

Exploitation DADS à destination des chercheurs

Création d'un fichier de production et de recherche (FPR)

Contexte

Les chercheurs exploitent actuellement le fichier DADS super-anonymisé disponible pour tout public sur le site www.insee.fr.

Il a été décidé lors du comité client produit de janvier 2010 de leur fournir un fichier avec non plus les tranches de rémunération mais les rémunérations réelles.

Objectif

Le but de la demande est d'obtenir deux tables SAS de niveau postes et salariés sur l'année 2010. Un fichier ne comprend qu'une validité.

Il s'agit des fichiers nationaux « postes » et « salariés » au 12^{ème} (échantillon au 12^{ème} constitué par tous les salariés nés au mois d'octobre).

Pour chacune des tables, mis à part pour les variables de rémunération, il s'agit de fournir des fichiers à l'image des fichiers détails insee.fr de l'année considéré.

NOUVEAUTES 2010

Les zones d'emploi sont désormais dans la nomenclature : Zone d'emploi 2011.

Les descriptifs des fichiers postes et salariés sont les suivants

2010 - Fichier postes 1/12^e

Informations concernant le salarié

L'âge du salarié (AGE) ;

Le sexe du salarié (SEXE) ;

La région de résidence (REGR) ;

Le département de résidence (DEPR) ;

La zone d'emploi de résidence (ZEMPR) : zone d'emploi 2011 ; ATTENTION

Le code étranger (IND_ETRANGER).

Informations sur l'établissement

La région d'implantation (REGT) ;

Le département d'implantation (DEPT) ;

La zone d'emploi d'implantation (ZEMPT) : : zone d'emploi 2011 ; ATTENTION

Le domaine d'emploi du lieu d'affectation (DOMEMPL) : **attention nouvelle définition depuis**

2009 et nouvelles modalités ;

Le domaine d'emploi de l'employeur (DOMEMPL_EMPL) : **attention nouvelle variable depuis 2009 ;**

L'activité en Naf, rév2 (A6, A17 et A38) ;

La tranche d'effectif (TREFF calculé à partir de TREFFECT).

Informations sur le poste

La date de début de période (DATDEB) ;

La date de fin de période (DATFIN) ;

La durée de la période (DUREE) ;
Le nombre d'années-travail (AT) calculé à partir de la variable DUREE (DUREE/360) ;
Le type de poste (annexe ou non annexe : FILT) ;
La convention collective du salarié (CONV_COLL) ;
La condition d'emploi (CPFD) ;
La PCS codée (PCS) ;
La catégorie socio-professionnelle (CS) : 2 1ers caractères de la PCS ;
Le type d'emploi (TYP_EMPLOI) ;
Le nombre d'heures salariées (NBHEUR) ;
L'équivalent temps-plein (ETP) ;
Le contrat de travail (CONT_TRAV) calculé à partir de la variable CONTRAT_TRAVAIL ;
La rémunération brute (S_BRUT).

Les modalités ou valeur des variables sont les mêmes que ceux des fichiers de diffusion à l'exception :

TREFF

. = Non disponible
'00' = 0 poste (salariés présents en cours d'année, mais pas au 31/12)
'01' = 1 à 9 postes
'02' = 10 à 19 postes
'03' = 20 à 49 postes
'04' = 50 à 99 postes
'05' = 100 à 249 postes
'06' = 250 postes et plus

CONT_TRAVAIL

'CDI' = CDI
'CDD' = CDD
'APP' = apprentissage
'TTP' = mission d'intérim, travail temporaire
'AUT' = autres
'ZZZ' = contrat de travail manquant

TYPEMPLOI

'O' = Ordinaire
'A' = Apprenti
'X' = Autre

2010 - Fichier salariés 1/12^e

Informations concernant le salarié

L'âge du salarié (AGE) ;
Le sexe du salarié (SEXE) ;
La région de résidence (REGR) ;
Le département de résidence (DEPR) ;
La zone d'emploi de résidence (ZEMPR) : zone d'emploi 2011 ; ATTENTION

Informations sur l'établissement du poste principal

La région d'implantation (REGT) ;
Le département d'implantation (DEPT) ;
La zone d'emploi d'implantation (ZEMPT) : : zone d'emploi 2011 ; ATTENTION
Le domaine d'emploi du lieu d'affectation (DOMEMPL) : **attention nouvelle définition depuis**

2009 et nouvelles modalités ;

L'activité en Naf, rév2 (A6, A17 et A38) ;
La tranche d'effectif (TREFF) calculé à partir de TREFEFFECT.

