



> Конспект > 5 урок > ВИЗУАЛИЗАЦИЯ

Альтернативные инструменты для визуализации

[Plot.ly](#)

[DataWrapper](#)

[Flourish](#)

[DataLens](#)

[Power BI](#)

[Qlik](#)

[Fine BI](#)

[D3.js](#)

[DataLens: подключение](#)

[DataLens: создание графиков](#)

[DataLens: организация дашборда](#)

[Полезные ссылки по DataLens](#)

[Superset: основной функционал](#)

[Superset: строим дашборд](#)

Альтернативные инструменты для визуализации

В рамках этого блока мы с вами изучали, как подходить к визуализации данных в Tableau. Однако Tableau - не единственное место, где можно делать красивые визуализации и строить на их основе дашборды! В зависимости от ваших целей, возможностей и предпочтений вы можете в итоге работать с совсем другим инструментом.

Это не означает, что знания этого блока пропадут - многие высказанные в нём идеи универсальны и остаются правдой вне зависимости от используемого инструмента. Тем не менее, небольшой обзор альтернатив может оказаться вам полезен.

Plot.ly

- Изначально библиотека для визуализации интерактивных графиков, реализованная на нескольких языках
- Можно строить графики прямо в браузере без написания кода через Chart Studio
- Если используете Chart Studio, то можно вдобавок к графику получить код, которым можно создать этот график в каком-либо из языков программирования
- Скорее заточено под визуализацию отдельных графиков, нежели создание полноценного дашборда (однако обратите внимание на Dash)

DataWrapper

- Инструмент для медийных изданий - чаще всего им пользуются дата-журналисты
- Очень красивый дизайн графиков, можно брать в качестве вдохновения для своих работ
- Шаблонов графиков не очень много
- Есть крутой блог про визуализацию

Flourish

- Другой инструмент для дата-журналистов и блогеров
- Большой акцент на интерактивность и подвижность в графиках - особенно в народе известны bar chart races
- Визуальный стиль не такой классный, как у DataWrapper

DataLens

BI-инструмент от Яндекса. По понятным причинам распространён в России, подробнее о нём - в других частях конспекта 😊

Power BI

- BI-система от Microsoft, среди коммерческих предложений - лидер на рынке
- Отличная интеграция с экосистемой Windows
- Несколько более шаблонная визуализация, чем в Tableau, но в целом довольно гибкий инструмент
- Имеет тенденцию выдавать не очень аккуратный дизайн по умолчанию и давать пользователям возможность сделать его ещё хуже
- Нет разделения на графики и дашборды, всё в одном месте - в результате график из одного дашборда сложно кинуть в другой
- Есть дополнительный магазин визуализаций + возможность создавать свои через JavaScript
- Полноценный инструмент для предобработки данных, гораздо более мощный по сравнению с Tableau (хотя в Tableau есть инструмент под названием Prep)

Qlik

- Есть в двух версиях - Qlik Sense для пользовательской работы в вебе и QlikView как корпоративное решение
- Фактически есть свой скриптовый язык, что даёт большую гибкость в создании кастомных визуализаций
- По причине выше имеет высокий порог вхождения, и часто под работу с Qlik выделяются отдельные рабочие должности
- Очень популярен в банках и ритейле

Fine BI

- Китайский аналог Tableau

- В демонстрациях очень причудливые дизайнерские решения
- Тем не менее, реально очень гибкий инструмент - имеет как возможность строить по шаблонам, так и увязывать отдельные визуальные элементы с какими-то параметрами
- Есть сквозное оформление, позволяющее настраивать стиль дашборда целиком

D3.js

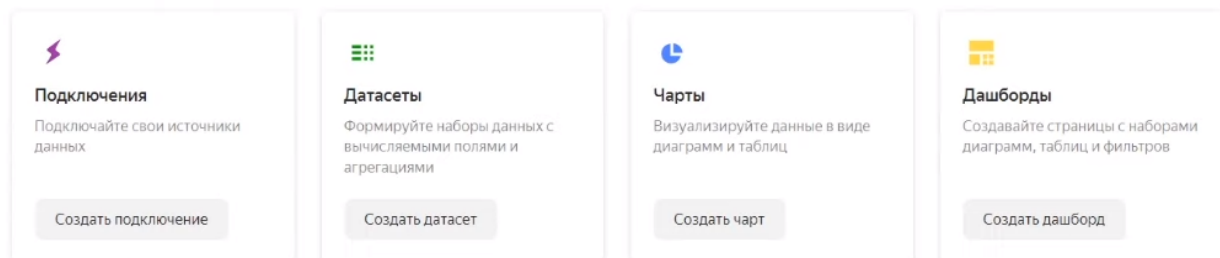
- Библиотека для визуализации данных на JavaScript
- Очень гибкая, можно управлять абсолютно любыми атрибутами графика
- Так как это JS, очень хорошо и плавно интегрируется в веб
- Больше не для аналитики, а для представления информации для конечного пользователя в вебе

Ещё Рома рекомендует посмотреть блог [LEFT JOIN](#), в нём также очень хорошо разбираются разные BI-инструменты.

DataLens: подключение

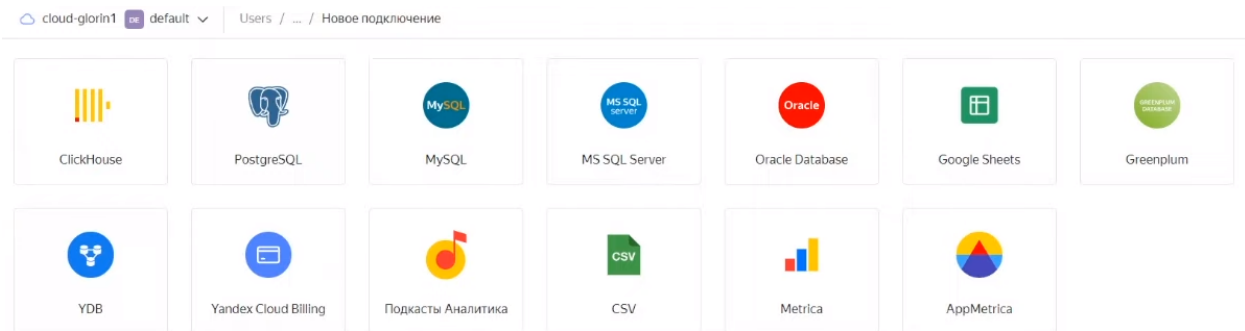
При заходе в DataLens вы увидите такой экран:

Yandex DataLens

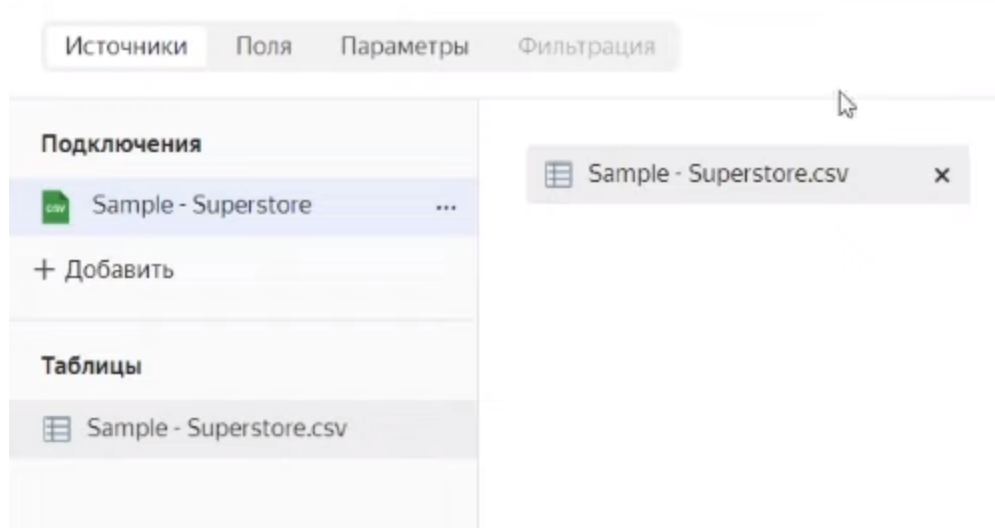


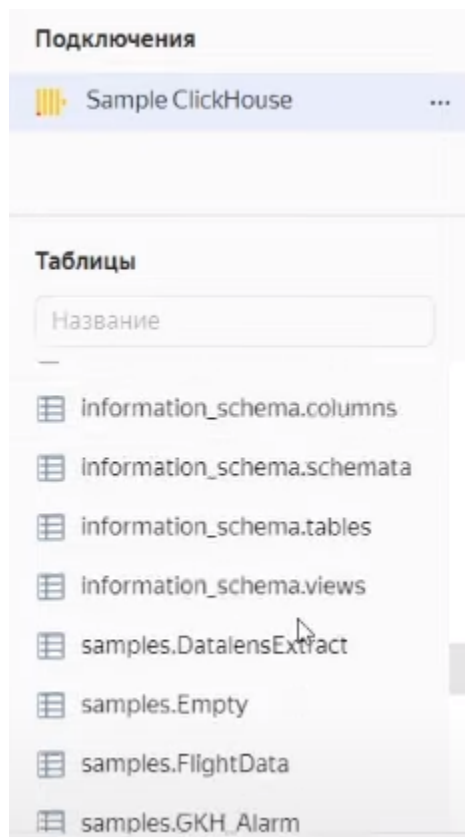
Первое, что нам надо сделать - это **подключиться** к какому-то источнику данных (крайняя левая сущность). Обычно это СУБД, но также подходят обычные csv-

файлы, а также есть некоторые совсем нестандартные опции:



После подключения вы создаёте некоторый **датасет** - набор данных. с которым вы и будете работать в DataLens. Для csv-файла он по определению будет всего один, для баз данных набор будет больше:

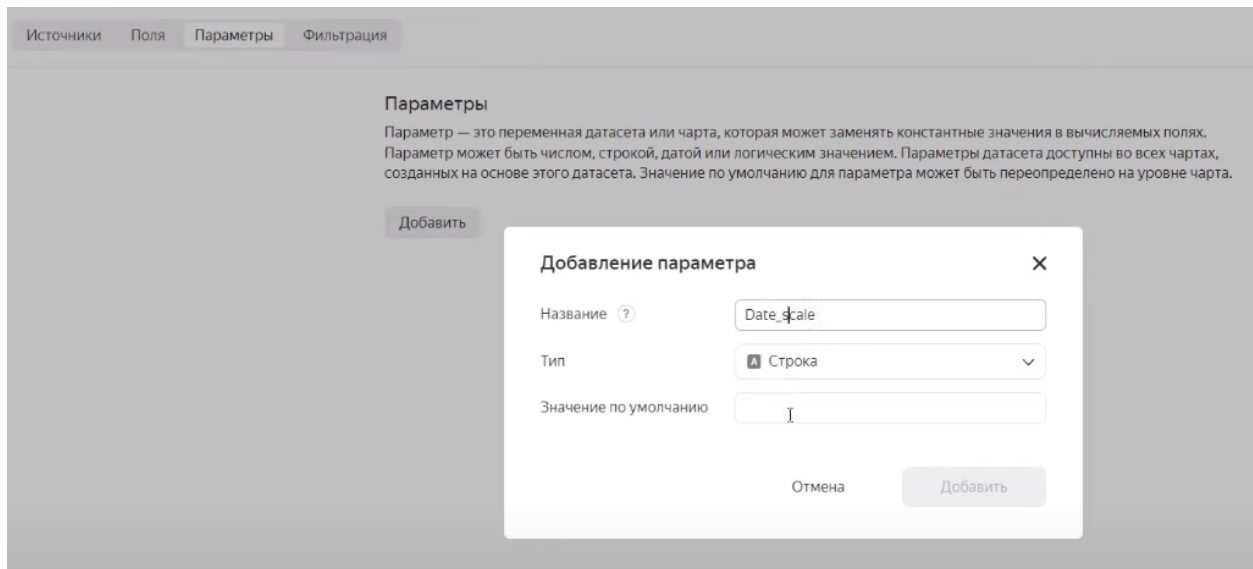




В этом же разделе можно проводить стандартные операции с таблицами - в частности, джойны:

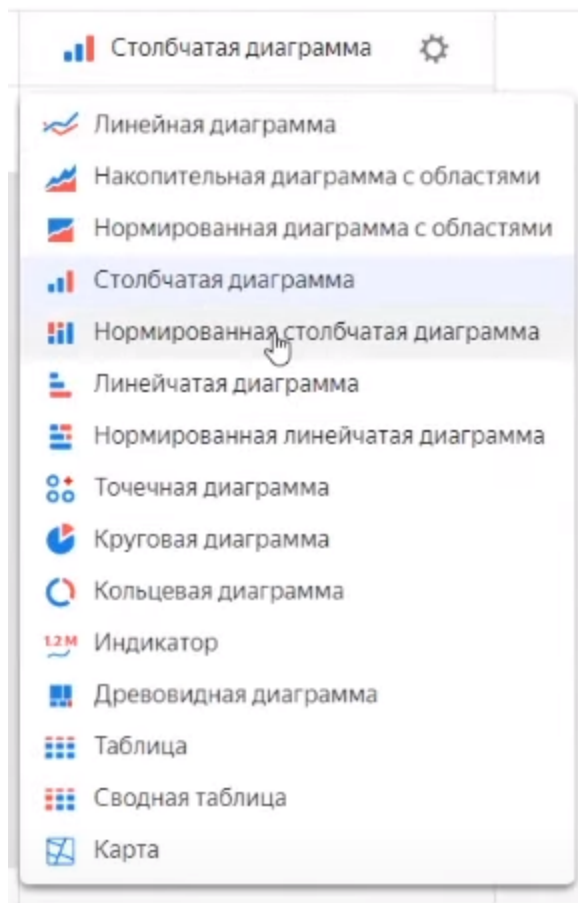


В разделе “Поля” можно задать некоторые параметры тех столбцов таблицы, с которыми мы будем работать. В частности, можно задать тип столбца, а также заранее задать агрегирующую функцию:

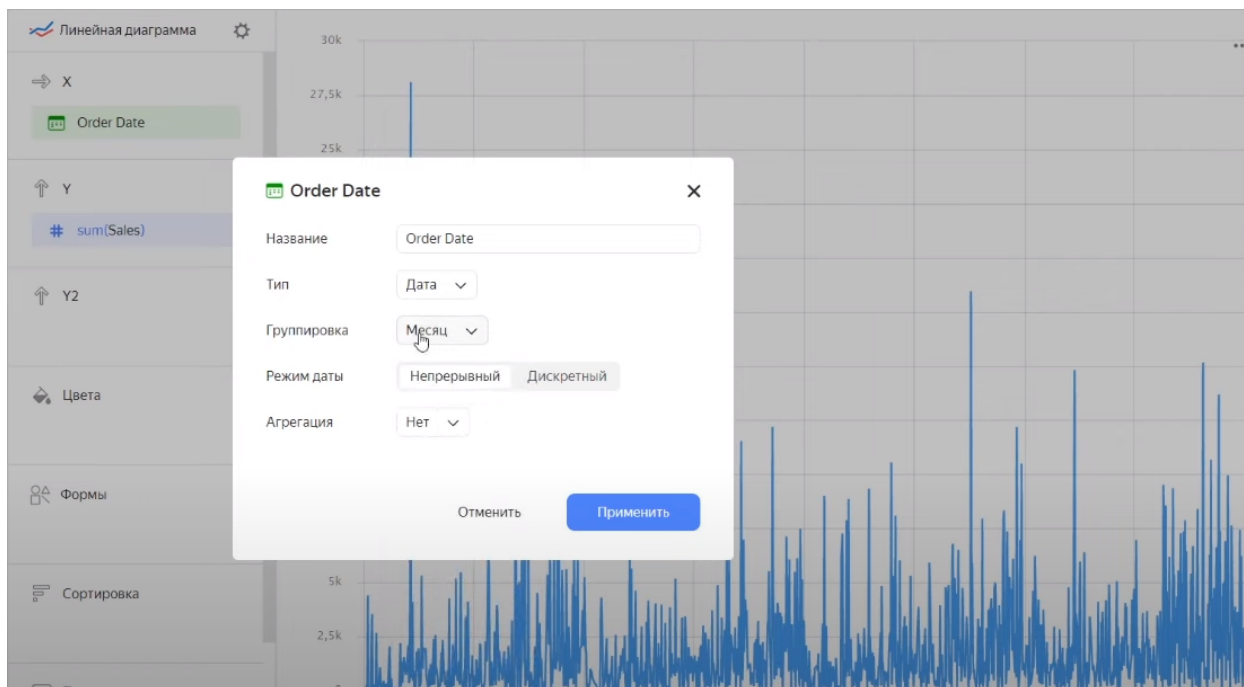


DataLens: создание графиков

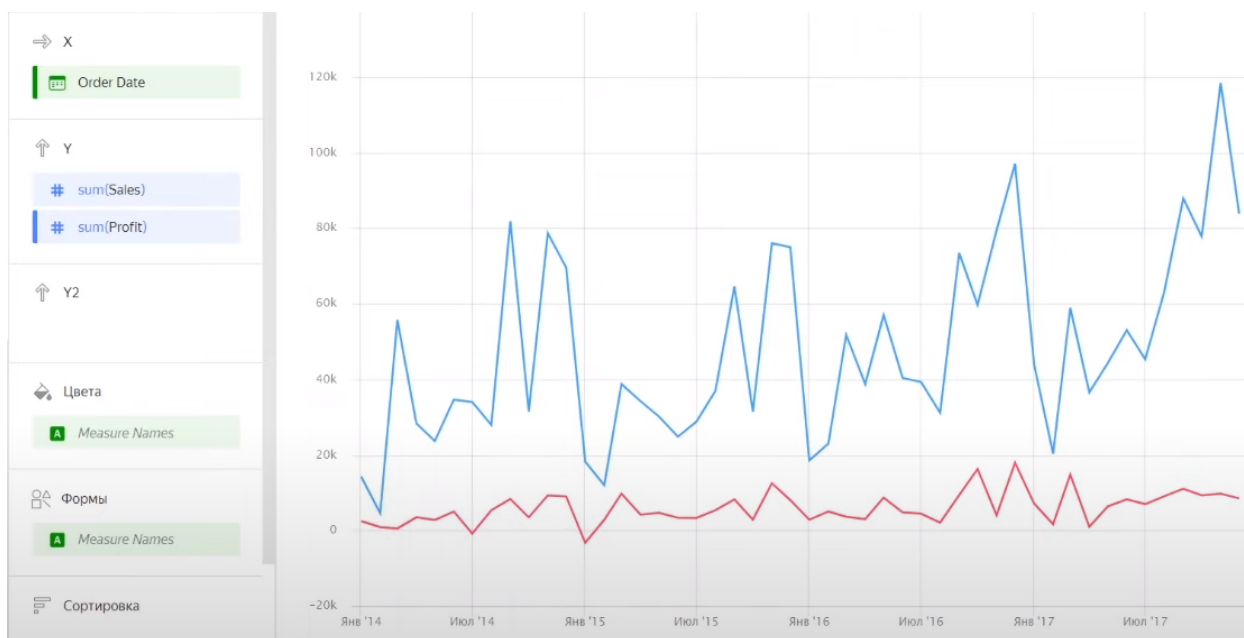
DataLens использует “шаблонный” подход к визуализации - то есть мы должны заранее выбрать какой-то тип графика, а уже потом подгонять наши данные под готовые параметры:



Построение графика происходит достаточно просто - мы перетягиваем нужные поля в те элементы шаблона, которые предоставляет данный тип графика. Сами переменные можно дополнительно настраивать:



В один и тот же раздел шаблона можно добавить сразу несколько полей - например, в лекции таким образом рисовалось несколько линейных диаграмм в одной системе координат:

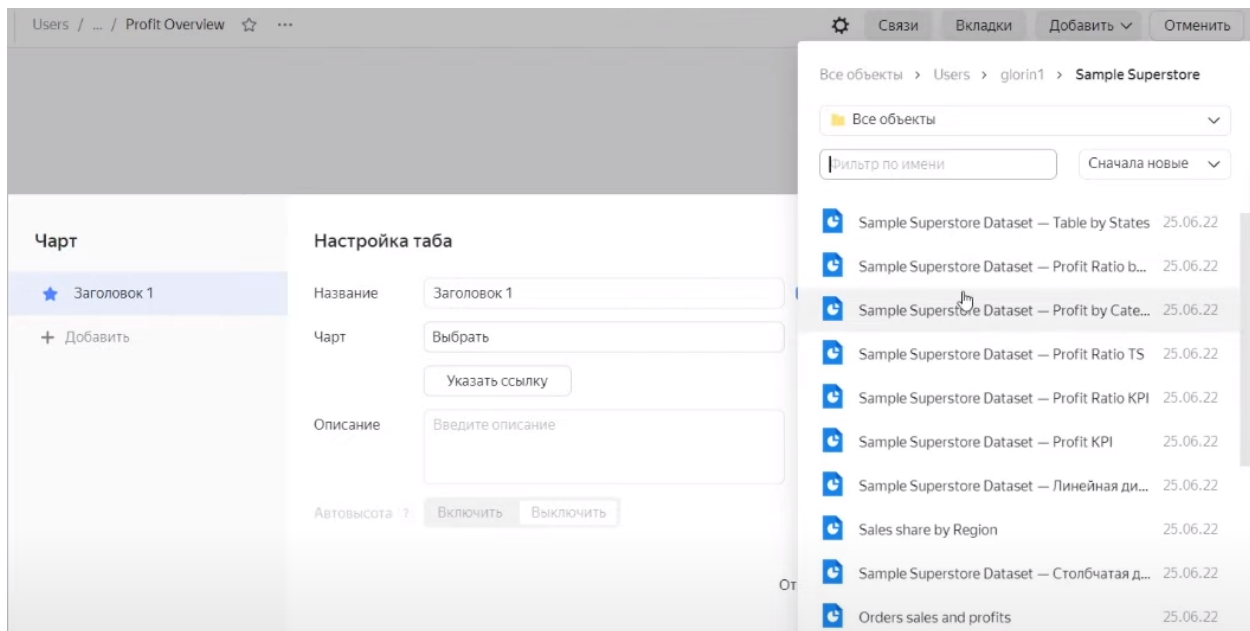


Полноценную демонстрацию построения графиков лучше наблюдать своими глазами в лекции, так как письменный конспект довольно плохо передаёт процесс

конструирования визуализации 😊 Это то, что нужно видеть в динамике своими глазами.

DataLens: организация дашборда

При создании дашборда ваш главный друг - кнопка “Добавить”. Оттуда вы выбираете диаграмму из того набора, который вы заранее создали. Опционально можно задать заголовок и описание:



С элементами дашборда работать достаточно просто - они очень интерактивно растягиваются и перетаскиваются простым нажатием кнопок мыши:

Profit Overview



В одно и то же место можно вставить несколько графиков - тогда они будут находиться в разных вкладках, и их можно будет переключать:

Чарт

Sample Superstore Data...

Sample Superstore Data...

+

Добавить

Настройка таба

Название

Sample Superstore Dataset — Table by States and Mo

Чарт

Sample Superstore Dataset — Table by States an...

Открыть

Указать ссылку

Описание

Введите описание

Автовысота ?

Включить

Выключить

Отменить

Добавить

Sample Superstore Dataset — Table by States

Sample Superstore Dataset — Table by States and Months

...

		Profit	Profit Ratio
Central	Illinois	-12 607,89	-15,7%
	Indiana	18 382,94	34,3%
	Iowa	1 183,81	25,8%
	Kansas	836,44	28,7%
	Michigan	24 463,19	32,1%
	Minnesota	10 823,19	36,2%
	Missouri	6 436,21	29,0%
	Nebraska	2 037,09	27,3%
	North Dakota	230,15	25,0%
	Oklahoma	4 853,96	24,7%
	South Dakota	394,83	30,0%

Sample Superstore Dataset — Table by States

Sample Superstore Dataset — Table by States and Months

...

