

Les bases statistiques SAE 2018 – Documentation

Sommaire :

A-	Accès aux données et contenu des dossiers	1
B-	Contenu des corrections	2
C-	Ruptures de séries.....	3
D-	Tables fournies	3
a-	Données d'activité et de capacités.....	3
b-	Données sur le personnel médical	6
c-	Données sur le personnel non médical	6
d-	Nouvelles bases mises à disposition	7

A- Accès aux données et contenu des dossiers

Pour consulter les bases statistiques SAE 2018, il faut dé-zipper le dossier contenant les tables. Pour cela, il faut cliquer sur *SAE 2018 – Bases statistiques.zip*, et enregistrer le fichier dans un répertoire. Il faut ensuite double-cliquer sur ce dossier, et cliquer sur « Extraire ».

Une fois dé-zippé, le répertoire « SAE 2018 Bases statistiques » contient 3 sous-répertoires :

- un sous-répertoire « Bases statistiques »
- un sous-répertoire « Séries longues »
- un sous-répertoire « Documentation »

Le sous-répertoire « Bases statistiques » contient deux dossiers :

- un dossier « Base SAS » avec les différentes tables redressées au format SAS ;
- un dossier « Base CSV » avec les différentes tables redressées au format CSV.

Chacun de ces dossiers contient l'intégralité des bases (dans deux formats différents). Lorsqu'elles sont suffixées par la lettre « a », en fin de nom de fichier, cela indique que la base est identique à la base administrative, contrairement au cas où elles sont suffixées par la lettre « r », qui indique que la base est redressée.

Le sous-répertoire « Séries longues » contient les tables de séries longues, proposées là encore dans deux formats différents, avec les deux dossiers suivants :

- un dossier « Base SAS » avec les tables de séries longues au format SAS ;
- un dossier « Base CSV » avec les tables de séries longues au format CSV.

Chacun de ces dossiers contient (dans deux formats différents) :

- trois tables contenant des données structurantes d'activité et de capacité (CAPACT18 pour l'année 2018, CAPACT0018 et CAPACT_TOT0018 pour la période 2000-2018) ;
- deux tables contenant des informations sur les effectifs et ETP de personnel pour la période 2001-2018 (PM_01_18 pour le personnel médical et PNM_01_18 pour le personnel non médical).

Le sous-répertoire « Documentation » contient :

- les formats SAS ;
- l'aide au remplissage de l'enquête (*SAE_AR_2018.pdf*) ;
- les nomenclatures de l'enquête (*Nomenclatures_2018.xls*) ;
- le fichier *SAE2018_Dictionnaire_variables_par_bordereaux.xlsx* : il établit, pour chaque bordereau de la SAE, une correspondance entre le visuel du questionnaire et les variables des fichiers CSV ou SAS ;
- le fichier *Liste Variables.xlsx* : il reprend les informations du fichier précédent (*SAE2018_Dictionnaire_variables_par_bordereaux.xlsx*) mais sous forme de tableau. Il contient, pour chaque table SAS, la liste des variables ainsi que leur libellé (et éventuellement la modalité associée), et la case du questionnaire à laquelle elle correspond ;
- ce document.

Vous pourrez trouver d'autres informations concernant la SAE sur la page de la DREES :

<http://drees.solidarites-sante.gouv.fr/etudes-et-statistiques/open-data/etablissements-de-sante-sociaux-et-medico-sociaux/article/la-statistique-annuelle-des-etablissements-sae>

Pour toute question, vous pouvez nous écrire à l'adresse : drees-sae@sante.gouv.fr

B- Contenu des corrections

Les corrections de la SAE 2018 sont de deux ordres : harmonisation et redressement des données.

L'harmonisation des données concerne la mise en cohérence des données entre bordereaux (cohérence entre les bordereaux Q20/Q21 et Q23/Q24 particulièrement, dont les ETP salariés et les effectifs de libéraux doivent être égaux), ainsi que l'harmonisation des autorisations avec le référentiel ARHGOS (notamment pour les autorisations des maternités, des urgences, en soin de suite et réadaptation, etc.).

Le redressement des données : contrairement aux bases administratives, qui doivent retranscrire exactement les réponses des établissements géographiques à la SAE telles que validées par eux suite à la phase de validation des données collectées, les bases statistiques permettent de réaliser, si nécessaire, des corrections de valeurs aberrantes, manquantes ou des corrections d'incohérences qui n'auraient pu être traitées pendant la phase de validation de la collecte.

