DaWaK '09

Visual Modelling of Data Warehousing Flows with UML Profiles

Jesús Pardillo[1], Matteo Golfarelli[2],
Stefano Rizzi[2], Juan Trujillo[1]

[1] University of Alicante, Spain
[2] University of Bologna, Italy

Introduction

- Data warehousing ≡ data transformations
  - Complexity
- "Model-driven development" ≈ Visualisation
  - Perception & cognition

Outline

- UML-based Visual Language for F^w
  - Formalisation
  - UML Mapping
  - Diagrams
    - data cubes + data flows

Definition

Let C be the set of all data cubes,

and X be the set of the data types
that are not data cubes,

A data-warehousing flow is
a function f^w: {C,X}→{C,X}

F^w is the set of all f^w's

F^W Characterisation

Subject	Class	Definition
OLAP	F^C	$C \rightarrow C$
ETL	F^{+C}	$X \rightarrow C$
Mining	F^{-C}	$C \rightarrow X$
Object	$F^{\varnothing}$	$X \rightarrow X$

F^W Characterisation

The archetypical data warehousing $f^W \in F^W$ holds

$$f^W = k^{\varnothing} \bullet j^{-c} \bullet i^{c} \bullet h^{+c} \bullet g^{\varnothing}$$

where

$k^{\varnothing}, g^{\varnothing} \in F^{\varnothing}$

$h^{+c}, i^{c}, j^{-c} \in F^{+c}, F^{c}, F^{-c}$

Visualisation Requirements

(i) supported by multidimensional diagrams

(ii) understandable

(iii) formal

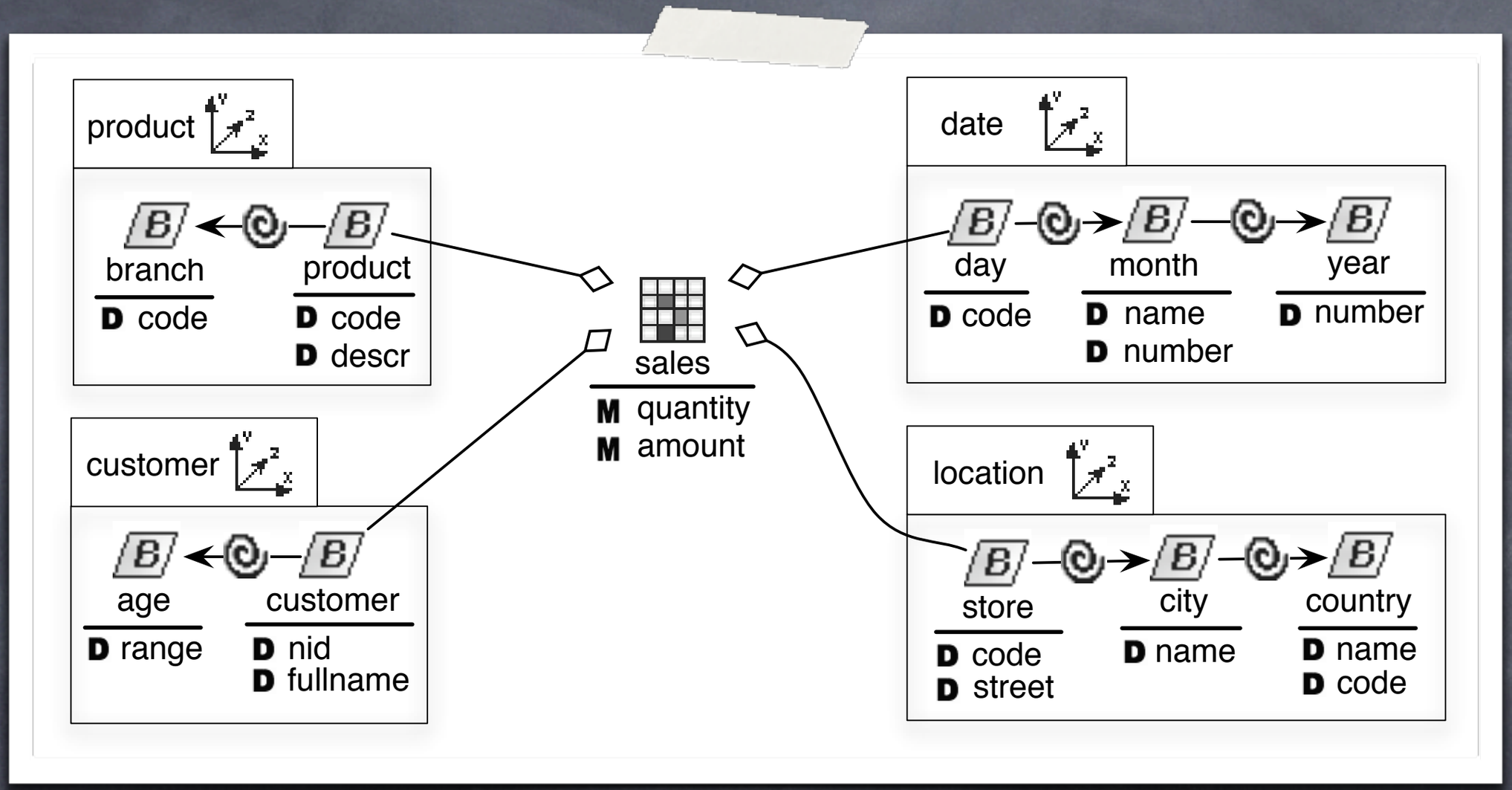
(iv) standard

Concerns

(i) Data Cubes
  @ Multidimensional Diagrams
  @ UML Class Diagrams

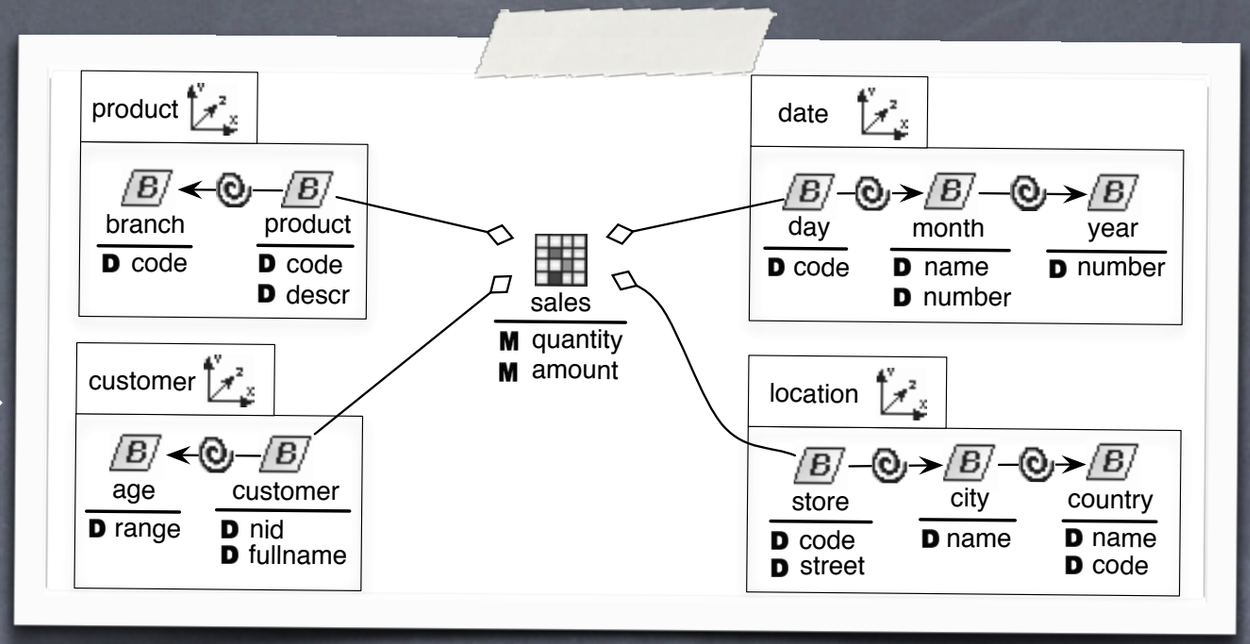
(ii) Data-warehousing Flows
  @ UML Activity Diagrams

