The University Of Sheffield.

Q-Step

A step-change in quantitative social science skills

Funded by the Nuffield Foundation, ESRC and HEFCE

Sheffield Methods Institute.

Learning R is difficult… but worth the investment!

Dr Todd Hartman
Senior Lecturer, Quantitative Social Science
Director, Sheffield Q-Step Centre
http://tkhartman.staff.shef.ac.uk/

Brief Biography

Why should you listen to me?

Sheffield Methods Institute.

- Teaching statistics 10+ years (yikes!)
  - (Mostly) introductory statistics / quantitative data analysis
  - UG and PG levels
  - To students with range of backgrounds but all R ‘newbies’ (e.g., ‘believers’ vs ‘skeptics’)
  - Made the switch from Stata to R 5 years ago

WELCOME
TO STATISTICS

Why We Teach Our Students R?

Sheffield Methods Institute.

- Top 5 Reasons
  1. Powerful, free software with large and helpful user base (e.g., Stack Overflow)
  2. Encourages creative problem solving
  3. Fosters transparency and the importance of replication
  4. Programming is a marketable skill / heuristic (data science / analysis jobs = R & MySQL)
  5. We use it to do our own research / consultancy (we're the converted)

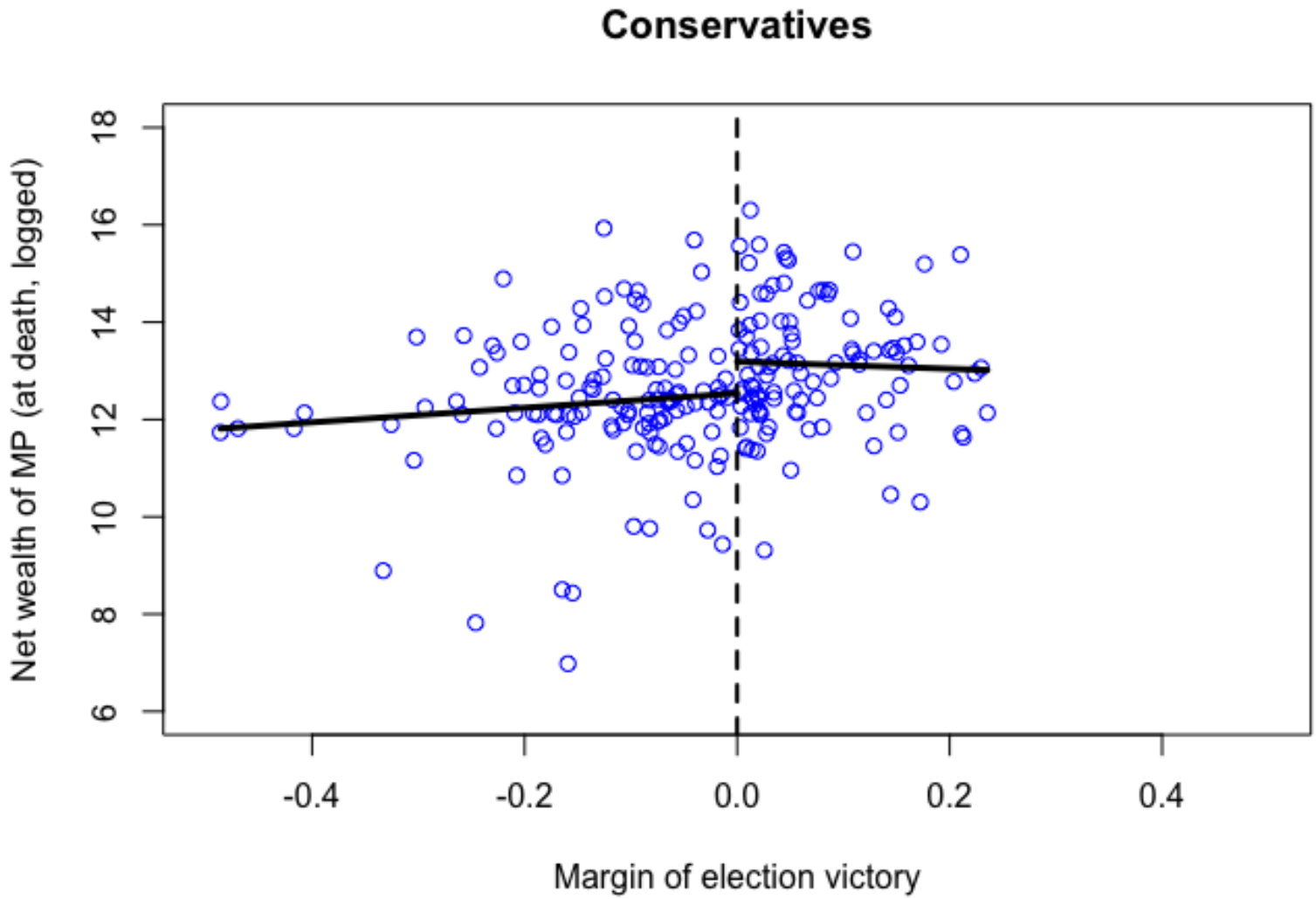
How We Teach R?

