

DEEP LEARNING INDABA CAMEROUN

TABLEAU DE BORD DE LA CAMPAGNE DE DON DE SANG



DEEP
LEARNING
INDABAX
CAMEROON



DEEP
LEARNING
INDABA

Optimisation des dons de sang au Cameroun, cas de la ville de Douala :

Réalisation d'un dashboard intelligent

ANABA OHANDZA Telesphore Rodrigue

KEGNE Bienvenu Landry Jean

NOULAYE LAMBU Merveille

TAGNE TCHINDA Rine Valdy

Douala 2025

Table des matières

Introduction	3
Contexte et justification	3
Méthodologie de travail	4
Préprocessing	4
Visualisation	5
Modélisation	6
Résultats	6
Prise en main de l'application	9
Présentation de l'interface	9
Visualisation des indicateurs	10
Produire un rapport	11
Formulaire	11
Nouveau tableau de bord	12
Fonctionnalités clés de l'application	12
Cartographie de la Répartition des Donneurs	12
Conditions de Santé et Éligibilité	13
Profilage des donneurs idéaux	14
Analyse de l'Efficacité des Campagnes	14
Modèle de prédition d'éligibilité au don de sang	15
Conclusion	16
Code QR et lien vers l'application	17
Références bibliographiques	18

Introduction

Contexte et justification

La transfusion sanguine joue un rôle essentiel dans la prestation de soins de santé d'importance vitale. Les personnes vulnérables telles que les mères pendant l'accouchement, les enfants sous-alimentés et touchés par le paludisme, les victimes de traumatismes et d'accidents, et les patients souffrant de drépanocytose et d'autres maladies chroniques bénéficient particulièrement de ces soins, (Message de la Directrice régionale de l'OMS pour l'Afrique, Dre Matshidiso Moeti). En Afrique, où les besoins en sang sont particulièrement critiques, les dons volontaires jouent un rôle vital dans la prise en charge des urgences médicales, des maladies chroniques, et des complications liées à la maternité. Selon l'Organisation mondiale de la santé (OMS), environ 5 millions de décès par an pourraient être évités en Afrique grâce à des transfusions sanguines sûres et suffisantes (OMS, 2021). Cependant, le continent reste confronté à des défis majeurs en matière de collecte, de gestion, et de distribution du sang.

En effet, il ressort du rapport de l'OMS de 2021 sur le don du sang en Afrique subsaharienne que la sous-région collecte seulement 5 à 10 dons de sang pour 1 000 habitants, ce qui est bien en dessous du seuil recommandé par l'OMS de 10 à 20 dons pour 1 000 habitants (OMS, 2021). Par ailleurs ; plus de 70 % des besoins en sang ne sont pas satisfaits, en raison d'un manque de donneurs réguliers et d'infrastructures adéquates (Croix-Rouge africaine, 2022). Pour le cas des femmes enceintes l'UNICEF montre que les complications liées à la grossesse et à l'accouchement représentent une part importante des besoins en sang. Afrique subsaharienne où 66 % des décès maternels mondiaux surviennent, l'accès à des transfusions sanguines rapides pourrait sauver des milliers de vies chaque année (UNICEF, 2020).

En 2023, les besoins annuels en produits sanguins au Cameroun étaient estimés à plus de 400 000 poches de sang, tandis que les 473 formations sanitaires suivies par le Centre National de Transfusion Sanguine (CNTS) n'en ont collecté que 147 034, couvrant ainsi 36,75 % des besoins. Cette couverture, bien qu'insuffisante, s'améliore progressivement au fil des années : de 103 359 poches collectées en 2019 (soit un taux de couverture de 25,84 %), elle est passée à 99 850 en 2020 (24,96 %), puis à 140 207 en 2021 (35,05 %), illustrant une tendance à la hausse malgré des fluctuations. (MINSANTE).

Selon les projections démographiques sur les populations cibles des différents programmes mis en

œuvre par le MINSANTE, l’Institut National de la Statistique du Cameroun (INS Cameroun), estime la population de Douala, capitale économique du Cameroun à 3 816 533 d’habitants (MINSANTE 2016). Ce qui fait de la capitale économique du Cameroun une ville à forte densité de population. Cependant, comme dans de nombreuses villes africaines, les besoins en sang dépassent largement les disponibilités. Selon les données de l’Établissement Français du Sang (EFS) et des hôpitaux locaux, seulement 30 % des besoins en sang sont couverts à Douala, laissant des milliers de patients sans accès à des transfusions vitales (EFS, 2023). La collecte est confrontée aux défis comme le manque de sensibilisation, les problèmes de sécurité et la rétention des donneurs notamment causé par le manque d’un système de suivi et de reconnaissance des donneurs réguliers.

Face à ce constat alarmant, avoir une vue globale sur les potentiels donneurs de sang et leurs conditions de santé devient une préoccupation majeure pour les autorités de santé publique afin de réduire cette situation. Ainsi, cette étude s’appuie sur des données collectées lors d’une campagne de don de sang en 2019 pour proposer un outil de suivi optimisé.

L’objectif principal de ce travail est de faire usage de ces données afin de proposer un outil de suivi des campagnes de même envergure. De façon spécifique, il sera question de :

- ✓ Visualiser la répartition géographique des donneurs pour identifier les zones les plus représentées et celles qui sont sous-représentées ;
- ✓ Analyser l’impact des conditions de santé sur l’éligibilité et adapter les campagnes de sensibilisation ;
- ✓ Identifier les profils de donneurs idéaux grâce à des techniques de clustering ;
- ✓ Améliorer la rétention des donneurs et des candidats au don notamment en gardant une trace de leurs facteurs démographiques et leurs situation géographique.

