```python
# -*- coding: utf-8 -*-
"""
Created on Wed Feb  2 10:24:57 2022

@author: Lucas R. Amaral

Course IBM Python Project for Data Engineering
Peer-graded Assignment: Peer Review Assignment
"""

#%clear
import os
import glob
import pandas as pd
from datetime import datetime

#============================================================================
# It sets the working directory.
#============================================================================
WORKING_DIRECTORY = """C:\\Users\\01278575677\\OneDrive - Receita Federal do Brasil\\3. Cursos e pós graduações\\3.39
os.chdir(WORKING_DIRECTORY)
os.getcwd()

#============================================================================
# It downloads messages to log file.
#============================================================================
def log(message):
    logfile    = 'logfile.txt'            # all event logs will be stored in this file
    timestamp_format = '%Y-%h-%d-%H:%M:%S' # Year-Monthname-Day-Hour-Minute-Second
    now = datetime.now()                  # get current timestamp
    timestamp = now.strftime(timestamp_format)
    message = timestamp + '\t' + message + '\n'
    with open(logfile, 'a') as f:
        print(message)
        f.write(message)

#============================================================================
# It downloads csv and jason files from urls and saves it to local disk.
#============================================================================
```

```
def download_json():
    file_url = ['https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-PY0221E
                , 'https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-PY022

    for f in file_url:
        file_name = f.split('/')
        file_name = file_name[len(file_name)-1]
        print(file_name)

        df = pd.read_json(f)
        df.to_json(file_name)

    return(0)

#-----------------------------------------------------------------------------
def download_csv():
    file_url = ['https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-PY0221E

    for f in file_url:
        file_name = f.split('/')
        file_name = file_name[len(file_name)-1]
        print(file_name)

        df = pd.read_csv(f)
        df.to_csv(file_name, index=False)

    return(0)

#-----------------------------------------------------------------------------
def download():
    download_json()
    download_csv()
    return(0)

#=============================================================================
# It extracts data from files.
#=============================================================================
def extract():
```

```
    # process exchange_rates.csv
    csvfile = 'exchange_rates.csv'
    extracted_data_csvfile = pd.read_csv(csvfile)


    #process all json files
    extracted_data_jsonfile = pd.DataFrame(columns=['Name', 'Market Cap (US$ Billion)']) # create an empty data frame
    for jsonfile in glob.glob('*.json'):

        extracted_data_jsonfile_temp = pd.read_json(jsonfile)
        extracted_data_jsonfile      = extracted_data_jsonfile.append(
            extracted_data_jsonfile_temp, ignore_index=True)

    return(extracted_data_csvfile, extracted_data_jsonfile)


#================================================================================
# It transforms data.
#================================================================================
def transform_exchange_rate(df_exchange_rate):
    # It renames columns.
    df_exchange_rate.columns = ['Symbol', 'Rate']
    df_exchange_rate         = df_exchange_rate[['Symbol', 'Rate']]
    return(df_exchange_rate)

#--------------------------------------------------------------------------------
def transform_bank_market_cap(df_bank_market_cap
                            , df_exchange_rate
                            , exchange_symbol):

    # Changes the Market Cap (US$ Billion) column from USD to GBP
    # Rounds the Market Cap (US$ Billion)` column to 3 decimal places
    # Rename Market Cap (US$ Billion) to Market Cap (GBP$ Billion)

    exchange_rate = float(df_exchange_rate[df_exchange_rate['Symbol']==exchange_symbol]['Rate'])
    exchange_symbol = 'Market Cap ('+ exchange_symbol + '$ Billion)'
    df_bank_market_cap[exchange_symbol] = round(df_bank_market_cap['Market Cap (US$ Billion)'] * exchange_rate, 3)

    return(df_bank_market_cap)
```

```
#-------------------------------------------------------------------------------
def transform(df_exchange_rate, df_bank_market_cap, exchange_symbol):

    df_exchange_rate   = transform_exchange_rate(df_exchange_rate)

    df_bank_market_cap = transform_bank_market_cap(df_bank_market_cap
                                                   , df_exchange_rate
                                                   , exchange_symbol='GBP')
    return(df_exchange_rate, df_bank_market_cap)

#===============================================================================
# It loads data.
#===============================================================================
def load(df):
    file_name = 'bank_market_cap_gbp.csv'
    df.to_csv(file_name, index=False)


#===============================================================================
# ETL Process.
#===============================================================================

#-------------------------------------------------------------------------------
# Download
#-------------------------------------------------------------------------------
log('Downloading files...')
download()
log('Files downloaded succesfully...')

