

Quiz 3

- Cryptography Engineering

1. Please write a python program to determine keyword length of the encrypted message using I.C. (hint: $3 < \text{the keyword length} < 8$)
Ex: keyword = CEISFUN, the output should be "7"
2. Then write a second python program to solve the encryption keyword letters for the message.
3. Finally, break these ciphertext and recover to the plaintext.

Note:

The first two programs will need to read the message from stdin and output the result followed by a newline to stdout.

The answer should be saved in a text file.

Encrypted message 1

- ZQQT K PQUWD PGMWD BQTX Y LFQWL SHAJB UCIPV KUQEJ RBAAC LRSIZ ZCRWT LDFMT PGYXF ISOSE ASZ XN PHTAY HHIIR ADDIJ LBFO E VKUWW VFFLV TCXG HFFXF ZVGXF BFQEI ZOSEZ UGFGF UJUGK PCZWZ UQQI VAFV CSDCX YOPYR SQTEI HQFII VTAYI LRGG R AWAR N LAGWK JCZ XZ UIMPC FTA V X LHMRU LAMRT PDMXV VIDWV SJQWW YCYOE VKXIU NSBVV CWAYJ SMMGH BWDIU DSY Y AGQXR ZWP IF SRZSK PCZWR URQQS YOOIW YSELF USEEE KOEAV SSMVE DSY Y APQHR PZKYE SSMVE PBSWF TSFLZ UUILZ JVUXY HGOSJ AIERF ZAMP C SONSL YOZHR ULUIK FHAET XIUVV HBPXY PGP MW MWOYC AMMXK HQTII PHEIC MAAVV JZAWV SMFSR UOSIZ UKTMT ODDSX YSEW Y HGSEZ USPEJ AFARX HGOIE KSZGP VJQVG YSVYU PQQEE KWZAY PQTTV YGARJ HBPXY PBSWR YSPEP IMPEP MWZHZ UUFLV PFDIR SZQ ZV SWZPZ LIAJK OSUVT VBHIE AWARR SJMPL LHTIJ HAQTI PBOMG SSEAY PQTLR CSEAV WHMAR FHDEU PHUSE HZMFL ZSEEE KKTMT O ODID HYURX YOBMU OOHST HAARX AVQVV CSZYV ZCRWZ USOYI PGFWR UREXI PDBME NHTIK OWZXR DRDCM LWXJI VAMXK YOOXZ C SEYG LFEXZ AWARJ HFQAF YYURX HGMGK PJQPP PBXMK LFMXL YSMWZ UGAGZ LHKXY LQDIU BZUXP VTARV DFUXV YCDXY LDMVK POX MK FCREE VHTII MWZHZ HGBSN LFRYC HHAYT OGFSE LOZHR ZKTSC LGAQV HQTEJ AWEID LBFME AVQLV HZFLP ZQQT K PQUWD VTMXV TDQVR ASOPR ZGAJR UHMKF UWEXJ HGFLV KQED ZCRGF UGQVM HHUWD VFFLV PABSJ AIDIJ VTBPL YOXMJ AGURV JIDIJ PBFLV JVGVT OVUWK VFKEE KHDEU PHUSE DVQXY LFAJR UQUIE ACDGF TDMVR AWHIC FFQGV UHFMD LGMVV ZINNV JHQHK VJQVP KWRJV YSZXY HBPPZ UURVF THTEK DVUGY AVQME KIXKV UQQSI JFQHL SWFCF MTAVD LFMKV ZQAYC KOXPF DAQVV ZHMXV TSXJ HFQNV HZAYJ SM IEK JVQHR URFLV TCFMM LGAJK OSIVZ ASDJF YAMWZ TDAVK HBFE E PBSVV KWQRK PBFLV HBMPP ZWESW OWELZ ZHAVP HGFLV MOO XJ OSDIT VFPWG YCNES PZUXP PGMTF DSDJL SOZHK YCGFC LGAQV ASEXR URUXZ ZPKXY PGFVF BPXIJ VAQWK HBPEI KHTEK HZMVX LD AVK PCZSW OWEXF YWOEC LIUHV UQQMJ ZWRXV KQARJ PGFIE JMUWE VZQWJ WSDXZ UOOMF BGMRU LLMGK PBSME PHEHV TOZJH PBNVZ LTF SN YWFIR OWEXF YMIID BGFOE VKYSI LHTEE TSDIW HQFWY BAMRE HHGVV CWQAV YOZME KIOXZ VBAJV EHQRU LRQ BG LFUIE JSUWK OSNIJ AVQPG ACFLV JFUXZ JWEQF MVGQR UVUWK VFKLZ ZHAVZ JOXGY HFMGK LFEGR UCZPP ISQWK PAMXV KPKXY L GFEE KODHN OWOLY BAMRV EDQVZ LBOIN OSFLV YOOXL HZAVK YOPMK PCZEI FVMWW BFZMJ OSPXF MCDQT VFDIT AJUIN ZCRME K WHMU BOXWN LAGWK YSSEI KHTID HGRSI TWZKG HFFWF MOSVV HHILF SSIID BGFQV HGGVV AVQQS FHTIZ YFQPR AWARK VHTID HGE SW ISURX ZPKAY VAFV FODIJ BFDSL URQHR URURT VBFID WZMXZ UUFLV PBOMU LBFWZ UHTIZ YZUZV ZCDGF URUXZ VBILZ JVFVR KW FMF UVMWY HBPIU KCIRK VIEAV TIE XI HHTII JCZWZ KSDXY LUQRV YOXFV HFURX VTFLV DVAPV UODVR AWHIK OOXZY LFQWG LQFMM LDDSS HPUPZ AMAJZ AGPIK HWXW

