

2024

АО «АМТ-ГРУП»
BI-система «Yandex DataLens»

Назначение и преимущества системы

Yandex DataLens – это российская система бизнес-аналитики для визуализации и построения графиков и отчетов

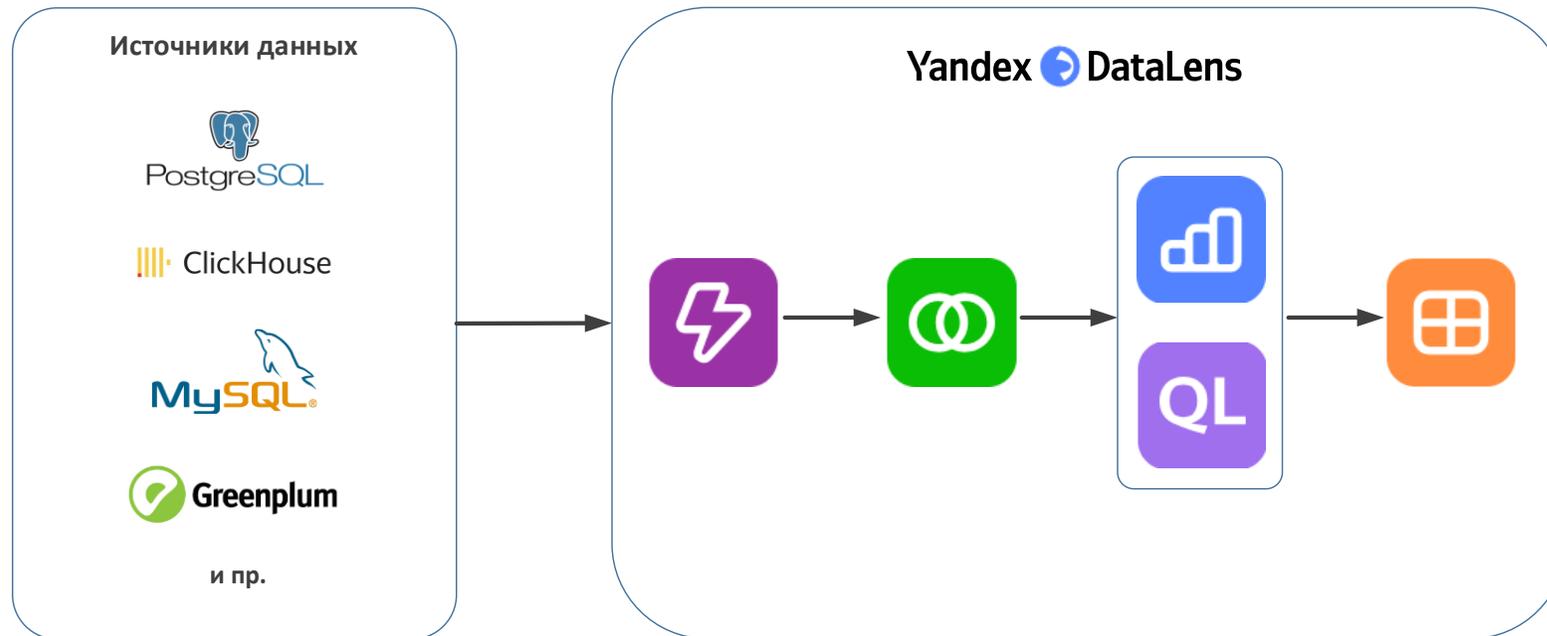


Преимущества:

- Позволяет полностью заменить иностранные аналоги (Power BI, Tableau и т.п.)
- Выполняет аналитические расчеты и построение визуализаций в режиме реального времени на основе данных, полученных напрямую из источников
- Обширный состав подключаемых источников данных
- Интерфейс адаптирован для просмотра на мобильных устройствах
- Может внедряться в двух конфигурациях – как локальный open source, и как облачный BI-сервис на платформе Yandex Cloud.
- Отсутствие ежемесячных платежей за пользование продуктом и возможность дополнительной модификации DataLens под нужды Заказчика в случае использования open source конфигурации
- Простой и понятный локализованный интерфейс
- Наличие документации по внедрению и настройке и бесплатных обучающих курсов по работе с системой

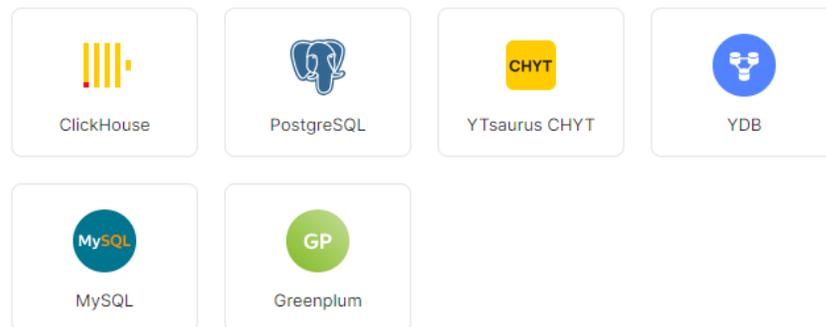
Ключевые сущности DataLens

-  Подключение – задает параметры для доступа к источнику данных
-  Датасет – описание набора данных из источника
-  Чарт в визарде – визуализация на основе данных из датасета
-  QL-чарт – визуализация данных путем прямого запроса в источник данных через подключение
-  Дашборд – набор страниц с чартами и дополнительными виджетами для фильтрации



Подключения к источникам данных

- Источники для open source версии - PostgreSQL, MySQL, ClickHouse, Greenplum, YTsaurus CHYT, YDB
- Расширенный набор типов источников для облачной версии (включая файлы, файловые облачные сервисы, партнерские подключения)
- Настройка кэширования
- Настройки уровня доступа для выполнения прямых SQL-запросов
- Продвинутые настройки подключения (использование сертификатов, управление доступом к экспорту данных)



PostgreSQL

Имя хоста

Порт

Путь к базе данных

Имя пользователя

Пароль

Время жизни кеша в секундах

Уровень доступа SQL запросов Разрешить подзапросы в датасетах и запросы из чартов
 Разрешить подзапросы в датасетах
 Запретить

Продвинутые настройки подключения

Указание collate в запросе

TLS

CA Certificate

Запрет на экспорт данных

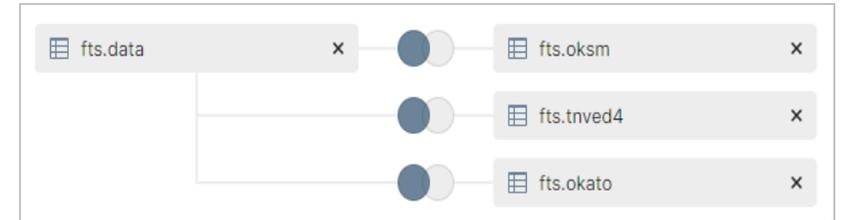
Модель данных датасета

- Объединение данных из нескольких таблиц по одному или нескольким полям
- Поддержка всех типов связи JOIN: inner, left, right и full
- Оптимизация связи между таблицами – применение оператора JOIN только в запросах к нескольким таблицам
- Создание дополнительных полей, в т.ч. вычисляемых
- Задание для поля типа агрегации по умолчанию
- Добавление параметров и фильтров на уровне датасета
- Предпросмотр данных датасета

#	Имя ↓	Источник поля ↓	Тип ↓	Агрегация ↓	Описание ↓
9	quantity	data.quantity	# Целое число	Нет	
10	region_code	data.okato	Строка	Нет	
11	district_code	data.district_code	Строка	Нет	
12	dt	data.dt	Дата	Нет	

Предпросмотр Количество строк: 3 | не больше 1 000

#	year	direction	month	country	code	measure	value	netto	quantity	region_code	district_code	dt	gruppa	tov_poz
1	2020	ИМ	10/2020	KR	2922290000	1	20790	8000	0	40000	02	2020-10-01	29	22
2	2020	ЭК	03/2020	BY	2007999708	1	26	52	0	36000	04	2020-03-01	20	07
3	2020	ИМ	12/2020	CN	8409990009	1	51812	14747	0	08000	07	2020-12-01	84	09



Связь

fts.data — inner — fts.oksm

country — alpha2

Добавить связь

Оптимизировать связь ?

