# Рабочая тетрадь №4. Визуализация данных в DataLens

## Введение

Практическая работа заключается в знакомстве с бесплатной платформой Yandex DataLens <https://datalens.yandex.ru> для визуализации данных в формате построения динамических отчетов. В данной работе студенту предлагается ознакомиться с базовыми возможностями платформы для визуализации данных на основе прохождения пошаговой демонстрации с построением визуализаций. Также для студентов представлено небольшое самостоятельное задание в конце настоящего материала.

Yandex DataLens – это инструмент, разработанный компанией Яндекс, который предоставляет возможность анализировать и визуализировать большие объемы данных.

DataLens позволяет пользователям создавать интерактивные дашборды и отчеты, которые помогают визуализировать и понять сложные данные. Он предоставляет различные инструменты для фильтрации, сортировки и анализа данных, а также функции прогнозирования и моделирования.

DataLens может использоваться в различных областях, включая бизнес-аналитику, маркетинг, финансы, науку и другие. Он помогает пользователям принимать более информированные решения на основе данных и выявлять скрытые закономерности и тренды.

## Практическое задание №2

Для выполнения демонстрационного задания студенту необходимо загрузить таблицы данных из облачного хранилища Yandex <https://disk.yandex.ru/d/2kvbzX8LOVCMOQ>, а также войти под своей учетной записью Yandex, или временно зарегистрировать её.

Зайдите в Yandex DataLens <https://datalens.yandex.ru/>. Выберите пункт «Создать подключение».

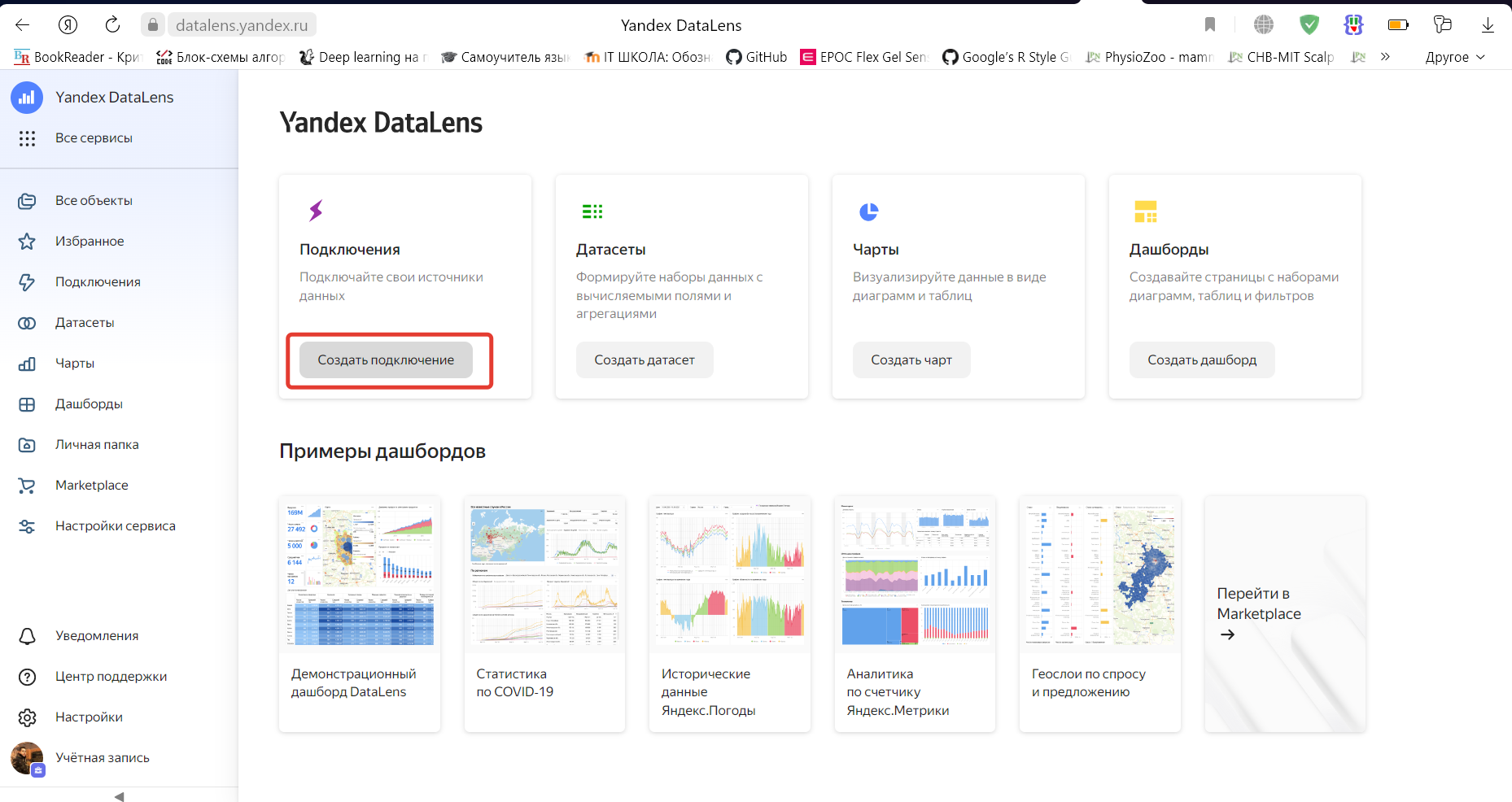


Рисунок . Начальная страница Yandex DataLens

Подключение дает возможность пользователю определять доступ к данным, на основе которых будет строиться визуализация. Выбираем «Файлы».

