ENV-TECH_green classification

June 9, 2022

[ ]:
```
import io
import sharepy
import getpass
import pandas as pd
import sys
import os
from pathlib import Path
import numpy as np
import csv
```

[ ]:
```
# # Authentication to sharepoint
# URL = 'https://areasciencepark.sharepoint.com'
# SHAREPOINT_USER = 'leyla.vesnic@areasciencepark.it'
# SHAREPOINT_PASSWORD =  getpass.getpass('inserire la password: ')
# s = sharepy.connect(URL, username=SHAREPOINT_USER,␣
 ↪password=SHAREPOINT_PASSWORD)
```

NB!!!!!!!! Da migliorare: il ciclo for va a prendere solo i file dell'ultima directory salvata nella lista finale, quindi serve che vada a pescare anche quelli prima! Nel mentre ho costruito, con i medesimi cicli, un dataframe per ogni cartella di CPC e IPC, per poi fare l'append su tutti (mooooolto lungo e poco efficiente).

[ ]:
```
lista_dir1 = ['1_CPC']
lista_dir2 = ['2_CPC']
lista_dir3 = ['3_CPC']
lista_dir4 = ['4_CPC']
lista_dir5 = ['5_CPC']
lista_dir6 = ['6_CPC']
lista_dir7 = ['7_CPC']
lista_dir8 = ['8_CPC']
lista_dir9 = ['1_IPC']
```

[ ]:
```
path_base = r"C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green␣
 ↪classification"
path_dati = path_base + "\\ENV-TECH codes\\"
lista_dd = []
for x in lista_dir1:
    lista_dd.append(path_dati + x + "\\")
```

```
print(f'directory dati: {lista_dd}')
```

```
directory dati: ['C:\\Users\\bincoletto\\OneDrive - Area Science
Park\\Documenti\\green classification\\ENV-TECH codes\\1_CPC\\']
```

[ ]:
```
def lista_file_csv(dd):
    lista_file = []
    for root, dirs, files in os.walk(dd):
        for name in files:
            if 'csv' in name:
                lista_file.append(os.path.join(root, name))
    n = len(lista_file)
    print(f'la directory contiene {n} file')
    return (lista_file)
```

[ ]:
```
for dd in lista_dd:
    print(f'>>> inizio elaborazione {dd}')
    file_da_elaborare = lista_file_csv(dd)
```

```
>>> inizio elaborazione C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\1_CPC\
la directory contiene 13 file
```

[ ]:
```
def carica_dati_da_csv(ff):
    print(f'Elaborazione dei file {ff} \n ')
    df = pd.read_csv(ff, sep = ',', encoding='utf-8-sig', dtype=str)
    return(df)
```

[ ]:
```
lista_df = []
for ff in file_da_elaborare:
    df = carica_dati_da_csv(ff)
    lista_df.append(df)
df_code_1 = pd.concat(lista_df)
```

```
Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\1_CPC\1.1.1. Emissions
abatement from stationary sources.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\1_CPC\1.1.2. Emissions
abatement from mobile sources.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\1_CPC\1.1.3. Air pollution
abatement - Not elsewhere classified.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\1_CPC\1.2.1. Water and
```

```
wastewater treatment.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\1_CPC\1.2.2. Fertilizers from
wastewater.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\1_CPC\1.2.3. Oil spill and
pollutant clean-up.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\1_CPC\1.3.1. Solid waste
collection.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\1_CPC\1.3.2. Material
recovery recycling and re-use.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\1_CPC\1.3.3. Fertilizers from
waste.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\1_CPC\1.3.4. Incineration and
energy recovery.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\1_CPC\1.3.6. Waste management
Not elsewhere classified.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\1_CPC\1.4.  Soil
remediation.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\1_CPC\1.5.  Environmental
monitoring.csv
```

[ ]:
```
path_base = r"C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green␣
  ↪classification"
path_dati = path_base + "\\ENV-TECH codes\\"
lista_dd = []
for x in lista_dir2:
    lista_dd.append(path_dati + x + "\\")
print(f'directory dati: {lista_dd}')
```

```
for dd in lista_dd:
    print(f'>>> inizio elaborazione {dd}')
    file_da_elaborare = lista_file_csv(dd)

lista_df = []
for ff in file_da_elaborare:
    df = carica_dati_da_csv(ff)
    lista_df.append(df)
df_code_2 = pd.concat(lista_df)
```

```
directory dati: ['C:\\Users\\bincoletto\\OneDrive - Area Science
Park\\Documenti\\green classification\\ENV-TECH codes\\2_CPC\\']
>>> inizio elaborazione C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\
la directory contiene 31 file
Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2. Climate change
mitigation technologies related to energy generation, transmission or
distribution.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.1. Renewable energy
generation.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.1.1. Wind energy.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.1.2. Solar thermal
energy.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.1.3. Solar
photovoltaic (PV) energy.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.1.4. Solar thermal-PV
hybrids.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.1.5. Geothermal
energy.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.1.6. Marine energy,
e.g. using wave energy or salinity gradient.csv
```

