Visualiser des données dans R

Introduction à ggplot2 (1)

Fousseynou Bah

28-Jan-2019

Introduction

Objectif de ce cours

Dans le flux de travail (*workflow*) du *data scientist*, la visualisation est sans aucun doute la phase la plus surprenante et aussi la plus artistique. D'un côté, elle permet au *data scientist* d'explorer ses données, de mieux cerner le problème posé, de poser les bases pour la modélisation et souvent même d'identifier les insuffisances du travail de nettoyage et de mise en forme. Elle peut aussi bien conforter le *data scientist* dans ses hypothèses qu'elle peut battre celles-ci en brêches ou y apporter des nuances. Quelle que soit l'issue, le *data scientist* s'en trouve plus informé, mieux outillé pour poursuivre son analyse.

De l'autre coté, elle présente une dimension artistique. Elle offre au *data scientist* l'occasion d'adresser à l'utilisateur du produit final - et ce travers un mélange de formes et de couleurs - un message, qui se veut souvent simple et précis. Elle invite ainsi la créativité artistique dans un domaine fort bien guidé par la logique.

Dans ce cours, nous allons explorer le célèbre package ggplot2.

La visualisation, un outil de narration

Dans ce cours, nous allons commencer par présenter la logique (et même la philosopie) de *ggplot2*, et ensuite nous allons démontrer sa richesse et sa versatilité à travers des illustrations. Nous allons partir de données (*data sets*), nous poser quelques questions et voir dans quelle mesure la visualisation peut nous aider à trouver des éléments de réponses. Nous allons voir comment est-ce que la nature et le volume des données en main peuvent eux-mêmes suggérer le type de visualisation qui sied.

Les exemples nous offriront aussi l'occasion de relever certaines récurrences dans la syntaxe de *ggplot2* et, par là, de mieux comprendre sa logique.

Que nous faut-il?

- R (évidemment) et RStudio (de préférence) installés sur le poste de travail;
- le package *ggplot2* installé;
- les fichiers fournis dans le cadre du module.

Données

Dans ce chapitre nous allons utiliser les données suivantes:

- Une compilation de données tirées des Recensements Généraux de la Population et de l'Habitat au Mali menés en (RGPH, 1976, 1987, 1998, 2009) ainsi que des projections faites par OCHA;
- Des séries tirées des Indicateurs Mondiaux du Développement du site Internet de la Banque Mondiale.

La grammaire des graphiques

Une deconstruction verbale des graphiques

La grammaire des graphiques est une approche particulière de la représentation visuelle des données. Elle est venue en prominence avec l'ouvrage de Lelan Wilkinson et a trouvé sa déclinaison dans R avec la package ggplot2 de Hadley Wickham.

Elle établit les grands principes de la représentationne visuelle des données à travers un cadre dans lequel sont décrites de façon concise toutes les composantes d'un graphique.

Les composantes du graphique

1. Les données
2. Les esthétiques
3. L'échelle
4. Les objets géométriques
5. Les statistiques
6. Les facettes
7. Le système de coordonnées.

Explorons ces éléments avec des illustrations.

Points: à la recherche d'une relation

Aperçu

Ici, nous allons examiner la relation entre deux variables: le PIB/habitant et l'espérance de vie. Nous allons utiliser deux séries - le PIB/habitant (dollars US constant de 2011) et l'espérance de vie - allant de 1990 à 2016 et couvrant les pays de l'UEMOA - Bénin, Burkina-Faso, Côte d'Ivoire, Guinée-Bissau, Mali, Niger, Sénégal, Togo. Voici la question que nous allons explorer: y a-t-il un lien entre la PIB/habitant et l'espérance de vie dans les pays de l'UEMOA ? Les donnnées ont été tirées du site Internet de la Banque Mondiale.

Voici un aperçu de notre *dataset* :

```
pibhbt_espvie <- read_csv("data/pibhbt_espvie.csv") # Importation du ficher CSV
glimpse(pibhbt_espvie) # Créer un aperçu de la structure des données
```

```
## Observations: 216
## Variables: 5
## $ iso2c         <chr> "BF", "BF", "BF", "BF", "BF", "BF", "BF", "BF", ...
## $ pays          <chr> "Burkina Faso", "Burkina Faso", "Burkina Faso", ...
## $ annee         <dbl> 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, ...
## $ pib_hbt       <dbl> 844.3422, 896.5985, 874.8064, 880.8915, 868.4635...
## $ esperance_vie <dbl> 49.454, 49.409, 49.374, 49.360, 49.380, 49.445, ...
```

Illustration

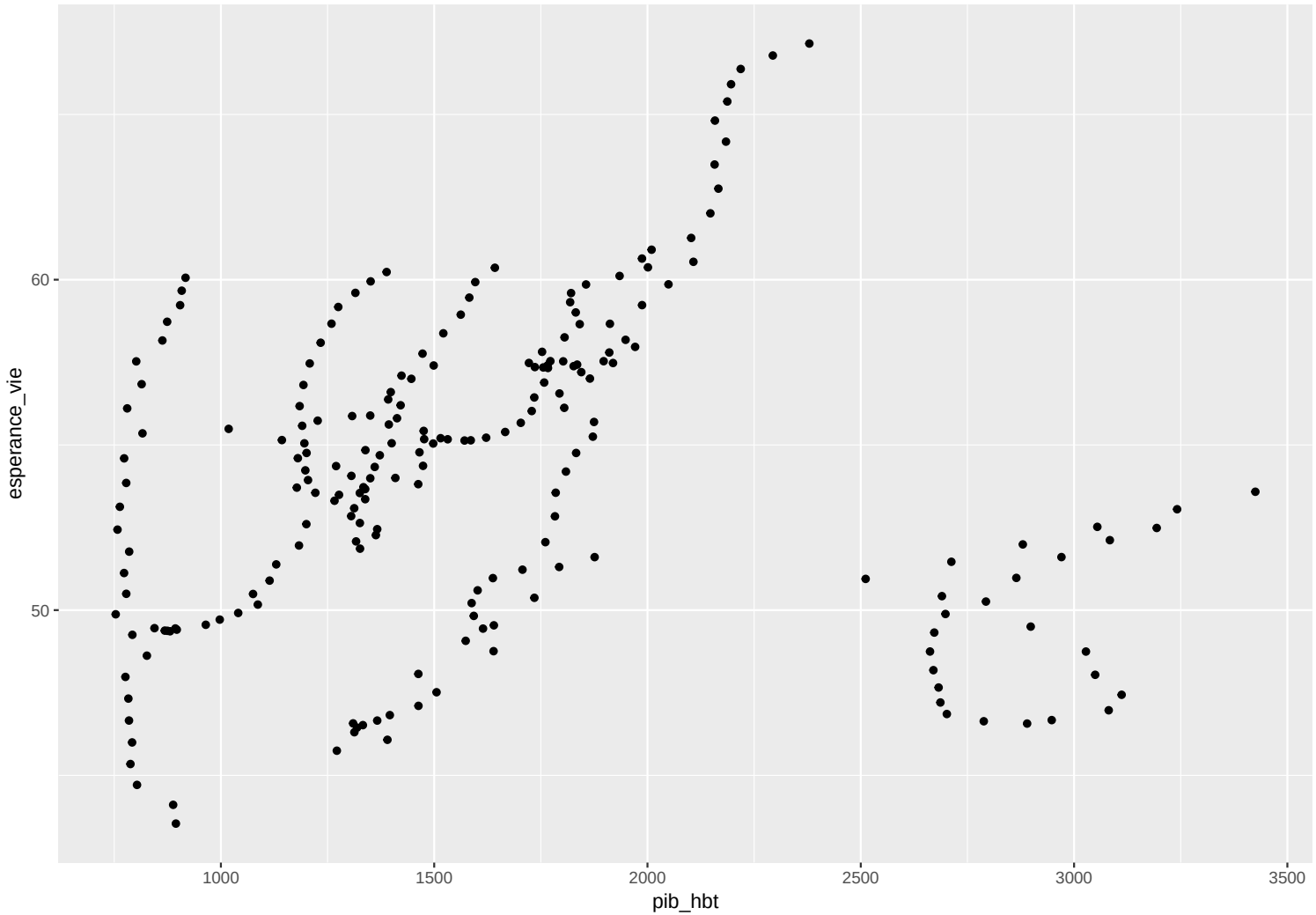
Quand on investigue la relation entre deux variables continues, les points constituent souvent le type de visualisation le plus commode. Ils permettent souvent de jetter la base de la narrative qu'on construira autour des données.

Commençons par faire une simple représentation de ces deux variables dans un plan à deux dimensions:

- le PIB/habitant sur l'axe des abscisses ;
- l'espérance de vie sur l'axe des ordonnées.

Illustration

```
ggplot(data = pibhbt_espvie, mapping = aes(x = pib_hbt, y = esperance_vie)) +
  geom_point()
```

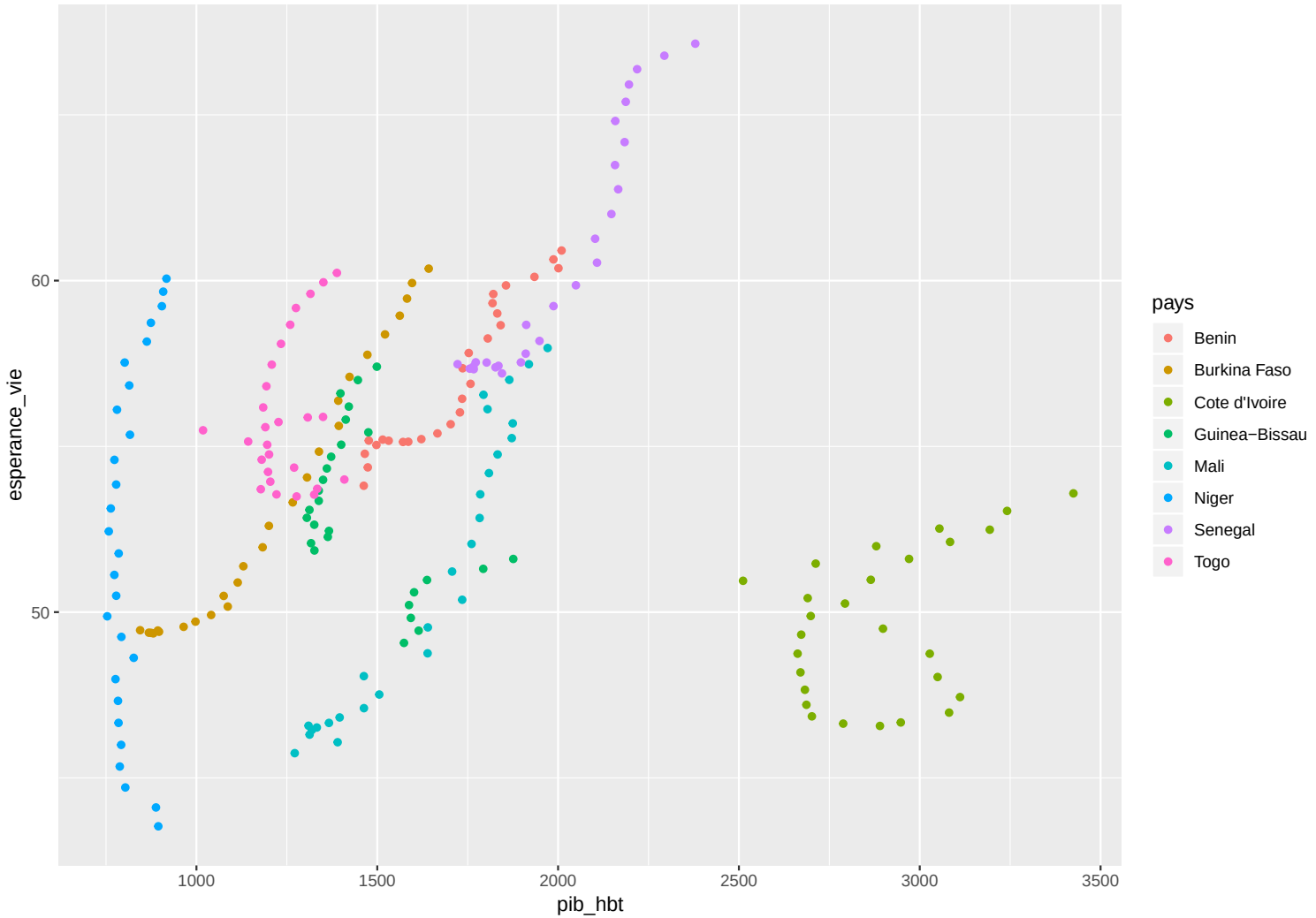


Illustration

Et si nous souhaitions distinguer entre les pays en assignant une couleur à chacun d'entre eux?

Illustration

```
ggplot(data = pibhbt_espvie, mapping = aes(x = pib_hbt, y = esperance_vie, color = pays)) +
  geom_point()
```

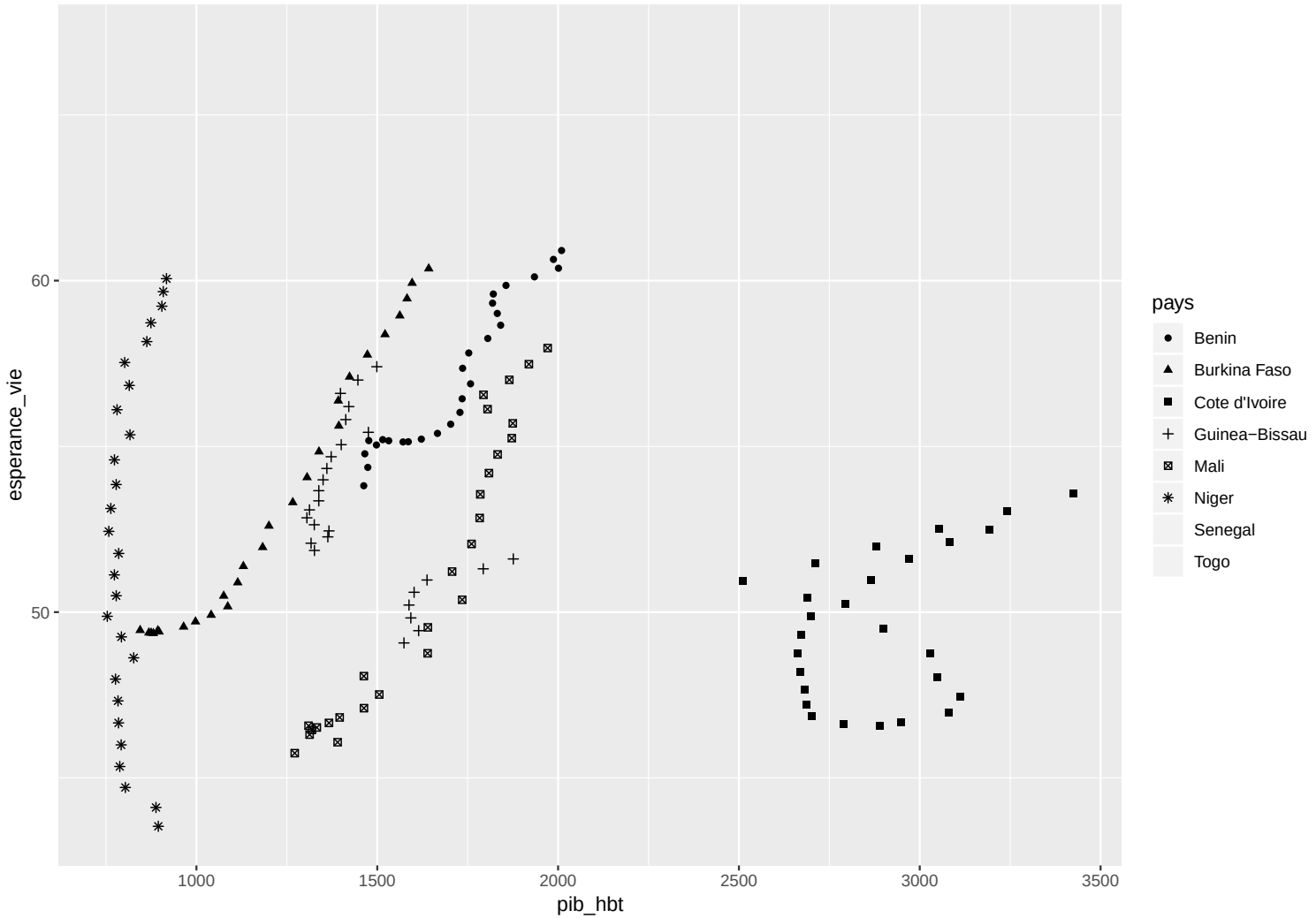


Illustration

Peut-être nous aimerions plutôt essayer des formes pour chaque pays?

