

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:
студент групи АІ-204
Бериславський В.Р.

Перевірив:
Блажко О. А.

Одеса – 2021

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

Хід роботи:

1) Обробка текстових даних результатів роботи команд

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

```
berislavskij_vladislav@vpsj3IeQ:~  
login as: berislavskij_vladislav  
berislavskij_vladislav@91.219.60.189's password:  
Last login: Sun Mar 14 17:31:46 2021 from 188.163.100.178  
[berislavskij_vladislav@vpsj3IeQ ~]$ find / -name "ber*" 2> /dev/null | head -n 10  
/var/spool/mail/berislavskij_vladislav  
/usr/share/locale/ber  
/usr/share/i18n/locales/ber_MA  
/usr/share/i18n/locales/ber_DZ  
/usr/local/apps/share/man/man3/ber_free.3  
/usr/local/apps/share/man/man3/ber_get_stringa.3  
/usr/local/apps/share/man/man3/ber_start_set.3  
/usr/local/apps/share/man/man3/ber_get_next.3  
/usr/local/apps/share/man/man3/ber_printf.3  
/usr/local/apps/share/man/man3/ber_bvstr.3  
[berislavskij_vladislav@vpsj3IeQ ~]$
```

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

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

```
berislavskij_vladislav@vpsj3IeQ:~  
[berislavskij_vladislav@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  
[berislavskij_vladislav@vpsj3IeQ ~]$
```

3. Модифікують рішення попереднього завдання так, щоб на екран було виведено імена користувачів, відсортовані в порядку зростання значення, та без дублювання цих значень (використовують конвеєр з команд cut, sort, uniq).

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

```
berislavskij_vladislav@vpsj3IeQ:~  
[berislavskij_vladislav@vpsj3IeQ ~]$ ls  
berislavskij_lab_3      Berislavskij_Vladislav.pdf  house      Operating-System.-Laboratory-Work-1  os.lab1.utf.  
Berislavskij_Vladislav.docx  flat.txt                  magazin    os.lab1.cpi251.html                wall.txt  
[berislavskij_vladislav@vpsj3IeQ ~]$ cut -d':' -f1,3,6 /etc/passwd | tr ':' ' ' > accounts.csv  
[berislavskij_vladislav@vpsj3IeQ ~]$ less 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,/var/lib/nfs  
systemd-network,192,/var/lib/nfs  
dbus,81,/var/lib/dbus  
polkitd,999,/var/lib/polkitd  
postfix,89,/var/spool/postfix  
chrony,998,/var/lib/chrony  
sshd,74,/var/empty/sshd  
soft,1000,/home/soft  
sasauth,997,/run/sasauthd  
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
```

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

```
berislavskij_vladislav@vpsj3IeQ:~  
[berislavskij_vladislav@vpsj3IeQ ~]$ grep berislavskij accounts.csv  
berislavskij_vladislav,54386,/home/berislavskij_vladislav  
[berislavskij_vladislav@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  
[berislavskij_vladislav@vpsj3IeQ ~]$
```

2. Обробка текстового файлу формату CSV

1. Створюють 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.

```
berislavskij_vladislav – Блокнот
Файл Правка Формат Вид Справка
Назва,Місто/Район,Забрана площа на зерно та зелений корм - га,Забрана площа - % від 2020,Забрана площа н
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
Стр 1, стлб 1 100% Windows (CRLF) UTF-8
```

2. Копіюють файл на сервер за адресою з IP = 91.219.60.189 у домашній каталог

```
Командная строка
Microsoft Windows [Version 10.0.19041.867]
(c) Корпорация Майкрософт (Microsoft Corporation), 2020. Все права защищены.

C:\Users\Lenovo\Desktop\Labs\OS\pscp.exe -P 22 Desktop\Labs\OS\4\berislavskij_vladislav.csv berislavskij_vladislav@91.219.60.189:/home/berislavskij_vladislav
berislavskij_vladislav@91.219.60.189's password:
berislavskij_vladislav.cs | 3 kB | 3.1 kB/s | ETA: 00:00:00 | 100%

C:\Users\Lenovo>
```

