

Разработка интерактивного дашборда для анализа данных о продажах в Yandex DataLens

Пастухов Никита Сергеевич

Приамурский государственный университет имени Шолом-Алейхема

Студент

Аннотация

В статье рассмотрены ключевые этапы: подключение источника данных, визуализация ключевых показателей, компоновка единого интерфейса. В результате разработки создан функциональный прототип информационно-аналитической системы, реализующей концепцию интерактивной бизнес-аналитики для исследования данных предметной области

Ключевые слова: Kaggle, Dataset, Yandex DataLens, визуализация данных, BI-инструменты, анализ данных, дашборд

Development of an interactive dashboard for analyzing sales data in Yandex DataLens

Pastukhov Nikita Sergeevich

Sholom-Aleichem Priamursky State University

Student

Abstract

The article discusses the key steps: connecting a data source, visualizing key indicators, and building a single interface. As a result of the development, a functional prototype of an information and analytical system has been created that implements the concept of interactive business intelligence for domain data research

Keywords: Kaggle, Dataset, Yandex DataLens, data visualization, BI tools, data analysis, dashboard

1 Введение

1.1 Актуальность

Актуальность исследования обусловлена растущей потребностью в инструментах оперативной бизнес-аналитики, сочетающих мощный функционал визуализации с низким порогом вхождения. Современные облачные BI-платформы, такие как Yandex DataLens, предоставляют научно обоснованные методы визуализации данных в интуитивно понятном интерфейсе, что делает продвинутую аналитику доступной для широкого круга специалистов.

1.2 Обзор исследований

Статья Е.П. Панина демонстрирует применение платформы Yandex DataLens для анализа данных о промышленном оборудовании, что позволяет визуализировать ключевые параметры, выявлять факторы отказов и создавать эффективный инструмент для мониторинга технического состояния [1]. Н.О. Плющ и Н.А. Федькова анализируют российский BI-сервис Yandex DataLens как пример системы поддержки принятия решений, подробно рассматривают его функциональные возможности, преимущества (простота, доступность, безопасность) и ключевые недостатки по сравнению с зарубежными аналогами [2]. П.В. Хан и Э.А. Таиров представляют метод автоматизации обработки и визуализации многопараметрических числовых данных научных экспериментов с использованием конвейеров данных, где скрипты для программ gnuplot, gawk и sort комбинируются для эффективного построения и обновления точечных диаграмм, что минимизирует ручной труд исследователя [3]. О.А. Нанишвили описывает методы учета влияния верхней части геологического разреза (включая зону малых скоростей и многолетнюю мерзлоту) на данные сейсморазведки, а именно методику «замещения слоя» и использование преломленных волн для компенсации искажений и повышения точности интерпретации [4]. Д.В. Исаев исследует архитектуру и возможности Apache Spark для распределенной обработки больших данных, проводя сравнительный анализ с Apache Hive и демонстрируя на практическом примере, что Spark является оптимальным решением для итеративной аналитики и сложных преобразований в памяти. В то время как Hive эффективен для быстрых разовых SQL-запросов к статичным данным [5]. Э.В. Талапина занимается анализом дискриминационных рисков, возникающих при обработке данных искусственным интеллектом (на примере профилирования, генетического тестирования и предиктивной аналитики), и предлагает контуры будущего правового регулирования, направленного на предотвращение таких рисков через сочетание юридических и технических мер [6]. Д.И. Агафонов и Е.А. Слепенко разрабатывают метод автоматизированного сбора и обработки экспериментальных данных для оценки эксплуатационных свойств автомобиля с помощью компьютера, аналого-цифрового преобразователя и комплекса датчиков, что повышает точность, оперативность и наглядность лабораторных работ и научных исследований [7].

1.3 Цель исследования

Цель статьи — продемонстрировать возможности платформы Yandex DataLens для построения эффективных аналитических решений на примере разработки интерактивного дашборда. Исследование направлено на практическую реализацию сложных визуализаций, раскрывающую мощный функционал инструмента для преобразования данных в стратегические инсайты.

2 Материалы и методы

В качестве основного инструмента исследования использовалась облачная BI-платформа Yandex DataLens. Материалом для анализа послужил открытый датасет о продажах с платформы Kaggle [<https://www.kaggle.com/datasets/bravehart101/sample-supermarket-dataset?resource=download>], методология работы включала его загрузку, трансформацию и последующую визуализацию с помощью интерактивных чартов.

3 Результаты и обсуждения

Скачиваем Dataset «Sample Superstore Dataset» с платформы Kaggle. Регистрируемся и заходим в DataLens (рисунок 1).

