

MINISTÈRE DE L'ÉDUCATION NATIONALE ET DE LA JEUNESSE,
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE ET DE L'INNOVATION

Secrétariat général

Paris, le 4 janvier 2021

Direction de
l'évaluation,
de la prospective
et de la performance

Note technique

Sous-direction des
évaluations et de la
performance
scolaire

Bureau
des études statistiques sur
les élèves

DEPP B1/EB

Affaire suivie par :

Evelyne Blanché
Pascal Dieusaert
e-mel :
evelyne.blanche@
education.gouv.fr
pascal.dieusaert@
education.gouv.fr

Téléphone :
01 55 55 74 35
01 55 55 72 57

61-65 rue Dutot
75732 Paris Cedex 15

Objet : Transmission de fichiers de données de la DEPP au réseau Quetelet

Comme convenu entre le directeur de la DEPP et le directeur du centre Maurice Halbwachs, unité partenaire du réseau Quetelet, la DEPP met à disposition de ce réseau les fichiers suivants :

- un fichier sur les élèves (caractéristiques socio-démographiques et scolaires), extraits de la Base Centrale SCOLARITE (BCS) pour l'année 2020-2021
- un fichier sur les établissements scolaires (caractéristiques administratives et géographiques) pour l'année 2020-21.

1 - Caractéristiques des fichiers

Les fichiers de données sont des fichiers texte, de longueur fixe, avec des « ; » pour séparateur de champ.

Les dessins d'enregistrements de chaque type de fichier (fichier « élèves »/fichier « établissements ») se trouvent en pages 2 et 3 de la présente note. L'indication de l'existence ou non d'une nomenclature associée figure en face de chaque variable.

2 - Les nomenclatures

Deux fichiers de nomenclatures (élèves/établissements) commentées accompagnent la transmission des fichiers de données.

Pour chaque nomenclature sont précisées le ou les variables associées, repérées par leur nom dans le dessin d'enregistrement.

3 - Liste des fichiers transmis

Le fichier « élèves » :	DEPP-elev-Qt2020	(année 2020-21)
Le fichier « établissements » :	DEPP-etab-Qt2020	(année 2020-21)
Les fichiers de nomenclatures :	DEPP-nomen-elev-Qt2020.xls	
	DEPP-nomen-etab-Qt2020.xls	

Les données de la Base Centrale SCOLARITE (BCS) à destination du centre Quetelet

1 – Fichier sur les élèves caractéristiques socio-démographiques et scolaires

Variables	Long.	Contenu	Observations	Nomenclature associée O/N
AN_SCO	4	Année scolaire	Année basse de rentrée (2004 pour 2004-2005)	N
SI_SCO	1	Source de l'info (ou mode de collecte)	SCOLARITE ='1'	N
SEXE	1	Caractéristiques de l'élève Sexe		O
AGE	2	Age en années révolues au 31.12 de l'année basse	(Année basse de rentrée) – (Année de naissance)	N
NATIO_G	1	Français/étranger		O
REGIME	1	Régime scolaire	Interne, externe, demi-pensionnaire, ...	O
RESP_G	1	Responsable de l'élève	Parent/ autre	O
PCS1	2	Origine sociale	Nomenclature INSEE recomposée en 32 postes	O
DRES	3	Département de résidence		O
NUM_ETAB	8	Scolarité de l'année en cours N Identifiant de l'établissement N	Permet de faire le lien avec le fichier « établissements »	N
MEFST6	6	Formation suivie l'année N	Nomenclature hiérarchisée	O
GR_NSF	3	Groupes de spécialités de formation	Nomenclature interministérielle (NSF)	O
OPT1 à OPT4	6*4	Langues vivantes	4 au maximum. Code sur 6 caractères	O
		Scolarité de l'année précédente N-1		
SECTEURP	1	Secteur de l'établissement N-1	Public/Privé	O
NUMETABP	8	Identifiant de l'établissement N-1	A défaut de l'identifiant, N° du département	N
GEOSTATP	8	Situation géographique de l'étab N-1	Nomenclature hiérarchisée	O
MEFST6P	6	Formation suivie l'année N-1	Nomenclature hiérarchisée	O
GR_NSFP	3	Groupes de spécialités de formation de l'année N-1	Nomenclature interministérielle (NSF)	O

