# 03-B
# A QMRT

## A Quick Multidimensional Data Modeling Run-Through (QMRT)

# A QMRT

- from the article, "A Quick Multidimensional Modeling Run-Through", Haug (2011)
  - this works through a "not-a-real-world" example
  - has **No** Prescription for integration across MDMs
  - I recommend you use this approach for your M2 Project activities, but you are free to use any other approach that works for you…
  - I suggest you look at the MDM examples in the Kimball books (required and optional) if you want further practice

# A QMRT

- this is not really a formal approach as much as my informal description of the common steps taken by many different sources…
  - no real "name" for the approach
  - similar to Kimball's approach but somewhat different
  - similar to traditional OLTP / ERD approaches
  - eight basic steps
  - better suited to our project, examples, and exams than for the real world…
- we will list the steps and then look at each step a little more closely

# A QMRT

1. Identify Business Process or Business Area
2. Identify and Classify Important Concepts
3. Further Classify Concepts from #2 as either:
   - **FACTs**, **DATTs**, **DIMs**, etc.
4. Identify Grain for Desired Analysis

# A QMRT

5. Define High-level (coarse) Contextual Detail Grouping [**DIMs**]
6. Define Lower-level (fine) Contextual Detail Groupings [**HIERs**, **LVLs**, and **PROPs**]
7. Identify Slowly Changing Dimensions (SCD) and choose strategies for handling them ..
   Then re-evaluate model.
8. Look at OLTP sources for any additional "good" **FACTs**, **DATTs**, **DIMs**, **HIERs**, **LVLs**, etc.

# A QMRT

1. Identify Business Process or Business Area
   - usually we know this already, it is merely a formality to document it.
   - Examples:
     - Customer makes a Purchase
     - Sales transaction at a Point of Sale
     - Company makes a Shipment
   - only One Process / Area per Star Schema
   - repeat these 8 steps for each Star Schema
   - Document and capture this in a usable format!

# A QMRT

## 2. Identify Important Concepts / Aspects

- don’t worry about levels of abstraction yet
- don’t worry about "is this a fact or not?" yet
- don’t worry about "where does this go?" yet
- if the concept can "take on a value directly"
  - determine numeric/text
  - determine discrete/continuous
  - determine bounded/unbounded
- we can clarify / complete this analysis in step #3 if necessary
- Document and capture this in a usable format!

# A QMRT

## 2. Identify Important Concepts / Aspects

### Example Format -A

| # | Category Name | Concept Name | Concept Type<br>DIM / HIER / LVL<br>MKEY / MNAM / MPROP / FACT | Type<br>Instance<br>or Larger Concept | N<br>—or—<br>T | C<br>—or—<br>D | B<br>—or—<br>U | O<br>—or—<br>K | Notes<br>this can contain any notes you want to add |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Product | Make | MPROP | | T | D | B | K | Example only |
| 2 | Product | Packrat | | | T | D | B | K | Delete the contents |
| 3 | | | | | | | | | If you Fill Down or Copy / Paste |
| 4 | | | | | | | | | Then the Data Validation |
| 5 | | | | | | | | | Will also be available |
| 6 | | | | | | | | | for new rows |
| 7 | | | | | | | | | You can also go to |
| 8 | | | | | | | | | Data->Validation |
| 9 | | | | | | | | | And Change the Validation |
| 10 | | | | | | | | | Details to match what |
| 11 | | | | | | | | | You want to use |
| 12 | | | | | | | | | |
| 13 | | | | | | | | | |
| 14 | | | | | | | | | |
| 15 | | | | | | | | | |
| 16 | | | | | | | | | |
| 17 | | | | | | | | | |
| 18 | | | | | | | | | |
| 19 | | | | | | | | | |
| 20 | | | | | | | | | |

