

ГРУППИРОВКА И ПРЕОБРАЗОВАНИЕ ДАТЫ

- ETL-ПРОЦЕСС
- КОМПОНЕНТ «ГРУППИРОВКА»
- КОМПОНЕНТ «ДАТА И ВРЕМЯ»

ETL-ПРОЦЕСС



Важнейшим процессом в системах хранения данных является извлечение данных из различных источников, их преобразование к единому формату и модели данных, очистка от дубликатов, противоречий и других факторов, которые могут помешать анализу, а также загрузка в единый интегрированный источник.

Для реализации данного процесса в систему включается специальный комплекс аппаратно-программных средств, называемых ETL (англ.: Extract, Transform, Load – извлечение, преобразование, загрузка).

ETL-ПРОЦЕСС

Рассмотрим преобразование данных подробнее. **Преобразование данных** (которое еще называют трансформацией) зависит от задач, алгоритмов и целей анализа. Таким образом, для разных задач потребуются различные методы преобразования.

Преобразование данных – очень широкое понятие, которое не имеет четко очерченных границ. Данный термин иногда распространяют на любые манипуляции с данными независимо от их целей и методов.

ETL-ПРОЦЕСС

Однако в контексте аналитики преобразование данных имеет вполне конкретные цели и задачи, а также использует достаточно стабильный набор методов.

Преобразование данных не ставит целью изменить информационное содержание данных. Его задача – представить информацию в таком виде, чтобы она могла быть использована наиболее эффективно с точки зрения решаемых задач аналитики данных.

ETL-ПРОЦЕСС

Преобразование данных в том или ином виде выполняется во многих компонентах информационных систем:

- В процессе переноса и загрузки данных в интегрированный источник или области временного хранения (ETL).
- Непосредственно при подготовке данных к анализу в бизнес-приложении (SKD).

Такая распределенность процесса преобразования обусловлена тем, что на каждом этапе он преследует различные цели.

ETL-ПРОЦЕСС

Операции преобразования могут проводиться для обеспечения технической и логической совместимости данных, их подготовки к извлечению, переносу в хранилище данных и так далее.

Например, адреса часто вводят одной строкой. В то же время для анализа могут представлять интерес отдельные компоненты адреса, которые имеют как **текстовый формат** (улица, город), так и **числовой** (номер дома, офиса). С помощью трансформации можно распределить соответствующие элементы по отдельным полям и преобразовать их в нужный формат.

ETL-ПРОЦЕСС

Существует две основные цели преобразования данных на этапе процесса ETL:

- Приведение их в соответствие с моделью данных, используемой в консолидированном источнике.
- Осуществление корректной консолидации данных и загрузка в консолидированный источник.

ETL-ПРОЦЕСС

В связи с этим возникает вопрос. Если преобразованию данных так много внимания уделяется на этапе интеграции данных, то зачем включать средства преобразования в аналитическое приложение? Ведь это, несомненно, усложняет и делает более дорогим его разработку.

Ответ прост. Не все данные поступают в бизнес-приложение из систем, где они прошли предварительную подготовку. Но главная причина заключается в том, что трансформация данных в этих системах в большей степени носит технический характер и слабо связана с возможными методами, алгоритмами и целями анализа.

ETL-ПРОЦЕСС

В связи с этим возникает вопрос. Если преобразованию данных так много внимания уделяется на этапе интеграции данных, то зачем включать средства преобразования в аналитическое приложение? Ведь это, несомненно, усложняет и делает более дорогим его разработку.

Ответ прост. Не все данные поступают в бизнес-приложение из систем, где они прошли предварительную подготовку. Но главная причина заключается в том, что трансформация данных в этих системах в большей степени носит технический характер и слабо связана с возможными методами, алгоритмами и целями анализа.

