Winning Wordle with Math

Benjamin Wilson
bwilson4@stevenson.edu

STEVENSON UNIVERSITY | BEVERLY K. FINE SCHOOL of the SCIENCES Mathematics

Fall 2023 MD-DC-VA Section Meeting
Stevenson University

How to play Wordle

Rules

- You have 6 tries to guess the five-letter word
- Each guess must be a valid 5-letter word
- The color of the tiles will change to show how close your guess was to the word

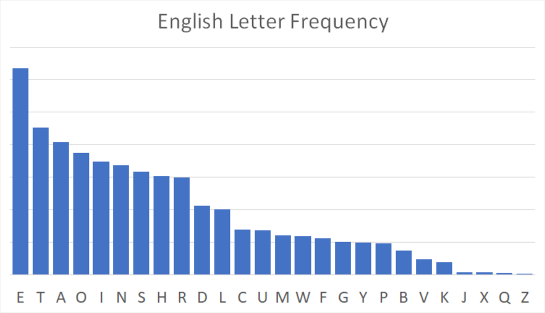
Try it!

https://mywordle.strivemath.com/?word=dcivp

Strategies

Strategies

Guess words with common letters



https://en.wikipedia.org/wiki/Letter_frequency

Strategies

Good Words

RATES	LANES	TEALS	RITES	PALES	SCANT	OTHER
SNARE	MARES	SOARE	RAILS	TONES	SHARD	NAILS
TEARS	ROLES	STARE	TIRES	DARES	TIDAL	BLIND
TALES	SIREN	SALET	SLATE	REALS	TRASH	RAINS
TASER	TARES	DATES	DEARS	READS	SLIDE	CLEAT
CARES	TILES	EARNS	MALES	CRANE	DANCE	CLEAN
CRATE	CHORD	EARTH	SANER	DREAM	CLOTH	ASCOT

Strategies

Good Words

RATES	LANES	TEALS	RITES	PALES	SCANT	OTHER
SNARE	MARES	SOARE	RAILS	TONES	SHARD	NAILS
TEARS	ROLES	STARE	TIRES	DARES	TIDAL	BLIND
TALES	SIREN	SALET	SLATE	REALS	TRASH	RAINS
TASER	TARES	DATES	DEARS	READS	SLIDE	CLEAT
CARES	TILES	EARNS	MALES	CRANE	DANCE	CLEAN
CRATE	CHORD	EARTH	SANER	DREAM	CLOTH	ASCOT

But which words are best?

Strategies

One method of determining which words to guess is by maximizing the (Shannon) entropy of the guess which is the expected value of the information gained by making the guess.

Strategies

One method of determining which words to guess is by maximizing the (Shannon) entropy of the guess which is the expected value of the information gained by making the guess.

Example (Playing Cards)

Suppose I pick a playing card at random and ask you to figure out my card by only asking me Yes/No questions. What might you ask? What questions are best? What questions gain you more information?

How many Yes/No questions do you expect to ask, on average?

Strategies

One method of determining which words to guess is by maximizing the (Shannon) entropy of the guess which is the expected value of the information gained by making the guess.

Example (Playing Cards)

Suppose I pick a playing card at random and ask you to figure out my card by only asking me Yes/No questions. What might you ask? What questions are best? What questions gain you more information?

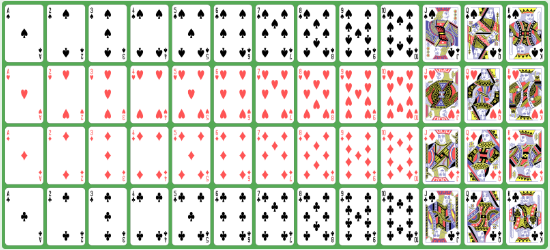
How many Yes/No questions do you expect to ask, on average?

Answer: ≈ 5.7 questions

Strategies

Example (Playing Cards)

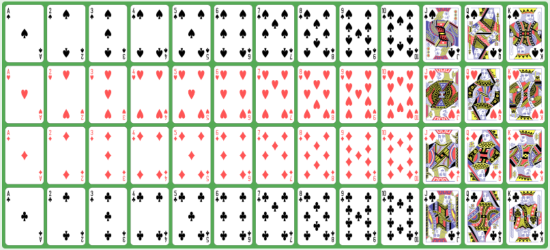
- A. Is the card a face card?
- B. Is the card a club?
- C. Is the card a red card?
- D. Is the card the 2 of diamonds?



Strategies

Example (Playing Cards)

- A. Is the card a face card? Probability $= 12/52 \approx 23\%$
- B. Is the card a club? Probability $= 13/52 = 25\%$
- C. Is the card a red card? Probability $= 26/52 = 50\%$
- D. Is the card the 2 of diamonds? Probability $= 1/52 \approx 2\%$



Information

The amount of information conveyed by data, $I(\text{data})$, is measured in *bits* and is given by

$$I(\text{data}) = -\log_2 (\text{Probability of learning that data}) .$$

This can be viewed as the number of times your sample space is cut in half if you learn that data.

Information

The amount of information conveyed by data, $I(\text{data})$, is measured in *bits* and is given by

$$I(\text{data}) = -\log_2 (\text{Probability of learning that data}) .$$

This can be viewed as the number of times your sample space is cut in half if you learn that data.

