### Лабораторная работа № 2

### Работа с визуализаторами и внешними библиотеками Loginom

### Практическое задание

### 1. Работа со стандартными компонентами и визуализаторами

1. Создайте новый пакет lab2. В нем настройте импорт из Loginom Data файл (\*.lgd) **Компоненты**  $\rightarrow$  **Импорт**  $\rightarrow$   $\bigcirc$  Loginom Data файл . Наведите мышкой на узел и увидите несколько Loginom Data

иконок, которые обозначают действия с файлом . Выберите **Настройка** и в окне **Импорт из Loginom Data файла** в строке **Имя файла** выберите файл *Торговля − товары для творчества 2009 год.lgd* (рис. 1).

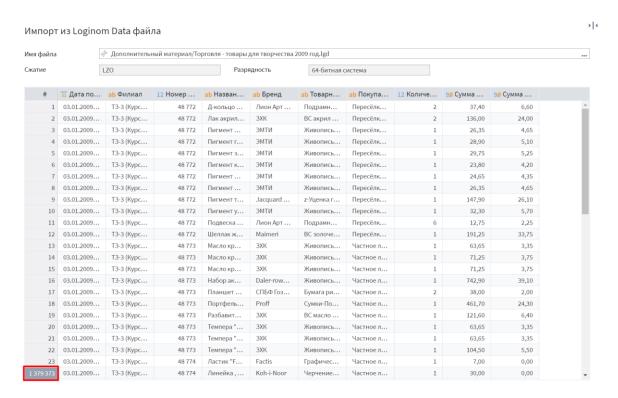


Рис. 1. Импорт из Loginom Data файла

**Loginom** распознал столбцы и строки данных, которых в этом наборе данных 1 379 373. Нажмите 2 раза **Далее** и в мастере настройки **Описание узла** задайте узлу метку *Товары для творчества*. Такие метки делаются для того чтобы в дальнейшем было легче разобраться в построенном сценарии.

Товары для творчества

Далее запустите узел нажав кнопку Выполнить. Все данные считались - узел стал

зеленым и его название изменилось . Иногда бывает полезно просмотреть какие данные попали для исследования, для этого сделаем Визуализацию данных.

2. Нажмите левой кнопкой мыши на узел **визуализатора** (рис. 2).

и выберите Настройка

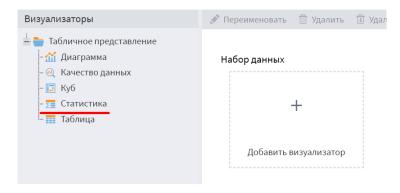


Рис. 2. Настройка визуализатора

Визуализаторы можно использовать разные (Диаграмма, Куб, Таблицы и др.) (рис 2). Выберем визуализатор Статистика, для этого с помощью мыши перетащите Статистика на поле Набор данных. Зайдите внутрь визуализатора щёлкнув на кнопке Войти. Посмотрите, как выглядят данные после загрузки (рис. 3).

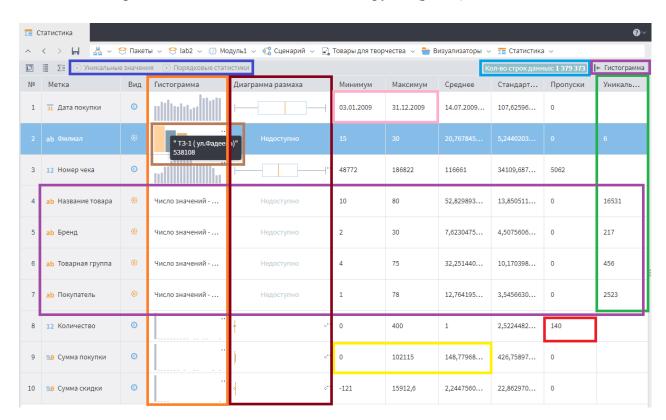


Рис. 3. Визуализатор Статистика

С помощью визуализатора **Статистика** можно посмотреть различные статистические показатели по каждому полю набора данных. По умолчанию отображается 8 показателей: *гистограмма значений поля*, *диаграмма размаха*, *минимальное*, *максимальное и среднее значения*, *стандартное отклонение*, *количество пропусков и* 

уникальных значений (рис. 3). Если часть данных не отобразилась щелкните на кнопке **Уникальные значения** и **Порядковые статистики** в адресной строке (рис. 3, синяя рамка).

В верхней части окна визуализатора Статистика отображается общее количество записей в наборе (рис. 3, голубая рамка).

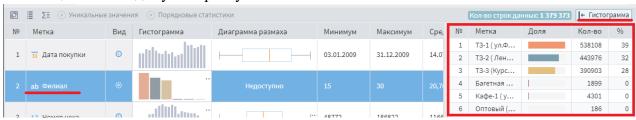
В столбце **Уникальные** можно видеть, что в наборе присутствуют продажи по 6 филиалам, всего в продажах участвовало 2523 покупателя, 16531 товар 217 брендов из 456 товарных группы (рис. 3, зеленая рамка).

