

**МИНОБРНАУКИ РОССИИ**

**федеральное государственное бюджетное образовательное учреждение**

**высшего образования**

**«Московский государственный технологический университет «СТАНКИН»**

**(ФГБОУ ВО «МГТУ «СТАНКИН»)**

|  |  |
| --- | --- |
| **Институт  информационных технологий** | **Кафедра информационных систем** |

|  |  |  |  |
| --- | --- | --- | --- |
| **Основная образовательная программа 09.03.02 «Информационные системы и технологии»**  **Отчет по дисциплине «Технологии интеллектуального анализа данных»** | | | |
| **по лабораторной работе №1-4** | | | |
| **Тема: «EXEL – УНИВЕРСАЛЬНАЯ СИСТЕМА ОБРАБОТКИ ДАННЫХ»** | | |
|  | | |
| **Проверил**  **преподаватель кафедры ИС** |  | **Саркисова И.О.** | |
|  | подпись |  | |
| **Выполнил**  **студент группы ИДБ-22-06** |  | **Мустафаева П.М.** | |
|  | подпись |  | |

**ОГЛАВЛЕНИЕ**

[**ЛАБОРАТОРНАЯ РАБОТА №1 3**](#_Toc162877615)

[**ЛАБОРАТОРНАЯ РАБОТА №2 8**](#_Toc162877616)

[**ЛАБОРАТОРНАЯ РАБОТА №3 16**](#_Toc162877617)

[**ЛАБОРАТОРНАЯ РАБОТА №4 20**](#_Toc162877618)

**ЛАБОРАТОРНАЯ РАБОТА №1**

**ЦЕЛЬ РАБОТЫ**

Изучение возможностей прикладного программного обеспечения для организации данных и их анализа.

**ПРЕДМЕТНАЯ ОБЛАСТЬ**

Предметная область: пункты упаковки багажа в аэропортах.

**ИНФОРМАЦИОННЫЕ ОБЪЕКТЫ, ИХ АТРИБУТЫ, ТИПЫ ЗНАЧЕНИЙ, СВЯЗИ**

Сущность: предоставление услуг по упаковке багажа в аэропортах.

Атрибуты: тип упаковки, производитель, уровень защиты.

Тип связи: многое ко многим.

**МОДЕЛЬ ОСТЕРВАЛЬДЕРА**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Ключевые партнеры**  -поставщики материалов  -авиакомпании  -аэропорт  -доставщики и логистические компании | **Ключевые деятельности**  -упаковка и защита багажа | **Ценностные предложения**  -различные варианты упаковки  -быстрота и эффективность  -разнообразие услуг  -профессиональное обслуживание | | **Отношения с клиентами**  -качество обслуживания  -доступность  -программа лояльности и бонусы  -обратная связь и улучшение сервиса | **Потребительские сегменты**  -клиенты, которые нуждаются в защите багажа во время его транспортировки  -туристы  - туристические группы |
| **Ключевые ресурсы**  -упаковочные материалы  -оборудование и инвентарь  -персонал | **Каналы сбыта**  -пункты упаковки внутри аэропорта  -туристические агентства |
| **Структура издержек**  -сырье и материалы  -оборудование  -аренда помещения | | | **Потоки доходов**  -оплата за упаковку багажа  -дополнительные услуги (хрупкие предметы)  -партнерство с авиакомпаниями | | |

**ВИЗУАЛЬНОЕ ОПИСАНИЕ ПРЕДМЕТНОЙ ОБЛАСТИ**



Рис. 1.1 Визуализация точек упаковки багажа

**СТРУКТУРА СПИСКА**

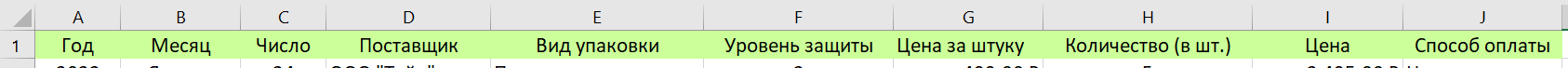
****

Рис. 1.2 Названия полей

**РАСКРЫВАЮЩИЕСЯ СПИСКИ**

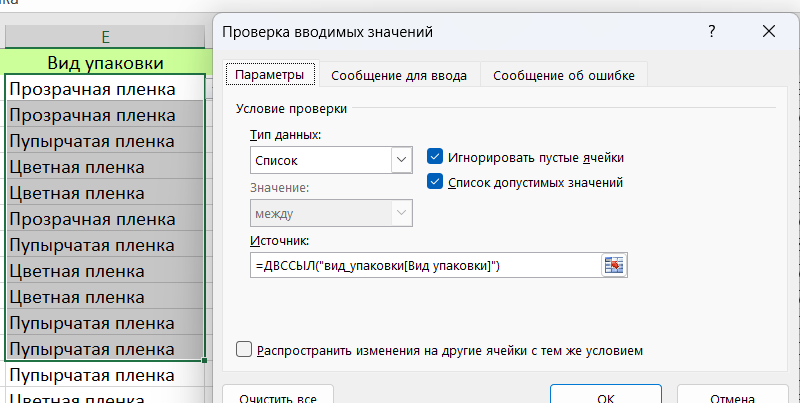
****

Рис. 1.3 Данные для раскрывающегося списка

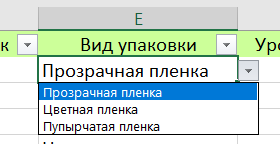
****

Рис. 1.4 Раскрывающийся список для поля «Вид упаковки»

