

**EXAMPLE** (From “*Computer Security: Art and Science*” by Matt Bishop, Chapter 9, Page 224)

Consider the Vigenere cipher

```
ADQYS  MIUSB  OXKKT  MIBHK  IZOOO  EQOOG  IFBAG  KAUMF
VVTAA  CIDTW  MOCIO  EQOOG  BMBFV  ZGGWP  CIEKQ  HSNEW
VECNE  DLA AV  RWKXS  VNSVP  HCEUT  QOIOF  MEGJS  WTPCH
AJMOC  HIUIX
```

Index of Coincidence: 0.043

Assume key length is greater than 1 and apply Kasiski method.

Letters	Start	End	Gap Length	Factors of Gap length
MI	5	15	10	2, 5
OO	22	27	5	5
OEQOOG	24	54	30	<u>2, 3</u> , 5
FV	39	63	24	2, 2, <u>2, 3</u>
AA	43	87	44	2, 2, 11
MOC	50	122	72	2, 2, <u>2, 3</u> , 3
QO	56	105	49	7, 7
PC	69	117	48	2, 2, 2, <u>2, 3</u>
NE	77	83	6	<u>2, 3</u>
SV	94	97	3	3
CH	118	124	6	<u>2, 3</u>

Longest Repetition: 6 characters long with gap of 30

Next longest Repetition: 3 characters with gap of 72

GCD of 30 and 72 = 6

Of 11 repetitions, six have gaps with factor 6

Arrange message into 6 columns:

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

---

Each column represents one alphabet. The indices of coincidence are as follows:

Column #1: IC = 0.069	Column #2: IC = 0.078	Column #3: IC = 0.078
Column #4: IC = 0.056	Column #5: IC = 0.124	Column #6: IC = 0.043

All indices of coincidence represent one alphabet except for the ICs associated with columns #4 (period between 1 and 2) and #6 (period between 5 and 10). Using the statistical nature of the measure, we assume that the key length 6.

Counting the characters in each column:

Column	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
1	3	1	0	0	4	0	1	1	3	0	1	0	0	1	3	0	0	1	1	2	0	0	0	0	0	0
2	1	0	0	2	2	2	1	0	0	1	3	0	1	0	0	0	0	1	0	4	0	4	0	0	0	0
3	1	2	0	0	0	0	0	2	0	1	1	4	0	0	0	4	0	1	3	0	2	1	0	0	0	0
4	2	1	1	0	2	2	0	1	0	0	0	0	1	0	4	3	1	0	0	0	0	0	0	2	1	1
5	1	0	5	0	0	0	2	1	2	0	0	0	0	0	5	0	0	0	3	0	0	2	0	0	0	0
6	0	1	1	1	0	0	2	2	3	1	1	0	1	2	1	0	0	0	0	0	0	3	0	1	0	1

Unshifted alphabets have following frequency characteristics

H M M M H M M H H M M M M H H M L H H H M L L L L L

H meaning HIGH Frequency

M meaning MEDIUM Frequency

L meaning Low Frequency

Compare the frequency counts in the six alphabets above with the frequency count of the unshifted alphabet.

ADIYS RIUKB OCKKL MIGHK AZOTO EIOOL IFTAG PAUEF  
 VATAS CIITW EOCNO EIOOL BMTFV EGGOP CNEKI HSSEW  
 NECSE DDAAA RWCXS ANSNP HHEUL QONOF EEGOS WLPCM  
 AJEOC MIUAX

In the last line, the group **AJE** suggests the word **ARE**. Assuming this, second alphabet maps A into S. Substituting back produces

ALIYS RICKB OCKSL MIGH S AZOTO MIOOL INTAG PACEF  
 VATIS CIITE EOCNO MIOOL BUTFV EGOOP CNESI HSSEE  
 NECSE LDAAA RECXS ANANP HHECL QONON EEGOS ELPCM  
 AREOC MICAX

The last block suggests MICAL, because AL is a common ending for adjectives. This means that the 4<sup>th</sup> alphabet maps O into A, and the cipher becomes

ALIMS RICKP OCKSL AIGHS ANOTO MICOL INTOG PACET  
 VATIS QIITE ECCNO MICOL BUTTV EGOOD CNESI VSSEE  
 NSCSE LDOAA RECLS ANAND HHECL EONON ESGOS ELDCM  
 AREOC MICAL

In English, a Q is always followed by a U, so the I in the second group of the second line must map to U. The fifth alphabet maps M to A. The cipher is solved:

ALIME RICKP ACKSL AUGHS ANATO MICAL INTOS PACET  
 HATIS QUITE ECCNO MICAL BUTTH EGOOD ONESI VESEE  
 NSOSE LDOMA RECLE ANAND THECL EANON ESSOS ELDOM  
 AREOC MICAL

With proper spacing and punctuation, we have

A LIMERICK PACKS LAUGHS ANATOMICAL  
 INTO SPACE THAT IS QUITE ECONOMICAL  
 BUT THE GOOD ONES I'VE SEEN  
 SO SELDOM ARE CLEAN,  
 AND THE CLEAN ONES SO SELDOM ARE COMICAL.

The key is ASIMOV.