

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»**

**Виконала:**  
Студентка групи AI-202  
Гребенік Анжеліка Олександрівна

Одеса-2021

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

## **2 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

4) 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/1ITMLdjcw7nboUo9nFw4NwkWcvjbZeZ ZIDVb2Hg4laUQ>

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 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.

## Хід роботи

### 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

*Використовуємо конвеєр з команд `find`, `head` та `grep`*

```
grebenik_anzhelika@vpsj3IeQ:~  
[grebenik_anzhelika@vpsj3IeQ ~]$ find / -name "gre*" 2>/dev/null | head -10 | gr  
ep gre  
/opt/oracle/product/18c/dbhomeXE/md/property_graph/dal/groovy/gremlin-groovy-2.4  
.5.jar  
/opt/oracle/product/18c/dbhomeXE/md/property_graph/dal/groovy/gremlin-opg-rdbms.  
sh  
/opt/oracle/product/18c/dbhomeXE/md/property_graph/dal/groovy/gremlin-opg-rdbms.  
bat  
/opt/oracle/product/18c/dbhomeXE/ctx/data/anl/lang/greek  
/opt/oracle/product/18c/dbhomeXE/perl/lib/5.22.0/User/grent.pm  
/var/spool/mail/grebenik_anzhelika  
/usr/lib/kbd/consolefonts/greek-polytonic.psfu.gz  
/usr/lib/modules/3.10.0-1127.19.1.el7.x86_64/kernel/net/ipv4/gre.ko.xz  
/usr/lib/modules/3.10.0-693.el7.x86_64/kernel/net/ipv4/gre.ko.xz  
/usr/share/man/man1/grep.1.gz  
[grebenik_anzhelika@vpsj3IeQ ~]$
```

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.

*Використовуємо команду `find` та всі помилки відправляємо до `/dev/null`*

```
[grebenik_anzhelika@vpsj3IeQ ~]$ find /home/ -name "g*" 2>/dev/null  
/home/guda_liliya  
/home/goshurenko_volodimir  
/home/golovanchuk_mikola  
/home/grebenik_anzhelika  
/home/grebenik_anzhelika/grebenik_lab_3  
/home/grebenik_anzhelika/grebenik_lab2.pdf  
/home/grebenik_anzhelika/grebenik_lab2.html  
/home/grebenik_anzhelika/.config/libreoffice/4/user/gallery  
/home/grebenik_anzhelika/grebenik_lab2.docx  
/home/grishenko_oleksandra  
[grebenik_anzhelika@vpsj3IeQ ~]$
```

3) Модифікуйте рішення попереднього завдання так, що на екран було виведено імена користувачів, відсортовані в порядку зростання значення, та без дублювання цих значень

*Використовуємо конвеєр з команд `find`, `sort`, `uniq`*

```
[grebenik_anzhelika@vpsj3IeQ ~]$ find /home -name "g*" 2>/dev/null | sort | uniq
/home/golovanchuk_mikola
/home/goshurenko_volodimir
/home/grebenik_anzhelika
/home/grebenik_anzhelika/.config/libreoffice/4/user/gallery
/home/grebenik_anzhelika/grebenik_lab2.docx
/home/grebenik_anzhelika/grebenik_lab2.html
/home/grebenik_anzhelika/grebenik_lab2.pdf
/home/grebenik_anzhelika/grebenik_lab_3
/home/grishenko_oleksandra
/home/guda_liliya
[grebenik_anzhelika@vpsj3IeQ ~]$
```

4) 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.

*З `/etc/passwd` ми беремо 1й, 4й та 6й стовпчик з ім'ям, ідентифікатором та початковим каталогом користувача. Командою `cut` вирізаємо ці стовпчики до файлу `accounts.csv`, командою `sed` змінюємо всі знаки `:` на `,`, Використовуємо конвеєр команд*

```
[grebenik_anzhelika@vpsj3IeQ ~]$ sed 's:/,/g' /etc/passwd | cut -d , -f 1,4,6 >
accounts.csv
[grebenik_anzhelika@vpsj3IeQ ~]$ ls
accounts.csv  grebenik_lab2.docx  Operating-System.-Laboratory-Work-1
dirl          grebenik_lab2.html  os.lab1.cpl251.html
dir2          grebenik_lab2.pdf   os.lab1.utf.html
file1.txt     grebenik_lab_3
[grebenik_anzhelika@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
sshd,74,/var/empty/sshd
soft,1000,/home/soft
sasauth,76,/run/sasauthd
mailnull,47,/var/spool/mqueue
smmsp,51,/var/spool/mqueue
emps,1001,/home/emps
```

