

Міністерство освіти і науки України
Одеський національний політехнічний університет
Інститут комп'ютерних систем
Кафедра інформаційних систем

Лабораторна робота №4
з дисципліни «Операційні Системи»

Тема: «Складна обробка текстових даних засобами оболонки Unix-подібних
ОС інтерфейсу командного рядка»

Виконав:
ст. гр. AI-204
Колесник К. В.

Перевірів:
Блажко О. А.

Мета роботи: придбання навичок складної обробки текстових даних роботи засобами оболонки Unix-подібних ОС інтерфейсу командного рядка.

Завдання для виконання:

1. Обробка текстових даних результатів роботи команд

Виконайте завдання, знаходячись на Linux-сервері за адресою з IP = 91.219.60.189

1.1) Виведіть на екран перші десять назв каталогів або файлів, назви яких починаються з перших трьох букв вашого прізвища в латиниці

1.2) Назва вашого домашнього каталогу містить ваші прізвище та ім'я в транслітерації. Виведіть на екран список назв домашніх каталогів користувачів, в назві яких присутня перша буква вашого прізвища.

1.3) Модифікуйте рішення попереднього завдання так, що на екран було виведено імена користувачів, відсортовані в порядку зростання значення, та без дублювання цих значень (використайте конвеєр з команд cut, sort, uniq).

1.4) В лабораторній роботі № 2 було описано структуру файлу /etc/passwd, який використовується для зберігання даних про облікові записи користувачів в Unix-подібних ОС. Створіть файл з назвою accounts.csv, який буде містити облікові записи з наступними колонками, розділеними символом «кома»: ім'я користувача (логін), ідентифікатор користувача, початковий каталог користувача.

1.5) З файлу accounts.csv отримайте на екран перелік облікових записів,

в яких четверта цифра ідентифікатора користувача співпадає з четвертою цифрою вашого ідентифікатора користувача.

2. Обробка текстового файлу формату CSV

Нехай існують два файли HTML-формату DOC з табличними даними, які визначено у відповідності із номером вашої команди та номером учасника в команді, як в лабораторній роботі № 3 -

<https://docs.google.com/spreadsheets/d/1ITMLdjcw7nboUo9nFw4NwkWcvjbZeZZlDVb2Hg4laUQ>

Необхідно створити файл формату CSV, який містить табличні данні, використовуючи наступні кроки:

2.1) базове створення файлу CSV-формату, використовуючи Google Docs за методичними рекомендаціями

[https://drive.google.com/file/d/17wTiwec-1QybKQje49VdjQFahkely1Ou:](https://drive.google.com/file/d/17wTiwec-1QybKQje49VdjQFahkely1Ou)

- кодування символів – UTF-8
- роздільник між колонками – символ коми;
- назви файлів визначити з урахуванням транслітерації свого прізвища та номеру файлу.

2.2) копіювання файлів на сервер за адресою з IP = 91.219.60.189 у домашній каталог учасника команди;

2.3) модифікація файлу CSV-формату:

- заміна в числах коми на точку з одночасним видаленням лапки;
- заміна підрядка « ... 1» на порожнє значення (якщо таке є);
- заміна символу – (дефіс) на порожнє значення (якщо таке є);

