

Семинарское занятие №5

Тема «Измерения в многомерных моделях»

Измерения в многомерных моделях

SQL Server 2014

[Другие версии](#)

Измерение базы данных является коллекцией объектов, называемых атрибутами, которые используются для предоставления сведений о данных фактов в одном или нескольких кубах. Например, типичным атрибутом измерения «Продукт» может быть название, категория, размер, цена продукта или линия товаров. Эти объекты привязаны к одному или нескольким столбцам в одной или нескольких таблицах в представлении источника данных. По умолчанию эти атрибуты отображаются как иерархии атрибутов и позволяют понять смысл данных фактов в кубе. Атрибуты могут быть организованы в пользовательские иерархии, которые обеспечивают различные пути доступа к данным и помогают пользователям при просмотре данных в кубе. Кубы содержат все измерения, которыми пользователи пользуются при анализе данных фактов. Экземпляр измерения базы данных в кубе называется измерением куба и относится к одной или нескольким группам мер в кубе. Измерение базы данных может использоваться в кубе несколько раз. Например, если таблица фактов содержит несколько зависимых от времени фактов, то для облегчения анализа каждого из них может быть определено отдельное измерение куба. Однако необходимо существование только одного зависимого от времени измерения базы данных, что означает также необходимость существования лишь одной зависимой от времени таблицы реляционной базы данных для поддержки нескольких зависимых от времени измерений куба.

Примечание

Дополнительные сведения о проблемах производительности, связанных с конструкцией измерений, см. в разделе [Руководство по управлению производительностью служб Analysis Services SQL Server 2008 R2](#).

Определение измерений, атрибутов и иерархий

Для определения измерений, атрибутов и иерархий базы данных и куба проще всего воспользоваться мастером кубов, который позволяет создавать измерения одновременно с определением куба. Мастер кубов создает измерения на основе обнаруженных или указанных пользователем таблиц измерений из представления источников данных, используемого для

куба. После этого мастер создает измерения базы данных и добавляет их к новому кубу, создавая измерения куба.

При создании куба в него можно также добавить любые измерения, которые уже существуют в базе данных. Эти измерения могли быть ранее определены мастером измерений для другого куба. После того как измерение базы данных определено, его изменение и настройка производятся в конструкторе измерений. Кроме того, в ограниченной степени измерение куба может быть настроено при помощи конструктора кубов.

Примечание

Разрабатывать и настраивать измерения, атрибуты и иерархии можно также программным способом с помощью XML для аналитики или объектов АМО. Дополнительные сведения см. в разделах [Справочник по языку ASSL](#) и [Разработка объектов управления аналитикой \(объекты АМО\)](#).

В этом разделе

В следующей таблице описаны подразделы, содержащиеся в этом разделе.

[Определение измерений базы данных](#)

Описывает, как производится изменение и настройка измерения базы данных при помощи конструктора измерений.

[Справочник по свойствам атрибута измерения](#)

Описывает определение, изменение и настройку атрибута измерения базы данных при помощи конструктора измерений.

[Определение связей атрибутов](#)

Описывает определение, изменение и настройку связи атрибутов при помощи конструктора измерений.

[Создание пользовательских иерархий](#)

Описывает определение, изменение и настройку пользовательской иерархии атрибутов измерений при помощи конструктора измерений.

[Использование мастера бизнес-аналитики для улучшения измерений](#)

Описывает, как улучшить измерение базы данных при помощи мастера бизнес-аналитики.

См. также

[Кубы в многомерных моделях](#)

Создание мер и групп мер в многомерных моделях

SQL Server 2014

[Другие версии](#)

Мера — это агрегат значений числовых данных, например сумма, количество, минимальное, максимальное, среднее или созданное вами многомерное выражение. *Группа мер* — это контейнер для одной или нескольких мер. Все меры находятся в группе мер даже в том случае, если имеется только одна мера. Куб должен иметь как минимум одну меру и группу мер.

Этот раздел включает следующие подразделы:

- [Подходы к созданию мер](#)
- [Компоненты меры](#)
- [Моделирование мер и групп мер по фактам и таблицам фактов](#)
- [Гранулярность группы мер](#)

Подходы к созданию мер

Меры могут быть статическим элементом куба, созданного во время разработки, и всегда присутствуют при доступе к кубу. Кроме того, меру можно определить как *вычисляемый элемент* с помощью многомерных выражений, когда ее значение вычисляется на основе других мер в кубе. Вычисляемый элемент можно распространить на сеанс или пользователя.

