

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 AI-203

Токарєв О.С.

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 ваше 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 учасника в команді, як в лабораторній роботі № 3 -

<https://docs.google.com/spreadsheets/d/1ITMLdjcw7nboUo9nFw4NwkWcvjbZeZZIDVb2Hg4laUQ>

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-1QybKQje49VdjFahkely1Ou>:

- 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 колонок за номером 2 - для 1-го учасника команди, або за номером 3 – для 2-го учасника команди.

## Хід роботи:

### 1. Обробка текстових даних результатів роботи команд

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

```
[tokarev_oleksandr@vpsj3IeQ ~]$ find / -name "tok*" 2> /dev/null | head
/var/spool/mail/tokarev_oleksandr
/usr/local/webuzo/enduser/webuzo/phpmyadmin/vendor/phpmyadmin/sql-parser/bin/tokenize-query
/usr/local/apps/php73/ext/tokenizer.so
/usr/local/apps/php73/ext/tokenizer.a
/usr/local/apps/python2/lib/python2.7/token.pyo
/usr/local/apps/python2/lib/python2.7/tokenize.pyc
/usr/local/apps/python2/lib/python2.7/test/tokenize_tests.txt
/usr/local/apps/python2/lib/python2.7/tokenize.py
/usr/local/apps/python2/lib/python2.7/token.py
/usr/local/apps/python2/lib/python2.7/tokenize.pyo
```

Виведіть на екран список назв домашніх каталогів користувачів, в назві яких присутня перша буква вашого прізвища:

```
[tokarev_oleksandr@vpsj3IeQ ~]$ find /home/ -maxdepth 1 -name "*t*" 2> /dev/null
/home/tretyakov_oleksandr
/home/yufitenko_andrij
/home/kebab_dmitro
/home/kovach_dmitro
/home/sherbakov_artur
/home/strinaglyuk_ivan
/home/koldunova_tetyana
/home/tsitkov_ruslan
/home/perets_sergij
/home/sultanov_artem
/home/shostak_roman
/home/koldunova_anastasiya
/home/nguen_trong
/home/vyatkīn_danilo
/home/katrich_evgen
/home/belobrov_artur
/home/stepanenko_gleb
/home/tomchuk_viktoriya
/home/kalarashan_kostyantın
/home/chobotar_oleksandra
/home/sidelnikov_mikita
/home/topuzanov_bogdan
/home/svetashov_danilo
/home/shapovalova_viktoriya
/home/tokarev_oleksandr
/home/borovets_vladislav
/home/baranyuk_dmitro
/home/voyakovskij_dmitro
/home/babich_artem
/home/panfilov_kristian
/home/tsigankova_anastasiya
/home/matnenko_stanislaw
/home/tarasov_daniil
/home/martinyuk_denis
/home/nesterenko_mikola
/home/soft-1
/home/kruts_viktor
/home/krutienko_andrij
/home/kostetskiy_bogdan
/home/demtsun_andrij
/home/lihmanov_artem
/home/arestov_yaroslav
/home/kelembet_evgen
```

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

```
[tokarev_oleksandr@vpsj3IeQ ~]$ find /home/ -maxdepth 1 -name "*t*" 2> /dev/null  
| sort | uniq | cut -c 7-  
arestov_yaroslav  
babich_artem  
baranyuk_dmitro  
belobrov_artur  
borovets_vladislav  
chobotar_oleksandra  
demsun_andrij  
kalarashan_kostyantyn  
katrich_evgen  
kebap_dmitro  
kelembet_evgen  
koldunova_anastasiya  
koldunova_tetyana  
kostetskiy_bogdan  
kovach_dmitro  
krutienko_andrij  
kruts_viktor  
lihmanov_artem  
martinyuk_denis  
matnenko_stanislav  
nesterenko_mikola  
nguen_trong  
panfilov_kristian  
perets_sergij  
shapovalova_viktoriya  
sherbakov_artur  
shostak_roman  
sidelnikov_mikita  
soft-1  
stepanenko_gleb  
strinaglyuk_ivan  
sultanov_artem  
svetashov_danilo  
tarasov_daniil  
tokarev_oleksandr  
tomchuk_viktoriya  
topuzanov_bogdan  
tretyakov_oleksandr  
tsigankova_anastasiya  
tsitkov_ruslan  
voyakovskij_dmitro  
vyatkin_danilo  
yufitenko_andrij
```

