NAME: Sarvesh Patil
CLASS: D15A
ROLL NO: 46

LAB 01

LAB 01: To understand the process of Breaking the Mono-alphabetic Substitution Cipher using the Frequency analysis method

ROLL NO	46
NAME	Sarvesh Patil
CLASS	D15A
SUBJECT	Internet Security Lab
LO MAPPED	LO1: To apply the knowledge of symmetric cryptography to implement classical ciphers

NAME: Sarvesh Patil
CLASS: D15A
ROLL NO: 46

AIM:
To understand the process of Breaking the Mono-alphabetic Substitution Cipher using Frequency analysis method

INTRODUCTION:
Mono-Alphabetic Substitution Cipher
A mono-alphabetic cipher (aka simple substitution cipher) is a substitution cipher where each letter of the plain text is replaced with another letter of the alphabet. It uses a fixed key which consists of the 26 letters of a “shuffled alphabet”.

Plain text alphabet:	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
Cipher text alphabet (key):	M	U	A	L	V	O	Z	K	R	N	J	X	Q	D	F	S	H	P	E	B	C

With the above key, all “A” letters in the plain text will be encoded to an “M”.

This type of cipher is a form of symmetric encryption as the same key can be used to both encrypt and decrypt a message.

A monoalphabetic substitution is a cipher in which each occurrence of a plaintext symbol is replaced by a corresponding ciphertext symbol to generate the ciphertext. The key for such a cipher is a table of the correspondence or a function from which the correspondence is computed.

Example: An affine cipher E(x) = (ax + b) MOD 26 is an example of a monoalphabetic substitution.

There are other ways to “generate” a monoalphabetic substitution.

Alphabet Mixing via a Keyword
A keyword or keyphrase can be used to mix the letters to generate the cipher alphabet.

Example: If the keyword is ANDREW DICKSON WHITE, then the cipher alphabet is given by

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

Perhaps a better keyword is EZRA CORNELL:
plain 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
cipher E Z R A C O N L B D F G H I J K M P Q S T U V W X Y

NAME: Sarvesh Patil
CLASS: D15A
ROLL NO: 46

Alphabet Mixing via a Columnar Transposition

The letters from the keyword form the headings of the columns, and the remaining letters of the alphabet fill in order in the rows below. Mixing is achieved by transcribing columns.

Example: If the keyword is CORNELL, then write
C O R N E L
A B D F G H
I J K M P Q
S T U V W X
Y Z
so that transcribing columns left-to-right gives the substitution

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

For instance, FAR ABOVE CAYUGA'S WATERS is enciphered as OCVCA NWYIC QPBCE LCGYE.

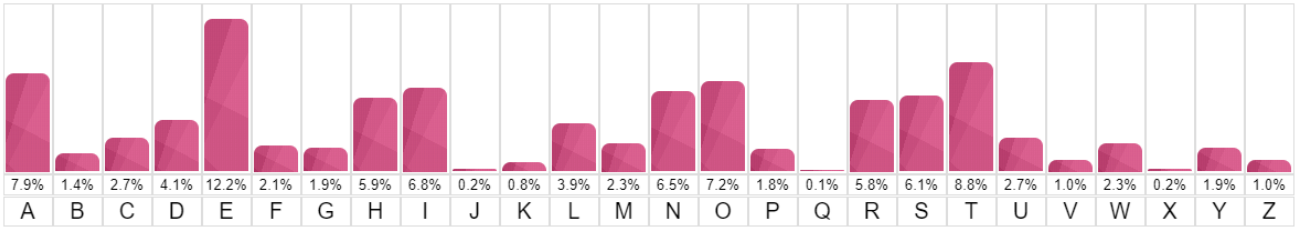
NAME: Sarvesh Patil
CLASS: D15A
ROLL NO: 46

Frequency Analysis

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 in breaking substitution ciphers (e.g. mono-alphabetic substitution cipher, Caesar shift cipher, Vatsyayana 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 the English language, letters E, T, A, and O are the most common, while letters Z, Q, and X are not as frequently used.

