

let's talk data

www.imsd.ch



"50 Shades of Truth:
Visualisation and Communication of Uncertainty"

Rodolphe Dewarrat & Vera Gärttling, IMSD

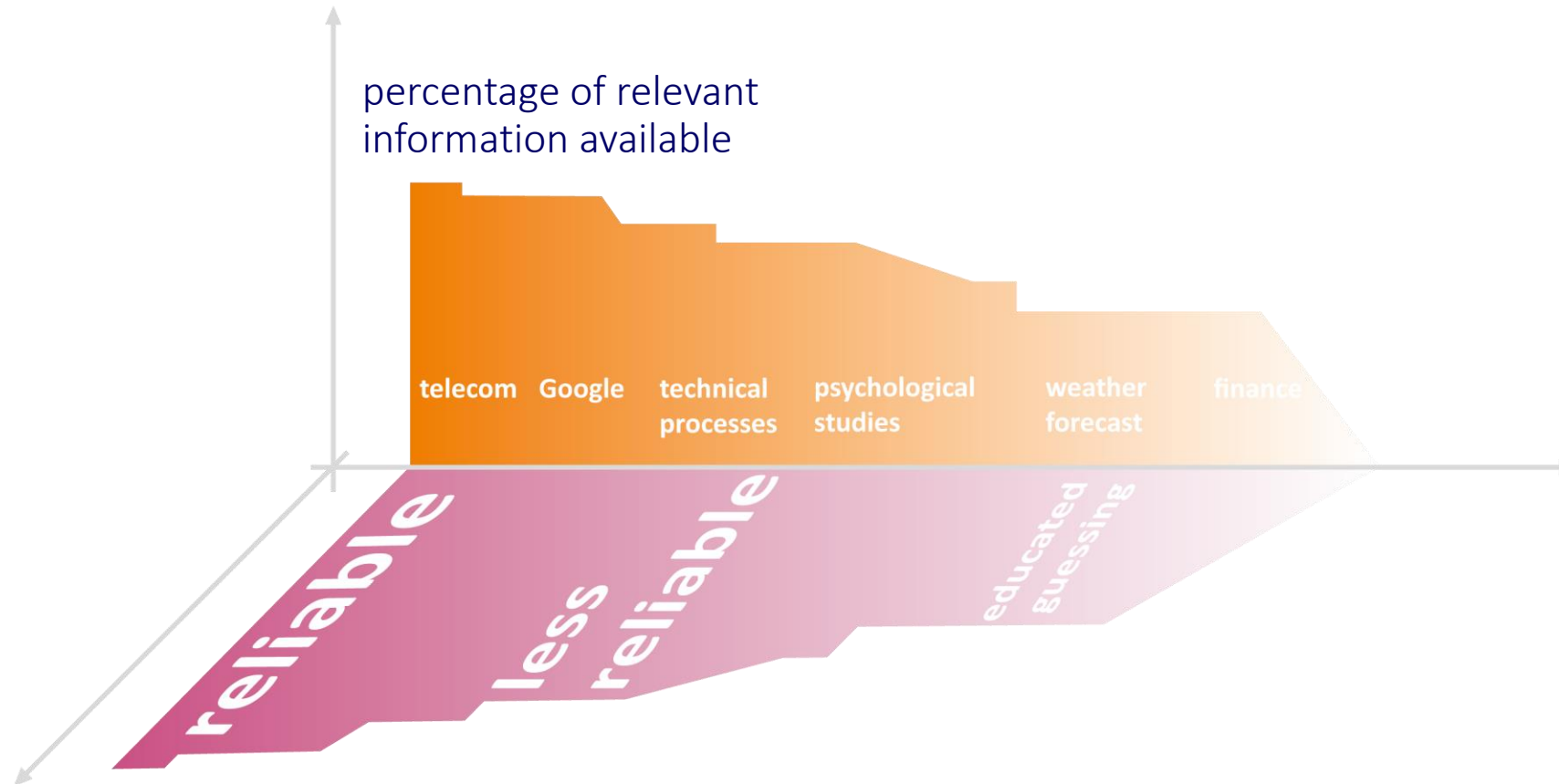
let's talk data

www.imsd.ch



- Issue
- Decision to use analytics
- Data
- Preprocessing
- Algorithms
- Raw results
- Interpretation of results
- Action

Available Data and Truth:



Even the Best Algorithms Available

- don't make the difference between these situations
- don't know what information could be interesting
- are based on violated assumptions

Visit the booth of the
swiss statistical society



The screenshot shows the top portion of a web page from the journal Nature. The header features the 'nature' logo in white on a dark red background, with the tagline 'International weekly journal of science' below it. A navigation bar includes links for Home, News & Comment, Research, Careers & Jobs, Current Issue, Archive, Audio & Video, and For Authors. Below this, a breadcrumb trail indicates the current page is an article within the 'Research Highlights: Social Selection' section of Volume 519, Issue 7541. The article title is 'Psychology journal bans P values', with a subtitle 'Test for reliability of results 'too easy to pass', say editors.' The author is identified as Chris Woolston, and the article is dated 26 February 2015, with a clarification date of 09 March 2015. The main text begins with 'A controversial statistical test has finally met its end, at least in one journal. Earlier this month, the editors of *Basic and Applied Social Psychology (BASP)* announced that the journal would no longer publish papers containing P values because the statistics were too often used to support lower-quality research¹.'

The best analytical method ADF

Presented on the 28th of May
at an expert round table of
statisticians at the ETH

Developed at the University of
Exeter (UK), tested and validated
by international experts

The best analytical method ADF

Understanding 3-way inte x

tomhouslay.com/2014/03/21/understanding-3-way-intera

teetotallers are slow, but tall, heavy teetotallers are really fast.

Remember, these lines are theoretical, and drawn from the regression coefficients rather than the data itself; it's up to you to make sure that short, heavy teetotallers are represented within your data set. I, for one, drink quite a lot of beer.

** My supervisor calls this the 'ADF method', as in 'ANY DAMN FOOL (can see what's happening if you just actually plot your data)'*

*** I originally had beer count, but my esteemed colleague [Lilly Herridge](#) (who knows what she is talking about much more than I do) pointed out that counts are discrete, not continuous, hence*

Follow

Staff profiles - Biosciences: x

biosciences.exeter.ac.uk/staff/index.php?web_id=Alastair_Wilson

UNIVERSITY OF EXETER

Home | Contact us | Staff | Students | MyExeter (Staff) | exehub (Students) | Site map | 中文

Search

Business and community | Working here | Alumni and supporters | Our departments | Visiting us | About us

College of Life and Environmental Sciences > Biosciences > Staff profiles

Biosciences

Dr Alastair Wilson

Associate Professor of Evolutionary Biology

Profile

I am an evolutionary ecologist and research in my group seeks to understand how genetic and environmental processes interact with each other to shape phenotypic evolution. To do this we use ecological with quantitative genetic methods to test evolutionary theory in a wide range of empirical systems both in the lab and in the field. Current projects range from estimating the genetic basis of life history variation in the wild Soay sheep of St Kilda, through to experimental studies of aggression and social dominance in lab populations of fishes.

The common themes running through all my research are those of trying to predict how traits should respond to natural selection, and trying to understand why sometimes they don't seem to respond at all.


I am a member of the [evolution](#) group.

Visit my [research group website](#).

Qualifications

1996 BA hon (Natural sciences), Cambridge University
1998 MSc (Aquatic Resource Management), King's College London
2002 PhD (Zoology), University of Guelph

[Links](#)

