# **Лекция 2 АиВД – Пакеты в Loginom**

## **2 Назначение, структура и создание пакета**

### **2.1 Определение и структура пакета**

**Реализация задачи анализа данных в системе Loginom принято (как и в других системах) называть проектом. Проект по анализу данных начинается с создания пакета. Поэтому последовательно проведём рассмотрение соответствующих понятий.**

**Пакет** – **важное понятие платформы** **Loginom**. Все действия с проектом **в** **Loginom** осуществляются в рамках **Пакета**, который является **минимальной единицей проекта** и представляет собой контейнер для компонентов, сценариев, подключений и т.д.

**Пакеты** сохраняются по-отдельности в виде файлов с расширением **\*.lgp**, и включают в себя **Ссылки** и **Модули**.

**Ссылки** применяются **для подключения других пакетов** с целью использования созданных в них **производных компонентов** и **подключений** в текущем проекте. Соответствующие объекты доступны только в том случае, когда они опубликованы для общего доступа. **Каждый пакет содержит хотя бы один модуль. Модуль включает в себя**:

* [Сценарий](https://help.loginom.ru/userguide/quick-start/first-workflow.html) — содержит последовательность узлов обработки данных.
* [Подключения](https://help.loginom.ru/userguide/integration/connections/) — в них представлен список внешних источников и приемников данных, к которым можно подключиться.
* [Компоненты](https://help.loginom.ru/userguide/processors/#standartnye-komponenty) — доступные для работы подмодели, как созданные в рамках текущего пакета, так и заимствованные из других пакетов через ссылки.

Примечание: при открытии любого пакета всегда создается файл с именем **<Название\_пакета>. lgp.lck**. **Этот файл защищает открытый пакет от возможности редактировать или удалить его другими пользователями**. После закрытия пакета этот файл удаляется.

### **2.2 Определение и назначение сценария**

***Сценарий***— последовательность действий, которые необходимо провести для анализа данных.

Сценарий обработки представляет собой комбинацию **узлов обработки** **данных**, настраиваемую пользователем для решения конкретной задачи. Вот как может выглядеть типичный и достаточно простой сценарий (рисунок 2.1).

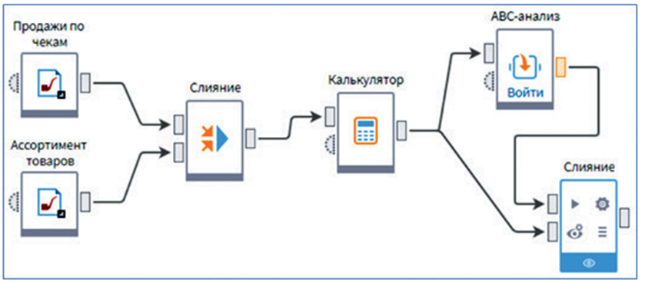


Рисунок 2.1. Пример сценария обработки данных

Последовательность обработки задается соединением выхода предыдущего узла сценария со входом последующего. Входом и выходом узла являются **входные** и **выходные порты**.

#### **2.3 Определение, назначение и логика работы узла**

**Порт** — важнейший элемент узла, определяющий способ (вид) коммуникации между узлами. В зависимости от структуры передаваемых/принимаемых данных различают **следующие типы портов**, **каждый из которых обозначается иконкой соответствующей формы**:

* Табличные данные;
* Переменные;
*  Данные в иерархической древовидной форме;
*  Параметры подключения к источнику данных;
*  Структура данных не определена.

Кроме **формы, соответствующей определенному типу порта**, используются цветовая индикация состояния и обозначение необязательности порта, а также включения/отключения автосинхронизации.

Данные перед обработкой узлом попадают в его входной порт, где они могут быть предварительно модифицированы.

Таким образом, порты являются важнейшей частью узла, обеспечивающей приём и передачу данных, их преобразование в случае необходимости, а также информирующей аналитика о характеристиках данных. Узлы Сценария принимают данные через входные порты, а отдают через выходные. Входные порты на пиктограмме узла сценария располагаются слева, а выходные — справа (см. рисунок 2.2).

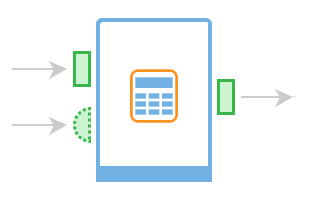


Рисунок 2.2. Узел сценария с двумя входными и одним выходным портом.

**Внутри** порта (см. рисунок 2.3) данные так же представлены в виде входного и выходного наборов полей или [переменных](https://help.loginom.ru/userguide/workflow/variables/), в зависимости от типа порта.

Входной набор представляет данные, которые порт принимает.

Выходной набор представляет данные, которые порт отдает. При настройке порта пользователь имеет возможность:

* задать произвольную структуру выходного набора — его перечень и параметры полей/переменных;
* сопоставить поля (или переменные) входного набора порта с полями выходного.

**2.3.1 Схема движения данных в узле**

Для узла, принимающего и отдающего набор данных (таблицу), схему движения данных можно представить следующим образом (см. рисунок 6):

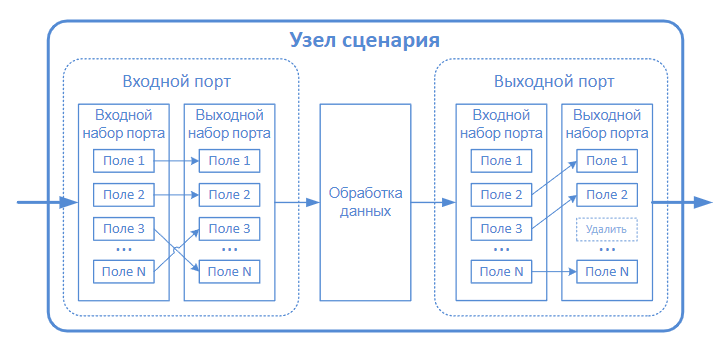


Рисунок 2.3. Схема движения данных для узла, принимающего и отдающего набор данных (таблицу).

Таким образом, для узла, схема движения данных которого представлена на рисунке 2.3, данные проходят следующие этапы преобразований:

1. **Изменение структуры данных во входном порту,** и передача их на обработку.
2. Обработка алгоритмом узла. В **результате обработки формируются новые данные**, которые поступают в выходной порт.
3. **Изменение структуры данных в выходном порту**.

**Внутри порта возможно произвести следующие изменения структуры данных**:

* изменение имен полей/переменных;
* изменение меток полей/переменных;
* задание порядка полей в таблице;
* удаление ненужных полей/переменных;
* задание назначения полей (см. [параметры поля набора данных](https://help.loginom.ru/userguide/data/datasetfieldfeatures.html)).

Эти действия, а также возможные **способы реагирования порта на изменение структуры данных**, поступающих в порт, задаются в настройках порта.

Одиночным кликом мыши по порту можно открыть диалог для последующего выбора одной из доступных команд:

* **Быстрый просмотр…** — открывает визуализатор [Быстрый просмотр](https://help.loginom.ru/userguide/visualization/preview/quick-view.html). Доступно только для выходных портов.
* **Настроить порт…** — открывает мастер настройки порта.
* **Редактирование метки порта…** — позволяет изменить текущую метку порта на любую другую.

Метка порта будет отображаться в Мастерах некоторых узлов, где **по логике возможен выбор входных наборов**. Она также будет отображаться и на панели выбора **визуализаторов для выходных портов** (но сами визуализаторы переименовывать нужно отдельно).