The following chart shows the frequency of each letter of the alphabet for the English language:



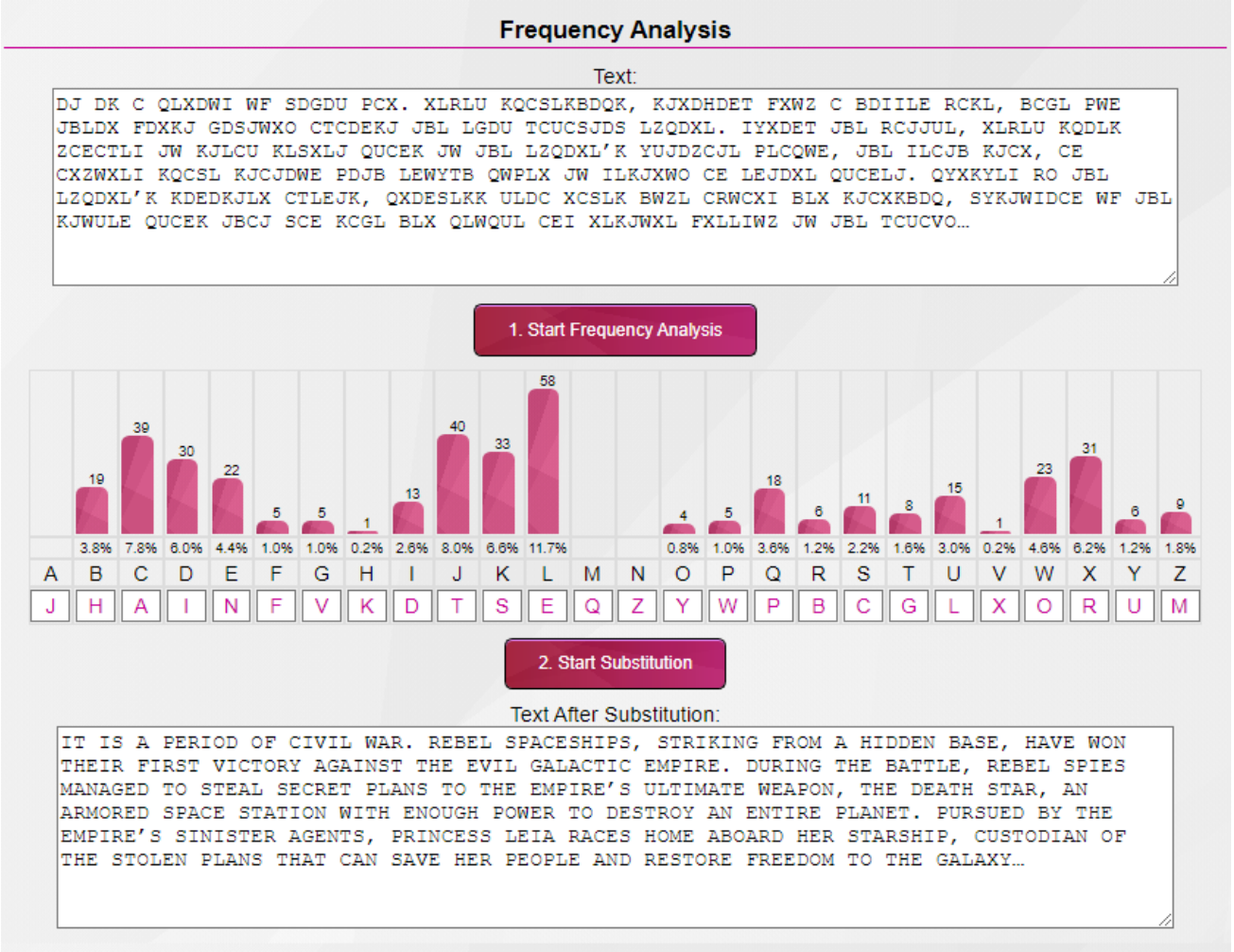
We can assume that most samples of text written in English would have a similar distribution of letters. However, this is only true if the sample of the text is long enough. A very short text may lead to a significantly different distribution.

When trying to decrypt a cipher text based on a substitution cipher, we can use frequency analysis to help identify the most recurring letters in a cipher text and hence make the hypothesis of what these letters have been encoded as (e.g. E, T, A, O, etc). This will help us decrypt some of the letters in the text. We can then recognize patterns/words in the partly decoded text to identify more substitutions.

