**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**

**ХАРКІВСЬКИЙ ПОЛІТЕХНІЧНИЙ ФАХОВИЙ КОЛЕДЖ**

**Циклова комісія інформаційних технологій**

**КУРСОВИЙ ПРОЕКТ**

**(РОБОТА)**

Програмування

(назва дисципліни)

на тему: аналіз даних з соціальних медіа

**Студента ІV курсу 407-АКIТ групи**

**спеціальності 151 Автоматизація та комп’ютерно-інтегровані технології**

спеціалізації 5.151.1 Обслуговування інтелектуальних інтегрованих систем

Горобець М.О.

Керівник: викладач ***Васильєв М.В.***

Національна шкала \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Кількість балів: \_\_\_\_ Оцінка: ECTS \_\_\_\_\_

Члени комісії \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ (Васильєв М.В.)

(підпис)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ (\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_)

(підпис)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ (\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_)

(підпис)

м. Харків – 2024 рік

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**

**ХАРКІВСЬКИЙ ПОЛІТЕХНІЧНИЙ ФАХОВИЙ КОЛЕДЖ**

|  |  |  |  |
| --- | --- | --- | --- |
| **РОЗГЛЯНУТО** | | **ЗАТВЕРДЖЕНО** | |
| на засіданні циклової комісії | | Зав. відділення | |
| Голова циклової комісії | |  | Олеся Твердохлібова |
| Марина ВЕЛИЧКО | |  | «31» серпня 2023 р. |
| «31» серпня 2023р. |  |  |  |

**Завдання**

на курсовий проєкт студенту гр. 407-АКІТ

спеціальність/освітньо-професійна програма 151 Автоматизація та комп’ютерно-інтегровані технології/5.151.1 Обслуговування

інтелектуальних інтегрованих систем

***Горобцю Максиму Олександровичу***

Тема: Аналіз соціальних медіа.

Термін виконання роботи з 18.11.2023 р. по 4.04.2024 р.

Вхідні дані для проектування: (Варіант №5)

Розробка програми для аналізу даних з соціальних медіа. Для зберігання інформації використовується CSV-файл, де міститься інформація про публікації, лайки, коментарі та репости. Програма повинна проводити аналіз популярності та взаємодії з контентом.

**Література та посібники для проектування:**

1. Head-First Python, 2nd edition: Paul Barry. - Sebastopol, California, U.S.: O'Reilly Media, 2016. – 622 с.
2. Think Python: How to Think Like a Computer Scientist, 2nd edition: Allen B. Downey. - Sebastopol, California, U.S.: O'Reilly Media, 2015. – 292 с.
3. Clean Code: A Handbook of Agile Software Craftsmanship: Robert C. Martin. - London, England: Pearson, 2008. – 464 c.
4. Python.org: веб-сайт. URL: https://www.python.org (дата звернення: 02.01.2023)
5. Python Tutorial: веб-сайт. URL: https://www.w3schools.com/python/ (дата звернення: 02.01.2023)
6. Learn to become a modern Python developer: веб-сайт. URL: https://roadmap.sh/python/ (дата звернення: 02.01.2023)

**Зміст розрахунково-пояснювальної записки**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| № п/п | Зміст | Планований термін виконання | Фактичний термін виконання | % |
|  | Вступ | 22.11.2023 | 22.11.2023 | 5 |
|  | Аналіз задачі, засобів та методів її вирішення | 30.11.2023 | 30.11.2023 | 15 |
|  | Проектування загального алгоритму роботи програми | 18.01.2024 | 18.01.2024 | 40 |
|  | Розробка програмного забезпечення | 01.03.2024 | 01.03.2024 | 80 |
|  | Керівництво користувача | 14.03.2024 | 14.03.2024 | 90 |
|  | Висновки | 28.03.2024 | 28.03.2024 | 95 |
|  | Список використаних джерел | 02.04.2024 | 02.04.2024 | 100 |

Керівник курсової роботи \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Микола Васильєв

Завдання до курсової роботи

Одержав(ла) студент(ка) гр. 407-АКІТ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_(\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_)

Дата «18» листопада 2023 р.

**Реферат**

Пояснювальна записка: 39 сторінок, 20 рисунків, 4 використаної літератури.

Ключові слова: ПРОГРАМУВАННЯ, ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ, PYTHON, CSV,АНАЛIЗ.

Об’єктом розробки є програмне забезпечення для аналiзу даних з соціальних медіа.

Мета розробки – розробка програмного забезпечення.

* процесі роботи проведено розробку блок-схеми алгоритму роботи програми та окремих функцій, розроблено програму, яка виконує алгоритм роботи.

Основні конструктивні та техніко-економічні показники: висока надійність, зручність, простота, масштабованість.

**Перелік використаних скорочень**

* ООП – Об’єктно-орiєнтоване програмування
* ПЗ – програмне забезпечення
* CSV – Comma-separated values
* VS Code – Visual Studio Code

**ЗМIСТ**

[Реферат 2](#_Toc159008429)

[Перелік використаних скорочень 3](#_Toc159008430)

[ВСТУП 5](#_Toc159008431)

[1. АНАЛІЗ ЗАДАЧІ, ЗАСОБІВ ТА МЕТОДІВ ЇЇ ВИРІШЕННЯ 7](#_Toc159008432)

[1.1 Розумiння вимог 7](#_Toc159008433)

[1.2 Аналіз вихідних даних 8](#_Toc159008434)

[1.3 Вибір технологій 10](#_Toc159008435)

[2. ПРОЕКТУВАННЯ ЗАГАЛЬНОГО АЛГОРИТМУ РОБОТИ ПРОГРАМИ 13](#_Toc159008436)

[3. РОЗРОБКА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ 18](#_Toc159008437)

[3.1 Розробка головного класу Post 18](#_Toc159008438)

[3.2 Класи для тестування 21](#_Toc159008439)

[3.3 Клас CSV\_Reader 22](#_Toc159008440)

[3.4 Клас Bcolors 24](#_Toc159008441)

[3.5 Сортування та клас Sorter 25](#_Toc159008442)

[3.6 Основний файл виконання програми 26](#_Toc159008443)

[4. КЕРІВНИЦТВО КОРИСТУВАЧА 28](#_Toc159008444)

[4.1 Основний функцiонал 28](#_Toc159008445)

[4.3 Iмпортування файлу 30](#_Toc159008446)

[4.4 Робота з файлом 31](#_Toc159008447)

[ВИСНОВКИ 33](#_Toc159008448)

[ЛIТЕРАТУРА 34](#_Toc159008449)

[ДОДАТКИ 35](#_Toc159008450)

[Додаток А 35](#_Toc159008451)

[Додаток Б 36](#_Toc159008452)

[Додаток В 39](#_Toc159008453)

**ВСТУП**

У наш час програмування є однією з найбільш динамічно розвиваючихся галузей, що відображає світові тенденції технологічного прогресу та потреби суспільства. Однією з найактуальніших галузей програмування є аналіз даних з соціальних медіа. Соціальні платформи, такі як Facebook, Instagram, Twitter та інші, стали не просто майданчиками для спілкування, але й важливим інструментом для бізнесу, політики, науки та культури. Вони об'єднують мільйони користувачів з усього світу, дозволяючи їм обмінюватися інформацією, думками та переживаннями.

Актуальність розробки програмного забезпечення для аналізу даних з соціальних медіа визначається не лише через загальний розвиток цих платформ, але й через необхідність ефективного управління контентом та взаємодії з аудиторією. Компанії, маркетологи, політики та дослідники активно використовують дані з соціальних медіа для аналізу тенденцій, прогнозування поведінки аудиторії, а також для створення та вдосконалення своїх стратегій.

Метою нашої роботи є розробка програми для аналізу даних з соціальних медіа, спрямованої на вивчення популярності та взаємодії з контентом. Важливими завданнями програми є виявлення ключових трендів, аналіз залученості аудиторії та визначення ефективності контенту.

Галузь застосування програми включає маркетингові дослідження, наукові аналізи, стратегії управління контентом та багато іншого. Розробка стане корисним інструментом для студентів, дослідників та фахівців у сфері соціальних медіа, сприяючи їхньому подальшому професійному розвитку та науковим досягненням.

Програмне забезпечення було створене в Visual Studio Code, інструменті для розробки, редагування та налагодження сучасних веб-застосунків і програм для хмарних систем. Visual Studio Code доступний безкоштовно для платформ Windows, Linux і OS X. Microsoft презентувала Visual Studio Code у квітні 2015 на конференції Build 2015 як перший кросплатформний продукт у лінійці Visual Studio.

Python - інтерпретована, об'єктно-орієнтована мова програмування з високорівневою синтаксичною структурою. Це мова загального призначення, яка найчастіше використовується для розробки веб-застосунків, наукових досліджень, штучного інтелекту та інших задач. Python має простий синтаксис, що дозволяє розробникам швидко писати чистий та зрозумілий код. Він підтримує модульність та розширюваність, що дозволяє легко інтегрувати з іншими мовами та бібліотеками.

