

МІНІСТЕРСТВО ОСВІТИ Й НАУКИ УКРАЇНИ
ОДЕСЬКИЙ НАЦІОНАЛЬНИЙ ПОЛІТЕХНІЧНИЙ УНІВЕРСИТЕТ
Інститут комп'ютерних систем
Кафедра інформаційних систем

Лабораторна робота № 4
За дисципліною: "Операційні системи"
Тема: «Складна обробка текстових даних засобами оболонки Unix-
подібних ОС інтерфейсу командного рядка»

Виконав:
Студент групи АІ-205
Шостак Р.С.
Перевірили:
Блажко О.А.
Дрозд М.О.

Одеса 2021

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

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

```
[shostak_roman@vpsj3IeQ ~]$ login as: shostak_roman
[shostak_roman@91.219.60.189's password:
Last login: Sun Mar 14 14:13:37 2021 from 62.16.26.1
[shostak_roman@vpsj3IeQ ~]$
```

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

```
[shostak_roman@vpsj3IeQ ~]$ find / -name "sho*" 2> /dev/null | head -n 10
/opt/oracle/product/18c/dbhomeXE/rdbms/admin/showlibc.sql
/opt/oracle/product/18c/dbhomeXE/rdbms/xml/em/orarep/config/show_params.xsd
/opt/oracle/product/18c/dbhomeXE/rdbms/xml/em/orarep/security/show_profiledetail.x
/opt/oracle/product/18c/dbhomeXE/rdbms/xml/em/orarep/security/show_objects.xsd
/opt/oracle/product/18c/dbhomeXE/rdbms/xml/em/orarep/security/show_roles.xsd
/opt/oracle/product/18c/dbhomeXE/rdbms/xml/em/orarep/security/show_profiles.xsd
/opt/oracle/product/18c/dbhomeXE/rdbms/xml/em/orarep/security/show_users.xsd
/opt/oracle/product/18c/dbhomeXE/rdbms/xml/em/orarep/security/show_privs.xsd
/opt/oracle/product/18c/dbhomeXE/rdbms/xml/em/orarep/security/show_userdetail.xsd
/opt/oracle/product/18c/dbhomeXE/rdbms/xml/em/orarep/security/show_roledetail.xsd
[shostak_roman@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.

```
[shostak_roman@vpsj3IeQ home]$ find . -maxdepth 2 -name "s*" 2> /dev/null
./sherbakov_georgij
./sherbakov_artur
./strinaglyuk_ivan
./sinyakov_igor
./sultanov_artem
./sokolov_oleksij
./shostak_roman
./shostak_roman/shostak_lab_3
./shostak_roman/shostak.txt
./sivash_andrij
./stepanenko_gleb
./sirenko_mariya
./sidelnikov_mikita
./svetashov_danilo
./shapovalova_viktoriya
./soft-1
./shulyak_mikola
[shostak_roman@vpsj3IeQ home]$
```

3) Модифікуйте рішення попереднього завдання так, що на екран було виведено імена користувачів, наприклад, maria, oleksandr та інші, відсортовані в порядку зростання значення, та без дублювання цих значень (використайте конвеєр з команд cut, sort, uniq).

```
[shostak_roman@vpsj3IeQ home]$ find -maxdepth 1 -name "s*" 2> /dev/null |cut -d '_' -f2 |sort|uniq
andrij
artem
artur
danilo
georgij
gleb
igor
ivan
mariya
mikita
mikola
oleksij
roman
./soft-1
viktoriya
[shostak roman@vpsj3IeQ home]$
```

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 з файлу /etc/passwd новий файл з назвою accounts.csv, який буде містити всі облікові записи користувачів, але лише з колонками «ім'я користувача (логін)», «ідентифікатор користувача», «початковий каталог користувача». Колонки повинні мати символом розділу символ «кома», як у стандартному CSV-форматі.

