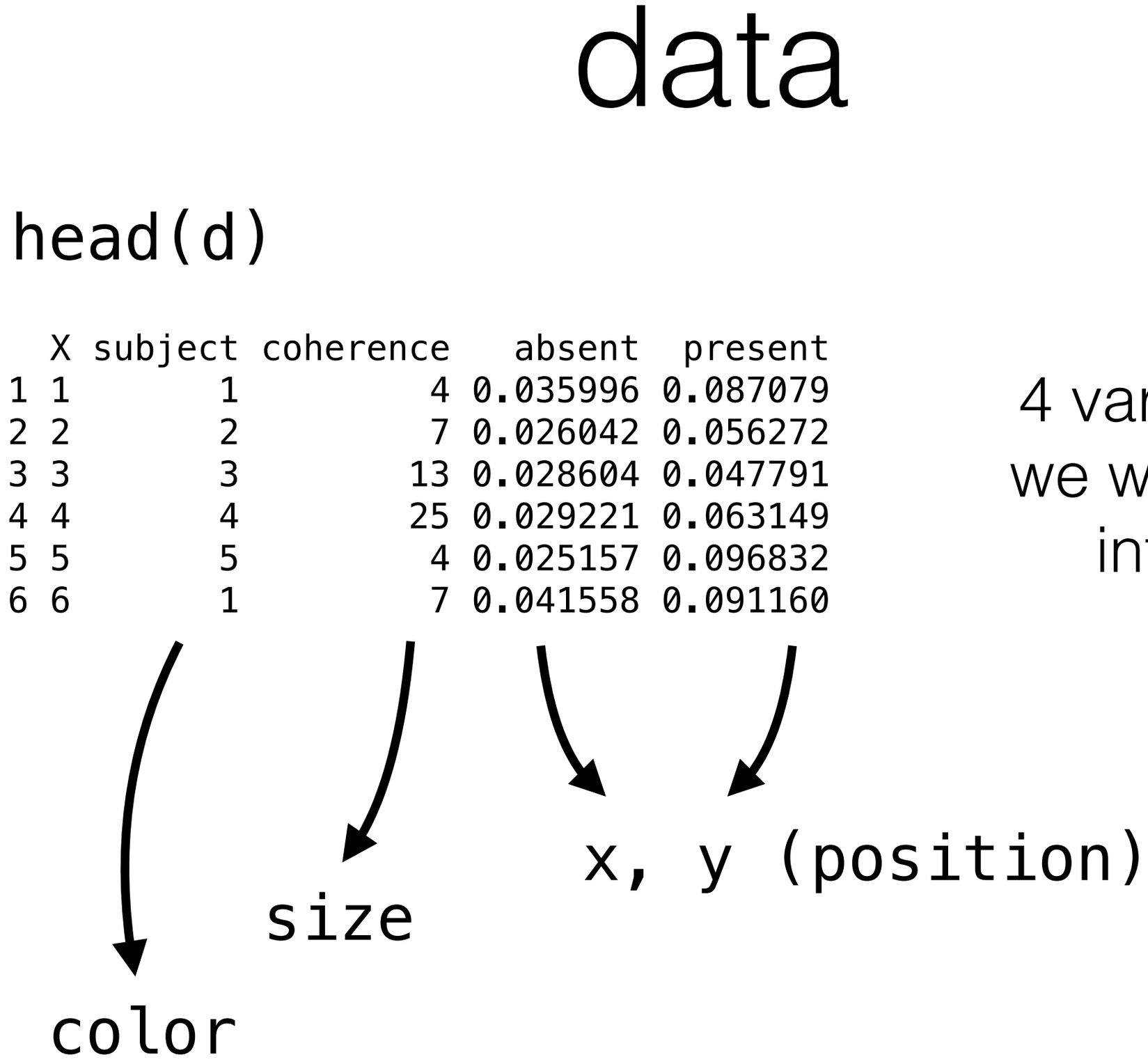


data

```
head(d)
```

```
  X subject coherence   absent  present
1 1       1         4 0.035996 0.087079
2 2       2         7 0.026042 0.056272
3 3       3        13 0.028604 0.047791
4 4       4        25 0.029221 0.063149
5 5       5         4 0.025157 0.096832
6 6       1         7 0.041558 0.091160
```

