

Langages d'exploitation des bases de données

Projet final

420-C42

Énoncé

Vous devez concevoir et réaliser une base de données permettant de répondre à une problématique donnée.

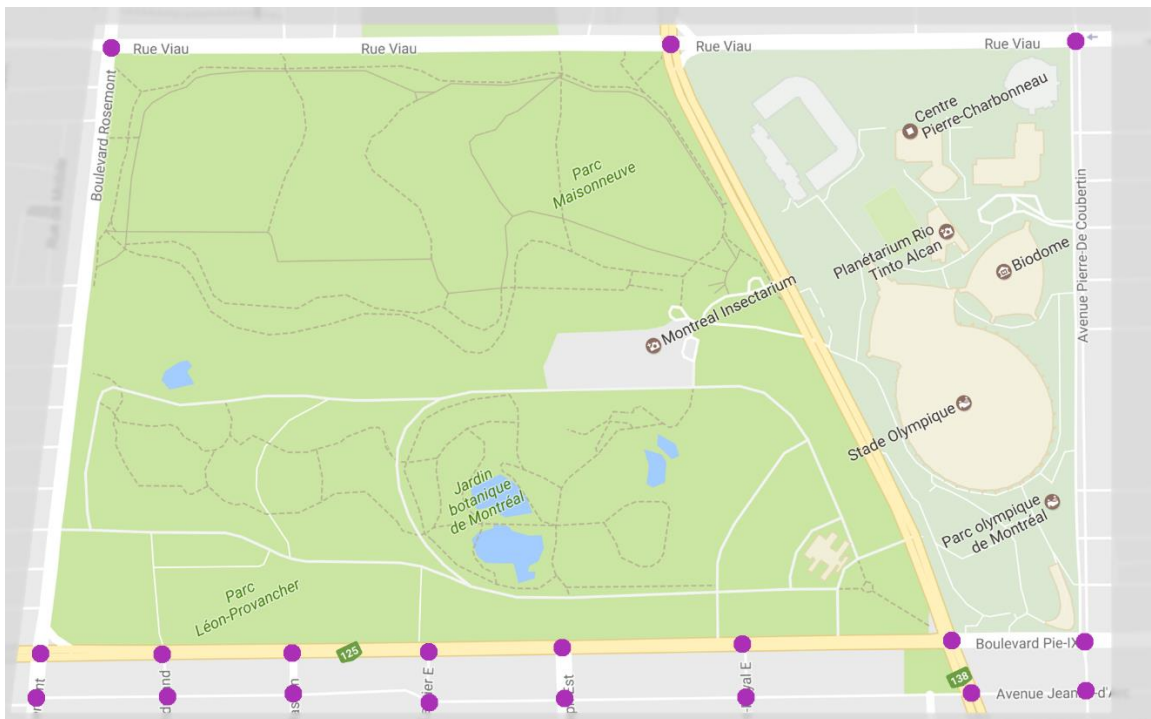
Ce projet se veut un deuxième contact avec ce que pourrait être un mandat en industrie (avec quelques simplifications évidemment). Pour y arriver, vous devrez réaliser 5 étapes distinctes, mais interdépendantes :

1. conception de la base de données;
2. réalisation de la structure (scripts créant les objets fondamentaux de la base de données : tables, contraintes, indexes, vues, séquences, etc.);
3. réalisation d'outils d'automatisation (scripts créant des objets plus avancés: fonctions, requêtes préparées, procédures et fonctions SQL et PL/pgSQL, déclencheurs, etc.);
4. réalisation des scripts permettant de populer la base de données;
5. réalisation des requêtes demandées.

