MAT7381 Exercices
Chapitre 4 Régression linéaire simple

Consigne concernant tous les numéros appliqués : *Toute analyse doit être comprise comme une réponse à une question concrète posée par un client dont les connaissances statistiques sont limitées. Vous devez donc lui expliquer le sens de vos résultats numériques, en des termes spécifiques au contexte et dans le langage de l'homme de la rue. Évitez (dans la mesure du possible) le langage technique.*

4.1 Les données suivantes, du Ministère de l'agriculture, des pêcheries et de l'alimentation, portent sur une expérience visant à déterminer la relation entre le taux de semis d'avoine (x, en kg/ha) et le nombre de plantules par mètre carré (y). Les données portent sur 17 parcelles de terre. Faites les exercices suivants à l'aide d'une calculatrice et non d'un ordinateur avec logiciel.

x	90	95	95	95	95	100	105	106	108	113	113	114	114	120	120	126	133
y	232	208	220	150	203	252	232	298	279	237	210	211	201	304	275	284	313

a) Faites un nuage de points
b) Déterminer la droite des moindres carrés
c) Tester l'hypothèse qu'il n'y a pas de relation entre x et y.
d) Déterminer un intervalle de confiance à 90% pour le nombre moyen de plantules obtenus avec un taux de semis de 90 kg/ha, de 100 kg/ha et de 150 kg/ha.
e) Déterminer des limites de prédiction à 90% pour le nombre de plantules obtenus avec un taux de semis de 90 kg/ha, de 100 kg/ha et de 150 kg/ha.
f) Dresser une table d'analyse de variance.

4.2 Les données dans le tableau 1 (ci-dessous) portent sur un certain nombre d'indiens qui ont immigré des montagnes vers les plaines. Le but de cette collecte de données était de savoir si le pouls est affecté par le nombre d'années passées dans les plaines.

a) Faites un nuage de points
b) Estimer les paramètres.
c) Tester l'hypothèse que la droite de régression est horizontale
d) Dresser une table d'analyse de variance.

4.3 Les données $(x_1 ; y_1), \ldots , (x_n ; y_n)$ forment un nuage qu'on pourrait décrire par une droite passant par l'origine. Déterminer l'équation de la droite des moindres carrés $y = bx$.

4.4 Considérer le modèle suivant : $y_1, \ldots , y_n$ sont des variables aléatoires indépendantes de loi normale, de variance $\sigma^2_{y.x}$, et d'espérance $E(y_i) = \beta x_i$, $i = 1, \ldots , n$.

a) Montrer que l'estimateur $\hat{\beta} = \dfrac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$ est sans biais.

b) Déterminer la variance de $\hat{\beta}$.

c) Montrer que $\dfrac{\sum_{i=1}^n (y_i - \hat{\beta} x_i)^2}{\sigma^2}$ est de loi χ^2. Déterminer le nombre de degrés de liberté et le paramètre de non centralité. [*Suggestion* : *Montrez que* $\sum_{i=1}^n (y_i - \hat{\beta} x_i)^2 = \sum_{i=1}^n y_i^2 - \hat{\beta} \sum_{i=1}^n x_i y_i$ *peut s'écrire sous la forme* $\boldsymbol{y}'(\mathbf{I}-\boldsymbol{x}(\boldsymbol{x}'\boldsymbol{x})^{-1}\boldsymbol{x}')\boldsymbol{y}$].

d) Montrer que $\hat{\sigma}^2_{y.x} = \dfrac{\sum_{i=1}^n (y_i - \hat{\beta} x_i)^2}{n-1}$ est un estimateur sans biais de σ^2

e) Montrer que $\hat{\sigma}^2_{y.x}$ est indépendante de $\hat{\beta}$.

f) Déterminer (et justifier) une formule d'un intervalle de confiance pour β.

4.5 Une usine fabrique des toiles métalliques pour des usines de pâtes et papier. Afin de mieux répartir son personnel, le gérant aimerait prévoir le temps, T, requis pour la finition des toiles. Ce temps pourrait être lié, entre autres variables, à la surface de la toile, S. On a obtenu les données du tableau 2 :

a) Quelle variable doit-on utiliser comme variable dépendante ? (Justifier ce choix).
b) Déterminer l'équation de régression correspondante.
c) Quel est le temps moyen de finition pour une toile de 20 m^2 ?
d) La régression est-elle utilisable pour des fins de prévision ? (Faire un test à 5%).
e) Quel est le pourcentage de variation expliquée par la droite de régression ?
f) Faire un graphique des données. Tracer la droite de régression. Le modèle est-il raisonnable ?

4.6 Un professeur de secondaire est responsable de l'enseignement de l'algèbre. Au début de l'année, il fait passer à 20 de ses étudiants un petit test mesurant les habiletés arithmétiques (H) de ses étudiants. À la fin du premier semestre, il examine les résultats (F) de ses étudiants à l'examen d'algèbre. Les résultats sont présentés au tableau 3:

a) Quelle variable doit-on utiliser comme variable dépendante ? (Justifier ce choix).
b) Déterminer l'équation de régression correspondante.
c) Quelle note à l'examen d'algèbre aurait un étudiant dont la note au premier aurait été 25 ?
d) La régression est-elle utilisable pour des fins de prévision ? (Faire un test à 5%).
e) Quel est le pourcentage de variation expliquée par la droite de régression ?
f) Faire un graphique des données. Tracer la droite de régression. Le modèle est-il raisonnable ?

4.7 Soit $a_i = \dfrac{x_i - \bar{x}}{\sum_j (x_j - \bar{x})^2}$, $i = 1, \ldots, n$. Montrer que $\hat{\beta}_1 = \sum_{i=1}^n a_i y_i$, $\sum_i a_i = 0$, $\sum_i a_i x_i = 1$, $\sum_{i=1}^n a_i^2 = \dfrac{1}{\sum (x_i - \bar{x})^2}$

4.8 (*Suite*) Soit $\boldsymbol{a} = [a_1, \ldots, a_n]'$ et $\mathbf{C} = \boldsymbol{I} - \dfrac{\boldsymbol{ee}'}{n}$ où $\boldsymbol{e} = (1, \ldots, 1)'$. Montrer que (i) $\boldsymbol{a} = \dfrac{\boldsymbol{Cx}}{\boldsymbol{x}'\boldsymbol{Cx}}$ et donc que $\hat{\beta}_1 = \dfrac{\boldsymbol{x}'\boldsymbol{Cy}}{\boldsymbol{x}'\boldsymbol{Cx}}$, (ii) $\sum_i (y_i - \bar{y})^2 = \boldsymbol{y}'\mathbf{C}\boldsymbol{y}$, (iii) $\hat{\beta}_1^2 \sum_i (x_i - \bar{x})^2 = \boldsymbol{y}'\boldsymbol{a}(\boldsymbol{a}'\boldsymbol{a})^{-1}\boldsymbol{a}'\boldsymbol{y}$, (iv) la matrice $\mathbf{A} = \mathbf{C} - \boldsymbol{a}(\boldsymbol{a}'\boldsymbol{a})^{-1}\boldsymbol{a}' = \mathbf{C} - \dfrac{\mathbf{C}\boldsymbol{xx}'\mathbf{C}}{\boldsymbol{x}'\mathbf{C}\boldsymbol{x}}$ est idempotente, et (v) $\mathbf{A}\boldsymbol{\mu} = \mathbf{0}$.

4.9 Démontrer la décomposition SCT = SCE + SCR.

4.10 Soit $\hat{\boldsymbol{\varepsilon}}$ le vecteur dont les composantes sont $\hat{\varepsilon}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$.

a) Montrer que $\hat{\boldsymbol{\varepsilon}}$ peut s'écrire comme $\hat{\boldsymbol{\varepsilon}} = \mathbf{C}\left(\mathbf{I} - \dfrac{\boldsymbol{xx}'}{\boldsymbol{x}'\mathbf{C}\boldsymbol{x}}\right)\mathbf{C}\boldsymbol{y}$, où $\boldsymbol{x} = (x_1, \ldots, x_n)'$ et $\boldsymbol{y} = (y_1, \ldots, y_n)'$.
b) Montrer que la matrice de covariance $\mathbf{V}$ de $\hat{\boldsymbol{\varepsilon}}$ est singulière; montrer en fait qu'elle est de rang n -2.
c) Montrer que $\boldsymbol{x}'\hat{\boldsymbol{\varepsilon}} = 0$ et $\boldsymbol{e}'\hat{\boldsymbol{\varepsilon}} = 0$ (ce sont les deux contraintes qui expliquent le rang).
d) Montrer que $\mathrm{Var}(\hat{\varepsilon}_i) = \sigma_{y.x}^2\left[1 - \dfrac{1}{n} - \dfrac{(x_i - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right]$, et que $\mathrm{Cov}(\hat{\varepsilon}_i, \hat{\varepsilon}_j) = -\sigma_{y.x}^2\left[\dfrac{1}{n} + \dfrac{(x_i - \bar{x})(x_j - \bar{x})}{\sum (x_i - \bar{x})^2}\right]$.

4.11 Considérons une nouvelle observation y qui correspond à une valeur x de la variable indépendante. Montrer que si y satisfait les conditions du modèle [$y \sim \mathcal{N}(\beta_o + \beta_1 x,\ \sigma_{y.x}^2)$], alors la statistique $\dfrac{y - \hat{y}}{\hat{\sigma}_{y.x}\sqrt{1 + \dfrac{1}{n} + \dfrac{(x - \bar{x})^2}{n\sigma_x^2}}} \sim t_{n-2}$, où $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$. On utilise cette statistique pour vérifier si une donnée particulière satisfait les conditions du modèle.

