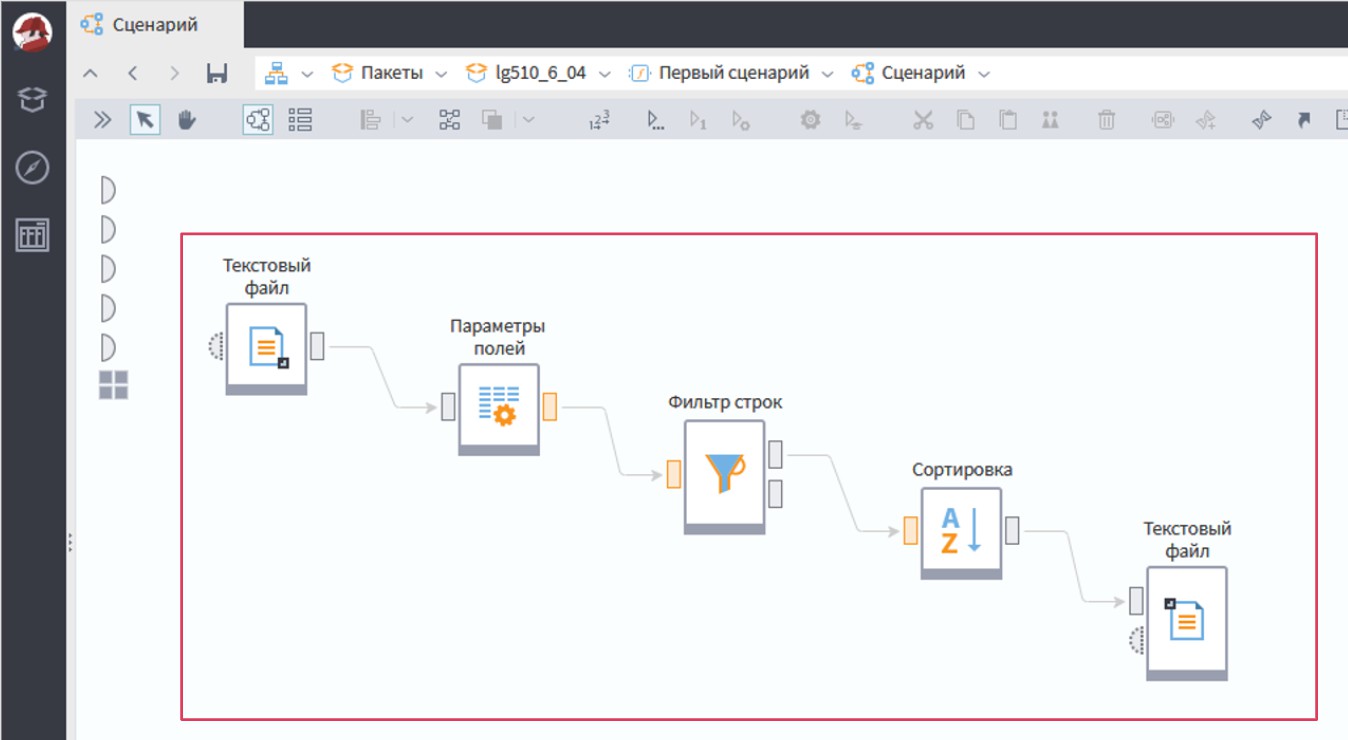
**Рабочая тетрадь №1. Часть 1**

**Первый сценарий**

**Сценарий** – набор узлов, связанных между собой, выстроенных в заданной последовательности, которая позволяет решить задачу. Каждый узел сценария – шаг алгоритма решения поставленной задачи.

# Типы сценария:

* импорт;
* обработка;
* экспорт.

# Текстовый файл (импорт)

Первый узел сценария – узел импорта. В сценарии может быть несколько узлов импорта, они могут быть разных видов, в зависимости от формата данных, который необходимо импортировать.

# Параметры полей, фильтр строк, сортировка

Далее идет группа узлов обработки данных. Это основная часть сценария, так как логически правильный набор узлов и грамотно построенная их последовательность влияют на корректное решение задачи.

# Текстовый файл (экспорт)

Последним в сценарии идёт узел экспорта, отвечающий за сохранение результатов обработки во внешний источник.

Любой из этих трёх видов действия может в сценарии отсутствовать. Например, если нужны только отчеты, которые удобно просмотреть в самом приложении – отсутствует узел экспорта. Если в расчетах используются только переменные – отсутствует узел импорта. А если необходимо только изменить формат данных – отсутствуют узлы обработки.

**Проектирование сценария** – описание режима работы, создание собственных компонентов, работа с портами и переменными.

**Loginom** позволяет разрабатывать сценарии «снизу вверх» и «сверху вниз». Это позволяет аналитику выбирать подход, который позволит построить лучшее решение для каждой конкретной задачи.

Создадим новый пакет и новый сценарий.

# Импорт данных

Существует 7 стандартных компонентов, предназначенных для создания узлов импорта

# 1С Запрос

Назначение компонента: импорт из базы данных 1С.

# База данных

Позволяет импортировать данные из баз данных различных типов.

# Текстовый файл

Назначение: импорт файлов в txt или csv-формате, в которых строки и столбцы разделены однотипными символами-разделителями.

# Deductor Warehouse

Собственный формат хранилища данных, специальная разработка для использования с **Loginom** и **Deductor** (предыдущее поколение аналитической платформы на базе **MS SQL**, **Oracle**, **Firebird**). Через компонент **Deductor Warehouse** осуществляется подключение к **Data Warehouse**, спроектированном на платформе **Deductor**. В **Loginom** возможно импортировать данные из хранилища данных **Deductor**

# Warehouse.

* **Excel файл**

Назначение: импорт **xls** или **xlsx**-файлов.