NAME: Sarvesh Patil
CLASS: D15A
ROLL NO: 46

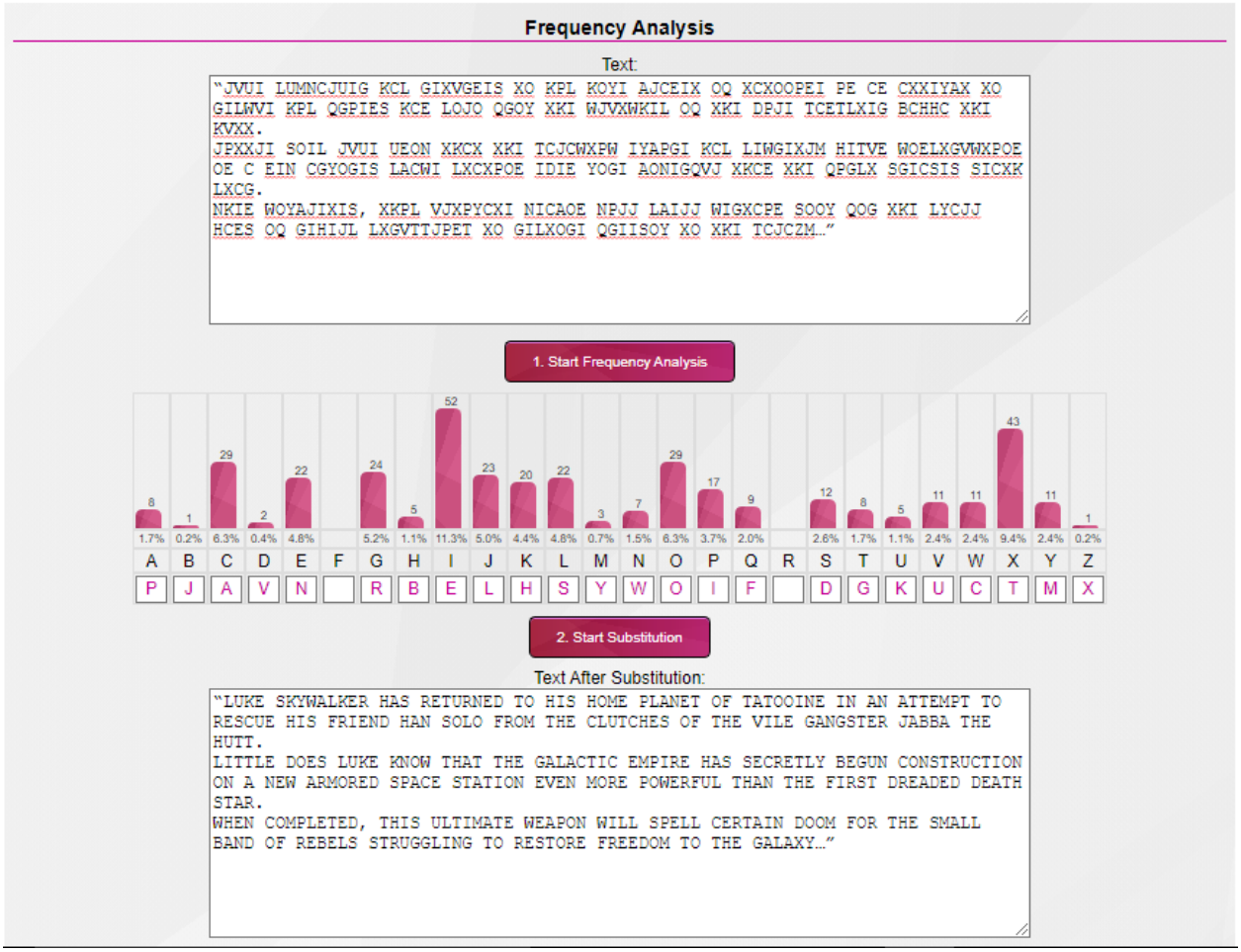
RESULTS:

1. **CIPHER #1**
   KEY: RANDOM



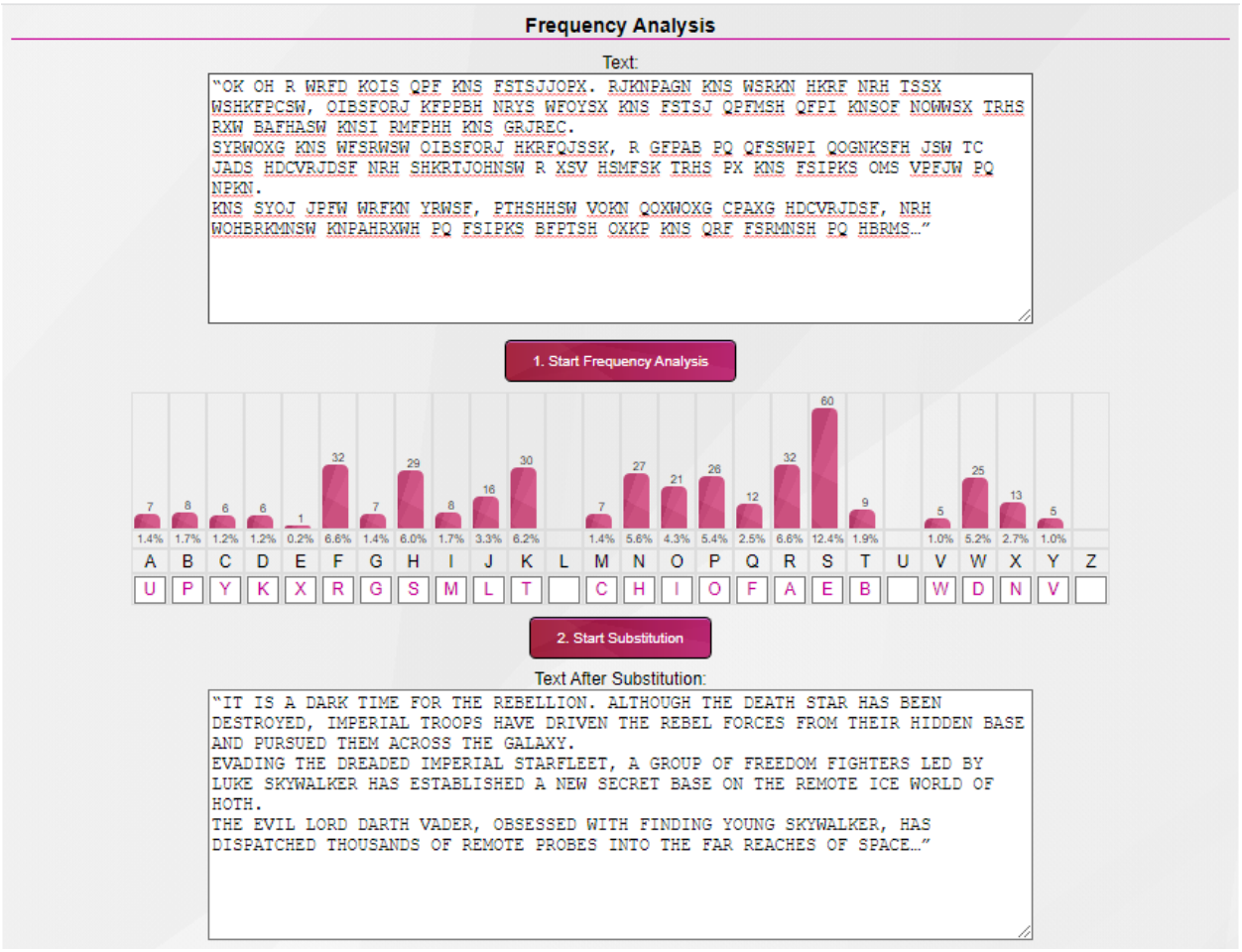
NAME: Sarvesh Patil
CLASS: D15A
ROLL NO: 46

