

SCI03 Automne 2014

TP 4 : Régression linéaire

On souhaite faire une étude statistique des données contenues dans le fichier `pub.txt`. Voici une description du contenu de ce fichier :

This data appeared in the Wall Street Journal. The advertisement were selected by an annual survey conducted by Video Board Tests, Inc., a New York ad-testing company, based on interviews with 20,000 adults who were asked to name the most outstanding TV commercial they had seen, noticed, and liked. The retained impressions were based on a survey of 4,000 adults, in which regular product users were asked to cite a commercial they had seen for that product category in the past week.

Number of cases : 21

Variable Names :

- *FIRM : Firm name*
- *SPEND : TV advertising budget, 1983 (\$ millions)*
- *MILIMP : Millions of retained impressions per week*

1. Semble-t-il exister une relation entre les variables SPEND et MILIMP ? (faire un ou plusieurs graphiques).
2. Faire une régression linéaire en prenant MILIMP comme variable à expliquer et SPEND comme variable explicative :
 - (a) Donner des estimations ponctuelle et par intervalle de confiance des coefficients de régression ;
 - (b) Représenter graphiquement la droite des moindres carrés.
 - (c) Le R^2 est-il significativement différent de zéro ?
 - (d) Quel serait l'impact d'une campagne de publicité de 80 millions de \$? (donner un intervalle de confiance et un intervalle de prédiction) ;
3. Le modèle linéaire est-il adéquat pour ces données ? (Analyser les résidus).
4. Recommencer l'analyse précédente en prenant comme variable à expliquer $\log(\text{MILIMP})$. Discuter la validité du modèle.