**1. АНАЛІЗ ЗАДАЧІ, ЗАСОБІВ ТА МЕТОДІВ ЇЇ ВИРІШЕННЯ**

**1.1 Розумiння вимог**

**Вхідні дані:**

Програма для аналізу даних з соціальних медіа призначена для роботи зі структурованими даними у форматі CSV. Вона приймає дані через командний рядок де користувачем вказується шлях до відповідного файлу. Передбачається, що формат CSV-файлу має стандартизовану структуру, де зберігається інформація про публікації в соціальних мережах, таку як текст публікації, кількість лайків, коментарів та репости.

**Обробка даних:**

Після отримання вхідних даних у форматі CSV програма проводить їх обробку для подальшого аналізу. Цей процес включає зчитування даних з файлу та їх трансформацію у об’єкти класу. Після завершення обробки даних формується масив, що містить інформацію про кожну публікацію, включаючи текст, кількість лайків, коментарів та репости.

**Аналіз популярності:**

Одним із основних функціональних елементів програми є аналіз популярності контенту, що дозволяє визначити відгуки аудиторії на публікації. Програма використовує метрики, такі як загальна кількість взаємодій (лайки, коментарі, репости разом) або сортування постiв для оцінки популярності контенту. Такий аналіз дозволяє користувачеві краще розуміти, який контент найбільш цікавий для аудиторії та як взаємодіється з нею.

**1.2 Аналіз вихідних даних**

CSV-файл містить дані про різні публікації, що є повідомленнями або оголошеннями на платформі соціальних медіа чи іншому веб-сайті. Ось основнi стовпчики що буде мати CSV-файл:

1. **Назва публікації (Post title)**: Це перше поле в кожному рядку і містить заголовок або опис публікації.
2. **Лайки (Likes)**: Це кількість користувачів, які виразили своє задоволення публікацією, натиснувши кнопку "лайк".
3. **Коментарі (Comments)**: Це кількість коментарів, які були залишені під публікацією.
4. **Репости (Reposts)**: Це кількість разів, коли публікація була поділена іншими користувачами на їх власних сторінках чи профілях.
5. **Відвідування профілю (Profile visits)**: Це кількість разів, коли користувачі переходили на профіль автора публікації, можливо, після перегляду публікації.

Ці дані можуть бути використані для аналізу популярності публікацій, їхньої взаємодії з аудиторією та впливу на сторінку або профіль у соціальних мережах. Наприклад, високі цифри лайків та репостів можуть свідчити про успішні публікації, які викликали зацікавленість аудиторії. Коментарі також можуть дати уявлення про взаємодію з аудиторією та відгуки на контент.

**Загальнi вiдомостi про CSV:**

CSV (Comma-Separated Values) - це текстовий формат для зберігання даних у вигляді таблиці, де кожен рядок представляє один запис, а кожний стовпець в цьому записі містить значення різних полів, розділені комами або іншими розділовими знаками. CSV файл може мати розширення .csv.

Принцип роботи CSV файлу досить простий:

1. **Розділення даних**: Дані у CSV файлі зазвичай розділяються комами, хоча іноді використовуються інші розділові знаки, такі як крапки з комою або табуляція.
2. **Рядки і стовпці**: Кожен рядок у CSV файлі відповідає окремому запису або рядку даних, а кожен стовпець в рядку містить певну інформацію, яка відноситься до цього запису. Наприклад, якщо CSV файл містить дані про співробітників, кожен рядок може містити прізвище, ім'я, посаду тощо, а кожен стовпець у рядку буде відповідати конкретному полю (прізвище, ім'я, посада тощо).
3. **Кодування даних**: Дані у CSV файлах можуть бути закодовані в різних форматах, таких як UTF-8 або ASCII, залежно від потреби і використовуваних програм.
4. **Заголовок**: У деяких випадках перший рядок CSV файлу може містити заголовки стовпців, які надають інформацію про те, які дані містяться в кожному стовпці. Це полегшує читання та обробку даних.
5. **Зберігання даних**: CSV файли легко зберігаються у текстовому форматі і можуть бути відкриті та редаговані з використанням різноманітних програм, включаючи електронні таблиці, текстові редактори та спеціалізовані програми для обробки даних.

CSV файли широко використовуються для обміну даними між програмами, так як вони є простими у використанні та зрозумілими як для комп'ютерів, так і для людей.

**1.3 Вибір технологій**

Python - це високорівнева, інтерпретована мова програмування загального призначення з простим синтаксисом та динамічною типізацією.

#### **Загальна інформація:**

Python був розроблений Гвідо ван Россумом та вперше випущений у 1991 році. Він швидко став популярним завдяки своїй простоті та читабельності коду. Python підтримує багато парадигм програмування, таких як процедурне, об'єктно-орієнтоване, функціональне та аспектно-орієнтоване програмування.

#### **Основні особливості Python:**

1. **Простота читання та написання коду**: Python має чітку та лаконічну синтаксичну структуру, що дозволяє легко читати та розуміти код.
2. **Динамічна типізація**: Змінні в Python не пов'язані з конкретними типами даних під час оголошення, а типи визначаються автоматично під час виконання програми.
3. **Широкий спектр бібліотек та фреймворків**: Python має велику кількість стандартних бібліотек та сторонніх модулів, які полегшують розробку різноманітних проектів.
4. **Портативність**: Python працює на різних операційних системах, таких як Windows, macOS, Linux та інших.
5. **Швидкість розробки**: Благодаря простоті та потужним інструментам, розробка програм на Python зазвичай займає менше часу порівняно з іншими мовами програмування.
6. **Широке застосування**: Python використовується у багатьох галузях, включаючи веб-розробку, наукові обчислення, штучний інтелект, обробку даних, машинне навчання та багато інших.

**Стороннi бiблiотеки:**

У цьому проекті використовується лише модуль operator та функція attrgetter. Модуль operator є частиною стандартної бібліотеки Python і містить функції, які дозволяють виконувати операції з типовими операторами. Функція attrgetter з цього модулю використовується для отримання значень атрибутів об'єктів, що дозволяє здійснювати сортування за цими значеннями.

Це рішення обрано для того, щоб зберегти проект простим та не додавати надмірної складності за рахунок використання сторонніх бібліотек.

**Редактор коду:**

Як редактор коду, в цьому проектi був обраний Microsoft Visual Studio Code. **Visual Studio Code (VS Code)** - це безкоштовний редактор коду, який надає широкі можливості для розробки програмного забезпечення. Дане середовище розробки засноване на платформі Electron, що дозволяє використовувати вбудований веб-браузер і забезпечує можливості розширення за допомогою плагінів. VS Code підтримує різні мови програмування, включаючи Python, та надає багато функцій, які сприяють продуктивності розробника.

Де використовують Visual Studio Code:

* Розробка веб-застосунків (Frontend та Backend).
* Розробка мобільних додатків.
* Розробка хмарних додатків.
* Розробка штучного інтелекту, машинного навчання та аналізу даних.
* Робота з системами керування версіями (наприклад, Git).
* Робота з Docker і Kubernetes.

Плюси Visual Studio Code:

1. **Безкоштовність**: VS Code є безкоштовним для використання та має відкритий вихідний код.
2. **Підтримка різних мов програмування**: Включаючи Python, JavaScript, TypeScript, C++, Java, і багато інших.
3. **Розширюваність**: За допомогою розширень можна розширити функціональність редактора.
4. **Широкі можливості налаштування**: Велика кількість налаштувань та можливість налаштування редактора під власні потреби.
5. **Зручний інтерфейс користувача**: Інтуїтивно зрозумілий та легкий у використанні інтерфейс.

**2. ПРОЕКТУВАННЯ ЗАГАЛЬНОГО АЛГОРИТМУ РОБОТИ ПРОГРАМИ**

Для зручності рекомендується розбити роботу на два етапи.

**Перший етап** – створення коду, який буде переводити кожен рядок CSV-файлу в об'єкт класу Experiment для подальшої роботи.

У цьому етапі розроблюється клас CSV\_Reader, який матиме два методи:

* Метод імпортування, що приймає задане користувачем ім'я файлу та відкриває його стандартними засобами Python.
* Метод, який розділяє дані, що знаходяться в файлі, та одразу заносить їх до ініціалізатора об'єкта класу Post.

**Другий етап** – сортування масиву перетворених об'єктів та виведення їх в консоль.

На другому етапі необхідно здійснити сортування об'єктів за їх атрибутами, наприклад, за датою або іншими параметрами. Після сортування об'єкти можуть бути виведені в консоль або оброблені іншим чином, залежно від потреб програми.

Ці два етапи дозволяють розділити роботу на більш зручні компоненти, що полегшує процес розробки та збереження коду в майбутньому.