2. CIPHER #2
KEY: RANDOM



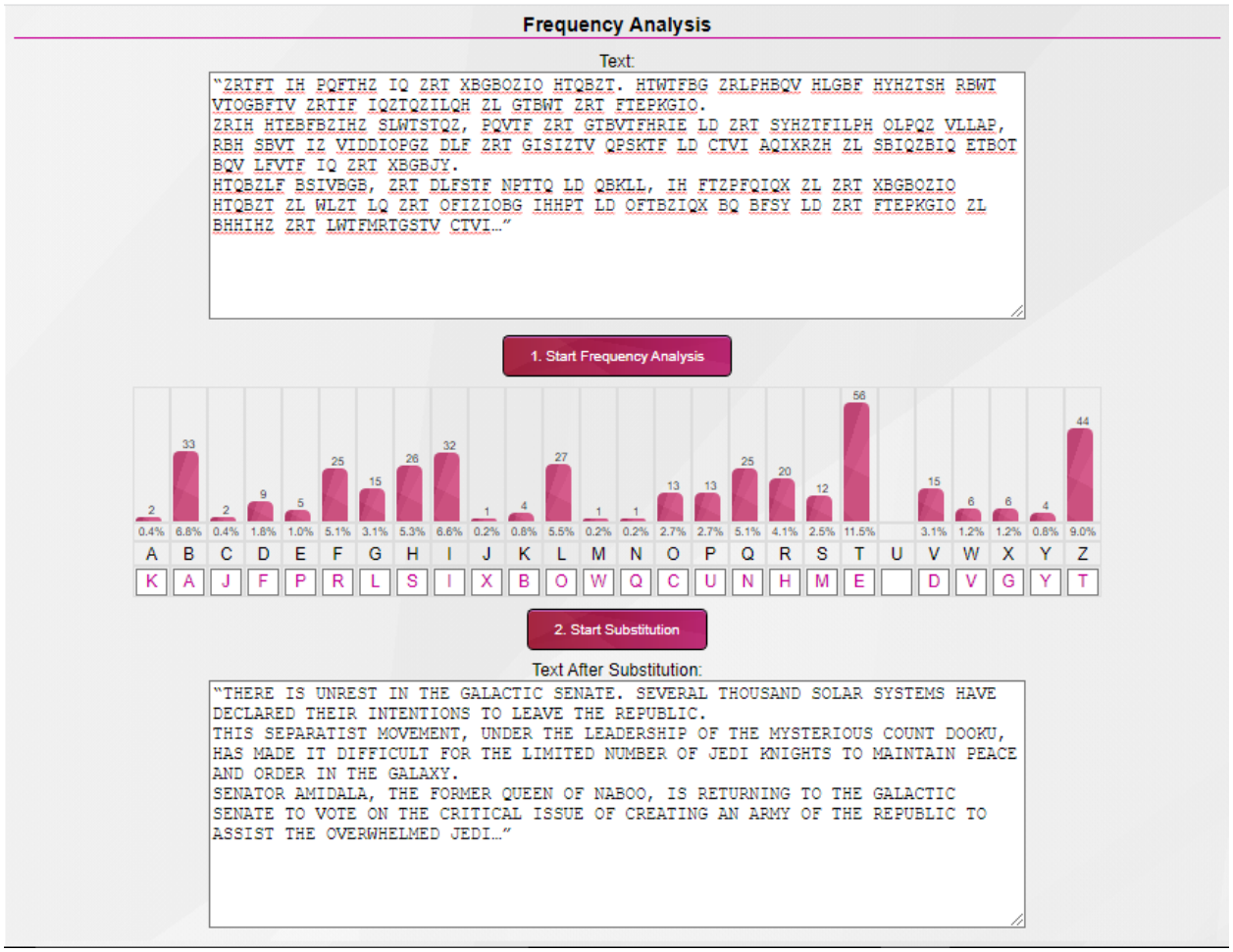
NAME: Sarvesh Patil
CLASS: D15A
ROLL NO: 46

3. CIPHER #3
KEY: RANDOM



NAME: Sarvesh Patil
CLASS: D15A
ROLL NO: 46

4. CIPHER #4
KEY: RANDOM



NAME: Sarvesh Patil
CLASS: D15A
ROLL NO: 46

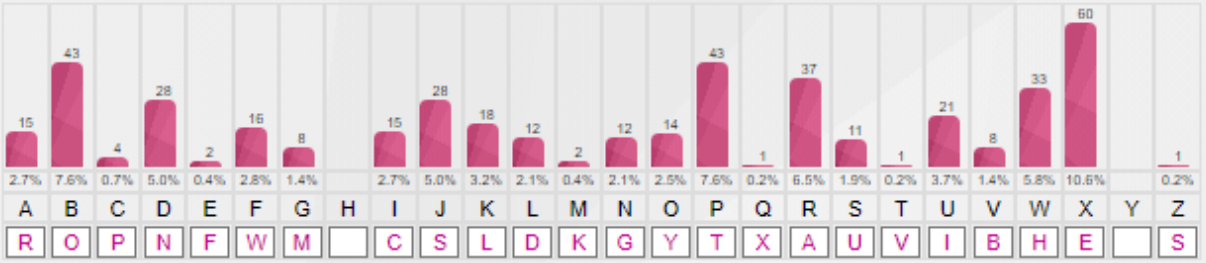
5. CIPHER #5
KEY: RANDOM

Frequency Analysis

Text:

"FX IWBBJX PB NB PB PWX GBBD. VSP FWO, JBGX JRO, PWX GBBD? FWO IWBBJX PWUJ RJ BSA NBRK? RDL PWXO GRO FXKK RJM FWO IKUGV PWX WUNWXJP GBSDPRUD? FWO, 35 OXRAJ RNB, EKO PWX RPKRDPUI? FWO LBXJ AUIX CKRO PXQRJ? FX IWBBJX PB NB PB PWX GBBD UD PWUJ LXIRLX RDL LB PWX BPWXA PWUDNJ, DBP VXIRSJX PWXO RAX XRJO, VSP VXIRSJX PWXO RAX WRAL, VXIRSJX PWRP NBRK FUKK JXATX PB BANRDUZX RDL GXRJSAX PWX VXJP BE BSA XDXANUXJ RDL JMUKKJ, VXIRSJX PWRP IWRKKXDNX UJ BDX PWRP FX RAX FUKKUDN PB RIIXCP, BDX FX RAX SDFUKKUDN PB CBJPCBDX, RDL BDX FWUIW FX UDPXDL PB FUD, RDL PWX BPWXAJ, PBB."

1. Start Frequency Analysis



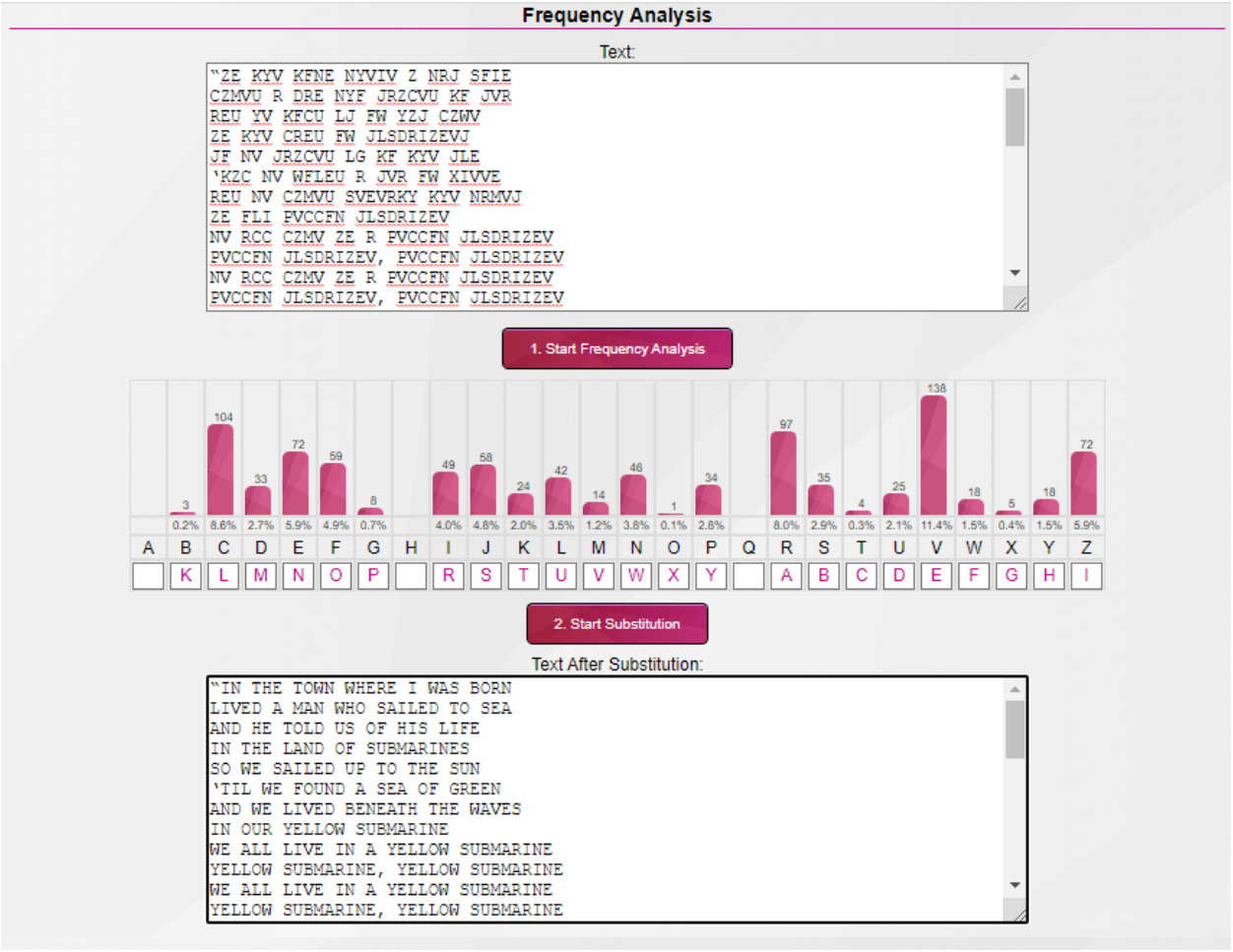
2. Start Substitution

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 ORGANISE 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."

NAME: Sarvesh Patil
CLASS: D15A
ROLL NO: 46

6. CIPHER #6
KEY: 10



NAME: Sarvesh Patil
CLASS: D15A
ROLL NO: 46

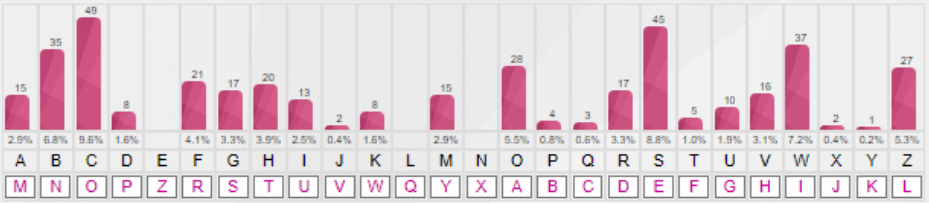
7. CIPHER #7
KEY: 12

Frequency Analysis

Text:

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

1. Start Frequency Analysis



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
