**Ключевые слова:**[многомерная модель](https://intuit.ru/studies/courses/599/455/lecture/10171?page=1#keyword1), [гранулированность](https://intuit.ru/studies/courses/599/455/lecture/10171?page=1" \l "keyword2), [запрос](https://intuit.ru/studies/courses/599/455/lecture/10171?page=1#keyword6), [аддитивные факты](https://intuit.ru/studies/courses/599/455/lecture/10171?page=1#keyword8), [average](https://intuit.ru/studies/courses/599/455/lecture/10171?page=1#keyword9), [quantity](https://intuit.ru/studies/courses/599/455/lecture/10171?page=1#keyword10), [hand](https://intuit.ru/studies/courses/599/455/lecture/10171?page=1#keyword11), [actual](https://intuit.ru/studies/courses/599/455/lecture/10171?page=1#keyword12), [fact](https://intuit.ru/studies/courses/599/455/lecture/10171?page=2#keyword27), [TIME](https://intuit.ru/studies/courses/599/455/lecture/10171?page=2#keyword30), [percentage](https://intuit.ru/studies/courses/599/455/lecture/10171?page=2#keyword34), [outlet](https://intuit.ru/studies/courses/599/455/lecture/10171?page=2#keyword35), [bold](https://intuit.ru/studies/courses/599/455/lecture/10171?page=2#keyword47), [BOND](https://intuit.ru/studies/courses/599/455/lecture/10171?page=2#keyword49), [total](https://intuit.ru/studies/courses/599/455/lecture/10171?page=2#keyword51), [AND](https://intuit.ru/studies/courses/599/455/lecture/10171?page=2#keyword52), [level](https://intuit.ru/studies/courses/599/455/lecture/10171?page=2#keyword56), [manufacturing](https://intuit.ru/studies/courses/599/455/lecture/10171?page=2#keyword59)

* Тема: на основе определенных в процессе анализа требований кандидатов в измерения, показатели и факты построить эскиз *многомерной модели* ХД.
* Задачи:
	+ Измерения;
	+ Показатели;
	+ Факты;
	+ Пересмотр показателей и измерений (агрегация и грануллированность);
	+ Объединение фактов.

**Зачем нужно анализировать бизнес-требования**

* Во-первых, это все равно придется делать, поскольку одной из главных целей проекта является удовлетворение таких требований.
* Во-вторых, очень важно, когда это делать и как. Если серьезный анализ требований отнести на стадию, скажем разработки приложений, то может возникнуть ряд обратных задач проектирования, решение которых может привести к изменению структуры модели.
* Например, показатель Средний остаток может быть заменен на Количество перемещенного запаса, что повлечет внесение изменений в результаты всех последующих стадий проектирования и разработки ХД.

**Построение эскиза многомерной модели ХД**

* Эта задача решается в пять этапов:
* Определение измерений;
* Определение показателей (агрегация и *гранулированность*);
* Определение фактов;
* Анализ фактов, показателей и измерений.
* Построение эскиза *многомерной модели*.

Определив потенциальные измерения, показатели и факты, проектировщик ХД на следующем этапе может приступить к созданию эскиза *многомерной модели* данных.

**Измерения**

* На предыдущем семинаре мы определили кандидаты в измерения.
* На этапе создания эскиза *многомерной модели* проектировщик ХД должен решить, какие измерения должны быть в модели.
* Учитывая тот факт, что ХД, как правило создается для хранения и анализа исторических данных о хозяйственной деятельности организации, необходимо добавить в модель измерение "Время".

**Измерения модели ХД**



**Факты**

* Факты строятся на основе выбранных измерений и показателей (метрик).
* Построение фактов есть процесс группировки измерений и показателей иаким образом, чтобы можно было удовлетворить бизнес – требованиям.
* Первоначально проектировщик ХД фожет создать один факт для всех запросов.
* Для каждого набора показателей точно связанных с одними и теми же измерениями создадим отдельный факт.

Заметим, что вопросы 6, 8 и 9 не имеют показателей, связанных с ними. Не объединить ли *запрос* 6 с запросами 5 и 7 в факт 4, а запросы 8 и 9 с запросом 2 в факт 2, тем самым избавившись от фактов без показателей. (такие факты называются factless facts, потому что они фиксируют событие: продажу товара в определенный момент времени (факты 2 и 3) в определенном месте (факт 2 только). Никакие показатели для них не требуются.

