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:

ст. гр. AI-201

Малофеєв Д. Ю.

Перевірив:

Блажко О. А.

Одеса – 2021

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

**Завдання для виконання:**

**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 другого

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

2.5) отримання CSV-файлу, створеного іншим учасником команди;

2.6) об`єднання двох файлів в один за першим стовпчиком;

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.

**Хід роботи:**

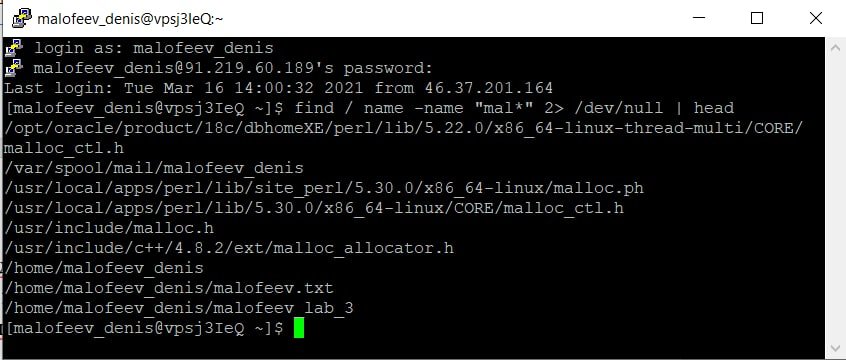
1.1) Виводять на екран перші десять назв каталогів або файлів, назви яких починаються з перших трьох букв прізвища в латиниці – bel (рис. 1.1)

Рис. 1.1

1.2) Назва домашнього каталогу містить прізвище та ім'я в транслітерації. Виводять на екран список назв домашніх каталогів користувачів, в назві яких присутня перша буква прізвища – b (рис. 1.2)

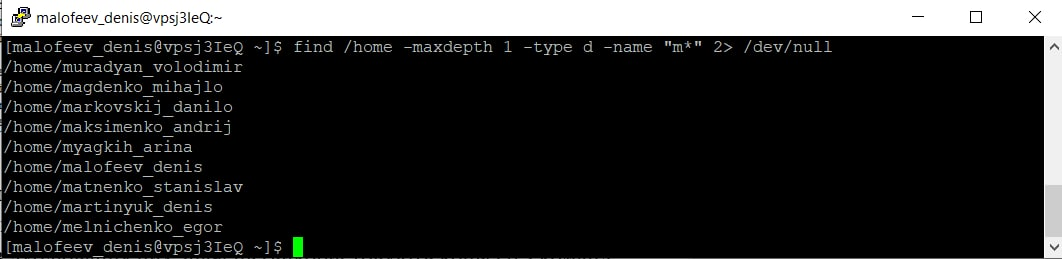


Рис. 1.2

1.3) Модифікують рішення попереднього завдання так, щоб на екран було виведено імена користувачів, відсортовані в порядку зростання значення, та без дублювання цих значень (використовують конвеєр з команд cut, sort, uniq).