# Loginom Data файл

Назначение: импорт файлов в собственном формате (**\*.lgd**).

Данный формат разработан для работы в **Loginom** в целях обеспечения максимальной скорости импорта и экспорта данных. Скорость обмена информацией мало зависит от объёма данных, зависит только от скорости работы дисковой подсистемы, в отличии от других форматов данных. Размер файла в **lgd**-формате меньше, чем такой же набор данных другого формата, благодаря встроенным алгоритмам сжатия. Экспорт и импорт данных выполняется с минимальными настройками в связи с тем, что **lgd**-файл содержит в себе информацию о параметрах полей набора данных (метка, тип данных, имя поля и др.).

# XML файл

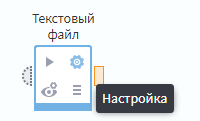
Назначение: импорт **xml**-файлов.

Необходимость предварительной настройки подключения делит

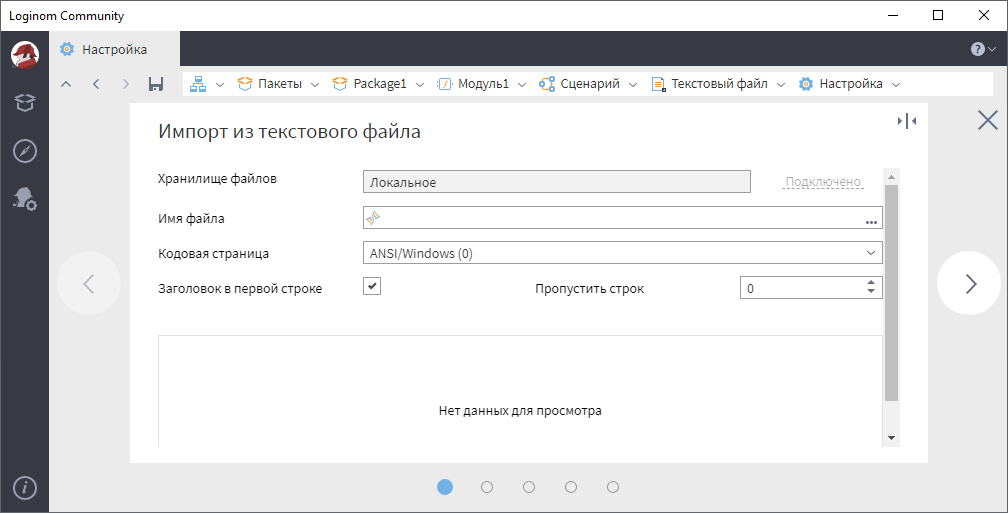
**Компоненты импорта** на *2 группы*:

* *Требуется* (1С Запрос, База данных, Deductor Warehouse, XML файл).
* *Не требуется* (Текстовый файл, Excel файл, Loginom Data файл).

# Практическое задание №1

Добавим в область построения компонент **Текстовый файл.** Создастся узел. Перейдём в настройку узла.

Появится окно настройки узла, в котором можно задать путь к файлу,

который необходимо импортировать (параметр **Имя файла**). Чтобы выбрать файл, необходимо нажать на многоточие справа.

В открывшемся окне выбираем файл **Потребление электрической энергии.txt** и нажмём кнопку **Открыть.**

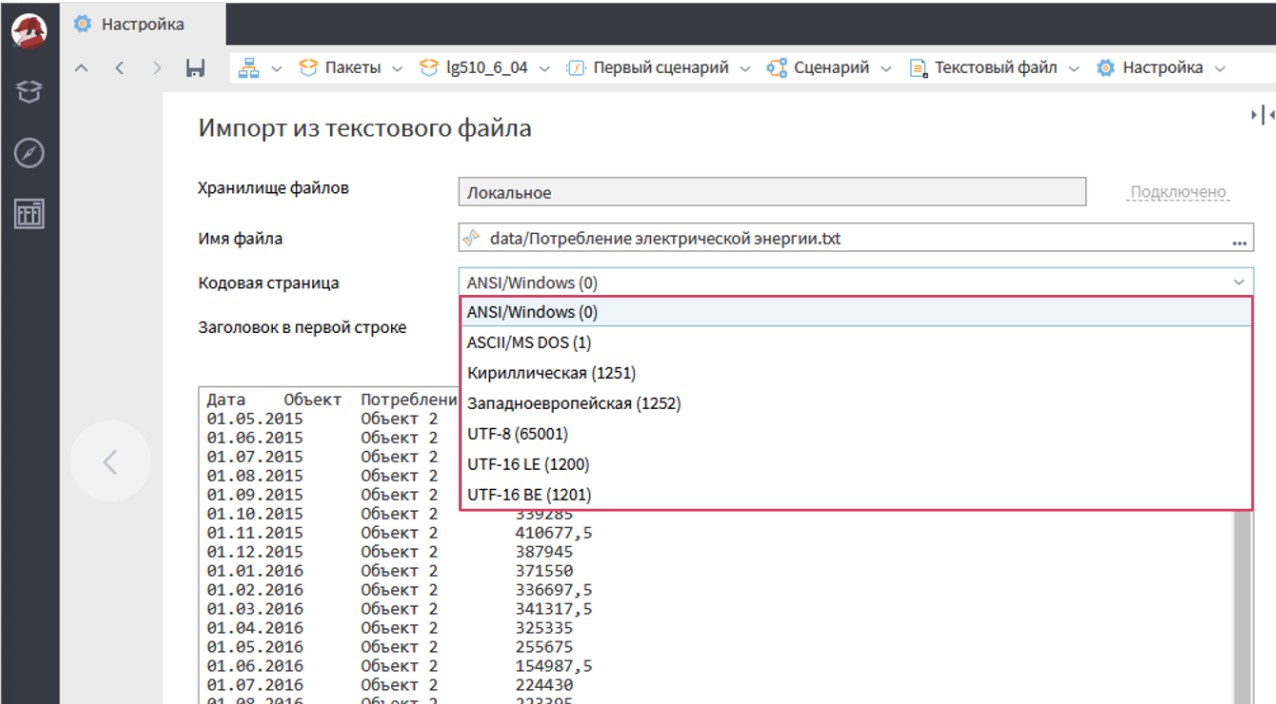
В **Loginom** есть 2 пути для файлов: абсолютный и относительный. **Абсолютный** путь – полный. **Относительный** – путь относительно текущей директории (файла пакета). Если набор данных, который необходимо загрузить и файл пакета находятся в одной папке, то при перемещении, переименовании папки или передаче пакета другому пользователю, не потребуется перенастройка узла импорта при использовании относительных путей.