À compter de la SAE 2018, nous apportons également une correction de la non réponse partielle et totale. Les établissements qui n'ont pas validé leur questionnaire mais ont renseigné une partie des données (statut de l'enquête « en cours de saisie ») ainsi que ceux n'ayant pas saisi leur SAE sont ajoutés aux bases statistiques.

Pour les établissements « en cours de saisie », les données complétées par l'établissement sont récupérées et intégrées aux bases statistiques, et les quelques données manquantes (notamment les capacités et les données de personnel) sont ajoutées en utilisant les données N-1. Pour les établissements n'ayant pas saisi leur SAE ou l'ayant validée mais n'ayant pas rempli un bordereau qu'ils auraient dû remplir (parce qu'ils ont l'autorisation par exemple), nous récupérons les données

PMSI pour l'activité et nous imputons les données manquantes en utilisant, le plus souvent, les données N-1 (pour les bordereaux et variables les plus importants). Par la suite, la dénomination « non réponse totale par bordereau » correspond à la correction de cette non réponse totale ou partielle.

Après intégration de ces établissements, le taux de couverture de l'enquête est de 99,9 % (contre 99,6 % de taux de réponse en base administrative).

Attention : un incendie au CHU de Pointe-à-Pitre/Abîmes en Guadeloupe fin 2017 a endommagé une partie des bâtiments, entraînant le déménagement temporaire des services impactés vers d'autres établissements. Afin de pouvoir séparer l'activité du CHU de ces autres établissements, des Finess « provisoires » ont été créés. Les données du CHU sont donc réparties sur 3 Finess au lieu d'un seul : les Finess 970100442 (CHU de Pointe-à-Pitre/Abîmes), 970112835 (Polyclinique de Guadeloupe - site CHU) et 970112850 (Clinique les eaux claires – site CHU).

C- Ruptures de séries

La SAE a subi une refonte très importante en 2013, sur laquelle vous trouverez des informations sur le site de la DREES ou dans la documentation relative aux bases administratives. **En raison de cette refonte, les évolutions 2012-2013 sont difficiles à analyser.**

Pour faciliter l'exploitation des données, nous avons malgré tout mis à disposition des tables qui regroupent les années 2000 à 2018. Cela ne signifie pas que l'évolution annuelle est interprétable : sur un grand nombre de variables, les concepts ont changé. Ainsi, les données d'activité importées du PMSI ne se basent pas exactement sur les mêmes notions que l'activité qui était déclarée dans la SAE précédemment : les évolutions 2012-2013 restent difficiles à analyser.

Attention : à partir de 2011, les données incluent Mayotte.

D- Tables fournies

a- Données d'activité et de capacités

Pour l'année 2018

La table **CAPACT18** (disponible dans le sous-répertoire « Séries longues ») regroupe les données structurantes d'activité et de capacités, par discipline d'équipement, pour chaque établissement interrogé en 2018.

Les bases auxquelles des corrections ont été apportées, par rapport aux bases administratives, sont listées ci-dessous en précisant les modifications effectuées :

La table **ID_2018R** : elle est au même format que la table ID_2018 des bases administratives, mais elle contient des corrections sur la variable ETAT_SAISIE.

La table **IDB_2018R** : elle est au même format que la table IDB_2018 des bases administratives, mais elle contient des corrections sur le champ.

La table **ST_2018R** : elle est au même format que la table ST_2018 des bases administratives, mais elle contient des corrections sur le champ.

La table **ORG_2018R** : elle est au même format que la table ORG_2018 des bases administratives, mais elle est corrigée de la non-réponse totale au bordereau.

La table **PN_2018R** : elle est au même format que la table PN_2018 des bases administratives, mais elle contient des corrections sur l'existence d'UHSA et sur le champ (correction de la non-réponse totale au bordereau).

La table **FILTRE_2018R** : elle est au même format que la table FILTRE_2018 des bases administratives, mais elle contient des corrections sur les activités et équipements disponibles. Elle est également corrigée de la non-réponse totale au bordereau.

La table **MCO_2018R** : elle est au même format que la table MCO_2018 des bases administratives, mais elle contient des corrections sur les capacités, le nombre de journées exploitables et sur l'activité (y compris l'activité externe). Elle est également corrigée de la non-réponse totale au bordereau.

La table **MCOGER_2018R** : elle est au même format que la table MCOGER_2018 des bases administratives, mais elle contient des corrections sur les capacités (lits et places) et sur l'activité.

La table **PSY_2018R** : elle est au même format que la table PSY_2018 des bases administratives, mais elle contient des corrections sur les capacités et l'activité (temps complet, partiel et ambulatoire).

La table **PSY2_2018R** : elle est au même format que la table PSY2_2018 des bases administratives, mais elle contient des corrections sur le nombre d'actes ambulatoire par mode légal de soins.

