1. **La corrélation entre les variables suivantes est-elle positive, négative ou nulle ?**
2. Revenu personnel et total des impôts personnels à payer. *Positive et fort*
3. Poids de l’homme et poids de la femme dans un couple. *Nulle (ou positif faible)*
4. Âge de l’homme et âge de la femme dans un couple. *Positive moyen*
5. **Quelle serait la valeur du coefficient de corrélation, r, calculé à partir d’un nuage de points ne contenant que 2 points ?** *Elle serait de 1 ou -1, car toutes les coordonnées (les deux) sont sur la droite de régression, car ces deux points forment la droite de régression (ils forment directement la pente).*
6. Cinq clients d’Énergie Nouveau-Brunswick ont payé les montants suivants pour le nombre de kilowattheures (kWh) consommés :

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| kWh consommés | 1600 | 800 | 1250 | 1700 | 940 |
| Coûts ($) | 109,40 | 66,72 | 90,53 | 114,33 | 74,13 |

1. **Calculer et interpréter le coefficient de corrélation.**

|  |  |  |  |
| --- | --- | --- | --- |
|  | *kWh* | *Coût* | *x\*y* |
|  | *1600* | *109,4* | *175040* |
|  | *800* | *66,72* | *53376* |
|  | *1250* | *90,53* | *113162,5* |
|  | *1700* | *114,33* | *194361* |
|  | *940* | *74,13* | *69682,2* |
| *moyenne* | *1258* | *91,022* |  |
| *écart-type* | *394,740421* | *20,9593814* |  |
| *somme*  |  |  | *605621,7* |

*Coefficient de corrélation, r :* $\frac{\sum\_{}^{}xy-n\*\overbar{x}\*\overbar{y}}{\left(n-1\right)\*S\_{x}\*S\_{y}}$ *=* $\frac{605621,7-5\*1258\*91,022}{4\*394,740421\*20,9593814}$ *= 0,99997763 ≈ 1*

*Le coefficient de corrélation est de 1, ce qui veut dire qu’il y a une corrélation parfaite et positive entre les variables. Les coordonnées se retrouvent sur la droite de régression. En d’autres mots, On peut estimer une quantité d’énergie utilisée en sachant le coût ou vice-versa.*

1. **À quoi ressemblerait le diagramme de dispersion ?**
2. **Trouver l’équation de la droite de régression.**

*m =*$\frac{\sum\_{}^{}xy-n\*\overbar{x}\*\overbar{y}}{\left(n-1\right)\*S\_{x}^{2}}$ *=* $\frac{605621,7-5\*1258\*91,022}{4\*394,740421^{2}}$ *≈ 0,053*

*b =* $\overbar{y}$ *- m*$\overbar{x}$ *= 91,022 - 0,053 x 1258 ≈ 24,2279*

*L’équation de la droite de régression serait : y = 0,053x + 24,2279*

1. **Estimer le montant qu’il faudrait payer pour une consommation de 1 400 kWh.**

*y = 0,053x + 24,2279 🡪 y = 0,053(1400) + 24,2279 = 98,4279.*

*Cela lui couterait environ 98,43$ pour une consommation de 1 400 kWh.*

1. **Estimer la consommation d’une cliente qui paye 132,05 $ pour sa facture d’électricité.**

*y = 0,053x + 24,2279 🡪 132,05 = 0,053x + 24,2279 🡪 107,8221 = 0,053x 🡪 x = 2 034,38 kWh*

*Pour une facture de 132,05$, elle devrait consommer environ 2 034,38 kWh.*

1. **Établir la signification des valeurs de m et de b de l’équation de la droite de régression dans le contexte de ce problème.**

*Le « b »,l’ordonnée à l’origine, représenterait le tarif de base, sans même utiliser des kWh cela lui coûte 24,23$ et le « m » représente le taux de variation, chaque kWh leur coûte environ 0,053 $.*

1. Jonathan est un artisan qui fabrique et vend des sculptures sur bois. Il note le nombre de sculptures vendues au cours des 11 dernières années. Dans l’ordre chronologique :

