**Resumo do Projeto: Avaliação dos Ativos de Média Tensão da Rede Subterrânea da Light**

**Objetivo Geral**

Realizar uma análise aprofundada dos ativos subterrâneos de média tensão para:

1. Identificar segmentos críticos da rede subterrânea.
2. Propor um ranking para priorização de manutenção e melhorias.

**Objetivos Específicos**

* Analisar dados históricos e de operação para mapear a contribuição de cada segmento nos índices regulatórios (DEC e FEC).
* Definir critérios de criticidade técnica, estrutural e operacional.
* Consolidar e tratar bases de dados, incluindo resultados de inspeções em campo.
* Desenvolver matrizes de priorização e modelos preditivos utilizando inteligência computacional.
* Gerar relatórios técnicos com recomendações práticas.

**Etapas Principais**

1. **Definição dos Critérios de Criticidade**
   * Identificar variáveis relevantes (frequência de falhas, tempo de operação, impacto regulatório).
   * Validar critérios com especialistas e ajustar pesos com base em análises.
2. **Coleta, Organização e Análise de Dados**
   * Estruturar dados técnicos e operacionais em bases padronizadas.
   * Utilizar Python para tratamento, limpeza e análise exploratória.
3. **Inspeções em Campo**
   * Planejar inspeções com base em amostragem estatística.
   * Desenvolver protocolos para avaliar condições estruturais e operacionais.
4. **Matriz de Priorização e Modelo Preditivo**
   * Aplicar algoritmos de Machine Learning (Random Forest, XGBoost) para prever falhas.
   * Desenvolver dashboards interativos para visualização de criticidade.
5. **Relatório Executivo**
   * Consolidar as análises e apresentar recomendações para tomada de decisão.

**Contribuições do Python**

O Python será uma ferramenta central no projeto, com aplicações em todas as etapas:

1. **Tratamento e Análise de Dados**
   * Bibliotecas: pandas, numpy, scipy para limpeza, análise e imputação de dados faltantes.
   * Geração de relatórios exploratórios com matplotlib e seaborn.
2. **Modelagem Preditiva**
   * Aplicação de modelos de Machine Learning usando scikit-learn e XGBoost.
   * Validação cruzada e análise de métricas de desempenho.
3. **Visualizações Interativas**
   * Desenvolvimento de dashboards dinâmicos com Dash e Plotly.
   * Integração de tabelas, gráficos de barras e mapas para insights rápidos.
4. **Simulações Estatísticas**
   * Uso de técnicas como Monte Carlo para inferências em unidades não inspecionadas.
5. **Automação de Relatórios**
   * Geração de documentos e gráficos embutidos usando ferramentas como Jinja2 e Matplotlib.

**Principais Ferramentas e Bibliotecas Já Instaladas**

1. **Manipulação e Análise de Dados**
   * pandas: Para manipulação de dados estruturados.
   * numpy: Para cálculos matemáticos e operações vetoriais.
   * scipy: Para análises estatísticas e operações científicas.
2. **Visualizações**
   * **Recomendação Plotly**: Excelente substituto para matplotlib (já instalado, mas não utilizado devido à incompatibilidade). A versão instalada de plotly é recente e suficiente para interatividade.
3. **Dashboards**
   * dash, dash-core-components, dash-html-components, dash-table: Ferramentas completas para criação de dashboards interativos.
4. **Machine Learning e Estatística**
   * scikit-learn: Para modelagem preditiva.
   * statsmodels: Análises estatísticas adicionais.
   * joblib, threadpoolctl: Para paralelismo e otimização.
5. **Outros**
   * cvxpy: Útil para otimização matemática.
   * openpyxl: Para manipulação de arquivos Excel.
   * SQLAlchemy: Suporte para bases de dados SQL.

**Sugestões Adicionais (Se Necessário)**

1. **Visualizações Geoespaciais**
   * **Recomendação**: Caso precise de mapas interativos, considere instalar folium ou explorar funcionalidades de mapas com plotly.
2. **Validação de Dados**
   * great\_expectations pode ser útil para validação estruturada de dados.
3. **Exploração de Dados**
   * pandas-profiling: Para relatórios automáticos de análise exploratória.
4. **Infraestrutura**
   * gunicorn (se necessário para o deployment de aplicações web em produção).

**Bibliotecas Adicionadas**

1. **Visualização**
   * **folium**: Para mapas interativos, ideal para representar geolocalização de ativos.
   * **altair**: Ferramenta flexível para gráficos declarativos.
   * **missingno**: Para análise visual de valores ausentes nos dados.
2. **Validação e Exploração de Dados**
   * **great-expectations**: Ótimo para garantir a qualidade e consistência das bases de dados.
   * **ydata-profiling**: Criação de relatórios exploratórios automáticos, útil para uma visão rápida do dataset.
3. **Back-end e Infraestrutura**
   * **gunicorn**: Para servir aplicações web baseadas em Flask ou Dash em produção.
4. **Outros Suportes**
   * **jsonschema e jsonschema-specifications**: Validação de dados baseados em JSON.
   * **marshmallow**: Serialização e validação de objetos.