**План першого етапу:**

Нижче наведенi блок-cхеми роботи алгоритмiв iмпортування CSV файлу та переведення його в об’єкти класу Post, що дасть змогу полегшити роботу з сортуванням кожного запису або посту соцiальних мереж та обчислювати загальний критерiй зацiкавленостi аудиторiї.

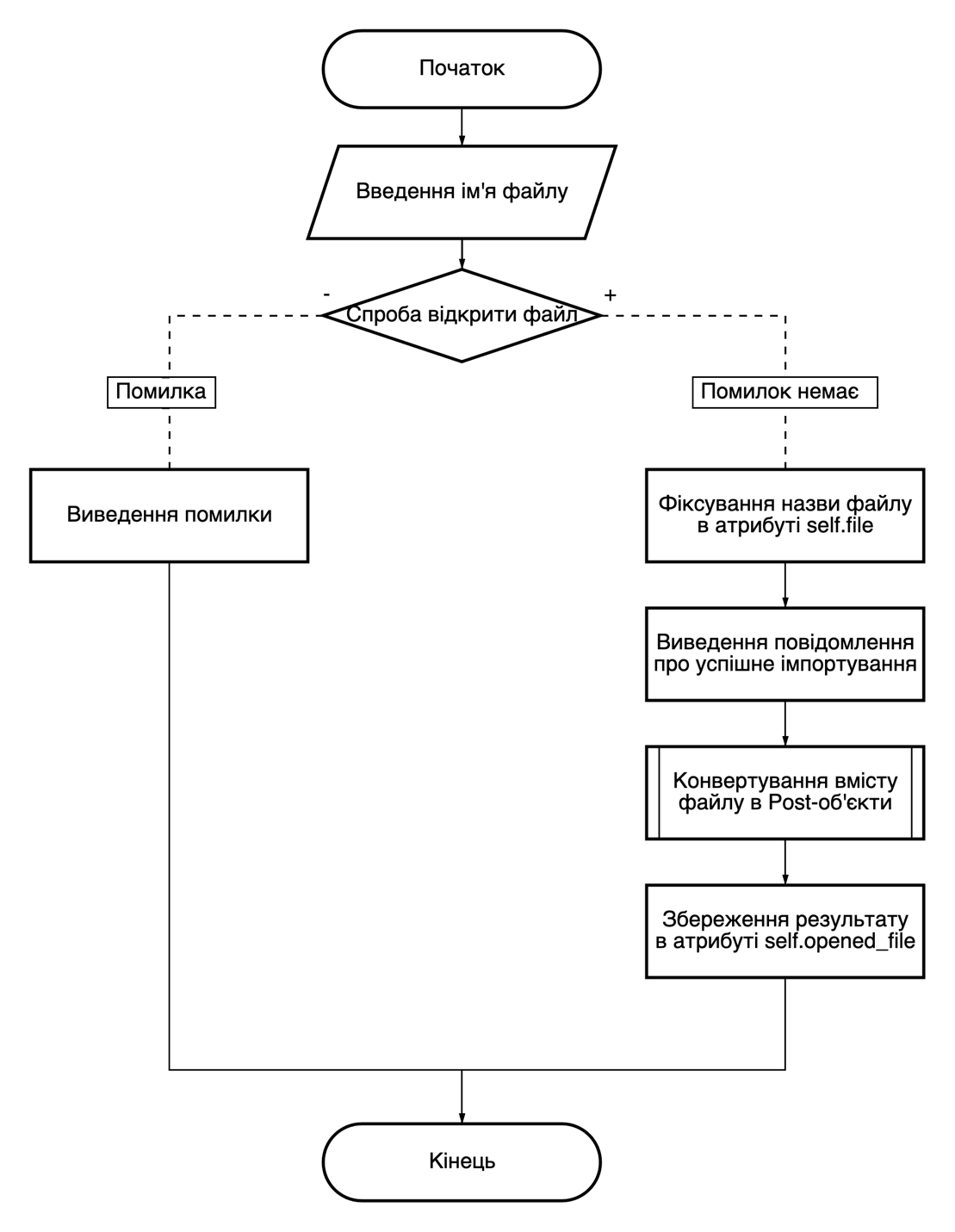
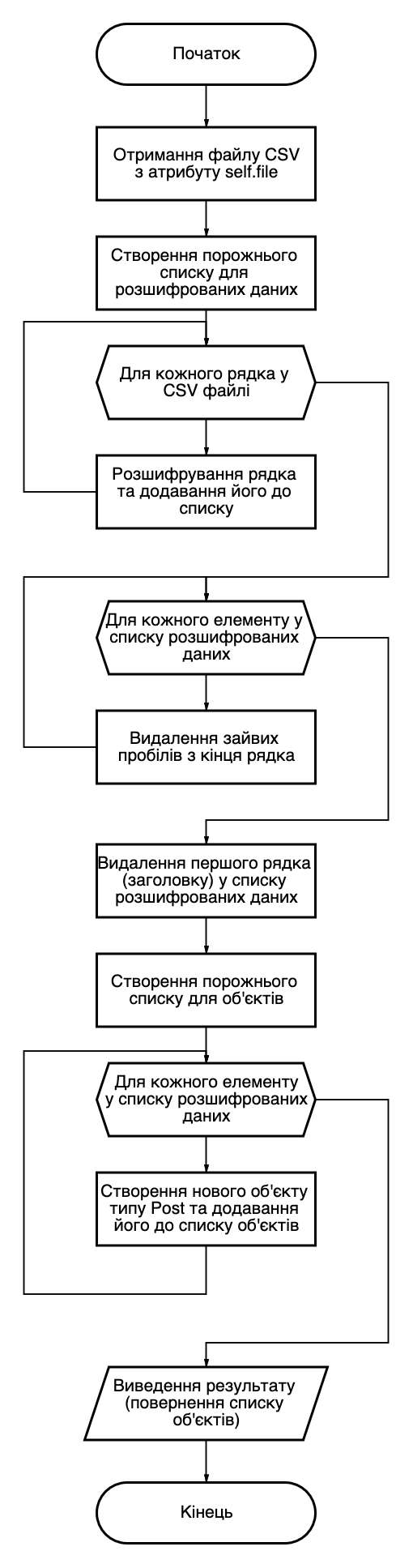


Рис. 2.1 – Схема функції iмпорту файлу.

Рис. 2.2 – Схема методу переводу даних CSV файлу в об’єкти Post.

**План другого етапу:**

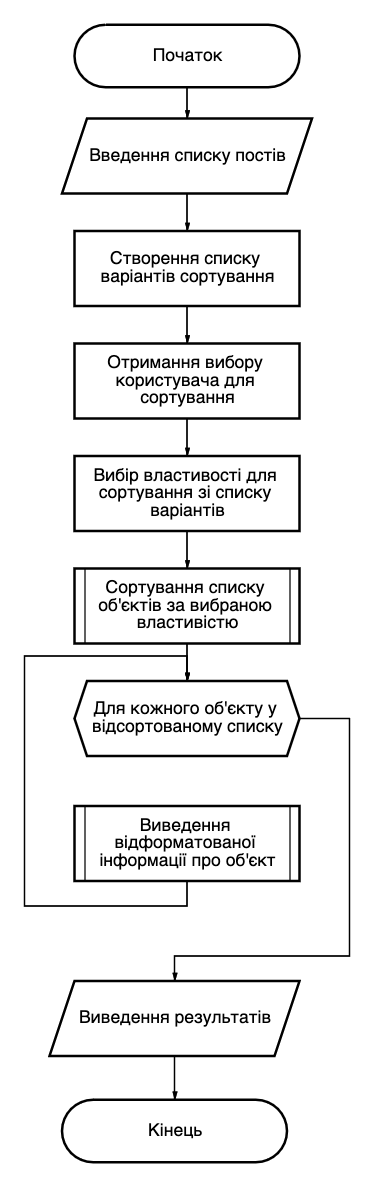
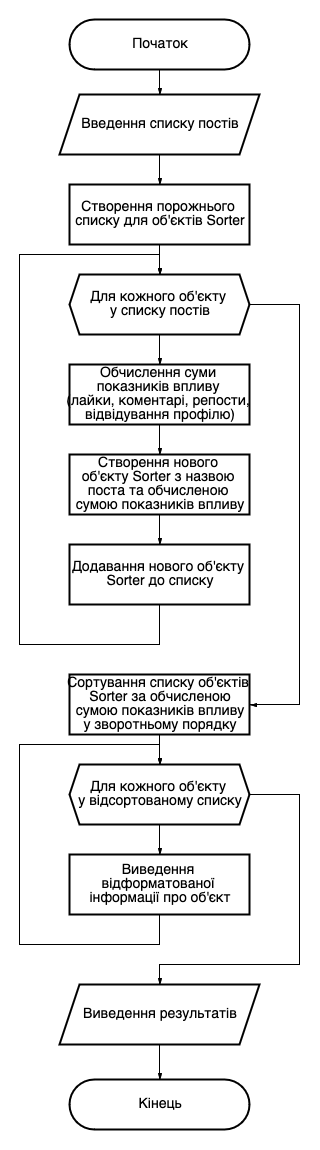
****Нижче наведенi блок-cхеми роботи алгоритмiв сортування об’єктiв класу Post та виведення загального значення популярностi посту, це допоможе мати бiльш краще уявлення про взаемодiю користувачiв з постом.

