

Analyse exploratoire des données multidimensionnelles

PUBLIC

Cet atelier s'adresse à toute personne souhaitant analyser et résumer des tableaux de données multivariées dans une optique exploratoire. Dans cet atelier, deux méthodes seront présentées : l'Analyse en Composantes Principales (ACP) et la Classification. Comme prérequis, il est supposé que les participants connaissent les notions générales de la statistique descriptive (moyenne, variance, coefficient de corrélation...).

Objectifs et Programme proposé :

L'Analyse en Composantes Principales (ACP) permet de synthétiser un tableau de données au moyen de représentations graphiques et d'outils d'aide à l'interprétation. Les méthodes de classification, permettent, quant à elles, de dégager des groupes d'observations homogènes. Ces méthodes répondent à des problèmes aussi divers que la segmentation de consommateurs en marketing, la recherche d'une taxonomie pour des données biologiques ou encore l'analyse des liaisons entre mesures physiques.

PARTIE I : Analyse en Composantes Principales.

1. Introduction à l'analyse des données multidimensionnelles
2. Fondement de l'analyse en composantes principales.
3. Mise en œuvre et outils d'aide à l'interprétation
4. Étude de cas et application logicielle.

PARTE II : Classification automatique

1. Qu'est-ce que la classification ?
2. Les deux grandes familles : classification hiérarchique et méthodes de partitionnement
3. Caractérisation des classes et aide à l'interprétation
4. Étude de cas et application logicielle.

La mise en œuvre, l'interprétation des résultats et la complémentarité ACP / classification seront illustrées sur la base d'exemples, à l'aide entre autres des logiciels XLStat et R.

Compétences à acquérir :

- Être en mesure d'identifier quelle(s) méthode(s) utiliser pour analyser un tableau de données issus d'une expérimentation, d'un plan d'expériences ou d'une enquête.
- Être capable de mettre en œuvre les techniques d'analyse en composantes principales et de classification et savoir interpréter les résultats.
- Pratiquer de façon autonome les traitements de données et l'interprétation des résultats.

Université ou École organisatrice : Oniris, Nantes – Unité StatSC

Lieu de la formation : Oniris, site de la Géraudière, Nantes

Durée de la formation: 2 jours, **les 4 et 5 juin 2018**

Volume horaire: 10 heures de cours / 6 heures de TD

Capacité d'accueil: de 8 à 16 étudiants

Responsable pédagogique : Véronique CARIOU et Pr. El Mostafa QANNARI

Contact : veronique.cariou@oniris-nantes.fr