

N° 11-633-X au catalogue — N° 036
ISSN 2371-3437
ISBN 978-0-660-40863-7

Études analytiques : méthodes et références

Base de données longitudinales sur l'immigration (BDIM) Rapport technique, 2020

Date de diffusion : le 6 décembre 2021

 Statistique
Canada

Statistics
Canada

Canada 

Comment obtenir d'autres renseignements

Pour toute demande de renseignements au sujet de ce produit ou sur l'ensemble des données et des services de Statistique Canada, visiter notre site Web à www.statcan.gc.ca.

Vous pouvez également communiquer avec nous par :

Courriel à STATCAN.infostats-infostats.STATCAN@statcan.gc.ca

Téléphone entre 8 h 30 et 16 h 30 du lundi au vendredi aux numéros suivants :

- | | |
|---|----------------|
| • Service de renseignements statistiques | 1-800-263-1136 |
| • Service national d'appareils de télécommunications pour les malentendants | 1-800-363-7629 |
| • Télécopieur | 1-514-283-9350 |

Programme des services de dépôt

- | | |
|-----------------------------|----------------|
| • Service de renseignements | 1-800-635-7943 |
| • Télécopieur | 1-800-565-7757 |

Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois. À cet égard, notre organisme s'est doté de normes de service à la clientèle que les employés observent. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1-800-263-1136. Les normes de service sont aussi publiées sur le site www.statcan.gc.ca sous « Contactez-nous » > « [Normes de service à la clientèle](#) ».

Note de reconnaissance

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population du Canada, les entreprises, les administrations et les autres organismes. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques exactes et actuelles.

Publication autorisée par le ministre responsable de Statistique Canada

© Sa Majesté la Reine du chef du Canada, représentée par le ministre de l'Industrie 2021

Tous droits réservés. L'utilisation de la présente publication est assujettie aux modalités de l'[entente de licence ouverte](#) de Statistique Canada.

Une [version HTML](#) est aussi disponible.

This publication is also available in English.

Base de données longitudinales sur l'immigration (BDIM) Rapport technique, 2020

Diversité et statistique socioculturelle

11-633-X — N° 036

ISSN 2371-3437

ISBN 978-0-660-40863-7

Décembre 2021

Direction des études analytiques : méthodes et références

Les documents de cette série traitent des méthodes utilisées pour produire des données qui seront employées pour effectuer des études analytiques à Statistique Canada sur l'économie, la santé et la société. Ils ont pour but de renseigner les lecteurs sur les méthodes statistiques, les normes et les définitions utilisées pour élaborer des bases de données à des fins de recherche. Tous les documents de la série ont fait l'objet d'un examen par les pairs et d'une révision institutionnelle, afin de veiller à ce qu'ils soient conformes au mandat de Statistique Canada et qu'ils respectent les normes généralement reconnues régissant les bonnes pratiques professionnelles.

Les documents peuvent être téléchargés gratuitement de www.statcan.gc.ca.

Table des matières

Glossaire	9
Introduction.....	10
2 Sources des données.....	11
2.1 Données sur l'immigration.....	11
2.1.1 Fichier intégré des résidents permanents et non permanents (FRPN 1980-2020).....	11
2.1.2 Admissions avant 1980 : Fichier intégré des résidents permanents et non permanents (FRPN) 1952-1979.....	12
2.1.3 Fichier des résidents non permanents (FRN).....	12
2.1.4 Nouveau: Entrée express (EE).....	12
2.1.5 Fichier d'admission du Québec	12
2.2 Les modules additionnels de la BDIM.....	13
2.2.1 Module de données sur les enfants	13
2.2.2 Module de données sur les salaires et traitements.....	13
2.2.3 Le module sur les services d'établissement	14
2.3 Fichier des familles T1	15
2.4 Fichiers auxiliaires	16
3 Concepts et variables	17
3.1 Les statuts de l'immigrant au Canada.....	17
3.1.1 Un aperçu de l'immigration au Canada.....	17
3.2 La population cible et la période de couverture.....	18
3.3 Variables relatives à l'admission.....	19
3.3.1 Catégorie d'admission	19
3.3.2 Type de demandeur	23
3.3.4 Changements au fil du temps	24
3.3.5 Variables de catégorie d'admission du FRPN.....	24
3.4 Variables d'intérêt.....	25
3.4.1 Variables géographiques.....	26
3.4.2 Variables temporelles	27
3.4.3 Variables de scolarité	28
3.4.4 Variables de profession envisagée.....	28
3.4.5 Autres variables de la BDIM.....	28
4 Couplage d'enregistrements	30
5 Traitement des données	32
5.1 Traitement.....	32
5.2 Couplage du fichier des résidents non permanents (FRN).....	33
5.3 Variables dérivés inclus dans le FFT1	34

5.4 Variables dérivées incluses dans le FRPN	35
5.5 Détection des valeurs aberrantes.....	35
6 Dissémination	37
6.1 Produits analytiques.....	37
6.2 Demande de fichiers analytiques	38
6.3 Autres programmes statistiques utilisant les données de la BDIM.....	38
6.4 Confidentialité	39
7 Évaluation des données et indicateurs de qualité	40
7.1 Sources d'erreurs	40
7.1.1 Erreurs de couplage d'enregistrements	40
7.1.2 Erreurs de mesure	40
7.1.3 Erreurs de couverture.....	40
7.2 Exactitude des données.....	41
7.2.1 BDIM de 2020 : Taux de couplage.....	41
7.2.2 Disponibilité de la date de décès.....	43
7.2.4 Indicateur de conjoint.....	45
7.3 Variables imputées	46
7.3.1 Imputation des variables relatives à la scolarité.....	46
7.3.2 Imputation des variables relatives à la langue	47
7.4 Couverture.....	47
7.4.1 Couverture du Fichier intégré des résidents permanents et non permanents (FRPN).....	47
7.4.2 Taille du Fichier sur la famille T1 et couverture par année	48
7.5 Évaluation de la qualité de la données sur l'immigration	50
7.5.1 Évaluation de la qualité du Fichier intégré des résidents permanents et non permanents (FRPN).....	50
7.5.2 Évaluation de la qualité du fichier des résidents non permanents (NRF).....	52
7.6 Évaluation de la qualité de la variable de la province de résidence (PRCO_)	54
8 Comparabilité	56
8.1 Variation historique de la couverture.....	56
8.2 Changements méthodologiques	56
8.3 Changements du contenu de la base de données historiques.....	57
8.4 Changements du contenu de la base de données historiques.....	57
8.4.1 Banque de données administratives longitudinales (DAL).....	57
8.4.2 Recensement	59
8.4.3 Enquête longitudinale auprès des immigrants de Canada (ELIC).....	59
8.5 Discussion sur la BDIM avec différents couplages	60
8.5.1 Recensement	60
8.5.2 Enquête sur la santé dans les collectivités canadiennes (ESCC)	61

8.5.3 Métadonnées de la Base de données sur les congés des patients (BDCP).....	61
8.5.4 Enquête sociale générale (ESG).....	62
8.5.5 Banque de données administratives longitudinales (DAL).....	62
8.5.6 Enquête longitudinale auprès des immigrants du Canada (ELIC).....	63
9 Analyses possibles grâce à la BDIM	64
9.1 Possibilités analytiques découlant des données relatives aux résidents non permanents.....	64
9.2 Possibilités analytiques grâce aux données relatives aux décès.....	64
9.3 Capacités analytiques avec les données sur la citoyenneté.....	64
9.4 Capacités analytiques avec les données sur les enfants.....	64
9.5 Capacités analytiques avec les données sur l'entrée express.....	64
9.6 Capacités analytiques avec les données des fichiers sur les salaires et traitements	65
10 Résumé	66
Annexes.....	67
A. Liens vers les principaux documents et pages Web relatifs à la BDIM	67
B. Couverture	67
C. Analyses précédentes.....	69
D. Pratiques exemplaires et conseils pour les analystes	69
D.1 Conseils de programmation	69
D.2 Création d'une cohorte.....	72
D.3 Calcul des taux de rétention.....	72
D.4 Calcul des trajectoires de revenu au cours du temps	77
D.5 Arrondissement des données.....	78
D.6 Détection de valeurs aberrantes.....	78
D.7 Ajustement du revenu selon l'Indice des prix à la consommation (IPC).....	79
D.8 Calcul de mesures de revenu clés.....	80
Bibliographie.....	82

Remerciements

La publication initiale du rapport technique de la BDIM a été co-écrite par **Rose Evra** et **Elena Prokopenko**.

Nous souhaitons souligner la contribution particulière des personnes suivantes : **Laetitia Martin** de la Diversité et statistique socioculturelle (DSS), qui a rédigé les sections 3.3.1 à 3.3.3 de ce rapport, ainsi que **Tristan Cayn**, **Stephanie Cheng**, **Allan Dawe**, **Ian Marrs** et **Eric Mongrain**, membres de l'équipe des données administratives, qui produisent la BDIM et ont contribué au contenu de plusieurs sections du rapport.

La BDIM est le fruit d'un partenariat entre Statistique Canada, Immigration, Réfugiés et Citoyenneté Canada (IRCC) et les provinces. IRCC finance la BDIM depuis ses débuts et collabore continuellement avec Statistique Canada pour élargir le contenu de la BDIM en plus de contribuer au développement de nouveaux outils d'analyse.

Du côté provincial, les ministères suivants ont fait partie du consortium qui finance la BDIM :

Department of Advanced Education, Skills and Labour (Terre-Neuve-et-Labrador)
Island Investment Development Inc (Île-du-Prince-Édouard)
Nova Scotia Office of Immigration (Nouvelle-Écosse)
Division de la croissance démographique (Nouveau-Brunswick)
Ministère de l'Immigration, de la Francisation et de l'Intégration (Québec)
Ontario Ministère du Travail, de la Formation et du Développement des compétences
Travail et Immigration Manitoba
Ministry of Immigration and Career Training (Saskatchewan)
Alberta Labour
Labour Market Information Office (Colombie-Britannique)

Ces partenaires provinciaux ont également participé aux consultations visant à remanier la BDIM et ses outils d'analyse.

De plus, nous remercions les personnes suivantes d'avoir relu le rapport avant sa publication : Margareta Dovgal, Benoît St-Jean, Winnie Chan, Hélène Maheux, Scott McLeish et Tiana Major (Statistique Canada), Yoko Yoshida (Département de sociologie de l'Université de Western Ontario), Michael Haan (Chaire de recherche du Canada en migration et relations ethniques et Département de sociologie de l'Université de Western Ontario) et Ian Clara (Centre de données de recherche du Manitoba).

Résumé

La base de données longitudinales sur l'immigration (BDIM) est une source importante et complète de données contribuant à mieux comprendre le comportement économique des immigrants. Il s'agit du seul ensemble annuel de données canadiennes permettant d'étudier les caractéristiques des immigrants au Canada au moment de leur admission ainsi que leur situation économique et leur mobilité régionale (interprovinciale) sur plus de 35 ans.

La BDIM comprend les dossiers administratifs d'Immigration, Réfugiés et Citoyenneté Canada (IRCC) qui contiennent des renseignements exhaustifs sur les immigrants admis au Canada depuis 1952. Elle comprend également des données sur les résidents non permanents à qui on a délivré des permis de séjour temporaire depuis 1980.

Ce rapport traitera des sources de données, des concepts et des variables de la BDIM, du couplage d'enregistrements, du traitement des données, de la diffusion, de l'évaluation des données et des indicateurs de qualité, de la comparabilité avec d'autres ensembles de données relatives à l'immigration ainsi que des analyses rendues possibles avec la BDIM.

Mots clés : données administratives, immigration, BDIM, données longitudinales, résidents non permanents, déclarants

Glossaire

Ci-après figure la signification des sigles utilisés dans le présent rapport.

Acronyme français	Acronyme anglais	Définition
ARC	CRA	Agence du revenu du Canada
BDIM	IMDB	Base de données longitudinales sur l'immigration
DAL	LAD	Banque de données administratives longitudinales
ECDS	SDLE	Environnement de couplage de données sociales
ELIC	LSIC	Enquête longitudinale auprès des immigrants du Canada
ENM	NHS	Enquête nationale auprès des ménages
FFT1	T1FF	Fichier sur la famille T1
FIR	ILF	Fichier des immigrants reçus
FRPN	PNRF	Fichier intégré des résidents permanents et non permanents
FRN	NRF	Fichier des résidents non permanents
FRN, permis	NRF, Permits	Fichier des résidents non permanents, permis
IPC	CPI	Indice des prix à la consommation
IRCC	IRCC	Immigration, Réfugiés et Citoyenneté Canada
RNP	NPR	Résident non permanent
RP	PR	Résident permanent

Introduction

La base de données longitudinales sur l'immigration (BDIM) est une source importante et complète de données contribuant à mieux comprendre le comportement économique des immigrants. Il s'agit de la seule source annuelle de données canadiennes permettant d'étudier les caractéristiques des immigrants au Canada au moment de leur admission ainsi que leur situation économique et leur mobilité régionale (interprovinciale) sur plus de 35 ans.

La BDIM combine des fichiers de données administratives sur les admissions d'immigrants et sur les permis de résidence non permanente provenant d'Immigration, Réfugiés et Citoyenneté Canada¹ (IRCC) à des fichiers de données fiscales provenant de l'Agence du revenu du Canada (ARC). Les dossiers administratifs de l'IRCC contiennent des renseignements exhaustifs sur les immigrants qui ont été admis au Canada depuis 1952. Les fichiers fiscaux de 1982 et des années suivantes sont disponibles pour les déclarants.

La BDIM vise à fournir des données détaillées et fiables sur l'efficacité et les répercussions des programmes d'immigration. En tant que base de données sur les immigrants (les déclarants de même que les non déclarants), la BDIM peut être utilisée pour répondre aussi bien à de vastes questions de recherche qu'à des questions plus précises. Cette base de données fournit également des renseignements sur l'expérience avant l'admission, comme les permis de travail ou d'études. Sa principale force est qu'elle permet d'analyser la situation socioéconomique sur une période suffisante pour pouvoir évaluer les répercussions des caractéristiques de l'immigration à l'admission (y compris la catégorie d'admission, le niveau de scolarité et les connaissances en français ou en anglais) sur les résultats. De plus, les renseignements annuels sur le lieu de résidence permettent d'étudier la migration secondaire (les déménagements ultérieurs des immigrants au Canada). Il est à noter que les mises à jour annuelles de la BDIM sont indépendantes les unes des autres. D'année en année, il y a des changements aux processus de création des données, incluant à l'Identificateur personnel unique (IMDB_ID).

Depuis son origine, la BDIM est constituée de plusieurs fichiers : un pour chaque année d'imposition depuis 1982 (e.g. IMDB_T1FF_2019), deux contenant les caractéristiques d'immigration au niveau personnel (e.g. PNRF_1952_1979 ET PNRF_1980_2020), un fichier sommaire sur les personnes ayant détenu au moins un permis de résidence non permanente (NRF_PERMIT_1980_2020) et un fichier sommaire de l'information au niveau de personne pour les résidents non permanents (NRF_PERSON_1980_2020)). La BDIM est mise à jour chaque année au moyen de techniques de couplage décrites dans le présent rapport. Chaque année, une nouvelle année d'imposition est ajoutée à la base de données ainsi que des données relatives aux nouvelles admissions et aux permis de résidence non permanente.

Récemment, le contenu de la BDIM a été élargi. La BDIM comprend maintenant les données sur l'acquisition de la citoyenneté depuis 2005, ainsi que les données d'entrée express. De nouveaux modules sur les salaires, les enfants et le processus d'établissement des immigrants ont été intégrés dans la BDIM. De plus, il y a eu un changement dans la structure des fichiers - les déclarants et les non déclarants sont rassemblés dans un seul fichier appelé PNRF_1980_2020. Les non déclarants peuvent être filtrés en utilisant la variable FIRST_TAX_YEAR.

Les fichiers de la BDIM sont uniquement accessibles aux membres des Centres de recherche de données (CRD), aux chercheurs de Statistique Canada et aux personnes réputées y être employées. Cette précaution permet que les mesures adéquates soient prises afin de protéger les renseignements personnels et la confidentialité. Les informations liées à la BDIM sont à la disposition du public sous forme de tableaux sommaires agrégés annuels produit par Statistique Canada et qui sont publiés sur son site Web. De plus, des chercheurs externes peuvent demander des tableaux et analyses ponctuels; Statistique Canada fournit ces services selon une formule de recouvrement des coûts.

Le présent rapport vise à discuter des sources de données de la BDIM (section 2), de ses concepts et variables (section 3), du couplage d'enregistrements (section 4), du traitement des données (section 5), de sa diffusion (section 6), de l'évaluation des données et des indicateurs de qualité (section 7), de sa comparabilité avec d'autres ensembles de données relatifs à l'immigration (section 8) ainsi que des analyses que la BDIM permet (section 9).

1. Auparavant Citoyenneté et Immigration Canada (CIC).

2 Sources des données

Plusieurs fichiers sont nécessaires pour créer la BDIM, incluant le FRPN et le FRN fichier de personne. Ces fichiers, décrits dans la présente section, comprennent des données d'immigration, des données fiscales des immigrants et des fichiers auxiliaires contenant l'information disponible pour des immigrants admis depuis 1952 et des résidents non permanents depuis 1980.

2.1 Données sur l'immigration

Chaque année, Statistique Canada (StatCan) reçoit d'IRCC les données d'admission relatives aux nouveaux titulaires de permis de résidence permanente et non permanente.

2.1.1 Fichier intégré des résidents permanents et non permanents (FRPN 1980-2020)

Chaque année les données d'admission sont ajoutées pour créer le Fichier des immigrants reçus (FIR). Ce fichier contient des renseignements tels que la date d'admission, la date de naissance et la catégorie d'immigration. Le FIR peut être considéré comme le recensement des personnes ayant immigré au Canada à titre de résidents permanents depuis 1980; il présente des renseignements sur leurs caractéristiques lors de leur admission. Ce fichier n'est cependant pas à la disposition des utilisateurs de la BDIM. Les données d'admission de ces immigrants déclarants sont disponibles au sein du Fichier intégré des résidents permanents et non permanents (FRPN). Ce fichier contient également des renseignements sur les non-déclarants pour l'identification : cependant, la première année d'imposition et la dernière année d'imposition seront absentes.

Le FIR, qui consigne les données administratives relatives à la résidence permanente, surestime le nombre d'immigrants vivant actuellement au Canada. Cet écart s'explique de deux façons. Tout d'abord, le FIR ne recense pas les personnes ayant quitté le pays. Des immigrants peuvent être en effet repartis du Canada après leur admission. Ensuite, les décès des immigrants admis à partir de 1980 ne sont que partiellement consignés. Des renseignements supplémentaires sur les données relatives à la mortalité sont présentés à la section 7.2.

Les chercheurs peuvent accéder au Fichier intégré des résidents permanents et non permanents (FRPN, l'acronyme anglais est PNRF), qui combine des données du FIR et du FRN au niveau des personnes. Le FRPN permet aux utilisateurs de suivre l'historique de la migration des immigrants, notamment leur expérience avant leur admission au Canada. Le FRPN contient des données détaillées sur les caractéristiques sociodémographiques des immigrants admis au Canada depuis 1980; cela permet, par exemple, de déterminer si une personne était un résident non permanent avant son admission. Il est important de noter que les enregistrements des résidents non permanents n'étant pas devenus des résidents permanents ne sont pas inclus dans ce fichier. Ce dernier contient en revanche le nombre de permis pour chaque résident non permanent devenu résident permanent ainsi que sa date d'admission. Le FRPN peut également comprendre une date de décès, lorsqu'un lien vers l'enregistrement de décès a été établi (voir la section 7.2.2). Pour de plus amples détails sur le contenu de ce fichier, reportez-vous à la composante relative à l'immigration du dictionnaire de la BDIM, sections 3.3 et 3.4 de ce rapport.

En outre, un fichier appelé PNRF_EXTRA_1980_2013 est à la disposition des utilisateurs de données ; il comprend des variables qui ont été retirées, ayant peu de valeur analytique ou pour lesquelles aucune métadonnée n'est disponible. La liste complète des variables se trouve dans le dictionnaire de données sur l'immigration de la BDIM.

Par le passé, le PNRF séparait les déclarants des non déclarants (PNRF_2016 et PNRF_NONFILERS_2016). Depuis la publication de la BDIM de 2020, les déclarants et les non-déclarants ont été fusionnés et le fichier s'appelle PNRF_1980_2020.

2.1.2 Admissions avant 1980 : Fichier intégré des résidents permanents et non permanents (FRPN) 1952-1979

Avant la diffusion de la BDIM 2018, les données sur les immigrants dans la BDIM étaient limitées aux admissions à partir de 1980. Dans le cadre de la diffusion de la publication 2018 de la BDIM, le PNRF inclura désormais les données de 1952, élargissant l'univers de la BDIM. Le nouveau dossier (PNRF_1952_1979) contient les admissions d'immigrants de 1952-1979. Cependant, le PNRF 1952-1979 comporte moins de variables que pour les personnes admises après 1980 (voir section 2.1.1), puisqu'il s'agit de données plus anciennes. Ses principales catégories sont les suivantes : Sexe, pays, année de naissance, année de débarquement et mois.

2.1.3 Fichier des résidents non permanents (FRN)

Le fichier des résidents non permanents (NRF_Permit) est créé à partir des données des personnes ayant obtenu des permis depuis 1980. Ce fichier comprend le type de permis (de travail ou d'études, par exemple) et la dernière date de validité d'un permis. Le fichier est mis à jour en y ajoutant les données annuelles relatives aux nouveaux permis de résidence non permanente. Pour la version 2020 de la BDIM, le fichier de permis FRN s'appelle NRF_PERMIT_1980_2020.

Les résidents peuvent posséder plusieurs permis au fil du temps. Ces permis comprennent les permis de travail, les permis d'études, les demandes d'asile et les autres permis délivrés, ainsi que la date à laquelle ils ont été délivrés et la date de leur expiration. Le FRN au niveau personne, NRF_PERSON_1980_2020, conserve une ligne d'information par personne ; ce fichier comprend entre autres la somme du nombre de permis et la première année du permis de séjour temporaire.

Ces données peuvent être liées au FRPN à l'aide de l'identificateur personnel unique (IMDB_ID). Pour les variables communes entre le FRPN (PNRF) et le FRN(NRF), en cas de différences utiliser les valeurs du FRPN. Pour de plus amples détails sur les variables incluses dans ces fichiers, veuillez consulter la composante relative à l'immigration du dictionnaire de la BDIM.

2.1.4 Nouveau: Entrée express (EE)

La Base de données longitudinales sur l'immigration (BDIM) comprend des données sur les immigrants admis à travers l'Entrée express² (EE), un système de gestion des applications. L'Entrée express (une extension du FRPN) est une procédure de sélection pour les immigrants économiques demandant un statut de résidence permanente. Ce processus de sélection a débuté le 1^{er} Janvier 2015 et le premier tirage (pour sélectionner les résidents permanents qualifiés) a eu lieu le 31 janvier 2015.

La BDIM comprend des données de 200 300 individus (les demandeurs principaux et leur famille) admis à travers EE. Ces individus peuvent être identifiés en utilisant l'indicateur EXPRESS_ENTRY_IND provenant du fichier intégré des résidents permanents et non permanents appelé PNRF_1980_2020.

Des données plus détaillées sur les demandeurs principaux admis dans le cadre de l'EE sont disponibles. Par exemple, la transférabilité des compétences et le niveau de scolarité le plus élevé.

Si vous souhaitez obtenir de l'information supplémentaire, veuillez consulter le rapport technique dédié à l'entrée express.

2.1.5 Fichier d'admission du Québec

Fourni par la province du Québec, le dossier contient des informations détaillées concernant le programme de sélection de la province pour les années d'admissions 2012 à 2020. Il comprend des variables telles que la catégorie de sélection du Québec, le code de famille du Québec, le domaine d'études du Québec et d'autres variables détaillées dans le dictionnaire d'admission de la BDIM. Le fichier contient 403 446 enregistrements. La province de Québec, en accord spécial avec le gouvernement du Canada, est responsable de ses niveaux d'immigration, programmes et politique d'immigration.

2. Pour plus de détails, veuillez référer au [entrée express](#).

Si vous souhaitez obtenir de l'information supplémentaire, veuillez consulter le rapport technique dédié au fichier d'admission du Québec

2.2 Les modules additionnels de la BDIM

La BDIM comprend des modules externes sur les enfants, les salaires et les services d'établissement.

La BDIM sera diffusé en plusieurs étapes, et les modules externes pourraient ne pas avoir été mis à jour au moment de la diffusion. Les modules externes diffusés précédemment peuvent être utilisés, comme l'identificateur unique de personne (IMDB_ID) de la BDIM n'a pas changé depuis le cycle précédent. Les sections suivantes seront mises à jour, lorsque les modules seront diffusés.

2.2.1 Module de données sur les enfants

Voici une brève introduction au module sur les enfants de la Base de données longitudinales sur l'immigration (BDIM), qui comprend un fichier appelé PNRF_CHILD_1980_2019 avec les fichiers d'immigration des enfants et les Fichiers T1 de la famille (T1FF) depuis 1982, appelé IMDB_CHILD_T1FF, pour les enfants immigrants durant leur enfance. Dans ces fichiers fiscaux, les parents d'enfants sont identifiés par IMDB_ID_PARENT, ce qui équivaut à l'IMDB_ID du parent si les parents sont présents dans les fichiers d'immigration (par exemple, s'ils avaient un ou des permis permanents ou non permanents).

Depuis 1980, plus de 2 millions d'immigrants admis au Canada étaient âgés de moins de 18 ans au moment de leur admission. Cela représente 25 % des immigrants admis pendant cette période. Ces enfants recevront probablement la totalité ou une partie de leur éducation au Canada et auront des défis différents de ceux des immigrants adultes. La Base de données longitudinales sur l'immigration (BDIM) contient peu de renseignements sur les enfants d'immigrants durant leur enfance, car ils ne sont probablement pas des déclarants.

Afin d'accroître la capacité d'analyse de la BDIM, un module pour enfants a été produit. La capacité d'étudier l'impact de la condition socioéconomique des enfants sur les résultats économiques à l'âge adulte est une valeur ajoutée pour la BDIM.

Différentes méthodes ont été utilisées afin d'ajouter des renseignements fiscaux pour les enfants immigrants. Une méthode consistait à utiliser le numéro de la demande d'immigration pour identifier un parent. La deuxième méthode a utilisé le Registre des personnes à charge de Statistique Canada, qui est le résultat des couplages d'enregistrements, pour identifier les tuteurs des enfants. Une fois le lien parent-enfant établi, les fichiers fiscaux durant les années d'enfance sont créés pour toutes les années depuis 1980.

Afin de déterminer le lien parent-enfant, l'information provenant d'un lien DIN-SIN (Numéro d'identificateur dépendant – Numéro d'assurance sociale) a été priorisée. Lorsque cette information n'était pas disponible, le numéro de demande d'immigration a été utilisé pour identifier les parents des enfants. Une fois le lien parent-enfant est établi, des dossiers fiscaux sont créés pour les années de l'enfance pour les enfants immigrants admis depuis 1980.

Les fichiers fiscaux relatifs aux parents de l'enfant comprennent un sous-ensemble des variables incluses dans les IMDB_T1FFs. Seules les principales variables de revenu (comme le revenu d'emploi), les avantages fiscaux et les déductions accordés aux familles et aux parents ont été conservées (comme les prestations fiscales pour enfants et le montant relatif aux études et les frais de scolarité transférés d'un enfant). Le fichier PNRF_CHILD_1980_2019 comprend un sous-ensemble de variables disponibles dans le FRPN et des informations sur le(s) parent(s) de l'enfant, telles que l'IMDB_ID et la première et la dernière année du dépôt pendant l'enfance de l'enfant.

Pour en savoir plus, veuillez consulter le rapport technique sur le module sur les enfants de la BDIM.

2.2.2 Module de données sur les salaires et traitements

Voici un résumé du couplage entre la Base de données longitudinales sur l'immigration (BDIM) et le fichier supplémentaire de l'État de rémunération payée (T4). Les fichiers préliminaires d'impôt sur les salaires et traitements sont tirés des fichiers d'impôt supplémentaires T4, qui contiennent les renseignements fiscaux sur l'emploi fournis par l'employeur du particulier. Les fichiers supplémentaires T4 servent à déclarer les traitements, salaires et avantages imposables versés aux employés pour les services rendus au cours de l'année, ainsi que le

facteur d'équivalence, les montants de rémunération des employés qui ont accumulé des prestations pour l'année en vertu d'un régime de pension agréé ou d'un régime de participation différée aux bénéficiaires. Les variables extraites de ces fichiers comprennent la province d'emploi, la province de l'employé, la rémunération T4 par année d'imposition et le nombre de feuillets T4 par année d'imposition. Les sur les salaires et traitements sont disponibles de 1997 à 2020.

Trois raisons principales justifient l'intégration des fichiers d'impôt T4 à la BDIM :

1. Pour mieux comprendre la couverture réelle de la BDIM en ce qui concerne les résidents temporaires travaillant au Canada en utilisant les NAS temporaires comme base d'analyse.
2. Pour avoir une couverture et une compréhension plus complètes des résidents temporaires qui travaillent au Canada en utilisant les feuillets T4 plutôt que de se fier aux déclarants T1, en particulier les résidents temporaires qui ne font pas la transition vers la résidence permanente.
3. Pour comprendre la faisabilité de diffuser les résultats de la BDIM plus tôt au moyen du fichier d'impôt T4, puisque les fichiers T4 sont disponibles environ six mois avant les fichiers T1FF.

L'intégration du T4 procure certains avantages au T4, particulièrement en ce qui concerne la couverture supplémentaire des travailleurs étrangers temporaires. Les valeurs fournies par le couplage au T4 ont été validées par rapport aux valeurs T1FF, correspondant à 93,0 % du temps, tout en constatant une différence globale moyenne de 1,8 % dans les gains T4.

Pour en savoir plus, veuillez consulter le rapport technique sur le module des salaires et traitements de la BDIM

2.2.3 Le module sur les services d'établissement

Le fichier intégré des résidents permanents et non permanents (FRPN) de la BDIM et les autres fichiers de la BDIM peuvent être intégrés au module des services d'établissement. Le numéro d'identification de personne non confidentielle (IMDB_ID) est inclus dans tous les fichiers, il doit être utilisé pour intégrer les immigrants et les résidents non permanents à leurs dossiers dans d'autres fichiers de la BDIM. Les dossiers DOM_CLIENT_SETTLEMENT (bénéficiaires de services nationaux) et FRN_CLIENT_SETTLEMENT (bénéficiaires de services étrangers) contiennent des informations sur le type et le nombre de services reçus au niveau de la personne. Ensuite, une série de fichiers, par type de service, avec plus de détails sur les services reçus sont mis à la disposition des utilisateurs. Au total, ce module comprend 15 fichiers. Un dictionnaire est à la disposition des utilisateurs pour plus de détails (en anglais seulement, informations en français disponibles sur demande).

Plusieurs dossiers reliés aux services d'établissement offerts aux résidents³ et non permanents sélectionnés pour devenir résidents permanents sont disponibles à Immigration, réfugiés et citoyenneté Canada (IRCC). L'Immigration – Environnement de Déclarations d'Ententes de Contribution (iEDEC) est l'endroit où toutes les données sur les règlements sont recueillies et stockées. iEDEC est un système de déclaration utilisé par les organismes qui fournissent des services d'établissement aux immigrants pour signaler leurs activités. Annuellement, Statistique Canada a reçu des fichiers générés par iEDEC, afin de produire un module de services d'établissement de la BDIM. Les données reçues couvraient les services fournis à partir de 2013. Selon les données reçues, 1 450 180 personnes ont reçu des services au Canada depuis 2013 et 105 020 personnes ont reçu des services avant leur arrivée depuis 2015.

Les services d'établissement sont reçus au Canada ou avant l'arrivée. Divers services sont offerts aux nouveaux immigrants et aux résidents non permanents, certains sont liés à l'emploi ou à l'évaluation des besoins et d'autres à l'information et à l'orientation. Des services de soutien, comme le transport et la garde d'enfants, sont également fournis.

Plusieurs composantes des données de l'ICARE ajoutent une puissance analytique à la BDIM qui combine les admissions d'immigrants et les permis de séjour non permanent avec leurs dossiers fiscaux. Les données actuellement disponibles ne permettent pas d'ajouter des données sur les services d'établissement reçus avant 2013 (et 2015 pour les services étrangers).

3. Pour connaître les critères d'admissibilité, veuillez consulter l'annexe 6.2.

La couverture pour les immigrants admis avant 2013 est partielle. Dans les cas d'admissions multiples, qui sont rares, les services de règlement se rapportent à la plus récente admission lorsque les caractéristiques d'admission (conservées dans la BDIM) se rapportent à la première admission. Les services d'établissement ne se limitent pas aux immigrants récents. Par exemple, 315 immigrants admis pour la première fois en 1980 avaient reçu des services d'établissement entre 2013 et octobre 2020, c'était également le cas pour 1 050 immigrants admis en 1990 et 4 245 immigrants admis en 2000. Les données de règlement non liées à une admission récente n'ont pas été supprimées de la BDIM.

Il est à noter que le module comprend des données sur 87 525 résidents non permanents qui ont reçu des services d'établissement, il peut s'agir de personnes qui ont été admises en 2020 (la BDIM comprend les admissions jusqu'à 2019) ou qui sont en voie de devenir résidents permanents. Les données provenant d'organismes situés au Québec ne sont pas recueillies, de sorte que seuls les services offerts aux immigrants hors Québec sont disponibles.

Le module de données de d'établissement se compose de plusieurs fichiers, car différents types de services se trouvent dans chacun. En outre, une distinction est faite entre les services reçus avant l'arrivée (étrangers) et après l'arrivée (nationaux). Les données sur ces services sont disponibles sur différents fichiers. Il n'a été possible d'intégrer que 18,5 % bénéficiaires de services étrangers à un enregistrement BDIM. La raison principale est que ces personnes ne sont pas arrivées au Canada avant 2019 ou ne sont toujours pas au Canada. La couverture des services étrangers commence en 2015, alors que celle des services nationaux commence en 2013.

