

Contrôle périodique

J.-S. Bourdeau et N. Saunier

3 octobre 2017

Veillez

- noter le barème (la note totale est sur 20) et le temps indicatif à consacrer à chaque exercice;
- indiquer clairement les numéros des questions que vous traitez et vos réponses correspondantes (et souligner ou encadrer les résultats numériques);
- apporter une attention particulière à la rédaction et à la définition des notations que vous employez;
- certains exercices nécessitent des fichiers disponibles sur moodle (Section “Contrôle périodique”).

Des tables statistiques sont disponibles sur moodle.

Exercice 1 (méthodes d'enquêtes)

15 min (/2.5 Pts)

1. À partir de quelle liste est constituée la base de sondage de l'enquête ménage Origine-Destination de la région de Montréal? (0.5 Pt)
2. À quels biais s'attend-on lors de la réalisation de l'enquête ménage Origine-Destination de la région de Montréal? Pourquoi? (1 Pt)
3. À quoi sert le facteur d'expansion? (1 Pt)

Solution

1. La liste de téléphones fixes
2. On s'attend aux biais suivants par exemple:
 - (a) sous-représentation des jeunes qui ont de moins en moins de téléphones fixes;
 - (b) sous-représentation des familles occupées, qui n'ont pas le temps de répondre au téléphone.
3. Le facteur d'expansion sert à pondérer l'échantillon afin qu'il représente la taille de la population de référence.

Exercice 2 (méthodes d'enquêtes)

25 min (/4 Pts)

Le Réseau de transport métropolitain Laurentides est la société de transport qui dessert un territoire de banlieue au nord de Montréal. Vous désirez recueillir des informations concernant le profil des usagers de la ligne d'autobus 23, ainsi que leur profil d'utilisation de cette ligne d'autobus. Le trajet de la ligne d'autobus relie la gare de train Sainte-Thérèse et la municipalité de Sainte-Anne-des-Plaines, en passant par le Collège Lionel-Groulx (cégep).

1. Quelle est la population de référence pour cette enquête? (0.5 Pt)
2. Quelle technique de collecte de données suggérez-vous? Expliquer le fonctionnement de cette collecte. (1 Pt)
3. Quel sera le format du questionnaire? (0.5 Pt)
4. Identifier un cadre temporel approprié pour cette enquête ainsi que l'unité temporelle. Justifier. (0.5 Pt)
5. Quelle est la taille minimale de l'échantillon que vous devez recueillir? Vous désirez un niveau de confiance de 95 % et vous acceptez une marge d'erreur de 4 %. Selon les données de transactions de carte à puce, vous savez que votre population de référence est composée de 2000 individus, et vous voulez vous assurer de respecter la proportion de 75 % d'étudiants dans la clientèle de la ligne. (1.5 Pts)

Solution

1. Les usagers qui utilisent la ligne d'autobus 23
2. Interception, enquête à bord des autobus de la ligne 23
3. Papier ou iPad
4. Un jour de semaine en automne (durant le calendrier scolaire et de travail)
5. La taille minimale de l'échantillon est $n = (1.96^2 \times 0.75(1 - 0.75))/0.04^2 = 450$ (correction pop limitée $n' = 450/(1 + 450/2000) = 367$ individus)

Exercice 3 (modèles de données)

35 min (/6 Pts)

Nous nous intéressons à créer un système d'information pour un intermédiaire en taxi, c'est-à-dire la compagnie qui prend les commandes des clients et les assigne à ses taxis. Pour cela, il est nécessaire de modéliser les différents objets et concepts nécessaires à son fonctionnement. Ces entités sont les suivantes: taxi, conducteur, compagnie (intermédiaire en taxi), course, client, appel (lorsqu'un client réserve une course par téléphone).

