

INFO-H-403 / INFO-F-301 : Bases de données

Projet : Outil d'analyse financière

Professeur : Esteban Zimányi

Assistants : Frédéric Servais et Boris Verhaegen

<http://cs.ulb.ac.be/public/teaching/infoh403>

Année académique 2007-2008

Étude de cas

Félicitations! Vous avez été engagés par *Goohoo!*, un portail Internet prometteur, pour développer un module d'analyse financière, *Goohoo! Finance*. Un analyste a compilé ci-dessous le résultat des diverses interviews.

Le patron de *Goohoo!* nous livre les informations suivantes :

Il s'agit d'une application graphique permettant d'analyser des actions et des indices. Pour chaque produit financier, on veut pouvoir consulter des informations générales comme son nom, son code et le marché auquel il appartient. Le code détermine univoquement le produit financier et se compose de maximum quatre caractères. L'application devra également permettre de consulter l'historique des cours d'un produit financier : le volume, le cours à l'ouverture et à la fermeture ainsi que le cours minimal et maximal pour chaque séance. On devra aussi être capable de consulter les actualités liées aux produits financiers. De plus, l'application devra permettre de créer, modifier ou supprimer un produit financier.

Un consultant spécialisé dans la finance nous donne plus de précisions sur les produits financiers :

Un produit financier est soit une action, soit un indice. Tous les deux appartiennent à un marché. Un marché est identifié par son nom, sa ville et son pays et travaille avec une seule devise. Un indice est un produit financier qui est un panier pondéré d'actions d'un même marché. Par exemple, l'indice "Big Three" contient 1000 actions dont 500 de Microsoft, 300 de Google et 200 d'Apple. Le cours d'un indice est la somme pondérée des cours des actions dont il est composé. Le volume du jour d'une action représente le nombre d'actions échangées ce jour.

Un informaticien de *Goohoo!* nous donne quelques informations d'ordre technique :

Les cours des actions sont reçues en fichiers CSV. Le nom du fichier est composé du code de l'action suivi de l'extension du fichier. Par exemple, pour l'action Apple, le fichier s'appellera "AAPL.csv". Le séparateur de champs est la virgule et le symbole décimal est le point. La date est au format "AAAA-MM-JJ". La première ligne du fichier contient uniquement des méta-données et chaque ligne est structurée de la manière suivante :

```
Date,Cours ouverture,Cours maximum,Cours minimum,Cours fermeture,Volume,Cours fermeture ajusté
```

Les actualités des produits financiers sont disponibles au format XML et suivent le modèle ci-dessous. Une actualité est composée d'un titre, des codes des produits financiers concernés, de la source de l'information, d'un texte descriptif, d'un lien et d'une date. Cette date est au format RFC 822, par exemple "Wed, 10 Oct 2007 13:05:02 GMT".

```
<news>
  <item>
    <title>Titre de l'actualité</title>
    <code>Code du produit financier 1</code>
    <code>Code du produit financier 2</code>
    ...
    <source>Nom de la source</source>
    <description>Texte</description>
    <link>Lien vers l'article original</link>
    <pubDate>Date au format RFC822</pubDate>
  </item>
  ...
</news>
```

Ces données devront être importées dans votre base de données.

Déroulement du projet

Première partie

Pour cette partie, on vous demande de modéliser le problème à l'aide du formalisme entité-association et de préciser les contraintes d'intégrité nécessaires. Ces contraintes doivent être exprimées en français et utiliser les mêmes noms d'entités, d'associations ou attributs que dans votre modèle conceptuel. Vous pouvez également exprimer et justifier des hypothèses sur votre modèle. Ces hypothèses peuvent résulter par exemple d'incomplétudes ou d'ambiguïtés dans l'énoncé du problème. Déduisez ensuite de ce modèle conceptuel le modèle relationnel correspondant ainsi que ses contraintes. N'hésitez pas à justifier vos choix de modélisation.

Deuxième partie

Création

On vous demande tout d'abord de déduire de votre modèle relationnel un script SQL DDL de création de la base de données et de ses différentes tables ainsi que de créer cette base de données.

Initialisation

Les informations concernant les produits financiers nécessaires pour initialiser votre base de données sont disponibles dans les annexes. On vous demande d'écrire à partir de ces tableaux un script SQL d'insertion de données dans vos tables. En ce qui concerne l'historique des cours et les informations relatives aux produits financiers, nous vous fournissons :

- plusieurs fichiers CSV contenant l'historique des cours pour les actions
- un fichier XML représentant les actualités concernant certains produits financiers

On vous demande d'écrire des scripts permettant d'importer ces fichiers dans votre base de données. Ces données devront être présentes dans votre base de données lors de la défense. Les fichiers CSV et XML seront disponible le 8 novembre sur la page web du projet.

Application

Ensuite, nous vous demandons de développer une interface graphique pour votre base de données permettant au minimum de réaliser les opérations ci-dessous :

- visualiser les actions et les indices, leurs actualités et l'historique de leurs cours
- éditer un produit financier
- créer un produit financier
- supprimer un produit financier

Pour ce dernier point, nous vous demandons d'expliquer et de justifier vos choix de stratégie de suppression. Votre application devra bien entendu veiller à ce que la base de données reste cohérente. Vous pouvez bien sûr ajouter des fonctionnalités à votre application comme par exemple un graphique du cours en fonction du temps. Ces apports personnels seront valorisés.