Encrypted message 2

- UTCEM KTGHF KJYHF WTFIA GICHN NKMUD PFUAX FXCPL MEMLE GUETB UFDHV GGRXV KJKIH DVADG VVLIP KKFLA CKUTT VGPTL GERZG QNGHY QIRWX OFQII CIRIH UYSIH WICPK UREPB PJRRH PMGRM KFLHB PTCUK QDRWX XVPNZ TRBJT NTFPK CTRTK QWMJK GUSRT VZMCP GDSHM EFLIB PLYAE AWMGZ GKVCW GDYCV KGYIX QLPXH NMCHY TFKZG QNJTW IVNGX XZMJL NPYRJ WZPTW YVKJL VJCIT UZBTH NULDM KFLHT PUCBU TRATY TVQWH PVQPG FRQLX NVEYG YVKJL VSCST KCWJG NVEYG KEEHH OVRWB PXUWB EYGIA CJADL VLQCH UDYAE NRZDN TRLST POGTM AKMPV SLGGX CEBIA KJBXY HZAJE VPYIM CTFTL KKQTE HDMGX ECMHX NPRDT PRETB PNFXV JGPDZ TVQHA CJEPB PVBPL VIMCZ CJATG FVLRR QMCGI TVHJW KTCPG FZLLA KTFEX TJMCL CEBIA KEEHT TVBPR DPBPR HZLSB PXRWX KIPTT NCCKX NZLAB GLMUM JVGCV QETTG VZMCT NMYAN GKFTL CDEK KEAXI NVQLA KTFWT XVQLX RYKLT AKPPW KKGDG CCYQN UVQPG FNFXV JRPTF CBGCZ TRNXW JRTDV CDMCZ VYCGX XLVJX UFDHB PVAJK KJRHT PUQIK KGNXG IKFTM JZLIT YUPNO GZJUK QDYIM TRAIB XVQJI GIQIB VZMCL CILH TBGCZ CJYRM KMCAR KEJXM GIYIN TVYHB PJMRB GKWIA GTPTW WCGIR QWMCX YIGIX TPIA GGYGM KRJXM AFDPG QKFTK HZLSL CJNDP GIDJE CKMJV JIRDG GRLST UNFDE GJMBX CTPL VZQTF GERXG VYCWX CCRWR UTCEM KTGHF QWYIX OGCCT VVAAT UJMUT PKYVH PZQIL CJRWX FICPF UFDHR PJCGO CKGHF QIRWX KDNDL VLPTL QWNAN TRJXL VJGCX ELPTL KERWX EYSGV JYGHM QIWPB FKPPW KKGDG YYCIA GIMUT PTGTG VFPRH OGYGT VZTTE AICRX PKRXF GJYX ULZYX EKCSM QMCGR FZDUX TVLIA CEBAB PXDGH OKFPM YYGRA VYCXG FLVX PTCCK EICSN NZRNH HWMGF GIYVX UTMJE FRJAH YDCGX UKYIX OVLIL CICYX CCMJL NPUPM EYCSM PURWX OFRXO GJMUM JVUGB VVPUH TDYHB OGMGM CERPG KEEGX FZCCM KERWX CEYAR UZQDY JZQWB UKMGR CJRWX HRAL JPTV QIBHI TFZPU KCGIR KJYEH YVPUN NLSM TFSQE GJMBX VVQIT PUGIB USWIA KJRGH WSJTL QDCHM CEBPK FKFPK CCYXZ GGMGM KFLDY JZQIH TZAPE GMSX PTCXL UZDIX FTMCL KJRTG EPGH QCCHL RVPB PRAXH WJYCW GOYRM KEEGX KKQX ORLSL KEZGB GWRDP TZRTT JZQIH TPUTF WJZG QNKDK GKFPG OVPTY CTRHA WDCG CKSGX XZCLX FLSX TRLXG FLAIB QEMUX ZKCCW GUCMI GIGTG EVGHM JVZTL VYCAI VFRWX EIGIB EZQBH HYSBT PYGHM QIWWB UKMGB ERJRA CIYRM GIQRT PFLAR DVCHM KDYIX FSWIA GJRPB FRPSP JZAWA WDCX ZGCB GEATP JVRWX TRAIN CCMGM TRBXM KFLPK AYYHY WILXL JVBIH HFPBV QIPTV VMGTP UFDXG FZTXW WRJHP GDSHM TVEPK FKFTF CJDDK OZLVI CIRHH HREGX CKUWH NVUTF WJRBX CJSXG VYCBU AKFTB TICAT VZMCM QKFTF CJQDY DVGZ USWLA QDRWX ARPTL WIPDN PUCST PUGCV QERTF RYIB PXRWX KEAXW GERHB PKFTB TCGKX UFRPH PUGIB QEUWB EYRGT FZRXH PYYHA CEBTW FFUCM QLQLX OLQIK CKFTK EFLHB FVPIA GXCCX TRJQX CIGZ QWRWX YYMAX PRPGT VZTMM JRLIA GICHI GTRXO GGPDU CSGAB VPMUB VJBTM CZJH

