Data 8, Lab 2

Causality, Expressions, and Table Operations

Hubert Luo
Fall 2019

6 September 2019

Berkeley
UNIVERSITY OF CALIFORNIA

Agenda

1. Causality
2. Expressions: Variables and Function Calls
3. Some Tables Operations
4. Lab Notebook

Lab Assistants

We have three lab assistants who will be joining us this semester!

Causality

- Treatment: Factor of interest
- Outcome: What is being measured

- Example: Study the effect of capital punishment on a state's murder rate
  - Treatment: Whether a state has capital punishment
  - Outcome: A state's murder rate

Causality

- Association: Any relationship between the treatment and outcome (i.e., there is a **correlation** between them)
- Causation: The treatment *causes* the outcome
  - **Association does not imply causation!!!!**

- Example: Study the effect of capital punishment on a state's murder rate
  - Association: Most states who have capital punishment also have high murder rate
  - Outcome: Having capital punishment causes states to have high murder rate

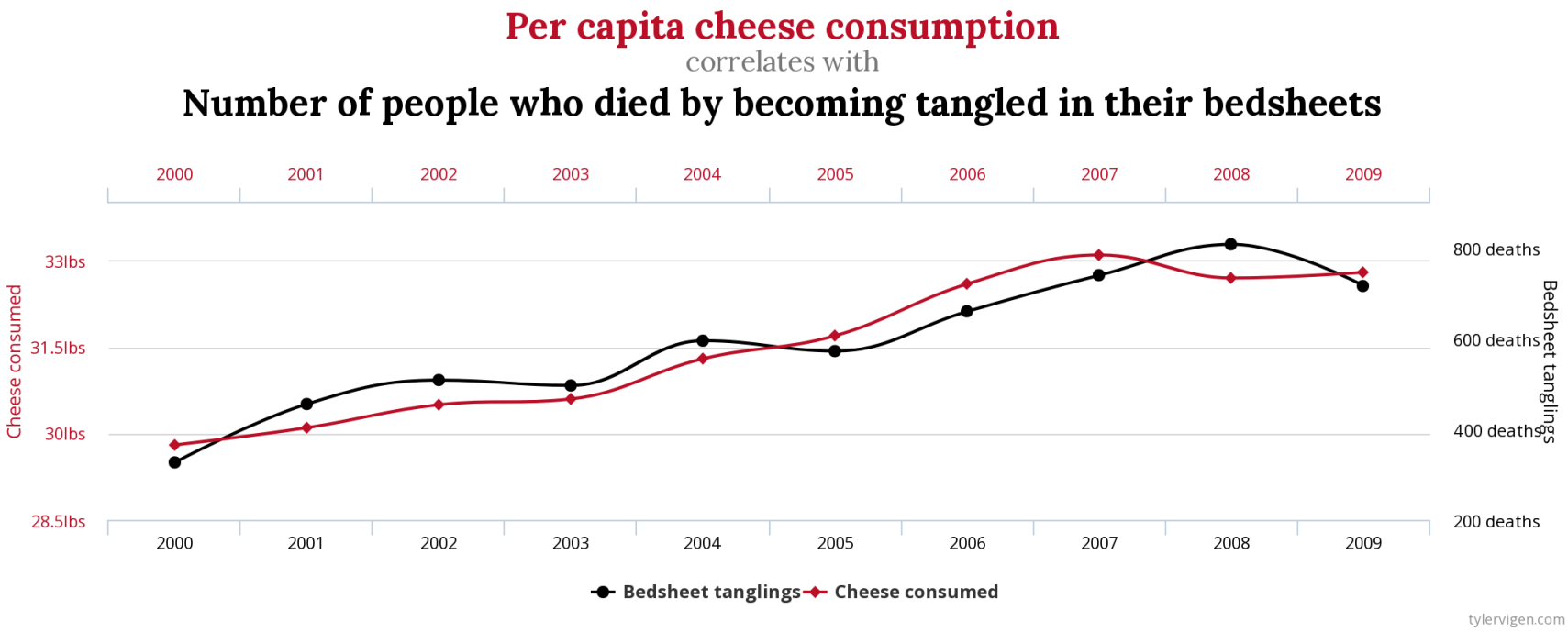
Association vs. Causation

- Two variables being associated does **not** mean one causes the other!!
- Example: Study the effect of capital punishment on a state's murder rate
  - Even if states with capital punishments also have high murder rates, it does **not** mean having capital punishment *causes* them to have high murder rates

