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»

Курсовая работа

по дисциплине «Структуры и Алгоритмы Обработки Данных»

Выполнил: студент группы БФИ №2203

Кубатин Д. А.

Проверил:

Симонов С. Е.

Москва 2024ОГЛАВЛЕНИЕ

[Введение 3](#_Toc167191290)

[Техническое задание 4](#_Toc167191291)

[Описание реализации 5](#_Toc167191292)

[Заключение 6](#_Toc167191293)

## Введение

Целью данной курсовой работы являлось создание программы, которая будет способна рассчитать общую выручку магазина, найти товары, которые были проданы наибольшее количество раз и принесли наибольшую выручку, а также составить отчёт, содержащий информацию о доле каждого товара в общей выручке.

Для решения поставленных задач были применены такие структуры данных, как хэш-таблицы и массивы, а также следующие алгоритмы обработки данных: быстрая сортировка, сортировка обменом.

## Техническое задание

Написать программу, которая будет считывать данныеиз 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.

## Описание реализации

Для решения поставленной задачи были использованы библиотеки csv, matplotlib и dateutil.

С помощью библиотеки csv данные будут считаны из файла и перемещены в массив для удобства дальнейшего использования.

Перед началом работы с полученными данными необходимо удостовериться в правильности заполнения исходного файла. Для этого была создана переменная errorless, в которой будет храниться результат проверки. В случае если исходный файл не будет совпадать с шаблоном | Номер заказа | Дата заказа | Название товара | Категория товара |Количество продаж | Цена за единицу | Общая стоимость |, одно из полей будет заполнено неправильно или произведение цены за единицу и количества продаж не будет совпадать с общей стоимостью, errorless примет значение «false», программа выведет пользователю сообщение об ошибке и место в файле с исходными данными, в котором была допущена ошибка и не будет производить никаких расчётов.

После получения и проверки данных, информация о каждом виде товаров будет распределена по хеш-таблице. Хэш-ключом служит остаток числа, полученного из переведённых в номера таблицы Unicode каждого из символов в названии товара, на количество позиций в файле с исходными данными.

Для нахождения общей выручки магазина программа складывает каждое значение в колонке общей стоимости из массива с информацией о товарах rows.

Поиск товара, который принёс больше всего выручки, происходит по хэш-таблице: в каждой категории товаров суммируется выручка с каждого чека, после чего эта сумма сравнивается с переменной maxPrice. Если полученная с одного вида товаров выручка будет больше, чем maxPrice, то значение этой переменной будет обновлено, а название товара будет сохранено. Поиск товара, который был продан максимальное количество раз происходит по тому же принципу.

Для нахождения максимальной выручки с одного чека каждая строка хэш-таблицы будет отсортирована с помощью алгоритма быстрой сортировки таким образом, чтобы первым находился товар с наибольшей выручкой. Далее, с помощью алгоритма сортировки обменом (пузырьком) стоки хэш-таблицы будут отсортированы так, чтобы на вершине находилась строка, в которой у первого элемента была максимальная выручка, после чего информация об этом элементе будет выведена пользователю. Поиск товара, который был продан наибольшее количество раз за один чек, происходит точно так же.

Наконец, для формирования отчёта о продажах магазина программа перебирает каждый вид товаров и суммирует количество продаж товара и выручку с каждого чека, после чего полученная информация выводится пользователю. Далее, информация о каждом товаре и его название сохраняются и с помощью библиотеки matplotlib выводятся пользователю в виде диаграмм.

## Заключение

В ходе данной курсовой работы была создана программа, которая способна предоставить отчёт, содержащий информацию о доле каждого товара в общей выручке и количестве их продаж (рис. 1), отобразить информацию о количестве проданных товаров и выручке с них в виде столбчатых диаграмм (рис. 2 и 3), а так же наглядно отображает долю каждого вида товаров в общей выручке магазина в виде круговой диаграммы (рис. 4).

