

Опыт интеграции **ETL** **Loginom** и **BI** **Visiology** в эпоху больших данных



Сергей Власов
Ведущий эксперт по
анализу данных



VISIOLOGY – ЭКСПЕРТИЗА ЗРЕЛОГО ВЕНДОРА

*Лучшая BI-платформа для
крупных предприятий с
лучшей визуализацией и
самая удобная платформа
для аналитиков**

14 лет
в визуализации
информации

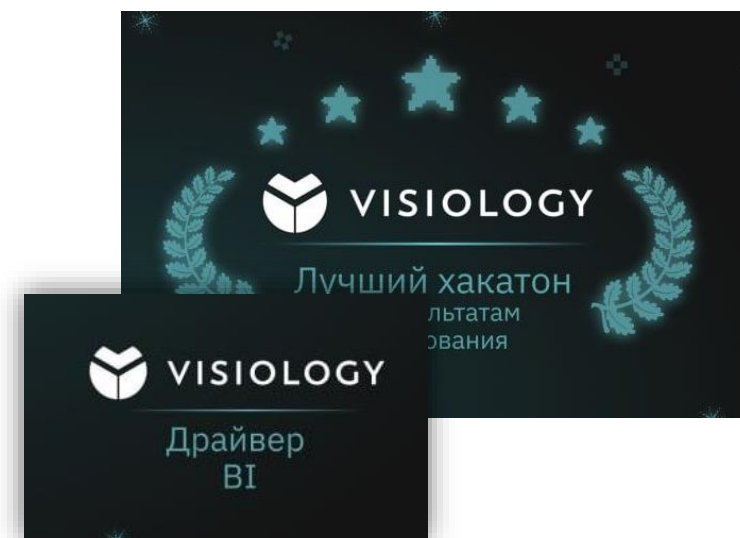
7 лет
в бизнес-
аналитике



Visiology – победитель конкурса «Лучшие цифровые решения», проведённого Аналитическим Центром при Правительстве РФ.

Visiology третий год подряд – в лидерах независимого исследования российских BI-платформ BI-КРУГ ГРОМОВА

15 ноября 2022 года Visiology победила в номинациях «Лучшее решение в рамках хакатона» и «Драйвер BI» на Битве BI, организованной Координационным центром при Правительстве РФ



200+ проектов
внедрения VI

НАШИ ЗАКАЗЧИКИ

Госуправление

Более 40 ЦУР по всей
России и СНГ



МИНИСТЕРСТВО
ЗДРАВООХРАНЕНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ



Аналитический центр Москвы



РОССИЙСКИЙ ФЕДЕРАЛЬНЫЙ
ГЕОЛОГИЧЕСКИЙ ФОНД



ФЕДЕРАЛЬНОЕ
КАЗНАЧЕЙСТВО

Росмолодежь

Коммерческие



Образование



КубГАУ

Кубанский государственный
аграрный университет
1921



Холдинги и ГК



РОСАТОМ



РОСКОСМОС



УГМК

Банки и страхование



РоссельхозБанк

КАК ПРЕВРАТИТЬ ДАННЫЕ В ЦЕННЫЕ ИДЕИ?

ЧТО НЕОБХОДИМО ДЛЯ ПРИНЯТИЯ ВЕРНЫХ РЕШЕНИЙ:



ЕДИНАЯ «ТОЧКА ПРАВДЫ»

Любой отчет, построенный на основе данных из BI, является эталонным и актуальным в любой момент времени



МГНОВЕННЫЙ ДОСТУП К ИНФОРМАЦИИ

Руководство сразу получает ответы на важные бизнес-вопросы, управленческие решения принимаются быстро, на основе фактов и цифр



ИНСТРУМЕНТЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

Вывод информации на мобильные устройства руководителя, использование предсказательных сценариев и обучаемых математических моделей

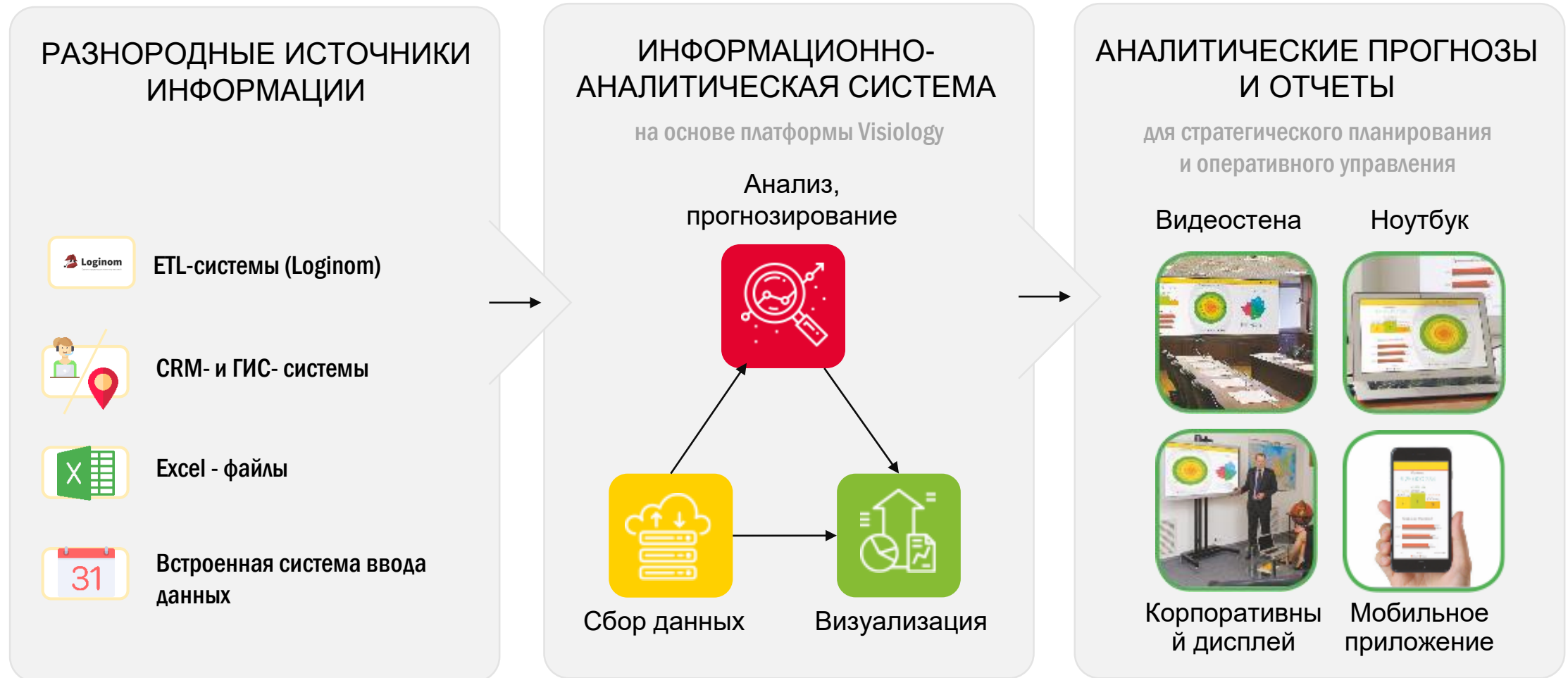


УНИВЕРСАЛЬНОЕ РЕШЕНИЕ

Сотрудники создают отчеты без привлечения ИТ-специалистов, доступ к информации разграничивается по ролевому признаку

АНАЛИТИЧЕСКАЯ ПЛАТФОРМА VISIOLOGY

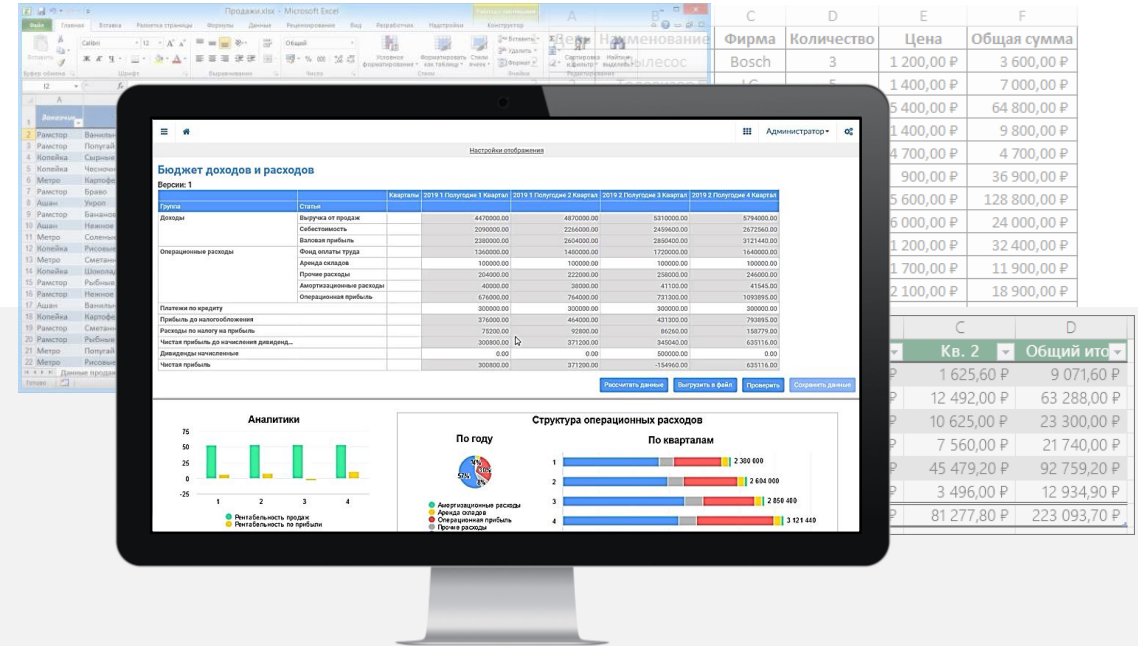
Принимайте решения на основе анализа данных



