A.Convert_Excel_to_Numpy

November 17, 2020

[65]:
```
#all Modules:
import pandas as pd
import numpy as np
import glob
import multiprocessing as mp
from tqdm import tqdm
```

[66]:
```
#definieer variabelen:
load_path = '/home/18005152/notebooks/zero/Data(xlsx)/'
save_path = '/home/18005152/notebooks/zero/A/'
#Get all the file names:
paths = glob.glob(load_path+"/*.xlsx")
```

[3]:
```
#definieer de functies:
def load_house(sheet, house):
    """Load the file and save the file."""
    dff = pd.read_excel(load_path + str(house) + '.xlsx', sheet_name=sheet,␣
 ↪engine="openpyxl")
    np.save(save_path + sheet + '_' + str(house), dff.values)

def nParse3(n):
    """output a string that is 3 decimals long. Levy en Jefry kunnen uitleg␣
 ↪geven"""
    number = str(n)
    if len(number) == 1:
        number = "00" + number
    elif len(number) == 2:
        number = "0" + number
    elif len(number) == 3:
        number = number
    return str(number)

def get_sheet_names(i):
    """Get the sheet names from the excel file."""
    return pd.ExcelFile(load_path + str(i) + '.xlsx',engine="openpyxl").
 ↪sheet_names
```

[53]:
```
#start the main loop:
for path in tqdm(paths[0:120]):

    #extract house number:
    house_number = path[-8:-5]

    #read the excel file:
    df = pd.read_excel(path, engine="openpyxl", sheet_name=None)

    #Get The sheet names:
    sheets_names = list(df.keys())

    #write the file to the desired Numpy file:
    for sheet in sheets_names:
        output_file_name = sheet + "_" + house_number
        np.save(save_path+output_file_name, df[sheet].to_numpy())

    print("House {} has {} sheets.".format(house_number, len(sheets_names)))
```

```
House 087 has 22 sheets.
```