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.1.7. Hydro energy.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.2. Energy generation from fuels of non-fossil origin.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.2.1. Biofuels, e.g. bio-diesel.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.2.2. Fuel from waste, e.g. synthetic alcohol or diesel.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.3. Combustion technologies with mitigation potential (e.g. using fossil fuels, biomass, waste, etc.).csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.3.1. Technologies for improved output efficiency (combined heat and power, combined cycles, etc.).csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.3.2. Technologies for improved input efficiency (efficient combustion or heat usage).csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.4. Nuclear energy.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.4.1. Nuclear fusion reactors.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.4.2. Nuclear fission reactors.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.5. Technologies for an efficient electrical power generation, transmission or distribution.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.5.1. Superconducting electric elements or equipment.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science

```
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.5.2. Smart grids as
climate change mitigation technology in the energy generation sector.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.5.3. Not elsewhere
classified.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.6. Enabling
technologies (technologies with potential or indirect contribution to GHG
emission mitigation).csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.6.1.1. Batteries.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.6.1.2. Capacitors.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.6.1.3. Thermal energy
storage.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.6.1.4. Mechanical
energy storage, e.g. flywheels or pressurised fluids.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.6.2. Hydrogen
technology.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.6.3. Fuel cells.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.6.4. High-voltage
direct current transmission.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\2_CPC\2.7. Other energy
conversion or management systems reducing GHG emissions.csv
```

[ ]:
```
path_base = r"C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green␣
 ↪classification"
path_dati = path_base + "\\ENV-TECH codes\\"
lista_dd = []
for x in lista_dir3:
```

```
    lista_dd.append(path_dati + x + "\\")
print(f'directory dati: {lista_dd}')

for dd in lista_dd:
    print(f'>>> inizio elaborazione {dd}')
    file_da_elaborare = lista_file_csv(dd)

lista_df = []
for ff in file_da_elaborare:
    df = carica_dati_da_csv(ff)
    lista_df.append(df)
df_code_3 = pd.concat(lista_df)
```

```
directory dati: ['C:\\Users\\bincoletto\\OneDrive - Area Science
Park\\Documenti\\green classification\\ENV-TECH codes\\3_CPC\\']
>>> inizio elaborazione C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\3_CPC\
la directory contiene 5 file
Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\3_CPC\3. Capture, storage,
sequestration or disposal of greenhouse gases.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\3_CPC\3.1. CAPTURE OR
DISPOSAL OF NITROUS OXIDE (N2O).csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\3_CPC\3.2. Capture or
disposal of methane (CH4).csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\3_CPC\3.3. Capture or
disposal of perfluorocarbons (PFC), hydrofluorocarbons (HFC) or sulfur
hexafluoride (SF6).csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\3_CPC\3.4. Capture or
disposal of carbon dioxide (CO2).csv
```

[ ]:
```
path_base = r"C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green␣
  ↪classification"
path_dati = path_base + "\\ENV-TECH codes\\"
lista_dd = []
for x in lista_dir4:
    lista_dd.append(path_dati + x + "\\")
print(f'directory dati: {lista_dd}')
```

```
for dd in lista_dd:
    print(f'>>> inizio elaborazione {dd}')
    file_da_elaborare = lista_file_csv(dd)

lista_df = []
for ff in file_da_elaborare:
    df = carica_dati_da_csv(ff)
    lista_df.append(df)
df_code_4 = pd.concat(lista_df)
```

```
directory dati: ['C:\\Users\\bincoletto\\OneDrive - Area Science
Park\\Documenti\\green classification\\ENV-TECH codes\\4_CPC\\']
>>> inizio elaborazione C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\4_CPC\
la directory contiene 11 file
Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\4_CPC\4. Climate change
mitigation technologies related to transportation.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\4_CPC\4.1. Road transport.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\4_CPC\4.1.1. Conventional
vehicles (based on internal combustion engine).csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\4_CPC\4.1.3. Electric
vehicles.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\4_CPC\4.1.4. Fuel efficiency-
improving vehicle design (common to all road vehicles).csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\4_CPC\4.2. Rail transport.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\4_CPC\4.3. Aeronautics or air
transport.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\4_CPC\4.4. Maritime or
waterways transport.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
```

```
Park\Documenti\green classification\ENV-TECH codes\4_CPC\4.5. Enabling
technologies in transport.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\4_CPC\4.5.1. Electric vehicle
charging.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\4_CPC\4.5.2. Application of
hydrogen technology to transportation, e.g. using fuel cells.csv
```

[ ]:
```
path_base = r"C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green␣
 ↪classification"
path_dati = path_base + "\\ENV-TECH codes\\"
lista_dd = []
for x in lista_dir5:
    lista_dd.append(path_dati + x + "\\")
print(f'directory dati: {lista_dd}')

for dd in lista_dd:
    print(f'>>> inizio elaborazione {dd}')
    file_da_elaborare = lista_file_csv(dd)

lista_df = []
for ff in file_da_elaborare:
    df = carica_dati_da_csv(ff)
    lista_df.append(df)
df_code_5 = pd.concat(lista_df)
```

```
directory dati: ['C:\\Users\\bincoletto\\OneDrive - Area Science
Park\\Documenti\\green classification\\ENV-TECH codes\\5_CPC\\']
>>> inizio elaborazione C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\5_CPC\
la directory contiene 9 file
Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\5_CPC\5. Climate change
mitigation technologies related to buildings.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\5_CPC\5.1. Integration of
renewable energy sources in buildings.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\5_CPC\5.2.1. Energy efficient
lighting.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
```

```
Park\Documenti\green classification\ENV-TECH codes\5_CPC\5.2.2. Energy efficient
heating, ventilation or air conditioning.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\5_CPC\5.2.3. Energy
efficiency in home appliances.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\5_CPC\5.2.4. Energy efficient
elevators, escalators and moving walkways, e.g. energy saving or recuperation
technologies.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\5_CPC\5.2.5. End-user
side.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\5_CPC\5.3. Architectural or
constructional elements improving the thermal performance of buildings.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\5_CPC\5.4. Enabling
technologies in buildings.csv
```

