

## Source et méthode

### Fichier statistique des sortants de prison

<b>ORIGINE DES DONNEES</b>	<b>FURNISSEUR DES DONNEES</b> : ministère de la justice / DAP et DACG
	<b>NOM de la source en entrée</b> : Genésis et Condamnations des personnes physiques
	<b>TYPE de SOURCE</b> : Appariement de sources administratives
	<b>CHAMP</b> : - France Entière (métropole et DOM) et collectivités d'outre-mer - Personnes écrouées condamnées ayant purgé tout ou partie de leur peine en prison et qui ont connu une libération de fin de peine (sortants de prison)
	<b>UNITE DE COLLECTE</b> : Personne placée sous main de justice prise en charge par l'Administration Pénitentiaire / sortie de prison
	<b>DESCRIPTIF</b> Le fichier statistique des sortants de prison est issu de l'appariement des fichiers Genésis (Gestion Nationale des personnes Ecrouées pour le Suivi Individualisé et la Sécurité) et du fichier statistique du Casier Judiciaire National (CJN) des personnes physiques. Genésis est l'application de suivi des personnes détenues en établissement pénitentiaire. Elle succède à l'appli GIDE et a été déployée progressivement dans les établissements pénitentiaires entre octobre 2013 et novembre 2016, date à laquelle les données sont considérées comme étant exhaustives. Genésis permet de compter et de caractériser les personnes écrouées ainsi que leurs conditions de détention pour celles qui sont hébergées. Quant au fichier statistique du CJN, il enregistre depuis 1984 les informations relatives aux condamnations définitives des personnes physiques. Il permet de calculer la récidive.
<b>TRAITEMENTS STATISTIQUES</b>	<b>RESPONSABLE du TRAITEMENT</b> : ministère de la justice / SG / SEM / SDSE
	<b>OBJECTIFS et FINALITES</b> Mesurer la récidive des sortants de prison. Identifier les déterminants de la récidive (caractéristiques des détenus, conditions de détention, nature des infractions pénalement sanctionnées...)
	<b>TRAITEMENTS et REDRESSEMENTS (cf détail en annexe)</b> La sous-direction de la statistique et des études (SDSE) du ministère de la justice, service statistique ministériel, réalise chaque année l'appariement des fichiers de Genésis et du fichier statistique du Casier sur la base d'un triplet de variables indirectement identifiantes : la date de naissance du détenu, la date du jugement et le lieu du jugement. Cette opération nécessite au préalable des transformations des tables en entrée (concaténations, jointures, opérations groupées) et des créations et redressements de variables. Des calculs d'indicateurs sont également réalisés portant sur la récidive après la sortie de prison, sur le passé pénal du détenu, la vie en détention et sur les aménagements de peine pour enrichir l'analyse de la récidive des sortants de prison.
<b>DIFFUSION</b>	<b>FICHIERS STATISTIQUES</b> - <b>Descriptif</b> : le fichier statistique des sortants de prison est produit annuellement en allongeant le délai disponible pour la saisie des condamnations au casier ; d'abord provisoires, les taux de récidive des sortants de l'année N jusqu'à T années après la sortie sont considérés définitifs après 5 années de recul, soit lors de la production du fichier en fin d'année N+T+5. Ainsi, pour la cohorte des sortants de prison de 2016, le taux de récidive à 1 an est définitif lors de la production en fin d'année 2022 ; le taux à 2 ans l'est en fin 2023, et ainsi de suite. - <b>Périodicité</b> : annuelle - <b>Historicité</b> : les données sont disponibles à partir des sortants de prison de 2015
	<b>PRODUITS DE DIFFUSION</b> Infos Rapides Justice – études de la collection Infostat Justice
	<b>CONCEPTS IMPORTANTS (cf. Glossaire pour les définitions précises)</b> - PPSMJ écrouées - milieu fermé - condamnation - infraction - aménagement de peine - récidive légale et réitération
	<b>LIMITES et PRECAUTIONS D'UTILISATION (cf détail en annexe)</b>

- Nécessité de disposer d'un recul suffisant pour évaluer un taux jusqu'à un horizon fixé pour une cohorte de sortants donnée ;
- Exclusion du champ des sortants non appariés avec les données du casier (environ 8 %) qui peut induire un biais dans les calculs de taux de récidive, notamment du fait d'un profil différent entre détenus appariés ou non.
- Exploitation de Genésis pour la détermination de la condamnation ayant conduit en détention qui est actuellement imparfaite, ce qui peut rendre moins fiable la détermination de l'infraction principale à l'origine de l'incarcération ainsi que le passé pénal du détenu concerné.
- Incomplétude des données : les données 2015 et 2016, du fait du déploiement progressif de Genésis entre octobre 2013 et novembre 2016, sont incomplètes. De plus, lorsqu'il y a eu reprise de données entre l'applicatif GIDE et Genésis, les informations peuvent être partielles. C'est notamment le cas des informations relatives aux activités effectuées par les détenus qui n'étaient pas renseignées dans GIDE.

