## Import Python libraries ##

import http.client

import json

import numpy as np

import pandas as pd

from google.colab import files

### --------- Please update this section only ------------- ###

## Define search variables ##

country\_list = ["Japan", "Thailand", "Australia", "Vietnam", "India", "Philippines", "Malaysia"] # Please update list of countries accordingly

### ------------------------------------------------------- ###

## Initialize other variables

country\_id = []

successful\_country = []

news\_id = []

## Establish initial connection ##

conn = http.client.HTTPSConnection("www.kensho.com")

headers = {'authorization': "Token 915a67f1a7e00a6170c50b2bae50fd1cc63cb6f5"} # Please update token if necessary

conn.request("GET", "/external/v1/list\_entities\_of\_class?class\_name=Region", headers = headers)

res = conn.getresponse()

## Import data using json and pandas libraries ##

temp = res.read()

dict = json.loads(temp) # dict is a dictionary. There are 2 keys in dict: 'data' (i.e. the data sources) and 'metadata' (i.e. data type of each column). Each of them has different length.

## Query unique ID for each country in country\_list ##

for country in range(len(country\_list)):

for kenshoCountry in range(len(dict['data'])):

if country\_list[country].replace(" ", "").upper() == dict['data'][kenshoCountry]['entity\_name'].replace(" ", "").upper() :

country\_id.append(dict['data'][kenshoCountry]['entity\_id'])

## Re-establish connection and query related M&A news using country\_id ##

for number in range(len(country\_id)):

conn.request("GET", "/external/v1/get\_related\_entities?entity\_id=" + country\_id[number] + "&relationship=RegionTimelines", headers = headers)

res = conn.getresponse()

temp = res.read()

dict = json.loads(temp)

for news in range(len(dict['data'])):

if dict['data'][news]['timeline\_type'] == "M&A":

successful\_country.append(country\_list[number])

news\_id.append(dict['data'][news]['entity\_id'])

## If there are no M&A news for all countries, print error message ##

if news\_id == []:

print("There are no country M&A news in Kensho database!")

## Else download news into Excel csv file ##

else:

for number in range(len(news\_id)):

conn.request("GET", "/external/v1/get\_timeline?timeline\_id=" + news\_id[number], headers = headers)

res = conn.getresponse()

temp = res.read()

dict = json.loads(temp)

## Create a Pandas Dataframe using 'data' index ##

df\_main = pd.DataFrame(dict['data'])

## Update starting index from 0 to 1 using numpy library ##

df\_main.index = np.arange(1,len(df\_main)+1)

## Download data as Main.csv in 'Download' folder of local directory ##

df\_main.to\_csv(successful\_country[number] + '\_MAnews.csv')

files.download(successful\_country[number] + '\_MAnews.csv')

print("These countries do not have any M&A news: " + str(set(country\_list).difference(successful\_country\_id)))