CS109: Probability for Computer Scientists

Jerry Cain
March 28, 2022

Quick slide reference

Today's discussion thread: https://edstem.org/us/courses/21301/discussion/1324037

Live Lecture!

- Gates B01, with overflow rooms until attendance drops off
- Lecture also broadcast live during 3:15 – 4:35 via Canvas
- Perfectly acceptable if you need to watch lecture videos later or prefer to watch from the comfort of your dorm room

Ask your questions in class and on Ed

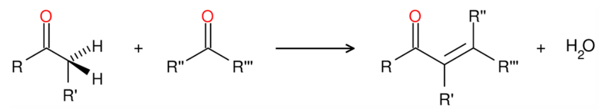
- Students in lecture are encouraged to interrupt me, ask questions, or even request I explain something a second time. Don't be shy!
- Ed can be used for questions, too (e.g., questions that arise while watching recordings)
- Persistent copy: Teaching staff and I can answer questions ***during*** and ***after*** lecture

If you were enrolled in the course as of Sunday night, you're already in the Ed forum.
Today's discussion thread: https://edstem.org/us/courses/21301/discussion/1324037

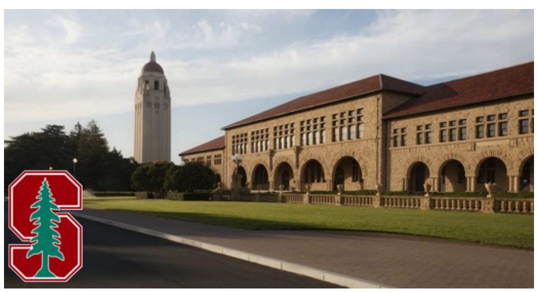
Welcome to CS109!

Jerry Cain

I went here from 1987 through 1991 and majored in chemistry.

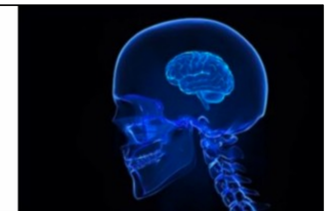


Then I came here for a PhD in chem, switched to CS



Received MSCS 1998
Lecturer: 26 years today!
Speculation: Have taught more classes than other lecturers.

<u>My interests over time</u>

Chemistry and Physics	
STEM Education	
Make Sense Need Nods	

Why Jerry likes probability

- I majored in chemistry, and my undergraduate research was rooted in surface science and ***statistical*** mechanics.

- When I switched to CS as a grad student here, I focused on CS theory and all the beautiful mathematics that comes with it. Math is brain food and nourishes the soul.

- Probability has revived parts of AI and information theory that were thought to be borderline dead when I was getting my MSCS degree here.

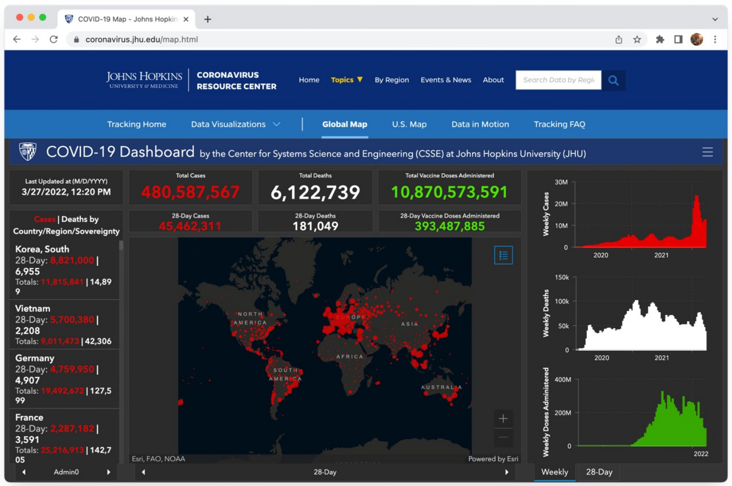
1974

1996

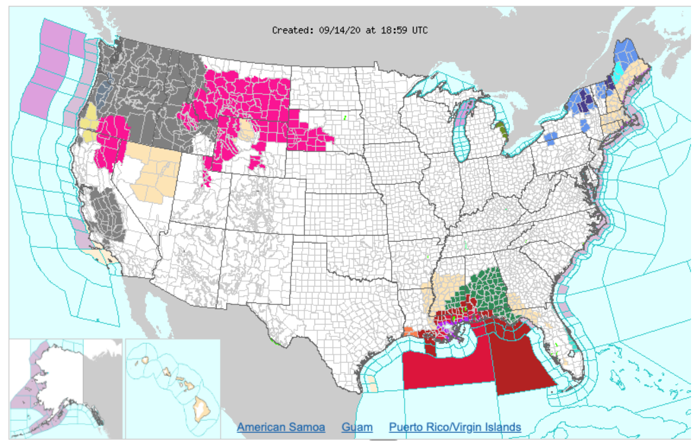
$$PV = \frac{1}{3} N m v_{\text{rms}}^2. \quad f(v) = 4\pi \left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} v^2 e^{-\frac{mv^2}{2kT}} \quad v_{\text{rms}}^2 = \int_0^\infty v^2 f(v)\, dv = 4\pi \left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} \int_0^\infty v^4 e^{-\frac{mv^2}{2kT}}\, dv$$

What makes this quarter important

We are seeing a huge surge in **statistics, predictions, and probabilistic models** shared through global news, governing bodies, and social media.



Global cases of COVID-19
as of March 27, 2022 (JHU)
https://coronavirus.jhu.edu/map.html



National Weather Service Alerts
https://www.weather.gov/

FiveThirtyEight 2020

The New York Times 2020

US Politics In Review
https://fivethirtyeight.com/
https://www.nytimes.com/

What makes this quarter important

We are seeing a huge surge in **statistics, predictions, and probabilistic models** shared through global news, governing bodies, and social media.

The challenge of delivering Stanford-class education reflects our university's commitment to fostering a **diverse body of students**.

What makes this quarter important

We are seeing a huge surge in **statistics, predictions, and probabilistic models** shared through global news, governing bodies, and social media.

The challenge of delivering Stanford-class education reflects our university's commitment to fostering a **diverse body of students**.

The **technological and social innovation** we develop during this time will strongly impact how we solve problems that impacts the **lives of countless people across the globe**.

What makes this quarter important

We are seeing a huge surge in **statistics, predictions, and probabilistic models** shared through global news, governing bodies, and social media.

The challenge of delivering Stanford-class education reflects our university's commitment to fostering a **diverse body of students**.

The **technological and social innovation** we develop during this time will strongly impact how we solve problems that impacts the **lives of countless people across the globe**.

Our teaching goals (at a minimum)

- To teach how probability applies to real life problems that really matter
- To help you foster and maintain human connection throughout the course

The CS109 Course Assistants

Course Mechanics

Course website

https://cs109.stanford.edu

And this URL will live on for the next several years, even after this quarter is over:

https://web.stanford.edu/class/archive/cs/cs109/cs109.1226

Prerequisites

CS106B/X

Programming
Recursion
Hash tables
Binary trees

MATH 51/CME 100

Multivariate differentiation
Multivariate integration
Basic facility with linear algebra (vectors)

CS103
(co-requisite OK)