---

## 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

<https://docs.google.com/spreadsheets/d/1ITMLdjcw7nboUo9nFw4NwkWcvjbZeZ ZIDVb2Hg4laUQ>

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>

*Використовуючи всі методичні матеріали створюємо цю таблицю згідно номера своєї команди та номера учасника. Зберігаємо цю таблицю у форматі csv*

OS\_lab\_4.pdf - Google

HTML2CSV.pdf - Google

Послання на сер...

Діальність підпр...

Діальність підпр...

OS\_lab\_3.pdf - Google

Мій диск - Google

Кількість найма...

← → ↺ 🏠

docs.google.com/spreadsheets/d/1r16euYsF06MYNRqQ\_oZIs6rA9mvFmV9x6V3P5AhoE1/edit?gid=0

☆ 🌐 📄 🗑️ ⚙️ 🔍

📄

Кількість найманих працівників у суб'єктів господарювання по містах обласного значення та районах у 2019 році

☆ 🌐 📄 🗑️ ⚙️ 🔍

Сохранение...

Настройки Доступа

👤

Файл Правка Вид Вставка Формат Данные Инструменты Дополнения Справка

Подключение...

🔍 100% 123x10

🔍 🔍

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

Використовуючи команду строку та `pscp.exe` копіюємо цей файл до віддаленого сервера.

```
D:\1 курс\Операционные системы>pscp.exe -P 22 Grebenik_4.csv grebenik_anzhelika@91.219.60.189:/home/grebenik_anzhelika
grebenik_anzhelika@91.219.60.189's password:
Grebenik_4.csv | 2 kB | 2.4 kB/s | ETA: 00:00:00 | 100%

D:\1 курс\Операционные системы>
```

Перевіряємо чи є цей файл:

```
grebenik_anzhelika@vpsj3IeQ:~
login as: grebenik_anzhelika
grebenik_anzhelika@91.219.60.189's password:
Last failed login: Tue Mar 23 14:53:28 EDT 2021 from 46.149.53.142 on ssh:notty
There was 1 failed login attempt since the last successful login.
Last login: Tue Mar 23 14:05:07 2021 from 78.26.152.154
[grebenik_anzhelika@vpsj3IeQ ~]$ ls
accounts.csv  Grebenik_4.csv  grebenik_lab_3
dir1          grebenik_lab2.docx  Operating-System.-Laboratory-Work-1
dir2          grebenik_lab2.html  os.lab1.cpl251.html
file1.txt     grebenik_lab2.pdf   os.lab1.utf.html
[grebenik_anzhelika@vpsj3IeQ ~]$
```

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;

Використовуючи команду `sed 's/"/\([0-9]\+\)[,]\([0-9]\+\)"/\1.\2 /g' Grebenik_4.csv > Grebenik_42.csv` відбувається заміна в числах коми на точку з одночасним видаленням лапки

`sed 's/(,\)\...1/1/g' Grebenik_4.csv > Grebenik_42.csv` відбувається заміна підрядка ... 1 на порожнє значення

```
[grebenik_anzhelika@vpsj3IeQ ~]$ sed 's/"/\([0-9]\+\)[,]\([0-9]\+\)"/\1.\2 /g' Grebenik_4.csv > Grebenik_42.csv
[grebenik_anzhelika@vpsj3IeQ ~]$ sed 's/(,\)\...1/1/g' Grebenik_4.csv > Grebenik_42.csv
[grebenik_anzhelika@vpsj3IeQ ~]$ cat Grebenik_42.csv
,місто/район,Усього осіб,осіб підприємства,"у відсотках
до загальної кількості
найманих працівників підприємства",осіб фізичних,"у відсотках
до загальної кількості
найманих працівників"
Одеська область1,,285524,100,57402,100
м.Одеса,місто,206857,172703,60.5,34154,59.5
м.Балта,місто,2067,1686,0.6,381,0.7
м.Білгород-Дністровський,місто,9165,7386,2.6,1779,3.1
м.Біляївка,місто,2542,2232,0.8,310,0.5
м.Ізмаїл,місто,12281,10089,3.5,2192,3.8
м.Чорноморськ,місто,19154,17563,6.1,,2.8
м.Подільськ,місто,2497,1716,0.6,781,1.4
м.Теплодар,місто,804,568,0.2,236,0.4
м.Южне,місто,7898,7229,2.5,669,1.2
Ананьївський,район,1412,867,0.3,545,0.9
Арцизький,район,3293,2685,0.9,608,1.1
Балтський,район,1134,933,0.3,201,0.3
Білгород-Дністровський,район,3768,3323,1.2,445,0.8
```

