

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»

Виконала:
ст.гр. АІ -204
Сіренко Марія

Перевірили:
Блажко О.А.

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

Завдання :

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. 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>:

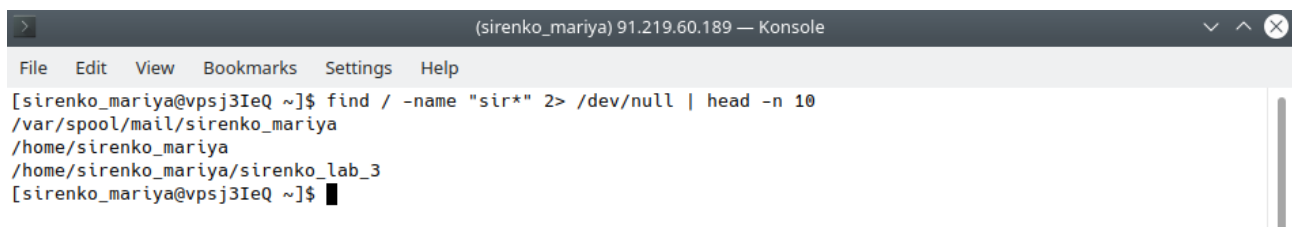
- кодування символів – UTF-8
 - роздільник між колонками – символ коми;
 - назви файлів визначити з урахуванням транслітерації свого прізвища та номеру файлу.
- 2) копіювання файлів на сервер за адресою з IP = 91.219.60.189 у домашній каталог учасника команди;
- 3) модифікація файлу CSV-формату:
- заміна 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.

Виконання завдань:

Обробка текстових даних результатів роботи команд

1) Виведено перші десять назв каталогів або файлів, назви яких починаються

з перших трьох букв вашого “sir”, команда `find / -name "sir*" 2> /dev/null | head -n 10`



```
(sirenko_mariya) 91.219.60.189 — Konsole
File Edit View Bookmarks Settings Help
[sirenko_mariya@vpsj3IeQ ~]$ find / -name "sir*" 2> /dev/null | head -n 10
/var/spool/mail/sirenko_mariya
/home/sirenko_mariya
/home/sirenko_mariya/sirenko_lab_3
[sirenko_mariya@vpsj3IeQ ~]$
```

2) Виведіено на екран список назв домашніх каталогів користувачів, в назві яких присутня буква s ,команда `find/ home -type d -maxdepth 1 -name "s*" 2> /dev/null`



```
(sirenko_mariya) 91.219.60.189 — Konsole
File Edit View Bookmarks Settings Help
[sirenko_mariya@vpsj3IeQ ~]$ find /home -type d -maxdepth 1 -name "s*" 2> /dev/null
/home/sherbakov_georgij
/home/sherbakov_artur
/home/strinaglyuk_ivan
/home/sinyakov_igor
/home/sultanov_artem
/home/sokolov_oleksij
/home/shostak_roman
/home/sivash_andrij
/home/stepanenko_gleb
/home/sirenko_mariya
/home/sidelnikov_mikita
/home/svetashov_danilo
/home/shapovalova_viktoriya
/home/soft-1
/home/shulyak_mikola
[sirenko_mariya@vpsj3IeQ ~]$
```

3) Введено імена користувачів, відсортовані в порядку зростання значення, та без дублювання цих значень ,команда `find/ home -type d -maxdepth 1 -name "s*" 2> /dev/null | cut -d '_' -f2 |sort|uniq`

```
(sirenko_mariya) 91.219.60.189 — Konsole
File Edit View Bookmarks Settings Help
[sirenko_mariya@vpsj3IeQ ~]$ find /home -type d -maxdepth 1 -name "s*" 2> /dev/null | cut -d '_' -f2 |sort|uniq
andrij
artem
artur
danilo
georgij
gleb
/home/soft-1
igor
ivan
mariya
mikita
mikola
oleksij
roman
viktoriya
[sirenko_mariya@vpsj3IeQ ~]$
```