4.12 Montrer que SCT = $\boldsymbol{y}'\mathbf{C}\boldsymbol{y}$, SCE = $\dfrac{\boldsymbol{y}'\mathbf{C}\boldsymbol{xx}'\mathbf{C}\boldsymbol{y}}{\boldsymbol{x}'\mathbf{C}\boldsymbol{x}}$ et SCR = $\boldsymbol{y}'\left[\mathbf{C} - \dfrac{\mathbf{C}\boldsymbol{xx}'\mathbf{C}}{\boldsymbol{x}'\mathbf{C}\boldsymbol{x}}\right]\boldsymbol{y}$.

4.13 Montrer que $\dfrac{\text{SCR}}{\sigma_{y.x}^2} \sim \chi_{n-2}^2$ *centrale*, et de là vérifier que $\hat{\sigma}_{y.x}^2$ = MCR est un estimateur sans biais de $\sigma_{y.x}^2$.

4.14 Montrer que $\frac{\text{SCE}}{\sigma^2_{y.x}} = \frac{1}{\sigma^2_{y.x}}\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 \sim \chi^2_1$ avec paramètre de non centralité $\lambda = \beta_1^2\sum(x_i - \bar{x})^2 / \sigma^2_{y.x}$.

4.15 Montrer que SCR et SCE sont indépendantes de $\bar{y}$ et que si $\beta_1 = 0$, alors la statistique $F = \frac{\text{SCE}}{\text{SCR}/(n-2)} = \frac{\text{MCE}}{\text{MCR}} =$ suit une loi $\mathcal{F}_{1,n\text{-}2}$.

4.16 Montrer que r^2 = SCE/SCT.

4.17 Vérifier que si on soustrait $\bar{y}$ de chaque côté de l'équation de la droite des moindres carrés $\hat{\boldsymbol{y}} = \hat{\beta}_o\boldsymbol{e} + \hat{\beta}_1\boldsymbol{x}$, et puis on divise par S_y, on obtient l'équation $(y\text{-}\ \bar{y}\boldsymbol{e})/S_y = r(x\text{-}\ \bar{x}\boldsymbol{e})/S_x$; en d'autres termes, le coefficient de corrélation r est la pente de la droite des moindres carrés calculée à partir des données centrées-réduites [c.à.d. on remplace les y_i par $(y_i - \bar{y})/S_y$ et les x_i par $(x_i - \bar{x})/S_x$].

4.18 Soit $\boldsymbol{y} = (y_1, \ldots, y_n)'$ et $\boldsymbol{x} = (x_1, \ldots, x_n)'$, $\mathbf{C} = \mathbf{I} - (1/n)\boldsymbol{ee}'$ où $\boldsymbol{e} = (1, \ldots, 1)'$, d'ordre n, $\boldsymbol{\beta} = \begin{bmatrix}\beta_o\\ \beta_1\end{bmatrix}$, $\hat{\boldsymbol{\beta}} = \begin{bmatrix}\hat{\beta}_o\\ \hat{\beta}_1\end{bmatrix}$ et $\mathbf{X} = [\boldsymbol{e} \,|\, \boldsymbol{x}]$.

a) Vérifier que $r = \frac{\boldsymbol{x}'\mathbf{C}\boldsymbol{y}}{\sqrt{\boldsymbol{x}'\mathbf{C}\boldsymbol{x}.\boldsymbol{y}'\mathbf{C}\boldsymbol{y}}}$ **.** [*Si on définit le "cosinus" entre deux vecteurs* ***u*** *et* ***v*** *par cos*[$\boldsymbol{u}$; $\boldsymbol{v}$] = $\frac{\boldsymbol{u}'\boldsymbol{v}}{\sqrt{\boldsymbol{u}'\boldsymbol{u}.\boldsymbol{v}'\boldsymbol{v}}}$, alors r est le cosinus entre la projection de $\boldsymbol{y}$ et la projection de $\boldsymbol{x}$ sur $[\mathcal{C}(\boldsymbol{e})]^{\perp}$].

b) Montrer que le vecteur des valeurs prédites $\hat{\boldsymbol{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \left[\frac{\boldsymbol{ee}'}{n} + \frac{\mathbf{C}\boldsymbol{xx}'\mathbf{C}}{\boldsymbol{x}'\mathbf{C}\boldsymbol{x}}\right]\boldsymbol{y}$ **.**

c) Montrer que $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \frac{\boldsymbol{ee}'}{n} + \frac{\mathbf{C}\boldsymbol{xx}'\mathbf{C}}{\boldsymbol{x}'\mathbf{C}\boldsymbol{x}}$ et donc que $\hat{\boldsymbol{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ est le projection orthogonale de $\boldsymbol{y}$ sur le champs de **X**. [On peut montrer ce résultat sans faire de calculs].

d) Montrer que le coefficient de corrélation entre $\boldsymbol{x}$ et $\boldsymbol{y}$ est égal au coefficient de corrélation entre $\boldsymbol{y}$ et $\hat{\boldsymbol{y}}$.

4.19 Les données suivantes présentent le nombre de bactéries N encore vivantes après avoir été exposées à des rayons X pendant un temps de durée t.

N	355	211	197	166	142	106	104	60	56	38	36	32	21	19	15
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

a) Ajuster le modèle $N = \beta_o + \beta_1 t + \varepsilon$. Faire un examen des résidus pour évaluer l'ajustement du modèle.

b) Supposons que le phénomène soit assez bien connu pour savoir que la relation entre N et t est de la forme $N = \beta_o e^{\beta_1|t|}$, d'où $\ln N = \ln \beta_o + \beta_1|t| + \varepsilon$ peut être considéré comme modèle pour ln N. Examiner ce modèle.

4.20 Le tableau ci-dessous présente la teneur en plomb telle que mesurée par le National Bureau of Standards (x_i) et par un laboratoire clinique (y_i) dans un échantillon sur 48 paires de prélèvements de sang humain.

x_i	y_i	x_i	y_i	x_i	y_i	x_i	y_i	x_i	y_i	x_i	y_i	x_i	y_i
15,21	10	17,16	15	19,96	20	21,27	17	22,96	21	24,32	18	27,69	31
15,48	11	18,22	15	19,8	25	21,59	19	23,38	24	24,96	28	30,43	37
15,90	16	18,32	14	20,11	21	21,8	23	23,48	20	25,27	32	32,54	35
16,64	20	19,06	17	20,11	21	22,75	21	23,59	28	25,57	19	32,64	32
16,74	11	19,17	15	20,64	21	22,75	30	23,69	19	26,54	17	33,17	32
16,74	16	19,43	15	20,85	26	22,85	27	23,8	16	26,54	19	33,9	32
		19,43	15	21,27	24	22,85	15	24,8	21	26,96	26	39,38	39

[Institute of applied Technology, National Bureau of Standards, reproduit dans Morrison, Donald F., *Applied Linear Statistical Methods*, (1983), Prentice-Hall, New York]

a) Construire un nuage de points. Déterminer l'équation de la droite de régression pour prédire les résultats du laboratoire predict (clinique à partir de celui du NBS. Commenter sur la qualité de l'ajustement.

b) Faire un graphique des résidus centrés-réduits (vs les valeurs produites par le NBS) et commenter sur les hypothèses de linéarité et d'homoscédasticité. Identifier les données aberrantes, s'il y en a, ainsi que les données particulièrement influentes.

c) Déterminer un intervalle de confiance et des limites de prédiction (à 80 % de confiance) pour $x = 20$. *Énoncer vos résultats dans le langage de l'homme de la rue*.

d) Déterminer un intervalle de confiance à 95% pour β_1 et servez-vous en pour tester l'hypothèse que $\beta_1 = 1$.

e) Supposons que nous voulons tester l'hypothèse que les deux mesures sont identiques et que les différences sont des erreurs purement aléatoires. Voici une façon d'exprimer cette hypothèse :

$$H_o : \beta_o + \beta_1 x = x \text{ pour tout } x$$

ce qui est équivalent à l'hypothèse

$$H_o : \beta_o = 0 \text{ et } \beta_1 = 1.$$

Tester cette hypothèse. [*Suggestion* : reformulez le modèle $u_i = y_i - x_i = \beta_o + \delta_1 x_i + \varepsilon_i$, où $\delta_1 = \beta_1 - 1$]

Le modèle de régression traite les valeurs de x comme des constantes et non des variables aléatoires. Supposons qu'en fait (comme c'est le cas ici) les valeurs de x sont elles aussi aléatoires. Pourquoi est-il permis de les traiter comme constantes ? La réponse : la régression est une analyse conditionnelle : *Les moyennes des y_i sont des moyennes* conditionnelles étant donné les x_i : *c'est* $E(y_i|x_i)$ *qui est égale à* $\beta_o + \beta_1 x_i$ *et* c'*est la variance* conditionnelle $\text{Var}(y_i|x_i)$ *que nous avons désignée par* σ^2 *dans le modèle de régression simple*.

f) Considérons une analyse non conditionnelle basée sur le modèle bidimensionnel suivant : $\begin{bmatrix} x_i \\ y_i \end{bmatrix} \sim \mathcal{N}(\mu;\Sigma)$

, où $\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$ et $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$. Montrer que les hypothèses du modèle de régression (normalité, linéarité des moyennes et homoscédasticité) s'appliquent à la distribution *conditionnelle* des y_i. Précisons :

i) Montrer que la distribution conditionnelle de y_i étant donné x_i est normale ;

ii) Montrer que l'espérance conditionnelle de y_i étant donné x_i est une fonction linéaire de x_i : $E(y_i|x_i) = \gamma_o + \gamma_1 x_i$; et

iii) Montrer que la variance conditionnelle $\sigma^2 = \text{Var}(y_i|x_i)$ est indépendante de x_i.

