

Обработка данных с использованием Loginom

Работа в аналитической low-code платформе Loginom

Кафедра прикладной математики ИИТ РТУ МИРЭА

Среда Logiном для работы с данными

Пакеты, сценарий, компоненты, узлы

Данные, переменные, входные и выходные порты

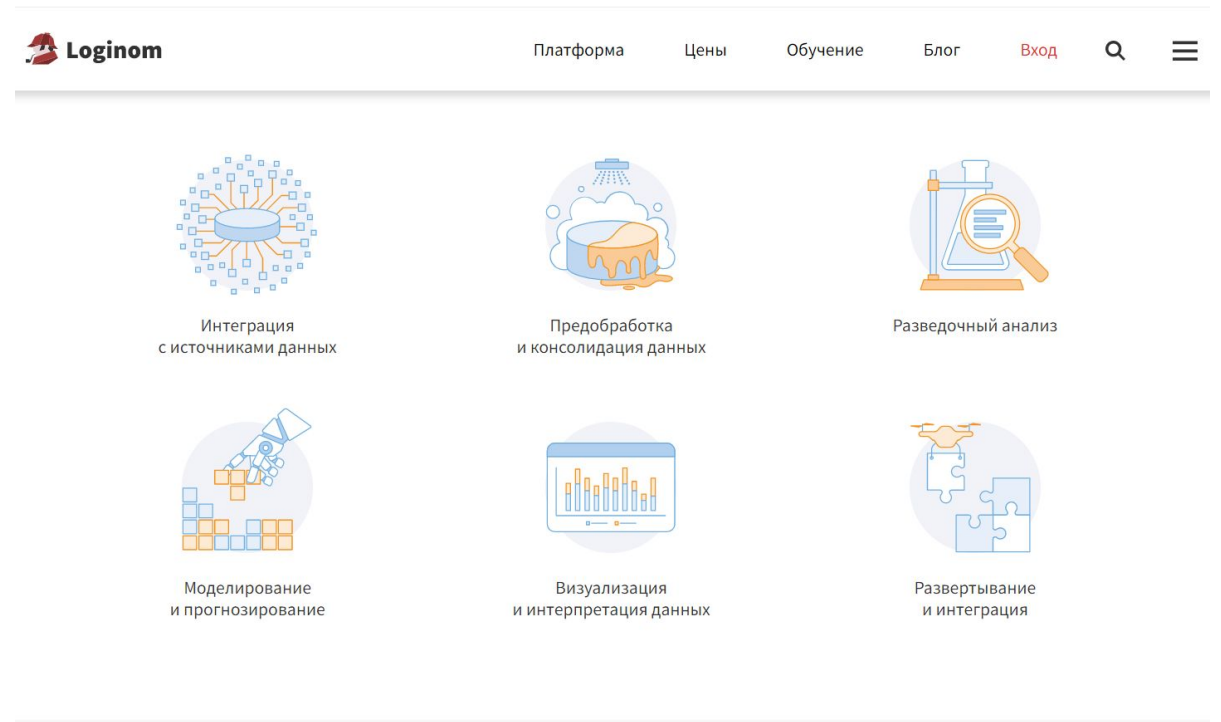
Настройка узлов, действия с узлами

Построение сценария на основе ориентированного графа

Low-code платформа Loginom

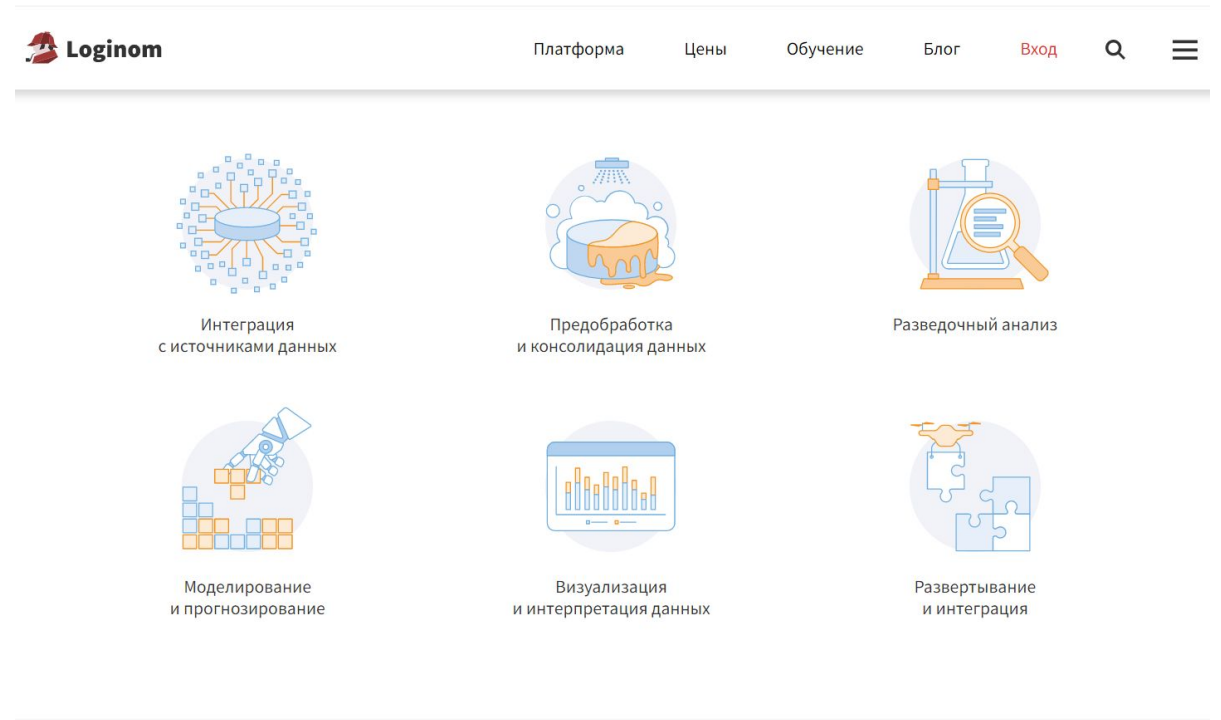
Loginom — интегрированная Low-code платформа для реализации всех аналитических процессов: от консолидации и подготовки данных до моделирования, развертывания и визуализации

- Визуальная настройка логики обработки и повторное использование компонентов
- Продвинутая аналитика: от простейших формул до машинного обучения
- Интеграция данных: файлы, базы данных, учетные системы, включая Excel и 1С: Предприятие
- Доступ к ClickHouse
- Экспорт в Tableau
- Подключение и вызов REST-сервисов
- Высокая производительность: in-memory, параллельная обработка, быстрые алгоритмы



Low-code платформа Loginom

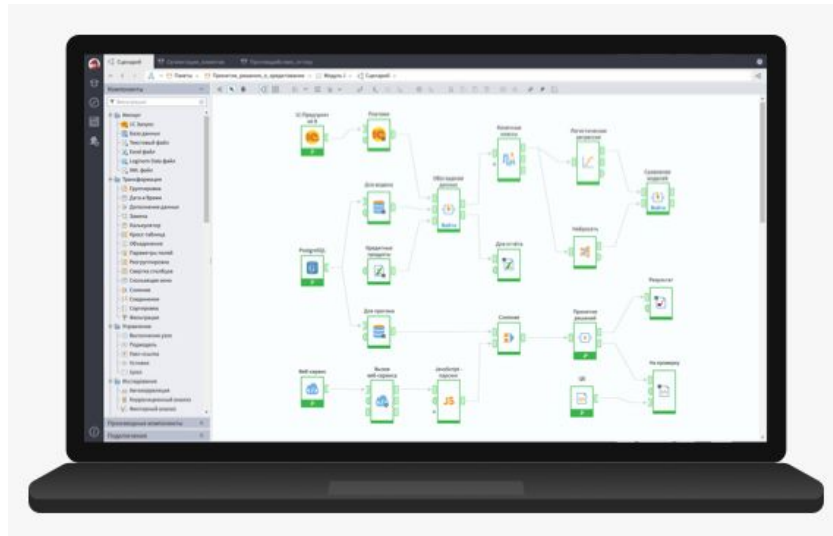
- Позволяет просто анализировать данные без использования кода
- Визуальный интерфейс и система сценариев
- Разработанная логика в блоках обработки данных
- Быстрая работа с большими объемами данных
- Бесплатная пользовательская версия
- Поддержка большого количества форматов данных



Как использовать Loginom

- Скачать версию Loginom Community Edition 6.5.4, заполнив форму по ссылке:
- <https://loginom.ru/download>

- Использовать демо-стенд на сайте:
- <https://demo.basegroup.ru/>
- Данные сбрасываются каждую неделю, общий доступ к файлам и проектам



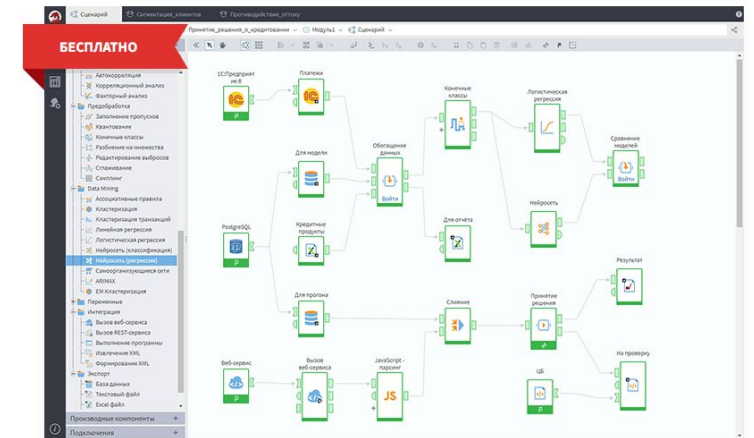
Открытый демостенд Loginom Enterprise

CPU Intel Xeon 12 ядер, 16 ГБ ОЗУ

НЕ ЗАГРУЖАЙТЕ КОНФИДЕНЦИАЛЬНЫЕ ДАННЫЕ!
Загруженные данные доступны любому пользователю.
Каждое воскресенье все данные и пакеты удаляются.

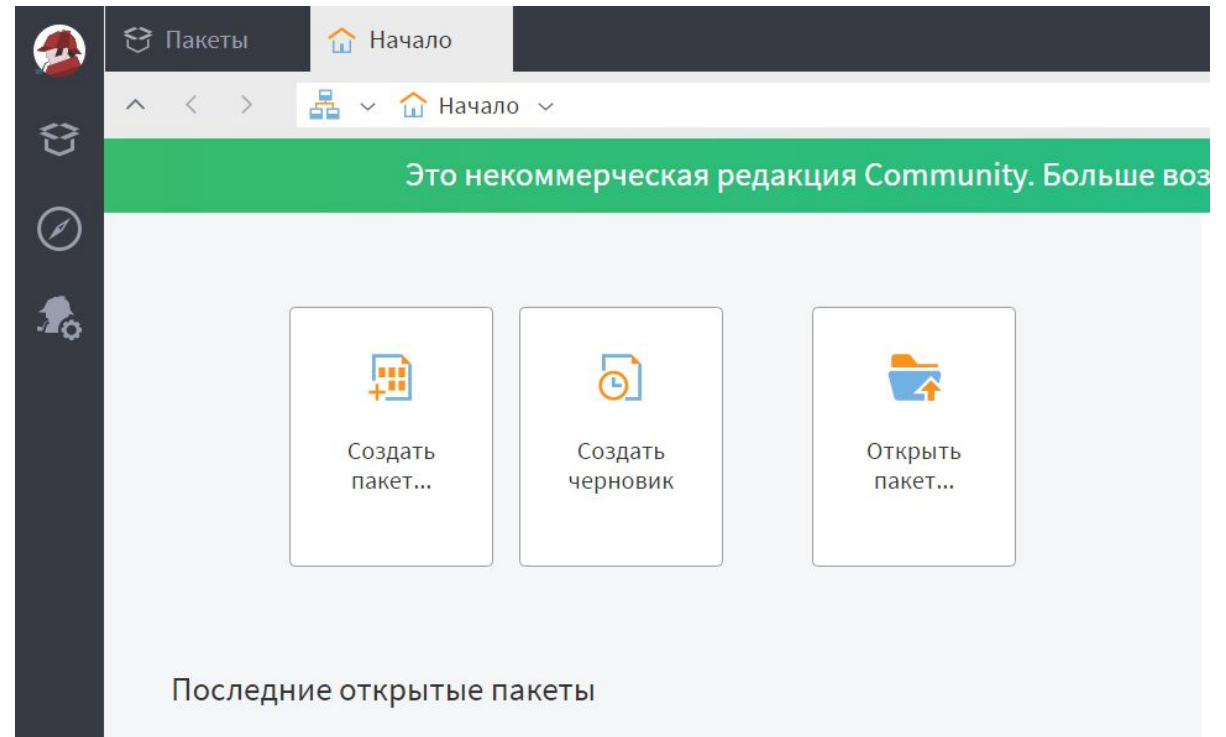
Перейти

Если автоматическая авторизация не работает —
логин `user` без пароля



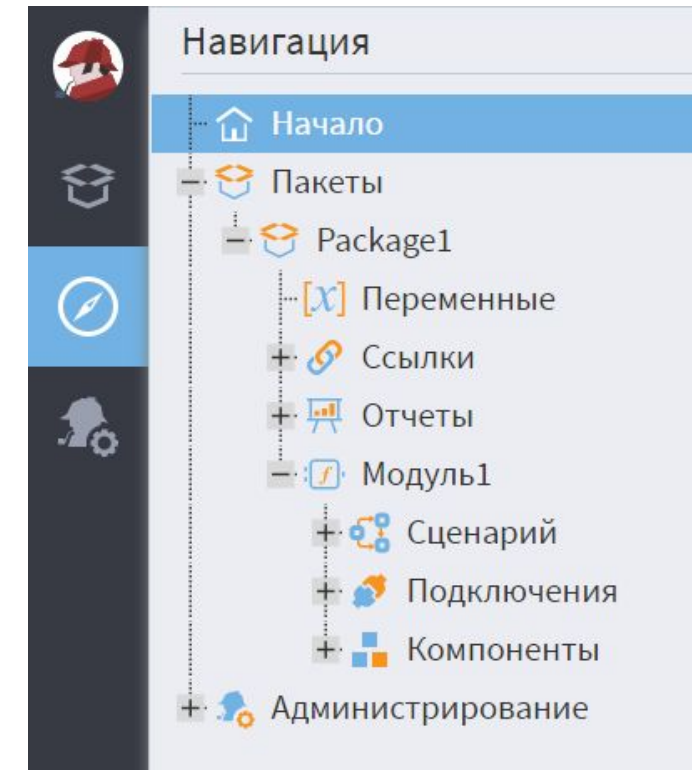
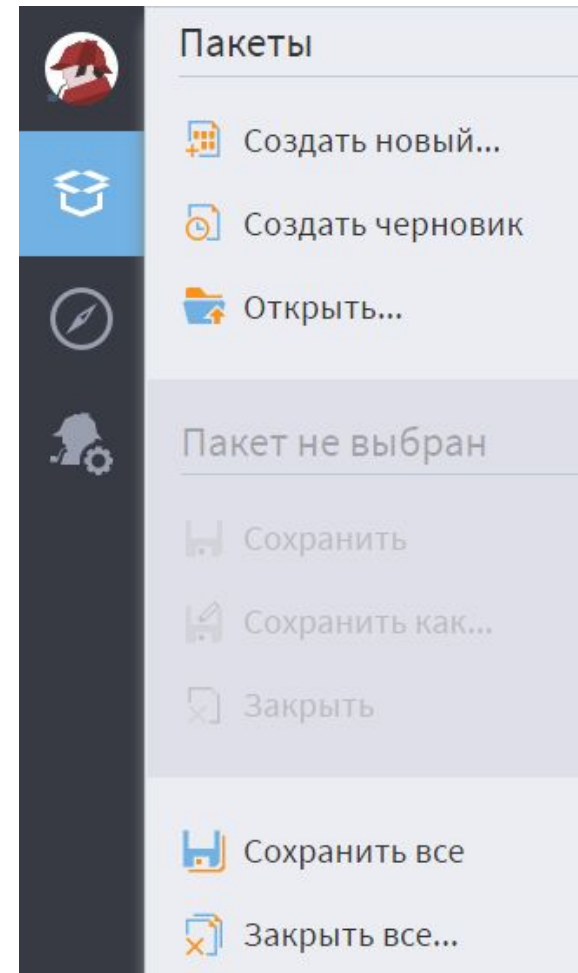
Визуальный интерфейс

- Для каждого проекта можно создать отдельный пакет со сценариями обработки данных, подключениями и визуализациями
- Для черновой работы можно производить все те же операции с помощью черновика. Для сохранения его как пакета вручную пройти процесс сохранения
- Открытие ранее созданных пакетов
- Обратная совместимость пакетов для различных версий Logiном



Навигация

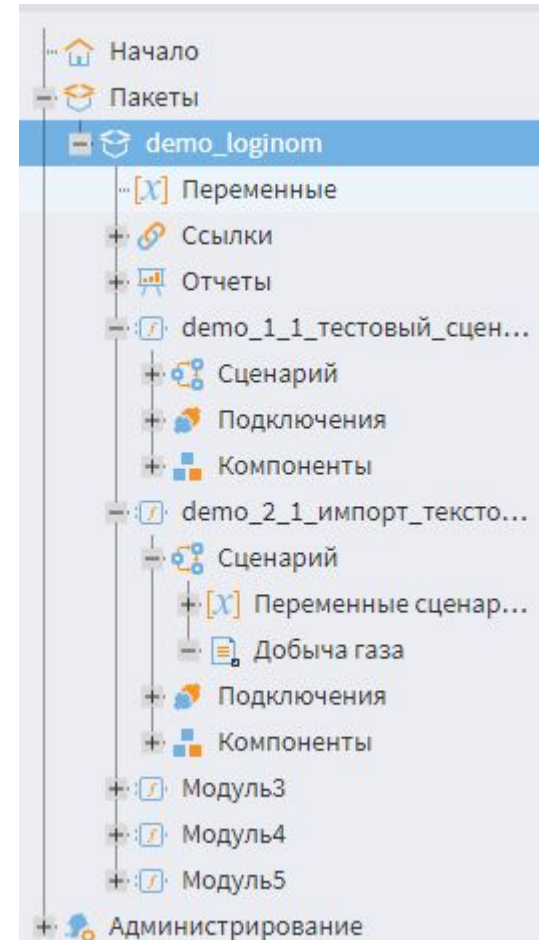
- В кнопке «Пакеты» можно найти интерфейс взаимодействия с пакетами Loginom
- По кнопке «Навигация» открывается древовидная структура проекта
- Проект (пакет) состоит из модулей, каждый из которых решает задачу обработки данных
- Модульная структура проекта позволяет отделять логику обработки данных на разных стадиях



Модули и сценарии

Каждый модуль состоит из:

- Вкладки **сценарий**, в котором производится обработка данных средствами Loginom;
- Вкладки **подключения**, которая отвечает за соединение с внешними источниками, типа РСУБД или REST-сервисов
- Вкладки **компоненты** для продвинутой работы с переиспользованием логики

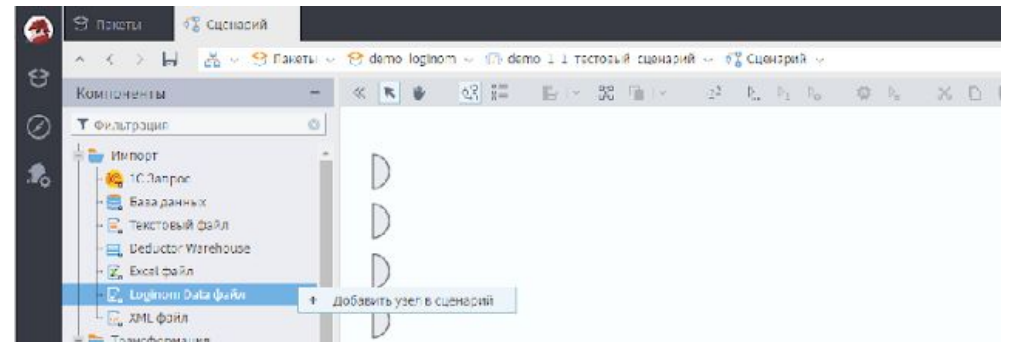
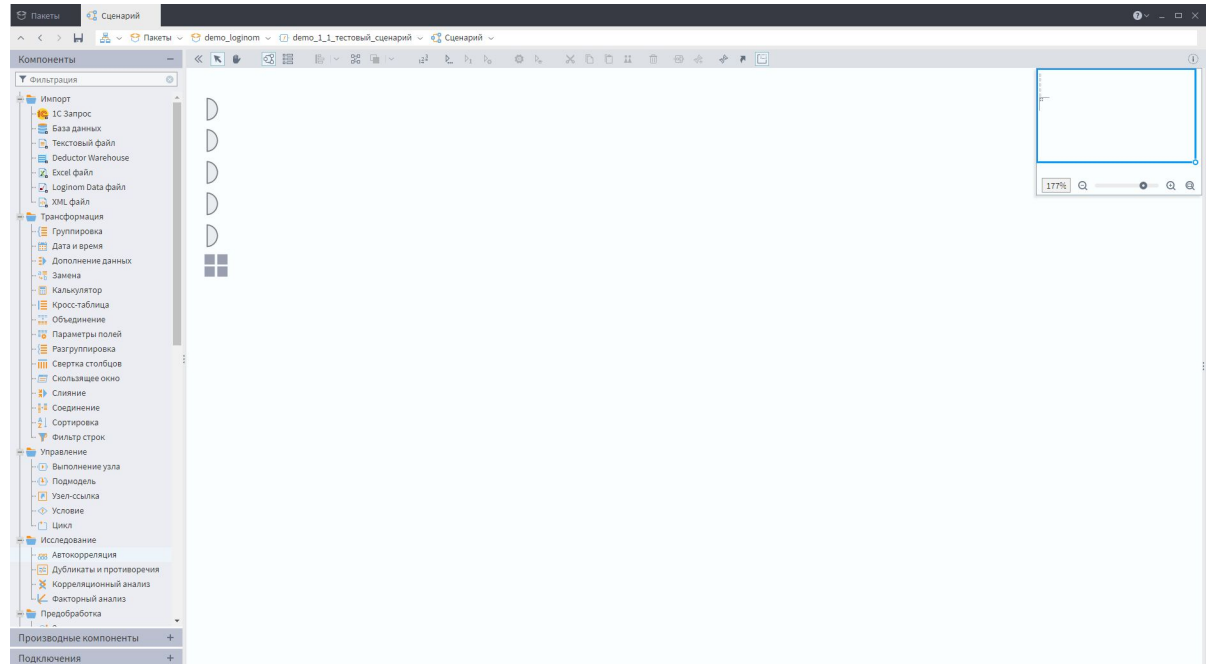


Модули и сценарии

Создав модуль и перейдя к сценарию открывается пустое полотно. Сценарий позволяет выносить на него узлы компонентов обработки данных.

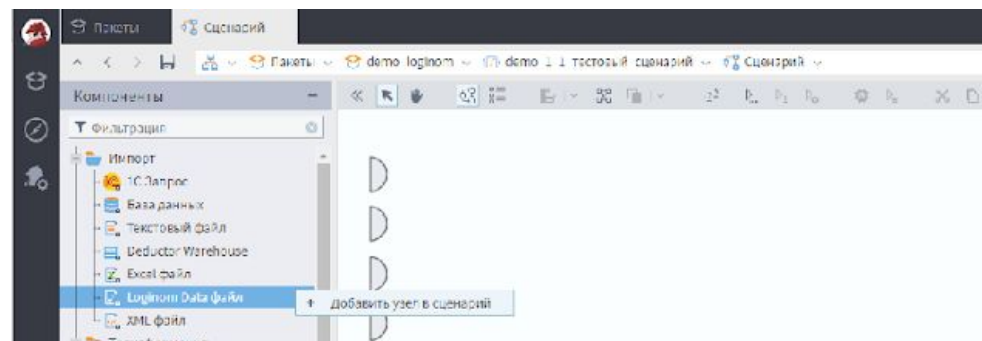
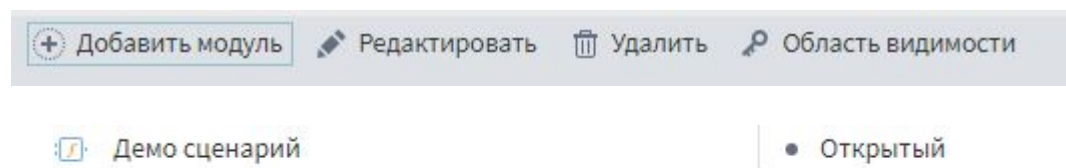
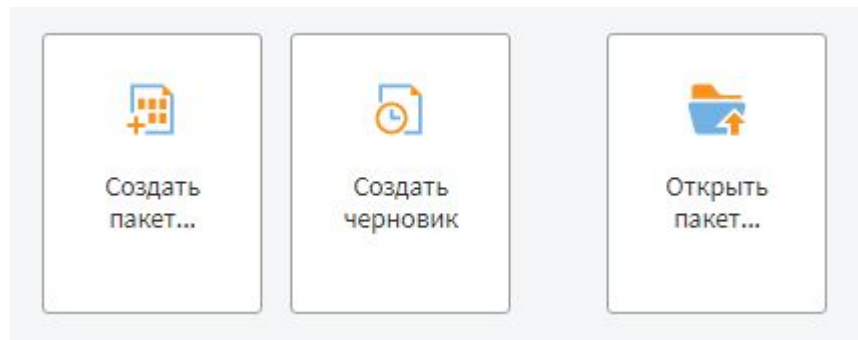
Слева в сценарии находятся компоненты обработки данных, которые можно с помощью перетаскивания мышью перенести на полотно.

Каждый компонент имеет предназначение и заранее реализованную логику работы с данными.



Задание №1. Создание пакета Loginom

1. Создайте пакет Loginom одним из удобных для вас способов
2. Сохраните пакет под названием «Практическая работа №3»
3. Создайте модуль внутри пакета с названием «Демо сценарий» (функция **редактировать**)
4. Перенесите на полотно компонент «Loginom Data файл»
5. Сохраните пакет.



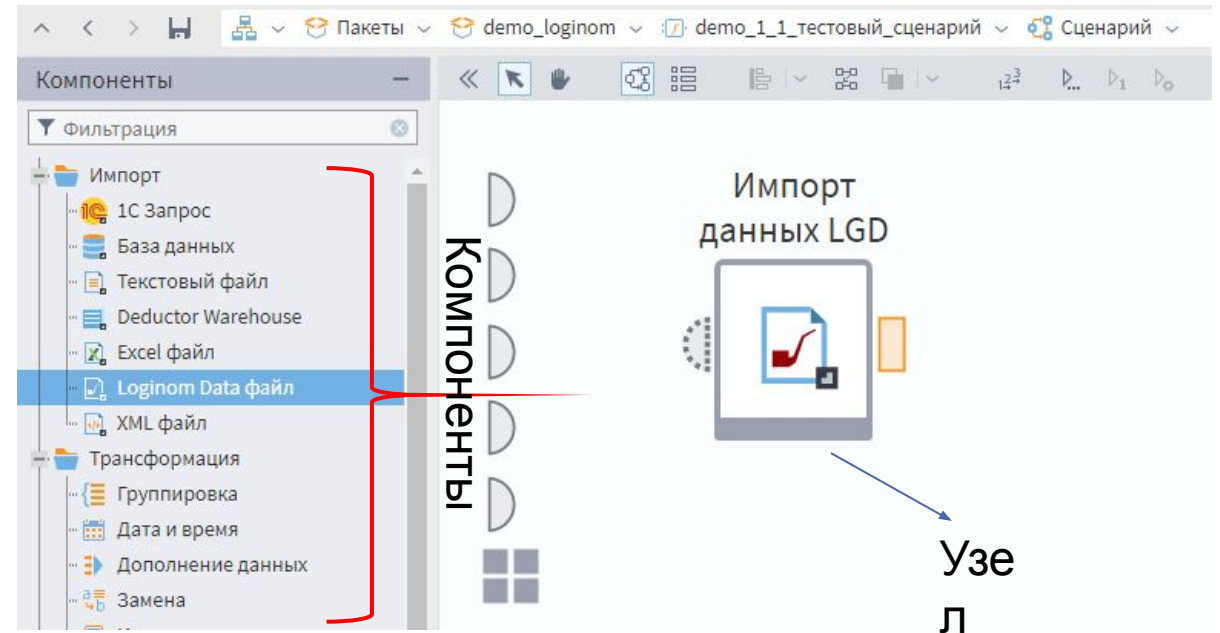
Компоненты и узлы

Компоненты обработки данных – разработанные базовые единицы логики обработки данных, на базе которых создаются узлы.

