

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:
Демцун А.О

Перевірили:
Блажко О.А

Одеса 2021

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

Хід роботи:

1. Перелік завдань до лабораторної роботи
2. Результатами виконання пунктів завдань
3. Висновки

Перелік завдань до лабораторної роботи

1.1 Обробка текстових даних результатів роботи команд

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

- 1) Виведіть на екран перші десять назв каталогів або файлів, назви яких починаються з перших трьох букв вашого прізвища в латиниці
- 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.
- 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).
- 4) В лабораторній роботі No 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.
- 5) З файлу accounts.csv отримайте на екран перелік облікових записів, в яких четверта

цифра ідентифікатора користувача співпадає з четвертою цифрою вашого ідентифікатора користувача.

1.2 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 No 3 –

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

Необхідно створити файл формату CSV, який містить табличні данні, використовуючи

наступні кроки:

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

<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) копіювання файлів на сервер за адресою з IP = 91.219.60.189 у домашній каталог

учасника команди;

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
- 4) передача створеного CSV-файлу в GitHub-репозиторій;
 - 5) отримання CSV-файлу, створеного іншим учасником команди;
 - 6) об'єднання двох файлів в один за першим стовпчиком;
 - 7) виведення на екран змісту CSV-файлу з рішення попереднього завдання із сортуванням рядків файлу-таблиці у зворотному напрямі значень колонки за номером 2 – для 1-го учасника команди, або за номером 3 – для 2-го учасника команди.

Результати виконання завдань:

1.1 Обробка текстових даних результатів роботи команд

```
-bash-4.2$ find / -name "dem*" 2> /dev/null | head
/opt/oracle/product/18c/dbhomeXE/demo
/opt/oracle/product/18c/dbhomeXE/sdk/demo
/opt/oracle/product/18c/dbhomeXE/sdk/demo/demo_procob_ic.mk
/opt/oracle/product/18c/dbhomeXE/sdk/demo/demo_proc_ic.mk
/opt/oracle/product/18c/dbhomeXE/R/library/ORE/Meta/demo.rds
/var/spool/mail/demtsun_andrij
/usr/local/webuzo/enduser/themes/default/images/webuzo/demos.png
/usr/local/webuzo/enduser/themes/default/images/demos.png
/usr/local/webuzo/enduser/themes/default/images/demos_small.gif
/usr/local/webuzo/enduser/themes/default/images/demos.gif
-bash-4.2$ █
```

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

```
-bash-4.2$ find /home -maxdepth 1 -name "d*" 2> \dem\null  
/home/dang_din  
/home/dorozhkin_mihajlo  
/home/demtsun_andrij
```

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

```
-bash-4.2$ find /home -maxdepth 1 -name "d*" 2> \dem\null | cut -d'_' -f2 | sort | uniq  
andrij  
din  
mihajlo  
-bash-4.2$
```

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

```

-bash-4.2$ cut -d':' -f1,3,6 /etc/passwd |tr ':' ',' > accounts.csv
-bash-4.2$ ls
accounts.csv  hata.html  Operating-System.-Laboratory-Work-1
demnull      lab2.docx  os.lab1.cpl251.html
dentsun_lab_3 lab2.pdf   os.lab1.utf.html
-bash-4.2$ 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
mail,8,/var/spool/mail
operator,11,/root
games,12,/usr/games
ftp,14,/var/ftp
nobody,99,/
systemd-network,192,/
dbus,81,/
polkitd,999,/
postfix,89,/var/spool/postfix
chrony,998,/var/lib/chrony
sshd,74,/var/empty/sshd
soft,1000,/home/soft
saslauth,997,/run/saslauthd
mailnull,47,/var/spool/mqueue
smmsp,51,/var/spool/mqueue
emps,996,/home/emps
rpc,32,/var/lib/rpcbind
rpcuser,29,/var/lib/nfs
nfsnobody,65534,/var/lib/nfs
oracle,54321,/home/oracle
apache,995,/home/apache
mysql,27,/var/lib/mysql
named,25,/var/named
exim,93,/var/spool/exim

```

Створимо 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.

```
-bash-4.2$ 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
-bash-4.2$
```

З файлу accounts.csv отримаємо на екран перелік облікових записів, в яких четверта цифра ідентифікатора користувача співпадає з четвертою цифрою вашого ідентифікатора користувача (54392)

1.2 Обробка текстового файлу формату CSV

Demtsun_Andrew – Блокнот

Файл Правка Формат Вид Справка

Назва,місто/район,"На зерно та зелений корм
га","На зерно та зелений корм
до площ під урожай 2020р.", "Усього на зернові культури
га", "Усього на зернові культури
у % до площ під урожай 2020р.", "На пшеницю
га", "На пшеницю
у % до площ під урожай 2020р.", "На ріпак
га", "На ріпак
у % до площ під урожай 2020р."