Exprimer les paramètres β_o, β_1 et σ^2 du modèle de *régression* en fonction de μ_1, μ_2, σ_1^2, σ_2^2, et ρ.

g) L'hypothèse définie en e) sur les paramètres β_o et β_1 de la distribution *conditionnelle* de y_i étant donné x_i entraîne, dans le modèle bidimensionnel, l'hypothèse $H_o : \mu_1 = \mu_2$. Testez cette hypothèse en vous basant sur la distribution des $u_i = y_i - x_i$, où $\begin{bmatrix} x_i \\ y_i \end{bmatrix}$ suit le modèle défini en f) : $\begin{bmatrix} x_i \\ y_i \end{bmatrix} \sim \mathcal{N}(\mu;\Sigma)$.

h) En fait, l'hypothèse définie en e) est équivalente à l'hypothèse $H_{oo} : \mu_1 = \mu_2$ *et* $\sigma_1 = \rho\sigma_2$, où ρ est le coefficient de corrélation entre x et y. Soit $\boldsymbol{w} = \begin{bmatrix} w_{1i} \\ w_{2i} \end{bmatrix} = \begin{bmatrix} x_i \\ y_i - x_i \end{bmatrix}$. Montrer que sous H_{oo}, $\boldsymbol{w}$ est normale de moyenne $\begin{bmatrix} \mu_1 \\ 0 \end{bmatrix}$ et de matrice de covariance $\boldsymbol{\Psi} = \begin{bmatrix} \psi_1^2 & 0 \\ 0 & \psi_2^2 \end{bmatrix}$, où $\psi_1^2 = \sigma_1^2 = \rho^2\sigma_2^2$ et $\psi_2^2 = \sigma_2^2 - \sigma_1^2 = (1-\rho^2)\sigma_2^2$. Montrer que les estimateurs du maximum de vraisemblance des paramètres sont $\hat{\psi}_1^2 = \dfrac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$; $\hat{\psi}_2^2 = \dfrac{\sum_{i=1}^n (x_i - y_i)^2}{n}$; $\hat{\mu}_1 = \bar{x}$. Déterminer L_{oo}, la valeur maximale de la fonction de vraisemblance sous H_{oo}. Vous pouvez tenir pour acquis le résultat suivant :

La fonction de vraisemblance $\dfrac{1}{(2\pi)^{n/2}(\sigma^2)^{n/2}} \exp\left(-\dfrac{\sum_{i=1}^n (y_i - \mu)^2}{2\sigma^2}\right)$ *est maximisée lorsque* $\mu = \hat{\mu} = \bar{y}$ et $\sigma^2 = \hat{\sigma}^2 = (1/n)\sum_{i=1}^n (y_i - \bar{y})^2$; lorsque $\mu = \mu_o$ est connue, on a plutôt $\hat{\sigma}^2 = (1/n)\sum_{i=1}^n (y_i - \mu_o)^2$.

i) Déterminer un test de H_{oo} en utilisant le fait $-2 \ln (L_{oo}/L)$ suit asymptotiquement une loi χ^2 sous H_{oo} (L est la valeur maximale de la fonction de vraisemblance maximisée sans contrainte).

j) Rappelons que x et y sont censées mesurer la même chose et l'on pourrait croire que dans ce cas, on aurait $\beta_o = 0$ et $\beta_1 = 1$: lorsque $x = 0$, il faut bien que y le soit ; et lorsque x augmente d'une unité, y devrait augmenter d'une unité aussi. Pourtant ce n'est pas le cas. Considérer encore le modèle décrit en f) :

$\begin{bmatrix} x_i \\ y_i \end{bmatrix} \sim \mathcal{N}(\mu;\Sigma)$, où $\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$ et $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$. L'hypothèse que « les deux variables mesurent la même chose » devrait plutôt correspondre à l'hypothèse $\mu_1 = \mu_2$ (= μ) et $\sigma_1 = \sigma_2$ (= σ). Déterminer β_o et β_1 sous cette hypothèse.

k) Tester l'hypothèse $\mu_1 = \mu_2$ (=μ) et $\sigma_1 = \sigma_2$ (= σ) dans le modèle bidimensionnel en considérant la distribution du vecteur aléatoire $\boldsymbol{v}_i = \begin{bmatrix} x_i - y_i \\ x_i + y_i \end{bmatrix}$. Utilisez la statistique -2 ln (L_{oo}/L). [Vous remarquerez que l'hypothèse énoncée en k) dans le modèle bidimensionnel est incompatible avec l'hypothèse $\beta_1 = 1$ dans l'analyse conditionnelle (c'est-à-dire, dans le modèle de régression) à moins que ρ soit égal à 1, une hypothèse manifestement fausse.]

ℓ) Une faiblesse des modèles ci-dessus est qu'elle ne tient pas compte des erreurs de mesure. Afin d'en tenir compte, supposons que les valeurs observées $\begin{bmatrix} x_i \\ y_i \end{bmatrix}$ sont égales à $\boldsymbol{z}_i + \boldsymbol{\varepsilon}_i$, où $\boldsymbol{z}_i$ est le vecteur des mesures des deux labos si celles-ci avaient été prises sans erreur (c'est donc l'espérance conditionnelle des mesures étant donné le spécimen) et $\boldsymbol{\varepsilon}_i = \begin{bmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{bmatrix}$ est un vecteur d'erreurs de loi normale (indépendante de x_i et de y_i) de moyenne $\mathbf{0}$ et de matrice de covariance $\sigma^2\mathbf{I}_2$. Sous l'hypothèse que les deux laboratoires mesurent exactement la même chose, $\boldsymbol{z}_i = z_i\boldsymbol{e}$, où z_i est la teneur réelle de plomb dans le spécimen. Si on suppose que z_i est de loi normale de moyenne μ et d'écart-type σ_z, alors chaque observation $z_i\boldsymbol{e} + \boldsymbol{\varepsilon}_i$ est de moyenne $\mu\boldsymbol{e}$ et de matrice de covariance $\sigma_z^2\boldsymbol{ee}' + \sigma^2\mathbf{I}$. Sous cette hypothèse, déterminer la distribution de $\boldsymbol{v}_i$ (défini en k)). Sachant que la matrice de covariance échantillonnale des $\boldsymbol{v}_i$ est sans biais pour la variance de $\boldsymbol{v}_i$, trouver un estimateur sans biais de μ, de σ^2 et de σ_z^2.

4.21 Les données suivantes portent sur un certain nombre de ménages. Les variables sont :

y : Facture d'électricité
x_1 : Revenu du ménage
x_2: Nombre de personnes
x_3: Superficie du logement

a) Supposons qu'on doive faire un choix d'une variables parmi x_1, x_2, et x_3 pour estimer y. Obtenez la matrice des corrélations à l'aide d'un logiciel. La variable « surface », ayant la plus forte corrélation avec y, est la meilleure candidate. Mais supposons qu'on veuille introduire une deuxième variable indépendante, et utiliser celle-ci en même temps que « surface » (en supposant connue la technique pour le faire). Dites pourquoi la variable « revenu », bien qu'elle soit fortement corrélée avec le montant de la facture, n'est probablement pas une aussi bonne candidate que « Nombre » (de personnes dans le ménage).

b) Concentrez-vous sur la variable $x = x_3$.
   i) Construisez un nuage de points ;
   ii) Déterminez la droite des moindres carrés ;
   iii) Dites si la régression est significative, et si on peut s'attendre à de bonnes prédictions ;
   iv) Déterminer un intervalle de confiance pour la moyenne des y et des limites de prédiction pour x = 1000.

ID	y	x_1	x_2	x_3	ID	y	x_1	x_2	x_3	ID	y	x_1	x_2	x_3
1	228	3220	2	1160	12	420	3600	3	1680	23	756	4420	2	1990
2	156	2750	1	1080	13	876	5370	1	2550	24	636	4480	2	2070
3	648	3620	2	1720	14	840	3180	7	1770	25	708	3820	4	1850
4	528	3940	1	1840	15	876	5910	2	2960	26	960	5740	2	2700
5	552	4510	3	2240	16	276	3020	2	1190	27	1080	5600	3	3030
6	636	3990	4	2190	17	1236	5920	3	3130	28	480	3950	2	1700
7	444	2430	1	830	18	372	3520	2	1560	29	96	2290	3	890
8	144	3070	1	1150	19	276	3720	1	1510	30	1272	5580	5	3270
9	744	3750	2	1570	20	540	4840	1	2190	31	1056	5820	2	2660
10	1104	4790	5	2660	21	1044	4700	6	2620	32	156	3160	2	1330
11	204	2490	1	900	22	552	3270	2	1350	33	396	2880	4	1280
										34	768	3780	3	1950

4.22 [Graybill, Franklin A. et Iyer, Hariharan K., *Regression Analysis, concepts and applications*, (1994), Duxbury Press] Dans le cadre d'une étude sur la perception visuelle, on soumet 10 sujets à l'expérience suivante. On présente à chaque sujet une image sur diapositive 100 fois. Le sujet doit déclarer à chaque fois s'il a perçu l'objet ou non. Les diapositives présentées aux dix sujets portaient la même image, mais le contraste optique entre l'objet et le fond variait. Le tableau ci-dessous montre les valeurs de x, une mesure du contraste optique, et y, la proportion des fois où l'objet a été perçu par le sujet.

y *Taux de détection de l'objet*	x *Contraste optique entre l'objet et le fond*	y *Taux de détection de l'objet*	x *Contraste optique entre l'objet et le fond*
0,02	0,000	0,56	0,025
0,06	0,005	0,78	0,030
0,10	0,010	0,86	0,035
0,18	0,015	0,94	0,040
0,35	0,020	0,99	0,045

a) Faites un graphique pour voir si une droite pourrait décrire la relation entre x et y.
b) Déterminer une droite de régression de y sur x, puis examiner le graphique des résidus normés.
c) Considérez le modèle logistique : déterminez une régression de $z = \ln\left(\frac{y}{1-y}\right)$ sur x. Ce modèle semble-t-il meilleur ? [*La somme des carrés* $\sum_{i=1}^{n}(z_i - \hat{z}_i)^2$ *ne peut pas être comparée à la somme* $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ *issue du premier modèle. Le taux de détection estimé par ce deuxième modèle est* $\hat{\hat{y}}_i = \frac{e^{\hat{z}_i}}{1+e^{\hat{z}_i}}$. *On doit donc comparer* $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ *à* $\sum_{i=1}^{n}(y_i - \hat{\hat{y}}_i)^2$.]