ETL-ПРОЦЕСС

Дата	Количество
01.03.2020	10 000
02.03.2020	12 000
01.04.2020	8 000
03.04.2020	13 000



Дата	Количество
01.03.2020	22 000
01.04.2020	21 000

Рассмотрим в качестве примера такую ETL-операцию, как **группировка**.

Иногда высокая степень детализации данных может мешать анализу данных. Ежедневные показатели некоторого бизнес-процесса, например, продаж, подвержены колебаниям под действием различных случайных факторов, поэтому если взять значения продаж по отдельности за каждый день, вряд ли они дадут надежную оценку, характеризующую развитие процесса. Но если сгруппировать ежедневные показатели за неделю или месяц, то оценка будет более обоснованной.

ETL-ПРОЦЕСС

Таким образом, группировка данных – полезный инструмент, применение которого в процессе подготовки данных к анализу позволяет более эффективно использовать содержащуюся в данных информацию.

Перед тем, как приступить к группировке, необходимо определить в исходной выборке данных поля групп и показателей.

Группы (или измерения) – это данные, характеризующие исследуемый процесс качественно, а показатели (их также называют фактами) – количественно.

Группа	Показатель
Клиент	Сумма
Клиент 1	15 000
Клиент 2	23 000

ETL-ПРОЦЕСС

Например, процесс продаж описывается **качественными** и **количественными** данными.

Качественные данные – наименования товаров и клиентов. Очевидно, что просто указать эти параметры недостаточно. С каждым проданным товаром или клиентом связан набор числовых показателей: цены, количество, суммы, скидки, наценки и так далее.

Каждое наименование товара или клиента, которое содержится в поле группы, называется **значением группы**. Суть группировки заключается в том, что все записи, содержащие одноименные значения группы, по которой проводится группировка, объединяются в одну, а соответствующие показатели агрегируются.

Качественные	Количественные
Клиент	Сумма
Клиент 1	15 000
Клиент 2	23 000

Значение группы

ETL-ПРОЦЕСС

Например, процесс продаж описывается **качественными** и **количественными** данными.

Качественные данные – наименования товаров и клиентов. Очевидно, что просто указать эти параметры недостаточно. С каждым проданным товаром или клиентом связан набор числовых показателей: цены, количество, суммы, скидки, наценки и так далее.

Каждое наименование товара или клиента, которое содержится в поле группы, называется **значением группы**. Суть группировки заключается в том, что все записи, содержащие одноименные значения группы, по которой проводится группировка, объединяются в одну, а соответствующие показатели агрегируются.

Клиент	Сумма
Клиент 1	15 000
Клиент 2	23 000

ETL-ПРОЦЕСС

Рассмотрим возможные ETL-операции с **датой и временем**.

Обычно каждый элемент временного ряда представляет собой соответствие дата-значение, а в случае многомерного ряда – дата-значение 1, значение 2 ...

Все методы обработки, рассмотренные ранее, были направлены на преобразование значений, сами же даты обычно оставались без изменений. Однако, как показывает практика, при решении различных задач анализа может быть полезным и преобразование даты.

Дата	Значение 1	Значение 2	Значение N
01.01.2020	10	25,6	...
01.02.2020	20	245,25	...
...
01.12.2020	560	54,34	...

ETL-ПРОЦЕСС

Дата	Дата (Неделя)	Дата (Месяц)	Дата (Квартал)	Дата (Год)
24.04.2020	17	04 Апрель	2	2020
12.08.2020	33	08 Август	3	2020
30.12.2020	53	12 Декабрь	4	2020

Чаще всего даты в исходных данных представлены в определенном формате, обычно **ДД.ММ.ГГ**, то есть день, месяц и год, при этом на отображение каждого элемента даты отводятся две цифры.


Возможны некоторые вариации: указывается полностью год (02.10.2014) или(и) месяц (12 сентября 2014) и т.д.

С точки зрения анализа наибольший интерес представляют не детализированные данные за каждый день, а агрегированные по неделям, месяцам, кварталам или годам. В этой связи может оказаться полезным извлекать из даты дополнительную информацию о временных интервалах, на которых проводится анализ.