**2.3.2 Настройка порта**

**Порт можно настраивать**. Фактически это означает настройку входных или выходных данных в зависимости от типа порта.

Командой **«Настроить порт…»** открывается окно настройки входных или выходных данных в зависимости от типа порта.

В случае, если структура входных данных, подаваемых в порт, **еще неизвестна** (например, если к входному порту связь не установлена), то интерфейс настройки порта предоставляет возможность задать только структуру выходных данных порта.

**Если структура входных данных известна**, то в интерфейсе настройки порта можно задать структуру выходных данных порта и сопоставить поля (или переменные) входного набора порта с выходным. **В этом случае для сопоставления можно выбрать два различных интерфейса**:

[Связи](https://help.loginom.ru/userguide/workflow/ports/connections-interface.html) — графически отображающий связи между полями/переменными входного и выходного наборов порта;

[Таблица](https://help.loginom.ru/userguide/workflow/ports/table-interface.html) — сопоставление осуществляется путем выбора входных полей/переменных в таблице.

В случае, если структура входных данных, подаваемых в порт, еще неизвестна (**например, если к входному порту связь не установлена**), то интерфейс настройки порта предоставляет возможность **задать только структуру выходных данных порта**: не знаем, что на входе, но знаем, что будет на выходе по структуре.

Если структура входных данных известна, то в интерфейсе настройки порта можно задать структуру выходных данных порта и сопоставить поля (или переменные) входного набора порта с выходным. В этом случае для сопоставления можно выбрать два различных интерфейса:

* [Связи](https://help.loginom.ru/userguide/workflow/ports/connections-interface.html) — графически отображающий связи между полями/переменными входного и выходного наборов порта;
* [Таблица](https://help.loginom.ru/userguide/workflow/ports/table-interface.html) — сопоставление осуществляется путем выбора входных полей/переменных в таблице.

В Loginom существует механизм сокрытия [портов порядка выполнения](https://help.loginom.ru/userguide/workflow/ports/service-ports.html) и [портов управляющих переменных](https://help.loginom.ru/userguide/workflow/variables/control-variables.html), т.е. по умолчанию они скрыты от пользователя:

* Для отображения портов порядка выполнения необходимо воспользоваться кнопкой панели инструментов области построения сценариев;
* Для отображения портов управляющих переменных необходимо в контекстном меню узла выбрать пункт *«Показать порт управляющих переменных*».

**2.3.3 Цветовая индикация портов**

На рисунке 7 показана цветовая индикация на примере [*интерфейса порта Таблица*](https://help.loginom.ru/userguide/workflow/ports/table-interface.html).

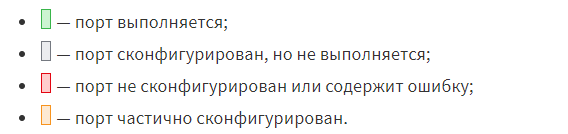


Рисунок 2.3. Цветовая индикация состояния порта

**2.3.4 Связывание портов**

Для связывания портов необходимо навести курсор мыши на порт и зажать левую кнопку мыши, после этого вести курсор к тому порту, с которым надо связать, и отпустить левую кнопку мыши.

При выделении порта появляется пунктирная линия с началом в исходном порту и концом в позиции курсора мыши, на его конце для удобства отображается тип порта, с которым можно связать исходный порт (см. рисунок 2.4):

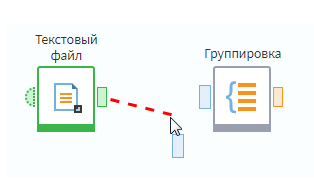


Рисунок 2.4. Соединение портов.

**Важно**: можно соединять только порты одного типа, то есть порт *Таблица* может быть соединен только с одноименным портом. Также нельзя соединять выходной порт узла с его же входным портом. Циклические связи запрещены!

**2.3.4 Определение и назначение подключений**

<https://help.loginom.ru/userguide/integration/connections/>

Перед тем как использовать в сценариях узлы [взаимодействия с SOAP-сервисами](https://help.loginom.ru/userguide/processors/integration/soap-request.html)/[REST-сервисами](https://help.loginom.ru/userguide/processors/integration/rest-request.html) и [импорта](https://help.loginom.ru/userguide/integration/import/)/[экспорта](https://help.loginom.ru/userguide/integration/export/) из бизнес приложений, хранилищ данных, баз данных или иных источников, в пакете необходимо создать подключение к источнику данных. Подключение содержит в себе все необходимые параметры для соединения с источником данных, такие как пароль/логин, расположение источника и др. Созданное подключение в дальнейшем может использоваться в пакете многократно.

**2.4 Создание пакета**

При выборе опции «Создать пакет» открывается диалоговое окно, предлагающее сохранить пакет в соответствующей папке с необходимым для Вас названием (рисунок 2.5).

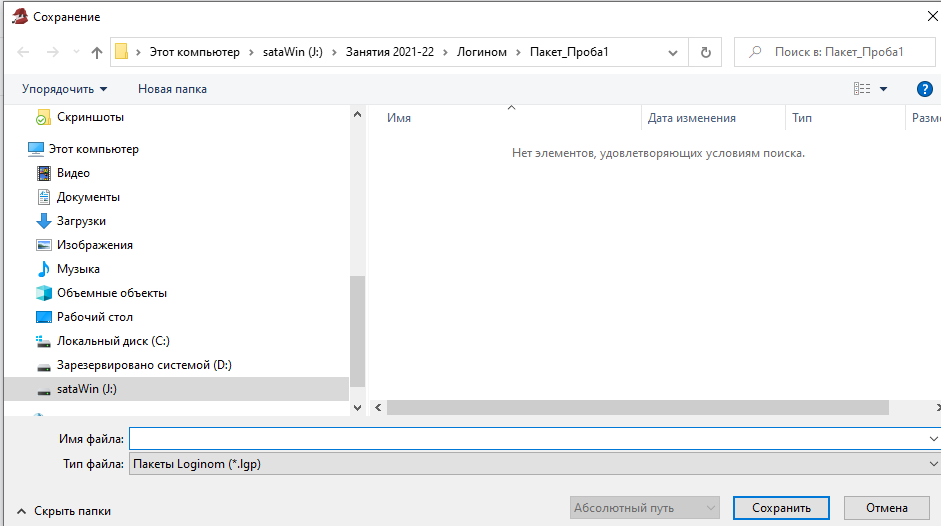


Рисунок 2.5. Предложение о создании пакета

Назовём пакет «Пакет\_Проба1» и сохраним его в одноимённой папке (Рисунок 2.6).

