Cryptology: Learn to Decipher

Grade Levels

This activity is intended for students in grades 5-7.

Objectives and Topics

Students will learn how to decipher a substitution cipher. To do this, students will need to understand how to change fractions into decimals and percentages, have basic data analysis skills, and use logical thinking.

Materials and Resources

- Encoded Message (see Notes)
- Vocabulary Sheet (see Worksheets)
- Four steps to decoding a substitution cipher (see Notes)
- One-word frequency analysis w/pie chart (see Notes)
- One-sentence frequency analysis w/frequency distribution table (see Notes)
- Frequency Activity (see Worksheet)
- Access to Excel or a spreadsheet application (optional for ease of graphing)

Activity Outline

Deciphering a Substitution Cipher

Introduce the encoded message (see Notes) to the students. Ask them to look it over and to try to decode it. No prior introduction is required; see what they can do without any help. At some point, bring the class back together and discussion how everyone tried to decipher the message. Discuss any possible solutions; see if there are any guesses.

Defining a Structure

- Handout and discuss the vocabulary list (see Worksheets). Have students become familiar with the terms. No memorization necessary, but familiarity and access to this vocabulary list is important.
- Read over the four steps to decoding a substitution cipher (see Notes)
- Go over Step 1, and show the "Honolulu" example (see Notes).
- Go over Step 2, and show the "Kumu Ali lives in Honolulu" example (see Notes).
- Break up into four groups and do the Frequency Activity (each group does a frequency distribution table of one step). (see Worksheets).

Frequency Distribution and Patterns

Once everyone is finished with the Frequency Activity, each group should go over Step 3 for their frequency distribution table (which can be created in Excel) and discuss the patterns. As a class, compare frequency distribution tables. For example, note that 'e' is the most frequent.

***Note: given any passage, paragraph, text in the English language, and the frequency distribution table will always look generally the same.

***Interesting fact: La Disparation, a 300-page French Novel, was written without the letter 'eÂ. Even more amazing, it was then translated into English, also without the letter 'e'.

Re-examine the encoded message. Create a frequency distribution table for this cipher text either as a class or in groups. Examine this table, and note the patterns. Discuss the possible substitutions and play around with the combinations. Use educated guesses to fill in the blanks.

If we get stuck, try asking these questions:

- The letters 'a' and 'd' occur the most amount of times. What letter(s) could 'a' and 'd' possibly represent? Try it out with different combinations. (Try 'e' first, then 't').
- For the word 'ymm', what could the 'mm' possibly represent? What 1-letter digraphs end a sentence?
- List all possible two letter words. What could 'yd' possibly be? What letter could possibly end a two-letter word (yd) but begin a three letter word (dra)?
- The letters 'n', 'a', show up both as 'na' (in yna) and 'an' (in lebran). What letters could possibly do this?
- The letter 'y' begins three words 'ymm', 'yna', 'yd'. What letter could 'y' possibly be? Is it more likely a vowel or consonant?
- Think of as many three-letter words as possible. There are two in this sentence that end with 'a'. What could 'a' be?
- The letter 'u' is in the middle of two three-letter words. Is 'u' more likely to be a vowel or a consonant?

Notes

Encoded Message

Here's an example of the encoded message you can present to your class:

Huf oui ymm yna aqbandc yd talutehk dra cipcdedideuh lebran!

Work through the activity to decode it yourself! To encode a message of your own, randomly assign letters to a new letter and rewrite the message using the encryption.