ETL-ПРОЦЕСС

Так, если базовый интервал, по которому осуществляется анализ деятельности предприятия, – **неделя**, то в некоторых случаях может представлять интерес отображение даты не на каждый день, а на первый или последний день недели.

Иными словами, все данные за неделю определяются по первому или последнему ее дню.

Дата	Кол-во		Дата	Дата (Год + Неделя, Первый день)	Кол-во
02.01.2017	250		02.01.2017	02.01.2017	250
03.01.2017	230		03.01.2017	02.01.2017	230
04.01.2017	345		04.01.2017	02.01.2017	345
05.01.2017	215		05.01.2017	02.01.2017	215
06.01.2017	312		06.01.2017	02.01.2017	312
07.01.2017	124		07.01.2017	02.01.2017	124
08.01.2017	321		08.01.2017	02.01.2017	321
09.01.2017	234		09.01.2017	09.01.2017	234
10.01.2017	243		10.01.2017	09.01.2017	243
11.01.2017	312		11.01.2017	09.01.2017	312
12.01.2017	321		12.01.2017	09.01.2017	321
13.01.2017	267		13.01.2017	09.01.2017	267
14.01.2017	351		14.01.2017	09.01.2017	351
15.01.2017	216		15.01.2017	09.01.2017	216
16.01.2017	187		16.01.2017	16.01.2017	187
17.01.2017	179		17.01.2017	16.01.2017	179
18.01.2017	261		18.01.2017	16.01.2017	261
19.01.2017	305		19.01.2017	16.01.2017	305
20.01.2017	156		20.01.2017	16.01.2017	156



КОМПОНЕНТ «ГРУППИРОВКА»

ЗАДАНИЕ

В файле **sales.lgd** хранятся данные о продажах строительных товаров: их названия и товарные группы, дата продажи, количество проданного и единица измерения, сумма с учетом скидки, а также информация о городе продажи и типе клиента-покупателя. Всего в наборе **98 471 запись**.

Нам необходимо получить набор данных с продажами каждого товара по дням.



Поскольку один и тот же товар мог быть продан в один день несколько раз в разных городах, нам потребуется провести группировку. Для этого создадим новый пакет и импортируем данный набор в Loginom.

Компонент **Группировка** позволяет объединять записи по выбранным полям в группы, агрегируя данные в полях, которые выбраны в качестве показателей, с помощью различных статистических функций. Для каждой группы возвращается одна строка.

Фильтрация

Рассмотрим два параметра настроек узла. Флаг **Кэшировать значения групп** включен по умолчанию и позволяет сохранить полученные после группировки значения групп. Это может существенно ускорить последующую обработку, особенно при подаче большого набора данных, однако необходимо учитывать, что чем больше набор, тем больше памяти займут сохраненные значения.

При установке флага **Сортировать результирующие данные** записи на выходе: будут отсортированы по полям в списке **Группы** с учетом их расположения в списке.

Выбранные поля		
Группа		
31 Дата		
ab Товар		
Показатели		
9.0 Количество (Сумма)		
9.0 Сумма с учетом скидки (Сумма)		

