

O Método da Thaís pretende fazer análises estatísticas em recorte amostral do banco de dados Inquérito de Saúde do Município de São Paulo em 2015. A ideia é recortar indivíduos de interesse (adultos, ambos os sexos, que possuísem respostas para IMC e diagnóstico de depressão/ansiedade) e explorá-los para investigação de associação entre sobrepeso, obesidade, ansiedade e depressão.

A descrição do método é bastante completa, contemplando até os detalhes da coleta, que pelo que eu entendi, fora realizada pela prefeitura. O texto está muito bem escrito e elucida o passo a passo da coleta e estruturação dos dados.

Não é possível saber o que foi esforço do pesquisador e o que já estava pronto no banco-fonte utilizado. Acho que o ponto principal, quais análises foram/serão feitas (regressões, tese de associação qui-quadrado, etc.), não foram descritos apropriadamente, com suas devidas fórmulas e descrição das VIs e da VD.

Acredito que você possa simplificar a descrição da amostra e direcioná-la apenas para seu estudo. Se você informar ao leitor que sua fonte de dados é o ISA-Capital 2015 e aonde obter tais dados e a documentação metodológica por trás dele, você não precisa replicar isso no seu artigo. Deste modo, você ganha espaço e fôlego para focar nos critérios do recorte da amostra para estudo. Sobre as variáveis de estudo, não fica claro se elas foram obtidas diretamente (se estavam prontas) ou se foi você quem fez, justamente por que não dá para separar o que foi esforço seu e o que não foi. Isso é facilmente resolvido explicando um exemplo da amostra, que pode ser apresentado em forma de tabela. Assim o leitor tem uma compreensão maior do que constitui seus dados e também o auxilia em uma possível reprodutibilidade. Em geral, acho que você poderia focar na explicação de cada uma das análises que você fará com estes dados: o que as compõe, qual o n para cada análise, qual a variável de interesse e explicar os por quês das variáveis independentes. Por fim, acho que você pode falar do Stata apenas uma vez: que ele é uma ferramenta estatística e foi utilizado para estruturação do conjunto de dados, construção de variáveis, análises estatísticas e o que mais for necessário.