Key Concepts: Use Real Data

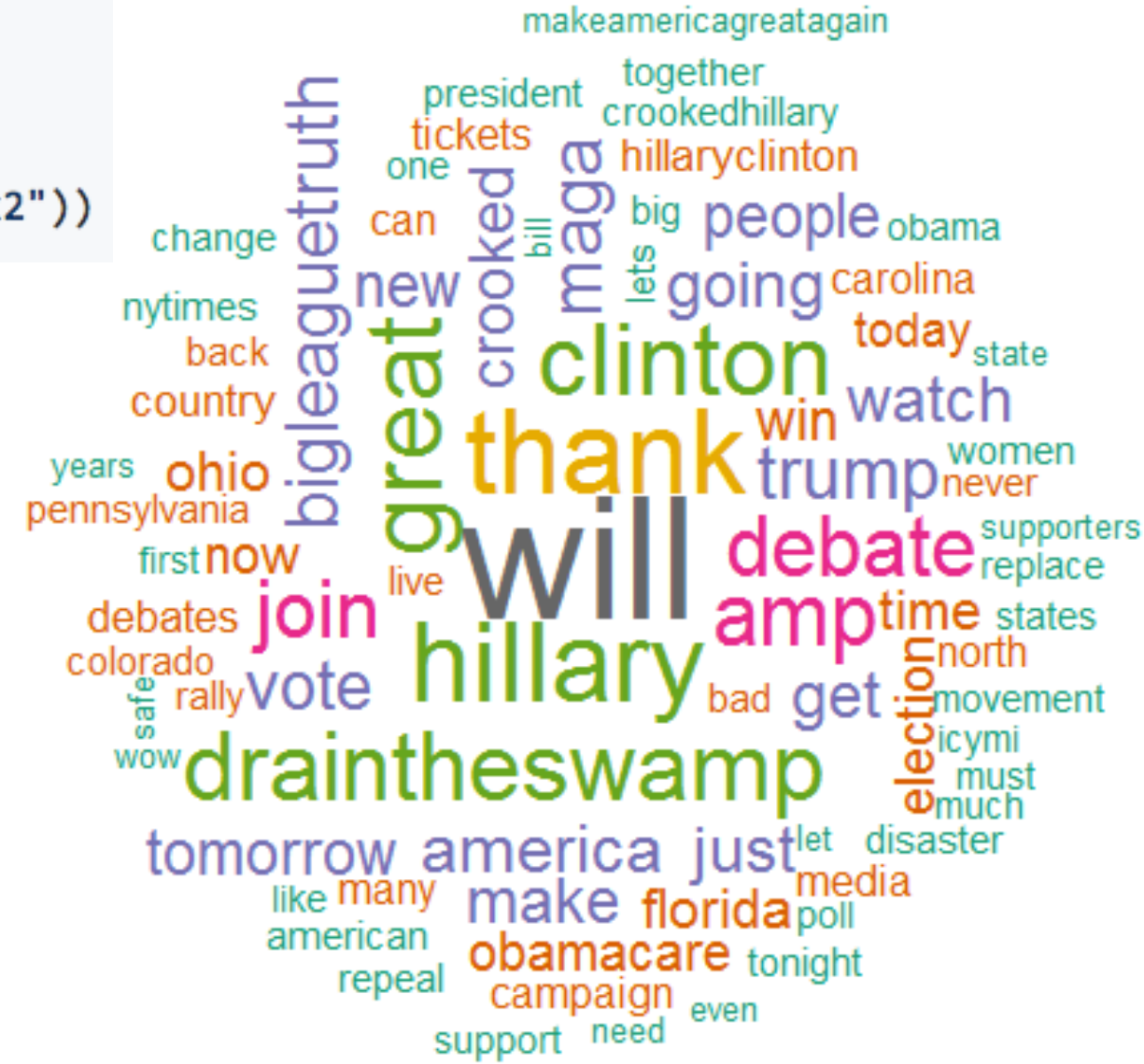
Sheffield Methods Institute.

