

Experiment 01

Breaking the Mono-alphabetic Substitution Cipher using Frequency analysis method.

--

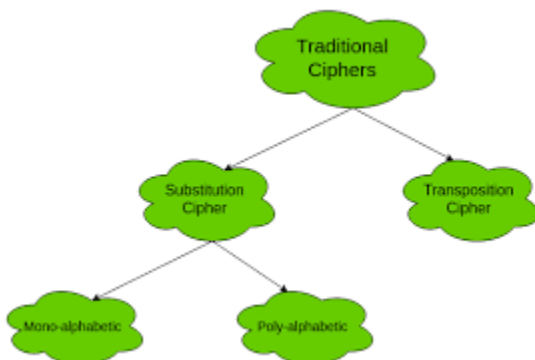
Roll No.	17
Name	Manav Jawrani
Class	D15-A
Subject	Internet Security Lab
LO Mapped	LO1: To apply the knowledge of symmetric cryptography to implement classical ciphers.

Aim: Write a program to understand Breaking the Mono-alphabetic Substitution Cipher using Frequency analysis method.

Introduction:

What is a Substitution Cipher?

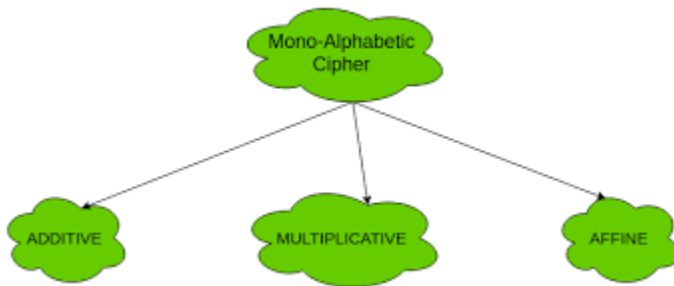
- In cryptography, a **substitution cipher** is a method of encrypting in which units of plaintext are replaced with the ciphertext, in a defined manner, with the help of a key; the "units" may be single letters (the most common), pairs of letters, triplets of letters, mixtures of the above, and so forth. The receiver deciphers the text by performing the inverse substitution process to extract the original message. Substitution ciphers can be compared with transposition ciphers. In a transposition cipher, the units of the plaintext are rearranged in a different and usually quite complex order, but the units themselves are left unchanged. By contrast, in a substitution cipher, the units of the plaintext are retained in the same sequence in the ciphertext, but the units themselves are altered.
- Types of substitution cipher are:
 1. Monoalphabetic Substitution Cipher
 2. Polyalphabetic Substitution Cipher



What is a Mono-alphabetic Substitution Cipher?

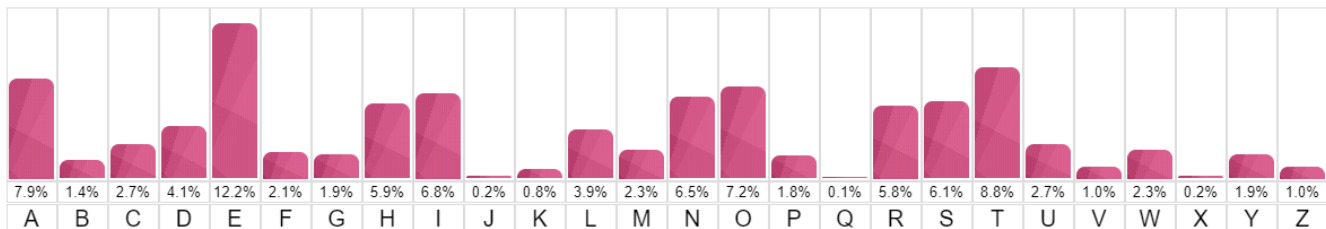
- Monoalphabetic cipher is a substitution cipher in which for a given key, the cipher alphabet for each plain alphabet is fixed throughout the encryption process. For example, if 'A' is encrypted as 'D', for any number of occurrences in that plaintext, 'A' will always get encrypted to 'D'.
- There are various techniques of Monoalphabetic cipher which are as follows

1. Additive Cipher – Additive cipher is one method of changing a permutation of the letters of the alphabet. Each letter in the alphabet is cyclically changed by the equivalent amount and the relative order of the letters is kept similar
2. Multiplicative Cipher – Multiplicative cipher is another method for creating a permutation of the letters of the alphabet. It can take a key value and each letter's position number is multiplied by 5 and thus the product is decreased by modulo 26.
3. Affine Cipher – The Affine cipher is a type of monoalphabetic substitution cipher, wherein each letter in an alphabet is mapped to its mathematical equivalent, encrypted using a simple mathematical function, and transformed back to a letter.



