

CURSO DE INTRODUÇÃO AO STATA

Professor: Ernesto Friedrich de Lima Amaral (DCP/UFMG)

amaral@fafich.ufmg.br

www.ernestoamaral.com/stata20092b.html

Data: 05/12/2009

Horário: 8:00 às 12:00 e 14:00 às 18:00

Público-alvo: estudantes e professores do Departamento de Ciência Política (DCP) da Faculdade de Filosofia e Ciências Humanas (FAFICH) da Universidade Federal de Minas Gerais (UFMG).

APRESENTAÇÃO:

Esse curso tem como objetivo introduzir a utilização do programa computacional Stata, o qual apresenta uma vasta possibilidade de exploração de dados e modelagem estatística. Com os conhecimentos adquiridos nesse curso, o aluno poderá analisar dados provenientes de diferentes fontes, tanto de bancos construídos com dados secundários, como de bancos elaborados pelo próprio pesquisador. Dessa forma, espera-se que após o curso os alunos poderão trabalhar com bases de dados no desenvolvimento de suas pesquisas, dissertações ou teses.

O curso está organizado de forma a introduzir ao aluno as principais ferramentas de programação do Stata. Esse instrumental será utilizado na análise de dados da Pesquisa Nacional por Amostra de Domicílios (PNAD), com ferramentas de gerenciamento de dados, além de uma série de comandos estatísticos e de elaboração de gráficos. Após receberem explicação e exemplificação de comandos de programação, os alunos poderão desenvolver seus próprios programas em Stata.

Primeiramente, a intenção é de propiciar ao aluno o conhecimento básico de como efetuar as principais manipulações de dados no Stata, desde a montagem do banco de dados até a exploração de estatísticas descritivas. Haverá então explicação sobre criação de nomes para variáveis e suas categorias, geração de novas variáveis, elaboração de tabelas e estatísticas descritivas, e geração de gráficos. Posteriormente, haverá um maior detalhamento de como os dados podem ser gerenciados no ambiente Stata. Será explicado como selecionar ou excluir variáveis do banco de dados, anexar dois ou mais bancos, recodificar um grupo de variáveis, reorganizar os dados em diferentes formatos, criar bancos de dados agrupados por categorias de variáveis, e analisar os dados com testes de correlação e significância. Comandos de regressão estatística e de análise de resíduos não são o foco desse curso. No entanto, poderá ser realizada uma rápida exemplificação das possibilidades existentes de análises inferenciais com o Stata.

MATERIAL:

O material que servirá de base para o acompanhamento do curso será disponibilizado na página www.ernestoamaral.com/stata20093.html, contendo instruções, bancos de dados, documentos, comandos e exemplos em Stata.

CONTEÚDO DO CURSO:

Introdução: explicação da estrutura do curso; características gerais do programa computacional Stata.

Importação dos dados: diferentes formas de importar dados para o programa Stata; visualização de dados.

Modificação dos dados: criação de nomes para variáveis e suas categorias; geração de novas variáveis.

Análise exploratória dos dados: comandos para elaborar tabelas e estatísticas descritivas.

Gerenciamento de dados: seleção de variáveis; exclusão de variáveis ou de categorias; inclusão de variáveis ou de observações de outros bancos.

Gráficos: ferramentas para geração de gráficos, tais como *scatter plots*, histogramas, *box-plot*, gráficos de linha.

Gerenciamento de dados: reorganização do arranjo (formato) dos bancos de dados; criação de bancos de dados agrupados por categorias de variáveis.

Gerenciamento de dados: utilização do comando *foreach* para criação e recodificação de grupos de variáveis; alocação de informações entre registros.

Testes de significância: testes de correlação e significância entre variáveis (*t-test*, qui-quadrado).

Regressão estatística: rápida exemplificação de modelos de mínimos quadrados ordinários (*ordinary least squares*), modelos logísticos e modelos multinomiais.

BIBLIOGRAFIA:

Agresti, Alan, e Barbara Finlay. 1999. **Statistical Methods for the Social Sciences**. Upper Saddle River, NJ: Prentice Hall.

Hamilton, Lawrence C. 1992. **Regression with Graphics: A Second Course in Applied Statistics**. Belmont, CA: Duxbury Press.

Long, J. Scott, e Jeremy Freese. 2001. **Regression Models for Categorical Dependent Variables Using Stata**. College Station, TX: Stata Press.

Powers, Daniel A., e Yu Xie. 2000. **Statistical Methods for Categorical Data Analysis**. San Diego, CA: Academic Press.

Raudenbush, Stephen W., Anthony S. Bryk, Yuk Fai Cheong, e Richard T. Congdon, Jr. 2001. **HLM 5: Hierarchical Linear and Nonlinear Modeling**. Lincolnwood, IL: SSI Scientific Software International.

Raudenbush, Stephen W., e Anthony S. Bryk. 2002. **Hierarchical Linear Models: Applications and Data Analysis Methods**. Thousand Oaks, CA: Sage Publications.

Singer, Judith D., e John B. Willett. 2003. **Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence**. New York, NY: Oxford University Press.

Wonnacott, Thomas H., e Ronald J. Wonnacott. 1986. **Regression: A Second Course in Statistics**. Malabar, FL: Krieger Publishing Company.