[ ]:
```python
path_base = r"C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green␣
 ↪classification"
path_dati = path_base + "\\ENV-TECH codes\\"
lista_dd = []
for x in lista_dir6:
    lista_dd.append(path_dati + x + "\\")
print(f'directory dati: {lista_dd}')

for dd in lista_dd:
    print(f'>>> inizio elaborazione {dd}')
    file_da_elaborare = lista_file_csv(dd)

lista_df = []
for ff in file_da_elaborare:
    df = carica_dati_da_csv(ff)
    lista_df.append(df)
df_code_6 = pd.concat(lista_df)
```

```
directory dati: ['C:\\Users\\bincoletto\\OneDrive - Area Science
Park\\Documenti\\green classification\\ENV-TECH codes\\6_CPC\\']
>>> inizio elaborazione C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\6_CPC\
la directory contiene 18 file
```

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\6_CPC\6.2. SOLID WASTE MANAGEMENT.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\6_CPC\6.2.1.Waste collection, transportation, transfer or storage.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\6_CPC\6.2.2. Waste processing or separation.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\6_CPC\6.2.3. Landfill technologies aiming to mitigate methane emissions.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\6_CPC\6.2.4. Bio-organic fraction processing; Production of fertilisers from the organic fraction of waste or refuse.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\6_CPC\6.2.5.1. Mechanical processing of waste for the recovery of materials, e.g. crushing, shredding, separation or disassembly.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\6_CPC\6.2.5.10. Packaging reuse or recycling, e.g. of multilayer packaging.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\6_CPC\6.2.5.11. Recycling of waste of electrical or electronic equipment (WEEE).csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\6_CPC\6.2.5.12. Recycling of batteries or fuel cells.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\6_CPC\6.2.5.13. Use of waste materials as fillers for mortars or concrete.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\6_CPC\6.2.5.2. Waste management of vehicles.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\6_CPC\6.2.5.3. Construction

```
or demolition C&D waste.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\6_CPC\6.2.5.4. Glass
recycling.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\6_CPC\6.2.5.5. Plastics and
rubber recycling.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\6_CPC\6.2.5.6. Paper
recycling.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\6_CPC\6.2.5.7. Disintegrating
fibre-containing textile articles to obtain fibres for re-use.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\6_CPC\6.2.5.9. Recycling of
wood or furniture waste.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\6_CPC\6.3. ENABLING
TECHNOLOGIES OR TECHNOLOGIES WITH A POTENTIAL OR INDIRECT CONTRIBUTION TO GHG
EMISSIONS MITIGATION.csv
```

[ ]:
```python
path_base = r"C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green␣
↪classification"
path_dati = path_base + "\\ENV-TECH codes\\"
lista_dd = []
for x in lista_dir7:
    lista_dd.append(path_dati + x + "\\")
print(f'directory dati: {lista_dd}')

for dd in lista_dd:
    print(f'>>> inizio elaborazione {dd}')
    file_da_elaborare = lista_file_csv(dd)

lista_df = []
for ff in file_da_elaborare:
    df = carica_dati_da_csv(ff)
    lista_df.append(df)
df_code_7 = pd.concat(lista_df)
```

```
directory dati: ['C:\\Users\\bincoletto\\OneDrive - Area Science
Park\\Documenti\\green classification\\ENV-TECH codes\\7_CPC\\']
```

```
>>> inizio elaborazione C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\
la directory contiene 31 file
Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7. CLIMATE CHANGE
MITIGATION TECHNOLOGIES IN THE PRODUCTION OR PROCESSING OF GOODS.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.1.  TECHNOLOGIES
RELATED TO METAL PROCESSING.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.1.1. Reduction of
greenhouse gas GHG emissions.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.1.2. Process
efficiency.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.2. TECHNOLOGIES
RELATING TO CHEMICAL INDUSTRY.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.2.1. Process
efficiency in chemical industry.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.2.2. Feedstock.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.2.3. Reduction of
greenhouse gas GHG emissions, e.g. CO2.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.2.4. Improvements
relating to chlorine production.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.2.5. Improvements
relating to adipic acid or caprolactam production.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.2.6. Improvements
relating to fluorochloro hydrocarbon, e.g. chlorodifluoromethane HCFC-22
production.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
```

Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.2.7. Improvements relating to the production of bulk chemicals.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.3. TECHNOLOGIES RELATING TO OIL REFINING AND PETROCHEMICAL INDUSTRY.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.3.1. Bio-feedstock.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.3.2. Ethylene production.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.4. TECHNOLOGIES RELATING TO THE PROCESSING OF MINERALS.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.4.1. Production of cement.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.4.2. Production or processing of lime.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.4.3. Glass production.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.5. TECHNOLOGIES RELATING TO AGRICULTURE, LIVESTOCK OR AGROALIMENTARY INDUSTRIES.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.5.1. Using renewable energies, e.g. solar water pumping.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.5.2. Measures for saving energy, e.g. in green houses.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.5.3. Reduction of greenhouse gas GHG emissions in agriculture.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science

```
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.5.4. Land use policy
measures.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.5.5. Afforestation or
reforestation.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.5.6. Livestock or
poultry management.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.5.7. Fishing;
Aquaculture; Aquafarming.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.5.8. Food processing,
e.g. use of renewable energies or variable speed drives in handling, conveying
or stacking.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.6. TECHNOLOGIES IN
THE PRODUCTION PROCESS FOR FINAL INDUSTRIAL OR CONSUMER PRODUCTS.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.7. CLIMATE CHANGE
MITIGATION TECHNOLOGIES FOR SECTOR-WIDE APPLICATIONS.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\7_CPC\7.8. ENABLING
TECHNOLOGIES WITH A POTENTIAL CONTRIBUTION TO GHG EMISSIONS MITIGATION.csv
```

