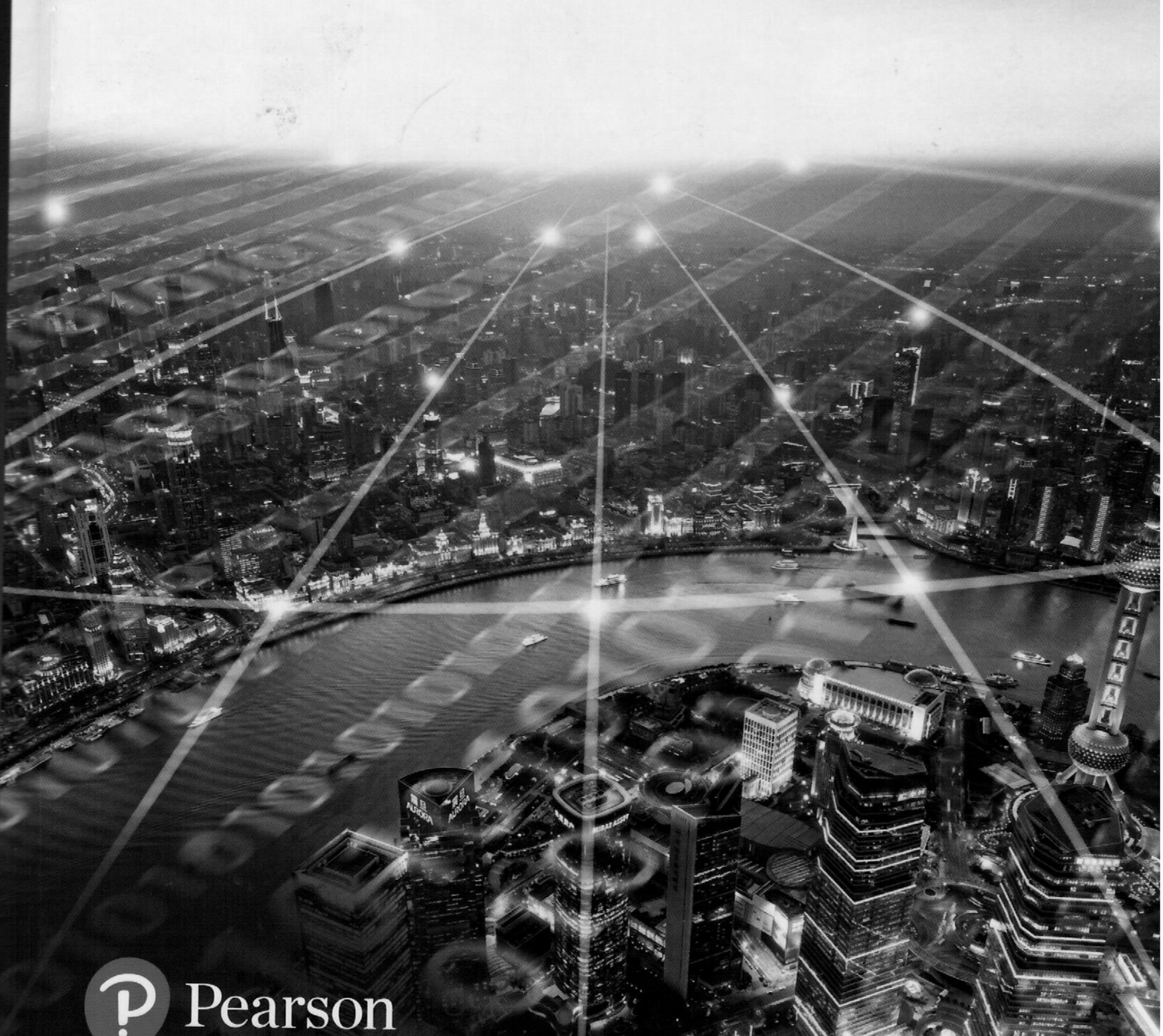
Intro STATS

FIFTH EDITION

De VEAUX | VELLEMAN | BOCK



Pearson

Director, Portfolio Management: *Deirdre Lynch*
Courseware Portfolio Manager: *Patrick Barbera*
Content Producer: *Sherry Berg*
Managing Producer: *Karen Wernholm*
Senior Producer: *Stephanie Green*
Manager, Courseware QA: *Mary Durnwald*
Manager, Content Development: *Bob Carroll*
Product Marketing Manager: *Emily Ockay*
Field Marketing Manager: *Andrew Noble*
Marketing Assistants: *Jennifer Myers, Erin Rush*
Senior Author Support/Technology Specialist: *Joe Vetere*
Manager, Rights and Permissions: *Gina Cheselka*
Manufacturing Buyer: *Carol Melville, LSC Communications*
Art Director: *Barbara Atkinson*
Production Coordination, Composition, and Illustrations: *Cenveo® Publisher Services*
Cover Image: *Liping/Shutterstock*