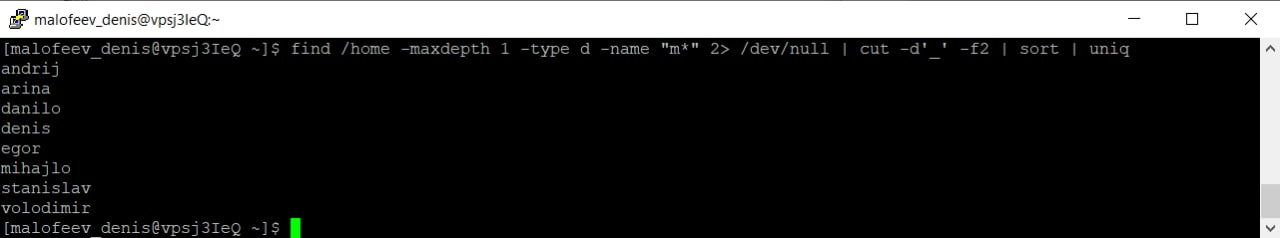
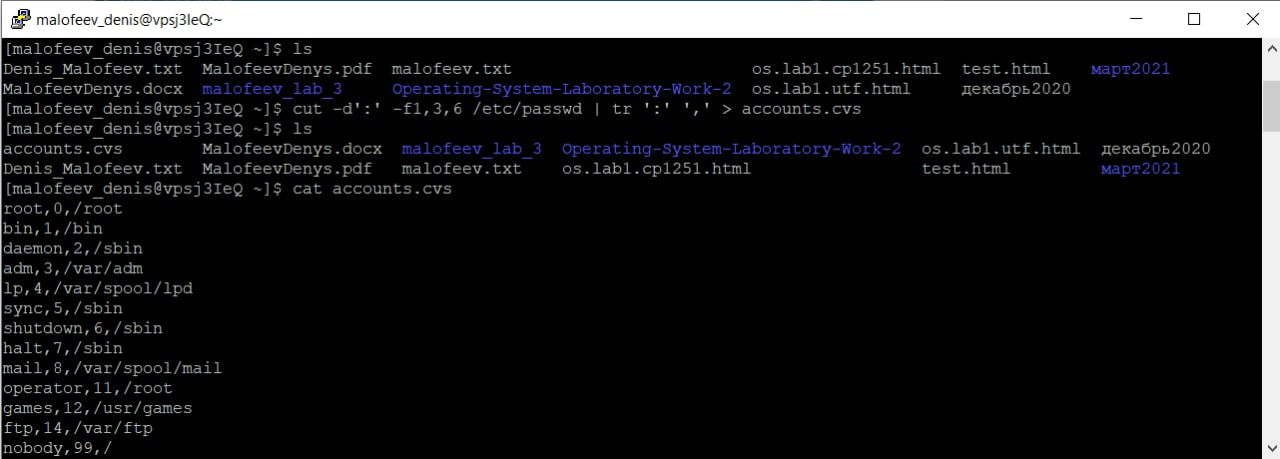
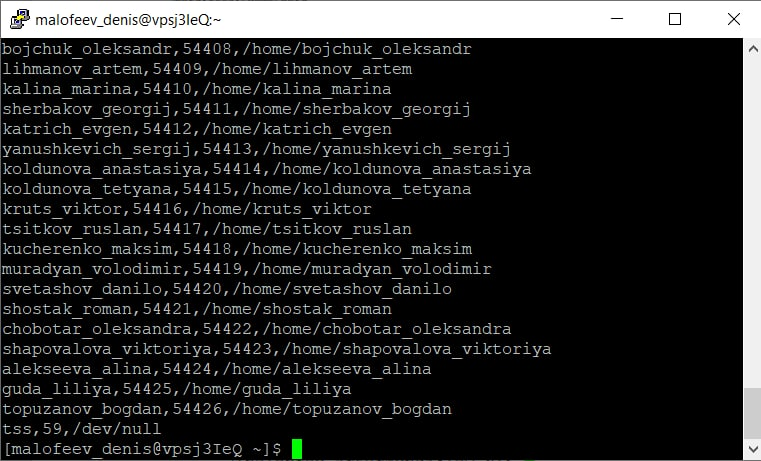


Рис. 1.3

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. Створюють 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.

Рис 1.4

1.5) З файлу accounts.csv отримують на екран перелік облікових записів, в яких четверта цифра ідентифікатора користувача співпадає із власною четвертою цифрою ідентифікатора користувача (54388).

