

Лекция №7

Кубы

Модель данных хранилища

Модель данных хранилища – набор кубов.

Куб (Cube) состоит из

измерений (Dimensions) – многомерной системы координат

показателей (Measures) – числовых значений, вычисляемых для каждой комбинации значений координат

Выбранные значения измерений задают ячейку куба (Cell), которая содержит соответствующие значения показателей

Год	(Все)								
Месяц	(Все)								
День	(Все)								
Группа товаров	Табачные изделия								
Товар	(Все)								
Группа товаров	Крупы					л	Единство	Норд	
Товар	(Все)					38		47 574	
						5		300	
Покупатель	Данные	Вест	Геракл	Единство	Норд	31	7 892		
Группа товаров	Кондитерские изделия					5	80		
Товар	(Все)					31			
						51			
Продавец						3 715			
						17			
Покупатель	Данные	Вест	Геракл	Единство	Норд				
Берба	Стоимость	3 263	11 224		2 892				
	Количество	117	271		96				
Витязь	Стоимость	13 507	788		10 555				
	Количество	229	22		254				
Запад	Стоимость	1 702	625						
	Количество	37	24						

Измерения

Измерение (Dimension) – набор значений одного типа для идентификации некоторого свойства бизнес-процесса. Обычно измерение является качественной характеристикой

Значения, измерений, называются метками (members).

Измерение может иметь иерархическую структуру – задавать некоторую классификацию значений, например, группа товаров – товар или год – квартал – месяц – день.

Задание метки верхнего уровня фиксирует несколько меток нижнего. Например, задание группы товаров «Крупы» означает выбор всех товаров входящих в эту группу

Показатели

Показатель является числовой характеристикой бизнес-процесса.

С каждым показателем связывается функция агрегирования (как правило, сумма), задающее правило вычисления одного значения для множества ячеек, по значениям показателя для каждой ячейки.

Значения показателя как правило имеет смысл для множества ячеек в целом, например, стоимость имеет смысл и для отдельной продаж и (ячейки) и для продаж (ячеек) некоторого года.

Иногда показатель (количество) имеет смысл для классов ячеек в отдельности (продажа единиц некоторого товара), но не имеет смысл для их объединения (бессмысленно находить суммарное количество проданных мешков риса и коробок конфет).

Некоторые показатели ячейки (средняя цена) вычисляются по другим показателям ячейки (стоимость, деленная на количество).

Операции над данными в кубах

Формирование среза (Slice) – подмножество куба – куб, получившийся в результате фиксации значения одного или более измерений.

Операция агрегирования (Drill Up) - переход от детализированных данных к агрегированным. Агрегирование получается

при «проекции» куба на меньшее число измерений («свертывании» прочих измерений)

при фиксации обобщающего уровня иерархии измерения.

Операция детализации (Drill Down) - переход от агрегированных к более детальным значениям показателей в результате «развертывания» измерения и выбора более детального уровня измерения.

Формирование среза (Slice)

Год	(Все)						
Месяц	(Все)						
День	(Все)						
Группа товаров	(Все)						
Товар	(Все)						
		Продавец					
Покупатель	Данные	Вест	Геракл	Гермес	Единство	Норд	Плюс
Верба	Стоимость	108 038	215 465			57 575	
	Количество	609	1 236			427	
Витязь	Стоимость	221 158	172 823		22 162		
	Количество	1 014	479				
Запад	Стоимость	145 868	41 417				
	Количество	435	167				
Квадрат	Стоимость		32 187	59 109			
	Количество		293	376			
Кузьмич	Стоимость			234 081			
	Количество			802			
Ландыш	Стоимость		272 968	35 568			
	Количество		1 085	132			
Север	Стоимость		19 650	37 933			
	Количество		213	306			
Экспресс	Стоимость	74 390	163 796	31 774			
	Количество	253	940	203			