- Use real data from the social sciences
  - Toy datasets = false sense of security
  - Students need to understand that 90% of data analysis is data cleaning, 10% analysis
  - And like to think that they're doing something important
    - e.g., discovering real relationships in the world
  - Don't reuse the same data over and over
  - Show them where to find it

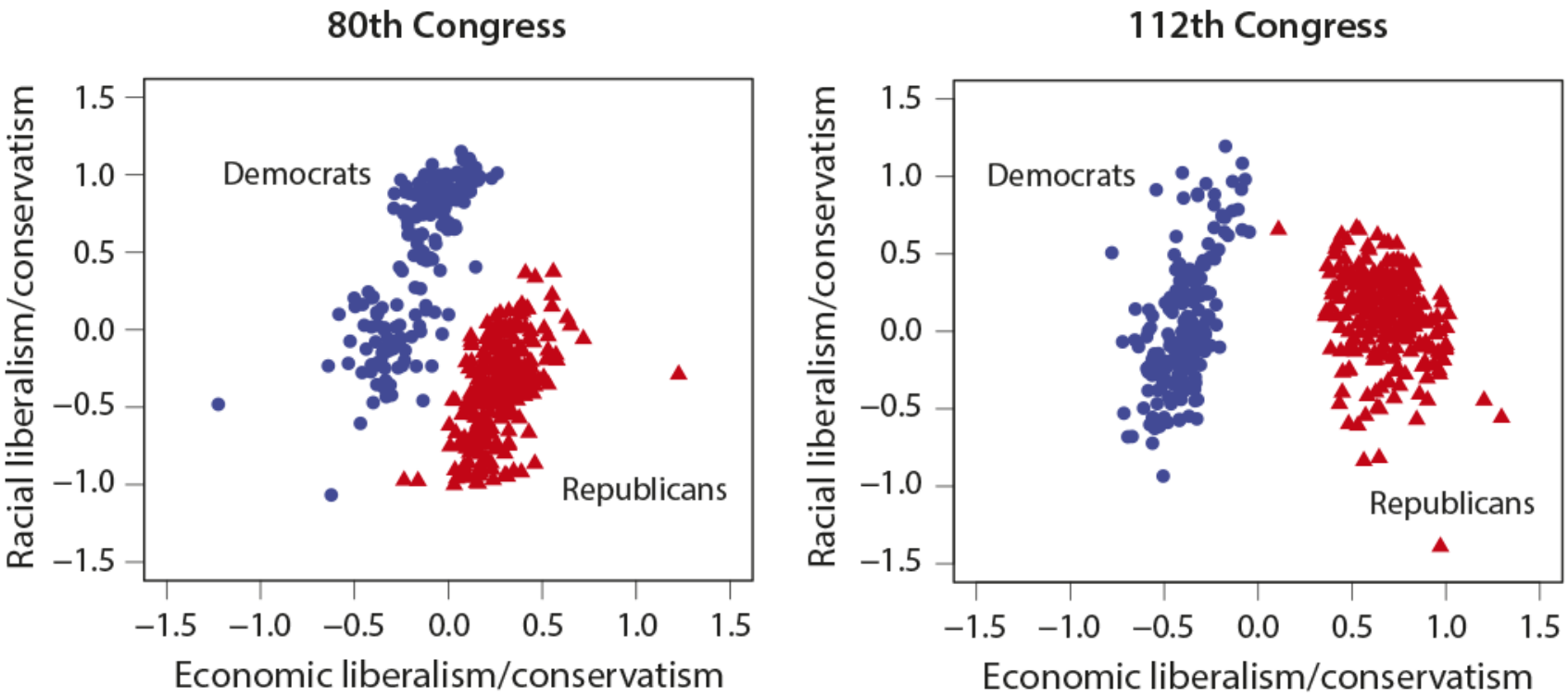
Regression


Conservatives
Net wealth of MP (at death, logged)
6
8
10
12
14
16
18
-0.4
-0.2
0.0
0.2
0.4
Margin of election victory

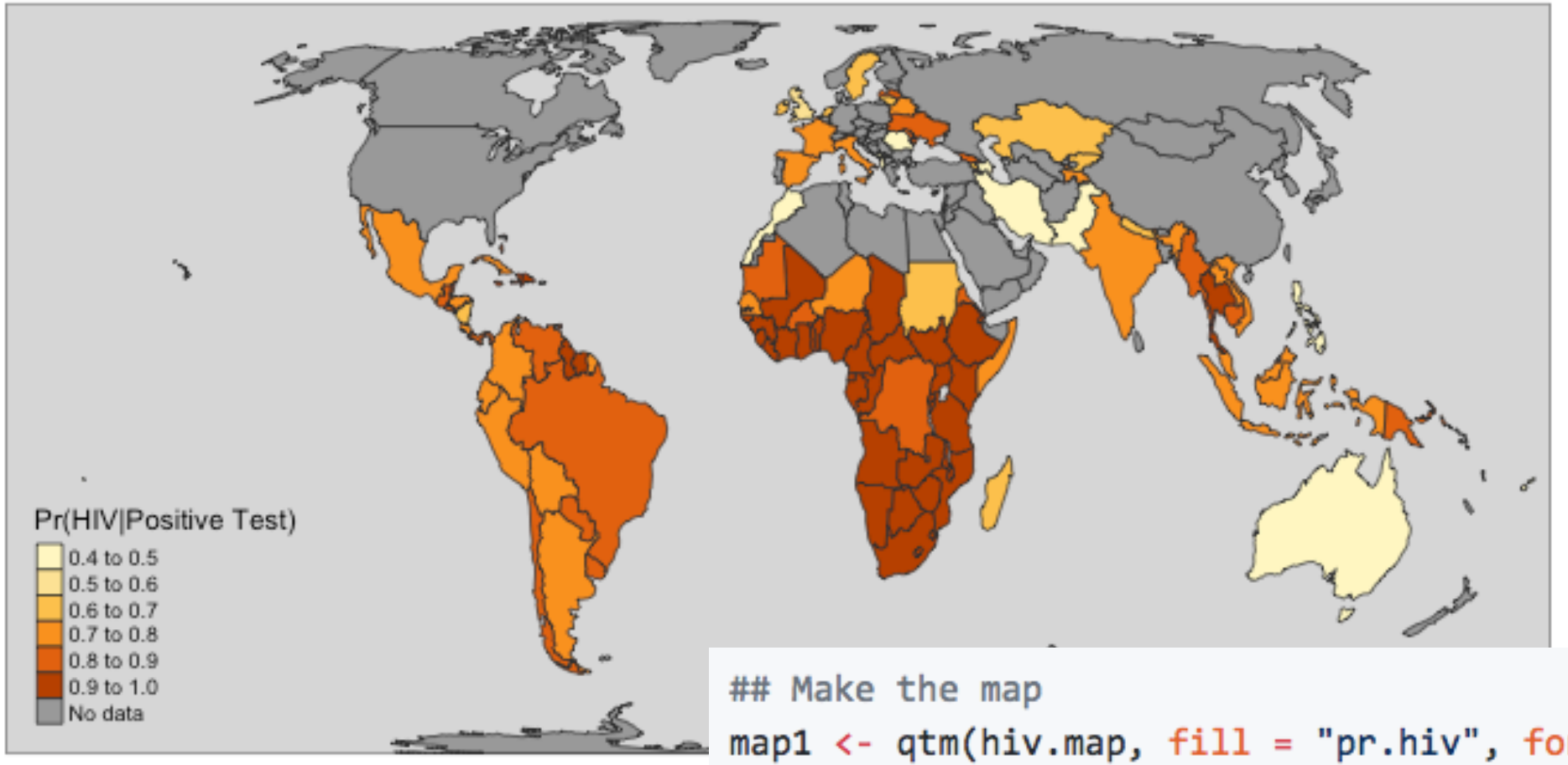
```
wordcloud(words = names(word.freq),
          freq = word.freq,
          min.freq = 10,
          random.order = FALSE,
          colors=brewer.pal(8, "Dark2"))
```



