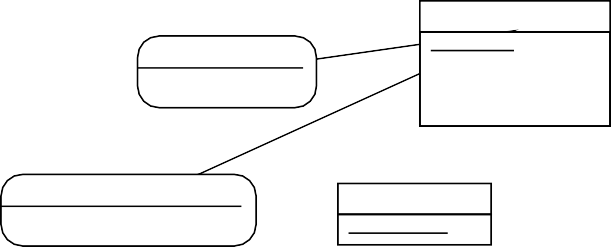
**DATABASE PROJECT INSTRUCTIONS**

French instructions are further. In case of incoherence or understanding difficulty regarding the present instructions, **English instructions prevail**.

The construction group BatiTP is a company that constructs and renovates buildings and has created a data warehouse. The conceptual schema of the data warehouse (in entity/relationship formalism) is:



COMPANY

**VATNum** NameC TypeC GroupC

(1,1)

(0,n)

EMPLOYS

(1,1)

HAS

(0,n)

EMPLOYEE

**CodeEE** LastNameEE FirstNameEE

(1,1)

QUALIFICATION

(0,n) **CodeQ**

NameQ

LevelQ

(1,n)

AUTHORIZE

OFFERS

LOCATED1

(0,n) (0,n)

INCLUDES

NbHoursPlanned

(1,n)

RANGEOFWORKS

**CodeRW**

NameRW

(1,n)

(0,n)

CITY

**CodeCity** NameCity DepartmentCity RegionCity

(0,n)

(1,1)

QUOTATION (1,n)

**CodeQ**

BELONGS

DateQ DateEndQ

(1,1)

(0,n)

CALENDAR1

**DateWQ**

(0,n)

(0,n)

(1,1)

TYPEOFWORKS

CONCERNS

LOCATED2

WORKQUOTE

NbHoursWQ

(0,n) **CodeTW**

NameTW HoulryRateCust

(0,n) HourlyRateEmp

(1,1)

(1,n)

SITE

**CodeS** AddrRoadS PostalCodeS

(0,n)

WORKOUTOFQUOTE

NbHoursWOQ

CALENDAR2

(0,n) **DateWOQ**

This data warehouse allows monitoring the activity of the different companies that work on construction sites from the initial quotation all the way to doing the works on the construction sites.

A construction work estimate is expressed using a quotation that is made by a company and concerns a construction site located in a city. Each quotation specifies a provisional number of hours for each type of construction work (such as: masonry, electricity, etc.). Each type of construction work is described by a name an hourly rate charged to the customer (to calculate the total price of the quotation) and another hourly rate used to pay the employee (to calculate the salary of the employee). Each type of construction work belongs to only one range of construction works (such as: structural work, finishing work, etc.).

Construction works are done either according to a quotation or out of a quotation. If it is according to a quotation, the daily number of hours done by an employee for each type of construction work is recorded, for each quotation. If the construction works falls out of a quotation, the daily number of hours done by an employee for each type of construction work is recorded for each construction site.

All employees are employed by only one company, and have only one qualification. Each qualification has a code, a name and belongs to only one level of qualification. This level corresponds to one of the following values: BEP, CAP, Bac Pro, BTS or Engineer (these are French diplomas). Moreover, all qualifications are associated to at least one range of construction works, allowing the employee to execute that range of construction works.

All companies are identified by a code (the company tax number), a name and a geographic location using the city, department number (French administrative division) and region. All companies are only of one type (for example: Plc—SA in French, Ltd—SARL in French, etc.) and belong to only one group that correspond to groups of companies.

The Entity/Relationship schema corresponds to the following relational schema where primary keys are underlined and foreign keys are followed by a #.

* **Company** (VATNum, NameC, TypeC, GroupC, CodeCity#)
* **Employee** (CodeEE, LastNameEE, FirstNameEE, CodeQ#, VATNum#)
* **City** (CideCity, NameCity, DepartmentCode, RegionCity)



* **Site** (CodeS, AddrRoadS, PostalCodeS, CodeCity#)
* **Qualification** (CodeQ, NameQ, LevelQ)
* **Authorize** (CodeQ#, CodeRW#)
* **RangeOfWorks** (CodeRW, NameRW)
* **TypeOfWorks** (CodeTW, NameTW, HourlyRateCust, HourlyRateEmp, CodeRW#)
* **Include** (CodeQ#, CodeTW#, NBHoursPlanned)
* **Quotation** (CodeQ, DateQ, DateEndQ, VATNum#, CodeS#)
* **WorkQuote** (CodeQ#, CodeEE#, DateWQ, CodeTW#, NBHourWQ)
* **WorkOutOfQuote** (CodeS#, CodeEE#, DateWOQ, CodeTW#, NBHourWOQ)

The attributes (or columns or fields) of this schema are defined in the data dictionary at the end of this document.

