

Исследование популярности проката велосипедов в городах России с помощью Yandex DataLens

Лапушкина Елена Павловна

Приамурский государственный университет имени Шолом-Алейхема

Студент

Аннотация

Целью данной статьи является, построение дашборда. Построение дашборда проходит на основании данных взятых из датасета. Результатом исследования динамики популярности проката велосипедов в России станет построение готового дашборда с подробным описанием и готовой визуализацией с помощью Yandex DataLens.

Ключевые слова: Дашборд, визуализация данных, чарт, диаграмма, графика.

A study of the popularity of bicycle rental in Russian cities using Yandex DataLens

Lapushkina Elena Pavlovna

Sholom-Aleichem Priamurskiy State University

Student

Abstract

The purpose of this article is to build a dashboard. The dashboard is built based on data taken from the dataset. The result of the study of the dynamics of the popularity of bicycle rental in Russia will be the construction of a ready-made dashboard with a detailed description and ready-made visualization using Yandex DataLens.

Keywords: Dashboard, data visualization, chart, chart, graph.

1 Введение

1.1 Актуальность

На сегодняшний день дашборд это-уникальное программное обеспечение для анализа и сбора данных в режиме реального времени. Ключевыми характеристиками дашбордов выступают наглядность, интерактивность, автоматизация.

1.2 Обзор исследований

Д.А. Рыжкина рассмотрела проблему визуализации данных на основном листе отчета в среде Yandex, а также сокращенные представления основных показателей из огромного массива данных. [1]. Н.О. Плющ, Н.А. Федькова рассмотрели главные возможности ресурса. Сферы деятельности,

пользующиеся данным инструментом визуализации. Показаны преимущества и недостатки при использовании данного сервиса. [2]. В.С. Есенкина в исследовании рассматривает применение инструмента визуализации Yandex DataLens при анализе и обработке данных. DataLens представляет собой систему визуализации данных, которая может быть применена как для стандартного анализа данных, так и в интеллектуальном анализе, связанном с решением задач бизнес-аналитики. [3]. Н.Л. Марголина, К.Е. Ширяев в своей статье рассмотрели влияние одной из тем математического анализа на формирование исследовательской компетенции. Особенностью задач обучения данной темы является не только овладение стандартными методами, но и их комплексное применение при построении графиков функций. [4]. П.С. Ростовцев, Х.М. Салех посвятили статью исследованию генерации датасет для извлечения данных из структурированных документов на основе машинного обучения. В данной работе рассматривается анализ методов и подходов для извлечения данных, разработка системы генерации данных, сбор и генерация данных для дальнейшей классификации с помощью машинного обучения. [5]. W Young Gareth, Kitchinc Rob провели исследование целью которого было создание конкретных требований пользователей и рекомендаций по проектированию новых городских информационных панелей на основе пользовательского опыта в контексте общественной информатики [6] Ици Сяо Кэ Мяо Инцзе в данной статье основное внимание уделили оценке величины зрительного восприятия, особенно суждению о ценности, закодированному выделенной областью диаграммы данных [7]. З.Т. Хасенова определила и рассмотрела современные инструменты для реализации и развертывания сервер приложения. Дано описание и результат создание тепловой карты на основе доступных библиотек [8]. А.В. Смирнов в статье обобщен опыт применения интерактивных веб-приложений (дашбордов) в демографической науке [9].

1.3 Цель исследования

Цель исследования – создать дашборд с помощью которого можно делать визуализацию, а также проводить более детальный анализ данных для любого города.

2 Материалы и методы

Для построения дашборда данные взяты из датасета аудитория <https://datalens.yandex.ru/datasets/udn3tgb319puj-auditoriya>. При помощи селекторов можно визуализировать данные чтоб наглядно видеть результаты по построению задачи в DataLens.

3 Результаты и дискуссии

На первом этапе необходимо создать подключение, а для этого нужно обратиться к данным из открытых источников, расположенных в Marketplace. Для этого нужно зайти <https://datalens.yandex.ru/>, авторизоваться под своим пользователем (рис. 1).

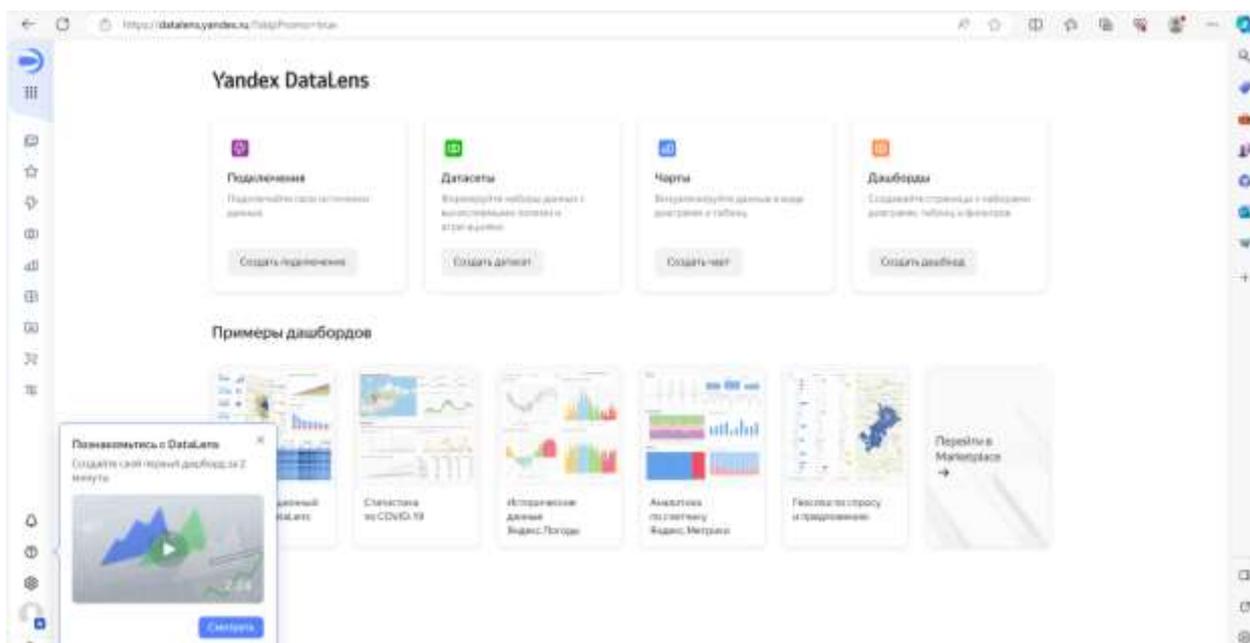


Рисунок 1 –Регистрация в Yandex DataLens

В навигационном меню нажать Marketplace (рис. 2).

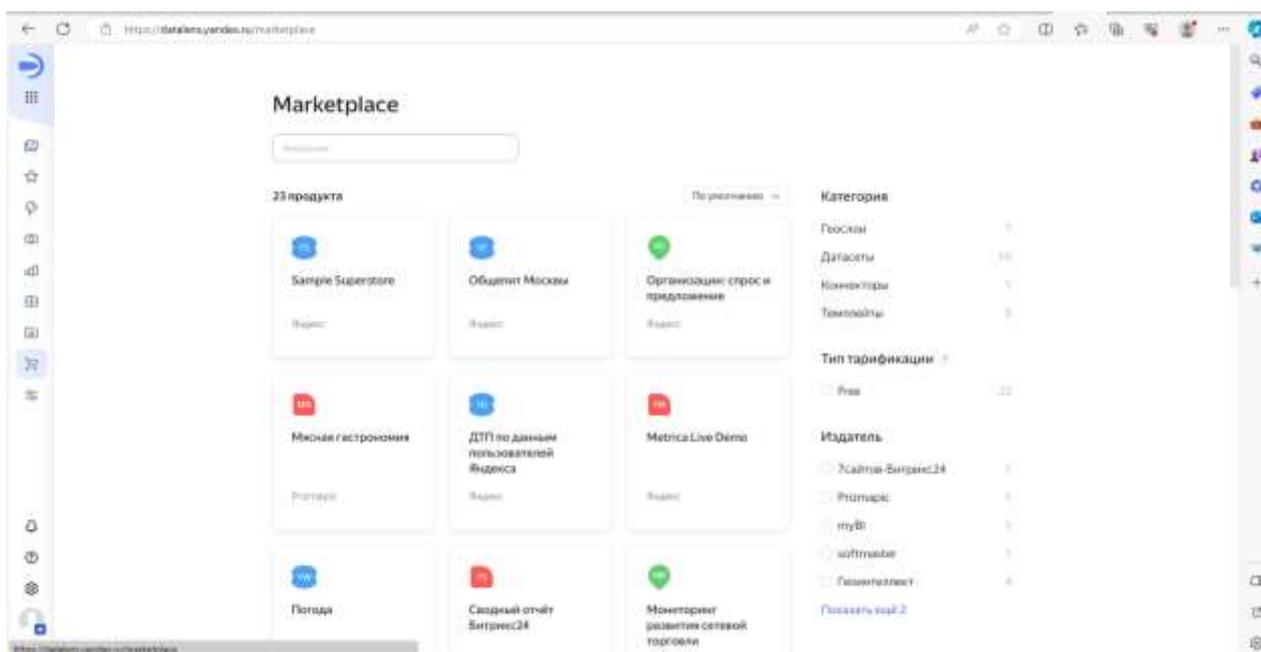


Рисунок 2- Меню Marketplace

На данной странице можно увидеть все открытые продукты Yandex DataLens. Необходимо выбрать категорию геослой (рис. 3).

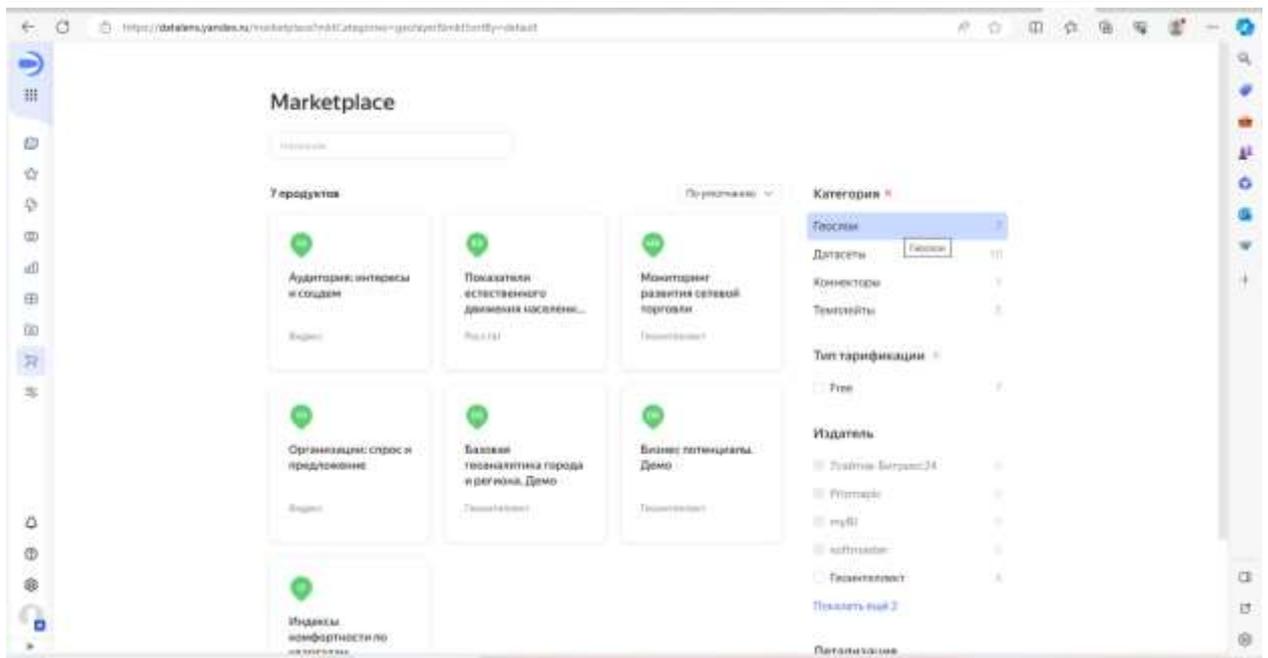


Рисунок 3- Выбор категории геослои

Выбрав категорию геослои выбрать продукт- Аудитория: интересы и соцдем (рис. 4).

