



caf.fr

**POUVOIR ADJUDICATEUR  
(organisme contractant)**

*ORGANISME DE DROIT PRIVE CHARGE DE LA GESTION D'UN SERVICE PUBLIC*  
CAISSE D'ALLOCATIONS FAMILIALES DES HAUTS-DE-SEINE  
70-88, rue Paul Lescop  
92023 NANTERRE CEDEX

## **CAHIER DES CLAUSES TECHNIQUES PARTICULIERES (C.C.T.P.)**

**OBJET DE LA PROCEDURE ADAPTEE N° 03/2026**

**PRESTATIONS DE FORMATIONS DANS LE DOMAINE DES  
OUTILS D'ANALYSE POUR LE DISPOSITIF « PERSPICAF »  
DESTINÉES À L'INSTITUT RÉGIONAL DE FORMATION  
DES ALLOCATIONS FAMILIALES (IRFAF) RELEVANT DE  
LA CAF DES HAUTS-DE-SEINE  
(CCTP TOUS LOTS)**

ACCORD-CADRE ALLOTI PASSÉ PAR PROCEDURE ADAPTEE EN APPLICATION DES DISPOSITIONS DES ARTICLES R 2123-1-3° (MARCHE DE SERVICES SPECIFIQUES), R.2131-15, R.2161-2 A R.2161-5, ET R2162-13 et 14 DU CODE DE LA COMMANDE PUBLIQUE

Date d'établissement : 20/01/2026

## SOMMAIRE

PREAMBULE - CONTEXTE OBJET .....	3
TITRE I : CLAUSES COMMUNES À TOUS LES LOTS.....	4
ARTICLE 1 - PRESTATIONS ATTENDUES.....	4
1.1 : CONCEPTION .....	4
1.2 : MAINTENANCE.....	4
1.3 : ANIMATION DES SESSIONS DE FORMATIONS .....	5
ARTICLE 2 - OBLIGATIONS GENERALES.....	5
ARTICLE 3 - INTERVENANTS - ORGANISATION MATERIELLE.....	6
ARTICLE 4 – DOCUMENTS DES FORMATIONS (PROGRAMMES -SUPPORTS) .....	6
ARTICLE 5 – MODALITES D’EVALUATION.....	7
TITRE II : CLAUSES TECHNIQUES RELATIVES À CHAQUE LOT .....	8
ARTICLE 1 – LOT N°1 – FORMATIONS POWER BI.....	8
1.1 LES PRODUITS DE FORMATION .....	8
1.2 VOLUMETRIE PREVISIONNELLE .....	9
ARTICLE 2 – LOT N°2 – FORMATIONS SQL ET PYTHON DANS DATABRICKS .....	10
2.1 : LES PRODUITS DE FORMATION .....	10
2.2 VOLUMETRIE PREVISIONNELLE .....	12
ARTICLE 3 : LOT N° 3 – FORMATIONS ANALYSE DE DONNEES AVEC PYTHON.....	12
3.1 : LES PRODUITS DE FORMATION .....	12
3.2 VOLUMETRIE PREVISIONNELLE .....	15
ARTICLE 4 : LOT N° 4 – QFORMATIONS CARTOGRAPHIE AVEC MAPINFO.....	15
4.1 : LES PRODUITS DE FORMATION .....	15
4.2 VOLUMETRIE PREVISIONNELLE .....	16

## **PREAMBULE - CONTEXTE - OBJET**

Le présent accord-cadre a pour objet la conception, l'animation et la maintenance d'actions de formations dans le domaine des outils d'analyse de données pour le dispositif national de formation du Projet d'Etudes en Réseau Statistique du Potentiel d'Information des CAF (« PERSPICAF »).

Les 101 Caisses d'allocations familiales sont réparties sur chaque département du territoire pour assurer, sous la tutelle de la Caisse nationale des allocations familiales, les missions de la branche Famille de la Sécurité Sociale. Ce sont des organismes de droit privé dotés de Conseils d'administration et qui sont chargés de la gestion d'un service public.

L'Institut régional de formation des allocations familiales (IRFAF) est un service commun, géré par la CAF des Hauts-de-Seine, qui déploie, met en place des formations pour les huit Caf d'Ile-de-France, mais également pour les Caf du réseau (province, DROM/COM, Corse) ou pour la Caisse Nationale des Allocations Familiales (CNAF).

