

Руководство по библиотеке компонентов для Loginom Data Monetization Pack (DMP)

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

О библиотеке.....	4
Обучение.....	4
Формат поставки библиотеки.....	5
Установка, настройка и обновление библиотеки.....	5
Установка.....	5
Указание лицензионного ключа.....	7
Настройка подключения к СУБД Clickhouse.....	9
Настройка Telegram-бота.....	11
Обновление библиотеки.....	12
Основные сценарии работы с библиотекой.....	13
Готовые шаблоны сценариев.....	13
Как правильно выгружать данные в КХД.....	15
ETL-процесс для новой таблицы в core-уровне КХД.....	17
Нейминг полей и меток для core-уровня.....	22
Рекомендации по организации ETL-сценариев множества таблиц.....	25
Оптимизация потребления памяти цепочками обновлений КХД.....	29
Clickhouse Kit.Igp.....	32
Шаблоны сценариев.....	32
📁 CH: DM / Stage Выгрузка.....	32
📁 CH: DM / Stage Выгрузка и публикация.....	32
📁 CH: DM / Stage Выгрузка с партициями.....	33
📁 CH: Core выгрузка.....	33
📁 CH: Core выгрузка ETL партициями.....	34
📁 CH: Core внешняя таблица.....	35
📁 CH: Загрузка таблицы +.....	35
📁 CH: Перенос локального КХД.....	36
📁 CH: ОВТ сборка.....	36
📁 CH: ОВТ с партициями.....	41
Базовые компоненты.....	43
CH: Создать/адаптировать МТ таблицу.....	43
🔑 CH: Создать денорм. звезду.....	45
🔑 CH: Публикация таблицы.....	48
CH: Загрузить таблицу.....	49
CH: Календарь модели.....	51
CH: Партиции таблицы.....	52
CH: Перенос партиций.....	52
CH: Удалить таблицу.....	52
Data Preparation Kit.Igp.....	53

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Базовые компоненты.....	53
1. Аудит таблицы для КХД (АТ).....	53
2. Экспорт мета-данных.....	56
3.1 Список таблиц (мета-справочник).....	57
Data Control Kit.lgp.....	58
Шаблоны сценариев.....	58
📁 Аудит качества данных.....	58
Базовые компоненты.....	58
🔑 Отчет DQ.....	58
SQLite Kit.lgp.....	59
Шаблоны сценариев.....	59
📁 Экспорт в лок. КХД.....	59
📁 Экспорт в SQLite.....	59
📁 SQLite: Сборка OBT.....	59
Базовые компоненты.....	60
SQLite: создать / адаптировать таблицу.....	60
SQLite: SQL-запрос.....	60
SQLite: генерация модели.....	61
ETL Kit.lgp.....	62
DE: 🔑 ABC комплементарный.....	62
DE: 🔑 RFM.....	68
DE: 🔑 Факторный анализ.....	70
DE: Факторный анализ - интерпретация.....	74
DE: 🔑 Стабильность продаж.....	74
DE: 🔑 Товарные рекомендации.....	76
DE: 🔑 Воронка Lite.....	78

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

О библиотеке

Data Monetization Pack - библиотека, созданная для упрощения разработки систематизированной аналитики, простой организации универсального аналитического хранилища (на базе Clickhouse или Arenadata Quick Marts) с возможностью переиспользования данных в любых BI и аналитических инструментах.

Главная цель библиотеки - ускорить и упростить разработку, внедрение и дальнейшее развитие корпоративных аналитических продуктов. Таких как BI-аналитика, личные кабинеты, системы принятия решений и др. А также автоматизировать многочисленные рутинные задачи, связанные с документированием и созданием моделей данных.

Обучение

Работа с DMP подразумевает наличие у разработчика базовых навыков работы с Logiном. Получить их можно на наших бесплатных открытых программах.

Если вы изучаете Logiном с нуля, то вам нужно пройти следующие обучающие программы:

- Марафон [“Генеральная уборка данных”](#) для изучения основ обработки данных в Logiном, а также очистки данных для аналитики;
- Марафон [“Рецепт продвинутой аналитики”](#) для понимания механики создания и использования дополнительных библиотек компонентов;

Если у вас есть опыт работы с Logiном и вы понимаете, как работают подмодели, производные компоненты, ссылки на пакеты, настройки подключения к базам данных, управление последовательностью выполнения узлов, тогда можете осваивать сразу DMP:

- Марафон [“На высоту enterprise аналитики”](#).

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Формат поставки библиотеки

Библиотека поставляется как набор пакетов Loginom, а также рядом вспомогательных файлов (примеры, настройки), распределенных по каталогам. Пакеты, содержащие компоненты, зашифрованы, и не могут быть открыты или изменены. Можно использовать ссылки на них.

Examples	25.10.2024 11:29	Папка с файлами	
Log	25.10.2024 11:29	Папка с файлами	
Meta Data	25.10.2024 11:29	Папка с файлами	
Sample Data	25.10.2024 11:29	Папка с файлами	
Settings	25.10.2024 11:29	Папка с файлами	
Clickhouse Kit.Igp	25.10.2024 11:28	Loginom Package ...	579 КБ
Data Control Kit.Igp	25.10.2024 11:28	Loginom Package ...	133 КБ
Data Preparation Kit.Igp	25.10.2024 11:28	Loginom Package ...	240 КБ
ETL kit.Igp	25.10.2024 11:28	Loginom Package ...	18 КБ
License.db	15.08.2024 17:12	Data Base File	8 КБ
License.Igp	25.10.2024 11:28	Loginom Package ...	58 КБ

Установка, настройка и обновление библиотеки

Установка

Библиотека должна быть установлена путем копирования каталога библиотеки в файловое хранилище Loginom. Рекомендуется создать отдельную рабочую папку для размещения в ней библиотеки и предоставления к ней доступа.

Также, необходимо разместить в рабочей папке библиотеку API Pack.

[Скачать готовый набор библиотек можно тут.](#)

Также, в рабочей папке должны находиться библиотеки от Loginom (они уже есть в готовом наборе):

[File Api Kit](#), [Silver Kit](#), [JSON Kit](#)

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Имя	Дата изменения	Тип	Размер
admin	04.12.2024 15:44	Папка с файлами	
alberto	20.11.2024 17:09	Папка с файлами	
AnastM	14.11.2024 16:01	Папка с файлами	
CSV	22.07.2024 15:31	Папка с файлами	
Data	22.07.2024 15:31	Папка с файлами	
ETL_processes	10.12.2024 8:19	Папка с файлами	
Examples	10.09.2024 16:45	Папка с файлами	
Libs	25.11.2024 15:52	Папка с файлами	
oat	12.11.2024 11:15	Папка с файлами	
OEM_prod	10.12.2024 8:19	Папка с файлами	
Qlik	23.07.2024 14:16	Папка с файлами	
service	25.03.2024 15:21	Папка с файлами	
services	22.11.2024 14:44	Папка с файлами	
user	25.03.2024 15:21	Папка с файлами	
viewer	25.03.2024 15:21	Папка с файлами	

Пример рабочей папки в файловом хранилище

Имя	Дата изменения	Тип	Размер
API Pack	10.12.2024 13:11	Папка с файлами	
Data Monetization Pack	10.12.2024 13:11	Папка с файлами	
fileapi_kit	10.12.2024 13:11	Папка с файлами	
json_kit	03.08.2024 13:41	Папка с файлами	
silver_kit	10.12.2024 13:11	Папка с файлами	

Расположение библиотек в рабочей папке

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

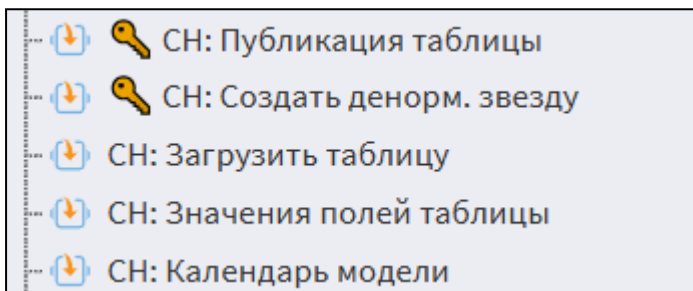
<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

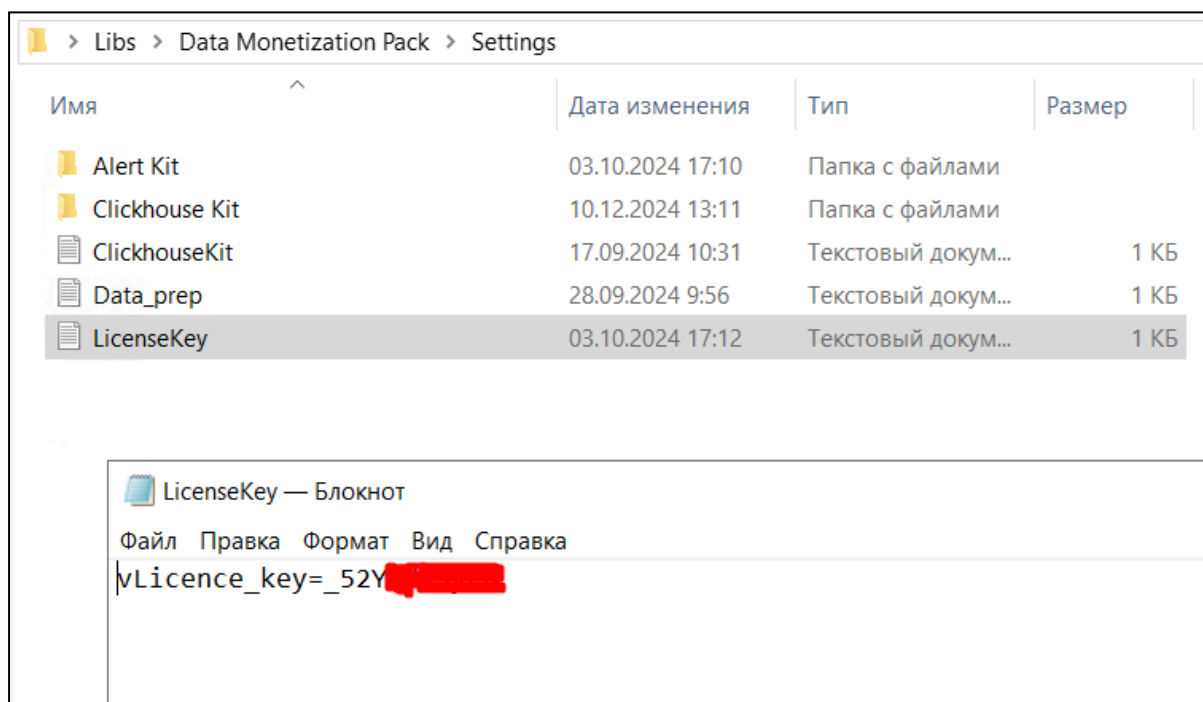
Указание лицензионного ключа

У DMP есть ряд глобальных настроек, которые задаются в папке Data Monetization Pack/Settings. На текущем этапе нужно указать ключ (приобретается у поставщика).

Без указанного ключа, некоторые компоненты пакета будут работать с ограничениями. Такие компоненты отмечены знаком ключа.



Ключ указывается в файле Data Monetization Pack/Settings/LicenseKey.txt



Сохраните файл и закройте его.

Вы можете проверить актуальность лицензии в пакете Data Monetization Pack\License check.lgp.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Активируйте компонент “Лицензия”, и откройте предпросмотр второго порта данных.

Если значение поля “Статус лицензии” = true, То значит ваша лицензия активна.

Иначе, лицензия не активна, и компоненты будут работать в демо-режиме. В поле “Описание статуса” можно прочитать, по какому критерию произошел сбой проверки.

The diagram shows a central component labeled "Лицензия" (License) with a play button icon. To its left, a vertical stack of ports is shown with values 44, 32, 3, 10, and 0. A line connects the port with value 10 to the license component. To the right of the license component, three ports are shown with values 1, 1, and 1. A red arrow points to the middle port with value 1.

Параметры лицензии		Определение доступа		Статус	
#	0/1 Статус лицензии	ab	Описание статуса		
1	true		Валидный ключ Валидный пользователь/сервер Актуальный ср...		

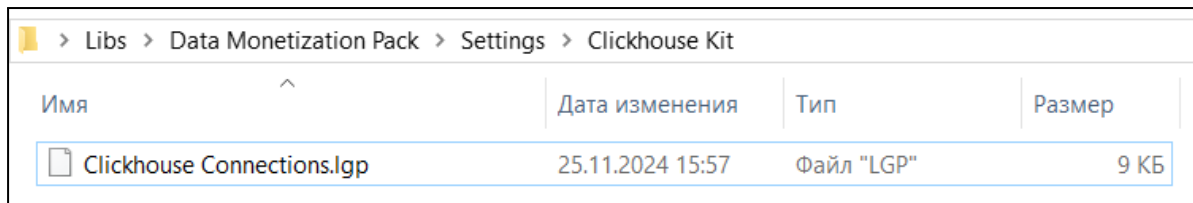
<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

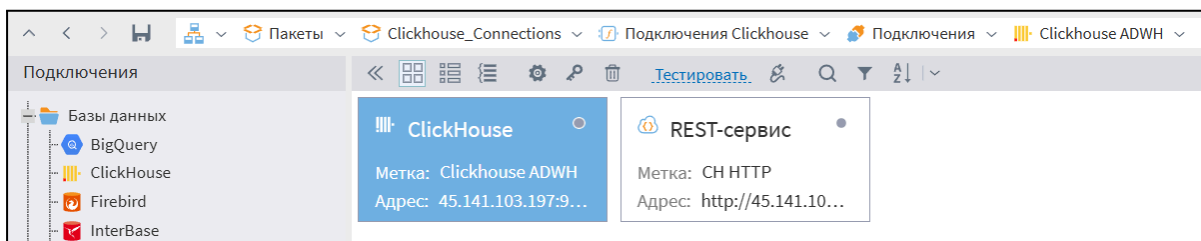
ИП Стучалкин Евгений Игоревич, 504011547861

Настройка подключения к СУБД Clickhouse

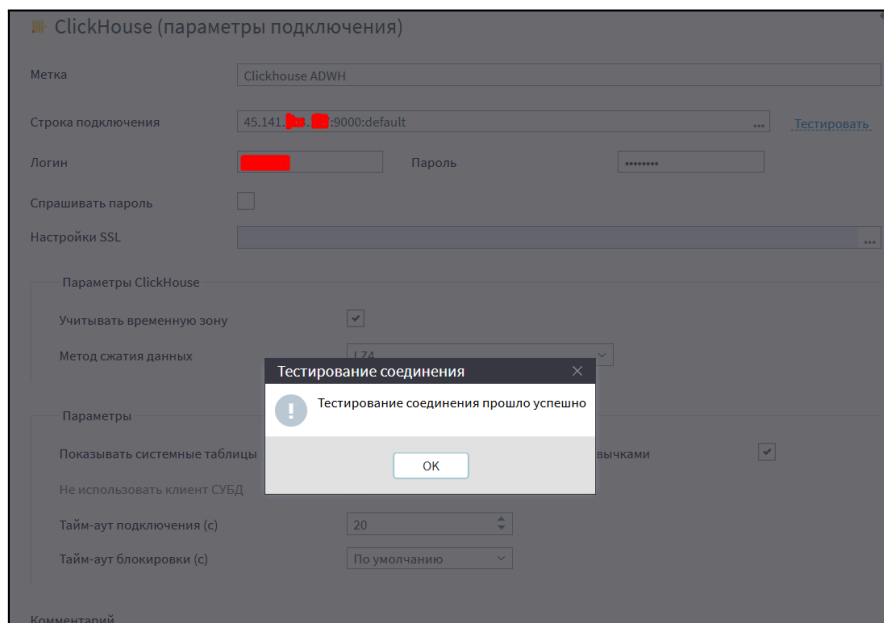
Откройте пакет для настройки основного подключения Data Monetization Pack\Settings\Clickhouse Kit\Clickhouse Connections.lgp



Перейдите в раздел “Подключения” и укажите реквизиты подключения через драйвер и через REST-сервис.



Драйвер сконфигурируйте произвольным образом под особенности вашего сервера Clickhouse. Убедитесь, что подключение работает.



<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

В REST-сервисе укажите только URL подключения, логин и пароль.

Параметры подключения

URL сервиса [Тестировать](#)

Метод

Content-Type запроса

Допустимый Content-Type ответа

Аутентификация для выполнения запроса

Тип аутентификации

Имя пользователя

Пароль

Рекомендуется поднимать отдельный инстанс Clickhouse для работы в DMP, т.к. работа должна выполняться из под пользователя с правами на создание, изменение и удаление БД/Таблиц. А также в наборе присутствуют компоненты очистки хранилища, способные удалить все БД кроме системных за одно действие.

Сохраните изменения в пакете Clickhouse Connections.lgp и закройте его.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>



ИП Стучалкин Евгений Игоревич, 504011547861

Настройка Telegram-бота

Создайте Telegram-бота по [инструкции](#), создайте группу для получения в нее оповещений о качестве данных, узнайте ИД этой группы (делаем все до раздела статьи “Настройка сценария в Loginom”).

Альтернативно, ИД группы можно узнать через бота @myidbot, добавив его в группу и написав в чате группы команду /getgroupid. Бота после этого можно удалить.

Перейдите в папку с настройками Telegram-бота API Pack\Telegram\Content. Там будет 2 файла для настроек.

UserStorage > Libs > Publicated > API Pack > Telegram > Content		
Имя	Дата изменения	Тип
 Alerts_channels.txt	17.10.2024 10:58	Файл "ТХТ"
 Basic_Settings.txt	17.10.2024 10:56	Файл "ТХТ"

В файле Basic_Settings.txt укажите в ссылке токен вашего бота. В файле Alerts_channels.txt укажите ИД созданной вами группы для получения оповещений.

Чеклист проверки установки:

Создан рабочий каталог на файловом сервере	
В него скопированы библиотеки, и размещены в папках: Data Monetization Pack, API Pack, fileapi_kit, silver_kit, json_kit	
Указан валидный лицензионный ключ в Data Monetization Pack/Settings/LicenseKey.txt	
Настроено подключение к Clickhouse	
Настроен Telegram-бот	

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

В пакете Data Monetization Pack\License Check.lgp можно проверить корректность настроек. Активируйте компонент проверки и убедитесь, что он выполнен без ошибок, а на выходе находится подтверждение того что все работает.

Выходная переменная License_status должна иметь значение true. Иначе, библиотека будет работать в демо-режиме.

Подмодели должны активироваться и выводить подтверждения об успешном выполнении. Проверьте статусы проверок на выходных портах

Лицензия

Проверка базовых настроек

Войти

Таблица 1			
#	ab Ответ коннектора БД Clickhouse	ab Ответ HTTP Clickhouse	ab Ответ Телеграм
1	Подключение к БД работает	REST-подключение к БД работает	{"ok":true,"result":{"message_id":361,"from":{"id":6893628387,"is_b...