Encrypted message 3 HINT key length could be 5 or 6

- IVIKDKDQMJGLPWLZGMPFBJIIDBBYSLJDXFGBIWWEHAPHEYSGNCCYOOTSTZABCOBVRTAZEYVWWWWAZAIDGAZPETHPVBPWOBVJXGFMD
OBCGPFKXKSZZAIGCJRPETACJHUTHPVHKJHPZHFPMEVZEQSBYOMHSDVFTASFGZTCOBZCGHFMD OBCWVNVBRVKGXDBMKFBTGBVGMP
TBVFM TGBLBMXZWESHGCBYSKDTBYSFOWARQH CJEQBCUIDCNCHWWGNEDWHPKQCZGDKIGDENHPZGIGWVTWIASBFHATQIJSBCD
WZBMPGQKKTHTQIGMEFMJSGISLKCFTHPVFXLSZVHAGSMGCLHWJCSXMDTRBTIWWEGHUHPVGXRZCJWHCCZZBVPFKVFTIWWECYIVQJU
XCHTVATCWVRBHJHPFILTCNYWLUOBYSKHAIEGBDBBYSKTIJHATSFGZTCOBZCGIVIKVXLOAZBAXRQEUYDFITFBBSWIHAPHPVKTHAIUOGS
HPRHMWSGNWLWLSLKCTKCQUOGPGGCIFDFBYOMWSPRRDAMUWLTOAVKAXQPTONHSLYWLHISOISZPHQFBBRCCCRMWWVBCYCCWKV
XGOLVENPHMJCEJHQFBLIVMJSMWSVYOWICJVGBUHMUOGSPICOGRLRUTXBAKSTRVWKVXG

Polyalphabetic cipher

- A polyalphabetic cipher is any cipher based on substitution, using **multiple substitution alphabets**.
- The Vigenere cipher is probably the best-known example of a polyalphabetic cipher, though it is a simplified special case.
- The **Enigma machine** and some **electric encryption device** are more complex but is still fundamentally a **polyalphabetic substitution cipher**.

Vigenere cipher

$k =$ C R Y P T O C R Y P T O C R Y P T (+ mod 26)

$m =$ W H A T A N I C E D A Y T O D A Y

C = Z Z Z J U C | L U D T U N | W G C Q S

suppose most common = "H" \rightarrow first letter of key = "H" - "E" = "C"

	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
A	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
B	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	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	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
G	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
H	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	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	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

	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
A	A	B	C	D	E	F	G	H	I									R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	I								S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	I								T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	I								U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	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
G	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
H	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	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	I										T	U	V	W												J
L	I										U	V	W	X												K
M	I										V	W	X	Y												L
N	I										W	X	Y	Z												M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	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

Vigener cipher

k = C R Y P T O C R Y P T O C R Y P T (+ mod 26)
m = W H A T A N I C E D A Y T O D A Y

c = Z Z Z J U C | L U D T U N | W G C Q S

↑ ↑ ↑

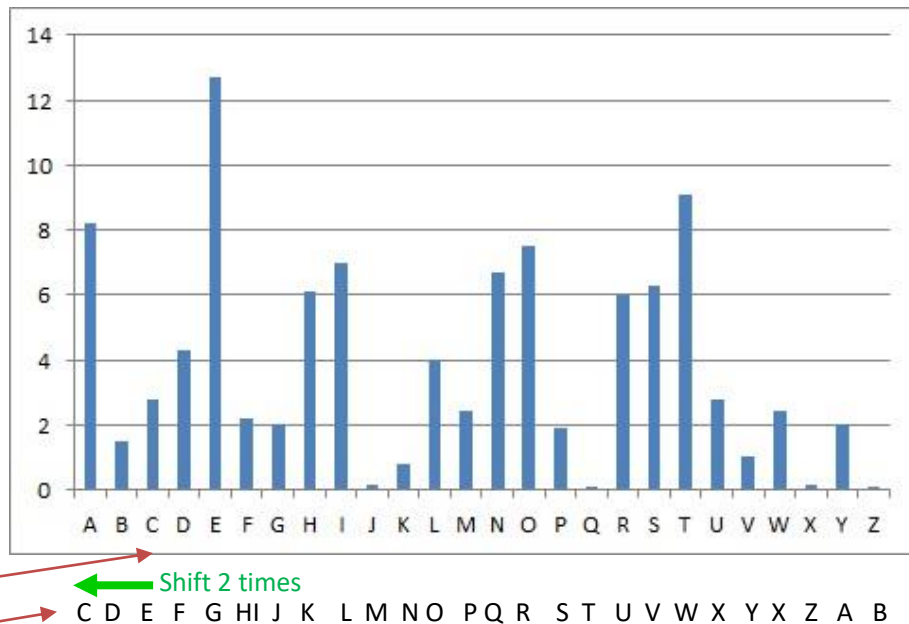
suppose most common = "H" → first letter of key = "H" - "E" = "C"