Afin de mener à bout ces objectifs, la section suivante présente notre méthodologie de travail.

Méthodologie de travail

Préprocessing

Au cours de cette étape, nous avons tout d’abord procédé à la correction des orthographies de l’ensemble des noms de quartiers présents dans la base de données. Ensuite, nous avons attribué à chacun de ces quartiers les coordonnées géographiques d’un point spécifique, généralement situé en son centre.

En ce qui concerne les religions et les professions, nous nous sommes appuyés sur la Nomenclature Camerounaise des Fonctions (NCF) établie par l’Institut National de la Statistique (INS) du Cameroun, ainsi que sur la classification des religions utilisée par cet institut lors de la cinquième Enquête Camerounaise auprès des Ménages (ECAM5).

Une correction est immédiatement appliquée à la variable « Age », en soustrayant 6 à chaque valeur afin d’ajuster l’indicateur d’âge par rapport à l’année 2019. Cette opération prépare le terrain

pour la suite des manipulations. Nous avons également procédé au regroupement des variables Âge et Classe d’âge.

Par ailleurs, nous avons créé une variable, dénommée *NB_Raison*, permettant de recenser le nombre de motifs de non-éligibilité pour chaque individu figurant dans la base. Cette opération a été effectuée à l'aide d'une macro VBA.

Par la suite, seules certaines colonnes jugées pertinentes pour l'analyse sont sélectionnées. Ainsi, le traitement se focalise sur des variables telles que l'éligibilité au don, le niveau d'étude, l'âge, le genre, la situation matrimoniale, la profession, la religion, l'historique de don et le taux d'hémoglobine. Cette sélection permet de réduire la complexité de la base et de concentrer l'analyse sur les indicateurs essentiels.

Ensuite, plusieurs étapes de nettoyage et de préparation des données sont entreprises. Tout d'abord, les enregistrements dépourvus de valeurs pour le taux d'hémoglobine sont éliminés, garantissant ainsi l'intégrité des analyses futures. Les individus dont l'âge correspond à des valeurs jugées non conformes aux normes réglementaires (12, 17 et 119 ans) sont ensuite retirés de la base, ce qui contribue à l'exactitude des données démographiques. Dans le même temps, les âges manquants sont remplacés par la médiane des âges, assurant ainsi que l'ensemble de la variable « Age » soit complet et cohérent.

L'étape suivante consiste à transformer la variable cible « ÉLIGIBILITÉ AU DON. ». Les modalités textuelles sont converties en valeurs numériques, avec la valeur 1 attribuée aux individus éligibles et la valeur 0 aux cas temporairement ou définitivement non éligibles. Ce recodage est fondamental pour préparer la variable en vue d'une utilisation dans des modèles de classification.

Une sélection spécifique de variables pour la modélisation est ensuite effectuée, en regroupant l'éligibilité, l'âge, le taux d'hémoglobine, le genre, la profession et l'historique de don. Pour assurer la robustesse des modèles, une correction des valeurs aberrantes est réalisée sur les variables quantitatives, à l'exception de la variable cible. Cette opération repose sur le calcul de l'intervalle interquartile, et les valeurs excédant ou se situant en deçà des bornes définies sont remplacées par ces limites respectives.

Visualisation

Après les traitements sur-cités sur la base des données collecter en 2019, un ensemble des graphiques interactifs ont été produit avec les Bibliothèques python Plotly et Folium. Notre outil de visualisation est une application conçue avec la bibliothèque streamlit de Python. Notre choix s'est porté particulièrement sur ces outils du fait des fonctionnalités qu'ils offrent, de leur portabilité et de leur intégration à d'autres interfaces. Une autre raison du choix est le dynamisme dont leur combinaison ensemble nous permet d'avoir dans l'application.

Les graphiques conçus ressortent de prime à bord les caractéristiques sociodémographiques des individus enregistrés en 2019, leur repère géographique et le statut d'éligibilité de chacun.

Modélisation

Nous élaborons un modèle de classification fondé sur l'algorithme des k plus proches voisins, communément désigné sous l'appellation K-Nearest Neighbors (KNN). Son principal objectif est de déterminer l'éligibilité d'un individu au don de sang en fonction de diverses caractéristiques qui lui sont propres.

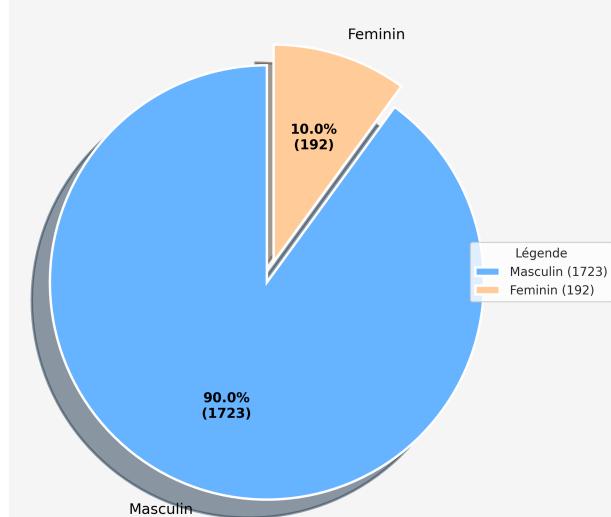
Toutefois, une fonction complémentaire a été conçue en amont du modèle afin de classer les individus potentiellement éligibles, sous réserve des résultats de certains examens et tests, tels que les TDR et l'évaluation du taux d'hémoglobine dans le sang.

