Example of Breaking a Vigenère cipher

```
FZFGWBOPFW LWKRASUQSY JHSIJDHFVW
ICCWAYHFRY GMEIJXWPXW WCKXZJPXRC
FBASXMOSMF LBLXZNBDXG ICLRUJCOXO
NQBWZJVXHH JSMIVNBQSL MSYSGPVBVK
NGQIJBOPVW FRFRYGIQML MOARGUWZXM
WSPSJHCKZW WGXXATBPMF NHXRVBVXXA
XHEIMXSLJS GCLOLMCRKZ YOIMUJKFXZ
TIQTAHHRVW XCOGGSJBVK FHFSFXGLWZ
JKXWUTBPMV JFFRYNBEIJ TKKQASRXWO
JZIEKXVBGG ZZAJGWHEIZ THAEQROAIZ
JFCIWQJBVQ XZBIHDOKHK YIMMVBVBXZ
JFQLWUZBEK ZFBSXROHMF LOAEAXMZLS
NBTSMQRYIO TFQLLMSQVG ZPIIGKUBXL
NBDYHFBATA HYFRYYVBHS NGFIKBVBRK
ZRAIFQMXAZ NHBVSGPFXO NHETASYBCW
XFXRUQCPIT DVBV
```

The digram HE occurs with the H in positions 182, 287 and 442. If these occurrences are encryptions of the same sequence of letters in the plain text, then the length of the keyword should divide

$$\gcd(287 - 182, 442 - 287) = \gcd(105, 155) = 5.$$

So we guess that the keyword has length five.

Rewrite the ciphertext in blocks of length five.

```
FZFGW BOPFW LWKRA SUQSY JHSIJ DHFVW
ICCWA YHFRY GMEIJ XWPXW WCKXZ JPXRC
FBASX MOSMF LBLXZ NBDXG ICLRU JCOXO
NQBWZ JVXHH JSMIV NBQSL MSYSG PVBVK
NGQIJ BOPVW FRFRY GIQML MOARG UWZXM
WSPSJ HCKZW WGXXA TBPMF NHXRV BVXXA
XHEIM XSLJS GCLOL MCRKZ YOIMU JKFXZ
TIQTA HHRVW XCOGG SJBVK FHFSF XGLWZ
JKXWU TBPMV JFFRY NBEIJ TKKQA SRXWO
JZIEK XVBGG ZZAJG WHEIZ THAEQ ROAIZ
JFCIW QJBVQ XZBIH DOKHK YIMMV BVBXZ
JFQLW UZBEK ZFBSX ROHMF LOAEA XMZLS
NBTSM QRYIO TFQLL MSQVG ZPIIG KUBXL
NBDYH FBATA HYFRY YVBHS NGFIK BVBRK
ZRAIF QMXAZ NHBVS GPFXO NHETA SYBCW
XFXRU QCPIT DVBVZ
```

Consider the Caeser cipher corresponding to letters in positions congruent to one modulo five i.e. the first letters in each block.

Positions congruent to one modulo five.

Exp Freq %	in plaintext	Obs Freq %	in ciphertext
a	8.15	0	A
b	1.37	5.3	B
c	2.21	0	C
d	4.58	3.2	D
e	12.61	0	E
f	1.86	5.3	F
g	2.36	4.3	G
h	6.85	3.2	H
i	6.97	2.2	I
j	0.14	11.8	J
k	1.07	1.1	K
l	4.37	3.2	L
m	1.96	5.3	M
n	6.52	11.8	N
o	7.58	0	O
p	1.40	1.1	P
q	0.19	4.3	Q
r	5.02	2.2	R
s	6.05	4.3	S
t	9.93	6.5	T
u	3.22	2.2	U
v	0.78	0	V
w	2.49	4.3	W
x	0.13	9.7	X
y	2.11	4.3	Y
z	0.07	4.3	Z

Positions congruent to one modulo five. Number of letters, $n = 93$.

	Exp. Freq. $100p_i$ %	Obs. Occ. a_i	Shift 0 np_i	Shift 5 np_{i-5}
A	8.15	0	7.58	0.73
B	1.37	5	1.27	2.32
C	2.21	0	2.06	0.12
D	4.58	3	4.26	1.96
E	12.61	0	11.73	0.07
F	1.86	5	1.73	7.58
G	2.36	4	2.20	1.27
H	6.85	3	6.37	2.06
I	6.97	2	6.48	4.26
J	0.14	11	0.13	11.73
K	1.07	1	1.00	1.73
L	4.37	3	4.06	2.20
M	1.96	5	1.82	6.37
N	6.52	11	6.06	6.48
O	7.58	0	7.05	0.13
P	1.40	1	1.30	1.00
Q	0.19	4	0.18	4.06
R	5.02	2	4.67	1.82
S	6.05	4	5.63	6.06
T	9.93	6	9.23	7.05
U	3.22	2	2.99	1.30
V	0.78	0	0.73	0.18
W	2.49	4	2.32	4.67
X	0.13	9	0.12	5.63
Y	2.11	4	1.96	9.23
Z	0.07	4	0.07	2.99

Chi-squared statisitic for shift of zero places

$$X_0 = \sum_{i=0}^{25} (a_i - np_i)^2 / np_i = 1949.79$$

Chi-squared statisitic for shift of five places

$$X_5 = \sum_{i=0}^{25} (a_i - np_{i-5})^2 / np_i = 23.99$$

If we calculate $X_0, X_1, \ldots, X_{25}$ we find that X_5 is by far the smallest. This indicates that the Caeser cipher for the first letter in each block is a shift five places to the right. This takes 'a' to F, so the first letter in the keyword is F.

We may use a similar analysis for the other letters in each block to deduce that the keyword is FOXES and the plain text is

> Alice was beginning to get very tired of sitting by her sister on the bank and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations in it, and "what is the use of a book," thought Alice, "without pictures or conversations?" So she was considering, in her own mind (as well as she could, for the hot day made her feel very sleepy and stupid), whether the pleasure of making a daisy-chain would be worth the trouble of getting up and picking the daisies, when suddenly a White Rabbit with pink eyes ran close by her.