`sed 's/(,\)\-1/g' Grebenik_4.csv > Grebenik_42.csv` відбувається заміна символу – на порожнє значення

```
Ширяївський,район,1425,,0.4,264,0.5[grebenik_anzhelika@vpsj3IeQ ~]$ sed 's/(,\)\-1/g' Grebenik_4.csv > Grebenik_42.csv
[grebenik_anzhelika@vpsj3IeQ ~]$ cat Grebenik_42.csv
,місто/район,Усього осіб,осіб підприємства,"у відсотках
до загальної кількості
найманих працівників підприємства",осіб фізичних,"у відсотках
до загальної кількості
найманих працівників"
Одеська область1,,285524,100,57402,100
м.Одеса,місто,206857,172703,60.5,34154,59.5
м.Балта,місто,2067,1686,0.6,381,0.7
м.Білгород-Дністровський,місто,9165,7386,2.6,1779,3.1
м.Біляївка,місто,2542,2232,0.8,310,0.5
м.Ізмаїл,місто,12281,10089,3.5,2192,3.8
м.Чорноморськ,місто,19154,17563,6.1,1591,2.8
м.Подільськ,місто,2497,1716,0.6,781,1.4
м.Теплодар,місто,804,568,0.2,236,0.4
м.Южне,місто,7898,7229,2.5,669,1.2
Ананьївський,район,1412,867,0.3,545,0.9
Арцизький,район,3293,2685,0.9,608,1.1
```

`cut -f1,3 -d, Grebenik_4.csv > Grebenik_42.csv` видаляємо з файлу окремі стовпчики з урахуванням символу- роздільника



```
[grebenik_anzhelika@vpsj3IeQ ~]$ cut -f1,3 -d, Grebenik_4.csv > Grebenik_42.csv
[grebenik_anzhelika@vpsj3IeQ ~]$ cat Grebenik_42.csv
,Усього осіб
до загальної кількості
найманих працівників підприємства", "у відсотках
до загальної кількості
найманих працівників"
Одеська область1,
м.Одеса,206857
м.Балта,2067
м.Білгород-Дністровський,9165
м.Біляївка,2542
м.Ізмаїл,12281
м.Чорноморськ,19154
м.Подільськ,2497
м.Теплодар,804
м.Южне,7898
Ананьївський,1412
Арцизький,3293
Балтський,1134
```

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

*Переходимо до Operating-System.-Laboratory-Work-1*

```
[grebenik_anzhelika@vpsj3IeQ ~]$ ls
accounts.csv  Grebenik_42.csv  grebenik_lab_3
dir1          Grebenik_4.csv   Operating-System.-Laboratory-Work-1
dir2          grebenik_lab2.docx os.lab1.cpl251.html
file1.txt     grebenik_lab2.html os.lab1.utf.html
Gr           grebenik_lab2.pdf
[grebenik_anzhelika@vpsj3IeQ ~]$ mv Grebenik_42.csv Operating-System.-Laboratory-Work-1/
```

