

**ÉTUDE LONGITUDINALE DU DÉVELOPPEMENT
DES ENFANTS DU QUÉBEC – (ELDEQ 1998-2019)**



**Guide de l'utilisateur de la
banque de données du volet 2019**

Jeunes âgés d'environ 21 ans

**Direction des enquêtes longitudinales et sociales
Institut de la statistique du Québec**

Septembre 2020

La documentation qui suit présente l'information nécessaire pour utiliser le plus efficacement possible les données de l'*Étude longitudinale du développement des enfants du Québec* (ELDEQ) du **volet 2019 (E22)**. Ce document, divisé en quatre sections, présente d'abord un aperçu de la collecte. La deuxième section traite, pour sa part, de la convention de désignation des variables présentes dans la banque de données. La structure et le contenu des fichiers sont abordés à la section 3 et, finalement, les remarques et mises en garde sont présentées à la section 4.

1. Collecte de données : Un aperçu

Le volet 2019 de l'*Étude longitudinale du développement des enfants du Québec* (ELDEQ) a été réalisé auprès de 1323 jeunes¹ du milieu du mois de février à la fin juin 2019. Ainsi, au moment de cette collecte, les jeunes² avaient environ 21 ans.

La banque de données du volet 2019 inclut les informations tirées du Questionnaire en ligne au jeune (QELJ)³, qui a été rempli en ligne et sans l'aide d'une intervieweuse. Un seul instrument de mesure a été administré en 2019.

2. Convention de désignation des variables

Pour désigner et retrouver facilement chacune des variables des différents fichiers de données, nous avons conçu une nomenclature qui s'appuie en grande partie sur celle utilisée pour l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) de Statistique Canada que nous avons adaptée pour l'ELDEQ. Généralement, toutes les variables contenues dans la banque de données ont au plus huit caractères.

2.1. Nomenclature générale des variables

Les noms de variables se présentent ainsi :

V SE C Q ppx, où :

V : la première position se rapporte au volet de l'étude. Ainsi,

- « A » correspond au volet 1998 (5 mois) ;
- « B » correspond au volet 1999 (17 mois) ;
- « C » correspond au volet 2000 (29 mois) ;
- « D » correspond au volet 2001 (41 mois) ;
- « E » correspond au volet 2002 (44 à 56 mois) ;
- « F » correspond au volet 2003 (environ 5 ans) ;
- « G » correspond au volet 2004 (environ 6 ans) ;
- « H » correspond au volet 2005 (environ 7 ans) ;
- « I » correspond au volet 2006 (environ 8 ans) ;

1 Pour obtenir plus de détails concernant la population et l'échantillon initial, consultez le numéro 1 des volumes 1 et 2 de la collection ELDEQ (disponible sur le site web de l'ELDEQ à l'adresse suivante : <http://www.iesuisieserai.stat.gouv.qc.ca>, sous l'onglet « Publications »).

2 À partir du volet 2013, les termes « jeune » et « enfant » sont tous deux utilisés dans la documentation technique et dans les bases de données de l'ELDEQ pour désigner l'enfant-cible.

3 Tous les instruments de collecte sont disponibles sur le site de l'ELDEQ (www.iesuisieserai.stat.gouv.qc.ca) sous l'onglet « Outils de collecte ».

- « K » correspond au volet 2008 (environ 10 ans) ;
- « M » correspond au volet 2010 (environ 12 ans) ;
- « N » correspond au volet 2011 (environ 13 ans) ;
- « P » correspond au volet 2013 (environ 15 ans) ;
- « R » correspond au volet 2015 (environ 17 ans) ;
- « T » correspond au volet 2017 (environ 19 ans) ;
- « V » correspond au volet 2019 (environ 21 ans).

SE : les deuxième et troisième positions se rapportent à une section du Questionnaire en ligne du jeune (QELJ) (voir le tableau 2).

C : correspond généralement à la personne à laquelle la variable se rapporte. Ainsi,

- « N » signifie que la variable se rapporte à l'enfant/jeune **et qu'elle est tirée du questionnaire informatisé ou en ligne (QIE et QELJ)**;
- « E » signifie que la variable se rapporte à l'enfant;
- « R » signifie que la variable se rapporte à la fratrie;
- « M » signifie que la variable se rapporte à la mère/conjointe présente;
- « J » signifie que la variable se rapporte au père/conjoint présent;
- « F » signifie que la variable se rapporte au ménage;
- « A » signifie que la variable se rapporte au père biologique absent du ménage;
- « I » signifie que la variable se rapporte à l'enseignant/e de l'enfant cible;
- « 1 » signifie que la variable se rapporte au premier enseignant de l'enfant (volet 2011);
- « 2 » signifie que la variable se rapporte au deuxième enseignant de l'enfant (volet 2011).

