

ANÁLISE DA CORRELAÇÃO DAS VARIÁVEIS EM UMA BASE DE DADOS DE BOLSISTAS DA UFC

I Encontro de Avaliação Científica e Acadêmica

André Alves Bezerra, Gisele Azevedo de Araújo Freitas, Walter da Cruz Freitas Junior

Mapa de calor é uma ferramenta visual para análise e detecção de tendências além de permitir identificar discrepâncias em um conjunto de dados. Este estudo utiliza a biblioteca seaborn, escrita na linguagem python, para gerar mapas de calor e mensurar a correlação das variáveis na base de dados. A base de dados utilizada neste estudo contém informações de alunos bolsistas da UFC, fomentada pela EIDEIA (Escola Integrada de Desenvolvimento e Inovação Acadêmica). O objetivo do estudo é realizar previsões de desempenho destes estudantes bolsistas e poder inferir se estes serão reprovados em disciplinas ou irão evadir dos cursos. Porém na etapa atual, de pré-processamento dos dados, será verificado quais variáveis têm maior relevância na situação final do aluno. Cada amostra da base de dados contém mais de 40 variáveis, sendo muitas delas não numéricas. Nesta etapa, utilizamos apenas as variáveis numéricas, pois a ferramenta utilizada para gerar o mapa de calor permite apenas variáveis numéricas. Em um próximo passo, variáveis não numéricas serão convertidas em variáveis numéricas para serem utilizadas na ferramenta. Ao utilizar o mapa de calor apenas com as variáveis numéricas da base de dados, observou-se que há alta correlação entre as variáveis ira geral do semestre e ira individual do aluno. Espera-se obter melhores resultados quando compararmos também com as variáveis não numéricas.

Palavras-chave: mapas de calor. correlação. alunos bolsistas.