

Лабораторна робота №1, Обробка та аналіз текстових даних на Python, Варіант 14

Виконав: студент групи ІП-11, Лошак Віктор Іванович

Перевірів: Юлія Тимофєєва Сергіївна

Тема роботи: Робота з текстовими даними в Python

Мета роботи: Ознайомитись з представленням тексту Python в та регулярними виразами

15.03.2024

Створіть програму, яка:

1. Читає текстовий файл як рядок. За допомогою фрагментів: виділіть частину тексту в окрему рядкову змінну та використовуйте функції та методи роботи з рядками, які описані в теоретичних відомостях.
2. За допомогою регулярних виразів виконайте завдання за варіантом:
Варіант 14. Знайти всі номери будинків і замінити цифри після першої цифри на будь-який символ.

Create a program that:

1. Reads the text file as a string. By using slices: select a part of the text into a separate string variable and use functions and methods for working with strings that are described in the theoretical information.
2. Use regular expressions to perform tasks according to variant:
Variant 14. Find all house numbers and replace digits after the first digit with any symbol.

```
In [24]: with open('text3.txt', 'r', encoding='utf-8') as file:
          text = file.read()

          print(text)
```

Slowly and solemnly he was borne into Briony LODge 145a and laid out in the principal room, while I (094)-44-33 still observed the proceedings from my post by the window 12/15/1892. The lamps had been lit, but the blinds had not been drawn, so that I could see Holmes sh_holmes221@mail.uk as he lay upon the couch h.... I do not know whether he was seized with compunction at that moment for the part he was playing, but I know that I never felt more heartily ashamed of myself in my life 34-34-32 than when I saw the beautiful creature against whom I wston23@mymail.com was conspiring, or the grace and kindness with which she w aited upon the injured man. And yet it would be the blackest treachery to 1895-11-17 Holmes to draw back now from the part which he had intrusted to me AveNUe 123. I hardened my heart Wtason Street 45b, and took the smoke-rocket from under my ul ster. After all,, 344-21-01 I thought, we are not injuring her. We are but preventing her from injuring another adlerrr@mail.com.

Task 1

```
In [5]: text[0]
```

```
Out[5]: 'S'
```

```
In [6]: text[0:100]
```

```
Out[6]: 'Slowly and solemnly he was borne into Briony LODgE 145a and laid out in the\np  
rincipal room, while I '
```

```
In [7]: text.count('t')
```

```
Out[7]: 66
```

```
In [8]: text.find("s")
```

```
Out[8]: 11
```

```
In [9]: text.index('laid')
```

```
Out[9]: 60
```

```
In [10]: text.upper()
```

```
Out[10]: 'SLOWLY AND SOLEMNLY HE WAS BORNE INTO BRIONY LODGE 145A AND LAID OUT IN THE\nP  
RINCIPAL ROOM, WHILE I (094)-44-33 STILL OBSERVED THE PROCEEDINGS FROM MY POST  
\nBY THE WINDOW 12/15/1892.    THE LAMPS HAD BEEN LIT, BUT THE BLINDS HAD NOT B  
EEN\nDRAWN, SO THAT I COULD SEE HOLMES SH_HOLMES221@MAIL.UK AS HE LAY UPON THE  
COUCH.... I DO NOT\nKNOW WHETHER HE WAS SEIZED WITH COMPUNCTION AT THAT MOMENT  
FOR THE PART\nHE WAS PLAYING, BUT I KNOW THAT I NEVER FELT MORE HEARTILY ASHAME  
D OF\nMYSELF IN MY LIFE 34-34-32 THAN WHEN I SAW THE BEAUTIFUL CREATURE AGAINST  
WHOM\nI WSTON23@MYMAIL.COM WAS CONSPIRING, OR THE GRACE AND KINDLINESS WITH WHI  
CH SHE WAITED\nUPON THE INJURED MAN. AND YET    IT WOULD BE THE BLACKEST TREACH  
ERY TO 1895-11-17\nHOLMES TO DRAW BACK NOW FROM THE PART WHICH HE HAD INTRUSTED  
TO ME AVENUE 123.\nI HARDENED MY HEART WTASON STREET 45B, AND TOOK THE SMOKE-RO  
CKET FROM UNDER MY ULSTER.\nAFTER ALL,, 344-21-01 I THOUGHT, WE ARE NOT INJURIN  
G HER. WE ARE BUT PREVENTING\nHER FROM INJURING ANOTHER ADLERRR@MAIL.COM.'
```

```
In [11]: text.lower()
```

```
Out[11]: 'slowly and solemnly he was borne into briony lodge 145a and laid out in the\np  
rincipal room, while i (094)-44-33 still observed the proceedings from my post  
\nby the window 12/15/1892.    the lamps had been lit, but the blinds had not b  
een\nndrawn, so that i could see holmes sh_holmes221@mail.uk as he lay upon the  
couch.... i do not\nknow whether he was seized with compunction at that moment  
for the part\nhe was playing, but i know that i never felt more heartily ashame  
d of\nmyself in my life 34-34-32 than when i saw the beautiful creature against  
whom\ni wston23@mymail.com was conspiring, or the grace and kindness with whi  
ch she waited\nnupon the injured man. and yet    it would be the blackest treach  
ery to 1895-11-17\nholmes to draw back now from the part which he had intrusted  
to me avenue 123.\ni hardened my heart wtason street 45b, and took the smoke-ro  
cket from under my ulster.\nafter all,, 344-21-01 i thought, we are not injurin  
g her. we are but preventing\nnher from injuring another adlerrr@mail.com.'
```

```
In [12]: is_title= text.title()
```

```
is_title
```