**ЗАПОЛНЕННАЯ БД**

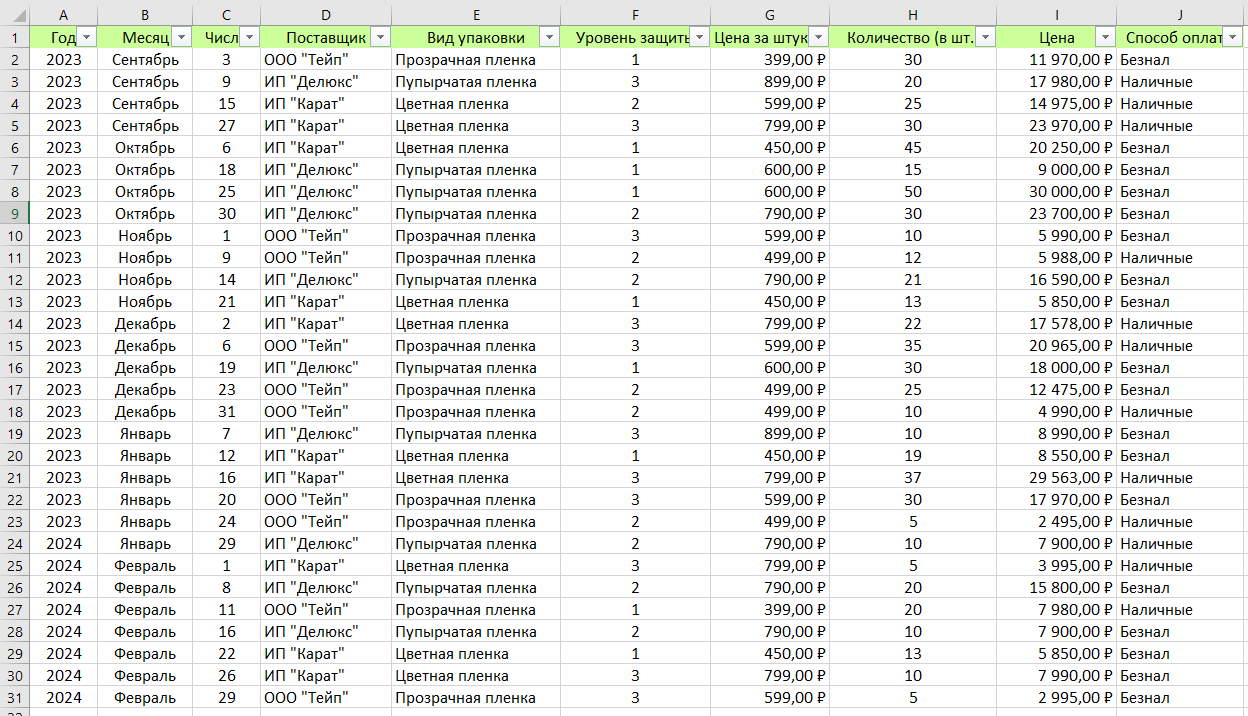


Рис. 1.5 Заполненная база данных

**ВЫБОРКА ПО ДВУМ УСЛОВИЯМ**

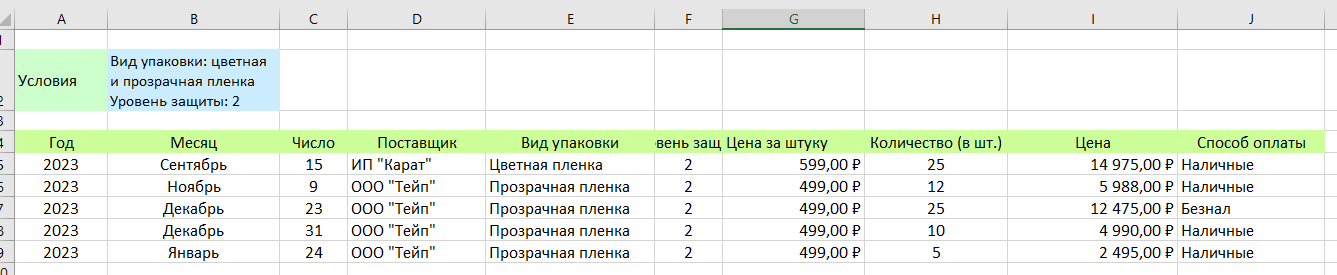
****

Рис. 1.6 Результат выборки по двум условиям

**АВТОФИЛЬТР**

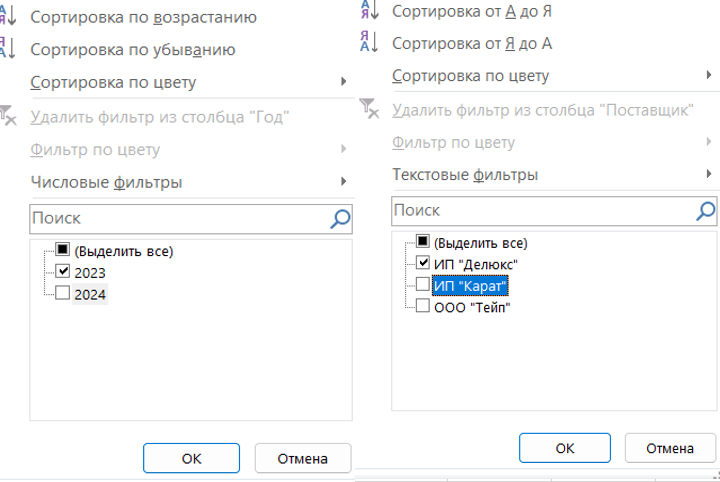
****

Рис. 1.7 Условия автофильтра

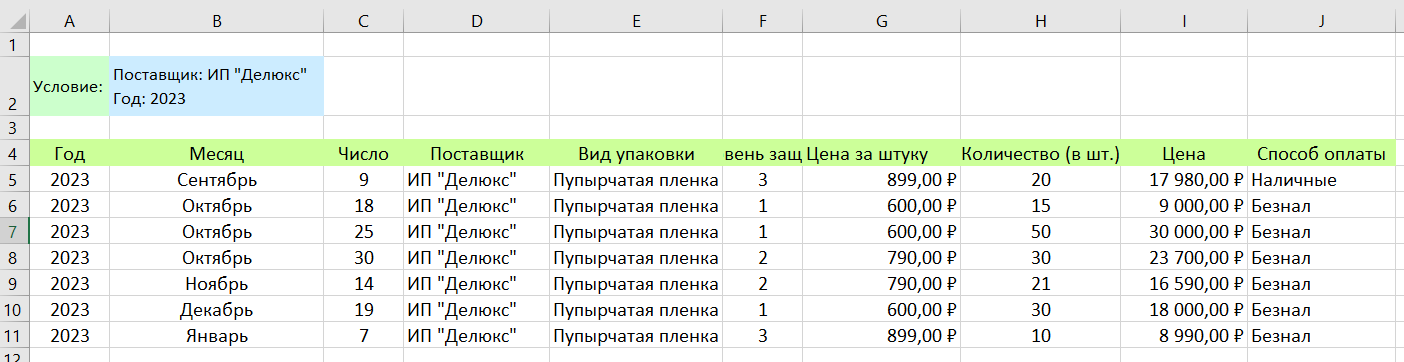
****

Рис. 1.8 Результат автофильтра

**РАСШИРЕННЫЙ ФИЛЬТР**

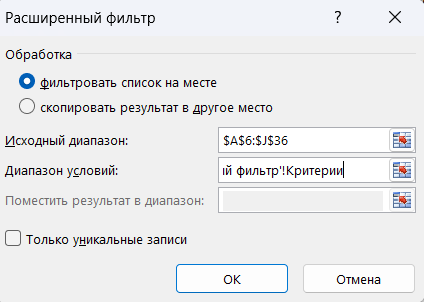
****

Рис. 1.9 Условия для расширенного фильтра

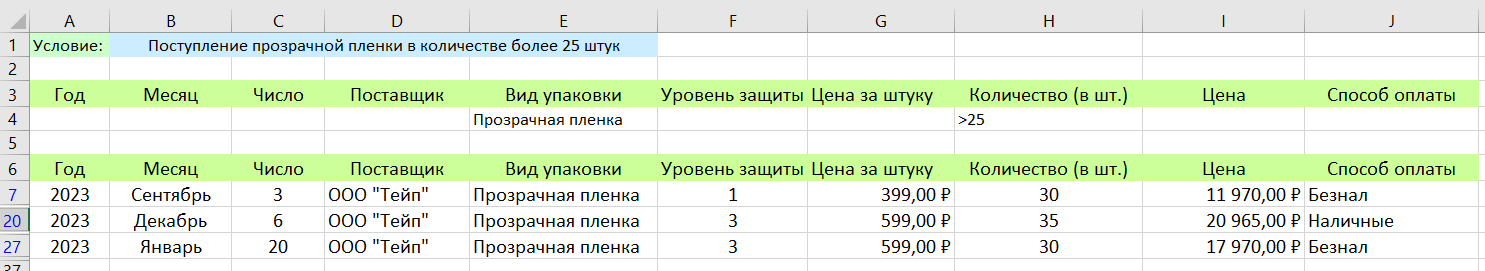
****

Рис. 1.10 Результат расширенного фильтра

**НАСТРАИВАЕМАЯ СОРТИРОВКА**

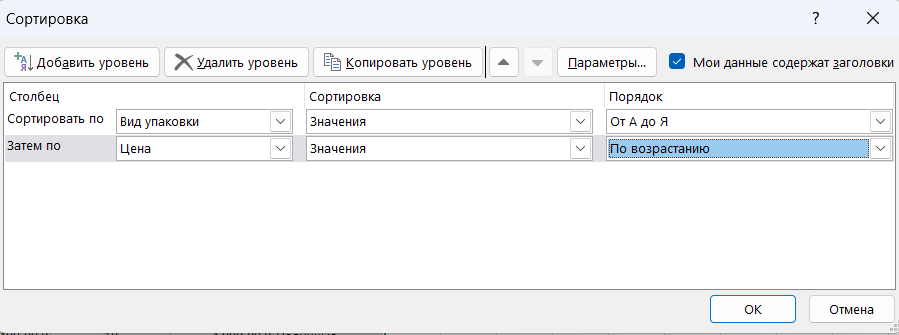
****

Рис. 1.11 Условия для настраиваемой сортировки

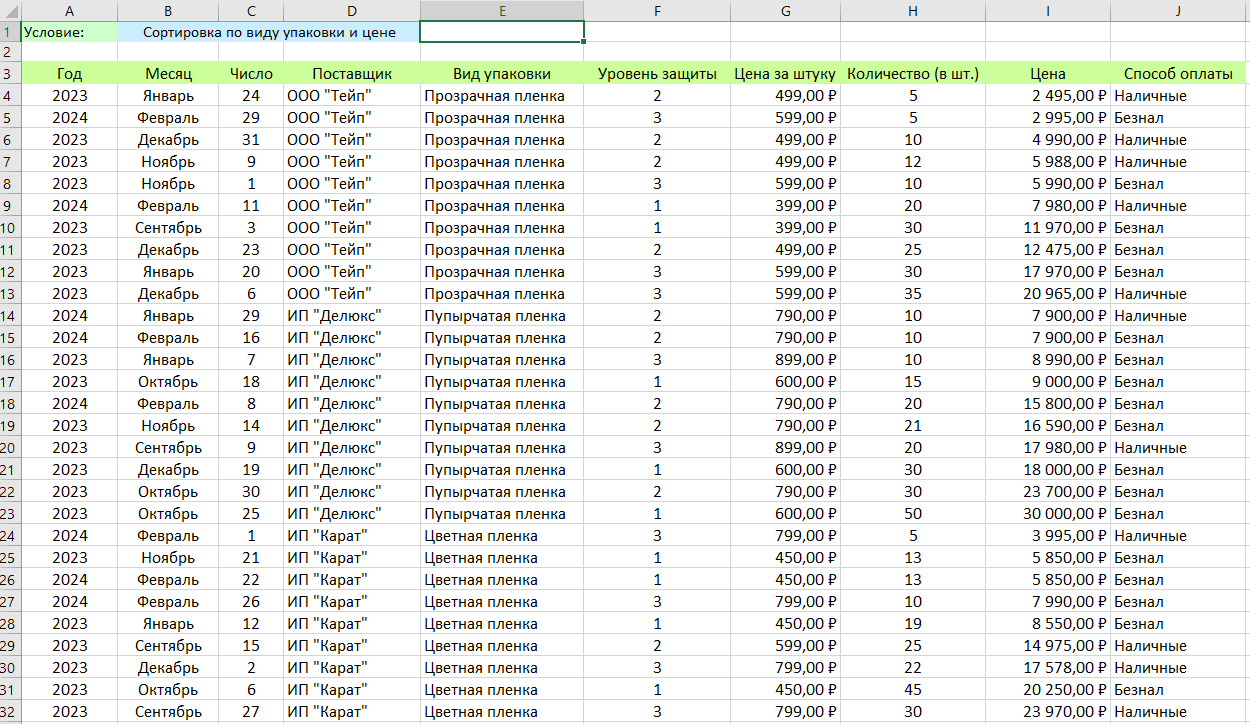
****

Рис. 1.12 Результат настраиваемой сортировки

**ЛАБОРАТОРНАЯ РАБОТА №2**

**ЦЕЛЬ РАБОТЫ**

Изучение OLAP-технологий и возможностей прикладного программного обеспечения для организации данных и их анализа.