Informations sur le poste principal

La date de début de période (DATDEB) ;

La date de fin de période (DATFIN) ;
La durée de la période (DUREE) ;
La convention collective du salarié (CONV_COLL) ;
La condition d'emploi (CPFD) ;
La PCS codée (PCS) ;
La catégorie socio-professionnelle (CS) : 2 1ers caractères de la PCS ;
Le type d'emploi (TYP_EMPLOI) ;
Le nombre d'heures salariées (NBHEUR) ;
Le contrat de travail (CONT_TRAV) calculé à partir de la variable CONTRAT_TRAVAIL.

Informations sur l'ensemble des postes effectués par le salarié

Le total du nombre d'heures (NBHEUR_TOT) ;
Le nombre de périodes effectuées (NB_PER) ;
Le nombre de périodes non annexes effectuées (NB_PER_N)
Le total des rémunérations nettes (S_NET_TOT) ;
Le total des indemnités ASSEDIC (ALCHO).

Les modalités ou valeur des variables sont les mêmes que ceux des fichiers de diffusion à l'exception :

TREFF

. = Non disponible
'00' = 0 poste (salariés présents en cours d'année, mais pas au 31/12)
'01' = 1 à 9 postes
'02' = 10 à 19 postes
'03' = 20 à 49 postes
'04' = 50 à 99 postes
'05' = 100 à 249 postes
'06' = 250 postes et plus

CONT_TRAVAIL

'CDI' = CDI
'CDD' = CDD
'APP' = apprentissage
'TTP' = mission d'intérim, travail temporaire
'AUT' = autres
'ZZZ' = contrat de travail manquant

TYPEMPLI

'O' = Ordinaire
'A' = Apprenti
'X' = Autre

Méthodologie

La Déclaration Annuelle des Données Sociales (DADS) est une formalité déclarative que doit accomplir toute entreprise employant des salariés, en application de l'article R243-14 du Code de la Sécurité Sociale (décret du 24 mars 1972) et des articles 87.240 et 241 de la loi 51-711 du 7 juin 1951 du Code Général des Impôts.

La DADS contient les informations destinées à l'usage de différentes institutions (URSSAF, CPAM, IRCANTEC, UNEDIC, CNRACL) et à l'INSEE pour l'observation des salaires et du volume de travail associé. Depuis 1993, l'exploitation porte sur l'ensemble des employeurs et de leurs salariés, dans le champ défini ci-dessous.

A partir de la validité 2009, sont ajoutées aux informations issues de la source DADS, celles construites à partir des fichiers de paye des agents de l'Etat et des fichiers de déclaration des particuliers qui emploient des salariés.

En résumé, à compter de la validité 2009, le champ de diffusion des DADS - grand format couvre l'ensemble des employeurs et de leurs salariés, sauf les activités extra-territoriales (division NAF Rev2 99).

L'exploitation des différentes sources est pour partie informatique et pour partie manuelle. Les traitements manuels sont effectués dans trois pôles, situés dans les directions régionales de Reims et de Dijon pour les salariés du secteur privé et dans la direction régionale de Rennes pour ceux des secteurs publics ; les traitements des salariés des particuliers-employeurs sont beaucoup plus automatisés et suivis uniquement à EFA.. La chaîne de traitement informatique des DADS a été profondément modifiée à partir de la validité des DADS 2002 et bien aménagée pour la validité 2009. La nouvelle chaîne d'exploitation repose sur un traitement complet de chaque DADS, entreprise par entreprise, au fur et à mesure de l'arrivée de celles-ci, l'objectif étant de terminer tous les traitements d'une entreprise le plus rapidement possible. La production et la maintenance des 3 chaînes de traitement sont assurées par le CNIO.

