import requests

from bs4 import BeautifulSoup

import pandas as pd

def extract\_store\_details(store\_elem):

    store\_details = {

        'web\_scraper\_order': store\_elem.get('web\_scraper\_order', 'missing'),

        'web-scraper-start-url': store\_elem.get('web-scraper-url', 'missing'),

        'location\_title': store\_elem.get('loc\_title', 'missing'),

        'city': store\_elem.get('store\_city', 'missing'),

        'location\_of\_store-irvine': store\_elem.get('store-irvine', 'missing'),

        'location\_of\_store-sandiego': store\_elem.get('store-sandiego', 'missing'),

        'location-losangeles': store\_elem.get('losangeles', 'missing'),

        'country': store\_elem.get('store\_country', 'missing'),

        'phone': store\_elem.get('store\_phone', 'missing'),

        'hoursofoperation': store\_elem.get('op\_hours', 'missing'),

        'latitude': store\_elem.get('lat', 'missing'),

        'longitude':store\_elem.get('long', 'missing'),

    }

    return store\_details

def scrape\_sephora\_stores(url):

    response = requests.get(url)

    response.raise\_for\_status()  # Check for request errors

    soup = BeautifulSoup(response.text, 'html.parser')

    stores = []

    # Update selector based on actual HTML structure

    store\_elements = soup.find\_all('div', class\_='store-item')  # Example class name

    for store\_elem in store\_elements:

        web\_scraper\_order = store\_elem.find('h3', class\_='web-scraper-ord')

        web\_scraper\_start\_url = store\_elem.find('p', class\_='web\_scraper\_url')

        location\_title= store\_elem.find('span', class\_='loc\_title')

        city= store\_elem.find('span', class\_='store\_city')

        location\_of\_store\_irvine = store\_elem.find('span', class\_='store\_irvine')

        location\_of\_store\_sandiego = store\_elem.find('div', class\_='store\_sandiego')

        location\_losangeles= store\_elem.find('div', class\_='store\_la')

        country = store\_elem.get('countr', 'N/A')

        phone = store\_elem.find('div', class\_='store\_phone')

        hoursofoperation = store\_elem.find('div', class\_='hrsofop')

        allstores = store\_elem.find('div', class\_='store\_all')

        latitude = store\_elem.find('div', class\_='lat')

        longitude = store\_elem.find('div', class\_='long')

        store = {

            'web\_scraper\_order': web\_scraper\_order.text if web\_scraper\_order else 'N/A',

            ' web\_scraper\_start\_url':  web\_scraper\_start\_url.text if  web\_scraper\_start\_url else 'N/A',

            ' location\_title':   location\_title.text if   location\_title else 'N/A',

            ' city': city.text if city else 'N/A',

            'country': 'USA',

            'location\_of\_store\_irvine': location\_of\_store\_irvine.text if location\_of\_store\_irvine else 'N/A',

            'location\_of\_store\_sandiego': location\_of\_store\_sandiego.text if location\_of\_store\_sandiego else 'N/A',

            'location\_losangeles': location\_losangeles.text if location\_losangeles else 'N/A',

            'phone': phone.text if phone else 'N/A',

            'hoursofoperation' : hoursofoperation.number if hoursofoperation else 'N/A',

            'allstores': allstores.text if allstores else 'N/A',

            'latitude':   latitude.text if   latitude else 'N/A',

            'longitude':  longitude.text if  longitude else 'N/A'

        }

        if store['State'] == 'CA':

            stores.append(store)

    return stores

url = 'https://www.sephora.com/happening/storelist'

store\_data = scrape\_sephora\_stores(url)

# Save data to CSV

df = pd.DataFrame(store\_data)

df.to\_csv('sephora\_california\_stores.csv')