

FORMATION 1. FULL PROGRAM / PROGRAMME COMPLET

Bootcamp in administrative data analysis (BADA) – Full program (in person at the QICSS, May 4-8, 2026)

The BADA is a bilingual, intensive bootcamp-style training combining theory and practice to master the analysis of Canadian administrative data. Develop expertise in Statistics Canada's main databases (tax data, survey-admin data linkages, and multi-jurisdictional data integration platforms) through morning lectures and hands-on afternoon lab workshops. Participants will learn analysis and programming techniques specific to longitudinal administrative data by working with real administrative databases.

Prerequisites: Active research project at the QICSS and valid security clearance; intermediate level in Stata or equivalent software.

Bootcamp en analyse de données administratives (BADA) - Programme complet (présentiel au CIQSS, 4-8 mai 2026)

Le BADA est une formation bilingue et intensive de type « bootcamp » combinant théorie et pratique pour maîtriser l'analyse des données administratives canadiennes. Développez une expertise des principales bases de données de Statistique Canada (données fiscales, couplages d'enquêtes et plateformes d'intégration de données multi-juridictionnelles) à travers des conférences matinales et des ateliers pratiques en laboratoire l'après-midi. Les participants acquerront des techniques d'analyse et de programmation spécifiques aux données administratives longitudinales en travaillant avec de vraies bases de données administratives.

Prérequis : Projet de recherche actif au CIQSS et cote de sécurité valide; niveau intermédiaire en Stata ou logiciel équivalent.

FORMATION 2. ADVANCED STATA WORKSHOP / ATELIER STATA AVANCÉ

Advanced Stata programming applied to the analysis of massive and longitudinal administrative datasets [BADA Day 1] (in person at the QICSS, May 4, 2026, 9am to 5pm)

This full day of bilingual technical training is designed for individuals with a working knowledge of Stata who wish to develop advanced skills necessary for the analysis of large and complex longitudinal datasets. Specifically, this Stata programming training will benefit researchers who wish to analyse Canadian administrative datasets and related data linkages. The workshop will provide detailed examples of Stata codes as well as exercises that will facilitate data preparation, data management, and the completion of data validation and description steps necessary before performing regression analysis.

Prerequisite: Prior knowledge of Stata is necessary. This activity will not include an introduction to Stata.

Programmation avancée Stata appliquée à l'analyse de bases de données administratives massives et longitudinales [BADA Jour 1] (en personne au CIQSS, 4 mai 2026, 9h à 17h)

Cette journée complète de formation technique bilingue est destinée aux personnes ayant une connaissance pratique de Stata qui souhaitent acquérir les compétences avancées nécessaires à l'analyse de bases de données longitudinales, massives et complexes. Plus précisément, cette formation à la programmation Stata sera utile aux chercheurs qui souhaitent analyser des bases de données administratives canadiennes et couplages de données. L'atelier fournira des exemples détaillés de codes Stata ainsi que des exercices qui faciliteront la préparation et la gestion des données, ainsi que la validation et la description des données nécessaires avant d'effectuer des analyses par régression.

Prérequis : Des connaissances préalables de Stata sont nécessaires. Cette activité ne comprendra pas d'introduction à Stata.

FORMATION 3. INTRODUCTION TO ADMINISTRATIVE DATA ANALYSIS ET REVENU / INTRODUCTION À L'ANALYSE DE DONNÉES ADMINISTRATIVES ET REVENU

Introduction to administrative data analysis with a focus on income data [BADA Day 2] (hybrid at the QICSS, May 5, 2026, 9am-12am)

During this bilingual training session, we will introduce key concepts and basic features of administrative data, with a focus on Canadian tax data such as the Longitudinal Administrative Databank (LAD) and other datasets derived from the T1 Family Files (T1FF). Emphasis will be put on the longitudinal and relational nature of the data. The session includes a detailed description of variables on income, program participation, and wealth in tax data, as well as a brief overview of place of residence information. It will conclude with a discussion of tax data quality and population coverage.

Databases discussed: Longitudinal administrative data (LAD) and other tax datasets.

Session open to all.

Introduction à l'analyse des données administratives axée sur les données relatives au revenu [BADA jour 2] (hybride au CIQSS, le 5 mai 2026, 9h-12h)

Au cours de cette séance de formation bilingue, nous présenterons les concepts clés et les caractéristiques fondamentales des données administratives, en mettant l'accent sur les données fiscales canadiennes telles que la Banque de données administratives longitudinales (DAL) et d'autres ensembles de données dérivés des fichiers familiaux T1 (T1FF). L'accent sera mis sur la nature longitudinale et relationnelle des données. La séance comprend une description détaillée des variables relatives au revenu, à la participation aux programmes et à la richesse dans les données fiscales, ainsi qu'un bref aperçu des informations sur le lieu de résidence. Elle se terminera par une discussion sur la qualité des données fiscales et la couverture populationnelle.

