Seminário

Modjissola Cheryl Adoutan

Sugere-se que na apresentação do seminário sejam executados os comandos um a um, com comentários e interpretação das saídas de cada comando.

Informações sobre o banco de dados

Os dados se referem a 231 idosos residentes no município de São Paulo no ano de 2010.

As variáveis que compõem o banco são:

|  |  |  |
| --- | --- | --- |
| Nome da variável | Detalhamento | Códigos |
| id | Número de identificação do participante |  |
| sexo | Sexo  | 1-masculino2-feminino |
| imc | índice de massa corporal |  |
| dcnt | Número de doenças crônicas não transmissíveis |  |
| triglic | Concentração de triglicerides |  |

Sugestão de análise do banco

1. Abrir o arquivo no Excel e transformar o banco de dados de xls para dta, utilizando o software Transfer
2. Abrir o arquivo no Stata
3. Criar um arquivo log
4. Descrever as variáveis (comando desc)
5. Criar uma nova variável sexo numérica (transformar a variável string em numérica)

Comandos: rsexo=real(sexo)

 desc rsexo

1. Construir um label (rótulo) para a variável rsexo
2. Fazer uma tabela de frequência das variáveis
3. Construir um box plot para a variável imc e investigar se existem valores aberrantes
4. Fazer a mesma investigação de valores aberrantes para a variável triglic
5. Resumir as variáveis quantitativas apresentando as medidas de tendência central e dispersão (comando sum)
6. Apresentar as variáveis imc, triglic e idade em gráficos apropriados (histogramas)
7. Avaliar se estas variáveis são distribuídas segundo uma distribuição normal (comando ladder)
8. Apresentar a variável imc em quatro partes utilizando os quartis.

Primeiro crie a nova variável, ex: gen imccat=imc

Use o comando recode para categorizar a variável imccat em 4 categorias.

1. Construir um rótulo para a nova variável imc
2. Criar uma variável para o triglicérides com duas categorias (abaixo de 150 mg/dL e 150 e mais mg/dL). Criar antes uma variável nova para triglicérides Ex: gen trigliccat=triglic
3. E agora, recodificar a nova variável em duas categorias.
4. Nomear estas categorias triglicérides normal e alto, respectivamente
5. Apresentar a tabela cruzada entre as variáveis sexo e triglicérides em duas categorias
6. Apresentar as porcentagens para a tabela do item 17 e o teste qui quadrado de Pearson para investigar associação. As variáveis estão associadas? Discutir os resultados
7. Fechar o arquivo log