1.

Examples:

Placebo Effect: In one clinical trial on the common cold, a volunteer population was randomly divided into a treatment group and control group. The treatment group were given vitamin C but told it was the placebo. The control group was given the placebo but told it was vitamin C. At the end of the trial the control group reported fewer colds.

The Physicians Health Study examined the claim that aspirin taken regularly can reduce the risk of heart attacks. More than 22,000 male physicians were divided into two groups. One group took a dose of aspirin every second day, while the other took a placebo pill that was identical in appearance. When 189 of the 11,034 subjects in the placebo group had suffered heart attacks compared to 104 of the 11,037 men in the treatment group, the researchers stopped the trial and announced the result. Deciding when evidence is conclusive is sometimes a difficult problem both statistically and ethically.

For the year 2009, the FBI reported that there were 778 murders and 2,586 forcible rapes in New York . For Utah in 2009, there were 37 murders and 905 forcible rapes. Are you safer in Utah? New York's population was 19,541,453; Utah's population was 2,784,572.

	New York	Utah
Murders per 100,000	4	1.3
Rapes per 100,000	13.2	32.5

$\frac{778}{19,541,453} \times 100,000$

$\frac{37}{2,784,572} \times 100,000$

$\frac{905}{19,541,453} \times 100,000$

$\frac{905}{2,784,572} \times 100,000$

2.

Chapter 2: Observational Studies

- In an observational study the subjects determine whether they get the treatment or the control ("self-selection") *e.g. smoking studies, health and fitness studies (usually)*
- Always find confounding factors
- Try to "control for" confounding factors by comparing small, homogeneous groups, *e.g. compare male smokers age 55-59 to male nonsmokers age 55-59*

3.

Example 1. Children of women who smoked during pregnancy scored 9 points less, on average, in IQ tests at ages 3 and 4 than children of nonsmokers. Does this imply that smoking during pregnancy causes the baby to have a lower IQ? Suggest a possible confounding factor.

4.

Example 2. People who exercise tend to live longer than those who don't exercise. Does this imply that exercising causes people to live longer? Suggest a possible confounding factor.

5a

2. Data from the National Health Interview Survey for 1985 are shown below. True or False: the data show that as people age, they become more health conscious and adopt more healthful lifestyles.

Age	Eats Breakfast	Current Drinker	Current Smoker
18–64	70%	40%	35%
65 & over	90%	10%	15%

Association is not Causation

Just because two things are associated, it does not necessarily imply that one of them causes the other one. Confounding factors might be driving the association.

- "Pellagra"
- Cervical cancer and circumcision
- Ultrasound and low birthweight
- Chocolate and car accidents

1. Recent census data (2008) indicates that for full-time workers, the median income for women was only 77% of the median income for men.

Discrimination?

2. Two studies conducted in Denver, Colorado found that children with leukemia were more likely to live in areas with large electric power lines than children without leukemia.

Confounding factors?

3. Admissions to the Graduate School at UC-Berkeley in 1973.

44% of male applicants were admitted

30% of female applicants were admitted

Bias?

#1 Overall : Men 44% Women 35%

Berkeley graduate admissions data

6 largest :

	Men		*Women*	
Major	*Number of applicants*	*Percent admitted*	*Number of applicants*	*Percent admitted*
A	825	62	108	82
B	560	63	25	68
C	325	37	593	34
D	417	33	375	35
E	191	28	393	24
F	373	6	341	7
Total	2,691	44	1835	30

6.8.

Berkeley graduate admissions

Major	*Male applicants*	*Female applicants*	*Total applicants*
A	825	108	933
B	560	25	585
C	325	593	918
D	417	375	792
E	191	393	584
F	373	341	714
Total	2,691	1,835	4,526

The weighted average for men is

$$62\% \times \frac{933}{4{,}526} + 63\% \times \frac{585}{4{,}526} + 37\% \times \frac{918}{4{,}526} + 33\% \times \frac{792}{4{,}526} + 28\% \times \frac{584}{4{,}526} + 6\% \times \frac{714}{4{,}526} \approx 39\%$$

The weighted average for women is

$$82\% \times \frac{933}{4{,}526} + 68\% \times \frac{585}{4{,}526} + 34\% \times \frac{918}{4{,}526} + 35\% \times \frac{792}{4{,}526} + 24\% \times \frac{584}{4{,}526} + 7\% \times \frac{714}{4{,}526} \approx 43\%$$