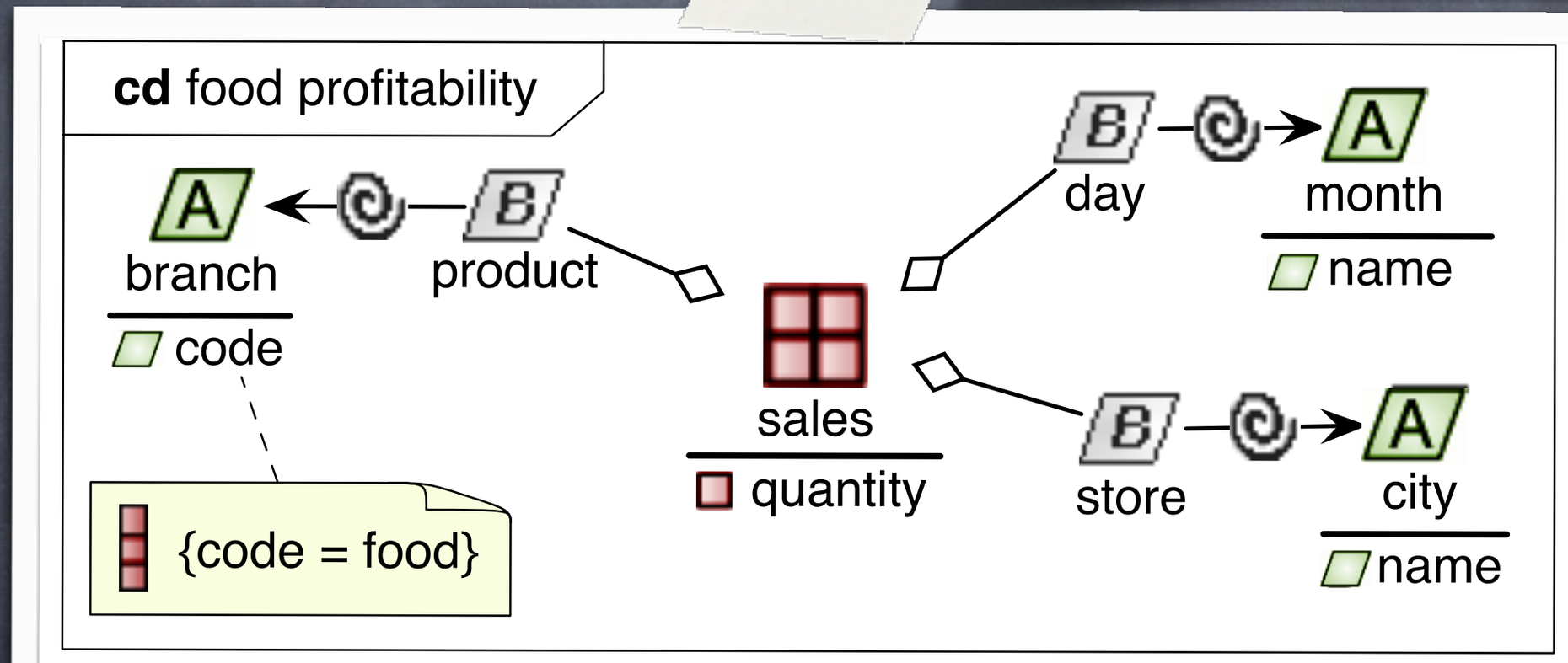
Visual Modelling of F^w with UML Profiles



Multidimensional Diagram

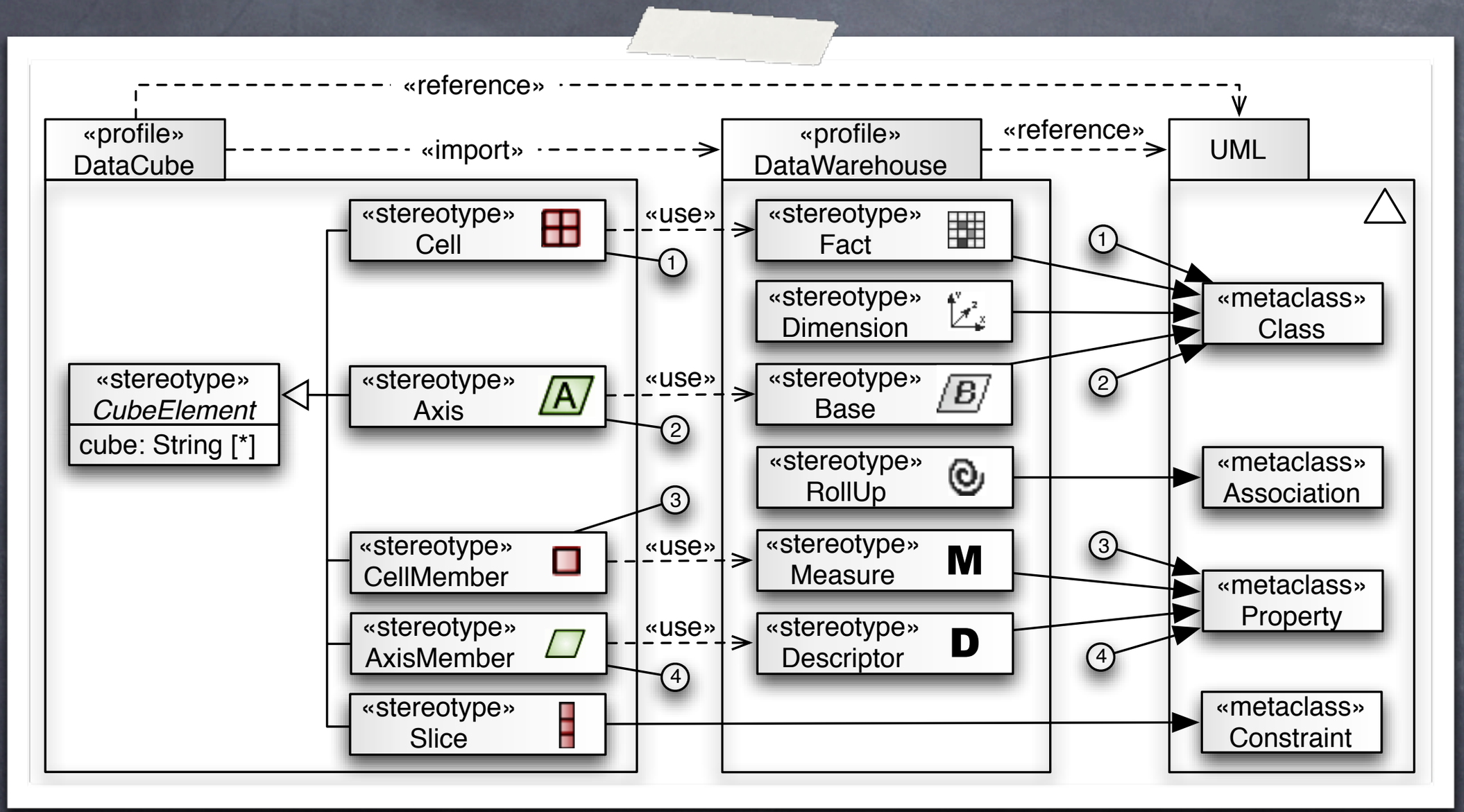
Visual Modelling of F^w with UML Profiles





