**Data science et détection des fraudes (35h)**

**Objectif**

A la fin de ce cours, l’étudiant découvre que grâce aux données, les outils de data science permettent de construire des indicateurs pour industrialiser les processus de détection, de correction et de sanction. Cette formation permet également de comprendre la gestion des anomalies des contrats, des comportements des clients et des intermédiaires d’assurance.

Ainsi l’étudiant :

* Aura une connaissance globale des méthodes de détection de fraudes,
* Aura les connaissances nécessaires pour construire un projet anti-fraudes.
* Pourra identifier des situations grâce à l’exploitation des données externes.

**Contenu de la formation**

****Chapitre I : Conception des processus****

* Cadre légal des fraudes en assurance
* Type de fraudes, types d’anomalies, et types de risques
* Processus anti-fraudes
* Actions prescriptives et types d’automisation

****Chapitre II : Sources de données et nettoyage****

* Données internes (problématique de qualité de données)
* Dark data (problématique de transformation des données)
* Utilisation des données externes (problématique de jointure des données)
* Construction d’indicateurs
* Transformation des données (données tabulaires, textes, images)

****Chapitre III : Méthodes de détection de fraudes****

* Construction de preuves de fraudes
* Règles déterministes
* Détection d’anomalies
* Méthodes non supervisées
* Classification des fraudes

****Chapitre IV : Cas pratiques****

* Détection des anomalies avec bases de référence
* Classification binaire
* Détection des aberrations
* Segmentation avec Kmeans

# Chapitre V-  Informations pratiques

* Etude de cas des différentes anomalies
* Revue des méthodologies statistiques
* Cas pratiques concrets
  + Manipulation des bases de données
  + Construction des algorithmes avec R