

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»

Виконала:
Студентка групи AI-205
Алексєєва А. О.
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

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

```
alekseeva_alina@vpsj3IeQ:~  
login as: alekseeva_alina  
alekseeva_alina@91.219.60.189's password:  
Last login: Tue Mar 30 12:54:53 2021 from odessa-rentaflat.tenet.odessa.ua  
[alekseeva_alina@vpsj3IeQ ~]$ find / -name "ale*" 2> /dev/null | head -n 10  
/opt/oracle/product/18c/dbhomeXE/jdk/lib/missioncontrol/plugins/com.jrockit.mc.c  
onsole.ui.notification_5.5.2.174165/icons/alert_obj.png  
/opt/oracle/product/18c/dbhomeXE/cv/remenv/pluggable/alert_log_file_size_analyze  
r2.sh  
/opt/oracle/product/18c/dbhomeXE/cv/remenv/pluggable/alert_log_file_size_analyze  
r.sh  
/opt/oracle/diag/clients/user_oracle/host_1552702111_110/alert  
/var/spool/mail/alekseeva_alina  
/var/softaculous/alegro  
/usr/share/locale/ale  
/usr/local/webuzo/enduser/softimages/top15/48/alegro.png  
/usr/local/webuzo/enduser/softimages/top15/70/alegro.png  
/usr/local/webuzo/enduser/softimages/top15/alegro.png  
[alekseeva_alina@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.

```
[alekseeva_alina@vpsj3IeQ home]$ find -maxdepth 2 -name "a*" 2> /dev/null  
./ahmadi_umran  
./ahmadi_muhamad  
./alekseeva_alina  
./alekseeva_alina/alekseeva_lab_3  
./arestov_yaroslav  
[alekseeva_alina@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).

```
[alekseeva_alina@vpsj3IeQ home]$ find -maxdepth 1 -name "a*" 2> /dev/null |cut -d ' ' -f2 | sort |uniq
alina
muhamad
umran
yaroslav
[alekseeva_alina@vpsj3IeQ home]$
```

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 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.

```
[alekseeva_alina@vpsj3IeQ ~]$ sed '1,33d; s/:x:/,/; s/:[0-9]*:/:/; s/:\bin\/ba
sh/' /etc/passwd | cat >accounts.csv
[alekseeva_alina@vpsj3IeQ ~]$ cat 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
kalarashan_kostyantyn,54330,/home/kalarashan_kostyantyn
nosov_andrii,54331,/home/nosov_andrii
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_andrii,54339,/home/yufitenko_andrii
cheremnih_grigoriy,54340,/home/cheremnih_grigoriy
barkar_karina,54341,/home/barkar_karina
loboda_oleksandra,54342,/home/loboda_oleksandra
borovets_vladislav,54343,/home/borovets_vladislav
matnenko_stanislaw,54344,/home/matnenko_stanislaw
bodnar_illya,54345,/home/bodnar_illya
nezhivih_mariya,54346,/home/nejhivih_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
```

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

```
[alekseeva_alina@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
[alekseeva_alina@vpsj3IeQ ~]$
```

2.2 Обробка текстового файлу формату CSV

Варіант 10-2

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/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-1QybKQje49VdjQFahkely1Ou>: – 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","Обсяг виробництва (валовий збір) у початково оприбуткованій масі, ц","Урожайність, з 1 га зібраної площі, ц"
Одеська область,605000,10580379,"17,5"
м.Одеса,місто,-,-,-
м.Балта,місто,18665,610518,32.7
м.Білгород-Дністровський,місто,3838,83690,21.8
м.Біляївка,місто,2717,36229,13.3
м.Ізмаїл,місто,450,3140,7
Ананьївський,район,23306,401449,17.2
Арцизький,район,26603,368101,13.8
Балтський,район,8458,270993,32
Білгород-Дністровський,район,31420,507911,16.2
Біляївський,район,21740,355424,16.3
Березівський,район,33019,566220,17.1
Болградський,район,24985,332300,13.3
Великомихайлівський,район,32136,373948,11.6
Іванівський,район,26606,330236,12.4
Ізмаїльський,район,25826,211119,8.2
Кілійський,район,15280,328931,21.5
Кодимський,район,14372,456355,31.8
Лиманський,район,32072,558800,17.4
Подільський,район,28111,671202,23.9
Окнянський,район,20019,389348,19.4
Лобашівський,район,24828,428324,17.3
Миколаївський,район,10798,255061,23.6
Овідіопольський,район,15167,416772,27.5
Роздільнянський,район,22709,314328,13.8
Ренійський,район,6751,85495,12.7
Савранський,район,12079,436385,36.1
Саратський,район,41304,614652,14.9
Тарутинський,район,31814,573658,18
Татарбунарський,район,13142,184200,14
Захарівський,район,12116,160353,"13,2"
Ширяївський,район,24579,255237,"10,4"