Le fichier échantillon national « postes » (échantillon au 12ème) : données individuelles ainsi que données « établissements » redescendues au niveau « postes » ;

Le fichier échantillon national « salariés » au 12ème ;

Particularités propres aux validités 2009 et 2010

2009

❖ Extension du champ

A partir de la validité 2009, le champ est étendu à l'ensemble des salariés, à l'exception des activités extra-territoriales. Le champ est donc étendu sur l'année 2009 aux :

- particuliers employeurs
- salariés de la fonction publique d'Etat

A noter :

- les particuliers employeurs possèdent des SIREN et NIC fictifs qui sont des identifiants calculés par l'application PE (Particuliers Employeurs). Il ne s'agit pas de véritables numéros Siren. Un SIREN correspond donc à un employeur et le NIC à un lieu de travail pour cet employeur. **Ces SIREN fictifs commencent par la lettre 'P'.**

Comme les fichiers des années précédentes et pour être cohérent avec les fichiers détail Insee.fr, les secteurs dont l'Apet commencent par 01 et 02 ont été retirés des fichiers FPR mis à disposition.

❖ Cas particulier de la fonction publique d'Etat : ajout de la notion d'employeur

Dans les données DADS, y compris sur la validité 2009, le SIREN de l'entreprise est toujours celui du lieu de travail du salarié. Sur les validités précédentes cette notion était confondue avec celle de SIREN employeur.

A partir de 2009, ce n'est plus le cas avec l'ajout de la fonction publique d'Etat. En effet, un professeur de lycée fonctionnaire sera déclaré par le SIREN de son établissement scolaire mais son employeur (Ministère de l'Éducation Nationale) possède un autre SIREN et relève d'une autre catégorie juridique.

C'est la raison pour laquelle il a été ajouté une **nouvelle variable relative à l'employeur du salarié** :

- DOMEMPL_EMPL = domaine d'emploi de l'employeur, donnée calculée notamment à partir de la catégorie juridique employeur mais au contour affiné pour certains SIREN (*voir paragraphe suivant concernant cette variable*)

Cette variable de niveau employeur est le pendant de DOMEMPL déjà disponibles dans les DADS et relative au lieu de travail. Si le SIREN employeur se confond avec le SIREN de lieu de travail, cette nouvelle variable est égales à DOMEMPL correspondant au SIREN lieu de travail.

- SIREN_EMPL = SIREN de l'entreprise employeur
→ si SIREN = SIREN_EMPL alors DOMEMPL_EMPL=DOMEMPL

❖ Nouvelle définition pour DOMEMPL

La variable « domaine d'emploi » a été entièrement redéfinie pour mieux s'adapter au grand format DADS. Voici la nouvelle définition de DOMEMPL : (qui se calcule à partir de la catégorie juridique et du SIREN)

| DOMEMPL= | Intitulé | Filtre sur catégorie juridique et SIREN |
|----------|--|---|
| 0 | Entreprises absentes du référentiel | CJ=00 |
| 1 | Fonction Publique d'Etat | CJ=71** et CJ=7331 et CJ de 7382 à 7389 (à l'exception des SIREN 180014045,180035016, 180035024, 180035032, 180035065, 180036063, 189400039) + siren 180020026 (CDC) |
| 2 | Fonction Publique Territoriale | CJ=72** et CJ=7312 à 7314 et CJ=7341 à 7363 et CJ=7371 à 7379 + siren=180014045 (CNFPT) |
| 3 | Fonction Publique Hospitalière | CJ=7364 à 7366 + siren 189400039 et 180036063 |
| 4 | Autres organismes publics administratifs | CJ=74** (sauf siren 180020026) et CJ=7321, 7322, 7323, 7381 + siren 180035016, 180035065,180035024 et 180035032 |
| 5 | Personnes morales de droit public soumises au droit commercial | CJ=4*** |
| 6 | Entreprises individuelles | CJ=1*** (sauf source PE) et CJ=5498 |
| 7 | Particuliers Employeurs | données marquées comme issues de PE |
| 8 | Organismes privés spécialisés et groupements de droit privé | CJ=8*** et CJ=9*** |
| 9 | Autres sociétés privées | CJ=2*** et CJ=3*** et CJ=5*** (sauf 5498) et CJ=6*** |

Remarques :

- la variable DOMEMPL_EMPL présente la même définition mais pour l'employeur.