4.23 [Draper, Norman, et Smith, Harry, *Applied Regression Analysis, Second Edition*, (1981), Wiley, New York]. Le tableau suivant présente les prix (y) de certains portos selon l'année.

Année	*Prix (Dollars US)*	*Année*	*Prix (Dollars US)*	*Année*	*Prix (Dollars US)*
1890	50	1935	13	1950	7
1900	35	1940	7	1952	5
1920	25	1941	10	1955	6
1931	12	1944	6	1960	5
1934	15	1948	9		

a) Déterminer une régression du prix sur l'âge (une variable exogène plus naturelle que l'année). La régression linéaire semble-t-elle adéquate ?
b) Déterminer une régression de ln *prix* sur l'âge. Obtenez-vous de meilleurs résultats ? Discuter.

4.24 Les données du tableau 4 portent sur le salaire à l'entrée d'un groupe d'employés et d'employées. Le but est de savoir s'il y a une discrimination systémique contre les femmes.

a) Faites le test d'égalité de moyennes afin de déterminer s'il y a une différence significative entre le salaire des hommes et celui des femmes.
b) À l'aide d'un modèle de régression, déterminer un intervalle de confiance pour le salaire moyen des femmes avec 12 années de scolarité. Faites la même chose pour les hommes. Les intervalles sont-ils disjoints ?
c) À l'aide des deux modèles de régression, tester l'hypothèse que le salaire moyen des femmes avec 12 années de scolarité est égal au salaire moyen des hommes avec 12 années de scolarité. [En attendant de développer les outils nécessaires pour ce problème, vous allez considérer vos variances estimées comme de vraies variances.]
d) Tester la même hypothèse qu'en c) en ne considérant que les sujets qui ont 12 années de scolarité. Pouvez-vous expliquer pourquoi l'approche en c) est meilleure ? [*En faisant abstraction du fait qu'en c) vous utilisez les variances estimées comme si elles étaient de vraies variances.*]

4.25 Le tableau 8 présente des données démographiques et économiques sur 96 pays.

Variables
pays : Appellation du pays
pib : PIB par habitant, en dollars US
espvie : Espérance de vie des femmes à 60 ans
mort : Taux de mortalité infantile – filles (nombre par 1000 naissances)

a) Faites un graphique pour exposer la relation entre le PIB et l'espérance de vie des femmes à 60 ans. Vous verrez qu'elle est loin d'être linéaire.
b) Considérer maintenant la relation entre l'espérance de vie et le *logarithme* du PIB.
   (i) Déterminer maintenant la droite de régression permettant de prédire l'espérance de vie des femmes à 60 ans à partir du *logarithme* du *PIB*.
   (ii) Faire un graphique des résidus centrés-réduits et commenter. Y a-t-il des données aberrantes ?
   (iii) Il y a deux données aberrantes. Éliminez-les et recommencez. La régression est-elle meilleure ?
   (iv) Peut-on affirmer avec confiance que la « vraie » droite des moindres carrés n'est pas horizontale ?
c) Montrer par un nuage de points ou par un graphique des résidus que la relation entre le taux de mortalité infantile et le PIB ne semble pas non plus linéaire.
d) Supposons que la relation entre Y, le taux de mortalité infantile, et x, le PIB, est de la forme suivante : $y = \alpha x^{\beta}$. Il s'ensuit que $\ln y = \ln \alpha + \beta \ln x$.
   (i) Examiner alors le lien entre le logarithme du PIB et le logarithme du taux de mortalité : vérifier visuellement que la relation est linéaire et calculer le coefficient de corrélation.
   (ii) Peut-on affirmer avec confiance que β est non nul ?
   (iii) Estimer le taux de mortalité d'un pays dont le PIB est de 1000 \$ par habitant.

4.26 Le tableau 5 présente des données sur les professeurs d'une certaine université. Voici le sens des variables

`ID`	*Identité*	`Salaire0`	*Salaire à l'entrée*
`Sexe`	*Sexe :* 0 = *femme*	`Salaire91`	*Salaire en* 1991
`Entrée`	*Date d'entrée*	`Expérience`	*Expérience antérieure à l'engagement*
`Département`	*Département*		

Dans ce qui suit, soustrayez de la date d'entrée la plus petite valeur, soit 1965, et définissez cela comme la date d'entrée

a) On s'intéresse au lien entre le salaire à l'entrée et l'expérience. Déterminer la droite de régression, présenter un nuage de points. Commenter vos résultats.
b) On s'intéresse ici à la relation entre le salaire à l'entrée et la date d'entrée.
   i) Faites un nuage de points pour montrer que la relation entre le salaire à l'entrée et la date d'entrée n'est pas linéaire.
   ii) Si on suppose que les salaires se sont accrus à un taux constant ces dernières décennies, alors la relation est en fait exponentielle : $y = \alpha e^{\beta x}$. Auquel cas, $\ln y = \ln \alpha + \beta x$. Examinez à l'aide d'un nuage de points la relation entre le logarithme du salaire et la date d'engagement.

iii) Bien que le graphique en (ii) demeure quelque peu convexe, faites une analyse de régression qui lie le logarithme du salaire à l'année d'engagement. Estimer le salaire moyen d'un professeur engagé en 1978. Commenter vos résultats.[*Commencez par estimer le* logarithme *de du salaire*]

iv) Déterminer un intervalle de confiance pour le salaire d'un professeur engagé en 1978 [*Commencez par déterminer un intervalle de confiance pour le logarithme du salaire et en déduire un intervalle de confiance pour le salaire.*)

v) Montrer que le pourcentage d'accroissement annuel dans une relation de la forme $y = \alpha e^{\beta x}$ est de $\theta = 100(e^{\beta}-1)$ %. Estimer la variance de l'estimateur $\hat{\theta} = \theta(\hat{\beta}) = e^{\hat{\beta}} - 1$ de θ à l'aide de la méthode delta.

[Un aperçu de la méthode delta : on développe la fonction θ jusqu'au terme linéaire, soit $\theta(\hat{\beta}) \approx \theta(\beta) + \theta'(\beta)(\hat{\beta}-\beta)$, d'où $\mathrm{Var}[\theta(\hat{\beta})] \approx [\theta'(\beta)]^2\sigma^2_{\hat{\beta}}$. On estime alors cette variance par $[\theta'(\hat{\beta})]^2\hat{\sigma}^2_{\hat{\beta}}$]

c) Revenons au salaire à l'entrée. Utiliser le taux d'accroissement θ (en décimales, pas en pourcentage) obtenu en b) (v) pour ajuster les salaires y à l'entrée, c'est-à-dire, pour les exprimer en dollars de 1991. Il suffit de multiplier y par $(1+\hat{\theta})^k = e^{k\hat{\beta}}$, où k est la différence entre 1991 et l'année d'engagement. Examiner à nouveau l'effet de l'expérience sur les salaires à l'entrée (comme en a)), en utilisant cette fois-ci le salaire ajusté. La relation ici est-elle plus forte qu'en a) ? Pouvez-vous expliquer pourquoi, en termes concrets ?

d) Calculer le coefficient de corrélation entre le salaire en 1991 et le salaire à l'entrée. Expliquer, en termes concrets, pourquoi ce coefficient est négatif. Confirmer votre interprétation en remplaçant le salaire brut à l'entrée par le salaire ajusté.

4.27 Le tableau 6 présente des données démographiques et économiques sur un certain nombre de pays. Les variables sont les suivantes :

`Naiss`	*Le taux de naissance par* 100000 *h.*
`Deces`	*Le taux de décès par* 100000 *h.*
`Mort`	*Le taux de mortalité infantile*
`Mvie`	*L'espérance de vie des hommes*
`Fvie`	*L'espérance de vie des hommes*
`PNB`	*Le produit national brut*
`Pays`	*Le nom du pays*

a) Faites un graphique pour exposer la relation entre le PNB et l'espérance de vie des hommes. Vous verrez qu'elle est loin d'être linéaire.

b) Déterminer maintenant la droite de régression permettant de prédire l'espérance de vie des hommes à partir du logarithme du PNB. Estimer l'espérance de vie des hommes d'un pays dont le PNB est de 1000 $ par habitant.

c) La relation entre le taux de mortalité infantile et le PNB ne semble pas non plus linéaire. Supposons que la relation entre y, le taux de mortalité et x, le PNB, est de la forme suivante : $y = \alpha x^{\beta}$. Il s'ensuit que $\ln y = \ln \alpha + \beta \ln x$. Examiner alors le lien entre le logarithme du PNB et le logarithme du taux de mortalité. Vérifier visuellement que la relation est approximativement linéaire et calculer le coefficient de corrélation. Estimer le taux de mortalité d'un pays dont le PNB est de 1000 $ per capita.