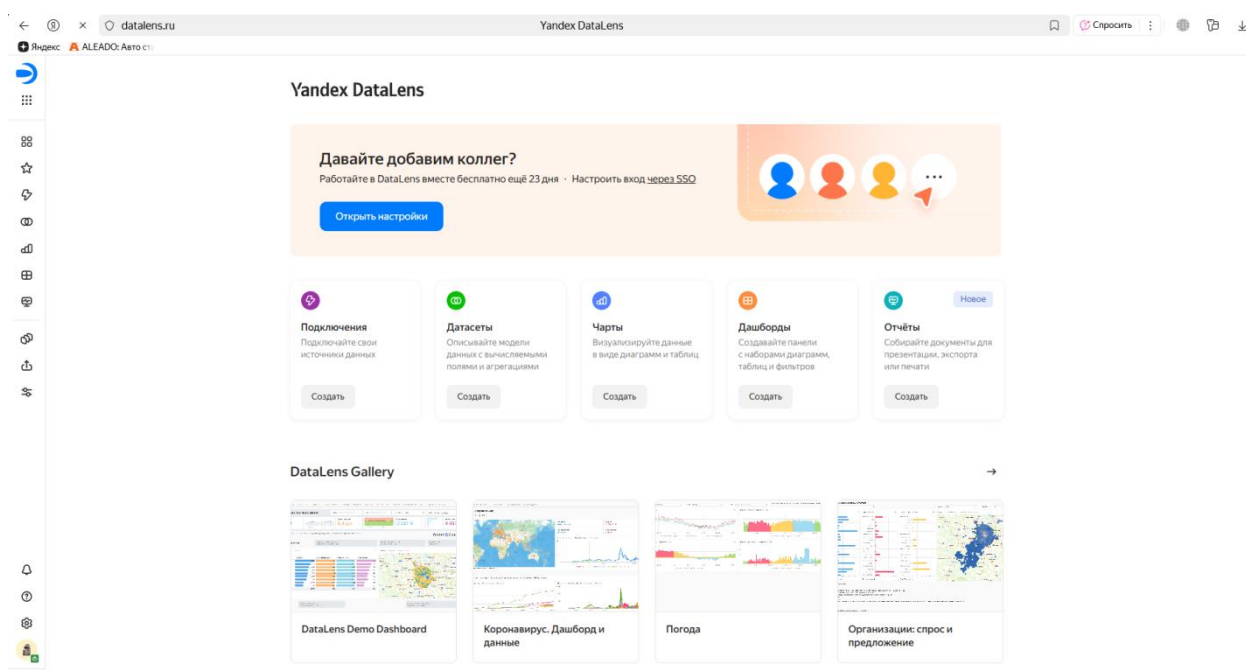


Рисунок 1 – Регистрация в DataLens

В левой панели переходим в раздел «Коллекция и воркбуки» и создаем воркбук под названием «Project» (рисунок 2).

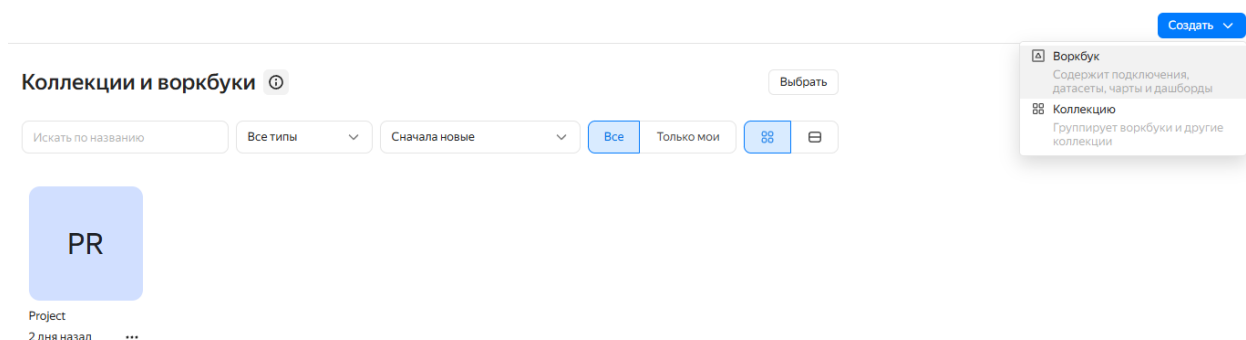


Рисунок 2 – Создание воркбука

Заходим в воркбук, «Создать подключение», «Файлы и сервисы», «Файлы» (рисунок 3). Выбираем файл «SampleSuperstore.csv.zip» и даем название подключению «Подключение_1» (рисунок 4).

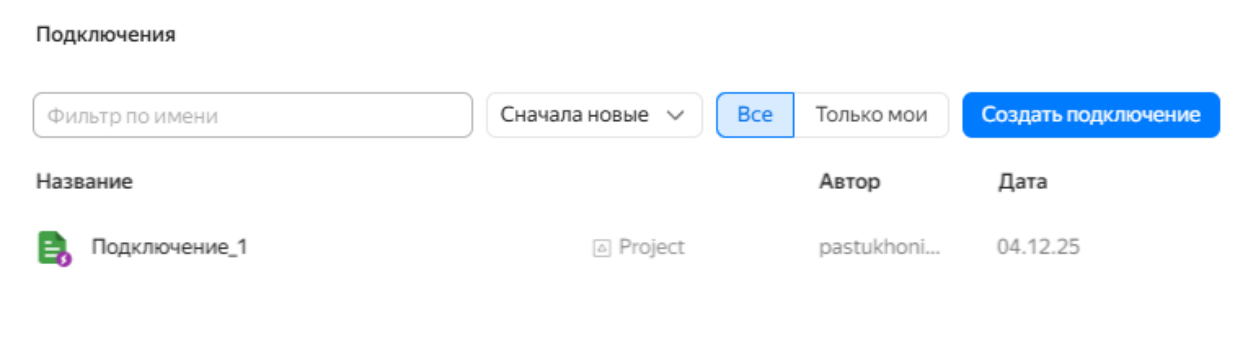


Рисунок 3 – Подключение таблицы данных о продажах

Режим отгрузки	Сегмент	Страна	Город	Государство	Почтовый индекс	Регион	Категория	Подкатегория	Продажи	Количество	Скидка	Прибыль
Second Class	Consumer	United States	Henderson	Kentucky	42420	South	Furniture	Bookcases	261.96	2	0.0	41.9136
Second Class	Consumer	United States	Henderson	Kentucky	42420	South	Furniture	Chairs	731.94	3	0.0	219.582
Second Class	Corporate	United States	Los Angeles	California	90036	West	Office Supplies	Labels	14.62	2	0.0	6.8714
Standard Class	Consumer	United States	Fort Lauderdale	Florida	33311	South	Furniture	Tables	957.5775	5	0.45	-383.031
Standard Class	Consumer	United States	Fort Lauderdale	Florida	33311	South	Office Supplies	Storage	22.368	2	0.2	2.5164
Standard Class	Consumer	United States	Los Angeles	California	90032	West	Furniture	Furnishings	48.86	7	0.0	14.1694
Standard Class	Consumer	United States	Los Angeles	California	90032	West	Office Supplies	Art	7.28	4	0.0	1.9656
Standard Class	Consumer	United States	Los Angeles	California	90032	West	Technology	Phones	907.152	6	0.2	90.7152
Standard Class	Consumer	United States	Los Angeles	California	90032	West	Office Supplies	Binders	18.504	3	0.2	5.7825
Standard Class	Consumer	United States	Los Angeles	California	90032	West	Office Supplies	Appliances	114.9	5	0.0	34.47
Standard Class	Consumer	United States	Los Angeles	California	90032	West	Furniture	Tables	1706.184	9	0.2	85.3092
Standard Class	Consumer	United States	Los Angeles	California	90032	West	Technology	Phones	911.424	4	0.2	68.3568
Standard Class	Consumer	United States	Concord	North Carolina	28027	South	Office Supplies	Paper	15.552	3	0.2	5.4432
Standard Class	Consumer	United States	Seattle	Washington	98103	West	Office Supplies	Binders	407.976	3	0.2	132.5922
Standard Class	Home Office	United States	Fort Worth	Texas	76106	Central	Office Supplies	Appliances	68.81	5	0.8	-123.858
Standard Class	Home Office	United States	Fort Worth	Texas	76106	Central	Office Supplies	Binders	2.544	3	0.8	-3.816
Standard Class	Consumer	United States	Madison	Wisconsin	53711	Central	Office Supplies	Storage	665.88	6	0.0	13.3176
Second Class	Consumer	United States	West Jordan	Utah	84084	West	Office Supplies	Storage	55.5	2	0.0	9.99
Second Class	Consumer	United States	San Francisco	California	94109	West	Office Supplies	Art	8.56	2	0.0	2.4824
Second Class	Consumer	United States	San Francisco	California	94109	West	Technology	Phones	213.48	3	0.2	16.011
Second Class	Consumer	United States	San Francisco	California	94109	West	Office Supplies	Binders	22.72	4	0.2	7.384
Standard Class	Consumer	United States	Madison	Wisconsin	53711	Central	Office Supplies	Storage	665.88	6	0.0	13.3176

