

Amélioration des cartographies épidémiques avec des annotations

GRAPH Network & WHO, supported by the Global Fund to fight HIV, TB & Malaria

October 2023

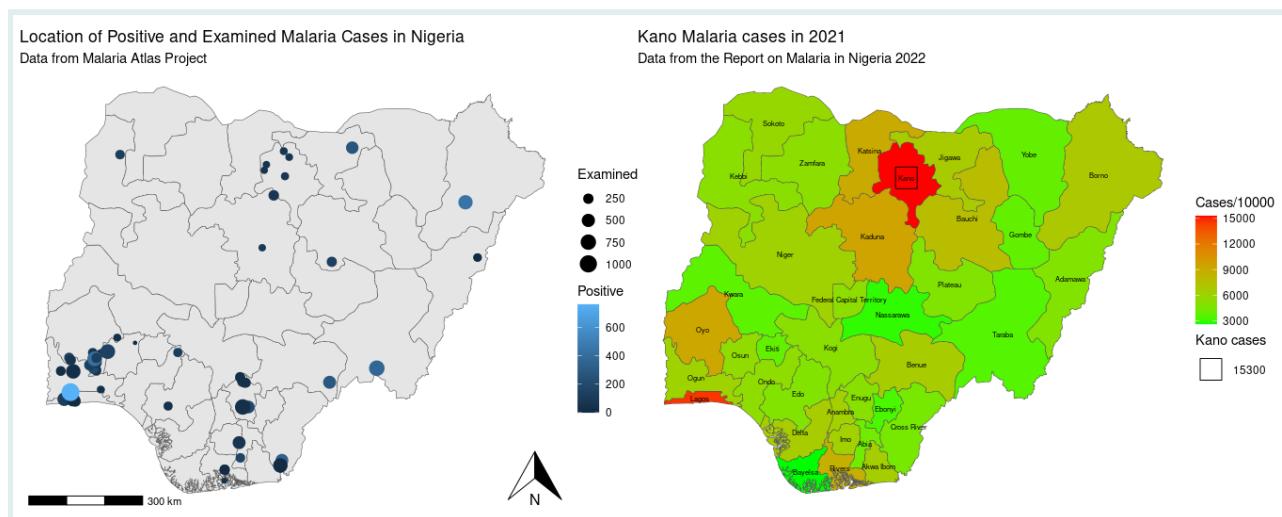
This document is a draft of a lesson made by the GRAPH Network, a non-profit headquartered at the University of Geneva Global Health Institute, in collaboration with the World Health Organization, under a Global Fund 2023 grant to create e-learning modules to build in-country data capacity for epidemiological and impact analysis for National HIV, TB and malaria programs

Introduction
Packages
Préparation des Données
Création d'une Carte Choroplète Simple
Ajout d'Indicateurs de Données Continues à la Carte Choroplète
Exploration des Taux d'Augmentation des Cas de Paludisme à l'aide de Cartes Choroplèthes
Étiquetage de la Carte Choroplète avec les Noms des États
Affichage des Noms des États Combinés et des Taux d'Augmentation sur la Carte Choroplète
Mettre en Évidence une Région Spécifique sur la Carte tout en Préservant le Contexte
Étiquetage des Emplacements des Points : Exploration du Taux de Positivité au Paludisme et de l'Incidence
Dernières Réflexions
Références
Récapitulatif
Réponses

Introduction

Dans la visualisation des données géospatiales, les cartes sont des outils puissants pour raconter des histoires. Cependant, une carte sans annotations claires et des étiquettes équivaut à un livre sans titres ni chapitres. Bien que l'histoire puisse toujours être là, il devient considérablement plus difficile de comprendre, d'interpréter et d'apprécier.

Dans cette leçon, nous mettons particulièrement l'accent sur l'importance de l'annotation et de l'étiquetage des cartes. Une annotation appropriée transforme une visualisation simple en un guide informatif, permettant aux spectateurs de saisir rapidement des données spatiales complexes. Grâce à un étiquetage précis, les zones d'intérêt peuvent être immédiatement reconnues, facilitant une compréhension plus claire de la narration des données.



Objectifs d'apprentissage

Objectifs d'apprentissage : Techniques Avancées de Visualisation Géospatiale

À la fin de cette section, vous devriez être capable de :

1. Incorporer des indicateurs de données continues dans les cartes choroplèthes pour une granularité accrue.
2. Calculer et visualiser les taux d'augmentation des cas de paludisme à l'aide de cartes choroplèthes.
3. Superposer efficacement les noms des États sur les cartes choroplèthes, en veillant à la clarté et à la lisibilité.
4. Intégrer sans heurts les noms des États avec les taux d'augmentation sur les cartes sans compromettre la lisibilité.
5. Appliquer des techniques pour mettre en évidence des régions spécifiques sur une carte tout en conservant le contexte global.
6. Déterminer des stratégies optimales de placement des points et les intégrer efficacement dans les visualisations géospatiales.

En maîtrisant ces compétences, vous disposerez des connaissances et outils nécessaires pour créer des visualisations géospatiales riches, détaillées et informatives.