Для показателя **Сумма покупки** можно посмотреть, на какую минимальную, максимальную и среднюю суммы произошли покупки (рис. 3, желтая рамка).

Гистограмма – наиболее универсальный показатель. Она отображает распределение значений по некоторым интервалам для полей непрерывного вида и распределение по уникальным значениям для дискретного вида данных (рис. 3, оранжевая рамка). При наведении курсора на столбец гистограммы можно увидеть количество значений, соответствующих данному интервалу или уникальному значению (рис. 3, коричневая рамка).

Гистограмма не отображается, если в поле большое количество уникальных значений, как в поле **Название Товара, Бренд, Товарная группа, Покупатель**, где их 16531, 217, 456, 2523 соответственно (рис. 3, фиолетовая рамка). В этом случае можно рассмотреть ее более детально, нажав кнопку **Гистограмма** в правом верхнем углу (рис. 3,

фиолетовая рамка). После нажатия на кнопку список интервалов/уникальных значений с количеством и процентом значений, относящихся к каждому интервалу.



Обратите внимание, что уникальные значения рассчитываются только для полей с дискретным видом данных . Знаком обозначен непрерывный тип данных. Для полей непрерывного вида также отображается минимум и максимум — границы первого и

последнего интервалов, и число интервалов. Все эти значения можно задавать вручную, и распределение значений по интервалам изменится в соответствии с новыми данными.

Ţ	🗊 🎚 ∑ 🕦 Уникальные значения 🕠 Порядковые статистики Кол-во строк данных: 1 379 373									тограмма		
Nο	Метка	Вид	Гистограмма	Диаграмма размаха	Минимум	Максимум	Cpe,	Nō	Метка	Доля	Кол-во	%
1	31 Дата покупки	0			03.01.2009	31.12.2009	1	1 2	<null></null>		53875	0 4
								3	20.01.2009		54415	4
2	ab Филиал	(6)		Недоступно	15	30	21	4	06.02.2009		69474	5
								5	23.02.2009		65525	5
3	12 Номер чека	0			48772	186822	1:	6	12.03.2009		75960	6
							1.	7	30.03.2009	_	63080	5
		0			40			8	16.04.2009		53815	4
4	<b>ab</b> Название товара	100	Число значений	Недоступно	10	80	5:		Минимум	03.01.	2009 00:00	
5	ab Бренд	*	Число значений	Недоступно	2	30	7,	_	Максимум	ум 31.12.2009 00:00		
6	аь Товарная группа	0	Число значений	Недоступно	4	75	3:	Числ	ю интервалов	1 100	00	<b>‡</b>

В целом доступность показателя для определенного поля зависит от типа или вида данных этого поля. Например, **Диаграмма размаха** доступна только для полей непрерывного вида (рис. 1, бордовая рамка).

В строке метки Дата покупки можно увидеть, что данные есть с января 2009 г. по конец 2009 г. (рис. 3, розовая рамка). В строке метки Количество указаны пропуски (140 пропусков) (рис. 3, красная рамка). Пропуски в номерах чеков скорее всего были или из-за отмены чеков или каких-то возвратов или закрытия смены, на данном этапе это не важно. Попробуем избавиться от этих пропусков с помощью компонента Заполнение пропусков.

### 3. Перетащите в рабочее пространство из панели Компоненты $\rightarrow$ Предобработка

компонент Заполнение пропусков . Соедините выходной порт набора данных с входным портом узла Заполнение пропусков:

Заполнение пропусков



Настройте его, щёлкнув левой кнопкой мыши по узлу Заполнение пропусков, и выберите . Далее выберите поле Количество, в качестве метода обработки Заменять наиболее вероятным. (рис. 4). Изучите настройки компонента Заполнение пропусков.

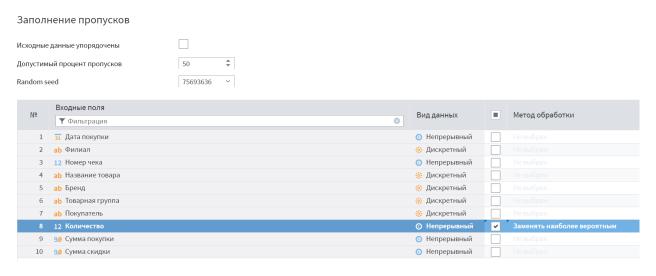


Рис. 4. Настройка узла Заполнение пропусков

Нажмите **Далее** и запустите выполнение компонента. Компонент заполнит пропуски наиболее подходящими данными. Что бы убедиться в этом добавьте визуализатор **Статистика** и посмотрите количество пропусков:

Nō	Метка	Вид	Гистограмма	Диаграмма размаха	Минимум	Максимум	Среднее	Стандарт	Пропуски
1					03.01.2009	31.12.2009	14.07.2009	107,62596	
2	ab Филиал	0		Недоступно	15	30	20,767845	5,2440203	0
3	12 Номер чека	<b>©</b>			48772	186822	116661	34109,687	5062
4	аь Название товара	0	Число значений	Недоступно	10	80	52,829893	13,850511	0
5	ab Бренд	0	Число значений	Недоступно	2	30	7,6230475	4,5075606	0
6	<b>ab</b> Товарная группа	(6)	Число значений	Недоступно	4	75	32,251440	10,170398	0
7	ab Покупатель	(6)	Число значений	Недоступно	1	78	12,764195	3,5456630	0
8	12 Количество	<u>©</u>		···	0	400	1	2,5234875	0

4. Рассмотрим работу с товарными группами. Для этого воспользуемся компонентом

**Группировка** (**Компоненты** → **Трансформация**). Соедините выходные и входные порты двух узлов и войдите в настройки узла **Группировка**. Изучите компонент Группировка.



В качестве **Группы** выберем *Товарная группа* и *Дата покупки* (рис. 5, красная рамка), в качестве **Показателей** выберем *Сумма покупки* с вариантом агрегации *Сумма* (рис. 5, синяя рамка). Вариант агрегации можно изменить выбрав кнопку **Редактировать (F2)** (рис. 5, зеленая рамка).

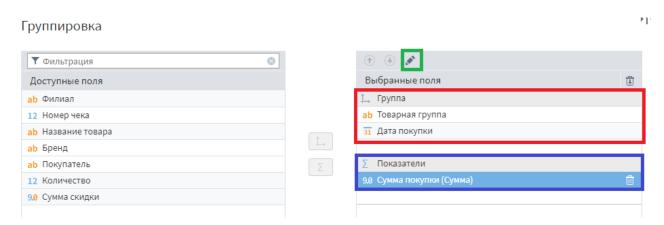
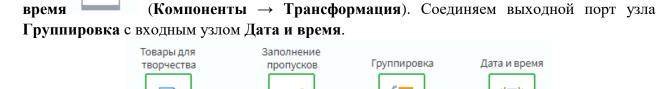


Рисунок 5 Настройка узла Группировка

Для того чтобы выделить временные периоды воспользуемся компонентом Дата и



Перейдем в настройку 🌼 компонента Дата и время. Выберем Месяц и нажимаем Далее. В Описании узла изменим метку на Выделение месяца (рис. 6).

Преобразование даты/времени

время

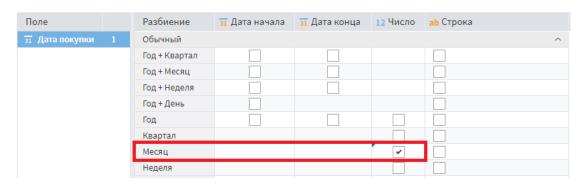
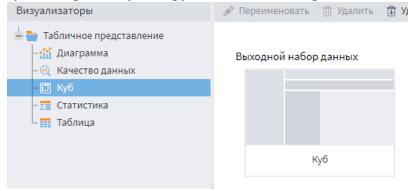


Рис. 6. Настройка узла Дата и время

Посмотрим, как изменяются продажи по товарным группам. Для этого к узлу Дата и время добавим визуализатор Куб из группы Табличное представление.



В левой части визуализатора находится Область списка полей, в правой – Область построения куба (рис. 7, синие рамки). Область построения в свою очередь разделена на 4 части (рис. 7, красные рамки): • - область списка измерений, значения которых располагаются в строках; 2 - область списка измерений, значение которых будут столбцами: 3 - область значений: 4 - область фильтрации, сюда также можно добавить измерения, они не будут отображаться в кубе, но могут участвовать в фильтрации, сортировке и вычисляемых фактах.

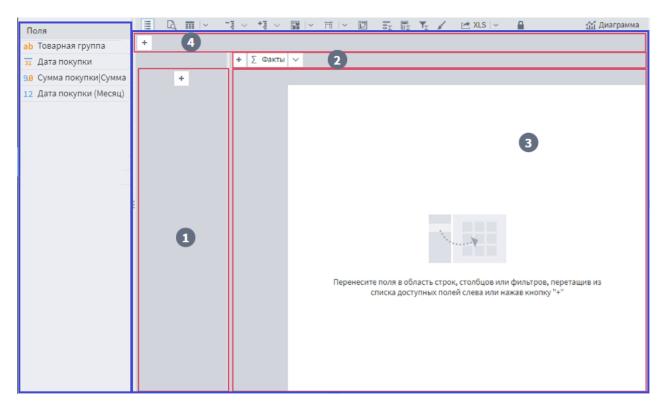


Рис. 7. Визуализатор Куб. Настройка

Для того чтобы настроить отчет, необходимо добавить нужные поля в соответствующие области. Это можно сделать, перетащив поля с помощью Drag & Drop или нажав кнопку + в нужной области и выбрав поле из списка. Настройте отчет. Добавьте измерения в строки. В качестве измерений будем использовать Дата покупки (Месяц), в качестве фактов Товарная группа (рис. 8).