❖ Nouveauté 2009 : codification de la PCS

Jusqu'à la validité 2008 des DADS, la variable de mesure de la qualification était la CS-ESE à 2 chiffres (catégorie socioprofessionnelle). A partir de la validité 2009 figure désormais la codification de la profession à 4 positions codée selon la nomenclature PCS-ES (il s'agit d'une nomenclature adaptée aux DADS disponible sur www.insee.fr). Les deux premières positions de la PCS permettent de se ramener à la CS.

La finalité d'une codification en PCS-ES est de disposer d'une information plus complète, notamment en acquérant une notion de profession, ce que ne permettait pas la CS.

La variable PCS est disponible sur l'ensemble du champ salarié.

Cette PCS-ES est déterminée selon différentes méthodes en fonction de l'application source :

- dans l'application DADS (secteur privé), elle est obtenue à partir du libellé d'emploi grâce à un système de codification utilisant le logiciel Sicore et bénéficie comme pour la CS d'un traitement manuel ciblé sur une partie de l'échantillon au 1/12^{ème}.

Dans l'application SIASP (fonctions publiques), elle se déduit de la PCS déclarée par l'employeur, redressée et contrôlée ensuite par l'information du grade.

- Dans l'application PE (particuliers employeurs), la PCS s'obtient à partir du métier déclaré complété par d'autres variables comme le code d'exonération de l'employeur ou le type de personnel.

➤ **Rupture sur la partie secteur privé (source=DADS)**

Une refonte du système de codification a donc été réalisée dans l'application de gestion DADS pour parvenir à une PCS-ES comme indication de la qualification. En conséquence, il y'a une **rupture de série** entre la CS 2009 (= 2 premiers chiffres de la PCS) et la CS 2008.

En particulier, la PCS 2009 code davantage d'employés (CS 5) et moins de professions intermédiaires (CS 4). Il y'a également une hausse du codage de cadres (du fait de l'utilisation du statut professionnel : variable plus fiable pour repérer les cadres).

➤ **Rupture sur le champ fonction publique territoriale et hospitalière (source=SIASP en 2009 et DADS en 2008)**

La variable PCS_1 réropolée n'a pas pu être calculée sur ce champ, elle est donc à valeur manquante dans les fichiers DADS 2009 sur cette population. Par ailleurs, une comparaison de la structure de la CS à l'aide des fichiers DADS de diffusion 2008 et des fichiers de diffusion 2009 sur le champ FPT ou FPH amènera à un risque de rupture, car ces deux CS ont été calculées selon deux méthodes différentes (application DADS en 2008 et application SIASP en 2009).

❖ **Nomenclature agrégée**

Pour la nomenclature agrégée A6 (issue de la NAF rév 2), la modalité « GU » a été corrigée en modalité « JU » sur les deux années (variables A6 et A6_1).

❖ **ETP**

L'utilisation de la variable ETP telle que décrite dans le guide méthodologique des DADS n'est possible que sur le secteur privé (hors PE). Pour un champ plus large il faut en redéfinir une au moyen des variables disponibles sur tout le champ (nombre d'heures, durée du poste, APET, etc...).

2010

Procédure CONTENTS

Répertoire

Libref SORTIE
Moteur V9
Nom physique X:\HAB-PSM-DADS\FPR
Nom du fichier X:\HAB-PSM-DADS\FPR

| N | Nom | Type de membre | Taille du fichier | Modifié(e) le |
|---|----------|----------------------|----------------------|---------------------------|
| 1 | POSTAN10 | DATA | 394236928 | 21 septembre 2012 16 h 01 |
| 2 | SALAN10 | DATA | 299312128 | 21 septembre 2012 16 h 03 |

Procédure CONTENTS

| Nom de la table | SORTIE.POSTAN10 | Observations |
|----------------------------|------------------------------------|----------------------------|
| 3496924 | | |
| Type de membre | DATA | Variables 29 |
| Moteur | V9 | Index 0 |
| Créée | vendredi 21 septembre 2012 15 h 58 | Longueur d'observation 112 |
| Dernière modification | vendredi 21 septembre 2012 15 h 58 | Observations supprimées 0 |
| Protection | | Compressée NON |
| Type de table | | Triée NON |
| Libellé | | |
| Représentation des données | WINDOWS_64 | |
| Codage | wlatin1 Western (Windows) | |