Абсолютный путь в серверной версии указывается относительно каталога, используемого в качестве хранилища. В настольных редакциях – полный путь, начиная с имени диска, где хранится файл.

После выбора файла вернёмся в окно настройки.

Можно настроить параметр Кодовая страница. По умолчанию используется

# ANSI/Windows(0).



В следующих окнах оставим все настройки по умолчанию:

* в поле **Заголовок в первой строке** оставим галочку.

# Пропустить строк – 0.

Следующее окно – окно **Настройка форматов импорта**. Большинство настроек по умолчанию берётся из операционной системы. Оставим настройки без изменений.

Окно **Параметры импорта с разделителями** (всего 4 окна для текстового файла).

Здесь настраиваются параметры импортируемых полей.

В окне можно увидеть корректность настроек, выставленных в предыдущем окне оценив несколько первых записей из файла и выбрать символ-разделитель, используемый в файле. По умолчанию используется знак табуляции.

Если необходимо, чтобы несколько последовательных символов- разделителей воспринимались как один, необходимо поставить флажок в поле **Принимать несколько подряд разделителей за один**. Применять данный флажок нужно осторожно, так как несколько разделителей подряд могут означать пустые значения в полях.

В следующем окне данного назначения можно изменить тип данных (если при автоматическом распознавании типа данных загружаемых полей была допущена ошибка), выбрав правильный из выпадающего списка, который

появляется при щелчке на типе нужного поля. В записях **Поля, не соответствующих выбранному типу**, будет указано значение **null**. В поле **Кол-во строк для анализа** по умолчанию для текстового файла указано

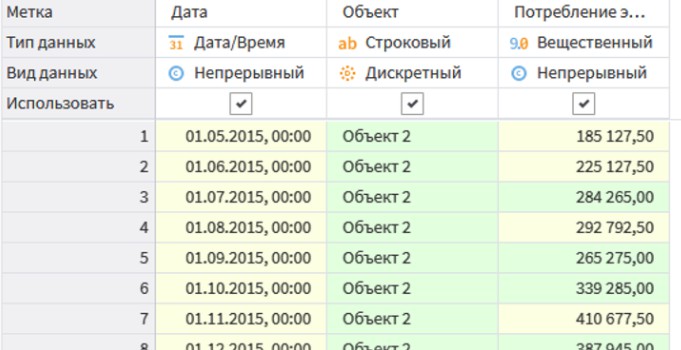
25 строк. Если необходимо увеличить количество строк, то данные нужно обновить. Если узел не определит тип данных, то будет указан неопределенный тип, а в записях Поля появится значение **null**.

При выборе типа данных происходит их конвертация. Для конвертации существует цветовая индикация:

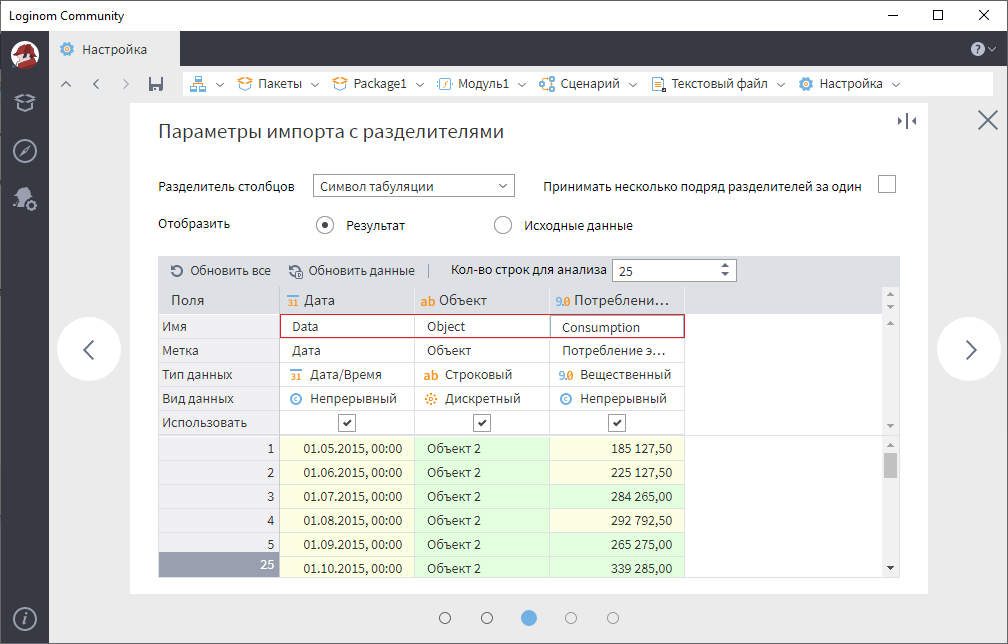
* + Красный – данные полностью утеряны, обычно запись имеет значение

null

* + Желтый – при конвертации возможно произошли потери данных
  + Зелёный – конвертация прошла успешно, без потерь.