Чтобы создать меру или группу мер, используйте один из следующих подходов:

Мастер кубов	Запустите мастер кубов в SQL Server Data Tools (SSDT) для создания куба. В обозревателе решений щелкните правой кнопкой мыши узел Кубы и выберите команду Создать куб . Обратитесь к разделу Многомерное моделирование (учебник по Adventure Works) , если вам нужна помощь по этим действиям. При создании куба на основе таблиц из существующего хранилища данных определения мер и групп мер материализуются в качестве части процесса создания куба. В мастере выберите, какие факты и таблицы фактов нужно использовать в качестве основы для объектов меры и группы мер в кубе.
Диалоговое окно создания меры	Если куб уже существует в SQL Server Data Tools (SSDT), дважды щелкните имя куба в обозревателе решений, чтобы открыть его в конструкторе кубов. В области "Меры" щелкните правой кнопкой мыши верхний узел, чтобы создать новую группу мер или новую меру с помощью указания исходной таблицы, столбца и типа агрегирования. Этот подход требует выбора метода агрегирования из фиксированного списка готовых функций. В разделе Использование агрегатных функций представлено описание наиболее часто используемых агрегатов.
Вычисляемый элемент	Вычисляемые элементы позволяют добиться гибкости и широких возможностей анализа кубов в Службы Analysis Services, так как можно управлять временем и способом их создания. Иногда мера требуется только временно — в течение сеанса пользователя или в среде Management Studio в рамках исследования.

	<p>В SQL Server Data Tools (SSDT) откройте вкладку "Вычисления", чтобы создать новый вычисляемый элемент.</p> <p>Выберите этот подход при создании меры на основе многомерного выражения. Дополнительные сведения см. в статьях Построение мер в многомерных выражениях, Вычисления, Вычисления в многомерных моделях и Основные принципы создания скриптов многомерных выражений (службы Analysis Services).</p>
Многомерные выражения или XMLA	<p>В SQL Server Management Studio можно выполнять многомерные выражения или XMLA, чтобы изменить базу данных для включения новой вычисляемой меры. Этот подход полезен для нерегламентированной проверки данных после развертывания решения на сервере. См. раздел Документирование и работа со скриптами в базе данных служб Analysis Services.</p>

Компоненты меры

Мера представляет собой объект со свойствами. Кроме имени мера должна иметь тип агрегата и исходный столбец или выражение, используемое для наполнения меры данными. Можно изменить определение меры, задавая ее свойства.

источник	<p>Большинство мер являются производными от числовых столбцов таблиц фактов во внешнем хранилище данных, например столбца Sales Amount в таблицах Sales и Reseller Sales из хранилища данных AdventureWorks, однако можно создать новые меры, основываясь целиком на определяемых вычислениях.</p> <p>Столбцы атрибутов из таблиц измерения можно использовать для определения мер, но такие меры обычно являются полуаддитивными или неаддитивными в отношении статистической обработки. Дополнительные сведения о режиме полуаддитивном см. в разделе Определение полуаддитивного режима.</p>
агрегат	<p>По умолчанию меры суммируются вдоль каждого измерения. Однако свойство AggregateFunction позволяет изменить это. См. список в разделе Использование агрегатных функций.</p>
Свойства	<p>См. описания дополнительных свойств в разделе Настройка свойств мер.</p>

Моделирование мер и групп мер по фактам и таблицам фактов

Перед запуском мастера важно знать принципы моделирования определения меры.

Меры и группы мер — многомерные объекты, представляющие факты и таблицы фактов во внешнем хранилище данных. В большинстве случаев меры и группы мер будет основываться на объектах в представлении источника данных, которые в свою очередь создаются из базового хранилища данных.

На следующей диаграмме показана таблица фактов **FactSalesQuota** и две связанные с ней таблицы измерений — **DimTime** и **DimEmployee**. В образце куба Adventure Works эти таблицы используются в качестве основы для группы мер Sales Quotas и измерений Time и Employee.

Таблица фактов содержит столбцы двух основных типов: столбцы атрибутов и столбцы мер.

- Столбцы атрибутов используются для создания связей между внешними ключами и таблицами измерений, чтобы количественные данные в столбцах мер можно было организовать по данным, содержащимся в таблицах измерений. Столбцы атрибутов используются также для определения гранулярности таблицы фактов и ее группы мер.
- Столбцы мер определяют меры, которые содержатся в группе мер.

При запуске мастера кубов внешние ключи отфильтровываются. В списке оставшихся столбцов для выбора находятся столбцы мер, а также столбцы атрибутов, которые не определены как внешний ключ. В примере **FactSalesQuote** мастер предложит **CalendarYear** и **CalendarQuarter** в дополнение к **SalesAmountQuota**. Только столбец меры **SalesAmountQuota** приведет к созданию поддающейся обработке меры для многомерной модели. Для определения суммы каждой квоты существуют другие столбцы на основе даты. Следует исключить из списка мер в мастере кубов другие столбцы — **CalendarYear** и **CalendarQuarter** (или позднее удалить их из группы мер в конструкторе).