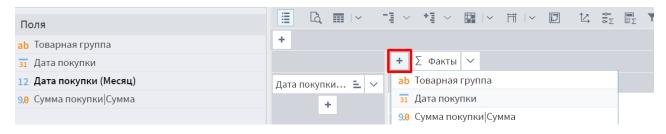
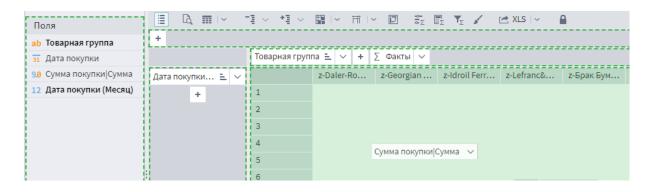


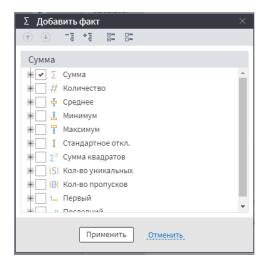
Рис. 8. Настройка визуализатора Куб

В область значений добавим Сумма покупки | Сумма, перетащив в центр куба 9.0 Сумма покупки Сумма (рис. 9).



#### Рис. 9. Настройка визуализатора Куб

При добавлении необходимо выбрать вариант агрегации. Оставим вариант по умолчанию – **Сумма**.



В результате получили многомерный отчет по продажам. Кроме значений фактов по заданным измерениям в кубе рассчитываются итоги по строкам и столбцам. Поля, задействованные отчете, выделились жирным в списке.

После того, как отчет сформирован, можно перенастраивать его структуру: добавлять, удалять измерения, менять из местами с помощью перетаскивания или соответствующих кнопок. Ограничений по количеству активных измерений нет, но для того что бы отчет был интерпретируемым, рекомендуется использовать не более 5-6.

7. Отфильтруем товарные группы по товару **Кисти**. Для этого нажмите кнопку товарная группа и в открывшемся списке уберите весь выбор с помощью и отфильтруйте их нажав на по слову **кист**, т.к. существует разные варианты написания слова (кисть, кисти, кистей) (рис. 10, 11)

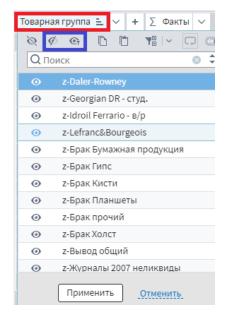


Рис. 10. Настройка фильтра

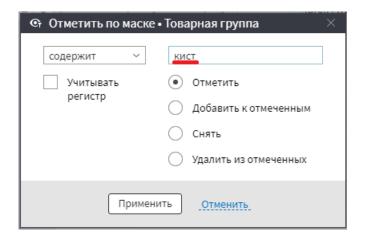
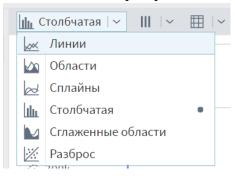


Рис. 11. Настройка маски

Убедитесь, что фильтр сработал, проверив заголовки оставшихся столбцов.

8. Рассмотрим полученный результат на диаграмме нажав на кнопку расположенную в правом верхнем угле. Получили диаграмму, построенную на основе куба. Ее отличие от обычной диаграммы в том, что она однозначно соответствует текущему состоянию куба и при любых его изменениях, перестраиваются. В раскрывающемся списке:



выберите Линии.

Чтобы рассмотреть какие-то линии более детально следует щёлкнуть правой

Показать все серии

кнопкой мыши внизу по сериям, выбрать **Скрыть все серии**— Скрыть все серии

— Инвертировать серии

и щёлкая левой кнопкой мыши выбрать нужные нам линии (рис. 12)

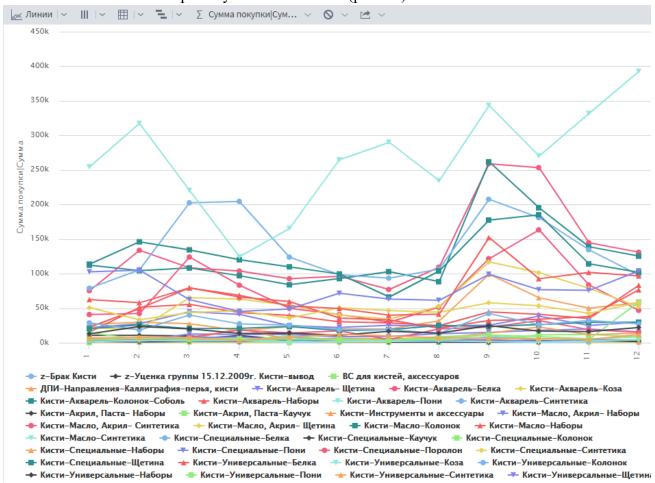
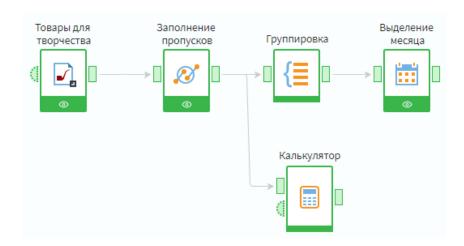