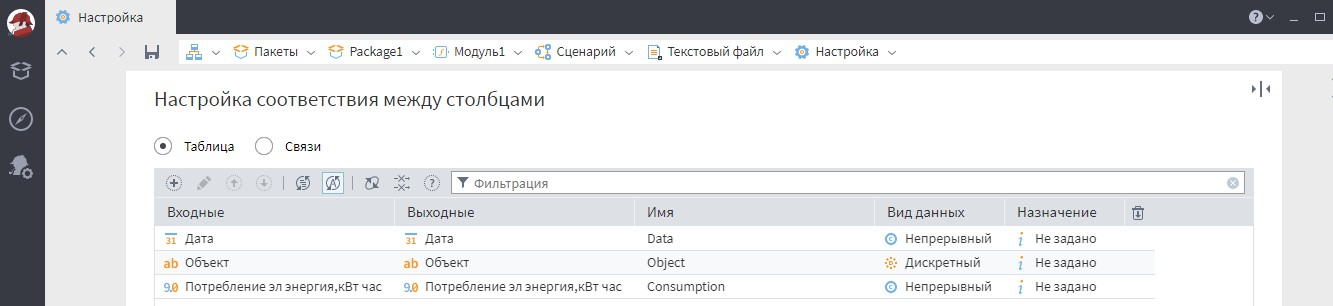


Здесь можно задать полям логичные имена.

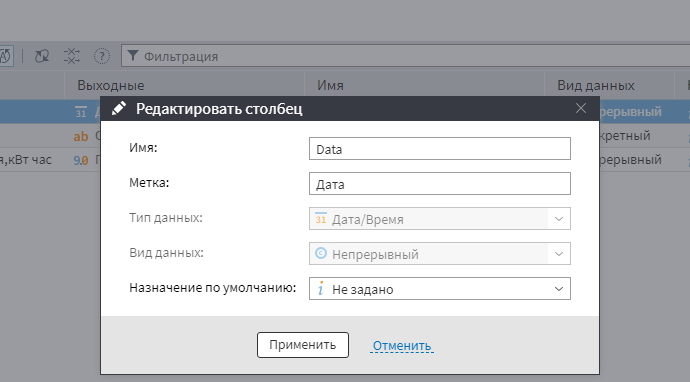


# Окно Настройка соответствия между столбцами.

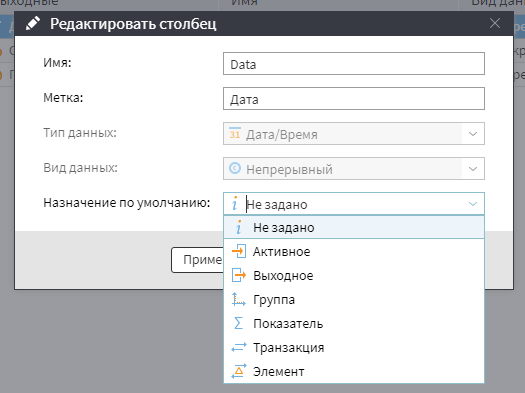
На данном шаге можно настроить соответствие между столбцами и ознакомиться с параметром поля **Назначение**, которое задаётся полям при настройке узлов или в настройках входных портов этих узлов. По умолчанию назначение при импорте не задано.



Откроем поля **Дата** (дважды щелкнув мышкой в любом месте, кроме Входные строки Дата). Откроется окно **Редактировать столбец**, в котором можно поменять имя, метку столбца и выбрать назначение в графе **Назначение по умолчанию**.



Существуют следующие варианты значений параметра **Назначение**, которые задают функциональность столбцов и зависят от узла обработки.



Вернёмся в мастер настройки. Перейдем к следующему шагу. Окно **Описание узла**.

В данном окне можно задать метку узла и комментарий. Узел будет отображен в сценарии с названием, указанным в поле Метка. Рекомендация: задавайте осмысленные метки для удобной ориентации в сценарии.

В поле Комментарий можно описать какие функции выполняет данный узел в сценарии.

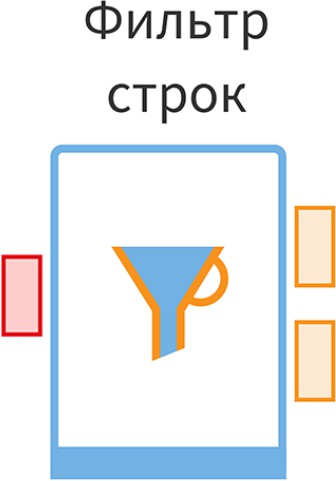
Это был последний шаг мастера. После всех настроек необходимо их сохранить. Нажать на кнопку сохранения в правом углу окна  .

# Компоненты фильтра строк и сортировки

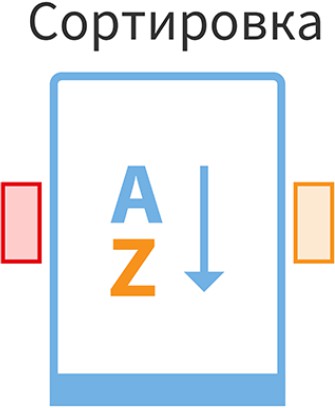
Раздел **Трансформация** – набор компонентов для первоначальной подготовки и простой обработки исходных наборов данных.

* Группировка
* Дата и время
* Дополнение данных
* Замена
* Калькулятор
* Кросс-таблица
* Объединение
* Параметры полей
* Разгруппировка
* Свёртка столбцов
* Скользящее окно
* Слияние
* Соединение
* Сортировка
* Фильтр строк

Рассмотрим 2 компонента: **Фильтр строк** и **Сортировка**.