Les formations objet de l'accord-cadre sont principalement destinées aux chargés d'études des CAF. Leur métier consiste à manipuler des données brutes dans le but de réaliser des analyses sociologiques, statistiques et démographiques afin notamment d'apporter les éléments pour aider à la décision. Ces analyses requièrent des compétences à la fois théoriques, méthodologiques et pratiques en programmation, en statistiques et en cartographie.

Le détail des prestations attendues figure ci-après.

# TITRE I : CLAUSES COMMUNES À TOUS LES LOTS

## ARTICLE 1 - PRESTATIONS ATTENDUES

### 1.1 : CONCEPTION

On entend par la prestation « Conception » les éléments suivants, en tenant compte des spécificités liées au contexte :

- Analyser le besoin et effectuer si nécessaire un diagnostic (réunion de cadrage (présentiel ou distanciel), entretien, recensement des besoins, réalisation d'enquête, questionnaire...),
- Elaborer et faire valider au pilote de la formation, une proposition pédagogique (parcours de formation en présentiel ou classe virtuelle, ou Blended Learning),
- Concevoir et faire valider au pilote de la formation, 15 jours ouvrés avant le début de l'action, un kit pédagogique qui comprend :
  - Un programme de formation détaillé (prérequis, titre, public, objectifs globaux et opérationnels, éléments de contenu et méthodes pédagogiques et d'évaluation)
  - Des supports formateurs comportant un déroulé pédagogique, un diaporama, des exercices et leurs corrigés, des études de cas, bibliographies, des vidéos...
  - Des supports stagiaires (format papier et/ou dématérialisé) comportant un livret stagiaire ou des fiches de synthèse, des exercices et des corrigés
  - Pour la mise en place de formations à distance, prévoir et effectuer des tests techniques nécessaires au bon déroulement de la formation.

Dans tous les cas, le titulaire devra respecter les instructions et procédures communiquées par l'Irfaf en termes de formalisation et de charte graphique et ajuster le cas échéant, le kit pédagogique selon les modifications demandées par l'Irfaf.

Le kit pédagogique reste la propriété du titulaire du marché, qui en concède, à titre non exclusif, l'usage au pouvoir adjudicateur pour les besoins découlant de l'objet du marché, conformément à l'article 11 du CCAP.

### 1.2 : MAINTENANCE

Le titulaire s'engage à la réalisation d'opérations de « maintenance » sur l'ensemble des éléments contenus dans la prestation « conception ».

La prestation de maintenance, réalisée à la demande du pouvoir adjudicateur comprend :

- L'ajustement du programme de formation et du déroulé pédagogique en fonction des évolutions de la réglementation, des techniques ou des outils, objets de la formation, ou des actions d'amélioration issues du bilan de la formation (prérequis, public, éléments de contenu et méthodes pédagogiques...),
- L'intégration de nouveaux exercices ou de nouveaux supports stagiaires ou formateurs,
- Toute autre demande de transformation qui pourra être demandée par l'Irfaf.

### 1.3 : ANIMATION DES SESSIONS DE FORMATIONS

On entend par la prestation « Animation » les éléments suivants :

- Préparer et animer les sessions de formation (en présentiel ou en classe virtuelle en utilisant les outils collaboratifs de l'Irfaf, exemple Teams...) aux dates ou échéances et lieux définis et selon le programme validé par l'Irfaf,
- Participer aux réunions préparatoires pour analyser les besoins de formation ou aux réunions pour la réalisation du bilan des sessions de formations,
- Faire compléter et transmettre à la fin de chaque session les documents administratifs dont notamment les feuilles d'émargement signées par l'intervenant et les stagiaires,
- Compléter et communiquer le bilan formateur de l'Irfaf au plus tard dans les 8 jours suivant la fin de l'action,
- S'assurer du transfert du kit pédagogique à l'ensemble des formateurs intervenant dans le cadre du marché,
- Restituer les travaux effectués en formation (compte-rendu écrit et/ou oral), et/ou les plans d'actions etc., à la demande.
- Transmettre à l'Irfaf, dans les 8 jours suivants la fin de la session, l'ensemble des documents administratifs liés à la session de formation. Le prestataire s'assure de la présence des stagiaires en leur faisant compléter les émargements au format papier ou en ligne en fonction des orientations de la structure, matin et après-midi. Les formateurs sont garants de la complétude par les stagiaires par demi-journées et doivent tracer les absences et retards.