Рис. 12. Диаграмма Линии

9. Чаще всего какие-либо расчёты производятся с помощью компонента Калькулятор, перенесём его в область построения сценариев. Соединим выходной порт узла Заполнение пропусков с входным портом узла Калькулятор:



10. Перейдите в настройки <sup>©</sup> узла **Калькулятор**. Создадим выражение с именем *Price* и меткой *Цена*, тип оставим *Вещественный*. *Цену* можно посчитать по формуле (*Сумма покупки* + *Сумма скидки*) / *Количество* (рис. 13).

Калькулятор

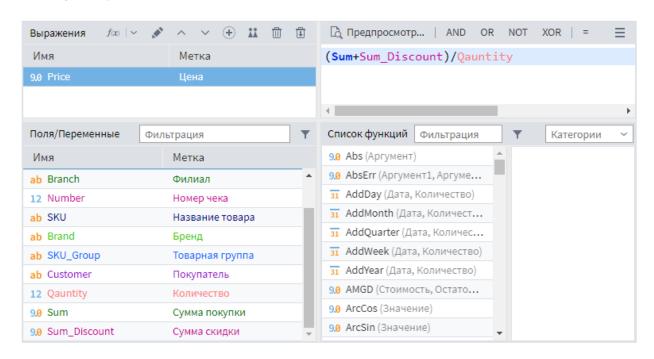


Рис. 13. Выражение Ргісе

Добавим еще одно выражение нажав на кнопку . Зададим имя *Discount*, метку *Процент скидки*, тип *Вещественный*. Рассчитаем по формуле (*Сумма скидки / Сумма покупки*) \* 100 (рис. 14).

Калькулятор

### Калькулятор

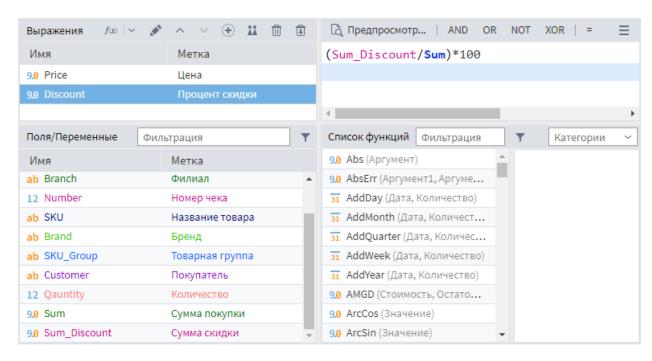


Рис. 14. Выражение Discount

В быстром просмотре узла **Калькулятор** видно, что появились два новых столбца ( в начале или конце строки заголовков столбцов) **Цена** и **Процент скидки** (рис. 15).

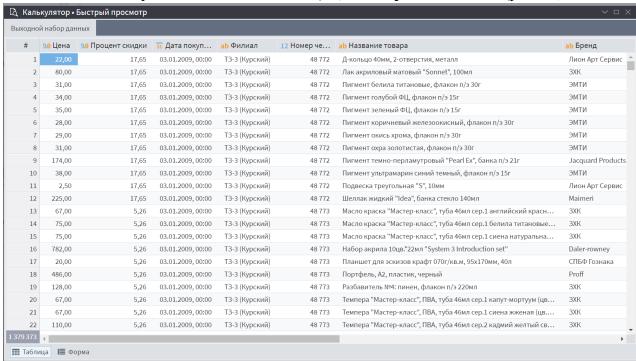


Рис. 15. Результат работы узла Калькулятор

# 2. Работа с библиотекой Loginom silver kit

### Краткая теория

В Loginom предусмотрена возможность повторного использования настроенных узлов или фрагментов сценария:

- Копирование фрагмента. Данный способ позволяет создавать полностью независимую копию узла/узлов. Никакой связи между исходным узлом и копией не формируется, при необходимости внесения изменений их придется вносить в каждую копию. По сути это то же самое, что создавать нужные узлы заново с такими же настройками. Таким образом, единственное преимущество копирования ускорение разработки сценария. Поэтому использовать этот способ при реализации повторной логики не рекомендуется.
- Использование стандартного компонента Выполнение узла. Данный способ позволяет использовать имеющиеся узлы повторно для обработки новых данных, но исключает возможность увидеть внутреннюю структуру/настройки узлов или вносить в них изменения. При этом существует связь с базовым узлом: при внесении изменений в его настройки, узел выполнения отнаследуют эти изменения.
- Создание производного компонента. Производный компонент и созданный на базе него узел также имеют связь с базовым узлом, но, в отличие от узла выполнения, в его логику обработки можно вносить изменения.