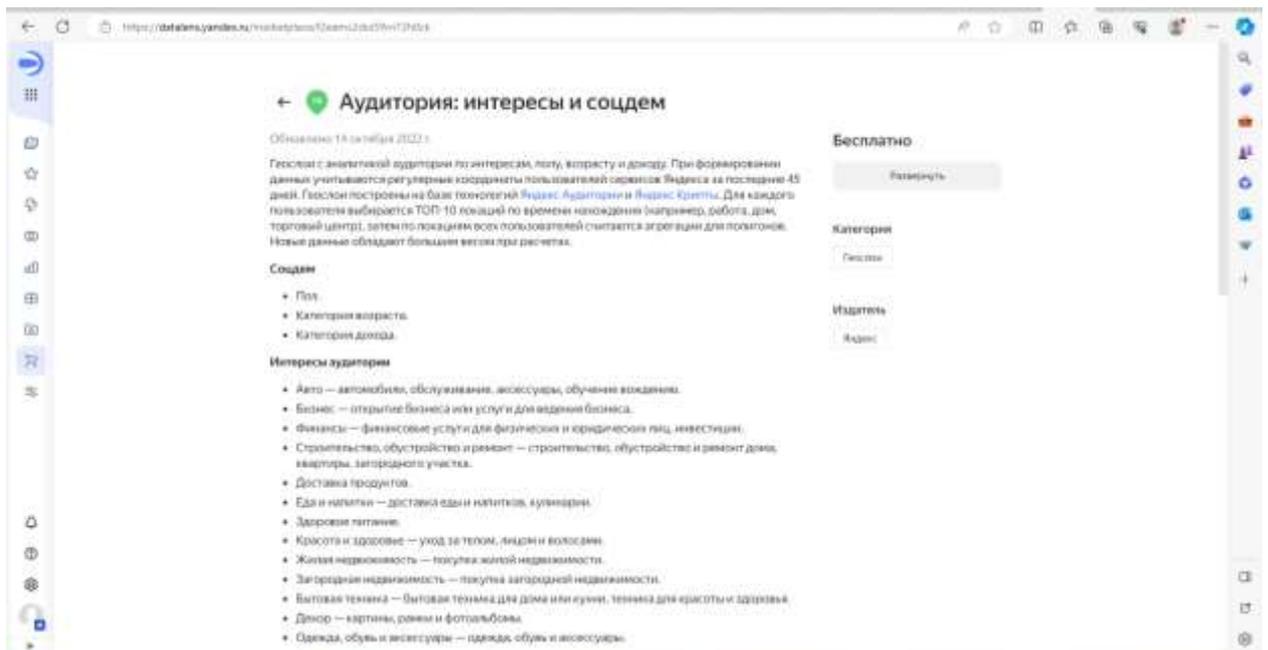


Рисунок 4- Выбор продукта- Аудитория: интересы и соцдем

При нажатии на кнопку подключить, открывается выбор расположения папки, выбрав нажать сохранить и открыть (рис. 5).

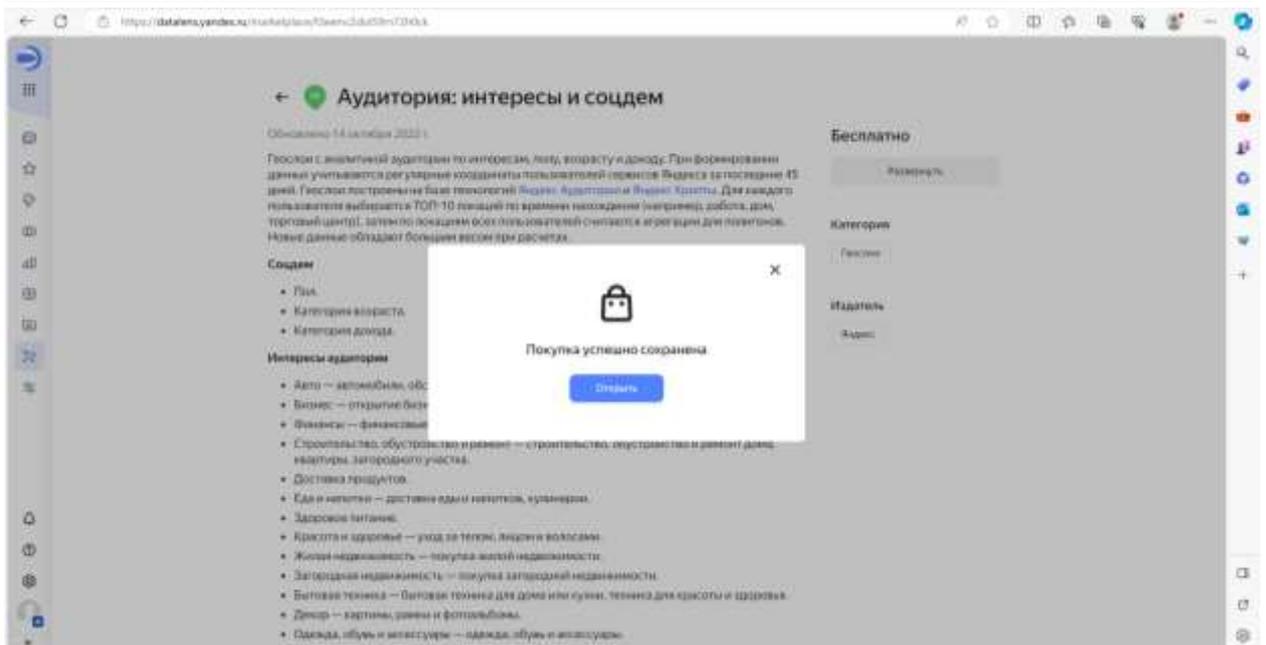


Рисунок 5- Открытие папки

Видно, что на основе открытых данных существует визуализация, необходимо будет создать новые чарты и дашборд (рис. 6).

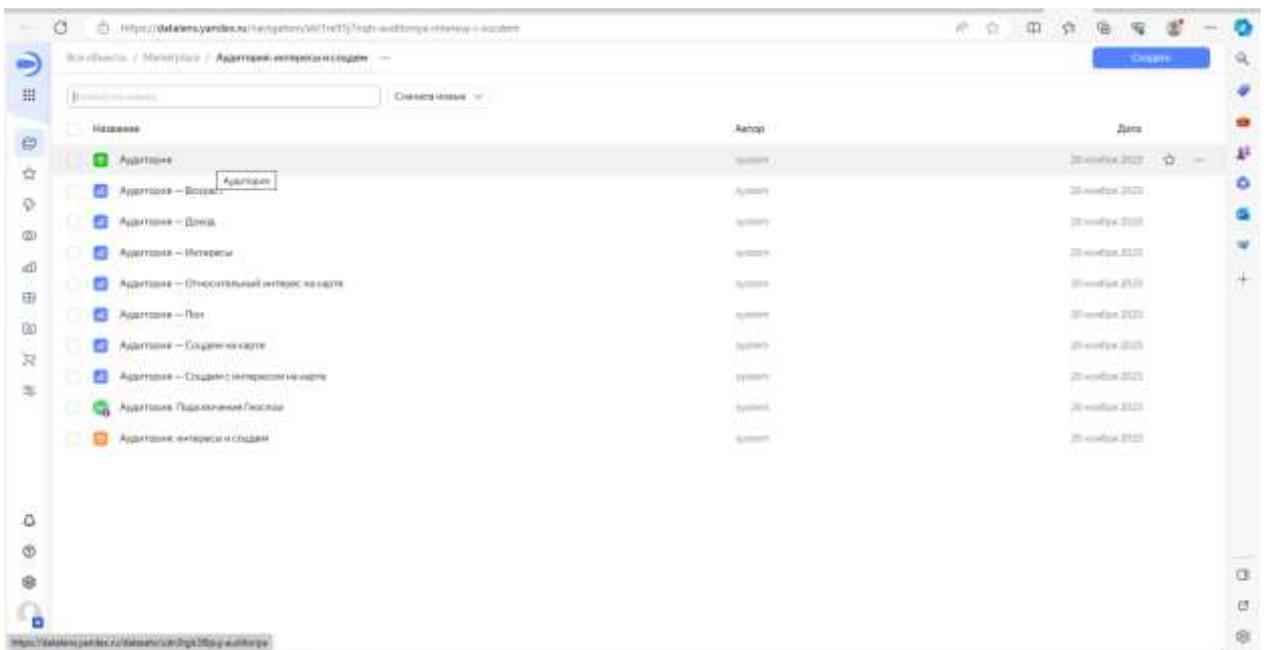


Рисунок 6- Визуализация данных

Следующим этапом будет изучение существующего датасета под названием аудитория (рис.7).

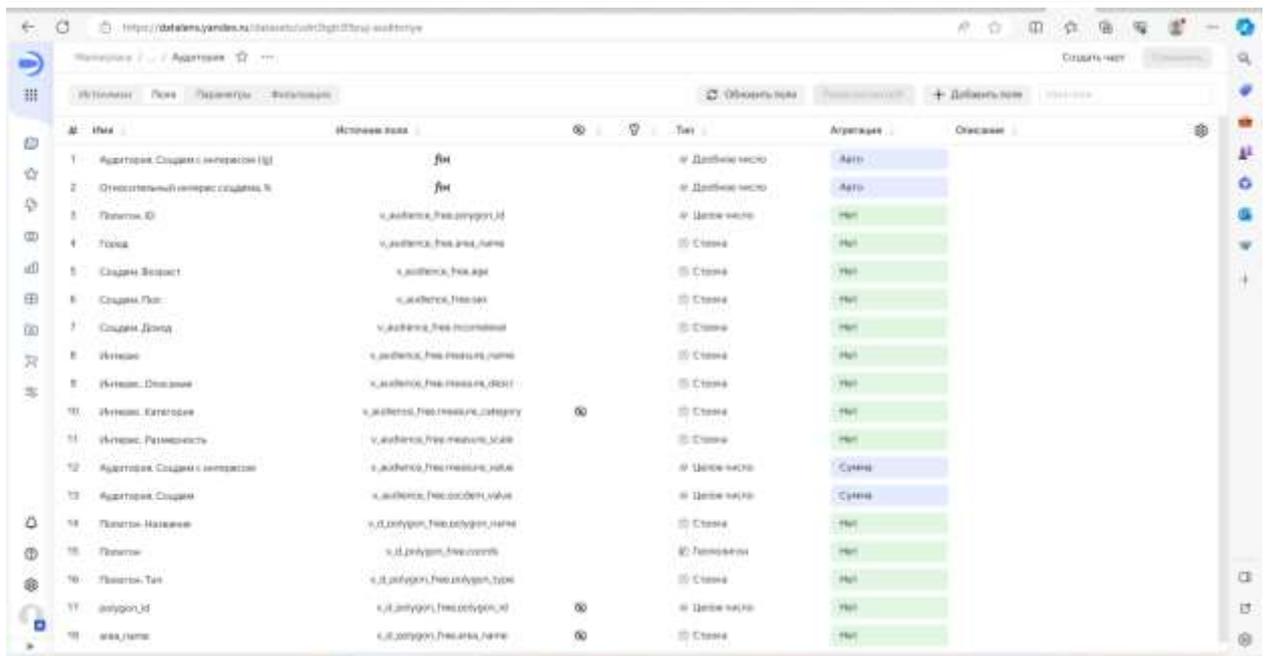


Рисунок 7- Выбор датасета- аудитория

Далее перейти на вкладку источники (рис. 8).

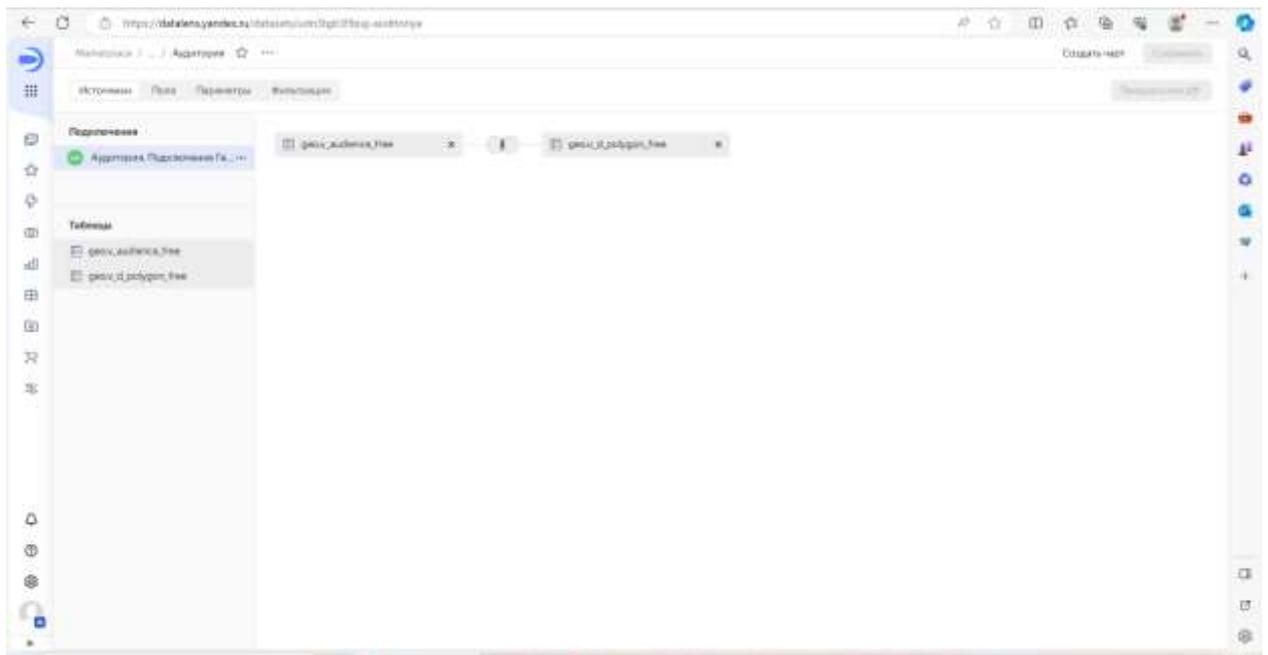


Рисунок 8- Вкладка источники

В подключении уже создана модель данных состоящая из двух таблиц, при нажатии на связь, можно проверить тип связи, и поля, по которым произошло объединение. В данном случае связь автоматически создалась корректно, поэтому применив данные условия нужно перейти к списку полей (рис.9).