Рисунок 4 – Просмотр данных таблицы

Создаем датасет, нажимаем на кнопку «Создать датасет» и даем ему имя «Датасет_1» (рисунок 5).

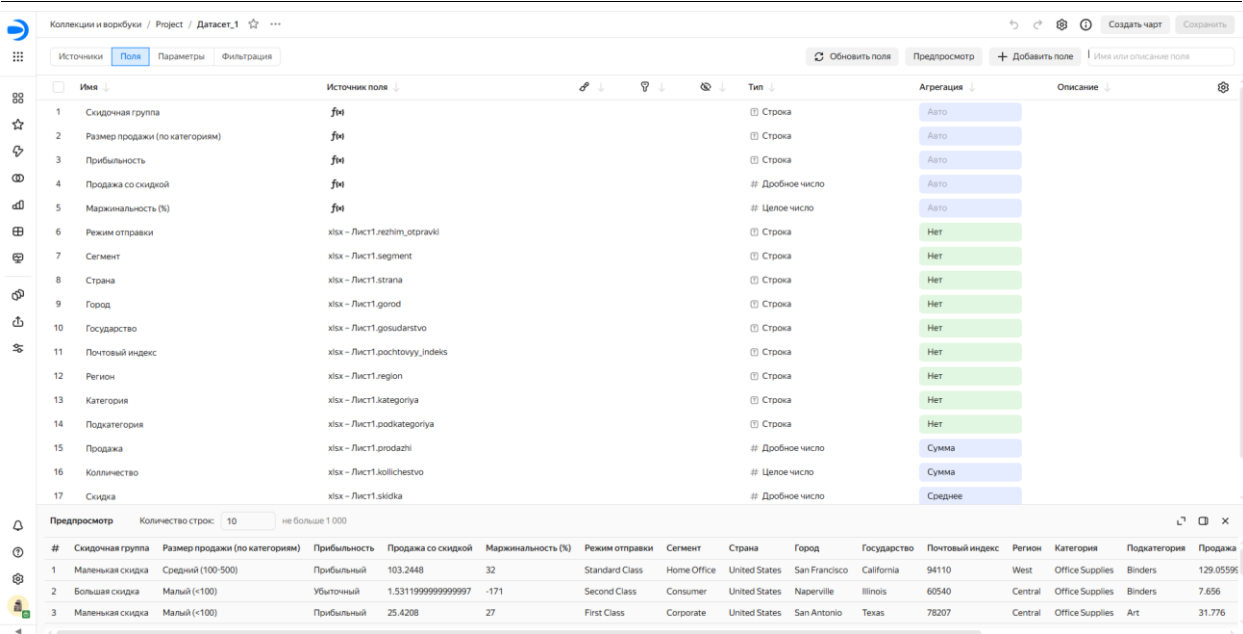


Рисунок 5 – Создание датасета

Настраиваем датасет. Переименовываем названия полей, изменяем их тип и агрегацию (рисунок 6).

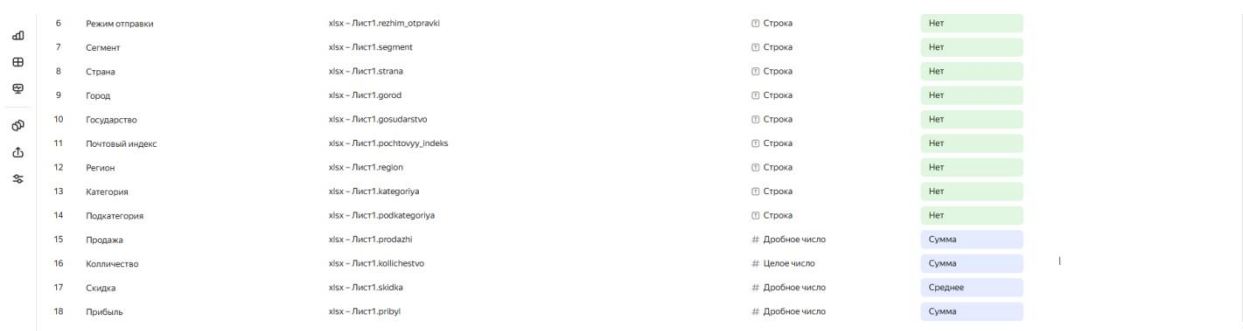


Рисунок 6 – Компановка датасета

Добавляем поля. Нажимаем на «Добавить поле». Создаем первое поле - «Маржинальность (%)» (рисунок 7).

Маржинальность (%)

Формула

Поле из источника

Помочь с формулами

Справочник

Поле

Скидочная группа

Размер продажи (по к...

Прибыльность

Продажа со скидкой

Маржинальность (%)

Режим отправки

Сегмент

Страна

Город

Государство

Почтовый индекс

Регион

Категория

Подкатегория

Продажа

Количество

Скидка

Прибыль

1 IF SUM([Продажа]) = 0

2 THEN 0

3 ELSE

4 SUM([Прибыль]) / SUM([Продажа]) * 100

5 END

Отменить

Сохранить

Рисунок 7 – Добавление поля «Маржинальность (%)»

Добавляем второе поле, назовем «Продажа со скидкой» (рисунок 8).

Продажа со скидкой

Формула

Поле из источника

Помочь с формулами

Справочник

Поле

Скидочная группа

Размер продажи (по к...

