**TD1 Statistiques multivariées et Programmation R**

**Corrélations et Régressions linéaires**

1) A l'aide de la fonction « read.csv », chargez les données des IRIS parisiens dans R.

2) Affichez les premières lignes des données attributaires de ces IRIS afin de vérifier le bon chargement des données.

3) Récupérez dans trois variables distinctes, les taux de familles monoparentales, les taux d'ouvriers et les revenus moyens des IRIS parisiens.

4) Affichez le nuage de points représentant en abscisse les taux d'ouvriers des IRIS parisiens et en ordonnée les revenus moyens des IRIS parisiens.

5) A l'aide de la fonction « lm », créez le modèle linéaire cherchant à expliquer les revenus moyens des IRIS parisiens (variable à expliquer) par les taux d'ouvriers des IRIS parisiens (variable explicative). Puis affichez le résumé du résultat obtenu.

6) En vous aidant de la fonction « coef » ou de la fonction « names », récupérez dans une variable nommée « a » le coefficient directeur de la droite de régression et dans une variable « b » la valeur d'intersection de l'axe des ordonnées.

7) Sur le nuage de points représentant en abscisse les taux d'ouvriers des IRIS parisiens et en ordonnée les revenus moyens des IRIS parisiens, représentez la droite de la régression linéaire à l'aide de la fonction « abline » (essayez de la représenter en pointillé et en rouge).

8) Appliquez la fonction « plot » directement sur la variable stockant la régression linéaire. Cliquez sur les graphiques.

9) Utilisez la fonction « cor » afin de comparer le $R^2$ calculé avec la méthode de Pearson et la méthode de Spearman.

10) A l'aide de la fonction « log », calculez les logarithmes des revenus moyens des IRIS parisiens, puis affichez un nuage de points avec cette variable en ordonnée et avec en abscisse les taux d'ouvriers des IRIS parisiens.

11) Etudiez la relation linéaire entre les logarithmes des revenus moyens des IRIS parisiens (variable à expliquer) et les taux d'ouvriers des IRIS parisiens (variable explicative).

12) Ajoutez à la relation linéaire précédente une deuxième variable explicative quantitative (les taux de familles monoparentales des IRIS parisiens) toujours à l'aide de la fonction « lm ».

13) Réalisez une Anova en utilisant la fonction « anova » afin de comparer les deux modèles explicatifs créés pour expliquer les logarithmes des revenus moyens des IRIS parisiens.