import pandas as pd

import numpy as np

import os

os.chdir('E:\\FCSE\\Introduction To Data science\\assignment 1')

*#question 1*

pd.set\_option('display.max\_rows', 5)

dataset=pd.read\_excel("Energy Indicators.xls")

dataset=dataset[17:244]

dataset = dataset.drop(['Unnamed: 0', 'Unnamed: 1'], *axis*=1)

dataset.columns=['Country', 'Energy Supply', 'Energy Supply per Capita', 'Renewable']

dataset['Energy Supply']\*=1000000

*#revise on this line*

dataset['Country'] = dataset['Country'].str.replace(*r*"\(.\*?\)","")

dataset['Country']=dataset['Country'].str.replace('\d+', '')

dataset.replace("Republic of Korea","South Korea",*inplace*=True)

dataset.replace("United States of America","United States", *inplace*=True)

dataset.replace("United Kingdom of Great Britain and Northern Ireland","United Kingdom",*inplace*=True)

dataset.replace("China, Hong Kong Special Administrative Region","Hong Kong", *inplace*=True)

*#dataset*

*#question 2*

GDP=pd.read\_csv('world\_bank.csv',*skiprows*=4)

*#GDP=GDP[4:]*

GDP.replace("Korea, Rep.","South Korea",*inplace*=True)

GDP.replace("Iran, Islamic Rep.","Iran",*inplace*=True)

GDP.replace("Hong Kong SAR, China","Hong Kong", *inplace*=True)

GDP.rename(*columns*={'Country Name':'Country'},*inplace*=True)

GDP

GDP=GDP.drop(GDP.columns[[3]],*axis*=1)

mergeDataSet=pd.merge(dataset,GDP,*on*='Country')

mergeDataSet=mergeDataSet[["Country","Energy Supply","Energy Supply per Capita","Renewable","Country Code","Indicator Name","2010","2011","2012","2013","2014","2015"]]

mergeDataSet.set\_index('Country',*inplace*=True)

mergeDataSet

*def* returnTop15():

    rows = ['2010', '2011', '2012', '2013', '2014', '2015']

    avg= mergeDataSet[rows].mean(*axis*=1)

    one\_to\_top=pd.Series(range(0,184))

    avg.sort\_values(*ascending*=False,*inplace*=True)

    countries=avg.index.values

    avg=avg.set\_axis(one\_to\_top)

    avg=avg.truncate(*after*=14)

    avg=avg.set\_axis(countries[0:15])

    avg

    return avg

*def* avgEnergyCapita():

    newDataSet=mergeDataSet.drop(mergeDataSet[mergeDataSet["Energy Supply per Capita"].str.contains("...")==True].index)

    return pd.to\_numeric(newDataSet ["Energy Supply per Capita"]).mean()

mage=avgEnergyCapita()

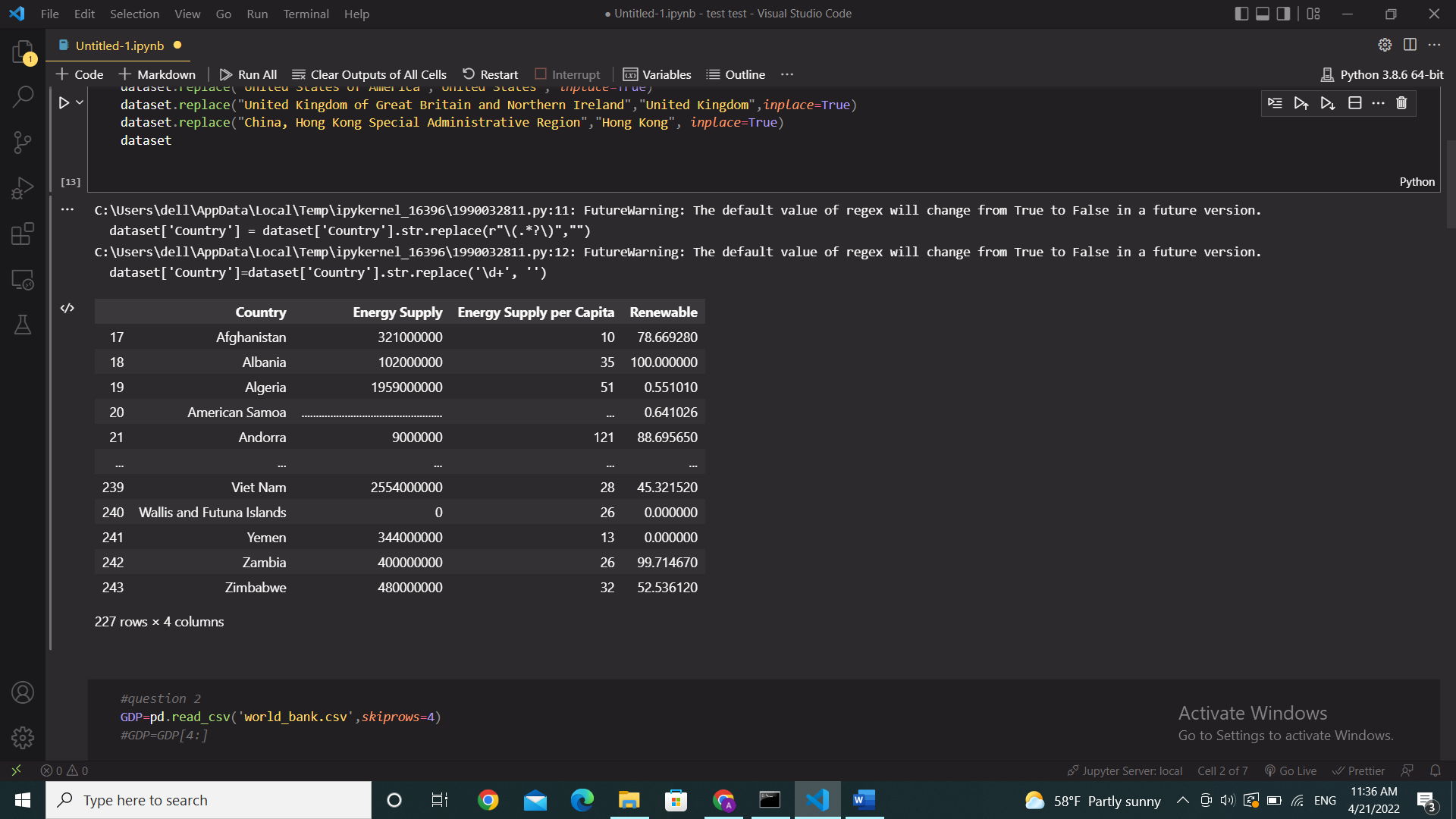
mage

*def* countryMinRenewable():

    newDataSet=mergeDataSet.loc[(mergeDataSet['Renewable'] == mergeDataSet['Renewable'].min())]

    minimum\_energy=newDataSet.iloc[:,[2]]

    return minimum\_energy



A screenshot of a computer

Description automatically generated with medium confidence

Graphical user interface, text

Description automatically generated

A screenshot of a computer

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated with medium confidence