[ ]:
```python
path_base = r"C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green␣
↪classification"
path_dati = path_base + "\\ENV-TECH codes\\"
lista_dd = []
for x in lista_dir8:
    lista_dd.append(path_dati + x + "\\")
print(f'directory dati: {lista_dd}')

for dd in lista_dd:
    print(f'>>> inizio elaborazione {dd}')
    file_da_elaborare = lista_file_csv(dd)

lista_df = []
for ff in file_da_elaborare:
```

```
    df = carica_dati_da_csv(ff)
    lista_df.append(df)
df_code_8 = pd.concat(lista_df)
```

```
directory dati: ['C:\\Users\\bincoletto\\OneDrive - Area Science
Park\\Documenti\\green classification\\ENV-TECH codes\\8_CPC\\']
>>> inizio elaborazione C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\8_CPC\
la directory contiene 3 file
Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\8_CPC\8. CLIMATE CHANGE
MITIGATION IN INFORMATION AND COMMUNICATION TECHNOLOGIES ICT.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\8_CPC\8.1.  Energy efficient
computing.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\8_CPC\8.2.  Energy efficiency
in communication networks.csv
```

[ ]:
```
path_base = r"C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green␣
 ↪classification"
path_dati = path_base + "\\ENV-TECH codes\\"
lista_dd = []
for x in lista_dir9:
    lista_dd.append(path_dati + x + "\\")
print(f'directory dati: {lista_dd}')

for dd in lista_dd:
    print(f'>>> inizio elaborazione {dd}')
    file_da_elaborare = lista_file_csv(dd)

lista_df = []
for ff in file_da_elaborare:
    df = carica_dati_da_csv(ff)
    lista_df.append(df)
df_code_9 = pd.concat(lista_df)
```

```
directory dati: ['C:\\Users\\bincoletto\\OneDrive - Area Science
Park\\Documenti\\green classification\\ENV-TECH codes\\1_IPC\\']
>>> inizio elaborazione C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\1_IPC\
la directory contiene 13 file
Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\1_IPC\1.1.1. Emissions
```

abatement from stationary sources e.g. SOx, NOx, PM emissions from combustion plants.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\1_IPC\1.1.2. Emissions abatement from mobile sources e.g. NOx, CO, HC, PM emissions from motor vehicles.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\1_IPC\1.1.3. Air pollution abatement - Not elsewhere classified.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\1_IPC\1.2.1. Water and wastewater treatment.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\1_IPC\1.2.2. Fertilizers from wastewater.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\1_IPC\1.2.3. Oil spill and pollutant clean-up.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\1_IPC\1.3.1. Solid waste collection.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\1_IPC\1.3.2. Material recovery, recycling and re-use.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\1_IPC\1.3.3. Fertilizers from waste.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\1_IPC\1.3.4. Incineration and energy recovery.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\1_IPC\1.3.6. Waste management Not elsewhere classified.csv

Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science Park\Documenti\green classification\ENV-TECH codes\1_IPC\1.4. Soil remediation.csv

```
Elaborazione dei file C:\Users\bincoletto\OneDrive - Area Science
Park\Documenti\green classification\ENV-TECH codes\1_IPC\1.5.  Environmental
monitoring.csv
```

Append dei diversi dataframe e controllo dimensioni

Solo CPC

[ ]:
```
CPC_code = df_code_1.append([df_code_2, df_code_3, df_code_4, df_code_5,␣
 ↪df_code_6, df_code_7, df_code_8])
CPC_code.shape
```

[ ]: (3613, 1)

[ ]:
```
CPC_code_univoci = CPC_code.drop_duplicates(subset=['Code'])
CPC_code_univoci.shape
```

[ ]: (3097, 1)

CPC + IPC

[ ]:
```
# append di tutti e 8 i vari dataframe
df_code = df_code_1.append([df_code_2, df_code_3, df_code_4, df_code_5,␣
 ↪df_code_6, df_code_7, df_code_8, df_code_9])
df_code.shape
```

[ ]: (4525, 1)

[ ]:
```
# eliminare i code ripetuti
df_code_univoco = df_code.drop_duplicates(subset=['Code'])
df_code_univoco.shape
```

[ ]: (3150, 1)

Salvataggio csv "ENV_TECH_univoco": tutti i codici CPC e IPC univoci, senza doppioni

[ ]:
```
path = "C:\\Users\\bincoletto\\OneDrive - Area Science Park\\Documenti\\green␣
 ↪classification\\ENV-TECH codes\\finale\\ENV_TECH_univoco.csv"
#df_code_univoco.to_csv(path, sep=';', encoding='utf-8', index=False)
df_code_univoco.to_csv(path, sep='|',encoding='utf-8-sig', index=False)
```

Salvataggio csv "ENV_TECH_level": codici senza doppioni per riga, ma con una colonna in più relativa al livello 2 (nome dei singoli csv, es. "1.1.1. Emissions abatement from stationary sources")

