# Computer Science Examples

# Experimental Design:
# Comparison of *Method-1* and *Method-2*

- A computer scientists wants to compare the running time of two sorting algorithms. He tells Joe to implement *Method-1* and Steve to implement *Method-2*. Joe's program runs in 4.1 seconds while Steve's program runs in 2.3 seconds.

- Is *Method-2* faster than *Method-1*?

- What are extraneous variables that influence the running time of a program?

# Experimental Design

- This finding would be meaningless if, for example, Joe programmed the algorithm in Java and ran the program on an old Windows computer while Steve programmed the algorithm in C++ and ran it on a new MacBook Pro.

- How should we design this experiment to accurately compare the two algorithms?

# Is the sample representative of the population?

- A sorting algorithm tested on *random* data does not necessarily tell you anything about its performance on *nearly sorted* data, or *nearly reverse-sorted* data.

- If one method performs better than another when implemented in C++, will it also be more efficient if implemented in Java?

- A Bing vs. Google study taken only (or mostly) by CS students does not necessarily tell you anything about the preferences of all college students, or all adults, etc.

# The wording of the question matters…

- These are all potentially different questions with different answers:
  - Which search engine results do you prefer?
  - Which search engine returns results most relevant to your query?
  - What search engine results are the most visually appealing?

# … as does the variable being measured

- These are all potentially different questions with different answers:
  - What is the optimal traffic light strategy for a city (not clear what *optimal* means)?
  - What traffic light strategy maximizes a vehicle's average velocity?
  - What traffic light strategy minimizes the average time of a person's commute.
  - What traffic light strategy minimizes the variation of a person's commute