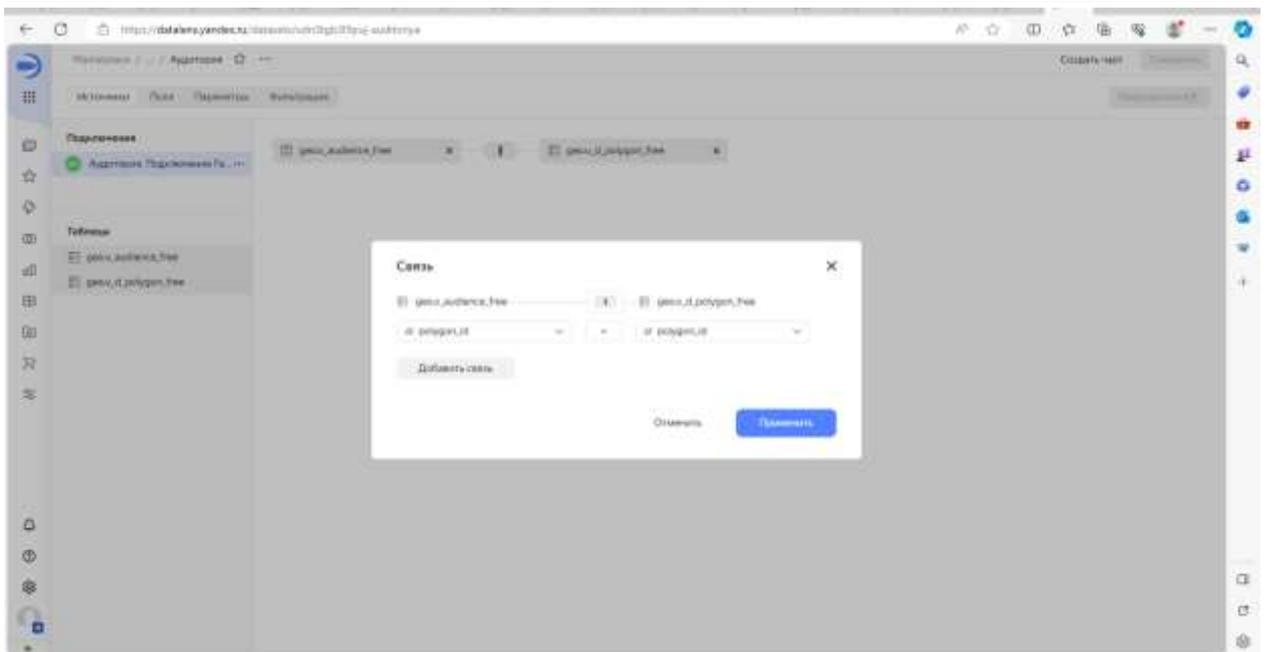


Рисунок 9- Проверка связи

На вкладке поля можно удостовериться в корректности назначенного типа для каждого поля (рис.10).

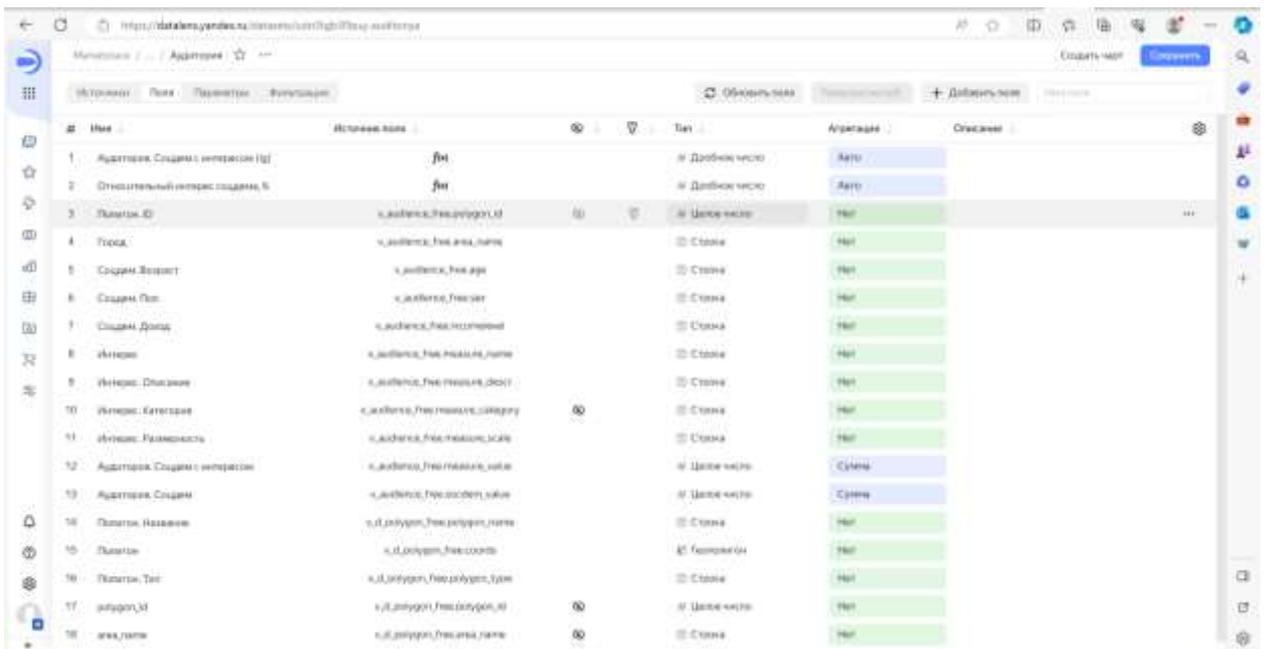


Рисунок 10- Проверка вкладки поля

Следующим действием будет создание различных визуализаций и диаграмм, индикаторов и т.д. Необходимо нажать кнопку- создать чарт (рис. 11).

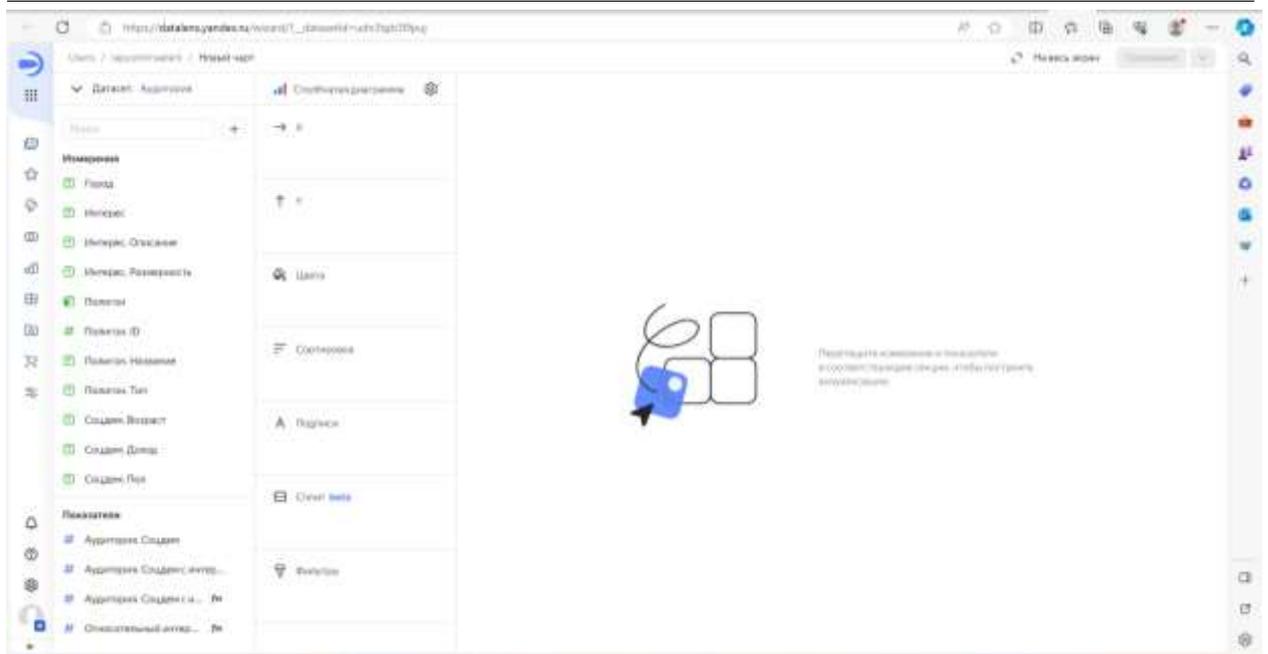


Рисунок 11-Создание чарт

На данной вкладке создать новую визуализацию, для этого необходимо выбрать тип-столбчатая диаграмма (рис.12).

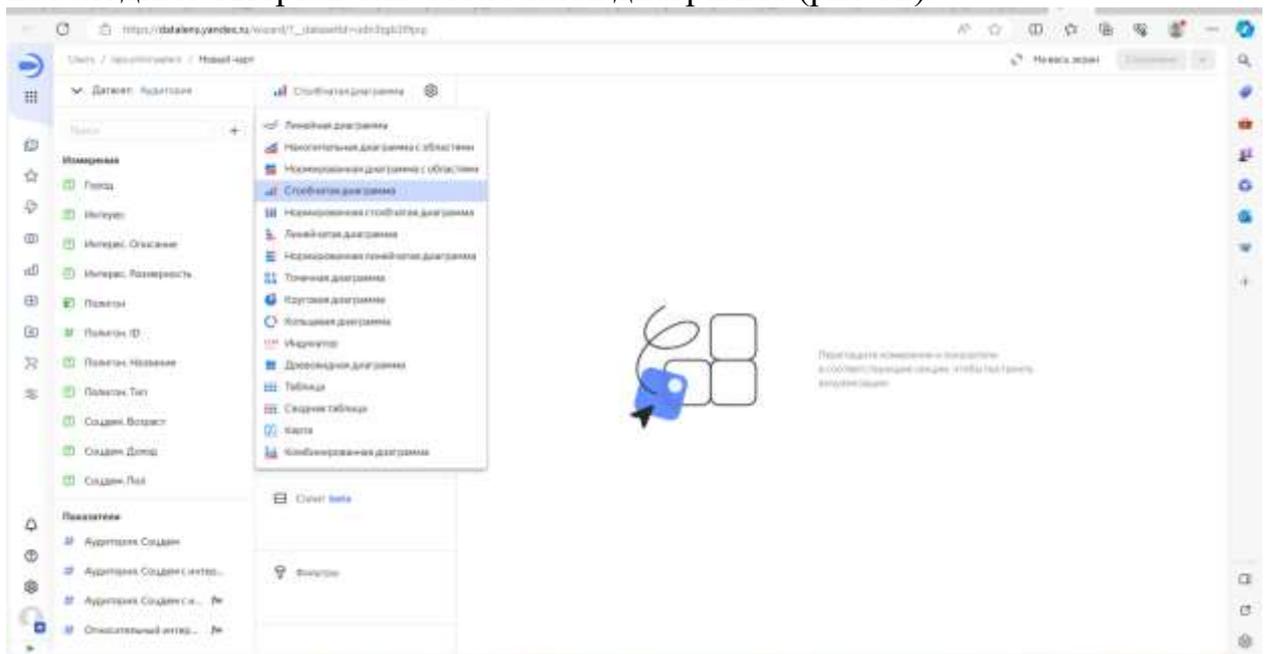


Рисунок 12-Выбор столбчатой диаграммы

По оси OX будет отображаться пол человека, при этом нужно перенести соответствующее поле на ось OX (рис.13).

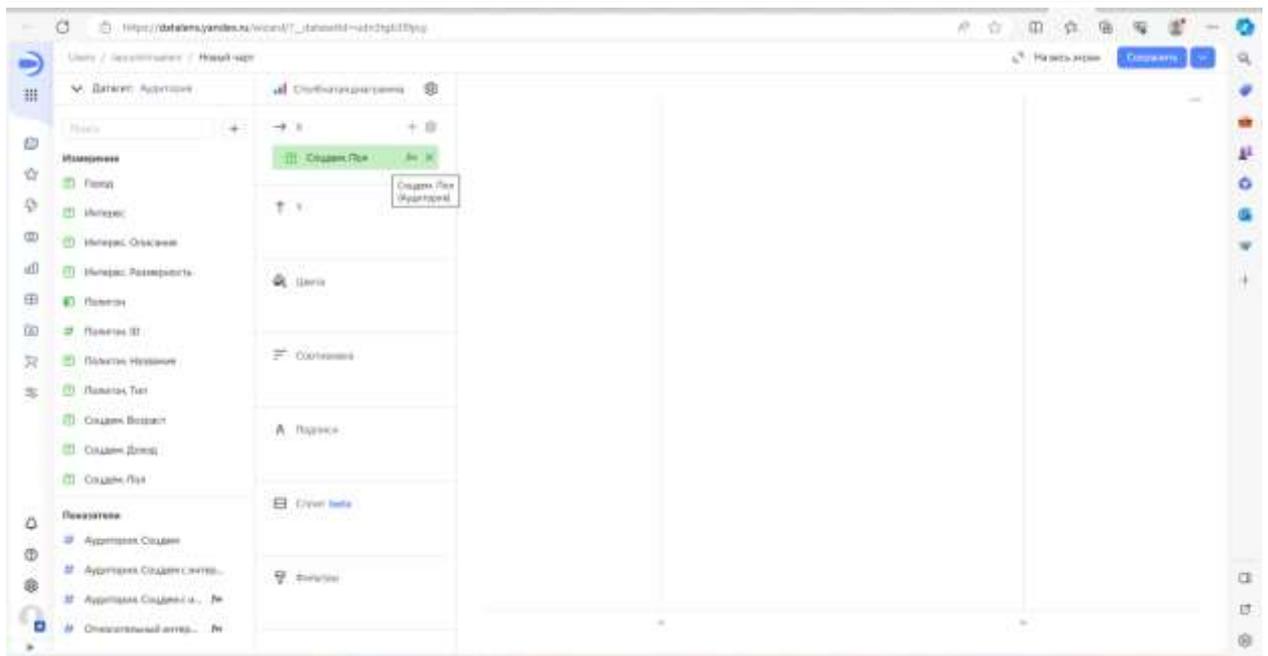


Рисунок 13- Перенос поля на ось x

По оси OY будет размещена сумма общего количества людей с указанным интересом, необходимое поле находится на вкладке- показатели, - Аудитория: соцдем с интересом (рис.14).