Отменить **Применить**

Настройка поля

year Формула Поле из источника Не показывать в визарде

1 | YEAR([dt])

year fn

direction

Датасет на основе SQL-запроса

- Позволяет расширить модель данных для построения визуализаций
- Можно объединять со стандартными датасетами
- Доступны все операции по работе со стандартным датасетом
- Диалект SQL-запроса соответствует диалекту источника из выбранного подключения

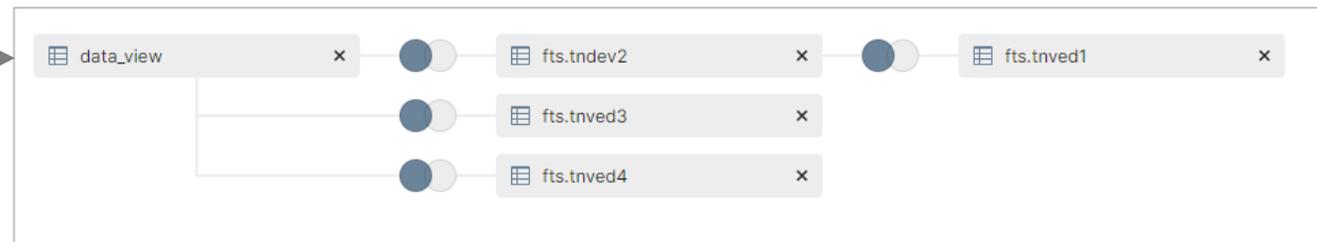
Источник ×

Название источника

data_view

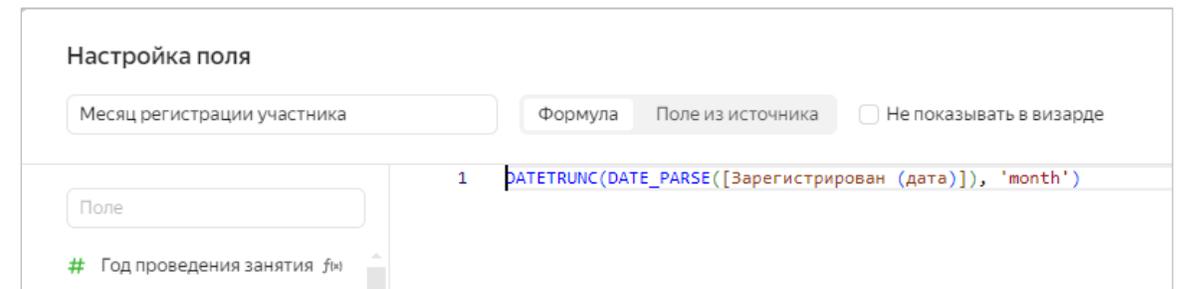
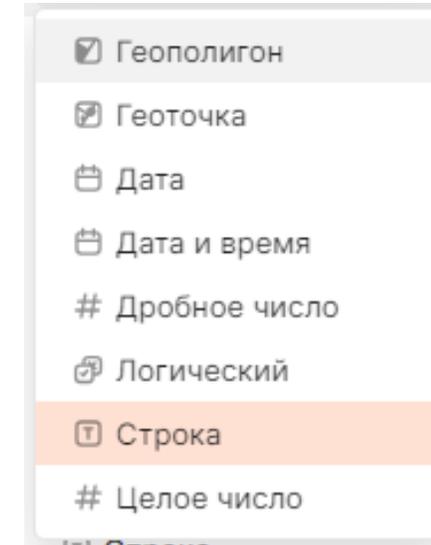
Подзапрос ⓘ

```
6 select substring(code for 2) as gruppa,  
7     substring(code from 3 for 2) as tov_poz,  
8     substring(code from 5 for 6) as sub_poz,  
9     *  
10 from fts.data  
11
```



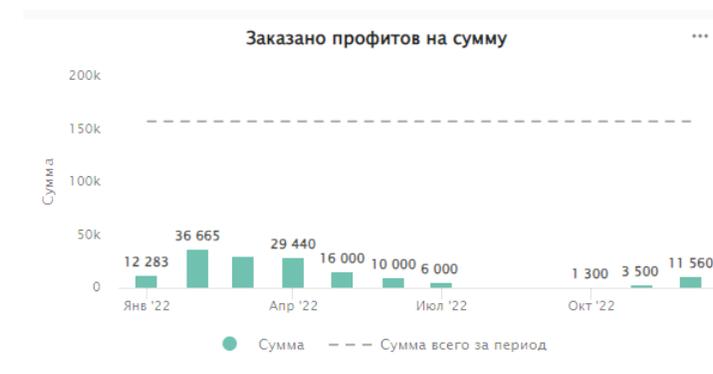
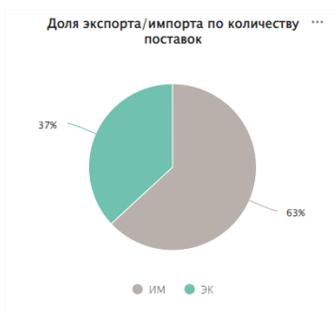
Типы данных DataLens

- Возможность работы с различными источниками данных, в которых типы данных полей могут отличаться
- Для унификации и более эффективной работы с данными DataLens конвертирует типы данных полей источника в собственные типы:
 - Логический
 - Дата
 - Дата и время
 - Дробное число
 - Целое число
 - Строка
 - ГеополYGON
 - Геоточка
 - Массивы целых чисел, дробных чисел, строк
- Изменение типа данных полей доступно через интерфейс системы и через функции по работе с полями



Возможности визуализации

- Обширный набор типов диаграмм - линейная, накопительная с областями, столбчатая, линейчатая, нормированная, точечная, круговая, кольцевая, комбинированная, древовидная, таблица (в т.ч. сводная), индикатор
- Построение визуализаций на карте (доступно в облачной версии продукта)
- Создание диаграмм на основе данных, полученных через прямой запрос в источник (QL-чарты)
- Поддержка мультидатасетных диаграмм
- Добавление параметров и фильтров на уровне диаграммы
- Сортировка значений внутри диаграммы, детализация во всплывающем окне, изменение цветов и формы линий
- Наличие инспектора чартов для оценки скорости выполнения запроса к источнику



Гибкая настройка чартов

- Общие настройки – заголовок чарта, отображение легенды, пагинация для таблиц, отображение суммы в тултипах и пр.
- Настройки полей - название для отображения на диаграмме, тип данных, правила агрегации и группировки, формат отображения значения, отображение группы разрядов, префикс, постфикс, размерность и пр.
- Настройки осей диаграмм – название, отображение сетки масштабирование и пр.
- Настройки цвета для полей и значений, включая задание цвета по пороговым значениям и градиентную заливку
- Настройки подписей, сортировки по значениям, предустановленные фильтры
- Доступный состав настроек зависит от типа чарта и типа поля данных

value

Название: value

Тип (до агрегации): Целое число

Агрегация: Сумма

Подпись: Скрыть Показать

Формат: Число Процент

Отображать группы разрядов: С разделителем Слитно

Префикс:

Постфикс: \$

Размерность: Миллионы, М

Отменить Применить

Северо-Западный ФО

Линейный индикатор:

Тип заливки: Один цвет Два цвета Градиент

Цвет: EAE6E6

Отображать значения:

Отображать в итогах:

Выравнивание: По умолчанию Слева Справа

Масштаб: Авто Вручную

мин макс

Настройки чарта

Заголовок: Общая стоимость поставок по подлози

Пагинация:

Лимит: 100

Группировка:

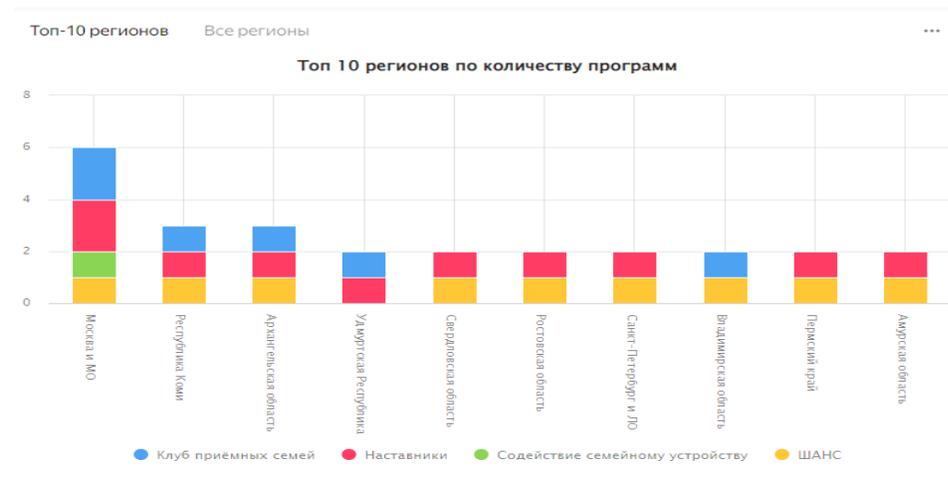
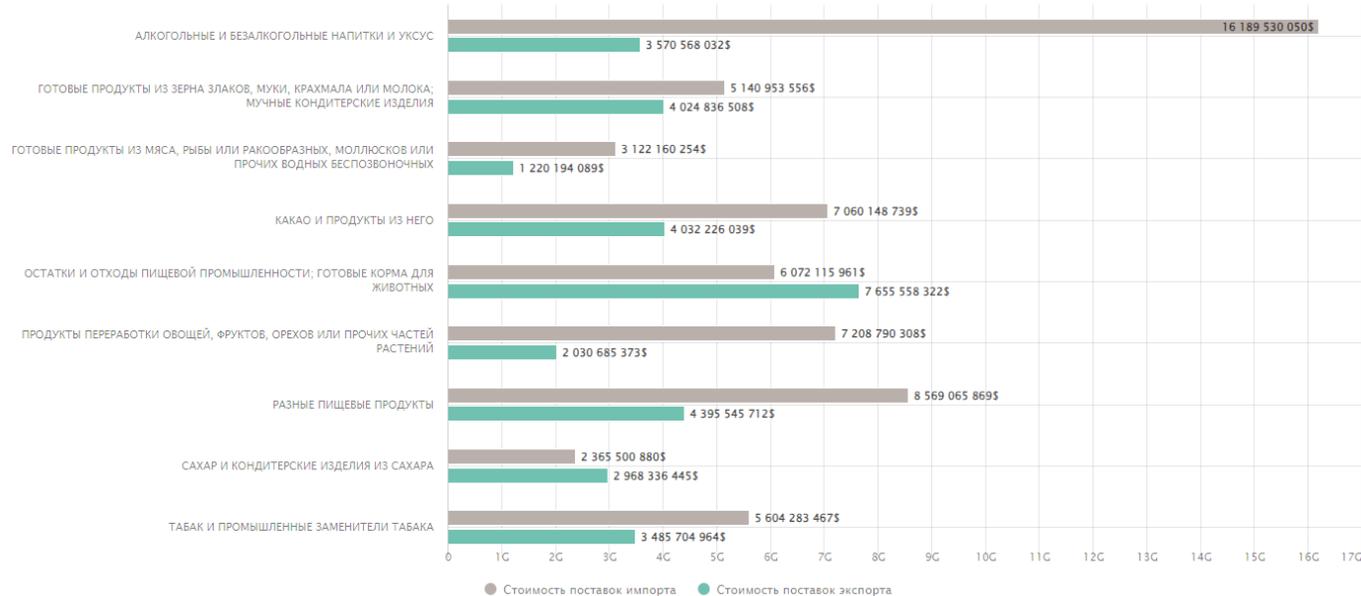
Итоги:

Отменить Применить

Линейчатые и столбчатые диаграммы

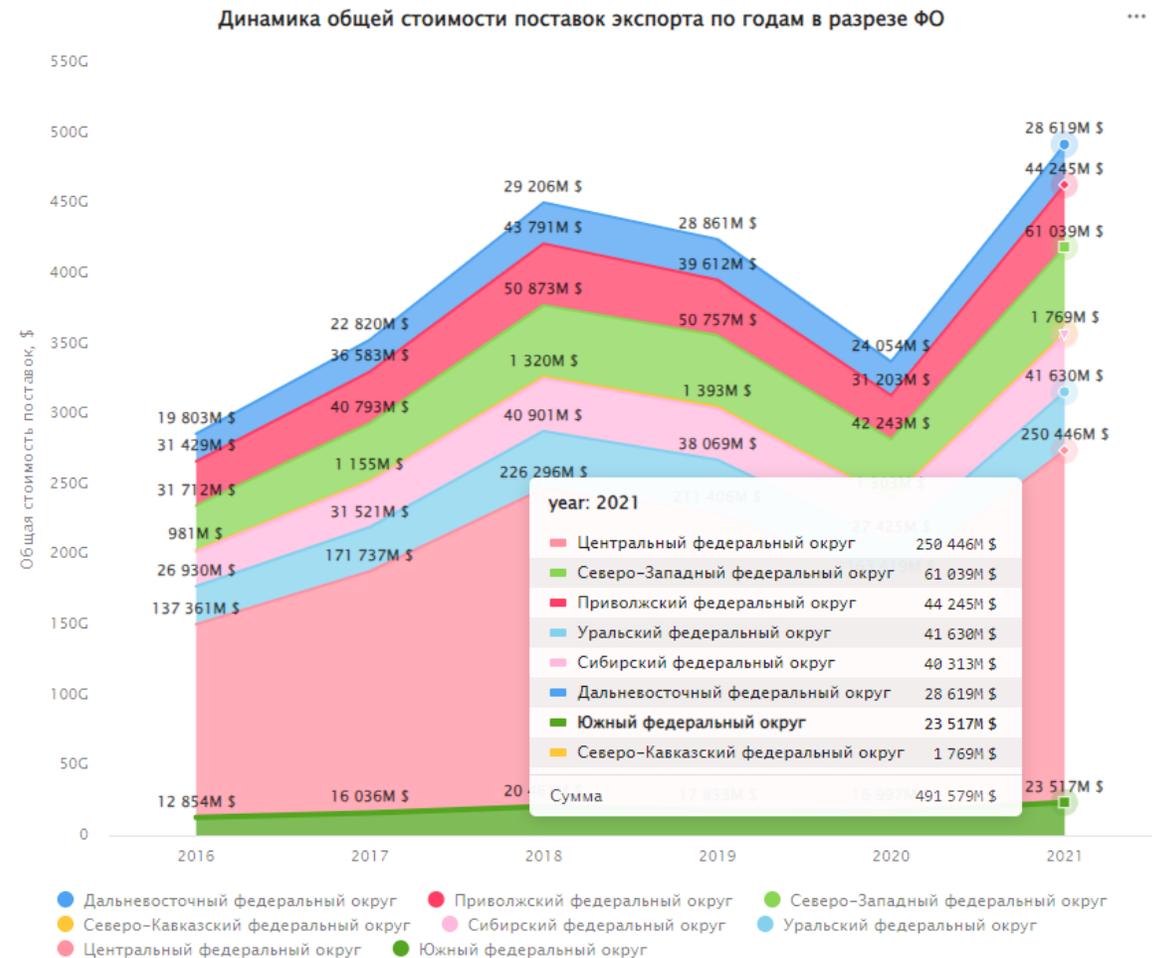
- Используются для отображения значения одного или нескольких показателей по периодам времени или по категориям
- Дополнительные возможности отображения:
 - С группировкой – для сравнения нескольких категорий по одному показателю
 - С накоплением – для оценки соотношения внутренних сегментов в итоговой величине

