Министерство цифрового развития, связи и массовых коммуникаций РФ ордена Трудового Красного Знамени федеральное государственное бюджетное образовательное учреждение высшего образования

«МОСКОВСКИЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ СВЯЗИ И

ИНФОРМАТИКИ»

Кафедра «Информатика»

Курсовая работа

по дисциплине «Структуры и алгоритмы обработки данных»

Выполнил: студент группы БФИ2204

Яроцкий М. В.

Проверил: Симонов С. Е.

Москва 2024

**Цель:** Изучить структуры данных, такие как массивы и хеш-таблицы, а так же алгоритмы сортировки данных и поиска для этих типов данных.

**Задача:** 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 товара и доле каждого товара в общей выручке.

**Ход работы**

1) Функция `read\_data` считывает данные из CSV-файла с разделителем `|` и преобразует их в список словарей, где каждый словарь представляет строку из файла с ключами: `order\_number`, `order\_date`, `product\_name`, `category`, `quantity`, `price\_per\_unit` и `total\_cost`. Она пропускает заголовок, удаляет лишние пробелы, преобразует типы данных для количественных и числовых полей, и возвращает структурированные данные для дальнейшего анализа и обработки.

Код функции

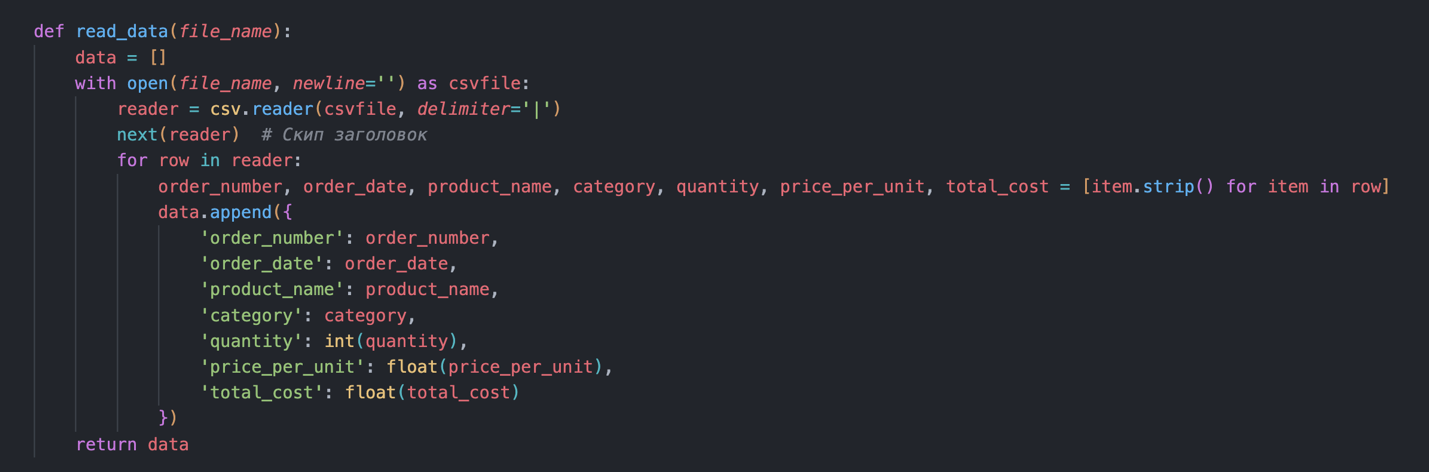


Рисунок 1 – функция read\_data

2) Функция `calculate\_total\_revenue` принимает на вход список словарей с данными о продажах и вычисляет общую выручку, суммируя значения поля `total\_cost` для каждой записи. Функция возвращает общую выручку как число с плавающей запятой, что позволяет использовать результат для дальнейшего анализа финансовых показателей.

Код функции



Рисунок 2 – функция calculate\_total\_revenue

3) Функция `find\_popular\_product` принимает на вход список словарей с данными о продажах и определяет самый продаваемый товар, основываясь на количестве проданных единиц. Функция создает словарь, в котором ключами являются названия товаров, а значениями - общее количество проданных единиц для каждого товара. Затем она находит товар с наибольшим количеством продаж и возвращает его название.

