

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-202

Сідельніков М. В.

Перевірив:

Блажко О. А.

Одеса-2021

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.
 - 1.1. Виведіть на екран перші десять назв каталогів або файлів, назви яких починаються з перших трьох букв вашого прізвища в латиниці
 - 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.
 - 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*).
 - 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.
 - 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.
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/1ITMLdjcw7nboUo9nFw4NwkWcvjbZeZZlDVb2Hg4laUQ>.

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

- 2.1. 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.

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

- 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 третього;

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

- 2.5. отримання CSV-файлу з GitHub-репозиторій, створеного іншим учасником команди;

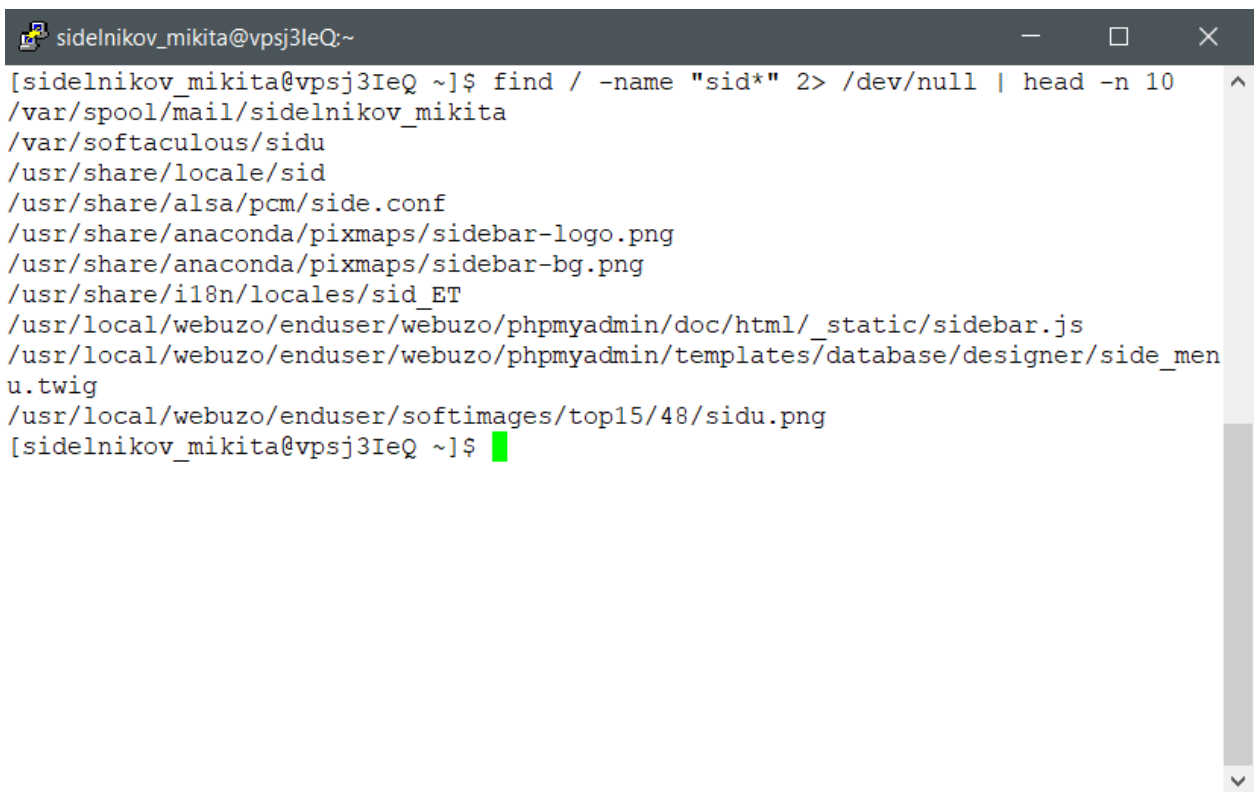
- 2.6. об'єднання двох файлів в один за першим стовпчиком;

- 2.7. 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.

