Introduction to Data Visualization in R

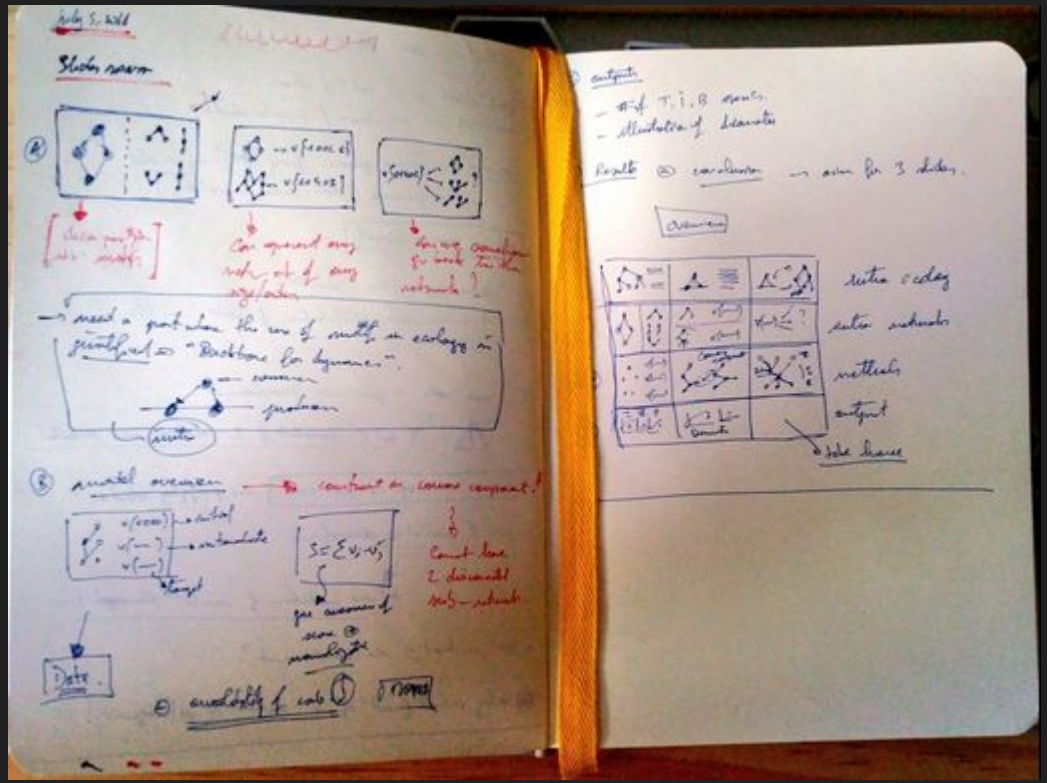
Scott McCain

Before diving into R...

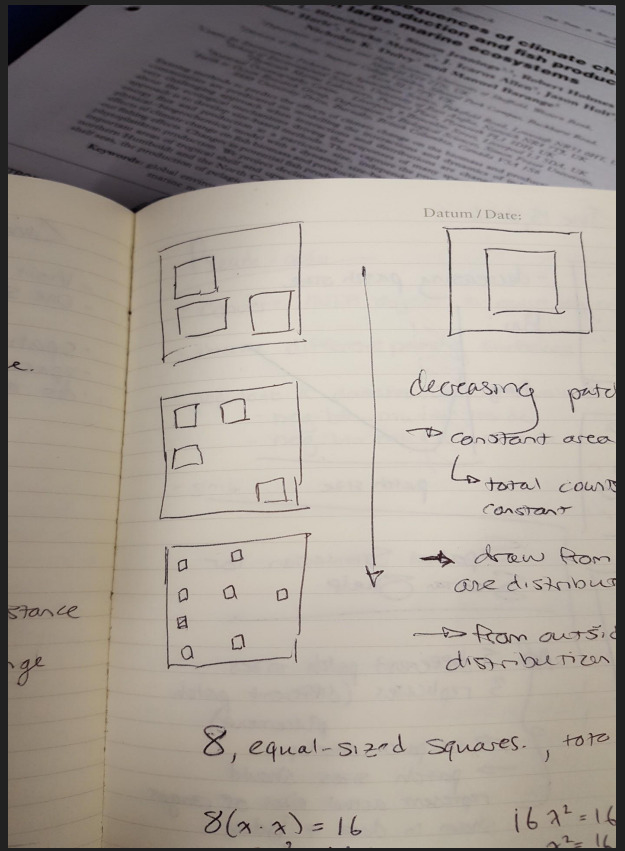
What is your message?