VISIOLOGY ver. 3



АНАЛИТИЧЕСКАЯ ПЛАТФОРМА VISIOLOGY



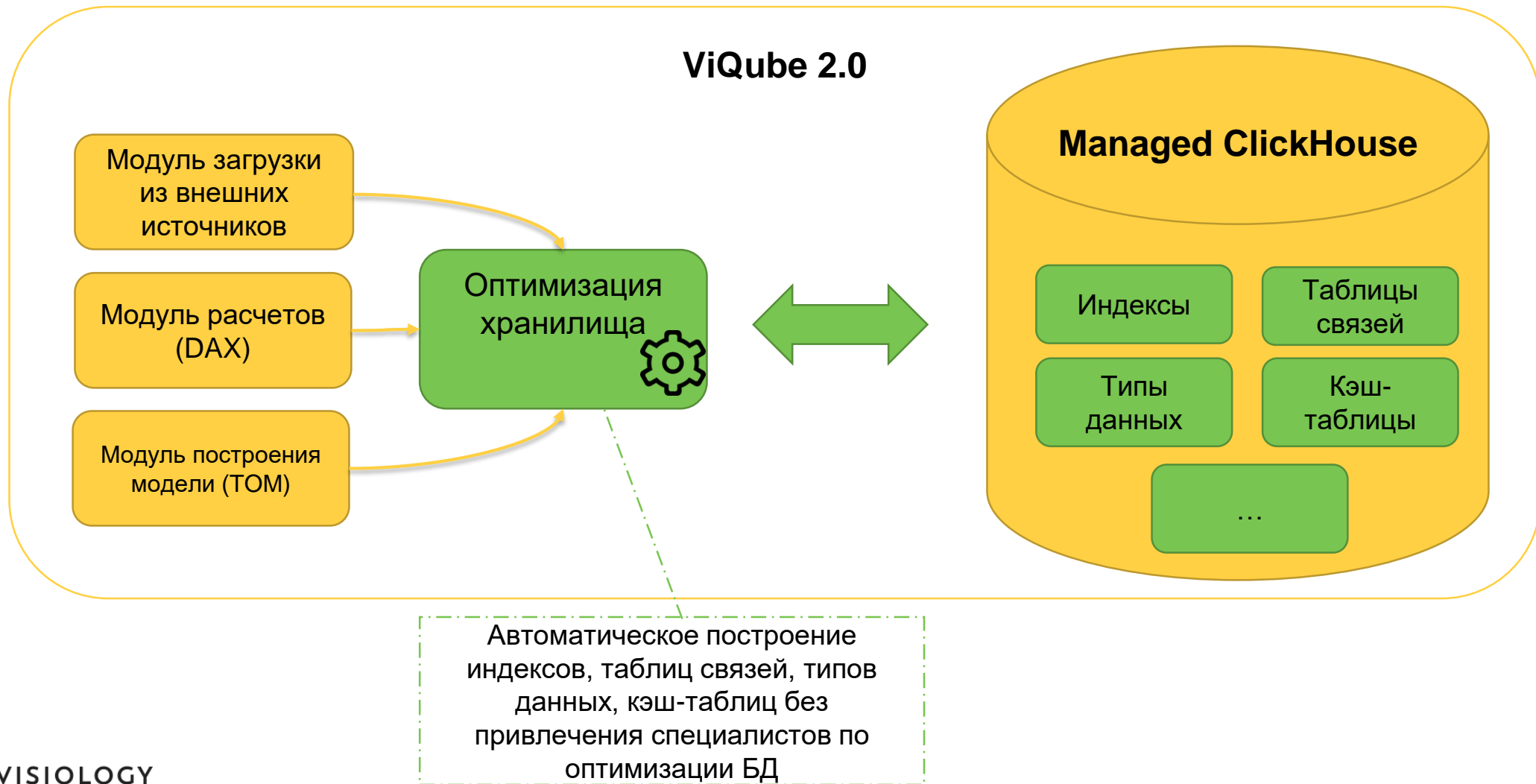
VISIOLOGY DASHBOARDS

Создавайте красивые и наглядные аналитические дашборды и регламентные отчеты для тысячи пользователей

VISIOLOGY SMART FORMS

Собирайте безошибочную отчетность за 1 день с сотен сотрудников и филиалов

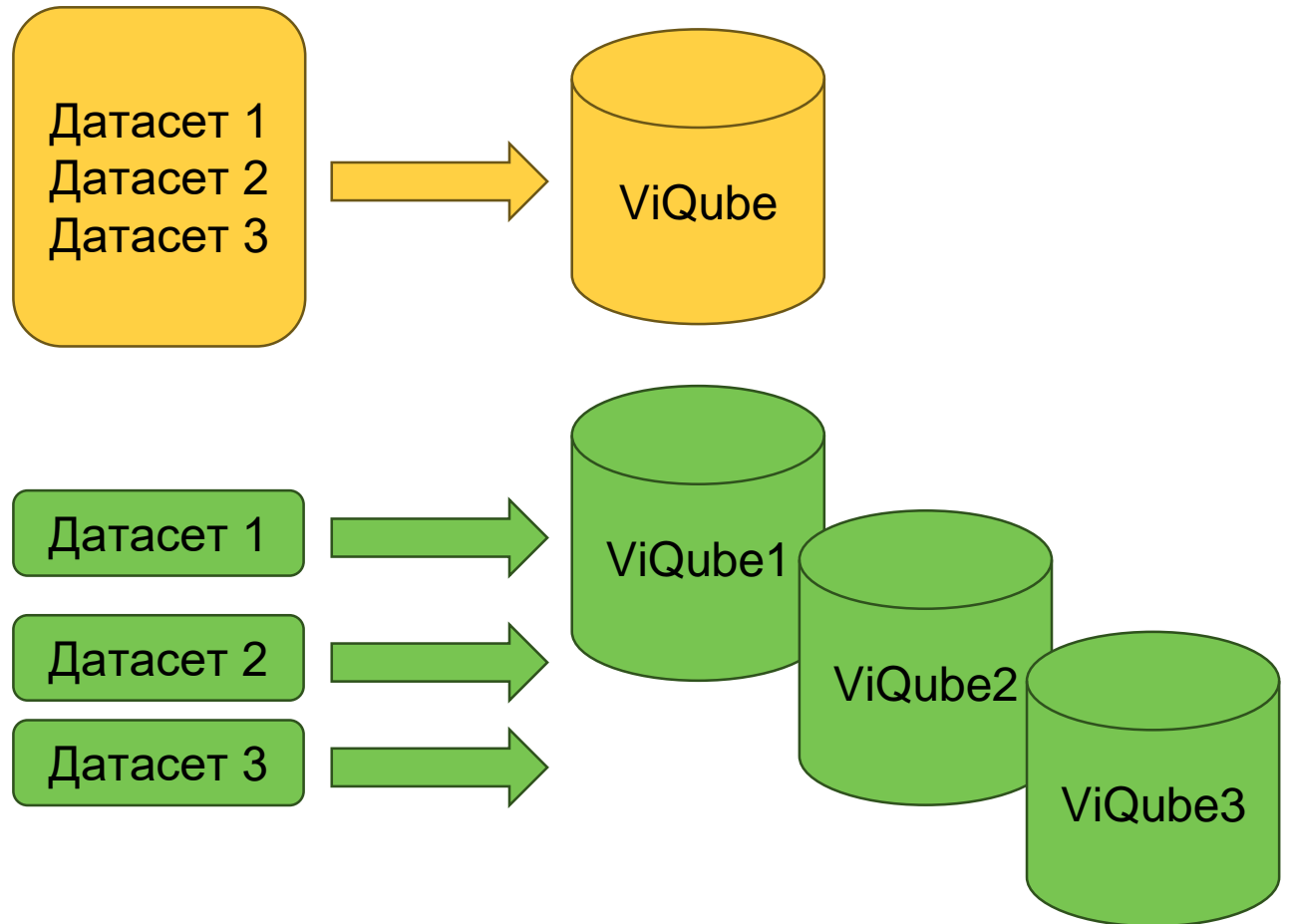
✓ Высокая скорость внутренней база данных на основе Clickhouse



✓ Увеличена скорость обработки запросов

Например, при наличии большого объема (>10ТБ) загружаемых данных и ограниченного времени на загрузку вы можете ее ускорить за счет распараллеливания загрузки по наборам данных.

