

# ASSIGNMENT 1

## CLASSICAL CIPHERS

```
In [1]: import re
        from collections import defaultdict

In [2]: def preprocess_text(text):
        """Remove non-letter characters and convert to lowercase."""
        return re.sub(r'^a-zA-Z', '', text).lower()

In [3]: def count_frequencies(text, n):
        """Count n-gram frequencies."""
        freq_dict = defaultdict(int)
        for i in range(len(text) - n + 1):
            ngram = text[i:i + n]
            freq_dict[ngram] += 1
        return freq_dict

In [4]: def sort_frequencies(freq_dict):
        """Sort frequencies in decreasing order."""
        return sorted(freq_dict.items(), key=lambda item: item[1], reverse=True)

In [5]: def print_frequencies(freq_list, n=None):
        """Print the top n frequencies or all if n is None."""
        for i, (ngram, count) in enumerate(freq_list[:n]):
            print(f"{ngram}: {count}")

In [6]: def analyze_text(text, ngram_sizes=[1, 2, 3], top_n=30):
        """Analyze and print the frequencies of n-grams in the text."""
        preprocessed_text = preprocess_text(text)

        for n in ngram_sizes:
            print(f"\nTop {top_n} {n}-grams:")
            freq_dict = count_frequencies(preprocessed_text, n)
            sorted_freq_list = sort_frequencies(freq_dict)
            print_frequencies(sorted_freq_list, top_n if n > 1 else None)
            if n == 1:
                print("\nAll single-letter frequencies:")
                print_frequencies(sorted_freq_list) # Print all single-letter frequencies

In [7]: def main():
        text1 = """To be, or not to be, that is the question: Whether 'tis nobler in the mind to suffer
        The slings and arrows of outrageous fortune, Or to take arms against a sea of troubles
        And by opposing end them."""

        text2 = """In the beginning God created the heaven and the earth. And the earth was without form, and void;
        and darkness was upon the face of the deep. And the Spirit of God moved upon the face of the waters."""

        print("Text 1 Analysis:")
        analyze_text(text1)

        print("\n\nText 2 Analysis:")
        analyze_text(text2)

In [8]: if __name__ == "__main__":
        main()
```

**Output:**

**Analyzing text 1**

Text 1 Analysis:

Top 30 1-grams:

t: 19  
o: 18  
e: 17  
s: 13  
n: 12  
r: 11  
a: 11  
i: 8  
h: 7  
u: 6  
b: 5  
f: 5  
d: 4  
g: 4  
l: 3  
m: 3

w: 2  
p: 2  
q: 1  
k: 1  
y: 1

All single-letter frequencies:

t: 19  
o: 18  
e: 17  
s: 13  
n: 12  
r: 11  
a: 11  
i: 8  
h: 7  
u: 6  
b: 5  
f: 5

d: 4  
g: 4  
l: 3  
m: 3  
w: 2  
p: 2  
q: 1  
k: 1  
y: 1

Top 30 2-grams:

th: 6  
he: 6  
in: 5  
to: 4  
rt: 4  
nd: 4  
ob: 3  
eo: 3

or: 3  
ti: 3  
st: 3  
es: 3  
er: 3  
sa: 3  
ou: 3  
be: 2  
no: 2  
ot: 2  
et: 2  
is: 2  
bl: 2  
le: 2  
em: 2  
dt: 2  
os: 2  
ng: 2  
an: 2

```
ar: 2
ro: 2
of: 2
```

Top 30 3-grams:

```
the: 5
tob: 2
obe: 2
eor: 2
eth: 2
tis: 2
ert: 2
ble: 2
hem: 2
ndt: 2
ing: 2
san: 2
and: 2
ort: 2
```

```
beo: 1
orn: 1
rno: 1
not: 1
ott: 1
tto: 1
bet: 1
tha: 1
hat: 1
ati: 1
ist: 1
sth: 1
heq: 1
equ: 1
que: 1
ues: 1
```

## Analyzing text 2

Text 2 Analysis:

Top 30 1-grams:

```
e: 23
t: 16
a: 15
n: 13
h: 13
d: 12
o: 11
i: 7
r: 7
s: 6
f: 6
g: 4
w: 4
p: 4
c: 3
v: 3
```

```
u: 3
m: 2
b: 1
k: 1
```