Résultats

Nous présentons ici les principaux résultats obtenus à partir des données de l'étude. La campagne de don de sang organisée en 2019 s'est tenue pendant x jours (du x au y). 1915 individus ont pris part à cette campagne.

En ce qui concerne la répartition par sexe, le graphique ci-dessous met en évidence une répartition inégale de la population selon le sexe. La majorité des individus sont de sexe masculin, représentant 90 % de l'effectif total, soit 1 723 personnes, tandis que les femmes ne constituent que 10 %, avec 192 individus.

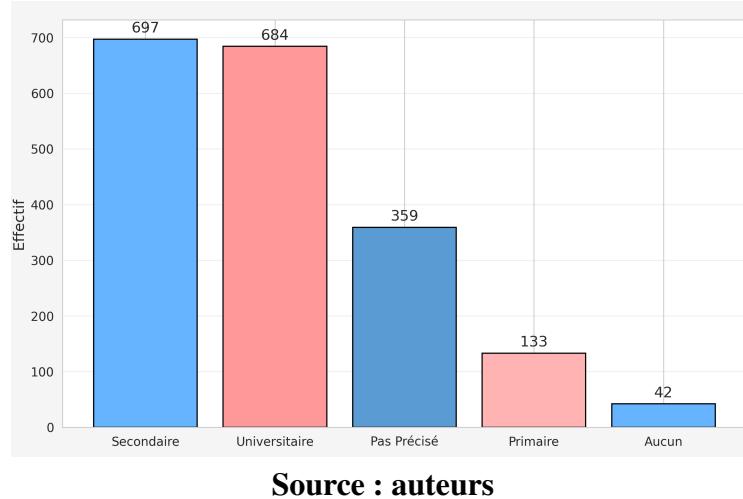
Graphique 1 – Répartition de la population étudiée selon le sexe



Source : auteurs

En ce qui concerne la répartition de la population selon le niveau d'étude, les catégories les plus représentées sont celles des personnes ayant un niveau secondaire (697 individus) et universitaire (684 individus), ce qui indique une forte proportion de personnes ayant poursuivi des études après le primaire. En revanche, une part significative de la population (359 individus) appartient à la catégorie Pas précisé, ce qui peut traduire un manque d'information ou une classification incomplète. Les niveaux primaire et aucun sont nettement moins fréquents, avec respectivement 133 et 42 individus.

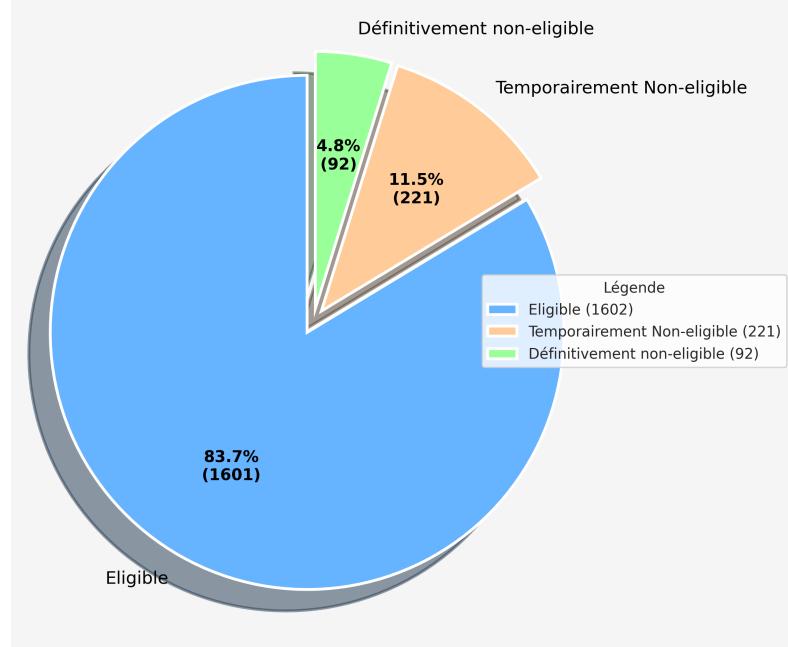
Graphique 2 – Répartition de la population étudiée selon le niveau d'étude



Source : auteurs

S’agissant du statut d’éligibilité au don de sang, une large majorité des individus, soit 83,7 % (1 601 personnes), est éligible au don de sang. Ce groupe constitue la principale part de la population étudiée. En revanche, 11,5 % (221 personnes) sont temporairement non éligibles, ce qui signifie que des facteurs momentanés, tels que des conditions médicales réversibles ou des tests en attente, les empêchent de donner leur sang pour le moment. Enfin, une petite fraction de la population, 4,8 % (92 personnes), est définitivement non éligible, probablement en raison de conditions médicales permanentes ou de critères stricts d’exclusion. Globalement, cette répartition montre que la majorité des personnes analysées peuvent donner leur sang, avec une minorité rencontrant des restrictions temporaires ou définitives.