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;

```
Ширяївський, район[alekseeva_alina@vpsj3ieQ ~]$ sed 's/(,)\...1/1/g' table.csv
> table2.csv
[alekseeva_alina@vpsj3ieQ ~]$ cat table2.csv
Назва,місто/район,"Зібрана площа, га","Обсяг виробництва (валовий збір) у початк
ово оприбуткованій масі, ц","Урожайність, га зібраної площі, ц"
Одеська область,,605000,10580379,"17,5"
м.Одеса,місто,-,-,-
м.Валта,місто,18665,610518,32.7
м.Білгород-Дністровський,місто,3838,83690,21.8
м.Біляївка,місто,2717,36229,13.3
м.Ізмаїл,місто,450,3140,7
Ананьївський,район,23306,401449,17.2
Арцизький,район,26603,01,13.8
Балтський,район,8458,270993,32
Білгород-Дністровський,район,31420,507911,16.2
Біляївський,район,21740,355424,16.3
Березівський,район,9,566220,
Болградський,район,24985,332300,13.3
Великомихайлівський,район,32136,373948,11.6
Іванівський,район,26606,330236,12.4
Ізмаїльський,район,25826,19,8.2
Кілійський,район,15280,328931,21.5
Кодимський,район,14372,456355,31.8
Лиманський,район,32072,558800,17.4
Подільський,район,1,671202,23.9
Окнянський,район,9,389348,19.4
Любашівський,район,24828,428324,17.3
Миколаївський,район,10798,255061,23.6
Овідіопольський,район,15167,416772,27.5
Роздільнянський,район,22799,314328,13.8
Ренійський,район,,85495,12.7
Савранський,район,12079,436385,
Саратський,район,41304,614652,14.9
Тарутинський,район,4,573658,18
Татарбунарський,район,13142,184200,14
Захарівський,район,6,160353,"13,2"
```

```
[alekseeva_alina@vpsj3ieQ ~]$ sed 's/(,)\...1/1/g' table.csv > table2.csv
[alekseeva_alina@vpsj3ieQ ~]$ cat table2.csv
Назва,місто/район,"Зібрана площа, га","Обсяг виробництва (валовий збір) у початк
ово оприбуткованій масі, ц","Урожайність, з 1 га зібраної площі, ц"
Одеська область,,605000,10580379,"17,5"
м.Одеса,місто,-,-,-
м.Валта,місто,18665,610518,32.7
м.Білгород-Дністровський,місто,3838,83690,21.8
м.Біляївка,місто,2717,36229,13.3
м.Ізмаїл,місто,450,3140,7
Ананьївський,район,23306,401449,17.2
Арцизький,район,26603,368101,13.8
Балтський,район,8458,270993,32
Білгород-Дністровський,район,31420,507911,16.2
Біляївський,район,21740,355424,16.3
Березівський,район,33019,566220,17.1
Болградський,район,24985,332300,13.3
Великомихайлівський,район,32136,373948,11.6
Іванівський,район,26606,330236,12.4
Ізмаїльський,район,25826,211119,8.2
Кілійський,район,15280,328931,21.5
Кодимський,район,14372,456355,31.8
Лиманський,район,32072,558800,17.4
Подільський,район,28111,671202,23.9
Окнянський,район,20019,389348,19.4
Любашівський,район,24828,428324,17.3
Миколаївський,район,10798,255061,23.6
Овідіопольський,район,15167,416772,27.5
Роздільнянський,район,22799,314328,13.8
Ренійський,район,6751,85495,12.7
Савранський,район,12079,436385,36.1
Саратський,район,41304,614652,14.9
Тарутинський,район,31814,573658,18
Татарбунарський,район,13142,184200,14
Захарівський,район,12116,160353,"13,2"
```

```
[alekseeva_alina@vpsj3ieQ ~]$ cut -f1,3 -d, table.csv > table2.csv
[alekseeva_alina@vpsj3ieQ ~]$ cat table2.csv
Назва,"Зібрана площа
Одеська область,605000
м.Одеса,-
м.Валта,18665
м.Білгород-Дністровський,3838
м.Біляївка,2717
м.Ізмаїл,450
Ананьївський,23306
Арцизький,26603
Балтський,8458
Білгород-Дністровський,31420
Біляївський,21740
Березівський,33019
Болградський,24985
Великомихайлівський,32136
Іванівський,26606
Ізмаїльський,25826
Кілійський,15280
Кодимський,14372
Лиманський,32072
Подільський,28111
Окнянський,20019
Любашівський,24828
Миколаївський,10798
Овідіопольський,15167
Роздільнянський,22799
Ренійський,6751
Савранський,12079
Саратський,41304
Тарутинський,31814
Татарбунарський,13142
Захарівський,12116
Ширяївський,24579
```