[ ]:
```
# Aggiungo una colonna con il nome del file .csv ad ogni ddataframe

# path_base = r"C:\Users\bincoletto\OneDrive - Area Science␣
 ↪Park\Documenti\green classification"
```

```
# path_dati = path_base + "\\ENV-TECH codes\\"

path = os.path.join(path_dati, '1_CPC')
files = [os.path.join(path,i) for i in os.listdir(path) if os.path.isfile(os.
 ↪path.join(path,i))]

df_1 = pd.DataFrame()

for file in files:
    _df = pd.read_csv(file)
    _df['level_2'] = os.path.split(file)[-1]
    df_1 = df_1.append(_df)

df_1['sublevel'] = df_1['level_2'].str.replace('.csv', '', regex=True)
df_1 = df_1.drop(columns=[
    'level_2'
])
df_1 = df_1.rename(columns={
    'sublevel': 'level_2'
})

df_1["Fonte"] = 'CPC'

df_1.value_counts('level_2')
```

[ ]:
```
level_2
1.1.2. Emissions abatement from mobile sources              898
1.2.1. Water and wastewater treatment                       621
1.1.3. Air pollution abatement - Not elsewhere classified   402
1.3.1. Solid waste collection                               277
1.3.2. Material recovery recycling and re-use               191
1.1.1. Emissions abatement from stationary sources          156
1.3.4. Incineration and energy recovery                      65
1.3.3. Fertilizers from waste                                55
1.2.3. Oil spill and pollutant clean-up                      35
1.3.6. Waste management Not elsewhere classified             29
1.4.  Soil remediation                                       15
1.5.  Environmental monitoring                                6
1.2.2. Fertilizers from wastewater                            4
dtype: int64
```

[ ]:
```
path = os.path.join(path_dati, '2_CPC')
files = [os.path.join(path,i) for i in os.listdir(path) if os.path.isfile(os.
 ↪path.join(path,i))]

df_2 = pd.DataFrame()
```

```
for file in files:
    _df = pd.read_csv(file)
    _df['level_2'] = os.path.split(file)[-1]
    df_2 = df_2.append(_df)

df_2['sublevel'] = df_2['level_2'].str.replace('.csv', '', regex=True)
df_2 = df_2.drop(columns=[
    'level_2'
])
df_2 = df_2.rename(columns={
    'sublevel': 'level_2'
})

df_2["Fonte"] = 'CPC'

df_2.value_counts('level_2')
```

[ ]:
```
level_2
2. Climate change mitigation technologies related to energy generation,
transmission or distribution            95
2.1. Renewable energy generation
38
2.6. Enabling technologies (technologies with potential or indirect contribution
to GHG emission mitigation)    23
2.1.3. Solar photovoltaic (PV) energy
15
2.1.1. Wind energy
11
2.3. Combustion technologies with mitigation potential (e.g. using fossil fuels,
biomass, waste, etc.)          10
2.5. Technologies for an efficient electrical power generation, transmission or
distribution                    8
2.4. Nuclear energy
8
2.2.  Energy generation from fuels of non-fossil origin
6
2.5.3. Not elsewhere classified
5
2.3.2. Technologies for improved input efficiency (efficient combustion or heat
usage)                         5
2.6.2. Hydrogen technology
4
2.3.1. Technologies for improved output efficiency (combined heat and power,
combined cycles, etc.)             4
2.1.2. Solar thermal energy
4
2.7. Other energy conversion or management systems reducing GHG emissions
```

```
2
2.6.4. High-voltage direct current transmission
1
2.6.3. Fuel cells
1
2.6.1.4. Mechanical energy storage, e.g. flywheels or pressurised fluids
1
2.6.1.3. Thermal energy storage
1
2.6.1.2. Capacitors
1
2.6.1.1. Batteries
1
2.1.7. Hydro energy
1
2.5.2. Smart grids as climate change mitigation technology in the energy
generation sector                          1
2.1.4. Solar thermal-PV hybrids
1
2.4.2. Nuclear fission reactors
1
2.4.1. Nuclear fusion reactors
1
2.1.5. Geothermal energy
1
2.1.6. Marine energy, e.g. using wave energy or salinity gradient
1
2.2.2. Fuel from waste, e.g. synthetic alcohol or diesel
1
2.2.1. Biofuels, e.g. bio-diesel
1
2.5.1. Superconducting electric elements or equipment
1
dtype: int64
```

[ ]:
```python
path = os.path.join(path_dati, '3_CPC')
files = [os.path.join(path,i) for i in os.listdir(path) if os.path.isfile(os.
 ↪path.join(path,i))]

df_3 = pd.DataFrame()

for file in files:
    _df = pd.read_csv(file)
    _df['level_2'] = os.path.split(file)[-1]
    df_3 = df_3.append(_df)

df_3['sublevel'] = df_3['level_2'].str.replace('.csv', '', regex=True)
```

```
df_3 = df_3.drop(columns=[
    'level_2'
])
df_3 = df_3.rename(columns={
    'sublevel': 'level_2'
})

df_3["Fonte"] = 'CPC'

df_3.value_counts('level_2')
```

[ ]:
```
level_2
3. Capture, storage, sequestration or disposal of greenhouse gases
7
3.1. CAPTURE OR DISPOSAL OF NITROUS OXIDE (N2O)
1
3.2. Capture or disposal of methane (CH4)
1
3.3. Capture or disposal of perfluorocarbons (PFC), hydrofluorocarbons (HFC) or
sulfur hexafluoride (SF6)    1
3.4. Capture or disposal of carbon dioxide (CO2)
1
dtype: int64
```

[ ]:
```
path = os.path.join(path_dati, '4_CPC')
files = [os.path.join(path,i) for i in os.listdir(path) if os.path.isfile(os.
 ↪path.join(path,i))]

df_4 = pd.DataFrame()

for file in files:
    _df = pd.read_csv(file)
    _df['level_2'] = os.path.split(file)[-1]
    df_4 = df_4.append(_df)

df_4['sublevel'] = df_4['level_2'].str.replace('.csv', '', regex=True)
df_4 = df_4.drop(columns=[
    'level_2'
])
df_4 = df_4.rename(columns={
    'sublevel': 'level_2'
})

df_4["Fonte"] = 'CPC'

df_4.value_counts('level_2')
```