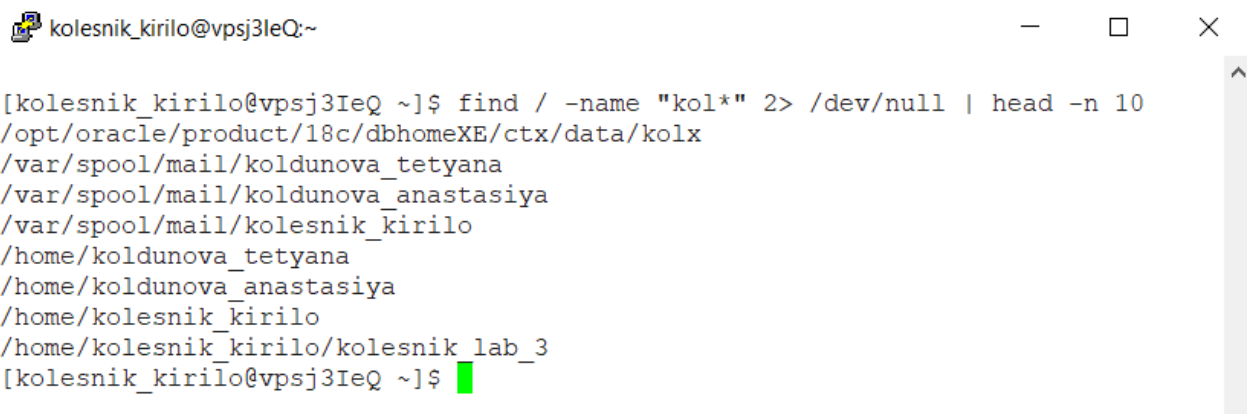
- видалення з файлу стовпчиків окрім першого та другого
- 2.4) передача створеного CSV-файлу в GitHub-репозиторій;
- 2.5) отримання CSV-файлу, створеного іншим учасником команди;
- 2.6) об'єднання двох файлів в один за першим стовпчиком;
- 2.7) виведення на екран змісту файлу з рішення попереднього завдання у зворотному напрямі для колонки за номером 2 - для 1-го учасника команди, або за номером 3 – для 2-го учасника команди.

Хід роботи:

1. Обробка текстових даних результатів роботи команд

Виконано завдання, знаходячись на Linux-сервері за адресою з IP = 91.219.60.189

1.1) Виведено на екран перші десять назв каталогів та файлів, назви яких починаються з перших трьох букв прізвища в латиниці – kol (рис. 1.1)



```
kolesnik_kirilo@vpsj3IeQ:~  
[kolesnik_kirilo@vpsj3IeQ ~]$ find / -name "kol*" 2> /dev/null | head -n 10  
/opt/oracle/product/18c/dbhomeXE/ctx/data/kolx  
/var/spool/mail/koldunova_tetyana  
/var/spool/mail/koldunova_anastasiya  
/var/spool/mail/kolesnik_kirilo  
/home/koldunova_tetyana  
/home/koldunova_anastasiya  
/home/kolesnik_kirilo  
/home/kolesnik_kirilo/kolesnik_lab_3  
[kolesnik_kirilo@vpsj3IeQ ~]$
```

Рис. 1.1

1.2) Виведено на екран список назв домашніх каталогів користувачів, в назві яких присутня перша буква мого прізвища – k (рис. 1.2).

```
[kolesnik_kirilo@vpsj3IeQ ~]$ find /home -maxdepth 1 -name "k*" 2> /dev/null
/home/kebab_dmitro
/home/kucherenko_maksim
/home/kovach_dmitro
/home/koldunova_tetyana
/home/kurgan_roman
/home/koldunova_anastasiya
/home/katrich_evgen
/home/kalarashan_kostyantyn
/home/kolesnik_kirilo
/home/kalina_marina
/home/karaulnij_daniil
/home/kruts_viktor
/home/krutienko_andrij
/home/kostetskiy_bogdan
/home/kelembet_evgen
```

Рис. 1.2

1.3) Модифіковано рішення попереднього завдання так, що на екран виведено імена користувачів, відсортовані в порядку зростання значення, та без дублювання цих значень, використано конвеєр з команд cut, sort, uniq (рис. 1.3).

```
[kolesnik_kirilo@vpsj3IeQ ~]$ find /home -maxdepth 1 -name "k*" 2> /dev/null | c
ut -d '_' -f2 | sort | uniq
anastasiya
andrij
bogdan
daniil
dmitro
evgen
kirilo
kostyantyn
maksim
marina
roman
tetyana
viktor
[kolesnik_kirilo@vpsj3IeQ ~]$
```

Рис. 1.3

1.4) Створено файл з назвою accounts.csv, який містить облікові записи з наступними колонками, розділеними символом «кома»: ім'я користувача (логін), ідентифікатор користувача, початковий каталог користувача (рис. 1.4).

```
[kolesnik_kirilo@vpsj3IeQ ~]$ cut -d ':' -f1,3,6 /etc/passwd | tr ':' ' ' > accounts.csv
[kolesnik_kirilo@vpsj3IeQ ~]$ ls
accounts.csv          lab3File.html          os.lab1.utf.html
KolesnikKirilo.docx   newCatalog              student.txt
KolesnikKirilo.pdf    Operating-System.-Laboratory-Work-1
kolesnik_lab_3        os.lab1.cp1251.html
[kolesnik_kirilo@vpsj3IeQ ~]$ cat accounts.csv
root,0,/root
bin,1,/bin
daemon,2,/sbin
adm,3,/var/adm
lp,4,/var/spool/lpd
sync,5,/sbin
shutdown,6,/sbin
halt,7,/sbin
```

