

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:
студентка групи AI-204
Томчук Вікторія Борисівна

Перевірив:
Блажко А.А.

Одеса-2021

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

2 Завдання до виконання

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. 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 № 3

Необхідно створити файл формату CSV, який містить табличні данні, використовуючи наступні кроки:

- 1) базове створення файлу CSV-формату, використовуючи Google Docs за методичними рекомендаціями:
 - 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.

Хід роботи

Завдання 1.1. Виведіть на екран перші десять назв каталогів або файлів, назви яких починаються з перших трьох букв вашого прізвища в латиниці.

```
tomchuk_viktoriya@vpsj3IeQ:~  
login as: tomchuk_viktoriya  
tomchuk_viktoriya@91.219.60.189's password:  
Last login: Sun Mar 14 13:55:07 2021 from 78.26.152.197  
[tomchuk_viktoriya@vpsj3IeQ ~]$ find / -name -name "tom*" 2> /dev/null | head  
/opt/oracle/product/18c/dbhomeXE/md/property_graph/lib/tomcat-embed-jasper-8.5.30.jar  
/opt/oracle/product/18c/dbhomeXE/md/property_graph/lib/tomcat-embed-core-8.5.30.jar  
/var/cache/yum/x86_64/7/updates/packages/tomcat-servlet-3.0-api-7.0.76-16.el7_9.noarch.rpm  
/var/spool/mail/tomchuk_viktoriya  
/var/softaculous/tomatocart  
/var/softaculous/apps/tomcat8  
/var/softaculous/apps/tomcat8/tomcatctl  
/var/softaculous/apps/tomcat9  
/var/softaculous/apps/tomee  
/var/softaculous/apps/tomcat6  
[tomchuk_viktoriya@vpsj3IeQ ~]$
```

Завдання 1.2. 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.

```
tomchuk_viktoriya@vpsj3IeQ:~  
[tomchuk_viktoriya@vpsj3IeQ ~]$ find /home -maxdepth 1 -type d -name "t*" 2> /dev/null  
/home/tretyakov_oleksandr  
/home/tsitkov_ruslan  
/home/tomchuk_viktoriya  
/home/topuzanov_bogdan  
/home/tokarev_oleksandr  
/home/tsigankova_anastasiya  
/home/tarasov_daniil  
[tomchuk_viktoriya@vpsj3IeQ ~]$
```

Завдання 1.3. 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).

```
tomchuk_viktoriya@vpsj3IeQ:~  
[tomchuk_viktoriya@vpsj3IeQ ~]$ find /home -maxdepth 1 -type d -name "t*" 2> /dev/null | cut -d'_' -f2 | sort | uniq  
anastasiya  
bogdan  
daniil  
oleksandr  
ruslan  
viktoriya  
[tomchuk_viktoriya@vpsj3IeQ ~]$
```