Simpson's Paradox

Example (hypothetical):

Job A: 19 men, 1 woman, all earn $70,000

Job B: 19 women, 1 man, all earn $50,000

Overall:

Average for men = $69,000

Average for women = $51,000

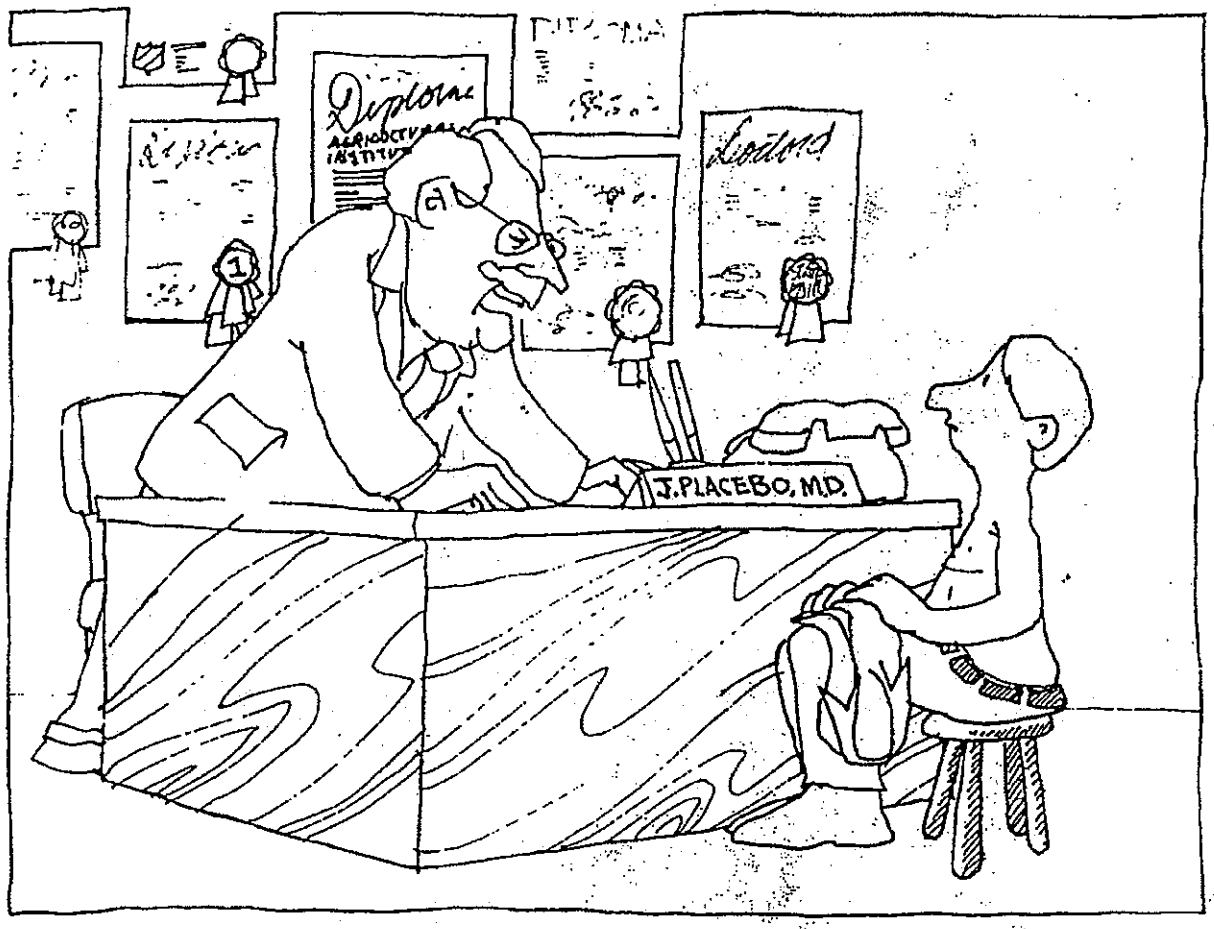
Why?

Women tend to have the lower-paying job.

Simpson's Paradox, Summary

- Overall averages or percentages can be misleading.

- If we want to understand what's going on, we need to look at the averages or percentages at the decision-making level, i.e. break up the data into homogeneous groups.



"I want to make one thing perfectly clear, Mr. Smith. The medication I prescribe will cure that run-down feeling."