EXO1 : Enseigner la statistique avec R

AFROMPA’R Bouaké 2025

Claude Grasland

2025-03-28

.

On se propoe de montrer l’intérêt de R pour l’enseignement de la statistique aux étudiants de licence de géographie… Cet exercice correspond typiquement au travail qu’on pourrait donner à des étudiants de licence 2 ou 3 ayant suivi des cours de statistique univariée et bivariée et ayant des connaissances de base en cartographie thématique.

On propose de réaliser une analyse de la distribution de l’Indicateur de Développement humain des régions de Côte d’Ivoire entre 2002 et 2022 à l’aide des données accessibles sur le site [Global Data Lab](https://globaldatalab.org/) . Le programme est rédigé de telle sorte qu’on puisse facilement refaire l’exercice en prenant d’autres dates (ex. Evolution entre 2012 et 2022) ou d’autres pays d’Afrique de l’Ouest.

## (1) PREPARATION DES DONNEES STATISTIQUES

Reading layer `mapreg' from data source   
 `/Users/claudegrasland1/worldregio/afromapr/data/exo\_stats/mapreg.shp'   
 using driver `ESRI Shapefile'  
Simple feature collection with 14 features and 3 fields  
Geometry type: MULTIPOLYGON  
Dimension: XY  
Bounding box: xmin: -8.599302 ymin: 4.361807 xmax: -2.494888 ymax: 10.73664  
Geodetic CRS: WGS 84

| code | pays | nom | pop2002 | hdi2002 | pop2022 | hdi2022 |
| --- | --- | --- | --- | --- | --- | --- |
| CIVr101 | CIV | Abidjan | 2728 | 0.564 | 5984 | 0.569 |
| CIVr102 | CIV | Yamoussoukro | 112 | 0.623 | 245 | 0.619 |
| CIVr103 | CIV | Bas Sassandra | 529 | 0.406 | 1160 | 0.520 |
| CIVr104 | CIV | Comoe | 749 | 0.378 | 1642 | 0.582 |
| CIVr105 | CIV | Denguele | 381 | 0.332 | 836 | 0.440 |
| CIVr106 | CIV | Goh-Djiboua | 686 | 0.424 | 1504 | 0.551 |
| CIVr107 | CIV | Lacs | 1098 | 0.389 | 2409 | 0.542 |
| CIVr108 | CIV | Lagunes | 1023 | 0.456 | 2245 | 0.566 |
| CIVr109 | CIV | Montagnes | 1365 | 0.380 | 2994 | 0.492 |
| CIVr110 | CIV | Sassandra-Marahoue | 1614 | 0.407 | 3541 | 0.512 |
| CIVr111 | CIV | Savanes | 628 | 0.409 | 1376 | 0.471 |
| CIVr112 | CIV | Vallee du Bandama | 520 | 0.443 | 1140 | 0.554 |
| CIVr113 | CIV | Woroba | 574 | 0.199 | 1259 | 0.423 |
| CIVr114 | CIV | Zanzan | 834 | 0.320 | 1829 | 0.491 |

## (2) STATISTIQUE UNIVARIEE

### Paramètres principaux

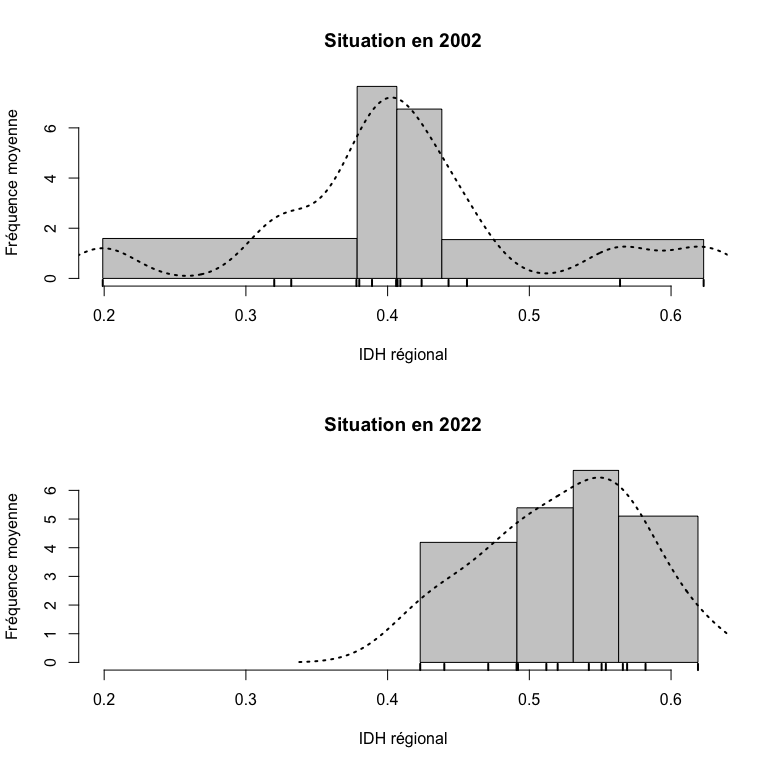
* **Consigne** : Etudiez l’évolution de l’IDH entre 2002 et 2022 en vous servant de paramètres principaux (valeurs centrales, paramètres de dispersion). Puis établissez deux histogrammes permettant de visualiser l’évolution.

