**Практикум 5:**

**Построение многомерных диаграмм**

**Ключевые слова:**[многомерная диаграмма](https://intuit.ru/studies/courses/599/455/lecture/10188?page=1#keyword1), [OLAP](https://intuit.ru/studies/courses/599/455/lecture/10188?page=1#keyword2), [CASE](https://intuit.ru/studies/courses/599/455/lecture/10188?page=1#keyword3), [multidimensional](https://intuit.ru/studies/courses/599/455/lecture/10188?page=1" \l "keyword4), [киоски данных](https://intuit.ru/studies/courses/599/455/lecture/10188?page=1#keyword5), [куб данных](https://intuit.ru/studies/courses/599/455/lecture/10188?page=1#keyword6), [физические модели данных](https://intuit.ru/studies/courses/599/455/lecture/10188?page=1#keyword11), [многомерная модель](https://intuit.ru/studies/courses/599/455/lecture/10188?page=1#keyword13), [drilling](https://intuit.ru/studies/courses/599/455/lecture/10188?page=1" \l "keyword15)

* *Многомерная диаграмма*;
* Определение кубов данных;
* Определение измерений;
* Определение атрибутов;
* Определение иерархий;
* Определение метрик;
* Определение ассоциаций.

Рассмотрим методику проектирования кубов данных для *OLAP* хранилищ данных с использованием *CASE* PowerDesigner компании Sybase.

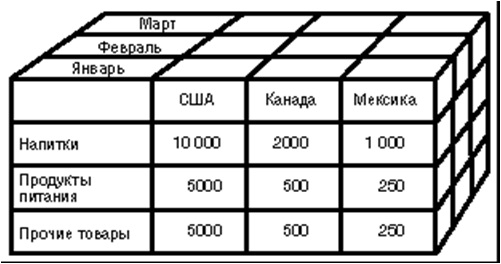
**Многомерная диаграмма**

* **Многомерная диаграмма** (*multidimensional* diagram) представляет собой модель хозяйственной деятельности организации в терминах кубов данных и измерений.
* Многомерные диаграммы используются для проектирования OLAP ХД.
* OLAP ХД заполняются данными из ХД и *киосков данных*. Эти данные преобразуются из реляционного представления в многомерное.

**Пример**

* Данные о продажах (Sales) имеют измерения "Товар" (product), "Регион" (region), "Покупатель" (customer) и "Магазин" (store). Факты, например итоговый объем продаж (sales totals), рассматриваются с точки зрения этих определенных пользователем измерений. Когда аналитик делает выборку об итоговых объемах продаж (sales total) по конкретному товару для конкретного региона, он изучает данные о продажах с точки зрения измерений "Товар" и "Регион". Наиболее часто используемым измерением является время, поскольку основной целью выполнения аналитических запросов является нахождение трендов в данных.

**Куб данных**

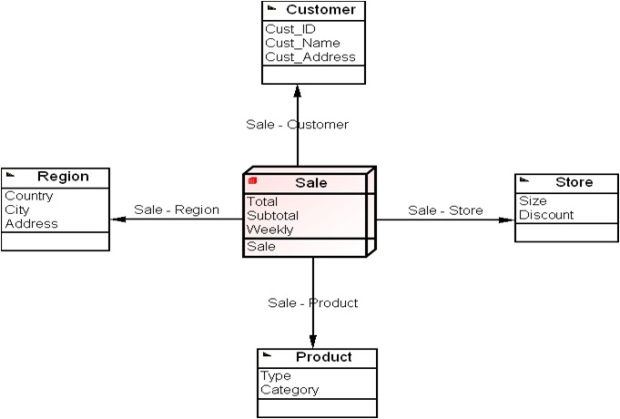


**Элементы многомерной диаграммы**

* **Кубы данных** (cube), содержат набор метрик, которые связаны с различными аспектами хозяйственной деятельности организации и используются для информационной поддержки принятия решений.
* **Измерения** (Dimension), являются своеобразными осями – аспектами исследования данных в кубе.
* **Атрибуты** (Attribute), используются для квалификации измерения.
* **Факты** (Fact), группируют метрики, используемые *кубом данных*.
* **Метрики** (Measure), переменные, как правило, числовые, связываемые с фактом.
* **Иерархии** (Hierarchy), представляет организационную структуру, которая описывает модель доступа к *кубу данных* через измерение.
* **Ассоциации** (Association), устанавливает связь между *кубом данных* и измерением.