4 variables that we want to **map** into a plot



geometry

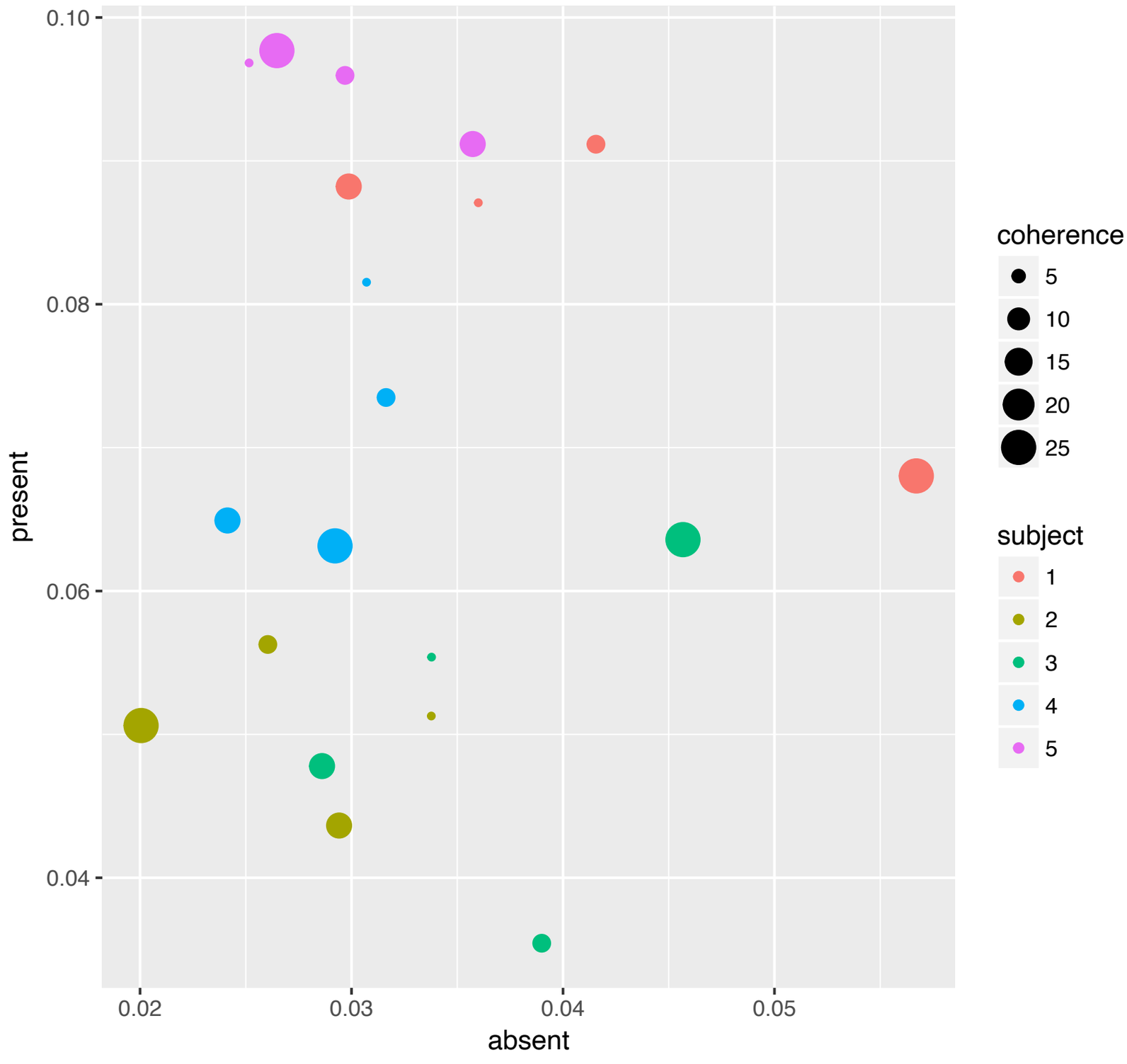
- 0d: points, text
- 1d: lines, paths
- 2d: polygons, intervals

Wickham (2010)