Le Titulaire rend compte à l'Irfaf des éventuels problèmes rencontrés en toute circonstance. Il doit, en particulier **informer immédiatement** l'Irfaf des désistements, absences des stagiaires ou les départs

## ARTICLE 2 - OBLIGATIONS GENERALES

De manière générale, le prestataire s'engage à :

- Concevoir et mettre à disposition de l'Irfaf des kits pédagogiques : supports stagiaires, guide animateur ou séquençement pédagogique, exercices stagiaires, évaluation stagiaire en amont de la formation et en aval, etc.
- Effectuer la maintenance des kits,
- Assurer l'animation des modules de l'offre sur le territoire national à la demande de l'Irfaf,
- Respecter la planification des sessions de formation,
- Transmettre, en version dématérialisée et/ou en version papier le bilan formateur complété dans son intégralité, au plus tard dans les 8 jours suivant la fin de l'action,
- Transmettre la feuille d'émargement à l'Irfaf signée par l'intervenant et les stagiaires, dans les 8 jours qui suivent la fin du stage. L'envoi se fera par scan suivi d'un envoi des originaux à l'Irfaf.
- Respecter les procédures et les règles de collaboration indiquées par l'Irfaf lors de la réunion de cadrage qui suivra la notification du marché et inscrites dans le guide de collaboration.

- Respecter les conditions mises en place par l'IRFAF pour répondre aux exigences du Label QUALIOPI. Des contrôles du respect de ces conditions pourront être effectués par l'IRFAF selon des modalités qui seront communiquées au prestataire.

## **ARTICLE 3 - INTERVENANTS - ORGANISATION MATERIELLE**

### **Intervenants**

Cf. articles 14.3 et 14.4 du CCAP.

En cas d'intervention dans les locaux d'une Caf cliente, l'intervenant devra contacter et faire valider par le manager de projet toute demande supplémentaire qui lui aurait été faite par la Caf, toutes difficultés rencontrées, ou toute demande de modification de programme.

### **Durée de la formation**

Une journée de formation équivaut à 7 heures, soit de 09h00 à 17h00 (une heure réservée à la pause déjeuner). Certaines actions seront (modules à distance de la formation initiale) et pourront (potentiels modules de perfectionnement, séances d'analyse de pratiques...) être animées en demi-journées, soit 3h30.

Ces horaires sont susceptibles de modifications à la demande de l'Irfaf.

### **Nombre de participants**

Voir spécificités propres à chaque lot.

### **Lieux de formation**

Les actions de formation se dérouleront se dérouleront soit à l'Irfaf au 67/69, avenue Jean Jaurès - 75927 PARIS CEDEX 19, soit directement en CAF (région Ile-de-France dans les Caisses d'Allocations Familiales ainsi que sur les autres départements et territoires (Province et DROM-COM, Corse) selon la demande). Le lieu sera confirmé lors de l'émission du bon de commande.

## **ARTICLE 4 – DOCUMENTS DES FORMATIONS (PROGRAMMES -SUPPORTS)**

Concernant l'élaboration du programme, le manager de projets de l'IRFAF transmet au prestataire la fiche catalogue ou les informations et documents nécessaires à une bonne compréhension de l'action de formation à mettre en place.

À la réception du déroulé pédagogique de l'action établi par le prestataire, un échange sur son contenu entre le manager de projets et le prestataire pourra s'effectuer par mail, par téléphone ou réunion si nécessaire.

Le prestataire aura la charge d'animer un groupe de travail composé d'experts métier du réseau et du comité pédagogique du groupe formation afin de recueillir les éléments nécessaires à la conception des exercices et réalisation des supports de formation. Le prestataire ajuste si nécessaire le déroulé pédagogique.

L'intervenant devra fournir le programme finalisé et le dossier documentaire en version dématérialisée au manager de projets, pilote de l'action, 15 jours ouvrés avant le début de l'action, selon la complexité du dossier. Ce délai pourra être révisé du fait de l'Irfaf et en cas de situations particulières.

Les documents produits par le prestataire seront validés par le manager de projets. Ils devront être datés et actualisés.

Le kit pédagogique est constitué du déroulé pédagogique et des supports documentaires remis aux formateurs et aux stagiaires pendant la formation et à l'issue de la formation.

Les coordonnées du prestataire ou du formateur (adresse postale, lien internet, numéro de téléphone) ne doivent pas figurer sur les documents remis aux stagiaires, ni leurs être communiquées.

L'IRFAF archive systématiquement, sous format dématérialisé, un exemplaire de chaque document remis ou présenté aux stagiaires.

Les supports pédagogiques dont l'IRFAF est propriétaire ne peuvent être reproduits ou utilisés à l'extérieur de l'Institut qu'avec l'autorisation expresse du Directeur ou de son représentant.