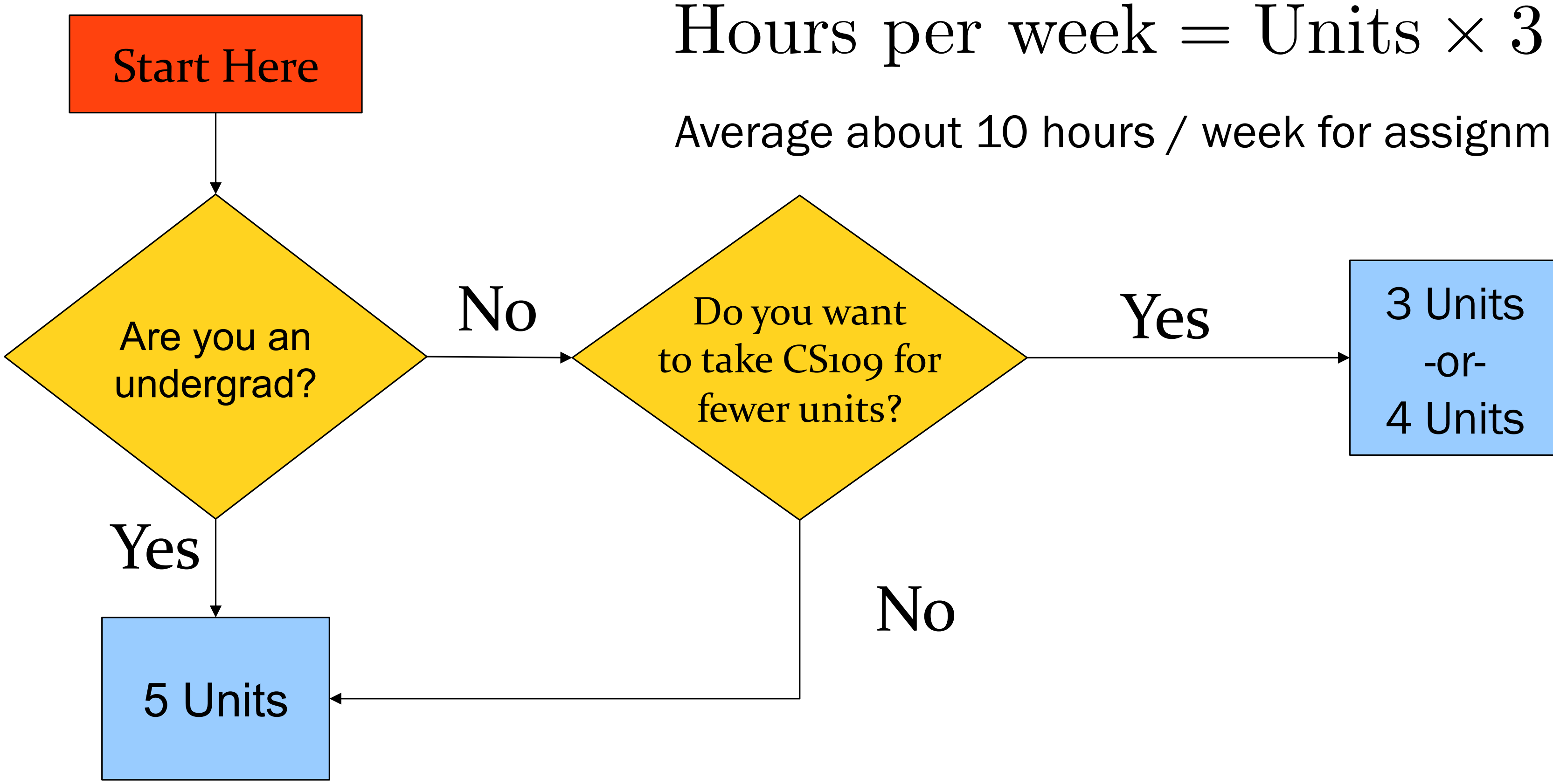
Proofs (induction)
Set theory
Mathematical maturity

Important!

How many units should I take?

$$\text{Hours per week} = \text{Units} \times 3$$

Average about 10 hours / week for assignments



Companion class: CS109A

- CS109A is an extra 1-unit "Pathfinders" or "ACE" section with additional support, practice, and instruction
- Meets for an additional weekly section and has additional review sessions
- Admission by application — see course website for details
- CS109A meetings are Tuesday evenings from 6:45 – 8:45pm
- Interested? Attend a special introduction this Tuesday at 6:45pm. This first meeting will be held online.

Ian Tullis

Course components

54%	6 Problem Sets
20%	Two take-home quizzes
20%	In-class Final Exam
6%	Section participation

Course components

54% **6 Problem Sets**

20% Two take-home quizzes

20% In-class Final Exam

6% Section participation

LaTeX

Written portion
- LaTeX for powerful typesetting
- Tutorial on CS109 website

python

Coding portion in Python
- Review session #1 on Monday 04/04 (time TBD)

Late policy
- Submit by deadline: **on-time bonus (~1%)**
- Grace period: **1 class day:**
  Miss the deadline? No on-time bonus. 😟
- Pre-granted grace period for all assignments.

Course components

54%	6 Problem Sets	
20%	**Two take-home quizzes**	
20%	In-class Final Exam	
6%	Section participation	

- 3-5 hours of individual work
  + typesetting for anyone on top of material
- Open-note, open-book, open-web
- 46-hour take-home window (between Wed/Fri lectures)
  - Week 4: Wed 4/20 – Fri 4/22
  - Week 7: Wed 5/11 – Fri 5/13

Course components

54%	6 Problem Sets
20%	Two take-home quizzes
20%	**In-class Final Exam**
6%	Section participation

- Scheduled for Tuesday, June 7th from 3:30pm until 6:30pm
- Closed-note, closed-book, traditional pencil and paper
- We'll supply a comprehensive list of all relevant formulas ahead of time
- No alternative final offered

Course components

54%	6 Problem Sets
20%	Two take-home quizzes
20%	In-class Final Exam
6%	**Section participation**

- Section meets Tuesdays and Wednesdays (times to be released soon)
- Sections start Week 2

More info about sections in coming days!

CS109 Contest

- Announced mid-quarter
- Boosts final course grades <u>after</u> letter grade buckets have been determined

Your baseline is CS109, and the sky is the limit.

Previous winning submissions:
- Recidivism Risk: Algorithmic Prediction and Racial Bias
- A Better Way to Reform the Electoral College
- Monte Carlo Tree Search for Tic Tac Toe
- COVID's impact on Student Interest in Post-Secondary School Education

Stanford Honor Code

Permitted

- Talk to the course staff
- Talk with classmates (cite collaboration)
- Look up general material online

Cite all references aside from instructors, staff, lecture slides/notes, course reader, and the optional Ross textbook.

NOT permitted:

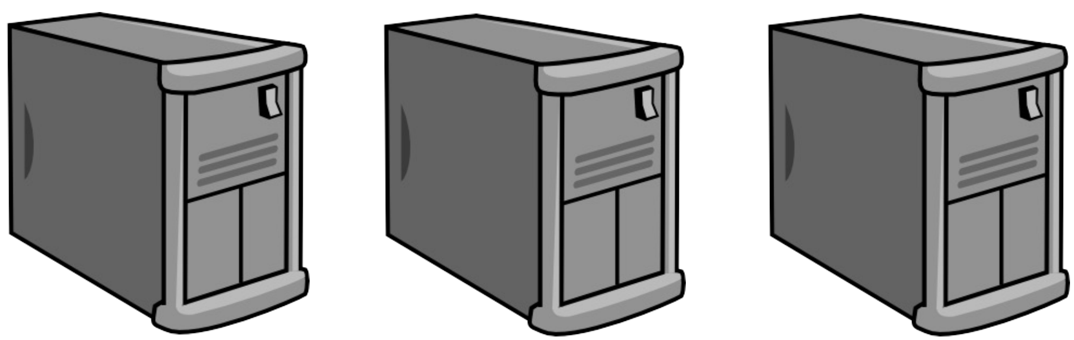
- Copy answers:
  - from classmates
  - from former students
  - from previous quarters
- Copy answers from the internet
  Ask for answers on the internet
  - Besides, these are usually incorrect