4) Створено 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,

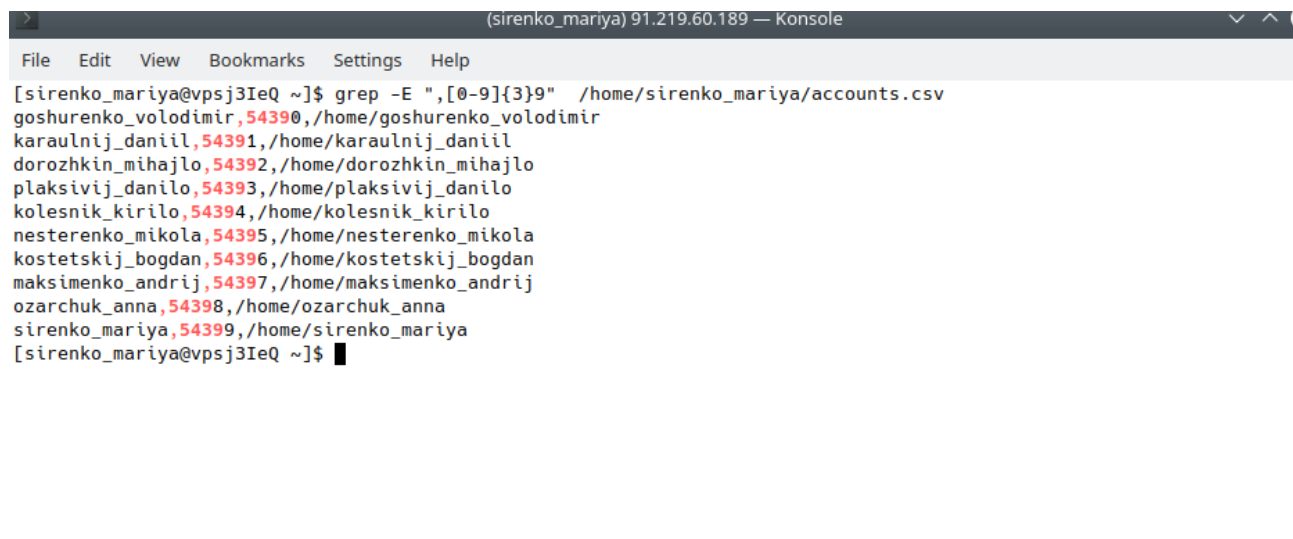
команда `cut -d ':' -f1,3,6 /etc/passwd |tr ':' ','>accounts.csv`

```
(sirenko_mariya) 91.219.60.189 — Konsole
File Edit View Bookmarks Settings Help
[sirenko_mariya@vpsj3IeQ ~]$ cut -d ':' -f1,3,6 /etc/passwd | tr ':' ',' > accounts.csv
[sirenko_mariya@vpsj3IeQ ~]$ ls
accounts.csv  lab_2.pdf          os.lab1.cp1251.html  sirenko_lab_3
lab_2.odt     Operating-System.Laboratory-Work-1  os.lab1.utf.html    test.html
[sirenko_mariya@vpsj3IeQ ~]$ cat account.csv
cat: account.csv: No such file or directory
[sirenko_mariya@vpsj3IeQ ~]$
```

Зміст строреного файлу *accounts.csv*

```
exim,93,/var/spool/exim
vmail,5000,/var/local/vmail
arestov_yaroslav,54322,/home/arestov_yaroslav
martinyuk_denis,54323,/home/martinyuk_denis
bogachik_egor,54324,/home/bogachik_egor
eliseev_danilo,54325,/home/eliseev_danilo
bochkarova_karina,54326,/home/bochkarova_karina
tsigankova_anastasiya,54327,/home/tsigankova_anastasiya
vyatkin_danilo,54328,/home/vyatkin_danilo
harkevich_leonid,54329,/home/harkevich_leonid
kalarashan_kostyantyn,54330,/home/kalarashan_kostyantyn
nosov_andrij,54331,/home/nosov_andrij
kebab_dmitro,54332,/home/kebab_dmitro
prihodko_egor,54333,/home/prihodko_egor
kelembet_evgen,54334,/home/kelembet_evgen
malofeev_denis,54335,/home/malofeev_denis
tarasov_daniil,54336,/home/tarasov_daniil
stepanenko_gleb,54337,/home/stepanenko_gleb
strinaglyuk_ivan,54338,/home/strinaglyuk_ivan
yufitenko_andrij,54339,/home/yufitenko_andrij
cheremnih_griqorij,54340,/home/cheremnih_griqorij
```