Формирование среза
выбором года и группы
товаров

Год	1997						
Месяц	(Все)						
День	(Все)						
Группа товаров	Крупы						
Товар	(Все)						
		Продавец					
Покупатель	Данные	Вест	Геракл	Гермес	Единство	Норд	Плюс
Верба	Стоимость	61 649	49 835			7 108	
	Количество	192	179			31	
Витязь	Стоимость	31 662					
	Количество	101					
Запад	Стоимость						8 411
	Количество						35
Квадрат	Стоимость			38 821	49 948		
	Количество			133	203		
Кузьмич	Стоимость			198 490		48 194	
	Количество			524		156	
Ландыш	Стоимость		237 758		81 868		57 293
	Количество		632		257		146
Север	Стоимость		7 197				
	Количество		24				
Экспресс	Стоимость	72 974				25 630	
	Количество	231				100	

Операция агрегирования (Drill Up)

The diagram illustrates the Drill Up operation in a data cube. It shows a hierarchy of data tables. A red arrow points from a detailed table (bottom-left) to a more aggregated table (top-right). Yellow circles highlight specific filters in the top-right table, and green circles highlight specific data points in the bottom-left table. Green arrows show the flow of data aggregation from the bottom-left table up to the top-right table.

Top-Right Table (Aggregated View):

Год	Месяц	День	Группа товаров	Товар	Продавец	Единство	Норд
(Все)	(Все)	(Все)	Табачные изделия	(Все)		47 574	300
			Крупы	(Все)		7 892	80
			Кондитерские изделия	(Все)		3 715	17

Bottom-Left Table (Detailed View):

Группа товаров	Товар	Покупатель	Данные	Вест	Геракл	Единство	Норд
Кондитерские изделия	Верба		Стоимость	3 263	11 224	2 892	
			Количество	117	271	96	
	Витязь		Стоимость	13 507	788	10 555	
			Количество	223	22	254	
	Запад		Стоимость	1 702	625		
			Количество	37	24		

Bottom-Right Table (Aggregated View):

Год	Месяц	День	Группа товаров	Товар	Продавец	Единство	Норд
(Все)	(Все)	(Все)	(Все)	(Все)		57 575	427
			Кондитерские изделия	(Все)		22 162	351
			Табачные изделия	(Все)		77 229	300

Операция детализации (Drill Down)

Выполняется детализация –
развертывается измерение
«Продавец»

Год	(Все)	▼		
Месяц	(Все)	▼		
День	(Все)	▼		
Группа товаров	(Все)	▼		
Товар	(Все)	▼		
Продавец	(Все)	▼		
Покупатель	▼	Данные	▼	Итого
Верба	Стоимость	495 135		
	Количество	3 119		
Витязь	Стоимость	651 992		
	Количество	3 208		
Запад	Стоимость	576 789		
	Количество	2 693		
Север	Стоимость	318 223		
	Количество	1 862		
Экспресс	Стоимость	417 698		
	Количество	2 518		

Год	(Все)	▼												
Месяц	(Все)	▼												
День	(Все)	▼												
Группа товаров	(Все)	▼												
Товар	(Все)	▼												
			Продавец	▼										
Покупатель	▼	Данные	▼	Вест	Геракл	Гермес	Единство	Норд	Плюс	Ромашка	СибАтом	Тор	Успех	Фокус
Верба	Стоимость			108 038	215 465			57 575		38 491		19 521	56 047	
	Количество			609	1 236			427		346		132	369	
Витязь	Стоимость			221 158	172 823		22 162				104 388		84 158	47 304
	Количество			1 014	479		351				418		660	286
Запад	Стоимость			145 868	41 417				12 524	175 752	80 347			120 881
	Количество			495	167				160	878	363			630
Север	Стоимость				19 650	37 933				9 080		133 634		117 926
	Количество				213	306				200		537		606
Экспресс	Стоимость			74 390	163 796	31 774	77 229	38 748		31 761				
	Количество			253	940	203	458	428		237				

