

Contrôle périodique

N. Saunier

20 octobre 2020

Introduction

Dans ce contexte inhabituel de contrôle périodique en ligne, veuillez:

- noter qu'il comporte 6 questions comptant respectivement pour 2, 3, 5, 1, 5 et 4 points (la note totale est sur 20), sans retour possible entre les questions, pour une durée de deux heures;
- noter que tous les documents sont autorisés;
- apporter une attention particulière à la rédaction et la clarté du résultat ou de la réponse que vous donnez à chaque question;
- noter que certains exercices nécessitent les données récemment publiées sur la fréquentation des voies actives et sécuritaires (VAS), mises dans une base de données SQLite pour les exercices du contrôle disponible dans la section du contrôle périodique;
 - j'ai ajouté dans les tables "pietons" et "velos" des colonnes extraites de la colonne d'horodatage du comptage: année (colonne "annee"), mois de l'année (colonne "mois"), jour de semaine (commençant par 0 pour dimanche) (colonne "jour") et heure de la journée (colonne "heure");
- noter que vous pouvez déposer des fichiers pour répondre à certaines questions.

Exercice 1

(/2 pts)

Proposer une méthode permettant de collecter les comptages des piétons/cyclistes enregistrés dans la table "pietons" de la base de données sur les VAS. Cette méthode peut-elle collecter des données de vitesse sur les piétons? Si non, proposer une méthode permettant de mesurer la vitesse des piétons.

Solution

Des méthodes permettant de compter les piétons sont: le comptage manuel, l'utilisation d'un capteur infrarouge ou d'un capteur vidéo avec algorithmes de vision par ordinateur. Pour les cyclistes, on peut mentionner les méthodes manuelle, l'utilisation de boucles magnétique ou de capteur vidéo avec algorithmes de vision par ordinateur.

Certaines de ces méthodes ne permettent pas de relever la vitesse. Les méthodes fonctionnant par détection, si elles sont doublées, permettent de déterminer la vitesse à partir des instants de passage. Les positions des usagers dans des images vidéo successives permettent de déterminer leur vitesse. Le seul capteur permettant de mesurer directement la vitesse instantanée est le radar (et il permet aussi de compter), quoiqu'il soit plus rarement utilisé pour ces usagers.

Exercice 2

(/3 pts)

La Ville de Montréal désire faire une enquête sur les motifs des piétons se déplaçant sur la VAS. Quelle est la population de référence pour cette enquête? De quel type d'enquête s'agit-il et quelle technique d'enquête recommandez-vous?

Solution

- La population de référence est constituée des piétons se déplaçant sur la VAS.
- Le type d'enquête est une enquête section, se déroulant sur une période donnée, par exemple quelques jours ou semaines.
- La technique d'enquête la plus naturelle est l'interception, consistant à arrêter les piétons sur la VAS pour leur poser des questions.

Exercice 3

(/5 pts)

Nous nous intéressons à créer un système d'information pour la collecte de données de comptages et de vitesses sur des rues de Montréal. Pour cela, il est nécessaire de modéliser les différents objets et concepts nécessaires à son fonctionnement. Ces entités sont (au minimum) les suivantes: usager, ligne de comptage, passage, rue, capteur, véhicule. Un usager, possiblement dans un véhicule, sera compté lorsqu'il franchit une ligne de comptage (virtuelle) par un capteur placé sur une certaine rue.

1. Proposer un modèle de données sous forme d'un diagramme Entité/Association impliquant toutes les entités listées ci-dessus. Ajouter des attributs (en indiquant les identifiants) et les associations entre entités avec leurs cardinalités, minimale et maximale, et les fonctionnalités.
2. Traduire le schéma Entité/Association en schéma relationnel. Indiquer clairement les clefs primaires et externes (et à quoi les clefs externes font référence), et proposer des types pour les attributs.

La réponse peut être faite dans le cadre de texte, dans un document rédigé ou écrit à la main et pris en photo.

Solution

1. Les entités et leurs attributs sont les suivants (l'identifiant de chaque entité est en **gras**):

usager id, nom, date de naissance, genre

ligne de comptage id, coordonnée origine, coordonnées fin, horodatage de début d'installation, horodatage de fin d'installation

passage id, horodatage

rue id, nom, ligne (champ vectoriel de type ligne), ville

capteur id, type de capteur, coordonnée de positionnement

véhicule numéro d'immatriculation, type de véhicule, marque, modèle

Les associations sont les suivantes (il est souhaitable de nommer les associations et de faire un schéma):

- rue-ligne de comptage: une rue a eu (historiquement) 0-n lignes de comptage, une ligne de comptage est installé sur 1-1 rue. La fonctionnalité est 1-n.
- ligne de comptage-capteur: une ligne de comptage nécessite 1-1 capteur, un capteur peut avoir été installé sur 0-n lignes de comptage. La fonctionnalité est n-1.
- passage-ligne de comptage: un passage a lieu à 1-1 ligne de comptage, une ligne de comptage peut compter 0-n passages. La fonctionnalité est n-1.
- passage-usager: un passage implique 1-1 usager, un usager peut faire 0-n passages. La fonctionnalité est 1-n.
- passage-véhicule: un passage implique 0-1 véhicule, un véhicule peut faire 0-n passages. La fonctionnalité est 1-n.