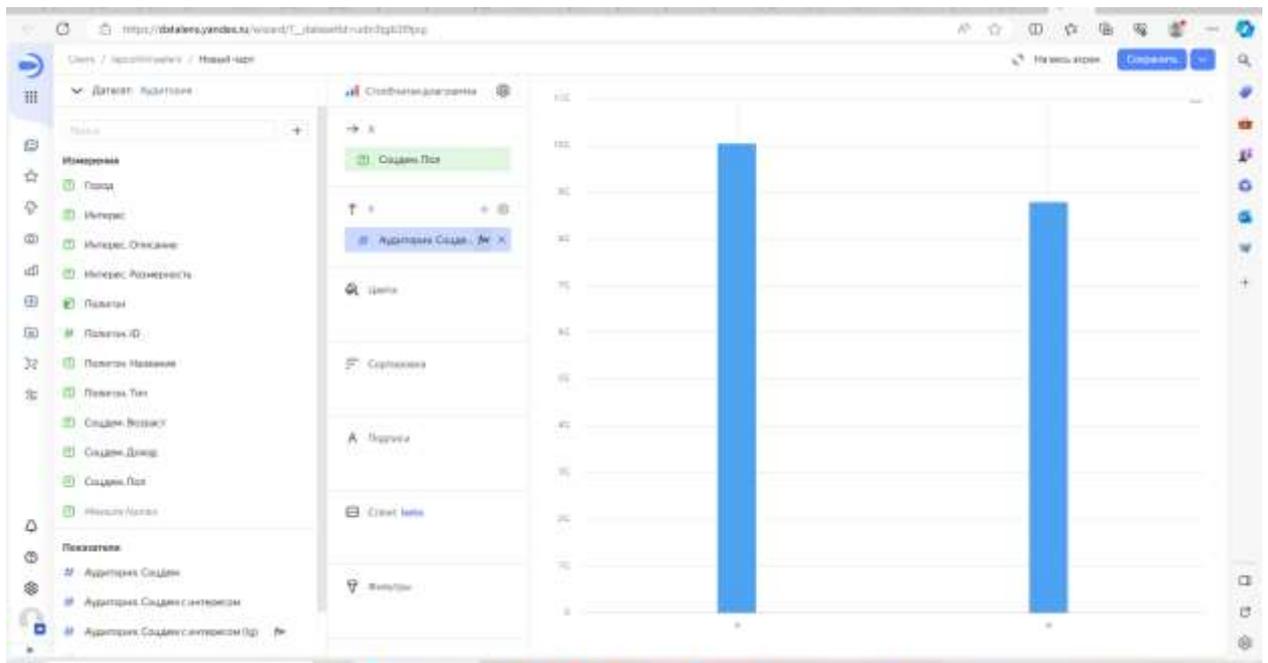


Рисунок 14 –Перенос данных на ось OY

Для красочного дашборда, нужно добавить цвет, сортировку. Для этого нужно перенести Соцдем. пол в поле цвета, теперь можно увидеть, что назначены достаточно яркие цвета (рис.15).

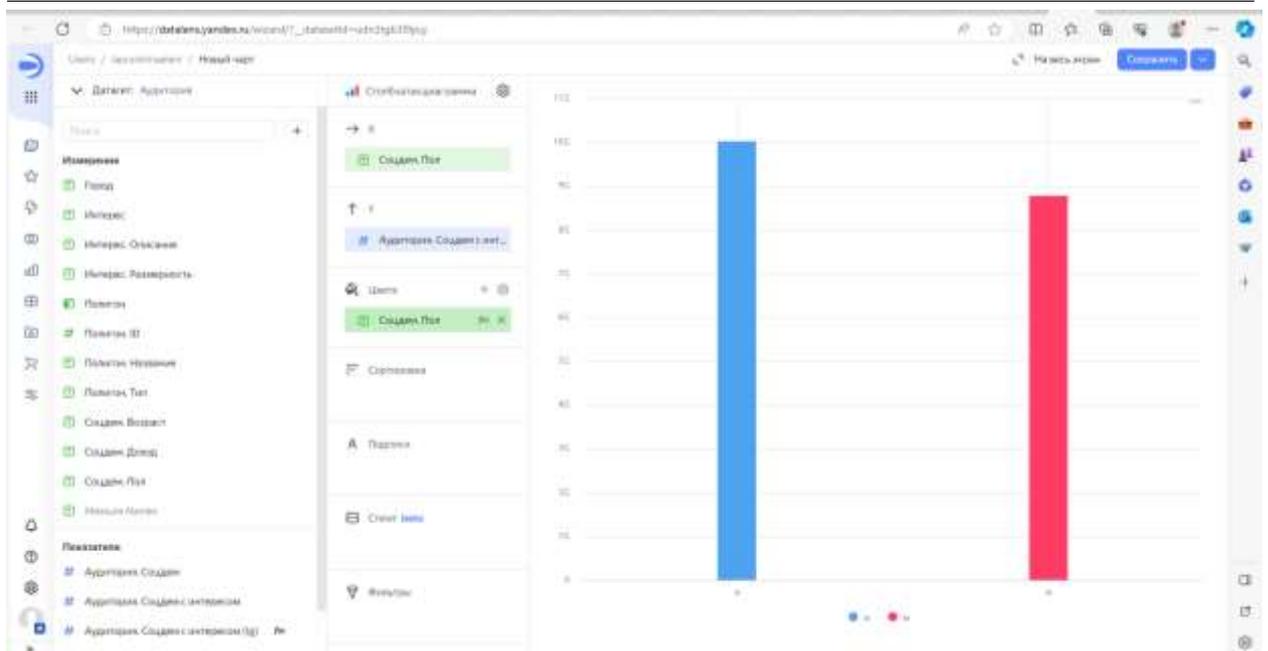


Рисунок 15-Добавление цвета в дашборд

Далее необходимо создать фильтр, для этого перенести измерение-Интерес в фильтры, и в данной вкладке выбираем велосипеды, то есть операция принадлежит множеству велосипедистов, нажать применить фильтр (рис.16).

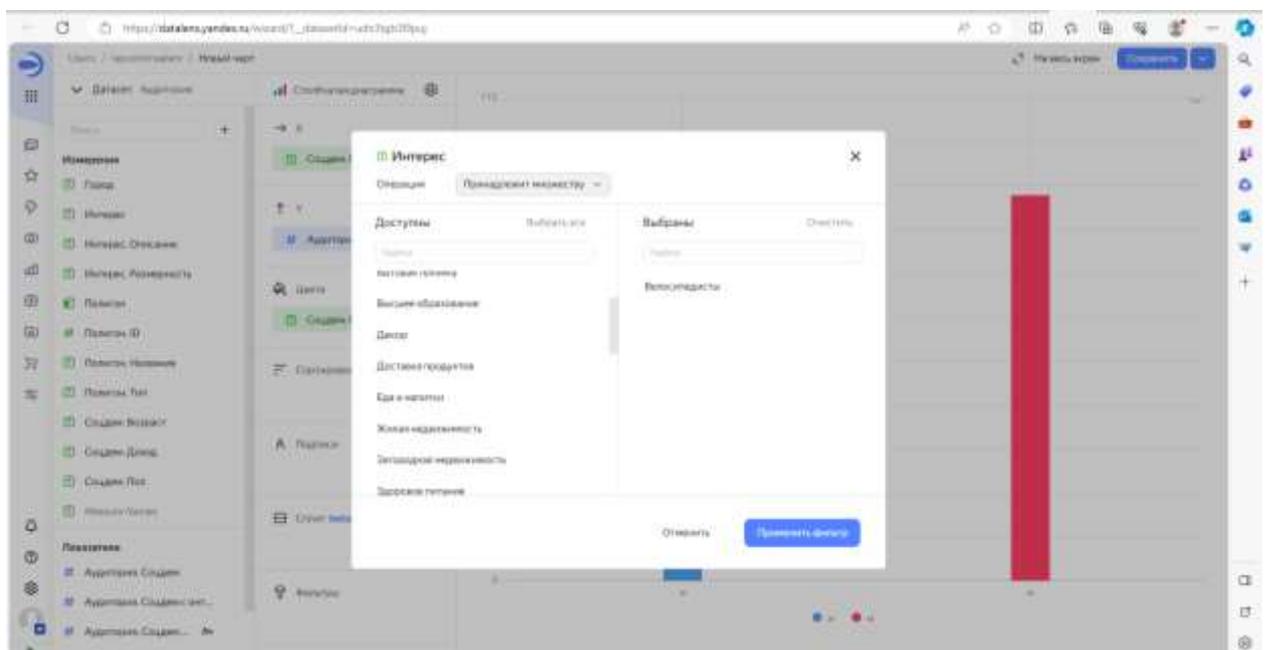


Рисунок 16- Выбор измерения- Интерес

Далее нажать сохранить, назвать чарт- Столбчатая диаграмма пол. (рис.17).

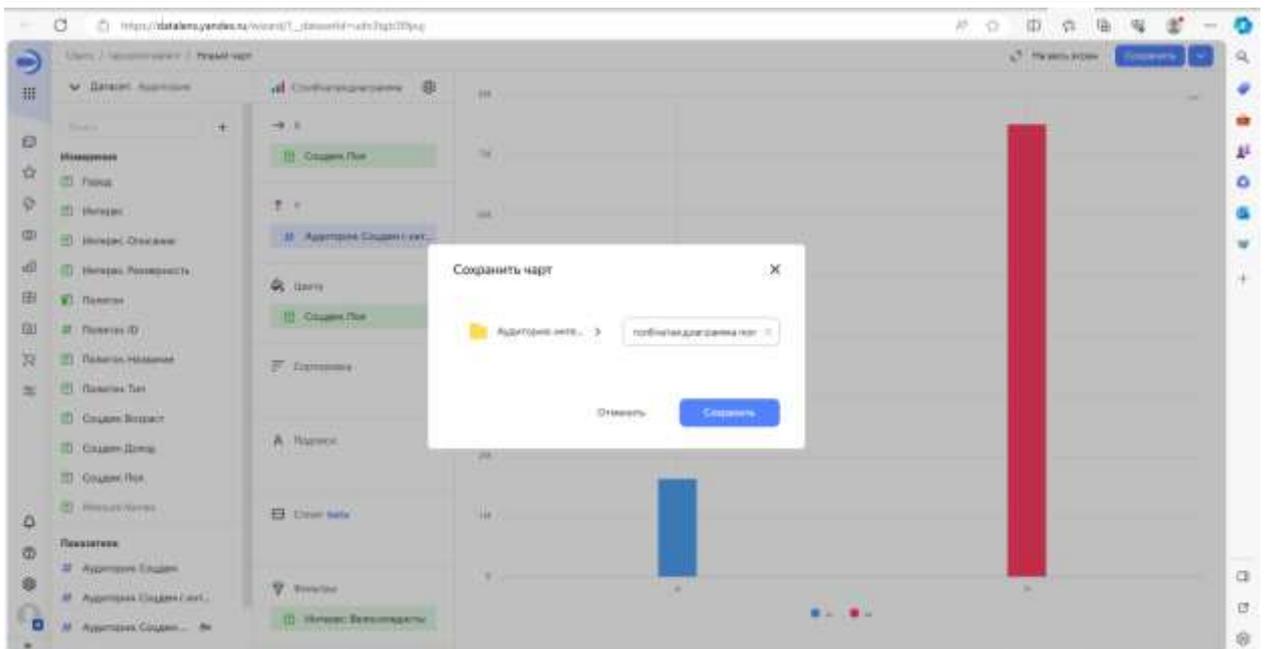


Рисунок 17-Сохранение диаграммы

Далее аналогичным образом создать диаграмму для анализа возрастных категорий велосипедистов. Зайти в навигационное меню, выбрать вкладку чарты, и создать новый чарт. (рис.18).

