GOBSHITEY

VOWEL　CONSONANT

The Best 100 words to know for Countdown

Building a puzzle book fast with Python and Generative AI

David Curran June 2025

This Talk: Hack a game and show others how to

- Moneyball

- Countdown Letters Round

- English (Natural Language Processing)

  - Finding the best words

  - Finding the best words given you know other words

  - Variants of words

- Making Word Games in Python

  - Word Search

  - Crosswords

  - Letters round, cloze

"Virtually everyone suffers from the deeply ingrained habit of considering language as a medium of communication."

— Dmitri Borgmann, in Language on Vacation

"Dacoity, that's a Countdown word"

— Susie Dent

Moneyball

By using data and analysis of the correct metric you can improve performance.

In baseball this was on base percentage. Before people thought in terms of how often you hit the ball, got home runs, scouting opinions

Find an area where the KPI being chased after is not the right one.

Basketball more 3 pointer

Popup pass when tackled in Rugby

In Countdown people think they are trying to find words they know. They should be looking to learn **words that will be useful**.
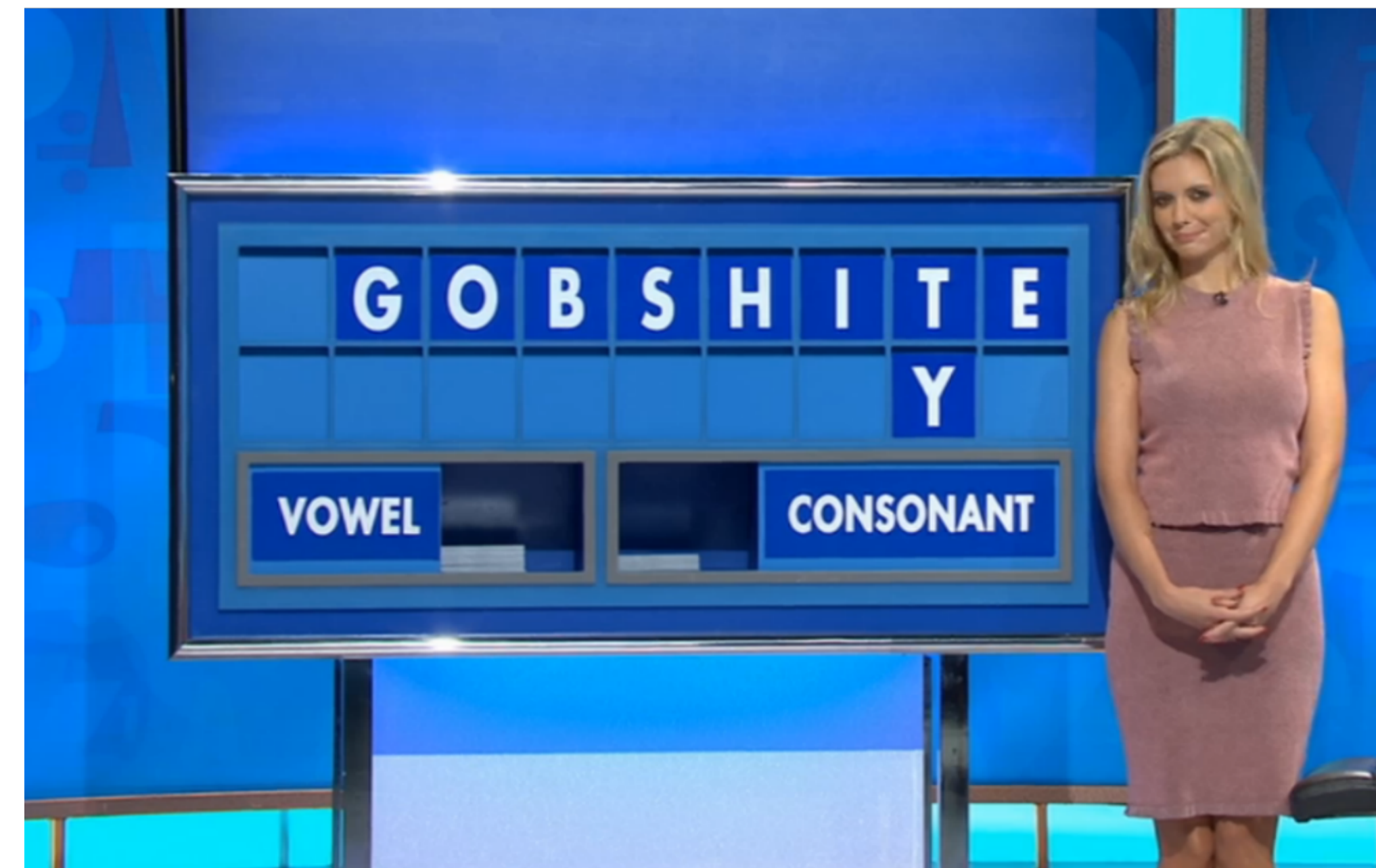
Countdown Letters Rules