Хід роботи

1. Обробка текстових даних результатів роботи команд. Виконайте завдання, знаходячись на Linux-сервері за адресою з IP = 91.219.60.189.
 - 1.1. Виведіть на екран перші десять назв каталогів або файлів, назви яких починаються з перших трьох букв вашого прізвища в латиниці.

A screenshot of a terminal window with a dark background. The title bar shows the user 'sidelnikov_mikita' and the host 'vpsj3IeQ'. The terminal displays the command `find / -name "sid*" 2> /dev/null | head -n 10` and its output, which lists ten files and directories starting with 'sid'. The output is as follows:

```
[sidelnikov_mikita@vpsj3IeQ ~]$ find / -name "sid*" 2> /dev/null | head -n 10
/var/spool/mail/sidelnikov_mikita
/var/softaculous/sidu
/usr/share/locale/sid
/usr/share/alsa/pcm/side.conf
/usr/share/anaconda/pixmaps/sidebar-logo.png
/usr/share/anaconda/pixmaps/sidebar-bg.png
/usr/share/i18n/locales/sid_ET
/usr/local/webuzo/enduser/webuzo/phpmyadmin/doc/html/_static/sidebar.js
/usr/local/webuzo/enduser/webuzo/phpmyadmin/templates/database/designer/side_menu.twig
/usr/local/webuzo/enduser/softimages/top15/48/sidu.png
```

The prompt `[sidelnikov_mikita@vpsj3IeQ ~]$` is visible at the bottom of the terminal output.

- 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.

```
sidelnikov_mikita@vpsj3IeQ:~  
[sidelnikov_mikita@vpsj3IeQ ~]$ find /home -name "s*" -maxdepth 1 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  
[sidelnikov_mikita@vpsj3IeQ ~]$
```

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*).

```
sidelnikov_mikita@vpsj3IeQ:~  
[sidelnikov_mikita@vpsj3IeQ ~]$ find /home -name "s*" -maxdepth 1 2> /dev/null | cut -b 7- | sort | uniq  
shapovalova_viktoriya  
sherbakov_artur  
sherbakov_georgij  
shostak_roman  
shulyak_mikola  
sidelnikov_mikita  
sinyakov_igor  
sirenko_mariya  
sivash_andrij  
soft-1  
sokolov_oleksij  
stepanenko_gleb  
strinaglyuk_ivan  
sultanov_artem  
svetashov_danilo  
[sidelnikov_mikita@vpsj3IeQ ~]$
```

- 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.

```
sidelnikov_mikita@vpsj3leQ:~  
[sidelnikov_mikita@vpsj3leQ ~]$ cut -f 1,4,6 -d: /etc/passwd | sed 's:/,/g' > accounts.csv  
[sidelnikov_mikita@vpsj3leQ ~]$ cat accounts.csv  
root,0,/root  
bin,1,/bin  
daemon,2,/sbin  
adm,4,/var/adm  
lp,7,/var/spool/lpd  
sync,0,/sbin  
shutdown,0,/sbin  
halt,0,/sbin  
mail,12,/var/spool/mail  
operator,0,/root  
games,100,/usr/games  
ftp,50,/var/ftp  
nobody,99,/var/ftp  
systemd-network,192,/var/lib/systemd/network  
dbus,81,/var/lib/dbus  
polkitd,997,/var/lib/polkitd  
postfix,89,/var/spool/postfix  
chrony,996,/var/lib/chrony  
sshd,74,/var/empty/sshd  
soft,1000,/home/soft  
sasauth,76,/run/sasauthd  
mailnull,47,/var/spool/mqueue
```

