What is the difference between these two questions? Which one is a statistical question?

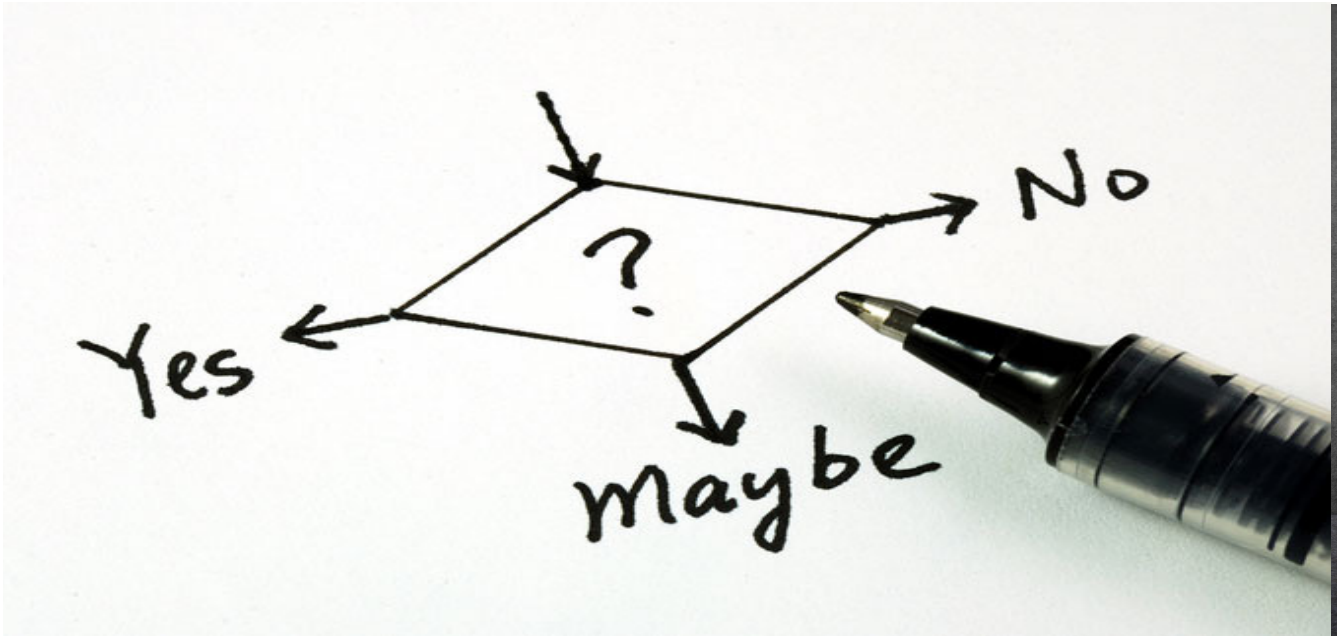
- Will it rain tomorrow?
- What is the probability that it will rain tomorrow?

Key concepts underlying statistics and statistical thinking

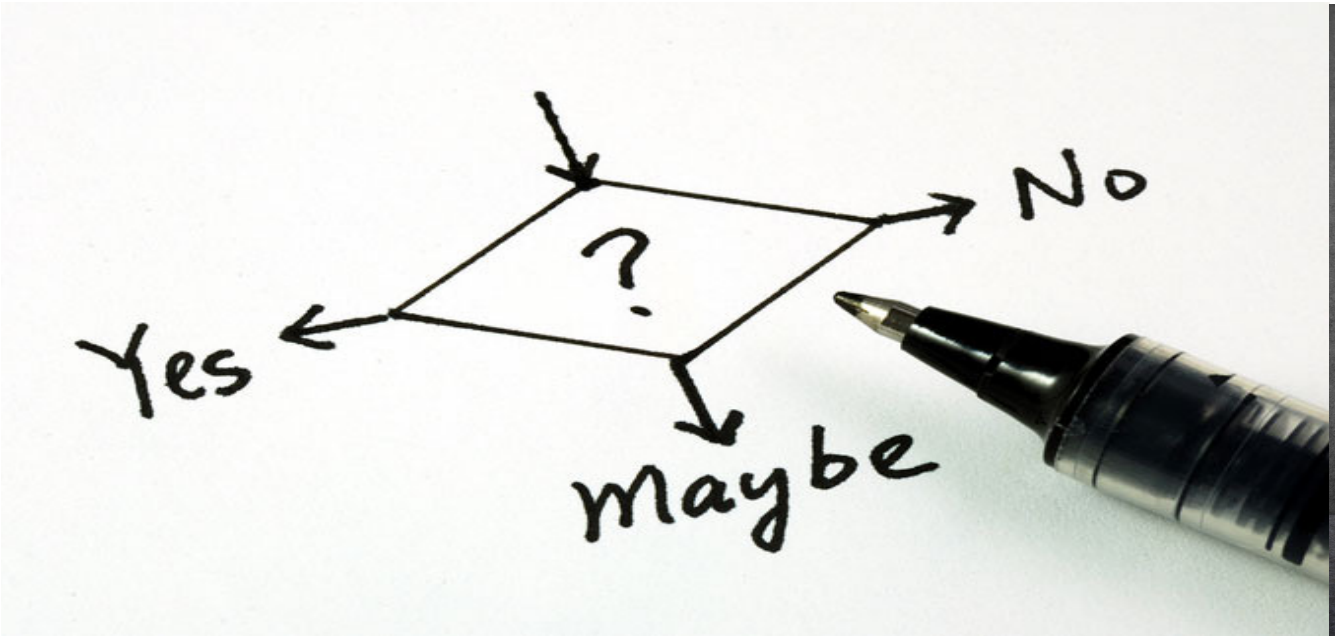
- Decision (e.g., based on probability of raining).
- Uncertainty (unpredictability).
- Risk of being wrong (error).
- Variability - Answer (estimates of the probability of raining) may change with more data (preferably towards more accurate answers, i.e., probabilities).
- Accuracy (close to reality, i.e., yes/no rain; models predicts correctly).
- Knowledge (accumulation of evidence, i.e., that the model that we used to predict rain becomes more and more accurate).

The science of aiding decision-making with incomplete information.

(or without complete knowledge)



Statistics is key in decision-making processes because most decisions are made without complete knowledge (i.e., decisions always carry some level of uncertainty).



“Statistics” as defined by the American Statistical Association (ASA) “is the science of learning from data, and of measuring, controlling and **communicating uncertainty**.”

“Statistics” as defined by the American Statistical Association (ASA) “is the science of learning from data, and of measuring, controlling and **communicating uncertainty**.”

“The statements of science are not of what is true and what is not true, but statements of what is known with different degrees of certainty.” (Richard Feynman)

Welcome to BIOL 322

Statistics for Biological Sciences (BioStatiscs)

Pedro Peres-Neto, PhD

Professor, Department of Biology, Concordia University

& Canada Research Chair

What do university professors do?

1. Research (conduct and publish; majority of research in the world is conducted at the university level).
2. Teaching.
3. Service (institutional, national, and international).
4. Professional Development.
5. Curriculum Development.
6. Graduate Student Supervision and Mentoring (MSc and PhD levels; and postdoctoral researchers).
7. Outreach and Public Scholarship.
8. Diversity, Equity, and Inclusion (DEI) Efforts.
9. Other Responsibilities (e.g., editorial work).

Don’t hesitate to raise your hand during lectures if you have any questions.

I’m also adept at 'reading the room' to notice when students have questions.

Statistics is key!

“Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write”

- Herbert George Wells

Statistics is key!

“Mathematics may rule the universe, but statistics rule societies”

- An inspiring moment during a BIOL322 lecture in 2018

Statistics is key!

Every discipline considers statistics!