2 – Fichier des établissements

caractéristiques administratives et géographiques

Variable	Long.	Contenu	Observations	Nomenclature associée O/N
AN_SCO	4	Année scolaire	Année basse de rentrée	
GEOSTAT	8	Situation Géographique de l'étabt scolaire	Nomenclature hiérarchisée (Académie, ...)	O
NUM_ETAB	8	Identifiant de l'établissement N	Permet de faire le lien avec le fichier « élèves »	N
MINISTER	2	Ministère de tutelle	Source : MEN – Base Centrale des Etablissements	O
SECTEUR	1	Secteur d'enseignement	Source : MEN –BCE (Public / Privé)	O
TYP_ETAB	1	Type d'établissement scolaire	Source : MEN –BCE (Lycée, LP, Collège)	O
CONTRA_E	1	Contrat de l'établissement scolaire	Source : MEN –BCE (Privé sous / hors contrat)	O
NATURE	3	Nature de l'Etablissement scolaire	Source : MEN – Base Centrale des Etablissements	O
ETABSENS	1	Etablissement sensible ou non	Source : MEN - Base Centrale des Etablissements	O
APP_ECLAIR	4	Appartenance à un ECLAIR	Variable construite à partir de la BCE	
APP_REPPLUS	1	Appartenance à un REPPLUS	Variable construite à partir de la BCE	
APP_REP2015	1	Appartenance à un REP créée en 2015	Variable construite à partir de la BCE	
SECT_SCO	7	Secteur de carte scolaire	Source : MEN - Base Centrale des Etablissements	Dept
ZONE_EMP	4	Zone d'Emploi (associée à commune)	Source : INSEE - France Métro seulement	N
COMINSEE	5	N° de la commune_INSEE d'implantation	Source : INSEE - France Métro et DOM	N
LIBINSEE	32	Libellé de la commune_INSEE	Paris, Lyon & Marseille par arrondissement	N
POPINSEE	9	Population de la commune_INSEE au RP99	Paris, Lyon & Marseille par arrondissement	N
TYP_COMM	2	Type de commune au RP99	Source : INSEE - France Métro seulement	O
TRANCHUU	1	Tranche d'Unité Urbaine au RP99	Source : INSEE - France Métro et DOM	O
SI_SOURCE	1	Source de l'info (ou mode de collecte)	SCOLARITE ='1'	N
APP_ZUS	1	Appartenance à une ZUS	Variable construite à partir de la BCE	O

ZUS : Zone urbaine sensible

ECLAIR : Ecoles, Collèges, Lycées Ambition, Innovation, Réussite
 REPPLUS Réseaux d'éducation prioritaire mis en place en 2014

REP2015 Réseaux d'éducation prioritaire mis en place en 2015 (pour ne pas confondre avec les anciens REP existant en 1995)

Le système d'information SCOLARITE

SCOLARITE est le système d'information (SI) sur la scolarité des élèves dans les établissements publics et privés du second degré (collèges, lycées, lycées professionnels) dépendant du ministère de l'éducation nationale.

C'est un système partagé entre les trois niveaux qui assurent la gestion et le pilotage du système éducatif, à savoir :

- l'établissement
- l'académie
- l'administration centrale (Ministère)

Il prend appui sur les bases de gestion des élèves implantées dans les établissements et sur les référentiels communs que sont le répertoire des établissements (immatriculation des établissements et informations relatives à leurs caractéristiques administratives et géographiques) et la base des nomenclatures. Ces référentiels et les règles de gestion associées assurent la cohérence du système.

Le cœur du système d'information est constitué à chacun des trois niveaux d'une base de données « élèves » qui alimente un ensemble complexe de bases ou d'applications connexes.

Les bases « élèves » fondamentales sont :

- la Base Elèves Etablissement (BEE), dans laquelle se fait la saisie des informations concernant les élèves. Il en existe autant que de collèges et lycées (environ 11 000).
- la base nationale du Système d'information Statistique Consolidé Académique (SYSCA), qui est alimentée par les données en provenance des BEE. S'agissant d'une base de gestion, elle est alimentée en continue.
- la Base Centrale SCOLARITE (BCS), qui est alimentée chaque année en octobre par les extractions de données en provenance de SYSCA. Elle est sous la responsabilité de la DEPP. Les données transmises au réseau Quételet sont issues de cette base.

Les échanges d'information entre les différents niveaux se font de manière montante et descendante.

Le SI SCOLARITE, qui comporte des données nominatives pour les besoins de gestion, a fait l'objet d'une déclaration à la CNIL. Dans l'arrêté de création, sont mentionnés de façon extrêmement précise quels sont les destinataires de chaque catégorie d'informations, et notamment qui a accès aux données nominatives. L'administration centrale, dans la mesure où elle n'est pas impliquée dans les actes de gestion, n'est pas habilitée actuellement à détenir des données nominatives (exception faite des élèves des classes post-bac, pour leur repérage éventuel dans le système d'information sur le suivi des étudiants – SISE). Cette absence d'identifiant quelconque est évidemment préjudiciable à l'analyse des parcours scolaires, et donc à certains aspects du pilotage du système éducatif lui-même. La mise au point du système FAERE de cryptage des données nominatives (identifiants), qui a recueilli l'agrément de la CNIL a fait évoluer la situation.

Historique de l'implantation de SCOLARITE

Expérimenté à la rentrée 90 dans une académie, SCOLARITE couvre tous les lycées et collèges publics de métropole et des DOM depuis la rentrée 1993. Il s'est étendu très progressivement aux établissements privés sous contrat (signature de protocoles d'accord entre l'état et les partenaires de l'enseignement privé sous contrat contraignant ces derniers à informatiser la gestion de leurs

enseignants et de leurs élèves), pour atteindre l'exhaustivité à la rentrée 2003. Par contre rien n'oblige les établissements privés hors contrat à remonter des données individuelles. Ils sont donc diversement couverts selon les académies.

Couverture du SI SCOLARITE

Globalement, SCOLARITE couvre 99% des élèves des établissements du second degré publics et privés (sous contrat et hors contrat) dépendant du MEN (France métropolitaine + DOM). Ce sont ces données qui sont mises à disposition du réseau Quêtelet.