*76, 78, 70, 71, 64, 62, 69, 56, 65, 49, 52*

**Prévois le nombre de sculpture qu’il vendra au cours des 3 prochaines années.**

|  |  |  |  |
| --- | --- | --- | --- |
|  | *Année* | *Nombre de sculptures vendus* | *x\*y* |
|  | *1* | *76* | *76* |
|  | *2* | *78* | *156* |
|  | *3* | *70* | *210* |
|  | *4* | *71* | *284* |
|  | *5* | *64* | *320* |
|  | *6* | *62* | *372* |
|  | *7* | *69* | *483* |
|  | *8* | *56* | *448* |
|  | *9* | *65* | *585* |
|  | *10* | *49* | *490* |
|  | *11* | *52* | *572* |
| *moyenne* | *6* | *64,72727273* |  |
| *écart-type* | *3,31662479* | *9,39245345* |  |
| *somme*  |  |  | *3996* |

**Équation de la droite de régression :**

*m =*$\frac{\sum\_{}^{}xy-n\*\overbar{x}\*\overbar{y}}{\left(n-1\right)\*S\_{x}^{2}}$ *=* $\frac{3996 - 11\*6\*64,72727273}{10\*3,31662479^{2}}$ *≈ - 2,509*

*b =* $\overbar{y}$ *- m*$\overbar{x}$ *= 64,72727273 - 2,509 x 6 ≈ 79,7818*

*y = 2,509x + 79,7818*

*12e année : y = -2,509(12) + 79,7818 = 49,6738 ≈50 sculptures*

*La 12e année, il devrait vendre environ 50 sculptures.*

*13e année : y = -2,509(13) + 79,7818 = 47,1648 ≈ 47 sculptures*

*La 13e année, il devrait vendre environ 47 sculptures.*

*14e année : y = -2,509(14) + 79,7818 = 44,6558 ≈ 45 sculptures*

*La 14e année, il devrait vendre environ 45 sculptures.*

1. Afin d’estimer les coûts de chauffage d’un immeuble, on note quotidiennement le nombre de litre de mazout consommé et la température extérieure moyenne.
2. **La corrélation entre X et Y devrait-elle être fortement positive, faiblement positive, fortement négative ou faiblement positive ?**

*Cette relation devrait être fortement négative, car plus il fait chaud, on utilise moins de mazout (ce qui fait que le prix est plus bas) et vice-versa.*

1. **Cette relation est-elle causale ou non-causale ?**

*Cette relation est causale, la température influence le thermostat de la maison, qui à son tour utilise le mazout.*

1. Le conseiller pédagogique du CCNB croit que les résultats des élèves dans leur cours de français est en relation avec leur résultats dans le cours de mathématiques. En choisissant un échantillon de dix dossiers d’élèves, il note les données suivantes :

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Élève* | *Note en français* | *Note en mathématiques* | *x\*y* |
|  | *1* | *70* | *74* | *5180* |
|  | *2* | *66* | *68* | *4488* |
|  | *3* | *82* | *90* | *7380* |
|  | *4* | *76* | *73* | *5548* |
|  | *5* | *91* | *92* | *8372* |
|  | *6* | *69* | *74* | *5106* |
|  | *7* | *81* | *86* | *6966* |
|  | *8* | *88* | *91* | *8008* |
|  | *9* | *63* | *71* | *4473* |
|  | *10* | *87* | *84* | *7308* |
| *moyenne* |  | *77,3* | *80,3* |  |
| *écart-type* |  | *9,9336689* | *9,20205291* |  |
| *somme*  |  |  |  | *62829* |

**Est-ce que le conseiller a raison en mentionnant qu’il y a une relation entre ces résultats ? Démontre à l’aide de calculs.**

*Coefficient de corrélation, r :* $\frac{\sum\_{}^{}xy-n\*\overbar{x}\*\overbar{y}}{\left(n-1\right)\*S\_{x}\*S\_{y}}$ *=* $\frac{62829-10\*77,3\*80,3}{9\*9,9336689\*9,20205291 }$ *= 0,92027226 ≈ 0,92*

