Group 11: Nathaniel Stone, Tashfin Hassan, Bobby Kestyn, Mario Gomez, Dian Rajic

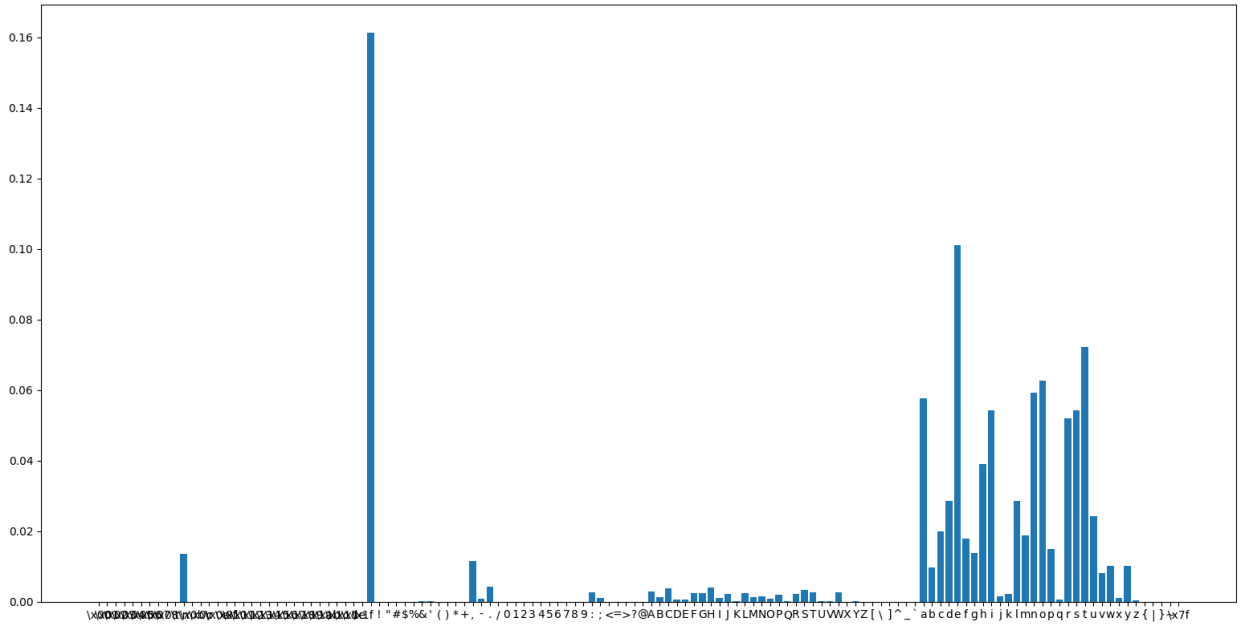
HOMEWORK 02

Part 1:

This program will create a bar chart and display a frequency list in the terminal for any text inserted. There will be two options; the first one will allow you to insert a text manually, and the second one will allow you to read a text from a file. After that, a message will appear in the terminal giving two modes; a TOTAL ascii analysis, or an alpha-numerical analysis (only letters and numbers). In the next step you will need to type the text manually or the name of the file in case you chose the other mode. Finally a plot will pop up on the screen showing the results and a list with the frequencies of every character will be printed in the terminal.