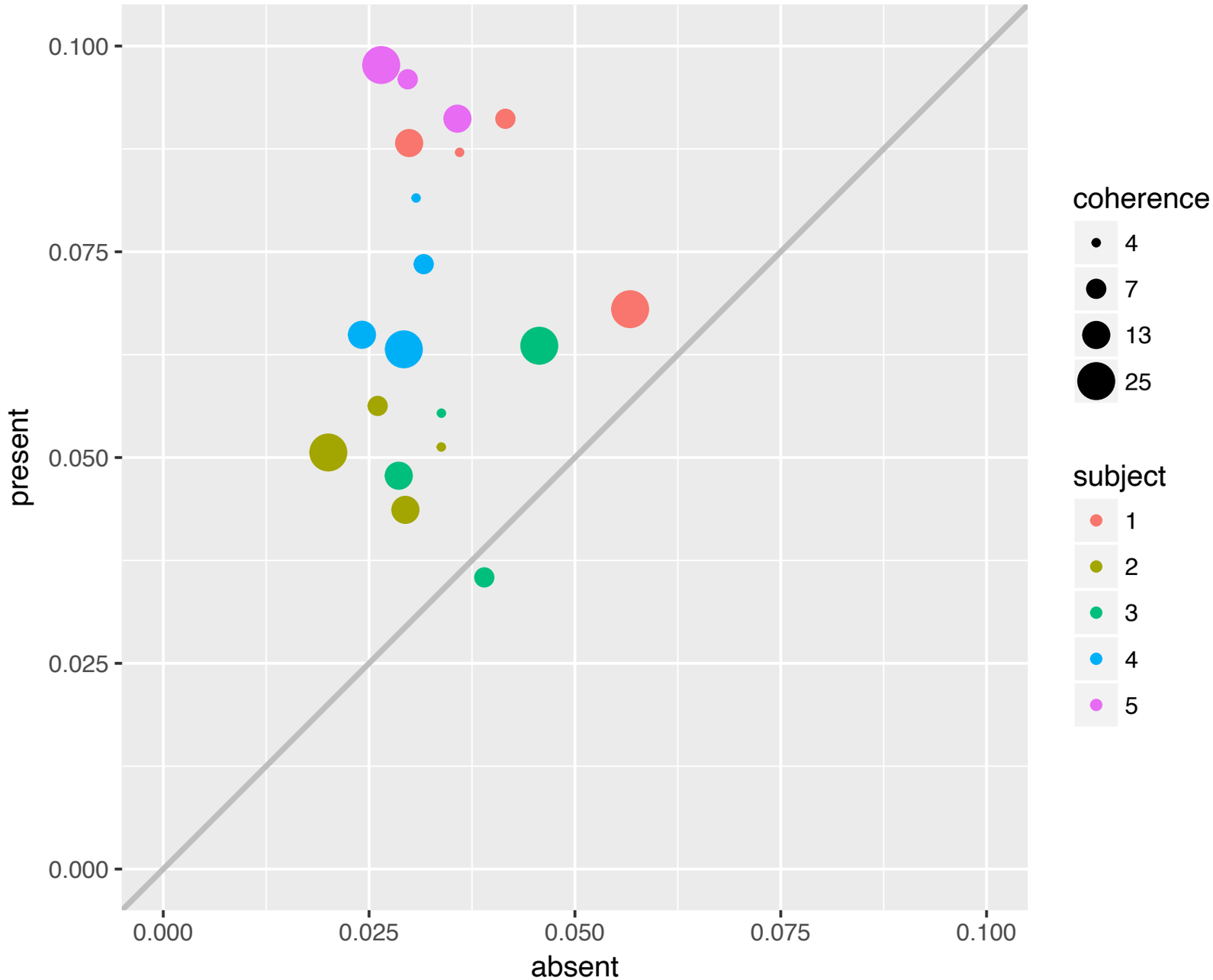
geometry

- 0d: points, text 
- 1d: lines, paths
- 2d: polygons, intervals

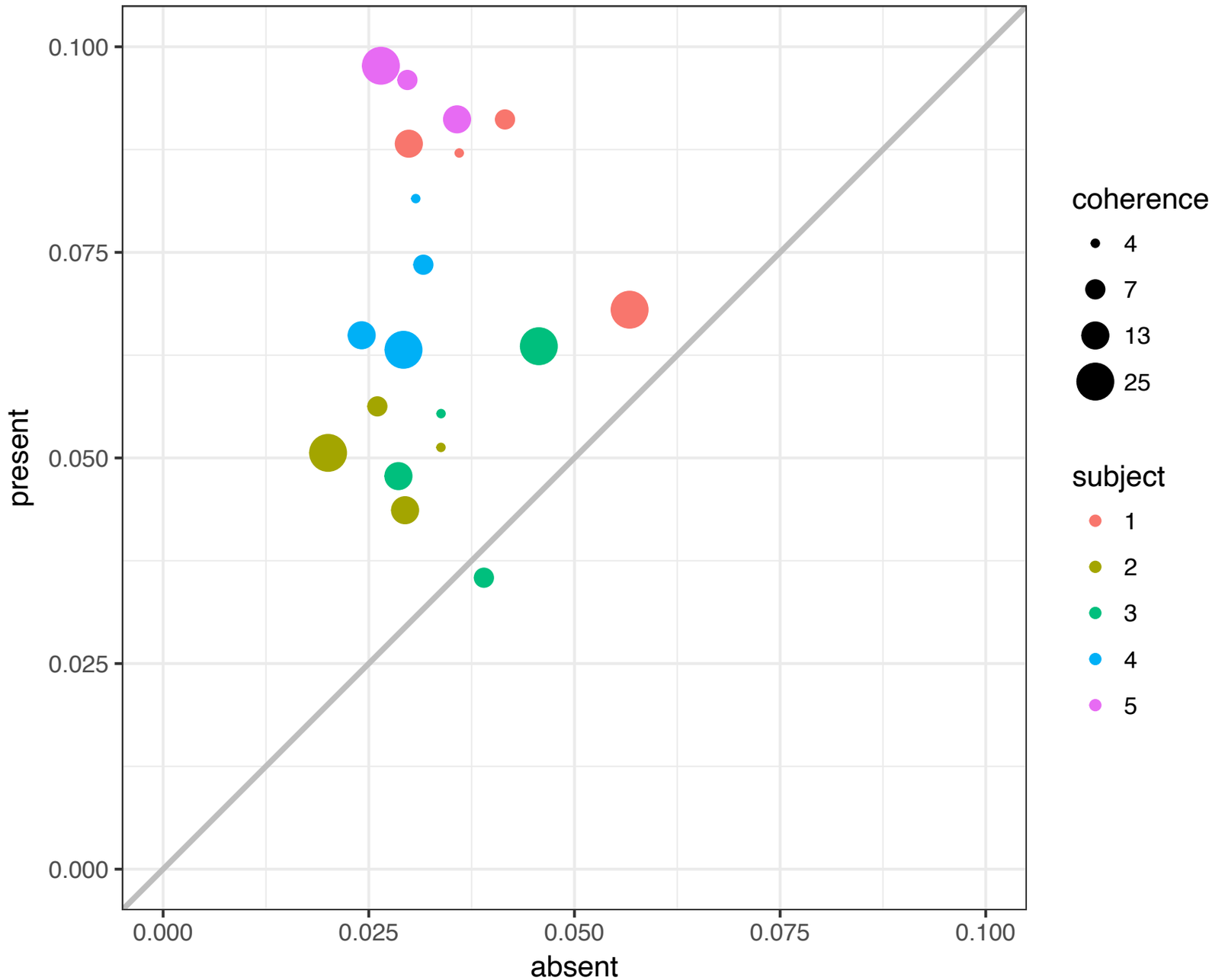
Wickham (2010)