✓
Увеличение скорости обработки запросов за счет распределения данных на нескольких узлах (**горизонтальное масштабирование**)

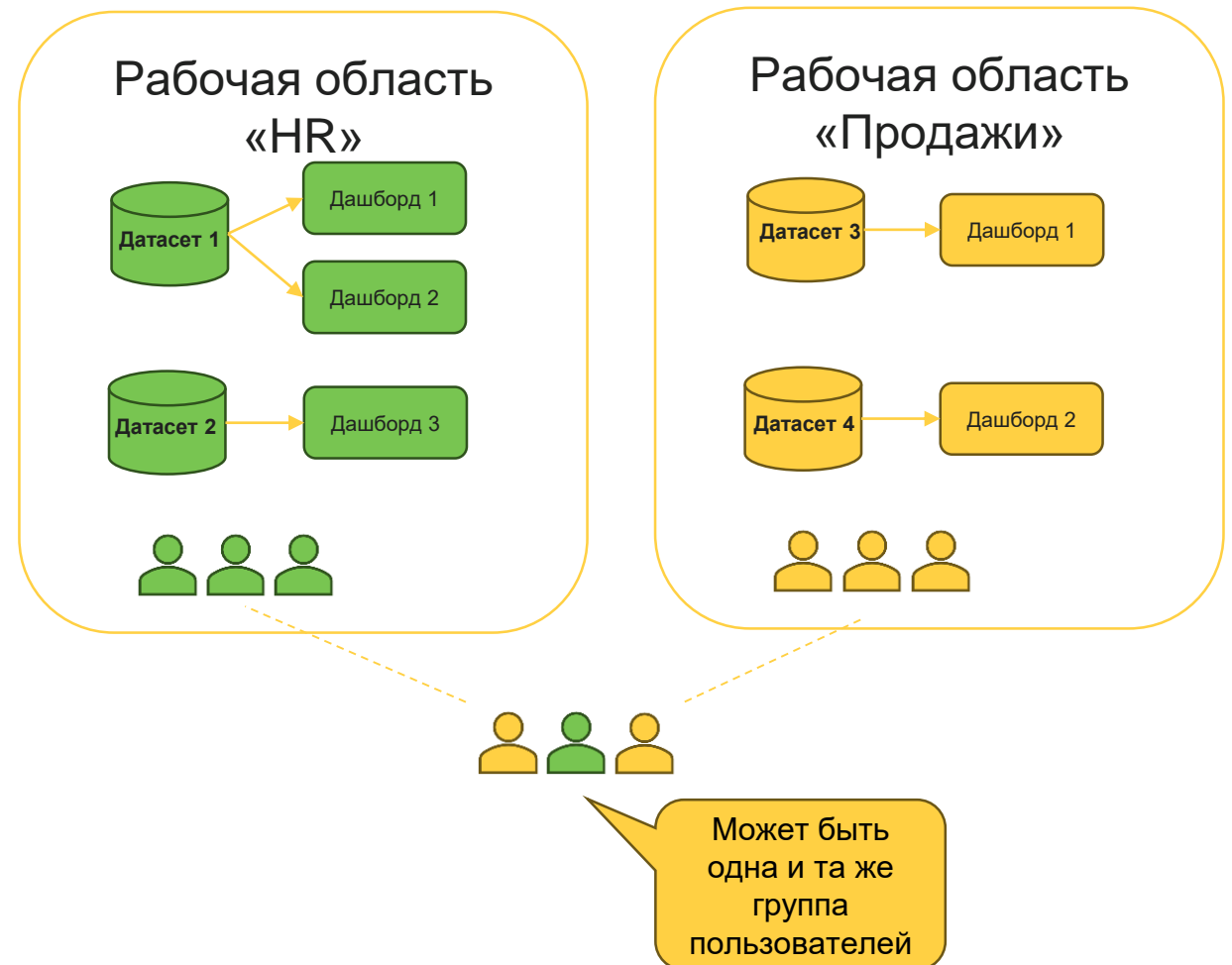


✓ Поддержка множества рабочих областей

Множество рабочих областей позволяет:

- Организовать работу разных подразделений в независимых пространствах – «Рабочих областях»
- **Разделить дашборды и датасеты по различным тематикам**
 - Разграничить доступ к данным

В списке доступны только те рабочие области, которые разрешены роли пользователя.

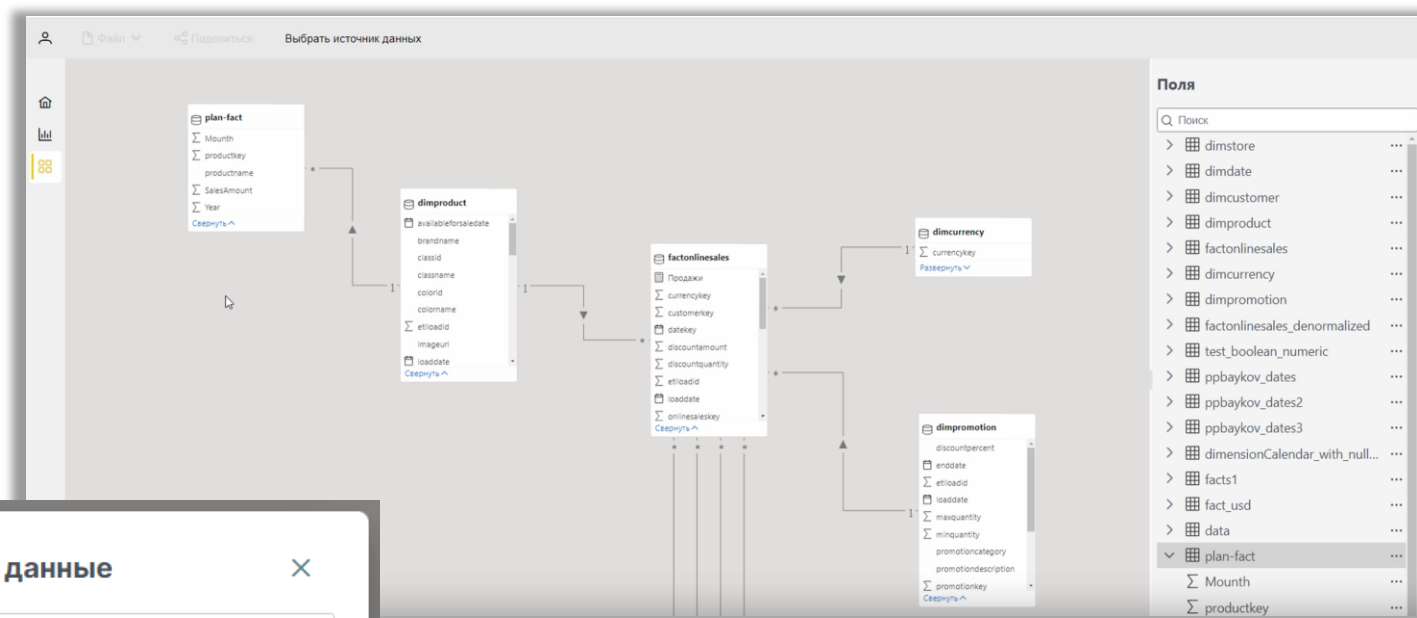


✔ «Созвездие фактов»

Появилась возможность связывать данные из разных источников через общее измерение.

