

ОДЕСЬКИЙ НАЦІОНАЛЬНИЙ ПОЛІТЕХНІЧНИЙ УНІВЕРСИТЕТ
ІНСТИТУТ КОМП'ЮТЕРНИХ СИСТЕМ
КАФЕДРА «ІНФОРМАЦІЙНИХ СИСТЕМ»

Лабораторна робота № 3

з дисципліни «Операційні системи»

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

Виконав:

Студент групи АІ-202
Матненко Станіслав Володимирович

Одеса 2020

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

Завдання до лабораторної роботи

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 № 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 участника 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:

1) базове створення файлу CSV-формату, використовуючи Google Docs за методичними рекомендаціями <https://drive.google.com/file/d/17wTiwec-1QybKQje49VdqFahkely1Ou>:

– 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 файла-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

```
[matnenko_stanislav@vpsj3IeQ ~]$ find / -name "mat*" 2> /dev/null |head -10
/var/spool/mail/matnenko_stanislav
/usr/lib/firmware/matrox
/usr/share/man/man8/matchpathcon.8.gz
/usr/share/joe/syntax/matlab.jsf
/usr/share/mime/application/mathematica.xml
/usr/share/mime/application/mathml+xml.xml
/usr/local/webuzo/enduser/webuzo/filemanager/libraries/geshi/geshi/matlab.php
/usr/local/apps/perl/lib/site_perl/5.30.0/x86_64-linux/bits/mathinline.ph
/usr/local/apps/perl/lib/site_perl/5.30.0/x86_64-linux/bits/mathcalls.ph
/usr/local/apps/perl/lib/site_perl/5.30.0/x86_64-linux/bits/mathdef.ph
[matnenko_stanislav@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.

```
[matnenko_stanislav@vpsj3IeQ ~]$ find /home -name "m*" -maxdepth 1 2> /dev/null
/home/muradyan.volodimir
/home/magdenko_mihajlo
/home/markovskij_danilo
/home/maksimenko_andrij
/home/myagkikh_arina
/home/malofeev_denis
/home/matnenko_stanislav
/home/martinyuk_denis
/home/melnichenko_egor
[matnenko_stanislav@vpsj3IeQ ~]$ █
```

- 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).

```
[matnenko_stanislav@vpsj3IeQ ~]$ find /home -name "m*" -maxdepth 1 2> /dev/null | cut -b 7-24 | sort -r | uniq
myagkikh_arina
muradyan.volodimir
melnichenko_egor
matnenko_stanislav
martinyuk_denis
markovskij_danilo
malofeev_denis
maksimenko_andrij
magdenko_mihajlo
[matnenko_stanislav@vpsj3IeQ ~]$ █
```

- 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.

```
[matnenko_stanislav@vpsj3IeQ ~]$ cut -f 1,4,6 -d: /etc/passwd > accounts.csv | 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
saslauthd,76,/run/saslauthd
mailnull,47,/var/spool/mqueue
smmsp,51,/var/spool/mqueue
emps,1001,/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
vmail,5000,/var/local/vmail
arestov_yaroslav,54331,/home/arestov_yaroslav
martinyuk_denis,54332,/home/martinyuk_denis
bogachik_egor,54333,/home/bogachik_egor
eliseev_danilo,54334,/home/eliseev_danilo
bochkarova_karina,54335,/home/bochkarova_karina
tsigankova_anastasiya,54327,/home/tsigankova_anastasiya
vyatkin_danilo,54328,/home/vyatkin_danilo
harkevich_leonid,54329,/home/harkevich_leonid
kalarashan_kostyantin,54336,/home/kalarashan_kostyantin
nosov_andrij,54337,/home/nosov_andrij
kebap_dmitro,54338,/home/kebap_dmitro
prihodko_egor,54339,/home/prihodko_egor
kelembet_evgen,54340,/home/kelembet_evgen
malofeev_denis,54341,/home/malofeev_denis
```

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