		01.01.2014	01.02.2014	01.03.2014	01.04.2014	01.05.2014	01.06.2014	01.07.2014	01.08.2014
Central	Illinois	-0,23	-0,27	-0,01	0,08	-0,23	-0,44	-0,22	-0,01
	Indiana	0,00	0,36	0,40	0,48		0,05		
	Iowa			0,09					
	Kansas								
	Michigan	0,21		0,48	0,09	0,46	0,42		0,39
	Minnesota			0,30			0,26		
	Missouri		0,20						0,28
	Nebraska							0,28	
	North Dakota								
	Oklahoma						0,33		
	South Dakota	0,29						0,42	

Лайфхак: если вам нужно пустое пространство в дашборде, то можно добавить в качестве элемента текст и ничего там не писать. В результате создастся пустой блок, который всё так же можно передвигать и растягивать:



Можно также добавлять селекторы - фактически фильтры данных, которые позволяют нам смотреть на них в определённом разрезе (диапазон дат или какая-либо страта):

Селектор

✕

На основе датасета

Ручной ввод

Общие настройки

Датасет

Sample Superstore Dataset

Поле

Order Date

Тип селектора

Список

Поле ввода

Календарь

Операция

—

Диапазон

☐

Значение по умолчанию

Не определено

Внешний вид

Название

☒

Order Date

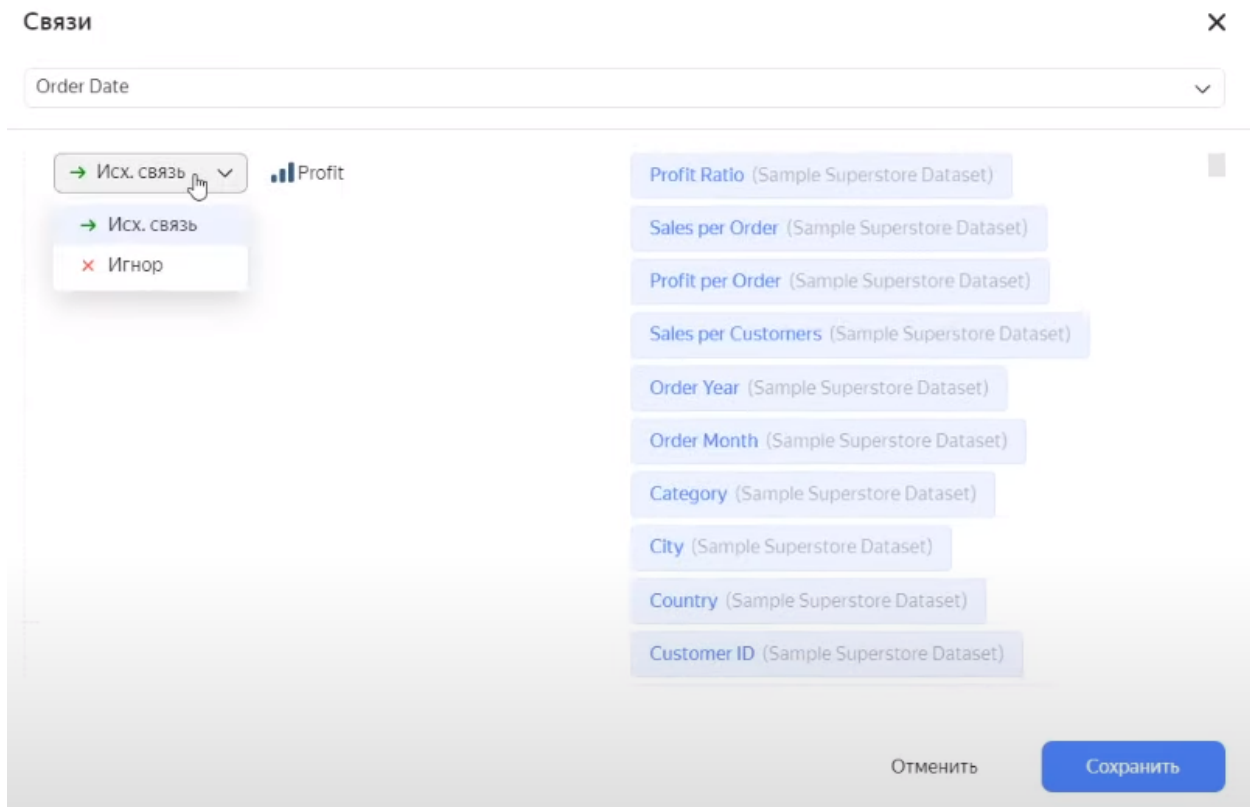
Внутренний заголовок

☐

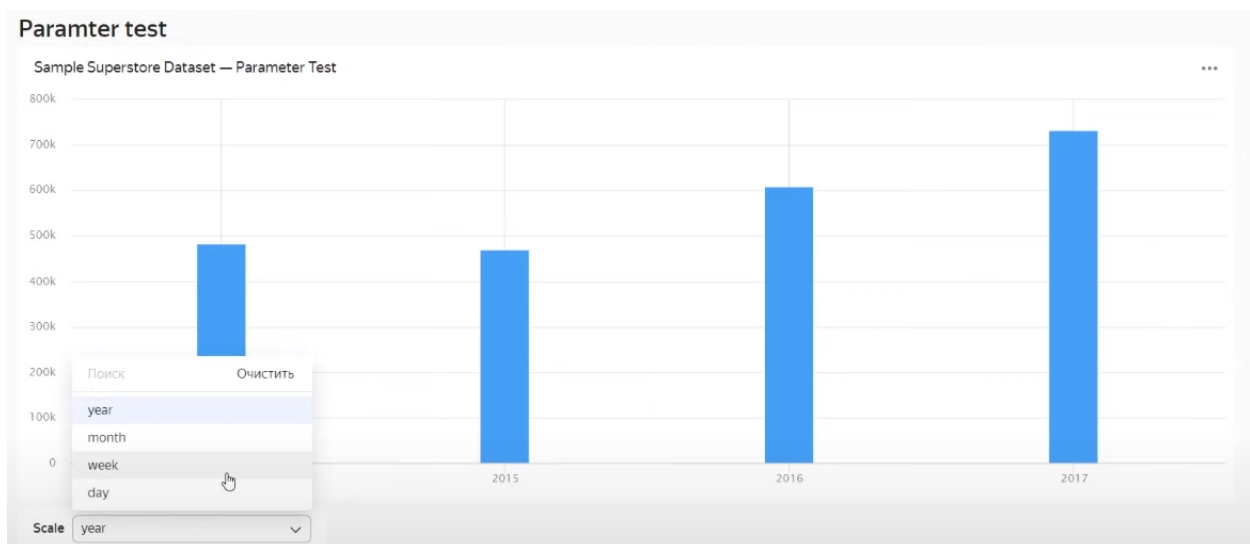
Отменить

Добавить

Через кнопку “Связи” можно задать, на какие именно диаграммы будет действовать наш фильтр:



Также в связке с параметрами селекторы могут менять какие-то характеристики нашего графика - например, уже упомянутое временное разрешение:



Полезные ссылки по DataLens

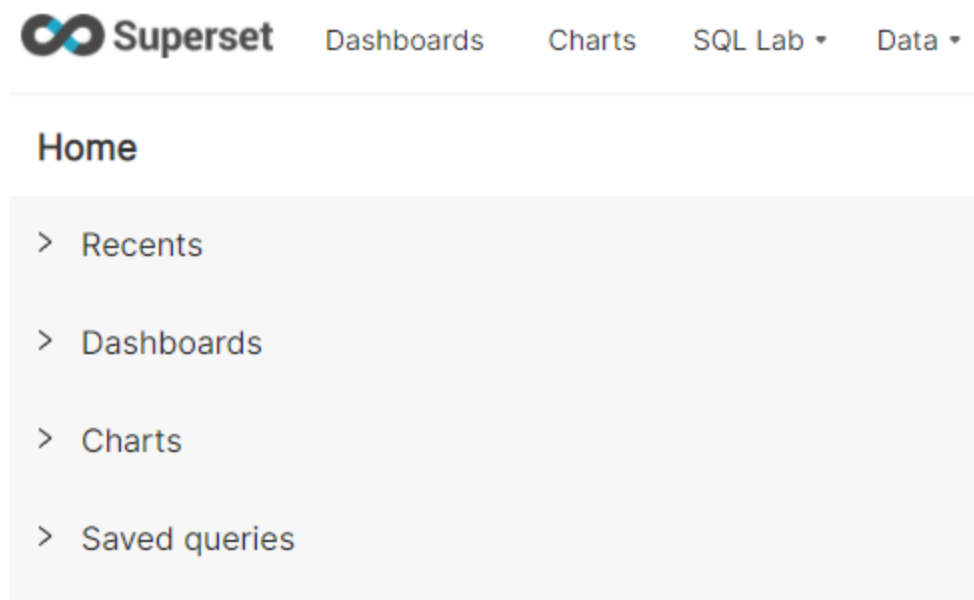
- [Официальная документация](#)
- [Чат в Telegram](#)
- [Учебный марафон](#)

Superset: основной функционал

Apache Superset — это BI-платформа для создания интерактивных визуализаций (по шаблону) и дашбордов. Может быть интегрирована с широким набором баз данных (полный список смотрите [в документации](#)) и развёрнута на собственном сервере. Инструмент **бесплатный**.

Документацию, примеры визуализаций, ссылки на источники информации и ресурсы для обучения можно найти на [официальном сайте](#).

Внутри Superset выглядит примерно так:

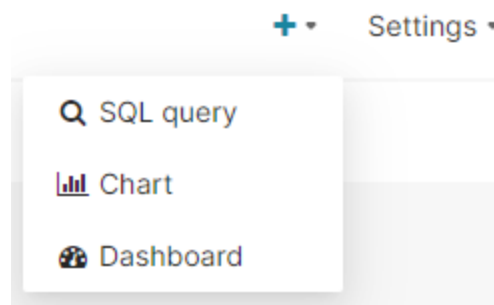


Здесь можно видеть то, что вы недавно создавали или просматривали, а также создавать новые объекты. Всего их три типа:

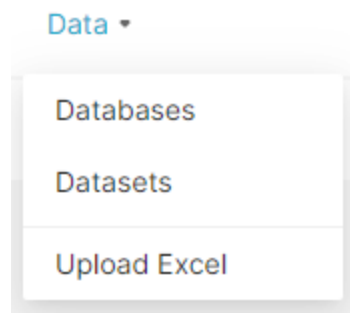
1. Дашборды — интерактивные доски со всей необходимой бизнесу информацией.

2. Диаграммы — основной контент дашборда; представляют собой графики, таблицы и другие интерактивные элементы.
3. Сохранённые запросы — SQL-запросы, которые вы делали ранее и решили оставить на потом.

Всё это можно создавать либо в трёх разделах снизу, либо на панели инструментов сверху, либо через кнопку в правом верхнем углу рядом с настройками:



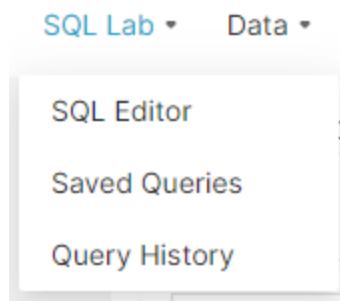
Загружать свои данные можно через раздел **Data**:



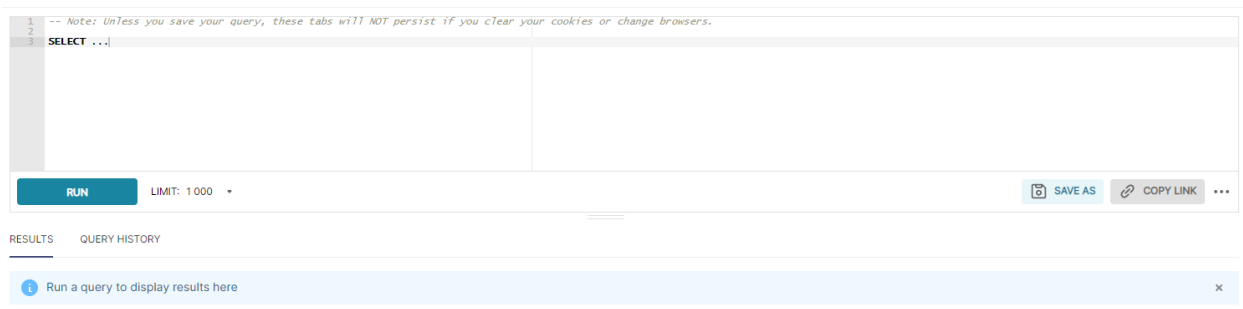
Он показывает:

1. Какие **базы данных** нам доступны — это зависит от того, к какой именно базе вы его подключили.
2. Какие **данные** нам доступны. Здесь можно найти все те наборы данных, которые **находятся** внутри вышеуказанной базы данных, которые вы **создали** в процессе работы над ними и которые вы **загрузили** с компьютера.
3. Также есть возможность загружать свои Excel-таблички и работать с ними.