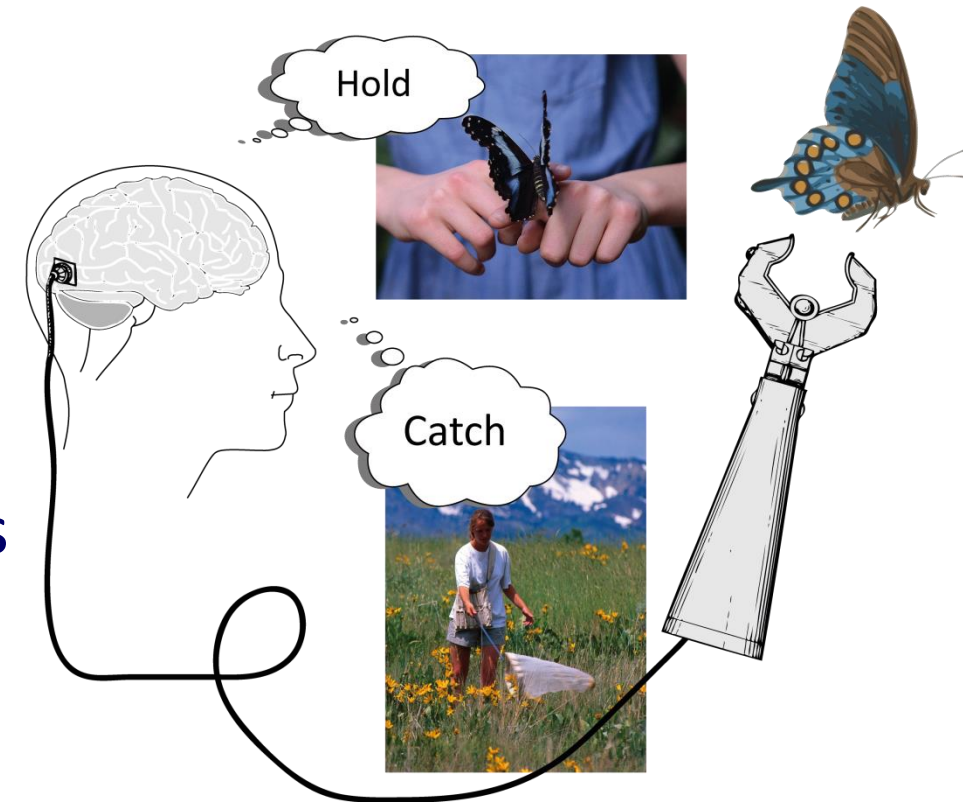


Dr Alastair Wilson

- Profile
- Research
- Supervision/Group
- Publications
- Teaching
- External responsibilities