Обновление библиотеки

Чтобы при обновлении библиотеки не происходил сброс системных настроек, обновлять нужно только файлы, содержащие компоненты (те, на которые вы добавляете ссылки). Пакет “Clickhouse connections.lgp” обновлять не надо.

Перед обновлением сделайте резервную копию всей папки Data Monetization Pack, чтобы в случае возникновения проблем откатиться. Если после обновления возникают сообщения об ошибках, свяжитесь с подставщиком или поддержкой.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

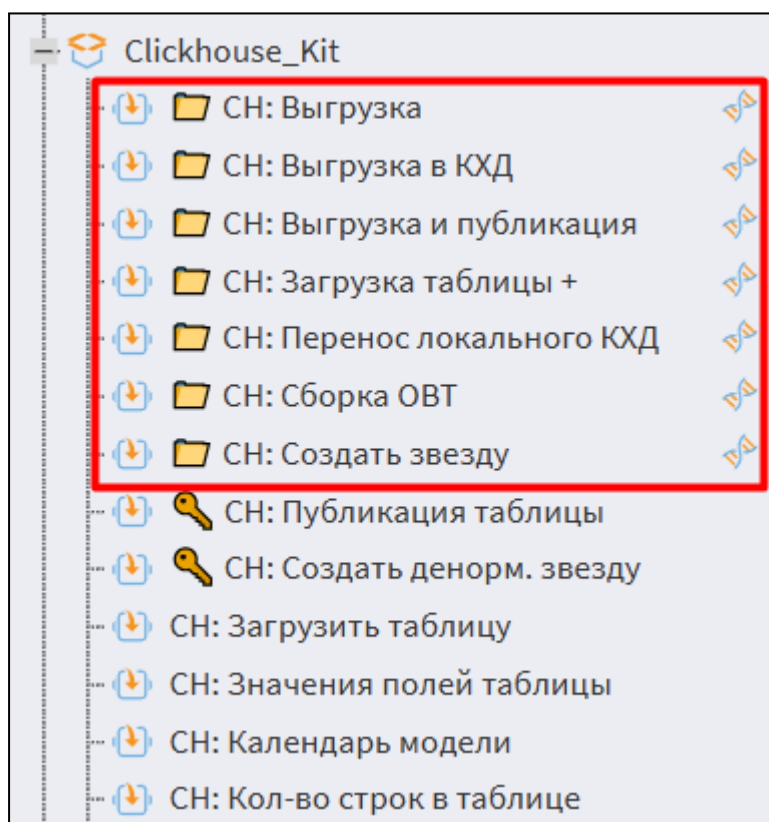
ИП Стучалкин Евгений Игоревич, 504011547861

Основные сценарии работы с библиотекой

Готовые шаблоны сценариев

На текущий момент DMP насчитывает несколько десятков компонентов, и их число будет увеличиваться. Также эти компоненты должны комбинироваться между собой, и между компонентами других пакетов.

Чтобы упростить и ускорить внедрение, в некоторых пакетах представлены готовые сценарии для ряда задач. Они отмечены знаком “📁”.



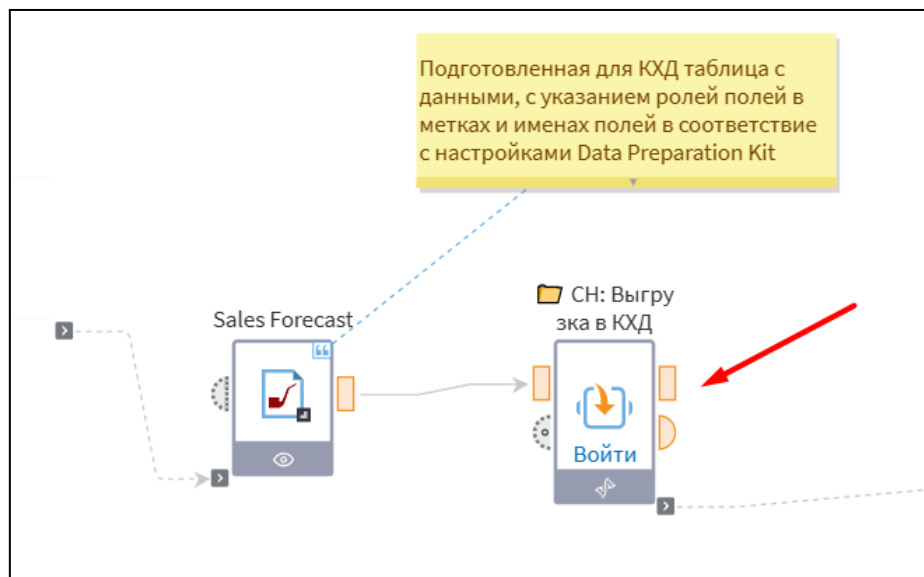
Изображение для примера, актуальный набор компонентов может отличаться

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

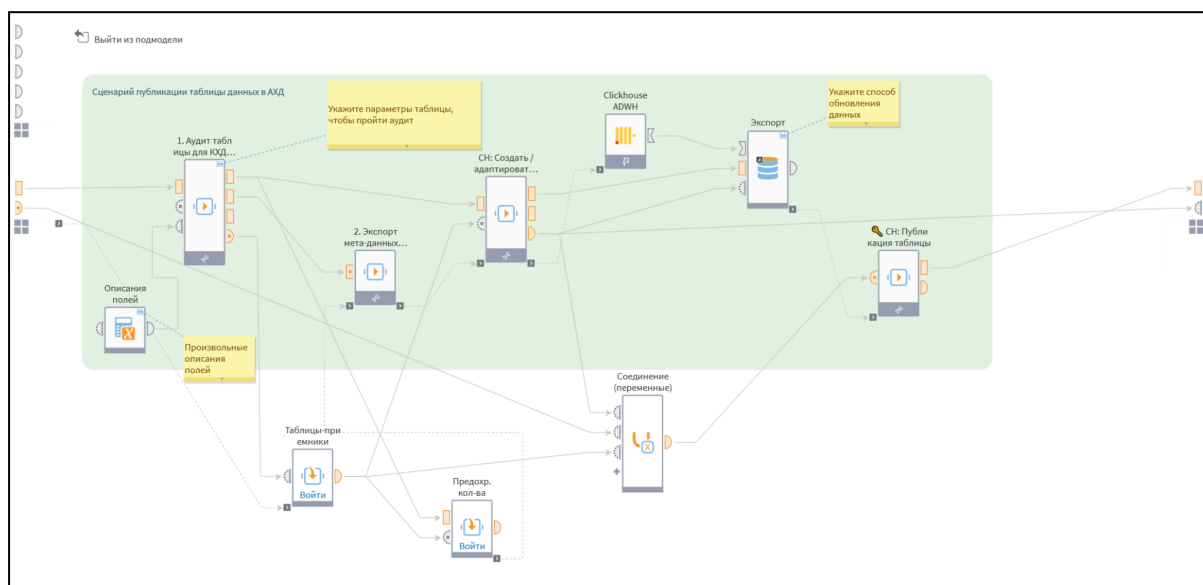
<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Такие компоненты действуют как производные узлы, т.е. компоненты, в которые можно зайти, настроить компоненты или добавить свои.



Производный узел с шаблоном **СН: Выгрузка в КХД**



Содержимое шаблона **СН: Выгрузка в КХД**

Ориентируйтесь в первую очередь на шаблоны сценариев для понимания основных способов работы в DMP.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Как правильно выгружать данные в КХД

За выгрузку данных в КХД на базе Clickhouse отвечает пакет “Clickhouse Kit.Igp”. В DMP мы придерживаемся следующей структуры организации КХД: все формируемые таблицы делятся на 3 типа (уровня хранилища):

Stage - предварительная выгрузка данных из источников “как есть”, для дальнейшей очистки и подготовки к анализу и построению витрин на последующих этапах. Как правило, используется для “сложных” источников, отдающих данные по API. Потому что если при каждом запуске ETL-процесса запрашивать данные напрямую из источников, то либо процесс будет долгим, либо будет отваливаться из-за превышения лимитов запросов.

Никаких особых требований к структуре и формату данных тут нет, в отличие от последующих этапов.

Core - это уровень, который хранит данные, очищенные после stage уровня через ETL-процесс. Таблицы этого уровня должны соответствовать определенному формату (описывается дальше). Каждая таблица должна содержать данные одной предметной области (сущности). Т.е. отдельно справочник товаров, отдельно справочник клиентов, отдельно таблица продаж, отдельно таблица остатков и т.д.

Таблицы core-уровня не используются для решения бизнес-задач по анализу данных в пользовательских инструментах. Они служат кирпичиками, из которых будут собраны витрины, или модели данных - те самые финальные таблицы, с которыми будут работать пользователи, подключая к ним BI и другие инструменты.

Data Mart - уровень, который хранит витрины или модели данных. В нашей интерпретации, модели / витрины это - наборы таблиц в топологии Звезда, полученные из объединения нескольких таблиц core-уровня.

Сборка звезды осуществляется за счет функционала генерации модели данных, автоматически. Также можно собрать витрину произвольной структуры в Loginom и выгрузить ее.

Главное правило: витрины / модели собираются только из таблиц core-уровня. Хотите добавить в витрину данные из новых таблиц? Сначала создавайте сущности в

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

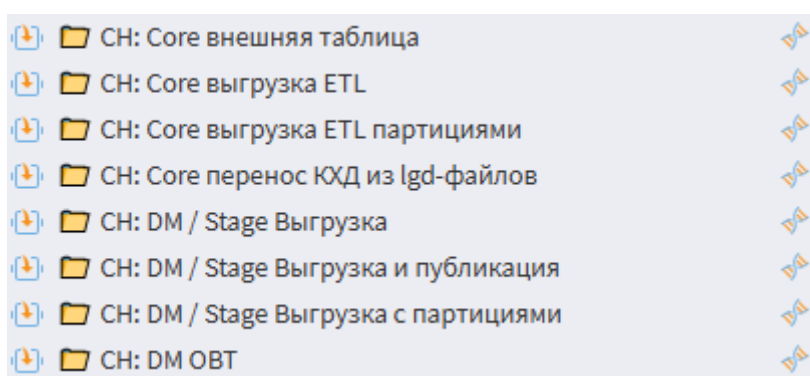
core-уровне, а потом используйте их в витрине.

Такой подход обеспечивает автоматическое соблюдение связности данных независимо от структуры связей, а также минимизирует трудозатраты на поддержку и развитие аналитики (за счет переиспользования данных с core-уровня, а не создания отдельной логики под каждую бизнес-задачу).

Закрепляем:

Stage (опционально) > ETL > Core > Data Mart

В Clickhouse Kit.Igr представлен целый ряд сценариев, связанных с выгрузкой данных:



В их названии отмечено, для каких уровней КХД предназначен шаблон (DM = Data Mart). Часть шаблонов отмечено как Stage / DM, потому что предназначена для выгрузки таблиц произвольной структуры, что может быть использовано как для формирования stage-уровня, так и для формирования витрины на базе сценария, собранного в Loginom с использованием таблиц core-уровня.

Работа генератора Звезд (на картинке CH: DM OBT) не требует загрузки данных в Loginom, а работает по принципу генерации SQL-запроса, который выполняется на стороне СУБД.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

ETL-процесс для новой таблицы в core-уровне КХД

Формат таблицы КХД, к которому надо прийти в результате трансформаций описан тут:

<https://loginom.ru/enterprise-analytics-day2>

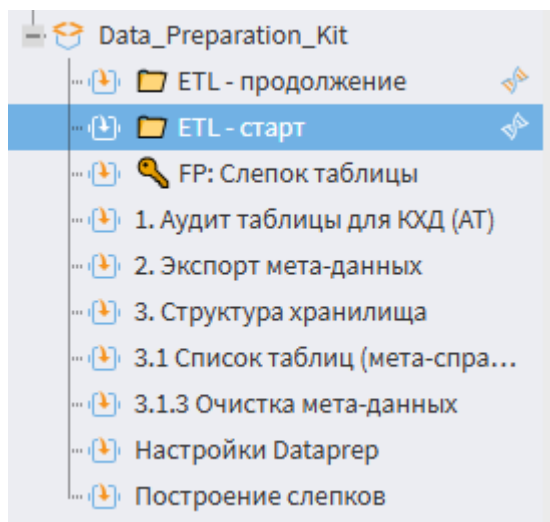
Несмотря на low-code подход и визуальное проектирование, недостаточная структурированность сценариев разработки является проблемой в т.ч. для Loginom. Как правило, сценарий трансформации данных наращивается стихийно, а момент, когда понятная схема превратилась в непроходимый лабиринт - упускается.

Написание пояснений и группировка в подмодели зачастую упускаются в горячке разработки. Поэтому мы предлагаем следующий подход созданию ETL-сценариев, чтобы:

1. Разработка изначально была структурирована;
2. Формировалась автоматическая документация процесса трансформации таблиц с пояснениями.

1. Добавляем в пакет с будущим ETL-сценарием ссылку на “[Data Preparation Kit.Igp](#)”

В нем нас интересуют 2 шаблона сценариев: ETL - старт и ETL - продолжение.



<https://marketplace.loginom.ru/solutions/data-monetization-pack>

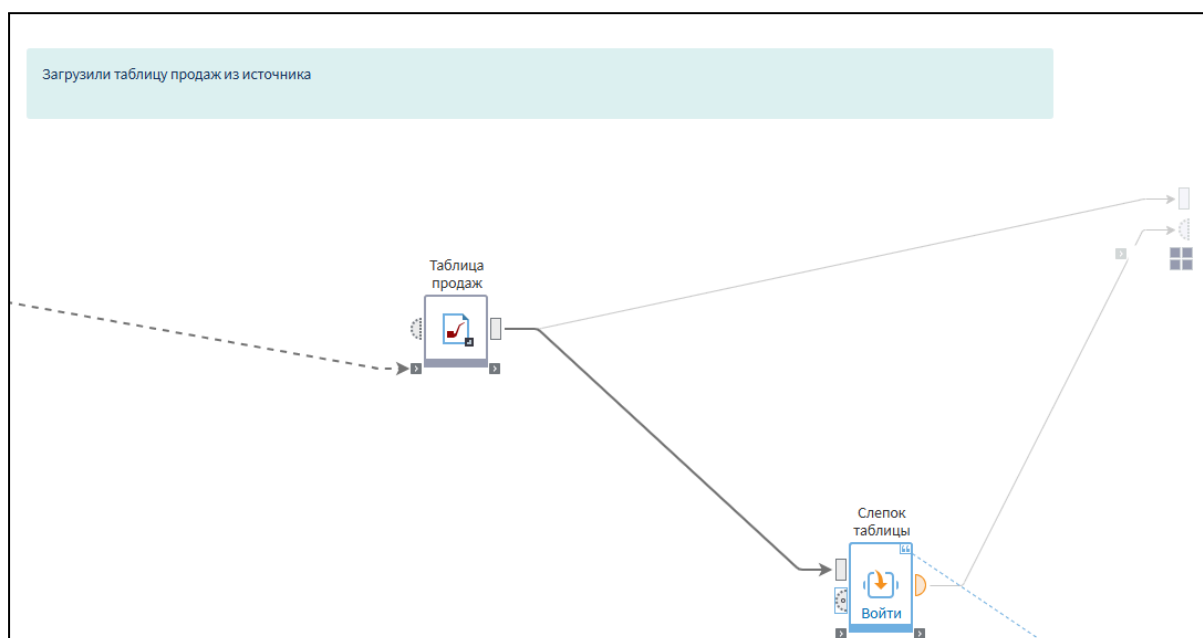
<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

2. Начинаем разработку с добавления ETL - страт

Внутри этого компонента добавляем загрузку из источников. Если таблиц-источников несколько (из БД берется несколько таблиц и соединяется в одну для дальнейшей обработки), делаем это здесь.

- Соедините загрузку данных через последовательность выполнения с началом подмодели (пунктир к LGD-файлу).
- Напишите в текстовой области что вы загружаете и зачем. Скопируйте этот текст.
- Соедините итоговую таблицу со входом в “Слепок таблицы” и с выходом из подмодели.



Критически важно уже на этапе импорта данных задать имена полей, которые дальше будут использоваться в сценарии! Хотя в LogiQL можно переименовать поле на любом этапе сценария, делать это нужно как можно ближе к импорту. Т.к. например у вас загрузится поле `Vyuchka`, вы будете использовать его в сценарии в разных формулах, а в какой-то момент поймете - `Vyuchka` выглядит не очень красиво, пусть будет `Revenue`. Если переименовать поле посреди сценария, то вам в итоге сложно будет отслеживать весь путь данных.

<https://marketplace.logiql.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стулалкин Евгений Игоревич, 504011547861

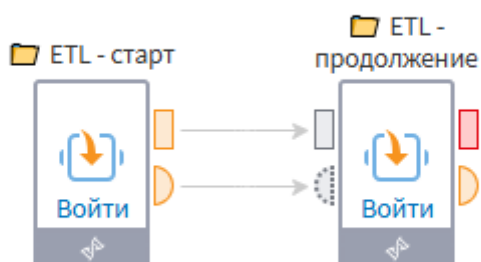
Зайдите в порт переменных слепка таблицы. Там нужно задать:

1. Название ветки - это будет будущее название таблицы, которая будет сформирована в КХД.
2. Комментарий слепка - сюда вставьте текст, который вы написали в примечании.

Метка	Имя	Назначение	Значение	
ab	Название ветки	split_name	Не задано	Sales_2
ab	Комментарий слепка	print_comment	Не задано	Загрузили таблицу продаж и...

Компонент слепка сохраняет образец вашей таблицы, и объединяет их в последовательность с остальными слепками с аналогичной меткой. Формируя таким образом наглядный процесс формирования таблицы для КХД. Для последующих этапов добавляем компонент ETL - продолжение.