Прибыльность

Продажа со скидкой

Маржинальность (%)

Режим отправки

Сегмент

Страна

Город

Государство

Почтовый индекс

Регион

Категория

Подкатегория

Продажа

Количество

Скидка

Прибыль

1 [Продажа] * (1 - [Скидка])

Отменить

Сохранить

Рисунок 8 – Добавление поля «Продажа со скидкой»

Создаем третье поле с названием «Прибыльность» (рисунок 9).

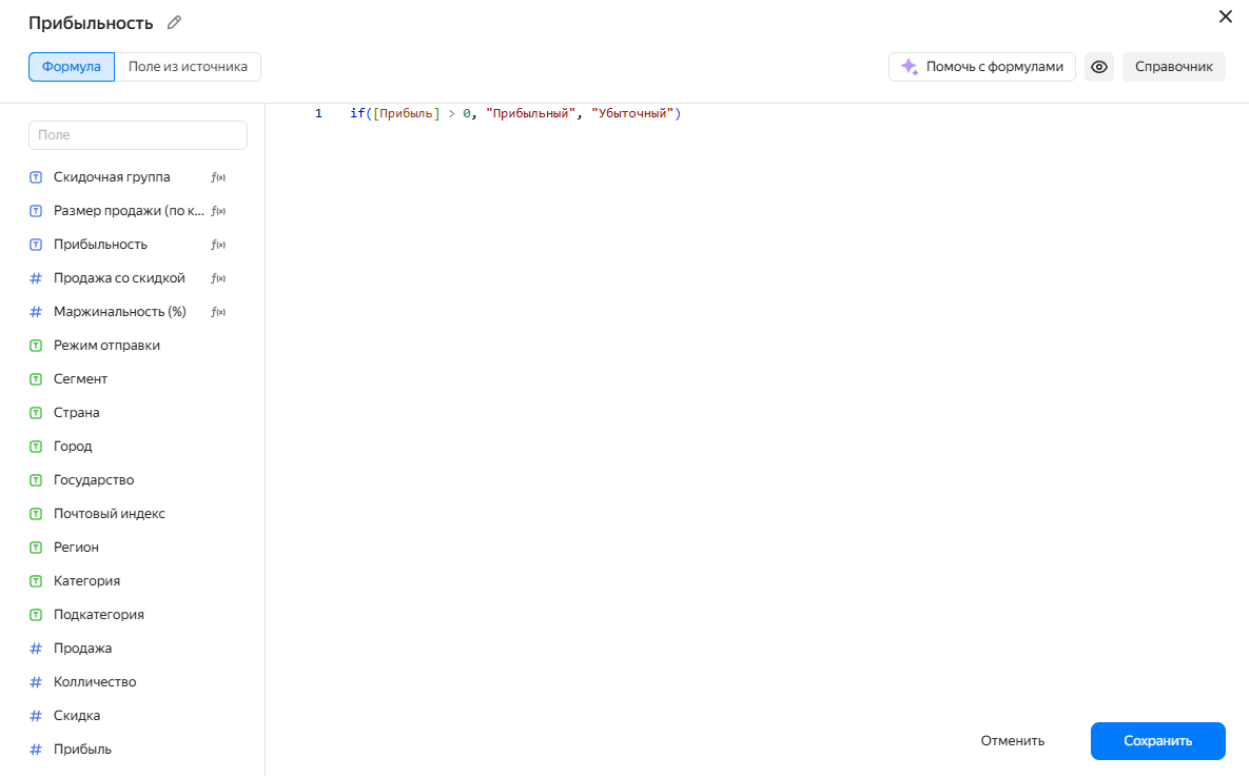


Рисунок 9 – Создание поля «Прибыльность»

Вносим в датасет четвертое поле «Размер продажи (по категориям)» (рисунок 10).

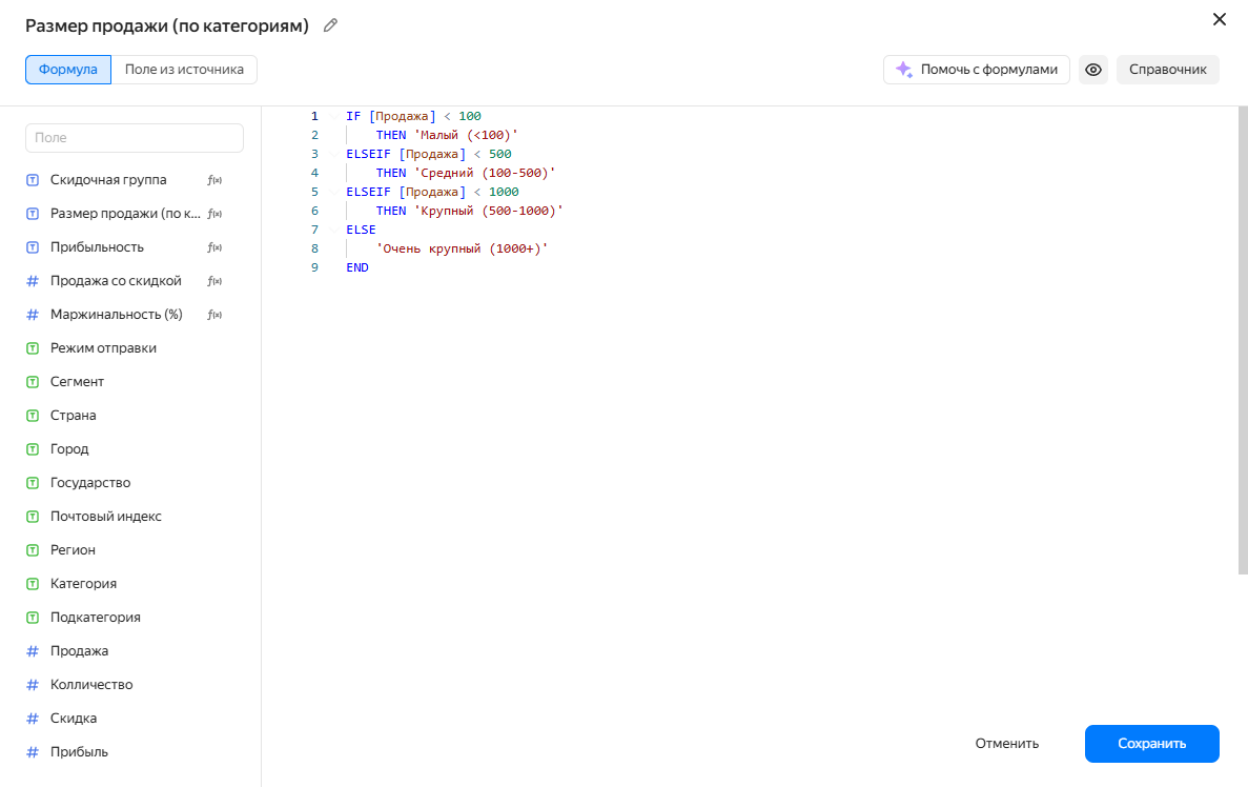


Рисунок 10 – Добавление поля «Размер продажи (по категориям)»

Создаем пятое поле «Скидочная группа» (рисунок 11).