```
[shostak_roman@vpsj3IeQ ~]$ sed '1,33d; s/:x::,/; s/:[0-9]*::/,; s:/\bin\/*bash//' /etc/passwd| cat >accounts.csv
[shostak_roman@vpsj3IeQ ~]$ cat accounts.csv
arestov_yaroslav,54322,/home/arestov_yaroslav
martinyuk_denis,54323,/home/martinyuk_denis
bogachik_igor,54324,/home/bogachik_igor
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_kostyantin,54330,/home/kalarashan_kostyantin
nosov_andrij,54331,/home/nosov_andrij
kebab_dmitro,54332,/home/kebab_dmitro
prihodko_igor,54333,/home/prihodko_igor
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
cherevnnih_grigorij,54340,/home/cherevnnih_grigorij
barkar_karina,54341,/home/barkar_karina
loboda_oleksandra,54342,/home/loboda_oleksandra
borovets_vladislav,54343,/home/borovets_vladislav
matnenko_stanislav,54344,/home/matnenko_stanislav
bodnar_illya,54345,/home/bodnar_illya
nezhivih_mariya,54346,/home/nezhivih_mariya
grebenik_anzhelika,54347,/home/grebenik_anzhelika
lobko_daniil,54348,/home/lobko_daniil
baranyuk_dmitro,54349,/home/baranyuk_dmitro
kovach_dmitro,54350,/home/kovach_dmitro
sidelnikov_mikita,54351,/home/sidelnikov_mikita
lukashak_daniil,54352,/home/lukashak_daniil
markovskij_danilo,54353,/home/markovskij_danilo
melnichenko_igor,54354,/home/melnichenko_igor
perets_sergij,54355,/home/perets_sergij
uzun_mihajlo,54356,/home/uzun_mihajlo
polyanskiy_mikola,54357,/home/polyanskiy_mikola
sherbakov_artur,54358,/home/sherbakov_artur
ershov_oleksandr,54359,/home/ershov_oleksandr
zagorujko_maksim,54360,/home/zagorujko_maksim
zijodulloev_ehson,54361,/home/zijodulloev_ehson
veselkova Anna,54362,/home/veselkova_Anna
```

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