Data Cube Diagram

Visual Modelling of F^w with UML Profiles



UML Profile for Data Cubes

Specification Levels

Space	Membership
<cell> @ <fact> <axis> @ <base>	<cellmember> @ <measure> <axismember> @ <descriptor>

Space Interpretation:

(a) all measures & descriptors retrieved
(b) measures & descriptors unknown

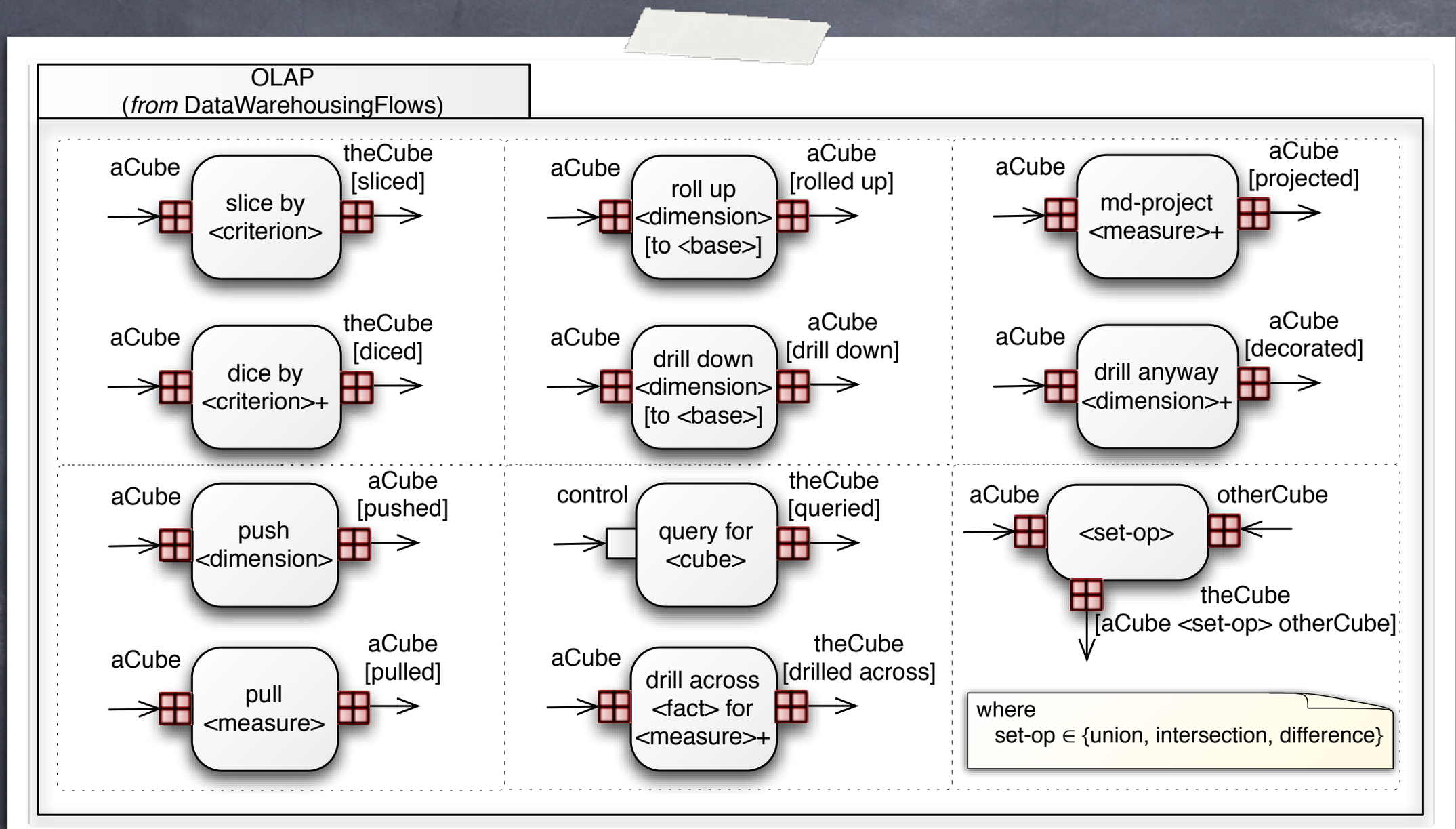
Diagramming Data Cubes

1. Copy the multidimensional diagram
2. Rename it as a data cube diagram
3. Specify data CubeElement's
   1. Stereotype multidimensional elements
   2. Tag them with the current cube
4. Hide undesired multidimensional elements

