

APLICAÇÃO ANALYTICS COM ESTUDO DE CORRELAÇÃO: FATORES QUE INFLUENCIAM NA VARIAÇÃO DE PREÇOS DE ALIMENTOS

João Edegar Dematê¹

Bruno Ibias Pereira²

Thiago do Vale Pereira Livramento³

Fabio Dejalma de Souza⁴

Samuel Jiukoski da Silva⁵

Denis Rodrigues Magalhães⁶

Tulio Bonilla⁷

Ibsem Agrello Dias⁸

RESUMO

Em 8 de maio de 2019 os alunos de todos os semestres de 2019 do curso de Análise e Desenvolvimento de Sistemas participaram do III Aula Aberta. No evento estes foram organizados em equipes com no máximo 10 integrantes, para cada equipe foi entregue um problema de ordem quantitativa. Os professores indicaram um Dataset (conjunto de dados disponibilizados para pesquisa) para cada grupo com dois objetivos a serem alcançados. O primeiro, desenvolver uma rotina automatizada para captura, tratamento e carga dos dados de um site da web para um Banco de Dados Não Relacional; o segundo, extrair os dados, armazená-los em uma planilha e realizar um estudo estatístico de correlação para identificar quais variáveis apresentam relação direta com a variável alvo determinada no problema apresentado. Para desenvolver a rotina automatizada utilizamos a linguagem de programação Javascript, com as bibliotecas Node e React, para armazenar os dados foi utilizado o Banco de Dados MongoDB, o problema apresentado para nosso grupo tem origem nos dados extraídos no site <https://www.kaggle.com/jboysen/global-food-prices> sob o título de **Global FoodPrices** tendo como alvo a identificação das variáveis que influenciam no valor dos alimentos em três países da América do Sul (Bolívia, Colômbia e Peru). O dataset adquirido, contém mais de 740 mil linhas de preços de vários produtos obtidos nos mercados de países considerados em desenvolvimento. Os dados incluem informações sobre o país, o mercado, o preço do bem em moeda local, a quantidade de bem e o mês registrado. Na rotina desenvolvida foi possível identificar o potencial para a automatização dos processos de captura, extração e análise de dados como forma de identificar soluções para problemas que tenham origem dados disponíveis na internet. A partir da implementação desta rotina e do conhecimento das tecnologias aplicadas os integrantes da equipe identificaram diversas outras aplicações. Após execução da rotina que capturou e armazenou os dados foram identificadas e extraídas as variáveis independentes: a província ou o estado, o tipo de mercado, o mês e o ano, do dataset original e a variável dependente: o preço pago. Os testes de correlação ainda estão sendo aplicados, por isso não apresentamos aqui os dados da correlação entre as variáveis.

Palavras-chave: Correlação. Analytics. Preços de alimentos. Javascript.

¹ Graduando de Análise e Desenvolvimento de Sistemas. Faculdade CESUSC. E-mail: joaodematejr@gmail.com

² Graduando de Análise e Desenvolvimento de Sistemas. Faculdade CESUSC. E-mail: brunoibias@gmail.com

³ Graduando de Análise e Desenvolvimento de Sistemas. Faculdade CESUSC. E-mail: evilskull05@gmail.com

⁴ Graduando de Análise e Desenvolvimento de Sistemas. Faculdade CESUSC. E-mail: fabio.dejalmaledesouza@gmail.com

⁵ Graduando de Análise e Desenvolvimento de Sistemas. Faculdade CESUSC. E-mail: samueljiukoski@gmail.com

⁶ Graduando de Análise e Desenvolvimento de Sistemas. Faculdade CESUSC. E-mail: denis_primavera@hotmail.com

⁷ Graduando de Análise e Desenvolvimento de Sistemas. Faculdade CESUSC. E-mail: tulibonilla@gmail.com

⁸ Professor Mestre em Administração. Faculdade CESUSC. E-mail: ibsem.dias@gmail.com