Это позволяет:

- Создавать новые меры на основе фактов из разных источников
- Выводить данные фактов из разных источников на одном виджете



Получить данные

- JDBC
- Файл CSV (UTF-8)
- Книга Excel (.XLSX)

Отмена **Выбрать**



Настройка мер и возможности DAX

- ✓ RowContext
- ✓ FilterContext

Вы можете:

- Рассчитать доли на разных уровнях иерархии
- Производить расчеты с учётом фильтров по значениям измерений
- Выбирать диапазон дат

		On
		Ratio
▾ A. Datum	A. Datum Advanced Digital Camera M300 Azure	0,29
	A. Datum All in One Digital Camera M200 Grey	0,29
	A. Datum Interchangeable lens Non-SLR Digital Camera X250 Pink	0,42
A. Datum Всере		1
▸ Adventure Works		1
▸ Contoso		1
▸ Fabrikam		1

```
SUM ( factonlinesales[salesamount] )  
/ CALCULATE (  
    SUM ( factonlinesales[salesamount] ),  
    REMOVEFILTERS ( dimproduct ),  
    SUMMARIZE ( dimproduct, dimproduct[br  
andname] )  
)
```

	On		
	d_Выручка	d_Выручка БЕЗ учета промоакции	d_Объем продаж только по Жен
▸ A. Datum	1 438 032,61	965 912,31	260 093,1
▸ Adventure Works	2 937 165,29	1 926 228,72	521 101,95
▸ Contoso	4 958 670,55	3 178 175,09	745 128,21
▸ Fabrikam	3 628 308,17	2 326 167,78	80 121,75
▸ Litware	2 386 650,95	1 573 654,39	326 051,04
▸ Northwind Traders	863 765,43	556 711,23	327 768,36
▸ Proseware	1 775 387,33	1 163 480,92	48 228,95
▸ Southridge Video	990 505,39	666 635,76	186 903,41
▸ Tailspin Toys	239 413,63	159 344,89	88 790,88

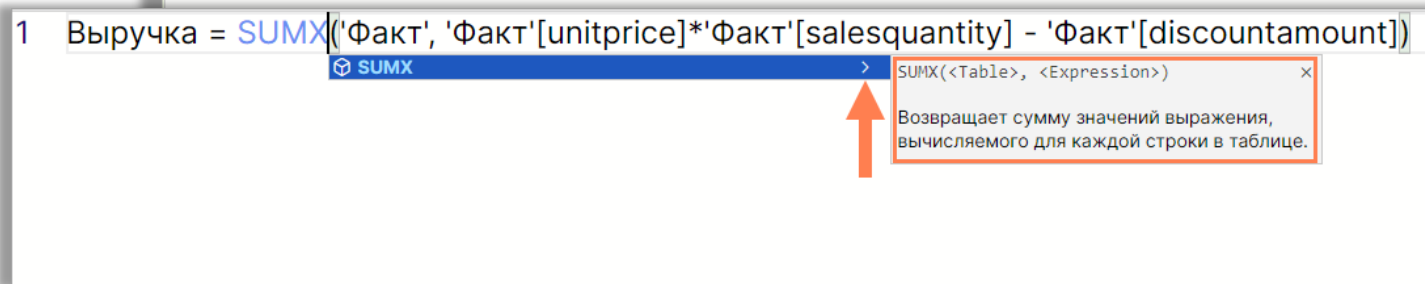
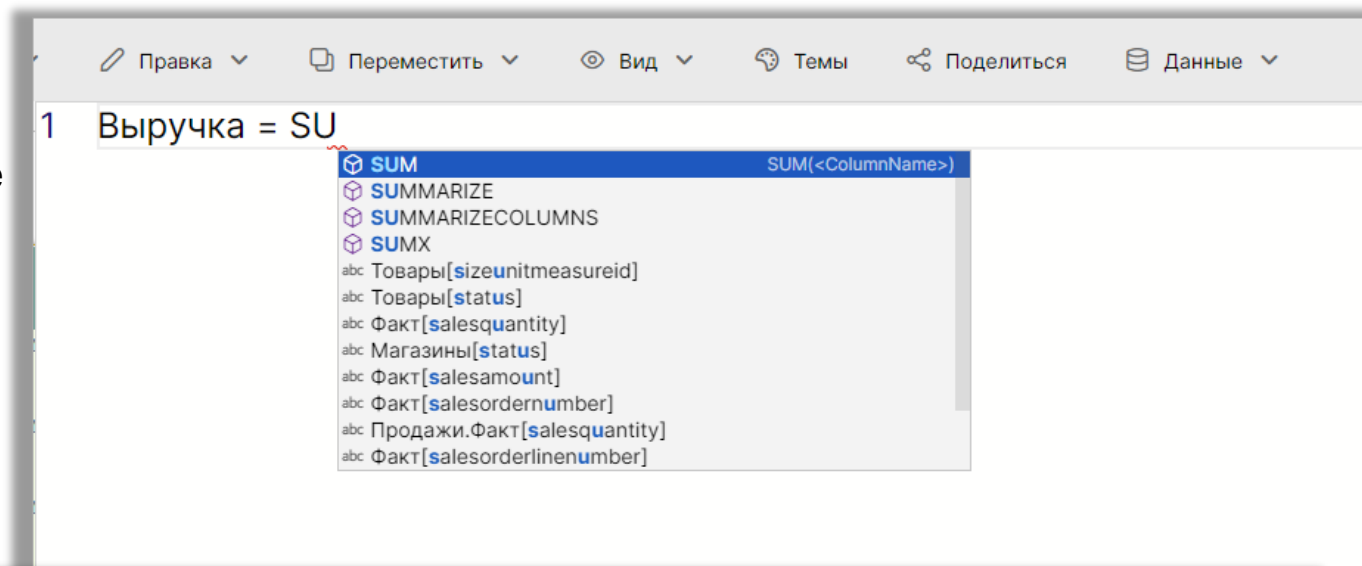
```
CALCULATE (  
    SUMX (  
        factonlinesales,  
        factonlinesales[unitprice] * factonlinesales[salesq  
uantity] - factonlinesales[discountamount]  
    ),  
    FILTER ( dimpromotion, dimpromotion[promotioncategory]  
IN { "Store" } )  
)
```

```
CALCULATE (  
    SUM ( factonlinesales[salesamount] ),  
    FILTER ( dimcustomer, dimcustomer[gender] IN { "F" } )  
)
```

✓ Помощь при наборе формул ("intellisense")

Добавлены:

- подсказки названия функций
- подсказки названия таблиц и столбцов которые можно использовать в конкретной функции
- пояснительный текст для каждой функции
- автодополнение по клику



Создавать меры стало ещё
проще и быстрее!



✓ Настройка мер и возможности DAX

- Функции DAX – полный перечень приведен в документации

The screenshot shows the documentation page for the **CALCULATE** function in the Visiology Platform 3.0. The page is in Russian and provides detailed information about the function's syntax, parameters, and usage.

Синтаксис

```
CALCULATE ( <Expression> [, <Filter> [, <Filter> [, ... ] ] )
```

Параметр	Атрибуты	Описание
Expression	Обязательный	Выражение для вычисления.
Filter	Обязательный, повторяемый	FILTER выражение.

Возвращаемое значение

Скалярное значение, которое является результатом вычисления выражения.

Примечание

FILTER переопределяет все существующие фильтры для того же столбца.

- Обязательно указывать выражение **FILTER** внутри **CALCULATE**.
- Не поддерживается вложенность.
- Мера может содержать **CALCULATE**.
- **CALCULATE** может оборачивать меру, которая не содержит **CALCULATE**.

<https://visiology-doc.atlassian.net/wiki/spaces/v3/pages/2393452/DAX>

VITALK GPT



Анна из Visiology admin
Forwarded messages

Visiology Official 13/06/23



Visiology запустила ViTalk GPT

ViTalk GPT - виртуальный ассистент от компании Visiology с искусственным интеллектом. Он прост в использовании, работает в диалоговом режиме и помогает писать DAX-запросы и ETL-скрипты для аналитической платформы.



ViTalk GPT

Привет! Я **ViTalk GPT**, виртуальный помощник Visiology. Могу помочь с **DAX**-запросами для **Visiology 3**. Знаю и учитываю все особенности Visiology при написании кода. В чем вам нужна помощь?



Ivan Vakhmyanin

Есть расчетный показатель на Qlik: `sum({$<Year={2000}, Region={"U*"}>} Sales)`
Напиши DAX запрос с таким-же смыслом.



ViTalk GPT

В DAX, чтобы получить сумму продаж для года 2000 и регионов, начинающихся на "U", вы можете использовать функции **CALCULATE** и **FILTER**.

```
Sales_2000_U_Regions =  
CALCULATE(  
    SUM('SalesTable'[Sales]),  
    FILTER('YearTable', 'YearTable'[Year] = 2000),  
    FILTER('RegionTable', LEFT('RegionTable'[Region], 1) = "U")  
)
```

В этом коде:

- `SUM('SalesTable'[Sales])` вычисляет сумму всех продаж.
- `FILTER('YearTable', 'YearTable'[Year] = 2000)` фильтрует данные так, чтобы они включали только год 2000.
- `FILTER('RegionTable', LEFT('RegionTable'[Region], 1) = "U")` фильтрует данные так, чтобы они включали только регионы, начинающиеся с "U".
- `CALCULATE` используется для применения этих фильтров к расчету суммы продаж.

Пожалуйста, убедитесь, что вы адаптировали имена таблиц и полей в соответствии с вашей схемой данных.

Помните, что в Visiology поддерживается только однонаправленная связь, поэтому убедитесь, что ваша модель данных соответствует этому требованию.