All single-letter frequencies:

```
e: 23
t: 16
a: 15
n: 13
h: 13
d: 12
o: 11
i: 7
r: 7
s: 6
f: 6
g: 4
```

w: 4  
p: 4  
c: 3  
v: 3  
u: 3  
m: 2  
b: 1  
k: 1

Top 30 2-grams:  
th: 12  
he: 10  
an: 5  
nd: 5  
ea: 4  
dt: 4  
in: 3  
nt: 3  
ed: 3

ee: 3  
ar: 3  
wa: 3  
of: 3  
go: 2  
od: 2  
at: 2  
te: 2  
ve: 2  
rt: 2  
as: 2  
sw: 2  
it: 2  
da: 2  
es: 2  
up: 2  
po: 2  
on: 2  
ef: 2

fa: 2  
ac: 2

Top 30 3-grams:  
the: 9  
and: 5  
dth: 4  
nth: 3  
ndt: 3  
god: 2  
ate: 2  
hee: 2  
eea: 2  
ear: 2  
art: 2  
rth: 2  
was: 2  
upo: 2  
pon: 2

ont: 2  
hef: 2  
efa: 2  
fac: 2  
ace: 2  
ceo: 2  
eof: 2  
oft: 2  
fth: 2  
int: 1  
heb: 1  
ebe: 1  
beg: 1  
egi: 1  
gin: 1

For the fourth question I tried manual mapping. So I found a few repetitive clusters of letters such as 'bt', 'jpx', 'ci' and 'm' which can be mapped as 'in', 'the', 'is' and 'a' respectively.

So after mapping the words the decrypted text is as follows.

```
# Part (d): Decrypt the given ciphertext using frequency analysis
ciphertext = """"bt jpx rmlx pcuv amlx icvjp ibtwxvr ci m lmt'r pmtn, mtn yvcjx cdxv mwmbtrj jpx amtngxrjbah uqct
jpx qgmrxv ci jpx ymgi ci jpx hbtw'r qmgmax; mtn jpx hbtw rmy jpx qmvj ci jpx pmtn jpmj yvcjx.
jpxt jpx hbtw'r acutjxtmtax ymr apmtwxn, mtn pbr jpcuwpjr jvcufgxn pbl, rc jpmj jpx scbtjr ci pbr
gcbtr yxvx gccrxn, mtn pbr htxxr rlcjx ctx mwmbtrj mtcjpxv. jpx hbtw avbxn mgcun jc fvbwt bt jpx
mrjvcgcwxvr, jpx apmgxnmtr, mtn jpx rccjprmexvr. mtn jpx hbtw qmhhx, mtn rmbn jc jpx ybrx lxt
ci fmfegct, ypcrcxdxv rpmgg vxmn jpbv yvbjbwt, mtn rpy lx jpx btjxvqxjmbct jpxvxci, rpmgg fx
agcjpjn ybjp ramvgxj, mtn pmx m apmbt ci wcn mfcuj pbr txah, mtn rpmgg fx jpx jpbvn vugxv
bt jpx hbtwncl. jpxt amlx bt mgi jpx hbtw'r ybrx lxt; fuj jpxe acugn tcj vxmn jpx yvbjbwt, tcv lmhx
htcyt jc jpx hbtw jpx btjxvqxjmbct jpxvxci. jpxt ymr hbtw fxgrpmoomv wvxmjge jvcufgxn, mtn
pbr acutjxtmtax ymr apmtwxn bt pbl, mtn pbr gcvnr yxvx mrjctbrpxn. tcy jpx kuxxt, fe vxmrct ci
jpx ycvnr ci jpx hbtw mtn pbr gcvnr, amlx btjc jpx fmkuxj pcux; mtn jpx kuxxt qmhhx mtn rmbn, c
hbtw, gbdx icvxdxv; gxj tcj jpe jpcuwpjr jvcufgxn jpxx, tcv gxj jpe acutjxtmtax fx apmtwxn; jpxvx br
m lmt bt jpe hbtwncl, bt ypc l br jpx qgbvj ci jpx pcge wcn; mtn bt jpx nmer ci jpe ybrncl ci jpx
wcnr, ymr icutn bt pbl; ypc l jpx hbtw txfuapmntxoomv jpe imjpxv, jpx hbtw, b rme, jpe imjpxv,
lmnx lmrjxv ci jpx lmbwabmtr, mrjvcgcwxvr, apmgxnmtr, mtn rccjprmexvr; icvmluap mr mt
xzaxgxtj qgbvj, mtn htcygxnw, mtn utnxvrjmtnbtw, btjxvqxjbtw ci nvxmlr, mtn rpybtw ci
pmvn rxtjxtaxr, mtn nbrccgdbtw ci ncufjr, yxvx icutn bt jpx rmlx nmtbxg, ypc l jpx hbtw tmlxn
fxgjxrpoomv; tcy gxj nmtbxg fx amggxn, mtn px ybgg rpy jpx btjxvqxjmbct""""

print("\n\nAnalyzing Ciphertext:")
single_letter_freqs = analyze_text(ciphertext)

# Example of a mapping based on frequency analysis
mapping = {

    'b': 'i',
    't': 'n',
    'j': 't',
    'p': 'h',
    'x': 'e',
    'c': 'o',
    'i': 'f',
    'm': 'a'
}

# Decrypt the ciphertext using the mapping
decrypted_text = decrypt(ciphertext, mapping)

print("\nDecrypted Text:\n")
print(decrypted_text)

if __name__ == "__main__":
    main()
```

