Database Management Systems

Practice N. 4

The practice purpose is to generate cubes, reports, and dashboards from a data warehouse that addresses the issues described in the following by using Pentaho Community Edition client-server suite and its plug-ins.

1. Problem description

The management of a chain of Italian seaside resorts wants to analyze the number of rented items in their seaside resorts. Currently, there is a separate operational database for each seaside resort and the management is not able to compare the number of rented items and the net income of the seaside resorts. Each operational database only stores the information about the number of rented items (beach umbrellas, sun beds, etc.) for one seaside resort and the price paid by clients to rent the items. The conceptual model of the operational database of one seaside resort (Stab1 source data) is reported in Figure 1, while its logical schema is depicted in Figure 2.

The management of the company needs a data warehouse to analyze the net income and the number of rented items for every seaside resort and compare them. The management frequently performs the following analyses.

- Monthly net income for each province and region of the seaside resorts.
- Quarterly net income for each province and region of the seaside resorts.
- Four-monthly net income for each province and region of the seaside resorts.
- Yearly net income for each province and region of the seaside resorts.
- Monthly, quarterly, four-monthly and yearly net income for each category of rented items, detailed for region and province of the seaside resorts.
- Number of rented items per month, quarter, four-month period and year for each category of rented items.
- Number of monthly, quarterly, four-monthly and yearly rented items for each province and region of the seaside resorts.

The database Stab1, stored on an Oracle DBMS, is the operational databases of one seaside resort. To carry out the practice a data warehouse solution is already given. All the tables and the corresponding data have already been loaded into Pentaho Community Edition. All the material related to data warehouse design and feeding (useful only to students who want to carry the practice on their own PC) are available at the course website http://dbdmg.polito.it/wordpress/teaching/database-management-systems/.

During the practice the following tasks have to be accomplished:
- *Cube creation*
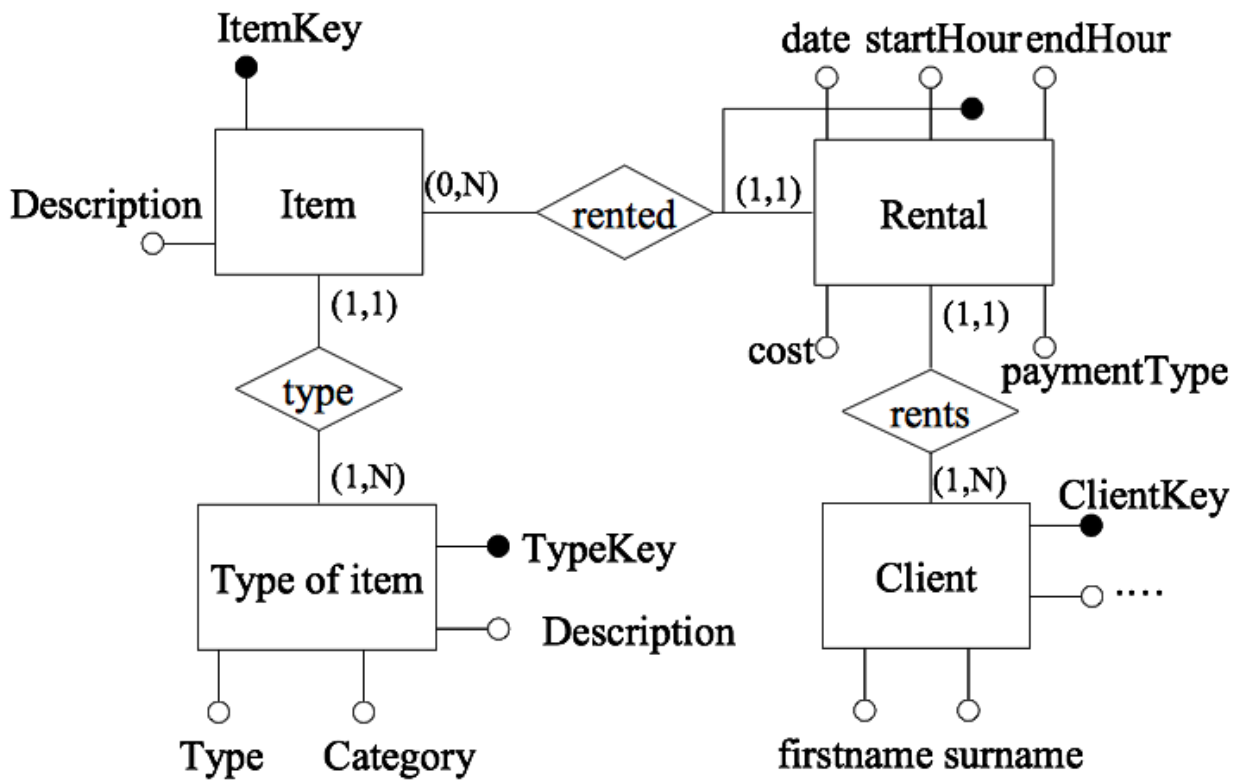- *Report creation*
- *Dashboard creation*