Завдання 1.4. 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.

```
tomchuk_viktoriya@vpsj3IeQ:~  
login as: tomchuk viktoriya  
tomchuk_viktoriya@91.219.60.189's password:  
Last login: Sat Mar 20 19:44:19 2021 from 78.26.152.232  
[tomchuk_viktoriya@vpsj3IeQ ~]$ cut -d':' -f1,3,6 /etc/passwd | tr ':' ',' > accounts.csv  
[tomchuk_viktoriya@vpsj3IeQ ~]$ ls  
accounts.csv  labtest_2.pdf                                os.lab1.cp1251.html  test.html  
labtest_2.odt  Operating-System.-Laboratory-Work-1  os.lab1.utf.html    tomchuk_lab_3  
[tomchuk_viktoriya@vpsj3IeQ ~]$ cat 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,/  
systemd-network,192,/  
dbus,81,/  
polkitd,999,/  

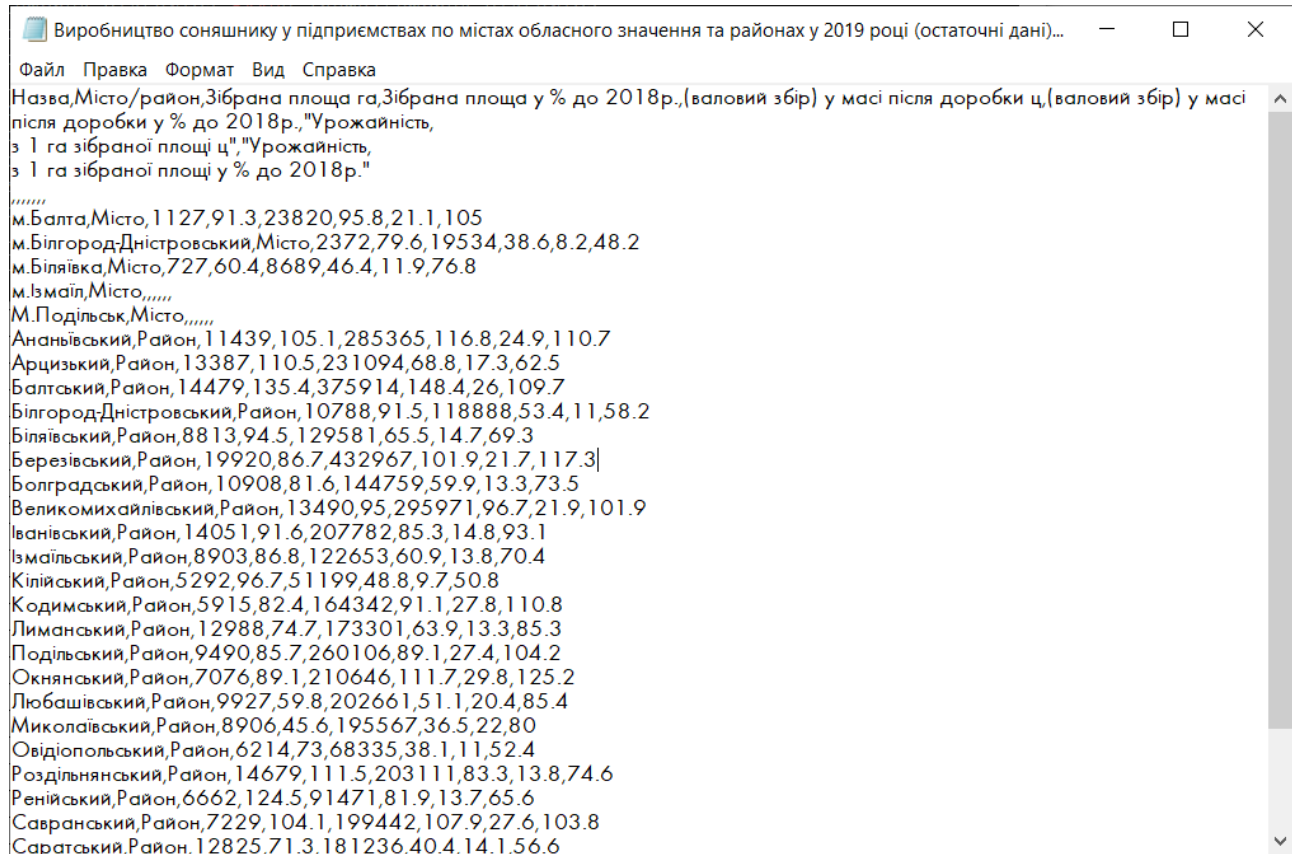
```

Завдання 1.5. З файлу `accounts.csv` отримайте на екран перелік облікових записів, в яких четверта цифра ідентифікатора користувача співпадає з четвертою цифрою вашого ідентифікатора користувача. (54400 через `cat /etc/passwd`)

```
tomchuk_viktoriya@vpsj3IeQ:~  
login as: tomchuk viktoriya  
tomchuk_viktoriya@91.219.60.189's password:  
Last login: Sun Mar 21 10:15:02 2021 from 78.26.152.232  
[tomchuk_viktoriya@vpsj3IeQ ~]$ grep -E ",[0-9]{3}0" "accounts.csv"  
soft,1000,/home/soft  
vmail,5000,/var/local/vmail  
tomchuk_viktoriya,54400,/home/tomchuk_viktoriya  
pargalova_sabina,54401,/home/pargalova_sabina  
chan_van,54402,/home/chan_van  
li_choan,54403,/home/li_choan  
nguen_trong,54404,/home/nguen_trong  
bezhanishvili_levan,54405,/home/bezhanishvili_levan  
bodnar_arsenij,54406,/home/bodnar_arsenij  
zelinko_igor,54407,/home/zelinko_igor  
bojchuk_oleksandr,54408,/home/bojchuk_oleksandr  
lihmanov_artem,54409,/home/lihmanov_artem  
[tomchuk_viktoriya@vpsj3IeQ ~]$ █
```