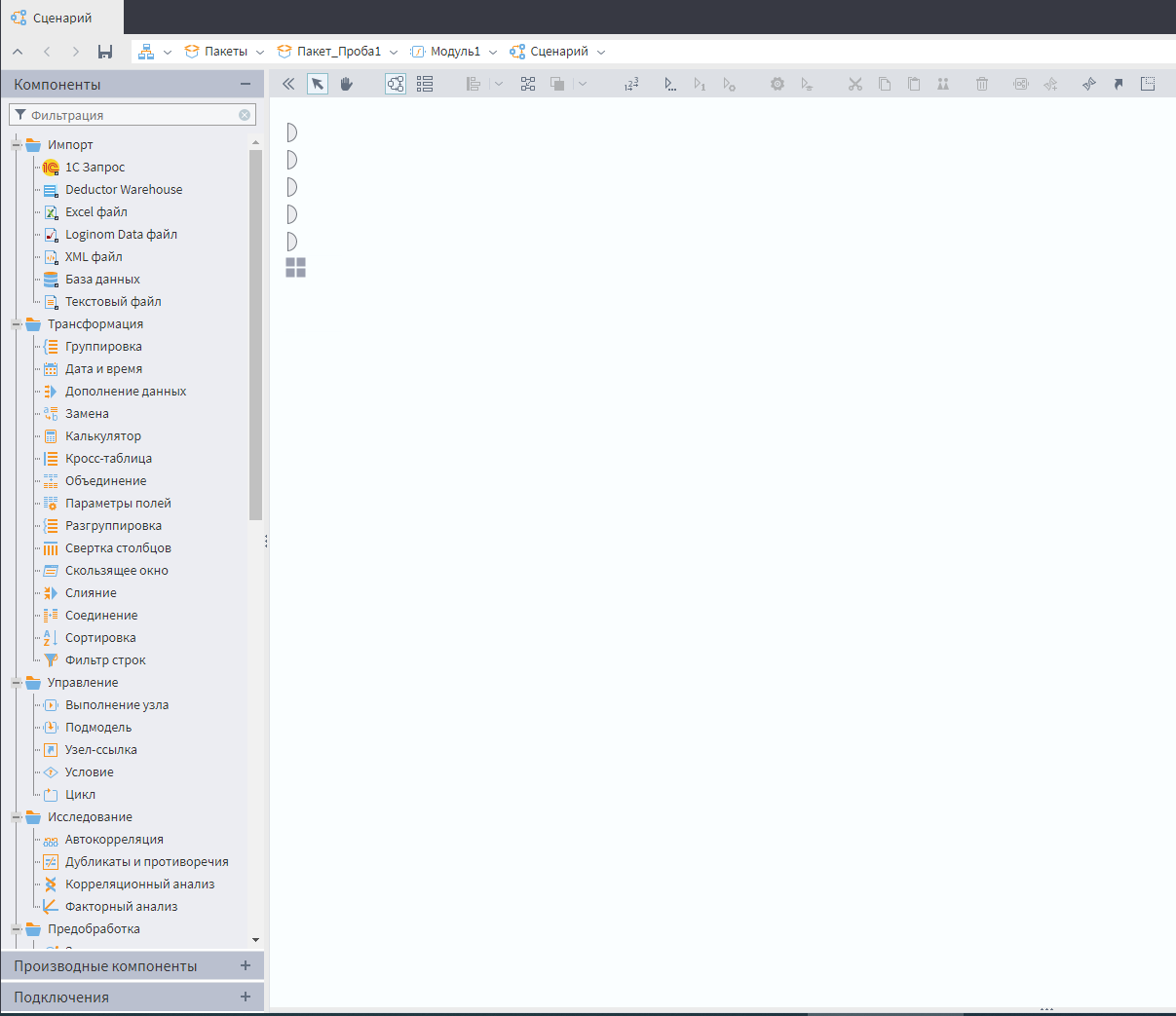


Рисунок 2.6. Открывающееся после создания пакета окно

Если, ничего на сделав во вновь созданном пакете, закончить работу с Loginom, и снова открыть его, увидим картину, отображённую на рисунке 2.7.

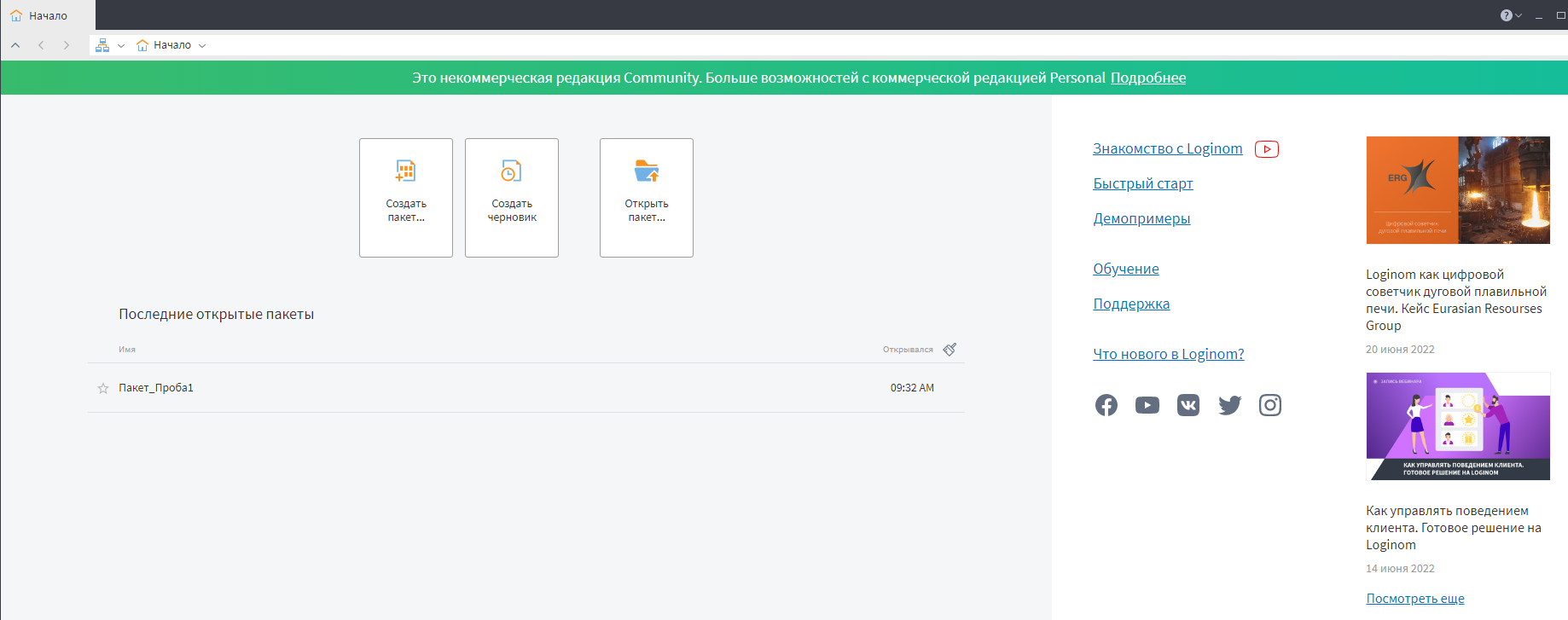


Рисунок 2.7. Окно открытия Loginom после создания и открытия пустого пакета

Щёлкаем по имени пакета и получаем картинку, полностью аналогичную рисунку 2.6.

Обратите внимание на крайнюю панель слева (на рисунке 2.6). На ней отображены возможные действия аналитика данных по их обработке. На рисунке 2.8 представлены все возможные опции. Они находятся в сворачиваемой в списке под общим названием «Компоненты», который сворачивается по клику по соответствующей кнопке (рисунок 2.9).

Компонент, назначенный для обработки данных,  является прообразом или шаблоном будущего узла сценария. Для того, чтобы создать узел сценария, выполняющий нужную операцию над данными, **необходимо мышью перенести соответствующий компонент из панели компонентов в область построения сценария**.

