

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ОДЕСЬКИЙ
НАЦІОНАЛЬНИЙ ПОЛІТЕХНІЧНИЙ УНІВЕРСИТЕТ

Інститут комп'ютерних систем
Кафедра інформаційних систем

З дисципліни «Операційні системи»

Лабораторна робота № 4

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

Виконав:
студент групи АІ-204
Бабич А.Ю.

Перевірили:
Блажко О.А.

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

Завдання до виконання:

2.1 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.

2.2 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 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:

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

```
[babich_artem@vpsj3IeQ ~]$ find / name -name "bab*" 2> /dev/null | head
/var/spool/mail/babich_artem
/var/lib/yum/yumdb/o/bab98c0aa6c8b65b19b880646d880b1ed895136e-openldap-2.4.44-21
.el7_6-x86_64
/home/babich_artem
/home/babich_artem/babich_lab_3
```

1.2) Назва домашнього каталогу містить прізвище та ім'я в транслітерації.

Виводимо на екран список назв домашніх каталогів користувачів, в назві яких присутня перша буква прізвища

```
[babich_artem@vpsj3IeQ ~]$ find /home -maxdepth 1 -type d -name "b*" 2> /dev/null
/home/bojchuk_oleksandr
/home/barkar_karina
/home/berislavskij_vladislav
/home/belobrov_artur
/home/bodnar_illya
/home/borovets_vladislav
/home/baranyuk_dmitro
/home/bodnar_arsenij
/home/babich_artem
/home/bochkarova_karina
/home/bezhanishvili_levan
/home/bogachik_egor
```

1.3) Модифікуємо рішення попереднього завдання так, щоб на екран було виведено імена користувачів, відсортовані в порядку зростання значення, та без дублювання цих значень

```
[babich_artem@vpsj3IeQ ~]$ find /home -maxdepth 1 -type d -name "b*" 2> /dev/nul
l | cut -d'_' -f2 | sort |uniq
arsenij
artem
artur
dmitro
egor
illya
karina
levan
oleksandr
vladislav
```

4) В лабораторній роботі № 2 було описано структуру файлу /etc/passwd, який використовується для зберігання даних про облікові записи користувачів в Unix-подібних ОС. Створюємо 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.

```
[babich_artem@vpsj3IeQ ~]$ ls
BabichArtem.docx  chto-to      Operating-System.-Laboratory-Work-1
BabichArtem.pdf   fail.html    os.lab1.cp1251.html
babich_lab_3      kogo-to.txt  os.lab1.utf.html
[babich_artem@vpsj3IeQ ~]$ cut -d':' -f1,3,6 /etc/passwd | tr ':' ' ',' > accounts
.csv
[babich_artem@vpsj3IeQ ~]$ ls
accounts.csv      chto-to      os.lab1.cp1251.html
BabichArtem.docx fail.html     os.lab1.utf.html
BabichArtem.pdf  kogo-to.txt
babich_lab_3     Operating-System.-Laboratory-Work-1
[babich_artem@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
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
```

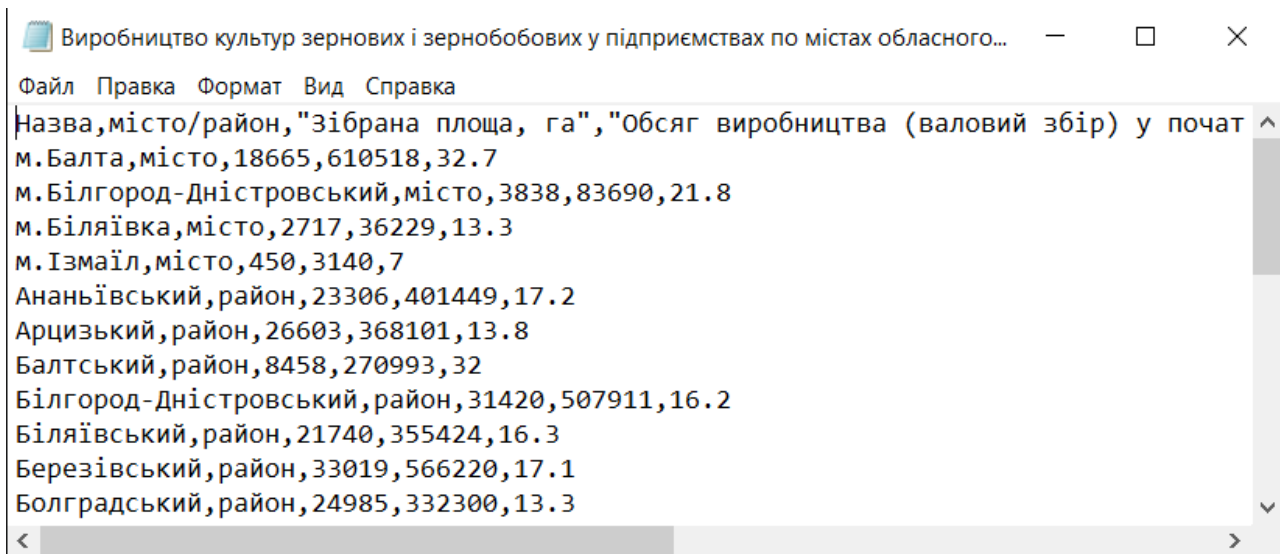
1.5) З файлу accounts.csv отримуємо на екран перелік облікових записів, в яких четверта цифра ідентифікатора користувача співпадає із власною четвертою цифрою ідентифікатора користувача

```
[babich_artem@vpsj3IeQ ~]$ grep -E ",[0-9]{3}8" "accounts.csv"
romanenko_oleksandr,54380,/home/romanenko_oleksandr
ahmadi_muhamad,54381,/home/ahmadi_muhamad
ahmadi_umran,54382,/home/ahmadi_umran
li_chuan,54383,/home/li_chuan
babich_artem,54384,/home/babich_artem
polovenko_illya,54385,/home/polovenko_illya
berislavskij_vladislav,54386,/home/berislavskij_vladislav
krutienko_andrij,54387,/home/krutienko_andrij
belobrov_artur,54388,/home/belobrov_artur
evchev denis,54389,/home/evchev denis
```

2.