## ARTICLE 5 – MODALITES D'EVALUATION

L'atteinte de chaque objectif pédagogique devra être mesurée au cours de chaque session conformément aux modalités définies dans le déroulé pédagogique validé par l'Irfaf.

Les prérequis ou les écarts de compétences devront systématiquement être vérifiés par le formateur au début de la formation. Les modalités de cette évaluation devront être précisées dans le déroulé pédagogique (quiz, QCM, exercices, mises en situation, tests, tour de table etc.).

### Modalités d'évaluation :

- **Evaluations de positionnement et des acquis** : communiquées dans la proposition pédagogique
- **Evaluation à chaud** : à la fin de la formation, un bilan oral est effectué par le formateur et une évaluation écrite adressée aux stagiaires permettent d'apprécier la qualité de la prestation et de mesurer l'efficacité de l'action au regard des objectifs globaux
- **Evaluation à froid** : réalisée avec un outil interne Caf (ou par l'Irfaf sur demande)
- **Attestation de suivi** : Feuille de présence
- **Certificats de réalisation** mentionnant la nature et la durée de l'action est remis aux stagiaires à l'issue de la formation.

Si l'un des moyens d'évaluation amène à constater un écart par rapport aux objectifs annoncés, le prestataire s'engage à :

- alerter pendant ou à l'issue de la formation le manager de projet pilote de l'action,
- réadapter le contenu de la formation,
- réitérer à ses seuls frais le stage concerné,
- remplacer immédiatement le formateur si sa responsabilité ou sa compétence est en cause.

## TITRE II : CLAUSES TECHNIQUES RELATIVES À CHAQUE LOT

### ARTICLE 1 – LOT N°1 – FORMATIONS POWER BI

#### 1.1 LES PRODUITS DE FORMATION

##### 1.1.1 : PUBLIC CONCERNE

**POWER BI – Module initiation** : Tout agent dont l'activité nécessite de construire des tableaux de bord : suivi d'activité analyse de données...

**POWER BI – Module perfectionnement** : Tout agent dont l'activité nécessite de construire des tableaux de bord avancés

##### Nombre de participants par session

- POWER BI – Module initiation : Entre 7 et 10 participants.
- POWER BI – Module perfectionnement : Entre 6 et 8 participants.

##### 1.1.2 : NATURE DES FORMATIONS

Formation sur un outil de visualisation de données.

##### 1.1.3 : THEMES ET CONTENUS DES FORMATIONS

Les formations ont pour objectifs de permettre aux publics concernés d'acquérir et/ou de développer des savoir-faire et savoir-faire technique pour utiliser Power BI dans le cadre de leur activité.

Les formations seront réalisées à partir des données du Système d'Information Décisionnel (SID) de la branche Famille.

La formation POWER BI perfectionnement sera réalisée en coanimation avec un expert métier de la branche Famille.

##### **POWER BI - Module Initiation**

##### Les prérequis :

- Connaître les données du système d'information de la branche Famille
- Usage courant d'un tableur
- Notions souhaitables sur les bases de données
- Avoir accès à Azure Databricks

##### Les objectifs généraux :

- Interroger et associer des sources de données hétérogènes
- Préparer et nettoyer les données
- Modéliser et optimiser le jeu de données
- Concevoir des mesures simples avec le langage DAX
- Réaliser et partager des tableaux de bord

##### Les objectifs pédagogiques :



- Identifier les types de sources, expliquer leurs différences, appliquer des requêtes, concevoir un processus d'association et justifier les choix d'intégration.
- Énumérer les techniques de nettoyage, décrire leur impact, appliquer des transformations, diagnostiquer les problèmes, élaborer un plan et évaluer la pertinence des règles.
- Définir les concepts de modélisation, expliquer les principes, construire un modèle, identifier les goulots d'étranglement, concevoir un modèle optimisé et comparer plusieurs modèles.
- Citer les fonctions DAX, expliquer les différences entre colonnes et mesures, écrire des mesures simples, décomposer une mesure complexe, créer des mesures personnalisées et vérifier leur cohérence.
- Identifier les composants, décrire les bonnes pratiques, construire un tableau interactif, évaluer la pertinence des visualisations, concevoir un tableau complet et critiquer/améliorer un tableau existant.