Последние два способа подразумевают не только работу с созданными фрагментами в рамках одного пакета, на также создание и подключение внешней библиотеки компонентов. Остановимся на этих способах подробнее.

Внешняя библиотека компонентов — это пакет с набором спроектированных подмоделей, каждая из которых выполняет какую-либо законченную обработку. Все подмодели библиотеки, ставшие производными компонентами, по умолчанию имеют область видимости Внутренний. Для того, чтобы компонент стал доступен в любом пакете, который будет его использовать, область видимости нужно вручную повысить до Открытый.

Библиотеки компонентов позволяет сократить время на разработку сценариев, многократно использовать логику обработки, ранее заложенную в компоненты. Механизм наследования обеспечивает актуальность версий компонентов. В библиотеку можно объединить компоненты для решения задач отдельной предметной области.

Внешние библиотеки условно можно классифицировать на закрытые и открытые:

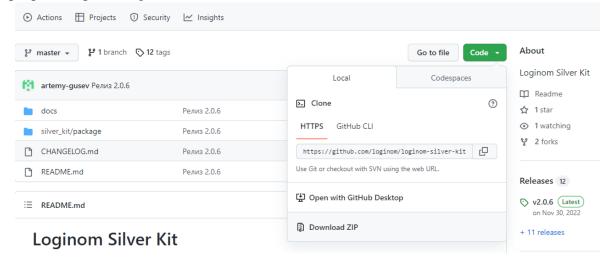
- Закрытые библиотеки. Закрытая библиотека представляет собой зашифрованный пакет, компоненты которого можно использовать в другом пакете только одним способом с помощью выполнения узла. Сам пакет библиотеки невозможно открыть в Loginom, можно использовать только ссылку на него. Соответственно, в компоненты библиотеки нельзя войти и увидеть, как они устроены внутри.
- Открытые библиотеки. Пакет открытой библиотеки можно открыть в Loginom точно также, как и любой другой пакет. На основе каждого из открытых узлов может быть создан производный компонент, что позволяет использовать возможности объектно-ориентированного моделирования (подробнее о производных компонентах и узлах далее). В этом случае можно увидеть, а, при желании, модифицировать внутреннее устройство компонентов как в пакете библиотеке, так и при подключении ссылки на него.

Существует и промежуточный вариант, когда пакет библиотеки зашифрован, но ее компоненты (или часть компонентов) доступны как производные.

Библиотека **Loginom Silver Kit** содержит компоненты для ETL и подготовки данных, Data Mining и прочие рутинные последовательности действий.

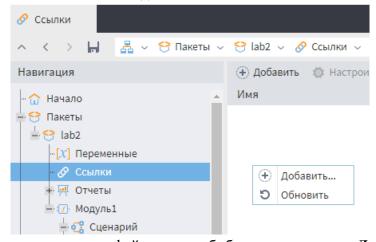
Рассмотрим на ее примере способы использования внешних компонентов. Библиотека расположена на GitHub. Для скачивания библиотеки перейдите по ссылке:

https://github.com/loginom/loginom-silver-kit - и нажмите кнопку **Download ZIP**. Там же можно познакомиться с документацией к библиотеке. После скачивания распакуйте архив с библиотекой на свой компьютер, либо в файловое хранилище (при использовании серверной версии Loginom).



Воспользуемся готовым сценарием, подключив библиотеку Loginom-silver-kit-master.

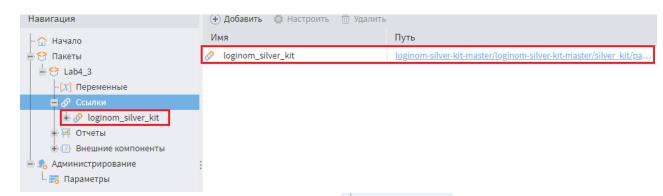
1. Щёлкните на Пакеты , перейдите в В Ссылки , выберите команду Добавить в контекстном меню раздела дерева пакетов, в контекстном меню области ссылок, либо можно воспользоваться одноименной кнопкой на панели инструментов



Далее следует указать путь к файлу пакета библиотеки и нажать Добавить.

Добавление ссылки на пакет						
Пакет	Путь					
Путь к файлу пакета:						
:er/loginom-silver-kit-n	naster/silver_kit/package/loginom_silver_kit.lgp	***				
(	Добавить					

Ссылка на библиотеку появилась в дереве пакетов, а подробная информация о ней – в рабочей области экрана.