What is frequency analysis in cryptography?

- In cryptography, frequency analysis is the study of the frequency of letters or groups of letters in a ciphertext. The method is used as an aid to breaking substitution ciphers for e.g. mono-alphabetic substitution cipher. Frequency analysis consists of counting the occurrence of each letter in a text. Frequency analysis is based on the fact that, in any given piece of text, certain letters and combinations of letters occur with varying frequencies. For instance, given a section of English language, letters E, T, A and O are the most common, while letters Z, Q and X are not as frequently used.



Results:

1. Problem 1

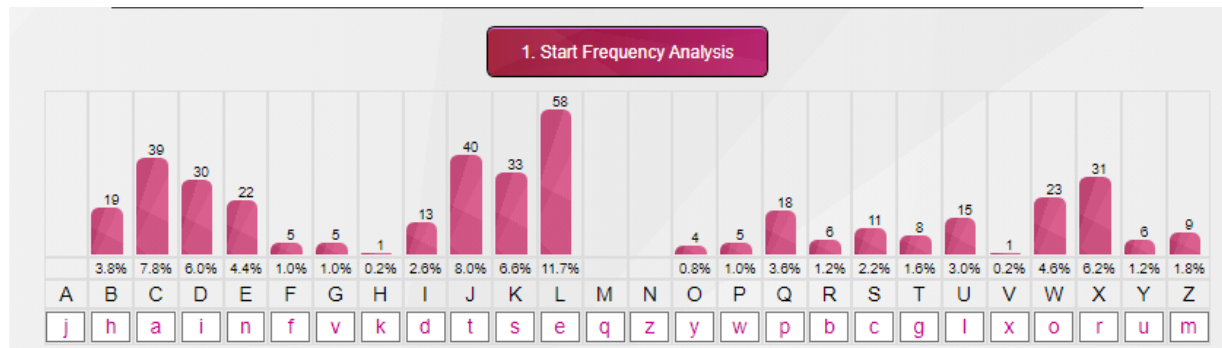
Cipher Text:

Frequency Analysis

Text:

DJ DK C QLXDWI WF SDGDU PCX. XLRLU KQCSLKBDQK, KJXDHDET FXWZ C BDIILE RCKL, BCGI
PWE JBLDX FDXKJ GDSJWKO CTCDEKJ JBL LGDU TCUCSJDS LZQDXL. IYXDET JBL RCJJUL,
XLRLU KQDLK ZCECTLI JW KJLCU KLSXLJ QUCEK JW JBL LZQDXL'K YUJDZCJL PLCQWE, JBL
ILCJB KJCK, CE CXZWXLI KQCSL KJCUDWE PDJB LEWYTB QWPLX JW ILKJXWO CE LEJDXL
QUCELJ. QYXKYLI RO JBL LZQDXL'K KDDEKJL CTLEJK, QXDESLKK ULDC XCSLK BWZL CRWCXI
BLX KJCKBDQ, SYKJWIDCE WF JBL KJWULE QUCEK JBCJ SCE KQGL BLX QLWQUL CEI XLKJWXL
FXLLIWZ JW JBL TCUCVO...

Frequency Analysis:



Plain Text:

2. Start Substitution

Text After Substitution:

IT IS A PERIOD OF CIVIL WAR. REBEL SPACESHIPS, STRIKING FROM A HIDDEN BASE, HAVE
WON THEIR FIRST VICTORY AGAINST THE EVIL GALACTIC EMPIRE. DURING THE BATTLE,
REBEL SPIES MANAGED TO STEAL SECRET PLANS TO THE EMPIRE'S ULTIMATE WEAPON, THE
DEATH STAR, AN ARMORED SPACE STATION WITH ENOUGH POWER TO DESTROY AN ENTIRE
PLANET. PURSUED BY THE EMPIRE'S SINISTER AGENTS, PRINCESS LEIA RACES HOME ABOARD
HER STARSHIP, CUSTODIAN OF THE STOLEN PLANS THAT CAN SAVE HER PEOPLE AND RESTORE
FREEDOM TO THE GALAXY...