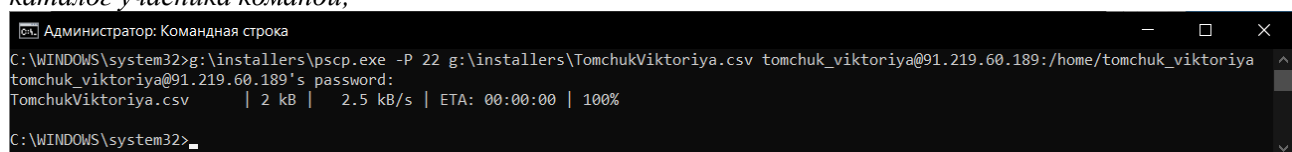
Завдання 2.1. базове створення файлу CSV-формату, використовуючи Google Docs за методичними рекомендаціями:

- кодування символів – UTF-8
- роздільник між колонками
- символ коми;
- назви файлів визначити з урахуванням транслітерації свого прізвища та номеру файлу.



Назва, Місто/район	Зібрана площа га	Зібрана площа у % до 2018р.	(валовий збір) у масі після доробки ц	(валовий збір) у масі після доробки у % до 2018р.	"Урожайність, з 1 га зібраної площі ц"	"Урожайність, з 1 га зібраної площі у % до 2018р."
м.Балта, Місто	1127,91.3	23820,95.8	21.1	105		
м.Білгород-Дністровський, Місто	2372,79.6	19534,38.6	8.2	48.2		
м.Біляївка, Місто	727,60.4	8689,46.4	11.9	76.8		
м.Ізмаїл, Місто						
М.Подільськ, Місто						
Ананьївський, Район	11439,105.1	285365,116.8	24.9	110.7		
Арцизький, Район	13387,110.5	231094,68.8	17.3	62.5		
Балтський, Район	14479,135.4	375914,148.4	26	109.7		
Білгород-Дністровський, Район	10788,91.5	118888,53.4	11	58.2		
Біляївський, Район	8813,94.5	129581,65.5	14.7	69.3		
Березівський, Район	19920,86.7	432967,101.9	21.7	117.3		
Болградський, Район	10908,81.6	144759,59.9	13.3	73.5		
Великомихайлівський, Район	13490,95	295971,96.7	21.9	101.9		
Іванівський, Район	14051,91.6	207782,85.3	14.8	93.1		
Ізмаїльський, Район	8903,86.8	122653,60.9	13.8	70.4		
Кілійський, Район	5292,96.7	51199,48.8	9.7	50.8		
Кодимський, Район	5915,82.4	164342,91.1	27.8	110.8		
Лиманський, Район	12988,74.7	173301,63.9	13.3	85.3		
Подільський, Район	9490,85.7	260106,89.1	27.4	104.2		
Окнянський, Район	7076,89.1	210646,111.7	29.8	125.2		
Любашівський, Район	9927,59.8	202661,51.1	20.4	85.4		
Миколаївський, Район	8906,45.6	195567,36.5	22	80		
Овідіопольський, Район	6214,73	68335,38.1	11	52.4		
Роздільнянський, Район	14679,111.5	203111,83.3	13.8	74.6		
Ренійський, Район	6662,124.5	91471,81.9	13.7	65.6		
Савранський, Район	7229,104.1	199442,107.9	27.6	103.8		
Саратський, Район	12825,71.3	181236,40.4	14.1	56.6		

Завдання 2.2. копіювання файлів на сервер за адресою з IP = 91.219.60.189 у домашній каталог учасника команди;



```

C:\WINDOWS\system32>g:\installers\pscp.exe -P 22 g:\installers\TomchukViktoriya.csv tomchuk_viktoriya@91.219.60.189:/home/tomchuk_viktoriya
tomchuk_viktoriya@91.219.60.189's password:
TomchukViktoriya.csv      | 2 kB | 2.5 kB/s | ETA: 00:00:00 | 100%

C:\WINDOWS\system32>
  
