

Caracterização numérica e gráfica de palavras no conto *A Cartomante*

Raiane Pereira e Sabrina Oliveira

Novembro 2025

1 Resumo

Este trabalho pretende demonstrar como métodos computacionais podem ser utilizados para analisar sistematicamente dados linguísticos, a fim de aproximar as técnicas quantitativas aprendidas na disciplina de *Ferramentas Computacionais de Modelagem* da interpretação linguística. No contexto dessa disciplina, a proposta dialoga diretamente com o objetivo de empregar recursos computacionais na exploração, visualização e interpretação de dados em diferentes contextos científicos.

Ao realizar uma análise exploratória do conto *A Cartomante*, o estudo evidencia o potencial das linguagens de programação — neste caso, R — para tratar textos como dados, permitindo operações que são difíceis de realizar manualmente, como contagem exata de ocorrências lexicais, identificação de padrões fonotáticos, cálculo de medidas estatísticas e produção de visualizações robustas. Isso não apenas exemplifica o uso prático das ferramentas apresentadas na disciplina, mas também demonstra como tais recursos fortalecem o diálogo entre ciências humanas e computacionais.

A importância da proposta também se manifesta na sua capacidade de revelar regularidades linguísticas que não são imediatamente perceptíveis por leitura intuitiva. A distribuição dos tamanhos das palavras, a identificação das palavras semânticas mais frequentes, os gráficos de estrutura fonotática (como a relação entre vogais e consoantes) e as representações visuais como treemaps e wordclouds oferecem uma visão ampla do comportamento lexical do texto. Tais análises facilitam o reconhecimento de tendências estilísticas, escolhas vocabulares, padrões

estruturais recorrentes e características gerais do português escrito.

Além disso, o trabalho contribui para consolidar habilidades fundamentais no uso do R, tais como: manipulação de dados; limpeza e normalização de dados textuais; criação de tabelas de frequência; construção de gráficos informativos e esteticamente claros; aplicação de funções estatísticas; uso de pacotes especializados (dplyr, ggplot2, wordcloud2, treemapify).

Essas competências são essenciais não apenas para análises linguísticas, mas para qualquer área que dependa de modelagem, padronização e visualização de dados, alinhando-se ao caráter transversal da disciplina.

Em termos mais amplos, o trabalho mostra como textos literários podem ser tratados como corpora estruturados, permitindo que pesquisadores explorem não apenas o conteúdo semântico, mas também propriedades formais e quantitativas que caracterizam o estilo de um autor ou a estrutura de um gênero. Essa abordagem fortalece práticas de pesquisa baseadas em dados, aproximando métodos da linguística computacional, análise exploratória e ciência de dados.

Assim, este estudo breve e introdutório pretende demonstrar a importância que reside tanto na aplicação concreta das ferramentas computacionais quanto na demonstração de como a modelagem quantitativa pode ampliar a compreensão de fenômenos linguísticos, servindo como exemplo sólido de interdisciplinaridade e de integração entre teoria, metodologia e prática analítica.