Monoalphabetic Substitution Cipher

- Encryption and decryption

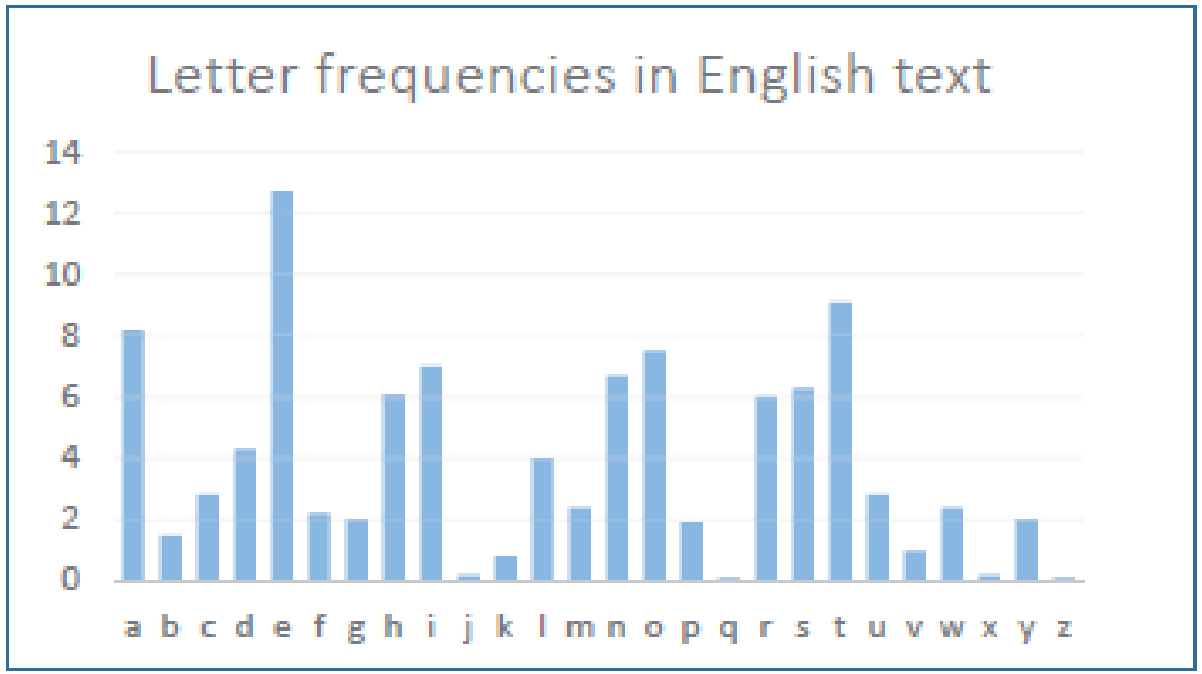
```
# Encryption
$ tr 'a-z' 'vgapnbrtmosicuxejhqyzflkdw' < plaintext  > ciphertext

# Decryption
$ tr 'vgapnbrtmosicuxejhqyzflkdw' 'a-z' < ciphertext > plaintext_new
```

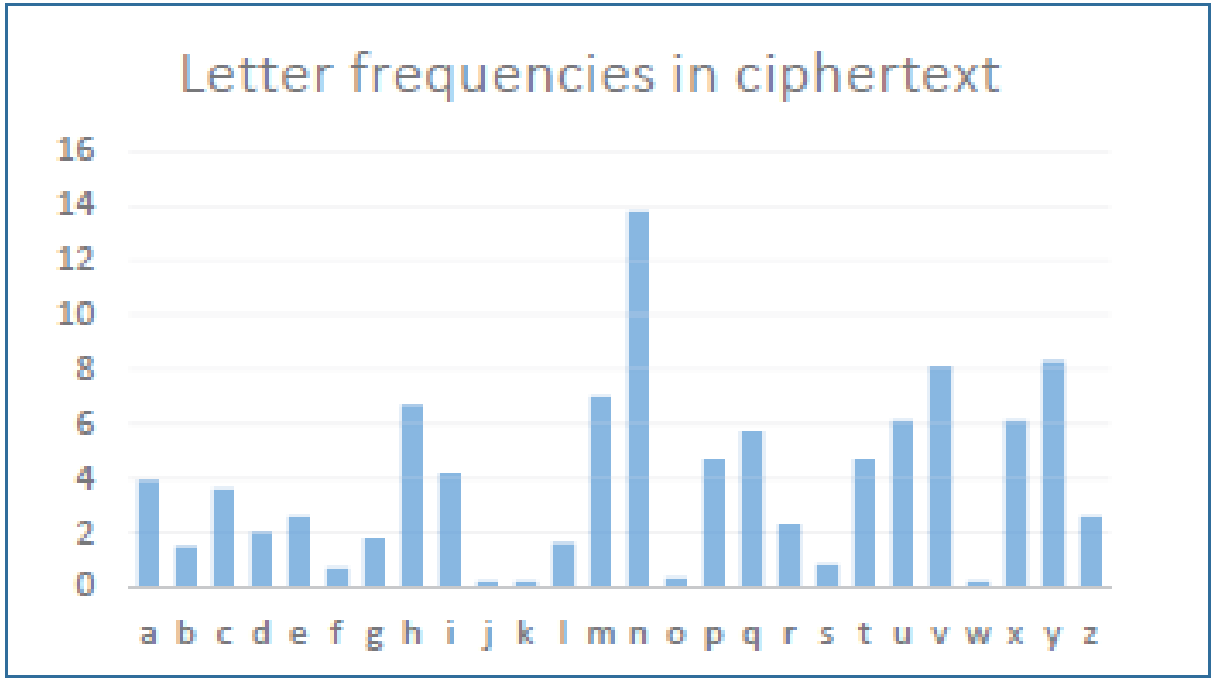
Breaking Monoalphabetic Substitution Cipher