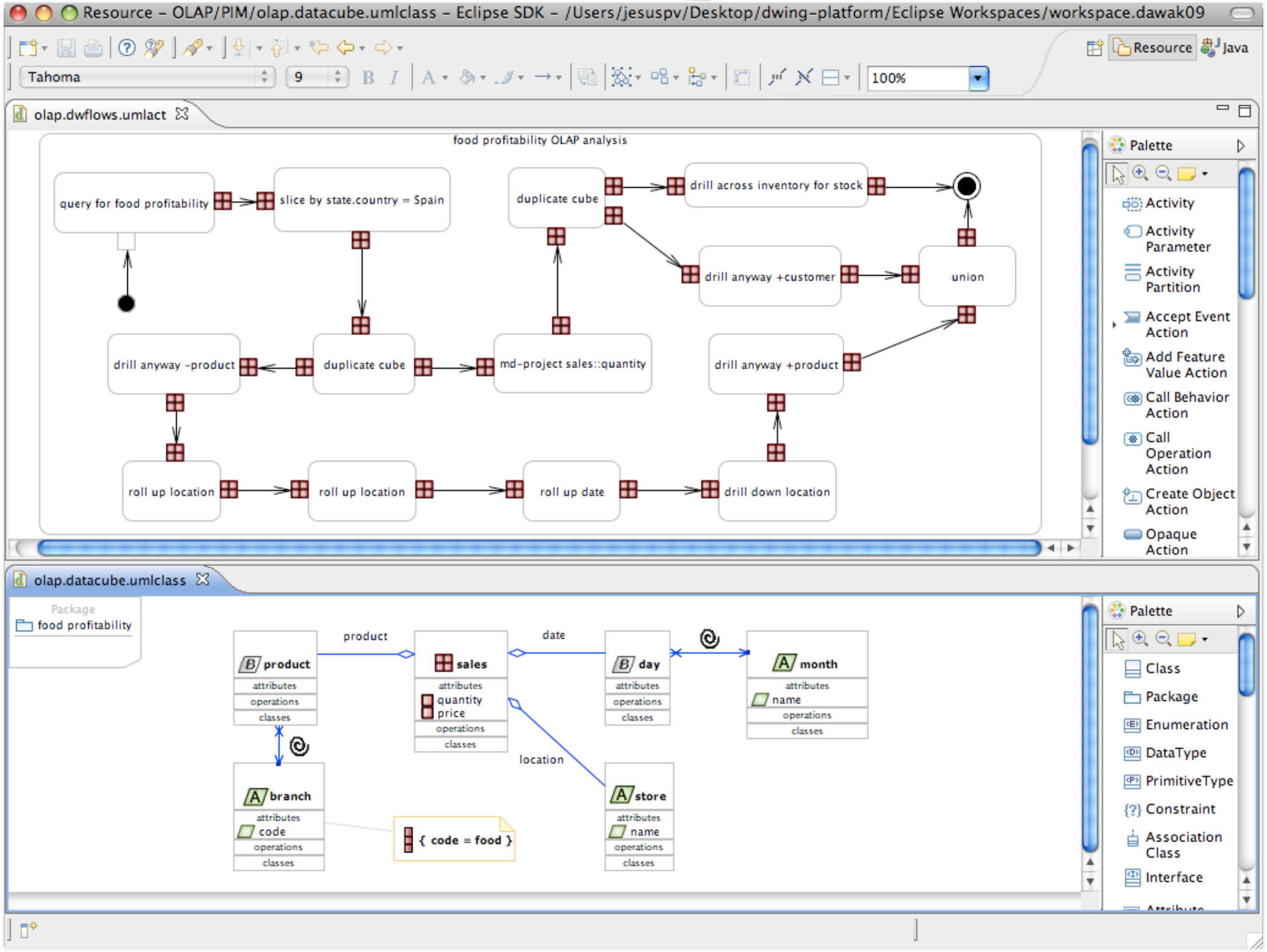
Fw Visual Library

- A UML Profile for Fw Library
  - naming patterns
  - iconography
- A Fw Library
  - Catalogues:
    ETL, OLAP, Mining, What-if, OLAM, etc.

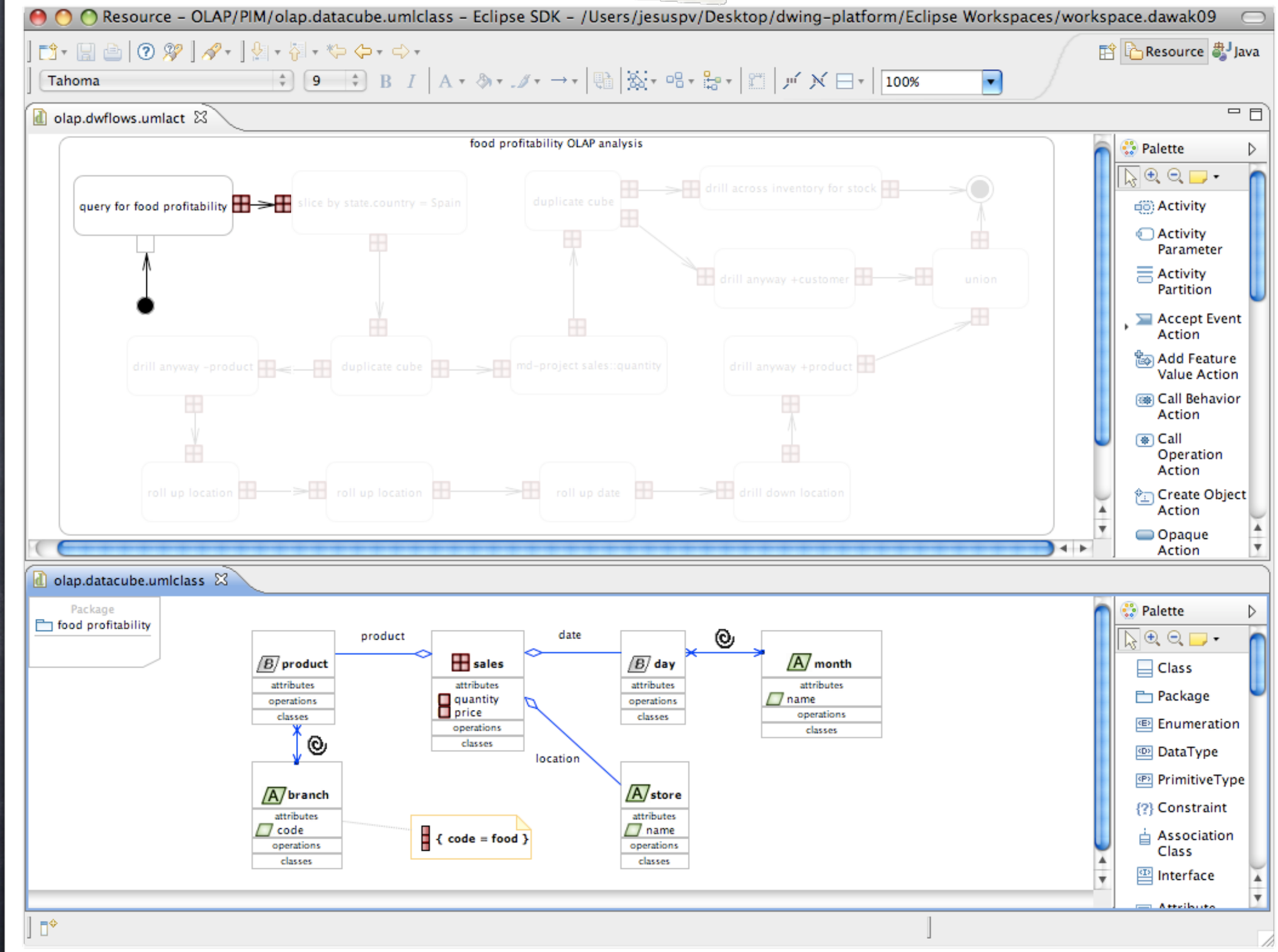
Visual Modelling of F^W with UML Profiles