4.28 [Reference: Moore, David S., and George P. McCabe (1989). *Introduction to the Practice of Statistics*. Source originale: étude dirigée par Jim Baumann et Leah Jones de Purdue University Education Department.]Une expérience a été menée pour déterminer l'effet de trois méthodes d'enseignement de la lecture à des enfants. On a fait passer aux sujets un test de compréhension avant et après la période d'enseignement. Nombre de cas : 66

Variables	1. ID:	Identification du sujet
	2. Groupe:	Type de formation reçue (nommé Basal, DRTA, ou Strat)
	3. PRE1:	Score au premier pré-test de compréhension
	4. PRE2:	Score au second pré-test de compréhension
	5. POST1:	Score au premier post test de compréhension
	6. POST2:	Score au second post test de compréhension
	7. POST3:	Score au troisième post test de compréhension

Les données sont présentées dans le tableau 7 en annexe.

Dans ce numéro, nous ne nous préoccupons que de la relation entre le pré-test PRE1 (x) et le post-test POST2 (y), sans tenir compte des différentes méthodes d'enseignement.

a) Faites un nuage de points (et expliquez ce qu'il vous apprend.)
b) Déterminer la droite de régression et le coefficient de corrélation.
c) Présentez le tableau fourni par la commande `summary()` et expliquez à vos clients les principales données du tableau (dans la mesure où vous les estimez pertinents pour votre client).
d) Déterminer une table d'analyse de variance et expliquer à vos clients les principales données qui s'y trouvent (dans la mesure où vous les estimez pertinents pour votre client).
e) Déterminer un intervalle de confiance à **75 %** pour $\beta_o+\beta_1 x$ pour $x = 10$ et pour $x = 18$. Expliquez pourquoi l'un des intervalles est beaucoup plus large que l'autre. Dites quelle règle générale cela illustre.
f) Déterminer des limites de prédiction à **75 %** pour $\beta_o+\beta_1 x$ pour $x = 10$ et pour $x = 18$. Expliquer clairement la différence entre ces limites et les intervalles de confiance déterminés au numéro précédent.

4.29 Supposons que, contrairement au modèle décrit dans ce chapitre, les valeurs $x_1, x_2,\ldots, x_n$ ne sont pas fixes : ce sont des réalisations de n variables aléatoires $X_1, X_2, \ldots , X_n$ de loi normale, de moyenne μ_x et d'écart-type σ_x.

L'expression $\mathrm{Var}(\hat\beta) = \dfrac{\sigma^2_{y.x}}{\sum_{i=1}^n (x_i - \bar x)^2}$ déterminée dans ce chapitre représente alors la variance *conditionnelle*, étant donné les valeurs observées des X_i.

a) Montrer que les estimateurs usuels $\hat\beta_o$ et $\hat\beta_1$ demeurent sans biais sous ce modèle.

a) Montrer que la variance *inconditionnelle* de $\hat\beta$ est $\dfrac{\sigma^2_{y.x}}{(n-3)\sigma^2_x}$.

b) Montrer que $\hat\sigma^2_{\hat\beta} = \dfrac{\sigma^2_{y.x}}{\sum_{i=1}^n (x_i - \bar x)^2}$ est un estimateur sans biais de $\mathrm{Var}(\hat\beta)$.

4.30 Supposons que dans le modèle de régression linéaire simple on remplace les valeurs x_i par les écarts $x_i - \bar x$, donc on remplace le modèle $y_i = \beta_o+\beta_1 x_i+\varepsilon_j$ par $y_i = \gamma_o+ \gamma_1(x_i - \bar x) +\varepsilon_i$.

a) Montrer que $\hat\gamma_1 = \hat\beta_1$

b) Montrer que $\hat\gamma_o = \bar y$. Déterminer la loi de $\hat\gamma_o$.

c) Montrer que la covariance entre $\hat\gamma_o$ et $\hat\gamma_1$ est nulle (et donc que $\hat\gamma_o$ et $\hat\gamma_1$ sont sous l'hypothèse de normalité).

4.31 Soit $Y_1, Y_2,\ldots,Y_n$ n variables aléatoires indépendantes de loi de Poisson de paramètres $\alpha x_1, \alpha x_2,\ldots,\alpha x_n$.

a) Montrer que l'estimateur des moindres carrés (non pondérés) est $\hat\alpha_1 = \dfrac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}$. Montrer que $\hat\alpha_1$ est sans biais et déterminer sa variance.

b) Montrer que l'estimateur des moindres carrés pondérés (qui minimise $\sum_{i=1}^n \dfrac{(Y_i - \alpha x_i)^2}{\alpha x_i}$) est

$$\hat\alpha_2 = \sqrt{\frac{\sum_{i=1}^n Y_i^2 / x_i}{\sum_{i=1}^n x_i}}.$$

c) Montrer que l'estimateur du maximum de vraisemblance est $\hat\alpha_3 = \dfrac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n x_i}$. Montrer que $\hat\alpha_3$ est sans biais et déterminer sa variance.

d) Montrer que la variance de $\hat\alpha_3$ atteint la borne de Cramér-Rao, soit $\dfrac{\left[\dfrac{d}{d\alpha}\mathrm{E}(\hat\alpha_3)\right]^2}{\mathrm{E}\left[\left(\dfrac{d}{d\alpha}\ln L\right)^2\right]}$, où L est la fonction de vraisemblance.

e) Considérer une étude conçue dans le but de déterminer si un certain polluant atmosphérique a un effet sur la probabilité d'avoir le cancer. À cette fin, on choisit 30 localités qui connaissent des taux de pollution variables. On compte alors le nombre de personnes d'habitants dans chacune des localités qui ont été atteint du cancer au courant des 5 dernières années. Si Y_i est le nombre de personnes ainsi atteintes dans la localité i, on supposera que Y_i est de loi de Poisson de paramètre $\alpha_i x_i$, où x_i est la population de la localité (en dizaines de milliers). Voici les valeurs de Y et de x :

y : 7 ; 1 ; 6 ; 6 ; 4 ; 5 ; 13 ; 7 ; 7 ; 5 ; 2 ; 1 ; 9 ; 6 ; 5 ; 12 ; 6 ; 7 ; 6 ; 3 ; 5 ; 3 ; 1 ; 2 ; 6 ; 9 ; 11 ; 12 ; 14 ; 9

x : 11 ; 15 ; 17 ; 18 ; 18 ; 18 ; 18 ; 19 ; 19 ; 19 ; 20 ; 20 ; 20 ; 20 ; 21 ; 21 ; 21 ; 21 ; 21 ; 21 ; 21 ; 22 ; 22 ; 23 ; 25 ; 25 ; 26 ; 27 ; 28 ; 29.

i) Déterminer les estimateurs du maximum de vraisemblance de α_i.

ii) Déterminer l'estimateur du maximum de vraisemblance de α sous l'hypothèse que $\alpha_1 = \alpha_2 = \ldots = \alpha_{30} = \alpha$.

iii) Tester l'hypothèse H_o : $\alpha_1 = \alpha_2 = \ldots = \alpha_{30}$ en utilisant le fait que la statistique $-2 \ln(L_o/L)$ est à peu près de loi χ^2 sous H_o (le nombre de degrés de liberté étant la réduction du nombre de paramètres imposée par H_o. L est la fonction de vraisemblance maximisée sans restriction ; L_o est la fonction de vraisemblance maximisée sous H_o.

Annexe

Tableau 1 ***Pouls (y) et Nombre d'années dans les plaines (x)***

ID	*Nombre d'années*	*Pouls*	*ID*	*Nombre d'années*	*Pouls*
1	1	88	21	18	80
2	6	64	22	11	76
3	5	68	23	11	60
4	1	52	24	21	64
5	1	72	25	24	64
6	19	72	26	14	68
7	5	64	27	25	76
8	25	80	28	32	60
9	6	76	29	5	76
10	13	60	30	12	88
11	13	68	31	25	72
12	10	72	32	26	68
13	15	88	33	10	60
14	18	60	34	19	74
15	2	60	35	18	72
16	12	72	36	10	56
17	15	84	37	1	64
18	16	64	38	43	72
19	17	72	39	40	92
20	10	64			

Tableau 2 ***Temps de finition d'une toile (T) et surface de la toile (S)***

i	*T*	*S*
1	5,50	9,30
2	5,90	13,50
3	5,80	11,10
4	6,30	14,90
5	7,00	16,70
6	7,50	23,20
7	5,50	11,10
8	7,20	20,40
9	6,50	15,80
10	6,50	14,90
11	7,10	18,60
12	7,00	15,80
13	6,90	16,70
14	6,80	15,80
15	6,60	16,70
Totaux	98,10	234,50

Tableau 3 Habileté mathématique (H) et résultat à un examen d'algèbre (F)

i	*F*	*H*	*i*	*F*	*H*
1	36	9	11	59	26
2	23	10	12	58	28
3	22	13	13	72	30
4	36	15	14	87	31
5	49	16	15	86	32
6	32	18	16	79	33
7	44	20	17	74	34
8	52	22	18	78	36
9	51	23	19	99	38
10	83	24	20	85	40
			Totaux	1 205	498