Рис. 1.5

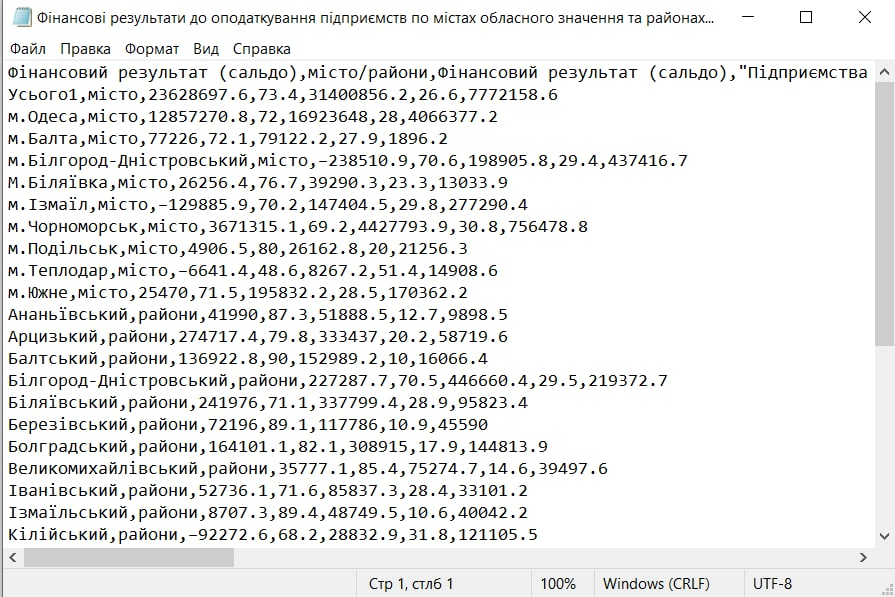
2. Обробка текстового файлу формату CSV

2.1) Створюють файл CSV-формату, використовуючи Google Docs за методичними рекомендаціями, із наступними характеристиками (рис. 2.1):

− кодування символів – UTF-8

− роздільник між колонками – символ коми;

− назви файлів визначити з урахуванням транслітерації свого прізвища та номеру файлу.

Рис 2.1

2.2) Копіюють файл на сервер за адресою з IP = 91.219.60.189 у домашній каталог (рисю 2.2)

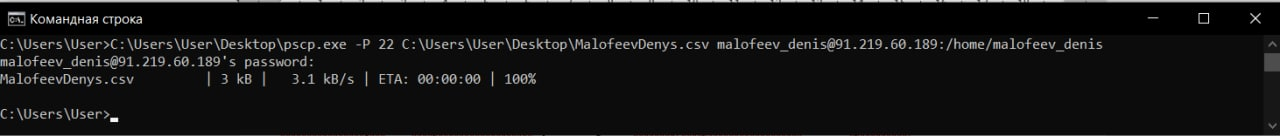


Рис. 2.2

2.3) модифікація файлу CSV-формату (рис. 2.3):