Packages

```
# Charger les bibliothèques
if(!require(pacman)) install.packages("pacman")
pacman::p_load(malariaAtlas,
  ggplot2,
  geodata,
  dplyr,
  here,
  readr,
  sf,
  patchwork)

# Désactiver la notation scientifique
options(scipen=100000)
```

Préparation des Données

Avant de plonger dans toute visualisation ou analyse, il est essentiel de charger et de prétraiter nos données. Cela comprend la lecture des datasets, la fusion des informations connexes et la filtration des entrées inutiles ou non pertinentes.

```
# Lecture des données géographiques au format shapefile
nga_adm1 <- sf::st_read(here::here("data/raw/NGA_adm_shapefile/NGA_adm1.shp"))
%>%
sf::st_simplify(dTolerance = 1000) # simplifier le shapefile pour accélérer
la visualisation.
```

```
## Reading layer `NGA_adm1` from data source
##
`/Users/kendavidn/Dropbox/tgc_github_projects/epi_reports_staging/data/raw/NGA_adm_shapefile'
##   using driver `ESRI Shapefile'
## Simple feature collection with 38 features and 9 fields
## Geometry type: MULTIPOLYGON
## Dimension:      XY
## Bounding box:  xmin: 2.668431 ymin: 4.270418 xmax: 14.67642 ymax: 13.89201
## Geodetic CRS:  WGS 84
```

Les données géographiques au format shapefile des limites administratives du Nigeria (comme les États ou les provinces) sont lues et stockées dans l'objet `nga_adm1`.

```
# Lecture des données attributives relatives aux cas de paludisme
malaria_cases <- read_csv(here::here("data/malaria.csv"))
```

Cette étape charge les données relatives aux cas de paludisme dans différents États du Nigeria.

```
# Filtrage des entrées 'Water body' des données géographiques
nga_adm1 <- filter(nga_adm1, NAME_1 != "Water body")

# Fusion des données géographiques avec les données sur le paludisme
malaria <- malaria_cases %>%
  left_join(nga_adm1, by = c("state_name" = "NAME_1")) %>%
  st_as_sf()
```

Ici, nous combinons les données sur les limites géographiques avec les données sur les cas de paludisme en utilisant les noms des États comme référence. Ces données fusionnées sont ensuite converties dans un format adapté aux visualisations géospatiales.

```
# Filtrage pour ne conserver que les colonnes essentielles pour notre analyse
malaria2 <- malaria %>%
  select(state_name, cases_2000, cases_2006, cases_2010, cases_2015,
cases_2021, geometry)
```

Nous réduisons notre dataset à des colonnes spécifiques, principalement les noms des États, les cas de paludisme de différentes années et les limites géographiques de ces États (c'est-à-dire `geometry`).

```
# Lecture des données de population pour différentes régions du Nigeria
population_nigeria <- read_csv(here::here("data/population_nigeria.csv"))
```

Cette étape charge les données indiquant la population des différents États ou régions du Nigeria.

```
# Combinaison des données de population avec nos données sur le paludisme
malaria3 <- malaria2 %>%
  left_join(population_nigeria, by = c("state_name")) %>%
  st_as_sf()
```

En fusionnant les données de population avec nos données sur les cas de paludisme, nous enrichissons notre dataset. Ces données combinées permettent des visualisations et des analyses plus complètes, telles que le calcul des taux d'incidence ou l'évaluation des tendances par rapport à la taille de la population.

Création d'une Carte Choroplète Simple

REMINDER



Les cartes choroplèthes sont des outils de visualisation puissants qui affichent des zones géographiques divisées ombrées ou hachurées en proportion de la valeur d'une variable.

Dans cet exemple, nous allons utiliser une carte choroplète pour visualiser la répartition des cas de paludisme dans différentes régions du Nigeria pour l'année 2021.

```
# Construction de la carte choroplète à l'aide de ggplot2
ggplot(data=malaria3) +
  geom_sf(aes(fill=cases_2021/1000)) + # La couleur de remplissage est
  # déterminée par le nombre de cas de paludisme en 2021, mis à l'échelle
  # par 1000
  labs(title = "Répartition des Cas de Paludisme au Nigeria en 2021", fill =
    "Cas pour 1000") + # Ajout de libellés et du titre à la carte
  scale_fill_continuous(low = "green", high = "red") + # Utilisation d'une
  # échelle de couleur continue allant du vert au rouge
  theme_void() # Utilisation d'un thème minimal pour une meilleure
  # visualisation de la carte
```

```
## Error in `calc_limits_bbox()`:
## ! Scale limits cannot be mapped onto spatial coordinates in `coord_sf()`
## i Consider setting `lims_method = "geometry_bbox"` or `default_crs =
NULL`.
```

Voici une explication détaillée du code :

- `ggplot(data=malaria3)`: Initialise un objet ggplot en utilisant le dataset malaria3.

- `geom_sf(aes(fill=cases_2021/1000))`: Ajoute les données géographiques de `malaria3` et remplit chaque région en fonction du nombre de cas de paludisme en 2021, mis à l'échelle par un facteur de 1000. Cela représente efficacement le nombre de cas pour mille habitants.
- `labs(title = "Répartition des Cas de Paludisme au Nigeria en 2021", fill = "Cas pour 1000")`: Spécifie le titre de la carte et l'étiquette de l'échelle de couleur.
- `scale_fill_continuous(low = "green", high = "red")`: Applique une échelle de couleur continue où les régions avec moins de cas sont colorées en vert et les régions avec plus de cas sont colorées en rouge.
- `theme_void()`: Supprime le texte des axes, les marques de graduations et autres éléments non essentiels pour mettre en évidence la carte.