Paramètres principaux

|  | Situation en 2002 | Situation en 2022 |
| --- | --- | --- |
| Minimum | 0.20 | 0.42 |
| Q1 | 0.38 | 0.49 |
| Médiane | 0.41 | 0.53 |
| Q3 | 0.44 | 0.56 |
| Maximum | 0.62 | 0.62 |
| Moyenne | 0.41 | 0.52 |
| Ecart-type | 0.10 | 0.06 |
| C.V. (%) | 24.77 | 10.65 |

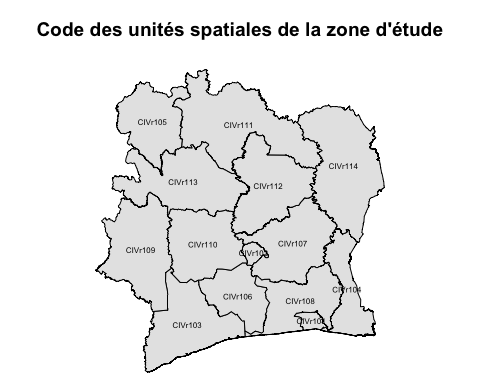
### Histogrammes

* **Consigne** : Etablissez deux histogrammes permettant de visualiser la forme de la distribution de l’IDH en 2002 et en 2022



## (3) PREPARATION DES DONNEES GEOMETRIQUES

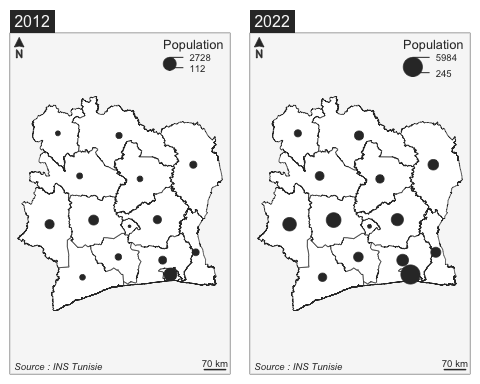
* **Consigne** : Après avoir chargé le fonds de carte affichez le contour des régions et le code des unités.



## (4) CARTOGRAPHIE THEMATIQUE

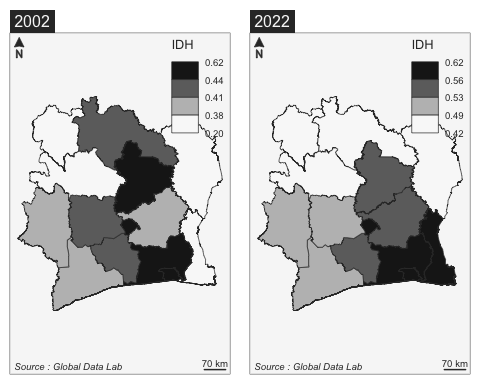
### Cartes de stock

* **Consigne** : Réalisez deux cartes de stock décrivant le nombre d’habitants en 2002 et 2022 Vous utiliserez la même échelle de taille pour rendre les deux cartes comparables.



### Cartes de ratio (choroplèthes)

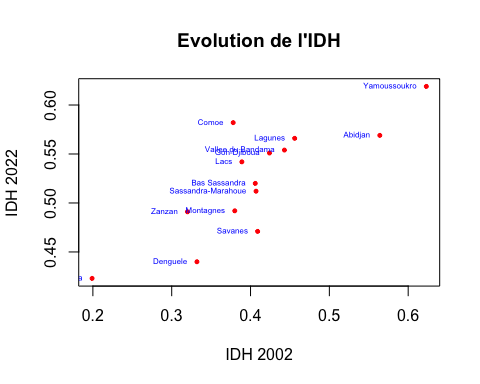
* **Consigne** : Réalisez deux cartes de taux décrivant le niveau de l’IDH en 2002 et 2022. Pour les rendre comparables vous utiliserez dans chaque carte une partition en quartiles (4 classes d’effectifs égaux)



## (5) STATISTIQUES BIVARIEES

### Nuage de points

* **Consigne** : Tracez un nuage de point montrant l’évolution de l’indicateur entre les deux dates.



### Analyse de la corrélation

* **Consigne** : calculez les coefficients de corrélation de Pearson et Spearman et testez leur sgnificativité.

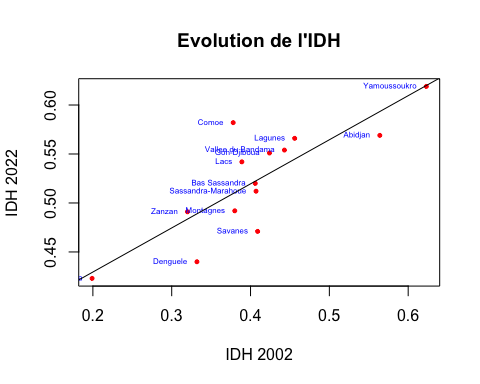
Pearson's product-moment correlation  
  
data: X and Y  
t = 4.9223, df = 12, p-value = 0.0003524  
alternative hypothesis: true correlation is not equal to 0  
95 percent confidence interval:  
 0.5073463 0.9403486  
sample estimates:  
 cor   
0.8177876

Spearman's rank correlation rho  
  
data: X and Y  
S = 134, p-value = 0.006405  
alternative hypothesis: true rho is not equal to 0  
sample estimates:  
 rho   
0.7054945

### Droite de régression

* **Consigne** : calculez l’equation de la droite de régression et tracez- là sur le graphique.

Call:  
lm(formula = Y ~ X)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.052586 -0.016570 -0.001574 0.019274 0.072368   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 0.33948 0.03848 8.823 1.36e-06 \*\*\*  
X 0.45013 0.09145 4.922 0.000352 \*\*\*  
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.03342 on 12 degrees of freedom  
Multiple R-squared: 0.6688, Adjusted R-squared: 0.6412   
F-statistic: 24.23 on 1 and 12 DF, p-value: 0.0003524



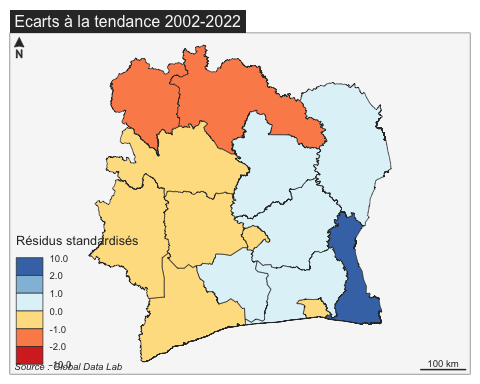
### Analyse des résidus

* **Consigne** : Calculez les valeurs théoriques prévus par le modèle de régression et les résidus. Affichez le tableau correspondant après l’avoir trié par ordre de résidus croissants.

|  | code | nom | X | Y | Y\_est | Y\_res |
| --- | --- | --- | --- | --- | --- | --- |
| 11 | CIVr111 | Savanes | 0.409 | 0.471 | 0.524 | -0.053 |
| 5 | CIVr105 | Denguele | 0.332 | 0.440 | 0.489 | -0.049 |
| 1 | CIVr101 | Abidjan | 0.564 | 0.569 | 0.593 | -0.024 |
| 9 | CIVr109 | Montagnes | 0.380 | 0.492 | 0.511 | -0.019 |
| 10 | CIVr110 | Sassandra-Marahoue | 0.407 | 0.512 | 0.523 | -0.011 |
| 13 | CIVr113 | Woroba | 0.199 | 0.423 | 0.429 | -0.006 |
| 3 | CIVr103 | Bas Sassandra | 0.406 | 0.520 | 0.522 | -0.002 |
| 2 | CIVr102 | Yamoussoukro | 0.623 | 0.619 | 0.620 | -0.001 |
| 14 | CIVr114 | Zanzan | 0.320 | 0.491 | 0.484 | 0.007 |
| 12 | CIVr112 | Vallee du Bandama | 0.443 | 0.554 | 0.539 | 0.015 |
| 6 | CIVr106 | Goh-Djiboua | 0.424 | 0.551 | 0.530 | 0.021 |
| 8 | CIVr108 | Lagunes | 0.456 | 0.566 | 0.545 | 0.021 |
| 7 | CIVr107 | Lacs | 0.389 | 0.542 | 0.515 | 0.027 |
| 4 | CIVr104 | Comoe | 0.378 | 0.582 | 0.510 | 0.072 |

### Cartographie des résidus

* **Consigne** : Cartographiez les résidus après les avoir standardisés.

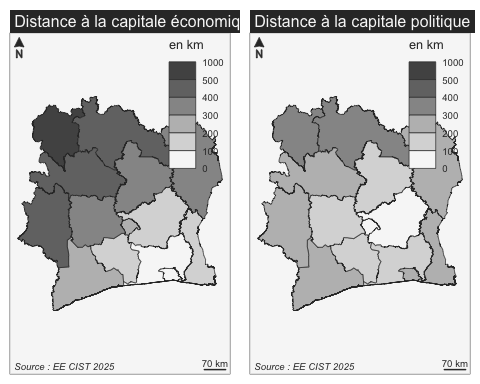


## (6) ANALYSE SPATIALE

On se propose d’analyser l’effet de la distance à la capitale économique (Abidjan) ou politique (Yamoussoukro) sur le niveau de développement en 2022.

### Distance aux deux capitales.

* **Consigne** : ajoutez au tableau de données une colonne correspondant à la distance en km à Abidjan ou Yamoussoukro et faites en une cartographie en prenant comme bornes de classes 0, 100, 200, 300, 400, 500, 1000 km.



### Relation entre développement et distance à la capitale

* **Consigne** : Déterminez le modèle qui décrit le mieux la relation entre l’IDH (Y) et la distance à la capitale (X), selon que l’on choisit la distance à la capitale économique (X1) ou la distance à la capitale politique (X2).