Общая стоимость поставок по разделам и группам ТНВЭД



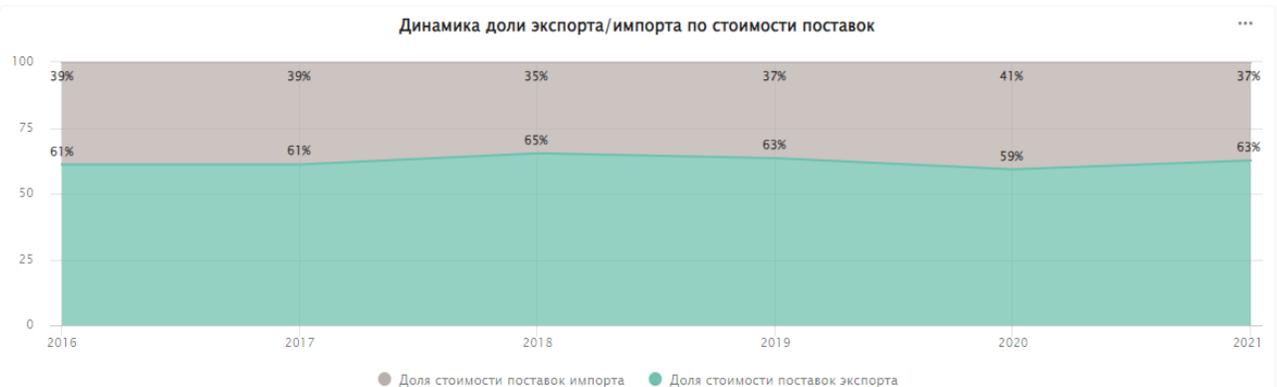
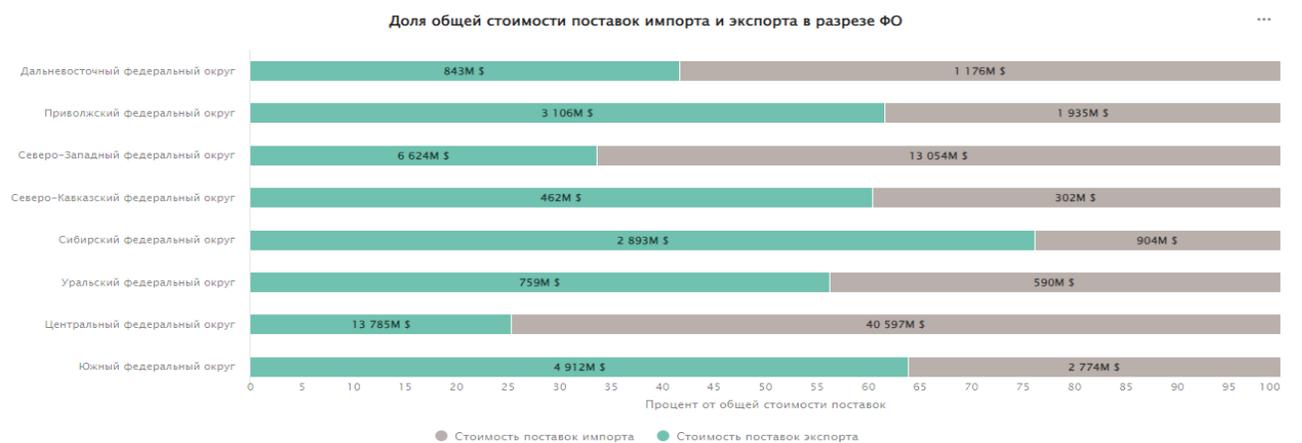
Линейные диаграммы и диаграммы с областями

- Показывают изменение одного или нескольких показателей во времени
- Линейная диаграмма может показывать несколько показателей разными линиями
- Накопительную диаграмму с областями используют для отображения суммарного итога в разбивке по нескольким категориям



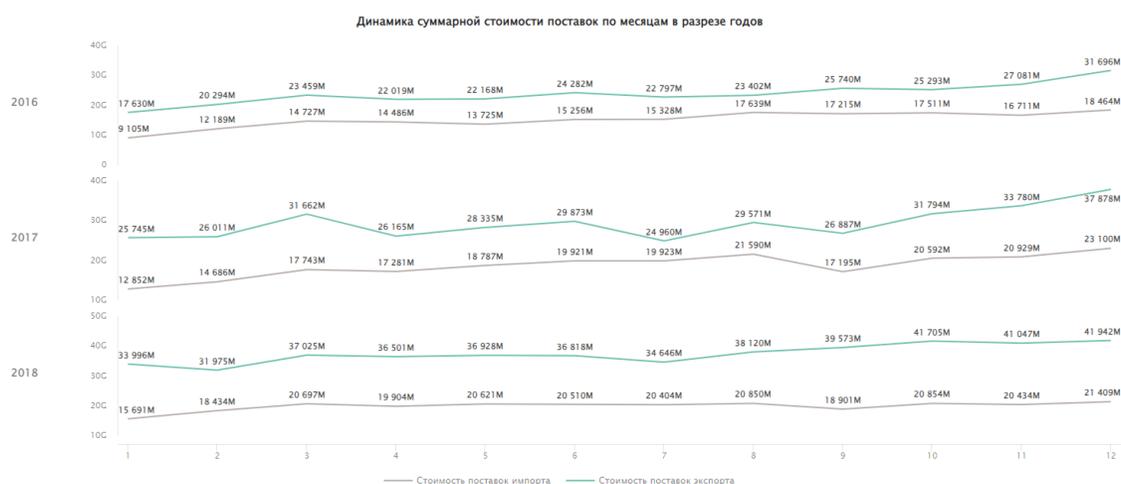
Нормированные диаграммы

- Используются для отображения процентного соотношения между показателями
- Ширина отдельной области показателя соответствует ее доле от общей суммы, приведенной к 100%
- Нормирование может быть применено в столбчатых, линейчатых диаграммах и диаграммах с областями



Функция сплита диаграммы

- Сплит диаграммы позволяет разделить ее на несколько небольших диаграмм по отдельному измерению
- Сплит доступен для диаграмм – линейная, столбчатая, диаграмма с областями



Доля стоимости поставок по ФО РФ – динамика по годам в разрезе импорта *** и экспорта



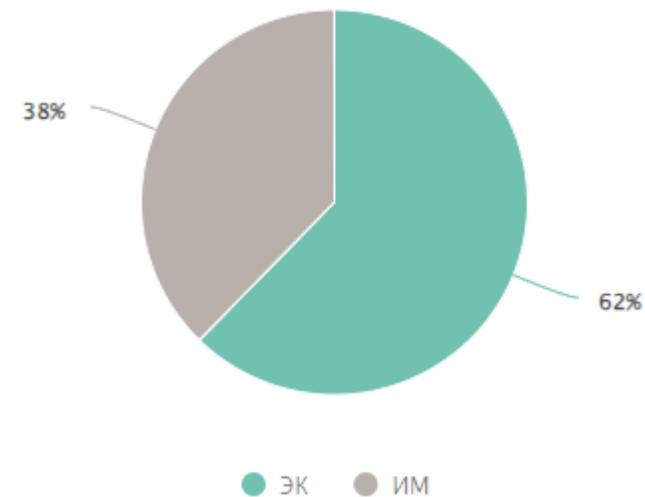
- Дальневосточный федеральный округ
- Северо-Западный федеральный округ
- Сибирский федеральный округ
- Центральный федеральный округ
- Приволжский федеральный округ
- Северо-Кавказский федеральный округ
- Уральский федеральный округ
- Южный федеральный округ

Круговая и кольцевые диаграммы

- Отображают пропорции между категориями при помощи сегментов и секторов
- Подходят для отображения небольшого количества категорий
- Размер каждого сегмента соответствует доле категории в общей сумме по выбранному показателю
- Кольцевая диаграмма в центре отображает общий итог выбранного показателя