## **POWER BI – Module perfectionnement**

### **Les prérequis :**

- Avoir accès à Azure Databricks et aux données du SID
- Connaître les données du système d'information de la branche Famille
- Usage courant d'un tableur
- Notions sur les bases de données
- Avoir suivi la formation « POWER BI - Module initiation : Concevoir des tableaux de bord avec les données du SID » ou disposer d'un niveau équivalent sur l'utilisation de POWER BI

### **Les objectifs généraux :**

- Transformer, créer des requêtes avancées et modéliser les données à l'aide du langage M
- Mettre en œuvre des fonctions DAX avancées pour réaliser des traitements plus poussés sur les données à analyser
- Maîtriser les jointures entre les tables
- Optimiser la performance du modèle

### **Les objectifs pédagogiques :**

- Décrire les concepts du langage M, expliquer les transformations avancées, appliquer des requêtes complexes, analyser la logique des scripts, concevoir des requêtes optimisées et évaluer leur performance.
- Identifier les fonctions DAX avancées (CALCULATE, FILTER, etc.), expliquer leur rôle, appliquer des calculs complexes, analyser des formules imbriquées, concevoir des mesures personnalisées et valider leur cohérence et performance.
- Définir les types de relations (1:N, N:N), expliquer les impacts des cardinalités, appliquer des jointures dans Power Query et le modèle, analyser les problèmes liés aux relations, concevoir un schéma relationnel optimal et évaluer la pertinence des choix.
- Identifier les facteurs influençant la performance, expliquer les bonnes pratiques, appliquer des optimisations (agrégations, compression), analyser les goulots d'étranglement, concevoir un modèle performant et justifier les choix techniques.

## **1.2 VOLUMETRIE PREVISIONNELLE**

Power Bi – Module initiation : 20 à 30 stagiaires par an.

Power Bi – Module perfectionnement : 30 à 35 stagiaires par an.

Ces chiffres sont donnés à titre indicatif et n'engagent pas la Caf des Hauts-de-Seine sur un volume de prestations.

## **ARTICLE 2 – LOT N°2 – FORMATIONS SQL ET PYTHON DANS DATABRICKS**

### **2.1 : LES PRODUITS DE FORMATION**

#### **2.1.1 : PUBLIC CONCERNE**

Tout agent amené à manipuler les données du SID nécessitant l'utilisation d'un langage de programmation.

Nombre de participants par session : entre 8 et 10 participants

#### **2.1.2 : NATURE DES FORMATIONS**

Formation sur des langages de requêtage, manipulation et analyse de données en vue de construire des analyses statistiques.

#### **2.1.3 : THEMES ET CONTENUS DES FORMATIONS**

Les formations ont pour objectifs de permettre aux publics concernés d'acquérir et/ou de développer des savoir-faire et savoir-faire technique pour utiliser Data bricks et les langages de programmation associés dans le cadre de leur activité.

Les formations seront réalisées à partir des données du SID et dans l'environnement Databricks de la branche famille.

#### **SQL dans Databricks**

##### **Les prérequis :**

- Connaître les données du système d'information de la branche Famille
- Notions souhaitables sur les bases de données
- Avoir accès à Azure Databricks

##### **Les objectifs généraux :**

- Consulter et manipuler les données dans Databricks
- Réaliser un décompte
- Appréhender l'algèbre relationnelle et le langage SQL (SELECT, FROM, WHERE, GROUP BY, ORDER BY, jointure, requêtes imbriquées, UPDATE)

##### **Les objectifs pédagogiques :**

- Identifier les commandes SQL disponibles, expliquer leur usage, exécuter des requêtes simples, analyser les résultats, concevoir des requêtes adaptées et évaluer leur pertinence.
- Reconnaître la structure d'une requête, expliquer les clauses, modifier une requête pour répondre à un besoin, analyser l'impact des changements, concevoir des requêtes optimisées et valider leur exactitude.

- Lister les fonctions de comptage, expliquer leur rôle, appliquer COUNT et variantes, analyser les résultats, concevoir des requêtes de synthèse et évaluer la cohérence des décomptes.
- Définir les concepts clés (SELECT, GROUP BY, jointure, requêtes imbriquées, UPDATE), expliquer leur logique, appliquer ces commandes, analyser des requêtes complexes, concevoir des solutions SQL avancées et évaluer leur performance.

### **PYTHON dans Databricks - Niveau 1**

#### **Les prérequis :**

- Connaître les données du système d'information de la branche Famille
- Notions souhaitables sur les bases de données
- Avoir accès à Azure Databricks

#### **Les objectifs généraux :**

- Utiliser les fonctionnalités de l'interface Databricks
- Employer les fonctions utiles pour manipuler les données du SID
- Appliquer les bonnes pratiques de la Branche famille sur la programmation d'un notebook.