What is a statistical question?

- ✓ What is the average size of Canadians?
- ✓ Is 10 a number?

What is the difference between these two types of questions?

What is a statistical question?

- ✓ What is the average size of Canadians?
- ✓ Is 10 a number?

More information (data) changes (hopefully improving) the answer; i.e., one requires statistics and the other doesn't.

Do we need statistics to calculate the number of female candidates in an election?

POLITICS October 16, 2017 6:26 pm Updated: October 18, 2017 11:35 am

Women make up 31% of candidates in Quebec municipal elections

By Raquel Fletcher

Quebec City Correspondent Global News

WEATHER Montre

26°

Weather Clear

Do we need statistics to calculate the number of female candidates?

NO! This question would be answered by simply counting the number of female candidates over the total number of candidates. This question is not answered by collecting more data that may change the results.

ARTICLE

doi:10.1038/nature14967

Mapping tree density at a global scale

T. W. Crowther[1], H. B. Glick[1], K. R. Covey[1], C. Bettigole[1], D. S. Maynard[1], S. M. Thomas[2], J. R. Smith[1], G. Hintler[1], M. C. Duguid[1], G. Amatulli[3], M.-N. Tuanmu[3], W. Jetz[1,3,4], C. Salas[5], C. Stam[6], D. Piotto[7], R. Tavani[8], S. Green[9,10], G. Bruce[9], S. J. Williams[11], S. K. Wiser[12], M. O. Huber[13], G. M. Hengeveld[14], G.-J. Nabuurs[14], E. Tikhonova[15], P. Borchardt[16], C.-F. Li[17], L. W. Powrie[18], M. Fischer[19,20], A. Hemp[21], J. Homeier[22], P. Cho[23], A. C. Vibrans[24], P. M. Umunay[1], S. L. Piao[25], C. W. Rowe[1], M. S. Ashton[1], P. R. Crane[1] & M. A. Bradford[1]

A study led by Yale University researchers has found that there are over 3 trillion trees on Earth - but they are disappearing at an alarming rate.

The study found that there are around 3.04 trillion trees on Earth, or around 422 for each person on the planet.

The number is a huge increase on the previous global estimate, which was just over 400 billion trees worldwide.

The study was based on on-the-ground data about the number of trees in more than 400,000 plots of forest from all continents except Antarctica.

Source - https://www.independent.co.uk/environment/how-many-trees-are-there-on-earth-10483553.html

"Statistics is the science of learning from data, and of measuring, controlling and **communicating uncertainty**."

We should become comfortable with the idea that the most interesting and useful results may change with new information (data) is gathered

Statistics: "The science of aiding decision-making with incomplete information."

Which one is a statistical question?

- How many students checked or used their cell phones during any class across Montreal universities?

- Have you checked your cell phone at the end of today's class?

Welcome to BIOL 322 (some initial thoughts on learning/teaching philosophy)

Learning is not a spectator sport. We do not learn much just sitting in classes listening to instructors & memorizing pre-packaged assignments.

We must talk about what we are learning, write about it, relate it to past experiences, and apply it to our daily lives. We must make what we learn part of ourselves.

- Chickering and Gamson

Let’s take a break – 1 minute

Statistics is a science in its own!

“Statistics is a science, not a branch of mathematics, but uses mathematical models as essential tools.”

- John Tukey

Statistics is a Mathematical Science
(not a branch of Mathematics)

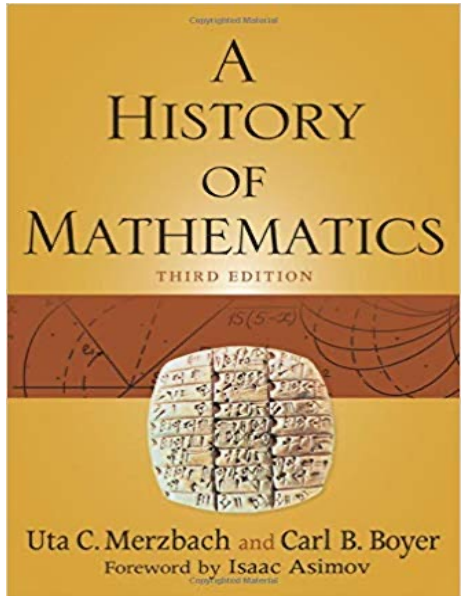
- We use the singular *is* and not the plural *are* to emphasize that statistics is a field of study, not just a "bunch" of methods.
- We use *mathematical* as an adjective because although statistics certainly makes use of much mathematics (another discipline), it is a separate discipline and not a branch of mathematics.
- We use the noun *science* because statistics is the science of gaining insight from data.

- From "Some Important Comparisons between Statistics and Mathematics, and Why Teachers Should Care" by Rossman, Chance, and Medina (2006).

Statistics is NOT a Branch of Mathematics

The book *A History of Mathematics* is the classic one-volume history of mathematics.

Statistics is not mentioned in the comprehensive index.



Statistical Thinking *versus* Mathematical Thinking

Mathematics is, by and large, a ***deterministic*** way of thinking and the way mathematics is taught in schools entrenches students into a deterministic way of viewing the quantitative world around them - *What is the size of our planet?*

Statistics is, by and large, a ***probabilistic*** or stochastic way of thinking (i.e., it considers uncertainty) - *What is the probability that it will rain tomorrow?*

Statistical Thinking *versus* Mathematical Thinking

Statistics is a separate discipline with its own unique ways of thinking and its own tools for approaching problems.

- J. Michael Shaughnessy, “Research on Students’ Understanding of Some Big Concepts in Statistics” (2006)

Statistics *versus* Data Science
(demystifying a trend)

"For statisticians, the entire data science trend seems a bit patronizing. No matter what your exact definition of data science is, it's going to sound pretty similar to the work that statisticians have been doing for decades."