Bases de données discutées : Données administratives longitudinales (DAL) et autres données fiscales.

Séance ouverte à tous.

FORMATION 4. FAMILIES IN TAX DATA / FAMILLES ET DONNÉES FISCALES

Studying families using tax data [BADA Day 3] (hybrid at the QICSS, May 6, 2026, 9am-12am)

Canadian tax data includes detailed information on family structure and family relationships. In this session based on recent research in demography, sociology and economics, we will discuss what types of family dynamics can be observed cross-sectionally and longitudinally in Canadian administrative data, with a focus on data derived from personal tax records. Topics include couple dissolution and reconstituted families, same-sex couples, measures of fertility and the presence of children, extended kinship networks and intergenerational transmission.

Databases discussed: Longitudinal administrative data (LAD) and other tax datasets.

Session open to all.

Étude des familles à partir des données fiscales [BADA jour 3] (hybride au CIQSS, 6 mai 2026, 9h-12h)

Les données fiscales canadiennes comprennent des informations détaillées sur la structure familiale et les relations familiales. Dans cette session basée sur des recherches récentes en démographie, sociologie et économie, nous discuterons des types de dynamiques familiales qui peuvent être observés de manière transversale et longitudinale dans les données administratives canadiennes, en mettant l'accent sur les données issues des dossiers fiscaux personnels. Les thèmes abordés comprennent la dissolution des couples et les familles recomposées, les couples de même sexe, les mesures de la fécondité et la présence d'enfants, les réseaux familiaux élargis et la transmission intergénérationnelle.

Bases de données discutées : Données administratives longitudinales (DAL) et autres données fiscales.

Séance ouverte à tous.

FORMATION 5. RECORD LINKAGES / COUPLAGE DE DONNÉES

Record linkages and data integration [BADA Day 4] (hybrid at the QICSS, May 7, 2026, 9am-12am)

This session will allow participants to develop a familiarity with record linkage approaches. The main characteristics, properties, and limitations of record linkages will be presented. We will focus on linkages between survey and tax data, federal immigration data (IMDB) and tax data. We will also provide an overview of various linkage platforms that integrate a large number of administrative, survey, and Census datasets, including in some cases provincial data. The potential and limitations of these platforms will be discussed, with a focus on the Education and Labour Market Longitudinal Platform (ELMLP), the CEEDD/BEAM business-level data and CanCHEC health data. The session will also include practical examples guiding participants who wish to perform and validate the quality of record linkages.

Databases discussed: Linkages between surveys and tax data, the Longitudinal Immigration Database (IMDB), the Education and Labour Market Longitudinal Platform (ELMLP), the Canadian Employer-Employee Dynamics Database (CEEDD/BEAM) and the Canadian Census Health and Environment Cohorts (CanCHECs).

Session open to all.

Couplages et intégration des données [BADA jour 4] (hybride au CIQSS, 7 mai 2026, 9h-12h)

Cette session permettra aux participants de se familiariser avec les approches de couplage entre bases de données. Les principales caractéristiques, propriétés et limites des méthodes de couplage seront présentées. Nous nous concentrerons sur les liens entre les données d'enquête et les données fiscales, les données fédérales sur l'immigration (IMDB) et les données fiscales. Nous fournirons également un aperçu des différentes plateformes de couplage qui intègrent un grand nombre de bases de données administratives, d'enquête et de recensement, y compris, dans certains cas, des données provinciales. Le potentiel et les limites de ces plateformes seront discutés, en mettant l'accent sur la Plateforme longitudinale sur l'éducation et le marché du travail (PLEMT), les données d'entreprises du BDCEE/MAEE et les données sur la santé du CSERCan. La séance comprendra également des exemples pratiques destinés à guider les participants qui souhaitent effectuer et valider la qualité des couplages de données.

Bases de données discutées : Couplage entre données d'enquête et données fiscales, la Banque de données longitudinales sur les immigrants (BDIM), la Plateforme longitudinale entre l'éducation et le marché du travail (PLEMT), la Base de données

canadienne sur les dynamiques des employeurs et employés (BCDEE/MAEE) et les Cohortes santé et environnement du recensement canadien (CSERCan).

Séance ouverte à tous.