Afin d'être incluses dans la BDIM, certaines de ces données ont été synthétisées au niveau des personnes. De nombreuses variables ont été dérivées, comme le nombre de services reçus par sujet. Comme pour `client_ID`, un numéro de service non confidentiel (`SERVIC_NUM`) a été créé.

Le fichier intégré des résidents permanents et non permanents (`FRPN`) de la Base de données longitudinales sur l'immigration (BDIM) et les autres fichiers de la BDIM peuvent être intégrés aux fichiers `DOM_CLIENT_SETTLEMENT` (une personne qui a reçu des services domestiques) et `FRN_CLIENT_SETTLEMENT` (une personne qui a reçu des services étrangers). Les fichiers `DOM_CLIENT_SETTLEMENT` et `FRN_CLIENT_SETTLEMENT` contiennent des renseignements sur le type et le nombre de services reçus au niveau de la personne. Ensuite, une série de fichiers, par type de service, avec plus de détails sur les services reçus sont mis à la disposition des utilisateurs. Au total, ce module comprend 15 fichiers. Le numéro d'identification de personne non confidentielle (`IMDB_ID`) est inclus dans tous les dossiers ; il devrait être utilisé pour intégrer les immigrants et les résidents non permanents à leurs dossiers dans d'autres fichiers BDIM.

Pour en savoir plus, veuillez consulter le rapport technique sur le module des salaires et traitements de la BDIM

2.3 Fichier des familles T1

Chaque année, Statistique Canada utilise le fichier annuel individuel T1, le fichier fiscal T4, et le fichier de l'allocation canadienne pour enfants de l'ARC (l'Agence du revenu du Canada) et crée un FFT1 analytique. Les données du FFT1 est disponible dès 1982 jusqu'à 2019.

Les fichiers de données fiscales utilisés pour créer les fichiers **IMDB_T1FF** sont ceux qui contiennent le fichier sur la famille T1⁴ (FFT1). Statistique Canada obtient les fichiers T1 individuels annuels, les fichiers fiscaux T4 et les fichiers de la Prestation fiscale canadienne pour enfants (PFCE)⁵ de l'ARC et crée le Fichier sur la famille T1 pour l'année en question. Le traitement comprend de nombreuses étapes, allant du codage géographique à la formation de familles (p. ex. lorsqu'un déclarant mentionne une conjointe également déclarante, cette dernière est associée au déclarant initial par un identificateur commun). Les données du FFT1 remontent jusqu'à l'année d'imposition de 1982. Grâce à l'expérience acquise au cours des nombreuses années de traitement du FFT1, des règles de vérification ont été créées pour réduire le nombre d'incohérences au sein de la base de données et améliorer la qualité des données.

4. Pour obtenir des renseignements sur le plus récent FFT1, consultez le [site Web](#) de Statistique Canada.

5. Le 1^{er} juillet 2016, la Prestation fiscale canadienne pour enfants a été remplacée par l'Allocation canadienne pour enfants.

La disponibilité des variables fiscales dépend des renseignements recueillis pour une année donnée. Le FFT1 produit chaque année pour la BDIM comprend les revenus individuels et familiaux ainsi que les variables de composition de la famille, telles que le nombre d'enfants et l'identificateur de conjoint. La BDIM contient les fichiers IMDB_T1FF à compter de 1982 pour les immigrants déclarants. Le processus de création de ces fichiers est décrit à la section 5.1. Pour de plus amples détails sur les variables incluses dans les fichiers IMDB_T1FF, veuillez-vous reporter à la composante relative aux données fiscales du dictionnaire de la BDIM.

2.4 Fichiers auxiliaires

Pour créer la BDIM, il est nécessaire d'utiliser des fichiers auxiliaires qui facilitent le couplage d'enregistrements et ajoutent des variables à la base de données. Ces fichiers auxiliaires ne sont pas disponibles pour les utilisateurs de la BDIM.

Le ECDS a été utilisé pour faciliter les couplages d'enregistrements.

L'Environnement de couplage de données sociales (ECDS) de Statistique Canada encourage l'utilisation novatrice des données administratives et des données d'enquête existantes pour répondre à des questions de recherche importantes et éclairer les politiques socio-économiques par le couplage d'enregistrements. L'ECDS élargit le potentiel d'intégration des données à de multiples domaines, comme la santé, la justice, l'éducation et le revenu, grâce à la création de fichiers de données analytiques intégrés, sans qu'il soit nécessaire de recueillir des données supplémentaires auprès des Canadiens.

3 Concepts et variables

3.1 Les statuts de l'immigrant au Canada

La BDIM fournit des données sur un sous-ensemble de la population des immigrants, tel qu'il est décrit à la section 2. Trouvez ci-après les définitions des termes « immigrants » et « résidents non permanents » selon Statistique Canada.

Le terme « **immigrant** » désigne des personnes qui sont, ou ont déjà été, des immigrants reçus ou des résidents permanents; c'est-à-dire auxquelles les autorités de l'immigration ont accordé le droit de résider au Canada en permanence. Les immigrants ayant obtenu la citoyenneté canadienne par naturalisation sont également inclus dans cette catégorie.

Les « **résidents non permanents** » ne sont pas considérés comme des immigrants, même s'il s'agit d'une population sur laquelle porte également la BDIM, comme le décrit la section 2. Dans la BDIM, le terme « **résident non permanent** » désigne une personne originaire d'un autre pays, titulaire d'un permis de travail ou d'études ou qui revendique le statut de réfugié. Ces personnes sont autorisées à demeurer au Canada pour la période indiquée sur leur permis.

3.1.1 Un aperçu de l'immigration au Canada

Un article de *Mégatendances canadienne*, *150 ans d'immigration au Canada*⁶, paru en 2016, résume la fluctuation des niveaux d'immigration et des pays d'origine au cours du dernier siècle. La migration au Canada a été continue depuis la fondation du pays. Plus de 17 millions d'immigrants se sont installés au Canada depuis 1867. Le nombre d'immigrants reçus a augmenté, passant d'environ 200 000 dans les années 1990 à plus de 250 000 au début des années 2010. La proportion des Canadiens nés à l'étranger a augmenté, passant de 14,7 % en 1951 (2,06 millions de personnes) à 21,9 % en 2016 (7,5 millions de personnes).

En vertu de l'article 95 de la *Loi constitutionnelle de 1867*, l'immigration est un domaine de pouvoir partagé entre le gouvernement fédéral et les gouvernements provinciaux. D'autres lignes directrices, définies dans la *Loi sur l'immigration et la protection des réfugiés* de 2002 (LIPR)⁷, précisent les objectifs et la direction stratégique de la politique d'immigration du gouvernement du Canada, que régit en partie Immigration, Réfugiés et Citoyenneté Canada (IRCC). Avant 2002, la Loi sur l'immigration de 1976 servait de principale loi en matière d'immigration au Canada.

Conformément à la LIPR, le gouvernement du Canada est chargé « de stipuler les exigences en matière d'admissibilité, de fixer les niveaux nationaux d'immigration, de définir les catégories d'immigration, de prendre des décisions concernant les demandes d'asile présentées au Canada, de réunifier les familles et d'établir les critères d'admissibilité pour les programmes d'établissement⁸ » pour toutes les provinces et les territoires, à l'exception du Québec. La province du Québec est entièrement responsable de ses niveaux, programmes et politiques d'immigration en vertu de l'*Accord Canada-Québec relatif à l'immigration et à l'admission temporaire des aubains*. Le gouvernement fédéral continue, cependant, à sélectionner et à traiter les immigrants que parrainent les familles, les personnes protégées au Canada et les personnes revendiquant le statut de réfugié au Québec.⁷

Les **résidents permanents** sont des personnes ayant été autorisées à vivre au Canada de façon permanente, qui ont le droit d'y travailler et d'y étudier, mais qui ne sont pas devenues des citoyens canadiens⁸. En vertu de la LIPR, il existe trois catégories générales d'immigrants : les immigrants économiques, les immigrants parrainés par la famille et les réfugiés.

6. Voir Statistique Canada 2016 dans la biographie

7. [Loi sur l'immigration et la protection des réfugiés](#) (L.C. 2001, ch. 27).

8. Définitions approuvées comme norme recommandée de Statistique Canada le 21 mars 2016.

Les résidents permanents peuvent devenir citoyens du Canada lorsqu'ils répondent à certaines exigences. La première est une exigence de résidence, selon laquelle le résident permanent doit avoir été présent au Canada pendant une période définie. Les résidents permanents doivent également être âgés de plus de 18 ans; dans le cas de mineurs, la demande doit être effectuée simultanément (concomitante) à celle de l'un ou des deux parents ou après (non concomitante) que l'un des parents ou les deux sont devenus citoyens canadiens. Les résidents permanents doivent avoir honoré leurs obligations de déclaration de revenus auprès du Canada. Les résidents permanents âgés de 18 à 54 ans doivent en outre présenter une preuve de maîtrise d'au moins l'une des langues officielles du Canada et avoir passé l'examen de citoyenneté (gouvernement du Canada, 2021).

La LIPR stipule que tous les ressortissants étrangers, à l'exception des résidents permanents, qui entrent au Canada doivent posséder un visa de résidence temporaire. Les visas de résidence temporaire sont délivrés aux travailleurs et étudiants « de manière à optimiser leur apport au développement économique, social et culturel du Canada tout en protégeant la santé des Canadiens et en assurant leur sécurité » (IRCC, 2015). Les résidents non permanents peuvent demander la résidence permanente dans le cadre de divers programmes et peuvent bénéficier d'un avantage par rapport aux demandeurs de l'étranger, s'ils ont étudié au Canada ou y ont acquis une expérience professionnelle.

Quant aux **personnes qui revendiquent le statut de réfugié**, « la Section de la protection des réfugiés (SPR) est la section de la Commission de l'immigration et du statut de réfugié du Canada (CISR) qui instruit les demandes d'asile présentées au Canada et qui décide de les accepter ou non » (CISR, 2015). Dans le cadre de la BDIM, ces demandeurs sont classés comme résidents non permanents possédant un permis de personne revendiquant le statut de réfugié.

3.2 La population cible et la période de couverture

La BDIM est une base de données qui comprend :

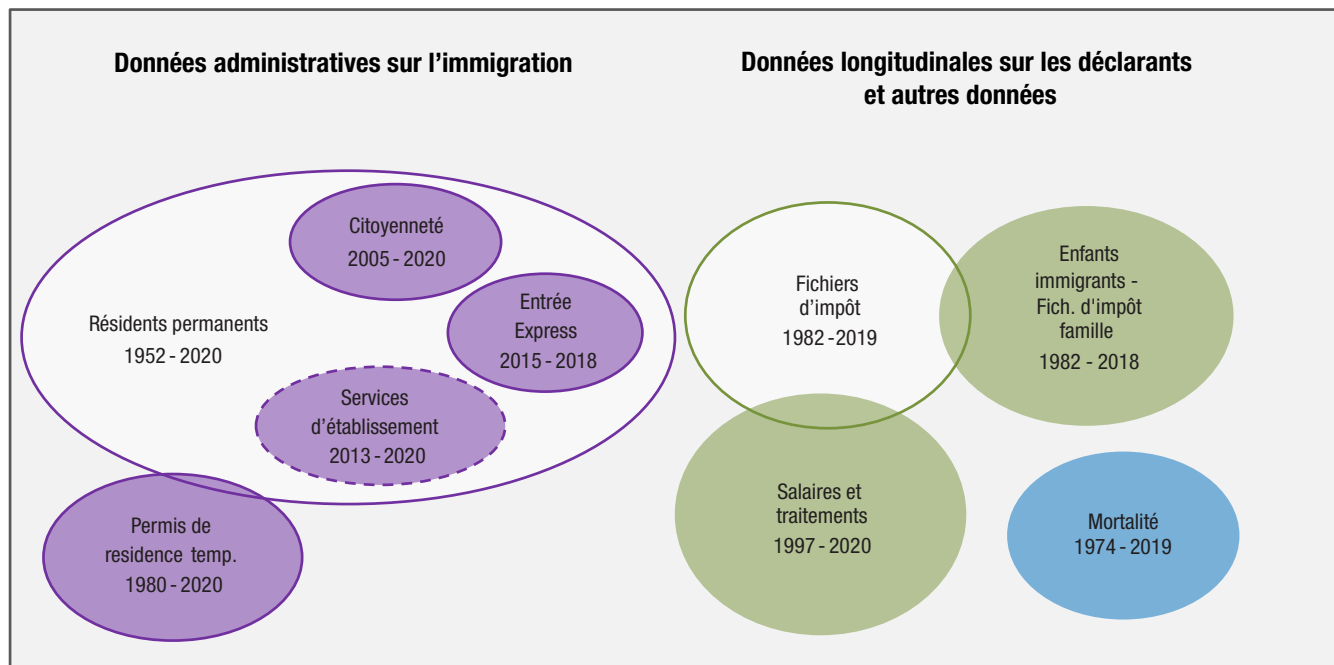
- Tous les immigrants admis au Canada depuis 1952;
- Tous les permis de résidents non permanents depuis 1980;
- Toutes les déclarations fiscales des immigrants depuis 1982;

Cette base de données fournit également des renseignements sur les permis des immigrants qui étaient résidents non permanents avant leur admission à titre de résidents permanents.

Le FFT1 porte sur « toutes les personnes ayant complété une déclaration de revenu T1 pour l'année de référence ou ayant reçu la Prestation fiscale canadienne pour enfants (PFCE), sur les conjoints faisant partie des non déclarants (y compris les renseignements sur les traitements et salaires figurant dans le fichier T4), sur les enfants non déclarants (données d'identification provenant de trois sources : fichier de la PFCE, fichier des naissances et fichier historique) et sur les enfants ayant produit une déclaration et indiqué l'adresse de leurs parents ». L'information familiale est basée sur le concept de la famille de recensement : parent(s) et les enfants vivant à la même adresse (Statistique Canada 2019).

La BDIM rassemble, par couplage d'enregistrements probabiliste (voir la section 4), les données administratives d'Immigration, Réfugiés et Citoyenneté Canada (IRCC) et les fichiers fiscaux du FFT1.

Figure 1
Source du contenu de la BDIM



Note : Voir le glossaire pour connaître les définitions des acronymes.

Source : Statistique Canada, description du déroulement du processus de création de la BDIM.

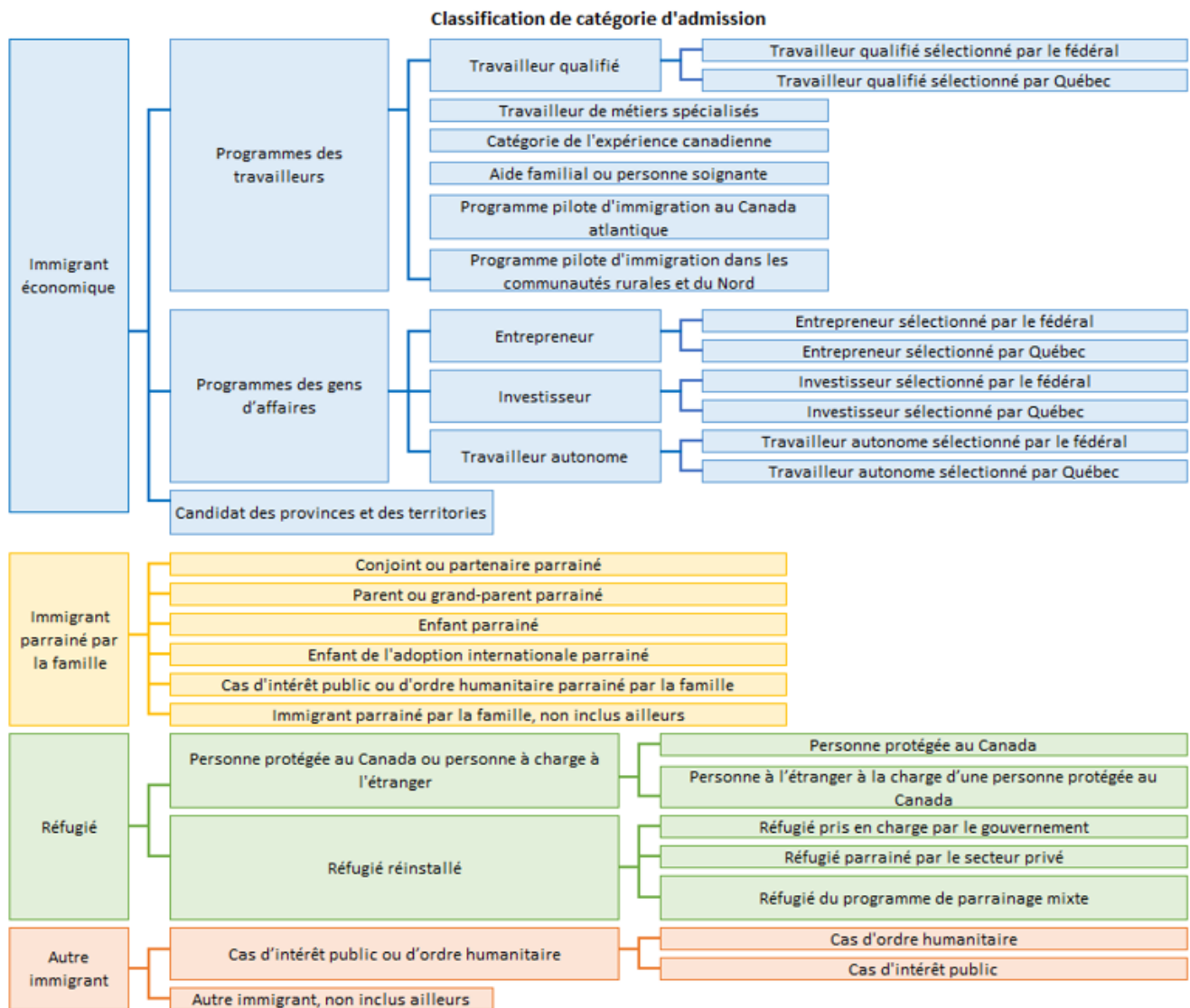
3.3 Variables relatives à l'admission

Les immigrants sont admis au Canada dans le cadre de différents programmes ayant tous des objectifs précis. Ces programmes sont des indicateurs des conditions sous lesquelles les immigrants sont admis au pays et du type de soutien reçu lors de leur établissement. Par conséquent, les analyses visant à orienter les politiques publiques devraient normalement en tenir compte. Afin de répondre à des questions de recherche variées, la BDIM comporte plusieurs variables relatives à l'admission des immigrants, toutes dérivées de deux concepts principaux, soit la catégorie d'admission et le type de demandeur.

3.3.1 Catégorie d'admission

La **catégorie d'admission** désigne le nom du programme ou du groupe de programmes d'immigration en vertu duquel un immigrant a obtenu pour la première fois depuis 1980 le droit de vivre au Canada en permanence par les autorités de l'immigration. Au fil des ans, les immigrants ont été admis au pays dans le cadre de plusieurs dizaines de programmes distincts. Afin de faciliter l'utilisation des données sur le sujet, la BDIM met à la disposition de ses utilisateurs différentes variables à l'intérieur desquelles les programmes ayant des buts similaires sont agrégés. Le plus haut niveau d'agrégation repose sur les trois objectifs principaux de la politique d'immigration du Canada, soit contribuer au développement économique du pays, réunir les familles et protéger les réfugiés.

Figure 2
Classification de catégorie d'admission



Note : Voir le glossaire pour connaître les définitions des acronymes.

Source : Statistique Canada.

3.3.1.1 Immigrant économique

L'admission d'immigrants économiques vise à atteindre le premier des objectifs susmentionnés de la politique d'immigration, c'est-à-dire celui de contribuer à l'économie canadienne. Les immigrants économiques sont répartis dans trois groupes de programmes principaux, soit les programmes des travailleurs, les programmes des gens d'affaires et les programmes de candidats des provinces et territoires.

Immigrant économique

Les programmes de travailleurs comprennent les immigrants qui ont été sélectionnés pour leur capacité à prendre part au marché du travail.

Ceux-ci sont évalués en fonction de leurs compétences et de leur expérience professionnelle et sont répartis en quatre grandes catégories :

1. Les travailleurs qualifiés sont sélectionnés pour leurs compétences et leur expérience de travail dans des postes de gestion, des postes professionnels, des emplois techniques ou des métiers spécialisés.
2. Les travailleurs exerçant un métier spécialisé sont sélectionnés plus particulièrement pour leurs compétences et leur expérience de travail dans le cadre d'un métier spécialisé admissible. Ils se distinguent des travailleurs qualifiés par l'exigence de détenir une offre d'emploi valide d'un employeur canadien ou un certificat de compétence délivré par un organisme provincial ou territorial canadien.
3. Les immigrants admis dans la catégorie de l'expérience canadienne se distinguent des deux premiers groupes par l'exigence de posséder de l'expérience de travail canadienne. Celle-ci doit avoir été acquise dans des postes de gestion, des postes professionnels, des emplois techniques ou des métiers spécialisés.
4. Les aides familiaux et les personnes soignantes obtiennent leur statut de résident permanent après avoir offert, au Canada et pour une période déterminée, des soins à domicile à des enfants ou à des personnes ayant des besoins spéciaux, telles que des personnes âgées ou des personnes souffrant d'un handicap ou d'une maladie chronique.
5. Programme pilote d'immigration au Canada atlantique (Veuillez trouver une courte description du programme ci-dessous)
6. Programme pilote d'immigration dans les communautés rurale et du Nord (Veuillez trouver une courte description du programme ci-dessous)

Programme pilote d'immigration au Canada atlantique

Ce programme aide à trouver des immigrants qualifiés pour immigrer au Canada atlantique afin de pourvoir des postes qui n'ont pas déjà été pourvus localement. Il s'agit d'un programme dirigé par l'employeur, ce qui signifie que les employeurs exigent des travailleurs qualifiés étrangers parce qu'il y a un besoin de main-d'œuvre étrangère qualifiée stable au Canada atlantique. Il s'agit de la voie (porte d'entrée) vers la résidence permanente pour les diplômés étrangers et les travailleurs étrangers qualifiés qui veulent vivre dans l'une des quatre provinces de l'Atlantique du Canada.

Il est également ouvert aux diplômés étrangers qui veulent aller au Canada atlantique après l'obtention de leur diplôme. Le candidat peut vivre à l'étranger ou temporairement au Canada.

Il existe trois programmes dans le cadre de ce programme :

- Programme des diplômés étrangers du Canada atlantique*
- Programme des travailleurs hautement qualifiés du Canada atlantique*
- Programme des travailleurs qualifiés intermédiaires du Canada atlantique*

Pour participer au programme susmentionné, les conditions doivent être remplies, tant au niveau de l'employeur qu'à celui du candidat.

[Programme pilote d'immigration dans les communautés rurales et du Nord](#)

L'objectif du projet pilote d'immigration dans les régions rurales et du Nord est de faire profiter les populations plus petites des avantages de l'immigration économique en attirant des travailleurs étrangers qualifiés qui veulent travailler et prospérer dans des communautés rurales. Contrairement au projet pilote de l'Atlantique, le projet pilote d'immigration dans les communautés rurale et du nord est une approche communautaire. Cela signifie que les communautés dans leur ensemble présentent une demande d'accueil et sont sélectionnées en fonction de trois grands critères :

- S'il y a un besoin économique d'immigration
- Si les ressources et les partenaires communautaires sont présents pour être investis dans le projet.
- Si le gouvernement fédéral dispose déjà de partenaires et de ressources en matière d'établissement dans la communauté

L'objectif du programme d'immigration est d'aider à répondre aux besoins du marché du travail des communautés locales et de soutenir le développement économique de cette région. En accueillant et en soutenant les immigrants, les communautés contribuent à les convaincre de demeurer dans une petite communauté rurale.

Les taux de rétention devraient être plus élevés en fonction:

- du travail avec leurs partenaires communautaires
- de la participation d'autres partenaires du gouvernement fédéral
- de la participation des gouvernements provinciaux et territoriaux

Les exigences doivent être satisfaites, tant à l'extrémité de la communauté qu'à celle du candidat.

Les immigrants économiques admis au Canada dans le cadre des [programmes des gens d'affaires](#) sont quant à eux répartis en trois grandes catégories :

1. Les entrepreneurs sont sélectionnés pour leurs compétences et leur capacité à soit posséder et gérer une entreprise, soit mettre sur pied une entreprise admissible au Canada. Certains possèdent un avoir net minimal alors que d'autres doivent avoir obtenu l'appui d'une organisation désignée pour leur idée d'entreprise.
2. Les investisseurs obtiennent leur statut de résident permanent à la condition d'investir une somme importante au Canada. Leur placement est attribué aux provinces et aux territoires participants afin de stimuler le développement économique et la création d'emplois.
3. Les travailleurs autonomes obtiennent leur statut de résident permanent à la condition d'être en mesure — et d'avoir l'intention — de créer leur propre emploi au Canada et de contribuer de façon appréciable à l'économie canadienne. Cette grande catégorie comprend aussi les personnes qui ont l'intention de contribuer de façon appréciable à la vie culturelle ou sportive du pays (p. ex. en tant qu'artiste, acteur, écrivain ou athlète professionnel).

Le dernier programme principal pour l'admission des immigrants économiques est celui des [programmes de candidats des provinces et territoires](#). Comme l'indique le nom, les immigrants sont sélectionnés par une province ou un territoire pour contribuer à l'économie locale en répondant à des besoins précis en matière de main-d'œuvre. Ils sont évalués en fonction de critères de sélection relatifs à leur scolarité, à leur expérience de travail et à leurs compétences particulières. Chaque province ou territoire participant possède ses propres critères pour sélectionner son groupe d'intérêt, par exemple les étudiants, les gens d'affaires, les travailleurs spécialisés ou les travailleurs semi-spécialisés.

3.3.1.2 Immigrant parrainé par la famille

L'admission d'immigrants parrainés par la famille vise la réunification des familles et permet aux citoyens canadiens et aux résidents permanents de parrainer des personnes qui leur sont apparentées. Les immigrants admis dans le cadre de tels programmes peuvent recevoir leur statut de résident permanent en raison de leur lien soit comme conjoint, partenaire, parent, grand-parent ou enfant.

Sous certaines conditions, un immigrant admis dans le cadre de tels programmes peut aussi être parrainé en raison d'un autre lien de parenté. C'est le cas par exemple des jeunes frères, sœurs, neveux, nièces et petits-enfants orphelins. Un citoyen canadien ou un résident permanent vivant au Canada peut aussi parrainer quelqu'un sur la base d'un autre lien si personne ne lui est apparenté par les liens décrits plus haut.

À cela s'ajoutent aussi les immigrants qui ont été parrainés par un citoyen canadien ou par un résident permanent vivant au Canada et qui n'auraient normalement pu se qualifier dans le cadre d'aucun programme, mais qui ont reçu, à titre exceptionnel, le statut de résident permanent pour des considérations d'ordre humanitaire.

Pour plus d'informations, visitez [Immigrant parrainé par la famille](#)

3.3.1.3 Réfugié

Le troisième et dernier objectif principal de la politique d'immigration du Canada est celui de protéger les réfugiés, c'est-à-dire les personnes qui ont une crainte fondée de retourner dans leur pays d'origine. Cette catégorie comprend les personnes qui craignent avec raison d'être persécutées pour des motifs liés à leur race, à leur religion, à leur nationalité, à leur appartenance à un groupe social particulier ou à leurs opinions politiques (les réfugiés au sens de la Convention de Genève), de même que les personnes qui ont subi des conséquences graves et personnelles en raison d'une guerre civile, d'un conflit armé ou d'une violation massive des droits de la personne. Certains réfugiés étaient au Canada lorsqu'ils ont demandé l'asile pour eux-mêmes et pour les membres de leur famille avec eux au Canada ou à l'extérieur du pays. D'autres étaient à l'étranger et ont été recommandés par le Bureau du Haut Commissariat des Nations Unies pour les réfugiés (HCR) ou par un autre organisme de recommandation aux fins d'établissement au Canada. Les immigrants qui ont été recommandés reçoivent un soutien pour leur rétablissement de la part de sources gouvernementales, d'organisations, d'individus ou de groupes d'individus du secteur privé, ou encore un soutien provenant à la fois du gouvernement du Canada et d'intervenants du secteur privé.

Pour plus d'informations, visitez [Réfugiés](#)

3.3.1.4 Autre immigrant

En plus des trois objectifs principaux susmentionnés, la politique d'immigration du Canada attribue aux autorités de l'immigration un certain pouvoir discrétionnaire pour octroyer le statut de résident permanent dans le cadre d'un programme qui ne vise ni les immigrants économiques, ni les immigrants parrainés par la famille, ni les réfugiés. Ce programme comprend notamment les immigrants qui ont reçu, à titre exceptionnel, le statut de résident permanent pour des considérations d'ordre humanitaire ou des raisons d'intérêt public.

Pour plus d'informations, visitez [Autre immigrant](#)

3.3.2 Type de demandeur

En plus de la catégorie d'admission, la BDIM fournit aux utilisateurs de données des renseignements sur le type de demandeur. Ce concept indique si l'immigrant est identifié comme demandeur principal, comme conjoint ou comme personne à charge dans la demande de visa de résident permanent.

En règle générale, les renseignements sur le type de demandeur sont utilisés aux fins d'analyse dans le cadre de l'étude sur les immigrants économiques. Comme les demandeurs principaux admis en vertu de ces programmes sont sélectionnés pour leur capacité à contribuer à l'économie canadienne, il est pertinent de les distinguer de leurs conjoints et de leurs personnes à charge qui, eux, n'ont pas été évalués en fonction de cette même capacité. En isolant les demandeurs principaux des autres demandeurs, l'efficacité de ces programmes peut alors être étudiée de façon plus directe.

Par contre, en ce qui concerne la réunification des familles et la protection des réfugiés, le but de la politique d'immigration demeure le même pour toutes les personnes qui présentent une même demande, et ce, sans égard au type de demandeur. Dans le cas des immigrants admis en vertu de ces deux objectifs, le concept du « type de demandeur » revêt donc plutôt une valeur administrative.

La valeur administrative du concept est particulièrement marquée pour les immigrants ayant le statut de demandeur principal, lequel ne dépend pas systématiquement du lien juridique entre les demandeurs de résidence permanente. Par exemple, dans le cas de la catégorie d'admission des conjoints et des partenaires parrainés, le conjoint est répertorié comme demandeur principal, alors que le type de demandeur « conjoint » n'apparaît pas sur la demande de résidence. De plus, dans le cas de la catégorie d'admission des enfants parrainés, le statut de demandeur principal sera désigné de façon nominative pour l'un des enfants, alors que les autres enfants seront désignés comme étant les personnes à charge. Enfin, dans certaines circonstances, la demande de résidence permanente se déroule en deux temps : à partir du Canada pour le demandeur principal, puis à partir de l'étranger pour les autres membres de la famille. Ce type de processus existe pour les aides familiaux résidants et pour les personnes protégées au Canada. Dans ces circonstances, l'un des membres de la famille faisant la demande à partir de l'étranger obtiendra le statut de demandeur principal, et ce, même s'il est le conjoint d'un immigrant approuvé antérieurement suite à une demande faite au Canada.

3.3.4 Changements au fil du temps

La BDIM comprend des renseignements sur les immigrants admis au pays depuis plus d'une trentaine d'années, au cours desquelles plusieurs changements relatifs aux politiques et aux programmes d'immigration sont survenus. De nouveaux programmes ont été créés, d'autres ont été abolis, alors que pour d'autres, les critères de sélection et d'admissibilité ont été modifiés. Ainsi, une personne admise en tant que travailleur qualifié en 1980 n'a pas nécessairement été évaluée selon les mêmes critères que ses homologues admis dans les années 2000. Même si tous les efforts ont été déployés pour tenter de créer les agrégations de programmes les plus similaires possible, l'utilisateur de la BDIM devrait garder en tête ces différences lorsqu'il tire des conclusions au sujet de différentes cohortes d'établissement.

Le changement le plus marquant survenu dans la période couverte par la BDIM est sans doute l'entrée en vigueur de la *Loi sur l'immigration et la protection des réfugiés de 2002*, qui a remplacé la *Loi sur l'immigration de 1976*. Même si ces deux lois visent les trois mêmes groupes principaux (immigrants économiques, immigrants parrainés par la famille et réfugiés), l'administration des programmes dans le cadre de celles-ci a subi plusieurs modifications. De plus, les programmes d'admission ont aussi été modifiés en fonction des besoins sociodémographiques et des priorités établies par les gouvernements successifs à l'intérieur même de ces deux cadres législatifs. Pour ces raisons, il est fortement conseillé aux utilisateurs de données s'intéressant à des programmes spécifiques ou à plusieurs cohortes d'établissement de se renseigner sur les changements apportés aux politiques ou aux programmes applicables à leur sujet d'étude.

Il est aussi important de souligner que quelques années peuvent s'écouler avant que les effets d'un changement administratif puissent être observés à l'intérieur de la base de données. Par exemple, lorsqu'un nouveau programme est créé, il peut s'écouler plusieurs mois — voire plusieurs années — entre le moment où celui-ci entre en vigueur (moment où il est possible de poser sa candidature) et le moment où les premiers immigrants sont admis au pays dans le cadre de celui-ci. Un scénario similaire existe pour les programmes abolis. Il peut s'écouler un certain temps entre le moment où toutes les demandes sont évaluées et où toutes les personnes admissibles sont reçues au pays et celui où les programmes abolis disparaissent complètement des statistiques sur les admissions annuelles.

3.3.5 Variables de catégorie d'admission du FRPN

Plusieurs catégories d'admission existent au sein du FRPN. Elles sont décrites dans la composante relative à l'immigration du dictionnaire de la BDIM. Cette section fournit de plus amples détails sur certaines de ces variables.