## **I – Traitements et redressements détaillés**

### **La transformation des données**

Les données extraites du logiciel Genésis sont retravaillées afin d'identifier le champ des sortants de prison, à savoir les personnes condamnées (y compris les condamnés-prévenus) ayant purgé tout ou partie de leur peine en prison et qui ont connu une libération de fin de peine. Il s'agit également de préparer les données en vue de l'appariement avec le fichier statistique du casier. Pour cela, il est nécessaire de reconstituer au niveau de la PPSMJ des informations disponibles au niveau de l'écrou et ainsi de récupérer des décisions de justice associées aux écrous, des mesures et pièces de justice associées aux décisions puis de déterminer l'infraction principale à mettre en face de chaque décision. Ces étapes permettent de disposer d'informations calculées au niveau du couple (PPSMJ, décision) qui seront utilisées pour l'appariement.

Les données du fichier statistique du casier sont elles aussi préparées en vue de l'appariement (concaténations de tables, normalisation des variables communes aux deux sources parmi lesquelles la date de naissance, la date de décision qui doit correspondre à une date de condamnation ou plus exactement à la date d'événement au casier, la commune de juridiction, le lieu de naissance, le sexe, les mesures associées à la décision, la nature d'affaire associée à l'infraction principale, l'infraction principale. Au total, une personne inscrite dans ce fichier suite à une condamnation est identifiée par son numéro de dossier qui est conservé tout au long de la vie de la personne.

### **L'appariement des données Genésis - CJN**

Il n'existe pas d'identifiant direct commun entre les deux sources Genésis et CJN. Après avoir repéré l'ensemble des sortants de prison dans Genésis, l'appariement avec les informations du CJN repose sur un identifiant indirect correspondant au triplet (date de naissance du détenu, date de jugement, commune de la juridiction de jugement). Dans la majorité des cas (91 %), une PPSMJ donnée est appariée à un seul numéro de dossier au casier. L'étape suivante consiste donc à départager les cas où plusieurs numéros de dossiers sont associés à une même PPSMJ. Ces correspondances multiples sont redressées à l'aide des variables communes aux deux sources non encore utilisées pour l'appariement. Au total, après une dernière étape de validation de l'appariement (score basé sur l'égalité de variables communes), le taux d'appariement atteint 92 % des sortants de prison repérés dans Genesis. Les sortants de prison non retrouvés dans les données du CJN ne sont pas présents dans le fichier statistique mis à disposition.

### **La création de variables complémentaires utiles à l'analyse de la récidive des sortants**

#### **Récidive des sortants de prison**

La récidive correspond dans ce fichier au fait de commettre, après la sortie de prison, une nouvelle infraction donnant lieu à condamnation, mais pas nécessairement pour des faits identiques à ceux ayant conduit en prison. Le délai de récidive est mesuré entre la date de libération de fin de peine et la date des faits de la nouvelle infraction. Différentes opérations sont menées, à partir du fichier statistique du casier, afin de filtrer dans un premier temps les condamnations pour des infractions commises après la levée d'écrou et dont la date d'événement intervient après la sortie, puis de calculer dans un second temps des indicatrices de première récidive à différents horizons (12, 18, 24 et 36 mois) ainsi que des nombres de condamnations survenues dans ces intervalles de temps (nombre total de condamnations à 12/18/24/36 mois après la sortie).

#### **Condamnation à l'origine de l'incarcération**

La condamnation à l'origine de la mise sous écrou est recherchée parmi les informations déjà calculées au niveau du couple (PPSMJ, décision) en s'appuyant sur la table des affaires dans Genésis et en conservant le premier code affaire. Seule la première date de décision est ensuite retenue et est supposée correspondre à la date de condamnation recherchée, tout comme l'infraction la plus grave qui est supposée être associée à la condamnation.

#### Passé pénal du détenu

L'affaire origine qui a conduit le détenu en prison ayant été retrouvée, il est possible de définir son passé pénal en exploitant l'historique des condamnations inscrites au casier depuis 1984 et en particulier de calculer le nombre de condamnations dans les 5 ans (respectivement 10 ans) avant la condamnation ayant entraîné la mise sous écrou.

#### Repérage des aménagements de peine

Le champ SDSE de la récidive des sortants exclut les personnes bénéficiant d'un aménagement de peine dès le début de l'exécution de peine, en placement sous surveillance électronique (PSE), placement à l'extérieur (PE) ou semi-liberté (SL). Des traitements sont nécessaires pour repérer ces personnes et les isoler, tout comme pour repérer les personnes ayant bénéficié d'un aménagement de peine après une période de détention et évaluer l'effet de ce dispositif sur la récidive. Les informations relatives aux aménagements (début de peine, fin de peine, aucun aménagement ou libération conditionnelle) sont ainsi constituées en travaillant principalement avec les tables de mouvements, de mesures ou de levées d'écrou dans Genésis et en opérant des redressements de certaines variables.