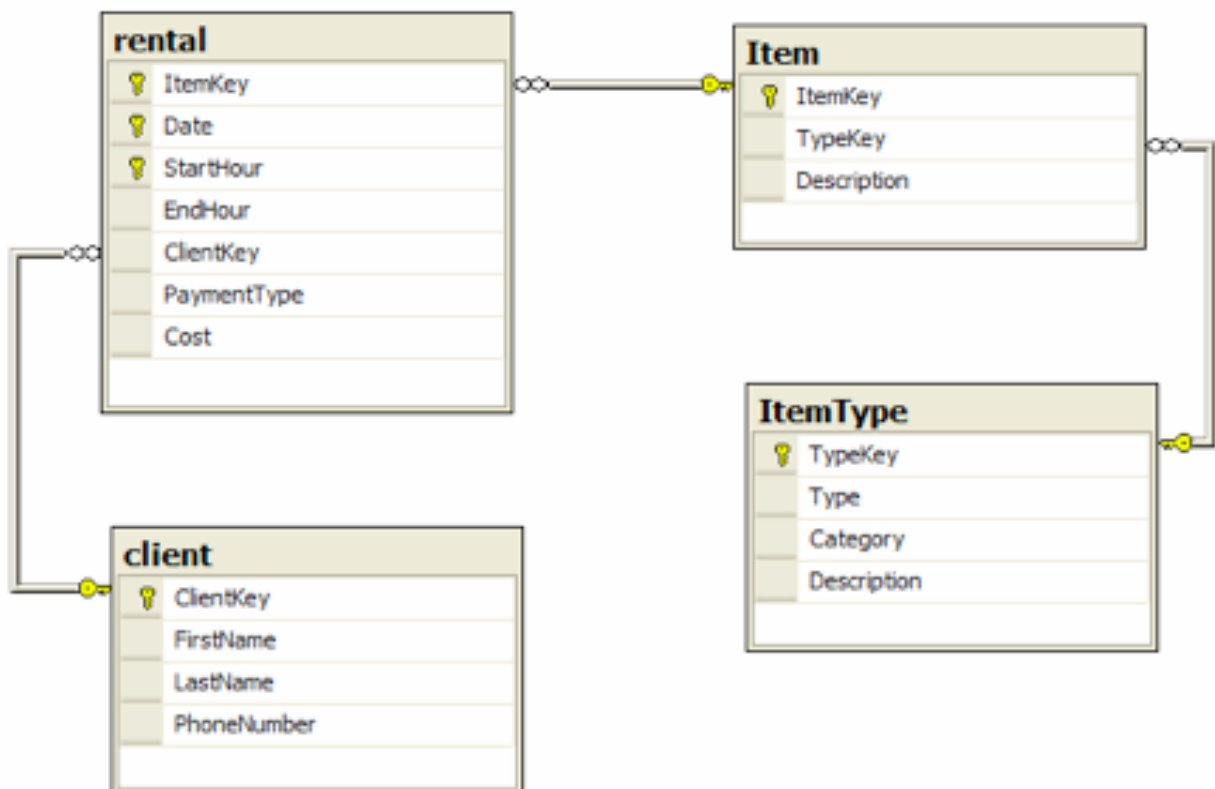**Figure 1. Conceptual model Stab1 data source**

Figure 2 – Logical schema of Stab1 data source

2. Cubes

Use the data warehouse already created and fed and generate cubes by considering the following frequently executed queries:

- Analyze the net income for each pair (province, quarter), the total net income for each province and the total net income for each quarter.
- Analyze the net income for each pair (item category, month), the net income for each item category (independently of the month) and the net income for each month (for all the item categories).