5) З файлу *accounts.csv* отримано на екран перелік облікових записів, в яких четверта цифра ідентифікатора користувача співпадає з четвертою цифрою власного (54399), команда *grep -E "[0-9]{3}9" /home/sirenko_mariya/accounts.csv*

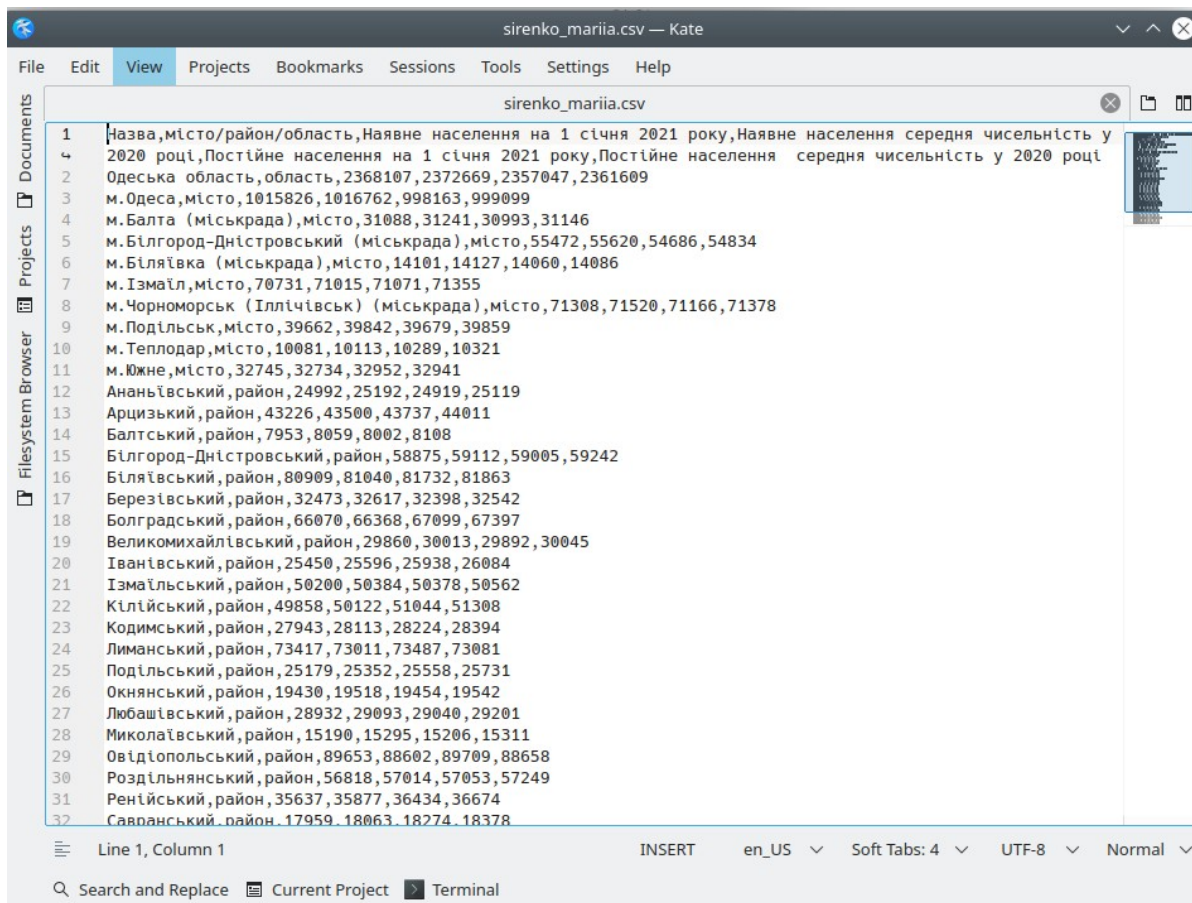


The screenshot shows a terminal window titled "(sirenko_mariya) 91.219.60.189 — Konsole". The terminal has a menu bar with "File", "Edit", "View", "Bookmarks", "Settings", and "Help". The command prompt is "[sirenko_mariya@vpsj3IeQ ~]\$". The command entered is "grep -E "[0-9]{3}9" /home/sirenko_mariya/accounts.csv". The output lists several users with their IDs, where the fourth digit of the ID is highlighted in red: "goshurenko_volodimir,54390,/home/goshurenko_volodimir", "karaulnij_daniil,54391,/home/karaulnij_daniil", "dorozhkin_mihajlo,54392,/home/dorozhkin_mihajlo", "plaksivij_danilo,54393,/home/plaksivij_danilo", "kolesnik_kirilo,54394,/home/kolesnik_kirilo", "nesterenko_mikola,54395,/home/nesterenko_mikola", "kostetskij_bogdan,54396,/home/kostetskij_bogdan", "maksimenko_andrij,54397,/home/maksimenko_andrij", "ozarchuk_anna,54398,/home/ozarchuk_anna", and "sirenko_mariya,54399,/home/sirenko_mariya". The prompt returns to "[sirenko_mariya@vpsj3IeQ ~]\$".

Обробка текстового файлу формату CSV

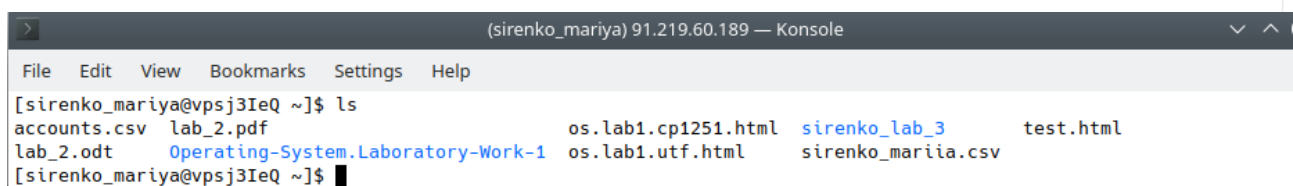
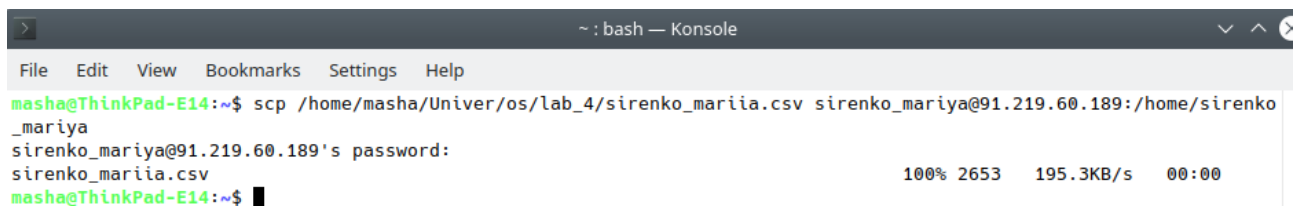
1) Створено файлу CSV-формату, використовуючи Google Docs за

методичними рекомендаціями



2) Скопійовано файл на сервер за адресою з IP = 91.219.60.189 у домашній каталог , команда

```
scp /home/masha/Univer/os /lab_4 sirenko_mariia.csv sirenko_mariya@91.219.60.189: /home/  
sirenko_mariya
```