Informations dépendantes de la machine/de l'hôte

| | |
|-------------------------------------|---------------------------------------|
| Taille de la page | 12288 |
| Nombre de pages | 32083 |
| Première page de données | 1 |
| Nb max. d'obs. par page | 109 |
| Obs. sur première page de données | 69 |
| Nombre de corrections dans la table | 0 |
| Nom du fichier | X:\HAB-PSM-DADS\FPR\postan10.sas7bdat |
| Version de création | 9.0202M3 |
| Hôte de création | X64_ESRV08 |

Liste alphabétique des variables et des attributs

| N | Variable | Type | Long. | Format | Informat | Libellé |
|----|--------------|-------|-------|--------|----------|--------------------------------------|
| 14 | A6 | Texte | 2 | \$2. | \$2. | NOMENCLATURE AGREGEE A6 |
| 15 | A17 | Texte | 2 | \$2. | \$2. | NOMENCLATURE AGREGEE A17 |
| 16 | A38 | Texte | 2 | \$2. | \$2. | NOMENCLATURE AGREGEE A38 |
| 2 | AGE | Num. | 8 | | | AGE DU SALARIE |
| 27 | AT | Num. | 8 | 8.2 | | ANNEE TRAVAIL |
| 29 | CONT_TRAV | Texte | 3 | \$3. | | CONTRAT DE TRAVAIL |
| 24 | CONV_COLL | Texte | 4 | | | CONVENTION COLLECTIVE NOUVELLE NORME |
| 3 | CPFD | Texte | 1 | | | CONDITION D'EMPLOI |
| 28 | CS | Texte | 2 | \$2. | | CATEGORIE SOCIOPROFESSIONNELLE |
| 4 | DATDEB | Num. | 8 | | | DATE DE DEBUT |
| 5 | DATFIN | Num. | 8 | | | DATE DE FIN |
| 6 | DEPR | Texte | 3 | \$3. | \$3. | DEPARTEMENT DE RESIDENCE |
| 7 | DEPT | Texte | 3 | \$3. | \$3. | DEPARTEMENT DE TRAVAIL |
| 8 | DOMEMPL | Texte | 1 | \$1. | \$1. | DOMAINE D'EMPLOI (LIEU DE TRAVAIL) |
| 23 | DOMEMPL_EMPL | Texte | 1 | | | DOMAINE D'EMPLOI (EMPLOYEUR) |
| 9 | DUREE | Num. | 8 | | | DUREE |
| 10 | ETP | Num. | 8 | | | EQUIVALENT TEMPS PLEIN |
| 12 | FILT | Texte | 1 | | | TYPE DE POSTE |
| 11 | IND_ETRANGER | Texte | 1 | | | CODE ETRANGER |
| 13 | NBHEUR | Num. | 8 | | | NOMBRE D'HEURES SALARIEES |
| 17 | PCS | Texte | 4 | | | PCS |
| 18 | REGR | Texte | 2 | \$2. | \$2. | REGION DE RESIDENCE |
| 19 | REGT | Texte | 2 | \$2. | \$2. | REGION DE TRAVAIL |
| 1 | SEXE | Texte | 1 | \$1. | \$1. | SEXE |
| 20 | S_BRUT | Num. | 8 | | | SALAIRE BRUT |
| 25 | TREFF | Texte | 2 | \$2. | | TRANCHE D'EFFECTIFS ETABLISSEMENT |
| 26 | TYPEMPL | Texte | 1 | \$1. | | TYPE D'EMPLOI |
| 21 | ZEMPR | Texte | 4 | \$4. | \$4. | ZONE D'EMPLOI DE RESIDENCE |
| 22 | ZEMPT | Texte | 4 | \$4. | \$4. | ZONE D'EMPLOI DE TRAVAIL |