# A QMRT

## 2. Identify Important Concepts / Aspects

### Example Format -B

| # | Category | Concept | Type or Instance | Numeric or Text | Continuous or Discrete | Bounded or Unbounded | Observed or Knowable | Notes |
|---|---|---|---|---|---|---|---|---|
| 1 | Product | Make | Type | Text | Discrete | Bounded | Knowable | Example only |
| 2 | Product | Packrat | Instance | Text | Discrete | Bounded | Knowable | Delete the contents |
| 3 | | | | | | | | If you Fill Down or Copy / Paste |
| 4 | | | | | | | | Then the Data Validation |
| 5 | | | | | | | | Will also be available |
| 6 | | | | | | | | for new rows |
| 7 | | | | | | | | You can also go to |
| 8 | | | | | | | | Data->Validation |
| 9 | | | | | | | | And Change the Validation |
| 10 | | | | | | | | Details to match what |
| 11 | | | | | | | | You want to use |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |
| 16 | | | | | | | | |
| 17 | | | | | | | | |
| 18 | | | | | | | | |
| 19 | | | | | | | | |
| 20 | | | | | | | | |

**or create your own format!**

# A QMRT

3. Classify as Observed / Analyzed Data or Contextual Detail
   - should be fairly automatic after steps #1 and #2
   - does not mean we should "stop thinking critically"
   - it is ok to think about **AGGs**, but **AGGs** are not **FACTs** and do not really belong here…
   - obviously this step combined with #2 allows us to categorize concepts as **FACTs**, **DATTs**, or "larger concepts"
   - Document and capture this in a usable format!

# A QMRT

## 4. Identify Grain for Desired Analysis

- often refined across multiple steps / iterations
- IOW, we probably need to redefine the grain as we change the **DIMs**, **LVLs**, etc. in the MDM
- our Goal in this step is to use as **Explicit** language as possible when identifying the grain
  - for example saying:

    "per Minute, per SKU, per Store_ID"

    is **BETTER THAN** saying:

    "by time, product, and location"
- Document and capture this in a usable format!

# A QMRT

5. Define High-level (coarse) Contextual Detail Grouping [**DIMs**]
   - similar to the way we differentiate Attribute Types and Entity Types in traditional ERD modeling
   - sometimes there will be synonyms here
   - we are **GROUPING** the attributes defined in Step#3 – **NOT** inventing new **DIMs** / **DATTs** out of thin air!
   - each **DATT** must be in one and only one **DIM**
   - each **DIM** must have at least one **DATT**
   - each **DIM** potentially changes the grain from Step#4
     - IOW, revisit Step#4 after this is completed
   - Document and capture this in a usable format!

# A QMRT

6. Define Lower-level (fine) Contextual Detail Groupings [hierarchies, levels, and member properties]
   - this is similar to the traditional modeling analysis:
     - ERD Attribute Types into Attribute Groups
     - EERD Entity Types into Sub Types and Super Types
     - Functional Dependencies and Normalization
   - each **DIM** has one or more **HIER**
   - each **HIER** has one or more **LVL**
   - each **LVL** has one or more **DATT**
   - each **DATT** is in one and only one **LVL** in a **HIER**
   - each **DATT** is in each and every **HIER**
   - Document and capture this in a usable format!

# A QMRT

7. Identify Slowly Changing Dimensions (SCD) and choose strategies for handling them ..
   Then re-evaluate model.
   - Document and capture this in a usable format!
   - Document the implications too!
   - revisit Step #4
   - reconsider Step #5 & #6
     - IOW, consider if **DIMs** need to be merged or split (e.g. Mini-Dimensions)
     - if they do, revisit Step #4, #5, #6, and #7 **YET AGAIN!**

# A QMRT

8. Look at OLTP sources for any additional "good" **FACTs**, **DATTs**, **DIMs**, **HIERs**, **LVLs**, etc.
   - but don’t go overboard here.
   - be careful! Consider the:
     - Grain
     - Additivity
     - Creeping Featurism
     - Creeping OLTP’isms
   - depending upon what you include
     - Revisit Steps #2, #3, #4, #5, #6, & #7!

# "Lifecycle Flowchart"