Graphique 3 – Répartition de la population étudiée selon le statut d’éligibilité

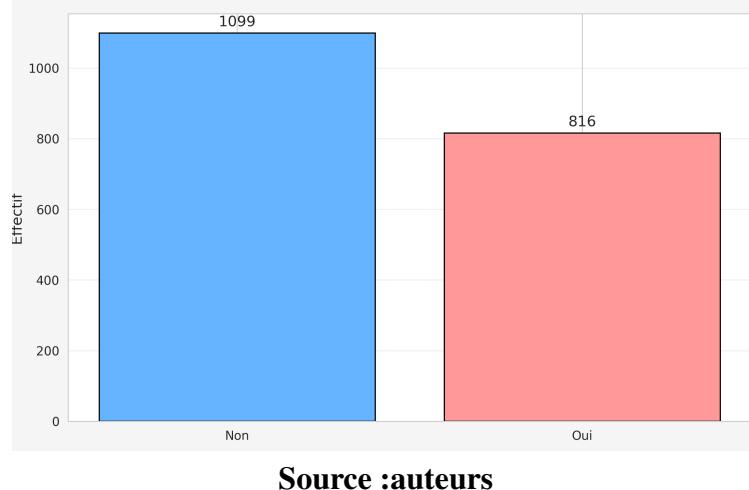


Source :auteurs

L’analyse de la répartition des individus selon leur historique de don de sang révèle une distribution significative. Sur l’ensemble de la population étudiée, 816 personnes ont un historique de don de

sang, tandis que 1099 individus n'ont jamais effectué de don. Cette distribution met en évidence que la majorité des personnes (57,4%) n'ont pas encore participé à un don de sang, contre 42,6% qui ont déjà fait un don. Ce résultat souligne potentiellement un besoin de sensibilisation et de promotion du don de sang au sein de cette population, car moins de la moitié des individus ont contribué à ce geste citoyen et solidaire qui peut sauver des vies.

Graphique 4 – Répartition de la population étudiée selon le statut d'éligibilité



Prise en main de l'application

Dans un souci de proposer une solution de visualisation de données intuitive, dynamique et accessible, nous avons développé une application de suivi des dons de sang avec Streamlit. Ce framework Python permet de transformer rapidement des scripts d'analyse en applications web interactives, sans nécessiter de compétences approfondies en développement frontend. Le tableau de bord que nous avons conçu offre quatre fonctionnalités clés qui permettent une exploration compréhensive des données :

- il offre une visualisation exhaustive et immédiate de tous les indicateurs pertinents relatifs aux dons de sang, permettant une compréhension rapide et globale des données ;
- il propose un outil de personnalisation graphique puissant permettant aux utilisateurs de générer leurs propres représentations croisées, accompagnées d'analyses statistiques avancées telles que des tests d'indépendance et de comparaison de moyennes ;
- L'application intègre un formulaire interactif d'éligibilité au don, qui détermine en temps réel le statut du donneur potentiel : éligible, temporairement éligible ou non-éligible ;
- Chaque formulaire rempli alimente automatiquement une base de données dynamique, créant ainsi un mécanisme d'enrichissement continu des informations. Cette approche permet de constituer progressivement un référentiel complet, où les données géographiques et personnelles sont automatiquement standardisées et intégrées, simplifiant radicalement les futures campagnes de collecte.

Chaque fonctionnalité a été pensée pour répondre aux besoins spécifiques des professionnels de santé et des chercheurs travaillant sur les problématiques de don de sang.

Présentation de l'interface

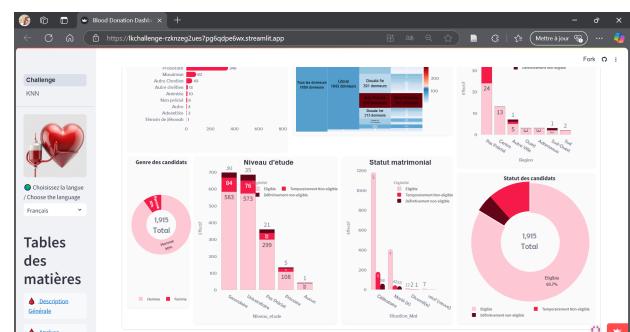
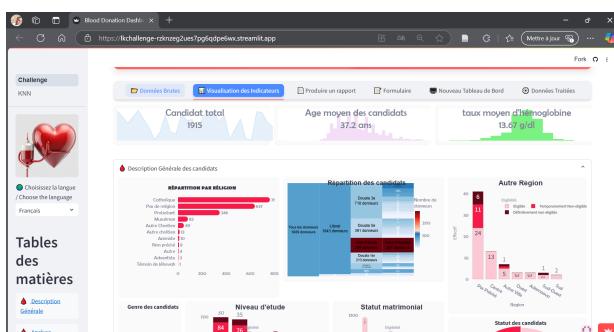
À l'entrée de notre tableau de bord, nous avons la page suivante. A gauche l'utilisateur a la possibilité de choisir la langue avec laquelle il veut interagir avec le tableau de bord. Nous proposons ainsi à l'utilisateur de switcher du français à l'anglais et vice versa. Sur le bord gauche également, l'utilisateur a une table des matières lui permettant de facilement se repérer et d'accéder aux différents onglets du tableau de bord.

Dans l'onglet **Données Brutes** l'utilisateur a un visuel sur trois différentes bases de données : la base de données traitées, les données géo spatialisées (qui reprend la première en ajoutant les coordonnées géographiques des quartiers), et les données sur les donneurs effectifs.