Доля экспорта/импорта по стоимости поставок (в \$) ***



Табличные визуализации

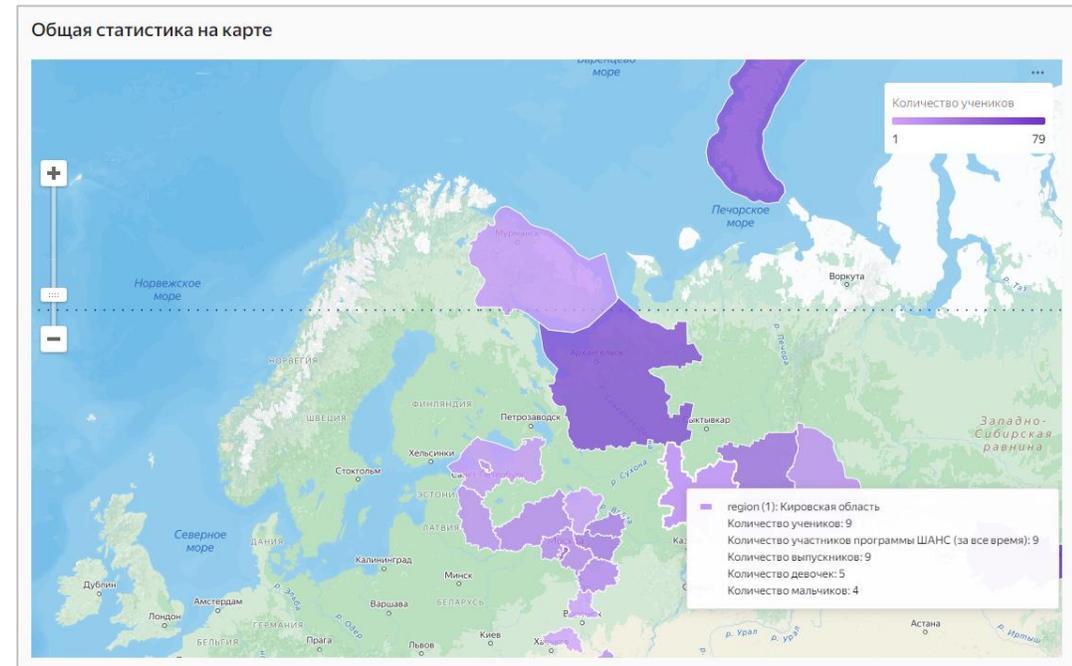
- Подходят для количественного сравнения нескольких категорий и показателей
- Могут быть простыми двумерными и сводными
- Возможность дополнительных настроек - изменение названий столбцов; настройка сортировки в таблице по нескольким столбцам, ширины столбцов таблицы, заливки полей и цвета строка; добавление строки с итогами, линейного индикатора в столбец с показателем

Страна	Количество поставок импорта, шт.	Количество поставок экспорта, шт.	Доля поставок импорта, %	Доля поставок экспорта, %	Общая стоимость импорта, \$	Общая стоимость экспорта, \$	Доля стоимости импорта, %	Доля стоимости экспорта, %
ГЕРМАНИЯ	1 367 014	175 297	89%	11%	144 678 530 539 \$	157 897 775 546 \$	48%	52%
ИТАЛИЯ	1 045 697	57 140	95%	5%	61 556 691 952 \$	85 485 984 340 \$	42%	58%
ДАНИЯ	156 118	12 812	92%	8%	6 299 323 827 \$	16 128 629 743 \$	28%	72%
ГОНКОНГ	56 580	18 397	75%	25%	2 225 729 681 \$	6 742 789 296 \$	25%	75%
АБХАЗИЯ	1 715	156 974	1%	99%	386 538 466 \$	1 266 073 444 \$	23%	77%
АНДОРРА	253	155	62%	38%	463 192 \$	31 358 631 \$	1%	99%
БУТАН	10	1	91%	9%	34 663 \$	329 \$	99%	1%

Раздел ТН ВЭД	Общая стоимость поставок экспорта по разделам и ФО								
	Дальневосточный федеральный округ	Приволжский федеральный округ	Северо-Западный федеральный округ	Северо-Кавказский федеральный округ	Сибирский федеральный округ	Уральский федеральный округ	Центральный федеральный округ	Южный федеральный округ	Итого
ЖИВЫЕ ЖИВОТНЫЕ; ПРОДУКТЫ ЖИВОТНОГО ПРОИСХОЖДЕНИЯ	19 134M \$	428M \$	6 070M \$	624M \$	525M \$	210M \$	3 819M \$	918M \$	31 729M \$
МИНЕРАЛЬНЫЕ ПРОДУКТЫ	89 919M \$	102 639M \$	131 368M \$	306M \$	93 720M \$	124 858M \$	572 606M \$	34 357M \$	1 149 774M \$
НЕДРАГОЦЕННЫЕ МЕТАЛЛЫ И ИЗДЕЛИЯ ИЗ НИХ	2 761M \$	15 793M \$	38 955M \$	381M \$	57 193M \$	49 992M \$	58 139M \$	10 146M \$	233 360M \$
ПЛАСТМАССЫ И ИЗДЕЛИЯ ИЗ НИХ; КАУЧУК, РЕЗИНА И ИЗДЕЛИЯ ИЗ НИХ	43M \$	15 552M \$	3 896M \$	438M \$	2 094M \$	2 245M \$	13 646M \$	1 383M \$	39 298M \$
ПРОДУКТЫ РАСТИТЕЛЬНОГО ПРОИСХОЖДЕНИЯ	1 653M \$	1 709M \$	4 511M \$	1 341M \$	2 564M \$	291M \$	17 179M \$	33 124M \$	62 372M \$
РАЗНЫЕ ПРОМЫШЛЕННЫЕ ТОВАРЫ	23M \$	1 269M \$	916M \$	104M \$	213M \$	233M \$	3 726M \$	330M \$	6 814M \$
ТЕКСТИЛЬНЫЕ МАТЕРИАЛЫ И ТЕКСТИЛЬНЫЕ ИЗДЕЛИЯ	26M \$	351M \$	1 016M \$	71M \$	142M \$	95M \$	4 092M \$	510M \$	6 304M \$
Итого	113 561M \$	137 741M \$	186 732M \$	3 266M \$	156 451M \$	177 924M \$	673 207M \$	80 769M \$	1 529 650M \$

Визуализации на карте*

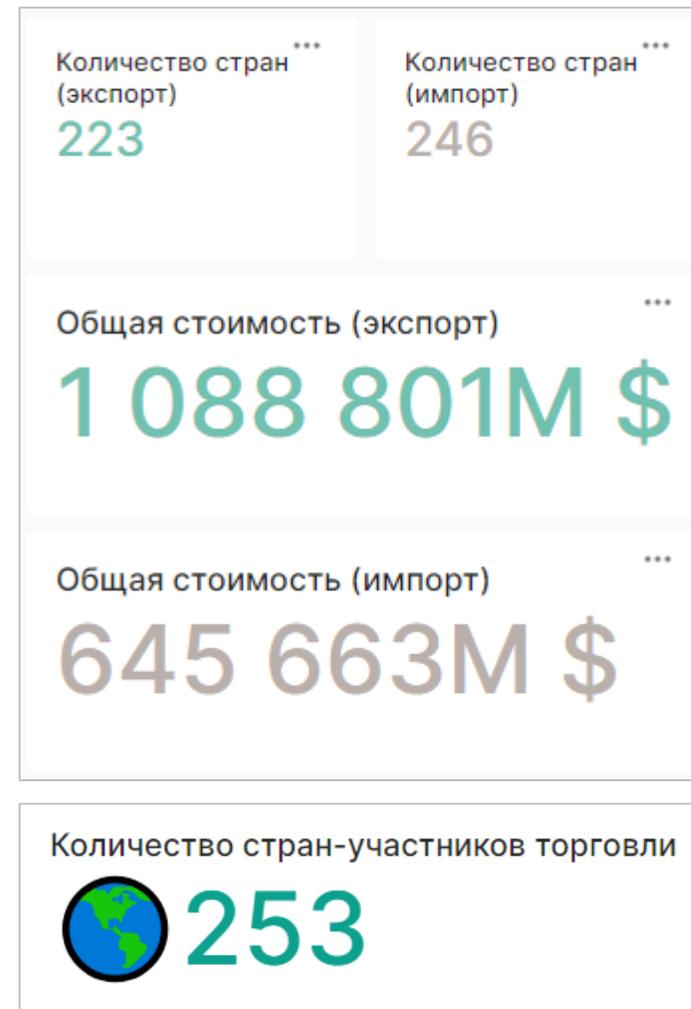
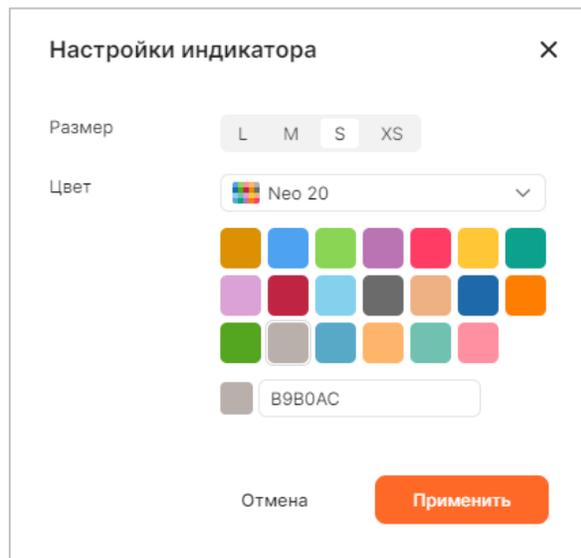
- Используются для отображения показателей на карте с использованием полей с типом данных Геополлигон или Геоточка
- Поддерживаются различные типы картографических визуализаций:
 - Точечная карта, в том числе с кластеризацией
 - Полилинейная карта
 - Фоновая карта
 - Тепловая карта
- На одной карте возможно размещение до 5 геослоев различных картографических типов
- Возможность задать цвета областей, точек, линий, так же размер точки в зависимости от значения показателя
- Вывод легенды и детальных данных во всплывающих подсказках



