

```
import pandas as pd

def main():
    df = pd.read_csv("dados/frequencia.csv")
    df.fillna({'quantidade': 0, 'cod_inst': 'UNKNOWN'},
inplace=True)
    df.rename(columns={'cod_inst': 'codigo_instituicao',
                        'freq': 'frequencia'}, inplace=True)
    df['data'] = pd.to_datetime(df['data'], dayfirst=True,
errors='coerce')
    df['frequencia'] = df['frequencia'].astype(int)
    df['codigo_instituicao'] = (
        df['codigo_instituicao'].str.strip().str.upper().str.zfill(
5)
    )
    df.to_parquet("output/frequencia_padronizada.parquet",
index=False)
    print("Parquet gerado com sucesso.")

if __name__ == "__main__":
    main()
```