



Collectif francophone pour l'enseignement libre de l'informatique

Analyse de données écologiques relatives à l'herbivorie du trille

ébauche d'énoncé de problème

CoFELI:Herbivorie_EPP

Luc Lavoie (luc.lavoie@usherbrooke.ca)

CoFELI/Exemple/Herbivorie_EPP, version 022a, en date du 2025-02-03

— version préliminaire —

Mise en garde

Simplification d'un système réel. Plusieurs éléments critiques ne sont pas pris en compte. Ce système ne doit pas être utilisé en pratique.

Historique

diffusion	resp.	description
2025-02-03	LL	Inclusion d'un dictionnaire des types de données et passage à AsciiDoc.
2022-01-30	LL	Revue.
2017-09-17	DAL+LL	Ébauche initiale.

1. Mise en contexte

Un groupe de chercheurs s'intéresse à la migration d'espèces végétales sous l'effet des changements climatiques et à son impact sur les écosystèmes. Par exemple, certaines plantes qui poussent sur les flancs de montagnes pourraient avoir tendance à migrer en altitude. Quel serait alors l'impact sur l'alimentation de certaines espèces animales et les conséquences qui s'ensuivraient probablement ?

En premier lieu, le groupe désire confirmer cette migration et, le cas échéant, la caractériser. Une étude en ce sens a été planifiée en prenant pour cas d'études le trille sur les flancs du mont Mégantic.

Le présent travail vise à doter les expérimentateurs d'un outil leur permettant de colliger puis d'analyser les données d'observation. Votre équipe a reçu le mandat de développer un tel outil en se concentrant tout d'abord sur les statistiques relatives aux attributs physiques des plants selon divers critères de recoupement de population. La première étape de ce travail consiste donc à développer le modèle de données requis.

2. Périmètre du problème

Un plant est décrit par son emplacement (donné par la placette observée et la parcelle au sein de celle-ci), la date de son identification (chaque plant est muni d'une étiquette au numéro unique). Une placette (une aire délimitée choisie pour son intérêt d'études) est décrite par son type de peuplement, par six niveaux d'obstruction latérale ($[\text{feuillue, coniférienne, totale}] \times [\text{à 1 m, à 2 m}]$), par l'importance de la couverture au sol des graminées, des mousses et des fougères ainsi que par les trois variétés d'arbres dominants. Les niveaux d'obstruction latérale et l'importance de la couverture au sol sont donnés selon les catégories

suivantes :

Tableau 1. Catégorisation des taux

Code	min	max
A	75	100
B	50	75
C	25	50
D	5	25
E	1	5
F	absent	

Les observations suivantes sont prises régulièrement par les « esclaves » écologistes tout au long de l'étude : dimensions des feuilles (longueur et largeur en millimètres) ; floraison (le plant est susceptible de participer à la reproduction ou pas, sur la base qu'il porte une fleur ou un fruit) ; état (convenu au sein d'une liste préétablie et codifiée par une lettre). Toutes les observations sont datées et peuvent être accompagnées d'une note (qui est un commentaire consigné par l'observateur dans une langue qu'on espère connue des utilisateurs de la base de données).

Cette présentation est très incomplète et sera affinée tout au long du développement du modèle.

3. Exemples de requêtes

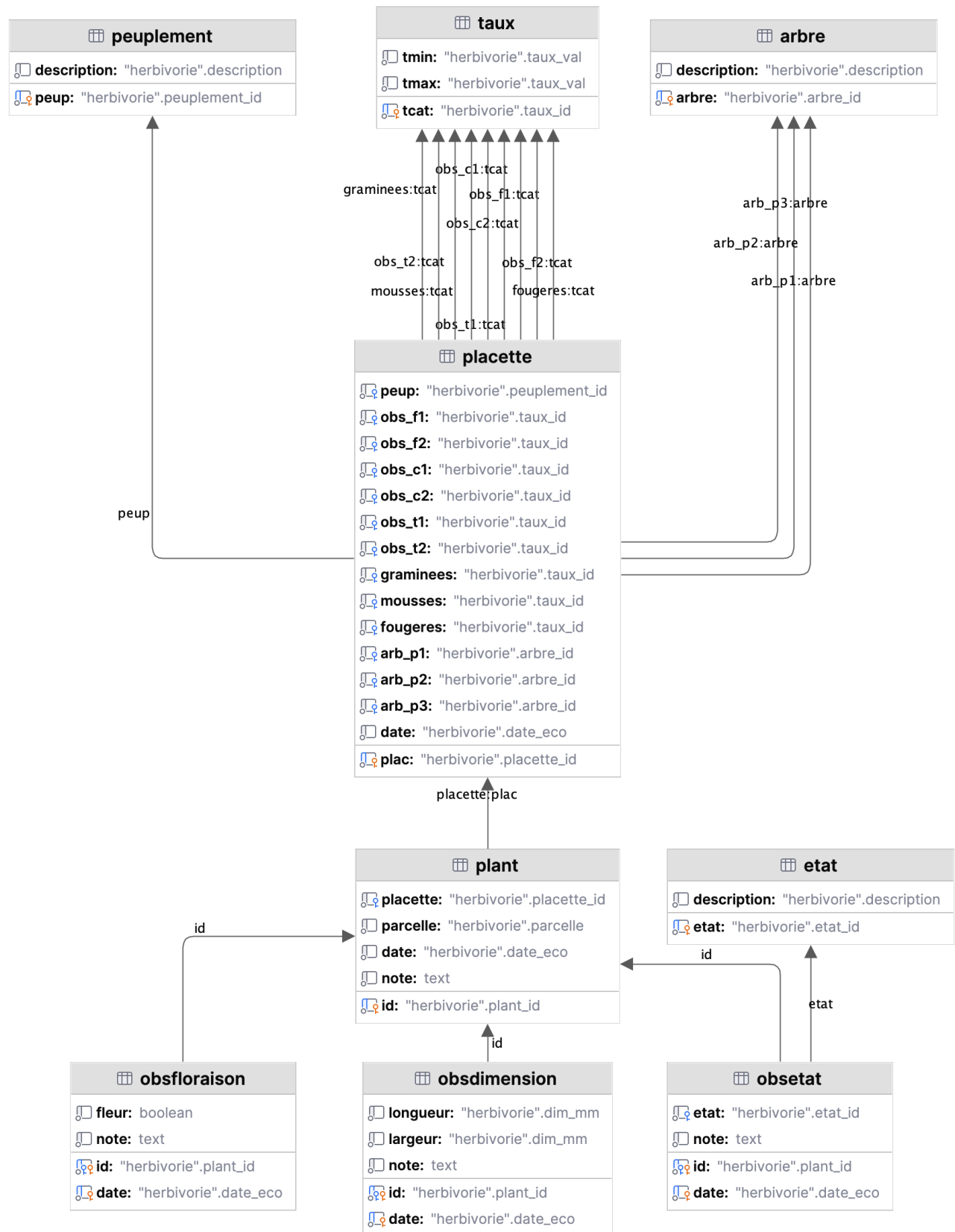
Voici quelques exemples de requêtes qui devront, à terme, pouvoir être exécutées :

1. Quels sont les plants de la placette A1 ?
2. Combien y a-t-il de plants dans la placette A2 ?
3. Combien d'observations ont été prises entre le 7 juin 2017 et le 7 juillet 2017 ?
4. Quels sont les plants en fleur au 7 juillet 2017 ?
5. Quelle est la largeur moyenne des feuilles des plants des placettes ayant une couverture au sol d'au moins 50 % de mousses et qui ont été observées entre le 10 et le 17 juillet 2017 ?
6. Calculer l'histogramme de l'état sur une période donnée (découpée en semaines).
7. Calculer l'histogramme de la surface de la feuille sur une période donnée (découpée en semaines).
8. Vérifier qu'aucune observation n'est antérieure à l'étiquetage du plant.

4. Première approche

4.1. Diagramme relationnel du modèle logique

Une première analyse a conduit à l'élaboration du schéma préliminaire présenté ci-après. Il est présenté sous la forme d'un diagramme relationnel.



Légende

Les rectangles représentent les relations (tables en SQL) et leurs attributs (colonnes en SQL). Les flèches représentent les clés étrangères (orientées de la relation référente vers la relation référée). Les icônes

décorées d'une clé jaune signalent les attributs faisant partie de la clé (candidate) primaire. Les icônes décorées d'une clé bleue signalent les attributs faisant partie d'une clé référentielle (*foreign key*).

Remarque

D'autres clés (candidates ou référentielles) peuvent exister. Il vous appartient de les découvrir.

4.2. Dictionnaire des types de données

Le présent dictionnaire est préliminaire ; il a été établi sur la base d'entretiens exploratoires avec les parties prenantes.

Dictionnaire des types de données

arbre_id

Code identifiant uniquement une variété d'arbres

- Un texte d'au plus 20 caractères

date_eco

Date d'une observation écologique.

- Voir la précision (d).
- La date du début du projet d'observation aurait pu être choisie.

description

Description textuelle consignée par l'observateur ; typiquement, une définition, une annotation ou un commentaire associé à une observation.

- Un texte d'au plus 60 caractères.

dim_mm

Dimension d'une feuille de trille exprimée en millimètre.

- La dimension maximale est d'un mètre.

etat_id

Code identifiant uniquement un état observable d'une plante.

- Une lettre majuscule.

parcelle_id

Code identifiant uniquement une parcelle au sein d'une placette.

- Un entier compris entre 0 et 99.

peuplement_id

Code identifiant uniquement un peuplement végétal de parcelle.

- Un texte composé de quatre lettres majuscules.

placette_id

Un code identifiant uniquement une placette.

- Un texte composé d'une lettre majuscule et d'un chiffre.

plant_id

Un code identifiant uniquement une plante

- Un texte composé du préfixe 'MM' suivi d'une lettre suivie de quatre chiffres ; la lettre doit être l'une des majuscules A, B ou C.

taux_id

Code identifiant uniquement un intervalle de couverture communément appelé « taux ».

- Une lettre majuscule.

taux_val

Valeur correspondant à la proportion d'une couverture à un centième près.

- Deux représentations ont été envisagées :
 1. Un nombre rationnel décimal compris entre 0.0 et 1.0 avec deux chiffres décimaux.
 2. Un entier compris entre 0 et 100.
- Le deuxième choix a été retenu après consultation des parties prenantes.
- Dès lors, une absence de couverture correspond à 0, une couverture totale à 100.

Produit le 2025-02-03 16:26:34 -0500



Collectif francophone pour l'enseignement libre de l'informatique