La variable la plus détaillée est celle des catégories d'immigration **IMMIGRATION_CATEGORY**, qui comprend plus de 100 catégories ayant existé de 1980 jusqu'à la BDIM actuelle. Des agrégats des renseignements disponibles dans la variable **IMMIGRATION_CATEGORY** figurent dans la variable dérivée de groupement de catégories d'immigration, **IMMIGRATION_CATEGORY_CENSUS**, qui comprend un moins grand nombre de catégories.

La variable agrégée **IMMIGRATION_CATEGORY_CENSUS** est une catégorisation conforme au standard de Statistique Canada utilisé pour le recensement de la population. Elle ne précise, cependant, pas que certains immigrants ont été admis dans le cadre des **programmes d'élimination de l'arriéré et d'examen administratif**. Ces programmes visaient à accélérer le traitement des dossiers d'immigration à la fin des années 1980, afin de faire face aux crises géopolitiques se déroulant à l'étranger et touchant la capacité de résidents temporaires à retourner dans leur pays (p. ex. manifestations de la place Tian'anmen, dissolution de l'URSS et de la Yougoslavie). Du fait de l'absence de distinction de ces catégories, ces personnes, dont le dossier a été traité rapidement selon des critères différents, ne peuvent être comparées aux autres immigrants dont les dossiers figurent dans les mêmes catégories. Pour recenser les immigrants admis dans le cadre de ces programmes, les utilisateurs doivent se reporter aux variables d'élimination de l'arriéré **BACKLOG_CLEARANCE_IND** et d'examen administratif **ADMINISTRATIVE_REVIEW_IND** (disponibles dans le fichier **PNRF_extra**, pour les années d'admission antérieures à 2014).

L'utilisateur peut également utiliser les renseignements agrégés sur l'immigration de la variable de groupement de catégories d'immigration **IMMIGRATION_CATEGORY_ROLLUP2** disponible dans le fichier **PNRF_extra** pour les immigrants admis jusqu'à 2013. Cette variable vise à fournir des informations cohérentes indépendamment des divers changements de politique et de réglementation (p. ex. *Loi sur l'immigration et la protection des réfugiés* de 2002 et *Loi sur l'immigration* de 1976) et pour maintenir au cours du temps des programmes particuliers relatifs à l'immigration (p. ex. travailleurs qualifiés). Cette variable offre des renseignements détaillés sur les initiatives d'élimination de l'arriéré et d'examen administratif. Des détails relatifs au groupement pour les variables dérivées figurent dans le dictionnaire des données sur l'immigration de la BDIM.

Un autre point à prendre en compte quant aux variables de catégories d'admission est leur relation avec le **type de demandeur** (variable de statut dans la famille **FAMILY_STATUS** du **FRPN**). En règle générale, les demandeurs principaux sont ceux étant évalués à l'aide des critères d'admission dans chacune des catégories, alors que le conjoint ou la conjointe et les personnes à charge qui les accompagnent sont automatiquement admis avec le demandeur principal (même si les aptitudes linguistiques du conjoint ou de la conjointe peuvent également être évaluées dans les demandes d'immigration de catégorie économique). Au sein des variables de groupement, certaines catégories d'admission précisent explicitement s'il s'agit 1) d'un demandeur principal ou 2) d'un conjoint ou d'une personne à charge, alors que pour d'autres catégories (p. ex. immigrants parrainé par la famille), il est nécessaire de se référer à la variable **FAMILY_STATUS** pour déterminer le statut de la personne en tant que demandeur principal ou conjoint/personnes à charge.

Deux catégories font exception à ce qui précède : la catégorie personnes à charge de type aide familial résidant et la catégorie réfugiés. Lorsqu'elles sont comparées à la catégorie de groupement de statut dans la famille **FAMILY_STATUS_ROLLUP**, ces variables font référence à des demandeurs principaux ainsi qu'à des personnes à charge. Cela peut se produire lorsque le demandeur principal se trouve déjà dans le pays et qu'une personne à sa charge soumet une demande distincte de résidence permanente de l'étranger. Puisque chaque demande distincte doit présenter un demandeur principal, même nominal, l'une des personnes à charge (généralement le conjoint) est considérée comme le « demandeur principal » de la demande des personnes à charge. Il n'existe, cependant, aucune différence de traitement entre les demandeurs principaux, les conjoints et les personnes à charge dans ces deux catégories d'admission.

3.4 Variables d'intérêt

La **BDIM** est une base de données complète fournissant aux chercheurs une myriade de variables permettant d'étudier les résultats relatifs aux caractéristiques d'immigration ainsi que diverses répercussions à long terme. Le nombre de variables au sein des fichiers de données fiscales dépasse 600 (la moitié environ au niveau individuel et l'autre moitié au niveau familial de l'agrégation). Le Fichier intégré des résidents permanents et non permanents (**FRPN**) contient plus de 50 variables. Même si les définitions exactes de ces variables figurent dans la composante relative à l'immigration du dictionnaire de la BDIM, certains des concepts plus nuancés sont

décrits dans le présent rapport. Les sections suivantes portent sur les variables géographiques, temporelles, de scolarité et de profession envisagée, afin d'en présenter plus précisément la signification et l'utilité. De plus amples détails sur les variables de revenu figurent dans la composante relative aux impôts du dictionnaire de la BDIM. De nouvelles variables ont été ajoutées à la BDIM 2016, celles-ci comprennent les variables Mouvements d'établissement des réfugiés syriens (SYRIAN_RRW), Entrée express (EXPRESS_ENTRY_IND) et le mois et l'année de citoyenneté (CITIZEN_YEAR et CITIZEN_MONTH).

Les variables du FRPN (PNRF) font référence aux caractéristiques des immigrants lors de leur admission ou à la réception d'un document de résidence temporaire, alors que les variables des fichiers de données fiscales font référence aux caractéristiques au cours de l'année d'imposition. Alors que certaines variables sont disponibles dans les deux fichiers, les variables relatives aux impôts changent au cours du temps. L'âge, par exemple, figure dans les deux fichiers et devrait changer chaque année dans le fichier de données fiscales. L'état matrimonial des immigrants (MARITAL_STATUS) et la province de destination (DESTINATION_PROVINCE) lors de la demande de résidence permanente peuvent également différer de l'état matrimonial (MSTCO) et de la province de résidence (PRCO) au moment de la déclaration de revenus. Pour les variables qui ne devraient pas changer au cours du temps, il convient d'utiliser le FRPN par souci de cohérence.

3.4.1 Variables géographiques

La BDIM permet d'étudier la mobilité et la rétention des immigrants déclarants au Canada. Il n'est cependant pas possible de capturer complètement les émigrations, car les immigrants ou déclarants ne sont pas tenus de déclarer qu'ils quittent ou comptent quitter le pays. Les fichiers FRPN et de données fiscales présentent diverses mesures du lieu géographique qui permet aux chercheurs d'établir une **destination envisagée** à l'admission et les lieux de résidences suivants des immigrants. Dans le FRPN, la destination envisagée est mesurée à l'échelle provinciale, de la région métropolitaine de recensement, de la division de recensement et de la subdivision de recensement. Ces variables proviennent de la destination auto déclarée à l'admission figurant sur la demande d'immigration. Contrairement aux variables géographiques du FFT1, les variables du fichier d'établissement sont uniquement disponibles pour les géographies définies dans le dernier recensement disponible; cela signifie qu'elles ne reflètent que les limites de recensement les plus récentes.

L'autre renseignement géographique figurant sur le fichier d'établissement est **la province de nomination**, disponible pour les candidats des provinces. La province indiquée est celle selon les critères de laquelle le demandeur a été admis; elle ne correspond cependant pas nécessairement à la variable de province de destination.

Sous **régions géographiques d'origine**, les variables de pays du fichier d'établissement indiquent le pays de naissance de la personne, son pays de citoyenneté et sa résidence la plus récente au moment de son admission. Il convient de préciser que ces régions géographiques peuvent ne pas demeurer comparables au fil du temps, car les frontières nationales peuvent changer d'une année à l'autre. On peut penser par exemple à la dissolution de l'URSS, de la Yougoslavie et de la Tchécoslovaquie, à l'union du Sikkim et de l'Inde et à celle du Vietnam et du Vietnam du Nord, ou encore à la création du Soudan du Sud.

Certaines personnes indiquent dans le fichier d'établissement le Canada comme pays de naissance. Normalement, les personnes nées au Canada reçoivent cette citoyenneté à la naissance et n'ont pas à faire de demande de résidence permanente. Les personnes indiquant être nées au Canada sur le fichier d'établissement sont probablement des enfants de diplomates étrangers nés lorsqu'ils résidaient au Canada et qui ont choisi ultérieurement de demander le statut de résident permanent.

Plusieurs variables géographiques des ensembles de données du FFT1 font référence à des notions de régions géographiques légèrement différentes de celles du fichier d'établissement. Le renseignement géographique le plus détaillé du FFT1 est disponible au niveau du secteur de recensement; il est dérivé **du code postal de l'adresse postale** (PSCO_). Le code postal indique généralement l'adresse de résidence lors de la déclaration de revenus au printemps de l'année suivante. L'adresse postale peut également faire référence à une entreprise, comme un cabinet comptable ou d'avocats, et n'est pas nécessairement l'adresse actuelle de la personne en question. La **province de résidence** au 31 décembre de l'année d'imposition (PRCO_) peut ne pas être identique à la province de l'adresse postale. Cette distinction est importante, car la variable dérivée de la région géographique de recensement peut ne pas correspondre à la province de résidence au 31 décembre (PRCO_);

elle devrait cependant correspondre au **code de province** (PR__). La variable PRHO_ indique une autre valeur d'adresse postale et existe uniquement à partir de 2008. De plus, alors que la variable de **code de province d'imposition** (TXPCO_) est, par définition, identique à la province de résidence au 31 décembre (PRCO_), ce code de province d'imposition (TXPCO_) est moins fiable (car des problèmes connus de qualité de données concernent cette variable, pour laquelle les valeurs manquantes et celles relatives à Terre-Neuve-et-Labrador reçoivent le code « 0 »). Pour en savoir davantage concernant la qualité de variables, s'il vous plaît consulter la section 7.

Utiliser les variables des fichiers de données fiscales pour étudier la mobilité géographique des immigrants exige de tenir attentivement compte du moment en question avant de déduire un déménagement et l'emplacement de travail. Un guide du chercheur portant sur l'étude de la mobilité et de la présence est inclus en annexe D.

3.4.2 Variables temporelles

L'« année d'admission » et l'« année d'imposition » sont des variables temporelles souvent utilisées pour produire des tableaux et effectuer des analyses à l'aide de la BDIM. L'année d'admission est l'année où l'immigrant a reçu le statut de résident permanent, alors que l'année d'imposition est l'année pour laquelle la déclaration de revenus est produite.

Il est recommandé de considérer la première année de revenu comme l'« année d'admission + 1 », puisqu'il s'agit de la première année complète où la personne se trouvera dans le pays. Il convient d'interpréter avec précaution les revenus déclarés pour l'année d'admission. Tout d'abord, environ 50 % de chaque cohorte d'admission produisent une première déclaration de revenus pour l'année d'admission (proportion calculée en fonction des déclarants à partir des immigrants couplés). Ensuite, les revenus déclarés l'année d'admission peuvent ne pas représenter une année de revenus complète. Une personne admise en octobre 2010 déclarera, au printemps 2011, des revenus pour seulement 3 mois, alors qu'une personne admise en janvier 2010 déclarera des revenus pour 12 mois.

Des déclarations de revenus sont également possibles après le décès d'une personne, lorsqu'une personne apparentée au décédé produit une déclaration de revenus en son nom. La variable de type de famille (FCMP_) du FFT1 doit être utilisée dans ce cas. Veuillez consulter le dictionnaire des données pour plus de précisions.

Le tableau 1 illustre, par exemple, les scénarios possibles et décrit les enregistrements à inclure dans une étude visant à évaluer les résultats socioéconomiques de la cohorte des immigrants de 1995 à 2000, cinq ans après leur admission. Pour qu'un enregistrement soit inclus, l'immigrant doit avoir été admis entre 1995 et 2000 et avoir produit une déclaration de revenus cinq ans après son admission. Cette analyse inclurait les enregistrements de la BDIM suivants : IM19952 et IM19963.

Tableau 1
Exemple de définition d'une cohorte d'intérêt

IMDB_ID	Année d'admission	Années d'imposition disponibles	Inclus dans le champ d'enquête
IM19801	1980	1982 à 2013	Non; admis avant 1995
IM19952	1995	1988 à 2011	Oui
IM19963	1996	1996 à 2013	Oui
IM19974	1997	2010 à 2013	Oui, mais pas de fichiers de données fiscales disponibles cinq ans après l'admission
IM20095	2009	2011 à 2013	Non; admis après 2000

Note : Cet exemple se fonde sur des données fictives.

Source : Statistique Canada, exemple provenant de la base de données longitudinales sur l'immigration.

L'un des inconvénients d'utiliser les données fiscales administratives est leur manque de précision quant aux dates. Outre l'année pour laquelle les revenus sont déclarés, aucune variable de temps n'existe dans le FFT1. Cela présente des difficultés pour étudier les durées d'emploi et de chômage, le moment des déménagements et les dates de mariage. Cette lacune ne permet pas non plus de distinguer facilement les travailleurs autonomes qui sont également des employés saisonniers de ceux qui reçoivent un revenu simultanément de deux sources. Malgré ces limites, il est tout de même possible de prendre des décisions en matière de date en tenant compte des points suivants.

Puisque les revenus de l'année précédente sont généralement déclarés au printemps de l'année civile suivante, une incertitude peut exister quant à l'année précise d'un changement d'adresse. Une personne A, par exemple, pourrait déclarer ses revenus de 2011 avec une adresse postale à Toronto, puis déclarer ses revenus de 2012 avec une adresse postale à Ottawa. On peut en déduire que cette personne A vivait à Toronto lors de sa déclaration de revenus de 2011 au printemps 2012 et vivait à Ottawa au printemps 2013. La personne A a pu déménager en 2012, après avoir produit sa déclaration de revenus pour l'année précédente, ou au cours des premiers mois de 2013, avant de produire sa déclaration de revenus pour 2012. Si cette personne A a changé de province, la variable de la province de résidence au 31 décembre (PRCO_) pourrait être utile pour déduire l'année du déménagement, puisqu'elle porte sur la province de résidence de cette personne en décembre de l'année d'imposition. L'adresse postale ne correspond pas nécessairement au lieu de résidence.

3.4.3 Variables de scolarité

Plusieurs variables du fichier d'établissement, comme le nombre d'années d'études et le niveau de scolarité, permettent de mesurer le niveau de scolarité à l'admission. La première prend la forme d'une réponse écrite à la question « Combien d'années d'études avez-vous terminées? » La deuxième est un choix parmi des options correspondant au niveau de scolarité le plus élevé atteint. La variable de « niveau de scolarité » dérivée combine les renseignements provenant de ces deux questions.

Des problèmes de qualité des données ont été relevés pour ces variables de scolarité depuis 2011. Par exemple, une proportion significative des personnes n'ont pas précisé leur niveau de scolarité ou leur nombre d'années d'études; le code « 0 » (« aucun ») a été attribué à ces variables EDUCATION_QUALIFICATIONS et YEARS_OF_SCHOOLING plutôt que « donnée manquante ». En 2011, jusqu'à 35 % des immigrants ont indiqué ne pas avoir suivi d'études, par rapport à environ 10 % dans les années 1990. Les variables relatives à la scolarité de 2011 à 2020 ont fait l'objet d'une imputation pour résoudre ce problème. Si vous souhaitez davantage de détails sur l'imputation, consultez la section 7.3.

3.4.4 Variables de profession envisagée

IRCC recueille la profession envisagée à partir de l'enregistrement d'admission et lui attribue une classification correspondant aux codes de 2011 Classification nationale des professions (CNP) du fichier d'établissement. Il s'agit de la classification la plus large en matière de compétences; les codes CNP à cinq chiffres étant les plus précis (voir le dictionnaire en annexe pour des définitions complètes).

Même si la profession envisagée est également considérée comme un remplacement valable de la profession de la personne dans son pays d'origine, des précautions sont recommandées. Pour pouvoir indiquer une profession envisagée, le demandeur doit démontrer qu'il possède les qualifications requises en matière d'études ainsi qu'au moins une année d'expérience dans le domaine. IRCC considère, par conséquent, ce renseignement comme une mesure prudente de la variable envisagée après l'admission, puisque les exigences qui l'accompagnent sont relativement strictes. Les étudiants venant d'obtenir un diplôme de génie, par exemple, ne peuvent pas indiquer « ingénieur » comme profession envisagée (du fait de leur manque d'expérience professionnelle) et sont plutôt classés comme « étudiants ». D'autres variables, comme **l'intention sur le marché du travail** (LM_INTENTION) et le **niveau de compétence** (SKILL_LEVEL), peuvent servir à obtenir des renseignements sur le domaine professionnel de la personne dans son pays d'origine. Il convient, en outre, de souligner que le champ de profession envisagée est obligatoire pour les immigrants demandeurs principaux au sein des catégories économiques. Pour les autres immigrants, ce renseignement pourrait être une mesure moins fiable de la profession envisagée.

3.4.5 Autres variables de la BDIM

Seules les variables nécessitant une explication détaillée et présentant d'éventuelles difficultés importantes pour les analystes ont été incluses à la section 3.4. Pour de plus amples détails sur les variables incluses dans la BDIM, veuillez-vous reporter aux dictionnaires de la BDIM. La composante relative aux impôts décrit les variables incluses dans les fichiers IMDB_T1FF, alors que la composante relative à l'immigration décrit les variables incluses dans les autres fichiers. Ces dictionnaires sont à la disposition des utilisateurs des données, mais il est également possible de les obtenir en faisant la demande par écrit à [Statistique Canada](https://www150.statcan.gc.ca/n1/pub/92-626-x/2016001/article/00001-eng.htm).

Les utilisateurs des données de la BDIM doivent garder à l'esprit que les données des fichiers relatifs à l'immigration et les données fiscales sont recueillies à divers moments et que, dans certains cas, les caractéristiques des personnes évoluent au cours du temps. Par conséquent, l'état matrimonial et la composition de la famille d'une personne peuvent changer au cours de l'année et différer entre le FRPN et le FFT1. Les variables à utiliser dans l'analyse dépendent de l'objet de l'étude.

4 Couplage d'enregistrements⁹

Tel que décrit dans ce document, la BDIM est le résultat de plusieurs couplages d'enregistrements. Elle a été créée pour fournir des renseignements statistiques anonymes. Cette section offre un aperçu des méthodes de couplage d'enregistrements utilisées pour créer la BDIM. Pour plus de détails concernant le traitement de données relatif au couplage d'enregistrement, veuillez consulter la section 5.

Le couplage d'enregistrements est le processus de mise en correspondance d'enregistrements entre ou au sein de bases de données. Cette approche est couramment utilisée pour combler des lacunes statistiques et créer un ensemble de données présentant de vastes applications (Rotermann and al. 2015).

Pour produire la BDIM, l'Environnement de couplage de données sociales (ECDS) est utilisé. L'ECDS est un environnement de couplage très sécurisé qui facilite la création de fichiers de données couplés sur la population pour l'analyse sociale.

Au centre de l'ECDS se trouve le Dépôt d'enregistrements dérivés (DED ou Dépôt), une base de données relationnelle dynamique nationale renfermant seulement des identificateurs personnels de base. Le DED est créé en couplant différents fichiers index sources de Statistique Canada dans le but de produire une liste de particuliers. Ces fichiers, qui renferment des identificateurs personnels sans variables d'analyse, sont intégrés à l'environnement, traités et couplés seulement une seule fois au DED. Les mises à jour de ces fichiers de données sont couplées au DED de façon régulière.

En 2020, le taux de couplage au dépôt pour les données d'immigration était de **97,4 %** (Cascagnette, 2021) et la méthode probabiliste a été utilisée pour coupler les données sur l'immigration d'IRCC aux données fiscales de l'ARC. Pour effectuer le couplage d'enregistrements. Le système généralisé de couplage d'enregistrement G-Coup a été utilisé.

Le système généralisé de couplage d'enregistrements qu'utilise Statistique Canada, G-Coup, se fonde sur la théorie mathématique du couplage d'enregistrements élaborée par Ivan P. Fellegi et Alan B. Sunter. La méthodologie de couplage d'enregistrements probabiliste compare des identificateurs non uniques (p. ex. nom et date de naissance) et estime la probabilité que les enregistrements correspondants portent sur la même entité (p. ex. personne). Le couplage d'enregistrements probabiliste s'avère particulièrement utile lorsque les identificateurs sont susceptibles de changer (p. ex. nom de famille des femmes qui se marient), peuvent comporter des erreurs et sont souvent absents.

La comparaison des enregistrements est faite champ par champ au moyen de règles de comparaison dont les niveaux de concordance sont correspondance exacte, proximité des chaînes, information manquante ou non-concordance entre les champs. Un niveau de concordance est produit pour chaque règle en fonction de la similitude des valeurs d'une paire d'enregistrements. Un modèle de comparaison des résultats est attribué à chaque paire d'enregistrements. Ce modèle est évalué afin de classer les paires comme étant liées, possiblement liées ou non liées.

La théorie du couplage d'enregistrements probabiliste se fonde sur la prémisse selon laquelle les résultats de certains modèles de comparaison des résultats sont caractéristiques de paires véritablement couplées, tandis que d'autres sont caractéristiques de paires véritablement non couplées. Ainsi, un poids est attribué à chaque niveau de concordance d'une règle en fonction du ratio de la probabilité estimée que le niveau de concordance soit obtenu pour les correspondances véritables à la probabilité estimée que le niveau soit obtenu pour les non-correspondances.

Comme on ne connaît pas d'avance la composition de l'ensemble couplé, on ne connaît pas les probabilités des modèles de résultats pour les enregistrements véritablement couplés. Les composantes de poids couplées sont estimées à partir de connaissances antérieures et de versions antérieures du processus de couplage. Elles sont précisées lors d'itérations successives du processus de couplage.

9. La majorité de cette section provient du rapport de méthodologie (Cascagnette, 2021).

Les composantes de poids non couplées sont calculées en fonction de la fréquence à laquelle les niveaux d'une règle ont été observés parmi des paires d'enregistrements qui ne vont pas ensemble, ce qui équivaut approximativement à la fréquence à laquelle les niveaux d'une règle seraient observés parmi des enregistrements associés de manière aléatoire. Après une itération répétée du processus de couplage, les composantes de poids couplées se stabilisent, et il est possible d'utiliser les poids finaux.

La stratégie du couplage d'enregistrements probabiliste comporte les six étapes suivantes :

1. Produire des paires potentielles au moyen d'un critère initial.
2. Créer des règles de comparaison et les appliquer aux paires potentielles pour obtenir des ratios de probabilité.
3. Appliquer des poids de fréquence.
4. Affecter des états de couplage aux paires au moyen de ratios de probabilité et de seuils.
5. Former des groupes.
6. Résoudre les conflits au moyen de la mise en correspondance.

Il faut répéter les étapes 2 à 4 de manière itérative. Chaque étape est décrite de manière plus détaillée ci-dessous.

Les utilisateurs d'un ensemble de données créé à partir d'un couplage d'enregistrements doivent être conscients que des erreurs de couplage sont possibles. Un couplage d'enregistrements peut produire quatre résultats : de réelles correspondances correctement classées comme des correspondances, de réelles correspondances classées à tort comme des non-correspondances, de réelles non-correspondances classées à tort comme des correspondances ou de réelles non-correspondances correctement classées comme des non-correspondances (Winkler, 2009). Comme le montre l'exemple du tableau 2, dans lequel les enregistrements du fichier 1 sont couplés aux enregistrements du fichier 2, le résultat du couplage d'enregistrements entre deux variables sera soit une correspondance soit une non-correspondance. Un bon couplage d'enregistrements maximisera la proportion des réelles correspondances correctement classées comme des correspondances et la proportion des réelles non-correspondances correctement classées comme des non-correspondances, et minimisera les autres résultats éventuels du couplage d'enregistrements.

Tableau 2
Exemple de résultats de couplage d'enregistrements

	Enregistrement	Fichier 2			Type de résultat
		A	B	D	
Fichier 1	A	Correspondance	Non-correspondance	Non-correspondance	Réelle correspondance
	C	Non-correspondance	Correspondance	Non-correspondance	Fausse correspondance
	D	Non-correspondance	Non-correspondance	Non-correspondance	Fausse non-correspondance
	E	Non-correspondance	Non-correspondance	Non-correspondance	Réelle non-correspondance

Source : Statistique Canada, exemple de résultats de couplage d'enregistrements.

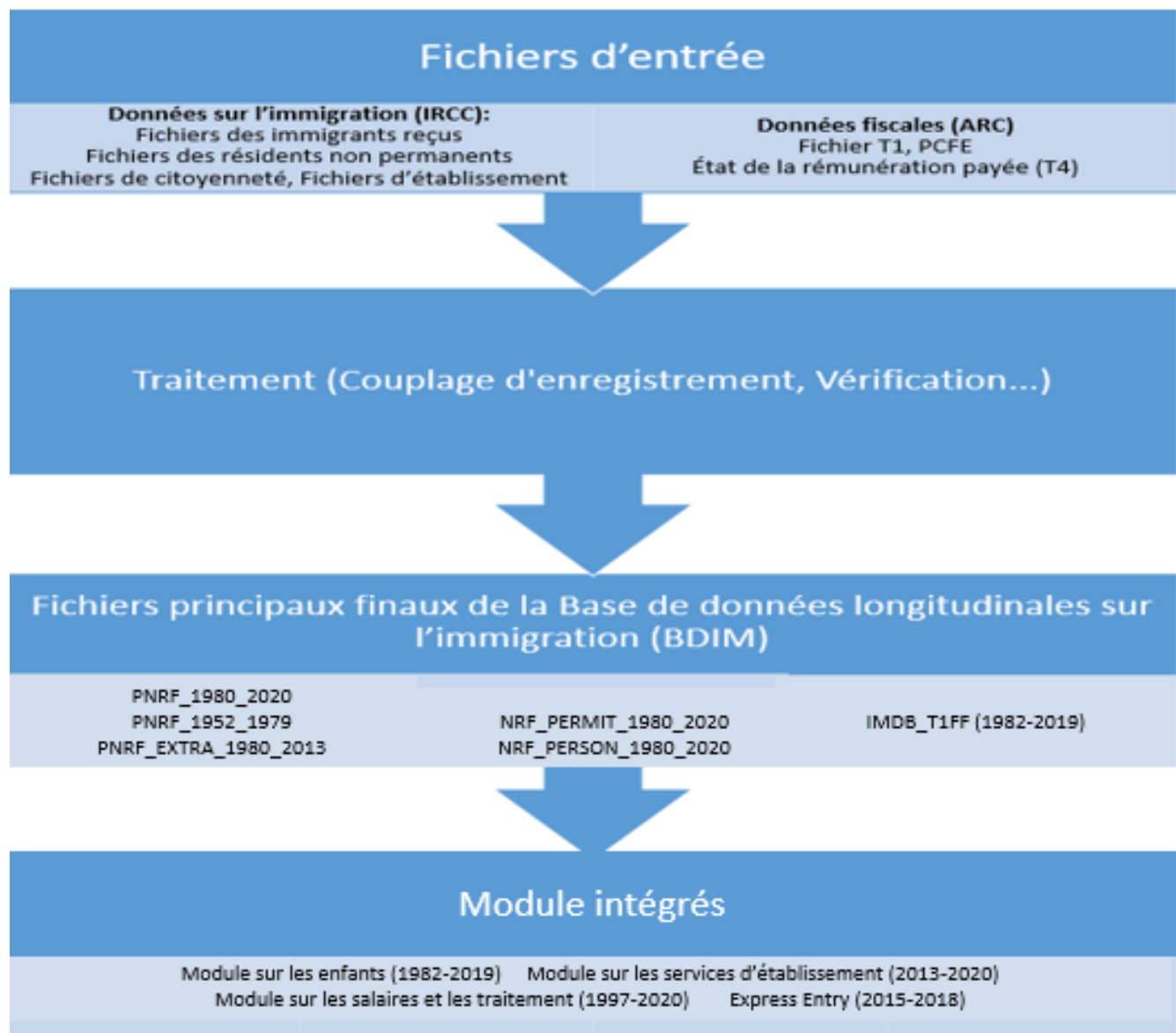
Les résultats d'un couplage d'enregistrements probabiliste dépendent de la qualité des variables de couplage. Des noms mal orthographiés, par exemple, ou des coquilles dans la date de naissance peuvent empêcher des correspondances ou créer des erreurs de correspondance. Une non-correspondance ne signifie pas nécessairement que la personne n'a pas produit de déclaration de revenus. Les taux de couplage d'enregistrements pour la BDIM la plus récente sont fournis à la section 7.2.1.

5 Traitement des données

5.1 Traitement

Plusieurs organismes gouvernementaux participent à la création et au traitement de la BDIM. De la collecte initiale des données à la diffusion en passant par le traitement, la coopération de ces organismes est nécessaire pour fournir la qualité élevée des données qu'attendent les utilisateurs des données de Statistique Canada. À chaque étape de la séquence de traitement, des vérifications attentives, manuelles et automatiques, de la qualité des données sont effectuées et des boucles de rétroaction servent à corriger toute erreur détectée à la source. La section suivante décrit brièvement le processus annuel de traitement qui permet de mettre la BDIM à jour.

Figure 3
Résumé du déroulement du processus de création de la BDIM



Note : Voir le glossaire pour connaître les définitions des acronymes.

Source : Statistique Canada, description du déroulement du processus de création de la BDIM.

Comme le montre la figure 3 (les fichiers finaux sont nommés avec l'acronyme anglais dans la BDIM), Statistique Canada reçoit d'abord les données de T1 de l'Agence du revenu du Canada (ARC), dans un fichier intitulé « Fichier maître des particuliers » et d'autres fichiers de données fiscales. Les fichiers de données fiscales servent alors à créer le FFT1, au sein duquel les personnes sont couplées à leurs conjoints et à leurs enfants par un identificateur commun et des variables géographiques sont créées. Statistique Canada effectue des vérifications manuelles de la qualité et compare les estimations du FFT1 à d'autres sources de données, comme le recensement (pour les années de recensement) et l'Enquête sur la dynamique du travail et du revenu, ainsi que des statistiques de revenu annuel que produit l'ARC.¹⁰

En matière d'immigration, IRCC fournit les données relatives aux immigrants admis, aux résidents non permanents et citoyens qui sont utilisées pour produire la BDIM. Ces données servent à créer le Fichier des immigrants reçus (FIR) et le Fichier des résidents non permanents (FRN). On considère que les fichiers FIR et FRN sont des recensements complets des permis de résidence permanente et temporaire qu'a délivrés IRCC depuis 1980.

Outre l'ajout des renseignements de l'année d'imposition la plus récente, une vérification complète des années précédentes a lieu afin d'ajouter des données fiscales pour toute nouvelle personne couplée. Cela peut signifier que les enregistrements de déclaration de revenus d'un immigrant admis ou d'un résident non permanent peuvent ne pas être couplés dans la BDIM une certaine année, mais que ses déclarations de revenus suivantes peuvent être couplées pour une année ultérieure. À mesure que la méthodologie s'améliore, la vérification rétroactive peut permettre de coupler également toutes ses déclarations de revenus antérieures, si elles figurent dans le FFT1. C'est ainsi qu'après le traitement des données fiscales les plus récentes, des personnes admises et ayant produit une déclaration de revenus plusieurs années auparavant peuvent encore être ajoutées à la BDIM. Lorsque plusieurs admissions depuis 1980 existent pour une personne, on conserve les données de sa première admission.

Bien que les déclarations de revenus d'une année donnée soient généralement produites au printemps de l'année suivante (p. ex. déclaration des revenus de 2013 en 2014), des exceptions existent. Parfois, une personne produit sa déclaration de revenus plus tard dans l'année et n'est alors pas incluse dans le traitement des T1 qu'effectue Statistique Canada pour cette année-là. Lorsque ce fichier est transmis aux fins de traitement pour la BDIM, ces déclarants retardataires ne sont pas inclus et ne le seront pas non plus lors du traitement de l'année suivante, puisque le FFT1 n'est pas mis à jour. De la même manière, les personnes qui produisent leur déclaration de revenus pour des années antérieures ne sont pas ajoutées à la BDIM pour ces années, puisque les FFT1 des années précédentes ne sont pas mis à jour. Dans ces cas, la première déclaration de ces personnes effectuée au moment attendu s'affichera comme leur première année dans la base de données.

Pour l'instant, on exécute une série de programmes pour évaluer la qualité des données et les taux de couplage, afin d'éviter tout doublon et de signaler les valeurs aberrantes. Une fois la base de données couplée, elle est jugée complète et prête à être mise à disposition.

Finalement, la base de données consiste en des fichiers SAS, un fichier de données fiscales par année depuis 1982 (IMDB_T1FF_&YEAR) et des fichiers sur les données relatives à l'immigration (**PNRFF_1980_2020**, **PNRF_EXTRAF_1980_2013**, **PNRF_1952_1979** et **NRF_PERMITF_1980_2020**). Tous ces fichiers sont décrits dans la section 2. L'identificateur personnel unique de la BDIM (IMDB_ID) sert à relier tous ces fichiers (voir l'annexe D.1 pour des conseils de programmation).

5.2 Couplage du fichier des résidents non permanents (FRN)

Le Fichier des résidents non permanents (FRN) qu'IRCC fournit couvre les enregistrements de permis de résidence temporaire délivrés à partir de 1980. Il fournit des données démographiques sur les résidents non permanents ainsi que des renseignements détaillés sur leur permis, tels que le type de permis et leur plage de validité.

10. Pour plus d'informations sur le [traitement et la qualité des données du FFT1](#).

Le FRN contient des millions d'observations. Elles comprennent cependant des doublons, selon lesquels une même personne peut compter plusieurs identifiants différents. Le problème provient principalement d'enregistrements à la fin des années 1980 pour lesquels le numéro d'identification de la personne initiale a été perdu. Ces enregistrements ont été supprimés en couplant le FRN à lui-même. Cela a permis de déterminer qu'environ 220 000 enregistrements (environ 400 000 observations) étaient des doublons. Dans les cas où les deux enregistrements de résidents non permanents possédaient chacun un enregistrement d'admission, le lien de doublon a été supprimé (ce qui s'applique à moins de 1 000 enregistrements), puisqu'on considère que le fichier d'établissement contient des identifiants uniques. Après nettoyage, il ne reste que des résidents non permanents distincts.

