

MARS 2024



STRUCTURE ET VARIABLES DU FICHER BALSAC

Service aux chercheurs

PRÉPARÉ PAR
Häxan Bondu

PROJET BALSAC
Université du Québec à Chicoutimi
555 boulevard de l'Université
Chicoutimi (Québec)
G7H 2B1
<https://balsac.uqac.ca/>
Courriel : balsac@uqac.ca

Introduction aux données BALSAC

Le fichier de population BALSAC est une base de données généalogiques en développement à l'Université du Québec à Chicoutimi (UQAC) depuis les années 1970.

Ce présent document se veut une description générale du fichier BALSAC, de sa structure et de son contenu pour les chercheuses et les chercheurs qui voudraient y accéder. Il s'agit d'une mise à jour du document I-C-245, Structure de la base de donnée BALSAC et contenu des principales tables (2012) de Michèle Jomphe.



1. Données disponibles

Le fichier de population BALSAC est construit principalement à partir des actes de l'état civil du Québec (naissances, mariages et décès). Ces actes ont été informatisés et reliés entre eux grâce à une méthodologie de jumelage des données nominatives permettant la reconstitution automatique des familles et des lignées généalogiques sur quatre siècles (début 17^e siècle jusqu'à la deuxième moitié du 20^e siècle) dans l'ensemble du Québec. Cela veut dire qu'avec le fichier BALSAC, il est possible de naviguer facilement d'une famille à une autre, d'une génération à une autre, et ce, tout en consultant une variété d'informations (prénom, nom, sexe, profession, âge à un événement, date d'événements, lieu d'événements, etc.). BALSAC, dont le développement se poursuit, contient aujourd'hui plus de 5 millions d'actes d'état civil sur l'ensemble du Québec. Ces actes contiennent des données sur plus de 6 millions d'individus.

Les biographies individuelles sont complètes jusqu'en 1971 pour la région du Saguenay-Lac-St-Jean. Pour les autres régions, elles sont complètes jusqu'en 1800 (données du RPQA accessibles via l'IMPQ décrite plus bas), partielles jusqu'en 1849 et en cours de traitement pour la période 1850-1916. Certaines données tirées des recensements canadiens de 1851 à 1921 sont aussi accessibles (données du CIEQ via l'IMPQ).

2. Variables et jeux de données

Les chercheuses et chercheurs désirant obtenir des données provenant de BALSAC doivent définir les variables nécessaires à la réalisation de leur projet. Le prêt de données peut prendre la forme d'un fichier généalogique de base ou personnalisé ou d'un jeu de données entièrement personnalisé.

- Le fichier généalogique de base comprend les variables suivantes : ID individu ; Sexe de l'individu ; ID du père de l'individu ; ID de la mère de l'individu ; Lieu de mariage des parents ; Date de mariage des parents ; Statut migratoire de l'individu ; Lieu d'origine de l'individu.
- Le fichier généalogique personnalisé contient des variables supplémentaires : Date de naissance de l'individu ; Date du premier mariage de l'individu ; Date de décès de l'individu ; Lieu de naissance de l'individu ; Lieu du premier mariage de l'individu ; Lieu de décès de l'individu et/ou des variables calculées pour chaque individu : Nombre de généalogies où il apparaît ; Nombre d'occurrences ; Contribution génétique totale ; Génération minimum ; Génération maximum ; Génération moyenne. Pour obtenir ces informations, vous devrez fournir une brève justification pour chacune des variables demandées.

Il est aussi possible de demander un jeu de données entièrement personnalisé. Vous trouverez la liste des variables pouvant être demandées à l'Annexe A de ce document.

2.1 DONNÉES CONFIDENTIELLES

Les projets de recherche qui concernent des données confidentielles sont ceux qui impliquent des données de naissances, mariages et décès de moins de 100 ans. Votre demande d'accès doit être accompagnée d'une approbation de votre comité d'éthique à la recherche (CER) institutionnel à jour si celle-ci concerne des données confidentielles.

2.2 DONNÉES DE L'IMPQ

Si les données que vous souhaitez obtenir sont disponibles sur le portail de l'Infrastructure intégrée des microdonnées historiques de la population du Québec (IMPQ), il vous sera suggéré de faire une demande d'accès directement auprès de l'IMPQ et de vérifier si l'accès au portail répond à vos besoins.

L'IMPQ contient les données d'état civil du Québec de 1621 à 1923 qui proviennent du Registre de population du Québec ancien (RPQA) et du fichier BALSAC. Il contient aussi les microdonnées tirées des recensements canadiens de 1851 à 1921 pour quelques villes (Québec et Trois-Rivières) et régions (Saguenay-Lac-St-Jean, Côte-Nord, Gaspésie) du Québec.

3.Échéancier pour une demande d'accès aux données

1

Présentation de votre demande d'accès

Une demande d'accès aux données doit être remplie sur le site de BALSAC.