Вывод из данного обсуждения: не все столбцы, предлагаемые мастером, полезны в качестве меры. Полагайтесь на собственное понимание данных и того, как они будут использоваться, при выборе столбцов для использования в качестве меры. Помните, что можно щелкнуть правой кнопкой мыши таблицу в представлении источника данных для нахождения данных, которые помогут определить, какие столбцы использовать в качестве мер. Подробнее см. в разделе [Просмотр данных в представлении источника данных \(службы Analysis Services\)](#).

Примечание

Не все меры извлекаются непосредственно из значения, хранимого в столбце таблицы фактов. Например, мера **Sales Person Count**, определенная в группе мер **Sales Quota** примера куба Adventure Works фактически основана на подсчете уникальных значений (или числа различных элементов) в столбце **EmployeeKey** таблицы фактов **FactSalesQuota**.

Гранулярность группы мер

Группы мер имеют гранулярность, которая описывает уровень детализации, поддерживаемой таблицей фактов. Гранулярность устанавливается через отношение внешнего ключа к измерению. Например, таблица фактов **FactSalesQuota** имеет связь по внешнему ключу с таблицей **DimEmployee**, каждая запись в таблице **FactSalesQuota** относится к одному сотруднику, таким образом, гранулярность группы мер с точки зрения измерения Employee находится на уровне индивидуального сотрудника.

Гранулярность группы мер не может быть задана мельче самого нижнего уровня измерения, из которого просматривается эта группа мер, а крупнее ее можно сделать при помощи дополнительных атрибутов. Например, в таблице фактов **FactSalesQuota** столбцы **TimeKey**, **CalendarYear** и **CalendarQuarter** используются для

установки гранулярности связи с таблицей **DimTime**. В результате этого гранулярность группы мер, как видно из измерения «Время», соответствует календарному кварталу, а не дню, который является нижним уровнем измерения «Время».

Можно задать гранулярность группы мер по определенному измерению с помощью вкладки **Использование измерений** конструктора кубов. Дополнительные сведения о связях между измерениями см. в разделе [Связи измерений](#).

См. также

[Кубы в многомерных моделях](#)

[Меры и их группы](#)

Секции в многомерных моделях

SQL Server 2014

[Другие версии](#)

В Службы Analysis Services физическое хранение фактических данных, загруженных в группу мер, обеспечивает *секция*. Для каждой группы мер автоматически создается одна секция, но существует стандартная практика создания дополнительных секций для дальнейшего сегментирования данных, что обеспечивает более эффективную обработку и более быстрое выполнение запросов.

Обработка более эффективна, поскольку секции могут обрабатываться параллельно на одном или нескольких серверах. Запросы выполняются быстрее, поскольку для каждой секции можно настроить режимы хранения и оптимизации агрегатов, что приводит к ускоренному получению ответов. Например, хранилище MOLAP для секций с новыми данными обычно работает быстрее, чем ROLAP. Также, если выполнено секционирование по дате, секции, содержащие новые данные, могут иметь больше оптимизаций, чем секции со старыми данными, обращение к которым происходит реже. Следует отметить, что применение разных режимов хранения и агрегирования в секциях негативно отразится на будущих операциях слияния. Перед оптимизацией отдельных секций учитывайте возможность слияния в своей стратегии управления секциями.

Примечание

Поддержка работы с несколькими секциями доступна в выпусках Business Intelligence и Enterprise. Выпуск Standard не поддерживает работу с множественными секциями. Дополнительные сведения см. в разделе [Возможности, поддерживаемые различными выпусками SQL Server 2014](#).

Важные соображения при проектировании стратегии секционирования

Целостность данных куба зависит от данных, распределенных по разделам куба, чтобы в секциях отсутствовали повторяющиеся данные. При сведении данных из секций любые элементы данных, присутствующие более чем в одной секции, будут сведены, как будто бы они были различными элементами данных. Это может привести к неправильным сводкам и ошибочным данным, которые представляются конечному пользователю. Например, если транзакция продаж для продукта X дублируется в таблице фактов для двух секций, сводные продажи продукта X могут включать двойные вхождения дублирующейся транзакции.

Секции могут быть объединены. Эту возможность можно использовать в общей стратегии хранения и обновления данных. Объединять можно только секции с одними и теми же режимами хранения и статистической схемой. Для создания секций для последующего объединения при создании секций можно скопировать статистическую схему другой секции. Кроме того, можно изменить секцию после ее создания, чтобы скопировать статистическую схему другой секции. Слияние секций следует выполнять осторожно во избежание повторения данных в результирующей секции, что может привести к неточности данных куба.