3. Достаиваем цепочку трансформаций через ETL - продолжение

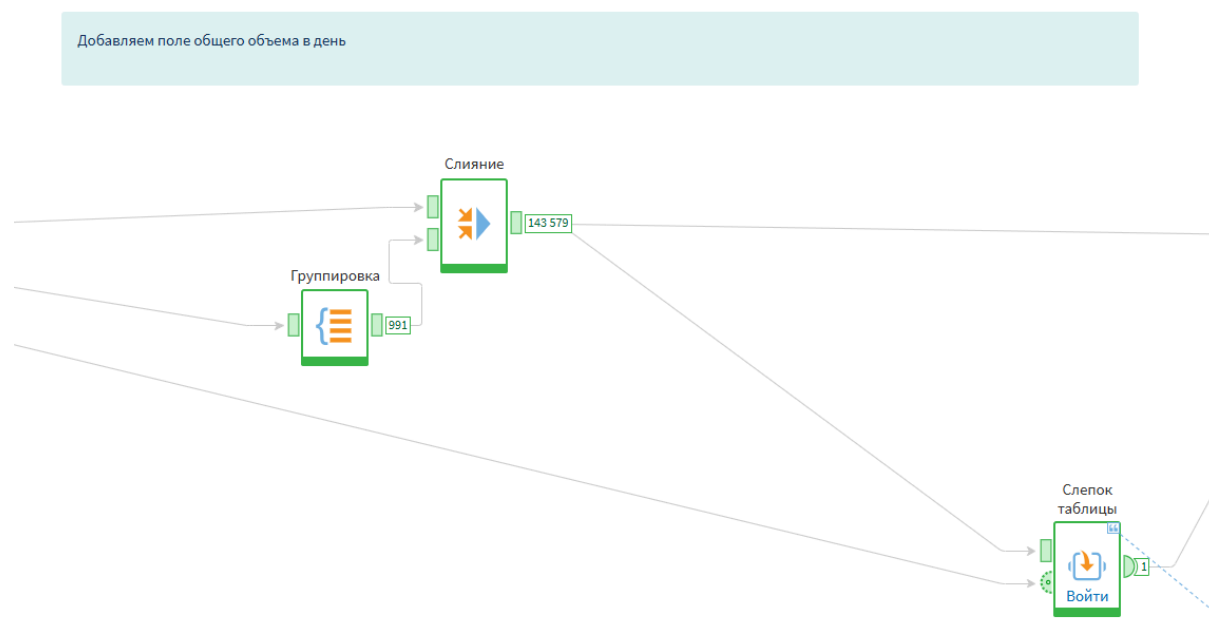


<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

В ней повторяете процесс, выполняя этап очистки или расчетов.



Небольшая разница будет в том, что настройки слепка не нужно будет писать его название, т.к. оно будет передаваться через переменные.

Входные	Выходные	Имя	Назначение	Значение
ab Группа слепка	ab Группа слепка	print_name	Не задано	Sales_2
	ab Название ветки	split_name	Не задано	
	ab Комментарий сле...	print_comment	Не задано	Добавляем поле об...

4. Используйте ветвление слепков, чтобы корректно фиксировать последовательность параллельных операций.

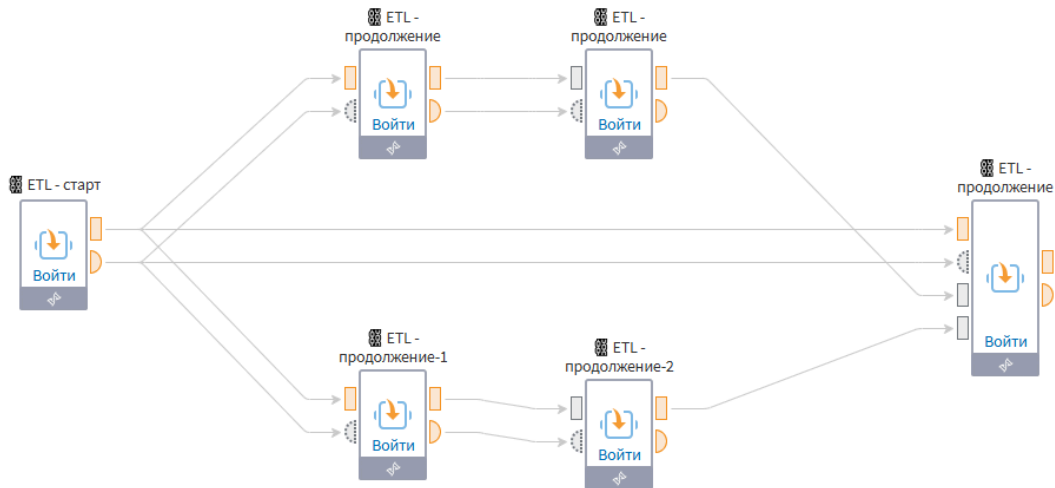
Если вы захотите запараллелить некоторые этапы вычисления, и при этом сформировать их слепки, то это можно сделать, выделив дополнительные ветки.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

На картинке основная ветка представлена ETL - старт и финальным узлом ETL - продолжение. Там формируются слепки



При этом у продолжений-веток в первом узле ветки пишется название ветки, чтобы их этапы трансформации друг с другом не перемешивались.

Ветка 1:

Входные	Выходные	Имя	Назначение	Значение
ab Группа слепка	ab Группа слепка	print_name	Не задано	linage_test
	ab Название ветки	split_name	Не задано	product_rating
	ab Комментарий сле...	print_comment	Не задано	Группировка по про...

Ветка 2:

Входные	Выходные	Имя	Назначение	Значение
ab Группа слепка	ab Группа слепка	print_name	Не задано	linage_test
	ab Название ветки	split_name	Не задано	managers_rating
	ab Комментарий сле...	print_comment	Не задано	Выручка по менедж...

Итоговый массив данных из слепков позволит вам создать упрощенную поэтапную

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

схему формирования таблицы, с отслеживанием возникновения полей на этих этапах.

			6	7	8	9	10	11
			> Загрузил дан...	> Выручка по м...	> Группировка ...	> Рейтинг по м...	> Рейтинг по п...	> Слияние рейт...
linage_test	linage_test	Amount_plan	1. Created					2. Exists
		key_ERP_User_Sales_ID	1. Created					2. Exists
		key_Product_ID	1. Created					2. Exists
		Managers_Rating						1. Created
		Plan_month	1. Created					2. Exists
		Product_Rating						1. Created
		Revenue_plan	1. Created					2. Exists
		Revenue_plan_1						1. Created
		Revenue_plan_2						1. Created
		linage_test__managers_rating	key_ERP_User_Sales_ID		1. Created			2. Exists
	Managers_Rating					1. Created		
	Revenue_plan		1. Created			2. Exists		
linage_test__product_rating	key_Product_ID			1. Created			2. Exists	
	Product_Rating					1. Created		
	Revenue_plan			1. Created			2. Exists	

Нет смысла делать слепки после каждого калькулятора. Это скорее инструмент фиксации крупных этапов преобразования, которые представляют единый смысловой этап. Например, загрузка > очистка > обогащение по одному сценарию > обогащение по второму сценарию.

5. Финализируем таблицу перед сохранением в КХД

Оберните вашу цепочку в подмодель, подайте итоговую таблицу на выход. Убедитесь что вы соединили первый узел цепочки с началом подмодели через последовательность выполнения.



Убедитесь, что вы задали нейминг полей по инструкции:

Нейминг полей и меток для core-уровня

Нейминг полей — ключевые поля

- Ключевые поля — это такие поля, по которым строится связь между таблицами, позволяя делать вычисления по полям одной таблицы в разрезе полей другой таблицы.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

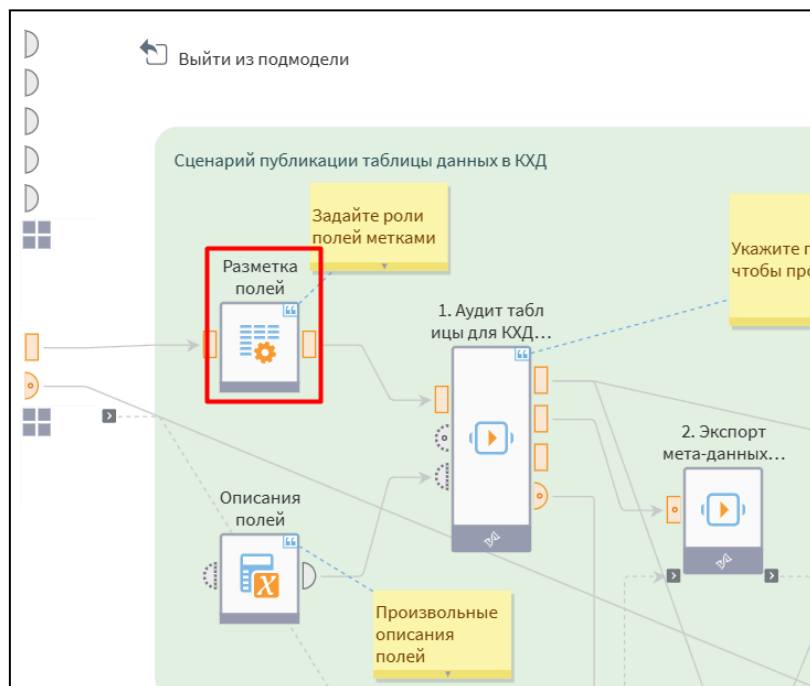
- Ключевые поля имеют свой префикс. В DMP (Data Monetization Pack) этот префикс определяется в файле настроек Data_prep.txt, по умолчанию key_;
- Ключевые поля называются одинаково во всех таблицах. Если идентификатор товара в справочнике товара называется key_Product_ID, он должен называться так во всех таблицах;
- Метка и имя ключевого поля должны быть одинаковыми.

Нейминг полей — поля данных:

Не ключевое поле может иметь произвольную метку и имя;

Для облегчения будущего сбора витрин, нужно заранее определить роли полей. Их может быть 3: **мера, измерение, каноническая дата**.

Лучше всего это делать с помощью параметров полей в шаблонах выгрузки в КХД, например “[СН: Выгрузка в КХД](#)” или “[Экспорт в лок. КХД](#)”, добавление которых описано в шаге 6.



Меры определяются добавлением «_M» в конец метки поля. Каноническая дата определяется через добавление «_CD» в конец метки поля. Каноническая дата — термин из моделирования данных. Обозначает единую ось данных в модели/витрине,

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

вокруг которой собираются показатели из разных таблиц, и их можно анализировать на одной временной оси (например, план и факт продаж). Одна таблица может содержать несколько канонических дат. Подробнее разберем эту тему позже.

Все поля, не являющиеся ключами, мерами или каноническими датами, становятся измерениями. Во время сборки витрины/модели, измерения могут быть связаны по ключевым полям с другими таблицами. Давая таким образом анализировать показатели одних таблиц в разрезе полей из других таблиц.

Форматы дат:

По возможности убирайте из полей дат значение времени (округляйте значение даты вниз до целого). Это оптимизирует работу будущих моделей/витрин;

Если поле содержит не дату (например, день продаж), а обозначение периода (например, месяц плана) — такое значение должно быть приведено к первой дате периода (первый день месяца плана). Это позволит на одной временной оси связать данные разных разрядностей во времени и анализировать их совместно в более укрупненных группировках календаря (например, месяц, квартал, год).

6. Добавьте компонент выгрузки в КХД

Например, “[СН: Выгрузка в КХД](#)” или “[Экспорт в лок. КХД](#)”. Настройте в нем формирование мета-данных в компоненте “[1. Аудит таблицы для КХД \(АТ\)](#)”. Выгрузите ваши данные в КХД, выполнив сценарий до конца.

7. Сформируйте итоговую подмодель выгрузки таблицы

Сверните 2 подмодели в единую подмодель. соединив начало и конец подмодели через последовательность выполнения. Не добавляйте выходные порты данных (потом узнаете почему).



8. Создайте из итоговой подмодели открытый производный компонент

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Активируйте возможность создания производных узлов (пригодится для создания дочерних компонентов на основе единого сценария, но с разными настройкам. Например, полная выгрузка данных, или только за последний месяц)

Новый производный компонент

Метка: ETL: Продажи

Описание:

Область видимости

- Закрытый (доступен только внутри текущего модуля)
- Защищенный (доступен только внутри текущего модуля и его подмоделей)
- Внутренний (доступен только внутри текущего пакета)
- Открытый (доступен во всех пакетах)

Целевые типы узлов

- Выполнение узла и Цикл
- Производные узлы

Добавить Отмена

На этом все. Повторяем процесс для остальных таблиц КХД.

Рекомендации по организации ETL-сценариев множества таблиц

Чтобы работа с ETL-сценариями оставалась удобной и управляемой при росте кол-ва таблиц в КХД, а также значительно сократить опасность безвозвратной потери сделанной работы, рекомендуется придерживаться следующих правил:

1. Для каждой отдельной таблицы КХД создаем отдельный пакет.

В демонстрационных сценариях показано, что все таблицы в КХД формируются в рамках одного пакета.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

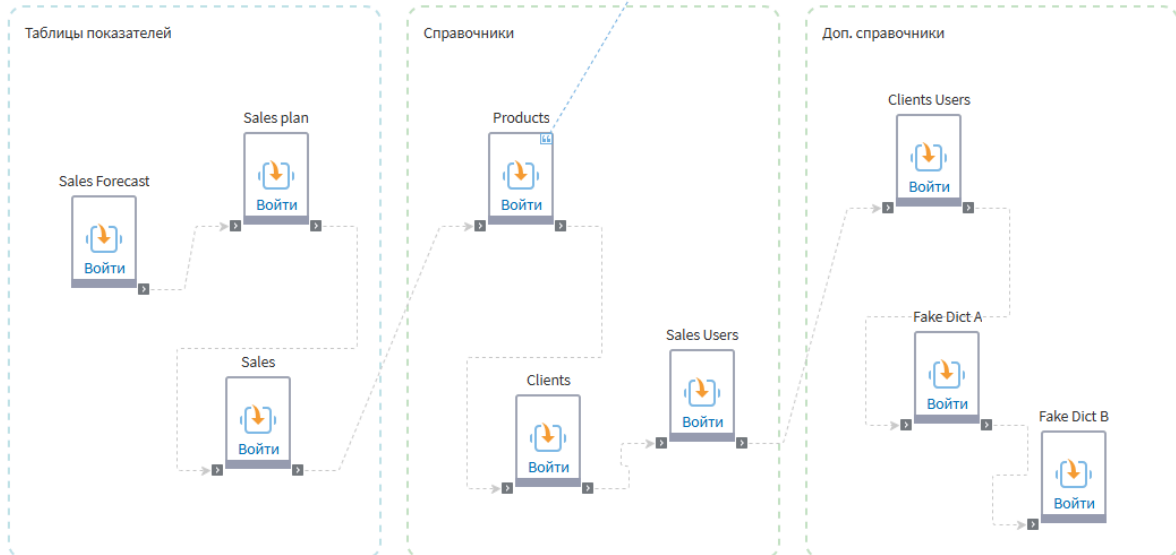
<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Процесс сохранения данных из источника в АХД для дальнейшего построения моделей данных

Пример подмодели (все подмодели идентичны), сохраняющей данные в АХД для дальнейшего построения моделей данных.

Из подобных подмоделей нужно создавать производные компоненты, и строить цепочки обновления данных, которые вы будете запускать с



Технически это не запрещено, но при росте проекта возникнут следующие проблемы:

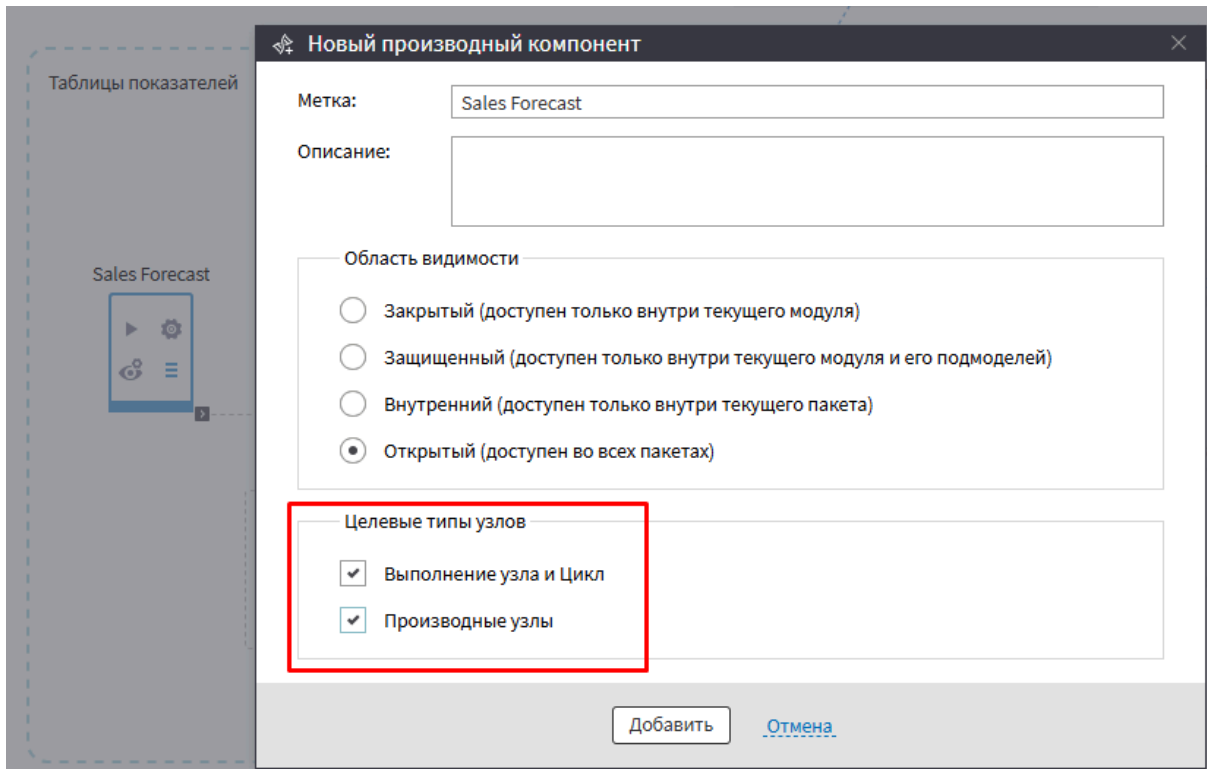
1. Пакет будет открываться очень долго (вплоть до нескольких минут);
2. Невозможна параллельная работа 2-х разработчиков над сценариями разных таблиц. Когда пакет открыт у одного пользователя, у второго он открывается только на чтение.

Поэтому важно следовать правилу: 1 таблица КХД = 1 пакет. Для каждой ETL-подмодели создаем открытый производный компонент, отмечая оба типа целевых узлов (производные узлы могут потребоваться для отладки).

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

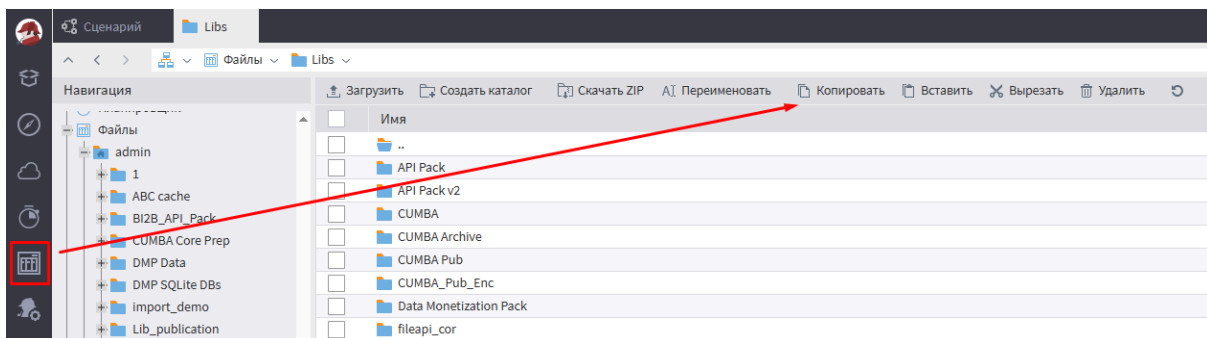
ИП Стучалкин Евгений Игоревич, 504011547861



2. Пакеты, в которых ведется разработка, нужно хранить в папке для разработки

Условно, это может быть папка “ETL dev”. Когда вы завершили разработку, сохранили изменения в dev-пакете, и хотите чтобы изменения поступили в прод, вам нужно скопировать пакет в другую папку, которая будет называться условно “ETL prod”.

