



Haute École en Hainaut
Av. Victor Maistriau 8 – 7000 Mons

MA1 Ingénieur industriel – Life Data Technologies

Programmation R

Description du projet

La ferme des Heureux Éleveurs d'Herbivores

Vincent Branders

Année académique 2023-2024

Version du 26 octobre 2023

Attention, il s'agit d'une pré-version. Une version finale, contenant des précisions sur les échéances et les modalités d'évaluation, sera communiquée rapidement.

1 Contexte

1.1 Bienvenue !

La ferme des Heureux Éleveurs d'Herbivores est une exploitation laitière et agricole. La gestion d'une telle exploitation ne s'improvise pas et nécessite une certaine organisation. Pour moderniser la gestion de la ferme, Mr. et Mme. VanDeBoer ont fait appel à vos services.

Vous avez à votre disposition une base de données décrivant tous les animaux qui ont un jour vécu à la ferme depuis sa création en 1990. Ces données comprennent la liste des animaux, la liste des naissances, la race de certains animaux, les familles des vaches, etc. Votre mission consistera à créer une application web locale qui présente certaines informations dérivées de ces données et utiles à Mr. et Mme. VanDeBoer. Votre application web sera responsable de l'extraction de données pertinentes de la base de données, du calcul de statistiques de ces données, et de l'affichage du résultat sur une page web via des graphes pertinents.



1.2 La vache, c'est pas de la tarte !

Pour pouvoir comprendre les tenants et aboutissants du projet, il est nécessaire de clarifier certaines informations sur la thématique des données. Ces informations sont décrites dans les paragraphes suivants.

1.2.1 Une vache

Lorsque l'on parle d'une vache, on ne fait pas spécifiquement référence à une femelle de l'espèce *Bos taurus*. Dans certains cas, la distinction mâle (taureau)/femelle (vache) peut avoir de l'importance. Dans ces cas, une distinction claire sera faite entre les individus mâle et femelle.

1.2.2 Les origines

Les vaches de la ferme des Heureux Éleveurs d'Herbivores ont deux origines possibles :

- Soit une vache est née dans la ferme des Heureux Éleveurs d'Herbivores d'une femelle **déjà présente** dans la ferme ;
- Soit une vache provient d'une autre ferme :
 - Dans le cas des femelles, ce sont des animaux achetés à une autre exploitation.
 - Dans le cas des mâles, il s'agit plutôt de paillette d'insémination (les taureaux ne sont pas amenés directement en ferme pour la reproduction, on procède à de l'insémination artificielle).

1.2.3 Le jeu des 7 familles

Pour identifier chaque animal, les vaches ont une carte d'identité. Sur cette carte se trouve un numéro unique d'identification, repris sur l'étiquette d'oreille de l'animal, la date de naissance, l'identification de la mère, ainsi que d'autres informations.

Dans la ferme, en plus de l'identification officielle, chaque vache appartient à une famille. Chez les humains, l'appartenance à une famille est indiquée à l'aide d'un nom, souvent appelé **le nom de famille**. Chez les vaches, l'appartenance à une famille est définie par la mère : une vache appartient toujours à la même famille que sa génitrice (ou vulgairement, sa mère).

Dans le cas de vaches achetées dans une autre exploitation, chaque nouvelle vache arrivée dans la ferme des Heureux Éleveurs d'Herbivores devient la **matriarche** d'une nouvelle famille. La famille sera constituée de la vache achetée, et de tous ses futurs veaux.

1.2.4 Les races de vaches

Il existe une multitude de races différentes de vaches :

- Holstein,
- Montbéliarde,
- Jersey,
- Blonde d'aquitaine,
- ...

À la ferme des Heureux Éleveurs d'Herbivores, il y a 3 races de vaches représentées :

- Holstein,
- Jersey,
- Blanc Bleu Belge.



Un comparatif de vaches créée par *Bing Image Generator*

1.2.5 Le mélange de sang

Dans sa chanson *Quand Les Sangs*, Maurane disait que lorsque les sangs se mélangent, tous les hommes s'arrangent et nous viennent les anges. Dans le cas des vaches, lorsqu'une vache de la race X et un taureau de la race Y donnent naissance à un veau, on considère simplement que le veau est à 50% de race X et à 50% de race Y. Les animaux achetés (taureaux ou

matriarches) seront *considérés* comme étant des animaux de pure race. En pratique, ce n'est pas nécessairement le cas, mais cela simplifie les statistiques.

1.2.6 Le vêlage

Lorsqu'une vache donne naissance à un ou plusieurs veaux, on dit qu'elle **vèle**. Comme vous l'aurez compris, chaque vêlage peut donner un ou plusieurs veaux.



La nature, pour belle qu'elle puisse paraître, peut également paraître cruelle : dans certains cas, un, plusieurs, ou tous les veaux meurent. Si la mort survient durant le vêlage, on parle de veau(x) **mort(s) né(s)**. Si la mort survient plus tard, par exemple dans les quelques semaines qui suivent la naissance, on parle de **décès** prématuré.

