

Présentation du jeu de données publiées sur le site « transparence.sante.gouv.fr »

Le jeu de données comprend 3 vues sous formes de fichiers au format .csv :

- Vue sur les avantages : c'est un tableau avec en ligne les déclarations et en colonne les informations décrivant chaque déclaration ; dans la colonne type, il est précisé « avantage ».
- Vue sur les conventions : c'est un tableau avec en ligne les déclarations et en colonne les informations décrivant chaque déclaration ; dans la colonne type, il est précisé « convention ».
- Vue sur les Entreprises : c'est un tableau listant les entreprises en ligne et en colonne les informations sur ces entreprises.

Les informations sur les bénéficiaires sont contenues dans les vues précédentes.

Les déclarations faisant l'objet d'une demande de rectification lors du téléchargement font l'objet d'une mention spécifique.

La fréquence de rafraîchissement du jeu de données est journalière.

I- Définition de la base de données

Une base de données est un entrepôt organisé d'informations.

La base de données relationnelle transparence-santé est une base de données informatique où l'information est structurée dans des tableaux à deux dimensions (lignes/colonnes) appelés des relations ou « tables ».

Une table contient 0 à n lignes. Par exemple, lorsqu'une entreprise transmet une nouvelle déclaration sur l'espace Entreprise de transparence-santé, elle enregistre une déclaration dans la base de données : cette déclaration vient s'enregistrer sous forme d'une nouvelle ligne de déclaration dans la table des Déclarations. Autre exemple : à chaque fois qu'une nouvelle entreprise vient s'inscrire dans le dispositif transparence-santé, une nouvelle ligne d'entreprise vient s'ajouter dans la table des Entreprises.

Dans une base relationnelle, chaque ligne a besoin d'être décrite par des dimensions temporelles, quantitatives ou qualitatives (n° ID, date [date], montant [euros TTC], nature [texte], ...). Par exemple, dans la table des Déclarations, comment décrire chaque déclaration ? Chaque ligne de déclaration est décrite grâce à des champs qui contiennent des informations précises : date de la déclaration, montant de la déclaration, bénéficiaire de la déclaration, n° d'identification de l'entreprise qui a déposé la déclaration, n° d'identifiant unique de la déclaration (n° de registre), Ces champs sont renseignés par l'application transparence-santé, après que le déclarant a saisi et validé les informations de la déclaration. Ils ne sont autres que les colonnes de la table des Déclarations qui décrivent chaque déclaration.

Certains champs ou colonne peuvent être partagés par 2 ou plusieurs tables, ainsi le champ n° d'identification de l'entreprise existe aussi bien dans la table des Entreprises que dans celle des Déclarations.

Table des Déclarations					
N° Id interne de la déclaration	Type de la déclaration	Date déclaration	montant	N°ID de l'entreprise	bénéficiaire
1	avantage	15/07/2014	10€ TTC	[code interne_entpeA]	Bénéficiaire x
2	convention	20/07/2015		[code interne_entpeB]	Bénéficiaire y

Table des Entreprises			
N° Id de la déclaration	Date d'inscription	Dénomination sociale	adresse
[code interne_entpeA]	11/07/2014	Entreprise A	Adresse de l'entpe A
[code interne_entpeB]	30/11/2016	Entreprise B	Adresse de l'entpe B

Il est possible d'importer le-s fichier-s d'extraction dans un système de gestion de bases de données relationnelles et d'y faire les requêtes.

II- Liste des champs présents dans les vues extraites de la base de données.

Vous trouverez ci-dessous la définition de chaque champ contenu dans chaque vue.

a) La vue Entreprise

Nom du champ	Unicité	Descriptif du champ	Format	Liste de valeurs / Commentaire
ENTREPRISE_IDENTIFIANT	OUI Le couple (ENTREPRISE_IDENTIFIANT, LIGNE_IDENTIFIANT) est unique	Identifiant unique de l'entreprise ayant signé cette convention ou versé cet avantage	DENTIFIANT	Cette valeur permet d'attacher cette ligne de déclaration à l'entreprise (mandant) concernée si une entreprise (mandataire) transmet cette ligne de déclaration. Cette colonne permet de gérer le cas d'une entreprise qui déclare pour d'autres entreprises lui ayant donné mandat. Sinon l'entreprise déclarante indique son propre identifiant récupéré lors de son inscription sur le site unique.
ENTREPRISE_PAYS_CODE	non	Pays de l'entreprise	texte	
SECTEUR_ACTIVITE_CODE	non	Secteur d'activité de l'entreprise	texte	