Одеська область,,827957.23,91.6,698217.52,97.6,476136.97,97.9,128670.82,68.6
м.Одеса,місто,,,,,,
м.Балта,місто,15899.12,727.1,13631.66,672.2,11491.7,696.9,2267.46,1427.9
м.Білгород-Дністровський,місто,6046.73,68.5,4785.78,73.6,3223.33,91,1260.95,54.3
м.Біляївка,місто,2933.87,138.1,2810.2,206.3,2026.62,326.9,,
м.Ізмаїл,місто,,,,,,
Ананьївський,район,22598.17,93.2,19815.86,87.6,15045.99,79.3,2782.31,173.1
Арцизький,район,47539.87,94.1,38934.93,105.5,22182.03,100.3,8604.94,63.7
Балтський,район,7488.43,35.1,6849.39.3,5226.23,37.7,639.43,16.3
Білгород-Дністровський,район,48609.2,86.8,35279.04,92.1,23338.51,91.7,13225.98,75.4
Біляївський,район,25090.64,88.9,24118.05,94.9,18195.95,100.3,652,25.9
Березівський,район,36208.85,78.9,28875.47,75.8,22608.3,78.8,7038.38,94.5
Болградський,район,39790.18,97.1,28518.98,76,16479.58,76.2,11271.2,325
Великомихайлівський,район,42379.5,97.7,35536.1,95.8,25018.56,101.8,6843.4,108.8
Іванівський,район,35693.09,98.8,33359.22,105.4,22653.22,104.7,2264.75,51.3
Ізмаїльський,район,39169.18,102.7,36541.77,116.6,21863.3,117.2,2627.41,38.7
Кілійський,район,19834.94,85.6,19145.01,87.4,10547.85,85.6,689.93,54.6
Кодимський,район,12720.82,126.4,12139.82,136.6,8769.62,144.6,581,49.2
Лиманський,район,56547.96,92.5,39041.93,100.4,27222.8,108.1,17506.03,78.7
Подільський,район,19982.96,90.8,19529.96,94.8,13669.99,91.7,453,32.3
Окнянський,район,28357.75,95.9,22627.71,101.3,17899.78,110.6,5730.04,79.3
Любашівський,район,20580.24,88,19125.02,94.6,14082.99,86.3,1455.22,45.9
Миколаївський,район,16773.68,76.9,16033.18,100.2,12419.39,95.3,740.5,12.7
Овідіопольський,район,21584.4,96.6,18782.4,104.8,11645.79,101.7,2522,60.3
Роздільнянський,район,36756.97,77.1,34614.33,102.8,27981.56,107.7,2142.64,15.3
Ренійський,район,9661.55,77.7,9043.55,73.4,3908.81,56.9,618,588.6
Савранський,район,8426.09,96.7,6489.09,84.8,4389.89,74.8,1937,183.8
Саратський,район,65448.89,89.9,50786.25,104.9,31667.06,110.6,14662.64,60.9
Тарутинський,район,57929.93,91.2,47767.09,97.8,32717.83,89.8,10162.84,69.1
Татарбунарський,район,31297.55,89.9,26993.04,99.1,13271.75,95.6,4304.51,56.9
Захарівський,район,13285.09,100,12094.28,102.2,9159.71,102.7,1190.81,82.3
Ширяївський,район,37478.73,99.3,33270.95,107.9,26678.88,112.1,4207.78,60.9

Створимо файл формату CSV, який містить табличні данні

```
C:\Windows\system32\cmd.exe
Microsoft Windows [Version 10.0.19042.867]
(c) Корпорация Майкрософт (Microsoft Corporation), 2020. Все права защищены.

C:\Users\Admin>C:\Users\Admin\Desktop\pscp.exe -P 22 C:\Users\Admin\Desktop\Demtsun_Andrew.csv demtsun_andrij@91.219.60.189:/home/demtsun_andrij
demtsun_andrij@91.219.60.189's password:
Demtsun_Andrew.csv      | 3 kB |   3.3 kB/s | ETA: 00:00:00 | 100%

C:\Users\Admin>_
```