Tableau 4 Salaire et scolarité d’un groupe de femmes et d’hommes

Femmes

Salaire	*Scol*	*Salaire*	*Scol*	*Salaire*	*Scol*	*Salaire*	*Scol*	*Salaire*	*Scol*
3.99	12	4.80	12	5.10	12	5.40	12	5.58	12
4.02	10	4.80	12	5.10	15	5.40	12	5.64	12
4.29	12	4.80	12	5.10	15	5.40	12	5.70	12
4.38	8	4.80	12	5.10	16	5.40	12	5.70	12
4.38	8	4.80	12	5.16	12	5.40	12	5.70	15
4.38	12	4.80	12	5.22	8	5.40	12	5.70	15
4.38	12	4.80	12	5.22	12	5.40	15	5.70	15
4.38	12	4.80	12	5.28	8	5.40	15	6.00	12
4.44	15	4.80	16	5.28	8	5.40	15	6.00	15
4.50	8	4.89	8	5.28	12	5.40	15	6.12	12
4.50	12	5.10	8	5.40	8	5.52	12	6.30	12
4.62	12	5.10	12	5.40	8	5.52	12	6.30	15
4.80	8								

Hommes

Salaire	*Scol*	*Salaire*	*Scol*	*Salaire*	*Scol*	*Salaire*	*Scol*	*Salaire*	*Scol*
4.62	12	5.40	12	6.00	12	6.00	15	6.60	15
5.04	15	5.40	15	6.00	12	6.00	15	6.60	15
5.10	12	5.40	15	6.00	12	6.00	15	6.84	15
5.10	12	5.70	15	6.00	12	6.00	16	6.90	12
5.22	12	6.00	8	6.00	15	6.30	15	6.90	15
5.40	12	6.00	12	6.00	15	6.60	15	8.10	16
5.40	12	6.00	12						

Tableau 5 Statistiques sur un échantillon de professeurs

Identité	*Sexe*	*Date d’entrée*	*Département*	*Salaire à l’en trée*	*Salaire en 19 91*	*Expérience*
1	F	1975	8	89	350	1
2	M	1968	8	75	430	4
3	F	1984	5	176	260	8
4	M	1972	1	91	511	4
5	M	1978	8	222	492	19,5
6	M	1975	8	140	449	3,5
7	M	1983	3	225	344	5
8	F	1979	7	177	406	5
9	M	1981	5	190	352	7,5
10	M	1972	3	95	406	3
11	M	1983	8	190	294	6
12	M	1973	4	123	438	4
13	M	1981	4	190	368	6
14	M	1968	4	105	492	5
15	M	1981	8	159	302	4,83
16	M	1971	3	115	475	4
17	M	1972	6	120	458	5
18	M	1971	8	120	468	5
19	F	1972	7	114	432	4
20	M	1986	1	190	245	0
21	F	1982	1	145	341	7,5
22	F	1986	4	230	296	6
23	M	1968	4	108	510	6,33
24	M	1968	5	96	476	6
25	M	1972	3	120	459	4

Tableau 5 (suite)

Identité	*Sexe*	*Date d'entrée*	*Département*	*Salaire à l'en trée*	*Salaire en 19 91*	*Expérience*
26	M	1967	8	110	515	3
27	F	1977	8	130	363	1
28	F	1983	4	170	276	4,5
29	M	1970	5	115	468	9,67
30	M	1969	8	108	476	6
31	F	1972	7	110	430	4,5
32	F	1985	5	210	287	8,5
33	M	1971	8	105	438	6,33
34	M	1969	6	110	484	6
35	M	1967	4	91	477	4,5
36	F	1967	3	120	539	10,5
37	M	1975	8	120	399	4
38	M	1967	1	88	470	3
39	M	1982	5	160	275	6
40	F	1987	2	210	255	6
41	M	1971	5	100	422	3
42	M	1967	8	85	465	1
43	M	1976	3	140	416	5
44	M	1972	7	100	410	2
45	M	1970	6	111	467	6
46	M	1971	3	96	416	4
47	M	1969	2	100	459	5
48	M	1971	2	105	431	5
49	M	1972	5	82	377	0,5
50	M	1983	5	220	351	11,33
51	M	1967	6	102	494	5
52	M	1969	4	103	476	6
53	F	1972	6	100	410	3
54	M	1976	5	133	403	6,33
55	F	1973	6	114	423	5
56	M	1972	7	105	432	3
57	M	1968	2	90	469	5
58	F	1969	3	126	523	11
59	M	1967	3	90	476	5,75
60	M	1969	7	102	474	5
61	F	1969	8	90	439	2
62	F	1968	2	80	455	4
63	M	1967	6	88	474	3
64	M	1969	1	115	511	12
65	F	1974	6	128	426	8
66	M	1966	3	82	482	3
67	M	1967	6	85	461	4
68	M	1969	6	85	431	1,5
69	M	1981	5	160	303	5
70	M	1972	6	110	424	9,67
71	M	1966	1	100	509	9,5
72	F	1983	5	180	283	4,48
73	M	1967	1	96	479	7,5
74	F	1966	5	98	513	7,5
75	M	1967	3	100	507	7
76	M	1969	3	85	428	1
77	F	1971	6	102	451	4
78	F	1970	8	98	446	5
79	M	1967	4	90	482	7
80	M	1966	4	80	481	10,17
81	M	1965	3	75	476	5
82	F	1970	3	93	444	5
83	M	1967	4	116	525	11

Tableau 5 (suite)

Identité	*Sexe*	*Date d'entrée*	*Département*	*Salaire à l'en trée*	*Salaire en 19 91*	*Expérience*
84	M	1971	6	95	434	2
85	F	1965	5	96	509	10,67
86	M	1972	1	95	398	4
87	M	1969	6	92	446	4
88	M	1966	6	98	538	10
89	F	1970	2	85	413	2,5
90	M	1970	3	85	415	1
91	M	1985	5	215	294	8,5
92	M	1969	4	83	426	1
93	F	1971	5	100	426	8,25
94	M	1976	8	150	411	7,5
95	F	1965	3	84	475	4,5
96	M	1968	2	80	441	5
97	M	1967	1	108	526	10
98	M	1970	8	100	443	8
99	F	1967	8	72	440	0
100	M	1966	1	105	510	11
101	F	1969	1	85	431	1,25
102	M	1966	2	83	467	5
103	M	1969	1	85	435	5,5
104	M	1968	3	82	444	5
105	F	1967	4	72	439	0
106	M	1967	3	93	483	9
107	M	1967	2	83	463	5
108	F	1968	5	82	445	5
109	M	1969	2	90	463	4
110	F	1967	6	75	444	1
111	M	1969	5	90	437	0
112	M	1968	6	115	516	11
113	F	1969	8	83	441	2
114	F	1966	8	85	472	4
115	F	1970	2	90	430	5,25
116	M	1968	7	85	456	4
117	F	1972	6	88	387	2
118	F	1971	2	85	395	1
119	M	1971	3	88	406	2,75
120	F	1975	2	107	364	6,83
121	F	1967	4	75	434	0
122	F	1970	3	85	410	4
123	F	1981	4	160	303	5,5
124	M	1970	8	90	424	6
125	F	1968	6	82	448	4
126	F	1978	4	130	331	4,57
127	F	1971	8	84	393	1
128	M	1967	8	103	509	14,5
129	M	1967	3	75	447	8
130	M	1968	8	75	440	6
131	M	1969	2	83	439	1
132	M	1980	8	145	312	5,5
133	F	1977	4	120	331	2,5
134	F	1980	2	150	328	7,75
135	M	1967	1	85	453	5,75
136	F	1970	8	82	397	0
137	M	1965	6	110	543	12
138	M	1965	3	82	473	4
139	M	1974	6	105	384	5
140	F	1972	6	88	401	1

Tableau 5 (suite)

Identité	*Sexe*	*Date d'entrée*	*Département*	*Salaire à l'en trée*	*Salaire en 19 91*	*Expérience*
141	M	1970	3	80	420	3,25
142	F	1985	4	170	253	5,17
143	M	1985	5	180	246	4,92
144	F	1974	8	94	356	1
145	F	1972	3	88	400	3
146	M	1969	6	83	431	3
147	M	1975	3	95	352	3,25
148	M	1982	2	155	278	2
149	M	1973	7	110	404	8
150	M	1970	3	85	398	2
151	F	1975	1	107	361	1
152	F	1980	5	130	287	6,5
153	F	1969	7	80	426	6
154	F	1981	5	165	321	10,5
155	F	1976	7	110	337	3
156	M	1981	5	165	318	15,1
157	M	1966	3	68	436	4
158	F	1975	2	102	346	9,75
159	F	1978	5	100	280	1
160	F	1979	8	133	295	8
161	M	1968	6	106	444	8
162	F	1981	5	150	302	9,5
163	M	1981	8	145	281	4
164	F	1975	2	117	376	11,83
165	F	1975	2	112	376	9
166	F	1976	5	100	317	5
167	M	1982	6	145	266	3
168	F	1983	5	140	226	6,33
169	F	1976	1	100	319	2
170	F	1984	1	160	240	1
171	M	1981	5	138	267	9

