Activity title

Cracking Codes

Time required

2 hours

Activity summary

Learn how to decode and encode communications

By the end of this activity, you will be able to:

- Understand how mathematics can be used to break codes

What equipment will you need?

None

How to do it

Codes have been used throughout the ages to pass secret messages. Most modern forms of communication, such as email or instant messaging, encode their message when it is sent, and decode their message where it is received.

In languages, some letters are used more often than others. Knowing this, we can use the frequency with which letters appear to help us crack codes.

Now try this

1. **Count the frequency with which each letter occurs in the passage below. In the table on the following page, write in the frequency of each letter, then list them in order from the most to least frequently used:**

   The pollution was carried by the wind and entered the town quickly. We could all see the brown cloud travelling over from Fox Corner. Homes, schools and offices were all put onto 'High Alert' status. Gas masks were issued – even some pets were given masks.

   Despite this, twelve people have died and eighty are seriously ill. Even now, a gas haze continues to pollute the town's atmosphere.

   Our priority is to keep the event Top Secret. We know that news travels fast and above all, we don't want any of our investors to jump to the conclusion that our Brown Box process is still dangerous.

Activity title: Cracking Codes: The Rosetta Stone

Letter Frequency

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**			

Most used to least used:

2. Count the frequency of the letters in the coded message below. As in step 2, you should then list the letters in order – from the most frequently used letters to the least frequently used letters.

New message in code:

Gsv vckozmzgrlm rh gsv evhhvo drviv gsv tzhvh ziv nrcvw

gl nzpv mrgiltvm wrlcrwv szh uzrovw.

Gsv lmob yildm tzh rg xzm yv rh mrgiltvm wrlcrwv. Gsviv rh ml

Yilnrmv rm gsv kilxvhh. Dsb wrw gsv tzh evhhvo yivxp hl vzhrob.

Hint:
Your list of letters should now help you to start deciphering some of the words. E.g. which letters might represent the word ‘the’? Once you get going, you’ll find you can decode some words without having to decipher every single letter.

Activity title: Cracking Codes: The Rosetta Stone

Code Letter Frequency

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**					

Most used to least used:

Hint:

1. The word ‘gsv’ that occurs several times is probably the code for ‘the’.
2. The pollution was probably nitrogen dioxide and the words for nitrogen dioxide occur in the second and third lines.

Meaning of the message:

Activity title: Cracking Codes: The Rosetta Stone

Here are some notes from other decoders, which may help you:

1. There is a frequent 3 letter word ‘gsv’ – you can guess that this is ‘the’ so that gives the code for g = T, s = H v = E

Rewrite the passage substituting for g, s and v

2. That decodes two of the most frequent letters used – T and E

The other 3 most frequent letters used are A, O and I

‘r’ occurs 15 times in the coded passage and ‘z’ occurs 11 times – both very frequent. These could be –A and –E or –I but which way round?

Clue 1 - There is a (mostly) decoded word THE-i-E so i = R or S

Clue 2 - There are two two-letter words with the same first letter (code r) and one of them has T as the second letter.

There is no word ‘ot’ So these words must be at or it and the other word must be as or is

Taking clue 1 and 2 together assume i = R. And r =I

(could be a big clue to the nature of the code, some kids will crack it at this point)

So if r = I, then the other frequent letter z must = A

Rewrite the passage substituting for i,r and z

3. There are two words THERE and dHERE So d = W

4. -hh occurs three times in the passage, twice at the end of words and once in the middle.

Most frequent double letters at ends of longer words are –SS and –LL so h = S or L.

The word Ih occurs – line 3, no word ‘il’ in English so h = S

Rewrite the passage substituting for d and h

5. The code letter –I (el) occurs frequently, the only frequent letter that has been not been decoded is O. So I = O

6. By inference from tASES, tAS = GASES, GAS So t = G

Rewrite the passage substituting for I and t

7. Now there are words ending in –ATIOm and –ROGEm So m=N

8. From yROWN so y = B

From EckoANATION so c = X, k = P, o = L

Rewrite the passage substituting for y, c, k and o

9. Should get by now

NIgROGEN wIOXIwE = NITROGEN DIOXIDE so w = D, g = T

And even GAS eESSEL = GAS VESSEL so e = V

Rewrite the passage substituting for w, g and e

10. These words now occur - nIXED , ONLb, uAILED, WHO xHExpED IT? So n = M, b = Y, u = F, x = C,

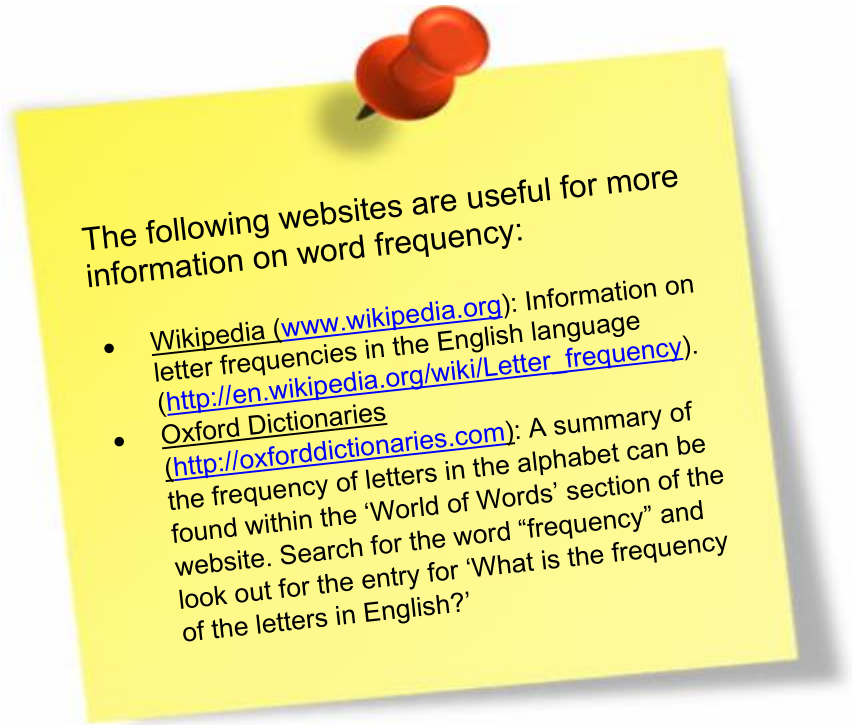
Rewrite the passage – it should now be readable.

You could also

Create your own code and write out a message – e.g. a = b, b = c, c = d, etc.
Then try using the above approach and see if it will break your code.

Further activities you could carry out

Research more information about word frequency in the English language.



Activity title: Cracking Codes: The Rosetta Stone

What results were expected?

This is the decoded code:

Gsv vckozmzgrlm rh gsv evhhvo drviv gsv tzhvh ziv nrcvw
The explanation is the vessel where the gases are mixed

gl nzpv mrgiltvm wrlcrwv szh uzrovw.
to make nitrogen dioxide has failed.

Gsv lmob yildm tzh rg xzm yv rh mrgiltvm wrlcrwv. Gsviv rh ml
The only brown gas it can be is nitrogen dioxide. There is no

Yilnrmv rm gsv kilxvhh. Dsb wrw gsv tzh evhhvo yivxp hl vzhrob.
bromine in the process. Why did the gas vessel break so easily?

Dsi xsvxpvw rg?
Who checked it?

The substitution code is simply:
abcdehijklmnopqrstuvwxyz = zyxwvutsrqponmlkjihgfedcba