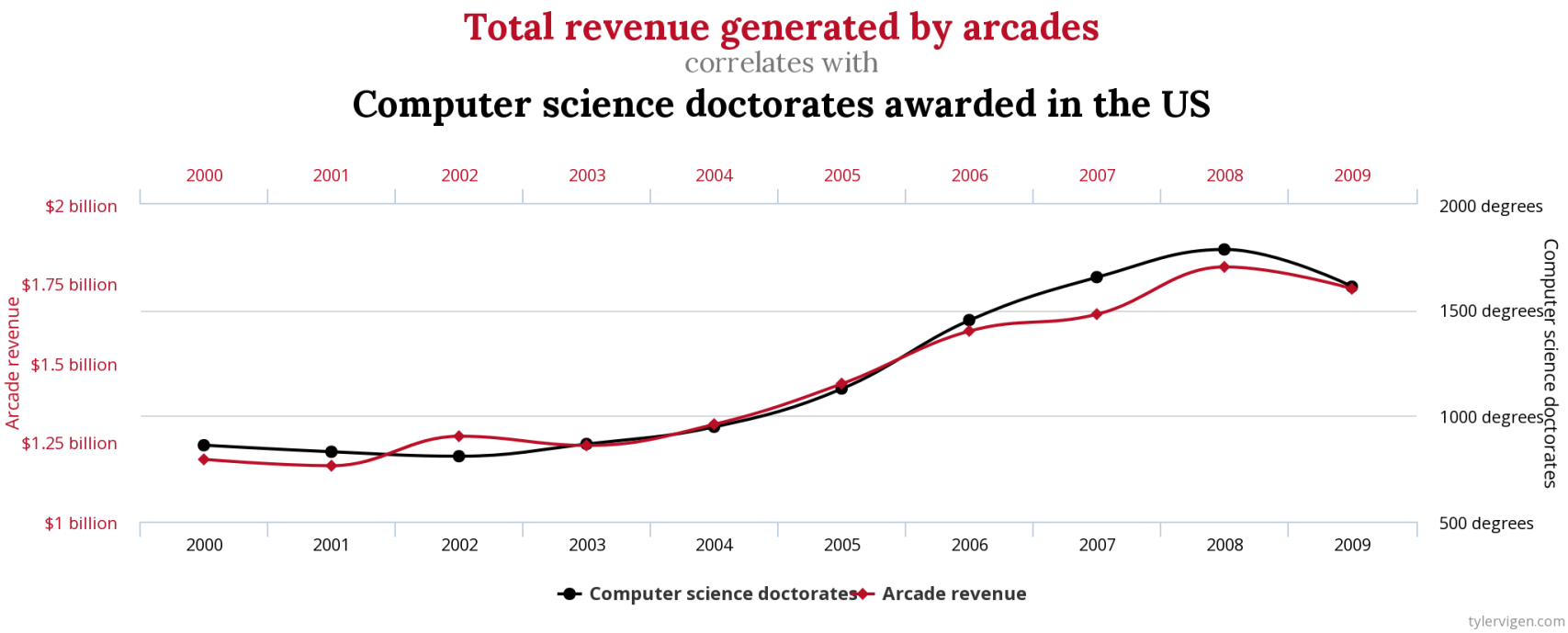
Association vs. Causation

- Confounding Factors: Other variables unaccounted for in relationship between treatment and outcome
- Example: Study the effect of capital punishment on a state's murder rate
  - Lots of other possible explanations: literacy rates, living standards, unemployment, gun control
  - We do not know which one *causes* higher murder rates