*Використовуючи команди git add додаємо файл у репозиторій*

```
nothing added to commit but untracked files present (use "git add" to track)
[grebenik_anzhelika@vpsj3IeQ Operating-System.-Laboratory-Work-1]$ git add Grebenik_42.csv
[grebenik_anzhelika@vpsj3IeQ Operating-System.-Laboratory-Work-1]$ git add Grebenik_Angelika.txt
[grebenik_anzhelika@vpsj3IeQ Operating-System.-Laboratory-Work-1]$ git commit -m "lab4 csv"
[main fa562f7] lab4 csv
```

```
[grebenik_anzhelika@vpsj3IeQ Operating-System.-Laboratory-Work-1]$ 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': AngelikaGrebenik
Password for 'https://AngelikaGrebenik@github.com':
Counting objects: 7, done.
Delta compression using up to 2 threads.
Compressing objects: 100% (5/5), done.
Writing objects: 100% (5/5), 1.16 KiB | 0 bytes/s, done.
Total 5 (delta 2), reused 0 (delta 0)
remote: Resolving deltas: 100% (2/2), completed with 1 local object.
To https://github.com/Daniil-Lobko/Operating-System.-Laboratory-Work-1.git
   54848d3..38f18ea  main -> main
[grebenik_anzhelika@vpsj3IeQ Operating-System.-Laboratory-Work-1]$
```

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

```
[grebenik_anzhelika@vpsj3IeQ ~]$ git pull
fatal: No remote repository specified. Please, specify either a URL or a
remote name from which new revisions should be fetched.
[grebenik_anzhelika@vpsj3IeQ ~]$ ls Operating-System.-Laboratory-Work-1
Combined_Grebenik_Lobko.pptx  LobkoDaniil.pdf
Grebenik_42.csv               Lobko_Daniil.txt
Grebenik_Angelika.txt         OSHW_Lobko.pptx
Grebenik_Anzhelika.txt        ГребенікАнжеліка.pdf
Lobko_3lab.pdf                Гребенік_Лабa№3.pdf
Lobko_42.csv                  Обчислювальний годинник шикарда(лаба№1).pptx
[grebenik_anzhelika@vpsj3IeQ ~]$
```

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

*Використовуємо команду join -t, -j1 1 -j2 1 Grebenik\_42.csv Lobko\_42.csv*

```
[grebenik_anzhelika@vpsj3IeQ ~]$ join -t, -j1 1 -j2 1 Grebenik_42.csv Lobko_42.c
sv
join: Grebenik_42.csv: No such file or directory
[grebenik_anzhelika@vpsj3IeQ ~]$ pushd Operating-System.-Laboratory-Work-1
~/Operating-System.-Laboratory-Work-1
[grebenik_anzhelika@vpsj3IeQ Operating-System.-Laboratory-Work-1]$ join -t, -j1
1 -j2 1 Grebenik_42.csv Lobko_42.csv
join: Grebenik_42.csv:4: is not sorted: до загальної кількості
join: Lobko_42.csv:3: is not sorted: м.Валта,1127
Подільський,1290,9490
Окнянський,1401,7076
Любашівський,1500,9927
Миколаївський,899,8906
Овідіопольський,12407,6214
Роздільнянський,2607,14679
Ренійський,2256,6662
Савранський,1383,7229
Саратський,3712,12825
Тарутинський,2681,19682
Татарбунарський,2082,6213
Захарівський,714,7232
Ширяївський,1425,15302
[grebenik_anzhelika@vpsj3IeQ Operating-System.-Laboratory-Work-1]$
```

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.

*Використовуючи конвеєр команд join та sort отримуємо результат.*

```
[grebenik_anzhelika@vpsj3IeQ Operating-System.-Laboratory-Work-1]$ join -t, -j1
1 -j2 1 Lobko_42.csv Grebenik_42.csv | sort -t, -r -nk2
join: Grebenik_42.csv:4: is not sorted: до загальної кількості
join: Lobko_42.csv:3: is not sorted: м.Балта,1127
Овідіопольський,6214,12407
Тарутинський,19682,2681
Ширяївський,15302,1425
Роздільнянський,14679,2607
Саратський,12825,3712
Любашівський,9927,1500
Подільський,9490,1290
Савранський,7229,1383
Окнянський,7076,1401
Ренійський,6662,2256
Татарбунарський,6213,2082
Миколаївський,8906,899
Захарівський,7232,714
[grebenik_anzhelika@vpsj3IeQ Operating-System.-Laboratory-Work-1]$
```

**Висновок:** під час лабораторної роботи ми отримали навички складної обробки текстових даних роботи засобами оболонки Unix-подібних ОС інтерфейсу командного рядка.

