

Q1: Extract URLs, Titles, and Headings using BeautifulSoup

```
from bs4 import BeautifulSoup
import requests
```

```
url = input("Enter website (without http://): ")
r = requests.get("http://" + url)
data = r.text
soup = BeautifulSoup(data, "html.parser")
for link in soup.find_all('a'):
    print(link.get('href'))
```

```
req = requests.get('https://en.wikipedia.org/wiki/Python_(programming_language)')
soup = BeautifulSoup(req.text, "lxml")
print(soup.title)
print(soup.title.name)
print(soup.title.string)
print(soup.h1.string)
```

Q2: Web Crawling using BeautifulSoup and MongoDB Integration

```
from pymongo import MongoClient
client = MongoClient("localhost", 27017)
db = client["lotto"]
collection = db["mega_millions"]
for page in range(1, 4):
    url = f'http://www.usamega.com/mega-millions-history.asp?p={page}'
    soup = BeautifulSoup(requests.get(url).text, "html.parser")
    for row in soup.find_all('table')[4].find_all('tr'):
        tds = row.find_all('td')
        if len(tds) > 3 and tds[1].a and tds[3].b:
            record = {'date': tds[1].a.string, 'numbers': [int(n) for n in tds[3].b.string.split('.')],
'mega_number': int(tds[3].strong.string)}
            collection.insert_one(record)
```

Q3: Processing MongoDB Data using Pandas

```
import pandas as pd
client = MongoClient("localhost", 27017)
db = client["test"]
collection = db["people"]
pipeline = [{"$group": {"_id": "$Country", "AvgAge": {"$avg": "$Age"}, "Count": {"$sum": 1}},
{"$sort": {"Count": -1, "AvgAge": 1}}]
agg_result = list(collection.aggregate(pipeline))
df = pd.DataFrame(agg_result).set_index("_id")
print(df.head())
```

Q4: Filtering and Visualizing MongoDB Data using Pandas and Folium

```
df_china = pd.DataFrame(list(collection.find({"Country": "China"})))
print(df_china)
import folium
world_map = folium.Map(location=[35, 100], zoom_start=4)
for _, row in df_china.iterrows():
    lat, lon = map(float, row["Location"].split(','))
    folium.Marker(location=[lat, lon], popup=f"{row['Name']}, Age: {row['Age']}").add_to(world_map)
world_map.save("china_people_map.html")
```