Только *кубы данных*, измерения и ассоциации представлены на палитре инструментов *CASE*.

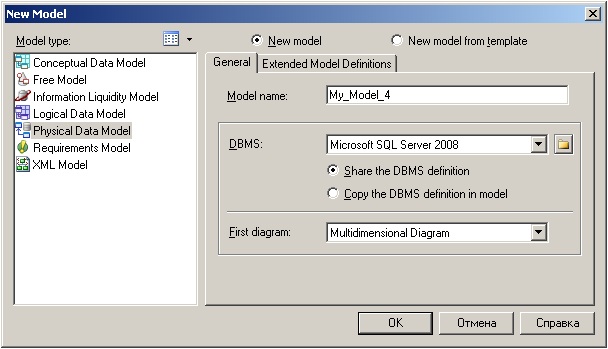
**Многомерная диаграмма**



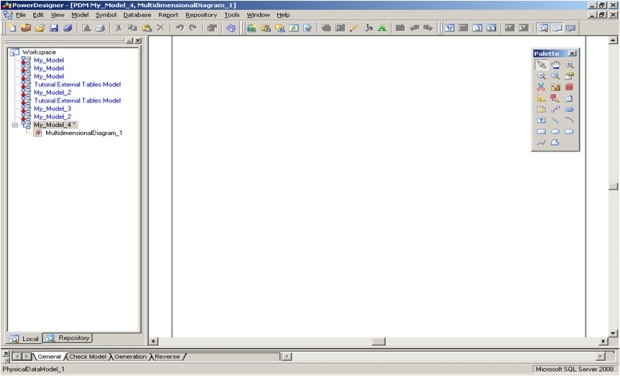
**Создание многомерной диаграммы**

* Выбрать в меню **File->New Model**. На рабочем пространстве появится диалоговое окно "New Model".
* В этом окне нужно выбрать тип модели "*физическая модель данных*", в качестве СУБД мы выберем MS SQL Server 2008 из выпадающего списка СУБД, укажем *многомерную диаграмму*, как класс физической модели, присвоим имя *многомерной модели* (My\_Model\_4) и нажжем кнопку "ОК".
* Таким образом, *многомерная модель* создана, рабочее пространство и палитра инструментов доступны.

**Создание многомерной модели**



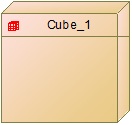
**Создание многомерной модели**



**Кубы данных**

* **Куб данных является набором метрик, соответствующее значение которых сохраняется в каждой его ячейки данных**. Метрики организованы в соответствие с измерениями для того, чтобы выполнять быструю выборку данных или операции сверки-развертки (*drill*-down).
* *Кубы данных* связаны с фактами, которые позволяют определить метрики для куба. На многомерной диаграмме *кубы данных* представляют OLAP кубы.

**Куб данных**



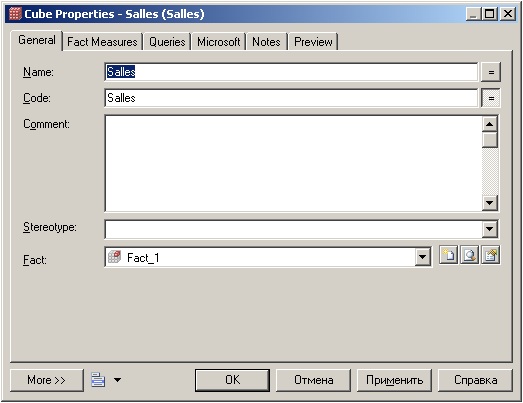
**Определение куба данных**

* Для построения *куба данных* можно использовать палитру инструментов. Выбираем на ней пиктограмму куба и щелчком левой кнопки мыши на рабочем пространстве создаем *куб данных*.
* Далее двойным щелчком мыши на *кубе данных* откроем диалоговое окно для определения свойств *куба данных*