**Факты модели ХД: запросы 1 и 3**



**Факты модели ХД: запросы 2,8 и 9**



**Факты модели ХД: запросы 5,6 и 7**



**Факты модели ХД: запрос 4**



**Грануллированность, аддитивность и объединение фактов**

* Граннулированность есть уровень детализации факта, с которой он будет храниться в ХД.
* Аддитивность – это возможность суммировать показатели в ХД. Тесно связана с *гранулированностью*.
* Факты бывают аддитивные, неаддитивные и полуаддитивные.

**Аддитивность фактов**

* Факты бывают аддитивные, неаддитивные и полуаддитивные:
	+ Проценты – неаддитивные факты (бессмысленно суммировать);
	+ Баланс - полуаддивный факт (можно складывать только в определенный момент времени, но не в различные);
	+ Доход – *аддитивный факт* (можно суммировать по всем измерениям).

**Гранулированность и аддитивность фактов**

* Рассмотрим FACT 1:
	+ average quantity on hand определяется ежемесячно;
	+ total cost и total revenue определяются ежедневно.
* Следовательно, либо факт нужно разбить на два факта, либо изменить измерение Time (ввести уровень гранулированности – день).

*average* *quantity* on *hand* – неаддивный показатель, необходимо хранить *actual* *quantity* on *hand* и позже в запросе вычислять среднее. .Лучше сделать этот показатель аддитивным.

**Гранулированность и аддитивность фактов**

* Рассмотрим FACT 2: он имеет проблему с измерением TIME для запроса 2 (ежедневно) и запросов 8 и 9 (ежемесячно):
* Следовательно, нужно изменить измерение Time (ввести уровень гранулированности – день). Это таже позволит суммировать по месяцам.

*average* *quantity* on *hand* – неаддивный показатель, необходимо хранить *actual* *quantity* on *hand* и позже в запросе вычислять среднее. .Лучше сделать этот показатель аддитивным.

**Гранулированность и аддитивность фактов**

* Рассмотрите факты FACT 3 и 4 самостоятельно.

*average* *quantity* on *hand* – неаддивный показатель, необходимо хранить *actual* *quantity* on *hand* и позже в запросе вычислять среднее. .Лучше сделать этот показатель аддитивным. Целесообразно все показатели сделать аддитивными.

**Факты. Промежуточный итог**



*FACT* 1

*Quantity* on *hand*

*Time*

*FACT* 2

*Time*

*FACT* 3

*percentage* of models sold througth retail *outlet* на *Quantity* of models sold througth retail *outlet*

*percentage* of models sold through sales office order desk на *Quantity* of models sold through sales office order desk

*percentage* of models sold through sales office salesperson на *Quantity* of models sold through sales office salesperson

*Time*

*FACT* 4

*percentage* of models eligible for discount на Number of models eligible for discount

*percentage* of models eligible for discount actually discounted на

*Quantity* of models eligible for discount actually *bold*

*Quantity* of models *bond* at a discount

**Объединение фактов**

* FACT 1 можно оставить без изменения (почему?).
* У FACT 2 те же измерения, что у 3 и 4. Можно исследовать возможность объединить эти факты.
* FACT 2 и FACT 3 можно объединить.
* Новый FACT 2 можно объединить с FACT 4.
* Добавим еще одно измерение "Продавец" (Seller).

Рассмотрим *fact* 1: ясно, что *total* и *and* *total* revenue не может быть испорчен измерением sales. Однако это не так для *quantity* on *hand* или reorder *level*. Фактически *quantity* on *hand* не следует трогать, поскольку с измерениями product и *manufacturing* для него все нормально. Не будем изменять этот факт.

*Fact* 2 можно объединить с *Fact* 3, добавляя измерение sales к *Fact* 2.

**Объединение фактов**



Рассмотрим *fact* 1: ясно, что *total* и *and* *total* revenue не может быть испорчен измерением sales. Однако это не так для *quantity* on *hand* или reorder *level*. Фактически *quantity* on *hand* не следует трогать, поскольку с измерениями product и *manufacturing* для него все нормально. Не будем изменять этот факт.

*Fact* 2 можно объединить с *Fact* 3, добавляя измерение sales к *Fact* 2.

**Факты модели ХД**



* Из факта 1 получим Inventory fact.
* Из объединенного факта 2 получим sales fact.
* Таким образом, мы имеем два киоска данных.

**Первый киоск данных**



**Второй киоск данных**



**Выводы**

* На этом семинаре мы закончили изучение простого бизнес-кейса по проектированию Хранилища данных, на котором рассмотрели, как на основе сбора и анализа бизнес - требований, корпоративной модели данных с помощью несложных приемов разработать эских схемы ХД.
* На последующих практикумах мы изучим методы проектирования ХД на основе графических нотаций CASE-средств проектирования.