Nom du champ	Unicité	Descriptif du champ	Format	Liste de valeurs / Commentaire
DENOMINATION_SOCIALE	Non	Dénomination sociale	Texte	
ADRESSE 1	Non			
ADRESSE 2	Non			
ADRESSE 3	Non			
ADRESSE 4	Non			
CODE_POSTAL	Non		CODEPOSTA L	
VILLE	Non	Ville du siège sociale de l'entreprise	TEXTE	

b) La vue déclaration (avantage ou convention)

Nom du champ	Unicité	Descriptif du champ	Format	Liste de valeurs / Commentaire
LIGNE_IDENTIFIANT	OUI Le couple (ENTREPRI SE_IDENTIF IANT, LIGNE_IDE	Identifiant unique de la ligne de déclaration (convention ou avantage) dans le système de l'entreprise déclarante	IDENTIFIANT	Cette valeur joue le rôle de clé pour identifier de manière unique cette ligne de déclaration (avantage ou convention). Elle

	NTIFIANT) est unique			est fournie en entrée par l'entreprise pour pouvoir par exemple procéder à une correction ultérieure.
LIGNE_ACTION	Non	Code action	LISTE	<ul style="list-style-type: none"> - [C] pour Création - [M] pour Modification - [S] pour Suppression - [V] pour Validation (cas d'un retour suite à alerte) <p>Important : lorsque le champ LIGNE_ACTION est égal à [V], aucun contrôle de type ALERTE n'est réalisé puisque l'entreprise valide les données de cette ligne de déclaration. Par contre les contrôle de type ERREUR sont toujours exécutés.</p>
LIGNE_DEMANDE_RECTIFICATION	Non	Ligne corrige une demande de rectification d'un bénéficiaire.	LISTE	<ul style="list-style-type: none"> - [O] Oui - [N] Non <p>Si [O], suppression du marqueur de "demande de rectification du bénéficiaire" sur le site grand public</p>
BENEF_CATEGORIE	Non	Catégorie du bénéficiaire	LISTE	<ul style="list-style-type: none"> - [PRS] Les professionnels de santé relevant de la quatrième partie du présent code - [APS] Les associations de professionnels de santé - [ETU] Les étudiants se destinant aux professions relevant de la quatrième partie du présent code ainsi que les associations et groupements les représentant - [AUS] Les associations d'utilisateurs du système de santé - [ETA] Les établissements de santé relevant de la sixième partie du présent code - [FON] Les fondations, les sociétés savantes et les sociétés ou organismes de conseil intervenant dans le secteur des produits

				ou prestations mentionnés au premier alinéa - [PRE] Les entreprises éditrices de presse, les éditeurs de services de radio ou de télévision et les éditeurs de services de communication au public en ligne - [LOG] Les éditeurs de logiciels d'aide à la prescription et à la délivrance - [PMO] Les personnes morales assurant la formation initiale des professionnels de santé mentionnés au 1° ou participant à cette formation - [ADU] Les associations d'étudiants
BENEF_NOM	Non	Nom du bénéficiaire	TEXTE	
BENEF_PRENOM	Non	Prénom du bénéficiaire	TEXTE	

BENEF_QUALITE	Non	Qualité / Profession du bénéficiaire	LISTE	-[10] Médecin -[40] Chirurgien-dentiste -[50] Sage-femme -[21] Pharmacien -[01] Préparateur en pharmacie et préparateur en pharmacie hospitalière -[60] Infirmier -[70] Masseur-kinésithérapeute -[80] Pédicure-podologue -[94] Ergothérapeute -[96] Psychomotricien -[91] Orthophoniste -[92] Orthoptiste -[98] Manipulateur d'électroradiologie médicale -[86] Technicien de laboratoire médical -[05] Audioprothésiste -[28] Opticien-lunetier -[82] Prothésiste et orthésiste pour l'appareillage des personnes handicapées -[95] Diététicien -[02] Aide soignant -[03] Auxiliaire de puériculture -[04] Ambulancier -[EDP] Editeur de presse
BENEF_ADRESSE1	Non	Adresse du bénéficiaire	TEXTE	- Adresse professionnelle si la valeur du champ BENEF_CATEGORIE est égale à [PRS] - Adresse siège social ou équivalent si la valeur du champ BENEF_CATEGORIE est égale à [APS] ou [AUS] ou [ETA] ou [FON] ou [PRE] ou [LOG] ou [PMO] ou [ADU]
BENEF_ADRESSE2	Non	Complément 1 de l'adresse du bénéficiaire	TEXTE	
BENEF_ADRESSE3	Non	Complément 2 de l'adresse du bénéficiaire	TEXTE	
BENEF_ADRESSE4	Non	Complément 3 de l'adresse du bénéficiaire	TEXTE	
BENEF_CODEPOSTAL	Non	Code postal du bénéficiaire	CODEPOSTAL	

