Trabalho Final: Análise de Energia de uma Instalação Elétrica

1. Objetivo

Você vai atuar como engenheiro(a) eletricista responsável pela **eficiência energética** de um prédio universitário (instalação trifásica 127/220 V). Chegaram os dados de um mês de medição e você precisa responder, com evidência numérica e visual: **a instalação está operando dentro dos limites de qualidade de energia e onde há mais oportunidade de economia?**

O trabalho usa `pandas` (manipulação de dados), `numpy` (cálculos) e `matplotlib` (visualização), além de reforçar três hábitos: especificar o problema antes de programar, usar IA de forma crítica e documentada, e testar o próprio código antes de confiar nele.

Grandezas que você vai usar. Em cada hora, a medição traz a potência ativa P (kW), a potência reativa Q (kvar), as tensões e as correntes das três fases. A potência aparente e o fator de potência são calculados por você:

$$S = \sqrt{P^2 + Q^2} \quad \text{(kVA)} \qquad\qquad FP = \frac{P}{S}$$

2. Arquivos fornecidos

Três arquivos CSV, referentes a setembro de 2026 (30 dias).

`medicoes_horarias.csv` — uma linha por hora (720 registros)

Coluna	Descrição
`timestamp`	data e hora da medição
`data`	data (AAAA-MM-DD)
`hora`	hora do dia (0 a 23)
`tensao_a_v`, `tensao_b_v`, `tensao_c_v`	tensão fase-neutro de cada fase, em volts
`corrente_a_a`, `corrente_b_a`, `corrente_c_a`	corrente de cada fase, em ampères
`potencia_ativa_kw`	potência ativa média na hora (kW)
`potencia_reativa_kvar`	potência reativa média na hora (kvar)
`energia_kwh`	energia ativa consumida na hora (kWh)

`consumo_circuitos.csv` — consumo diário por circuito (180 registros)

Coluna	Descrição
`data`	data
`circuito`	Climatização, Iluminação, Laboratórios, Servidores, Elevadores ou Bombas d'água
`energia_kwh`	energia consumida no dia pelo circuito

`cargas_instaladas.csv` — cadastro dos circuitos (6 registros)

Coluna	Descrição
`circuito`	nome do circuito (mesmo de `consumo_circuitos.csv`)
`potencia_instalada_kw`	potência instalada (kW)
`tipo_carga`	motores, iluminação, eletrônica, equipamentos mistos
`setor`	setor atendido

As fontes se relacionam por `data` e `circuito`. Parte do trabalho exige **combinar fontes** (`merge`), como na prática, em que medições, cadastro e faturamento raramente vêm de uma tabela só.

Carregue com `pd.read_csv(...)`. A coluna `data` vem como texto: converta para `datetime` antes de qualquer análise temporal.

> **Atenção aos dados.** Em dois momentos do mês as leituras saíram "estranhas". Um deles é uma **falta de energia** real; o outro é uma **falha do medidor**. Descubra qual é qual e trate cada um de forma diferente e justificada. Não aplique `dropna()` nem `fillna(0)` sem entender o que está fazendo.

3. O que vocês devem entregar

Etapa 0 — Especificação (entrega separada, antes do código). Escrevam meia página respondendo: quais perguntas específicas a análise vai responder (mínimo 4, Seção 4); para cada uma, que gráfico ou tabela responde melhor; e que suposições vocês estão fazendo sobre os dados. Essa etapa é entregue e avaliada separadamente.

Etapa 1 — Preparação dos dados. Carreguem os arquivos, corrijam os tipos (datas), calculem S e FP e tratem valores ausentes e anômalos de forma justificada.

Etapa 2 — Análise. Respondam às perguntas da Seção 4 com `pandas` (`groupby`, `merge`, agregações) e `numpy`.

Etapa 3 — Visualizações. Produzam os gráficos da Seção 5, todos com título, rótulos de eixo (com unidade) e legenda quando fizer sentido.

Etapa 4 — Log de uso de IA. Se usaram alguma IA, registrem o prompt, o que ela gerou e o que vocês corrigiram ou rejeitaram, e por quê. Isso **não é opcional e não é penalizado**: faz parte da nota. O que se penaliza é usar código gerado sem saber explicá-lo, ou sem documentar.

Etapa 5 — Testes de sanidade. Incluam no código pelo menos **3 verificações (`assert`)** que confirmem os números com uma conferência simples. Exemplos: $P \leq S$ em todas as

linhas; a soma de `energia_kwh` do mês em `medicoes_horarias.csv` é igual (dentro de 1%) à soma de `consumo_circuitos.csv`; nenhum *FP* fora do intervalo [0, 1].