Illustration

```
ggplot(data = pibhbt_espvie, mapping = aes(x = pib_hbt, y = esperance_vie, shape = pays)) +
  geom_point()
```

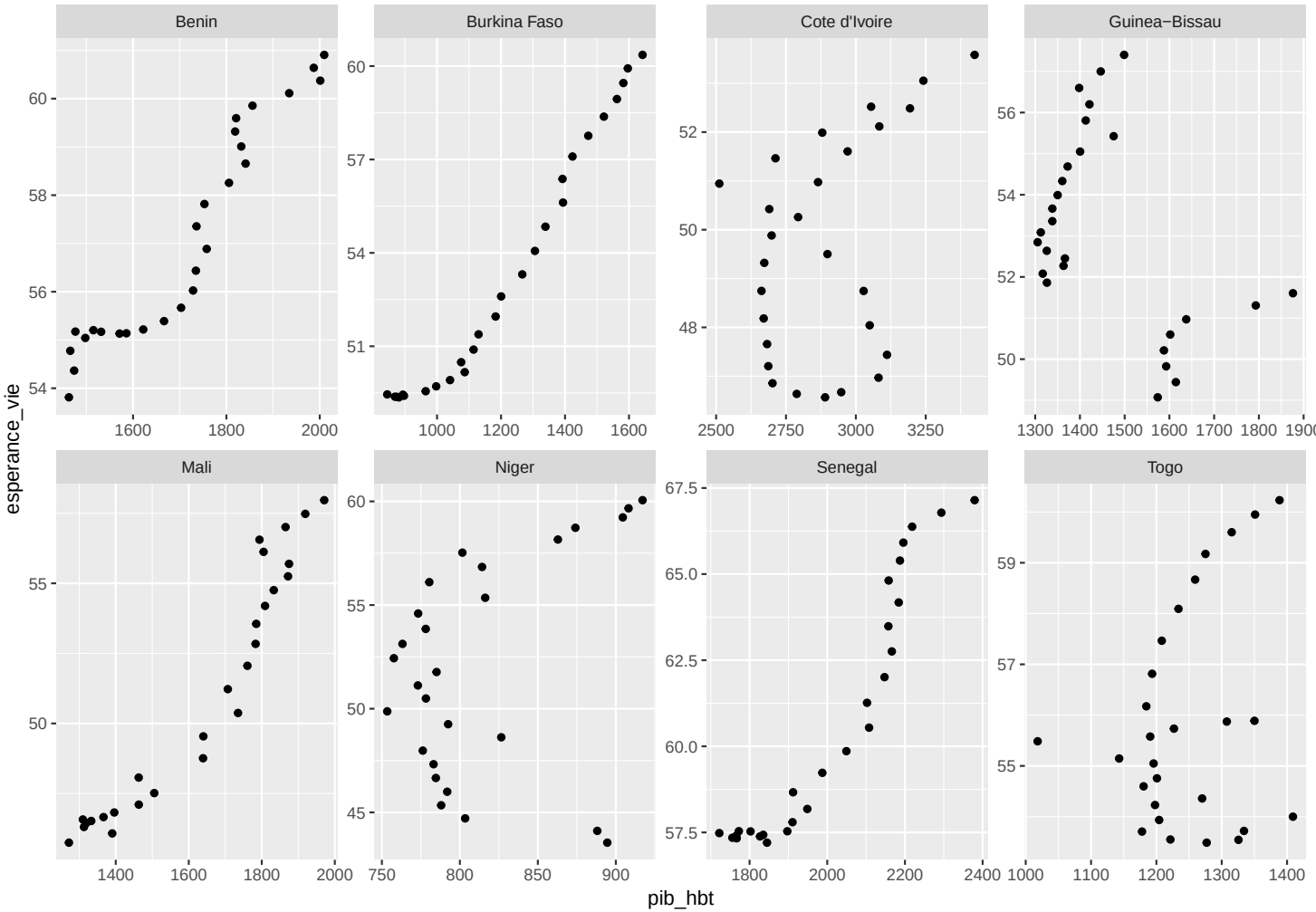


Illustration

Ou peut-être, nous voudrions les séparer, faire en sorte que chaque pays soit à part.

Illustration

```
ggplot(data = pibhbt_espvie, mapping = aes(x = pib_hbt, y = esperance_vie)) +
  geom_point() +
  facet_wrap(facets = ~pays, nrow = 2, scales = "free")
```



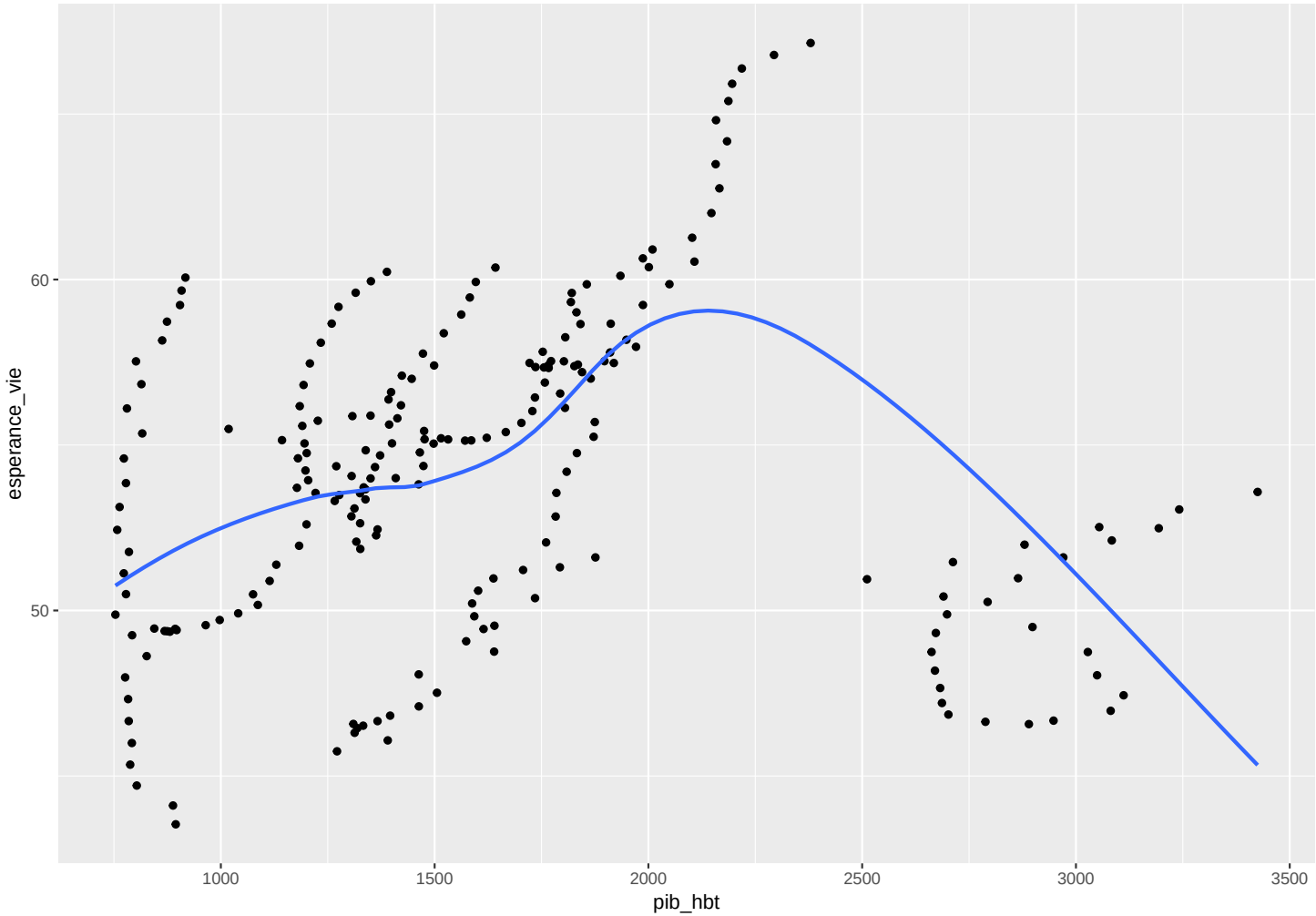
Illustration

Comme dans beaucoup de cas, une représentation de ce genre sert souvent à tester l'existence d'une relation entre les deux variables. Algébriquement, l'on déduit souvent des coefficients par la méthode des moindres carrés ordinaires pour tracer une ligne qui minimise la somme du carré des erreurs (distance entre les points observés et la droite tracée). *ggplot2* est bâti de sorte à pouvoir effectuer ce genre de calcul et projeter le résultat dans le plan.

Le calcul peut être fait au niveau global avec une droite commune à toutes les entités (les pays, ici)...

Illustration

```
ggplot(data = pibhbt_espvie, mapping = aes(x = pib_hbt, y = esperance_vie)) +
  geom_point() +
  geom_smooth(method = "loess", se = FALSE)
```

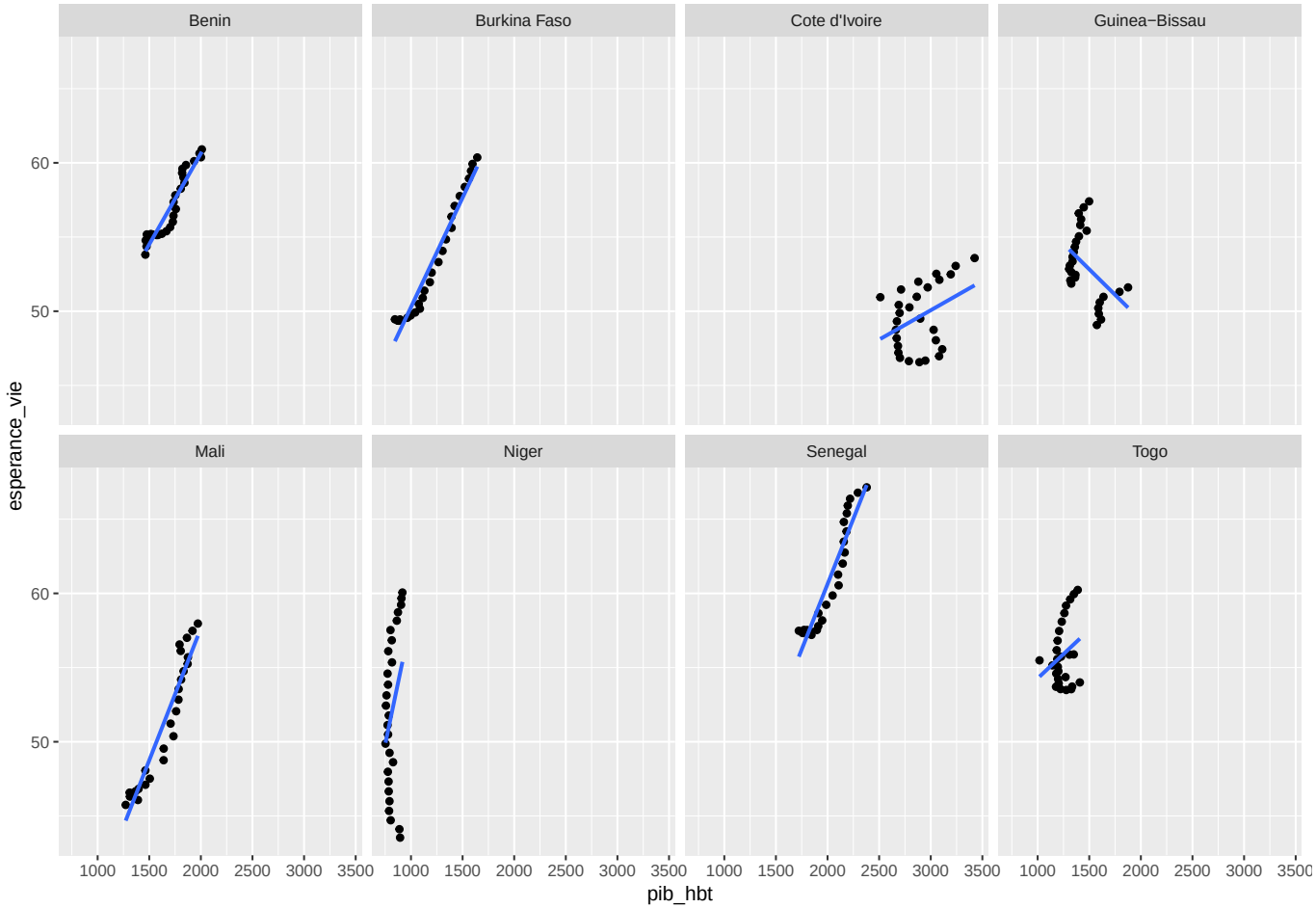


Illustration

... ou de façon spécifique à chaque entité.

Illustration

```
ggplot(data = pibhbt_espvie, mapping = aes(x = pib_hbt, y = esperance_vie)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE) +
  facet_wrap(facets = ~pays, nrow = 2)
```



A vous de jouer

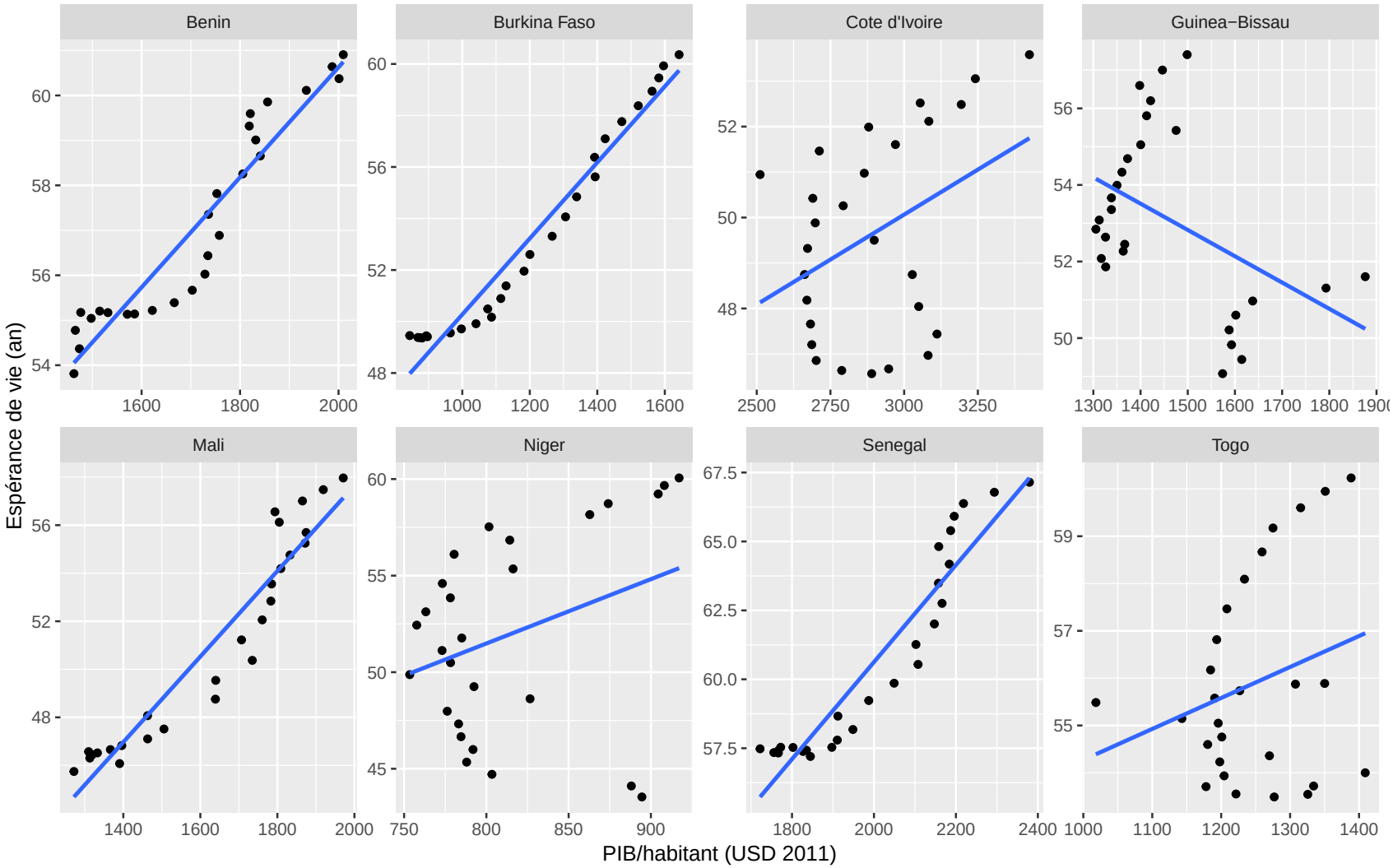
A partir des tableau et graphique suivants, narrez ce que vous observez. Les données soutiennent-elles l'idée que l'espérance de vie augmente avec le niveau de revenu d'un pays?

