

ESTATÍSTICA DESCRITIVA: EXPLORANDO DADOS COM GGLOT2

XXVIII Encontro de Iniciação à Docência

Irlene dos Santos Rabelo, Eric Oliveira Rocha, Juvêncio Santos Nobre, Maria Jacqueline Batista

A Estatística é uma ciência que proporciona métodos adequados de coleta e tratamento de dados, de forma a permitir a extração e a interpretação de informações que forneçam o apoio necessário nas tomadas de decisão. Entretanto, a habilidade em apresentar estas informações de forma efetiva se faz tão importante quanto saber extrair informações relevantes a partir de um conjunto de dados brutos. Uma rápida consulta aos artigos, periódicos e revistas especializadas das mais diversas áreas de conhecimento mostra que a Estatística tem se confirmado como linguagem predominante na comunicação de resultados. Somando esses fatores à rápida expansão do Big Data, o aprendizado de ferramentas de linguagem estatística se faz cada vez mais imprescindível. Um exemplo de ferramenta estatística poderosa é o software R, por ser uma linguagem de programação gratuita e com alta expansibilidade através do uso de pacotes. Outro destaque do R é que este está se tornando linguagem padrão em projetos de Data Science. Este trabalho tem por objetivo o fornecimento de um guia prático para a elaboração e a apresentação de análises descritivas de dados. Partindo de um exercício proposto, são apresentadas as principais medidas resumo de posição e variabilidade e elaborado um script, usando a funcionalidade R Notebook, que abrange desde a criação do banco de dados até a elaboração de gráficos mais atraentes gerados com a função ggplot do pacote ggplot2. Este script pode ser adaptado a outros experimentos e ainda exportado para PDF, HTML, ou até mesmo arquivos do MS Word.

Palavras-chave: Estatística. Medidas descritivas. Programa R. ggplot2.