Problématique

Mise en situation

Vous travaillez pour une jeune firme cherchant à décrocher un important contrat pour le Ministère du Transport du Québec (MTQ). Votre entreprise doit mettre de l'avant une solution d'ingénierie complexe visant à faire l'inspection de l'état des routes du réseau routier du Québec. Tout un défi!!! Votre patron a réussi à franchir les premières étapes du processus d'appel d'offres et l'entreprise se trouve parmi les finalistes. La décision définitive se prendra sur l'évaluation d'une preuve de concept qui doit être présentée. La démonstration à réaliser doit inclure la prise de mesures réelles d'une portion de la ville de Montréal. Voici la carte routière qui doit être inspectée pour l'évaluation :



L'entreprise divise son groupe de développement en quatre sous-groupes : mécanique, électrique, science et informatique. Du groupe informatique, on a créé un petit sous-groupe dont la tâche consiste à développer une base de données simplifiée qui servira justement de prototype et de preuve de concept pour décrocher le contrat.

Attention, la solution présentée doit être générique et correspondre à une solution se rapprochant d'un produit réel. Autrement dit, la solution ne doit pas être limitée à la zone géographique imposée pour la démonstration.

Données à traiter

Voici les données à stocker :

- employé
 - nom 32 caractères max, obligatoire
 - prénom 32 caractères max, obligatoire
 - genre (un choix entre F, H, ou X) obligatoire
 - numéro d'assurance social 9 caractères fixes, obligatoire
 - date d'embauche date, limité au 1^{er} janvier 2018 à aujourd'hui
 - salaire horaire nombre variant entre 15.00 et 250.00, obligatoire, défaut 27.50
 - un poste selon les choix suivants : professionnel, technicien, ingénieur, scientifique, manutentionnaire et soutien (on veut pouvoir ajouter d'autres postes) obligatoire
 - un département selon les choix suivants : administration, ventes et représentation, achats, mécanique, électrique, informatique et recherche (on veut pouvoir ajouter d'autres départements) obligatoire
- réseau routier :
 - pour chaque tronçon¹ de rue entre deux intersections :
 - nom de la rue 32 caractères max, obligatoire
 - les deux intersections reliées (il doit être possible de différencier l'intersection qui débute le tronçon et celle qui termine le tronçon) obligatoires
 - longueur du tronçon [m] nombre variant entre 0.0 à 100000.0, obligatoire
 - limite de vitesse [km/h] nombre variant de 25 à 120, obligatoire
 - nombre de voies nombre variant entre 1 à 8, obligatoire, défaut 1
 - on désire localiser différents types d'infrastructures tout au long du tronçon :
 - n panneaux de signalisation ($n \geq 0$) :
 - le type : arrêt, céder, limite de vitesse, ... on veut pouvoir en ajouter d'autres, obligatoire
 - une position relative sur le tronçon en % où 0% est le début et 100% la fin de 0.00% à 100.00%, obligatoire
 - n dispositifs particuliers ($n \geq 0$) :
 - le type : accès fauteuil roulant, signal audio pour piéton, etc. on veut pouvoir en ajouter d'autres, obligatoire
 - une position relative sur le tronçon en % où 0% est le début et 100% la fin de 0.00% à 100.00%, obligatoire

¹ Un tronçon est défini comme étant une section de rue à sens unique partant d'une première intersection vers une seconde. Si la rue est à double sens, 2 tronçons doivent être créés (de l'intersection A à l'intersection B et de B à A).