Помогает:

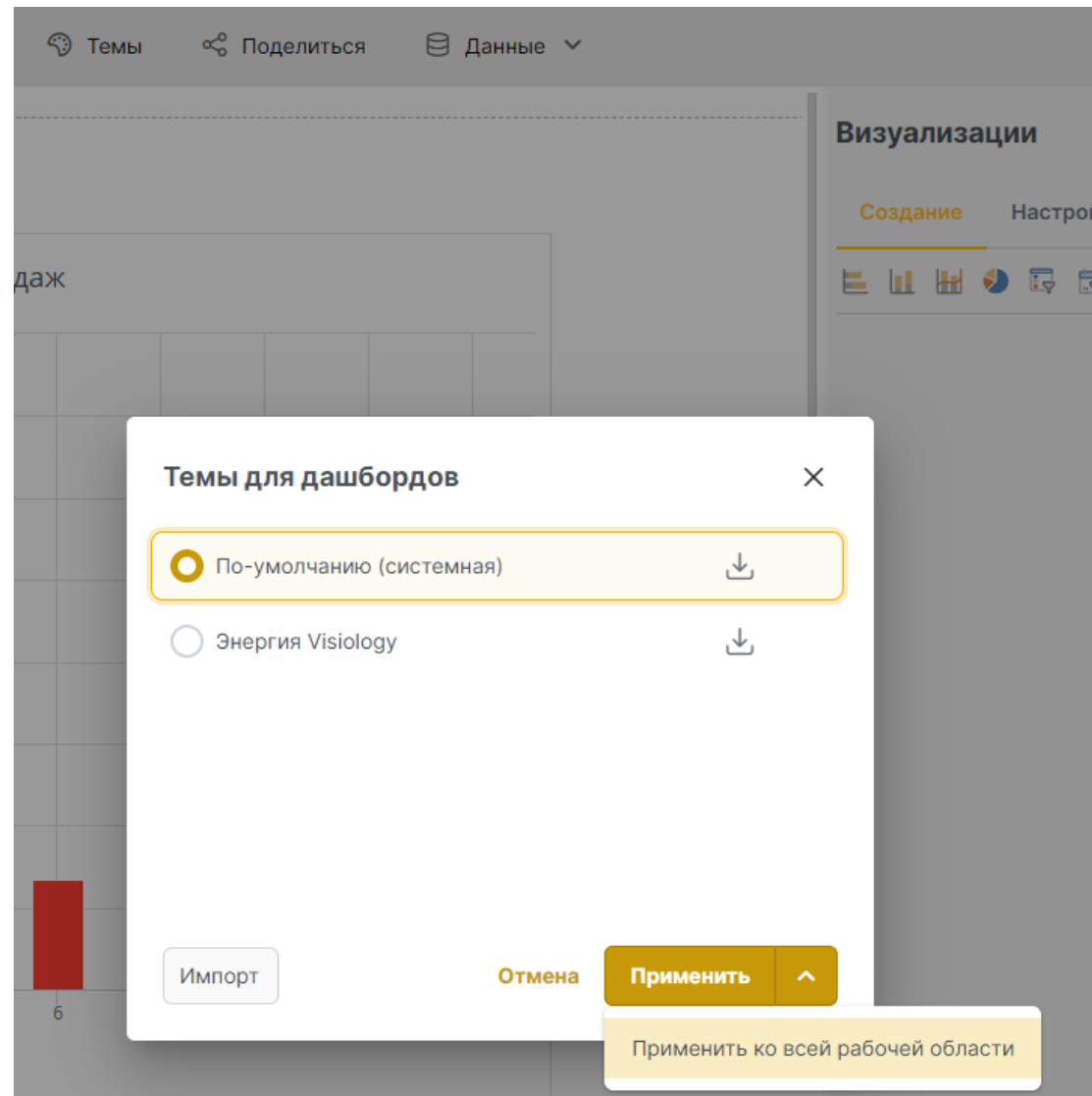
- Адаптировать запрос под Visiology 3 (с учетом ограничений)
- Подготовить запрос с другой платформы на DAX-е
- Написать запрос на DAX без знания языка на основе реальной модели
- Написать Python-скрипт для загрузки данных для ETL-инструмента ViXtract

✓ Настройка тем дашбордов в WEB

- ✓ Работа с платформой полностью в WEB! desktop больше не нужен!

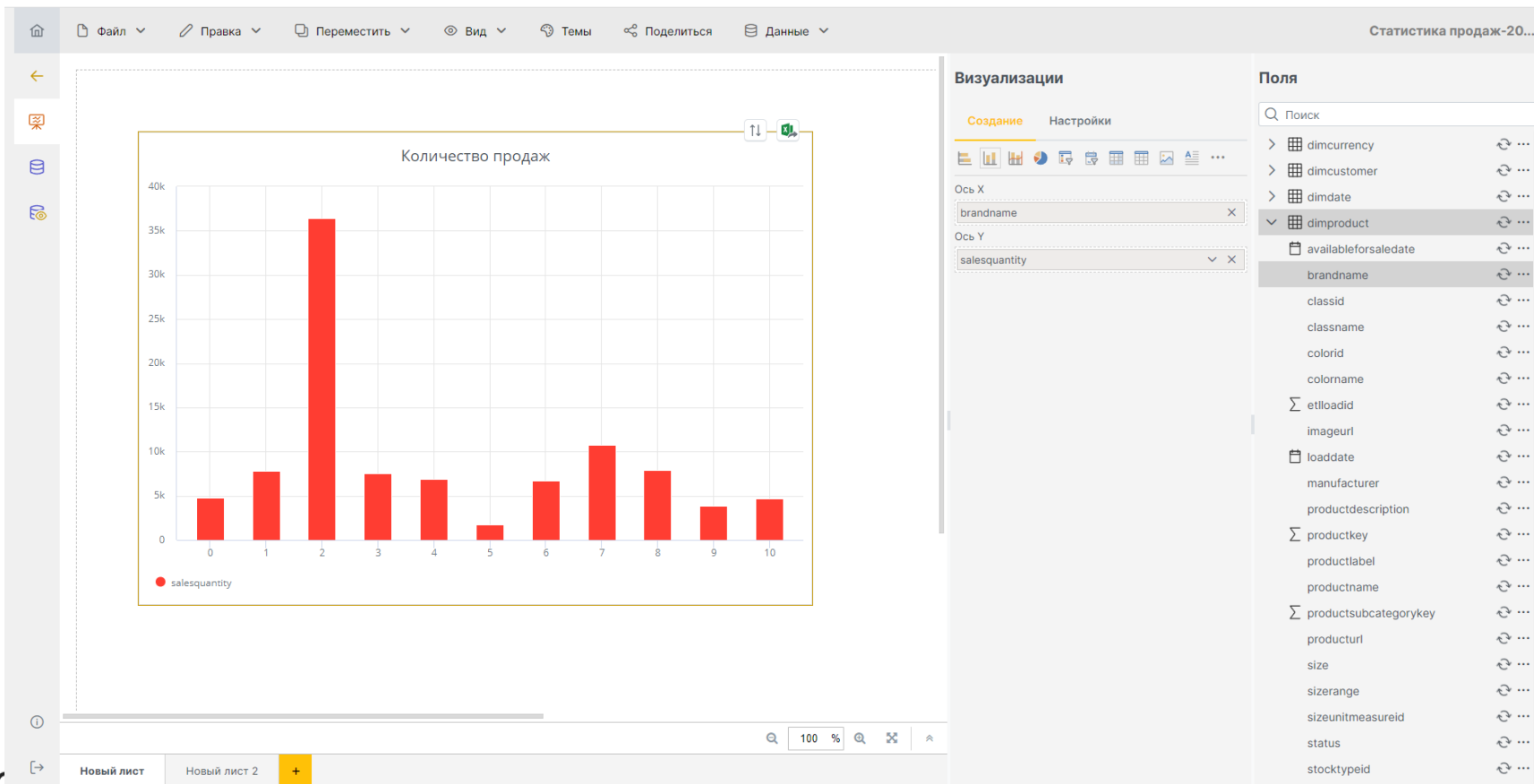
Возможности:

- Для разных рабочих областей можно настроить свои темы
- Для разных дашбордов в рамках рабочей области можно настроить свои темы



✓ Привычный интерфейс

- ✓ Привычный интерфейс для аналитиков PowerBI, не потребуются глубокие обучение, проще переход с Power BI.
- ✓ Вы можете просматривать или редактировать дашборды во время процедуры обновления данных. Это особенно актуально, если обновление данных занимает много времени.



✓ КАСТОМИЗАЦИЯ ВИДЖЕТОВ

Любое требование может быть реализовано и тиражировано на этапе реализации проекта, без привлечения вендора

Продвинутый JS API позволяет создать визуальный элемент любой сложности

Программируемый UI для переиспользования кастомных виджетов без написания кода

The screenshot shows a web application interface with a Sankey diagram titled "Объем продаж" (Sales Volume). The diagram illustrates the flow of sales from products to regions. The products listed are "Бакалея, сахар, соль" (24.9%) and "Консервы" (23.8%). The regions listed are "Барнаул" (16.5%), "Владиво" (10.4%), "Иркутск" (10.4%), and "Москва".