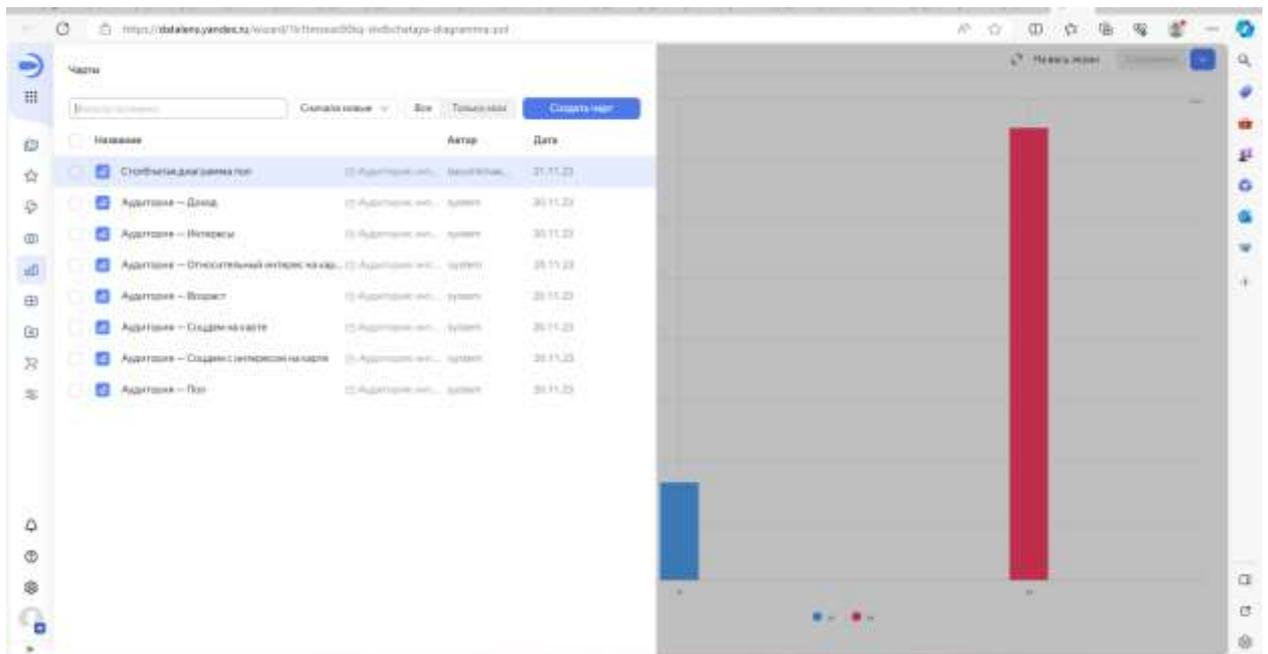


Рисунок 18- Создание чарта

Далее выбрать датасет, выбрать вкладку все объекты далее Marketplace и так как ранее скачивание было с Marketplace то чарты автоматически высвечиваются, перейти в Аудитория и соцдем и нажать на датасет-Аудитория (рис.19).

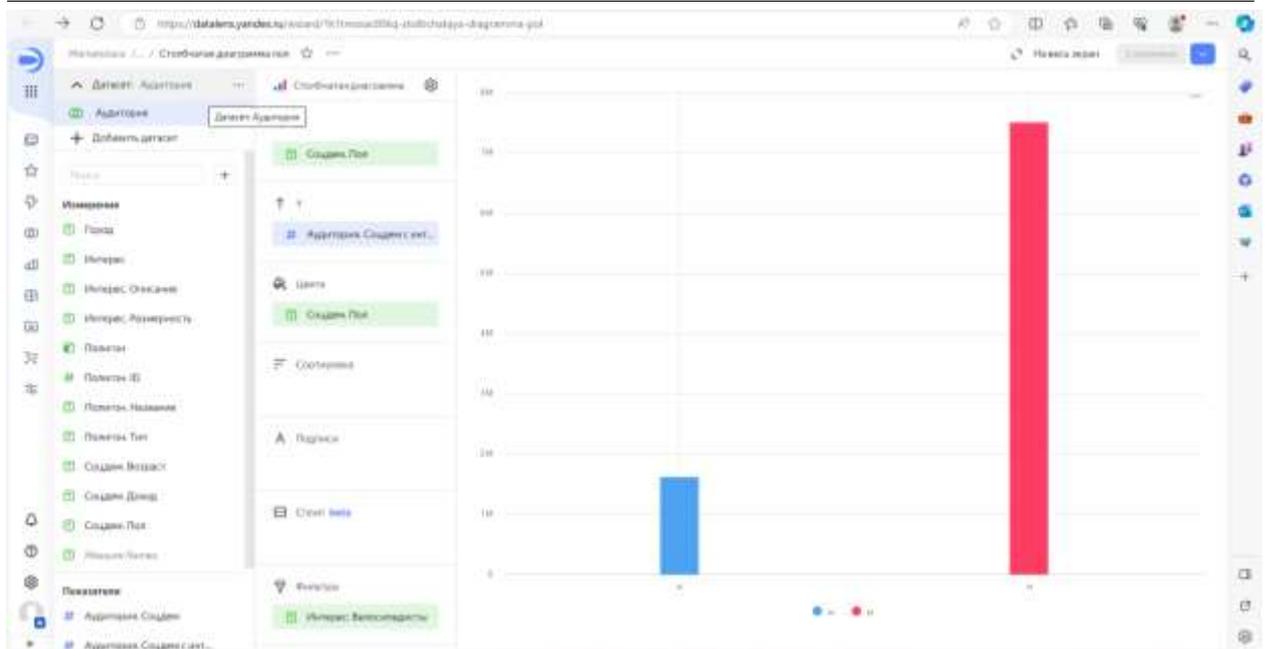


Рисунок 19- Выбор датасета аудитория

В левой части отображаются автоматически все измерения и показатели из данного датасета, по горизонтали оси ОХ вынести возраст а по вертикали оси ОУ сумму общей Аудитории:соцдем и интерес также добавить на сортировку цвета измерение-соцдем возраст. Далее в вкладке фильтр нажать плюс и выбрать интерес и принадлежит множеству велосипедисты, видно что диаграмма настраивается автоматически. Сохранить диаграмму назвав столбчатая диаграмма-возраст (рис. 20).

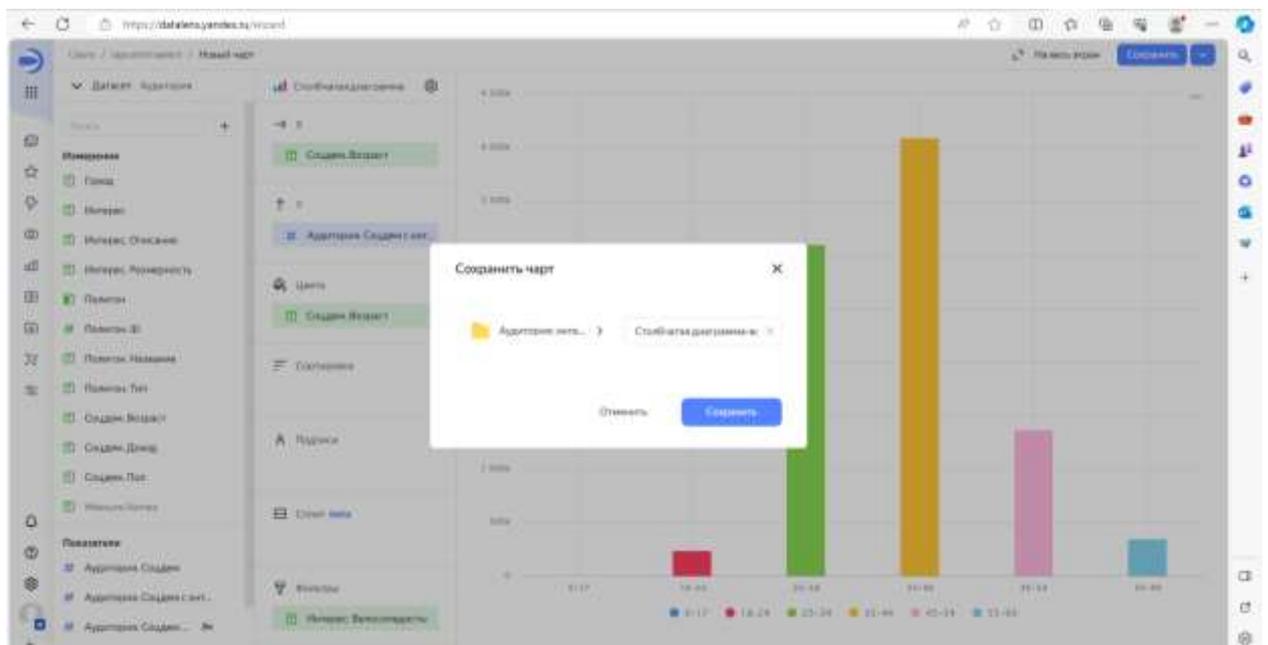


Рисунок 20- Создание столбчатой диаграммы возраст

Создать новую диаграмму аналогичным образом выбрав также датасет Аудитория, далее создать столбчатую диаграмму которая будет

отображать доходность категории от А до С2 (где А-это низкий доход, С2-значение выше среднего) аналогично категорию доход перенести на ось ОХ
Аудитория: соцдем интерес перенести на ось ОУ, добавить краски дашборду перенеся категорию соцдем.доход на вкладку цвета, далее выбрать фильтр интерес, принадлежит множеству велосипедисты, сохранить фильтр назвав его столбчатая диаграмма-доходы (рис.21).

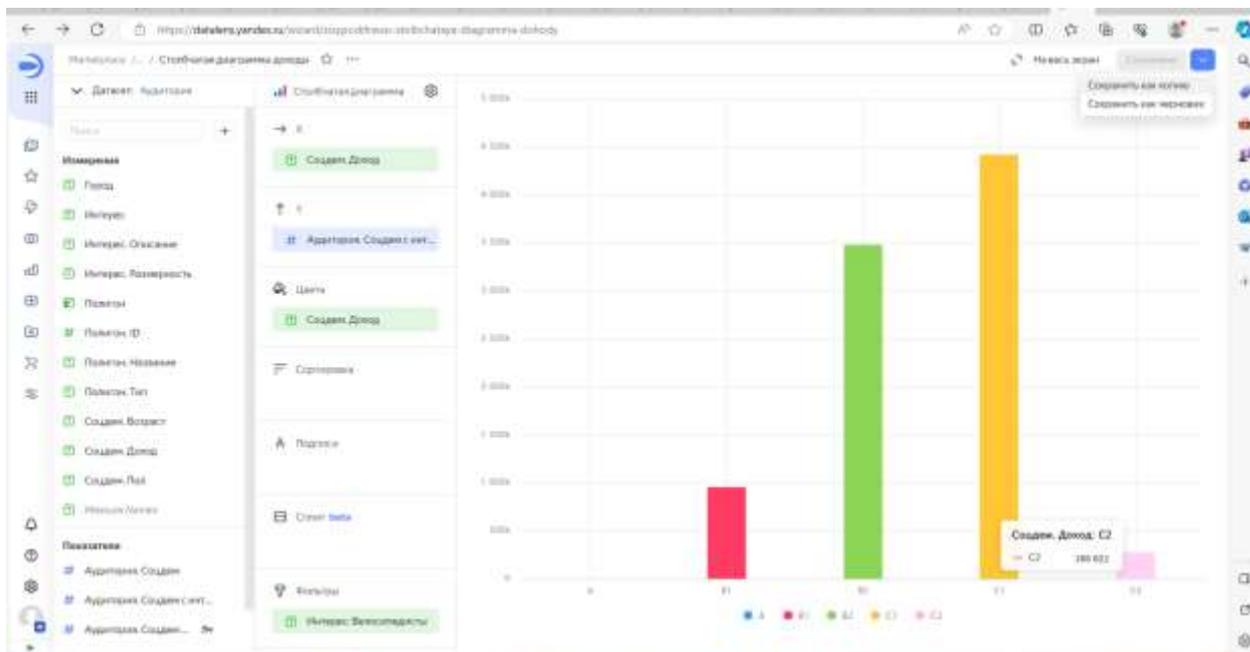


Рисунок 21-Создание столбчатой диаграммы-доходы

На следующем этапе нужно создать три индикатора, в навигационном меню выбрать чарты-создать чарт, среди диаграмм выбираем индикатор и также выбрать датасет- Аудитория. Впервые индикатор будет показывать численность аудитории, перенести категорию Аудитория:соцдем в вкладку показатель, и видно что автоматически назначаются оператор суммы, по индивидуальному усмотрению размер можно изменить а также цвет суммы. Далее нужно назначить фильтр выбрать интерес принадлежит множеству велосипедисты и применить фильтр сохранить под названием индикатор1 (рис.22).

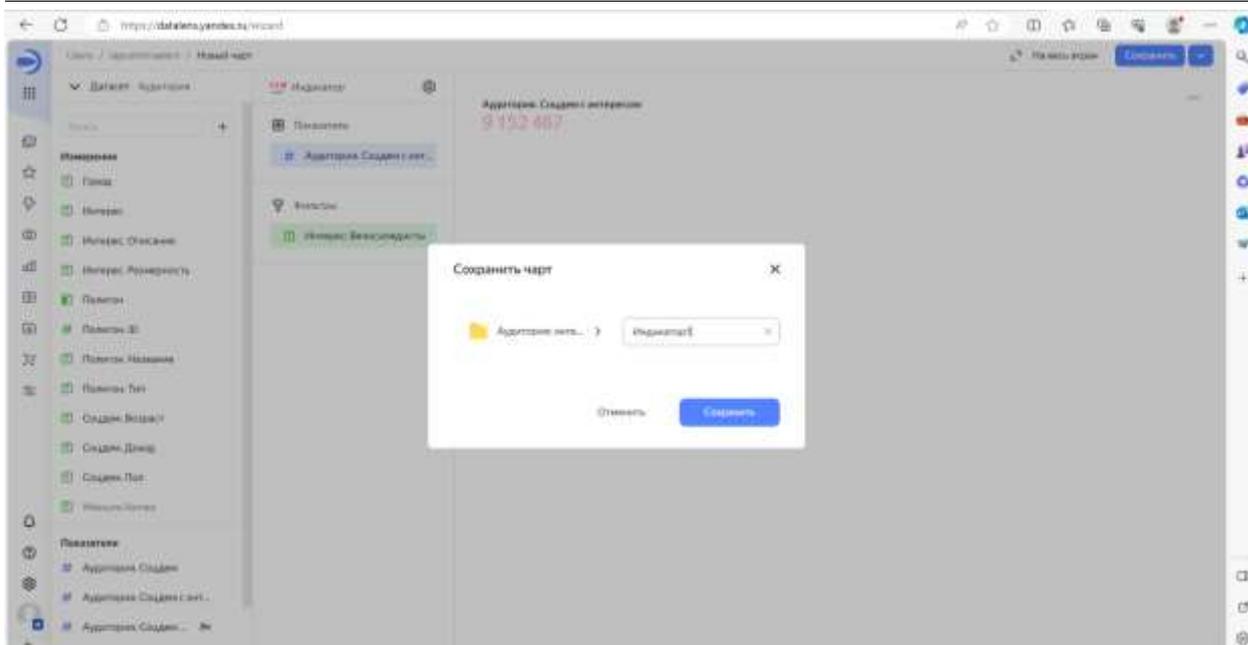


Рисунок 22-Создание индикатор численности

Следующие два чарта создаются аналогичным образом. Индикатор 2 показывает сумму по аудитории с интересом (рис.23). Индикатор 3 считает показатель относительно заинтересованных к общему числу аудитории (рис.24), ранее созданные анализы способствуют анализу общих показателей аудитории селекции по городам России, а последняя визуализация по счету а не по важности- это карта, по которой можно проводить более детальную аналитику по участкам или полигонам разного размера внутри города.