- Select 9 letters. 3, 4 or 5 have to be Vowels

-

- Then make the longest word you can out of these 9 letters.

- 7 or 8 is a good score

Questions

- What do 9 letters draws look like

- What's the average number of vowels asked for

- What's the best 100 words to know to ge the most points?

- Turn these 100 into a puzzle book for learning

Draw from this Letter Pile
Roughly same probability for each letter as in English

Consonant pile								
Letters:	RST	N	D	L	GMP	C	BFHVW	JKQXYZ
Frequency:	9	8	6	5	4	3	2	1

Vowel pile					
Letters:	E	A	I	O	U
Frequency:	20	15	13	13	7

Vowel Counts

- The percentage of games with 3, 4, or 5 vowels changes what words are likely to come out.

- Make a scraper for the countdown site and calculate

wiki.apterous.org/Episode_5219   New Chrome availabl

Michelle Nevitt played Craig Laurenson, with Michelle Nevitt winning 89 – 69. The Dictionary Corner guest was John Parrott, and the lexicographer was Susie Dent.

Rounds

Rnd	Selection	Michelle Nevitt	Score	Craig Laurenson	Others	Max.
1	S O A W G O H D E	WASHED	6 − 6	WASHED	WOOSHED*	7
2	G P F A O V D E S	PODGES	12 − 12	SOAPED	DAGOES*, DOSAGE, GASPED, VADOSE*	13
3	R A L T O D M U E	ROULADE	19 − 12	ROUTE	EMULATOR, MODULATE	21
4	Z T C I A T E R F	CATTIER	26 − 19	CATTIER	TRIFECTA*	29
5	100 1 7 9 3 4 → 892	892 $100 \times 9 - 7 - 1$	36 − 29	892 $100 \times 9 - 7 - 1$		39
TTT	AMANSHIP	It's not a man's hip, it's a woman's scarf.			PASHMINA	
6	S N G M U O A D E	MOUNDS	36 − 36	MANGOES	DAEMONS, MASONED*, MENUDOS*, OSMUNDA*	46
7	C A R N N I E Y P	CANNIER	43 − 43	CANNIER	CANNERY, CAPRINE*, PANNIER, PINNACE*	53
8	C V S A E D T I H	DITCHES	50 − 50	DITCHES	ACHIEST, ACTIVES*, ADVECTS*, ADVICES*, AITCHES, DATIVES*, SCATHED*	60
9	S E R N I O T R P	PROTEINS	58 − 58	POINTERS	INTRORSE*, PRINTERS, PRISONER*, REPOINTS, REPRINTS*, SPORTIER*, SPRINTER*	68
10	100 2 4 3 2 8 → 969	—	58 − 63	976 $(8 + 2) \times 100 - 4 \times 2 \times 3$	RR: 970 = (8 + 2) × (100 − 3)	75
TTT	LOSESOAR	There are no genies in these, just little squirts.			AEROSOLS	
11	X E A S P O T W U	EXPATS	64 − 69	EXPATS	SEXPOT*	81
12	L N N G E A I B N	ENABLING	72 − 69	LEANING		89
13	L O E R M R E U N	MOURNER	79 − 69	LEMON		96
14	25 8 8 6 7 7 → 266	—	79 − 69	—	RR: 266 = (25 + 7 + 6) × 7	106
15	D R I E D S A G E	DISAGREED (24.25 SECONDS)	89 − 69			116

Vowel Counts in Games

- 41% of letters round games have 3 vowels

- 54% have 4 vowels

- 5% have 5 vowels

- You need to know these to decide which is more useful
  Etaerio or *Standee*

  This scraping and aggregating is an less than an hour with Cursor.