How to Decipher

Four Steps to Decode a Substitution Cipher:

1. We need to understand the process of changing fractions into decimals and percents before we learn to decrypt substitution ciphers.
2. How do we decrypt a secret message enciphered with a substitution cipher? We create a frequency distribution table to use frequency analysis.
3. A frequency distribution table explains how often a letter is most likely to appear in a passage.
    - The frequency distribution table above shows how often, or frequent, a letter occurs in the sentence. In other words, the whole sentence is broken down into letters (or parts) a, e, h, i, k, l, m, n, o, s, u, v. Each letter occurs constitutes a certain percent of the sentence.
4. Use patterns to identify common groups of letters and then play the guessing game.
    - Examine the table above. Which letter(s) occur(s) most often? Least often? Which letters of the alphabet do not even show up in this sentence? Which letters are more likely to occur in a sentence? Why? Which letters are least likely to occur in a sentence? Are there letter pairs (digraphs) or triples (trigraphs) that are likely to occur? How many two letter words are there? How many three letter words are there?

Notes

Honolulu Example

Step 1: We need to understand the process of changing fractions into decimals and percents before we learn to decrypt substitution ciphers.

Consider the following one-word message: HONOLULU

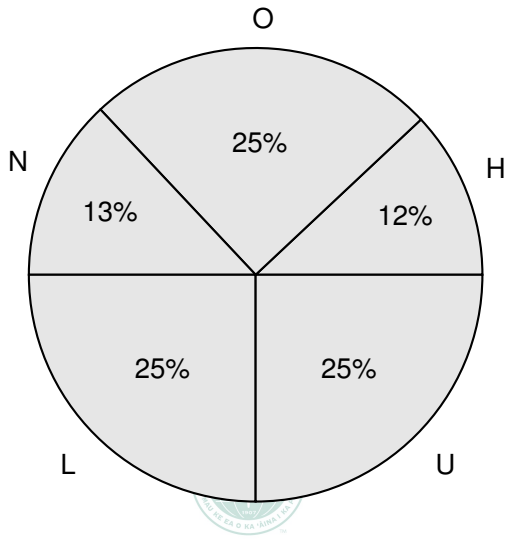
To begin deciphering substitution ciphers, we should ask ourselves the following questions:

- What is the total number of letters in this word?
- How many times does each letter occur?
- What fraction represents how many H's occur in the word HONOLULU? Express this fraction as a decimal.
- The letter 'H' occurs what percentage of the time?

The students (or you) can create the following table and fill-in the values:

	H	O	N	L	U	Total
Frequency	1	2	1	2	2	8
Fraction	$\frac{1}{8}$	$\frac{2}{8} = \frac{1}{4}$	$\frac{1}{8}$	$\frac{2}{8} = \frac{1}{4}$	$\frac{2}{8} = \frac{1}{4}$	$\frac{8}{8} = 1$
Decimal	0.125	0.25	0.125	0.25	0.25	1.0
Percent	12.5%	25%	12.5%	25%	12.5%	100%

Using Excel, another spreadsheet application, or by hand, the students (or you can walk them through it) must create the following pie chart for a visualization of the data/frequency:



Notes

Kumu Ali Example

Step 2: How do we decrypt a secret message enciphered with a substitution cipher? We create a frequency distribution table to use frequency analysis.

Consider the sentence: Kumu Ali lives in Honolulu.

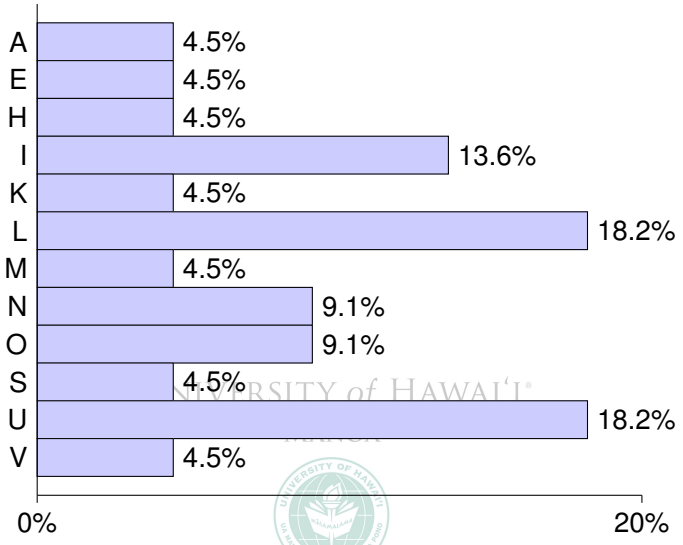
Like last time, we have similar questions to ask ourselves:

- How many letters are there total in this sentence?
- How often does each letter occur?
- What fraction of the sentence is the letter 'U'? Represent this fraction as a decimal.
- What percent of the word is the letter 'U'?