Le graphique résultant offre une vue claire de la répartition des cas de paludisme au Nigeria en 2021, avec l'intensité des couleurs indiquant l'ampleur des cas dans chaque région.

Exercice 1 : Modifiez la carte choroplèthe fournie pour visualiser la répartition des cas de paludisme au Nigeria pour l'année 2015

Instructions

1. Mettez à jour la correspondance des données dans la fonction `geom_sf()` pour refléter les cas de paludisme pour l'année 2015.
2. Ajustez le titre dans la fonction `labs()` pour indiquer que la visualisation concerne l'année 2015.
3. Modifiez le dégradé de couleurs dans `scale_fill_continuous()` pour passer du bleu (cas "faibles") au jaune (cas "elevés").

Ci-dessous, un code de départ :



```
# Construction de la carte choroplèthe pour 2015 à l'aide de
# ggplot2
ggplot(data=malaria3) +
  geom_sf(aes(fill=_____)) + # Remplissez la colonne de
  # données correcte pour 2015
  labs(title = "_____ ", fill = "Cas pour
  1000") + # Mettez à jour le titre de manière appropriée
  scale_fill_continuous(_____) + # Modifiez l'échelle de
  # couleurs
  theme_void()
```

Ajout d'Indicateurs de Données Continues à la Carte Choroplète

Lors de l'analyse de données sur les maladies, il est souvent utile de regarder au-delà des chiffres bruts de cas et de se concentrer sur les taux, en particulier les taux d'incidence. Le taux d'incidence fournit une mesure normalisée qui peut prendre en compte les différences de taille de population entre les régions, rendant les comparaisons plus significatives.

Compréhension de l'Incidence

L'incidence d'une maladie est un pilier de la recherche épidémiologique. Elle quantifie le nombre de nouveaux cas d'une maladie survenant au cours d'une période spécifique, généralement une année, par rapport à une population à risque.

Mathématiquement, il est donné par :

$$\text{Incidence} = \frac{\text{Nombre de nouveaux cas au cours de la période}}{\text{Population à risque au début de la période}}$$

Dans ce contexte, nous examinons l'incidence du paludisme dans différents États du Nigeria pour l'année 2021. Plus précisément, nous calculerons le taux d'incidence en divisant le nombre de nouveaux cas de paludisme en 2021 par la population de chaque État issue du dernier recensement disponible en 2019.

Le code R suivant réalise cela et visualise les données :

```
# Trouver les coordonnées du centre de chaque État pour positionner les étiquettes
centroid_coords <- st_coordinates(st_centroid(malaria2$geometry))

# Visualisation du Taux d'Incidence du Paludisme en 2021 à l'aide d'une Carte Choroplète
ggplot(data = malaria3) +
  # Remplir chaque État avec une couleur représentant le taux d'incidence en 2021.
  geom_sf(aes(fill = round(cases_2021/population_2019, 2))) +
  # Étiqueter chaque État avec son taux d'incidence spécifique.
  geom_text(aes(x = centroid_coords[, 1], y = centroid_coords[, 2], label =
    round(cases_2021/population_2019, 2)), size = 2) +
  # Ajout de titre et de titre de légende.
  labs(title = "Taux d'Incidence du Paludisme au Nigeria en 2021", fill =
    "Incidence") +
  # Utilisation d'une échelle de couleur continue du vert (faible incidence) au rouge (forte incidence).
  scale_fill_continuous(low = "green", high = "red") +
  # Utilisation d'un thème minimaliste pour une visualisation plus claire.
  theme_void()
```

```
## Error in `calc_limits_bbox()`:
## ! Scale limits cannot be mapped onto spatial coordinates in `coord_sf()`
```

```
## i Consider setting `lims_method = "geometry_bbox"` or `default_crs =  
NULL`.
```

Voici une brève explication de la visualisation :

- La fonction `geom_sf()` crée une carte choroplète où la couleur de chaque État représente son taux d'incidence du paludisme en 2021.
- La fonction `geom_text()` étiquète chaque État avec son taux d'incidence spécifique, positionnant chaque étiquette au centre de l'État.
- L'échelle de couleurs, passant du vert au rouge, met en évidence visuellement les régions présentant des taux d'incidence plus élevés.

Cette visualisation offre une compréhension immédiate de la situation du paludisme au Nigeria, révélant les zones à forte incidence qui pourraient nécessiter des interventions de santé publique plus ciblées.

Exploration des Taux d'Augmentation des Cas de Paludisme à l'aide de Cartes Choroplèthes

Comprendre l'évolution du nombre de cas de maladies au fil du temps peut fournir des informations sur l'efficacité des interventions, la progression de la maladie et les zones où des ressources accrues pourraient être nécessaires. Dans ce contexte, nous cherchons à visualiser l'augmentation en pourcentage des cas de paludisme de 2015 à 2021 dans différents États du Nigeria.

Calcul du Taux d'Augmentation

Le taux d'augmentation pour chaque État est calculé comme suit :

$$\text{Taux d'Augmentation} = \left(\frac{\text{Cas en 2021} - \text{Cas en 2015}}{\text{Cas en 2015}} \right) \times 100\%$$

Cette formule fournit la croissance en pourcentage (ou la diminution) des cas de paludisme de 2015 à 2021.