#-------------------------------------------------------------------------------
# Extract
#-------------------------------------------------------------------------------
log('Extracting data from files...')
df_exchange_rate, df_bank_market_cap = extract()
log('Data extracted from files succesfully...')

#-------------------------------------------------------------------------------
```

```
# Transform
#-------------------------------------------------------------------------------
log('Transforming data...')
exchange_symbol='GBP'
df_exchange_rate, df_bank_market_cap = transform(df_exchange_rate
                                                 , df_bank_market_cap
                                                 , exchange_symbol)
log('Data transformed succesfully...')

#-------------------------------------------------------------------------------
# Load
#-------------------------------------------------------------------------------
log('Loading results in files...')
load(df_bank_market_cap[['Name', 'Market Cap (GBP$ Billion)']])
log('Results loaded succesfully...')


#-------------------------------------------------------------------------------
# Question 1.
#-------------------------------------------------------------------------------
exchange_rate = float(df_exchange_rate[df_exchange_rate['Symbol']==exchange_symbol]['Rate'])
log('The exchange rate for Great British Pounds with the symbol '
+ exchange_symbol + ' is:' + str(exchange_rate))
df_exchange_rate[df_exchange_rate['Symbol']==exchange_symbol]

#-------------------------------------------------------------------------------
# Question 2.
#-------------------------------------------------------------------------------
log('First 5 exchange rates...')
log('\n' + str(df_exchange_rate.head(5)))
log('First 5 bank market cap...')
log('\n' + str(df_bank_market_cap[['Name', 'Market Cap (US$ Billion)']].head(5)))

#-------------------------------------------------------------------------------
# Question 3.
#-------------------------------------------------------------------------------
log('First 5 bank market cap in GBP$...')
df_bank_market_cap_gbp = df_bank_market_cap[['Name', 'Market Cap (GBP$ Billion)']]
log('\n' + str(df_bank_market_cap_gbp.head(5)))
```

```
#------------------------------------------------------------------------------
# Log results
#------------------------------------------------------------------------------
"""
2022-Feb-02-23:55:38 Downloading files...
2022-Feb-02-23:55:41 Files downloaded succesfully...
2022-Feb-02-23:55:41 Extracting data from files...
2022-Feb-02-23:55:41 Data extracted from files succesfully...
2022-Feb-02-23:55:41 Transforming data...
2022-Feb-02-23:55:41 Data transformed succesfully...
2022-Feb-02-23:55:41 Loading results in files...
2022-Feb-02-23:55:41 Results loaded succesfully...
2022-Feb-02-23:55:41 The exchange rate for Great British Pounds with the symbol GBP is:0.7323984208000001
2022-Feb-02-23:55:41 First 5 exchange rates...
2022-Feb-02-23:55:41
  Symbol      Rate
0    AUD  1.297088
1    BGN  1.608653
2    BRL  5.409196
3    CAD  1.271426
4    CHF  0.886083
2022-Feb-02-23:55:41 First 5 bank market cap...
2022-Feb-02-23:55:41
                                  Name  Market Cap (US$ Billion)
0                         JPMorgan Chase                   390.934
1  Industrial and Commercial Bank of China                   345.214
2                        Bank of America                   325.331
3                           Wells Fargo                   308.013
4                 China Construction Bank                   257.399
2022-Feb-02-23:55:41 First 5 bank market cap in GBP$...
2022-Feb-02-23:55:41
                                  Name  Market Cap (GBP$ Billion)
0                         JPMorgan Chase                    286.319
1  Industrial and Commercial Bank of China                    252.834
2                        Bank of America                    238.272
3                           Wells Fargo                    225.588
4                 China Construction Bank                    188.519
```

"""