# Work to be done

The project is divided into two parts which have to be handed in at two different dates. The first part consists in **Implementation and manipulation of the data warehouse (DW)**; and the second part consists in **Implementation of a multidimensional data mart based on Business Objects**.

1. **Part 1: Implementation and manipulation of the data warehouse (DW)**
   1. Implement this data warehouse as a database using Microsoft Access (the data dictionary detailing the different columns is provided hereafter);
   2. Enter data within the warehouse (i.e. fill in the tables with data) using the same technique as suggested during the course (enough data to test your queries—see below);
   3. A first group of users is in charge of studying the benefits and margins of the different construction works done on the construction sites. Their requirements as users correspond to implementing the following analyses using SQL queries:
      1. Display the construction work ranges (code and name) with the average margin made on the construction works (whether it corresponded to a quotation or not);
      2. Give the code and date of quotations that contain only structural works;
      3. List the names, codes and address of the construction sites that should be the most profitable. To be profitable for this company is to have a ratio of (customer hourly rate – employee hourly rate) divided by the employee hourly rate greater than 80%);
   4. Write also a new SQL query of similar complexity that would complete the current decisional analyses of the users of this data warehouse. Justify your answer by stating the advantage(s) of this new analysis compared to the user’s requirements (the 3 queries described above).

# 

# Elements that have to be handed in for part 1

Elements to hand in are: a **dossier** (a pdf file that contains a **design report** and a **user guide**) that has the layout described hereafter and **Files** that contain the different implementations (Access database).

Structure of the **Design report** (sections 1 to 3 and 5 to 15 pages at most): this report will detail the implemented solution. The report describes the solution and recalls the specifications (the objectives specified in the present instructions). The following layout is mandatory for writing the report:

1. Present the **architecture** of the decision support system using text and a schema
   1. Description: For the decision maker (users), specify the requirements (at least on paragraph). Recall that those user-requirements are stated within the project description. Moreover, in addition to this informal part, it is necessary to provide a synthetic view of the user requirements using a table with the following columns (the column names are given in italic):
      * *Name* of the data mart: a short name to be used in the global schema (see below);
      * *Description* of the data mart: a long title describing the data mart such as “this data mart allows analysing…”
      * *Desired tool*: mention if it is a spread sheet tool, a database (DBMS), an OLAP query tool (such as Business Object), etc.
      * *Decision-maker* that are concerned by that data mart.
   2. Schema: a schema of the decision support system will have to be provided using explicit logos for the different components of the system as well as using the names described in the table in [1.1.](#_bookmark0) This schema must be associated to a comment (composed of one or several small paragraphs) and describes the architecture (software and interconnections) of the decision support system. You can base your schema on the architecture schemas that were provided during the first course.
2. Details of the data mart: you will need to provide **an analysis of the user requirements** as well as the

**detailed design** (as done in course and during exercises).

* For a data mart that uses a database (with SQL queries):
  + State the precise and detailed requirements of the decision makers for specifying their analyses;
  + Data: provide one conceptual schema that answers to these requirements, i.e. all queries (entity relationship and relational schemas) as well as the data dictionary. Recall that the whole data warehouse schema may not be necessary. All this information is provided within the present instructions (note that to draw the schemas you can use PowerPoint);
  + Processes: SQL queries associated to the decisional analyses with possible comments as well as the analysis requirements (the sentence that describes each query).

1. For the data warehouse, it is necessary to provide the global and detailed design. Recall that, within this project, the description of the data warehouse is described in the beginning of the present document.

* Data: the conceptual schema of the whole data warehouse (in entity relationship and relational format) along with a data dictionary. All this information is provided in this document;

1. Structure of the **User guide** (section 4 and 1 to 3 pages at most):
   1. Provide a brief description of all the files handed in, using the same names and notations as in the schema in [1.2.](#_bookmark1)
   2. Give the specific name of the file to open and the different actions that have to be taken in order to run each of the different decisional analyses in your project (only if these differ from what was done during practical work sessions).
   3. The user knows how to manipulate the software (Access or BO) and how to connect them. However, you will have to guide him to which query correspond to what analysis, and in each report, which graphic and/or table corresponds to what analysis.

The dossier should contain as many screen captures as necessary (screen captures may be done using ALT+PrintScreen key combination and then using MS Paint to edit the image—for minimising sizes, JPG or

PNG file formats are recommended). Note that text and arrows may be added to an image using PowerPoint or Impress (and pasting the result into Word or Writer).

# Constraints

The project has the following constraints:

* Project has to be done in pairs;
* The design report and user guide have to be in one file in pdf format (non-editable format—no .doc,

.odt or .docx files);

* All documents (the design report and the user guide) and files, Access databases (.mdb or .accdb files), BO universe and reports have to be handed in before **a date that will be given later on the university Moodle server before 5 PM**. A drop zone will be opened on the server for this purpose (warning, do not wait for the last moment as the server may go down for maintenance after 5 PM).
* A forum will be open on Moodle for questions so that all may see the discussions. **No individual questions will be answered by mail**.