#### **Les objectifs pédagogiques :**

- Identifier les éléments de l'interface, expliquer leur rôle, appliquer les fonctionnalités pour exécuter des scripts, analyser les interactions et évaluer la pertinence des actions.
- Lister les fonctions disponibles, expliquer leur usage, appliquer des fonctions pour transformer les données, analyser les résultats et concevoir des scripts optimisés.
- Décrire les bonnes pratiques, expliquer leur importance, appliquer les standards dans un notebook, analyser la conformité et évaluer la qualité du code.

### **PYTHON dans Databricks - Niveau 2**

#### **Les prérequis :**

- Connaître les données du système d'information de la branche Famille
- Notions souhaitables sur les bases de données
- Avoir accès à Azure Databricks
- Avoir suivi la formation PYTHON niveau 1 ou disposer des connaissances équivalentes

#### **Les objectifs généraux :**

- Programmer et manipuler des paramètres (Widgets)
- Créer et transformer des colonnes dans une table et manipuler les données
- Mettre en forme des résultats statistiques
- Manipuler les dates
- Maîtriser les boucles et les itérations en Python

#### **Les objectifs pédagogiques :**

- Définir les widgets, expliquer leur usage, appliquer des paramètres dans un script, analyser leur impact et concevoir des notebooks interactifs.

- Lister les méthodes de création et transformation, expliquer leur logique, appliquer des opérations sur les colonnes, analyser les résultats et optimiser les scripts.
- Identifier les fonctions statistiques, expliquer leur rôle, appliquer des mises en forme, analyser la pertinence des résultats et concevoir des présentations optimisées.
- Lister les fonctions de gestion des dates, expliquer leur usage, appliquer des transformations, analyser les impacts et concevoir des scripts robustes.
- Définir les structures de boucle, expliquer leur fonctionnement, appliquer des itérations pour automatiser des tâches, analyser la performance et optimiser le code.

## **2.2 VOLUMETRIE PREVISIONNELLE**

SQL : 10 à 15 stagiaires par an.

PYTHON Niveau 1 : 20 à 25 stagiaires par an.

PYTHON Niveau 2 : 15 à 20 stagiaires par an.

Ces chiffres sont donnés à titre indicatif et n'engagent pas la Caf des Hauts-de-Seine sur un volume de prestations.

## **ARTICLE 3 : LOT N° 3 – FORMATIONS ANALYSE DE DONNEES AVEC PYTHON**

### **3.1 : LES PRODUITS DE FORMATION**

#### **3.1.1 : PUBLIC CONCERNE**

Chargés d'études, Data Analyste, Data Scientiste, Contrôleurs de gestion ayant des connaissances initiales en statistiques.

Nombre de participants par session : entre 5 et 8 participants.

#### **3.1.2 : NATURE DES FORMATIONS**

##### **Analyses statistiques multidimensionnelles**

Le métier de chargés d'études consiste à manipuler des données brutes dans le but de réaliser des analyses sociologiques, statistiques et démographiques afin notamment d'apporter les éléments pour aider à la décision. Ces analyses requièrent des compétences à la fois théoriques, méthodologiques et pratiques en programmation.

En outre, les fonctionnalités de Databricks offrent de nouvelles perspectives d'analyse de données plus avancées (régression, modèles de prévision, machine Learning, analyse géospatiale, ...). Les chargés d'études doivent donc être formés sur les méthodes statistiques théoriques et pratiques afin de réaliser des Analyse des Correspondances Multiples (ACM) et Analyse Factorielle des Correspondances (AFC) sous Databricks (utilisation du langage Python et des bibliothèques Pandas et matplotlib).

##### **Les prérequis :**

- Maîtriser le langage python dans Databricks,
- Maîtriser les méthodes d'échantillonnage aléatoire simple et stratifié,
- Maîtriser les concepts de statistique descriptive (variable quali/quant, moyenne, quantiles, variance, écart-type...),

- Connaître les grands principes et méthodes de statistique décisionnelle (tests, estimations, régression) et d'analyse des données multidimensionnelles
- Utiliser les données du Sid de la branche Famille.

#### Les objectifs professionnels :

- Analyser les caractéristiques socio-économiques des allocataires à l'aide d'une Analyse des Correspondances Multiples (ACM), afin d'identifier des profils types et d'orienter les actions de communication ou d'accompagnement vers les publics les plus vulnérables.
- Mettre en œuvre des outils d'exploration et de visualisation des données (avec Pandas et Matplotlib) pour produire des tableaux de bord destinés à suivre l'évolution des prestations familiales par territoire ou par typologie de ménage.

