

Statistical Thinking

- statistics is the analysis and interpretation of data
- to have meaning data must have context
- SO meaningless
 - \$SO good
 - SO%. bad
 - SOlb ?

Principles

1) Data illuminates

"What % of American population is black? white?"

White Americans

23.8% black 49.9% white

Actually

12.3% black 75.1% white

2) Data beats anecdotes

anecdote - story that sticks in our mind.

- very misleading

* a statistically literate person knows data is more reliable than anecdotes because they describe an overall picture not an isolated incident.

- 3) Beware of lurking variables
variable that is not the response or the explanatory
yet affects the outcome

<u>Soph</u>	<u>Senior</u>
3.84 GPA	2.57 GPA

Conclude sophomores are smarter

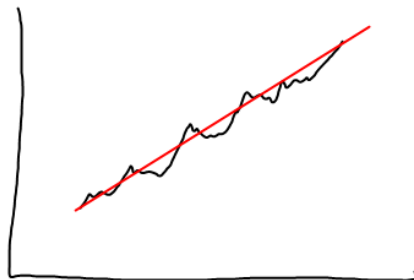
Problems?

level of student (lurking variable)

- 4) Where data comes from is vital
* How it is obtained is also vital?

- 5) Variation is everywhere
individuals vary

economy



c) Conclusions are not certain

Because of variation, conclusions are not certain

Introduction

individuals - objects described by a set of data
- do not have to be people

variable - characteristic of the individual

Two types of variables

1) Categorical - place individuals into categories
race, gender, eye color

2) Quantitative - numeral description of the individual
SAT score, GPA, blood pressure

Distribution

- patterns of variation
- what values the variable takes and how often it takes these values.