```
[matnenko_stanislav@vpsj3IeQ ~]$ grep 5435 accounts.csv
matnenko_stanislav:54350:/home/matnenko_stanislav
bodnar_illya:54351:/home/bodnar_illya
nezhivih_mariya:54352:/home/nezhivih_mariya
grebenik_anzhelika:54353:/home/grebenik_anzhelika
lobko_daniil:54354:/home/lobko_daniil
baranyuk_dmitro:54355:/home/baranyuk_dmitro
kovach_dmitro:54356:/home/kovach_dmitro
sidelnikov_mikita:54357:/home/sidelnikov_mikita
lukashak_daniil:54358:/home/lukashak_daniil
markovskij_danilo:54359:/home/markovskij_danilo
[matnenko_stanislav@vpsj3IeQ ~]$ █
```

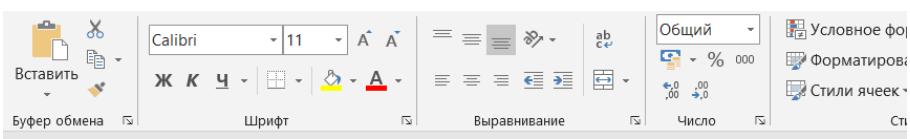
2.2 Обробка текстового файлу формату CSV

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.



A	B	C	D	E	F	G	H	I	J
7	Арцизький район	3950	1033	–					
8	Балтський район	–	к	–					
9	Білгород-, район	к		8900	–				
10	Біляївський район	143	1406	–					
11	Березівський район	552	4436	–					
12	Болградський район	5956	1076	–					
13	Великомихайлівський район	42	1052	–					
14	Іванівський район	1812	к	–					
15	Ізмаїльський район	к	к	к					
16	Кілійський район	71	598	к					
17	Кодимський район	–	к	–					
18	Лиманський район	к	–	к					
19	Подільський район	к	к	–					
20	Окнянський район	–	–	–					
21	Любашівський район	к	–	–					
22	Миколаївський район	–	–	–					
23	Овідіопольський район	к	к	–					

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

```
D:\Users\ThinkPad\Desktop\Универ\Лабы ОП>pscp.exe -P 22 matnenko_2_2.csv matnenko_stanislav@91.219.60.189:/home/matnenko_stanislav
matnenko_stanislav@91.219.60.189's password:
matnenko_2_2.csv          | 1 kB |   1.0 kB/s | ETA: 00:00:00 | 100%
D:\Users\ThinkPad\Desktop\Универ\Лабы ОП>
```

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;

```
[matnenko_stanislav@vpsj3IeQ ~]$ sed 's/\"([0-9]\+\)[,]\([0-9]\+\)\"/\1.\2 /g' matnenko1.csv > matnenko2.csv
[matnenko_stanislav@vpsj3IeQ ~]$ sed 's/(,)\...1/\1/g' matnenko2.csv > matnenko3.csv
[matnenko_stanislav@vpsj3IeQ ~]$ sed 's/(,)-/\1/g' matnenko3.csv > matnenko4.csv
[matnenko_stanislav@vpsj3IeQ ~]$ cut -f1,3 -d, matnenko4.csv > matnenko5.csv
[matnenko_stanislav@vpsj3IeQ ~]$
```

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

```
[matnenko_stanislav@vpsj3IeQ ~]$ git add matnenko5.csv
[matnenko_stanislav@vpsj3IeQ ~]$ git commit -m '4.2.4'
[master (root-commit) 3c455e9] 4.2.4
 1 file changed, 31 insertions(+)
 create mode 100644 matnenko5.csv
[matnenko_stanislav@vpsj3IeQ ~]$ git push
Username for 'https://github.com': ThaumBear
Password for 'https://ThaumBear@github.com':
Counting objects: 3, done.
Delta compression using up to 2 threads.
Compressing objects: 100% (2/2), done.
Writing objects: 100% (3/3), 625 bytes | 0 bytes/s, done.
Total 3 (delta 0), reused 0 (delta 0)
To https://github.com/ThaumBear/Operating-System-Laboratory-Work-4.git
```

 ThaumBear 4.2.4	3c455e9 5 minutes ago	 1 commit
 matnenko5.csv	4.2.4	5 minutes ago

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

```
[matnenko_stanislav@vpsj3IeQ ~]$ git clone https://github.com/ThaumBear/Operating-System-Laboratory-Work-4
Cloning into 'Operating-System-Laboratory-Work-4'...
remote: Enumerating objects: 6, done.
remote: Counting objects: 100% (6/6), done.
remote: Compressing objects: 100% (4/4), done.
remote: Total 6 (delta 0), reused 6 (delta 0), pack-reused 0
Unpacking objects: 100% (6/6), done.
```

Копіювання файлу другого учасника команди до домашнього репозиторію:

```
[matnenko_stanislav@vpsj3IeQ Operating-System-Laboratory-Work-4]$ cp Borovets5.csv /home/matnenko_stanislav
```