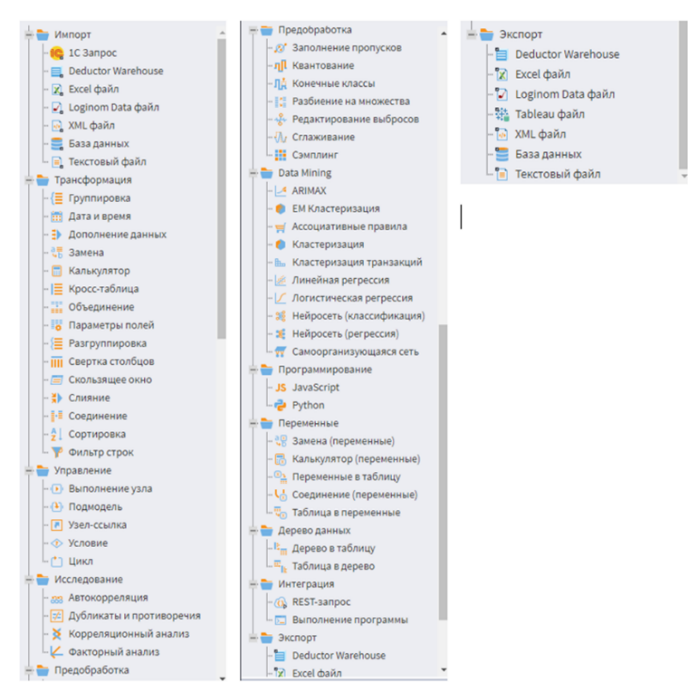


Рисунок 2.8. Компоненты работы с пакетом и его составляющими

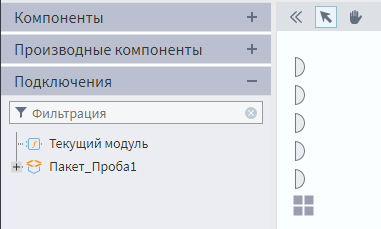


Рисунок 2.9. Вид свёрнутой панели «Компоненты»

# **2.5 Сценарий как основа работы при анализе данных**

Вся работа по анализу данных в Loginom Community базируется на выполнении следующих действий:

* проектирование сценариев; ([**https://help.loginom.ru/userguide/workflow/index.html**](https://help.loginom.ru/userguide/workflow/index.html));
* обработка данных;
* визуализация данных.

## **2.5.1 Создание первого сценария**

После создания нового пакета/черновика пользователь попадает на страницу [«Сценарий»](https://help.loginom.ru/userguide/interface/workflow.html), на которой увидит панель Компоненты (1), содержащую стандартные компоненты, предоставляемые платформой, и область построения сценария (2), в которую перетаскиваются (добавляются) компоненты для построения сценария (см. рисунок 2.10):

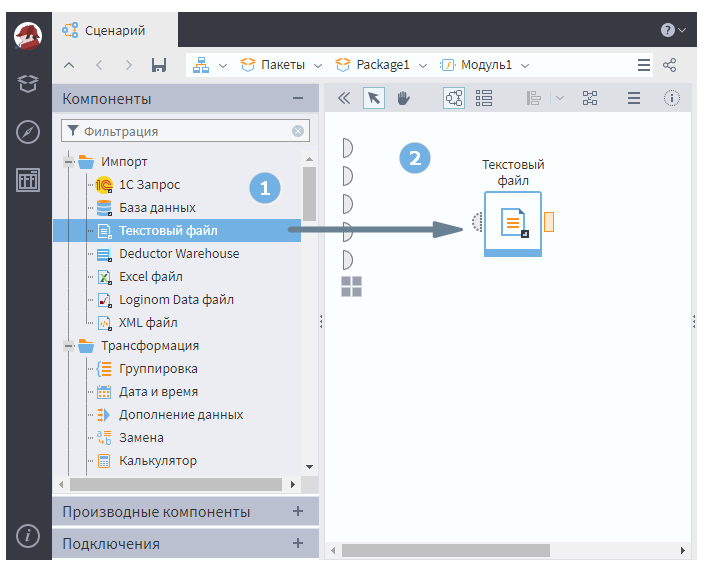


Рисунок 2.10.Часть страница сценария.

В простейшем случае сценарий импортирует данные из внешних источников, либо преобразовывает их и экспортирует (**выводит в отчет**).

### ***2.5.1.1 Создание простого сценария с использованием текстового файла примеров Loginom***

**Создадим простой сценарий, формирующий ТОП10 лучших клиентов**.

**Сценарий должен выполнить действия**:

* Импорт из файла [Sales.txt](https://help.loginom.ru/userguide/attach/Sales.txt) информации о продажах;
* Выделение 10 клиентов с наибольшими суммами покупок;
* Экспорт полученных результатов.

#### **Шаг 1. Помещение компонента в сценарий**

Для того чтобы использовать в сценарии какой-либо компонент, его необходимо перенести мышью из панели компонентов в область построения сценария.

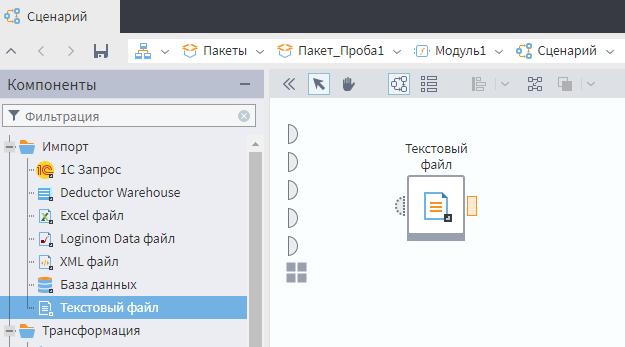


Рисунок 2.11. Внедрение узла в пакет «Пакет\_Проба1»

При этом создастся узел сценария, выполняющий действие импорта. При клике мышкой на узле отобразятся иконки возможных действий. Вызовем [Мастер настройки](https://help.loginom.ru/userguide/integration/import/txt-csv.html) (см. рисунок 2.12).

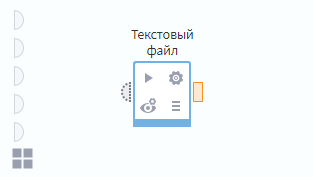


Рисунок 2.12. Вызов мастера настройки узла

При нажатии на значок шестерёнку появляется окно мастера настройки (рисунок 2.13).