Скопіюємо файл на сервер за адресою з IP = 91.219.60.189 у домашній каталог
учасника команди


```

-bash-4.2$ sed 's/\\"([0-9]\\+\\)[,]\\\"([0-9]\\+\\)\"/\\1\\.\\2 /g' Demtsun_Andrew.csv >
1.csv
-bash-4.2$ sed 's/\\(,\\)\\.\\.\\.1/\\1/g' 1.csv > 11.csv
-bash-4.2$ sed 's/\\(,\\)\\-/\\1/g' 11.csv > 111.csv
-bash-4.2$ cut -f1,2 -d, 111.csv > 1.csv
-bash-4.2$ █

-bash-4.2$ cat 1.csv
Назва,місто/район
га", "На зерно та зелений корм
до площ під урожай 2020р.", "Усього на зернові культури
га", "Усього на зернові культури
у % до площ під урожай 2020р.", "На пшеницю
га", "На пшеницю
у % до площ під урожай 2020р.", "На ріпак
га", "На ріпак
у % до площ під урожай 2020р."
Одеська область,
м.Одеса,місто
м.Балта,місто
м.Білгород-Дністровський,місто
м.Біляївка,місто
м.Ізмаїл,місто
Ананьївський,район
Арцизький,район
Балтський,район
Білгород-Дністровський,район
Біляївський,район
Березівський,район
Болградський,район
Великомихайлівський,район
Іванівський,район
Ізмаїльський,район
Кілійський,район
Кодимський,район
Лиманський,район
Подільський,район
Окнянський,район
Любашівський,район
Миколаївський,район
Овідіопольський,район
Роздільнянський,район
Ренійський,район
Савранський,район
Саратський,район
Тарутинський,район
Татарбунарський,район
Захарівський,район
Ширяївський,район
-bash-4.2$ █

```

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

```

login as: demtsun_andrij
demtsun_andrij@91.219.60.189's password:
Last login: Sun Mar 21 16:57:11 2021 from 188.163.103.236
-bash-4.2$ git config --global user.name "Andrew Demtsun"
-bash-4.2$ git config --global user.email demtsun.8864993@stud.opu.ua
-bash-4.2$ git init
Reinitialized existing Git repository in /home/demtsun_andrij/.git/
-bash-4.2$ ^C
-bash-4.2$ git config --global user.name "Andrew Demtsun"
-bash-4.2$ git config --global user.email demtsun.8864993@stud.opu.ua
-bash-4.2$ git init
Reinitialized existing Git repository in /home/demtsun_andrij/.git/
-bash-4.2$ git clone https://github.com/NikolayHolovanchuk/Operating-System.-Lab
oratory-Work-1
Cloning into 'Operating-System.-Laboratory-Work-1'...
remote: Enumerating objects: 54, done.
remote: Counting objects: 100% (54/54), done.
remote: Compressing objects: 100% (52/52), done.
remote: Total 54 (delta 15), reused 3 (delta 0), pack-reused 0
Unpacking objects: 100% (54/54), done.
-bash-4.2$ cd Operating-System.-Laboratory-Work-1
-bash-4.2$ git pull
Already up-to-date.
-bash-4.2$ ls
l.csv                Holovanchuk_mykola.csv    Слайд 1.png  Слайд 5.png
Demstun.lab3.pdf     Holovanchuk Mykola.txt    Слайд 2.png
demtsun_andrij.txt   ЛабораторнаРобота№4.docx  Слайд 3.png
Demtsun.lab2.pdf     OC протокол №2.docx       Слайд 4.png
-bash-4.2$ cd ..
-bash-4.2$ cp l.csv Operating-System.-Laboratory-Work-1/l.csv
cp: cannot stat 'l.csv': No such file or directory
-bash-4.2$ cd Operating-System.-Laboratory-Work-1
-bash-4.2$ git add *
-bash-4.2$ git commit -m "dem lab4"
[main 51e263b] dem lab4
 1 file changed, 41 insertions(+)
 create mode 100644 l.csv
-bash-4.2$ git push
warning: push.default is unset; its implicit value is changing in
Git 2.0 from 'matching' to 'simple'. To squelch this message
and maintain the current behavior after the default changes, use:

    git config --global push.default matching

To squelch this message and adopt the new behavior now, use:

    git config --global push.default simple

See 'git help config' and search for 'push.default' for further information.
(the 'simple' mode was introduced in Git 1.7.11. Use the similar mode
'current' instead of 'simple' if you sometimes use older versions of Git)

Username for 'https://github.com': lstrebitelPiva
Password for 'https://lstrebitelPiva@github.com':
Counting objects: 4, done.
Delta compression using up to 2 threads.
Compressing objects: 100% (3/3), done.
Writing objects: 100% (3/3), 713 bytes | 0 bytes/s, done.
Total 3 (delta 1), reused 0 (delta 0)
remote: Resolving deltas: 100% (1/1), completed with 1 local object.
To https://github.com/NikolayHolovanchuk/Operating-System.-Laboratory-Work-1
 2f770a3..51e263b  main -> main
-bash-4.2$ git pull
Already up-to-date.