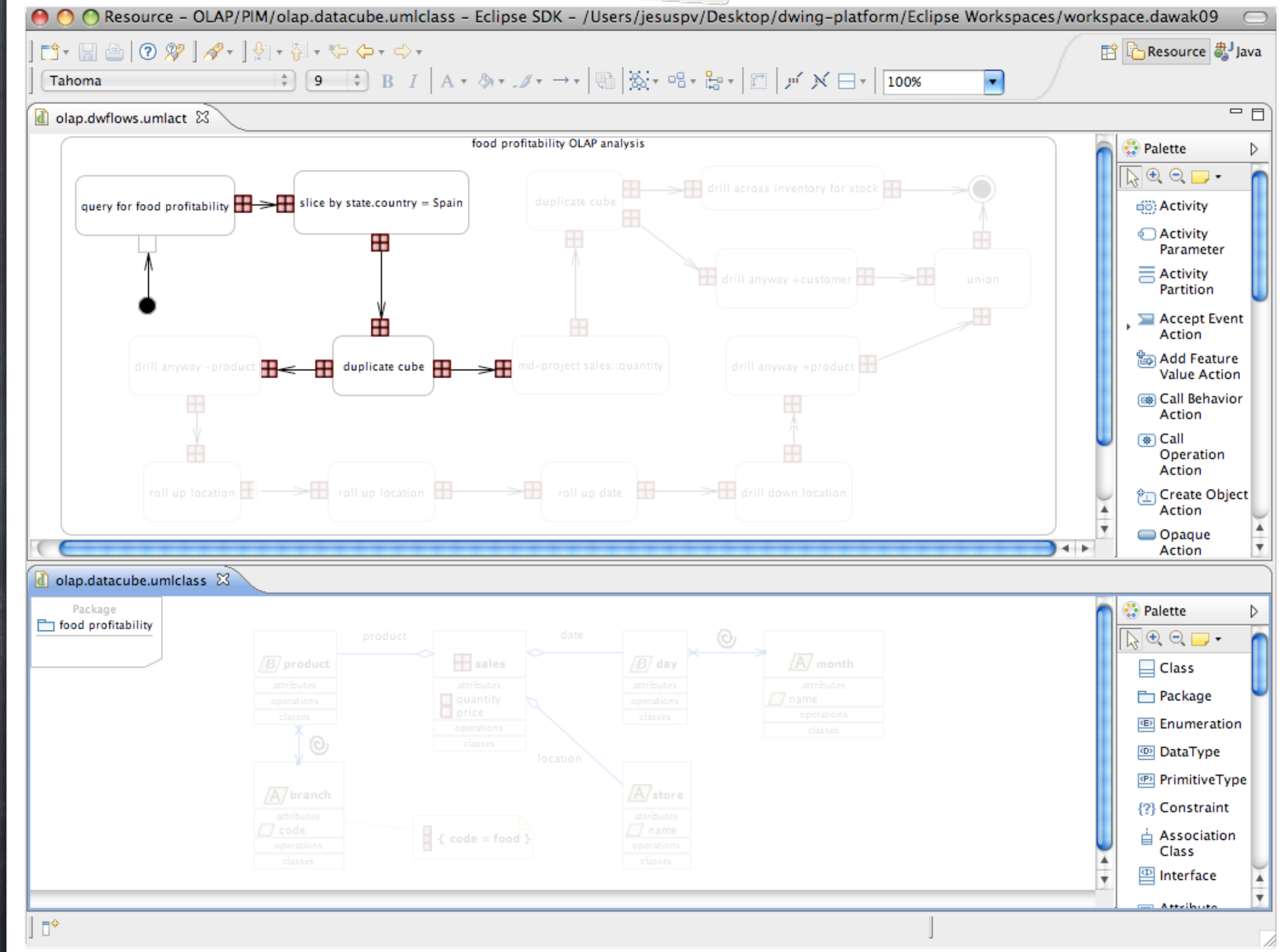
F^C Catalogue


Resource – OLAP/PIM/olap.datacube.umlclass – Eclipse SDK – /Users/jesuspv/Desktop/dwing-platform/Eclipse Workspaces/workspace.dawak09
Resource
Java
Tahoma
9
100%
olap.dwflows.umlact
food profitability OLAP analysis
query for food profitability
slice by state.country = Spain
duplicate cube
drill across inventory for stock
drill anyway +customer
union
drill anyway -product
duplicate cube
md-project sales::quantity
drill anyway +product
roll up location
roll up location
roll up date
drill down location
Palette
Activity
Activity Parameter
Activity Partition
Accept Event Action
Add Feature Value Action
Call Behavior Action
Call Operation Action
Create Object Action
Opaque Action
olap.datacube.umlclass
Package
food profitability
product
sales
date
day
month
location
branch
store
{ code = food }
Palette
Class
Package
Enumeration
DataType
PrimitiveType
Constraint
Association Class
Interface

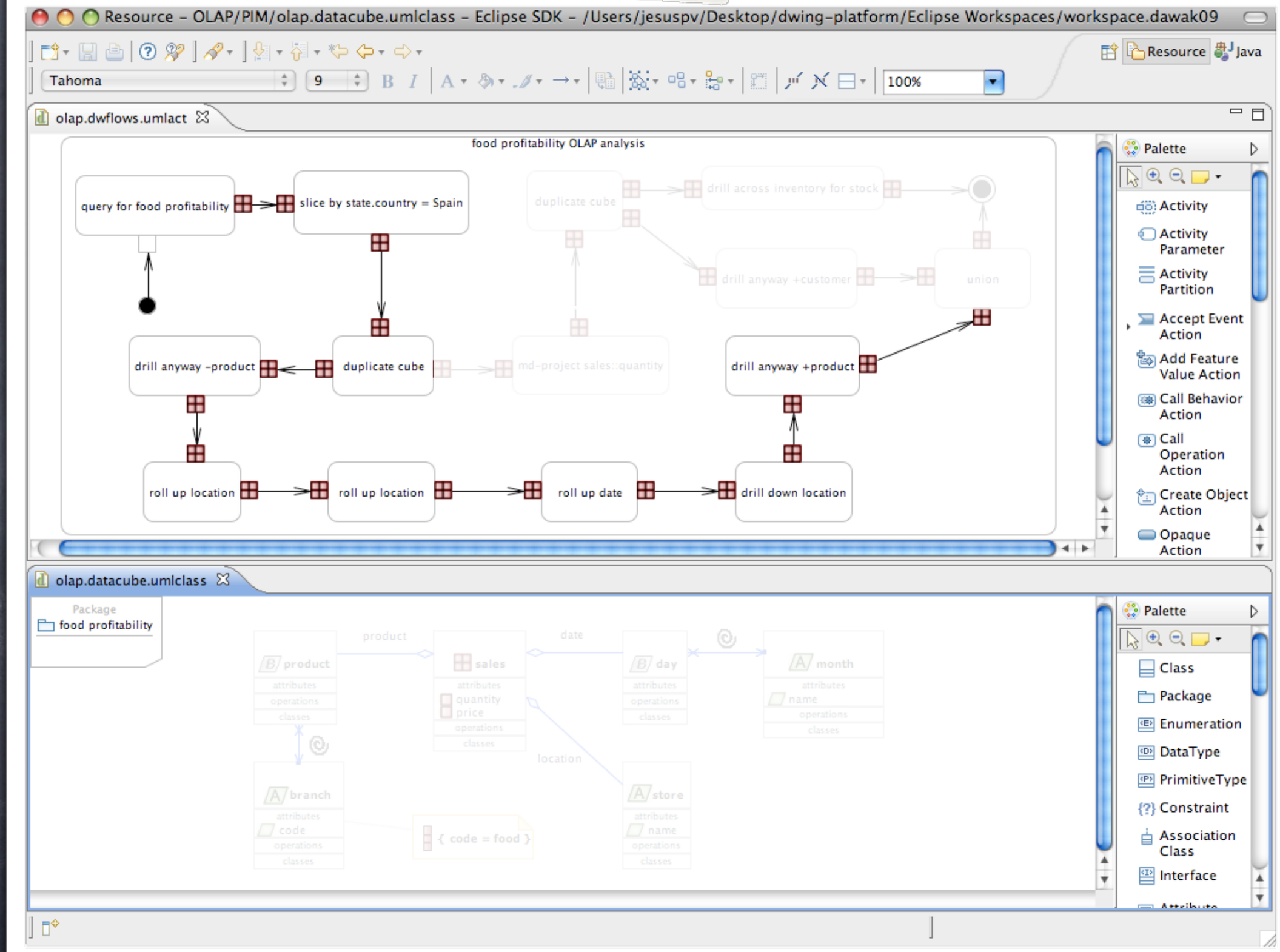
Prototype


Resource - OLAP/PIM/olap.datacube.umlclass - Eclipse SDK - /Users/jesuspv/Desktop/dwing-platform/Eclipse Workspaces/workspace.dawak09
Resource
Java
Tahoma
9
100%
olap.dwflows.umlact
food profitability OLAP analysis
query for food profitability
slice by state.country = Spain
duplicate cube
drill across inventory for stock
drill anyway +customer
union
drill anyway -product
duplicate cube
md-project sales::quantity
drill anyway +product
roll up location
roll up location
roll up date
drill down location
Palette
Activity
Activity Parameter
Activity Partition
Accept Event Action
Add Feature Value Action
Call Behavior Action
Call Operation Action
Create Object Action
Opaque Action
olap.datacube.umlclass
Package
food profitability
product
attributes
operations
classes
sales
attributes
quantity
price
operations
classes
date
day
attributes
operations
classes
month
attributes
name
operations
classes
location
branch
attributes
code
operations
classes
{ code = food }
store
attributes
name
operations
classes
Palette
Class
Package
Enumeration
DataType
PrimitiveType
Constraint
Association Class
Interface