Для этого в файловом менеджере выделяете нужные пакеты, нажимаете “Копировать”, переходите в папку “ETL prod”, нажимаете “Вставить” - новые файлы просто вставятся, одноименные будет предложено заменить.



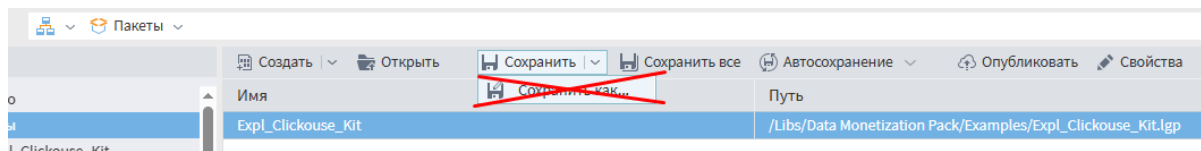
Не используйте опцию “Сохранить как” из диспетчера пакетов. При создании пакета

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

таким способом у него меняется внутренний GUID, и это не позволит использовать его производные компоненты при дальнейшем переписывании пакета через копирование (т.к. его гайд заменится гайдом копируемого пакета) или повторное действие “Сохранить как” (т.к. сгенерируется новый гайд).



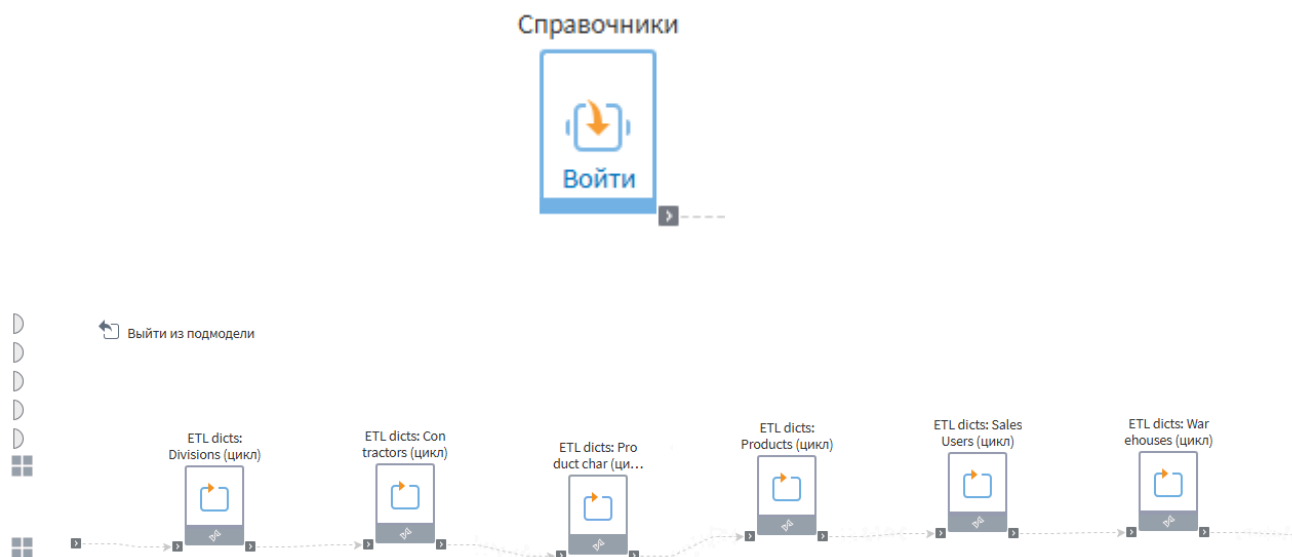
3. Каждый ETL-сценарий должен выполняться автономно

Т.е. не должно быть ситуаций, когда вы сделали ETL-процесс одной таблицы, потом через порт данных перекинули результат в ETL-процесс другой таблицы.

Если переиспользование данных другой таблицы КХД необходимо, вы можете обратиться к ним внутри процесса через запрос к БД, или через промежуточный LGD-файл (если объем данных позволяет).

4. Создавайте отдельные пакеты, в которых будете строить ETL-цепочки.

Добавляйте в них ссылки на пакеты с ETL-компонентами, и стройте цепочки из производных компонентов.



<https://marketplace.loginom.ru/solutions/data-monetization-pack>

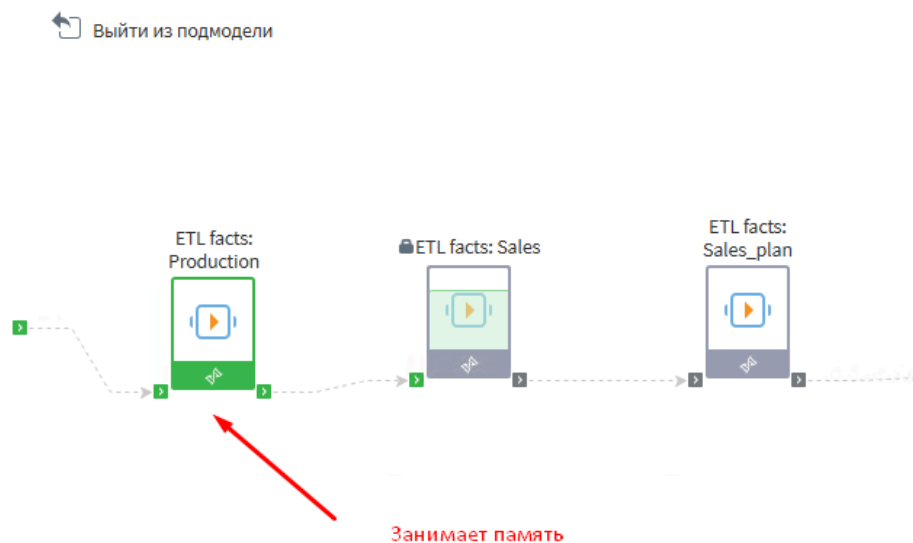
<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Оптимизация потребления памяти цепочками обновлений КХД

Если вы собрали цепочку обновления из нескольких производных компонентов, обернули ее в подмодель, и поставили выполнение подмодели на расписание, то при выполнении задачи расписания будет происходить следующая картина:

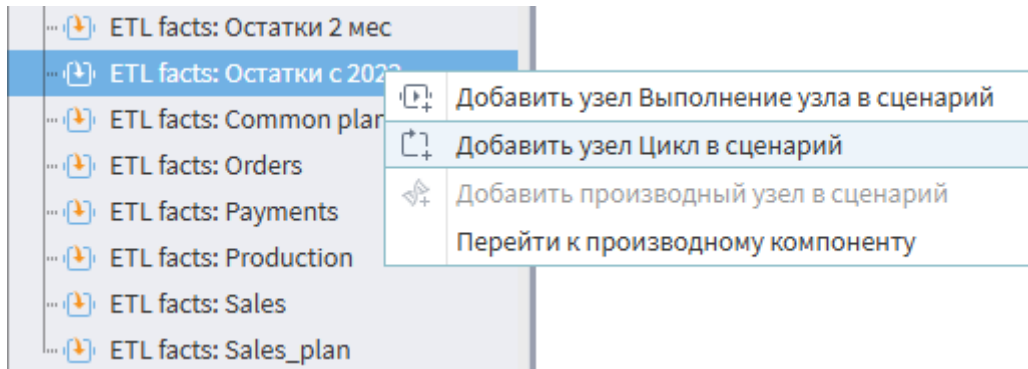
Каждый компонент цепочки будет хранить в себе загруженные данные, пока не выполнится вся цепочка.



Логично, но это не то поведение, которое подходит для ETL-цепочек. Потому что:

1. Хранение данных выполненных компонентов в принципе не нужно для работы сценария.
2. При обновлении цепочек в таблицах фактов большими фрагментами расход оперативной памяти может быть большим, вплоть до получения ошибки “Out of memory”.

На текущий момент эта ситуация решается прибеганием к хитрости: вместо цепочки из обычных подмоделей, делаем цепочку из циклов, добавляя производный компонент в сценарий через соответствующую опцию.



В добавленном цикле производим настройку, чтобы он выполнялся фиксировано 1 раз.

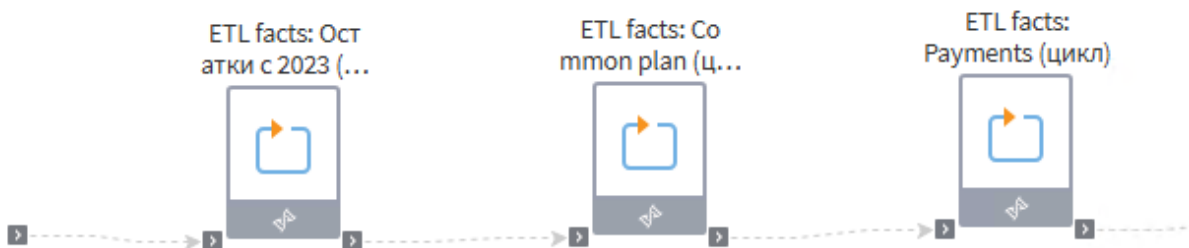
Настройка вида цикла ⌵

Исходный узел

Вид цикла Заданные итерации Цикл с постусловием Групповая обработка

Количество итераций

Повторяем для всех компонентов цепочки. Выглядеть она будет как-то так:



Важно: у производного компонента не должно быть выходных портов данных.

В чем фокус: при выполнении цикла без выходных портов данных, внутренний сценарий выполняется, но по завершении данные внутреннего сценария выгружаются из памяти. Т.к. смысл ETL-сценария в сохранении в БД Таким образом, ваша цепочка

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стулалкин Евгений Игоревич, 504011547861

не будет накапливать отработанные данные.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Clickhouse Kit.Igp

Библиотека для взаимодействия с СУБД Clickhouse, в т.ч. организации КХД и формирования слоя витрин / моделей с помощью генератора моделей данных.

Шаблоны сценариев

CH: DM / Stage Выгрузка

(Ранее  CH: Выгрузка)

Использует узлы DMP:

[CH: Создать/адаптировать MT таблицу](#)

Используется для простой выгрузки в Clickhouse произвольных таблиц, созданных в Loginom. Также содержит преднастроенный узел экспорта, автоматически сопоставляющий поля таблицы Loginom и принимающей таблицы в Clickhouse.

CH: DM / Stage Выгрузка и публикация

(Ранее  CH: Выгрузка и публикация)

Использует узлы DMP:

[CH: Создать/адаптировать MT таблицу](#)

 [CH: Публикация таблицы](#)

Позволяет выгрузить и опубликовать произвольную таблицу, сформированную в Loginom. Обычно используется для публикации спец. витрин, созданных с помощью low-code в Loginom, с требованием к консистентности данных (чтобы в опубликованной витрине всегда был полный объем данных).

Процесс работы: выгрузка данных во временную таблицу, перенос временной таблицы на место рабочей через команды RENAME или EXCHANGE (т.е. атомарно).

Есть опция создания клона временной таблицы через команду SELECT AS после завершения выгрузки и обмена. Таким образом, полный набор данных будет храниться

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

и во временной таблице, и в рабочей. Это дает возможность обновлять временную таблицу инкрементально, с переброской полного объема данных на место рабочей таблицы.

СН: DM / Stage Выгрузка с партициями

Использует узлы DMP:

[СН: Создать/адаптировать МТ таблицу](#)


 [СН: Публикация таблицы](#)

[СН: Перенос партиций](#)

[СН: Партиции таблицы](#)

Позволяет выгрузить произвольную таблицу, и обновлять ее посредством перезаписи партиций (хорошо подходит для больших объемов данных, но менее гибко по сравнению с Выгрузкой и публикацией)

СН: Core выгрузка

Ранее ( СН: Выгрузка в КХД)

Использует узлы DMP:

[1. Аудит таблицы для КХД \(АТ\)](#)

[2. Экспорт мета-данных](#)

[СН: Создать/адаптировать МТ таблицу](#)

 [СН: Публикация таблицы](#)

Используется для сохранения очищенной таблицы в core уровень КХД, для дальнейшего использования в генерации витрин. Также наполняет справочник мета-данных.

В качестве принимающей базы указывается база из настроек в файлах библиотеки (по умолчанию ADWH_storage).

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Процесс выгрузки выглядит так: выгрузка данных во временную таблицу, перенос временной таблицы на место рабочей через команды RENAME или EXCHANGE (т.е. атомарно).

Есть опция создания клона временной таблицы через команду SELECT AS после завершения выгрузки и обмена. Таким образом, полный набор данных будет храниться и во временной таблице, и в рабочей. Это дает возможность обновлять временную таблицу инкрементально, с переброской полного объема данных на место рабочей таблицы.

СН: Core выгрузка ETL партициями

Использует узлы DMP:

[1. Аудит таблицы для КХД \(АТ\)](#)

[2. Экспорт мета-данных](#)

[СН: Создать/адаптировать МТ таблицу](#)

 [СН: Публикация таблицы](#)

[СН: Перенос партиций](#)

[СН: Партиции таблицы](#)

Используется для сохранения очищенной таблицы в core уровень КХД, для дальнейшего использования в генерации витрин. Также наполняет справочник мета-данных.

В качестве принимающей базы указывается база из настроек в файлах библиотеки (по умолчанию ADWH_storage).

Позволяет создать таблицу и осуществить первичную выгрузку, с дальнейшим обновлением данных через перезапись партиций. Партиции выгружаются во временную таблицу, и переносятся в рабочую. Хорошо подходит для больших объемов данных.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

СН: Core внешняя таблица

Использует узлы DMP:

[СН: загрузить таблицу](#)

[1. Аудит таблицы для КХД \(АТ\)](#)

[2. Экспорт мета-данных](#)

Позволяет добавить в справочник мета-данных core-уровня таблицу, не пропуская ее через ETL-процесс Loginom.

Для этого таблица должна (одно из двух):

1. Существовать в Clickhouse на КХД сервере;
2. Быть доступной из Clickhouse на КХД сервере средствами подключения Clickhouse к сторонним базам.

Это позволяет не пересохранять большие объемы данных, которым не требуется очистка, и использовать их в рамках единой модели с другими таблицами core-уровня. Также, этот шаблон дает возможность разнести core-уровень на несколько баз (если это почему-то требуется).

СН: Загрузка таблицы +

Использует узлы DMP:

[СН: загрузить таблицу](#)

Позволяет загрузить произвольную таблицу, доступную через [мастер-подключение к Clickhouse](#), используя настраиваемый сценарий. Сценарий позволяет выбрать посредством фильтрации списков целевую таблицу и ее набор полей. А также сформировать условие отбора WHERE IN на основе значений полей таблицы.

При загрузке таблицы из базы КХД, есть возможность установить именование полей по меткам и с префиксами. Получив таким образом уникализированные понятные имена полей. Удобно использовать для работы с данными КХД в произвольных сценариях Loginom.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

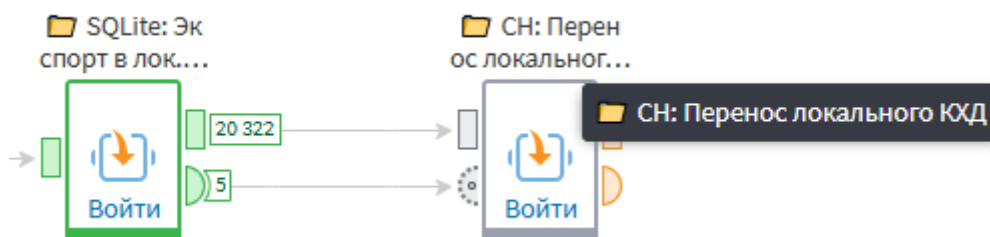
<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

📁 СН: Перенос локального КХД

Комбинируется с компонентом “📁 SQLite: Экспорт в лок. КХД” из SQLite Kit, чтобы перейти от тестового проекта к полноценному КХД без перенастройки сценариев.

Пример соединения:



Использует узлы DMP:

[СН: Создать/адаптировать МТ таблицу](#)

[🔑 СН: Публикация таблицы](#)

📁 СН: ОВТ сборка

Использует узлы DMP

[🔑 СН: Создать денорм. звезду](#)

[СН: Календарь модели](#)

Формирует из таблиц аналитического хранилища на базе мета-данных витрину (модель данных) в формате One Big Table. Витрина представляет из себя полностью автономную таблицу MergeTree, содержащую все данные внутри себя. принцип сборки модели описан тут: <https://loginom.ru/enterprise-analytics-day4>.

Первая часть сценария содержит гибкие настройки отбора полей модели. Сначала определяется перечень таблиц модели, после чего можно детально настроить перечень полей с их разбивкой по ролям. Например, сформировать модель со всеми полями показателей, но только с некоторыми полями измерений.

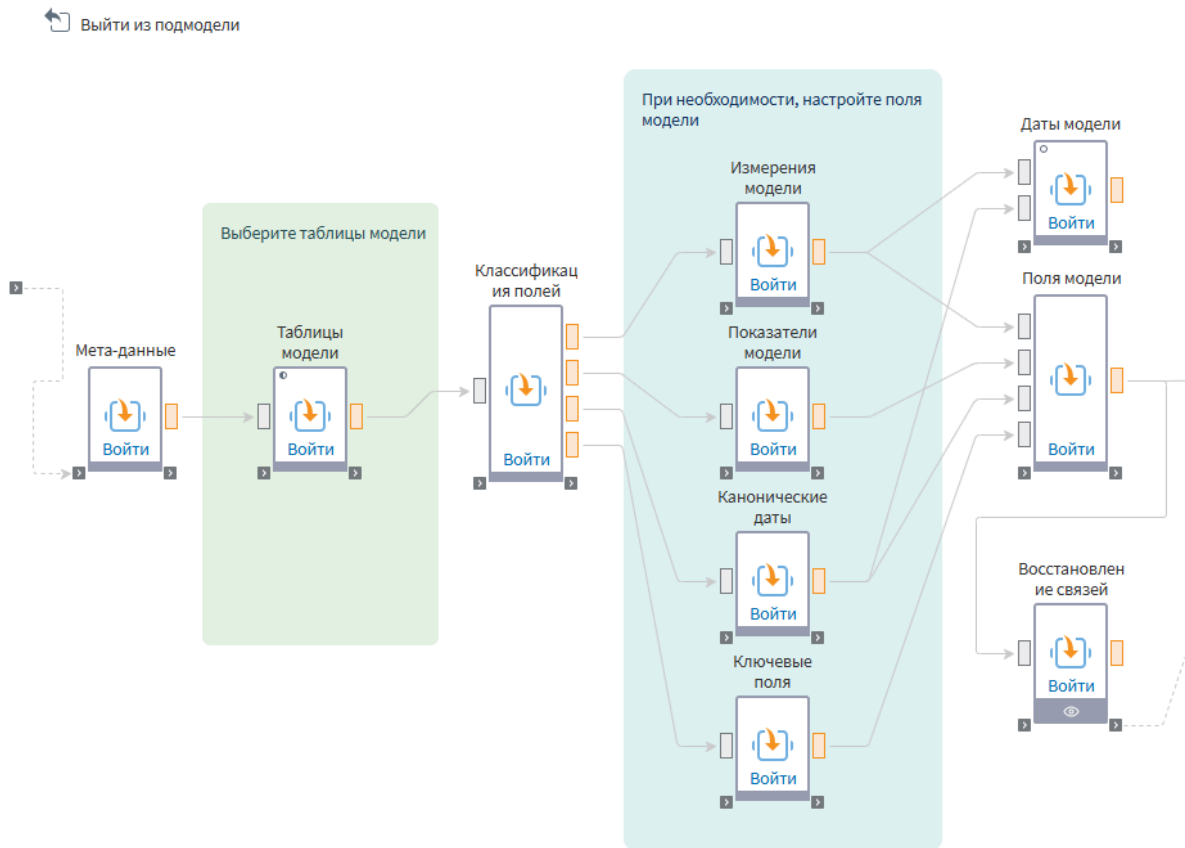
Базовый сценарий подразумевает указание в фильтра списка полей / таблиц, но

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

пользователь может сделать полностью собственную логику внутри соответствующих подмоделей.



