

Social Data Science

Department of Economics, Faculty of Social Sciences
University of Copenhagen

Summer 2016

Lecturers

Sebastian Barfort

David Dreyer Lassen

Occasional teaching assistant

Snorre Ralund

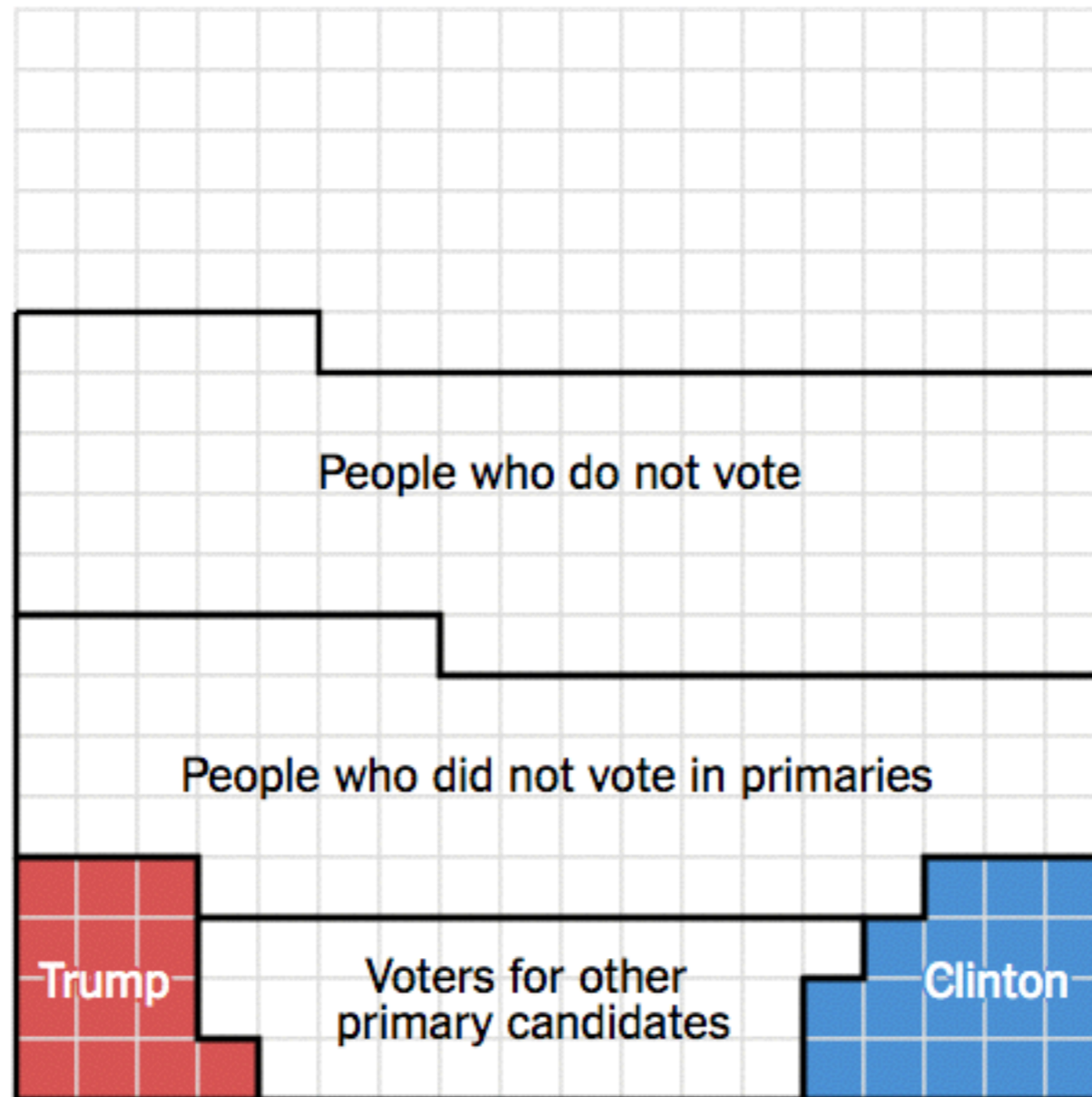
Welcome!