Узлы – конкретные блоки сценария обработки данных, реализующие функции компонента в своей конкретной реализации.

Узлы – физически расположенные блоки компонентов из панели компонентов.

Узел – объект, **компонент** – класс.

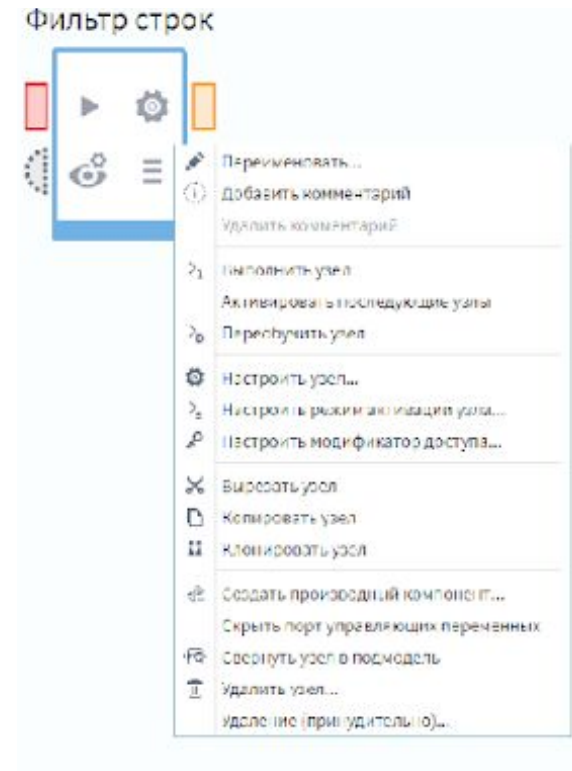


Настройка узлов, входные и выходные порты

Каждый узел обладает входными и выходными портами разных типов:

- Порт **таблица** (прямоугольный) хранит табличную структуру данных
- Порт **переменных** (полукруглый) хранит единичные значения для каждого имени. Переменные используются для параметризации сценария

Настроить узел можно с помощью **шестеренки**, нажав на него ЛКМ.



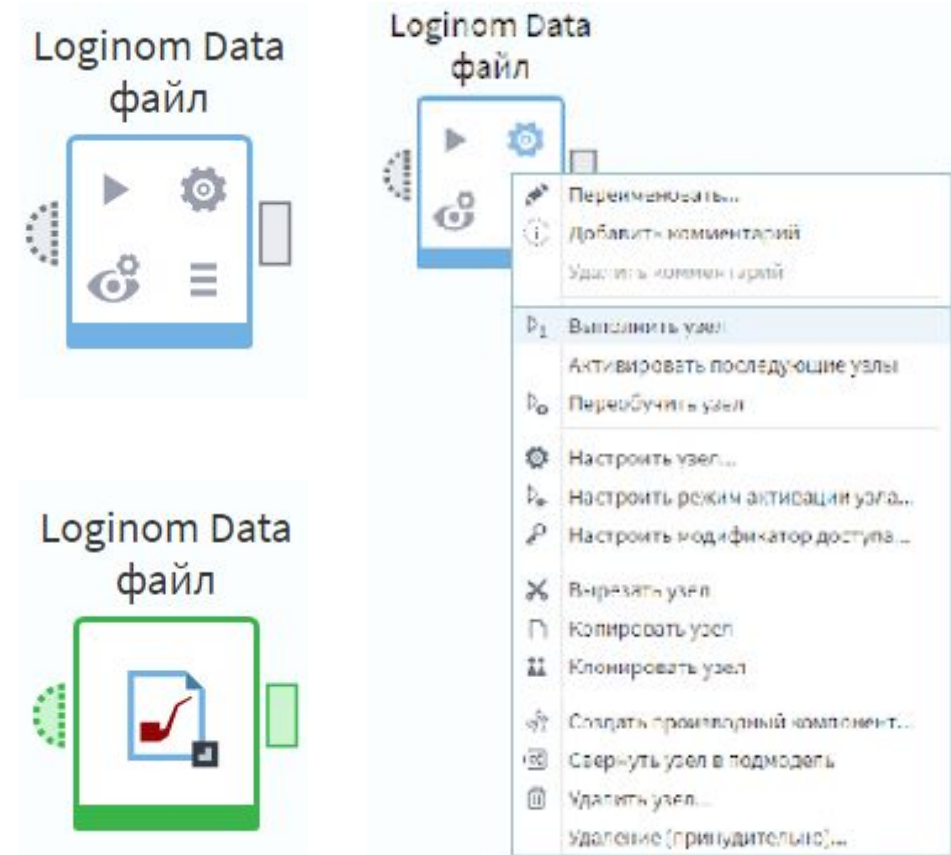
Активный узел и приостановка работы узла

Перед началом работы узла его необходимо настроить.

Настроенный узел является активным в том случае, если он подсвечен зелёным.

Запустить узел и сделать его активным можно с использованием кнопки «Выполнить узел» при нажатии ЛКМ на сам узел или нажатии ПКМ и выбора соответствующей функции.

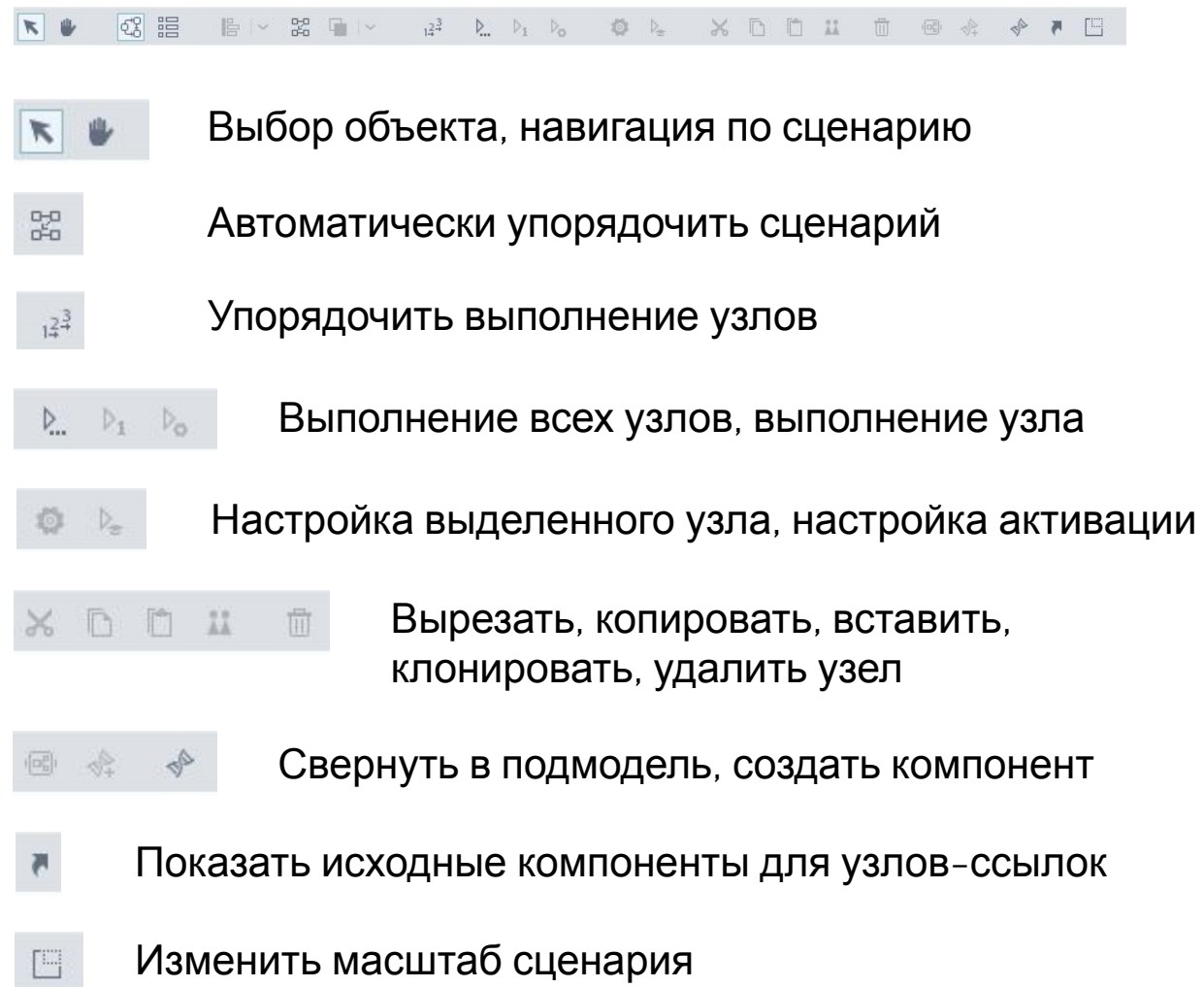
Остановить выполнение можно кнопкой «Деактивировать узел».



Возможности сценария Loginom

На верхней панели сценария расположены полезные функции управления сценарием и объектами в нём.

Во всех вкладках Loginom реализованы похожие панели, экспериментируйте, чтобы освоить возможности работы с инструментом.

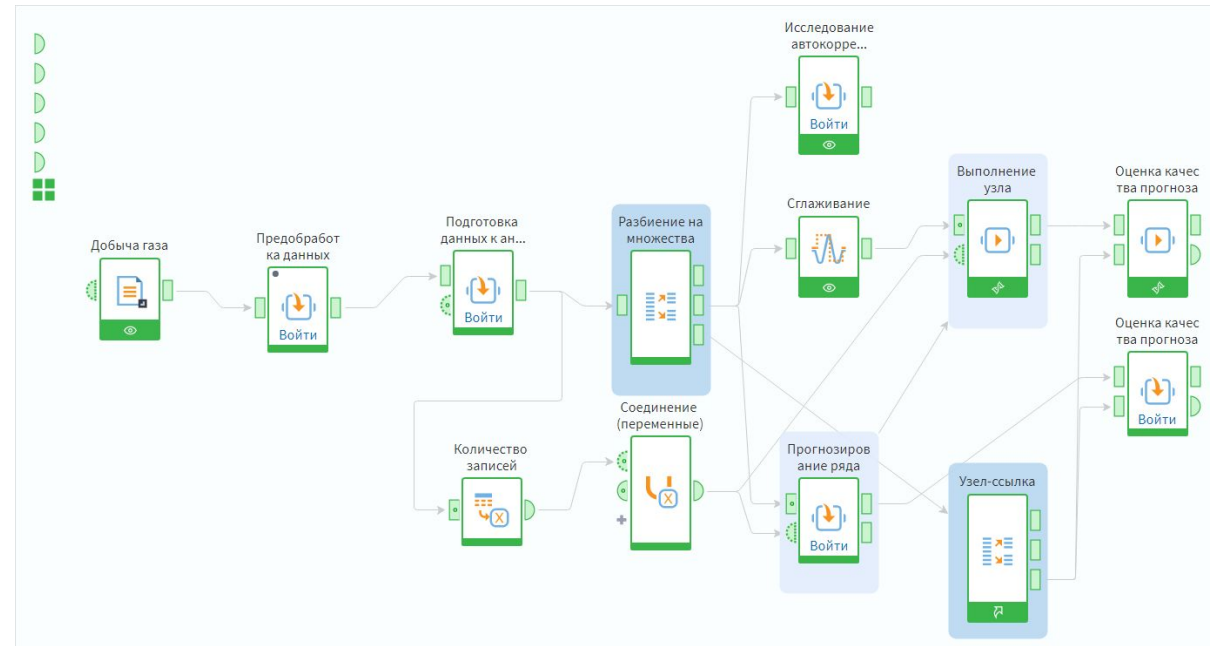


Сценарий обработки данных

LogiNot предназначен для конвейерной пакетной обработки данных с использованием эффективных реализаций алгоритмов.

Сценарий на основе LogiNot позволяет гибко решать различные задачи обработки и анализа данных на основе методов трансформации, анализа, моделирования и экспорта.

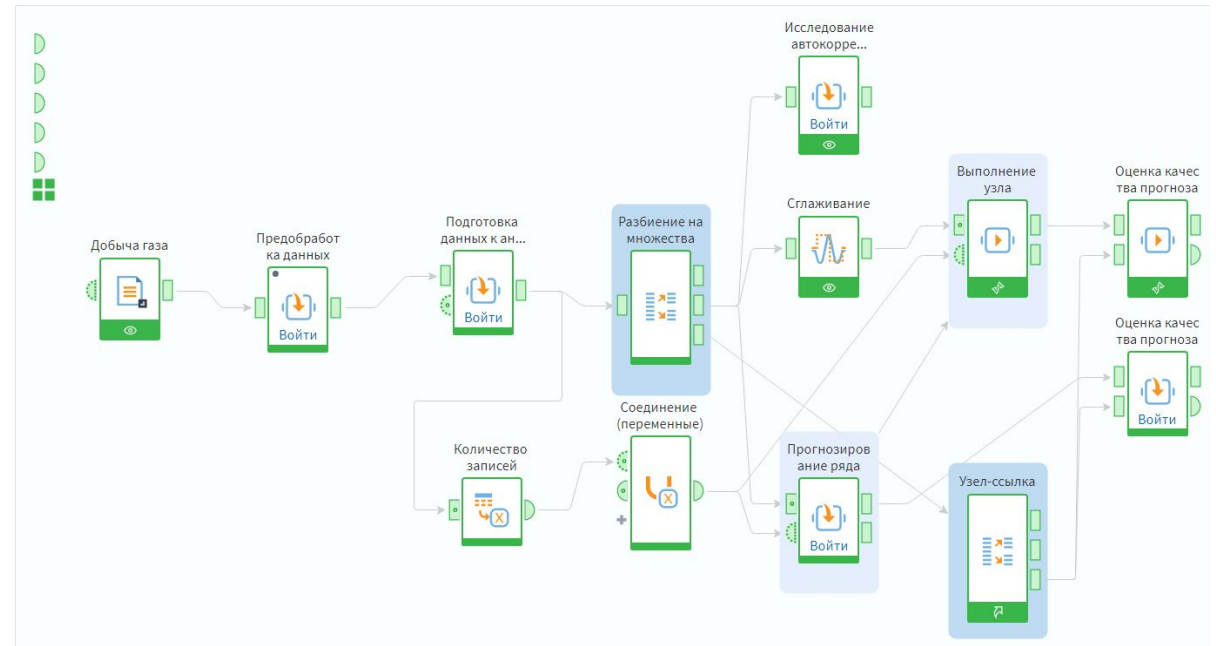
Для составления сценария обработки данных достаточно знать методы обработки данных на уровне ВХОД-ВЫХОД.



Сценарий обработки данных

Узлы сценария Logiном можно последовательно соединять портами схожего типа от выходных портов ко входным портам, начиная с импорта данных, заканчивая экспортом данных.

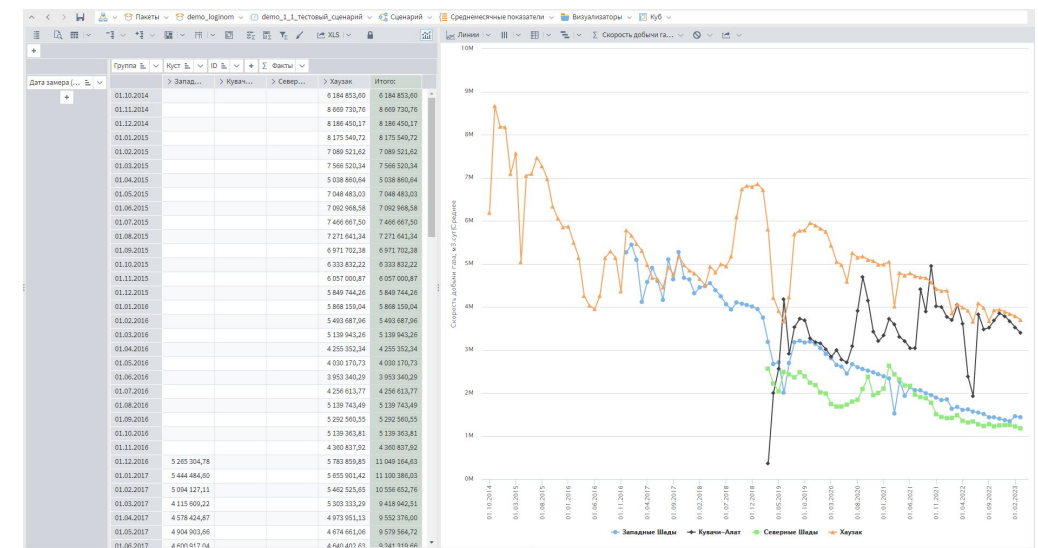
Сценарий Logiном представляет собой ориентированный ациклический граф (DAG) без циклов, который реализует процесс обработки данных на основе узлов заранее реализованных компонентов.



Задание №2. Демо сценарий Loginom

1. Откройте пакет demo_loginom.lgd
2. Откройте модуль demo_1_1_тестовый_сценарий
3. Запустите все узлы кнопкой «**ВЫПОЛНИТЬ ВСЕ**», убедитесь в том, что сценарий выполняется без ошибок
4. Исследуйте выходные и входные порты узлов на основе **настройки входных портов и предварительного просмотра порта данных**
5. Нажмите на «глаз» под узлом **среднемесячные показатели**, просмотрите визуализацию

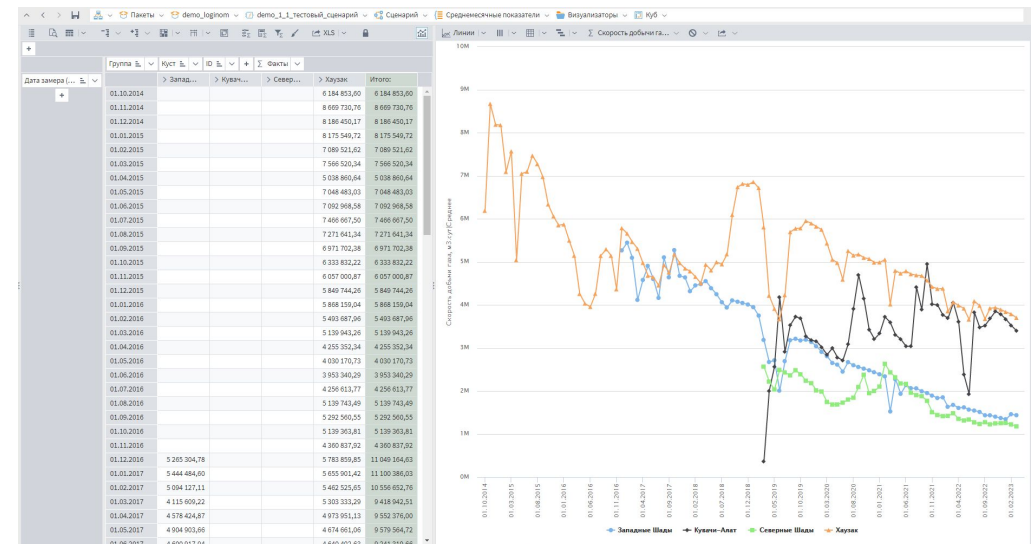
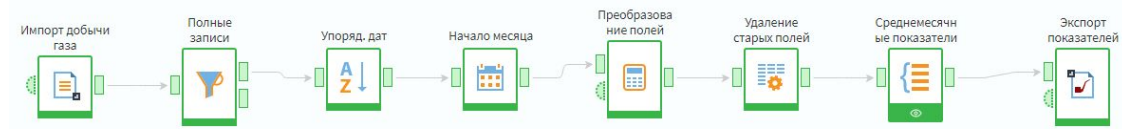
Метка	Область видимости
Переменные	
Ссылки	
Отчеты	
demo_1_1_тестовый_сценарий	● Открытый
demo_2_2_импорт_текстовый_файл	● Открытый
demo_2_1_импорт_lgd	● Открытый
demo_2_3_импорт_бд	● Открытый
demo_3_1_Фильтрация_данных	● Открытый



Задание №2. Демо сценарий Loginom

- Сколько полей имеет таблица данных на выходном порте «Импорт добычи газа»?
- Есть ли пропуски в записях импортируемой таблицы данных?
- Сколько полных записей на верхнем порте узла «Полные записи»?
- Какое поле появилось после выполнения узла «Начало месяца»?
- Какие поля добавляются в калькуляторе «Преобразование полей»?
- Сколько записей остается после агрегации на выходе из узла «Среднемесячные показатели»?

Метка	Область видимости
Переменные	
Ссылки	
Отчеты	
demo_1_1_тестовый_сценарий	Открытый
demo_2_2_импорт_текстовый_файл	Открытый
demo_2_1_импорт_lgd	Открытый
demo_2_3_импорт_бд	Открытый
demo_3_1_Фильтрация_данных	Открытый



Импорт и экспорт данных в Loginom

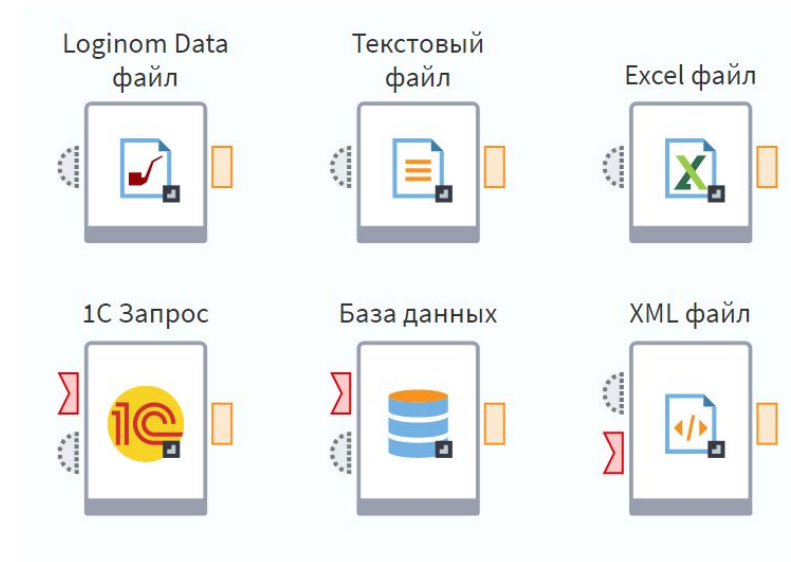
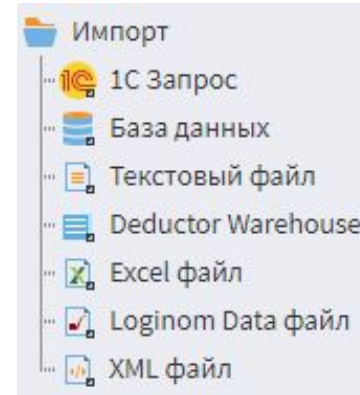
Импорт и экспорт данных .lgd файлов

Импорт и экспорт данных .txt, .csv, .tsv файлов

Импорт из баз данных

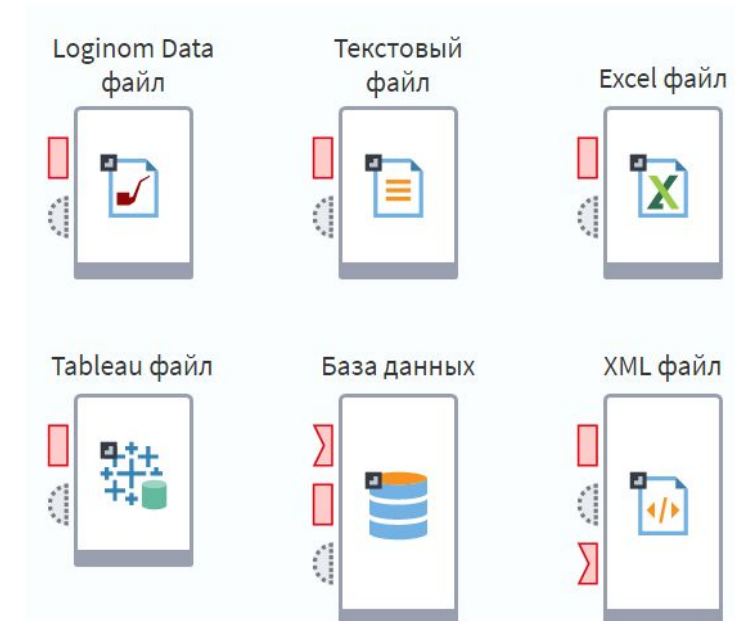
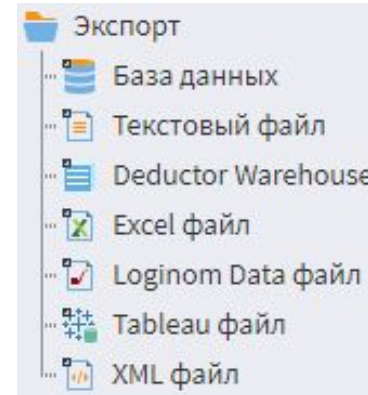
Импорт данных в Loginom

- Импорт данных в Loginom осуществляется посредством использования узлов компонентов из вкладки «Импорт».
- Можно осуществлять доступ к структурированным данным из различных источников, настраивая соединение или указывая тип разметки импортируемого файла.
- Можно получать доступ к данным из файлов и из внешних баз данных.



Экспорт данных из Loginom

- Экспорт данных из Loginom осуществляется посредством использования узлов компонентов из вкладки «Экспорт».
- Имеется возможность создать файл выбранного типа и структуры из созданного в Loginom набора данных
- Также можно создать таблицы в имеющихся базах данных на основе соединения с ними



Импорт из .lgd файлов

Файлы .lgd формата – проприетарный формат хранения данных Loginom ускоряющий запись и чтение данных на диск с использованием платформы Loginom.

Данные полей в файлах lgd формата автоматически определяют имена, метки, типы данных при импорте, а также вид этих данных в столбце.

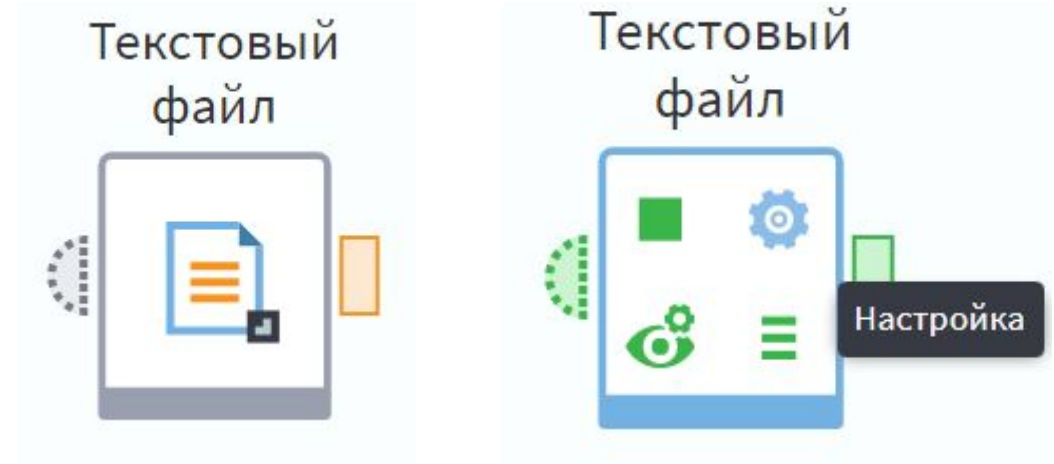
Файлы с данными архивируются и занимают меньше места на диске по сравнению с текстовыми файлами данных.



Импорт из текстовых файлов

С помощью компонента импорта «Текстовый файл» можно импортировать в Loginom данные из файлов в формате с условной разметкой, к ним относятся:

- «.txt» – текстовый файл,
- «.CSV» – колонки, разделенные запятой,
- «.tsv» – колонки, разделенные табуляцией,
- и т.д..



