

BASES DE DONNÉES ET LANGAGE SQL

SITUATION D'APPRENTISSAGE ET D'ÉVALUATION (SAÉ) 2-01

CONCEPTION ET IMPLÉMENTATION D'UNE BASE DE DONNÉES

Silvia Di Gregorio
silvia.digregorio@univ-paris13.fr

Table des matières

| | | |
|----------|---|----------|
| 1 | Contexte de la SAÉ | 1 |
| 2 | Description du travail demandé | 2 |
| 2.1 | Modélisation entités-associations et relationnelle | 2 |
| 2.2 | Script de création SQL et peuplement des tables de la base de données | 2 |
| 2.3 | Interrogation de la base de données et visualisation des résultats | 2 |
| 3 | Modalités, ressources et barème | 2 |
| 3.1 | Travail à réaliser | 2 |
| 3.2 | Individualisation des rapports de SAÉ | 3 |
| 3.3 | Présentation orale des rapports | 3 |
| 3.4 | Format du rapport | 3 |
| 3.5 | Échéances | 3 |
| 3.6 | Barème | 3 |
| 3.7 | Ressources | 4 |

1 Contexte de la SAÉ

Contexte général

L'objectif de cette SAÉ est de concevoir puis d'alimenter une base de données permettant de stocker toutes les informations issues de deux jeux de données fournis. Le travail principal consiste à récupérer des données officielles et à utiliser le langage SQL pour les stocker dans une base de données, puis les explorer afin de produire des analyses.

Commerce des technologies à faible empreinte carbone : situation à modéliser

La structure de la base de données doit modéliser toutes les informations incluses dans les jeux de données fournis.

Vous devez réfléchir à toutes les contraintes possibles liées à un cas réel et les inclure dans votre modèle. Celui-ci doit être normalisé autant que cela est pertinent.

La source de données que vous devez télécharger concerne le commerce des technologies à faible empreinte carbone : *Trade in Low Carbon Technology Products* et *Bilateral Trade in Low Carbon Technology Products*. Pour en savoir plus sur ces ensembles de données, vous pouvez lire *About: Trade in Low Carbon Technology Products* et *About: Bilateral Trade in Low Carbon Technology Products*.

Mise en place de l'environnement

Pour bien faire cette SAÉ, vous aurez besoin des logiciels suivants :

1. Le SGBDR *PostgreSQL* (choix imposé) ;

2. Un logiciel permettant de dessiner des modèles entités-associations comme ceux vus en cours. Vous pouvez, par exemple, utiliser le logiciel libre de création de diagrammes Dia ou LaTeXDraw (choix libre).
3. Un logiciel pour visualiser les résultats de vos requêtes (choix libre). Par exemple, vous pouvez utiliser *Metabase* (comme pendant le premier semestre) ou *Tableau*, que vous utilisez déjà dans le cours de Reporting et Data Visualisation.

2 Description du travail demandé

2.1 Modélisation entités-associations et relationnelle

L'objectif de cette première partie est de réaliser un modèle entités-associations de la base de données en respectant le cahier des charges de la Section 1. Ce modèle doit respecter la syntaxe vue en cours. Vous pouvez utiliser un logiciel de création de diagrammes. Ensuite, il faut déduire de cette modélisation le schéma relationnel.

Travail demandé

1. Modèle entités-associations respectant la syntaxe du cours.
2. Schéma relationnel (normalisé le plus possible).
3. Le rapport doit contenir aussi la description des entités, des associations et des tables (les cardinalités, les attributs, les clés, etc).

2.2 Script de création SQL et peuplement des tables de la base de données

L'objectif de la deuxième partie est de produire manuellement un script SQL de génération de la structure de la base de données. Ensuite, à partir du fichier plat fourni, il vous est demandé de réaliser un script SQL de projection des données contenues dans ce fichier afin de peupler les tables de votre base de données.

Travail demandé

1. Script SQL de création des tables.
2. Script SQL de peuplement des tables.
3. Description commentée des différentes étapes de votre script de peuplement.

2.3 Interrogation de la base de données et visualisation des résultats

L'objectif de la dernière partie est de démontrer votre capacité à interroger une base de données de manière raisonnable et cohérente. Vous devez créer des requêtes SQL pour interroger la base de données et visualiser les résultats (amusez-vous).

Travail demandé

1. Interrogation variée de la base de données.
2. Script SQL concernant les requêtes.
3. Visualisation et commentaire des résultats.

3 Modalités, ressources et barème

3.1 Travail à réaliser

Vous devez réaliser un rapport constitué des trois parties décrites précédemment :

1. Modélisation entités-associations et relationnelle (cf. Section 2.1)
2. Script de création SQL et peuplement des tables de la base de données (cf. Section 2.2)
3. Interrogation de la base de données et visualisation des résultats (cf. Section 2.3)

Pour chacune de ces sections, vous devrez préciser la contribution de chaque membre du groupe.