Les deux fichiers relatifs à l'immigration (FIR et FRN) contiennent des données démographiques. Toutefois, ces données démographiques ne sont pas toujours cohérentes. Cela se produit lorsque plusieurs sources sont disponibles ou en cas de conflit. Il a été décidé que les données du FIR relatives au Fichier intégré des résidents permanents et non permanents (FRPN) seraient conservées, du fait de problèmes de qualité concernant les données du FNR au cours de ses premières années.

5.3 Variables dérivés inclus dans le FFT1

Une fois que les couplages d'enregistrements se sont effectués, on ajoute au FFT1 des variables relatives à l'immigration pour les immigrants et les résidents temporaires.

Pour identifier le statut d'immigration d'un déclarant, l'année d'admission (LANDING_YEAR) et la première année de validité (FIRST_EFFECTIVE_YEAR) ont été créés. Par conséquent, l'indicateur de résident non permanent **TR_IND**, a été enlevé.

Des variables dérivées indiquant et décrivant les familles sont également créées. Dans chaque FFT1 annuel, il est possible d'estimer le nombre d'immigrants dans une famille admise depuis 1980 (variable **IMM80F&year**). Cependant, ceci peut être une sous-estimation car cette variable inclut seulement les déclarants et non les enregistrements imputés, donc les enfants sont sous-estimés. Il est également possible de déterminer si l'immigrant est un conjoint (pour l'année d'imposition donnée) et si ce conjoint est immigrant ou résident non permanent (variable **SP_IDI&year**). Les utilisateurs de données peuvent recenser les immigrants d'une même famille, pour chaque année d'imposition, à l'aide de la variable de **numéro d'identification de la famille** (FIN_). Cette variable présente la même valeur pour tous les membres d'une famille, c'est-à-dire l'identifiant IMDB_ID du membre de la famille le plus âgé admis depuis 1980. La qualité de ces variables dépend de la qualité du couplage d'enregistrements et des fichiers FFT1, puisque seules les personnes couplées sont comptabilisées (voir la section 7.5).

Les variables ayant le préfixe **TNK** représentent le nombre d'enfants déclarés d'un certain âge dans les familles d'immigrants et de résidents non permanents (voir la composante relative aux impôts du dictionnaire des données pour de plus amples détails). Le terme « enfant » désigne toute personne célibataire vivant avec un ou deux parents; un enfant peut être de tout âge. Au tableau 3, par exemple, la famille d'immigrants dont le numéro est IM19801 a deux enfants de 1 an en 2011 (TNK01I2011), alors que la famille IM19873 a trois enfants en 2011 (TNKIDI2011), un de 0 an (TNK00I2011), un de 1 an (TNK01I2011) et un de plus de 18 ans (TNK19I2011). L'immigrant IM20105 n'avait pas d'enfant en 2011.

Tableau 3
Exemple de variables relatives au nombre d'enfants dans une famille

IMDB_ID	TNK00I2011	TNK01I2011	TNKxxI2011	TNK19I2011	TNKIDI2011
	nombre				
IM19801	0	2	0	0	2
IM19802	0	1	0	0	1
IM19873	1	1	0	1	3
IM19994	0	0	0	1	1
IM20105	0	0	0	0	0

Note : Toutes les variables ne figurent pas dans ce tableau. Cet exemple se fonde sur des données fictives.

Source : Statistique Canada, exemple provenant de la base de données longitudinales sur l'immigration.

Une autre variable ajoutée au FFT1 est l'identificateur de valeur aberrante **OUTLIER_IND** (1: aberrante; 0: non). Il s'agit d'un indicateur ajouté pour détecter les enregistrements présentant des résultats extrêmes (voir la section 5.5 pour de plus amples détails) et devant être supprimés de tout tableau ou calcul. Les enregistrements marqués comme aberrants présentent des revenus extrêmes qui ajouteraient un biais aux résultats de l'analyse.

5.4 Variables dérivées incluses dans le FRPN

Lorsque le FRPN est produit, certaines variables relatives aux tendances de déclaration de revenus sont dérivées et incluses au fichier. La variable **FIRST_TAX_YEAR** indique la première année pour laquelle les données fiscales sont disponibles pour une personne donnée, alors que la variable **LAST_TAX_YEAR** indique la dernière année pour laquelle les données fiscales sont disponibles. Il convient de noter que des données fiscales ne sont pas nécessairement disponibles chaque année entre la première et la dernière année d'imposition. Par exemple, un cas où les variables `First_tax_year=1982` et `Last_tax_year=2012` existent ne signifie pas nécessairement que le déclarant a produit une déclaration de revenus de manière continue; les données fiscales pour 2006 peuvent manquer, par exemple. Lorsqu'il manque des variables pour la première et la dernière année d'imposition, **FIRST_TAX_YEAR** et **LAST_TAX_YEAR**, c'est pour désigner les non-déclarants ou les personnes qui n'ont jamais produit de déclaration de revenus auparavant. Il s'agit d'une mise à jour depuis la BDIM de 2018, les déclarants et les non-déclarants ont été fusionnés depuis.

La variable de déclarant avant l'admission **PREFILER_IND** sert à signaler les immigrants pour lesquels des données existent dans le FFT1 avant leur année d'admission. La plupart de ces cas ont été couplés à des enregistrements de résidents non permanents, comme on s'y attend (voir la section 7.2.4 pour de plus amples détails).

5.5 Détection des valeurs aberrantes

Après la création des fichiers `IMDB_T1FF`, on procède à la détection des valeurs aberrantes dans tous les fichiers de données fiscales, afin de relever tous les enregistrements aberrants. Un enregistrement est jugé aberrant lorsque l'on détermine qu'il contient une valeur ou plus de revenus extrême par rapport aux autres enregistrements. Les critères utilisés pour déterminer les valeurs aberrantes sont confidentiels. La variable **OUTLIER_IND** est créée pour recenser les enregistrements présentant des valeurs extrêmes.

L'indicateur de valeur aberrante **OUTLIER_IND** figure dans les fichiers de données fiscales, mais pas dans le FRPN. Cet indicateur peut être attribué à l'enregistrement d'une personne donnée pour une année particulière, sans que cela ne soit nécessairement le cas pour toutes les années pour lesquelles la personne a produit une déclaration de revenus. Toutes les valeurs aberrantes doivent être supprimées de toute analyse. Comme le montre le tableau 4, pour la personne `IM19802`, l'indicateur de valeur aberrante a seulement été attribué à l'enregistrement de 1983, alors qu'aucun des fichiers de données fiscales de la personne `IM19801` ne présente cet indicateur. Aucun indicateur de valeur aberrante n'apparaît en 2012 pour la personne `IM19994`, car aucune donnée fiscale n'est disponible pour cette personne cette année-là.

Tableau 4
Exemple relatif à l'indicateur de valeur aberrante

IMDB_ID	OUTLIER_IND1982	OUTLIER_IND1983	OUTLIER_INDyyyy nombre	OUTLIER_IND2012	OUTLIER_IND2014
IM19801	0	0	0	0	0
IM19802	0	1	0	0	0
IM19873	1	1	1	0	...
IM19994	0	0	0	...	0

... n'ayant pas lieu de figurer

Note : Toutes les variables ne figurent pas dans ce tableau. Cet exemple se fonde sur des données fictives.

Source : Statistique Canada, exemple provenant de la base de données longitudinales sur l'immigration.

Les valeurs aberrantes peuvent être supprimées des totalisations et de toute analyse. La BDIM exclut (les très rares) revenus importants, car ils influeraient sur les moyennes et donneraient une impression incorrecte de la situation en matière de revenus pour certains types d'immigrants. Prenons l'exemple fictif d'un revenu moyen de 40 000 \$ pour les Canadiens tchèques pour une année donnée; l'année suivante, cette moyenne atteint subitement 500 000 \$ parce qu'un joueur de hockey tchèque a été admis. Cela inclurait un biais dans la situation « réelle » des Canadiens tchèques en matière de revenus. Ainsi, le revenu « non représentatif » du joueur de hockey tchèque serait retiré des calculs. De plus, cet exemple présente une composante de confidentialité. Si une telle hausse du revenu était observée, il serait possible de déduire le revenu du joueur de hockey; ce qui constituerait une atteinte à la vie privée. Dans certains produits de la BDIM, le revenu médian, qui résiste mieux à l'influence d'importantes valeurs individuelles, est également fourni.

Lorsqu'une personne produit des tableaux ou analyse des données, les enregistrements jugés être des valeurs aberrantes pour une année donnée doivent être retirés des calculs pour l'année en question pour les raisons susmentionnées ou une règle de dominance doit être appliquée. Pour de plus amples détails, voir l'annexe D.6.

6 Dissémination

Une fois le couplage terminé, les fichiers de données (voir la section 6.3) sont enregistrés sur les serveurs de Statistique Canada, afin que les utilisateurs de données puissent créer des tableaux et des modèles à partir des résultats. Statistique Canada diffuse les résultats au moyen de produits tabulaires et analytiques, tout en respectant strictement les normes de confidentialité des données. Les membres des centres de données de recherche ont accès aux fichiers de micro données (voir section 6.2). Les règles de confidentialité sont maintenues afin d'assurer la sécurité et la sûreté des données (voir section 6.4).

[Accès aux données via le Programme CDR](#)

Les centres de données de recherche (CDR) sont des endroits sécuritaires où on peut consulter les micro données de STC qui se trouvent partout au Canada. Les CDR sont hébergés par 29 universités canadiennes et sont dirigés par des analystes de Statistique Canada, tandis que le CFDR (Centre fédéral de données de recherche), situé à Ottawa, a été créé pour répondre aux besoins analytiques des ministères fédéraux. Les CDR offrent des conditions sûres, régissant tous les aspects de l'utilisation des données, de l'accès aux données à leur publication.

6.1 Produits analytiques

Statistique Canada utilise un répertoire commun pour la diffusion de données en ligne. Ce répertoire de données comprend des tableaux qui présentent le revenu des immigrants selon diverses caractéristiques individuelles et géographiques. À partir de la page principale du moteur de recherche des données du site Web de Statistique Canada, on peut accéder aux tableaux de la BDIM en sélectionnant « Immigration et diversité ethnoculturelle », « Immigrants et résidents non permanents », puis « Base de données longitudinales sur l'immigration » sous « Enquête ou programme statistique ». Il est à noter que Statistique Canada a remplacé les tableaux du Système canadien de gestion de l'information socioéconomique (CANSIM) par ce répertoire commun de données en juin 2018. Les données de la BDIM qui étaient auparavant disponibles par l'entremise de CANSIM y sont désormais accessibles. Les mises à jour annuelles de la BDIM sont indépendantes les unes des autres. D'année en année, il peut y avoir des changements aux processus de création des données. Les mesures du revenu (moyennes et médianes) disponibles dans les tableaux sont les salaires, les traitements et les commissions, l'assurance-emploi, les revenus de placement, les revenus de travail indépendant et les prestations d'aide sociale (pour plus de détails sur la façon dont ces mesures sont calculées, voir l'annexe D.8).

Pour la BDIM 2020, dix tableaux ont été diffusés à l'échelle provinciale et nationale :

[Tableau 1 \(43-10-0026\)](#) : Revenu des déclarants immigrants selon le sexe, l'expérience avant l'admission, la connaissance des langues officielles à l'admission, la catégorie d'admission, l'année d'admission et l'année d'imposition, pour le Canada et les provinces;

[Tableau 2 \(43-10-0027\)](#) : Revenu des immigrants déclarants, selon le sexe, la catégorie d'admission d'immigrant, le profil socio-démographique, l'année d'admission et l'année d'imposition, pour Canada et les provinces;

[Tableau 3 \(43-10-0028\)](#) : Revenu des demandeurs d'asile selon le sexe, les groupes d'âge, la région de naissance, le statut de résidence, l'année de demande et l'année d'imposition, pour le Canada et les provinces;

[Tableau 4 \(43-10-0029\)](#) : Revenu des immigrants, selon le sexe, l'expérience avant l'admission, les connaissances des langues officielles à l'admission, la catégorie d'admission d'immigrant, l'année d'admission et l'année d'imposition, pour Canada, provinces et régions économiques;

[Tableau 5 \(43-10-0030\)](#) : Revenu des immigrants, selon le sexe, l'expérience avant l'admission, connaissance des langues officielles à l'admission, la catégorie d'admission d'immigrant, l'année d'admission et l'année d'imposition, pour Canada, provinces et régions métropolitaines de recensement;

[Tableau 6 \(43-10-0017\)](#) : Migration interprovinciale des déclarants immigrants selon la province d'admission, la province de résidence, les groupes d'âge durant l'année d'imposition par sexe, expérience avant l'admission, connaissance des langues officielles, catégorie d'admission, année d'admission et année d'imposition, Canada;

[Tableau 7 \(43-10-0018\)](#) : Migration interprovinciale des déclarants immigrants selon les groupes d'âge durant l'année d'imposition par sexe, connaissance des langues officielles, catégorie d'admission d'immigrant, expérience avant l'admission, année d'admission et année d'imposition, Canada et provinces;

[Tableau 8 \(43-10-0022\)](#) : Mobilité des immigrants déclarants, selon l'âge et sexe, les connaissances des langues officielles à l'admission, l'expérience avant l'admission, la catégorie d'admission d'immigrant, l'année d'admission et l'année d'imposition, Canada, provinces et régions métropolitaines de recensement

[Tableau 9 \(43-10-0024\)](#) : Mobilité des immigrants déclarants, selon le sexe, l'expérience avant l'admission, les connaissances des langues officielles à l'admission, la catégorie d'admission d'immigrant, l'année d'admission et l'année d'imposition, pour Canada, provinces et régions économiques

[Tableau 10 \(43-10-0031\)](#) : Résultats économiques des déclarants immigrants admis entre 0 et 14 ans en comparaison aux déclarants canadiens, selon l'âge à l'imposition, le sexe, la catégorie d'admission, la période d'admission, les groupes d'âge à l'admission, et le statut de faible revenu durant l'enfance, pour le Canada et les provinces

Tous les tableaux présentent le détail selon la province de résidence. La valeur de la province se fonde sur la province de résidence au 31 décembre de l'année d'imposition (variable PRCO).

Des tableaux supplémentaires seront diffusés en mars au niveau sous-provinciaux et sur les résultats des enfants immigrants.

Plusieurs articles analytiques relatifs à la BDIM ont été publiés au cours des années (voir l'annexe C). Les analystes de Statistique Canada répondent aussi aux demandes de données spécifiques de chercheurs et d'utilisateurs de données. Ils y répondent selon une formule de recouvrement des coûts.

6.2 Demande de fichiers analytiques

Une fois la BDIM publiée, tous les fichiers analytiques décrits dans le présent rapport (**ex. IMDB_T1FF_YEAR, PNRF_1980_2020, NRF_PERSON_1980_2020 and NRF_PERMIT_1980_2020**) sont également mis à la disposition des chercheurs sur place, qui obtiennent un droit d'accès lorsqu'ils reçoivent le statut de personne réputée être employée de Statistique Canada. Tout renseignement permettant l'identification est supprimé de ces micro données individuelles (comme une date de naissance exacte, la date d'admission, le numéro d'assurance sociale et le nom). Les chercheurs qui ne peuvent se rendre physiquement au bureau central de Statistique Canada peuvent accéder aux fichiers dans les centres de données de recherche (CDR) à travers le pays. Les centres de données de recherche permettent aux chercheurs un accès à de nombreuses enquêtes sur la population et les ménages, ainsi qu'aux données administratives, dans un environnement sécurisé au sein des universités. Les utilisateurs de la BDIM peuvent demander des totalisations personnalisées à Statistique Canada; ces demandes sont traitées selon une formule de recouvrement des coûts et ces coûts peuvent varier selon la nature et le type de chaque demande.

Avant la parution de tout résultat, Statistique Canada vérifie la confidentialité. Des exigences de taille de cellule minimale et l'arrondissement limitent les risques d'atteinte à la vie privée.

6.3 Autres programmes statistiques utilisant les données de la BDIM

Les données de la BDIM sont utilisées à diverses fins dans le cadre de nombreux programmes de Statistique Canada. La **Banque de données administratives longitudinales** (DAL) inclut dans son échantillon un échantillon de 20 % des enregistrements de la BDIM. La DAL utilise également les enregistrements de la BDIM pour ajouter des variables propres aux immigrants (comme l'année d'admission) à sa banque de données.

La **Base de données canadienne sur la dynamique employeurs-employés** (BDCDEE) est un ensemble de fichiers de données analytiques longitudinales que Statistique Canada maintient à jour pour fournir des correspondances de données entre employés et employeurs du marché du travail canadien depuis l'an 2001. Les fichiers de la BDCDEE couvrent toutes les personnes dont l'identité peut être établie dans les fichiers T1 et T4, ainsi que les renseignements sur les employeurs ou tout travail autonome auxquels les personnes peuvent être reliées. La BDIM est l'un des fichiers composant la BDCDEE et ce couplage permet aux chercheurs d'effectuer une analyse de la situation sur le marché du travail et des dynamiques de l'emploi relativement à la population immigrante au Canada.

L'Enquête sociale générale (ESG) de 2013 sur l'identité sociale (IS) recueille des renseignements détaillés sur les réseaux sociaux, ainsi que sur la participation et l'engagement civiques de la population canadienne. L'ESG de 2013 sur l'identité sociale a été couplée à la BDIM dans le but de sélectionner un échantillon représentatif de la population immigrante, afin de soutenir et d'évaluer les politiques et programmes relatifs aux immigrants. Immigration, Réfugiés et Citoyenneté Canada (IRCC) a en particulier utilisé ce fichier de données couplées pour élaborer un profil descriptif des relations sociales et de l'engagement civique des immigrants selon les catégories d'admission.

DEMOSIM,¹¹ modèle de microsimulation de Statistique Canada, utilise le fichier BDIM-DAL pour créer des projections de population pour les provinces, territoires, régions métropolitaines de recensement et certains découpages géographiques moins étendus, selon plusieurs caractéristiques. Les **programmes de recensement** ont recours à la base de données pour attester les données sur l'immigration.

La BDIM a été appariée à plusieurs enquêtes, notamment l'Enquête Longitudinale sur les Immigrants au Canada (ÉLIC), l'Étude Longitudinale et Internationale des Adultes (ELIA), et l'Enquête sur la Santé dans les Collectivités Canadiennes (ESCC). Suite à l'intégration de la BDIM dans l'ECDS d'autres programmes statistiques seront appariées à l'avenir.

6.4 Confidentialité¹²

Statistique Canada s'engage à respecter la vie privée des personnes. C'est la raison pour laquelle la sécurité et la sûreté des données sont des préoccupations majeures. Statistique Canada s'efforce de protéger les données des Canadiens. Tous les renseignements personnels créés, détenus ou recueillis par Statistique Canada sont protégés par la [Loi sur la protection des renseignements personnels](#) et par la [Loi sur la statistique](#) dans le cas des répondants à nos enquêtes. La Loi sur la statistique, la Loi sur l'accès à l'information et la Loi sur la protection des renseignements personnels font respecter la confidentialité des données. Pour plus d'information, visitez [Utilisation de nouvelles données et de données existantes pour les statistiques officielles](#). Pour obtenir de plus amples renseignements sur la confiance, la sécurité et la confidentialité des données, ainsi que la transparence et l'ouverture, veuillez consulter [Le Centre de confiance de Statistique Canada](#).

Du fait de son mandat unique à titre d'agence statistique nationale de collecte de renseignements personnels aux seules fins d'études statistiques et de recherche, Statistique Canada a préparé des [évaluations des facteurs relatifs à la vie privée](#) qui portent sur les problèmes de protection de la vie privée associés aux activités d'enquête.

Statistique Canada a mis en place une évaluation des facteurs relatifs à la vie privée¹³ lorsque le Comité des politiques (le haut comité exécutif de l'agence que préside le statisticien en chef) a approuvé les modifications importantes apportées à la base de données longitudinales sur l'immigration. L'objectif de cette évaluation était de déterminer s'il existait des problèmes de confidentialité, de sécurité et de protection de la vie privée liés à ces changements et, le cas échéant, de formuler des recommandations en vue de les résoudre ou de les atténuer.

Cette évaluation a conclu que, grâce aux mesures de protection existantes au sein de Statistique Canada ainsi qu'aux mesures additionnelles mises en place pour la Base de données longitudinales sur l'immigration, le risque de divulgation involontaire était extrêmement faible. L'importance des données sur le plan des politiques publiques l'emporte sur les préoccupations relatives à la vie privée. Les mécanismes de gouvernance en place constituent des mesures de protection contre l'utilisation inappropriée des données. Statistique Canada évalue régulièrement la pertinence de la BDIM au moyen d'examen périodiques qu'effectue le Comité des politiques, soutesant chaque fois la pertinence des renseignements par rapport à l'atteinte implicite portée à la vie privée.

Le travail statistique de l'agence comprend des projets de couplage d'enregistrements qui rassemblent des renseignements relatifs aux répondants individuels à des fins de recherche. Il s'agit d'une source reconnue de données statistiques précieuses, mais le couplage doit toujours être d'utilité publique. Afin de faire face à d'éventuelles atteintes à la vie privée dues à ce type de recherche, Statistique Canada respecte non seulement une directive, mais procède en outre à un examen bien défini et exige l'approbation de tous les [couplages d'enregistrements](#).

Pour protéger la vie privée, il est ainsi obligatoire d'arrondir les résultats tabulaires et descriptifs lors de la production de tableaux à partir des données de la BDIM (voir l'annexe D.5).

11. [DEMOSIM](#).

12. [Source](#).

13. [Site web](#).

7 Évaluation des données et indicateurs de qualité

7.1 Sources d'erreurs

Parce que la BDIM est produite à partir de plusieurs couplages d'enregistrements, elle comprend différentes sources d'erreurs, notamment des erreurs de couplage d'enregistrements, de mesure et de couverture. Nous expliquons, dans cette section, les sources d'erreurs et indiquerons l'importance de certaines d'entre elles.

Il convient de noter que, puisqu'il s'agit du recensement des immigrants déclarants admis depuis 1980, aucune pondération n'a été créée au sein de la BDIM. Aucune rectification n'est apportée aux années d'imposition manquantes ni aux erreurs de couplage, aucun échantillonnage n'a lieu et tous les déclarants couplés sont conservés dans l'ensemble de données final. Cependant, le couplage lui-même présente une forme d'erreur d'échantillonnage dans le cas d'échecs de couplage.

7.1.1 Erreurs de couplage d'enregistrements

Les ensembles de données produits à partir des résultats des couplages d'enregistrements peuvent comprendre des erreurs de couplage d'enregistrements. Deux **types d'erreurs** sont ainsi possibles : des faux positifs (fausses correspondances) et des faux négatifs (fausses non correspondances). Un couplage est un faux positif lorsque deux enregistrements ne concernant pas la même personne sont comptés comme une correspondance. Un couplage est un faux négatif lorsque deux enregistrements concernant la même personne sont comptés comme une non correspondance.

Il est possible de manquer une partie de l'historique d'un immigrant, puisque certains possèdent plusieurs numéros d'assurance sociale (NAS) au cours du temps (NAS temporaire attribué à l'admission au résident non permanent, puis NAS permanent attribué après l'admission). Ces deux NAS sont nécessaires pour constituer l'historique fiscal complet depuis l'admission de la personne au Canada. Le FCC et l'ECDS (décrit à la section 2.3) permettent de recenser ces deux NAS. Il est possible que, dans quelques cas, des liens entre les NAS échouent ou que des liens erronés soient créés.

7.1.2 Erreurs de mesure

Une **erreur de mesure** est la différence entre la valeur mesurée d'une variable et sa valeur réelle. On peut attribuer ce type d'erreur à plusieurs facteurs, notamment la saisie de données (p. ex. coquilles) et les erreurs des répondants (p. ex. mauvaise interprétation de la question posée). On a tenu compte de ce type d'erreur lors de la création du Fichier intégré des résidents permanents et non permanents (FRPN) pour éviter les renseignements contradictoires pour toute personne. Lorsqu'un enregistrement existe pour une personne dans le FIR et dans le FRN, par exemple, et que les variables sociodémographiques présentent des valeurs incohérentes, les valeurs à l'admission (celles du FIR) sont conservées. Voir les sections 7.2 et 7.5 qui fournissent certains chiffres.

7.1.3 Erreurs de couverture

Les **erreurs de couverture** proviennent d'omissions, de l'inclusion d'unités erronées, d'enregistrements en double et d'erreurs de classification d'enregistrements dans la base de données. Ces erreurs de couverture peuvent provenir d'une couverture inadéquate de la population. Elles peuvent se traduire par des estimations biaisées et les répercussions peuvent varier pour différents sous-groupes de la population. Ces erreurs entraînent souvent un sous-dénombrement. Le **sous-dénombrement** dans le cadre de la BDIM provient en partie de l'exclusion de la base de données des fichiers de données fiscales d'immigrants déclarants. Aucun enregistrement ne figure dans le fichier IMDB_T1FF pour les immigrants qui ne produisent pas de déclaration de revenus pour une année donnée ou qui le font en retard, même s'ils sont couplés aux données fiscales et font partie de la population d'intérêt. Si, pour une raison quelconque, l'enregistrement d'un immigrant n'est pas inclus dans le Fichier des immigrants reçus (FIR), il ne fera pas partie de la BDIM. Un **surdénombrement** survient lors de l'ajout à la base de données d'enregistrements exclus de la population cible. Un immigrant peut faire l'objet de plusieurs enregistrements dans le FIR du fait de plusieurs admissions non reconnues comme telles, par exemple. Veuillez-vous reporter à la section 7.4 et à l'annexe B pour de plus amples détails relatifs à la couverture de la BDIM.

7.2 Exactitude des données

Cette section porte sur l'exactitude des données relatives à l'immigration. Pour de plus amples détails sur l'exactitude du FFT1, veuillez-vous reporter au document sur le [FFT1](#) (numéro d'enregistrement 4105).

L'exactitude de la BDIM dépend de la représentativité de la population qui y est incluse. Une étude menée au cours des premières années de la BDIM a conclu que cette dernière « semble être représentative de la population la plus susceptible de produire des déclarations de revenus. Par conséquent, les résultats obtenus à partir de la BDIM ne devraient pas être appliqués à la population immigrante dans son ensemble, mais plutôt à l'univers des immigrants qui produisent des déclarations de revenus » (Carpentier et Pinsonneault, 1994).

Les raisons expliquant les différences entre les déclarants et la population totale des personnes nées à l'étranger sont expliquées dans un article de Badets et Langlois (2000) décrivant les défis d'utilisation de la BDIM :

Les caractéristiques de la population immigrante qui produit des déclarations de revenus diffèrent de celles de l'ensemble de la population née à l'étranger, étant donné que la probabilité ou la nécessité de produire une déclaration de revenus varie selon l'âge, la situation familiale de la personne et d'autres facteurs. On pourrait s'attendre à ce qu'un pourcentage plus élevé d'immigrants de sexe masculin produisent une déclaration de revenus, notamment parce qu'ils ont des taux de participation au marché du travail plus élevés que les femmes. Le niveau de « saisie » des immigrants dans la BDIM est aussi influencé par les changements touchant l'impôt sur le revenu. Par exemple, la mise en place de programmes de crédits d'impôt non remboursables à l'échelle fédérale et provinciale incite les personnes qui n'ont pas de revenus imposables à produire une déclaration de revenus pour être admissibles à certains crédits d'impôt.

Ces commentaires sur la représentativité de la BDIM s'appliquent toujours.

7.2.1 BDIM de 2020 : Taux de couplage

Cette section porte sur la **BDIM de 2020**. Le taux global de couplage entre les données d'immigration d'IRCC et le Dépôt d'enregistrements dérivés de l'ECDS était de **97,4 %** (voir section 4). Un lien ne signifie pas nécessairement qu'une déclaration fiscale est disponible puisqu'il est possible de coupler les personnes à charge des déclarants ou des immigrants qui n'ont pas encore produit leur déclaration de revenus. Ce taux de couplage théorique informe principalement sur la façon dont les fichiers d'IRCC pourraient être associés dans un environnement référentiel plus large.

Parmi les immigrants admis entre 1980 et 2019, 85,6 % étaient liés à au moins un FFT1. Ce taux représente la couverture effective du couplage entre les déclarations fiscales et les fichiers d'immigration. Tel que présenté dans les statistiques suivantes, ce taux de couverture peut varier selon le sexe et l'âge.

La proportion de déclarants couplés selon le groupe d'âge au moment de l'admission et le sexe est indiquée au tableau 5. On s'attend à ce que les taux plus faibles pour le groupe d'âge de 0 à 14 ans soient observés puisque les personnes de ce groupe d'âge ne sont pas en âge de travailler. Voir l'annexe B pour les taux selon le sexe, le groupe d'âge et la cohorte d'admission.

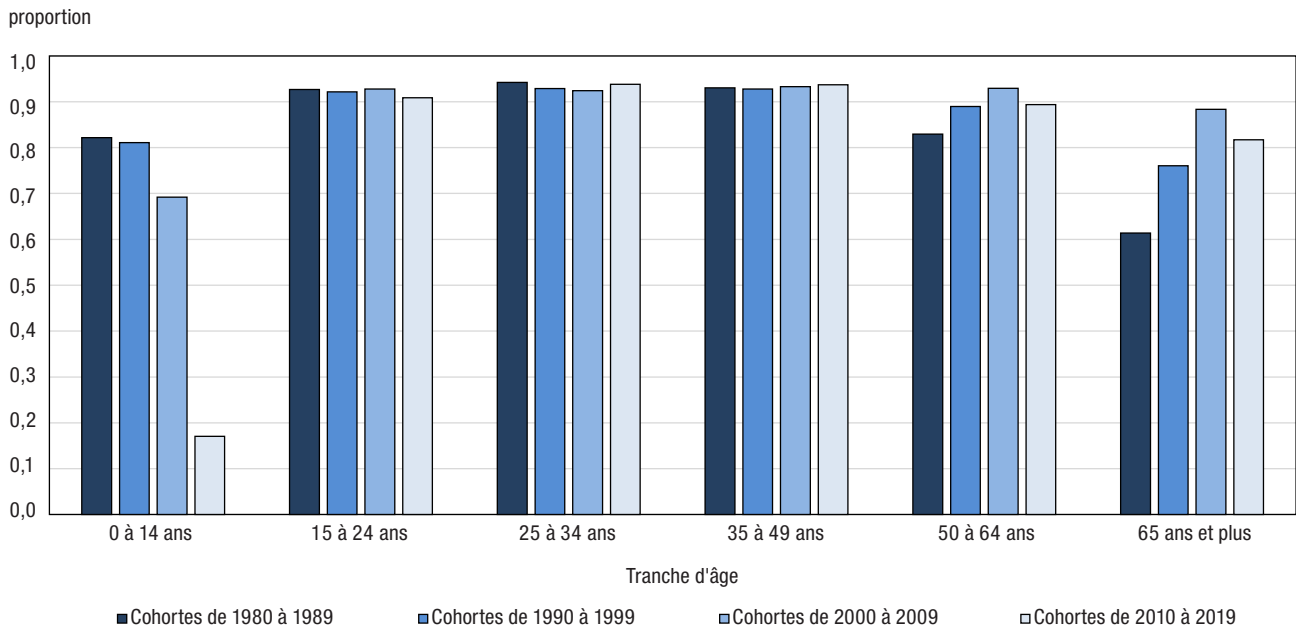
Tableau 5
Proportion de déclarants couplés, par tranche d'âge à l'admission et par sexe

	Âge à l'admission						Total
	0 à 14 ans	15 à 24 ans	25 à 34 ans	35 à 49 ans	50 à 64 ans	65 ans et plus	
	pourcentage						
Hommes	60,6	93,2	93,6	93,3	90,3	79,0	85,5
Femmes	59,7	92,2	93,7	93,8	88,7	77,3	85,7
Total	60,2	92,7	93,6	93,5	89,4	78,1	85,6

Source : Statistique Canada, Base de données longitudinales sur l'immigration de 2020.

À mesure que les immigrants vieillissent, ils commencent à produire des déclarations de revenus et sont inclus dans la BDIM. Le graphique 1 montre que, parmi les immigrants admis à tout âge de 0 à 14 ans, la proportion de déclarants couplés est supérieure pour les immigrants admis avant 2000 que pour ceux admis depuis 2000. Les immigrants récents ont des taux de couplage inférieurs. Voir l'annexe B pour un tableau présentant la proportion de déclarants couplés par groupe d'âge à l'admission, sexe et décennie d'admission.