```

Завдання 2.3. 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.

```
tomchuk_viktoriya@vpsj3ieQ:~$ ls
accounts.csv      Operating-System.-Laboratory-Work-1  test.html
labtest_2.odt     os.labl.cpl251.html                  tomchuk_lab_3
labtest_2.pdf     os.labl.utf.html                     TomchukViktoriya.csv
tomchuk_viktoriya@vpsj3ieQ:~$ sed 'a/"\([0-9]\+\)[,]\([0-9]\+\)\\"/1\2 /g' |
sed 's/(/,)\...1/1/g' | sed 's/(,)\-/1/g' | cut -d, -f1,2 TomchukViktoriya.csv
> 1.csv

^[
^C
tomchuk_viktoriya@vpsj3ieQ:~$ ls
1.csv      Operating-System.-Laboratory-Work-1  tomchuk_lab_3
accounts.csv  os.labl.cpl251.html                  TomchukViktoriya.csv
labtest_2.odt  os.labl.utf.html
labtest_2.pdf  test.html
tomchuk_viktoriya@vpsj3ieQ:~$ less 1.csv
Назва,Місто/район
з 1 га зібраної площі ц", "Урожайність
з 1 га зібраної площі у % до 2018р."
/
м.Балта,Місто
м.Вінгород-Дністровський,Місто
м.Вілявка,Місто
м.Гемал,Місто
М.Подільськ,Місто
Ананьївський,Район
Арцизький,Район
Балтський,Район
Вінгород-Дністровський,Район
Вілявський,Район
Березівський,Район
Болградський,Район
Великомихайлівський,Район
Гванівський,Район
Гемальський,Район
Кілійський,Район
Кодимський,Район
Лиманський,Район
Подільський,Район
Очнянський,Район
Лубашівський,Район
Миколаївський,Район
Овідіопольський,Район
Росдільнянський,Район
Ренійський,Район
Савранський,Район
Саратський,Район
Тарутинський,Район
Татарбунарський,Район
```

Завдання 2.4. передача створеного CSV-файлу в GitHub-репозиторій;

```
tomchuk_viktoriya@vpsj3IeQ:~  
[tomchuk_viktoriya@vpsj3IeQ ~]$ cp /home/tomchuk_viktoriya/1.csv /home/tomchuk_viktoriya/Operating-System-Laboratory-Work-4  
[tomchuk_viktoriya@vpsj3IeQ ~]$ git add 1.csv  
[tomchuk_viktoriya@vpsj3IeQ ~]$ git commit -m "lab 4"  
# On branch master  
# Untracked files:  
#   (use "git add <file>..." to include in what will be committed)  
#  
#       .bash_history  
#       .bash_logout  
#       .bash_profile  
#       .bashrc  
#       .cache/  
#       .config/  
#       .gitconfig  
#       .kshrc  
#       .local/  
#       Operating-System-Laboratory-Work-4/  
#       Operating-System.-Laboratory-Work-1/  
#       TomchukViktoriya.csv  
#       accounts.cvs  
#       labtest_2.odt  
#       labtest_2.pdf  
#       origin/  
#       os.lab1.cp1251.html  
#       os.lab1.utf.html  
#       test.html  
#       tomchuk_lab_3/  
nothing added to commit but untracked files present (use "git add" to track)
```

```
tomchuk_viktoriya@vpsj3IeQ:~  
[tomchuk_viktoriya@vpsj3IeQ ~]$ git push -u origin master  
Username for 'https://github.com': rxndx  
Password for 'https://rxndx@github.com':  
Counting objects: 3, done.  
Delta compression using up to 2 threads.  
Compressing objects: 100% (2/2), done.  
Writing objects: 100% (3/3), 572 bytes | 0 bytes/s, done.  
Total 3 (delta 0), reused 0 (delta 0)  
remote:  
remote: Create a pull request for 'master' on GitHub by visiting:  
remote:   https://github.com/rxndx/Operating-System-Laboratory-Work-4/pull/new/master  
remote:  
To https://github.com/rxndx/Operating-System-Laboratory-Work-4.git  
 * [new branch]      master -> master  
Branch master set up to track remote branch master from origin.  
[tomchuk_viktoriya@vpsj3IeQ ~]$
```

rxndx / **Operating-System-Laboratory-Work-4** Unwatch 1 Star 0 Fork 0

[Code](#) [Issues](#) [Pull requests](#) [Actions](#) [Projects](#) [Wiki](#) [Security](#) [Insights](#) [...](#)

master had recent pushes 1 minute ago [Compare & pull request](#)

master [Go to file](#) [Add file](#) [Code](#)

This branch is 1 commit ahead, 1 commit behind main. [Pull request](#) [Compare](#)

rxndx Lab 4 ... 1 hour ago 1

1.csv Lab 4 1 hour ago

Help people interested in this repository understand your project by adding a README. [Add a README](#)

About

No description, website, or topics provided.