La table **SSR_2018R** : elle est au même format que la table SSR_2018 des bases administratives, mais elle contient des corrections sur les autorisations ainsi que sur le nombre de lits, places, journées et séjours. Elle est également corrigée de la non-réponse totale au bordereau.

La table **SSR2_2018R** : elle est au même format que la table SSR2_2018 des bases administratives, mais elle contient des modifications sur les équipements disponibles suite à la correction de la non-réponse totale au bordereau.

La table **SSR_P_2018R** : elle est au même format que la table SSR_P_2018 des bases administratives, mais elle contient des modifications sur les effectifs et ETP suite à la correction de la non-réponse totale au bordereau.

La table **USLD_2018R** : elle est au même format que la table USLD_2018 des bases administratives, mais elle contient des corrections sur l'activité et les capacités suite à la correction de la non-réponse totale au bordereau.

La table **USLD_P_2018R** : elle est au même format que la table USLD_P_2018 des bases administratives, mais elle contient des modifications sur les effectifs et ETP suite à la correction de la non-réponse totale au bordereau.

La table **HAD_2018R** : elle est au même format que la table HAD_2018 des bases administratives, mais elle contient des corrections sur le nombre de patients pouvant être pris en charge simultanément par la structure HAD, de séjours et de journées ainsi que sur la file active. Elle est également corrigée de la non-réponse totale au bordereau.

La table **BLOCS_2018R** : elle est au même format que la table BLOCS_2018 des bases administratives, mais elle contient des corrections sur le nombre de salles. Elle est également corrigée de la non-réponse totale au bordereau.

La table **IMAGES_2018R** : elle est au même format que la table IMAGES_2018 des bases administratives, mais elle est corrigée de la non-réponse totale au bordereau.

La table **IMAGES_P_2018R** : elle est au même format que la table IMAGES_P_2018 des bases administratives, mais elle contient des modifications sur les effectifs et ETP suite à la correction de la non-réponse totale au bordereau.

La table **PHARMA_2018R** : elle est au même format que la table PHARMA_2018 des bases administratives, mais elle est corrigée de la non-réponse totale au bordereau

La table **BIO_2018R** : elle est au même format que la table BIO_2018 des bases administratives, mais elle contient des corrections sur les effectifs et ETP de personnel concourant à l'activité.

La table **INFOMED_2018R** : elle est au même format que la table INFOMED_2018 des bases administratives, mais elle est corrigée de la non-réponse totale au bordereau.

La table **TELEMED_2018R** : elle est au même format que la table TELEMED_2018 des bases administratives, mais elle est corrigée de la non-réponse totale au bordereau.

La table **PERINAT_2018R** : elle est au même format que la table PERINAT_2018 des bases administratives, mais elle contient notamment des corrections sur les types de maternités et les capacités. Elle est également corrigée de la non-réponse totale au bordereau.

La table **CHIRCAR_2018R** : elle est au même format que la table CHIRCAR_2018 des bases administratives, mais elle contient des corrections sur les autorisations. Contrairement aux années passées, l'activité déclarée entièrement en prestation inter-établissement (PIE) par l'établissement a également été supprimée.

La table **DOULEUR_2018R** : elle est au même format que la table DOULEUR_2018 des bases administratives, mais elle contient des corrections sur le type structure (centre ou consultation).

La table **DIALYSE_2018R** : elle est au même format que la table DIALYSE_2018 des bases administratives, mais elle contient des corrections sur les capacités et le nombre de séances notamment. Elle est également corrigée de la non-réponse totale au bordereau.

La table **CANCERO_2018R** : elle est au même format que la table CANCERO_2018 des bases administratives, mais elle est corrigée de la non-réponse totale au bordereau.

La table **URGENCE2_2018R** : elle est au même format que la table URGENCE2_2018 des bases administratives, mais elle contient des corrections sur le nombre de passages aux urgences générales notamment.

La table **SMURSAMU_2018R** : elle est au même format que la table SMURSAMU_2018 des bases administratives, mais elle contient des corrections sur les autorisations SMUR et SAMU et sur les carences ambulancières.

La table **SMURSAMU_P_2018R** : elle est au même format que la table SMURSAMU_P_2018 des bases administratives, mais elle contient des corrections sur les effectifs et ETP de personnel concourant à l'activité.

La table **REA_2018R** : elle est au même format que la table REA_2018 des bases administratives, mais elle est corrigée de la non-réponse totale au bordereau.

