

PROJET DE REGRESSION LINEAIRE

Licence 3-Université des Lagunes

- Groupe alpha (α) : KOUASSI AKRE / ANOH LINDAH
- Groupe beta (β) : KRA KOUAME GERARD / TOGBE JOSEPH

INSTRUCTIONS :

A rédiger sous forme de rapport avec **R Markdown**. Prière de présenter toutes les réponses et résultats dans le rapport. La rédaction et la présentation seront notées. Prière d'imprimer et faire relier le rapport.

Date limite : Vendredi 14 Juillet 2023 à 13h30 mn.

Email : jptchiekre@gmail.com

L'entreprise INFORMATEX se spécialise dans l'analyse de systèmes et la programmation sur ordinateur de problèmes techniques et de gestion. Elle veut utiliser la régression dans une étude sur le temps requis, par ses analystes-programmeurs, pour programmer des projets complexes. Cette étude pourrait permettre à la firme d'établir des normes quant au temps requis pour programmer certains projets et d'assurer éventuellement une meilleure planification des ressources humaines. Les données du tableau suivant représentent le temps total en heures requis pour programmer différents projets en fonction du nombre d'instructions dans chaque programme.

Temps total en heures	40	55	62	58	82	94	120
Nombre d'instructions	60	82	100	142	190	220	285
Temps total en heures	134	128	140	152	174	167	218
Nombre d'instructions	354	400	425	440	500	530	640

1. Si nous voulons expliquer les fluctuations dans le temps requis pour programmer les projets quelle variable devons-nous identifier comme variable dépendante ? Comme variable explicative ?
2. Qu'est-ce qui peut renseigner l'entreprise sur la forme de liaison statistique qui peut exister entre ces deux variables ?
3. Quelle méthode d'ajustement linéaire devons-nous utiliser pour obtenir les estimateurs des coefficients de la droite de régression ?
4. Donner les valeurs de $\hat{\beta}_0$ et de $\hat{\beta}_1$.
5. Quelle est l'équation de la droite des moindres carrés ?
6. Si nous ne tenons pas compte du nombre d'instructions, quelle valeur pourrions-nous utiliser comme estimation du temps moyen de programmation des projets ?
7. Quelle correction pouvons-nous apporter à l'estimation obtenue en 6., en tenant compte du nombre d'instructions par l'entremise de la droite des moindres carrés ?
8. D'après la droite des moindres carrés, à quelle augmentation du temps de programmation pouvons-nous nous attendre lorsque le nombre d'instructions augmente de 50 ?
9. Pour chaque nombre d'instructions suivant, estimer le temps de programmation à l'aide de la droite des moindres carrés
10. Selon les résultats observés, quels sont les écarts de prévision de l'équation des moindres carrés pour le nombre d'instruction en 9. ?
11. Si nous avons utilisé l'estimation obtenue en 6. au lieu de celles déduites de l'équation des moindres carrés pour effectuer les prévisions selon le nombre d'instructions spécifié en 9., quels auraient été alors, dans chaque cas, les écarts de prévision ?
12. Pour chaque valeur x_i spécifié en 9., vérifier la relation

$$y_i - \bar{y} = (\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i).$$

13. Calculer la variation totale, la variation expliquée par la droite des moindres carrés et la variation résiduelle.
14. Quelle proportion de la variation totale dans le temps de programmation est expliquée par la droite des moindres carrés ? Quelle proportion demeure inexpliquée par la droite ?
15. Nous avons fixé le R^2 à 0.90 comme valeur minimale pour considérer la droite des moindres carrés d'utilité pratique. D'après les résultats obtenus, devrions-nous utiliser la droite des moindres carrés comme outil de prévision ?