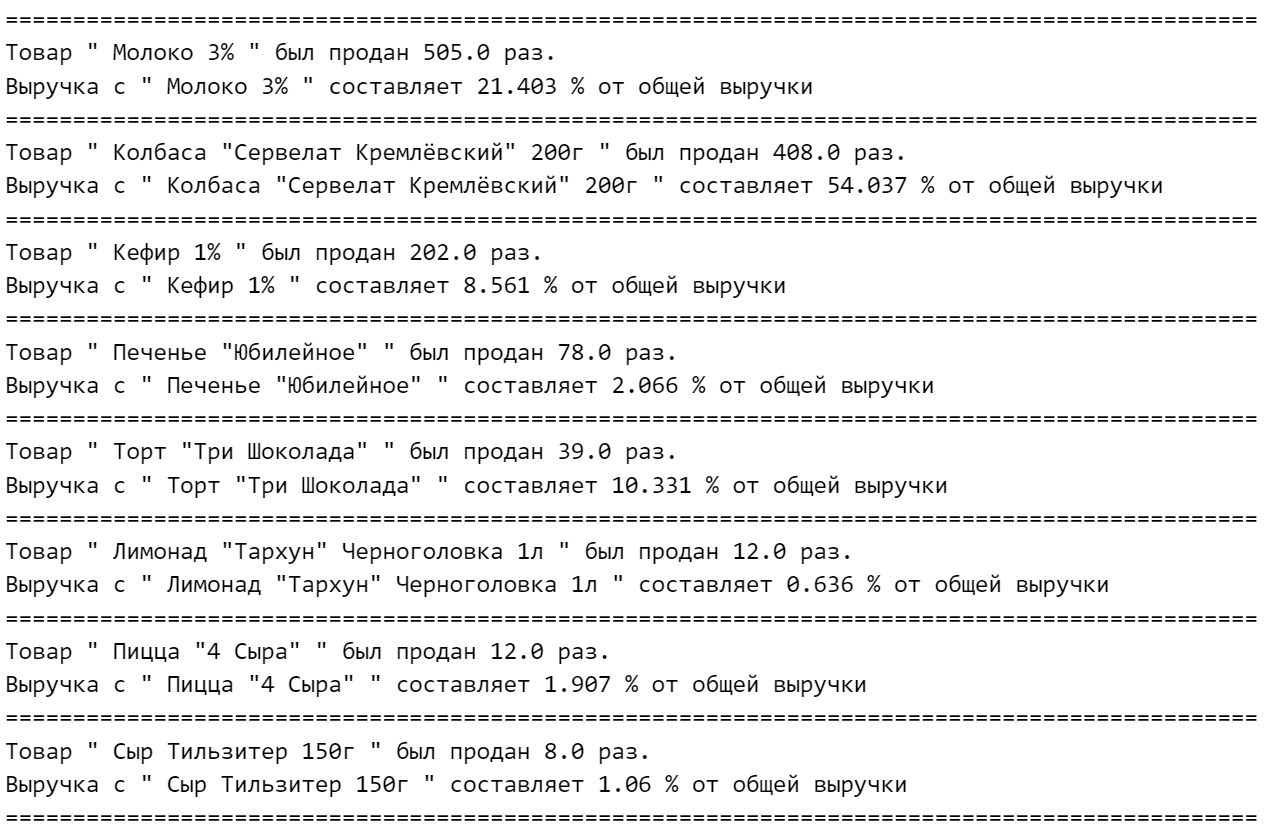


Рисунок 1 – отчёт о продажах магазина

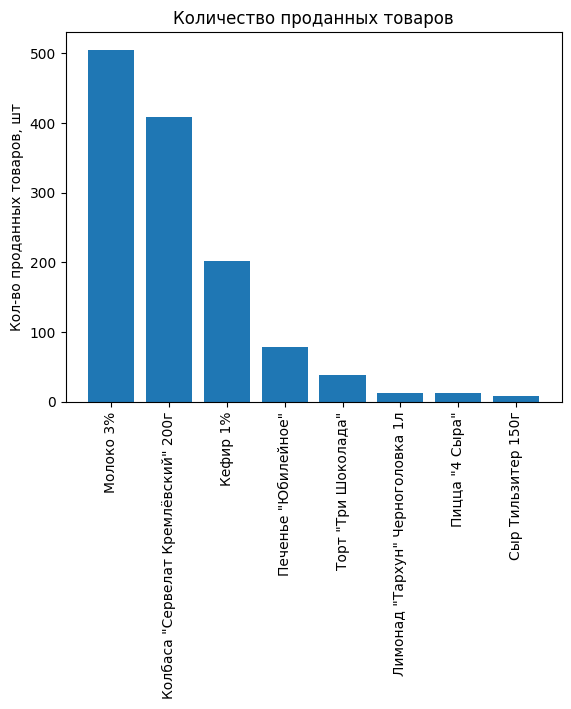


Рисунок 2 – диаграмма продаж товаров

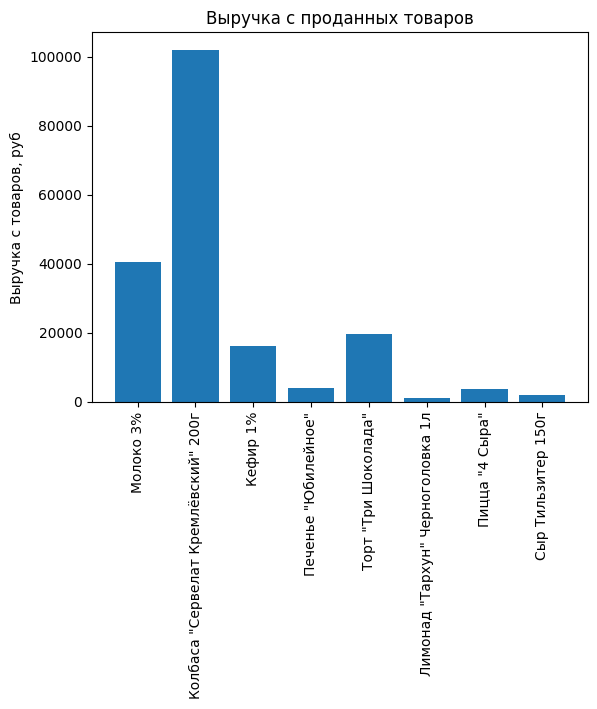