[ ]:
```
level_2
4. Climate change mitigation technologies related to transportation
75
4.1. Road transport
38
4.5. Enabling technologies in transport
15
4.3. Aeronautics or air transport
13
4.5.1. Electric vehicle charging
11
4.1.4. Fuel efficiency-improving vehicle design (common to all road vehicles)
7
4.4. Maritime or waterways transport
6
4.1.1. Conventional vehicles (based on internal combustion engine)
4
4.1.3. Electric vehicles
4
4.2. Rail transport
1
4.5.2. Application of hydrogen technology to transportation, e.g. using fuel
cells     1
dtype: int64
```

[ ]:
```
path = os.path.join(path_dati, '5_CPC')
files = [os.path.join(path,i) for i in os.listdir(path) if os.path.isfile(os.
 ↪path.join(path,i))]

df_5 = pd.DataFrame()

for file in files:
    _df = pd.read_csv(file)
    _df['level_2'] = os.path.split(file)[-1]
    df_5 = df_5.append(_df)

df_5['sublevel'] = df_5['level_2'].str.replace('.csv', '', regex=True)
df_5 = df_5.drop(columns=[
    'level_2'
])
df_5 = df_5.rename(columns={
    'sublevel': 'level_2'
})

df_5["Fonte"] = 'CPC'

df_5.value_counts('level_2')
```

[ ]:
```
level_2
5. Climate change mitigation technologies related to buildings
62
5.2.2. Energy efficient heating, ventilation or air conditioning
12
5.2.5. End-user side
10
5.1. Integration of renewable energy sources in buildings
7
5.2.1. Energy efficient lighting
7
5.4. Enabling technologies in buildings
6
5.3. Architectural or constructional elements improving the thermal performance
of buildings                              4
5.2.3. Energy efficiency in home appliances
3
5.2.4. Energy efficient elevators, escalators and moving walkways, e.g. energy
saving or recuperation technologies     1
dtype: int64
```

[ ]:
```
path = os.path.join(path_dati, '6_CPC')
files = [os.path.join(path,i) for i in os.listdir(path) if os.path.isfile(os.
 ↪path.join(path,i))]

df_6 = pd.DataFrame()

for file in files:
    _df = pd.read_csv(file)
    _df['level_2'] = os.path.split(file)[-1]
    df_6 = df_6.append(_df)

df_6['sublevel'] = df_6['level_2'].str.replace('.csv', '', regex=True)
df_6 = df_6.drop(columns=[
    'level_2'
])
df_6 = df_6.rename(columns={
    'sublevel': 'level_2'
})

df_6["Fonte"] = 'CPC'

df_6.value_counts('level_2')
```

[ ]:
```
level_2
6.2.  SOLID WASTE MANAGEMENT
28
```

```
6.3. ENABLING TECHNOLOGIES OR TECHNOLOGIES WITH A POTENTIAL OR INDIRECT
CONTRIBUTION TO GHG EMISSIONS MITIGATION               2
6.2.2. Waste processing or separation
1
6.2.3. Landfill technologies aiming to mitigate methane emissions
1
6.2.4. Bio-organic fraction processing; Production of fertilisers from the
organic fraction of waste or refuse                  1
6.2.5.1. Mechanical processing of waste for the recovery of materials, e.g.
crushing, shredding, separation or disassembly     1
6.2.5.10. Packaging reuse or recycling, e.g. of multilayer packaging
1
6.2.5.11. Recycling of waste of electrical or electronic equipment (WEEE)
1
6.2.5.12. Recycling of batteries or fuel cells
1
6.2.1.Waste collection, transportation, transfer or storage
1
6.2.5.2. Waste management of vehicles
1
6.2.5.3. Construction or demolition C&D waste
1
6.2.5.4. Glass recycling
1
6.2.5.5. Plastics and rubber recycling
1
6.2.5.6. Paper recycling
1
6.2.5.7. Disintegrating fibre-containing textile articles to obtain fibres for
re-use                                               1
6.2.5.9. Recycling of wood or furniture waste
1
6.2.5.13. Use of waste materials as fillers for mortars or concrete
1
dtype: int64
```

[ ]:
```python
path = os.path.join(path_dati, '7_CPC')
files = [os.path.join(path,i) for i in os.listdir(path) if os.path.isfile(os.
 ↪path.join(path,i))]

df_7 = pd.DataFrame()

for file in files:
    _df = pd.read_csv(file)
    _df['level_2'] = os.path.split(file)[-1]
    df_7 = df_7.append(_df)
```

```
df_7['sublevel'] = df_7['level_2'].str.replace('.csv', '', regex=True)
df_7 = df_7.drop(columns=[
    'level_2'
])
df_7 = df_7.rename(columns={
    'sublevel': 'level_2'
})

df_7["Fonte"] = 'CPC'

df_7.value_counts('level_2')
```