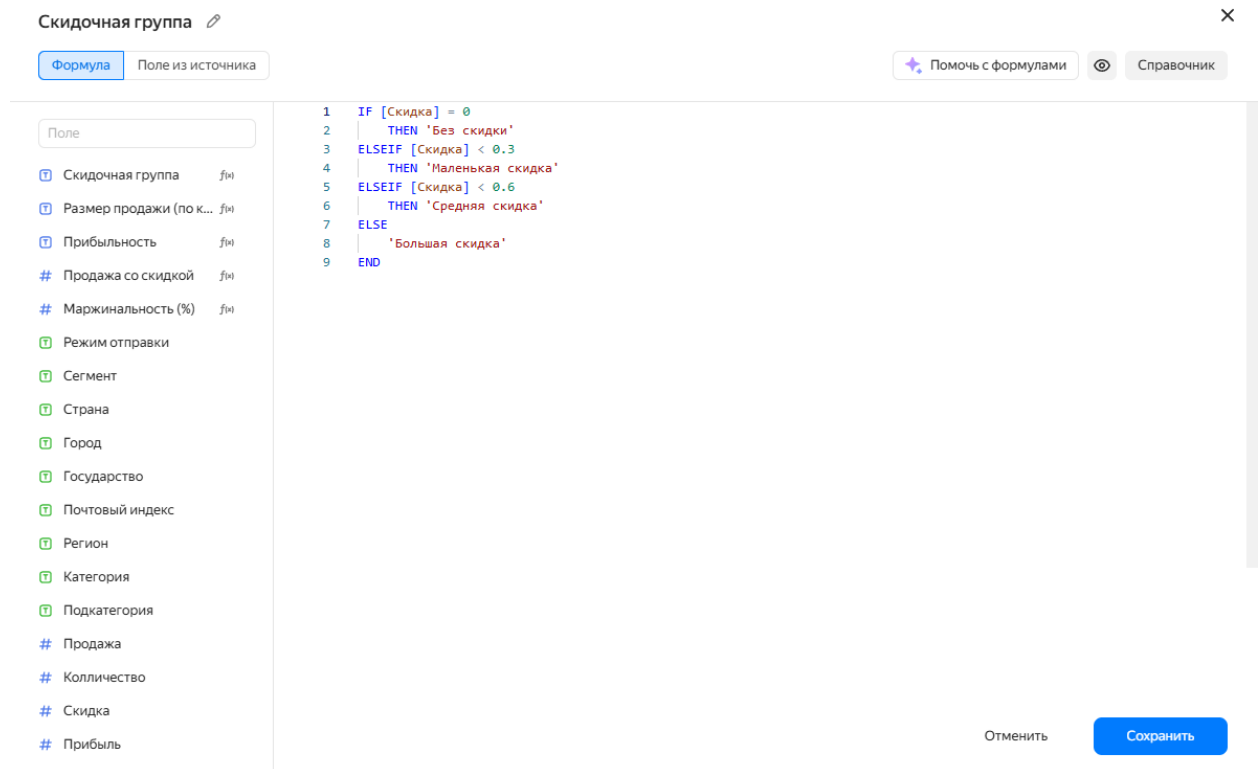


Рисунок 11 – Добавление поля «Скидочная группа»

Корректируем тип новых полей в датасете (рисунок 12).

Имя	Источник поля	Тип	Агрегация	Описание
1 Скидочная группа	fm	Строка	Авто	
2 Размер продажи (по категориям)	fm	Строка	Авто	
3 Прибыльность	fm	Строка	Авто	
4 Продажа со скидкой	fm	Дробное число	Авто	
5 Маржинальность (%)	fm	Целое число	Авто	

Рисунок 12 – Корректировка типа новых полей

Создаем чарт, в правом верхнем углу нажимаем на кнопку «Создать чарт».

Задаем настройки: Тип – Индикатор, показатель – «Продажа». Сохраняем как «Индикатор – Продажа» (рисунок 13).

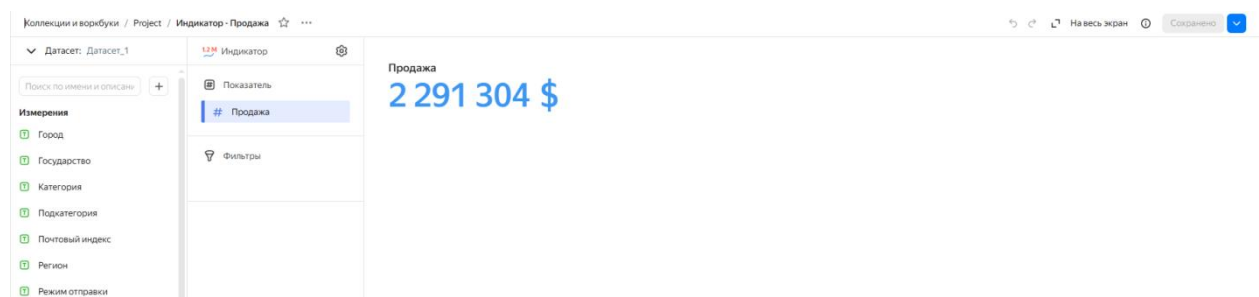


Рисунок 13 – Создание чарта «Индикатор – Продажа»

Создаем чарт «Индикатор - Прибыль». Перетаскиваем поле «Прибыль» в параметр показатель и сохраняем как копию (рисунок 14). Для создания

чарта «Индикатор – Маржинальность» повторяем порядок действий, но заменяем «Прибыль» на «Маржинальность (%)» (рисунок 15).