- 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.

```
sidelnikov_mikita@vpsj3leQ:~  
[sidelnikov_mikita@vpsj3leQ ~]$ grep [0-9][0-9][0-9]5[0-9] accounts.csv  
matnenko_stanislaw,54350,/home/matnenko_stanislaw  
bodnar_ilya,54351,/home/bodnar_ilya  
nezhivih_mariya,54352,/home/nezhivih_mariya  
grebenik_anzhelika,54353,/home/grebenik_anzhelika  
lobko_daniil,54354,/home/lobko_daniil  
baranyuk_dmitro,54355,/home/baranyuk_dmitro  
kovach_dmitro,54356,/home/kovach_dmitro  
sidelnikov_mikita,54357,/home/sidelnikov_mikita  
lukashak_daniil,54358,/home/lukashak_daniil  
markovskij_danilo,54359,/home/markovskij_danilo  
[sidelnikov_mikita@vpsj3leQ ~]$
```

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/1ITMLdjcw7nboUo9nFw4NwkWcvjbZeZZlDVb2Hg4laUQ>.

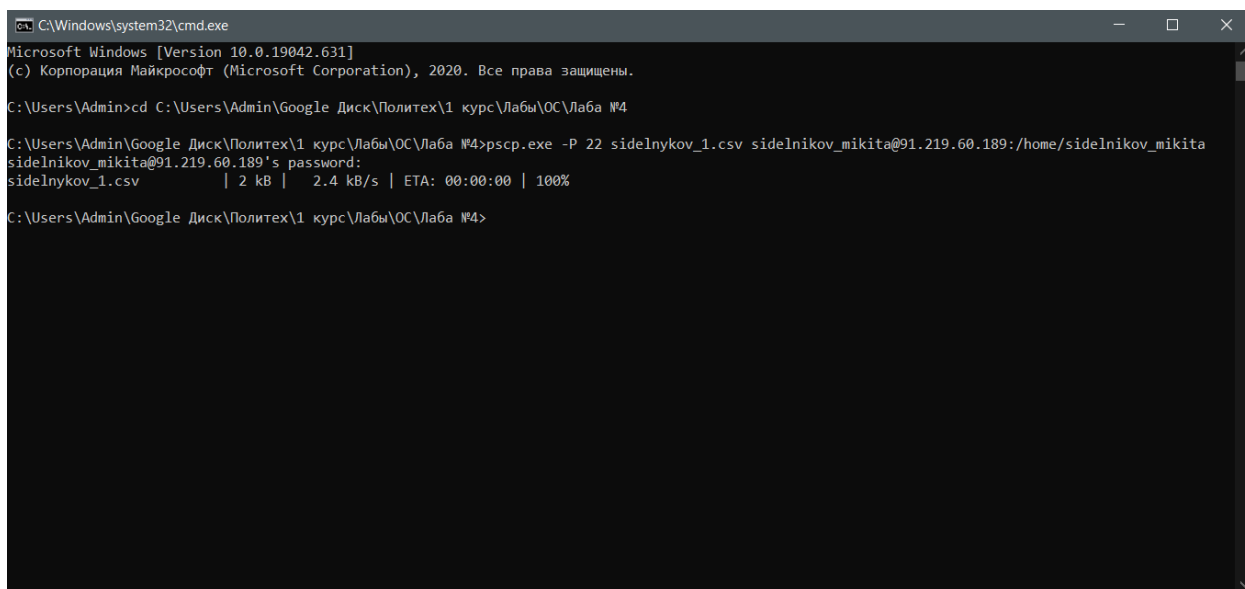
Необхідно створити файл формату CSV, який містить табличні данні, використовуючи наступні кроки:

- 2.1. 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.