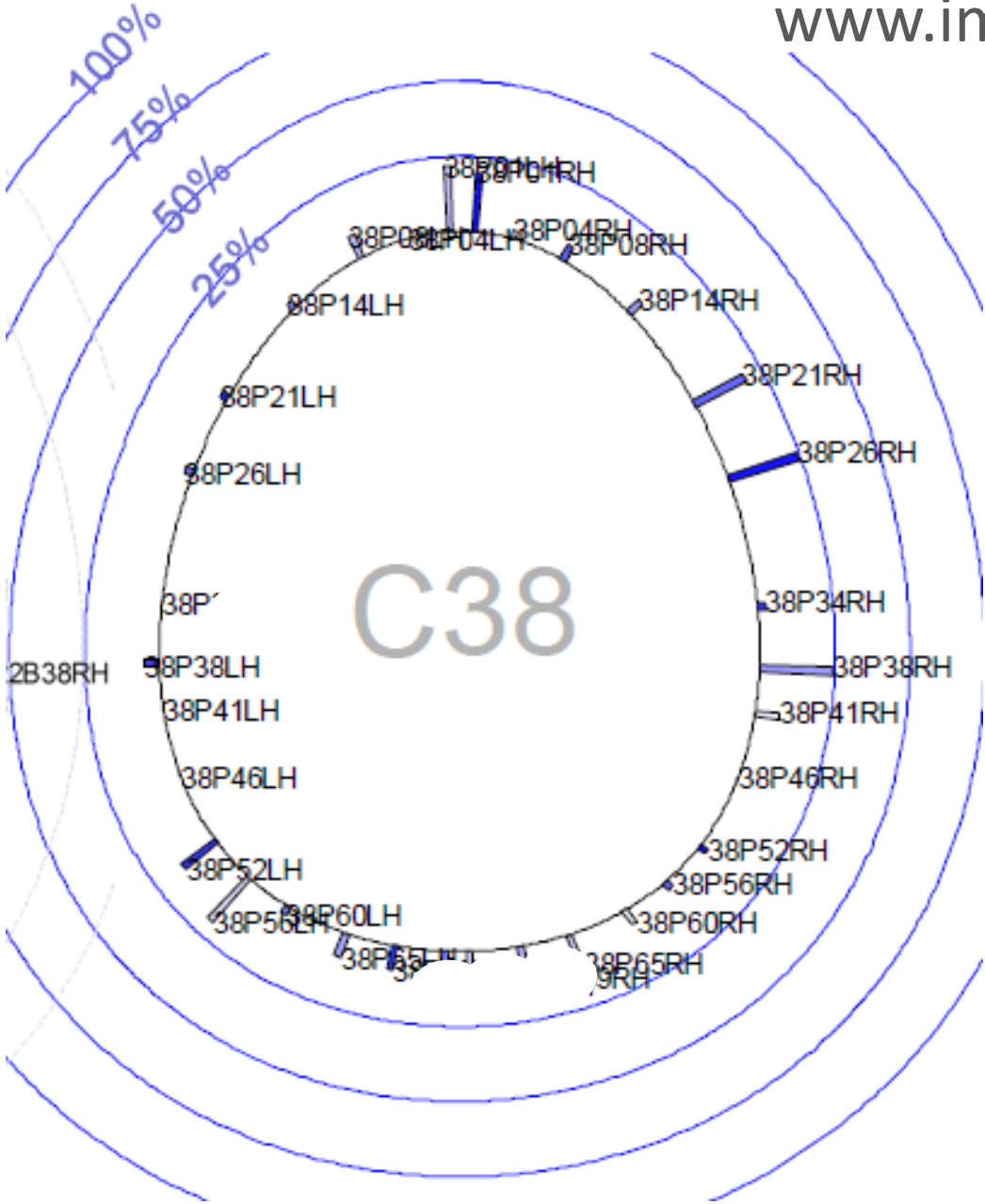
The human visual system

- Today probably our best brain to machine interface
- Trained for millenaries to recognize patterns instantly