with default settings

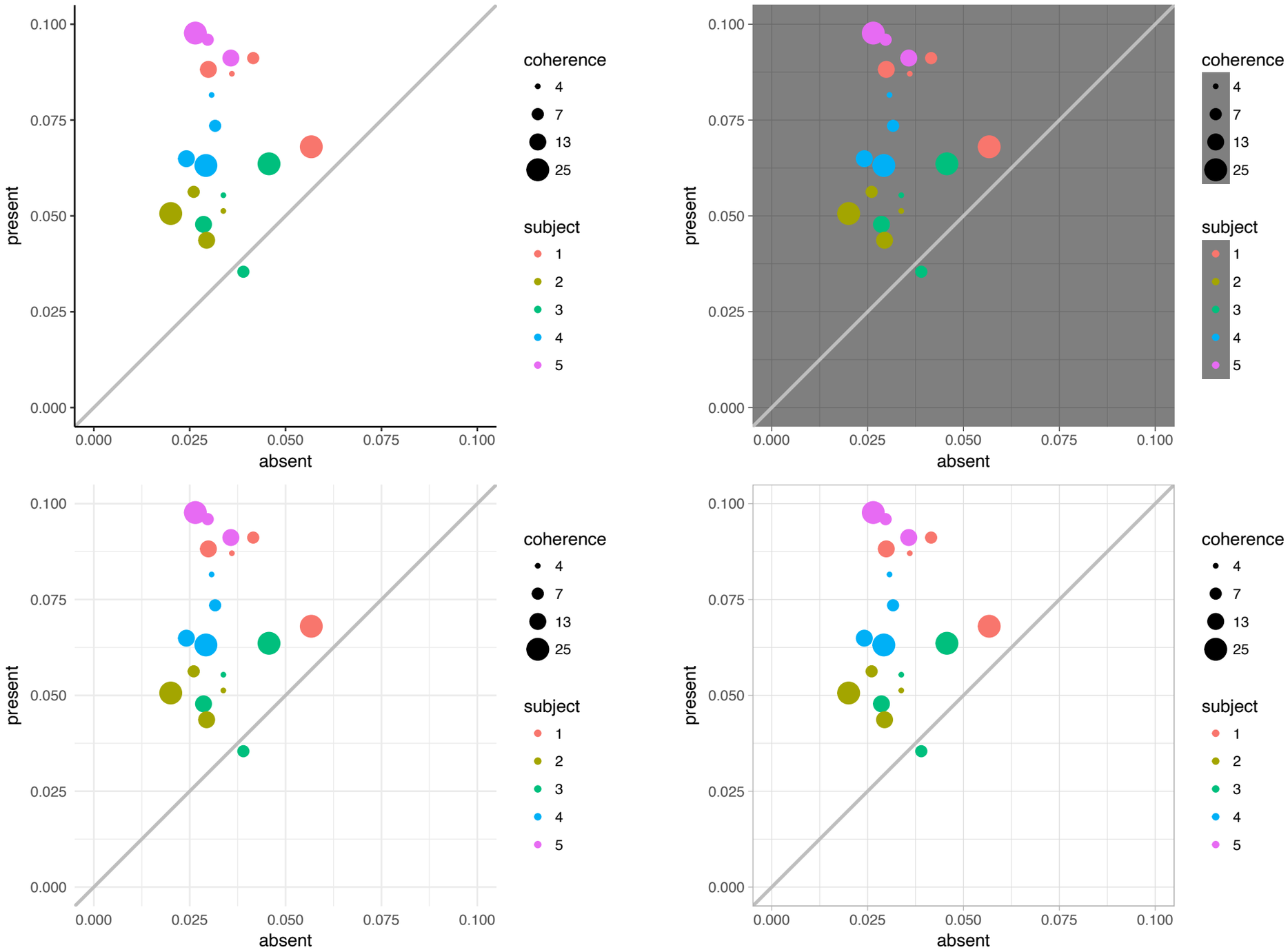


+ scale, coord (aspect ratio), unity line



+ scale, coord (aspect ratio), unity line

themes / look of plot



worth the hassle?

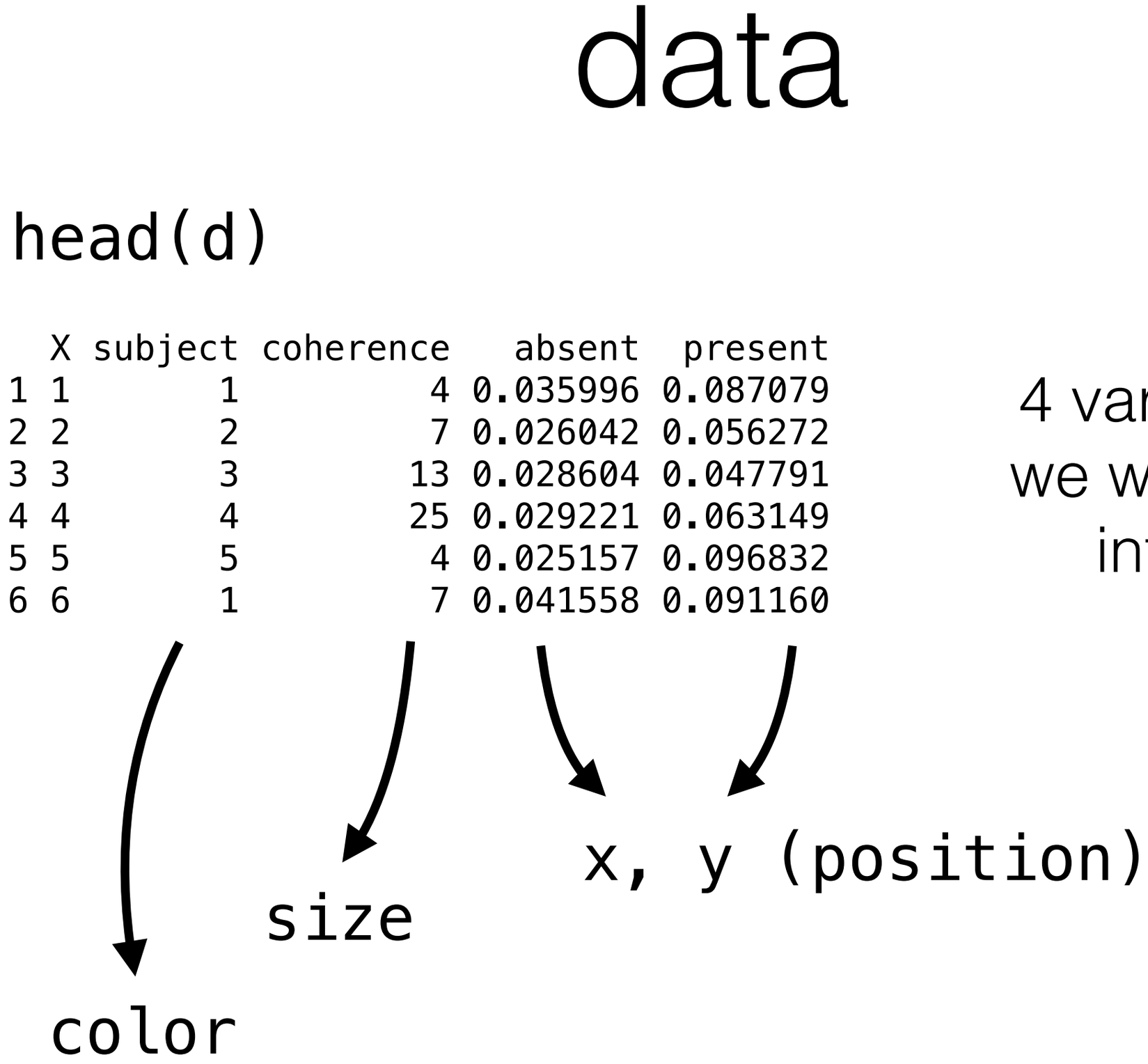
- I think yes: already for basic plotting
- for data exploration we often slice across different dimensions:
  - subjects, regions of interest, …
  - measures: RT, % correct, fMRI response amplitude, …

data

```
head(d)
```

```
  X subject coherence   absent  present
1 1       1         4 0.035996 0.087079
2 2       2         7 0.026042 0.056272
3 3       3        13 0.028604 0.047791
4 4       4        25 0.029221 0.063149
5 5       5         4 0.025157 0.096832
6 6       1         7 0.041558 0.091160
```

