

**ÉTUDE LONGITUDINALE DU DÉVELOPPEMENT
DES ENFANTS DU QUÉBEC – (ELDEQ 1998-2021)**



**Guide de l'utilisateur de la
banque de données du volet 2021**

Jeunes âgés d'environ 23 ans

**Direction principale des statistiques sociales et de santé (DPSSS)
Institut de la statistique du Québec**

Juin 2022

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 2021 (E24)**. 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 2021 de l'*Étude longitudinale du développement des enfants du Québec* (ELDEQ) a été réalisé auprès de 1368 jeunes¹ du milieu du mois de février à la fin juin 2021. Ainsi, au moment de cette collecte, les jeunes² avaient environ 23 ans.

La banque de données du volet 2021 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 2021.

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 :

X S E C Q ppx, où :

X : 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) ;
« X »	correspond au volet 2021 (environ 23 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 2021 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 2021. Le tableau 2 décrit les sections du questionnaire QELJ.

Tableau 1. Fichiers inclus dans la banque de données du volet 2021 de l'ELDEQ (E24)

Nom du fichier	Nombre de lignes contenues dans le fichier	Données contenues dans le fichier
QELJ2401f	1368	Données colligées à l'aide du QELJ et concernant le jeune.
INDI2411f	18	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 – 1998-2021⁵ ».</i>
POIDS2401	1368	Variable de poids échantillonnaux généraux transversaux relative aux données recueillies sur le ménage (PEGENT24 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 2021⁷ ».</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 – 1998-2021, 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 2021 » est disponible sur le site de l'ELDEQ, sous l'onglet « Documentation technique ».

Tableau 2. Composition du fichier QELJ2401 de la banque de données du volet 2021 de l'ELDEQ, jeunes d'environ 23 ans

Questionnaires et sections	Lettres indiquant la section du questionnaire informatisé ¹
Structure et mouvement familiaux	MF
Relations interpersonnelles ²	RI/IDG
Parcours et engagement scolaires	SC
Stress et résilience	ST
Santé et bien-être	HL
Relation conjugale et histoire reproductive	HL
Parcours socioprofessionnels	TV
Situation financière	SI
Habitudes de vie	HD
Activités physiques et de loisirs	AP
Sommeil	SO
Confiance envers les institutions	CE
Engagement social et communautaire - bénévolat	ES

1. Ces lettres se retrouvent à la 2^e et 3^e position du nom des variables.

2. Cette section comprend 21 questions sur les relations interpersonnelles (identifiées par les lettres RI) et 1 question sur l'identité de genre (identifiées par IDGx dans les dernières positions du nom de la variable).

Source : Institut de la statistique du Québec, ELDEQ 1998-2021.

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 t24 "le chemin du répertoire où se trouve vos fichiers";
```

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

```
PROC FORMAT cntlin=t24.sas_fmfs library =t24.formats; RUN;  
PROC CATALOG catalog=t24.formats; copy out=work.formats; RUN; QUIT;
```

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

4. Remarques et mises en garde

4.1. Volet 2021

Mode de collecte – Fichier QELJ2401f

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 (XQJNMODE) 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 QELJ2401f

La variable AGE24 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 (XQJNDATM).

Indice de masse corporelle – Fichier QELJ2401f

L'indice de masse corporelle (IMC) a été dérivé à partir du poids et de la taille déclarés par le jeune. Les répondantes enceintes étaient exclues du calcul de l'IMC. La classification standard internationale chez les adultes a été utilisée. Deux variables sont donc disponibles dans le fichier QELJ2401 (XQJNIMCA et XQJNIMCAC).

Codification des emplois – Fichier QELJ2401f

Une variable XTVND01 est disponible dans le fichier QELJ2401 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 ».

Statut de fumeur – Fichier QELJ2401f et INDI2411f

Tout comme au volet 2019, les répondants ayant répondu « Non » à la question HDV_Q2 ont été considérés « Non-fumeurs ».

Nombre de contact avec les amis par téléphone - Fichier QELJ2401f

Après la variable XRINQ15B, la variable ERR_XRINQ15b a été ajoutée pour identifier les répondants qui ont vu une erreur de traduction dans la question XRINQ15B en anglais, avant que celle-ci n'ait pas été corrigée. Des analyses n'ont pas montré de différence ou risque de biais potentiel.

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 2021 » 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
2013 (E16)	QELJ	PTRN8EM1 à PTRN8EM11	Modification du libellé de la variable
2015 (E18)	QELJ	RTRN8EM1 à RTRN8EM11	Modification du libellé de la variable
2017 (E20)	INDI	TQEED35L	Ajout d'une variable dérivée activité physique de loisir suite à une nouvelle définition de la variable
2019 (E22)	QELJ	VHLNQ01	Renommée pour VHLNQ1
2019 (E22)	QELJ	VTVNQ31ABA à VTVNQ31HBM	Corrections apportées au dénominateur des variables
2019 (E22)	QELJ	VTVNQ35ABA à VTVNQ35HBM	Corrections apportées au dénominateur des variables
2019 (E22)	INDI	VQEED35	Changement de nom de variable pour VQEES41
		VQEES35	Changement de nom de variable pour VQEED41