Рис. 2.3 – Схема методу сортування об’єктiв Post.

****Рис. 2.4 – Схема методу узагальнення взаемодiй з постом – show\_popular.

**3. РОЗРОБКА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ**

**3.1 Розробка головного класу Post**

Для створення програмного забезпечення з метою аналізу даних з соціальних медіа з CSV-файлу, який містить інформацію про публікації, лайки, коментарі та репости, потрібно виконати такі кроки:

1. **Визначення вимог до програми**:
   * Розробка функціоналу для аналізу популярності та взаємодії з контентом.
   * Опрацювання даних з CSV-файлу: визначення структури файлу, вибір необхідних полів для аналізу.
2. **Розробка архітектури програми**:
   * Визначення компонентів програми: модулі для обробки даних, аналізу популярності, взаємодії з контентом тощо.
   * Встановлення взаємозв'язків між компонентами, їх взаємодія для забезпечення потрібного функціоналу.
   * Вибір оптимальних методів обробки та аналізу даних для забезпечення швидкості та ефективності програми.
3. **Реалізація компонентів програми**:
   * Розробка коду компонентів відповідно до принципiв ООП.
   * Забезпечення коректної роботи кожного компонента та їх взаємодії з даними з CSV-файлу.
4. **Тестування програми**:
   * Проведення тестів для перевірки коректності роботи програми, відповідно до заданих вимог.
   * Виявлення та усунення будь-яких помилок або недоліків.
5. **Виправлення та перевірка правильності роботи**:
   * Внесення виправлень у програму, якщо необхідно після тестування.
   * Повторна перевірка програми для забезпечення її коректної роботи після внесених змін.

З вже визначеними вимогами та розробленою архiтектурою можна приступити до розробки програмного коду:

Для легкого манiпулювання даними що знаходяться в CSV файлi їх рекомендуеться перевести в об’єкти, для цього створюємо клас Post, який має всi необхiднi атрибути, такi як:

self.post\_title

self.likes

self.comments

self.reposts

self.profile\_visits

Рис. 3.1.1 – Класс об’єкту Post.

**Методи**

**\_\_str\_\_:** Цей метод визначає рядкове представлення об'єкта типу "пост". Він викликається при використанні функції str() або при виведенні об'єкта на екран. У цьому випадку він повертає рядок, який містить інформацію про заголовок поста, кількість лайків, коментарів, репостів і відвідувань профілю.

**print\_out**: Цей метод створений для форматованого виведення інформації про об'єкт типу "пост" у терміналі з використанням кольорів, які імпортовані з класу Bcolors. Він повертає рядок, який містить інформацію про заголовок поста, кількість лайків, коментарів, репостів і відвідувань профілю, причому кожен параметр окремо забарвлюється відповідним кольором.

В даному прикладi у кожного атрибуту є параметри за замовчуванням, якi використовують клас PostTitle та Generator щоб створити начальний CSV файл для тестування програми, давайте розглянемо в пунктi 3.2:

**3.2 Класи для тестування**

Клас Generator має в собi атрибут rows(поля) та метод **generate()** який за кiлькiстю полiв створює випадковi об’єкти класу Post та записує їх в файл.



Рис. 3.2.1 – Клас Generator.

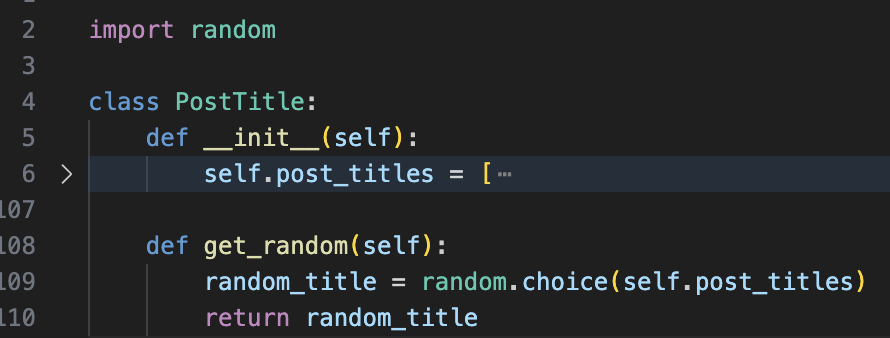
Клас PostTitle має список вже заданих назв постiв та метод який повертає один з цих назв – **get\_random()**.

Рис. 3.2.2 – Клас PostTitle.

Всi цi класи були створенi для спрощення тестування програми i не мають прямого впливу на її роботу, хоча й дуже допоможуть в майбутньому, коли програма буде повнiстю готова.

**3.3 Клас CSV\_Reader**

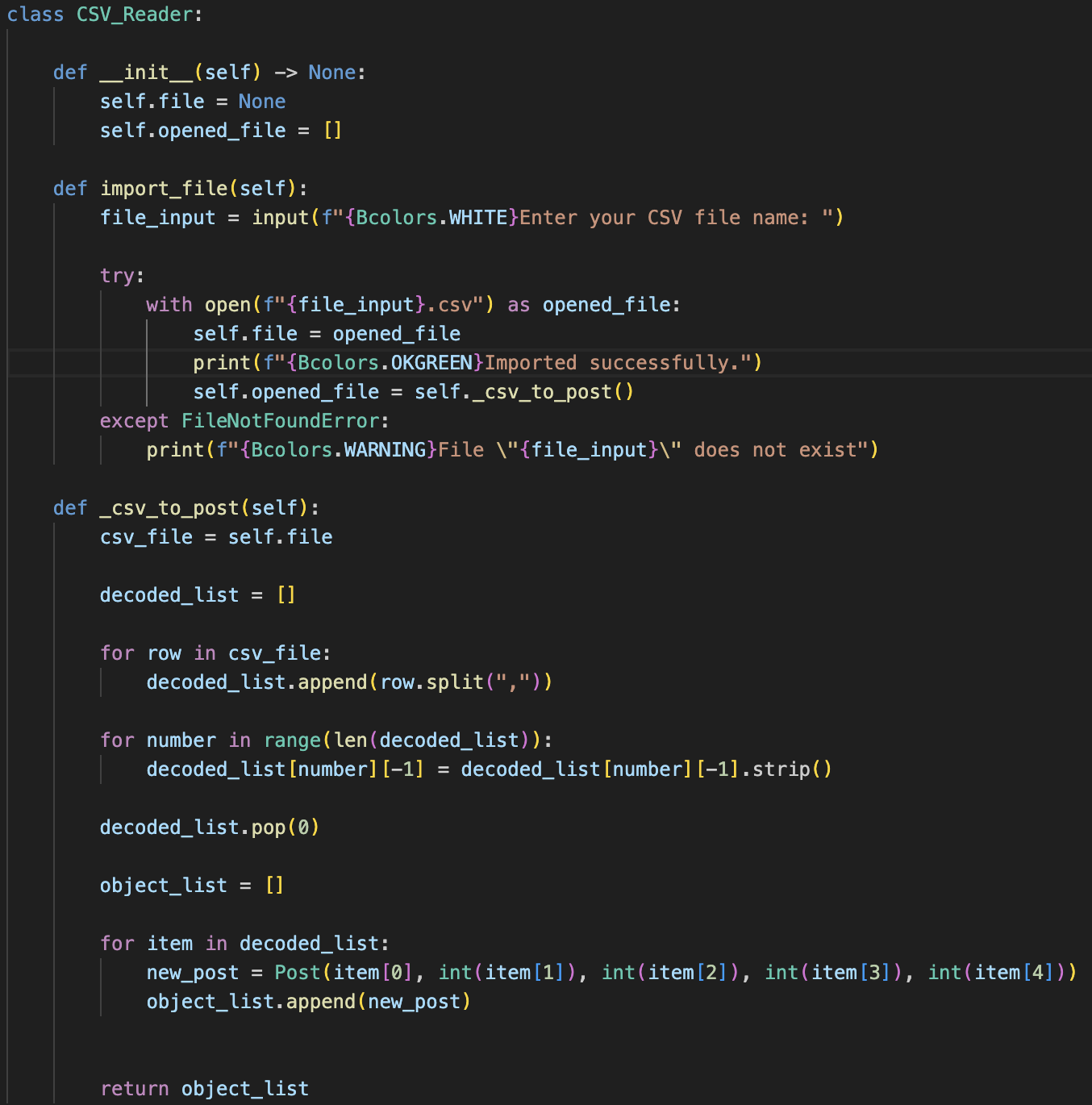
Перейдемо до створення класу який буде вiдповiдати за iмпортування самого CSV-файлу та переведенню його в об’єкти Post та буде зручно зберiгати їх в атрибутi opened\_file який є масивом.