Cluster Analysis

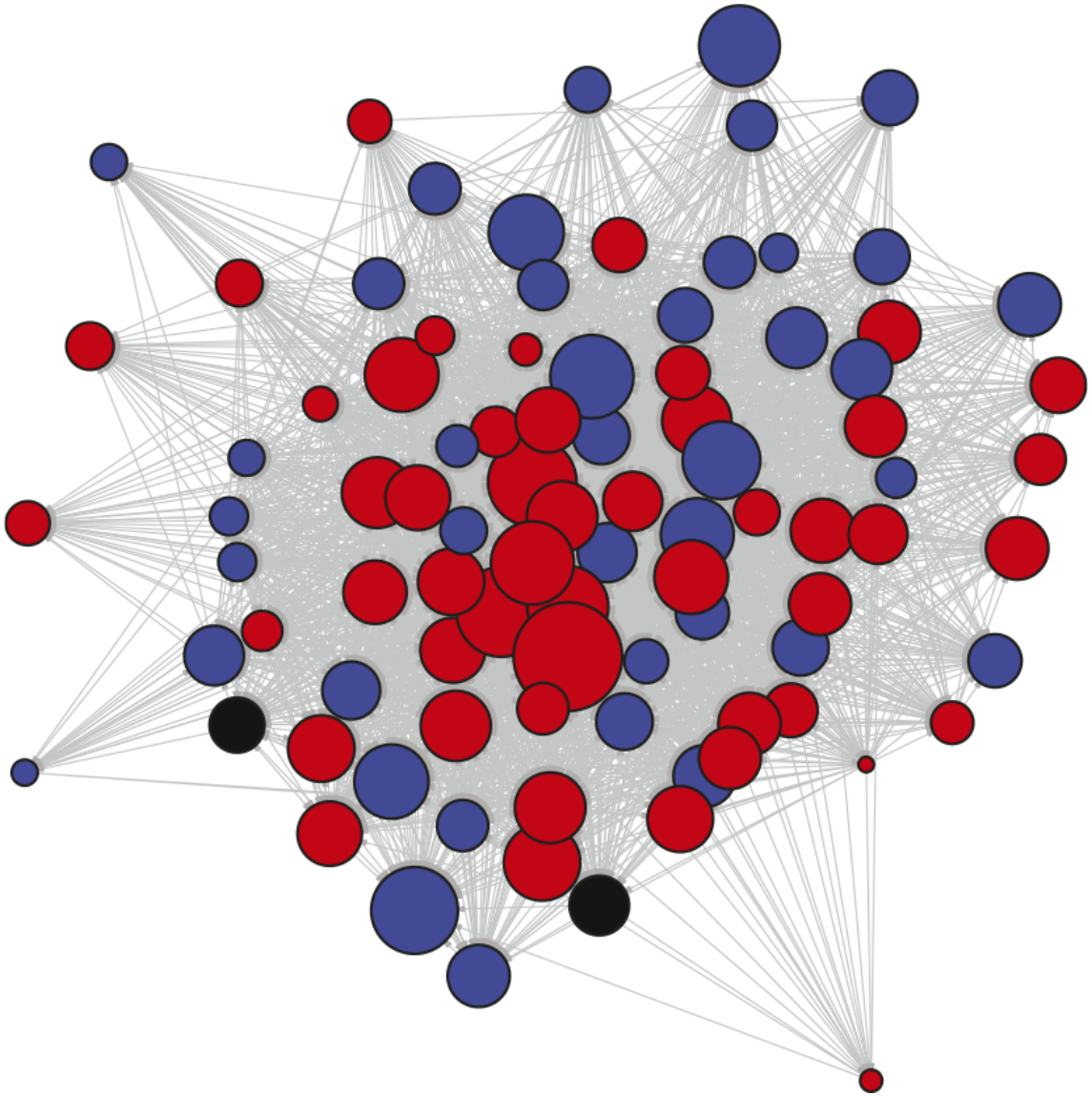


Global Probabilities of HIV Given a Positive Test

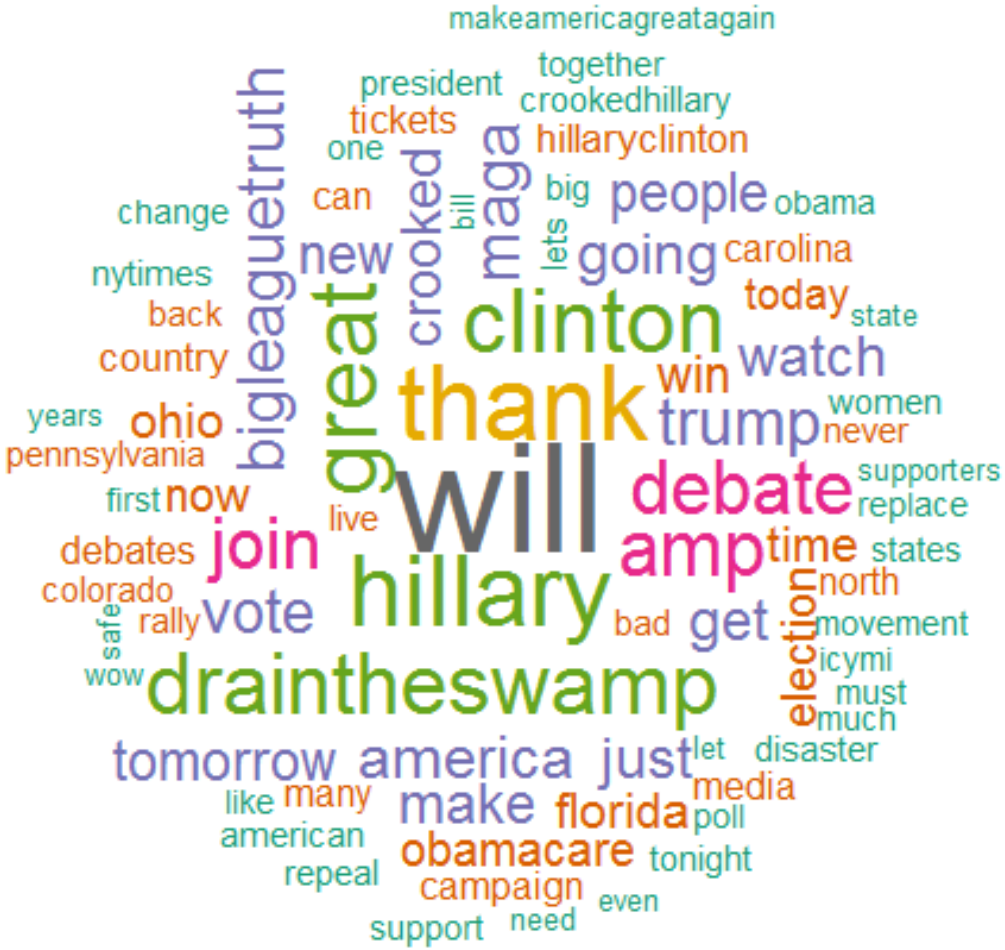


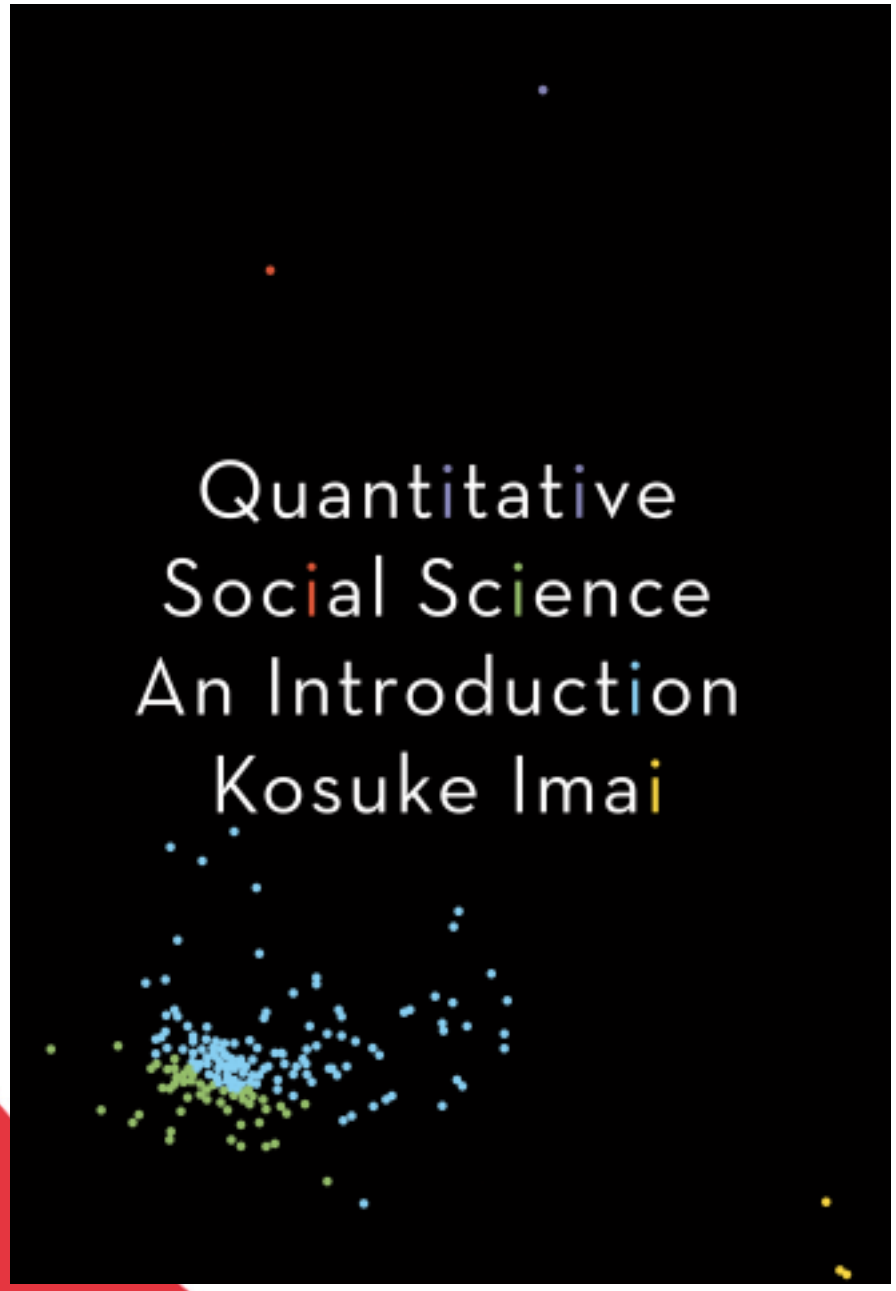

```
## Make the map
map1 <- qtm(hiv.map, fill = "pr.hiv", format = "World",
            style = "gray", text.root = 5,
            fill.title = "Pr(HIV|Positive Test)",
            fill.textNA = "No data")
map1
```

Social Network Analysis



Text Analysis





Sheffield Methods Institute.

- Find a non-traditional ‘statistics’ textbook
- Pros: uses real data from published social science research; lots of example code
- Cons: focuses on breadth over depth; for more advanced users (exercises are quite difficult)