- Nate Silver

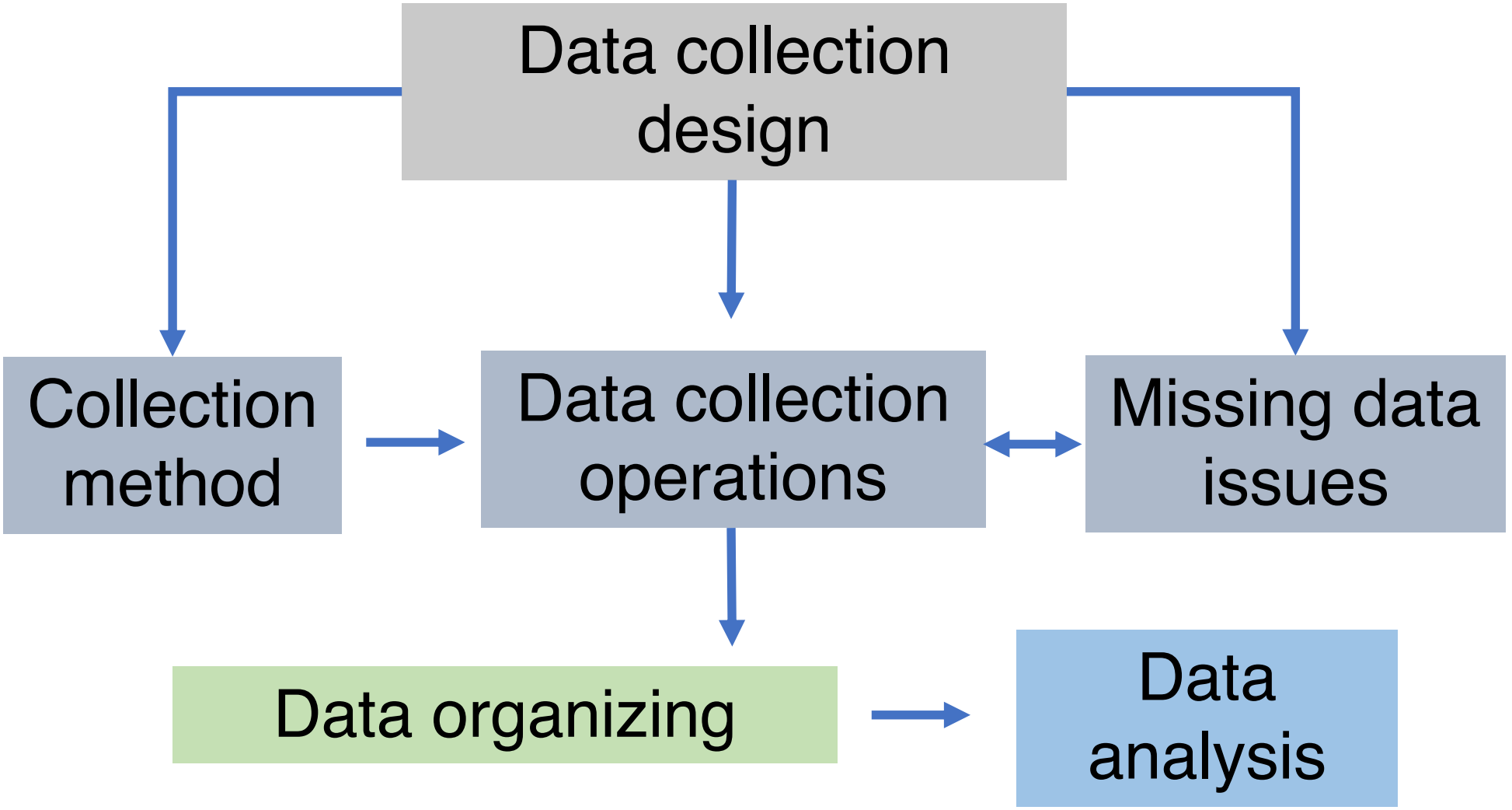
Roles of statistics

Statistics is a discipline that:

1) Designs data collection protocols (observational and experimental).
2) Summarizes information to aid understanding.
3) Draws conclusions from data.
4) Communicate uncertainty.
5) Estimates the present or predict the future.

- adapted from: http://www.scc.ms.unimelb.edu.au/whatisstatistics/

Roles of statistics: [1] - Design data collection and protocols (experimental and observational)



Roles of statistics: [2] - Summarize information to aid understanding

NCSS Data - [Example Data: Pain]

File Edit View Data Analysis Graphics Tools Window Help

Column Info | Kerlosin

	Drug	Patient	Time	Cov	Pain	C6	C7	C8	C9	C10	C11	C12	C13
1	Kerlosin	1	0.5	125	68								
2	Kerlosin	1	1	196	67								
3	Kerlosin	1	1.5	189	61								
4	Kerlosin	1	2	135	57								
5	Kerlosin	1	2.5	128	43								
6	Kerlosin	1	3	151	37								
7	Kerlosin	2	0.5	215	75								
8	Kerlosin	2	1	151	68								
9	Kerlosin	2	1.5	191	62								
10	Kerlosin	2	2	212	47								
11	Kerlosin	2	2.5	127	46								
12	Kerlosin	2	3	133	42								
13	Kerlosin	3	0.5	161	77								
14	Kerlosin	3	1	155	77								
15	Kerlosin	3	1.5	178	72								
16	Kerlosin	3	2	172	46								
17	Kerlosin	3	2.5	143	34								
18	Kerlosin	3	3	154	41								
19	Kerlosin	4	0.5	212	71								
20	Kerlosin	4	1	177	71								
21	Kerlosin	4	1.5	120	65								
22	Kerlosin	4	2	216	51								
23	Kerlosin	4	2.5	112	43								
24	Kerlosin	4	3	143	37								
25	Kerlosin	5	0.5	113	83								
26	Kerlosin	5	1	178	80								
27	Kerlosin	5	1.5	211	59								

Row: 1 Col: 1 100%

From raw (primary) data……to summaries

Roles of statistics: [3] & [4] - Draw conclusions from data & communicate uncertainty (estimate error)

Example: Voting polls in the news which make a claim about precision; example:

> "43% of the voting intention goes to the XXX party. The sample size was 1020; for a sample of this size the maximum margin of error is about 3%."
>
> *Do you know what that means?*

- Source - http://www.scc.ms.unimelb.edu.au/whatisstatistics/ssize.html

Roles of statistics: [3] & [4] - Draw conclusions from data & communicate uncertainty (estimate error)

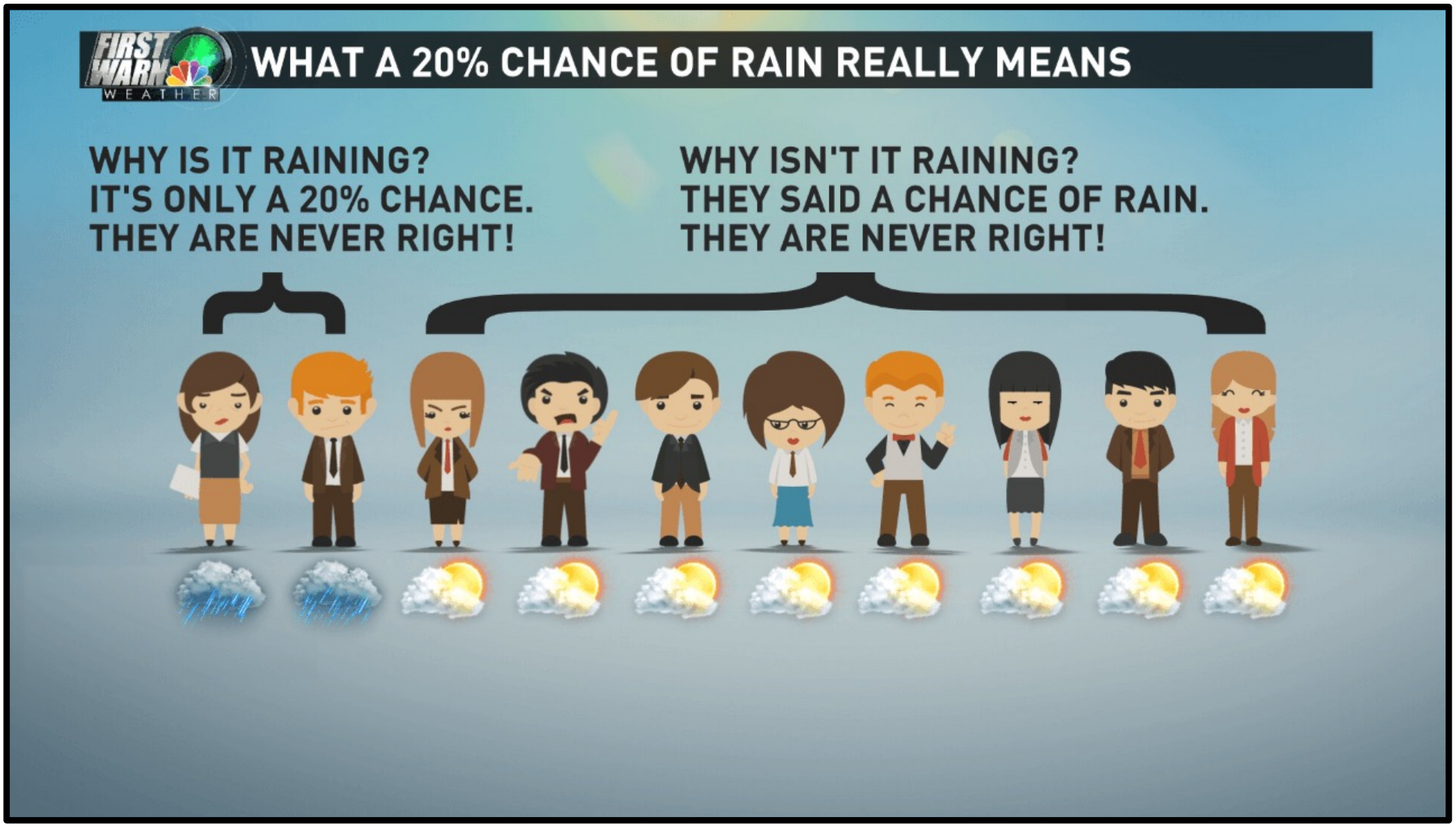
Example: Voting polls in the news which make a claim about precision; example:

"43% of the voting intention goes to the XXX party. The sample size was 1020; for a sample of this size the maximum margin of error is about 3%."