− 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

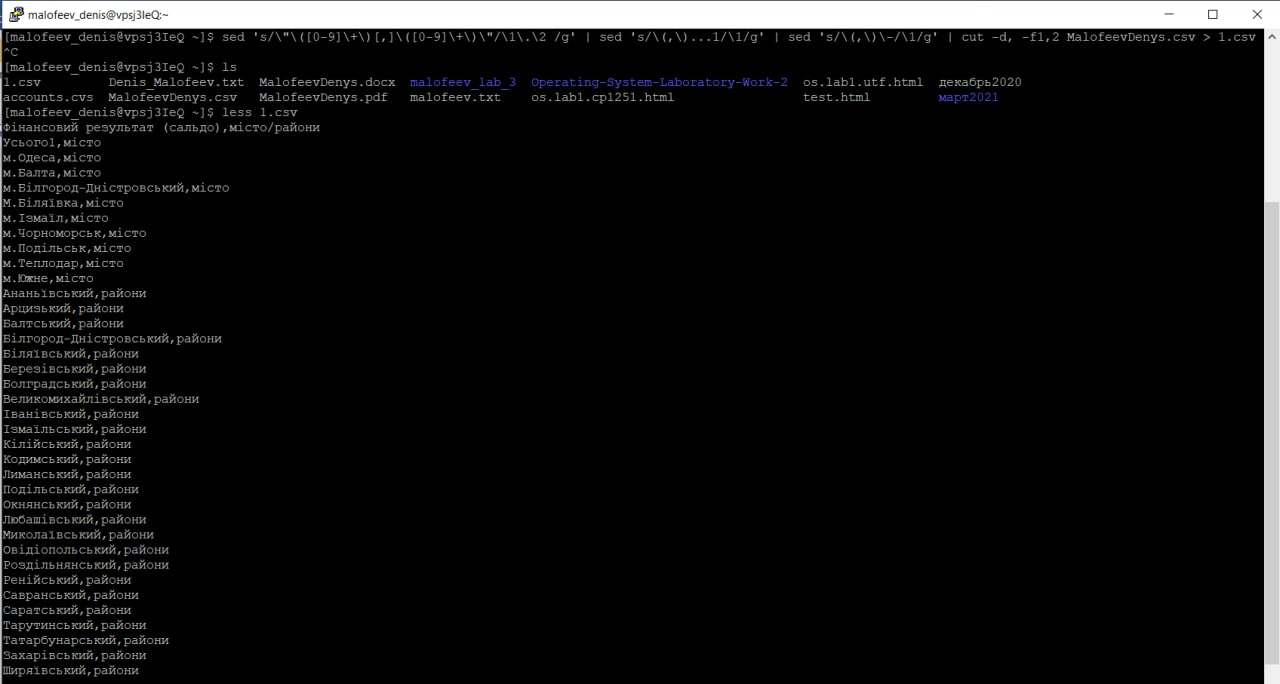


Рис 2.3

2.4) передача створеного CSV-файлу в GitHub-репозиторій (рис 2.4);

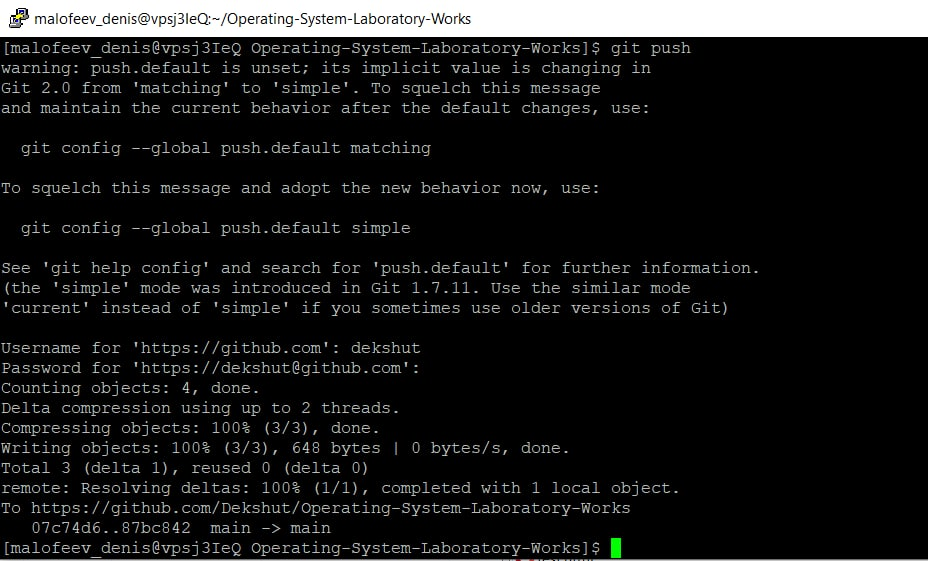


Рис. 2.4

2.5) отримання CSV-файлу, створеного іншим учасником команди (рис 2.5)

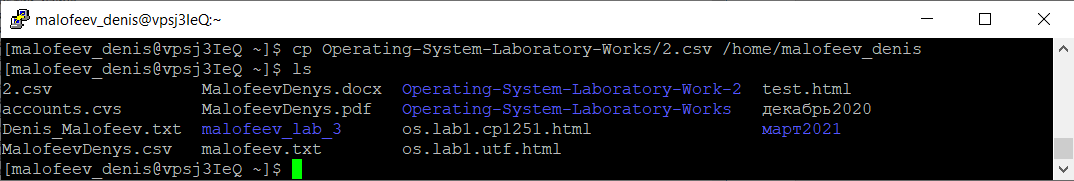


Рис 2.5

2.6) об`єднання двох файлів в один за першим стовпчиком (рис 2.6);