- 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;

```
[alekseeva_alina@vps31eq ~]$ join -t, -a 1 -a 2 -o 1,2,2,2 -e '' table1.csv gida_1.csv | cat > text4.csv
join: gida_1.csv:4: is not sorted: м.Балта,21498
join: table2.csv:3: is not sorted: м.Одеса,-
[alekseeva_alina@vps31eq ~]$ cat text4.csv
Назва,"Зібрана площа,Зібрана площа
,,га
м.Одеса,,
м.Балта,,21498
м.Вінгород-Дністровський,,4076
м.Вілявка,,3106
м.Ізмаїл,,608
м.Подільськ,,
Ананьївський,,23829
Арцизький,,31751
Валтський,,10810
Вінгород-Дністровський,,33227
Вілявський,,23298
Верев'яський,,36266
Волградський,,30735
Великомиколайківський,,33570
Ікманівський,,27603
Ізмаїльський,,28958
Кілійський,,15712
Кодимський,,16586
Леманський,,32072
Одеська область,605000,
м.Одеса,-,
м.Балта,18665,
м.Вінгород-Дністровський,3838,
м.Вілявка,2717,
м.Ізмаїл,450,
Ананьївський,23306,
Арцизький,24603,
Валтський,8458,
Вінгород-Дністровський,31420,
Вілявський,21740,
Верев'яський,33819,
Волградський,24985,
Великомиколайківський,32136,
Ікманівський,26606,
Ізмаїльський,25826,
Кілійський,15280,
Кодимський,14372,
Леманський,32072,
```

- 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.

```
[alekseeva_alina@vps31eq ~]$ sort --field-separator="," --key=3 -r text4.csv
Назва,"Зібрана площа,Зібрана площа
,,га
Кілійський,6751,7076
м.Ізмаїл,,608
Саратський,41304,42845
м.Вінгород-Дністровський,,4076
Верев'яський,,36266
Тарутинський,31814,34116
Леманський,,32772
Великомиколайківський,,33570
Вінгород-Дністровський,,33227
Арцизький,,31751
м.Вілявка,,3106
Волградський,,30735
Ширівський,24579,29397
Подільський,28111,28959
Ізмаїльський,,28958
Львівський,24828,27997
Ікманівський,,27603
Роздільнянський,22799,24867
Ананьївський,,23829
Вілявський,,23298
Омнянський,20019,21815
м.Балта,,21498
Татарбунарський,13142,18013
Овідіопольський,15167,16970
Кодимський,,16586
Савранський,12079,16114
Кілійський,,15712
Миколайівський,10798,14046
Захарівський,12116,13233
Валтський,,10810
Одеська область,605000,
м.Подільськ,,
м.Одеса,-,
м.Одеса,,
м.Ізмаїл,450,
м.Вілявка,2717,
м.Вінгород-Дністровський,3838,
м.Балта,18665,
Леманський,32072,
Кодимський,14372,
Кілійський,15280,
Ізмаїльський,25826,
```

Висновок: в ході цієї лабораторної роботи ми придбали навички складної обробки текстових даних роботи засобами оболонки Unix-подібних ОС інтерфейсу командного рядка.