№	Название ФО	Сокр.	Площадь (км²)	Население	
01	Центральный	ЦФО	650205,00	40229958	
02	Северо-Западный	СЗФО	1686972,00	13867757	
03	Южный	ЮФО	447821,00	16635215	37
04	Северо-Кавказский	СКФО	170439,00	10	
05	Приволжский	ПФО	1036975,00	28683811	
06	Уральский	УрФО	1818497,00	12255848	
07	Сибирский	СФО	4361727,00	16645682	
08	Дальневосточный	ДФО / ДВФО	6952555,00	790	

Импорт из текстовых файлов

Настройка импорта данных из текстового файла заключается в выполнении двух последовательных действий:

- **Настройка импорта из файла**, в которую входит определение пути к файлу, указание кодовой страницы данных, факт наличия заголовка в первой строке и количество строк для пропуска.
- **Настройка формата импорта**, в котором определяются ключевые значения для разделителей столбцов, числовых значений, даты и времени и т.д.

Импорт из текстового файла

Хранилище файлов: Подключено

Имя файла:

Кодовая страница:

Заголовок в первой строке: Пропустить строк:

№	Название ФО	Сокр.	Площадь (км²)	Население (чел.) 2023 г.	Плотность населения (чел/км²)
01	Центральный	ЦФО	650205,00	40229958	61,87
02	Северо-Западный	СЗФО	1686972,00	13867757	8,22
03	Южный	ЮФО	447821,00	16635215	37,15
04	Северо-Кавказский	СКФО	170439,00	10203845	59,87
05	Приволжский	ПФО	1036975,00	28683811	27,66
06	Уральский	УрФО	1818497,00	12255848	6,74
07	Сибирский	СФО	4361727,00	16645682	3,82
08	Дальневосточный	ДФО / ДВФО	6952555,00	7902613	1,14

Настройка форматов импорта

Десятичный разделитель:

Разделитель столбцов:

Формат даты:

Считать последовательные разделители одним:

Разделитель даты:

Ограничитель строк:

Разделитель времени:

Пусто:

Истина:

Переменный тип:

Ложь:

Импорт из текстовых файлов

В настройке форматов импорта необходимо для полей определить имена и метки полей данных, а также типы и вид данных в столбцах.

Имя поля – используемое в Loginom имя переменной столбца, название которого должно соответствовать правилам именования переменной в C++.

Метка поля – название столбца, используемое для отображения при работе с визуализаторами данных.

При корректной настройке форматов импорта чистые данные импортируются без ошибок.

В случае наличия пропусков в данных столбец подсветится желтым. В случае полного несоответствия столбца формату – подсветится красным.

Поля	12 №	ab Название ФО	ab Сокр.	9.0 Площадь (км?)	12 Население (...)	9.0 Пл...
Имя	COL1	Nazvanie_FO_	Sokr_	Ploschad__km__	Naselenie__chel...	Plotn
Метка	№	Название ФО	Сокр.	Площадь (км?)	Население (чел...	Плотн
Тип данных	12 Целый	ab Строковый	ab Строковый	9.0 Вещественный	12 Целый	9.0 Ве
Вид данных	⊙ Непрерывный	⊙ Дискретный	⊙ Дискретный	⊙ Непрерывный	⊙ Непрерывный	⊙ Не
Использовать	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
1	1	Центральный	ЦФО	650 205,00	40 229 958	
2	2	Северо-Западный	СЗФО	1 686 972,00	13 867 757	
3	3	Южный	ЮФО	447 821,00	16 635 215	
4	4	Северо-Кавказс...	СКФО	170 439,00	10 203 845	
5	5	Приволжский	ПФО	1 036 975,00	28 683 811	
6	6	Уральский	УрФО	1 818 497,00	12 255 848	
7	7	Сибирский	СФО	4 361 727,00	16 645 682	
8	8	Дальневосточн...	ДФО / ДВФО	6 952 555,00	7 902 613	

Задание №3. Импорт из текстовых файлов

1. Откройте пакет с расширением «.lgr», созданный вами ранее;
2. Создайте новый модуль и измените его название на «Импорт из текстового файла»
3. Зайдите в сценарий и перенесите из панели компонентов на полотно сценария компонент «текстовый файл»
4. Настройте импорт данных из файла «data\GAZ\GAZ_n.csv»:
 - Кодировка UTF-8;
 - Разделитель табуляция;
 - Пусто – null.
5. Убедитесь в корректности импорта данных воспользовавшись предпросмотром выходного порта

Добыча газа

Импорт из текстового файла

Хранилище файлов: Локальное

Имя файла: Data/GAZ/GAZ_n.csv

Кодовая страница: UTF-8 (65001)

Заголовок в первой строке: Пропустить строк: 0

Настройка форматов импорта

Определить автоматически:

Десятичный разделитель: Запятая (,)

Разделитель столбцов: Символ табуляции

Формат даты: dd/ММ/yyyy

Считать последовательные разделители одним:

Разделитель даты: Точка (.)

Ограничитель строк: Двойная кавычка (")

Разделитель времени: По умолчанию

Пусто: null

Истина: True

Переменный тип:

Ложь: False

Добыча газа

Быстрый просмотр

Настроить порт...

Редактирование метки порта...

Задание №3. Импорт из текстовых файлов

Сколько полей (столбцов) получилось в выходной таблице в результате выполнения узла?

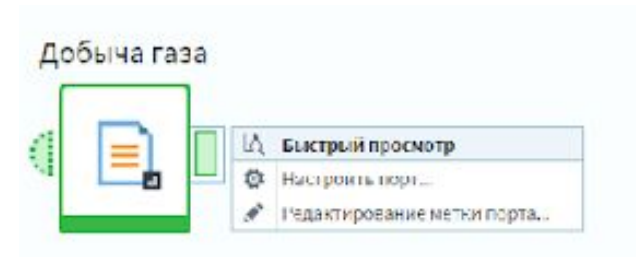
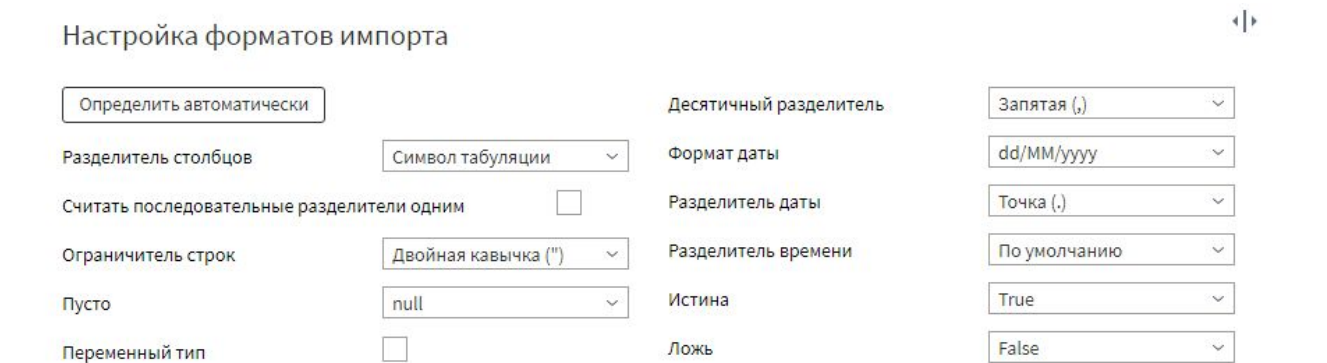
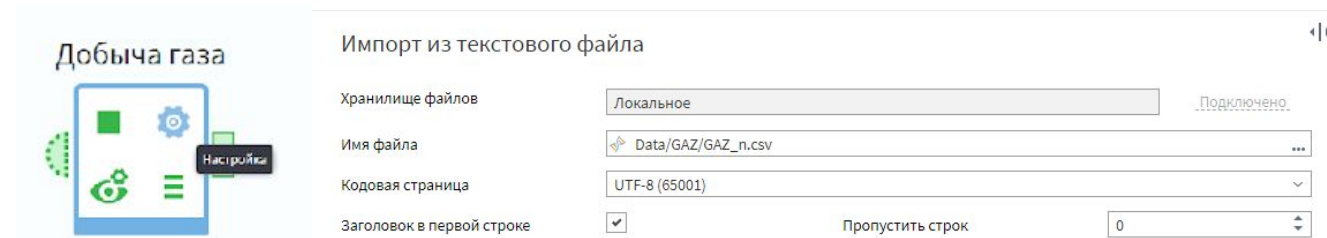
Какие типы данных имеет каждое поле?

- Перейдите в настройку форматов импорта, найдите поле ID и измените его тип данных на строковый. Сохраните изменения . Выполните узел.

Как изменился вид данных в поле ID?

Сколько строк данных получилось в результате импорта данных из внешнего файла GAZ_n.csv?

Каким образом мы настроили чтение данных из файла .csv с разделением табуляцией?



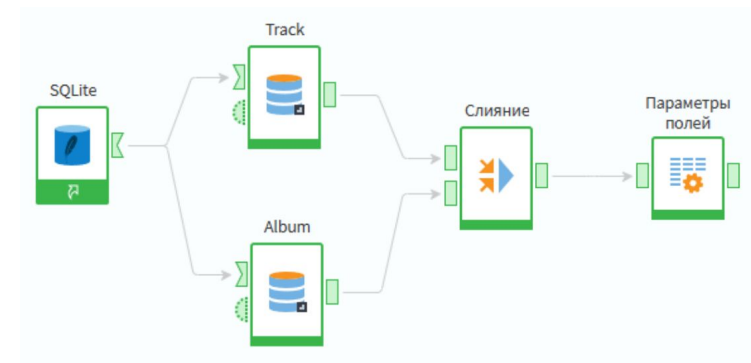
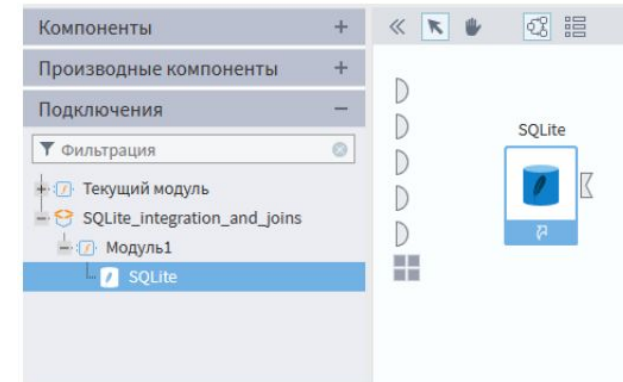
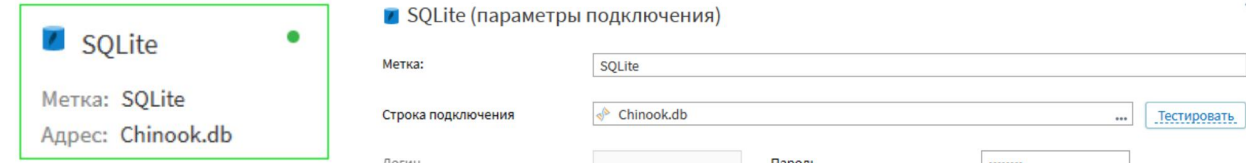
Импорт из баз данных

Loginom поддерживает импорт и экспорт данных из внешних РСУБД.

Подключение к базе данных настраивается посредством вкладки «**Подключения**» внутри модуля пакета Loginom.

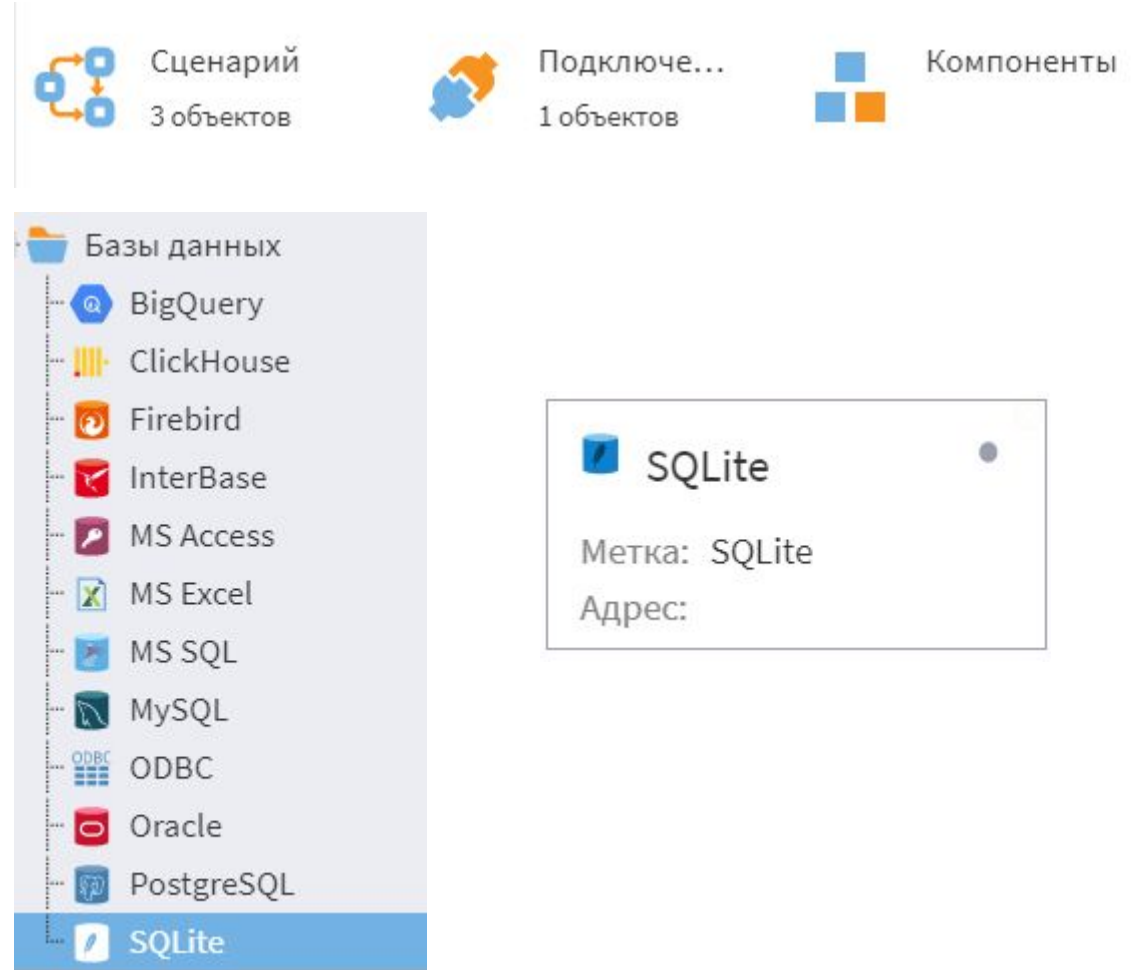
Настроив подключение, имеется возможность импортировать данные из и экспортировать данные в базу данных по подключению, выведенному на сценарий из раздела «Подключения», ниже раздела «Компоненты».

Подробнее в материалах «Loginom консолидация»



Задание №4. Импорт данных из РСУБД

- Модуль demo_2_3_импорт_бд
- Создайте новый модуль для данного задания перейдите в него
- Внутри модуля перейдите во вкладку «Подключения»
- Из папки «Базы данных» перенесите на полотно справа или ПКМ->Добавить подключение компонент SQLite
- Два раза нажмите на новое созданное подключение для его настройки



Задание №4. Импорт данных из РСУБД

- В настройках укажите в поле «строка подключения» путь к экземпляру базы данных Data/SQLite/Chinook.db
- Нажмите на кнопку «Тестировать» для проверки корректности соединения с SQLite
- Нажмите на кнопку «Подключить»
- Перейдите во вкладку сценарий
- Снизу раскройте «Подключения»
- В раскрытых подключениях найдите свой пакет и данный модуль, где и будет расположено ваше новое соединение с базой данных

The screenshot displays the configuration for a SQLite database connection. At the top, the title is "SQLite (параметры подключения)". Below it, the "Метка:" (Label) field contains "SQLite". The "Строка подключения" (Connection string) field contains "Data/SQLite/Chinook.db" and includes a "Тестировать..." (Test...) button. A summary card below shows the SQLite icon, the label "SQLite", and the connection string "Data/SQLite/Chi...". To the right is a "Подключить" (Connect) button with a plug icon.

Below the configuration, a tree view shows the following structure:

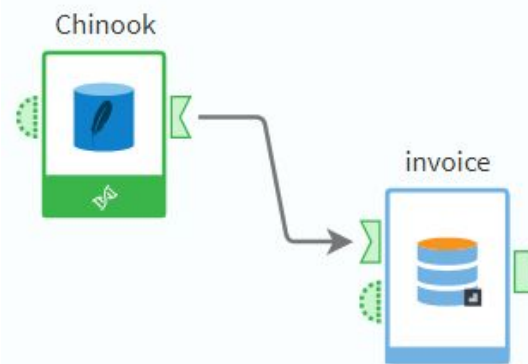
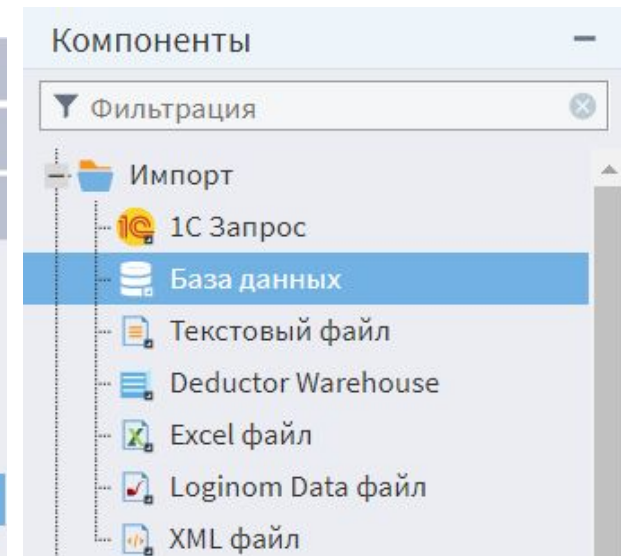
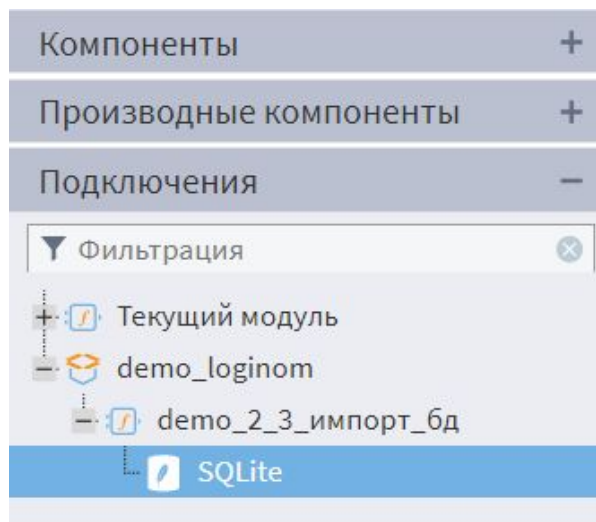
- Выполнение узла
- Подмодель
- Производные компоненты +
- Подключения +

On the right side, a panel titled "Подключения" (Connections) is shown with a filter field. The list includes:

- Текущий модуль
- demo_loginom
- demo_2_3_импорт_бд
- SQLite** (highlighted)

Задание №4. Импорт данных из РСУБД

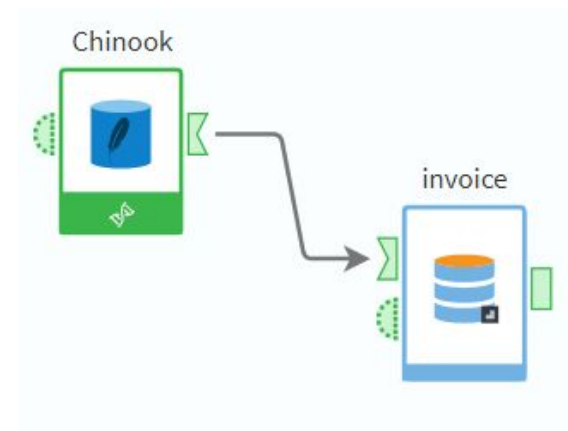
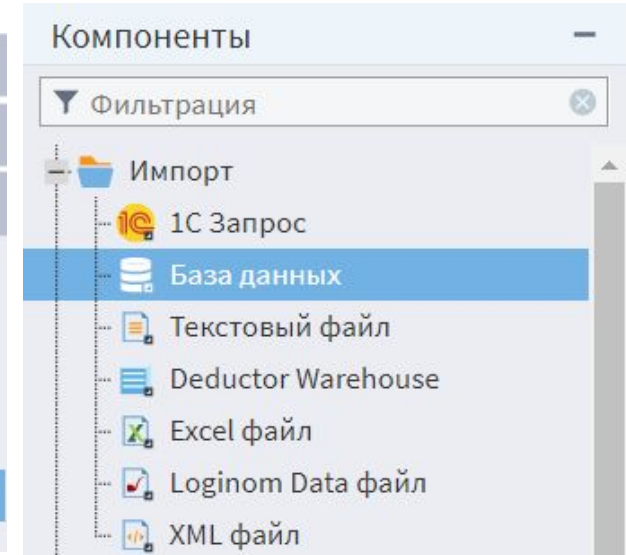
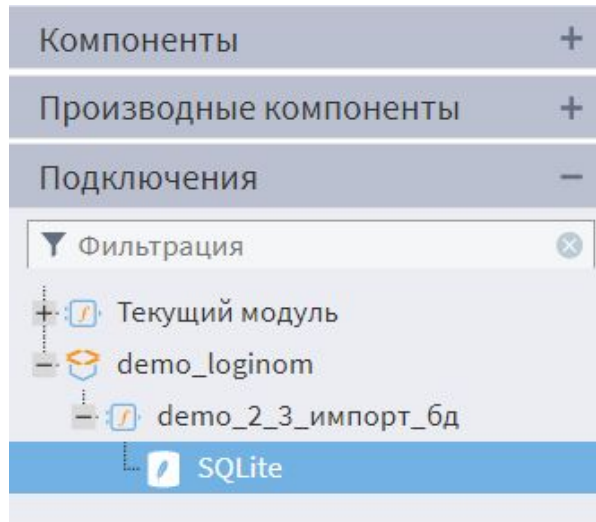
- Нажмите на него правой кнопкой мыши и выберите опцию «Добавить узел подключения в сценарий»
- Раскройте вкладку «Компоненты» и перенесите на сценарий компонент «База данных»
- Перетащите выходной порт подключения ко входному порту базы данных схожего вида. Активируйте соединение
- Настройте базу данных на импорт таблицы Invoice. Проверьте данные на выходном порте.



#	InvoiceId	CustomerId	InvoiceDate	BillingAddress	BillingCity	BillingState	BillingCountry	BillingPostalCode
1	1	2	01.01.2007, 00:00	Theodor-Heuss-Straße 34	Stuttgart	<null>	Germany	70174
2	2	4	02.01.2007, 00:00	Ullevøien 14	Oslo	<null>	Norway	0171
3	3	8	03.01.2007, 00:00	Grêtrystraat 63	Brussels	<null>	Belgium	1000
4	4	14	06.01.2007, 00:00	8220 111 ST NW	Edmonton	AB	Canada	T6G 2C1
5	5	23	11.01.2007, 00:00	69 Salem Street	Boston	MA	USA	2113
6	6	37	19.01.2007, 00:00	Berger Straße 10	Frankfurt	<null>	Germany	60316
7	7	38	01.02.2007, 00:00	Barbarossastraße 19	Berlin	<null>	Germany	10779
8	8	40	01.02.2007, 00:00	8, Rue Hanovre	Paris	<null>	France	75002
9	9	42	02.02.2007, 00:00	9, Place Louis Barthou	Bordeaux	<null>	France	33000
10	10	46	03.02.2007, 00:00	3 Chatham Street	Dublin	Dublin	Ireland	<null>
11	11	52	06.02.2007, 00:00	202 Howton Street	London	<null>	United Kingdom	N1 5LH
12	12	2	11.02.2007, 00:00	Theodor-Heuss-Straße 34	Stuttgart	<null>	Germany	70174
13	13	16	19.02.2007, 00:00	1600 Amphitheatre Parkway	Mountain View	CA	USA	94043-1
14	14	17	04.03.2007, 00:00	1 Microsoft Way	Redmond	WA	USA	98052-8
15	15	19	04.03.2007, 00:00	1 Infinite Loop	Cupertino	CA	USA	95014
16	16	21	05.03.2007, 00:00	801 W 4th Street	Reno	NV	USA	89503
17	17	25	06.03.2007, 00:00	319 N. Frances Street	Madison	WI	USA	53703
18	18	31	09.03.2007, 00:00	1944 Chain Lake Drive	Halifax	NS	Canada	B3K 3C2
19	19	40	14.03.2007, 00:00	8, Rue Hanovre	Paris	<null>	France	75002
20	20	54	22.03.2007, 00:00	110 Raeburn Pl	Edinburgh	<null>	United Kingdom	EH4 1JH
21	21	55	04.04.2007, 00:00	421 Bourke Street	Sidney	NSW	Australia	2010
22	22	57	04.04.2007, 00:00	Calle Lisa, 198	Santiago	<null>	Chile	<null>
23	59	05.04.2007, 00:00	3 Raj Bhavan Road	Bangalore	<null>	India	560001	

Задание №4. Импорт данных из РСУБД

- Существует ли возможность быстрого просмотра информации для выходного порта узла подключения к базе данных?
- Какое количество записей находится в таблице Invoice?
- Какую метку имеет поле первичного ключа для таблицы Invoice?