Table 1: PIB/habitant et espérance de vie dans les pays de l'UEMOA

pays	PIB/hbt: minimum	PIB/hbt: maximum	Espérance de vie: minimum	Espérance de vie: maximum
Benin	1462.68	2009.62	53.81	60.91
Burkina Faso	844.34	1642.32	49.36	60.36
Cote d'Ivoire	2511.42	3424.96	46.57	53.58
Guinea-Bissau	1305.43	1876.13	49.07	57.40
Mali	1271.79	1971.08	45.75	57.97
Niger	753.43	917.14	43.54	60.06
Senegal	1722.09	2379.45	57.20	67.15
Togo	1018.24	1409.02	53.49	60.23

A vous de jouer

PIB/habitant et espérance de vie dans les pays de l'UEMOA



Source: Données tirées de http://data.worldbank.org

Lignes: comprendre une évolution

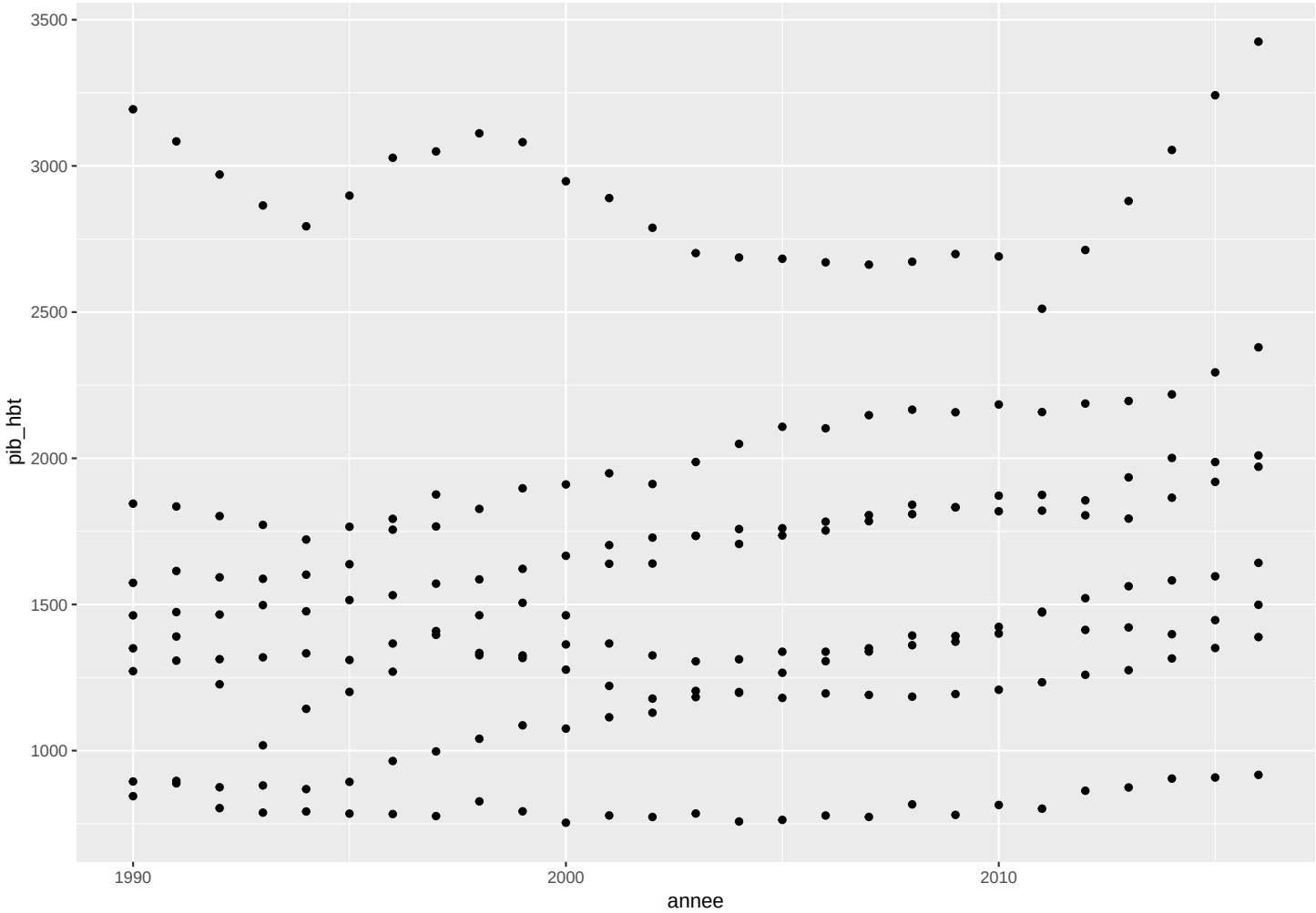
Aperçu

Ici, nous allons reconduire les données sur le PIB/habitant. Cette fois-ci, nous allons prendre en compte la dimension temporelle. Avec les séries temporelles, les lignes constituent la visualisation la plus commode. Considérons d'abord le PIB/habitant.

Commençons par regarder les points...

Illustration

```
ggplot(data = pibhbt_espvie, mapping = aes(x = annee, y = pib_hbt)) +
  geom_point()
```

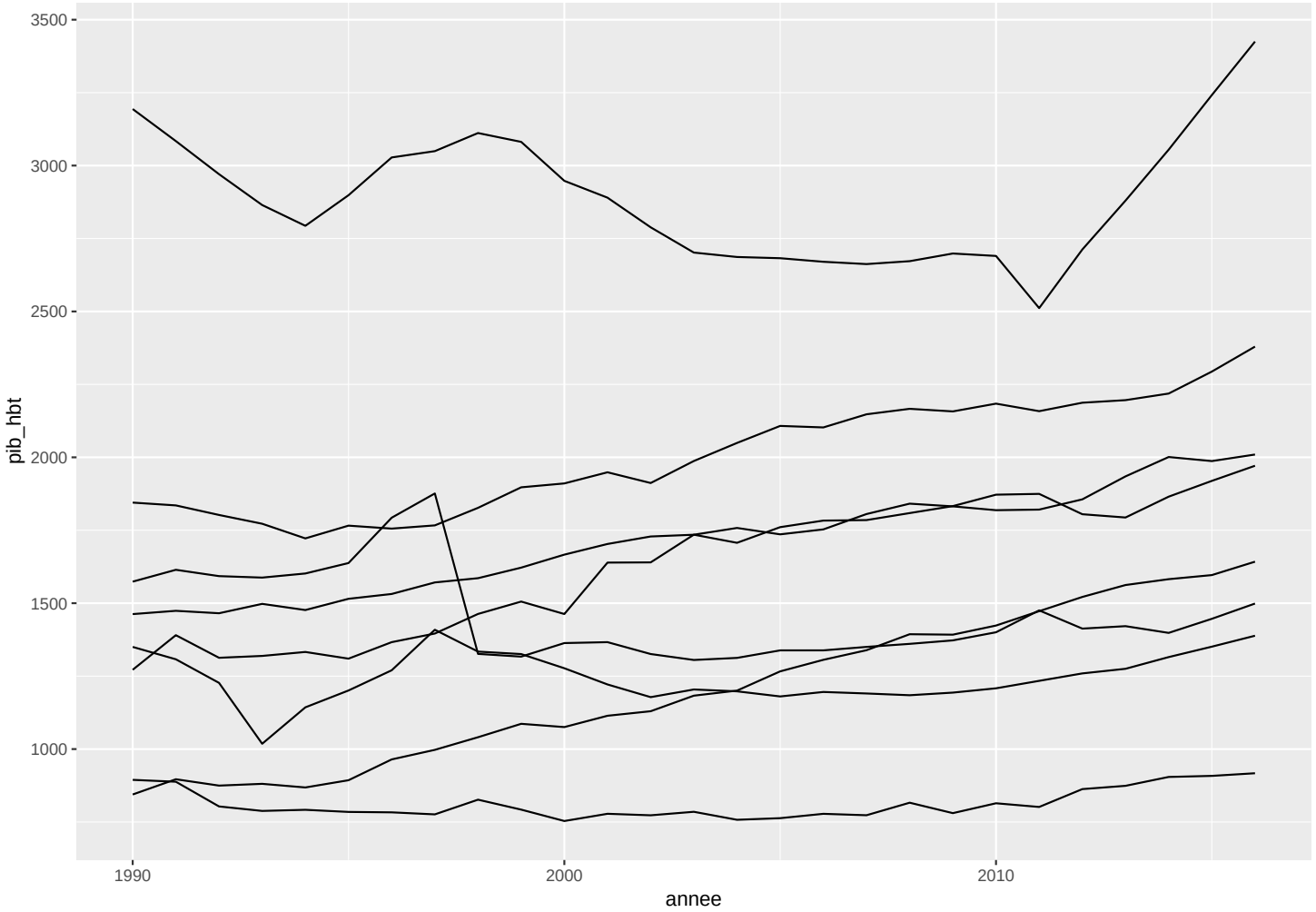


Illustration

On voit que les points affichent plusieurs trajectoires. Sur la base des données, l'on en déduit qu'il faut une ligne pour chaque pays.
Groupons-les donc par pays pour en faire des lignes.

Illustration

```
ggplot(data = pibhbt_espvie, mapping = aes(x = annee, y = pib_hbt, group = pays)) +
  geom_line()
```

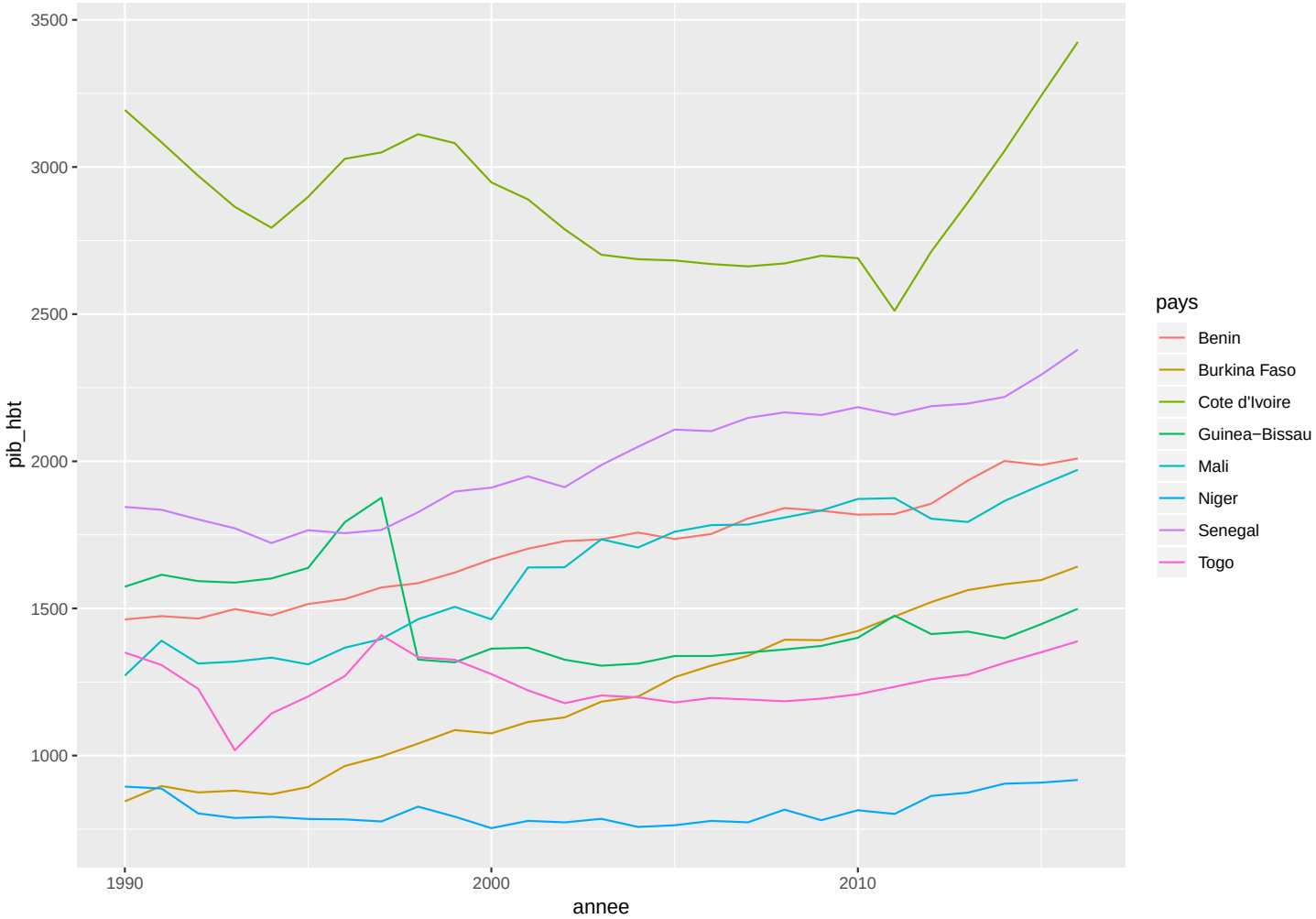


Illustration

Comme avant, il peut être utile d'introduire des couleurs...

Illustration

```
ggplot(data = pibhbt_espvie, mapping = aes(x = annee, y = pib_hbt, color = pays)) +
  geom_line()
```

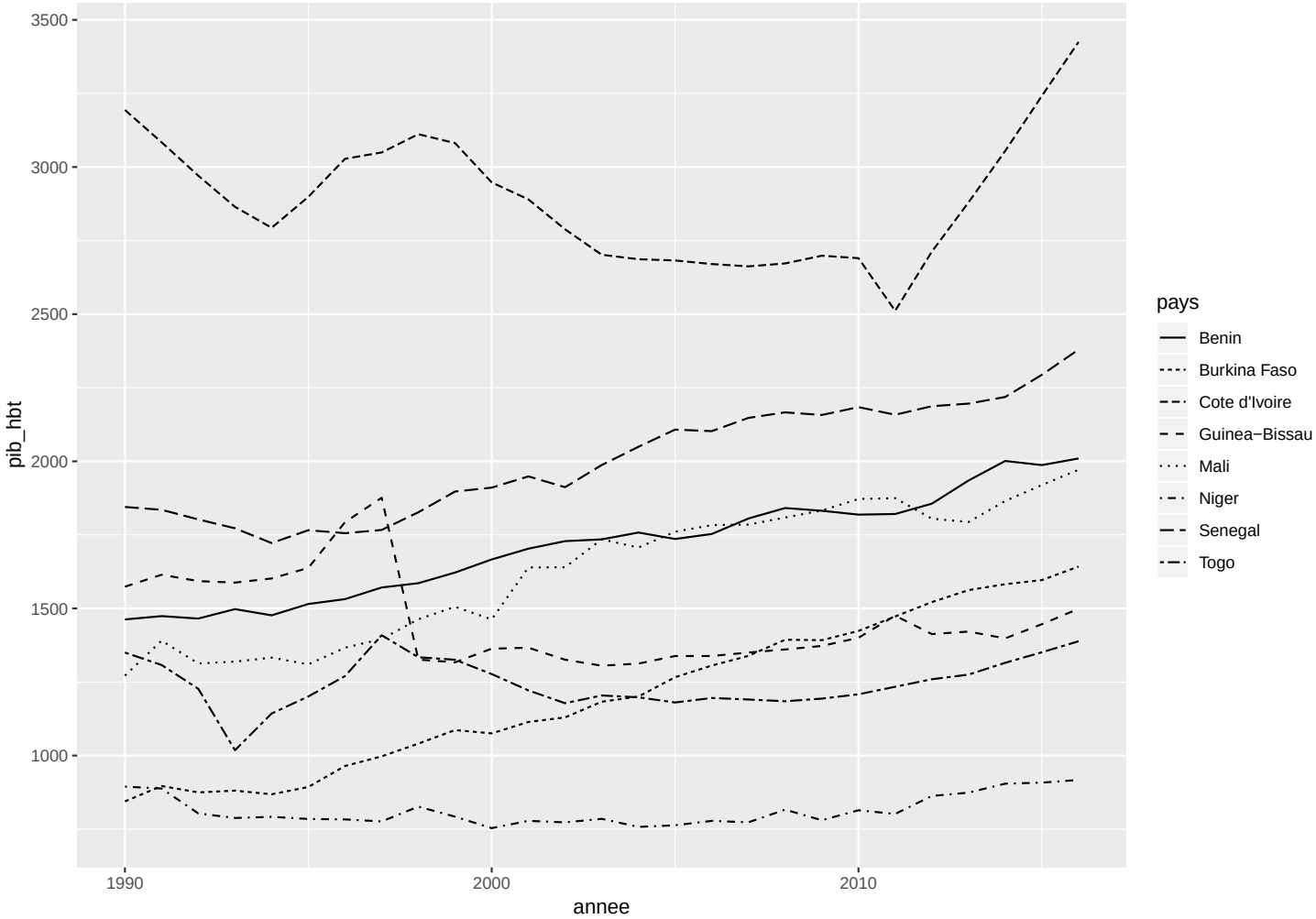


Illustration

... ou de différencier entre les lignes par type...