Find what words we can make from a random draw

- Make a million draws

- Find 7,8,9 letter words

- Get an English dictionary. Complete pain for these sorts of problems

- Count how often a word can be found

```python
from itertools import combinations

# Assume anagram_dict is {sorted_letters: [word1, word2, ...]}
def find_words_for_draw(draw, anagram_dict, word_lengths=[7,8,9]):
    found_words = set()
    for length in word_lengths:
        for combo in combinations(draw, length):
            key = ''.join(sorted(combo))
            if key in anagram_dict:
                found_words.update(anagram_dict[key])
    return found_words
```

Most Found Words

Canonical	Total Hit Count
OTARIES	5373
ERASION	4868
OTARINE	4799
ATONIES	4777
ATONERS	4356
ANESTRI	4254
ROSEATE	4171
AERIEST	4069
ROADIES	3850
NORITES	3837
ARENOSE	3817
IODATES	3725
ARENITE	3718
EROTISE	3668
ETESIAN	3666
EARNEST	3604
ADONISE	3485
ANEROID	3463

Mutual Information

Naive vs. True Coverage: The Impact of Word Interference
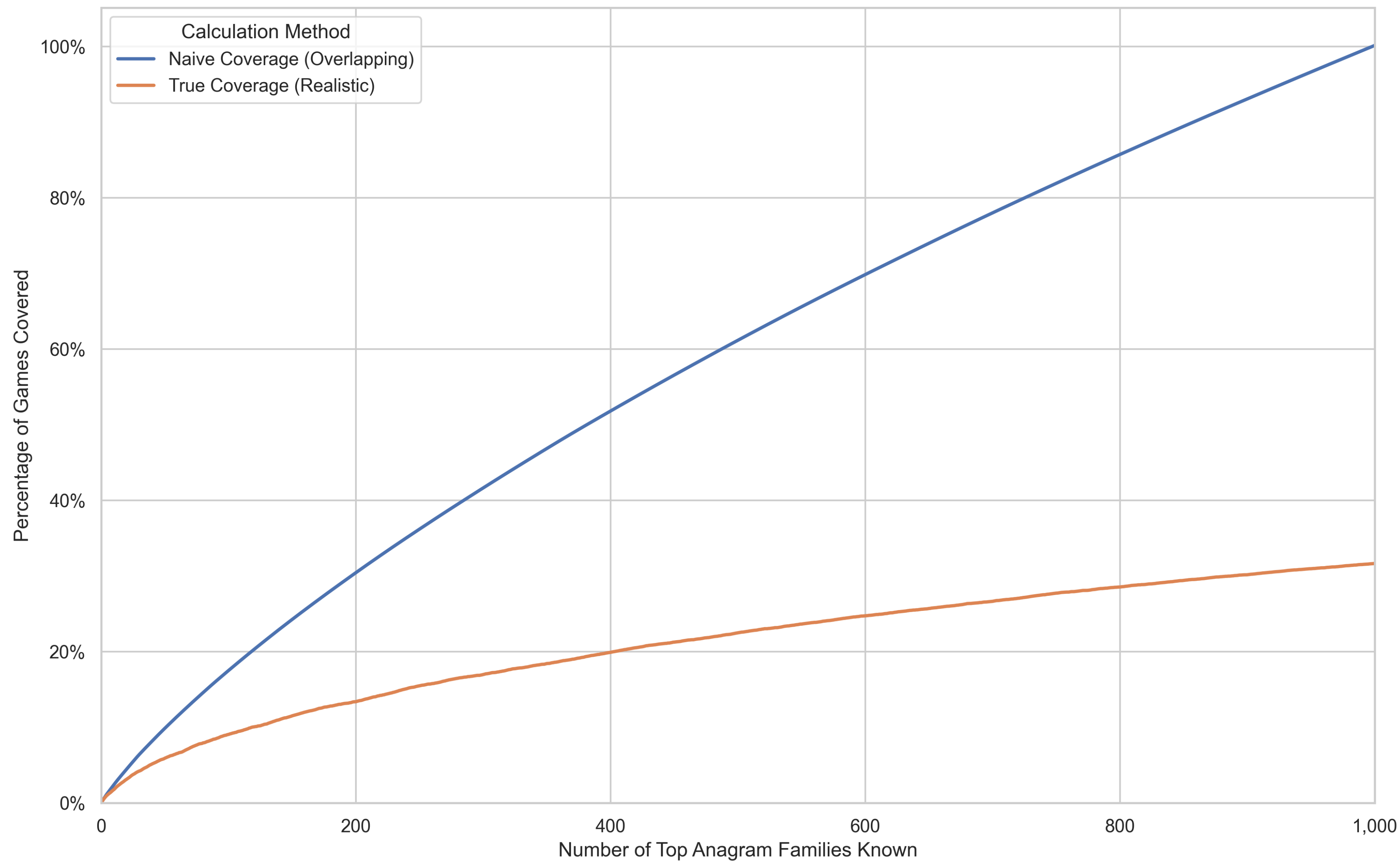
Double Count

