Звіт з лабораторної роботи №1 З курсу «Комп'ютерні системи» Студентки 3 курсу КІ-СА Уваренко Олени

1) Оберіть 3 текстових файла різного тематичного та лінгвістичного спрямування (наприклад, вірш Тараса Шевченка "Мені тринадцятий минало", "Казка про рєпку" Леся Подерв'янського та специфікацію інерфейсу РСІ)

Мої файли:

- i) Стаття «Події поза планами кандидатів у президенти: Зеленський, Порошенко»
- іі) «Убога мова в нас» А.Фета
- ііі) Класифікація штучних клапанів для серця





- 2) Створіть програму (будь-якою зручною для вас мовою), яка в якості вхідних даних приймає текстовий файл, та аналізуючи його вміст:
 - і) обраховує частоти (імовірності) появи символів в тексті
 - іі) обраховує середню ентропію алфавіту для даного тексту
 - ііі) виходячи з ентропії визначає кількість інформації та порівнює її з розмірами файлів
 - іv) виводить на екран значення частот, ентропії та кількості інформації

```
File to analize: fet.txt
Total summ of sumbols in file: 355
Relative frequency of appearing letter "a" in the text = 0.0563380281690141; Concrete letter present in the text: 20 times. Relative frequency of appearing letter "6" in the text = 0.0112676056338028; Concrete letter present in the text: 4 times. Relative frequency of appearing letter "a" in the text = 0.0535211267605634; Concrete letter present in the text: 19 times. Relative frequency of appearing letter "r" in the text = 0.0169014084507042; Concrete letter present in the text: 6 times.
Relative frequency of appearing letter "r"
                                                              in the text = 0; Concrete letter present in the text: 0 times.
Relative frequency of appearing letter "A" in the text = 0.0647887323943662; Concrete letter present in the text: 8 times. Relative frequency of appearing letter "e" in the text = 0.0647887323943662; Concrete letter present in the text: 23 times.
 Relative frequency of appearing letter "e" in the text
                                                                                   0.00563380281690141 ; Concrete letter present in the text: 2 times.
 Relative frequency of appearing letter "x" in the text
                                                                                  0.0140845070422535; Concrete letter present in the text: 5 times. 0.0225352112676056; Concrete letter present in the text: 8 times.
                                              letter "3" in the text =
 Relative frequency of appearing
Relative frequency of appearing letter 'β' in the text = 0.0929577464788732 ; Concrete letter present in the text: 33 times. Relative frequency of appearing letter "i" in the text = 0.0507042253521127 ; Concrete letter present in the text: 18 times.
 Relative frequency of appearing letter "i" in the text
                                                                                = 0; Concrete letter present in the text: 0 times
 Relative frequency of appearing letter "й" in the text
                                                                                   0.0112676056338028 ; Concrete letter present in the text: 4 times.
                                              letter "k" in the text =
 Relative frequency of appearing
                                                                                   0.0197183098591549; Concrete letter present in the text: 7 times.
 Relative frequency of appearing letter "n" in the text = 0.0394366197183099; Concrete letter present in the text: 14 times.
Relative frequency of appearing letter "m" in the text = 0.0450704225352113 ; Concrete letter present in the text: 16 times.
 Relative frequency of appearing letter "H" in the text
                                                                                = 0.0619718309859155 ; Concrete letter present in the text: 22 times.
 Relative frequency of appearing letter "o" in the text
                                                                                   0.0704225352112676; Concrete letter present in the text: 25 times.
Relative frequency of appearing letter "n" in the text = Relative frequency of appearing letter "p" in the text = Relative frequency of appearing letter "c" in the text =
                                                                                   0.0366197183098592 ; Concrete letter present in the text:
                                                                                                                                                                       13 times.
                                                                                   0.047887323943662 ; Concrete letter present in the text: 17 times.
Relative frequency of appearing letter "c" in the text = 0.0422535211267606; Concrete letter present in the text: 15 times. Relative frequency of appearing letter "\tau" in the text = 0.0507042253521127; Concrete letter present in the text: 18 times. Relative frequency of appearing letter "y" in the text = 0.028169014084507; Concrete letter present in the text: 10 times.
Relative frequency of appearing letter "o" in the text Relative frequency of appearing letter "x" in the text
                                                                                  0; Concrete letter present in the text: 0 times
                                                                                   0.0225352112676056 ; Concrete letter present in the text: 8 times.
Relative frequency of appearing letter "u" in the text = 0.00845070422535211; Concrete letter present in the text: 3 times.
 Relative frequency of appearing letter "4" in the text = 0.00845070422535211; Concrete letter present in the text: 3 times.
Relative frequency of appearing letter "w" in the text = 0.0140845070422535; Concrete letter present in the text: 5 times. Relative frequency of appearing letter "w" in the text = 0.0128169014084507; Concrete letter present in the text: 1 times.
Relative frequency of appearing letter "w" in the text Relative frequency of appearing letter "b" in the text
                                                                                   0.0197183098591549 ; Concrete letter present in the text: 7 times.
Relative frequency of appearing letter "w" in the text = 0.005/3380281690141; Concrete letter present in the text: 2 times.
Relative frequency of appearing letter "w" in the text = 0.0169014084507042; Concrete letter present in the text: 6 times.