Library of Congress Cataloging-in-Publication Data

Names: De Veaux, Richard D. | Velleman, Paul F., 1949— | Bock, David E.
Title: Intro stats.
Description: Fifth edition / Richard D. De Veaux, Williams College, Paul F. Velleman, Cornell University, David E. Bock, Cornell University | Boston: Pearson, [2018] | Includes index.
Identifiers: LCCN 2016016733 | ISBN 9780134210223 ((hardcover)) | ISBN 0134210220 ((hardcover))
Subjects: LCSH: Statistics—Textbooks.
Classification: LCC QA276.12.D4 2018 | DDC 519.5—dc23
LC record available at **https://lccn.loc.gov/2016016733**

1 18

Pearson

www.pearson.com

Student Edition
ISBN 13: 978-0-13-421022-3
ISBN 10: 0-13-421022-0

same way as blocking. For example, a retrospective study of music education and grades might match each student who studies an instrument with someone of the same sex who is similar in family income but didn't study an instrument. We could then compare grades of music students with those of non-music students. The matching would reduce the variation due to income and sex differences.

Blocking for experiments is the same idea as stratifying is for sampling. Both methods group together subjects that are similar and randomize within those groups as a way to remove unwanted variation. (But be careful to keep the terms straight. Don't say that we "stratify" an experiment or "block" a sample.) We use blocks to reduce variability so we can see the effects of the factors; we're not usually interested in studying the effects of the blocks themselves.

EXAMPLE 11.5 Blocking

RECAP: In 2007, pet food contamination put cats at risk, as well as dogs. Our experiment should probably test the safety of the new food on both animals.

QUESTIONS: Why shouldn't we randomly assign a mix of cats and dogs to the two treatment groups? What would you recommend instead?

ANSWERS: Dogs and cats might respond differently to the foods, and that variability could obscure my results. Blocking by species can remove that superfluous variation. I'd randomize cats to the two treatments (test food and safe food) separately from the dogs. I'd measure their responses separately and look at the results afterward.

Confounding

Professor Stephen Ceci of Cornell University performed an experiment to investigate the effect of a teacher's classroom style on student evaluations. He taught a class in developmental psychology during two successive terms to a total of 472 students in two very similar classes. He kept everything about his teaching identical (same text, same syllabus, same office hours, etc.) and modified only his style in class. During the fall term, he maintained a subdued demeanor. During the spring term, he used expansive gestures and lectured with more enthusiasm, varying his vocal pitch and using more hand gestures. He administered a standard student evaluation form at the end of each term.

The students in the fall term class rated him only an average teacher. Those in the spring term class rated him an excellent teacher, praising his knowledge and accessibility, and even the quality of the textbook. On the question "How much did you learn in the course?," the average response changed from 2.93 to 4.05 on a 5-point scale.[9]

How much of the difference he observed was due to his difference in manner, and how much might have been due to the season of the year? Fall term in Ithaca, New York (home of Cornell University), starts out colorful and pleasantly warm but ends cold and bleak. Spring term starts out bitter and snowy and ends with blooming flowers and singing birds. Might students' overall happiness have been affected by the season and reflected in their evaluations?

Unfortunately, there's no way to tell. Nothing in the data enables us to tease apart these two effects, because all the students who experienced the subdued manner did so during the fall term and all who experienced the expansive manner did so during the spring. When the levels of one factor are associated with the levels of another factor, we say that these two factors are **confounded**.

[9]But the two classes performed almost identically well on the final exam.

In some experiments, such as this one, and some observational studies as well, it's just not possible to avoid some confounding. Professor Ceci could have randomly assigned students to one of two classes during the same term, but then we might question whether mornings or afternoons were better, or whether he really delivered the same class the second time (after practicing on the first class). Or he could have had another professor deliver the second class, but that would have raised more serious issues about differences in the two professors and concern over more serious confounding.

EXAMPLE 11.6 Confounding

RECAP: After many dogs and cats suffered health problems caused by contaminated foods, we're trying to find out whether a newly formulated pet food is safe. Our experiment will feed some animals the new food and others a food known to be safe, and a veterinarian will check the response.

QUESTION: Why would it be a bad design to feed the test food to some dogs and the safe food to cats?

