

Report of Frequency Analysis Attack

● Idea

一開始先寫一個計算每個字母次數的程式，去觀察字母出現的頻率。頻率高的前幾個字母通常可以推論為母音，而母音 a, e, i, o, u 中，又以 e 出現的頻率最高。之後再去猜測單字可能出現的頻率，如 the, an, is, are...，去找出各自轉換後相對應的字母。之後再寫另一個程式去解密。

● Process

1. 計算每個字母次數的程式：觀察字母出現的頻率。將字元出現的次數由小到大列印出來後，除了空白，出現次數最多的是 y，猜測 y 可能對應到原本的字元為 e。

```
[('d', 1), ('', 2), ('g', 2), ('l', 2), ('c', 2), ('?', 2), ('3', 2), ('^', 3), ('5', 5), (';', 6), ('/', 6), ('t', 6), ('-', 6), ('6', 6), ('1', 7), ('4', 7), ('k', 7), ('2', 8), ('8', 12), ('"', 18), (':', 20), ('r', 21), ('0', 28), ('n', 45), ('e', 51), ('p', 70), ('.', 118), ('j', 158), ('v', 160), ('s', 167), ('y', 169), ('z', 209), ('q', 223), ('a', 259), ('g', 262), ('o', 265), ('w', 278), ('f', 321), ('x', 486), ('b', 553), ('m', 563), ('l', 586), ('c', 636), ('1', 679), ('h', 783), ('u', 819), ('n', 886), ('y', 1196), ('', 2158)]
```

2. 之後再從單字推測，nby 可能對應到的是 the，之後則是發現其他單字可推測如下：

Encrypt	nbyg	ni	qbi	nbun	nuey
Decrypt	them	to	who	that	take

3. 因此，推測其解密的方法為向右移 6 個字母。

Encrypt	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
Decrypt	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f

4. 解密

一開始先讀檔，之後建立一個如上表所示的 encrypt-decrypt 的字典，之後再加密的文字依照這個字典進行解密，最後則是將解密的文字寫入檔案中。