Mastère BI **Série 1**

Problem 1

Student	GPA	ACT
1	2.8	21
2	3.4	24
3	3.0	26
4	3.5	27
5	3.6	29
6	3.0	25
7	2.7	25
8	3.7	30

The above table contains the ACT scores and the GPA (grade point average) for eight college students. Grade point average is based on a four-point scale and has been rounded to one digit after the decimal.

1. Estimate the relationship between GPA and ACT using OLS; that is, obtain the intercept and slope estimates in the equation.

$$GPA = \beta_0 + \beta_1 ACT.$$

   Comment on the direction of the relationship. Does the intercept have a useful interpretation here? Explain. How much higher is the GPA predicted to be if the ACT score is increased by five points?

2. Compute the fitted values and residuals for each observation, and verify that the residuals (approximately) sum to zero.

3. What is the predicted value of GPA when ACT=20?

4. How much of the variation in GPA for these eight students is explained by ACT? Explain.

Problem 2

The data set BWGHT.RAW contains data on births to women in the United States. Two variables of interest are the dependent variable, infant birth weight in ounces (*bwght*), and an explanatory variable, average number of cigarettes the mother smoked per day during pregnancy (*cigs*). The following simple regression was estimated using data on $n = 1388$ births:

$$\widehat{bwght} = 119.77 - 0.514 cigs$$

1. What is the predicted birth weight when *cigs* = 0? What about when *cigs* = 20 (one pack per day)? Comment on the difference.
2. Does this simple regression necessarily capture a causal relationship between the child's birth weight and the mother's smoking habits? Explain.
3. To predict a birth weight of 125 ounces, what would *cigs* have to be? Comment.
4. The proportion of women in the sample who do not smoke while pregnant is about .85. Does this help reconcile your finding from part (iii)?

Problem 3

On s'intéresse dans un secteur de production à la relation entre les bénéfices réalisés par les entreprises © et le budget annuel qu'elles consacrent à la publicité (p).

P	15	8	36	41	16	8	21	21	53	10	32	17	58	6	20
C	48	43	77	89	50	40	56	62	100	47	71	58	102	35	60

1. Si on veut faire établir une régression linéaire entre les deux variables, quelle doit être la variable endogène.
2. On admet l'existence d'une relation linéaire. Calculer alors les estimations des coefficients du modèle.
3. Calculer le coefficient de corrélation de Pearson.
4. Donner l'équation de la variance et en déduire le coefficient de détermination.
5. Sachant que la variance estimée des termes d'erreur est de 10,155, procéder à l'estimation des variances des coefficients.
6. Déterminer au seuil de 0,05 un intervalle de confiance aux coefficients et pour la variance des termes d'erreur estimée.
7. Peut-on affirmer que les coefficients sont sifnificativement différents de 0.
8. Déterminer un intervalle de confiance pour le bénéfice prévisible relatif à une entreprise qui consacre un budget de 48 à son programme publicitaire.