The goal of the next two classes is to learn to identify and describe errors in statistical reasoning. Some of the more common types of errors:

1. Sample bias
2. Ignoring base rates
3. Misuse of aggregated/disaggregated data
4. Interpreting correlation as causation
5. Ignoring regression to the mean
6. Inappropriately extrapolating outside of a sample.

A random sample of 20,000 Americans found that those whose primary information source is NPR are 12 times more likely to be able to answer simple questions about US history than those who prefer Fox News. Therefore, NPR better educates its consumers than Fox.

None of Michael Jordan's kids made the NBA. Growing up rich destroys the incentive for hard work!

How could Obama have won? I only know 1 person who voted for him!

Gay/lesbian congressmen vote in an overwhelmingly liberal manner on economic issues. This implies there is a natural axis between social and economic issues.

85% of severe car accidents occur within 10 miles of the drivers' homes. You should always wear your seatbelt, even if just driving down the street!

Among women who carry their babies to full term, birthweight among babies of smokers and birthweight among babies of nonsmokers are the same on average. Among women who give birth prematurely, babies of smokers are slightly *heavier* than babies of nonsmokers. Do these facts together imply that smoking during pregnancy does not result in lower birthweight on average, as is commonly asserted?

UC-Berkeley was sued in the 1970's for bias against women applying to graduate school. The plaintiffs' alleged that Berkeley was more likely to admit men than women. In 1973, the following data described the school's graduate admissions:

	Applicants	**Admitted**
Men	8,442	44%
Women	4,321	35%

Is this evidence of systematic bias against women?

UC-Berkeley was sued in the 1970's for bias against women applying to graduate school. The plaintiffs' alleged that Berkeley was more likely to admit men than women. In 1973, the following data described the school's graduate admissions:

	Applicants	**Admitted**
Men	8,442	44%
Women	4,321	35%

Is this evidence of systematic bias against women?

Not necessarily. In particular, not if women are disproportionately likely to apply to departments that are more competitive. We need to disaggregate the data:

Department	**Men**	**Women**		
	Applicants	**Admitted**	**Applicants**	**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	272	6%	341	**7%**

We see that in 4 out of the 6 departments with the highest number of applicants, a higher proportion of women than men were admitted. In the two that admit a higher proportion of men, the difference is quite small.

A university president wants to admit more minority students to increase campus diversity. He offers funds to the deans of each of 4 colleges for recruiting more minority candidates. 5 years later, he is shown the following data:

College	Before	After		
	Minority applicants	Accepted	Minority applicants	accepted
A	102	51 (50%)	180	108 (60%)
B	400	300 (75%)	280	225 (80%)
C	120	30 (25%)	340	112 (30%)
D	500	350 (70%)	350	280 (80%)

Each college admitted a higher proportion of minority applicants. The total number of minority applicants increased. Was the president's program a success?

A university president wants to admit more minority students to increase campus diversity. He offers funds to each of 4 colleges for deans to recruit and admit more minority candidates. 5 years later, he is shown the following data:

College	Before	After		
	Minority applicants	Accepted	Minority applicants	accepted
A	102	51 (50%)	180	108 (60%)
B	400	300 (75%)	280	225 (80%)
C	120	30 (25%)	340	112 (30%)
D	500	350 (70%)	350	280 (80%)

Each college admitted a higher proportion of minority applicants. The total number of minority applicants increased. Was the president's program a success?

No. Let's look at the aggregated data:

College	Before	After		
	Minority applicants	Accepted	Minority applicants	accepted
Total	1,122	731 (65.2%)	1,160	725 (62.5%)