Рис. 1.4

1.5) З файлу accounts.csv отримано на екран перелік облікових записів, в яких четверта цифра ідентифікатора користувача співпадає з четвертою цифрою власного ідентифікатора користувача (543**9**4) (рис. 1.5).

```
[kolesnik_kirilo@vpsj3IeQ ~]$ grep -E "[0-9]{3}9" "accounts.csv"
goshurenko_volodimir,54390,/home/goshurenko_volodimir
karaulnij_daniil,54391,/home/karaulnij_daniil
dorozhkin_mihajlo,54392,/home/dorozhkin_mihajlo
plaksivij_danilo,54393,/home/plaksivij_danilo
kolesnik_kirilo,54394,/home/kolesnik_kirilo
nesterenko_mikola,54395,/home/nesterenko_mikola
kostetskij_bogdan,54396,/home/kostetskij_bogdan
maksimenko_andrij,54397,/home/maksimenko_andrij
ozarchuk_anna,54398,/home/ozarchuk_anna
sirenko_mariya,54399,/home/sirenko_mariya
```

Рис. 1.5

2. Обробка текстового файлу формату CSV

Створено файл формату CSV (використовуючи таблицю – команда 6 учасник 1), який містить табличні дані, використовуючи наступні кроки:

2.1) Створено базовий файл CSV-формату, використовуючи Google Docs за методичними рекомендаціями (рис. 2.1):

- кодування символів – UTF-8
- роздільник між колонками – символ коми;
- назви файлів визначено з урахуванням транслітерації свого прізвища

та номеру файлу.

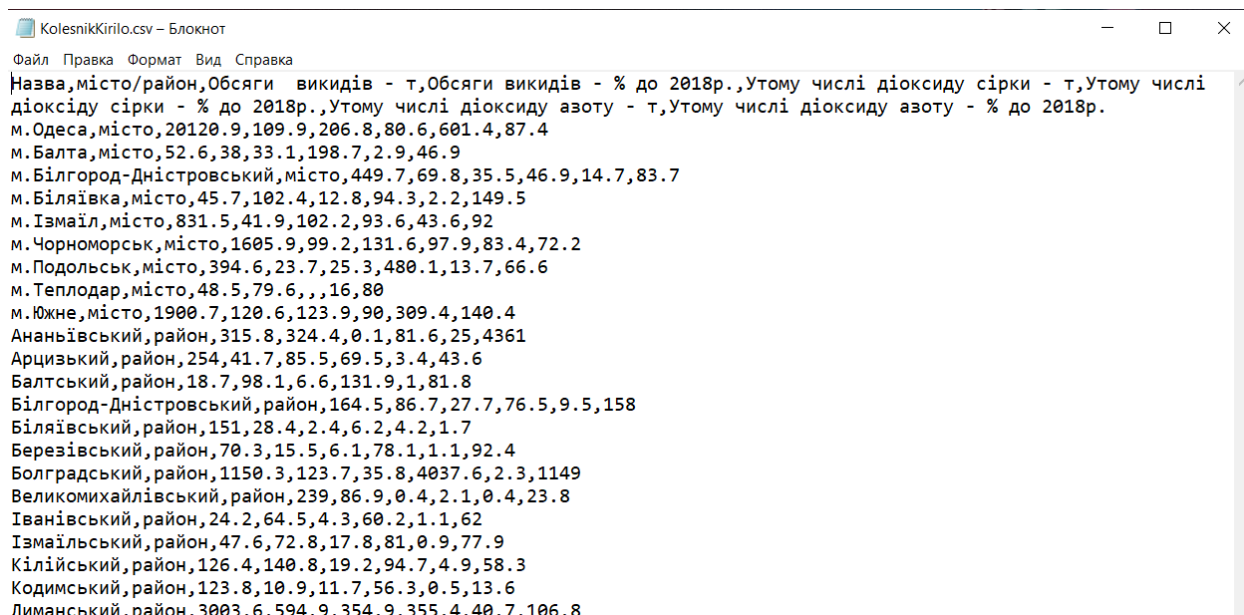


Рис. 2.1

2.2) Скопійовано файл на сервер за адресою з IP = 91.219.60.189 у домашній каталог учасника команди (рис. 2.2);