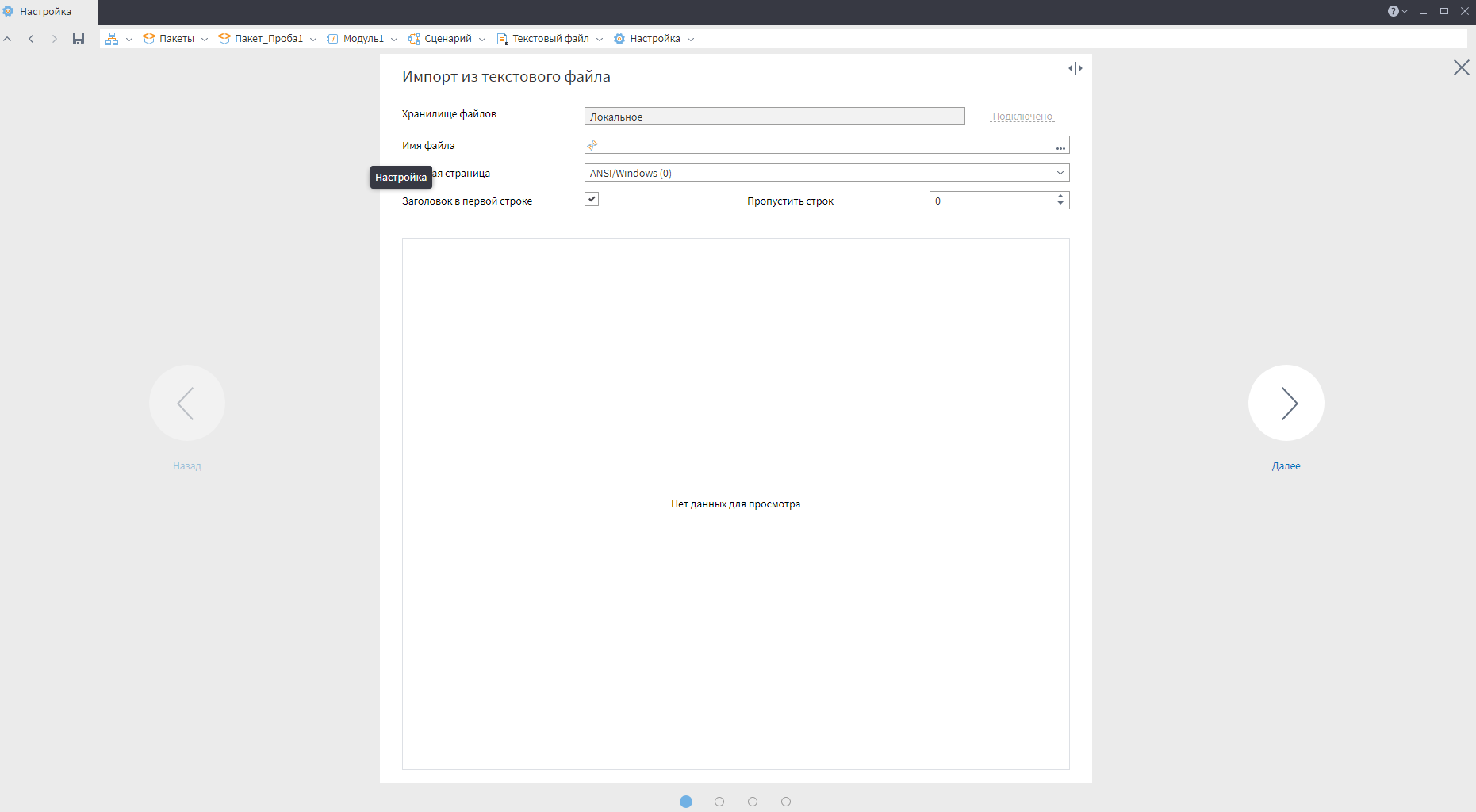


Рисунок 2.13. Окно настройки для указания места входного файла для обработки.

Таким образом, возникает вопрос о «добыче» нужного файла. Адрес нужного файла находится по адресу [**https://help.loginom.ru/userguide/attach/Sales.txt**](https://help.loginom.ru/userguide/attach/Sales.txt). На этой странице нужный файл предлагается к скачиванию, что и было сделано в соответствующую локальную папку. Этот адрес и надо указать в мастере настройки (рисунок 2.14а)[[1]](#footnote-1). При этом показывается только часть файла. С помощью кнопки в правом верхнем углу окна можно развернуть изображение таблицы, содержащейся в используемом файле. Тогда получим вид, показанный на рисунке 2.14.б.

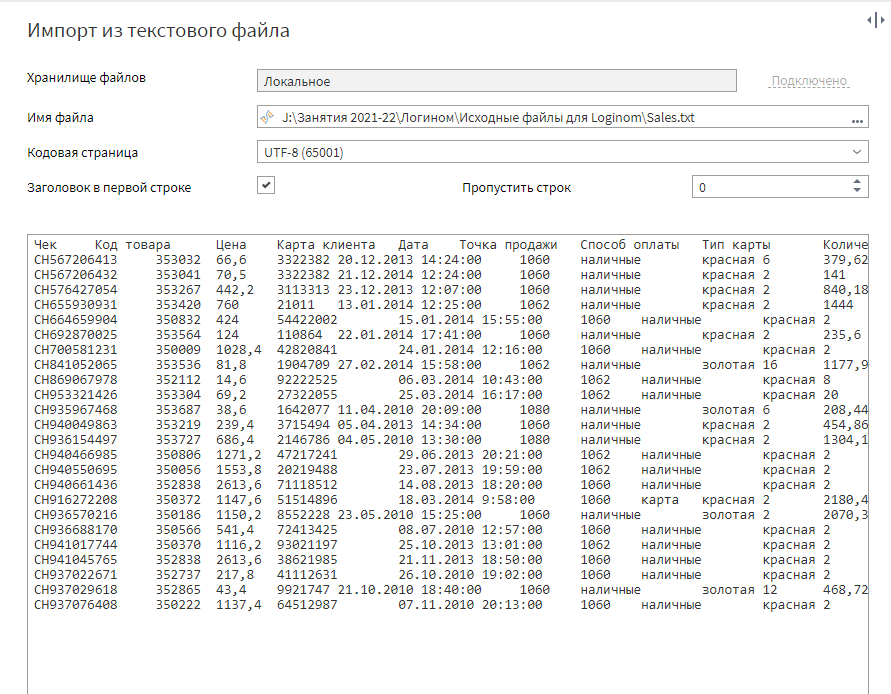


Рисунок 2.14а. Отображение файла при настройке входного узла после указания пути до файла.

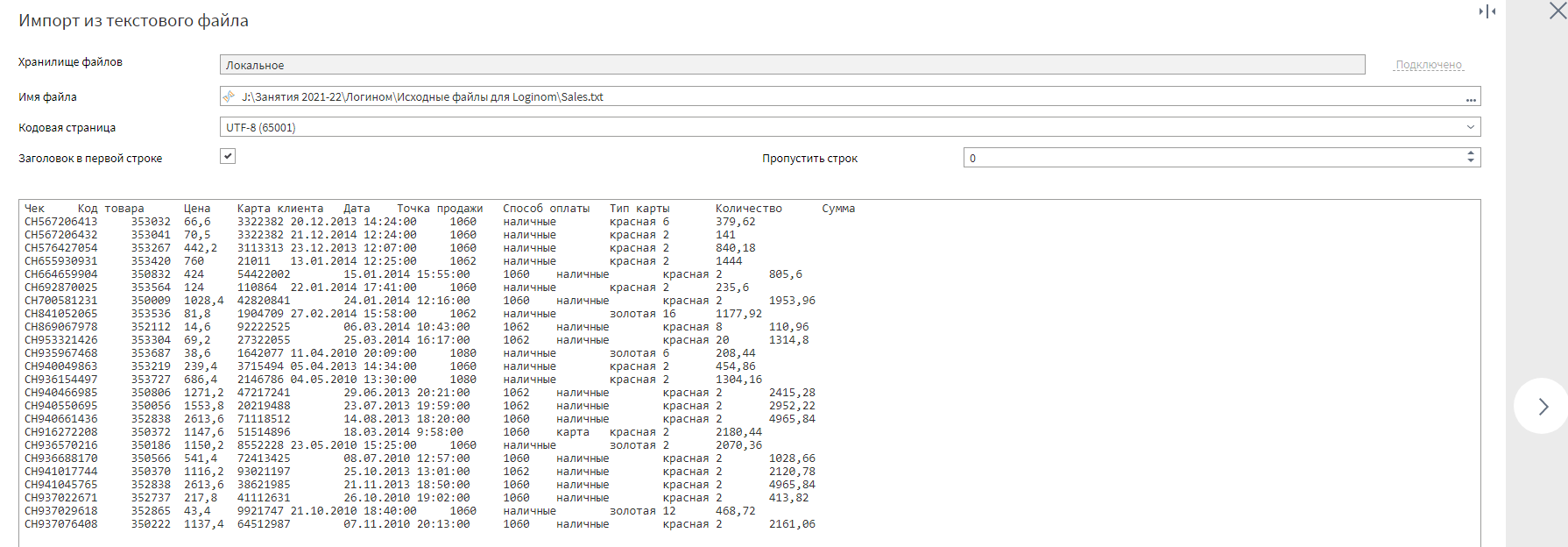
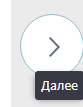


Рисунок 2.14б. Расширенное отображение файла при настройке входного узла.