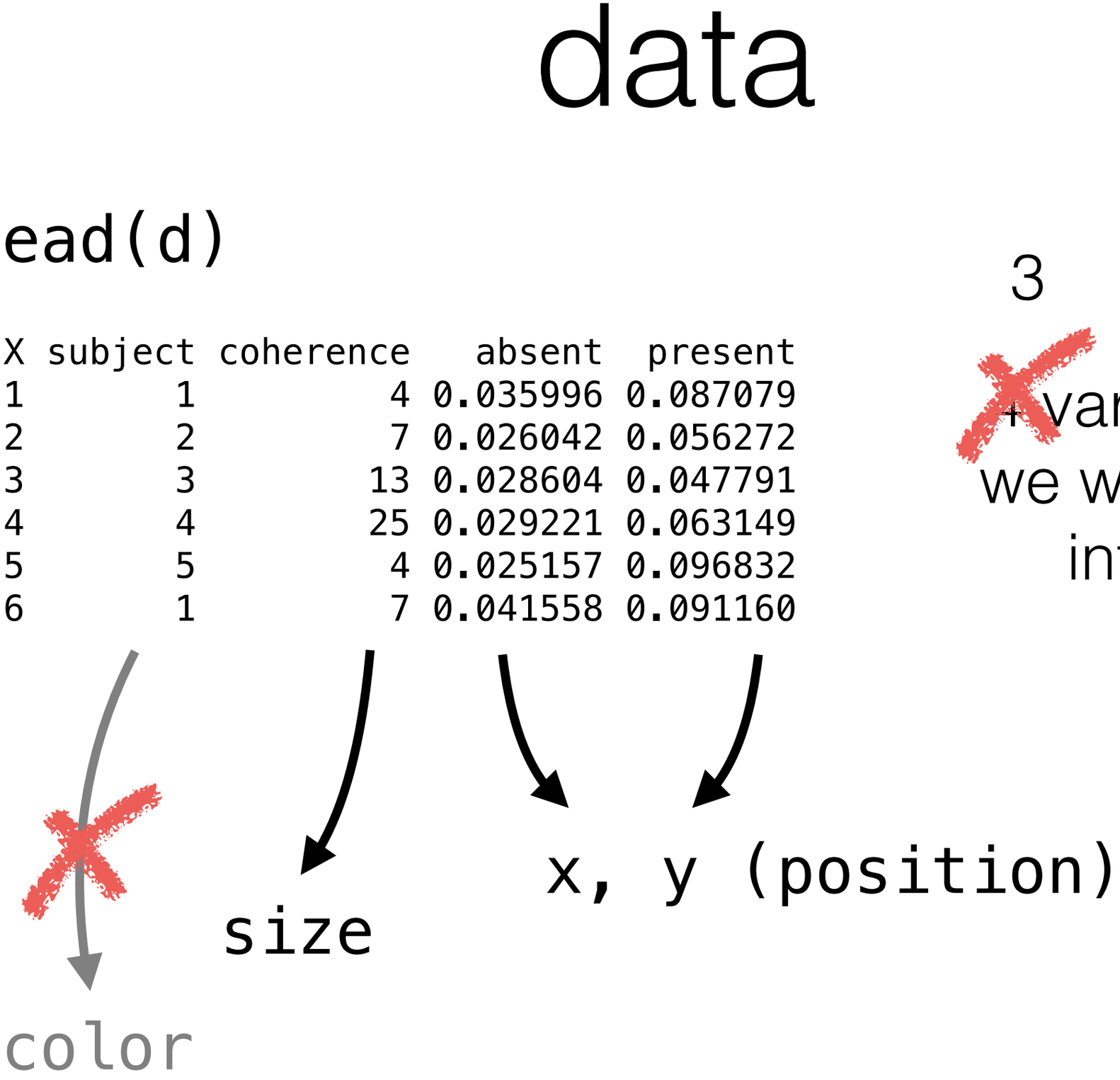
4 variables that we want to **map** into a plot



data

```
head(d)
```

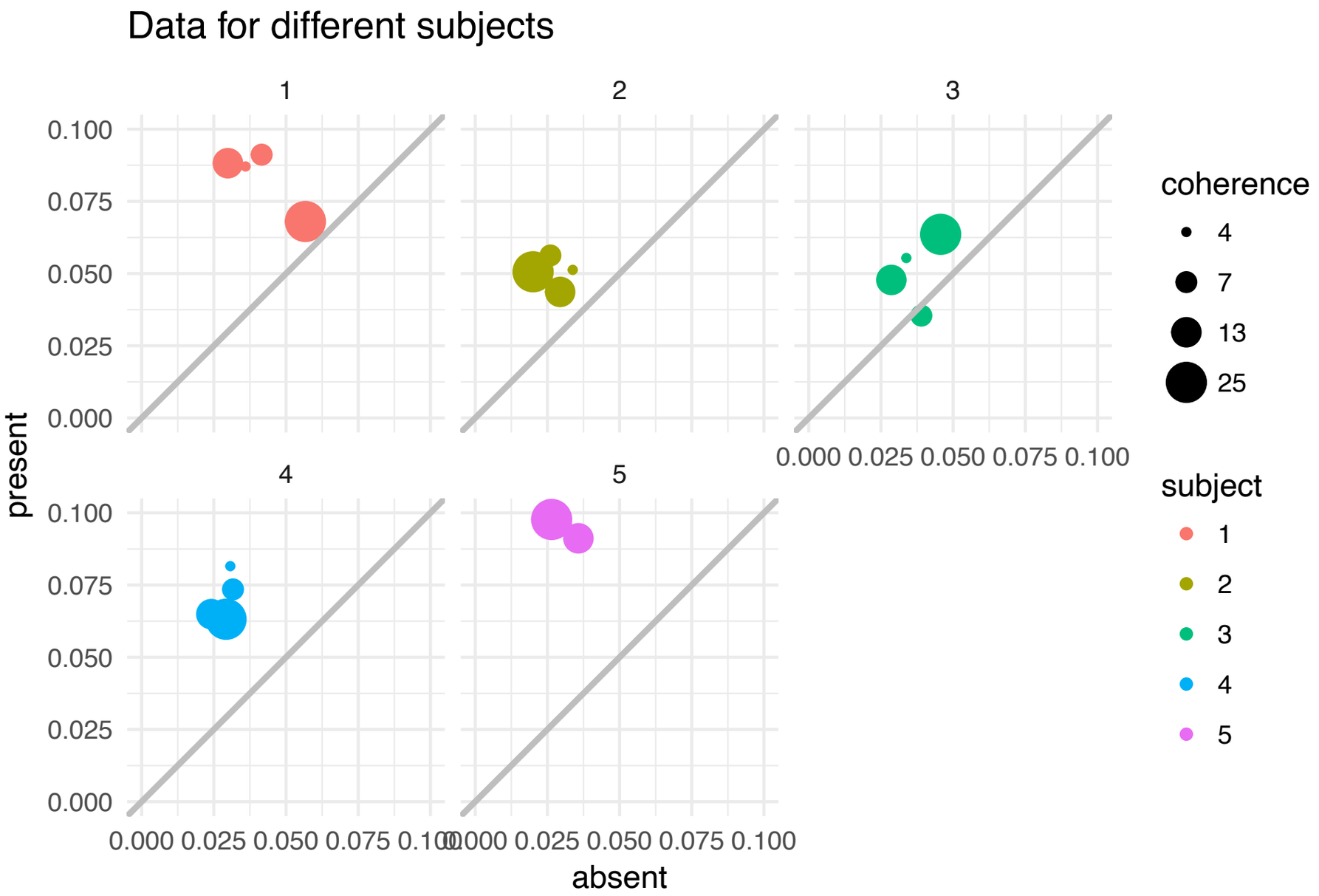
```
  X subject coherence   absent  present
1 1       1         4 0.035996 0.087079
2 2       2         7 0.026042 0.056272
3 3       3        13 0.028604 0.047791
4 4       4        25 0.029221 0.063149
5 5       5         4 0.025157 0.096832
6 6       1         7 0.041558 0.091160
```