Кэшировать значения групп

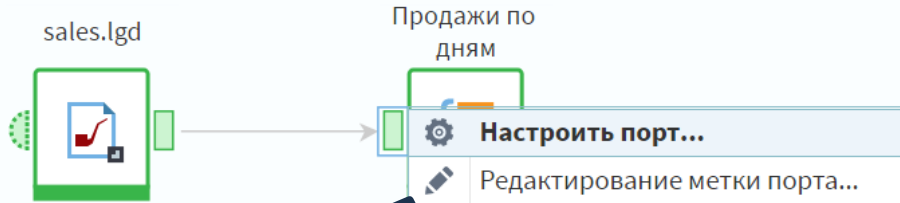


Сортировать результирующие данные



#	31 Дата	ab Товар	9.0 Количество Сум...	9.0 Сумма с учетом скидки Сумма
1	01.03.2016, 00:00	Доска паркетная UPOFLOOR Лос 2085x188x14 дуб красный преми...	28,00	13 708,80
2	01.03.2016, 00:00	Гвоздь оцинкованный 1.4x25, 300 штук, 10 категория, уп	112,00	3 311,84
3	01.03.2016, 00:00	Гипсокартон АБС 9.5x1200x2500, кв.м.	44,00	2 271,28
4	01.03.2016, 00:00	Гипсокартон БЕЛГИПС 12.5x1200x2500, кв.м	44,00	2 259,40
5	01.03.2016, 00:00	Гипсокартон БЕЛГИПС 9.5x1200x2500, влагостойкий, мал.уп., кв.м	44,00	2 820,40
6	01.03.2016, 00:00	Гипсокартон БЕЛГИПС 9.5x1200x2500, кв.м	44,00	2 135,76
7	01.03.2016, 00:00	Гипсокартон КНАУФ 12.5x1200x2500, кв.м.	91,00	4 945,78
8	01.03.2016, 00:00	Гипсокартон КНАУФ 12.5x1200x2500, огн., кв.м.	44,00	2 772,88
9	01.03.2016, 00:00	Лист стекломгнезитовый СМЛ 1220x2440x8, шт	44,00	17 442,48
10	01.03.2016, 00:00	Пигмент блескообразующий TICIANA золото, 0.1 кг, шт	60,00	11 160,00
11	01.03.2016, 00:00	Пигмент блескообразующий TICIANA серебро, 0.1 кг, шт	60,00	11 160,00
12	01.03.2016, 00:00	Покрытие декоративное OPTIMIST ELITE Манна, 1000 мл, шт	60,00	1 100,40
13	01.03.2016, 00:00	Покрытие декоративное OPTIMIST ELITE Старый дуб, 1000 мл, шт	60,00	1 100,40
14	01.03.2016, 00:00	Решетка радиаторная КОСТРОМА 600x1200мм, белая, шт	60,00	2 641,20
15	01.03.2016, 00:00	Решетка радиаторная КОСТРОМА 600x600мм, белая, шт	60,00	2 641,20
16	01.03.2016, 00:00	Решетка радиаторная КОСТРОМА 600x900мм, белая, шт	60,00	2 641,20
17	01.03.2016, 00:00	Шпатлевка ТЕКС РЕ-файн по дереву белая, 0.25 кг, шт	60,00	1 100,40
18	01.03.2016, 00:00	Шпатлевка ТЕКС РЕ-файн по дереву белая, 0.75 кг, шт	60,00	2 641,20
19	01.03.2016, 00:00	Шпатлевка ТЕКС РЕ-файн по дереву береза, 0.75 кг, шт	60,00	2 641,20
20	01.03.2016, 00:00	Шпатлевка ТЕКС РЕ-файн по дереву бук, 0.75 кг, шт	60,00	2 641,20
21	01.03.2016, 00:00	Шпатлевка ТЕКС РЕ-файн по дереву дуб, 0.25 кг, шт	60,00	1 100,40
75 455	01.03.2016, 00:00	Шпатлевка ТЕКС РЕ-файн по дереву дуб, 0.25 кг, шт	120,00	2 200,80

Мы получили набор данных по продажам товаров по дням. Необходимо обратить внимание, что при группировке мы потеряли информацию о значениях показателей в размене исключенных групп.