Если после этого нажать на кнопку  увидим окно, позволяющую проводить настройку форматов отображения полей (рисунок 2.14в).

Если после этого снова нажать на кнопку «Далее» появится окно, показывающее окончательный вид таблицы (рисунок 2.14г) и позволяющее настроить связи между. В режиме «Связи» (рисунок 2.14д) можно каким-то образом устанавливать соответствие между столбцами, а в режиме «Таблица» можно добавлять новые колонки.

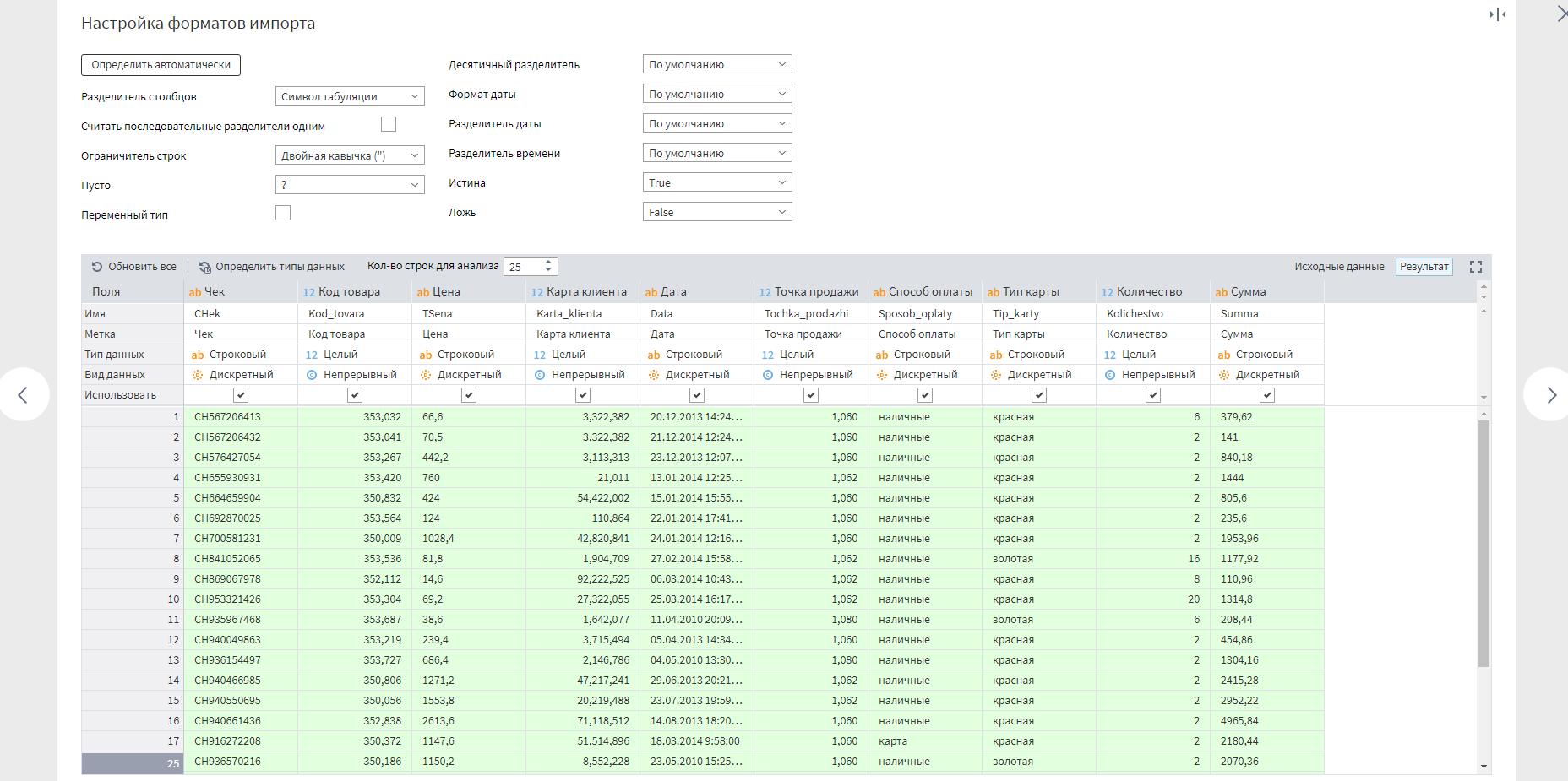


Рисунок 2.14в. Окно настройки форматов отображения полей.

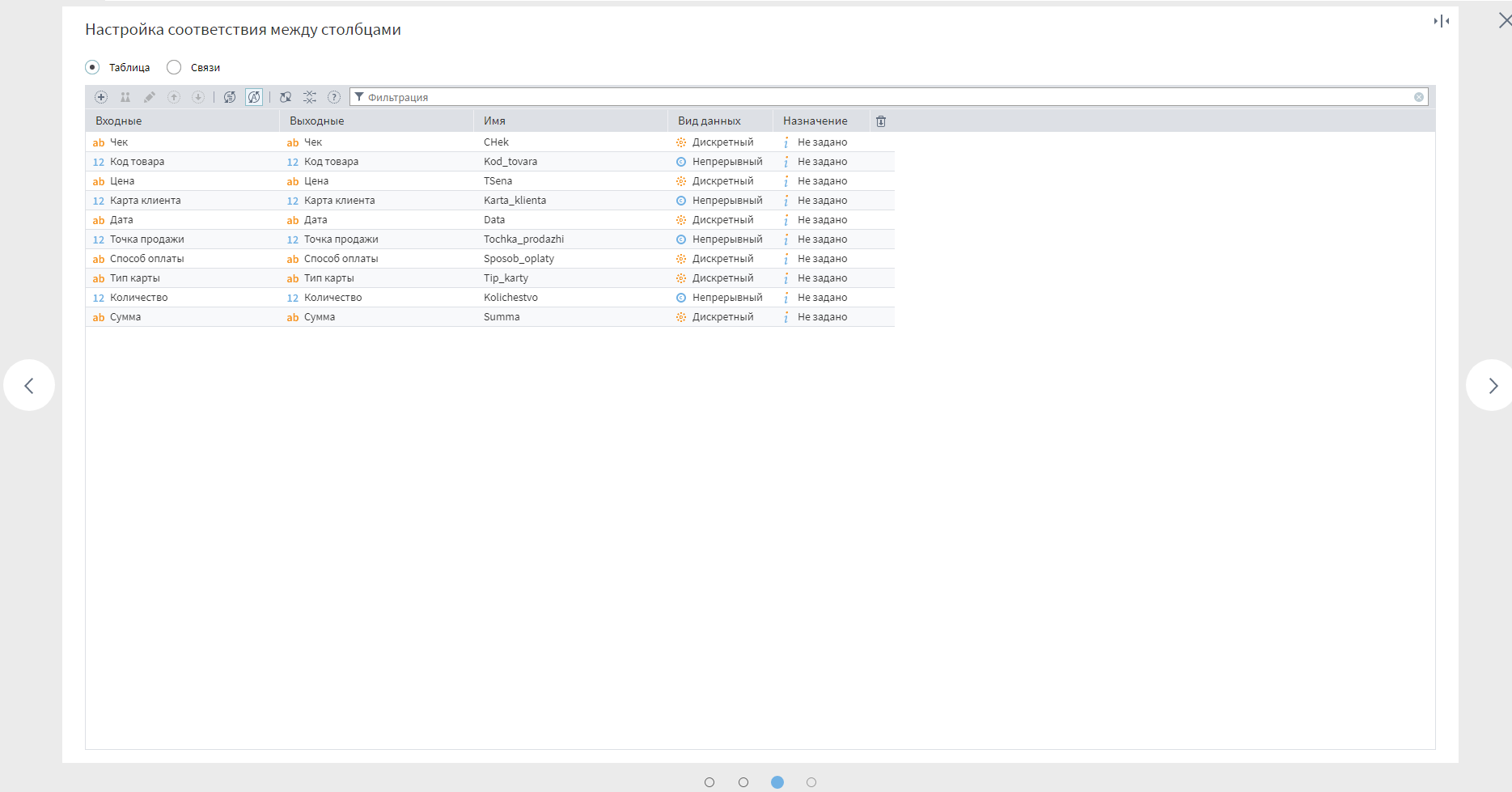


Рисунок 2.14г. Окно настройки соответствия между столбцами в опции «Таблица».