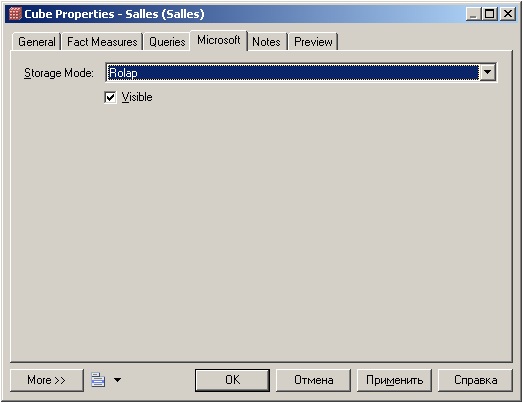
**Определение куба данных**

* Для *куба данных* можно определить следующие свойства:
  + Имя (Name) определяет имя *куба данных*, желательно в терминах понятных пользователям.
  + Код (Code) определяет техническое имя *куба данных*, которое будет использовано при генерировании скрипта.
  + Комментарий (Comment) определяет дополнительное описание *куба данных*.
  + Факт (Fact) определяет факт, используемый в *кубе данных*.
* Присвоим *кубу данных* имя "Продажа" (Sale).

**Определение куба данных**



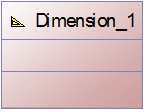
**Определение куба данных**



**Измерения**

* Измерения являются осями для анализа данных в многомерной структуре данных.
* Измерение состоит из упорядоченного списка атрибутов, которые совместно определяют общий семантический смысл (своими значениями) в моделируемой предметной области.
* Каждый атрибут определяет единственную позицию вдоль оси *куба данных*.

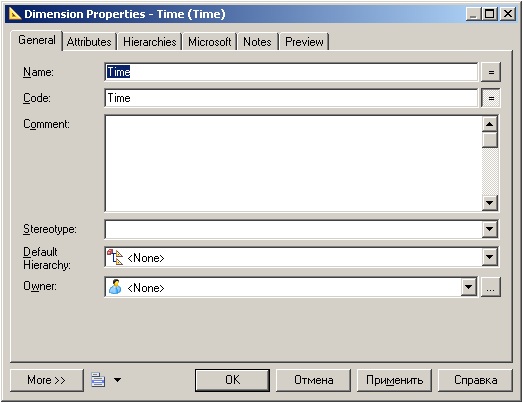
**Измерение**



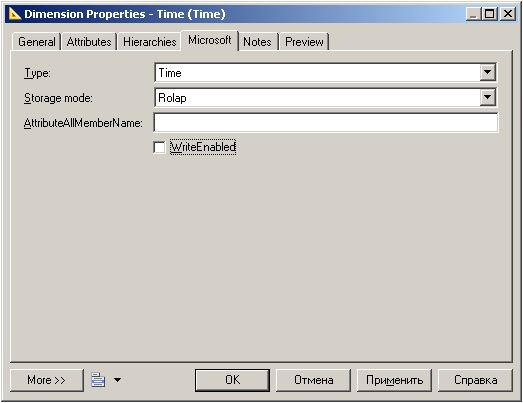
**Определение измерения**

* Для построения измерения можно использовать палитру инструментов. Выбираем на ней пиктограмму измерения и щелчком левой кнопки мыши на рабочем пространстве создаем измерение.
* Далее двойным щелчком мыши на измерении откроем диалоговое окно для определения свойств измерения.

**Определение измерения**



**Определение измерения**



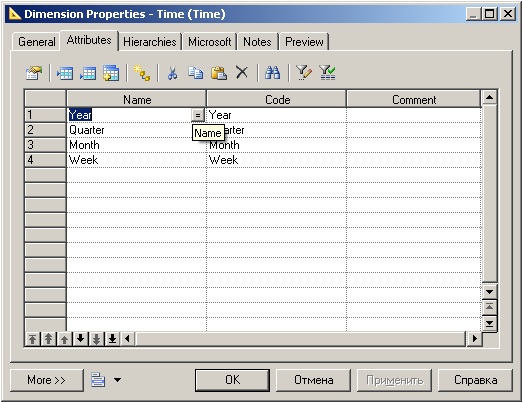
**Определение измерения**



**Атрибуты**

* Атрибуты являются квалификаторами измерений в запросах. Например, Измерения "Время" (Time) может содержать атрибуты "Год". "Квартал", "Месяц", "Неделя". Атрибуты могут быть организованы в иерархии.
* Для создания атрибутов измерения можно использовать диалоговое окно свойств измерения. Например, для измерения "Время" определим следующие атрибуты "Год" (Year), "Квартал" (Quarter), "Месяц" (Month) и "Неделя" (Week).