Visualisation des Taux d'Augmentation

Plongeons dans le code qui réalise cette visualisation :

```

# Calculer le taux d'augmentation pour chaque État
malaria3 %>%
  mutate(increase_rate = round(((cases_2021 - cases_2015) / cases_2015) *
    100)) -> malaria3

# Visualisation des taux d'augmentation à l'aide d'une carte choroplète
ggplot(data = malaria3) +
  # Colorer chaque État en fonction de son taux d'augmentation
  geom_sf(aes(fill = increase_rate), color="white", size = 0.2) +
  # Utilisation d'une échelle de couleur continue pour représenter les taux
  # d'augmentation, passant du vert au rouge
  scale_fill_continuous(name="Taux d'augmentation (%)", limits=c(0,70), low =
    "green", high = "red", breaks=c(0, 20, 40, 60))+

  # Étiquetage de chaque État avec son taux d'augmentation en pourcentage
  geom_text(aes(x = centroid_coords[, 1], y = centroid_coords[, 2], label =
    paste0(increase_rate, "%")), size = 2)+

  # Ajout d'un titre à la carte
  labs(title = "Taux d'Augmentation du Paludisme au Nigeria en 2021 par
    rapport à 2015")+
  theme_void()

```

```

## Error in `calc_limits_bbox()`:
## ! Scale limits cannot be mapped onto spatial coordinates in `coord_sf()`
## i Consider setting `lims_method = "geometry_bbox"` or `default_crs =
NULL`.

```

Dans cette visualisation :

- La fonction `geom_sf()` crée la carte choroplète où l'intensité de couleur de chaque État représente son taux d'augmentation du paludisme.
- `scale_fill_continuous()` définit l'échelle de couleur pour les taux d'augmentation, mettant en évidence davantage les zones à forte augmentation.
- `geom_text()` ajoute des étiquettes à chaque État, fournissant des valeurs exactes en pourcentage.
- La carte résultante nous permet d'identifier rapidement les régions avec une forte croissance des cas de paludisme sur la période sélectionnée.

Grâce à cette visualisation, les parties prenantes peuvent cibler les régions où le paludisme est en augmentation et potentiellement allouer plus efficacement des ressources.

Étiquetage de la Carte Choroplète avec les Noms des États

Dans une carte choroplète, bien que les dégradés de couleurs offrent un indice visuel pour comprendre la répartition d'une variable dans différentes régions, l'ajout d'étiquettes peut considérablement améliorer la clarté de la visualisation. Cela est particulièrement vrai lorsque le public n'est peut-être pas familier avec toutes les limites géographiques affichées. Dans le cas de notre dataset sur le paludisme, l'ajout des noms des États à la carte rend les données plus accessibles et compréhensibles.

Analysons le code :

```
# Construction d'une carte choroplèthe avec les noms des États
ggplot(data = malaria3) +
  # Remplir chaque État en fonction du nombre de cas de paludisme en 2021,
  # normalisé par 10 000
  geom_sf(aes(fill = cases_2021/10000)) +
  # Ajouter le nom de chaque État à son centre
  geom_text(aes(x = centroid_coords[, 1], y = centroid_coords[, 2], label =
    state_name), size = 2, check_overlap = TRUE) +
  # Ajouter des titres et des étiquettes
  labs(title = "Cas de Paludisme au Nigeria en 2021", fill = "Cas/10 000") +
  # Utiliser une échelle de couleur continue du vert (faible nombre de cas) au
  # rouge (nombre élevé de cas)
  scale_fill_continuous(low = "green", high = "red") +
  theme_void()
```

```
## Error in `calc_limits_bbox()`:
## ! Scale limits cannot be mapped onto spatial coordinates in `coord_sf()`
## i Consider setting `lims_method = "geometry_bbox"` or `default_crs =
NULL`.
```

Voici une explication détaillée de la visualisation :

- `geom_sf(aes(fill = cases_2021/10000))` : Cela crée la carte choroplèthe où la couleur de chaque État représente le nombre de cas de paludisme en 2021, normalisé par un facteur de 10 000.
- `geom_text(aes(...))` : Cette fonction place du texte sur le graphique. Dans ce cas, elle est utilisée pour ajouter le nom de chaque État à son centre géographique. Le paramètre `check_overlap = TRUE` permet à ggplot de vérifier les chevauchements de libellés de texte et de tenter d'éviter les chevauchements.
- `labs(title = "Cas de Paludisme au Nigeria en 2021", fill = "Cas/10 000")` : Cette fonction ajoute un titre au graphique et un libellé à l'échelle de couleur.
- `scale_fill_continuous(low = "green", high = "red")` : Cela définit l'échelle de couleur de la carte, en transitionnant du vert pour les États avec moins de cas au rouge pour ceux avec plus de cas.

La visualisation résultante est une carte claire et informative des cas de paludisme au Nigeria en 2021, avec chaque État étiqueté pour une référence facile.

Affichage des Noms des États Combinés et des Taux d'Augmentation sur la Carte Choroplèthe

Les visualisations peuvent transmettre une multitude d'informations lorsqu'elles intègrent de multiples points de données de manière intuitive. En associant les noms des États à leurs taux d'augmentation respectifs, nous pouvons fournir une vue plus riche et plus détaillée des données sans submerger le public.