Tableau 6 Données statistiques sur un ensemble de pays

	Pays	PIB	EspVie	MortInf	Décès
1	Afghanistan	900	45.02	149.20	17.39
2	Albania	8000	77.41	14.61	6.15
3	Algeria	7300	74.50	25.81	4.69
4	American_Samoa	8000	74.21	9.66	4.10
5	Andorra	46700	82.43	3.80	6.35
6	Angola	8200	38.76	175.90	23.40
7	Anguilla	12200	80.87	3.47	4.44
8	Antigua_and_Barbuda	16400	75.48	14.63	5.73
9	Argentina	14700	76.95	10.81	7.38
10	Armenia	5700	73.23	18.85	8.46
11	Aruba	21800	75.72	12.92	7.84
12	Australia	41000	81.81	4.61	6.88
13	Austria	40400	79.78	4.32	10.14
14	Azerbaijan	10900	67.36	51.08	8.25
15	Bahamas_The	28700	71.18	13.49	6.88
16	Bahrain	40300	78.15	10.43	2.61
17	Bangladesh	1700	69.75	50.73	5.75
18	Barbados	21800	74.34	11.86	8.39
19	Belarus	13600	71.20	6.25	13.77
20	Belgium	37800	79.51	4.33	10.57
21	Belize	8400	68.23	21.95	5.87
22	Benin	1500	59.84	61.56	9.00
23	Bermuda	69900	80.71	2.47	7.57
24	Bhutan	5500	67.30	44.48	7.12
25	Bolivia	4800	67.57	42.16	6.85
26	Bosnia_and_Herzegovina	6600	78.81	8.67	8.80
27	Botswana	14000	58.05	11.14	10.57
28	Brazil	10800	72.53	21.17	6.36
29	British_Virgin_Islands	38500	77.63	13.63	4.49
30	Brunei	51600	76.17	11.51	3.35
31	Bulgaria	13500	73.59	16.68	14.32
32	Burkina_Faso	1200	53.70	81.40	12.74
33	Burma	1400	64.88	49.23	8.16
34	Burundi	300	58.78	61.82	9.61
35	Cambodia	2100	62.67	55.49	8.07
36	Cameroon	2300	54.39	60.91	11.83
37	Canada	39400	81.38	4.92	7.98
38	Cape_Verde	3800	70.70	26.94	6.34
39	Cayman_Islands	43800	80.68	6.63	5.10
40	Central_African_Republic	700	50.07	99.38	15.01
41	Chad	1600	48.33	95.31	15.47
42	Chile	15400	77.70	7.34	5.97
43	China	7600	74.68	16.06	7.03
44	Colombia	9800	74.55	16.39	5.26
45	Comoros	1000	64.20	62.63	7.23
46	Congo_Democratic_Republic_of_the	300	55.33	78.43	11.06
47	Congo_Republic_of_the	4100	54.91	76.05	11.49
110	Lebanon	14400	75.01	15.85	6.54
111	Lesotho	1700	51.63	55.04	15.19
112	Liberia	500	57.00	74.52	10.62
113	Libya	14000	77.65	20.09	3.40
114	Liechtenstein	141100	80.31	4.15	7.61
115	Lithuania	16000	75.34	6.27	11.33
116	Luxembourg	82600	79.61	4.44	8.48
117	Macau	33000	84.41	3.18	3.72
118	Macedonia	9700	75.14	8.54	8.91
119	Madagascar	900	63.63	51.45	7.79
120	Malawi	800	51.70	81.04	13.22
121	Malaysia	14700	73.79	15.02	4.93
122	Maldives	6900	74.45	27.45	3.71
123	Mali	1200	52.61	111.35	14.29
124	Malta	25600	79.72	3.69	8.60
125	Marshall_Islands	2500	71.76	23.74	4.38
126	Mauritania	2100	61.14	60.42	8.83
127	Mauritius	14000	74.48	11.52	6.68
128	Mexico	13900	76.47	17.29	4.86
129	Micronesia_Federated_States_of	2200	71.52	24.34	4.35
130	Moldova	2500	71.37	12.43	10.74
131	Monaco	30000	89.73	1.79	8.28
132	Mongolia	3600	68.31	37.26	6.04
133	Montserrat	3400	73.16	15.23	7.20
134	Morocco	4800	75.90	27.53	4.75
135	Mozambique	1000	51.78	78.95	13.00
136	Namibia	6900	52.19	45.59	12.95
137	Nauru	5000	65.35	8.66	6.11
138	Nepal	1200	66.16	44.54	6.81
139	Netherlands	40300	79.68	4.59	8.85
140	New_Caledonia	15000	76.75	5.71	5.28
141	New_Zealand	27700	80.59	4.78	7.15
142	Nicaragua	3000	71.90	22.64	5.03
143	Niger	700	53.40	112.22	14.11
144	Nigeria	2500	47.56	91.54	16.06
145	Northern_Mariana_Islands	12500	77.08	5.79	3.28
146	Norway	54600	80.20	3.52	9.24
147	Oman	25600	74.22	15.47	3.45
148	Pakistan	2500	65.99	63.26	6.92
149	Palau	8100	71.78	12.43	7.87
150	Panama	13000	77.79	11.64	4.65
151	Papua_New_Guinea	2500	66.24	43.29	6.58
152	Paraguay	5200	76.19	23.02	4.57
153	Peru	9200	72.47	22.18	5.93
154	Philippines	3500	71.66	19.34	5.02
155	Poland	18800	76.05	6.54	10.17
156	Portugal	23000	78.54	4.66	10.80

Tableau 6 (suite)

	Pays	PIB	EspVie	MortInf	Décès
48	Cook_Islands	9100	74.70	15.81	7.37
49	Costa_Rica	11300	77.72	9.45	4.33
50	Cote_d'Ivoire	1800	56.78	64.78	10.16
51	Croatia	17400	75.79	6.16	11.91
52	Cuba	9900	77.70	4.90	7.47
53	Cyprus	21000	77.82	9.38	6.45
54	Czech_Republic	25600	77.19	3.73	10.86
55	Denmark	36600	78.63	4.24	10.19
56	Djibouti	2800	61.14	54.94	8.23
57	Dominica	10400	75.98	12.78	8.06
58	Dominican_Republic	8900	77.31	22.22	4.35
59	Ecuador	7800	75.73	19.65	5.00
60	Egypt	6200	72.66	25.20	4.82
61	El_Salvador	7200	73.44	20.30	5.62
62	Equatorial_Guinea	36600	62.37	77.30	9.03
63	Eritrea	600	62.52	41.33	8.08
64	Estonia	19100	73.33	7.06	13.55
65	Ethiopia	1000	56.19	77.12	11.04
66	Faroe_Islands	32900	79.72	6.06	8.67
67	Finland	35400	79.27	3.43	10.24
68	France	33100	81.19	3.29	8.76
69	French_Polynesia	18000	77.10	7.27	4.87
70	Gabon	14500	52.49	49.95	13.00
71	Gambia_The	1900	63.51	71.67	7.65
72	Georgia	4900	77.12	15.17	9.92
73	Germany	35700	80.07	3.54	10.92
74	Ghana	2500	61.00	48.55	8.75
75	Gibraltar	43000	78.68	6.69	8.18
76	Greece	29600	79.92	5.00	10.70
77	Greenland	36500	70.96	10.05	8.12
78	Grenada	10200	73.04	11.43	7.94
79	Guatemala	5200	70.88	26.02	4.98
80	Guernsey	44600	82.16	3.55	8.44
81	Guinea	1000	58.11	61.03	10.45
82	Guinea-Bissau	1100	48.70	96.23	15.27
83	Guyana	7200	67.08	36.76	7.20
84	Haiti	1200	62.17	54.02	8.21
85	Honduras	4200	70.61	20.44	5.02
86	Hong_Kong	45900	82.04	2.90	7.07
87	Hungary	18800	74.79	5.31	12.68
88	Iceland	38300	80.90	3.20	6.96
89	India	3500	66.80	47.57	7.48
90	Indonesia	4200	71.33	27.95	6.26
91	Iran	10600	70.06	42.26	5.94
92	Iraq	3800	70.55	41.68	4.82
93	Ireland	37300	80.19	3.85	6.34
94	Isle_of_Man	35000	80.64	4.32	9.92

