

# Projet de Traitement Données sur la régression. 01/12/16. BG et GO.

---

1. Etudier les données de pollution airquality.
  - a. Commenter Pairs.
  - b. Calculer les moyennes, écart-types, médianes, écart-interquartiles, histogrammes,
  - c. Calculer les corrélations des variables numériques. [extension passer à la corrélation partielle]
  - d. Etudier les données manquantes

Pour les variables de DS ou celles de scale(DS) répondre aux questions suivantes

2. Régresser l'ozone sur la température par une régression linéaire simple.
3. Les grandes valeurs sont mal modélisées passer à une régression par un polynôme de degré 2 ou bien faire une modélisation par zones. Dans ce dernier cas, sépare les jours de pollution inférieure à 80 microgrammes/m<sup>3</sup> des autres jours. Comparer les coefficients des deux régressions. Utiliser le summary.
4. Régression linéaire sur toutes les variables. Faire les modèles, étudier et commenter les résultats. Faire les représentations visuelles possibles.
5. Extensions éventuelles
  - a. Vers la régression quantile. Utiliser le package rq.
  - b. Vers la régression robuste. Utiliser robustReg.
  - c. Vers l'analyse de la variance.
6. Commenter l'intérêt de l'ensemble de l'étude.