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



```
C:\Windows\system32\cmd.exe
Microsoft Windows [Version 10.0.19042.631]
(c) Корпорация Майкрософт (Microsoft Corporation), 2020. Все права защищены.

C:\Users\Admin>cd C:\Users\Admin\Google Диск\Политех\1 курс\Лабы\ОС\Лаба №4

C:\Users\Admin\Google Диск\Политех\1 курс\Лабы\ОС\Лаба №4>pscp.exe -P 22 sidelnikov_1.csv sidelnikov_mikita@91.219.60.189:/home/sidelnikov_mikita
sidelnikov_mikita@91.219.60.189's password:
sidelnikov_1.csv      | 2 kB |   2.4 kB/s | ETA: 00:00:00 | 100%

C:\Users\Admin\Google Диск\Политех\1 курс\Лабы\ОС\Лаба №4>
```

- 2.3. модифікація файлу CSV-формату:

- заміна в числах коми на точку з одночасним видаленням лапки;

```
sidelnikov_mikita@vpsj3IeQ:~  
[sidelnikov_mikita@vpsj3IeQ ~]$ sed 's/\([0-9]\+\)[,]\([0-9]\+\)\"/\\1\\.\\2 /g' sidelnikov_1.csv > sidelnikov_2.csv  
[sidelnikov_mikita@vpsj3IeQ ~]$ cat sidelnikov_2.csv  
,"Обсяги викидів, т",У % до 2018р., "діоксиду сірки, т",діоксиду сірки у % до 2018р., "діоксиду азоту, т",діоксиду азоту у % до 2018р.  
Одеська область,33106.20,88.50,1482.80,90.80,1252.80,54.90  
м.Одеса,20120.90,109.90,206.80,80.60,601.40,87.40  
м.Балта,52.60,38.00,33.10,198.70,2.90,46.90  
м.Білгород-Дністровський,449.70,69.80,35.50,46.90,14.70,83.70  
м.Біляївка,45.70,102.40,12.80,94.30,2.20,149.50  
м.Ізмаїл,831.50,41.90,102.20,93.60,43.60,92.00  
м.Чорноморськ,1605.90,99.20,131.60,97.90,83.40,72.20  
м.Подольськ,394.60,23.70,25.30,480.10,13.70,66.60  
м.Теплодар,48.50,79.60,,,16.00,80.00  
м.Кюне,1900.70,120.60,123.90,90.00,309.40,140.40  
райони,,,,,  
Ананьївський,315.80,324.40,0.10,81.60,25.00,4361.00  
Арцизький,254.00,41.70,85.50,69.50,3.40,43.60  
Балтський,18.70,98.10,6.60,131.90,1.00,81.80  
Білгород-Дністровський,164.50,86.70,27.70,76.50,9.50,158.00  
Біляївський,151.00,28.40,2.40,6.20,4.20,1.70  
Березівський,70.30,15.50,6.10,78.10,1.10,92.40  
Болградський,1150.30,123.70,35.80,4037.60,2.30,1149.00  
Великомихайлівський,239.00,86.90,0.40,2.10,0.40,23.80  
Іванівський,24.20,64.50,4.30,60.20,1.10,62.00
```

- заміна підрядка « ... 1» на порожнє значення (якщо таке є);

```
sidelnikov_mikita@vpsj3IeQ:~  
[sidelnikov_mikita@vpsj3IeQ ~]$ sed 's/\(,\)\...1\\1/g' sidelnikov_2.csv > sidelnikov_3.csv  
[sidelnikov_mikita@vpsj3IeQ ~]$ cat sidelnikov_3.csv  
,"Обсяги викидів, т",У % до 2018р., "діоксиду сірки, т",діоксиду сірки у % до 2018р., "діоксиду азоту, т",діоксиду азоту у % до 2018р.  
Одеська область,33106.20,88.50,1482.80,90.80,1252.80,54.90  
м.Одеса,20120.90,109.90,206.80,80.60,601.40,87.40  
м.Балта,52.60,38.00,0,198.70,2.90,46.90  
м.Білгород-Дністровський,449.70,69.80,35.50,46.90,14.70,83.70  
м.Біляївка,45.70,102.40,12.80,94.30,2.20,149.50  
м.Ізмаїл,831.50,41.90,102.20,93.60,43.60,92.00  
м.Чорноморськ,1605.90,99.20,131.60,97.90,83.40,72.20  
м.Подольськ,394.60,23.70,25.30,480.10,13.70,66.60  
м.Теплодар,48.50,79.60,,,16.00,80.00  
м.Кюне,1900.70,120.60,123.90,90.00,309.40,140.40  
райони,,,,,  
Ананьївський,315.80,324.40,0.10,81.60,25.00,.00  
Арцизький,254.00,41.70,85.50,69.50,3.40,43.60  
Балтський,18.70,0,6.60,131.90,1.00,81.80  
Білгород-Дністровський,164.50,86.70,27.70,76.50,9.50,158.00  
Біляївський,151.00,28.40,2.40,6.20,4.20,1.70  
Березівський,70.30,15.50,6.10,0,1.10,92.40  
Болградський,1150.30,123.70,35.80,4037.60,2.30,1149.00  
Великомихайлівський,239.00,86.90,0.40,2.10,0.40,23.80  
Іванівський,24.20,64.50,4.30,60.20,1.10,62.00
```