Why you should take CS109

Traditional View of Probability

CS view of probability

`http://www.site.com`



Machine Learning

= Machine (compute power)
+ Probability
+ Data

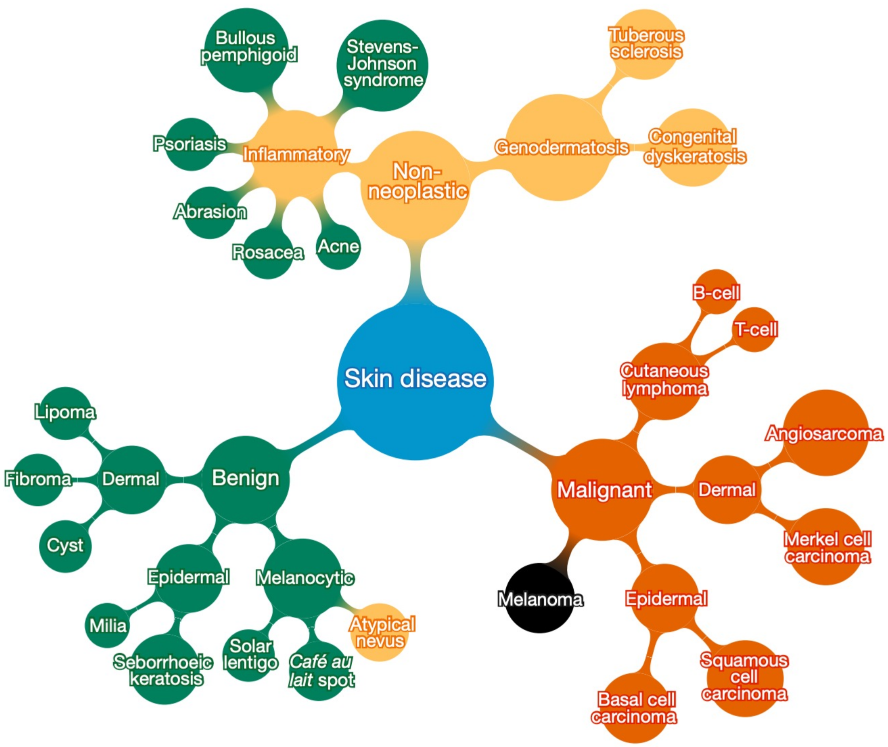
Machine Learning Algorithm

Data → Build a *probabilistic model* → Do one thing

Classification

Chihuahua or muffin?

Classification: Where is this useful?

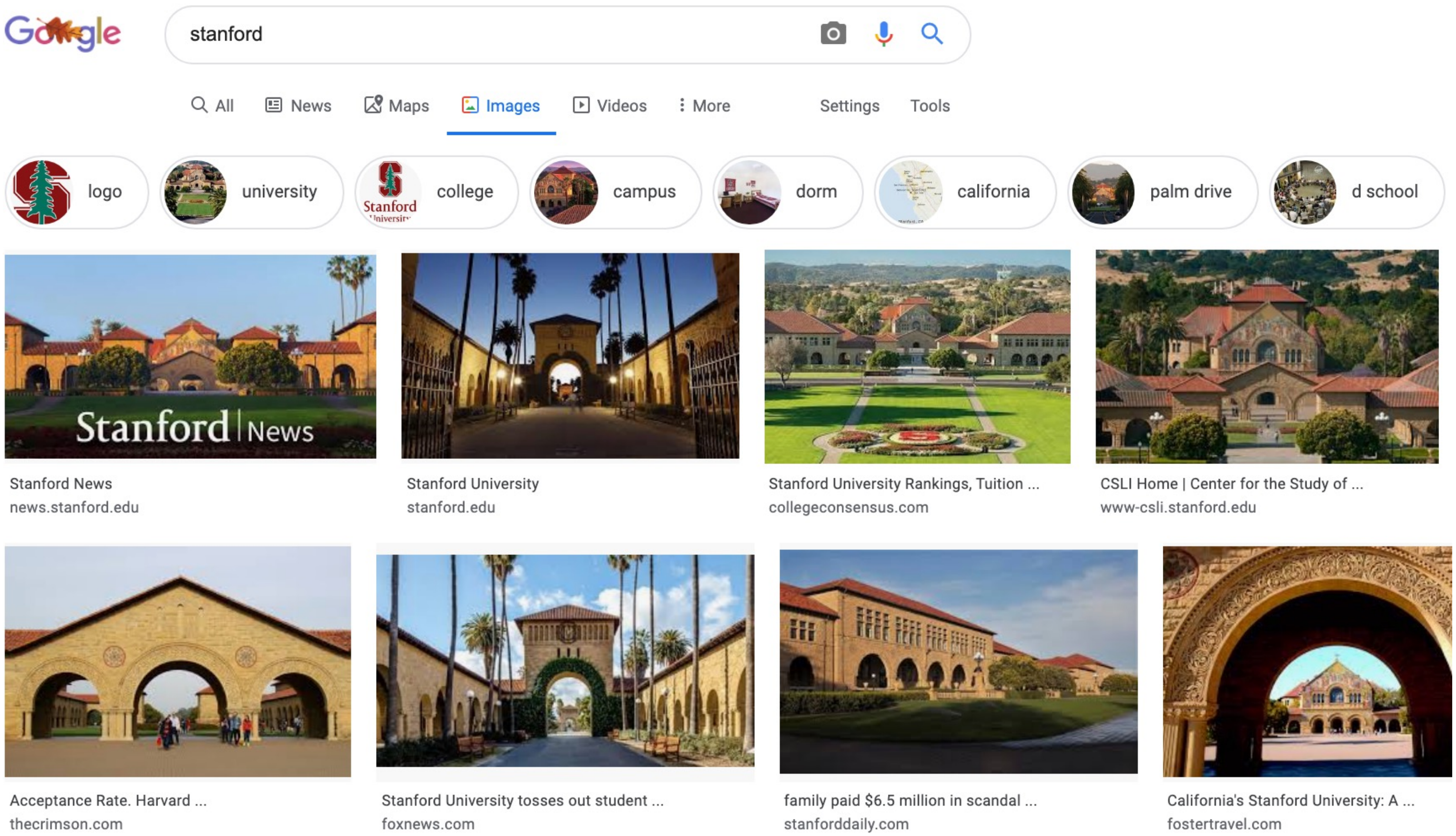


A machine learning algorithm performs **better than** the best dermatologists.

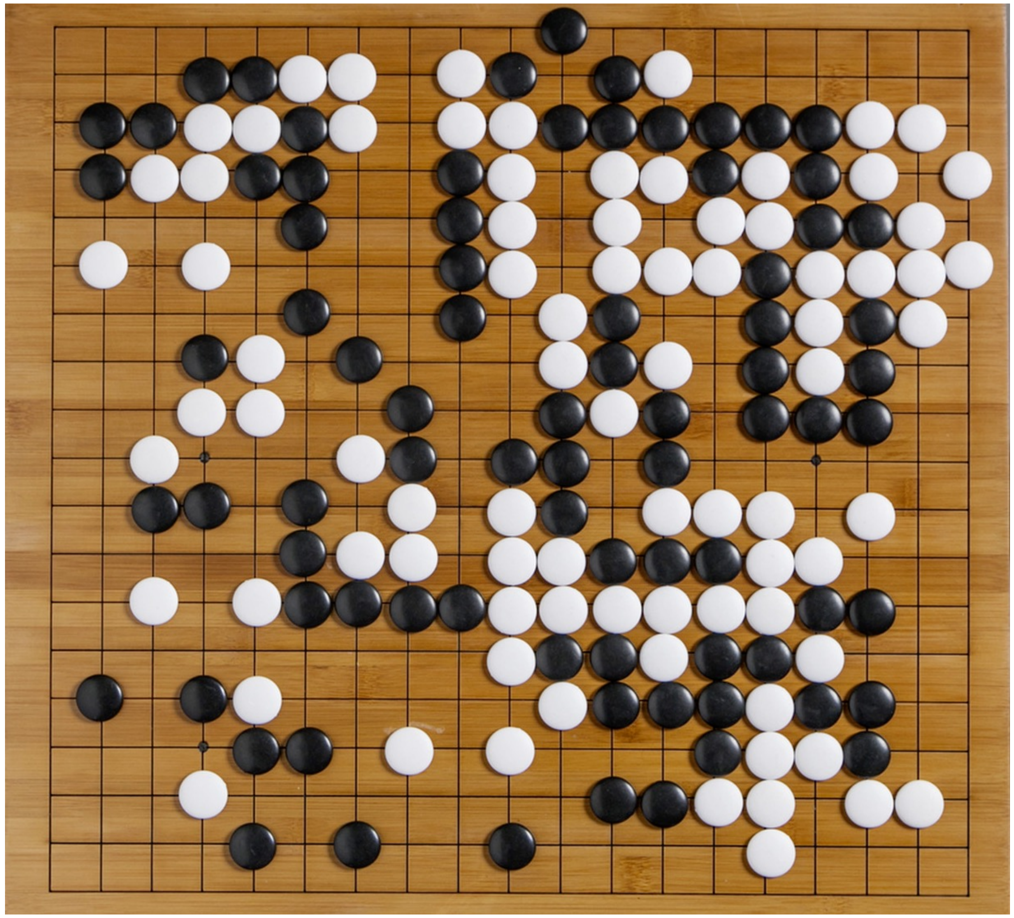
Developed in 2017 at Stanford.

