МИНИСТЕРСТВО ЦИФРОВОГО РАЗВИТИЯ, СВЯЗИ И

МАССОВЫХ КОММУНИКАЦИЙ РОССИЙСКОЙ ФЕДЕРАЦИИ

Ордена Трудового Красного Знамени

федеральное государственное бюджетное образовательное

учреждение высшего образования

«Московский технический университет связи и информатики»

Кафедра «Математическая кибернетика и информационные технологии»

Дисциплина «Структуры и алгоритмы обработки данных»

Курсовая работа

Выполнил:

студент группы БФИ2201

Амири Ясин

Проверил:

Павликов А. Е.

Москва 2024

Оглавление

[1 Введение 3](#_Toc135165313)

[2 Техническое задание 4](#_Toc135165314)

[3 Ход выполнения работы 5](#_Toc135165315)

[4 Заключение 7](#_Toc135165316)

Введение

В современном мире компьютерные технологии играют ключевую роль в различных сферах деятельности, в том числе и в розничной торговле. Эффективное управление магазином требует точного учета и анализа продаж, что может быть достигнуто с помощью использования различных структур данных и алгоритмов.

В данной курсовой работе предлагается разработать программу, которая позволит рассчитать общую выручку магазина, найти товар, который был продан наибольшее количество раз, и товар, который принес наибольшую выручку. Для решения этих задач будет использована такая структура данных, как хеш-таблица, а также методы сортировки.

Хеш-таблица представляет собой структуру данных, в которой элементы хранятся в виде пар "ключ-значение". Ключ является уникальным числом, используемым для индексации значений. Когда происходит конфликт, то есть несколько ключей генерируют один и тот же индекс, для его разрешения будет использован метод цепочек, который предполагает хранение элементов в одном индексе с помощью двусвязного списка.

Кроме того, в работе будут применены методы сортировки для определения товара, который был продан наибольшее количество раз.

Таким образом, разработанная программа позволит эффективно управлять магазином, анализируя продажи и принимая обоснованные решения.

Техническое задание

Написать программу, которая будет считывать данные из CSV файла, содержащего информацию о продажах товаров в магазине. Данные в файле содержатся в следующем формате: | Номер заказа | Дата заказа | Название товара | Категория товара | Количество продаж | Цена за единицу | Общая стоимость | Необходимо:

1. Рассчитать общую выручку магазина.
2. Найти товар, который был продан наибольшее количество раз.
3. Найти товар, который принес наибольшую выручку.
4. Составить отчет, содержащий информацию об общей выручке магазина, количестве проданных единиц каждого товар и доле каждого товара в общей выручке. Для решения задач необходимо использовать структуры данных, такие как массивы и хеш-таблицы, а также различные алгоритмы обработки данных, например, сортировку и поиск. Также необходимо учитывать возможные ошибки ввода-вывода и обрабатывать их в соответствии с требованиями

Ход выполнения работы

Подключим необходимые библиотеки для работы: pandas, matplotlib.pyplot, csv. После подключения модулей, напишем код для работы с хеш-таблицей. В хеш-таблице мы создаем класс узла списка (Node), после чего создаем класс хэш-таблицы (HashTable). В классе хеш-таблицы пропишем методы для работы: инициализация (\_\_init\_\_), добавление элемента в таблицу (add), получение значения по ключу (get), изменение емкости таблицы (resize).

Создадим функцию для сортировки. Использовать мы будет быструю сортировку, так как она быстрее всех выполняет фильтрацию элементов от наименьшего к наибольшему. Чтобы отсортировать элементы, мы должны создать переменную, в которой будет храниться опорный элемент. С этим опорным элементом будут сравниваться остальные элементы. Для реализации быстрой сортировки создадим 3 массива, в которые будут отправляться элементы, которые меньше опорного, равны опорному, больше опорного. Далее напишем функцию для выполнения расчетов.

Объявим переменные, которые будут хранить: общую выручку и максимального количества проданного товара (rev, max\_tovar); переменную типа string, в которую мы запишем названия товара, проданный максимальное количество раз (max\_tovar\_name); переменную для хранения количество строк в csv файле (count\_row) и три массива: массив для ключей / название товаров продаж и массив для выручки (names\_array, reven\_array).

Используем оператор try-except, который будет обрабатывать возможные ошибки. Объявим переменную, в которой будет наш исходный файл.

Вызовем хеш-таблицу, записав ее в переменную (h1, h2). Будем проходить по всем строкам в файле. В names\_array добавляем товар/ключ. Используя функцию append, добавим в массив строку с индексом 2. Строка с индексом 2 содержит данные о названии товара. В rev добавим строку с индексом 6, которая отвечает за общую стоимость. Добавлять мы будем, преобразовав значение в тип int.