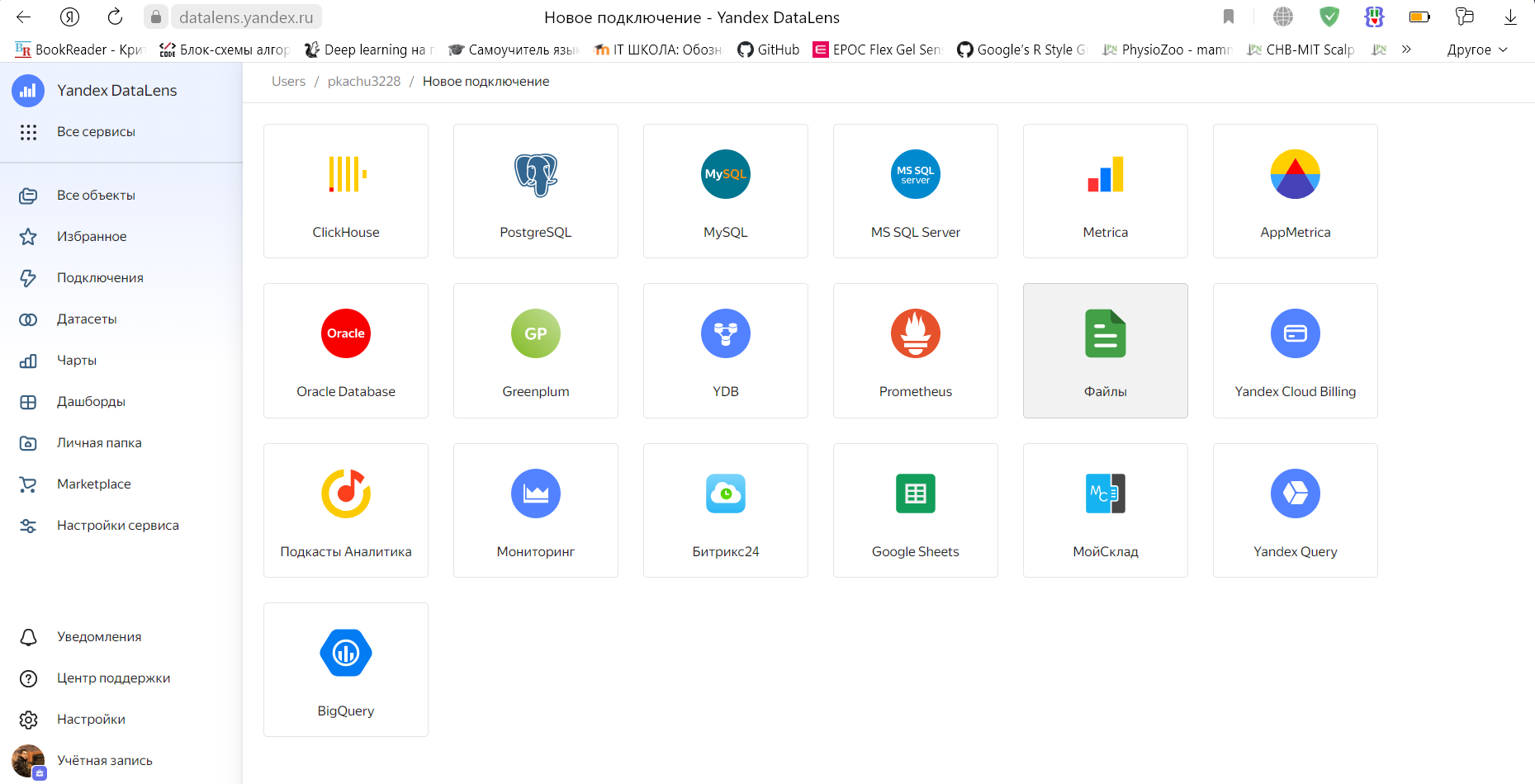


Рисунок . Подключения Yandex DataLens

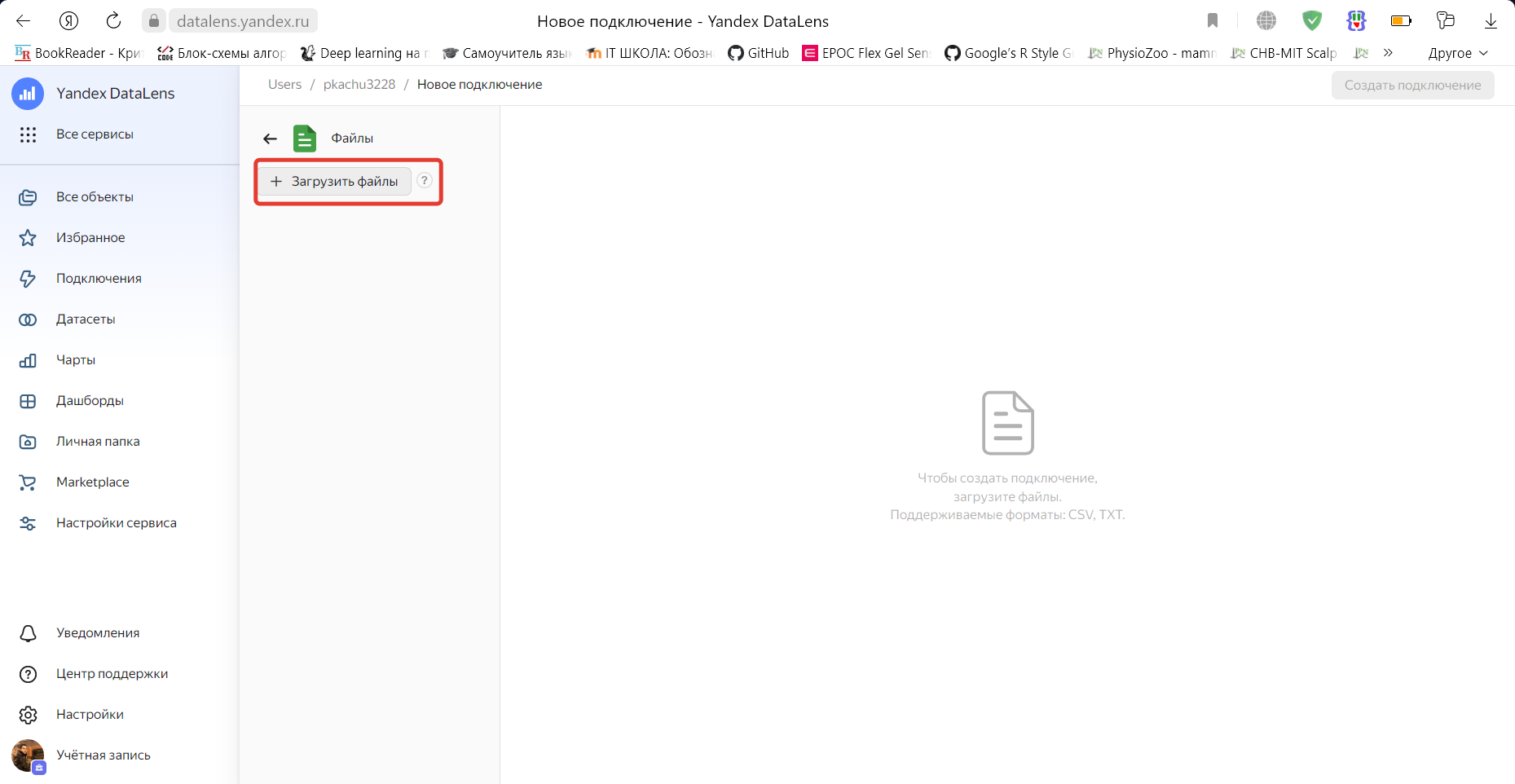


Рисунок . Загрузка данных в подключение файла

В загрузке файла выберем три «.csv» файла из приложенной выше ссылки на Yandex Disk.

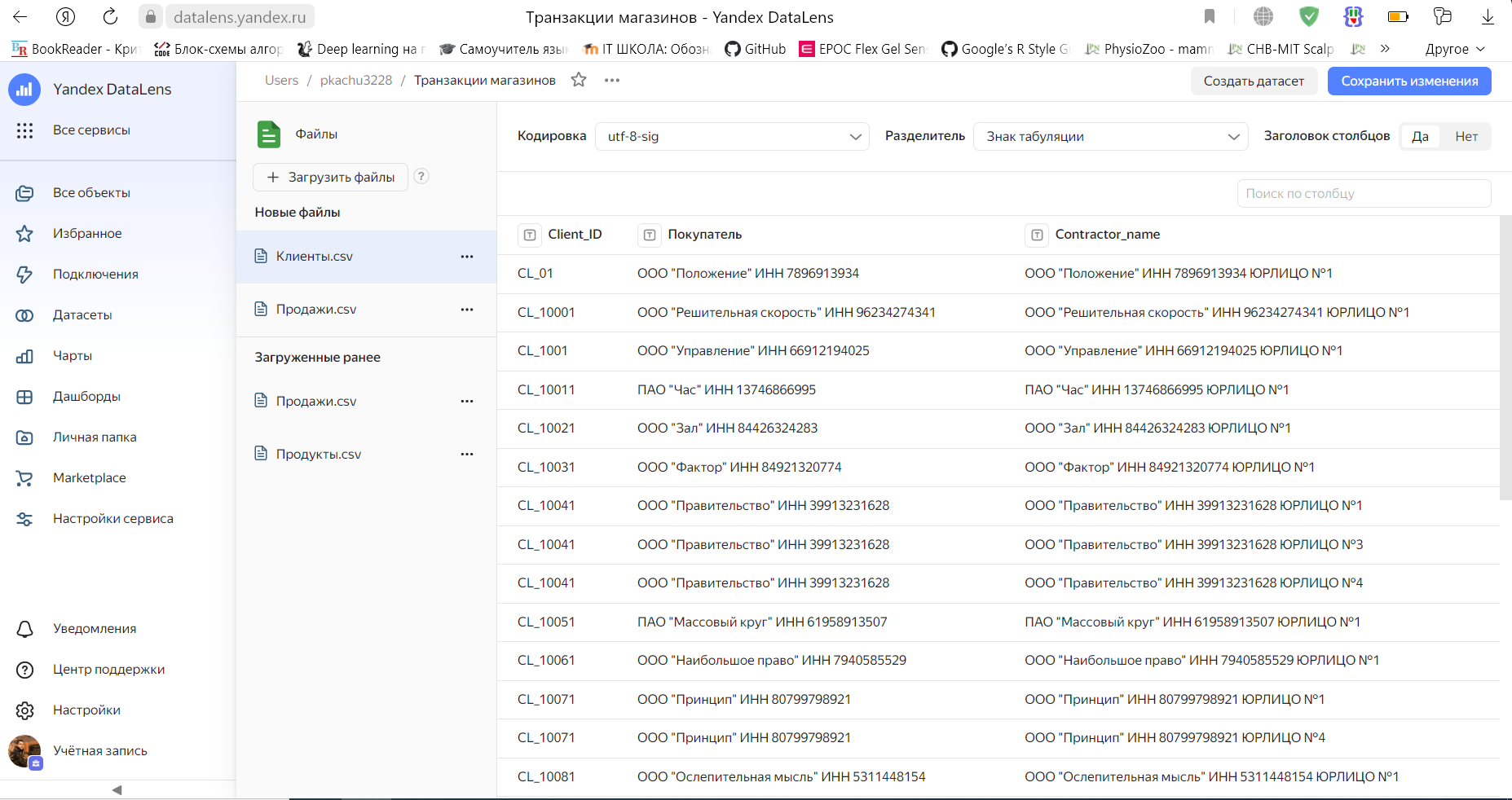


Рисунок . Предпросмотр загружаемой выборки данных

После загрузки данных нам доступен предпросмотр загруженных таблиц.

Во всех таблицах, введенных в систему визуализации необходимо проверить поле «Заголовок столбцов». Если поля в двух таблицах будут иметь разные имена или несовпадающие ключи, то связи между двумя таблицами могут не установиться. Если в таблице не стоит «Да» в «Заголовок столбцов», то нужно поставить его.

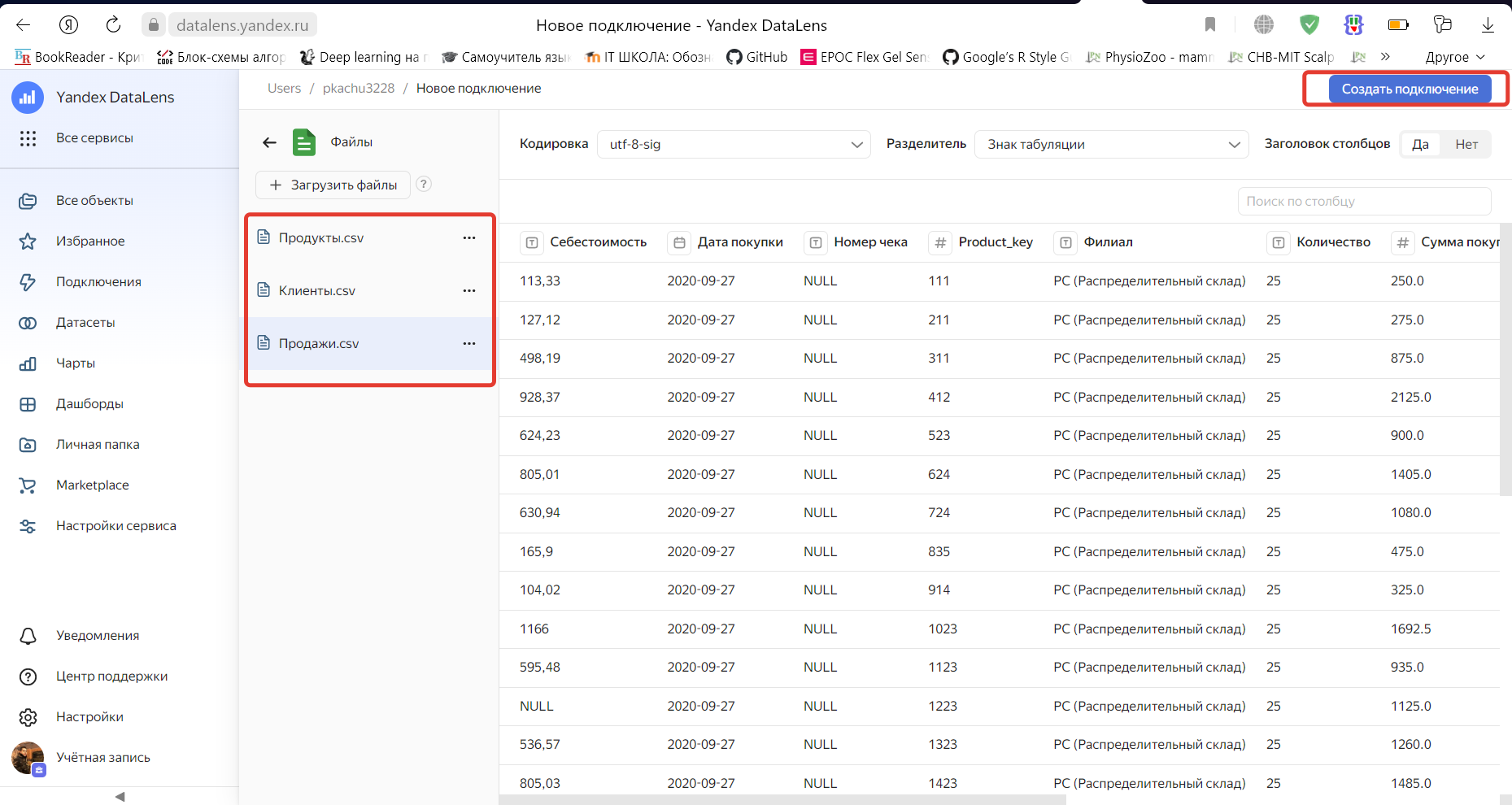


Рисунок . Создание подключения

Сохраним изменения и продолжим. Данному подключению дайте наименование «Транзакции магазинов».