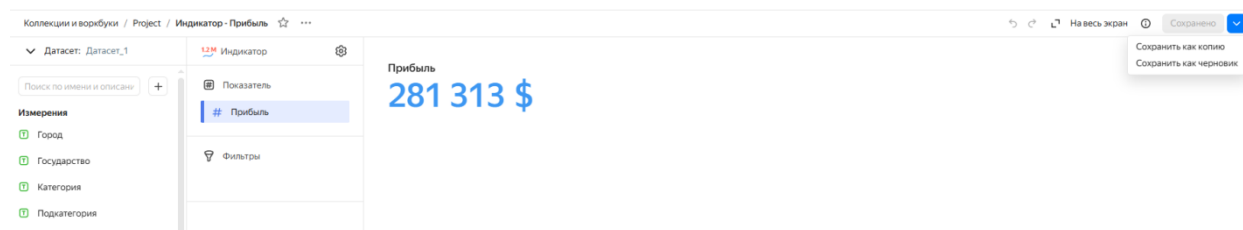


Рисунок 14 – Создание индикатора «Индикатор – Прибыль»

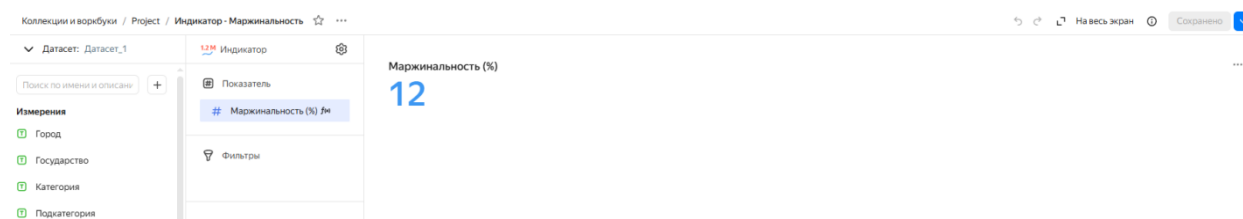


Рисунок 15 – Создание индикатора «Индикатор – Маржинальность»

Заменяем индикатор на древовидную диаграмму.

В параметре «Измерения» добавляем поле «Категория» и поле «Подкатегория». В параметр «Размер» перетаскиваем поле «Прибыль». Сохраняем чарт как «Древовидная диаграмма – Прибыль по категориям» (рисунок 16).

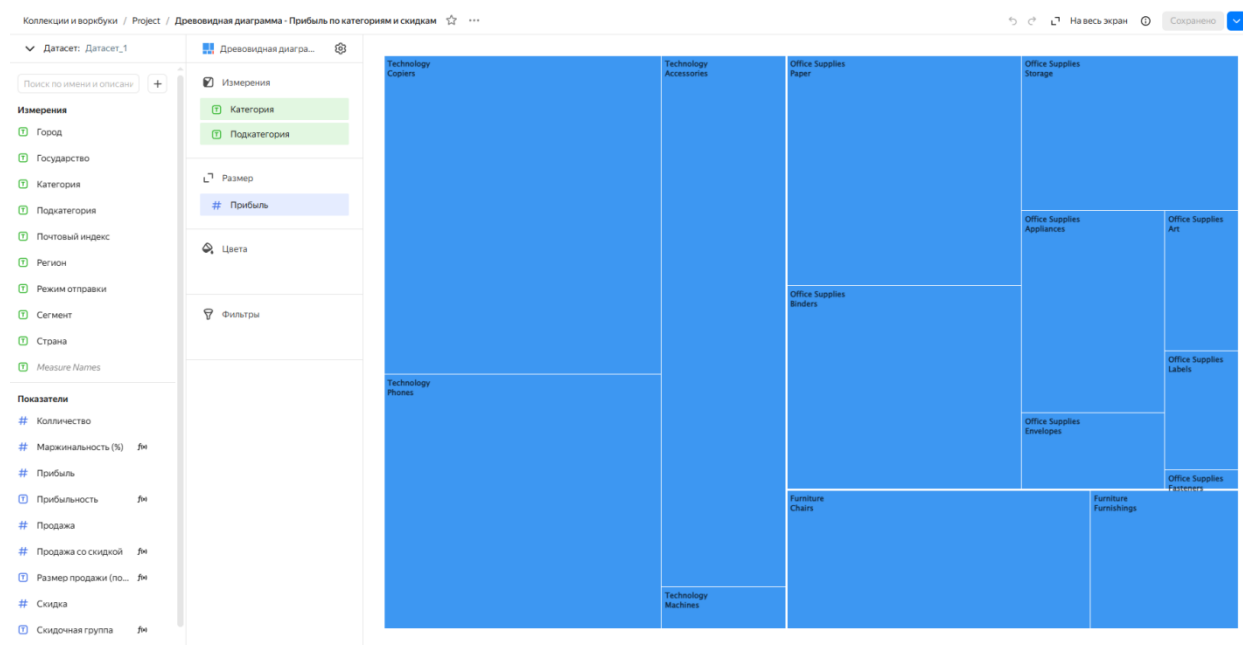


Рисунок 16 – Построение древовидной диаграммы

Выбираем Таблицу. Параметр «Столбцы»заполняем полем «Регион», а поле «Продажа со скидкой» заносим в параметры «Цвет», «Сортировка». Сохраняем как «Таблица – Регион – Продажа со скидкой». Данная таблица показывает цену со скидкой в разных регионах по убыванию (рисунок 17).

