SEMINÁRIO DE INICIAÇÃO CIENTÍFICA E TECNOLÓGICA DO IFG

Desenvolvimento de uma Ferramenta Automatizada para Leitura, Extração e Análise de Dados de Contas de Energia Elétrica em PDF

MENDONÇA, Gustavo Heitor Lima
Instituto Federal de Educação, Ciência e Tecnologia de Goiás, Câmpus Itumbiara

MARQUES, Leonardo Garcia
Instituto Federal de Educação, Ciência e Tecnologia de Goiás, Câmpus Itumbiara

As faturas de energia elétrica do Câmpus Itumbiara do Instituto Federal de Goiás (IFG) são disponibilizadas em PDF, e seus dados de consumo, demanda e valores precisam ser transcritos antes de qualquer análise ao longo do tempo. Este trabalho teve como objetivo desenvolver uma ferramenta em Python que lê essas faturas, extrai seus dados, armazena-os em banco de dados e os analisa de forma automatizada, para subsidiar a gestão energética do câmpus. Foram utilizadas 65 faturas mensais, de março de 2020 a agosto de 2026, obtidas dos processos de pagamento do câmpus, de acesso público no Sistema Unificado de Administração Pública (SUAP). A extração utiliza a biblioteca PyMuPDF para obter cada palavra com suas coordenadas na página: palavras-âncora do próprio documento delimitam as tabelas, as linhas são reconhecidas pelo conteúdo, como a unidade de medida ou o posto horário, e as colunas numéricas são calibradas pelas linhas completas de cada fatura. Como a distribuidora adotou três diagramações diferentes no período, a ferramenta identifica automaticamente a diagramação de cada documento e aplica o extrator correspondente. Os dados são convertidos para tipos numéricos, gravados em banco SQLite com nove tabelas e analisados quanto a sazonalidade, correlações e meses atípicos. A corretude foi verificada por uma identidade interna da própria fatura: em todas as 65, a soma dos itens de fornecimento com os lançamentos financeiros

reproduziu o valor a pagar, sem diferença de um centavo. A análise da série produziu três resultados de interesse para a gestão do câmpus. A demanda contratada, de 250 kW, nunca foi atingida: o maior valor medido foi 177,12 kW, em dezembro de 2023, e a readequação para 195 kW, o pico acrescido de 10%, representaria economia estimada em R$ 20.148,58 por ano, mantido o perfil de uso. O saldo de créditos de geração distribuída passou de 921 kWh, em setembro de 2020, para 412.470 kWh, em agosto de 2026, mês em que a fatura registrou pela primeira vez créditos a expirar, no total de 4.846,66 kWh. Nas seis faturas com adicional de bandeira tarifária, a energia injetada compensou quase todo o acréscimo: dos R$ 6.066,46 de adicional bruto, foram efetivamente pagos R$ 277,48. Conclui-se que delimitar as tabelas por âncoras e reconhecer as linhas pelo conteúdo, em vez de recorrer a posições fixas na página, permitiu aplicar o mesmo método às três diagramações da série, e que a base obtida já oferece subsídios para decisões sobre o contrato de fornecimento de energia do câmpus.

Palavras-chave

faturas de energia elétrica; extração de dados; PyMuPDF; demanda contratada; geração distribuída.

Agradecimentos

O presente trabalho foi realizado com apoio do Instituto Federal de Educação, Ciência e Tecnologia de Goiás (Edital nº 21/2025-PROPPG).