Association vs. Causation

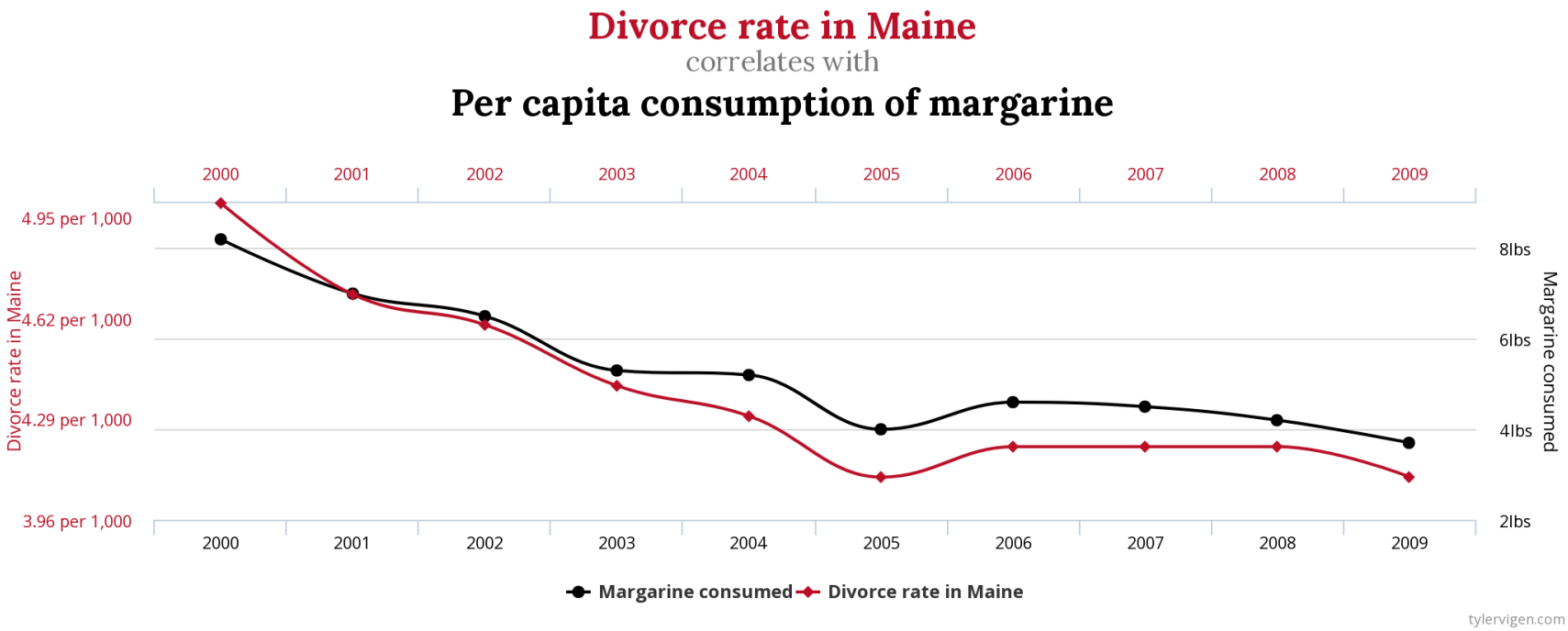


Association vs. Causation

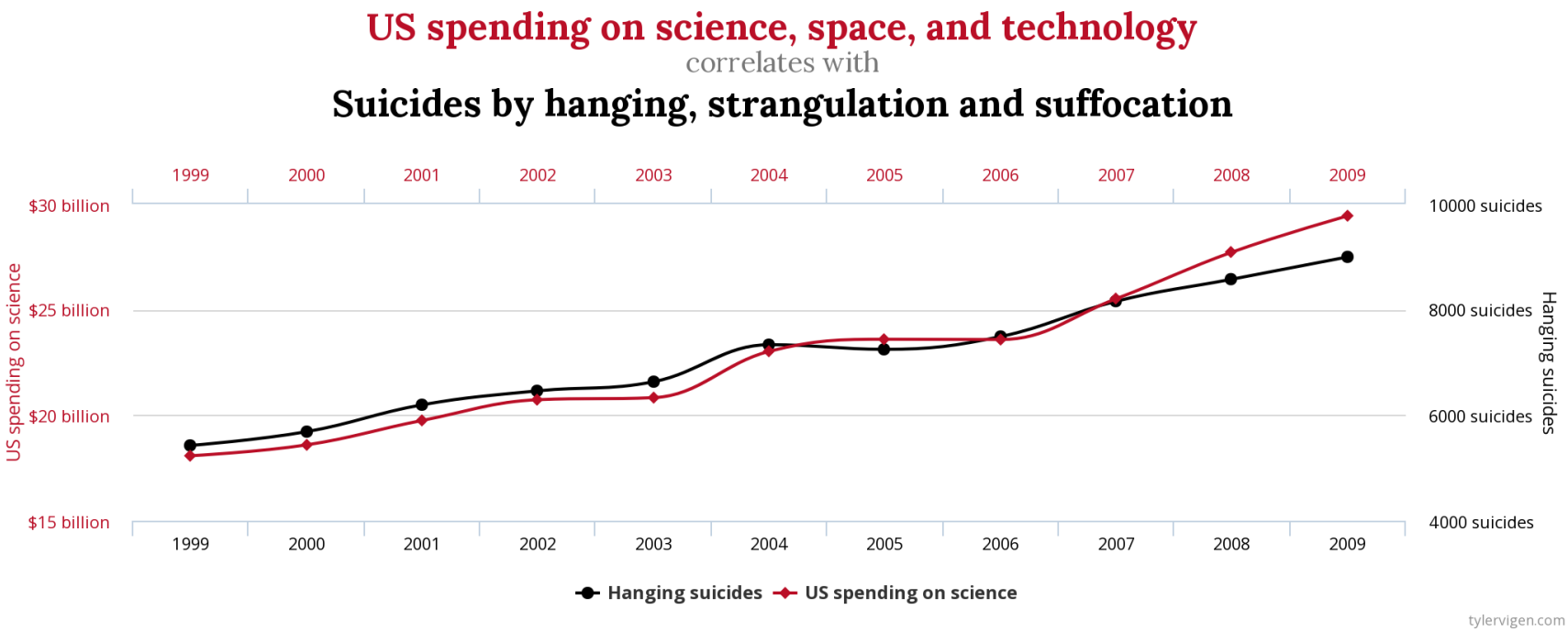


Berkeley
UNIVERSITY OF CALIFORNIA

Association vs. Causation



Association vs. Causation



Establishing Causality

- Treatment Group: Treatment applied
- Control Group: Treatment not applied
- Example: Study the effect of capital punishment on a state's murder rate
  - Treatment Group: States with capital punishment
  - Control Group: States without capital punishment

Establishing Causality

- Observational Study: Experimenter has no ability to divide study participants into treatment and control groups
- Randomized Control Experiment: Randomly divide participants into treatment and control group
- Example: Study the effect of capital punishment on a state's murder rate
  - Observational study since researchers cannot just randomly decide which states have capital punishment or not

Establishing Causality: Examples

- Researcher test the effect of a new drug on lung cancer by giving the drug to 75% of participants selected at random and a placebo drug to the other subjects