Example 1 of 4

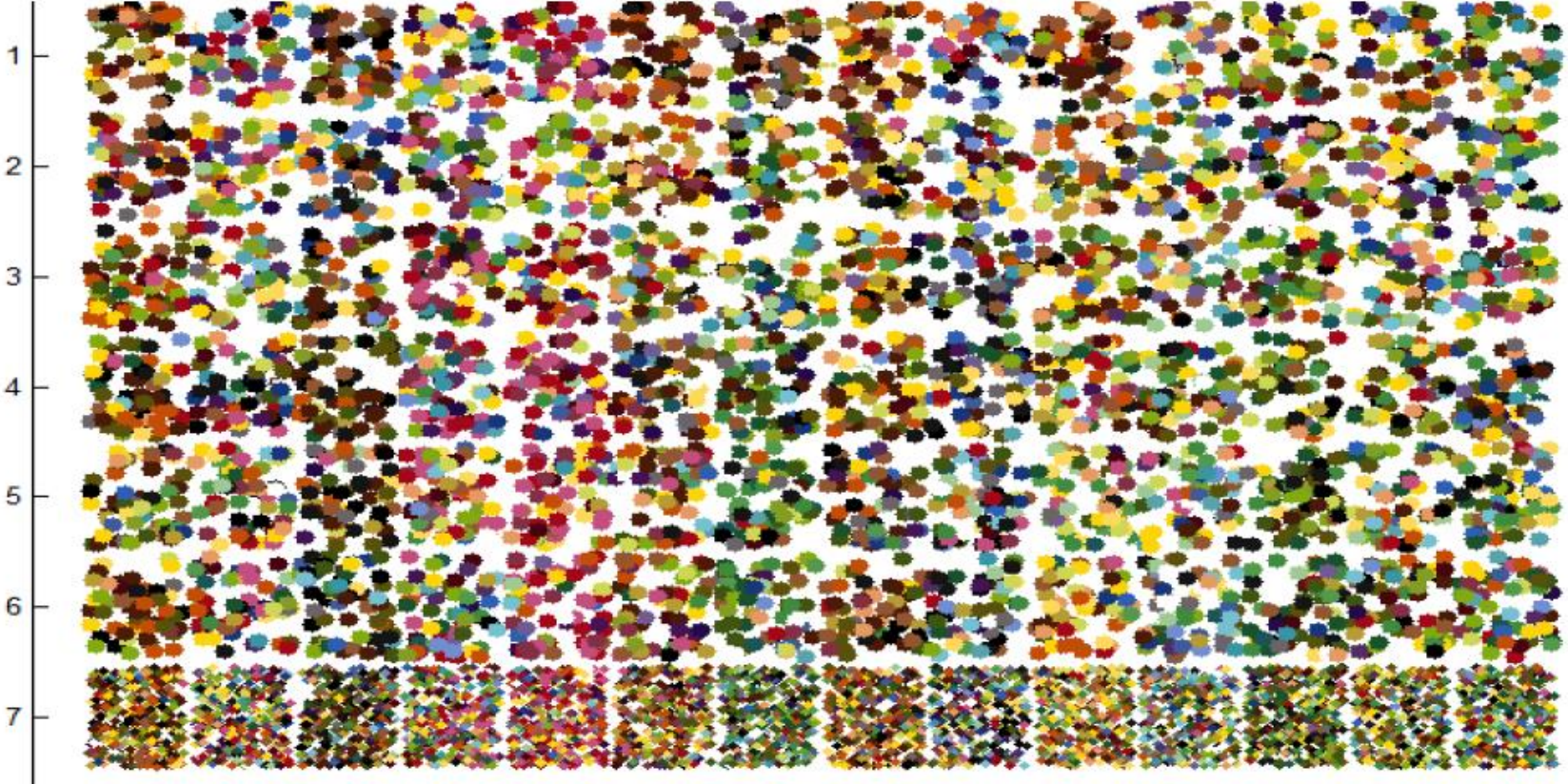
Aircraft section plot



Example 2 of 4

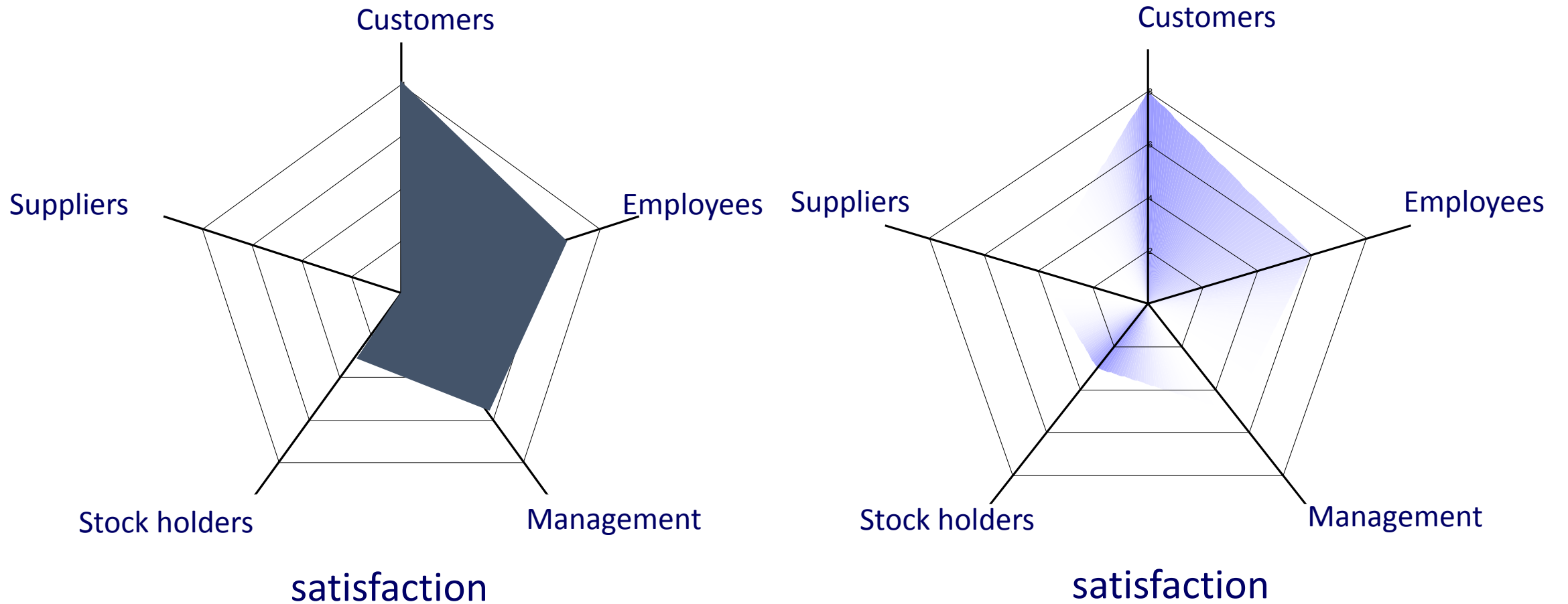


| | burnt | candy | fish | flower | fruity | meat | musty | hazelnut | plastic | rice | soap | vegetable | vinegar | woody |
|---------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 1 Dutch | 58.4320789 | 80.0719684 | 57.7367404 | 81.9177965 | 67.7326035 | 50.5574088 | 60.3787158 | 54.840714 | 71.3119772 | 55.4762246 | NaN | 71.8571947 | 72.602786 | 68.8559632 |
| 2 Dutch-Chinese | 70.5822053 | 81.8492737 | 70.055807 | 75.4783544 | 73.7187105 | 62.862493 | 72.3152526 | 70.0165246 | 77.9557895 | 59.3655965 | 64.4686175 | 78.0999789 | 80.9884228 | 75.5764035 |
| 3 German | 47.9815737 | 72.3784263 | 58.7672509 | 70.537293 | 73.1603947 | 61.1121842 | 61.5286158 | 55.1184982 | 71.4308667 | 54.5474526 | 55.6190158 | 64.3347351 | 68.7651789 | 68.4171175 |
| 4 Malaysian | 63.4237702 | 73.0827018 | 42.4107105 | 72.8734965 | 64.3420053 | 59.6748386 | 65.3073474 | 55.9291719 | 67.324307 | 64.7403912 | 64.3707684 | 53.9171193 | 65.3446228 | 68.1135 |
| 5 Malaysian-Chinese | 60.083907 | 77.7489719 | 58.9153211 | 70.8780316 | 71.2605632 | 60.6907281 | 58.9403281 | 65.3705579 | 68.0639579 | 55.2268947 | 67.3090772 | 53.4215754 | 64.9527298 | 65.2645825 |
| 6 USA | 55.5846614 | 80.8872263 | 62.7684211 | 85.4085228 | 63.4449842 | 70.4763596 | 59.514214 | 43.8068386 | 68.3753281 | 50.6796281 | 52.1157614 | 56.6332316 | 72.7129035 | 72.4103842 |
| 7 OVERALL | 50.9448838 | 81.2669547 | 82.2533489 | 96.8136448 | 65.8193551 | 47.6442567 | 60.0024763 | 44.3445518 | 62.0755148 | 62.1733501 | NaN | 56.2239787 | 71.1094301 | 76.5334009 |