Illustration

```
ggplot(data = pibhbt_espvie, mapping = aes(x = annee, y = pib_hbt, linetype = pays)) +
  geom_line()
```

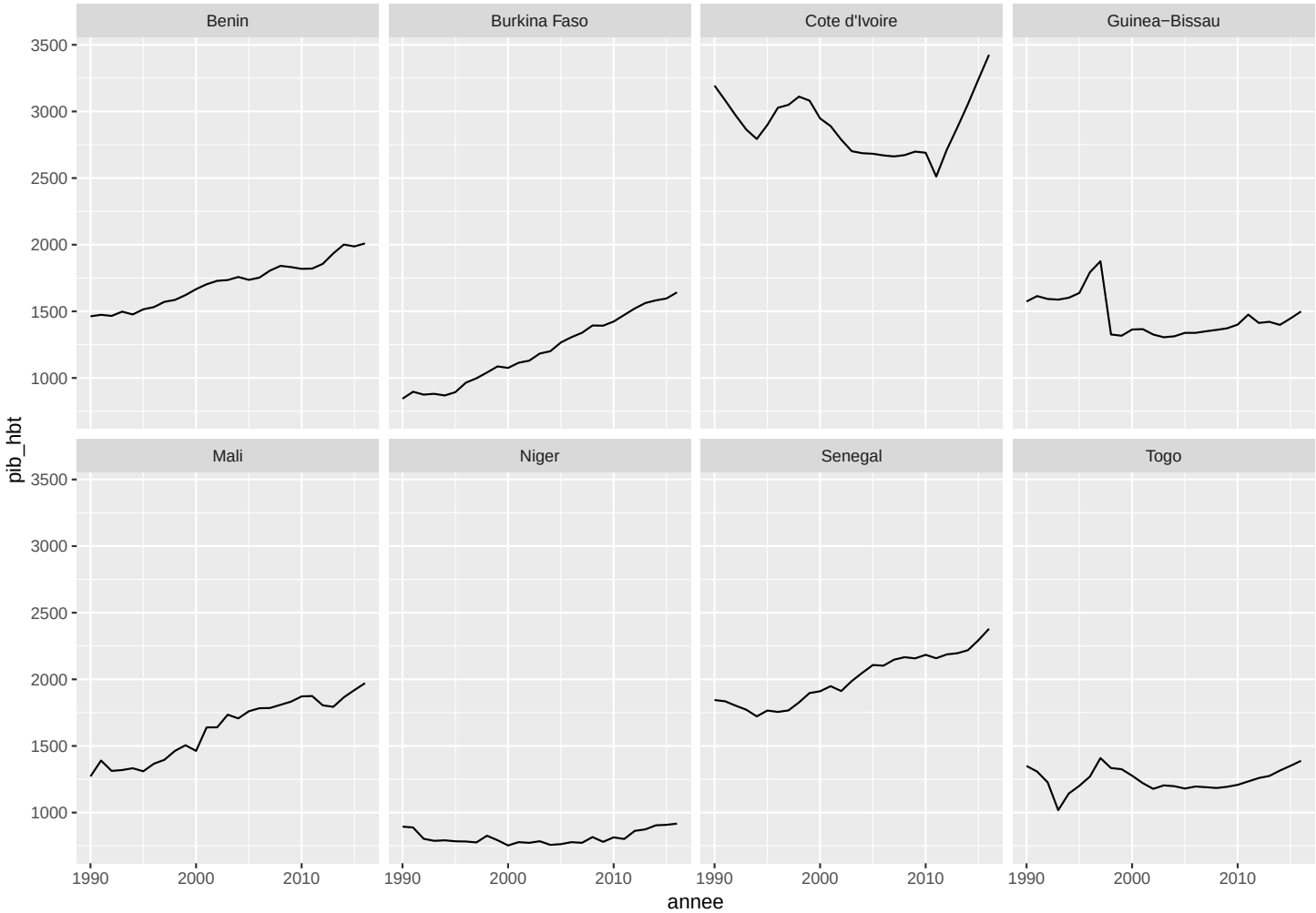


Illustration

... ou tout simplement de séparer par pays.

Illustration

```
ggplot(data = pibhbt_espvie, mapping = aes(x = annee, y = pib_hbt)) +
  geom_line() +
  facet_wrap(facets = ~pays, nrow = 2)
```



Illustration

Comme avant, l'on peut partir de la visualisation pour poser les bases de la modélisation. L'on peut revenir au point pour ensuite introduire en droite de lissage de la série.

Illustration

```
ggplot(data = pibhbt_espvie, mapping = aes(x = annee, y = pib_hbt)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE, color = "red") +
  facet_wrap(facets = ~pays, nrow = 2)
```

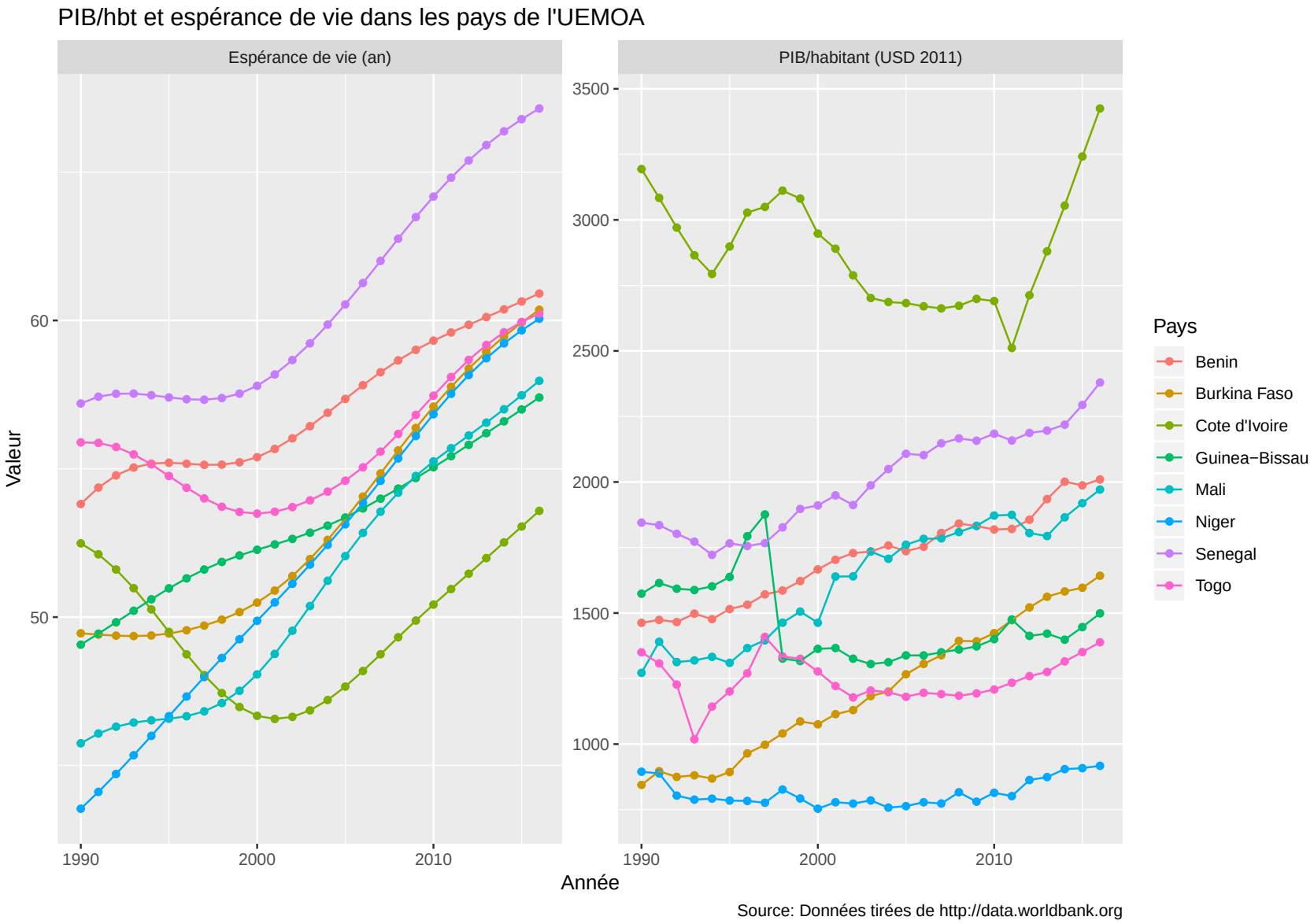
A vous de jouer

A partir des tableau et graphique suivants, narrez ce que vous observez. Comment est-ce que le PIB/habitant et l'espérance de vie ont évolué dans les pays de l'UEMOA sur la période couverte par les données? Si vous souhaitez, vous pouvez appliquer à l'espérance de vie les codes utilisés sur le PIB/habitant.

Table 2: PIB/habitant et espérance de vie dans les pays de l'UEMOA: taux de croissance/an

Pays	PIB/hbt (%)	Espérance de vie (%)
Benin	1.23	0.48
Burkina Faso	2.59	0.77
Cote d'Ivoire	0.27	0.08
Guinea-Bissau	-0.19	0.61
Mali	1.70	0.91
Niger	0.10	1.24
Senegal	0.98	0.62
Togo	0.11	0.29

A vous de jouer



Barres: allier le discret au continu (et au discret)

Aperçu

Très souvent dans le processus exploratoire de données, le *data scientist* est amené à procéder à des comparaisons entre différentes entités (pays, par exemple). Généralement ces entités sont en nombre fini (on ne crée pas un nouveau pays tous les jours!) tandis que les attributs ou caractéristiques sur lequels ces comparaisons portent sont souvent continus (le PIB/habitant, par exemple). Dans ce genre de cas, les barres apparaissent comme le meilleur moyen de visualiser l'étendu des écarts ou des similarités.

Les barres sont aussi utiles pour visualiser la composition d'une entité. Par exemple, la répartition de la population d'un pays entre les groupes d'âge ou entre les deux sexes. Là, elles se prêtent à l'examen d'un groupe à travers le croisement de deux attributs catégoriels.

Illustrons tout ça avec quelques exemples!

Aperçu

Ici, nous allons utiliser des données tirées des Recensements Généraux de la Population et de l'Habitat conduits au Mali en 1976, 1987, 1998 et 2009. Ces données sont tirées du site de l' Institut National de la Statistique (INSTAT).

```
groupage_rgph <- read_csv("data/groupage_rgph.csv") # Importation du ficher CSV: RGPH: 1976, 1987, 1998, 2009
groupage_rgph <- groupage_rgph %>%
  mutate(groupage = factor(groupage,
                           levels = unique(groupage_rgph$groupage),
                           labels = unique(groupage_rgph$groupage),
                           ordered = TRUE)) # Classement ordonné des groupes d'âge
glimpse(groupage_rgph) # Créer un aperçu de la structure des données
```

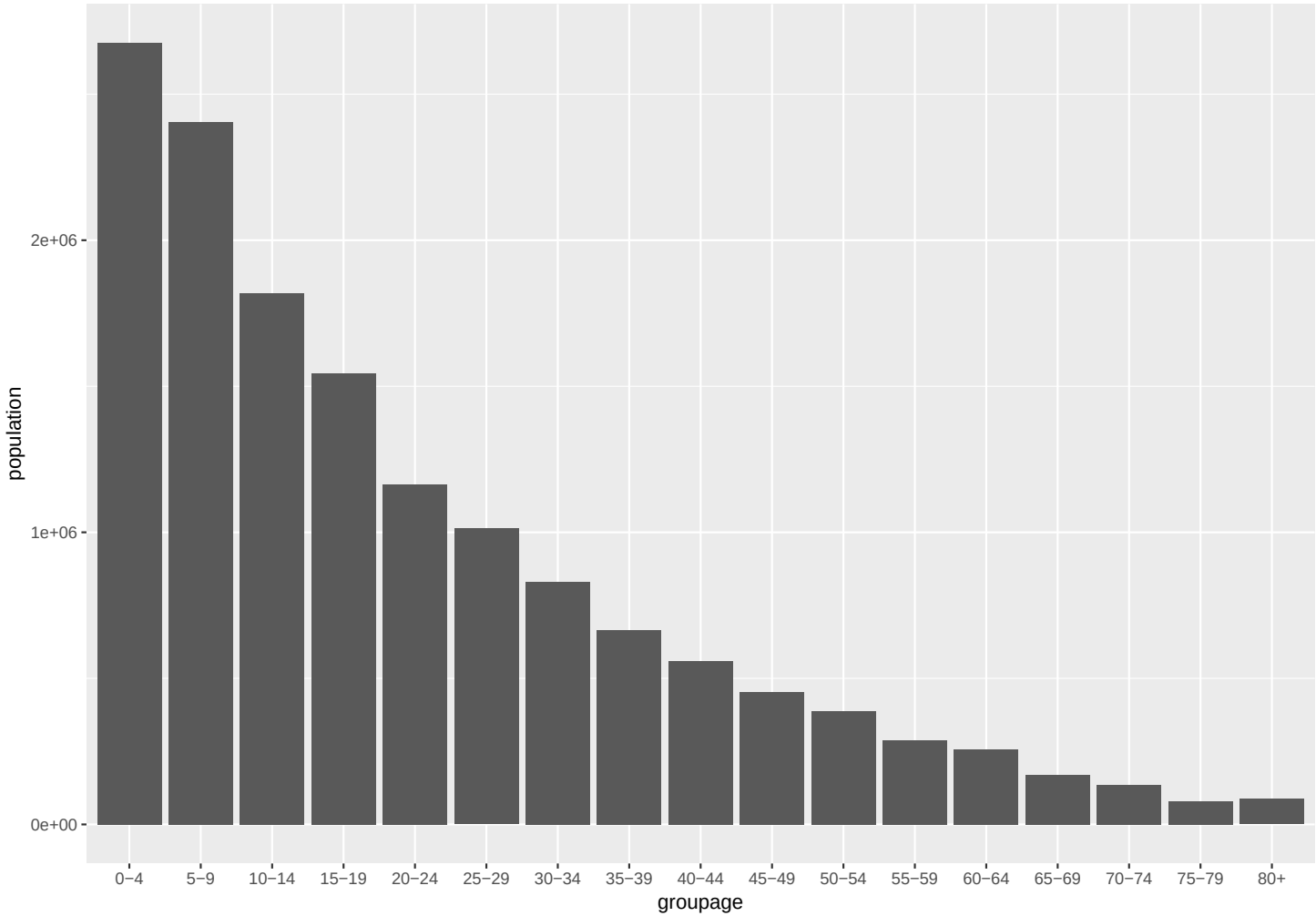
```
## Observations: 304
## Variables: 8
## $ annee      <dbl> 1976, 1976, 1976, 1976, 1976, 1976, 1976, 1976, 197...
## $ id         <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, ...
## $ groupage   <ord> 0-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39,...
## $ sexe       <chr> "Homme", "Homme", "Homme", "Homme", "Homme", "Homme...
## $ milieu     <chr> "Urbain", "Urbain", "Urbain", "Urbain", "Urbain", "...
## $ population <dbl> 100416, 80816, 60647, 58662, 46239, 36390, 30895, 2...
## $ source     <chr> "RGPH", "RGPH", "RGPH", "RGPH", "RGPH", "RGPH", "RG...
## $ office     <chr> "DNSI", "DNSI", "DNSI", "DNSI", "DNSI", "DNSI", "DN...
```

Illustration

Commençons par regarder la répartition de la population entre les groupes d'âge sur l'année 2009.