Рис. 3.3.1 – Клас CSV\_Reader.

**Принцип дiї:**

1. **Ініціалізація:** У конструкторі \_\_init\_\_, він ініціалізує дві змінні: file (яка представлятиме відкритий файл) та opened\_file (яка буде зберігати об'єкти, що відображають дані з CSV файлу).
2. **Імпорт файлу:** Метод import\_file отримує назву файлу від користувача, як введення. Він відкриває цей файл (якщо він існує), читає його вміст і перетворює його на список об'єктів Post. Якщо файл не знайдено, виводиться відповідне повідомлення.
3. **Перетворення CSV у пости:** Метод \_csv\_to\_post відповідає за перетворення вмісту CSV файлу на об'єкти Post. Він читає CSV файл, розділяє кожен рядок за комами, видаляє зайві пробіли, видаляє перший рядок (який зазвичай містить заголовки стовпців), і створює об'єкти Post з кожного рядка, використовуючи дані з цього рядка.

**3.4 Клас Bcolors**

Ще один клас який можна вiднести до допомiжних – Bcolors, ось що про нього потрiбно знати:

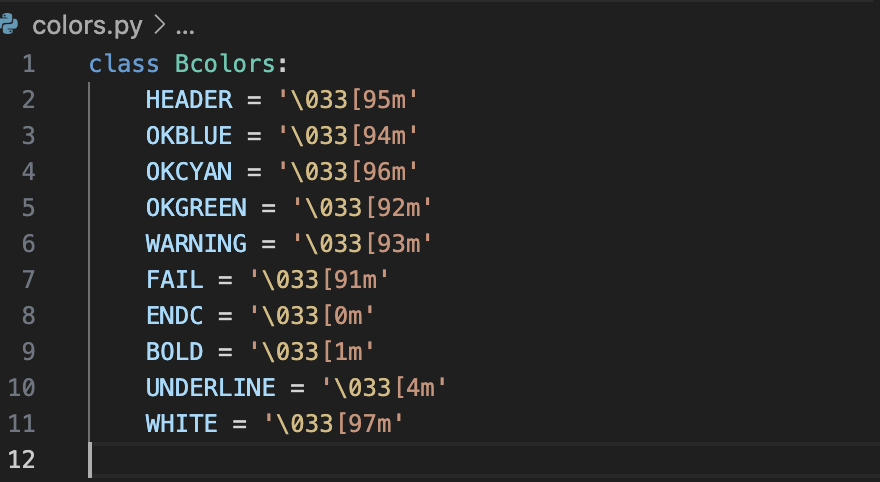
1. **Статичні атрибути:** Клас Bcolors має ряд статичних атрибутів, які представляють різні кольори. Кожен атрибут є рядком, що містить спеціальний код кольору ANSI для відображення тексту в відповідному кольорі у терміналі.
2. **Використання кольорів:** Ці атрибути можна використовувати, додавши їх до виведеного тексту за допомогою спеціальних кодів у форматі Escape-послідовностей. Наприклад, Bcolors.OKGREEN додасть код кольору, щоб текст виводився зеленим кольором. Це може бути корисно для покращення зручності виведення інформації в консольному інтерфейсі вашої програми.
3. **Нестатичні атрибути та конструктор:** Оскільки всі атрибути цього класу є статичними, йому не потрібен конструктор. Статичні атрибути доступні через ім'я класу, а не через екземпляр класу. Тому немає потреби в ініціалізації цих атрибутів для кожного об'єкта класу. Вони доступні всюди в програмі, де імпортований цей клас.

Рис. 3.4.1 – Клас Bcolors.

**3.5 Сортування та клас Sorter**

Одним із найважливіших завдань у рамках цього проекту є організація та обчислення загальних показників для кожного поста. Для цього був створений окремий клас під назвою Sorter, який допомагає в сортуванні та аналізі інформації.

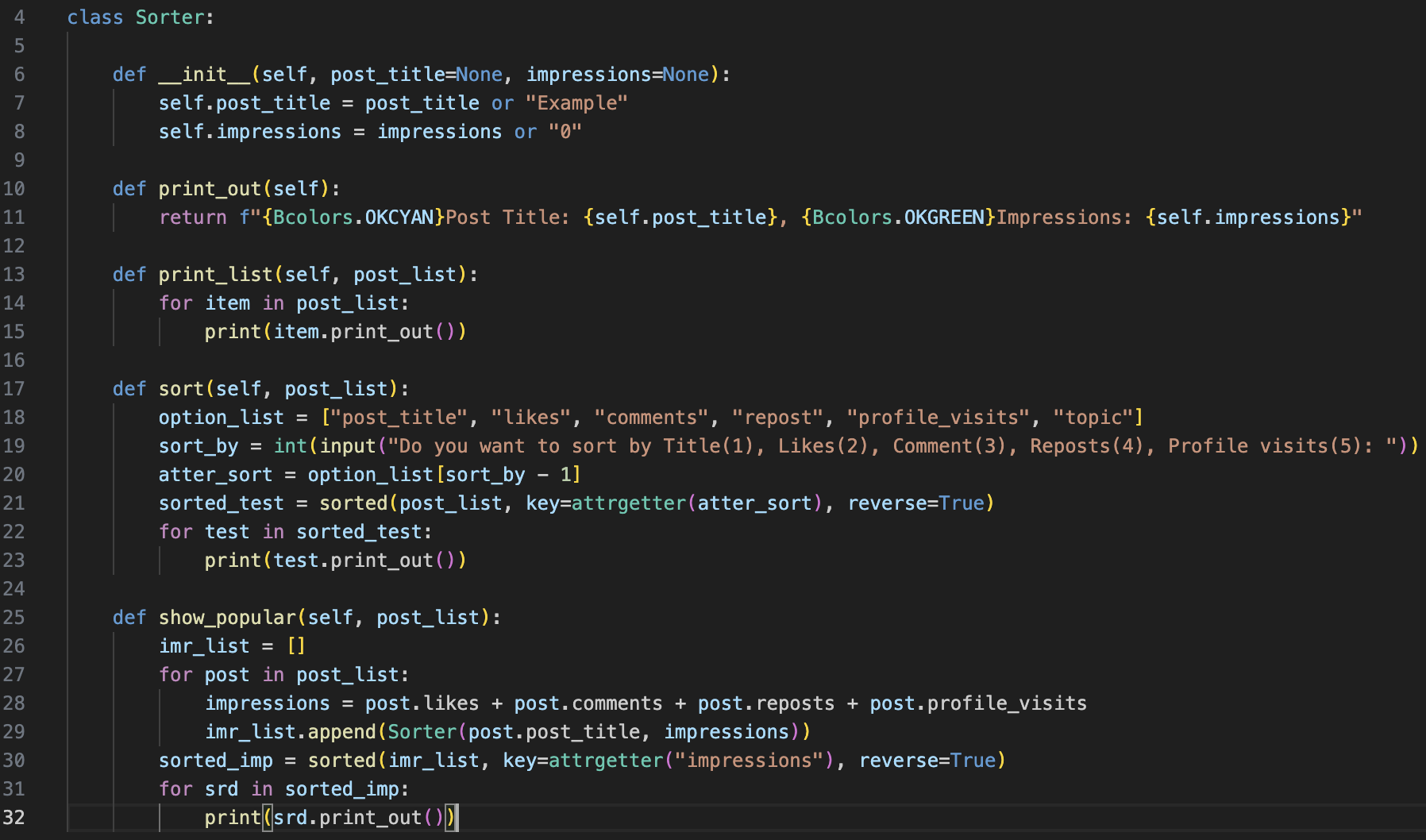
Клас Sorter має ряд методів, серед яких конструктор (**init**), який ініціалізує атрибути об'єктів, такі як назва поста (post\_title) та кількість переглядів (impressions). Метод print\_out виводить інформацію про пост, включаючи назву та кількість переглядів. Метод print\_list приймає список постів і виводить інформацію про кожен з них. Метод sort сортує список постів за вибраним параметром (наприклад, назвою, кількістю лайків тощо) та виводить відсортований список. Метод show\_popular аналізує популярність кожного поста, розраховуючи загальну кількість взаємодій (лайки, коментарі, репости, відвідування профілю) і виводить пости в порядку спадання популярності.

Рис. 3.5.1 – Клас Sorter.

Цей клас був створений таким, як він є, для того, щоб забезпечити зручний та ефективний інструмент для обробки даних про пости у соціальних мережах. Основна мета цього класу полягає в тому, щоб упорядковувати та аналізувати інформацію про пости з різних кутів зору, таких як загальна популярність, кількість лайків, коментарів, репостів тощо.

Створення окремого класу дозволяє легко маніпулювати та аналізувати дані, використовуючи спеціалізовані методи та функції. Наприклад, метод sort дозволяє швидко сортувати список постів за різними критеріями, тоді як метод show\_popular робить можливим виділення найпопулярніших постів.