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 ,

команди `sed 's/\("[0-9]\+\)/[,]/([0-9]\+)\)/"^\1\.\ 2 g' sirenko_mariia.csv | sed 's/(,)\ ...1/1/g' | sed 's/(,)\-/1/g' | cut -d , -f1,2 > sirenko_mariia_2.csv`

```
(sirenko_mariya) 91.219.60.189 — Konsole
File Edit View Bookmarks Settings Help
[sirenko_mariya@vpsj3IeQ ~]$ sed 's/\("[0-9]\+\)/[,]/([0-9]\+)\)/"^\1\.\ 2 g' sirenko_mariia.csv | sed 's/(,)\ ...1/1/g' |
sed 's/(,)\-/1/g' | cut -d , -f1,2 > sirenko_mariia_2.csv
[sirenko_mariya@vpsj3IeQ ~]$ ls
accounts.csv  lab_2.pdf  os.lab1.cp1251.html  sirenko_lab_3  sirenko_mariia.csv
lab_2.odt     Operating-System.Laboratory-Work-1  os.lab1.utf.html  sirenko_mariia_2.csv  test.html
[sirenko_mariya@vpsj3IeQ ~]$ cat sirenko_mariia_2.csv
Назва,місто/район/область
Одеська область,область
м.Одеса,місто
м.Балта (міськрада),місто
м.Білгород-Дністровський (міськрада),місто
м.Біляївка (міськрада),місто
м.Ізмаїл,місто
м.Чорноморськ (Іллічівськ) (міськрада),місто
м.Подільськ,місто
м.Теплодар,місто
м.Южне,місто
Ананьївський,район
```

4) передача створеного CSV-файлу в GitHub-репозиторій

```
(sirenko_mariya) 91.219.60.189 — Konsole
File Edit View Bookmarks Settings Help
ariya/Operating-System.Laboratory-Work-1
[sirenko_mariya@vpsj3IeQ Operating-System.Laboratory-Work-1]$ ls
Ozarchuk_Anna.txt  Sirenko_Mariia.pdf  Sirenko_Mariia.txt  Машина_Перейри_2_3.pptx  ОС лаба2.pdf
sirenko_mariia_2.csv  SirenkoMariia.pdf  Машина_Перейри_1_4.pptx  Машина_Перейри_full.pptx  ОС лаба 3.pdf
[sirenko_mariya@vpsj3IeQ Operating-System.Laboratory-Work-1]$ git add sirenko_mariia_2.csv
[sirenko_mariya@vpsj3IeQ Operating-System.Laboratory-Work-1]$ git commit -m 'Завдання Лаб.родоти №4 2.2.4'
[main 20405de] Завдання Лаб.родоти №4 2.2.4
1 file changed, 37 insertions(+)
create mode 100644 sirenko_mariia_2.csv
[sirenko_mariya@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': mariiasirenko2
Password for 'https://mariiasirenko2@github.com':
Counting objects: 4, done.
Delta compression using up to 2 threads.
Compressing objects: 100% (3/3), done.
Writing objects: 100% (3/3), 721 bytes | 0 bytes/s, done.
Total 3 (delta 1), reused 0 (delta 0)
remote: Resolving deltas: 100% (1/1), completed with 1 local object.
To https://github.com/mariiasirenko2/Operating-System.Laboratory-Work-1.git
fd85467..20405de main -> main
[sirenko_mariya@vpsj3IeQ Operating-System.Laboratory-Work-1]$
```


5) отримання CSV-файлу, створеного іншим учасником команди, команда *git pull*

```
(sirenko_mariya) 91.219.60.189 — Konsole
File Edit View Bookmarks Settings Help
[sirenko_mariya@vpsj3IeQ Operating-System.Laboratory-Work-1]$ ls
Ozarchuk_Anna.txt      Sirenko_Mariia.pdf  Sirenko_Mariia.txt      Машина_Перейри_2_3.pptx  ОС лаба2.pdf
sirenko_mariia_2.csv  SirenkoMariia.pdf  Машина_Перейри_1_4.pptx  Машина_Перейри_full.pptx  ОС лаба 3.pdf
[sirenko_mariya@vpsj3IeQ Operating-System.Laboratory-Work-1]$ 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/mariiasirenko2/Operating-System.Laboratory-Work-1
 20405de..a7bd805  main    -> origin/main
Updating 20405de..a7bd805
Fast-forward
 1.csv | 36 ++++++
 1 file changed, 36 insertions(+)
 create mode 100644 1.csv
[sirenko_mariya@vpsj3IeQ Operating-System.Laboratory-Work-1]$ ls
1.csv      Sirenko_Mariia.pdf  Машина_Перейри_1_4.pptx  ОС лаба2.pdf
Ozarchuk_Anna.txt  SirenkoMariia.pdf  Машина_Перейри_2_3.pptx  ОС лаба 3.pdf
sirenko_mariia_2.csv  Sirenko_Mariia.txt  Машина_Перейри_full.pptx
[sirenko_mariya@vpsj3IeQ Operating-System.Laboratory-Work-1]$
```