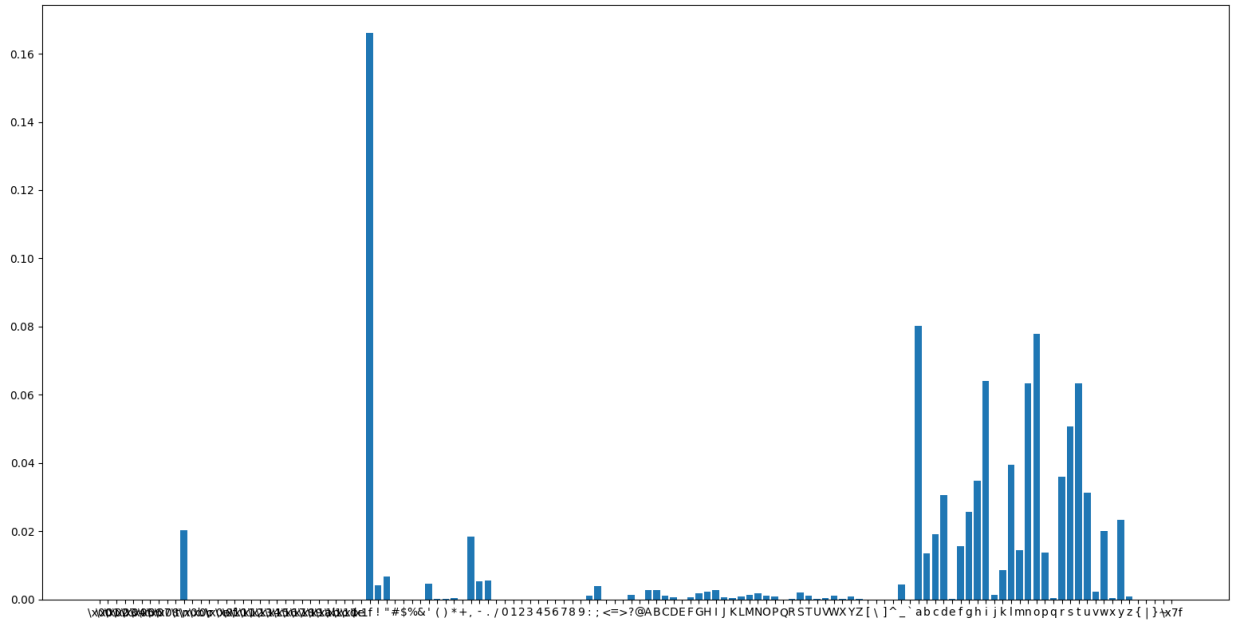
Part 2:

Doi all ascii:

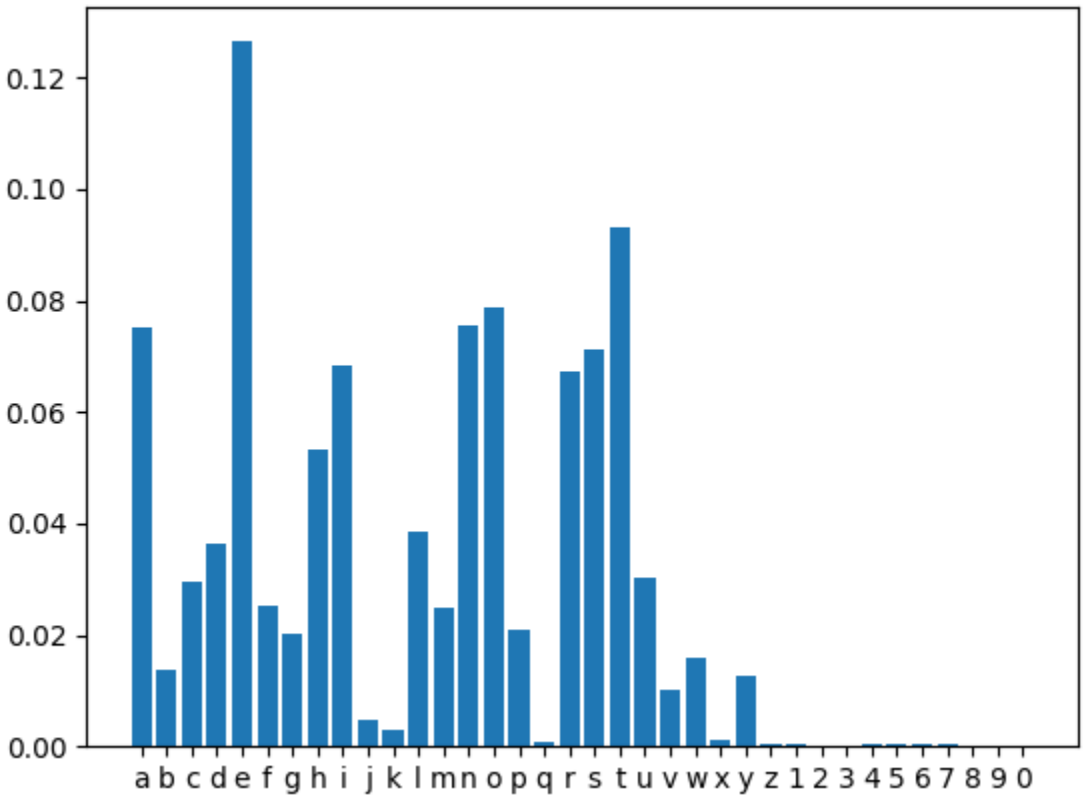


Group 11: Nathaniel Stone, Tashfin Hassan, Bobby Kestyn, Mario Gomez, Dian Rajic

Gadsby all ascii:

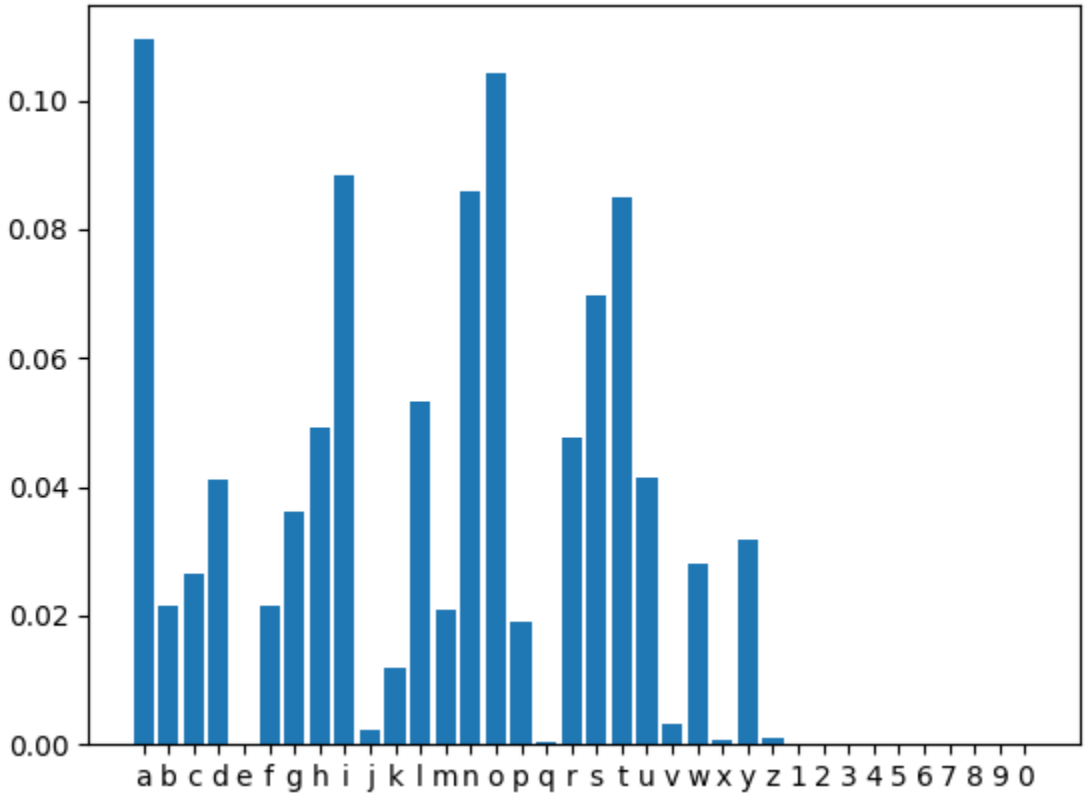


Doi alphanumeric:



Group 11: Nathaniel Stone, Tashfin Hassan, Bobby Kestyn, Mario Gomez, Dian Rajic

Gadsby alphanumeric:



Observations:
Despite the amount of sheer text in Gadsby being much greater than the Declaration of Independence, the frequency distributions of the two texts are generally the same; besides the fact that Gadsby has almost no e's.

Also, the frequency tables from all the above histograms are in the appendices at the end of the document.