**СТРУКТУРА ТАБЛИЦЫ**

****

Рис. 2.1 Структура таблицы

**ОПИСАНИЕ OLAP-КУБА**

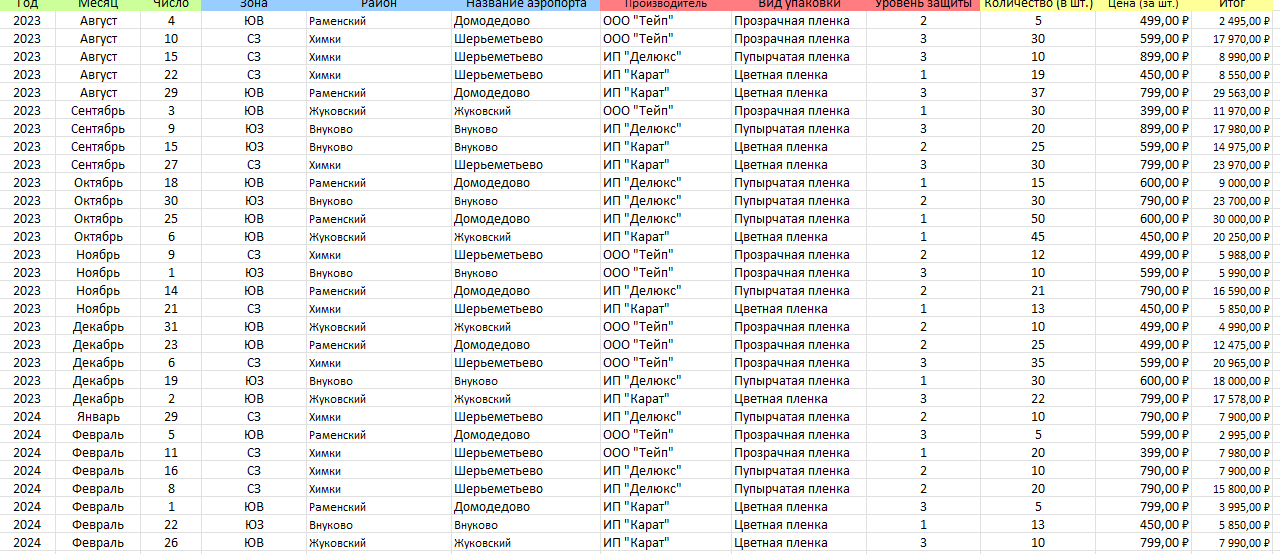
****

Рис. 2.2 Таблица, описывающая OLAP-куб с учетом иерархии измерений

Метрика – Итог (Цена (за шт.) \* Количество (в шт.)).

Заданные иерархии:

* Измерение «Товар»: производитель / тип товара / уровень защиты
* Измерение «Местоположение»: зона / район / название аэропорта
* Измерение «Дата»: год / месяц / число

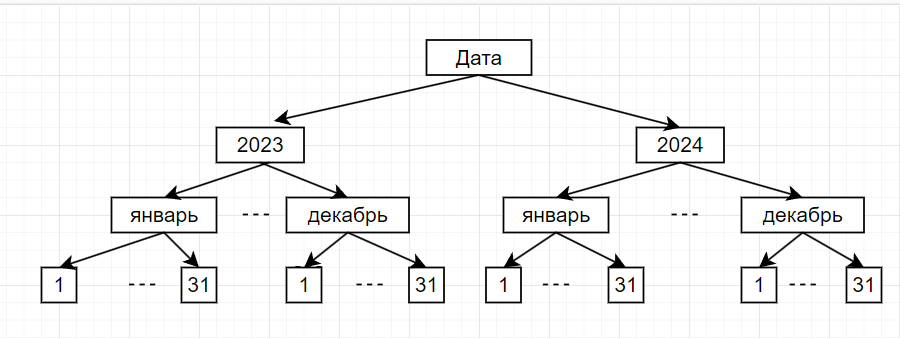


Рис. 2.3 Иерархия «Дата»

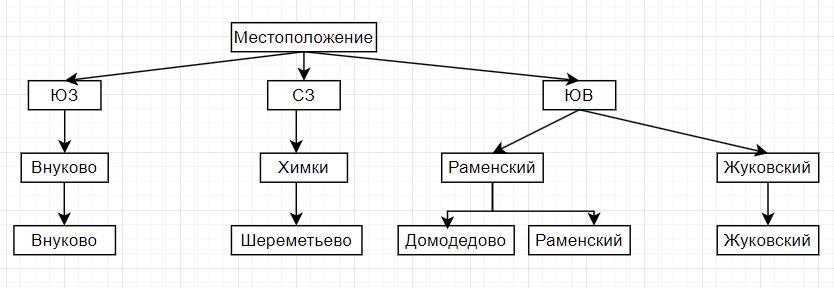


Рис. 2.4 Иерархия «Местоположение»

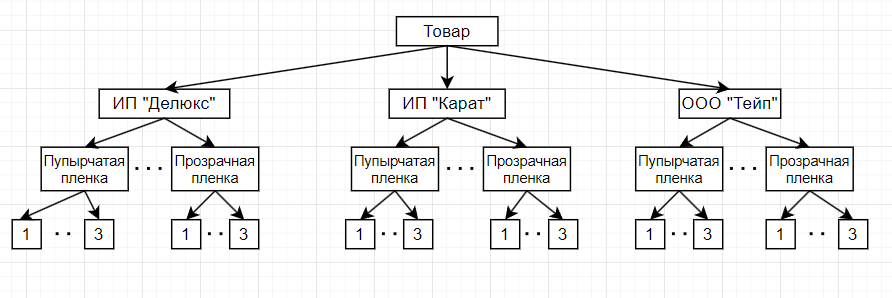


Рис. 2.5 Иерархия «Товар»

