

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:

Студент групи АІ-202

Полянський М.О.

Перевірили:

Блажко О. А.

Дрозд М.О.

Одеса 2021

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

План

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/1ITMLdjcw7nboUo9nFw4NwkWcvjbZeZzIDVb2Hg4laUQ>

Необхідно створити файл формату CSV, який містить табличні дані, 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 третього;
- 4) передача створеного CSV-файлу в GitHub-репозиторій;
- 5) отримання CSV-файлу з GitHub-репозиторій, 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.

2.1 Обробка текстових даних результатів роботи команд

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

```
[polyanskiy_mikola@vpsj3IeQ ~]$ find / -name "pol*" 2> /dev/null |head -n10
/sys/fs/cgroup/devices/system.slice/polkit.service
/sys/fs/cgroup/systemd/system.slice/polkit.service
/sys/kernel/security/ima/policy
/sys/module/drm_kms_helper/parameters/poll
/sys/module/pci_aspm/parameters/policy
/opt/oracle/product/18c/dbhomeXE/ctx/data/an1/lang/polish
/opt/oracle/product/18c/dbhomeXE/jdk/jre/lib/security/policy
/opt/oracle/product/18c/dbhomeXE/jdk/jre/bin/policytool
/opt/oracle/product/18c/dbhomeXE/jdk/bin/policytool
/opt/oracle/product/18c/dbhomeXE/dv/admin/policy_it.dlf
```

- 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.

```
[polyanskiy_mikola@vpsj3IeQ ~]$ find /home/ -name "p*" 2> /dev/null
/home/plaksivij_danilo
/home/polovenko_illya
/home/polyanskiy_mikola
/home/polyanskiy_mikola/Operating-System/.git/hooks/pre-applypatch.sample
/home/polyanskiy_mikola/Operating-System/.git/hooks/post-update.sample
/home/polyanskiy_mikola/Operating-System/.git/hooks/pre-rebase.sample
/home/polyanskiy_mikola/Operating-System/.git/hooks/pre-commit.sample
/home/polyanskiy_mikola/Operating-System/.git/hooks/pre-push.sample
/home/polyanskiy_mikola/Operating-System/.git/hooks/prepare-commit-msg.sample
/home/polyanskiy_mikola/Operating-System/.git/packed-refs
/home/polyanskiy_mikola/Operating-System/.git/objects/pack
```

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).

```
[polyanskiy_mikola@vpsj3IeQ ~]$ find /home/ -name "p*" 2> /dev/null|sort|uniq
/home/panfilov_kristian
/home/pargalova_sabina
/home/perets_sergij
/home/plaksivij_danilo
/home/polovenko_illya
/home/polyanskiy_mikola
```

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.

```
[polyanskiy_mikola@vpsj3IeQ ~]$ cut -f 1,4,6 -d: /etc/passwd>accounts.csv
[polyanskiy_mikola@vpsj3IeQ ~]$ cat accounts.csv
root:0:/root
bin:1:/bin
daemon:2:/sbin
adm:4:/var/adm
lp:7:/var/spool/lpd
sync:0:/sbin
shutdown:0:/sbin
halt:0:/sbin
mail:12:/var/spool/mail
operator:0:/root
games:100:/usr/games
ftp:50:/var/ftp
nobody:99:/
systemd-network:192:/
dbus:81:/
polkitd:997:/
postfix:89:/var/spool/postfix
chrony:996:/var/lib/chrony
```

```
[polyanskiy_mikola@vpsj3IeQ ~]$ sed 's:/:/g' accounts.csv
root,0,/root
bin,1,/bin
daemon,2,/sbin
adm,4,/var/adm
lp,7,/var/spool/lpd
sync,0,/sbin
shutdown,0,/sbin
halt,0,/sbin
mail,12,/var/spool/mail
operator,0,/root
games,100,/usr/games
ftp,50,/var/ftp
nobody,99,/
systemd-network,192,/
dbus,81,/
polkitd,997,/
postfix,89,/var/spool/postfix
chrony,996,/var/lib/chrony
sshd,74,/var/empty/sshd
soft,1000,/home/soft
```

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

```
[polyanskiy_mikola@vpsj3IeQ ~]$ grep "[0-9][0-9][0-9][6][0-9]" "accounts.csv"
melnychenko_egor:54360:/home/melnychenko_egor
perets_sergij:54361:/home/perets_sergij
uzun_mihajlo:54362:/home/uzun_mihajlo
polyanskiy_mikola:54363:/home/polyanskiy_mikola
sherbakov_artur:54364:/home/sherbakov_artur
ershov_oleksandr:54365:/home/ershov_oleksandr
zagorujko_maksim:54366:/home/zagorujko_maksim
zijodulloev_ehson:54367:/home/zijodulloev_ehson
veselkova_anna:54368:/home/veselkova_anna
voyakovskij_dmitro:54369:/home/voyakovskij_dmitro
```

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/1ITMLdjcw7nboUo9nFw4NwkWcvjbZeZlIDVb2Hg4laUQ>

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);

```
[polyanskiy_mikola@vpsj3IeQ ~]$ sed 's/\\(,\\)\\-\\/\\1/g' polyanskiy_3_2.csv>polyanskiy_4_2.csv
[polyanskiy_mikola@vpsj3IeQ ~]$ cat polyanskiy_4_2.csv
Назва,місто/район,Кількість живонароджених,Кількість померлих,Природний приріст скорочення (-)
м.Одеса,місто,542,1454,912
Березівський,район,43,48
Білгород-Дністровський,район,140,290,150
Болградський,район,106,214,108
Ізмаїльський,район,117,306,189
Подільський,район,146,364,218
```

– видалення з файлу стовпчиків окрім першого та третього;

```
[polyanskiy_mikola@vpsj3IeQ ~]$ cut -f1,3 -d, polyanskiy_4_2.csv>polyanskiy_5_2.csv
[polyanskiy_mikola@vpsj3IeQ ~]$ cat polyanskiy_5_2.csv
Назва,Кількість живонароджених
м.Одеса,542
Березівський,43
Білгород-Дністровський,140
Болградський,106
Ізмаїльський,117
Подільський,146
Роздільнянський,60
```

4) передача створеного CSV-файлу в GitHub-репозиторій;

5) отримання CSV-файлу з GitHub-репозиторій, створеного іншим учасником команди;

6) об'єднання двох файлів в один за першим стовпчиком;

```
[polyanskiy_mikola@vpsj3IeQ ~]$ join -t, -j1 1 -j1 1 polyanskiy_5_2.csv laba4_5.csv > laba4_end.csv
[polyanskiy_mikola@vpsj3IeQ ~]$ cat laba4_end.csv
Назва,Кількість живонароджених,Кількість прибутих
м.Одеса,542,945
Березівський,43,52
Білгород-Дністровський,140,118
Болградський,106,40
Ізмаїльський,117,10
Подільський,146,58
Роздільнянський,60,38
```

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.

```
[polyanskiy_mikola@vpsj3IeQ ~]$ sort -t, -r -nk2 laba4_end.csv
м.Одеса,542,945
Білгород-Дністровський,140,118
Подільський,146,58
Ізмаїльський,117,10
Болградський,106,40
Роздільнянський,60,38
Березівський,43,52
Назва,Кількість живонароджених,Кількість прибутих
```

Висновок: в ході лабораторії придбали навички роботи з оболонкою Unix-подібних ОС інтерфейсу командного рядка та утилітами обробки текстових даних.