Компонент **Фильтр строк** предназначен для выделения из набора данных записей, соответствующих заданным при настройке узла условиям. Записи разделяются на те, что удовлетворяют заданным условиям и те, что не удовлетворяют.

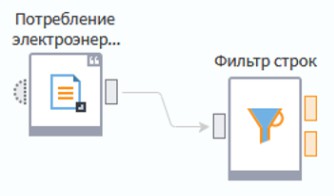


Компонент Сортировка предназначен для сортировки набора данных по возрастанию/убыванию значений по одному/ нескольким полям.

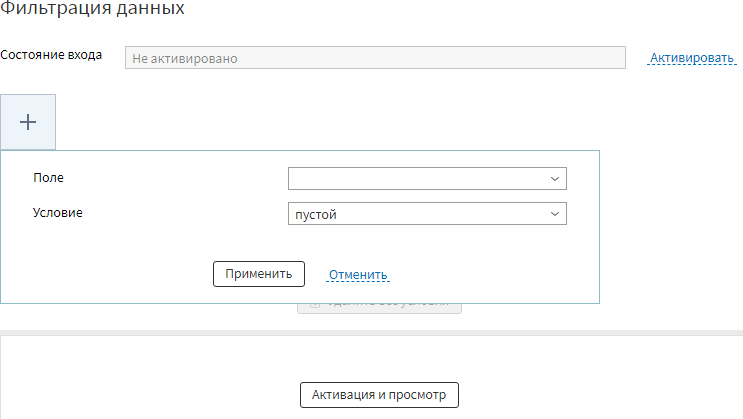
# Практическое задание №2

Добавим в область построения узел **Фильтр строк**.

Создадим связь между узлом импорта и узлом фильтр строк.

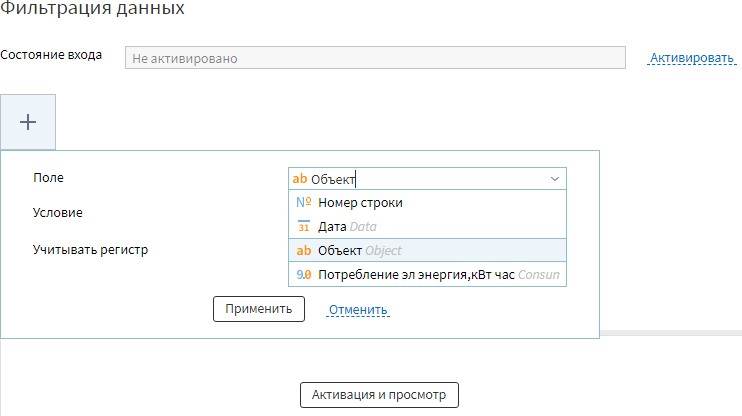


Перейдем в настройки узла.



Здесь предлагается задать условие фильтрации. Во вкладке **Поле** выберем пол, по которому будет проводиться фильтрация, например, **Объект**.

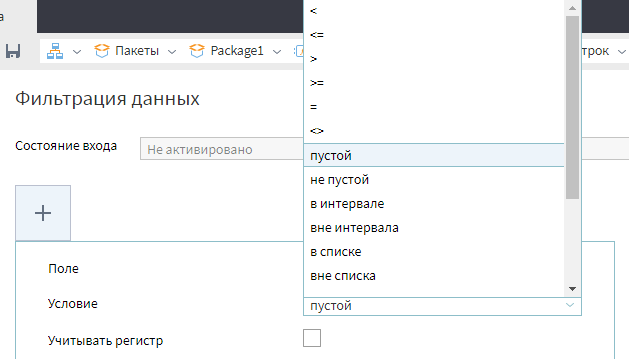
Одно поле можно использовать сразу в нескольких условиях.



Во вкладке **Условие** задаётся условие фильтрации. По умолчанию: пустой.

Для строковых полей доступен флажок **Учитывать регистр**.

Условия фильтрации:



* **<, <=, >, >=, =, <>**

Если выбрать одно из этих условий, появится вкладка **Значение для сравнения**. Будут отбираться те значения поля, участвующего в фильтре, которые удовлетворяют заданному условию.

# Пустой / не пустой

Отбираются записи, для которых в поле содержатся пустые/не пустые значения.

# В интервале / Вне интервала

Отбираются записи, значения которых лежат в указанном интервале/ вне указанного интервала. Используется для числовых полей, полей типа Дата/время.

При выборе данного условия появится 2 вкладки: **Нижняя граница** и

# Верхняя граница.

* **В списке / Вне списка**

Отбираются записи, которые содержатся в выбранном списке/вне выбранного списка.

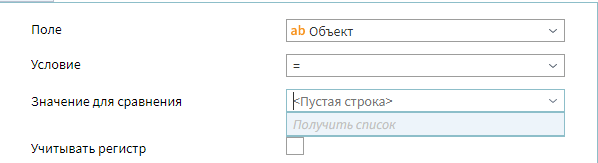
При выборе данного условия появится вкладка **Список**.

# Содержит/Не содержит

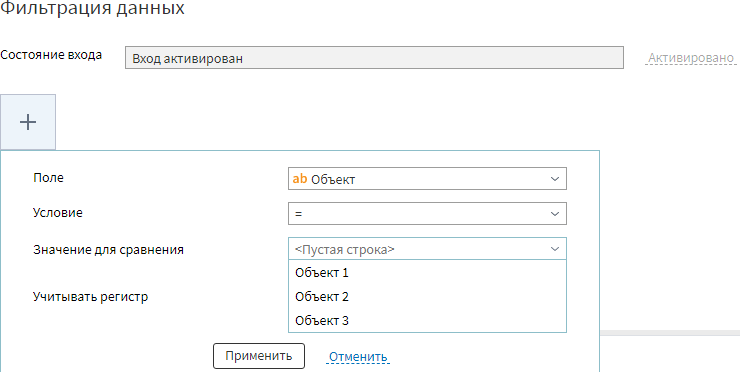
Это условие используется для строковых полей и позволяет отбирать записи, значения которых в выбранном поле содержат/не содержат указанную последовательность символов.

