Exercice 1

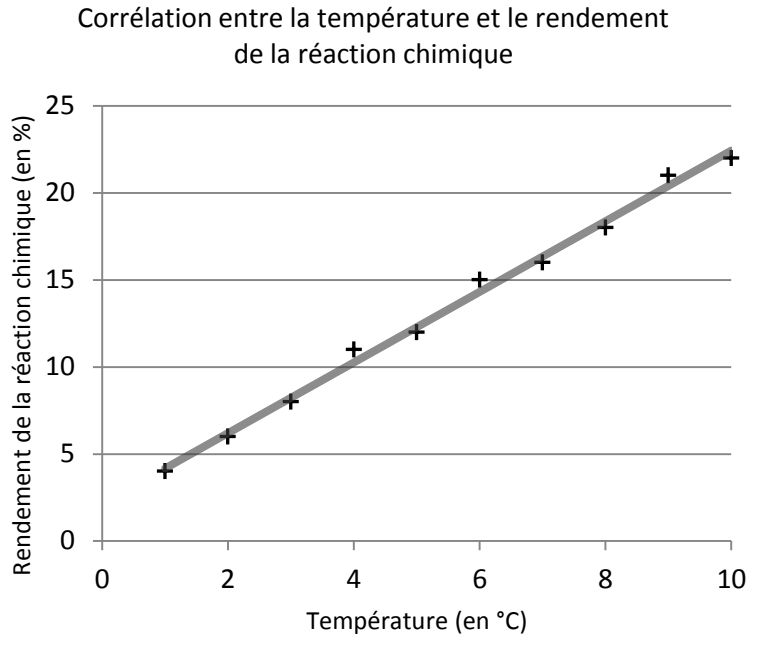
Question	Proposition de correction
1	82,5 % des salariés gagnent moins de 1050 euros (mais on ne sait pas exactement quelle est la proportion de salaires inférieurs ou égaux à 1050 euros, comprise entre 82,5 % et 92 % des salariés d'après les informations disponibles).
2	Le salaire moyen est de 993 euros par mois.
3	La médiane est comprise entre 1000 et 1050 euros. On peut approximer le salaire médian à environ 1006,08 euros : $$Me = 1000 + \frac{0{,}5 - 0{,}455}{0{,}37} * 50 \approx 1006$$ La moyenne est supérieure à la médiane de 1,3 % environ (signe d'une distribution légèrement asymétrique, étalée à gauche).
4	25 % des salaires sont supérieurs à 916 euros et 25 % supérieurs à 1040 euros.
5	Titre : Distribution des salaires dans l'entreprise étudiée.

Exercice 2

Question	Proposition de correction
1	Comparer la dispersion des salaires dans les deux distributions implique l'utilisation du coefficient de variation. Pour l'établissement E1 : $$CV_{\mathrm{E1}} = \frac{800}{1500} \approx 53{,}3\%$$ Pour le second établissement, le coefficient de variation des salaires vaut 44 %. Les salaires sont donc plus dispersés dans l'établissement E1.
2	Etendue = 5 clients Ecart interquartile = 2 clients Ecart absolu moyen = 1,1 client Écart-type = 1,4 clients *(Rappel : les calculs doivent apparaître sur la copie.)* L'étendue donne l'écart entre le plus petit nombre et le plus grand nombre d'arrivées de clients observées, qui est de 5 clients, et ne tient pas compte de la répartition des arrivées de clients entre ces valeurs extrêmes. Lors de la moitié environ des périodes on observe entre 2 et 4 clients qui arrivent au bureau de poste. L'EIQ valant 2 arrivées de clients est l'ampleur de l'intervalle dans lequel se situent 50 % des observations, en éliminant le quart des nombres de clients observés les plus faibles et le quart des observations les plus élevées. L'avantage est que cette mesure ne repose pas sur les valeurs extrêmes de la distribution mais elle tient imparfaitement compte de l'ensemble des valeurs. On observe en moyenne 3 entrées de clients par période de 10 minutes et le nombre de clients observé au cours d'une période s'écarte en moyenne de 1,1 par rapport à 3. Cet indicateur a l'avantage de prendre en compte les écarts entre chaque observation et la moyenne. En moyenne le nombre de clients qui arrivent en 10 minutes a tendance à se situer entre 2 et 4 (conclusion comparable à celle obtenue avec l'EIQ). L'écart-type est une moyenne quadratique des écarts à la moyenne. La valeur de l'EAM repose sur une moyenne arithmétique dont le sens est plus facile à interpréter. Cependant l'écart-type est un indicateur de dispersion plus fréquemment utilisé que l'EAM.

Exercice 3

Question	Proposition de correction
1	Faux. Par exemple les 20 % les mieux payés touchent 60 % de la masse salariale. Autrement dit le 1/5^{e} des salaires les plus hauts représentent 3/5^{e} de la masse salariale, ce n'est pas une distribution « très peu concentrée ».
2	Vrai. Comme la concentration de la masse salariale est nette (question 1), il y a des différences relatives importantes entre ces salaires.
3	Vrai. La surface sous la courbe de Lorenz est à peu près égale à la surface de concentration.
4	Vrai. La concentration est faible avec une courbe de Lorenz presque confondue avec la bissectrice représentant une répartition parfaitement égalitaire des salaires. Ou : Faux. La concentration est faible mais la répartition des salaires n'est pas parfaitement égalitaire puisque la courbe de Lorenz n'est pas confondue avec la bissectrice.
5	Faux. L'indice de Gini est proche de zéro puisque la concentration est quasi-nulle (question 4).
6	Vrai. Entreprises de grande taille, avec hiérarchie marquée et donc postes à haute responsabilité relativement bien payés, entreprises qui ont des activités dans différents pays où les salaires peuvent être très inégaux, etc. Ou : On ne sait pas. Aucune indication ne figure sur les graphiques.

Exercice 4

Question	Proposition de correction
1	Coefficient de corrélation linéaire : $$r = \frac{16{,}75}{\sqrt{8{,}25} * \sqrt{34{,}21}} \approx 0{,}997$$ Une relation très forte existe entre les deux variables puisque *r* est très proche de 1. Presque une relation fonctionnelle.
2	Régression de y en x : $$b_1 = \frac{16{,}75}{8{,}25} \approx 2{,}03$$ $$b_0 = 13{,}3 - b_1 * 5{,}5 \approx 2{,}133$$ $$\hat{y}_i \approx 2{,}03x_i + 2{,}135$$ *(Résultats utilisés pour les deux questions suivantes.)* Ou : Régression de x en y : $$b_1{}' = \frac{16{,}75}{34{,}21} \approx 0{,}49$$ $$b_0{}' = 5{,}5 - b_1 * 13{,}3 \approx -1{,}01$$ $$\hat{x}_i \approx 0{,}49y_i - 1{,}01$$
3	Quand la température est de 0°C, alors le rendement de la réaction est estimé, grâce à l'équation de régression, à 2 % environ. Cependant nous n'avons fait aucun test pour ce niveau de température, l'estimation est peu fiable. Quand la température croît de 1°C, le rendement de la réaction augmente de 2 points de pourcentage.
4