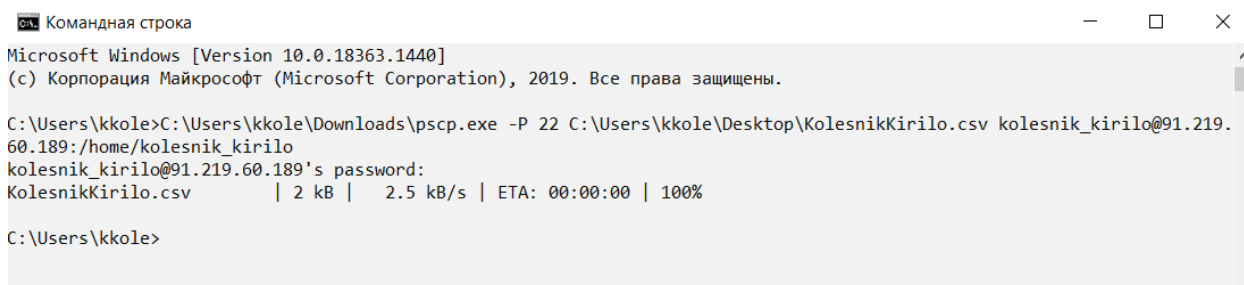


Рис. 2.2

2.3) Модифіковано файл CSV-формату (рис. 2.3):

- замінено в числах коми на точку з одночасним видаленням лапки;
- замінено підрядок « ... 1» на порожнє значення (якщо таке є);
- замінено символ – (дефіс) на порожнє значення (якщо таке є);
- видалено з файлу стовпчики окрім першого та другого

```

kolesnik_kirilo@vpsj3IeQ:~
[kolesnik_kirilo@vpsj3IeQ ~]$ ls
accounts.csv      kolesnik_lab_3      os.lab1.cp1251.html
KolesnikKirilo.csv  lab3File.html       os.lab1.utf.html
KolesnikKirilo.docx newCatalog           student.txt
KolesnikKirilo.pdf Operating-System.-Laboratory-Work-1
[kolesnik_kirilo@vpsj3IeQ ~]$ sed 's/\\"([0-9]\+\\)\[,]\\"([0-9]\+\\)\\"/\\1\\.\\2 /g' |
sed 's/\(,\)...\1/g' | sed 's/\(,\)\\"-/\\1/g' | cut -f1,3 -d, KolesnikKirilo.csv
> 1.csv
^C
[kolesnik_kirilo@vpsj3IeQ ~]$ ls
1.csv      KolesnikKirilo.pdf  Operating-System.-Laboratory-Work-1
accounts.csv  kolesnik_lab_3      os.lab1.cp1251.html
KolesnikKirilo.csv  lab3File.html       os.lab1.utf.html
KolesnikKirilo.docx newCatalog           student.txt
[kolesnik_kirilo@vpsj3IeQ ~]$ less 1.csv
Назва,Обсяги  викидів - т
м.Одеса,20120.9
м.Валта,52.6
м.Вінгород-Дністровський,449.7
м.Віляївка,45.7
м.Ізмаїл,831.5
м.Чорноморськ,1605.9
м.Подольськ,394.6
м.Теплодар,48.5
м.Южне,1900.7
Ананьївський,315.8
Арцизький,254
Балтський,18.7
Вінгород-Дністровський,164.5
Віляївський,151
Березівський,70.3
Волградський,1150.3
Великомихайлівський,239
Іванівський,24.2
Ізмаїльський,47.6