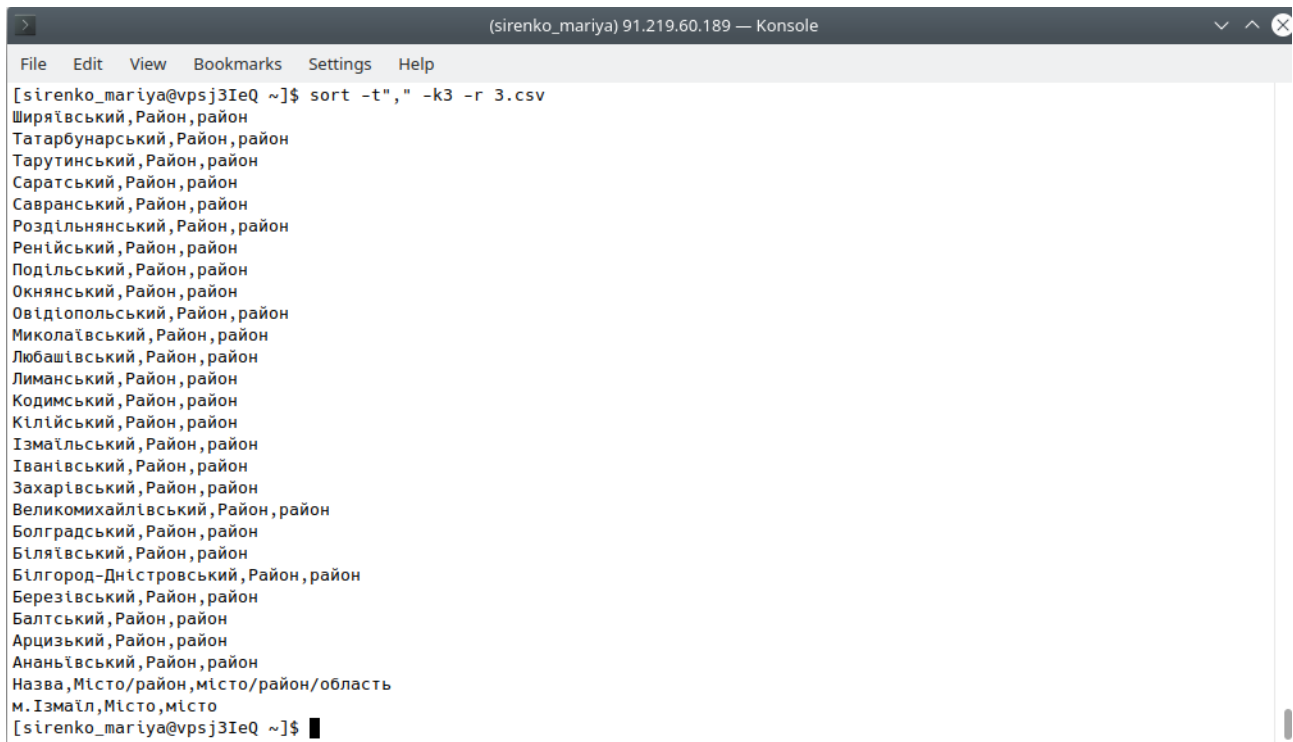
6) об'єднання двох файлів в один за першим стовпчиком ,команда

