Week 2: Cryptanalysis and Statistical Modelling

Jay Daigle

Occidental College

September 6, 2018

Definition

Let $\mathbf{s} = c_1 c_2 \ldots c_n$ *be a string of* n *letters. The index of coincidence of mathbfs is denoted* $\mathsf{IndCo}(\mathbf{x})$ *and is defined to be the probability that two randomly chosen characters in the string* $\mathbf{s}$ *are identical.*

Definition

Let $\mathbf{s} = c_1 c_2 \ldots c_n$ *be a string of* n *letters. The index of coincidence of mathbfs is denoted* $\mathrm{IndCo}(\mathbf{x})$ *and is defined to be the probability that two randomly chosen characters in the string* **s** *are identical.*

Proposition

Let $\mathbf{s} = c_1 c_2 \ldots c_n$ *be a string of* n*, and let* F_i *be the frequency with which the letter* i *appears in the string* **s**. *Then*

$$\mathrm{IndCo}(\mathbf{s}) = \frac{1}{n(n-1)} \sum_{i=0}^{25} F_i(F_i - 1). \tag{1}$$

Two important values of the index of coincidence

Two important values of the index of coincidence

Proposition

1. *If* $\mathbf{s}$ *is a string of letters generated uniformly at random, then* $\mathsf{IndCo}(\mathbf{s}) \approx .038$.

Two important values of the index of coincidence

Proposition

1. *If* $\mathbf{s}$ *is a string of letters generated uniformly at random, then* $\mathrm{IndCo}(\mathbf{s}) \approx .038$.
2. *If* $\mathbf{s}$ *is a string of letters with the frequencies common in written English, then* $\mathrm{IndCo}(\mathbf{s}) \approx .068$.

Recall the Vigenère Cipher

Recall the Vigenère Cipher

Choose a keyword. Write the keyword down repeatedly until you have a string of letters as long as your message. This is the *keystream*.

Recall the Vigenère Cipher

Choose a keyword. Write the keyword down repeatedly until you have a string of letters as long as your message. This is the *keystream*.
For each letter of your plaintext, add the corresponding letter from your keystream. This gives the ciphertext.

Recall the Vigenère Cipher

Choose a keyword. Write the keyword down repeatedly until you have a string of letters as long as your message. This is the *keystream*.
For each letter of your plaintext, add the corresponding letter from your keystream. This gives the ciphertext.
To decrypt, generate the keystream as before, and then subtract it from the ciphertext to get the plaintext.

Example

Example

Plaintext | I L O V E C R Y P T O L O G Y

Example

Plaintext	I	L	O	V	E	C	R	Y	P	T	O	L	O	G	Y
Keystream	M	A	T	H	M	A	T	H	M	A	T	H	M	A	T

Example

Plaintext	I	L	O	V	E	C	R	Y	P	T	O	L	O	G	Y
Keystream	M	A	T	H	M	A	T	H	M	A	T	H	M	A	T
Plaintext	8	11	14	21	4	2	17	24	15	19	14	11	14	6	24
Keystream	12	0	19	7	12	0	19	7	12	0	19	7	12	0	19

Example

Plaintext	I	L	O	V	E	C	R	Y	P	T	O	L	O	G	Y
Keystream	M	A	T	H	M	A	T	H	M	A	T	H	M	A	T
Plaintext	8	11	14	21	4	2	17	24	15	19	14	11	14	6	24
Keystream	12	0	19	7	12	0	19	7	12	0	19	7	12	0	19
Ciphertext	20	11	7	2	16	2	10	3	1	19	7	18	0	6	17

Example

Plaintext	I	L	O	V	E	C	R	Y	P	T	O	L	O	G	Y
Keystream	M	A	T	H	M	A	T	H	M	A	T	H	M	A	T
Plaintext	8	11	14	21	4	2	17	24	15	19	14	11	14	6	24
Keystream	12	0	19	7	12	0	19	7	12	0	19	7	12	0	19
Ciphertext	20	11	7	2	16	2	10	3	1	19	7	18	0	6	17
Ciphertext	U	L	H	C	Q	C	K	D	B	T	H	S	A	G	R