Крім того, використання класу дозволяє зберігати інформацію про кожен пост у відокремлених об'єктах, що спрощує доступ до цієї інформації та робить код більш структурованим та зрозумілим для подальшого розширення та модифікації. Такий підхід сприяє підтримці чистого та організованого коду, що є важливим для подальшого розвитку проекту та пiдтримання принципiв ООП.

**3.6 Основний файл виконання програми**

Файл main.py - це головний файл програми, де відбувається взаємодія з іншими частинами програми. У парадигмі об'єктно-орієнтованого програмування (ООП) main.py виконує роль координатора роботи об'єктів та визначає послідовність їх викликів та дій.

У моему коді, main.py відповідає за створення об'єктів Post, CSV\_Reader і Sorter, тобто за їх створення та управління.

Структура файлу main.py досить проста і пряма:

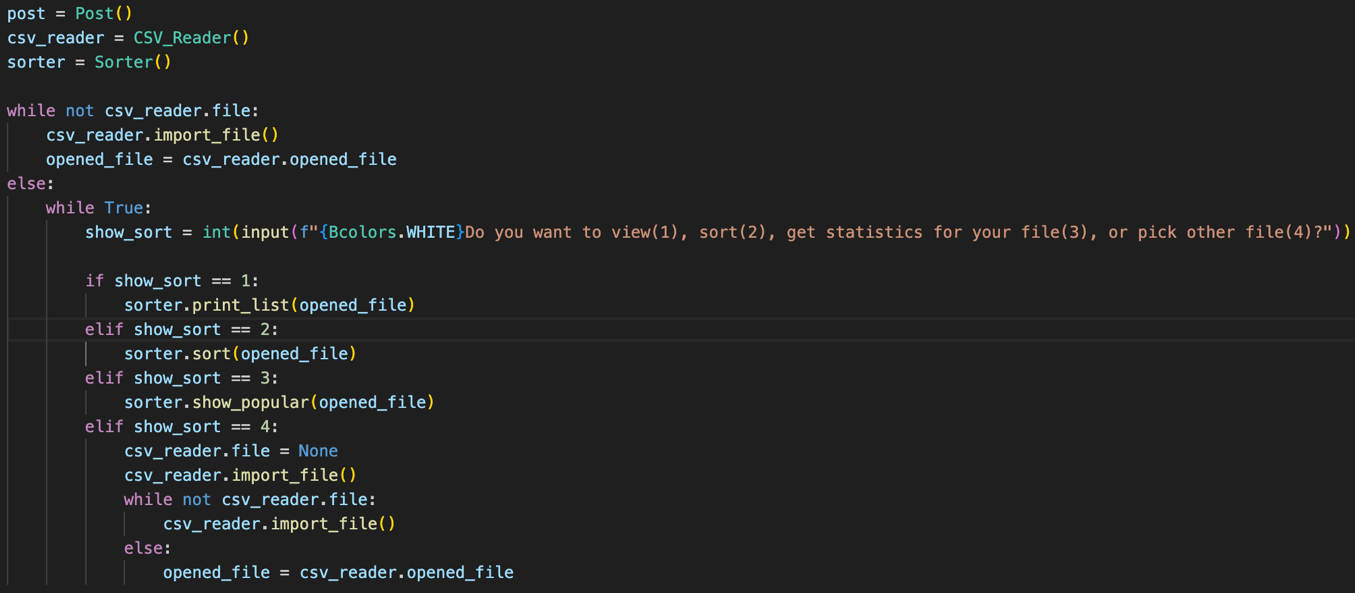
1. Спочатку створюються об'єкти для постів (Post), CSV читача (CSV\_Reader) та сортувальника (Sorter).
2. Потім main.py імпортує файл (з використанням методу import\_file() з CSV читача) та отримує відкритий файл.
3. Далі створюється цикл, який перевіряє, чи існує файл. Якщо файл не існує, програма намагається знову імпортувати його.
4. Після цього відбувається безкінечний цикл, де користувач може обирати, що саме він хоче зробити з файлом: переглянути, відсортувати, отримати статистику або вибрати інший файл. Залежно від вибору користувача, викликаються відповідні методи об'єкта Sorter, який здійснює відповідні операції над відкритим файлом.

Рис. 3.6.1 – Файл main.py.

На цьому етапi розробку коду ПО можна вважати закiнченною, всi потрiбнi класи, методи та модулi були створенi та всi вимоги до функцiоналу програми задовiльненi.

1. **КЕРІВНИЦТВО КОРИСТУВАЧА**

**4.1 Основний функцiонал**

**Програма має наступний функцiонал:**

* Перегляд файлу в форматованому виглядi (view)
* Сортування за назвою (Title)
* Сортування за кiлькiстю лайкiв (Likes)
* Сортування за кiлькiстю коментарiв (Comment)
* Сортування за кiлькiстю репостiв (Reposts)
* Сортування за кiлькiстю вiзитiв профiлю з посту (Profile visits)
* Отримання загальної статистики для кожного посту (get statistics for your file)
* Вибiр iншого файлу (pick other file)

**4.2 Запуск програми**

Для запуску програми треба виконати наступнi дiї:

1. Перейти в внутрiшнiй каталог програми.
2. Вiдкрити файл main.py.
3. В Visual Studio Code, в правому верхньому кутку натиснути на кнопку “▶” (Run Pyhon File).

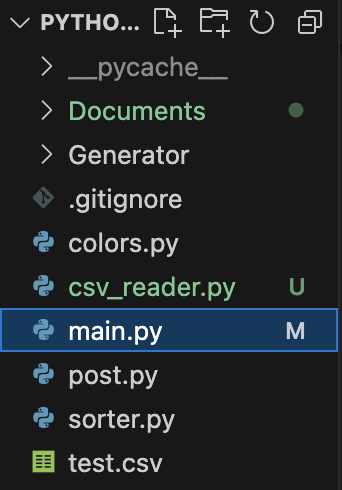


Рис. 4.2.1 – Файл main.py в кореневому каталозi.

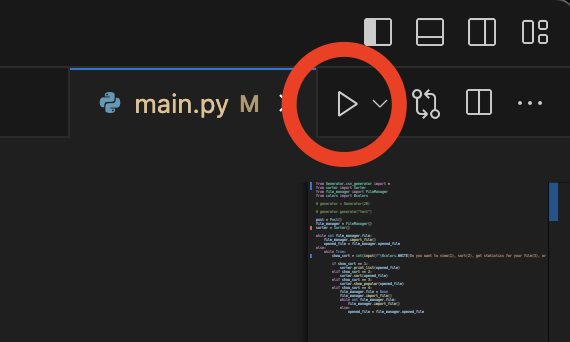
****

Рис. 4.2.2 – Кнопка запуску файлу (вiдмiчена червоним кольором).

**4.3 Iмпортування файлу**

Перед використанням програми треба переконатися в наявностi сумiсного CSV файлу в корневому каталозi, сумiсний файл повинен мати наступнi колонки в зазначеному порядку:

**Post title, Likes , Comments, Reposts, Profile visits**

1. Пiсля переконання в тому, що файл мiститься в каталозi програми, запустiть програму:

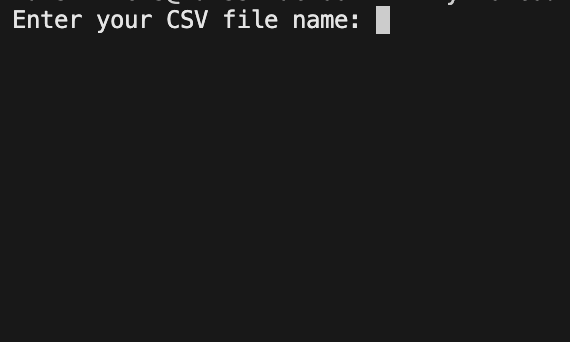


Рис. 4.3.1 – Початковий екран вибору файлу.

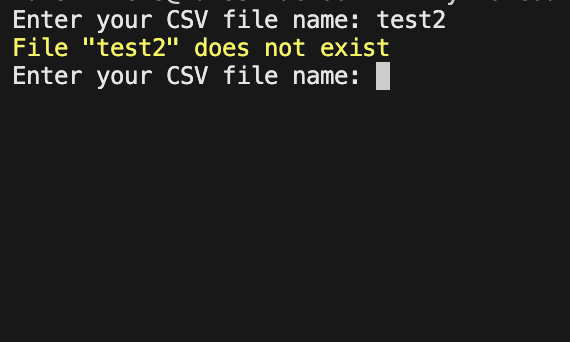
2. Пicля запуску, на початковому екранi (Рис. 4.3.1), введiть iм’я файлу у форматi лише назви (без формату файла).

Рис. 4.3.2 – Помилка при знаходженнi файлу.