Do you know what that means? ("we're pretty sure the true value regarding voting intention for party XXX in the population is 43 ±3%, or somewhere between 40% and 46%")

- Source - http://www.scc.ms.unimelb.edu.au/whatisstatistics/ssize.html

Roles of statistics: [5] - Estimate the present or predict the future

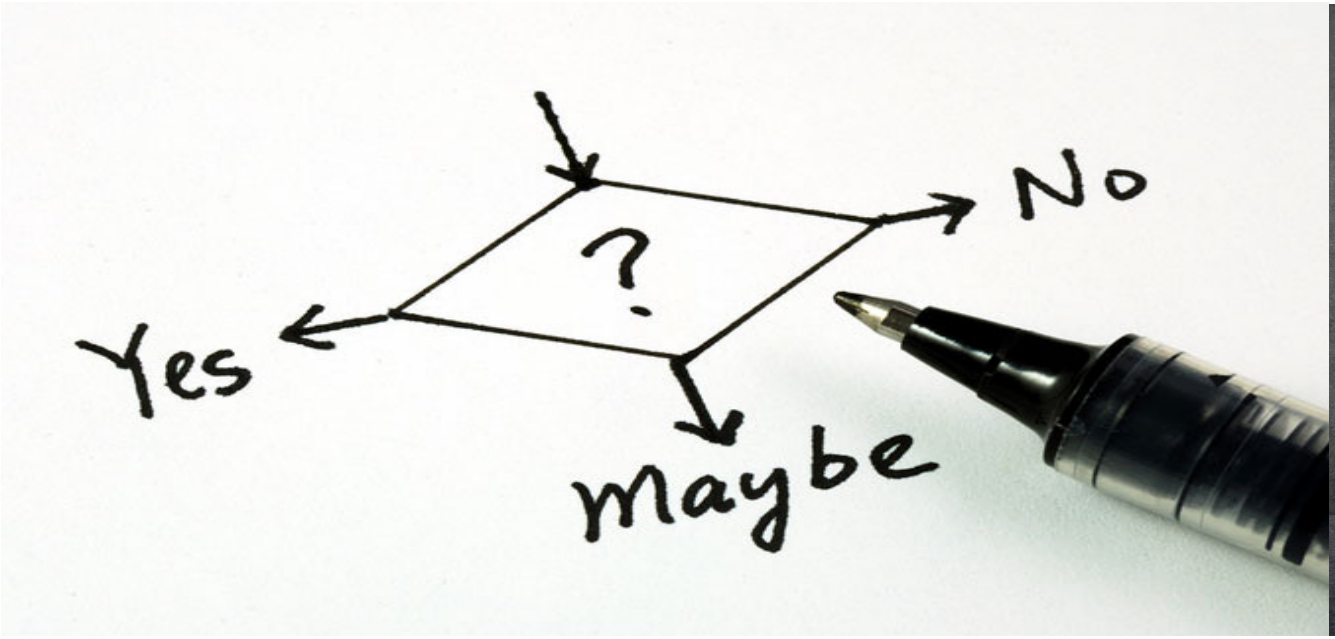


- Source: http://wxbrad.com/why-a-50-chance-of-rain-usually-means-a-100-chance-of-confusion/

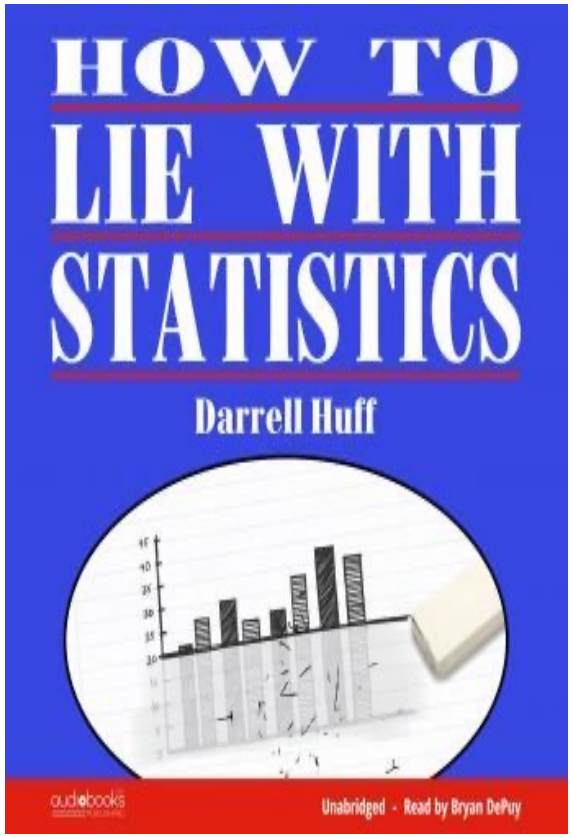
What is the role of statistics?

Convince you and others!

Statistics is key in decision-making processes because most decisions are made without complete knowledge (i.e., decisions always carry some level of uncertainty).



What is the role of statistics? **Convince you and others!**



HOW TO LIE WITH STATISTICS
(Huff, D. 1954)

