

Introduction pratique à l'utilisation du logiciel STATA

Mercredi 24 et jeudi 25 novembre 2010, de 9h30 à 16h30

Lieu de la formation

Laboratoire du CIQSS, 3535 chemin Queen-Mary, bureau 420, Montréal.

Objectifs

Présenter aux étudiants une introduction à la syntaxe de base du progiciel Stata. À la fin de cet atelier pratique, les étudiants seront capables d'utiliser la programmation pour manipuler des fichiers, transformer des données et réaliser des traitements statistiques de base afin d'obtenir des résultats qui seront visualisés dans le journal du programme. La formation se déroulera entièrement en laboratoire informatique. Les participants auront à réaliser des exercices qui utilisent le fichier de microdonnées à grande diffusion du recensement 2001 et de l'Enquête sociale générale 2001 de Statistique Canada.

Responsable

María Constanza Street, candidate au doctorat en démographie
Centre - Urbanisation Culture Société, Institut national de la recherche scientifique.

Contenu

Plan détaillé

1^{er} jour

Séance de l'avant-midi

- L'environnement de Stata
- La logique d'une séance sous Stata
 - Les programmes « do file »
 - Le journal du programme « log file »
 - L'utilisation de l'éditeur de texte
- Préparation d'une séance de travail : version, capture, set memory, clear, set more, log using, log close, local
- Ouvrir et sauvegarder une base de données : use, save
- La syntaxe de base d'une instruction
- Exploration d'une base de données : browse, edit, describe, list, codebook, count

Séance de l'après-midi

- Les types de variables : numériques « numeric » et alphanumériques « string »
- Définition du nom et du descripteur d'une variable : label define, label values, label variable, label copy, label list, label drop.
- Changement du nom d'une variable : rename, renpfix
- Les opérateurs logiques.
- Poser des conditions à une instruction : « if »
- Création d'une nouvelle variable : generate, replace, egen
- Tableau de fréquences et statistiques descriptives : tabulate, summarize, tabstat, table
- Exécuter une instruction pour un groupe d'observations : bysort

2^e jour

Séance de l'avant-midi

- Transformation d'une variable : recode
- Traitement des valeurs manquantes « missing values » : mvencode
- Boucles : forvalues, foreach, while
- Transformation d'une variable alphanumérique en variable numérique : destring, encode
- Transformation d'une variable numérique en variable alphanumérique : tostring, decode

Séance de l'après-midi

- Concaténation, fusion et jointure des bases de données : merge, append, joinby.
- Remodeler une base de données : reshape long, reshape wide
- Éliminer ou conserver une variable / une observation : drop, keep
- Ordonner les variables : order, aorder
- Ordonner les observations : sort, gsort, asort

- Récapitulation : Quelques astuces pour rédiger un programme de traitement de données destiné à l'exécution avec Stata.