# Практическое задание №3

Выберите условие равно =.

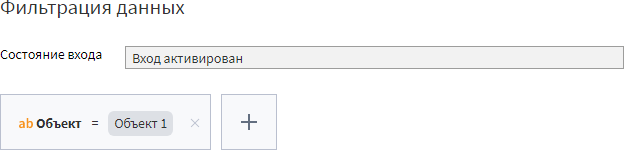
Значение для сравнения можно внести в ручную, а можно получить список всех значений столбца нажатием на **Получить список**.

После того, как список получен, состояние входа меняется на **Активировано**, что означает успешное выполнение всех предшествующих узлов сценария, т.е. набор данных подан на вход узла Фильтр строк.



Выберем Объект 1 и нажмём кнопку **Применить**.

Заданное нами условие сохранилось.

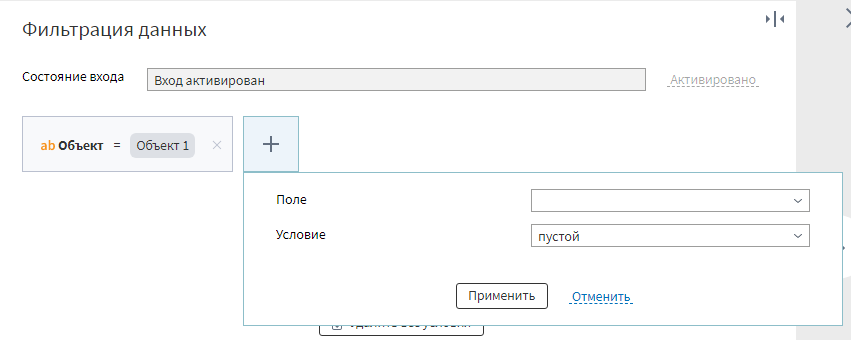


Применяя фильтр в режиме предпросмотра, обработается первые 1000 строк исходного набора данных. Если среди них нет записей, которые удовлетворяют условиям фильтрации, то появится предупреждение

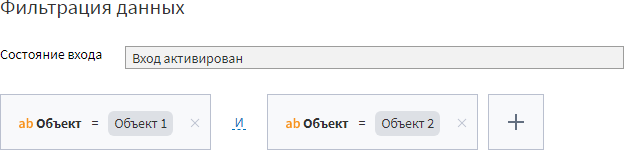
# «Достигнуто максимальное количество строк сканирования: 1000».

Можно добавить другие условия с помощью соответствующей кнопки +,

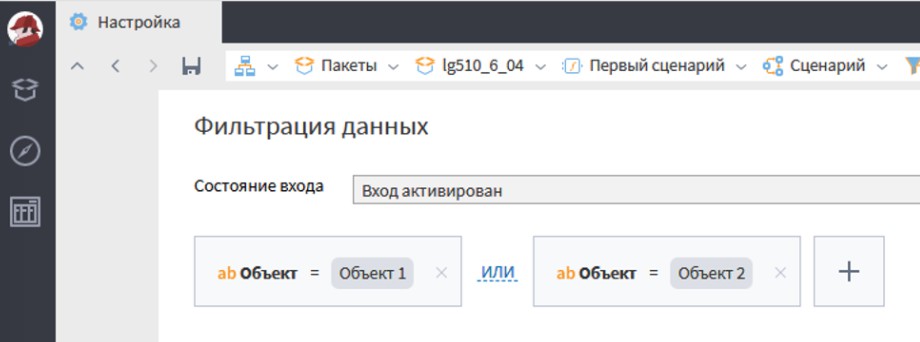
расположенной справа от уже созданного условия фильтрации.



Создадим такое же условие фильтрации, но в качестве значения для сравнения возьмём **Объект 2**.



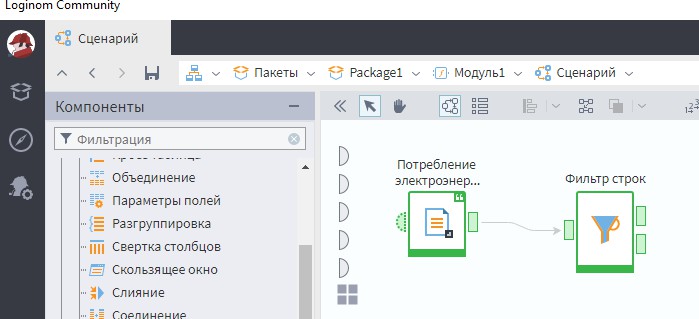
По умолчанию между условиями фильтра установится оператор **И**. Чтобы изменить его на оператор **ИЛИ** необходимо щёлкнуть по нему левой кнопкой мыши.



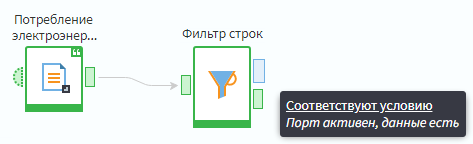
Сохраним изменения. Перейдём к просмотру результатов.

Запустим обработку, нажав на **Активировать узел** в узле **Фильтр строк**.