## **II – Limites et précautions d'utilisation détaillées**

### **Les limites liées au déploiement progressif de Genésis**

Le passage de Gide à Genésis a été réalisé progressivement d'octobre 2013 à novembre 2016. Pour la plupart des établissements pénitentiaires, la date de bascule a eu lieu avant le 31 décembre 2015 (environ 70 % des établissements dont 59 % au cours de la seule année 2015), les déploiements résiduels ayant eu lieu courant 2016 (à l'exception d'un centre de détention où le transfert s'est fait en mai 2017). Au total, les sortants de prison qui n'ont pas été détenus, au moins pour une partie de leur peine, dans un établissement équipé de Genésis sont ainsi hors du champ du fichier statistique. Cela concerne surtout les sortants de 2015 mais aussi dans une moindre mesure les sortants de 2016.

Néanmoins, certains sortants ont été incarcérés dans un établissement dans lesquels le suivi passait par Genésis, d'autres ont été inscrits dans Gide au moment de leur entrée en prison. Pour ces derniers, une reprise de données a eu lieu au moment où l'établissement est passé de Gide à Genésis, et implique des informations seulement partiellement disponibles. C'est notamment le cas des activités effectuées en prison (sport, formation, travail...) qui n'étaient pas renseignées dans Gide. Au total, une indicatrice permet de repérer les PPSMJ concernées par au moins une reprise de données, ce qui permet de connaître celles pour lesquelles une information complète est disponible (lorsque l'écrou initial est présent dans Genésis) et celles où l'information est partielle.

### **Nécessité de disposer d'un recul suffisant pour évaluer un taux jusqu'à un horizon fixé pour une cohorte donnée**

Compte tenu des délais parfois importants entre la commission d'une infraction et l'inscription au CJN (constatation, délais d'enquêtes, juridiques, administratifs, de traitements...), les indicateurs produits sont en principe révisables presque sans limite de temps. De fait, le délai associé à la récidive des sortants, c'est-à-dire la durée entre la date de libération et la date des faits de la nouvelle infraction (crime ou délit) condamnée, ne peut être évalué qu'à partir du moment où la condamnation de la nouvelle infraction est inscrite au casier, ce qui nécessite un recul temporel important. En pratique, pour concilier fiabilité de la mesure de la récidive et disponibilité de statistiques définitives dans un délai raisonnable, la SDSE retient le processus de diffusion suivant : d'abord provisoires, les taux de récidive des sortants de l'année N jusqu'à T années après la sortie sont considérés définitifs lors de la production du fichier statistique des sortants de prison en fin d'année N+T+5.

### **Non prise en compte des sortants non appariés dans le fichier statistique**

Environ 8 % des sortants de prison ne sont pas appariés (numéro de dossier de la PPSMJ non retrouvé au CJN). Actuellement, on les retire de l'analyse et du fichier statistique des sortants. Or, ces sortants ont des caractéristiques particulières pouvant influencer sur leur récidive (profil davantage jeune et de nationalité étrangère), d'où un biais potentiel dans les calculs de taux de récidive. Une prochaine avancée consistera à redresser l'appariement de ce biais induit par la non prise en compte au sein des cohortes des sortants non appariés.

### **Unité de compte dans les cohortes de sortants**

La SDSE suit la récurrence de cohortes de sortants de prison d'une année donnée : la cohorte des sortants de prison de l'année N correspond à l'ensemble des sortants appariés de l'année N et appartenant au champ SDSE d'analyse de la récurrence. Une même personne peut ainsi appartenir à deux cohortes distinctes et elle peut être comptée à plusieurs reprises au sein d'une même cohorte lorsqu'elle a connu plusieurs sorties la même année. Pour illustrer ce choix, la définition initialement retenue d'une cohorte impliquait de ne retenir que la première sortie sur la période 2015 – 2017, si bien qu'un détenu ayant connu une première sortie en 2016 et une seconde en 2017 faisait partie de la cohorte des sortants de 2016 mais pas de la cohorte des sortants de 2017. Ce choix conduisait à sous-estimer le taux de récurrence au fil des cohortes successives, en particulier du fait de l'exclusion des cohortes de détenus multirécidivistes.

### **La détermination de la condamnation à l'origine de la mise sous écrou est à améliorer**

Les traitements actuels, exploitant Genésis pour la détermination de la condamnation qui a conduit en détention, peuvent conduire pour certains sortants à une détermination imparfaite de l'infraction principale à l'origine de l'incarcération (et donc de la nature d'affaire associée à l'infraction), pouvant entraîner elle-même des imperfections dans le passé pénal du détenu. En attendant de mener des travaux complémentaires pour améliorer la qualité de ces variables, les incohérences directement identifiables entre l'infraction principale et la condamnation à une peine de prison restent mineures en ce qu'elles concernent moins de 0,1 % de l'ensemble des sortants (c'est le cas notamment lorsque la peine principale pour l'infraction affichée est une amende).