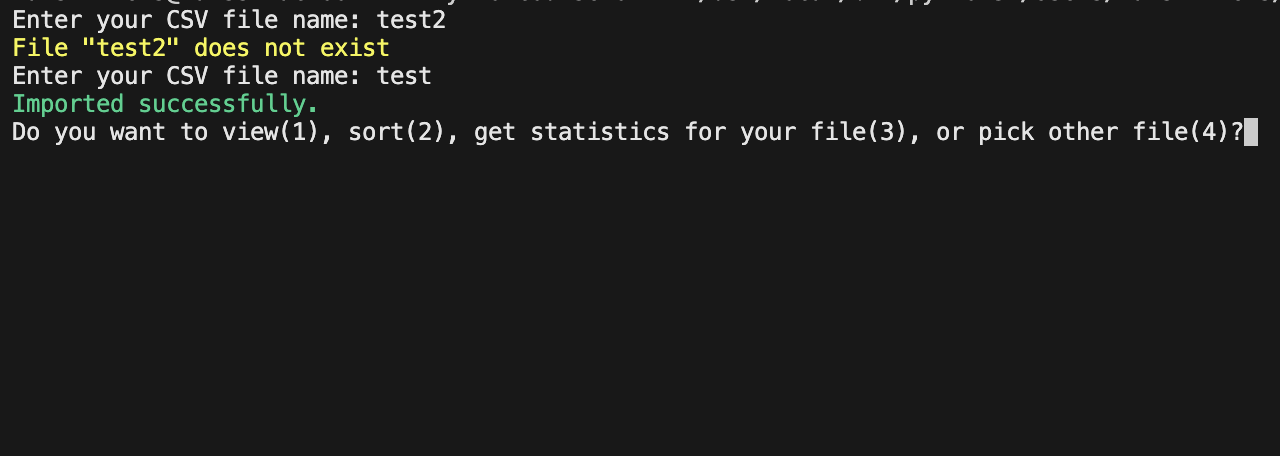
3. Помилка “File “**iм’я файлу**” does not exist” (Рис 4.3.2) наголошує користувачевi на тому, що файлу з введеною назвою не iснує, в такому випадку рекомендується перевiрити наявнiсть файлу в корневiй папцi програми або впевнитися що назва була вказана правилльно.

Рис. 4.3.3 – Приклад коректно введеного файлу.

1. Якщо файл було введено коректно програма засвiдчить про це написом “Imported successfully” (Рис 4.3.3).

**4.4 Робота з файлом**

**Функцiя view:**

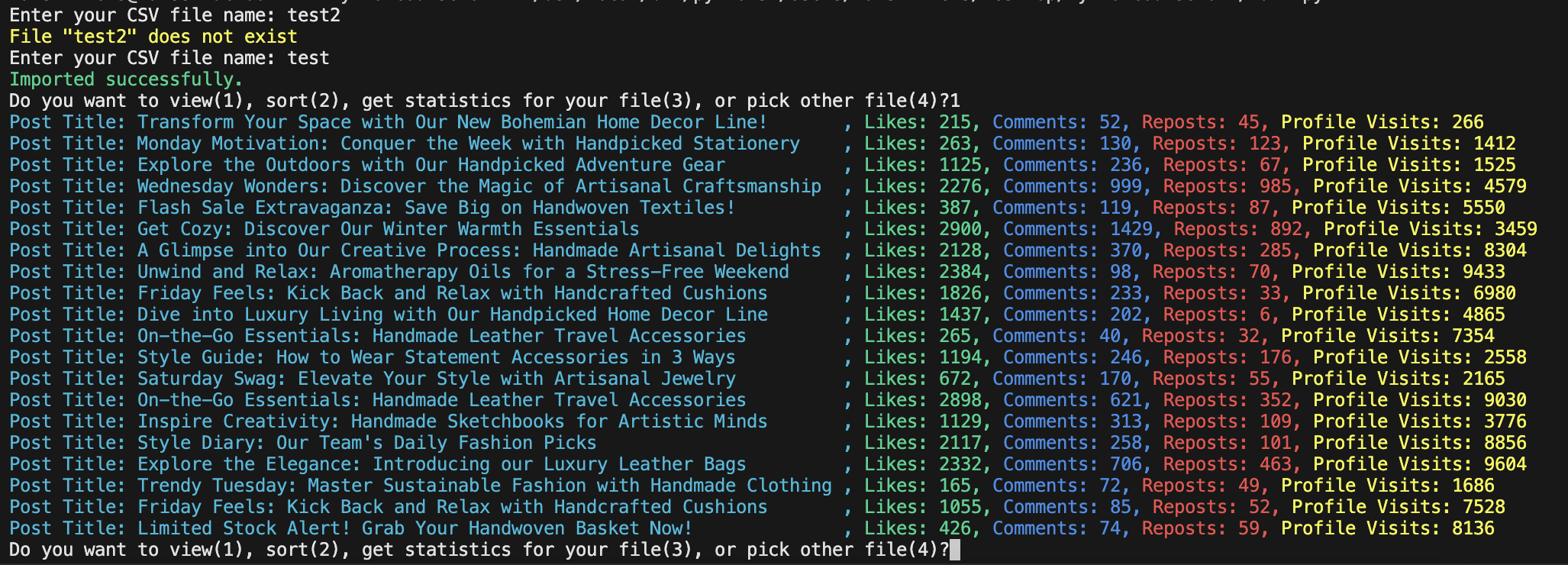


Рис. 4.4.1 – Перегляд файлу за допомогою опцiї “view”.

Фнкцiя “view” (Рис 4.3.4) виводить всi пости в файлi CSV в оригiнальному їх порядку, видiляючи данi рiзними кольорами для бiльш комфортного перегляду.

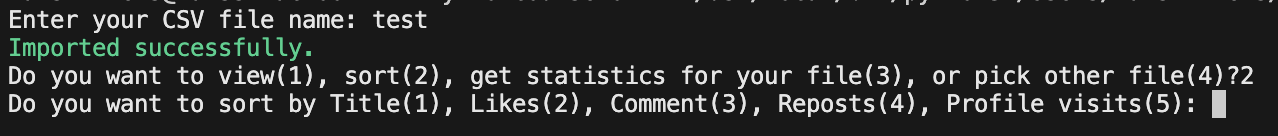
**Функцiя sort:**

Рис. 4.4.2 – Вибiр стовпчика сортування.

При виборi функцiї sort користувачу пропонується обрати стовпчик по якому буде проходити сортування (вiд бiльшого числа до меньшого).

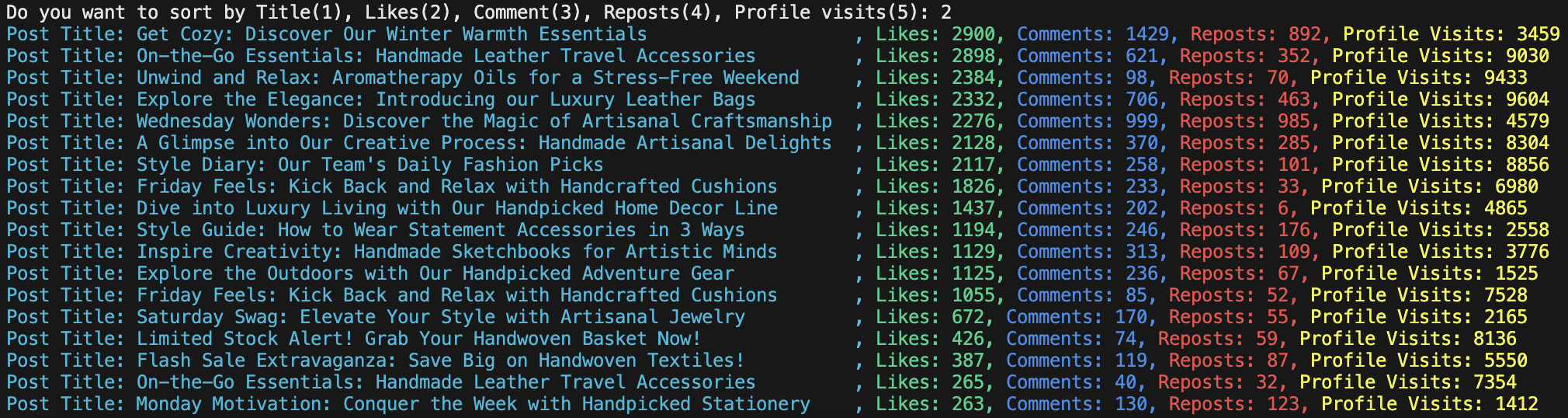


Рис. 4.4.3 – Приклад сортування за кiлькiстю лайкiв (2).