Below the diagram is a code editor with the following JavaScript code:

```
43 var value = Math.round(Number(innerItem)
44   _item.push(w.data.rows[index][0]);
45   _item.push(w.data.rows[index][1]);
46   _item.push(value);
47   return _item;
48 });
49
50 //Убираем нулевые значения если необходимо
51 if (!w.props.showZeros){
52   // Транспируем матрицу, фильтруем, транспируем обратно
53   _data = _zip.apply(null, [_zip.apply(null, data).filter(function(dataRow){
54     return !_every(dataRow, function(dataRow){
55       return dataRow[2] > 0
56     })
57   })
58   });
59 }
60
```

On the right side of the screenshot, there is a settings panel for the widget, including options for "Показывать название" (Show name), "Размер заголовка" (Header size), "Шрифт заголовка" (Header font), "Размер шрифта подли" (Subfont size), "Цвет шрифта подли" (Subfont color), "Шрифт подли" (Subfont), "Цвет шрифта подли" (Subfont color), "Размер шрифта подли" (Subfont size), "Шрифт подли" (Subfont), "Расширенные свойства (JSON)" (Advanced properties (JSON)), and "Цветовая палитра" (Color palette).

Разработка портала «Цифровая Статистика» для Корпорации МСП (мсп.рф)

The collage displays several screenshots from the "Digital Statistics" portal. The top left screenshot shows a map of Russia with a data table for "Сибирский округ" (Siberian Federal District) for the period 2019.02.01 - 2019.02.01. The table includes columns for "Всего" (Total), "Иркутск" (Irkutsk), "Новосибирск" (Novosibirsk), "Томск" (Tomsk), and "Челябинск" (Chelyabinsk).

The top right screenshot shows a "Статистика" (Statistics) dashboard with a bar chart and a table of data for "Московская область" (Moscow Region).

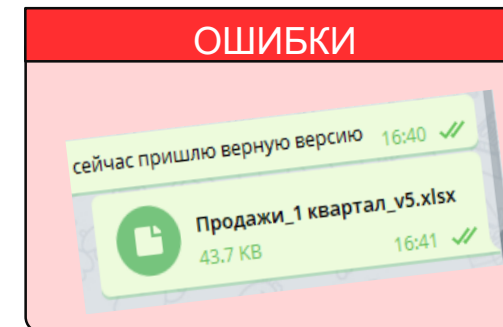
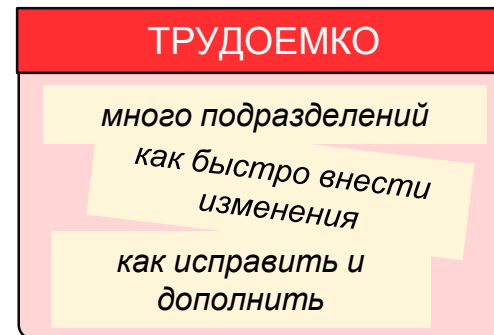
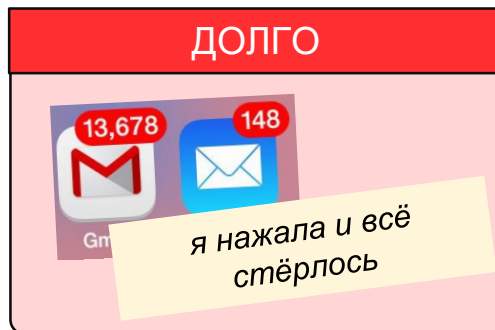
The middle left screenshot shows a "Бизнес России" (Business in Russia) dashboard with a globe and several summary cards for "Крупнейшие компании" (Largest companies), "Средняя зарплата" (Average salary), "Число МСП" (Number of MSPEs), and "Оборот МСП" (MSPE turnover).

The middle right screenshot shows a "Статистика по РФ" (Statistics by RF) dashboard with a globe and several summary cards for "ВВП" (GDP), "Индекс цен производителей" (Producer Price Index), "Средняя зарплата" (Average salary), and "Валовой региональный продукт" (Gross regional product).

The bottom left screenshot shows a "Сводные показатели" (Summary indicators) dashboard with several summary cards for "Число МСП" (Number of MSPEs), "Оборот МСП" (MSPE turnover), "Средняя зарплата" (Average salary), and "Топ 3 по давлению" (Top 3 by pressure).

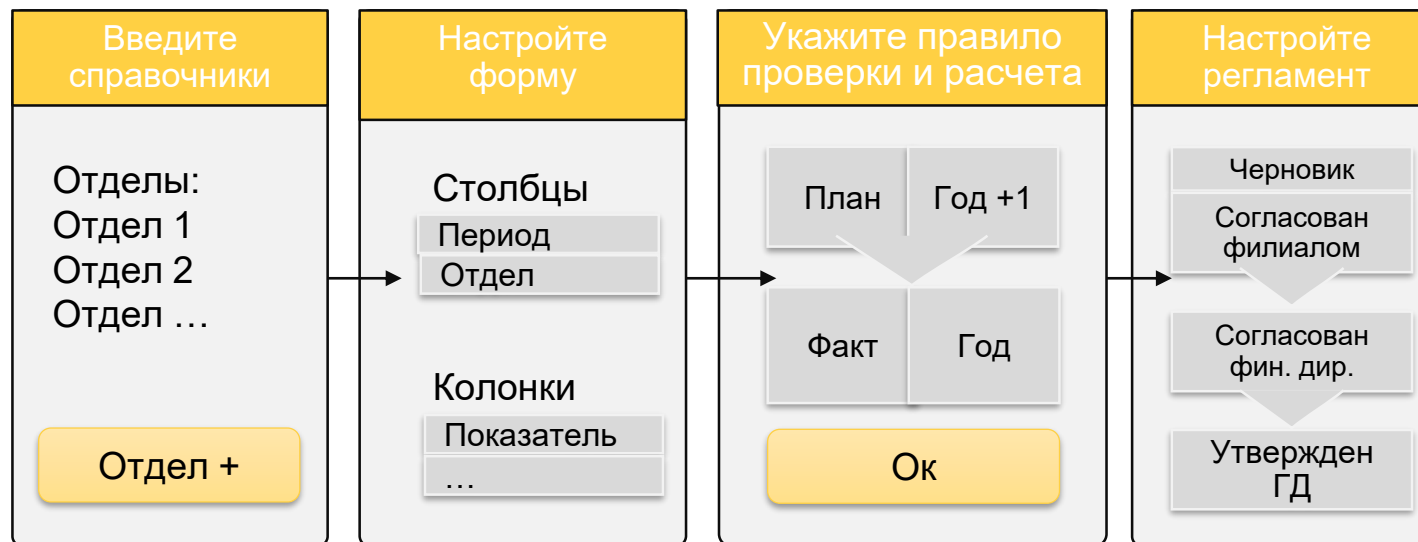
VISIOLOGY SMART FORMS

ПРОБЛЕМЫ СБОРА ДАННЫХ В EXCEL



Что такое Smart Forms? Это модуль для гибкой, прозрачной, управляемой автоматизации сбора отчетности по подразделениям и с контрагентов

- 1** ГИБКОСТЬ (AGILE) – создавайте формы без программирования и меняйте в режиме онлайн
- 2** ПРОЗРАЧНОСТЬ – подробный аудит действий пользователя и истории формирования значений в отчёте
- 3** УПРАВЛЯЕМОСТЬ – уровни доступа, бизнес-процессы согласования, автопроверки



VISIOLOGY SMART FORMS

С помощью данного компонента можно быстро реализовывать такие задачи как, например:

- ✓ Сбор отчётности с различных подразделений
- ✓ Управление нормативно-справочной информацией (НСИ)
- ✓ Автоматизация процессов планирования и бюджетирования
- ✓ Минимизировать риски и ошибки при заполнении форм
- ✓ Сделать процесс сбора данных максимально прозрачным и эффективным

ВАЖНОСТЬ ETL

Преобразование
данных

80%

Анализ и
визуализация

КАКИМ ДОЛЖЕН БЫТЬ ETL ИНСТРУМЕНТ?