Используя оператор if, найдем товар, который принес который принес наибольшую выручку, а также найдем название этого товара.

Если max\_tovar меньше чем сравниваемый элемент из строки, то обновляем значения в max\_tovar и max\_tovar\_name, которое содержится в строках под индексом 2. В массив для выручки добавим значение выручки из i-ой строки 6 индекса, который отвечает за общую стоимость.

В h2 записываем значение выручки/ключ и товар. Добавляем в h1 товар/ключ и количество его продаж и добавляем товар/ключ и его выручку. Выводим наши значения из хеш-таблиц, используя метод get, который прописали в классе хеш-таблицы. Сортируем массив с выручками. А возвращать мы будем общую выручку, названия товара, который был продан наибольшее количество раз и товар, который принес наибольшую выручку (rev, max\_tovar\_name, ht2.get(str(result[-1]))). Последнее мы будем выводим из хеш-таблицы методом get последний элемент, преобразованный в тип string.

Используем оператор try-except, который будет отвечать за ввод/вывод без ошибок. Объявим переменную, в которой будет наш исходный файл. Вызовем нашу функцию для расчета.

Итоговый вывод оказался таким:

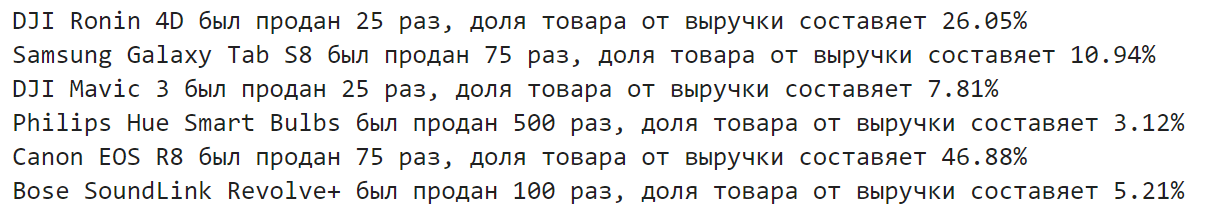


Рисунок 1 – Отчет содержащий информацию количестве проданных единиц каждого товар и доле каждого товара в общей выручке

