Data Cleaning Code

November 22, 2024

```python
[3]: import pandas as pd
     from sklearn.preprocessing import StandardScaler

     # read csv
     file_path = 'sort_v2.csv'
     data = pd.read_csv(file_path)

     # the columns which you need to standized
     score_columns = [
         'Children Score',
         'Single Parents Score',
         'After-tax Income Score',
         'Education Score',
         'Movers Score',
         'Total Score'
     ]

     # initialize Z-score tol
     scaler = StandardScaler()

     # z-score application
     data[score_columns] = scaler.fit_transform(data[score_columns])

     # save
     data.to_csv('z_score_cleaned_data.csv', index=False)

     print("Z-Score is finished")
```

```
Z-Score is finished
```

```python
[5]: # keep 2 decimals
     data[score_columns] = data[score_columns].round(2)

     # save
     data.to_csv('z_score_cleaned_data_rounded.csv', index=False)
```

```python
[ ]:
```