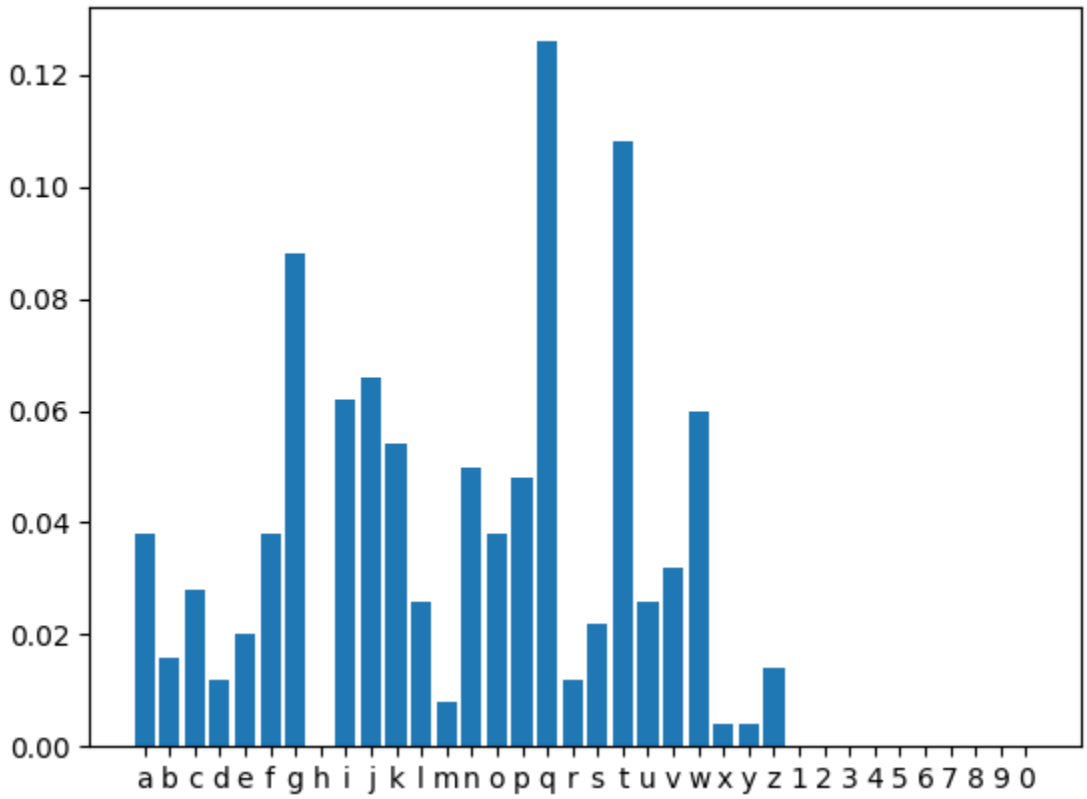
Part 3:

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
u	b	f	v j	w	l d	o	z	a	n	i	m p	k	r	l d	h	e	v j	g	t	m p	c	s	x q	x q	y

```
FTAIC TIZWN UTWPU GTSWK KJKJI UQRJI GTSFT AIKWS QJPOG VQTPK FCJFG TLQGI AKFNJ
GWNSW QOWNG TOTCV JGFTA LQKWQ BQWNU IQJGC TIZJP OGVQT PKFCJ FGTOT UIQJG CTIZW
NGTKT BQCVJ GFTAO TWSFT AVJBQ PTGST APOWG FQGZQ QRKTT ZWPUO TPTGN QGGKQ JNCWG
VJKKE JGGQI NTSGV QVQJI GFTAC WKKZP TCCVQ PFTAS WPOWG YATGQ LFNGQ BQXTL NDDDD
```

Group 11: Nathaniel Stone, Tashfin Hassan, Bobby Kestyn, Mario Gomez, Dian Rajic

```
QBQIF GVWPU GVJGE JZQNT AIOJW KFKWB WPUQJ NWQIE TIQRI TOAMG WBQET IQQPX TFJLK
QJPOE TIQRK QJNAI JLKQC JNMIQ JGQOL QMJAN QTSJR ITLKQ EJPOL QMJAN QNTEQ OQNWU
PQITI GQJET SOQNW UPQIN NTEQC VQIQT AGGVQ IQWPG VQCTI KONTA UVGGT NTKBQ GVJGR
ITLKQ EYTGQ OLFLW KKLAI PQGGT POQNW UPWPU FTAIK WSQDD
```



Cipher frequency table (alpha-numerical) *as there are only letters, we did not perform the ascii analysis:
QTGJIWKNP(AFO)VC(IU)SEBZ(RD)M(XY)H
{'a': 0.038, 'b': 0.016, 'c': 0.028, 'd': 0.012, 'e': 0.02, 'f': 0.038, 'g': 0.088, 'h': 0, 'i': 0.062, 'j': 0.066, 'k': 0.054, 'l': 0.026, 'm': 0.008, 'n': 0.05, 'o': 0.038, 'p': 0.048, 'q': 0.126, 'r': 0.012, 's': 0.022, 't': 0.108, 'u': 0.026, 'v': 0.032, 'w': 0.06, 'x': 0.004, 'y': 0.004, 'z': 0.014, '1': 0, '2': 0, '3': 0, '4': 0, '5': 0, '6': 0, '7': 0, '8': 0, '9': 0, '0': 0}