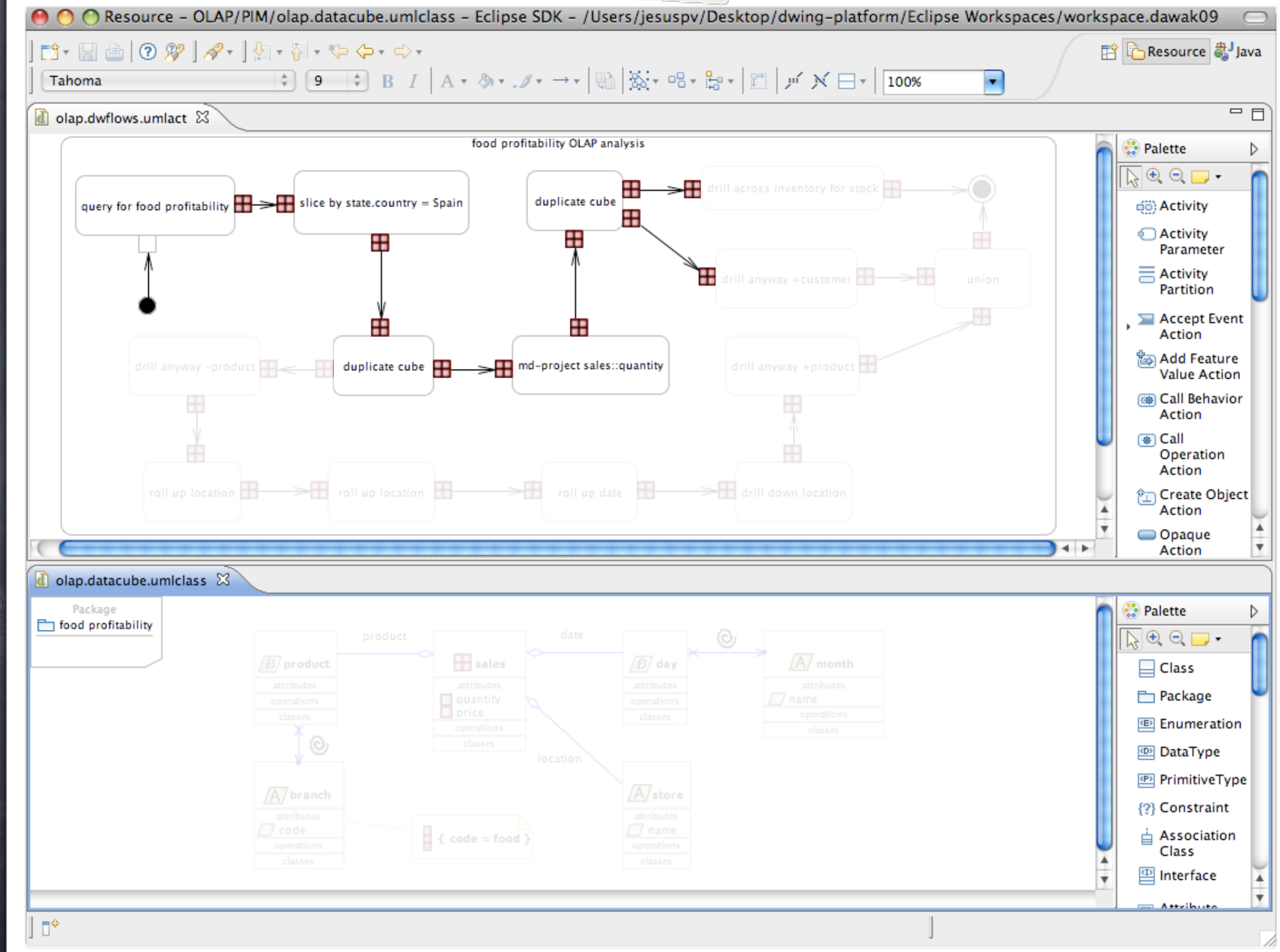
Prototype


Resource - OLAP/PIM/olap.datacube.umlclass - Eclipse SDK - /Users/jesuspv/Desktop/dwing-platform/Eclipse Workspaces/workspace.dawak09
Resource
Java
Tahoma
9
100%
olap.dwflows.umlact
food profitability OLAP analysis
query for food profitability
slice by state.country = Spain
duplicate cube
drill across inventory for stock
drill anyway +customer
union
drill anyway -product
md-project sales::quantity
drill anyway +product
roll up location
roll up location
roll up date
drill down location
Palette
Activity
Activity Parameter
Activity Partition
Accept Event Action
Add Feature Value Action
Call Behavior Action
Call Operation Action
Create Object Action
Opaque Action
olap.datacube.umlclass
Package
food profitability
product
sales
day
month
branch
store
{ code = food }
location
Palette
Class
Package
Enumeration
DataType
PrimitiveType
Constraint
Association Class
Interface