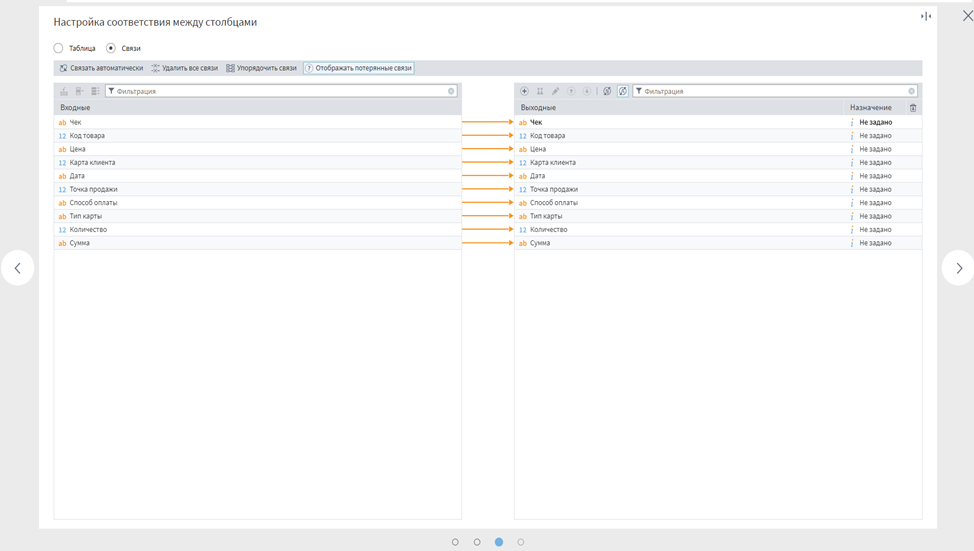


Рисунок 14.д. Окно настройки соответствия между столбцами в опции «Связи».

Следующий переход по «Далее» приводит к возможности создания комментария и сохранения результата (рисунок 2.14е).

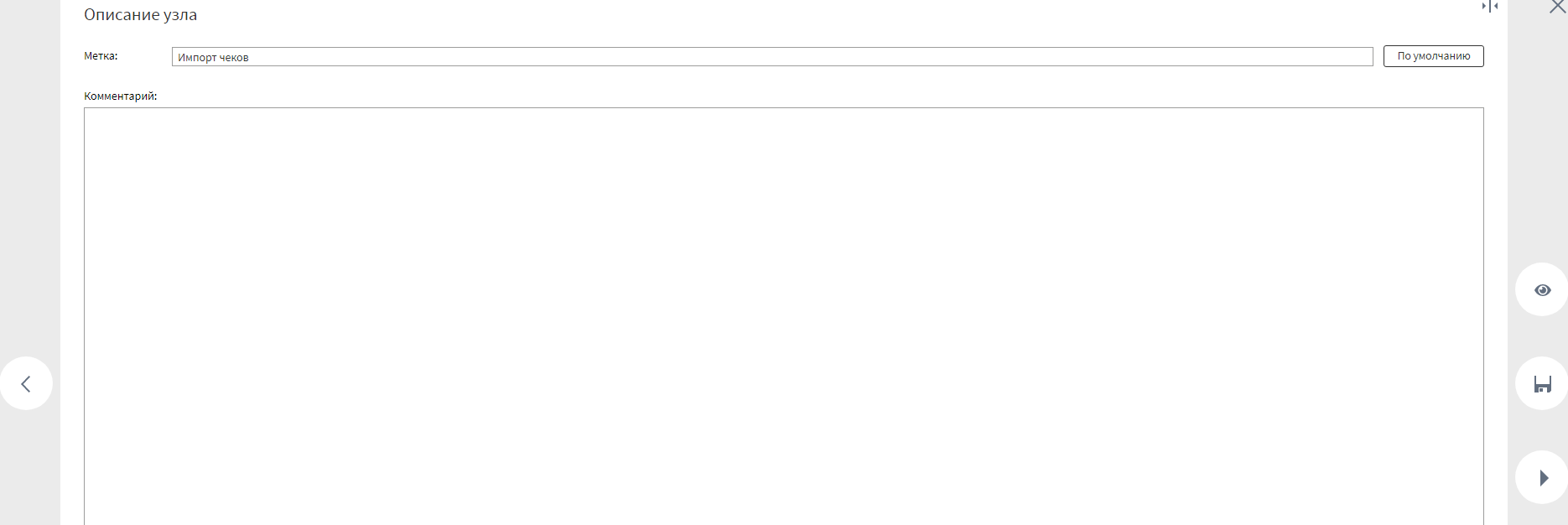


Рисунок 2.14е.

На заключительном этапе остаётся выбрать одну из трёх возможностей: визуализацию, сохранение или выполнение (кнопки справа на рисунке 2.14е). О визуализации будем говорить позже.

Удобнее всего нажать кнопку выполнения. При этом произойдёт выполнение всех отмеченных особенностей и переход на полотно сценария. **Если реализовать сохранение и выйти из системы, то для выполнения этих особенностей потребует активации узлов**. После этого надо активировать узел, как показано на рисунке 2.15.

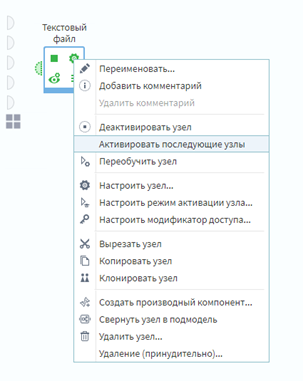


Рисунок 2.15. Активация узлов пакета

#### **Шаг 2. Выполнение узла**

После настройки узла выполним его, используя меню возможных действий. Теперь в выходном порте узла присутствуют импортированные данные, которые можно увидеть, выбрав [Быстрый просмотр…](https://help.loginom.ru/userguide/visualization/preview/quick-view.html) в контекстном меню порта (см. рисунок 2.16).

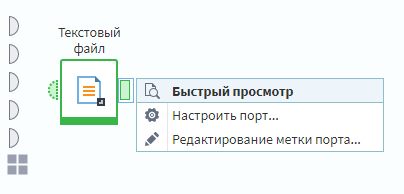


Рисунок 2.16. Возможность просмотра содержимого узла

Результат в режиме «Таблица» показан на рисунке 2.17. Результат просмотра в режиме «Форма» показан на рисунке 2.18. В этом режиме можно последовательно просматривать строки импортированной таблицы.

Как следует из контекстного меню можно изменить название (метку) узла. Переименуем узел с помощью контекстного меню узла (рисунок 2.19) и назовём его «Импорт чеков».

