

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from scipy.stats import pearsonr, linregress, norm

# Carregar os dados
file_path = '/content/drive/MyDrive/Python/Frequencia
Igreja/Frequência_Ceia IBCA 2022-2024.xlsx'
data = pd.read_excel(file_path)

# Remover colunas desnecessárias
data = data[['Dias', 'Manhã', 'Noite']]

# Converter a coluna 'Dias' para datetime
data['Dias'] = pd.to_datetime(data['Dias'], errors='coerce')

# Filtrar apenas os domingos
sundays_data = data[data['Dias'].dt.dayofweek == 6].copy()
sundays_data['Ano'] = sundays_data['Dias'].dt.year

# Selecionar colunas de interesse e remover valores nulos
sundays_data = sundays_data[['Dias', 'Ano', 'Manhã',
'Noite']].dropna()

# Garantir que os dados são numéricos
sundays_data['Manhã'] = pd.to_numeric(sundays_data['Manhã'],
errors='coerce')
sundays_data['Noite'] = pd.to_numeric(sundays_data['Noite'],
errors='coerce')
sundays_data = sundays_data.dropna()

# Salvar os dados de frequência em Parquet
sundays_data.to_parquet('frequencias_domingo.parquet', index=False)
print("Arquivo 'frequencias_domingo.parquet' salvo com sucesso.")

# Regressão linear
morning_service = sundays_data['Manhã']
evening_service = sundays_data['Noite']
slope, intercept, r_value, p_value, std_err =
linregress(morning_service, evening_service)

# Resultados da regressão em DataFrame
regression_results = pd.DataFrame([
    'Coeficiente Angular (slope)': slope,
    'Intercepto': intercept,
    'Correlação (R)': r_value,
    'Coeficiente de Determinação (R²)': r_value**2,
    'Valor-p': p_value,

```

```

        'Erro Padrão': std_err
    ]])

# Salvar os resultados da regressão em Parquet
regression_results.to_parquet('regressao_resultados.parquet',
index=False)
print("Arquivo 'regressao_resultados.parquet' salvo com sucesso.")

# Gráfico de dispersão + linha de regressão
plt.figure(figsize=(10, 6))
plt.scatter(morning_service, evening_service, alpha=0.7,
label="Dados de Frequência")
plt.plot(morning_service, slope * morning_service + intercept,
color="red", label=f"y = {slope:.2f}x + {intercept:.2f}")
plt.title('Gráfico de Dispersão: Culto da Manhã vs. Culto da Noite
com Regressão Linear')
plt.xlabel('Frequência do Culto da Manhã')
plt.ylabel('Frequência do Culto da Noite')
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()

# Exibir resultados
print(f"Equação da linha de regressão: y = {slope:.2f}x +
{intercept:.2f}")
print(f"Coeficiente de correlação (R²): {r_value**2:.2f}")
print(f"Correlação (R): {r_value:.2f}, Valor-p: {p_value:.4f}")

```