  - Randomized control experiment since researcher randomly selected treatment and control groups
- Psychologist stood outside an elementary school and asked the first 100 students they saw whether the student played basketball
  - Observational study since researcher doesn't control who gets asked the question

Establishing Causality

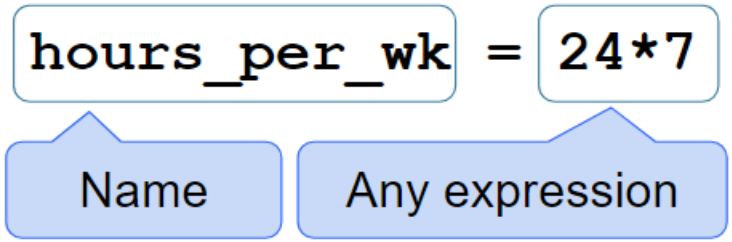
- Observational Study: Cannot prove causation!!
- Randomized Control Experiment: Can sometimes prove causation, but also other potential problems:
  - Selection bias: Study's subjects don't represent a typical member of the population
    - Example: Drug trial subjects are only white males
  - Measurement error: People don't answer truthfully
    - Example: Farmers may underestimate crop numbers to get more aid in a government agriculture survey
  - Questionnaire design: Leading questions, order of questions
    - Example: "Do you agree with America's pointless intervention in Afghanistan?"

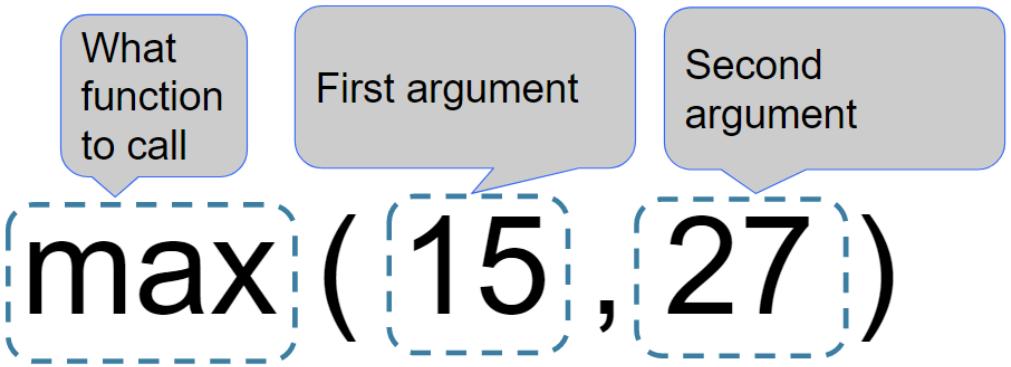
Come to OH to hear more
or check out Stat 152:
Sampling Surveys!