#	InvoiceId	CustomerId	InvoiceDate	BillingAddress	BillingCity	BillingState	BillingCountry	BillingPostalCode
1	1	2	01.01.2007, 00:00	Theodor-Heuss-Straße 34	Stuttgart	<null>	Germany	70174
2	2	4	02.01.2007, 00:00	Ullevøien 14	Oslo	<null>	Norway	0171
3	3	8	03.01.2007, 00:00	Grêtrystraat 63	Brussels	<null>	Belgium	1000
4	4	14	06.01.2007, 00:00	8210 111 ST NW	Edmonton	AB	Canada	T6G 2C1
5	5	23	11.01.2007, 00:00	69 Salem Street	Boston	MA	USA	2117
6	6	37	19.01.2007, 00:00	Berger Straße 10	Frankfurt	<null>	Germany	60316
7	7	38	01.02.2007, 00:00	Barbarossastraße 19	Berlin	<null>	Germany	10779
8	8	40	01.02.2007, 00:00	8, Rue Hanovre	Paris	<null>	France	75002
9	9	42	02.02.2007, 00:00	9, Place Louis Barthou	Bordeaux	<null>	France	33000
10	10	46	03.02.2007, 00:00	3 Chatham Street	Dublin	Dublin	Ireland	<null>
11	11	52	06.02.2007, 00:00	202 Howton Street	London	<null>	United Kingdom	N1 5LH
12	12	2	11.02.2007, 00:00	Theodor-Heuss-Straße 34	Stuttgart	<null>	Germany	70174
13	13	16	19.02.2007, 00:00	1600 Amphitheatre Parkway	Mountain View	CA	USA	94043-1
14	14	17	04.03.2007, 00:00	1 Microsoft Way	Redmond	WA	USA	98052-8
15	15	19	04.03.2007, 00:00	1 Infinite Loop	Cupertino	CA	USA	95014
16	16	21	05.03.2007, 00:00	801 W 4th Street	Reno	NV	USA	89503
17	17	25	06.03.2007, 00:00	319 N. Frances Street	Madison	WI	USA	53703
18	18	31	09.03.2007, 00:00	1944 Chain Lake Drive	Halifax	NS	Canada	B3S 1C2
19	19	40	14.03.2007, 00:00	8, Rue Hanovre	Paris	<null>	France	75002
20	20	54	22.03.2007, 00:00	110 Raeburn Pl	Edinburgh	<null>	United Kingdom	EH4 1JH
21	21	55	04.04.2007, 00:00	421 Bourke Street	Sydney	NSW	Australia	2010
22	22	57	04.04.2007, 00:00	Calle Lisa, 198	Santiago	<null>	Chile	<null>
23	59	59	05.04.2007, 00:00	3 Raj Bhavan Road	Bangalore	<null>	India	560001

Трансформация данных в Loginom

Фильтрация, сортировка, замена, данных

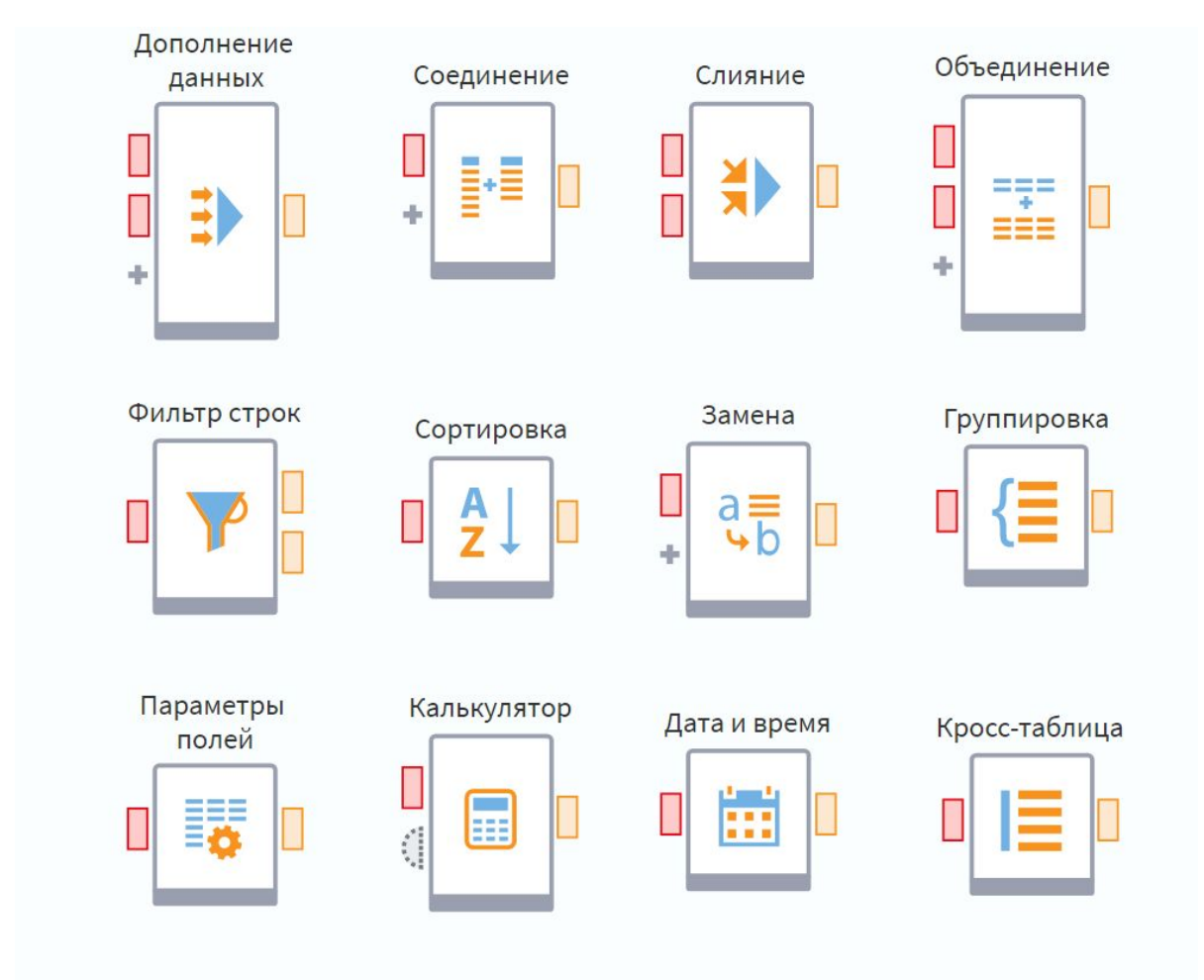
Преобразования данных, вычисление, работа с датами и временем

Обогащение данных (слияние, дополнение, объединение и соединение)

Агрегация данных, расчет агрегаций по группам

Трансформация данных

- Трансформация данных преследует цель преобразовать информацию в удобный для анализа вид
- Стандартно различают методы преобразования типов, вычислений, фильтрации, сортировки, агрегации, обогащения и транспонирования данных
- Реализации методов трансформации присутствуют в качестве стандартных обработчиков Loginom



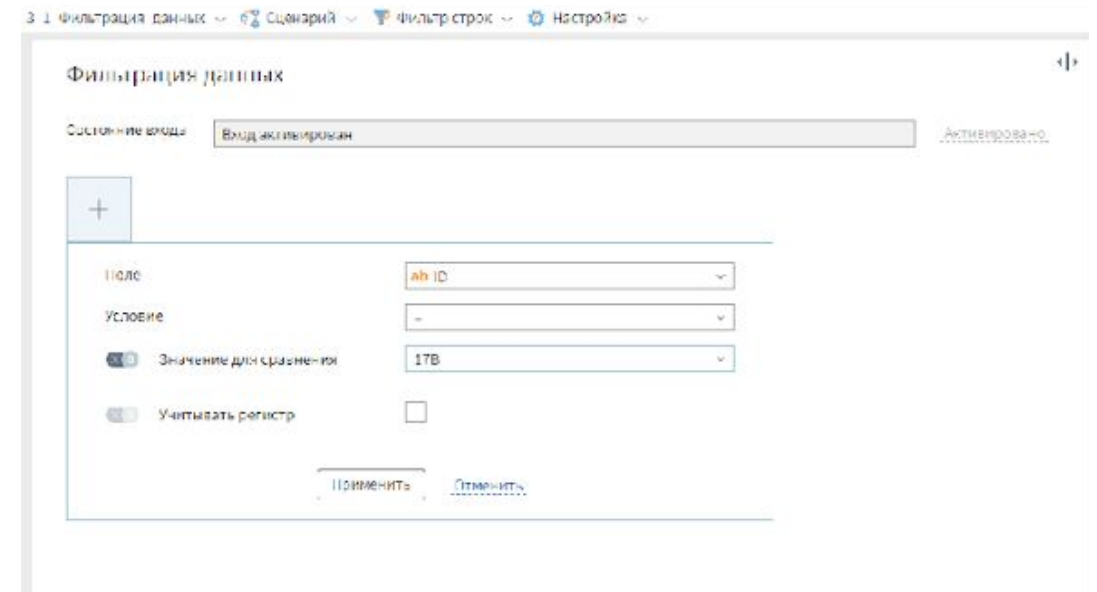
Компонент «Фильтр строк»

- Позволяет разделять выборку данных на две части, удовлетворяющие условию (верхний порт) и не удовлетворяющие условию (нижний порт)
- Условие задается алгеброй логики на основе логических выражений над полями
- В качестве значений для сравнения можно указывать порт управляющих переменных
- Ссылка на документацию <https://help.loginom.ru/userguide/processors/transformation/row-filter/index.html>

Фильтр строк

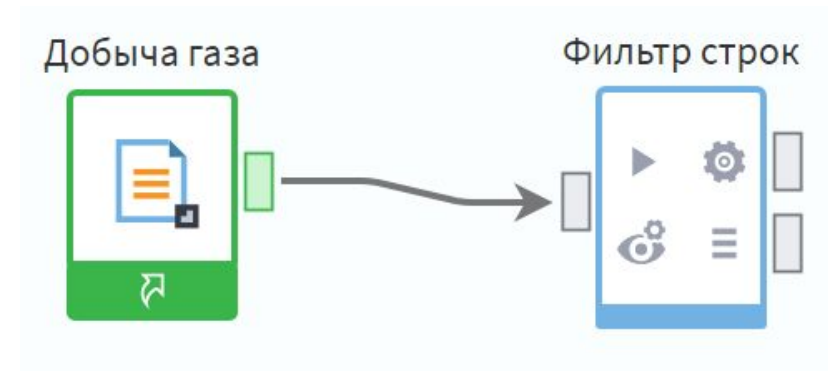


Фильтр строк



Задание №5. Фильтрация пустых полей

- Модуль
demo_3_1_Фильтрация_данных
- Для импортированных данных из задания №3 проведите фильтрацию импортированных данных на основе фильтрации записей с пустыми полями
- Соедините выход порта узла **импорта данных из текстового файла** со входом в узел **фильтр строк**
- Настройте узел фильтр строк для фильтрации данных



Фильтрация данных

Состояние входа Не активировано

[Активировать](#)

ab ID не пустое × .И.

31 Дата замера не пустое × .И.

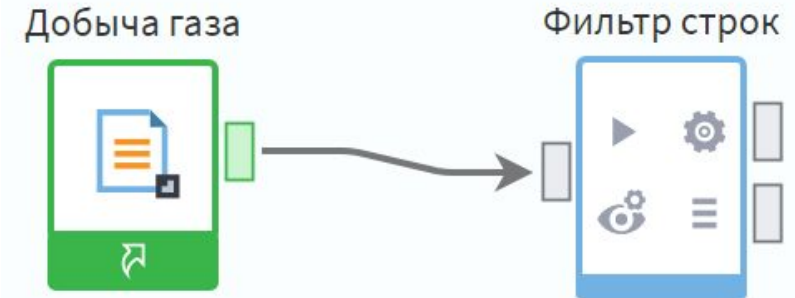
9.0 Ртр МПа не пустое ×

.И. 9.0 Туст °С не пустое × .И. 9.0 газ м3/сут не пустое × .И. 9.0 конд-т м3/сут не пустое ×

.И. 9.0 вода м3/сут не пустое × .И. ab Куст не пустое × .И. ab Группа не пустое × +

Задание №5. Фильтрация пустых полей

- При настройке фильтрации данных нажмите на «+» и выберите первое поле после «**Номер строки**»
- Задайте для него условие «**не пустой**»
- Нажмите **применить** и повторите действие для всех остальных столбцов
- Нажмите **далее** и настройте название метки узла как «полные записи» и нажмите **выполнить**
- Для нижнего выходного порта с помощью быстрого просмотра скажите, сколько в нем находится записей



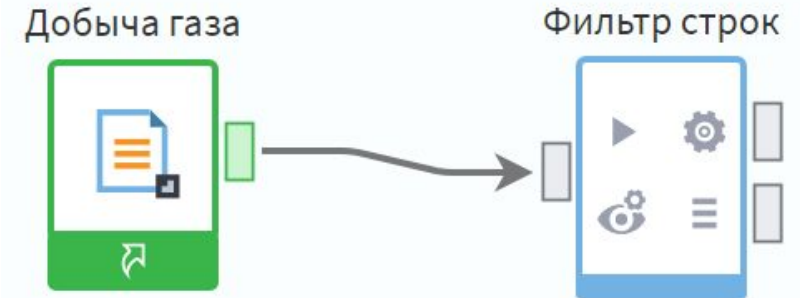
Фильтрация данных

Состояние входа [Активировать...](#)

ab ID не пустое	и	31 Дата замера не пустое	и	9.0 Ртр МПа не пустое		
и	9.0 Туст °С не пустое	и	9.0 газ м3/сут не пустое	и	9.0 конд-т м3/сут не пустое	
и	9.0 вода м3/сут не пустое	и	ab Куст не пустое	и	ab Группа не пустое	+

Задание №5. Фильтрация пустых полей

- Сколько строк данных удовлетворяют условию полных записей?
- Сколько строк данных не удовлетворяют условию, т.е. содержат пропуски в данных?
- Какой процент от общего количества строк занимают неполные записи?
- На основе предпросмотра убедитесь, присутствуют ли в выборке неполных записей те, которые имеют пропуски в нескольких полях?

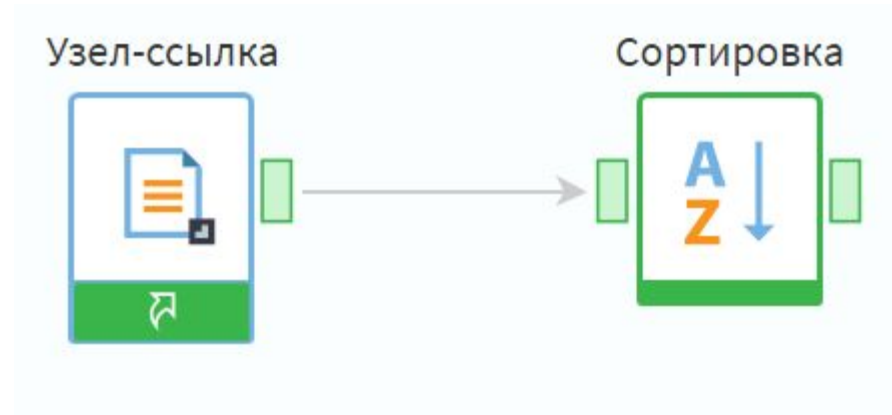


Фильтрация данных

Состояние входа: [Активировать](#)

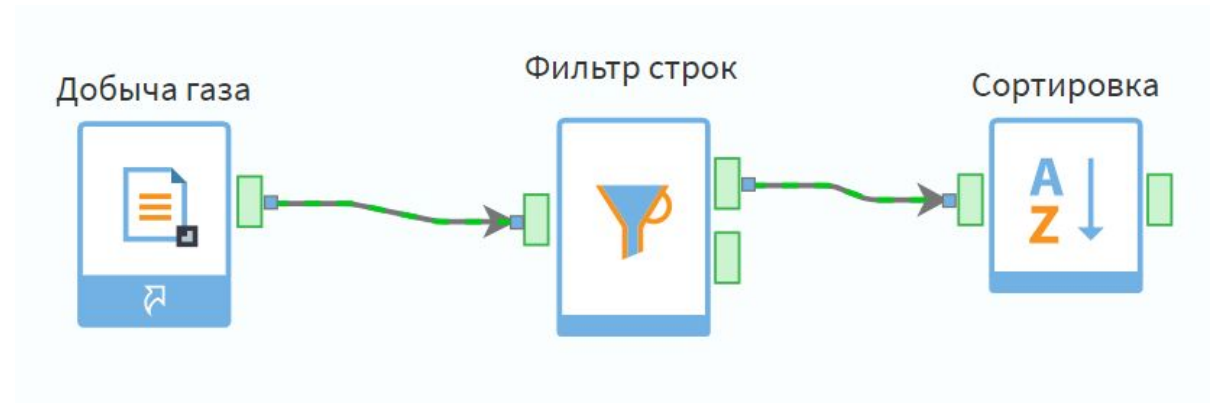
Компонент «Сортировка»

- Упорядочивает записи по значению выбранных полей
- Упорядочивает строки всей таблицы по значениям выбранных полей
- Можно выбирать несколько полей для сортировки в лексикографическом порядке
- Порядок сортировки таблицы можно изменять
- Ссылка на документацию <https://help.loginom.ru/userguide/processors/transformation/sorting.html>



Задание №6. Сортировка записей

- Модуль demo_3_2_Сортировка_данных
- Для данных, удовлетворяющих условию фильтрации из предыдущего задания произведите сортировку записей по дате замера (по возрастанию) и скорости добычи газа (по убыванию)
- Перенесите с помощью мыши верхний выходной порт фильтра строк и входной порт сортировки
- Перейдите в настройку узла сортировки



Задание №6. Сортировка записей

- Перенесите в настройках сортировки поля «Дата замера» и «газ м3/сут» в **поля сортировки**
- Переключателем настройте порядок сортировки как указано ранее
- Сохраните настройки узла и выполните его в сценарии
- Убедитесь в корректности сортировки данных на основе быстрого просмотра выходного порта сортировки

Сортировка

Фильтрация
Доступные поля
9.0 Ртр МПа
9.0 Туст °С
9.0 конд-т м3/сут
9.0 вода м3/сут
ab Куст
ab Группа
ab ID



Переместить вверх	Переместить вниз		
Поля сортировки	Порядок	Регистр	
31 Дата замера	↕		
9.0 газ м3/сут	↕		

Задание №6. Сортировка записей

- Какая дата замера оказывается самой ранней в таблице добычи газа?
- Какой ID скважины соответствует самой ранней записи замеров скорости добычи?
- Изменилось ли количество записей в выходной таблице по итогам сортировки таблицы по сравнению с количеством записей на входе?
- Какая дата замера является самой последней?

Добыча газа → Фильтр строк → Сортировка

Сортировка

Фильтрация

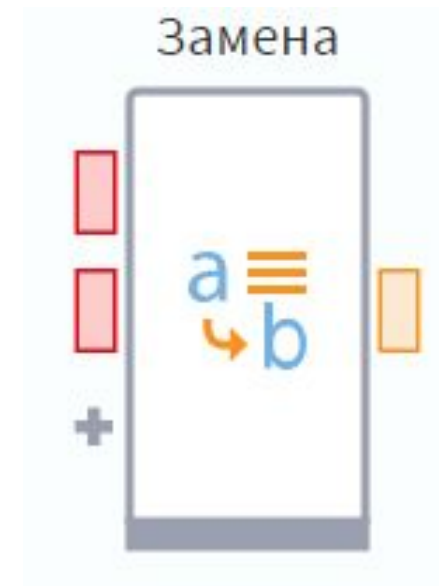
Доступные поля

- 9.0 Ртр МПа
- 9.0 Туст °С
- 9.0 конд-т м3/сут
- 9.0 вода м3/сут
- ab Куст
- ab Группа
- ab ID

Поля сортировки	Порядок	Регистр	
31 Дата замера	↔		
9.0 газ м3/сут	↔		

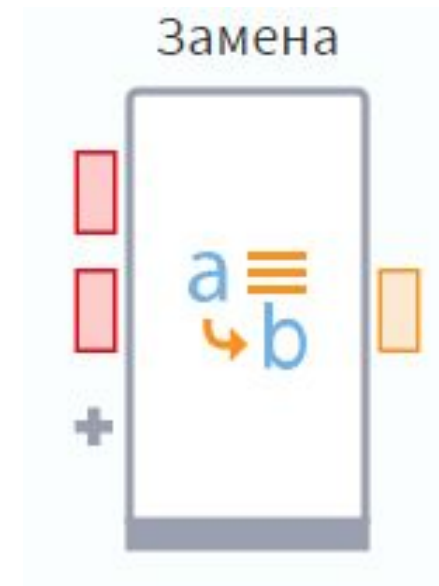
Компонент «Замена»

- **Назначение:** используется для замены данных в исходном наборе с помощью таблиц замен
- **Таблица замен** – справочник, содержащий пары значений: те, что необходимо заменить и соответствующие им новые значения.
- Таблицы замен или подготавливаются заранее и подаются на вход узла, или создаются вручную внутри узла.



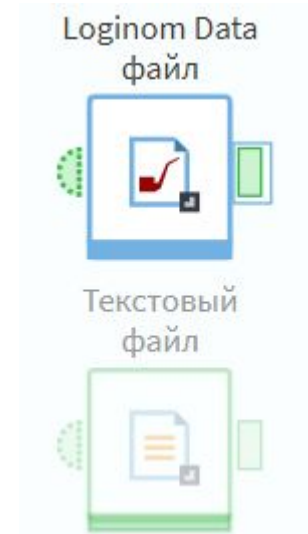
Компонент «Замена»

- Замена значений другими значениями с совпадающими типами данных
- Можно заменять значение по заданной таблице замен
- Можно заменять целые числа на строковое выражение диапазона с заданной точностью, шириной окна (дискретизация данных)
- Ссылка на документацию <https://help.loginom.ru/userguide/processors/transformation/substitution/index.html>



Задание №7. Использование таблицы замен

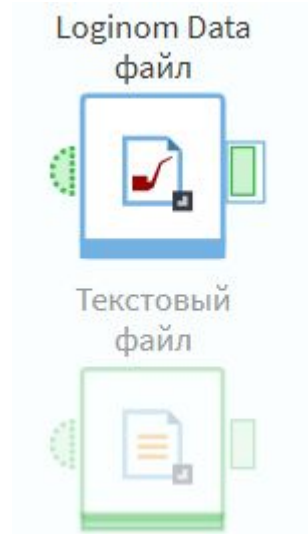
- Модуль demo_3_3_замена
- Необходимо заменить краткие названия месяцев на полные.
- В поле построения сценария добавим 2 компонента импорта: Loginom data файл (Исходники) и Текстовый файл (Замена).
- Подать на вход набор данных из файлов `Data/Demo/data_months.lgd` и `Data/Demo/ months.txt` соответственно.
- Таблица исходных данных `Data/Demo/ data_months.lgd` содержит Идентификатор – Краткое название месяца.



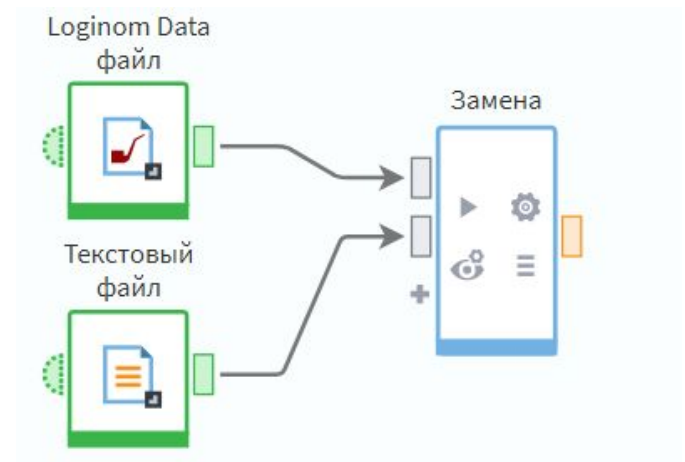
ab Идентификат...	ab Месяц регистрации
01	Окт
02	Июнь
03	Ноябрь
04	Авг
05	Янв
06	Июль
07	Сен
08	Дек
09	Февраль
10	Март
11	Октябрь
12	Апр
13	Июн
14	Май
15	Апрель
16	Июл
17	Декабрь
18	Фев
19	Мар
20	Сентябрь

Задание №7. Использование таблицы замен

- Таблица замен
Data/Demo/months.txt содержит пары значений: **Краткое название месяца – Полное название месяца.**
- Добавим в область построения сценария компонент **Замена.**
- Подключим наборы данных к соответствующим портам узла и откроем настройку узла **Замена.**



ab Краткое название	ab Полное назван...
Янв	Январь
Фев	Февраль
Мар	Март
Апр	Апрель
Май	Май
Июн	Июнь
Июл	Июль
Авг	Август
Сен	Сентябрь
Окт	Октябрь
Ноя	Ноябрь
Дек	Декабрь



Задание №7. Использование таблицы замен

Для полей исходных данных можно настроить **Способ замены**:

- **Не заменять.**
- **Ввод вручную** – формируем внутреннюю таблицу замен.
- **Таблица замен** – используем внешнюю таблицу замен.

Выберем для поля **Месяц регистрации** способ **Таблица замен**

Замена

Поля	Способ замены
ab Идентификатор	Не заменять
ab Месяц регистрации	Таблица замен
	Не заменять
	Ввод вручную
	Таблица замен (Таблица замен 1)

№	Название поля	Назначен
1	ab Краткое название	Не используемое
2	ab Полное название	Не используемое

Задание №7. Использование таблицы замен

Назначение поля *Краткое название* установим **Значение**, поля *Полное название* – **Замена**.

Если в исходных данных встретятся значения, для которых не будет совпадений в таблице замен, можно выбрать вариант их замены во вкладке **Заменять остальные**.

На следующем шаге настройки, в окне *Настройка соответствия между столбцами* по умолчанию используется соответствие столбцов Замена: столбец с исходными значениями заменится на столбец с новыми значениями.

Замена

Поля	Способ замены
ab Идентификатор	Не заменять
ab Месяц регистрации	Таблица замен 1

№	Название поля	Назначение
1	ab Краткое название	<i>i</i> Не используемое
2	ab Полное название	<i>i</i> Не используемое

*A** Значение
+B Замена
i Информационное

Настройка соответствия между столбцами

Таблица Связи

Связать автоматически | Удалить все связи | Упорядочить связи | Отображать потерянные связи

Входные	Выходные	Назначение
ab Идентификатор	ab Идентификатор	<i>i</i> Не задано
ab Месяц регистрации Замена	ab Месяц регистрации Замена	<i>i</i> Не задано
0/1 Месяц регистрации Заменен	0/1 Месяц регистрации Заменен	<i>i</i> Не задано

Задание №7. Использование таблицы замен

Сохраним настройки. Выполним узел. Посмотрим что получилось в результате.

Всем ли записям в выходной таблице данных соответствует значение true в поле **Месяц регистрации заменен**?

Есть ли во входной таблице данных такие записи, заменять которые было не нужно?