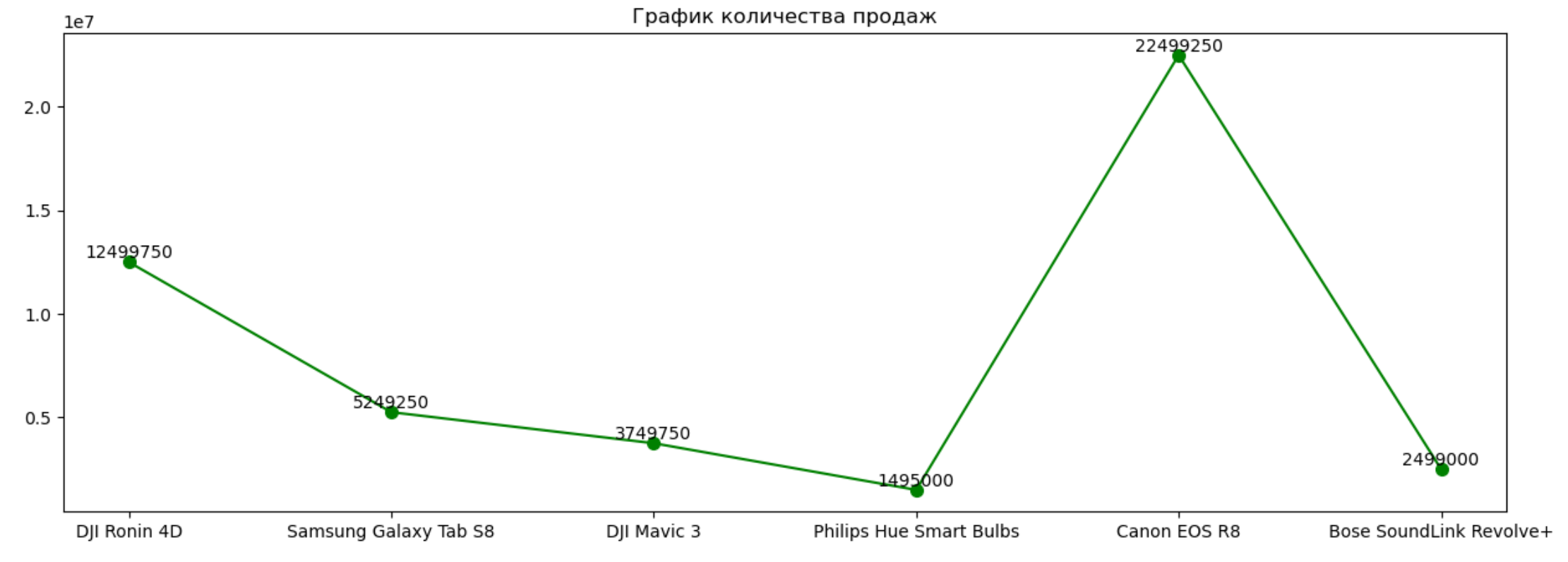


Рисунок 2 – График выручки товаров

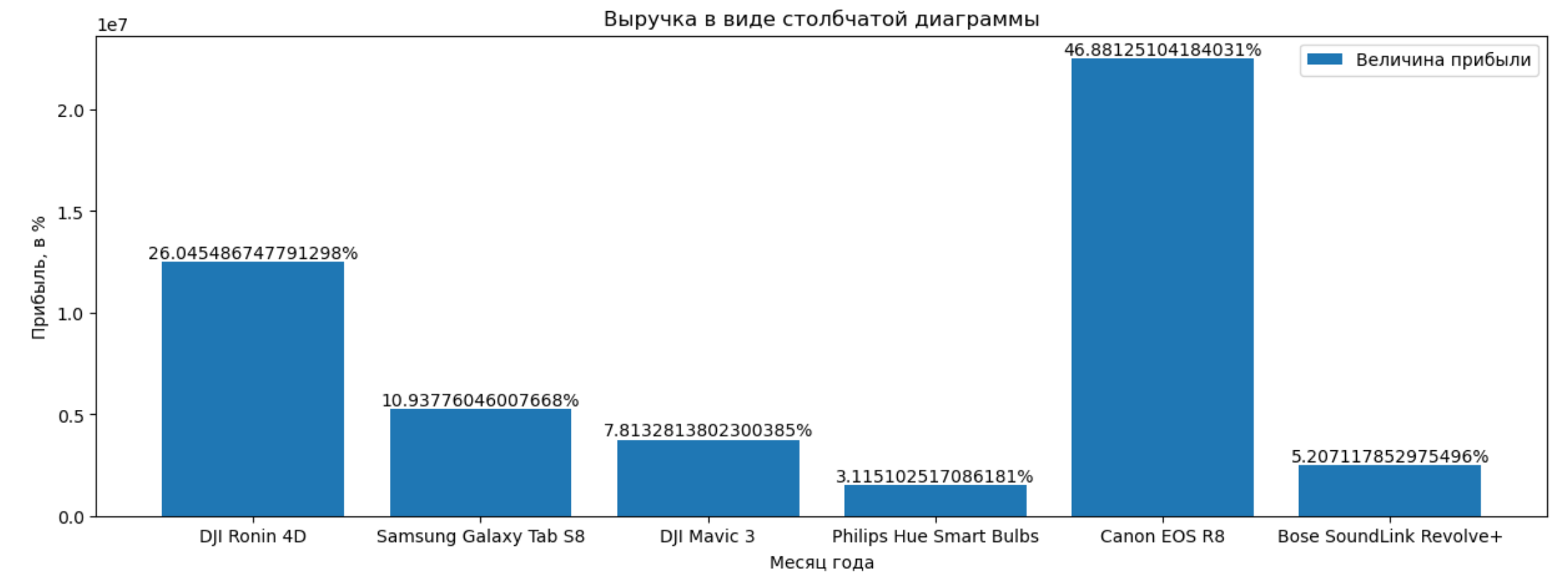


Рисунок 3 – Диаграмма выручки товаров

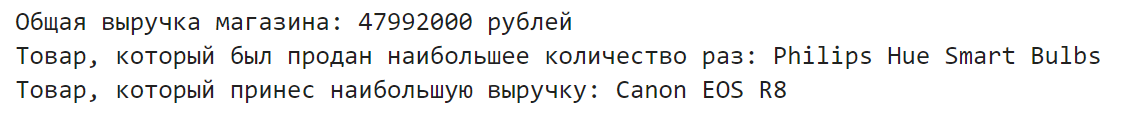


Рисунок 4 – Конечный результат программы.

Заключение

Написал программу, которая будет считывать данные из CSV файла, содержащего информацию о продажах товаров в магазине. Рассчитал общую выручку магазина. Нашёл товар, который был продан наибольшее количество раз. Нашёл товар, который принес наибольшую выручку. Составил отчет, содержащий информацию об общей выручке магазина, количестве проданных единиц каждого товар и доле каждого товара в общей выручке.

В целом, разработанная программа демонстрирует возможность эффективного анализа продаж в магазине на основе данных из CSV файла. Полученные результаты могут быть использованы для оптимизации ассортимента, ценовой политики и других аспектов деятельности магазина, что повышает его конкурентоспособность на рынке.