- n dispositifs de signalisation lumineuse ($n \geq 0$) :
 - m lumières ($1 \leq m \leq 8$) et pour chacune d'entre elle :
 - forme : ronde, carrée, losange, flèche selon 8 directions, humains, main, vélo, barre verticale, barre horizontale
on veut pouvoir en ajouter d'autres, obligatoire
 - couleur : rouge, jaune, vert, blanche ou personnalisée rouge/vert/bleu encodé en texte hex
on veut pouvoir ajouter d'autres couleurs nommées, obligatoire
 - mode : solide, clignotant, contrôlé, intelligente
choix immuables, obligatoire
 - orientation : horizontale, verticale ou autre
choix immuables, obligatoire
 - une position relative sur le tronçon en % où 0% est le début et 100% la fin
de 0.00% à 100.00%, obligatoire
 - type de pavage selon ces choix : asphalte, ciment, pavé brique, pavé pierre, non pavé, indéterminé
choix immuables, obligatoire
 - pour chaque intersection² du réseau routier à couvrir :
 - identifiant
nombre de 7 chiffres, obligatoire
 - les coordonnées en latitude et longitude (par exemple, l'intersection Pie-IX et Sherbrooke est : 45.554618, -73.554570)
attention, on vous demande de déterminer vous-même le type et format requis, obligatoire
 - type de pavage selon ces choix : asphalte, ciment, pavé brique, pavé pierre, non pavé, indéterminé
choix immuables, obligatoire
- profileur laser :
 - marque du capteur
32 caractères max, obligatoire
 - numéro de série
16 caractères fixes, obligatoire
 - date de fabrication
date
 - date d'acquisition
date
 - toutes les calibrations effectuées sur le capteur
obligatoire
 - date et heure du début de la procédure de calibration
date, obligatoire
 - date et heure de la fin de la procédure de calibration
date, obligatoire
 - employé qui a fait la calibration
obligatoire
 - les trois valeurs de calibration (v1, v2 et v3)
nombres variant entre -1000 et 1000 avec une résolution de 0.0001, obligatoire
- véhicule d'inspection :
 - marque
32 caractères max, obligatoire
 - modèles
32 caractères max, obligatoire
 - date d'acquisition
date
 - immatriculation
6 caractères fixes, obligatoire
- inspection :
 - date et heure du début de l'inspection
date, obligatoire
 - date et heure de la fin de l'inspection
date, obligatoire

² Une intersection correspond à la zone transitoire entre 2 tronçons. Les intersections possèdent entre 2 à 8 tronçons liés.

- les informations reliées au véhicule utilisé pour l'inspection obligatoire
 - le véhicule utilisé obligatoire
 - l'employé qui a conduit le véhicule obligatoire
 - le kilométrage au compteur au début de l'inspection nombres variant entre 1 et 500 000, obligatoire
 - le kilométrage au compteur à la fin de l'inspection nombres variant entre 1 et 500 000, obligatoire
- les informations reliées au profileur laser obligatoire
 - le profileur laser utilisé obligatoire
 - l'employé qui a opéré le profileur laser obligatoire
- tous les voies routières parcourues pour cette inspection (pour chaque tronçon inspecté, vous devez indiquer la voie inspectée où 1 est la voie la plus à droite) obligatoire
- chemin d'accès du fichier de données 1024 caractères max, obligatoire
- nom du fichier de données : PZ2_00000000_YYMMDDHHMMSS.exte où :
 - 00000000 indique nombre de 8 caractères incrémenté globalement pour toutes les inspections (la première valeur doit être 20)
 - YYMMDDHHMMSS indique la date et l'heure de l'inspection par groupe de 2 nombres pour l'année, le mois, le jour, l'heure, la minute et la seconde
 - exte indique l'extension limité à ces choix : xdat, jdat, bdat, kdat.

Le nom du fichier doit être généré automatiquement par le serveur. L'extension doit être générée aléatoirement parmi les valeurs proposées.

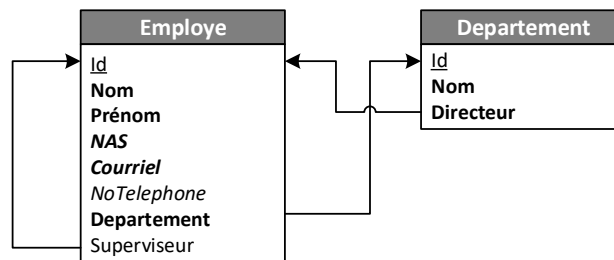
30 caractères fixes, obligatoire

Détails sur les tâches à réaliser

Conception

Vous devez prendre le temps de réfléchir à ce projet. Le défi est de le faire en équipe et de trouver une solution adéquate qui assure l'intégrité des données. Vous devez vous assurer de n'avoir aucune redondance structurelle et de respecter la 3^e forme normale.

