Tools of the trade

10 / 09 – 2019

Matthias Stahl

matthias.stahl@ki.se

data → biology

Intro

R

missing values

Show

Science

data science process

tools of the trade

visualization

examples
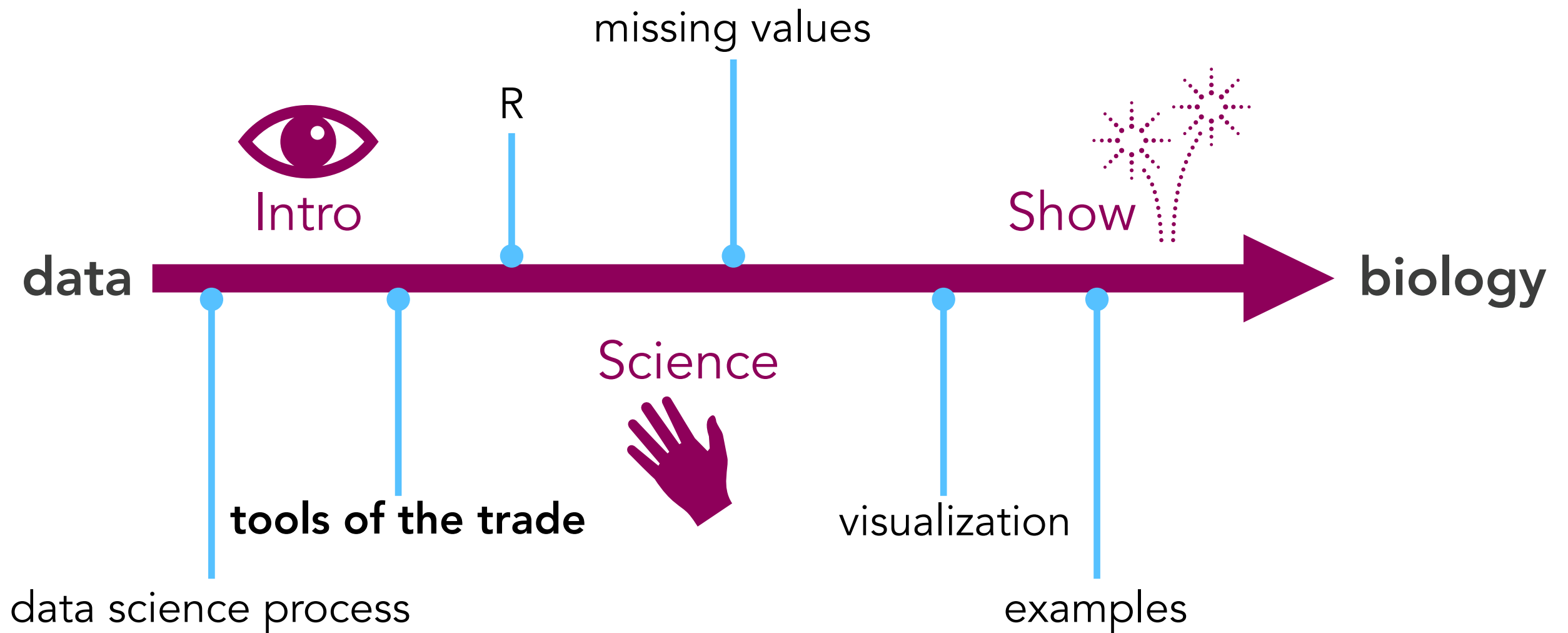
We need tools for each step

Data sourcing

Data value chain