Замена

Поля	Способ замены
ab Идентификатор	Не заменять
ab Месяц регистрации	Таблица замен 1

№	Название поля	Назначение
1	ab Краткое название	Не используемое
2	ab Полное название	Не используемое

- А+ Значение
- +B Замена
- i Информационное

Настройка соответствия между столбцами

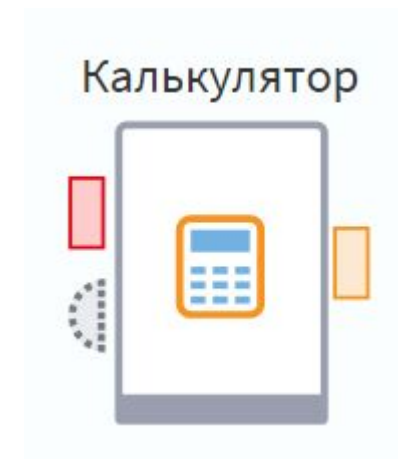
Таблица Связи

Связать автоматически | Удалить все связи | Упорядочить связи | Отображать потерянные связи

Входные	Выходные	Назначение
ab Идентификатор	ab Идентификатор	Не задано
ab Месяц регистрации Замена	ab Месяц регистрации Замена	Не задано
0/1 Месяц регистрации Заменен	0/1 Месяц регистрации Заменен	Не задано

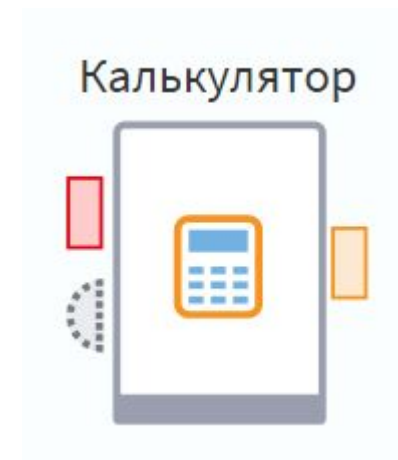
Компонент «Калькулятор»

- Вычисляет новые поля данных на основе существующих полей и входных переменных на основе поэлементных математических операций между столбцами
- Оперирует стандартными математическими операциями и встроенными функциями, позволяющими рассчитывать значения или работать с подстроками на основе регулярных выражений



Компонент «Калькулятор»

- Для нового поля значение в каждой строке данных вычисляется на основе соответствующих ему значений из других полей, определенных на уровне выражения, записанного аналитиком в калькуляторе
- Значения переменных из порта «переменные» для каждой строки данных остаются постоянными
- Выходная таблица имеет большее число полей на количество вычисляемых полей



Компонент «Калькулятор»

Настройка калькулятора имеет 4 рабочие области:

- Верхняя левая – вычисляемые поля, можно создать новые, задать им тип, имя и метку
- Нижняя левая – входные данные и переменные на портах
- Верхняя правая – поле выражения для выбранного вычисляемого поля
- Нижняя правая – справка встроенных функций

Калькулятор

The screenshot shows the LogiCAD calculator interface with the following components:

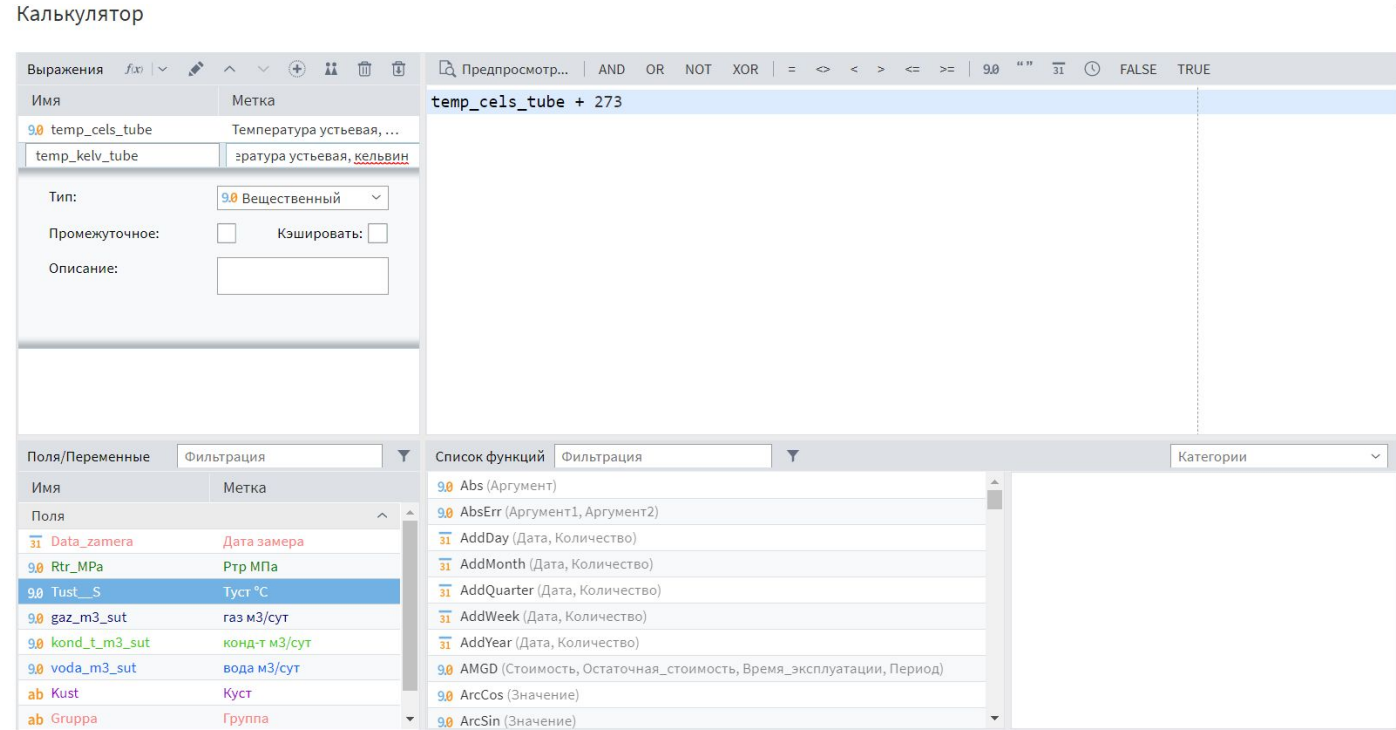
- Top Left Panel (Fields/Variables):** A table with columns 'Имя' (Name) and 'Метка' (Label). It contains one entry: 'temp_cels_tube' with the label 'Температура устьева...'. Below this is a section for 'Поля/Переменные' (Fields/Variables) with a 'Фильтрация' (Filter) dropdown. It lists several variables: 'Data_zamera' (Дата замера), 'Rtr_MPa' (Ртр МПа), 'Tust__S' (Туст °C), 'gaz_m3_sut' (газ м3/сут), 'kond_t_m3_sut' (канд-т м3/сут), 'voda_m3_sut' (вода м3/сут), 'Kust' (Куст), and 'Gruppa' (Группа).
- Top Right Panel (Expression Editor):** A text area for entering expressions. The current expression is `IFF(Tust__S > 100, Tust__S / 10, Tust__S)`. Above the text area is a toolbar with logical operators (AND, OR, NOT, XOR), comparison operators (<, >, <=, >=), and constants (9.0, "", TRUE, FALSE).
- Bottom Left Panel (Fields/Variables):** A table with columns 'Имя' (Name) and 'Метка' (Label). It lists several variables: 'Data_zamera' (Дата замера), 'Rtr_MPa' (Ртр МПа), 'Tust__S' (Туст °C), 'gaz_m3_sut' (газ м3/сут), 'kond_t_m3_sut' (канд-т м3/сут), 'voda_m3_sut' (вода м3/сут), 'Kust' (Куст), and 'Gruppa' (Группа).
- Bottom Right Panel (Function List):** A list of functions with a 'Фильтрация' (Filter) dropdown and a 'Категории' (Categories) dropdown. The functions listed include: 'Abs (Аргумент)', 'AbsErr (Аргумент1, Аргумент2)', 'AddDay (Дата, Количество)', 'AddMonth (Дата, Количество)', 'AddQuarter (Дата, Количество)', 'AddWeek (Дата, Количество)', 'AddYear (Дата, Количество)', 'AMGD (Стоимость, Остаточная_стоимость, Время_эксплуатации, Период)', 'ArcCos (Значение)', and 'ArcSin (Значение)'.

Компонент «Калькулятор»

При создании нового вычисляемого поля для него создается свое пустое поле для вычислений.

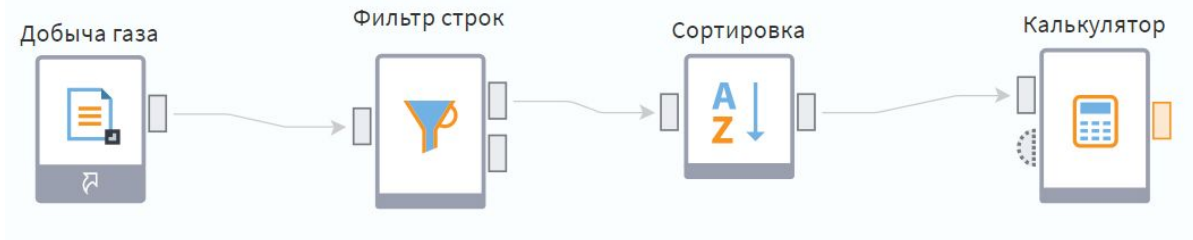
Выражение пишется в одну строку с применением числовых констант, арифметических операций (+, -, *, /, ...), входных полей и переменных, а также встроенных функций

Если поле вычисляется как промежуточное, оно используется только калькулятором и не подается в выходную таблицу данных



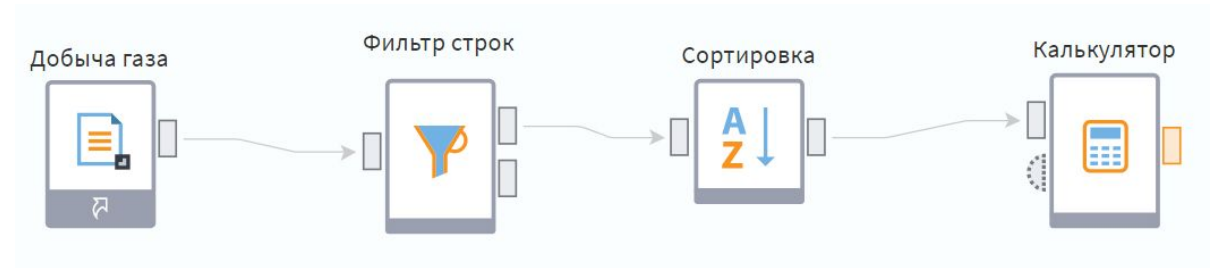
Задание №8. Вычисление новых полей

- Модуль demo_3_4_калькулятор
- В данном задании для упорядоченной выборки необходимо подкорректировать числовые значения некоторых полей
- Для задания с сортировкой данных продолжим построение тестового сценария обработки данных добычи газа
- Добавим в область построения сценария компонент калькулятор



Задание №8. Вычисление новых полей

- Соединим входной порт калькулятора с выходным портом сортировки из задания №6. Перейдем к настройке узла «Калькулятор»
- В поле выражения с помощью кнопки *добавить выражение* добавим новое вычисляемое поле с именем **Temp_tube**, и меткой *Температура устьевая, цельсий*
- Запишем для его вычисления следующее условное выражение, которое в поле **Tust__S** проверяет условие и в случае успешного прохождения производит деление значения в записи на 10
- При написании выражений можно просто два раза кликнуть ЛКМ на нужное входное поле, чтобы оно переместилось в выражение



Имя	Метка
Temp_tube	Температура устьевая, цел

Тип: 9.0 Вещественный

Промежуточное: Кэшировать:

Описание: Скорректированное поле устьевой температуры в скважине

```
Предпросмотр... | AND OR NOT XOR | = <> <
```

```
IFF(Tust__S > 100, Tust__S / 10, Tust__S)
```

Задание №8. Вычисление новых полей

- Также создадим поля **Pressure_tube** и **Volume_tube** с соответствующими выражениями для их вычисления
- Данные выражения основаны на анализе предметной области и обзора данных с точки зрения анализа качества выборки
- Сохраните настройки узла и выполните узел
- С помощью быстрого просмотра выходного порта просмотрите какие поля получились в итоге

Имя	Метка
9.0 Temp_tube	Температура устьева, ...
Pressure_tube	Давление устьевое, МПа

Тип: 9.0 Вещественный

Промежуточное: Кэшировать:

Описание: Скорректированные значения устьевого давления с поправкой на опечатки при сборе данных

Имя	Метка
9.0 Temp_tube	Температура устьева, ...
9.0 Pressure_tube	Давление устьевое, МПа
Volume_tube	Скорость добычи газа, м3.

Тип: 9.0 Вещественный

Промежуточное: Кэшировать:

Описание: Скорректированная скорость добычи газа из скважин

```
Предпросмотр... | AND OR NOT XOR | = <
```

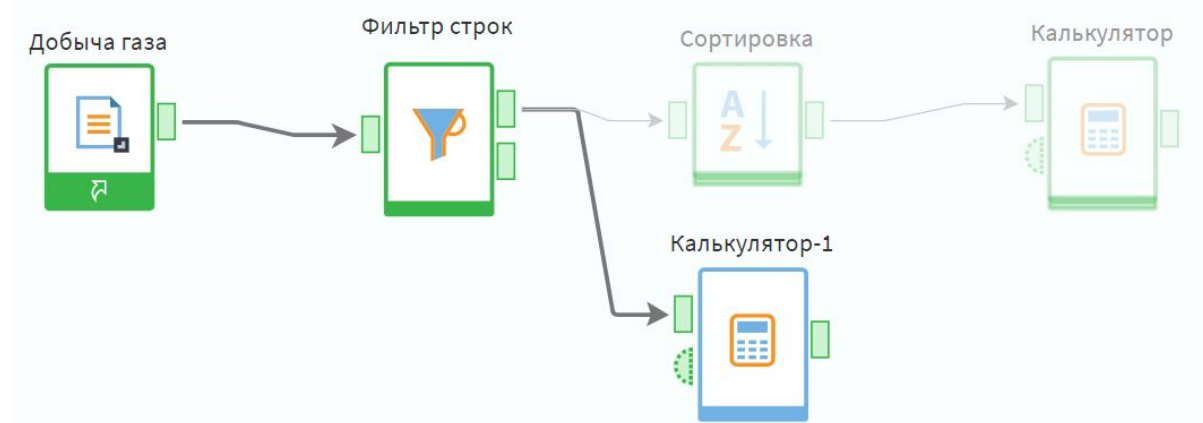
```
IFF(Rtr_MPa > 20, Rtr_MPa / 10, Rtr_MPa)
```

```
Предпросмотр... | AND OR NOT XOR | = <> < > <= >
```

```
IFF(gaz_m3_sut < 1000, gaz_m3_sut * 1000, gaz_m3_sut)
```


Задание №8. Вычисление новых полей

- Сколько новых полей появилось в выходной таблице данных после вычисления выражений в узле калькулятор?
- Скопируйте настроенный узел калькулятора, вставьте его в сценарий и подайте на его вход данные из верхнего выходного порта фильтра строк.
- Выполните данную копию узла.
- Просмотрите выходную таблицу данных у копии калькулятора.
- Сравните поля: **Скорость добычи газа, м3.сут** и **газ м3/сут**. В какой строчке произошли изменения?
- Также сравните другие вычисленные поля с их базовыми

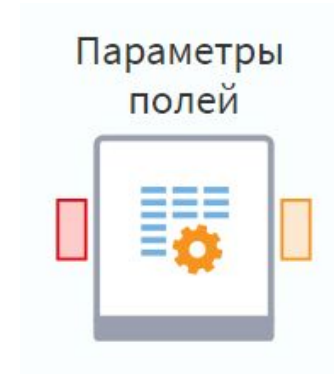


9.0 Скорость добычи газа, м3.сут	9.0 газ м3/...	9.0 Давление устьевое, МПа	9.0 Ртр М...
361 939,00	361 939,00	17,50	17,50
349 748,00	349 748,00	17,60	176,00
310 185,00	310 185,00	17,30	17,30
265 176,00	265 176,00	17,50	17,50
263 621,00	263 621,00	17,40	17,40
446 708,00	446 708,00	16,00	16,00
438 621,00	438 621,00	15,70	15,70
372 252,00	372 252,00	15,90	15,90
383 300,00	383,30	15,80	158,00
499 651,00	499 651,00	14,80	14,80
483 540,00	483 540,00	14,70	14,70

Компонент «Параметры полей»

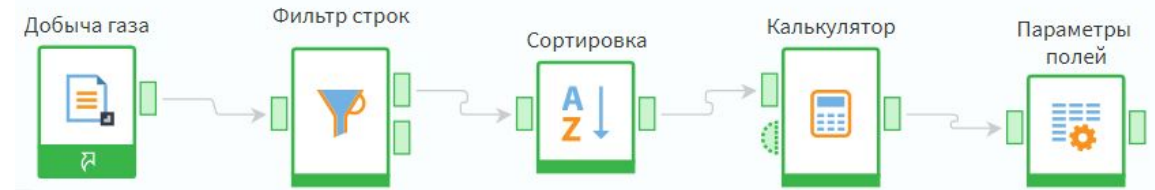
- Позволяет изменять поля, их имена, метки, типы и виды данных
- Позволяет исключать поля из выборки данных и передает изменения в выходную выборку
- Преобразования данных должны быть валидными, без потери информации при преобразовании
- Пример корректной цепочки преобразований:

Вещественные -> Целые ->
Строковые



Задание №9. Параметры полей

- Модуль demo_3_5_параметры_полей
- Для предыдущего задания добавьте в сценарий обработки данных компонент параметры полей, чтобы явно преобразовать структуру данных после проведения расчетов в калькуляторе
- Подайте выходной порт калькулятора на входной порт «Параметры полей»



Метка	Имя	Вид данных	Назначение	Кэши...	Исключить
9.0	Температура устьева, цельсий	Непрерывный	Не задано	Отклю...	<input type="checkbox"/>
9.0	Давление устьево, МПа	Непрерывный	Не задано	Отклю...	<input type="checkbox"/>
9.0	Скорость добычи газа, м3.сут	Непрерывный	Не задано	Отклю...	<input type="checkbox"/>
9.0	газ м3/сут	Непрерывный	Не задано	Отклю...	<input checked="" type="checkbox"/>
9.0	Ртр МПа	Непрерывный	Не задано	Отклю...	<input checked="" type="checkbox"/>
9.0	Туст °С	Непрерывный	Не задано	Отклю...	<input checked="" type="checkbox"/>
31	Дата замера	Непрерывный	Не задано	Отклю...	<input type="checkbox"/>
ab	ID	Дискретный	Не задано	Отклю...	<input type="checkbox"/>
9.0	Скорость добычи конденсата...	Непрерывный	Не задано	Откл...	<input type="checkbox"/>
9.0	Скорость добычи воды, м3.сут	Непрерывный	Не задано	Откл...	<input type="checkbox"/>
ab	Куст	Дискретный	Не задано	Отклю...	<input type="checkbox"/>
ab	Группа	Дискретный	Не задано	Отклю...	<input type="checkbox"/>

Задание №9. Параметры полей

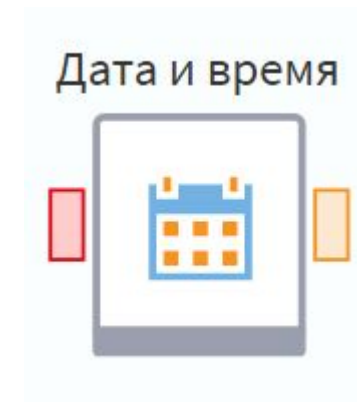
- Настройте параметры полей как показано на слайде, исключив поля, на основе которых проводилось вычисление новых, а также изменив метку для полей **конд-т** и **вода** на новые метки
- Выполните узел и выполните быстрый просмотр выходного порта из узла «Параметры полей»



Метка	Имя	Вид данных	Назначение	Кэши...	Исключить
9.0	Температура устьевая, цельсий	Непрерывный	Не задано	Отклю...	<input type="checkbox"/>
9.0	Давление устьевое, МПа	Непрерывный	Не задано	Отклю...	<input type="checkbox"/>
9.0	Скорость добычи газа, м3.сут	Непрерывный	Не задано	Отклю...	<input type="checkbox"/>
9.0	газ м3/сут	Непрерывный	Не задано	Отклю...	<input checked="" type="checkbox"/>
9.0	Ртр МПа	Непрерывный	Не задано	Отклю...	<input checked="" type="checkbox"/>
9.0	Туст °С	Непрерывный	Не задано	Отклю...	<input checked="" type="checkbox"/>
31	Дата замера	Непрерывный	Не задано	Отклю...	<input type="checkbox"/>
ab	ID	Дискретный	Не задано	Отклю...	<input type="checkbox"/>
9.0	Скорость добычи конденсата...	Непрерывный	Не задано	Откл...	<input type="checkbox"/>
9.0	Скорость добычи воды, м3.сут	Непрерывный	Не задано	Откл...	<input type="checkbox"/>
ab	Куст	Дискретный	Не задано	Отклю...	<input type="checkbox"/>
ab	Группа	Дискретный	Не задано	Отклю...	<input type="checkbox"/>

Компонент «Дата и время»

- Обработчик Дата и время производит трансформацию данных формата Дата/Время.
- Суть такого преобразования заключается в том, что на основе столбца с информацией о дате/времени формируется один или несколько дополнительных столбцов, в которых выделяется определенная информация о дате/времени.
- Вычисляет новые столбцы данных на основе преобразований даты и времени. Позволяет вычислять:
 - даты начала и конца периодов,
 - календарные сущности (день, неделя, месяц, квартал, год, ...)



Пример преобразования «Дата и время»

На основе исходного столбца формируются остальные. В каждом сформированном столбце выделена определенная информация из даты, например, день месяца, минута.

Дата (исходный столбец)	(Год + Месяц)	Дата (Год + День)	Дата (День месяца)	Дата (Минута)	Дата (Свой формат)
03.04.2012 00:04	01.04.2012	03.04.2012	3	4	произвольный текст - 03.04.12
17.04.2012 19:39	01.04.2012	17.04.2012	17	39	произвольный текст - 17.04.12
09.05.2012 19:42	01.05.2012	09.05.2012	9	42	произвольный текст - 09.05.12
16.05.2012 12:13	01.05.2012	16.05.2012	16	13	произвольный текст - 16.05.12
28.05.2012 20:35	01.05.2012	28.05.2012	28	35	произвольный текст - 28.05.12

Настройка преобразования «Дата и время»

В области Поле отображаются поля входного набора данных с типом **Дата/Время**. Эти поля можно отсортировать по количеству выбранных преобразований. Для этого нужно нажать на заголовок столбца с количеством преобразований. **Сортировка** возможна по возрастанию и по убыванию.

В области **Разбиение** при помощи флажков чекбокса необходимо выбрать формат и тип данных для преобразования.

Возможны варианты преобразования в следующие типы данных:

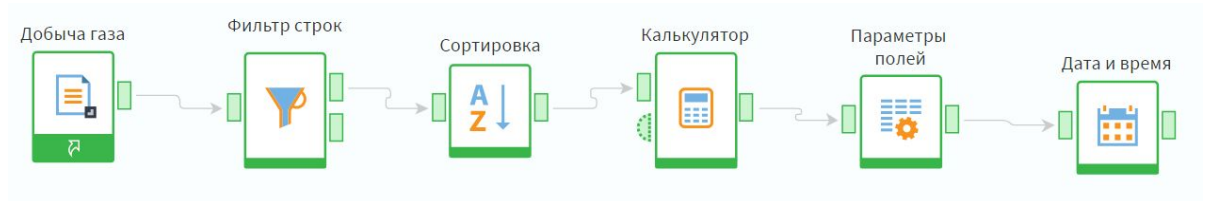
- Дата/Время;
- Число;
- Строка.

При преобразовании в тип данных Дата/Время для ряда форматов имеется возможность выбрать начало и/или конец периода.

Разбиение	31 Дата начала	31 Дата конца	12 Число	ab Строка
Обычный				
Год + Квартал	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
Год + Месяц	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
Год + Неделя	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
Год + День	<input type="checkbox"/>			<input type="checkbox"/>
Год	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Квартал			<input type="checkbox"/>	<input type="checkbox"/>
Месяц			<input type="checkbox"/>	<input type="checkbox"/>
Неделя			<input type="checkbox"/>	<input type="checkbox"/>
День года			<input type="checkbox"/>	<input type="checkbox"/>
День квартала			<input type="checkbox"/>	<input type="checkbox"/>
День месяца			<input type="checkbox"/>	<input type="checkbox"/>
День недели			<input type="checkbox"/>	<input type="checkbox"/>
Часы			<input type="checkbox"/>	<input type="checkbox"/>
Минуты			<input type="checkbox"/>	<input type="checkbox"/>
Секунды			<input type="checkbox"/>	<input type="checkbox"/>
Миллисекунды			<input type="checkbox"/>	<input type="checkbox"/>
Дата	<input type="checkbox"/>			<input type="checkbox"/>
Время	<input type="checkbox"/>			<input type="checkbox"/>
Свой формат				<input type="checkbox"/>

Задание №10. Выделение начала месяца

- Модуль demo_3_6_дата_время
- Для созданного сценария в задании №9 добавим компонент даты и времени для выделения даты начала месяца добычи газа.
- Даты замеров в нашей выборке данных размещены неравномерно, что мешает обрабатывать данные временных рядов.
- Добавим в сценарий модуля обработки данных компонент «**Дата и время**»
- Соединим выходной порт «Параметров полей» с входным портом «Дата и время»



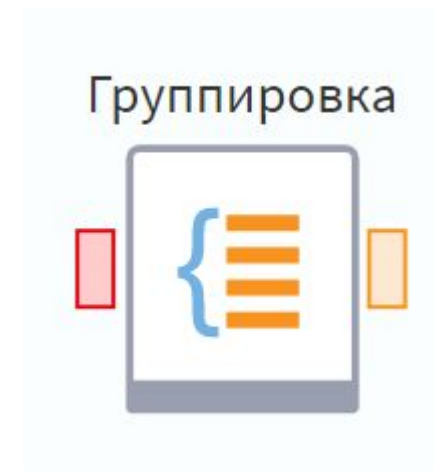
Задание №10. Выделение начала месяца




- Для поля «**Дата замера**» выберем разбиение по дате начала *Год+месяц*
- Сохраним настройки узла и выполним его
- На выходном порте узла активируем быстрый просмотр, чтобы убедиться в результате работы компонента

Разбиение	31 Дата начала	31 Дата конца	12 Число	ab Строка
Обычный				
Год + Квартал	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
Год + Месяц	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
Год + Неделя	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
Год + День	<input type="checkbox"/>			<input type="checkbox"/>
Год	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Квартал			<input type="checkbox"/>	<input type="checkbox"/>
Месяц			<input type="checkbox"/>	<input type="checkbox"/>
Неделя			<input type="checkbox"/>	<input type="checkbox"/>
День года			<input type="checkbox"/>	<input type="checkbox"/>
День квартала			<input type="checkbox"/>	<input type="checkbox"/>
День месяца			<input type="checkbox"/>	<input type="checkbox"/>
День недели			<input type="checkbox"/>	<input type="checkbox"/>
Часы			<input type="checkbox"/>	<input type="checkbox"/>
Минуты			<input type="checkbox"/>	<input type="checkbox"/>
Секунды			<input type="checkbox"/>	<input type="checkbox"/>
Миллисекунды			<input type="checkbox"/>	<input type="checkbox"/>
Дата	<input type="checkbox"/>			<input type="checkbox"/>
Время	<input type="checkbox"/>			<input type="checkbox"/>
Свой формат				<input type="checkbox"/>

Компонент «Группировка» и агрегация данных








- Компонент «Группировка» осуществляет операцию агрегации данных в таблице по полям групп с использованием агрегационных функций для полей показателей
- При настройке группировки можно выбрать два типа полей:
- **Группы** – те поля, в которых содержатся общие для нескольких записей значения групп
- **Показатели** – те поля, на основе которых вычисляется агрегационная функция для записей со схожими значениями в поле **групп**



Выбранные поля		
	Группа	
	Показатели	






Компонент «Группировка» и агрегация данных

- Компонент осуществляет обработку данных по схожему принципу с оператором GROUP BY языка SQL
- Необходимо сначала выбрать агрегационные функции над полями, а затем выбрать поля, по которым происходит группировка вычислений
- Ограничения на поля для групп сохраняются, это должны быть дискретные поля с перечнем уникальных значений

Выбранные поля		
	Группа	
	Дата замера (Год + Месяц, Первый день)	
	ID	
	Куст	
	Группа	
	Показатели	
	Температура устьевая, цельсий (Среднее)	
	Давление устьевое, МПа (Среднее)	
	Скорость добычи газа, м3.сут (Среднее)	
	Скорость добычи конденсата, м3.сут (Среднее)	
	Скорость добычи воды, м3.сут (Среднее)	

Компонент «Группировка» и агрегация данных

- Чтобы перенести поля в группы можно их перетащить туда с помощью ЛКМ одно за другим, или выделив все через **ctrl** нажать на кнопку «**Переместить в группы**» (Alt+C). Поля в группах можно менять местами, меняя порядок вывода групп в выходной таблице
- Показатели переносятся схожим образом, но с помощью другой кнопки «**Переместить в показатели**» (Alt+S)

Выбранные поля		
	Группа	
	Дата замера (Год + Месяц, Первый день)	
	ID	
	Куст	
	Группа	
	Показатели	
	Температура устьевая, цельсий (Среднее)	
	Давление устьевое, МПа (Среднее)	
	Скорость добычи газа, м3.сут (Среднее)	
	Скорость добычи конденсата, м3.сут (Среднее)	
	Скорость добычи воды, м3.сут (Среднее)	

Компонент «Группировка» и агрегация данных

- Настройка вычисляемых статистик доступна на этапе выбора списка показателей. Для выбора вычисляемой статистики для конкретного поля достаточно нажать на него два раза или с помощью ПКМ перейти в раздел выбора показателей.
- Для **строковых данных** недоступны среднее, медиана, СКО и сумма

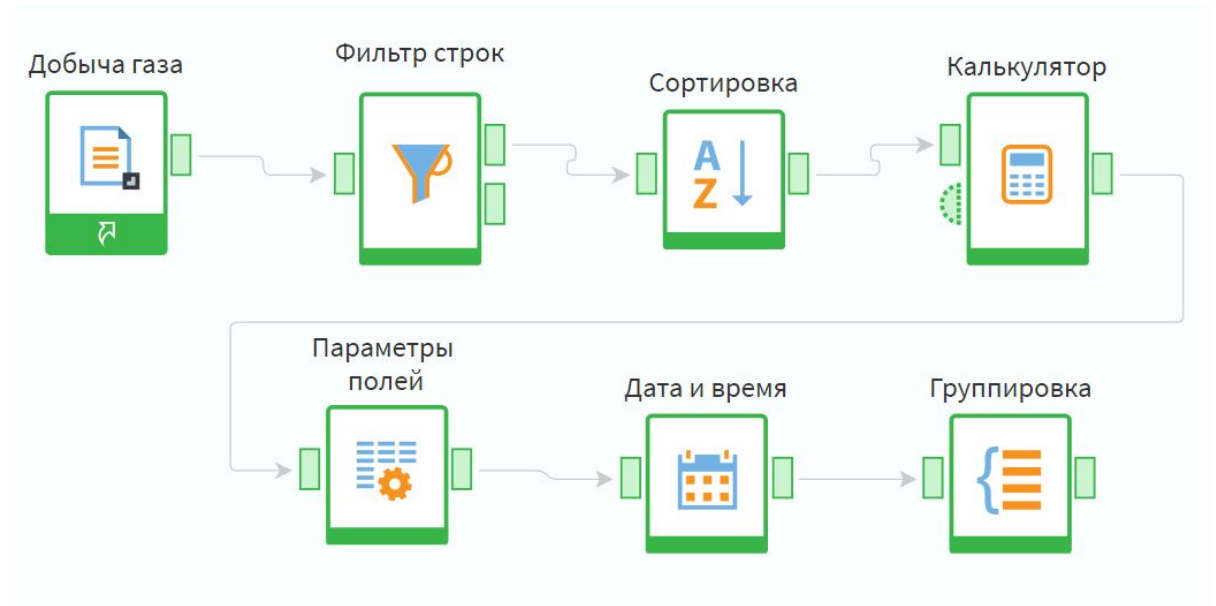
Σ Варианты агрегации

<input type="checkbox"/>	Σ	Сумма	<input type="checkbox"/>	↑↓	Стандартное откл.
<input type="checkbox"/>	#	Количество	<input type="checkbox"/>	S	Кол-во уникальных
<input type="checkbox"/>	↓	Минимум	<input type="checkbox"/>	0	Кол-во пропусков
<input type="checkbox"/>	↑	Максимум	<input type="checkbox"/>	1...	Первый
<input checked="" type="checkbox"/>	—	Среднее	<input type="checkbox"/>	...n	Последний
<input type="checkbox"/>	—	Медиана			

Применить [Отменить](#)

Задание №11. Группировка данных в добыче

- Модуль demo_3_7_группировка_данных
- Добавим в сценарий обработки данных расчёт среднемесячных показателей для того, чтобы нормировать временные показатели измерений в одну интервальную шкалу
- Добавим в сценарий из задания №10 компонент «Группировка»
- Соединим выходной порт «Дата и время» с входным портом «Группировка»



Задание №11. Группировка данных в добыче

В настройке группировки выберем следующие поля в качестве групп:

- Дата замера (Год+месяц)
- ID
- Куст
- Группа

В настройке группировки выберем все числовые поля в качестве показателей и рассчитаем для них **среднее**

Выбранные поля		
	Группа	
	Дата замера (Год + Месяц, Первый день)	
	ID	
	Куст	
	Группа	
	Показатели	
	Температура устьевая, цельсий (Среднее)	
	Давление устьевое, МПа (Среднее)	
	Скорость добычи газа, м3.сут (Среднее)	
	Скорость добычи конденсата, м3.сут (Среднее)	
	Скорость добычи воды, м3.сут (Среднее)	

Задание №11. Группировка данных в добыче

Сколько записей мы получили в результате операции группировки данных?