During the practice the following activities must be performed:
- *Builds the cubes by using Pentaho schema workbench*
- *Visualize and interact with the previously created cubes by using Pentaho BI User Console and Saiku Analytics plug-in*

2.1   Instructions

To *create and publish cubes* in Pentaho Community Edition use client *Schema WorkBench.*

Steps to create cubes in Schema WorkBench:

- Run the Java application SchemaWorkBench
    - C:\Pentaho\clients\schema-workbench\ workbench.bat

- Establish a connection with the data warehouse
    - Select *Options -> Connections* from the drop-down menu
    - Insert the following data into panel *General*
        - Connection name: *SeasideResort*
        - Connection type: *Oracle*
        - Access: *Native (JDBC)*
        - Hostname: *130.192.27.12*
        - Database Name: *xe*
        - Tablespace for Data: *xe*
        - Port number: *1521*
        - Username: *seasideDW*
        - Password: *password*

- Create schemata
    - Select *File -> New -> Schema* from the drop-down menu
    - Schema name: SeasideSchema*[Student ID number] (*e.g., *SeasideSchema190000* if the Student ID is 190000*)*

- Create cubes
    - Left-hand side panel -> right-click on icon *Schema -> Add cube*
    - Cube name: *SeasideCube*

- Create links to the fact table
    - Left-hand side panel -> right-click on icon *SeasideCube -> Add Table*
    - Schema: select *SeasideDW* from the drop-down menu
    - Table name: *RENTAL*

- Create dimensions
    - Left-hand side panel -> right-click on icon *SeasideCube -> Add Dimension*

- o Example: temporal dimension
  - ▪ Name: *TimeDim*
  - ▪ Foreign key *(dalla fact table)*: *COD_T*
  - ▪ Type: *TimeDimension*

- ● Create hierarchies on dimensions
  - o Left-hand side panel -> right-click on icon *Dimension* previously inserted (e.g., *TimeDim*) -> *Add Hierarchy*
  - o Example: hierarchy on the temporal dimension
    - ▪ Name: *TimeDimHier*
    - ▪ Primary key *(*of the dimension table*)*: *COD_T*

- ● Creare the dimension tables
  - o Left-hand side panel -> right-click on icon *Hierarchy* previously inserted (e.g., *TimeDimHier*) -> *Add Table*
  - o Example: table for a temporal dimension
    - ▪ Schema: *SeasideDW*
    - ▪ Name: *TIMEDIM*

- ● Create hierarchy levels
  - o Left-hand side panel -> right-click on icon *Hierarchy* previously inserted (e.g., *TimeDimHier*) -> *Add Level*
  - o Example: level of a temporal dimension
    - ▪ Name: *YEAR*
    - ▪ Column: *YEAR*
    - ▪ nameColumn: *YEAR*
    - ▪ Type: *String*
    - ▪ levelType*: TimeYears*
    - ▪ HideMemberIf*: Never*
  - o N.B. Levels have to be create in order of decreasing granularity (e.g., *YEAR* first, *TRIMESTER* secondly)

- ● Create measures
  - o Left-hand side panel -> right-click on icon SeasideCube -> *Add Measure*
  - o Example of measure
    - ▪ Name: *NetIncomeMeasure*
    - ▪ Aggregator: *sum*
    - ▪ Column: *NETINCOME*
    - ▪ Datatype: *Integer*

NOTE: Create multiple instances of the temporal dimension with AT MOST one hierarchy per dimension rather than a single dimension with multiple hierarchies.

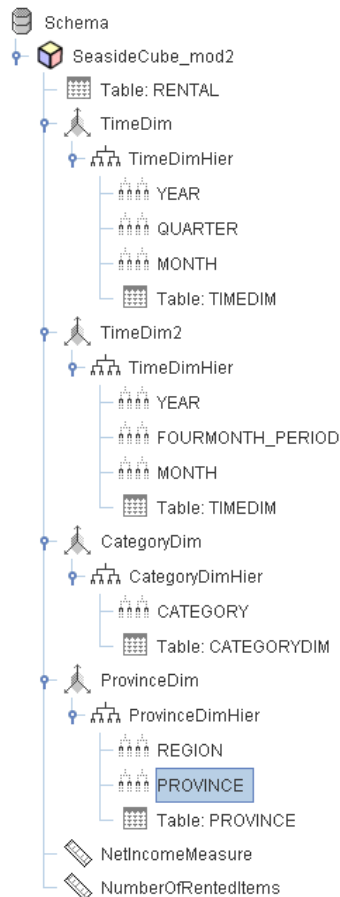The final cube structure should look like those depicted in Figure 1.