Illustration

```
ggplot(data = groupage_rgph %>% filter(annee == 2009, !(groupage %in% c("Total", "ND"))),
       mapping = aes(x = groupage , y = population)) +
  geom_bar(stat = "identity")
```



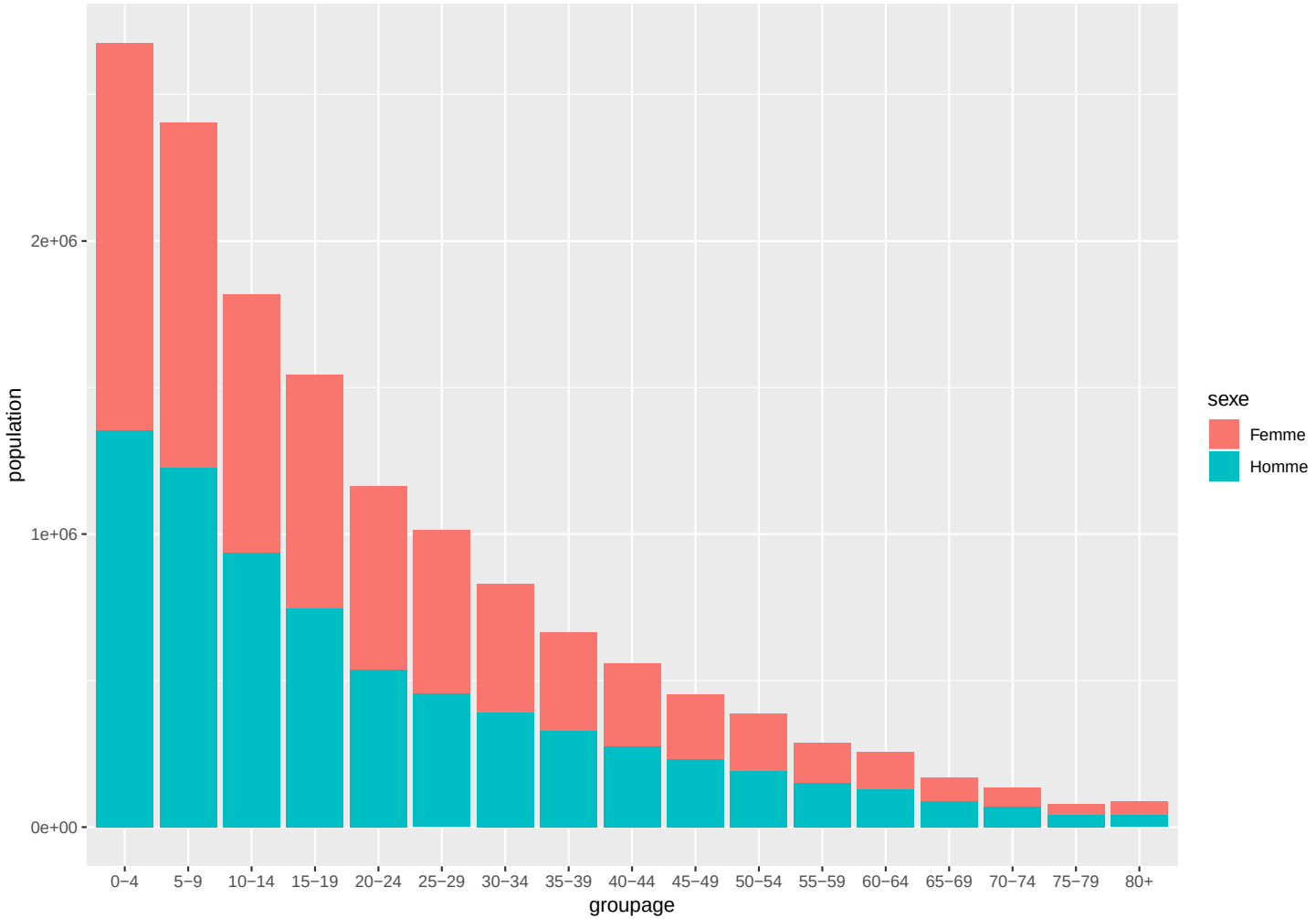
Illustration

Maintenant, sachant qu'à l'intérieur des groupes d'âge, nous avons la distinction faite entre hommes et femmes et entre habitants en milieu rural et citadins, essayons de rendre compte de ces distinctions dans les barres.

Commençons par la différence en termes de sexe...

Illustration

```
ggplot(data = groupage_rgph %>% filter(annee == 2009, !(groupage %in% c("Total", "ND"))),
       mapping = aes(x = groupage , y = population, fill = sexe)) +
  geom_bar(stat = "identity")
```

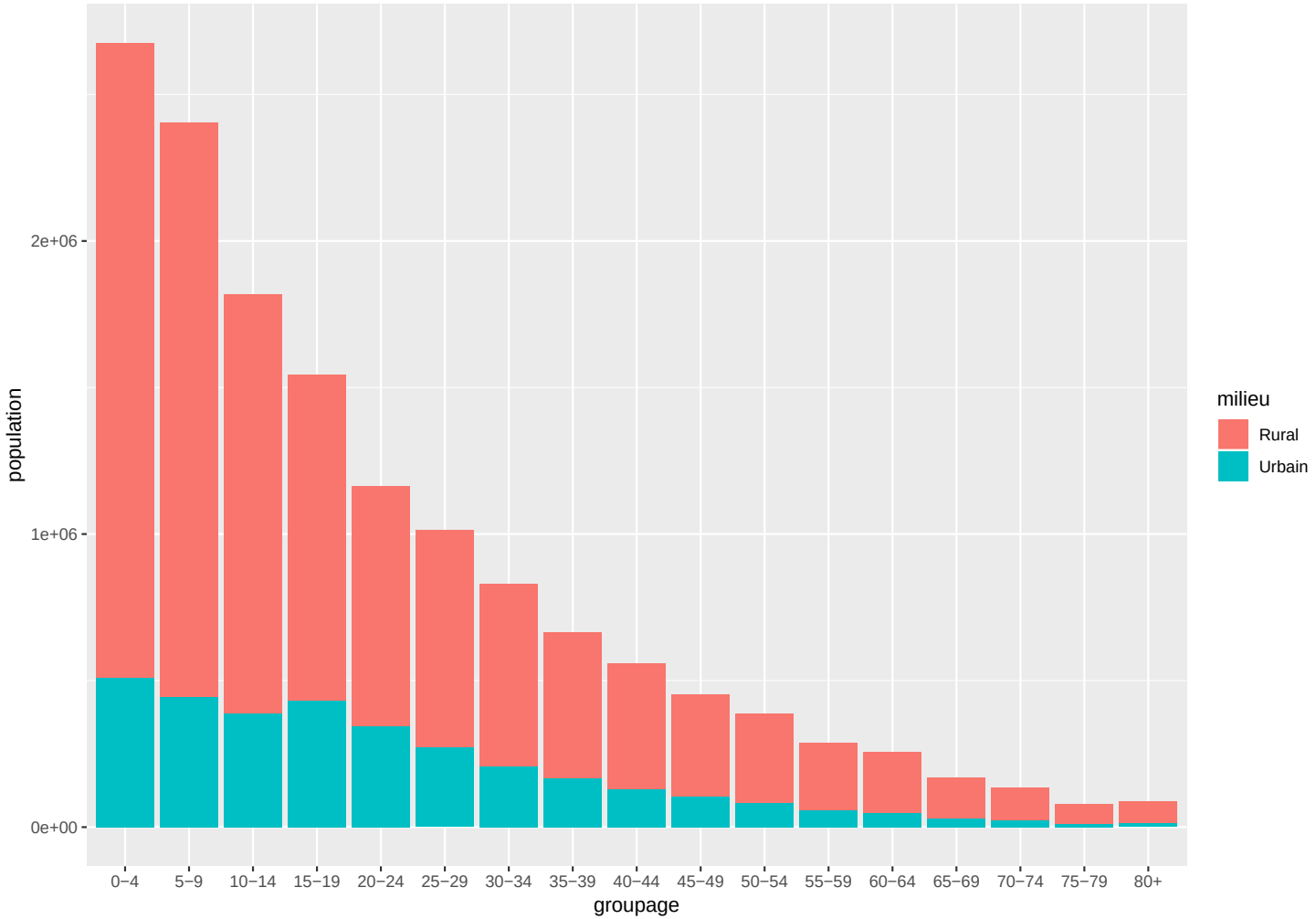


Illustration

... et en termes de milieu

Illustration

```
ggplot(data = groupage_rgph %>% filter(annee == 2009, !(groupage %in% c("Total", "ND"))),
       mapping = aes(x = groupage , y = population, fill = milieu)) +
  geom_bar(stat = "identity")
```

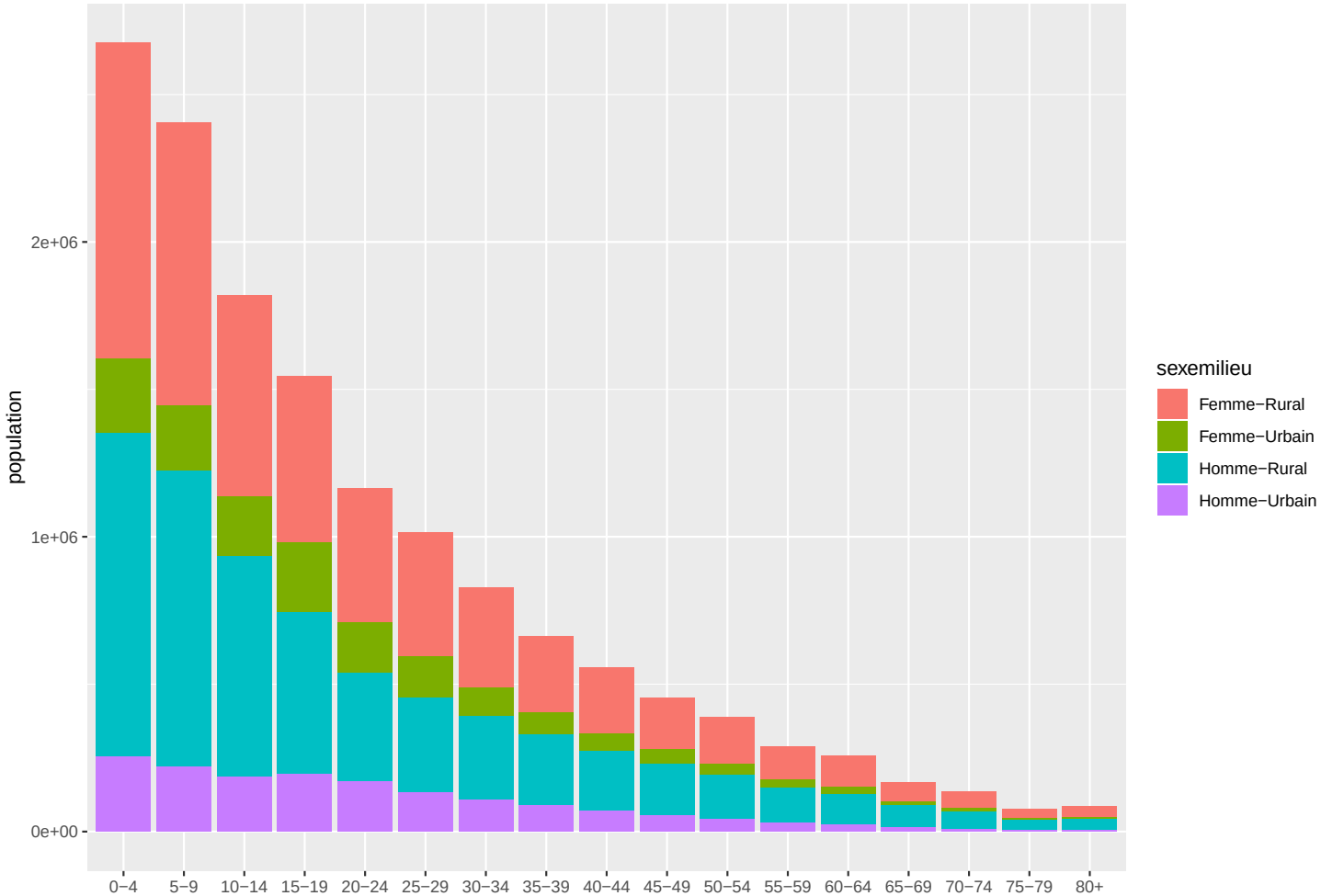


Illustration

Maintenant, considérons les deux.

Illustration

```
ggplot(data = groupage_rgph %>% filter(annee == 2009, !(groupage %in% c("Total", "ND"))) %>%
         unite("sexemilieu", c(sexe, milieu), sep = "-"),
       mapping = aes(x = groupage , y = population, fill = sexemilieu)) +
  geom_bar(stat = "identity")
```

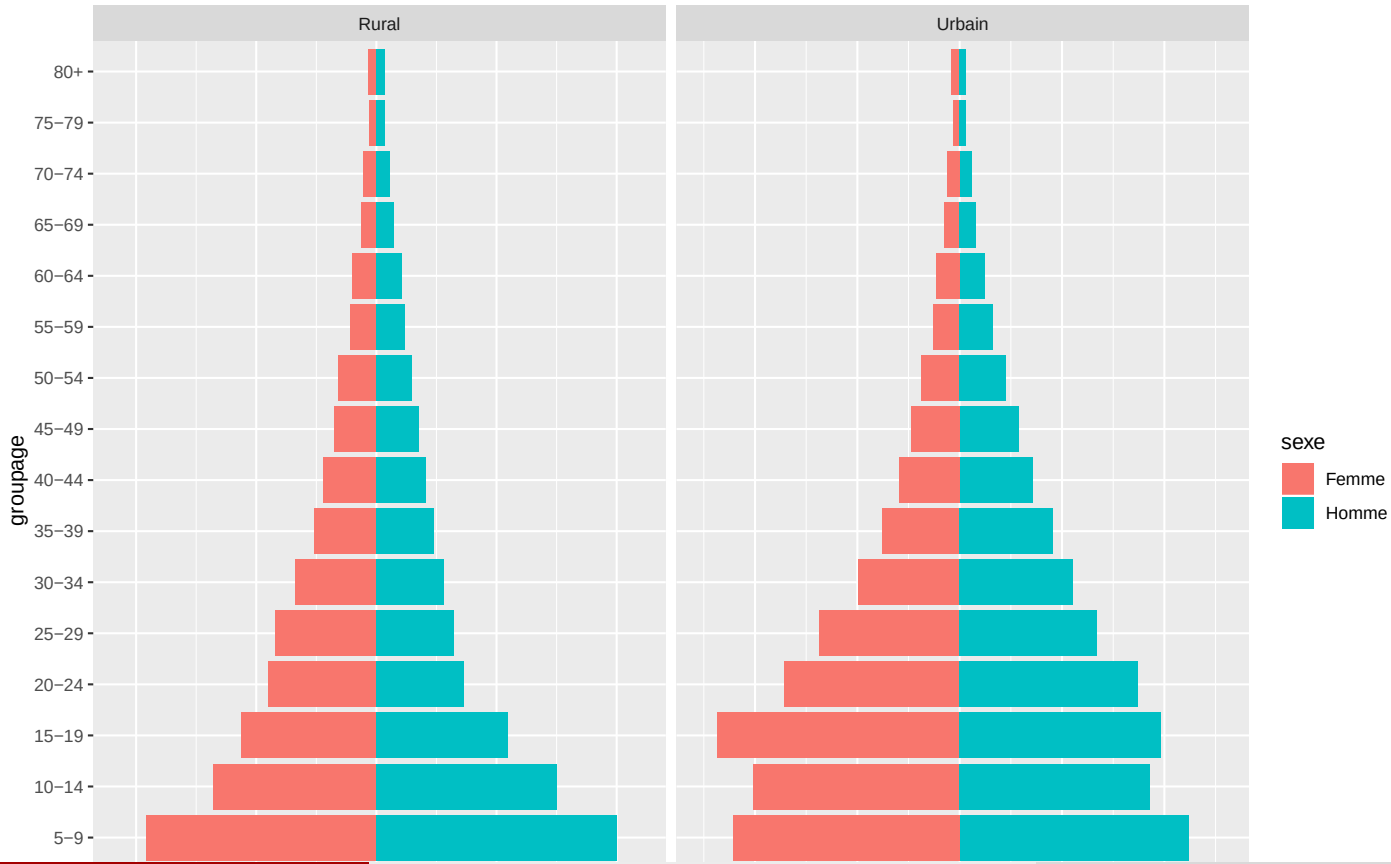


Illustration

Avec quelques modifications, l'on peut augmenter le pouvoir informatif de ces barres. Transformons-les en pyramides des âges!