always bring computer!

http://sebastianbarfort.github.io/sds_summer/
+ Absalon homepage

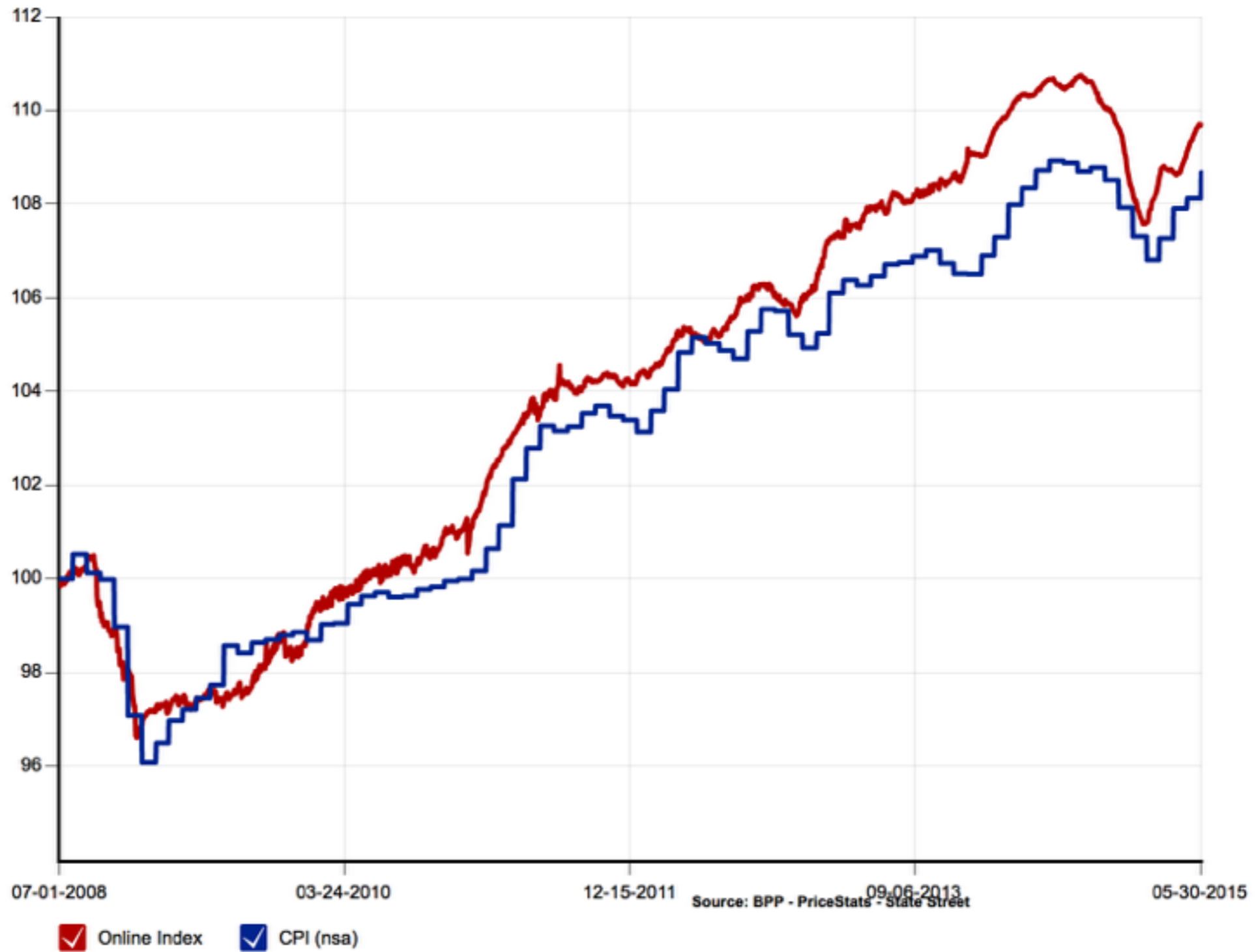
Today

1. New course: Why and (so) What?
2. Who are we — and you?
3. Logistics and Plumbing
R, Absalon vs. Github, groups, assignments??, exam project, course evaluation, Q&As
4. Course culture and ethics
5. Reading list and Lecture plan
6. Groups and details
7. Computer stuff