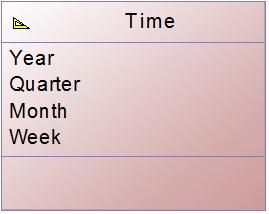
**Определение атрибутов**



**Свойства атрибутов**

* Атрибуты измерения имеют следующие свойства:
  + Имя (Name) определяет имя атрибута в терминах пользователя.
  + Код (Code) определяет техническое имя атрибута, используемое при генерировании скрипта.
  + Комментарий (Comment) определяет дополнительное описание атрибута
  + Измерение (Dimension) определяет измерение для атрибута.

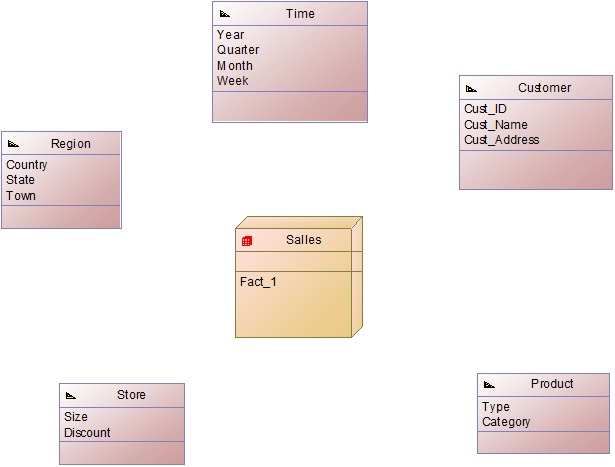
**Определение атрибутов**



**Создание измерений и атрибутов**

* Самостоятельно определите измерения и их атрибуты:
  + "Покупатель" (Customer): атрибуты "Имя покупателя" (Cust\_Name) и "Адрес покупателя" (Cust\_Address) и "Идентификатор покупателя" (Cust\_ID).
  + "Товар" (Product): атрибуты "Тип товара" (Type) и "Категория" (Category).
  + "Регион" (Region): атрибуты "Страна" (Country), "Область" (State) и "Город" (Town).
  + "Магазин" (Store): атрибуты "Размер" (Size) и "Скидка" (Discount).

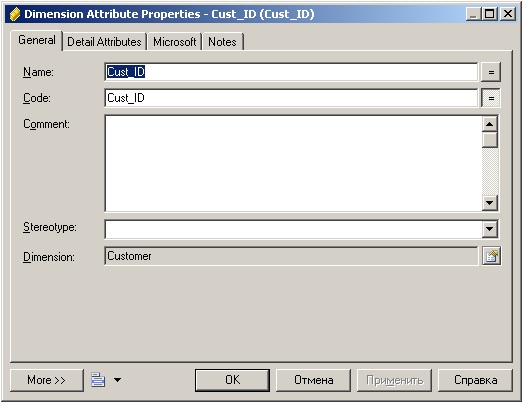
**Задание**



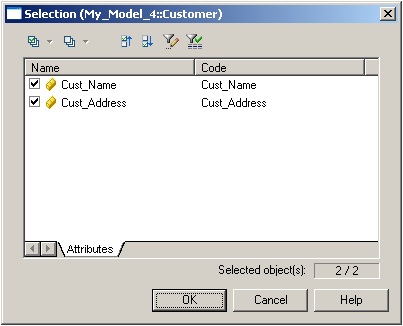
**Определение атрибута через другие атрибуты**

* Атрибут может участвовать в определении другого атрибута, чем самым дополняя определение последнего. Уточняющие атрибуты находятся в списке атрибутов измерения и могут быть использованы в определении другого атрибута.
* В измерении "Покупатель" (Customer) атрибуты "Имя покупателя" (Cust\_Name) и "Адрес покупателя" (Cust\_Address) могут быть использованы для детализации атрибута "Идентификатор покупателя" (Cust\_ID).
* Для этого, нужно в диалоговом окне свойств измерения выбрать список атрибутов, в списке атрибутов выбрать атрибут Cust\_ID, открыть для него диалоговое окно свойств атрибута, на нем выбрать вкладку "Detail Attributes" и занести на нее атрибуты Cust\_Name и Cust\_Address.