*Le conseiller peut affirmer ceci, car la corrélation entre les deux résultats est assez forte, avec un coefficient de corrélation d’environ 0,92.*

1. Un sociologue fait une étude dans un quartier de la ville. Il vérifie s’il y a un lien entre le nombre d’années de scolarité et le salaire en milliers de dollars. Voici les données recueillies chez neuf personnes de ce quartier.
2. **Trace un diagramme de dispersion.**
3. **Trouver l’équation de la droite de régression et tracer cette droite.**

|  |  |  |  |
| --- | --- | --- | --- |
|  | *Scolarité* | *Salaire (millier de $)* | *x\*y*  |
|  | *10* | *35* | *350* |
|  | *17* | *51* | *867* |
|  | *12* | *39* | *468* |
|  | *13* | *44* | *572* |
|  | *6* | *31* | *186* |
|  | *18* | *57* | *1026* |
|  | *11* | *40* | *440* |
|  | *7* | *26* | *182* |
|  | *14* | *46* | *644* |
| *Moyenne* | *12* | *41* |  |
| *Écart-type* | *4,0620192* | *9,72111105* |  |
| *Somme* |  |  | *4735* |

*m =*$\frac{\sum\_{}^{}xy-n\*\overbar{x}\*\overbar{y}}{\left(n-1\right)\*S\_{x}^{2}}$ *=* $\frac{4735-9\*12\*41}{8\*4,0620192^{2}}$ *= 2,32575758 ≈ 2,326*

*b =* $\overbar{y}$ *- m*$\overbar{x}$ *= 41 - 2,32575758 x 12 = 13,0909091 ≈ 13,09*

*Équation de la droite de regression : y = 2,326x + 13,09*

1. **Prévoir le salaire d’une personne qui a 15 ans de scolarité**

*y = 2,326(15) + 13,09 ≈ 47,98. Une personne ayant 15 années de scolarité possèdera un salaire approximatif de 47 980 $.*

1. **Prévoir la scolarité d’une personne qui a 42 000 $ de salaire.**

*42 = 2,326x + 13,09 🡪 28,91 = 2,326x 🡪* *x ≈12,42906. Une personne ayant un salaire de 42 000 $ devrait avoir un niveau de scolarité d’environ 12,4 ans.*

1. **Calculer le coefficient de corrélation. Que peut-on conclure ?**

*Coefficient de corrélation, r :* $\frac{\sum\_{}^{}xy-n\*\overbar{x}\*\overbar{y}}{\left(n-1\right)\*S\_{x}\*S\_{y}}$ *=*$\frac{4735-9\*12\*41}{8\*4,0620192\*9,72111105}$ *= 0,97183047 ≈ 0,97*

*Le coefficient de corrélation est d’environ 0,97, ce qui veut dire que la corrélation est très forte. Les points sont situés assez proche de la droite de régression. Les approximations sont assez près de la réalité et on peut se fier les ces approximations.*

1. **Qu’est-ce qui est la différence entre une relation causale et une relation non-causale ? Explique la différence et donne quelques exemples.**

***Une relation causale est une relation qu’une variable est la « cause de l’autre ».*** *Par exemple, la scolarité et le salaire est causale, plus tu as d’années d’études, plus ton salaire devrait être élevé. Le salaire est la cause du nombre d’année d’étude. Le temps d’étude / pratique le résultat sur une évaluation / statistique est aussi causale. Une personne qui étudie/se pratique, devrait mieux performer. Le résultat est causé par la préparation d’une personne (en général).*

***Une relation non causale est une relation qui est en lien avec un facteur externe ou un autre facteur, la relation est illusoire.*** *Il n’y a pas de cause à* ***effet directe*** *entre ces deux variables. Par exemple, le nombre d’école dans une région et la quantité d’ours dans cette région. Ou encore, le nombre de décès dans une région et le prix d’une pomme.*