Etapa 6 — Relatório final. De 2 a 4 páginas (ou notebook comentado) com as respostas, os gráficos e a **interpretação**: não basta mostrar o gráfico, expliquem o que ele significa para quem vai decidir (investir em correção de fator de potência? reforçar a alimentação? mudar horários de operação?). O log de IA entra como seção separada.

4. Perguntas obrigatórias

1. **Qual foi a demanda máxima da instalação (kW) e quando ocorreu?** Como difere a curva de carga média (por hora do dia) entre dias úteis e fins de semana?
2. **Em quantas horas o fator de potência ficou abaixo de 0,92** (limite de referência da regulamentação para unidades consumidoras)? Em quais faixas de horário isso se concentra?
3. **As tensões das três fases ficaram dentro da faixa adequada de 117 V a 133 V?** Qual fase teve mais horas fora dos limites, e existe relação com a demanda (tensão × potência)?
4. **Qual circuito mais consome, e qual tem o maior fator de carga?** Combine `consumo_circuitos.csv` com `cargas_instaladas.csv`, usando:

$$FC = \frac{\text{energia consumida no mês}}{\text{potência instalada} \times \text{horas do mês}}$$

Opcionais (pontuação extra):

5. **Desequilíbrio entre fases.** Calcule, hora a hora, o desequilíbrio de corrente $(I_{\text{max}} - I_{\text{média}})/I_{\text{média}}$ e verifique em quantas horas passa de 10%.
6. **Correção do fator de potência.** Para a hora de maior demanda, calcule a potência reativa do banco de capacitores para elevar o *FP* a 0,95:

$$Q_c = P\,(\tan\varphi_1 - \tan\varphi_2), \qquad \varphi = \arccos(FP)$$

7. **Tarifa.** Compare o custo mensal na tarifa convencional (R$ 0,80/kWh) com uma tarifa por horário (ponta, das 18h às 20h59: R$ 1,30; intermediária, 17h–17h59 e 21h–21h59: R$ 0,95; fora de ponta: R$ 0,60). *Valores fictícios, só para o exercício.*

5. Gráficos obrigatórios

- **Série temporal:** demanda (kW) hora a hora ao longo do mês (ou a demanda máxima diária), identificando dias atípicos.
- **Curva de carga:** demanda média por hora do dia, com duas linhas (dias úteis e fins de semana).
- **Histograma ou boxplot da tensão:** distribuição das tensões das três fases, com os limites de 117 V e 133 V marcados.
- **Barras (ou Pareto):** consumo mensal por circuito, em kWh e em percentual do total.

Gráfico adicional sugerido: dispersão tensão × demanda, com reta de ajuste (`numpy.polyfit`).

6. Como vocês serão avaliados

Critério	Peso
Especificação prévia (Etapa 0): clareza das perguntas e das suposições	15%
Correção técnica do código (`pandas`, `numpy`, `matplotlib`)	25%
Interpretação: conectar os números a uma decisão de engenharia	25%
Log de uso de IA: honestidade e criticidade	15%
Testes de sanidade (Etapa 5)	10%
Defesa oral	10%

Sobre a defesa oral (5 a 7 minutos por aluno ou dupla): o professor poderá pedir que expliquem qualquer célula do código, inclusive trechos gerados com apoio de IA. Não é uma pegadinha: é o que qualquer engenheiro enfrenta ao revisar um código que não escreveu sozinho.

7. Regras sobre uso de IA

O uso é **permitido e incentivado**, com duas condições:

1. **Tudo precisa ser documentado** (Etapa 4). Um notebook inteiro gerado por IA e entregue sem o log é penalizado, mesmo que esteja correto.
2. **Vocês precisam ser capazes de explicar qualquer linha** na defesa oral.

Pedir ajuda para lembrar a sintaxe do `pandas` ou do `matplotlib` é uso legítimo. O que não se delega é a parte central do trabalho: decidir quais perguntas fazer aos dados e o que os resultados significam.

8. Formato de entrega

- Notebook (`.ipynb`) ou script (`.py`) com o código completo e comentado;
- Documento de especificação (Etapa 0), em arquivo separado;
- Log de uso de IA (Etapa 4), como seção do relatório ou arquivo separado;
- Relatório final (Etapa 6), em PDF ou Markdown.