Data acquisition

Data storage

Data analysis

Data visualization
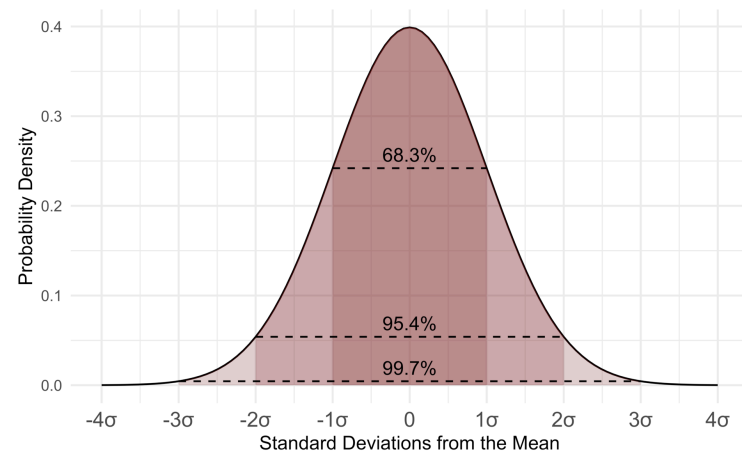
We need tools for each step

Data sourcing

Data acquisition

Data storage

Data analysis
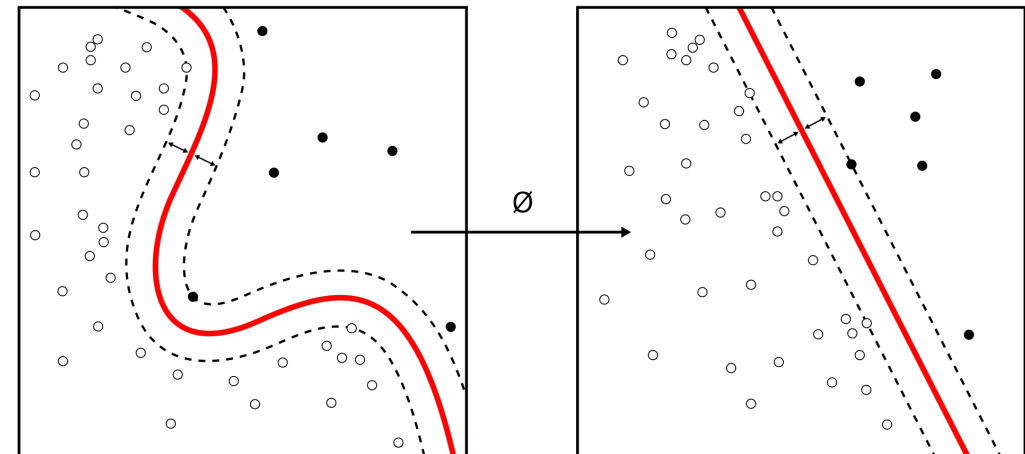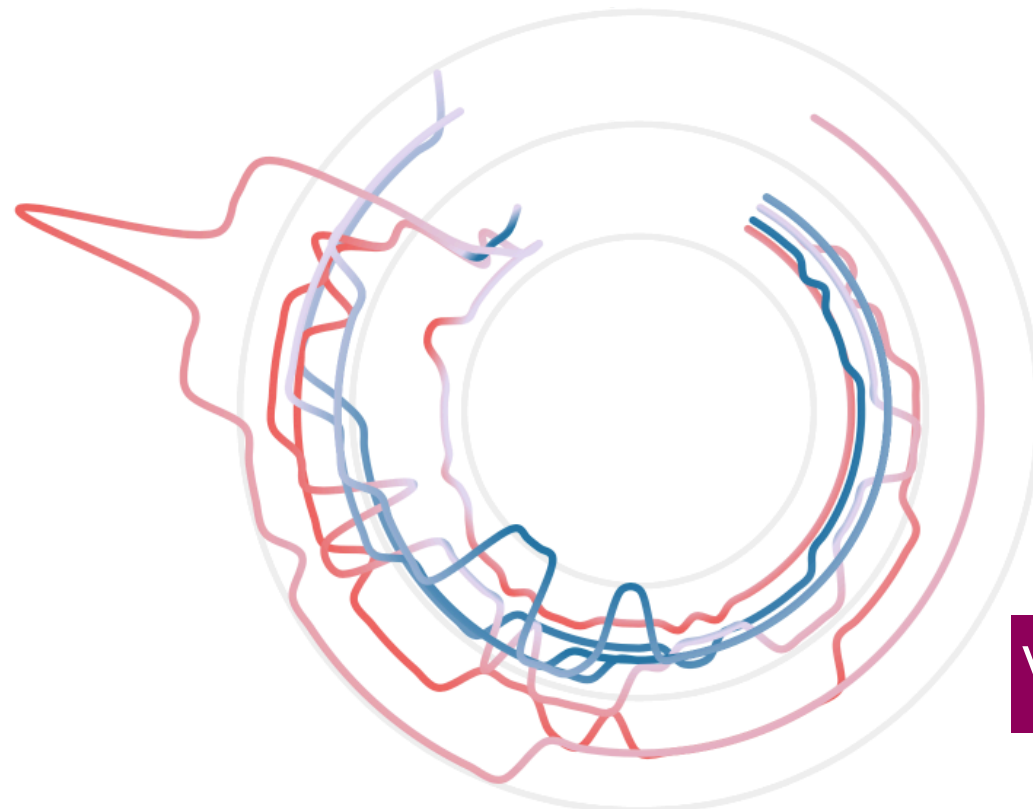
Data visualization

