

| US Students taking College Statistics by Major | | | | |
|--|--------------------|--------|-------|-----|
| Business/Econ | Sociology /SocWork | Health | Pysch | Bio |
| 40% | 20% | 18% | 12% | 10% |

2011 US 914K: https://nces.ed.gov/programs/digest/d13/tables/dt13_322.10.asp

1) Based on their majors, 57% of four-year college students take introductory statistics: statistical inference.

Most college students taking introductory statistics (inference) are in majors that deal mainly with observational studies.

2) Today's students are interested in using disparities as evidence of discrimination. Almost all disparities are based on observational data.

3) Confounding is the “elephant in the room” in introductory statistics.

As teachers, we know about it.

Most textbooks (80%) don't include confounding in their index.

If they do mention confounding, they limit it to one page where they tell students that “association is not causation”.



4) Confounding is not a big deal in the second course that focuses on multivariate analysis for prediction. In predicting you don't care about confounding. You are more concerned about getting a higher R-squared. Predictive analytics is much more common than descriptive analytics.

Confounding: Survey Results

2016 GIASE update recommends teaching multivariate thinking (and confounding) in intro statistics.

a. strongly disagree b. Disagree c. Neutral d. Agree. e. Strongly agree

1. Confounding is as important as randomness in studying variation.

a. 0 S-Dis b. 0 Dis c. 1 Neutral d. 6 Agree . e. 4 S-Agree

2. Confounding should be featured (15-20%) in an introductory statistical-inference class.

a. 0 S-Dis b. 0 Dis c. 3 Neutral d. 5 Agree . e. 3 S-Agree

3. Confounding should be featured in a separate introductory statistics class.

a. 1 S-Dis b. 3 Dis c. 4 Neutral d. 1 Agree . e. 2 S-Agree

4. Which of the following are your main **hesitations** about *including confounding in your* statistics course?* (Enter "1" for main reason; check all others that apply)

2 Need knowledge/training to do it.

2 Not aware of suitable curriculum or textbook.

3 Syllabus too full.

0 Not relevant to course goals or audience

2 Not supported internally by department/chair/teachers

0 Not supported by the statistics education field.

1 Too close to causation.

1 No hesitation

0 OTHERn _____

5. Which of the following are (would be) your main **motivations** for *including (or considering including) confounding* in your* statistics course? (Enter "1" for main reason; check all others that apply)

2 All or almost all

5 Increase relevance to course audience.

1 Increase fit with course goals.

0 Called for by the statistics education field.

0 Chance to be at the cutting edge.

0 Is ultimately easier to teach.

3 Is more interesting or creative to teach.

0 close connection with causation.

0 OTHER _____

* Should have said "in your introductory statistics course".