Esteva, Andre, et al. "Dermatologist-level classification of skin cancer with deep neural networks." *Nature* 542.7639 (2017): 115-118.

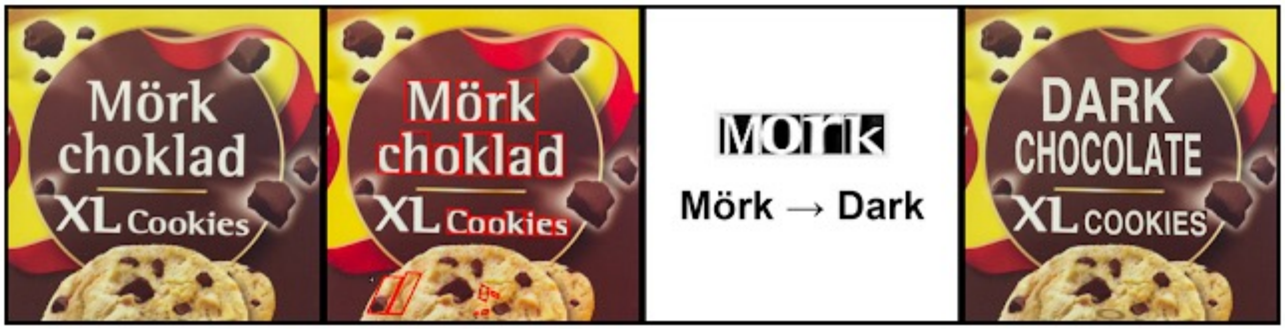
Classification: Image tagging



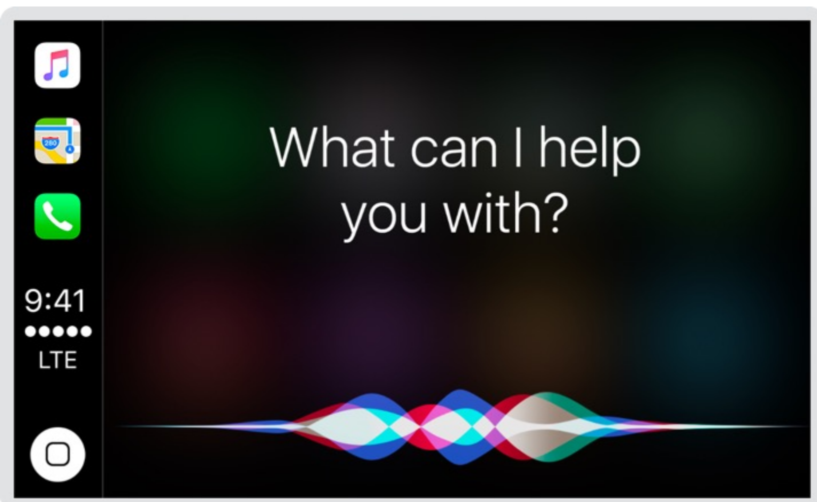
Decision-making: The last remaining board game



Natural language processing



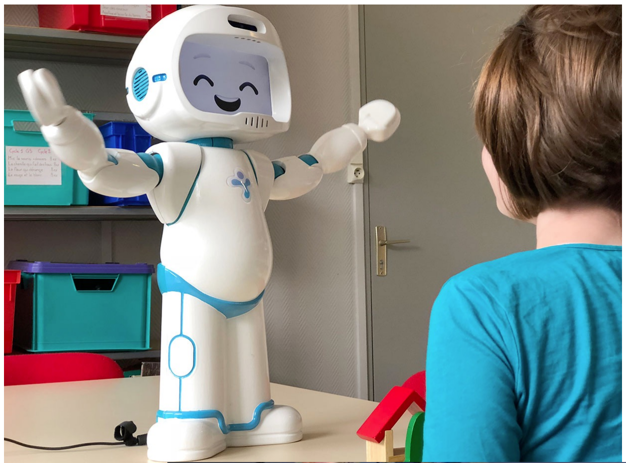
Augmented reality machine translation on Google Translate

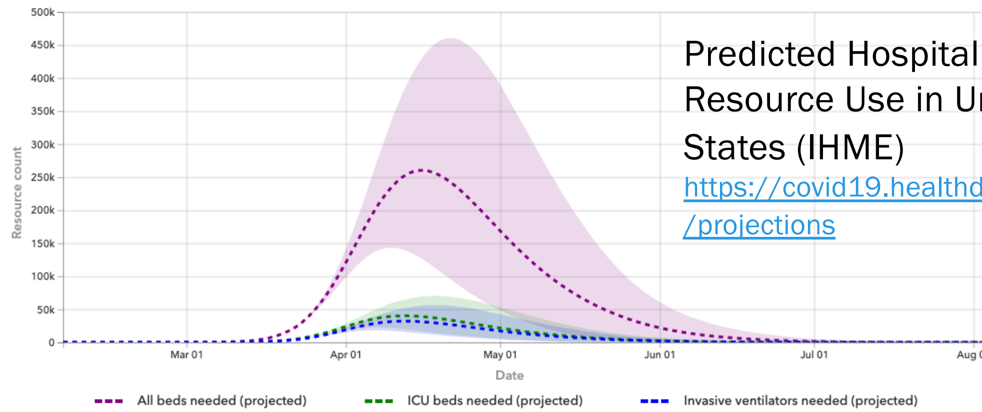


Voice assistants:
voice to text to answer

Probability is *more* than just machine learning.