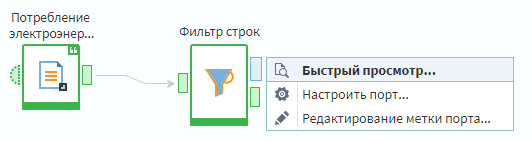
Быстрый просмотр.

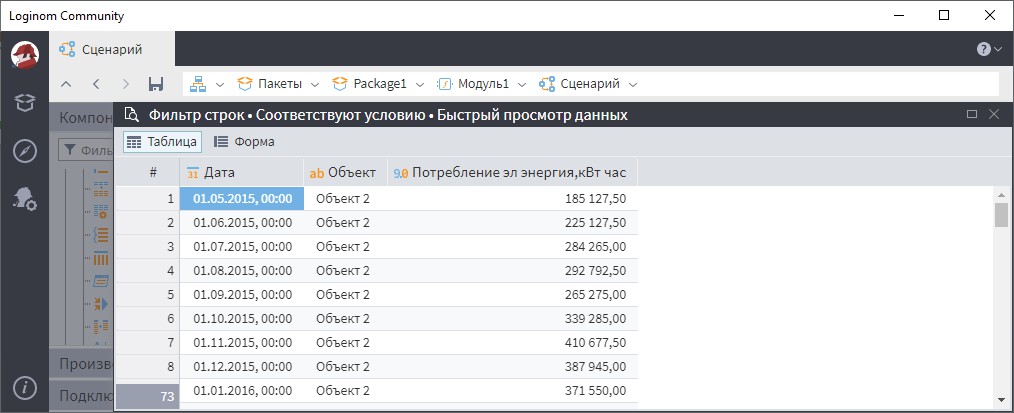


Проверим соответствует ли условию набор данных на выходном порте.



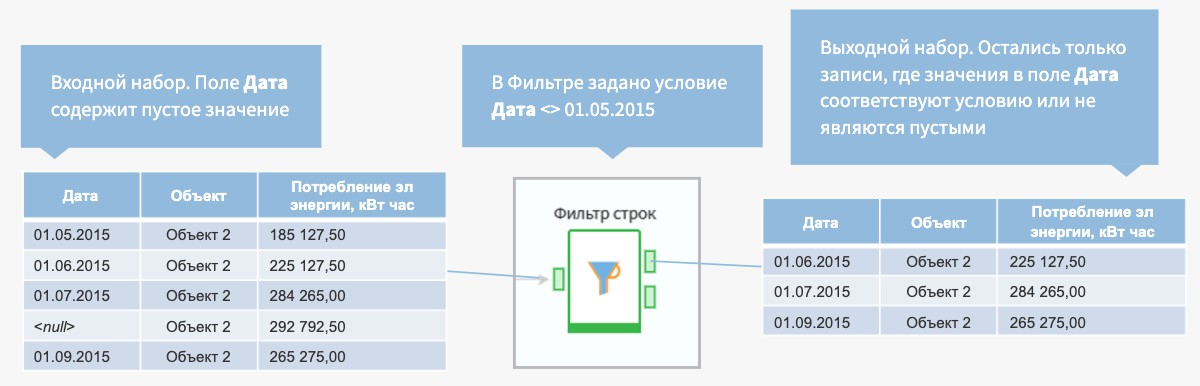
Чтобы просмотреть набор данных, удовлетворяющий заданному фильтру, необходимо щелкнуть левой кнопкой мыши на соответствующем порте и выбрать команду **Быстрый просмотр**.



Появится набор данных, состоящий из 73 строк, содержащий те записи, у которых в поле Объект было указано значение Объект 1 или Объект 2. Остальные записи можно посмотреть на втором порте фильтра.

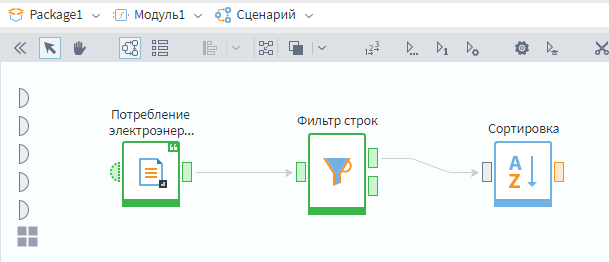
# Пустые значения в фильтре

Если в фильтруемом поле присутствует значение **null**, в выходной набор **Соответствует условию** не попадут записи с пустыми значениями в данном поле если выбрать в качестве условия **<>** или **не содержит**.

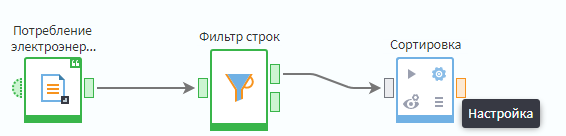


# Практическое задание №4

Добавим к сценарию компонент Сортировка, позволяющий отсортировать

набор данных последовательно по нескольким полям. Свяжем узел **Фильтр строк** с узлом **Сортировка**.

Перейдём в настройки узла **Сортировка**.

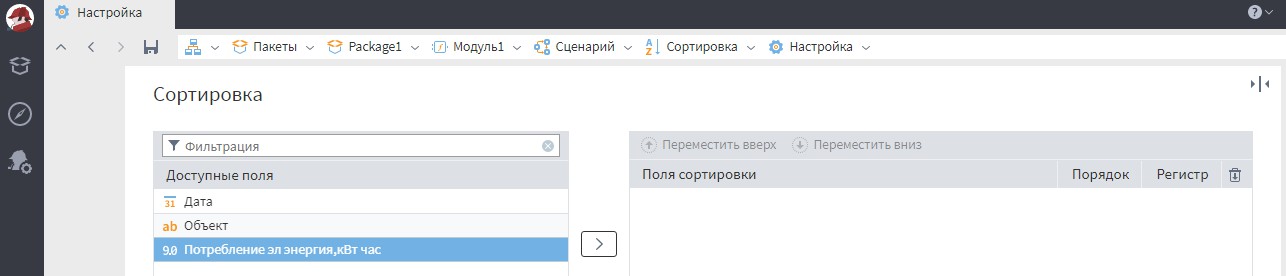


В левой части окна настройки находятся поля входного набора данных. Чтобы отсортировать значения поля, необходимо перенести его в правую часть окна одним из 4-х способов:

* Двойным щелчком по полю левой кнопки мыши;
* Используя механизм Drag&Drop;
* Нажатием кнопки Добавить;
* Используя команду Добавить контекстного меню.