Q : la cinquième position se rapporte au genre de variable. Ainsi,

- « Q » signifie que la variable se rapporte à une question qui a été posée directement dans l'un des questionnaires;
- « S » signifie que la variable représente une note calculée pour une échelle utilisée dans le questionnaire;
- « D » signifie que la variable a été calculée à partir d'autres questions qui ont été posées dans le questionnaire;
- « T » signifie que la note calculée pour une échelle utilisée dans le questionnaire a été standardisée, c'est-à-dire ramenée à une valeur se situant entre 0 et 10;
- « L » signifie que la note calculée pour une échelle à un volet donné est basée uniquement sur les items longitudinaux, c'est-à-dire ceux communs à tous les volets;
- « M » signifie qu'il s'agit d'une information de type date (en mois);
- « A » signifie qu'il s'agit d'une information de type date (en année) ;
- « P » a plusieurs significations. Il importe de vérifier l'étiquette de la variable pour déterminer avec précision le genre de variable dont il est question.

ppx : les dernières positions se rapportent à l'identification de la question. De façon générale, **pp** est un numéro séquentiel attribué à la variable et **x** est un indicateur alphabétique séquentiel pour une série de variables similaires ou pour les variables à choix multiples.

3. La structure et le contenu des fichiers

L'ensemble des questions posées durant la collecte se trouvent dans la base de données, exception faite des variables permettant d'identifier un ou plusieurs membres du ménage. La grande majorité des variables ont fait l'objet d'une validation transversale, alors que certaines ont également été validées de façon longitudinale⁴.

Un seul instrument de mesure a été utilisé au volet 2019 et il s'adressait seulement au jeune (QELJ). Afin de faciliter l'utilisation de la banque de données, le tableau 1 présente la structure des fichiers inclus dans la banque de données du volet 2019. Le tableau 2 décrit les sections du questionnaire QELJ.

Tableau 1. Fichiers inclus dans la banque de données du volet 2019 de l'ELDEQ (E22)

Nom du fichier	Nombre de lignes contenues dans le fichier	Données contenues dans le fichier
QELJ2201	1 323	Données colligées à l'aide du QELJ et concernant le jeune.
INDI2211	1 323	Variables dérivées construites par les partenaires et l'équipe de l'ELDEQ à partir des données issues du QELJ. <i>Elles sont présentées en détail dans la partie B du document des « Variables dérivées de l'ELDEQ – 1998-2019⁵ ».</i>
POIDS2201	1 323	Variable de poids échantillonnaux généraux transversaux relative aux données recueillies sur le ménage (PEGENT22 ⁶). <i>Le choix d'une pondération appropriée dépend des variables considérées dans le modèle statistique. Pour en connaître davantage sur la création et l'utilisation optimale de la variable de pondération, référez-vous au document « Pondération et analyse de la non-réponse des données du volet 2019⁷ ».</i>

4 Pour un aperçu des méthodes de validation utilisées, consulter le numéro 12 (section 2) des volumes 1 et 2 de la collection ELDEQ disponible sur le site de l'ELDEQ (www.iesuisieserai.stat.gouv.qc.ca) sous l'onglet « Publications ».

5 Le document « Variables dérivées de l'ELDEQ – 1998-2019, Partie B » est disponible sur le site de l'ELDEQ, sous l'onglet « Documentation technique ».

6 Il est à noter que, comme au volet 2017, une seule pondération a été calculée en prévision de produire davantage de pondérations sur mesure aux diverses analyses. N'hésitez pas à communiquer avec l'équipe de l'ELDEQ si vous désirez obtenir de telles pondérations.

7 Le document « Pondération et analyse de la non-réponse des données du volet 2019 » est disponible sur le site de l'ELDEQ, sous l'onglet « Documentation technique ».