Releases

No releases published [Create a new release](#)

Packages

No packages published [Publish your first package](#)

Завдання 2.5. отримання CSV-файлу, створеного іншим учасником команди;

```
tomchuk_viktoriya@vpsj3leQ:~
[ tomchuk_viktoriya@vpsj3leQ ~ ]$ git pull
Already up-to-date.
[ tomchuk_viktoriya@vpsj3leQ ~ ]$
```

```
tomchuk_viktoriya@vpsj3leQ:~
login as: tomchuk_viktoriya
tomchuk_viktoriya@91.219.60.189's password:
Last login: Sun Mar 21 14:19:15 2021 from 78.26.152.232
[ tomchuk_viktoriya@vpsj3leQ ~ ]$ git init
Reinitialized existing Git repository in /home/tomchuk_viktoriya/.git/
[ tomchuk_viktoriya@vpsj3leQ ~ ]$ ls
1.csv          Operating-System.-Laboratory-Work-1  test.html
2.csv          Operating-System-Laboratory-Work-4    tomchuk_lab_3
accounts.cvs   origin                               TomchukViktoriya.csv
labtest_2.odt  os.lab1.cp1251.html
labtest_2.pdf  os.lab1.utf.html
```


Завдання 2.6. об'єднання двох файлів в один за першим стовпчиком;

```
tomchuk_viktoriya@vpsj3leQ:~  
[tomchuk_viktoriya@vpsj3leQ ~]$ join -t, -a 1 -a 2 -o0,1,2,2.2 -e '' 1.csv 2.csv  
Назва,Місто/район,Місто/район  
з 1 га зібраної площі ц", "Урожайність", "Урожайність  
,,1 га зібраної площі у % до 2018р."  
,,  
м.Балта,Місто,Місто  
м.Білогород-Дністровський,Місто,Місто  
м.Біляївка,Місто,Місто  
м.Ізмаїл,Місто,Місто  
м.Подільськ,Місто,Місто  
Ананьївський, Район, Район  
Арцизький, Район, Район  
Валтський, Район, Район  
Білогород-Дністровський, Район, Район  
Біляївський, Район, Район  
Березівський, Район, Район  
Болградський, Район, Район  
Великомихайлівський, Район, Район  
Іванівський, Район, Район  
Ізмаїльський, Район, Район  
Кілійський, Район, Район  
Кодимський, Район, Район  
Лиманський, Район, Район  
Подільський, Район, Район  
Окнянський, Район, Район  
Любашівський, Район, Район  
Миколаївський, Район, Район  
Овідіопольський, Район, Район  
Роздільнянський, Район, Район
```

Завдання 2.7. 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.

```
tomchuk_viktoriya@vpsj3leQ:~  
[tomchuk_viktoriya@vpsj3leQ ~]$ sort --field-separator="," --key=2 -r 1.csv 2.csv  
з 1 га зібраної площі ц", "Урожайність  
з 1 га зібраної площі ц", "Урожайність  
Ширяївський, Район  
Ширяївський, Район  
Татарбунарський, Район  
Татарбунарський, Район  
Тарутинський, Район  
Тарутинський, Район  
Саратський, Район  
Саратський, Район  
Савранський, Район  
Савранський, Район  
Роздільнянський, Район  
Роздільнянський, Район  
Ренійський, Район  
Ренійський, Район  
Подільський, Район  
Подільський, Район  
Окнянський, Район  
Окнянський, Район  
Овідіопольський, Район  
Овідіопольський, Район  
Миколаївський, Район  
Миколаївський, Район  
Любашівський, Район  
Любашівський, Район  
Лиманський, Район  
Лиманський, Район
```

ВИСНОВОК: під час роботи над придбання навичок складної обробки текстових даних роботи засобами оболонки Unix-подібних ОС інтерфейсу командного рядка найскладнішим

завданням було саме відправка файлу через гіт в відповідний репозиторій.