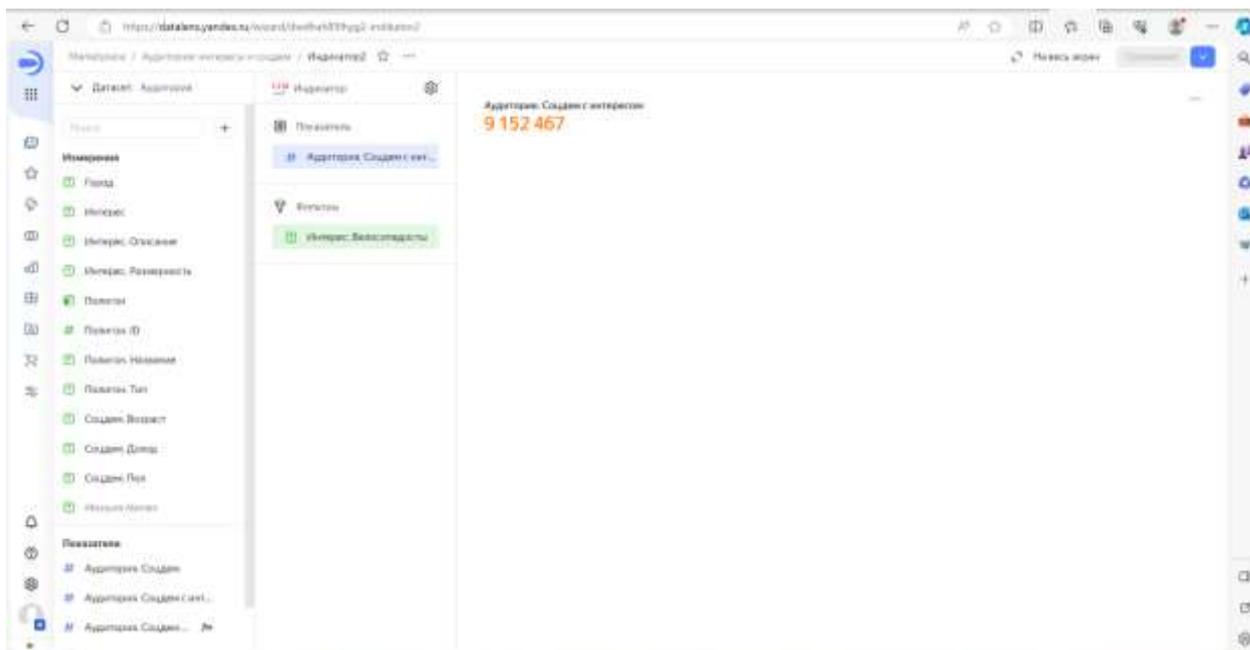


Рисунок 23- Создание индикатор суммы аудитории с интересом

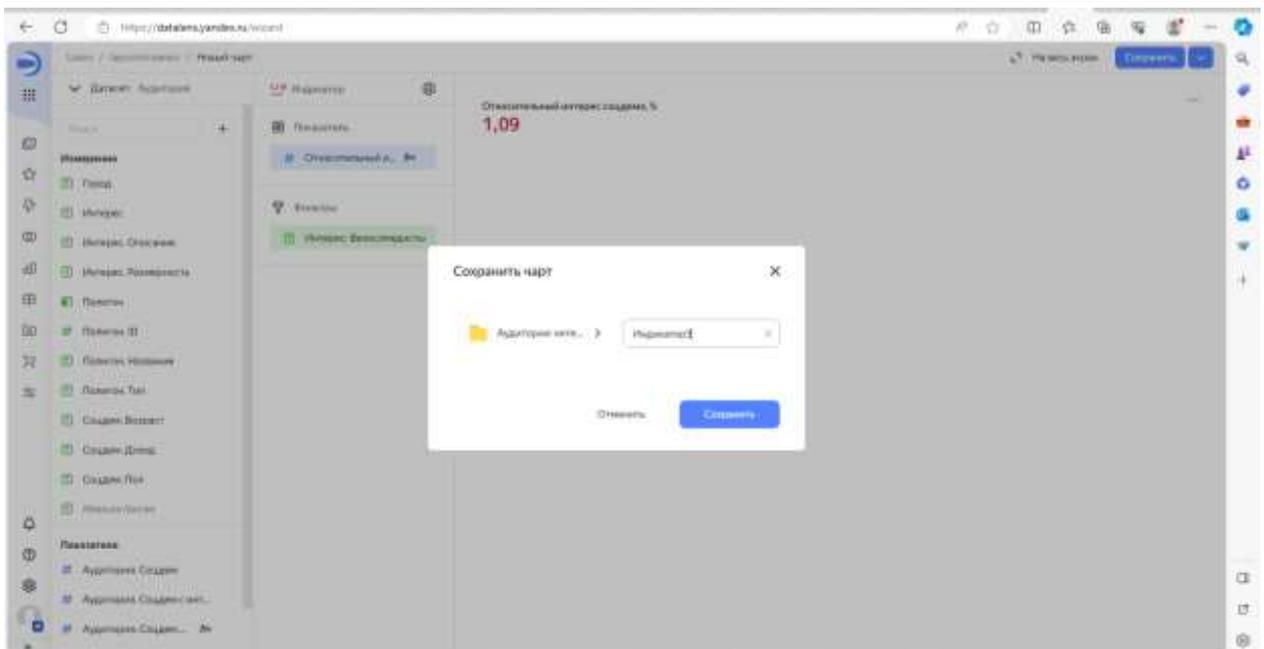


Рисунок 24- Создание индикатор показатель заинтересованных к общему числу аудитории

Далее нужно создать новый чарт, датасет выбрать Аудитория, в тип диаграмма выбрать карта, что бы на карте отображались показатели именно на участках, точнее на полигонах необходимо выбрать вкладку точки не точки а полигоны либо геополигоны, перенести соответствующее поле в окошко (рис. 25).

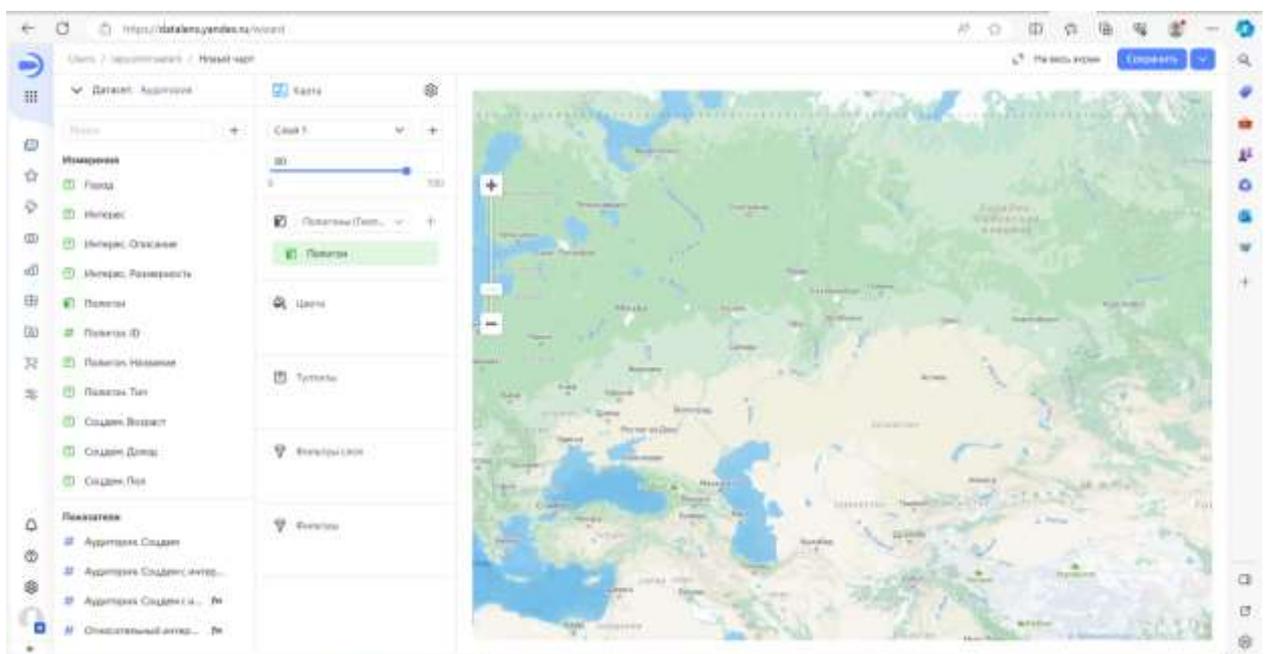


Рисунок 25- создание чарта полигон

Далее нужно сделать постройку цвета для отображения динамики популярности проката, для этого перенести показатель аудитории с

интересом на поле цвета, нажать на шестеренку выбрать цвет на свое усмотрение (рис. 26).

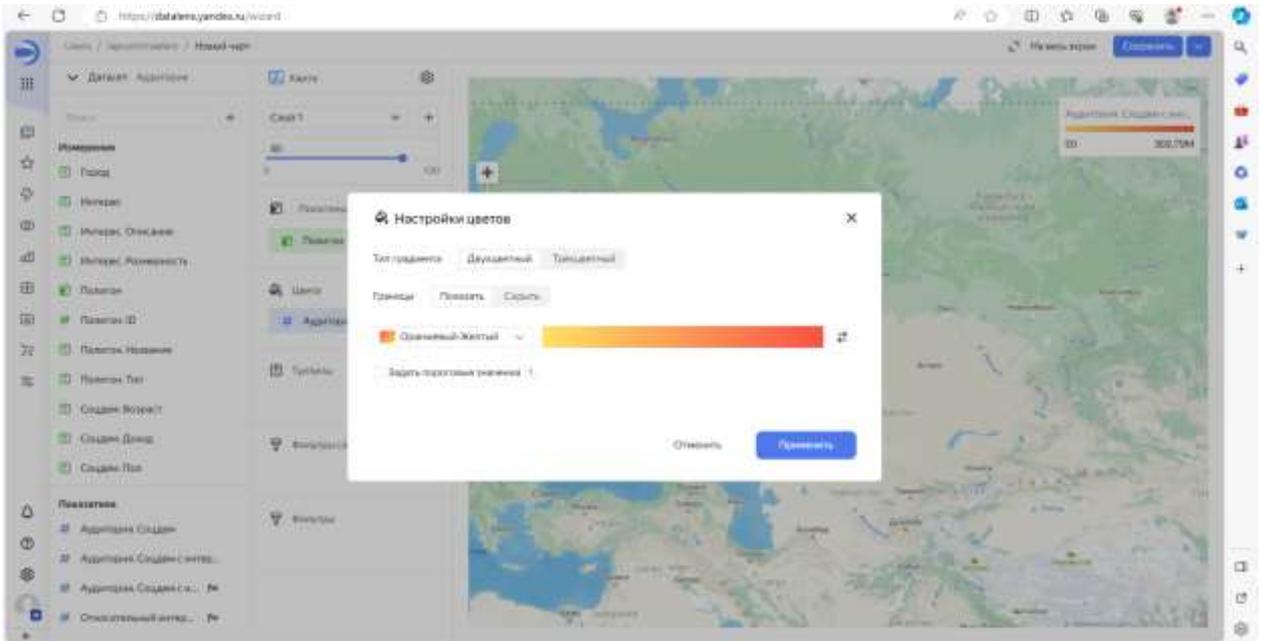


Рисунок 26- Выбор цвета

Увеличив карту с помощью показателя, можно разглядеть изменения на пример в Москве. Для этого нужно настроить фильтр, указав группу по интересам принадлежавшим велосипедистам и добавить фильтр по типу полигона. На полигоне крупные участки для наглядности это-hash6, а hash7 это- более мелкие (рис. 27).

