# spec_functions

October 17, 2023

[ ]:

```
import pandas as pd
import matplotlib.pyplot as plt
import os

def file_extract(spec_file):
    """

    Extracts from a given txt file the 2 axis from the spectra.

    Parameters
    ----------
    spec_file : str
        Path to the txt spectra file.

    Returns
    -------
    df : pd.DataFrame
        Dataframe with one x 'wavelength' column and onde y 'CPS' column.

    Examples
    --------
    Reading the fluorescence emission txt example file, storing the axis in
    the dataframe

    >>> file_extract('../example_data/aunu_fl.txt')
        Wavelenght (nm)            CPS
    0           1200.00  1.03100000E+3
    1          1201.000  1.03400000E+3
    2          1202.000  1.09800000E+3
    3          1203.000  1.12200000E+3
    4          1204.000  1.07900000E+3
    ..              ...            ...
    196        1396.000  9.37000000E+2
    197        1397.000  8.98000000E+2
    198        1398.000  9.12000000E+2
    199        1399.000  8.44000000E+2
    200         1400.00  9.53000000E+2
    <BLANKLINE>
    [201 rows x 2 columns]

    """
    data = []

    # Open the text file for reading
    with open(spec_file, 'r') as file:
        for line in file:
            # Split each line into two elements, assuming space as the
            # separator
            elements = line.split()

            # Check if the first element is a string (you can modify
            # this condition)
            if not elements[0].isalpha():
                # Append the first element to 'axis' and the second
                # element to 'cps'
                data.append([elements[0], elements[1]])

    # Create a Pandas DataFrame from the collected data
    df = pd.DataFrame(data, columns=['Wavelenght (nm)', 'CPS'])
    df['CPS'] = df['CPS'].astype('float')
    df['Wavelenght (nm)'] = df['Wavelenght (nm)'].astype('float')
    df.round({'Wavelenght (nm)': 0})
    return df


def spc_plot(df, input_path, title= None):
    """

    Plots a matplotlib spectra from a given dataframe extracted from a txt.

    Parameters
    ----------
    df : pd.Dataframe
        Dataframe extracted from a spectrum in .txt format
    input_path : str
        String of the path from the parsed input txt file
    title : str
        Plot title

    Returns
    -------
    None

    """
    df_names = df.columns.values.tolist()
    col1 = df_names[0]
    col2 = df_names[1]
    fig, ax = plt.subplots()
    # plt.rcParams.update(
        # {
            # "xtick.major.pad": 50,
            # "xtick.major.size": 50,
            # "ytick.major.pad": 50,
            # "ytick.major.size": 50,
        # }
    # )
    # ax.set_xlim(1200, 1400)
    ax.grid(True)
    plt.xlabel(col1)
    plt.ylabel(col2)
    if title == None:
        plt.title((os.path.basename(input_path))[:-4])
    else:
        plt.title(title)
    ax.plot(df[col1], df[col2])
    plt.savefig(os.path.splitext(input_path)[0])


def mean_baseline(df):
    baseline = df['CPS'].mean()
    df['CPS'] = df['CPS'] - baseline
    df['CPS'] = df['CPS'].clip(lower=0)
    return df
```

[ ]:

```
df = file_extract('/home/eduardotc/Programação/my_gits/materials_chempy/
 ↪example_data/aunu_fl.txt')
ttt = df['CPS'].mean()
print(ttt)
```