*карты доступны в облачной версии системы; реализация визуализаций карт в open source версии запланирована в родмапе по развитию

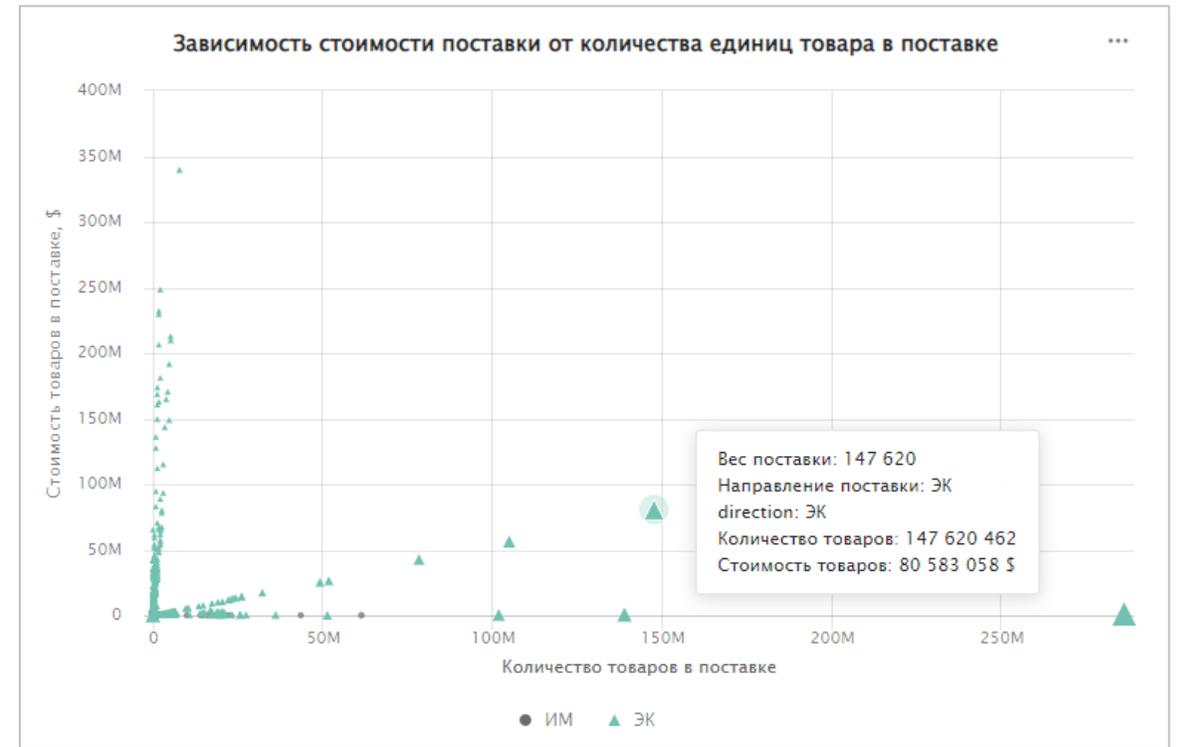
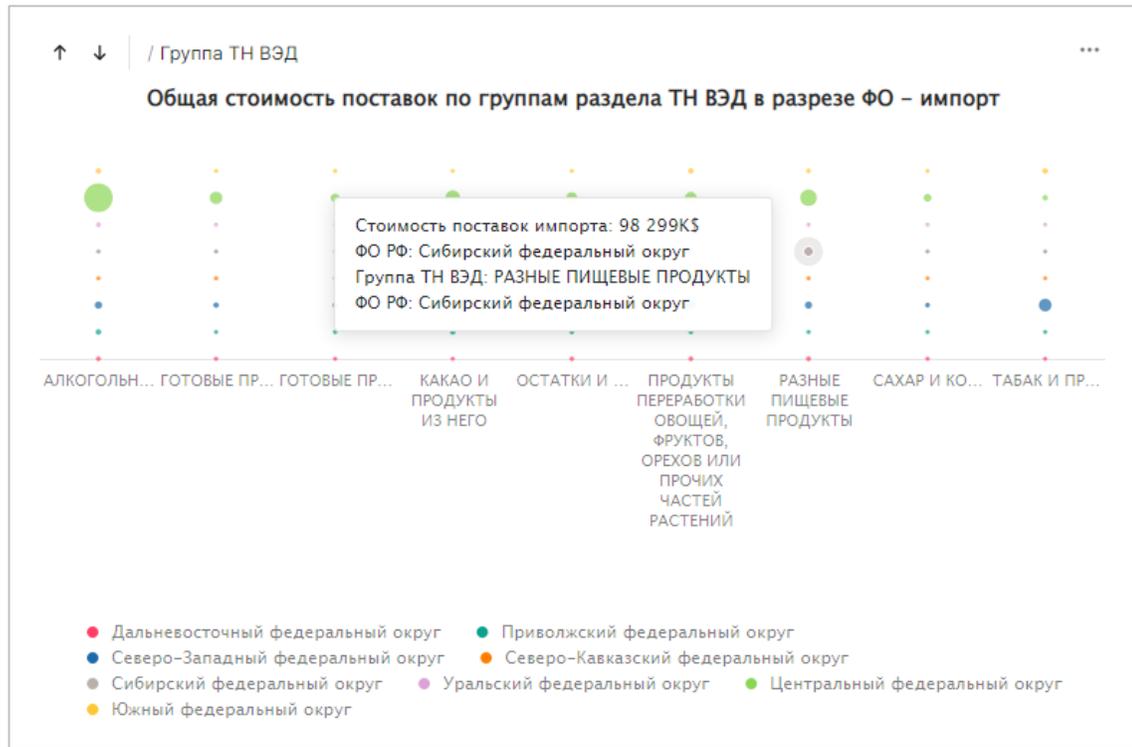
Индикаторы

- Используются для отображения одного ключевого показателя
- Возможность настройки размера и цвета индикатора
- Возможность использования графических символов при вычислении значений индикатора



Точечные диаграммы

- Используются для визуализации отношения между двумя величинами в виде разброса точек
- Возможность определения размера точек в зависимости от значения заданного показателя
- Возможность определения цвета и внешнего вида точек в зависимости от категории