Локальные секции

Локальные секции представляют собой секции, определяемые, обрабатываемые и хранимые на одном сервере. При наличии больших групп мер в кубе может появиться необходимость разделить их, чтобы обработка производилась параллельно в секциях. Преимуществом этой параллельной обработки является более быстрое выполнение. Поскольку одно задание обработки секции необязательно должно закончиться до запуска другого, они могут выполняться параллельно. Дополнительные сведения см. в разделе [Создание локальной секции и управление ей \(Analysis Services\)](#).

Удаленные секции

Удаленные секции представляют собой секции, определенные на одном сервере, но обрабатываемые и хранимые на другом. При необходимости распределения хранилища данных и метаданных по нескольким серверам используйте удаленные секции. Обычно при переходе от разработки к производству объем анализируемых данных возрастает в несколько раз. При наличии таких больших объемов данных одной из возможных альтернатив является распределение этих данных по нескольким компьютерам. Это связано не с тем, что один компьютер не может хранить все данные, а с необходимостью параллельной обработки этих данных несколькими компьютерами. Дополнительные сведения см. в разделе [Создание удаленной секции и управление ей \(Analysis Services\)](#).

Агрегаты

Агрегатами называются предварительно вычисленные сводки по данным куба, помогающие службам Службы Analysis Services ускорить выполнение запросов. Можно управлять количеством агрегатов, создаваемых для группы мер, задавая пределы для хранилища и значения прироста производительности, а также останавливая процесс построения агрегатов. Наличие большего числа агрегатов не всегда приводит к лучшим результатам. Каждый новый агрегат затрачивает дополнительные ресурсы, включая место на диске и процессорное время. Рекомендуется создавать агрегаты, чтобы достичь повышения производительности примерно на 30 %, а затем увеличивать их число только на основании тестирования или опыта использования. Дополнительные сведения см. в разделе [Проектирование агрегатов \(службы Analysis Services — многомерные данные\)](#).

Слияние и правка секций

Если две секции используют одну и ту же статистическую схему, то можно осуществить слияние этих двух секций в одну. Например при наличии измерения описи, разбитого на секции по месяцам, в конце каждого календарного месяца можно осуществлять слияние секции за месяц с секций на дату года. Таким образом, секция текущего месяца может быть быстро обработана и проанализирована, в то время как остальную часть года по месяцам необходимо обработать повторно только при слиянии. Эта повторная обработка требует более длительного времени, но может выполняться реже. Дополнительные сведения об управлении процессом слияния секций см. в разделе [Объединение секций в службах Analysis Services \(SSAS — многомерные данные\)](#). Описание изменения секции куба с помощью вкладки **Секции** в конструкторе кубов см. в разделе [Изменение и удаление секций \(Analysis Services — многомерные данные\)](#).

См. также

Раздел	Описание
Создание локальной секции и управление ей (Analysis Services)	Содержит сведения о секционировании данных с помощью фильтров и различных таблиц фактов без дубликации данных.
Определение хранилища секции (Analysis Services — многомерные данные)	Содержит описание настройки хранилища для секций.
Изменение и удаление секций (Analysis Services — многомерные данные)	Содержит описание способа просмотра и редактирования секций.
Объединение секций в службах Analysis Services (SSAS — многомерные данные)	Содержит сведения о слиянии секций, имеющих различные таблицы фактов или срезы данных без дубликации данных.
Настройка обратной записи секции	Предоставляет инструкции по разрешению записи в секции.
Создание удаленной секции и управление ей (Analysis Services)	Содержит описание способа создания и управления удаленной секцией.

Создание пользовательских иерархий

SQL Server 2014

[Другие версии](#)

Службы SQL Server Службы Analysis Services позволяют создавать пользовательские иерархии. Иерархия — это набор уровней на основе атрибутов. Например, иерархия, связанная со временем, может содержать такие уровни, как «Год», «Месяц» и «День». В некоторых иерархиях каждый атрибут однозначно задает атрибут родительского элемента. Такую иерархию иногда называют естественной. Конечные пользователи могут использовать иерархии для просмотра данных в кубе. Иерархии задаются с помощью панели «Иерархии» конструктора измерений в среде разработки SQL Server Data Tools (SSDT).

Вновь созданная пользовательская иерархия может оказаться *неровной*. Неровная иерархия — это иерархия, в которой по крайней мере у одного элемента логический родительский элемент не находится ровно на один уровень выше самого элемента. Существуют настройки, управляющие видимостью недостающих элементов в неровных иерархиях и способом их отображения. Дополнительные сведения см. в разделе [Неоднородные иерархии](#).

Примечание

Дополнительные сведения о проблемах производительности, связанных с конструированием и настройкой пользовательских иерархий, см. в [Руководстве по производительности служб SQL Server 2005 Analysis Services](#).

См. также

[Свойства пользовательской иерархии](#)

[Свойства уровня \[занятого\]](#)

[Измерения «родители-потомки»](#)