Statistical consulting
My experiences

Karl Broman
Biostatistics & Medical Informatics
University of Wisconsin–Madison

biostat.wisc.edu/~kbroman
My situation

- Applied statistician at a university
  - No formal consulting role
  - Do my best to answer colleagues' questions
  - Can always say no

- Almost no experience with industry

- Almost no experience with consulting companies
Consulting vs. collaboration

- **Consulting**: short-term
  - Try to answer a specific question within a week or two

- **Collaboration**: long-term
  - Get to know the scientist's grander schemes
My experience

- Consulting class in graduate school at UC-Berkeley
  - Three times, with Andrew Gelman, Terry Speed, and David Freedman

- Postdoc with a geneticist
  - Relatively isolated from other statisticians

- Postdoc to faculty member
  - Steady stream of consulting experiences
  - From courses taught, referred by dept chair, other friends, software/papers
Why do it?

► Learn more biology
► See new data
► Think about new problems
  – And it's nice to get an easy one, sometimes
► Be useful (help people)
► Publications
► Leads to long-term collaboration
  – Find out if you like the person
► Could lead to new statistical methods research
  – But that's more often from longer-term collaborations
The initial meeting

▶ Listen

▶ Verify that you're following

▶ Don't be embarrassed to admit ignorance

▶ Focus particularly on:
  – Scientific questions
  – Form of the data
  – Where the data came from

▶ I'll usually delay giving specific advice to a second session
Key difficulties

- Admitting ignorance
- Getting the scientist to back way up
- "I have a quick question…"
- You don't have time
- You see a big problem
- Co-authorship on publications
- They don't want anything fancy, and you think it's needed
How to say no

► Can you find someone to take over?
► Give at least a rough guide
► My two-week rule
Disadvantages

► You don't have time to do things properly
► You don't know what you're getting into
► You don't see long-term results
► I hate power/sample size calculations
Be self-sufficient

- Learn multiple programming languages
  - R
  - C/C++
  - Perl/Ruby/Python

- "How would you like to receive the data?"
  - "In its current form."
Time management

► Don't let short-term things crowd out your long-term projects

► Save big blocks of time for yourself.
  - I reserve a full day each week for my own work
Summary

▶ Statistical consulting is fabulous
  – Lots of new problems
  – Learn a lot of science
  – Help people

▶ Don't need to know the answer
  – You just need to connect the scientist to the answer