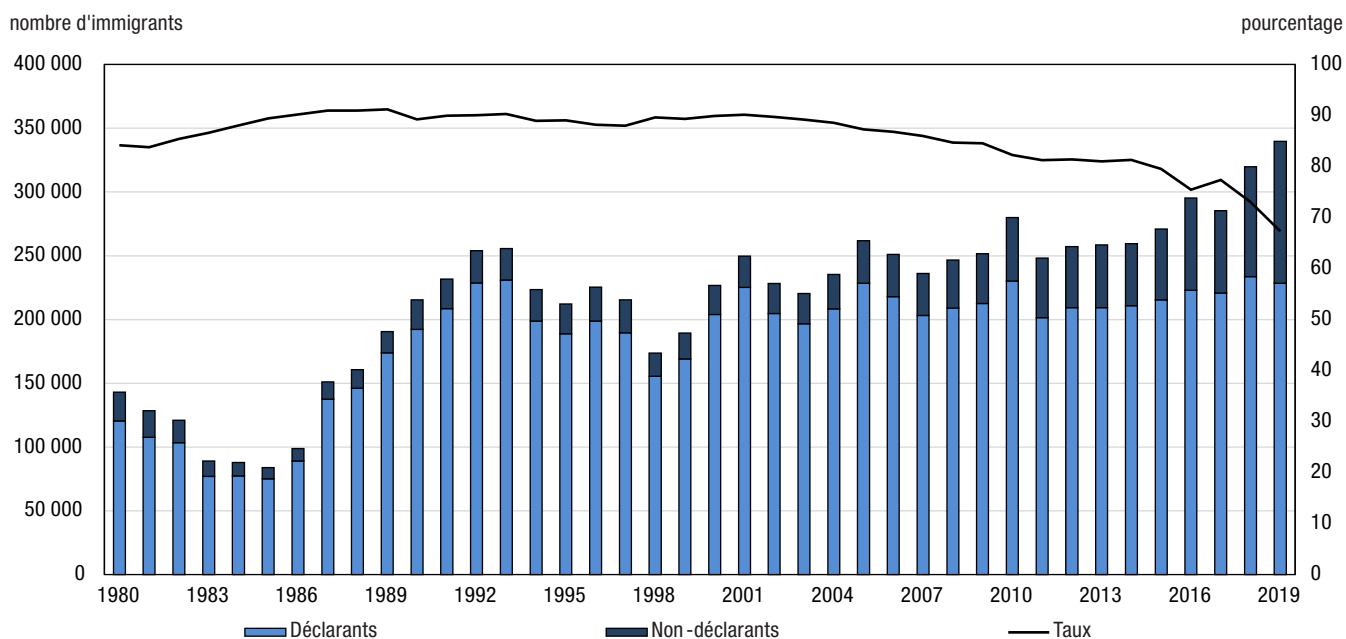
Graphique 1
Proportion de déclarants couplés, par tranche d'âge à l'admission et décennie d'admission



Source : Statistique Canada, Base de données longitudinales sur l'immigration de 2020.

Le graphique 2 illustre la proportion de déclarants ainsi que le nombre de déclarants et de non-déclarants par année d'admission; la mention « non-déclarant » indique qu'aucun enregistrement dans le FFT1 n'est disponible. Dans la **BDIM de 2020**, le taux de déclaration varie en fonction de l'année d'admission, de **67,3 %** pour les personnes admises en 2019 à **91,2 %** pour celles admises en **1989**. Le taux de déclaration augmente à mesure du nombre d'années de présence des immigrants au Canada; cela peut expliquer pourquoi les taux de couplage sont supérieurs pour les immigrants admis dans les années 1990 et le début des années 2000. Voir l'annexe B, tableaux 14 et 15, pour connaître des chiffres de répartition détaillés par année d'admission.

Graphique 2
Répartition des déclarants et des non-déclarants par année d'admission



Source : Statistique Canada, Base de données longitudinales sur l'immigration, 2020.

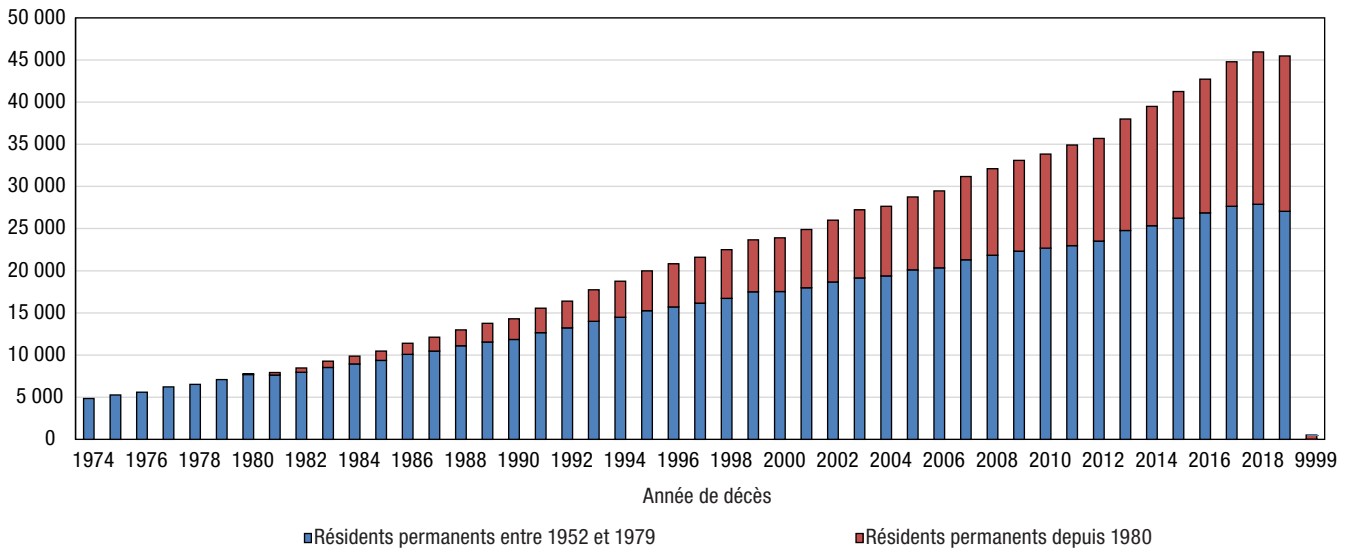
7.2.2 Disponibilité de la date de décès

L'année et le mois d'un décès, ainsi que l'indicateur de décès, sont inclus dans le FRPN. Dans la BDIM de 2020, ces variables ont été couplées à l'aide de la Base canadienne de données sur la mortalité (BCDM). Dans le passé, ces variables étaient fondées sur la Base de mortalité amalgamée (BMA) de Statistique Canada, qui est un ensemble de données archivé qui combine des enregistrements entre la BCDM, les statistiques de l'état civil, et des fichiers fiscaux. La BCDM est une base de données administratives qui recueille des renseignements sur les dates et les causes de décès dans tous les registres provinciaux et territoriaux de l'état civil au Canada. Un certain sous-dénombrement, bien que minime, existe dans la base de données, car il ne comprend pas les décès de Canadiens (1) qui sont morts à l'extérieur du Canada, à l'exception des États-Unis ; (2) qui ont servi comme membres des Forces canadiennes, ou (3) dont les corps n'ont pas été identifiés. Il est à noter que la BCDM ne comprend pas les décès qui ont été déclarés dans les dossiers fiscaux.

Le graphique 3 décrit la tendance générale du nombre de décès par an depuis 1974 pour les immigrants arrivés depuis 1952. Les données antérieures à 1980 ont été ajoutées récemment à la BDIM. La valeur « 9999 » représente les enregistrements d'immigrants décédés pour lesquels l'année de décès n'est pas disponible.

Graphique 3 Résidents permanents et non permanents par année de décès

nombre de décès

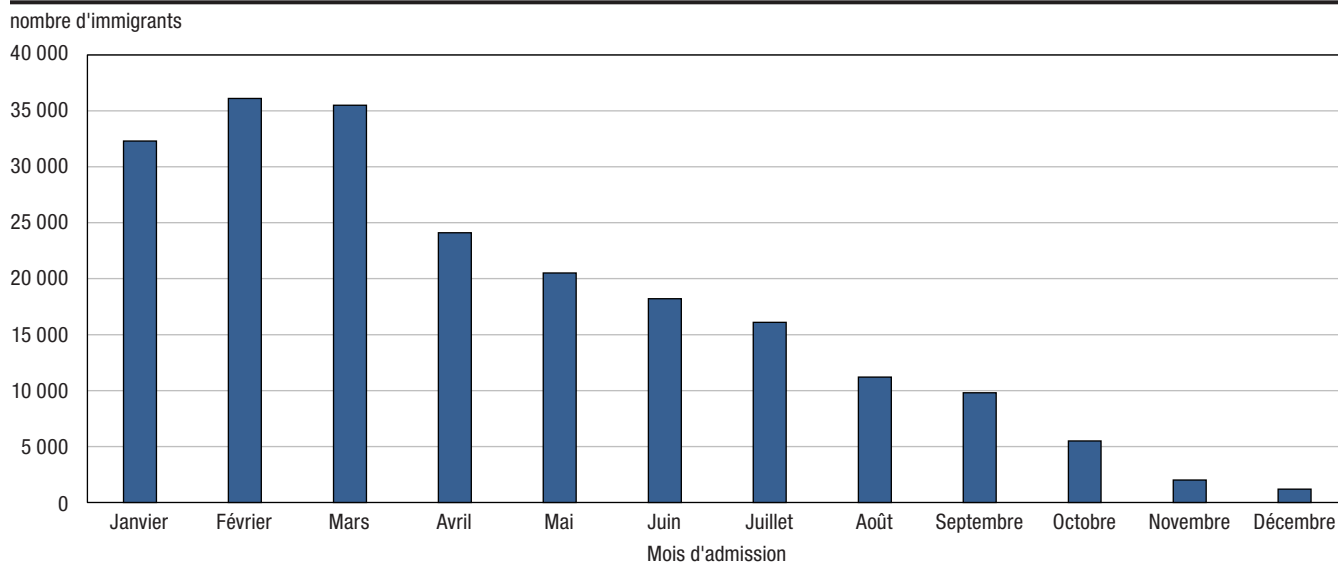


Note : La valeur 9999 a été utilisée lorsque la date de décès n'était pas disponible.
Source : Statistique Canada, Base de données longitudinales sur l'immigration, 2020.

7.2.3 Déclarants avant l'admission par rapport aux enregistrements du Fichier des résidents non permanents (FRN)

Les résultats inclus dans cette section sont tirés d'une étude fondée sur la BDIM de 2014. Les **déclarants avant l'admission** sont des immigrants qui ont produit une déclaration de revenus avant leur année d'admission. Il est parfois déduit que tous les déclarants avant l'admission sont des immigrants qui étaient des résidents non permanents avant leur admission. Cette section explique pourquoi ce n'est pas le cas. Au total, 1,26 million de personnes ont produit une déclaration de revenus avant leur admission officielle depuis 1980; sur ce total, 212 500 ne sont pas couplées à un enregistrement de résident non permanent, contrairement à ce à quoi l'on pourrait s'attendre. Après une enquête approfondie, on a découvert que la plupart des déclarants avant l'admission qui étaient des résidents permanents non couplés à un enregistrement de résident non permanent étaient probablement des immigrants ayant produit une déclaration de revenus lorsque cela n'était pas requis : 96 % de ces déclarants avant l'admission ont uniquement produit une déclaration de revenus pour l'année précédant leur année d'admission et 75 % n'ont déclaré aucun revenu (96 % n'avaient aucun salaire). Comme le montre le graphique 4, la plupart de ces déclarants avant l'admission ont été admis au cours des premiers mois de l'année, avant la date d'échéance de la production d'une déclaration de revenus pour l'année précédente. Il semble que certains immigrants admis avant le mois de mai ont produit une déclaration de revenus pour l'année précédant leur année d'admission; ce qu'ils n'étaient pas tenus de faire.

Selon ces résultats, il est justifié de supprimer les enregistrements présentant les variables de déclarant avant l'admission `Prefiler_ind=1` et `FIRST_EFFECTIVE_YEAR=.` des études sur les immigrants ayant une expérience avant leur admission dépend de l'analyse, puisque `FIRST_EFFECTIVE_YEAR=.` signifie qu'aucun enregistrement ne figure dans le fichier de permis de résidents non permanents.

Graphique 4**Répartition des déclarants avant l'admission sans permis de résidence non permanente, par mois d'admission**

Source : Statistique Canada, Base de données longitudinales sur l'immigration, 2014.

Tous les immigrants ayant une expérience avant leur admission ne sont pas considérés comme étant des déclarants avant l'admission : 478 100 immigrants disposent d'enregistrements de résidents non permanents avec la variable `Prefiler_ind=0`. En fonction du sujet étudié, utiliser la variable `FIRST_EFFECTIVE_YEAR<>`, ou le nombre de permis de résidence temporaire (variable `NUMBER_ALL_PERMITS`) est plus approprié pour étudier les immigrants ayant une expérience avant leur admission. La variable `Prefiler_ind=0` indique qu'aucune déclaration de revenus n'a été produite avant l'admission, mais elle ne signifie pas que la personne n'a pas d'expérience au Canada avant son admission.

7.2.4 Indicateur de conjoint

La BDIM contient des variables permettant aux utilisateurs de données d'obtenir des renseignements sur l'état matrimonial et les conjoints. La section suivante présente les résultats d'une étude menée sur la BDIM de 2012. Aucune modification majeure n'a été apportée depuis aux codes d'état matrimonial ni à l'indicateur relatif à la famille.

L'**identificateur de conjoint** (`SP__IDI`) est dérivé des fichiers de données fiscales. Ce renseignement peut uniquement être dérivé lorsque le répondant déclare son conjoint ou sa conjointe lors de sa déclaration de revenus; cela entraîne une sous-estimation des couples par rapport à l'état matrimonial déclaré dans les fichiers de données fiscales. Il est également possible à partir du FFT1 d'obtenir l'état matrimonial au moment de la déclaration.

Avant 1991, la catégorie « **célibataire** » n'était pas proposée comme **état matrimonial** (`MSTCO`). La catégorie « **conjoint de fait** » est disponible depuis 1992 pour tous les ensembles de données (1982 à 2012). Depuis 1992, la proportion des enregistrements de la BDIM indiquant un état matrimonial « célibataire » varie entre 20 % et 30 %. La proportion de personnes « **séparées** » a diminué, passant de 30 % avant 1992 à 4 % par la suite. Des modifications de tendance n'ont pas influé sur les autres catégories d'état matrimonial.

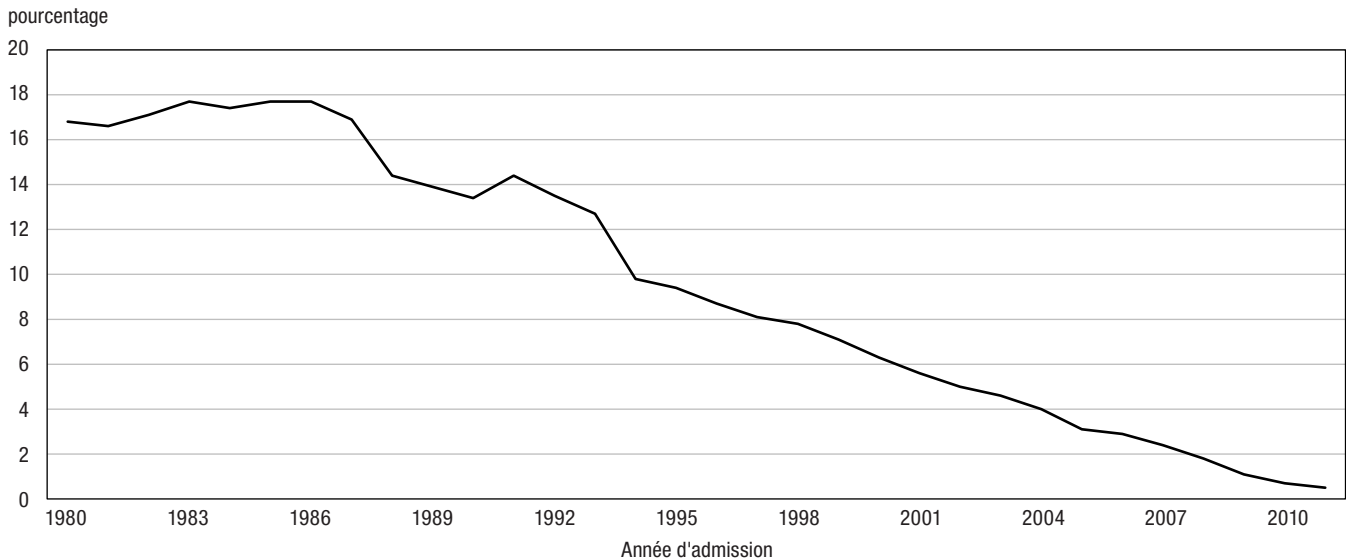
Une analyse portant sur la répartition de l'état matrimonial (`MSTCO` des fichiers de données fiscales) et l'identificateur de conjoint (`SP__IDI`) a révélé des différences entre ces deux variables. Ces différences sont dues à l'absence de valeurs d'état matrimonial dans certains enregistrements. Dans une situation parfaite, les enregistrements de toutes les personnes mariées présenteraient des renseignements sur le conjoint et ceux de toutes les personnes célibataires n'en indiqueraient pas. L'analyse effectuée a indiqué que la qualité des données s'était améliorée après 1992, lorsque les états matrimoniaux distincts « conjoint de fait » et « célibataire » ont été introduits.

Lacunes d'indication de conjoint

À la suite d'un examen de l'historique longitudinal des immigrants sur la BDIM de 2012, on a relevé des cas où le conjoint ou conjoint de fait manquait (ou différait) pour une année donnée et où le même conjoint était déclaré deux ou trois ans plus tard. Le graphique 5 fournit un résumé de ces lacunes.

Graphique 5

Proportion des cas présentant un identificateur de conjoint (SP__IDI) incohérent, par année d'admission



Source : Statistique Canada, Base de données longitudinales sur l'immigration de 2012.

La plupart des immigrants du fichier ont un conjoint ou non entre 1980 et 2012 selon les fichiers IMDB_T1FF. Il faut souligner qu'aucun état matrimonial (et aucun renseignement sur le conjoint) n'est disponible pour 1,2 million d'immigrants sur environ 6 millions d'immigrants.

7.3 Variables imputées

7.3.1 Imputation des variables relatives à la scolarité

Un problème de qualité des données a en outre été relevé quant aux variables relatives au niveau de scolarité et aux années d'études. Une proportion non négligeable de personnes n'ayant pas déclaré leur niveau de scolarité ou d'années d'études ont reçu le code « 0 » soit « aucun » plutôt que « manquant » pour les variables **EDUCATION_QUALIFICATIONS** et **YEARS_OF_SCHOOLING**. Ce problème est devenu courant de 2011 à 2014. En 2011, 35 % des immigrants ont indiqué ne pas avoir suivi d'études, par rapport à environ 10 % dans les années 1990.

Ce problème a été résolu en imputant des variables de scolarité à partir des moyennes des valeurs des variables de scolarité de 2008 à 2010, pour refléter les variables de scolarité de l'année la plus récente. Pour cette imputation, on a utilisé des variables telles que l'âge à l'admission, le groupement de catégories d'immigration 2, la profession envisagée, le sexe et le pays de dernière résidence. La méthodologie d'imputation du plus proche voisin a été employée. La variable d'imputation de la scolarité **Education_imputation_ind** (0: non; 1: oui), disponible dans le FRPN, a été créée pour déterminer les enregistrements au sein desquels les variables de scolarité ont été imputées.

Pour les immigrants admis en 2016, le nombre de cas où une scolarité non-déclarée a reçu le code « 0 » ou « aucun » plutôt que « manquant » a été réduit. Cependant, un nombre non négligeable d'enregistrements avait un niveau de scolarité manquant avec des années d'études valides. Pour ces enregistrements, les années d'études ont été utilisées pour imputer une valeur pour le niveau de scolarité.

Pour les demandeurs principaux admis depuis 2015, sous le processus d'entrée express, les années de scolarisation sont sous-estimées.

Pour la BDIM 2020, les personnes admises entre 2015 et 2018 qui ont été connectées au dossier Express Entry ont vu leur éducation imputée en utilisant les valeurs trouvées dans le fichier Express Entry. Une variable d'éducation plus complète, Education_Derived, a été créée, combinant les données de Education_Qualification pour ceux qui n'ont pas été trouvés dans le fichier Express Entry, et les nouvelles valeurs provenant de Express Entry.

Les utilisateurs devraient noter que les valeurs du fichier d'Express Entry sont basées sur l'évaluation par l'agent d'immigration des qualifications scolaires du demandeur dans le contexte d'une équivalence canadienne, alors que les valeurs de Education_Qualification sont auto-déclarées par le demandeur et ne reflète pas nécessairement l'équivalence canadienne.

Education_Derived a été imputée à 'manquant' pour les personnes admises en 2019 ou plus tard.

7.3.2 Imputation des variables relatives à la langue

Deux variables linguistiques ont été ajoutées à La BDIM, English_IND et French_IND, indiquant la première langue officielle connue à l'admission. Pour ceux qui ont été admis en 2018 ou avant, ils sont définis comme des résidents permanents ayant français (anglais) comme langue maternelle ou ayant une langue maternelle autre que l'anglais ou le français et déclarant le français seulement (anglais seulement) comme la langue officielle connue à l'admission.

Pour les personnes admises en 2019 ou plus tard, elles sont définies comme des résidents permanents ayant déclaré français seulement (anglais seulement) comme langue officielle à l'admission ou ayant déclaré l'anglais et le français comme leur connaissance de la langue officielle à l'admission et déclarant le français (l'anglais) comme la langue dans laquelle ils sont le plus à l'aise.

7.4 Couverture

7.4.1 Couverture du Fichier intégré des résidents permanents et non permanents (FRPN)

Le Fichier intégré des résidents permanents et non permanents (FRPN) de 2020 contient plus de 8,0 millions d'enregistrements (tableau 6); parmi ceux-ci, plus de **7,3 millions (84,9 %)** sont couplés à au moins un fichier de données fiscales. Il faut noter que les données sur l'immigration des non-déclarants et les déclarants sont incluses dans un fichier intitulé PNR_ 1980_2020. Le tableau suivant présente la répartition des enregistrements en fonction de leur présence dans les différents fichiers d'immigration et de données fiscales. Environ **2 million** d'enregistrements portent sur des immigrants qui étaient des résidents temporaires avant de devenir résidents permanents; plus de **1,9 million** de ces enregistrements sont couplés à au moins un fichier de données fiscales. Voir l'annexe B pour des chiffres de répartition détaillés par année d'admission.

Tableau 6
Couverture des résidents permanents

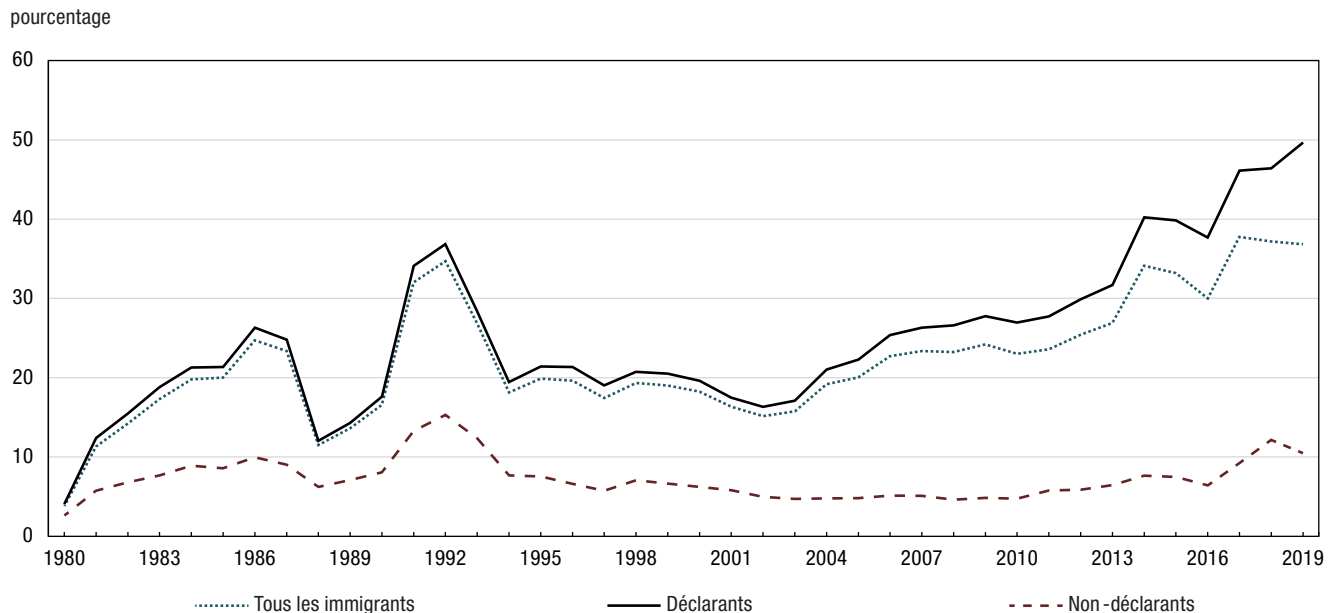
	Résidents permanents	Résidents permanents ayant un permis de résidence non permanente	Nombre de déclarants
	nombre		
Total des déclarants	5 412 220	1 950 160	7 362 380
Total des non-déclarants	1 213 900	97 910	1 311 810
Total	6 626 120	2 048 070	8 674 190
	pourcentage		
Pourcentage de déclarants	81,7	95,2	84,9

Source : Statistique Canada, Base de données longitudinales sur l'immigration de 2020.

Les données relatives aux immigrants possédant des permis de résidence non permanente sont disponibles. La proportion d'immigrants ayant une expérience avant l'admission varie en fonction de l'année d'admission (graphique 6); elle va de **3,8 %** en **1980** à **37,8 %** en **2017**. Par conséquent, la proportion d'immigrants ayant une

expérience avant l'admission au début des années 1980 est sous-représentée. La proportion des immigrants déclarants possédant une expérience avant leur admission (ligne pleine) est supérieure à la proportion globale des immigrants possédant une expérience avant leur admission (ligne pointillée), car le taux de couplage de ces immigrants est supérieur à celui des immigrants sans expérience avant leur admission.

Graphique 6
Répartition des immigrants avec permis de résidence non permanente, par année d'admission



Source : Statistique Canada, Base de données longitudinales sur l'immigration, 2020.

7.4.1.2 Couverture des résidents non permanents

Cette section décrit la couverture des individus qui ont seulement des permis de résidence non permanente depuis 1980, les données fiscales sont disponibles pour 32,0 % de ces personnes. Parmi les individus qui ne sont pas devenus des résidents permanents, les demandeurs d'asile ont le plus haut taux de couverture, des données fiscales sont disponibles pour 46,3 % de cette population (tableau 6.B). Il y a une grande variété de permis de résidence non permanente. Certains sont d'une durée aussi courte qu'un jour.

Tableau 6b
Couverture des résidents non permanent qui n'ont jamais obtenu la résidence permanente par type de permis

	Avec permis de travail	Avec permis d'étude	Demandeur d'asile	Total
	nombre			
Total déclarants	1 512 810	733 810	207 290	1 780 720
Total non déclarants	2 171 200	1 522 830	240 600	3 784 560
Total	3 684 010	2 256 650	447 890	5 565 280
	pourcentage			
Pourcentage de déclarants	41,1	32,5	46,3	32,0

Source : Statistique Canada, Base de données longitudinales sur l'immigration de 2020.

7.4.2 Taille du Fichier sur la famille T1 et couverture par année

Les fichiers de données fiscales sont disponibles à partir de 1982 pour les résidents permanents couplés. Certains résidents permanents étaient des résidents non permanents avant leur admission. Le tableau 7 fournit des détails sur la répartition des résidents permanents couplés, avec ou sans permis de résidence non permanente avant leur admission, par année d'imposition. Au moins un fichier de données fiscales est disponible pour **81,7 %** de résidents permanents sans permis de résidence non permanente avant leur admission et pour

95,2 % de résidents permanents qui étaient résidents non permanents avant leur admission. Le fait que les résidents permanents possédant des permis temporaires avant leur admission présentent un taux plus élevé de déclarations de revenus que les résidents permanents sans permis avant leur admission peut s'expliquer par une exigence relative au processus de demande de résidence permanente pour les résidents non permanents. Les résidents non permanents qui demandent la résidence permanente doivent en effet respecter l'obligation de déclaration de revenus au Canada. Le nombre de déclarants dans le fichier IMDB_T1FF augmente à mesure des années, puisque la taille de la population dans le champ d'enquête augmente.

Tableau 7
Résidents permanents et non permanents par année d'imposition

	Résidents permanents admis avant 1980	Résidents permanents depuis 1980	Résidents permanents ayant un permis non-permanent	Résidents non-permanents uniquement	Nombre de déclarants
	nombre				
1982	1 642 310	188 850	55 630	25 160	1 907 140
1983	1 623 560	224 610	65 560	23 300	1 932 230
1984	1 619 210	263 840	79 950	23 660	1 981 630
1985	1 600 350	298 100	94 900	22 360	2 010 490
1986	1 655 910	356 240	124 900	26 370	2 157 440
1987	1 636 020	415 890	158 070	26 550	2 229 950
1988	1 656 440	508 220	199 540	35 570	2 392 190
1989	1 683 790	622 130	262 500	47 930	2 607 710
1990	1 692 930	743 550	309 820	51 270	2 788 060
1991	1 690 080	841 220	358 670	51 040	2 930 980
1992	1 696 540	947 330	402 590	50 530	3 086 500
1993	1 729 010	1 091 400	441 990	50 480	3 301 590
1994	1 713 720	1 212 810	467 540	49 620	3 432 190
1995	1 697 670	1 324 970	492 780	51 980	3 555 740
1996	1 678 700	1 432 520	513 580	53 900	3 667 000
1997	1 654 210	1 546 910	534 080	55 910	3 779 450
1998	1 630 200	1 646 770	554 200	54 820	3 874 360
1999	1 623 770	1 771 140	590 030	59 450	4 032 740
2000	1 606 040	1 914 750	630 540	67 330	4 207 120
2001	1 592 600	2 071 730	679 610	77 440	4 409 790
2002	1 562 620	2 201 810	717 330	83 100	4 553 370
2003	1 542 640	2 324 170	754 520	86 190	4 696 100
2004	1 526 590	2 455 490	796 490	87 700	4 854 930
2005	1 501 430	2 571 730	830 300	93 050	4 985 330
2006	1 483 690	2 721 280	887 720	97 970	5 179 540
2007	1 464 850	2 844 960	956 800	109 940	5 365 540
2008	1 444 590	2 970 160	1 035 250	132 300	5 571 420
2009	1 424 130	3 085 680	1 102 640	142 250	5 743 950
2010	1 399 320	3 211 760	1 160 310	151 290	5 912 060
2011	1 380 260	3 342 540	1 226 990	159 050	6 098 310
2012	1 353 600	3 461 020	1 296 570	170 260	6 271 070
2013	1 335 160	3 594 130	1 373 870	190 610	6 483 500
2014	1 313 450	3 713 650	1 454 690	213 250	6 684 940
2015	1 286 100	3 841 070	1 516 220	235 650	6 869 180
2016	1 257 870	3 962 770	1 574 600	281 850	7 067 410
2017	1 231 890	4 066 420	1 624 280	402 480	7 315 600
2018	1 208 530	4 206 820	1 658 620	593 650	7 658 370
2019	1 169 610	4 339 640	1 667 300	832 600	8 000 150
Total des déclarants	2 059 950	5 412 220	1 950 160	1 774 990	...
Total des non-déclarants	2 053 730	1 213 900	97 910	3 621 700	...
	pourcentage				
Pourcentage de déclarants	50,1	81,7	95,2	32,9	...

... n'ayant pas lieu de figurer

Note : Les résidents permanents depuis 1980 excluent les cohortes d'immigrants après 2019.

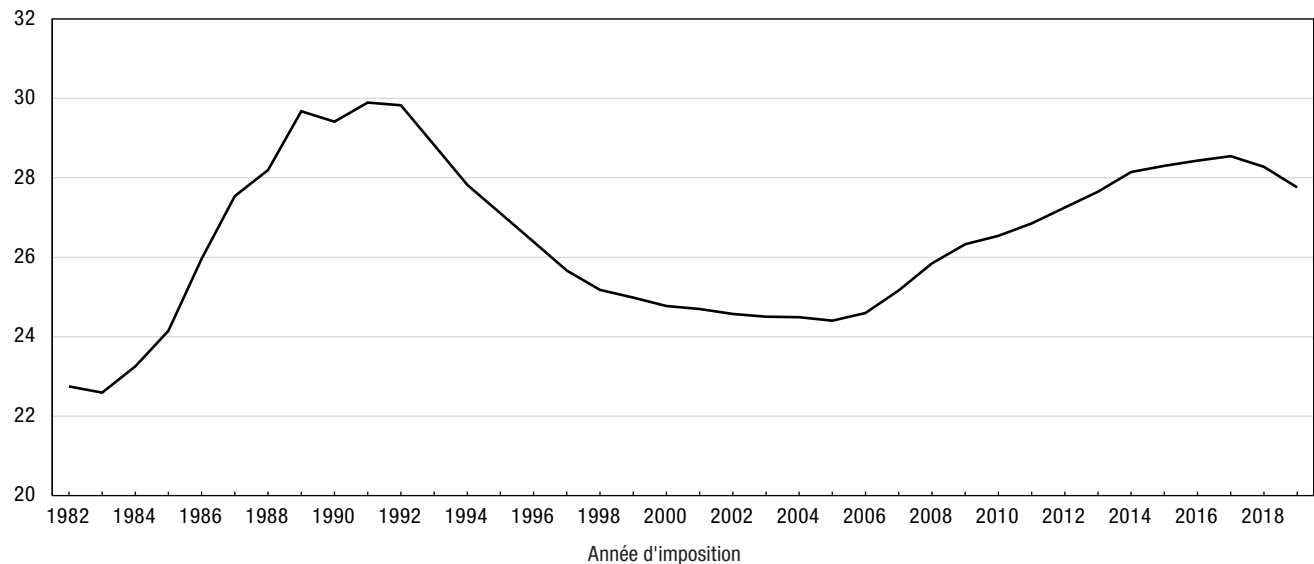
Les statistiques sur les résidents non permanents portent sur les personnes qui ont obtenu leur premier permis entre 1980 et 2019.

Source : Statistique Canada, Base de données longitudinales sur l'immigration 2020.

Le graphique 7 montre que la proportion des résidents permanents qui étaient résidents non permanents avant leur admission, par année, varie de 22,6 %, le plus bas niveau pour l'année d'imposition 1983, à 29,9 %, niveau le plus élevé correspondant à l'année d'imposition de 1991. Depuis les années 2000s, cette proportion est demeurée stable, à environ 26,4 %.

Graphique 7**Pourcentage de résidents permanents qui étaient résidents non permanents avant leur admission, par année d'imposition**

pourcentage



Source : Statistique Canada, Base de données longitudinales sur l'immigration, 2020.