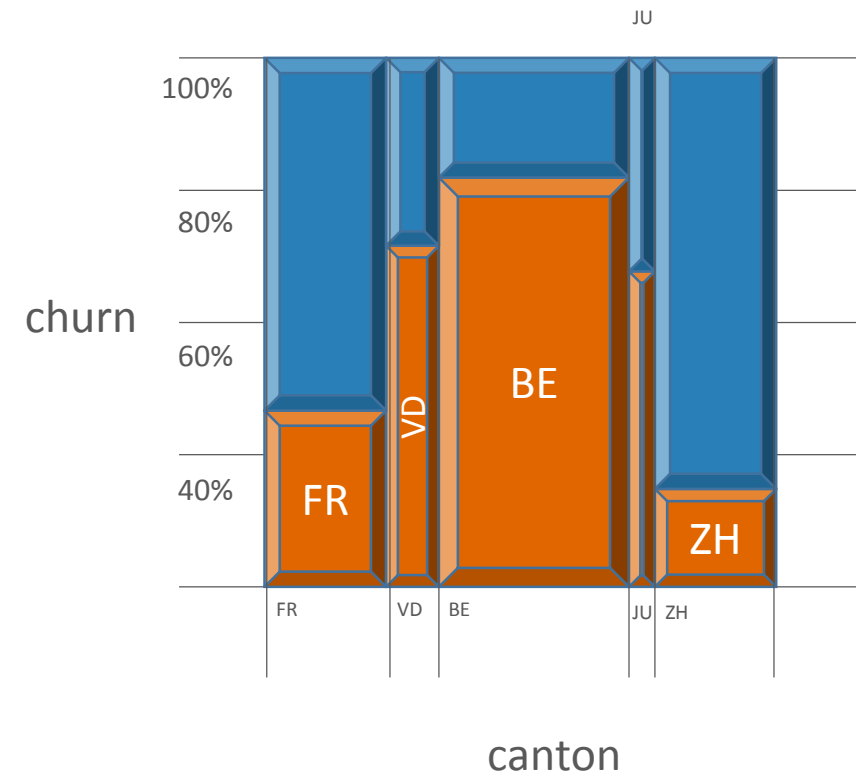
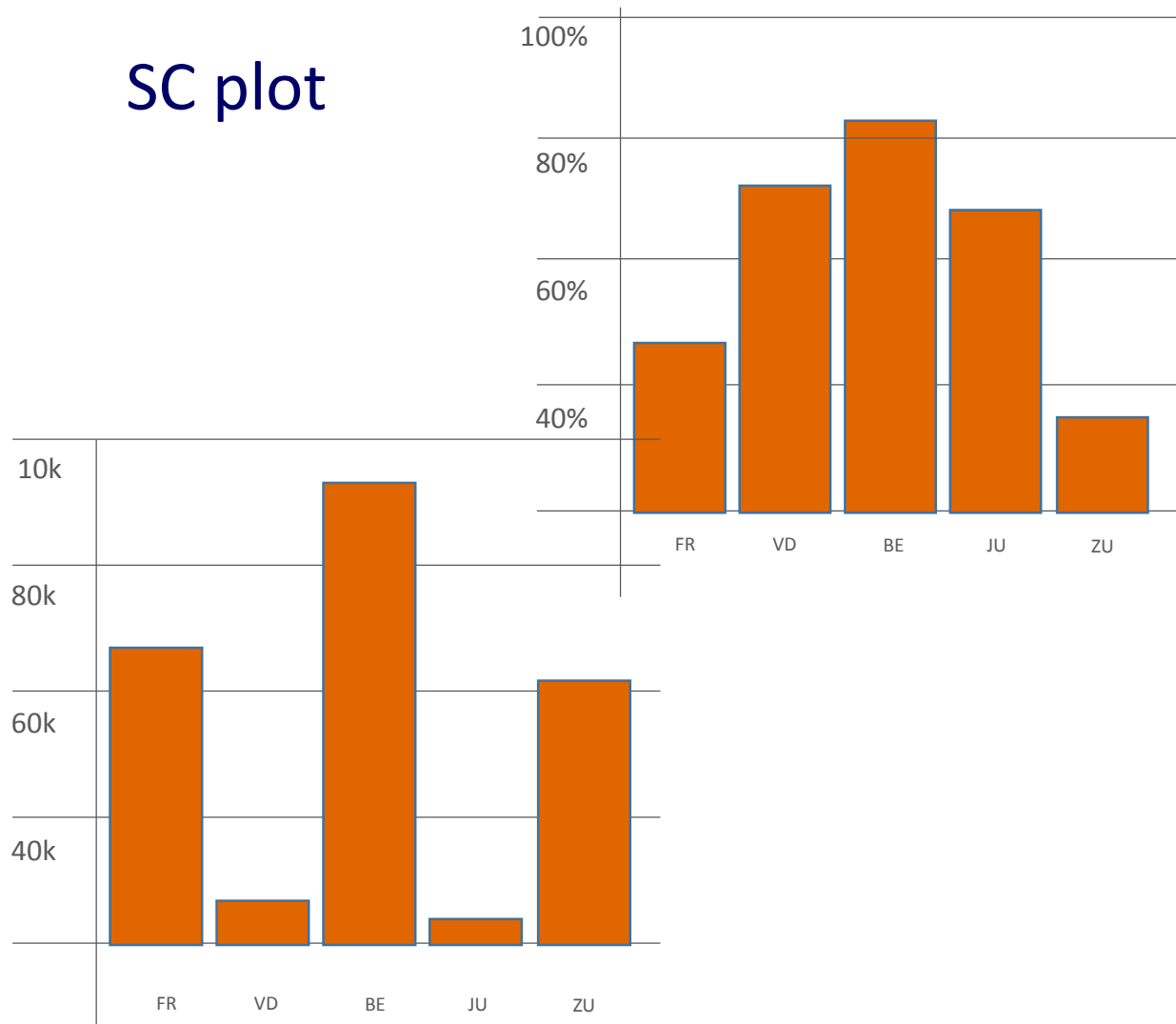
Example 3 of 4

Uncertain radar plot



Example 4 of 4

SC plot



1cm² = 10k customers

Conclusion

- Before starting anything (e.g. buying a data cluster, doing an analysis) look at your data and consider its validity and truth.
- This step cannot be replaced by a machine.
- Strongly consider working with partners and tools (software, team) that allow you a high flexibility in your data representation.

let's talk data



Thank you for your attention
contact: rodolphe.dewarrat@imsd.ch
twitter: [@imsd_analytics](https://twitter.com/imsd_analytics)