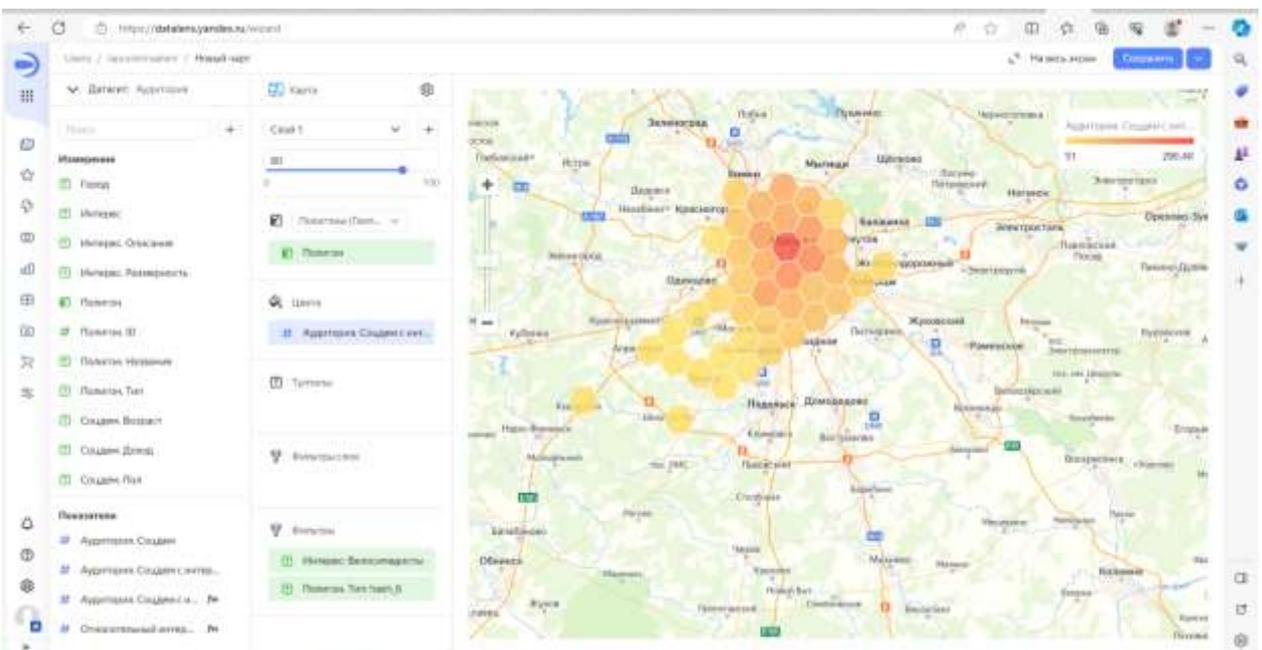


Рисунок 27- Настройка фильтра карты

В данной карточке необходимо настроить тултипы, нажав плюс добавить показатель Аудитория. Соцдем и интерес, Аудитория. Соцдем чтобы увидеть просто общую численность, относительный интерес соцдем и сам интерес, сохранить под названием аудитория карта (рис. 28).

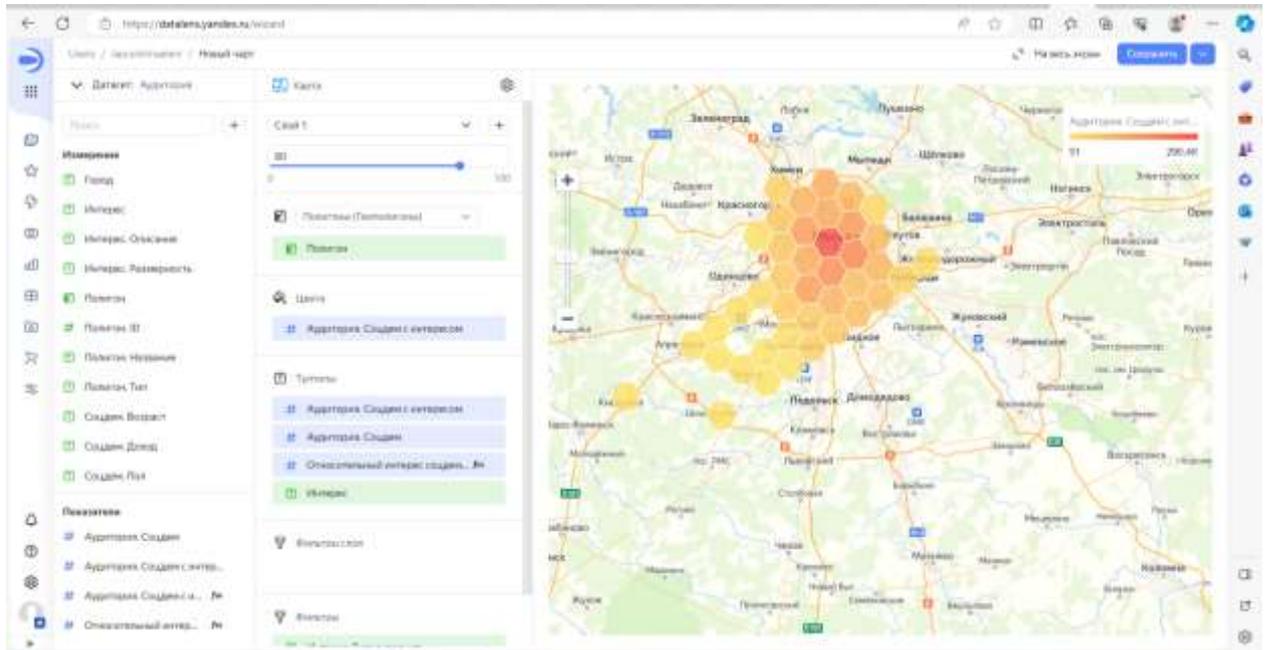


Рисунок 28-Редактирование вкладки тултипы

Заключительный этап создания отчетного дашборда. Необходимо нажать Marketplace, выбрать датасет- Аудитория, в правом углу экрана нажать кнопку создать дашборд (рис. 29).

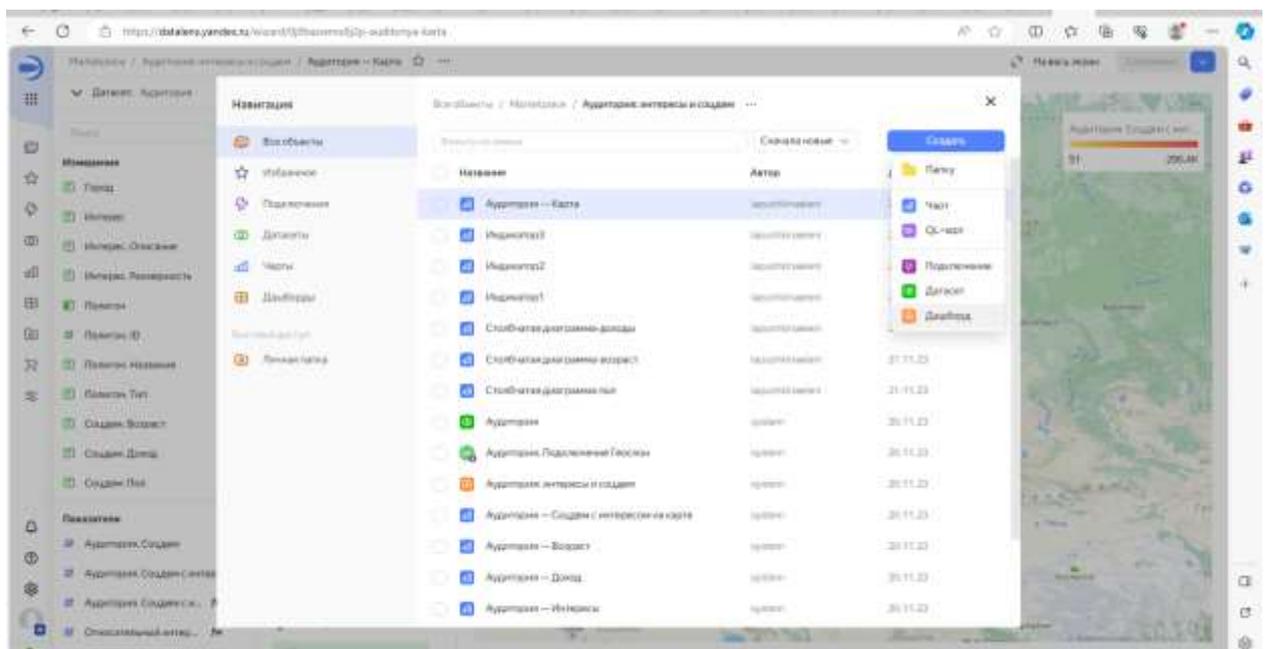


Рисунок 29- Создание дашборда

Далее в новом дашборде нажать кнопку чарт, и последовательно вынести на основную вкладку все ранее сохраненные чарты (рис.30).

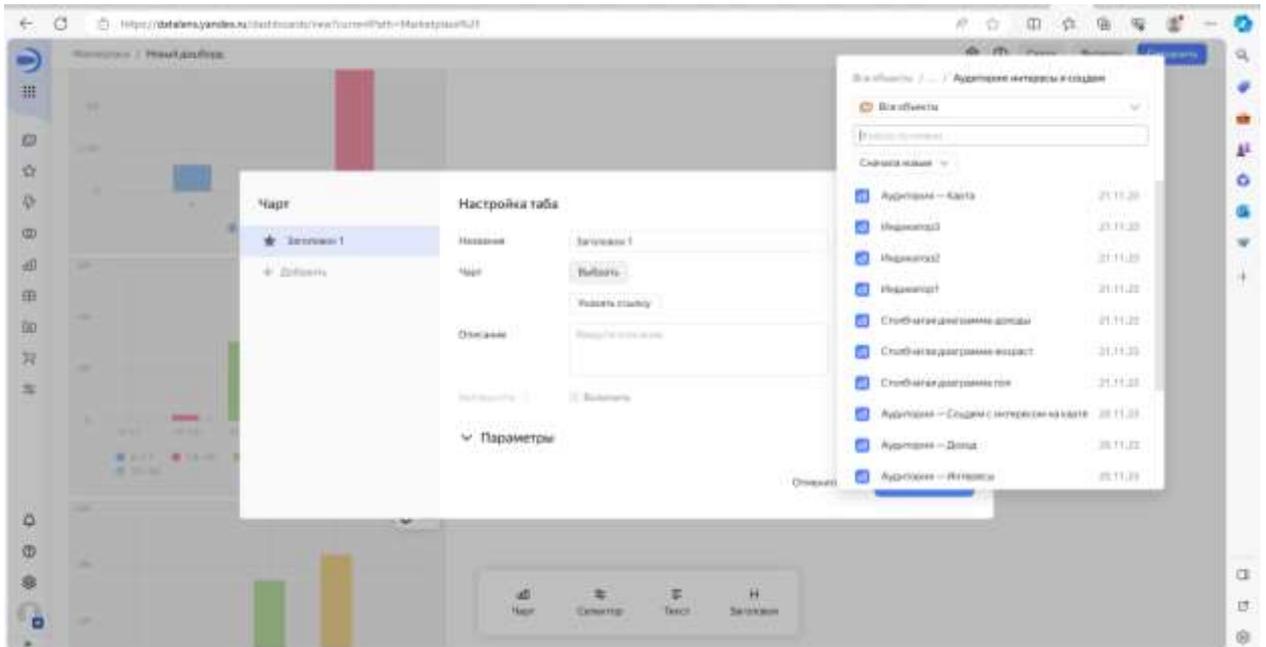


Рисунок 30- Вынос чартов на дашборд

Необходимо настроить селектор, для этого перейти на вкладку селектор настроив его, выбрать датасет- Аудитория, в вкладке поле нужно выбрать три чарта это- пол, доход и возраст, добавить множественный выбор, нажать кнопку добавить, для удобства лучше разместить селектор над диаграммой (рис. 31).

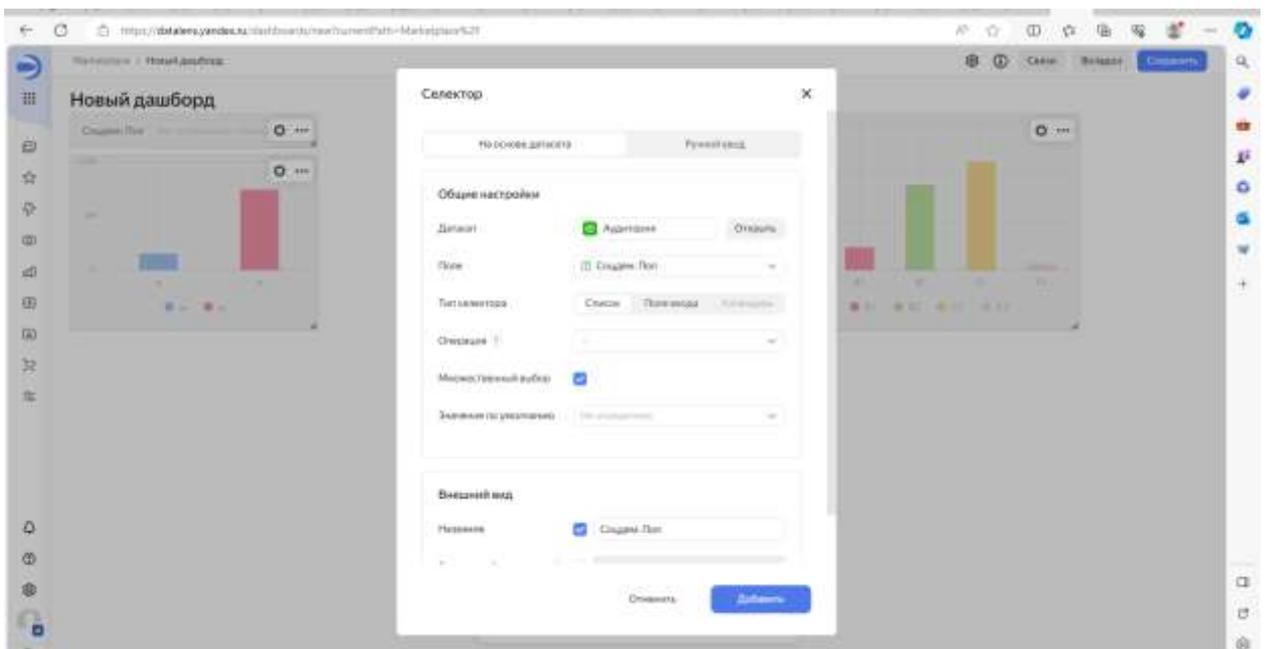


Рисунок 31- Настройка селектора

Далее вынести индикаторы основных показателей, которые создавались ранее, это- индикатор1, индикатор2, индикатор3 (рис.32).