Probability and medicine



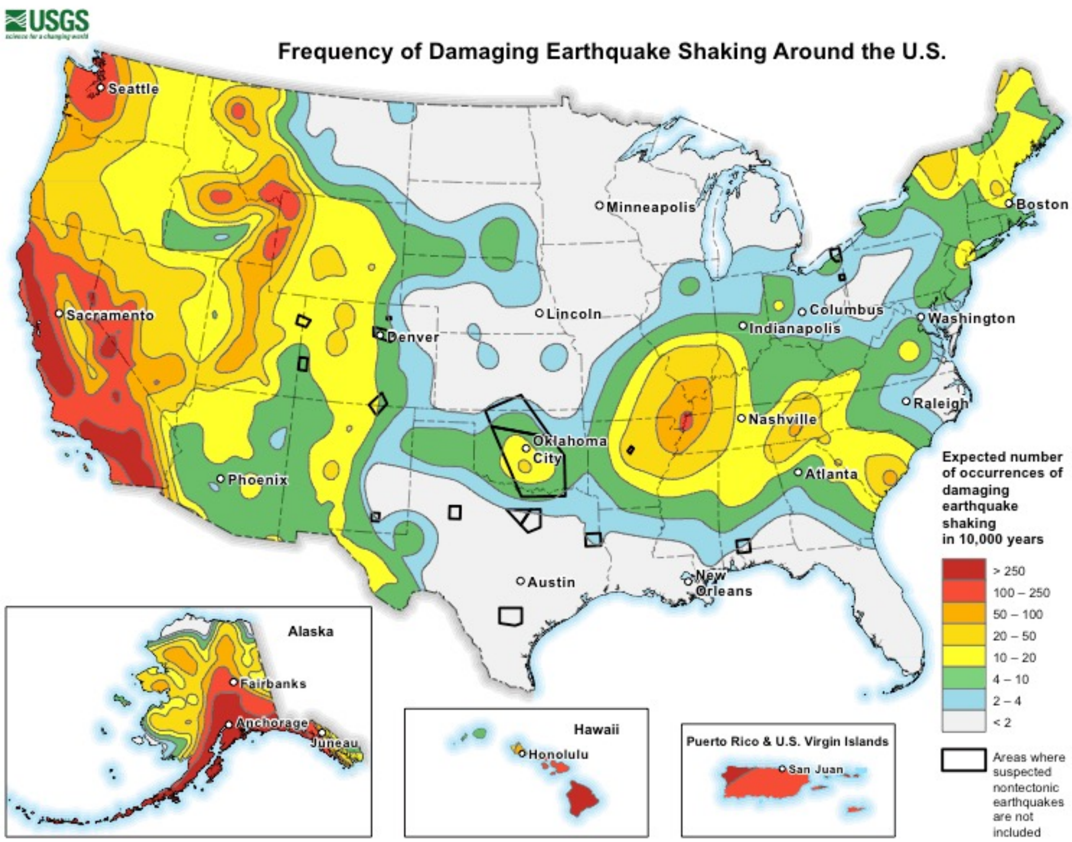


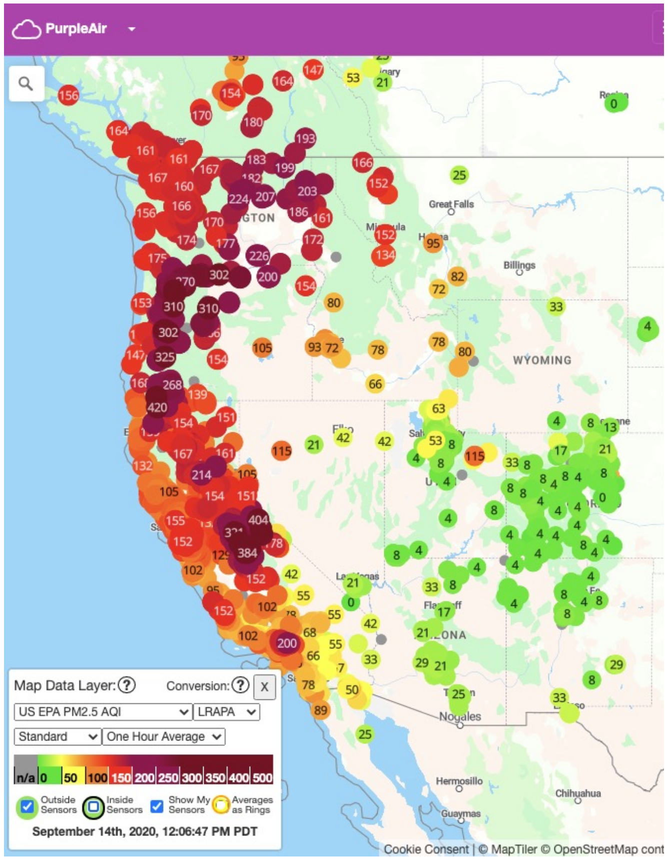
Predicted Hospital Resource Use in United States (IHME)
https://covid19.healthdata.org/projections

How do COVID-19 testing rates in a region correlate with the actual spread of the disease?