ПРОСТОЙ для ИЗУЧЕНИЯ

Кривая входа не должна быть крутой



НА ВСЕ ЕСТЬ КОННЕКТОРЫ

Минимум повторной разработки, для всех систем должны быть готовые подключения



УНИВЕРСАЛЬНЫЙ

Не надо переучиваться



НЕ ОГРАНИЧИВАЕТ РАЗВИТИЕ

Усложнение кейсов не требует замены инструмента



НЕДОРОГОЙ или БЕСПЛАТНЫЙ

Не временно, а навсегда

КЕЙС POSITIVE TECHNOLOGIES



Постановка



- 1** BI платформа для внутренней отчетности РТ
- 2** BI платформа для возрождения продуктов РТ SIP и РТ RP
- 3** Российский производитель или Open Source решение
- 4** Низкий уровень зависимости от поставщика

Просмотренные участники и кандидаты

- 1. Fsight (Форсайт)**
- 2. Proceset**
- 3. Loginom**
- 4. Modus BI**
- 5. Metabase**
- 6. Polymatica**
- 7. Apache superset**
- 8. Visiology**



- 1. Visiology**
- 2. Apache Superset**
- 3. Fsight (Форсайт)**

Всего 8 решений и 11 вендоров

Возможная архитектура корп. ВІ

Источники данных



Инструменты извлечения данных (ETL)



Или



Хранение и очистка данных (DWH)



Или



Или



Или



Витрины данных



Финальная архитектура

Источники данных



Creatio



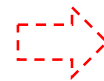
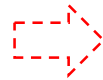
Инструменты извлечения данных (ETL)



Хранение и очистка данных (DWH)



Витрины данных



Задачи проекта



- Создать Корпоративное Хранилище Данных (КХД \ DWH)
- Победить Разрозненность данных.
- Автоматизировать сбор данных.
- Простота доступа к данным (Демократизация данных)
- Построение BI витрин с возможностью самостоятельного развития их в подразделениях (Self-service)
- Единая версия правды
- Наличие полноценного Data Quality контроля
- Создание новых источников

Переход к Data-driven подходу в управлении бизнесом!

1

Разработка

Опираясь на визуализацию промежуточных результатов, лиды и менеджеры могут эффективно перераспределять ресурсы между разработчиками и командой поддержки.

2

Продажи

По запросам бизнеса происходит декомпозиция воронки продаж по отраслям и продуктам. BI-платформа позволяет в реальном времени отслеживать динамику выручки, а также ее соответствие запланированным показателям. Благодаря этому руководители могут оперативно корректировать стратегию продаж.

3

Работа с кадрами

Визуализация процесса набора персонала и динамические визуализации опросов среди сотрудников являются важными инструментами для HR.

4

Поддержка пользователей

Визуализация количества задач по типам и клиентам, оперативное реагирование на отклонение от заявленных SLA и усиление команд еще при первых признаках перегрузки.

ОБРАЗОВАНИЕ



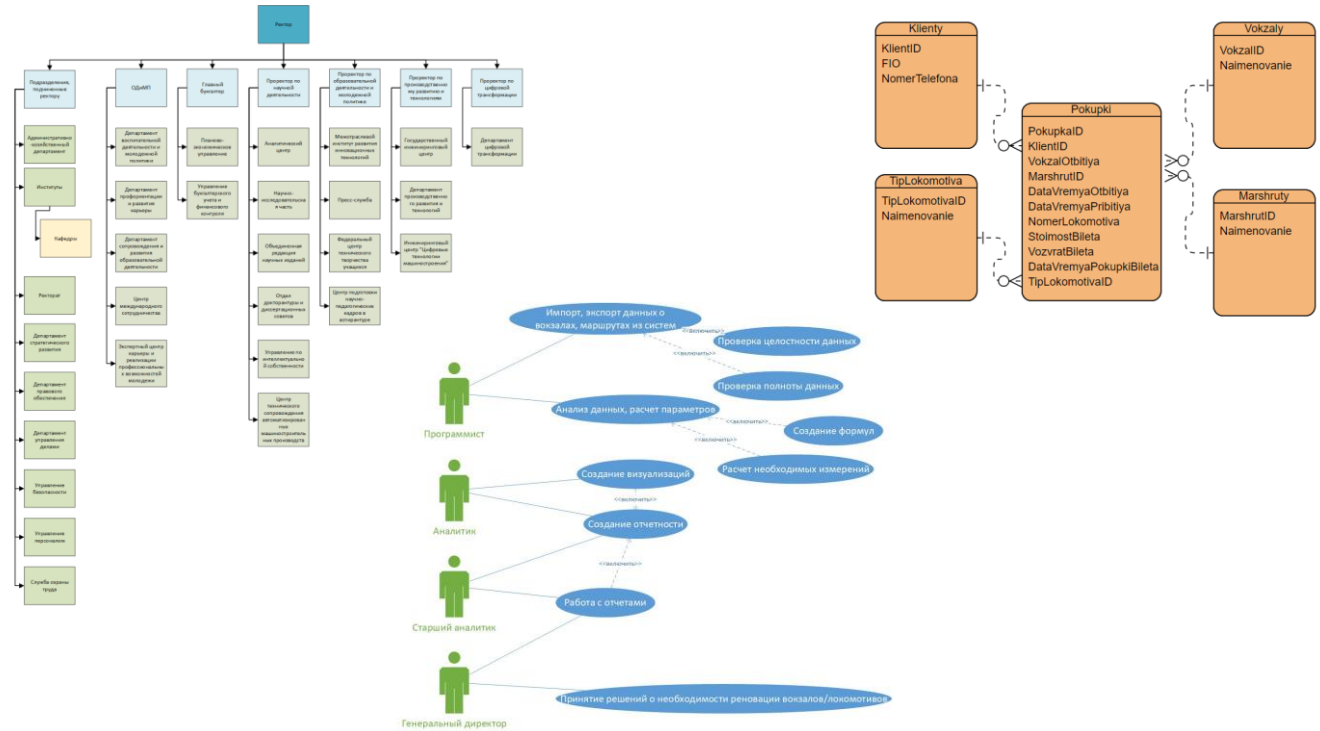


РОССИЙСКИЙ ЭКОНОМИЧЕСКИЙ УНИВЕРСИТЕТ имени Г. В. ПЛЕХАНОВА

Кафедра Прикладной информатики и информационной безопасности

50 студентов в потоке

Курсовые работы





РОССИЙСКИЙ
ЭКОНОМИЧЕСКИЙ
УНИВЕРСИТЕТ
имени Г. В. ПЛЕХАНОВА

Кафедра Прикладной
информатики и
информационной
безопасности

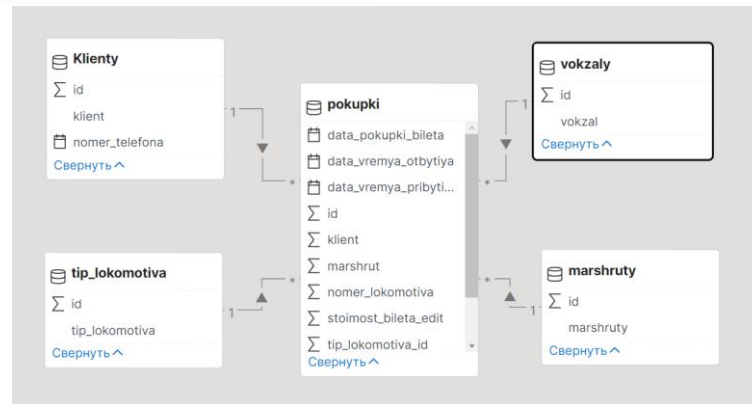
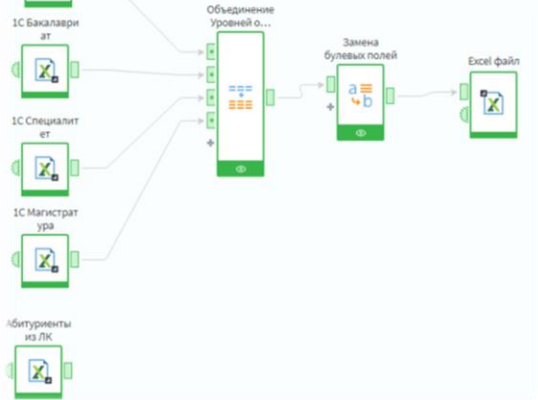
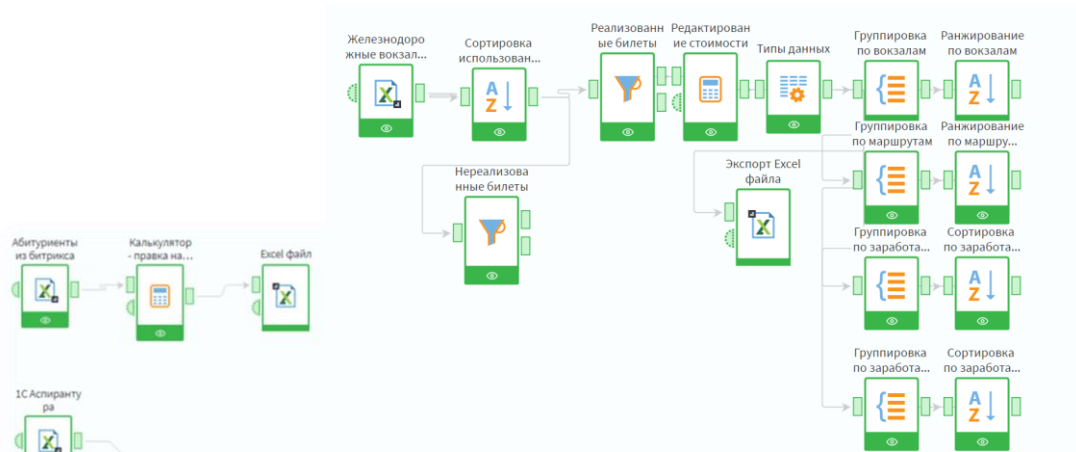