Indépendamment de la survie ou de la mort d'un veau, des complications, parfois multiples, peuvent survenir pendant le vêlage. On compte par exemple les complications suivantes :

- Aide vétérinaire au vêlage
- Mère avec post complication
- Veau avec post complication
- Veau mal formé
- Veau mal placé
- Veau trop gros à la naissance
- Veau trop petit à la naissance
- Vêlage difficile
- Vêlage plus tôt que prévu

2 Schéma de la base de données

Les données sont disponibles sur Moodle : sql-data.zip.

Cette section décrit le schéma de la base de données que vous allez utiliser pour le projet. Vous devrez extraire les informations de ce texte pour créer le schéma de la base de données. Dans ce schéma, vous ne devez pas vous contenter des tables, mais vous devez également prendre en compte toutes les contraintes possibles, c'est-à-dire que chaque clé primaire et chaque clé étrangère doivent être précisées.

2.1 Description de la base de données

Dans la ferme laitière et agricole, la ferme des Heureux Éleveurs d'Herbivores, nous avons des **animaux**, et plus précisément des vaches. Chaque vache doit être identifiable par un identifiant (*id*). D'autres informations complèteront cette table, comme l'identifiant de sa famille (*famille_id*), son sexe (*sexe*), si elle est présente ou non dans la ferme (*presence*), et si elle est apprivoisée (*apprivoise*). Enfin pour avoir des données plus complètes, nous enregistrerons également des informations sur leur mort, c'est-à-dire si l'animal est mort-né

(*mort_ne*) ou décédé prématurément (*decede*). (Ces 2 dernières informations devront être deux attributs différents.)

Le nom de famille de chaque vache sera enregistré dans une table **familles**, où pour chaque ligne, nous aurons l'identifiant du nom de famille (*id*) et son nom (*nom*).

Afin de connaître les races présentes dans la ferme, nous retiendrons l'ensemble des races présentes dans celle-ci dans une table **types**. Pour chaque ligne, nous aurons un identifiant unique (*id*) et une race donnée (*type*). Pour rappel, un exemple de race présente dans la ferme est “Blanc Bleu Belge”. Il nous faudra aussi retenir le ou les race(s) de chaque vache. Pour cela, nous créerons une troisième table **animaux_types**. Nous identifierons de manière unique chacune de ces lignes via une clé composée, c'est-à-dire l'identifiant de l'animal (*animal_id*) et l'identifiant de la race (*type_id*), et ajouterons le pourcentage (*pourcentage*) de cette race pour cette vache.

Dans une telle ferme, les naissances sont nombreuses. Nous les répertorierons dans une table **velages**. Chaque vêlage doit être identifié par un identifiant (*id*), et nous stockerons les informations suivantes : l'identifiant de la mère (*mere_id*), l'identifiant du père (*pere_id*), et la date de naissance (*date*). Afin de pouvoir retrouver à quel animal chaque vêlage correspond, nous associerons chaque animal étant né dans la ferme à son vêlage grâce à leurs identifiants respectifs (*animal_id*, *velage_id*) dans une table **animaux_velages**.

Malheureusement, il peut y avoir des complications lors du vêlage d'une vache. Nous enregistrerons chaque complication dans une table **complications**, où nous conserverons un identifiant de la complication (*id*) et la description de celle-ci (*complication*). Pour rappel, une complication peut être de type “Veau trop petit à la naissance” ou encore “Aide vétérinaire au vêlage”. Afin de pouvoir faire des statistiques sur ces complications lors des vêlages, nous conserverons celles-ci dans une table **velages_complications**, dans laquelle nous associerons quel vêlage a eu quelle complication grâce à leurs identifiants respectifs (*velage_id*, *complication_id*).

3 Instructions

Pour ce projet, il vous est demandé :

- a) D'insérer les données décrites dans la Section 2 dans votre base de données, et de calculer, pour chaque animal, son héritage génétique. L'**héritage génétique** de chaque animal sera ajouté comme nouvelle entrée de la table **animaux_types** dans la base de données.
- b) Ensuite, vous devez créer une page web, en utilisant le langage *R* et plus particulièrement le package *Shiny*, qui affiche une figure illustrant des quantités extraites de la base de données, à l'aide d'attributs modifiables directement dans l'interface web.
 - (a) Votre page web sera responsable de l'extraction de données pertinentes de la base de données, du calcul de quantités pertinentes à l'exploitation des Heureux Éleveurs d'Herbivores, et de l'affichage du ou des résultats sous formes de graphiques et de tables.
 - (b) Vous implémenterez une interface qui permettra à la ferme des Heureux Éleveurs d'Herbivores de sélectionner un ensemble d'attributs de son choix. L'interface utilisateur n'est pas contrainte.
 - (c) La page web doit proposer au moins deux graphiques, présentés dans des onglets (*tabs*) différents, et réalisés avec le package *ggplot2*, parmi les propositions ci-dessous. Ces graphes doivent être dynamiquement redessinés selon divers paramètres que vous rendrez modifiables pour les utilisateurs. Lorsque cela semble

pertinent, les données peuvent en plus être présentées sous forme de tableaux, dans des onglets séparés.