С данными можно работать, используя соответствующий диалект SQL в разделе **SQL Lab**:



Здесь можно просто писать любой запрос, внешне это очень похоже на Redash.



При выполнении запроса появляется опция **Explore**, через которую и можно перейти к самой визуализации:

EXPLORE

DOWNLOAD TO CSV

COPY TO CLIPBOARD

Filter results

100 rows returned The number of rows displayed is limited to 100 by the limit dropdown.

user_id	post_id	action	time	gender	age	country	city	os	source	exp_group
33138	2484	view	2021-10-03T05:44:11	1	15	Russia	Volgograd	Android	ads	4
25919	2385	view	2021-10-03T05:44:53	1	19	Russia	Moscow	iOS	ads	1
33138	2484	like	2021-10-03T05:45:03	1	15	Russia	Volgograd	Android	ads	4
33138	2502	view	2021-10-03T05:45:22	1	15	Russia	Volgograd	Android	ads	4
133957	2419	view	2021-10-03T05:45:27	0	24	Russia	Biysk	iOS	organic	3
33138	2418	view	2021-10-03T05:45:31	1	15	Russia	Volgograd	Android	ads	4
25919	2266	view	2021-10-03T05:45:33	1	19	Russia	Moscow	iOS	ads	1
133957	2498	view	2021-10-03T05:45:34	0	24	Russia	Biysk	iOS	organic	3
25919	2491	view	2021-10-03T05:45:35	1	19	Russia	Moscow	iOS	ads	1
133957	2485	view	2021-10-03T05:45:35	0	24	Russia	Biysk	iOS	organic	3
133957	2321	view	2021-10-03T05:45:39	0	24	Russia	Biysk	iOS	organic	3
139336	2520	view	2021-10-03T05:45:43	0	30	Russia	Voronezh	Android	organic	3

При переходе к визуализации Superset позволяет сохранить результат в виде виртуального набора данных либо перезаписать существующий.

Save or Overwrite Dataset



Save this query as a virtual dataset to continue exploring

☒ Save as new

Query simulator.feed_actions 10/30/2021 18:44:32

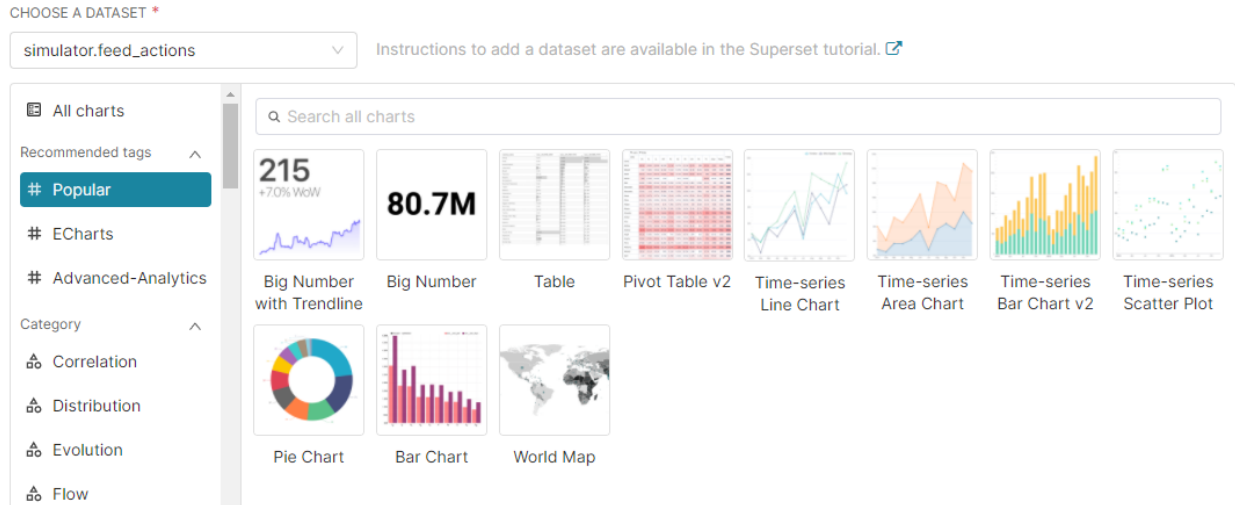
☐ Overwrite existing

Select or type dataset name

SAVE & EXPLORE

Superset: строим дашборд

На дашборде мы размещаем наши диаграммы - **Charts**. Шаблонов достаточно много:

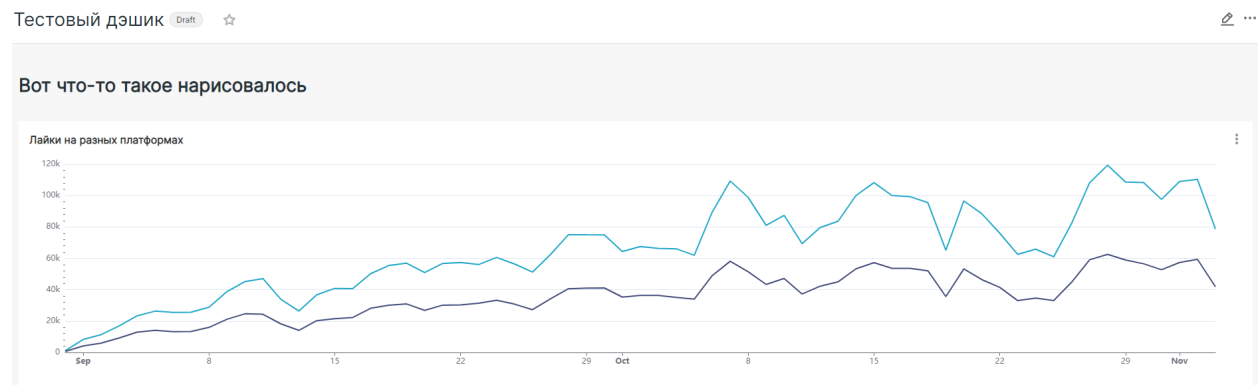


Диаграммы делятся по **категориям** и по **тегам**.