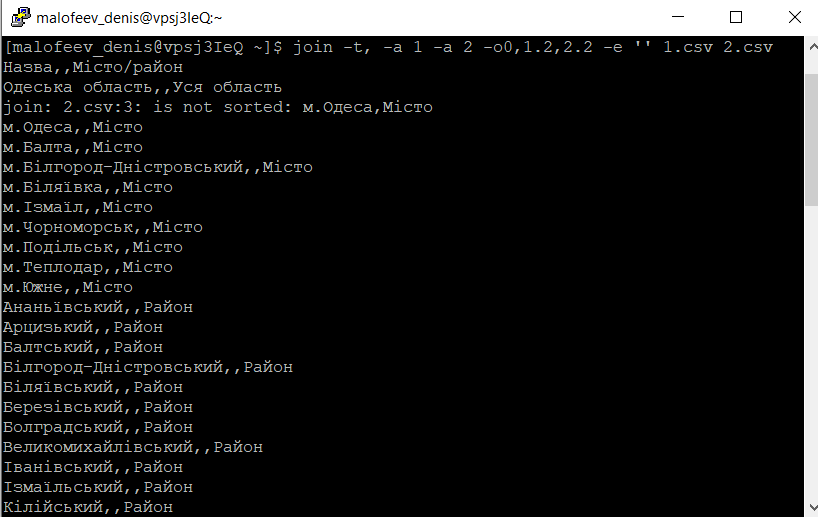


Рис. 2.6

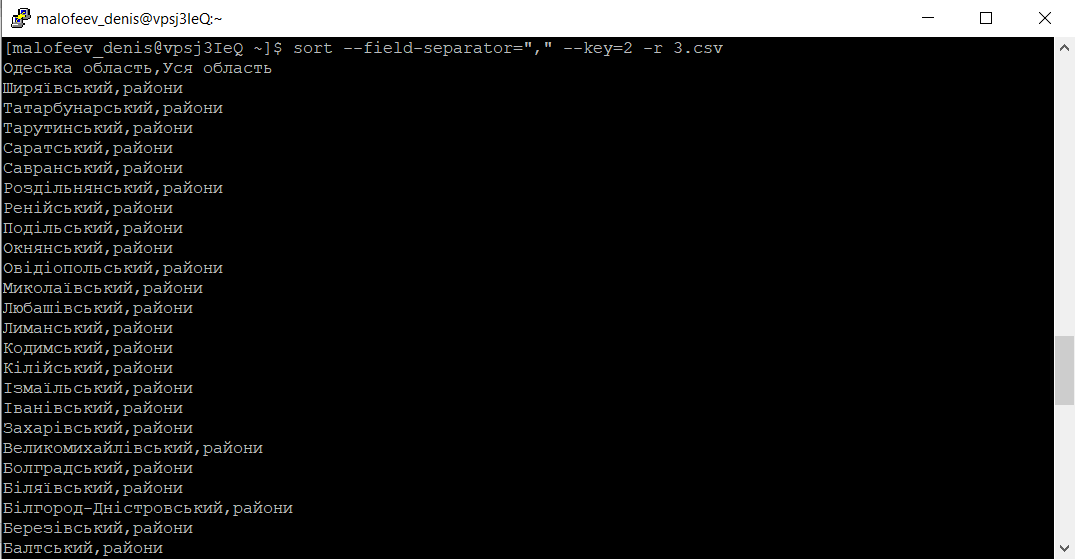
2.7) виведення на екран змісту файлу з рішення попереднього завдання у зворотному напрямі для колонки за номером 2 - для 1-го учасника команди, або за номером 3 – для 2-го учасника команди (рис 2.7).

Рис 2.7

**Висновки:** в ході виконання лабораторної роботи, були придбані навички складної обробки текстових даних роботи засобами оболонки Unix-подібних ОС інтерфейсу командного рядка. Найскладнішим завданням було склеювання двох csv файлів.