



# Small Data Science

Microconf EU 2015

Marc Hoffmann  
[mh@h-rd.org](mailto:mh@h-rd.org)

# Small Data Science

How to extract **actionable knowledge** from less or no data?



# Small Data Science

How to leverage:

- your **HYPOTHESIS**,
- some data, and
- basic psychology





# Small Data Science

A hypothesis is a **testable** assumption.

Hyp 1: Shopify shops **need** adwords help.

→ **difficult** to test, big data.

Hyp 2: Shopify shops **use** adwords help.

→ **easier** to test, see shopify marketplace!



# Why Small Data Science?



**Small amount of data**  
high uncertainty  
many alternatives

**Big data**  
low uncertainty  
few alternatives

**Make it simple**

**Make it complex**



# Small Data Science

Use all background knowledge you can get.

“Solve Problems where People are already paying for solutions” (Dan Norris)

➡ **BUT offer different solution!**

Example: Wpcurve.com vs. agency

# Small Data Science

gmail	free \$
mailchimp	cheap \$
contsant contact	\$
salesforce	\$\$\$\$\$
<b>followup.cc</b>	\$



# Small Data Science

## Reason Backwards / Pattern Matching

ex. appointment reminder, Patrick McKenzie, 2014

MRR ~ \$6500/month, 142 accounts, ...

~ 140 new customers in 2014

~ 0.7% conversion rate

→ bid up to \$ 1.8 per adwords click (J. Cohen)



calculated

Can you build a business based on these numbers ?



# Small Data Science

## Use the “right” statistics

Median

for “small” samples

What’s most frequent?

- 1) Type of customer
- 2) Contract size
- 3) Length of contract

Mean

for “large” samples

What’s the “average”?

- 1) average \$\$
- 2) average age

Depends on data distribution

Fluctuates with small samples



# Small Data Science

## Heuristics

**Find most important reason and ignore the rest**





# Small Data Science

A photograph of two brown goats with yellow ear tags locking horns in a fight. The goats are positioned in the center of the frame, facing each other with their heads lowered and horns interlocked. The background is a dry, dusty ground. The image is used as a background for the text overlay.

## Small Data Thinking

- 1) Bet against yourself – Would I bet \$\$ for/against?
- 2) Use counts to estimate unknowns, not percentages.



# Small Data Science



Slides, reading list at:

<http://h-rd.org/mceu2015>

Questions? I like to help, just ask the guy  
on the photo!

**mh@h-rd.org**