Un immigrant qui produit une déclaration de revenus pour une année donnée n'en produira pas nécessairement l'année suivante. Si une personne A, par exemple, a été admise en 1983, elle peut figurer dans les fichiers de données fiscales de 1984 à 1999, mais pas dans le fichier de 2000, et y figurer ensuite à nouveau de 2001 à 2013. Ainsi, des fichiers de données fiscales étaient disponibles tous les ans pour **34,7 %** des déclarants de la cohorte de 1980. Une émigration externe, un décès ou un retard de déclaration peut expliquer, en outre, que des immigrants déclarants aient pu ne plus produire de déclaration de revenus ou ne pas en produire certaines années.

La plupart des immigrants déclarent leurs revenus pour la première fois l'année de leur admission ou l'année précédente ou suivante. Ainsi, sur les **251 100** immigrants admis en 2006, **100 900 (40,2 %)** ont produit une déclaration de revenus pour la première fois en 2006, alors que **15 570 (6,2 %)** l'ont fait en 2007 et **3 200 (1,3 %)** en 2015.

7.5 Évaluation de la qualité de la données sur l'immigration

7.5.1 Évaluation de la qualité du Fichier intégré des résidents permanents et non permanents (FRPN)

Le contenu du FRPN (PNRF_1980_2020) a été évalué. Alors que la collecte des données fiscales et relatives à l'admission est obligatoire pour les personnes dans le champ d'enquête, ces dernières peuvent ne pas remplir tous les champs. Ces champs ont pu être laissés vides parce que le répondant ne connaissait pas la réponse ou pour d'autres raisons, que les utilisateurs des données ne peuvent connaître (p. ex. refus) (McLeish, 2011). L'absence de réponse à certaines questions peut présenter des problèmes lorsque la BDIM est utilisée à des fins statistiques, notamment dans les cas suivants :

1. Si l'utilisateur de la base de données souhaite produire un échantillon fondé sur les caractéristiques pour lesquelles des enregistrements manquent, une erreur de couverture peut survenir (c.-à-d. les personnes incluses dans la base d'échantillon peuvent ne pas être représentatives de la population cible).
2. S'il s'agit d'une absence de données informatives (p. ex. l'absence du renseignement n'est pas un cas aléatoire; l'absence de réponse indique ce que la réponse aurait été), toute analyse utilisant ces variables serait biaisée.

On a évalué la présence de variables manquantes et de valeurs non valides. Les chiffres présentés dans cette section sont arrondis. Les valeurs non valides sont soit incohérentes soit elles ne figurent pas dans les tableaux de métadonnées à la disposition des utilisateurs (voir la composante sur l'immigration du dictionnaire de données en annexe). La plupart des problèmes de qualité figurant dans le tableau 8 concernent des données recueillies dans les années 1980 et 1990. Il convient de souligner que des valeurs semblant valides peuvent également être erronées.

La variable de **numéro d'identification du cas** (CASE_ID) présente des taux de réponse aux éléments se situant généralement dans la frange supérieure des 90 % (généralement plus de 99 %). Cependant, pour certaines années d'admission, le taux de réponse diminue significativement (jusqu'à 80 % en 1991 et 1992). Par conséquent, toute analyse fondée sur cette variable pour toutes les années d'admission sous-représentera ces années, pour lesquelles le manque de réponse à l'élément est plus élevé (p. ex. 1986, 1987, 1990, 1991, 1992, 1993). Aucune détection des valeurs non valides n'a eu lieu pour la variable de numéro d'identification du cas (CASE_ID).

La variable d'âge à l'admission **Landing_age** est définie comme non valide lorsqu'elle dépasse 100, même s'il est possible dans certains cas que ces valeurs soient exactes. Il convient de souligner que, selon les valeurs de cette variable, le nombre d'immigrants admis après l'âge de 100 ans était bien plus élevé entre 1986 et 1994 que pour les autres années d'admission. Il pourrait s'agir d'un problème de saisie de données.

Dans le FRPN de la BDIM de **2020**, 25 enregistrements présentaient une **année de naissance** antérieure à 1880 et 18 avaient une année de naissance de 1753 avec une année d'admission après 1985, même jusqu'en 2012.

Les variables relatives au pays présentent également des problèmes de qualité. Le **pays de naissance** manque dans certains enregistrements pour pratiquement toutes les années d'admission. Des valeurs manquent, par exemple, pour plus de 100 enregistrements dans chacune des années de 1985 à 1993. Le **pays de citoyenneté** est manquant pour moins de 20 enregistrements par année d'admission pour la plupart des cohortes (à l'exception des cohortes de 2005, 2006, 2016 et 2017, où plus de 80 valeurs étaient manquantes par année d'admission). Le **pays de résidence** manque pour de nombreux enregistrements d'admission pour 2013 (cette valeur manque pour 1190 enregistrements, soit 0,5 % des admissions de cette année-là), pour 2014 (cette valeur manque pour 5845 enregistrements, soit 2,3 % des admissions de cette année-là) et pour 2015 (manquant pour 7360 enregistrements, soit 2,8 % des admissions de cette année-là).

Les **variables de scolarité** avant la cohorte de 2017, après imputation (voir la section 6.3), présentent plus de 150 valeurs manquantes par année d'admission de 1980 à 1984; cela équivaut à un taux de valeurs manquantes par année d'admission de moins de 0,5 %. Une nouvelle variable a été dérivée en utilisant les données d'entrée express pour imputer le niveau de scolarité des personnes admises entre 2015 et 2018.

Le pourcentage de réponses valides pour les **variables de profession** est plus de 99 % pour toutes les années d'admission.

Les variables de statut dans la famille **Family_Status** et, d'immigrants sélectionnés par le Québec **CSQ_IND** ont la plupart de leurs valeurs manquants pour les enregistrements des années d'admission antérieures à 1999.

La variable de langue maternelle **Mother_Tongue** manque pour 550 enregistrements relatifs aux admissions de 2011.

La langue officielle **Official_Language** a un nombre croissant de valeurs manquantes ; 2016 à 2020, entre 1810 et 13045 par cohorte ont une valeur manquante.

La variable de l'état matrimonial **Marital_Status** a eu plus de 200 valeurs manquantes par cohorte depuis 2012.

Les variables **Destination_CD**, **Destination_CMA**, et **Destination_CSD** ont moins de valeurs manquantes dans la BDIM **2020** que lors des années précédentes suite à la mise à jour avec la Classification Géographique Type (CGT) de 2016.

Les variables d'**année et de mois de décès** manquent pour certaines personnes reconnues comme décédées (Death_Indicator=1). La valeur « 9999 » a été attribuée à la variable d'année de décès Death_Year et la valeur « 99 » à celle du mois de décès Death_Month dans les cas où l'année et le mois de décès étaient inconnus.

Tableau 8
Évaluation de la qualité du Fichier intégré des résidents permanents et non permanents

Variables du PNRF	Réponses valides		Champs vides		Réponses non valides	
	nombre	pourcentage	nombre	pourcentage	nombre	pourcentage
Case_ID	8 629 920	97,43	227 940	2,57	0	0,00
Landing_age	8 856 420	99,98	720	0,01	720	0,01
Birth_Year	8 857 810	100,00	30	0,00	30	0,00
Gender	8 857 860	100,00	0	0,00	0	0,00
Country_Birth	8 854 740	99,96	3 120	0,04	0	0,00
Country_Citizenship	8 856 930	99,99	940	0,01	0	0,00
Country_Residence	8 841 460	99,81	16 410	0,19	0	0,00
Education_Qualification	8 472 550	95,65	385 310	4,35	0	0,00
Level_of_Education	8 794 500	99,28	63 360	0,72	0	0,00
Years_of_Schooling	8 783 590	99,16	74 280	0,84	0	0,00
Landing_age_6_groups	8 217 210	92,77	640 650	7,23	0	0,00
Landing_age_9_groups	8 857 140	99,99	720	0,01	0	0,00
Occupation_CD	8 857 140	99,99	720	0,01	0	0,00
NOC5-NOC2	8 851 190	99,92	6 670	0,08	0	0,00
Family_Status	8 851 190	99,92	6 670	0,08	0	0,00
Family_Status_rollup	8 851 140	99,92	6 720	0,08	0	0,00
Marital_status	8 855 270	99,97	2 590	0,03	0	0,00
Marital_status_rollup	8 855 270	99,97	2 590	0,03	0	0,00
Mother_Tongue	8 853 050	99,95	4 820	0,05	0	0,00
Official_Language	8 853 050	99,95	4 820	0,05	0	0,00
Skill_level_CD11	8 855 110	99,97	2 760	0,03	0	0,00
Special_Program	8 816 570	99,53	41 290	0,47	0	0,00
CSQ_ind	1 390 050	15,69	7 467 820	84,31	0	0,00
Destination_CD	8 857 640	100,00	230	0,00	0	0,00
Destination_CMA	8 857 510	100,00	350	0,00	0	0,00
Destination_CSD	8 857 510	100,00	350	0,00	0	0,00
Destination_ER	8 857 510	100,00	350	0,00	0	0,00
Destination_Province	8 857 520	100,00	350	0,00	0	0,00
Permits and NPR-specific variables	2 047 724	100,00	0	0,00	0	0,00
Death_Year	8 857 370	99,99	490	0,01	0	0,00
Death_Month	8 857 350	99,99	510	0,01	0	0,00

Notes : FRPN: Fichier intégré des résidents permanents et non permanents. RNP: Résident non permanent. Seules les variables ayant des valeurs manquantes ou invalides ont été incluses dans ce tableau. Tous les chiffres sont arrondis.

Source : Statistique Canada, Base de données longitudinales sur l'immigration 2020.

7.5.2 Évaluation de la qualité du fichier des résidents non permanents (NRF)

Une validation du contenu des fichiers NRF_PERMIT_1980_2020 et NRF_PERSON_1980_2020 a été effectuée. Ces fichiers contiennent des ensembles de variables différents les uns des autres. Dans le tableau 8b, les variables “Landing_Year” à “Number_All_Permits) apparaissent dans le fichier des personnes, tandis que les autres apparaissent dans le fichier des permis. Alors que la collecte des données fiscales et relatives à l'admission est obligatoire pour les personnes dans le champ d'enquête, ces dernières peuvent ne pas remplir tous les champs. Ces champs ont pu être laissés vides parce que le répondant ne connaissait pas la réponse ou pour d'autres raisons, que les utilisateurs des données ne peuvent connaître (p. ex. refus) (McLeish 2011). L'absence de réponse à certaines questions peut présenter des problèmes lorsque la BDIM est utilisée à des fins statistiques, notamment dans les cas suivants :

1. Si l'utilisateur de la base de données souhaite produire un échantillon fondé sur les caractéristiques pour lesquelles des enregistrements manquent, une erreur de couverture peut survenir (c.-à-d. les personnes incluses dans la base d'échantillon peuvent ne pas être représentatives de la population cible).
2. S'il s'agit d'une absence de données informatives (p. ex. l'absence du renseignement n'est pas un cas aléatoire; l'absence de réponse indique ce que la réponse aurait été), toute analyse utilisant ces variables serait biaisée.

La présence de variables manquantes et de valeurs non valides a été évaluée. Les chiffres présentés dans cette section sont arrondis. Les valeurs non valides sont soit incohérentes, soit non répertoriées dans les tableaux de métadonnées mis à la disposition des utilisateurs (voir la composante immigration de l'annexe du dictionnaire des données). Il convient de noter que certaines valeurs apparemment valides peuvent également être erronées.

La variable `Landing_year` présente un pourcentage élevé de valeurs manquantes (72,27 %). Ceci est normal puisque seuls les immigrants admis ont une année d'admission et que la NRF inclut tous les résidents non permanents, qu'ils soient devenus résidents permanents ou non.

Dans le fichier `NRF_PERSON` de la BDIM 2020, 350 enregistrements ont une année de naissance antérieure à 1880, et 331 enregistrements ont une année de naissance antérieure à 1753.

Alors que la plupart des enregistrements ont un pays de naissance (`Country_Birth`), ceux dont le pays de naissance est manquant ont également une année d'admission manquante.

Les variables `Effective_Date` et `Valid_Date` n'ont pas de réponses invalides en elles-mêmes, mais lorsqu'elles sont comparées ensemble, elles peuvent produire des réponses invalides. Par exemple, la `Valid_Date` doit toujours arriver après la `Effective_Date`. Les enregistrements dont la date de validité apparaît avant la date d'entrée en vigueur peuvent être considérés comme des réponses non valides pour l'une de ces variables ou les deux. De plus, tout enregistrement dont la durée entre `Effective_Date` et la `Valid_Date` est de 5 ans ou plus peut être considéré comme suspect et comprenant probablement une valeur non valide pour une des deux variables. 0,05 % des comparaisons entre les variables `Effective_Date` et `Valid_Date` pourraient être considérées comme non valides en raison de ces deux problèmes.

Plus de 99 % des valeurs manquantes de la variable `Valid_Date` se produisent lorsque la variable `Document_Type` est 46 (demande d'asile), ce qui s'explique par le fait qu'aucune date de fin n'est attribuée aux demandes d'asile.

Les variables `Destination_CD`, `Destination_ER`, `Destination_CMA`, `Destination_CSD` et `Destination_province` présentent une proportion de valeurs manquantes plus faible que les autres variables, mais beaucoup plus importante que le dans le fichier `PNRF`. La plupart des années (sur la base de la variable `Effective_Date`) ont un taux de variables manquantes très faible, autour de 1%. Cependant, les années 1989, 2018 et 2019 ont un taux de variables manquantes de près de 12 %, et les années 2014, 2015, 2016 et 2017 ont un taux de variables manquantes compris entre 32 % et 37 %. La BDIM 2020 utilise la Classification géographique standard (CGS) pour mettre à jour la région et le code géographiques.

Tableau 8b
Évaluation de la qualité des Fichiers résidents non permanents

Variables du NRF	Réponses valides		Champs vides		Réponses non valides	
	nombre	pourcentage	nombre	pourcentage	nombre	pourcentage
Landing Year	2 135 030	27,73	5 565 280	72,27	0	0,00
birth_year	7 698 670	99,98	1 280	0,02	350	0,00
birth_month	7 698 970	99,98	1 330	0,02	0	0,00
gender	7 700 300	100,00	0	0,00	0	0,00
COUNTRY_BIRTH	7 680 380	99,74	19 920	0,26	0	0,00
NUMBER_OTHER_PERMITS	7 700 300	100,00	0	0,00	0	0,00
NUMBER_REFUGEE_CLAIMS	7 700 300	100,00	0	0,00	0	0,00
NUMBER_WORK_PERMITS	7 700 300	100,00	0	0,00	0	0,00
NUMBER_STUDY_PERMITS	7 700 300	100,00	0	0,00	0	0,00
NUMBER_ALL_PERMITS	7 700 300	100,00	0	0,00	0	0,00
COUNTRY_RESIDENCE	18 156 760	96,59	641 360	3,41	0	0,00
COUNTRY_CITIZENSHIP	18 741 590	99,70	56 530	0,30	0	0,00
LEVEL_OF_STUDY_ROLLUP	6 196 150	32,96	12 601 970	67,04	0	0,00
LEVEL_OF_STUDY	6 196 150	32,96	12 601 970	67,04	0	0,00
SKILL_LEVEL_CD11	12 051 230	64,11	6 746 890	35,89	0	0,00
OCCUPATION_CD	12 058 430	64,15	6 739 680	35,85	0	0,00
NOC5_CD11	12 051 230	64,11	6 746 890	35,89	0	0,00
NOC4_CD11	12 051 230	64,11	6 746 890	35,89	0	0,00
NOC3_CD11	12 051 230	64,11	6 746 890	35,89	0	0,00
NOC2_CD11	12 051 230	64,11	6 746 890	35,89	0	0,00
DESTINATION_CSD	17 375 250	92,43	1 422 860	7,57	0	0,00
DESTINATION_CMA	17 375 250	92,43	1 422 860	7,57	0	0,00
DESTINATION_PROVINCE	17 375 280	92,43	1 422 840	7,57	0	0,00
DESTINATION_CD	17 375 250	92,43	1 422 860	7,57	0	0,00
DESTINATION_ER	17 375 250	92,43	1 422 860	7,57	0	0,00
effective_date	18 798 120	100,00	0	0,00	0*	0,00
valid_date	17 863 380	95,03	934 730	4,97	0*	0,00
DOCUMENT_TYPE	18 798 120	100,00	0	0,00	0	0,00
SPECIAL_PROGRAM	3 102 770	16,51	15 695 340	83,49	0	0,00
CLASSIFICATION_ID	6 688 900	35,58	12 109 210	64,42	0	0,00
LMIA_EXEMPTIONS	7 894 700	42,00	10 903 420	58,00	0	0,00

* valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,05$)

Notes : NPR : résident non permanent. Seules les variables dont les valeurs sont manquantes ou invalides ont été incluses dans le tableau. Tous les chiffres sont arrondis. Les variables Effective_Date et Valid_Date peuvent être invalides lorsqu'elles sont comparées entre elles. Voir le paragraphe de la section 7.5.2 pour plus de détails.

Source : Statistique Canada, Base de données longitudinales sur l'immigration 2020.

7.6 Évaluation de la qualité de la variable de la province de résidence (PRCO_)

Une validation des variables géographiques incluses dans les fichiers fiscaux de la BDIM a été effectuée. Cette section explique comment la variable Province de résidence (PRCO_) a été dérivée et sa qualité.

La province de résidence (PRCO_) est fondée sur les renseignements fournis par les déclarants, lorsque disponible. Les renseignements manquants pour la province de résidence sont remplacés par les renseignements recueillis sur le code postal de l'adresse postale de l'individu (PSCO_I), si disponible, sinon celui de la famille est utilisé (PSCO_F).

Tableau 9
Concordance entre le PRCO et le PSCO__

PRCO	Province et Territoires	Premier caractère du code postal (PSCO)
0	Terre-Neuve-et-Labrador	A
2	Île-du-Prince-Édouard	B
1	Nouvelle-Écosse	C
3	Nouveau-Brunswick	E
4	Québec	G, H, J
5	Ontario	K, L, M, N, P
6	Manitoba	R
7	Saskatchewan	S
8	Alberta	T
9	Colombie-Britannique	V
10	Territoires du Nord-Ouest	X
11	Yukon	Y
12	Non-résidents manquant	manquant
14	Nunavut	X

Note: Les valeurs de certain codes postales sont U ou F pour les valeurs manquantes, respectivement États-Unis est U et, étranger est marqué par F.

Bien que la province de résidence (PRCO_) soit plus fiable que la province d'imposition (TXPCO_), certaines anomalies ont été observées principalement sur le code des non-résidents dans les déclarations des années d'imposition 1989, 1993 et 1998. Celles-ci peuvent avoir une incidence sur certaines provinces.

Pour l'année d'imposition 1993, la IMDB_T1FF comprend des anomalies pour la province du Manitoba qui compte un nombre inhabituel de résidents (48 130 en 1993, comparativement à 33 650 l'année d'imposition précédente et à 37 365 l'année d'imposition suivante). Des changements semblables sont observés dans les Territoires du Nord-Ouest. De plus, 740 personnes sont codées comme résidant au Nunavut alors que le Nunavut a été créé en 1998. 725 personnes sont codées comme résidant dans plusieurs juridictions. Les utilisateurs peuvent utiliser l'information de la variable PSCO_F pour diminuer l'effet des anomalies sur leurs analyses qui incluent la province de résidence. Toutefois, comme nous l'avons indiqué plus haut, le moment est différent entre PSCO (basé sur le lieu de résidence lors de la transmission de la déclaration) et PRCO_ (lieu de résidence au 31 décembre).

Les enregistrements des non-résidents (PRCO_=12) semblent surestimés dans le fichier IMDB_T1FF de 1989. Il comprend 79 210 non-résidents du Canada, dont bon nombre ont un statut de résident non permanent. Les utilisateurs peuvent décider d'utiliser le code postal de l'adresse postale (PSCO_ au niveau individuel ou familial) pour calculer la valeur du PRCO_ ou retirer les non-résidents de leur analyse.

Dans le fichier FFT1 de la BDIM de 1998, IMDB_T1FF_1998, un nombre plus élevé que prévu d'enregistrements est attribué à Terre-Neuve-et-Labrador (PRCO_). Dans ces cas, le lieu de résidence de la famille au moment du dépôt de la demande est également Terre-Neuve selon la variable PSCO_F.

8 Comparabilité

8.1 Variation historique de la couverture

Au cours des années, la couverture et le contenu de la BDIM ont évolué. Les fichiers *IMDB_T1FF* initiaux comprenaient uniquement des données sur les immigrants admis au Canada à partir de 1980. Depuis la parution de la BDIM de 2013, les résidents non permanents déclarants ont été ajoutés aux fichiers *IMDB_T1FF* pour les années d'imposition de 1982 et ultérieures. Du fait de ce changement, on peut désormais disposer de renseignements sur les permis de résidence temporaire des immigrants ayant une expérience au Canada avant leur admission.

En 2012, la BDIM a fait l'objet d'une refonte. La couverture de la BDIM a été modifiée pour inclure à la base de données des immigrants ayant obtenu leur statut d'immigrant reçu à partir de 1980 et ayant produit au moins une déclaration de revenus depuis 1982, qu'ils en aient produit une ou non après leur admission. La BDIM incluait initialement uniquement des personnes ayant obtenu le statut d'immigrant reçu à partir de 1980 et ayant produit au moins une déclaration de revenus après avoir été reçues comme immigrants. Avant ce cycle, la BDIM incluait jusqu'aux seize premières années des fichiers de données fiscales relatifs à un résident permanent donné (Dryburgh, 2004). Cette limite du nombre de fichiers de données fiscales pour une personne donnée ne s'applique plus.

Les données fiscales incluses dans la BDIM provenaient à l'origine principalement des formulaires T1 et seulement quelques variables fiscales importantes au niveau personnel étaient conservées. Dans la BDIM de 2006 et les BDIM suivantes, on a utilisé les fichiers FFT1 relatifs à 1982 et aux années ultérieures; cela a produit un taux de couplage initial de 80 %. À partir de là, la BDIM a exclu les fichiers de données fiscales de 1980 et 1981, puisque les renseignements pour ces années ne figurent pas dans le FFT1.

On utilisait initialement le Système de soutien des opérations des bureaux locaux (SSOBL) pour rassembler les données d'immigration incluses dans la BDIM. Pour l'année d'immigration de 2013 et les suivantes, le système mondial de gestion des cas (SMGC) sera utilisé. IRCC a, par conséquent, cessé de fournir certaines variables. Ces anciennes variables seront disponibles dans le fichier *PNRF_EXTRA*; elles figurent dans la composante sur l'immigration du dictionnaire des données de la BDIM.

Avec la version de 2018 de la BDIM, de nouveaux changements ont été apportés en ce qui concerne les personnes intégrées dans la BDIM. Premièrement, il y a eu un changement dans la structure des fichiers : les déclarants et les non-déclarants ont fusionné. De plus, il y a eu un changement de couverture : L'univers de la BDIM s'est élargi pour inclure tous les immigrants depuis 1952 et tous les résidents non permanents depuis 1980. Les fichiers fiscaux des résidents non permanents (temporaires) pour les visas d'étudiant et les permis de travail sont maintenant aussi inclus dans les *IMDB_T1FFs*.

8.2 Changements méthodologiques

La méthodologie employée pour effectuer le couplage d'enregistrements a fait l'objet de changements au cours des années.

Le taux de couplage initial de la BDIM était de 55 % pour la BDIM de 1995 (Langlois et Dougherty, 1997). Les méthodes et outils employés pour le couplage d'enregistrements ont depuis évolué; ce qui explique l'amélioration des taux de couplage au cours des années.

À la fin des années 2000, le taux de couplage était d'environ 81 %. Pour la BDIM de 2012, les renseignements sur les personnes à charge ont servi pour procéder au couplage d'enregistrements; cela a permis de coupler une plus grande proportion d'enfants immigrants. Ces renseignements figuraient dans le fichier de la prestation fiscale canadienne pour enfants (PFCE). Il convient de noter que l'ajout des enfants n'améliore pas le taux de déclarants. Grâce aux changements méthodologiques, le taux de couplage pour la BDIM de 2014 a atteint 89 %.

Pour la BDIM de 2015, en incluant dans le processus de couplage le Registre d'assurance sociale (RAS) – une base de données spécifiquement pour les données de NAS – le taux de couplage a augmenté à 97 %. Le Registre d'assurance sociale fournit des données de très haute qualité, et environ 730 000 numéros d'assurance sociale se trouvent exclusivement sur ce registre (Diaz-Papkovich 2017).

Pour la BDIM 2016, un nouveau processus de couplage d'enregistrements a été utilisé afin de faciliter le couplage avec d'autres sources de données. Les données d'immigration ont été couplées aux données fiscales par l'intermédiaire de l'ECDS (voir la section 4). À partir de maintenant, le couplage se fera au Dépôt d'enregistrements dérivés de Statistique Canada.

Pour la BDIM 2018, afin d'améliorer les résultats du couplage d'enregistrements, une combinaison de L'Environnement de couplage de données sociales (ECDS) et du Fichier de contrôle des couplages (FCC) a été employé comme pour l'année 2015.

8.3 Changements du contenu de la base de données historiques

Veillez-vous reporter aux dictionnaires de la BDIM (composantes sur l'immigration et les impôts) pour une description complète du contenu des fichiers. Certaines importantes modifications récentes du contenu de la BDIM sont énumérées ci-après.

En 2012, la BDIM de 2009 a fait l'objet d'un remaniement; un indicateur a en effet été ajouté pour détecter les valeurs aberrantes dans les fichiers FFT1 créés. La variable d'identificateur de conjoint (SP_IDI) a été introduite dans la BDIM de 2010, permettant de recenser les immigrants ayant un conjoint immigrant. Les années et mois de décès ont été ajoutés à la BDIM de 2013; cela a permis de recenser les immigrants admis au Canada à partir de 1980 et décédés depuis. À la suite de l'ajout à la BDIM des données relatives aux résidents non permanents, certains détails sur les permis de résidence temporaire (type, dates d'entrée en vigueur, etc.) sont désormais disponibles depuis la BDIM de 2013.

Pour la BDIM de 2016, des drapeaux identifiant les immigrants sélectionnés selon l'entrée express, les vagues de réfugiés syriens rétablis, ainsi que l'année et le mois de citoyenneté ont été ajoutés.

Pour la BDIM de 2018, il y a eu un changement de couverture. La BDIM s'est élargie pour inclure les résidents permanents admis de 1952 à 1979, ainsi que les fichiers fiscaux de tous les résidents non permanents depuis 1980. En outre, plusieurs modules de données ont été ajoutés à la BDIM : sur les enfants, l'établissement, les salaires, ainsi que des détails sur l'entrée express.

8.4 Changements du contenu de la base de données historiques

La BDIM est l'un des nombreux programmes statistiques pouvant servir à produire des estimations relatives à la population immigrante. Dans certains cas, ces estimations différeront du fait de plusieurs facteurs, comme la couverture et des limites liées au type de données (données administratives, données d'enquête, données de recensement). Certains des programmes statistiques et différences relatives à la BDIM sont décrits dans la présente section. On utilise la BDIM de 2013 pour procéder à des comparaisons.

8.4.1 Banque de données administratives longitudinales (DAL)

La Banque de données administratives longitudinales¹⁴ (DAL) est un échantillon longitudinal de 20 % des déclarants canadiens. Elle est couplée à la BDIM pour inclure un échantillon de 20 % des enregistrements de la BDIM et pour ajouter des variables propres aux immigrants, comme l'année d'admission, la catégorie d'immigration et l'état matrimonial à l'admission. Elle contient des renseignements sur les particuliers et sur les familles de recensement. Elle est utile aux analyses longitudinales, qui comparent le revenu et la mobilité des immigrants à ceux des déclarants canadiens. Toute analyse comparant des immigrants déclarants à la population de déclarants canadiens devrait se servir de cet ensemble de données.

La DAL contient un moins grand nombre de variables relatives à l'immigration que la BDIM. Les renseignements portant sur la période précédant l'admission, par exemple (comme le nombre de permis de travail et d'études), ne figurent pas dans la DAL. Les renseignements sur l'admission, notamment la profession envisagée et la province de destination, ne figurent pas non plus dans la DAL.

14. Pour de plus amples détails sur la DAL, reportez-vous à la description figurant sur le [site Web de Statistique Canada](#) ou renseignez-vous sur un rapport technique détaillé.

Le tableau 10 présente le revenu total moyen et médian (XTIRC) pour l'année d'imposition de 2012 des immigrants admis entre 1982 et 2013, par sexe, afin d'illustrer le degré de comparaison des estimations produites à partir de ces bases de données. Le revenu total moyen et médian par sexe, comme nous nous y attendons, est similaire pour les deux sources de données. Les différences peuvent s'expliquer par le fait que la DAL est un échantillon de 20 % de la population canadienne et que la BDIM est un recensement des immigrants déclarants couplés admis au Canada depuis 1980. Les chiffres de population diffèrent, mais ni l'une ni l'autre de ces sources ne devrait être utilisée pour dénombrer la population; la DAL étant un échantillon de déclarants et la BDIM étant limitée aux immigrants déclarants. La population de la DAL est estimée en multipliant les enregistrements par un facteur de pondération de 5.

Tableau 10
Comparabilité du revenu total de 2012 entre la DAL et la BDIM pour les immigrants admis entre 1982 et 2013

	Hommes			Femmes			Total		
	Population	Moyenne	Médiane	Population	Moyenne	Médiane	Population	Moyenne	Médiane
	nombre	dollars		nombre	dollars		nombre	dollars	
Particulier									
BDIM	2 776 700	41 900	29 400	2 906 000	28 700	20 600	5 682 690	35 000	24 200
DAL	2 686 300	41 700	29 200	2 803 100	28 700	20 500	5 489 390	34 900	24 100
Famille									
BDIM	...	73 700	56 300	...	69 900	51 400	...	71 700	53 700
DAL	...	73 800	56 500	...	69 700	51 500	...	71 700	53 900

... n'ayant pas lieu de figurer

Note : BDIM : Base de données longitudinales sur l'immigration; DAL : Banque de données administratives longitudinales.

Source : Statistique Canada, Base de données longitudinales sur l'immigration de 2013 et Banque de données administratives longitudinales de 2013.

Dans le tableau 11, la comparabilité a été limitée au revenu total pour 2012 des immigrants admis en 2011. Les différences estimées observées entre la BDIM et la DAL pour ce groupe sont supérieures à celles observées pour la population immigrante admise pour toute année entre 1980 et 2013. Cela peut s'expliquer par le fait que la population d'intérêt est moins importante et plus précise. Les estimations de la DAL sont dérivées des enregistrements inclus dans l'échantillon de 20 % des immigrants admis en 2011. Ces enregistrements ne correspondent pas toujours au 20 % de la population précise de la BDIM. Ils constituent probablement une proportion plus réduite de la population précise de la BDIM, puisque l'échantillon n'a pas été prélevé pour être représentatif de cette population particulière. Les estimations de la BDIM sont dérivées de la population immigrante couplée admise en 2011 et ayant produit une déclaration de revenus en 2012. Ainsi, les estimations de la DAL peuvent prendre des valeurs légèrement différentes de celles de la BDIM lors de l'examen de sous-ensembles de population.

Tableau 11
Comparabilité du revenu total moyen et médian pour 2012 des immigrants admis en 2011

	Hommes		Femmes		Total	
	Moyenne	Médiane	Moyenne	Médiane	Moyenne	Médiane
	dollars					
Particulier						
BDIM	30 100	22 400	18 900	14 100	24 300	17 800
DAL	29 500	22 100	18 700	13 900	23 900	17 500
Famille						
BDIM	49 900	39 300	48 200	37 100	49 000	38 200
DAL	49 300	39 000	48 000	36 600	48 600	37 800

Note : BDIM : Base de données longitudinales sur l'immigration; DAL : Banque de données administratives longitudinales.

Source : Statistique Canada, Base de données longitudinales sur l'immigration de 2013 et Banque de données administratives longitudinales de 2013.

8.4.2 Recensement

Le questionnaire détaillé du recensement et l'Enquête nationale auprès des ménages (ENM) de 2011 permettent de recueillir des données relatives aux immigrants. Ces données sont recueillies pour une portion de la population (reportez-vous à la description du programme du recensement pour connaître la proportion exacte; cette valeur a varié au cours du temps). On y recueille ainsi le lieu de naissance, le lieu de naissance des parents, le statut d'immigrant, l'année d'immigration, l'âge à l'immigration et la citoyenneté. Depuis le Recensement de 2016, la catégorie d'immigration est également disponible. Le recensement recueille des données sur les Canadiens de la première, deuxième et plus anciennes générations, alors que la BDIM recueille uniquement des données sur les nouveaux arrivants et leur famille. Le recensement comprend également des données sur les minorités visibles, la scolarité, le logement et la langue pour l'année de recensement; à moins que l'année d'admission ne soit également une année de recensement, ce dernier ne présente pas ces renseignements pour l'année d'admission. Le recensement ne permet pas d'effectuer d'études longitudinales sur la situation économique ou la mobilité à long terme des immigrants. De plus amples détails sur le programme du recensement figurent sur le [site Web de Statistique Canada](#). Il y a aussi un rapport technique pour obtenir plus d'informations.

L'Enquête nationale auprès des ménages (ENM) de 2011 a permis d'estimer que plus de 4,6 millions d'immigrants vivant au Canada en 2011 avaient été admis entre 1981 et 2011. Le tableau 11 compare les estimations de la population immigrante par décennie d'admission selon l'ENM et le FRPN. Le FRPN de 2013 ne devrait pas être utilisé pour estimer les chiffres de population, même après suppression des enregistrements des personnes décédées. Le faire contribuerait à une surestimation de la population immigrante vivant au Canada admise entre 1981 et 2011, car le FRPN ne tient pas compte de l'émigration. De plus, le FRPN est un sous-ensemble de la population immigrante, puisque seuls les déclarants y sont inclus. Cela peut expliquer les chiffres de population plus faibles pour le FRPN que pour l'ENM pour la cohorte d'immigrants la plus récente (de 2001 à 2011). Le nombre d'immigrants au tableau 11 prend en considération le nombre de décès en utilisant la variable Death_ indicator (décrit à la section 7.2.2).

