• Left over from last time
  – How are things going with reproducing Remy results?
• Overall goal of experiment design:
  – Learn as much as possible from as few experiments as possible
• Some terminology
  – Response variable: outcome
    * Why call it a variable?
    * Examples of non-performance variables?
  – Factors: things you change
    * Why call them predictor variables?
  – Primary / secondary factors
    * How to decide which ones to use?
  – Replication: How many reactions
• Important properties of experiment design
  – Every experiment should get you closer to answering one of the questions
  – You should be able to explain all behavior in the results—if not, you may need more experiments
  – Control all variables you can
  – Measure the variables you can’t control
• Interacting factors
  – Understand which of your factors interact, and which are independent
  – Saves you a lot of time not running experiments that don’t reveal more information
  – May take a few experiments to determine
  – If you know for sure they are independent, make sure to say so in the paper
• Sensitivity analysis
  – Is your system sensitive to processor speed? Disk speed? Bandwidth?
  – The world will change in the future (or even on other systems right now)
  – Understand what changes would, and would not, affect your system
    * What if disks get faster?
* Inevitable increases in bandwidth
* Greater parallelism
  – Obviously there are limits to what you can/should actually test
  – Usually just report in a sentence
    * Might seem unsatisfying to you, but it’s helpful to you readers

• Common mistakes
  – Ignoring variation in experimental error
  – Not controlling params
  – One factor at a time experiments
  – Not isolating effects
  – Too many experiments
    * Break into several sub-evals to answer questions, evaluate particular pieces of the SUT

• Method for determining if effects are significant
  – From chapter 20 in the book if you want to read it
  – Set up: you have a categorical variable, you want to understand what the affects of each choice are and whether they are significant
  – Basic idea: try to explain the numbers in terms of two components:
    * Actual effects
    * Experimental error (randomness)
    – See rcommands.txt

• For next time
  – lab1 due Friday at midnight
  – Note that papers3 was posted
    * Due Tuesday after spring break
    * One paper looks long but isn’t really
  – For Tuesday: stay tuned, will post tonight or tomorrow AM