Код функции

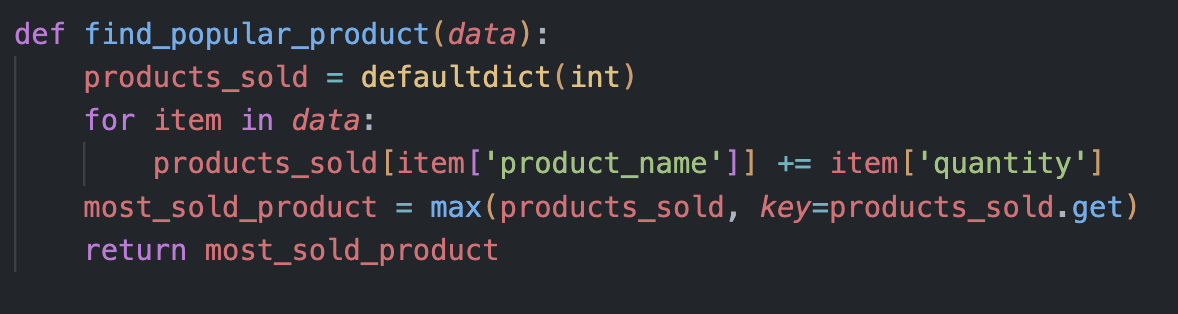


Рисунок 3 – функция find\_popular\_product

4) Функция `find\_most\_revenue\_product` принимает на вход список словарей с данными о продажах и определяет товар, который принес наибольшую выручку. Функция создает словарь, где ключами являются названия товаров, а значениями - суммарная выручка для каждого товара. Затем она находит товар с наибольшей суммарной выручкой и возвращает его название.

Код функции

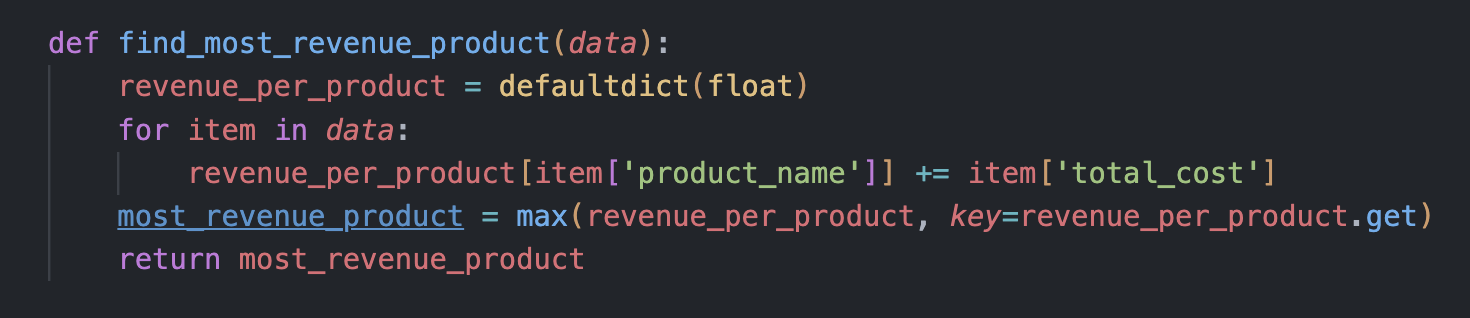


Рисунок 4 – функция find\_most\_revenue\_product

5) Функция `print\_sales\_info` принимает на вход список словарей с данными о продажах и выводит на экран сводную информацию о продажах. Она вычисляет и выводит общую выручку магазина, определяет и выводит самый продаваемый товар и товар, принесший наибольшую выручку. Затем функция выводит подробную информацию о продажах каждого товара, включая количество проданных единиц, выручку и долю в общей выручке. Для этого она использует вспомогательные функции `calculate\_total\_revenue`, `find\_popular\_product` и `find\_most\_revenue\_product`.

