

Estatísticas descritivas e técnicas de visualização na análise de textos

Alfredo José Barreto Luiz¹, Carolina Borges da Silva Luiz²

Resumo: *O uso de métodos quantitativos para a análise de textos possui inúmeras aplicações em diversas áreas, como na identificação de autores, na crítica e no estudo da história da literatura, no mapeamento de vocabulário e ainda na análise sociolinguística. A identificação e classificação automática de textos também é beneficiada pelas técnicas da linguística quantitativa. Por outro lado, sabe-se da dificuldade no ensino da estatística e dos métodos quantitativos, especialmente para alunos das áreas biológicas e humanas, não familiarizados às ciências exatas. Assim, o presente trabalho demonstra o uso da estatística descritiva e da visualização de dados aplicadas na análise de texto, poéticos e científicos, com o objetivo de oferecer ferramentas para profissionais que se deparem com a necessidade desse tipo de análise, mas também para disponibilizar exemplos de aplicação que possam facilitar o ensino dos conceitos estatísticos para alunos mais afeitos às ciências humanas e biológicas. O uso de estatísticas simples (como médias, frações e proporções) e de gráficos, se mostrou mais útil na distinção de poemas e resumos científicos do que apenas as suas características quantitativas absolutas.*

Palavras-chave: Linguística quantitativa; métodos quantitativos; visualização de dados.

¹ Engenheiro Agrônomo, Dr., pesquisador na Embrapa Meio Ambiente. E-mail: alfredo.luiz@embrapa.br.

² Historiadora, mestranda em História Social na FFLCH/USP. E-mail: cbsluiz@gmail.com.