Plaintext m
frequency

$$M + K = C$$

key
 $C=3$
Shift 2 times

Ciphertext C
frequency



Substitution Cipher

Encryption: Substitution cipher

Similar to Caesar cipher, but you can choose which letters substitute the alphabet

ABCDEFGHIJKLMNOPQRSTUVWXYZ
MRBGSLOAEFYWDKUQHPCJTZXIN

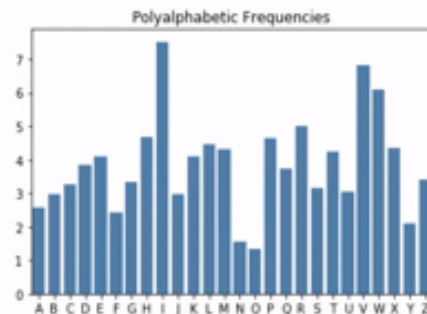
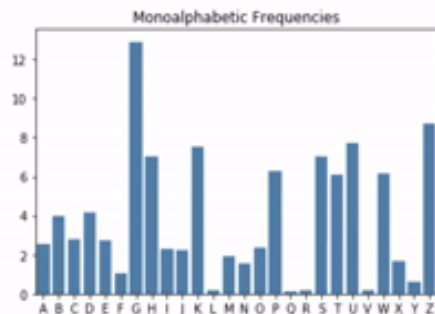
HI WORLD



AE VUPWG

Identify a Polyalphabetic Cipher

Identifying a Polyalphabetic Cipher



North Carolina
School of Science
and Mathematics

Calculating Probabilities

For the monoalphabetic ciphertext, the probability of picking two letters at random from the text and having them both be A's (using the distribution above) would be:

$$M_{\text{both A's}} = \frac{73}{1000} \cdot \frac{72}{999} \approx 0.00526$$

The probability of picking two letters at random from the text and having them both be B's

$$M_{\text{both B's}} = \frac{9}{1000} \cdot \frac{8}{999} \approx 0.0000721$$

and in general, where n_i denotes the number of character i and n_{text} denotes the total number of characters in the text:

$$M_{\text{both i's}} = \frac{n_i}{n_{\text{text}}} \cdot \frac{n_i - 1}{n_{\text{text}} - 1}$$

For long messages, $\frac{n_i}{n_{\text{text}}} \approx \frac{n_i - 1}{n_{\text{text}} - 1}$, so the formula can be simplified to:

$$M_i^2 \approx \left(\frac{n_i}{n_{\text{text}}} \right)^2$$

The sum of all these probabilities, which is equivalent to asking "What's the probability of picking two letters at random from the text, and having them be the same letter?" works out to be:

$$\sum_{i=A}^Z M_i^2 \approx 0.066$$

Doing the same calculation on our very evenly distributed polyalphabetic ciphertext yields:

$$\sum_{i=A}^Z P_i^2 \approx 0.038$$

This score, calculated by summing the squares of the letter frequencies is called the **Index of Coincidence**. When presented with an unknown ciphertext, its index of coincidence will suggest if it was enciphered with a polyalphabetic cipher (if the score is close to 0.038) or a monoalphabetic cipher (if the score is close to 0.066). This gives a quick and easy way to determine with a single number the most likely type of cipher used to create a ciphertext.

Identifying a Polyalphabetic Cipher

Monoalphabetic

A	73	J	2	S	63
B	9	K	3	T	93
C	30	L	35	U	130
D	44	M	25	V	13
E	27	N	5	W	16
F	28	O	74	X	78
G	16	P	27	Y	19
H	35	Q	3	Z	1
I	74	R	77	

Polyalphabetic

A	38	J	38	S	39
B	39	K	39	T	38
C	38	L	38	U	39
D	38	M	39	V	39
E	39	N	39	W	38
F	38	O	38	X	38
G	38	P	39	Y	39
H	39	Q	38	Z	38
I	38	R	39	



North Carolina
School of Science
and Mathematics

Identifying a Polyalphabetic Cipher

Monoalphabetic

A	73	J	2	S	63
B	9	K	3	T	93
C	30	L	35	U	130
D	44	M	25	V	13
E	27	N	5	W	16
F	28	O	74	X	78
G	16	P	27	Y	19
H	35	Q	3	Z	1
I	74	R	77	