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

```
berislavskij_vladislav@vpsj3ieQ:~  
[berislavskij_vladislav@vpsj3ieQ ~]$ sed 's/\\"([0-9]\\+)[,]\"([0-9]\\+)"\"/\\1\\.\\2 /g' | sed 's/\\(,\\)\\.\\.\\.1/\\1/g' ^  
| sed 's/\\(,\\)-/\\1/g' | cut -d, -f1,2 berislavskij_vladislav.csv  
Назва,Місто/Район  
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
```

4. Передача створеного CSV-файлу в GitHub-репозиторій

```
berislavskij_vladislav@vpsj3IeQ ~]$ mc

1)+ Stopped /usr/bin/mc -P "$MC_PWD_FILE" "$@"
berislavskij_vladislav@vpsj3IeQ ~]$ pwd
home/berislavskij_vladislav
berislavskij_vladislav@vpsj3IeQ ~]$ pushd /Operating-System.-Laboratory-Works
bash: pushd: /Operating-System.-Laboratory-Works: No such file or directory
berislavskij_vladislav@vpsj3IeQ ~]$ pushd Operating-System.-Laboratory-Works
/Operating-System.-Laboratory-Works ~
berislavskij_vladislav@vpsj3IeQ Operating-System.-Laboratory-Works]$ ls
berislavskij_vladislav.csv Бериславський Владислав.pdf Бериславський Владислав Лаборвторна 3.pdf
berislavskij_vladislav@vpsj3IeQ Operating-System.-Laboratory-Works]$ git add berislavskij_vladislav.csv
berislavskij_vladislav@vpsj3IeQ Operating-System.-Laboratory-Works]$ git commit -m AddingTheFile
main 7597cb9] AddingTheFile
1 file changed, 32 insertions(+)
create mode 100644 berislavskij_vladislav.csv
berislavskij_vladislav@vpsj3IeQ Operating-System.-Laboratory-Works]$ git push
arning: push.default is unset; its implicit value is changing in
it 2.0 from 'matching' to 'simple'. To squelch this message
nd maintain the current behavior after the default changes, use:

    git config --global push.default matching

or 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)

sername for 'https://github.com': Ratchet34702
assword for 'https://Ratchet34702@github.com':
ounting objects: 4, done.
elta compression using up to 2 threads.
ompressing objects: 100% (3/3), done.
riting objects: 100% (3/3), 1.59 KiB | 0 bytes/s, done.
otal 3 (delta 1), reused 0 (delta 0)
emote: Resolving deltas: 100% (1/1), completed with 1 local object.
o https://github.com/Ratchet34702/Operating-System.-Laboratory-Works
688477f..7597cb9 main -> main
berislavskij_vladislav@vpsj3IeQ Operating-System.-Laboratory-Works]$
```