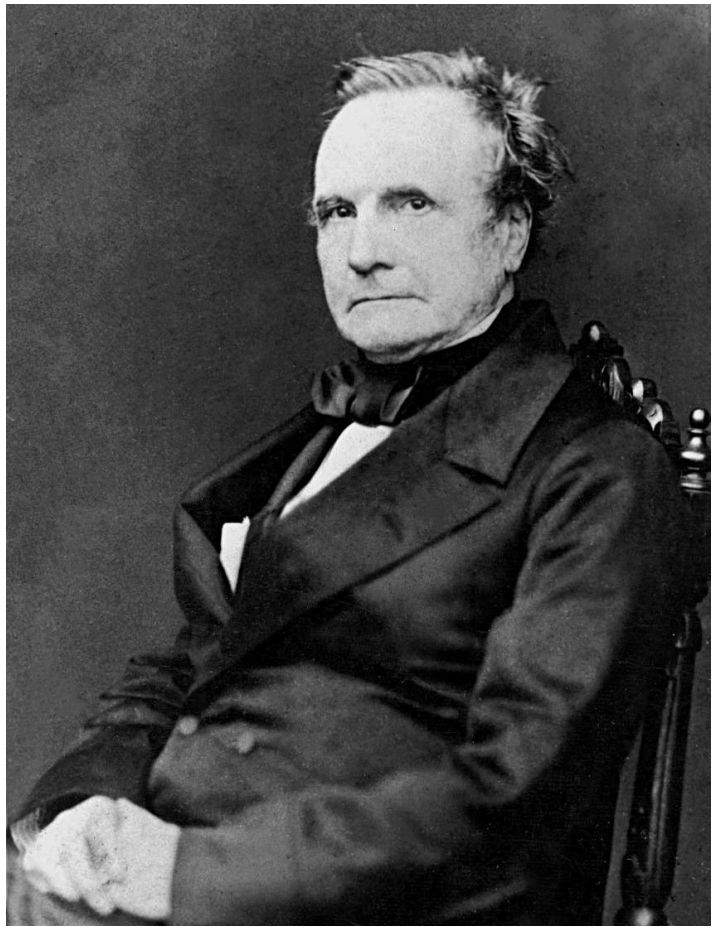
Example

Plaintext	I	L	O	V	E	C	R	Y	P	T	O	L	O	G	Y
Keystream	M	A	T	H	M	A	T	H	M	A	T	H	M	A	T
Plaintext	8	11	14	21	4	2	17	24	15	19	14	11	14	6	24
Keystream	12	0	19	7	12	0	19	7	12	0	19	7	12	0	19
Ciphertext	20	11	7	2	16	2	10	3	1	19	7	18	0	6	17
Ciphertext	U	L	H	C	Q	C	K	D	B	T	H	S	A	G	R