Далее создадим датасет на базе которого мы будет строить чарт. Войдем на вкладку «Датасеты» и создадим датасет.

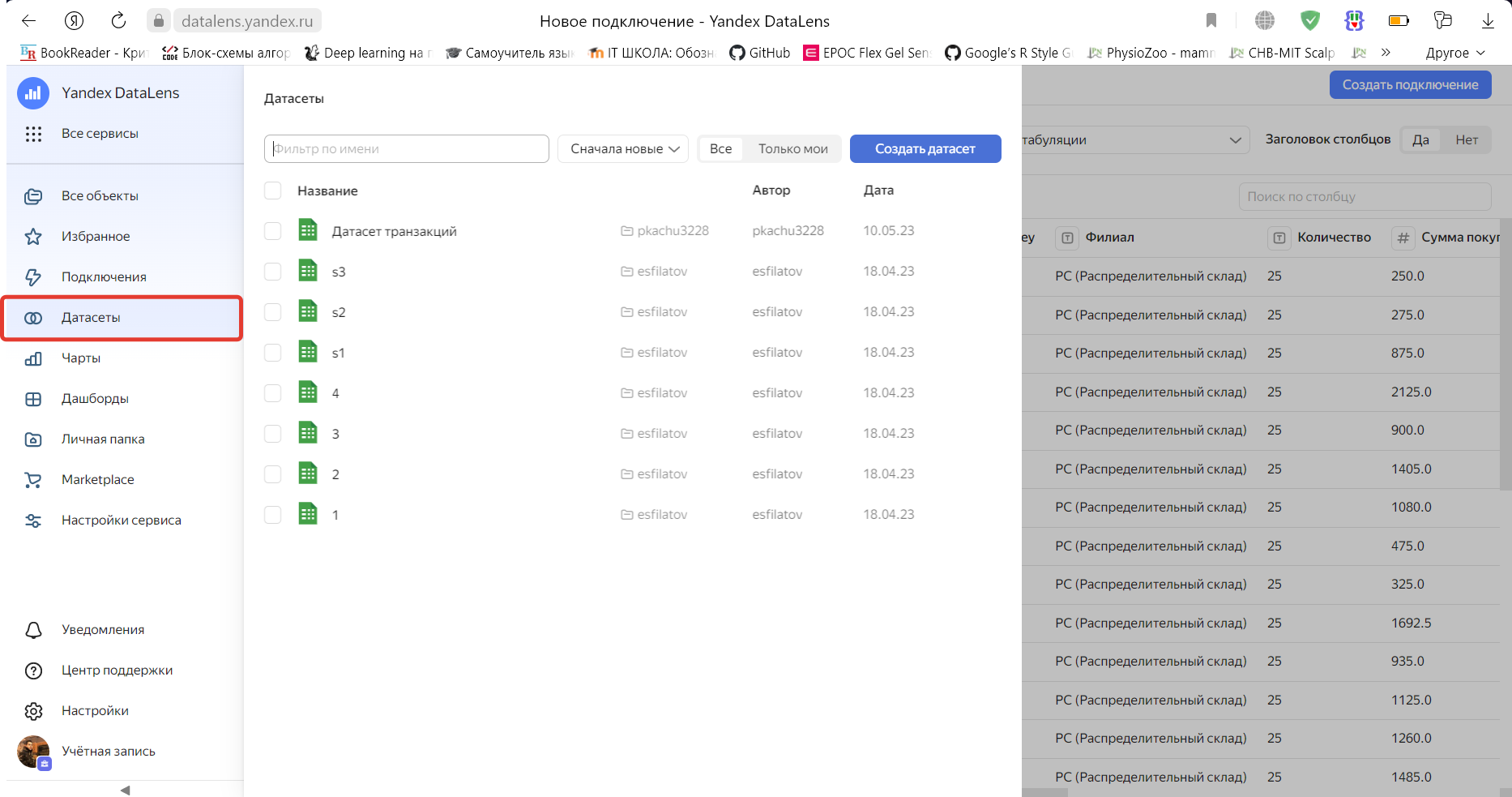


Рисунок . Создание датасета в DataLens

Из раздела существующих подключений выберем свое.