**ВИЗУАЛИЗАЦИЯ OLAP-КУБА**

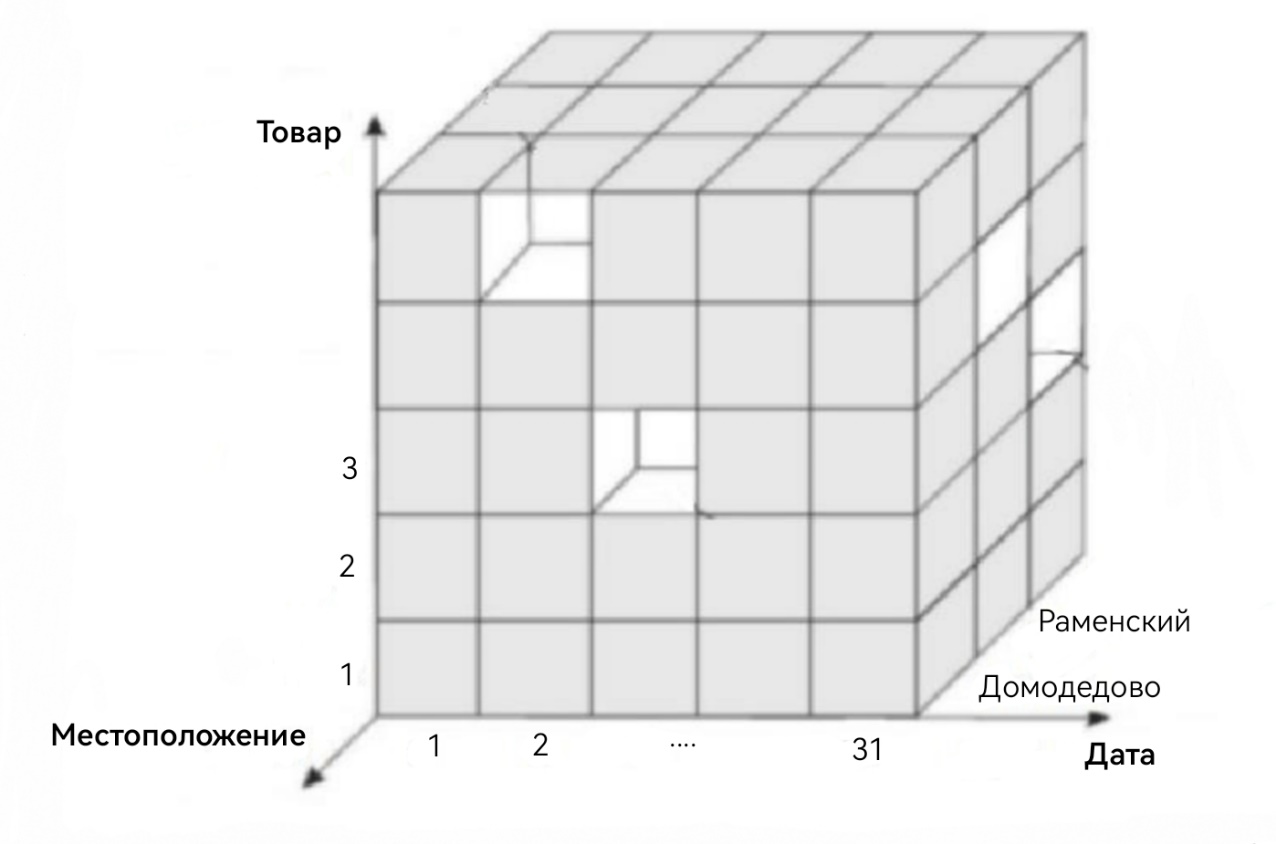
****

Рис. 2.6 Визуализация OLAP-куба

**СВЯЗЬ МЕЖДУ ИЗМЕРЕНИЯМИ И ФАКТАМИ**

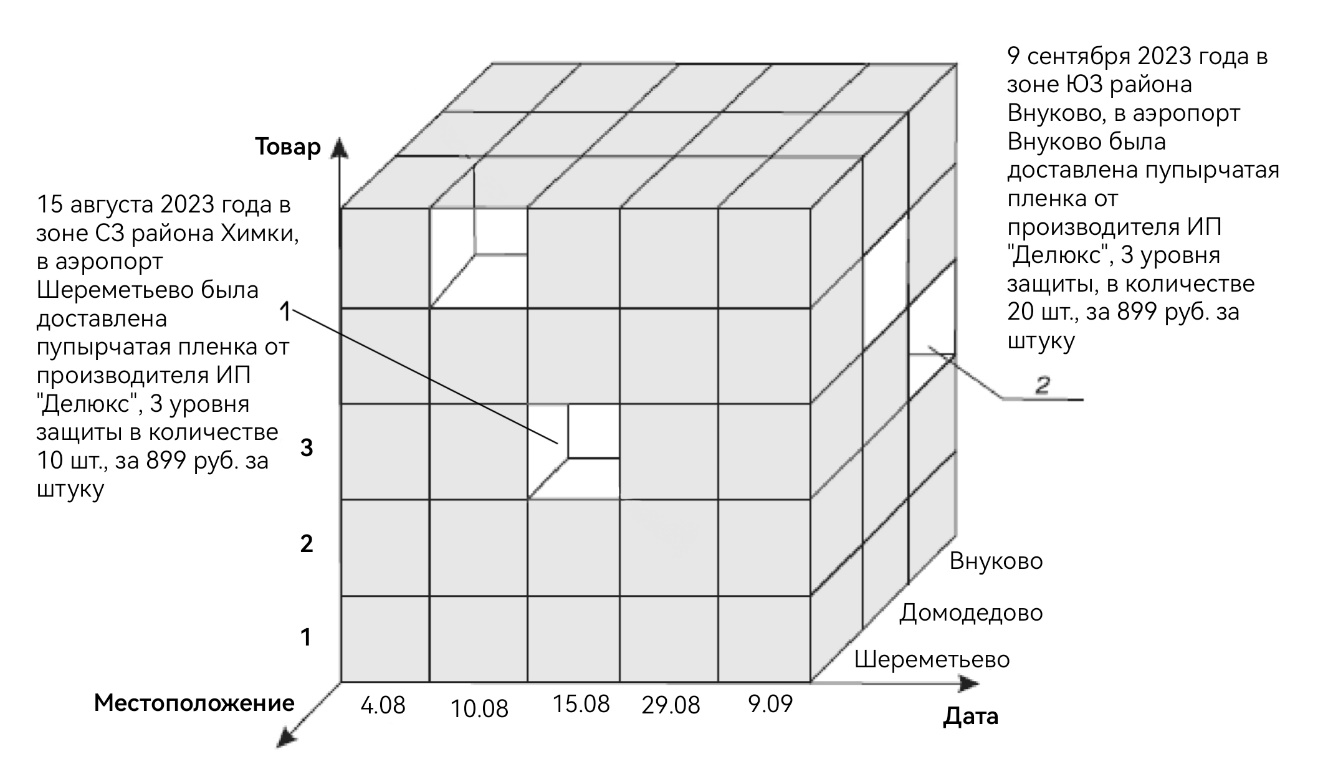


Рис. 2.7 Визуализация связи между измерениями и фактами

**РЕАЛИЗАЦИЯ ПРОСТОГО СРЕЗА**

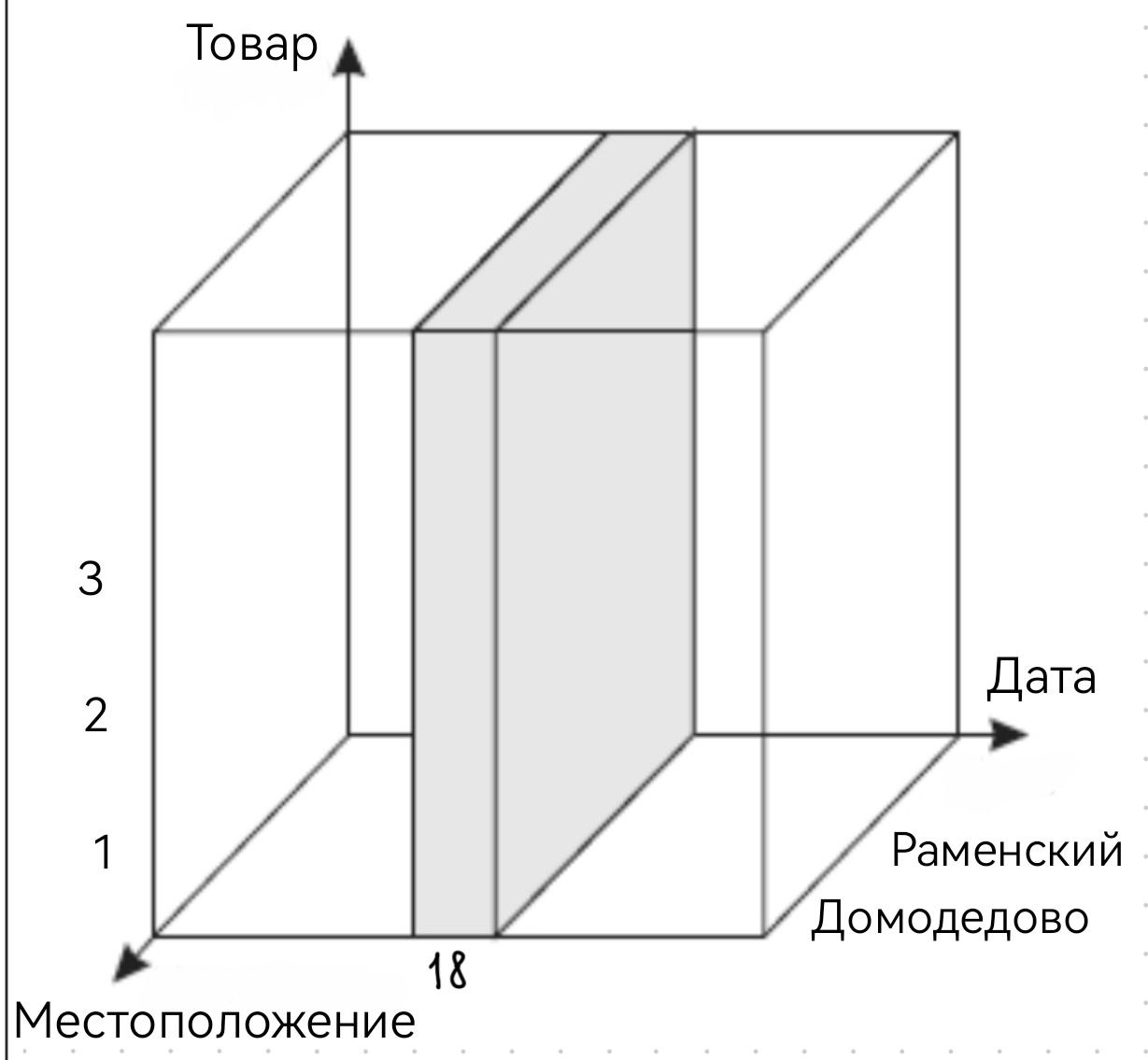


Рис. 2.8 Визуализация простого среза

В результате простого среза получаем сводную таблицу с информацией о прибыли за 18 октября 2023 года.