**Определение атрибута через другие атрибуты**



**Определение атрибута через другие атрибуты**



**Иерархии**

* Иерархия определяет один или несколько путей доступа к данным через измерение.
* Различают два основных типа иерархий:
  + Организационный путь описывает шаблон обхода измерения, от самого общего к наиболее частному атрибуту. Этот путь упорядочивает множество всех атрибутов измерения.
  + Консолидирующий путь представляет консолидацию атрибутов измерения. Например, у измерения "Время" с основной периодичностью в один день могла бы быть иерархия, определяющая объединение дней в недели, недели в месяцы, месяцы в кварталы, и кварталы в годы.

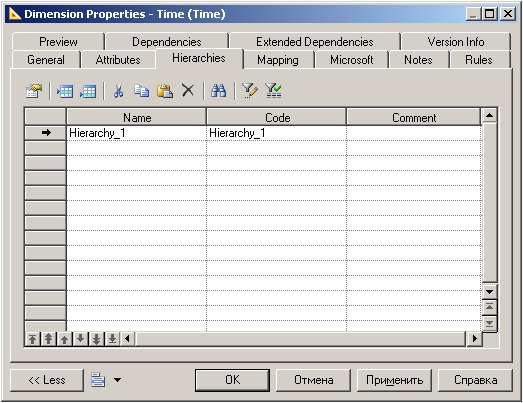
**Свойства иерархии**

* Иерархия имеет следующие свойства:
  + Имя (Name) определяет имя иерархии в терминах пользователя.
  + Код (Code) определяет техническое имя, используемое пр генерировании скрипта.
  + Комментарий (Comment) определяет описание иерархии.
  + Измерение (Dimension) определяет родительское имя иерархии.

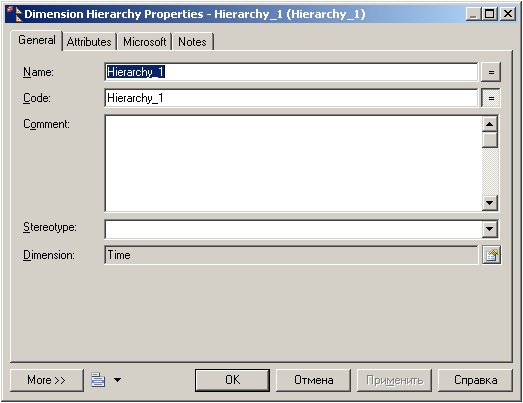
**Определение иерархии в диаграмме**

* Добавить иерархию в измерение можно открыв окно свойств измерения на вкладке "Иерархии" (Hierarchies) с помощью кнопки "Добавить строку".
* Рассмотрим измерение "Время" и создадим для нее иерархию.
* Отрыв диалоговое окно свойств иерархии добавим атрибуты в иерархию.