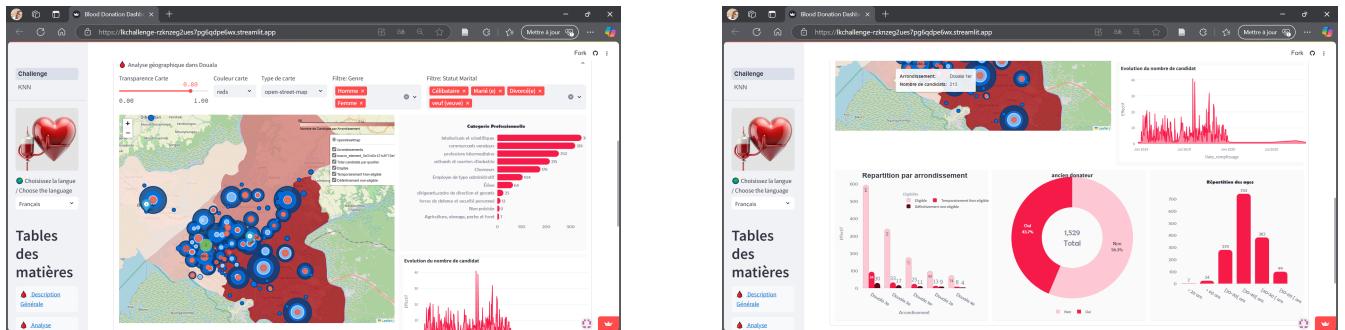
Visualisation des indicateurs

Cet onglet regroupe plusieurs sections telles que présentées par les figures suivantes. D'entrée de jeu, nous pouvons avoir quelques statistiques générales notamment le nombre total de participant à la campagne, l'âge moyen des participants et le taux d'hémoglobine moyen des participants.

Il est important de noter que les sections suivantes sont déroulantes et regorgent différents graphiques selon le contenu. Par exemple, lorsqu'on clique sur la section description générale des candidats, nous avons le visuel suivant :



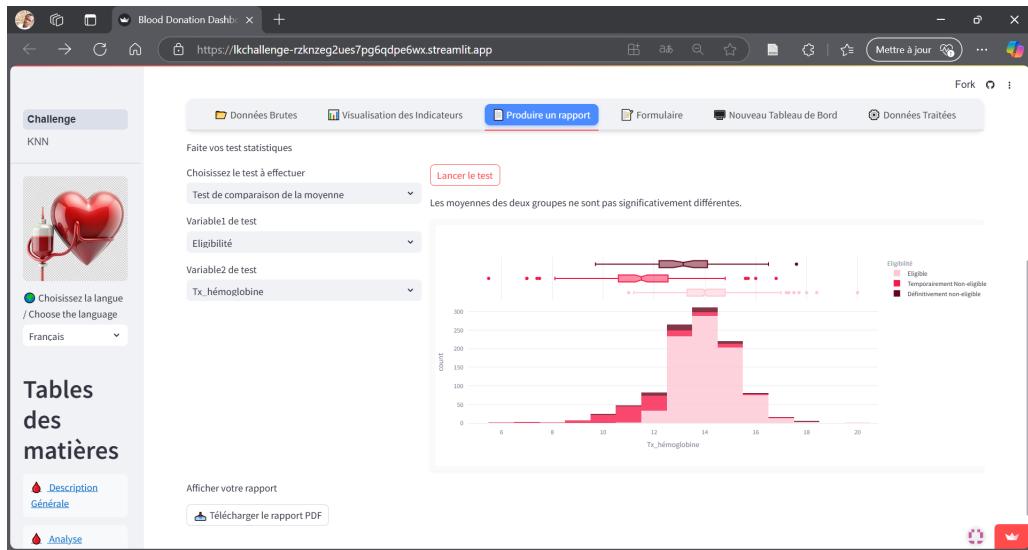
De même, en cliquant sur la section **analyse géographique dans douala**, nous avons le visuel suivant



Notons également que dans cette section comme dans plusieurs autres, l'utilisateur a la possibilité d'avoir la répartition géographique des participants à la campagne selon les arrondissements de Douala. Il a la possibilité de filtrer selon le sexe, le genre, de choisir le type de cartes à afficher et la possibilité de personnaliser la carte. Les graphiques sont dynamiques : à chaque choix de l'utilisateur, les graphiques s'ajustent automatiquement.

Produire un rapport

Dans cet onglet, nous proposons à l'utilisateur d'effectuer certains tests statistiques et leurs interprétations. Dans cet onglet, l'utilisateur a la possibilité de télécharger le rapport contenant du projet tout entier. Cet onglet se présente comme suit



Formulaire

Cet onglet présente un formulaire qui une fois rempli, nous donne indique si le candidat remplit les conditions préalables pour passer les tests de sang approfondis. Les informations renseignées sont

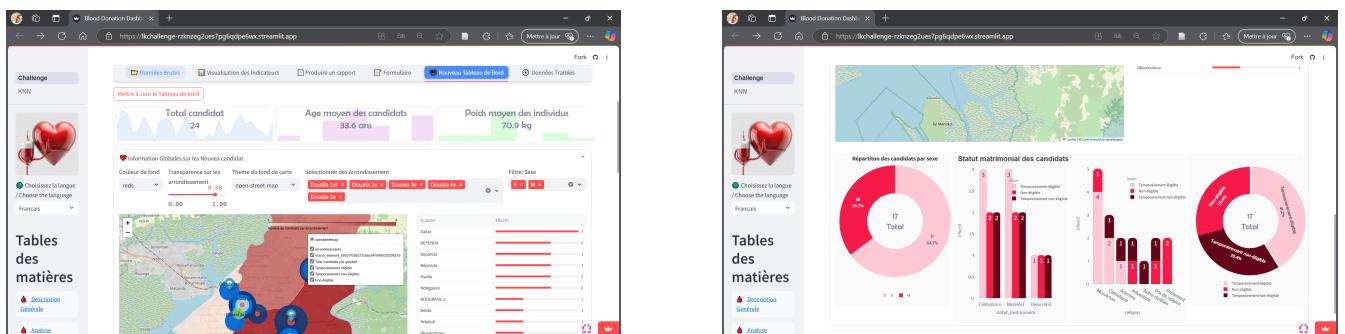
directement stockées dans une base de données pour des utilisations ultérieures. Visuellement nous avons :