Tableau 2. Composition du fichier QELJ2201 de la banque de données du volet 2019 de l'ELDEQ, jeunes d'environ 21 ans

Questionnaires et sections	Lettres indiquant la section du questionnaire informatisé ¹
Structure et mouvement familiaux	MF
Relations interpersonnelles ²	RI/IDG/VI
Choix de carrière	CC
Parcours et engagement scolaires	SC
Stress	ST
Santé et bien-être	HL
Intérêt politique	PO
Parcours socioprofessionnels	TV
Situation financière	SI
Habitudes de vie	HD
Sommeil	SO

1. Ces lettres se retrouvent à la 2^e et 3^e position du nom des variables.
 2. Cette section comprend 4 questions sur l'identité de genre (identifiées par IDGx dans les dernières positions du nom de la variable) et une question sur la victimisation (identifiée par les lettres VI)
- Source : Institut de la statistique du Québec, ELDEQ 1998-2019.

Dans une optique longitudinale, il est conseillé de se référer aux questionnaires⁸ de chacun des volets pour s'assurer de la comparabilité des questions dans le temps. En effet, certaines questions ou certains choix de réponse peuvent avoir été modifiés, inversés ou supprimés d'un volet à l'autre.

Comme les volets précédents, chacun des fichiers contient une variable nommée IDME (numéro de ménage). Cette variable doit absolument être utilisée lors de la fusion de deux fichiers ou plus. Tous les fichiers sont constitués d'une ligne par sujet.

Finalement, tous les fichiers ont été développés avec le logiciel SAS 9.4 sur LINUX. Un fichier de formats SAS est également fourni. Le fichier de format permet d'obtenir le catalogue des formats sur n'importe quelle plateforme (par exemple sur WINDOWS ou LINUX). La commande suivante peut être utilisée pour créer le catalogue des formats en SAS :

```
LIBNAME t22 "le chemin du répertoire où se trouve vos fichiers";

PROC FORMAT cntlin=t22.sas_fmts library =t22.formats; RUN;
PROC CATALOG catalog=t22.formats; copy out=work.formats; RUN; QUIT;
```

Une version SPSS WINDOWS (.sav) des fichiers est également fournie.

⁸ Les questionnaires de l'étude sont disponibles sur le site de l'ELDEQ, sous l'onglet « Outils de collecte ».

4. Remarques et mises en garde

4.1. Volet 2019

Mode de collecte – Fichier QELJ2201

La collecte de données a été réalisée en multimode (web ou téléphonique). Le web était privilégié mais le jeune avait la possibilité de répondre aux questions au téléphone. Une variable (VQJNMODE) est disponible pour identifier le mode de collecte. Certaines questions pourraient présenter un biais de désirabilité sociale. Des analyses supplémentaires seront nécessaires dans certains cas.

Âge du jeune et mois de passation – Fichier QELJ2201

La variable AGE22 identifie l'âge en années du jeune lors de la passation du questionnaire. Une variable qui identifie le mois de passation du questionnaire est aussi présente dans le fichier (VQJNDATM).

Indice de masse corporelle – Fichier QELJ2201

L'indice de masse corporelle (IMC) a été dérivé à partir du poids et de la taille déclarés par le jeune. La classification standard internationale chez les adultes a été utilisée. Deux variables sont donc disponibles dans le fichier QELJ2201 (VQJNIMCA et VQJNIMCAC). Il faut noter qu'au volet 2019, il n'a pas été demandé aux répondantes si elles étaient enceintes au moment de l'entrevue. Il n'était donc pas possible d'exclure les répondantes enceintes du calcul de l'IMC.

Codification des emplois – Fichier QELJ2201

Une variable VTVND01 est disponible dans le fichier QELJ2201 et présente la codification de l'emploi principal du jeune. Cette variable est basée sur la classification nationale des professions (CNP-2016) en 10 catégories. En raison de petits effectifs, la catégorie « Gestion » a été regroupée avec la catégorie « Indéterminé ».

Horaire des cours la fin de semaine – Fichier QELJ2201

Les variables VSCNQ13BM1 à VSCNQ13BM3 indiquent si le répondant a des cours actuellement la fin de semaine. Habituellement, lorsqu'un répondant oublie de répondre à certaines questions, un avertissement apparaît et il a le choix de continuer sans répondre ou de compléter sa réponse. L'avertissement pour ces questions (horaire des cours en semaine et fin de semaine) a été oublié lors de la programmation du questionnaire en ligne. Ainsi, plusieurs jeunes ont seulement répondu à la partie de la question portant sur les cours en semaine. Cela a introduit de la non-réponse partielle supplémentaire. Comme la non-réponse à ces questions lors du volet 2017 n'était pas très importante, on peut s'attendre à ce que la réponse réelle lors du volet de 2019 soit en fait « Je n'ai pas de cours la fin de semaine ». Il est suggéré d'effectuer des analyses supplémentaires en utilisant les réponses du volet 2017 et celle des items sur les horaires en semaine avant d'imputer des valeurs aux données manquantes en 2019.