Tableau 12

Comparabilité des estimations de population entre la Base de données longitudinales sur l'immigration et l'Enquête nationale auprès des ménages

Décennie d'admission	Estimations de l'ENM	Estimations du FRPN pour 2013
	nombre	
1981 à 1990	949 890	1 052 650
1991 à 2000	1 539 055	1 896 235
2001 à 2011	2 154 985	2 120 290
Total	4 643 930	5 069 175

Note : BDIM : Base de données longitudinales sur l'immigration; ENM : Enquête nationale auprès des ménages; FRPN : Fichier intégré des immigrants et résidents non permanents.

Source : Statistique Canada, Base de données longitudinales sur l'immigration de 2013 et Enquête nationale auprès des ménages de 2011.

8.4.3 Enquête longitudinale auprès des immigrants de Canada (ELIC)

L'Enquête longitudinale auprès des immigrants du Canada (ELIC) vise à fournir des renseignements sur la manière dont les nouveaux immigrants s'adaptent à la vie au Canada et à comprendre les facteurs pouvant faciliter ou empêcher cette adaptation. L'ELIC est conçue pour examiner les quatre premières années de l'établissement. On a recueilli en trois vagues les données relatives à des immigrants de 15 ans et plus admis au Canada entre le 1^{er} octobre 2000 et le 30 septembre 2001. L'ELIC permet d'étudier les compétences linguistiques, le logement, la scolarité, la reconnaissance des diplômes étrangers, l'emploi, la santé, les valeurs et les attitudes, l'établissement et l'utilisation des réseaux sociaux, le revenu et les perceptions relatives à l'établissement au Canada. La BDIM contient des caractéristiques comme la scolarité et la langue uniquement à l'admission, alors que l'ELIC permet une évaluation de l'évolution au cours du temps. De plus amples détails sur l'ELIC figurent sur le [site Web de Statistique Canada](#).

Selon l'ELIC, il a été estimé que 164 200 immigrants âgés de 15 ans et plus ont été admis au Canada entre le 1^{er} octobre 2000 et le 30 septembre 2001. L'estimation de cette même population est 156 670 selon la BDIM de 2013 en utilisant des calculs basés sur le FRPN (tableau 12). Une partie de la différence est due à la fois à l'exclusion des non-déclarants de l'estimation du FRPN et à l'émigration non considéré par la BDIM. Une autre partie s'explique par le fait que l'ELIC est une enquête introduisant des estimations de variance. Comme le montre le tableau 13, les proportions de la couverture selon la tranche d'âge varient en fonction de la tranche d'âge, même si la population de l'ELIC est en âge de déclarer des revenus. Il convient de noter que l'âge dans l'ELIC est celui environ six mois après l'admission, alors que l'âge dans la BDIM est celui à l'admission. De plus, le calcul des estimations de l'ELIC est fondé sur les facteurs de pondération de la première vague, conçus pour estimer le nombre d'immigrants dans cette cohorte vivant toujours au Canada six mois après leur admission. La proportion inférieure d'immigrants âgés de 65 ans et plus pourrait être due à une proportion inférieure de déclarants pour cette tranche d'âge. Le nombre d'immigrants âgés de 15 à 24 ans plus élevé dans la BDIM que dans l'ELIC s'explique probablement par la non-prise en compte de l'émigration.

Tableau 13**Répartition par sexe : comparaison de la Base de données longitudinales sur l'immigration et de l'Enquête longitudinale auprès des immigrants du Canada**

	ELIC de 2003		FRPN de 2013	
	nombre	pourcentage	nombre	pourcentage
Hommes	81 550	49,7	77 640	49,6
Femmes	82 650	50,3	78 830	50,4
Total	164 200	100,0	156 470	100,0

Note : ELIC : Enquête longitudinale auprès des immigrants du Canada; FRPN : Fichier intégré des résidents permanents et non permanents.

Source : Statistique Canada, Base de données longitudinales sur l'immigration de 2013 et Enquête longitudinale auprès des immigrants du Canada, 1^{re} vague, 2003.

Tableau 14**Répartition par tranche d'âge : comparaison de la Base de données longitudinales sur l'immigration et de l'Enquête longitudinale auprès des immigrants du Canada**

Tranche d'âge	ELIC de 2003		FRPN de 2013	
	nombre	pourcentage	nombre	pourcentage
15 à 24 ans	26 730	16,3	27 990	17,9
25 à 34 ans	65 500	39,9	63 050	40,3
35 à 49 ans	53 970	32,9	49 030	31,3
50 à 64 ans	12 890	7,8	12 280	7,8
65 ans ou plus	5 100	3,1	4 120	2,6
Total	164 200	100,0	156 470	100,0

Note : ELIC : Enquête longitudinale auprès des immigrants du Canada; FRPN : Fichier intégré des résidents permanents et non permanents.

Source : Statistique Canada, Base de données longitudinales sur l'immigration de 2013 et Enquête longitudinale auprès des immigrants du Canada, 1^{re} vague, 2003.

8.5 Discussion sur la BDIM avec différents couplages

Afin d'améliorer la capacité d'analyse de la BDIM, plusieurs sources de données ont été intégrées, notamment le Recensement, l'Enquête sur la santé dans les collectivités canadiennes (ESCC), Métadonnées de la Base de données sur les congés des patients (BDPCP), l'Enquête sociale générale (ESG), la Banque de données administratives longitudinales (DAL) et l'Enquête longitudinale auprès des immigrants du Canada (ELIC). Vous trouverez ci-dessous un bref aperçu de chacun d'eux.

8.5.1 Recensement

Réalisé tous les cinq ans, le Recensement de la population est la principale source de données sociodémographiques pour des groupes de population particuliers comme les familles monoparentales, les peuples autochtones, les immigrants, les aînés et les groupes linguistiques. Les chiffres de population rajustés tirés du recensement servent de base au Programme d'estimation de la population.

Le recensement est présenté sous forme de deux questionnaires, le questionnaire abrégé et le questionnaire détaillé. Le formulaire abrégé sert à énumérer tous les résidents habituels de tous les logements privés du Recensement de 2016 et les résidents qui sont à l'étranger (en 2016, cela comprenait les employés du gouvernement canadien (fédéral et provinciaux) et leurs familles, et les membres des Forces canadiennes et leurs familles). Il contient des questions sur des renseignements démographiques de base, comme l'âge, le sexe, la connaissance des langues officielles, la composition du ménage, et plus encore.

En 2016, un échantillon de 25% des ménages canadiens a reçu un questionnaire détaillé. Il contient des sujets allant du niveau d'éducation, des limitations d'activités, de l'origine ethnique, et plus. Les données sur le revenu ont été tirées des fichiers d'impôt sur le revenu des particuliers et de prestations. D'autres données sur l'immigration concernant la catégorie d'admission ont été obtenues à partir des fichiers administratifs d'Immigration, Réfugiés et Citoyenneté Canada.

Le Recensement fait l'objet d'un processus complexe de conception de la base de sondage et de l'échantillon, de collecte, de codage, de vérification et d'imputation, et de certification avant diffusion. Pour de plus amples renseignements sur tout aspect du Recensement de 2016, veuillez consulter le site Web de la [Guide du Recensement de la Population, 2016](#), ou [les produits de référence disponibles](#).

8.5.2 Enquête sur la santé dans les collectivités canadiennes (ESCC)

L'Enquête sur la santé dans les collectivités canadiennes est un projet conjoint de Statistique Canada et de Santé Canada. La composante annuelle de l'Enquête sur la santé dans les collectivités canadiennes (ESCC) recueille des données **transversales** sur la santé, les comportements liés à la santé et l'utilisation des soins de santé chez les ménages non institutionnalisés âgés de 12 ans ou plus.

L'enquête exclut les membres à temps plein des Forces canadiennes, les résidents des réserves et de certaines régions éloignées, qui représentent ensemble environ 4 % de la population cible. L'ESCC a été menée pour la première fois en 2001 (cycle 1.1) et a été répétée tous les deux ans jusqu'en 2005 (cycle 3.1), chaque fois avec un échantillon d'environ 130,000 personnes. À partir de 2007, l'enquête a été menée annuellement (taille de l'échantillon de 65 000 personnes). Les taux de réponse variaient de 69,8 % à 78,9 %. Des détails sur la stratégie et le contenu de l'échantillonnage sont disponibles dans le guide de l'utilisateur de l'ESCC et dans la documentation des données, qui sont disponibles auprès de votre analyste de CDR.

Les enquêtes sur le contenu de l'ESCC sont conçues pour fournir des résultats transversaux à l'échelle provinciale sur des sujets de santé précis et ciblés. Deux cycles de contenu ont été utilisés dans le cadre de ce projet de couplage. L'ESCC sur la santé mentale et le bien-être (2002 et 2012) a recueilli de l'information sur les troubles mentaux, l'utilisation du système de santé mentale et l'incapacité associée aux problèmes de santé mentale dans la population à domicile âgée de 15 ans et plus. Un rapport technique détaillé est disponible pour obtenir plus d'informations.

Pour plus d'informations sur l'ESCC, veuillez consulter le site internet de Statistique Canada.

8.5.3 Métadonnées de la Base de données sur les congés des patients (BDCP)

La Base de données sur les congés des patients (BDCP) est une base de données nationale qui recueille des renseignements administratifs, cliniques et démographiques sur toutes les sorties des établissements de soins actifs, y compris les congés, les décès, les sorties et les transferts, au cours d'une **année fiscale** (1^{er} avril au 31 mars). Avec le temps, la BDCP a été étendue à la saisie de données sur les interventions chirurgicales d'un jour, la réadaptation, les soins de longue durée et d'autres types de soins. Notez que la BDCP est basée sur des événements, ce qui signifie qu'il y aura plus d'un dossier pour une personne hospitalisée plus d'une fois au cours **d'une année fiscale**. Les exigences en matière de collecte varient selon l'année de données et la province ou le territoire.

Plus de 3,2 millions de résumés sont soumis chaque année à la BDCP, ce qui représente environ 75 % de toutes les sorties de patients hospitalisés en phase aiguë au Canada. Le Québec ne soumet pas de données à la BDCP ; les sorties de patients hospitalisés en phase aiguë au Québec sont déclarées dans la Base de données sur la morbidité hospitalière (BDMH) et représentent habituellement 25 % du total des sorties de patients hospitalisés au Canada. Environ 2,4 millions de résumés de chirurgie d'un jour sont soumis à l'Institut canadien d'information sur la santé (ICIS) chaque année ; environ 35 % sont envoyés à la BDCP et 65 % sont envoyés au Système national d'information sur les soins ambulatoires (SNISA).

La population de référence comprend habituellement toutes les séparations de soins actifs aux malades hospitalisés et de chirurgie d'un jour au Canada (à l'exclusion des morts infantiles et des cas de donneurs cadavériques) du 1^{er} avril au 31 mars. Toutes les données sur les soins de courte durée, sauf celles du Québec, sont soumises à la BDCP ; les données sur les soins de courte durée du Québec sont soumises par l'entremise du ministère de la Santé et des Services sociaux du Québec une fois par année et sont incluses dans la BDMH. Les données sur les chirurgies d'un jour de l'Ontario, de l'Alberta et de la Nouvelle-Écosse sont soumises à la SNISA.

La population de référence comprend habituellement toutes les sorties des établissements de soins actifs pour malades hospitalisés et de chirurgie d'un jour au Canada (à l'exclusion des morts à la naissance et des cas de donneurs cadavériques) du 1^{er} avril au 31 mars. Toutes les données sur les soins de courte durée, sauf celles du Québec, sont soumises à la BDCP ; les données sur les soins de courte durée du Québec sont soumises par l'entremise du ministère de la Santé et des Services sociaux du Québec une fois par année et sont incluses dans la BDMH. Les données sur les chirurgies d'un jour de l'Ontario, de l'Alberta et de la Nouvelle-Écosse sont soumises au SNISA.

Pour de plus amples renseignements sur la BDCP, veuillez consulter le [site web ICIS](#). Il y a un rapport technique disponible pour plus d'informations.

8.5.4 Enquête sociale générale (ESG)

L'ESG (Enquête sociale générale du Canada) est une enquête transversale annuelle indépendante qui examine en profondeur un sujet afin de suivre l'évolution des conditions de vie. L'ESG est une mesure du bien-être des Canadiens et peut fournir de l'information sur des questions précises de politique sociale. Chaque enquête recueille des données sociodémographiques approfondies sur l'âge, le sexe, l'éducation, la religion, l'origine ethnique, le revenu, etc.

L'ESG est un examen exhaustif de plusieurs sujets essentiels, notamment les familles, les soins, l'emploi du temps, la victimisation, le bénévolat, etc. Chacun des six thèmes du sondage est répété de façon exhaustive environ tous les cinq ans.

Jusqu'en 1998, la taille de l'échantillon était fixée à 10.000. Il a augmenté en 1999 pour atteindre un objectif de 25.000 personnes. La plus grande taille de l'échantillon permet d'obtenir les estimations de base à l'échelle provinciale, nationale et de certaines régions métropolitaines de recensement (RMR).

Pour plus d'informations sur l'[Enquête sociale générale](#) veuillez consulter le site internet de Statistique Canada.

8.5.5 Banque de données administratives longitudinales (DAL)

La DAL est un échantillon aléatoire de 20% de la base de données fiscales du Fichier de familles T1 (FFT1). La sélection pour la DAL est fondée sur le NAS d'une personne. Il n'y a pas de limite d'âge, mais les personnes sans NAS ne peuvent être incluses que dans la composante familiale. Une fois qu'une personne est choisie pour la DAL, elle demeure dans l'échantillon et est sélectionnée chaque année de la FFT1 si elle apparaît sur la T1 cette année-là. Les personnes sélectionnées pour la DAL sont reliées d'une année à l'autre par un numéro d'identification non confidentiel unique de la DAL (LIN__I) généré à partir du NAS, afin de créer un profil longitudinal de chaque personne.

La DAL est augmentée chaque année d'un échantillon de nouveaux déclarants, de sorte qu'elle se compose d'environ 20 % des déclarants pour chaque année. L'échantillon de 2 0% est passé de 3.227.485 personnes en 1982 à 5.579.280 en 2016 (soit une augmentation de 73 %). Cette augmentation reflète l'accroissement de la population canadienne et l'augmentation de la fréquence des déclarations de revenus par suite de l'instauration du crédit de taxe de vente fédérale en 1986 et du crédit pour taxe sur les produits et services en 1989.

Pour de plus amples renseignements sur la DAL, veuillez consulter le site Web de la [Banque de données administratives longitudinales \(DAL\)](#) ou le [Dictionnaire des données administratives longitudinales](#)

8.5.6 Enquête longitudinale auprès des immigrants du Canada (ELIC)

L'Enquête longitudinale auprès des immigrants du Canada (ELIC) sert à recueillir de l'information pour mieux comprendre la vie des immigrants récents au Canada. L'ELIC est conçu pour saisir les quatre premières années de leur établissement au Canada, en fournissant des indicateurs sur la façon dont les immigrants relèvent des défis tels que la connaissance ou la maîtrise de l'une des deux langues officielles du Canada, la participation au marché du travail, l'accès à la formation ou les études. Les quatre premières années sont une période où les immigrants tissent des liens économiques, sociaux et culturels avec le Canada.

L'objectif de l'enquête est d'étudier la vie des nouveaux immigrants au Canada, leur adaptation au fil du temps et de voir ce qui facilite et entrave leur intégration dans la société canadienne.

La population cible de l'enquête est constituée d'immigrants qui doivent satisfaire aux exigences suivantes :

- Doit avoir été âgé de 15 ans et plus au moment de l'admission.
- Doivent être arrivés au Canada entre le 1^{er} octobre 2000 et le 30 septembre 2001.
- Avoir été admis à l'étranger et avoir présenté une demande par l'entremise d'une mission canadienne à l'étranger.

La population d'intérêt de l'ELIC est constituée des immigrants qui vivaient encore au Canada au moment de l'entrevue. L'enquête se déroule en trois vagues, soit trois questionnaires distincts qui ont chacun été soumis à un processus d'essai rigoureux.

Pour de plus amples renseignements sur l'ELIC, veuillez consulter le site Web [Enquête longitudinale auprès des immigrants du Canada](#).

9 Analyses possibles grâce à la BDIM

La BDIM a été créée pour permettre d'effectuer une analyse de sujets reliés à l'immigration; cette section présente un aperçu des analyses possibles grâce aux renseignements supplémentaires désormais disponibles dans la BDIM. Comme le décrit le présent rapport, le contenu de la BDIM a évolué; cela a accru ses capacités analytiques. Trouvez ci-après des exemples d'analyses que la BDIM permet d'effectuer.

9.1 Possibilités analytiques découlant des données relatives aux résidents non permanents

L'ajout des données relatives aux résidents non permanents élargit la portée de l'analyse que permet actuellement la BDIM. Les renseignements de nombre et de type de permis obtenus avant l'admission peuvent permettre d'établir le profil pré migratoire des immigrants. En comparant ces populations (avec ou sans expérience au Canada avant l'admission), il est possible d'évaluer les répercussions de l'expérience au Canada avant l'admission sur la situation économique et sur les tendances de mobilité des immigrants. Le profil socioéconomique particulier au moment de la délivrance du permis de résidence temporaire permet aussi d'évaluer la situation économique et la mobilité avant l'admission. L'évolution de la profession envisagée, du niveau de compétence et du niveau de scolarité en fonction de la délivrance de permis de résidence temporaire est également disponible.

9.2 Possibilités analytiques grâce aux données relatives aux décès

L'ajout au FRPN des variables d'indicateur, de mois et d'année de décès permet d'estimer la proportion d'enregistrements inclus dans la BDIM qui concernent des immigrants décédés. La variable d'année de décès (YOD) est uniquement présente dans les cas où un formulaire T1 a été rempli de façon posthume au nom de la personne décédée, alors que l'année et le mois du décès figurent dans tout enregistrement couplé à l'ensemble de données sur la mortalité, quel que soit le profil de déclaration de revenus. De nouvelles analyses possibles peuvent inclure l'évaluation de profils économiques avant le décès et l'étude des attentes relatives à la vie après l'admission, par catégorie d'immigration et profil économique.

9.3 Capacités analytiques avec les données sur la citoyenneté

L'ajout de renseignements sur la citoyenneté à la BDIM offre plus de possibilités d'analyse. L'addition de l'année et du mois de citoyenneté canadienne permet de savoir si et quand les immigrants ont obtenu leur citoyenneté. Elle sert également de variable explicative supplémentaire pour l'étude des résultats socioéconomiques ou de la démographie. Par exemple, l'obtention de la citoyenneté donne une idée du taux d'obtention au fil du temps et des caractéristiques associées à l'obtention de la citoyenneté. Veuillez noter que les données la citoyenneté sont disponibles à partir de 2005.

9.4 Capacités analytiques avec les données sur les enfants

L'ajout du module sur les enfants à la BDIM permet d'approfondir l'analyse des données sur les conditions socioéconomiques des enfants immigrants durant leur enfance. Les enfants (personnes de moins de 18 ans) représentent environ 25 % des immigrants admis au Canada depuis 1980. Les enfants immigrants font face à des difficultés et à des défis différents de ceux de leurs parents ou tuteurs. Ce module permet d'analyser les résultats économiques des enfants d'immigrants en fonction de leur statut économique pendant l'enfance. Les enfants immigrants issus de familles à faible revenu et leurs résultats économiques futurs par rapport aux enfants immigrants issus de familles à revenu élevé. Des données sont également disponibles pour les enfants réfugiés syriens.

9.5 Capacités analytiques avec les données sur l'entrée express

L'entrée express est un processus de demande pour les immigrants économiques qui veulent s'établir au Canada de façon permanente et prendre part à notre économie. L'ajout de l'entrée express augmente la capacité d'analyse en ce qui concerne les caractéristiques précises de sélection et les résultats économiques qui s'en suivent.

Par exemple, en 2017, les immigrants admis en 2015 à titre de travailleurs qualifiés (fédéral) ont déclaré les salaires les plus élevés parmi les immigrants admis dans le cadre du système d'EE. Parmi leurs caractéristiques nous pouvons vérifier s'ils avaient une offre d'emploi au moment de leur application.

9.6 Capacités analytiques avec les données des fichiers sur les salaires et traitements

Les **fichiers préliminaires des salaires et traitements** fournissent la rémunération et les avantages imposables versés aux employés. Les variables extraites de ces fichiers incluent la province de l'emploi, la province de l'employé(e), les revenus d'emploi pour l'année de d'imposition, et le nombre de feuillet T4.

L'ajout de ces données permet l'analyse d'immigrants et surtout de résidents non permanents qui n'auraient pas rempli de rapport d'impôt T1. De plus, cette information est fournie plus tôt que celle des rapports annuels d'impôt, permettant d'obtenir une information plus à jour sur la situation économique ou l'emplacement géographique des résidents permanents et non permanents. Finalement, la présence de plusieurs feuillets T4 informe aussi de la stabilité en emploi.

10 Résumé

La BDIM est un ensemble de données combinant des enregistrements sur l'immigration et les impôts à des fins d'analyses socioéconomiques et relatives à la mobilité des immigrants (avec ou sans expérience avant l'admission) admis au Canada depuis 1952 ou plus tard ainsi que les résidents non permanents depuis 1980. La BDIM permet d'effectuer des analyses sur les déclarants et les non déclarants. Les déclarants peuvent présenter un profil différent des non-déclarants. Le présent rapport technique a été élaboré pour fournir une description complète de la qualité des données de la BDIM ainsi que des modifications récemment apportées à cette base de données.

Avec la sortie de la BDIM 2018, le contenu s'est enrichi avec l'intégration de nouveaux modules. Ces modules comprennent l'établissement, les données sur les enfants immigrants et les salaires. De plus, la BDIM comprend également des données d'entrée express, ainsi que l'acquisition de la citoyenneté depuis 2005.

Il y a eu un changement dans la structure des fichiers - les déclarants et les non-déclarants ont fusionné. Dans le passé, c'était PNRF_YEAR pour les déclarants et PNRF_NONFILERS_YEAR pour les non-déclarants ; maintenant, le fichier fusionné s'appelle PNRF_1980_2020.

Comment accéder à la BDIM :

Comme le décrit la section 6, plusieurs produits sont à la disposition des chercheurs. Ils sont accessibles sur le site Web de Statistique Canada en sélectionnant « Diversité ethnique et immigration » comme sujet, puis [Base de données longitudinales sur l'immigration \(BDIM\)](#) sous Produits vedettes.

Il est à noter que les mises à jour annuelles de la BDIM sont indépendantes les unes des autres. D'année en année, il y a des changements aux processus de création des données, incluant à l'Identificateur personnel unique (IMDB_ID).

Annexes

A. Liens vers les principaux documents et pages Web relatifs à la BDIM

Dictionnaires (composantes sur les impôts et l'immigration) :

Disponible aux utilisateurs de données ou sur demande en communiquant avec Statistique Canada par courriel à l'adresse STATCAN.infostats-infostats.STATCAN@statcan.gc.ca

[Portail sur les statistiques sur les immigrants et résidents non permanents](#) : Le portail sur les immigrants et résidents non permanents regroupe les données les plus demandées, des outils et des rapports sur une seule page.

Historique de la [BDIM](#) :

Publications de la BDIM dans [Le Quotidien](#) :

[Analyses](#) fondées sur la [BDIM](#) :

Cornelissen, L. 2021. [Portrait des personnes immigrantes qui exercent une profession infirmière ou d'aide aux soins de santé](#). Statistique Canada : Regards sur la société canadienne.

Evra, R. et Kazemipur, A. 2019. [Rôle du capital social et des caractéristiques ethnoculturelles dans le revenu d'emploi des immigrants au fil du temps](#). *Statistique Canada* : Regards sur la société canadienne.

Hou, F., Crossman, E. et Picot, G. 2020. [Sélection des immigrants en deux étapes : l'expérience de travail spécialisé par rapport à un emploi réservé](#). Statistique Canada : Aperçus économiques.

Huystee, M. 2016. [Mobilité interprovinciale : Taux de rétention et taux d'entrée nets immigrants admis de 2008 à 2013](#). IRCC Rapports de recherche.

Ng, E. et Zhang, H. 2020. [La santé mentale des immigrants et des réfugiés : données canadiennes provenant d'une base de données couplée au niveau national](#). Statistique Canada : Rapports sur la santé.

Ng, E et al. 2019. [Utilisation des services hospitaliers liée à la tuberculose chez les nouveaux immigrants au Canada](#). Statistique Canada: Rapports sur la santé.

Picot, G, et Lu, Y. 2017. [Faible revenue chronique chez les immigrants au Canada et dans ses collectivités](#). Statistique Canada.

[Indice des prix à la consommation](#) (62-001-X).

[Description des estimations annuelles du revenu des familles de recensement et des particuliers](#) (Fichier sur la famille T1) :

B. Couverture

La BDIM de 2020 a été utilisée pour produire ces chiffres. Les déclarants sont des immigrants couplés ayant produit au moins une déclaration de revenus depuis 1982. Les statistiques au-dessous exclues les admissions de 2019.

Tableau 15
Répartition des déclarants et des non déclarants par année d'admission

	Déclarants ¹	Non-déclarants	Total	
	Immigrants	Immigrants	Immigrants	Déclarants
	nombre		pourcentage	
1980	120 450	22 670	143 120	84,2
1981	107 710	20 870	128 580	83,8
1982	103 440	17 660	121 090	85,4
1983	77 110	11 920	89 030	86,6
1984	77 500	10 530	88 030	88,0
1985	75 070	8 870	83 940	89,4
1986	89 070	9 710	98 770	90,2
1987	137 540	13 630	151 170	91,0
1988	146 250	14 500	160 760	91,0
1989	173 910	16 750	190 660	91,2
1990	192 260	23 170	215 430	89,2
1991	208 560	23 260	231 820	90,0
1992	228 650	25 280	253 930	90,0
1993	230 890	24 790	255 680	90,3
1994	198 860	24 740	223 600	88,9
1995	188 890	23 270	212 150	89,0
1996	198 740	26 630	225 360	88,2
1997	189 630	25 820	215 460	88,0
1998	155 630	18 060	173 680	89,6
1999	169 150	20 220	189 370	89,3
2000	203 960	22 790	226 750	90,0
2001	225 220	24 550	249 770	90,2
2002	204 750	23 460	228 210	89,7
2003	196 710	23 820	220 530	89,2
2004	208 470	26 880	235 350	88,6
2005	228 450	33 330	261 780	87,3
2006	217 980	33 130	251 110	86,8
2007	203 120	33 060	236 180	86,0
2008	208 900	37 710	246 610	84,7
2009	212 670	38 920	251 590	84,5
2010	230 300	49 770	280 070	82,2
2011	201 580	46 540	248 120	81,2
2012	209 320	47 920	257 240	81,4
2013	209 420	49 080	258 500	81,0
2014	210 940	48 560	259 510	81,3
2015	215 470	55 530	271 000	79,5
2016	222 860	72 530	295 390	75,4
2017	220 760	64 570	285 320	77,4
2018	233 690	86 100	319 790	73,1
2019	228 510	111 240	339 750	67,3
Total	7 362 380	1 311 810	8 674 190	84,9

1. Les déclarants sont des immigrants qui ont produit des déclaration d'impôt au moins une fois depuis 1982.

Note : Tous les chiffres sont arrondis.

Source : Statistique Canada, Base de données longitudinales sur l'immigration 2020.

Tableau 16
Proportion de déclarants couplés par tranche d'âge à l'admission, sexe et décennie d'admission

Sexe et cohortes	Âge à l'admission						Total
	0 à 14 ans	15 à 24 ans	25 à 34 ans	35 à 49 ans	50 à 64 ans	65 ans et plus	
	pourcentage						
Cohortes de 1980 à 1989							
Hommes	82,7	93,9	95,0	93,5	84,8	62,6	89,4
Femmes	81,6	91,5	93,4	92,6	81,5	60,5	87,2
Total	82,2	92,7	94,2	93,0	82,9	61,4	88,3
Cohortes de 1990 à 1999							
Hommes	82,0	93,0	93,0	92,7	90,0	76,8	89,7
Femmes	80,2	91,5	92,7	92,8	88,1	75,5	88,9
Total	81,1	92,1	92,9	92,8	89,0	76,0	89,3
Cohortes de 2000 à 2009							
Hommes	69,8	93,1	91,9	92,8	93,1	89,3	87,3
Femmes	68,5	92,5	92,9	93,7	92,8	87,6	88,0
Total	69,2	92,8	92,4	93,3	92,9	88,3	87,6
Cohortes de 2010 à 2018							
Hommes	17,1	90,5	93,7	93,2	90,0	82,0	76,6
Femmes	17,0	91,2	93,8	94,2	88,9	81,5	78,5
Total	17,1	90,9	93,8	93,7	89,4	81,7	77,6

Source : Statistique Canada, Base de données longitudinales sur l'immigration de 2020.

C. Analyses précédentes

Depuis sa création, la BDIM a permis de produire plusieurs analyses. Trouvez ci-après un résumé de certaines des études de Statistique Canada fondées sur la BDIM.

Ces dernières années, plusieurs publications du *Quotidien* se sont basées sur la BDIM. Les sujets discutés ont inclus l'évolution de la répartition régionale des nouveaux immigrants au Canada, le revenu et la mobilité des immigrants, les immigrants dans les régions et les immigrants qui quittent le Canada. Ces articles sont accessibles sur le site Web de Statistique Canada.

Les séries de publications *L'emploi et le revenu en perspective* (75-001-X) et de [documents de recherche de la Direction des études analytiques](#) ont proposé plusieurs documents fondés sur la BDIM. Parmi les sujets couverts, on peut citer le revenu des immigrants qui entreprennent des études postsecondaires au Canada et les avantages en matière de revenu des immigrants admis ayant été auparavant résidents temporaires au Canada.

D. Pratiques exemplaires et conseils pour les analystes

D.1 Conseils de programmation

Cette section offre des renseignements relatifs à la programmation pour les personnes souhaitant mieux comprendre la structure de programmation utilisée pour accéder aux données des fichiers de la BDIM. Toute personne peut utiliser sa propre programmation. Il existe deux types de fichiers au sein de la BDIM : les fichiers de données annuelles de la BDIM et les données sur l'immigration (pour de plus amples détails sur les fichiers constituant la BDIM, reportez-vous à la section 3). Le nom des variables relatives à l'impôt dans la BDIM consiste en trois parties : 1) le sigle indiqué dans le dictionnaire sur les données fiscales de la BDIM, 2) le niveau d'agrégation (I ou F) et 3) l'année (l'extension de l'année à quatre chiffres existe dans la plupart des cas, mais pas tous).

Exemple : La variable de revenus d'intérêts et d'investissements au niveau individuel pour 2014 serait intitulée INVI_I2014.

Les observations contenues dans les fichiers de la BDIM sont triées en fonction d'une variable, **IMDB_ID** (qui ne comprend pas d'extension d'année), permettant aux utilisateurs de maintenir un lien entre les années. L'accès aux données est rendu possible par le langage de programmation SAS. Un exemple de programme SAS permettant d'accéder aux données de la BDIM est fourni ci-dessous. Les exemples ci-dessous permettent d'effectuer la tâche suivante :

« obtenir le nombre de **bénéficiaires d'aide sociale (AS)** parmi les immigrants admis entre 2007 et 2012, vivant en Ontario entre 2015 et 2017 et pour lesquels aucun revenu ne figurait sur leur feuillet T4, par sexe et année (de 2015 à 2017) ».

Nous encourageons les chercheurs utilisant nouvellement la BDIM à passer cet exemple de programme SAS en revue. L'exemple comprend généralement trois composantes :

1. la configuration de la bibliothèque : les attributions de bibliothèque sur les deux premières lignes sont les emplacements des fichiers d'entrée (première ligne) et de sortie (deuxième ligne);
2. les étapes permettant de générer un ensemble de données de travail :
 - a. Les fichiers d'entrée étant au format SAS, ils sont accessibles au moyen d'une instruction **SET** ou **MERGE**.
 - b. Ce programme vise à obtenir le nombre de **bénéficiaires d'aide sociale (AS)** parmi les immigrants :
 - i. admis entre 2007 et 2012;
 - ii. vivant en Ontario de 2015 à 2017;
 - iii. dont le feuillet T4 n'indiquait aucun revenu;et génère le nombre de bénéficiaires d'aide sociale, par **sexe et année** (dans ce cas, de 2015 à 2017);
3. l'ensemble de données utilisé pour produire le nombre de bénéficiaires d'aide sociale. Cette partie, qui commence par « proc freq », produit les chiffres souhaités, tels qu'ils sont précisés dans le reste du document. À la fin du programme, quatre tableaux sont générés à partir du fichier de données créé.

Il est généralement recommandé que les programmes utilisent des variables figurant dans le FRPN plutôt que dans les fichiers de données fiscales annuelles, par souci de cohérence. L'exemple de programme utilise, par exemple, la variable de groupement par sexe (**GENDER**), variable figurant dans le **FRPN (PNRF)**, plutôt que **SXCO_I&YEAR**, variable figurant dans le fichier **IMDB_T1FF** annuel. Dans ce programme, seules les personnes ayant rempli une déclaration chaque année entre 2015 et 2017 sont sélectionnées.

Lors de toute programmation dans SAS, il est important de distinguer les valeurs manquantes des zéros présents dans les champs numériques. Dans SAS, la plupart des opérations mathématiques effectuées avec des valeurs manquantes produiront des valeurs manquantes. Dans la BDIM, pour les années au cours desquelles une personne est présente, les variables numériques qui ne s'appliquent pas à cette personne prennent la valeur zéro « 0 ». Par exemple, si une personne sans conjoint a déclaré un revenu en 2015, la valeur de la variable **RRSPSI2015** (contributions au REER d'un conjoint) devrait être zéro « 0 ». Si cette personne n'avait pas produit de déclaration de revenus en 2015, la valeur serait manquante.

