**Statistique Multidimensionnelle 2018-2019**

**Session de juin**

**BAB3 Informatique**

Pour tous les groupes d’étudiants, le travail à soumettre sur la plateforme Moodle pour la veille de l’examen à minuit consiste dans l’analyse et le commentaire des fichiers de données. L’analyse est à faire par des méthodes de statistique descriptive élémentaire, d’analyses en composantes principales (ACP), d’analyse de correspondances simples (AC) ou multiples (ACM) et/ou de classification. Les étudiants doivent justifier leur(s) choix et la pertinence de la démarche suivie.

Les étudiants doivent soumettre :

* (Q1 et Q2) Le rapport d’analyse (sous forme de texte et/ou de présentation) au format PDF
* (Q1 et Q2) Les scripts correspondants à l’analyse effectuée
* (Q2) Une présentation de 10 slides (approx.) décrivant une technique d’analyse non vue au cours et comportant un exemple de celle-ci. L’exemple peut être un des exemples décrit dans le paquet R utilisé.

**QUESTION 1 :**

Fichier : **pourcentage\_chercheurs\_belgique.xls**

Le fichier contient des données relatives à l’évolution de 2006 à 2015 du pourcentage de chercheurs employés dans les différentes Provinces belge et la Région de Bruxelles Capitale.

* + Effectuez une analyse en composante principale des données.
  + Le profil est-il stable dans le temps ?
  + Dégagez des groupements de provinces/région possédant des comportements similaires au moyen d’une classification hiérarchique ascendante.

**QUESTION 2 :**

Technique non vue au cours : **analyse linéaire discriminante (LDA)**