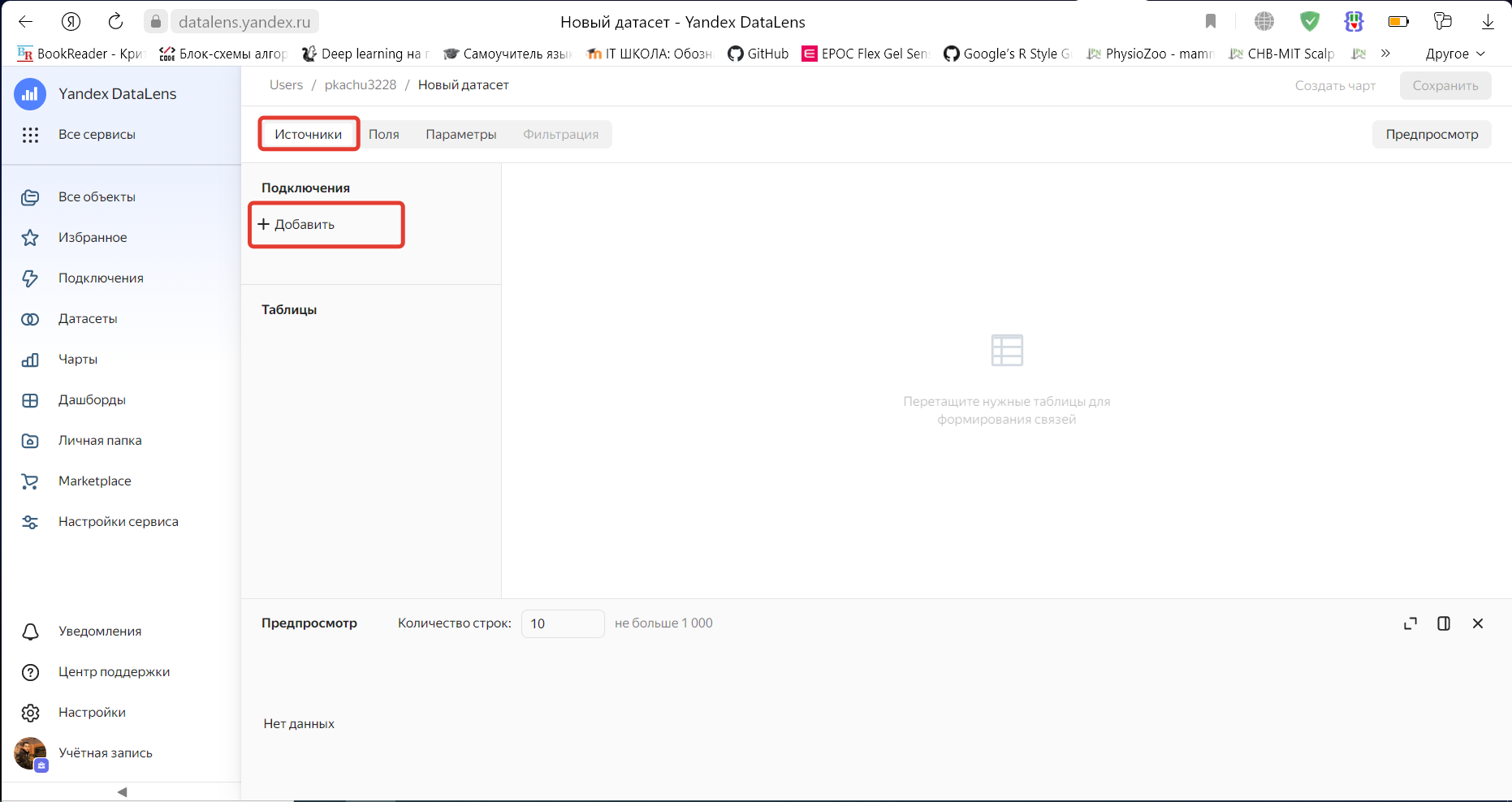


Рисунок . Определение данных в подключении

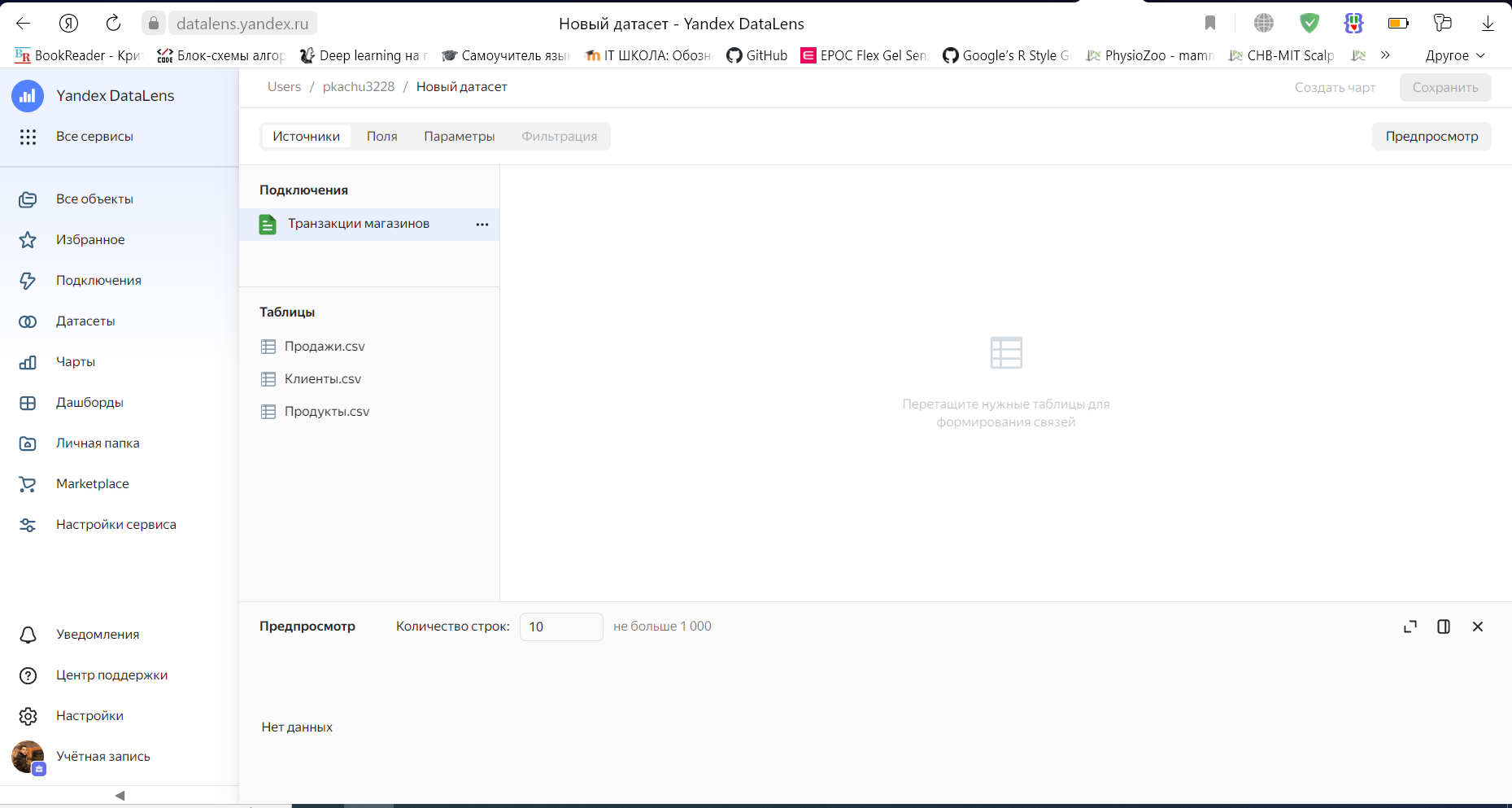


Рисунок . Определение датасета из подключения

Перетаскиванием потащите плашку с таблицей «Продажи.csv» на основное поле. Затем перетащите остальные плашки перенесением их прямиком на него. Установите тип объединения таблиц правым и левым, где в качестве главной таблицы выступает таблица «Продажи.csv».



Рисунок . Возможные проблемы при импорте таблиц из подключения в датасет

В случае, если при соединении таблиц возникают проблемы, связь между таблицами можно настроить вручную.