Illustration

```
ggplot(data = groupage_rgph %>% filter(annee == 2009, !(groupage %in% c("Total", "ND"))) %>%
         mutate(population = ifelse(sexe == "Femme", -population, population)),
       mapping = aes(x = groupage , y = population, fill = sexe)) +
  geom_bar(stat = "identity") +
  coord_flip() +
  facet_wrap(facets = ~milieu, scales = "free_x") +
  scale_y_continuous(labels = abs)
```



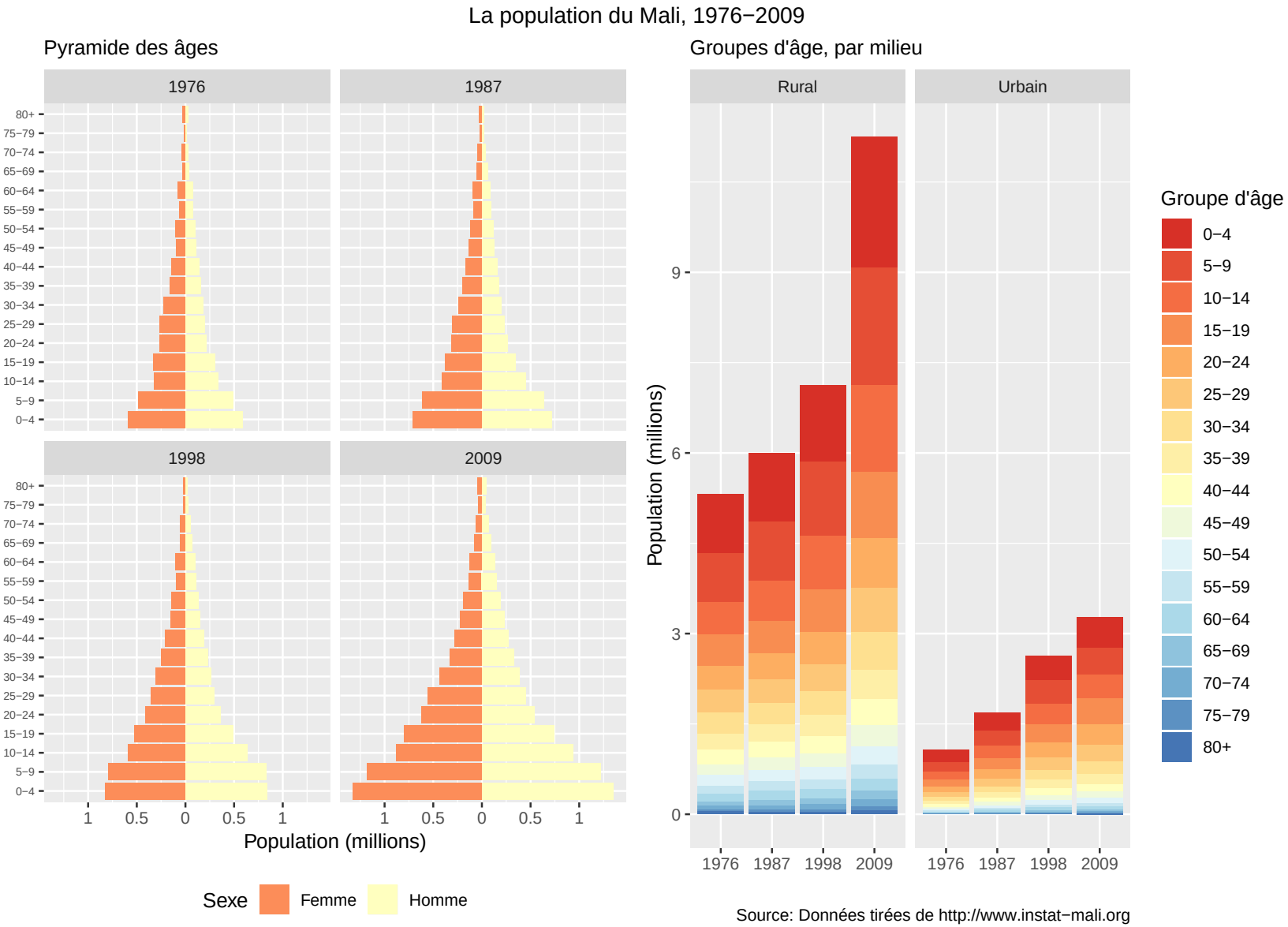
A vous de jouer

A partir des tableau et graphiques suivants, narrez ce que vous observez. Quel commentaire peut-on en faire en ce qui concerne l'évolution démographique du Mali?

Table 3: Population du Mali, 1976-2009

Sexe	Milieu	1976	1987	1998	2009
Femme	Rural	2723571	3082636	3634679	5692616
Femme	Urbain	545614	853002	1320210	1631056
Homme	Rural	2594518	2923423	3530816	5561319
Homme	Urbain	529215	837288	1325207	1643671
Total	Total	6392918	7696349	9810912	14528662

A vous de jouer



Source: Données tirées de http://www.instat−mali.org

Histogrammes et densités: comprendre une distribution

Aperçu

Nous allons utiliser ici des données issues des Recensements Généraux de la Population et de l'Habitat conduits au Mali en 1998 et 2009. Il s'agit de la population par commune. Commençons par importer les données dans notre environnement de travail.

```
adm3_pop_1998 <- read_csv("data/adm3_pop_1998.csv") # Importation du ficher CSV: RGPH-1998
adm3_pop_2009 <- read_csv("data/adm3_pop_2009.csv") # Importation du ficher CSV: RGPH-2009
adm3_pop_rgph <- bind_rows(adm3_pop_1998, adm3_pop_2009) # Concatenation des deux fichiers
glimpse(adm3_pop_1998) # Créer un aperçu de la structure des données
```

```
## Observations: 703
## Variables: 13
## $ Admin0_Nam <chr> "Mali", "Mali", "Mali", "Mali", "Mali", "Mali", "Ma...
## $ Pcode_Ad_0 <chr> "ML", "ML", "ML", "ML", "ML", "ML", "ML", "ML", "ML...
## $ Admin1_Nam <chr> "Kayes", "Kayes", "Kayes", "Kayes", "Kayes", "Kayes...
## $ Pcode_Ad_1 <chr> "ML01", "ML01", "ML01", "ML01", "ML01", "ML01", "ML...
## $ Admin2_Nam <chr> "Kayes", "Kayes", "Kayes", "Kayes", "Kayes", "Kayes...
## $ Pcode_Ad_2 <chr> "ML0103", "ML0103", "ML0103", "ML0103", "ML0103", "...
## $ Admin3_Nam <chr> "Bangassi", "Colimbine", "Diamou", "Djelebou", "Fal...
## $ Pcode_Ad_3 <chr> "ML010301", "ML010302", "ML010303", "ML010304", "ML...
## $ annee      <dbl> 1998, 1998, 1998, 1998, 1998, 1998, 1998, 1998, 199...
## $ homme      <dbl> 3351, 4828, 6361, 7903, 3232, 1333, 2682, 1295, 671...
## $ femme      <dbl> 3263, 4913, 6189, 8403, 3309, 1355, 2614, 1294, 752...
## $ total      <dbl> 6614, 9741, 12550, 16306, 6541, 2688, 5296, 2589, 1...
## $ source     <chr> "RGPH", "RGPH", "RGPH", "RGPH", "RGPH", "RGPH", "RG...
```

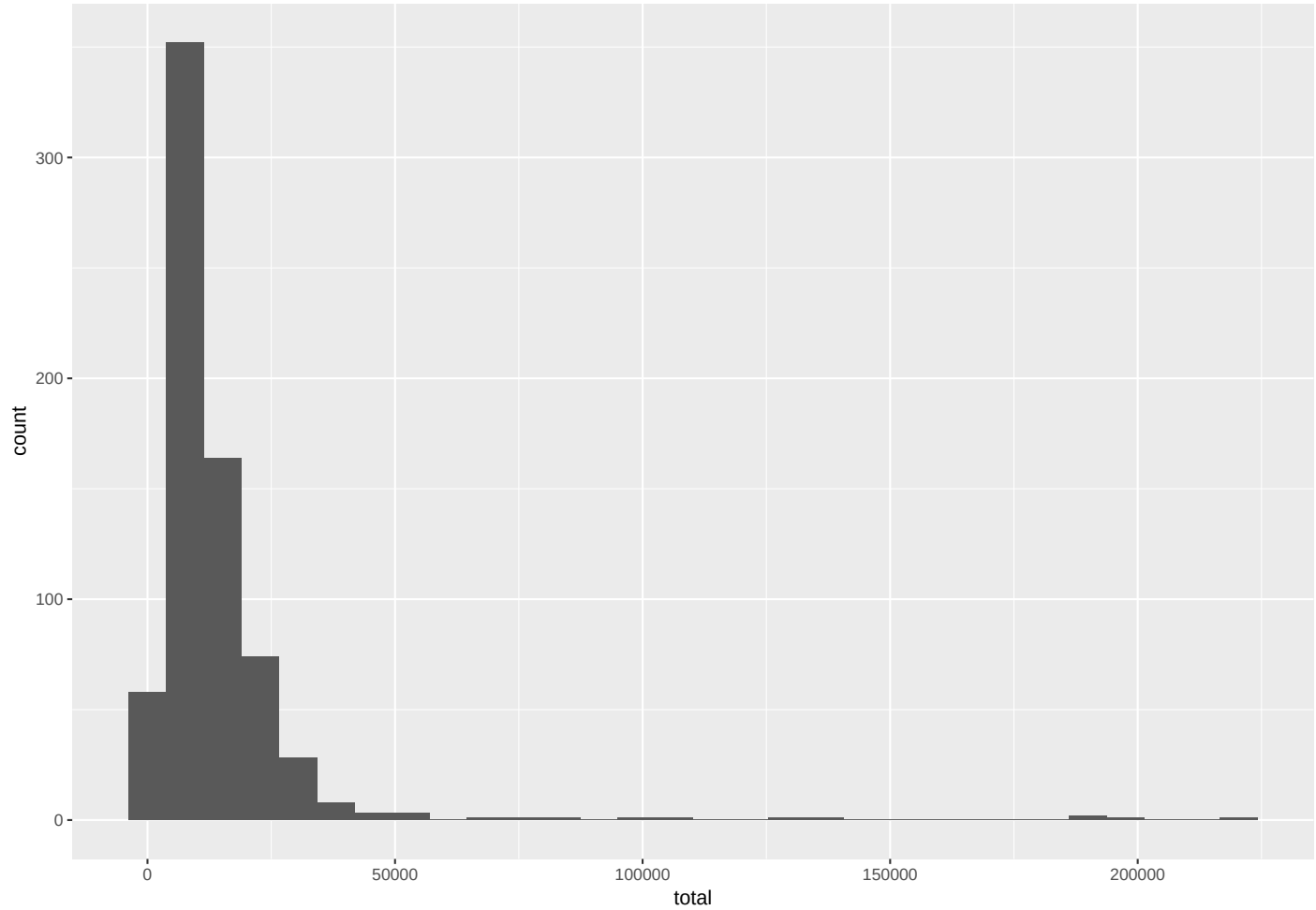
Illustration

Maintenant, explorons les données sur la population à travers la visualisation. Quelle est la taille moyenne? Quelle est la taille médiane? Quelle est l'étendu de la dispersion autour de la moyenne? Quelle sont les cercles et/ou régions où se trouvent les communes les plus peuplées? Les communes les moins peuplées? Dans quelle mesure la situation a-t-elle changé entre les deux recensements, 1998 et 2009?

Voici quelques questions qui peuvent nous servir de point de départ.

Illustration

```
ggplot(data = adm3_pop_1998,
       mapping = aes(x = total)) +
  geom_histogram()
```



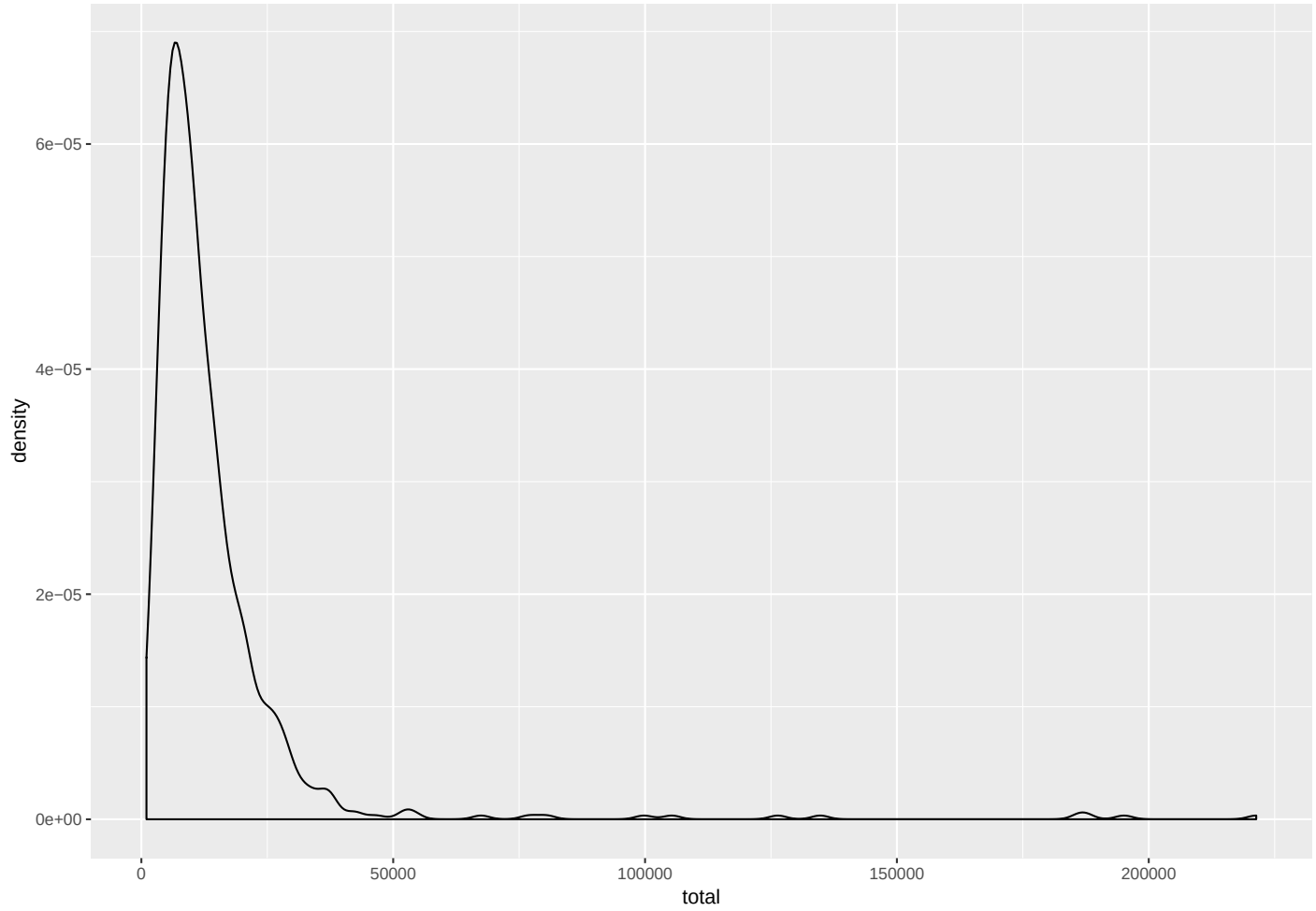
Illustration

Là nous voyons que la majorité des communes du Mali comptaient un population inférieure à 50000 personnes en 1998.

Avec la densité, l'on peut montrer la même information, mais cette fois-ci en terme de proportion.