NY Times August 1,

<http://www.nytimes.com/interactive/2016/08/01/us/elections/nine-percent-of-america-selected-trump-and-clinton.html>



US Inflation

The Billion Prices Index: <http://bpp.mit.edu>



BIG DATA

IN
ACTION
FOR
DEVELOPMENT

 **THE WORLD BANK**
IBRD • IDA | WORLD BANK GROUP
Latin America & the Caribbean
Opportunities for All

 **WORLD BANK GROUP**
Macroeconomics & Fiscal Management

SECONDMUSE 

New course I

- Background: Why Social Data Science
 - Big Data / Deep Data / New Data
 - Social Fabric / Taking Data Science Back



● "data science"
Søgeterm

● "big data"
Søgeterm

+ Tilføj sammenligning

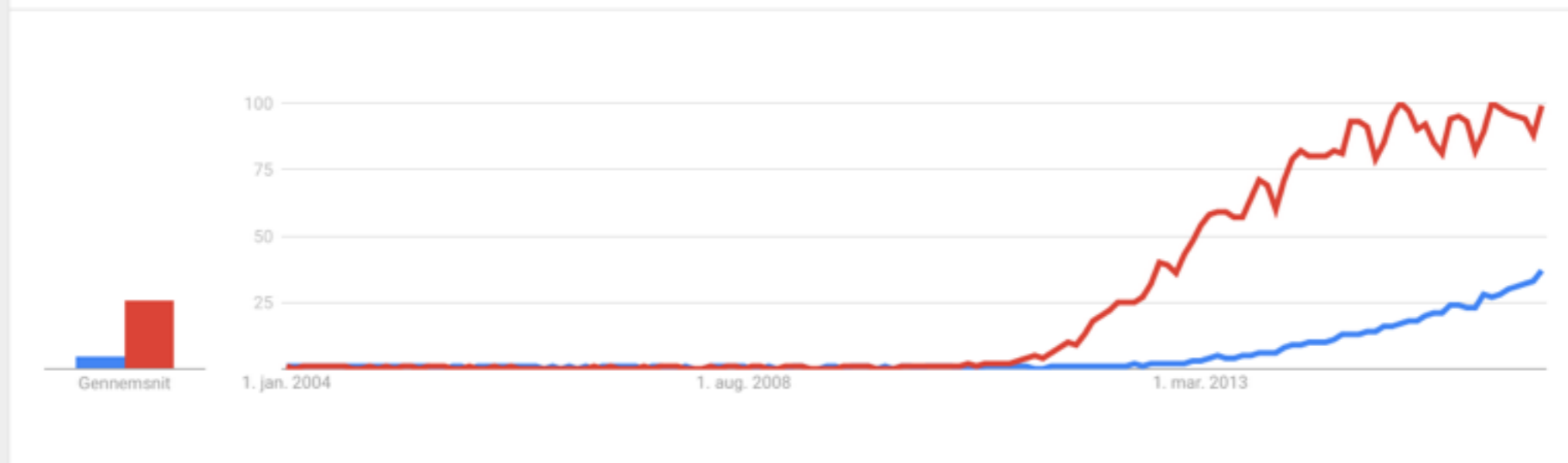
Hele verden ▾

2004 - i dag ▾

Alle kategorier ▾

Websøgning ▾

Interesse over tid ?