(d) Enfin vous serez libre d'ajouter de nouvelles améliorations de votre choix.

3.1 Propositions de graphiques à afficher sur la page web

- a) Afficher le nombre de vêlages par jour sur une période. Donner l'option à l'utilisateur d'affiner sa recherche en ajoutant un champ famille qui est optionnel.
- b) Une croyance populaire chez certains éleveurs est que les naissances arrivent plus souvent à la pleine lune.
 - Est-ce vrai ou pas ? Produire une *heatmap* ou *bar chart* sur 30 jours (ou par phase lunaire) avec le pourcentage de vêlage.
- c) Une vache fait généralement son premier veau à l'âge de 2 ans. Avant ce premier veau, elle n'est pas économiquement utile, puisqu'elle ne produit pas encore de lait. Une exploitation cherche généralement à garder ses vaches le plus longtemps possible, pour maximiser le nombre de lactations (une lactation correspond à un vêlage).
 - Pour chaque année, montrer graphiquement comment tous les vêlages de l'année se répartissent entre premier vêlage, second vêlage, etc.
- d) Il arrive malheureusement qu'un veau soit mort-né. Pour en diminuer le nombre, il est utile de savoir quelles peuvent en être certaines causes.
 - Y a-t-il un moment particulier durant l'année lors duquel plus de mort-nés arrivent ?
 - Y a-t-il une famille particulière chez qui cela arrive plus souvent ? (Afficher, famille par famille, le nombre/la proportion de vivants et de mort-nés, en choisissant une unité de temps adéquate.)
 - Y a-t-il un père pour lequel cela arrive souvent ? (Afficher, par père, le nombre/la proportion de vivants et de mort-nés, en choisissant une unité de temps adéquate.)
- e) Il arrive qu'un veau décède dans ses premières semaines de vie. On parle alors de décès prématuré. Pour en diminuer le nombre, nous voulons savoir quelles peuvent en être les causes.
 - Y a-t-il un moment particulier durant l'année lors duquel plus de décès prématurés arrivent ?
 - Y a-t-il une famille particulière chez qui cela arrive plus souvent ? (Afficher, famille par famille, le nombre/la proportion de vivants et de décès prématurés, en choisissant une unité de temps adéquate. Les décès prématurés peuvent être divisés en catégories pour en refléter la cause.)
 - Y a-t-il un père pour lequel cela arrive souvent ? (Afficher, par père, le nombre/la proportion de vivants et de décès prématurés, en choisissant une unité de temps adéquate.)
- f) Les vendeurs de doses pour l'insémination artificielle proposent des doses sexées (promettant 90% de veaux femelles). Comme elles sont également plus chères, il faut bien choisir les animaux sur lesquelles les mettre pour s'assurer que la dose n'est pas perdue. On va préférer les vaches provenant de familles qui font moins de femelle de base, et pour lesquelles il y a peu de jumeaux (il est moins probable que la gestation aille jusqu'au

bout dans des cas de jumeaux) et le moins de morts à la naissance ou prématurées, afin de maximiser l'utilité de la dose.

- Quelles familles sont les meilleures suivant ces critères-là ?
- g) Occasionnellement, un veau mâle est vendu à une autre ferme. Pour pouvoir le vendre, nous devons fournir des informations sur sa généalogie, ainsi que sur sa *qualité*.
- Quelles sont les généalogies des taureaux de la ferme ? (afficher l'arbre généalogique jusqu'à la matriarche pour les taureaux présents dans la ferme.)
 - Pour chacun, afficher si la mère, la grand-mère, les surs, tantes, cousines, nièces sont encore présentes dans la ferme.
 - Pour chacun, vient-il d'une bonne famille ou une famille avec problèmes ? (afficher au sein d'une famille les statistiques de femelles, jumeaux, morts à la naissance ou prématurées, ou autres problèmes)
- h) Le vendeur de dose d'insémination propose de racheter des doses d'un taureau ou d'un enfant d'un taureau que nous avons déjà eu. Nous pouvons utiliser les données des veaux déjà nés pour évaluer le taureau au sein de notre ferme.
- Quel fut son utilisation ? (graphe par mois du nombre de veau né de ce taureau, veau de 2ème génération, de 3ème génération)
 - Quelles sont ses statistiques ? (% femelle, % jumeaux, % mort-nés/décédés, % problèmes, ...)
 - Combien de descendantes avons-nous encore (filles, petites filles, arrière petites filles, *stack bar chart* par taureau) ?

Vous veillerez, bien entendu, à écrire du code *R* clair et documenté.

Chaque groupe (de 2 *étudiants* réalisera son travail en utilisant la plateforme GitHub de façon à garder une trace des contributions des différents membres du groupe. Pensez à organiser votre GitHub proprement et utilisez des messages de commit clairs (effectuez des commits régulièrement). Votre dépôt doit être privé. Vous ajouterez **vbranders** comme contributeur.