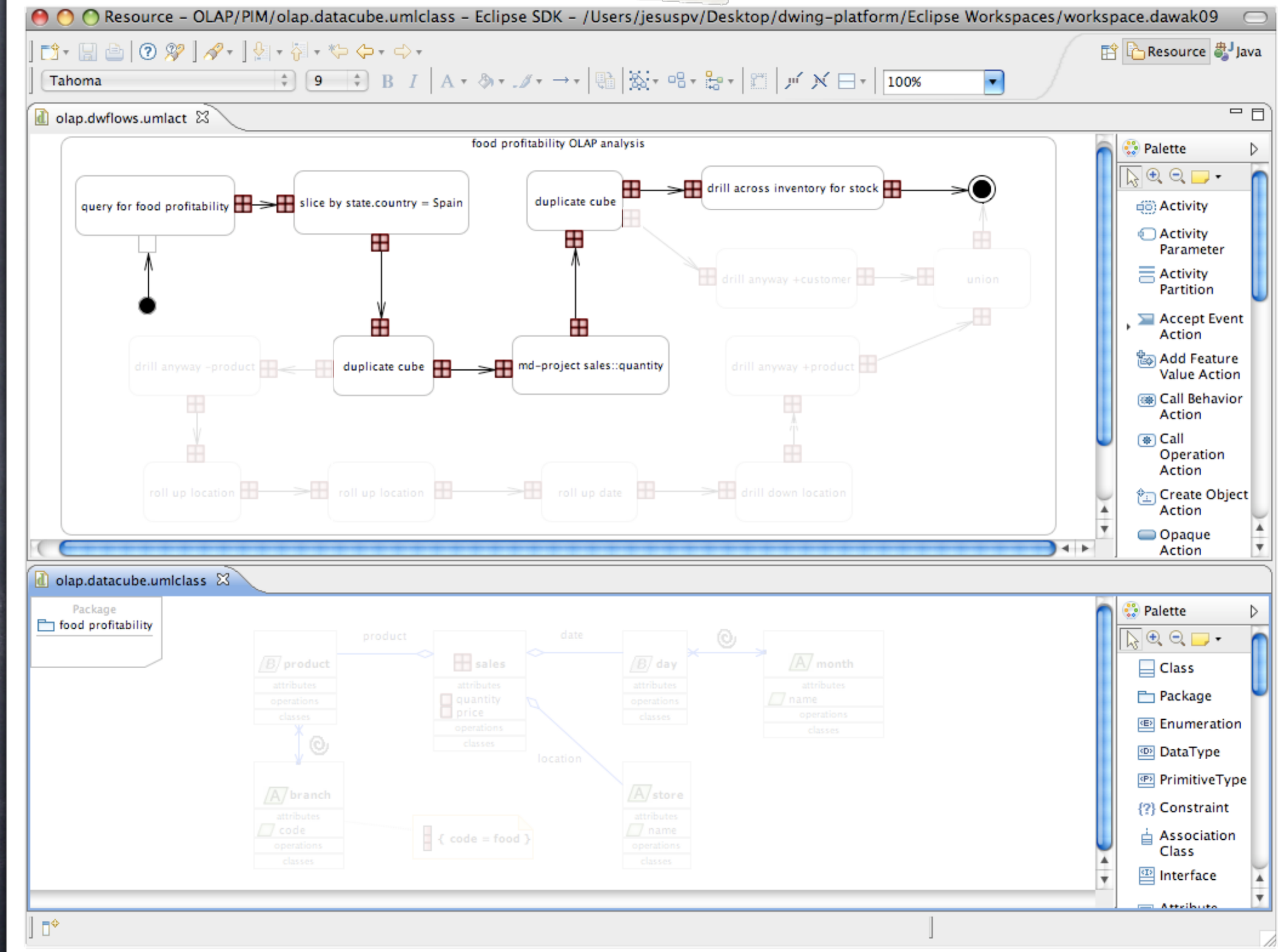
Prototype


Resource - OLAP/PIM/olap.datacube.umlclass - Eclipse SDK - /Users/jesuspv/Desktop/dwing-platform/Eclipse Workspaces/workspace.dawak09
Resource
Java
Tahoma
9
100%
olap.dwflows.umlact
food profitability OLAP analysis
query for food profitability
slice by state.country = Spain
duplicate cube
drill across inventory for stock
drill anyway +customer
union
drill anyway -product
duplicate cube
md-project sales::quantity
drill anyway +product
roll up location
roll up location
roll up date
drill down location
Palette
Activity
Activity Parameter
Activity Partition
Accept Event Action
Add Feature Value Action
Call Behavior Action
Call Operation Action
Create Object Action
Opaque Action
olap.datacube.umlclass
Package
food profitability
product
sales
day
month
branch
store
{ code = food }
location
Palette
Class
Package
Enumeration
DataType
PrimitiveType
Constraint
Association Class
Interface

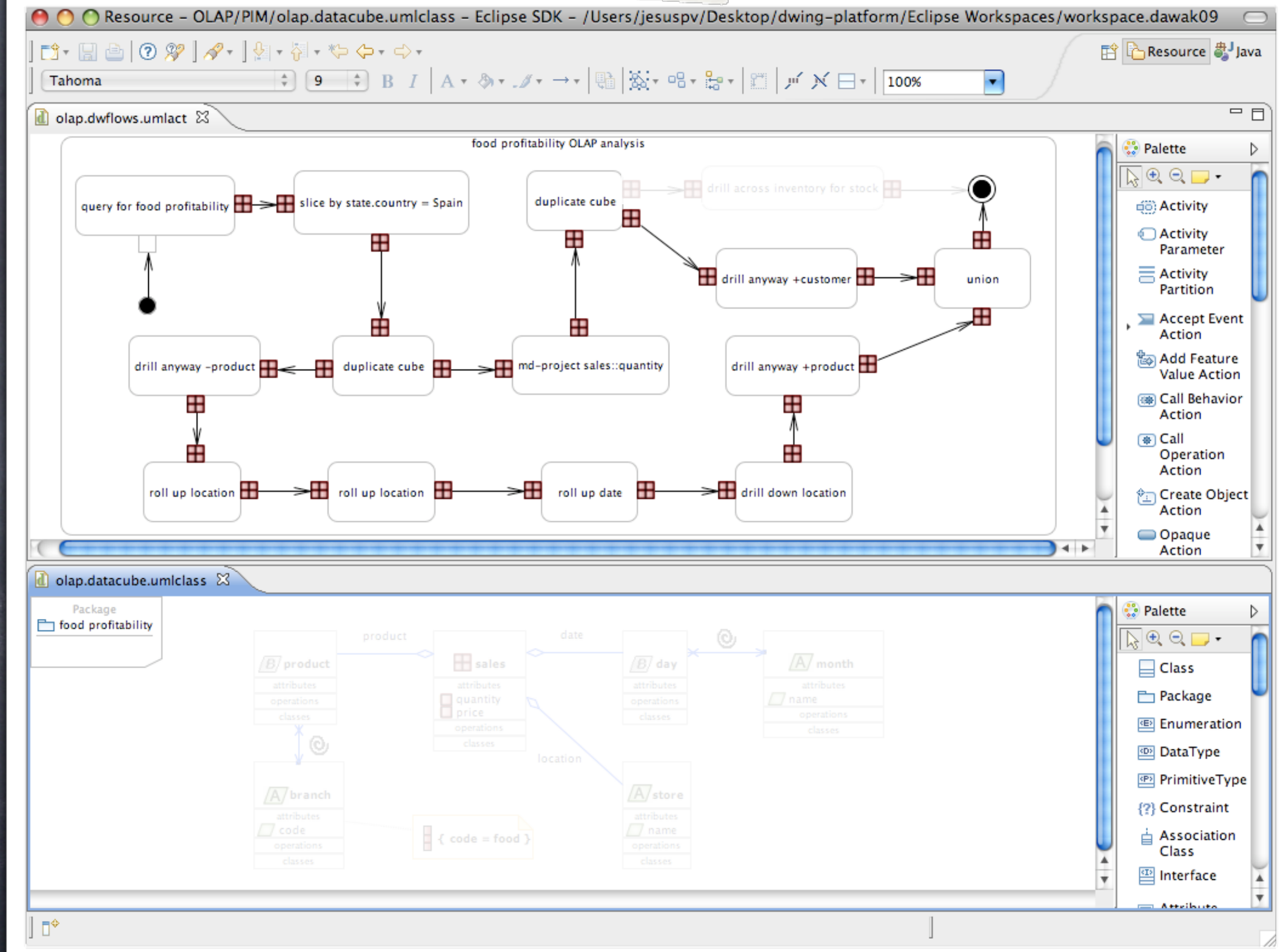
Prototype


Resource - OLAP/PIM/olap.datacube.umlclass - Eclipse SDK - /Users/jesuspv/Desktop/dwing-platform/Eclipse Workspaces/workspace.dawak09
Resource
Java
Tahoma
9
100%
olap.dwflows.umlact
food profitability OLAP analysis
query for food profitability
slice by state.country = Spain
duplicate cube
drill across inventory for stock
drill anyway +customer
union
drill anyway -product
duplicate cube
md-project sales::quantity
drill anyway +product
roll up location
roll up location
roll up date
drill down location
Palette
Activity
Activity Parameter
Activity Partition
Accept Event Action
Add Feature Value Action
Call Behavior Action
Call Operation Action
Create Object Action
Opaque Action
olap.datacube.umlclass
Package
food profitability
product
date
location
product
attributes
operations
classes
sales
attributes
quantity
price
operations
classes
day
attributes
operations
classes
month
attributes
name
operations
classes
branch
attributes
code
operations
classes
store
attributes
name
operations
classes
{ code = food }
Palette
Class
Package
Enumeration
DataType
PrimitiveType
Constraint
Association Class
Interface