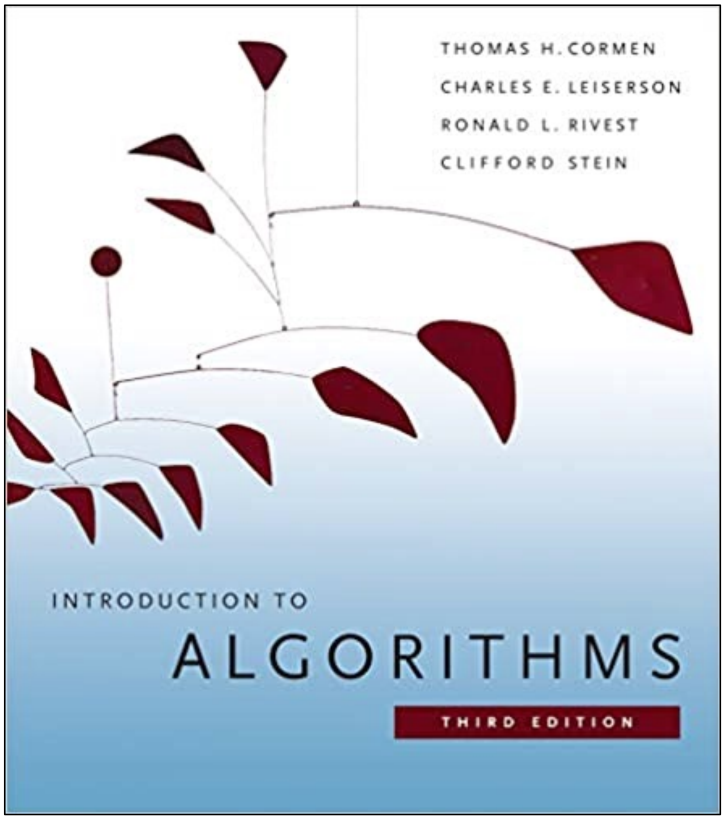
Probability and art

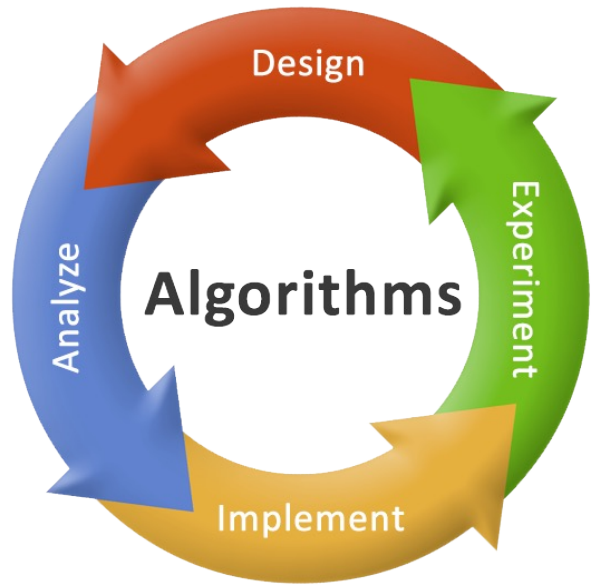
Probability and climate



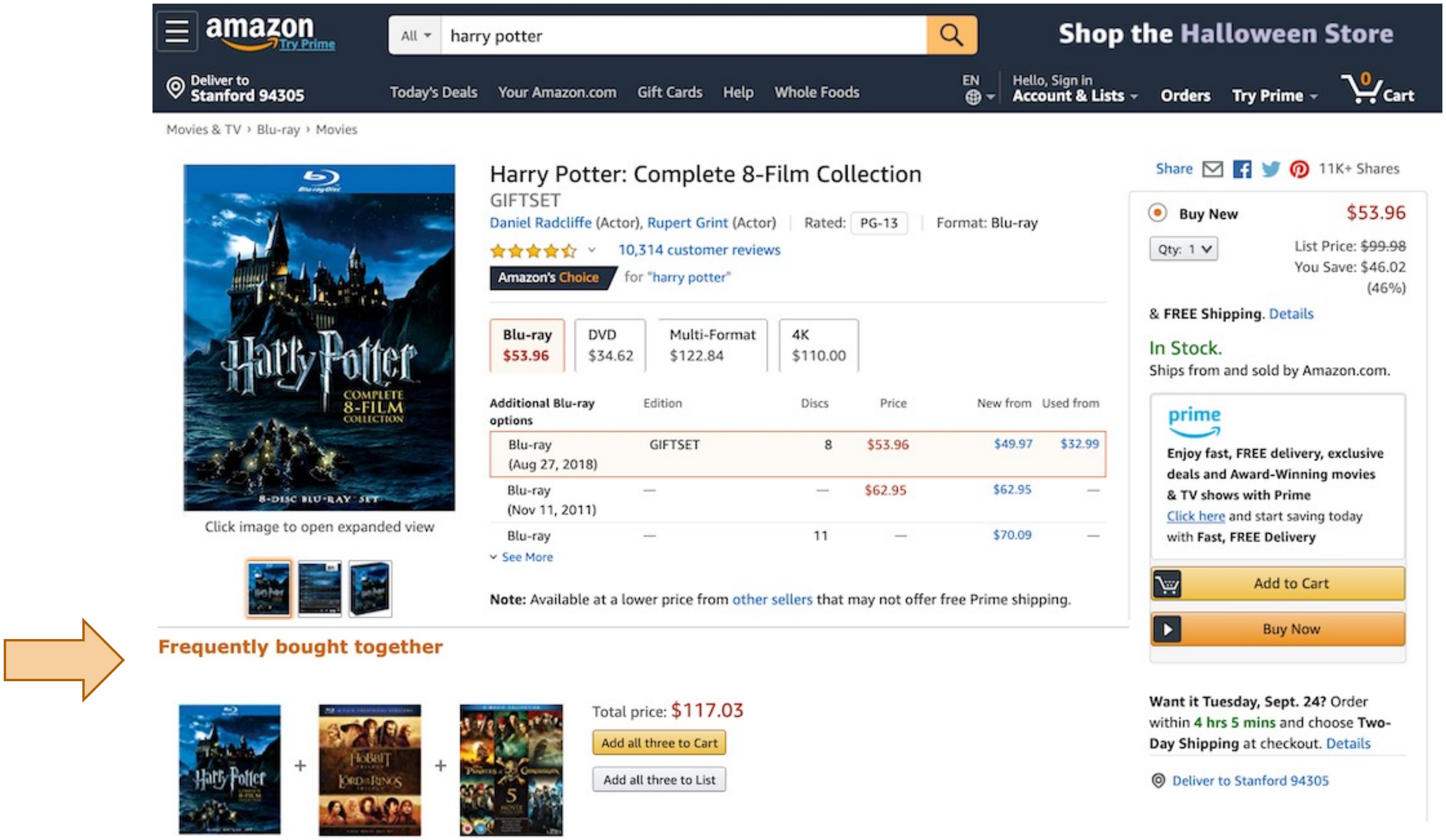


Probabilistic analysis of algorithms

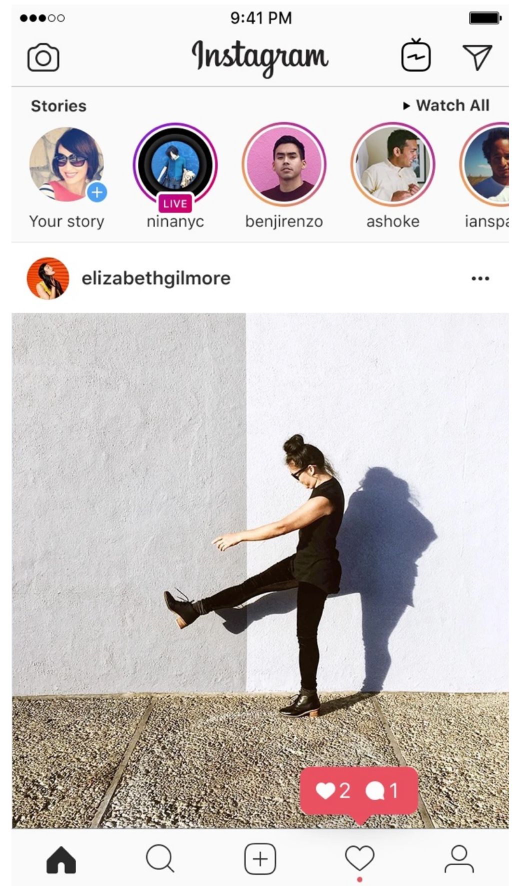




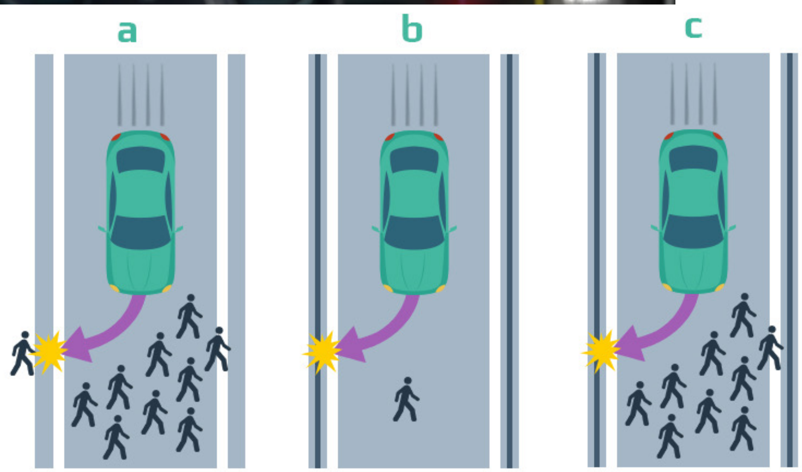
Probability in practice



Probability at your fingertips



Probability and ethics



The golden rule for autonomous car ethics doesn't exist

So far, there are no unified ethical standards and [certifications](#) for autonomous cars. The big [Moral Machine study](#) conducted by MIT showed that it's hard to identify universal ethical values. The moral choices that people made in the MIT survey were different and varied even at a local level. That's why it's hard to create a universal ethics of self-driving cars that won't be controversial. [[source](#)]

We'll get there!

Probability is not always intuitive.

Disease testing

A patient takes a virus test that returns positive.

What is the probability they really have the virus?

- 0.03% of people have the virus
- Test has 99% positive rate for people with the virus
- Test has 7% positive rate for people without the virus

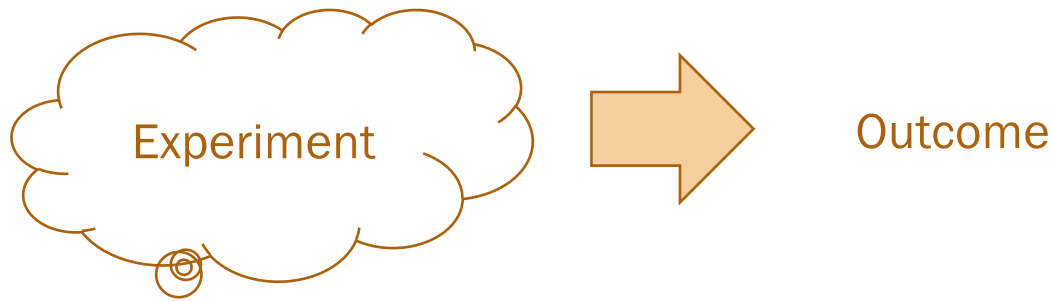
Answer: 0.42%

Probability = Important
+ Needs Studying

Counting I

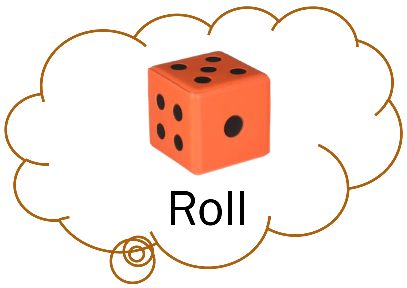
What is Counting?

An experiment
in probability:



Counting: How many possible outcomes can occur by performing this experiment?