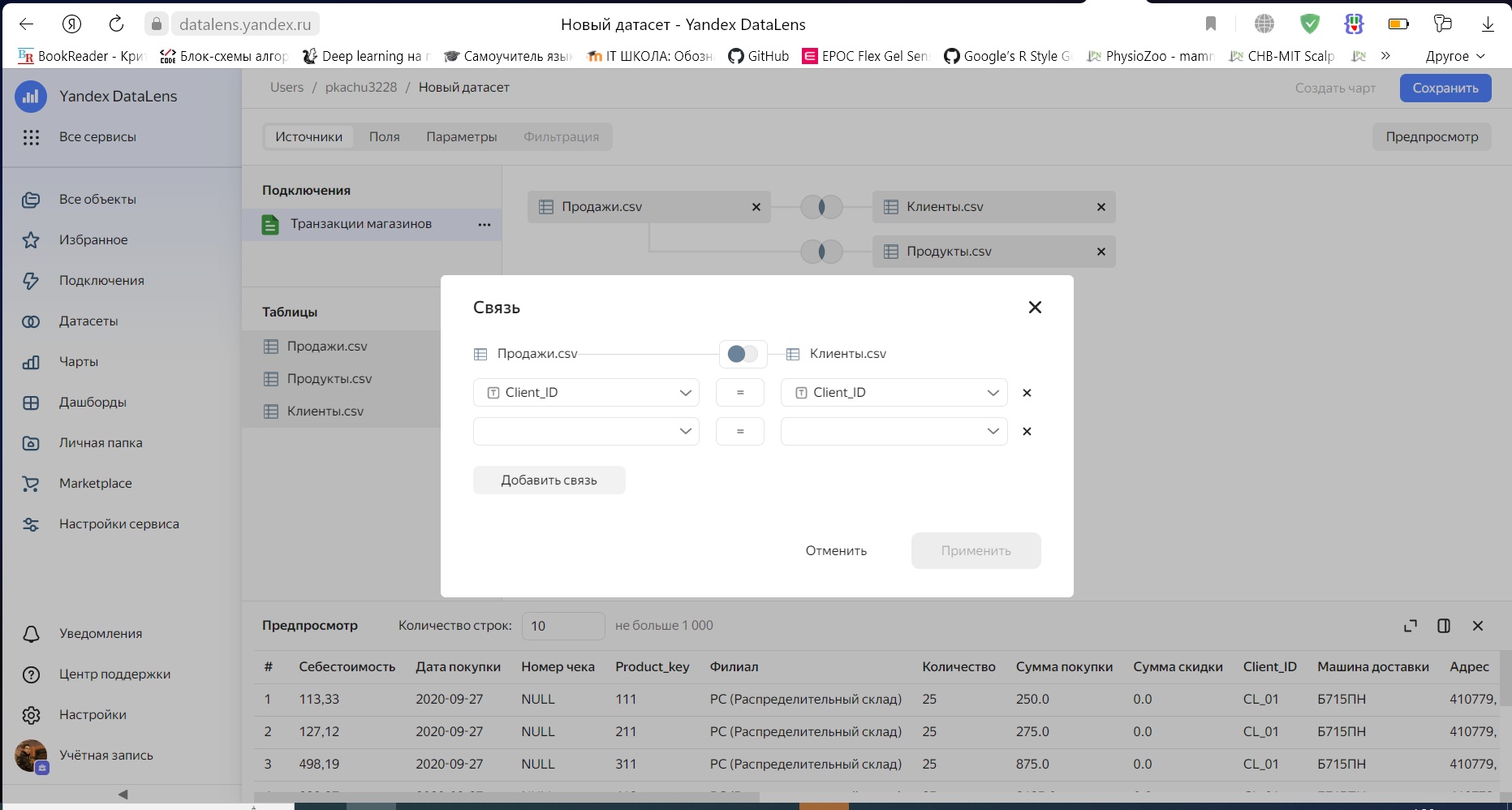


Рисунок . Меню настройки связей между таблицами

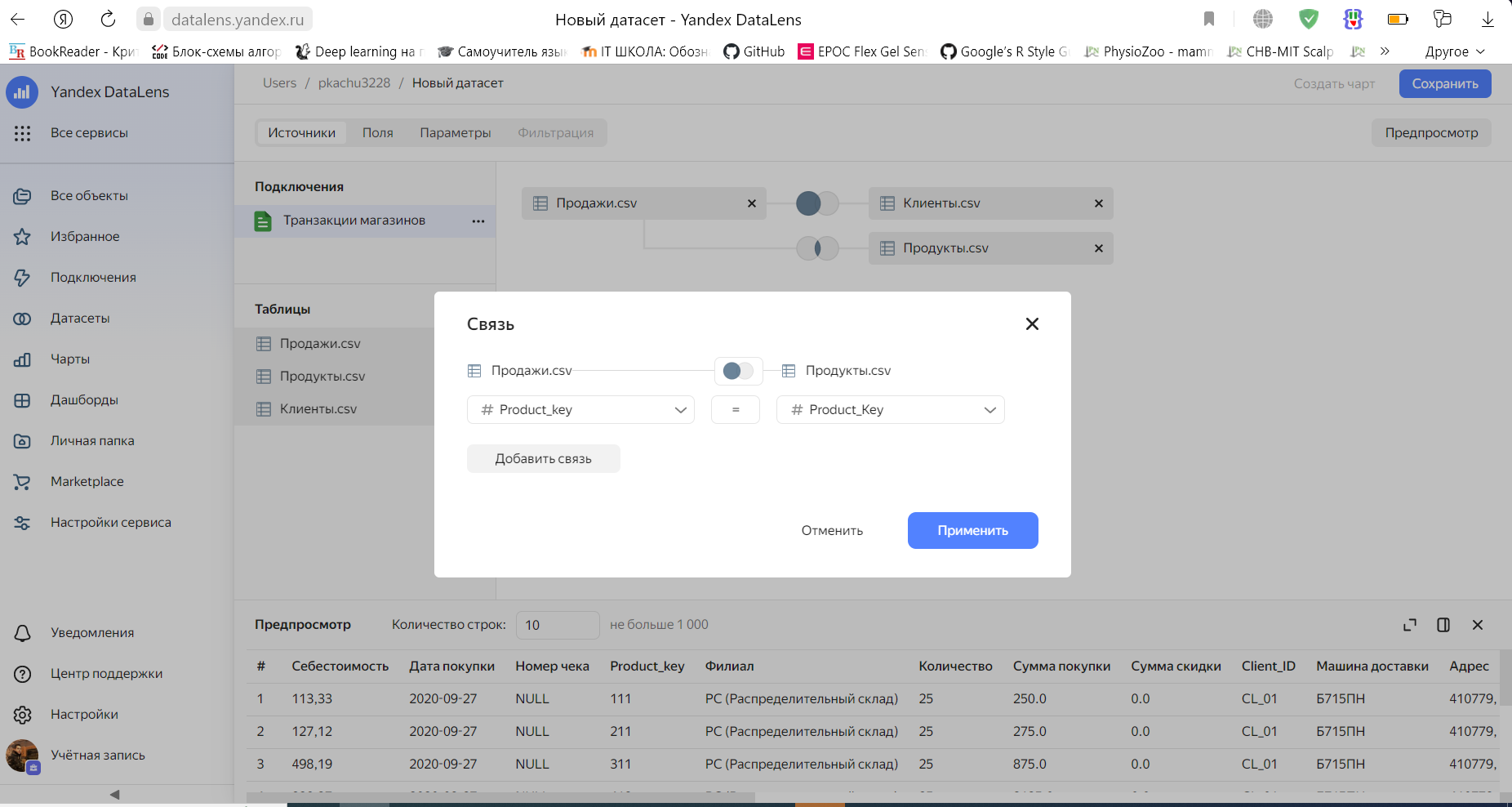


Рисунок . Настроенные связи между таблицами

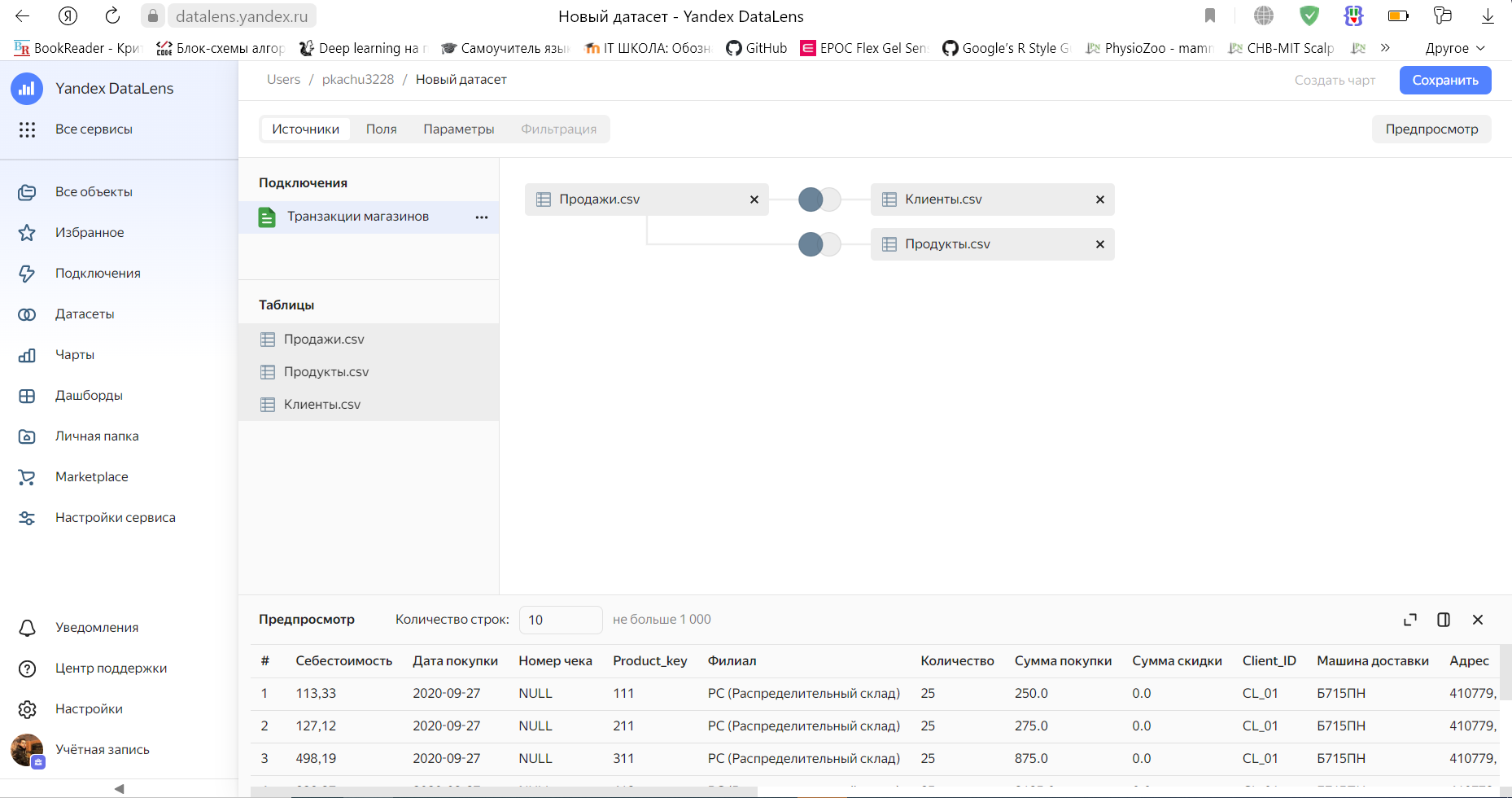


Рисунок . Настроенное левое соединение информационных таблиц в одну транзакционную таблицу

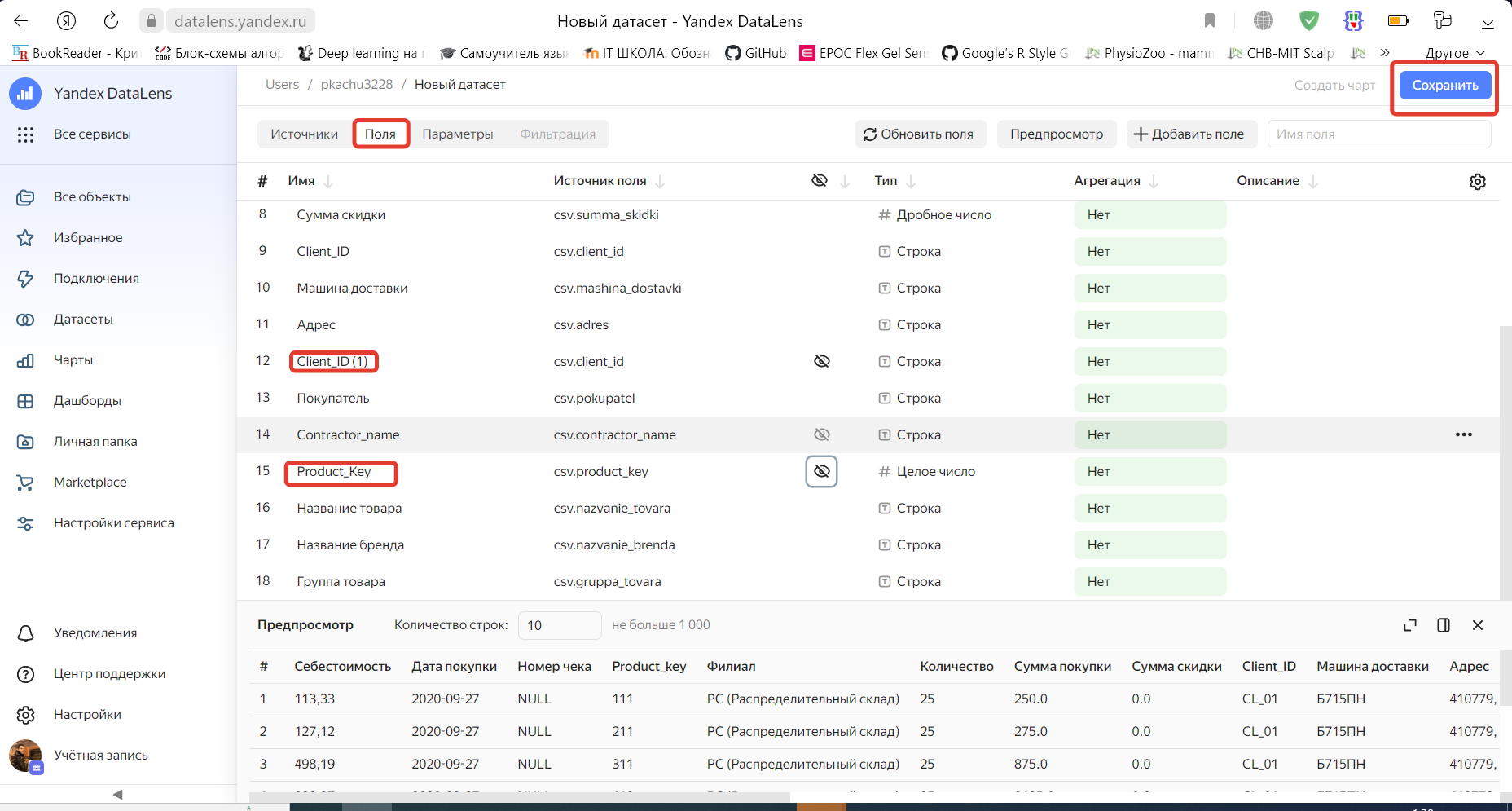


Рисунок . Изменение доступа к просмотру столбцов

Спрячьте поля Client\_ID и Product\_Key, поскольку это внутренние ключи для соединения полей, и они уже повторяются в таблице единожды.

Созданный датасет назовем «Датасет транзакций».

Создадим чарт на основе созданного датасета, в котором сделаем простейшую столбчатую визуализацию.

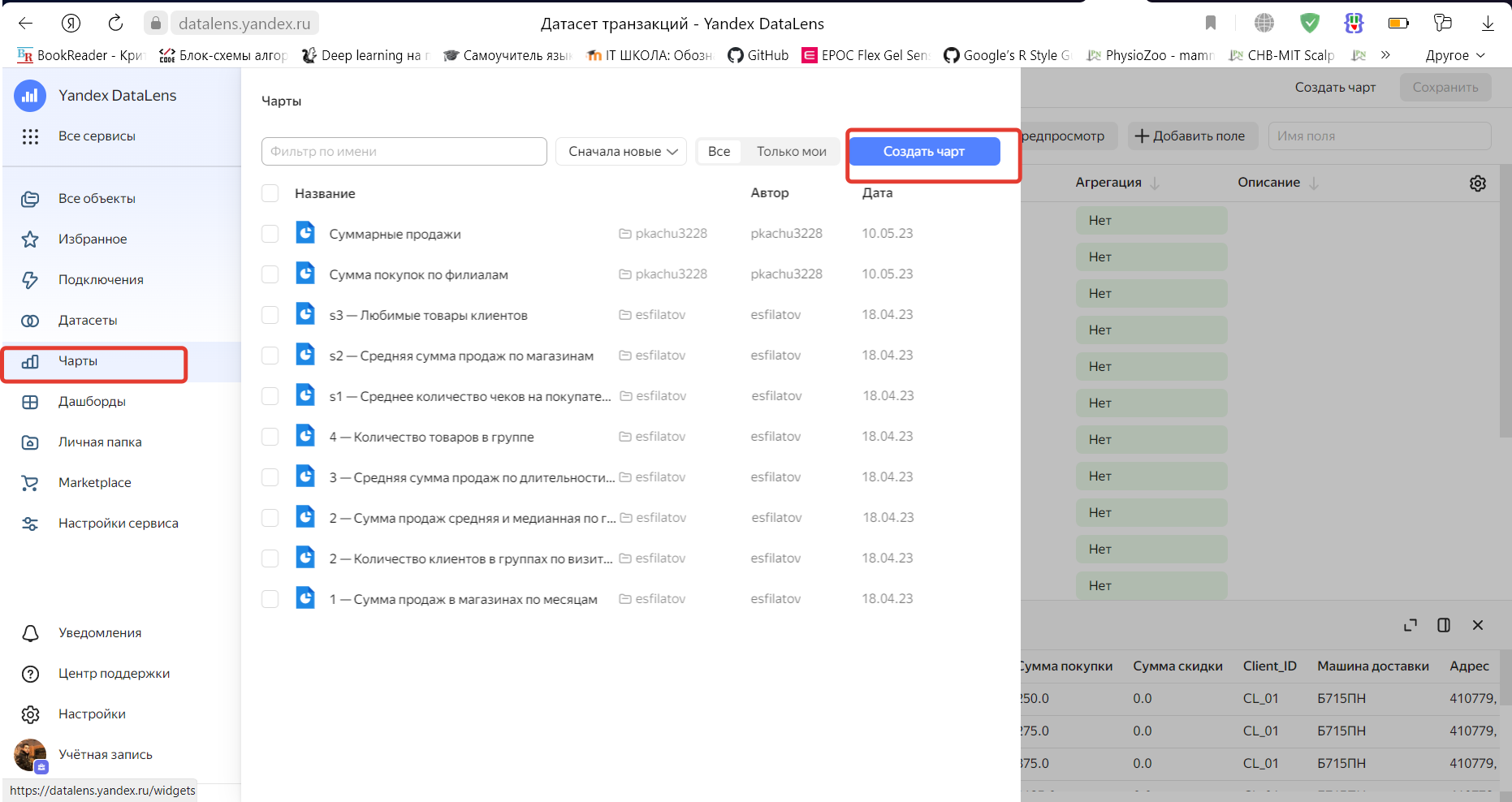


Рисунок . Возможность создания чарта из полученного датасета

Выберем созданный нами датасет транзакций.