Рисунок 3 – диаграмма выручки с товаров

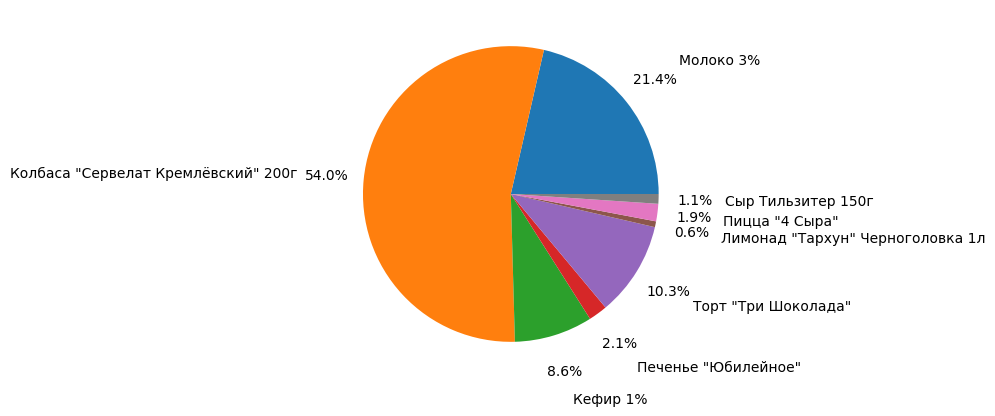


Рисунок 4 –доля каждого вида продуктов в общей выручке магазина в виде круговой диаграммы