```
[matnenko_stanislav@vpsj3IeQ Operating-System-Laboratory-Work-4]$ cd -
/home/matnenko_stanislav
[matnenko_stanislav@vpsj3IeQ ~]$ ls
accounts.cs           Lab2_html_79f37802d21ca400.png   Lab2_html_f23085cb6fd11072.png
accounts.csv          Lab2_html_82c25e6e07b4d4a3.png   Lab2_html_f8483ac3ebe4b9e1.png
Borovets5.csv         Lab2_html_88327d35b7579b10.png  Lab2.pdf
.git                  Lab2_html_89c5a2b950de484.png  matnenko
Lab2.docx             Lab2_html_8b5fdd150ced5f94.png  matnenko1.csv
Lab2.html              Lab2_html_8d9004669466bfcd.png  matnenko_2_2.csv
Lab2.html_2111992e4828fa27.png Lab2_html_93220d88e129bc55.png  matnenko2.csv
Lab2.html_2845477d4c2d458b.png Lab2_html_9e055c3e0ad51988.png  matnenko3.csv
Lab2.html_2b2166f69121ca5b.png Lab2_html_9fe3fb0a664ecc5.png  matnenko4.csv
Lab2.html_33e85aa5b39fbff8.png Lab2_html_b4311fa5a62c767.png  matnenko5.csv
Lab2.html_3b12d34fec7e9c63.png Lab2_html_b629b6e7d4d43137.png  matnenko_lab_3
Lab2.html_472a46cc16d8c2fc.png Lab2_html_d1ce501624b27830.png  Operating-System-Laboratory-Work-4
Lab2.html_574df2bcce7e978.png Lab2_html_d39a289e664b5a0a.png  os.lab1.cp1251.html
Lab2.html_69ef761206c65cd5.png Lab2_html_d3aef4fcf65aff1.png  os.lab1.utf.html
Lab2.html_73e159d1c181d4e.png Lab2_html_e5e55e3e01266105.png
```

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

```
[matnenko_stanislav@vpsj3IeQ ~]$ join -t, -j1 1 -j2 1 Borovets5.csv matnenko5.csv
join: Borovets5.csv:3: is not sorted: м.Одеса,к
join: matnenko5.csv:3: is not sorted: м.Балта,к
Одеська область,827957.23,20258
райони.,
Ананьївський,22598.17,к
Арцизький,47539.87,3950
Валтський,7488.43,
Білгород-Дністровський,48609.2,к
Біляївський,25090.64,143
Березівський,36208.85,552
Болградський,39790.18,5956
Великомихайлівський,42379.5,052
Іванівський,35693.09,1812
Ізмаїльський,39169.18,к
Кілійський,19834.94,71
Кодимський,12720.82,
Лиманський,56547.96,к
Подільський,19982.96,к
Окнянський,28357.75,
Любашівський,20580.24,к
Миколаївський,16773.68,
Овідіопольський,21584.4,к
Роздільнянський,36756.97,к
Ренійський,.55,
Савранський,8426.09,
Саратський,65448.89,758
Тарутинський,57929.93,3917
Татарбунарський,31297.55,к
Захарівський,13285.09,к
Ширяївський,37478.73,701
[matnenko stanislav@vpsj3IeQ ~]$ █
```

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 участника команди, або за номером 3 – для 2-го участника команди.

```
[matnenko_stanislav@vpsj3IeQ ~]$ join -t, -j1 1 -j2 1 Borovets5.csv matnenko5.csv | sort -r -nk3 -t,
join: Borovets5.csv:3: is not sorted: м.Одеса,к
join: matnenko5.csv:3: is not sorted: м.Балта,к
Одеська область,827957.23,20258
Болградський,39790.18,5956
Арцизький,47539.87,3950
Тарутинський,57929.93,3917
Іванівський,35693.09,1812
Саратський,65448.89,758
Ширяївський,37478.73,701
Березівський,36208.85,552
Біляївський,25090.64,143
Кілійський,19834.94,71
Великомихайлівський,42379.5,052
Татарбунарський,31297.55,к
Савранський,8426.09,
Роздільнянський,36756.97,к
Ренійський,.55,
райони.,
Подільський,19982.96,к
Окнянський,28357.75,
Овідіопольський,21584.4,к
Миколаївський,16773.68,
Любашівський,20580.24,к
Лиманський,56547.96,к
Кодимський,12720.82,
Ізмаїльський,39169.18,к
Захарівський,13285.09,к
Білгород-Дністровський,48609.2,к
Балтський,7488.43,
Ананьївський,22598.17,к
[matnenko stanislav@vpsj3IeQ ~]$ █
```

Висновок. В процесі виконання роботи ми опанували методи методи складного редактування даних в оболонці Unix. Найскладнішим вважаю пункт 4 завдання 2.2.