How can you be as transparent as possible?

Sketch it out!



^Tim Poisot (captured from @tpoi)



Transparency

1) If possible, plot your *raw* data

    Maximizing transparency,

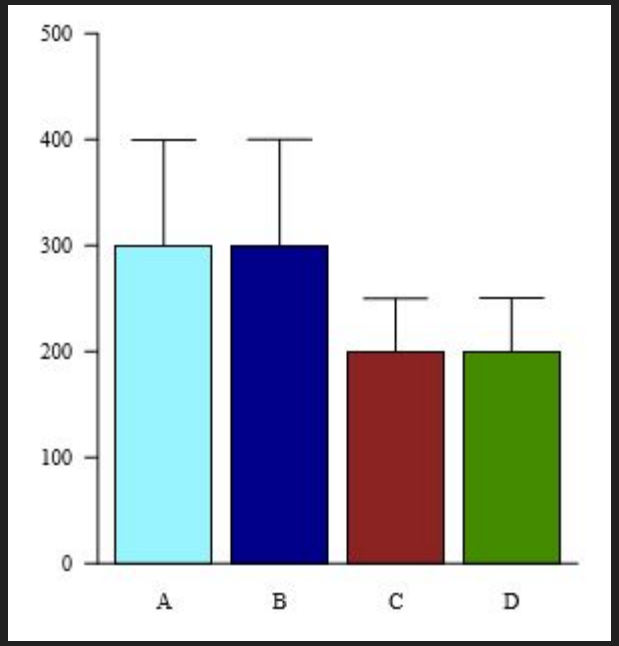
    Easier for people to conduct meta-analyses (citations!),

    Open science!

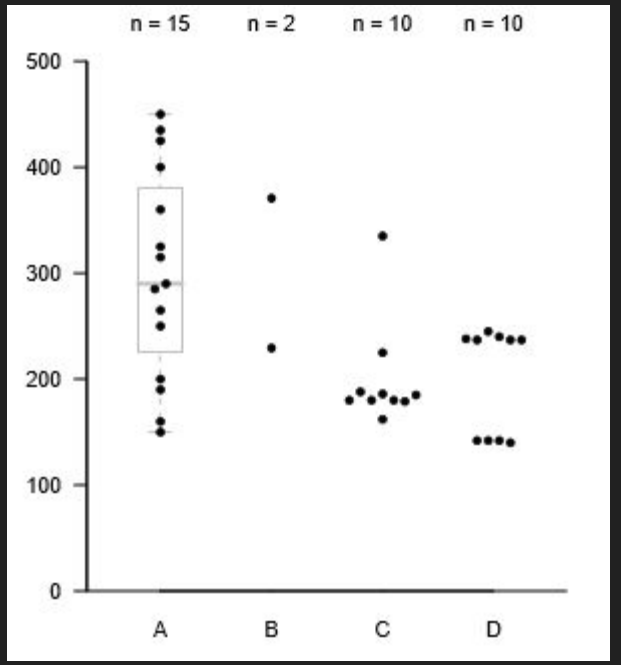
2) If not possible (i.e. looks messy), plot figures that reflect the distribution of your data