Observations:

In this problem we have observed that there is a random mono-alphabetic substitution where the letters A,M,N are not used to hide other letters and letter L is used the most.

2. Problem 2

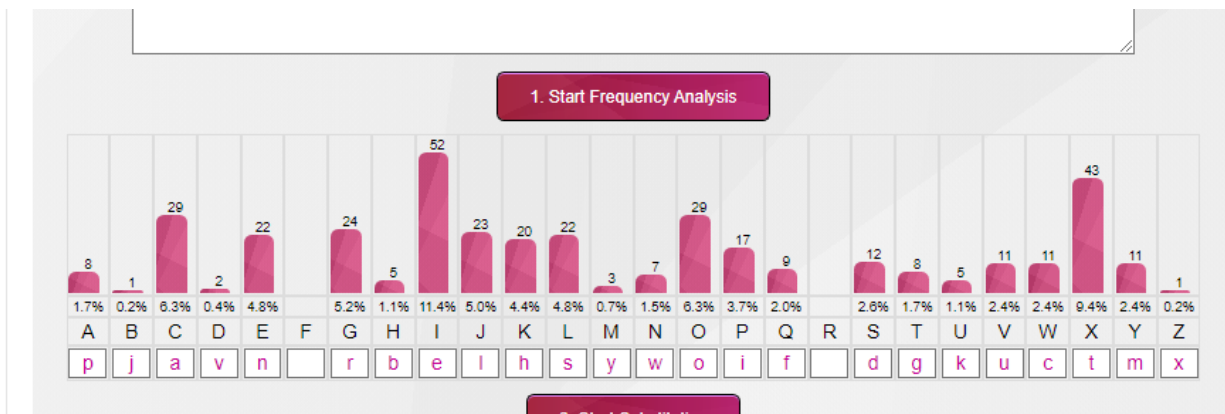
Cipher Text:

Frequency Analysis

Text:

"JVUI LUMNCJUIG KCL GIXVGEIS XO KPL KOYI AJCEIX QQ XCXOPEI PE CE CXXIYAX XO
GILWVI KPL QGPIES KCE LOJO QGOY XKI WJYXWKIL QQ XKI DPJL TCETLXIG BCHHC XKI KVXX.
JFXXJI SOIL JVUI UEON KXCX XKI TCJCWXPW IYAPGI KCL LIWGIXJM HITVE WOELXGVWXPGE OE
C EIN CGYOGIS LACWI LXCXPOE IDIE YOGI AONIGQVJ XKCE XKI QPGLX SGICIS SICXK LXCXG.
NKIE WOYAJIXIS, XKPL VJXPYCXI NICAOE NPJL LAIJJ WIGXCPG SOOY QOG XKI LYCJJ HCES
QQ GIHIJL LXGVTTJPEI XO GILXOGI QGIISOY XO XKI TCJCZM..."

Frequency Analysis:



Plain Text:

2. Start Substitution

Text After Substitution:

"LUKE SKYWALKER HAS RETURNED TO HIS HOME PLANET OF TATOOINE IN AN ATTEMPT TO
RESCUE HIS FRIEND HAN SOLO FROM THE CLUTCHES OF THE VILE GANGSTER JABBA THE HUTT.
LITTLE DOES LUKE KNOW THAT THE GALACTIC EMPIRE HAS SECRETLY BEGUN CONSTRUCTION ON
A NEW ARMORED SPACE STATION EVEN MORE POWERFUL THAN THE FIRST DREADED DEATH STAR.
WHEN COMPLETED, THIS ULTIMATE WEAPON WILL SPELL CERTAIN DOOM FOR THE SMALL BAND
OF REBELS STRUGGLING TO RESTORE FREEDOM TO THE GALAXY..."

Observations:

In this problem we have observed that there is a random mono-alphabetic substitution where the letters F,R are not used to hide other letters and letter I is used the most.