Какие месяц и год были первыми в списке после группировки?

Какие скважины по ID наблюдались в первый месяц в выходной выборке данных?

Какова максимальная скорость добычи конденсата в первый месяц наблюдения? (до 2 знака после запятой)

Выбранные поля	
↑	Группа
31	Дата замера (Год + Месяц, Первый день)
ab	ID
ab	Куст
ab	Группа
Σ Показатели	
9.0	Температура устьевая, цельсий (Среднее)
9.0	Давление устьевое, МПа (Среднее)
9.0	Скорость добычи газа, м3.сут (Среднее)
9.0	Скорость добычи конденсата, м3.сут (Среднее)
9.0	Скорость добычи воды, м3.сут (Среднее)

Обогащение данных

Компоненты обогащения данных решают задачу увеличения числа строк или столбцов, посредством соединения двух различных таблиц данных

По смыслу операторы обогащения повторяют стандартные операторы JOIN, UNION, MERGE языка SQL

- Слияние – JOIN различного вида
- Дополнение – множественный LEFT JOIN
- Объединение – UNION строк данных
- Соединение – конкатенация таблиц по столбцам

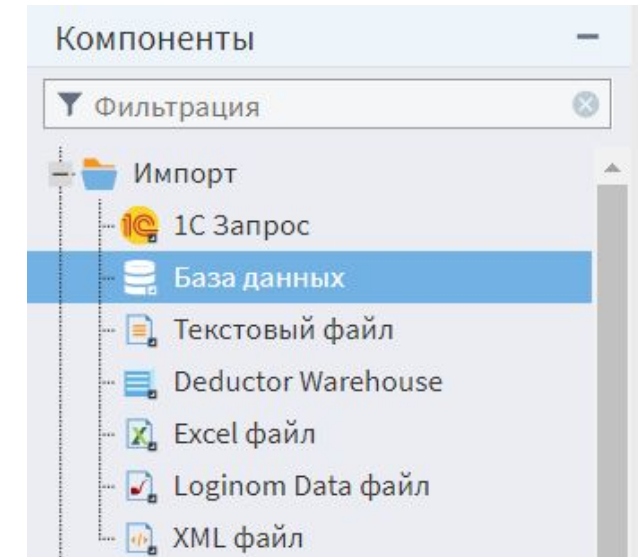
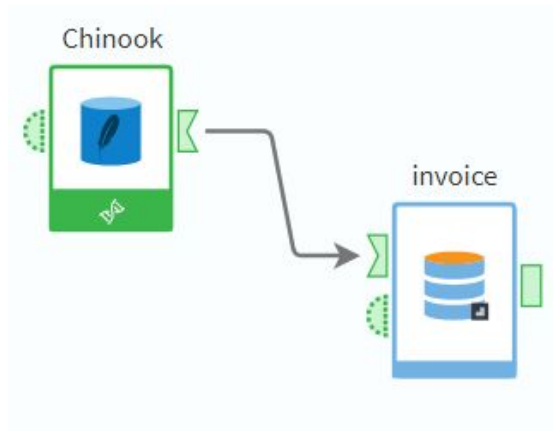


Обогащение данных

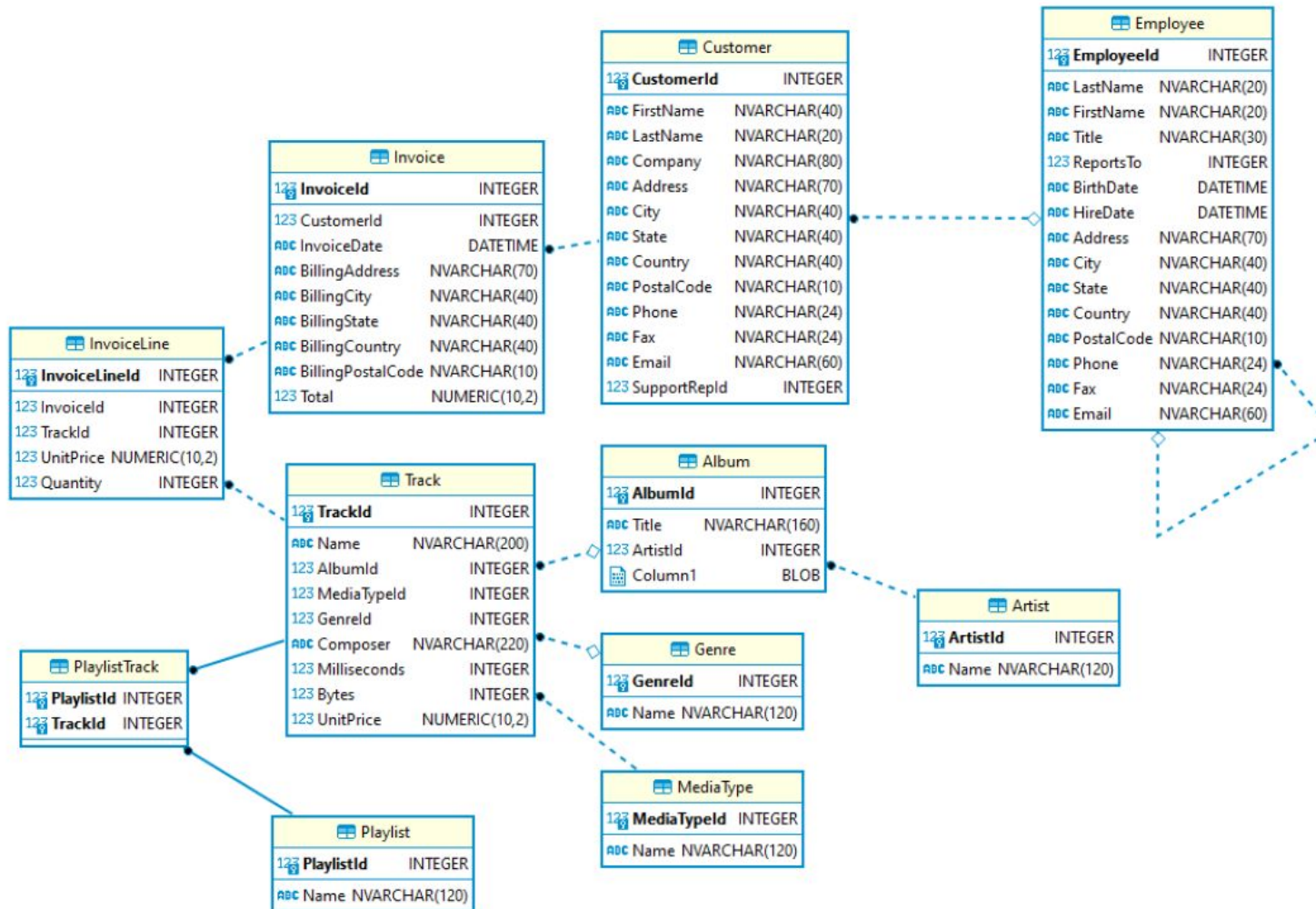
Приведем далее примеры обогащения данных на основе импортированной в Loginom базы данных из задания №4.

Из базы данных Chinook.db будут импортированы таблицы данных и соединены посредством различных методов обогащения данных в Loginom

Примеры обогащения данных представлены в пакете demo_loginom.lgr в модулях demo_4_*

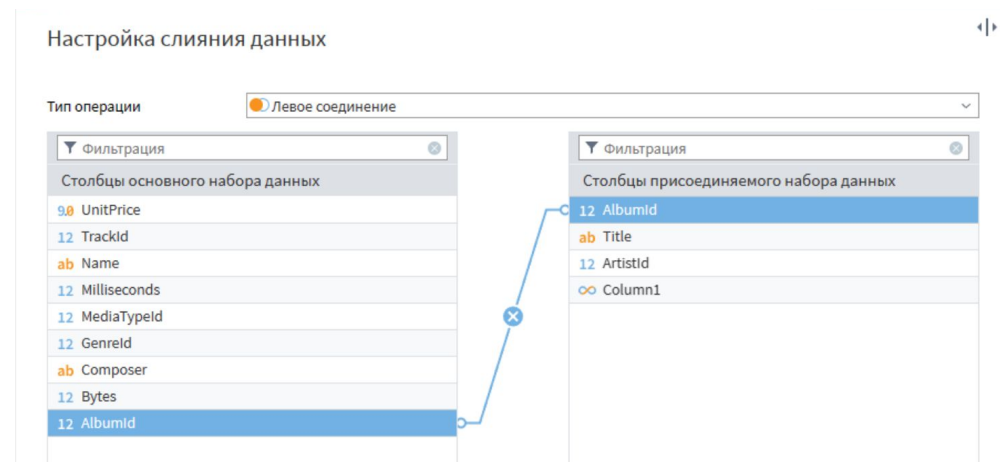
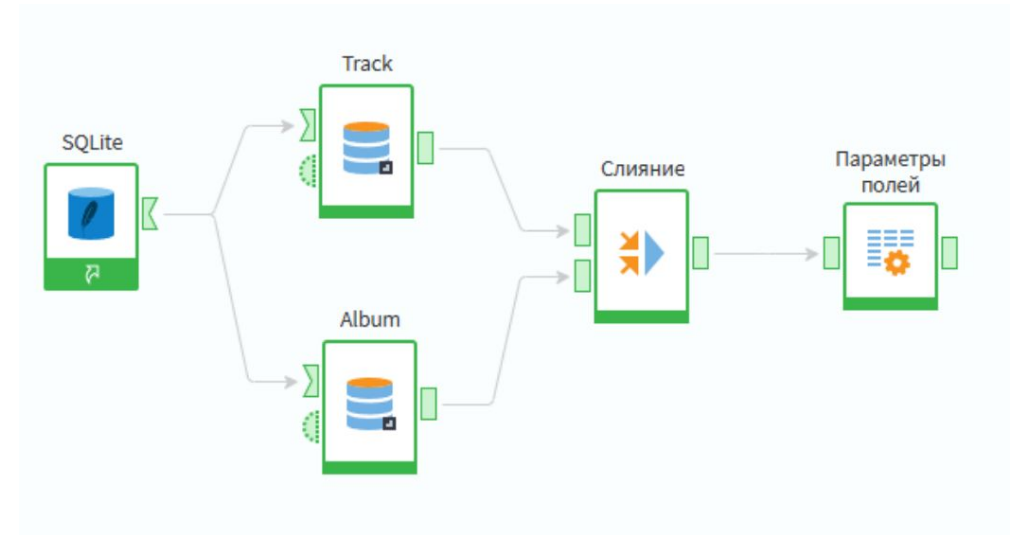


Обогащение данных



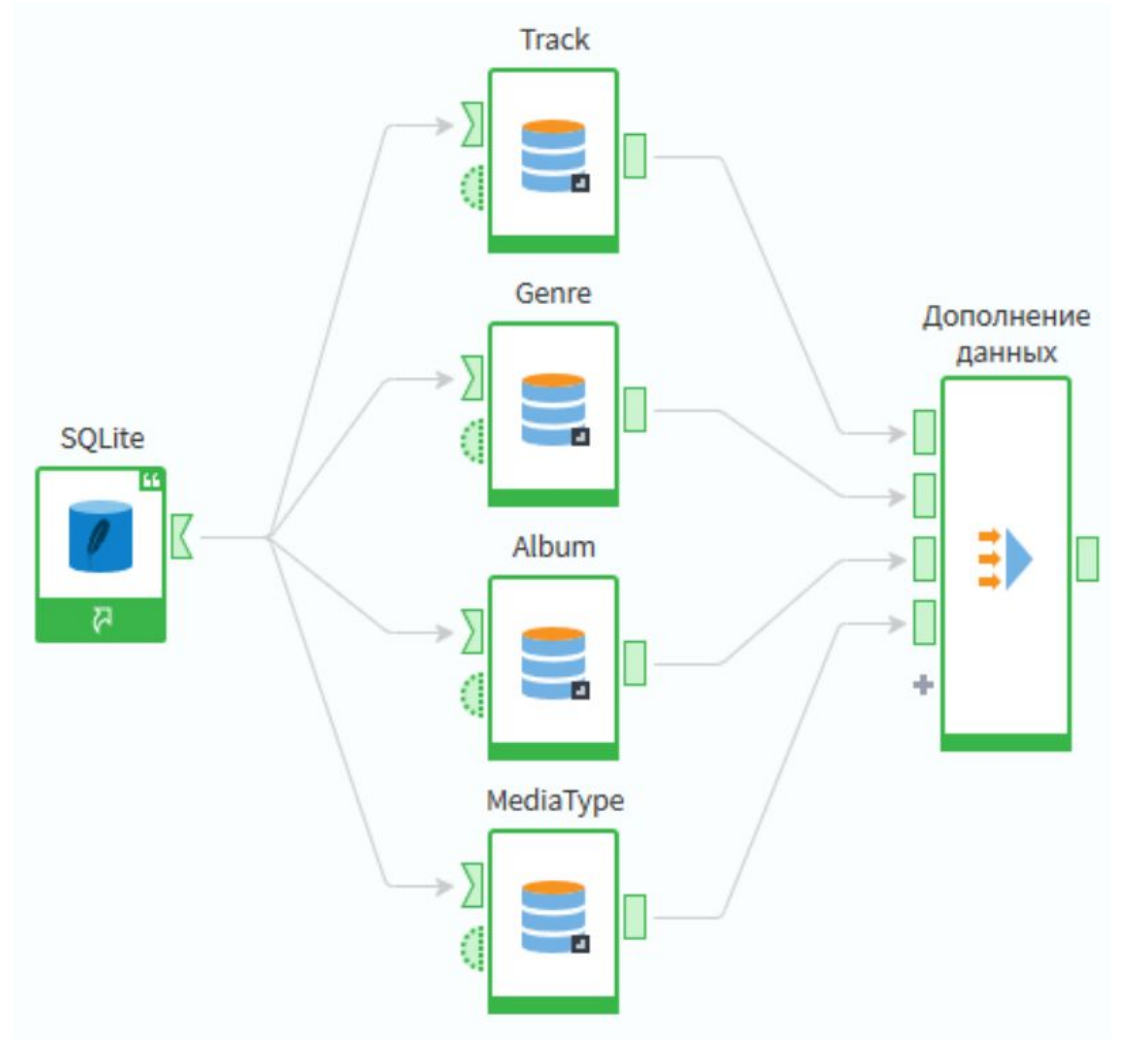
Компонент «Слияние»

- Компонент «**Слияние**» принимает на вход строго 2 таблицы в каждый из входных портов
- Верхний порт отвечает за основной набор данных, нижний порт отвечает за присоединяемый набор данных
- На этапе настройки слияния предлагается выбрать тип операции соединения (левое, внутреннее, ...), а также перетаскиванием полей настроить соответствие соединения по ключам таблиц данных



Компонент «Дополнение данных»

- Компонент «Дополнение данных» соединяет таблицы с помощью левого соединения главной таблицы (верхний порт) с присоединяемыми таблицами (нижние порты)
- Для каждой присоединяемой таблицы в настройках предлагается выбрать по какому полю для каждой входной таблицы стоит соединять поля с данными в единый выходной набор
- Применяется для отношений многие-к-одному



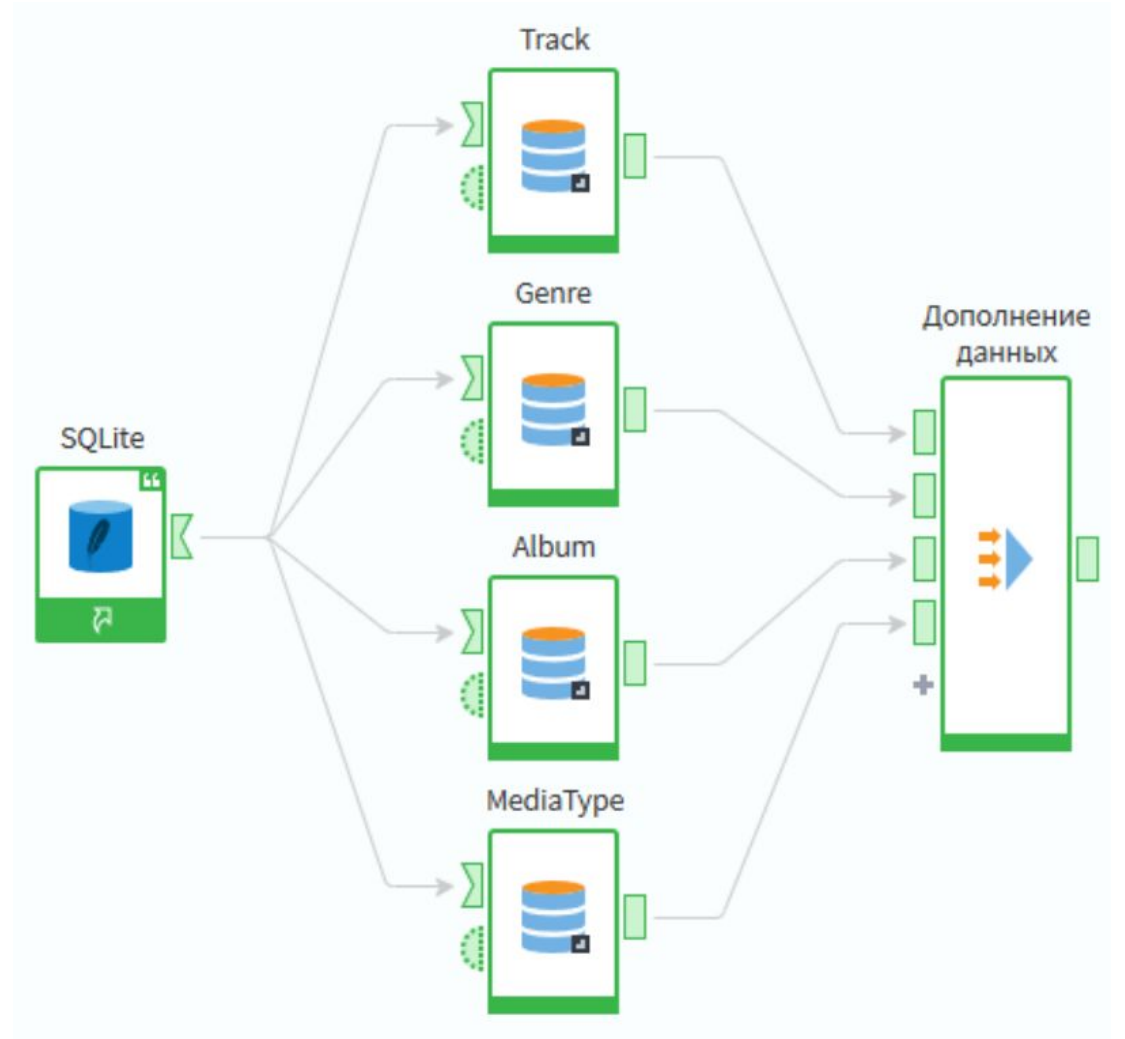
Компонент «Дополнение данных»

Дополнение данных

№	Главная таблица		Присоединяем...		Присоединя...		Присоединяемая таблица 3
1	12 TrackId	<input type="checkbox"/>	Не выбрано	<input type="checkbox"/>	Не выбрано	<input type="checkbox"/>	Не выбрано
2	ab Name	<input type="checkbox"/>	Не выбрано	<input type="checkbox"/>	Не выбрано	<input type="checkbox"/>	Не выбрано
3	12 AlbumId	<input type="checkbox"/>	Не выбрано	<input checked="" type="checkbox"/>	12 AlbumId	<input type="checkbox"/>	Не выбрано
4	12 MediaTypeId	<input type="checkbox"/>	Не выбрано	<input type="checkbox"/>	Не выбрано	<input checked="" type="checkbox"/>	12 MediaTypeId
5	12 GenreId	<input checked="" type="checkbox"/>	12 GenreId	<input type="checkbox"/>	Не выбрано	<input type="checkbox"/>	Не выбрано
6	ab Composer	<input type="checkbox"/>	Не выбрано	<input type="checkbox"/>	Не выбрано	<input type="checkbox"/>	Не выбрано
7	12 Milliseconds	<input type="checkbox"/>	Не выбрано	<input type="checkbox"/>	Не выбрано	<input type="checkbox"/>	Не выбрано
8	12 Bytes	<input type="checkbox"/>	Не выбрано	<input type="checkbox"/>	Не выбрано	<input type="checkbox"/>	Не выбрано
9	90 UnitPrice	<input type="checkbox"/>	Не выбрано	<input type="checkbox"/>	Не выбрано	<input type="checkbox"/>	Не выбрано

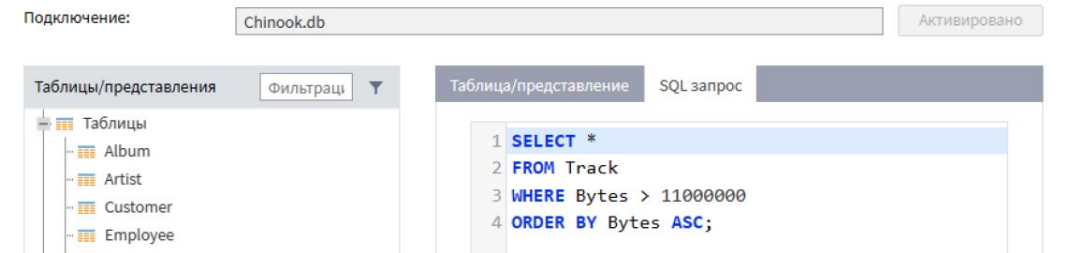
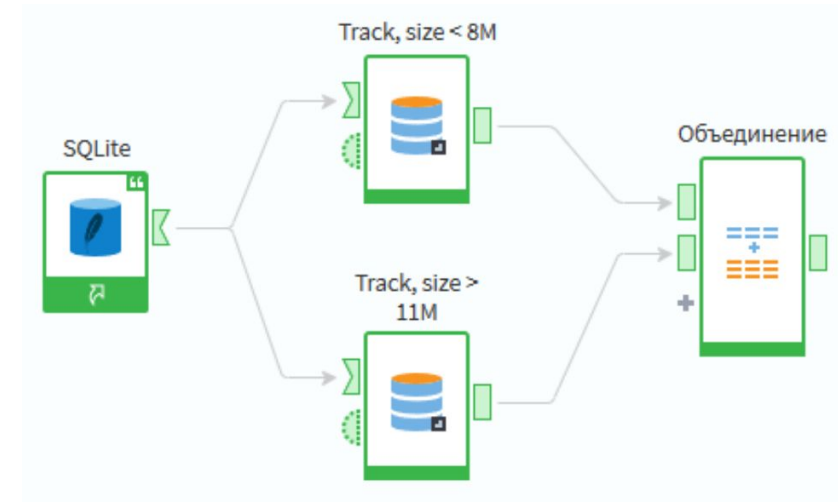
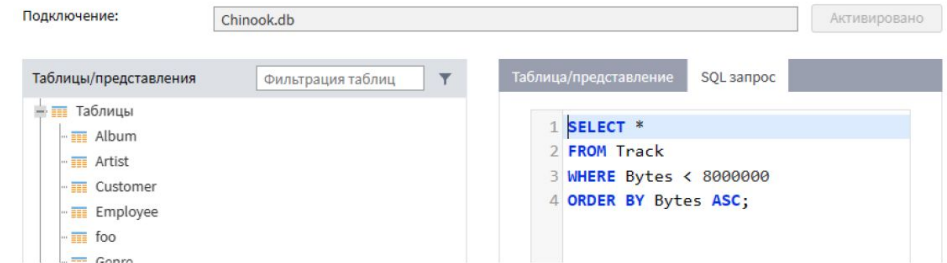
Компонент «Дополнение данных»

- Чаще всего компонент «Дополнение данных» используется для обогащения данных таблицы фактов данными из внешних нормативно-справочных источников данных
- Например для номера заказа мы можем в таблице детализации заказов восстановить дату и время заказа, а также восстановить название блюд из меню по ключам позиций в детализации чеков



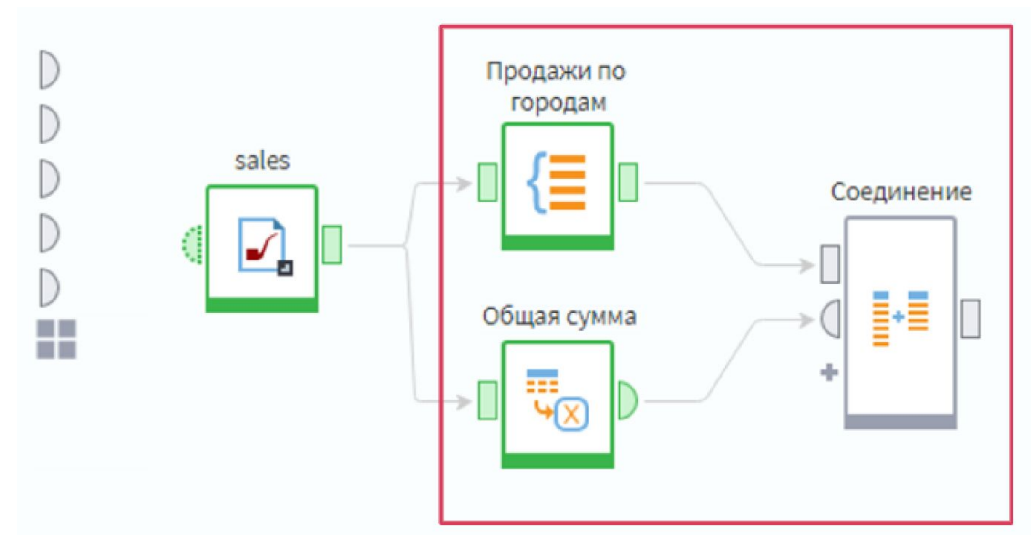
Компонент «Объединение»

- Компонент «Объединение» выполняет роль UNION из языка SQL для обработки данных и соединяет данные из нескольких таблиц строками, дополняя записи данных снизу.
- При настройке узла «Объединение» требуется выбрать общие поля для всех таблиц во избежание нарушения целостности данных
- Выходная таблица обладает настроенной структурой выбранных общих полей для объединяемых таблиц данных



Компонент «Соединение»

- Компонент «Соединение» конкатенирует столбцы данных без привязки к ключу или структуре таблиц
- Необходим для «Сшивания» таблиц по столбцам, если таблицы не имеют общих полей и нужно расположить поля упорядоченных наборов рядом в одной таблице
- При настройке можно указать принцип дополнения строк в меньшей таблице до таблицы с БОльшим числом записей



Компонент «Соединение»

С помощью компонента **Соединение** возможно, к примеру, добавить в набор данных поле с позициями объектов в рейтинге.