    Boxplots, violin plots

    Bar plots with mean +/- SD are not recommended



<

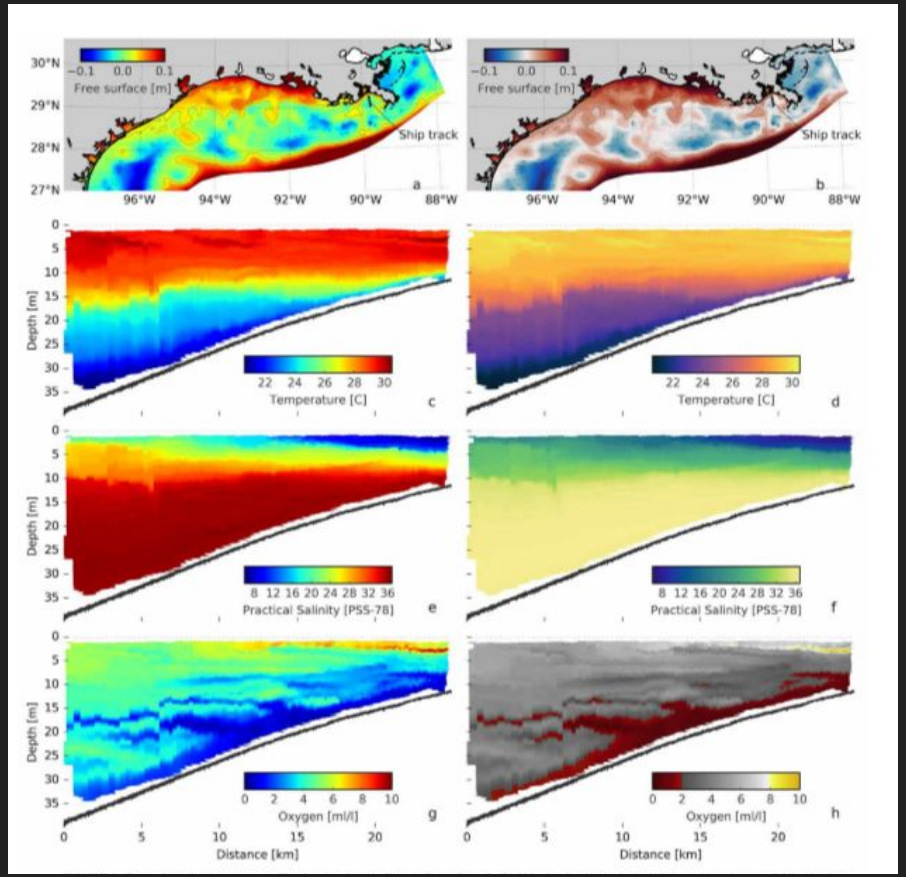


Images captured from Tatsuki Koyama poster,
“Beware of Dynamite”
biostat.mc.vanderbilt.edu/wiki/pub/Main/TatsukiRcode/Poster3.pdf

Colour Choice

1) Colour scheme should reflect the data values
    a) If the colours reflect categorical, unordered variables, it might not be best to have them as a gradient!

2) Colour-blind friendly, people!



Thyng *et al.* 2016. True colours of oceanography: Guidelines for effective and accurate colourmap selection, *Oceanography*.

Colour Resources

R Colour Brewer:

http://colorbrewer2.org/#

R Colour Brewer Tutorial:

https://www.r-bloggers.com/r-using-rcolorbrewer-to-colour-your-figures-in-r/

Using Colour in R (Dr. Jenny Bryan, R and statistics guru, UBC)

https://www.stat.ubc.ca/~jenny/STAT545A/block14_colors.html

List of all colours in R:

www.stat.columbia.edu/~tzheng/files/**Rcolor**.pdf

A (*very*) Brief Intro to R

1) R is a *statistical programming language*
2) Key Terminology:
    a) Objects (= things)
    b) Classes (= type of thing)
    c) Functions (= changes things)
        i) Arguments (= tells a function how to change things)
    d) Package (= add-on)

3) The number sign/hashtag/pound sign = #

    a) Anything after #, the computer ignores!
    b) It is *very* useful in coding, for *commenting* on what your code is doing!

Name the objects, functions, and arguments:

```
> numbers <- c(1, 2, 3, 4)
> class(x = numbers)
[1] "numeric"
> print(x = numbers)
[1] 1 2 3 4
> sum(numbers)
[1] 10
```

How to get help

1) Read documentation.

    ?sum (= ?function)

    help(sum) (= help(function))

2) Google the error message. Use the error message!

3) Stackoverflow!

    Somebody else probably had the exact same problem you did.