Example (Playing Cards)

A. Is the card a face card? Information $= -\log_2(12/52) \approx 2.1$ bits

B. Is the card a club? Information $= -\log_2(13/52) = 2$ bits

C. Is the card a red card? Information $= -\log_2(26/52) = 1$ bit

D. Is the card the 2 of diamonds? Information $= -\log_2(1/52) \approx 5.7$ bits

Entropy

The *entropy*, $H(X)$, of a discrete random variable X (like the specific card you choose from a 52-card deck) is the average amount of information contained in each piece of data received about the value of X.

$$H(X) = -\sum_{i=1}^{N} p_i \log_2(p_i)$$

Notice, that something with low probability gives a lot of information, but contributes less to the entropy.

Entropy

The *entropy*, $H(X)$, of a discrete random variable X (like the specific card you choose from a 52-card deck) is the average amount of information contained in each piece of data received about the value of X.

$$H(X) = -\sum_{i=1}^{N} p_i \log_2(p_i)$$

Notice, that something with low probability gives a lot of information, but contributes less to the entropy.

Maximizing entropy

To maximize entropy, if there are N outcomes, we want p_i to be as close to $1/N$ as possible.

Entropy

Example (Playing Cards)

- A. Is the card a face card? Probability $= 12/52 \approx 23\%$
- B. Is the card a club? Probability $= 13/52 = 25\%$
- C. Is the card a red card? Probability $= 26/52 = 50\%$
- D. Is the card the 2 of diamonds? Probability $= 1/52 \approx 2\%$

In the card example, there are two outcomes to each question (Yes or No), so to maximize entropy, we want to ask a question that has as close as possible to a 50% chance of being true. A great first question to ask is "Is the card red?". What's a good second question?

Entropy and Wordle

Each guess in Wordle has 3^5 possible outcomes (since there are 5 letters and each can be Green, Yellow, or Black for a given guess). To maximize entropy, it would be great to guess a word where each of these outcomes is equally likely (leading to an entropy of about 7.9 bits). In practice, this won't be possible, so we just want to get as close to this scenario as we can.

High entropy starting words

Word	TARES	RATES	SOARE	TALES	TEARS	SALET
Entropy (bits)	6.196	6.098	6.062	6.056	6.034	6.018

Entropy and Wordle

Each guess in Wordle has 3^5 possible outcomes (since there are 5 letters and each can be Green, Yellow, or Black for a given guess). To maximize entropy, it would be great to guess a word where each of these outcomes is equally likely (leading to an entropy of about 7.9 bits). In practice, this won't be possible, so we just want to get as close to this scenario as we can.

High entropy starting words

Word	TARES	RATES	SOARE	TALES	TEARS	SALET
Entropy (bits)	6.196	6.098	6.062	6.056	6.034	6.018

Since each of these words has an entropy around 6 bits, on average, guessing one of these words will cut the number of possibilities in half 6 times (from about 13000 to about 200). The lowest entropy starting words are QAJAQ, XYLYL, IMMIX, FUFFY, and PZAZZ all with an entropy around 2 bits.

Excel sheet with more results

Other Strategies

Beyond maximizing entropy, what other strategies might we employ?

- Brute force all games with a specific starter - SALET is best starter (See here for more)
- Brute force all games with 2-start - PARSE, CLINT is best (See here for more)
- Brute force all games with 3-start - BLIND, CHAPT, MORSE is best (See here for more)
- Take advantage of common English words in entropy computation (Watch this video by math YouTuber 3Blue1Brown to get a feel for this and also because it's awesome!)

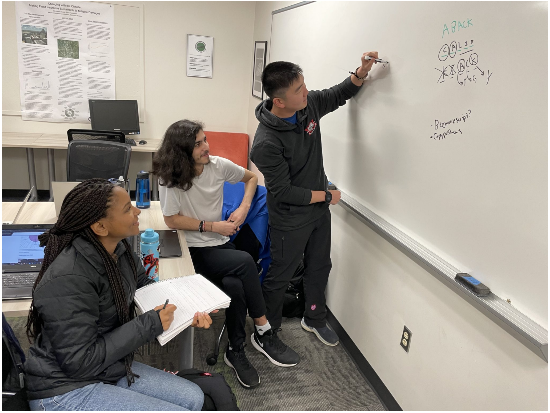
Other Questions

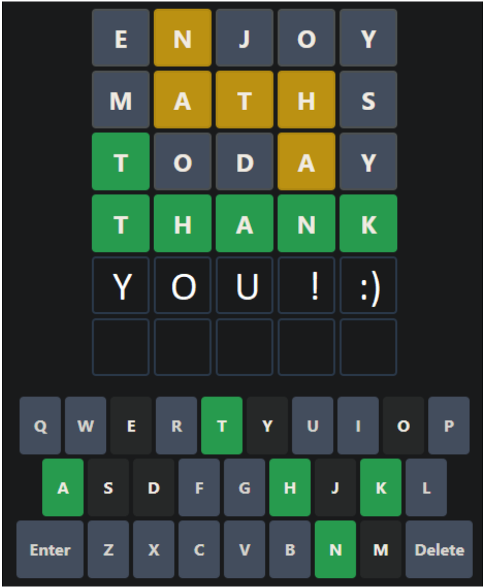
Several Stevenson Applied Mathematics students explored other questions related to math and Wordle. Some of their research questions included:

- Are humans or computers better at solving Wordle puzzles? Does it depend on the word?
- Can we use Scrabble or Morse Code to help solve Wordle?
- Could Shakespeare guess today's Wordle?
- If your only knowledge of the English language consisted of Taylor Swift lyrics, could you solve a Wordle?

Other Questions

To learn more, check out my students' research posters in MAC S150 or reach out to my students: Julian Cha, Brittney Johnson, or Chris Villanueva!




E N J O Y
M A T H S
T O D A Y
T H A N K
Y O U ! :)
Q W E R T Y U I O P
A S D F G H J K L
Enter Z X C V B N M Delete