Figure 1. Final cube structure

To publish cubes select *File -> Publish* from the drop-down menu
- URL: *http://130.192.27.12:8090/pentaho/*
- Publish password: *publish*
- Pentaho credential
    - Username: *bdati[LabInf PC ID number] (e.g., bdati97 if the PC ID is 97)*
    - Password: *password*
- Location: *bdati[LabInf PC ID number]*
- Pentaho or JNDI Data Source: *SeasideResort*

To visualize cubes use the *Pentaho BI User Console*
- To run the Pentaho BI User Console connect through a Web browser (Mozilla Firefox or Google Chrome) to the following address: *http://130.192.27.12:8090*
- Login
    - Username: *bdati[LabInf PC ID number]*
    - Password: *password*

- New Analysis
    - Schema: SeasideSchema*[Student ID number]*
    - Cube: *SeasideCube*

- Use Saiku Analytics (*File -> New -> Saiku Analytics*) to browse cubes

3. Reports

Use the data warehouse already created and fed to:

- Create a report that contains for each **year** a table showing the **monthly total income** for each **item category**
- Publish the reports on Pentaho BI User Console
- Visualize the reports and export them in PDF format

During the practice the following activities have to be performed:
- *Create a report with Pentaho report designer*
- *Visualize and export the report with Pentaho BI User Console*

3.1 Instructions

To create and publish reports in Pentaho Community Edition use client *Report Designer*

To create reports with the Report Designer Wizard:
- Run Pentaho Report Designer
  - *C:\Pentaho\clients\report-designer \report-designer.bat*

Steps to complete wizard creation:

STEP "Look and feel"
- Select an existing template
- STEP "Data source" : Create a new Data Source for the report
  - Click on icon "+" at the top-right hand corner of panel *Data Source*
  - Type: *JDBC*

STEP "Data source"
- Establish a new connection
  - Connection name: *SeasideResort*
  - Connection type: *Oracle*
  - Access: *Native (JDBC)*
  - Hostname: *130.192.27.12*
  - Database Name: *xe*
  - Tablespace for Data: *xe*
  - Port number: *1521*
  - Username: *seasideDW*
  - Password: *password*

- Create a new query using panel *Available queries*
  - Query example
    - Query name: *Query1*
    - SQL code: *SELECT * FROM RENTAL*
    - Note: omit semicolon (;) at the end of SQL instructions
  - Click on *preview* to preview the query output

STEP "Layout"
- Define the output layout
  - How items should be grouped (*Group Items by*) (if any)
  - Which attributes visualize for each group (*Selected Items*)

- E.g., *NETINCOME*
  - o Click on *preview* to preview the query result

STEP "Format"
- ● Define the format and the aggregation type
  - o Example of details about *NETINCOME*
    - ▪ Display name: *NETINCOME*
    - ▪ Date format: *NONE*
    - ▪ Aggregation: *SUM*
  - o Click on *preview* to preview the query result

At the end of the guided creation procedure the report can be further personalized through the graphical interface. An example of result is reported in Figure 2.

To publish reports select *File -> Publish* from the drop-down menu
- URL: *http://130.192.27.12:8090/pentaho/*
- Timeout: *30*
- Pentaho credential
  - o Username: *bdati[LabInf PC ID] (*e.g. *bdati97* if the PC ID is 97*)*
  - o Password: *password*
- Location: *bdati[LabInf PC ID]*
- Publish password: *publish*
- Pentaho or JNDI Data Source: *SeasideResort*
- Output type: *PDF*

Now the report published can be accessed through the *Pentaho BI User Console.* Alternatively, click on *File -> Export* from the drop-down menu to export the report in different formats (e.g., PDF, HTML).

Report Title
Sub Title 1
Sub Title 2

NETINCOME
$70,500.00
$39,750.00
$9,150.00
$63,600.00
$70,500.00
$39,750.00
$9,150.00
$63,600.00
$89,300.00
$50,350.00
$11,590.00
$80,560.00
$70,500.00
$39,750.00
$9,150.00
$63,600.00
$70,500.00
$39,750.00
$9,150.00
$63,600.00
$89,300.00
$50,350.00
$11,590.00
$80,560.00
$131,600.00
$74,200.00
$17,080.00
$118,720.00
$70,500.00
$1,607,700.00

Wed Oct 30 12:04:10 CET 2013

Figure 2. Report Preview example

4. Dashboards

Use the data warehouse already created and fed to:

- Create a dashboard that allows us to select an **year** of activity of the seaside resort and produces a bar chart showing the **number of rented items** during the specified year for each **item category**

During the practice the following activities have to be performed:
- *Create the dashboard with Community Dashboard Editor (CDE) plug-in*
- *Visualize the dashboard with Pentaho BI User Console*

4.1 Instructions

To create dashboards use the *Community Dashboard Editor (CDE)* plug-in embedded into the Pentaho BI User Console.