```

Передача створеного CSV-файлу в GitHub-репозиторій та отримання CSV-файлу, створеного іншим учасником команди

```
-bash-4.2$ cd Operating-System.-Laboratory-Work-1
-bash-4.2$ cat 2.csv > m2.csv
-bash-4.2$ cat 3.csv | -f2 -d,
-bash: -f2: command not found
-bash-4.2$ cat 3.csv | cut -f2 -d,
На зерно та зелений корм
га
```

```
827957.23
к
15899.12
6046.73
2933.87
к
```

```
22598.17
47539.87
7488.43
48609.2
25090.64
36208.85
39790.18
42379.5
35693.09
39169.18
19834.94
12720.82
56547.96
19982.96
28357.75
20580.24
16773.68
21584.4
36756.97
.55
8426.09
65448.89
57929.93
31297.55
13285.09
37478.73
```

```

-bash-4.2$ cat 3.csv | cut -f2 -d, > 3m.csv
-bash-4.2$ paste -d ',' 2.csv 3m.csv
,На зерно та зелений корм
Одеська область,827957.23,га
м.Одеса,,
м.Балта,15899.12,
м.Білгород-Дністровський,6046.73,827957.23
м.Біляївка,2933.87,к
м.Ізмаїл,,15899.12
Ананьївський,22598.17,6046.73
Арцизький,47539.87,2933.87
Балтський,7488.43,к
Білгород-Дністровський,48609.2,
Біляївський,25090.64,22598.17
Березівський,36208.85,47539.87
Болградський,39790.18,7488.43
Великомихайлівський,42379.5,48609.2
Іванівський,35693.09,25090.64
Ізмаїльський,39169.18,36208.85
Кілійський,19834.94,39790.18
Кодимський,12720.82,42379.5
Лиманський,56547.96,35693.09
Подільський,19982.96,39169.18
Окнянський,28357.75,19834.94
Любашівський,20580.24,12720.82
Миколаївський,16773.68,56547.96
Овідіопольський,21584.4,19982.96
Роздільнянський,36756.97,28357.75
Ренійський,9661.55,20580.24
Савранський,8426.09,16773.68
Саратський,65448.89,21584.4
Тарутинський,57929.93,36756.97
Татарбунарський,31297.55,.55
Захарівський,13285.09,8426.09
Ширяївський,37478.73,65448.89
,57929.93
,31297.55
,13285.09
,37478.73
-bash-4.2$ paste -d ',' 2.csv 3m.csv > m2.csv

```

Об'єднання двох файлів в один за першим стовпчиком

```

-bash-4.2$ sort -t, -k2 m2.csv
м.Одеса,,
Кодимський,12720.82,42379.5
,13285.09
Захарівський,13285.09,8426.09
м.Ізмаїл,,15899.12
м.Балта,15899.12,
Миколаївський,16773.68,56547.96
Кілійський,19834.94,39790.18
Подільський,19982.96,39169.18
Любашівський,20580.24,12720.82
Овідіопольський,21584.4,19982.96
Ананьївський,22598.17,6046.73
Біляївський,25090.64,22598.17
Окнянський,28357.75,19834.94
м.Біляївка,2933.87,к
,31297.55
Татарбунарський,31297.55,.55
Іванівський,35693.09,25090.64
Березівський,36208.85,47539.87
Роздільнянський,36756.97,28357.75
,37478.73
Ширяївський,37478.73,65448.89
Ізмаїльський,39169.18,36208.85
Болградський,39790.18,7488.43
Великомихайлівський,42379.5,48609.2
Арцизький,47539.87,2933.87
Білгород-Дністровський,48609.2,
Лиманський,56547.96,35693.09
,57929.93
Тарутинський,57929.93,36756.97
м.Білгород-Дністровський,6046.73,827957.23
Саратський,65448.89,21584.4
Балтський,7488.43,к
Одеська область,827957.23,та
Савранський,8426.09,16773.68
Ренійський,9661.55,20580.24
,На зерно та зелений корм
-bash-4.2$ █

```

Виведення на екран змісту CSV-файлу з рішення попереднього завдання із сортуванням рядків файлу-таблиці у зворотному напрямі значень колонки за номером 2

Висновок: Під час виконання лабораторної роботи, ми отримали практичні навички складної обробки текстових даних роботи засобами оболонки Unix-подібних ОС інтерфейсу командного рядка. . Напевно, найскладнішими завданнями була обробка текстового файлу формату CSV, а саме модифікація файлу формату CSV.