The left screenshot shows the 'Challenge' section with a heart icon and a 'Tables des matières' sidebar. It includes fields for 'Informations générales' (Sexe: Féminin, Poids: 60kg, Dernière donation: Plus de 2 mois), 'Informations socio-démographiques' (Niveau d'étude: Primaire, Profession: Chaudronnier, Quartier de résidence: Dakar, Nationalité: Cameroun), and 'Statut matrimonial' (Célibataire). The right screenshot shows the 'Challenge' section with a heart icon and a 'Tables des matières' sidebar. It includes fields for 'Informations générales' (Sexe: Féminin, Poids: 60kg, Dernière donation: Plus de 2 mois), 'Informations socio-démographiques' (Niveau d'étude: Universitaire 1er cycle, Profession: Chaudronnier, Quartier de résidence: Dakar, Nationalité: Cameroun), and 'Statut matrimonial' (Célibataire). Both screenshots show a 'Tables des matières' sidebar with 'Description Générale' and 'Analyse' sections.

Les données reçues ici alimentent un nouveau tableau de bord pour la préparation des futures campagnes. Ainsi, plus besoin de collecter les coordonnées, plus besoin de chercher les vrais noms des quartiers ou arrondissement, ils sont automatiquement calculés à partir des informations entrées dans formulaire et celles de la base de initiale.

Nouveau tableau de bord

Ce second tableau de bord est produit à partir des données collectées dans la section précédente. Ce tableau permet à l'utilisateur d'avoir une vue globale sur les individus qui se sont présentées pour donner du sang. Ici, on peut déjà identifier les individus temporairement non éligibles et les contacter le moment opportun.

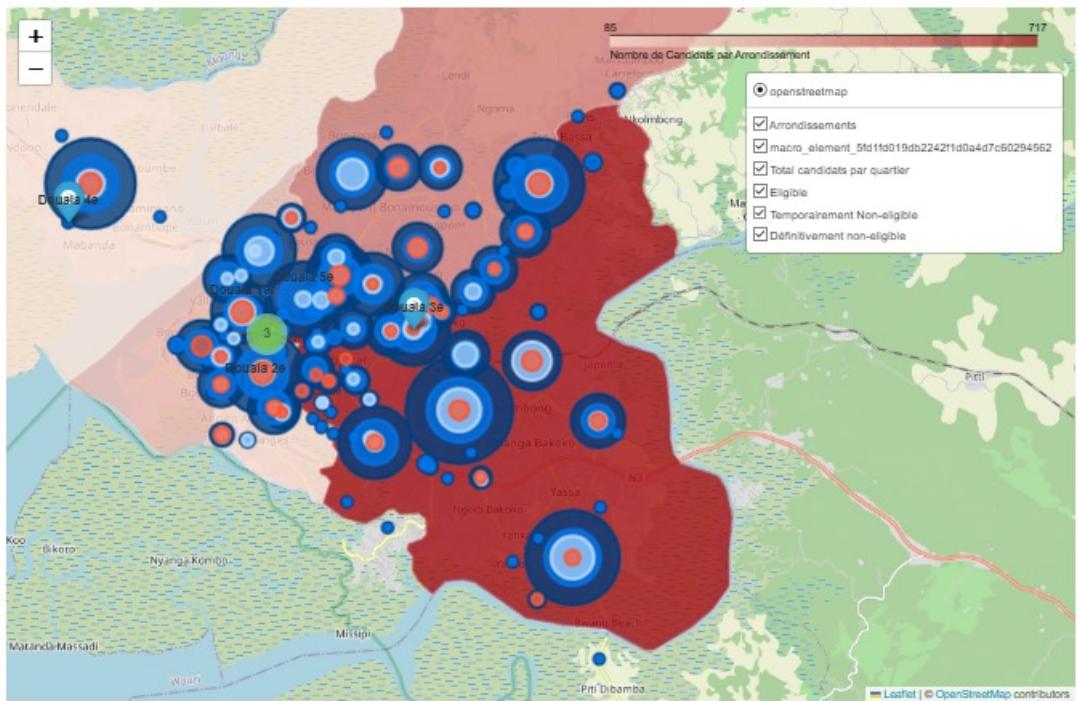


Fonctionnalités clés de l'application

L'application développée permet de répondre aux différents points suivants

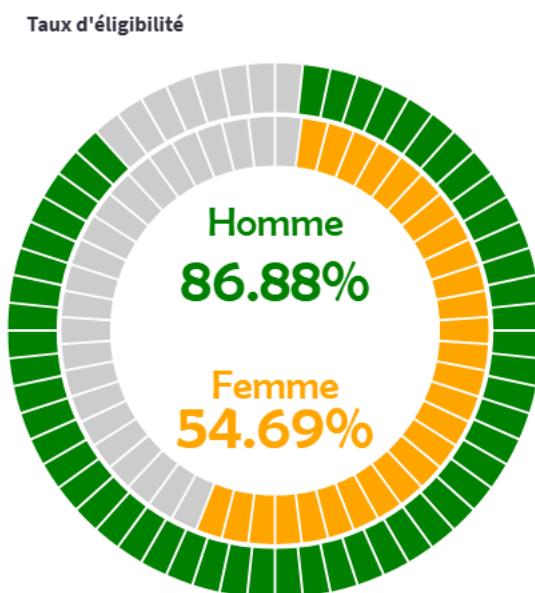
Cartographie de la Répartition des Donneurs