Код функции

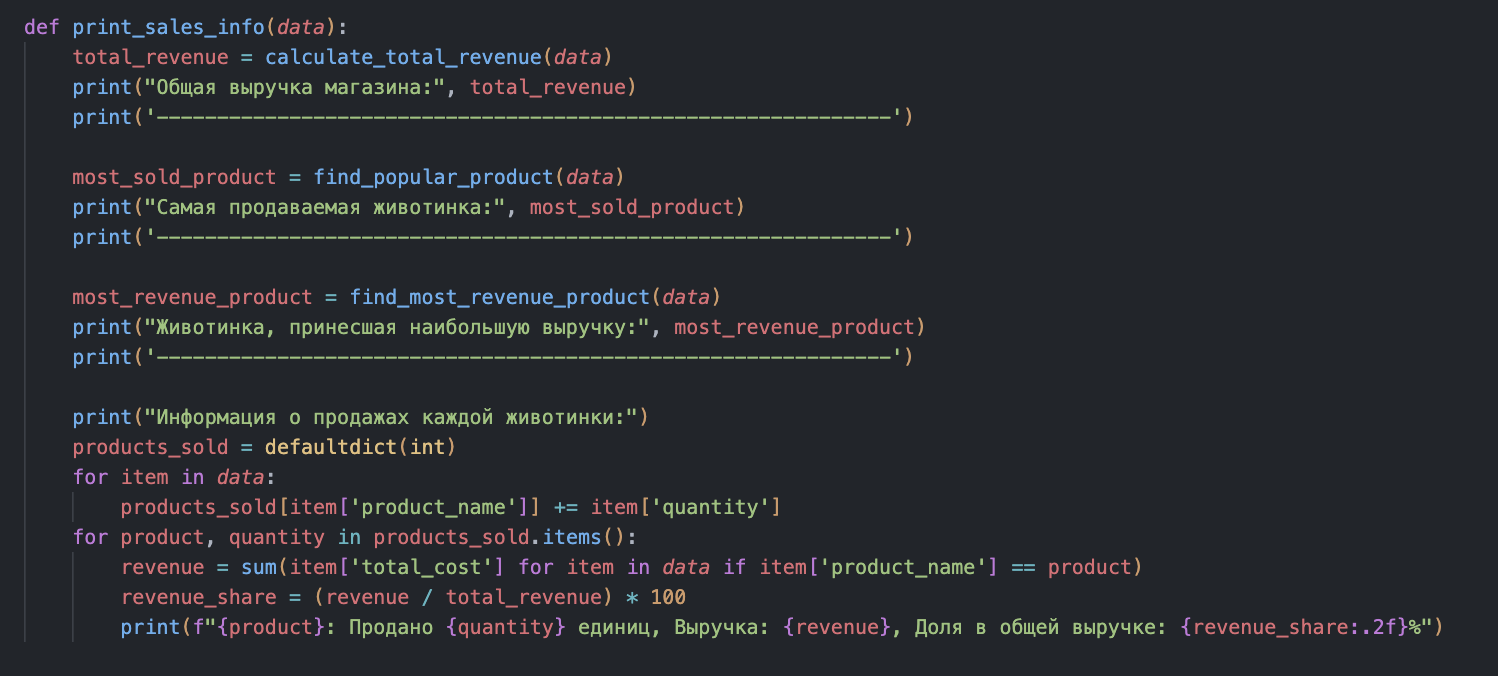


Рисунок 4 – функция print\_sales\_info

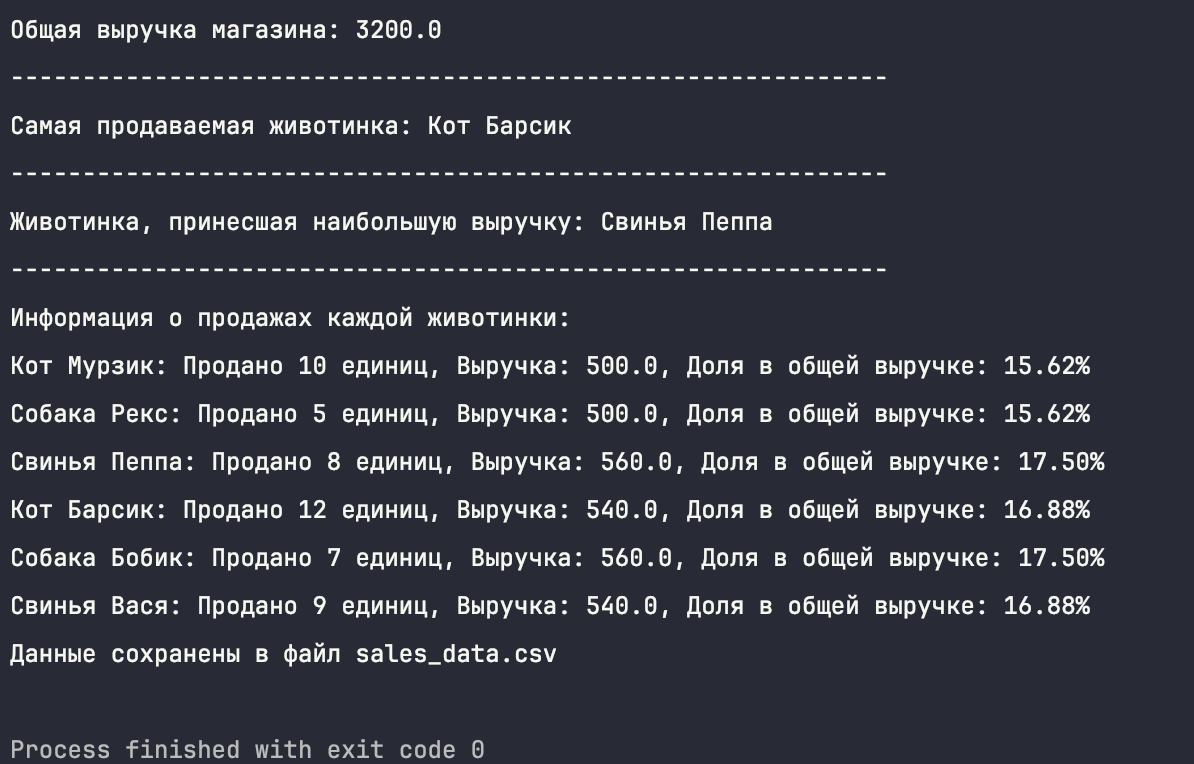


Рисунок 5 – пример вывода

Также в программе присутствует обработка исключений на чтение файла

Код

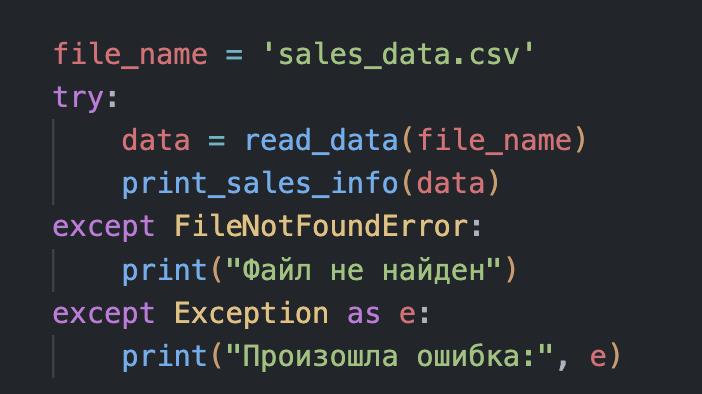


Рисунок 6 – обработка исключений

Данный код пытается считать данные из CSV-файла с именем `sales\_data.csv` с помощью функции `read\_data`, затем передает эти данные функции `print\_sales\_info` для вывода сводной информации о продажах. Если файл не найден, выводится сообщение "Файл не найден". В случае возникновения любой другой ошибки выводится сообщение "Произошла ошибка" с подробностями об ошибке.