```
[shostak_roman@vpsj3IeQ ~]$ grep -E [^0-9][0-9]{3}2 "accounts.csv"
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
svetashov_danilo,54420,/home/svetashov_danilo
shostak_roman,54421,/home/shostak_roman
chobotar_oleksandra,54422,/home/chobotar_oleksandra
shapovalova_viktoriya,54423,/home/shapovalova_viktoriya
alekseeva_alina,54424,/home/alekseeva_alina
guda_liliya,54425,/home/guda_liliya
topuzanov_bogdan,54426,/home/topuzanov_bogdan
[shostak_roman@vpsj3IeQ ~]$
```

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 файла 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.

Назва,місто/район,Зібрана площа -га,Зібрана площа -% від 2018р.,Валовий збір у масі після доробки -ц,Валовий збір у масі після доробки - % з 1 га зібраної плоші - ц","Урожайність з 1 га зібраної плоші - % 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,14479,135.4,375914,148.4,26,189.7
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,9490,85.7,260186,89.1,27.4,104.2
Окнянський,район,7076,89.1,219646,111.7,29.8,125.2
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
Тарутинський,район,19682,110.1,377099,91.1,19.2,82.8
Татарбунарський,район,6213,45.9,54165,14.8,8.7,32.1
Захарівський,район,7232,89.3,134210,92.7,18.6,103.9
Ширяївський,район,15302,80.8,277174,78.5,18.1,96.8

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

Имя	Размер	Тип	Изменено
..		Родительский кат...	28.03.2021 21:05:42
lab1oc.doc	239 KB	Документ Micros...	21.03.2021 23:52:21
lab1oc.docx	217 KB	Документ Micros...	21.03.2021 23:48:13
lab1oc.pdf	158 KB	Microsoft Edge P...	21.03.2021 23:52:44
lab1oc_html_2e030c1...	76 KB	Файл "PNG"	21.03.2021 23:56:37
lab1oc_html_6fb8178...	48 KB	Файл "JPG"	21.03.2021 23:56:37
lab1oc_html_21a2723...	12 KB	Файл "PNG"	21.03.2021 23:56:37
lab1oc_html_51c8800...	2 KB	Файл "PNG"	21.03.2021 23:56:37
lab1oc_html_83b4707...	2 KB	Файл "PNG"	21.03.2021 23:56:37
lab1oc_html_19287cb...	4 KB	Файл "PNG"	21.03.2021 23:56:37
lab1oc_html_b506487...	46 KB	Файл "PNG"	21.03.2021 23:56:37
lab1oc_html_c6deac7...	2 KB	Файл "PNG"	21.03.2021 23:56:37
lab1oc_html_d8f7403...	5 KB	Файл "PNG"	21.03.2021 23:56:37
lab1oc_html_ffccc678...	2 KB	Файл "PNG"	21.03.2021 23:56:37
os.lab.cp1251.html	31 KB	Chrome HTML Do...	22.03.2021 0:11:19
photo_2021-03-24_13...	24 KB	Файл "JPG"	24.03.2021 14:02:17
photo_2021-03-26_16...	52 KB	Файл "JPG"	26.03.2021 16:26:47
МУ ЛР Комп схем и ...	1 268 KB	Документ Micros...	28.03.2021 21:05:41

Имя	Размер	Изменено	Права	Владелец
..		28.02.2021 20:22:15	rwxr-xr-x	root
ferma		20.03.2021 19:42:24	rwxrwxr-x	shostak_roman
Operating-System-L...		15.03.2021 9:11:05	rwxrwxr-x	shostak_roman
shostak_lab_3		20.03.2021 20:37:39	rwxrwxr-x	shostak_roman
accounts.csv	5 KB	28.03.2021 21:50:05	rw-rw-r--	shostak_roman
hey2	0 KB	22.03.2021 8:44:21	rw-rw-r--	shostak_roman
lab1oc.doc	239 KB	21.03.2021 23:52:21	rw-rw-r--	shostak_roman
lab1oc.html	34 KB	21.03.2021 23:56:37	rw-rw-r--	shostak_roman
lab1oc.pdf	158 KB	21.03.2021 23:52:44	rw-rw-r--	shostak_roman
shostak81.csv	3 KB	28.03.2021 22:42:46	rw-rw-r--	shostak_roman

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 ε);
- видалення з файлу стовпчиків окрім першого та третього;

```
[shostak_roman@vpsj3IeQ ~]$ sed 's/\"([0-9]+\+[,]\([0-9]+\+)\\"/\1.\2 /g' shostak81.csv | cut -f1,2 -d, 1>shostak81ver2.csv
[shostak_roman@vpsj3IeQ ~]$ cat shostak81ver2.csv
Назва,місто/район
з 1 га зібраної плоші - ц", "Урожайність
з 1 га зібраної плоші - % до 2018р."
,
м.Балта,місто
м.Білгород-Дністровський,місто
м.Біллівка,місто
м.Ізмаїл,місто
М.Подільськ,місто
Ананьївський,район
Арцизький,район
Балтський,район
Білгород-Дністровський,район
Білінівський,район
Березівський,район
Болградський,район
Великомихайлівський,район
Іванівський,район
Ізмаїльський,район
Кілійський,район
Кодимський,район
Лиманський,район
Подільський,район
Окницький,район
Любашівський,район
Миколаївський,район
Овідопольський,район
Роздільнянський,район
Ренийський,район
Савранський,район
Саратський,район
Тарутинський,район
Татарбунарський,район
Захарівський,район
Ширіївський,район
[shostak_roman@vpsj3IeQ ~]$ █
```

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

