# MASTER1 ASA – Méthodes statistiques 2020-2021

1. Ouvrir le fichier disponible à l’adresse suivante :

<http://www.fabricemonna.com/wp-content/uploads/2019/11/bresilL3.txt>

Il représente l’analyse de plusieurs lichens collectés au Brésil, aux alentours de Rio de Janeiro (les concentrations sont exprimées en µg g-1). RJ : centre de Rio, PB et PH (deux sites distincts) un parc naturel à 30 km au nord de Rio, et LUM : site de référence à plus de 100 km au Nord.

1. Combien le tableau crée contient-ils d’observations et de variables (utiliser une fonction R) ?
2. Donner, à l’aide d’une fonction R, la moyenne, MAX, MIN, Q1, la médiane, et Q3 des concentrations en Al, Ba, Ca, Cd, Cu, Fe, La, Mg, Mn, Na, Pb, S, Sr, Ti, Zn, Zr (et seulement ces variables).
3. Quel est l’écart type de Al, de Ti?
4. Représenter sous forme d’histogramme la distribution en Al et Ti (On prendra soin de nommer correctement les axes et de donner un titre aux deux graphes). Quelle est l’allure des deux distributions ?
5. Une transformée log peut être appliquée sur des données asymétriques pour récupérer un peu de normalité. Opérer cette transformation sur Al et Ti. Représenter à nouveau les distributions. Qu’observez-vous ?
6. Il existe en statistiques une fonction qui mesure l’asymétrie d’une distribution. Quelle est-elle ? Appliquer cette fonction aux données Al et Ti puis aux mêmes données transformées. Qu’observez-vous ?
7. On s’intéresse maintenant plus particulièrement à RJ. Trouver l’intervalle de confiance de la teneur en Pb sur ce site en considérant que les données sont normalement distribuées, d’abord à 95%, puis à 99% de confiance.
8. Si la population sous-jacente n’est pas normalement distribuée, et que le nombre d’individus est petit, on peut toujours estimer l’intervalle de confiance par bootstrap. Faire cette estimation pour 95% de confiance.
9. Faire un ‘stripchat’ représentant la teneur en Pb en fonction des 4 sites étudiés (l’ajout d’un argument jitter sera apprécié). Est-il plus intéressant d’utiliser la teneur en Pb pour ce type de diagramme ou sa transformée log ? Pourquoi ?
10. Envoyer le script R à l’adresse suivante : Fabrice.Monna@u-bourgogne.fr. Le nom du fichier doit être NOM\_PRENOM.R, transmis dans la limite du temps imparti. Le sujet du mail doit être M1ASA\_CC\_2020.