La ligne de comptage peut être vue comme une association entre rue et capteur (positionnement du capteur pour compter sur une rue pendant une période donnée). Si plusieurs usagers passent la ligne dans le même véhicule, on ferait autant de passages que d'usagers, en les liant au même véhicule.

2. Chaque entité devient une table (Usagers, LignesComptage, Passages, Rues, Capteurs, Vehicules). Il n'est pas nécessaire d'ajouter de tables puisqu'il n'y a pas d'associations n-m Il faut ajouter des clefs externes suivantes pour les associations 1-n:

- rueId et capteurId dans LignesComptage pour faire référence respectivement à Rues.id et Capteurs.id;
- ligneId dans Passages pour faire référence à LignesComptage.id;
- usagerId et vehiculeId dans Passages pour faire référence respectivement à Usagers.id et Vehicules.id.

Exercice 4

(/1 pt)

Quel est la clef primaire de la table piétons/velos? Quel(s) champs devrait-on déclarer comme clef étrangère?

Solution

La clef primaire de la table "piétons"/"velos" est la clef composite (id_station, timestamp_local). Le champ id_station est une clef étrangère qui fait référence au champ id_station de la table "stations".

Exercice 5

(/5 pts)

Veuillez extraire l'information suivante de la base de données sur les VAS à l'aide de requête SQL. Donner la réponse (un extrait si elle est trop longue) et la requête SQL utilisée pour obtenir la réponse.

1. quelle est la station avec l'achalandage (piétons/vélos) le plus grand par jour au mois de juin/juillet/août/septembre?

2. quel est le VAS (défini par la colonne "nom_vas" de la table "station") avec l'achalandage (piétons/vélos) le plus élevé par jour en semaine/fin de semaine?
3. pour un mois de votre choix, créer une vue des comptages horaires moyens par jour de semaine pour chaque station (avec son "nom_station" dans le résultat).
4. pour une station de votre choix (à préciser), calculer le rapport entre les deux directions de déplacement selon les jours de semaine.
5. afficher le nombre moyen de piétons/cyclistes par jour de semaine et mois pour les stations ayant une moyenne de piétons/cyclistes par période de 15 min supérieure à 150/40.

Solution

Voici les requêtes pour obtenir la bonne réponse:

1. pour l'achalandage cycliste en juillet:

```
SELECT id_station, jour, SUM(compte_total) FROM velos
WHERE mois = 7 GROUP BY id_station, jour
ORDER BY AVG(compte_total) DESC
```
2.

```
SELECT stations.nom_vas, jour, AVG(compte_total) FROM velos, stations
WHERE (jour BETWEEN 1 and 5) AND (velos.id_station = stations.id_station)
GROUP BY stations.id_station, jour ORDER BY AVG(compte_total)
DESC
```
3. pour le mois de juin:

```
CREATE VIEW profil_horaire AS
SELECT stations.nom_vas, heure, AVG(compte_total) FROM velos,
stations
WHERE (mois = 6) AND (velos.id_station = stations.id_station)
GROUP BY stations.id_station, jour, heure
```
4. pour la station V07, avec les deux façons de calculer le ratio moyen comme la moyenne des ratios ou le ratio des sommes:

```
SELECT jour, SUM(compte_dir_nord) / SUM(compte_dir_sud),
AVG(compte_dir_nord / compte_dir_sud)
FROM velos WHERE id_station='V07' GROUP BY jour
```
5. (notez la sous-requête)

```
SELECT V.id_station, V.mois, V.jour, AVG(V.compte_total)
FROM velos V, (SELECT id_station, AVG(compte_total) as n15
FROM velos GROUP BY id_station) V15
WHERE V15.n15>40 AND V.id_station = V15.id_station
GROUP BY V.id_station, V.mois, V.jour
```

Exercice 6

(/4 pts)

1. Quel est le type de données spatiales utilisé pour enregistrer la localisation des stations (données matricielles ou vectorielles)?
2. Si vous avez à votre disposition une couche de données pour les arrondissements de Montréal (dont la géométrie est un polygone) avec des attributs utiles pour votre analyse, expliquer la procédure à suivre pour combiner les information des arrondissements avec chaque station.

3. En prenant en compte le système de référence spatiale, décrire une procédure pour calculer la distance en ligne droite entre toutes les stations.
4. Discuter le caractère approprié de la géométrie choisie pour la localisation des points de comptage.

Notez que cette question ne requiert aucune manipulation de données.

Solution

1. La localisation des stations est représentée par un point, il s'agit donc de données vectorielles.
2. Il faut faire une jointure spatiale entre les deux couches de données spatiales, qui permettra d'associer les attributs des arrondissements (données zonales) à chaque station (point) selon l'arrondissement dans laquelle la station se retrouve.
3. Il faut d'abord convertir les stations dans un système de référence des coordonnées orthonormé où on peut calculer des distances, par exemple le système de coordonnées MTM. On suit ensuite un algorithme de calcul de la matrice des distance euclidienne (ligne droite) entre chaque pair de station. On initialise une matrice carrée de taille $n \times n$ pour les n stations, puis avec une double boucle sur les stations, on calcule la distance entre chaque station (voir algorithme de calcul de distance vu en cours).
4. La position des lieux de comptage est représentée par un point. Concrètement, sur le terrain, il est inhabituel et peu pratique de compter les usagers de la route passant sur un point. On considère généralement une ligne (virtuelle), qui pourrait être représentée ainsi dans la base de données.