

## DONNEES DU POLY "STATISTIQUES POUR DONNEES DE LA POLLUTION ATMOSPHERIQUE"

Les données ont été mises à disposition par l'association *AirNormand* (<http://http://www.airnormand.fr/>), sauf `ozone-Rennes.txt` qui est issu du livre "Statistiques avec R".

- `NO2_trafic.xls` : mesures de NO<sub>2</sub> au cours de différents trajets en voiture, selon le type de routes empruntées ("A" : Autoroute, "P" : Périurbain, "T" : Tunnel, "U" : Urbain, "V" : Voie rapide urbaine) et selon l'intensité du trafic (de "A" : fluide à "D" : congestionné).
- `AmmoniacRouen.txt` : mesures d'ammoniac ("pqvnh") à Rouen, de la vitesse du vent ("ws") et de la direction du vent ("ws", ou classée par secteurs "DV8" et "DV16").
- `NO2AirNormand.txt` : mesures de NO<sub>2</sub> par heure au centre-ville de Rouen (2 mesures en doublon "JUS\_NO2" et "JUSa\_NO2"), en zone urbaine de Rouen ("CHS\_NO2") et à Evreux ("EVT\_NO2").
- `NO2-France-07_09.txt` : mesures quotidiennes de NO<sub>2</sub> en différents sites en France et pour différentes années.
- `EauxDePluieBR0.xls` : Mesures de la concentration en différents ions des eaux de pluie tombées de 2001 à 2006 dans la forêt de Brotonne (Normandie).
- `PM10_AirNormand.txt` : mesures journalières de particules (PM10) en différents sites de normandie (voir leur situation avec le fichier `cartePM10.emf`).
- `PM10_AirNormand_corrige.txt` : idem que ci-dessus mais les mesures ont été corrigées en utilisant comme station de référence "PQVP2r" (centre-ville de Rouen) pour tous les sites de Rouen, et comme station de référence "HRIP2r" (centre-ville du Havre) pour tous les autres sites.
- `ozone.txt` : mesures journalières d'ozone, de température, de radiation solaire et de vitesse du vent (dans une direction privilégiée) durant un été.
- `ozone-Rennes.txt` : ozone maximum à Rennes pour chaque jour de l'été 2001 ("maxO3"), ainsi que de la veille ("maxO3v"), et différentes mesures météo (température, nébulosité et vitesse du vent à 9h, 12h et 15h ; direction du vent ; présence de pluie ou non).
- `NOx_2001_2009_Rouen_LeHavre.xls` : mesures (resp. quotidiennes, mensuelles ou annuelles selon les onglets) de NO et NO<sub>2</sub> en différents sites de Rouen et du Havre entre 2001 et 2009.

Tous les fichiers s'ouvrent avec R Commander grâce au menu [Données\Importer des données](#) :

- pour les `.txt` : choisir ensuite [depuis un fichier texte](#) sans changer les choix par défaut,
- pour les `.xls` : choisir ensuite [depuis Excel](#) et choisir l'onglet à ouvrir.

On peut ensuite visualiser le tableau en cliquant sur [Visualiser](#).

## Utilisation des données par chapitre du poly :

1. pas d'illustrations dans le poly.  
Exercice : ouvrir et visualiser différents jeux de données mis à disposition.
2. p17-21 : Illustrations à partir du fichier `NO2_trafic.xls` (importé sous le nom "tab" dans le poly)  
Exercice : refaire le même type d'analyses à partir de `AmmoniacRouen.txt`.
3. *Variable quantitative/Facteur*.  
Illustrations à partir de : p28 : `AmmoniacRouen.txt` (importé sous le nom "a" dans le poly), p29 : `NO2_trafic.xls`; p30-34 : `AmmoniacRouen.txt`.  
*Variable quantitative/ Variable quantitative*.  
Illustrations à partir de : p38 : `NO2AirNormand.txt`; p39 : `NO2-France-07_09.txt`.  
Exercice : Etudier le lien entre les variables du fichier `ozone.txt`.
4. p44-49 : Illustrations à partir du fichier `EauxDePluieBR0.xls`.  
Exercice : Faire une classification des sites (c'est à dire des variables) selon leurs mesures de PM10 dans le fichier `PM10_AirNormand.txt`, puis faire la même classification pour les mesures corrigées dans `PM10_AirNormand_corrige.txt`. Comparer les deux classifications.
5. pas d'illustrations.
6. p56-57. Illustrations basées sur des simulations (non fournies).
7. *t-test indépendant*.  
p62-63 : illustrations à partir du fichier `NO2_trafic.xls` où un nouveau facteur a été créé, qui ne contient que la modalité "D" (trafic congestionné) ou "O" (toutes les autres conditions de trafic), ce qui peut se faire en utilisant le menu [Données\Gérer les variables dans le jeu de données actif\Recoder des variables](#).  
Exercice : dans `NO2_trafic.xls` tester l'égalité de la concentration moyenne en NO2 selon que le type de route emprunté est "T" (tunnel) ou non.  
*t-test apparié*.  
p66-67 : illustrations à partir du fichier `NO2AirNormand.txt`  
Exercice : Créer une variable correspondant à la différence, en valeur absolue, entre les variables "JUS" et "JUSa", puis tester si la moyenne de cette différence est nulle, enfin tester si la moyenne de cette différence est inférieure à un seuil de tolérance donné (par exemple 1). Pour effectuer ces tests, on pourra utiliser le menu [Statistiques\Moyennes\t-test univarié](#).

8. p69-71 : illustrations à partir du fichier `NO2_trafic.xls`.  
Exercice : Pour le même jeu de données, tester si la moyenne de NO2 est la même quelles que soient les conditions de circulation.
  
9. p77-83 : illustrations à partir du fichier `ozone.txt`.  
Exercice : A partir du jeu de données `ozone-Rennes.txt`, essayer d'expliquer au mieux la concentration en ozone en fonction des autres variables à disposition, à l'aide d'un modèle de régression.
  
10. *Séries temporelles et dépendance temporelle, Tirer profit de la dépendance.*  
p86-89 : Illustrations à partir du fichier `ozone-Rennes.txt`.  
Exercice : Dans le même jeu de données, créer une variable contenant le maximum d'ozone de l'avant veille, et évaluer la pertinence d'inclure ou non cette variable dans le modèle de régression précédent.  
  
*Analyser les cycles et la tendance.*  
p91-95 : Illustrations à partir du fichier `NOx_2001_2009_Rouen_LeHavre.xls`, onglet "annuel" ou "mensuel" selon les cas.  
Exercice : Effectuer le même type d'analyses pour les autres sites du fichier précédent. On pourra gérer la présence des valeurs manquantes comme expliqué p96.