This gave me the following partially decrypted code.

```
in the rale houv aale fovth finwevr of a lan'r hann, ann yvote odev awainrt the aanngertiah uqon
the qgartev of the yagg of the hinw'r qagaa; ann the hinw ray the qavt of the hann that yvote.
then the hinw'r aountenanae yar ahanwen, ann hir thouwhtr tvoufgen hil, ro that the sointr of hir
goinr yeve gooren, ann hir hneer rlote one awainrt anothev. the hinw avien agoun to fvinw in the
artvogovevr, the ahagneanr, ann the roothraeevr. ann the hinw rqahe, ann rain to the yire len
of fafegon, yhoroedev rhagg vean thir yvitinw, ann rhou le the intevqvetaion theveof, rhagg fe
agothen yith raavget, ann hade a ahain of wogn afout hir neah, ann rhagg fe the thivn vugev
in the hinwnol. then aale in agg the hinw'r yire len; fut thee aougn not vean the yvitinw, nov lahe
hnoyn to the hinw the intevqvetaion theveof. then yar hinw fegrhaoav wveatge tvoufgen, ann
hir aountenanae yar ahanwen in hil, ann hir govnr yeve artonirhen. noy the kuen, fe vearon of
the yovnr of the hinw ann hir govnr, aale into the fankuet houre; ann the kuen rqahe ann rain, o
hinw, gide fovedev; get not the thouwhtr tvoufge thee, nov get the aountenanae fe ahanwen; theve ir
a lan in the hinwnol, in yhol ir the rqivit of the hoge wonr; ann in the naer of the yirrol of the
wonr, yar founn in hil; yhol the hinw nefuahanneoav the fathev, the hinw, i rae, the fathev,
lane lartev of the lawiaianr, artvogovevr, ahagneanr, ann roothraeevr; fovarluah ar an
ezaeggent rqivit, ann hnoygenwe, ann unnevrtanninw, intevqvetaion of nvealr, ann rhouinw of
havn rentenaer, ann nirrogdinw of nouftr, yeve founn in the rale nanieg, yhol the hinw nalen
fegterhaoav; noy get nanieg fe aaggen, ann he yigg rhou the intevqvetaion
```

In the words such as 'houv', 'fovth', 'fathev' letter 'v' can be replaced with 'r' which gives us 'hour', 'forth' and 'father' respectively. And the letter 'n' can be replaced with letter 'd' which converts the words 'ann' as 'and'.

After this mapping, the below is the decrypted code.

```
in the rale hour aale forth finwerr of a lan'r hand, and yrote oder awainrt the aandgertiah upon
the qgarter of the yagg of the hinw'r qagaae; and the hinw ray the qart of the hand that yrote.
then the hinw'r aountenanae yar ahanwed, and hir thouwhtr troufged hil, ro that the sointr of hir
goindr yere goored, and hir hneer rrote one awainrt another. the hinw aried agoud to frinw in the
artrogowerr, the ahagdeanr, and the roothraeerr. and the hinw rqahe, and raid to the yire len
of fafegon, yhoroeder rhagg read thir yritinw, and rhou le the interqretation thereof, rhagg fe
agothed yith raarget, and hade a ahain of wogd afout hir neah, and rhagg fe the third ruger
in the hinwdol. then aale in agg the hinw'r yire len; fut thee aougnd not read the yritinw, nor lahe
hnoyn to the hinw the interqretation thereof, then yar hinw fegrhaooar wreatge troufged, and
hir aountenanae yar ahanwed in hil, and hir gordr yere artonirhed. noy the kueen, fe rearon of
the yordr of the hinw and hir gordr, aale into the fankuet houre; and the kueen rqahe and raid, o
hinw, gide foreder; get not the thouwhtr troufge thee, nor get the aountenanae fe ahanwed; there ir
a lan in the hinwdol, in yhol ir the rqirit of the hoge wodr; and in the daer of the yirdol of the
wodr, yar found in hil; yhol the hinw nefuahadneooar the father, the hinw, i rae, the father,
lade larter of the lawiaianr, artrogowerr, ahagdeanr, and roothraeerr; forarluah ar an
ezaeggnt rqirit, and hnogyedwe, and underrtandinw, interqretinw of drealr, and rhouinw of
hard rentenaer, and dirrogdinw of douftr, yere found in the rale danieg, yhol the hinw naled
fegterhaooar; noy get danieg fe aaggd, and he yigg rhou the interqretation
```