More Resources

- http://data8.org/materials-sp18/lec/ch2notes.pdf

Variables and Function Calls





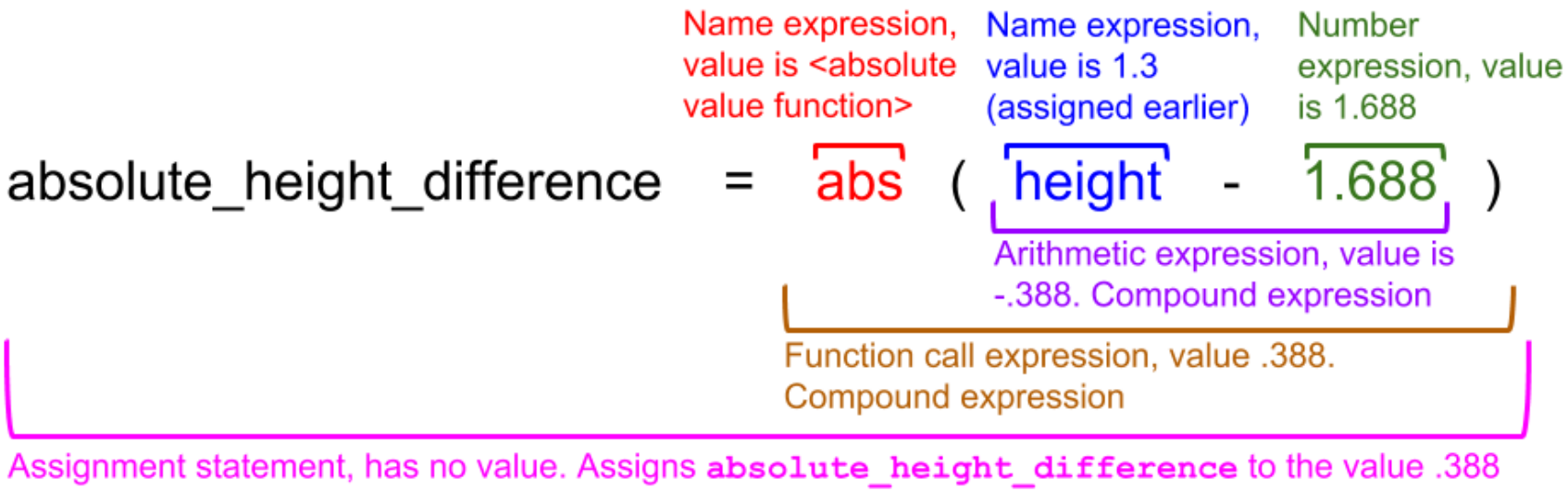
Expressions: Common Functions

Expression Type	Operator	Example	Value
Addition	+	2 + 3	5
Subtraction	-	2 - 3	-1
Multiplication	*	2 * 3	6
Division	/	7 / 3	2.66667
Remainder	%	7 % 3	1
Exponentiation	**	2 ** 0.5	1.41421

Expressions: Common Functions

Function	Description	Example
abs	Returns the absolute value of its argument	abs(-3) -> 3
max	Returns the maximum of all its arguments	max(1,2) -> 2
min	Returns the minimum of all its arguments	min(1,3) -> 1
pow	Raises its first argument to the power of its second argument	pow(2,0) -> 1
round	Rounds its argument to the nearest integer	round(0.6) -> 1

Expressions



Expressions: Demo

- (Demo on Notebook)

Some Table Operations

Name	Purpose	Example
sort	Create a copy of a table sorted by the values in a column	tbl.sort("N")
where	Create a copy of a table with only the rows that match some *predicate*	tbl.where("N", are.above(2))
num_rows	Compute the number of rows in a table	tbl.num_rows
num_columns	Compute the number of columns in a table	tbl.num_columns
select	Create a copy of a table with only some of the columns	tbl.select("N")
drop	Create a copy of a table without some of the columns	tbl.drop("N")

Some Where Predicates

Predicate	Result	Example
are.equal_to	Find rows with values equal to 50	are.equal_to(50)
are.not_equal_to	Find rows with values not equal to 50	are.not_equal_to(50)
are.above	Find rows with values above (and not equal to) 50	are.above(50)
are.above_or_equal_to	Find rows with values above 50 or equal to 50	are.above_or_equal_to(50)
are.below	Find rows with values below 50	are.below(50)
are.between	Find rows with values above or equal to 2 and below 10	are.between(2, 10)

Tables: Demo

- (Demo on Notebook)

Announcements

- HW2 due Thursday 9/12
- Tutoring section sign-ups have been released
  - Work with a tutor and other students in a small group setting

Lab Notebook

TBD