Была добавлена возможность вывода отчета о продажах в xlsx файл

Код функции

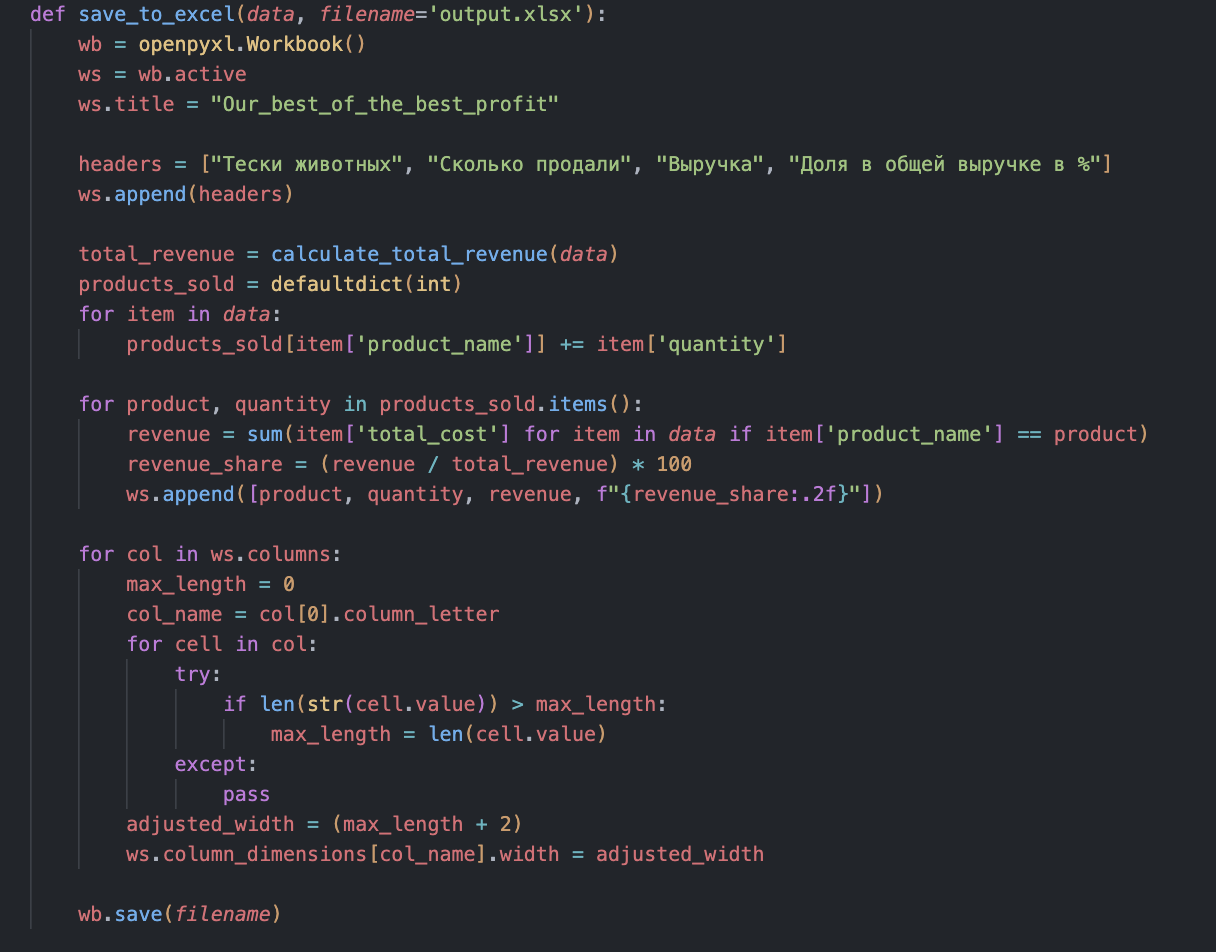


Рисунок 7 – вывод в xlsx

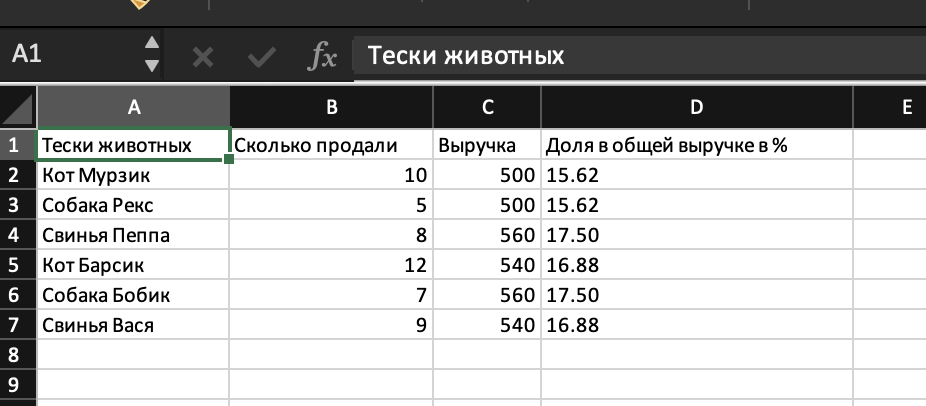


Рисунок 8 – вывод в excel

Функция `save\_to\_excel` принимает данные о продажах в формате списка словарей и сохраняет их в файл Excel с указанным именем `output.xlsx`. Функция сначала создает новую книгу Excel и активирует первый лист, задавая ему название "Our\_best\_of\_the\_best\_profit". Затем добавляет заголовки столбцов, рассчитывает общую выручку, количество проданных единиц для каждого продукта и заполняет лист соответствующими данными. Для каждого продукта функция вычисляет долю его выручки в общей выручке и добавляет эту информацию в лист. После этого функция автоматически подгоняет ширину столбцов под содержимое для улучшения читаемости. В конце функция сохраняет книгу Excel на диск под заданным именем.

**Вывод :**

В процессе написания данной программы были изучены и применены основные методы работы с CSV-файлами в Python, такие как чтение и обработка данных с использованием модуля `csv`. Были реализованы функции для вычисления общей выручки, определения самого продаваемого товара и товара с наибольшей выручкой, что способствовало пониманию принципов агрегирования и анализа данных. Также была усвоена работа с коллекциями, такими как списки и словари, включая использование `defaultdict` для упрощения подсчета и суммирования значений. Дополнительно был изучен подход к обработке исключений для обеспечения устойчивости кода при возникновении ошибок, таких как отсутствие файла. Была добавлена возможность вывода данных в excel.