2. Перейдите в сценарий нажав на кнопку Сценарий

воспользоваться

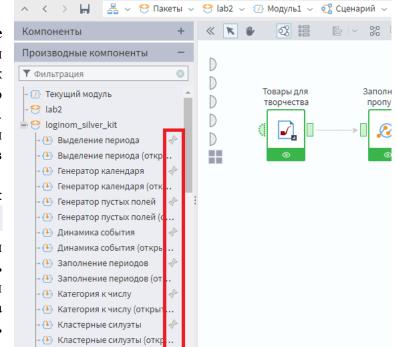
чтобы

Производные компоненты +
Во вкладке Производные компоненты появился подключённый пакет и список компонентов, которые можно использовать в текущем пакете. Все производные компоненты сгруппированы по пакетам, в

Раскройте список

loginom\_silver\_kit. Значок указывает на то, что компоненты библиотеки можно использовать как производные узлы. При отсутствии данного значка компонент можно использовать только через выполнение узла.

которых они созданы.



добавленной

библиотекой

- 3. Пусть стоит целью узнать, какие продукты и услуги компании приносят наибольшую прибыль, а от каких лучше отказаться. Для этого обычно используются Базовые методы:
  - 1. Анализ рентабельности ассортиментной группы товаров
  - 2. АВС-анализ

2. ADC anasir

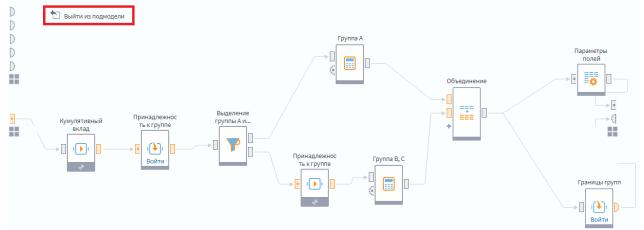
- 3. Анализ адаптивной матрице BCG
- 4. Анализ по методу Дибба-Симкина
- 5. Анализ по матрице совместных покупок

Остановим свое внимание на методе ABC-анализа, основанному на законе Парето<sup>1</sup>. Согласно принципу Парето "Небольшая часть усилий приводит к большим результатам", если говорим о товарах, это означает, что достаточная небольшая часть товаров приносит наибольшую прибыль. Для реализации этого метода можно воспользоваться известными расчетными формулами, а можно взять готовый компонент из библиотеки Loginom. Для нашей задачи воспользуемся компонентом **ABC-анализ** (метод касательных)

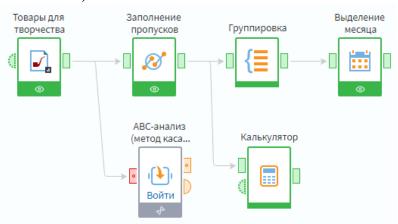
 $<sup>^1</sup>$  Закон Парето (принцип Парето, принцип 80/20) — эмпирическое правило, названное в честь экономиста и социолога Вильфредо Парето, в наиболее общем виде формулируется как «20 % усилий дают 80 % результата, а остальные 80 % усилий — лишь 20 % результата»



рабочее поле сценария и нажмите Войти . Компонент АВС-анализ (метод касательных) откроется как подмодель и можно увидеть его уже готовую для использования структуру:



Соединим выходной порт узла Товары для творчества с входным портом узла АВС-анализ (метод касательных).



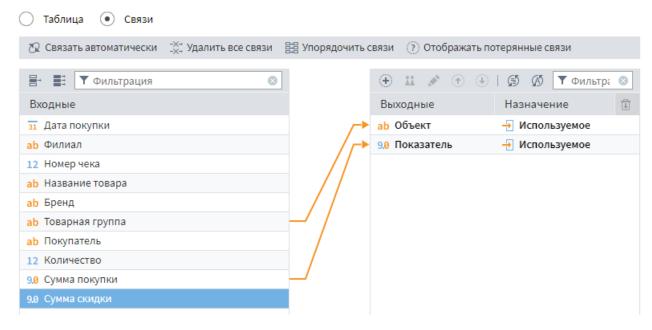
Входной порт узла подсвечен красным , следовательно его нужно настроить. Для это щёлкнем правой кнопкой мыши по порту и выберем **Настроить порт** настроить порт...

Редактирование метки порта...

На этапе **Настройка соответствия между столбцами** 

выберем связи Таблица • Связи и соединим Товарная группа и Объект, Сумма покупки и Показатель методом перетаскивания (рис. 16) и нажмите Выполнить.

### Настройка соответствия между столбцами



Ри. 16. Настройка входного порта узла АВС-анализ

4. В быстром просмотре выходного узла мы можем наблюдать, что всего 129 товарных групп приносят нам основную прибыль (рис. 17), остальные товары могут быть так же важны, чтобы поддерживать ассортимент.