Average entropy of the unequipotent alphabet in the text: 4.43257982826731
 mount of information in the text: 196.695729879362
 ile to analize: heart.txt
Relative frequency of appearing letter "r"
                                                              in the text = 0.0145152928978745; Concrete letter present in the text: 28 times.
```

```
Total summ of sumbols in file: 1929
Relative frequency of appearing letter "a" in the text = 0.0762052877138414 ; Concrete letter present in the text: 147 times.
Relative frequency of appearing letter "6" in the text = 0.00933125972006221 ; Concrete letter present in the text: 18 times.
Relative frequency of appearing letter "B" in the text = 0.046656298600311 ; Concrete letter present in the text: 90 times.
Relative frequency of appearing letter "r" in the text = 0.0145152928978745; Concrete letter present in the text: 28 times. Relative frequency of appearing letter "r" in the text = 0; Concrete letter present in the text: 0 times. Relative frequency of appearing letter "A" in the text = 0.0207361327112494; Concrete letter present in the text: 40 times. Relative frequency of appearing letter "e" in the text = 0.0580611715914982; Concrete letter present in the text: 112 times. Relative frequency of appearing letter "e" in the text = 0.00881285640228098; Concrete letter present in the text: 17 times.
 elative frequency of appearing letter "x" in the text = 0.0046656298600311; Concrete letter present in the text: 9 times.
Relative frequency of appearing letter "3" in the text =
                                                                                      {\tt 0.0217729393468118} ; Concrete letter present in the text: 42 times.
 Relative frequency of appearing letter "и"
                                                                                      0.0585795749092794; Concrete letter present in the text: 113 times. 0.0570243649559357; Concrete letter present in the text: 110 times.
                                                                 in the text =
Relative frequency of appearing letter "i" in the text =
 delative frequency of appearing letter "i" in the text
                                                                                      0.00673924313115604; Concrete letter present in the text: 13 times. 0.00829445308449974; Concrete letter present in the text: 16 times.
 elative frequency of appearing letter "й" in the text =
 Relative frequency of appearing letter "κ" in the text =
                                                                                      0.0451010886469673; Concrete letter present in the text: 87 times.
Relative frequency of appearing letter "л"
                                                                 in the text =
                                                                                      0.0368066355624676 ; Concrete letter present in the text: 71 times.
Relative frequency of appearing letter "m" in the text =
                                                                                      0.0285121824779679; Concrete letter present in the text: 55 times.
 Relative frequency of appearing letter "н" in the text
                                                                                       0.0787973043027475; Concrete letter present in the text: 152 times.
 Relative frequency of appearing letter "o" in the text = Relative frequency of appearing letter "n" in the text =
                                                                                      0.0855365474339036 ;
                                                                                                                       Concrete letter present in the text: 165 times.
                                                                                      0.0393986521513738; Concrete letter present in the text: 76 times.
Relative frequency of appearing letter "p" in the text = 0.055469155002592; Relative frequency of appearing letter "c" in the text = 0.0378434421980301;
                                                                                                                     Concrete letter present in the text: 107 times.
                                                                                      0.0378434421980301; Concrete letter present in the text: 73 times.
 Relative frequency of appearing letter "T"
                                                                 in the text
                                                                                       0.0502851218247797 ; Concrete letter present in the text: 97 times.
 Relative frequency of appearing letter "y" in the text = Relative frequency of appearing letter "\p" in the text =
                                                                                      0.0305857957490928;
                                                                                                                       Concrete letter present in the text: 59 times.
                                                                                      0.0031104199066874;
                                                                                                                       Concrete letter present in the text: 6 times.
Relative frequency of appearing letter "x" in the text = 0.0139968895800933 ;
                                                                                                                       Concrete letter present in the text: 27 times.
Relative frequency of appearing letter "q" in the text = Relative frequency of appearing letter "q" in the text =
                                                                                      0.0202177293934681;
                                                                                                                       Concrete letter present in the text: 39 times.
                                                                                      0.0171073094867807; Concrete letter present in the text: 33 times.
 kelative frequency of appearing letter "w" in the text = 0.00622083981337481;
Relative frequency of appearing letter "w" in the text = 0.00259201658890617;
Relative frequency of appearing letter "b" in the text = 0.0134784862623121;
                                                                                       0.00622083981337481; Concrete letter present in the text: 12 times.