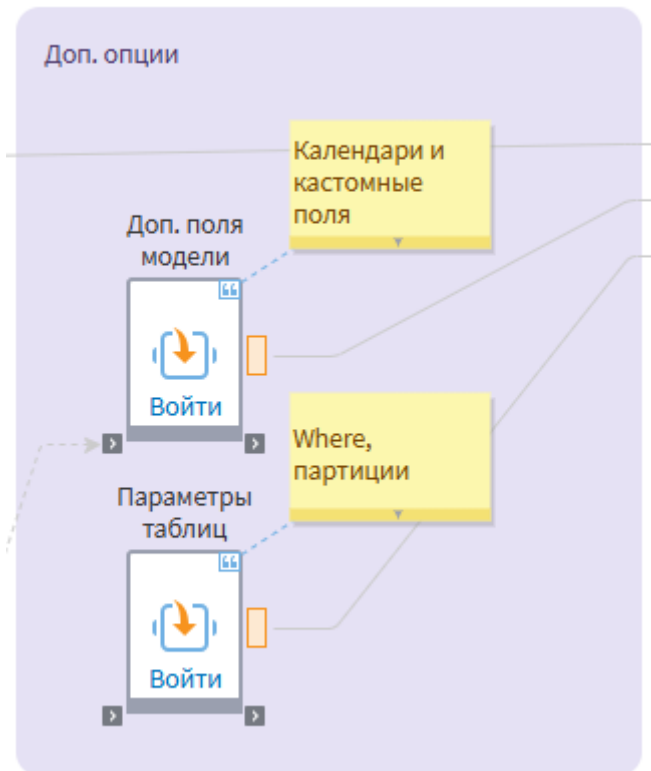
Узел **восстановления связей** показывает, какие дополнительные слияния ключевых полей будут выполнены для обеспечения целостности модели при кросс-табличных связях. Слияния ключевых полей определяются в зависимости от конкретного набора таблиц и ключей.

Далее, идут шаблоны для тонкой настройки будущей модели.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

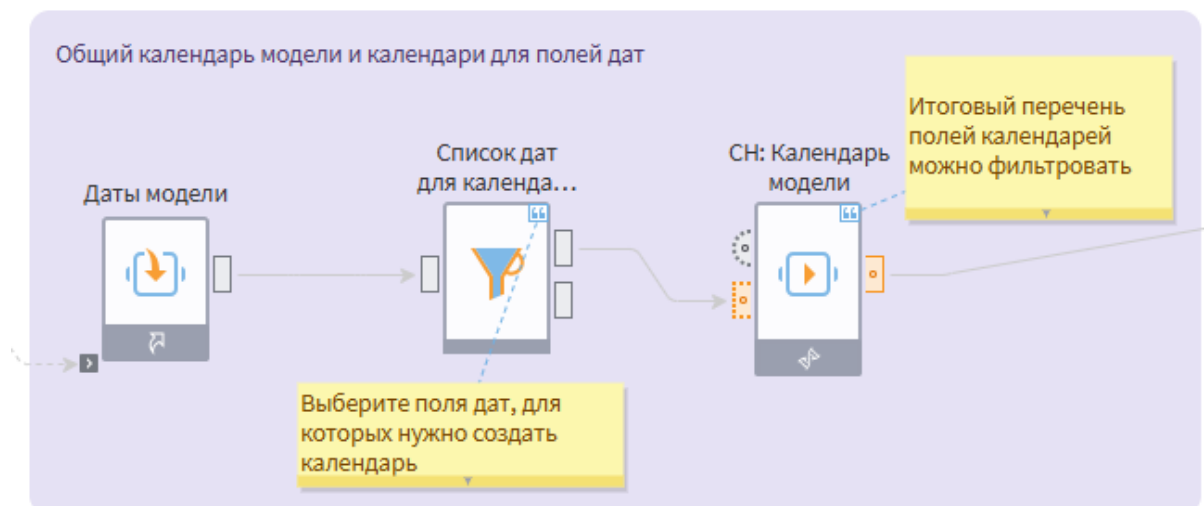
<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861



Доп. поля позволяет создать перечень дополнительных расчетных полей, которые изначально отсутствовали в таблицах модели. Самый распространенный способ их использования - создание календарных полей.

В календарях можно сгенерировать перечень выражений для каждого поля даты модели (в т.ч. для единой временной оси) на основе шаблонов календарных выражений (подробнее в описании СН: Календарь модели).



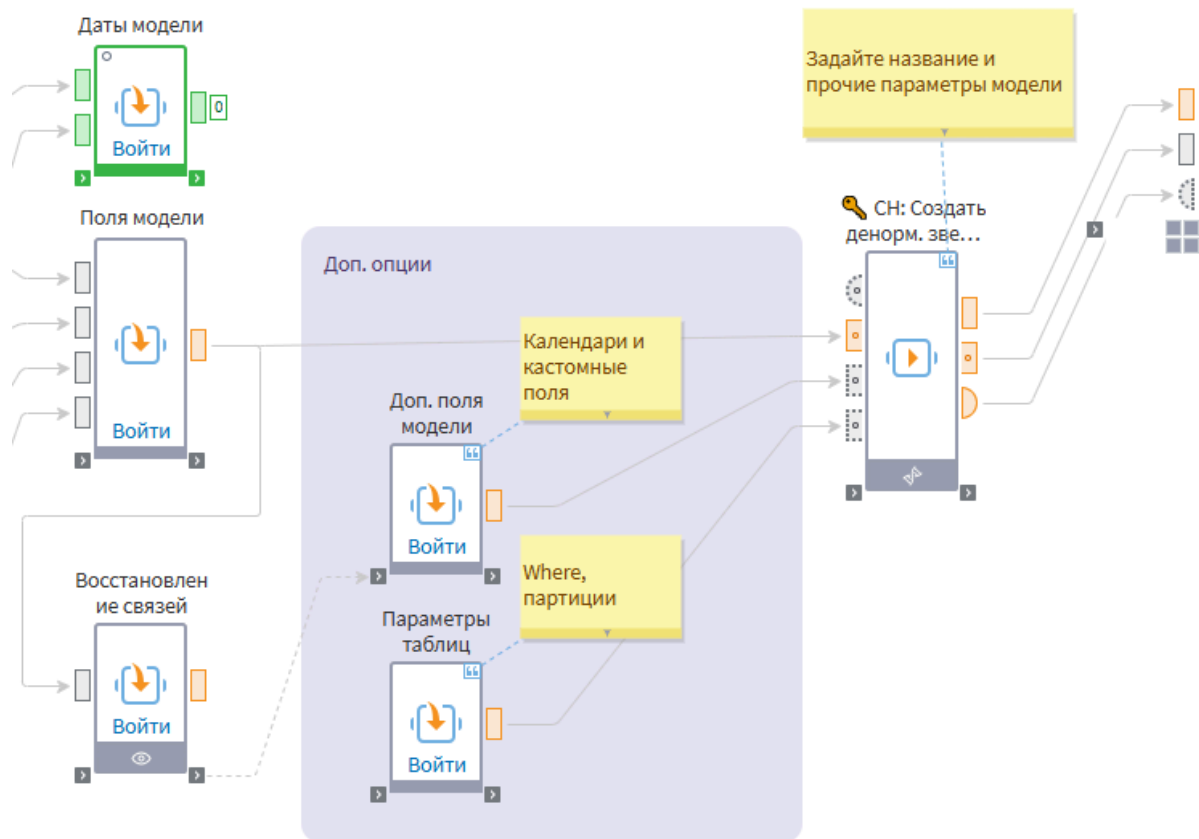
Подмодель “Параметры таблиц” содержит шаблоны для создания выражений WHERE и указания партиций, индивидуально для каждой таблицы АХД из которых строится модель.

Итоговый набор параметров поступает в компонент “🔑 [СН: Создать денорм. звезду](#)”, отвечающий за генерацию модели.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861



<https://marketplace.loginom.ru/solutions/data-monetization-pack>
<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

СН: ОВТ с партициями


Использует узлы DMP

[СН: ОВТ сборка](#)

[СН: Перенос партиций](#)

[СН: Удалить таблицу](#)

Шаблон предназначен для простой настройки частичного обновления витрин на уровне данных отдельных таблиц соге-уровня хранилища и периодов времени внутри них. Используется для поддержания в актуальном состоянии витрин с большим объемом данных (сотни млн. строк).

Содержит преднастроенный сценарий создания витрины на базе компонента “ [СН: ОВТ сборка](#)”, а также производный узел от него, который:

- 1) Наследует перечень полей и настроек витрины;
- 2) Создает витрину во временной БД;
- 3) Позволяет указать дополнительные условия загрузки данных поверх основной витрины, в частности условия WHERE;

Таким образом:

- 1) Во временной БД создается фрагмент витрины со структурой, идентичной основной модели;
- 2) Далее, происходит перенос партиций в основную витрину через серию команд REPLACE PARTITION, что обеспечивает атомарную замену фрагментов модели;
- 3) Временная таблица с партициями удаляется.

Имейте в виду, что поля даты, на базе которого строятся партиции блока модели, должны быть неизменными. Иначе один объект при изменении даты может появиться сразу в 2-х партициях, потому что:

1. Он останется в старой партиции, которая не менялась обновлением;
2. Он появится в новой партиции из-за изменившейся даты.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

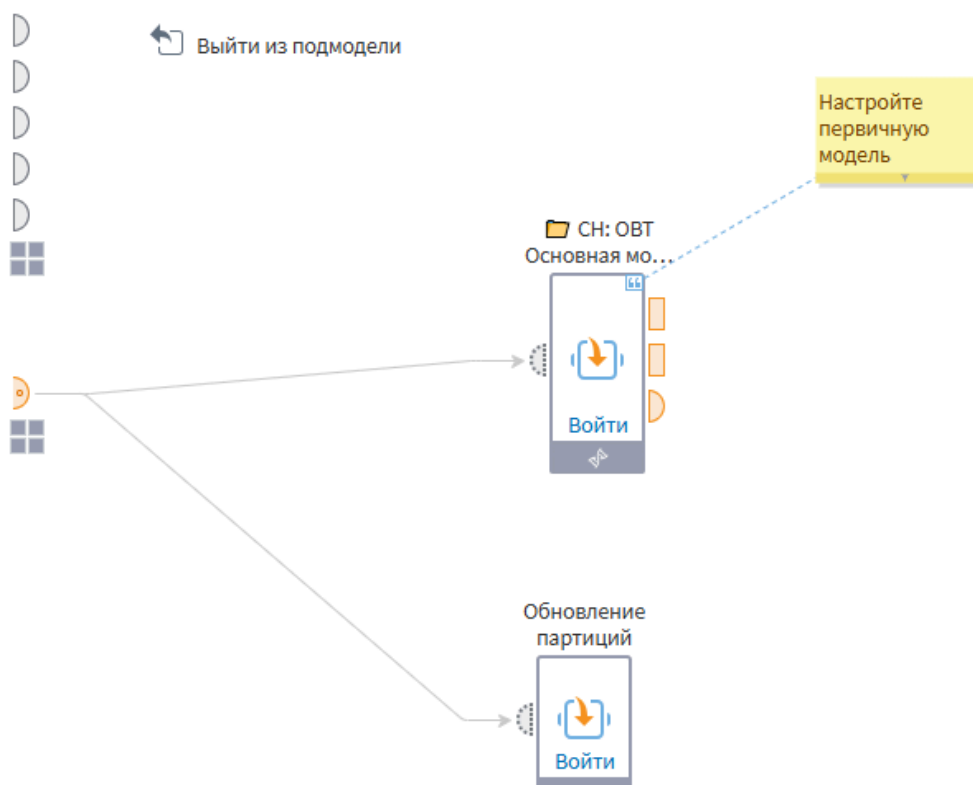
ИП Стучалкин Евгений Игоревич, 504011547861

Например: в таблице Сделка есть 2 даты: дата создания сделки, и дата закрытия. Если партицию строить по дате закрытия с группировкой по месяцам, то в случае если сделка будет закрыта в месяце 1, а потом открыта и закрыта снова в месяце 2, то при обновлении партиции только за месяц 2 в ваших данных также останется запись о сделке в партиции 1. В этом случае, партиции надо строить по дате создания сделок, т.к. эта дата остается неизменной всегда.

Обновление через партиции имеет смысл применять:

1. Если базовая таблица содержит слишком большой объем данных (сотни миллионов строк), и полное обновление начинает занимать десятки минут;
2. Если частично обновляемые данные имеют ограниченную глубину изменений задним числом.

Изначально шаблон содержит 2 подмодели.



После настройки основной модели, нужно донастроить сценарий частичной загрузки в

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

обновлении партиций. Вы можете клонировать подмодель “Обновление партиций”, чтобы создавать разные фрагменты данных для частичного обновления. Эти подмодели можно поставить на выполнение по расписанию через планировщик.

Базовые компоненты

СН: Создать/адаптировать МТ таблицу

Создает на стороне Clickhouse таблицу для экспорта данных на основе структуры таблицы, выгружаемой из Loginom. Если таблица в Clickhouse уже существует, то происходит сверка структур таблицы в Loginom и Clickhouse, и при выявлении расхождений таблица в Clickhouse меняется под выгружаемую таблицу из Loginom.

Настройки:

Имя БД - Название БД, в которой будет создана / находится принимающая таблица.

Имя таблицы - Название таблицы, которая будет создана / адаптирована.

Сопоставление типов полей в Loginom и Clickhouse настраивается в файле Data Monetization Pack\Settings\ClickhouseKit.txt.

Если таблицы не существует, она будет создана. Если таблица существует, но отличается от выгружаемой из Loginom, она будет приведена к структуре таблицы Loginom через команду ALTER TABLE, включая:

- Добавление полей;
- Изменение типа данных поля;
- Удаление поля (переименование поля на стороне Loginom интерпретируется как создание поля с новым именем и удаление старого поля)

Движок - можно указывать движки семейства MergeTree.

Поле партиции - список полей через запятую, опционально.

Поля сортировки - поля, используемые в ORDER BY при создании МТ-таблицы.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Список полей через запятую, опционально. При пустом значении, будет использован ORDER BY Tuple()

Поля первичного ключа - список полей через запятую, опционально.

Порядок полей по частотности - если **Поля сортировки** пусты, то для более оптимальной работы таблицы можно расположить поля таблицы по возрастанию частотности значений в них. При этом будет проведен анализ частотности полей в выгружаемой Loginom-таблицы, что может занять некоторое время при первичном создании таблицы. Если опция деактивирована, то порядок полей будет взят из таблицы Loginom.

Упорядочивание по частотности происходит внутри типов данных, порядок типов определяется в файле Data Monetization Pack\Settings\ClickhouseKit.txt

Nullable mode - Помимо типов данных, при создании полей указывается, могут ли они принимать значения null. Есть 3 режима работы этой опции.

- 0 - все поля создаются как nullable, и в них можно писать значения null
- 1 - все поля создаются как не nullable, и при попытке записать в них null значение Loginom будет выдавать ошибку.
- 2 - автоопределение. Возможность хранения null значений в таблице CH определяется на основе наличия null значений в полях Loginom при создании таблицы. Первичный аудит займет некоторое время. Учитывайте, что если при создании таблицы в поле не было null значений, и оно создалось как не-nullable, а потом этом поле появится null - экспорт выдаст ошибку.

Удалить и пересоздать - иногда выгружаемая структура данных в Loginom меняется таким образом, что принимающая таблица в Clickhouse не может быть адаптирована через ALTER TABLE (в этом случае вы получите соответствующую ошибку).

В таком случае, при установке этой опции в режим true, при каждой активации компонента целевая таблица будет удаляться и создаваться заново. Используйте эту операцию и установите снова режим false, чтобы преодолеть невозможность выполнения ALTER TABLE.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Содержимое выходных портов

Порт Данные - пропускает насквозь таблицу Loginom без изменений.

Порт Describe Table - результат запроса DESCRIBE TABLE, который возвращает список полей и типы данных таблицы непосредственно из СУБД, отражая ее текущее состояние.

Порт переменных Результат - содержит результат активации

Данные		Describe Table	Результат	
№	Имя	Метка	Значение	
1	ab Status	Статус операции	✔ Таблица создана	
2	ab Generated_query	Созданный запрос	CREATE TABLE "New_DB"."New_Table_tmp" ("key_ERP_User_Sales...	
3	ab DB_name	Имя БД	New_DB	
4	ab Table_name	Имя таблицы	New_Table_tmp	
5	ab Full_table_name	Полное имя таблицы	"New_DB"."New_Table_tmp"	
6	ab Select	SELECT	SELECT "key_ERP_User_Sales_ID", "key_Product_ID", "Revenue_pla...	
7	ab Select_as	SELECT AS	SELECT "key_ERP_User_Sales_ID", "key_Product_ID", "Revenue_pla...	

Статус операции - результат активации узла. Варианты: таблица создана / изменена / без изменений.

Созданный запрос - запрос, выполненный при активации компонента. В зависимости от состояния принимающей таблицы (существует / не существует), может иметь варианты: CREATE TABLE (таблица не существует), ALTER TABLE (таблица существует, но структура не совпадает с Loginom, таблица в БД приводится к новой структуре), отсутствие запроса (таблица есть, структура совпадает с Loginom).

Имя БД, Имя Таблицы, Полное имя таблицы - сквозные параметры для передачи в последующие узлы, в т.ч. узел Экспорт.

SELECT, SELECT AS - шаблоны запросов SELECT, с подстановкой алиасов (меток) или загрузкой исходных имен полей.