Procédure CONTENTS

| Nom de la table | SORTIE.SALAAAN10 | Observations |
|----------------------------|------------------------------------|----------------------------|
| 2192177 | | |
| Type de membre | DATA | Variables 30 |
| Moteur | V9 | Index 0 |
| Créée | vendredi 21 septembre 2012 16 h 01 | Longueur d'observation 136 |
| Dernière modification | vendredi 21 septembre 2012 16 h 01 | Observations supprimées 0 |
| Protection | | Compressée NON |
| Type de table | | Triée NON |
| Libellé | | |
| Représentation des données | WINDOWS_64 | |
| Codage | wlatin1 Western (Windows) | |

Informations dépendantes de la machine/de l'hôte

| | |
|-------------------------------------|---------------------------------------|
| Taille de la page | 12288 |
| Nombre de pages | 24358 |
| Première page de données | 1 |
| Nb max. d'obs. par page | 90 |
| Obs. sur première page de données | 57 |
| Nombre de corrections dans la table | 0 |
| Nom du fichier | X:\HAB-PSM-DADS\FPR\salaan10.sas7bdat |
| Version de création | 9.0202M3 |
| Hôte de création | X64_ESRV08 |

Liste alphabétique des variables et des attributs

| N | Variable | Type | Long. | Format | Informat | Libellé |
|----|--------------|-------|-------|--------|----------|--------------------------------------|
| 15 | A6 | Texte | 2 | | | NOMENCLATURE AGREGEE A6 |
| 16 | A17 | Texte | 2 | | | NOMENCLATURE AGREGEE A17 |
| 17 | A38 | Texte | 2 | | | NOMENCLATURE AGREGEE A38 |
| 2 | AGE | Num. | 8 | | | AGE |
| 24 | ALCHO | Num. | 8 | | | TOTAL DES INDEMNITES ASSEDIC |
| 29 | CONT_TRAV | Texte | 3 | \$3. | | CONTRAT DE TRAVAIL |
| 25 | CONV_COLL | Texte | 4 | | | CONVENTION COLLECTIVE NOUVELLE NORME |
| 3 | CPFD | Texte | 1 | | | CONDITION D'EMPLOI |
| 30 | CS | Texte | 2 | \$2. | | CATEGORIE SOCIOPROFESSIONNELLE |
| 5 | DATDEB | Num. | 8 | | | DATE DE DEBUT |
| 6 | DATFIN | Num. | 8 | | | DATE DE FIN |
| 7 | DEPR | Texte | 3 | | | DEPARTEMENT DE RESIDENCE |
| 8 | DEPT | Texte | 3 | | | DEPARTEMENT DE TRAVAIL |
| 9 | DOMEMPL | Texte | 1 | | | DOMAINE D'EMPLOI |
| 26 | DOMEMPL_EMPL | Texte | 1 | | | DOMAINE D'EMPLOI (EMPLOYEUR) |
| 10 | DUREE | Num. | 8 | | | DUREE |
| 11 | FILT | Texte | 1 | | | TYPE DE POSTE |
| 14 | NBHEUR | Num. | 8 | | | NOMBRE D'HEURES SALARIEES |
| 23 | NBHEUR_TOT | Num. | 8 | | | TOTAL DES NOMBRES D'HEURES |
| 12 | NB_PER | Num. | 8 | | | NOMBRE DE PERIODES |
| 13 | NB_PER_N | Num. | 8 | | | NOMBRE DE PERIODES NON ANNEXES |
| 4 | PCS | Texte | 4 | | | PCS-ESE DECLAREE |
| 18 | REGR | Texte | 4 | | | REGION DE RESIDENCE |
| 19 | REGT | Texte | 4 | | | REGION DE TRAVAIL |
| 1 | SEXE | Texte | 1 | \$1. | \$1. | SEXE |
| 22 | S_NET_TOT | Num. | 8 | | | TOTAL DES S-NET |
| 27 | TREFF | Texte | 2 | \$2. | | TRANCHE D'EFFECTIC ETABLISSEMENT |
| 28 | TYPEMPL | Texte | 1 | \$1. | | TYPE D'EMPLOI |
| 20 | ZEMPR | Texte | 4 | | | ZONE D'EMPLOI DE RESIDENCE |
| 21 | ZEMPT | Texte | 4 | | | ZONE D'EMPLOI DE TRAVAIL |