Категории обозначают основную идею — для визуализации чего используют именно такой график или таблицу.

Теги обозначают какое-то свойство диаграммы — она 2D или 3D, изображает тренд, используется для сравнения величин и т.д.

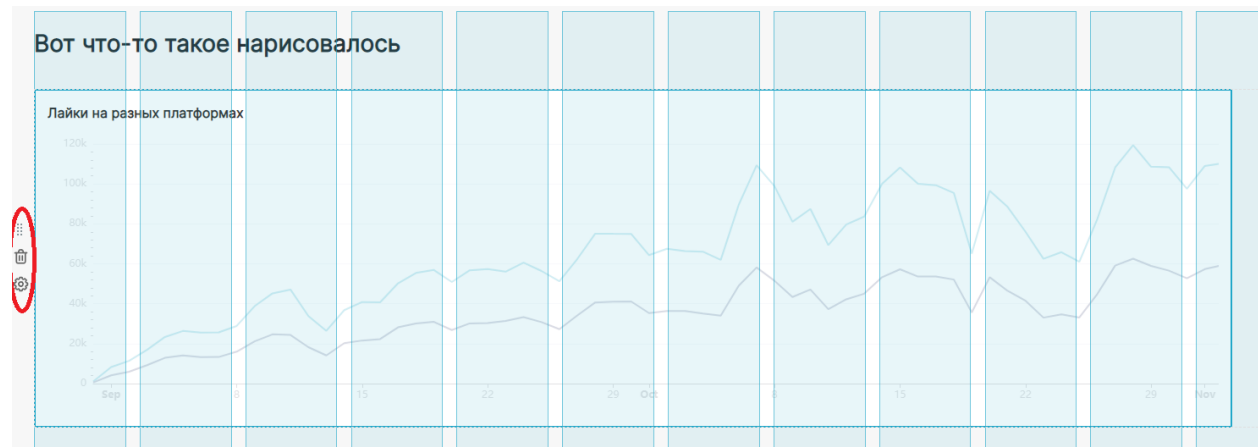
Сами дашборды создаются на вкладке **Dashboards**:



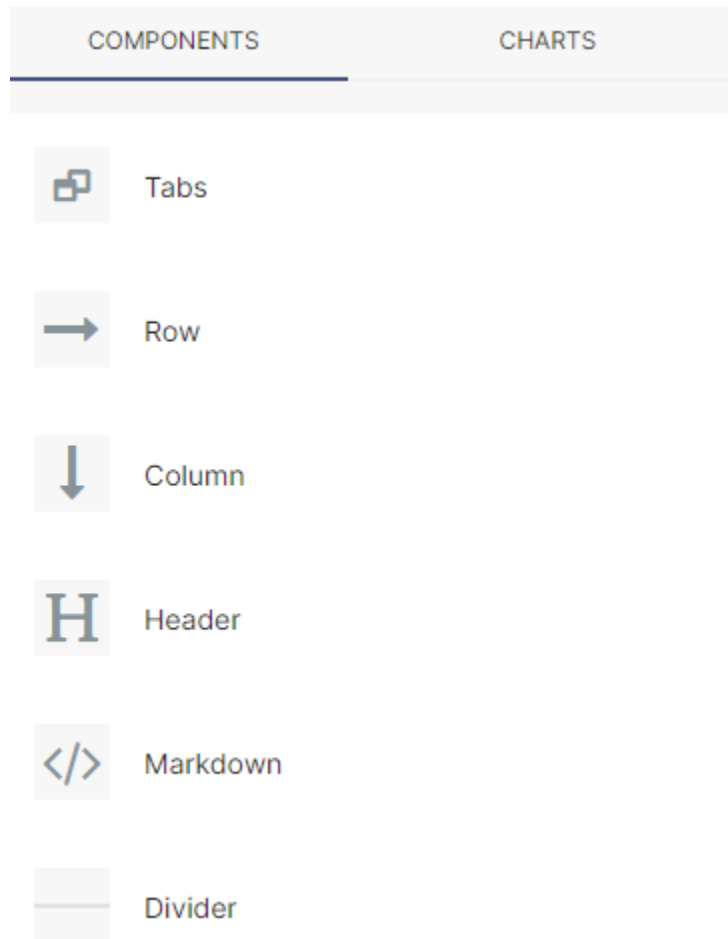
Обратите внимание на тег Draft рядом с названием дашборда — при нажатии на него дашборд становится доступным для просмотра остальным! В любой момент можно вернуться обратно, если не хотите, чтобы дашборд отображался в общем списке.

Размеры и положение графиков можно регулировать их перетаскиванием и растягиванием. Также при выборе объекта слева посередине есть небольшая

панель инструментов, через которую можно удалять и менять настройки отображения:



В окошке справа вы можете выбирать диаграммы, которые хотите разместить на дашборде (вкладка Charts), а также разные сопроводительные элементы:



- **Вкладки (Tabs)** — позволяет разместить графики на нескольких вкладках, которые можно переключать; полезен для экономии места, особенно если графики объединены каким-то общим принципом (например, DAU, WAU и MAU из лекции вполне можно было разместить в трёх разных вкладках и переключать их в зависимости от необходимого временного разрешения).
- **Ряд (Row)** и **колонка (Column)** — создают свободное место под новые диаграммы; полезны для организации элементов дашборда (горизонтально либо вертикально).
- **Заголовок (Header)** — имя раздела дашборда; можно менять размер текста, но только указанными в Superset градациями.
- **Маркдаун-разметка (Markdown)** — позволяет писать текст произвольного форматирования, используя соответствующую разметку; применяется для аннотации дашборда текстом.

- **Разделитель (Divider)** — незаменимый компонент, если графики нужно как-то отделить друг от друга в целях более удачной компоновки.

После создания дашборд можно снова редактировать (нажав на иконку в виде карандаша). Также можно сохранять его копию, сохранять его как изображение и давать ссылку на него:

