

describes non-data ink. Design elements!

The plotting space you are using

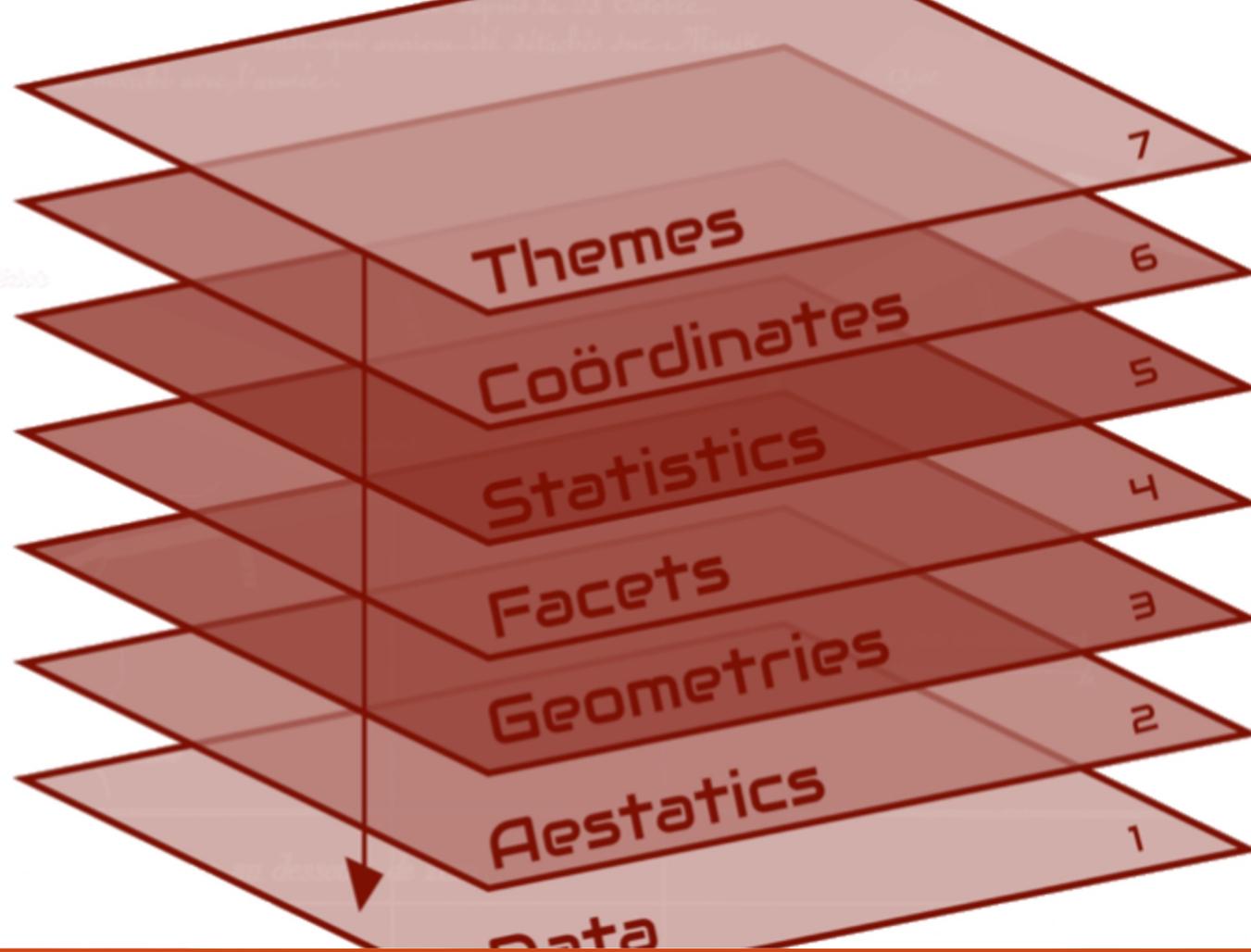
Statistical models & summaries

Rows and columns of sub-plots

Shapes used to represent your data

scales on which the data is mapped

The actual variables to be plotted



10. La grammaire des graphiques

La grammaire des graphiques

C'est une chose de reconnaître quand les graphiques sont **efficaces** et que leur esthétique les rend **faciles à lire**, et quand ils sont disposés dans un tableau de bord qui raconte une histoire visuelle **convaincante** (et quand ils ne le sont pas).

C'est une toute autre chose d'apprendre à **construire** de tels graphiques.

La **grammaire des graphiques** [Wilkinson, 1999 ; Wickham, 2009] fournit une voie fiable pour y parvenir.

La grammaire des graphiques

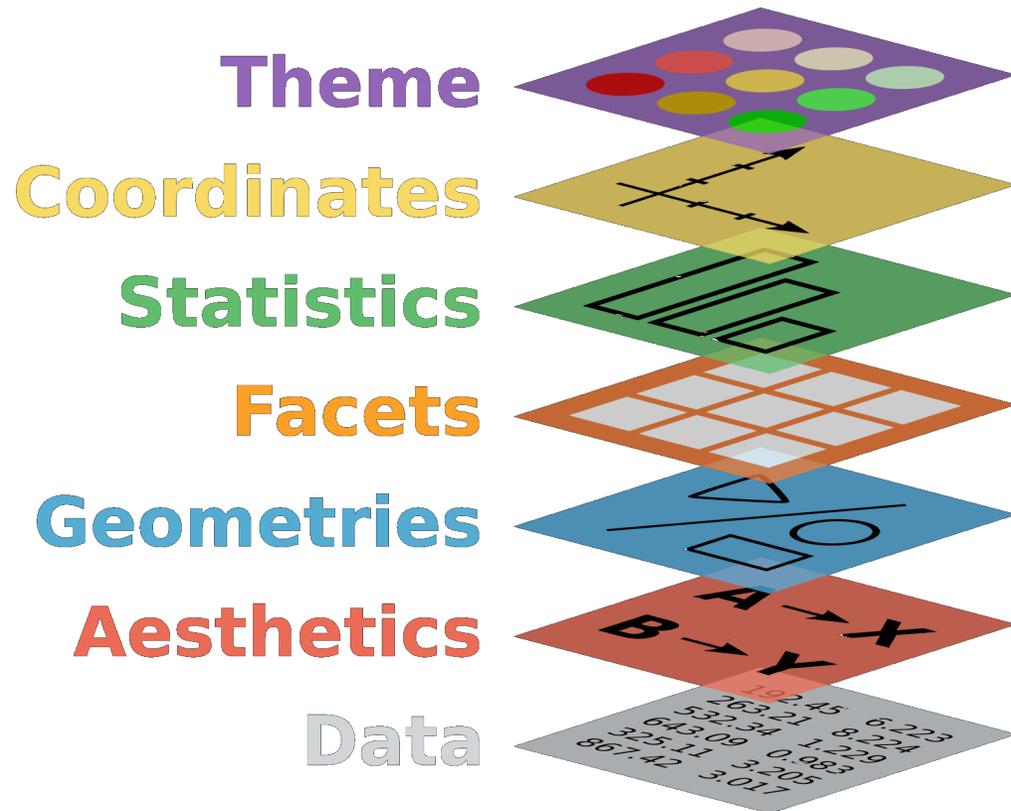
"Une **grammaire** est un ensemble de règles structurelles qui permet de définir et d'établir les **composantes d'une langue**.

Le système/structure d'une langue se compose généralement de la **syntaxe** (règles) et de la **sémantique** (signification).

Une grammaire graphique est un cadre qui suit une approche **en couches** pour décrire et construire des visualisations ou des graphiques de manière **structurée**.

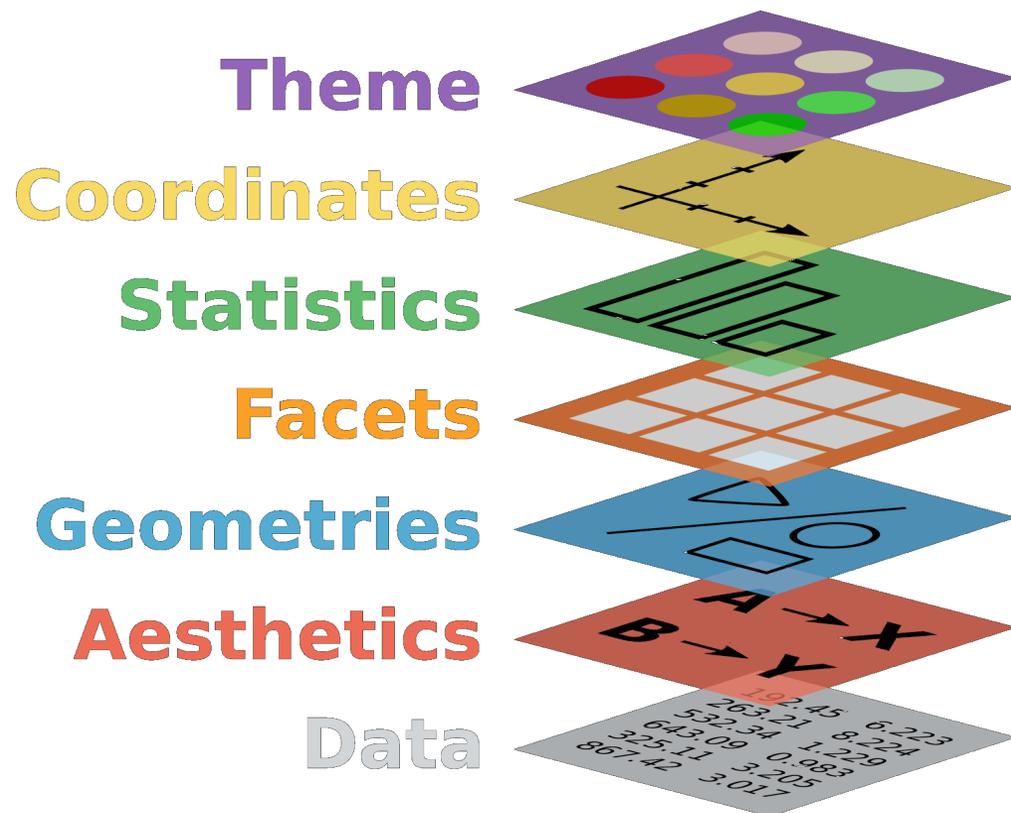
La grammaire **stratifiée** des graphiques utilise des composantes prédéfinies pour construire des diagrammes (au lieu d'essais et d'erreurs aléatoires)."

La grammaire stratifiées des graphiques



- 1. Données** (obligatoire) : les observations se retrouvent en lignes, les variables en colonnes.
- 2. Esthétique** (obligatoire) : le lien explicite de correspondance des variables aux échelles du graphique (position, forme, taille, couleur, etc.)
- 3. Géométrie** (obligatoire) : le type de graphique sur lequel les données sont représentées (barres, lignes, points, etc.)
- 4. Facettes** (facultatif) : les sous-ensembles de données représentés sur le graphique (niveaux)

La grammaire stratifiées des graphiques



5. **Statistiques** (facultatif) : les mesures qui pourraient fournir un contexte au graphique (centralité, dispersion, tendance, etc.)
6. **Coordonnées** (obligatoire) : l'espace de traçage du graphique (axes, échelle, etc.)
7. **Thèmes** (obligatoire) : les choix de design/ conception utilisés pour créer une identité visuelle (polices, couleurs, etc.)

Exemples – Gapminder

L'ensemble de données Gapminder (<https://gapminder.org>) contient des informations socio-démographiques (plus de 500 variables) au sujet des nations de la Terre, pour des années allant de 1800 à 2020.

Nous déconstruisons 8 graphiques construits à partir de cet ensemble de données en utilisant la grammaire stratifiée des graphiques.



Données : Gapminder, 2009

Géométrie : diagramme de densité empilé

Esthétique :

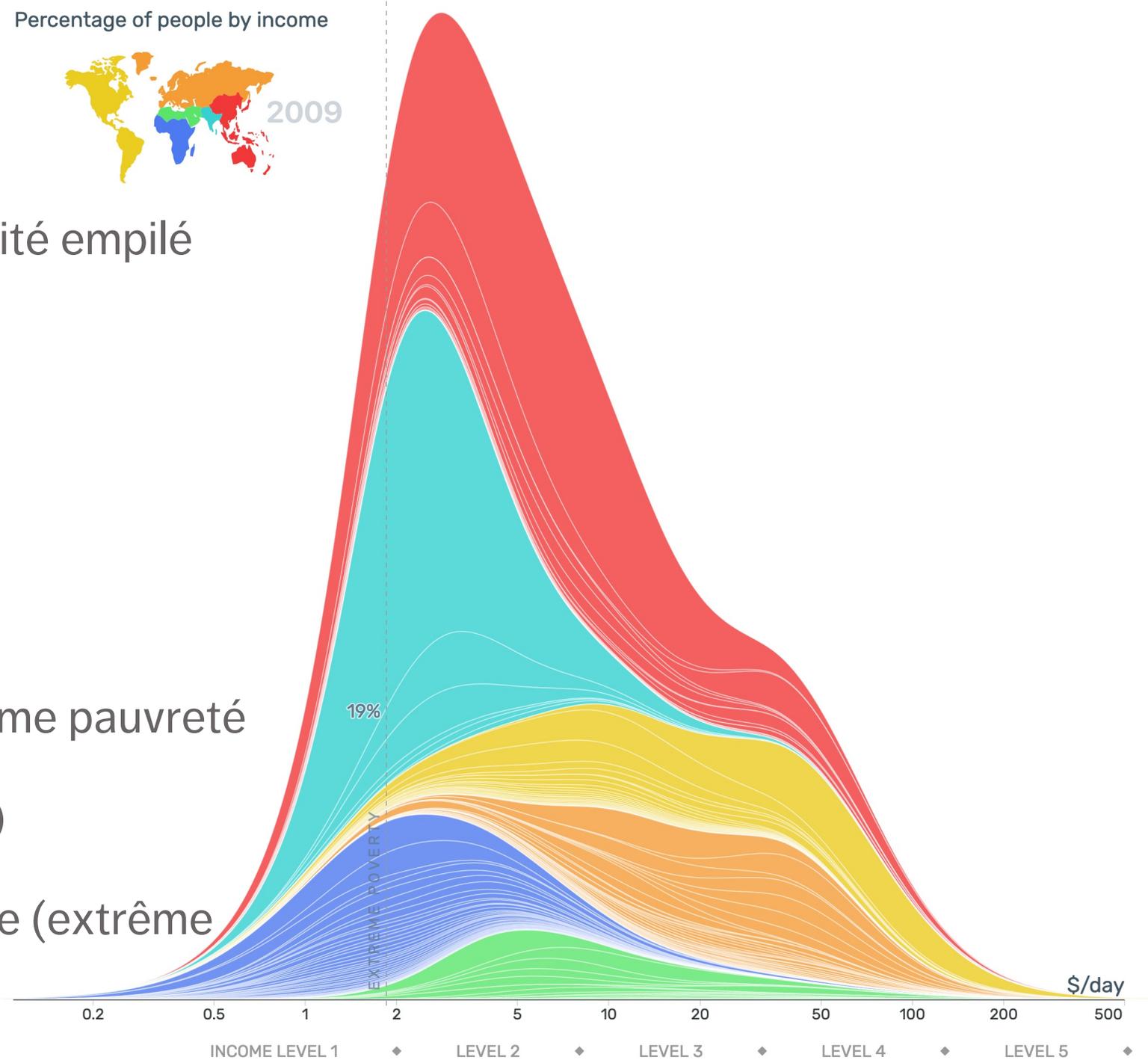
- x : revenu journalier
- (y : pourcentage par pays)
- “fill” : région

Facettes : aucune

Statistiques : proportion d’extrême pauvreté

Coordonnées : logarithmique (x)

Thème : Gapminder Tools; parure (extrême pauvreté)



Données : Gapminder, 2009

Géométrie : diagramme à bulles

Esthétique :

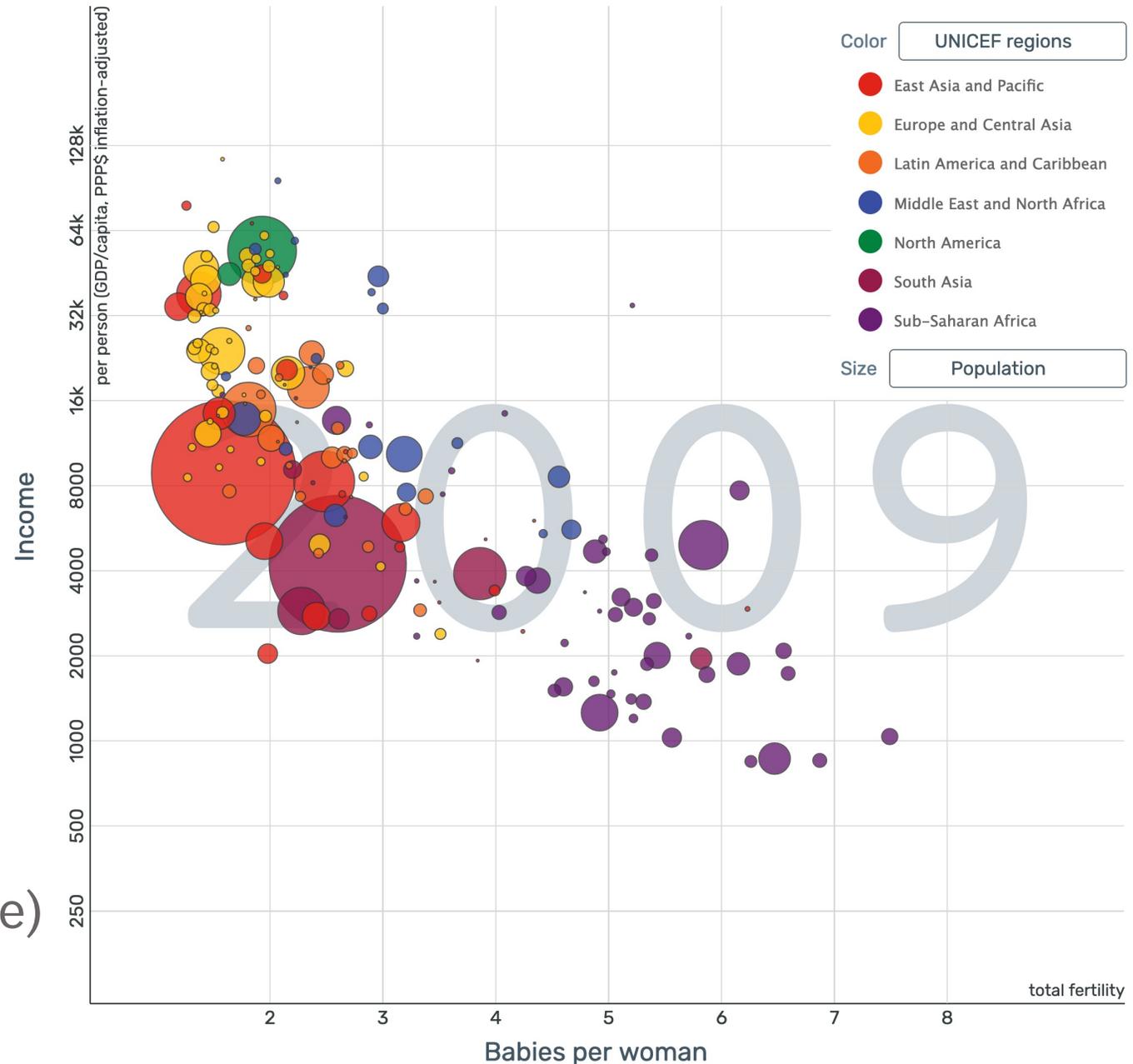
- x : fécondité totale
- y : revenu par personne
- “fill” : région de l'UNICEF
- taille : population

Facettes : aucune

Statistiques : aucune

Coordonnées : logarithmique (x, y, taille)

Thème : Gapminder Tools



Données : Gapminder, 2009

Géométrie : diagramme de dispersion

Esthétique :

- x : fécondité totale
- y : espérance de vie

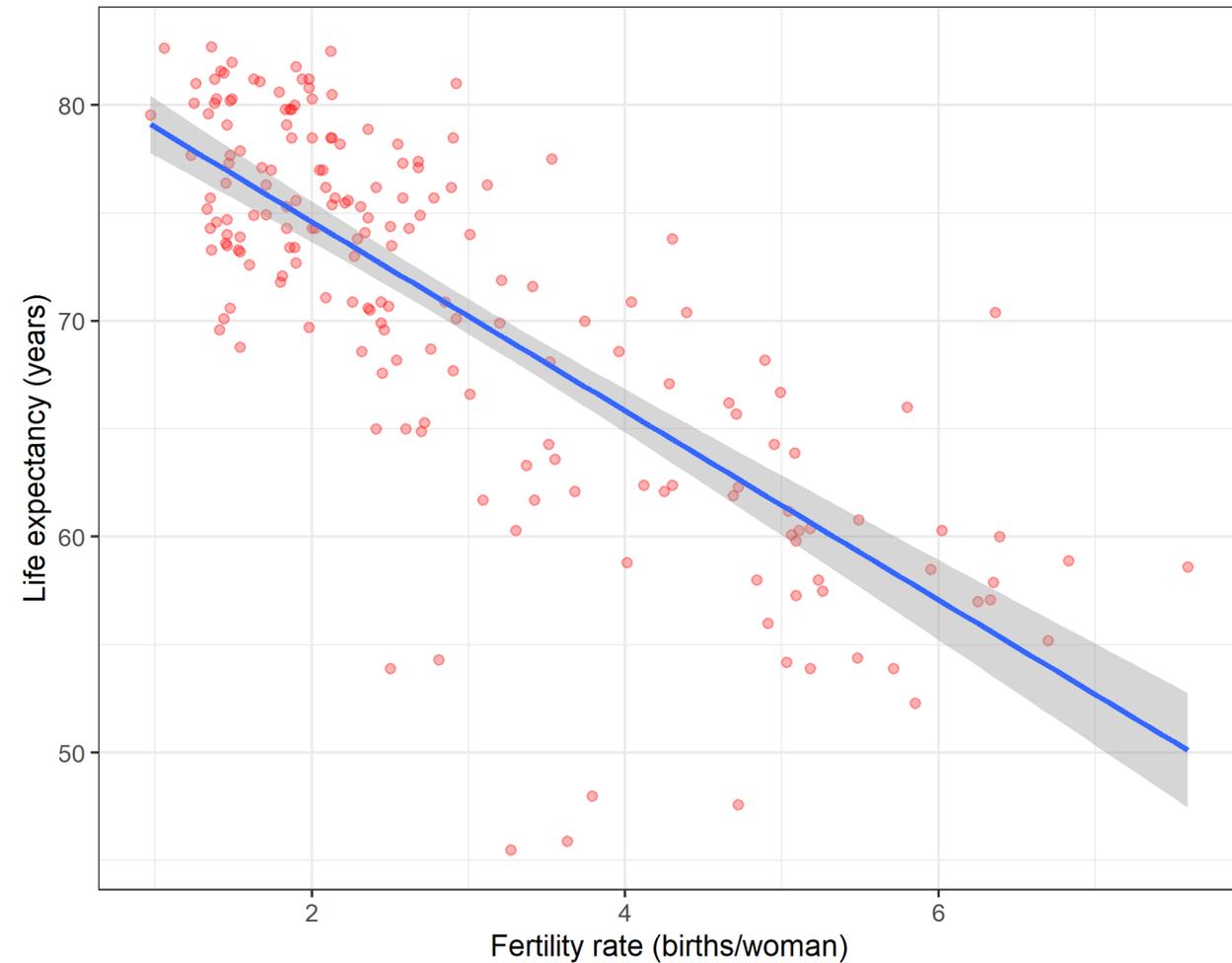
Facettes : aucune

Statistiques : ligne de meilleur ajustement, intervalle de confiance

Coordonnées : linéaire (x, y)

Thème : ggplot2, par défaut

Fertility and Life Expectancy
Gapminder (2009)



Données : Gapminder, 2009

Facettes : continent

Géométrie : diagramme de densité

Statistiques : aucune

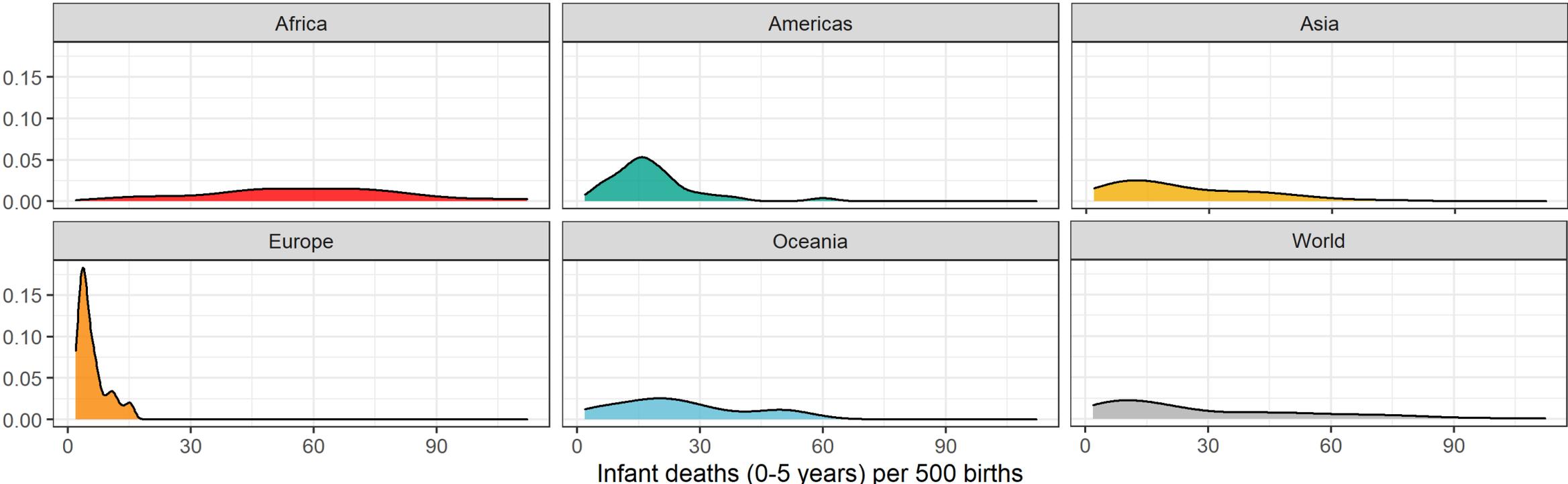
Esthétique :

Coordonnées : linéaire (x)

- x : mortalité infantile
- "fill" : continent

Thème : Darjeeling1

Infant Mortality by Continent Gapminder (2009)



Données : Gapminder, 2009

Géométrie : diagramme à bulles

Esthétique :