[ ]:
```
level_2
7. CLIMATE CHANGE MITIGATION TECHNOLOGIES IN THE PRODUCTION OR PROCESSING OF
GOODS                                        85
7.2. TECHNOLOGIES RELATING TO CHEMICAL INDUSTRY
21
7.5. TECHNOLOGIES RELATING TO AGRICULTURE, LIVESTOCK OR AGROALIMENTARY
INDUSTRIES                                   14
7.8. ENABLING TECHNOLOGIES WITH A POTENTIAL CONTRIBUTION TO GHG EMISSIONS
MITIGATION                                   13
7.4. TECHNOLOGIES RELATING TO THE PROCESSING OF MINERALS
10
7.1.  TECHNOLOGIES RELATED TO METAL PROCESSING
9
7.7. CLIMATE CHANGE MITIGATION TECHNOLOGIES FOR SECTOR-WIDE APPLICATIONS
8
7.6. TECHNOLOGIES IN THE PRODUCTION PROCESS FOR FINAL INDUSTRIAL OR CONSUMER
PRODUCTS                                     7
7.1.1. Reduction of greenhouse gas GHG emissions
5
7.4.1. Production of cement
4
7.5.8. Food processing, e.g. use of renewable energies or variable speed drives
in handling, conveying or stacking     3
7.5.3. Reduction of greenhouse gas GHG emissions in agriculture
3
7.2.1. Process efficiency in chemical industry
3
7.1.2. Process efficiency
3
7.3. TECHNOLOGIES RELATING TO OIL REFINING AND PETROCHEMICAL INDUSTRY
3
7.2.2. Feedstock
3
7.2.3. Reduction of greenhouse gas GHG emissions, e.g. CO2
3
```

```
7.4.2. Production or processing of lime
2
7.4.3. Glass production
2
7.5.6. Livestock or poultry management
2
7.5.5. Afforestation or reforestation
1
7.2.4. Improvements relating to chlorine production
1
7.5.7. Fishing; Aquaculture; Aquafarming
1
7.3.2. Ethylene production
1
7.5.4. Land use policy measures
1
7.3.1. Bio-feedstock
1
7.5.2. Measures for saving energy, e.g. in green houses
1
7.5.1. Using renewable energies, e.g. solar water pumping
1
7.2.6. Improvements relating to fluorochloro hydrocarbon, e.g.
chlorodifluoromethane HCFC-22 production                1
7.2.7. Improvements relating to the production of bulk chemicals
1
7.2.5. Improvements relating to adipic acid or caprolactam production
1
dtype: int64
```

[ ]:
```
path = os.path.join(path_dati, '8_CPC')
files = [os.path.join(path,i) for i in os.listdir(path) if os.path.isfile(os.
 ↪path.join(path,i))]

df_8 = pd.DataFrame()

for file in files:
    _df = pd.read_csv(file)
    _df['level_2'] = os.path.split(file)[-1]
    df_8 = df_8.append(_df)

df_8['sublevel'] = df_8['level_2'].str.replace('.csv', '', regex=True)
df_8 = df_8.drop(columns=[
    'level_2'
])
df_8 = df_8.rename(columns={
    'sublevel': 'level_2'
```

```
})

df_8["Fonte"] = 'CPC'

df_8.value_counts('level_2')
```

[ ]:
```
level_2
8. CLIMATE CHANGE MITIGATION IN INFORMATION AND COMMUNICATION TECHNOLOGIES ICT
33
8.1.  Energy efficient computing
11
8.2.  Energy efficiency in communication networks
3
dtype: int64
```

[ ]:
```
path = os.path.join(path_dati, '1_IPC')
files = [os.path.join(path,i) for i in os.listdir(path) if os.path.isfile(os.
 ↪path.join(path,i))]

df_9 = pd.DataFrame()

for file in files:
    _df = pd.read_csv(file)
    _df['level_2'] = os.path.split(file)[-1]
    df_9 = df_9.append(_df)

df_9['sublevel'] = df_9['level_2'].str.replace('.csv', '', regex=True)
df_9 = df_9.drop(columns=[
    'level_2'
])
df_9 = df_9.rename(columns={
    'sublevel': 'level_2'
})

df_9["Fonte"] = 'IPC'

df_9.value_counts('level_2')
```

[ ]:
```
level_2
1.1.2. Emissions abatement from mobile sources e.g. NOx, CO, HC, PM emissions
from motor vehicles        310
1.2.1. Water and wastewater treatment
178
1.1.3. Air pollution abatement - Not elsewhere classified
137
1.3.2. Material recovery, recycling and re-use
76
```

```
1.1.1. Emissions abatement from stationary sources e.g. SOx, NOx, PM emissions
from combustion plants     71
1.3.3. Fertilizers from waste
46
1.3.4. Incineration and energy recovery
40
1.3.1. Solid waste collection
29
1.4.  Soil remediation
7
1.2.3. Oil spill and pollutant clean-up
6
1.3.6. Waste management  Not elsewhere classified
6
1.2.2. Fertilizers from wastewater
3
1.5.  Environmental monitoring
3
dtype: int64
```

[ ]:
```
Code_level_all = df_1.append([df_2, df_3, df_4, df_5, df_6, df_7, df_8, df_9])
Code_level_all.shape
```

[ ]:
```
(4525, 3)
```

[ ]:
```
# Uso il groupby per mettere tutti i valori della colonna 'Fonte' in un'unica␣
 ↪cella per codice
# (che dovrà apparire solo una volta per riga)
fonte_all_subset = Code_level_all.copy()
fonte_all_subset= fonte_all_subset.groupby('Code')['Fonte'].apply(';'.join).
 ↪reset_index()
print(f"I progetti del dataframe sono: {fonte_all_subset.shape}")
print(fonte_all_subset.head(5))
print(fonte_all_subset.tail(5))
```

```
I progetti del dataframe sono: (3150, 2)
          Code    Fonte
0  A23K  10/26  CPC;IPC
1  A23K  10/28  CPC;IPC
2  A23K  10/32  CPC;IPC
3  A23K  10/33  CPC;IPC
4  A23K  10/37  CPC;IPC
             Code    Fonte
3145  Y02W  30/84  CPC;CPC
3146  Y02W  30/91  CPC;CPC
3147  Y02W  30/97      CPC
3148  Y02W  90/00      CPC
```

```
3149  Y02W  90/10      CPC
```