Out[12]: 'Slowly And Solemnly He Was Borne Into Briony Lodge 145A And Laid Out In The\np
rincipal Room, While I (094)-44-33 Still Observed The Proceedings From My Post
\nBy The Window 12/15/1892. The Lamps Had Been Lit, But The Blinds Had Not B
een\nDrawn, So That I Could See Holmes Sh_Holmes221@Mail.Uk As He Lay Upon The
Couch.... I Do Not\nKnow Whether He Was Seized With Compunction At That Moment
For The Part\nHe Was Playing, But I Know That I Never Felt More Heartily Ashame
d Of\nMyself In My Life 34-34-32 Than When I Saw The Beautiful Creature Against
Whom\nI Wston23@Mymail.Com Was Conspiring, Or The Grace And Kindliness With Whi
ch She Waited\nUpon The Injured Man. And Yet It Would Be The Blackest Treach
ery To 1895-11-17\nHolmes To Draw Back Now From The Part Which He Had Intrusted
To Me Avenue 123.\nI Hardened My Heart Wtason Street 45B, And Took The Smoke-Ro
cket From Under My Ulster.\nAfter All,, 344-21-01 I Thought, We Are Not Injurin
g Her. We Are But Preventing\nHer From Injuring Another Adlerrr@Mail.Com.'

```
In [13]: text.capitalize()
```

```
Out[13]: 'Slowly and solemnly he was borne into briony lodge 145a and laid out in the\np  
rincipal room, while i (094)-44-33 still observed the proceedings from my post  
\nby the window 12/15/1892. the lamps had been lit, but the blinds had not b  
een\nndrawn, so that i could see holmes sh_holmes221@mail.uk as he lay upon the  
couch.... i do not\nknow whether he was seized with compunction at that moment  
for the part\nhe was playing, but i know that i never felt more heartily ashamed  
d of\nmyself in my life 34-34-32 than when i saw the beautiful creature against  
whom\ni wston23@mymail.com was conspiring, or the grace and kindness with whi  
ch she waited\nupon the injured man. and yet it would be the blackest treach  
ery to 1895-11-17\nholmes to draw back now from the part which he had intrusted  
to me avenue 123.\ni hardened my heart wtason street 45b, and took the smoke-ro  
cket from under my ulster.\nafter all,, 344-21-01 i thought, we are not injurin  
g her. we are but preventing\nher from injuring another adlerrr@mail.com.'
```

```
In [14]: split_text = text.split('a')
         '?'.join(split_text)
```