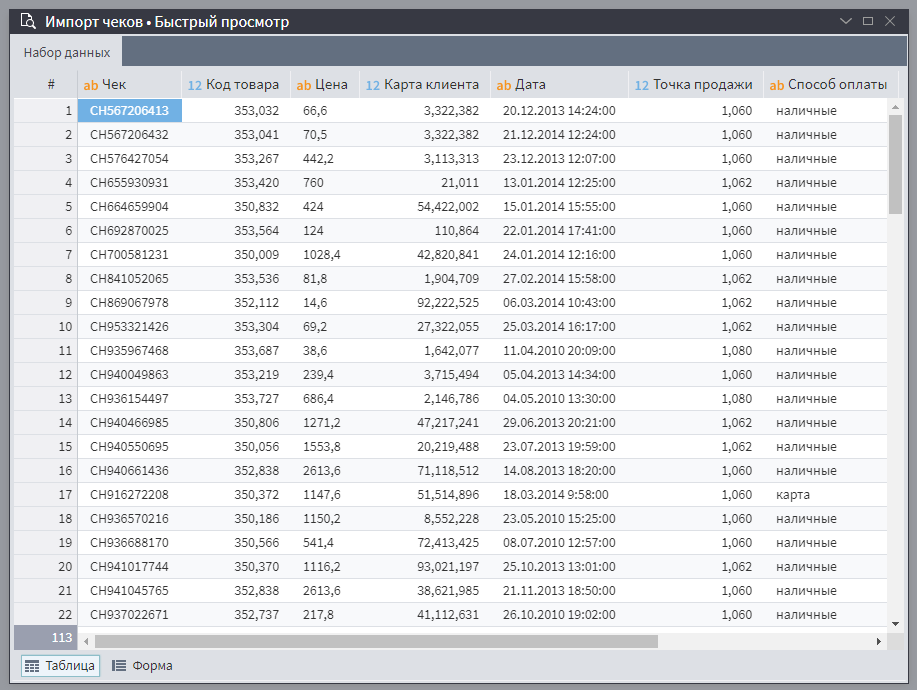


Рисунок 2.17. Просмотр содержимого узла в режиме «Таблица». Теперь можно увидеть всю таблицу.

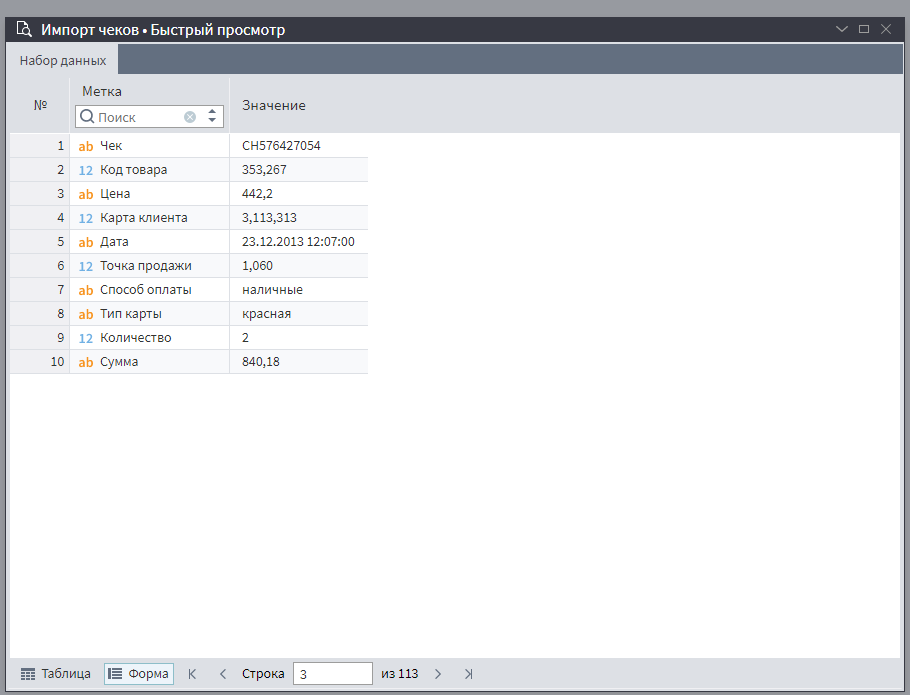


Рисунок 2.18. Представление 3-ей строки таблицы в режиме просмотра «Форма».

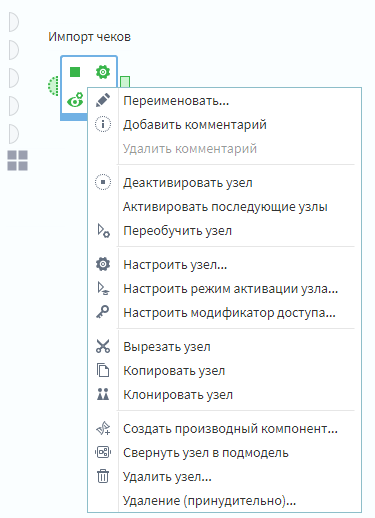


Рисунок 2.19. Переименование узла

#### **Шаг 3. Группировка данных**

Наблюдение и фиксация отдельных экономических и других данных обычно не позволяет охарактеризовать закономерности их появления так как при этом отображаются сведения по каждой единице исследуемой ситуации. Такие данные не являются обобщающими показателями. Поэтому требуется их предварительная обработка, которая позволила бы сделать определённый вывод о тенденции или даже сформулировать некую закономерность. **Это и есть один из этапов анализа данных!**

Поэтому важнейшим этапом анализа данных является **систематизация первичных данных** и получение на этой основе сводной характеристики всего объекта при помощи обобщающих показателей. Другими словами, надо создать ***сводку первичного статистического материала.***

**Определение понятия «сводка» при анализе данных**

***Сводка – это научная обработка первичных данных с целью получения обобщённых характеристик изучаемого социально-экономического явления по ряду существенных для него признаков с целью выявления типичных черт и закономерностей, присущих изучаемому явлению в целом.***

Сводка может быть различной в зависимости от ряда характеризующих её признаков. По глубине и точности обработки материала различают сводку простую и сложную.

**Простая сводка** – это операция по подсчёту общих итогов по совокупности единиц наблюдения и оформление этого материала в статистических таблицах.

**Сложная сводка** – это комплекс последовательных операций, включающих группировку полученных при наблюдении материалов, составление системы показателей для характеристики типичных групп и подгрупп изучаемой совокупности явлений, подсчёт числа единиц и итогов по каждой группе и подгруппе, и по всему объекту и представление результатов в виде статистических таблиц.

Проведению сводки предшествует разработка её программы по её созданию, которая включает следующие этапы:

* выбор группировочного признака;
* определение порядка формирования групп;
* разработка системы статистических показателей для характеристики групп и объекта в целом;
* разработка макетов статистических таблиц для представления результатов сводки.

**Группировкой** называется разбиение общей совокупности единиц объекта наблюдения по одному или нескольким существенным признакам на **однородные группы**, различающиеся между собой в количественном и качественном отношении и позволяющие выделить социально-экономические типы, изучить структуру совокупности и проанализировать связи между отдельными признаками.

**Группировки являются важнейшим статистическим методом обобщения статистических данных**, **основой для правильного исчисления статистических показателей**. Группировка проводится по определённому **признаку**.

Первое знакомство с Loginom – вебинар

<https://yandex.ru/video/preview/8029020725068850392>

1. В скопированный файл была вставлена дополнительная строка [↑](#footnote-ref-1)