gadsby:
AOINTSLHRUDGYWCFBMPKVJZXQE
{''a: 0.109, 'b': 0.0214, 'c': 0.0266, 'd': 0.0411, 'e': 0, 'f': 0.021516, 'g': 0.03606, 'h': 0.0490, 'i': 0.0882, 'j': 0.002294, 'k': 0.01177, 'l': 0.05319, 'm': 0.020748, 'n': 0.08597, 'o': 0.1041, 'p': 0.019046, 'q': 0.000513, 'r': 0.047574, 's': 0.06966, 't': 0.08492, 'u': 0.0415, 'v': 0.0032373, 'w': 0.0279, 'x': 0.00079165, 'y': 0.0317511, 'z': 0.0010838}

Group 11: Nathaniel Stone, Tashfin Hassan, Bobby Kestyn, Mario Gomez, Dian Rajic

doi:
ETONASIRHLDUCFMPGWBYVJKXQZ
{'a': 0.07527, 'b': 0.013745, 'c': 0.02945, 'd': 0.0363, 'e': 0.12645634245320067, 'f': 0.02513, 'g': 0.0202906, 'h': 0.0534101 'i': 0.0684, 'j': 0.004581, 'k': 0.00287, 'l': 0.0384, 'm': 0.025003, 'n': 0.075533, 'o': 0.07880, 'p': 0.02081, 'q': 0.000785, 'r': 0.0671553, 's': 0.071344, 't': 0.093205, 'u': 0.0303704, 'v': 0.0102107, 'w': 0.015970, 'x': 0.0013090, 'y': 0.0126979, 'z': 0.0005236, }

Challenges:
With Gadsby being an anomaly in terms of frequency distribution due to how few e's there are, it was difficult comparing it to frequency distributions within the cipher. With the cipher being formatted to not give away even the location of spaces in the plain-text, it was particularly difficult formulating words and extrapolating potential letters out of what little we could decode.

Key:
As we could not solve the cypher, we did not actually get a valid key for this problem. But we used a tool (https://planetcalc.com/8047/) from the internet to at least try to understand the code and get the actual key:

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
u	v	w	x	m	y	t	z	r	a	l	b	c	s	d	n	e	p	f	o	g	h	i	j	q	k

And the plaintext is:

YOUR WORK IS GOING TO FILL A LARGE PART OF YOUR LIFE AND THE ONLY WAY TO BE TRULY SATISFIED IS TO DO WHAT YOU BELIEVE IS GREAT WORK AND THE ONLY WAY TO DO GREAT WORK IS TO LOVE WHAT YOU DO IF YOU HAVE NOT FOUND IT YET KEEP LOOKING DO NOT SETTLE AS WITH ALL MATTERS OF THE HEART YOU WILL KNOW WHEN YOU FIND IT QUOTE BY STEVE JOBS XXXX EVERYTHING THAT MAKES OUR DAILY LIVING EASIER MORE PRODUCTIVE MORE ENJOYABLE AND MORE PLEASURABLE WAS CREATED BECAUSE OF A PROBLEM AND BECAUSE SOME DESIGNER OR TEAM OF DESIGNERS SOMEWHERE OUT THERE IN THE WORLD SOUGHT TO SOLVE THAT PROBLEM QOTED BY BILL BURNETT ON DESIGNING YOUR LIFE XX

APPENDIX 1 Frequency table for ASCII analysis doi.txt:

{'\\x00': 0, '\\x01': 0, '\\x02': 0, '\\x03': 0, '\\x04': 0, '\\x05': 0, '\\x06': 0, '\\x07': 0, '\\x08': 0, '\\t': 0, '\\n': 0.0134850400337126, '\\x0b': 0, '\\x0c': 0, '\\r': 0, '\\x0e': 0, '\\x0f': 0, '\\x10': 0, '\\x11': 0, '\\x12': 0, '\\x13': 0, '\\x14': 0, '\\x15': 0, '\\x16': 0, '\\x17': 0, '\\x18': 0, '\\x19': 0, '\\x1a': 0, '\\x1b': 0, '\\x1c': 0, '\\x1d': 0, '\\x1e': 0, '\\x1f': 0,