**Функцiя get statistics for your file:**

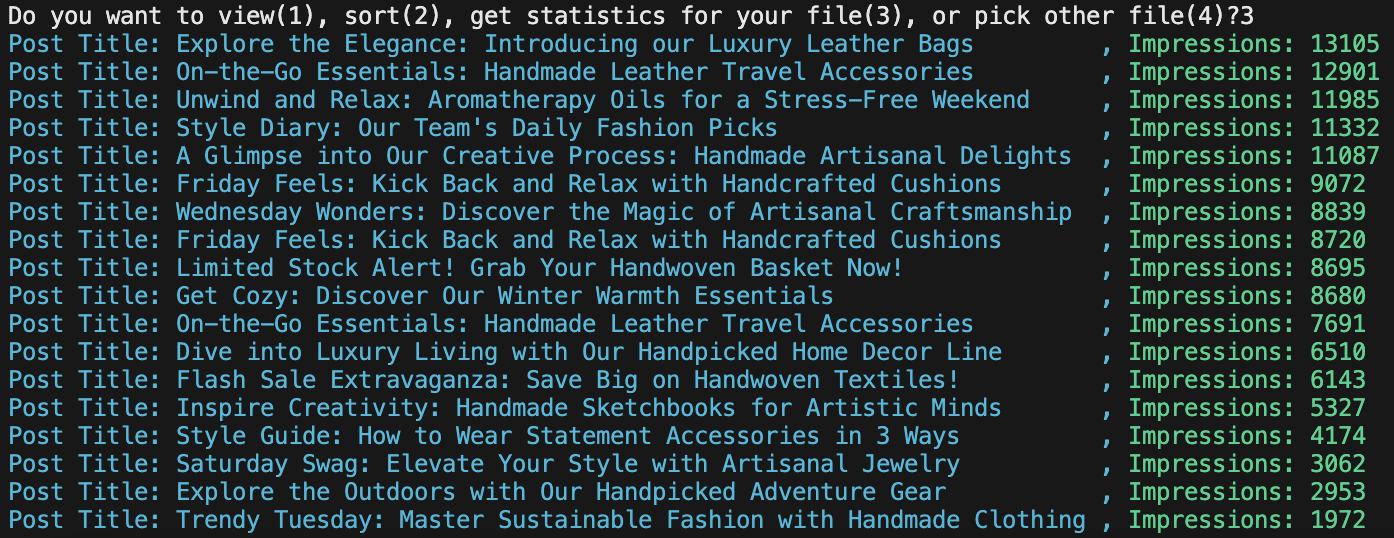
Функцiя get statistics for your file показує загальну кiлькiсть взаємодiй користувачiв з постом щоб отримати об’єктивний список самих популярних постiв.

Рис. 4.4.4 – Приклад роботи функцiї get statistics for your file.

**Функцiя pick another file:**

Функцiя pick another file повертає користувача в меню вибору файлу (див роздiл “4.3 Iмпортування файлу”)

ВИСНОВКИ

Під час виконання цієї курсової роботи було успішно створено застосунок для роботи з CSV-файлами та їх вмiстом. Проведено аналіз процесу розробки програмного забезпечення та визначено необхідність такого програмного забезпечення для ефективного сортування та аналiзу постiв соцiальних мереж. Основним методом для вирішення цієї задачі стало створення програмного забезпечення за допомогою мови Python, яке включало в себе роботу з класами та парадигмою ООП.

Був розроблений загальний алгоритм роботи програми, в якому були описані класи, які було доцільно створити, та їх призначення. Також була створена блок-схема основних функцій. Додатково були розроблені функціональні алгоритми роботи програми, де детально описано вимоги до додатку, коротко описаний його функціонал.

Під час розробки програмного забезпечення роботу було розбито на кроки для спрощення та зручності виконання завдання, і написаний відповідний код. У керівництві користувача детально описано необхідні дії для запуску додатку та його використання.

В результаті виконання курсової роботи було успішно досягнуто мети - розроблено програмне забезпечення для аналізу даних з соціальних медіа на мові Python, яке дозволяє ефективно аналiзувати данi та виводити статистику постiв.

**ЛIТЕРАТУРА**

1. CSV – Вiкiпедiя: веб-сайт. URL: https://uk.wikipedia.org/wiki/CSV
2. How to read CSV files with Python #1 - Without using python libraries: веб-сайт. URL:

<https://youtu.be/GBD4na8AQjY?si=x9-NbacAcLiKra48> (дата звернення 19.01.2024)

1. W3Schools Online Web Tutorials: веб-сайт. URL:

htpps://www.w3schools.com (дата звернення 14.01.2024)

1. Пришвидшений курс Python. Практичний проєктно-орiєнтований

вступ до програмування – Ерiк Маттерс, 2021 – 600 c

**ДОДАТКИ**

**Додаток А**

post = Post()

csv\_reader = CSV\_Reader()

sorter = Sorter()

while not csv\_reader.file:

csv\_reader.import\_file()

opened\_file = csv\_reader.opened\_file

else:

while True:

show\_sort = int(input(f"{Bcolors.WHITE}Do you want to view(1), sort(2), get statistics for your file(3), or pick other file(4)?"))

if show\_sort == 1:

sorter.print\_list(opened\_file)

elif show\_sort == 2:

sorter.sort(opened\_file)

elif show\_sort == 3:

sorter.show\_popular(opened\_file)

elif show\_sort == 4:

csv\_reader.file = None

csv\_reader.import\_file()

while not csv\_reader.file:

csv\_reader.import\_file()

else:

opened\_file = csv\_reader.opened\_file

**Додаток Б**

class CSV\_Reader:

def \_\_init\_\_(self) -> None:

self.file = None

self.opened\_file = []

def import\_file(self):

file\_input = input(f"{Bcolors.WHITE}Enter your CSV file name: ")

try:

with open(f"{file\_input}.csv") as opened\_file:

self.file = opened\_file

print(f"{Bcolors.OKGREEN}Imported successfully.")

self.opened\_file = self.\_csv\_to\_post()

except FileNotFoundError:

print(f"{Bcolors.WARNING}File \"{file\_input}\" does not exist")

def \_csv\_to\_post(self):

csv\_file = self.file

decoded\_list = []

for row in csv\_file:

decoded\_list.append(row.split(","))

for number in range(len(decoded\_list)):

decoded\_list[number][-1] = decoded\_list[number][-1].strip()

decoded\_list.pop(0)

object\_list = []

for item in decoded\_list:

new\_post = Post(item[0], int(item[1]), int(item[2]), int(item[3]), int(item[4]))

object\_list.append(new\_post)

return object\_list

fake = Faker()

class Post:

def \_\_init\_\_(self, post\_title=None, likes=None, comments=None, reposts=None, profile\_visits=None):

self.post\_title = post\_title or PostTitle.get\_random()

self.likes = likes or random.randint(50, 3000)

self.comments = comments or random.randint(1, self.likes // 2)

self.reposts = reposts or random.randint(0, self.comments)

self.profile\_visits = profile\_visits or random.randint(self.likes, 10000)

def \_\_str\_\_(self):

return f"{self.post\_title}, {self.likes}, {self.comments}, {self.reposts}, {self.profile\_visits}"

def print\_out(self):

return f"{Bcolors.OKCYAN}Post Title: {self.post\_title}, {Bcolors.OKGREEN}Likes: {self.likes}, {Bcolors.OKBLUE}Comments: {self.comments}, {Bcolors.FAIL}Reposts: {self.reposts}, {Bcolors.WARNING}Profile Visits: {self.profile\_visits}"

class Sorter:

def \_\_init\_\_(self, post\_title=None, impressions=None):

self.post\_title = post\_title or "Example"

self.impressions = impressions or "0"

def print\_out(self):

return f"{Bcolors.OKCYAN}Post Title: {self.post\_title}, {Bcolors.OKGREEN}Impressions: {self.impressions}"

def print\_list(self, post\_list):

for item in post\_list:

print(item.print\_out())

def sort(self, post\_list):

option\_list = ["post\_title", "likes", "comments", "repost", "profile\_visits", "topic"]

sort\_by = int(input("Do you want to sort by Title(1), Likes(2), Comment(3), Reposts(4), Profile visits(5): "))

atter\_sort = option\_list[sort\_by - 1]

sorted\_test = sorted(post\_list, key=attrgetter(atter\_sort), reverse=True)

for test in sorted\_test:

print(test.print\_out())

def show\_popular(self, post\_list):

imr\_list = []

for post in post\_list:

impressions = post.likes + post.comments + post.reposts + post.profile\_visits

imr\_list.append(Sorter(post.post\_title, impressions))

sorted\_imp = sorted(imr\_list, key=attrgetter("impressions"), reverse=True)

for srd in sorted\_imp:

print(srd.print\_out())

**Додаток В**

class Bcolors:

HEADER = '\033[95m'

OKBLUE = '\033[94m'

OKCYAN = '\033[96m'

OKGREEN = '\033[92m'

WARNING = '\033[93m'

FAIL = '\033[91m'

ENDC = '\033[0m'

BOLD = '\033[1m'

UNDERLINE = '\033[4m'

WHITE = '\033[97m'