Plongeons dans cette visualisation :

Nous souhaitons présenter une carte choroplèthe montrant les cas de paludisme pour 10 000 habitants dans les États nigérians en 2021, avec des étiquettes combinant les noms des États et leurs taux d'augmentation respectifs depuis 2015.

```
# Combinez les noms des États et leurs taux d'augmentation respectifs en une
# seule étiquette
malaria3$label_text <- paste(malaria3$state_name, malaria3$increase_rate, "%")

# Calculez le centre de chaque État pour servir de points de référence pour
# les étiquettes
centroid_coords <- st_coordinates(st_centroid(malaria3$geometry))

# Visualisez les données
ggplot(data = malaria3) +
  # Créez une carte choroplèthe ombrée en fonction du nombre de cas de
  # paludisme en 2021 pour 10 000 habitants
  geom_sf(aes(fill = cases_2021/10000)) +
  # Ajoutez des étiquettes combinées (nom de l'État et taux d'augmentation) au
  # centre de chaque État
  geom_text(aes(x = centroid_coords[, 1],
                y = centroid_coords[, 2],
                label = label_text), size = 2) +
  # Ajoutez des titres et personnalisez la légende des couleurs
  labs(title = "Cas de Paludisme au Nigeria en 2021", fill = "Cas/10 000") +
  scale_fill_continuous(low = "green", high = "red") +  # Définir la gradation
  # des couleurs
  theme_void()  # Appliquez un thème minimal pour plus de clarté
```

```
## Error in `calc_limits_bbox()`:
## ! Scale limits cannot be mapped onto spatial coordinates in `coord_sf()`
## i Consider setting `lims_method = "geometry_bbox"` or `default_crs =
NULL`.
```

Dans cette visualisation :

- La ligne `malaria3$label_text <- paste(malaria3$state_name, malaria3$increase_rate, "%")` construit nos étiquettes combinées en concaténant le nom de l'État avec son taux d'augmentation, suivi d'un signe de pourcentage.

- `geom_text()` ajoute ces étiquettes combinées au graphique. Il positionne chaque étiquette au centre de l'État correspondant, en veillant à ce que les étiquettes soient centrées et facilement associées à leurs régions respectives.
- `scale_fill_continuous(low = "green", high = "red")` attribue une gradation de couleurs en fonction du nombre de cas de paludisme. Les États avec moins de cas seront colorés en vert, passant au rouge pour les États avec plus de cas.

Cette approche nous permet de communiquer efficacement deux informations essentielles (cas pour 10 000 habitants et taux d'augmentation) dans la même visualisation tout en conservant le nom des États pour une identification aisée.

Mettre en Évidence une Région Spécifique sur la Carte tout en Préservant le Contexte

Parfois, vous souhaitez attirer l'attention sur une zone ou une région particulière de votre carte sans omettre les détails des régions environnantes. En utilisant des éléments graphiques spécifiques, tels que des marqueurs plus grands ou des couleurs distinctes, vous pouvez mettre l'accent sur certaines régions tout en mettant en valeur les données plus larges. Dans cet exemple, nous mettons l'accent sur la région de "Kano" au Nigéria.

```

# Calcul des coordonnées du centroïde pour l'étiquetage
centroid_coords <- st_coordinates(st_centroid(malaria3$geometry))

# Visualisation des cas de paludisme au Nigeria avec une mise en avant de Kano
ggplot(data = malaria3) +
  # Création d'une carte choroplète avec des couleurs basées sur les cas de
  # paludisme en 2021
  geom_sf(aes(fill = cases_2021/1000)) +
  # Définition d'une échelle de couleurs continues du vert au rouge
  scale_fill_continuous(low = "green", high = "red") +
  # Superposition d'un point sur Kano pour le mettre en avant. La taille du
  # point correspond au nombre de cas
  geom_point(data = subset(malaria3, state_name == "Kano"),
             aes(x = st_coordinates(st_centroid(geometry))[1],
                  y = st_coordinates(st_centroid(geometry))[2], size =
round(cases_2021/1000)),
             color = "black", shape = 22, fill = "transparent") +
  # Étiquetage de chaque région avec son nom
  geom_text(aes(x = centroid_coords[, 1], y = centroid_coords[, 2], label =
state_name), size = 2, check_overlap = TRUE) +
  # Personnalisation de l'échelle de taille du point mis en avant
  scale_size_continuous(range = c(1, 12)) +
  # Ajout du titre et des légendes
  labs(title = "Cas de paludisme à Kano en 2021", subtitle = "Données du
Rapport sur le Paludisme au Nigeria 2022", size = "Cas à Kano", fill =
"Cas/10 000",) +
  # Application d'un thème minimal pour plus de clarté
  theme_void()

```

```

## Error in `calc_limits_bbox()`:
## ! Scale limits cannot be mapped onto spatial coordinates in `coord_sf()`
## i Consider setting `lims_method = "geometry_bbox"` or `default_crs =
NULL`.