                                                                                                                     ; Concrete letter present in the text: 5 times.
                                                                                                                       Concrete letter present in the text: 26 times.
Relative frequency of appearing letter "w" in the text = 0.013478402023121; Concrete letter present in the text: 20 times. Relative frequency of appearing letter "w" in the text = 0.013478402623121; Concrete letter present in the text: 26 times. Relative frequency of appearing letter "a" in the text = 0.0212545360290306; Concrete letter present in the text: 41 times.
 werage entropy of the unequipotent alphabet in the text: 4.55328529664514
 mount of information in the text: 1097.91091715356
```

```
ile to analize: zelen.txt
Total summ of sumbols in file: 2221
Relative frequency of appearing letter "a" in the text = 0.0805943268797839 ; Concrete letter present in the text: 179 times.
Relative frequency of appearing letter "6" in the text = 0.015959457817965; Concrete letter present in the text: 37 times. Relative frequency of appearing letter "6" in the text = 0.0495272399819901; Concrete letter present in the text: 110 times. Relative frequency of appearing letter "r" in the text = 0.00945520036019811; Concrete letter present in the text: 21 times.
Relative frequency of appearing letter "r" in the text = 0; Concrete letter present in the text: 0 times.

Relative frequency of appearing letter "r" in the text = 0.3036217919855921; Concrete letter present in the text: 88 times.

Relative frequency of appearing letter "g" in the text = 0.036217919855921; Concrete letter present in the text: 103 times.
Relative frequency of appearing letter "A" in the text = 0 , Concrete letter present in the text: 0 times. Relative frequency of appearing letter "A" in the text = 0.039621791855921; Concrete letter present in the text: 88 times. Relative frequency of appearing letter "e" in the text = 0.00315173345339937; Concrete letter present in the text: 103 times. Relative frequency of appearing letter "e" in the text = 0.00315173345339937; Concrete letter present in the text: 7 times.
 lelative frequency of appearing letter "x" in the text = 0.00945520036019811; Concrete letter present in the text: 21 times.
Relative frequency of appearing letter "3"
Relative frequency of appearing letter ""
                                                                                                 in the text = 0.0162089149031968 ; Concrete letter present in the text: 36 times.
                                                                                                 in the text = 0.0675371454299865; Concrete letter present in the text: 150 times.
 Relative frequency of appearing letter "i" in the text = 0.00/33/14259803 ; Concrete letter present in the text: 130 times. Relative frequency of appearing letter "i" in the text = 0.00495272399819901 ; Concrete letter present in the text: 11 times.
Relative frequency of appearing letter "\kappa" in the text = 0.0135074290859973; Concrete letter present in the text: 30 times. Relative frequency of appearing letter "\kappa" in the text = 0.0409725348941918; Concrete letter present in the text: 91 times. Relative frequency of appearing letter "\kappa" in the text = 0.040072039621792. Concrete letter present in the text = 0.040072039621792.
 Relative frequency of appearing letter "й"
 Relative frequency of appearing letter "m" in the text = 0.0292660963529941; Concrete letter present in the text: 65 times.
Relative frequency of appearing letter "m" in the text = 0.0630346690679874; Concrete letter present in the text: 140 times.
Relative frequency of appearing letter "o" in the text = 0.0855470508779829; Concrete letter present in the text: 190 times Relative frequency of appearing letter "n" in the text = 0.0238631247185952; Concrete letter present in the text: 53 times. Relative frequency of appearing letter "p" in the text = 0.045024763619991; Concrete letter present in the text: 100 times. Relative frequency of appearing letter "c" in the text = 0.0328680774425934; Concrete letter present in the text: 72 times. Relative frequency of appearing letter "c" in the text = 0.0328680774425934; Concrete letter present in the text: 72 times.
                                                                                                                                                                                Concrete letter present in the text: 190 times.
                                                                                                in the text = 0.0540297163439892; Concrete letter present in the text: 120 times. in the text = 0.0315173345339937; Concrete letter present in the text: 70 times.
 Relative frequency of appearing letter "T" in the text =
Relative frequency of appearing letter "f" in the text = 0.054029/15439892; Concrete letter present in the text: 120 times. Relative frequency of appearing letter "y" in the text = 0.0015173345339937; Concrete letter present in the text: 70 times. Relative frequency of appearing letter "y" in the text = 0.00720396217919856; Concrete letter present in the text: 16 times. Relative frequency of appearing letter "x" in the text = 0.00810445745159838; Concrete letter present in the text: 18 times.
Relative frequency of appearing letter "u" in the text = 0.000104437139082; Concrete letter present in the text: 10 times. Relative frequency of appearing letter "u" in the text = 0.00010445745159838; Concrete letter present in the text: 10 times.
Relative frequency of appearing letter "u" in the text = 0.0103556956325979; Concrete letter present in the text: 23 times.
Relative frequency of appearing letter "u" in the text = 0.00720396217919856; Concrete letter present in the text: 16 times.
Relative frequency of appearing letter "b" in the text = 0.0193606483565961; Concrete letter present in the text.
 Relative frequency of appearing letter "m" in the text = 0.007700396217919856; Concrete letter present in the text: 45 times. Relative frequency of appearing letter "m" in the text = 0.0202611436289959; Concrete letter present in the text: 45 times.
 werage entropy of the unequipotent alphabet in the text: 4.47647389634111
   mount of information in the text: 1242.7810654717
```