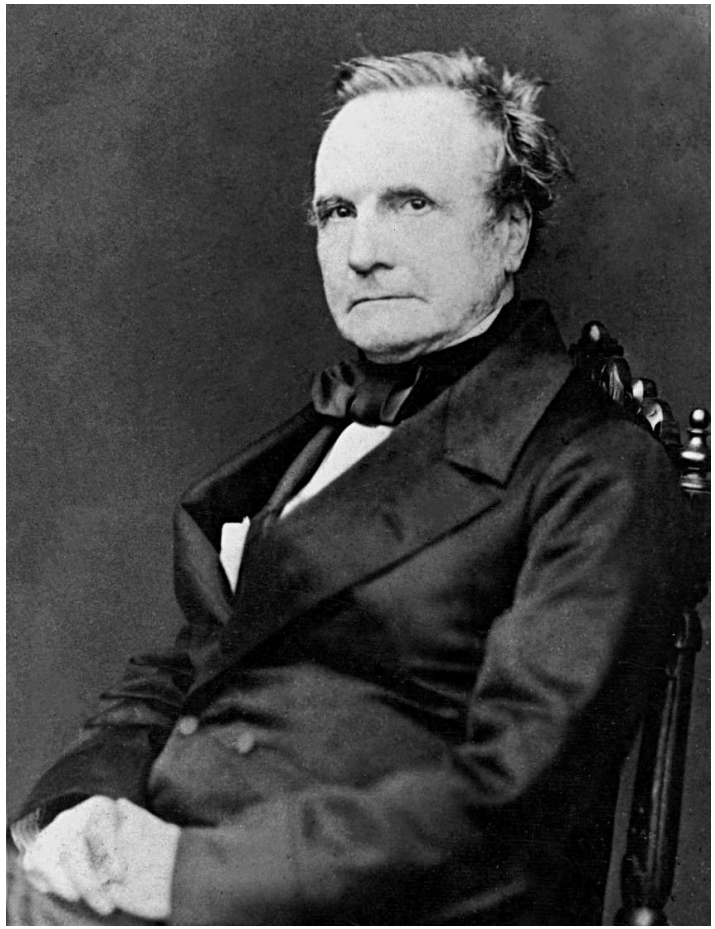
```
ILOVECRYPTOLOGY
ULHCQCKDBTHSAGR
```

The Kasiski Method

The Kasiski Method

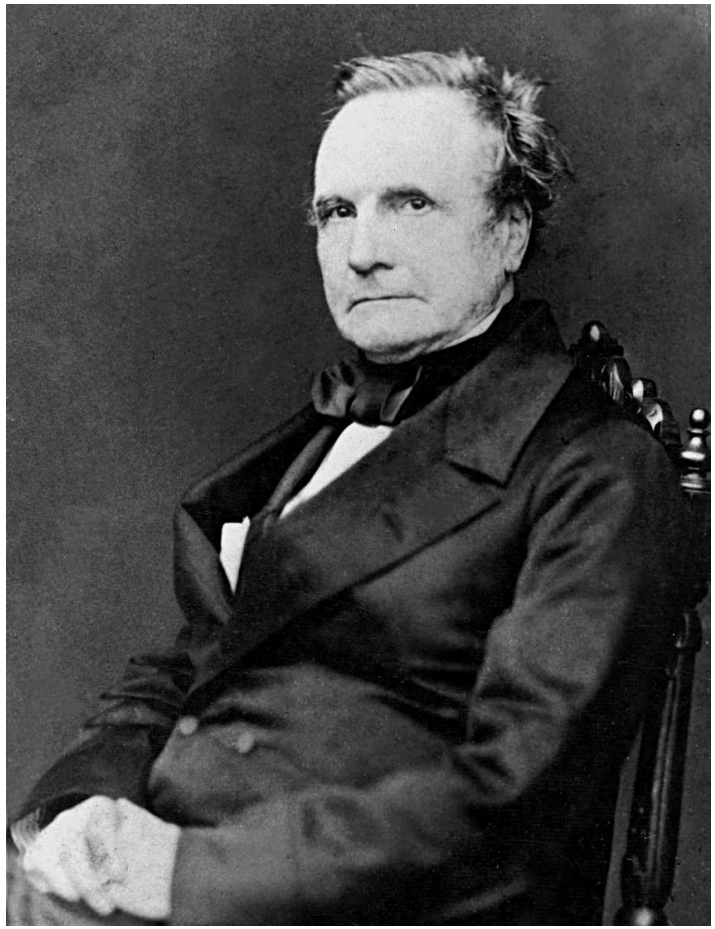


The Kasiski Method



The Kasiski method was first discovered by Charles Babbage in 1854.

The Kasiski Method



The Kasiski method was first discovered by Charles Babbage in 1854.

It was first published by Friedrich Kasiski in 1863.

```
zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
zvreg kwivs saolt nliuw oldie aqewf iiykh bjowr hdogc qhkwa
jyagg emisr zqoqh oavlk bjofr ylvps rtgiu avmsw lzgms evwpc
dmjsv jqbrn klpcf iowhv kxjbj pmfkr qthtk ozrgq ihbmq sbivd
ardym qmpbu nivxm tzwqv gefjh ucbor vwpcd xuwft qmoow jipds
fluqm oeavl jgqea lrkti wvext vkrrg xani
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
1: zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz yrigz h
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
1: zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz yrigz h
```

1 coincidence

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
1: zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz yrigz h
```

1 coincidence

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
2:  zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr zyrig zh
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
1: zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz yrigz h
```

1 coincidence

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
2:  zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr zyrig zh
```

1 coincidence

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
1: zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz yrigz h
1 coincidence
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
2:  zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr zyrig zh
1 coincidence
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
3:   zp gdlrj lajkp ylxzp yyglr jgdlr zhzqy jzqre pvmsw rzyri gzh
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
1: zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz yrigz h
1 coincidence
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
2:  zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr zyrig zh
1 coincidence
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
3:   zp gdlrj lajkp ylxzp yyglr jgdlr zhzqy jzqre pvmsw rzyri gzh
3 coincidences
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
1: zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz yrigz h
1 coincidence
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
2:  zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr zyrig zh
1 coincidence
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
3:   zp gdlrj lajkp ylxzp yyglr jgdlr zhzqy jzqre pvmsw rzyri gzh
3 coincidences
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
4:    z pgdlr jlajk pylxz pyygl rjgdl rzhzq yjzqr epvms wrzyr
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
1: zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz yrigz h
```
1 coincidence
```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
2:  zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr zyrig zh
```
1 coincidence
```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
3:   zp gdlrj lajkp ylxzp yyglr jgdlr zhzqy jzqre pvmsw rzyri gzh
```
3 coincidences
```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
4:    z pgdlr jlajk pylxz pyygl rjgdl rzhzq yjzqr epvms wrzyr
```
3 coincidences