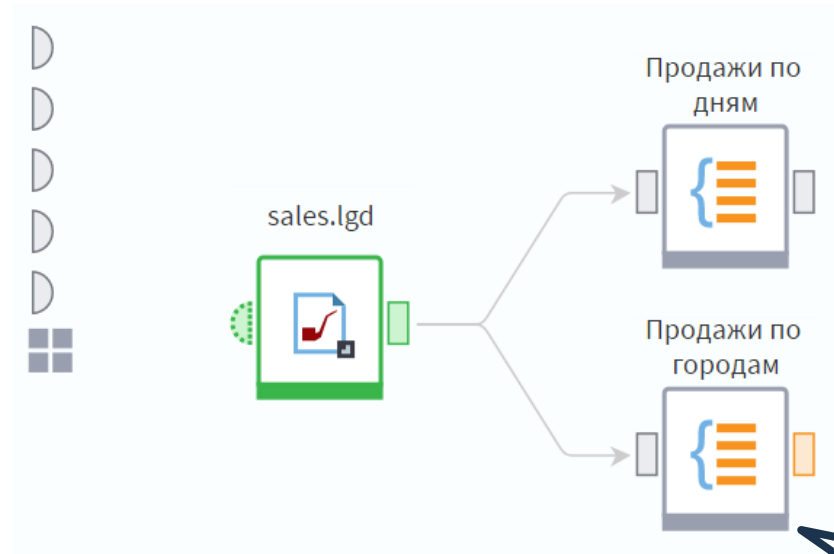


Перейдем в настройку входного узла **Продажи по дням**. После того, как узел настроен, назначение полей во входном порте автоматически меняется в соответствии с заданными настройками. Чтобы изменить этот параметр необходимо дважды щелкнуть на нужном поле левой кнопкой мыши.

Таблица Связи

Входные	Имя	Вид данных	Назначение
31 Дата	Date	Непрерывн...	Группа
9.0 Количество		Непрерывн...	Показатель
9.0 Сумма с учетом скидки		Непрерывн...	Показатель
ab Группа товара		Дискретный	Не задано
ab Товар		Дискретный	Группа
ab Единица измерения		Дискретный	Не задано
ab Город		Дискретный	Не задано
ab Группа клиента		Дискретный	Не задано

Обратите внимание, что при включенной **автосинхронизации** работать с узлом группировки нужно с осторожностью – если имена полей в новых входных данных не совпадут с заданными на входе узла, новые поля, даже если у них настроено назначение, не свяжутся с имеющимися, а добавятся к ним.



Добавим в сценарий еще один узел группировки и переименуем оба узла, как показано на слайде.

Фильтрация

Доступные поля

- 31 Дата
- ab Группа товара
- ab Товар
- ab Единица измерения
- ab Группа клиента



Выбранные поля

- Группа
- ab Город
- Показатели
- 9.0 Количество (Сумма)
- 9.0 Сумма с учетом скидки (Сумма)



В новом узле настроим группировку продаж по сумме и количеству по городам. Для этого нам понадобится указать в списке групп только одно поле – **Город**, а показатели останутся такими же.