3. Problem 3

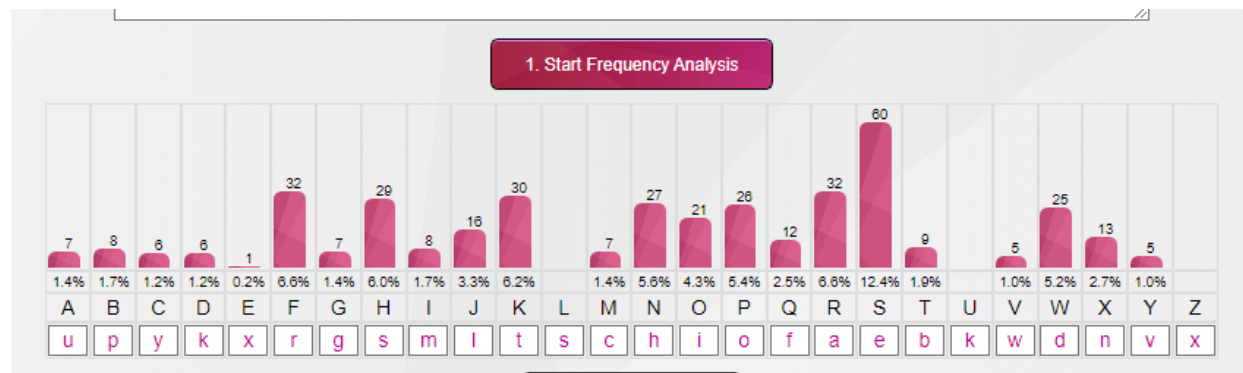
Cipher Text:

Frequency Analysis

Text:

"OK OH R WRFD KOIS QFF KNS FSTJSJJOPX. RJKNFAGN KNS WSRKN HKRF NRH TSSX WSHKFPCSW, OIBSFORJ KFFPBH NRYS WFOYSX KNS FSTSJ QPFMSH QFPI KNSOF NOWWSX TRHS RXW BAFHASW KNSI RMFPHH KNS GRJREC. SYRWXG KNS WFSRWSW OIBSFORJ HKRFQJSSK, R GFFAB PQ QFSSWPI QOGNKSEH JSW TC JADS HDCVRJDSE NRH SHKRTJOHNSW R XSV HSMFSK TRHS PX KNS FSIKPS OMS VPFJW PQ NPKN. KNS SYOJ JPFW WRFKN YRWSE, PTHSHHSW VOKN QOXWOXG CPAXG HDCVRJDSE, NRH WOHBKRMNSW KNPAHRXWH PQ FSIKPS BFPTSH OXKP KNS QRF FSRMNSH PQ HBRMS..."

Frequency Analysis:



Plain Text:

2. Start Substitution

Text After Substitution:

"IT IS A DARK TIME FOR THE REBELLION. ALTHOUGH THE DEATH STAR HAS BEEN DESTROYED, IMPERIAL TROOPS HAVE DRIVEN THE REBEL FORCES FROM THEIR HIDDEN BASE AND PURSUED THEM ACROSS THE GALAXY. EVADING THE DREADED IMPERIAL STARFLEET, A GROUP OF FREEDOM FIGHTERS LED BY LUKE SKYWALKER HAS ESTABLISHED A NEW SECRET BASE ON THE REMOTE ICE WORLD OF HOTH. THE EVIL LORD DARTH VADER, OBSESSED WITH FINDING YOUNG SKYWALKER, HAS DISPATCHED THOUSANDS OF REMOTE PROBES INTO THE FAR REACHES OF SPACE..."

Observations:

In this problem we have observed that there is a random mono-alphabetic substitution where letters L,U are not used to hide other letters and letter S is used the most.