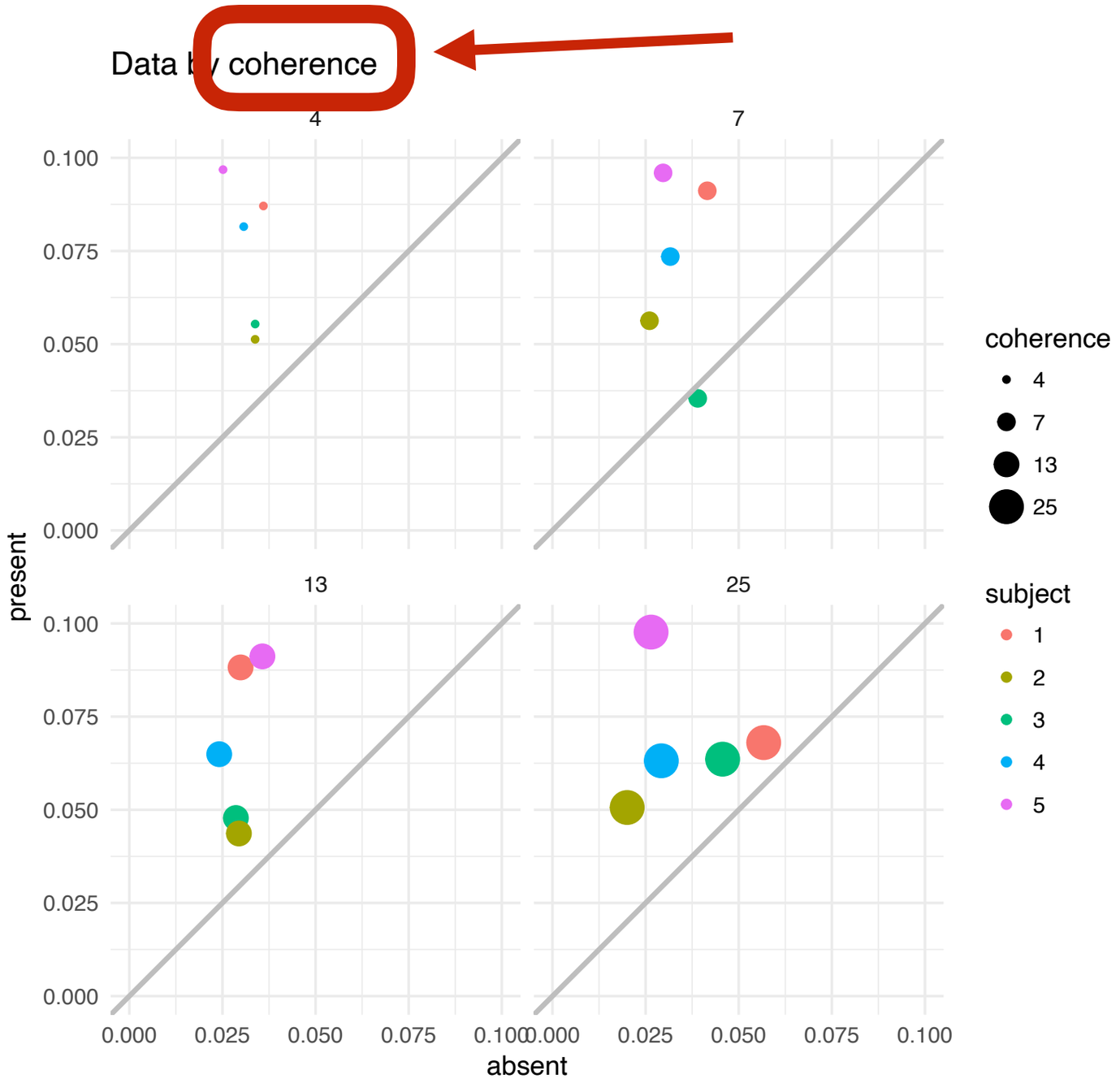
3

~~x~~ variables that we want to **map** into a plot

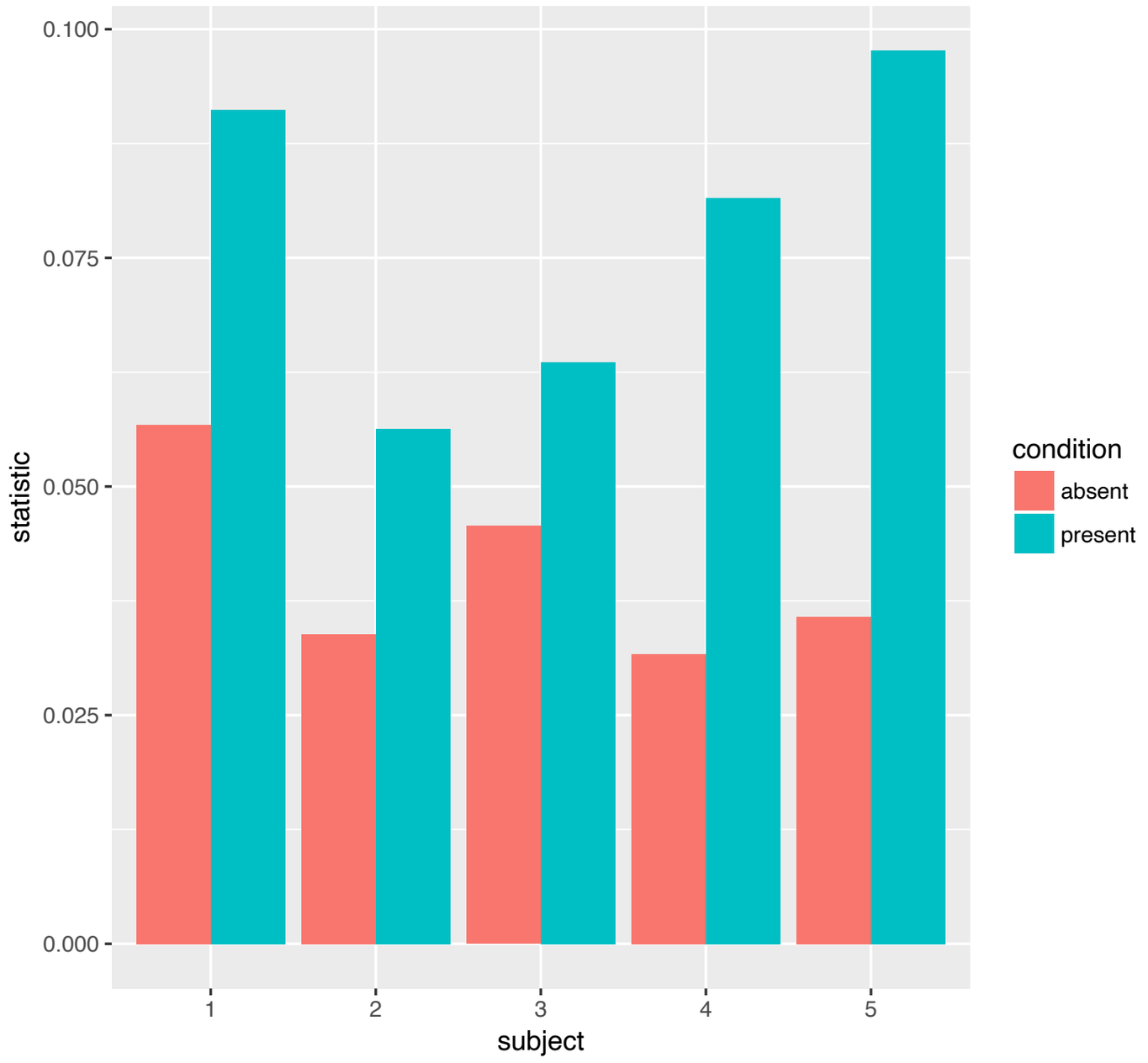
facet (lattice)



facet (lattice)