🔑 СН: Создать денорм. звезду

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Формирует из таблиц аналитического хранилища на базе мета-данных витрину (модель данных) в формате One Big Table. Витрина представляет из себя полностью автономную таблицу MergeTree, содержащую все данные внутри себя. Принцип сборки модели описан тут: <https://loginom.ru/enterprise-analytics-day4>.

Настройки компонента

Домен - БД, которая будет содержать витрину. К имени будет добавлен префикс из настроек Data Monetization Pack\Settings\ClickhouseKit.txt

Название ядра - Таблица, в которой будет создана витрина. К имени будет добавлен префикс из настроек Data Monetization Pack\Settings\ClickhouseKit.txt

Для каждой модели во временной базе создается служебная таблица именем вида “название модели_create_query”, где можно найти сгенерированный SQL запрос создания витрины.



dm_Partitions_tmp	
Таблицы	
core_P_test_part	1,8M
P_test_create_query	9,3K
P_test_part_create_query	9,4K

ORDER BY - выражение сортировки витрины, обязательное для MergeTree таблиц. Если оставить пустым, сформируется автоматически на основе возрастания частотности значений полей-измерений модели (с учетом WHERE-условий).

Есть возможность указать конструкцию table:имя_бд.имя_таблицы - тогда выражение ORDER BY будет взято из соответствующей таблицы. Подход используется, когда нужно взять сортировку конкретной таблицы, например чтобы создать идентичную модель с ограниченным набором данных и потом добавить их в целевую модель.

Включать в модель только уровни с мерами - модель строится по принципу этажей, где данные каждой таблицы лежат в своем этаже. Некоторые таблицы (Товар) могут не иметь своих полей показателей.

Если опция не активна, то для таких таблиц будут созданы свои этажи:

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Сущность	Сумма продажи	Сумма плана	ИД товара	Товар
Продажи	5000	0	1	Молоко
План	0	4000	1	Молоко
Товар	0	0	1	Молоко
Товар	0	0	2	Хлеб

Это позволяет использовать в аналитике записи справочников, которые не представлены в фактах.

Если опция активна, то в витрине будут данные только тех таблиц, в которых есть поля показателей:

Сущность	Сумма продажи	Сумма плана	ИД товара	Товар
Продажи	5000	0	1	Молоко
План	0	4000	1	Молоко

Имена полей в нижнем регистре - итоговые имена полей витрины будут в нижнем регистре (требуется для некоторых систем-потребителей)

Алиасы в нижнем регистре - для каждой витрины формируется 2 шаблона SELECT: с исходными именами полей, и с подстановкой алиасов из мета-данных.

Публиковать - во время создании витрины она создается во временной таблице, а потом единомоментно переносится в целевое расположение с подменой существующей таблицы, если она есть. Отключение этой опции остановит процесс генерации на создании временной таблицы.

Заполнить пропущенные даты - если в таблицах показателей с единой временной осью есть пропущенные даты, то такие даты будут добавлены в модель дополнительно.

Входные порты данных

CT2 Метаданные - список полей модели в формате, который отдает компонент [“3.1](#)

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

[Список таблиц \(мета-справочник\)](#)

Дополнительные поля - принимает таблицу со списком выражений (SQL), который будет создан поверх моделей. Входные поля:

- Выражение: SQL-формула поля;
- Тип данных: Тип данных поля;
- Алиас поля: как оно будет называться в модели.

Параметры таблиц - позволяет передавать дополнительные параметры при загрузке данных из КХД во время формирования модели. Содержит поля:

- Тип параметра: Партиция или WHERE
- Название таблицы: таблица КХД, к которой применяется параметр
- Значение параметра: значение параметра

При заполненном WHERE во все запросы SELECT к конкретной таблице будет добавлено данное условие.

При указании партиции (преобразование в число поля даты таблицы, например toYYYYMM(Date_field)), в модели создается дополнительное поле "partition", содержащее указанное число. Сами партиции формируются из комбинации полей "entity" и "partition". Это позволяет обновлять данные в модели частично, путем замены партиций.

Выходные порты данных

Результат публикации - запрос DESCRIBE TABLE к таблице модели, подтверждающий ее существование в БД.

Шаблоны показателей - шаблоны выражений в синтаксисах разных BI-систем.

Шаблоны запросов - содержит шаблоны SELECT запросов для загрузки данных витрины, некоторые сквозные параметры по месту нахождения модели, а также код статуса генерации (должен быть 200).

СН: Публикация таблицы

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Компонент мгновенно переносит таблицу из текущего размещения в новое (БД + имя таблицы), через команду RENAME или EXCHANGE, в зависимости от того существует ли в целевом расположении таблица, или еще нет.

Используется для того, чтобы при обращении к таблице в новом расположении всегда иметь полные данные на момент последнего обновления, и не попасть в ситуацию когда чтение аналитикой происходит из опустошенной таблицы, которая находится в процессе записи.

Настройки компонента:

Публиковать таблицу - выполнять ли процесс перемещения таблицы при активации компонента.

Клонировать опубликованную таблицу - после перемещения, создает в старом размещении клон таблицы через CREATE AS SELECT.

БД таблицы - источника - БД, из которой берем таблицу

Таблица - источник - Таблица, которую берем

БД, куда публикуем - БД, в которую делаем перенос. Можно оставить пустым, тогда значение возьмется из БД таблицы - источника.

Таблица, куда публикуем - Имя таблицы после переноса. Можно оставить пустым, тогда значение возьмется из Таблица - источник

СН: Загрузить таблицу

Позволяет загрузить произвольную таблицу, доступную через [мастер-подключение к Clickhouse](#), используя настраиваемый сценарий.

Настройки компонента:

Название БД - БД, из которой загружаем таблицу.

Название таблицы - Название таблицы, которую загружаем.

adwh_alias_mode - если таблица из базы КХД, можно указать способ именования полей: 0 - имена из БД, 1 - имена из БД + префикс имени, 2 - метка из Loginom, 3 -

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

метка из Loginom с префиксом.

WHERE - строка для условия отбора WHERE. Пишется сразу выражение отбора, например Sales_Sum>0 and Product_ID=3.

SELECT DISTINCT - загружать только уникальные строки.

LIMIT - если больше 0, загружает указанное число первых строк

SAMPLE - если для таблицы настроено выражение семплирования, можно указать кол-во записей для отбора семпла.

Дополнительные порты:

Список полей - позволяет опционально подать список полей таблицы. Если ничего не подано, загружаются все поля.

WHERE IN - позволяет подать массив вида (Имя поля|Значение), которое будет трансформировано в условие WHERE IN со списком значений для соответствующих полей. При указании условия отбора в параметре **WHERE**, полученное выражение отбора будет объединено через оператор AND

Например, мы указали условие WHERE Sales_Sum>0 and Product_ID=3, а в WHERE IN подали список полей Client_ID = 1,2,3 и City=Moscow. Итоговое выражение WHERE при запросе будет выглядеть так:

WHERE (Sales_Sum>0 and Product_ID=3) AND ClientID IN ('1','2','3') AND City IN ('Moscow')

CH: Календарь модели

Подмодель, формирующая SQL-выражения для создания календарных полей для компонента “🔑 [CH: Создать денорм. звезду](#)”. Работает по следующему принципу:

1. Из файла Data Monetization Pack\Settings\ClickhouseKit_cal.txt берутся шаблоны выражений календаря. Шаблон содержит: группу измерений, шаблон имени, шаблон выражения, тип данных. Например:

День; День_Недели_\$\$1_N; toDayOfWeek(\$1); Nullable(Int64)

Группа измерений (День) используется для того чтобы отфильтровать ненужные группы (например, если вам в модели нужны только разрезы, относящиеся к месяцам).

Шаблон имени (День_Недели_\$\$1_N) содержит будущее название поля в модели, где \$1 - метка, которая будет заменена на название поля, на базе которого генерируется измерение. Например, День_Недели_прд_Дата_продажи_N. Для удобства в конце имени стоит буква, определяющая тип данных в поле (чтобы удобнее было выбирать в BI нужный вариант из списков). S - строковое значение, D - дата, N - число. Вы можете изменить эти шаблоны на свое усмотрение. Например, если есть требование что в модели поля должны быть на латинице. Вместо “День_Недели_\$\$1_N” можно написать “Day_of_Week_\$\$1_N”.

Шаблон выражения (toDayOfWeek(\$1)) содержит формулу, в которую на место \$1 подставляется название поля, на базе которого вычисляется календарное измерение.

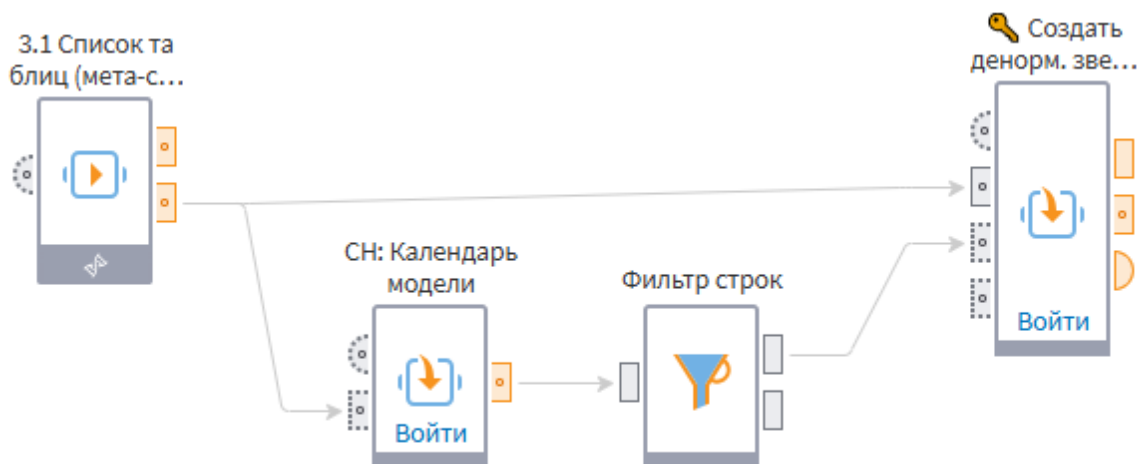
Тип данных (Nullable(Int64)) - тип данных будущего поля. Используем nullable, чтобы в модели не появлялись значения типа “Январь 1970”.

При активации компонента без входных данных, происходит генерация выражений для поля canonical_date (единая временная ось витрины / модели). Также, на вход можно подать мета-данные из компонента “3.1 Список таблиц (мета-справочник)”. Тогда для полей с типом “Дата” будут сформированы дополнительные шаблоны выражений. Итоговый список можно отфильтровать и подать на вход “Дополнительные поля” компоненту “🔑 [CH: Создать денорм. звезду](#)”.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861



СН: Партиции таблицы

Возвращает список партиций таблицы, если они в ней есть.

СН: Перенос партиций

Переносит все партиции из таблицы-источника в целевую таблицу посредством операции REPLACE PARTITION. Таблицы должны иметь одинаковые структуры (партиции, ORDER BY, Primary Key, набор полей, типы данных). На выходе показывается перечень партиций целевой по убыванию даты изменения для контроля выполнения обновления.

СН: Удалить таблицу

Удаляет таблицу из БД

Data Preparation Kit.Igp

Содержит компоненты для формирования мета-данных и data lineage на основе слепков таблиц, а также для аудита пригодности данных к сохранению в КХД. Кроме того, позволяет вызывать существующие мета-данные КХД для использования в автоматизациях.

Базовые компоненты

1. Аудит таблицы для КХД (АТ)

Проверяет пригодность таблицы на сохранения в core-уровень КХД, для использования в дальнейшей генерации витрин / моделей данных.

Формирует мета-данные сохраняемой таблицы. На вход принимает произвольные данные. Базово, для соответствия стандартам КХД, должны быть заполнены настройки:

Название таблицы - имя таблицы в БД КХД. Само название БД, в которой будет расположено КХД, определяется в файле настроек соответствующей СУБД, например Data Monetization Pack\Settings\ClickhouseKit.txt

Описание таблицы - произвольный текст, сохраняется в мета-данные как доп. сведения о таблице.

Первичный ключ - поле первичного ключа таблицы, если оно существует. Иначе, можно оставить пустым. Первичный ключ должен быть уникальным для конкретной таблицы.

Префикс имени поля - аббревиатура, которая будет подставлена к именам полей этой таблицы в моделях или при загрузке через специализированный компонент импорта. Должна быть уникальная для конкретной таблицы.

Префикс метки поля - аббревиатура, которая будет подставлена к меткам полей этой таблицы в моделях или при загрузке через специализированный компонент импорта. Должна быть уникальная для конкретной таблицы.

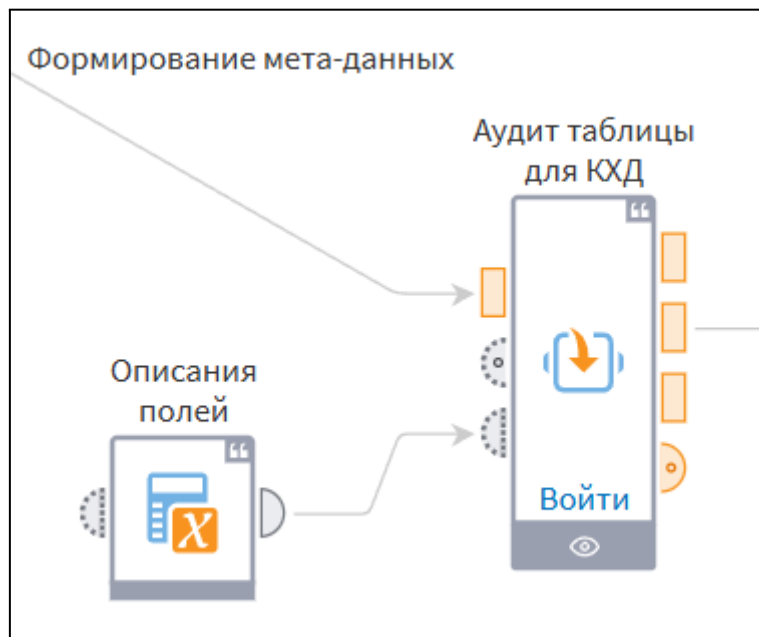
<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

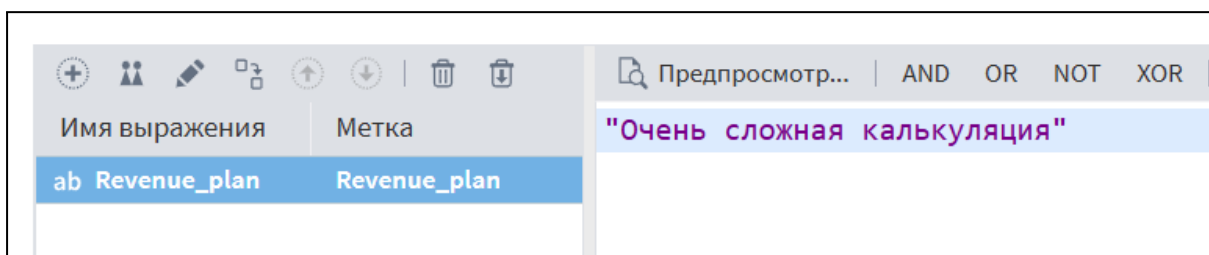
ИП Стучалкин Евгений Игоревич, 504011547861

Источник в БД - при заполнении этого параметра, можно сформировать мета-данные по существующей таблице в БД, не пересохраняя ее в КХД. При этом данную таблицу можно будет использовать при генерации моделей. Опция используется для сценариев, когда нецелесообразно пересохранять в КХД большие таблицы только ради их добавления в единую модель.

Третий порт переменных используется, чтобы задать полям таблицы произвольное описание.



Для этого в калькуляторе переменных нужно создать переменные с именами полей, к которым вы хотите добавить описание. Обратите внимание, такие описания могут быть динамическими.



Выходные порты

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

AT1 - Данные - данные, поданные на вход. Проходят насквозь без изменений.

AT2 - Мета-данные - массив с описанием полей таблицы, используется для формирования справочника мета-данных КХД. Каждому полю сопоставляются роль в будущей модели данных: Первичный ключ, вторичный ключ, каноническая дата, поле показателя, поле измерения.

Определение роли происходит за счет наличия специальных символов в метках полей, которые подаются на вход в аудит.

ab key_ERP_User_Sales_ID	ab key_Product_ID	12 План выручки _M	31 Месяц плана _CD	12 План по количеству _M
2d5c07da-0e15-11eb-81e4-005056915f96	6ad5c66e-febb-11e8-8...	8 700 000	01.02.2024, 00:00	13 200
926b216b-ec2d-11ea-81de-005056915f96	594a5ebd-b36a-4e0a-8...	38 600 000	01.09.2023, 00:00	31 800
926b216b-ec2d-11ea-81de-005056915f96	73d9ac40-26af-11ec-82...	6 300 000	01.11.2024, 00:00	5 400
2e5da389-33d0-11eb-81ed-005056915f96	3b7dec60-ae2d-491a-b...	9 100 000	01.04.2024, 00:00	6 600
9ceb6322-edbb-11ea-81de-005056915f96	3bf2c4e4-87bb-4fa8-91...	4 800 000	01.11.2023, 00:00	11 900

Изменить символы можно в файле Data Monetization Pack\Settings\Data_prep.txt.

Чтобы поле считалось **ключевым**, оно должно начинаться с префикса key_, метка и имя должны быть одинаковыми. Ключевое поле будет считаться первичным ключом, если его название указано в настройках аудита. Иначе, ключ считается вторичным.

Поле канонической даты - будет использовано как единая временная ось, на которой можно объединить несколько таблиц фактов. Одна таблица может иметь несколько канонических дат (например, Дата создания и Дата закрытия в сделках). Определяется через добавление "_CD" в метку поля.

Поле показателя - будет использовано в витринах / моделях в центральной таблице связей. Определяется через добавление "_M" в метку поля.

Поля измерений - к ним будут отнесены все поля, не являющиеся ключами, показателями или каноническими датами. Это сделано специально, т.к. поля измерений обычно самые многочисленные, и расставлять в них метки было бы утомительно. Поля измерений при генерации модели / витрины будут проброшены ко всем связанным показателям, включая связи сквозь другие таблицы.

[Подробнее о генерации моделей](#)

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

АТ3 - Аудит - порт содержит результаты аудита. Критические ошибки отмечаются красным крестом. При наличии критических ошибок, будет невозможно сохранить мета-данные, и продолжить процесс выгрузки данных в КХД.

АТ1 - Данные		АТ2 - Мета-данные		АТ3 - Аудит		Сквозные настройки	
#	ab	Проверка	ab	Описание проверки			
1	⚠	Первичный ключ не задан		Справочные поля будут работать только для собственных показ...			
2	✓	Все параметры заполнены	OK				
3	✓	Есть связи с другими таблицами	Sales_managers, Products, Sales, Sales_forecast				
4	✓	Префикс имени/поля уникален					

В аудите возможен ряд ситуаций, которые не являются явной ошибкой, но могут потенциально ей быть. Такие вещи отмечаются знаком лампочки. Контролируйте их на свое усмотрение в зависимости от ваших целей. Например, отсутствие первичного ключа в таблице фактов далеко не всегда является ошибкой. В то время как отсутствие первичного ключа в справочнике приведет к неправильной работе генератора моделей.