On vous demande de créer un graphique typique de base de données. Vous pouvez inventer votre propre notation pourvu que vous ajoutiez une légende explicative. Voici un exemple typique :



Exemple de schéma relationnel

Ce schéma est un bon exemple, car on vous demande d'ajouter clairement toutes les contraintes suivantes sur le graphique : clé primaire (soulignée), clé étrangère (illustré par une flèche directionnelle),

unicité (en caractères italiques) et obligatoire (en caractères gras). Dans tous les cas, vous devez ajouter une légende explicative de votre notation.

Cette partie du travail, souvent sous-estimée et bâclée, est probablement la plus importante de tout le projet. Vous pouvez remettre un document réalisé avec un programme de dessin vectoriel comme Visio ou produit à la main (à condition qu'il soit propre, numérisé et intégré dans votre remise en version électronique – le document papier n'est pas à remettre).

Création de la structure (base de données et objets fondamentaux)

Vous devez créer un script (on format `pgsql`) permettant de créer automatiquement tous les objets fondamentaux de la base de données :

- tables
- contraintes
- séquences (vous devez avoir au moins une séquence)
- indexes (vous devez avoir au moins un index)
- vues (vous devez avoir au moins une vue)

Création d'outils d'automatisation

Vous devez créer un script permettant de créer au minimum les objets PL/pgSQL suivants :

1. vous devez créer au moins 4 procédures permettant de faciliter les insertions tout en assurant l'intégrité de la base de données (dans les tables de votre choix);
2. vous devez créer au moins 4 fonctions utilitaires pertinentes pour vos requêtes et scripts;
3. vous devez créer au moins un déclencheur.

Dans tous les cas, vous devez faire l'utilisation de ces procédures/fonctions dans le script d'insertion là où pertinent.

Insertion d'enregistrements dans la base de données

On vous demande de rédiger un script permettant de populer les différentes tables avec des données pertinentes et cohérentes les unes avec les autres. Vous devez respecter ces nombreuses contraintes :

- au moins 4 [3]³ employés différents (les employés doivent avoir vos noms);
- tel que stipulé dans le mandat d'évaluation, le réseau routier identifié par les points magenta de la carte donnée sauf pour toutes les intersections de la rue Jeanne-d'Arc (vous y trouverez 11 intersections et les données doivent être aussi réalistes que possible);
- au moins 3 [2]³ profileurs laser :
 - le premier avec une seule calibration
 - le second avec 3 [2]³
 - le dernier avec plus de 5 [3]³ calibrations et il est important que la calibration la plus récente respecte cette égalité : $v_1 = v_2 = v_3$
- au moins 2 véhicules;
- au moins 100 inspections totalisant au moins 1000 tronçons.

³ Les indications entre crochets [x] représentent la quantité pour les groupes de taille réduite.

Rédaction de requêtes spécifiques

On vous demande de créer un script avec les requêtes répondant aux questions suivantes :

1. Donner la liste des employés : nom, prénom, poste, nom du département, ancienneté (en année et mois), leur salaire annuel (considérant qu'ils travaillent 35 heures par semaine et 52 semaines par année) et leur salaire annuel augmenté de 15%.
2. Donner le nombre de calibrations que chaque employé a fait.
3. Donner le nombre d'inspections que chaque employé a fait.
4. Pour chaque véhicule, combien de kilomètres ont été parcourus pour réaliser les inspections.
5. Pour chaque véhicule, combien de kilomètres de tronçons ont été parcourus pour réaliser les inspections.
6. Pour chacune des inspections, on désire savoir quels ont été les frais associés (vous devez tenir compte du temps passé pour les deux employés lors de l'inspection, des coûts d'exploitation du véhicule à 4.79\$ par kilomètre et de 1/25 du temps requis pour faire la calibration du capteur utilisé).
7. On veut la liste des profileurs laser ayant besoin d'être calibrés. La formule suivante permet de valider cette information. Si cet énoncé est vrai, une calibration est requise :

$$\sqrt{\left| \frac{v_1 v_2}{v_3^2} - 1 \right|} \leq \frac{1}{\pi^2}$$

8. Chaque membre de l'équipe doit réaliser une requête **pertinente** de son cru. Considérant toutes les requêtes produites par l'équipe, vous devez respecter ces contraintes :
 - il doit avoir au minimum 3, 4 et 5 tables dans l'une de vos requêtes,
 - vous devez utiliser ces clauses: GROUP BY, HAVING, ORDER BY, LIMIT,
 - vous devez avoir au moins une requête corrélée.