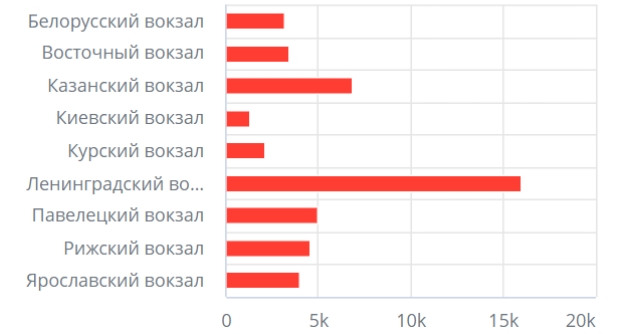
РОССИЙСКИЙ
ЭКОНОМИЧЕСКИЙ
УНИВЕРСИТЕТ
ИМЕНИ Г. В. ПЛЕХАНОВА

Кафедра Прикладной
информатики и
информационной
безопасности

Примеры работ



Линейчатая диаграмма

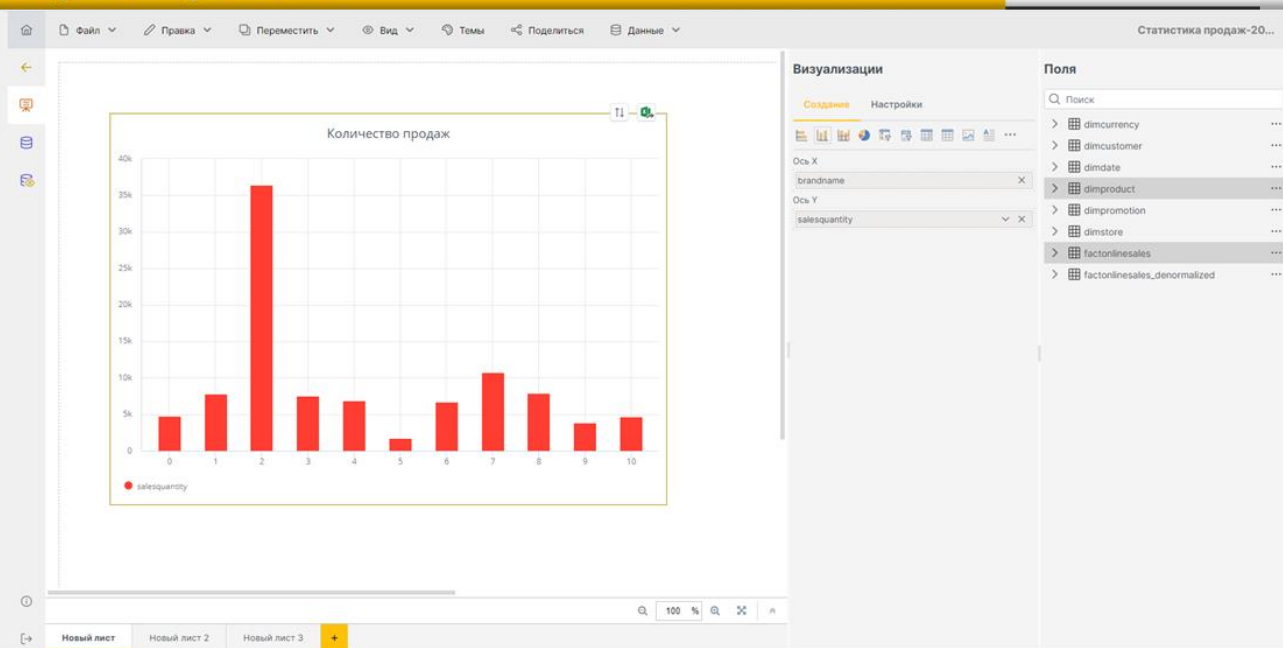


Круговая диаграмма



VISIOLOGY 3

Пример DAX



```
SUMX ( Sales, Sales[Price] - Sales[TotalCost] )  
/ CALCULATE (  
    SUMX ( Sales, Sales[Price] - Sales[TotalCost] ),  
    PREVIOUSMONTH ( 'Date'[Date] )  
)
```



Поддержка аналитического языка запросов DAX, меньше расчетов на уровне витрин DWH, меньше JS и Python кода, большой шаг в сторону Self-Service.



Привычный интерфейс для аналитиков PowerBI, не потребуется глубокое обучение, проще переход с Power BI.



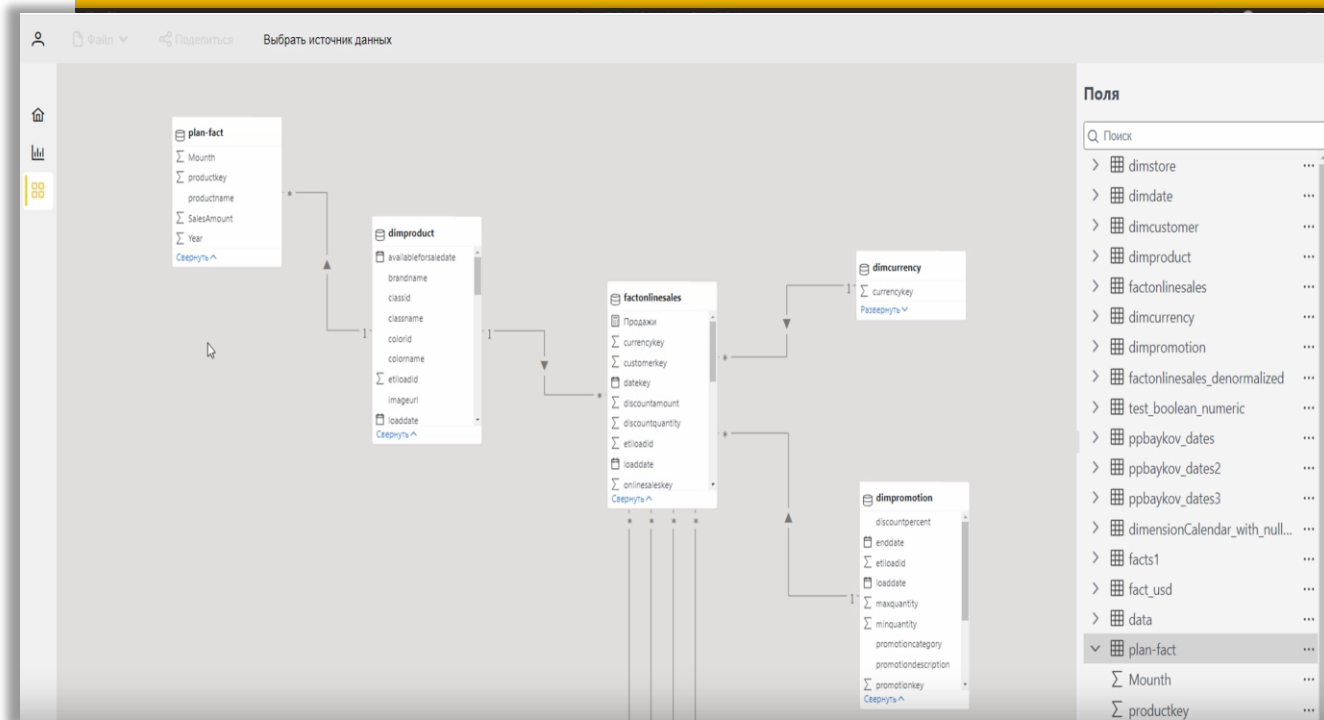
Визуальная модель данных, работающая по принципам PowerBI, которую очень просто настроить.



Незаметная смена интерфейса для бизнес-пользователя

VISIOLOGY 3

Пример модели данных



Гарантированно быстрая аналитика

Быстродействие не зависит от качества построения аналитического хранилища.

Быстрее бизнес результат: экономия времени и ресурсов на построение внешнего хранилища.

Комфортная работа пользователей на млрд'ах строк без трудозатрат на построение и оптимизацию высокоскоростного внешнего хранилища.



ViQube 2.0 с ядром на основе Clickhouse



Дизайнер модели данных и дашбордов в виде Web приложения



Поможем сделать первый шаг
к управлению
на основе данных

<https://ru.visiology.su/>



@visiology_official



VISIOLOGY