- заміна символу – (дефіс) на порожнє значення (якщо таке є);


```

sidelnikov_mikita@vpsj3leQ:~
[sideInikov_mikita@vpsj3leQ ~]$ sed 's/(,)\|~/\1/g' sidelnikov_3.csv > sidelnikov_4.csv
[sideInikov_mikita@vpsj3leQ ~]$ cat sidelnikov_4.csv
,"Обсяги викидів, т",У % до 2018р., "діоксиду сірки, т",діоксиду сірки у % до 2018р., "діоксиду азоту, т",діоксиду азоту у % до 2018р.
Одеська область,33106.20,88.50,1482.80,90.80,1252.80,54.90
м.Одеса,20120.90,109.90,206.80,80.60,601.40,87.40
м.Балта,52.60,38.00,0,198.70,2.90,46.90
м.Білгород-Дністровський,449.70,69.80,35.50,46.90,14.70,83.70
м.Біляївка,45.70,102.40,12.80,94.30,2.20,149.50
м.Ізмаїл,831.50,41.90,102.20,93.60,43.60,92.00
м.Чорноморськ,1605.90,99.20,131.60,97.90,83.40,72.20
м.Подольськ,394.60,23.70,25.30,480.10,13.70,66.60
м.Теплодар,48.50,79.60,,16.00,80.00
м.Кіжне,1900.70,120.60,123.90,90.00,309.40,140.40
райони,,,,,
Ананьївський,315.80,324.40,0.10,81.60,25.00,.00
Арцизький,254.00,41.70,85.50,69.50,3.40,43.60
Балтський,18.70,0,6.60,131.90,1.00,81.80
Білгород-Дністровський,164.50,86.70,27.70,76.50,9.50,158.00
Біляївський,151.00,28.40,2.40,6.20,4.20,1.70
Березівський,70.30,15.50,6.10,0,1.10,92.40
Болградський,1150.30,123.70,35.80,4037.60,2.30,1149.00
Великомихайлівський,239.00,86.90,0.40,2.10,0.40,23.80
Іванівський,24.20,64.50,4.30,60.20,1.10,62.00

```

— видалення з файлу стовпчиків окрім першого та третього;

```

sidelnikov_mikita@vpsj3leQ:~
[sideInikov_mikita@vpsj3leQ ~]$ cut -f1,3 -d, sidelnikov_4.csv > sidelnikov_5.csv
[sideInikov_mikita@vpsj3leQ ~]$ cat sidelnikov_5.csv
," т"
Одеська область,88.50
м.Одеса,109.90
м.Балта,38.00
м.Білгород-Дністровський,69.80
м.Біляївка,102.40
м.Ізмаїл,41.90
м.Чорноморськ,99.20
м.Подольськ,23.70
м.Теплодар,79.60
м.Кіжне,120.60
райони,
Ананьївський,324.40
Арцизький,41.70
Балтський,0
Білгород-Дністровський,86.70
Біляївський,28.40
Березівський,15.50
Болградський,123.70
Великомихайлівський,86.90
Іванівський,64.50
Ізмаїльський,72.80