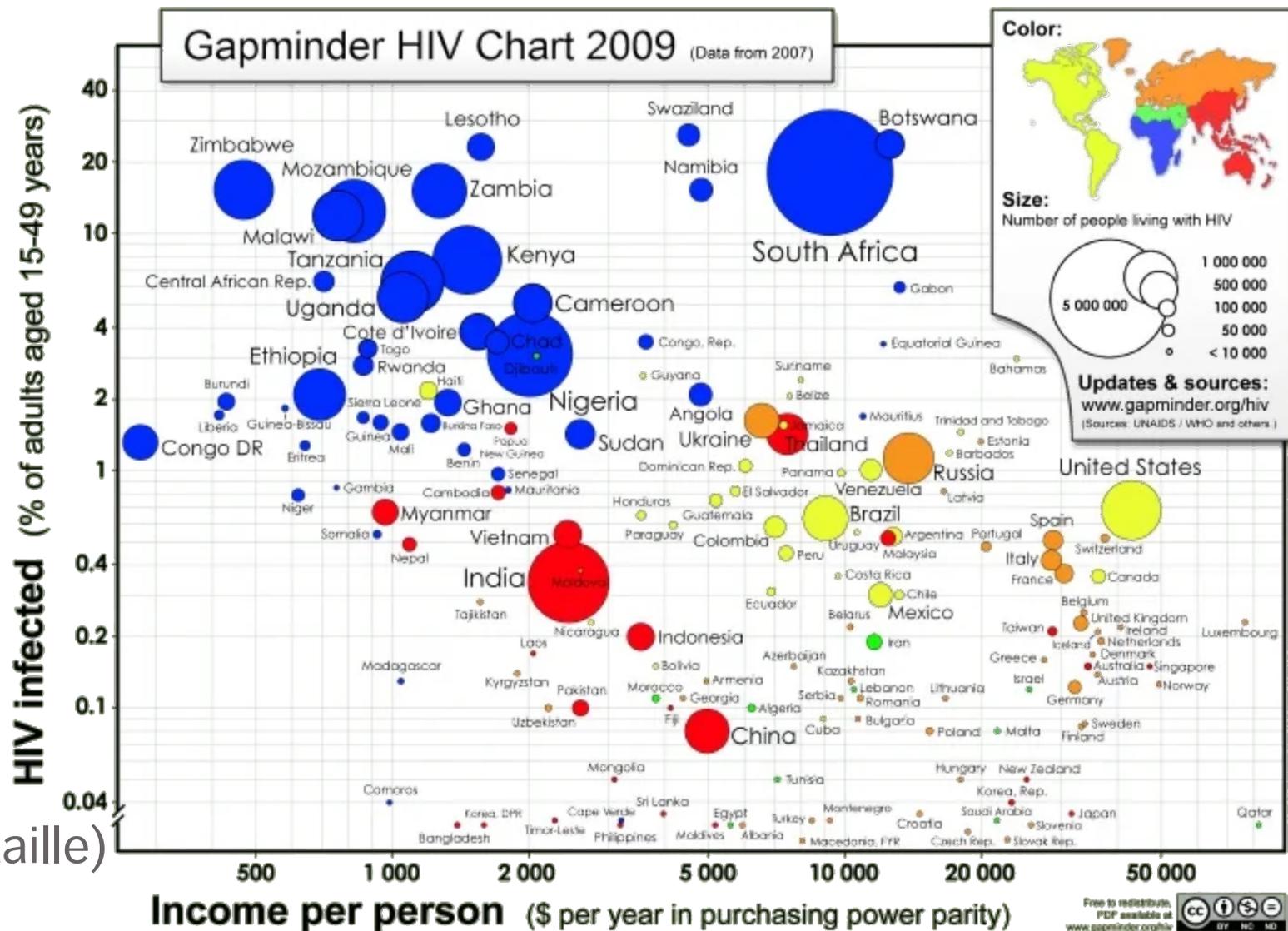
- x : revenu par personne
- y : taux d'infection par le VIH
- "fill" : région de l'OMS
- taille : population infectée par le VIH

Facettes : aucune

Statistiques : aucune

Coordonnées : logarithmique (x, y, taille)

Thème : ancien Gapminder Tools



Données : Gapminder, 2009

Géométrie : graphique boxplot

Esthétique :

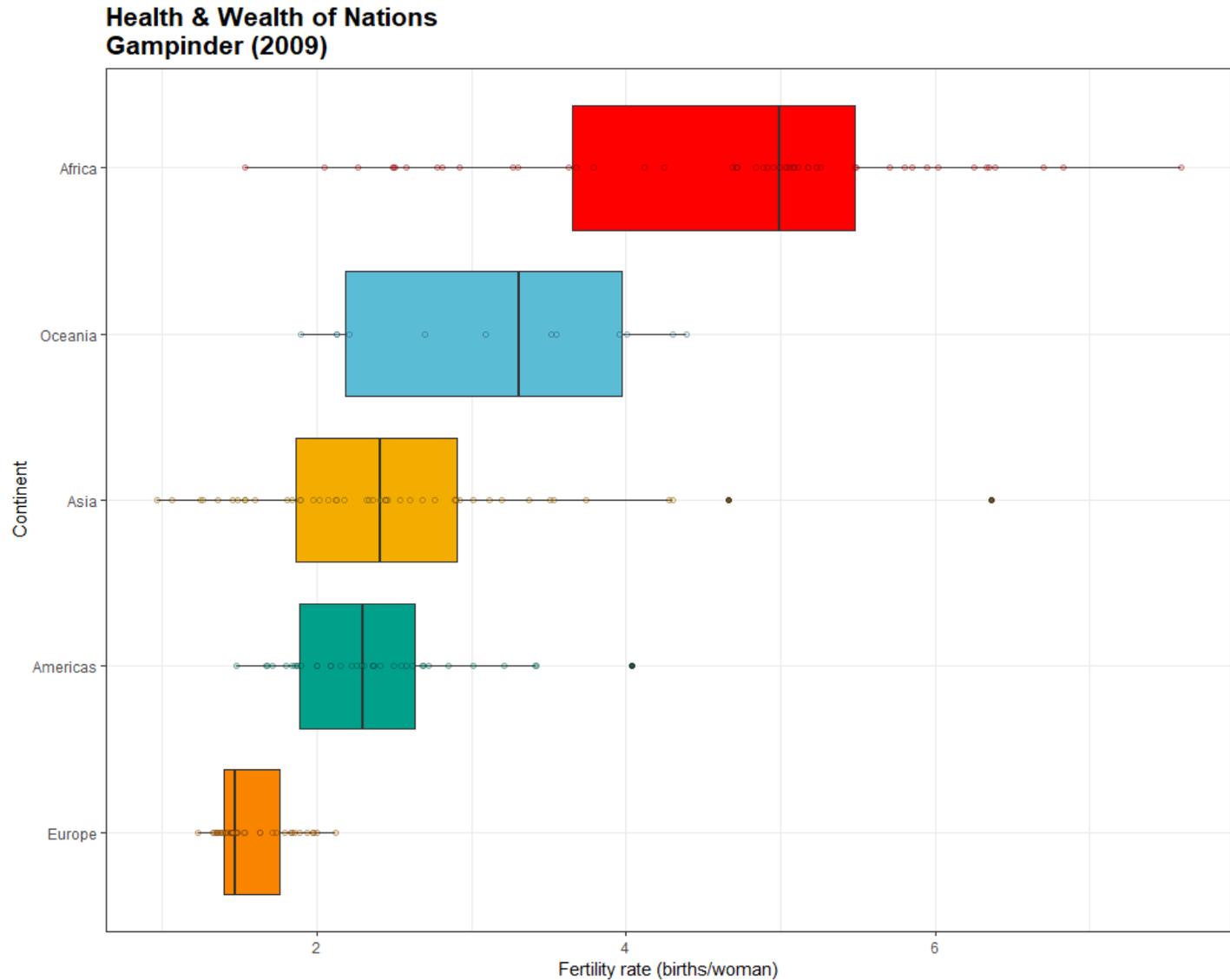
- x : taux de fécondité
- y : continent
- "fill" : continent

"Facettes" : continent

Statistiques : résumé à 5 points

Coordonnées : linéaire (x)

Thème : Darjeeling1



Données : Gapminder, 1960-2011

Géométrie : diagramme en lignes

Esthétique :

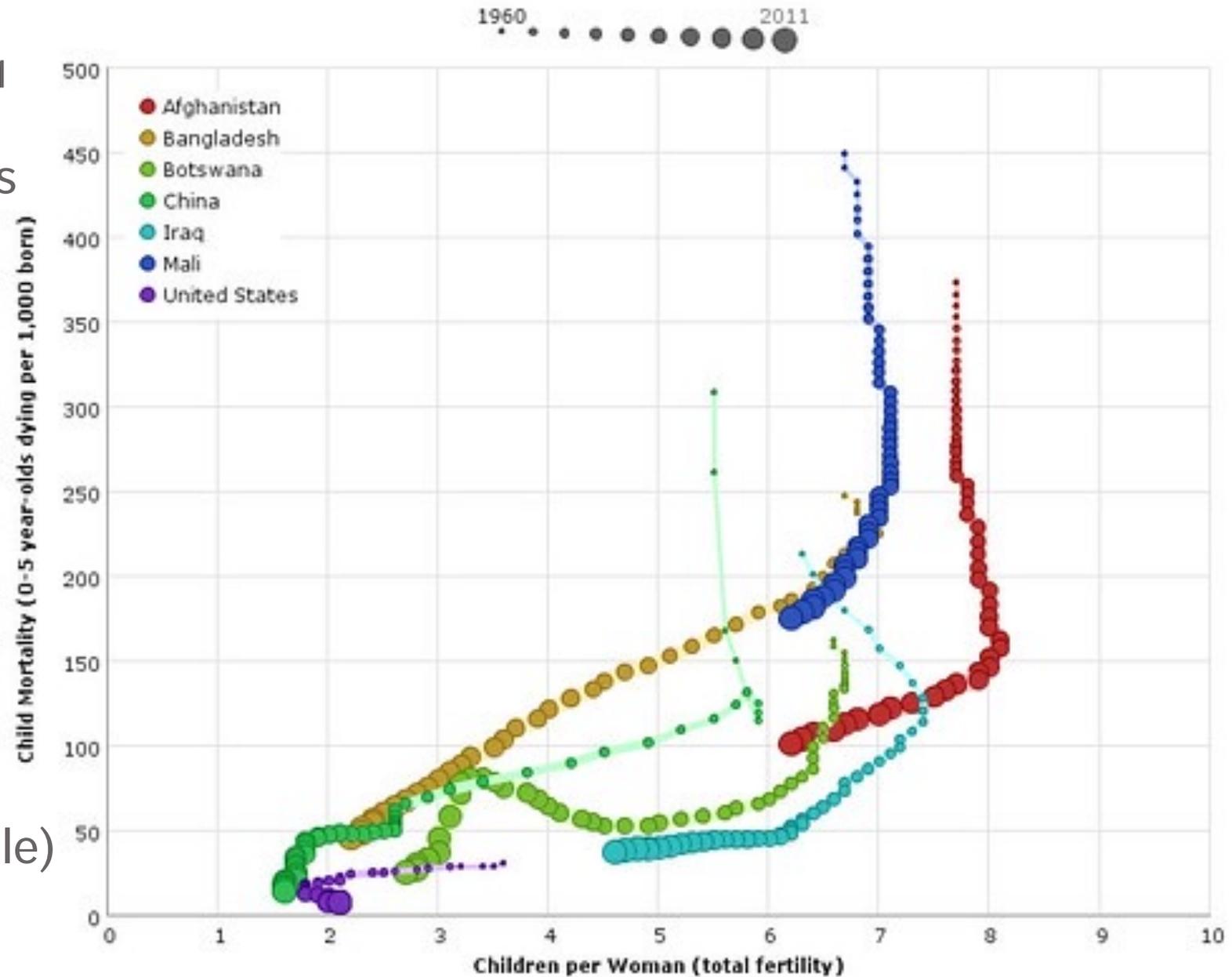
- x : fécondité totale
- y : pourcentage par pays
- couleur : pays
- taille : année

Facettes : aucune

Statistiques : aucune

Coordonnées : linéaires (x, y, taille)

Thème : personnalisé



Données : Gapminder, 2012

Géométrie : diagramme à bulles

Esthétique :

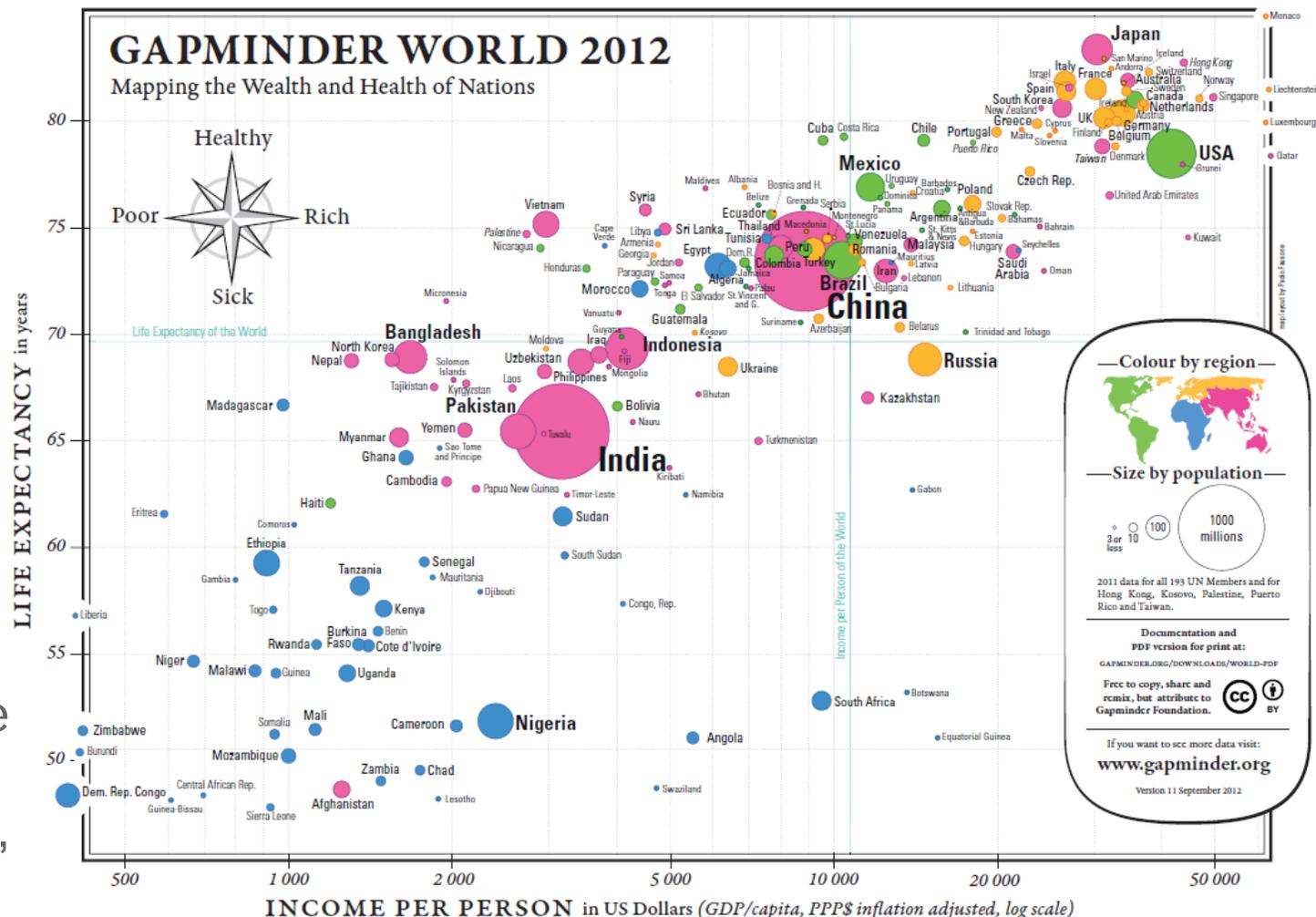
- x : revenu par personne
- y : espérance de vie
- "fill" : région
- taille : population

Facettes : aucune

Statistiques : espérance de vie mondiale, revenu mondial par personne

Coordonnées : logarithmique (x, taille), linéaire (y)

Thème : ancient Gapminder Tools



Lectures suggérées

La grammaire des graphiques

The Practice of Data Visualization
Essentials of Visual Design

Visual Design

- The Grammar of Graphics

Exercices

La grammaire des graphiques

Déconstruisez les graphiques introduits dans les 9 premiers modules en termes de la grammaire des graphiques.

Qu'ont en commun certains des graphiques les plus efficaces ? Qu'en est-il de ceux qui sont les moins efficaces ?

Cela suggère-t-il une stratégie ?