- 3) Проведіть стиснення кожного вхідного файлу за допомогою 5 різних алгоритмів стиснення (zip, rar, gzip, bzip2, xz, або будь-які інші на ваш вибір, можна використовувати готові програмні засоби для стиснення).
- 4) Порівняйте результуючі обсяги архівів з обчисленою кількістю інформації та наведіть у звіті висновки щодо кореляції цих величин для обраних вами файлів (яка відмінність, що вийшло більше і чому)

Кількість інформації - мінімальна кількість біт, необхідна для кодування всіх можливих значень повідомлення. Вона буде меншою за розмір наший файлів.

3 розповсюджених визначень, можна навести таке:

Кількість інформаціі можна розгядати як міру зменшення невизначенності знання

При отриманні інформаційних поідомень.

Файл	fet.txt, bytes	heart.txt, bytes	zelen.txt, bytes
rar	524	1,527	1,994
tar	2,560	6,144	6,656

bz2	383	1,201	1,601
zip	577	1,579	2,051
XZ	508	1,452	1,880
ORIGINAL	817	4,222	4,982
Кількість інформації	196	1097	1242



Розміри архівованих файлів у трохи більші за кількість інформації. Однак, коли ми прицюємо з великим текстом, помітне сильніше стиснення.

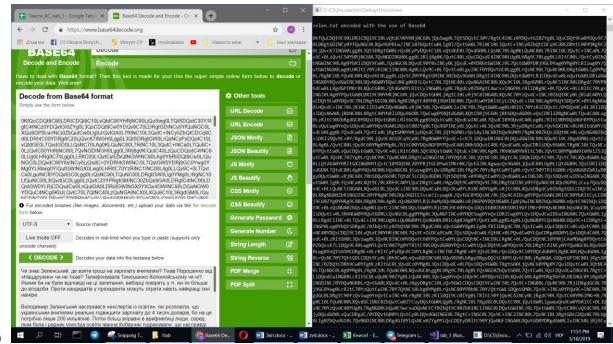
2. Дослідження способів кодування інформації на прикладі Base64

- 1) Ознайомтесь зі стандартом RFC4648
- 2) Для практичного засвоєння методу кодування, створіть програму, що кодує довільний файл в Base64 (шляхом реалізації алгоритму вручну, а не виклику бібліотечної функції) перевірте коректність роботи програми, порівнявши результат з існуючими програмними засобами (наприклад, openssl enc -base64)

1) Закодуйте в Base64 обрані вами текстові файли Texcriv_KC_na6_1 - Google Ta6::: X

