

Kod do konwersji danych z pliku .txt do csv dla wybranego (w tym wypadku 2023) roku

In [ ]:

```
import pandas as pd

# Parametry populacji
KRAJ = "Denmark"
ROK = 2023
PLEC = "Male"
WIEK_START = 30
N_LAT = 10

# Wczytywanie danych bez nagłówków
df = pd.read_csv('data/Mortality.txt', sep=r'\s+', skiprows=2)
df.columns = df.columns.str.strip()

# Konwersja kolumny Age na liczby, błędy (np: '110+') zamieniamy na NaN
df['Age'] = pd.to_numeric(df['Age'], errors='coerce')

# Filtrowanie danych dla wybranej populacji
data_clean = df[(df['Year'] == ROK) &
                 (df['Age'] >= WIEK_START) &
                 (df['Age'] < WIEK_START + N_LAT)].copy()

# Sprawdzenie poprawności
if len(data_clean) == N_LAT:
    print(f"Wczytano dane dla {len(data_clean)} lat (wiek {WIEK_START}-{WIEK_START+N_LAT})")
    print(data_clean[['Age', 'qx']])
    data_clean.to_csv('data/processed_mortality.csv', index=False)
else:
    print(f"BŁĄD: Znaleziono tylko {len(data_clean)} wierszy zamiast {N_LAT}.")
```

Sukces! Wczytano dane dla 10 lat (wiek 30–39).

	Age	qx
20898	30.0	0.00064
20899	31.0	0.00043
20900	32.0	0.00066
20901	33.0	0.00072
20902	34.0	0.00064
20903	35.0	0.00082
20904	36.0	0.00081
20905	37.0	0.00086
20906	38.0	0.00095
20907	39.0	0.00080