BENEF_VILLE	Non	Ville du bénéficiaire	TEXTE	
BENEF_PAYS	Non	Pays du bénéficiaire	LISTE	- Norme ISO 3166-1 alpha-2 (entre crochet) Exemple : [FR] pour France
BENEF_TITRE	Non	Titre du bénéficiaire	LISTE	-[PR] Professeur -[MG] Médecin Général -[PG] Pharmacien Général -[PC] Pharmacien Chef -[MC] Médecin chef -[DR] Docteur -[AUTRE] Autre
BENEF_SPECIALITE	Non	Spécialité (ou discipline) du bénéficiaire	LISTE	- [SCD01] Orthopédie dento-faciale - [SCD02] Chirurgie Orale - [SCD03] Médecine Bucco-Dentaire - [SM01] Anatomie et cytologie pathologiques - [SM02] Anesthésie-réanimation - [SM03] Biologie médicale - [SM04] Cardiologie et maladies vasculaires - [SM05] Chirurgie générale - [SM06] Chirurgie maxillo-faciale - [SM07] Chirurgie maxillo-faciale et stomatologie - [SM08] Chirurgie orthopédique et traumatologie - [SM09] Chirurgie infantile - [SM10] Chirurgie plastique reconstructrice et esthétique - [SM11] Chirurgie thoracique et cardio-vasculaire - [SM12] Chirurgie urologique - [SM13] Chirurgie vasculaire - [SM14] Chirurgie

			viscérale et digestive - [SM15] Dermatologie et vénéréologie - [SM16] Endocrinologie et métabolisme - [SM17] Génétique médicale - [SM18] Gériatrie - [SM19] Gynécologie médicale - [SM20] Gynécologie- obstétrique - [SM21] Hématologie - [SM22] Hématologie (option Maladie du sang) - [SM23] Hématologie (option Onco-hématologie) - [SM24] Gastro- entérologie et hépatologie - [SM25] Médecine du travail - [SM26] Qualifié en Médecine Générale - [SM27] Médecine interne - [SM28] Médecine nucléaire - [SM29] Médecine physique et réadaptation - [SM30] Néphrologie - [SM31] Neuro- chirurgie - [SM32] Neurologie - [SM33] Neuro- psychiatrie - [SM34] O.R.L et chirurgie cervico faciale - [SM35] Oncologie (option onco- hématologie) - [SM36] Oncologie option médicale - [SM37] Oncologie option radiothérapie - [SM38] Ophtalmologie - [SM39] Oto-rhino- laryngologie - [SM40] Pédiatrie - [SM41] Pneumologie - [SM42] Psychiatrie - [SM43] Psychiatrie option enfant & adolescent - [SM44] Radio- diagnostic
--	--	--	---

				<ul style="list-style-type: none"> - [SM45] Radio-thérapie - [SM46] Réanimation médicale - [SM47] Recherche médicale - [SM48] Rhumatologie - [SM49] Santé publique et médecine sociale - [SM50] Stomatologie - [SM51] Gynéco-obstétrique et gynécologie médicale option 1 - [SM52] Gynéco-obstétrique et gynécologie médicale option 2 - [SM53] Spécialiste en Médecine Générale - [SM54] Médecine Générale - [SM55] Radio-diagnostic et Radio-Thérapie - [SM99] ORL et ophtalmologie - [SP01] Radio-pharmacie - [SP02] Hygiène - [SP03] Pharmacovigilance - [SP04] Hémovigilance - [AUTRE] Autre
BENEF_QUALIFICATION	Non	Qualification du bénéficiaire	TEXTE	Champ volontairement libre
BENEF_IDENTIFIANT_TYPE	Non	Type d'identifiant utilisé pour identifier le bénéficiaire	LISTE	<p>- Si le champ BENEF_CATEGORIE est égale à [PRS], les valeurs autorisées sont : ' [RPPS] pour l'identifiant personnel du bénéficiaire dans le Répertoire Partagé des Professionnel de Santé, [ORDRE] pour le numéro d'inscription à l'ordre, [AUTRE] pour les autres cas</p> <p>- Si le champ BENEF_CATEGORIE est égale à [APS] ou [AUS] ou [ETA] ou [FON] ou [PRE] ou [LOG] ou [PMO] ou [ADU], les valeurs autorisées sont : ' [SIREN] pour le numéro SIREN, [FINESS] pour le</p>