Cette carte interactive visualise la répartition géographique des candidats au don de sang par arrondissement, avec une segmentation des niveaux d'éligibilité représentée par des cercles propor-

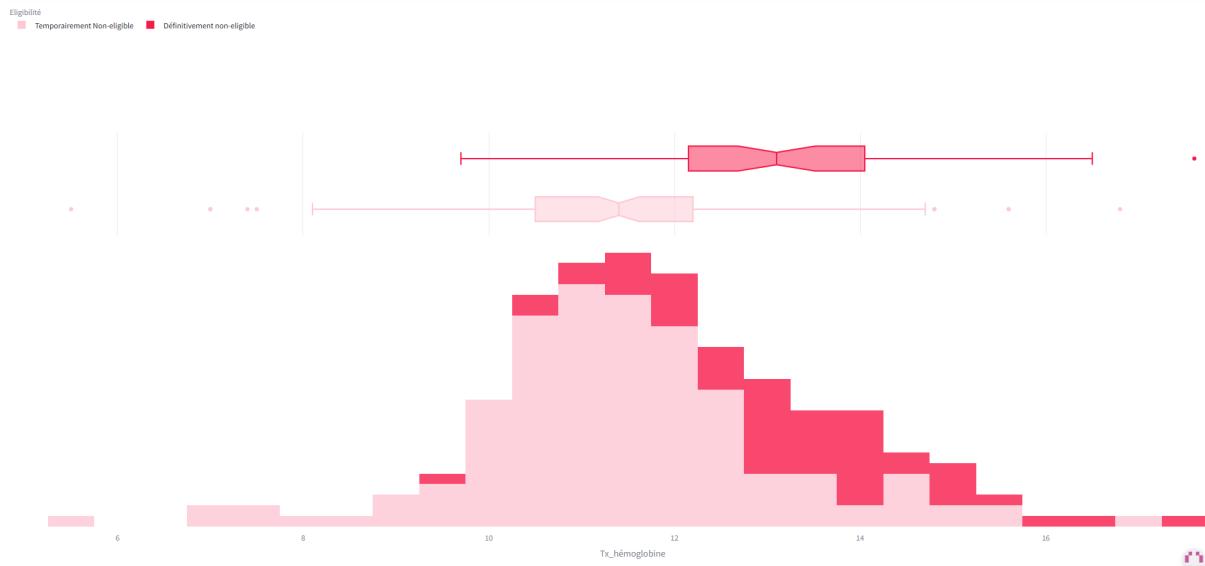


tionnels, offrant ainsi une lecture immédiate des disparités spatiales et des zones de forte ou faible admissibilité.

Conditions de Santé et Éligibilité

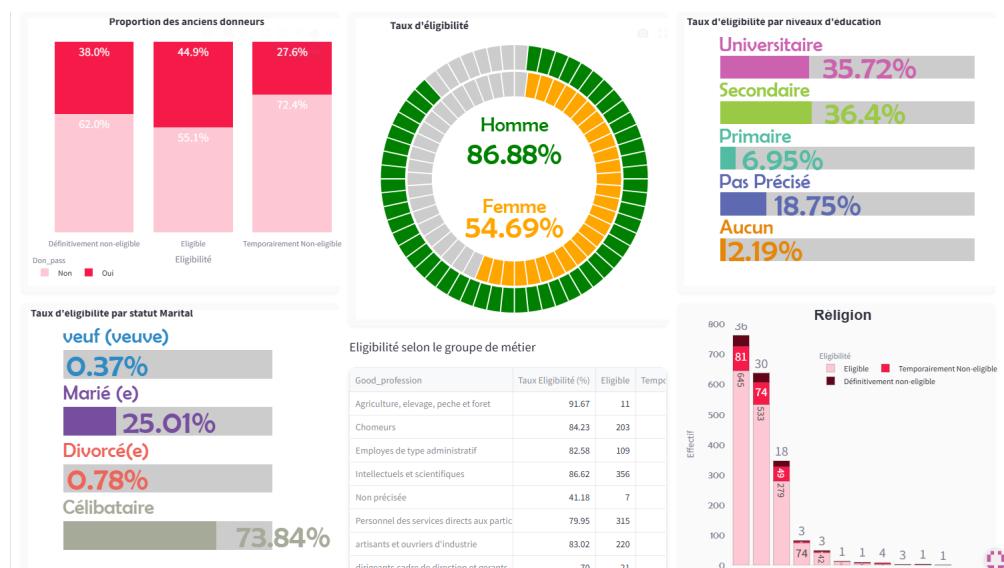


Le graphique ci-dessus montre que L'éligibilité au don de sang est significativement plus élevée chez les hommes (86,88 %) que chez les femmes (54,69 %), ce qui pourrait refléter des différences physiologiques ou des critères de sélection plus stricts pour ces dernières.



L'analyse du graphique met en évidence une différence dans la distribution du taux d'hémoglobine entre les donneurs éligibles et non éligibles, avec une médiane plus élevée chez ces derniers. Cela suggère que, dans certains cas, l'inéligibilité au don de sang pourrait être associée à des taux d'hémoglobine plus élevés, probablement en raison de conditions médicales sous-jacentes.