Again, the students (or you) should create the following table and fill-in the values:

	A	E	H	I	K	L	M	N	O	S	U	V	Total
Frequency	1	1	1	3	1	4	1	2	2	1	4	1	22
Fraction	$\frac{1}{22}$	$\frac{1}{22}$	$\frac{1}{22}$	$\frac{3}{22}$	$\frac{1}{22}$	$\frac{2}{11}$	$\frac{1}{22}$	$\frac{1}{11}$	$\frac{1}{11}$	$\frac{1}{22}$	$\frac{2}{11}$	$\frac{1}{22}$	$\frac{22}{22} = 1$
Decimal	0.045	0.045	0.045	0.136	0.045	0.182	0.045	0.091	0.091	0.045	0.182	0.045	$.997 \approx 1.0$
Percent	4.5%	4.5%	4.5%	13.6%	4.5%	18.2%	4.5%	9.1%	9.1%	4.5%	18.2%	4.5%	$99.7\% \approx 100\%$

Using Excel, another spreadsheet application, or by hand, the students (or you can walk them through it) should create the following frequency distribution table for a visualization of the data/frequency:



Vocabulary

Cipher: Any general system for hiding the meaning of a message by replacing each letter in the original message with another letter.

Ciphertext: The secret message that needs to be discovered through decryption.

Code: A system for hiding the meaning of a message by replacing each word or phrase in the original message with another character or set of characters.

Cryptanalysis: The study of breaking (deciphering) secret messages.

Cryptography: The study of creating (enciphering) secret messages.

Cryptology: The art of creating and breaking secret messages.

Decrypt/decipher/decode: The act of discovering the original message which has been hidden through a cipher or code.

Encrypt/encipher/encode: The act of hiding a message by using a code or cipher.

Frequency Analysis: A system for deciphering substitution ciphers by analyzing how often letters occur in a given passage.

Frequency Distribution Table: A table which displays how often each letter occurs in a passage.

Letter Frequency: How often a letter occurs in a passage

Plaintext: The message that needs to be hidden through encryption.

Substitution Cipher: A type of cipher which matches the actual alphabet with a randomly mixed alphabet.

Worksheets

Frequency Activity

In groups of 4-5, each group will take a sentence (or sentences) and create a frequency table. Here are the steps:

1. Count how many letters there are total.
2. Create a table with 28 columns and 5 rows (or use the attached table below).
3. For each letter, count how often the letter occurs in the sentence. Fill in the frequency row.
4. For the fraction row, we want to know what fraction of the sentence is made up of A's, B's, . . . , Z's.
5. For the decimal row, change the fraction into a decimal.
6. For the percent row, we want to know what percent is A's, B's, . . . , Z's in the sentence.
7. Discuss as a table what patterns you see in the table.
8. Discuss as a class overall patterns that occur in all four tables.
9. Create a frequency table of the encoded sentence: Huf oui ymm yna aqbandc yd talutehk dra cipcde-dideuh lebran!
10. Discuss as a group or class the patterns that occur in this table.
11. Mix and match, which letter occurs most often in the ciphertext? What are the possibilities for this letter in "plaintext"? Think about patterns in the English language. What possible combination of vowels could occur as a digraph? What possible combination of consonants could occur as a trigraph?

	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	Total
Frequency																											
Fraction																											
Decimal																											
Percent																											