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.:

```
[tokarev_oleksandr@vpsj3IeQ ~]$ sed -n '34,138p' /etc/passwd | sed 's/\(^\\)/,/g' | sed 's/,/\bin\/bash:/\bin\/bash/g' | cut -f1,3,5,6 -d, | cat 1>accounts.csv
[tokarev_oleksandr@vpsj3IeQ ~]$ cat accounts.csv
arestov_yaroslav:x:54322:54331::/home/arestov_yaroslav:/bin/bash
martinyuk_denis:x:54323:54332::/home/martinyuk_denis:/bin/bash
bogachik_egor:x:54324:54333::/home/bogachik_egor:/bin/bash
eliseev_danilo:x:54325:54334::/home/eliseev_danilo:/bin/bash
bochkarova_karina:x:54326:54335::/home/bochkarova_karina:/bin/bash
tsigankova_anastasiya:x:54327:54327::/home/tsigankova_anastasiya:/bin/bash
vyatkin_danilo:x:54328:54328::/home/vyatkin_danilo:/bin/bash
harkevich_leonid:x:54329:54329::/home/harkevich_leonid:/bin/bash
kalarashan_kostyantyn:x:54330:54336::/home/kalarashan_kostyantyn:/bin/bash
nosov_andrij:x:54331:54337::/home/nosov_andrij:/bin/bash
kebab_dmitro:x:54332:54338::/home/kebab_dmitro:/bin/bash
prihodko_egor:x:54333:54339::/home/prihodko_egor:/bin/bash
kelembet_evgen:x:54334:54340::/home/kelembet_evgen:/bin/bash
malofeev_denis:x:54335:54341::/home/malofeev_denis:/bin/bash
tarasov_daniil:x:54336:54342::/home/tarasov_daniil:/bin/bash
stepanenko_gleb:x:54337:54343::/home/stepanenko_gleb:/bin/bash
strinaglyuk_ivan:x:54338:54344::/home/strinaglyuk_ivan:/bin/bash
yufitenko_andrij:x:54339:54345::/home/yufitenko_andrij:/bin/bash
cheremnih_grigoriy:x:54340:54346::/home/cheremnih_grigoriy:/bin/bash
barkar_karina:x:54341:54347::/home/barkar_karina:/bin/bash
loboda_oleksandra:x:54342:54348::/home/loboda_oleksandra:/bin/bash
borovets_vladislav:x:54343:54349::/home/borovets_vladislav:/bin/bash
matnenko_stanislav:x:54344:54350::/home/matnenko_stanislav:/bin/bash
bodnar_ilya:x:54345:54351::/home/bodnar_ilya:/bin/bash
nezhivih_mariya:x:54346:54352::/home/nezhivih_mariya:/bin/bash
grebenik_anzhelika:x:54347:54353::/home/grebenik_anzhelika:/bin/bash
lobko_daniil:x:54348:54354::/home/lobko_daniil:/bin/bash
baranjuk_dmitro:x:54349:54355::/home/baranjuk_dmitro:/bin/bash
kovach_dmitro:x:54350:54356::/home/kovach_dmitro:/bin/bash
sidelnikov_mikita:x:54351:54357::/home/sidelnikov_mikita:/bin/bash
lukashak_daniil:x:54352:54358::/home/lukashak_daniil:/bin/bash
markovskij_danilo:x:54353:54359::/home/markovskij_danilo:/bin/bash
melnichenko_egor:x:54354:54360::/home/melnichenko_egor:/bin/bash
perets_sergij:x:54355:54361::/home/perets_sergij:/bin/bash
uzun_mihajlo:x:54356:54362::/home/uzun_mihajlo:/bin/bash
polyanskiy_mikola:x:54357:54363::/home/polyanskiy_mikola:/bin/bash
sherbakov_artur:x:54358:54364::/home/sherbakov_artur:/bin/bash
```

```
[tokarev_oleksandr@vpsj3IeQ ~]$ grep tokarev accounts.csv
tokarev_oleksandr:x:54369:54375::/home/tokarev_oleksandr:/bin/bash
[tokarev_oleksandr@vpsj3IeQ ~]$ sed -n '/[0-9]\{3\}6[0-9]/p' accounts.csv
melnichenko_egor:x:54354:54360::/home/melnichenko_egor:/bin/bash
perets_sergij:x:54355:54361::/home/perets_sergij:/bin/bash
uzun_mihajlo:x:54356:54362::/home/uzun_mihajlo:/bin/bash
polyanskiy_mikola:x:54357:54363::/home/polyanskiy_mikola:/bin/bash
sherbakov_artur:x:54358:54364::/home/sherbakov_artur:/bin/bash
ershov_oleksandr:x:54359:54365::/home/ershov_oleksandr:/bin/bash
zagorujko_maksim:x:54360:54366::/home/zagorujko_maksim:/bin/bash
zijodulloev_ehson:x:54361:54367::/home/zijodulloev_ehson:/bin/bash
veselkova_anna:x:54362:54368::/home/veselkova_anna:/bin/bash
voyakovskij_dmitro:x:54363:54369::/home/voyakovskij_dmitro:/bin/bash
golovanchuk_mikola:x:54364:54370::/home/golovanchuk_mikola:/bin/bash
demtsun_andrij:x:54365:54371::/home/demtsun_andrij:/bin/bash
grishenko_oleksandra:x:54366:54372::/home/grishenko_oleksandra:/bin/bash
myagkih_arina:x:54367:54373::/home/myagkih_arina:/bin/bash
kurgan_roman:x:54368:54374::/home/kurgan_roman:/bin/bash
tokarev_oleksandr:x:54369:54375::/home/tokarev_oleksandr:/bin/bash
```