Out[14]: 'Slowly ?nd solemnly he w?s borne into Briony LODgE 145? ?nd l?id out in the\np
rincip?l room, while I (094)-44-33 still observed the proceedings from my post
\nby the window 12/15/1892. The l?mps h?d been lit, but the blinds h?d not b
een\ndr?wn, so th?t I could see Holmes sh_holmes221@m?il.uk ?s he l?y upon the
couch.... I do not\nknow whether he w?s seized with compunction ?t th?t moment
for the p?rt\nhe w?s pl?ying, but I know th?t I never felt more he?rtily ?sh?me
d of\nmyself in my life 34-34-32 th?n when I s?w the be?utiful cre?ture ?g?inst
whom\nI wston23@mym?il.com w?s conspiring, or the gr?ce ?nd kindness with whi
ch she w?ited\nupon the injured m?n. And yet it would be the bl?ckest tre?ch
ery to 1895-11-17\nHolmes to dr?w b?ck now from the p?rt which he h?d intrusted
to me AveNUe 123.\nI h?rdened my he?rt Wt?son Street 45b, ?nd took the smoke-ro
cket from under my ulster.\nAfter ?ll,, 344-21-01 I thought, we ?re not injurin
g her. We ?re but preventing\nher from injuring ?nother ?dlerrr@m?il.com.'

```
In [15]: replaced_text = text.replace('a', '?')
replaced text
```

```
Out[15]: 'Slowly and solemnly he was borne into Briony LODGE 145 and laid out in the principal room, while I (094)-44-33 still observed the proceedings from my post by the window 12/15/1892. The lamps had been lit, but the blinds had not been drawn, so that I could see Holmes sh_holmes221@mail.uk as he lay upon the couch.... I do not know whether he was seized with compunction at that moment for the part he was playing, but I know that I never felt more heartily ashamed of myself in my life 34-34-32 than when I saw the beautiful creature against whom I wston23@mymail.com was conspiring, or the grace and kindness with which she waited upon the injured man. And yet it would be the blackest treachery to 1895-11-17 Holmes to draw back now from the part which he had intrusted to me Avenue 123. I hardened my heart Wtason Street 45b, and took the smoke-rocket from under my ulster. After all,, 344-21-01 I thought, we are not injuring her. We are but preventing her from injuring another adlerr@mail.com.'
```

```
In [16]: text.startswith('S')
```

```
Out[16]: True
```

Task 2

```
In [17]: import re
```

```
In [18]: address_pattern = r'\b[A-Z][a-z]+\s*(?:[A-Za-z]+\s+)?\d+[a-zA-Z]?\b'
addresses = re.findall(address_pattern, text)
addresses
```

```
Out[18]: ['Briony LODGE 145a', 'Avenue 123', 'Wtason Street 45b']
```

```
In [22]: def replace_digits(match):
            address = match.group(0)
            first_digit_index = re.search(r'\d', address).start()
            modified_address = address[:first_digit_index+1] + re.sub(r'\d', 'X', address[first_digit_index+1:])
            return modified_address

        modified_text = re.sub(address_pattern, replace_digits, text)

        print(modified_text)
```

Slowly and solemnly he was borne into Briony LODGE 1XXa and laid out in the principal room, while I (094)-44-33 still observed the proceedings from my post by the window 12/15/1892. The lamps had been lit, but the blinds had not been drawn, so that I could see Holmes sh_holmes221@mail.uk as he lay upon the couch.... I do not know whether he was seized with compunction at that moment for the part he was playing, but I know that I never felt more heartily ashamed of myself in my life 34-34-32 than when I saw the beautiful creature against whom I wston23@mymail.com was conspiring, or the grace and kindness with which she waited upon the injured man. And yet it would be the blackest treachery to 1895-11-17 Holmes to draw back now from the part which he had intrusted to me Avenue 1XX. I hardened my heart Wtason Street 4Xb, and took the smoke-rocket from under my ulster. After all,, 344-21-01 I thought, we are not injuring her. We are but preventing her from injuring another adlerr@mail.com.

Висновок:

В ході виконання даної лабораторної роботи я ознайомився з основними методами та функціями для роботи з текстовими даними в Python, зокрема з представленням тексту у вигляді рядків та використанням регулярних виразів для пошуку та заміни в тексті. Лабораторна робота дала можливість практично застосувати набуті теоретичні знання для вирішення конкретного завдання: знаходження всіх номерів будинків у тексті та заміна всіх цифр після першої на вказаний символ.

Процес включав читання текстового файлу, використання регулярних виразів для ідентифікації відповідних фрагментів тексту, та модифікацію знайдених збігів згідно з умовами задачі. Це дозволило глибше розуміти, як можна маніпулювати текстовими даними, а також важливість регулярних виразів у процесі обробки тексту.

Однією з ключових вигод лабораторної роботи було освоєння використання функцій `re.findall` для знаходження всіх відповідностей певному шаблону в тексті та `re.sub` для заміни знайдених фрагментів на нові, модифіковані рядки. Ці навички будуть корисними для подальших завдань по обробці та аналізу текстових даних.

Завдяки виконанню цієї роботи я набув цінного досвіду в роботі з текстовими даними на Python, що безперечно знадобиться мені у майбутньому для реалізації проєктів, пов'язаних із обробкою та аналізом тексту.