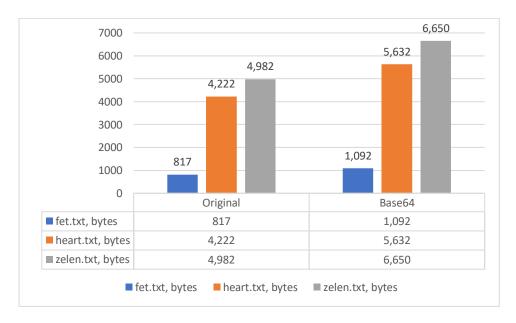
Base64 Decode and Encode - On: X ☆ ⁽¹⁾ : ← → C · https://www.base64decode.org Have to deal with Base64 format? Then this tool is made for you! Use the super simple online form below to decode or encode your data. Welcome! Decode from Base64 format 0KPQsdC+0LPQsCDQvNC+0LLQsCDQsiDQvdCw0YENINCi0L7RgiGW0YLQuCDQ uSDQvdC1NC80L7Qs9GC0LgsiOKAA0KQUSRiDQnNGA0YPQnSGPQLwsiNC902Y g0VRgpG20LjQxCDvvdC1NCi0LXRghC10LdSnC2DLXRGDC1DLGsiC2DlXrGDCvdC+0L PQvwg0YnCvDvgDcSqSG9A0YPQmSGPUYQbLCaCQSiNC90LXVX+0-L1gQuRg NC+0LQvdGA0ZYg0YX3dcXdLxRsiNciCgNC80LDRghC90LAg0LRgdGPNC2 0LDqSSGw1NC90LXdSiCQL7Qsyc10L3QuNGFRG9DLRghC90LAg0LRgdGPNC2 0LDqSSGw1NC90LXdSiCQL7Qsyc10L3QuNGFRG9DLRghC90LAg0LRghC9PNC2 0ALXRsiNgblCpdCaCquSGVC9CPC2gVC4PNC90L3QvdCnC3DxiSCDNC9PQRQ 6ALXRsiNgblCpdCaCquSGVC9CPC3CQPC4PNC90L3QvdCoV2gbSMDC4CCPCPC4 40LRqsiNcgNcriCquSGVC9CPC3CPC4PNC90L3QvdCyCl2DdSMDC4CCPCPC4 40LRqsiNcgNcriCquSGVC9CPC4PNC9CPC3PNC9CPC4PNC9CPC3CPSCPC4 40LRqsiNcgNcriCquSGVC9CPC4PNC9CPC4PNC9CPC4PNC9CPC4PNC9CPC4 40LRqsiNcgNcriCquSGVC9CPC4PN For encoded binaries (like images, documents, etc.) upload your data via the file deci-UTF-8 • Source charset. Live mode OFF Decodes in real-time when you type or paste (supports only ✓ DECODE ➤ Decodes your data into the textarea below. Убога мова в нас! Хотти й не могти. — Ні друзям, ні чужим не перекажеш ти Того, що в грудях б'є, немов прозорі хвилі. І марна вся жага невтолених сердець, І перед венною изманю мудрець Схиляє сивину в несилі. Тексти_КС_ваб_1 - Google Ta6.n X Ваse64 Decode and Encode - On X + ☆ 🕕 : атки 📳 (1) Oksana Boriych... 🦻 Интуит С# 🚪 modnakasta 🔼 🥛 Намиста мені 🦠 Decode from Base64 format • For encoded binaries (like images, documents, etc.) upload your data via the file de UTF-8 • Source charset. Live mode OFF Decodes in real-time when you type or paste (supports only **〈 DECODE 〉** Decodes your data into the textarea below. При захворюванні або дисфункції, що виникла внаслідок будь якої вади розвитку одного з чотирьох клаганів серця з метою відновлення його працедатности може бути проведена заміна приорацного клаганам на його протез. Як правилю, це потребує оперативного втручання на відкритому серці.

P ☐ □ ■ Snipping T. I Ilab □ ■ Snipping T. II Ila



Правильність кодування перевірено онлайн. Тексти співпадають.

- i) Обрахуйте кількість інформації в base64-закодованому варіанті файлу
- іі) Порівняйте отримане значення з кількістю інформації вихідного файлу
- ііі) Зробіть висновки з отриманого результату



Бачимо, що кількість інформації у закодованих файлах зросла приблизно на третину (особливості алгоритму base64).

1. Закодуйте в Base64 стиснені кращим з алгоритмів текстові файли 1.Обрахуйте кількість інформації в base64-закодованому варіанті стисненого файлу 2. Порівняйте отримане значення з кількістю інформації вихідного файлу та base64-закодованого файлу

3. Зробіть висновки з отриманого результату 2500 2,204 2000 1,656 1,601 1500 1,201 1000 528 383 500 bzip2 bzip2base64 fet.txt, bytes 383 528 ■ heart.txt, bytes 1,201 1,656 1,601 2,204 ■ zelen.txt, bytes

Аналогічно до попереднього завдання, розмір файлів зріс на третину.

■ zelen.txt, bytes

■ heart.txt, bytes

Посилання на код програм:

https://github.com/helenuvarenko/Computer-Systems/tree/master/1lab

Висновки до лабораторної роботи:

fet.txt, bytes

Під час виконання завдань лабораторної роботи №1 було здобуто навички із аналізування ефективності застосування різних типів архіваторів, застосовано знання для обрахунку ентропії та кодування за різними стандартами.