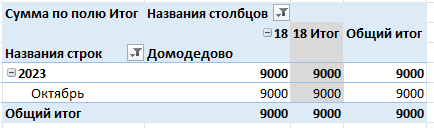


Рис. 2.9 Сводная таблица для простого среза

**РЕАЛИЗАЦИЯ ДВОЙНОГО СРЕЗА**

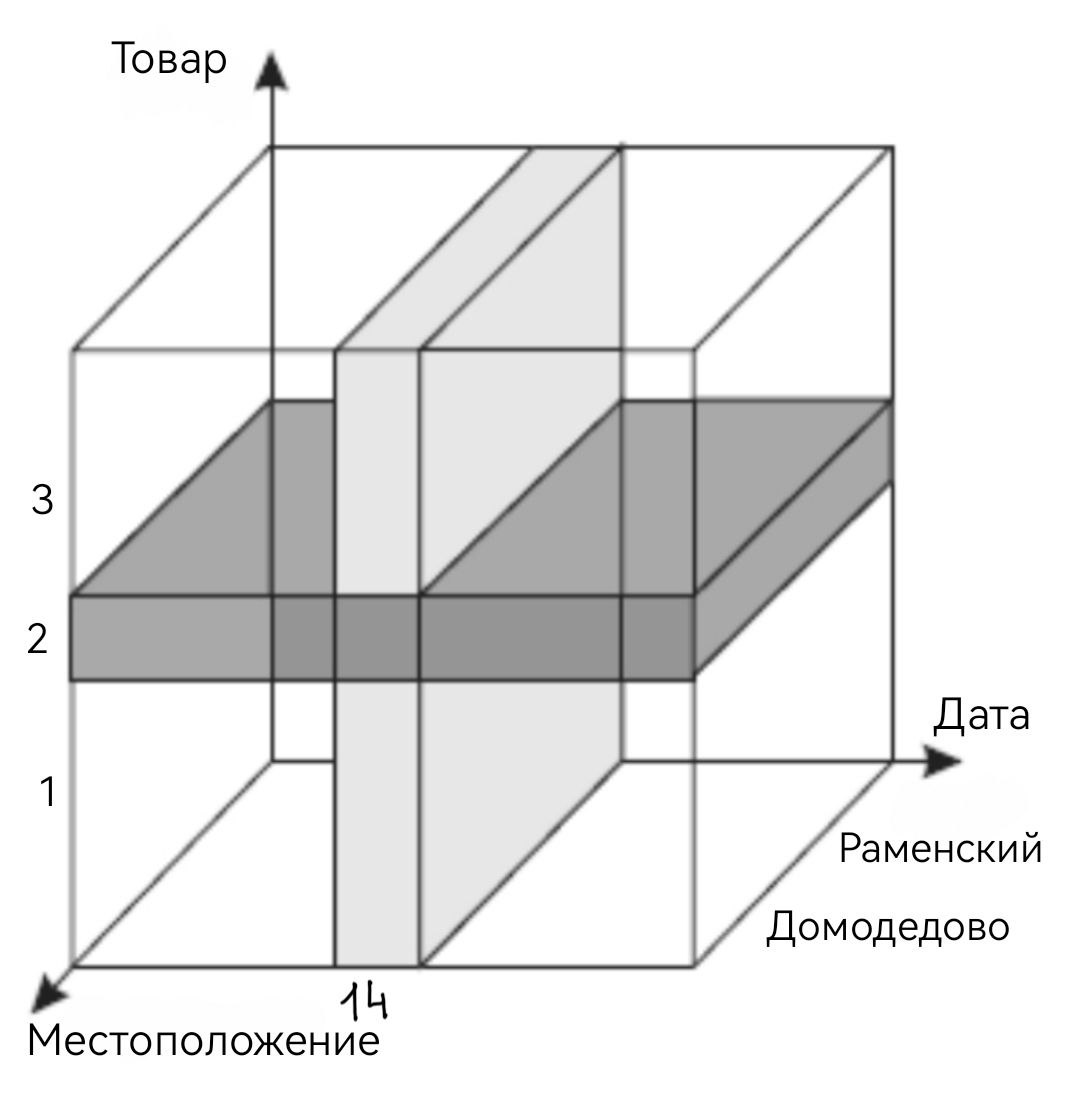
****

Рис. 2.10 Визуализация двойного среза

В результате двойного среза получаем сводную таблицу с информацией о прибыли от продажи пупырчатой пленки 2 уровня защиты, 14 ноября 2023 года.