' ': 0.1611883691529709, '!': 0, '"': 0, '#': 0, '$': 0, '%': 0, '&': 0.00010535187526337969, "'": 0.00010535187526337969, '(': 0, ')': 0, '*': 0, '+': 0, ',': 0.011588706278971767, '-': 0.0007374631268436578, '.': 0.004214075010535188, '/': 0,

0: 0, 1: 0, 2: 0, 3: 0, 4: 0, 5: 0, 6: 0, 7: 0, 8: 0, 9: 0,

':': 0.002739148756847872, ';': 0.001053518752633797, '<': 0, '=': 0, '>': 0, '?': 0, '@': 0,

'A': 0.0028445006321112516, 'B': 0.001369574378423936, 'C': 0.0037926675094816687, 'D': 0.0006321112515802782, 'E': 0.0006321112515802782, 'F': 0.002317741255794353, 'G': 0.0025284450063211127, 'H': 0.0038980193847450483, 'I': 0.0009481668773704172, 'J': 0.0022123893805309734, 'K': 0.00021070375052675939, 'L': 0.0025284450063211127, 'M': 0.001369574378423936, 'N': 0.0014749262536873156, 'O': 0.0007374631268436578, 'P': 0.0018963337547408343, 'Q': 0.00010535187526337969, 'R': 0.002107037505267594, 'S': 0.0032659081331647705, 'T': 0.002739148756847872, 'U': 0.00010535187526337969, 'V': 0.00010535187526337969, 'W': 0.002739148756847872, 'X': 0, 'Y': 0.00010535187526337969, 'Z':

0, '[': 0, '\\': 0, ']': 0, '^': 0, '_': 0, '`': 0,

'a': 0.05773282764433207, 'b': 0.00969237252423093, 'c': 0.01991150442477876, 'd': 0.028655710071639275, 'e': 0.1011378002528445, 'f': 0.017909818794774546, 'g': 0.01380109565950274, 'h': 0.03908554572271387, 'i': 0.05415086388537716, 'j': 0.0014749262536873156, 'k': 0.002107037505267594, 'l': 0.028445006321112517, 'm': 0.018752633796881586, 'n': 0.05931310577328276, 'o': 0.06268436578171091, 'p': 0.014854614412136535, 'q': 0.0005267593763168984, 'r': 0.05193847450484618 4, 's': 0.05415086388537716, 't': 0.07227138643067847, 'u': 0.024336283185840708, 'v': 0.008112094395280236, 'w': 0.010113780025284451, 'x': 0.001053518752633797, 'y': 0.010113780025284451, 'z': 0.0004214075010535187 7,

'{': 0, '|': 0, '}': 0, '~': 0, '\\x7f': 0}

APPENDIX 2 Frequency table for ASCII analysis gadsby_text.txt:

{'\\x00': 0, '\\x01': 0, '\\x02': 0, '\\x03': 0, '\\x04': 0, '\\x05': 0, '\\x06': 0, '\\x07': 0, '\\x08': 0, '\\t': 0, '\\n': 0.020333548748745888, '\\x0b': 0, '\\x0c': 0, '\\r': 0, '\\x0e': 0, '\\x0f': 0, '\\x10': 0, '\\x11': 0, '\\x12': 0, '\\x13': 0, '\\x14': 0, '\\x15': 0, '\\x16': 0, '\\x17': 0, '\\x18': 0, '\\x19': 0, '\\x1a': 0, '\\x1b': 0, '\\x1c': 0, '\\x1d': 0, '\\x1e': 0, '\\x1f': 0,
' ': 0.1659888170837514, '!': 0.004238089967473463, '"': 0.006651694700423809, '#': 0, '$': 0, '%': 0, '&': 0, "'": 0.00454514619089614, '(': 0.0001999435873449991, ')': 0.0001999435873449991, '*': 0.0003391900142459806, '+': 0, ',': 0.01851977477783052, '-': 0.005387765594707207, '.': 0.0055805683396470285, '/': 0,

0: 0, 1: 0, 2: 0, 3: 0, 4: 0, 5: 0, 6: 0, 7: 0, 8: 0, 9: 0,
':': 0.0009854362519146384, ';': 0.0038774774260119465, '<': 0, '=': 0, '>': 0, '?': 0.001206802366475173, '@': 0,