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
1: zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz yrigz h
1 coincidence
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
2:  zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr zyrig zh
1 coincidence
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
3:   zp gdlrj lajkp ylxzp yyglr jgdlr zhzqy jzqre pvmsw rzyri gzh
3 coincidences
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
4:    z pgdlr jlajk pylxz pyygl rjgdl rzhzq yjzqr epvms wrzyr
3 coincidences
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
5:      zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
5:      zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy
```

5 coincidences

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
5:      zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy
```

5 coincidences

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
6:       zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
5:      zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy
```

5 coincidences

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
6:       zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz
```

3 coincidences

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
5:      zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy
5 coincidences
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
6:       zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz
3 coincidences
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
7:        zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
5:      zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy
```

5 coincidences

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
6:       zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz
```

3 coincidences

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
7:        zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr
```

2 coincidences

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
5:      zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy
5 coincidences
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
6:       zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz
3 coincidences
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
7:        zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr
2 coincidences
```

Shift	1	2	3	4	5	6	7	8	9
Coincidences	6	6	9	5	8	13	15	11	11

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
5:      zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy
5 coincidences
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
6:       zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz
3 coincidences
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
7:        zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr
2 coincidences
```

Shift	1	2	3	4	5	6	7	8	9
Coincidences	6	6	9	5	8	13	15	11	11

Trigrams

Trigrams

Trigram	Places	Offset	Trigram	Places	Offset
avl	117 and 258	$141 = 3 \cdot 47$	bjo	86 and 121	$35 = 5 \cdot 7$
dlr	4 and 25	$21 = 3 \cdot 7$	gdl	3 and 24	$16 = 2^4$
lrj	5 and 21	$98 = 2 \cdot 7^2$	msw	40 and 138	$84 = 2^2 \cdot 3 \cdot 7$
pcd	149 and 233	$13 = 13$	qmo	241 and 254	$98 = 2 \cdot 7^2$
vms	39 and 137	$84 = 2^2 \cdot 3 \cdot 7$	vwp	147 and 231	$84 = 2^2 3 \cdot 7$
wpc	148 and 232	$21 = 3 \cdot 7$	zhz	28 and 49	$21 = 3 \cdot 7$

Trigrams

Trigram	Places	Offset	Trigram	Places	Offset
avl	117 and 258	$141 = 3 \cdot 47$	bjo	86 and 121	$35 = 5 \cdot 7$
dlr	4 and 25	$21 = 3 \cdot 7$	gdl	3 and 24	$16 = 2^4$
lrj	5 and 21	$98 = 2 \cdot 7^2$	msw	40 and 138	$84 = 2^2 \cdot 3 \cdot 7$
pcd	149 and 233	$13 = 13$	qmo	241 and 254	$98 = 2 \cdot 7^2$
vms	39 and 137	$84 = 2^2 \cdot 3 \cdot 7$	vwp	147 and 231	$84 = 2^2 3 \cdot 7$
wpc	148 and 232	$21 = 3 \cdot 7$	zhz	28 and 49	$21 = 3 \cdot 7$

It looks like the offset is 7.

Using the Index of Coincidence

Using the Index of Coincidence

Shift	indices

Using the Index of Coincidence

Shift	indices	
2	.038	0.40

Using the Index of Coincidence

Shift	indices			
2	.038	0.40		
3	0.39	0.42	0.38	

Using the Index of Coincidence

Shift	indices				
2	.038	0.40			
3	0.39	0.42	0.38		
4	0.34	0.42	0.39	0.35	

Using the Index of Coincidence

Shift	indices				
2	.038	0.40			
3	0.39	0.42	0.38		
4	0.34	0.42	0.39	0.35	
5	0.38	0.39	0.43	0.28	0.36

Using the Index of Coincidence

Shift	indices					
2	.038	0.40				
3	0.39	0.42	0.38			
4	0.34	0.42	0.39	0.35		
5	0.38	0.39	0.43	0.28	0.36	
6	0.38	0.40	0.39	0.38	0.32	0.33

Using the Index of Coincidence

Shift	indices						
2	.038	0.40					
3	0.39	0.42	0.38				
4	0.34	0.42	0.39	0.35			
5	0.38	0.39	0.43	0.28	0.36		
6	0.38	0.40	0.39	0.38	0.32	0.33	
7	0.62	0.57	0.65	0.60	0.60	0.64	0.64

Using the Index of Coincidence

Shift	indices							
2	.038	0.40						
3	0.39	0.42	0.38					
4	0.34	0.42	0.39	0.35				
5	0.38	0.39	0.43	0.28	0.36			
6	0.38	0.40	0.39	0.38	0.32	0.33		
7	0.62	0.57	0.65	0.60	0.60	0.64	0.64	
8	0.37	0.29	0.38	0.33	0.34	0.57	0.40	0.39

Using the Index of Coincidence

Shift	indices							
2	.038	0.40						
3	0.39	0.42	0.38					
4	0.34	0.42	0.39	0.35				
5	0.38	0.39	0.43	0.28	0.36			
6	0.38	0.40	0.39	0.38	0.32	0.33		
7	0.62	0.57	0.65	0.60	0.60	0.64	0.64	
8	0.37	0.29	0.38	0.33	0.34	0.57	0.40	0.39

Frequency Counts on Substrings

Frequency Counts on Substrings

