Final Project

August 7, 2022

[1]:
```
!pip install glob
!pip install pandas
!pip install requests
!pip install datetime
!pip install wget
```

```
ERROR: Could not find a version that satisfies the requirement glob (from
versions: none)
ERROR: No matching distribution found for glob
Requirement already satisfied: pandas in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (1.3.5)
Requirement already satisfied: python-dateutil>=2.7.3 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from pandas)
(2.8.2)
Requirement already satisfied: pytz>=2017.3 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from pandas)
(2022.1)
Requirement already satisfied: numpy>=1.17.3 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from pandas)
(1.21.6)
Requirement already satisfied: six>=1.5 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from python-
dateutil>=2.7.3->pandas) (1.16.0)
Requirement already satisfied: requests in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (2.28.1)
Requirement already satisfied: charset-normalizer<3,>=2 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from requests)
(2.1.0)
Requirement already satisfied: certifi>=2017.4.17 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from requests)
(2022.6.15)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from requests)
(1.26.11)
Requirement already satisfied: idna<4,>=2.5 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from requests)
```

```
(3.3)
Collecting datetime
  Downloading DateTime-4.5-py2.py3-none-any.whl (52 kB)
                              52.0/52.0 kB
8.4 MB/s eta 0:00:00
Requirement already satisfied: pytz in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from datetime)
(2022.1)
Requirement already satisfied: zope.interface in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from datetime)
(5.4.0)
Requirement already satisfied: setuptools in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
zope.interface->datetime) (63.2.0)
Installing collected packages: datetime
Successfully installed datetime-4.5
Requirement already satisfied: wget in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (3.2)
```

[24]:
```
import glob
import pandas as pd
from datetime import datetime
import requests
```

[25]:
```
r = requests.get("https://cf-courses-data.s3.us.cloud-object-storage.appdomain.
 ↪cloud/IBMDeveloperSkillsNetwork-PY0221EN-SkillsNetwork/labs/module%206/
 ↪Lab%20-%20Extract%20Transform%20Load/data/bank_market_cap_1.json",␣
 ↪allow_redirects=True)
open("bank_market_cap_1.json", 'wb').write(r.content)

r = requests.get("https://cf-courses-data.s3.us.cloud-object-storage.appdomain.
 ↪cloud/IBMDeveloperSkillsNetwork-PY0221EN-SkillsNetwork/labs/module%206/
 ↪Lab%20-%20Extract%20Transform%20Load/data/bank_market_cap_2.json",␣
 ↪allow_redirects=True)
open("bank_market_cap_2.json", 'wb').write(r.content)

r = requests.get("https://cf-courses-data.s3.us.cloud-object-storage.appdomain.
 ↪cloud/IBMDeveloperSkillsNetwork-PY0221EN-SkillsNetwork/labs/module%206/
 ↪Final%20Assignment/exchange_rates.csv", allow_redirects=True)
open("exchange_rates.csv", 'wb').write(r.content)
```

[25]:
```
590
```

[17]:
```
def get_exchange_rate():
    rates_df = pd.read_csv("exchange_rates.csv", index_col=0)
    exchange_rate = rates_df.loc["GBP"]["Rates"]
    return exchange_rate
```

[18]:
```
exchange_rate = get_exchange_rate()
exchange_rate
```

[18]:
```
0.7323984208000001
```

[8]:
```
def extract_from_json(file_to_process):
    dataframe = pd.read_json(file_to_process)
    return dataframe
```

[19]:
```
def extract():
    # Write your code here
    extracted_data = pd.DataFrame(columns=['Name','Market Cap (US$ Billion)'])
    jsons = []
    # Process json file
    files = glob.glob(r"*.json")
    files
    for jsonfile in files:
        jsons.append(extract_from_json(jsonfile))

    extracted_data = pd.concat(jsons, ignore_index=True)

    return extracted_data
```

[20]:
```
# files = glob.glob(r"*.json")
# files
extracted_data.head()
```

[20]:

	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

[21]:
```
def transform(data):
    # Write your code here
    data['Market Cap (US$ Billion)'] = round(exchange_rate * data['Market Cap␣
 ↪(US$ Billion)'], 3)
    data.rename(columns={'Market Cap (US$ Billion)': 'Market Cap (GBP$␣
 ↪Billion)'}, inplace=True)
    return data

transformed_data = transform(extracted_data)
transformed_data.head()
```

[21]:

	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
```

[22]:
```
def load(data_to_load, target_file):
    # Write your code here
    data_to_load.to_csv(target_file, index=False)
```

[23]:
```
load(transformed_data,'market_cap.csv')
```

[ ]:
```

```