Обробка текстового файлу формату CSV

2.1) Створюємо файл CSV-формату, використовуючи Google Docs за методичними рекомендаціями, із наступними характеристиками – кодування символів – UTF-8 – роздільник між колонками – символ коми; – назви файлів



визначити з урахуванням транслітерації свого прізвища та номеру файлу.

2.2) Копіюємо файл на сервер за адресою з IP = 91.219.60.189 у домашній каталог

```
Командная строка
C:\Users\artem>C:\users\artem\desktop\pscp.exe -P 22 C:\users\artem\desktop\BabichArtem.csv babich_artem@91.219.60.189:/
home/babich_artem
babich_artem@91.219.60.189's password:
BabichArtem.csv          | 1 kB | 1.8 kB/s | ETA: 00:00:00 | 100%
C:\Users\artem>
```

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

```
[babich_artem@vpsj3IeQ ~]$ sed 's/\"([0-9]\+)[,]\([0-9]\+\)\"/\1.\2/g' | sed 's/(\,)\...1/\1/g' | sed 's/(\,)\-/\1/g' | cut -d, -f1,2 BabichArtem.csv > 1.csv
^C
[babich_artem@vpsj3IeQ ~]$ ls
1.csv          BabichArtem.pdf  kogo-to.txt
accounts.csv   babich_lab_3     Operating-System.-Laboratory-Work-1
BabichArtem.csv chto-to          os.lab1.cp1251.html
BabichArtem.docx fail.html        os.lab1.utf.html
[babich_artem@vpsj3IeQ ~]$ less 1.csv
Назва,місто/район
м.Балта,місто
м.Вілгород-Дністровський,місто
м.Віляївка,місто
м.Ізмаїл,місто
Ананьївський,район
Арцизький,район
Балтський,район
Вілгород-Дністровський,район
Віляївський,район
Березівський,район
Болградський,район
Великомихайлівський,район
Іванівський,район
Ізмаїльський,район
```

2.4) Передача створеного CSV-файлу в GitHub-репозиторій

```
[babich_artem@vpsj3IeQ Operating-System.-Laboratory-Work-4]$ git push -u origin --all
Username for 'https://github.com': aptem9015
Password for 'https://aptem9015@github.com':
Counting objects: 3, done.
Delta compression using up to 2 threads.
Compressing objects: 100% (2/2), done.
Writing objects: 100% (3/3), 502 bytes | 0 bytes/s, done.
Total 3 (delta 0), reused 0 (delta 0)
To https://github.com/aptem9015/Operating-System.-Laboratory-Work-4
 * [new branch]      master -> master
Branch master set up to track remote branch master from origin.
```

2.5) Отримання CSV-файлу, створеного іншим учасником команди

```
[babich_artem@vpsj3IeQ ~]$ cp Operating-System.-Laboratory-Work-4/Operating-System.-Laboratory-Works/berislavskij_vladislav.csv /home/babich_artem
[babich_artem@vpsj3IeQ ~]$ ls
accounts.csv          fail.html
BabichArtem.csv       kogo-to.txt
BabichArtem.docx      Operating-System.-Laboratory-Work-1
BabichArtem.pdf       Operating-System.-Laboratory-Work-4
babich_lab_3          os.lab1.cp1251.html
berislavskij_vladislav.csv  os.lab1.utf.html
chto-to
```

2.6) Об'єднання двох файлів в один за першим стовпчиком

```
[babich_artem@vpsj3IeQ Operating-System.-Laboratory-Works]$ join -t, -a 1 -a 2 -o0,1.2,2.2 -e ' ' test.csv berislavskij_vladislav.csv
Назва,місто/район,Місто/Район
м.Балта,місто,
м.Вілгород-Дністровський,місто,
м.Віляївка,місто,
м.Ізмаїл,місто,
join: test.csv:6: is not sorted: Ананьївський,район
Ананьївський,район,
Арцизький,район,
Балтський,район,
Вілгород-Дністровський,район,
Віляївський,район,
Верезівський,район,
Болградський,район,
Великомихайлівський,район,
Іванівський,район,
Ізмаїльський,район,
Кілійський,район,
Кодимський,район,
Лиманський,район,
м.Одеса,,місто
join: berislavskij_vladislav.csv:3: is not sorted: м.Балта,місто,15899.12,727.1,13631.66,672.2,11491.7,696.9,2267.46,1427.9
```

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


```
[babich_artem@vpsj3IeQ Operating-System.-Laboratory-Works]$ sort --field-separator="," --key=2 -r test.csv berisclavskij.csv
```

Ширяївський, район
Ширяївський, район
Татарбунарський, район
Татарбунарський, район
Тарутинський, район
Тарутинський, район
Саратський, район
Саратський, район
Савранський, район
Савранський, район
Роздільнянський, район
Роздільнянський, район
Ренійський, район
Ренійський, район
Подільський, район
Подільський, район
Окнянський, район
Окнянський, район

Висновки: під час виконання лабораторної роботи, були придбані навички складної обробки текстових даних роботи засобами оболонки Unixподібних ОС інтерфейсу командного рядка. Найскладнішим завданням було 2.6 - склеювання двох csv файлів