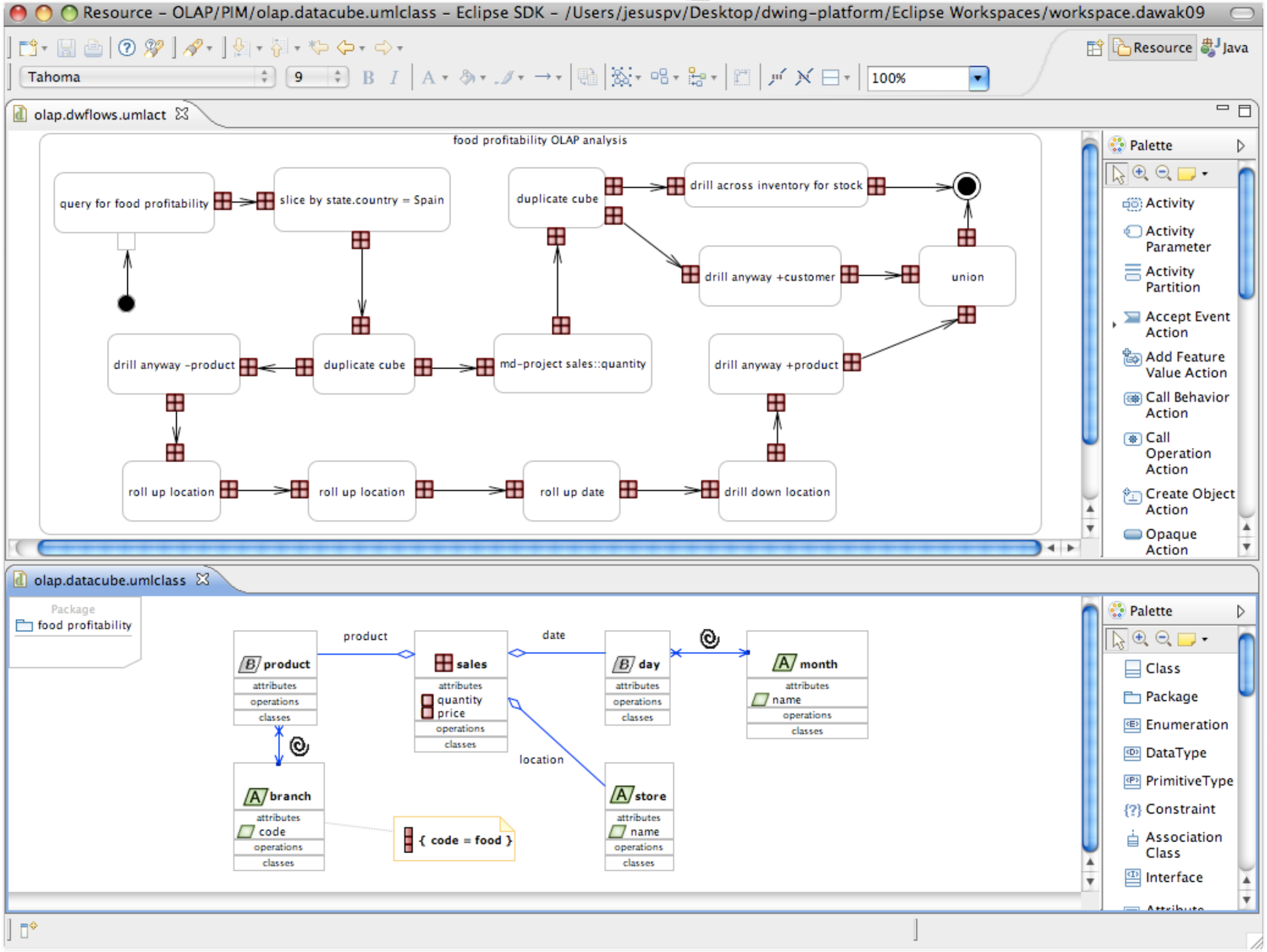
Prototype


Resource - OLAP/PIM/olap.datacube.umlclass - Eclipse SDK - /Users/jesuspv/Desktop/dwing-platform/Eclipse Workspaces/workspace.dawak09
Resource
Java
Tahoma
9
100%
olap.dwflows.umlact
food profitability OLAP analysis
query for food profitability
slice by state.country = Spain
duplicate cube
drill across inventory for stock
drill anyway +customer
union
drill anyway -product
duplicate cube
md-project sales::quantity
drill anyway +product
roll up location
roll up location
roll up date
drill down location
Palette
Activity
Activity Parameter
Activity Partition
Accept Event Action
Add Feature Value Action
Call Behavior Action
Call Operation Action
Create Object Action
Opaque Action
olap.datacube.umlclass
Package
food profitability
product
date
location
product
attributes
operations
classes
sales
attributes
quantity
price
operations
classes
day
attributes
operations
classes
month
attributes
name
operations
classes
branch
attributes
code
operations
classes
store
attributes
name
operations
classes
{ code = food }
Palette
Class
Package
Enumeration
DataType
PrimitiveType
Constraint
Association Class
Interface

Prototype


Resource – OLAP/PIM/olap.datacube.umlclass – Eclipse SDK – /Users/jesuspv/Desktop/dwing-platform/Eclipse Workspaces/workspace.dawak09
Resource
Java
Tahoma
9
100%
olap.dwflows.umlact
food profitability OLAP analysis
query for food profitability
slice by state.country = Spain
duplicate cube
drill across inventory for stock
drill anyway +customer
union
drill anyway -product
duplicate cube
md-project sales::quantity
drill anyway +product
roll up location
roll up location
roll up date
drill down location
Palette
Activity
Activity Parameter
Activity Partition
Accept Event Action
Add Feature Value Action
Call Behavior Action
Call Operation Action
Create Object Action
Opaque Action
olap.datacube.umlclass
Package
food profitability
product
sales
day
month
branch
store
{ code = food }
location
Palette
Class
Package
Enumeration
DataType
PrimitiveType
Constraint
Association Class
Interface

Prototype


Resource – OLAP/PIM/olap.datacube.umlclass – Eclipse SDK – /Users/jesuspv/Desktop/dwing-platform/Eclipse Workspaces/workspace.dawak09
Resource
Java
Tahoma
9
100%
olap.dwflows.umlact
food profitability OLAP analysis
query for food profitability
slice by state.country = Spain
duplicate cube
drill across inventory for stock
drill anyway +customer
union
drill anyway -product
duplicate cube
md-project sales::quantity
drill anyway +product
roll up location
roll up location
roll up date
drill down location
Palette
Activity
Activity Parameter
Activity Partition
Accept Event Action
Add Feature Value Action
Call Behavior Action
Call Operation Action
Create Object Action
Opaque Action
olap.datacube.umlclass
Package
food profitability
product
sales
date
day
month
location
branch
store
{ code = food }
Palette
Class
Package
Enumeration
DataType
PrimitiveType
Constraint
Association Class
Interface

Prototype

Conclusion

- Wide range of visualisation techniques for F^W
  - Heterogeneous diagrams
  - Isolated frameworks
- Challenges
  - Handle data cubes (complex types)
  - Unify modelling requirements
  - Drive perception & cognition
- Contribution
  - Characterisation of F^W
  - UML Diagrams for F^W + C
  - Prototype on the Eclipse platform

Open Questions

- Systematic Review of Literature
- Study of the Cognitive Aids
  - Theoretical foundation
  - Empirical validation
- Make Diagrams Executable

DaWaK '09

Visual Modelling of Data Warehousing Flows with UML Profiles

Jesús Pardillo[1], Matteo Golfarelli[2], Stefano Rizzi[2], Juan Trujillo[1]

Thank you very much for your attention!

[1] University of Alicante, Spain
[2] University of Bologna, Italy