Requêtes

Nous vous demandons également d'écrire en algèbre relationnelle et calcul relationnel tuple les requêtes R1, R2 et R3 ainsi que toutes les requêtes en SQL :

- R1 : Les actions du marché X qui ont monté pendant toutes les 5 dernières séances.
- R2 : Les noms des actions qui sont associées à toutes les actualités auxquelles est associée l'action portant le nom X.
- R3 : Le nom des actions qui sont associées à au moins trois actualités provenant de la source 'Le canard du Solbosch'.
- R4 : La moyenne du cours d'un produit financier X sur les 30 derniers jours.
- R5 : Le minimum et le maximum du cours d'une action X depuis 12 mois.
- R6 : Les 5 actions ayant monté le plus et les 5 actions ayant baissé le plus au cours des 12 derniers mois ainsi que le pourcentage de leur hausse/baisse.

Enfin, un rapport devra être rendu aux assistants et vous devrez présenter vos scripts d'insertions et votre application pendant 10 minutes selon un scénario que vous déterminerez.

Rapports

Pour la première partie, on vous demande les documents suivants :

- un diagramme entité-association modélisant le projet ainsi que ses contraintes
- une traduction relationnelle de ce diagramme et ses contraintes
- vos hypothèses et vos choix de modélisation justifiés

Le formalisme utilisé doit être un de ceux vu au cours ou au TP.

Pour la deuxième partie, vous rendrez un rapport contenant :

- les documents de la première partie tenant compte des remarques des assistants
- le script SQL DDL de création de la base de données
- des scripts d'importation pour :
 - les données d'initialisation (en SQL)
 - les historiques des cours
 - les actualités des produits financiers
- les requêtes demandées
- les instructions d'installation de votre application
- un scénario de démonstration de votre application
- les explications et justifications de vos choix et hypothèses
- le tableau des responsabilités complété équitablement (voir annexes)

Le code source de votre application devra être rendu par mail à ulb.code@gmail.com avant votre défense.

Informations pratiques

- Le projet se fera obligatoirement par groupes de deux.
- La première partie, en version papier, devra être rendue aux assistants pour le 6 novembre entre 10h et 16h au UB4.129 (bureau de Boris Verhaegen) et être présentée aux assistants le 8 novembre entre 14h et 18h suivant l'horaire établi sur la page web du projet.
- Le rapport de la deuxième partie devra être rendu en version papier et une archive contenant tous les codes sources devra être envoyé à l'adresse ulb.code@gmail.com avant votre défense suivant l'horaire établi sur la page web du projet les 17 et 18 décembre.
- Le projet comptera pour 25% de la note finale du cours.
- Les apports personnels au projet seront valorisés à hauteur de 4 points sur 20. En d'autres mots, un projet parfait sans apport personnel aura une valeur de 16 points sur 20.
- Sauf mention explicite, vous pouvez utiliser les langages et outils de votre choix (MySQL, PostgreSQL, Java, Python, PHP, C++, Ruby, ...).
- Vous pouvez développer sur votre propre machine et présenter vos projets sur un ordinateur portable que vous apporterez lors de la défense. Au besoin, nous mettrons à votre disposition une base de données MySQL ainsi qu'un espace web PHP sur un serveur de l'Université.

Annexes

Marchés

Les marchés suivants devront être encodés dans votre base de données.

Code	Nom	Ville	Pays	Devise
NASD	NASDAQ	New York City	United States	USD
EURO	EURONEXT	Paris	France	EUR

Actions

Les actions suivantes devront être encodées dans votre base de données.

Code	Nom	Marché
MSFT	Microsoft Corporation	NASDAQ
AAPL	Apple Inc.	NASDAQ
GOOG	Google	NASDAQ
DELL	Dell	NASDAQ
RIMM	Research in Motion	NASDAQ
VIV	Vivendi	EURONEXT
FTE	France Telecom	EURONEXT
EN	Bouygues	EURONEXT
AF	Air France - KLM	EURONEXT
EAD	EADS	EURONEXT

Indice

L'indice "Big Three" devra être encodé dans votre base de données. Son code est "BIG3" et appartient au marché "NASDAQ". Cet indice est composé de 1000 actions réparties de la manière suivante :

Code	Quantité
MSFT	500
GOOG	300
AAPL	200

Historique et actualités

Les fichiers CSV et XML contenant respectivement l'historique des cours et les actualités seront publiés sur la page web des TP le 8 novembre 2007. Ceux-ci suivront les spécifications données dans cet énoncé.

Tableau des responsabilités

Dans votre rapport final, nous vous demandons d'insérer le tableau suivant rempli équitablement.

Module	Responsable
Script DDL de création	
Script d'insertion des données d'initialisation	
Script d'insertion de l'historique	
Script d'insertion des actualités	
Requêtes	Les deux membres du groupe
Visualisation des produits financiers	
Visualisation de l'historique	
Visualisation des actualités	
Édition d'un produit financier	
Création d'un produit financier	
Suppression d'un produit financier	

Bon travail !