NO NEED FOR GREED OR HUNGER
```

NAME: Sarvesh Patil
CLASS: D15A
ROLL NO: 46

8. CIPHER #8

KEY: 12

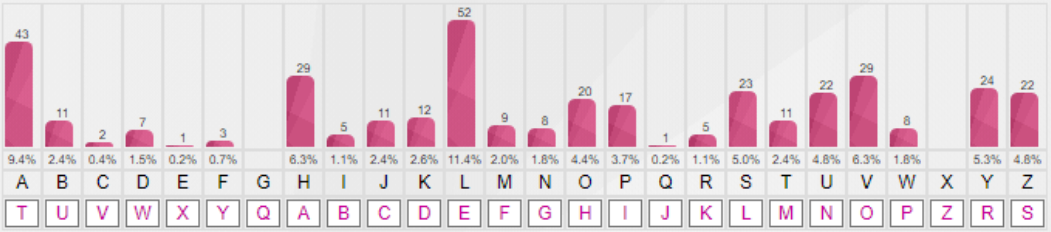
Frequency Analysis

Text:

```
SBRL ZRFDHSRLY OHZ YLABYULK AV OPZ OVTL WSHULA VM AHAVVPUL PU HU HAALTWA AV YLZJBL OPZ
MYPLUK OHU ZVSV MYVT AOL JSBAJOLZ VM AOL CPSL NHUNZALY QHIIH AOL OBAA.
SPAASL KVLZ SBRL RUVD AOHA AOL NHSHJAPJ LTWPYL OHZ ZLJYLASF ILNBU JVUZAYBJAPVU VU H ULD
HYTVYLK ZWHJL ZAHAPVU LCLU TVYL WVDLYMBS AOHU AOL MPYZA KYLHKLK KLHAO ZAHY.

DOLU JVTWSLALK, AOPZ BSAPTHAL DLHWVU DPSS ZWLSS JLYAHPU KVVT MVY AOL ZTHSS IHUK VM YLILSZ
ZAYBNNSPUN AV YLZAVYL MYLLKVT AV AOL NHSHEF…
```

1. Start Frequency Analysis



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…
```

NAME: Sarvesh Patil
CLASS: D15A
ROLL NO: 46

ANALYSIS:

In most of the problems, it is observed that “e” is being repeated the most number of times. and “a” is repeated the 2nd most number of times. Many problems had random keys but some had keys like 10, 12, and 19

CONCLUSION:

We have successfully understood the process of Breaking the Mono-alphabetic Substitution Cipher using the Frequency analysis method