import requests

from bs4 import BeautifulSoup

import pandas as pd

# Set the base URL and number of pages to scrape

base\_url = "http://www.pakistanjobsbank.com/Jobs/"

num\_pages = 10150

# Create lists to store the scraped data

title\_list = []

category\_list = []

date\_position\_list = []

location\_list = []

# Iterate over the pages and scrape the data

for page in range(1, num\_pages + 1):

    url = base\_url + str(page)

    response = requests.get(url)

    soup = BeautifulSoup(response.content, "html.parser")

    # Find the elements you want to scrape

    title\_elements = soup.find\_all("h1")

    category\_elements = soup.find\_all("a")

    date\_position\_elements = soup.find\_all("td")

    location\_elements = soup.find\_all(class\_="job-locations")

    for title in title\_elements:

        title\_list.append(title.text.strip())

    for category in category\_elements:

        category\_list.append(category.text.strip())

    for date\_position in date\_position\_elements:

        date\_position\_list.append(date\_position.text.strip())

    for location in location\_elements:

        location\_list.append(location.text.strip())

# Combine the lists into a list of tuples

data = list(zip(title\_list, category\_list, date\_position\_list, location\_list))

df = pd.DataFrame(data, columns=["Title", "Category", "Date\_Position", "Location"])

output\_file = "dfs.xlsx"

df.to\_excel(output\_file, index=False)

print("Data saved to", output\_file)