## 2. Обробка текстового файлу формату CSV

Копіювання файлів на сервер за адресою з IP = 91.219.60.189 у домашній каталог учасника команди:

```
D:\os_lab4>pscp -P 22 tokarev_4_2.csv tokarev_oleksandr@91.219.60.189:/home/tokarev_oleksandr/
tokarev_oleksandr@91.219.60.189's password:
tokarev_4_2.csv | 2 kB | 2.5 kB/s | ETA: 00:00:00 | 100%
```

## Модифікація файлу CSV-формату:

```
[tokarev_oleksandr@vpsj3IeQ ~]$ sed 's/\\\"([0-9]\\+)[,]\\\"([0-9]\\+)[\\\"/\\\\1\\\\2 /g' tokarev_4.csv | cut -f1,3 -d, |>tokarev_4_2.csv
[tokarev_oleksandr@vpsj3IeQ ~]$ cat tokarev_4_2.csv
Назва,всього
фізичні особи підприємці"
Одеська область,464772
м.Одеса,273720
м.Валта,3166
м.Вілгород-Дністровський,13123
м.Вілянка,3089
м.Ізмаїл,16246
м.Чорноморськ,23059
м.Подільськ,4149
м.Теплодар,1359
м.Южне,9242
Ананьївський,2218
Арцизький,5057
Валтський,1707
Вілгород-Дністровський,5474
Вілянівський,10985
```

## Передача створеного CSV-файлу в GitHub-репозиторій:

```
[tokarev_oleksandr@vpsj3IeQ ~]$ cp tokarev_4.csv Operating-System.-Laboratory-Work-1./tokarev_4_2.csv
[tokarev_oleksandr@vpsj3IeQ ~]$ ls
accounts.csv          os.labl.cp1251.html  tokarev_4_2.csv      Tokarev.doc          tokarev_lab_3
Operating-System.-Laboratory-Work-1.  os.labl.utf.html    tokarev_4_2.csv      Tokarev.html         Tokarev.pdf
[tokarev_oleksandr@vpsj3IeQ ~]$
[tokarev_oleksandr@vpsj3IeQ ~]$ cd Operating-System.-Laboratory-Work-1.
[tokarev_oleksandr@vpsj3IeQ Operating-System.-Laboratory-Work-1]$ git add *
```

```
[tokarev_oleksandr@vpsj3IeQ Operating-System.-Laboratory-Work-1]$ git commit -m "tokarev"
[main 19bae7e] tokarev
6 files changed, 75 insertions(+)
create mode 100644 kurgan_4_1.csv
create mode 100644 tokarev_4_2.csv
create mode 100644 "\320\232\321\203\321\200\320\263\320\260\320\275\320\240\320\276\320\274\320\260\320\275.pdf"
create mode 100644 "\320\233\320\260\320\261\_342\204\2263\_320\236\320\241\_320\240\320\276\320\274\320\260\320\275\320\232\321\203\321\200\320\263\320\260\320\275.docx"
create mode 100644 "\320\236\320\241\_320\233\320\260\320\261\_3\_320\242\320\276\320\272\320\260\321\200\320\265\320\262.pdf"
create mode 100644 "\320\242\320\276\320\272\320\260\321\200\320\265\320\262\320\220\320\273\320\265\320\272\321\201\320\260\320\275\320\264\321\200.pdf"
```

```
[tokarev_oleksandr@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': Senpail69
Password for 'https://Senpail69@github.com':
Counting objects: 8, done.
Delta compression using up to 2 threads.
Compressing objects: 100% (6/6), done.
Writing objects: 100% (6/6), 1.56 KiB | 0 bytes/s, done.
Total 6 (delta 3), reused 0 (delta 0)
remote: Resolving deltas: 100% (3/3), completed with 1 local object.
To https://github.com/MonX94/Operating-System.-Laboratory-Work-1.
9464547..965a079 main -> main
```

```
[tokarev_oleksandr@vpsj3IeQ Operating-System.-Laboratory-Work-1]$ git pull
remote: Enumerating objects: 11, done.
remote: Counting objects: 100% (11/11), done.
remote: Compressing objects: 100% (8/8), done.
remote: Total 9 (delta 4), reused 5 (delta 1), pack-reused 0
Unpacking objects: 100% (9/9), done.
From https://github.com/MonX94/Operating-System.-Laboratory-Work-1.
f4b21d1..7745c91 main -> origin/main
Updating f4b21d1..7745c91
Fast-forward
 tokarev_4_2.csv | 37 +++++
 1 file changed, 37 insertions(+)
[tokarev_oleksandr@vpsj3IeQ Operating-System.-Laboratory-Work-1]$
```