	Pays	PIB	EspVie	MortInf	Décès
157	Puerto_Rico	16300	78.92	8.07	7.95
158	Qatar	179000	75.70	12.05	2.43
159	Romania	11600	73.98	11.02	11.81
160	Russia	15900	66.29	10.08	16.04
161	Rwanda	1100	58.02	64.04	9.88
162	St_Helena_Ascen._&_Tristan_da_Cunha	2500	78.76	16.38	6.88
163	Saint_Kitts_and_Nevis	13700	74.60	9.66	7.10
164	Saint_Lucia	11200	76.84	12.72	7.00
165	Saint_Pierre_and_Miquelon	7000	79.87	7.47	8.83
166	Saint_Vincent_&_the_Grenadines	10300	74.15	14.27	6.98
167	Samoa	5500	72.40	22.74	5.34
168	San_Marino	36200	83.01	4.72	7.89
169	Sao_Tome_and_Principe	1800	63.11	53.21	8.18
170	Saudi_Arabia	24200	74.11	16.16	3.33
171	Senegal	1900	59.78	56.42	9.26
172	Serbia	10900	74.32	6.52	13.85
173	Seychelles	23200	73.52	11.66	6.91
174	Sierra_Leone	900	56.13	78.38	11.73
175	Singapore	62100	82.14	2.32	4.95
176	Slovakia	22000	75.83	6.59	9.60
177	Slovenia	28200	77.30	4.17	10.87
178	Solomon_Islands	2900	74.18	17.82	3.93
179	Somalia	600	50.40	105.56	14.87
180	South_Africa	10700	49.33	43.20	17.09
181	Spain	29400	81.17	3.39	8.80
182	Sri_Lanka	5000	75.73	9.70	5.92
183	Sudan	2300	55.42	68.07	11.00
184	Suriname	9700	74.22	17.61	5.54
185	Swaziland	4500	48.66	63.09	14.60
186	Sweden	39100	81.07	2.74	10.20
187	Switzerland	42600	81.07	4.08	8.72
188	Syria	4800	74.69	15.62	3.68
189	Taiwan	35700	78.32	5.18	7.00
190	Tajikistan	2000	66.03	38.54	6.60
191	Tanzania	1400	52.85	66.93	12.09
192	Thailand	8700	73.60	16.39	7.29
193	Timor-Leste	2600	67.95	38.01	5.89
194	Togo	900	62.71	51.48	7.96
195	Tonga	6100	75.16	13.65	4.90
196	Trinidad_and_Tobago	21200	71.37	27.69	8.29
197	Tunisia	9400	75.01	25.92	5.83
198	Turkey	12300	72.50	23.94	6.10
199	Turkmenistan	7500	68.52	42.34	6.24
200	Turks_and_Caicos_Islands	11500	79.11	11.97	2.99
201	Tuvalu	3400	64.75	34.52	9.20
202	Uganda	1300	53.24	62.47	11.71
203	Ukraine	6700	68.58	8.54	15.74

Tableau 6 (suite)

	Pays	PIB	EspVie	MortInf	Décès
95	Israel	29800	80.96	4.12	5.47
96	Italy	30500	81.77	3.38	9.84
97	Jamaica	8300	73.45	14.60	6.54
98	Japan	34000	82.25	2.78	10.09
99	Jersey	57000	81.38	3.98	7.52
100	Jordan	5400	80.05	16.42	2.69
101	Kazakhstan	12700	68.51	24.15	9.38
102	Kenya	1600	59.48	52.29	8.93
103	Kiribati	6200	64.39	38.89	7.40
104	Korea_North	1800	68.89	27.11	9.08
105	Korea_South	30000	79.05	4.16	6.26
106	Kuwait	48900	77.09	8.07	2.11
107	Kyrgyzstan	2200	70.04	29.27	6.79
108	Laos	2500	62.39	59.46	8.13
109	Latvia	14700	72.68	8.42	13.60

	Pays	PIB	EspVie	MortInf	Décès
204	United_Arab_Emirates	49600	76.51	11.94	2.06
205	United_Kingdom	34800	80.05	4.62	9.33
206	United_States	47200	78.37	6.06	8.38
207	Uruguay	13700	76.21	9.69	9.58
208	Uzbekistan	3100	72.51	21.92	5.29
209	Vanuatu	5100	64.70	46.85	7.43
210	Venezuela	12700	73.93	20.62	5.17
211	Vietnam	3100	72.18	20.90	5.96
212	Virgin_Islands	14500	79.33	7.24	7.17
213	Wallis_and_Futuna	3800	78.98	4.67	4.68
214	West_Bank	2900	75.01	14.92	3.58
215	Western_Sahara	2500	61.13	60.44	8.96
216	Yemen	2700	63.74	55.11	7.02
217	Zambia	1500	52.36	66.60	12.61
218	Zimbabwe	500	49.64	29.50	13.58

Tableau 7 Résultats d'une étude pédagogique

ID	Groupe	PRE1	PRE2	POST1	POST2	POST3	ID	Groupe	PRE1	PRE2	POST1	POST2	POST3
1	Basal	4	3	5	4	41	34	DRTA	6	2	7	0	55
2	Basal	6	5	9	5	41	35	DRTA	8	4	10	6	57
3	Basal	9	4	5	3	43	36	DRTA	9	6	8	6	53
4	Basal	12	6	8	5	46	37	DRTA	9	4	8	7	37
5	Basal	16	5	10	9	46	38	DRTA	8	4	10	11	50
6	Basal	15	13	9	8	45	39	DRTA	9	5	12	6	54
7	Basal	14	8	12	5	45	40	DRTA	13	6	10	6	41
8	Basal	12	7	5	5	32	41	DRTA	10	2	11	6	49
9	Basal	12	3	8	7	33	42	DRTA	8	6	7	8	47
10	Basal	8	8	7	7	39	43	DRTA	8	5	8	8	49
11	Basal	13	7	12	4	42	44	DRTA	10	6	12	6	49
12	Basal	9	2	4	4	45	45	Strat	11	7	11	12	53
13	Basal	12	5	4	6	39	46	Strat	7	6	4	8	47
14	Basal	12	2	8	8	44	47	Strat	4	6	4	10	41
15	Basal	12	2	6	4	36	48	Strat	7	2	4	4	49
16	Basal	10	10	9	10	49	49	Strat	7	6	3	9	43
17	Basal	8	5	3	3	40	50	Strat	6	5	8	5	45
18	Basal	12	5	5	5	35	51	Strat	11	5	12	8	50
19	Basal	11	3	4	5	36	52	Strat	14	6	14	12	48
20	Basal	8	4	2	3	40	53	Strat	13	6	12	11	49
21	Basal	7	3	5	4	54	54	Strat	9	5	7	11	42
22	Basal	9	6	7	8	32	55	Strat	12	3	5	10	38
23	DRTA	7	2	7	6	31	56	Strat	13	9	9	9	42
24	DRTA	7	6	5	6	40	57	Strat	4	6	1	10	34
25	DRTA	12	4	13	3	48	58	Strat	13	8	13	1	48
26	DRTA	10	1	5	7	30	59	Strat	6	4	7	9	51
27	DRTA	16	8	14	7	42	60	Strat	12	3	5	13	33
28	DRTA	15	7	14	6	48	61	Strat	6	6	7	9	44
29	DRTA	9	6	10	9	49	62	Strat	11	4	11	7	48
30	DRTA	8	7	13	5	53	63	Strat	14	4	15	7	49
31	DRTA	13	7	12	7	48	64	Strat	8	2	9	5	33
32	DRTA	12	8	11	6	43	65	Strat	5	3	6	8	45
33	DRTA	7	6	8	5	55	66	Strat	8	3	4	6	42

Tableau 8 Données statistiques sur un ensemble de pays

Pays	Pib	Espvie	Mort	Pays	Pib	Espvie	Mort
Algérie	1726	20	40	Luxembourg	44797	24	6
Australie	21319	25	5	Malawi	174	17	128
Autriche	25748	24	4	Malaisie	3613	19	9
Bahamas	13302	22	15	Maldives	1382	17	43
Bahrain	9369	20	13	Mali	254	23	116
Bangladesh	291	16	68	Mauritanie	313	16	89
Barbades	9380	22	10	Mexique	5036	22	26
Bolivie	1032	18	51	Namibie	1734	18	61
Botswana	3625	18	63	Népal	218	16	73
Brésil	3525	20	32	Pays-bas	24929	24	4
Bulgarie	1543	19	13	Antil._holl.	11783	22	10
Burkina_Faso	234	16	81	Nlle_Zélande	14754	24	6
Burundi	128	16	103	Nigéria	473	17	76
Canada	20822	24	5	Norvège	34377	24	4
Cap_Vert	1400	21	47	Pakistan	487	17	90
Rép._cent._af.	277	17	82	Panama	3397	22	17
Chile	4505	22	11	Paraguay	1445	20	32
Chine	798	20	42	Pérou	2060	20	33
Colombie	2093	21	22	Philippines	1032	18	25
Costa_Rica	2942	22	9	Pologne	3991	21	9

Tableau 8 (Suite)

Pays	Pib	Espvie	Mort	Pays	Pib	Espvie	Mort
Côte_d'Ivoire	808	17	75	Portugal	11229	23	6
Croatie	4242	21	7	Qatar	21220	15	9
Chypres	11715	23	8	Roumanie	1392	19	21
Danemark	32853	22	5	Rouanda	217	16	112
Équateur	1109	21	37	Singapoure	22072	23	4
Égypte	1307	17	38	Slovaquie	3492	23	5
Le_Salvador	2007	21	24	Slovénie	10052	23	5
Éthiopie	101	16	99	Î._Salomon	801	17	21
Fiji	2275	18	19	Afr._du_sud	3067	18	55
Finlande	25112	24	4	Espagne	14939	25	5
France	24267	26	5	Sri_Lanka	836	20	14
Allemagne	25749	24	4	Suriname	1657	19	20
Grèce	11811	24	6	Swaziland	1304	17	84
Guatemala	1637	19	37	Suède	26968	25	3
Haïti	496	16	57	Suisse	36031	25	4
Honduras	856	20	29	Syrie	2525	19	21
Hongrie	4813	20	8	Thaïlande	2000	20	16
Indonésie	674	18	34	Trin._et_Tob.	5119	21	10
Iran	3445	18	37	Tunisie	2247	19	24
Iraq	3144	19	62	Turquie	2813	20	33
Irelande	24825	22	6	Ouganda	301	16	87
Israël	17564	24	6	Ém._ar._unis	19700	23	11
Italie	20355	24	5	Roy.-Uni	24323	23	5
Japon	34276	27	3	É.-U.A.	32778	24	7
Kazakstan	982	19	34	Uruguay	5891	23	11
Koweït	16244	22	11	Vénézuéla	4312	21	16
Lesotho	460	17	108	Zambie	323	16	76
Porto_Rico	16868	24	9	Zimbabwe	436	18	51