ANSWER: This would create confounding. We would not be able to tell whether any differences in animals' health were attributable to the food they had eaten or to differences in how the two species responded.

A Two-Factor Example

Confounding can also arise from a badly designed multifactor experiment. Here's a classic. A credit card bank wanted to test the sensitivity of the market to two factors: the annual fee charged for a card and the annual percentage rate charged. Not wanting to scrimp on sample size, the bank selected 100,000 people at random from a mailing list. It sent out 50,000 offers with a low rate and no fee and 50,000 offers with a higher rate and a $50 annual fee. Not surprising, people preferred the low-rate, no-fee card. In fact, they signed up for that card at over twice the rate as the other offer. And because of the large sample size, the bank was able to estimate the difference precisely. But the question the bank really wanted to answer was "how much of the change was due to the rate, and how much was due to the fee?" Unfortunately, there's simply no way to separate out the two effects. If the bank had tested all four treatments—low rate with no fee, low rate with $50 fee, high rate with no fee, and high rate with $50 fee—each to 25,000 people, it could have learned about both factors and could have also seen what happens when the two factors occur in combination.

Lurking and Confounding?

A lurking variable creates an association between two other variables that tempts us to think that one may cause the other. Recall from the example in Chapter 8, that people in countries with more TV sets per capita tend to have longer lives. You shouldn't conclude it's the TVs "causing" longer life. It's more likely that a generally higher standard of living allows people to afford more TVs and get better health care, too. Our data revealed an association between TVs and life expectancy, but economic conditions were a likely lurking variable. A lurking variable, then, is usually thought of as a variable associated with both y and x that makes it appear that *x* may be causing *y*.

Confounding and lurking variables are very similar. Imagine an observational study hoping to understand the relationship between herbal supplements and patient health finds that patients who take the supplements report fewer colds. However, if they find from their survey that the patients who take the herbal supplements also tend to take larger doses of Vitamin C, we would say that taking Vitamin C is a confounder of herbal supplements. Had we not asked the question at all, and we later found that taking Vitamin C was more effective in preventing colds than the herbal supplement, we might call Vitamin C a lurking variable in the original study.

Both confounding and lurking variables are outside influences that make it harder to understand the relationship we are modeling. It's important to realize that in any observational study or even in a carefully designed experiment, there may be variables that influence the relationship between that variable and the response other than the ones being studied. You should always be alert for the possible effects of other variables on the coefficients you care about. Be especially wary of variables that you might not have considered.

WHAT CAN GO WRONG?

- **Don't give up just because you can't run an experiment.** Sometimes we can't run an experiment because we can't identify or control the factors. Sometimes it would simply be unethical to run the experiment. (Consider randomly assigning students to take—and be graded in—a statistics course deliberately taught to be boring and difficult or one that had an unlimited budget to use multimedia, real-world examples, and field trips.) If we can't perform an experiment, an observational study may be a good choice.
- **Beware of confounding.** Be aware of variables that may be confounded. In a prospective study, it may be possible to stratify the subjects by levels of one variable. In an experiment, unmeasured confounders will be balanced (on average) by randomization. To include a variable that may be a confounder, it is a good idea to block by the potential confounder to ensure that the levels are balanced. And always think about possible lurking variables that may be influencing the response that aren't in your study as well.
- **Bad things can happen even to good experiments.** Protect yourself by recording additional information. An experiment in which the air conditioning failed for 2 weeks, affecting the results, was saved by recording the temperature (although that was not originally one of the factors) and estimating the effect the higher temperature had on the response.[10]

  It's generally good practice to collect as much information as possible about your experimental units and the circumstances of the experiment. For example, in the nail polish experiment, it would be wise to record details (temperature, humidity) that might affect the durability of the polish on the acrylic nails. Sometimes we can use this extra information during the analysis to reduce biases.
- **Don't spend your entire budget on the first run.** Just as it's a good idea to pretest a survey, it's always wise to try a small pilot experiment before running the full-scale experiment. You may learn, for example, how to choose factor levels more effectively, about effects you forgot to control, and about unanticipated confoundings.

[10] R. D. De Veaux and M. Szelewski, "Optimizing Automatic Splitless Injection Parameters for Gas Chromatographic Environmental Analysis." *Journal of Chromatographic Science* 27, no. 9 (1989): 513–518.

APPENDIX
C Indexes

Datasets Index

BE = Boxed Example; E = Exercise; IE = In-Text Example; JC = Just Checking; RM = Random Matters; SBS = Step-by-Step examples.

T

U

V

Z