Profilage des donneurs idéaux



L'analyse des graphiques ci-dessus indique que les donneurs idéaux sont principalement des personnes célibataires, ayant un niveau d'éducation supérieur au secondaire et, pour la plupart, n'ayant jamais donné de sang auparavant.

Analyse de l'Efficacité des Campagnes

Avec un taux de conversion de 99,22 % (1 901 dons/1 915 inscriptions), la campagne 2019 démontre une efficacité opérationnelle remarquable, minimisant les pertes entre inscription et don ef-

fectif.

Analyse de l'Efficacité des Campagnes

Avec un score exceptionnel de 99,77% sur l'échantillon d'entraînement et 92,52% en validation, ce modèle démontre une capacité prédictive remarquable pour l'éligibilité au don de sang, combinant une précision cliniquement fiable (proche de l'excellence sur les données connues) et une robustesse généralisable (performance stable sur de nouvelles données). Son déploiement dans les centres de transfusion pourrait optimiser le pré-tri des donneurs en réduisant les tests inutiles (gain de temps/moyens) et en ciblant les candidats à haut potentiel d'éligibilité, tout en maintenant un taux de faux négatifs minimal – un atout crucial pour des systèmes de santé soucieux d'efficience et de sécurité transfusionnelle.

Conclusion

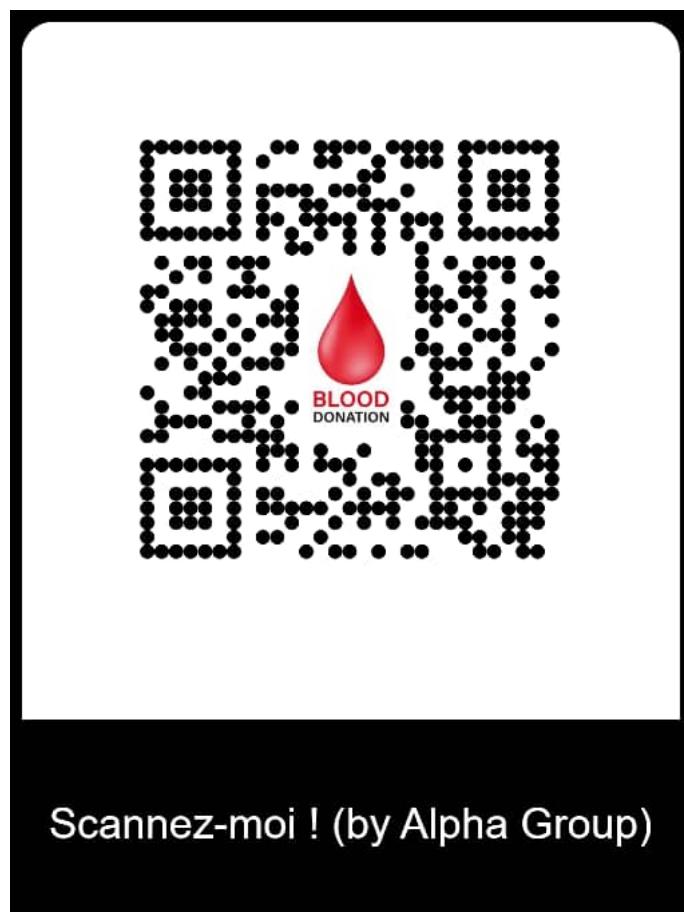
La présente étude a permis de mettre en lumière les enjeux cruciaux liés au don de sang dans le contexte camerounais, en particulier dans la ville de Douala. Les analyses réalisées à partir des données de la campagne de 2019 montrent clairement que, malgré une majorité de donneurs potentiellement éligibles, la répartition inégale des participants et la faible couverture de certains segments de la population traduisent un besoin urgent de renforcement des actions de sensibilisation et de suivi. L'outil de visualisation et le modèle de classification développés offrent une approche innovante pour mieux comprendre les profils des donneurs et orienter les stratégies d'intervention de manière ciblée et efficace.

Il est recommandé de renforcer les efforts de communication et de sensibilisation auprès des populations sous-représentées afin d'élargir le vivier des donneurs. Une meilleure coordination entre les établissements de santé et les organismes de collecte pourrait permettre d'optimiser les ressources et d'améliorer la réactivité des campagnes. Par ailleurs, l'intégration de l'outil numérique interactif dans le processus de suivi constitue une opportunité de centraliser les données et de dynamiser la gestion des campagnes. La mise en place de programmes d'engagement durable, adaptés aux spécificités locales, favoriserait la fidélisation des donneurs et contribuerait à une amélioration progressive du taux de couverture, garantissant ainsi une meilleure prise en charge des urgences médicales et une réponse efficace aux besoins en sang.

Code QR et lien vers l'application

En scannant ce code QR ou en cliquant sur le lien, vous serez redirigé vers le site de l'application

► Cliquez ici



Références bibliographiques

- ◆ Organisation mondiale de la santé (OMS). (2021). Blood safety and availability.
- ◆ Établissement français du sang (EFS). (2023). Chiffres clés du don de sang.
- ◆ Croix-Rouge française. (2022). L'importance du don de sang.
- ◆ Nomenclature camerounaise des métiers (INS Cameroun 2013)
- ◆ <https://www.minsante.cm/site/?q=en/node/4612#:~:text=Au%20Cameroun%2C%20les%20besoins%20annuels,besoins%20de%2036%2C75%25>
- ◆ <https://datacameroon.com/don-de-sang-le-cameroun-en-pénurie-permanente/>