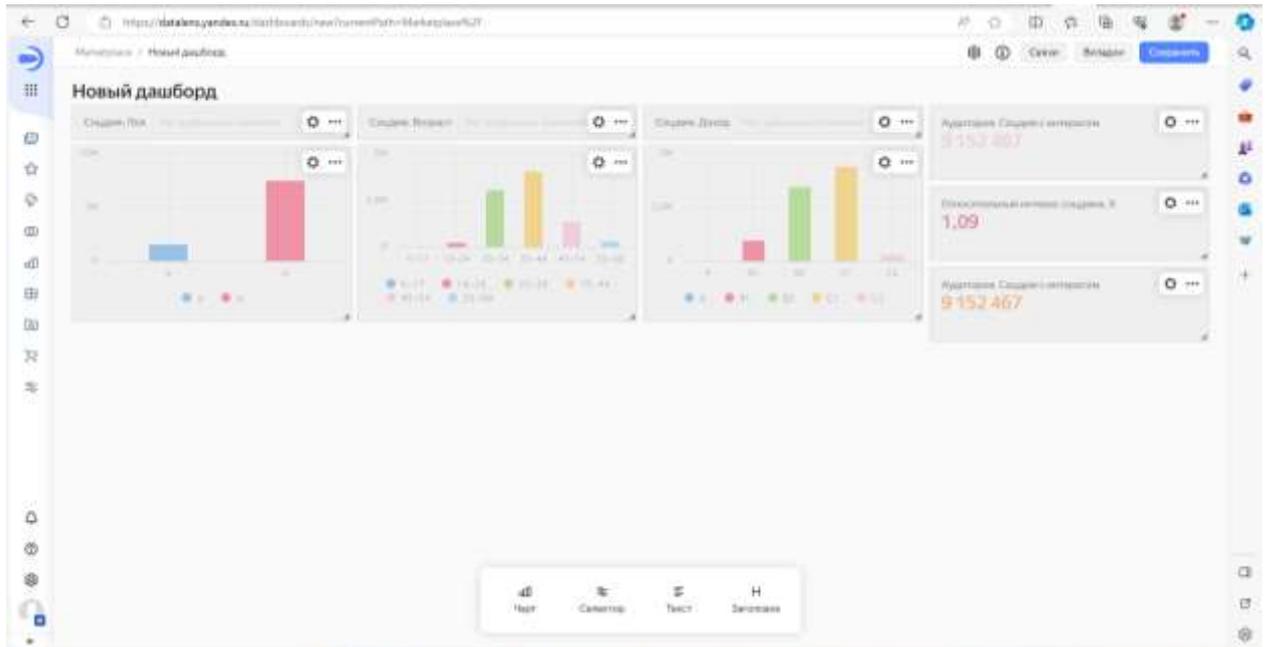


Рисунок 32- Добавление основных индикаторов в дашборд

В дашборд нужно так же добавить селектор по городам, селектор по полигонам (рис. 32).

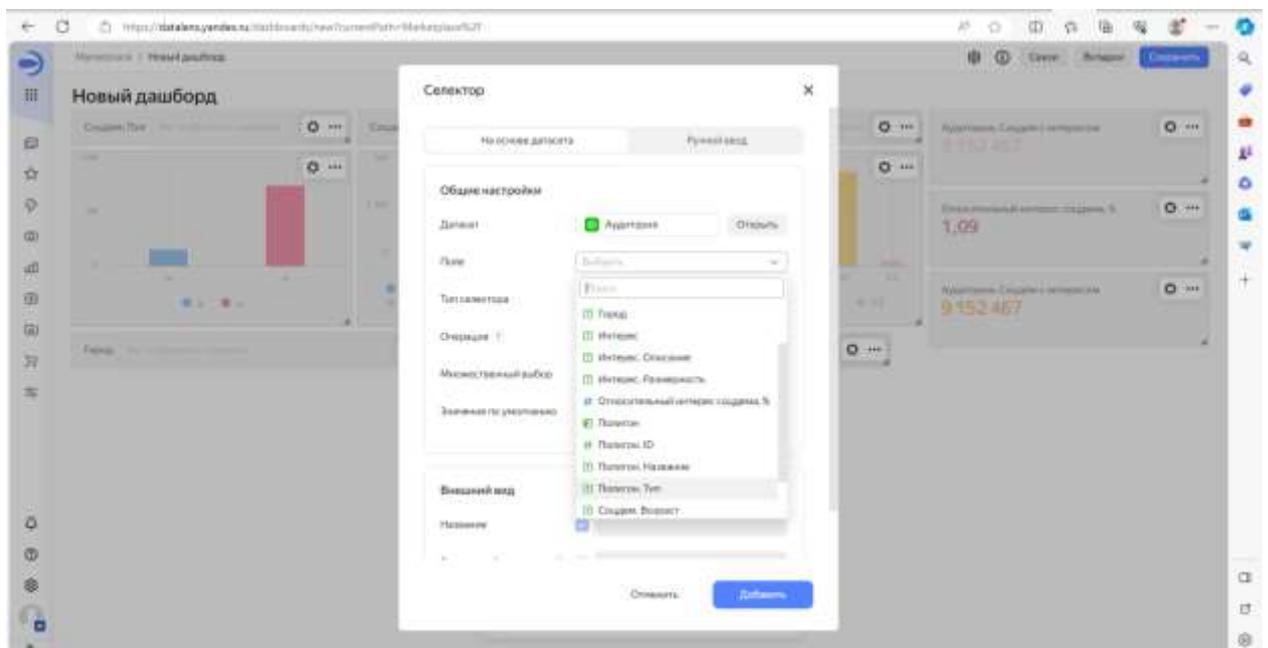


Рисунок 32- Добавление селектора город, полигон

Финальный этап это- выведение визуализации карты, нажать кнопку добавить чарт, выбрав Аудитория-карта, переименовав ее для большего удобства- в Исследование динамики, дашборд готов к тестированию (рис. 33).

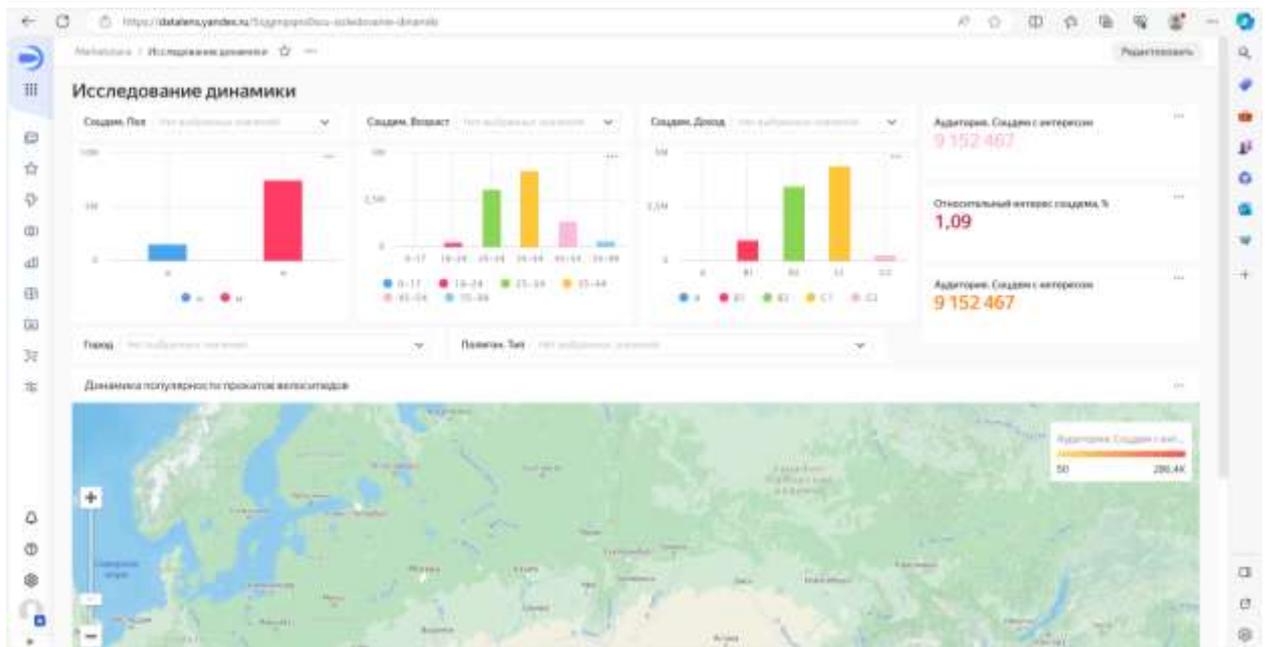


Рисунок 33- Тестирование дашборда

Для тестирования выбрав все показатели можно увидеть, что с каждым дополнительным выбором селекторов перестраиваются все визуализации, поэтому можно проводить более детальный анализ. Можно провести анализ популярности проката велосипедов, негатив на карте отображает увеличение числа заинтересованных людей в использовании проката для велопрогулок, красный индикатор говорит о том, что центральная часть города является самой популярной, но и стоит отметить что в удаленных участках тоже есть насыщенные цвета (рис. 34).

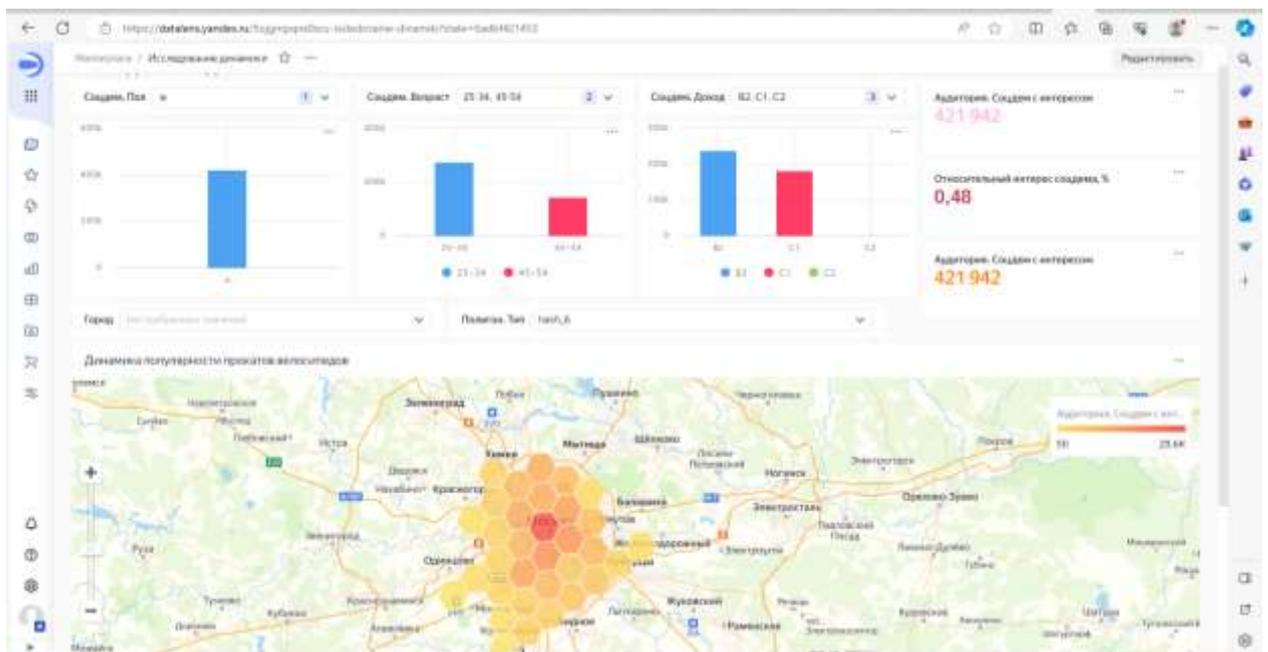


Рисунок 34- Проведение анализа популярности проката велосипедов в городе Москва

4 Выводы

В данном исследовании был создан дашборд в Yandex DataLens, с помощью которого можно сделать визуализацию любого города на предмет проката велосипедов.

Библиографический список

1. Рыжкина Д. А. Дашборд. Полезные советы //Аллея науки. 2017. Т. 1. №. 9. С. 341-343.
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.
3. Есенкина В.С. Визуализация данных как инструмент анализа и обработки больших объемов информации. // Безопасность. Управление. Искусственный интеллект. 2022. №2(2). С. 13-18.
4. Марголина Н. Л., Ширяев К. Е. Построение графиков функций в свете формирования исследовательских навыков //Образовательная деятельность вуза в современных условиях. 2016. С. 21-21.
5. Ростовцев П. С., Салех Х. М. Система генерации датасет //Современная школа России. Вопросы модернизации. 2021. №. 2-2. С. 115-117.
6. Young G. W., Kitchin R. Creating design guidelines for building city dashboards from a user's perspectives //International Journal of Human-Computer Studies. 2020. Т. 140. С. 102429.
7. Сяо Ю., Мяо К., Хуан Ю. Влияние графического кодирования на эффективность чтения диаграммы данных, отображаемой на периферии внимания //Дисплей. 2023. Т.77. С.102378.
8. Хасенова З. Т. Создание тепловой карты для визуализации данных //Фундаментальные и прикладные исследования в современном мире. – 2017. – №. 20-1. – С. 24-26.
9. Смирнов А. В. Опыт разработки демографических дашбордов //Демографические факторы адаптации населения к глобальным социально-экономическим вызовам. 2023. С. 550-559.