Illustration

```
ggplot(data = adm3_pop_1998, mapping = aes(x = total)) +
  geom_density()
```

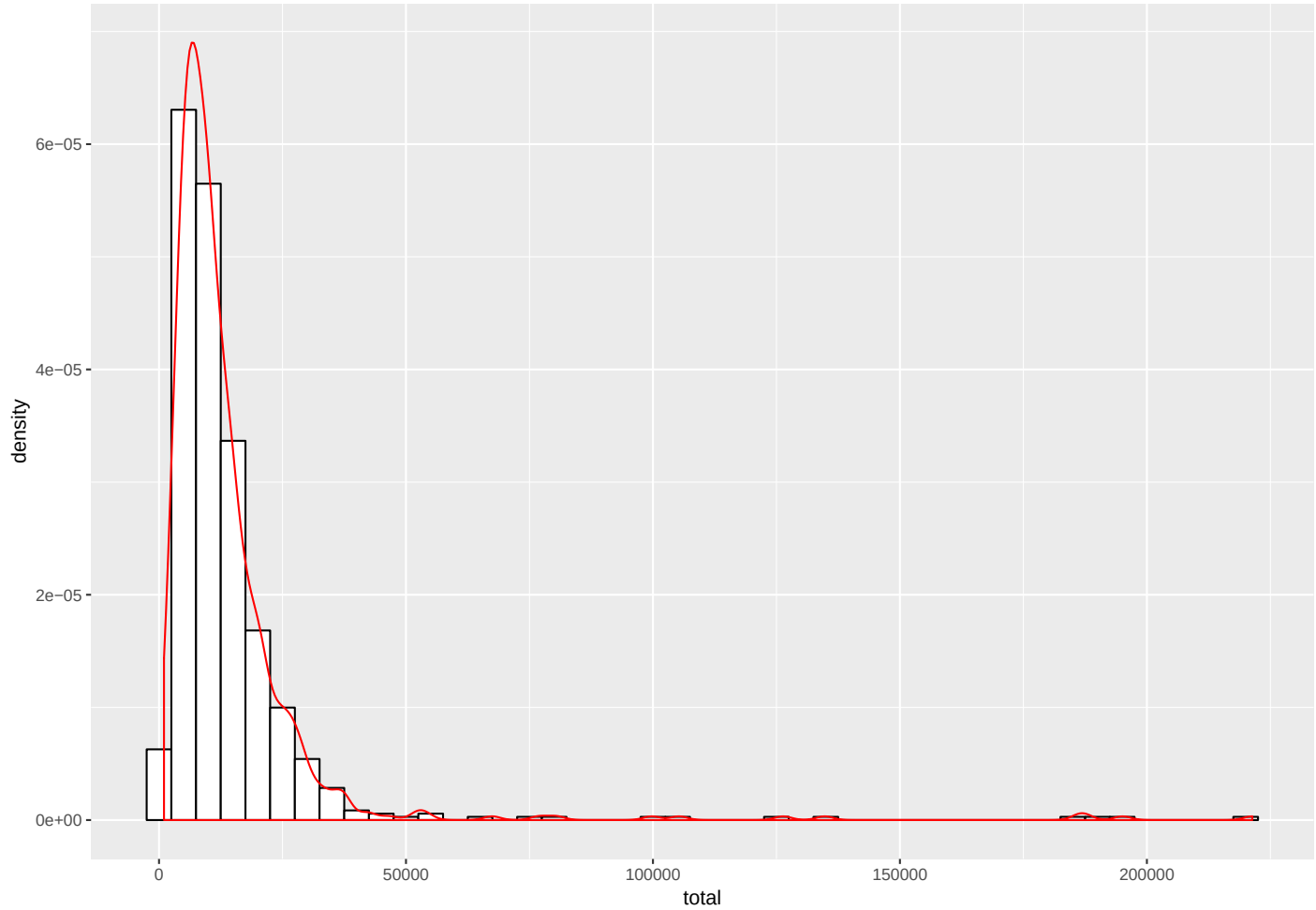


Illustration

Imaginons maintenant que l'on veuille voir à la fois les histogrammes et la densité. On procède d'abord à une transformation des histogrammes en proportions et on superpose les deux informations.

Histogrammes et densités: comprendre une distribution

```
ggplot(data = adm3_pop_1998, mapping = aes(x = total)) +
  geom_histogram(mapping = aes(y =..density..), binwidth = 5000,color = "black", fill = "white") +
  geom_density(color = "red")
```

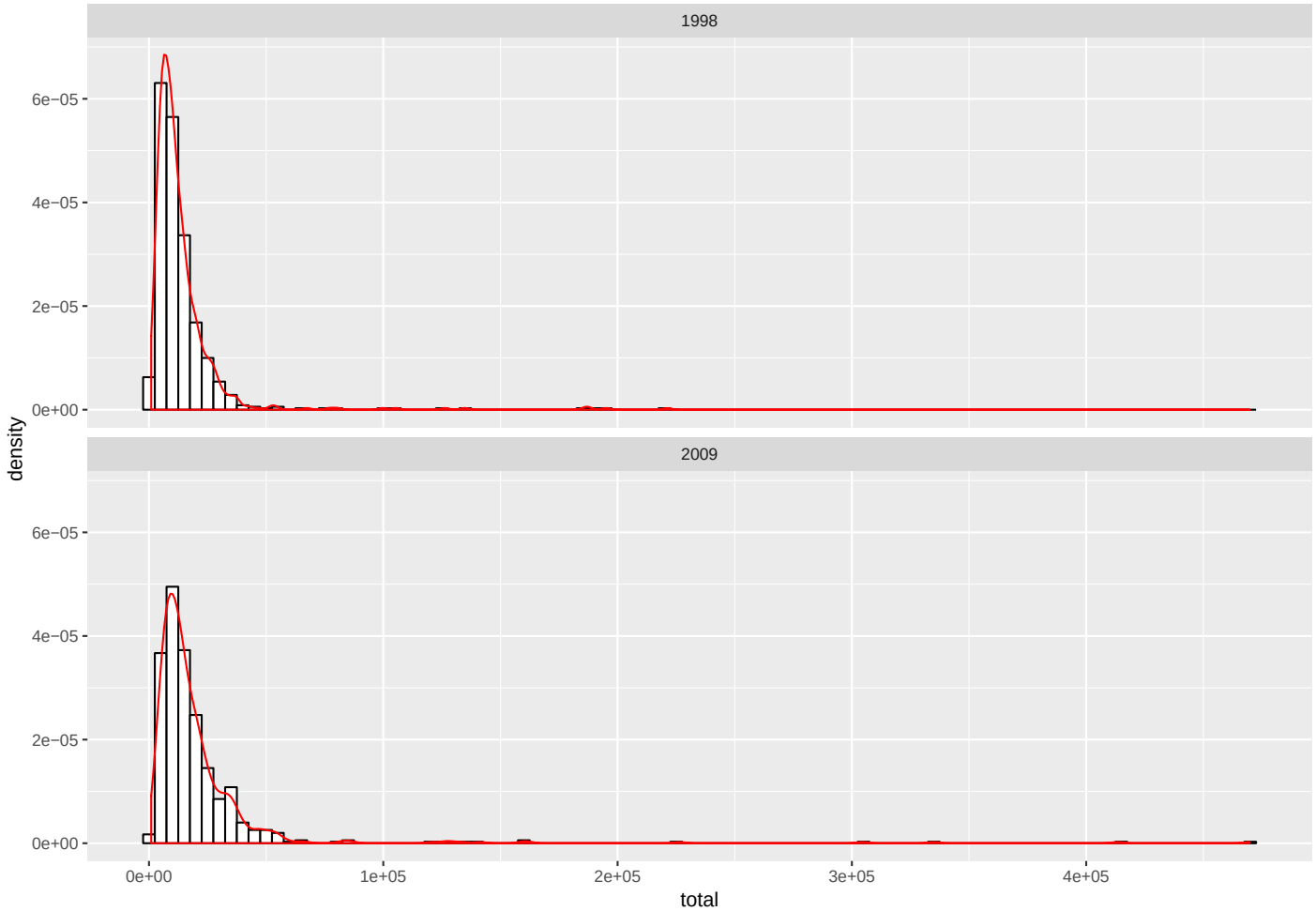


Illustration

Maintenant, combinons les informations des deux années pour avoir une lecture temporelle.

Illustration

```
ggplot(data = adm3_pop_rgph, mapping = aes(x = total)) +
  geom_histogram(mapping = aes(y =..density..), binwidth = 5000,color = "black", fill = "white") +
  geom_density(color = "red") +
  facet_wrap(facets = ~ annee, nrow = 2)
```

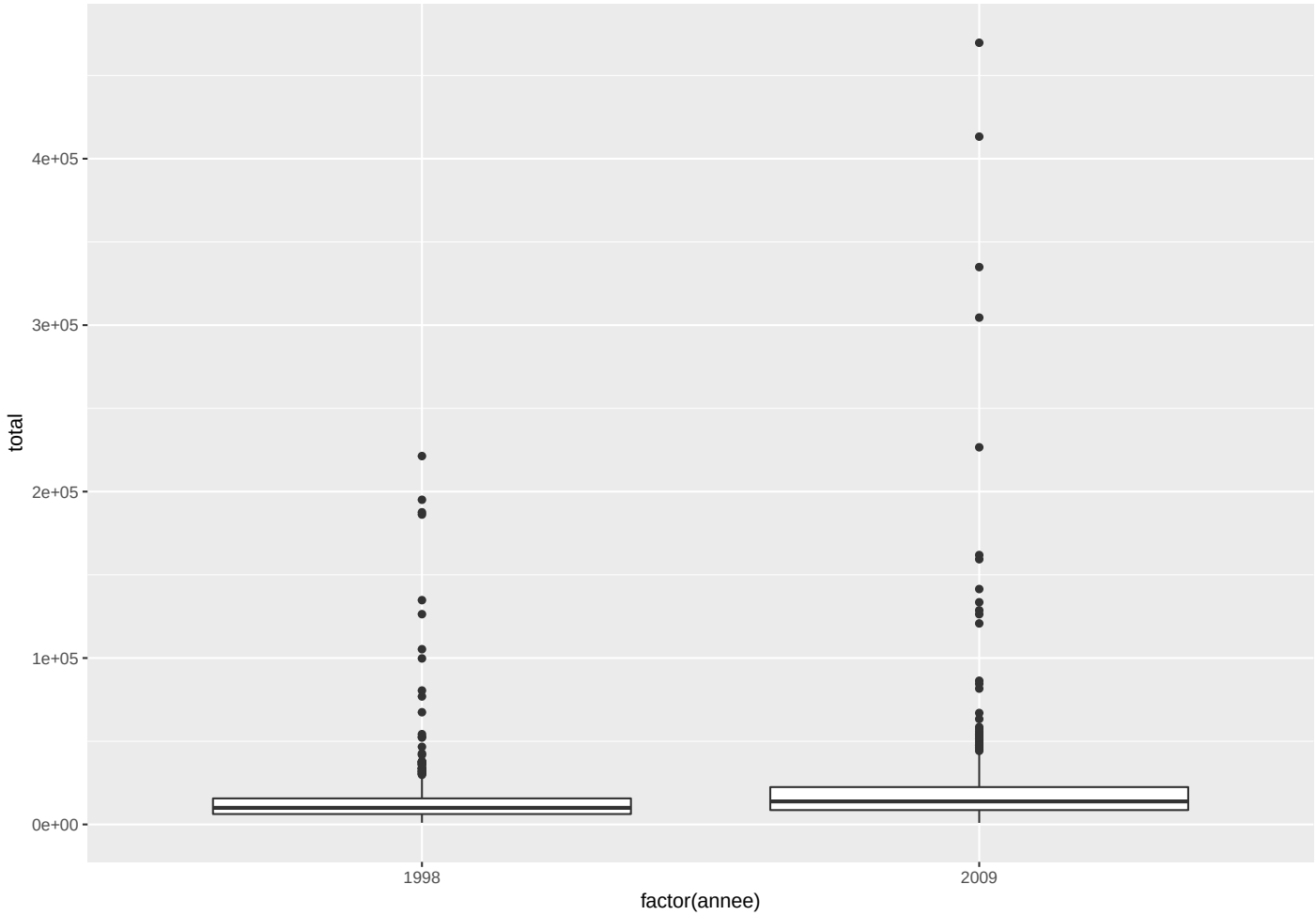


Illustration

Une alternative aux histogrammes est le boxplot. Qu'est-ce qu'un boxplot? Une simple représentation des informations majeures d'une distribution: la médiane (le 2ème quartile), les premier et troisième quartiles, ainsi que les valeurs abérantes.

Illustration

```
ggplot(data = adm3_pop_rgph, mapping = aes(x = factor(annee), y = total)) +
  geom_boxplot()
```



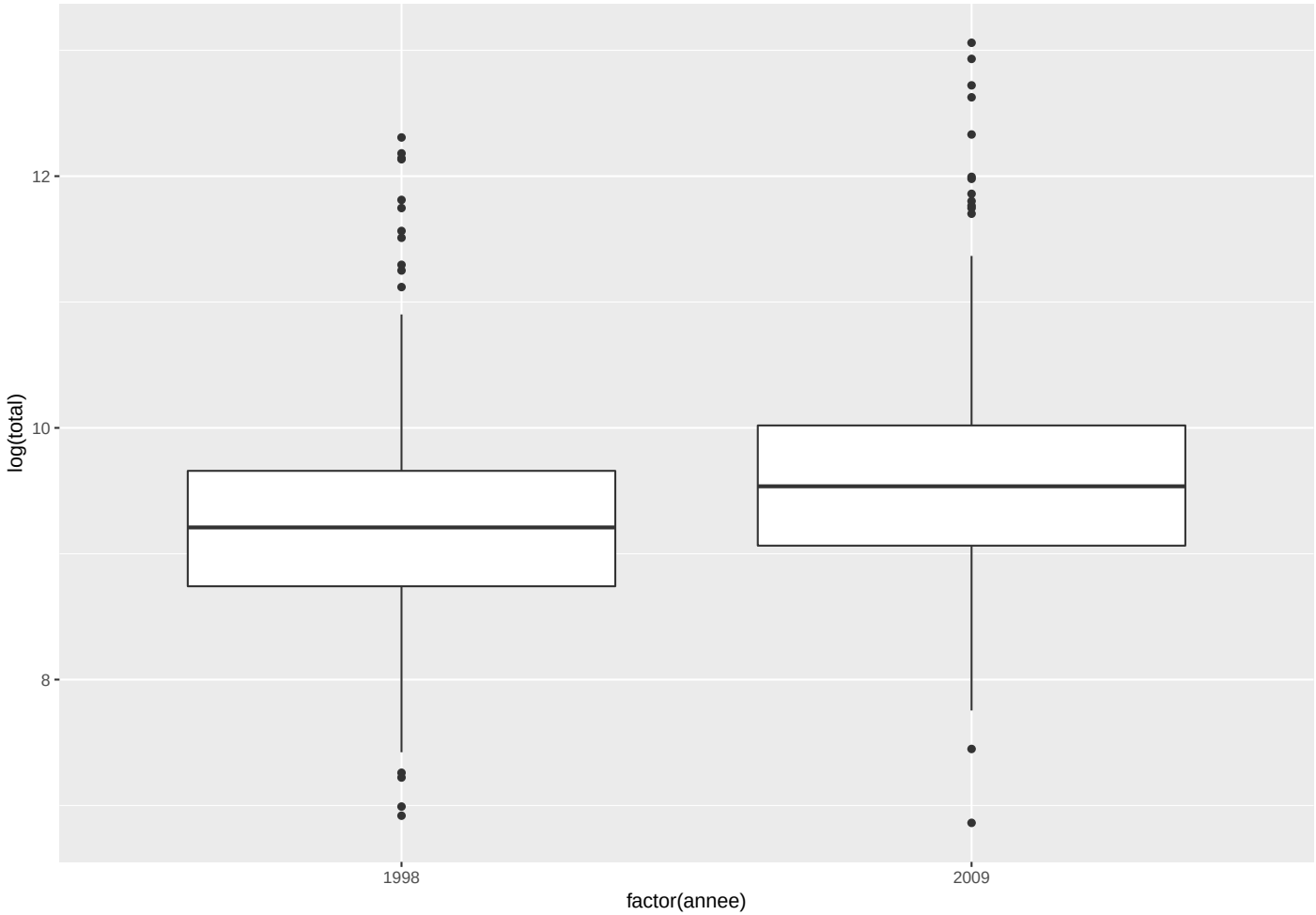
Illustration

Sur la base de ce premier jet, l'on peut retenir qu'en 1998, il n'y avait aucune commune au Mali dont la population dépassait 250000 habitants. Toutefois, en 2009, ce seuil a été dépassé et l'on comptait des communes qui allaient jusqu'à 400000 habitants. Au-delà de cette information, il est difficile de percevoir les disparités entre les communes d'une année à l'autre. Ceci est dû au fait que la distribution est concentrée en bas. Pour améliorer la visibilité, diverses techniques sont possibles.

En premier lieu, l'on peut procéder à la transformation logarithmique des valeurs...

Illustration

```
ggplot(data = adm3_pop_rgph, mapping = aes(x = factor(annee), y = log(total))) +
  geom_boxplot()
```



Illustration

...ou zoomer sur une partie du graphique sans altérer l'échelle de représentation des valeurs.

Supposons que nous souhaitions zoomer sur les communes ayant moins de 50000 habitants, en nous disant que la majorité est en dessous de ce seuil Pour mieux guider le choix, l'on peut calculer des statistiques sommaires.