`character_frequency(message)`

[0.073, 0.009, 0.030, ..., 0.019, 0.001]
 M_a M_b M_c M_f M_g

$$M_{\text{both i's}} = \frac{n_i}{n_{\text{text}}} \cdot \frac{n_i - 1}{n_{\text{text}} - 1}$$

So for long messages...

$$M_i^2 \approx \left(\frac{n_i}{n_{\text{text}}} \right)^2$$

Identifying a Polyalphabetic Cipher

Monoalphabetic

A	73	J	2	S	63
B	9	K	3	T	93
C	30	L	35	U	130
D	44	M	25	V	13
E	27	N	5	W	16
F	28	O	74	X	78
G	16	P	27	Y	19
H	35	Q	3	Z	1
I	74	R	77	

This number is around 0.066
for monoalphabetic ciphertexts



North Carolina
School of Science
and Mathematics

character_frequency(message)

[0.073, 0.009, 0.030, ..., 0.019, 0.001]
 M_a M_b M_c M_z M_x

Index of Coincidence

$$\sum_{i=A}^Z M_i^2 \approx 0.066$$

Probability of picking 2 A OR 2 B OR 2 C ...

To identify a Polyalphabetic Cipher

Identifying a Polyalphabetic Cipher

Ciphertext:

RHVST TEYSJ KMHUM BBCLC GLKBM HBSJH HDAYC PPWHD UUTAP STJAI YMXKA OKARN NATNG CVRCH BNGJU
EMXWH UERZE RLIMX MASRT LAHRJ KIILJ BQCTI BVFEW TKBQE OPKEQ OEBMU NUTAK ZOSLD MKXVO YELLX
SGHTT PNROY MORRW BWZKX FFIQJ HVDZ2 JGJZY IGYAT KWVIB VDBRM BNVFC MAXAM CALZE AYAZK HAOAA
ETSGZ AAJFX HUEKZ IAKFM FWXTO EBUGN THMYH FCEKY VRGZA QWAXB RSMSI INHQM HXRNR XMOEU ALYHN
ACLFH AYDPP JBAHV MXPNF LNWQB WUGOU LGFMO BJGJB FEYVR GLAQW ANZCL XZSVF BISM B KUOTZ TUWUO
WHFIC EBAHR JPCWG CVVEO LSSGN EFGCC SWHYK BJHMF ONHUE BYDRS NVFMR JRCHB NGJUB TYROU TYVRG
ZAXWX CSADX YIAKL INGXF FEEST UNIAJ EESFT HAHRT WEGTM CRS

Index of Coincidence ≈ 0.041709 --> Most Likely Polyalphabetic



North Carolina
School of Science
and Mathematics

keyword length	Index of Coincidence
1	0.066
2	0.052
5	0.044
10	0.041
large	0.038

Example

JAKXQ SWECW MMJBK TQMCM LWCXJ BNEWS XKRBO IAObI NOMLJ GUIMH
YTACF ICVOE BGOVC WYRCV KXJZV SMRXY VPOVB UBIJH OVCVK RXBOE
ASZVR AOXQS WECVO QJHSG ROXWJ MCXQF OIRGZ VRAOJ RJOMB DBMVS
CIESX MBDBM VSKRM GYFHA KXQSW ECWME UWXHD QDMXB KPUCN HWIWF
NFCKA SKXNF DLJBY RNOBI YFSQN HRIYV IWRQS WCGKC BHRVN SSWYF
SQNTS ZNWCT AWWIB SFIWW CTAWW IWWXI RGKRN LZIAW WIWHK PNFBS
ASVIE SXMBD BMVSK RMGYC NGKPU CNHWI WFNFC KASKX NFDLJ BYRNO
BIYFS QNHRI NBQMW SOVBO IWCVB INWCT AWWIO WFIRG ZVRAO WNJOR
RGZVR AORRB OMBDB MVSOP NJORR GZVRA OXQWB XNSXM BDBMV SPMOH
OIWWC TAWWI

Example

ABABA BABAB ABABA BABAB ABABA BABAB ABABA BABAB ABABA BABAB
JAXXQ SWECW MMJBK TQMCM LWCXJ BNEWS XKRBO IAObI NOMLJ GUIMH

Group 1 (A): JKQWCMJKQCLCJNWXROABNMJUM... I.C. = 0.06060

Group 2 (B): AXSEWMBTMMWXBESKBIOIOLGIH... I.C. = 0.05624

Avg. = 0.05842



North Carolina
School of Science
and Mathematics

Example

ABCAB CABCA BCABC ABCAB CABCA BCABC ABCAB CABCA BCABC ABCAB
JAKXQ SWECW MMJBK TQMCM LWCXJ BNEWS XKRBO IAQBI NQMLJ GUIMH

Group 1 (A):	JXWWJTCWJEXBAIMG...	I.C. = 0.04405
Group 2 (B):	AQEMBQMCBWKQONLUH...	I.C. = 0.05108
Group 3 (C):	KSCMKMLXNSRIBOJI ...	I.C. = 0.04782
		Avg. = 0.04765



North Carolina
School of Science
and Mathematics

Example

Key Length	Average I.C.
2	0.05842
3	0.04765
4	0.08340
5	0.04539
6	0.04539
7	0.04814
8	0.08125

A little higher than expected, but most likely candidate for the correct key length

This is also a likely candidate, but so are lengths of 12, 16, 20, etc. Why does that make sense if the key has length 4?