Google Trends

Sammenlign



"data science"

Søgeterm



"big data"

Søgeterm



trump

Søgeterm



Tilføj sammenligning

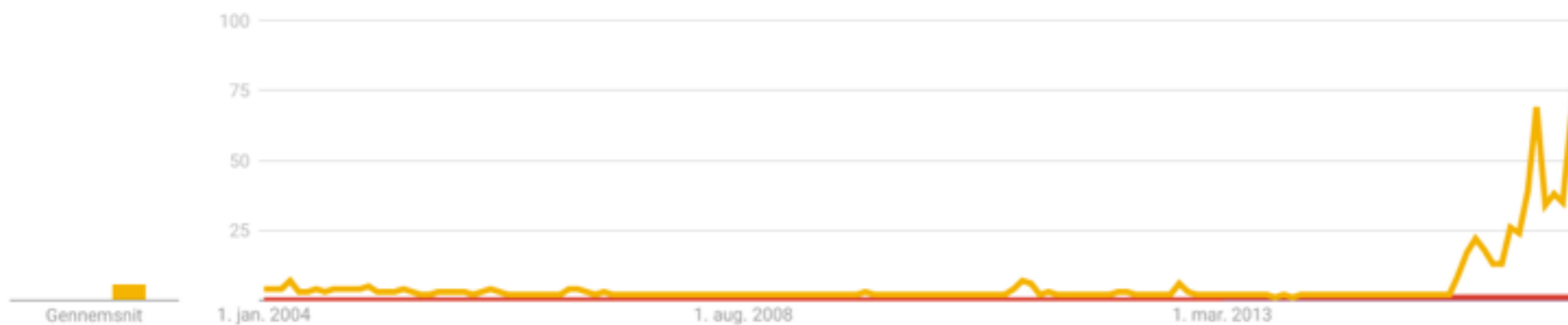
Hele verden

2004 - i dag

Alle kategorier

Websøgning

Interesse over tid



New course I

- Background: Why Social Data Science
 - Big Data / Deep Data / New Data
 - Social Fabric / Taking Data Science Back
- Econ: Not Econometrics, not standard Methods
- Social science methods: data collection, data construction
 - Sociology, political science, anthropology
 - Why important: research/substantive decisions taken along the way

New course II

- Internet/digital data allows for more/new/realtime data: consumer prices, Uber, Facebook
- Goal: ability to construct new data aimed at answering social science questions
- Prerequisites: Interest, willingness to program, will refer to regression models occasionally
- New course - hard to know time table

Some topics

We will present a social science view on data science methods needed for **collecting** and **analyzing real-world data**. Focus points: **generating new data** (collecting, scraping, working with APIs), **data manipulation tools** (transforming, cleaning), **visualization tools** (visualizing raw data and model results), **reproducibility tools** (git, github), an introduction to statistical techniques for predicting and classification, known as **statistical learning**

Meta and non-meta: What is data, types of data & types of questions, ethics, privacy, costs and benefits of data driven research / big data

Who are we?

- We are: Sebastian, David and Snorre
- Part of sodas.ku.dk
Copenhagen Centre for Social Data Science

Who are you?

- 7 Q survey NOW!

<https://sebastianbarfort.typeform.com/to/Rs3G3f>

(link at github, too)

Logistics and Plumbing I

- We meet every day: lectures in the morning, lab/exercises in the afternoon
- Bring computer - R!
- Absalon vs. Github

Logistics and Plumbing II

- Groups - we decide, more shortly
- Assignments to help you through the material
- Week three: Exam project (see Github post)
- Course evaluation - formal and informal
- Discussion forum - Absalon

Course culture and ethics

- Philosophy: Open source, everyone contributes
- Help each other: within groups, across groups
 - Discussion forum
- But don't free ride :-) Only fun if y'all pitch in
- Share, but don't copy (really, don't)

Reading list / Lecture plan

- Reading list at Github
 - New and fast moving topic - no textbook
 - Chapters (at Absalon), links to papers, blogs (UCPH domain)
- Lecture plan
 - Adjustable

Contact points

- For most questions, try
 - your group
 - Other groups / Absalon discussion forum
 - in-class
- In rare cases: email
- Don't call us (and we won't call you)

The next couple of weeks