There are three kinds of lies: lies, damned lies, and statistics.
—Disraeli

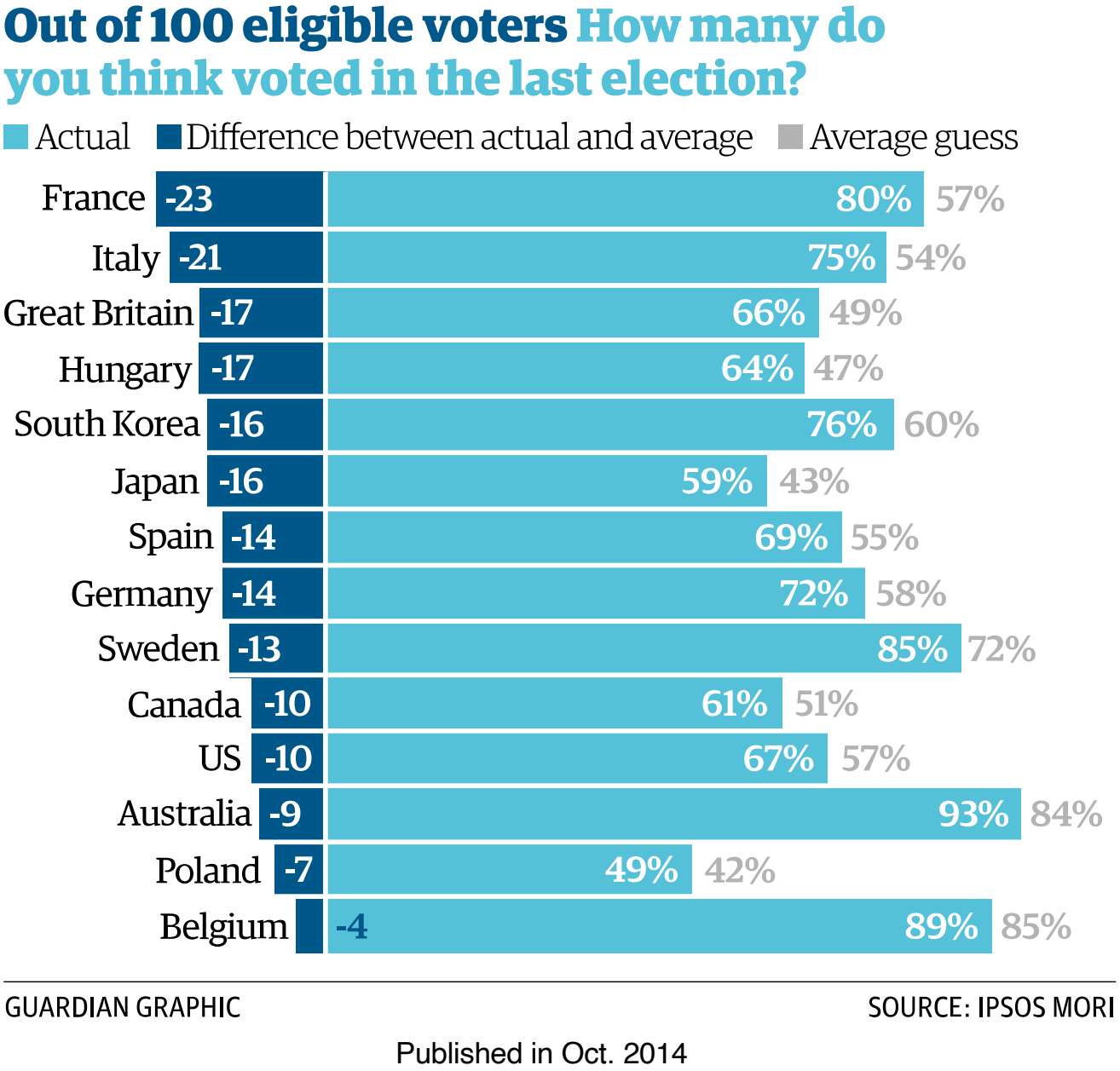
Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write.
—H. G. Wells

It ain't so much the things we don't know that get us in trouble. It's the things we know that ain't so.
—Artemus Ward

Round numbers are always false. —Samuel Johnson

I have a great subject [statistics] to write upon, but feel keenly my literary incapacity to make it easily intelligible without sacrificing accuracy and thoroughness.
—Sir Francis Galton

We often make decisions based on wrong guesses!



Published in Oct. 2014

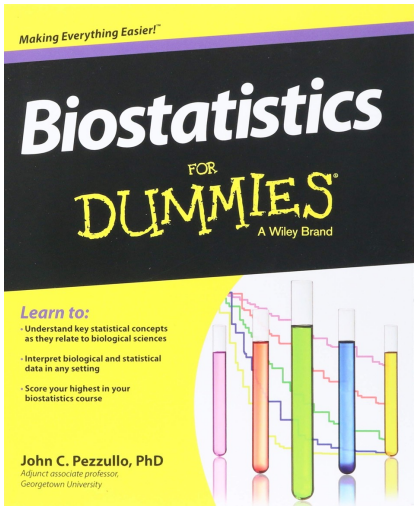
We often make decisions based on wrong guesses!

Probabilistic thinking is not always intuitive

The Monty Hall Problem (from "Let's make a deal"): In search of a new car, you pick a door, say 1. The game host then opens one of the other doors, say 2, to reveal a goat and offers to let you pick door 3 instead of door 1 if you want to. ***Would you switch to door 3 or keep door 1?***

Biostatistics is the branch of statistics that study data related to living organisms

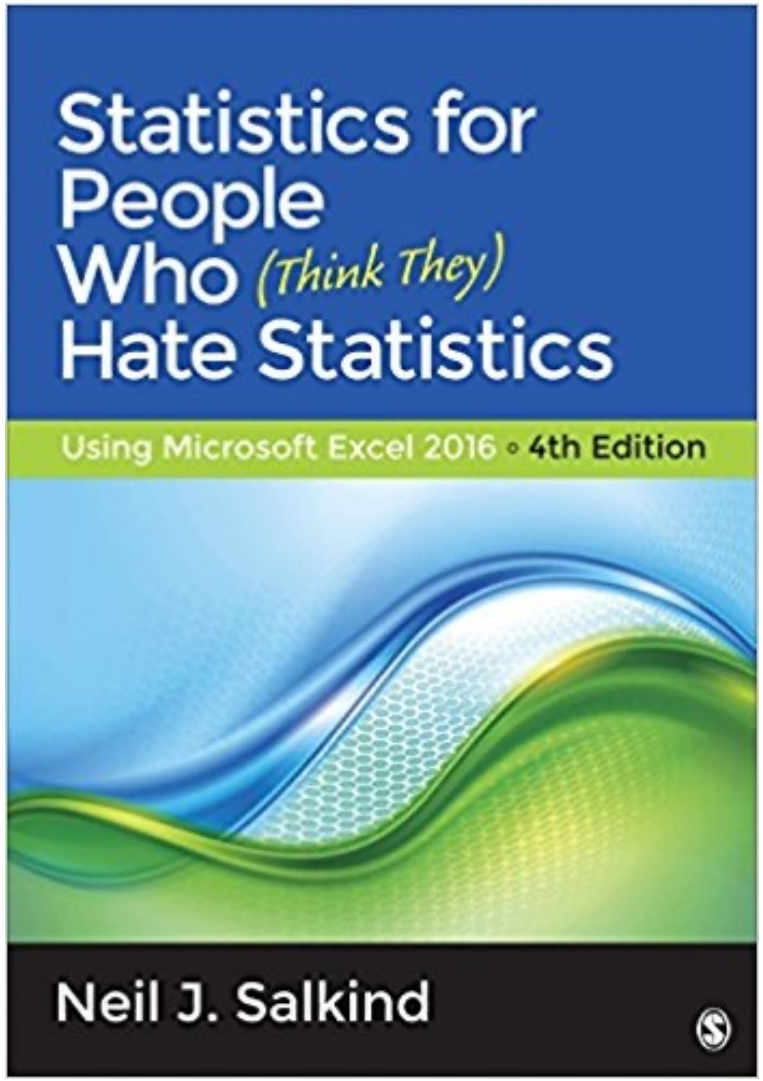
Biostatistics is an extremely active field of statistics



Let’s take a small break – 1 minute

The fear of statistics (bio or not)

The fear of statistics



Bottom line:

There is fear of statistics…but *there is no need to!*

The challenges in understanding numbers and numerical thinking

“Mathematics (and statistics) is not primarily a matter of plugging numbers into formulas and performing rote computations.

It is a way of questioning and thinking that may be unfamiliar to many of us but is available to almost all of us."

- John Allen Paulos, A Mathematician Reads the Newspaper.

