Cryptology

Goal: To apply various mathematical concepts to the study of *cryptology*. *Cryptology* is the science of secret communication. ***When submitting this project, include all newspaper articles that you use to complete it.***

Part 1: **Determining letter and word frequencies**. In this activity you will explore the relationship between the frequency of letters occurring in a document. Find an relatively long news article in a newspaper or on the internet. Count the number of times each different letter or word occurs in your paragraph. This information is called the *frequency*. When you are finished, total the frequency column to get the number of letters or words in your article.

Letter	Frequency	Percent
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 number of letters		

Word	Frequency	Percent
a		
an		
and		
are		
as		
at		
by		
do		
for		
go		
I		
if		
in		
is		
it		
of		
on		
or		
so		
that		
the		
then		
this		
to		
too		
will		
Total number of words		

After counting the occurrences of each letter or word, calculate the percent for each one by dividing each frequency by the total number and then multiplying by 100 (calculators may be used – round to the nearest tenth of a percent).

1.	What letter occurred the most frequently? The least?	
2.	Do some letters occur with about the same frequency?	
3.	Compare the percent of A's to D's. Describe your findings.	
4.	Compare the percent of T's to C's. Describe your findings.	
5.	Are there some words not listed on the chart that occurred as frequently as some of those on the common list? Why?	
6.	How could you obtain more accurate frequencies?	

Part 2: **Making the words fit**. People who work with the printed word have unique names for the height of letters with which they work. This special measurement is called a *point*. A point is $\frac{1}{72}$ of an inch. Seventy-two points are needed to make an inch! You have probably used ten-point (font size) or twelve-point while typing on your computer. This number designates the size of the uppercase letters on your document. A pica, which is a also a unit of measure, is equal to twelve-points.

7.	What part of an inch is a twelve-point letter?	
8.	If you found and measured a three-inch uppercase letter in a newspaper advertisement, how many points would it be?	
9.	From a newspaper, cut out a word that uses an uppercase letter. Measure its height. Convert this measurement to both points and picas. Glue the word on this paper somewhere.	

Part 3: **Check out that Fog**! The Gunning-Mueller Fog Index is a technique to measure how easy a written piece is to read and understand. One early use of this index was to help newspaper writers improve what they wrote and make it easier to understand. The Fog Index gives the approximate grade level needed to read and understand the book or article.

10.	Choose a front-page newspaper article. The longer the article, the more accurate the result. Count the number of words in the article.	
11.	Count the number of sentences in the article.	
12.	Find the average sentence length by dividing the number of words by the number of sentences.	
13.	Words of three or more syllables are labeled "big" words. Use your best judgment when you count these words. Skip capitalized words, compound words (like skateboarder), and verbs with three or more syllables ending in **es** or **ed** (likes surprises or bisected).	
14.	Divide the number of hard words by the total number of words, and multiply the result by 100. This is the percent of big words.	
15.	Add the average sentence length from problem #12 to the percent of big words from problem #14.	
16.	Multiply #15 by 0.4. This is the Fog Index.	
A Fog Index of 7 or 8, for example, means that a person would have to be at the seventh or eighth grade level to read and understand the article.		

Part 4: **Mode Code**. Now that you have an understanding of how our alphabet works, you are ready to move on to cryptology, the science of secret communication. The word derives from the Greek kryptos, meaning "hidden", and logos, meaning "word". People have wanted to communicate secretly for a long time. People who decode messages are aided by the fact that in English, certain letters of the alphabet occur more frequently than others. You calculated percentages in part 1. Here are the overall percentages in ordinary written English.

E	13%	D	4%	G	1.5%
T	9%	L	3.5%	W	1.5%
A	8%	C	3%	V	1%
O	8%	M	3%	J	0.5%
N	7%	U	3%	K	0.5%
I	6.5%	F	2.5%	X	0.5%
R	6.5%	P	2%	Q	0.3%
S	6%	Y	2%	Z	0.2%
H	5.5%	B	1.5%		

17.	In a large passage written in English, which letter would you expect to be the mode?	
18.	In a large passage written in English, about what percent of the letters would you expect to be vowels, excluding Y?	
19.	How many times would you expect the letter E to occur in a message containing 1270 letters?	
20.	How many vowels, excluding Y, would you expect to occur in a message containing 1270 letters?	

The next secret message was created using numerals to replace letters of the alphabet. A given numeral in the secret message stands for the same letter each time it occurs. Determine the frequency and percentage of use (to the nearest tenth) for each numeral in the secret passage. Use tally marks to record the results in the frequency table. The secret message uses only 21 letters of the alphabet.

21.	Which numeral in the secret message is the mode?	
22.	Which numeral in the secret message occurs with the second-greatest frequency?	

SECRET MESSAGE

___	___	___	___	___	___		___	___		___	___	___	___	___
13	15	12	11	8	16		13	19		7	15	17	2	10
___	___	___	___		___	___	___	___		___	___		___	
13	6	11	18		14	19	1	17		11	18		2	
___	___	___	___	___	___		___	___	___	___		___	___	___
15	17	2	21	21	12		8	17	2	13		2	8	1
___	___	___	___	___	___	___	___	___	___	___				
11	8	13	17	15	17	18	13	11	8	16				
___	___	___		___	___		___	___	___	___	___	___	___	
3	2	12		13	19		17	20	4	21	19	15	17	
___	___	___		___	___	___	___	___	___	___		___	___	
13	6	17		14	19	8	14	17	4	13		19	9	
___	___	___	___											
5	19	1	17											

	Frequency	Percent
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		

	Frequency	Percent
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		

Decode the secret message. It will require some guessing and checking!

Part 5: **Create your code**! Create your own code using numbers or others symbols to represent each of the alphabet. Turn in the following items:

1. [On one sheet of paper] A complete key for your code (i.e. the letter of the alphabet and the "code" that signifies that letter).
2. [On a separate sheet of paper] A NEATLY written secret message (like the one above) for me to try to solve (without looking at your key). It should be at least 50 characters long. Good luck!

Taken from Mathematics Teaching in the Middle School February 1998 and April 1999