What is ~~Counting~~ Combinatorial Analysis?

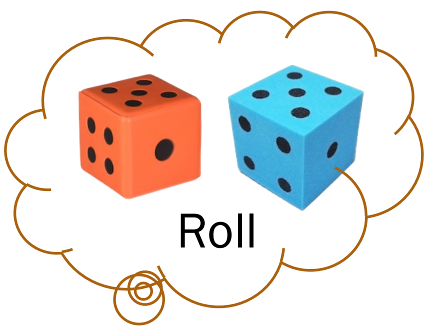


6

{1, 2, 3,
4, 5, 6}

3

{2, 4, 6}



36

{(1, 1) , (1, 2), (1, 3), (1, 4), (1, 5), (1, 6),
(2, 1) , (2, 2), (2, 3), (2, 4), (2, 5), (2, 6),
(3, 1) , (3, 2), (3, 3), (3, 4), (3, 5), (3, 6),
(4, 1) , (4, 2), (4, 3), (4, 4), (4, 5), (4, 6),
(5, 1) , (5, 2), (5, 3), (5, 4), (5, 5), (5, 6),
(6, 1) , (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)}

Sum Rule of Counting

If the outcome of an experiment can be either from

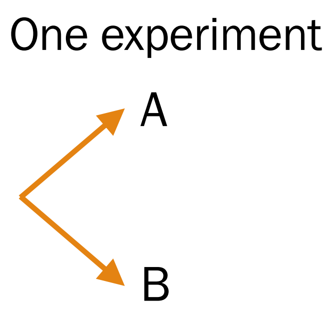
Set A, where $|A| = m$,

<u>or</u> Set B, where $|B| = n$,

where $A \cap B = \emptyset$,

Then the number of outcomes of the experiment is

$|A| + |B| = m + n$.



Product Rule of Counting

If an experiment has two parts, where

the first part's outcomes are drawn from A, where $|A| = m$,
<u>and</u> the second part's outcomes are drawn from B, where $|B| = n$,

Then the number of outcomes of the experiment is

$$|A||B| = mn.$$

Two-step experiment

$\longrightarrow$ A $\longrightarrow$ B

CS109 Mixer

Introduce yourself to you neighbors.

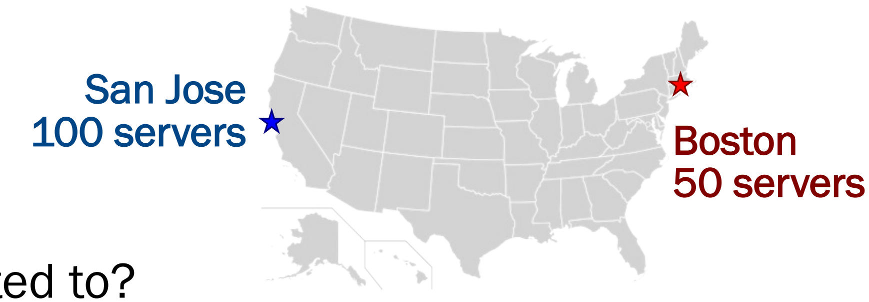
Then check out the question on Slide 52.
Post any clarifications here!

https://edstem.org/us/courses/21301/discussion/1324037

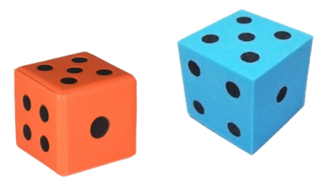
Let's try it out

Sum Rule, Product Rule, or something else? How many outcomes?

1. Video streaming application
   - Your application has distributed servers in 2 locations (SJ: 100, Boston: 50).
   - If a web request is routed to a server, how large is the set of servers it can get routed to?



2. Dice
   - How many possible outcomes are there from rolling two 6-sided dice?



3. Strings
   - How many *different* orderings of letters are possible for the string BOBA?

BOBA, ABOB, OBBA…

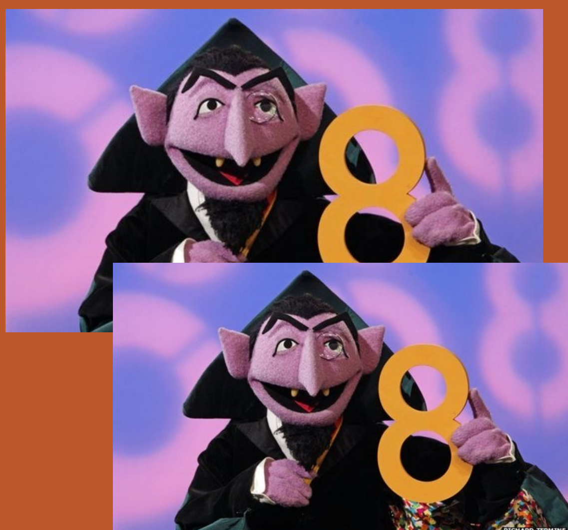
Let's try it out

Sum Rule, Product Rule, or something else? How many outcomes?

1. Video streaming application
   - Your application has distributed servers in 2 locations (SJ: 100, Boston: 50).
   - If a web request is routed to a server, how large is the set of servers it can get routed to?
2. Dice
   - How many possible outcomes are there from rolling two 6-sided dice?
3. Strings
   - How many *different* orderings of letters are possible for the string BOBA?

Builtin Pause

Questions?



Counting II

TOP DEFINITION

kick it up a notch

To make things more intense, exciting, or interesting.

(introduced by **chef Emeril Lagasse** in reference to **spicing** up his recipes :)

Inclusion-Exclusion Principle

If the outcome of an experiment can be either from

Set A <u>or</u> set B,

where A and B may overlap,

Then the total number of outcomes of the experiment is

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

One experiment

A

B only

Sum Rule of Counting:
A special case

Transmitting bytes over a network

An 8-bit string is sent over a network.

- The receiver only accepts strings that either start with 01 or end with 10.

$$\underbrace{\texttt{\underline{01}0011\underline{00}}}_{\textbf{byte}\text{ (8 bits)}}$$

How many 8-bit strings will the receiver accept?

Define

A : 8-bit strings starting with 01

B : 8-bit strings ending with 10

Transmitting bytes over a network

An 8-bit string is sent over a network.

- The receiver only accepts strings that either start with 01 or end with 10.

How many 8-bit strings will the receiver accept?

$$\underbrace{\underline{\texttt{01}}\texttt{0011}\underline{\texttt{00}}}_{\textbf{byte}\text{ (8 bits)}}$$

Define

A : 8-bit strings starting with 01

B : 8-bit strings ending with 10

General Principle of Counting

If an experiment has r steps, such that

Step i has n_i outcomes for all $i = 1, \ldots, r$,

Then the number of outcomes of the experiment is

$$n_1 \times n_2 \times \cdots \times n_r = \prod_{i=1}^{r} n_i \,.$$

Multi-step experiment

$\rightarrow 1 \rightarrow 2 \rightarrow \ldots$

Product Rule of Counting: A special case

License plates

How many CA license plates are possible if…

(pre-1982)

(present day)

License plates

How many CA license plates are possible if…

(pre-1982)

(present day)

Pigeonhole Principle

Floors and ceilings

Floor function

$\lfloor x \rfloor$

The largest integer $\leq x$

Ceiling function

$\lceil x \rceil$

The smallest integer $\geq x$

Check it out:

$\lfloor 1/2 \rfloor$ $\lfloor 2.9 \rfloor$ $\lfloor 8.0 \rfloor$ $\lfloor -1/2 \rfloor$

$\lceil 1/2 \rceil$ $\lceil 2.9 \rceil$ $\lceil 8.0 \rceil$ $\lceil -1/2 \rceil$

Pigeonhole Principle

For positive integers m and n,

if m objects are placed in n buckets,

then at least one bucket must contain at least $\lceil m/n \rceil$ objects.