North Carolina
School of Science
and Mathematics

The Algorithm

1. Assume key length, n , starting with a value of 2
2. Split ciphertext into n groups so that characters in the same group would have been enciphered using the same character of the keyword
3. Calculate the index of coincidence of each group
4. Calculate the average index of coincidence of all groups
5. If the average index of coincidence is "close" to the English value of ≈ 0.068 then assume n is the correct length
6. If not, increase n by 1 and start the process over



North Carolina
School of Science
and Mathematics

How can we crack a Vigenere cipher!!

Ciphertext

JAKXQ SWECW MMJBK TQMCM LWCXJ BNEWS XKRBO IAObI NOMLJ GUIMH
YTACF ICVOE BGOVC WYRCV KXJZV SMRXY VPOVB UBIJH OVCVK RXBOE
ASZVR AOXQS WECVO QJHSG ROXWJ MCXQF OIRGZ VRAOJ RJOMB DBMVS
CIESX MBDBM VSKRM GYFHA KXQSW ECWME UWXHD QDMXB KPUCN HWIWF
NFCKA SKXNF DLJBY RNOBI YFSQN HRIYV IWRQS WCGKC BHRVN SSWYF
SQNTS ZNWCT AWWIB SFIWW CTAWW IWWXI RGKRN LZIAW WIWHK PNFBS
ASVIE SXMBD BMVSK RMGYC NGKPU CNHWI WFNFC KASKX NFDLJ BYRNO
BIYFS QNHRI NBQMW SOVBO IWCVB INWCT AWWIO WFIRG ZVRAO WNJOR
RGZVR AORRB OMBDB MVSOP NJORR GZVRA OXQWB XNSXM BDBMV SPMOH
OIWWC TAWWI



North Carolina
School of Science
and Mathematics

Find out key length is 4 by I.C.

Ciphertext

1234

JAKXQ SWECW MMJBK TQMCM LWCXJ BNEWS XKRBO IAObI NOMLJ GUIMH
YTACF ICVOE BGOVC WYRCV KXJZV SMRXY VPOVB UBIJH OVCVK RXBOE
ASZVR AOXQS WECVO QJHSG ROXWJ MCXQF OIRGZ VRAOJ RJOMB DBMVS
CIESX MBDBM VSKRM GYFHA KXQSW ECWME UWXHD QDMXB KPUCN HWIWF
NFCKA SKXNF DLJBY RNOBI YFSQN HRIYV IWRQS WCGKC BHRVN SSWYF
SQNTS ZNWCT AWWIB SFIWW CTAWW IWWXI RGKRN LZIAW WIWHK PNFBS
ASVIE SXMBD BMVSK RMGYC NGKPU CNHWI WFNFC KASKX NFDLJ BYRNO
BIYFS QNHRI NBQMW SOVBO IWCVB INWCT AWWIO WFIRG ZVRAO WNJOR
RGZVR AORRB OMBDB MVSOP NJORR GZVRA OXQWB XNSXM BDBMV SPMOH
OIWWC TAWWI

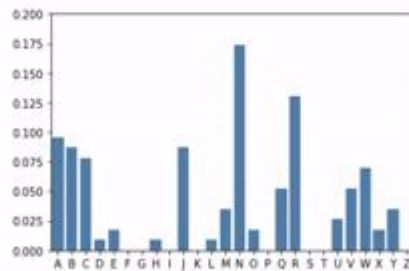


North Carolina
School of Science
and Mathematics

and Mathematics

Group 1

JQCJQ LJWRA NJMAC BCCJM VBJCX ARQCJ RJQRR RBVEB VMHQC UDXUW
NANUN YNYRC BNYNN ABWAW RNAWN AEBVM NUWNA NJNYN NWBCN AORRN
RRREV NRRQN BVOWA

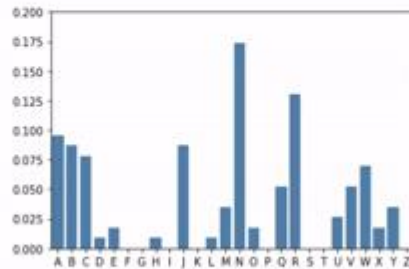


North Carolina
School of Science
and Mathematics

A

Group 1

JQCJQ LJWRA NJMAC BCCJM VBJCX ARQCJ RJQRR RBVEB VMHQC UDXUW
NANUN YNYRC BNYNN ABWAW RNAWN AEBVM NUWNA NJNYN NWBCN AORRN
RRREV NRRQN BVOWA

