We will use Kasiski's method to crack the following Vigenère cipher.

```
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 DVBV
```

Step 1. The digram `HE` occurs at positions 182, 287 and 442 in the ciphertext, and in the first two of these it is part of the trigram `HEI`. Guess: the key length is a common factor of $287-182=105$ and $442-287=155$, hence it should divide $\gcd(105,155)=5$. (I have anticipated this by writing the cipher in blocks of 5; usually Alice will not be so helpful!)

The substring consisting of letters in positions $\equiv 1 \pmod 5$:

```
FBLSJDIYGXWJFMLNIJNJJNMPNBFGMUWHWTNBXXGMYJTHXSFXJTJNTS
JXZWTRJQXDYBJUZRLXNQTMZKNFHYNBZQNGNSXQD.
```

This is a Caesar shift. We use the frequency table on the next page to determine the shift. After breaking this substring (which is a Caesar cipher) we have to do the same with substrings $\equiv 2,3,4,0 \pmod 5$, but they are similar so we don't do it here.

The second column in the table on the next page shows frequencies in standard English and the third column is the frequency of letters in substring 1. You can see that after shifting the column 2 down by 5 it frequencies more or less match those of column 3, so the shift is 5, corresponding to `F`. (See last page for the meaning of the last two columns.)

Letter	Frequency %	Observed	Expected Shift 0	Expected Shift 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	0.65
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	0.65	2.99
$\sum(o-e)^2/e$			1949.79	23.99

Table 1: A chi-squared calculation (note: the substring has length 93)

By the same method (and the results are as clear-cut in all cases), we find the shifts for the other substrings to be 14, 23, 4, 18, so that the keyword is

FOXES=(5, 14, 23, 4, 18).

Subtracting this key from the ciphertext we obtain the plaintext below. (This is the same thing as applying reverse shifts of 5, 14, 23, 4, 18 to the substrings $1, 2, \ldots, 5$, respectively, and then putting the substrings back together.)

The decrypted 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.

Chi-squared method (non-examinable, but useful, especially for Cipher Challenge)

The Chi-squared method is a statistical method that can be used to determine the shift in a Caesar cipher. You can write a simple code to do it automatically. Even though not examinable, it could be very useful for Cipher Challenge. Read pages 31-32 of Prof. Cameron's notes to see how it works.

Let me now clarify what the last two columns in the above table are. First note that the fifth column is the shift of the fourth column by 5. The values of these columns are used as the e_i values in the chi-square formula (see the notes). Finally, the fourth column is the expected frequency of letters in a standard English text of length 93. For instance, $7.58 = (8.15)(\frac{93}{100})$ is the frequency of the letter A.

Where did 93 come from? It is the length of Substring 1 above. So, what we are saying is that if Substring 1 were standard English (i.e., no shift), then we would expect to see 7.58 A's in it.