Exemple de programme pour la BDIM

* Exemple de programme SAS pour la BDIM

```
libname sourcel 'FILEFOLDER1'; * emplacement des fichiers de la BDIM ;
libname Out 'FILEFOLDER2'; * répertoire de l'utilisateur ;
```

* L'objectif de ce programme est d'utiliser la BDIM pour obtenir le nombre de personnes recevant des prestations d'aide sociale, vivant en Ontario, pour lesquelles aucun revenu d'emploi ne figurait sur leur feuillet T4, selon le sexe et l'année (dans ce cas, de 2015 à 2017). Les données relatives aux provinces et aux revenus proviennent des fichiers annuels de la BDIM, alors que la variable relative au sexe provient du fichier PNRF_1980_2019 (FRPN). ;

* La première étape est de créer un fichier de données contenant toute l'information nécessaire à la création de nos tableaux. Ce fichier est nommé SAOnt et est enregistré dans le répertoire de sortie. L'identifiant longitudinal (IMDB_ID) est utilisé pour fusionner les ensembles de données annuels de la BDIM. ;

```
data out.SAOnt;
merge
sourcel.imdb _ t1ff _ 2015(where=(prco _ i2015 = 5 and outlier _ ind2015=0) in=a
keep=imdb _ id prco _ i2015 saspyf2010 t4e _ _ i2015 outlier _ ind2015)

sourcel.imdb _ t1ff _ 2016(where=(prco _ i2016 = 5 and outlier _ ind2016=0) in=b
keep= imdb _ id prco _ i2016 saspyf2016 t4e _ _ i2016 outlier _ ind2016)

sourcel.imdb _ t1ff _ 2017(where=(prco _ i2017 = 5 and outlier _ ind2017=0) in=c
keep= imdb _ id prco _ i2017 saspyf2012 t4e _ _ i2017 outlier _ ind2017)

sourcel.pnrf _ 1980 _ 2019(keep= imdb _ id gender landing _ year immigration _
category);

by IMDB _ id ;
```

```
If a and b and c and (landing _ year>=2007 and landing _ year<=2012);
```

* La personne doit avoir produit une déclaration de revenus pour chacune des trois années, ne pas être une valeur aberrante, et avoir été admise entre 2007 et 2012 (population d'intérêt).;

* On crée une variable dichotomique identifiant les bénéficiaires d'aide sociale pour chaque année.

Cela produit trois variables :

```
flag _ sa2015, flag _ sa2016 et flag _ sa2017, prenant la valeur 1 ou 0.;
```

```
If (t4e _ i2015=0 and saspyf2015>0) then flag _ sa2015 = 1 ;
```

```
else flag _ sa2015 = 0 ;
```

```
if (t4e _ i2016=0 and saspyf2011>0) then flag _ sa2016 = 1 ;
```

```
else flag _ sa2016 = 0 ;
```

```
if (t4e _ i2017=0 and saspyf2012>0) then flag _ sa2017 = 1 ;
```

```
else flag _ sa2017 = 0 ;
```

```
run;
```

* La procédure 'freq' de SAS permet de produire nos tableaux. On doit ensuite s'assurer que les règles relatives à la confidentialité sont respectées. ;

```
proc freq data = out.SAOnt;
tables immigration _ category*flag _ sa2015*flag _ sa2016*flag _ sa2017
gender*flag _ sa2015*flag _ sa2016*flag _ sa2017 /missing ;
run;
* Fin de l'exemple de programmation.;
```

D.2 Création d'une cohorte

Avant de commencer l'analyse, il est essentiel de déterminer la cohorte d'intérêt. Cette cohorte peut être limitée par l'année d'admission, des variables géographiques ou toute autre variable d'intérêt (p. ex. la catégorie d'admission ou le sexe), selon les besoins du chercheur. Une cohorte unique clairement définie doit être choisie pour assurer la comparabilité. Un chercheur peut s'intéresser, par exemple, aux femmes admises en 2000 et vivant au sein d'une famille ayant reçu l'aide sociale en 2001 (tableau 17). Une question d'étude relative à cette cohorte pourrait être « Quelle portion de cette cohorte a reçu l'aide sociale au cours des deux années suivantes (2002 et 2003)? » Il convient de noter que l'Agence du revenu du Canada (ARC) exige que le conjoint recevant le revenu net le plus élevé déclare le versement de l'aide sociale. Par conséquent, il est conseillé de générer la mesure de l'aide sociale (SASPY_F), y compris pour les particuliers, à partir des données au niveau familial.

Tableau 17

Exemple : Femmes admises en 2002 ayant reçu l'aide sociale (SASPY_F) en 2001

IMDB_ID	Année d'admission	Sexe	SASPY_F2001	SASPY_F2002	SASPY_F2003
			dollars		
IM583	2000	Femme	20 500	19 000	14 000
IM145	2000	Femme	3 000	0	0
IM548	2000	Femme	11 500	13 800	0
IM798	2000	Femme	16 000	18 000	8 000
IM961	2000	Femme	10 000	0	0
IM967	2000	Femme	9 500	0	0
IM110	2000	Femme	5 000	2 000	1 000
IM125	2000	Femme	1 000	0	200

Source : Statistique Canada, exemple provenant de la base de données longitudinales sur l'immigration (BDIM).

D.3 Calcul des taux de rétention

L'un des principaux avantages de la BDIM est qu'elle comprend des variables géographiques permettant d'étudier la mobilité et la rétention. Aucune autre base de données ne présente un niveau comparable de détails annuels relatifs aux déclarants, en particulier pour des régions plus réduites. Disposer de mises à jour annuelles au niveau des provinces, des divisions de recensement, des régions métropolitaines de recensement, des agglomérations de recensement, des subdivisions de recensement et des secteurs de recensement permet une vaste gamme d'analyses.

Il est possible d'étudier les trajectoires de mobilité individuelle en attribuant simplement des indicateurs aux changements de code postal et de calculer des tendances de mobilité en étudiant les déménagements à des niveaux géographiques donnés. La mobilité au niveau des subdivisions de recensement (changements de subdivision de recensement d'une année à l'autre) et la mobilité à l'échelle provinciale (changement de province d'une année à l'autre) varie de manière significative en fonction de plusieurs caractéristiques d'immigrant, comme l'âge et la catégorie d'admission. Ces données géographiques sont dérivées du code postal (variable PSCO dans la BDIM au niveau individuel et familial). Le code postal est un code alphanumérique à six caractères définissant le lieu de livraison du courrier adressé aux clients du service postal au Canada. Voir la section 3.4.1 pour une description des variables géographiques.

Dans l'exemple ci-dessous (tableau 18), le chercheur s'intéresse à la mobilité jusqu'en 2002. Les personnes IM798, IM961, IM967 et IM110 peuvent être exclues de l'étude sur la mobilité du fait de l'absence de données (ou de fichiers).

Tableau 18
Exemple : Mobilité jusqu'en 2002 des immigrants admis en 2000

IMDB ID	Année d'admission	Province de destination	PRCO 2000	PRCO 2001	PRCO 2002
IM583	2000	C.-B.	C.-B.	C.-B.	C.-B.
IM145	2000	Alb.	Alb.	Sask.	Sask.
IM548	2000	Alb.	Ont.	Ont.	Ont.
IM798	2000	Ont.	..	Ont.	Ont.
IM961	2000	N.-B.	N.-B.	N.-B.	..
IM967	2000	Ont.	..	Alb.	Ont.
IM110	2000	..	Qc	..	Qc

.. indisponible pour une période de référence précise

Note : PRCO : province de résidence.

Source : Statistique Canada, exemple provenant de la base de données longitudinales sur l'immigration (BDIM).

Alors que la mobilité au niveau individuel est relativement simple, le taux de rétention des immigrants dans un territoire peut être calculé de diverses manières. La façon de calculer cette rétention est une décision analytique fondée sur les besoins particuliers du chercheur. Il est relativement simple de définir le nombre de personnes restant : il s'agit du nombre de personnes produisant une déclaration de revenus dans le territoire d'intérêt à un moment donné. Il est nécessaire de prendre une décision quant à la cohorte initiale d'admission pour laquelle le chiffre de rétention est calculé (le dénominateur du taux de rétention).

Les taux de rétention indiquent la proportion des immigrants déclarants résidant dans la province où ils ont été admis (définie comme la province de destination envisagée) à un moment donné. Pour une cohorte donnée (p. ex. l'année d'admission) et une année d'imposition donnée (ou un certain nombre d'années depuis l'admission), le dénominateur est le nombre de déclarants correspondant à la province d'admission sélectionnée. Le numérateur est le nombre de déclarants correspondant à la province d'admission sélectionnée qui résident également dans cette province.

Pour calculer les taux de rétention de la cohorte de 2011 trois ans après l'admission, un chercheur choisirait, par exemple, toutes les provinces d'admission (p. ex. la province de destination envisagée), toutes les provinces de résidence, l'année d'admission = 2011 et l'année de référence = 2014. Ce tableau ressemblerait au suivant.

Tableau 19
Province de résidence en 2014 et province d'admission, cohorte de 2011

Province d'admission	Province de résidence					
	Total – Province de résidence	Terre-Neuve-et-Labrador	Île-du-Prince-Édouard	Nouvelle-Écosse	Nouveau-Brunswick	Québec
	nombre d'immigrants					
Total – Province d'admission	174 740	405	330	1 365	880	31 505
Terre-Neuve-et-Labrador	515	325	0	5	0	5
Île-du-Prince-Édouard	1 245	0	265	25	10	30
Nouvelle-Écosse	1 460	10	5	1 080	10	25
Nouveau-Brunswick	1 340	0	10	35	750	55
Québec	36 275	10	10	35	15	30 200
Ontario	69 135	35	25	115	70	875
Manitoba	11 190	0	0	15	0	55
Saskatchewan	6 360	0	0	0	0	20
Alberta	21 940	10	0	20	0	95
Colombie-Britannique	25 000	5	0	30	5	140
Autre	280	0	0	0	0	0

Province d'admission	Province de résidence					
	Ontario	Manitoba	Saskatchewan	Alberta	Colombie-Britannique	Autre lieu de résidence
	nombre d'immigrants					
Total – Province d'admission	70 590	9 695	6 120	26 965	26 390	500
Terre-Neuve-et-Labrador	75	5	0	60	30	0
Île-du-Prince-Édouard	560	0	0	50	295	0
Nouvelle-Écosse	185	0	5	90	30	10
Nouveau-Brunswick	275	0	10	80	120	0
Québec	3 255	40	75	1 190	1 400	45
Ontario	63 145	275	335	2 815	1 325	115
Manitoba	645	9 170	80	825	380	10
Saskatchewan	295	45	5 370	445	165	10
Alberta	810	65	140	20 170	590	35
Colombie-Britannique	1 330	85	100	1 200	22 030	70
Autre	15	0	0	35	20	200

Source : Statistique Canada, Base de données longitudinales sur l'immigration de 2014.

Les résultats pour la Nouvelle-Écosse donnent des renseignements à ce sujet. Au total, 1 460 personnes ont été admises en Nouvelle-Écosse en 2011 et ont produit une déclaration de revenus en 2014. Pour 1 080 de ces personnes, la Nouvelle-Écosse figurait comme province de résidence en 2014. Le taux de rétention sur trois ans pour la Nouvelle-Écosse serait 1 080/1 460, soit environ 74 %. Tableau 18 fournit également des renseignements sur les migrants secondaires¹⁵ : 1 365 personnes admises en 2011 résidaient en Nouvelle-Écosse en 2014; parmi ces dernières, 1 080 envisageaient de résider en Nouvelle-Écosse et 285 avait indiqué une autre province de destination.

La définition susmentionnée de rétention part du principe que le nombre de déclarants pour la province donnée de destination envisagée est la population totale pouvant y rester au cours d'une année (c.-à-d., si les 1 460 personnes ayant envisagé de résider en Nouvelle-Écosse y avaient produit une déclaration de revenus en 2014, la province aurait enregistré un taux de rétention de 100 %). Cette méthode ne tient pas compte des interruptions de déclarations ou des émigrants ayant quitté le Canada, pour lesquels aucune déclaration fiscale n'était disponible en 2014.

15. Personnes ayant déménagé au sein du Canada après avoir atteint leur destination initiale au Canada.

Une option est une approche purement longitudinale, selon laquelle on sélectionne une seule cohorte d'admission (selon la province de destination envisagée, la province de déclaration de revenus initiale, ou les deux) et le taux de rétention est calculé comme la portion de cette cohorte déclarant toujours ses revenus dans cette province. Lorsque la province de déclaration de revenus initiale est utilisée pour définir la cohorte d'admission, il est recommandé que la première déclaration de revenus ait lieu l'année d'admission des immigrants (année d'admission = année d'imposition), afin d'exclure les personnes arrivées ailleurs et ayant ensuite migré dans cette région avant de déclarer leurs revenus pour la première fois. Il est possible d'ajouter une autre restriction, si le chercheur s'intéresse à la population dont le lieu de destination correspond au lieu géographique de la première déclaration de revenus.

Étant donné qu'une portion de chaque cohorte ne déclare pas de revenus pour l'année d'admission, il peut être nécessaire d'accroître la taille de la population pour une région en définissant la cohorte d'admission comme toute personne ayant déclaré ses revenus pour la première fois dans la région au cours des deux années suivant son admission (c.-à-d., $\text{first_tax_year} = \text{landing_year}$ ou $\text{landing_year}+1$). Il n'est pas recommandé d'inclure à la « cohorte d'admission » les personnes dont la première déclaration de revenus a eu lieu plusieurs années après leur admission, car il est possible que ces personnes soient d'abord arrivées ailleurs sans toutefois avoir déclaré de revenus. Il est également suggéré d'exclure les déclarants intermittents de ces analyses, car leur lieu de résidence n'est pas connu pour les années pour lesquelles aucune donnée fiscale n'existe. La rétention calculée de cette manière indiquera une diminution graduelle du nombre; cette diminution est due aux immigrants cessant de déclarer leurs revenus, émigrant ailleurs ou décédant.

Si un chercheur s'intéresse aux migrations secondaires vers une région, il est possible de l'évaluer en supprimant les personnes de la cohorte d'admission définie du nombre total d'immigrants déclarant des revenus dans la région au moment souhaité. Toutefois, à nouveau, il est suggéré de limiter ces analyses aux personnes ayant déclaré leurs revenus pour la première fois au cours de cette période (année 0 ou année 1), afin de ne pas prendre des déclarants tardifs pour des migrants dans la région. Si la cohorte d'admission est limitée aux immigrants dont la destination correspond au lieu de la première déclaration de revenus, il est suggéré de faire une autre distinction entre les migrants secondaires ayant produit leur première déclaration de revenus ailleurs (puis dans la région d'intérêt) et les immigrants ayant produit leur première déclaration de revenus dans la région d'intérêt, puis ayant été admis dans une autre région (ou lorsque toute donnée manque quant à leur destination envisagée).

Le tableau suivant présente un exemple d'une approche longitudinale de rétention provinciale utilisant des données fictives, selon diverses définitions de la cohorte d'admission initiale.

Tableau 20
Nombre d'immigrants déclarants au sein de la population en question résidant en Colombie-Britannique et taux de rétention associé, par année depuis l'admission

Années depuis l'admission	Déclarants ayant pour la première fois déclaré des revenus en C.-B. au cours de l'année 0		Déclarants ayant pour la première fois déclaré des revenus en C.-B. au cours de l'année 0 ou 1		Déclarants ayant pour la première fois déclaré des revenus en C.-B. au cours de l'année 0 ou 1 et pour lesquels la destination envisagée était la C.-B.	
	nombre	Taux de rétention pourcentage	nombre	Taux de rétention pourcentage	nombre	Taux de rétention pourcentage
0	20 000	100	20 000	...	17 500	...
1	18 000	90	25 000	100	19 000	100
2	17 000	85	23 000	92	18 000	95
3	16 500	83	22 000	88	17 500	92

... n'ayant pas lieu de figurer

Source : Statistique Canada, exemple provenant de la base de données longitudinales sur l'immigration.

Dans l'exemple ci-dessus, il est possible de calculer la rétention en Colombie-Britannique selon trois définitions de la population; le taux de rétention sur trois ans varie alors en fonction de la définition adoptée. Il est important que toutes les personnes de l'échantillon déclarent des revenus à chaque moment d'intérêt.

Avec la publication de la **BDIM de 2020**, un tableau récapitulatif sur la mobilité est disponible sur le site Web de Statistique Canada. Les mesures de mobilité comparent la destination prévue des fichiers d'immigration à la province de résidence obtenue à partir des déclarations de revenus. Par exemple, le tableau 21 présente les mesures de mobilité en fonction des différences entre la province de destination prévue des immigrants admis en 2010 et leur province de résidence en 2015 selon leur déclaration de revenus.

Tableau 21
Mesures de mobilité pour l'année d'admission 2010 selon la province, année d'imposition 2015

	Total destination (a)	Total résidence (b)	Migration externe (c)	Migration interne (d)	Destiné et toujours résident (e=a-c)	Taux de croissance de la population (f=b/a-1)	Taux de rétention (g=e/a)	Taux de migration externe (h=1-e/a)	Taux de migration interne (i=d/a)
Canada	200 600	200 600	27 260	27 260	173 340	0,0	86,4	13,6	13,6
Terre-Neuve-et-Labrador	525	410	245	130	280	-21,9	53,3	46,7	24,8
Île du Prince Edward	1 930	370	1 630	70	305	-80,8	15,8	84,5	3,6
Nouvelle Écosse	1 630	1 405	570	340	1 065	-13,8	65,3	35,0	20,9
Nouveau Brunswick	1 535	920	795	180	740	-40,1	48,2	51,8	11,7
Québec	38 050	33 900	5 955	1 805	32 095	-10,9	84,4	15,7	4,7
Ontario	83 355	84 965	7 725	9 335	75 630	1,9	90,7	9,3	11,2
Manitoba	11 475	9 785	2 420	730	9 055	-14,7	78,9	21,1	6,4
Saskatchewan	5 620	5 410	1 220	1 015	4 400	-3,7	78,3	21,7	18,1
Alberta	24 255	29 850	2 360	7 955	21 895	23,1	90,3	9,7	32,8
Colombie Britannique	31 820	32 790	4 250	5 215	27 575	3,1	86,7	13,4	16,4
Autre	405	420	95	115	305	3,7	75,3	23,5	28,4
Manquants	..	365	..	365	..	0,0	0,0	0,0	0,0

.. indisponible pour une période de référence précise

Source : Statistique Canada, Base de données longitudinales sur l'immigration 2015 (BDIM).

Le nouveau tableau fournit les mesures de mobilité suivantes :

- La **destination totale** (colonne a) représente le nombre d'immigrants admis en 2010 et déclarant des taxes en 2015, au Canada;
- La **résidence totale** (colonne b) représente le nombre d'immigrants déclarants en 2015 dans la province spécifiée;
- La **migration externe** (colonne c) représente le nombre d'immigrants déclarants originaires de la province spécifiée et ayant produit leur déclaration dans une autre province, en 2015;
- La **migration interne** (colonne d) représente le nombre d'immigrants originaires d'une province de destination différente et qui déclare des impôts dans la province spécifiée pour l'année 2015;
- Le nombre de personnes **destinés et toujours résidents** (colonne e) représente le nombre d'immigrants déclarants qui a continué leur résidence dans la province de destination, pour l'année 2015 ;
- Le **taux de croissance de la population** (colonne f) représente le pourcentage d'immigrants déclarants gagnés ou perdus par la province spécifiée. Cela tient compte des immigrants qui migrent à l'extérieur de la province et qui migrent dans la province spécifiée;

- Le **taux de rétention** (colonne g) représente le pourcentage d'immigrants déclarants qui continueront de résider dans la province de destination, en 2015. Cela ne tient pas compte des immigrants qui arrivent d'une autre province de destination;
- Le **taux de migration externe** (colonne h) représente le pourcentage d'immigrants déclarants originaires de la province spécifiée et qui déclare des impôts dans une autre province au cours de l'année 2015;
- Le **taux de migration interne** (colonne i) représente le pourcentage d'immigrants originaires d'une autre province de destination et qui déclare des impôts dans la province spécifiée en 2015.

Le tableau 21 montre que 200 600 immigrants ont été admis au Canada en 2010 et ont produit une déclaration de revenus en 2015.

Des 83 355 immigrants déclarants qui avaient l'intention de résider en Ontario, 75 630 y sont demeurés en 2015, ce qui représente un taux de rétention de 90,7 %.

Alors que 7 725 immigrants déclarants ont quitté l'Ontario, 9 335 immigrants déclarants ont déménagé en Ontario à partir d'autres provinces de destination. Ainsi, pour cette cohorte de 2010, le nombre total de résidents de l'Ontario en 2015 était de 84 965, soit 1,9 % de plus que le nombre d'immigrants déclarants qui comptaient résider en Ontario.

Enfin, les analystes doivent procéder avec précaution lors de l'étude de divisions géographiques de faible niveau sur une longue période, car les frontières des agglomérations et des régions métropolitaines de recensement évoluent et des subdivisions de recensement sont abandonnées et ajoutées. Lorsque cela est possible, les analystes devraient exécuter le programme de Fichier de conversion des codes postaux (FCCP+) pour uniformiser les codes postaux en divisions géographiques constantes.

D.4 Calcul des trajectoires de revenu au cours du temps

Comme pour le taux de rétention, calculer l'évolution des revenus de salaires, traitements et commissions d'une année à l'autre (ou, d'ailleurs, toute variable économique) nécessite des données consécutives. Si un chercheur, par exemple, souhaite comparer le revenu de salaires, traitements et commissions médian de la cohorte de femmes âgées de 24 ans à 54 ans en 2000, un an après leur admission et cinq ans après leur admission (tableau 22), les enregistrements présentant des fichiers FFT1 manquants devraient être supprimés de l'analyse. La décision de supprimer ces enregistrements se baserait sur le souhait d'évaluer le revenu médian de la cohorte par rapport au revenu médian des déclarants de la cohorte.

Tableau 22
Revenu de salaires, traitements et commissions médian de la cohorte de femmes âgées de 24 à 54 ans en 2000, 1 an et 5 ans après l'admission

IMDB_ID	Année d'admission	Âge à l'admission	Sexe	Revenu de salaires	Revenu de salaires
				2001	2005
				dollars	
IM583	2000	34	Femme	20 500	49 000
IM145	2000	53	Femme	..	56 000
IM548	2000	29	Femme	11 500	33 800
IM798	2000	31	Femme	36 000	0
IM961	2000	42	Femme	10 000	..
IM967	2000	40	Femme
IM110	2000	35	Femme	0	59 000

.. indisponible pour une période de référence précise

Source : Statistique Canada, exemple provenant de la base de données longitudinales sur l'immigration.

Des précautions doivent être prises lors du calcul du revenu pour la « première année au Canada », car il peut ne pas représenter une année d'imposition complète. Une personne admise, par exemple, en novembre 2013 et ayant déclaré des revenus pour 2013 ne déclarerait que deux mois de revenus en 2013. Une pratique exemplaire est d'utiliser la première année de revenu complète (année d'admission +1, voir le tableau 20). Une exception concerne les déclarants précoces, ayant déclaré des revenus au Canada avant leur admission ainsi que pour l'année d'imposition, car ils sont plus susceptibles de déclarer des revenus pour une année entière.

Il est également important d'étudier le revenu au cours du temps en dollars constants. Par conséquent, il est nécessaire d'apporter des ajustements à l'aide de l'Indice des prix à la consommation (IPC) (voir l'annexe D.7). Cet ajustement a été apporté dans les tableaux de la BDIM.

D.5 Arrondissement des données

Statistique Canada s'engage à respecter la vie privée des Canadiens. Par conséquent, tous les tableaux produits à partir des fichiers IMDB_TIFF font l'objet d'un arrondissement. L'intérêt de cet arrondissement est de veiller à ce qu'aucune cellule réduite ne soit divulguée pouvant révéler des renseignements sur des personnes données ou de petits groupes de personnes. En général, les macros prélèveront un ensemble de données d'entrée non arrondies de diverses statistiques (nombres, moyennes, médianes, etc.) et créeront un ensemble de données arrondi.

Les règles d'arrondissement sont mises à la disposition de l'ensemble des chercheurs ayant accès aux microdonnées dans les Centres de données de recherche (CDR).

D.6 Détection de valeurs aberrantes

La variable de valeur aberrante OUTLIER_IND a été créée pour détecter les valeurs aberrantes au sein du fichier FFT1 (voir la section 5.5). Elle devrait être utilisée pour supprimer les données aberrantes de tout calcul (p. ex. moyenne, médiane ou régression) utilisant des données fiscales. Les valeurs aberrantes varient d'une année à l'autre; ce qui signifie que les données relatives à une personne peuvent être détectées comme valeur aberrante pour une année, mais pas pour une année ultérieure.

Le tableau suivant (Tableau 23) présente la répartition des valeurs aberrantes dans les fichiers de données fiscales à partir de 1982, par type de résident, pour la **BDIM de 2020**. Moins de 0,06 % des enregistrements ont été relevés comme étant des valeurs aberrantes par année d'imposition. La proportion de valeurs aberrantes a augmenté de 1995 à 1996, du fait de mises à jour de la méthode de détection des valeurs aberrantes appliquée aux fichiers de données fiscales à partir de l'année d'imposition de 1997.

Tableau 23
Répartition des déclarants avec des valeurs aberrantes par année d'imposition

Année	Total	
	nombre	pourcentage
1982	500	0,03
1983	390	0,02
1984	560	0,03
1985	500	0,02
1986	460	0,02
1987	600	0,03
1988	960	0,04
1989	900	0,03
1990	710	0,03
1991	780	0,03
1992	990	0,03
1993	850	0,03
1994	500	0,01
1995	730	0,02
1996	1 540	0,04
1997	1 930	0,05
1998	2 490	0,06
1999	1 810	0,04
2000	2 040	0,05
2001	1 970	0,04
2002	2 050	0,04
2003	2 240	0,05
2004	1 950	0,04
2005	2 250	0,05
2006	2 470	0,05
2007	2 350	0,04
2008	2 480	0,04
2009	2 750	0,05
2010	2 430	0,04
2011	2 710	0,04
2012	2 320	0,04
2013	2 330	0,04
2014	2 340	0,03
2015	2 420	0,04
2016	2 290	0,03
2017	2 340	0,03
2018	2 810	0,04
2019	2 870	0,04

Source : Statistique Canada, Base de données longitudinales sur l'immigration de 2020.

D.7 Ajustement du revenu selon l'Indice des prix à la consommation (IPC)

Afin de tenir compte du coût de la vie, il est nécessaire d'ajuster tous les revenus en fonction de l'Indice des prix à la consommation (IPC) du Canada. « L'Indice des prix à la consommation (IPC) est un indicateur de la variation des prix à la consommation payés par les Canadiens. Pour l'établir, on compare au fil du temps le coût d'un panier fixe de biens et services achetés par les consommateurs. Puisque le panier comprend des biens et services d'une quantité et d'une qualité constante ou équivalente, l'indice reflète uniquement la fluctuation pure des prix¹⁶. » Les facteurs de correction pour 2017 sont présentés au tableau 24. Pour convertir les données en dollars constants d'une année donnée (année de référence), les utilisateurs de données doivent multiplier les valeurs en dollars de toutes les années, à l'exception de l'année de référence, par un facteur de correction propre à une année. Pour obtenir ces facteurs de correction, les utilisateurs de données doivent diviser l'IPC de l'année de référence par l'IPC de l'année étudiée. L'année de référence dans le tableau 24 est 2017.

16. [Site web](#).

Tableau 24
Facteur de correction de l'Indice des prix à la consommation de 2020

Année	L'ajustement de l'Indice des prix à la consommation de 2020 est égal à 137.0 divisé par
	nombre
1982	54,9
1983	58,1
1984	60,6
1985	63,0
1986	65,6
1987	68,5
1988	71,2
1989	74,8
1990	78,4
1991	82,8
1992	84,0
1993	85,6
1994	85,7
1995	87,6
1996	88,9
1997	90,4
1998	91,3
1999	92,9
2000	95,4
2001	97,8
2002	100,0
2003	102,8
2004	104,7
2005	107,0
2006	109,1
2007	111,5
2008	114,1
2009	114,4
2010	116,5
2011	119,9
2012	121,7
2013	122,8
2014	125,2
2015	126,6
2016	128,4
2017	130,4
2018	133,4
2019	136,0
2020 ¹	137,0

1. En 2020, l'IPC était 137.0. Pour transformer les montants d'une année en dollars constant de 2020, on doit multiplier le montant en dollars de cette année par le ratio d'inflation, qui est l'IPC de 2020 (xxxx) divisé par l'IPC de cette année.

Source : Statistique Canada, tableau 18-10-0005-01.

D.8 Calcul de mesures de revenu clés

Les tableaux de la BDIM contiennent plusieurs mesures du revenu. Le tableau 25 décrit les variables du FFT1 incluses dans ces calculs.

Tableau 25**Description des principales mesures de revenu dans la Base de données longitudinales sur l'immigration**

Mesure	Composantes	Formule
Revenu de salaires, traitements et commissions	Revenu des feuillets T4	T4E__i + OEI_i
Revenu de travail indépendant		
Depuis 1988	Revenu provenant d'un travail indépendant comme des revenus d'entreprise, de profession libérale, de commissions, d'agriculture ou de pêche; revenu de société en commandite	SEI__i + LTPL_i
Avant 1988	Revenu provenant d'un travail indépendant comme des revenus d'entreprise, de profession libérale, de commissions, d'agriculture ou de pêche	SEI__i
Revenu de placements	Revenus d'intérêts et d'investissements; dividendes; gains et pertes en capital, net imposable	INVI_i + XDIV_i + CLKGX
Indemnités d'assurance-emploi	Indemnités d'assurance-emploi	EINS__i
Allocations sociales	Allocations sociales (utiliser le niveau familial)	SASPYf
Revenu total	Somme de toutes les mesures décrites ci-dessus	

Source : Statistique Canada, Base de données longitudinales sur l'immigration de 2020.

Il convient de noter que toutes les valeurs aberrantes sont supprimées de ces calculs (Outlier_ind=1), que l'on utilise la variable de province de résidence à la fin de l'année (PRCO_) pour déterminer la province et que tous les revenus sont corrigés en fonction de l'Indice des prix à la consommation (IPC) de l'année la plus récente pour laquelle le FFT1 est disponible. « Moyenne, avec revenu » désigne le revenu moyen des immigrants déclarants ayant un revenu du type donné. « Médiane, avec revenu » désigne le revenu médian des immigrants déclarants ayant un revenu du type donné.

Bibliographie

- A. Carpentier et G. Pinsonneault. 1994. *Representativeness Study of Immigrants Included in the Immigrant Data Bank (IMDB Project)*. Ministère des Affaires internationales, de l'Immigration et des Communautés culturelles. Gouvernement du Québec.
- A. Diaz-Papkovich, *IMDB Linkage Summary 2017*. Document non publié. Ottawa, Statistique Canada, 2017.
- C. Langlois et C. Dougherty, *La Base de données longitudinale sur l'immigration (BDIM): Une introduction*. Recueil de la Conférence 1997 sur l'immigration, l'emploi et l'économie de Citoyenneté et Immigration Canada, 1997.
- Commission de l'immigration et du statut de réfugié du Canada, [Section de la protection des réfugiés](#), 2015. (site consulté le 13 janvier 2016).
- Gouvernement du Canada, [Demander la citoyenneté : Qui peut présenter une demande, 2021](#). (site consulté le 12 octobre 2021).
- Immigration, Réfugiés et Citoyenneté Canada, [Rapport annuel au Parlement sur l'immigration](#), 2018. (site consulté le 19 septembre 2019).
- M. Rotermann, C. Sanmartin, R. Trudeau, and H. St-Jean, « Couplage des données du Recensement de 2006 et des données hospitalières au Canada » [Rapports sur la santé](#). Vol. 26, n° 10 (n° 82-003-X au catalogue) Ottawa, Statistique Canada, 2015. (site consulté le 13 mars 2017).
- H. Dryburgh, [La Banque de données administratives longitudinales \(DAL\)](#) et la Banque de données longitudinales sur les immigrants (BDIM) : création de la DAL_BDIM - un document technique (n° 89-612-XIE au catalogue) Ottawa, Statistique Canada, 2004. (site consulté le 13 mars 2017).
- J. Badets, et C. Langlois. 2000. « Les défis de l'utilisation de données administratives à l'appui de la recherche en matière de politiques: l'exemple de la base de données longitudinale sur l'immigration (BDIM) » *Au Symposium 99 - Combiner des données de sources différentes, 1999*. Symposium international de Statistique Canada: Recueil, produit n° 11-522-XPE au catalogue de Statistique Canada.
- P. Cascagnette et la section de production d'ECDS. 2021. « Social Data Linkage Environment (SDLE) Methodology Report – Linkage between the Immigration File (1952- January 2021) and the SDLE Derived Record Depository (version 47).» Document inédit. Ottawa: Statistique Canada.
- S.B. Dusetzina, S. Tyree, A.M. Meyer, A. Meyer, L. Green et W.R. Carpenter, [Linking Data for Health Services Research: A Framework and Instructional Guide \[Internet\]](#). Préparé par l'Université de la Caroline du nord à Chapel Hill sous le contrat n° 290-2010-000141. AHRQ Publication n° 14-EHC033-EF. Rockville, MD: Agency for Healthcare Research and Quality, 2014 (site consulté le 14 mars 2017).
- S. McLeish, *2008 IMDB Landing File: Data Quality Working Paper*. Document non publié. Ottawa, Statistique Canada, 2011.
- Statistique Canada, [Estimations annuelles du revenu des familles de recensement et des particuliers](#) (Fichier des familles T1), 2018. (site consulté le 19 septembre 2019).
- Statistique Canada, [150 ans d'immigration au Canada](#). Mégatendances canadiennes (n° 11-630-X au catalogue) Ottawa, Statistique Canada, 2016. (site consulté le 14 mars 2017).
- W.E. Winkler, « Record linkage » *Sample Surveys: Design, Methods and Applications* 29A, 2009, p. 351–380