Например, в ежемесячном рейтинге по продажам сети магазинов достаточно будет отсортировать магазины по убыванию суммы продаж, после чего можно добавить поле с позициями.



The diagram illustrates the 'Join' component. It shows a table of sales data being joined with a ranking table to produce a final table with positions.

Магазин	Сумма
Магазин 4	69 821
Магазин 1	55 071
Магазин 2	42 800
Магазин 3	31 212

+

Позиция
1 место
2 место
3 место
4 место

=

Магазин	Сумма	Позиция
Магазин 4	69 821	1 место
Магазин 1	55 071	2 место
Магазин 2	42 800	3 место
Магазин 3	31 212	4 место

Работа с переменными в Logiном

Создание пользовательских переменных

Использование переменных в узлах

Соединение, таблица в переменные, калькулятор переменных

Использование переменных для параметризации сценария обработки данных

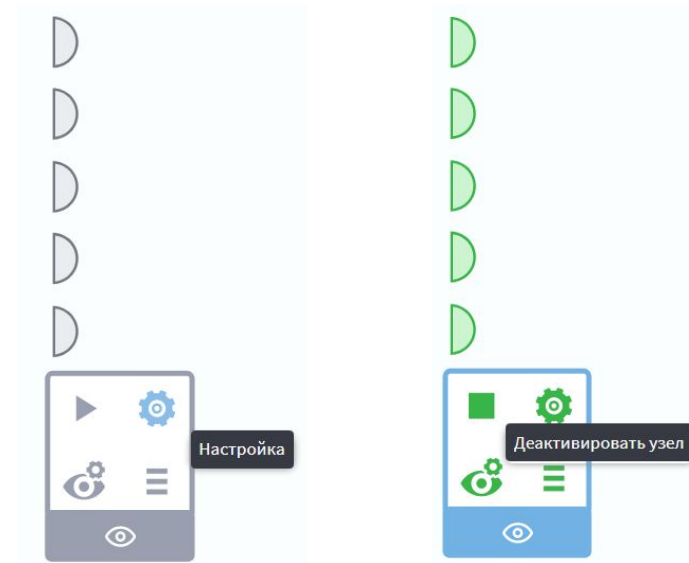
Работа с переменными в Logiном

- Переменные в Logiном используются для создания динамических сценариев.
- **Переменная** – объект, содержащий одно значение.
- Переменная имеет такие же параметры, что и поле набора данных: Метку, Имя, Тип, Значение.
- Некоторые узлы Logiном могут обладать **портом входных переменных** и **портом управляющих переменных**.
- При открытии нового пакета в области построения сценария есть 5 портов переменных – **Переменные сценария**.
- Порты переменных в сценарии расположены в следующем порядке:

Работа с переменными в Loginom

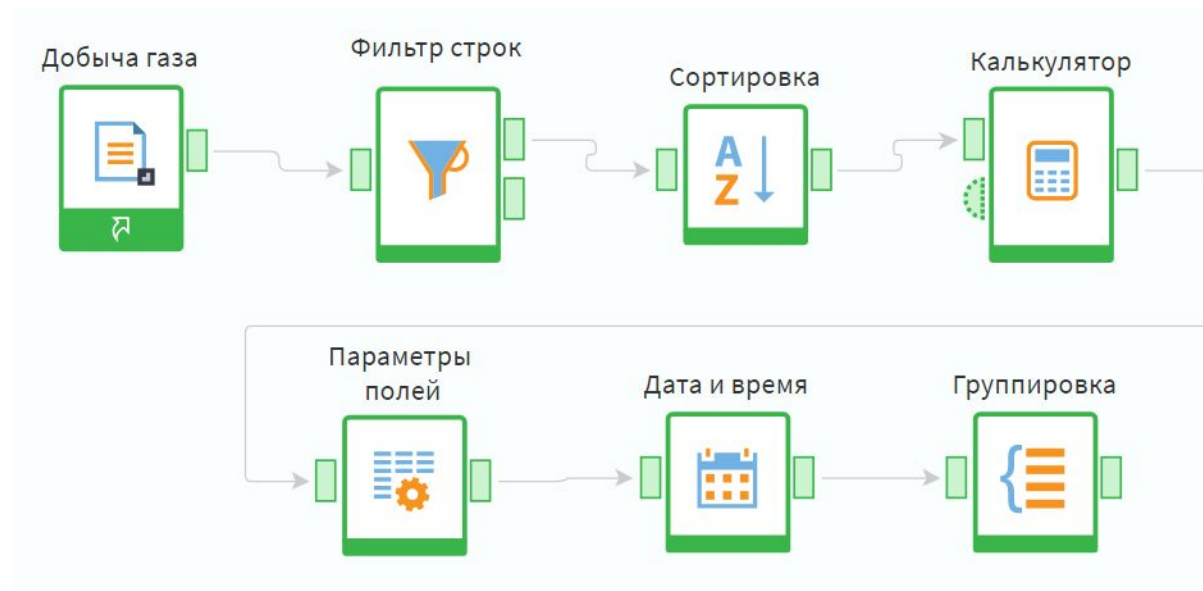
- **Переменные** – порт содержит все переменные сценария с портов 2-5.
- **Переменные системы** – переменные окружения ОС, считываются из настроек ОС.
- **Переменные сессии** – содержат сведения о текущей сессии.
- **Переменные пакета** – содержат информацию о текущем пакете.
- **Переменные пользователя** – переменные, созданные пользователем.

Порт переменных пользователя можно изменять и добавлять в него новые переменные и их значения, остальные порты изменять нельзя.



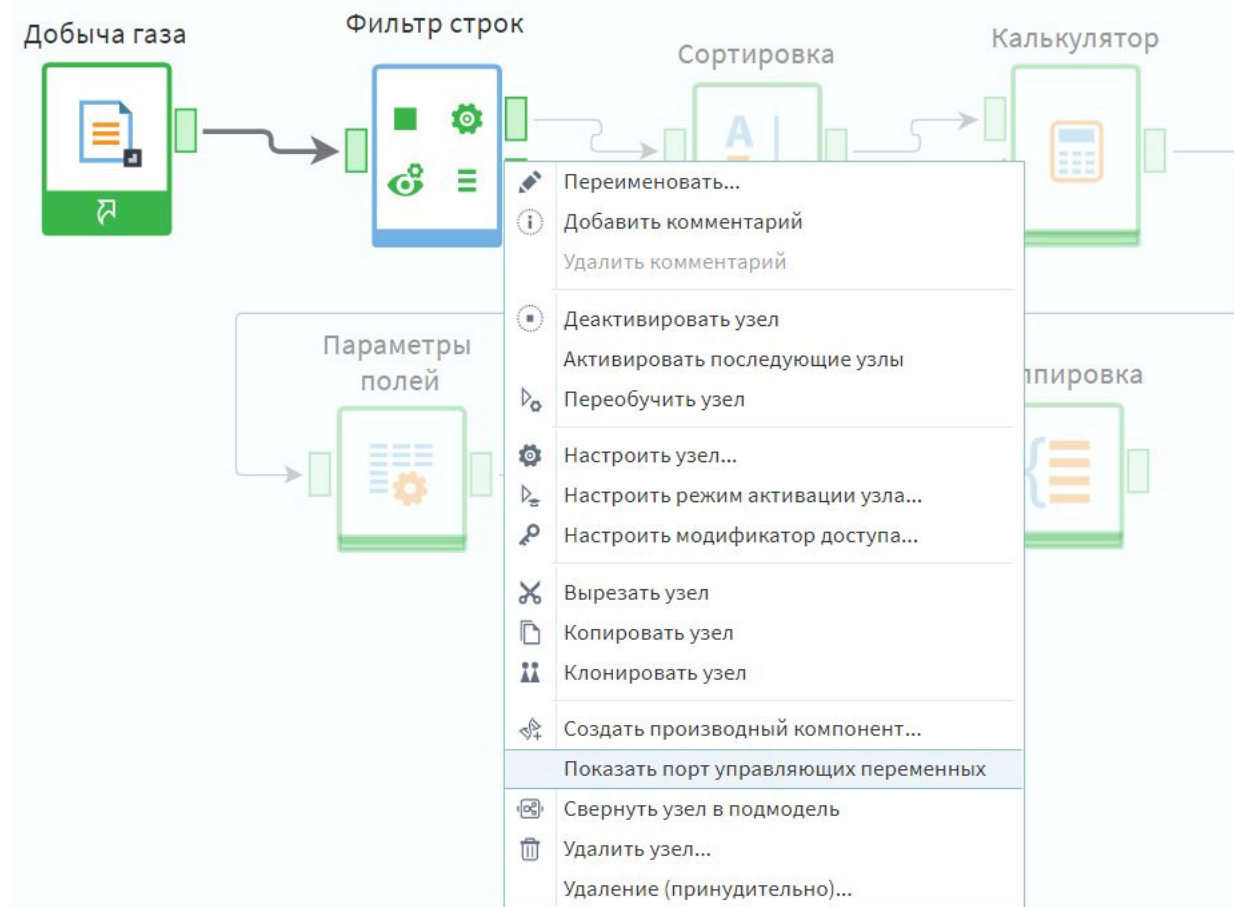
Задание №12. Работа с переменными Loginom

- Зададим параметризацию нашего сценария обработки данных добычи газа на месторождениях
- Добавим в сценарий возможность считывать имя файла из входной переменной
- Осуществлять фильтрацию данных по параметру ID скважины
- Задавать ограничения в калькуляторе с помощью переменных



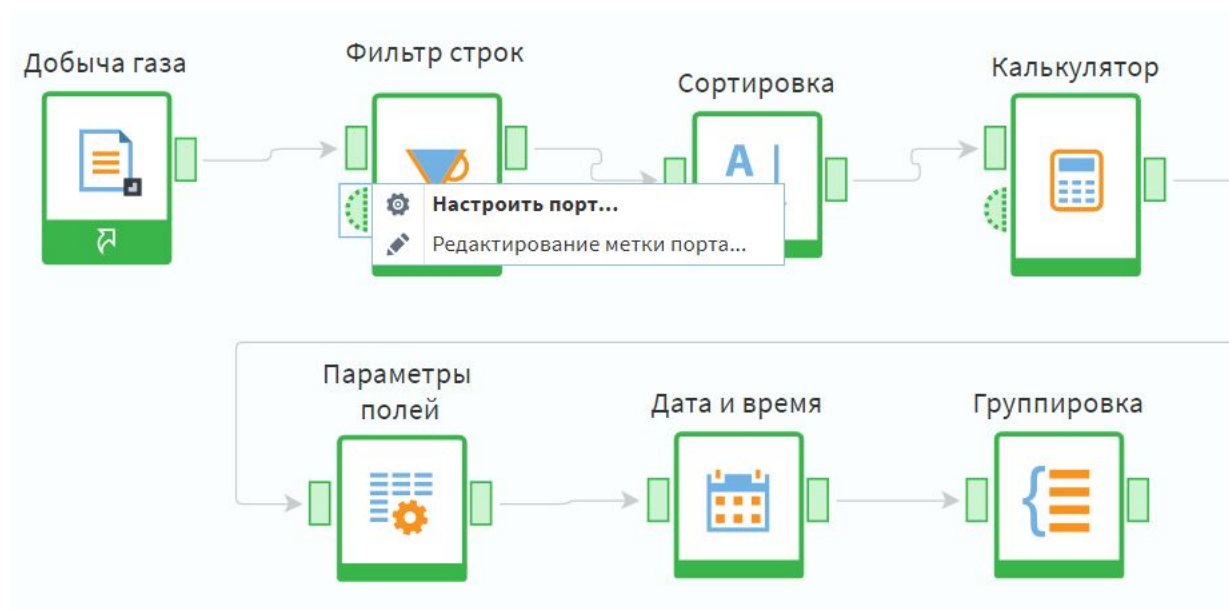
Задание №12. Работа с переменными Loginom

С помощью ПКМ по узлу и выбора опции меню «Показать порт управляющих переменных» для фильтра строк создадим у него порт переменных, который сможет принимать входные параметры



Задание №12. Работа с переменными Loginom

С помощью ПКМ по появившемуся порту перейдем в его настройки



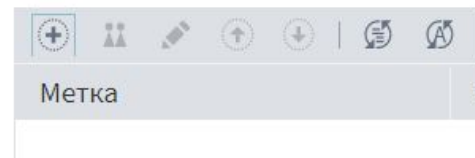
Задание №12. Работа с переменными Loginom

С помощью кнопки **добавить** мы сможем задать новую подаваемую переменную заданного типа с значением по умолчанию

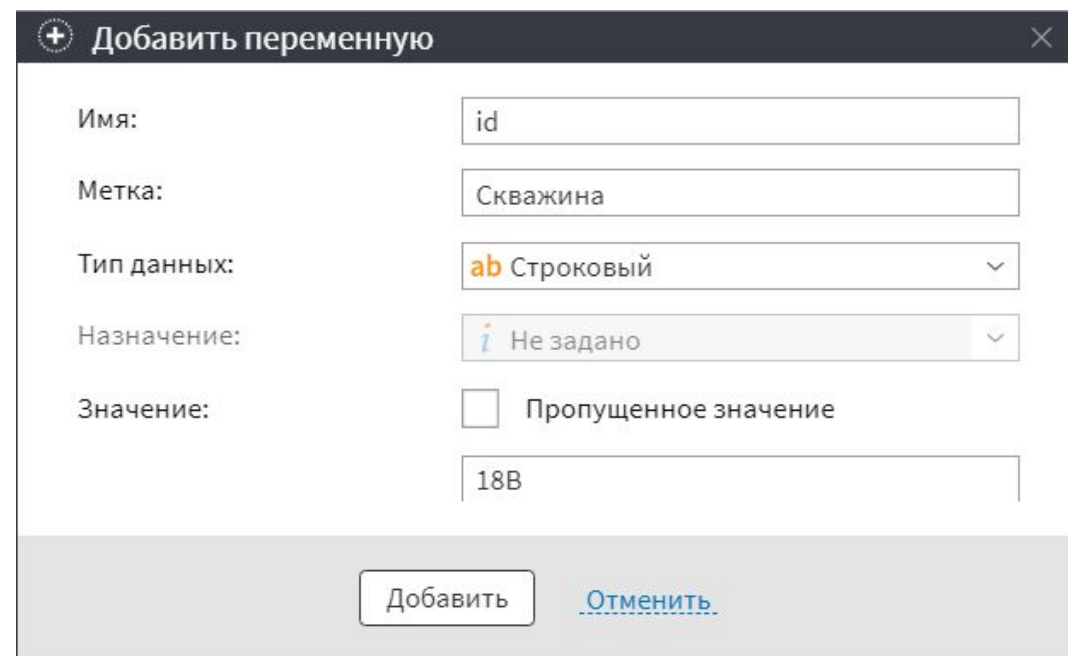
Создадим новую переменную как показано на слайде

Отключим **автосинхронизацию** с помощью кнопки рядом с кнопкой **добавить**

Настройка переменных



Метка	Имя
-------	-----



Добавить переменную

Имя:

Метка:

Тип данных:

Назначение:

Значение: Пропущенное значение

[Отменить](#)

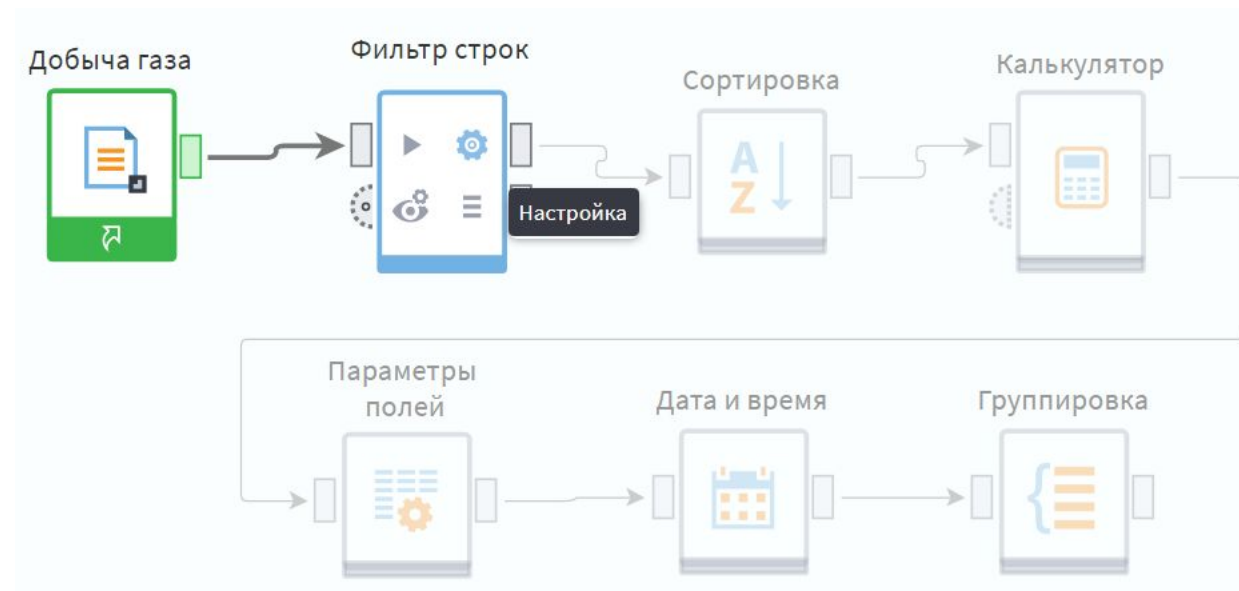
Задание №12. Работа с переменными Loginom

После создания новой переменной на входе порта управляющей переменной сохраним все настройки порта и вернемся в область построения сценария

Настроим узел «Фильтр строк»



Метка	Имя	Назначение	Значение
ab Скважина	id	Не задано	18В



Задание №12. Работа с переменными Loginom

В перечень условий добавим ещё одно условие по совпадению ID скважины со значением из внешней переменной

Если порт управляющих переменных был настроен корректно, то в значении для сравнения будет успешно добавлен переключатель, позволяющий передавать в данное поле значение из внешней переменной

The screenshot shows the Loginom interface with a list of conditions and a configuration dialog for a new condition.

Conditions list:

- 1. Ртр МПа не пустое
- 2. И. 9.0 конд-т мЗ/сут не пустое
- 3. ab Группа не пустое

Configuration dialog for condition 3:

- Поле: ab ID
- Условие: =
- Значение для сравнения: Скважина (18В)
- Учитывать регистр:
- Buttons: Применить, Отменить

Задание №12. Работа с переменными Loginom

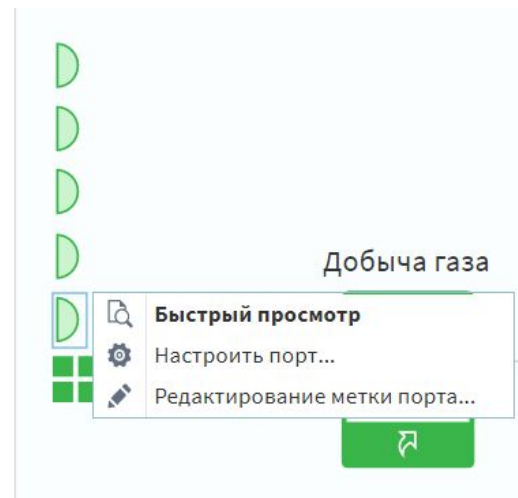
На выходном порте данных, соответствующих условию наблюдаем ожидаемую функциональность, что поданное значение по умолчанию для фильтрации данных обрабатывает в штатном режиме

Далее добавим в сценарий пользовательские переменные, которые можно будет использовать для эффективной параметризации сценария

Соответствуют условию		Не соответствуют условию								
#	ab ID	Дата замера	Ртр М...	Туст °С	газ м3/...	конд-т м3/сут	вода м3/сут	ab Куст	ab Груп...	
1	18B	29.10.2014, 00:00	6,20	64,90	478 689,10	0,45	0,48	КП 19	Хаузак	
2	18B	09.11.2014, 00:00	6,20	64,70	522 921,40	0,81	0,61	КП 19	Хаузак	
3	18B	01.12.2014, 00:00	61,00	63,90	452 048,00	0,56	1,59	КП 19	Хаузак	
4	18B	23.12.2014, 00:00	60,00	62,30	460 916,20	0,56	1,59	КП 19	Хаузак	
5	18B	24.12.2014, 00:00	60,00	62,60	460 171,40	0,56	1,59	КП 19	Хаузак	
6	18B	12.01.2015, 00:00	60,00	65,80	437 359,20	0,76	1,07	КП 19	Хаузак	
7	18B	13.01.2015, 00:00	60,00	65,20	436 971,20	0,68	1,24	КП 19	Хаузак	
8	18B	14.02.2015, 00:00	60,00	65,00	417 788,60	0,57	1,52	КП 19	Хаузак	
9	18B	15.02.2015, 00:00	60,00	64,10	418 137,60	0,60	1,59	КП 19	Хаузак	
10	18B	05.03.2015, 00:00	60,00	65,00	405 222,30	0,49	1,66	КП 19	Хаузак	
11	18B	06.03.2015, 00:00	60,00	66,30	404 942,40	0,54	1,60	КП 19	Хаузак	
12	18B	20.03.2015, 00:00	60,00	66,50	394 253,40	0,49	1,45	КП 19	Хаузак	
13	18B	21.03.2015, 00:00	60,00	65,60	393 910,20	0,49	1,45	КП 19	Хаузак	
14	18B	02.04.2015, 00:00	60,00	66,70	398 265,80	0,35	1,18	КП 19	Хаузак	
15	18B	03.04.2015, 00:00	60,00	66,90	395 436,50	0,40	1,09	КП 19	Хаузак	
16	18B	04.05.2015, 00:00	60,00	66,60	362 321,30	0,41	0,44	КП 19	Хаузак	
17	18B	05.05.2015, 00:00	60,00	67,10	360 730,40	0,44	0,56	КП 19	Хаузак	
18	18B	09.06.2015, 00:00	6,40	65,50	351 632,40	0,63	0,30	КП 19	Хаузак	
19	18B	10.06.2015, 00:00	6,40	65,30	352 763,60	0,39	0,29	КП 19	Хаузак	
20	18B	11.06.2015, 00:00	6,40	66,10	352 160,20	0,42	0,29	КП 19	Хаузак	
21	18B	18.07.2015, 00:00	5,90	64,90	365 683,10	0,42	1,36	КП 19	Хаузак	
22	18B	19.07.2015, 00:00	5,80	66,00	367 265,40	0,38	1,39	КП 19	Хаузак	
23	18B	03.08.2015, 00:00	6,00	70,00	349 566,30	0,63	1,44	КП 19	Хаузак	
24	18B	04.08.2015, 00:00	6,00	70,00	340 208,60	0,37	1,45	КП 19	Хаузак	

Задание №12. Работа с переменными Loginom

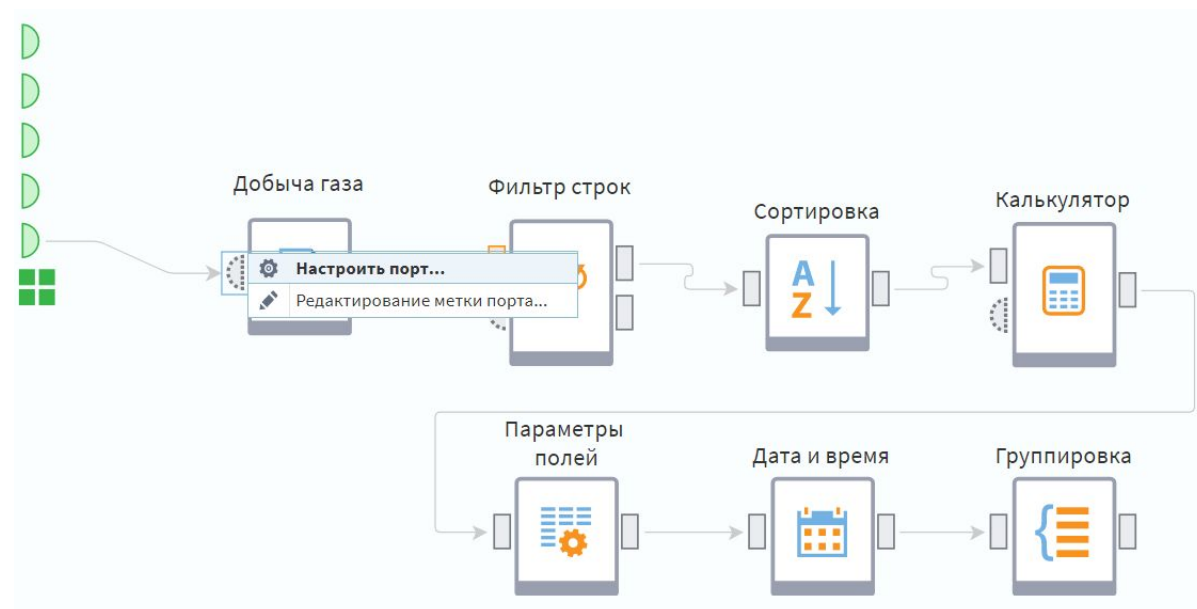
- На порте переменных щелкнем ПКМ и перейдем к настройке порта
- Создадим на порте набор переменных, показанных на слайде
- Сохраним настройки



Метка	Имя	Назначение	Значение
ab Путь к файлу импорта	path_import	Не задано	Data/GAZ/GAZ_n.csv
ab Путь к файлу экспорта	path_export	Не задано	Data/GAZ/GAZ_export.lgd
ab Скважина	id	Не задано	17B
9.0 Верхняя граница для давле...	pressure_upper	Не задано	20,00
9.0 Верхняя граница температу...	temp_upper	Не задано	100,00
9.0 Нижняя граница добычи газа	gaz_lower	Не задано	1 000,00

Задание №12. Работа с переменными Loginom

- Соединим порт переменных пользователя с входной переменной добычи газа
- Перейдем к настройке порта входных переменных узла импорта данных



Задание №12. Работа с переменными Loginom

- В настройках порта удалим все переменные, полученные от порта переменных пользователя и оставим только переменную строкового типа `path_import`
- Автосинхронизация отключается автоматически при удалении полей на порте узла
- Сохраним настройки порта и перейдем к настройкам узла импорта данных

Настройка соответствия между переменными

Таблица Связи

Входные	Выходные	Имя	Назначение	Значение	
ab Путь к файлу имп...	ab Путь к файлу имп...	path_import	Не задано	Data/GAZ/GAZ_n.csv	
ab Путь к файлу эксп...	ab Путь к файлу эксп...	path_export	Не задано	Data/GAZ/GAZ_export...	🗑
ab Скважина	ab Скважина	id	Не задано	17B	🗑
9.0 Верхняя граница ...	9.0 Верхняя граница ...	pressure_upper	Не задано	20,00	🗑
9.0 Верхняя граница ...	9.0 Верхняя граница ...	temp_upper	Не задано	100,00	🗑
9.0 Нижняя граница ...	9.0 Нижняя граница ...	gaz_lower	Не задано	1 000,00	🗑

- Добавить...
- Клонировать...
- Редактировать... (F2)
- Удалить выбранные (Delete)
- Переместить вверх (Ctrl+Up)
- Переместить вниз (Ctrl+Down)
- Синхронизировать переменные
- Автоматическая синхронизация
- Связать автоматически
- Удалить все связи
- Отображать потерянные связи

Задание №12. Работа с переменными Loginom

- В настройках узла импорта данных включим передачу значения для **Имя файла** из внешней переменной, нажав на переключатель рядом с полем
- Заметим, что значение из переменной успешно передано
- Остальные настройки импорта оставляем без изменений