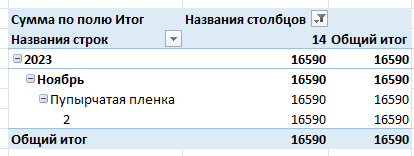


Рис. 2.11 Сводная таблица для двойного среза

**ТРАНСПОНИРОВАНИЕ КУБА**

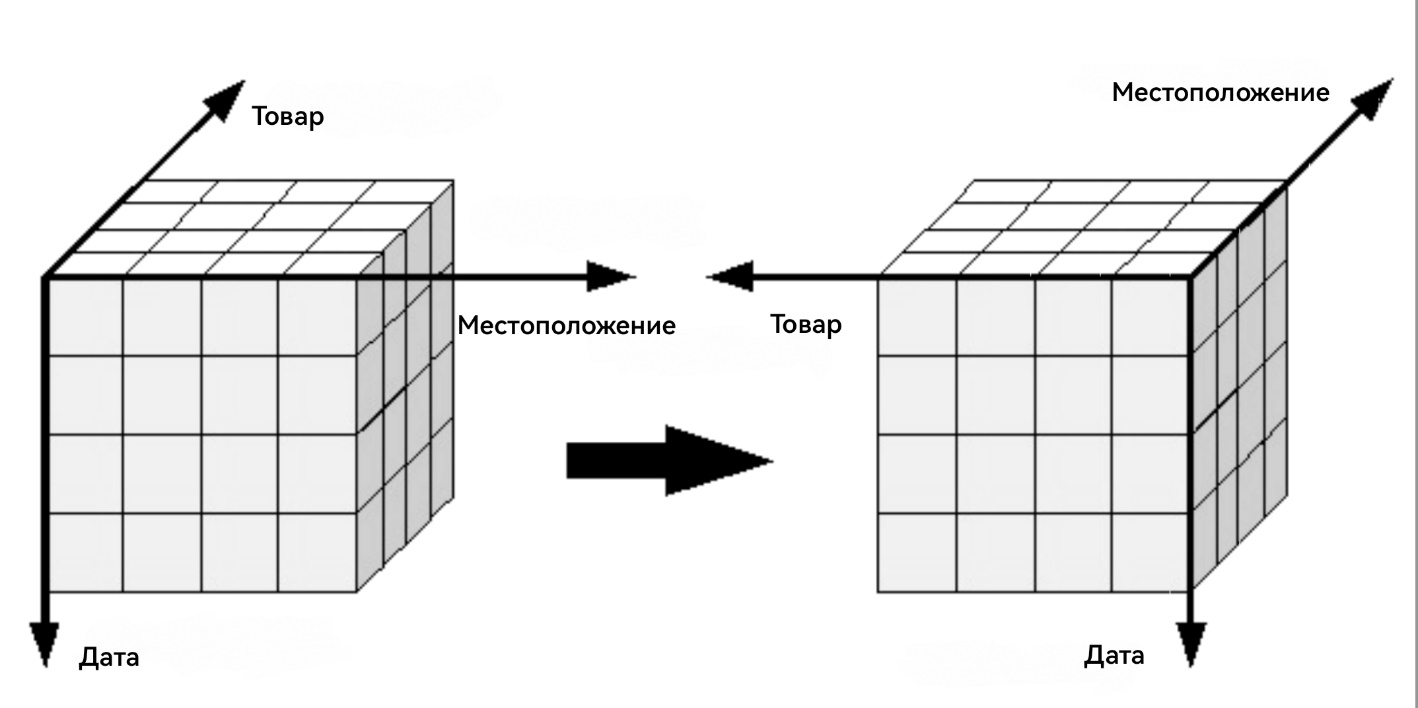
****

Рис. 2.12 Транспонирование куба

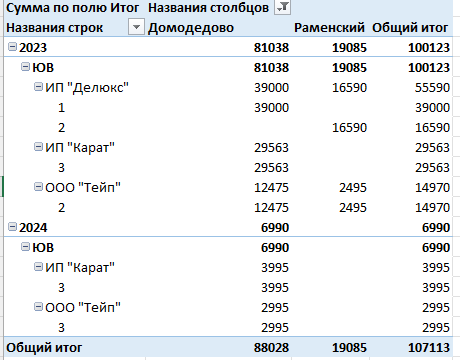


Рис. 2.13 Сводная таблица для не транспонированного куба

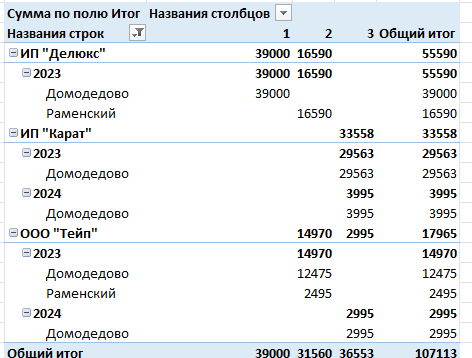


Рис. 2.14 Сводная таблица для транспонированного куба

**СВЁРТКА ПО ОДНОМУ ИЗМЕРЕНИЮ**

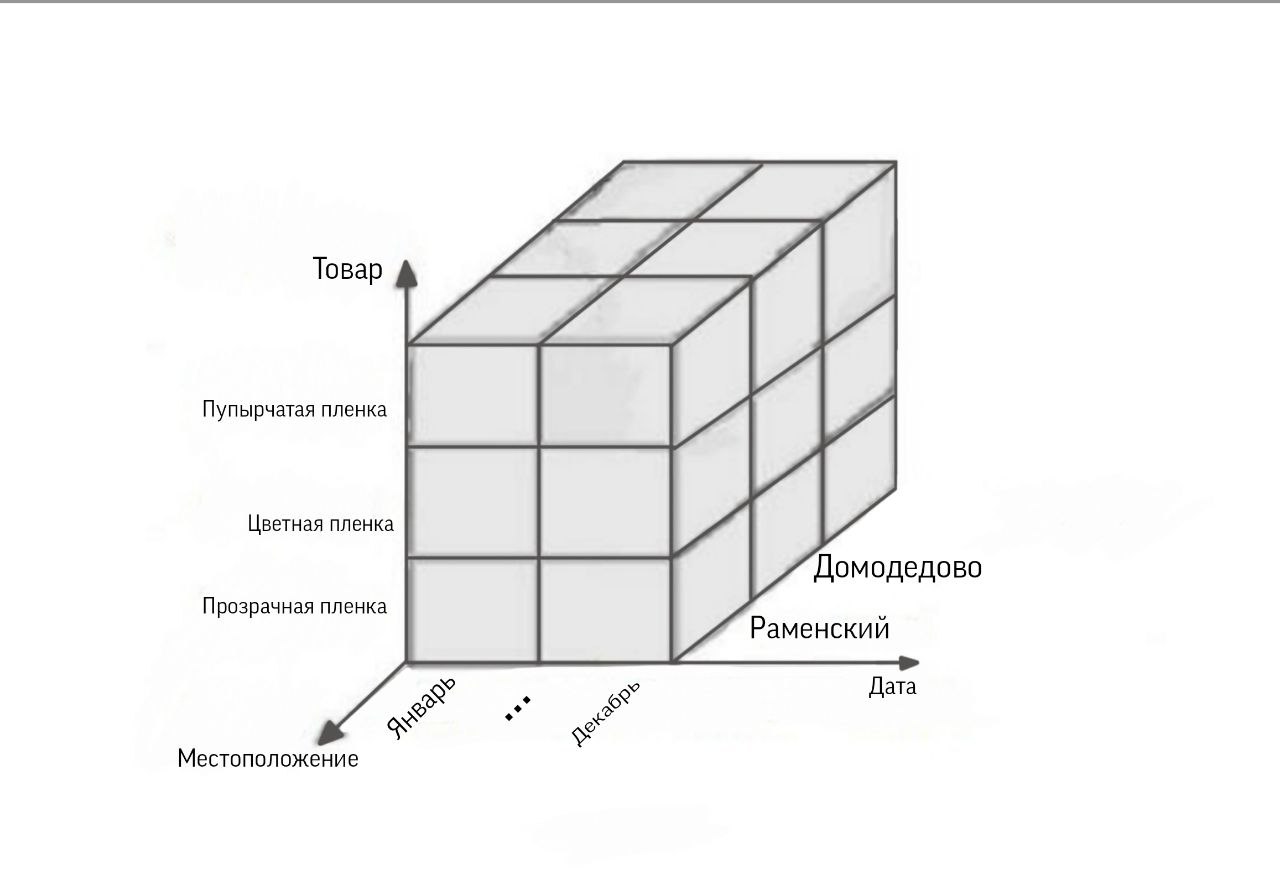


Рис. 2.15 Свертка по измерению Дата

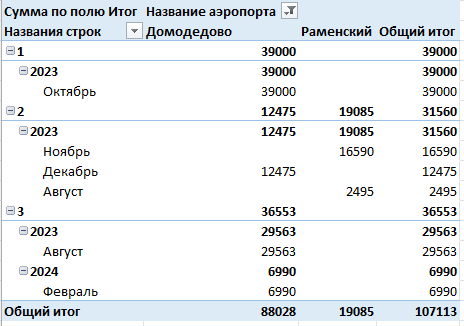


Рис. 2.16 Сводная таблица для свёртки

**OLAP-КУБ ПО ВЕРХНИМ УРОВНЯМ ИЕРАРХИИ ИЗМЕРЕНИЙ**

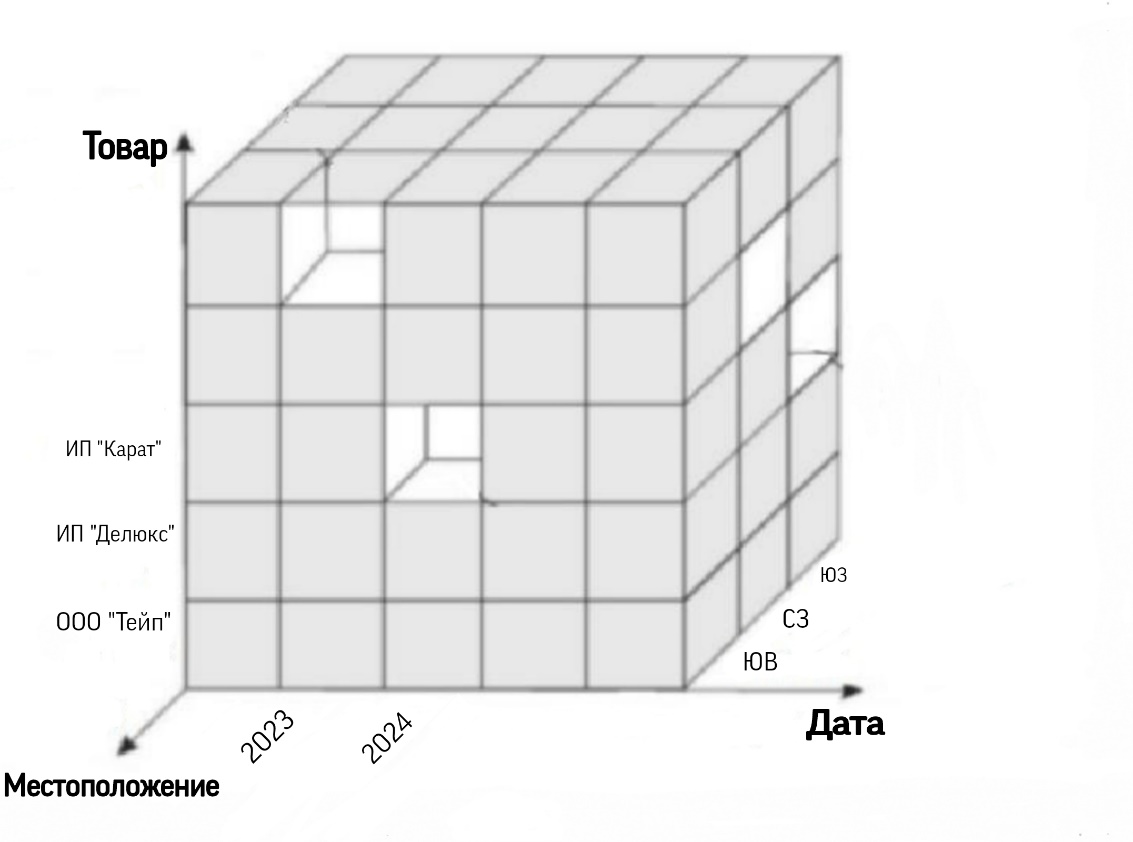


Рис. 2.17 Визуализация OLAP-куба

В результате построения OLAP-куба по верхним уровням иерархий измерения получаем сводную таблицу с информацией о прибыли от продаж пленок всех поставщиков, во всех зонах, в 2023/2024 году.

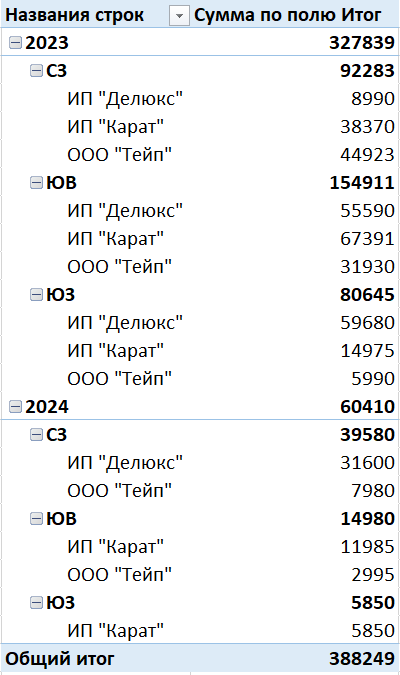


Рис. 2.18 Сводная таблица для OLAP-куба

**ДЕТАЛИЗАЦИЯ ПО ОДНОМУ ИЗМЕРЕНИЮ**

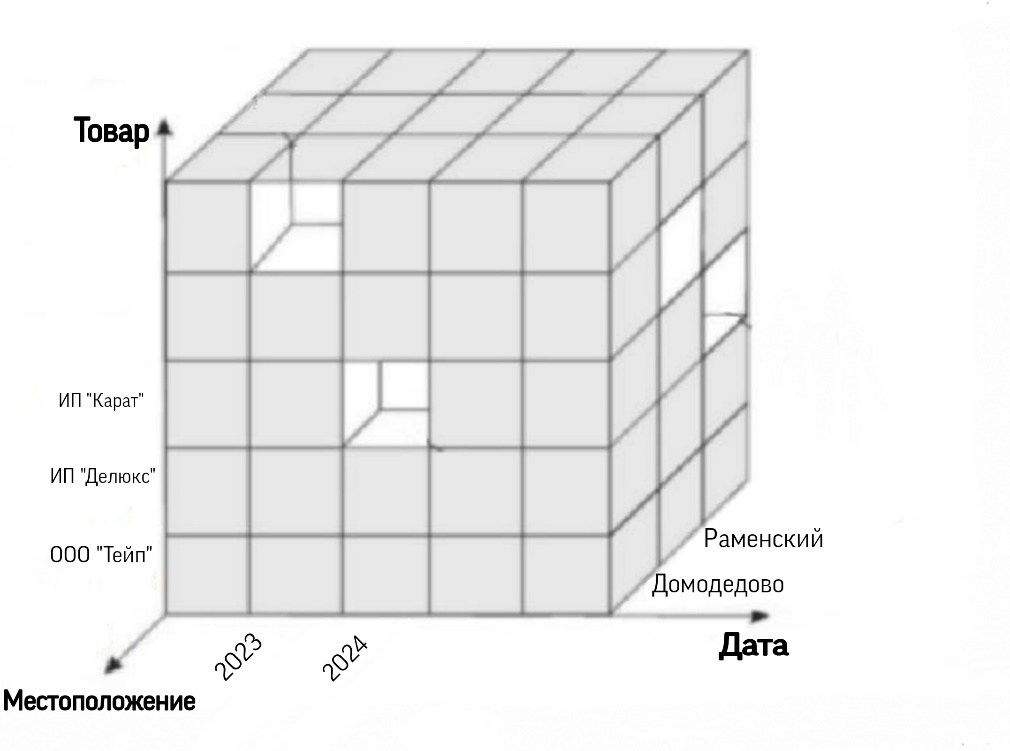


Рис. 2.19 Визуализация детализации

В результате детализации по местоположению получаем сводную таблицу с информацией о прибыли от продаж пленок всех поставщиков в аэропортах, в 2023/2024 году.

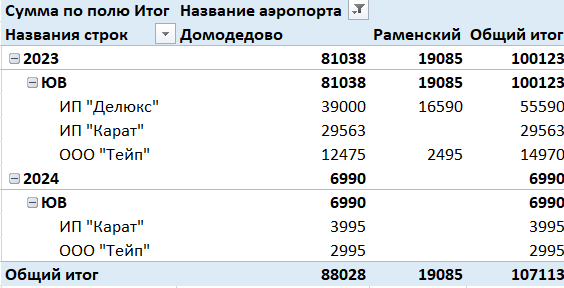


Рис. 2.20 Сводная таблица для детализации

# **ЛАБОРАТОРНАЯ РАБОТА №3**

**ЦЕЛЬ РАБОТЫ**

Изучение регрессионного анализа, как метода ИАД.