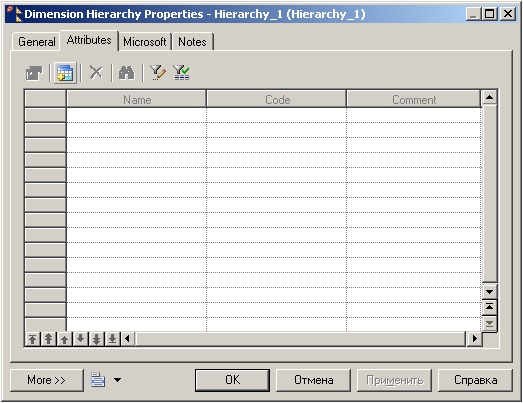
**Определение иерархии**



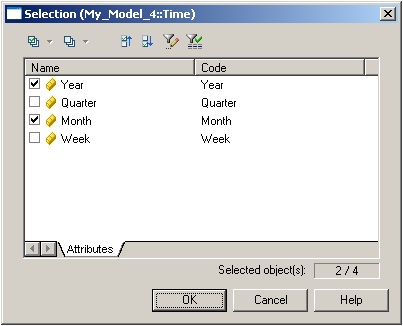
**Определение атрибутов иерархии**



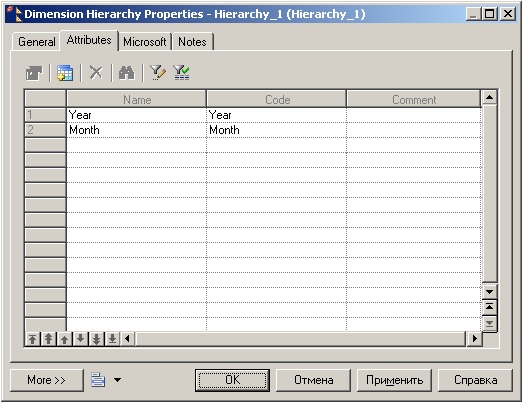
**Определение атрибутов иерархии**



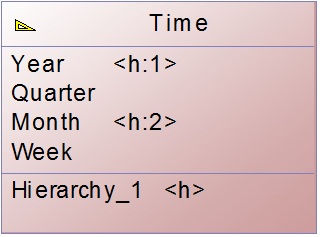
**Определение атрибутов иерархии**



**Определение атрибутов иерархии**



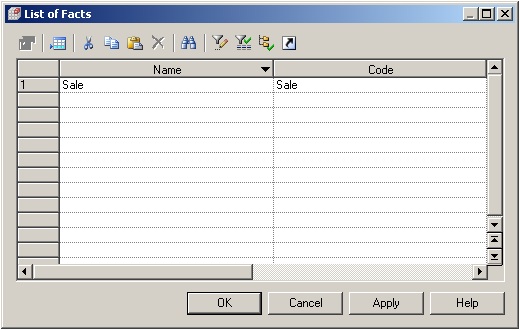
**Определение атрибутов иерархии**



**Факты**

* **Факт соответствует фокусу исследования данных для поддержки принятия решений руководство организации**. Факт – это набор метрик *куба данных*.
* Фактами могут быть, как в нашем примере "Продажи" (Sale), или доходы и бюджет. Одни и те же факты могут использовать в различных *кубах данных*.
* Факт может быть добавлен в *многомерную диаграмму* следующим образом. Выберите пункт меню **Model->Facts** и добавьте его в список в появившемся диалоговом окне "Список фактов" (List of Facts).

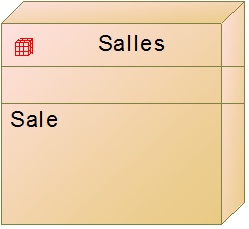
**Определение фактов**



**Свойства фактов**

* Факт имеет следующие свойства:
  + Имя (Name) определяет имя факта в терминах пользователя.
  + Код (Code) определяет техническое имя факта, используемое при генерировании скрипта.
  + Комментарий (Comment) определяет описание факта.

**Определение фактов**



**Метрики**

* Метрика является переменной, которая соответствует фокусу исследования данных. Метрика описывает значение ячейки *куба данных*.
* Например, метрикой часто бывает цена товара или итоговое значение. Метрики могут быть результатом вычислений.

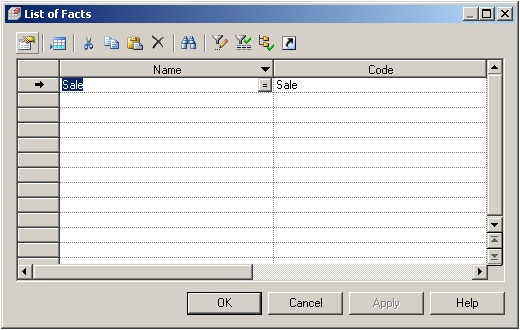
**Свойства метрик**

* Метрика имеет следующие свойства:
  + Имя (Name) определяет имя метрики в терминах пользователя.
  + Код (Code) определяет техническое имя, которое используется при генерировании скрипта.
  + Комментарий (Comment) определяет описание метрики.
  + Факт (Fact) определяет факт, которому принадлежит метрика.
  + Формула (Formula) определяет выражение, посредством которого метрика вычисляется.