```

Dans cette visualisation :

- La fonction `geom_point()` est utilisée pour placer un cercle sur la région de Kano. La taille du cercle représente le nombre de cas à Kano en 2021. Le cercle est conçu en mode transparent (`fill = "transparent"`) avec une bordure noire audacieuse (`color = "black"`) pour le mettre en évidence.
- `geom_text()` ajoute les noms de toutes les régions sur la carte. Il place chaque étiquette près du centre géographique de la région respective. Le paramètre `check_overlap = TRUE` empêche les étiquettes de se chevaucher autant que possible, garantissant ainsi la lisibilité.



Tandis que "Kano" est mis en évidence, toutes les autres régions sont également affichées avec leur ombrage de couleur respectif basé sur les cas de paludisme. Cela offre une vue holistique de la situation à travers le Nigéria, permettant aux téléspectateurs de comparer Kano avec les autres régions.

Une telle approche est inestimable lorsque vous souhaitez mettre en lumière des détails spécifiques ou des zones d'intérêt sans négliger le jeu de données plus large, enrichissant ainsi vos présentations de données.

Étiquetage des Emplacements des Points : Exploration du Taux de Positivité au Paludisme et de l'Incidence

La cartographie et la visualisation de données spécifiques sur une carte géographique peuvent fournir des informations cruciales, notamment lors de la manipulation de données épidémiologiques. Plongeons dans le code pour comprendre les processus et la visualisation que nous cherchons à réaliser :

```
# Récupération des Données
# Le package malariaAtlas fournit la fonction `getPR` pour accéder au taux de
# parasite (PR).
# Le PR est un indicateur essentiel de la prévalence du paludisme.

# Récupération des données pour le Nigéria pour les deux espèces de paludisme
nigeria_pr <- malariaAtlas::getPR(ISO = "NGA", species = "both") %>%
  # Filtrer les enregistrements sans valeurs de PR
  filter(!is.na(pr)) %>%
  # Supprimer toutes les lignes avec des valeurs manquantes de longitude ou de
  # latitude
  drop_na(longitude, latitude) %>%
  # Convertir les données en un dataframe spatial pour faciliter la
  # cartographie
  st_as_sf(coords = c("longitude", "latitude"), crs = 4326)
```

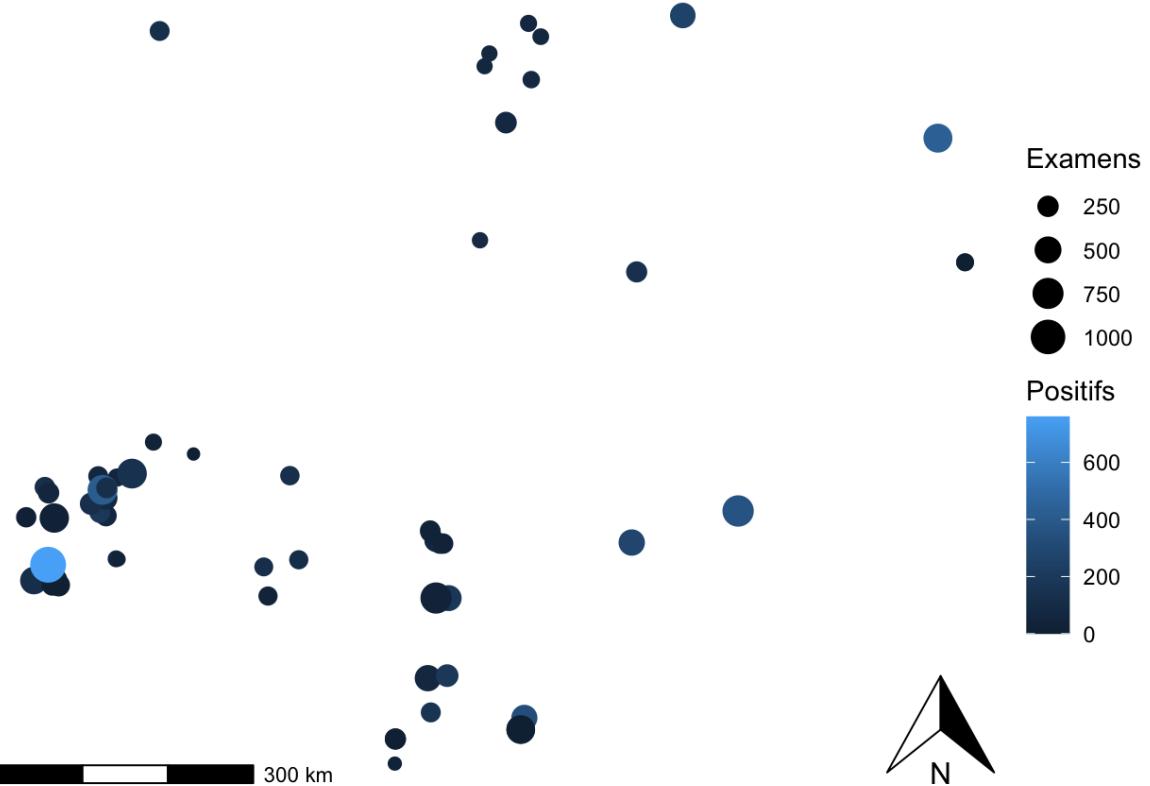
Ce bloc de code récupère les données de prévalence du paludisme pour le Nigéria. Après la récupération, les données sont nettoyées en filtrant les valeurs manquantes. Ensuite, elles sont transformées en un format adapté à la visualisation géospatiale (objet `sf`).

```

# Configuration d'une visualisation géospatiale avec ggplot2
ggplot() +
  # Traçage des limites administratives du Nigeria
  geom_sf(data = nga_adm1) +
  # Ajout de points pour chaque site de test
  # La couleur de chaque point indique si le test est positif, et la taille
  # représente le nombre de personnes testées
  geom_sf(aes(size = examined, color = positive), data = nigeria_pr) +
  # Définition des titres, des étiquettes des axes et des légendes du
  # graphique
  labs(title = "Emplacement des cas de paludisme positifs et examinés au
    Nigeria",
    subtitle = "Données du Projet Atlas du Paludisme",
    color = "Positifs",
    size = "Examens") +
  # Ajout d'étiquettes pour la longitude et la latitude
  xlab("Longitude") +
  ylab("Latitude") +
  # Intégration d'une flèche indiquant le nord pour l'orientation
  ggspatial::annotation_north_arrow(location = "br") +
  # Intégration d'une barre d'échelle pour aider à l'interprétation des
  # distances
  ggspatial::annotation_scale(location = "bl") +
  # Application d'un thème minimaliste pour une clarté visuelle
  theme_void()

