

# **Getting Help**

Computing for Data Analysis

Roger Peng, Associate Professor Johns Hopkins Bloomberg School of Public Health

#### **Asking Questions**

- · Asking questions via email is different from asking questions in person
- People o nthe other side do not have the background information you have
  - they also don't know you personally (usually)
- · Other people are busy; their time is limited
- The instructor (me) is here to help in all circumstances but may not be able to answer all questions!

#### **Finding Answers**

- · Try to find an answer by searching the archives of the forum you plan to post to.
- Try to find an answer by searching the Web.
- Try to find an answer by reading the manual.
- Try to find an answer by reading a FAQ.
- Try to find an answer by inspection or experimentation.
- Try to find an answer by asking a skilled friend.
- · If you're a programmer, try to find an answer by reading the source code.

#### **Asking Questions**

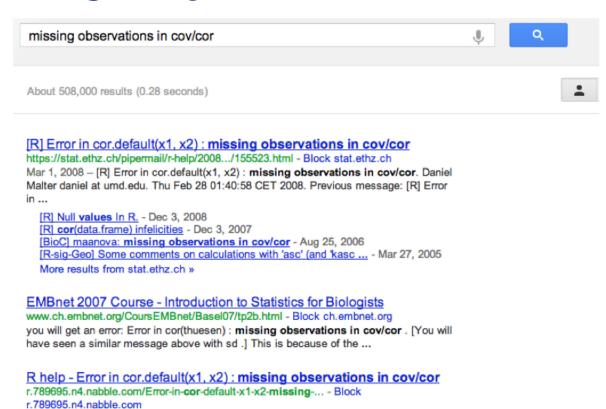
- It's important to let other people know that you've done all of the previous things already
- · If the answer is in the documentation, the answer will be "Read the documentation"
  - one email round wasted

# **Example: Error Messages**

- > library(datasets)
- > data(airquality)
- > cor(airquality)

Error in cor(airquality): missing observations in cov/cor

# Google is your friend



Error in cor.default(x1, x2): missing observations in cov/cor. Hello, I'm trying to do

cor(x1,x2) and I get the following error: Error in cor.default(x1, ...

13 posts - 5 authors - Feb 27, 2008

#### **Asking Questions**

- · What steps will reproduce the problem?
- · What is the expected output?
- What do you see instead?
- What version of the product (e.g. R, packages, etc.) are you using?
- What operating system?
- · Additional information

#### **Subject Headers**

- · Stupid: "Help! Can't fit linear model!"
- · Smart: "R 2.15.0 lm() function produces seg fault with large data frame, Mac OS X 10.6.3"
- · Smarter: R 2.15.0 lm() function on Mac OS X 10.6.3 -- seg fault on large data frame

#### Do

- · Describe the goal, not the step
- · Be explicit about your question
- · Do provide the minimum amount of information necessary (volume is not precision)
- Be courteous (it never hurts)
- Follow up witht the solution (if found)

#### Don't

- Claim that you've found a bug
- Grovel as a substitute for doing your homework
- Post homework questions on mailing lists (we've seen them all)
- · Email multiple mailing lists at once
- Ask others to debug your broken code without giving a hint as to what sort of problem they should be searching for

# Case Study: A Recent Post to the R-devel Mailing List

Subject: large dataset - confused

Message:

I'm trying to load a dataset into R, but

I'm completely lost. This is probably

due mostly to the fact that I'm a

complete R newb, but it's got me stuck

in a research project.!

#### Response

Yes, you are lost. The R posting guide is at http://www.r-project.org/posting-guide.html and will point you to the right list and also the manuals (at e.g. http://cran.r-project.org/manuals.html, and one of them seems exactly what you need).!

# **Analysis: What Went Wrong?**

- Question was sent to the wrong mailing list (R-devel instead of R-help)
- · Email subject was very vague
- Question was very vague
- · Problem was not reproducable
- No evidence of any effort made to solve the problem
- RESULT: Recipe for disaster!

#### **Places to Turn**

- · Class discussion board; your fellow students
- r-help@r-project.org
- Other project-specific mailing lists (This talk inspired by Eric Raymond's "How to ask questions the smart way")