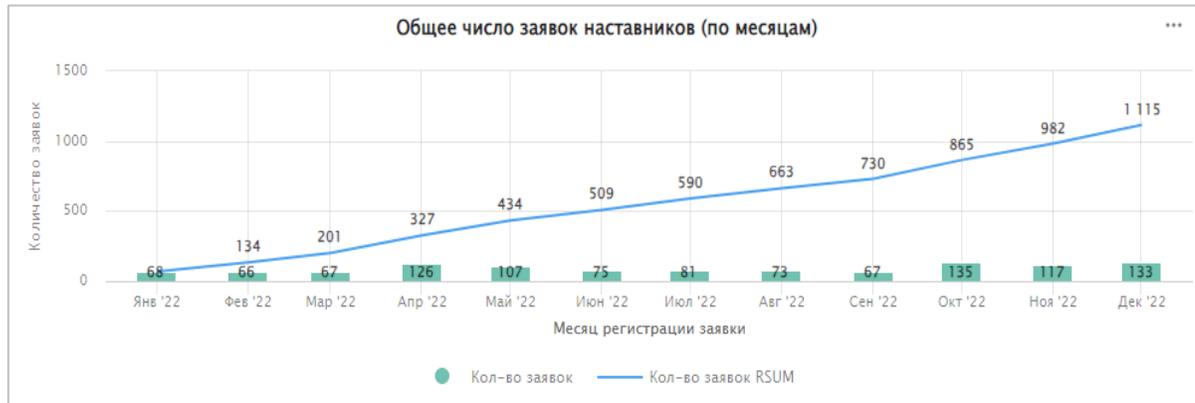
Древоподобные диаграммы

- Отражают иерархические данные как набор вложенных прямоугольников разных цветов
- Размеры прямоугольников показывают пропорции между значениями показателей и суммарным целым значением
- Могут использоваться для отображения пропорций категорий внутри одноуровневого показателя
- Возможность определения цвета в зависимости от принадлежности к категории или измеряемого значения

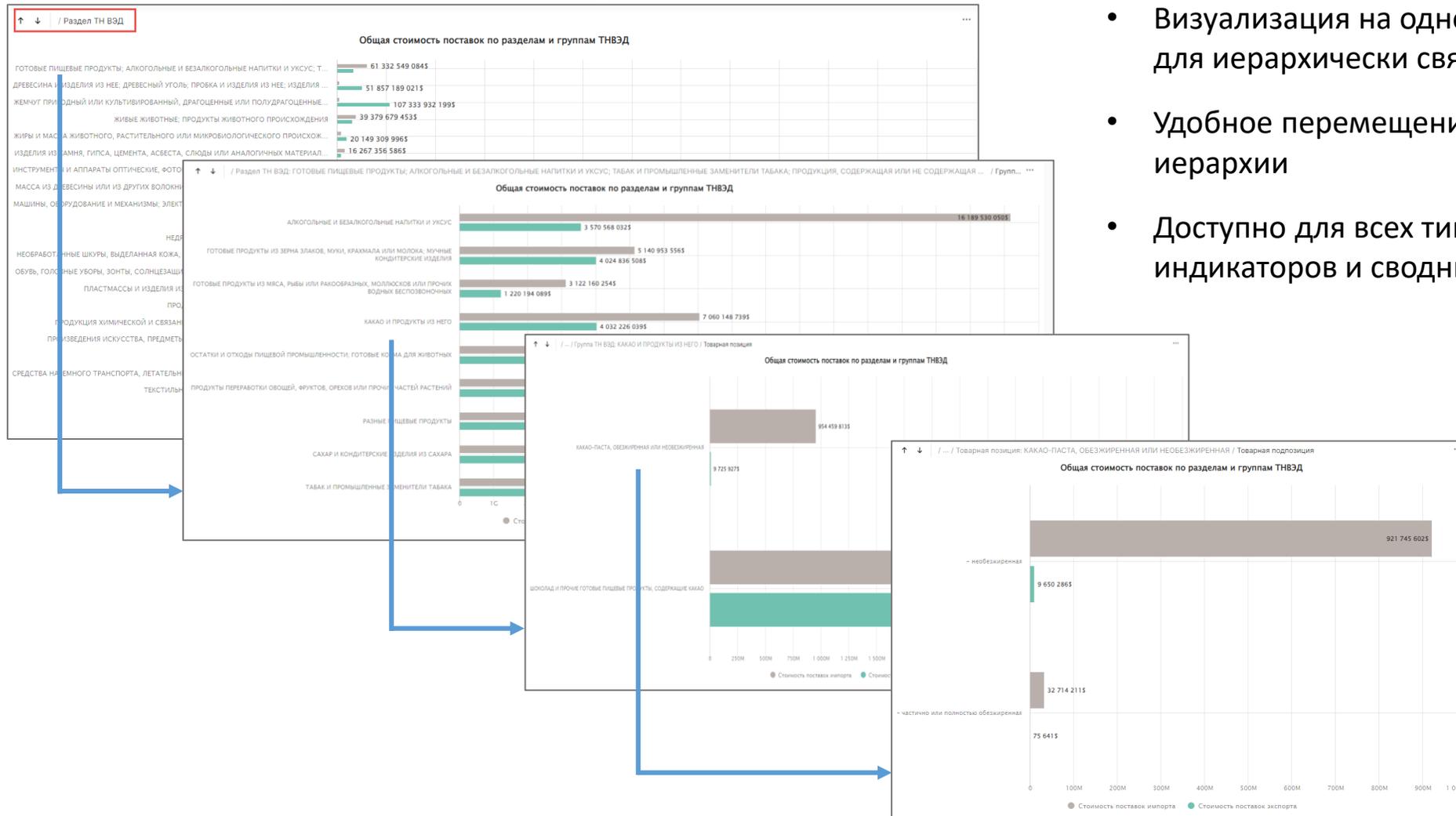


Комбинированные диаграммы

- Отдельный тип визуализации, который состоит из слоев
- Каждый слой представлен отдельным типом диаграммы
- Можно комбинировать линейные, столбчатые и диаграммы с областями
- Могут использоваться для сравнения плана и факта



Визуализация иерархий



- Визуализация на одной диаграмме значений для иерархически связанных показателей
- Удобное перемещение между уровнями иерархии
- Доступно для всех типов диаграмм, кроме карт, индикаторов и сводных таблиц

QL-чарты

- Визуализация на основе прямого запроса в источник на SQL-диалекте исходной БД
- Позволяет оптимизировать скорость выполнения запроса и расширить возможности визуализации с помощью специфических операций
- Возможность использования параметризованных запросов
- QL-чарты поддерживают все типы визуализаций, кроме карт, сводных таблиц и комбинированных диаграмм
- Визуализация генерируется сразу без создания промежуточного датасета и отображает данные из результата запроса

The screenshot displays a SQL query editor on the left and a table of results on the right. The query is a complex SQL statement with multiple filters and calculations. The table shows data for various countries, including the number of import and export shipments and the percentage of imports.

Страна	Количество поставок импорта, шт.	Количество поставок экспорта, шт.
АФГАНИСТАН	480	
АЛБАНИЯ	15 162	
АНТАРКТИДА	32	
АЛЖИР	819	1
АМЕРИКАНСКОЕ САМОА	66	
АНДОРРА	253	
АНГОЛА	60	
АНТИГУА И БАРБУДА	17	
АЗЕРБАЙДЖАН	17 950	236
АРГЕНТИНА	27 837	
АВСТРИЯ	37 403	1
АВСТРАЛИЯ	229 760	16
БАГАМЫ	421	
БАХРЕЙН	787	
БАНГЛАДЕШ	49 941	1
АРМЕНИЯ	66 055	37
БАРБАДОС	537	
БЕЛЬГИЯ	215 738	36
БЕРМУДЫ	229	

The screenshot shows a parameter configuration window with two input fields. The first field is labeled 'date-interval' and contains the date range '01.01.2010 — 01.01.2025'. The second field is labeled 'string' and contains the text 'name_country'. There is a 'Добавить значение' button below the fields and a 'Добавить параметр' button at the bottom.