```
zlxrh rrhwl oehdw eokli lwvlh phqby nwhwf julrx x
```

Frequency Counts on Substrings

```
zlxrh rrhwl oehdw eokli lwvlh phqby nwhwf julrx x
```

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Frequency	0	1	0	1	2	1	0	6	1	1	1	6	0	1	2	1	1	4	0	0	1	1	5	3	1	1

Frequency Counts on Substrings

```
zlxrh rrhwl oehdw eokli lwvlh phqby nwhwf julrx x
```

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Frequency	0	1	0	1	2	1	0	6	1	1	1	6	0	1	2	1	1	4	0	0	1	1	5	3	1	1

Frequency Counts on Substrings

```
zlxrh rrhwl oehdw eokli lwvlh phqby nwhwf julrx x
```

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Frequency	0	1	0	1	2	1	0	6	1	1	1	6	0	1	2	1	1	4	0	0	1	1	5	3	1	1

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

Or

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

Frequency Counts on Substrings

```
zlxrh rrhwl oehdw eokli lwvlh phqby nwhwf julrx x
```

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Frequency	0	1	0	1	2	1	0	6	1	1	1	6	0	1	2	1	1	4	0	0	1	1	5	3	1	1

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

Or

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

Mutual Index of Coincidence

Definition

Let $\mathbf{s} = c_1 c_2 \ldots c_n, \mathbf{t} = d_1 d_2 \ldots d_m$ *be two strings of letters. Then we define the mutual index of coincidence to be* $\mathrm{MutIndCo}(\mathbf{s}, \mathbf{t})$*, the chance that a randomly selected letter of* $\mathbf{s}$ *is the same as a randomly selected letter of* $\mathbf{t}$*.*

Mutual Index of Coincidence

Proposition

Let $\mathbf{s} = c_1 c_2 \dots c_n, \mathbf{t} = d_1 d_2 \dots d_m$ *be two strings of letters, and let* $F_i(\mathbf{s})$ *be the number of times the* i*th letter appears in the string* $\mathbf{s}$*. Then:*

Mutual Index of Coincidence

Proposition

Let $\mathbf{s} = c_1 c_2 \dots c_n, \mathbf{t} = d_1 d_2 \dots d_m$ *be two strings of letters, and let* $F_i(\mathbf{s})$ *be the number of times the* i*th letter appears in the string* $\mathbf{s}$*. Then:*

1. $$\text{MutIndCo}(\mathbf{s}, \mathbf{t}) = \frac{1}{nm} \sum_{i=0}^{25} F_i(\mathbf{s}) F_i(\mathbf{t}).$$

Mutual Index of Coincidence

Proposition

Let $\mathbf{s} = c_1 c_2 \dots c_n, \mathbf{t} = d_1 d_2 \dots d_m$ *be two strings of letters, and let* $F_i(\mathbf{s})$ *be the number of times the* i*th letter appears in the string* $\mathbf{s}$*. Then:*

1. $$\text{MutIndCo}(\mathbf{s}, \mathbf{t}) = \frac{1}{nm} \sum_{i=0}^{25} F_i(\mathbf{s}) F_i(\mathbf{t}).$$
2. *If the letters of* $\mathbf{s}$ *and* $\mathbf{t}$ *are drawn from the same distribution, given by taking English frequencies and permuting the letters, then* $\text{MutIndCo}(\mathbf{s}, \mathbf{t}) \approx .068$.

Mutual Index of Coincidence

Proposition

Let $\mathbf{s} = c_1 c_2 \ldots c_n, \mathbf{t} = d_1 d_2 \ldots d_m$ *be two strings of letters, and let* $F_i(\mathbf{s})$ *be the number of times the* i*th letter appears in the string* $\mathbf{s}$*. Then:*

1. $$\text{MutIndCo}(\mathbf{s}, \mathbf{t}) = \frac{1}{nm} \sum_{i=0}^{25} F_i(\mathbf{s}) F_i(\mathbf{t}).$$
2. *If the letters of* $\mathbf{s}$ *and* $\mathbf{t}$ *are drawn from the same distribution, given by taking English frequencies and permuting the letters, then* $\text{MutIndCo}(\mathbf{s}, \mathbf{t}) \approx .068$*.*
3. *If the letters of* $\mathbf{s}$ *and* $\mathbf{t}$ *are drawn from different such distributions, then* $\text{MutIndCo}(\mathbf{s}, \mathbf{t}) \approx .038$*.*

i	j	σ	MutIndCo$(i, j+\sigma)$	Relative shift equation
1	3	1	.067	$\beta_1 - \beta_3 = 1$
3	7	10	.069	$\beta_3 - \beta_7 = 10$
1	4	19	.071	$\beta_1 - \beta_4 = 19$
1	6	16	.071	$\beta_1 - \beta_6 = 16$
3	4	18	.073	$\beta_3 - \beta_4 = 18$
3	5	24	.067	$\beta_3 - \beta_5 = 24$
3	6	15	.074	$\beta_3 - \beta_6 = 15$
4	6	23	.066	$\beta_4 - \beta_6 = 23$
4	7	18	.071	$\beta_4 - \beta_7 = 18$
6	7	21	.069	$\beta_6 - \beta_7 = 21$

$$\beta_3 = \beta_1 + 25$$
$$\beta_6 = \beta_1 + 10$$
$$\beta_5 = \beta_3 + 2 = \beta_1 + 1$$

$$\beta_4 = \beta_1 + 7$$
$$\beta_7 = \beta_3 + 16 = \beta_1 + 15$$

$$\beta_3 = \beta_1 + 25 \qquad \beta_4 = \beta_1 + 7$$
$$\beta_6 = \beta_1 + 10 \qquad \beta_7 = \beta_3 + 16 = \beta_1 + 15$$
$$\beta_5 = \beta_3 + 2 = \beta_1 + 1$$

$$\text{MutIndCo}(2, 4 + 24) = .061 \Rightarrow \beta_2 = \beta_4 + 24 = \beta_1 + 5$$

$$\beta_3 = \beta_1 + 25 \qquad \beta_4 = \beta_1 + 7$$
$$\beta_6 = \beta_1 + 10 \qquad \beta_7 = \beta_3 + 16 = \beta_1 + 15$$
$$\beta_5 = \beta_3 + 2 = \beta_1 + 1$$

$$\text{MutIndCo}(2, 4 + 24) = .061 \Rightarrow \beta_2 = \beta_4 + 24 = \beta_1 + 5$$

$$\text{Key} = \text{AFZHBKP}$$

$$\beta_3 = \beta_1 + 25 \qquad \beta_4 = \beta_1 + 7$$
$$\beta_6 = \beta_1 + 10 \qquad \beta_7 = \beta_3 + 16 = \beta_1 + 15$$
$$\beta_5 = \beta_3 + 2 = \beta_1 + 1$$

$$\text{MutIndCo}(2, 4 + 24) = .061 \Rightarrow \beta_2 = \beta_4 + 24 = \beta_1 + 5$$

$$\text{Key} = \text{AFZHBKP} + \text{shift}$$

β_1	Keyword	Potential plaintext
0	AFZHBKP	`zkhwkhulvkdoowxuq`

β_1	Keyword	Potential plaintext
0	AFZHBKP	`zkhwkhulvkdoowxuq`
1	BGAICLQ	`yjgvjgtkujcnnvwtp`

β_1	Keyword	Potential plaintext
0	AFZHBKP	`zkhwkhulvkdoowxuq`
1	BGAICLQ	`yjgvjgtkujcnnvwtp`
2	CHBJDMR	`xifuifsjtibmmuvso`

β_1	Keyword	Potential plaintext
0	AFZHBKP	`zkhwkhulvkdoowxuq`
1	BGAICLQ	`yjgvjgtkujcnnvwtp`
2	CHBJDMR	`xifuifsjtibmmuvso`
3	DICKENS	`whetherishallturn`

β_1	Keyword	Potential plaintext
0	AFZHBKP	`zkhwkhulvkdoowxuq`
1	BGAICLQ	`yjgvjgtkujcnnvwtp`
2	CHBJDMR	`xifuifsjtibmmuvso`
3	DICKENS	`whetherishallturn`
4	EJDLFOT	`vgdsgdqhrgzkkstqm`

β_1	Keyword	Potential plaintext
0	AFZHBKP	zkhwkhulvkdoowxuq
1	BGAICLQ	yjgvjgtkujcnnvwtp
2	CHBJDMR	xifuifsjtibmmuvso
3	DICKENS	whetherishallturn
4	EJDLFOT	vgdsgdqhrgzkkstqm
5	FKEMGPU	ufcrfcpgqfyjjrspl

β_1	Keyword	Potential plaintext
0	AFZHBKP	zkhwkhulvkdoowxuq
1	BGAICLQ	yjgvjgtkujcnnvwtp
2	CHBJDMR	xifuifsjtibmmuvso
3	DICKENS	whetherishallturn
4	EJDLFOT	vgdsgdqhrgzkkstqm
5	FKEMGPU	ufcrfcpgqfyjjrspl
6	GLFNHQV	tebqebofpexiiqrok

β_1	Keyword	Potential plaintext
0	AFZHBKP	zkhwkhulvkdoowxuq
1	BGAICLQ	yjgvjgtkujcnnvwtp
2	CHBJDMR	xifuifsjtibmmuvso
3	DICKENS	whetherishallturn
4	EJDLFOT	vgdsgdqhrgzkkstqm
5	FKEMGPU	ufcrfcpgqfyjjrspl
6	GLFNHQV	tebqebofpexiiqrok

```
wheth erish alltu rnout tobet heher oofmy ownli feorw hethe
rthat stati onwil lbehe ldbya nybod yelse these pages musts
howto begin mylif ewith thebe ginni ngofm ylife ireco rdtha
tiwas borna sihav ebeen infor medan dbeli eveon afrid ayatt
welve ocloc katni ghtit wasre marke dthat thecl ockbe ganto
strik eandi began tocry simul taneo usly
```

```
wheth erish alltu rnout tobet heher oofmy ownli feorw hethe
rthat stati onwil lbehe ldbya nybod yelse these pages musts
howto begin mylif ewith thebe ginni ngofm ylife ireco rdtha
tiwas borna sihav ebeen infor medan dbeli eveon afrid ayatt
welve ocloc katni ghtit wasre marke dthat thecl ockbe ganto
strik eandi began tocry simul taneo usly
```

"Whether I shall turn out to be the hero of my own life, or whether that station will be held by anybody else, these pages must show. To begin my life with the beginning of my life, I record that I was born (as I have been informed and believe) on a Friday, at twelve oclock at night. It was remarked that the clock began to strike, and I began to cry, simultaneously."

```
wheth erish alltu rnout tobet heher oofmy ownli feorw hethe
rthat stati onwil lbehe ldbya nybod yelse these pages musts
howto begin mylif ewith thebe ginni ngofm ylife ireco rdtha
tiwas borna sihav ebeen infor medan dbeli eveon afrid ayatt
welve ocloc katni ghtit wasre marke dthat thecl ockbe ganto
strik eandi began tocry simul taneo usly
```

"Whether I shall turn out to be the hero of my own life, or whether that station will be held by anybody else, these pages must show. To begin my life with the beginning of my life, I record that I was born (as I have been informed and believe) on a Friday, at twelve oclock at night. It was remarked that the clock began to strike, and I began to cry, simultaneously."

From *David Copperfield*, by Charles Dickens

Definition

A crib is a known or guessed portion of the plaintext, which can be used to help cryptanalyze a ciphertext.

OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW

OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD IRF SNI JAM GPW

Let's guess the word "the" is in the message somewhere.

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

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

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

Let's assume the "aso" was real and see what we can conclude.

Key Length of Four

Key Length of Four

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
--- --- --- --- --- --f bn- the -as o-- --- --- --- --- ---
--- --- --- ---
--- --- --- --- --- --t he- aso -gu s-- --- --- --- --- ---
--- --- --- ---
```

Key Length of Four

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
--- --- --- --- --- --f bn- the -as o-- --- --- --- --- ---
--- --- --- ---
--- --- --- --- --- --t he- aso -gu s-- --- --- --- --- ---
--- --- --- ---
```

"fbn" isn't very likely.

Key Length of Five

Key Length of Five

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
--- --- --- --- --- -er e-- the --a so- --- --- --- --- ---
--- --- --- ---
--- --- --- --- --- -th e-- aso --m ob- --- --- --- --- ---
--- --- --- ---
```

Key Length of Five

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
--- --- --- --- --- -er e-- the --a so- --- --- --- --- ---
--- --- --- ---
--- --- --- --- --- -th e-- aso --m ob- --- --- --- --- ---
--- --- --- ---
```

This looks better...

Key Length of Five

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
--- --- ono --m di- -er e-- the --a so- -mo b-- auo --- ---
--- --- --- ---
--- --- mdi --e re- -th e-- aso --m ob- -au o-- ncj --- ---
--- --- --- ---
```

This looks better...but this doesn't.

Key Length of Six

Key Length of Six

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
--- --- --- --- --- gqu --- the --- aso --- --- --- --- ---
--- --- --- ---
--- --- --- --- --- the --- aso --- gxw --- --- --- --- ---
--- --- --- ---
```

Key Length of Six

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
--- --- --- --- --- gqu --- the --- aso --- --- --- --- ---
--- --- --- ---
--- --- --- --- --- the --- aso --- gxw --- --- --- --- ---
--- --- --- ---
```

"gxw" and "gqu" both look bad.