**ЗАДАНИЕ 1**

В качестве переменных для проверки гипотезы будут выступать количество товара (в шт.) и количество брака.

**ЗАДАНИЕ 2**

Гипотеза: чем больше количество товара, тем больше количество брака.

**ЗАДАНИЕ 3**

В качестве зависимой переменной Y будет выступать количество брака, а в качестве независимой переменной X – количество товара.

**ЗАДАНИЕ 4**

Для данных, включенных в регрессионную модель была создана отдельная таблица.

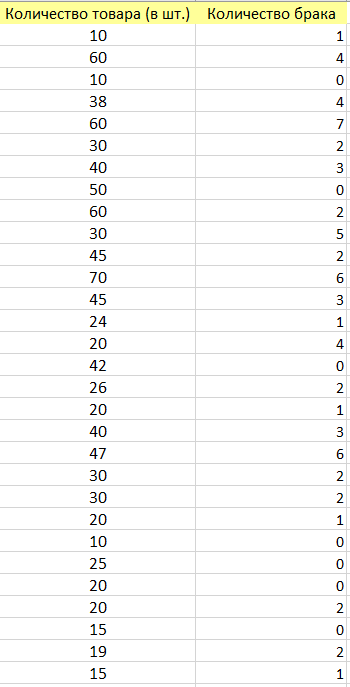


Рис. 3.1 Данные для регрессионной модели

**ЗАДАНИЕ 5**

Для найденной пары переменных была построена диаграмма рассеивания.

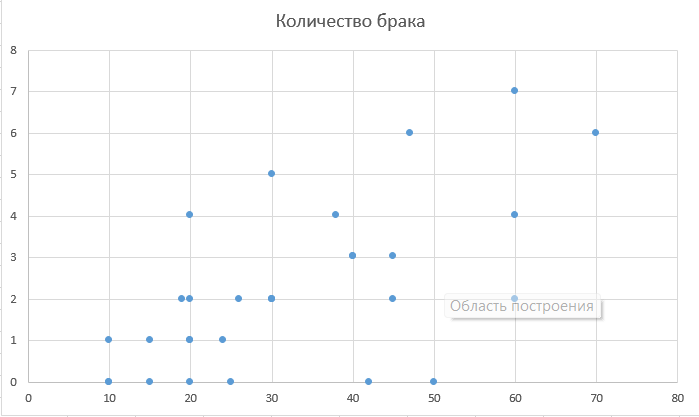


Рис. 3.2 Диаграмма рассеивания

**ЗАДАНИЕ 6**

Гипотеза о форме связи: форма связи – линейная, положительная.

**ЗАДАНИЕ 7**

Гипотеза о форме связи подтверждена.

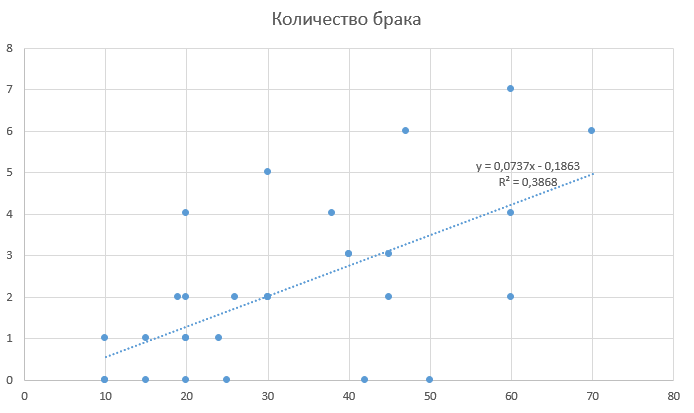


Рис. 3.3 Диаграмма с линией тренда

**ЗАДАНИЕ 8**

Выберем входные интервалы для X –столбец «Количество товара (в шт.)», а для Y – столбец «Количество брака».

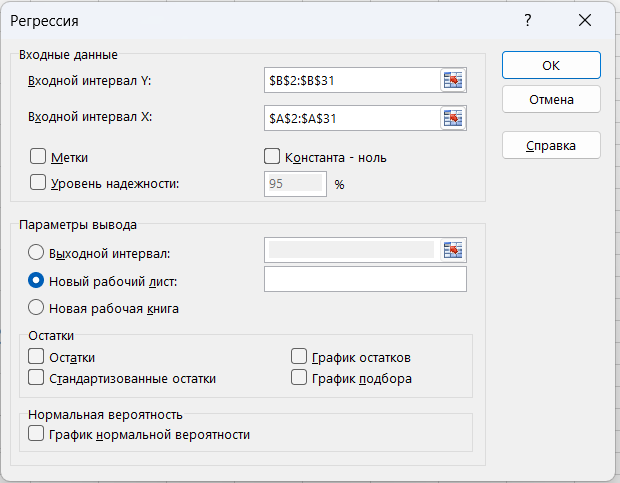


Рис. 3.4 Входные интервалы X и Y

**ЗАДАНИЕ 9**

Значение R2=0,3868 приемлемое и совпадает со значением из пункта 7. Значение переменной X1=0,0737 также совпадает, как и нулевая точка a = -0,1863.

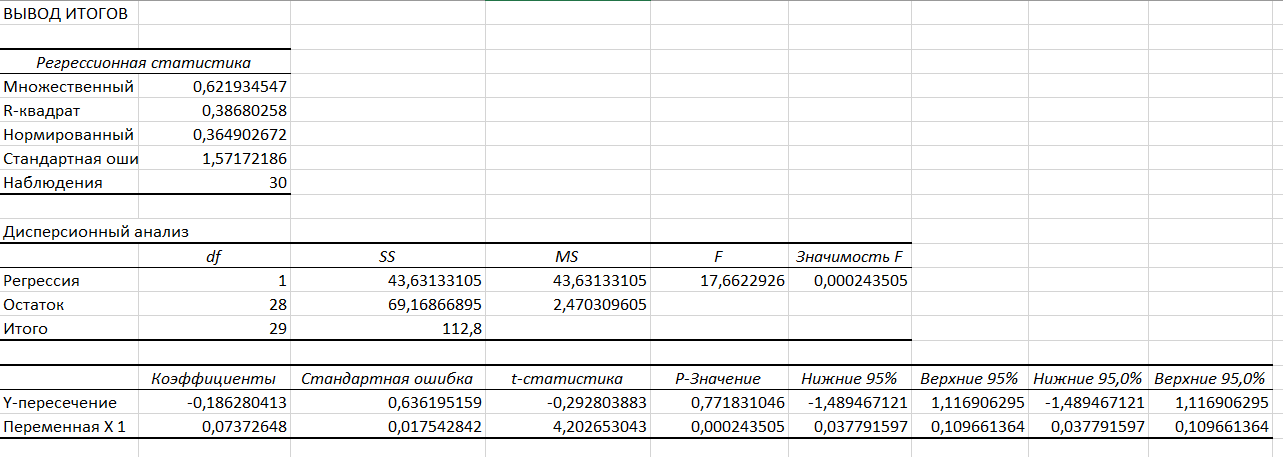


Рис. 3.5 Полученные данные

**ЗАДАНИЕ 10**

Коэффициенты уравнения регрессии, полученные с помощью функций НАКЛОН и ОТРЕЗОК:

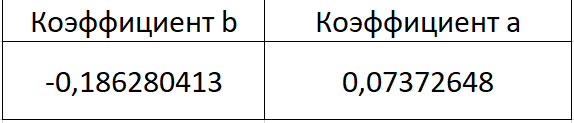


Рис. 3.6 Коэффициенты регрессии

Полученные величины совпадают с полученными в графике.

**ЗАДАНИЕ 11**

Коэффициент регрессии (a = 0,0737) – коэффициент, который характеризует степень влияния независимой переменной на зависимую. То есть количество товара влияет на количество брака. Коэффициент положительный, следовательно, влияние будет положительным (чем больше доставлено товара, тем больше количество брака, и наоборот).

Параметр сдвига – коэффициент, который показывает, какой будет зависимая переменная, если все используемые в данной модели факторы будут равны 0. Для данного графика значение графика b = -0,1863. Это означает, что с ростом количества товара, количество брака может уменьшаться.

**ЗАДАНИЕ 12**

Анализ достаточно точен и подтверждает изначальную гипотезу.

**ЗАДАНИЕ 13**

При полученном уравнении y = 0,0737x – 0,1863 можно предсказать количество брака для 80 и 100 товаров:

x = 80: y = 5 бракованных товаров;

x = 100: y = 7 бракованных товаров.

# **ЛАБОРАТОРНАЯ РАБОТА №4**

**ЗАДАНИЕ 1**

Вставим срез по измерению «Уровень защиты» по значению «1» и сделаем выборку «Название аэропорта» по значениям «Домодедово» и «Жуковский».