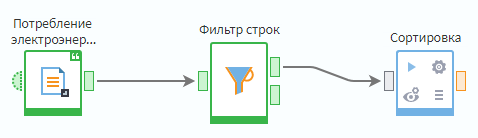
# Практическое задание №5

Выделим поле **Потребление эл энергии** и нажмём кнопку **Добавить**.

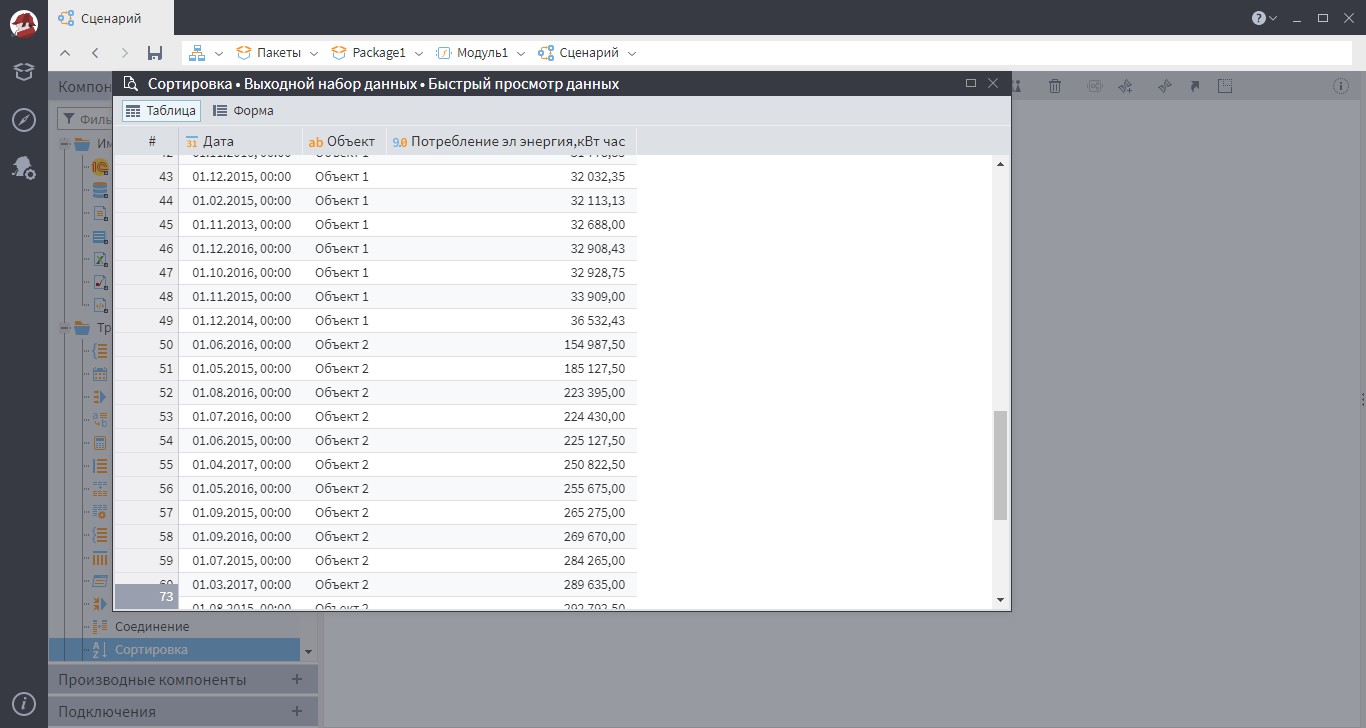


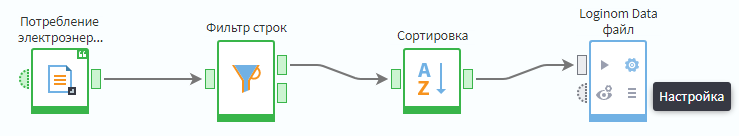
Записи сортируются алгоритмом сортировки в соответствии с заданным порядком по первому в очереди полю. Если в первом поле существуют одинаковые значения, то содержащие их строки сортируются по второму в очереди полю сортировки и т.д. согласно очередности полей сортировки.

Сохраним изменения. Выполним узел сортировка.



Посмотрим результирующий набор данных после двух узлов. Набор данных отсортирован в порядке возрастания потребления электроэнергии.



Создадим узел экспорта в Loginom Data файл. При экспорте по умолчанию выгружаются все поля. Если же необходимо выгрузить какие-то конкретные поля, это можно настроить с помощью настройки входного поля.

Откроется окно настройки **Экспорт в Loginom Data файл**.

В поле **Имя файла** можно выбрать папку для сохранения файла или сам файл, используя многоточие в строке. При экспорте можно сохранить данные в существующий или создать новый файл.

Поле Сжатие используется для уменьшения размера файла, если это необходимо. По умолчанию стоит значение **LZO** (рекомендуется). Но можно выбрать второй доступный вариант: **Без сжатия**.

Укажем в имени файла значение, выполним экспорт. Создастся новый файл с заданным именем в той же папке, в которой сохранён файл пакета.