Кэшировать значения групп

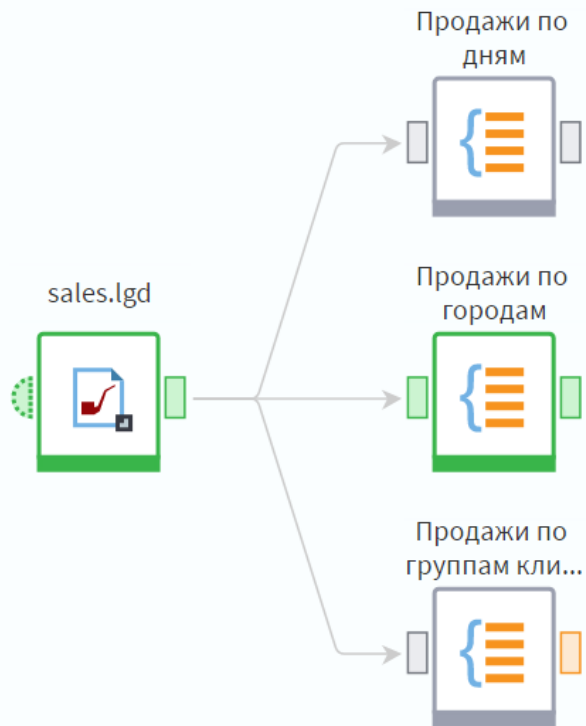


Сортировать результирующие данные



🔍 Продажи по городам • Быстрый просмотр			
Выходной набор данных			
#	ab Город	9.0 Количество Сум...	9.0 Сумма с учетом скидки Сумма
1	Балашиха	43 631,00	9 823 339,29
2	Барнаул	95 211,00	16 145 202,69
3	Великие Луки	77 556,00	14 877 644,50
4	Екатеринбург	10 436,00	1 882 384,41
5	Зеленоград	144 000,00	26 587 113,84
6	Иваново	141 437,00	30 147 324,08
7	Иркутск	3 669,00	572 372,28
8	Москва	2 499 028,00	476 055 694,18
9	Нальчик	16 654,00	2 733 682,58
10	Нижний Новгород	977 495,00	202 129 396,56
11	Новороссийск		
12	Пермь		
13	Петрозаводск		
14	Самара		
15	Сергиев Посад		
16	Тольятти		
17	Уссурийск		
18	Уфа	82 112,00	20 805 622,08
19	Владимир	351 686,00	71 025 229,83
20	Казань	57 217,00	9 894 487,86
21	Кемерово	53 492,00	12 118 969,91
73	Киров (Кировской обл.)	265 926,00	76 044 628,22

На выходе мы получили общие продажи по каждому из 73 городов.



Иногда требуется получить список всех значений какого-либо поля. В этом случае можно провести группировку без показателей. Обратите внимание: если выбрать несколько полей, то сохранится иерархия, которая присутствовала в исходных данных. Проведем группировку по полю **Группа клиентов**.

Фильтрация

Доступные поля

31	Дата
9.0	Количество
9.0	Сумма с учетом скидки
ab	Группа
ab	Товар
ab	Единица измерения
ab	Город

Выбранные поля

Группа
ab Группа клиента
Показатели

Мы получили список всех существующих в нашем наборе групп клиентов. Можно также использовать группировку без измерений, тогда мы получим агрегированное значение по выбранному полю.

Продажи по группам клиентов • Быстрый просмотр

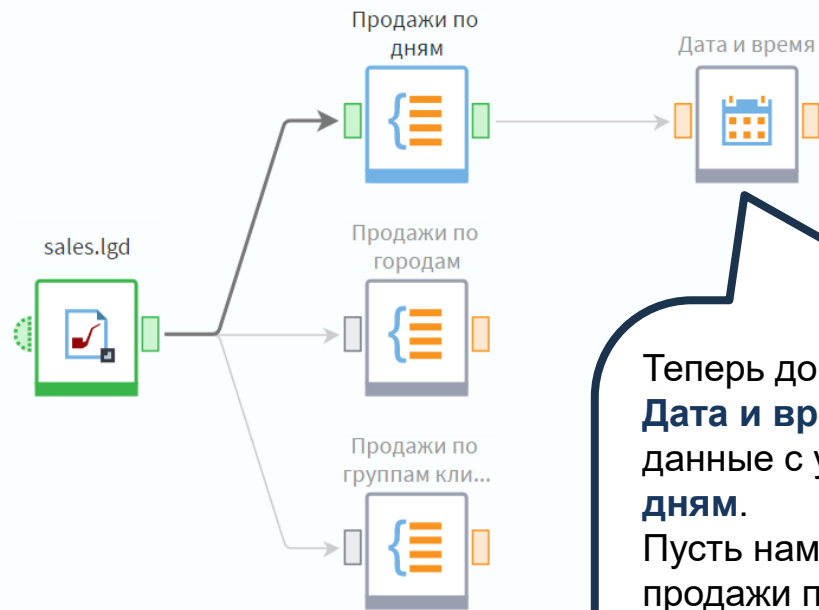
Выходной набор данных

#	ab Группа клиента
1	Клиент
2	VIP клиент
3	Постоянный клиент

- Кэшировать значения групп
- Сортировать результирующие данные



КОМПОНЕНТ «ДАТА И ВРЕМЯ»



Теперь добавим в наш сценарий узел **Дата и время** и подадим на его вход данные с узла группировки **Продажи по дням**.