Séries longues (disponibles dans le sous-répertoire « Séries longues »)

La table **CAPACT0018** contient, par discipline d'équipement, les données relatives à l'activité et aux capacités pour les années 2000 à 2018. Les données sont disponibles au niveau établissement géographique pour les établissements privés et sont regroupées au niveau entité juridique pour les établissements publics (pour être en cohérence avec le mode d'interrogation pré-refonte de 2013).

Les données couvrant 19 années, il existe potentiellement plusieurs ruptures de séries, dont celle (forte) correspondant à la refonte de l'enquête en 2013.

La table **CAPACT_TOT0018** reprend les données disponibles dans la table CAPACT0018, mais cette fois agrégées pour toutes les disciplines d'équipement, au niveau établissement (géographique pour le privé et entité juridique pour le public).

b- Données sur le personnel médical

Pour l'année 2018

La table **Q20_2018R** (personnel médical par spécialité exercée) : elle est au même format que la table Q20_2018 des bases administratives, mais elle contient des corrections sur les ETP et les effectifs suite à la correction de la non-réponse totale au bordereau et à la mise en cohérence avec le bordereau Q21.

La table **Q21_2018R** (personnel médical par groupe de disciplines d'équipement) : elle est au même format que la table Q21_2018 des bases administratives, mais elle contient des corrections sur les ETP et les effectifs suite à la correction de la non-réponse totale au bordereau et à la mise en cohérence avec le bordereau Q20.

La table **Q22_2018R** (internes et faisant fonction d'internes) : elle est au même format que la table Q22_2018 des bases administratives, mais elle contient des corrections sur les effectifs.

Séries longues (disponibles dans le sous-répertoire « Séries longues »)

La table **PM_01_18** contient, pour chaque établissement privé et entité juridique publique, le nombre d'ETP salariés et les effectifs des libéraux, pour chaque année de 2001 à 2018.

Les données couvrant 18 années, il existe potentiellement plusieurs ruptures de séries, dont celle (forte) correspondant à la refonte de l'enquête en 2013.

c- Données sur le personnel non médical

Pour l'année 2018

La table **Q23_2018R** (personnel non médical) : elle est au même format que la table Q23_2018 des bases administratives, mais elle contient des corrections sur les effectifs et ETP suite à la correction de la non-réponse totale au bordereau.

La table **Q24_2018R** (personnel non médical par disciplines d'équipement) : elle est au même format que la table Q24_2018 des bases administratives, mais elle contient des corrections sur les ETP de personnel non médical suite la correction de la non-réponse totale au bordereau et à la mise en cohérence avec le bordereau Q23.

Séries longues (disponibles dans le sous-répertoire « Séries longues »)

La table **PNM_01_18** contient, pour chaque établissement privé et entité juridique publique, le nombre d'ETP salariés par catégorie de personnel, pour chaque année de 2001 à 2018.

Les données couvrant 18 années, il existe potentiellement plusieurs ruptures de séries, dont celle (forte) correspondant à la refonte de l'enquête en 2013.

Les données de la table **PNM_01_18** bénéficient toutefois d'un travail de rétopolation des données antérieures à 2011, qui permet de s'affranchir de plusieurs ruptures de série. En effet, le champ du personnel dans la SAE a subi plusieurs évolutions entre 2008 et 2011, si bien qu'il n'est pas possible de calculer les évolutions annuelles en utilisant directement les tables de chaque année. Par

exemple, la SAE 2010 incluait les personnels des établissements sociaux, médico-sociaux et de formation, tandis qu'à partir de la SAE 2011, ceux-ci en sont exclus.

d- Nouvelles bases mises à disposition

Depuis la SAE 2016, deux nouvelles tables sont mises à disposition : les tables SYGEN_AAAAR et GHT_AAAAR.

La table **SYGEN_2018R** reprend les informations contenues dans la partie « Activité générale et activités autorisées » du bordereau SYGEN. Elle est au même format que la table SYGEN_2018 des bases administratives, mais elle est mise à jour des corrections apportées aux données des différents bordereaux du questionnaire.

La table **GHT_2018R** contient des données agrégées sur le contour des groupements hospitaliers de territoires (GHT) au 31 décembre tels que le nombre de lits, de places, de journées et de séjours et des informations sur les équipements, les plateaux techniques et le personnel des établissements de santé. Elle est au même format que la table GHT_2018 des bases administratives, mais elle est mise à jour des corrections apportées aux données des différents bordereaux du questionnaire.

À compter de la SAE 2018, une table supplémentaire est mise à disposition : la table HPR_AAAAR.

La table **HPR_2018R** contient les informations de la table SYGEN_2018R sur le contour des hôpitaux de proximité uniquement, selon la liste définie par l'arrêté d'avril 2018.