Рисунок . Выбор датасета для использования в визуализации

Перетаскиванием нужных полей в определение визуализатора мы можем получить подобный столбчатый график для отображения выручки, полученной с поставки товаров.

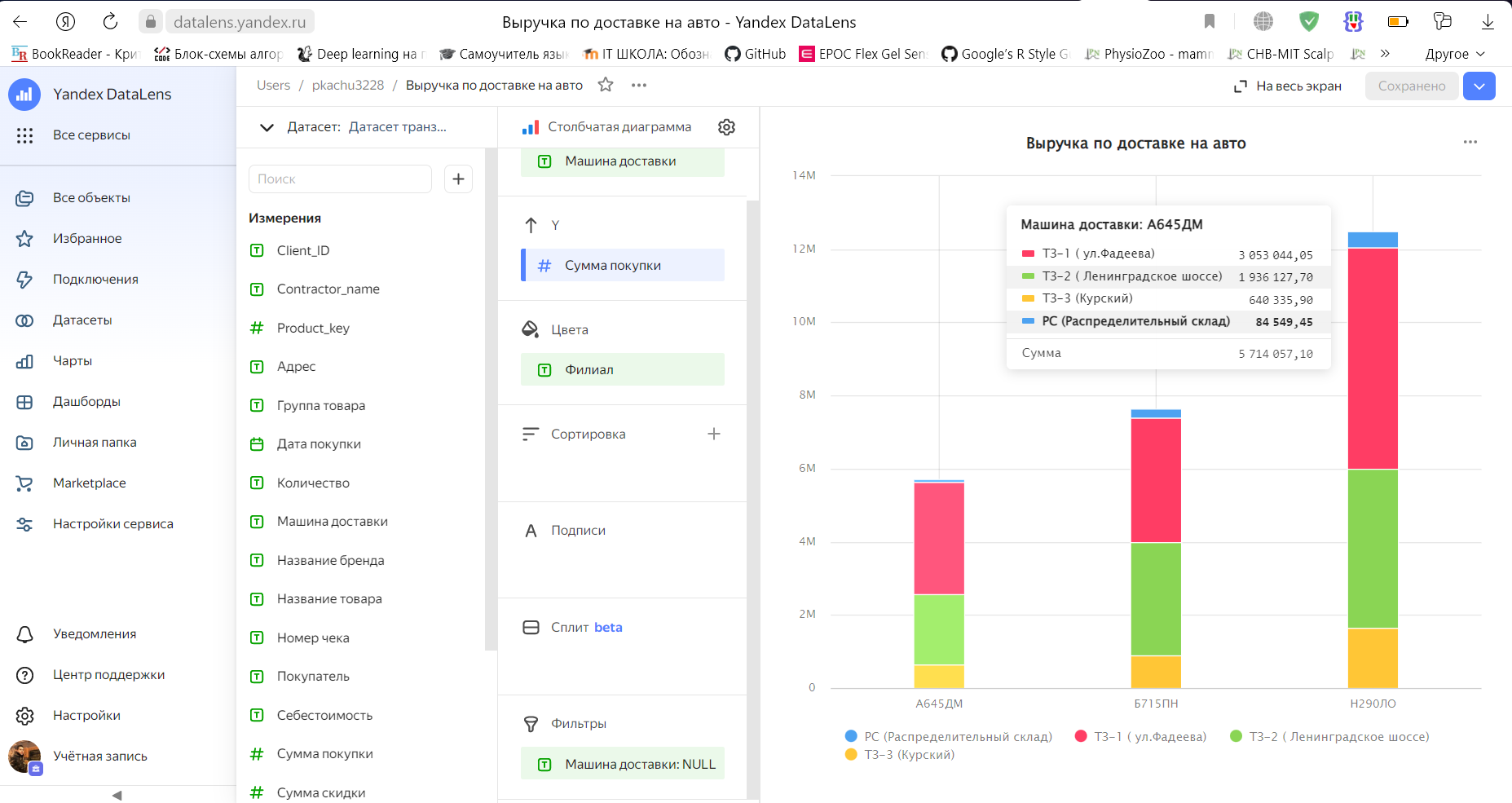


Рисунок . Построение столбчатой диаграммы

Далее нам необходимо создать дашборд на котором будет базироваться ваша текущая работа.

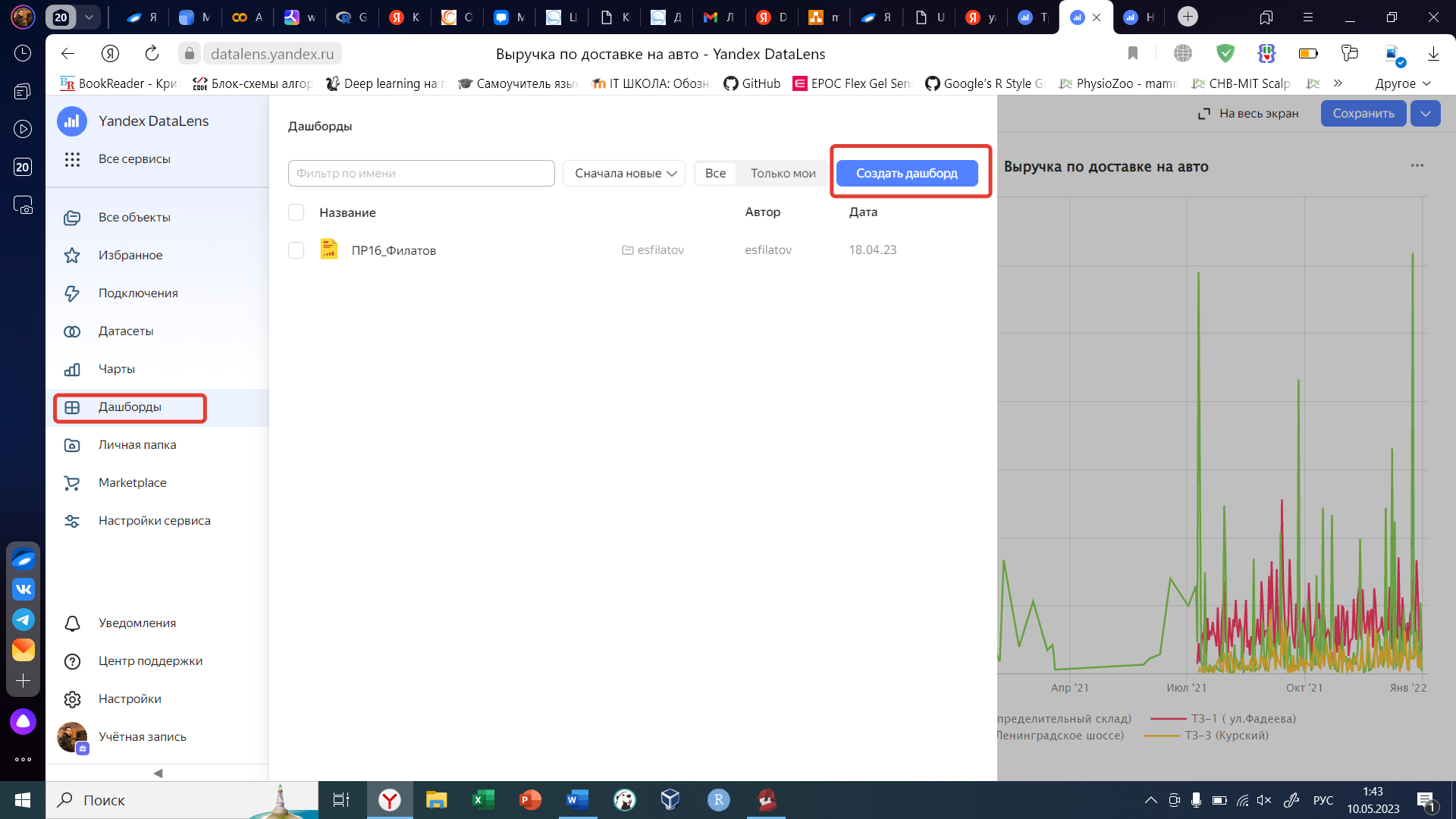


Рисунок . Создание дашборда в Yandex DataLens

Создайте дашборд «Дашборд для продаж товаров», или как для вас будет удобнее.