Pigeons in holes

21st century pigeons

Example:

m objects = 10 pigeons

n buckets = 9 pigeonholes

At least one pigeonhole must contain $\lceil m/n \rceil = 2$ pigeons.

Bounds: an important part of CS109

Balls and urns

n balls

r urns (buckets)

~~Balls and urns~~ Hash Tables and strings

Consider a hash table with 100 buckets.

950 strings are hashed and added to the table.

1. Is it guaranteed that at least one bucket contains ***at least*** 10 entries?

2. Is it guaranteed that at least one bucket contains **at least** 11 entries?

3. Is it possible to have a bucket with ***no entries***?

~~Balls and urns~~ Hash Tables and strings

Consider a hash table with 100 buckets.

950 strings are hashed and added to the table.

$n = 100$
$m = 950$

1. Is it guaranteed that at least one bucket contains ***at least*** 10 entries?

   Yes

2. Is it guaranteed that at least one bucket contains **at least** 11 entries?

   No

3. Is it possible to have a bucket with ***no entries***?

   Sure

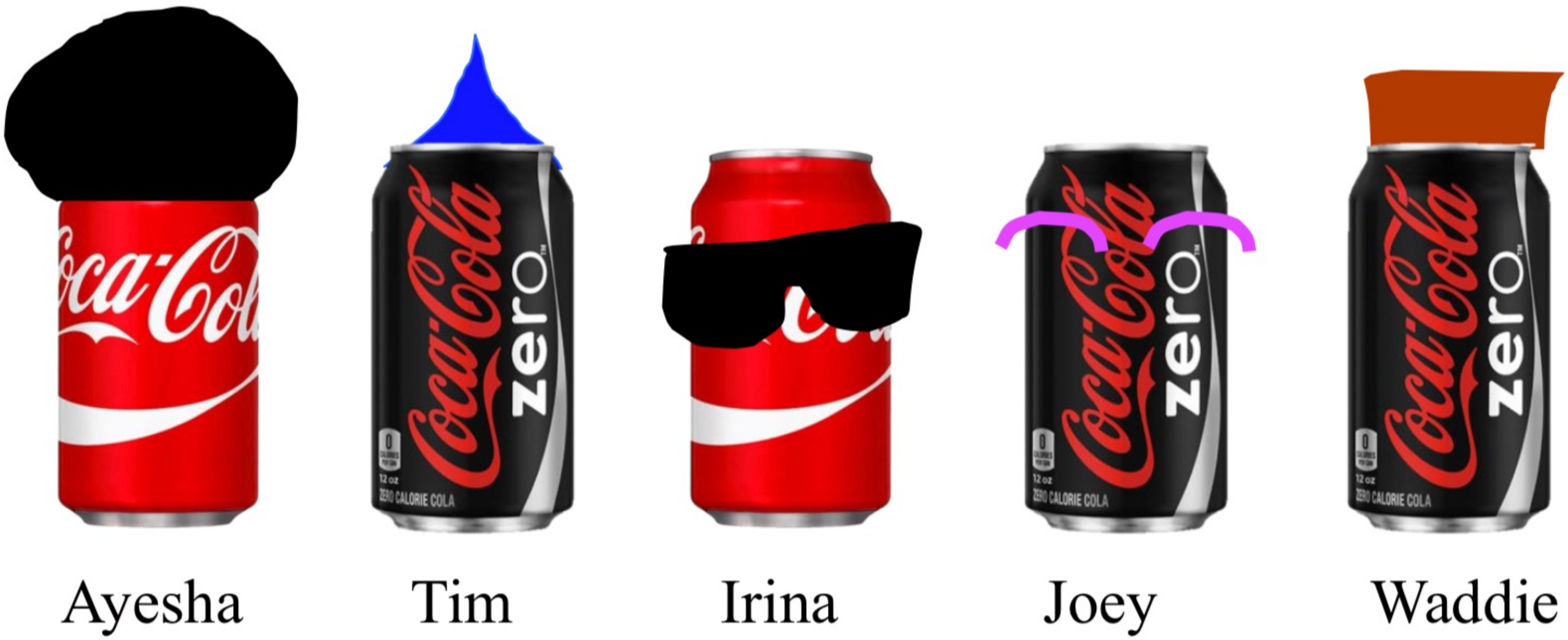
Permutations I

Unique 6-digit passcodes with six smudges

How many unique 6-digit passcodes are possible if a phone password uses each of six distinct numbers?

Sort n indistinct objects

Sort n distinct objects



Sort n distinct objects

1^{st} 2^{nd} 3^{rd} 4^{th} 5^{th}

Steps:

1. Choose 1^{st} can — 5 options
2. Choose 2^{nd} can — 4 options

...

5. Choose 5^{th} can — 1 option

$$\text{Total} = 5 \times 4 \times 3 \times 2 \times 1$$
$$= 120$$

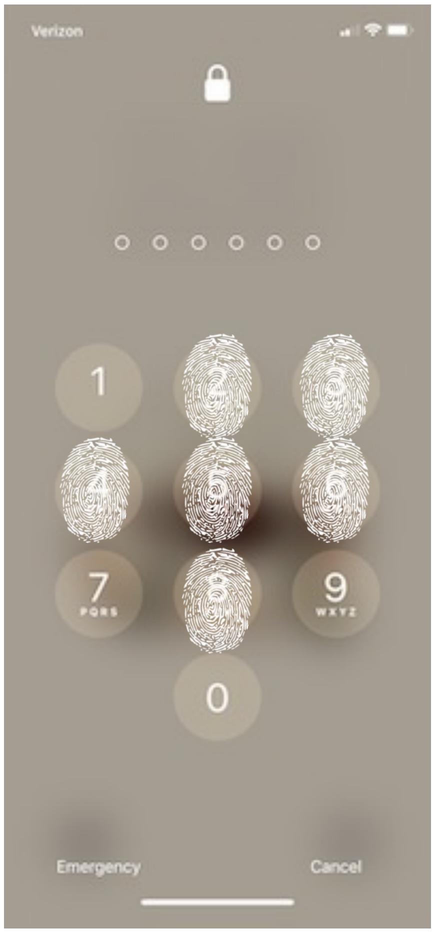
Permutations

A permutation is an ordered arrangement of objects.

The number of unique orderings (permutations) of n distinct objects is

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1.$$

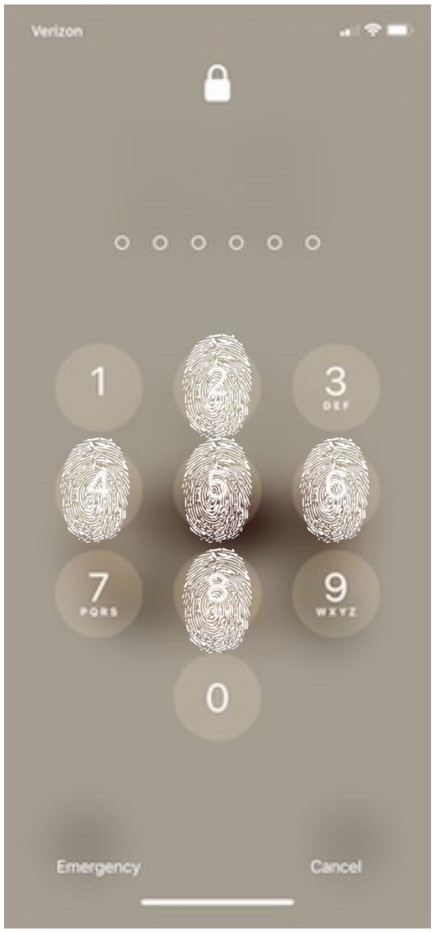
Unique 6-digit passcodes with six smudges



How many unique 6-digit passcodes are possible if a phone password uses each of **six** distinct numbers?

$$\text{Total} = 6!$$
$$= 720 \text{ passcodes}$$

Unique 6-digit passcodes with five smudges



How many unique 6-digit passcodes are possible if a phone password uses each of five distinct numbers?