```
[shostak_roman@vpsj3IeQ svetashov_shostack]$ git push
Username for 'https://github.com': DLuckYD
Password for 'https://DLuckYD@github.com':
Counting objects: 4, done.
Delta compression using up to 2 threads.
Compressing objects: 100% (3/3), done.
Writing objects: 100% (3/3), 679 bytes | 0 bytes/s, done.
Total 3 (delta 0), reused 0 (delta 0)
To https://github.com/daniil-svetashov/svetashov_shostack
  a4061ff..e6d6979  main -> main
[shostak_roman@vpsj3IeQ svetashov_shostack]$ █
```

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

```
[shostak_roman@vpsj3IeQ svetashov_shostack]$ ls
shostak81ver2.csv  start  table2.csv
[shostak_roman@vpsj3IeQ svetashov_shostack]$ █
```

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

```
[shostak_roman@vpsj3IeQ svetashov_shostack]$ join -t, -a 1 -a 2 -o0,1.2,2.2 -e '' shostak81ver2.csv table2.csv | cat > tot.csv
join: shostak81ver2.csv:4: is not sorted: ,
join: table2.csv:4: is not sorted: Березівський,район
[shostak_roman@vpsj3IeQ svetashov_shostack]$ cat tot
cat: tot: No such file or directory
[shostak_roman@vpsj3IeQ svetashov_shostack]$ ls
shostak81ver2.csv  start table2.csv  tot.csv
[shostak_roman@vpsj3IeQ svetashov_shostack]$ cat tot.csv
Назва,місто/район,місто/район
з 1 га зібраної плоші - ц","Урожайність,
,,1 га зібраної плоші - % до 2016р."
```
м.Балта,місто,
м.Білгород-Дністровський,місто,
м.Біляївка,місто,
м.Ізмаїл,місто,
м.Одеса,,місто
м.Подільськ,місто,
м.Ананьївський,район,
м.Арцизький,район,
м.Балта,район,
м.Білгород-Дністровський,район,
м.Біляївський,район,
м.Верезівський,район,
м.Волградський,район,
м.Великомихайлівський,район,
м.Іванівський,район,
м.Ізмаїльський,район,
м.Кілійський,район,
м.Кодимський,район,
м.Пиманський,район,
м.Одеський район,район
м.Верезівський,,район
м.Білгород-Дністровський,,район
м.Волградський,,район
м.Ізмаїльський,,район
м.Подільський,район,район
м.Скнианський,район,
м.Любашівський,район,
м.Миколаївський,район,
м.Світловопольський,район,
м.Роздільнянський,район,район
м.Ренійський,район,
м.Савранський,район,
м.Саратський,район,
м.Гаргузинський,район,
м.Гатарбунарський,район,
м.Захарівський,район,
м.Ширяївський,район,
[shostak_roman@vpsj3IeQ svetashov_shostack]$
```

---

```
join -t, -a 1 -a 2 -o0,1.2,2.2 -e " shostak81ver2.csv table2.csv | cat > tot.csv
```

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

```
[shostak_roman@vpsj3IeQ svetashov_shostack]$ sort --field-separator="," --key=2 -r tot.csv
з 1 га зібраної площі - ц","Урожайність,
Роздільнянський, район, район
Подільський, район, район
Одеський район, район
Ізмаїльський,, район
Болградський,, район
Білгород-Дністровський,, район
Березівський,, район
Ширяївський, район,
Гатарбунарський, район,
Гарутинський, район,
Саратський, район,
Завранський, район,
Ренійський, район,
Окнянський, район,
Світловодський, район,
Миколаївський, район,
Любашівський, район,
Пиманський, 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,
м. Одеса,, місто
Назва, місто/район, місто/район
м. Подільськ, місто,
м. Ізмаїл, місто,
м. Біляївка, місто,
м. Білгород-Дністровський, місто,
м. Валта, місто,
,,1 га зібраної площі - % до 2018р."
"
[shostak_roman@vpsj3IeQ svetashov_shostack]$
```

---