5. отримання CSV-файлу, створеного іншим учасником команди

```
berislavskij_vladislav@vpsj3IeQ:~
[berislavskij_vladislav@vpsj3IeQ ~]$ ls
accounts.csv          magazin
berislavskij_lab_3    Operating-System.-Laboratory-Work-1
berislavskij_vladislav.csv Operating-System.-Laboratory-Works
Berislavskij_Vladislav.docx os.lab1.cp1251.html
Berislavskij_Vladislav.pdf os.lab1.utf.html
flat.txt              wall.txt
house                 work.html
[berislavskij_vladislav@vpsj3IeQ ~]$ pushd Operating-System.-Laboratory-Works
~/Operating-System.-Laboratory-Works ~
[berislavskij_vladislav@vpsj3IeQ Operating-System.-Laboratory-Works]$ ls
berislavskij_vladislav.csv Бериславський Владислав Лаборвторна 3.pdf
Бериславський Владислав.pdf
[berislavskij_vladislav@vpsj3IeQ Operating-System.-Laboratory-Works]$ git pull
remote: Enumerating objects: 4, done.
remote: Counting objects: 100% (4/4), done.
remote: Compressing objects: 100% (2/2), done.
remote: Total 3 (delta 1), reused 3 (delta 1), pack-reused 0
Unpacking objects: 100% (3/3), done.
From https://github.com/Ratchet34702/Operating-System.-Laboratory-Works
   7597cb9..69alc20 main      -> origin/main
Updating 7597cb9..69alc20
Fast-forward
 test.csv | 31 +++++
 1 file changed, 31 insertions(+)
 create mode 100644 test.csv
[berislavskij_vladislav@vpsj3IeQ Operating-System.-Laboratory-Works]$ ls
berislavskij_vladislav.csv Бериславський Владислав.pdf
test.csv                  Бериславський Владислав Лаборвторна 3.pdf
```

```
[berislavskij_vladislav@vpsj3IeQ Operating-System.-Laboratory-Works]$ cp test.csv /home/berislavskij_vladislav
[berislavskij_vladislav@vpsj3IeQ Operating-System.-Laboratory-Works]$ popd
~
[berislavskij_vladislav@vpsj3IeQ ~]$ ls
accounts.csv          Berislavskij_Vladislav.docx  house                  Operating-System.-Laboratory-Works  test.csv
berislavskij_lab_3    Berislavskij_Vladislav.pdf  magazin                os.lab1.cpl251.html               wall.txt
berislavskij_vladislav.csv  flat.txt                    Operating-System.-Laboratory-Work-1  os.lab1.utf.html                  work.html
```

6. Об'єднання двох файлів в один за першим стовпчиком

```
berislavskij_vladislav@vpsj3IeQ:~
[berislavskij_vladislav@vpsj3IeQ ~]$ join -t, -a 1 -a 2 -o0,1,2,2.2 -e '' berislavskij_vladislav.csv test.csv
Назва,Місто/Район,місто/район
м.Балта,,місто
м.Білгород-Дністровський,,місто
м.Біляївка,,місто
м.Ізмаїл,,місто
join: test.csv:6: is not sorted: Ананьївський,район
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
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
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,район
Окнянський,район,район
Любашівський,район,район
Миколаївський,район,район
Овідіопольський,район,район
Роздільнянський,район,район
Ренійський,район,район
Савранський,район,район
Саратський,район,район
Тарутинський,район,район
```

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

```
berislavskij_vladislav@vpsj3leQ:~  
[berislavskij_vladislav@vpsj3leQ ~]$ sort --field-separator="," --key=2 -r berislavskij_vladislav.csv test.csv  
Татарбунарський, район  
Тарутинський, район  
Саратський, район  
Савранський, район  
Роздільнянський, район  
Ренійський, район  
Подільський, район  
Окнянський, район  
Овідіопольський, район  
Миколаївський, район  
Любашівський, район  
Лиманський, район  
Кодимський, район  
Кілійський, район  
Ізмаїльський, район  
Іванівський, район  
Захарівський, район  
Великомихайлівський, район  
Болградський, район  
Біляївський, район  
Білгород-Дністровський, район  
Березівський, район  
Балтський, район  
Арцизький, район  
Ананьївський, район  
Ширяївський, район  
Ширяївський, район  
Татарбунарський, район  
Тарутинський, район  
Саратський, район  
Савранський, район  
Роздільнянський, район  
Ренійський, район  
Подільський, район  
Окнянський, район  
Овідіопольський, район  
Миколаївський, район  
Любашівський, район  
Лиманський, район  
Кодимський, район  
Кілійський, район  
Ізмаїльський, район  
Іванівський, район  
Захарівський, район  
Великомихайлівський, район  
Болградський, район  
Біляївський, район  
Білгород-Дністровський, район
```

Висновки: в ході виконання лабораторної роботи, були придбані навички складної обробки текстових даних роботи засобами Unix. Найскладнішим завданням було об'єднання csv файлів.