Key Length of Six

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
--- --- --- --- --- gqu --- the --- aso --- --- --- --- ---
--- --- --- ---
--- --- --- --- --- the --- aso --- gxw --- --- --- --- ---
--- --- --- ---
```

"gxw" and "gqu" both look bad.

We could keep trying longer keywords. We won't get anywhere.

A new offset?

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

A new offset?

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

Key Length of Four

Key Length of Four

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
--- --- --- -wj q-t he- est --- --- --- --- --- --- --- ---
--- --- --- ---
--- --- --- -th e-e st- ezr --- --- --- --- --- --- --- ---
--- --- --- ---
```

Key Length of Four

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
--- --- --- -wj q-t he- est --- --- --- --- --- --- --- ---
--- --- --- ---
--- --- --- -th e-e st- ezr --- --- --- --- --- --- --- ---
--- --- --- ---
```

Nope.

Key Length of Five

Key Length of Five

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
--- --- --- tim --t he- -es t-- --- --- --- --- --- --- ---
--- --- --- ---
--- --- --- the --e st- -ns a-- --- --- --- --- --- --- ---
--- --- --- ---
```

Key Length of Five

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
--- --- --- tim --t he- -es t-- --- --- --- --- --- --- ---
--- --- --- ---
--- --- --- the --e st- -ns a-- --- --- --- --- --- --- ---
--- --- --- ---
```

Promising....

Key Length of Five

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
--- --s o-- tim --t he- -es t-- nsa --- --- --- --- --- ---
--- --- --- ---
so- --i m-- the --e st- -ns a-- com --- --- --- --- --- ---
--- --- --- ---
```

Promising....

Key Length of Five

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
wa- --s o-- tim --t he- -es t-- nsa --c om- -ic a-- dan --h
ea- -we r-- res
so- --i m-- the --e st- -ns a-- com --i ca- -da n-- hea --w
er- -re s-- ple
```

Promising....And now it's a fill-in-the-blank puzzle.

Key Length of Five

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
wa- --s o-- tim --t he- -es t-- nsa --c om- -ic a-- dan --h
ea- -we r-- res
so- --i m-- the --e st- -ns a-- com --i ca- -da n-- hea --w
er- -re s-- ple
```

Promising....And now it's a fill-in-the-blank puzzle.

Key Length of Five

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
wa- --s o-- tim --t he- -es t-- nsa --c om- -ic a-- dan --h
ea- -we r-- res
so- --i m-- the --e st- -ns a-- com --i ca- -da n-- hea --w
er- -re sam ple
```

Promising....And now it's a fill-in-the-blank puzzle.

Key Length of Five

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
wa- --s o-- tim --t he- -es t-- nsa --c om- -ic a-- dan --h
eaf swe raa res
so- --i m-- the --e st- -ns a-- com --i ca- -da n-- hea --w
era are sam ple
```

Key Length of Five

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
wa- --s o-- tim --t he- -es t-- nsa --c om- -ic a-- dan --h
eaf swe raa res
so- --i m-- the --e st- -ns a-- com --i ca- -da n-- hea --w
era are sam ple
```

Doesn't look so good.

Key Length of Five

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
wa- --s o-- tim --t he- -es t-- nsa --c om- -ic a-- dan --h
ea- -we r-- res
so- --i m-- the --e st- -ns a-- com --i ca- -da n-- hea --w
er- -re sim ple
```

Let's try "simple" instead.

Key Length of Five

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
wa- --s o-- tim --t he- -es t-- nsa --c om- -ic a-- dan --h
ea- -we rsa res
so- --i m-- the --e st- -ns a-- com --i ca- -da n-- hea --w
ers are sim ple
```

Let's try "simple" instead.

Key Length of Five

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
wa- --s o-- tim --t he- -es t-- nsa --c om- -ic a-- dan --h
ean swe rsa res
so- --i m-- the --e st- -ns a-- com --i ca- -da n-- hea nsw
ers are sim ple
```

Let's try "simple" instead.

Key Length of Five

```
OOF IKA AQW MPQ UMX ZXY IRK TZS PGM GPK QMI PLC NWX KEN QLD
IRF SNI JAM GPW
wat ers ome tim est heq ues tio nsa rec omp lic ate dan dth
ean swe rsa res
som eti mes the que sti ons are com pli cat eda ndt hea nsw
ers are sim ple
```

Done!

"Sometimes the questions are complicated, and the answers are simple."
Theodor Geisel

"Sometimes the questions are complicated, and the answers are simple."
Theodor Geisel a.k.a. Dr. Seuss