rearrange

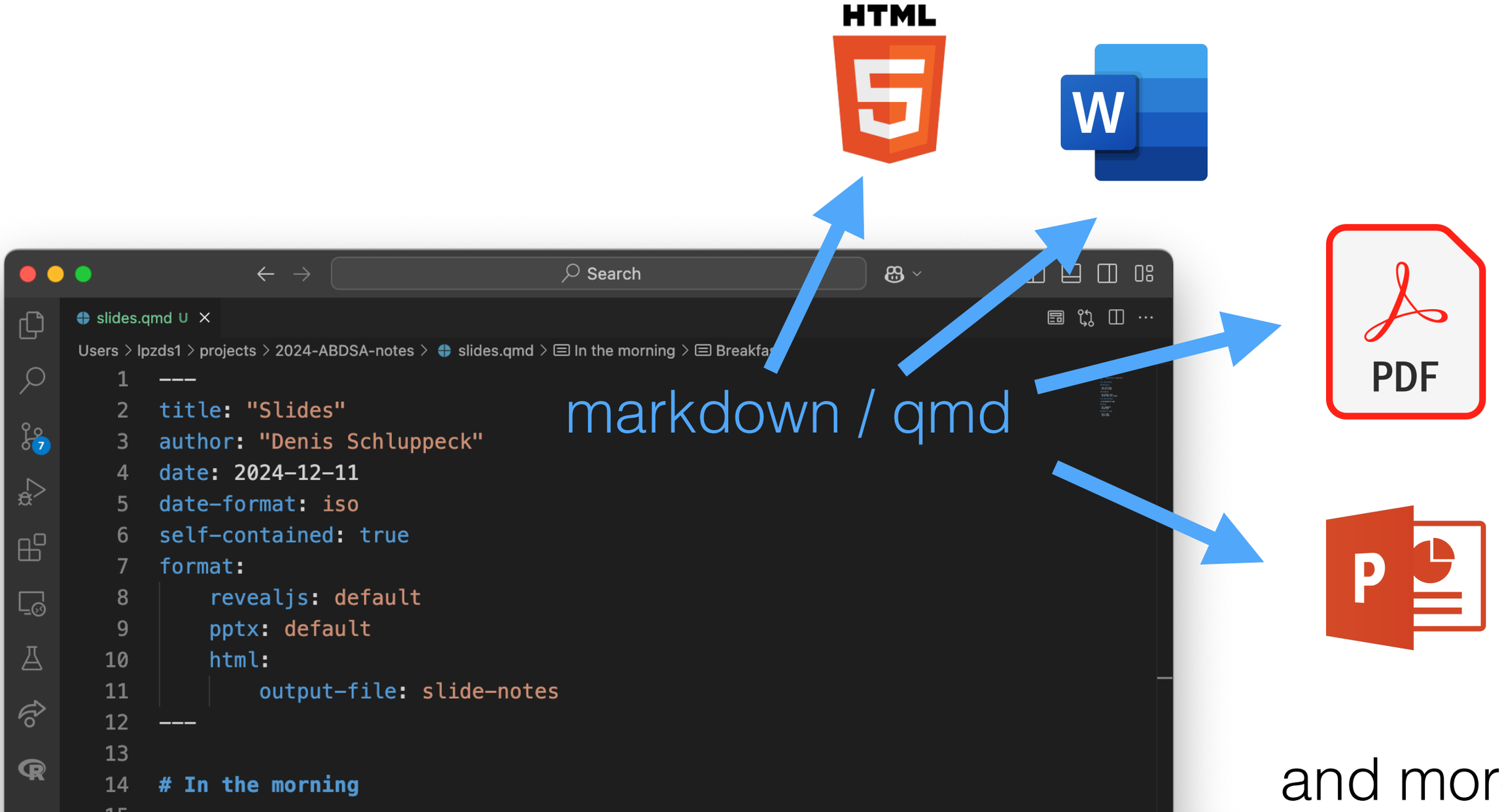

0.100
0.075
0.050
0.025
0.000
statistic
1
2
3
4
5
subject
condition
absent
present

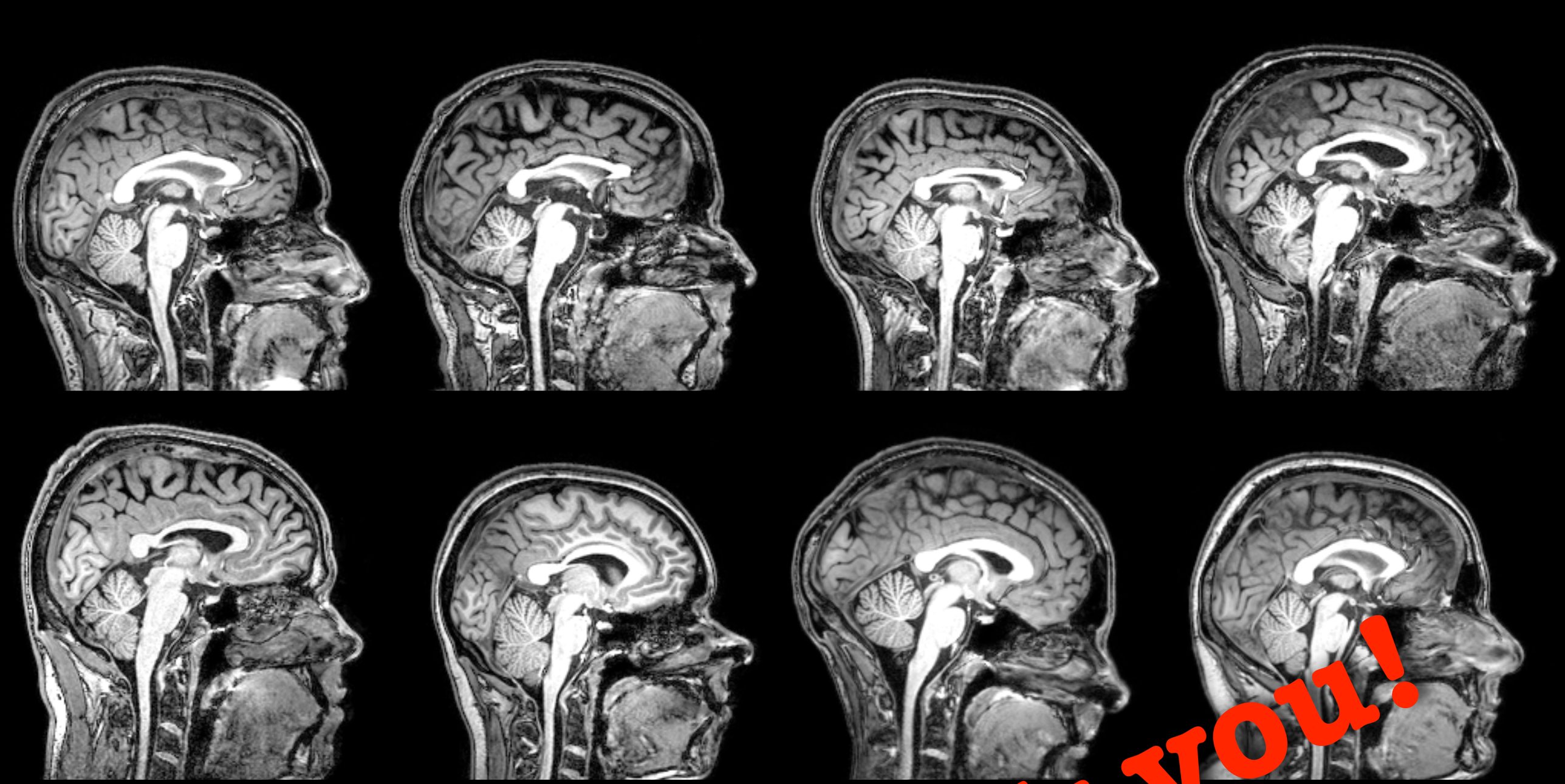
Examples

https://github.com/schluppeck/2024-ABDSA-notes

demo: editing on github website

https://quarto.org/


HTML
W
PDF
Search
slides.qmd U
Users > lpzds1 > projects > 2024-ABDSA-notes > slides.qmd > In the morning > Breakfa
1 ---
2 title: "Slides"
3 author: "Denis Schluppeck"
4 date: 2024-12-11
5 date-format: iso
6 self-contained: true
7 format:
8 revealjs: default
9 pptx: default
10 html:
11 output-file: slide-notes
12 ---
13
14 # In the morning
15
16 ## Getting up
17
18 - Turn off alarm
19 - Get out of bed
20
21 ## Breakfast
22
23 - Eat Kenyan fruit
24 - Drink *lots of* coffee
25
26 # In the afternoon
markdown / qmd
and more


Thank you!