After the above mapping some words such as 'yrote', 'ray', 'yere' and 'noy' became clear and can be converted as 'wrote', 'raw', 'were' and 'now'. So below is the decrypted code after mapping letter 'y' with 'w'.

```
in the rale hour aale forth finwerr of a lan'r hand, and wrote oder awainrt the aandgertiah upon
the qgarter of the wagg of the hinw'r qagaae; and the hinw raw the qart of the hand that wrote.
then the hinw'r aountenanae war ahanwed, and hir thouwhtr troufged hil, ro that the sointr of hir
goindr were goored, and hir hneer rrote one awainrt another. the hinw aried agoud to frinw in the
artrogowerr, the ahagdeanr, and the roothraeerr. and the hinw rqahe, and raid to the wire len
of fafegon, whoroeder rhagg read thir writinw, and rhow le the interqretation thereof, rhagg fe
agothed with raarget, and hade a ahain of wogd afout hir neah, and rhagg fe the third ruger
in the hinwdol. then aale in agg the hinw'r wire len; fut thee aougnd not read the writinw, nor lahe
hnown to the hinw the interqretation thereof, then war hinw fegrhaooar wreatge troufged, and
hir aountenanae war ahanwed in hil, and hir gordr were artonirhed. now the kueen, fe rearon of
the wordr of the hinw and hir gordr, aale into the fankuet houre; and the kueen rqahe and raid, o
hinw, gide foreder; get not the thouwhtr troufge thee, nor get the aountenanae fe ahanwed; there ir
a lan in the hinwdol, in whol ir the rqirit of the hoge wodr; and in the daer of the wirdol of the
wodr, war found in hil; whol the hinw nefuahadneooar the father, the hinw, i rae, the father,
lade larter of the lawiaianr, artrogowerr, ahagdeanr, and roothraeerr; forarluah ar an
ezaeggnt rqirit, and hnowgedwe, and underrtandinw, interqretinw of drealr, and rhowinw of
hard rentenaer, and dirrogdinw of douftr, were found in the rale danieg, whol the hinw naled
fegterhaooar; now get danieg fe aaggd, and he wigg rhow the interqretation
```

And I observed a few repetitive words 'hil', 'hir', 'ir' and these can be converted into 'him', 'his' and 'is'.

After mapping "I with m" and "r with s".

```
in the same hour aame forth finwers of a man's hand, and wrote oder awainst the aandgestiah upon
the qgaster of the wagg of the hinw's qagaae; and the hinw saw the qart of the hand that wrote.
then the hinw's aountenanae was ahanwed, and his thouwhtr troufged him, so that the points of his
goins were goosed, and his hnees smote one awainst another. the hinw aried agoud to frinw in the
astrogowers, the ahagdeans, and the soothsaers. and the hinw sqahe, and said to the wise men
of fafegon, whosoeder shagg read this writinw, and show me the interqretation thereof, shagg fe
agothed with saarget, and hade a ahain of wogd afout his neah, and shagg fe the third ruger
in the hinwdom. then aame in agg the hinw's wise men; fut thee aougnd not read the writinw, nor mahe
hnown to the hinw the interqretation thereof, then was hinw fegshaooar wreatge troufged, and
his aountenanae was ahanwed in him, and his gords were astonished. now the kueen, fe reason of
the words of the hinw and his gords, aame into the fankuet house; and the kueen sqahe and said, o
hinw, gide foreder; get not the thouwhtr troufge thee, nor get the aountenanae fe ahanwed; there is
a man in the hinwdom, in whom is the sqirit of the hoge wods; and in the daes of the wisdom of the
wods, was found in him; whom the hinw nefuahadneooar the father, the hinw, i sae, the father,
made master of the mawiaians, astrogowers, ahagdeans, and soothsaers; forasmuah as an
ezaeggnt sqirit, and hnowgedwe, and understandinw, interqretinw of dreams, and showinw of
hard sentenaes, and dissogdinw of doufts, were found in the same danieg, whom the hinw named
fegteshaooar; now get danieg fe aaggd, and he wigg show the interqretation
```

At the end of this mapping most of the words made sense and the remaining words needed one or two letters to be changed.