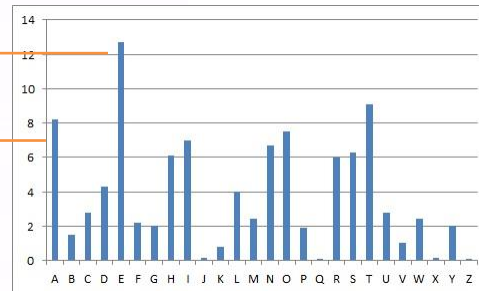
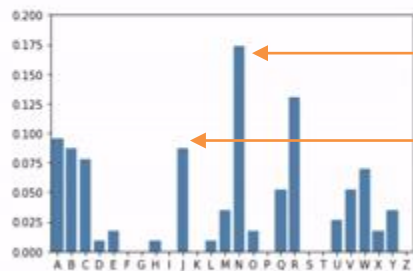


North Carolina
School of Science
and Mathematics

A

Group 1

JQCJQ LJWRA NJMAC BCCJM VBJCX ARQCJ RJQRR RBVEB VMHQC UDXUW
 NANJN YNYRC BNYNN ABWAW RNAWN AEBVM NUWNA NJNYN NWBCN AORRN
 RRRBV NRRQN BVOWA

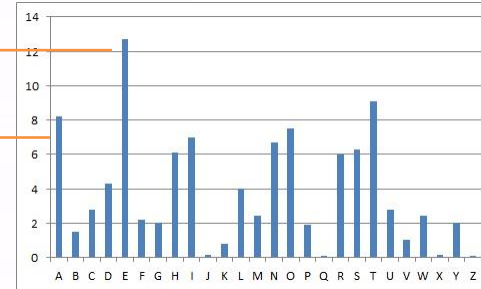
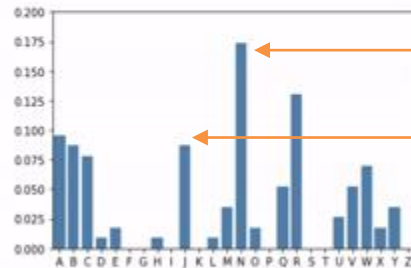


Shit 9 like a Caesar cipher

How can we find out the right shift

Group 1

JQCJQ LJWRA NJMAC BCCJM VBJCX ARQCJ RJQRR RBVEB VMHQC UDXUW
NANJN YNYRC BNYNN ABWAW RNAWN AEBVM NUWNA NJNYN NWBCN AORRN
RRREB NRRQN BVOWA



Shift 9 like a Caser cipher

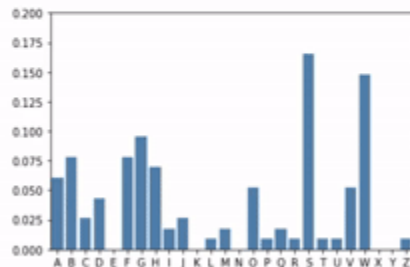
$$\text{I.C.} = \frac{\sum_{i=A}^{i=Z} (f_i + f_{i+k'}) (f_i + f_{i+k'} - 1)}{(N+N')(N+N'-1)} \text{ where } 0 \leq k < 26$$

$$\begin{aligned} & \sum_{i=A}^{i=Z} (f_i^2 + f_{i+k'}^2 + 2f_i f_{i+k'} - f_i - f_{i+k'}) \\ &= \sum_{i=A}^{i=Z} f_i^2 + \sum_{i=A}^{i=Z} f_{i+k'}^2 + \mathbf{2 \sum_{i=A}^{i=Z} f_i f_{i+k'}} - \sum_{i=A}^{i=Z} f_i - \sum_{i=A}^{i=Z} f_{i+k'} \end{aligned}$$

$$\sum_{i=A}^{i=Z} f_i f_{i+k'} = 3(2) + 4(2) + 5(1) + \cdots + 0(2) + 5(6)$$

Group 2

ASWBM WBSBO OGHCV GWVZR PUHVB SASVH OMFGA JDSSD SGASW WQBCI
 FSFBO FHVQG HSFTW WSWWW GLWHF SSDSG GCIFS FBOFH BSOVW WWGAJ
 GABDS JGAWS DSHWW



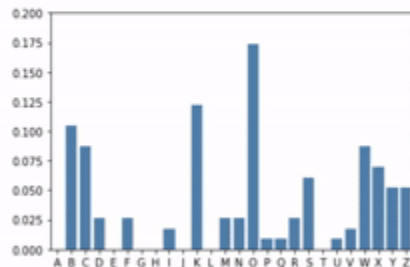
North Carolina
 School of Science
 and Mathematics

0 (A): 325.7668
 1 (B): 883.3711
 2 (C): 4317.3209
 3 (D): 611.5731
 4 (E): 893.7738
 5 (F): 751.9344
 6 (G): 3896.8727
 7 (H): 1589.9641
 8 (I): 1479.7620
 9 (J): 2843.9011
 10 (K): 1107.8685
 11 (L): 991.7172
 12 (M): 751.9465
 13 (N): 2039.9947
 14 (O): 50.3009
 15 (P): 1265.0717
 16 (Q): 1414.7607
 17 (R): 1465.7353
 18 (S): 553.8066
 19 (T): 4314.7729
 20 (U): 459.8573
 21 (V): 2343.4717
 22 (W): 1096.8136
 23 (X): 4275.2646
 24 (Y): 1108.1107
 25 (Z): 1788.0592