```

Emplacement des cas de paludisme positifs et examinés au Nigeria
 Données du Projet Atlas du Paludisme





Ce bloc de code visualise les données sur le paludisme sur une carte. Il met en évidence les régions en fonction du nombre de tests de dépistage du paludisme et de leurs résultats. Des outils spécifiques du package `ggspatial` sont utilisés pour ajouter des éléments cartographiques, garantissant que la carte est informative.

Exercice 3 : Réfléchir sur le taux de positivité par la taille et la couleur



Créez une nouvelle visualisation qui représente le taux de positivité (nombre de cas positifs divisé par le nombre d'examens) pour chaque emplacement. Modifiez la taille des points pour refléter le taux de positivité.

Dernières Réflexions

La visualisation des données géospatiales, c'est bien plus que de représenter des données sur une carte. C'est raconter une histoire ancrée dans un lieu et un espace. Cette leçon s'est plongée dans une approche narrative basée sur les couches, de l'importance d'annotations claires à l'intégration de différents types de données pour une compréhension plus approfondie.

Références

1. Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis (3e)*. Available at: <https://ggplot2-book.org/>
2. Hadley Wickham and Garrett Grolemund. *R for Data Science*. Available at: <https://r4ds.had.co.nz/>
3. Robin Lovelace, Jakub Nowosad, and Jannes Muenchow. *Geocomputation with R*. Available at: <https://r.geocompx.org/>

Récapitulatif

Dans cette leçon sur l'étiquetage des cartes, nous avons exploré en profondeur les points suivants :



RECAP



- **Annotations et Étiquettes** : Des annotations claires transforment les cartes, rendant les données spatiales complexes facilement compréhensibles.
- **Techniques de Cartographie Avancées** : Nous avons exploré comment intégrer des indicateurs continus, superposer des étiquettes, mettre l'accent sur des régions et déterminer les placements optimaux des points sur les cartes.
- **Application Pratique** : En utilisant un dataset sur le paludisme, nous avons démontré la préparation des données, la visualisation et l'interprétation dans R.
- **Packages R** : Des packages tels que `ggplot2`, `sf` et `ggspatial` facilitent les visualisations géospatiales avancées.
- **Mettre en Évidence avec le Contexte** : Il est essentiel de fournir une vue d'ensemble, même lorsque l'on met l'accent sur des zones spécifiques.

Réponses

Solution pour l'exercice 1

Pour visualiser la répartition des cas de paludisme au Nigéria en 2015, suivez les instructions fournies dans l'exercice. Voici la solution complète :

```
# Construction de la carte choroplète pour 2015 avec ggplot2
ggplot(data=malaria3) +
  geom_sf(aes(fill=cases_2015/1000)) + # Colonne de données mise à jour pour
  # refléter l'année 2015
  labs(title = "Répartition des cas de paludisme au Nigéria en 2015", fill =
    "Cas pour 1000") + # Titre mis à jour pour 2015
  scale_fill_continuous(low = "blue", high = "yellow") + # Échelle de couleurs
  # modifiée pour un dégradé de bleu à jaune
  theme_void()
```

Solution pour l'exercice 2

Pour combiner la visualisation des taux d'augmentation avec les noms des états, suivez les étapes ci-dessous :

```

# Combinaison de la visualisation des taux d'augmentation avec les noms des états
ggplot(data = malaria3) +
  # Création d'une carte choroplète avec une couleur de remplissage basée sur le taux d'augmentation
  geom_sf(aes(fill = increase_rate), color="white", size = 0.2) +
  # Étiquetage de chaque région avec son nom d'état
  geom_text(aes(x = centroid_coords[, 1],
                y = centroid_coords[, 2],
                label = state_name), size = 2, check_overlap = TRUE) +
  # Spécification de l'échelle de couleurs pour les taux d'augmentation
  scale_fill_continuous(name="Taux d'augmentation (%)",
                        limits=c(0,70),
                        low = "green",
                        high = "red",
                        breaks=c(0, 20, 40, 60)) +
  # Ajout d'un titre et d'une légende à la carte
  labs(title = "Taux d'augmentation du paludisme au Nigéria de 2015 à 2021",
       fill = "Taux d'augmentation (%)") +
  theme_void()  # Application d'un thème minimaliste pour la clarté

```

```

## Error in `calc_limits_bbox()`:
## ! Scale limits cannot be mapped onto spatial coordinates in `coord_sf()`
## i Consider setting `lims_method = "geometry_bbox"` or `default_crs =
NULL` .