# Notes concerning the grades

The **clarity** of the explanations of the report and more specifically of the user guide, **observing** the present

**instructions**, will be taken into account when correcting the projects.

# Technical notes

Date manipulation functions: some of these functions are available in Microsoft environments. Under Access, they are similar to those of Excel. Among them, one may find:

* *Month(table\_name.date\_filed\_name)* or *Month(a\_date)* gives the month of a date;
* *Date()* gives the date of today;
* *Year(table\_name.date\_field\_name)* or *Year(a\_date)* gives the year of a date.

In a French version of Microsoft Office, these functions are *Mois*(), *Maintenant*() and *Année*(). Moreover, take extra care when manipulating dates with Microsoft applications as some might use a French (DD/MM/YYYY) representation while others might use an English one (MM/DD/YYYY).

# Data dictionary of the data warehouse

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Description** | **Format** | **Constraints** |
| VATNum | Unique code of a company (Siret) | 14 digits (displayed as 3 groups of 3 digits followed by a group of 4 digits--groups are  separated by space) |  |
| NameC | Name of a company | 40 charatcter string |  |
| TypeC | Type of a company | 5 character string | Enumeration (SA, SARL, SCOP,  EURL) |
| GroupC | Company Group Name (owner) | 40 charatcter string |  |
| CodeCity | Unique code of a city | 5 character string | always 5 digits |
| NameCity | Name of a city | 40 charatcter string |  |
| DepartmentCity | Department code of a city | 3 character string | 2 digits that may be followed by  one character |
| RegionCity | Region name of a city | 40 charatcter string |  |
| CodeEE | Unique code of an employee | Number (Integer) | automatically generated |
| LastNameEE | Last name of an employee | 40 charatcter string |  |
| FirstNameEE | First name of an emplyee | 40 charatcter string |  |
| CodeQ | Unique code\* of a qualification | Number (Integer) | automatically generated |
| NameQ | Name of a qualification | 40 charatcter string |  |
| LevelQ | Level of a qualification | 9 character string | Enumeration (BEP, CAP, BAC  PRO, BTS, INGENIEUR) |
| CodeRW | Unique code of a range of construction works | Number (Integer) | automatically generated |
| NameRW | Name of a range of construction works | 40 charatcter string | Ennumeration (GROS OEUVRE, SECOND OEUVRE, AMENAGEMENT, FINITION /  EMBELLISSEMENT, SOL) |
| CodeS | Unique code of a construction site | Number (Integer) | automatically generated |
| AddrRoadS | Address (number and road name)of a  construction site | 40 charatcter string |  |
| PostalCodeS | Postal code of a construction site | 5 character string | always 5 digits |
| CodeTW | Unique code of a type of construction work | Number (Integer) | automatically generated |
| NameTW | Name of a type of construction work | 40 charatcter string | Ennumeration (TERRASSEMENT & VRD, MACONNERIE, ELEMENTS DE STRUCTURE METALLIQUES, METALLERIE, BOIS, COMPOSITE, CLOISENNEMENT, FAUX PLAFOND, REVETEMENT DE SOL, VENTILATION/CLIMATISATION, ELECTRICITE, CHAUDRONNERIE, PLOMBERIE, MENUISERIE, MOBILIER, PLAFONNAGE, PARQUET, PEINTURE, MOQUETTE, SOL SYNTHETIQUE) |
| HourlyRateCust | Hourly rate charged to the customer | Number (currency format: two  digits after decimal point, unit: Euro) | >= 0 AND >=  HourlyRateEmployee |
| HourlyRateEmp | Hourly rate of the employee working on the type of construction work | Number (currency format: two  digits after decimal point, unit: Euro) | >= 0 |
| CodeQ | Unique code\* of a quotation for construction  works | Number (Integer) | automatically generated |
| DateQ | Date when the quotation is created | Date (format: DD/MM/YYYY) | <= current date |
| DateEndQ | Date of end of validity of the quotation | Date (format: DD/MM/YYYY) | > DateQ |
| NbHoursPlanned | Number of hours planned for a type of work in a quotation | Number (real) | >= 0 |
| DateWQ | Date when a type of construction work in a quotation is done | Date (format: DD/MM/YYYY) | <= current date |
| NBHoursWQ | Number of hours spent on a type of  construction work in quotation | Number (real) | >= 0 |
| DateWOQ | Date when a type of construction work out of a  quotation is done | Date (format: DD/MM/YYYY) | <= current date |
| NBHoursWOQ | Number of hours spent on a type of  construction work out of a quotation | Number (real) | >= 0 |
| \*Warning: there are two CodeQ attributes that correspond to different different attributes. | | | |