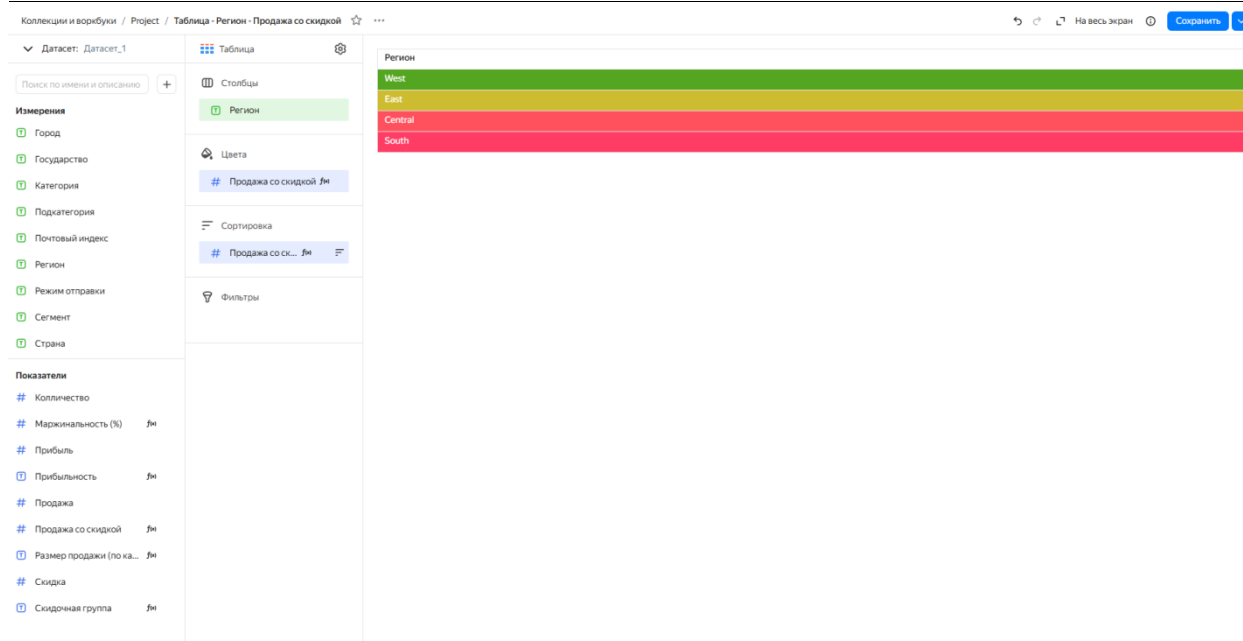


Рисунок 17 – Создание таблицы «Таблица – Регион – Продажа со скидкой»

Добавляем круговую диаграмму, выбираем «Круговая диаграмма».

Поле «Сегмент» переносим в показатели «Категории» и «Подписи». Поле «Продажа» переносим в показатели «Цвет», «Показатель», «Сортировка» (рисунок 18).

Дополнительно, изменяем настройки. Наводим курсор мыши на значок, находящийся слева от показателя «Сегмент» и открываем его. Изменяем свойства сегмента. Выставляем индикатор «Агрегация» на «нет» (рисунок 19).