```

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

```

sidelnikov_mikita@vpsj3leQ:~
[sideInikov_mikita@vpsj3leQ ~]$ git add sidelnikov_5.csv
[sideInikov_mikita@vpsj3leQ ~]$ git commit -m "sidelnikov - csv file"
[master (root-commit) 1e9da4c] sidelnikov - csv file
1 file changed, 38 insertions(+)
create mode 100644 sidelnikov_5.csv

```

```
sidelnikov_mikita@vpsj3IeQ:~  
hint: See the 'Note about fast-forwards' in 'git push --help' for details.  
[sidelnikov_mikita@vpsj3IeQ ~]$ git pull origin main  
From https://github.com/DanBearLuk/Operating-System.-Laboratory-Work-1  
* branch      main      -> FETCH_HEAD  
Merge made by the 'recursive' strategy.  
LukashakDaniil.pdf | Bin 0 -> 294904 bytes  
LukashakDaniilLab3.pdf | Bin 0 -> 628565 bytes  
Lukashak_Daniil.txt | 0  
SidelnikovNikita.pdf | Bin 0 -> 670286 bytes  
Sidelnikov_Nikita.txt | 0  
lab.pptx | Bin 0 -> 2844738 bytes  
6 files changed, 0 insertions(+), 0 deletions(-)  
create mode 100644 LukashakDaniil.pdf  
create mode 100644 LukashakDaniilLab3.pdf  
create mode 100644 Lukashak_Daniil.txt  
create mode 100644 SidelnikovNikita.pdf  
create mode 100644 Sidelnikov_Nikita.txt  
create mode 100644 lab.pptx  
[sidelnikov_mikita@vpsj3IeQ ~]$ git push origin main  
Username for 'https://github.com': HighstoneGD  
Password for 'https://HighstoneGD@github.com':  
Counting objects: 4, done.  
Delta compression using up to 2 threads.  
Compressing objects: 100% (2/2), done.
```

2.5. отримання CSV-файлу з GitHub-репозиторій, створеного іншим учасником команди;

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

```
sidelnikov_mikita@vpsj3IeQ:~  
[sidelnikov_mikita@vpsj3IeQ ~]$ join -t, -j1 1 -j2 1 sidelnikov_5.csv lukashak5.csv > common.csv  
join: lukashak5.csv:2: is not sorted: м.Одеса,390722.9  
join: sidelnikov_5.csv:3: is not sorted: м.Одеса,109.90  
[sidelnikov_mikita@vpsj3IeQ ~]$ cat common.csv  
Роздільнянський,10.60,903.6  
Ренійський,20.40,3469.2  
Савранський,116.00,166.8  
Саратський,45.50,49.6  
Тарутинський,0,8395.6  
Татарбунарський,35.70,250  
Захарівський,,426  
Ширяївський,101.60,125
```

2.7. 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.

```
sidelnikov_mikita@vpsj3IeQ:~  
[sidelnikov_mikita@vpsj3IeQ ~]$ sort -t, -rnk2 common.csv  
Тарутинський,0,8395.6  
Захарівський, ,426  
Савранський,116.00,166.8  
Ширяївський,101.60,125  
Саратський,45.50,49.6  
Татарбунарський,35.70,250  
Ренійський,20.40,3469.2  
Роздільнянський,10.60,903.6  
[sidelnikov_mikita@vpsj3IeQ ~]$
```

Висновки: під час виконання лабораторної роботи було придбано навичок складної обробки текстових даних роботи засобами оболонки Unix-подібних ОС інтерфейсу командного рядка. Найскладнішими виявились 1.4 (створення файлу accounts) та робота з github.