Пусть нам необходимо получить продажи по месяцам года. Очевидно, что для этого нужно сделать группировку, но поля с месяцем продажи в нашем наборе нет. Однако его легко получить из поля **Дата** с помощью узла **Дата и Время**.

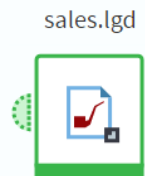
Поле	Разбиение	31 Дата начала	31 Дата конца	12 Число	ab Строка
31 Дата	Обычный				
	Год + Квартал	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
	Год + Месяц	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input type="checkbox"/>
	Год + Неделя	<input 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 checked="" type="checkbox"/>	<input type="checkbox"/>
	Неделя			<input type="checkbox"/>	<input checked="" type="checkbox"/> %W
	День года			<input type="checkbox"/>	<input type="checkbox"/>
				<input type="checkbox"/>	<input type="checkbox"/>
				<input type="checkbox"/>	<input type="checkbox"/>
				<input type="checkbox"/>	<input type="checkbox"/>
				<input type="checkbox"/>	<input type="checkbox"/>
				<input type="checkbox"/>	<input type="checkbox"/>
				<input type="checkbox"/>	<input type="checkbox"/>
				<input type="checkbox"/>	<input type="checkbox"/>
				<input type="checkbox"/>	<input type="checkbox"/>

Нам необходимо поле с месяцем года, поэтому выберем разбиение **Год+Месяц** с типом **Дата/Время** (выберем оба варианта: **Дата начала** и **Дата конца**). Дополнительно выделим из поля **Дата** номер месяца и неделю продажи с типами **Число** и **Строка** соответственно

Выходной набор данных

#	31 Дата	31 Дата (Год + Месяц, Последний день)	31 Дата (Год + Месяц, Первый де...	12 Дата (Месяц)	ab Дата (Неделя)	ab Товар
1	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Доска паркетная UFLOOR Loc 2085x188
2	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Гвоздь оцинкованный 1.4x25, 300 штук, 10
3	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Гипсокартон АБС 9.5x1200x2500, кв.м.
4	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Гипсокартон БЕЛГИПС 12.5x1200x2500, кв.
5	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Гипсокартон БЕЛГИПС 9.5x1200x2500, вла
6	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Гипсокартон БЕЛГИПС 9.5x1200x2500, кв.
7	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Гипсокартон КНАУФ 12.5x1200x2500, кв.м.
8	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Гипсокартон КНАУФ 12.5x1200x2500, огн.,
9	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Лист стекломгнезитовый СМЛ 1220x2440
10	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Пигмент блескообразующий TICIANA зол
11	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Пигмент блескообразующий TICIANA сер
12	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Покрытие декоративное OPTIMIST ELITE I
13	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Покрытие декоративное OPTIMIST ELITE I
14	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Решетка радиаторная КОСТРОМА 600x120
15	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Решетка радиаторная КОСТРОМА 600x600
16	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Решетка радиаторная КОСТРОМА 600x900
17	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Шпатлевка ТЕКС РЕ-файн по дереву бела
18	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Шпатлевка ТЕКС РЕ-файн по дереву бела
19	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Шпатлевка ТЕКС РЕ-файн по дереву бере
20	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Шпатлевка ТЕКС РЕ-файн по дереву бук,
21	01.03.2016, 00:00	31.03.2016, 00:00	01.03.2016, 00:00	3	10	Шпатлевка ТЕКС РЕ-файн по дереву дуб,

В наш набор данных добавились четыре новых поля, которые характеризуют поле **Дата**. Обратите внимание, что поля **Дата (Год+Месяц, Последний день)** и **Дата (Год+Месяц, Первый день)** обозначают один и тот же период, только в одном случае указывается дата конца этого периода, а в другом – дата начала.