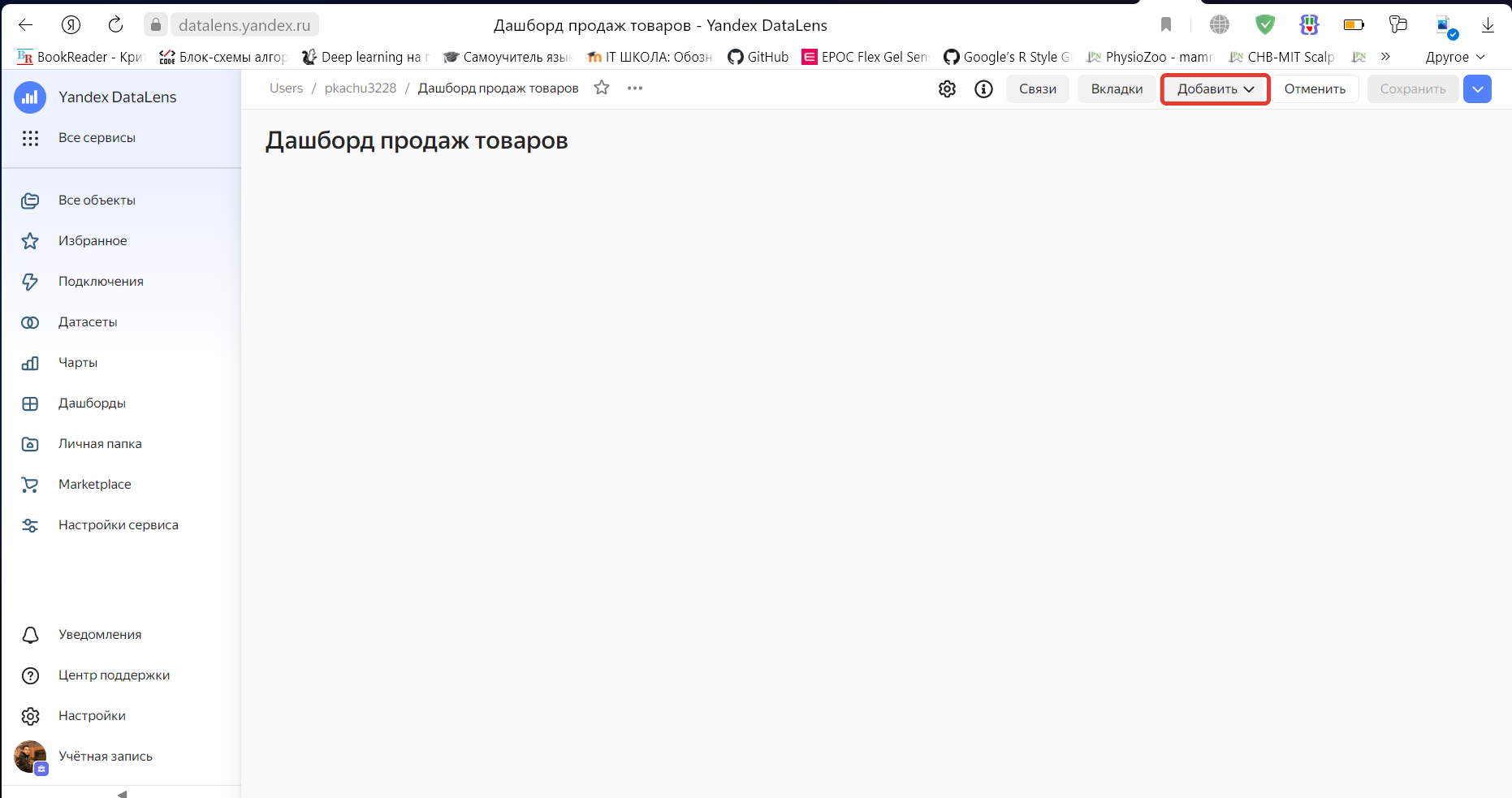


Рисунок . Добавление элементов в отчет

В дашборд прикрепите созданный чарт.

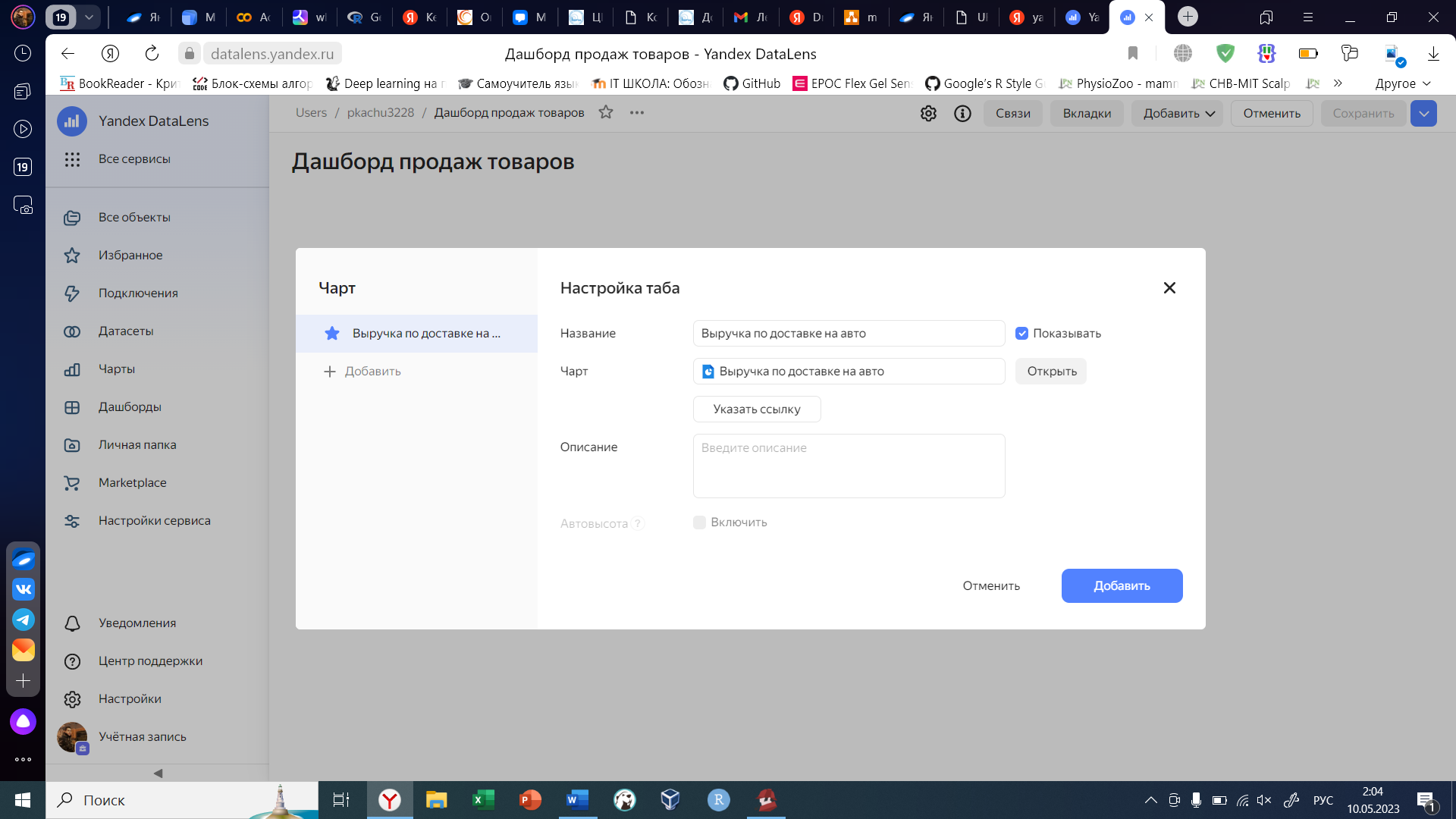


Рисунок . Добавление чарта в дашборд

Селекторы – фильтры для страниц дашборда, позволяющие в интерактивном формате изменять значение выбранного поля датасета. Дает возможность изменять перспективу отчета в реальном времени по требованию пользователя дашборда. Использование и объявление селектора будет показано далее.