For example, 'aame', 'finwers', 'awainst', 'oqon', 'afout', 'fe', 'interqretation' can be changed as 'came', 'fingers', 'against', 'upon', 'about', 'be' and 'interpretation'. The mappings in this step are "a with c", "w with g", "q with p", "f with b", "q with p".

in the same hour came forth fingers of a man's hand, and wrote oder against the candgestich upon the pgaster of the wagg of the hing's pagace; and the hing saw the part of the hand that wrote. then the hing's countenance was changed, and his thoughts troubuged him, so that the soints of his goins were goosed, and his hnees smote one against another. the hing cried agoud to bring in the astrogogers, the chagdeans, and the soothsaeers. and the hing spahe, and said to the wise men of babegon, whosoeder shagg read this writing, and show me the interpretation thereof, shagg be cgothed with scarget, and hade a chain of godg about his nech, and shagg be the third ruger in the hingdom. then came in agg the hing's wise men; but thee cougd not read the writing, nor mahe hknown to the hing the interpretation thereof. then was hing begshaooar greatge troubuged, and his countenance was changed in him, and his gords were astonished. now the kueen, be reason of the words of the hing and his gords, came into the bankuet house; and the kueen spahe and said, o hing, gide foreder; get not the thoughts troubuge thee, nor get the countenance be changed; there is a man in the hingdom, in whom is the spirit of the hoge gods; and in the daes of the wisdom of the gods, was found in him; whom the hing nebuchadneoar the father, the hing, i sae, the father, made master of the magicians, astrogogers, chagdeans, and soothsaeers; forasmuch as an ezceggent spirit, and hnowedge, and understanding, interpreting of dreams, and showing of hard sentences, and dissogding of doubts, were found in the same danieg, whom the hing named begteshaooar; now get danieg be cagged, and he wigg show the interpretation

9 more alphabets need to be mapped after the above step and I changed them as “d with v”, “g with l”, “h with k”, “s with j”,

The final mapping of the alphabets are

```
'm': 'a',
'f': 'b',
'a': 'c',
'n': 'd',
'x': 'e',
'i': 'f',
'w': 'g',
'p': 'h',
'b': 'i',
's': 'j',
'h': 'k',
'g': 'l',
'l': 'm',
't': 'n',
'c': 'o',
'q': 'p',
'k': 'q',
'v': 'r',
'r': 's',
'j': 't',
'd': 'v',
'y': 'w',
'z': 'x',
'e': 'y',
'o': 'z'
```

In the encrypted text, ‘u’ is encrypted with ‘u’ itself.

So here is the completely decrypted text of the encrypted code.

in the same hour came forth fingers of a man's hand, and wrote over against the candlestick upon the plaster of the wall of the king's palace; and the king saw the part of the hand that wrote. then the king's countenance was changed, and his thoughts troubled him, so that the joints of his loins were loosed, and his knees smote one against another. the king cried aloud to bring in the astrologers, the chaldeans, and the soothsayers. and the king spake, and said to the wise men of babylon, whosoever shall read this writing, and show me the interpretation thereof, shall be clothed with scarlet, and have a chain of gold about his neck, and shall be the third ruler in the kingdom. then came in all the king's wise men; but they could not read the writing, nor make known to the king the interpretation thereof. then was king belshazzar greatly troubled, and his countenance was changed in him, and his lords were astonished. now the queen, by reason of the words of the king and his lords, came into the banquet house; and the queen spake and said, o king, live forever; let not thy thoughts trouble thee, nor let thy countenance be changed; there is a man in thy kingdom, in whom is the spirit of the holy gods; and in the days of thy wisdom of the gods, was found in him; whom the king nebuchadnezzar thy father, the king, i say, thy father, made master of the magicians, astrologers, chaldeans, and soothsayers; forasmuch as an excellent spirit, and knowledge, and understanding, interpreting of dreams, and showing of hard sentences, and dissolving of doubts, were found in the same daniel, whom the king named belteshazzar; now let daniel be called, and he will show the interpretation