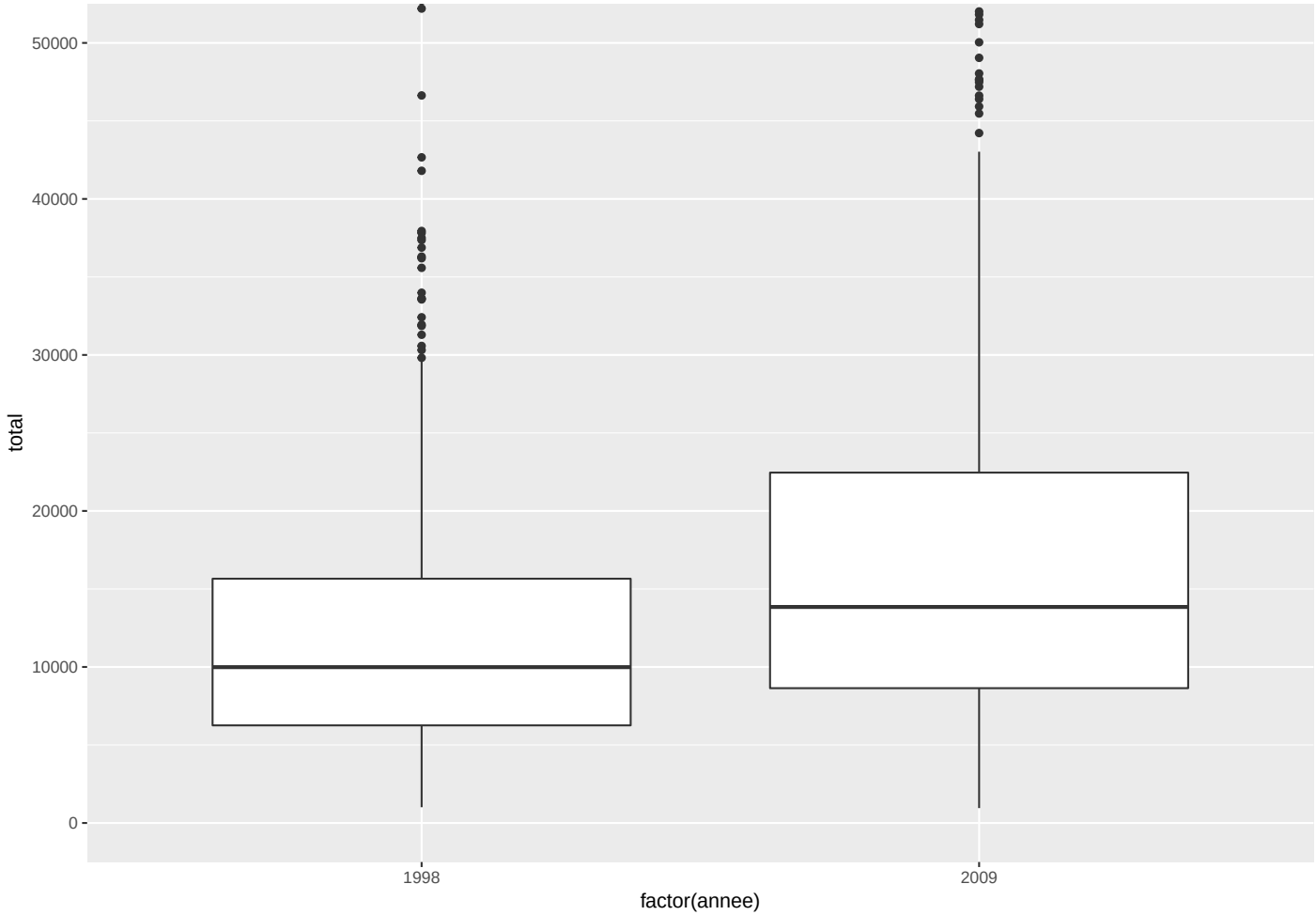
Table 4: Population des communes au Mali: quelques statistiques

Année	Moyenne	Quartile 1	Médiane	Quartile 3
1998	13855.34	6257.0	9988	15657.0
2009	20666.66	8641.5	13847	22458.5

Sur la base des résultats, le seuil choisi semble raisonnable: 75% (voire plus) des communes ont moins de 50 000 habitants. Maintenant, visualisons!

Illustration

```
ggplot(data = adm3_pop_rgph, mapping = aes(x = factor(annee), y = total)) +
  geom_boxplot() +
  coord_cartesian(ylim = c(0, 50000))
```

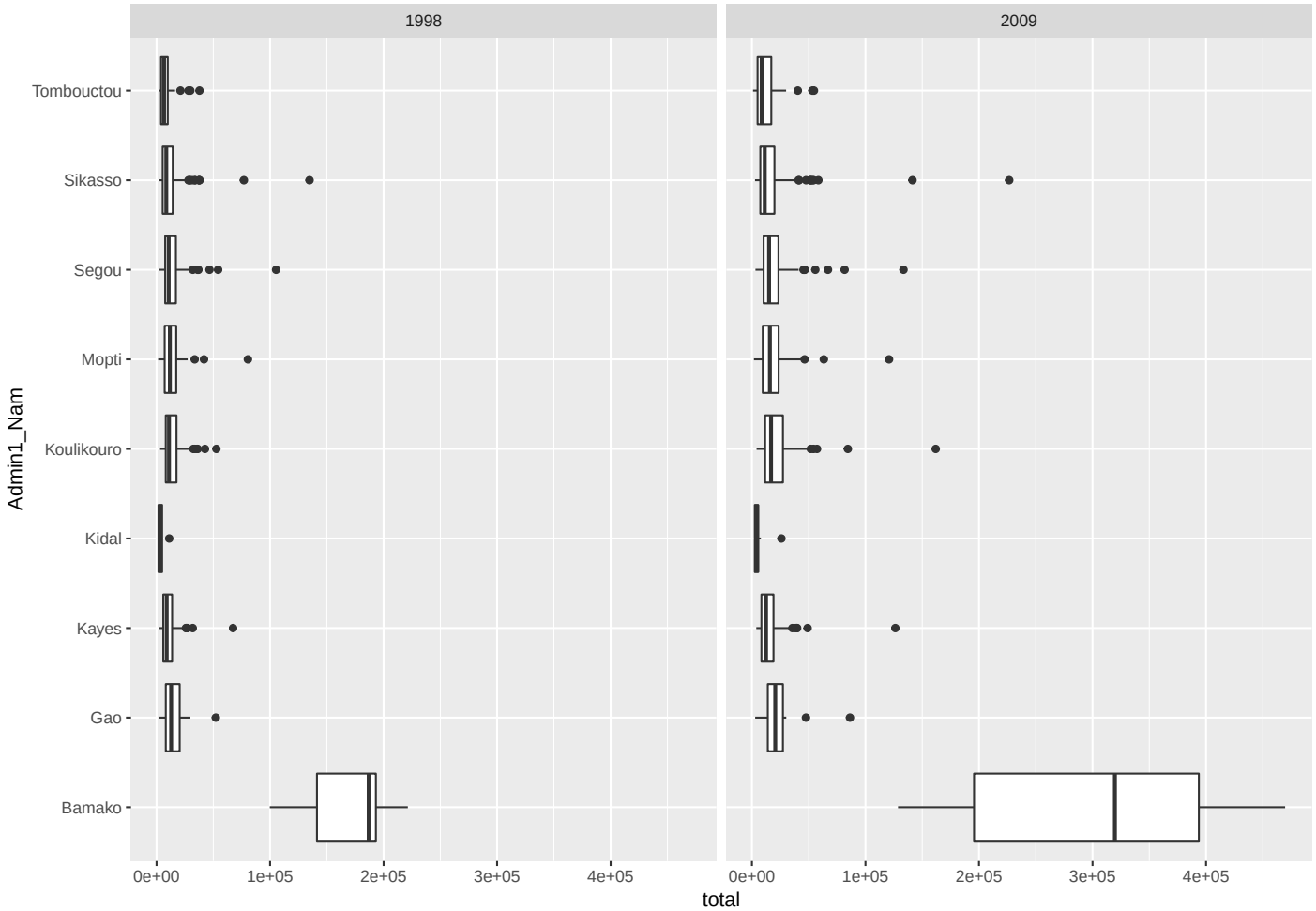


Illustration

Avec le graphique précédent, nous avons nettement améliorer la visibilité de notre graphique. Toutefois, nous avons focalisé notre attention sur un groupe - les communes de moins de 50000 habitants - sans chercher à savoir quels sont les entités qui en font partie et celles qui en sont exclues. Posons-nous la question suivante: où se trouve les communes de plus de 50000 habitants? Nous pouvons explorer cette question en profondeur en regardant la distribution de la population des communes par région.

Illustration

```
ggplot(data = adm3_pop_rgph, mapping = aes(x = Admin1_Nam, y = total)) +
  geom_boxplot() +
  coord_flip() +
  facet_wrap(facets = ~annee)
```

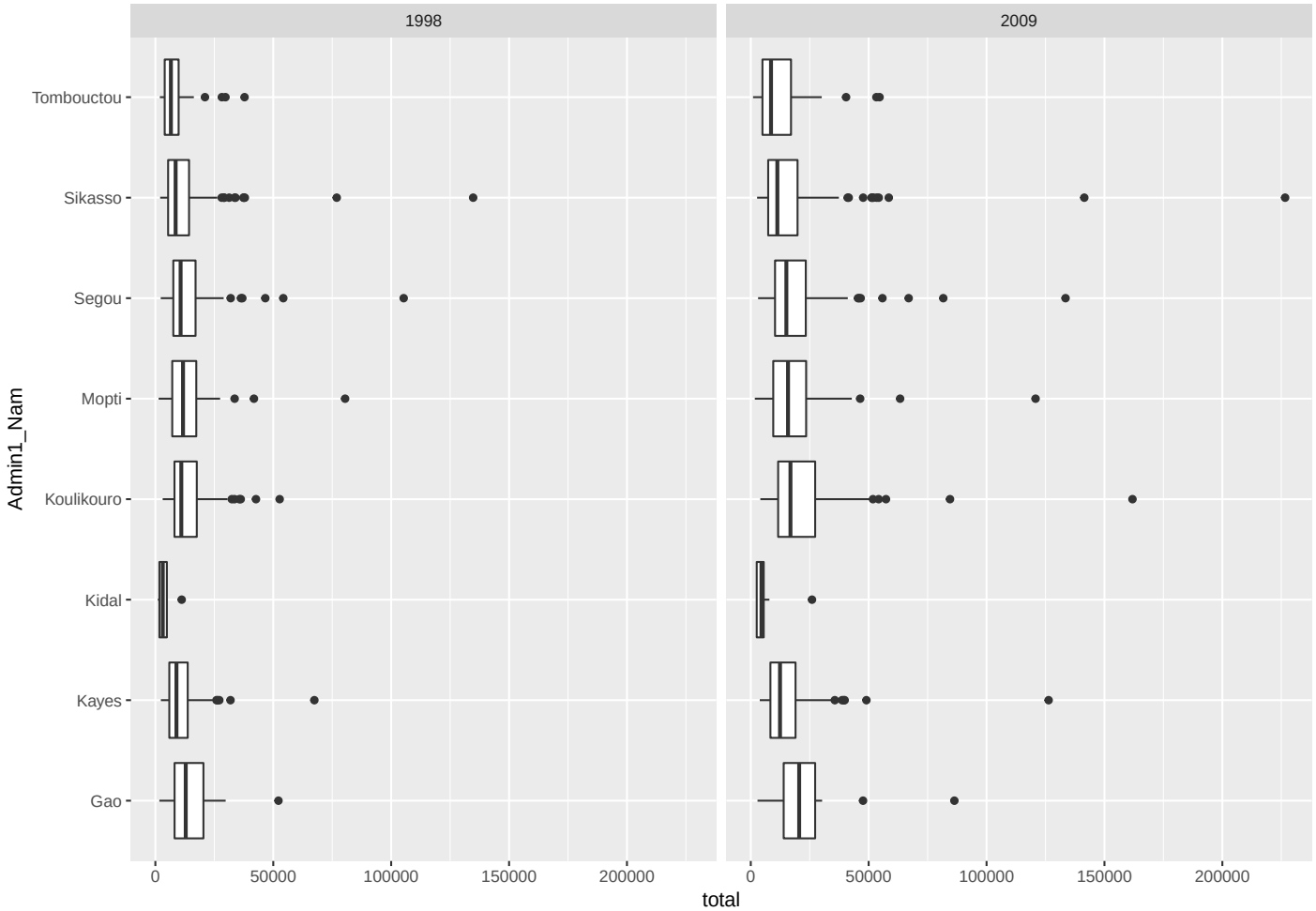


Illustration

Il apparait que les communes les plus peuplés sont en réalité dans le District de Bamako. Sur la base de cette information, l'on peut choisir de traiter Bamako à part et de faire un graphique qui serait focalisé sur les autres régions.

Illustration

```
ggplot(data = adm3_pop_rgph %>% filter(!(Admin1_Nam == "Bamako")), mapping = aes(x = Admin1_Nam, y = total)) +
  geom_boxplot() +
  coord_flip() +
  facet_wrap(facets = ~annee)
```



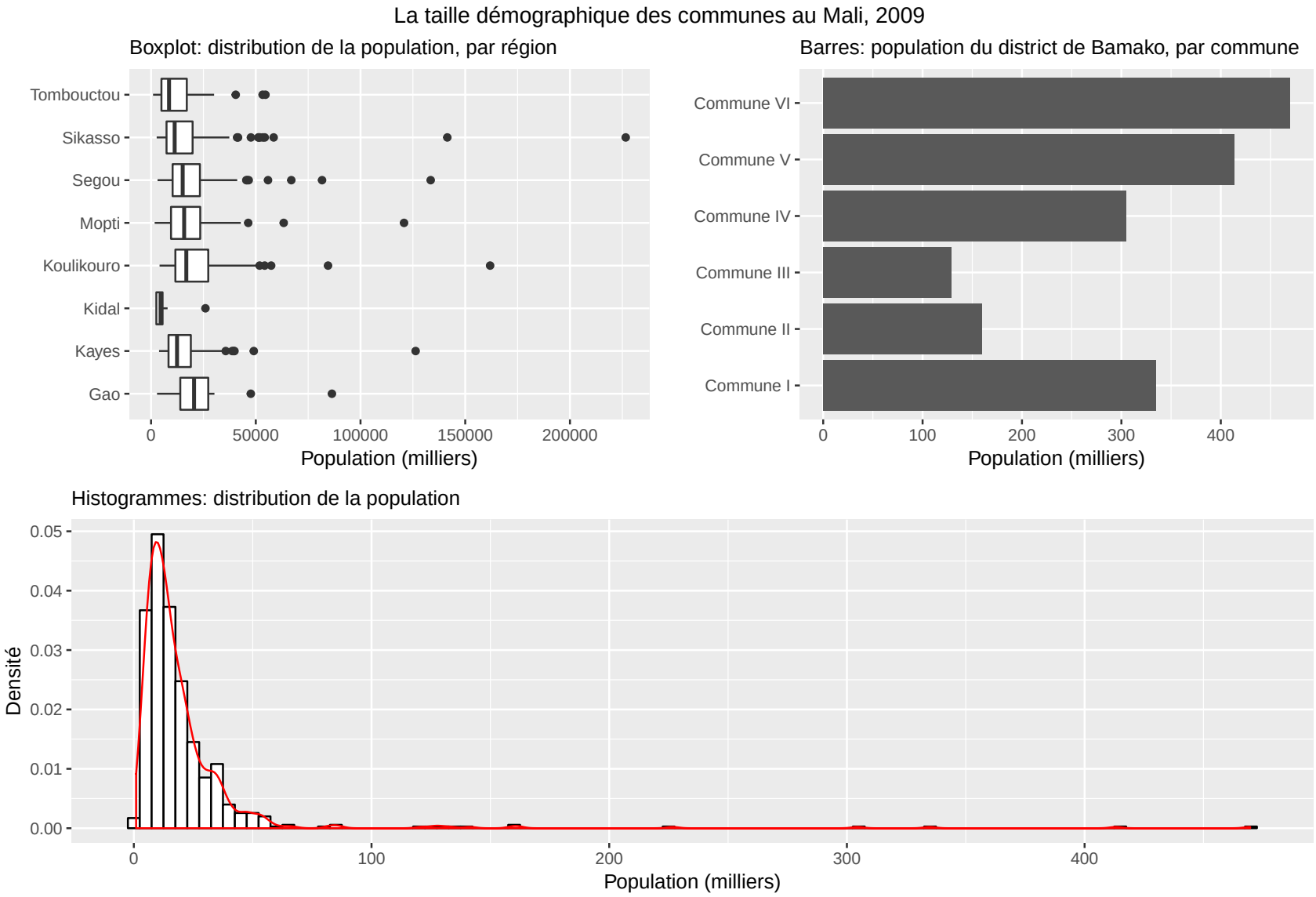
A vous de jouer

A partir des tableau et graphiques suivants, narrez ce que vous observez.

Table 5: Population des communes au Mali en 2009: quelques statistiques

Région	Moyenne	Quartile 1	Médiane	Quartile 3
Bamako	301727.667	195651.50	319710.5	393673.25
Gao	22596.000	13958.75	20537.5	27308.75
Kayes	15605.209	8339.00	12411.0	18994.00
Kidal	6158.091	2532.00	4549.0	5505.00
Koulikouro	22400.111	11613.75	16841.5	27314.25
Mopti	18878.287	9536.75	15807.5	23486.00
Segou	19675.008	10294.25	15102.0	23334.75
Sikasso	17989.510	7392.00	11276.0	19839.50
Tombouctou	12903.942	4955.50	8594.5	17069.00

A vous de jouer



Source: A partir de données des RGPH 199 et 2009