

- 3.1 Le jeu de donnée `croissance` contient la masse (en livres) et la taille (en pouces) de 19 adolescents. Notre objectif est de prédire la masse d'adolescents sur la base de leur taille à l'aide d'un modèle linéaire simple.
- (a) Calculez les statistiques descriptives de l'échantillon (moyenne, écart-type, quartiles, etc.) des deux variables et inspectez visuellement l'effet de la taille sur la masse à l'aide d'un nuage de points.
 - (b) Calculez le coefficient de corrélation entre la taille et la masse et interprétez ce dernier.
 - (c) Ajustez un modèle de régression linéaire pour prédire la masse en fonction de la taille. Écrivez l'équation du modèle de régression ainsi que celle du modèle ajusté.
 - (d) Interprétez les paramètres du modèle, β_0 et β_1 .
 - (e) Quelles sont les hypothèses nulles et alternatives des tests- t de la sortie? Écrivez les conclusions de ces tests.
 - (f) Donne un intervalle à 95% pour la pente et interprétez ce dernier.
 - (g) Prédisez la masse de deux adolescents hypothétiques de 70,5 et 71,5 pouces. Que représente la différence entre ces valeurs?
 - (h) Fournissez des intervalles de confiance pour la masse moyenne et la masse prédite pour deux individus de taille 70,5 et 71,5 pouces. Commentez sur les différences entre ces deux intervalles.
 - (i) Supposez qu'on ajuste un modèle de régression pour la masse (en kg) et la taille (en cm). Est-ce que le modèle linéaire serait toujours adéquat? Expliquez en quoi le changement d'unité affecterait les estimés de vos paramètres.
- 3.1 Les données `eolienne` contiennent des mesures de la production électrique d'éoliennes sur 25 périodes non-consécutives de 15 minutes. Nous sommes intéressés à modéliser la relation entre la production électrique et la vitesse du vent moyenne (mesurée en miles à l'heure) pendant la période de mesure.
- (a) Ajustez un modèle linéaire avec la vitesse du vent comme covariable et produisez un graphique des résidus contre les valeurs ajustées. Est-ce que vous remarquez une structure résiduelle qui n'est pas prise en compte dans votre modèle? Essayez aussi un modèle avec la réciproque de la vitesse du vent comme variable explicative. Commentez sur l'adéquation des deux modèles.
 - (b) Prédisez, en utilisant les deux modèles à tour de rôle, la production électrique sachant que la vitesse du vent moyenne dans une période donnée est de 5 miles à l'heure. Fournissez également des intervalles de prédiction pour vos estimés.
 - (c) [★] La production électrique de l'éolienne devrait être inexistante en l'absence de vent, mais cette réalité n'est pas capturée par le premier modèle liant la production électrique à la vitesse du vent. Mettez votre modèle à jour en retirant l'ordonnée à l'origine (avec `~-1` dans R ou l'option `no int` dans SAS avec `proc glm`). Qu'arrive-t-il si vous retirez l'ordonnée à l'origine?
 - (d) Produisez un diagramme quantile-quantile des résidus studentisés externes et commentez sur l'hypothèse de normalité.
- 3.1 Le jeu de données `auto` contient les caractéristiques de 392 voitures. Nous sommes intéressés à modéliser la consommation d'essence (en miles par gallon) de voitures en fonction de leur puissance (en watts).
- (a) Ajustez un modèle linéaire pour les données, regardez le diagramme quantile-quantile des résidus studentisés externes et commentez sur l'adéquation du postulat de normalité.
 - (b) Produisez un nuage de points des résidus contre la variable puissance et commentez sur l'adéquation de l'hypothèse de linéarité.