Statut de fumeur – Fichier QELJ2201 et INDI2211

Comparativement au volet 2017, il y a eu une erreur de filtre dans le questionnaire après la question HDV_Q2 et une erreur d'univers à la HDV_Q2f. Les répondants n'ayant pas répondu ou ayant répondu « Non » à la question HDV_Q2 (déjà fumé une cigarette au complet) ne devaient pas sauter la question HDV_Q2f (statut de fumeur actuellement) et l'univers de cette dernière devait être « Tous ». Pour éviter une trop grande proportion de données manquantes lors de la création de la variable dérivée VQEED29 (Type de fumeurs), les répondants ayant répondu « Non » à la question HDV_Q2 ont été considérés « Non-fumeurs ». Au volet 2017, la très grande majorité des répondants n'ayant pas fumé

une cigarette au complet au cours de leur vie ne fumait pas lors de l'entrevue (HDV_Q2f = « Jamais ») ce qui a justifié notre choix pour la création de la variable dérivée au volet 2019. Il est suggéré d'effectuer des analyses supplémentaires pour valider cette construction.

Genre de dettes – Fichier QELJ2201

À la question SIF_Q08, les précisions dans le choix «Autres, précisez» sont en grande majorité liées à une dette envers les parents/famille. Comme il s'agit d'une question à choix multiple, une nouvelle catégorie n'a pas été ajoutée pour éviter un biais potentiel mais ce choix de réponse sera tenu en compte au prochain volet. La liste des précisions fournies par les répondants peut être disponible aux utilisateurs sous certaines conditions.

Raison de ne pas avoir voté à l'élection provinciale de 2018 – Fichier QELJ2201

À la question POL_Q5, la catégorie «Autre - À l'ext. de la circonscription jour du vote» (code 8) a été ajoutée afin de diminuer le nombre de cas de la catégorie «Autre».

Variables ayant un taux de non-réponse partielle supérieur à 5 %

Le nombre de variables comportant un taux de non-réponse partielle supérieur à 5 % varie en fonction de l'instrument de collecte et de la section observée. Ces variables doivent être traitées avec prudence. Le document « Pondération et analyse de la non-réponse des données du volet 2019 » analyse la non-réponse partielle pour les variables qui posent problème.

4.2. Modifications aux variables des années antérieures

Des corrections ont été apportées à certains fichiers de la banque de données depuis la dernière livraison. Voici la liste exhaustive de ces modifications.

Il est de la responsabilité des utilisateurs d'utiliser les données les plus à jour.

Volet	Fichier	Nom des variables visées par la modification, l'ajout ou la suppression	Commentaires ou justifications
2010 (E13)	QAAENS	MAEIQ08	Modification du libellé de la variable
	QIE	MDCNQ1	Modification du libellé de la variable
2011 (E14)	QAAM	NQMMQ36D	Modification du libellé de la variable
	QAAP	NQPJQ03D	Modification du libellé de la variable
	QPABS	NQPAQ03D	Modification du libellé de la variable
	QAAENS	NWE1Q16, NWE2Q16	Modification du libellé de la variable
	QIE	NDCNQ1	Modification du libellé de la variable
2013 (E16)	QELJ	PDCNQ1	Modification du libellé de la variable
		PEVNQ1B à PEVNQ1N	Correction du format des variables
	QAAM	PQMMQ24D	Modification du libellé de la variable
2015 (E18)	QAAM	RQMMQ14D	Modification du libellé de la variable
	QELJ	RDCNQ1	Modification du libellé de la variable
		RPUNQ7A	Correction du format de la variable (fichier SPSS seulement)
2017 (E20)	QELJ	TQJNLANG	Ajout d'un libellé
		TSINQ02	Correction pour un répondant afin de correspondre au reclassement du « Autre » qui avait été fait à la TSINQ01
		TTVNQ34A	Modification du libellé de la variable