#### Les objectifs pédagogiques :

- Lister les bibliothèques utiles, expliquer leur rôle, appliquer des opérations de transformation et filtrage, analyser les résultats et optimiser le code.
- Identifier les types de graphiques adaptés, expliquer leur usage, créer des visualisations avec Matplotlib, analyser la lisibilité et améliorer la présentation.
- Définir les concepts d'AFC et ACM, expliquer leur logique, appliquer les méthodes avec Python, analyser les résultats et interpréter les dimensions principales.
- Lister les étapes d'interprétation, expliquer les relations entre modalités, appliquer des visualisations des axes factoriels et évaluer la pertinence des conclusions.

### **Modélisation LOGIT et typologie avancées**

Le métier de chargés d'études consiste à manipuler des données brutes dans le but de réaliser des analyses sociologiques, statistiques et démographiques afin notamment d'apporter les éléments pour aider à la décision. Ces analyses requièrent des compétences à la fois théoriques, méthodologiques et pratiques en programmation.

En outre, les fonctionnalités de Databricks offrent de nouvelles perspectives d'analyse de données plus avancées (régression, modèles de prévision, machine Learning, analyse géospatiale, ...). Les chargés d'études doivent donc être formés sur les méthodes statistiques théoriques et pratiques afin de maîtriser la construction de modèles LOGIT mais aussi d'approfondir l'analyse des données à travers des techniques de réduction de dimensionnalité et de segmentation.

Ils devront être capable de préparer et encoder des variables explicatives, construire et ajuster un modèle LOGIT et, enfin, segmenter une population via une Analyse en Composante Principale (ACP) et une Classification Ascendante Hiérarchique (CAH).

#### Les prérequis :

- Maîtriser le langage python dans Databricks
- Maîtriser les méthodes d'échantillonnage aléatoire simple et stratifié
- Maîtriser les concepts de statistique descriptive (variable quali/quant, moyenne, quantiles, variance, écart-type...),
- Avoir une approche théorique de la régression linéaire simple (Moindre carrés ordinaires MCO)
- Utiliser les données du Sid de la branche Famille

#### Les objectifs professionnels :

##### 1.Préparer et encoder des variables explicatives

- Sélectionner les variables pertinentes et traiter les données (valeurs manquantes, doublons, normalisation),
- Encoder les variables qualitatives et créer des variables dérivées si nécessaire,
- Détecter et corriger la colinéarité entre variables explicatives.

## 2. Construire et ajuster un modèle LOGIT

- Estimer un modèle LOGIT et interpréter ses coefficients,
- Évaluer la performance du modèle (AUC, courbe ROC, tests de significativité),
- Comparer et ajuster les modèles à l'aide de critères tels que AIC et BIC.

## 3. Segmenter une population via ACP et CAH,

- Réaliser une ACP pour réduire la dimensionnalité et interpréter les axes principaux,
- Appliquer une CAH sur les composantes principales pour créer des groupes homogènes,
- Analyser et interpréter les segments obtenus à l'aide de visualisations adaptée.

### Les objectifs pédagogiques :

- Lister les étapes de préparation, expliquer les techniques d'encodage, appliquer le traitement des données, analyser la colinéarité et optimiser la sélection des variables.
- Définir la méthode LOGIT, expliquer les coefficients, appliquer la modélisation avec Python, analyser la performance via AUC et ROC, comparer les modèles avec AIC/BIC.
- Décrire les concepts d'ACP et CAH, expliquer leur logique, appliquer les méthodes avec Python, analyser les segments et interpréter les visualisations.
- Interpréter un odds ratio et analyser la courbe ROC pour évaluer la performance d'un modèle statistique
- Mesurer la qualité d'un modèle (AIC, BIC, AUC)
- Effectuer une Analyse en Composantes Principales (ACP) pour réduire la dimensionnalité des données et faciliter leur interprétation
- Construire des typologies grâce à l'ACP et à la classification ascendante hiérarchique (CAH) pour regrouper des observations similaires

### **3.1.3 : THEMES ET CONTENUS DES FORMATIONS**

- Rappel des concepts statistiques théoriques utiles à la compréhension des méthodes abordées.
- Définition des problématiques, choix des données, écueils à éviter.
- Détermination des méthodes d'analyse en fonction des populations et des objectifs d'étude.
- Les tableaux de variables qualitatives (tableaux de contingence, tableau de Burt, tableaux disjonctifs, etc...).
- Présentation des outils et techniques statistiques propres à l'analyse multidimensionnelle sur des variables qualitatives : (AFC et AFCM).
- Mise en perspective de la méthodologie pour développer un modèle LOGIT.
- Présentation et mise en œuvre du modèle économétrique LOGIT.
- Interprétation des résultats.
- Rappel des concepts statistiques utiles à la compréhension des méthodes abordées.
- Définition et présentation d'une typologie comme outil d'aide à la décision.
- Réflexion sur la problématique, l'intérêt et les écueils de la typologie à partir d'exemples concrets.

- Présentation et mise en œuvre des outils et techniques statistiques propres à l'analyse multidimensionnelle : Corrélation, Régression linéaire multiple, Analyse en composantes principales (ACP), Classification Ascendante Hiérarchique (CAH), segmentation.
- Choix des variables à retenir pour la réalisation d'une typologie, en fonction des attentes de l'analyse.
- Interprétation et analyse critique des résultats produits par les logiciels.
- Réalisation d'applications métier avec Python et les données du SID de la branche Famille.

### **3.2 VOLUMETRIE PREVISIONNELLE**

Analyse statistique multidimensionnelle : 20 à 25 stagiaires par an.

Modélisation LOGIT : 20 à 25 stagiaires par an.

Ces chiffres sont donnés à titre indicatif et n'engagent pas la Caf des Hauts-de-Seine sur un volume de prestations.

## **ARTICLE 4 : LOT N° 4 – FORMATIONS CARTOGRAPHIE AVEC MAPINFO**

### **4.1 : LES PRODUITS DE FORMATION**

#### **4.1.1 : PUBLIC CONCERNE**

Cartographe, chargés d'études, statisticiens, chefs de projet.

Nombre de participants par session : entre 7 et 8 participants

#### **4.1.2 : NATURE DES FORMATIONS**

Formation sur un outil de cartographie permettant d'analyser des données géoréférencées.

#### **4.1.3 : THEMES ET CONTENUS DES FORMATIONS**

##### **MAPINFO Initiation :**

Cette action propose aux professionnels des CAF d'acquérir et de développer des compétences techniques pour utiliser Mapinfo dans le cadre de leur activité.

##### **Les prérequis :**

- Connaître les données du système d'information de la branche Famille
- Avoir accès au logiciel Mapinfo
- Notions souhaitables sur les bases données

##### **Les objectifs généraux :**

- Utiliser Mapinfo pour réaliser des cartes
- Appliquer les techniques d'analyse spatiale des données
- Travailler avec des données externes (partenaires, institutions, prestataires...).

Les objectifs pédagogiques :

- Identifier les fonctionnalités de MapInfo, expliquer leur usage, appliquer la création de cartes, analyser la lisibilité et optimiser la présentation.
- Lister les méthodes d'analyse spatiale, expliquer leur logique, appliquer des traitements dans MapInfo, analyser les résultats et évaluer leur pertinence.
- Décrire les formats de données externes, expliquer les étapes d'intégration, appliquer l'importation et la mise à jour, analyser la cohérence et optimiser la gestion des données.

**MAPINFO Perfectionnement**

Cette action de formation propose aux professionnels des CAF d'acquérir une expertise en cartographie dans le cadre de leur activité.

Les prérequis :

- Utilisation courante de Mapinfo
- Connaître les données du système d'information de la branche
- Notions souhaitables sur les bases données.

Les objectifs généraux :

- Utiliser Mapinfo pour réaliser des cartes
- Appliquer les techniques avancées d'analyse spatiale des données
- Savoir créer une carte thermique
- Présentation de Mapinfo Marketplace

Les objectifs pédagogiques :

- Identifier les fonctionnalités avancées de MapInfo, expliquer leur usage, appliquer la création de cartes complexes, analyser la lisibilité et optimiser la présentation.
- Lister les méthodes avancées d'analyse spatiale, expliquer leur logique, appliquer des traitements dans MapInfo, analyser les résultats et évaluer leur pertinence.
- Définir le concept de carte thermique, expliquer son utilité, appliquer la création dans MapInfo, analyser la visualisation et optimiser les paramètres.
- Décrire les fonctionnalités de Marketplace, expliquer son intérêt, appliquer l'installation d'extensions, analyser leur usage et évaluer leur valeur ajoutée.

**4.2 : VOLUMETRIE PREVISIONNELLE**

MAPINFO Initiation : 10 à 15 stagiaires par an

MAPINFO Perfectionnement : 10 à 15 stagiaires par an

Ces chiffres sont donnés à titre indicatif et n'engagent pas la Caf des Hauts-de-Seine sur un volume de prestations.

\*\*\*