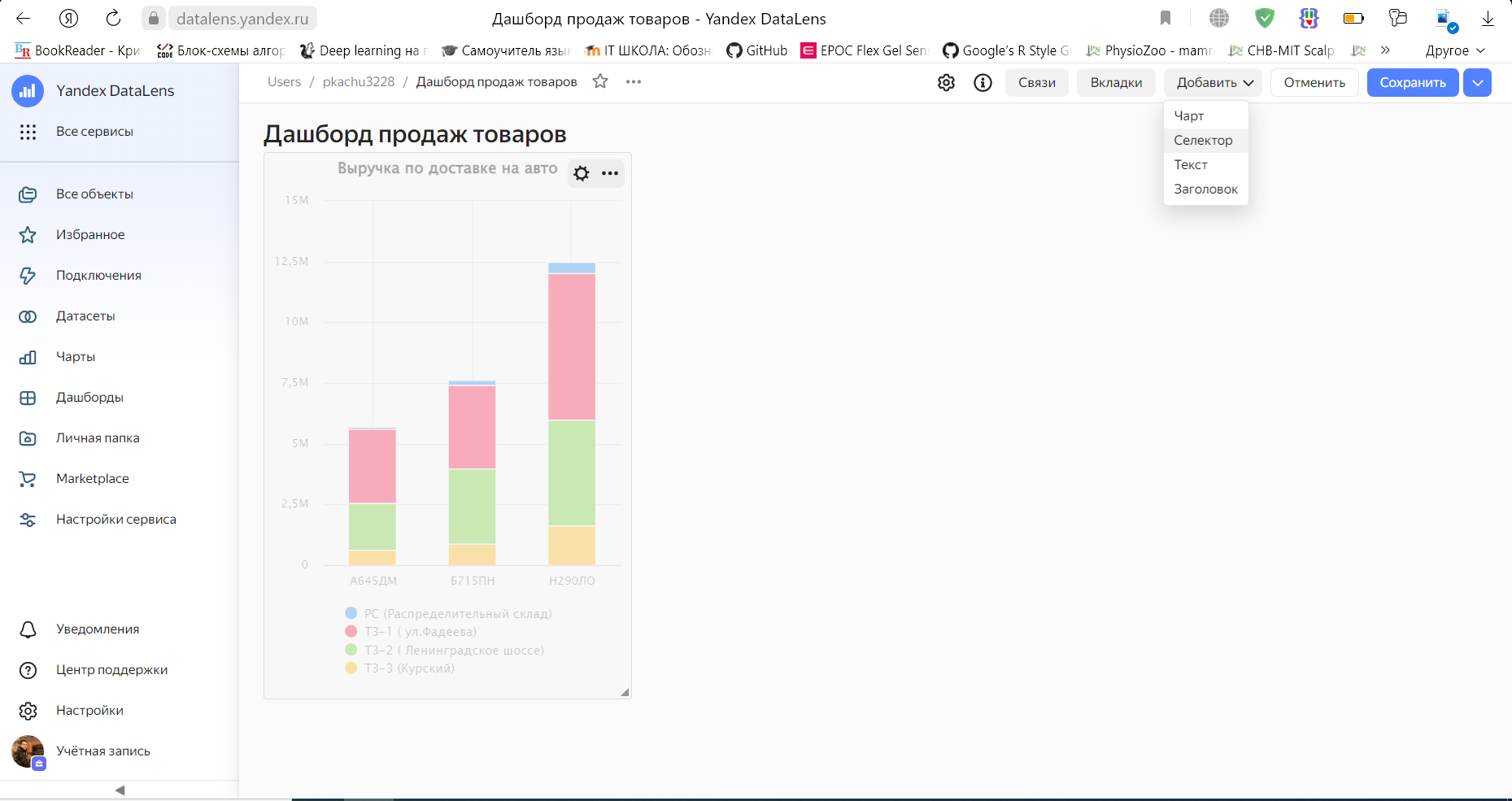


Рисунок . Редактирование дашборда и выбор селектора

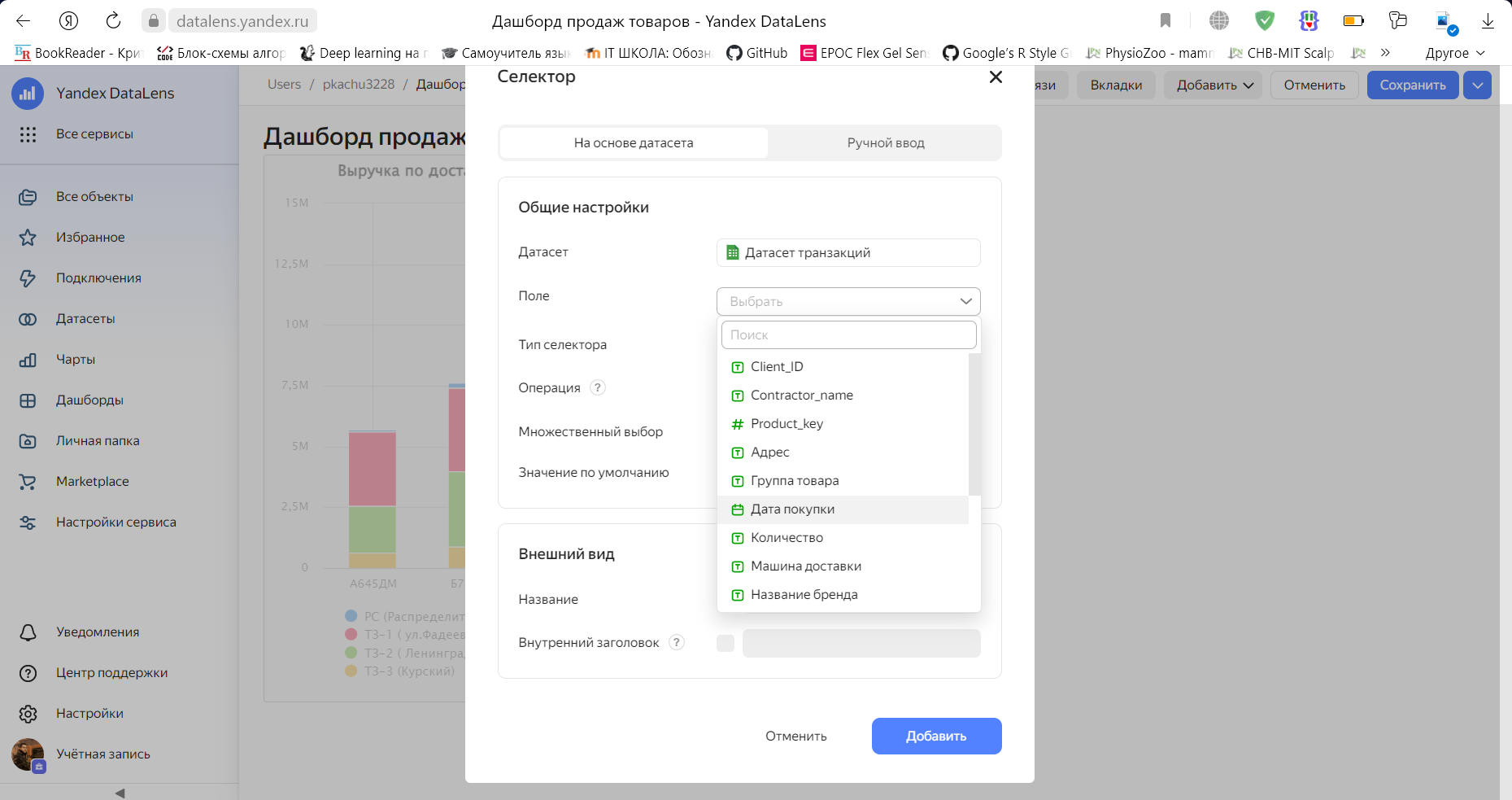


Рисунок . Выбор поля для фильтрации с помощью селектора

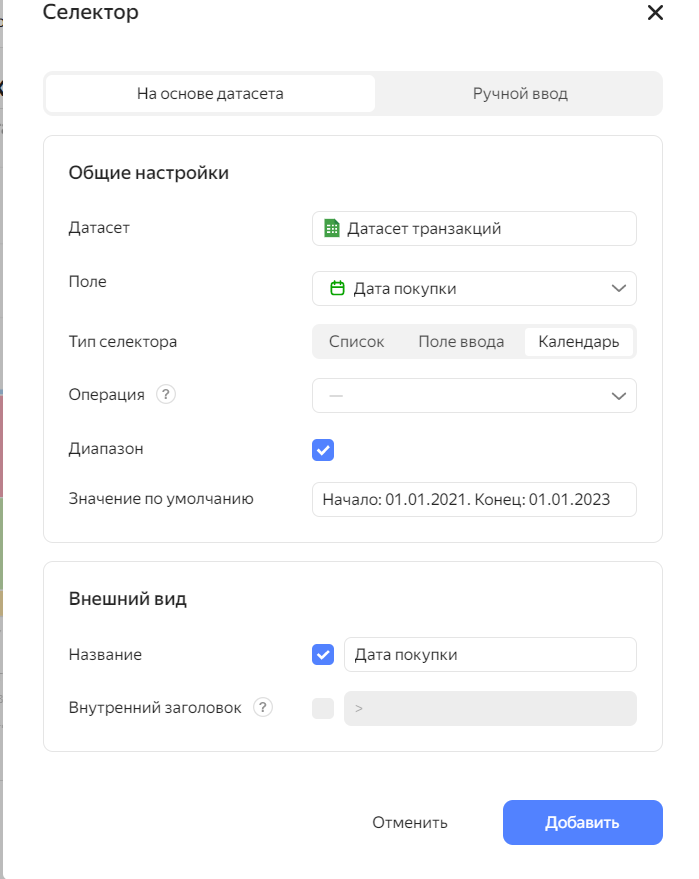


Рисунок . Определение диапазона значений дат для преобразования дашборда по дате заказов

С помощью селектора в дашбордах мы можем исполнять сложные динамические фильтрации и изменять графики по нашему усмотрению без настройки графиков.

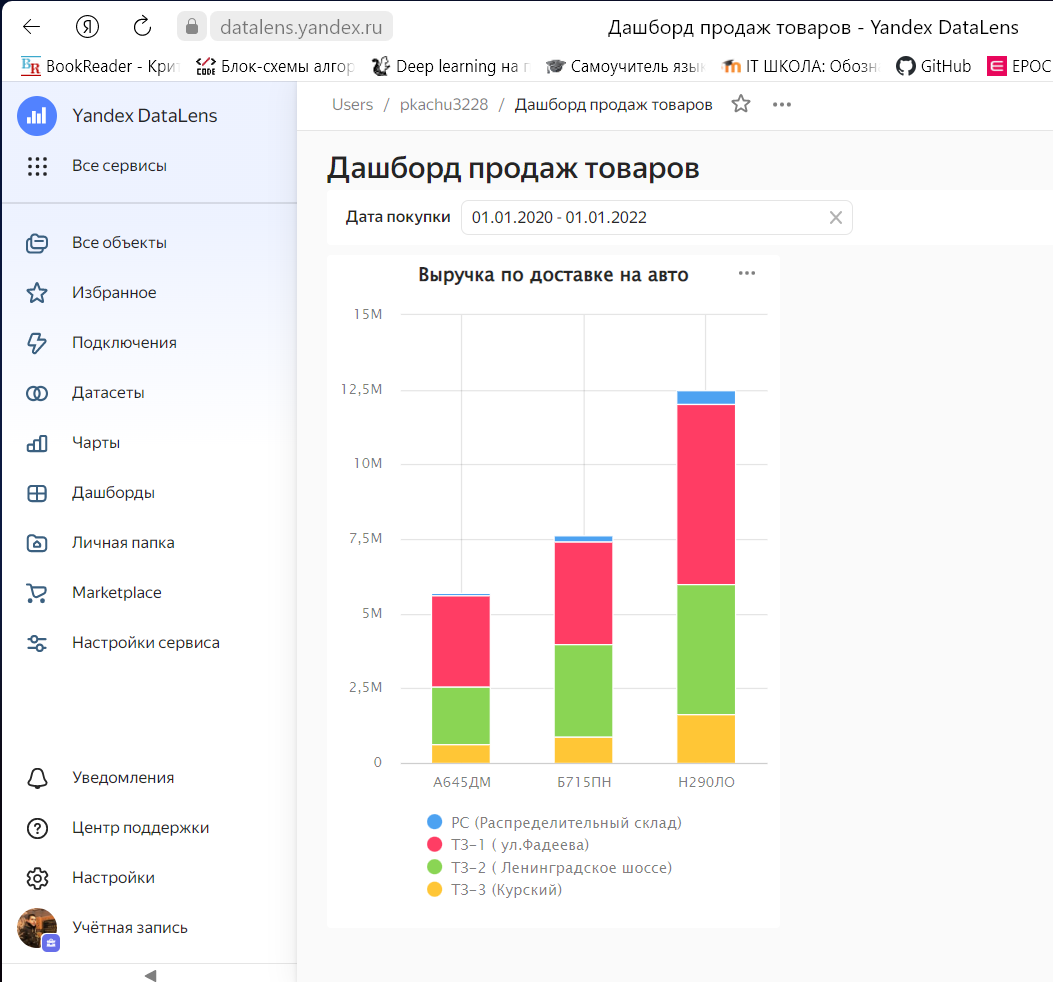


Рисунок . Начальный дашборд

## Задание

В случае, если у вас возникают проблемы с подключением к DataLens, покажите проблемы своему преподавателю и выполните серию визуализаций используя возможности Loginom по обработке данных.

1. Создайте линейную диаграмму показателей покупки на каждый день. Отобразите линейные диаграммы «суммы покупки» от «даты покупки» с цветом «Филиал». Установите, какие филиалы работают в данный момент, какие уже не работают. Установите, какой филиал самый прибыльный, а какой менее по средней покупке.
2. Сгруппируйте данные по чекам, датам и клиентам. Сформируйте в системе обработки данных выборку, для использования в визуализации данных среднего чека по столбцу сумма покупки.
3. С помощью столбчатой диаграммы сформируйте график количества адресов для каждой машины доставки.
4. Создайте график индикатора средней покупки.
5. Создайте, на свое усмотрение, 2 визуализации, помогающие понять эффективность работы службы доставки товаров. Если нужны дополнительные расчеты, воспользуйтесь Loginom или Python для вычисления группировок и агрегаций, после чего результат в виде таблицы данных погрузите отдельной таблицей в подключения и создайте датасет.
6. Вывести все графики на один дашборд с фильтрацией по дате покупки.

Представленной работой поделиться через ссылку с помощью функции «публичный доступ».