À même le script et avant ces requêtes, vous devez ajouter un commentaire expliquant précisément ce que fait la requête. Le pointage associé à chacune de ces requêtes est autant basé sur la pertinence que la qualité de la réalisation technique.

9. [Boni] La distance entre deux intersections (votre requête doit prendre 4 noms de rue en entrée). Par exemple, la distance entre l'intersection de la rue Viau et du Boulevard Rosemont et l'intersection de la rue Viau et de la rue Sherbrooke est.

Pour chacune des requêtes du script, vous devez respecter la structure suivante en y ajoutant les informations adéquates :

```
-- =====
-- Requête #
-- Objectif :    ...
--              ...
--              ...
-- Évaluation :  ...
--              ...
--              ...
-- Réalisé par : ...
-- Aidé par    : ...
-- =====
votre requête
-- =====
```

Vous devez donc indiquer :

- le numéro de la requête créée;
- l'objectif : ce que fait la requête;
- votre évaluation du résultat obtenu : certaines requêtes sont complexes et il est possible que vous ayez de la difficulté à la réaliser; dans ce cas, mettez ce que vous avez réussi à faire et expliquer ce qui fonctionne et ce qui ne fonctionne pas – soyez concis et précis.

ATTENTION : les informations demandées doivent être complétées adéquatement sinon vous aurez automatiquement zéro.

Contraintes

Plusieurs contraintes sont à respecter pour ce travail :

- Vous devez travailler en équipe de 4 ou 3.
- Il est très important que tous les membres de l'équipe travaillent équitablement, car l'évaluation finale tient compte de plusieurs critères, dont la répartition de la tâche de travail.
- Vous devez travailler avec le SGBDR PostgreSQL utilisé tout au long de la session.
- Les données que vous devez insérer doivent être le plus réaliste possible.
- Puisque le travail est pour le MTQ, on vous demande de créer tous vos objets et votre documentation en français.
- La date de remise est non négociable.

Stratégie d'évaluation

L'évaluation se fera en 2 parties. D'abord, l'enseignant évaluera le projet remis et assignera une note de groupe pour le travail. Ensuite, chaque équipe devra remettre un fichier Excel dans lequel sera soigneusement reportée une cote représentant la participation de chaque étudiant dans la réalisation du projet. Cette évaluation est faite en équipe et un consensus doit être trouvé.

Une pondération appliquée sur ces deux évaluations permettra d'assigner les notes finales individuelles.

Ce projet est long et difficile. Il est conçu pour être réalisé en équipe. L'objectif est que chacun prenne sa place et que chacun laisse de la place aux autres.

Ainsi, trois critères sont évalués :

- **participation** (présence en classe, participation active, laisse participer les autres, pas toujours en train d'être sur Facebook ou sur son téléphone, concentré sur le projet, pas en train de faire des travaux pour d'autres cours, ...)
- **réalisation** (répartition du travail réalisé : conception, modélisation, rédaction de script, documentation, ...)
- **impact** (débrouillardise, initiative, amène des solutions pertinentes, motivation d'équipe, ...)

Remise

Vous devez créer un fichier de format `zip` dans lequel vous insérez :

- votre schéma de conception schema.pdf
- le script de création de la base de données creation_structure.sql
- le script des objets complexes creation_outils.sql
- le script permettant d'ajouter des données insertions.sql
- le script des requêtes demandées requetes.sql
- optionnellement, vos impressions et commentaires sur ce travail commentaires.txt
- le fichier Excel rempli sur la participation active des membres du groupe evaluation.xlsx

Vous devez remettre votre projet une seule fois sur Lea après avoir nommé votre fichier :

NomPrenomEtudiant1_NomPrenomEtudiant2[_NomEtudiantN].zip