Продажи по
дням



Продажи по
городам



Продажи по
группам кли...



Дата и время



Группировка



Проведем группировку по месяцам года.

Фильтрация

Доступные поля

- 31 Дата
- 31 Дата (Год + Месяц, Последний день)
- 12 Дата (Месяц)
- ab Дата (Неделя)



Выбранные поля

Выбранные поля

- Группа
- 31 Дата (Год + Месяц, Первый день)
- ab Товар
- Показатели
- 9.0 Количество|Сумма (Сумма)
- 9.0 Сумма с учетом скидки|Сумма (Сумма)

Проведем группировку как показано на слайде.

Кэшировать значения групп

Сортировать результирующие данные

#	31 Дата (Год + Месяц, Первый де...	ab Товар	9.0 Количество Сумма Сум...	9.0 Сумма с учетом скидки Сумма Сумма
1	01.03.2016, 00:00	Доска паркетная UPOFLOOR Loc 2085x188x14 дуб красный преми...	1 162,00	560 415,74
2	01.03.2016, 00:00	Гвоздь оцинкованный 1.4x25, 300 штук, 10 категория, уп	228,00	6 741,96
3	01.03.2016, 00:00	Гипсокартон АБС 9.5x1200x2500, кв.м.	893,00	45 330,10
4	01.03.2016, 00:00	Гипсокартон БЕЛГИПС 12.5x1200x2500, кв.м	1 492,00	75 673,98
5	01.03.2016, 00:00	Гипсокартон БЕЛГИПС 9.5x1200x2500, влагостойкий, мал.уп., кв.м	680,00	42 393,18
6	01.03.2016, 00:00	Гипсокартон БЕЛГИПС 9.5x1200x2500, кв.м	1 616,00	77 394,60
7	01.03.2016, 00:00	Гипсокартон КНАУФ 12.5x1200x2500, кв.м.	1 822,00	99 640,38
8	01.03.2016, 00:00	Гипсокартон КНАУФ 12.5x1200x2500, огн., кв.м.	832,00	51 754,54
9	01.03.2016, 00:00	Лист стекломagneзитовый СМЛ 1220x2440x8, шт	1 977,00	765 126,28
10	01.03.2016, 00:00	Пигмент блескообразующий TICIANA золото, 0.1 кг, шт	2 110,00	425 567,38
11	01.03.2016, 00:00	Пигмент блескообразующий TICIANA серебро, 0.1 кг, шт	1 047,00	232 326,31
12	01.03.2016, 00:00	Покрытие декоративное OPTIMIST ELITE Манна, 15 кг, шт	300,00	199 569,00
13	01.03.2016, 00:00	Покрытие декоративное OPTIMIST ELITE Старый замок, 15 кг, шт	2 913,00	1 826 181,56
14		Кебая, шт	2 518,00	441 034,60
15		Кебая, шт	3 216,00	277 059,23
16		Кебая, шт	3 501,00	490 272,13
17		шт	531,00	9 621,16
18		шт	1 215,00	52 881,51
19		Шпатлевка ТЕКС РЕ-файн по дереву береза, 0.75 кг, шт	60,00	2 641,20
20	01.03.2016, 00:00	Шпатлевка ТЕКС РЕ-файн по дереву бук, 0.75 кг, шт	737,00	31 991,09
21	01.03.2016, 00:00	Шпатлевка ТЕКС РЕ-файн по дереву дуб, 0.25 кг, шт	594,00	10 742,10
8 410				

В результате количество записей сильно сократилось (до 8 410), так как мы получили продажи по месяцам.