A

Group 3

KWMC CNXOB MUYFO OYKVX OBOKO ZOWOS XCOZO OBCXB KYKWM XDKNW
CKDYB SRISK RSSSC WFCWX KZWKB VXBKY KNWCK DYBSR QOIBC WFZOO
ZOBO OZOBX BPOCW



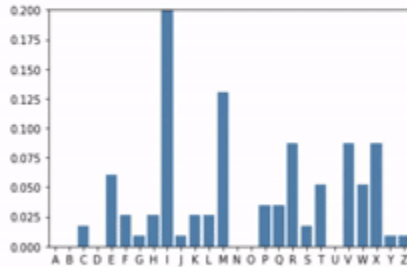
North Carolina
School of Science
and Mathematics

0 (A): 1046.2090
1 (B): 1561.0377
2 (C): 2470.9050
3 (D): 1353.3253
4 (E): 1357.3504
5 (F): 2957.6687
6 (G): 1232.2920
7 (H): 797.8171
8 (I): 705.6488
9 (J): 886.3809
10 (K): 58.9651
11 (L): 3622.6714
12 (M): 1265.4848
13 (N): 2159.6061
14 (O): 552.4254
15 (P): 5049.3125
16 (Q): 549.0219
17 (R): 2475.0480
18 (S): 1151.4112
19 (T): 1490.7523
20 (U): 2040.6587
21 (V): 491.1330
22 (W): 340.3210
23 (X): 1335.5712
24 (Y): 4497.8480
25 (Z): 1250.2771

A

Group 4

XEMTM XEKII LITIE VRXSY VIVRE VXEQG WXIVJ MMIMM RFXEE HMPHF
KXLRI QIWWC VQZT IITII RIIPS IMMRC PHFKX LRIQI MVWIT IIVWR
VRMMP RVXXM MMITI



North Carolina
School of Science
and Mathematics

A

0 (A):	921.3981
1 (B):	1117.8622
2 (C):	549.3962
3 (D):	1852.3085
4 (E):	38.6190
5 (F):	1604.6688
6 (G):	844.3241
7 (H):	1705.8365
8 (I):	789.9132
9 (J):	6101.7858
10 (K):	579.6742
11 (L):	3365.3096
12 (M):	917.9620
13 (N):	3284.8424
14 (O):	1296.2255
15 (P):	1477.7176
16 (Q):	391.4905
17 (R):	536.4596
18 (S):	6139.4866
19 (T):	439.6708
20 (U):	1199.8803
21 (V):	525.0241
22 (W):	3700.5745
23 (X):	638.0083
24 (Y):	2118.7299
25 (Z):	3423.9338

Ciphertext: Keyword = JOKE

JAKXQ SWECW MMJBK TQMCM LWCXJ BNEWS XKRBO IAObI NOMLJ GUIMH
YTACF ICVOE BGOVC WYRCV KXJZV SMRXY VPOVB UBIJH OVCVK RXBOE
ASZVR AOXQS WECVO QJHSG ROXWJ MCXQF OIRGZ VRAOJ RJOMB DBMVS
CIESX MBDBM VSKRM GYFHA KXQSW ECWME UWXHD QDMXB KPUCN HWIWF
NFCKA SKXNF DLJBY RNOBI YFSQN HRIYV IWRQS WCGKC BHRVN SSWYF
SQNTS ZNWCT AWWIB SFIWW CTAWW IWWXI RGKRN LZIAW WIWHK PNFBS
ASVIE SXMBD BMVSK RMGYC NGKPU CNHWI WFNFC KASKX NFDLJ BYRNO
BIYFS QNHRI NBQMW SOVBO IWCVB INWCT AWWIO WFIRG ZVRAO WNJOR
RGZVR AORRB OMBDB MVSOP NJORR GZVRA OXQWB XNSXM BDBMV SPMOH
OIWWC TAWWI



North Carolina
School of Science
and Mathematics

Plaintext

amath emati ciana physi cista ndane ngine erare eacha skedt
oprov ethea ssert ionth atall oddnu mbers great ertha nonea
repri methe mathe matic iansa ysthr eeisp rimef iveis prime
seven ispri meand sobym athem atica lindu ction allod dnumb
ersgr eater thano neare prime theph ysici stsay sthre eispr
imefi veisp rimes eveni sprim enine isane xperi menta lerro
relev enisp rimea ndsoy esall oddnu mbers great ertha nonea
repri methe engin eersa ysthr eeisp rimef iveis prime seven
ispri menin eispr imeel eveni sprim ethir teeni sprim eift
eenis prime



North Carolina
School of Science
and Mathematics