4. Problem 4

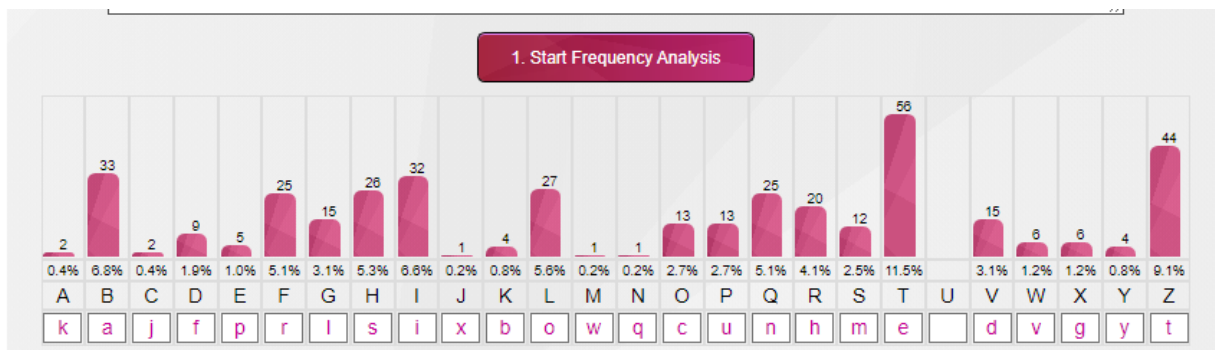
Cipher Text:

Frequency Analysis

Text:

"ZRTFT IH PQFTHZ IQ ZRT XBGBOZIO HTQBZT. HTWTFBG ZRLPHEQV HLGBF HYHZTSH RBWT VIOGBFTV ZRTIF IQZTQZILQH ZL GTBWT ZRT FTEPKGIO. ZRIH HTEBFBZIHZ SLWTSTQZ, PQVTF ZRT GTBVTFHRIE LD ZRT SYHZTFILPH OLPQZ VLLAP, RBH SBVT IZ VIDDIOPGZ DLF ZRT GISIZTV QPSKTF LD CTVI AQIXRZH ZL SBIQZBIQ ETBOT BQV LEVTF IQ ZRT XGBBJY. HTQBZLF BSIVBGB, ZRT DLFSTF NPTTQ LD QBKLL, IH FTZPFQIQX ZL ZRT XBGBOZIO HTQBZT ZL WLZT LQ ZRT OFIZIOBG IHHPT LD OFTBZIOX BQ BFSY LD ZRT FTEPKGIO ZL BHHIHZ ZRT LWTFMRTGSTV CTVI..."

Frequency Analysis:



Plain Text:

2. Start Substitution

Text After Substitution:

"THERE IS UNREST IN THE GALACTIC SENATE. SEVERAL THOUSAND SOLAR SYSTEMS HAVE DECLARED THEIR INTENTIONS TO LEAVE THE REPUBLIC. THIS SEPARATIST MOVEMENT, UNDER THE LEADERSHIP OF THE MYSTERIOUS COUNT DOOKU, HAS MADE IT DIFFICULT FOR THE LIMITED NUMBER OF JEDI KNIGHTS TO MAINTAIN PEACE AND ORDER IN THE GALAXY. SENATOR AMIDALA, THE FORMER QUEEN OF NABOO, IS RETURNING TO THE GALACTIC SENATE TO VOTE ON THE CRITICAL ISSUE OF CREATING AN ARMY OF THE REPUBLIC TO ASSIST THE OVERWHELMED JEDI..."

Observations:

In this problem we have observed that there is a random mono-alphabetic substitution where letter U is not used to hide other letters and letter T is used the most.

5. Problem 5

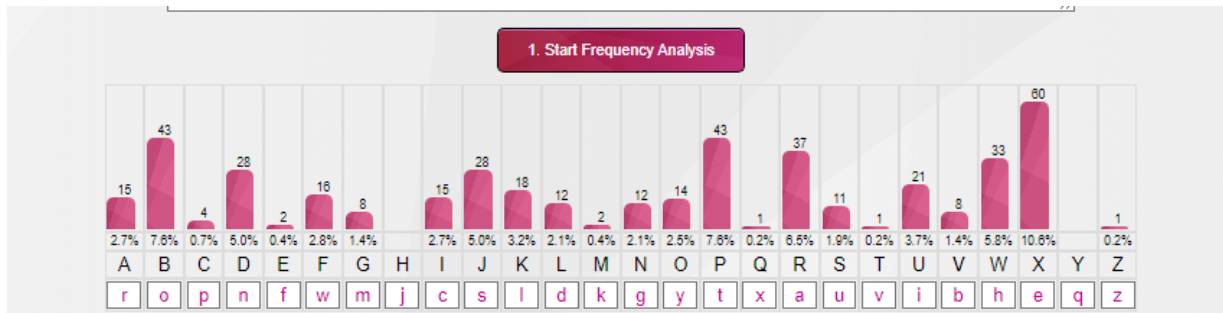
Cipher Text:

Frequency Analysis

Text:

"FX IWBBJX PB NB PB FWX GBBB. VSP FWO, JBGX JRO, FWX GBBB? FWO IWBBJX FWUJ RJ BSA NBRK? RDL FWXO GRO FXKK RJM FWO IKUGV FWX WUNWJF GBSDFRUD? FWO, 35 OKRAJ RNB, EKO FWX REKDFUI? FWO LBXJ AUIX CKRO EXQRJ? FX IWBBJX PB NB PB FWX GBBB UD FWUJ LXIRLX RDL LB FWX BFWXA FWUDNJ, DBF VXIRSJX FWXO RAX XRJO, VSP VXIRSJX FWXO RAX WRAL, VXIRSJX FWRP NBRK FUKK JXATX PB BANRDUZX RDL GXRJSAX FWX VXJP BE BSA XDXANUXJ RDL JMUUKJ, VXIRSJX FWRP IWRKKXDNX UJ BDX FWRP FX RAX FUKKUDN PB RIIXCP, BDX FX RAX SDFUKKUDN PB CBJPCBDX, RDL BDX FWUW FX UDFXDL PB FUD, RDL FWX BFWXAJ, PBB."

Frequency Analysis:



Plain Text:

2. Start Substitution

Text After Substitution:

"WE CHOOSE TO GO TO THE MOON. BUT WHY, SOME SAY, THE MOON? WHY CHOOSE THIS AS OUR GOAL? AND THEY MAY WELL ASK WHY CLIMB THE HIGHEST MOUNTAIN? WHY, 35 YEARS AGO, FLY THE ATLANTIC? WHY DOES RICE PLAY TEXAS? WE CHOOSE TO GO TO THE MOON IN THIS DECADE AND DO THE OTHER THINGS, NOT BECAUSE THEY ARE EASY, BUT BECAUSE THEY ARE HARD, BECAUSE THAT GOAL WILL SERVE TO ORGANIZE AND MEASURE THE BEST OF OUR ENERGIES AND SKILLS, BECAUSE THAT CHALLENGE IS ONE THAT WE ARE WILLING TO ACCEPT, ONE WE ARE UNWILLING TO POSTPONE, AND ONE WHICH WE INTEND TO WIN, AND THE OTHERS, TOO."

Observations:

In this problem we have observed that there is a random mono-alphabetic substitution where letters H,Y is not used to hide other letters and letter X is used the most.

6. Problem 6

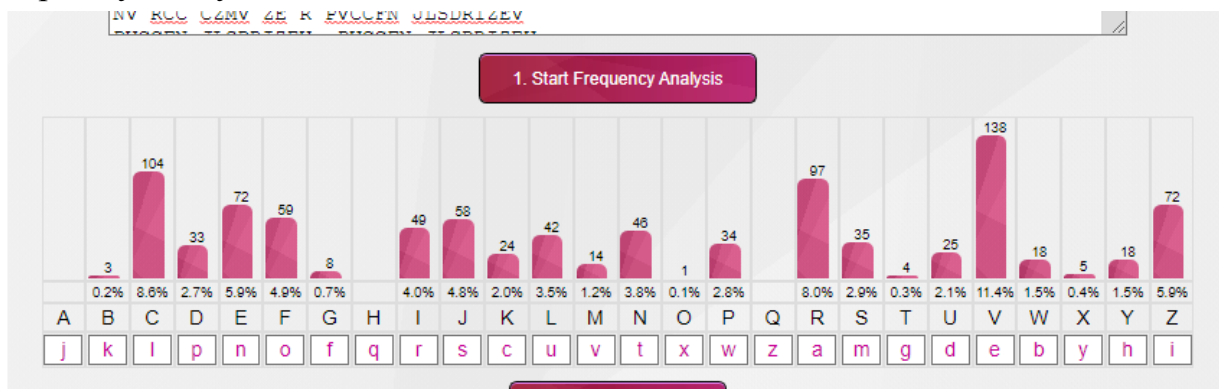
Cipher Text:

Frequency Analysis

Text:

```
"ZE KYV KFNE NYVIV Z NRJ SFIE
CZMVU R DRE NYF JRZCVU KF JVR
REU YV KFCU LJ FW YZJ CZNV
ZE KYV CREU FW JLSDRIZEVJ
JF NV JRZCVU LG KF KYV JLE
'KZC NV WFLEU R JVR FW XIVVE
REU NV CZMVU SVEVRKY KYV NRMVJ
ZE FLI FVCCFN JLSDRIZEV
NV RCC CZMV ZE R FVCCFN JLSDRIZEV
FVCCFN JLSDRIZEV, FVCCFN JLSDRIZEV
NV RCC CZMV ZE R FVCCFN JLSDRIZEV
FVCCFN JLSDRIZEV, FVCCFN JLSDRIZEV
```

Frequency Analysis:



Plain Text:

2. Start Substitution

Text After Substitution:

```
"IN CHE COTN THERE I TAS MORN
LIVED A PAN THO SAILED CO SEA
AND HE COLD US OB HIS LIBE
IN CHE LAND OB SUMPARINES
SO TE SAILED UF CO CHE SUN
'CIL TE BOUND A SEA OB YREEN
AND TE LIVED MENEACH CHE TAVES
IN OUR WELLOT SUMPARINE
TE ALL LIVE IN A WELLOT SUMPARINE
WELLOT SUMPARINE, WELLOT SUMPARINE
TE ALL LIVE IN A WELLOT SUMPARINE
WELLOT SUMPARINE, WELLOT SUMPARINE
```

Observations:

In this problem we have observed that there is a random mono-alphabetic substitution where letters A,H,Q are not used to hide other letters and letter V is used the most.

7. Problem 7

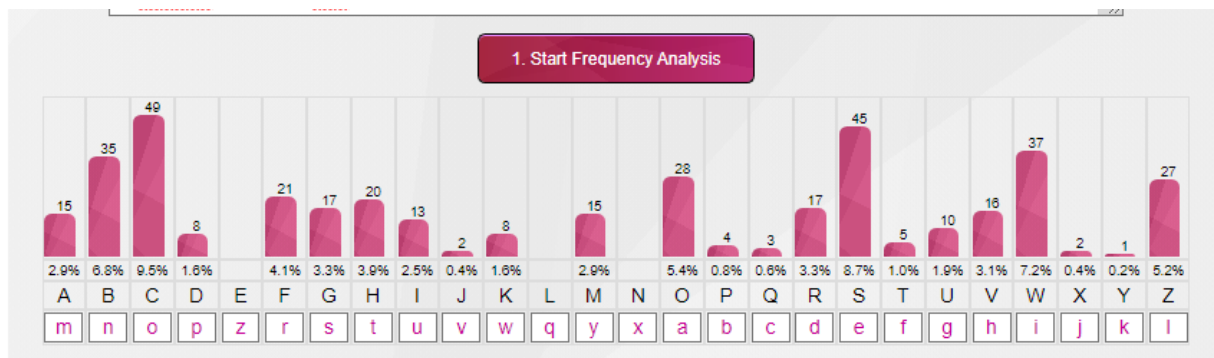
Cipher Text:

Frequency Analysis

Text:

```
"WAOUWBS HVSFS'G BC QCIBHFWSG
WH WGB'H VOFR HC RC
BCHVWBU HC YWZZ CF RWS TCF
QBR BC FSZWUNCB, HCC
WAOUWBS OZZ HVS DSCDZS
ZWJWBU ZWTB WB DSOQS
MCI, MCI AOM GOM W'A O RFSOASF
PIH W'A BCH HVS CBZM CBS
W VCDS GCASROM MCI KWZZ XCWB IG
QBR HVS KCFZR KWZZ PS OG CBS
WAOUWBS BC DCGSGGWCBCG
W KCBRSF WT MCI QOB
```

Frequency Analysis:



Plain Text:

2. Start Substitution

Text After Substitution:

```
"IMAGINE THERE'S NO COUNTRIES
IT ISN'T HARD TO DO
NOTHING TO KILL OR DIE FOR
AND NO RELIGION, TOO
IMAGINE ALL THE PEOPLE
LIVING LIFE IN PEACE
YOU, YOU MAY SAY I'M A DREAMER
BUT I'M NOT THE ONLY ONE
I HOPE SOMEDAY YOU WILL JOIN US
AND THE WORLD WILL BE AS ONE
IMAGINE NO POSSESSIONS
I WONDER IF YOU CAN
```

Observations:

In this problem we have observed that there is a random mono-alphabetic substitution where letters E,L,N are not used to hide other letters and letter C is used the most.

