

Objectifs pédagogiques

Ce stage permet de choisir la meilleure approche de programmation lorsque vous devez faire face à des gros volumes de données et que se posent des problèmes de performance, de mémoire ou d'espace.

A l'issue de ce stage vous serez capables d'arbitrer en faveur de la meilleure méthode de programmation

Pré-requis: Avoir une bonne pratique du logiciel SAS

Pédagogie: 40% apports théoriques, 60% exercices d'application

Durée : 2 jours

Contenu

Efficacité : de quoi parle t on ?

Efficacité, lisibilité à quel coût ? i

BENCHMARKS : Comparer les techniques de programmation pour se faire une idée

SAS aime les tables "grandes et minces"

Réduire l'espace occupé par des variables

Compresser les tables ? une bonne idée ?

PROC DATASETS : l'outil du programmeur efficace

Afficher les caractéristiques d'une table sas

Créer des indexes simples ou multiples

LES FORMATS : un patrimoine précieux dans la culture SAS

Table de dimensions (codes) : penser format

OTHER= pour repérer les mal codés

RAPPELS (inutiles ?) sur les best practices dans l'étape DATA

Keep= en abuser

WHERE mieux que IF quand on a le choix

SORT : la grande question des tris

Eviter les tris inutiles : INDEX/ NOTSORTED /GROUPFORMAT/CLASS/ SORTED BY

Les options SAS qui impactent le tri

Mémoire : comment contrôler la mémoire utilisée

PAGESIZE et Nombre de Buffers

L'instruction SASFILE

Techniques de jointure : laquelle choisir

HASHAGE /Utiliser des ARRAYS /SQL / MERGE/Formats ?

Impact des indexes

Techniques de filtre: laquelle choisir

Requête SQL imbriquée

Liste de codes dans une macro variable (SQL INTO)

Filtrer au moment d'un tri (WHERE)

Techniques pour agréger les données

SUMMARY ou MEANS : la fin d'un mythe

COMPTEURS dans une étape DATA (avec FIRST. LAST.)

GROUP BY de SQL