Сквозные настройки - содержат параметры для передачи дальше по сценарию.

АТ1 - Данные		АТ2 - Мета-данные		АТ3 - Аудит		Сквозные настройки	
№	Имя	Метка	Значение				
1	ab Table_name	Название таблицы	Sales_plan				
2	ab vTableDesc	Описание таблицы	План продаж				
3	ab vPrimaryKey	Первичный ключ					
4	ab vField_name_prefix	Префикс имени поля	slp_				
5	ab vField_alias_prefix	Префикс метки поля	ппр_				

2. Экспорт мета-данных

Сохраняет мета-данные из порта АТ2 - Мета-данные в локальное хранилище в папке Data Monetization Pack\Meta Data. Используется как стопор сценария при критических ошибках аудита.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

3.1 Список таблиц (мета-справочник)

Возвращает текущий справочник мета-данных, сформированный на основе компонентов “[1. Аудит таблицы для КХД \(АТ\)](#)” и “[2. Экспорт мета-данных](#)”.

У компонента 2 выхода, в первый выводится общая информация по таблицам, во второй - полная информация по полям.

Data Control Kit.Igp

Содержит различные компоненты для вывода различных статистик по данным в материализованном виде.

Шаблоны сценариев

Аудит качества данных

Использует узлы DMP:

[Отчет DQ](#)

Содержит преднастроенный сценарий проверки качества данных, позволяющий проверять:

1. Любые условия на уровне строк таблицы и значений полей;
2. Сравнение кол-ва строк в таблице с контрольным значением.

При наличии критических ошибок, фиксирует их в логах, а также опционально может останавливать сценарий.

При наличие любых ошибок, настроена отправка отчета через Телеграм-бота.

Базовые компоненты

Отчет DQ

Считает кол-во значений "ИСТИНА" в полях с префиксом "DQ_" в метке, выводя отчет в виде таблицы с кол-вом значений по каждому полю, а также процент от общего числа строк таблицы. Если в метке поля поставить восклицательный знак "!" в любом месте, то наличие хотя бы одного значения "ИСТИНА" в поле будет классифицирована как критическая ошибка.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

SQLite Kit.Igp

Операции в библиотеке SQLite производятся во временных базах данных, которые автоматически создаются и размещаются в папке DMP SQLite DBs:

- в папке пользователя в серверных версиях Loginom;
- в папке Data Monetization Pack в десктопных версиях Loginom.

Шаблоны сценариев

Экспорт в лок. КХД

Использует узлы DMP:

[1. Аудит таблицы для КХД \(АТ\)](#)

[2. Экспорт мета-данных](#)

Шаблон сценария для пилотирования проекта хранилища и генерации моделей. Формирует каталог мета-данных, и сохраняет таблицы в LGD-файлы в папку Data Monetization Pack\LGD storage, для использования в сценарии “[📁 SQLite: Сборка ОВТ](#)”.

Экспорт в SQLite

Использует узлы DMP:

[SQLite: создать / адаптировать таблицу](#)

Создает таблицу в SQLite в соответствие со структурой данных в Loginom, либо изменяет существующую таблицу под соответствие новой структуре данных. Шаблон содержит преднастроенный узел экспорта для бесшовной выгрузки данных.

SQLite: Сборка ОВТ

Использует узлы DMP:

[3.1 Список таблиц \(мета-справочник\)](#)

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

[SQLite: генерация модели](#)

Содержит шаблон для упрощенного набора полей мета-данных, которые будут сформированы в модель данных (витрину). Принцип сборки модели описан тут: <https://loginom.ru/enterprise-analytics-day4>. Модель может быть сформирована только из данных, сохраненных через “[📁 Экспорт в лок. КХД](#)”.

Первая часть сценария содержит гибкие настройки отбора полей модели. Сначала определяется перечень таблиц модели, после чего можно детально настроить перечень полей с их разбивкой по ролям. Например, сформировать модель со всеми полями показателей, но только с некоторыми полями измерений.

Базовый сценарий подразумевает указание в фильтра списка полей / таблиц, но пользователь может сделать полностью собственную логику.

Базовые компоненты

SQLite: создать / адаптировать таблицу

Создает файл БД в пользовательской папке, и формирует в нем таблицу со структурой, соответствующей введенной в него таблице с данными Loginom. Если таблица уже существует, производится сверка полей и типов данных между Loginom и таблицей. В случае расхождения, выполняется запрос ALTER TABLE для приведения таблицы в БД к структуре данных из Loginom.

Т.к. SQLite не поддерживает смену типа данных поля через ALTER TABLE, поле с измененным типом данных будет удалено из БД, и вместо него будет создано новое поле с соответствующим типом.

SQLite: SQL-запрос

Позволяет выполнить SQL-запрос к таблице в сценарии Loginom без ручной разработки сценария промежуточной выгрузки данных в БД. SQL-запрос подается через переменную. В нем нужно оперировать именами полей (не метками). В SQL-запросе нужно использовать конструкцию <ТТ> для обозначения таблицы, с которой происходит работа. Например, SELECT * FROM <ТТ>; загрузит вам все данные, как есть.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

В итоговой таблице метки исходных полей будут заменены на имена. Имена новых полей будут получены транслитерацией от меток.

Как работает процесс:

1. Таблица с данными сохраняется в служебную БД;
2. К ней выполняется SQL-запрос из переменной;
3. Итоговая таблица подается на выход.

Учитывайте, что обработка таким образом больших массивов может быть достаточно медленной, и потреблять много места на диске. Целевой сценарий для данного компонента - выполнение динамически формируемых запросов, логику которых нельзя уложить в фиксированный сценарий Loginom, для небольших (до 1 млн) таблиц.

SQLite: генерация модели

Формирует модель данных в виде плоской таблице с немедленной загрузкой результата в Loginom, на основе перечня мета-данных из компонента “[3.1 Список таблиц \(мета-справочник\)](#)”.

Принцип сборки модели описан тут: <https://loginom.ru/enterprise-analytics-day4>. Модель может быть сформирована только из данных, сохраненных через “[📁 Экспорт в лок. КХД](#)”.

Модель формируется посредством пересохранения во временную БД SQLite отобранных таблиц, после чего к ним выполняется SQL-запрос, результат которого загружается в Loginom. Результирующую витрину можно экспортировать куда-либо, или обработать и проанализировать в Loginom

Целевой сценарий - тестирование создания витрин через подход DMP с минимальной инфраструктурой. Ограничений на объем данных и кол-во таблиц отсутствует, однако с ростом объемов данных эффективность решения будет понижаться. Обращивать большие объемы (десятки млн. строк) может быть нецелесообразным. Для промышленного использования механики, нужно работать с набором компонентов для СУБД Clickhouse.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

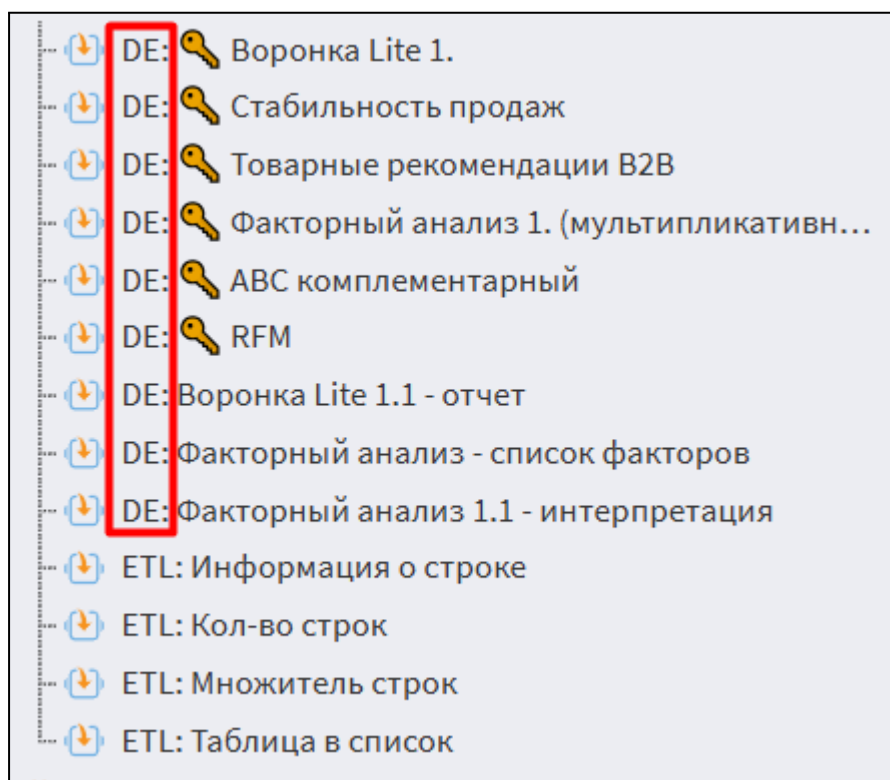
<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

ETL Kit.lgp

Библиотека ETL Kit содержит компоненты, упрощающие трансформации данных для решения ряда технических вопросов, а также обогащения данных для организации ряда сценариев бизнес анализа.

Сценарии для бизнес-анализа отмечены префиксом DE (Data Enrichment).



DE: 🗝️ ABC комплементарный

Для Loginom существует множество готовых компонентов для ABC анализа (в частности, в рамках бесплатной библиотеки Loginom Silver Kit). Наш вариант отличается тем, что он проводит ABC анализ **элементов**, разбитых по **группам**.

Самый распространенный пример:

- Элемент - Товарная номенклатура;
- Группа - Чек;

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

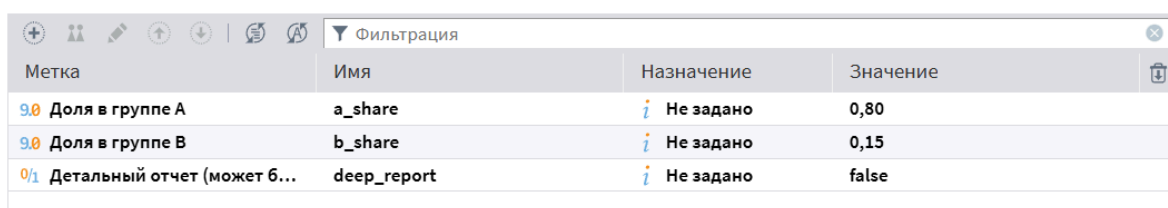
Итоговый результат анализа показывает не только ABC-сегментацию по элементам, но и различную статистику по вхождению элементов в группы, соседство с другими элементами и т.д.

Входы компонента

Настройки

Настройки компонента позволяют выбрать границы ABC-сегментов. Пример на картинке: в группу А попадет 80% лучших элементов, в группу В попадет 15% элементов после худшего из А. Остальные 5% попадут в С.

Параметр “Детальный отчет” влияет на вывод в итоговый отчет детализированных значений по комплементарным элементам (детально разберем далее). Отключение детализированного отчета может существенно повысить скорость работы компонента.



Метка	Имя	Назначение	Значение
90 Доля в группе А	a_share	Не задано	0,80
90 Доля в группе В	b_share	Не задано	0,15
01 Детальный отчет (может б...	deep_report	Не задано	false

Данные

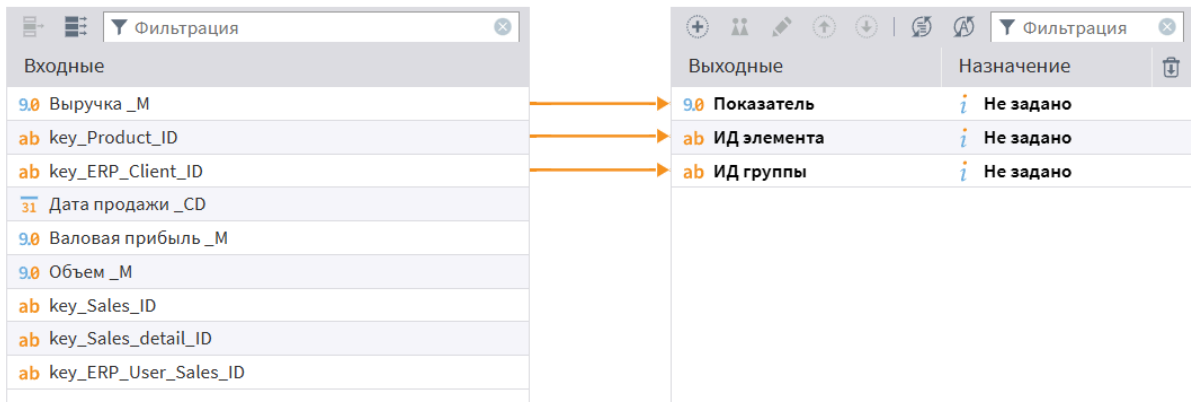
Основной массив анализируемых данных. В качестве анализируемых данных принимаются следующие поля:

- **Показатель** - числовое значение для определения значимости элемента (чем выше значение, тем выше значимость);
- **ИД элемента** - основной элемент, по которому строится ABC анализ. Например, номенклатура или элементы ее иерархии.
- **ИД группы** - группировка элементов, по которой определяется бизнес-логика соседстве элементов. Например, идентификатор чека, чтобы понять какие товары покупаются вместе. Или идентификатор клиента, чтобы понять структуру спроса клиентов.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

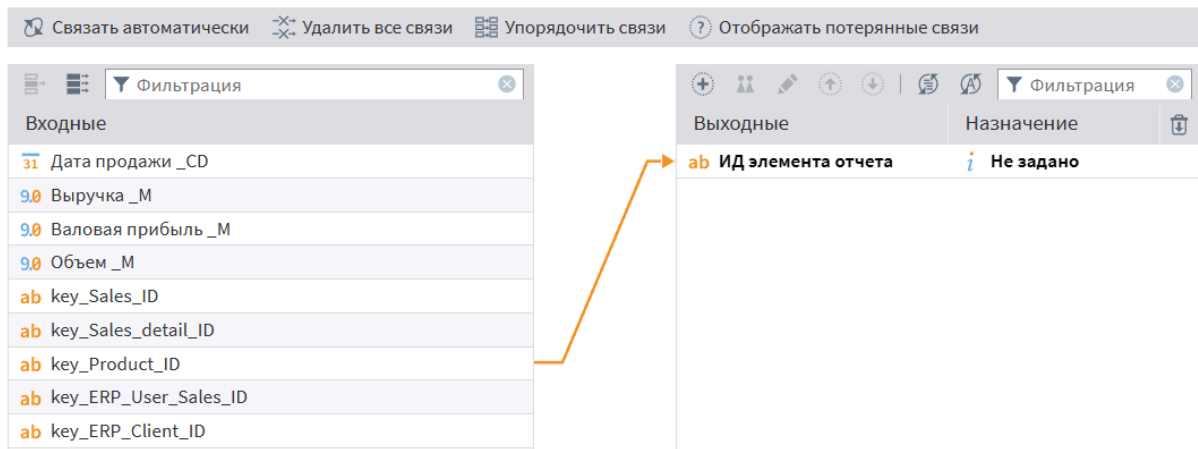
ИП Стучалкин Евгений Игоревич, 504011547861



Элементы отчета

Опциональный порт, куда можно подать список элементов отчета, которые будут включены в итоговый массив данных и расширенный отчет. При этом анализ проводится по всему массиву данных.

Используется для того, чтобы снизить выходной объем данных, оставив в нем только целевые элементы, при этом проведя анализ по всему объему данных.



<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Выходы компонента



ABC-отчет

Содержит агрегированный отчет со статистикой пересечений элементов и групп. Каждому элементу соответствует одна строка. Полям заданы технические наименования, т.к. элементы, показатели и группы могут быть любыми. Для простоты примеров, давайте считать что:

- 1) Показатель отчета - это выручка;
- 2) Элемент - это продукт;
- 3) Группа - это чек.

Звездочками отмечены значения, которые не заполняются при отключенном детальном отчете.

Перечень полей отчета:

- **ИД элемента** - Идентификатор товара;
- **ABC элемента** - Принадлежность товара к ABC сегментам, в соответствии с настройками компонента;
- **Показатель** - Сумма выручки;
- **Показатель сопутствующих элементов*** - Выручка по другим товарам, которые были в чеках с товаром в строке;
- **Доля сопутствующего показателя*** - Какой % выручка сопутствующих товаров составляет от выручки товара в строке;
- **Доля зависимого показателя от итога*** - Какой % выручка сопутствующих товаров составляет от общей выручки;
- **Показатель сопутствующих А-элементов*** - Выручка сопутствующих товаров категории А
- **Показатель сопутствующих В-элементов*** - Выручка сопутствующих товаров категории В

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

- **Показатель сопутствующих С-элементов*** - Выручка сопутствующих товаров категории С
- **Кол-во сопутствующих А-элементов*** - Кол-во уникальных сопутствующих товаров категории А;
- **Кол-во сопутствующих В-элементов*** - Кол-во уникальных сопутствующих товаров категории В;
- **Кол-во сопутствующих С-элементов*** - Кол-во уникальных сопутствующих товаров категории Х;
- **Кол-во групп** - Кол-во чеков с этим товаром;
- **Группы с другими А-элементами*** - Кол-во чеков, в которых были другие товары категории А;
- **Группы с другими В-элементами*** - Кол-во чеков, в которых были другие товары категории В;
- **Группы с другими С-элементами*** - Кол-во чеков, в которых были другие товары категории С;
- **Частота соседства А*** - В каком проценте чеков с данным товаром содержатся другие товары категории А;
- **Частота соседства В*** - В каком проценте чеков с данным товаром содержатся другие товары категории В;
- **Частота соседства С*** - В каком проценте чеков с данным товаром содержатся другие товары категории С;
- **Кол-во А-групп** - Товар входит в данное кол-во чеков с категорией А;
- **Кол-во В-групп** - Товар входит в данное кол-во чеков с категорией В;
- **Кол-во С-групп** - Товар входит в данное кол-во чеков с категорией С;
- **Доля А-групп** - Товар встречается в этом проценте чеков с категорией А;

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

- **Доля В-групп** - Товар встречается в этом проценте чеков с категорией В;
- **Доля С-групп** - Товар встречается в этом проценте чеков с категорией С;

Взаимосвязи

Содержит полный массив данных сопоставлений каждого элемента в каждой группе со всеми другими сочетаемыми элементами. В зависимости от структуры подаваемых на вход данных, кол-во строк может существенно отличаться от исходного в большую сторону. Для уменьшения кол-ва выходных строк рекомендуется ограничивать кол-во элементов в отчете через соответствующий входной порт.

Полям заданы технические наименования, т.к. элементы, показатели и группы могут быть любыми. Для простоты примеров, давайте считать что:

1. Показатель отчета - это выручка;
2. Элемент - это продукт;
3. Группа - это чек.