join -t , <(sort -k1 1.csv) <(sort -k1 sirenko_mariia_2.csv) >/home/sirenko_mariya /3.csv

```
(sirenko_mariya) 91.219.60.189 — Konsole
File Edit View Bookmarks Settings Help
[sirenko_mariya@vpsj3IeQ Operating-System.Laboratory-Work-1]$ join -t , <(sort -k1 1.csv) <(sort -k1 sirenko_mariia_2.csv)
> /home/sirenko_mariya/3.csv
[sirenko_mariya@vpsj3IeQ Operating-System.Laboratory-Work-1]$ pushd
-bash: pushd: no other directory
[sirenko_mariya@vpsj3IeQ Operating-System.Laboratory-Work-1]$ pushd ~
~ ~/Operating-System.Laboratory-Work-1
[sirenko_mariya@vpsj3IeQ ~]$ ls
3.csv      lab_2.odt  Operating-System.Laboratory-Work-1  os.lab1.utf.html  sirenko_mariia_2.csv  test.html
accounts.csv  lab_2.pdf  os.lab1.cp1251.html                sirenko_lab_3     sirenko_mariia.csv
[sirenko_mariya@vpsj3IeQ ~]$
```

```
(sirenko_mariya) 91.219.60.189 — Konsole
File Edit View Bookmarks Settings Help
[sirenko_mariya@vpsj3IeQ ~]$ ls
3.csv      lab_2.odt  Operating-System.Laboratory-Work-1  os.lab1.utf.html  sirenko_mariia_2.csv  test.html
accounts.csv  lab_2.pdf  os.lab1.cp1251.html                sirenko_lab_3     sirenko_mariia.csv
[sirenko_mariya@vpsj3IeQ ~]$ cat 3.csv
Ананівський, Район, район
Арцизький, Район, район
Балтський, Район, район
Березівський, Район, район
Білгород-Дністровський, Район, район
Біляївський, Район, район
Болградський, Район, район
Великомихайлівський, Район, район
Захарівський, Район, район
Іванівський, Район, район
Ізмаїльський, Район, район
Кілійський, Район, район
Кодимський, Район, район
Лиманський, Район, район
Любашівський, Район, район
Миколаївський, Район, район
м. Ізмаїл, Місто, місто
Назва, Місто/район, місто/район/область
Овідіопольський, Район, район
Окнянський, Район, район
Подільський, Район, район
Ренійський, Район, район
Роздільнянський, 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 ,команда `sort -t "," -k3 -r 3.csv`



```
(sirenko_mariya) 91.219.60.189 — Konsole
File Edit View Bookmarks Settings Help
[sirenko_mariya@vpsj3IeQ ~]$ sort -t"," -k3 -r 3.csv
Ширяївський, Район, район
Татарбунарський, Район, район
Тарутинський, Район, район
Саратський, Район, район
Савранський, Район, район
Роздільнянський, Район, район
Ренійський, Район, район
Подільський, Район, район
Окнянський, Район, район
Овідіопольський, Район, район
Миколаївський, Район, район
Любашівський, Район, район
Лиманський, Район, район
Кодимський, Район, район
Кілійський, Район, район
Ізмаїльський, Район, район
Іванівський, Район, район
Захарівський, Район, район
Великомихайлівський, Район, район
Болградський, Район, район
Біляївський, Район, район
Білгород-Дністровський, Район, район
Березівський, Район, район
Балтський, Район, район
Арцизький, Район, район
Ананьївський, Район, район
Назва, Місто/район, місто/район/область
м. Ізмаїл, Місто, місто
[sirenko_mariya@vpsj3IeQ ~]$
```

Висновок: було отримано навички обробки текстових даних засобами оболонки Unix-подібних ОС інтерфейсу командного рядка. Завдання були виконані на віддаленому сервері з встановленою ОС Linux (Rat Hat 4.8.5). Найскладнішим виявилось завдання 2.2.6 з використання команди `join` . Для правильної роботи команди ,файли повинні бути відсортованими .