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. (4 Pts)
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. (2 Pts)

Solution

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

taxi **numéro d'immatriculation**, marque, modèle

compagnie **id**, nom, adresse du siège social, date de création

conducteur **id**, nom, date de naissance, numéro de permis

course **id**, jour, heure de début, heure de fin, date de début, date de fin, prix

appel **id**, date, heure, origine, destination

client **id**, nom, date de naissance

Les associations sont les suivantes:

- taxi-compagnie: n à m (historiquement)
- taxi-conducteur: n à m (historiquement)
- course-taxi: 1 à n
- course-conducteur: 1 à n
- course-client: 1 à n
- client-appel: 1 à n
- appel-course: 1 à 1

2. Chaque entité devient une table (Taxis, Compagnies, Conducteurs, Courses, Appels, Clients). Il faut ajouter deux tables pour représenter les associations n-m, TaxiCompagnies et TaxiConducteurs, qui sont constituées de deux, clefs externes vers les clefs primaires des tables concernées par les associations. Il faut ajouter des clefs externes: taxiId, clientId et conducteurId dans Courses faisant références respectivement à Taxis.id, Clients.id et Conducteurs.id, clientId dans Appels faisant référence à Clients.id et courseId dans Appels faisant référence à Courses.id. Les types des id seraient numériques (nombre entier), ainsi que numéro de permis, prix numérique (nombre réel), origine et destination en fait deux coordonnées (numérique, nombre réel), les heures et dates leurs types respectifs, nom, marque, modèle, adresse des chaînes de caractères.

Exercice 4 (requêtes SQL)

20 min (/3.5 Pts)

Télécharger la base de données `enquete-od.sqlite` de la section du contrôle périodique et indiquer les requêtes SQL qui permettent de répondre aux questions suivantes (il n'est pas demandé de donner la réponse à la question). Le logiciel SQLitebrowser est aussi disponible si vous désirez tester vos requêtes.

1. Calculer le nombre moyen d'automobiles par ménage. (0.5 Pt)
2. Calculer le nombre déplacements effectués par mode. (1 Pt)
3. Afficher toutes les informations en une seule table, c'est-à-dire avec tous les déplacements effectués par toutes les personnes d'un ménage). (1 Pt)
4. Calculer la distance maximale parcourue par sexe et par mode (1=masculin, 2=féminin). (1 Pt)

Solution

1. `SELECT AVG(Menages.autolog) FROM Menages`
2. `SELECT mode, COUNT(*) FROM Deplacements GROUP BY mode`
3. `SELECT * FROM Menages, Personnes, Deplacements
WHERE Menages.feuillet = Personnes.feuillet
AND Personnes.rang = Deplacements.rang
AND Personnes.feuillet = Deplacements.feuillet`
4. `SELECT mode, sexe, MAX(dist) FROM Personnes, Deplacements
WHERE Personnes.rang = Deplacements.rang
AND Personnes.feuillet = Deplacements.feuillet
GROUP BY mode, sexe`

Exercice 5 (traitement de données)

25 min (/4 Pts)

1. Indiquer les entrées et les sorties (avec les structures de données possibles) d'un algorithme de planification d'itinéraire (du type de celui disponible dans Google Maps). (2 Pts)
2. Sachant qu'on dispose d'une fonction *CompteVehicules(I)* qui compte le nombre de véhicules apparaissant dans une image *I*, indiquer la sortie de l'algorithme suivant (ce qui est calculé): (1 Pt)

entrée: n images I_i **sortie:** ?**début** $x = 0$ $i = 1$ **tant que** $i \leq n$ $x = x + \text{CompteVehicules}(I_i)$ $i = i + 1$ **renvoyer** x **fin**

3. Modifier l'algorithme précédent de sorte qu'il retourne le nombre d'images contenant des véhicules. (1 Pt)

Solution

1. Solution possible:

- Entrées: origine et destination (points, un point est une liste de dimension 2), mode (variable catégorielle), réseau de transport (pour tous les modes autorisés) (un graphe par réseau de transport)

- Sortie: une liste de segments (identifiants) dans le graphe du mode de transport choisi
2. L'algorithme présenté calcule le nombre de véhicules (sortie).
 3. **entrée:** n images I_i
sortie: nombre d'images contenant des véhicules

début

$nImages = 0$

$i = 1$

tant que $i \leq n$

$x = CompteVehicules(I_i)$

si $x > 0$

$nImages = nImages + 1$

$i = i + 1$

renvoyer $nImages$

fin