Вычисляемые поля

- Возможность создания дополнительного поля данных, значения которого вычисляются при помощи формул
- В формулах могут использоваться существующие поля датасета, параметры, константы и функции
- Обширный набор доступных функций:
 - математические
 - агрегатные
 - логические
 - строковые
 - преобразование типов данных
 - функции для работы с датой и временем
 - оконные функции
 - функции для работы массивами, временными рядами и LOD-выражениями
 - функции разметки

Настройка поля

Участников ШАНС 2022 2023 RSUM

Поле

- ID семьи
- ID ученика

```
1 RSUM(COUNTD([ID ученика]))
2 ORDER BY [Месяц регистрации ученика]
3 BEFORE FILTER BY [Месяц регистрации ученика])
```

Настройка поля

Регион

Формула Поле из источника Не показывать в визарде

Поле

- # Год проведения занятия f(n)
- # Год выпуска f(n)
- # Возраст участника на М... f(n)

```
1 CASE [Регион исходный]
2 WHEN 'г. Москва' THEN 'Москва и МО'
3 WHEN 'Московская область' THEN 'Москва и МО'
4 WHEN 'Санкт-Петербург' THEN 'СПб и ЛО'
5 WHEN 'Ленинградская область' THEN 'СПб и ЛО'
6 ELSE [Регион исходный]
7 END
```

Инспектор чартов

- Предназначены для анализа работы построенных диаграмм и оптимизации процесса получения данных и отрисовки чарта
- Содержит техническую информацию для диагностики проблем в загрузке данных:
 - время загрузки конфигурации чарта и формирование запроса для него в источник
 - время выполнения запроса в источнике и передача их в DataLens,
 - время обработки в DataLens данных, полученных из источника,
 - время отрисовки данных в чарте,
 - идентификатор запроса в источник,
 - текст сформированного SQL-запрос, отправленного в источник данных

The screenshot shows the 'Инспектор' (Inspector) window with the following details:

- Performance Metrics:**
 - 0,1978 мс: Загрузка конфига
 - 2,43 с: Загрузка данных
 - 2,99 мс: Выполнение
 - 0,9 мс: Рендеринг
- Request ID:** d1.31123_ck.yr7fmt8m1o
- Источники данных:** DataLens BI Datasets: data_union (2,43 с, 5,92 КБ)
- Запрос в источник данных:**

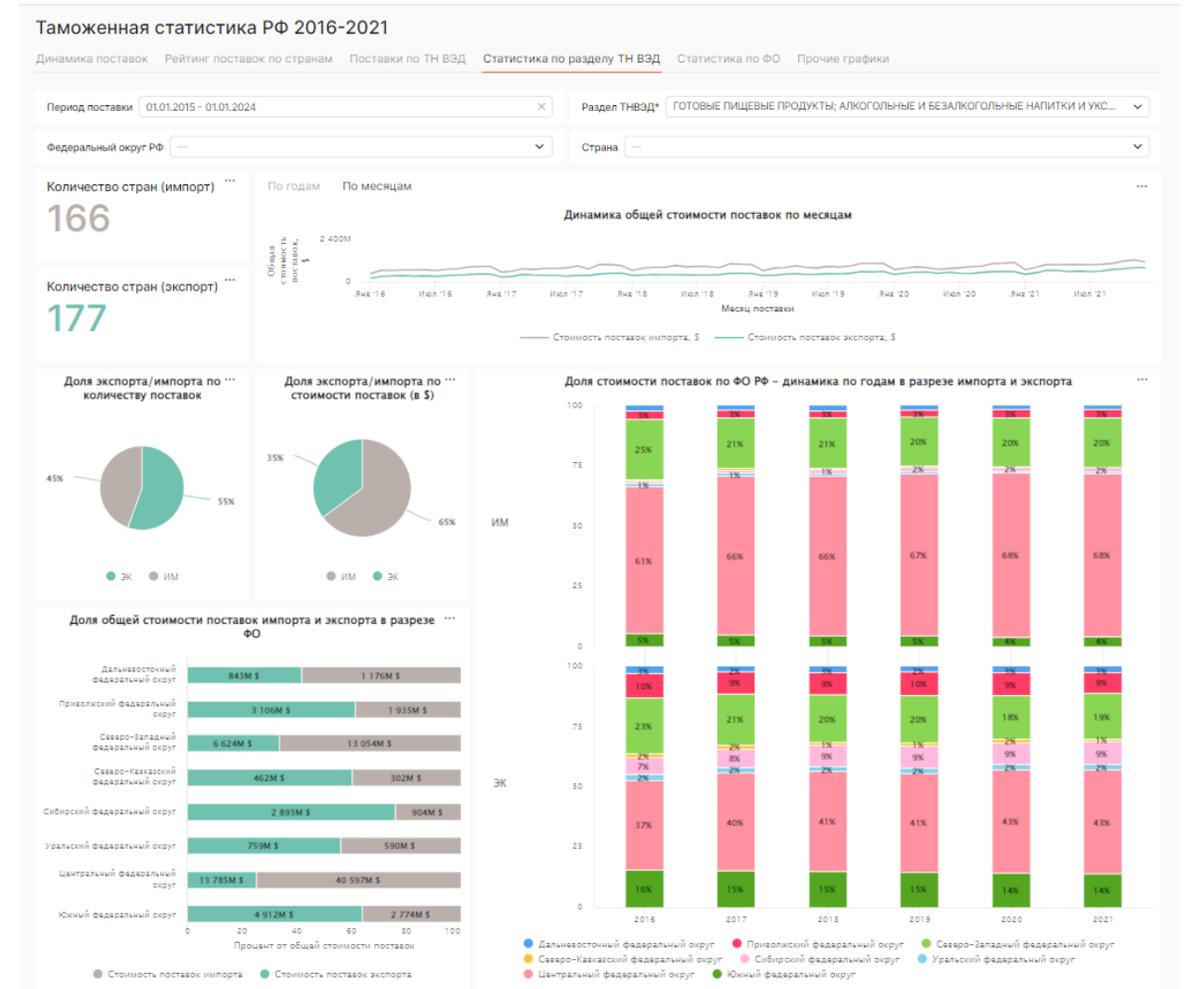
```
SELECT sum(CASE WHEN (t1.direction = 'ИМ') THEN t1.value ELSE 0 END) AS res_0
FROM datalens.fts.data AS t1
WHERE t1.dt BETWEEN '2016-01-01'::date AND '2017-01-01'::date
LIMIT 1000001
```
- Тело запроса:**

```
{
  "fields": [
    {
      "ref": {
        "type": "id",
        "id": "82be6bd0-0c71-11ef-87db-f9e35d3bfb45"
      },
      "block_id": 0,
      "role_spec": {
        "role": "measure"
      }
    }
  ],
  "filters": [
```

A 'Закреть' (Close) button is located at the bottom right of the window.

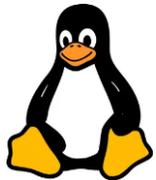
Возможности построения дашборда

- Может включать несколько страниц для размещения чартов
- Добавление разных типов виджетов – заголовки для формирования оглавления, чарты для визуализации данных, селекторы для фильтрации, текстовые поля с поддержкой языка разметки Markdown
- Гибкое управление влиянием селекторов на чарты и другие селекторы (двухстороннее, входящее, исходящее, не влияет)
- Возможность связи полей разных датасетов через алиасы
- Различные типы селекторов для фильтрации данных чарта – список, поле ввода, календарь, чекбокс
- Кроссфильтрация чартов (доступна не для всех типов диаграмм)
- Возможность добавления параметров на уровне дашборда для дополнительной фильтрации данных



DataLens в open-source версии

- Российский продукт с открытым исходным кодом, распространяемым под лицензией Apache 2.0
<https://github.com/datalens-tech/datalens>
- Возможность доработки под требования заказчика
- Установка на разные типы ОС – Linux, Windows, MacOS
- Язык разработки – Python, TypeScript, JavaScript
- Наличие сообщества на русском языке <https://t.me/YandexDataLens>
- Непрерывная доработка продукта и регулярный выпуск новых релизов



Опыт внедрения BI-систем АО АМТ-ГРУП

- Системный интегратор с опытом комплексной заказной разработки более 20 лет
- Опыт работы с построением системам отчетности и BI-решений для ОА Банк ГПБ, ДИТ города Москвы, АО «СО ЕЭС»
- Наличие демонстрационного дашборда на open source версии продукта, развернутого в собственной лаборатории