```

Dans cette solution, la carte choroplète est créée en fonction du taux d'augmentation des cas de paludisme de 2015 à 2021. Chaque état du Nigéria est étiqueté par son nom, et le dégradé de couleurs (du vert au rouge) met en évidence l'ampleur du taux d'augmentation. La carte est enrichie d'un titre et d'une légende pour assurer la clarté et la compréhension.

Solution pour l'exercice 3

Créez une visualisation qui représente le taux de positivité :

```

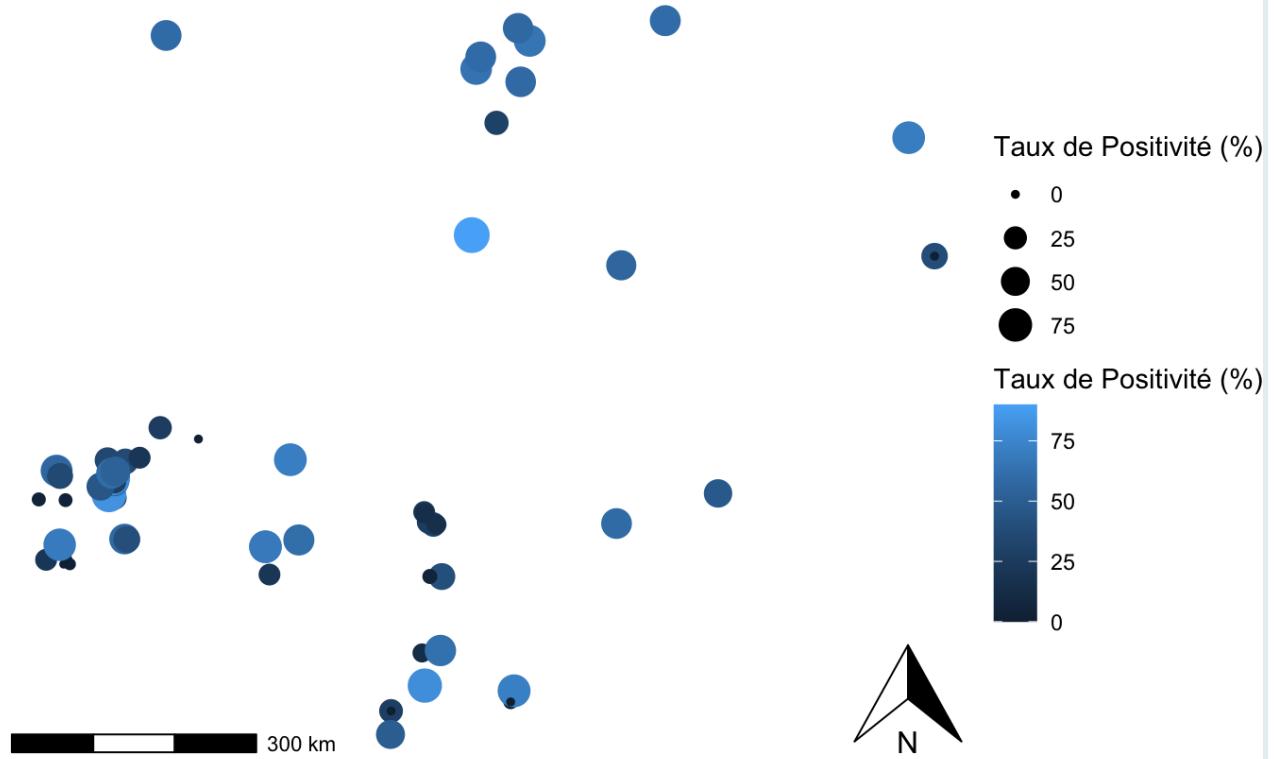
# Ajout d'une colonne de taux de positivité
nigeria_pr$positive_rate <- (nigeria_pr$positive / nigeria_pr$examined) * 100

ggplot() +
  geom_sf(data = nga_adm1) +
  geom_sf(aes(size = positive_rate, color = positive_rate), data =
    nigeria_pr) +
  labs(title = "Emplacement et Taux de Positivité des Cas de Paludisme au Nigéria",
       subtitle = "Données du Projet Atlas du Paludisme",
       color = "Taux de Positivité (%)",
       size = "Taux de Positivité (%)") +
  xlab("Longitude") +
  ylab("Latitude") +
  ggspatial::annotation_north_arrow(location = "br") +
  ggspatial::annotation_scale(location = "bl") +
  theme_void()

```

Emplacement et Taux de Positivité des Cas de Paludisme au Nigéria

Données du Projet Atlas du Paludisme



Contributeurs

Les membres de l'équipe suivants ont contribué à cette leçon :



IMAD EL BADISY

Data Science Education Officer

Deeply interested in health data



JOY VAZ

R Developer and Instructor, the GRAPH Network
Loves doing science and teaching science