[ ]:
```python
# verifica e modifica dei valori della colonna "Fonte" in:
# - CPC
# - IPC
# - CPC & IPC

# verifico i valori della colonna Fonte da semplificare secondo la␣
 ↪categorizzazione sopra
conteggio = fonte_all_subset.value_counts('Fonte')
print('Le righe della colonna "Fonte" sono categorizzate nel seguente modo:')
print(f'{conteggio}')

# ciclo che vado a trovare la presenza dei valori cpc e ipc e sostituisca i␣
 ↪valori correttamente secondo le catoegorie sopra identificate:

fonte_riclassification = fonte_all_subset.copy()
fonte_riclassification['Fonte'] = fonte_riclassification['Fonte'].map({
    'CPC': 'CPC',
    'CPC;IPC': 'CPC & IPC',
    'CPC;CPC': 'CPC',
    'CPC;CPC;CPC': 'CPC',
    'IPC': 'IPC',
    'CPC;CPC;IPC;IPC': 'CPC & IPC',
    np.nan:'Null'
}, na_action=None)
fonte_riclassification
```

```
Le righe della colonna "Fonte" sono categorizzate nel seguente modo:
Fonte
CPC                1877
CPC;IPC             839
CPC;CPC             236
CPC;CPC;CPC         135
IPC                  53
CPC;CPC;IPC;IPC      10
dtype: int64
```

[ ]:
```
             Code      Fonte
0     A23K  10/26  CPC & IPC
1     A23K  10/28  CPC & IPC
2     A23K  10/32  CPC & IPC
3     A23K  10/33  CPC & IPC
4     A23K  10/37  CPC & IPC
…           …        …
3145  Y02W  30/84        CPC
3146  Y02W  30/91        CPC
```

```
3147  Y02W  30/97        CPC
3148  Y02W  90/00        CPC
3149  Y02W  90/10        CPC

[3150 rows x 2 columns]
```

[ ]:
```
# Uso il groupby per mettere tutti i valori della colonna 'level_2' in un'unica␣
 ↪cella per codice
# (che dovrà apparire solo una volta per riga)
level_all_subset = Code_level_all.copy()
level_all_subset= level_all_subset.groupby('Code')['level_2'].apply('; '.join).
 ↪reset_index()
level_all_subset = level_all_subset.rename(columns={
    'level_2':'Level_2'
})
print(level_all_subset.head(10))
print(f"I progetti del dataframe sono: {level_all_subset.shape}")
```

```
           Code                                                Level_2
0    A23K  10/26  1.3.2. Material recovery recycling and re-use;…
1    A23K  10/28  1.3.2. Material recovery recycling and re-use;…
2    A23K  10/32  1.3.2. Material recovery recycling and re-use;…
3    A23K  10/33  1.3.2. Material recovery recycling and re-use;…
4    A23K  10/37  1.3.2. Material recovery recycling and re-use;…
5    A23K  10/38  1.3.2. Material recovery recycling and re-use;…
6    A43B   1/12  1.3.2. Material recovery recycling and re-use;…
7    A61L  11/00  1.3.6. Waste management Not elsewhere classifi…
8    B01D  46/00  1.1.3. Air pollution abatement - Not elsewhere…
9  B01D  46/0001  1.1.3. Air pollution abatement - Not elsewhere…
I progetti del dataframe sono: (3150, 2)
```

[ ]:
```
# merge dei due sottogruppi creati con il groupby
df_final = pd.merge(level_all_subset, fonte_riclassification, how="left",␣
 ↪on='Code')
df_final
```

[ ]:
```
             Code                                                Level_2  \
0     A23K  10/26  1.3.2. Material recovery recycling and re-use;…
1     A23K  10/28  1.3.2. Material recovery recycling and re-use;…
2     A23K  10/32  1.3.2. Material recovery recycling and re-use;…
3     A23K  10/33  1.3.2. Material recovery recycling and re-use;…
4     A23K  10/37  1.3.2. Material recovery recycling and re-use;…
…          …                                                   …
3145  Y02W  30/84  6.2.  SOLID WASTE MANAGEMENT; 6.2.5.12. Recycl…
3146  Y02W  30/91  6.2.  SOLID WASTE MANAGEMENT; 6.2.5.13. Use of…
3147  Y02W  30/97                        6.2.  SOLID WASTE MANAGEMENT
3148  Y02W  90/00  6.3. ENABLING TECHNOLOGIES OR TECHNOLOGIES WIT…
```

```
3149  Y02W  90/10  6.3. ENABLING TECHNOLOGIES OR TECHNOLOGIES WIT…

          Fonte
0     CPC & IPC
1     CPC & IPC
2     CPC & IPC
3     CPC & IPC
4     CPC & IPC
…         …
3145        CPC
3146        CPC
3147        CPC
3148        CPC
3149        CPC

[3150 rows x 3 columns]
```

[ ]:
```
path = "C:\\Users\\bincoletto\\OneDrive - Area Science Park\\Documenti\\green␣
 ↪classification\\ENV-TECH codes\\finale\\ENV_TECH_final.csv"
#df_code_univoco.to_csv(path, sep=';', encoding='utf-8', index=False)
df_final.to_csv(path, sep='|',encoding='utf-8-sig', index=False)
```