8. Problem 8

Cipher Text:

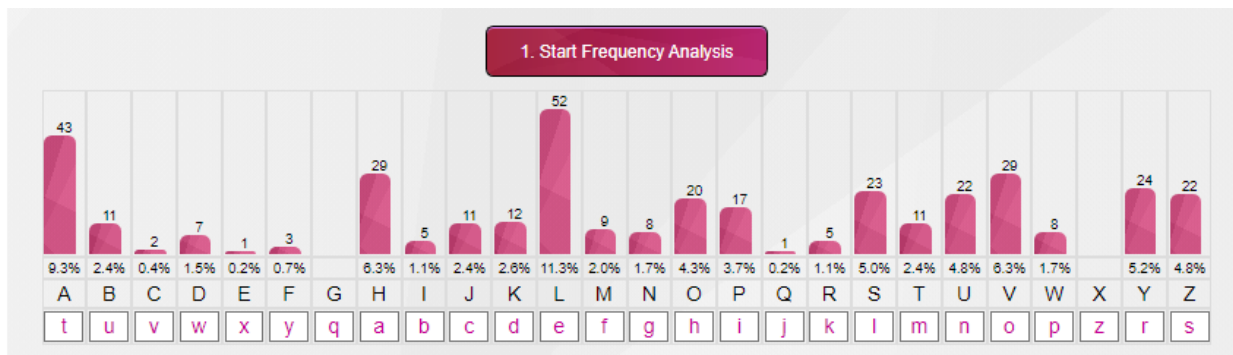
Frequency Analysis

Text:

"SBRL ZRFDSRLY OHZ YLABYULK AV OPZ OVTL WSHULA VM AHAVVPUL PU HU HAALTWA AV
YLZJBL OPZ MYPLUK OHU ZVSV MYVT AOL JSBAJOLZ VM AOL CPSP NHUNZALY QHIIH AOL OBAA.
SPAASL KVLZ SBRL RUVD ACHA AOL NHSHJAPJ LTWPYL OHZ ZLJYLASF ILNBV JVUZAYBJAPVU VU
H ULD HYTVYLK ZWHJL ZAHAPVU LCLU TVYL WVDLYMBS ACHU AOL MPYZA KYLHKLK KLHAO ZAHY.

DOLU JVTWSLALK, AOPZ BSAPTHAL DLHWVU DPSS ZWLSS JLYAHPU KVVV MVY AOL ZTHSS IHUK
VM YLILSZ ZAYBNNSPUN AV YLZAVYL MYLLKVT AV AOL NHSHEF..."

Frequency Analysis:



Plain Text:

2. Start Substitution

Text After Substitution:

"LUKE SKYWALKER HAS RETURNED TO HIS HOME PLANET OF TATOOINE IN AN ATTEMPT TO
RESCUE HIS FRIEND HAN SOLO FROM THE CLUTCHES OF THE VILE GANGSTER JABBA THE HUTT.
LITTLE DOES LUKE KNOW THAT THE GALACTIC EMPIRE HAS SECRETLY BEGUN CONSTRUCTION ON
A NEW ARMORED SPACE STATION EVEN MORE POWERFUL THAN THE FIRST DREADED DEATH STAR.

WHEN COMPLETED, THIS ULTIMATE WEAPON WILL SPELL CERTAIN DOOM FOR THE SMALL BAND
OF REBELS STRUGGLING TO RESTORE FREEDOM TO THE GALAXY..."

Observations:

In this problem we have observed that there is a random mono-alphabetic substitution where letters G,X are not used to hide other letters and letter L is used the most.

Conclusion:

We have successfully broken the mono-alphabetic substitution cipher using frequency analysis and understood how the mono-alphabetical substitution cipher works.