				<p>numéro FINESS, [AUTRE] pour les autres cas</p> <p>- Si le champ BENEFCATEGORIE est égale à [ETU], les valeurs autorisées sont : [RPPS], [AUTRE]</p>
BENEF_IDENTIFIANT_VALEUR	Non	Valeur de l'identifiant du bénéficiaire	TEXTE	<p>- Si le champ BENEFCATEGORIE_TYPE égal à [RPPS], identifiant personnel du bénéficiaire dans le Répertoire Partagé des Professionnel de Santé</p> <p>- Si le champ BENEFCATEGORIE_TYPE égal à [ORDRE], numéro d'inscription à l'ordre</p> <p>- Si le champ BENEFCATEGORIE_TYPE égal à [SIREN], le numéro siren</p> <p>- Si le champ BENEFCATEGORIE_TYPE égal à [FINESS], le numéro finess.</p> <p>- Valeur libre si le champ BENEFCATEGORIE_TYPE égal à [AUTRE] et la valeur est renseignée. Ne rien mettre si inconnu.</p>
BENEF_ETABLISSEMENT	Non	Nom de l'établissement d'enseignement	TEXTE	
BENEF_ETABLISSEMENT_CODEPOSTAL	Non	Code postal de l'établissement d'enseignement	CODEPOSTAL	
BENEF_ETABLISSEMENT_VILLE	Non	Ville de l'établissement d'enseignement	TEXTE	

BENEF_DENOMINATIONSOCIA LE	Non	Dénomination sociale du bénéficiaire	TEXTE	
BENEF_OBJETSOCIAL	Non	Objet social du bénéficiaire	TEXTE	
LIGNE_TYPE	Non	Avantage ou convention	LISTE	- [A] pour Avantage - [C] pour Convention
CONV_DATE_SIGNATURE	Non	Date de signature de la convention	DATE	Format JJ/MM/AAAA
CONV_OBJET	Non	Objet de la convention	TEXTE	
CONV_DATE_DEBUT	Non	Date de début de couverture de la convention	DATE	Format JJ/MM/AAAA
CONV_DATE_FIN	Non	Date de fin de couverture de la convention	DATE	Format JJ/MM/AAAA
CONV_EVENEMENT_DATE	Non	Date de l'événement lié à la convention	DATE	Format JJ/MM/AAAA A compléter si la convention concerne une manifestation (remplace le dépôt du programme sous forme de PDF)
CONV_EVENEMENT_NOM	Non	Nom de l'événement lié à la convention	TEXTE	A compléter si la convention concerne une manifestation (remplace le dépôt du programme sous forme de PDF)
CONV_EVENEMENT_LIEU	Non	Lieu de l'événement lié à la convention	TEXTE	A compléter si la convention concerne une manifestation (remplace le dépôt du

				programme sous forme de PDF)
CONV_EVENEMENT_ORGANISATEUR	Non	Organisateur de l'événement lié à la convention	TEXTE	A compléter si la convention concerne une manifestation (remplace le dépôt du programme sous forme de PDF)
AVANT_DATE_SIGNATURE	Non	Date où l'avantage a été accordé	DATE	Format JJ/MM/AAAA
AVANT_MONTANT_TTC	Non	Montant T.T.C. arrondi à l'euro le plus proche	MONTANT	
AVANT_NATURE	Non	Nature de l'avantage accordé	TEXTE	
AVANT_CONVENTION_LIE	Non	Intitulé ou référence de la convention éventuellement liée à cet avantage (pas de lien informatique avec la convention)	TEXTE	
Avant_CONV_SEMESTRE	oui	correspond au semestre de la déclaration. Elle est utile pour les avantages et les conventions		

III- Exploitation des vues

L'exploitation des données doit s'effectuer sur un progiciel permettant d'afficher plus de 2 millions de lignes.

A l'aide de nombreux logiciels, notamment libres de droit, il vous appartient de charger les vues par une requête. La démarche est la suivante :

- créer une base de données
- créer 3 tables : avantages, conventions, entreprises ; ajouter les colonnes qu'elles contiennent en vérifiant bien le format (texte, date, ...).
- loader les 3 tables (en retirant la 1^{ère} ligne)

Les tags des demandes de rectification des bénéficiaires, si elles existent au moment de la demande d'extraction sont également compris dans ces vues.

- <https://www.transparence.sante.gouv.fr/exports-etalab/exports-etalab.zip>
- Volumétrie de l'archive zip: 380Mo