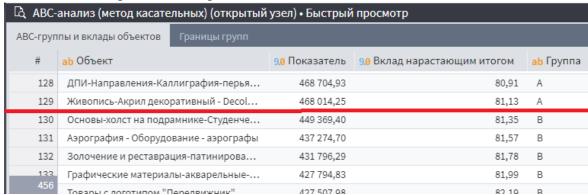


Рис. 17. Результат АВС-анализа

# Задания для самостоятельной работы

#### Задание 1.

- 1. Создайте новый файл и назовите его Задание 1.
- 2. Импортируйте файл Товары для творчества филиал ТЗ 1.lgd
- 3. Добавьте компонент **Фильтр строк** (<u>Помощь</u>) и отсортируйте данные с условием, что Дата покупки была совершена после 01.01.2009г.
- 4. Добавьте компонент **Группировка** (<u>Помощь</u>) и узнайте, сколько принес каждый бренд

На выходе должно получиться как на рис. 18, а.

### Задание 2.

Доработайте файл Задание 1.

1. Добавьте в рабочую область компонент Калькулятор

- 2. С помощью Калькулятора рассчитайте:
  - а. Общую сумму по всему филиалу (Воспользуйтесь функцией Stat)
  - b. Рассчитайте % по бренду (Задействуйте общую сумму по всему филиалу)
- 3. Добавьте визуализатор Таблица и отсортируйте % по бренду по убыванию.

На выходе должно получиться как на рис. 18, б.

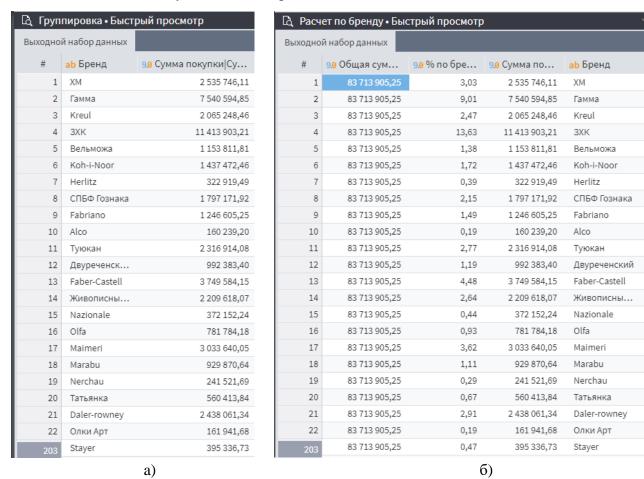


Рис. 18. Результат заданий 1 и 2

### Контрольные вопросы

- 1. Опишите назначение иконок внутри компонента.
- 2. Какие типы визуализаторов есть в Loginom?
- 3. Как добавить визуализатор к компоненту Loginom?
- 4. Какие статистические показатели отображаются в визуализаторе Статистика?
- 5. Где в визуализаторе Статистика можно увидеть общее количество записей в наборе?
- 6. Для данных какого типа могут быть рассчитаны уникальные значения? Где их модно посмотреть?
- 7. Опишите показатель Гистограмма в визуализаторе Статистика, что этот показатель отображает для разных типов данных?
- 8. В каких случаях показатель гистограмма не отображается в визуализаторе Статистика?
- 9. Где показатель Статистика можно посмотреть детально? Для полей какого типа отображается минимум, максимум и число интервалов. Можно ли их менять?
- 10. Для какого типа данных доступен показатель Диаграмма размаха?
- 11. Где можно посмотреть временной диапазон загруженных данных?

- 12. Какой компонент позволяет заполнить пропуски данных в Loginom? Что считается пропуском в Loginom?
- 13. Опишите настройки компонента Заполнение пропусков (исходные данные упорядочены, допустимый процент пропусков, Random seed пр.)?
- 14. Какие методы обработки пропусков предусмотрены в Loginom?
- 15. Опишите область построения куба для визуализатора Куб в Loginom?
- 16. Как настроить фильтрацию данных в визуализаторе Куб?
- 17. Какие типы диаграммы можно использовать в визуализаторе Куб?
- 18. Как настроить отображение тех или иных серий данных в визуализаторе Куб?
- 19. Какими способами в Loginom можно организовать повторное использование настроенных узлов или фрагментов сценария?
- 20. Какие из этих способов предусматривают подключение внешних библиотек?
- 21. Что из себя представляет внешняя библиотека компонентов?
- 22. На какие два типа можно классифицировать внешние библиотеки, охарактеризуйте каждый тип.
- 23. Какие последовательности действий содержит библиотека Loginom Silver Kit?
- 24. Где хранится библиотека Loginom Silver Kit и как ее подключить?
- 25. Где можно посмотреть содержимое, подключенной библиотеки?
- 26. Что означает значок
- 27. Как посмотреть содержимое какого-либо компонента библиотеки Loginom Silver Kit?