The challenges in understanding statistics

Let me describe my goal & teaching philosophy with a single problem: "Pretend that you need to learn a new number system"

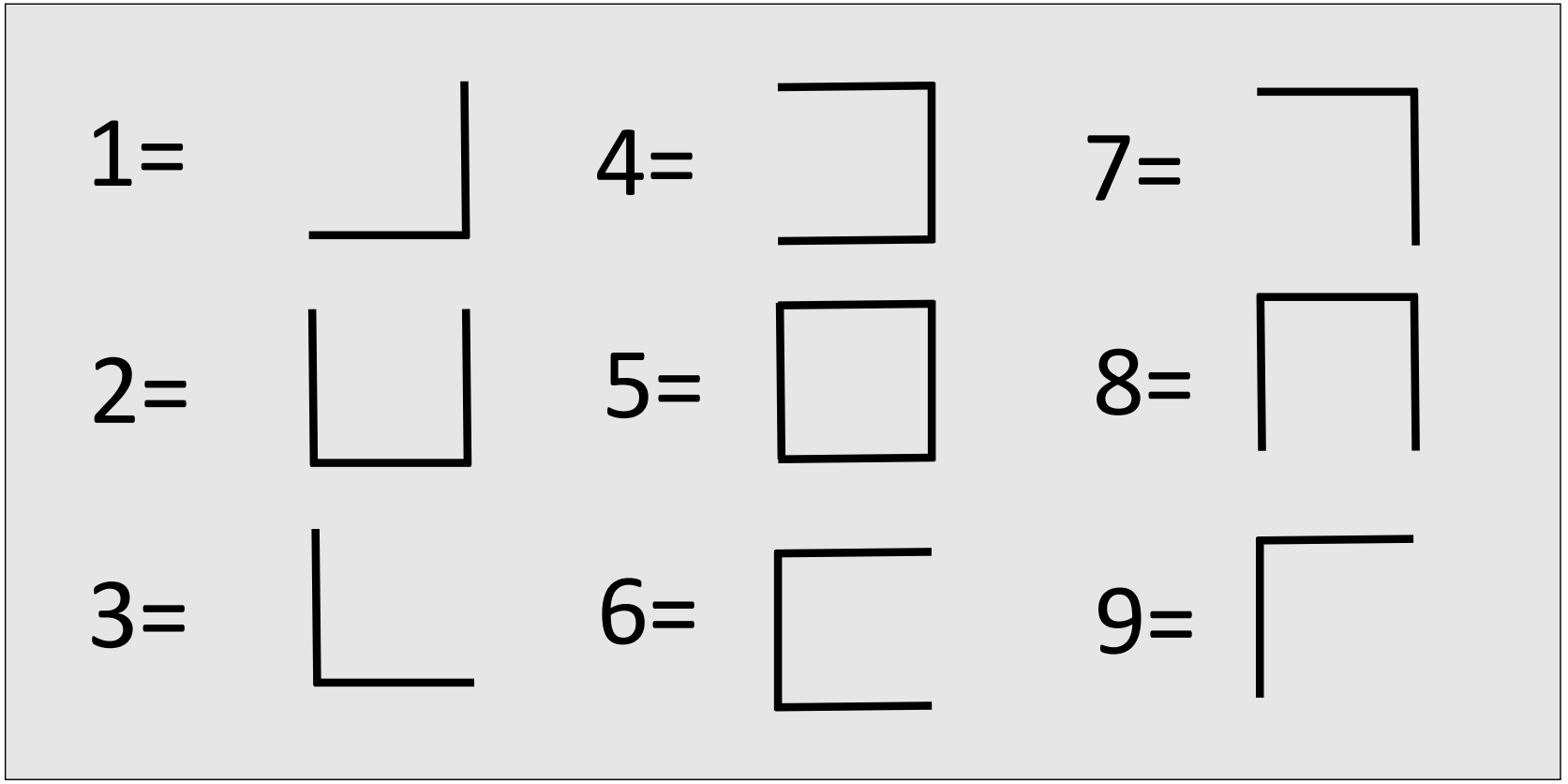


Image by P. Newbury

The challenges in understanding statistics

Let me describe my goal & teaching philosophy with a single problem: "Pretend that you need to learn a new number system"

What is this number?

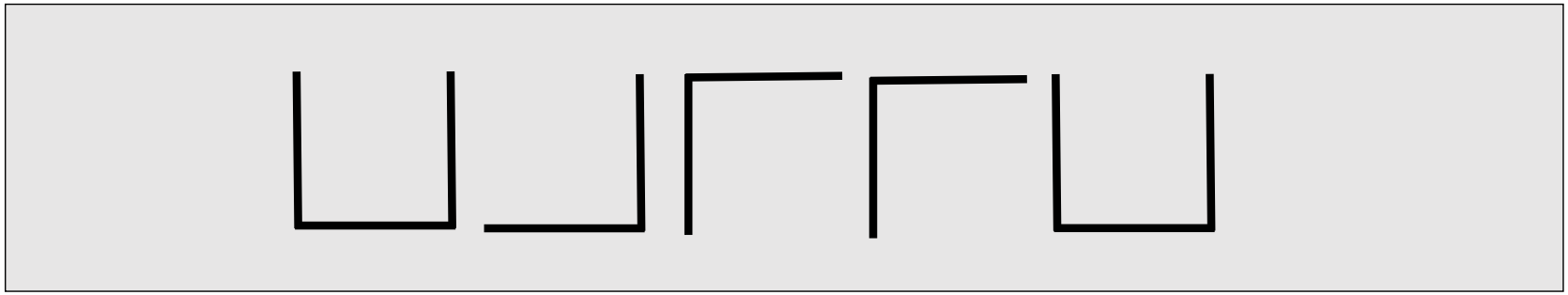
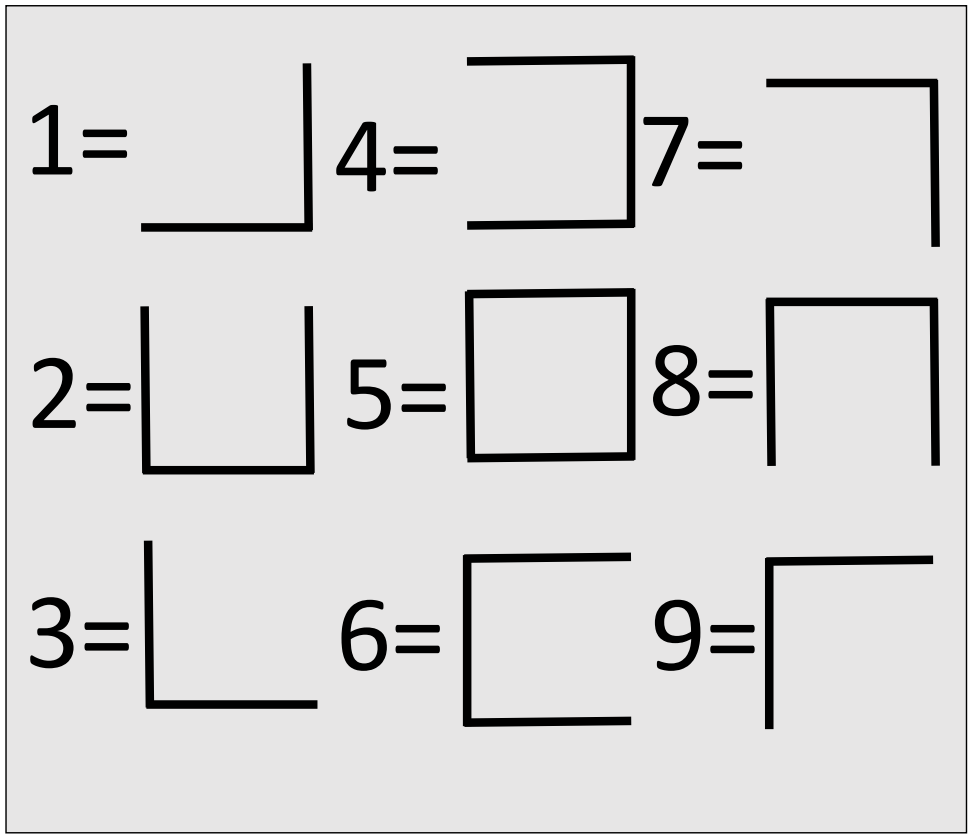