2

Prise de contact de BALSAC

À cette étape, on vous contactera, notamment pour s'assurer que les informations demandées soient exactes et pour modifier ou compléter votre demande d'accès, au besoin. Si votre demande comporte des données confidentielles, on vous demandera d'envoyer votre certificat d'approbation du comité d'éthique de la recherche (CER) de votre institution.

3

Analyse de votre demande d'accès par le comité directeur de BALSAC

Cette procédure prend en général autour de 4 semaines. Si le comité d'évaluation des demandes d'accès soulève des questions, vous serez recontacté.

4

Production de votre jeu de données

Si votre demande est acceptée, le jeu de données sera produit dans un délai de 3 à 4 semaines. Ces délais peuvent être plus longs si votre projet contient des reconstructions généalogiques ou des opérations de jumelage de données.

5

Signature de votre contrat

À cette étape, vous signerez le contrat avec BALSAC. Ce contrat confirme vos engagements éthiques auprès de BALSAC ainsi que les variables de votre jeu de données.

6

Transfert de vos données

Les fichiers à transférer sont transmis par l'intermédiaire du Dataverse de l'UQAC.

7

Envoi de votre facture

La facturation est établie selon les taux en vigueur qui sont présentés sur le site Web de BALSAC. La facture est envoyée à la chercheuse ou au chercheur responsable (ou à la chercheuse ou chercheur répondant.e dans le cas d'étudiant.es et de stagiaires postdoctoraux).

8

Renouvellement/modification de votre accès aux données (optionnel)

Lorsque la période pour laquelle un accès a été accordé viendra à échéance, vous devrez déposer une demande de prolongation au moins trente (30) jours avant l'échéance prévue.

Annexe A

Liste des variables

La table « Individus » contient l'ensemble des données personnelles d'un individu BALSAC. Elle cible les informations telles que les parents, le nom, le prénom, le sexe et les dates de naissance et de décès de l'individu. Chaque individu BALSAC possède une fiche « Individu ».

Individus	
Champ	Signification
IndID	Identifiant de l'individu (encodé)
PereID	Identifiant du père (encodé)
MereID	Identifiant de la mère (encodé)
NomID	Nom de l'individu
PrénomID	Prénom de l'individu
Sexe	Sexe de l'individu
LieuNaissance	Lieu de naissance de l'individu (paroisse, municipalité, province/état, pays)
JourNaissance	Jour de naissance de l'individu

STRUCTURE ET VARIABLES DU FICHIER BALSAC

MoisNaissance	Mois de naissance de l'individu
AnneeNaissance	Année de naissance de l'individu
LieuDeces	Lieu de décès de l'individu (paroisse, municipalité, province/état, pays)
JourDeces	Jour du décès de l'individu
MoisDeces	Mois de décès de l'individu
AnneeDeces	Mois de décès de l'individu
StatutNaissance	Variable d'identification pour les individus nés hors mariage (illégitimes) ou adoptés
StatutMigratoire	Variable d'identification pour les individus nés hors du Québec
Origine	Origine de l'individu (lorsqu'immigrant)

La table « Événements » contient les données des événements (naissance, union, décès, recensement) dans BALSAC. Elle compile l'ensemble des informations d'un événement donné.

Événements	
Champ	Signification
ActeID	Identifiant de l'acte (encodé)
TypeActe	Naissance (baptême), mariage ou décès (sépulture)
JourEvenement	Jour de l'événement
MoisEvenement	Mois de l'événement
AnneeEvenement	Année de l'événement
LieuID	Lieu de l'événement (paroisse, municipalité, province/état, pays)

La table « Mentions » concerne les individus mentionnés lors d'un événement. Elle compile des données telles que le prénom et le nom déclarés dans l'acte, leur lieu de résidence et leur profession au moment de l'événement, etc.

Mentions individuelles dans un événement (acte)	
Champ	Signification
ActeID	Identifiant de l'acte (encodé)
IndID	Identifiant de l'individu (encodé)
Role	Rôle de l'individu dans l'acte
NomID	Nom de l'individu (tel qu'écrit dans l'acte)
PrenomID	Prénom de l'individu (tel qu'écrit dans l'acte)
Sexe	Sexe de l'individu
ResidenceID	Résidence de l'individu déclarée dans l'acte (paroisse, municipalité, province/état, pays)
ProfessionID	Profession de l'individu déclarée dans l'acte
Origine	Origine (si déclarée dans l'acte)
EtatMatrimonial	État matrimonial déclaré dans l'acte (s'applique seulement aux sujets des actes)

STRUCTURE ET VARIABLES DU FICHIER BALSAC

Signature	Présence ou absence de la signature de l'individu dans l'acte
Presence	Mention de la présence de l'individu (présent, absent, décédé)
Age	Âge déclaré dans l'acte (s'applique seulement aux sujets de l'acte)