Импорт из текстового файла

Хранилище файлов: Подключено

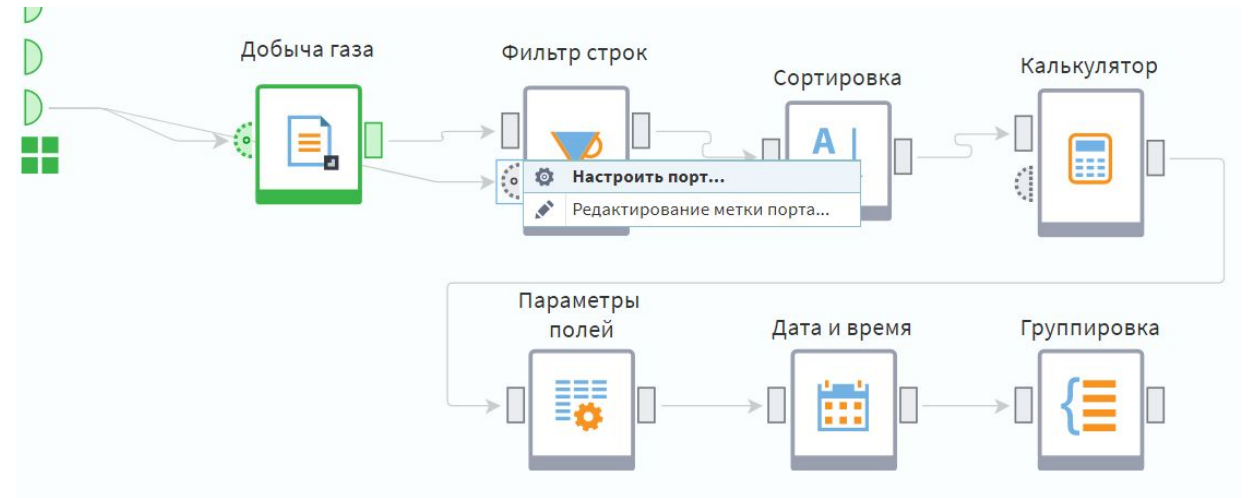
Имя файла:

Кодовая страница:

Заголовок в первой строке: Пропустить строк:

Задание №12. Работа с переменными Loginom

- Включим импорт данных, на основе быстрого просмотра убедимся в том, что данные успешно импортированы в сценарий обработки данных
- Передадим переменные пользователя на вход фильтра строк и перейдем к его настройкам



Задание №12. Работа с переменными Loginom

- В порте фильтра строк укажем в качестве входного значения для настроенной входной переменной поле **Скважина (id)**.
- Сохраним настройки и запустим выполнение узла

Настройка соответствия между переменными

Таблица Связи

Входные	Выходные	Имя	Назначение	Значение	
Скважина	ab Скважина	id	Не задано	17B	
ab Путь к файлу импорт.					
ab Путь к файлу экспорт					
ab Скважина <i>id</i>					

Задание №12. Работа с переменными Loginom

- На основе быстрого просмотра поля данных, соответствующих условию убедимся в том, что фильтр строк работает в соответствии с поданным на него значением для сравнения из порта входных переменных

ab ID	Дата замера	Ртр М...	Туст °С	газ м3/...	конд-т м3/сут	вода м3/сут	Куст	Груп...
17B	13.10.2014, 00:00	6,40	48,70	247 620,20	0,61	0,31	КП 19	Хаузак
17B	04.11.2014, 00:00	6,20	48,60	239 202,10	0,63	0,42	КП 19	Хаузак
17B	29.11.2014, 00:00	6,30	49,30	233 614,30	0,56	1,09	КП 19	Хаузак
17B	22.12.2014, 00:00	61,00	48,30	235 814,40	0,56	1,09	КП 19	Хаузак
17B	01.01.2015, 00:00	62,00	47,30	240 051,70	0,56	1,09	КП 19	Хаузак
17B	15.01.2015, 00:00	62,00	44,80	230 984,60	0,56	0,89	КП 19	Хаузак
17B	16.01.2015, 00:00	62,00	44,70	231 896,60	0,63	0,88	КП 19	Хаузак
17B	03.02.2015, 00:00	62,00	45,80	232 061,10	0,65	0,76	КП 19	Хаузак
17B	04.02.2015, 00:00	62,00	45,20	232 202,10	0,63	0,75	КП 19	Хаузак
17B	05.02.2015, 00:00	62,00	46,80	232 620,30	0,63	0,72	КП 19	Хаузак
17B	18.02.2015, 00:00	62,00	49,30	227 370,40	0,49	1,01	КП 19	Хаузак
17B	18.02.2015, 00:00	62,00	50,30	227 855,10	0,48	1,02	КП 19	Хаузак
17B	08.03.2015, 00:00	62,00	50,30	213 618,20	0,51	0,95	КП 19	Хаузак
17B	09.03.2015, 00:00	62,00	49,50	213 681,00	0,50	0,96	КП 19	Хаузак
17B	06.05.2015, 00:00	62,00	51,90	200 398,10	0,36	0,19	КП 19	Хаузак
17B	23.05.2015, 00:00	62,00	50,20	211 716,80	0,33	0,19	КП 19	Хаузак
17B	24.05.2015, 00:00	62,00	50,60	212 654,20	0,43	0,22	КП 19	Хаузак
17B	25.05.2015, 00:00	62,00	50,50	212 770,10	0,43	0,22	КП 19	Хаузак
17B	14.06.2015, 00:00	6,60	51,70	201 726,20	0,80	0,09	КП 19	Хаузак
17B	16.07.2015, 00:00	6,10	47,90	221 409,30	0,59	0,80	КП 19	Хаузак
17B	17.07.2015, 00:00	6,10	49,50	218 340,20	0,54	0,89	КП 19	Хаузак
17B	02.08.2015, 00:00	6,00	64,00	203 264,40	0,40	1,00	КП 19	Хаузак
17B	27.08.2015, 00:00	6,00	63,00	179 075,40	0,22	0,94	КП 19	Хаузак
17B	28.08.2015, 00:00	5,00	64,00	182 261,70	0,25	0,72	КП 19	Хаузак

Задание №12. Работа с переменными Loginom

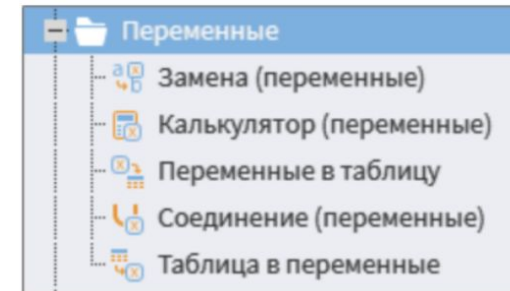
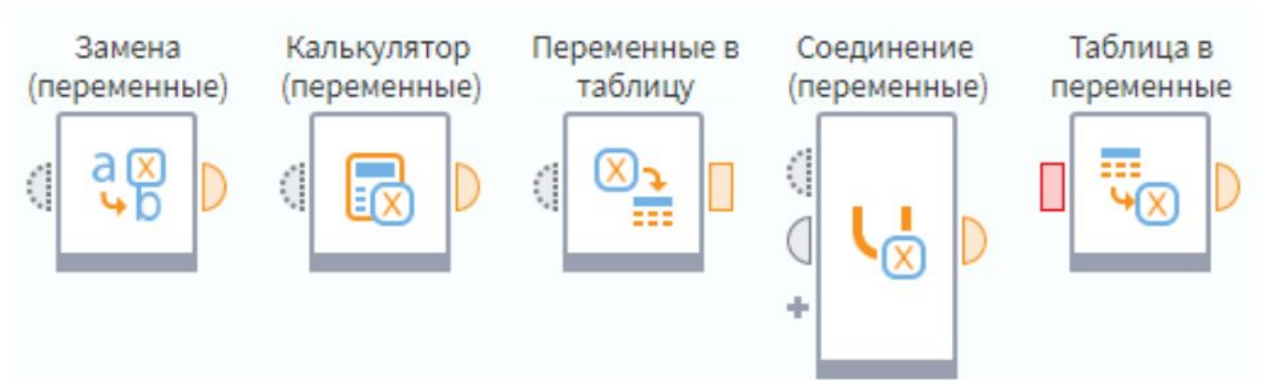
- Добавьте на входном порте переменных калькулятора переменные, отвечающие за прохождение условия для вычисляемых в нем значений в соответствии с названием и меткой созданных переменных пользователя и полей калькулятора из задания №8
- Добавьте узел экспорта данных, имя файла для записи в котором будет определяться внешней переменной пользователя

ab ID	31 Дата замера	90 Ртр М...	90 Туст °С	90 газ м3/...	90 конд-т м3/сут	90 вода м3/сут	ab Куст	ab Груп...
17B	13.10.2014, 00:00	6,40	48,70	247 620,20	0,61	0,31	КП 19	Хаузак
17B	04.11.2014, 00:00	6,20	48,60	239 202,10	0,63	0,42	КП 19	Хаузак
17B	29.11.2014, 00:00	6,30	49,30	233 614,30	0,56	1,09	КП 19	Хаузак
17B	22.12.2014, 00:00	61,00	48,30	235 814,40	0,56	1,09	КП 19	Хаузак
17B	01.01.2015, 00:00	62,00	47,30	240 051,70	0,56	1,09	КП 19	Хаузак
17B	15.01.2015, 00:00	62,00	44,80	230 984,60	0,56	0,89	КП 19	Хаузак
17B	16.01.2015, 00:00	62,00	44,70	231 896,60	0,63	0,88	КП 19	Хаузак
17B	03.02.2015, 00:00	62,00	45,80	232 061,10	0,65	0,76	КП 19	Хаузак
17B	04.02.2015, 00:00	62,00	45,20	232 202,10	0,63	0,75	КП 19	Хаузак
17B	05.02.2015, 00:00	62,00	46,80	232 620,30	0,63	0,72	КП 19	Хаузак
17B	18.02.2015, 00:00	62,00	49,30	227 370,40	0,49	1,01	КП 19	Хаузак
17B	18.02.2015, 00:00	62,00	50,30	227 855,10	0,48	1,02	КП 19	Хаузак
17B	08.03.2015, 00:00	62,00	50,30	213 618,20	0,51	0,95	КП 19	Хаузак
17B	09.03.2015, 00:00	62,00	49,50	213 681,00	0,50	0,96	КП 19	Хаузак
17B	06.05.2015, 00:00	62,00	51,90	200 398,10	0,36	0,19	КП 19	Хаузак
17B	23.05.2015, 00:00	62,00	50,20	211 716,80	0,33	0,19	КП 19	Хаузак
17B	24.05.2015, 00:00	62,00	50,60	212 654,20	0,43	0,22	КП 19	Хаузак
17B	25.05.2015, 00:00	62,00	50,50	212 770,10	0,43	0,22	КП 19	Хаузак
17B	14.06.2015, 00:00	6,60	51,70	201 726,20	0,80	0,09	КП 19	Хаузак
17B	16.07.2015, 00:00	6,10	47,90	221 409,30	0,59	0,80	КП 19	Хаузак
17B	17.07.2015, 00:00	6,10	49,50	218 340,20	0,54	0,89	КП 19	Хаузак
17B	02.08.2015, 00:00	6,00	64,00	203 264,40	0,40	1,00	КП 19	Хаузак
17B	27.08.2015, 00:00	6,00	63,00	179 075,40	0,22	0,94	КП 19	Хаузак
17B	28.08.2015, 00:00	5,00	64,00	182 261,70	0,25	0,72	КП 19	Хаузак

Работа с переменными Loginom

В Loginom есть 5 специальных компонентов для работы с переменными.

- Замена (переменные)
- Калькулятор (переменные)
- Переменные в таблицу
- Соединение (переменные)
- Таблица в переменные

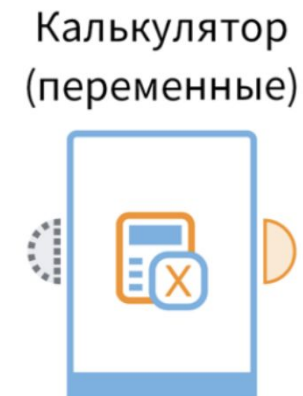


Замена (переменные) и Калькулятор (переменные)

Замена (переменные) –
заменяет значения выбранных
переменных на значения, взятых
в таблице замен



Калькулятор (переменные) –
Создаёт новые переменные в
соответствии с введёнными
формулами



Переменные в таблицу и Соединение (переменные)

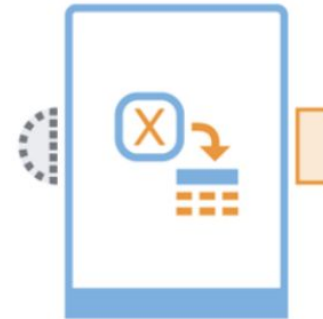
Переменные в таблицу – преобразует набор переменных в таблицу данных

Соединение (переменные) – объединяет наборы переменных в один (может объединять переменные с разных ветвей узла Условие)

У компонента динамическое количество портов, которые можно добавлять с помощью кнопки Добавить ещё один порт +.

По умолчанию – 2 порта: Входные переменные, Добавляемые переменные.

Переменные в таблицу



Соединение (переменные)

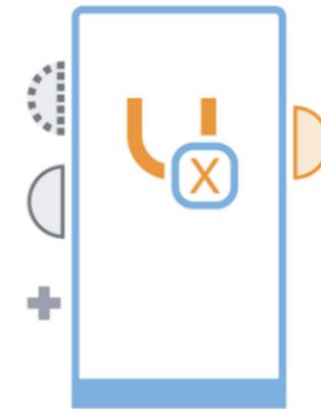


Таблица в переменные

Преобразует табличные данные в переменные выбранным способом.

Позволяет получить агрегированное значение по нужному полю, сохранить его как переменную (можно использовать в дальнейших расчетах).

При подаче новых данных значение переменной пересчитывается во время обработки, оставив сценарий актуальным.

Таблица
в переменные



Самостоятельная работа. Таможенные данные

Исследование качества данных

Обогащение данных

Агрегация данных

Задание на выполнение

Импортировать данные таможенной статистики по экспорту и импорту в регионах РФ за 2016 год

- Сколько строк имеет выборка данных за 2016 год?
- Какое количество столбцов имеет данная выборка?
- Какую информацию хранит каждое поле в отдельности? Какие данные необходимо еще добыть, чтобы расшифровать значения в полях `country`, `code`, `region`, `district`?
- Имеются ли в данных пропуски? На основе компонентов **таблица в переменные, группировка**, визуализатор **статистика**, ответьте на вопрос о количестве пропущенных значений в каждом из столбцов выборки данных (можно использовать что-то одно).
- Сколько уникальных стран участников торговли перечислено в поле `country`?
- Сколько уникальных регионов России участвовало в торговле (поле `region`)?

Задание на выполнение

Импортировать данные таможенной статистики по экспорту и импорту в регионах РФ за 2018 год

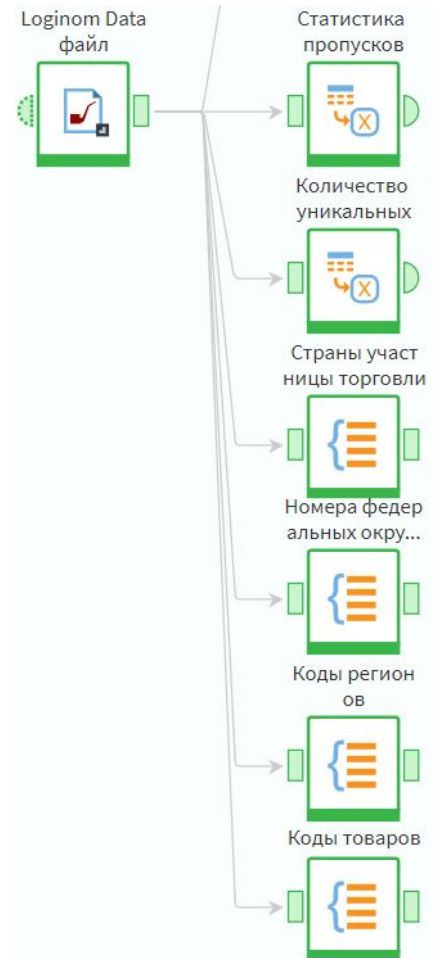
- Сколько строк имеет выборка данных за 2018 год?
- Какое количество столбцов имеет данная выборка?
- Имеются ли в данных пропуски? На основе компонентов **таблица в переменные, группировка, визуализатор статистика**, ответьте на вопрос о количестве пропущенных значений в каждом из столбцов выборки данных (можно использовать что-то одно).
- Сколько уникальных стран участников торговли перечислено в поле **country**?
- Сколько уникальных регионов России участвовало в торговле (поле **region**)?
- Сформируйте вывод о динамике качества данных собираемой статистики во времени и перечислите поля, дальнейший анализ по которым будет репрезентативен и валиден.

Задание на выполнение

Для импортированной выборки за 2016 год на основе компонентов трансформации данных отобразите:

- Перечень уникальных стран-участниц торговли в поле `country` с помощью компонента группировка (отсортируйте значения по возрастанию)
- Перечень уникальных регионов России участников торговли в поле `region` с помощью компонента группировка (отсортируйте значения по возрастанию)
- Перечень уникальных категорий товаров по ТН ВЭД, перевозимых через таможню в поле `code` с помощью компонента группировка (отсортируйте значения по возрастанию)
- Перечень уникальных кодов федеральных округов по полю `district` с помощью компонента группировка (отсортируйте значения по возрастанию)

Задание на выполнение



Задание на выполнение

Импортируйте в сценарий следующие выборки данных:

- Файл данных с кодами стран мира ОКСМ в формате alpha2 по пути:
Data/Alpha2/alpha2_codes.csv,
- в кодировке **UTF-8**,
- с разделителем столбцов **табуляцией**.
- Исключите из полей импортируемой таблицы: **fullname, english, alpha3, iso**.
- Источником данных классификатора выступил ресурс ОКСМ alpha2
<https://classifikators.ru/oksm>

Задание на выполнение

Импортируйте в сценарий следующие выборки данных:

- Файл данных с информацией о федеральных округах Data/ОКАТО/fed_codes.tsv.txt,
- в кодировке ANSI,
- с разделителем столбцов **табуляцией**.
- Измените тип поля **№** на строковый тип данных.
- Исключите из рассмотрения все поля кроме полей с метками: **№, Название ФО, Сокр., Кол-во субъектов, Админ. Центр**.
- Перечень федеральных округов <https://geostudy.ru/fo.html>

Задание на выполнение

Импортируйте в сценарий следующие выборки данных:

- Файл данных с информацией о регионах РФ в соответствии с ОКАТО
Data/ОКАТО/okato_regions.lgd,
- Исключите из рассмотрения все поля кроме полей с метками: **Код региона, Код района, Наименование территории.**
- ОКАТО <https://classifikators.ru/okato>

Задание на выполнение

Импортируйте в сценарий следующие выборки данных:

- Файл данных с информацией о разделах товарной номенклатуры ТН ВЭД
Data/ТН ВЭД/tnved_codes12.txt
- в кодировке ANSI,
- с разделителем столбцов **табуляцией**.
- Оставим остальные настройки импорта данных по умолчанию.

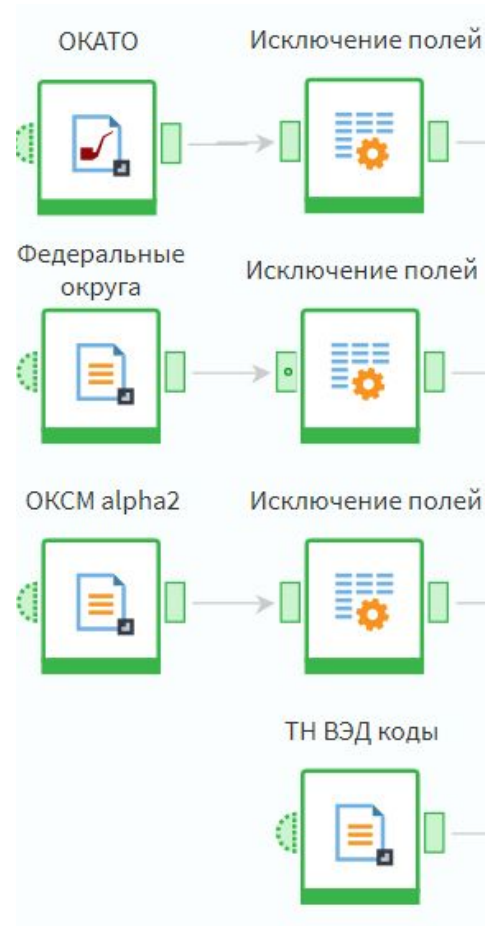
ТН ВЭД

<https://vdnk.ru/site/ru/tnved>

<https://vdnk.ru/site/ru/tnved/directory-tnved>

<https://vdnk.ru/site/ru/tnved/basis-classification-commodities>

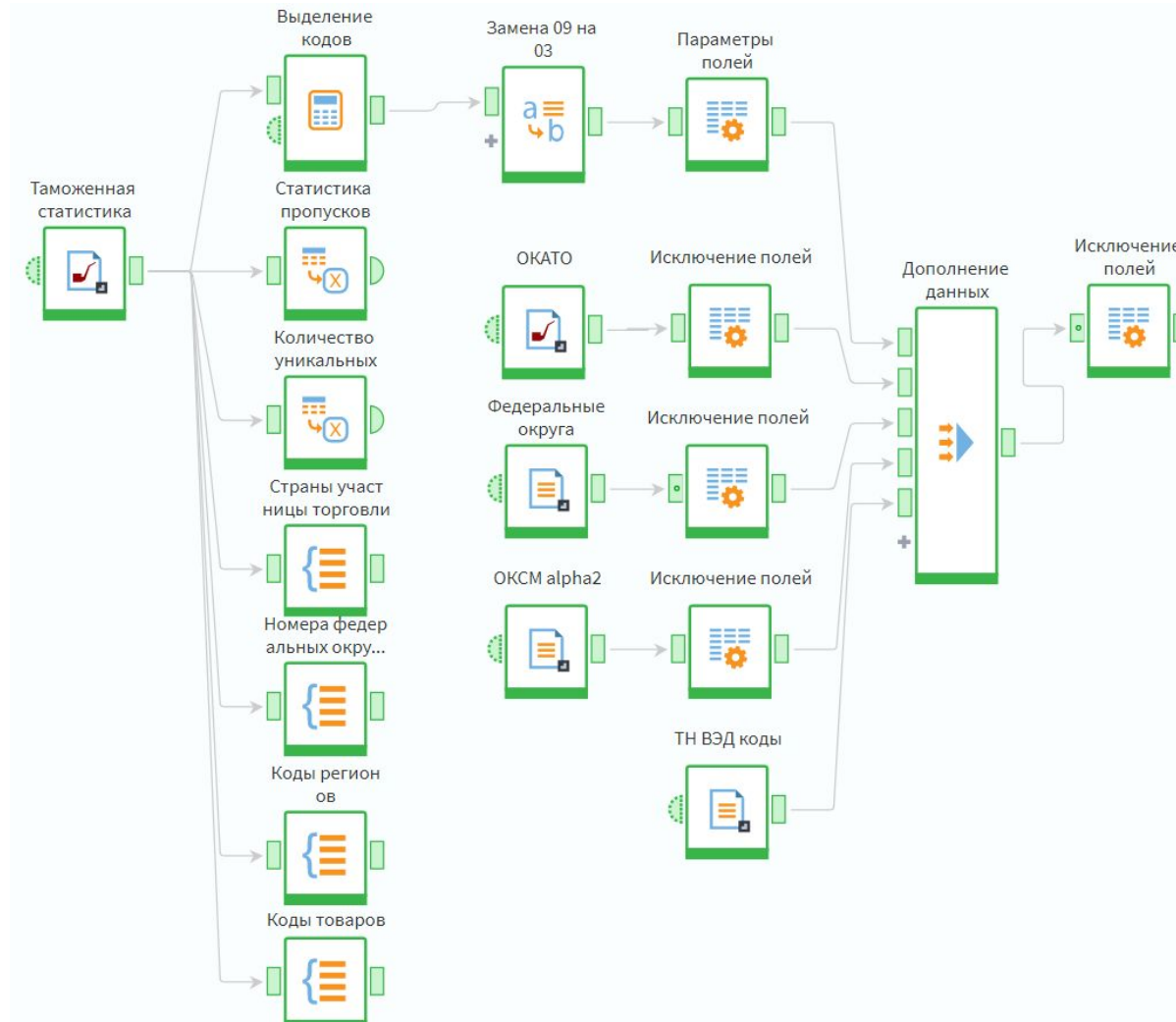
Задание на выполнение



Задание на выполнение

- В датасете в поле `district` присутствует федеральный округ с кодом «09», что есть девятый федеральный округ «Крымский ФО», упраздненный в 2016 году
- Замените в исходном наборе данных в поле с названием `district` значения «09» на «03» (Южный федеральный округ) с помощью компонента «Замена».
- С помощью компонента «Параметры полей» приведите выборку после замены к исходной структуре (количество полей и их названия).
- Получите из поля `region` *новые поля* с первыми двумя цифрами из `region` и последующими тремя цифрами из `region`, применив функции `Left()` и `Right()` из компонента калькулятор.
- **Дополните преобразованную выборку данных новыми полями из импортированных таблиц с помощью компонента «Дополнение данных»**
- **Исключите с помощью «Параметры полей» все промежуточные ключи (`district`, `region`, `country`, новые поля из калькулятора)**

Задание на выполнение



Задание на выполнение

- Сформируйте серию группировок данных на основе компонента «Группировка»:
- Какое количество уникальных кодов ТН ВЭД (поле **code**) и количество уникальных стран участниц торговли было задействовано в отчете об импорте и экспорте за 2016 код по: 1) каждому региону России, 2) Каждому Федеральному округу России
- Найдите самое популярное направление экспорта за рубеж по количеству уникальных позиций экспортируемых товаров за 2016 год. Считаем общее число уникальных кодов по всем регионам России
- Найдите самое популярное направление импорта в регион РФ из зарубежа. В какой регион поступает самый разнообразный спектр товаров
- Найдите десятку самых популярных пар **регион-страна** по количеству уникальных позиций в: 1) экспорте товаров, 2) импорте товаров.

Задание на выполнение

- Убедитесь в правильности построенного конвейера данных, указав в узле импорта данных таможенной статистики данных за отчетный 2021 год.
- В результате все процесс обработки данных должен выполняться правильно согласно логике обработки данных.
- Студенту необходимо сформировать отчет о выполнении задания, проиллюстрировав все основные стадии и ответы на поставленные вопросы в рамках самостоятельной работы.
- Задание считается выполненным после проверки работы преподавателем.

- В данной работе мы изучили базовые возможности обработки данных в Logiном и на примере освоили основные принципы построения low-code сценария обработки данных на основе базовых компонентов логики обработки табличных данных
- На примере самостоятельной работы были показаны эффективность и удобство работы с различными форматами табличных данных в рамках одного проекта.
- Для реализации проекта по обработке данных с использованием Logiном не требуется специальных технических знаний реализации конвейера обработки данных с помощью языков программирования, что резко понижает порог вхождения в обработку данных со стороны аналитиков.

Неосвещенные вопросы

- Кросс-таблица, скользящее окно
- Узел-ссылка, выполнение узла, подмодели, условие
- Блок исследования данных
- Блок предобработки данных
- Создание производных компонентов
- Импорт компонентов из внешних пакетов
- Блок data mining
- Соединение с внешними сервисами данных на основе REST и SOAP
- Визуализация данных с использованием отчетов

Вопросы на защиту работы

1. Какие данные помогает обрабатывать платформа Loginom?
2. В каком режиме обработки данных функционирует платформа Loginom?
3. Какие данные называются структурированными?
4. Как называются реализованные логические блоки обработки данных?
5. Как называются реализации компонентов обработки данных в Loginom, которые используются при построении сценария обработки данных?
6. Что представляет собой разработанный сценарий Loginom с точки зрения графовой структуры

Вопросы на защиту работы

1. Какие данные могут быть импортированы в Logiном и какие внешние системы хранения данных поддерживаются?
2. Какие правила существуют по именованию полей в таблицах и переменных Logiном?
3. Каким образом можно настроить компонент Logiном?
4. Можно ли параметризовать работу узла Logiном на основе переменных, и если да, то как?
5. Каким образом можно задать соединение с внешней базой данных в Logiном?