Об'єднання двох таблиць:

```
[kurgan_roman@vpsj3IeQ Operating-System.-Laboratory-Work-1.]$ cat kurgan_4_1.csv  
> merged_lab4.csv  
[kurgan_roman@vpsj3IeQ Operating-System.-Laboratory-Work-1.]$ cat tokarev_4_2.csv  
v | cut -f2 -d,  
всього  
464772  
273720  
3166  
13123  
3089  
16246  
23059
```

```
[kurgan_roman@vpsj3IeQ Operating-System.-Laboratory-Work-1.]$ cat tokarev_4_2.csv  
v | cut -f2 -d, > tokarev_4_2_temp.csv  
[kurgan_roman@vpsj3IeQ Operating-System.-Laboratory-Work-1.]$ paste -d ',' kurgan_4_1.csv tokarev_4_2_temp.csv  
Назва,всього,всього  
Одеська область,342926,464772  
м.Одеса,206857,273720  
м.Балта,2067,3166  
м.Білгород-Дністровський,9165,13123  
м.Біляївка,2542,3089  
м.Ізмаїл,12281,16246  
м.Чорноморськ,19154,23059  
м.Подільськ,2497,4149  
м.Теплодар,804,1359  
[kurgan_roman@vpsj3IeQ Operating-System.-Laboratory-Work-1.]$ paste -d ',' kurgan_4_1.csv tokarev_4_2_temp.csv > merged_lab4.csv
```

```
[tokarev_oleksandr@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/MonX94/Operating-System.-Laboratory-Work-1.  
7745c91..2c3a9f2 main -> origin/main  
Updating 7745c91..2c3a9f2  
Fast-forward  
merged_lab4.csv | 37 +++++++++++++++++++++++++++++++++++++  
1 file changed, 37 insertions(+)  
create mode 100644 merged_lab4.csv
```

Виведення на екран змісту CSV-файлу з рішення попереднього завдання із сортуванням рядків файлу-таблиці у зворотному напрямі значень колонки за номером 3 для 2-го учасника команди

```
[tokarev_oleksandr@vpsj3IeQ Operating-System.-Laboratory-Work-1.1]$ sort -t, -k3 merged_lab4.csv
Біляївський,6911,10985
Захарівський,714,1154
м.Вілгород-Дністровський,9165,13123
м.Теплодар,804,1359
Миколаївський,899,1380
Лиманський,11321,14556
м.Ізмаїл,12281,16246
Балтський,1134,1707
Подільський,1290,1764
Овідіопольський,12407,17659
Савранський,1383,1982
Окнянський,1401,1989
Ананьївський,1412,2218
Ширяївський,1425,2274
Кодимський,1404,2299
Любашівський,1500,2302
м.Чорноморськ,19154,23059
Великомихайлівський,1806,2607
м.Одеса,206857,273720
Березівський,1765,2865
Іванівський,2068,2997
м.Біляївка,2542,3089
м.Балта,2067,3166
Ренійський,2256,3282
Татарбунарський,2082,3444
Тарутинський,2681,3842
м.Подільськ,2497,4149
Ізмаїльський,2778,4308
Роздільнянський,2607,4423
Одеська область,342926,464772
Саратський,3712,5034
Арцизький,3293,5057
Вілгород-Дністровський,3768,5474
Кілійський,3699,5504
Болградський,3945,6513
м.Южне,7898,9242
Назва,всього,всього
```

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