```

Рис. 2.3

2.4) передано створений CSV-файл в GitHub-репозиторій (рис. 2.4);

```

[kolesnik_kirilo@vpsj3IeQ ~]$ ls
1.csv      KolesnikKirilo.pdf  Operating-System.-Laboratory-Work-1
accounts.csv  kolesnik_lab_3      os.lab1.cp1251.html
KolesnikKirilo.csv  lab3File.html       os.lab1.utf.html
KolesnikKirilo.docx newCatalog           student.txt
[kolesnik_kirilo@vpsj3IeQ ~]$ git clone https://github.com/kirillkolesnyk/Operat
ing-System.-Laboratory-Work-4
Cloning into 'Operating-System.-Laboratory-Work-4'...
warning: You appear to have cloned an empty repository.
[kolesnik_kirilo@vpsj3IeQ ~]$ git init
Reinitialized existing Git repository in /home/kolesnik_kirilo/.git/
[kolesnik_kirilo@vpsj3IeQ ~]$ git add 1.csv
[kolesnik_kirilo@vpsj3IeQ ~]$ git commit -m "KolesnikKiriloM1"
[master (root-commit) 497104c] KolesnikKiriloM1
1 file changed, 36 insertions(+)
 create mode 100644 1.csv
[kolesnik_kirilo@vpsj3IeQ ~]$ git remote add origin https://github.com/kirillkol
esnyk/Operating-System.-Laboratory-Work-4
[kolesnik_kirilo@vpsj3IeQ ~]$ git push origin master
Username for 'https://github.com': kirillkolesnyk
Password for 'https://kirillkolesnyk@github.com':
Counting objects: 3, done.
Delta compression using up to 2 threads.
Compressing objects: 100% (2/2), done.
Writing objects: 100% (3/3), 695 bytes | 0 bytes/s, done.
Total 3 (delta 0), reused 0 (delta 0)
To https://github.com/kirillkolesnyk/Operating-System.-Laboratory-Work-4
 * [new branch]      master -> master
[kolesnik_kirilo@vpsj3IeQ ~]$

```


Рис. 2.4

2.5) отримано CSV-файл, створеного іншим учасником команди (рис. 2.5);

```
[kolesnik_kirilo@vpsj3IeQ ~]$ cp Operating-System.-Laboratory-Work-4/2.csv /home/kolesnik_kirilo
[kolesnik_kirilo@vpsj3IeQ ~]$ ls
1.csv          newCatalog
2.csv          Operating-System.-Laboratory-Work-1
accounts.csv   Operating-System.-Laboratory-Work-4
KolesnikKirilo.csv  os.lab1.cp1251.html
KolesnikKirilo.docx os.lab1.utf.html
KolesnikKirilo.pdf  student.txt
kolesnik_lab_3      Нестеренко.pdf
lab3File.html
[kolesnik_kirilo@vpsj3IeQ ~]$
```

Рис. 2.5

2.6) об'єднано два файли в один за першим стовпчиком (рис. 2.6);

```
[kolesnik_kirilo@vpsj3IeQ ~]$ join -t, -a 1 -a 2 -o0,1.2,2.2 -e ' ' 1.csv 2.csv > 3.csv
join: 1.csv:11: is not sorted: Ананьївський,315.8
join: 2.csv:11: is not sorted: Ананьївський,440.2
[kolesnik_kirilo@vpsj3IeQ ~]$
```

Рис. 2.6

2.7) виведено на екран зміст файлу з рішення попереднього завдання у зворотному напрямі для колонки за номером 2 (рис. 2.7).

```
[kolesnik_kirilo@vpsj3IeQ ~]$ sort -t, -rnk 2 3.csv
Арцизький,254,6617.6
Великомихайлівський,239,1027.5
Тарутинський,202,8395.6
Біляївський,151,2525.9
Ширяївський,90,125
м.Одеса,20120.9,390722.9
```

Рис. 2.7

Висновок: Під час виконання цієї лабораторної роботи, було на практиці розглянуто принципи взаємодії типу “клієнт – сервер” із детальним розгляненням функцій та можливостей для складної обробки текстових даних у Unix-подібних ОС інтерфейсу командного рядка. Виявлено, що користувач

отримує доступ до більш складних функцій для створення та обробки файлів із текстовими даними для того із можливістю здійснювати детальніші операції. Також за допомогою Google таблиць було створено файл у потрібному табличному форматі – csv.

На мою думку, найбільш складними виявилися завдання із отриманням файлу партнера з репозиторію (виникали складності у оновленні вмісту репозиторію), також об'єднання та сортування таблиць (через складність реалізації операцій). Тим не менш, методичні вказівки до роботи та поради з мережі інтернет допомогли краще зрозуміти процес.