Image by P. Newbury

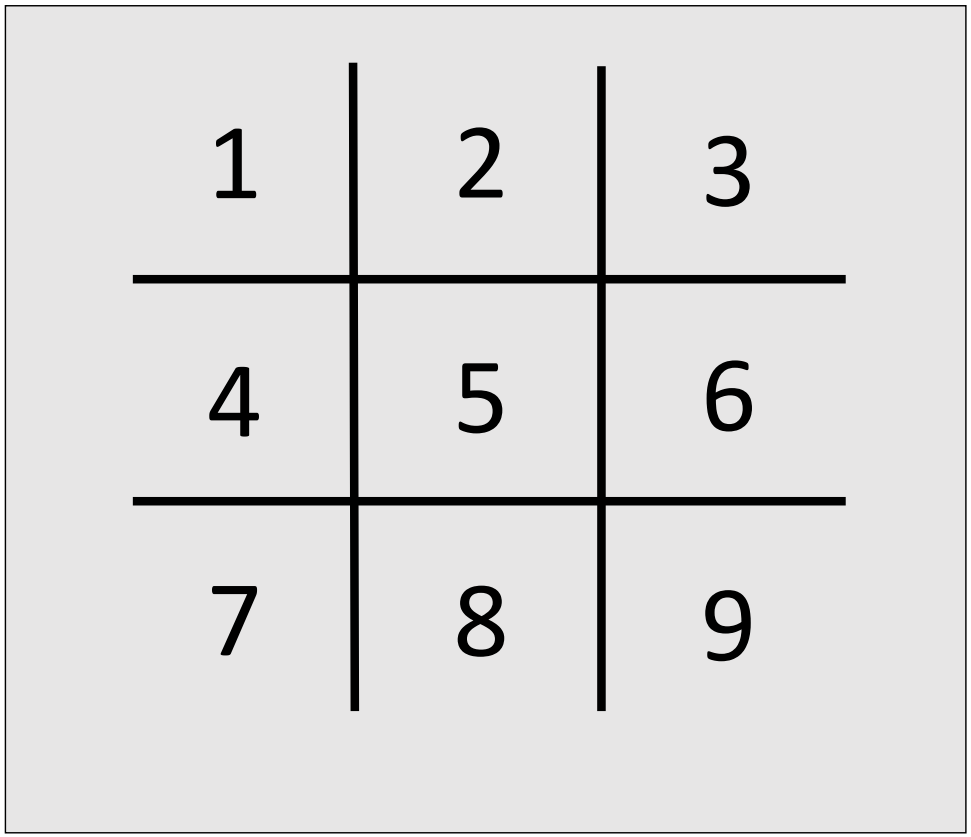
Teaching styles

Transmissionist



Unsupported content

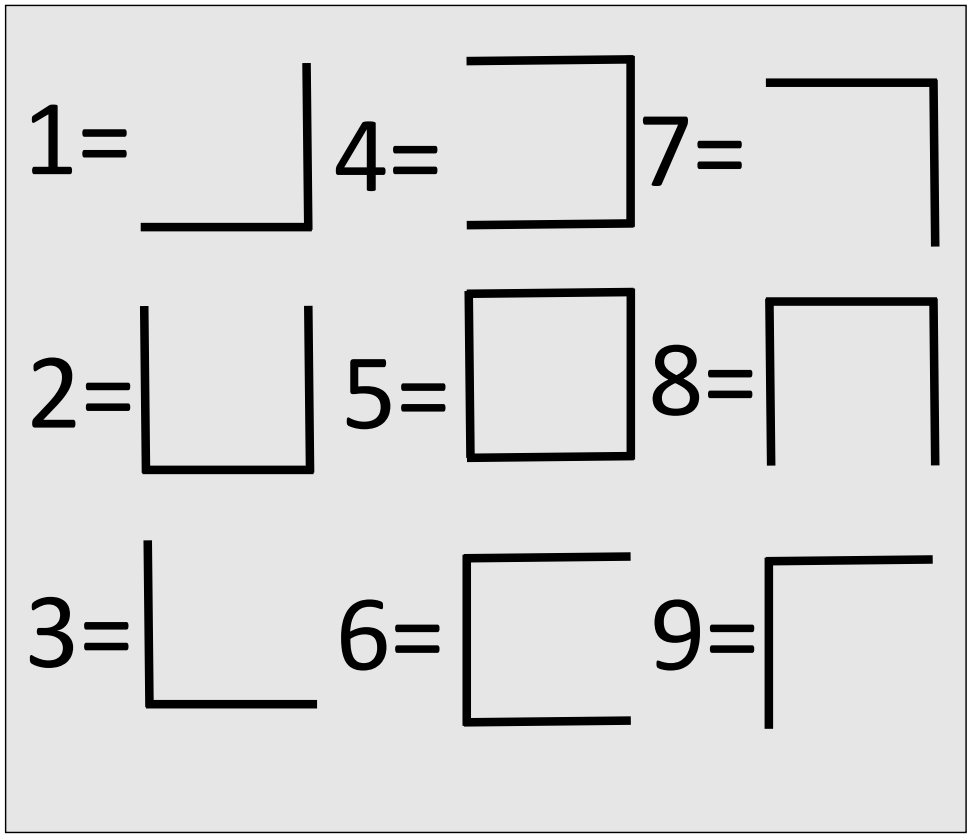
Constructivist



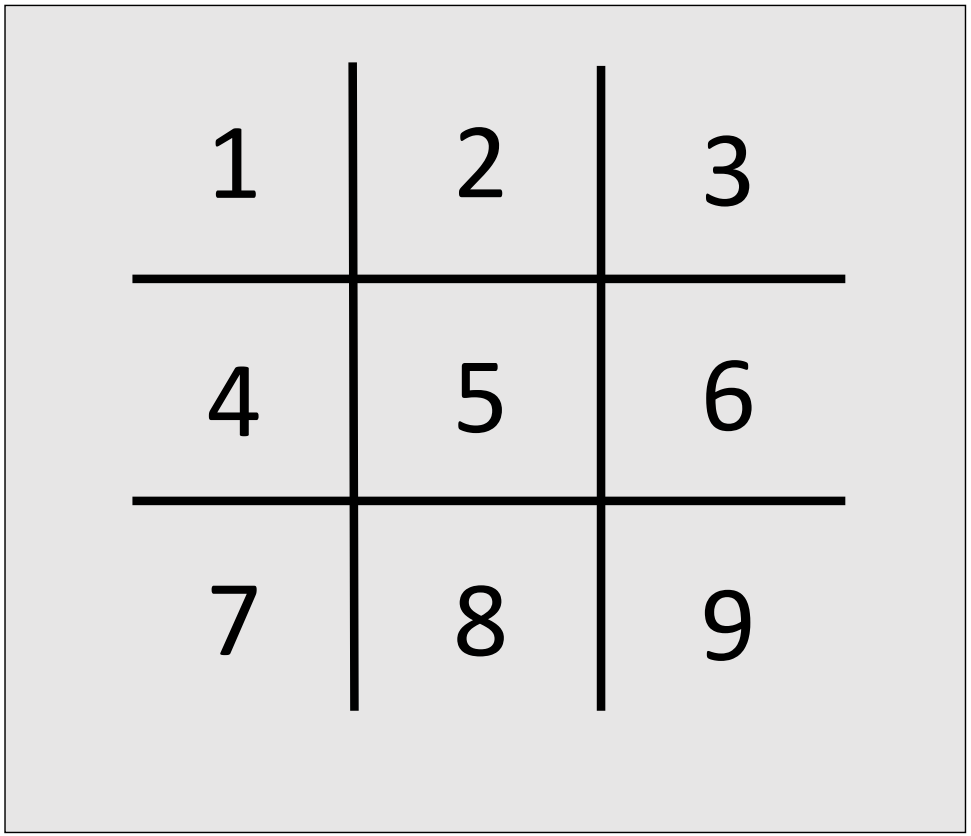
Built on familiar content

Teaching styles

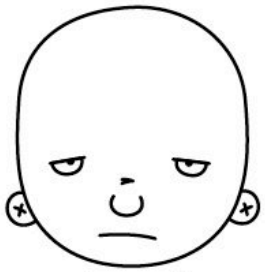
Transmissionist



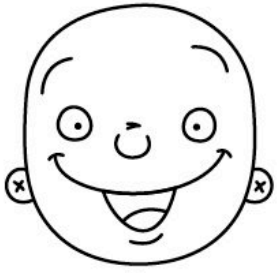
Constructivist



Unsupported content



Built on familiar content



The challenges in understanding statistics

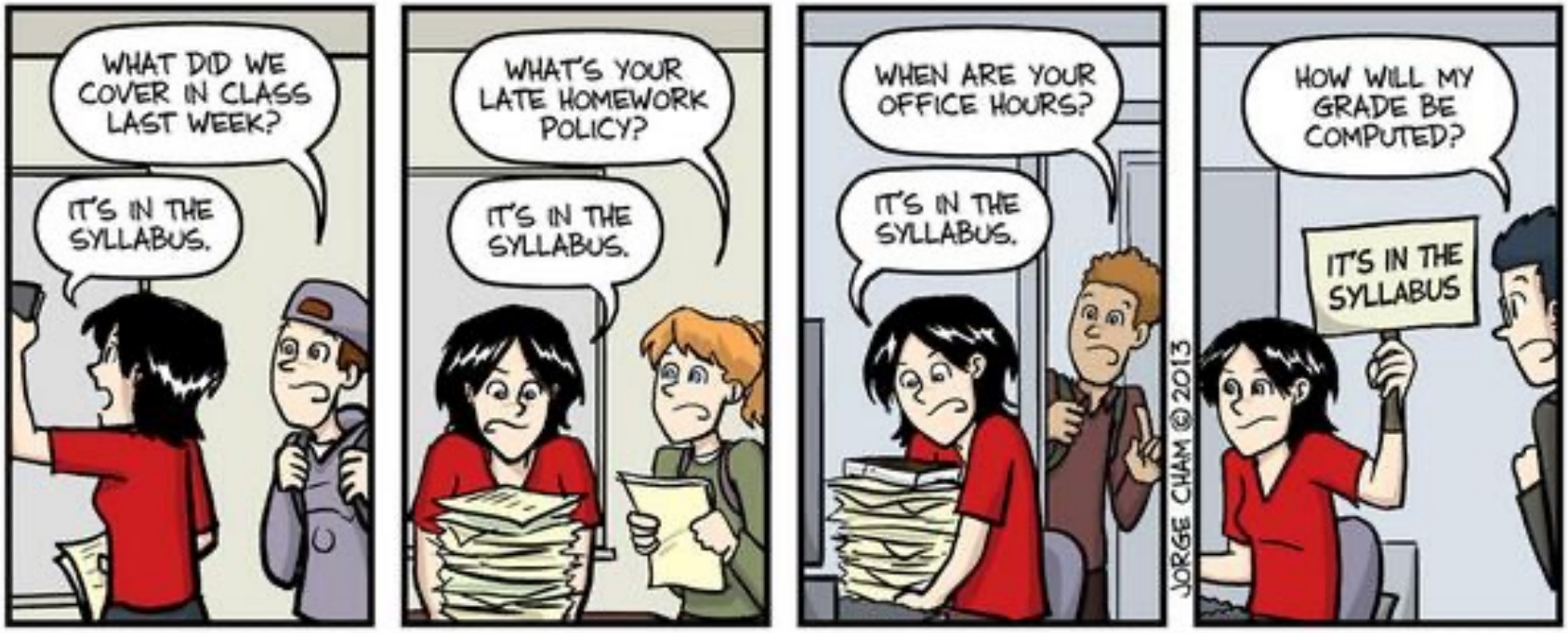
Imagination is our job

Let’s connect concepts

Let’s work together!

Communication, communication & communication

PLEASE READ THE SYLLABUS



But if it's not in there…let me know

Please use our Moodle FORUMs to ask questions: the answer to your question can help everyone!

What to study?

1) What do you need to study for assignments and exams?

ANSWER: You need to study everything listed under the lecture chapters in the WebBook. I provide a range of materials, including lecture slides (which can be downloaded from the WebBook), written text, figures, and videos.

2) Do exams cover detailed knowledge of equations and R?

ANSWER: They don't cover detailed knowledge but rather some basic commands that you will routinely use in your tutorials and reports. Additionally, I will ask for some calculations in exams and assignments that can be done by hand, along with a basic understanding of R commands. Examples of these will be provided in the quizzes.

3) Do we cover material in the lectures not covered in the lecture slides?

ANSWER: Yes! Lectures are designed to weave together concepts and knowledge in context. While lecture slides are comprehensive, they are not meant to replace a textbook. **In short, attending lectures and tutorials is essential for success in this course.**

Sharing any course material is an academic infraction

ISSUE: Sharing course materials, including but not limited to slides, tutorials, reports, assignments, recordings, and any other educational content provided in this course, **is strictly prohibited and considered an academic infraction**. This policy applies to:

[a] Sharing with current or future students of Concordia University, as well as students from any other institution.

[b] Posting or sharing any course materials on online digital platforms. Students who share course materials on these platforms can face penalties related to appropriate academic infractions (see below).

