

```
import pandas as pd

# 1. Leitura dos dados agregados
df = pd.read_csv('total_diario_mensal.csv')

# 2. Tratamento de valores nulos
df['precipitacao'] = df['precipitacao'].fillna(0)
df.fillna({'temperatura': df['temperatura'].mean()}, inplace=True)

# 3. Renomeação de colunas
df = df.rename(columns={
    'estacao_id': 'codigo_estacao',
    'precipitacao': 'chuva_mm',
    # ...
})

# 4. Ajuste de tipos de dados
df['data'] = pd.to_datetime(df['data'], dayfirst=True)
df['chuva_mm'] = df['chuva_mm'].astype(float)
df['codigo_estacao'] = df['codigo_estacao'].astype(str)

# 5. Padronização de códigos institucionais
mapping = {
    'EST001': 'RJ_EST001',
    'EST2': 'RJ_EST002',
    # ...
}
df['codigo_estacao'] =
df['codigo_estacao'].map(mapping).fillna(df['codigo_estacao'])

# 6. Salvamento em Parquet
df.to_parquet('dados_tratados.parquet', index=False)
```