3.2 Individualisation des rapports de SAÉ

Le travail se fera par groupe de deux ou trois étudiants. Un seul rapport devra être soumis par groupe.

L'originalité de votre rapport sera examinée qualitativement par votre enseignant et quantitativement par un logiciel anti-plagiat. En cas de similarité trop importante avec un autre rapport (de cette promotion ou d'une autre promotion), une pénalité pouvant aller jusqu'à 12 points sera appliquée.

Le diplôme du BUT est un diplôme individuel qui atteste du niveau individuel de l'étudiant. Toute évaluation doit donc garder cette double caractéristique d'individualité et de représentativité. Hélas, un travail en groupe et en autonomie peut permettre de s'écarter de cette double caractéristique. Par conséquent, pour éviter tout dévoiement excessif, la note de ce travail est plafonné à 200% de la meilleure note obtenue lors de la présentation orale de la SAÉ ou de la ressource sur laquelle elle s'appuie.

3.3 Présentation orale des rapports

Lors de la dernière séance, tous les groupes présenteront leur travail. Chaque groupe disposera de 15 minutes. Tous les membres doivent intervenir pendant une durée à peu près équivalente. La répartition des sujets sera décidée en interne par les membres du groupe.

Durant cette présentation, vous devrez :

- Expliquer les choix réalisés et les éventuelles difficultés rencontrées.
- Indiquer explicitement la contribution de chaque membre.

Ensuite, l'enseignante posera des questions pendant 5 minutes.

3.4 Format du rapport

Votre rapport doit être justifié, clair, bien écrit, bien présenté et bien structuré. Il doit comporter tous les éléments d'identification nécessaires (nom, prénom, groupe, promotion, nom de la SAÉ...). Vous devez utiliser les styles, faire apparaître un petit sommaire, indenter correctement et utiliser une police à chasse fixe pour tout ce qui concerne du code (instructions SQL par exemple).

Tout le texte ainsi que le code SQL doivent pouvoir être extractibles facilement. Bien entendu, vous pouvez insérer des images, des copies d'écran de résultats de requêtes SQL dans votre rapport, mais pas des copies d'écran de requêtes SQL elles-mêmes.

Le rapport doit être rendu dans un format PDF (Portable Document Format) standard. Le nom du rapport doit être de la forme suivante : <nom1>_<nom2>_sae201.pdf pour un groupe de deux étudiants, ou <nom1>_<nom2>_<nom3>_sae201.pdf pour un groupe de trois étudiants. Les noms doivent apparaître dans l'ordre alphabétique. Le nombre maximum de page du rapport est fixé à 15. Les rapports dépassant cette limite ne seront pas lus.

3.5 Échéances

Le rapport ainsi que les scripts SQL doivent être envoyés avant le 10 juin 2025 à minuit sur Moodle. Une seule personne par groupe peut soumettre les fichiers.

Remarque : une pénalité de 2 points par jour de retard sur la date de remise sera appliquée.

3.6 Barème

Forme sur 6 points

Trois aspects de la forme seront pris en considération :

- Respect des consignes (pdf standard, code SQL en clair et indenté, ...);
- Qualité de l'expression (orthographe, syntaxe, qualité des phrases, ...);
- Qualité de la présentation (clarté, utilisation des styles, structuration, ...).

La notation reflétera principalement l'aspect le plus faible parmi ces trois critères.

Modélisation entités-associations et relationnelle sur 6 points

Seront évaluées la syntaxe, la correction et la pertinence du modèle et du schéma. Une bonne explication est fondamentale.

Script de création SQL et peuplement des tables de la base de données sur 4 points

Seront évaluées la syntaxe, la réussite du peuplement des tables, la pertinence des instructions du script, la pédagogie et la qualité des explications données.

Interrogation de la base de données et visualisation des résultats sur 4 points

Seront évaluées la syntaxe, la pertinence de vos requêtes, l'intérêt de vos requêtes, la visualisation de leur résultats, ainsi que la qualité de vos commentaires.

3.7 Ressources

Fichier des données

Les sources de données que vous devez télécharger :

1. <https://climatedata.imf.org/datasets/1d33174e9e46429d9e570d539556f66a/explore>
2. https://climatedata.imf.org/datasets/975bc577fe7342c2a3651e8841959c47_0/explore

Pour plus d'informations :

1. https://climatedata.imf.org/datasets/1d33174e9e46429d9e570d539556f66a_0/about
2. https://climatedata.imf.org/datasets/975bc577fe7342c2a3651e8841959c47_0/about

Installation de PostGreSQL

<https://www.postgresql.org/download/>

<http://gurau-audibert.hd.free.fr/josdblog/2015/03/mise-en-oeuvre-postgresql/>

DIA ou LaTeXDraw (Choix libre de l'éditeur de diagrammes)

<http://dia-installer.de/>

<https://latexdraw.sourceforge.net>

Metabase ou Tableau (Choix libre du logiciel pour la visualisation)

<https://www.metabase.com/start/oss/>

<https://www.tableau.com/resource/data-visualization>