'A': 0.0026849567443471304, 'B': 0.002756365168398916, 'C': 0.001099689730397495, 'D': 0.0006640983436816041, 'E': 1.071126360776781e-05, 'F': 0.0007176546617204432, 'G': 0.0016745275440143673, 'H': 0.0023636188361140964, 'I': 0.002809921486437755, 'J': 0.0005069998107676763, 'K': 0.00034990127785374843, 'L': 0.0008247672977981212, 'M': 0.0012210840512855302, 'N': 0.0017994922861049919, 'O': 0.0011175418364104412, 'P': 0.0007640701373541036, 'Q': 3.5704212025892696e-06, 'R': 0.0001678079652169565, 'S': 0.0020315696642732944, 'T': 0.001060415097169013, 'U': 0.0001106830572802736, 'V': 0.0003034858022200879, 'W': 0.001046133412358656, 'X': 0.00024278864177607032, 'Y': 0.0007497884525437465, 'Z': 6.426758164660686e-05,

'[': 0, '\\': 0, ']': 0, '^': 0, '_': 0.004255942073486409, '`': 0,

'a': 0.08027020947661195, 'b': 0.013478340039774493, 'c': 0.019076760485434465, 'd': 0.030519960439733075, 'e': 3.5704212025892695e-05, 'f': 0.015584888549302161, 'g': 0.025649905919401313, 'h': 0.03482588841005573, 'i': 0.0640676380592618 6, 'j': 0.001231795314893298, 'k': 0.008572581307416836, 'l': 0.03948171765823214, 'm': 0.014499480503715023, 'n': 0.06334284255513623, 'o': 0.07780661884682535, 'p': 0.013667572363511724, 'q': 0.0003856054898796411, 'r': 0.03587916266481957, 's': 0.05075353739480647, 't': 0.06328928623709738, 'u': 0.03138757279196227, 'v': 0.0021493935639587403, 'w': 0.020151457267413835, 'x': 0.0003570421202589269 6, 'y': 0.023307709610502752, 'z': 0.0007569292949489251,

'{': 0, '|': 0, '}': 0, '~': 0, '\\x7f': 0}

APPENDIX 3 Frequency table for alpha-numerical analysis doi.txt:

{'a': 0.07527163241261946, 'b': 0.013745254614478335, 'c': 0.029454117031025004, 'd': 0.036392197931666447, 'e': 0.12645634245320067, 'f': 0.02513417986647467, 'g': 0.020290613954706112, 'h': 0.05341013221625867, 'i': 0.06846445869878257, 'j': 0.004581751538159445, 'k': 0.0028799581097002226, 'l': 0.03848671292053934, 'm': 0.025003272679670115, 'n': 0.07553344678622856, 'o': 0.07880612645634245, 'p': 0.020814242701924335, 'q': 0.0007854431208273334, 'r': 0.067155386830737, 's': 0.07134441680848279, 't': 0.09320591700484357, 'u': 0.030370467338656892, 'v': 0.010210760570755335, 'w': 0.01597067679015578, 'x': 0.0013090718680455558, 'y': 0.01269799712004189, 'z': 0.0005236287472182222,

'1': 0.0002618143736091111, '2': 0.00013090718680455556, '3': 0.00013090718680455556, '4': 0.0002618143736091111, '5': 0.0002618143736091111, '6': 0.0003927215604136667, '7': 0.0002618143736091111, '8': 0, '9': 0, '0': 0}

APPENDIX 4 Frequency table for alpha-numerical analysis gadsby_text.txt:

{'a': 0.10948433884823267, 'b': 0.021426585553194197, 'c': 0.026628905863448515, 'd': 0.04115676230956633, 'e': 6.125920655190776e-05, 'f': 0.021516118239693138, 'g': 0.03606282367244231, 'h': 0.049082761188051625, 'i': 0.08826509214798339, 'j': 0.002294864122367621, 'k': 0.011775904397939806, 'l': 0.05319655252034512, 'm': 0.020748022034465373, 'n': 0.0859749402722736, 'o': 0.10416421237153238, 'p': 0.019046900990985473, 'q': 0.0005136348857044573, 'r': 0.04757484225754313, 's': 0.06966585458949263, 't': 0.08492882151423334, 'u': 0.04157144001545617, 'v': 0.003237313453935433, 'w': 0.027976608407590486, 'x': 0.0007916574385169618, 'y': 0.03175111798051957, 'z': 0.0010838167313029833,

'1': 4.712246657839058e-06, '2': 0, '3': 4.712246657839058e-06, '4': 0, '5': 0, '6': 0, '7': 0, '8': 0, '9': 9.424493315678116e-06, '0': 0}