Поля отчета:

- **ИД группы** - ИД чека
- **ABC группы** - ABC чеков
- **ИД элемента** - ИД товара
- **ИД элемента_comp** - ИД сопутствующего товара
- **Показатель_comp** - Выручка сопутствующих товаров
- **ABC_comp** - ABC сопутствующих товаров
- **ABC** - ABC товара
- **Показатель** - Выручка товара

Этот массив данных нужен для динамического детального анализа результативности сопутствующих элементов.

Типовая структура отчета:

- Измерения: ИД Элемента > ИД элемента_comp > ABC_comp;
- Показатели: Показатель_comp, Кол-во уникальных ИД групп.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

DE: 🗝️ RFM

Компонент создает RFM-сегментацию клиентов и дает расшифровки сегментов

Входы компонента

Настройки

Настройки используются для установления границ сегментов, а также глубины анализируемых продаж.

Метка	Имя	Назначение	Значение
12 1 - день, 2 - неделя, 3 - месяц	Sales_group	Не задано	1
12 Кол-во продаж в анализе	Sales_count	Не задано	5
12 Давность (дней)	sales_prescription	Не задано	5
12 Частые покупки (до X дней)	Freq_control_1	Не задано	15
12 Редкие покупки (более X дней)	Freq_control_2	Не задано	60
12 Маленький чек (до X)	Sum_control_1	Не задано	200 000
12 Большой чек (более X)	Sum_control_2	Не задано	500 000

- **Sales_group** - что считать одной продажей. 1 - все покупки за день, 2 - все покупки за неделю, 3 - все покупки за месяц;
- **Sales_count** - Какое кол-во продаж участвует в анализе;
- **sales_prescription** - Опорный параметр для сегментации давности продажи, относительно даты последней продажи. Нижняя граница динамически определяется как средняя давность продажи. Верхняя давность определяется как средняя + sales_prescription.
- **Freq_control_1** - нижняя граница частоты покупок, то что ниже ее, считается частыми покупками;
- **Freq_control_2** - верхняя граница частоты покупок, то что ниже ее, считается редкими покупками;

Пояснение: давность от частоты отличается тем, что частота проверяется между покупками, а давность считается от последней покупки. Так, вы можете получить сегмент клиентов, которые покупали в среднем часто, но последняя продажа была давно.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

- **Sum_control_1** - нижняя граница среднего чека, то что ниже ее, считается низким чеком;
- **Sum_control_2** - верхняя граница среднего чека, то что выше ее, считается высоким чеком.

Данные

Данные для построения сегментации. На вход подается выгрузка продаж, необходимые поля:

- Дата продажи;
- Сумма продажи;
- ИД клиента.

Входные	Выходные	Назначение
31 Дата продажи_CD	31 Дата продажи	Не задано
9.0 Выручка_M	9.0 Сумма продажи	Не задано
ab key_ERP_Client_ID	ab ИД клиента	Не задано
9.0 Валовая прибыль_M		
9.0 Объем_M		
ab key_Sales_ID		
ab key_Sales_detail_ID		
ab key_Product_ID		
ab key_ERP_User_Sales_ID		

Выходы компонента

RFM

На выходе получаем список клиентов и их RFM-сегментации. По классике, частота покупок, давность покупок, чек покупок кодируются цифрами от 1 до 3. Таким образом каждый клиент получает код вида 312, 123 и т.д.

В нашем отчете представлена расшифровка статусов на основе статьи

<https://mindbox.ru/journal/education/chto-takoe-rfm-analiz/>. Значения сегментов можно

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

подправить под вас в файле Settings/ETL Kit/RFM Классы.xlsx.

DE: 🗝️ Факторный анализ

Компонент для мультипликативного факторного анализа. Позволяет сравнить между собой до 5 сценариев с произвольным набором параметров, перемножение которых дает анализируемый показатель. С помощью полученного датасета можно строить вот такие диаграммы:



А теперь на простом языке. Для мультипликативного факторного анализа нужно выбрать показатель, который декомпозируется на множители (факторы). Например:

Выручка = Кол-во заявок * Конверсию в продажу * Кол-во продукции * Цена единицы продукции

Множителей может быть сколько угодно, но мы возьмем пример выше для простоты. Цель мультипликативного факторного анализа - показать, какой финансовый эффект дало отклонение факторов от эталонного сценария.

Обычно такими эталонными сценариями являются:

- 1) Планы;
- 2) Разные версии бюджетов;

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

3) Фактические значения прошлых периодов.

По умолчанию всем доступен именно третий сценарий, т.к. планирование фин. результатов очень редко идет на супер-детальном уровне.

Таким образом, в факторном анализе должно присутствовать минимум 2 набора множителей: итоговый результат и сценарий, с которым сравниваем. Далее идет особая математика, которая считает, как отклонения факта от эталона повлияло на значение показателя в единицах этого показателя.

Входы компонента



Порядок факторов - перечисление через запятую названия множителей, **в порядке соответствующему бизнес-логике**. Хотя математически порядок множителей не влияет на результат, для осмысленности динамики факторов они должны идти по логическому порядку.

Наименование анализируемого показателя - как называется показатель, который получается в результате перемножения множителей.

Название набора 2-5 - Название эталонных наборов, с которыми будет сравниваться результат. Просто задайте название тем сценариям, которые будут использоваться, а остальные оставьте как есть.

Метка	Имя	Назначение	Значение
ab Порядок факторов	factors_order	Не задано	leads, CV, amount, price
ab Наименование анализиру...	Measure_name	Не задано	Выручка
ab Название набора 2	set2	Не задано	Прошлый год
ab Название набора 3	set3	Не задано	План
ab Название набора 4	set4	Не задано	Бюджет
ab Название набора 5	set5	Не задано	Оперплан



В этот порт может заходить произвольный набор полей, поэтому на нем работает

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

автосинхронизация. Тем не менее, для правильной работы компонента данные должны соответствовать определенной структуре:

- 1) Множители фактического результата должны называться в соответствии с порядком факторов из настроек, и иметь на конце метку набора “_1”: leads_1, CV_1, amount_1, price_1;
- 2) Множители базовых сценариев должны называться аналогично, и иметь метки в конце имени полей “_2” - “_5”. Например: leads_2, CV_2, amount_2, price_2;
- 3) Поля без числовых меток на конце будут использованы как разрезы отчета. Т.е. вы можете подать на вход произвольное число разрезов;
- 4) Порядок полей в таблице не имеет значения.

Структура нейминга для 3-х сценариев и одного разреза (товары):

№	Метка	Имя
	<input type="text" value="Поиск"/>	<input type="text" value="Поиск"/>
1	12 Заявки_1	leads_1
2	90 Конверсия_1	CV_1
3	90 Колво_1	amount_1
4	90 Цена_1	price_1
5	12 Заявки_2	leads_2
6	90 Конверсия_2	CV_2
7	90 Колво_2	amount_2
8	90 Цена_2	price_2
9	12 Заявки_3	leads_3
10	90 Конверсия_3	CV_3
11	90 Колво_3	amount_3
12	90 Цена_3	price_3
13	ab Товары	Products

После выполнения компонента, вы получите на выходе прежнюю таблицу с дополнительными полями, содержащими абсолютные значения отклонений для каждого сценария.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

№	Метка	Имя	Значение
	<input type="text" value="Q Поиск"/>	<input type="text" value="Q Поиск"/>	
4	90 price_1	price_1	521,00
5	12 leads_2	leads_2	50
6	90 CV_2	CV_2	0,20
7	90 amount_2	amount_2	20,00
8	90 price_2	price_2	472,00
9	12 leads_3	leads_3	80
10	90 CV_3	CV_3	0,25
11	90 amount_3	amount_3	40,00
12	90 price_3	price_3	300,00
13	ab Products	Products	Огурцы
14	90 Выручка_Результат	Vyruchka_Rezultat	199 282,50
15	90 Выручка_Прошлый год	Vyruchka_Proshliyi_god	94 400,00
16	90 Выручка_План	Vyruchka_Plan	240 000,00
17	90 Дельта_Прошлый год	Delta_Proshliyi_god	104 882,50
18	90 Дельта_План	Delta_Plan	-40 717,50
19	12 leads_2_D	leads_2_D	25
20	90 CV_2_D	CV_2_D	-0,03
21	90 amount_2_D	amount_2_D	10,00
22	90 price_2_D	price_2_D	49,00
23	12 leads_3_D	leads_3_D	-5
24	90 CV_3_D	CV_3_D	-0,08
25	90 amount_3_D	amount_3_D	-10,00
26	90 price_3_D	price_3_D	221,00
27	90 Заявки_Прошлый год	Zayavki_Proshliyi_god	47 200,00
28	90 Конверсия_Прошлый год	Konversiya_Proshliyi_god	-21 240,00
29	90 Колво_Прошлый год	Kolvo_Proshliyi_god	60 180,00
30	90 Цена_Прошлый год	TSena_Proshliyi_god	18 742,50
31	90 Заявки_План	Zayavki_Plan	-15 000,00
32	90 Конверсия_План	Konversiya_Plan	-72 000,00
33	90 Колво_План	Kolvo_Plan	-38 250,00
34	90 Цена_План	TSena_Plan	84 532,50

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стулалкин Евгений Игоревич, 504011547861

В из данных показателей можно составить визуализацию влияния на динамику, расположив показатели в следующем порядке:

1. Имя_показателя_Имя_сценария;
2. Имя_фактора_Имя_сценария, все по порядку;
3. Имя_показателя_Результат.

Есть также контрольное поле Дельта_Имя_сценария, оно содержит сумму всех факторов.

	Выручка_План	Заявки_План	Конверсия_План	Колво_План	Цена_План	Выручка_Результат	Дельта_План
Огурцы	240 000,00	-15 000,00	-72 000,00	-38 250,00	84 532,50	199 282,50	-40 717,50
Помидоры	309 825,00	-51 637,50	-86 062,50	-37 125,00	75 000,00	210 000,00	-99 825,00
Итого:	549 825,00	-66 637,50	-158 062,50	-75 375,00	159 532,50	409 282,50	-140 542,50

Выходные данные можно подать в компонент **DE: Факторный анализ - интерпретация**, чтобы получить с агрегированный отчет, пригодный для рассылки.

DE: Факторный анализ - интерпретация

Создает на основе данных факторного анализа html-форматированное сообщение с настраиваемой выжимкой.

DE: Стабильность продаж

Компонент обогащает данные продаж сегментами давности работы с клиентами. Таким образом позволяя сегментировать продажи на те, которые идут по новым клиентам, по вернувшимся, и по постоянным клиентам с разной давностью работы (год-два-три и более).

Особенность отчета в том, что состояние клиентской базы определяется на каждый месяц данных, т.е. клиент в первый месяц является новым, а во второй месяц переходит в разряд “первый год работы”.

Также, компонент выдает статистику по оттоку клиентской базы.

Входы компонента



Настройки

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

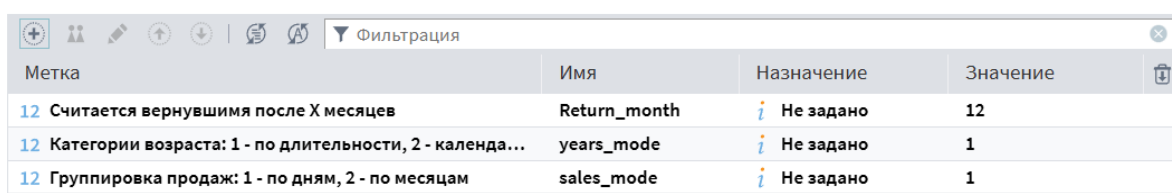
<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Return_month - через какое кол-во месяцев пропущенных продаж история сотрудничества клиента обнуляется, и он при покупке будет считаться как “вернувшийся”, с наработкой длительности сотрудничества с нуля.

years_mode - как определяется кол-во лет клиента. 1 - считается относительно первой покупки клиента. 2 - считается по календарным годам. Так, клиент купивший в декабре 2025 года, в 2026 будет считаться двухлетним.

sales_mode - как группируются продажи. 1 - по дням, 2 - по месяцам.



Метка	Имя	Назначение	Значение
12 Считается вернувшимся после X месяцев	Return_month	Не задано	12
12 Категории возраста: 1 - по длительности, 2 - календа...	years_mode	Не задано	1
12 Группировка продаж: 1 - по дням, 2 - по месяцам	sales_mode	Не задано	1

Даты первых продаж

По умолчанию, дата первой продажи определяется как минимальная дата покупки из данных. Можно переопределить эту дату, подав таблицу сопоставлений клиентов и дат первых продаж в этот порт. Если данные поданы и клиенту не присваивается дата из этого порта, она будет взята из данных продаж.

Данные

ИД клиента, Дата продажи, Сумма продажи.

Выходы компонента

Результат

Выводит сумму продаж, сегментированную по месяцам, клиентом и давности работы с ними. Дополнительно выводит месяц первой продажи, что позволяет строить отчеты когортного анализа.

Факторный анализ YoY

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Заготовка для влияние изменений возраста клиентской базы на выручку.



Анализ оттока

Выводит состояние клиентской базы на каждый месяц, включая клиентов которые выпали из активной работы.

DE: Товарные рекомендации

За счет обработки истории продаж, компонент позволяет получить на выходе список клиентов и перечень продуктов, который рекомендуется им к допродажам. Список допродаж формируется по различным правилам.

Принцип работы компонента:

- 1) Берется история продаж, которая содержит список продукции, которая считается купленной и не должна попасть в рекомендации;
- 2) Берется история продаж, на базе которой будут строиться рекомендации;
- 3) Определяется, для какого кол-ва клиентов будут сформированы рекомендации (по убыванию фин. результата);

Далее, по этой создается несколько рекомендательных списков товаров:

- 1) Топ X товаров по фин. результату;
- 2) Топ X товаров у топ Y клиентов по фин. результату;
- 3) Выборка сопутствующих товаров категории A по фин. результату при условии, что данные товары встречаются более чем у X клиентов;
- 4) Список MML (Minimum Must List) - ваш список товаров, которые должны быть куплены каждым клиентом;
- 5) Наборы товаров - ваш список сопоставлений товаров вида "взял товар A - значит должен взять товар B".

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

Входы компонента

Настройки






Top_x_products - кол-во топ продуктов для списка рекомендаций;

Top_x_clients - кол-во топ клиентов для дополнительной выборки топ продуктов;

Top_x_clients_recommendations - кол-во топ клиентов, для которых будут созданы рекомендации;

accompanying - включать в рекомендации сопутствующие товары категории А по фин. результату (может замедлить формирование рекомендаций);

accomp_threshold - сколько клиентов должны купить набор продуктов, чтобы он считался сопутствующим.

Метка	Имя	Назначение	Значение
12 1. Топ X товаров для рекомен...	Top_x_products	 Не задано	10
12 2. Топ X клиентов для эталона	Top_x_clients	 Не задано	10
12 3. Рекомендации для топ X кл...	Top_x_clients_recommendations	 Не задано	100
0/1 Анализ сопутствующих	accompanying	 Не задано	true
12 4. Порог сопутствующих (кол...	accomp_threshold	 Не задано	100

Данные для рекомендаций

Сюда подается история продаж, которая будет использована для построения рекомендаций.

История продаж

Сюда подается история продаж, которая будет содержать исключения из рекомендаций. Товары из истории продаж не будут включены в рекомендации.

Примечание: есть смысл разбивать продажи на однородные клиентские сегменты, чтобы рекомендации были логически обоснованы.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861



MML

Сюда подается список ИД товаров, которые должен купить каждый клиент, независимо не от чего.



Сопутствующие товары

Сюда подается список пар ИД товара - ИД сопутствующего товара. Если клиент купил первый товар и не купил второй, второй будет включен в рекомендацию.

Выходы компонента



Товарные рекомендации

Возвращает таблицу со списком ИД клиентов, ИД товаров для этого клиента, а также пометкой, по какой логике дана рекомендация.

DE: Воронка Lite

Компонент анализирует статистику закрытия сделок из CRM, для того чтобы:

- Сформировать обогатить данные для визуализации в BI-системах и отчетах динамического отчета по воронке;
- Провести сравнительный анализ менеджеров по продажам на разных этапах воронки, подсветить отклонения;
- Вывести итоговый отчет по воронке.

Отчет строится на упрощенных данных (учитывается только статус успешного закрытия сделки, история сделки в других статусах не представлена).

Входы компонента



Настройки

Позволяют задать границы критичных и некритичных отклонений при сравнении

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

конверсий менеджеров при переходе с этапов на этап в процентах. 1% = 0,01. Границы используются для оценки отклонений как вниз так и вверх. Отклонение задается как относительное, а не абсолютное. Т.е. если у одного менеджера конверсия с этапа на этап 10%, а в среднем по воронке у всех менеджеров 20%, то у первого по сравнению со средним отклонение -50%, а не -10%.

Конверсии считаются от даты создания сделок.

Сделки

Сюда подается список закрытых сделок. Для сделок, закрытых неуспешно, нужно указывать идентификатор статуса проигрыша. Входные поля:

- Воронка
- ИД сделки
- ИД менеджера
- ID статуса проигрыша
- Бюджет
- Дата создания
- Дата закрытия

Воронки

Сюда подается справочник воронок для обогащения отчета. Входные поля:

- Название воронки
- ИД статуса
- Название статуса
- Порядок статуса

Выходы компонента

Данные

Массив данных, подходящий для визуализации в BI, для построения отчета по воронкам и отчету по циклу сделок.

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

- Воронка
- ИД статуса проигрыша
- Статус проигрыша
- Порядок статуса
- Кол-во успешных
- Бюджет успешных
- Кол-во проигранных
- Бюджет проигранных
- ИД сделки
- Бюджет
- Дата создания
- Дата закрытия
- ИД статуса
- Название статуса
- Цикл успешных сделок
- Цикл проигранных сделок
- ИД успешной сделки
- ИД проигранной сделки
- Статусов в воронке
- Успешная сделка
- ИД менеджера



Отчет по менеджерам

Агрегированный отчет по менеджерам, со сравнением их отклонений из конверсий со средними значениями воронки. А также с различными вариантами расчета конверсии.

- Отклонение конверсии от среднего
- Статус отклонения - уровень критичности отклонения на основе настроек;
- % переходов
- % переходов total
- Воронка
- Название статуса
- Кол-во сделок

<https://marketplace.loginom.ru/solutions/data-monetization-pack>

<https://dm-lab.pro/>

ИП Стучалкин Евгений Игоревич, 504011547861

- % от всех сделок - какой % от всех сделок дошел до этапа;
- % успешных - какой % из сделок, дошедших до этапа станет успешным;
- ИД менеджера
- Кол-во сделок в воронке
- Кол-во успешных сделок
- % от всех сделок total
- % успешных total
- Кол-во сделок total
- Кол-во сделок в воронке total
- Кол-во успешных сделок total
- Описание этапа



Отчет по воронке

Как отчет по менеджерам, только с итоговыми показателями всей воронки.