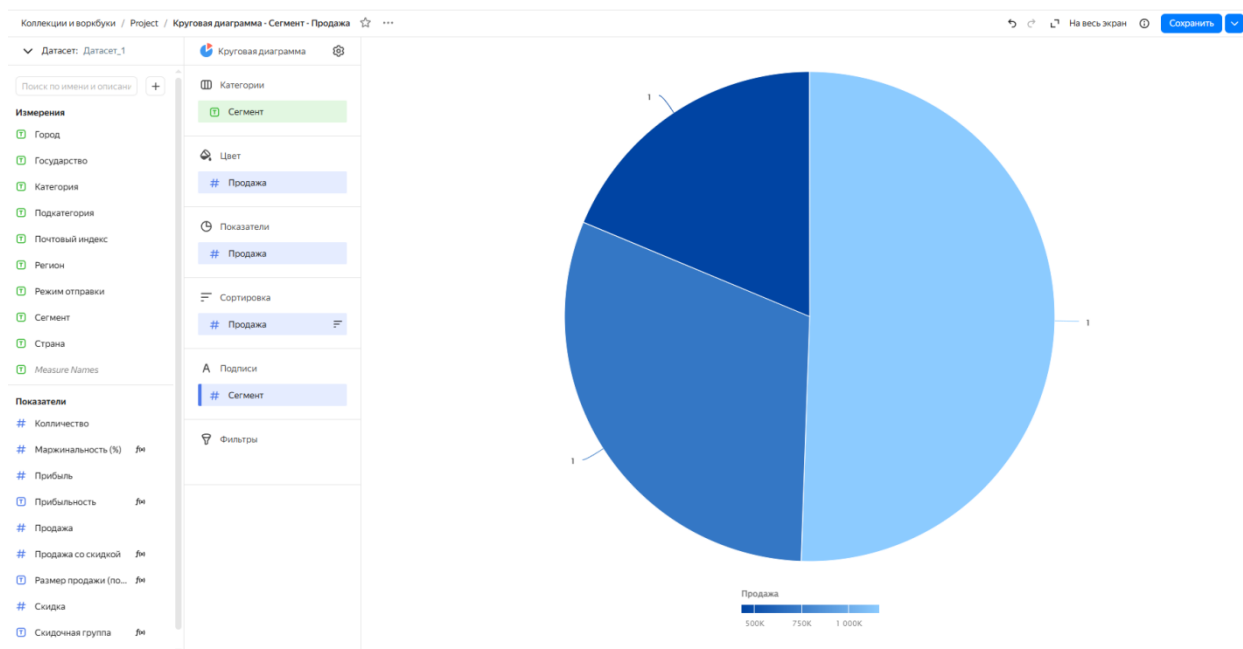


Рисунок 18 – Построение круговой диаграммы

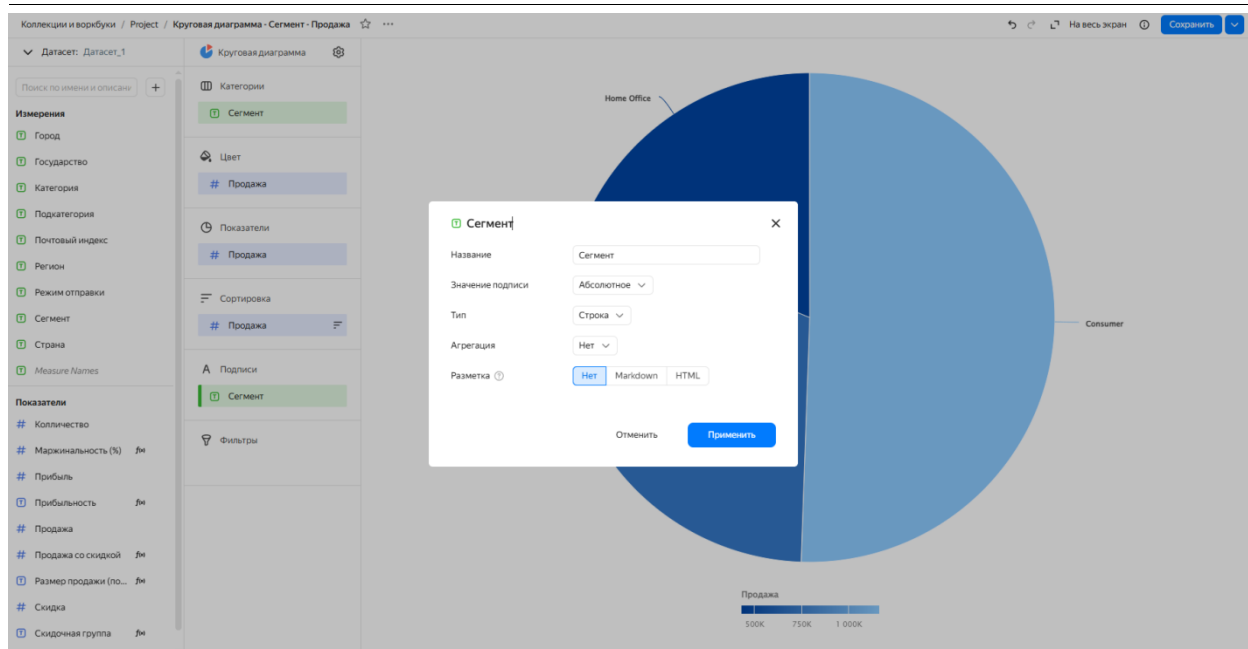


Рисунок 19 – Настройка поля «Сегмент»

Переносим чарты на дашборд. Переходим в левую панель, нажимаем «Дашборды», «Создать дашборд».

Переходим в воркбук «Project».

В нижнем перечне находим «Чарт», при переносе чарта на дашборд снимаем галочку, рядом с настройкой «Заголовок», только для индикаторов.

Нажимаем выбрать, переносим все созданные чарты на дашборд.

Сохраняем дашборд как «Дашборд_1» (рисунок 20).

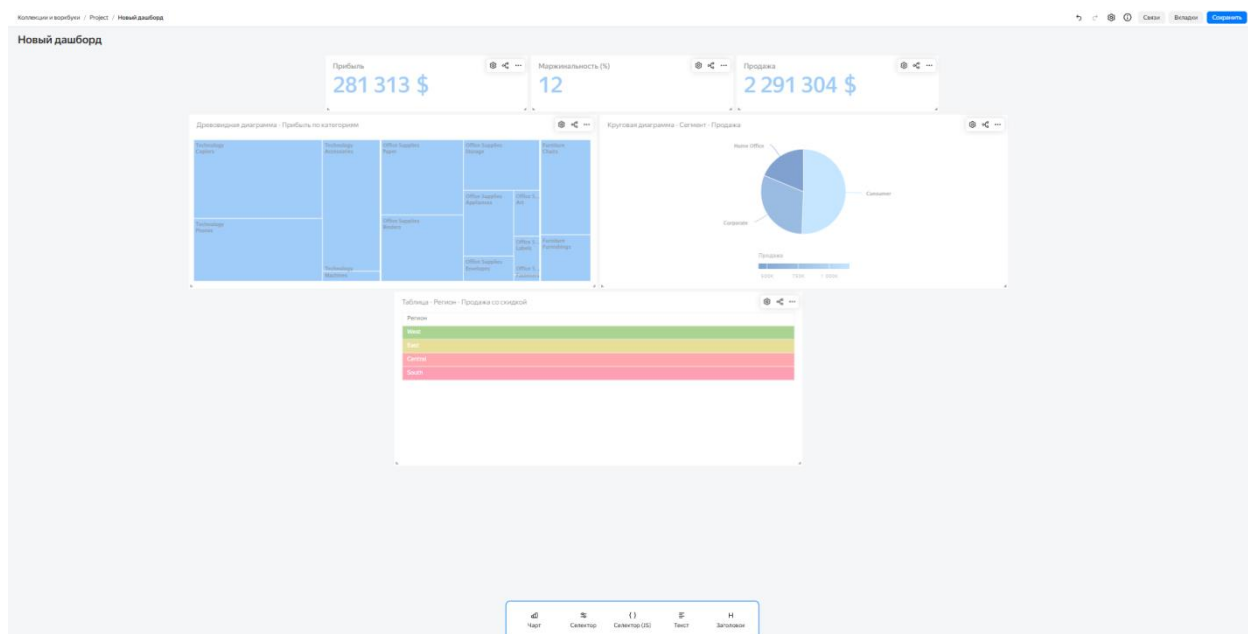


Рисунок 20 – Создание дашборда

4 Выводы

В результате на основе данных о продажах разработан и протестирован интерактивный дашборд анализа продаж в среде Yandex DataLens. Данная

платформа является эффективным и доступным инструментом, позволяющим преобразовывать сырые данные в наглядные интерактивные отчёты для поддержки принятия решений.

Библиографический список

1. Панин Е.П. Применение BI-платформы Yandex DataLens для анализа технических данных // Тенденции развития науки и образования. 2025. № 122-5. С. 123-126. URL: <https://elibrary.ru/item.asp?id=82645390>
2. Плющ Н.О., Федькова Н.А. Проблемы и перспективы Yandex DataLens на рынке бизнес-аналитики // Вестник образовательного консорциума Среднерусский университет. Информационные технологии. 2022. № 2 (20). С. 22-26. URL: <https://elibrary.ru/item.asp?id=50036404>
3. Хан П.В., Таиров Э.А. Автоматизация обработки и визуализации многопараметрических числовых данных при помощи Gnuplot, Gawk и конвейеров данных // Информационные и математические технологии в науке и управлении. 2019. № 1 (13). С. 114-124. URL: <https://elibrary.ru/item.asp?id=39557500>
4. Нанишвили О.А. Учет влияния верхней части разреза при обработке данных сейсморазведки // Национальная Ассоциация Ученых. 2017. № 6 (33). С. 21-25. URL: <https://elibrary.ru/item.asp?id=32373021>
5. Исаев Д.В. Распределенная обработка данных при помощи Apache Spark // Современная математика и концепции инновационного математического образования. 2020. Т. 7. № 1. С. 208-215. URL: <https://elibrary.ru/item.asp?id=43009237>
6. Талапина Э.В. Обработка данных при помощи искусственного интеллекта и риски дискриминации // Право. Журнал Высшей школы экономики. 2022. № 1. С. 4-27. URL: <https://elibrary.ru/item.asp?id=48307241>
7. Агафонов Д.И., Слепенко Е.А. Разработка метода автоматического сбора и обработки экспериментальных данных с помощью эвм при оценке эксплуатационных свойств автомобиля // Механики XXI века. 2007. № 6. С. 162. URL: <https://elibrary.ru/item.asp?id=20653748>