

# [SCOCHURN] Scoring pour la rétention client

Ce stage est destiné aux chargés d'étude marketing qui ont besoin de bâtir des scores d'attrition / churn. Pour cela, ils utiliseront des techniques de modélisation classiques (régression logistique, arbres de décision) pour estimer la probabilité de départ d'un client. Une ouverture est faite sur les modèles de durée qui sont complémentaires.

Cette formation traite aussi de la phase exploratoire et de l'utilisation du modèle (courbe de lift, courbe ROC, suivi dans le temps).

**Durée : 2 ou 3 jours selon les attentes et le niveau du public**

**Pré-requis :  
STAT101SAS**

**Logiciels possibles : SAS, SAS Enterprise Guide (code), Spad**

## 1. Définition de la problématique et préparation de la base

- Cerner la variable cible : qu'est-ce qu'un départ ? à quel horizon ?
- Pourquoi équilibrer un échantillon ?
- Scission en deux ou trois corpus de données

## 2. Statistique exploratoire

- Graphiques : courbes de densité, weight of evidence
- Indicateurs de liaison : Cramer, Fisher, Wilcoxon, Kolmogorov-Smirnov
- Caractérisation par les variables qualitatives

## 3. Mise au pas des données

- Ecrêtement de valeurs extrêmes
- Gestion de données manquantes
- Mise en tranches des variables quantitatives

## 4. Modélisations

- Régression logistique : sélection du meilleur modèle
- Régression logistique : simplification et interprétation du modèle
- Arbre de décision
- Modélisation d'une durée : modèles paramétriques AFT

## 5. Comparaison et utilisation de modèles

- Stabilité du modèle : AUC sur les trois corpus de données
- Performance du modèle : courbe de lift
- Prise de décision : seuil optimal, tableaux de classement