[Labour, ethical, moral, and educational concerns]: All materials distributed or created as part of this course are the intellectual property of the instructor and the institution. Instructors invest hundreds of hours, often including personal and family time, to prepare these materials for their students. These resources are intended exclusively for the educational use of students enrolled in this course.

Generative AI like ChatGPT - policy for BIOL322 Biostatistics

(adapted from Tristan Long, Wilfried Laurier University)

AI tools are NOT allowed for your assignments unless explicitly permitted; if so, it will be indicated in the assignment. If they are used, unless permitted in the assignment, it will be considered as an **academic misconduct.**

Why shouldn't you use generative AI for your assignments?

is not a reactionary stance rooted in fear of progress, but rather a thoughtful concern from an educator about what we are losing by becoming overly reliant on this technology.

First off, in a class like BIOL322, where your job is to learn (and learn from your mistakes), relying on an AI robs you of the human experience of growth and personal development..

How can you improve and grow as a student of biostatistics (or as a person), if what we instructors are reviewing does not represent your true efforts or thoughts?

Even though it's often easy to recognize its use in assignments, this may create a hostile dynamics between instructors and students. That said, we reserve the right that, if detected, it will be considered as an **academic misconduct.**

Why drop-in hours?
(applies to any course)

Seek clarification on topics that were unclear in lectures or readings.

Ask questions about assignments to ensure understanding.

Share and exchange ideas on how to improve the course experience for yourself and others.

Build a rapport with the instructor, which can help in finding an honors thesis advisor and obtaining strong letters of recommendation.

We should all be cordial

Instructors enjoy to be greeted cordially; for example:

Hello Pedro
Hello Dr. Peres-Neto; or Hello Prof. Peres-Neto
Hello could be replaced by Hi or Dear depending on the occasion.

Try to avoid being impersonal:

Hello,
Hi,

Hello sir/Hi sir - “If you forgot your instructor’s name, then please look over the course syllabus.”

Thank you ☺

Some closing thoughts for today

“Understand. Don't memorize. Learn principles, not formulas.”

“Learning is underrated. Grades are overrated”

Richard Feynman

Next class - Sampling

We biologists are relatively small

From Chris Lortie