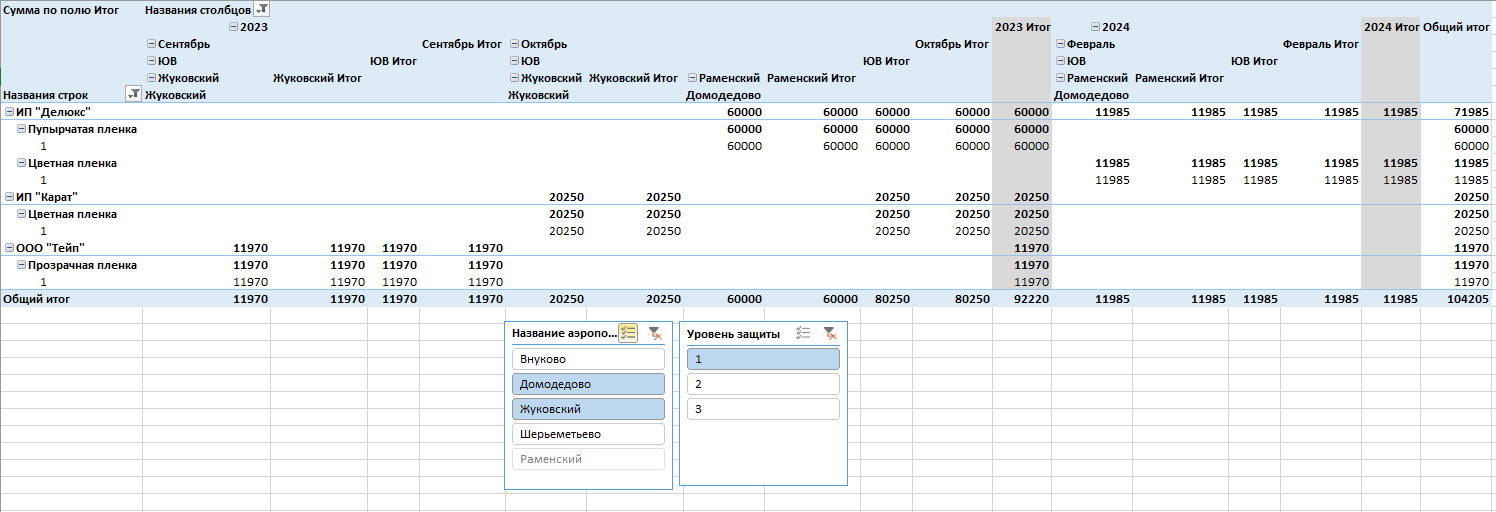


Рис. 4.1 Настроенный срез

**ЗАДАНИЕ 2**

Cпарклайны – это мини-диаграммы, построенные внутри ячейки Excel. Они позволяют вписать диаграмму в ячейку и тем самым сэкономить рабочее пространство. Ниже представлена визуализация спарклайнов:

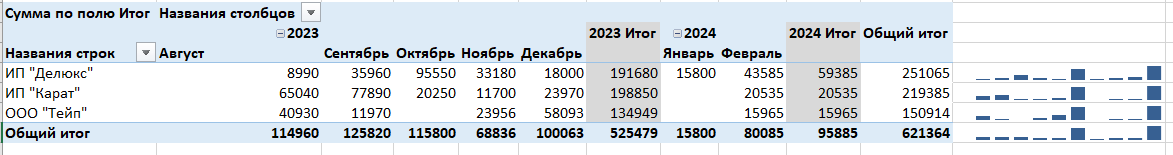


Рис. 4.2 Спарклайны

**ЗАДАНИЕ 3**

Добавим к столбцу «Цена (за шт.)» значки-индикаторы и отформатируем его, чтобы значения ниже среднего отображались на красном фоне.

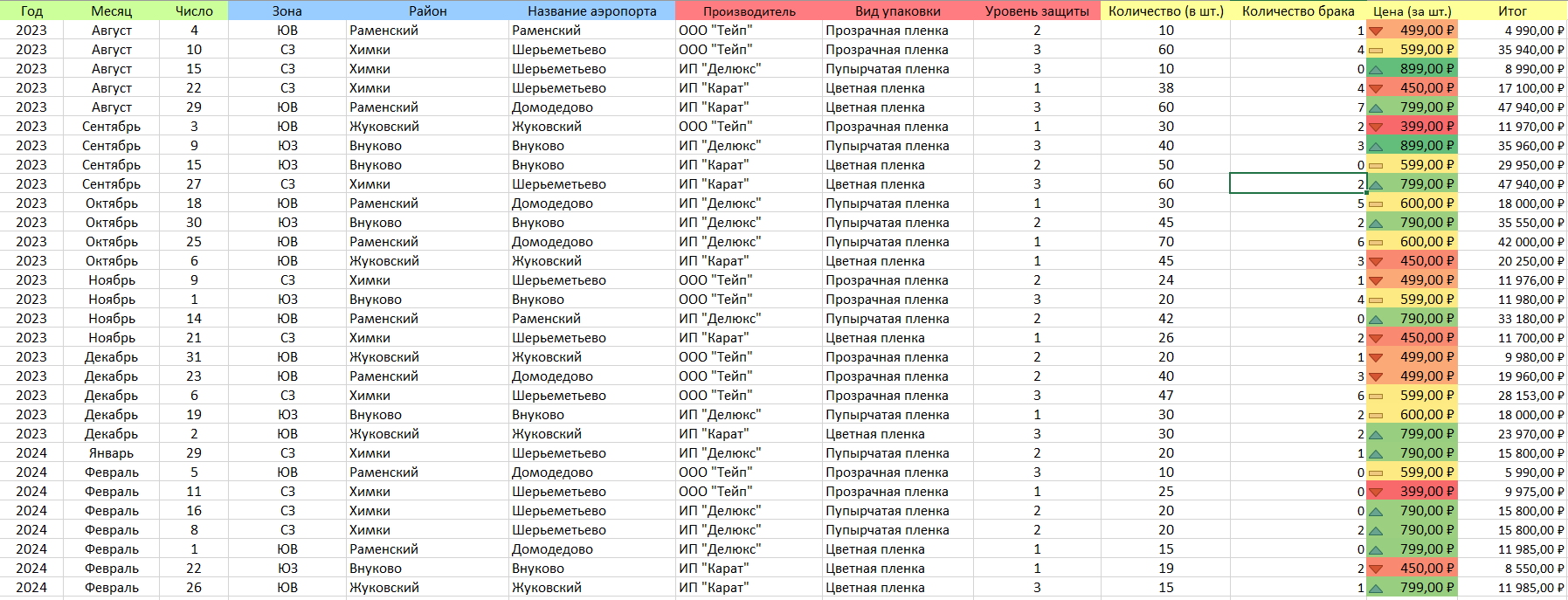


Рис. 4.3 Столбец «Цена (за шт.) с форматированием

**ЗАДАНИЕ 4-5**

Создадим правила форматирования для столбцов «Производитель» и «Зона».

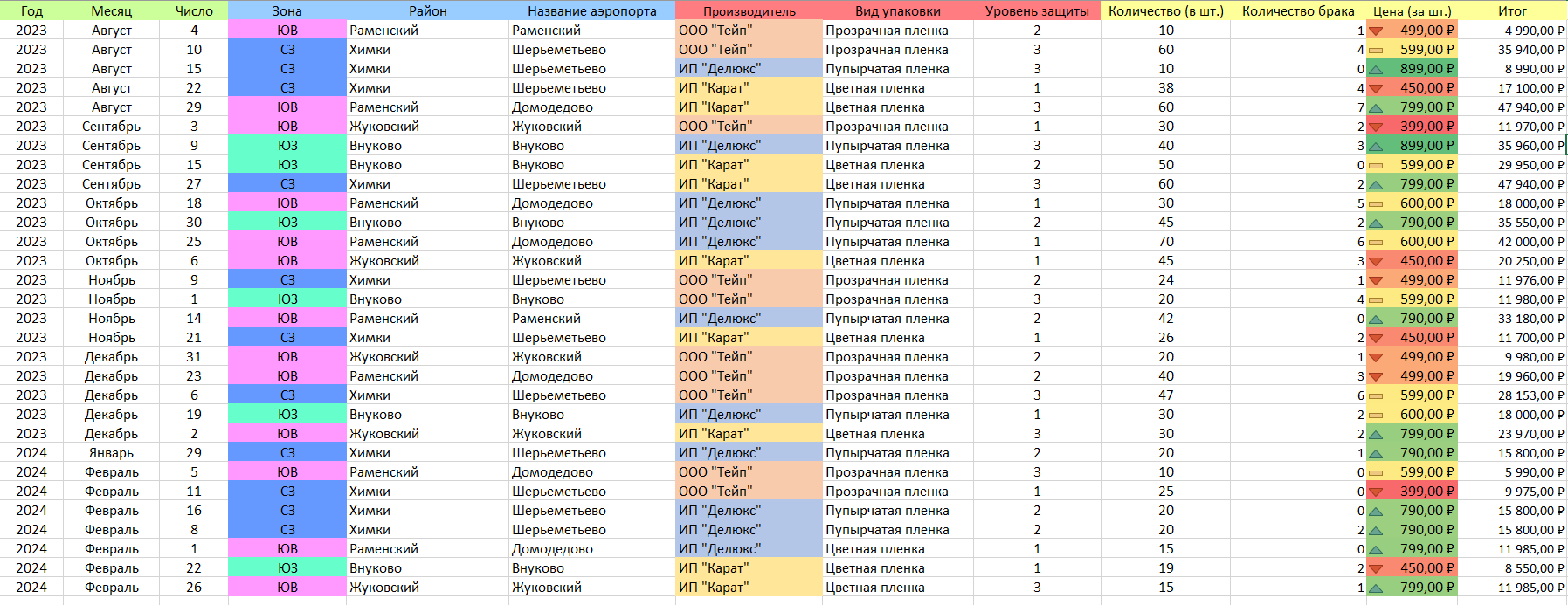


Рис. 4.4 Созданные правила