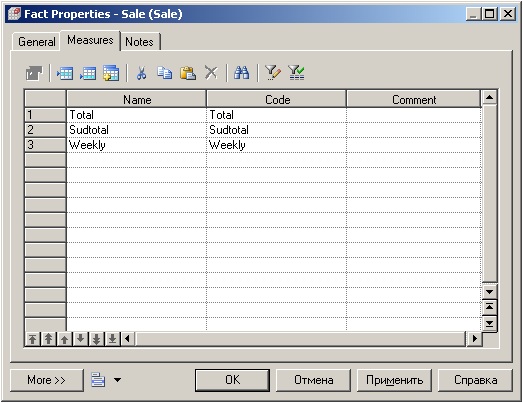
**Определение метрики на диаграмме**

* Для создания метрики можно из главного меню **Model->Facts** в диалоговом окне "Список фактов" (List of Facts) выбрать факт (в нашем случае Sale).
* Вызвать диалоговое окно "Свойства факта" (Fact Properties) , в котором на вкладке "Метрика" (Measure) добавить необходимые метрики.
* В нашем пример добавим следующие метрики "Итог" (Total). "Промежуточный итог" (Sudtotal) и "Еженедельный итог" (Weekly).

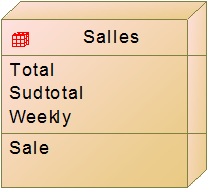
**Определение метрик**



**Определение метрик**



**Определение метрик**



**Ассоциации**

* **Ассоциация связывает куб данных с измерением, которое его определяет**. Ассоциация показывает аспект исследования *куба данных* по указанному измерению.
* Можно связать *куб данных* "Продажи" (Sale) с измерением "Покупатель" (Customer) посредством ассоциации "Продажи"-"Покупатель" (Sale-Customer).
* Допускается только одна ассоциация между измерением и *кубом данных*.

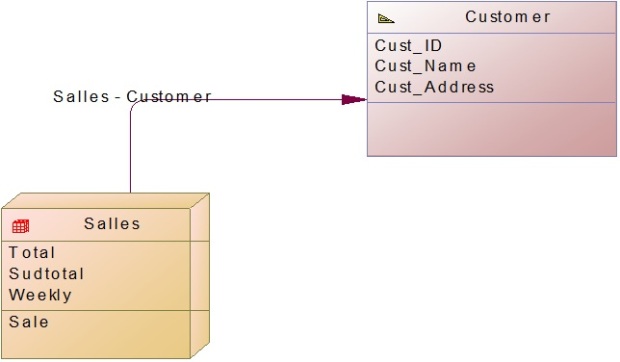
**Свойства ассоциации**

* Ассоциация имеет следующие свойства:
  + Куб (Cube) определяет *куб данных*, который является источников ассоциации.
  + Измерение (Dimension) определяет измерение, с которым связана ассоциация.
  + Иерархия (Hierarchy) определяет иерархию, используемую *кубом данных* для консолидации вычислений.

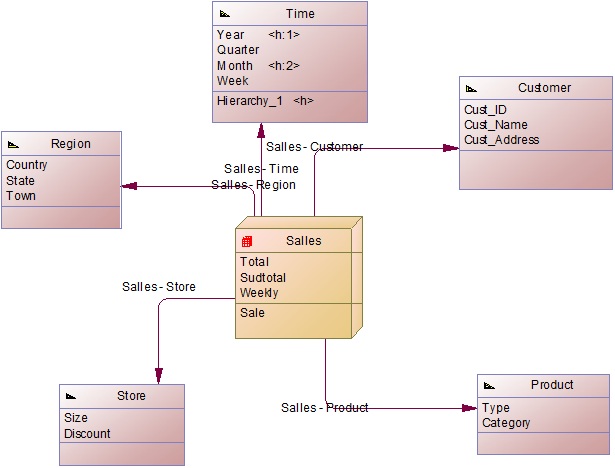
**Определение ассоциации на диаграмме**

* Ассоциацию можно создать при помощи палитры инструментов, выбрав пиктограмму ассоциации.
* Для нашего примера свяжем *куб данных* "Продажи" (Sale) с измерением "Покупатель" (Customer).
* Задание – установить ассоциации между всеми элементами диаграммы.

**Определение ассоциации на диаграмме**



**Многомерная диаграмма**



**Выводы**

* На этом занятии Вы научились строить многомерные диаграммы, которые нужных для создания OLAP хранилищ данных посредством:
  + Определение кубов данных;
  + Определение измерений;
  + Определение атрибутов;
  + Определение иерархий;
  + Определение метрик;
  + Определение ассоциаций.

Однако это далеко не все, что нужно еще сделать для завершения построения модели. Полученное графическое преставление многомерной диаграммы пока еще практически бесполезно. Практические аспекты его использования мы рассмотрим на следующем занятии.