- Current Words use letters that are already covered

- AEIO NRSST

  - ROASTIE

  - NASTIER

Greedy Algorithm
Find the word that matches the most draws

```
{
    "AEINRST": {0, 5, 12, 100, ...},     # All draws where this anagram family (e.g., RETAINS,
STAINER, etc.) can be made
    "AELNOST": {2, 7, 15, 22, ...},
    ...
}
```

- Find the next word that matches the most draws

- For draws that are not already matched

- Words are sorted vowel/consonant and alphabetically

- Then later find best word with those characters (anagrams)

Greedy Results
Top ones the same. 59 out of 100 different

NASTIER
SENATOR
STONIER
DIASTER
OESTRAL
ATONIES
ENTRAIL
ERASION
TORNADE
STANDEE
OERSTED
LEANERS
RETILES

....

Is this Really Natural Language Processing?
All in Data Science

- Ran a simulation of letter draws

- Scraped the wiki to get vowel number percentages

- Searched the English dictionary

- Greedy search to find optimal additions to a set

Now Turn this into a Book

- All the text I write myself. Though I will ask an LLM to critique it

- Group these 100 words into categories

- Around 8 words a day is reasonable to learn

  - 1. Add definition

  - 2. Find related words

```python
prefixes = ["UN", "CO", "RE", "DE", "IN"]
suffixes = ["ING", "Y", "ED", "LY", "NESS", "ABLE", "ER", "IEST", "EST", "ATE", "AL","s"]
and all
pattern = r".*ROTA.*"
matching_words = df_cdwords[df_cdwords[0].str.contains(pattern, case=False, na=False)]
print(matching_words.to_string(index=False))
```

Games

Made with Python. Very fast in Cursor

- Cloze

Puzzle 2.2 Cloze Puzzle – Fill in the Vowels

#								Vowels
1.			T	L			R	EIOU
2.	R			N			T	AIOU
3.	S			T		N		AEOU
4.			T		N		D	EEOU
5.			T	W			R	AEOU
6.	R		D				T	EIOU
7.			T	S			N	EEOU
8.	R			D			T	AEOU
9.			T		G			AEOU

Wordsearch
Python library from github

WORD SEARCH

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

Find words going NE, SW, N, SE, S, W, NW, and E:

OUTAGE OUTENED OUTLIER OUTSEEN OUTWEAR RAINOUT
READOUT RIDEOUT SOUTANE

Simple Letters Round

Puzzle 2.3 Letters Round – Find the Word

1. L N R T A E I O
2. D N T A E O O
3. G R T A E I O
4. D N R T A E I O
5. G T A E E O
6. G T A I O U
7. D L R A E I O

Realistic Letters Round
This is also for spaced repetition

Puzzle 2.3 Letters Round – Find the Word

1. | D | Q | R | R | T | A | E | O | U |

2. | N | J | Q | S | T | A | E | O | U |

3. | D | F | G | N | T | E | E | O | U |

4. | B | C | N | R | T | A | I | O | U |

5. | M | N | P | S | T | E | E | O | U |

6. | D | H | K | R | T | E | I | O | U |

7. | C | G | L | S | T | A | E | O | U |

8. | B | D | R | T | W | A | E | O | U |

9. | L | R | S | T | W | E | I | O | U |

Letters Reminders 0,1, 2, 7, 14 days
Leitner System

The Forgetting Curve
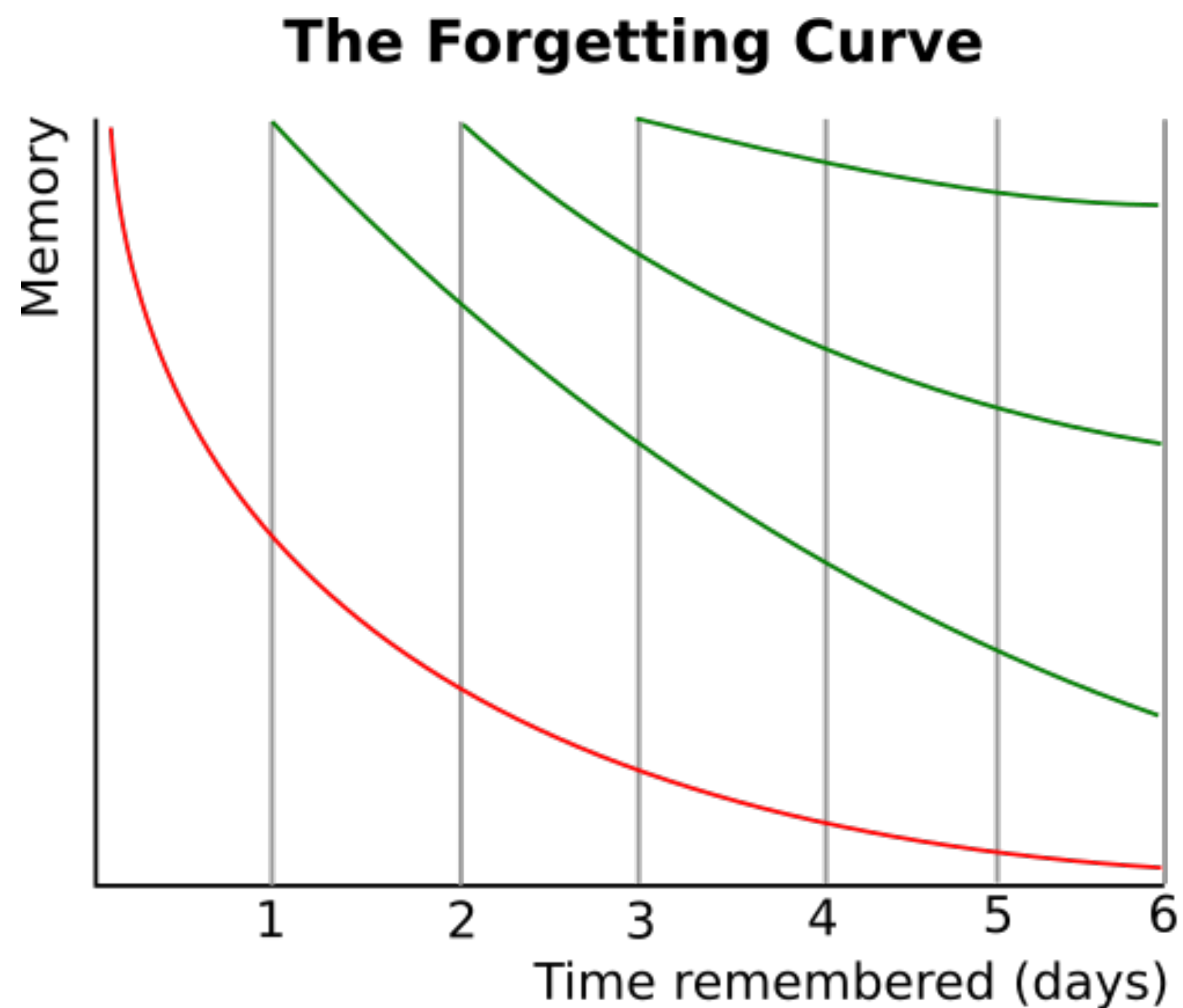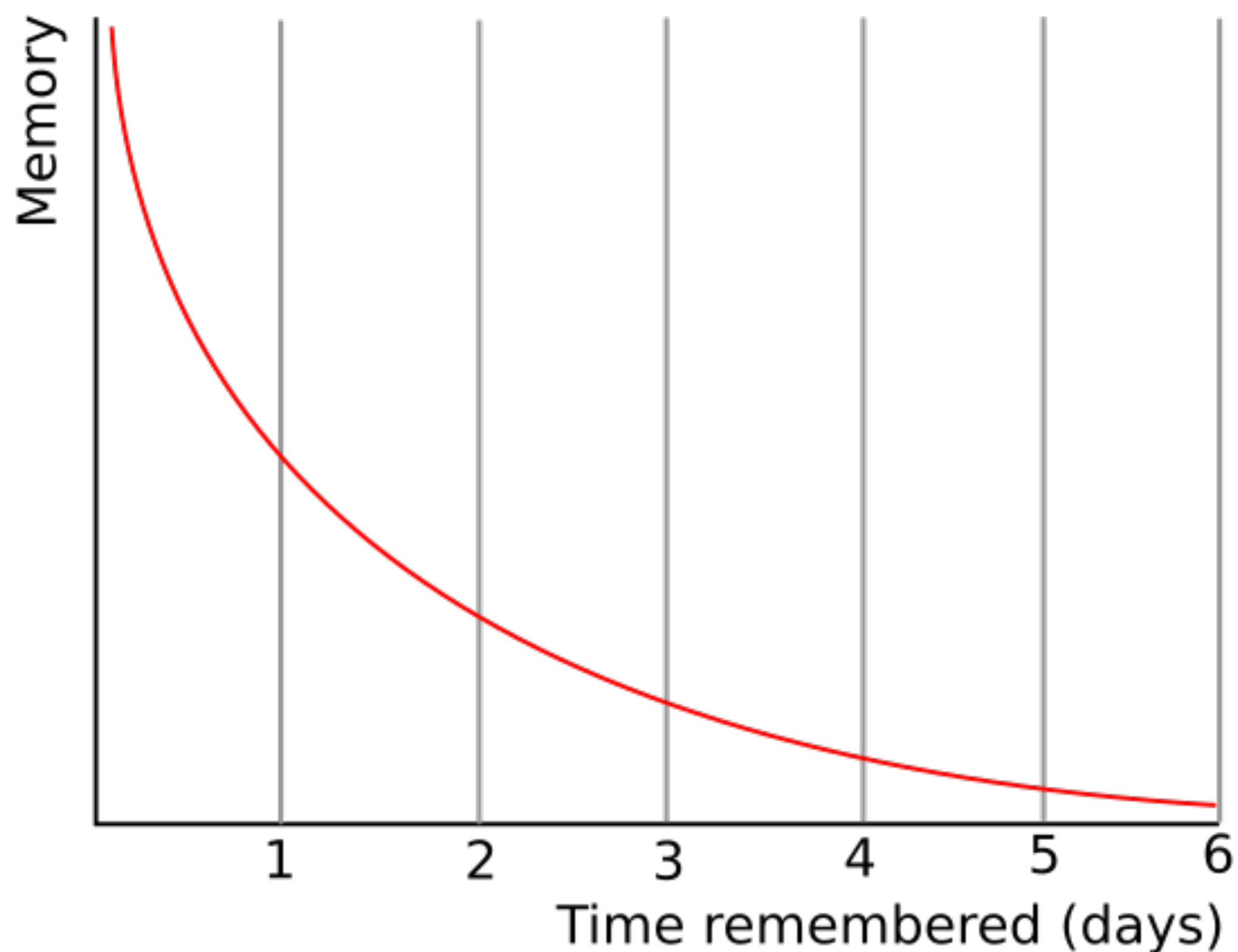
Cursor Made Anki Deck 10 minutes

Left screen:

Countdown Anagr...
33 minutes left

98 2 0

NRST EIO

Show answer

Right screen:

Countdown Anagr...
33 minutes left

98 2 0

NRST EIO

STONIER
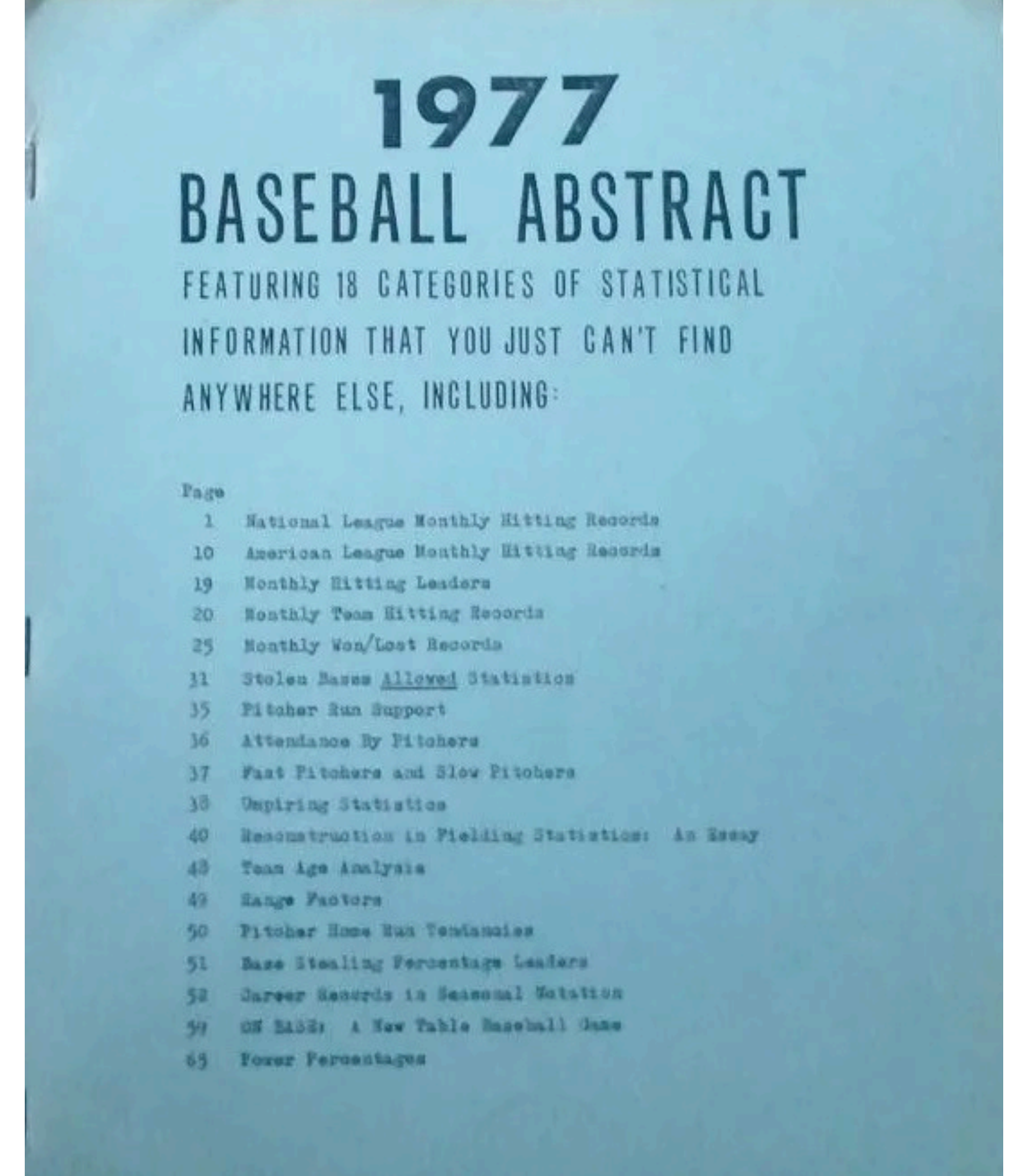
<1m	<6m	<10m	3d
Again	Hard	Good	Easy

What Next

- Have the 100 words classified

- Have code for finding related words

- Have game generating code for a list of words

- Make up the 12 chapters

- I want to print this and sell 75 copies like the 1977 Bill James Baseball Abstract

- All code will be public on github