Data value chain

Multiple faces

Statistics

Machine Learning

Visualization

Multiple tools

Perseus = Interface for R

Microsoft Azure
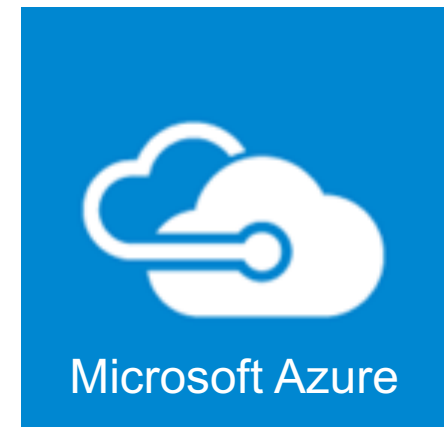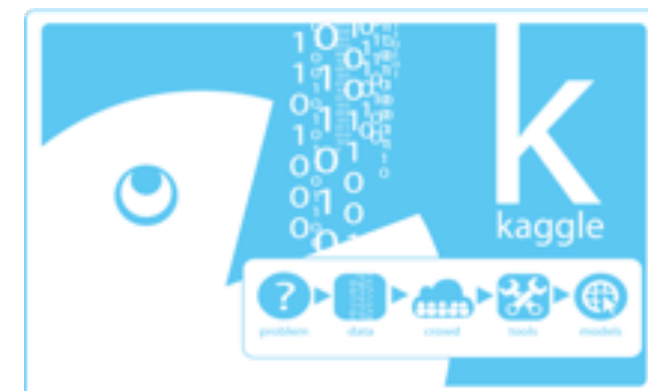
Why *R*
rather than *Perseus*?

Limitations

You find what is allowed to find.

You find *

```
41   # determine protein overlap to cell line panel
42   proteins.meltome <- row.names(meltome.avg)
43   panel <- removeNAsFromESet(eSet = panel.raw, na_ratio = 0)
44   proteins.panel <- fData(panel)$gene_symbol
45   intersection <- intersect(proteins.meltome, proteins.panel)
```

Damn ****, isn't that hard to learn?

How I learned R

1. Find a problem.

2. Divide it in smaller pieces.

3. Google solutions for them or ask seniors.

4. Implement!

That's exactly what we do now!

Quick R crash course

Goals

Try, try, try, …

get a feeling of R's power

preparation to missing value imputation

Please start up RStudio on your machine.

Find all the code here:

https://github.com/higsch/bioinfo-workshop