- Frequency analysis is the study of the frequency of letters or groups of letters in a ciphertext.
- Common letters : T, A, E, I, O
- Common 2-letter combinations (bigrams): TH, HE, IN, ER
- Common 3-letter combinations (trigrams): THE, AND, and ING

Breaking Monoalphabetic Substitution Cipher

- **Letter** Frequency Analysis results:





Breaking Monoalphabetic Substitution Cipher

- **Bigram** Frequency Analysis results:

```
Bigram frequency in English
-------------------------------------------
TH :  2.71        EN :  1.13        NG :  0.89
HE :  2.33        AT :  1.12        AL :  0.88
IN :  2.03        ED :  1.08        IT :  0.88
ER :  1.78        ND :  1.07        AS :  0.87
AN :  1.61        TO :  1.07        IS :  0.86
RE :  1.41        OR :  1.06        HA :  0.83
ES :  1.32        EA :  1.00        ET :  0.76
ON :  1.32        TI :  0.99        SE :  0.73
ST :  1.25        AR :  0.98        OU :  0.72
NT :  1.17        TE :  0.98        OF :  0.71
```

```
Bigram frequency in chiphertext (The top-10 patterns)
-----------------------------------------------------
tn  :  77         np  :  50
yt  :  76         hn  :  45
nh  :  61         nu  :  44
nq  :  51         mu  :  42
vu  :  51         cv  :  42
```

Breaking Monoalphabetic Substitution Cipher

- **Trigram** Frequency analysis results:

Trigram frequency in English

THE :	1.81	ERE :	0.31	HES :	0.24
AND :	0.73	TIO :	0.31	VER :	0.24
ING :	0.72	TER :	0.30	HIS :	0.24
ENT :	0.42	EST :	0.28	OFT :	0.22
ION :	0.42	ERS :	0.28	ITH :	0.21
HER :	0.36	ATI :	0.26	FTH :	0.21
FOR :	0.34	HAT :	0.26	STH :	0.21
THA :	0.33	ATE :	0.25	OTH :	0.21
NTH :	0.33	ALL :	0.25	RES :	0.21
INT :	0.32	ETH :	0.24	ONT :	0.20

Trigram frequency in chiphertext (The top-10 patterns)

ytn :	60	tnh :	13
vup :	26	pyt :	13
nhc :	16	hcv :	13
nhn :	15	tne :	13
nuy :	14	mrc :	13

Breaking Monoalphabetic Substitution Cipher

- Applying the partial mappings…

```
$ tr ntyhqu EHTRSN < ciphertext
THE ENmrcv cvaHmNES lERE v SERmES xb EiEaTRxcEaHvNmavi RxTxR ameHER
cvaHmNES pEfEixeEp vNp zSEp mN THE EvRid Tx cmpTH aENTzRd Tx
eRxTEaT axccERamvi pmeixcvTma vNp cmimTvRd axcczNmavTmxN ENmrcv lvS
mNfENTEp gd THE rERcvN ENrmNEER vRTHzR SaHERgmzS vT THE ENp xb
lxRip lvR m EvRid cxpEiS lERE zSEp axccERamviid bRxc THE EvRid S
vNp vpxeTEp gd cmimTvRd vNp rxfERNcENT SERfmaES xb SEfERvi
axzNTRmES cxST NxTvgic
SEfERvi pmbbERENT ENmr
cmimTvRd cxpEiS HvfmNr
vNp mTvimvN cxpEiS lER
```

```
$ tr ntyhquvmxbpz EHTRSNAIOFDU < ciphertext
THE ENIrcA cAaHINES lERE A SERIES OF EiEaTROcEaHANIaAi ROTOR aIeHER
cAaHINES DEfEiOeED AND USED IN THE EARid TO cIDTH aENTURd TO
eROTEaT aOccERaIAi DIeiOcATIa AND cIiITARd aOccUNIaATION ENIrcA lAS
INfENTED gd THE rERcAN ENrINEER ARTHUR SaHERgIUS AT THE END OF
lORiD lAR I EARid cODEiS lERE USED aOccERaIAiid FROc THE EARid S
AND ADOeTED gd cIiITARd AND rO
aOUNTRIES cOST NOTAgid NAwI rE
SEfERAi DIFFERENT ENIrcA cODEi
```

```
$ tr ntyhquvmxbpzfrcei EHTRSNAIOFDUVGMPL < ciphertext
THE ENIGMA MAaHINES lERE A SERIES OF ELEaTROMEaHANIaAL ROTOR aIPHER
MAaHINES DEVELOPED AND USED IN THE EARLd TO MIDTH aENTURd TO
PROTEaT aOMMERaIAL DIPLOMATIa AND MILITARd aOMMUNIaATION ENIGMA lAS
INVENTED gd THE GERMAN ENGINEER ARTHUR SaHERgIUS AT THE END OF
lORLD lAR I EARLd MODELS lERE USED aOMMERaIALLd FROM THE EARLd S
AND ADOPTED gd MILITARd AND GOVERNMENT SERVIaES OF SEVERAL
aOUNTRIES MOST NOTAgLd NAwI GERMANd gEFORE AND DURING lORLD lAR II
SEVERAL DIFFERENT ENIGMA MODELS lERE PRODUaED gUT THE GERMAN
MILITARd MODELS HAVING A PLUGgOARD lERE THE MOST aOMPLEk oAPANESE
AND ITALIAN MODELS lERE ALSO IN USE ...
```