Steps to create dashboards with CDE:

- Open Pentaho BI User Console through a browser Web (Mozilla Firefox o Google Chrome) connecting to address 130.192.27.12:8090
- Login
    - Username: *bdati[LabInf PC ID]*
    - Password: *password*

- Launch the Community Dashboard Editor by clicking on icon "CDE" at the top-right hand corner (Alternatively, select *File -> New -> CDE DashBoard* from the drop-down menu)
- Select an existing dashboard layout (e.g., *Two Columns Template*) by clicking on icon *Apply Template* of panel *Layout Structure*
- Personalize the layout by modifying the corresponding attribute values (Note: rows and columns are organized in a hierarchical fashion)
- Create a drop down menu through which a parameter could be set
- Select Data Source *Wizards -> OLAP Selector Wizard*
    - drag and drop levels as in Figure 3
    - give a name to the selector
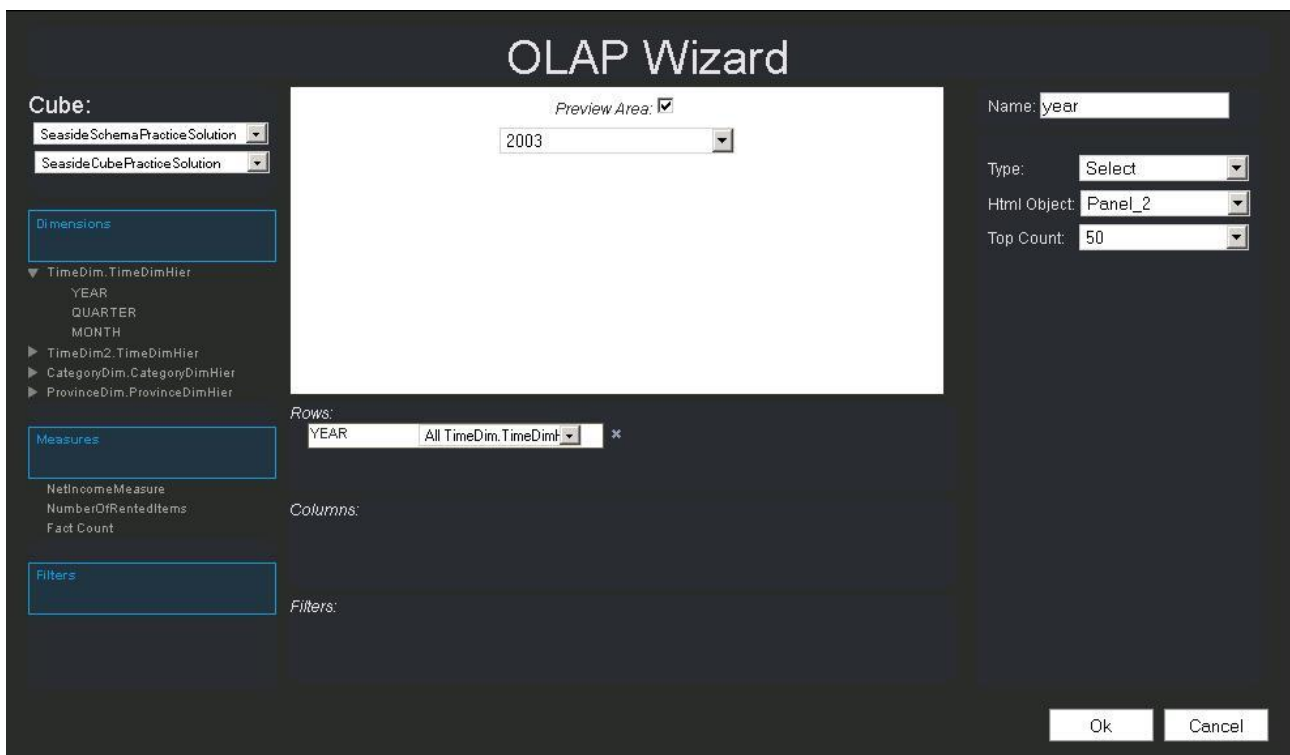   - Select the element of the layout through which the selector will be depicted (e.g., *Html object Panel_2*)

Figure 3. Year selector definition in CDE Dashboard

- Create the data source for the chart
- Select Data Source *Wizards -> OLAP Chart Wizard*
    - drag and drop levels, measures, and filters as in Figure 4
    - give a name and a title to the OLAP chart
    - select a chart type (e.g., *Bar Chart*)

o Select the element of the layout through which the chart will be depicted (e.g., *Html object Panel_1*)

Figure 4. Data Source definition in CDE Dashboard

● Check that Data Source parameter *JINDI* is set to *SeasideResort*

An example of result after preview is depicted in Figure 5.

Figure 5. Example of Dashboard with dynamic Bar Chart and drop-down menu